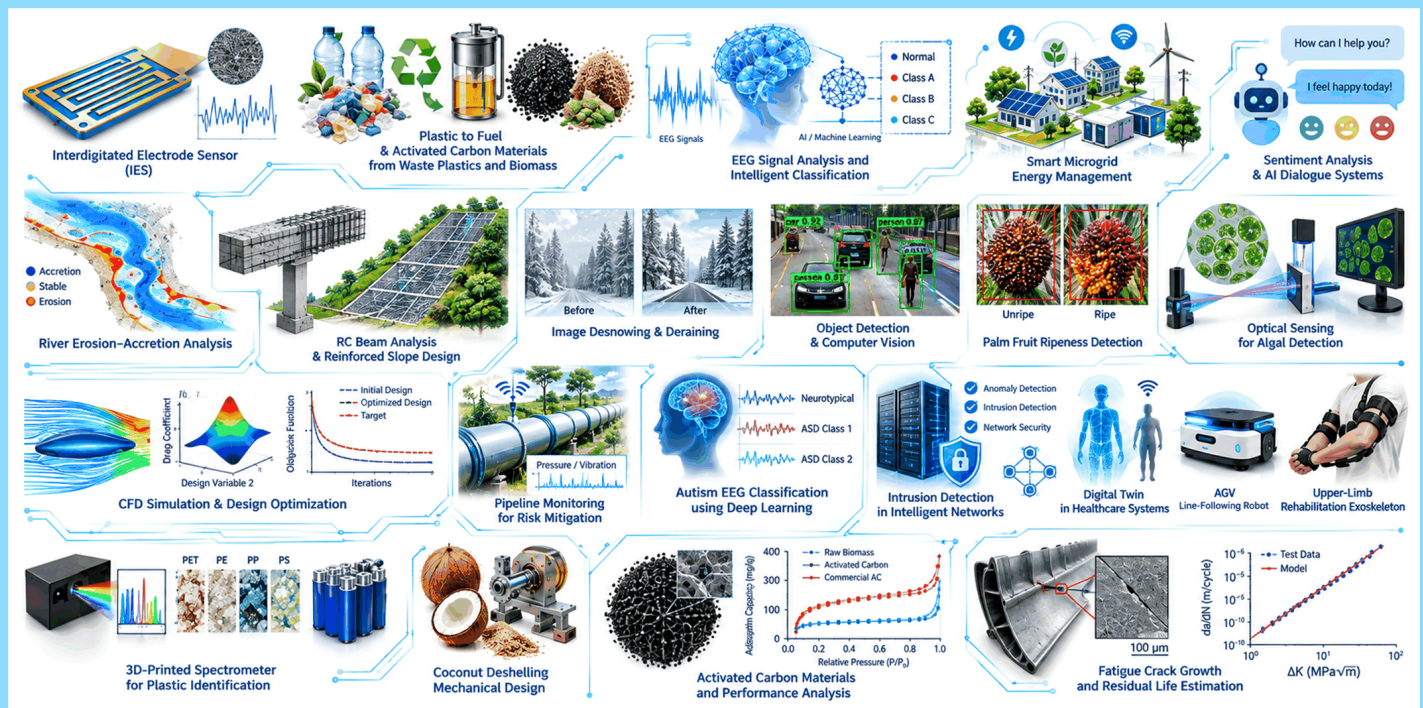


IUM ENGINEERING JOURNAL

Volume 27

September 2026

Number 3



IIUM ENGINEERING JOURNAL

Volume 27, Issue 3, Sep 2026

<https://doi.org/10.31436/iiumej.v27i3>

Copyright Notice

Consent to publish: The Author(s) agree to publish their articles with IIUM Press.

Declaration: The Author(s) declare that the article has not been published before in any form, that it is not concurrently submitted to another publication, and that it does not infringe on anyone's copyright. The Author(s) hold the IIUM Press and Editors of the journal harmless against all copyright claims.

Transfer of copyright: The Author(s) hereby agree to transfer the article's copyright to IIUM Press, which shall have the non-exclusive and unlimited right to publish the article in any form, including in electronic media. For articles with more than one author, the corresponding author confirms that he/she is authorized by his/her co-author(s) to grant this copyright transfer.

The IIUM Engineering Journal follows the open access policy.

All articles published for open access will be immediately and permanently free for everyone to read, download, copy, and distribute for noncommercial purposes.



IIUM Engineering Journal at <https://journals.iium.edu.my/ejournal> is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

Published by IIUM Press, International Islamic University Malaysia
Published triannually in January, May, and September.

IIUM ENGINEERING JOURNAL EDITORS

CHIEF EDITOR

Ahmad Faris Ismail, IIUM, Malaysia

TECHNICAL EDITOR

Khairul Azami Sidek, IIUM, Malaysia

EXECUTIVE EDITOR

Teddy Surya Gunawan, IIUM, Malaysia

ASSOCIATE EDITOR

Nor Farahidah Za'bah, IIUM, Malaysia

COPY EDITOR

Aliza 'Aini Md Ralib, IIUM, Malaysia

Suriza Ahmad Zabidi, IIUM, Malaysia

MALAY TRANSLATOR

Nurul Arfah Che Mustapha, IIUM, Malaysia

EDITORIAL BOARD MEMBERS

Abdumalik Rakhimov, IIUM, Malaysia

Ani Liza Asnawi, IIUM, Malaysia

Konstantin Khanin, University of Toronto, Canada

Md Rafiqul Islam, IIUM, Malaysia

Mohamed B. Trabia, United States

Mohammad S. Alam, Texas A&M University-Kingsville, United States

Mohammed Saedi Jami, IIUM, Malaysia

Mohd Hanafi Ani, IIUM, Malaysia

Mohd Sultan Ibrahim Shaik Dawood, IIUM Malaysia

Mustafizur Rahman, National University Singapore, Singapore

Nor Fadhillah Mohamed Azmin, IIUM, Malaysia

Norun Fariyah Abdul Malik, IIUM, Malaysia

Ossama Abdulkhalik, Michigan Technological University, United States

Sheroz Khan, Saudi Arabia

Tanveer Saleh, IIUM, Malaysia

AIMS & SCOPE OF IIUMENGINEERING JOURNAL

The **IIUM Engineering Journal**, published biannually (January and July), is a carefully refereed international publication of the International Islamic University Malaysia (IIUM). Contributions of high technical merit within the span of engineering disciplines covering the main areas of engineering: Electrical and Computer Engineering; Mechanical and Manufacturing Engineering; Automation and Mechatronics Engineering; Material and Chemical Engineering; Environmental and Civil Engineering; Biotechnology and Bioengineering; Engineering Mathematics and Physics; and Computer Science and Information Technology are considered for publication in this journal. Contributions from other areas of engineering and applied science are also welcome. The IIUM Engineering Journal publishes contributions under *Regular papers and Invited review papers*. It also welcomes contributions that address solutions to the specific challenges of the developing world, as well as science and technology issues, from an Islamic and multidisciplinary perspective.

REFEREES' NETWORK

All papers submitted to the IIUM Engineering Journal will be subjected to a rigorous reviewing process through a worldwide network of specialized and competent referees. Each accepted paper should have at least two positive referees' assessments.

SUBMISSION OF A MANUSCRIPT

A manuscript should be submitted online to the IIUM-Engineering Journal website at <https://journals.iium.edu.my/ejournal>. The journal website can also be used for further correspondence on the paper's status.

INTERNATIONAL ADVISORY COMMITTEE

A. Anwar, United States
Abdul Latif Bin Ahmad, Malaysia
Farzad Ismail, USM, Pulau Pinang, Malaysia
Hanafy Omar, Saudi Arabia
Hany Ammar, United States
Idris Mohammed Bugaje, Nigeria
K.B. Ramachandran, India
Kunzu Abdella, Canada
Luis Le Moyne, ISAT, University of Burgundy, France
M Mujtaba, United Kingdom
Mohamed Al-Rubei, Ireland
Mohamed B Trabia, United States
Syed Kamrul Islam, United States
Tibor Czigany, Budapest University of Technology and Economics, Hungary
Yiu-Wing Mai, The University of Sydney, Australia.

Published by:



**IIUM
Press**

IIUM Press,

International Islamic University Malaysia
Jalan Gombak, 53100 Kuala Lumpur, Malaysia
Phone (+603) 6421-5014, Fax: (+603) 6421-6298

Whilst every effort is made by the publisher and editorial board to see that no inaccurate or misleading data, opinion or statement appears in this Journal, they wish to make it clear that the data and opinions appearing in the articles and advertisement herein are the responsibility of the contributor or advertiser concerned. Accordingly, the publisher and the editorial committee accept no liability whatsoever for the consequence of any such inaccurate or misleading data, opinion, or statement.

ISSN 1511 - 788X



IIUM Engineering Journal

ISSN: 1511-788X E-ISSN: 2289-7860

Editorial Message Vol. 27 No. 3 2026

From Intelligent Models to Responsible Engineering Systems

It is our pleasure to present the September 2026 issue of the IIUM Engineering Journal. This issue captures an important transition in contemporary engineering: the movement from isolated models, devices, and analyses toward integrated systems that are expected to be intelligent, efficient, sustainable, reliable, and useful in real operating environments. Across the papers in this volume, advanced computation is increasingly coupled with physical engineering, while experimental validation, resource awareness, safety, and practical deployment remain central.

A defining feature of this issue is the breadth of data-driven engineering. Machine learning and artificial intelligence are applied to EEG classification, autism-spectrum assessment, intrusion detection, pipeline condition monitoring, sentiment analysis, object detection, microgrid energy management, image restoration, agricultural vision, asphalt characterization, and plastic identification. These studies illustrate how AI is becoming embedded within engineering workflows rather than treated as a separate computational layer. Equally important, several contributions address explainability, efficiency, synthetic-data augmentation, knowledge distillation, real-time operation, and compute-aware evaluation: issues that increasingly determine whether intelligent methods can be trusted and deployed responsibly.

The issue also demonstrates the continuing importance of modeling, optimization, and digital engineering. Finite-element analysis supports the design of interdigitated electrodes and strengthened concrete structures; numerical and predictive approaches address progressive collapse, reinforced slopes, hydraulic-drive stability, fatigue life, and microgrid operation; and digital-twin and surrogate-assisted optimization frameworks extend advanced computation into manufacturing and autonomous systems. Together, these works show how simulation and optimization can reduce development cost, improve design confidence, and accelerate the transition from concept to implementation.

Sustainability is another strong thread. Research on plastic-waste pyrolysis, activated carbon derived from palm empty fruit bunches, riverbank erosion and accretion, low-cost spectroscopy for plastic identification, and energy management for residential microgrids reflects engineers' growing responsibility to address resource use, environmental risk, and circularity. Sustainability here is not presented as an abstract objective, but as a measurable engineering problem requiring better materials, processes, sensing, control, and decision support.

The human and societal dimensions of engineering are equally visible. Work on rehabilitation exoskeletons, mobility and stability systems, biomedical signal analysis, and safe infrastructure reminds us that engineering performance must ultimately be judged by its contribution to human well-being. Even as systems become more autonomous and computationally sophisticated, usability, robustness, interpretability, and safety remain indispensable.

Taken together, the papers in Vol. 27 No. 3 demonstrate a mature and increasingly integrated view of engineering research: one that combines intelligence with physical understanding, sustainability with performance, and innovation with responsibility. This is the direction the IIUM Engineering Journal seeks to advance: rigorous research that crosses disciplinary boundaries, responds to real-world needs, and contributes meaningfully to technological and societal progress.

On behalf of the editorial team, I thank all authors, reviewers, and editorial board members for their contributions and commitment to scholarly quality. We hope this issue stimulates new ideas, strengthens interdisciplinary collaboration, and encourages engineering solutions that are not only technically advanced, but also responsible, resilient, and relevant.

Prof. Ir. Ts. Dr. Teddy Surya Gunawan

Executive Editor

IIUM Engineering Journal

IIUM ENGINEERING JOURNAL

Volume 27, Issue 3, September 2026
<https://doi.org/10.31436/iiumej.v27i3>

Table of Contents

Copyright	i
Editorial Board	ii
Editorial Message	iv
 Chemical and Biotechnology Engineering	
3191: Optimization Of Interdigitated Electrode Through Finite Element Analysis For Juice Adulteration Detection	1
<i>Sakinah Abdul Hamid, Mohd Asyraf Mohd Razib, Marmeezee Mohd Yusoff</i>	
3489: Pretreatment of Plastic Waste and Optimization of the Pyrolysis Process for Liquid Fuel Production	14
<i>Norsyamira Liyana Sulaiman, Yusilawati Ahmad Nor</i>	
 Civil and Environmental Engineering	
4347: GIS-Based Geospatial Assessment of Net Land Surface Change for Identifying Erosion-Accretion Hotspots Along the Padma River in Bangladesh.....	29
<i>Md. Tarequl Islam, Md. Sameer Hossain Talukder, Md. Sakib Hossain Talukder, Md. Sultanul Islam</i>	
4517: Repeated Loading Behavior of CFRP-Strengthened Continuous RC Beams: Experimental Investigation, Finite Element Analysis, and Predictive Modeling	47
<i>Zeinab Moubarek, Ehab Lotfy, Manar Ahmed, Erfan Latef, Ezzaat Sallam</i>	
4533: Machine-Learning Prediction of Marshall Mix-Design Properties and Mechanical Characterization of DCR-Treated Asphaltic Concrete.....	70
<i>Nur Atikah Kamaruzaman, Wan Nur Aifa Wan Azahar, Altamashuddin Khan Nadimalla, Norhidayu Kasim, Siti Aliyah Masjuki, Assayidatul Laila Nor Hairin, Norfarah Nadia Ismail, Mastura Bujang</i>	
4607: Progressive Collapse Analysis on 2D RC Frames under Impact of Single Interior and Exterior Columns	89
<i>Tahara Ramadhan Md Kassim, Muhammad Faiz Mohamad Nor, Taksiah A. Majid</i>	
4705: Numerical Evaluation of Geogrid-Reinforced Slopes under Various Soil and Geometric Configurations.....	102
<i>Adibah Farhana Hairul Hisham, Norhidayu Kasim, Wan Nur Aifa Wan Azahar</i>	
 Electrical, Computer and Communications Engineering	
4079: EEG Signal Classification under RF Exposure Using Synthetic Data–Augmented Feedforward Neural Network	117
<i>Muhammad Rasyid Rosli, Roshakimah Mohd Isa, Nur Irsalina Huda Nazri</i>	
4187: Optimal Operating Cost for A Grid-connected Microgrid Using Quadratic Interpolation Optimization	131
<i>Thach Chi Phan, Thuan Thanh Nguyen, Thang Trung Nguyen</i>	
4191: A Hybrid Framework for Real-Time Sentiment Analysis and Adaptive Conversational Outcome Prediction in Domain-Specific Dialogue Systems	149
<i>Mejdl Safran</i>	

4333: Regression-Based Optical Intensity and Image Processing Techniques for Diatom Algae Content Estimation	173
<i>Rahmat Maghrobi Alwafi, Agus Indra Gunawan, Setiawardhana</i>	
4344: Efficient Edge-Based Object Detection via Multi-Signal Teacher–Student Knowledge Distillation.....	186
<i>Richa Sharma, Ashutosh Gupta, Prasun Chakrabarti, Arvind Sharma</i>	
4355: Pipe-to-Soil Potential State Classification For Equatorial Gas Pipeline ICCP Systems Under Geomagnetically Induced Current Influence Using XGBoost Algorithm	214
<i>Maizura Mohd Sani, Ahmad Ihsan Mohd Yassin, Mohamad Huzaimy Jusoh, Nurhani Kasuan, Nur Dalila Khirul Ashar, Muhammad Asraf Hairuddin</i>	
4522: A CatBoost-Based Framework for Autism Spectrum Disorder Classification from EEG Signals with Blackman–Harris FIR Filtering.....	230
<i>Buleun Waladah, Melinda, Yunidar, Aufa Rafiki, Siti Rusdiana, Wei Kitt Wong</i>	
4534: Performance Evaluation of Boosting Algorithms on Intrusion Detection System Based on Real-Time CICIOT2023 Dataset	250
<i>Muhammad Hafiz Amrullah, Favian Dewanta, Muhamad Erza Aminanto</i>	
4635: A Predictive Energy Management Framework for Malaysian Residential Microgrid Using LSTM-Driven Adaptive Rule-Based Control.....	266
<i>Kamal Akashah Che Kamaruddin, Siti Hajar Yusoff, Teddy Surya Gunawan, Sarah Yasmin Mohamad, Aisha Hassan Abdalla Hashim, Nur Shahida Midi</i>	
4662: DesnowOracle: Quantifying Mask-Oracle Gap Recovery by a Training-Free Snow Prior in Compact Image Desnowing	296
<i>Teddy Surya Gunawan, Kartiwi Mira</i>	
4689: Compute-Aware Evaluation of Severity Conditioning and Lightweight Architectures for Real-World Image Deraining	319
<i>Teddy Surya Gunawan, Mira Kartiwi</i>	
4693: Similar Accuracy, Different Explanations in YOLO-Based Palm Fruit Ripeness Detection	340
<i>Queenta Paradissa Ramadhani, Favian Dewanta, Teddy Surya Gunawan, Mira Kartiwi, Bambang Setia Nugroho</i>	

Engineering Mathematics and Applied Science

4503: CFD Analysis of the Effect of a Ball Throttle on the Hydraulic Drive Stability	366
<i>Gulsara Annakulova, Sarvar Saidov, Bekzod Astanov</i>	
4618: A TRIZ-Based Conceptual Design of an Automatic Footrest Stand to Enhance Lateral Stability of the Flexible Unicycle Prototype During Mounting and Dismounting	376
<i>Noorhidayah Ramli, Muhammad Azim Azizi, Muhammad Amin Azman, Nuraini Ab Aziz</i>	

Materials and Manufacturing Engineering

4388: A Modular, 3D-Printed, and Cost-Effective Near-Infrared Spectrometer with Machine Learning for Commodity Plastic Identification	403
<i>Wei Jim Tan, Zulkifli Mohamad Ariff, Arjulizan Rusli, Muhamad Sharan Musa</i>	
4576: Kansei-Based AHP Framework for Feeding Mechanism Concept Selection in Agroindustrial Coconut Deshelling Machine Design	416
<i>Hety Handayani Hidayat, Bambang Purwantana, Radi, Mirwan Ushada</i>	
4626: Surface Treatment for Improved Electrochemical Performance of Activated Carbon Derived from Palm Empty Fruit Bunches.....	431
<i>Hasan Marzuki, Alya Naili Rozhan, Mohd Hanafi Ani, Hadi Purwanto</i>	

Mechanical and Aerospace Engineering

- 3409: Comparison of Fatigue Crack Growth Models for Predicting Residual Life 443
Muhamad Husnain Mohd Noh, Mohd Akramin Mohd Romlay, Mohd Shamil Shaari, Takahashi Akiyuki

Mechatronics and Automation Engineering

- 4276: Digital Twin Framework for Integrated Sequential Laser Beam Micromachining (LBMM) and Micro-EDM 460
Wan Ahmad Wan Azhar, Tanveer Saleh, Khairul Affendy Md Nor, Muhammad Mukhtar Noor Awalludin, Abd Halim Embong, Ali Sophian, Mohd Asyraf Mohd Razib, Azhar Mohd Ibrahim
- 4508: Budget-Efficient PID Tuning of a Line-Following Differential-Drive AGV Using Extra Trees Surrogate-Assisted CMA-ES..... 481
Mastang, Bakr Ahmed Taha, Retna Apsari, Muhammad Mukhlisin, Eni Dwi Wardihani, Norhana Arsad
- 4578: Proof-of-Concept Validation in Healthy Participants of an sEMG-Controlled 4-DOF Upper Limb Exoskeleton for Post-Stroke Rehabilitation 505
Ang Ling Weay, Tang Ming

Blank page.

Optimization of Interdigitated Electrode through Finite Element Analysis for Juice Adulteration Detection

SAKINAH ABDUL HAMID¹, MOHD ASYRAF MOHD RAZIB^{1,2*},
MARMEEZEE MOHD YUSOFF¹

¹Department of Mechatronic Engineering, Kulliyah of Engineering, International Islamic University of Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

²VLSI-MEMS Research Unit, Department of Electrical and Computer Engineering, Kulliyah of Engineering, International Islamic University of Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

*Corresponding author: asyrafr@iium.edu.my

(Received: 02 April 2024; Accepted: 6 August 2026; Published online: 11 September 2026)

ABSTRACT: Recently, interdigitated electrodes (IDEs) have been applied in various industries because they are easy to fabricate and low-cost. However, the optimal electrode geometry for maximizing sensitivity remains unclear, as conflicting reports exist on the effects of finger width, spacing, and thickness under varying operating conditions. Therefore, a simulation study was conducted using COMSOL Multiphysics 5.6 software to determine which IDE offers the best sensitivity. The IDEs operate with two parallel coplanar electrodes, with periodic repetition of the design (i.e., length, width, and electrode spacing) to generate a high signal and reduce the signal-to-noise ratio. Finite element analysis via COMSOL revealed that an optimized configuration featuring a smaller gap size (0.1 mm), a wider finger width (0.4 mm), and an electrode count of 8 yielded an ideal sensor design, achieving a peak electric field of 2582.52 V/m and a current density of 1.57×10^{10} A/m². Following the computational analysis, the optimized IDE sensor was physically fabricated. Experimental verification using this low-cost, screen-printed silver sensor successfully demonstrated a distinct contrast in electrical resistance; pure fruit juices exhibited a low baseline resistance between 2.5 and 6 Ω , whereas commercialized variants reached significantly higher values of up to 35 Ω over a 5-hour testing window. This strong correlation highlights the platform's stability and its immediate capacity to distinguish juice purity. These findings confirm that the optimized IDE platform provides a highly viable, inexpensive solution for detecting food and beverage adulteration.

ABSTRAK: Kebelakangan ini, elektrod berselang-seli (IDE) telah digunakan secara meluas dalam pelbagai industri kerana mudah difabrikasi dan berkos rendah. Walau bagaimanapun, geometri elektrod yang optimum bagi memaksimum kepekaan masih tidak jelas, kerana terdapat laporan bercanggah mengenai kesan lebar jari, jarak, dan ketebalan di bawah keadaan operasi berbeza. Oleh itu, satu kajian simulasi telah dijalankan menggunakan perisian COMSOL Multiphysics 5.6 bagi menentukan konfigurasi IDE berkepekaan terbaik. IDE beroperasi berdasarkan dua elektrod koplanar selari, dengan pengulangan berkala parameter reka bentuk iaitu panjang, lebar, dan jarak antara elektrod, bagi menjana isyarat tertinggi dan mengurangkan nisbah isyarat-ke-bunyi. Analisis unsur terhingga melalui COMSOL mendapati bahawa konfigurasi optimum dengan saiz jurang lebih kecil (0.1 mm), lebar jari lebih lebar (0.4 mm), dan bilangan elektrod sebanyak 8 telah menghasilkan reka bentuk sensor ideal, dengan mencapai medan elektrik puncak sebanyak 2582.52 V/m dan ketumpatan arus sebanyak 1.57×10^{10} A/m². Selepas analisis pengiraan, sensor IDE yang dioptimumkan telah difabrikasi secara fizikal. Pengesahan eksperimen menggunakan sensor perak cetakan skrin berkos rendah ini berjaya menunjukkan perbezaan ketara dalam rintangan elektrik; jus buah-

buahen tulen menunjukkan rintangan asas yang rendah antara 2.5Ω dan 6Ω , manakala varian komersial mencapai nilai yang lebih tinggi sehingga 35Ω dalam tempoh ujian selama 5 jam. Korelasi yang kuat ini menyerlahkan kestabilan platform dan keupayaannya yang segera bagi membezakan ketulenan jus. Penemuan ini mengesahkan bahawa platform IDE yang dioptimumkan menyediakan penyelesaian praktikal dan murah bagi mengesan pemalsuan makanan dan minuman.

KEYWORDS: *Interdigitated electrode (IDE), Finite Element Analysis, Biosensor*

1. INTRODUCTION

Interdigitated Electrodes (IDEs) are universal transducers in modern technologies that fall under the category of electrochemical biosensors featuring micro- or nanoscale terminal arrays [1]. These biosensors can be readily integrated into mobile lab-on-chip devices, enabling rapid point-of-care diagnosis. Compared with other impedimetric biosensing configurations, this architecture offers excellent sensitivity, low cost, and ease of fabrication, which are the primary advantages of these sensors [2]. Furthermore, IDEs eliminate the operational requirement for a reference electrode while providing a straightforward methodology for obtaining a steady-state current response. Consequently, such sensors have been widely utilized to monitor capacitance, impedance, dielectric constant, and mass conductivity in biological and chemical media [3].

Sensing devices utilizing IDEs can operate in either faradaic or non-faradaic modes. The distinction between faradaic and non-faradaic processes is essential for characterizing sensor performance. A faradaic process is fundamentally governed by transient electron transfer across the electrode-electrolyte interface via chemical oxidation and reduction reactions, whereas non-faradaic modes rely on changes in capacitance such as double-layer capacitance and physical molecular binding events at the electrode surface without charge transfer [4]. When implementing capacitive or impedimetric sensors, IDEs offer an exceptionally effective planar structure. Conventionally, an IDE has a coplanar architecture composed of two individually addressable, comb-like electrode structures that form a precise gap between them. In these configurations, fringe or edge field effects become less dominant when the finger-length dimensions significantly exceed the electrode layer thickness [5].

Mathematical modeling and finite element simulation are essential engineering tools before physical fabrication, as they minimize the time and cost of prototyping. Recent studies have demonstrated the effectiveness of finite element analysis in optimizing IDE performance by evaluating the influence of key parameters such as gap size, finger width, and electrode count on sensor sensitivity [6]. Furthermore, modeling enables a deeper evaluation of system physics under varying operational constraints, providing a clear pathway for design optimization. To this end, this study aims to optimize the geometric configuration of an IDE sensor via finite element analysis to achieve maximum sensitivity, followed by physical realization using a cost-effective manufacturing route. This work systematically evaluates the influence of gap size, finger width, electrode count, and material types on localized electric-field distribution and current-density profiles. By translating optimized simulation topologies into actual, flexible silver-paste electrodes screen-printed onto polyethylene terephthalate (PET) substrates, this study demonstrates a scalable framework for rapid, portable beverage quality assurance. Ultimately, this work provides an affordable, high-sensitivity electrochemical testing platform capable of robustly distinguishing pure fruit juices from commercialized or adulterated alternatives, addressing a critical need for accessible quality control in the beverage industry [7].

2. METHODOLOGY

2.1. Biosensor Structure

An IDE is a coplanar structure composed of two individually addressable interdigitated combs that form a micro-sized gap between them, as shown in Fig. 1. The comb-like structure is a convenient, simple, and low-cost design that can be used in a variety of applications [8]. IDEs operate based on two parallel-plate capacitors and a sample under test. Capacitance describes how two conductive objects separated by a space respond to a voltage difference applied to them. An electric field is established between them, causing positive and negative charges to collect on each object. In IDEs, the term "interdigit" refers to the repeated coplanar, digit-like or finger-like pattern used to generate a high signal and reduce the signal-to-noise ratio [9, 10]. The three major measurement concepts are changes in dielectric properties, charge distribution, and conductivity when the analyte forms on the IDE surface.

Fig. 1 shows the structure of the IDE, in which one plate serves as the ground (negative terminal). In contrast, the other serves as the electric potential (positive terminal), to which an alternating current (AC) voltage signal is applied to create an electric field between these two terminals [11]. The individual electrode finger width is defined as W_f , which is separated by a specific gap distance, G , from the opposing electrode finger. The total number of alternating finger structures across the active sensing area constitutes the electrode count parameter, N . Fig. 1 shows all parameters. Any medium placed on the sensors increases the electric potential, creating an electric field above the sensor that penetrates the sample under test [12].

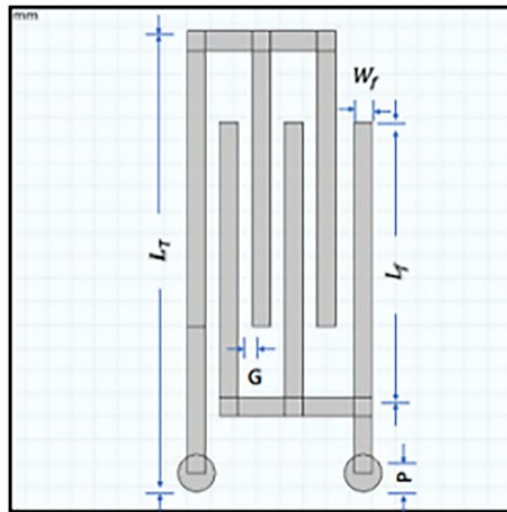


Figure 1. IDE geometry parameters

2.2. Finite Element Simulation

The IDE structure is simulated using COMSOL Multiphysics version 5.6 to investigate the electric field distribution and current density for the proposed parameters [13]. COMSOL is a software platform used to model and simulate multiphysics applications using advanced numerical methods. The simulation results help distinguish the effects of the electric field for various parameters and determine which design is best for this project.

The performance of the IDE's geometrical parameters is evaluated by analyzing electric-field strength and current-density flow. Various IDE geometric layouts are designed in 2D. The electric-field strength generated by the capacitive sensor is a key factor in evaluating sensitivity performance. To perform numerical calculations of electric fields, AC and DC modules need

to be simulated [14]. The working and counter electrodes were simulated with different geometries, such as width, thickness, and electrode gap. The electric field will be generated between the working and ground electrodes [15]. The Ohm's law will be used to begin the numerical calculation as Eq. (1) where J is the current density, σ is the conductivity of DMEM, E is the electric field, D is the electric displacement, and Je is the external current density.

$$J = \sigma E + j\omega D + Je \quad (1)$$

The electric field (E) and electric displacement (D) were obtained from the gradient of voltage (V) at each static point, as in Eq. (2) and Eq. (3), where ϵ_0 and ϵ_r are the permittivity of free space and the permittivity of the electrode, respectively. The equation used for numerically calculating the electrode spacing and capacitance is Eq. (4), where C is the capacitance of the IDE structure, A is the area of the surface sensor, and d is the distance between the electrodes.

$$E = -\nabla V \quad (2)$$

$$D = \epsilon_0 \epsilon_r E \quad (3)$$

$$C = \epsilon_0 \epsilon_r \frac{A}{d} \quad (4)$$

In this study, four simulations are being performed: the gap size between the electrodes, the width of the fingers, the number of electrodes, and the material used. The first simulation involves adjusting the gap between the electrodes from 0.1 mm to 0.5 mm (increasing by 0.1 mm each time). The relationship between finger width and performance varied from 0.2 mm to 0.4 mm (increasing by 0.05 mm each time). The simulation then increased the number of fingers from four to twelve (increasing by two fingers each time). The final simulation will test silver (Ag), gold (Au), and platinum (Pt) substrate materials. Table 1 shows the geometry parameters of the IDEs simulated in this study.

Table 1. IDE Design Parameters

Parameter	Symbol	Values	Unit
Total IDE Length	L_T	10.0	mm
Padding Size	P	$\pi * 0.4^2$	mm ²
Gap between Fingers	G	0.1 – 0.5	mm
Finger Width	W_f	0.2 – 0.4	mm
Number of Fingers	N	4 -12 fingers	
Finger Length	L_f	6	mm

The sequential optimization workflow, proceeding from electrode gap size (G), to finger width (W_f), electrode finger numbers (N), and finally electrode material selection, was deliberately structured based on a physics-driven hierarchy of parameter dominance. The local electric field intensity (E) and primary charge transport dynamics are most sensitively governed by the immediate spatial proximity between opposing conductors (G), as dictated by the electrostatic field gradient ($E = -\nabla V$) [16]. Once the critical field-generating gap was established, the electrode width (W_f) was optimized to maximize the active surface area and current density vectors [17]. Following these micro-scale boundary configurations, the macro-scale structural parameter (N) was evaluated to optimize total capacitance and signal-to-noise ratio within the physical footprint constraints. Finally, material types were assessed as a secondary factor to balance processing stability and electrical performance. This sequence approach ensures that primary local-field behaviors are stabilized before secondary-scale parameters are tuned.

2.2.1. Computational Preprocessing and Boundary Conditions

To execute the numerical physics calculations, the following preprocessing parameters and boundary settings were established within the COMSOL Multiphysics environment. For the physics interface, the Electrostatics (*es*) and Electric Currents (*ec*) interfaces within the AC/DC module were utilized to resolve the domain equations [18]. For boundary conditions, a terminal voltage of 1 V ($V_0 = 1$ V) was applied as a Dirichlet boundary condition to the working electrode terminal array. The opposing counter electrode array was assigned a ground boundary condition ($V = 0$ V). All outer boundaries of the surrounding computational domain matrix were set to Charge Conservation with Zero Charge/Electric Insulation constraints. For material properties, the electrodes were modeled using the bulk electrical conductivities and isotropic relative permittivities (ϵ_r) corresponding to silver (Ag), gold (Au), and platinum (Pt) domains. The fluid medium domain filling the electrode gaps was assigned the material properties of the target juice matrix with predefined baseline electrical conductivities (σ) and relative dielectric constants. Lastly, a physics-controlled, triangular finite element mesh sequence was implemented. To ensure accurate convergence and capture extreme field gradients, localized mesh refinement was enforced along all sharp boundaries, fingertips, and inter-electrode gap spaces (G), yielding a highly dense grid configuration within the active sensing area.

For boundary conditions, a terminal voltage of 1 V ($V_0 = 1$ V) was applied as a Dirichlet boundary condition to the working electrode terminal array. The opposing counter electrode array was assigned a ground boundary condition ($V = 0$ V). All outer boundaries of the surrounding computational domain matrix were set to Charge Conservation with Zero Charge/Electric Insulation constraints. For material properties, the electrodes were modeled using the bulk electrical conductivities and isotropic relative permittivities (ϵ_r) corresponding to silver (Ag), gold (Au), and platinum (Pt) domains. The fluid medium domain filling the electrode gaps was assigned the material properties of the target juice matrix with predefined baseline electrical conductivities (σ) and relative dielectric constants. Lastly, a physics-controlled, triangular finite element mesh sequence was implemented. To ensure accurate convergence and capture extreme field gradients, localized mesh refinement was enforced along all sharp boundaries, fingertips, and inter-electrode gap spaces (G), yielding a highly dense grid configuration within the active sensing area.

2.3. Fabrication of Interdigitated Electrode (IDE) sensor.

Screen printing appeared to be the least expensive technique compared to other printing methods. The mesh with an interdigitated layout is used to transfer silver paste onto the substrate, polyethylene terephthalate (PET), as shown in Fig. 2(a). A squeegee is moved across the screen to fill the open mesh apertures with ink. Lift the mesh quickly to avoid excessive ink absorption. To achieve the optimum curing time for the ink, the printed electrodes will need to cure for 15 minutes in the oven at 80°C temperatures. The physically realized, low-cost silver IDE sensor arrays resulting from this fabrication workflow are displayed in Fig. 2(b).

The experimental testing involved four distinct samples: pure orange juice, commercial orange juice, pure apple juice, and commercial apple juice. For each test, a 1 mL droplet was precisely dispensed onto the fabricated device via a microsyringe, after which an LCR meter measured the resulting resistance and conductance.

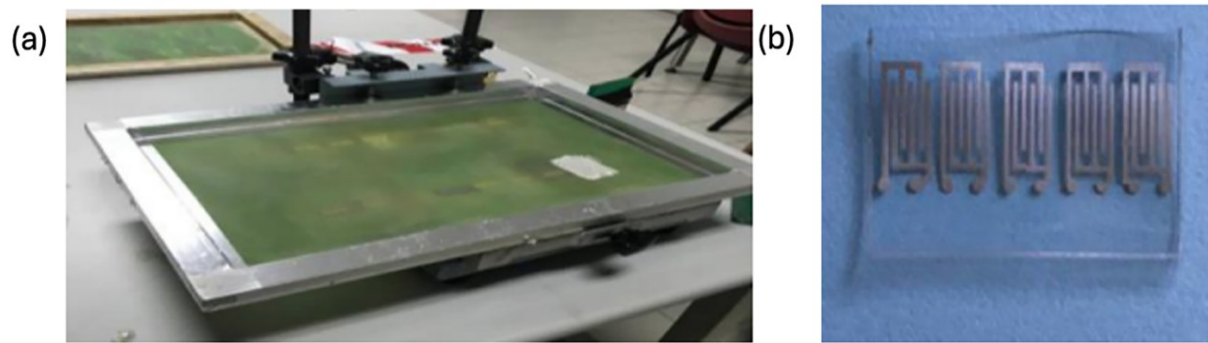


Figure 2. a) Silver ink absorbs onto the substrate through a mesh. b) Screen-printed silver interdigitated electrode (IDE) sensor chips on flexible PET substrates prior to fluid testing.

3. RESULT AND DISCUSSION

Fig. 3 depicts the colormap of the electric field and current density distribution, along with the electric field direction, for the IDE sensor design simulated in COMSOL Multiphysics. The simulation results are based on four research variables: electrode gap size, finger width, number of electrodes, and material type.

The standard unit of the electric field is the volt per meter (V/m). The electric field between the working and counter electrodes is excited when a DC voltage is applied to the working electrode. A direct spatial correlation is observed between the maximum electric field norm and the peak current density. The highest values for both parameters are concentrated along the vertical edges and sharp corners of the interdigitated fingers, particularly within the narrow gaps separating the opposing electrodes (ranging from $y = 0$ mm to $y = 4.5$ mm). As shown in Fig. 3 (left), the electric field norm peaks significantly at approximately 5×10^3 V/m at the tips and localized sharp boundaries of the inner finger channels. This intense localized field drives charge transport. Correspondingly, Fig. 3 (right) reveals that the current density norm reaches its maximum of roughly 3.5×10^{11} A/m² precisely within these high-field channels. The arrow surface plots confirm that the current density vectors are densely packed and tightly confined within the high-field gaps, indicating that the primary conductive pathways are restricted to these regions. Conversely, the peripheral areas, the outer boundary of the substrate, and the regions surrounding the bottom contact pads display both a negligible electric field (~ 0 V/m) and a minimal current density (~ 0 A/m²). This confirms excellent signal confinement to the active interdigitated area with virtually no parasitic leakage current in the surrounding matrix.

The primary current pathways are strictly confined to the narrow vertical channels between the interdigitated electrode fingers. The current density arrows (Fig. 3, right) demonstrate that charge transport flows laterally across the gaps from the high-potential fingers to the grounded fingers [19]. Negligible current flows through the bulk substrate or the outer margins, indicating excellent geometric confinement of the electrical signal. Severe electric field concentrations (hotspots) are observed at the sharp corners and tips of the interdigitated fingers, as well as along the innermost vertical interfaces. The tight confinement of current density to the finger gaps maximizes the device's sensitivity and active-area utilization. However, intense localized electric-field gradients increase the risk of localized dielectric breakdown or arcing across the gaps if the operational voltage exceeds design thresholds.

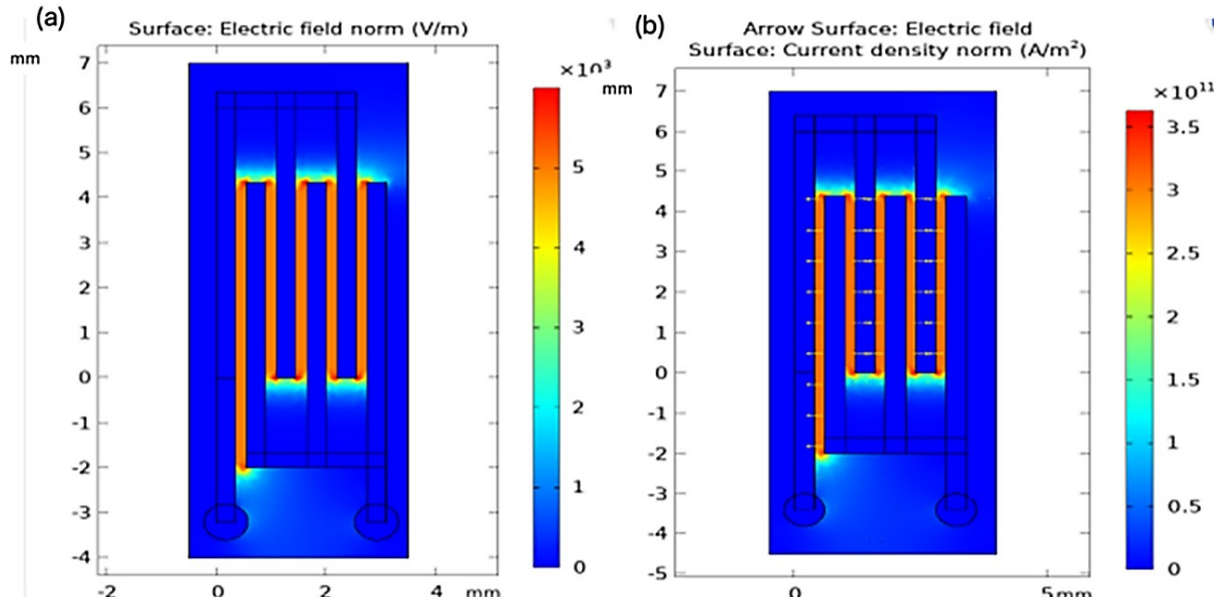


Figure 3. a) Color map of the electric field distribution of IDE. b) Color map of the current density of IDE.

3.1. Dependence on the gap between electrodes

Table 2 summarizes the calculated averages of the current density and the electric field magnitude across varying inter-electrode spacing gaps. As illustrated by the parameter trends in Fig. 4, expanding the gap size from 0.1 mm to 0.5 mm induces a dramatic, non-linear degradation in the average current density. This inverse relationship highlights that a larger spatial gap limits charge transport across the channels. Specifically, the maximum current density is achieved at the minimum evaluated gap of 0.1 mm, yielding a peak value of 1.82×10^{11} A/m², whereas the lowest current density drops sharply to 5.34×10^{10} A/m² at a gap of 0.5 mm.

Concurrently, the average electric field magnitude drops steeply as the finger spacing increases. The minimum electric field intensity of 866.87 V/m is recorded at the largest gap size of 0.5 mm, while the 0.1 mm spacing produces the maximum electric field of 2295.94 V/m. This pronounced field enhancement at smaller gaps is governed by electrostatic field-scaling laws: narrowing the conductor boundaries concentrates the potential drop, intensifying local flux lines [20]. Consequently, the IDE topology featuring a 0.1 mm gap represents the optimal design configuration, providing the maximum localized field intensity and current throughput required to maximize electrochemical sensing sensitivity. Based on these performance advantages, the 0.1 mm gap configuration was selected as the fixed baseline framework for all subsequent optimization stages.

Table 2. The average current density of various gap sizes

Gap size between fingers (mm)	Average Current Density (A/m ²)	Average Electric Field (V/m)
0.1	1.82×10^{11}	2295.94
0.2	1.46×10^{11}	1372.10
0.3	8.74×10^{11}	966.15
0.4	6.17×10^{11}	953.5
0.5	5.34×10^{11}	866.87

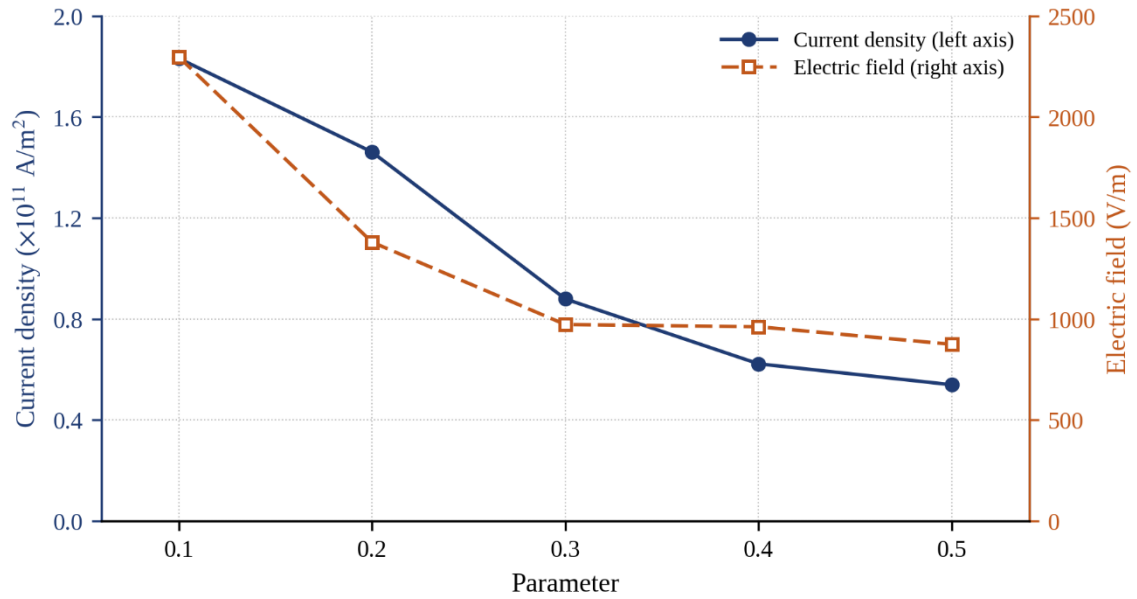


Figure 4. Graph of the average current density and electric field by varying gap sizes.

3.2. Dependence on the width of fingers

The average current density and electric field for IDEs with varying finger widths are presented in Table 3. Overall, the average current density distribution exhibits a gradual increasing trend from the smallest width (0.2 mm) to the largest width (0.4 mm), as demonstrated in Fig. 5. In this simulation, the 0.4 mm width yielded the highest current density value, indicating that wider fingers enhance charge transport efficiency. This behavior can be attributed to the increased electrode surface area available for charge accumulation and electron transfer, which facilitates higher current throughput [21].

The average electric field magnitude as a function of finger width is shown in Fig. 5. Overall, the electric field magnitude increases significantly with increasing electrode width. The highest average electric field of 2468.9 V/m was recorded at 0.4 mm, while the lowest value of 1371.53 V/m was observed at 0.2 mm. This trend follows electrostatic field-scaling principles: wider electrodes provide larger conductive surfaces for charge accumulation, intensifying the potential gradient and local flux lines. Ismail et al. [21] similarly reported that increasing electrode dimensions enhances electric field distribution and penetration depth in interdigitated electrode systems.

Thus, in the second parametric study, the 0.4 mm finger width was identified as the optimal configuration, producing both the highest current density and electric field magnitude. Consequently, this width was selected as the fixed baseline for subsequent simulation studies.

Table 3. The average current density of varying finger widths

Width of fingers (mm)	Average Current Density (A/m ²)	Average Electric Field (V/m)
0.20	8.45×10^{11}	1371.53
0.25	1.25×10^{11}	2023.48
0.30	1.46×10^{11}	2362.85
0.35	1.50×10^{11}	2441.78
0.40	1.52×10^{11}	2468.90

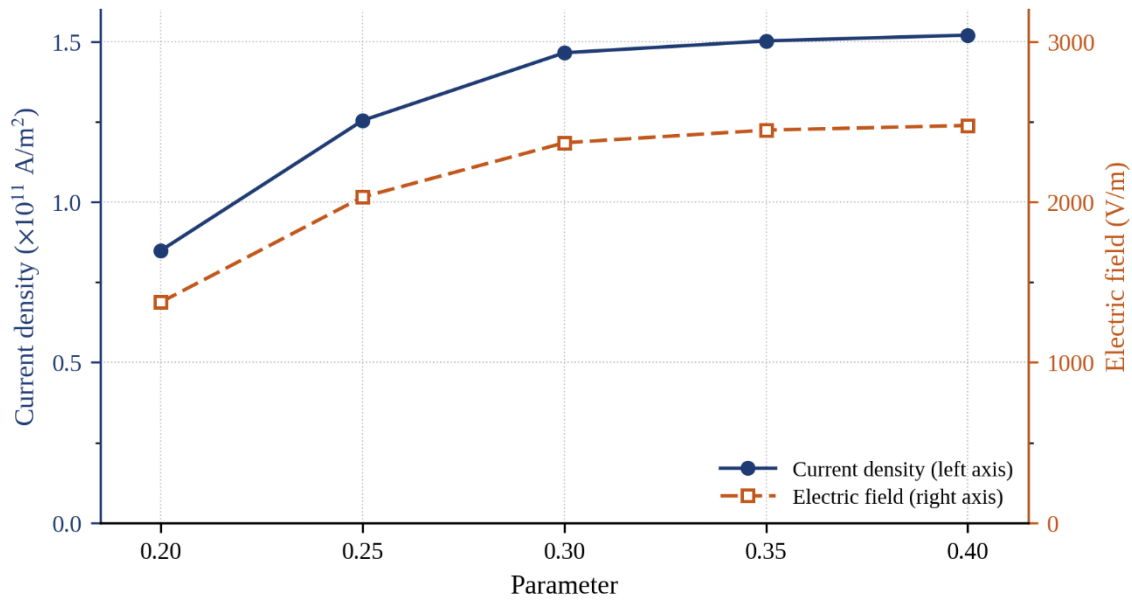


Figure 5. Graph of the average current density and electric field by varying the width of the fingers.

3.3. Dependence on the number of electrodes

Table 4 shows the simulation results for the distribution of current density and electric field magnitude for sensors with different numbers of fingers. Initially, the current density graph (Fig. 6) rose significantly from 4 to 8 electrodes. Regarding the number of fingers, it began at $1.78 \times 10^{10} \text{ A/m}^2$ at 4 electrodes and then dramatically climbed to $8 \times 10^{10} \text{ A/m}^2$ at 8 electrodes before declining to $1.57 \times 10^{10} \text{ A/m}^2$ at 12 electrodes. Thus, the current density is highest with 8 fingers. The graph in Fig. 6 illustrates the electric field increasing steadily between 4 and 8 electrodes, from about 288.67 V/m to 2581.52 V/m. Then, the graph dropped from 8 to 10 electrodes. After that, the number rose slightly to 1486.46 V/m at 12 electrodes. Thus, 8 electrodes recorded the highest electric field value, while 4 electrodes recorded the lowest.

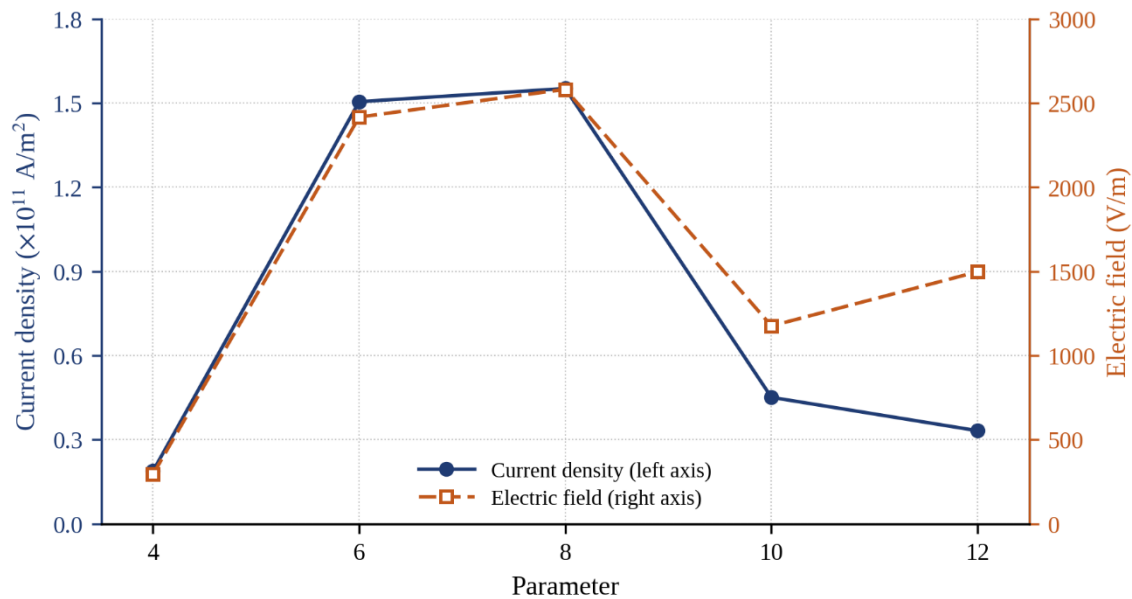


Figure 6. Graph of the average current density and electric field by varying the number of electrodes.

Table 4. The average current density for varying numbers of electrodes

Number of electrodes	Average Current Density (A/m ²)	Average Electric Field (V/m)
4	1.78×10^{10}	288.67
6	1.49×10^{11}	2414.55
8	1.57×10^{11}	2581.52
10	4.41×10^{10}	1160.89
12	3.22×10^{10}	1486.46

3.4. Dependence on the Type of Material

The simulation results for silver, gold, and platinum electrodes show identical current density and electric field distributions, as shown in Fig. 7. Both the average current density and electric field (2581.52 V/m) are the same across all three materials. This outcome is expected because in steady-state electrostatic simulations, the electric field distribution is determined solely by electrode geometry and applied boundary conditions, not by intrinsic material properties [22]. Material conductivity influences current density magnitude only in more complex, time-dependent models.

These three electrode materials are widely used in sensing applications. Platinum offers the highest stability but is prohibitively expensive. Gold provides excellent stability and conductivity; however, its high cost limits practicality for disposable devices. Silver delivers comparable electrical performance at a significantly lower cost, making it the most economically viable choice.

Based on these findings, silver was selected as the optimal electrode material. Despite identical simulation results, silver provides the best balance of high conductivity, cost-effectiveness, and sufficient stability for electrochemical sensing applications. The simulation confirms no performance advantage from using gold or platinum. Therefore, silver was chosen to maximize device performance while maintaining practical economic feasibility for future commercialization.

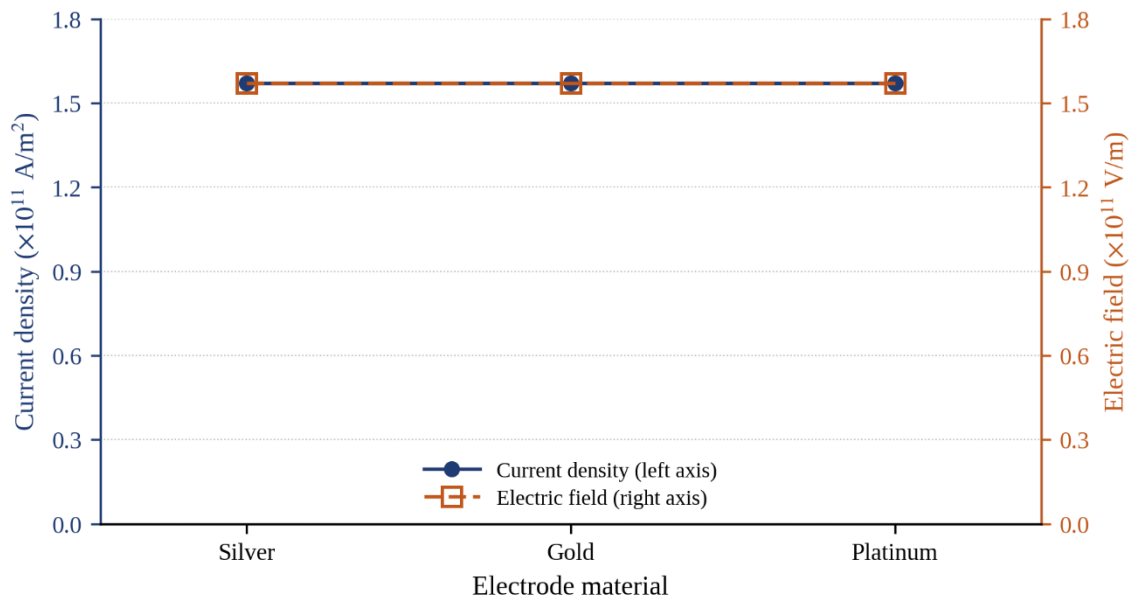


Figure 7. Graph of the average current density and electric field by varying the type of materials

3.5. Experimental Result

All samples were tested at 1 kHz frequency for 5 hours via an LCR meter with a 1 V supply voltage. Fig. 8 shows the resistance of pure and commercialized fruit juices at 1 kHz. Commercialized apple juice exhibits the most significant resistance increase, rising from 16 Ω to 35 Ω , while commercialized orange juice increases modestly from 20 Ω to 23 Ω . Pure apple and orange juices maintain consistently low resistance values between 2.5 Ω and 6 Ω throughout the experiment. This divergence is attributed to compositional differences; commercial juices contain preservatives, added sugars, and stabilizers that alter ionic conductivity and undergo electrochemical degradation over time, causing resistance drift [23]. Pure juices, with natural composition and minimal additives, exhibit stable resistance profiles.

Deionized (DI) water was deliberately excluded from this experiment. DI water has extremely low ionic conductivity, resulting in resistance values that exceed the measurable range of the LCR meter. Even with 1 V excitation, the sensor would produce near-infinite resistance, rendering the graph unreadable and providing no meaningful comparative data. Moreover, DI water lacks the ionic species and organic compounds necessary for electrochemical interaction with the IDE electrodes, making it unsuitable as a reference medium for juice adulteration studies. The fabricated IDE sensor successfully distinguishes between pure and commercial juices based on their distinct resistance signatures, validating its potential for detecting adulteration in commercial fruit juice products.

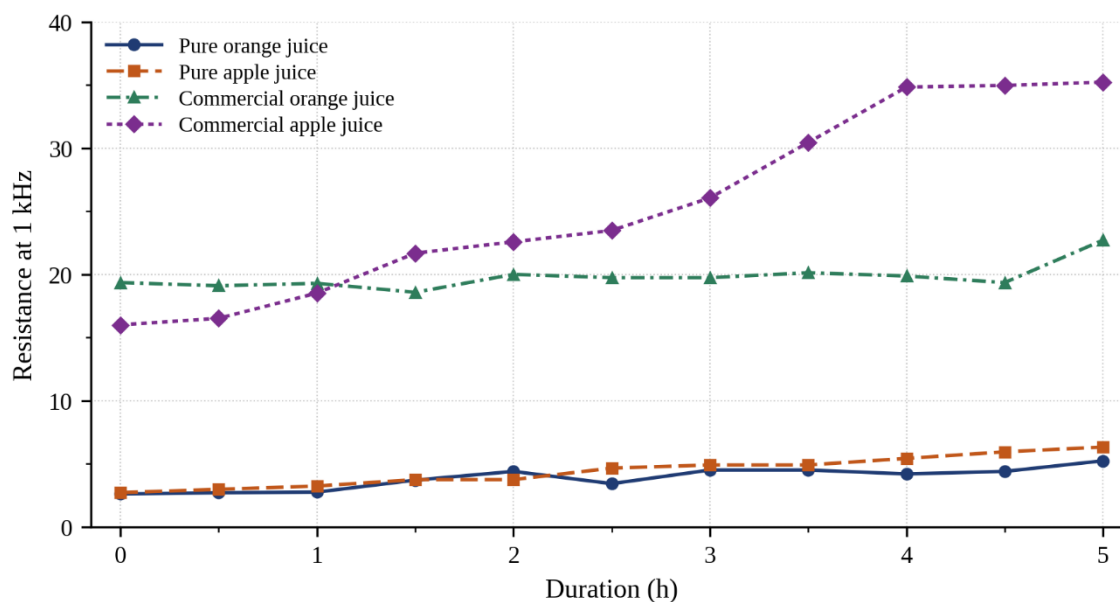


Figure 8. The resistance of pure and commercial fruit juices at 1kHz.

4. CONCLUSION

The electric field distribution was generated using an electrostatics study within the COMSOL AC/DC module for the interdigitated electrode (IDE) simulations. The primary objective of this simulation was to evaluate the influence of key design parameters—namely finger width, gap spacing, finger number, and electrode material—on the resulting electric field and current density distributions. As shown by the simulation results, smaller gap spacings and optimized finger widths produce significantly higher electric field intensities because potential gradients concentrate in confined regions. Therefore, systematically optimizing these geometric parameters can effectively maximize IDE performance for sensing applications.

Among the four parametric studies, finger width and gap spacing were the dominant factors influencing IDE performance, whereas the number of fingers and electrode material type had negligible effects on electric field magnitude. The optimal configuration, featuring a gap spacing of 0.1 mm, a finger width of 0.4 mm, and eight finger pairs, yielded the highest electric field (2582.52 V/m) and current density (1.57×10^{10} A/m²). This enhanced electric field and current density directly correlate with improved sensitivity, as stronger localized fields facilitate greater charge accumulation and redox current at the electrode-electrolyte interface, making this design highly suitable for sensitive electrochemical sensor platforms.

ACKNOWLEDGEMENT

The authors acknowledge funding from the Research Management Center, International Islamic University Malaysia. The authors are very grateful to IIUM for providing machining and testing facilities to carry out the experiments and for their valuable support in this research.

REFERENCES

- [1] S. N. Hosseini et al., "Recent advances in CMOS electrochemical biosensor design for microbial monitoring: Review and design methodology," *IEEE Trans. Biomed. Circuits Syst.*, vol. 17, no. 4, pp. 123–135, Aug. 2023.
- [2] J.-C. Lim, S. C. B. Gopinath, H. Krishnan, and S.-M. Tan, "Faradaic and non-faradaic depletions by anodic electrode corrosion with strong electrolytes on tautochrone curve interdigitated electrodes," *Sci. Direct*, vol. 123, pp. 45–56, 2024.
- [3] A. K. Assaifan et al., "Roles of interdigitated electrode geometry in non-faradaic impedimetric biosensors," *J. Electrochem. Soc.*, vol. 171, no. 8, p. 087515, 2024.
- [4] A. Hedayatipour et al., "CMOS based whole cell impedance sensing: Challenges and future outlook," *Biosens. Bioelectron.*, vol. 143, Art. no. 111600, Oct. 2019.
- [5] M. Z. Iskandarani, "Analytical and numerical modeling of electric field in coplanar devices with neural predictor," *Taylor & Francis Online*, vol. 56, no. 2, pp. 89–102, 2025.
- [6] T. Zhang et al., "A special two-working-electrode system: A summary of recent development in fabrication and application of interdigitated array electrodes," *Nanotechnology*, vol. 36, no. 13, p. 132001, 2025.
- [7] T. Americo da Silva et al., "Application of a microfluidic electronic tongue based on impedance spectroscopy for coconut water analysis," *Food Res. Int.*, vol. 178, p. 113925, 2024.
- [8] T. Zhang, Y. Jian, Z. Deng, W. Wu, and F. Liu, "A special two-working-electrode system: A summary of recent development in fabrication and application of interdigitated array electrodes," *Nanotechnology*, vol. 36, no. 13, p. 132001, 2025.
- [9] Habboush, S., Rojas, S., Rodríguez, N., & Rivadeneyra, A. (2024). The Role of Interdigitated Electrodes in Printed and Flexible Electronics. *Sensors*, vol. 24, no. 9, 2717.
- [10] B. Sarkar, M. T. Rahman, M. M. Rahman, S. Hossen, L. Zhenguo, and N. H. Bhuiyan, "Amplification of electrochemical signal by moving interdigitated electrode array for highly sensitive detection of Alzheimer's disease markers in blood," *Biosens. Bioelectron.*, vol. 269, p. 116939, 2025.
- [11] Hashim, U., Rijal, S., Dhahi, Th. S., Afnan Uda, M. N., Uda, M. N. A., Adam, T., & Pamungkas, Y. "Surface-specific reactivity and ionic selectivity in aluminium interdigitated electrode platforms for advanced electronic sensing interfaces," *Results Eng.*, vol. 28, p. 107325, Dec. 2025.
- [12] L. M. A. Melo et al., "Electrochemistry of the synthetic tryptamine 5-MeO-MiPT at glassy carbon and screen-printed electrodes: A rapid and simple screening method for application in forensic analysis," *Sens. Actuators B Chem.*, vol. 393, p. 134183, Oct. 2023.

-
- [13] M. B. Kulkarni and S. Goel, "Recent advancements in finite element analysis of interdigitated electrode sensors: A review," *Sens. Actuators A Phys.*, vol. 345, p. 113823, 2022. doi: 10.1016/j.sna.2022.113823.
 - [14] M. F. Hossain, J. W. Park, and M. Kim, "Simulation and optimization of interdigitated electrode geometry for enhanced electrochemical biosensing performance using COMSOL Multiphysics," *J. Electrochem. Sci. Technol.*, vol. 14, no. 2, pp. 165–175, 2023. doi: 10.33961/jecst.2022.00845.
 - [15] P. Sharma and A. Kumar, "Finite element analysis of interdigitated capacitive sensors: Effect of electrode geometry on electric field distribution and sensitivity," *IEEE Sensors J.*, vol. 24, no. 8, pp. 11892–11901, Apr. 2024. doi: 10.1109/JSEN.2024.3365237.
 - [16] A. K. Assaifan, R. A. Almansour, J. A. Alessa, S. Alhudaithy, A. S. Fakhouri, and A. A. Alsaleh, "Roles of interdigitated electrode geometry in non-faradaic impedimetric biosensors," *J. Electrochem. Soc.*, vol. 171, no. 8, p. 087515, 2024. doi: 10.1149/1945-7111/ad6a99.
 - [17] A. K. Assaifan, R. A. Almansour, J. A. Alessa, A. Altinawi, and S. Alhudaithy, "Impact of interdigital electrode count on non-faradaic electrochemical biosensor performance for detecting pathogens associated with newborn disabilities," *J. Electrochem. Soc.*, vol. 172, no. 5, p. 057515, 2025. doi: 10.1149/1945-7111/addb85.
 - [18] COMSOL Multiphysics® v. 5.6 AC/DC Module User's Guide. Stockholm, Sweden: COMSOL AB, 2020.
 - [19] S. S. Deshpande, G. Bacher, and S. Bhand, "Design and Development of Micro Interdigitated Electrodes Devices for Sensitive and Reliable Biosensors," Ph.D. dissertation, Birla Institute of Technology and Science, Pilani, India, 2023.
 - [20] V. R. Naganaboina, S. Bonam, M. Anandkumar, A. S. Deshpande, and S. G. Singh, "Improved chemiresistor gas sensing response by optimizing the applied electric field and interdigitated electrode geometry," *Sensors Actuators A: Phys.*, 2023.
 - [21] J. Zhang, V. Selamneni, B. P. Yalagala, M. Amjadi, and H. Heidari, "Development of Dielectrophoresis Electrodes for Nanowire Alignment," *Eng. Proc.*, vol. 127, no. 1, p. 9, Mar. 2026, doi: 10.3390/engproc2026127009.
 - [22] F. Li, C. Zhou, and A. Klinkova, "Simulating electric field and current density in nanostructured electrocatalysts," *Phys. Chem. Chem. Phys.*, vol. 24, no. 42, pp. 25695–25719, Oct. 2022, doi: 10.1039/d2cp02846h.
-

Pretreatment of Plastic Waste and Optimization of the Pyrolysis Process for Liquid Fuel Production

NORSYAMIRA LIYANA SULAIMAN, YUSILAWATI AHMAD NOR*

*Department of Chemical Engineering and Sustainability, Kulliyyah of Engineering,
International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia.*

**Corresponding author: yusilawati_ahmadnor@iium.edu.my*

(Received: 30 October 2024; Accepted: 21 June 2026; Published online: 11 September 2026)

ABSTRACT: This study investigates the conversion of polypropylene (PP), polyethylene terephthalate (PET), and high-density polyethylene (HDPE) into liquid fuel through pyrolysis, emphasizing the mechanistic effects of mechanical and chemical pretreatment. As global plastic waste generation reaches critical levels, sustainable upcycling technologies are urgently required to support the circular economy and mitigate escalating environmental contamination. Therefore, this study aimed to optimize pyrolysis parameters and maximize liquid-fuel yield by comparing untreated plastic feedstocks with ozonated, crushed, and combined ozonated-crushed samples. Pretreatment methods included ozonation and crushing, followed by pyrolysis under varying conditions of temperature (300 to 400 °C), heating rate (5 to 15 °C/min), holding time (1 to 3 h), and ozonation duration (0.5 to 1.5 h). For the first objective, the highest yield (46.6%) was achieved using ozone-treated PP at 400 °C, a heating rate of 5 °C/min, a holding time of 1 h, and an ozonation time of 1 h. Mechanistically, ozonation increased liquid-fuel yield by introducing oxygenated functional groups that facilitate chain scission at lower activation energies, whereas crushing plastic gave a lower liquid yield than ozonation by accelerating volatilization into non-condensable gases. Furthermore, FTIR analysis revealed that liquid fuel from treated plastics shared chemical similarities with diesel, gasoline, and kerosene, particularly in the C-H and aromatic C=C bands, although PP-derived fuels exhibited unique olefinic signatures. A Box-Behnken design and ANOVA were used to optimize the process conditions for the second objective. A maximum yield of 38.8% was achieved under the following conditions: 400 °C, a heating rate of 10 °C/min, a holding time of 2 h, and an ozonation time of 0.5 h. The PP-derived liquid fuel contained hydrocarbons similar to commercial fuels, indicating its potential as an alternative fuel source. This study concludes that chemical pretreatment significantly enhanced fuel yield and quality, making PP-derived liquid fuel a potential substitute for conventional petroleum fuels and thereby advancing the technological readiness of plastic-to-fuel biorefineries.

ABSTRAK: Kajian ini mengkaji penukaran polipropilena (PP), polietilena tereftalat (PET), dan polietilena berketumpatan tinggi (HDPE) kepada bahan api cecair melalui pirolisis, dengan penekanan terhadap kesan mekanistik prarawatan mekanikal dan kimia. Memandangkan penjanaaan sisa plastik global telah mencapai tahap kritikal, teknologi kitar tinggi nilai sisa secara mampan amat diperlukan bagi menyokong ekonomi kitaran dan mengurangkan pencemaran alam sekitar yang semakin meningkat. Oleh itu, kajian ini bertujuan mengoptimum parameter pirolisis dan memaksimum hasil bahan api cecair dengan membanding bahan suapan plastik tanpa rawatan dengan sampel rawatan menggunakan ozon, dihancurkan, serta gabungan ozonasi dan penghancuran. Kaedah prarawatan melibatkan ozonasi dan penghancuran, diikuti oleh pirolisis pada keadaan suhu berbeza (300–400 °C), kadar pemanasan (5–15 °C/min), masa pegangan (1–3 jam), dan tempoh ozonasi (0.5–1.5 jam). Bagi objektif pertama, hasil tertinggi sebanyak 46.6% diperoleh daripada PP rawatan ozon pada suhu 400 °C, kadar pemanasan 5 °C/min, masa pegangan 1 jam, dan tempoh

ozonasi 1 jam. Secara mekanistik, ozonasi meningkatkan hasil bahan api cecair dengan memperkenal kumpulan oksigen yang memudahkan pemutusan rantai pada tenaga pengaktifan yang lebih rendah, manakala penghancuran plastik menghasilkan perolehan cecair lebih rendah berbanding ozonasi akibat pemeruapan yang dipercepatkan kepada gas tidak terkondensasi. Selain itu, analisis inframerah transformasi Fourier (FTIR) menunjukkan bahawa bahan api cecair daripada plastik terawat mempunyai persamaan kimia dengan diesel, petrol, dan kerosin, khususnya pada jalur C–H dan C=C aromatik, walaupun bahan api terbitan PP menunjukkan ciri olefin tersendiri. Reka bentuk Box-Behnken dan analisis varians (ANOVA) digunakan untuk mengoptimumkan keadaan proses bagi objektif kedua. Hasil maksimum sebanyak 38.8% dicapai pada suhu 400 °C, kadar pemanasan 10 °C/min, masa pegangan 2 jam, dan tempoh ozonasi 0.5 jam. Bahan api cecair terbitan PP mengandungi hidrokarbon serupa bahan api komersial, sekaligus menunjukkan potensinya sebagai sumber bahan api alternatif. Kajian ini menyimpulkan bahawa prarawatan kimia meningkatkan hasil dan kualiti bahan api dengan ketara, menjadikan bahan api cecair terbitan PP berpotensi sebagai pengganti kepada bahan api petroleum konvensional serta meningkatkan kesiapsiagaan teknologi biopenapisan plastik kepada bahan api.

KEYWORDS: *Plastic Wastes, Pyrolysis, Liquid Fuel, Pretreatment Methods, Ozonation*

1. INTRODUCTION

Plastic products significantly enhance human welfare and support innovations in healthcare, automotive technology, electronics, construction, and medical packaging. However, global plastic production, which amounts to approximately 413 million tonnes annually, has caused significant environmental issues, including land and marine pollution and risks to human health [1, 2]. Recent global projections indicate that, in 2025, approximately 130 million tonnes (Mt) of plastic is estimated to pollute the environment each year, affecting land, air, and water [2]. Without ambitious global intervention, this figure is projected to rise to 280 Mt by 2040, which is equivalent to dumping nearly one garbage truck worth of plastic waste every second. Furthermore, approximately 31.9% of this global volume is mismanaged, mainly due to the unchecked proliferation of single-use packaging [1, 2]. Traditional disposal methods, such as landfilling and incineration, exacerbate these problems by causing ecological degradation and emitting harmful greenhouse gases, necessitating sustainable, environmentally friendly solutions [3, 4].

In Malaysia, the plastic pollution crisis is particularly acute because of infrastructure limitations and high consumption rates. The nation's per capita plastic consumption is 16.78 kg, placing it among the highest consumers in the region. Annual post-consumer plastic waste generation increased from 1.07 million tonnes in 2016 to 1.4 million tonnes in 2019, with only 21% recycled and the remainder disposed of in landfills or incinerated. Malaysia currently operates 135 landfills, most of which are nearing full capacity and are potential sources of methane emissions [3]. If current trends continue, the world could be burdened with 12 billion metric tonnes of plastic waste by 2050 [5]. The slow decomposition of plastic over centuries poses a severe long-term environmental threat [5, 6]. Recognizing this unsustainable trajectory, the Malaysian government has mandated a paradigm shift through the Circular Economy Blueprint for Solid Waste (2025–2035) and the impending enforcement of Extended Producer Responsibility (EPR) regulations slated for 2026. These legislative frameworks demand robust technological solutions to reclaim the estimated 81% of material value that is currently lost during conventional plastic disposal [3, 7, 8, 9].

Plastic recycling faces challenges such as feedstock segregation, processing cost, contamination, and labor-intensive separation [4, 10]. Therefore, converting plastic waste into

fuel for the petrochemical industry, transportation, and heating applications has been explored as an alternative waste valorization route [6, 10, 11, 12]. Pyrolysis is an emerging thermochemical technology that converts low-value municipal plastic waste into higher-value char, tar, gas, and liquid fuel [6, 10, 12]. This process decomposes organic substances in the absence of oxygen, producing carbon-rich solids, liquid oil, and gases [6, 11]. Previous studies have documented the pyrolysis of various feedstocks, including LDPE, HDPE, PP, and PS, yielding approximately 50% carbon solid from the original organic material [13]. For example, LDPE can produce 10–15% gas, 45–50% liquid oil, and 35–40% solid residue, depending on operating conditions [6, 14]. Liquid fuels derived from plastic waste can be used in engines, power generators, and heating systems. Moreover, studies indicate that plastic waste pyrolysis can generate crude-oil-like products with properties comparable to standard crude oil [10, 11].

Although pyrolysis offers several benefits, it also presents energy-consumption challenges because conventional pyrolysis requires high temperatures (300–1000 °C), which can increase operating costs [6, 10, 12]. This process is also constrained by the low oil yield and poor oil quality, and it often requires extensive preprocessing of plastic waste [10, 15]. Furthermore, the liquid-fuel quality depends on the macromolecular structure of the feed plastics and the selected pyrolysis parameters [12, 16, 17]. Catalysts and pretreatment processes have shown promise for improving liquid-fuel yield and quality, but further work is needed to refine these methods [10, 15, 18]. Historically, pretreatment has been limited to methods such as mechanical sorting, thermal drying, and basic solvent extraction, which do little to fundamentally alter the thermodynamic energy barriers required for subsequent depolymerization [19].

Chemical pretreatment methods have also been explored to modify the molecular structure of plastics prior to pyrolysis. Among these methods, ozonation has gained attention due to its ability to induce oxidative degradation under relatively mild conditions. Ozonation breaks plastic molecules into smaller fragments through oxidative chain scission. During ozonation, reactive oxygen species attack the polymer backbone, targeting tertiary carbons and unsaturated bonds and forming unstable ozonides and hydroperoxides. These intermediate structures may have lower activation energies for thermal cracking than virgin polymer chains, theoretically allowing for higher liquid yields at lower pyrolysis temperatures [20]. Despite these promising attributes, the quantitative effects of ozonation pretreatment on pyrolysis performance across different plastic types have not yet been comprehensively investigated. A detailed comparison of the responses of various plastics such as polypropylene (PP), polyethylene terephthalate (PET), and high-density polyethylene (HDPE) to ozonation pretreatment is therefore needed to better understand its influence on fuel yield, product quality, and the formation of undesirable byproducts [10, 15]. By modifying the molecular structure of plastic waste prior to thermal conversion, ozonation pretreatment may further enhance the environmental and operational benefits of plastic waste pyrolysis.

To maximize fuel yield and quality, the pyrolysis conditions, including temperature, heating rate, and holding time, must be optimized [12, 21]. However, previous optimization studies on liquid-fuel yield have not fully considered the combined effects of pretreatment processes. Therefore, the full potential of combining optimum pyrolysis conditions with suitable pretreatment has yet to be exploited. Although statistical optimization techniques, such as Central Composite Design (CCD), are widely used in the literature, they frequently require testing at extreme axial points [21]. In thermochemical pyrolysis, subjecting polymers simultaneously to maximum temperatures, maximum holding times, and maximum heating rates can induce hazardous thermal runaway or secondary cracking that inadvertently favors non-condensable gas production over liquid fuels [10, 12]. Consequently, a Box-Behnken

design (BBD) offers a suitable methodological alternative. BBD constructs a response surface using only the midpoints of the experimental cube and avoids extreme operational boundaries [21]. This ensures that the optimization process remains within a safe and energetically viable operating zone while requiring fewer experimental runs.

The aim of this research is to identify the optimal combinations of temperature, heating rate, holding time, and ozonation conditions for producing high-quality liquid fuel. Optimizing these parameters can make the pyrolysis process more economical, efficient, and sustainable. This study aims to increase liquid-fuel production from plastic waste while supporting environmental protection and resource conservation. This research can reduce costs, support sustainability, and make plastic-waste recycling more environmentally friendly through process optimization.

2. MATERIALS AND METHODS

This section defines the experimental techniques used in this study, including the project workflow, experimental design, materials and apparatus, and detailed analysis procedures.

2.1. Experimental Materials

In this study, polypropylene (PP), polyethylene terephthalate (PET), and high-density polyethylene (HDPE) plastic wastes were collected from households and areas surrounding the IIUM campus. Dichloromethane (DCM) (Sigma-Aldrich, >99% purity) was used for oil extraction. During pyrolysis, DCM was used to separate the liquid fuel (oil) from the carbon material produced. Scissors and knives were used to cut the plastic waste. Standard analytical laboratory equipment, including crucible boats, spatulas, sealer bags, centrifuge tubes, beakers, filter paper, conical flasks, weighing boats, micropipettes, and measuring cylinders, was used for sample preparation and measurement. The essential equipment included a laboratory-scale pyrolysis furnace, an analytical weighing scale, a Nano 15 ozone generator, a mechanical crusher (HITOP SY-20), and an FTIR instrument (Bruker INVENIO S). Safety equipment, including gloves, lab coats, and a ventilation system, was also used to ensure a safe working environment.

2.2. Experimental Procedures

2.2.1. Preparation of Plastic Waste and Pretreatment

Plastic waste was sorted, cleaned, and shredded into 1.25 cm by 1.25 cm pieces. Ozonation was performed on the sorted plastic waste at an ozone flow rate of 1 L/min for 1 h using the Nano 15 ozone generator to modify the plastic's molecular structure through oxidative degradation. For the optimization runs, ozonation time was varied as described in Section 2.2.3. Alternatively, the plastic waste was fed into the HITOP SY-20 crusher machine and crushed into smaller particles using a 3 mm × 6 mm mesh to increase the surface area available for treatment.

2.2.2. Plastic Waste Screening and Pretreatment Methods

PP, PET, and HDPE plastic wastes were each weighed at 5 g. In a nitrogen atmosphere (maintained via a continuous purge to ensure strictly non-oxidative conditions and prevent combustion), the pyrolysis process was carried out at 400 °C with a heating rate of 5 °C/min and a holding time of 1 h. To ensure experimental consistency and statistical reliability, all pyrolysis screening runs were performed in triplicate. Under identical conditions, the liquid-fuel yields were reported as the mean of the three replicates. All masses were measured with a

calibrated analytical balance (± 0.001 g), and the same furnace temperature program, nitrogen flow, and extraction protocol were applied throughout to maintain experimental consistency. Solid products, mainly carbonaceous material, were collected, weighed, and placed in centrifuge tubes. Liquid fuel retained in the carbon material was extracted using DCM. The liquid fuel was separated from the carbon material by sonication and filtration using filter paper. The carbon material was dried overnight at 60 °C in an oven before weighing. The extracted liquid fuel was also weighed, and the experiment was repeated for each plastic waste type under different pretreatment conditions.

2.2.3. Optimization of Pyrolysis for Liquid Fuel Production

A Box-Behnken design, a response-surface methodology, was used to optimize the pyrolysis of ozone-treated PP plastic waste. The effects of temperature (300–400 °C), heating rate (5–15 °C/min), holding time (1–3 h), and ozonation time (0.5–1.5 h) on the liquid-fuel yield during macromolecular breakdown were investigated using four independent variables. As previously justified, the BBD approach was selected to construct a response-surface model without extrapolating into extreme and potentially hazardous axial permutations. The process variables were optimized, and liquid-fuel yield was maximized using 24 experimental runs, including the required center-point replications to estimate the system's pure error.

2.3. Experimental Analysis

FTIR analysis was performed to identify the functional groups present in the liquid fuel and provide insight into its molecular structure. Liquid fuel samples from PP plastic waste with different pretreatments (non-treated, ozone-treated, crushed, and crushed and ozone-treated) were analyzed. The FTIR instrument emitted infrared light across a standard mid-infrared range (4000 to 400 cm^{-1} at a resolution of 4 cm^{-1} , with 32 scans accumulated per sample), and the liquid fuel absorbed specific frequencies associated with functional groups. The resulting absorption spectra were recorded and analyzed to identify characteristic functional-group peaks. The spectra were subsequently interpreted against published characterization of plastic-derived and conventional fuel components to assess the need for further upgrading.

3. RESULTS AND DISCUSSION

3.1. Screening of Plastic Waste and Pretreatment Methods

The screening process involved pyrolyzing three types of plastic waste, PP, PET, and HDPE, under four pretreatment conditions: ozone-treated, crushed, ozone-treated-and-crushed, and untreated. Table 1 shows that liquid-fuel yield was highest for ozone-treated PP (46.6%), followed by PET (35.0%) and HDPE (8.2%). Promising results were also seen in the combined ozone-treated-and-crushed conditions, in which PP yielded 35.4%, PET yielded 32.4%, and HDPE yielded 5.4%. All plastic types showed lower liquid-fuel yields after crushing alone than after ozonation, with PP at 31.0%, PET at 24.0%, and HDPE at 3.2%. Non-treated plastics gave the lowest liquid-fuel yields, with 21.6% for PP, 17.2% for PET, and 2.2% for HDPE.

The results show that ozone treatment can efficiently enhance liquid-fuel production from plastic waste, especially PP, which exhibited the highest yield under all conditions. The ozone-treated PP yield of 46.6% is notable compared with previous studies on liquid-fuel conversion in the slow-pyrolysis process, which reported a maximum conversion rate of 50.1% even under highly optimized catalytic environments [18, 22, 23]. Although the yield was below the maximum values reported in some studies, the ozone pretreatment effect was large compared with untreated and other pretreated samples, indicating the potential of ozone pretreatment to improve the efficiency of liquid-fuel production from plastic waste.

Compared with alternative pretreatment strategies reported in the literature, such as dry washing, solvent extraction, or thermal drying, ozonation can alter the polymer's fundamental reaction kinetics rather than merely removing surface contaminants [10, 15, 20]. This distinction is evident from the present screening results, in which ozone pretreatment increased the PP liquid-fuel yield from 21.6% to 46.6%, representing a 25.0 percentage-point increase and a 2.16-fold improvement relative to untreated PP. By comparison, crushing alone increased the PP liquid-fuel yield to 31.0%, indicating that increased surface area and improved heat transfer contributed to conversion but did not achieve the same enhancement as chemical oxidation. Physical preprocessing methods, including washing, drying, crushing, and sieving, mainly improve feed cleanliness, moisture control, particle homogeneity, and thermal accessibility; however, they do not intentionally generate chemically weak sites along the polymer backbone [10, 12, 15].

The limited effect of cleaning-type pretreatment is also supported by previous PP pyrolysis work, in which washed and unwashed waste PP differed by only 1.85 percentage points in liquid yield at 300 °C (81.47 and 79.62 wt%, respectively) and 1.92 percentage points at 400 °C (80.55 and 78.70 wt%, respectively) [24]. In contrast, ozonation acts as an advanced oxidative preprocessing route. Ozone-derived reactive oxygen species preferentially attack tertiary carbon sites in PP, generating peroxy and hydroperoxide intermediates that undergo β -chain scission to form lower-molecular-weight fragments with olefinic and ketone end groups [20]. These oxidized defects facilitate thermal cracking initiation and are therefore consistent with lower effective thermal resistance to depolymerization than in untreated PP. Although activation energies were not directly measured in the present study, the substantial yield increase and the subsequent appearance of oxygenated FTIR bands support facilitated early-stage chain scission rather than a simple surface-area effect.

Solvent-based approaches should be interpreted differently because they are primarily separation or product-recovery methods rather than true depolymerization pretreatments. For example, sequential solvent extraction has been reported to recover 63–81% of pyrolytic oil trapped in crude char and to remove 64–86% of inorganic contaminants; however, this treatment does not chemically activate the polymer backbone before pyrolysis [15]. Selective dissolution can also improve polymer separation, but solvent cost, toxicity, and recovery requirements constrain its implementation [15]. Hydrogen-peroxide-based treatment may generate alkoxy, peroxy, and hydroxyl radical species that can promote PP chain fragmentation; nevertheless, it is a liquid-phase oxidation route that requires chemical storage, post-treatment separation, and drying [20]. Therefore, the present results indicate that ozonation is particularly advantageous as a gas-phase chemical preprocessing method because it introduces chain-scission-prone oxygenated sites without using an organic solvent, although safe ozone generation and off-gas destruction remain necessary.

The visual differences in Fig. 1 further support the influence of polymer structure and pretreatment on product distribution. The recovered product mixtures are consistent with the quantitative trend in Table 1, with PP producing the greatest condensable liquid fraction, followed by PET and HDPE. Fig. 1 therefore provides qualitative evidence of feedstock-dependent product distribution. To quantify these differences, the liquid-fuel yield was calculated as the mass of recovered liquid fuel relative to the initial mass of plastic feedstock, as expressed in Eq. (1). This calculation enables direct comparison of the liquid products obtained from PP, PET, and HDPE under the different pretreatment conditions. The following discussion relates these differences to the molecular structures and thermal-degradation behavior of the three polymers.

$$\text{Percentage of liquid fuel yield (\%)} = \frac{\text{Mass of liquid fuel (g)}}{\text{Initial mass of plastic (g)}} \times 100\% \quad (1)$$

Table 1. Experimental results for different plastic types and pretreatment conditions. Values are means of three replicate runs

Type of plastic	Weight before pyrolysis (g)	Weight of carbon after pyrolysis (g)	Weight of carbon after oil extraction and drying (g)	Total mass of liquid fuel (g)	Liquid-fuel conversion yield (%)
No pretreatment					
PP	5.00	4.11	3.92	1.08	21.6
PET	5.00	4.23	4.14	0.86	17.2
HDPE	5.00	4.90	4.89	0.11	2.2
Ozone only					
PP	5.00	3.17	2.67	2.33	46.6
PET	5.00	3.53	3.25	1.75	35.0
HDPE	5.00	4.66	4.59	0.41	8.2
Crush only					
PP	5.00	3.75	3.45	1.55	31.0
PET	5.00	3.93	3.80	1.20	24.0
HDPE	5.00	4.87	4.84	0.16	3.2
Ozone and crush					
PP	5.00	3.62	3.23	1.77	35.4
PET	5.00	3.64	3.38	1.62	32.4
HDPE	5.00	4.78	4.73	0.27	5.4

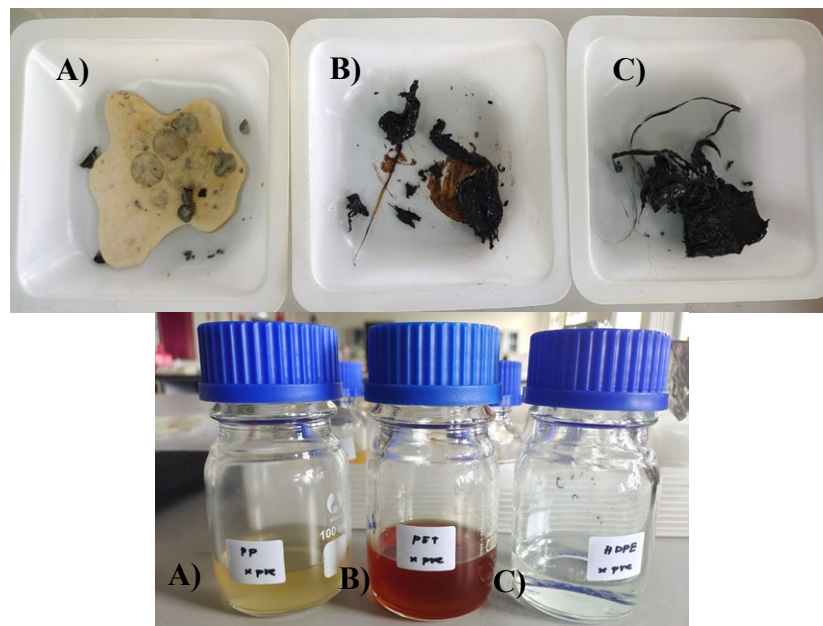


Figure 1. Mixtures of products from pyrolysis of several types of plastic waste. A: Polypropylene (PP). B: Polyethylene terephthalate (PET). C: High-density polyethylene (HDPE).

The higher liquid-fuel content from polypropylene (PP) is associated with its composition and molecular structure, which favor the formation of liquid-fuel products during pyrolysis [16, 17, 22]. The favorable decomposition characteristics of PP are related to its lower crystallinity and more amorphous structure, resulting in the thermal degradation of the shorter hydrocarbon chains. The relatively weak tertiary carbon-hydrogen bonds, which break down

more easily under pyrolytic conditions, explain the high liquid-fuel yield observed in this conversion. In essence, the tertiary carbons act as active sites where thermal cleavage initiates with a considerably lower activation energy than the primary or secondary carbons found in linear polyolefins [16, 22].

The high crystallinity and thermal stability of PET, in which ethylene terephthalate ($C_{10}H_8O_4$) repeating units are linked, mean that PET breakdown during pyrolysis produces more solid residue than liquid hydrocarbon products [12, 17]. The aromatic structure can promote gaseous by-product and solid-residue formation during pyrolysis, leading to a lower liquid-fuel yield. In contrast, HDPE consists of long chains of ethylene units $(C_2H_4)_n$ with a linear and highly crystalline structure due to the absence of side branches. Therefore, HDPE has higher thermal stability and greater potential for carbon-rich residue formation during pyrolysis, and it generally produces more solid residue than liquid hydrocarbons [11, 12, 17].

Crushing plastic before pyrolysis produces particles with a larger surface area, which facilitates heat transfer and more uniform thermal degradation. Nevertheless, this process can also decrease the liquid-fuel yield because of inherent factors in the pyrolysis process [10, 15]. These factors include the increased surface area, non-uniform pyrolysis, and shorter residence times of liquid intermediates in the reactor, resulting in a higher proportion of gaseous byproducts volatilizing before they can be effectively condensed into the target liquid phase. The present PP results support this interpretation: crushing improved conversion relative to untreated PP but did not reproduce the higher liquid yield achieved by ozone treatment, showing that physical size reduction and oxidative chain activation contribute differently to the overall product distribution.

3.2. Experimental Analysis

3.2.1. Chemical Composition

The liquid-fuel samples derived from PP plastic were characterized using Fourier-transform infrared spectroscopy (FTIR), as FTIR can identify specific molecular groups within the liquid fuel [24, 25]. Fig. 2 presents the FTIR spectra of liquid fuel produced from PP plastic under various pretreatment procedures, showing clear peaks associated with chemical changes induced by the pretreatment procedures.

The spectrum of the untreated sample is dominated by peaks for aliphatic hydrocarbons with C-H stretching bands at $2800\text{--}3000\text{ cm}^{-1}$ and C-C stretching at 1370 cm^{-1} , indicating the presence of long-chain hydrocarbons characteristic of PP-derived oil [24, 25]. However, the crushed-PP sample showed intensified peaks in this region, suggesting that hydrocarbons became more available because mechanical disruption exposed a larger surface area.

The spectrum of the ozone-treated PP shows strong peaks in the carbonyl ($C=O$) region ($1700\text{--}1750\text{ cm}^{-1}$) and hydroxyl (O-H) region ($3200\text{--}3600\text{ cm}^{-1}$) due to the oxidation of polymer chains. This spectroscopic evidence directly confirms the mechanistic hypothesis that ozone introduces reactive oxygen species onto the polyolefin backbone [20]. The combination of crushing and ozone treatment further intensified carbonyl and hydroxyl peaks, indicating stronger oxidative degradation when the two pretreatment methods were combined.

The FTIR results were interpreted by comparison with published characterization studies of plastic-derived liquid fuels. The strong aliphatic C-H stretching bands at approximately $2800\text{--}3000\text{ cm}^{-1}$ and the C-H deformation bands at $1460\text{--}1370\text{ cm}^{-1}$ indicate that the PP-derived oils were predominantly composed of hydrocarbons. Similar aliphatic hydrocarbon bands have been reported for liquid fuels produced from catalytic LDPE conversion, indicating that plastic-derived oils share important functional-group characteristics with conventional

fuels [26]. Therefore, the present results suggest chemical similarity with gasoline, kerosene, and diesel at the functional-group level, particularly in the aliphatic hydrocarbon region.

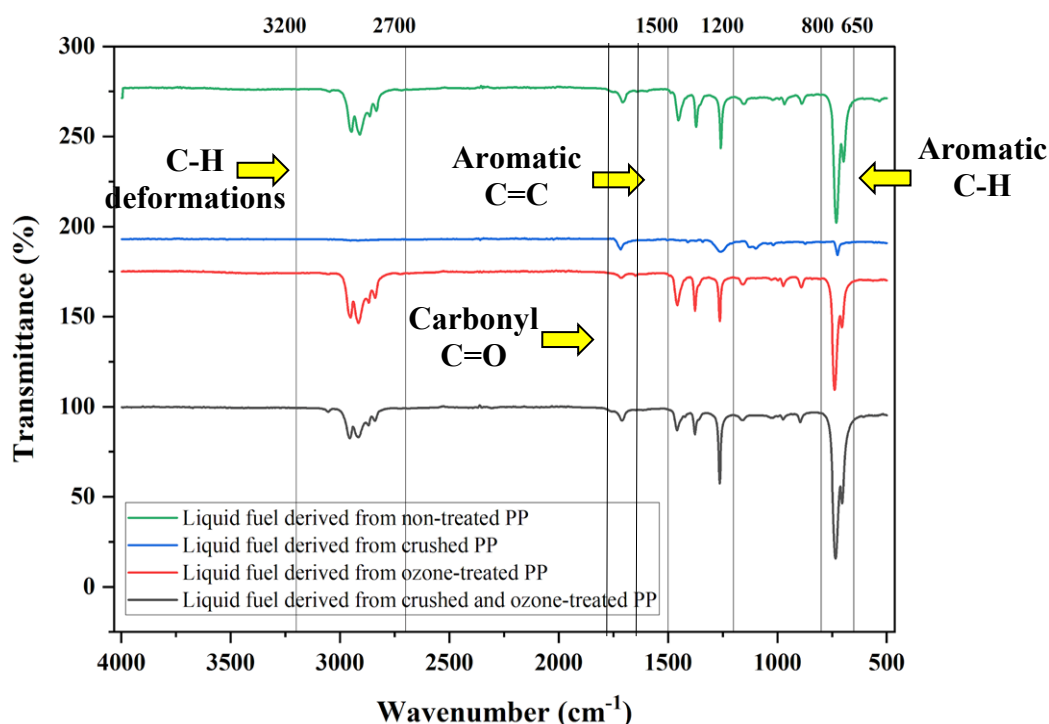


Figure 2. FTIR spectra of liquid fuel derived from PP plastic under four pretreatment conditions: non-treated, crushed, ozone-treated, and ozone-treated with crushing.

Nevertheless, similarity in the C-H region does not mean that PP-derived oil is chemically identical to commercial fuel. PP pyrolysis oil commonly contains paraffinic and olefinic hydrocarbons, while the amounts of unsaturated, aromatic, and oxygen-containing compounds depend on the feedstock composition and pyrolysis conditions [24, 25]. Consequently, differences are expected in the fingerprint region below 1500 cm^{-1} , where branched hydrocarbons, olefins, aromatics, and oxygenated compounds contribute differently to the spectral pattern. In the present study, the stronger carbonyl band at approximately $1700\text{--}1750\text{ cm}^{-1}$ and the broader hydroxyl band at approximately $3200\text{--}3600\text{ cm}^{-1}$ in the ozone-treated samples indicate that oxidation occurred before pyrolysis. This observation is consistent with the ozone-oxidation pathway of PP, in which tertiary-carbon sites form peroxy and hydroperoxide intermediates that promote chain shortening and oxygenated products [20].

In contrast, crushing mainly improves the physical accessibility of the plastic particles and is not expected to generate strong carbonyl or hydroxyl bands. Thus, the more intense C=O and O-H absorption bands in the ozone-treated oils provide spectroscopic evidence that ozonation modified the PP structure before pyrolysis. Although the oil contains hydrocarbon fractions within the ranges commonly associated with gasoline ($\text{C}_4\text{--C}_{12}$), kerosene ($\text{C}_9\text{--C}_{16}$) and diesel ($\text{C}_{10}\text{--C}_{22}$) [16, 24], further analysis using GC-MS, simulated distillation and standard fuel-property tests is required before confirming its suitability as a conventional fuel substitute.

3.3. Optimization of Pyrolysis for Liquid Fuel Formation

The Box-Behnken design was used to evaluate the effects of four process factors, namely temperature, heating rate, holding time, and ozonation time, on liquid-fuel yield [21]. The experimental design comprised 24 runs, as presented in Table 2. Temperature was investigated

over the range of 300–400 °C, with a center point of 350 °C. The heating rate ranged from 5 to 15 °C/min, with a center point of 10 °C/min. Holding time and ozonation time were evaluated over the ranges of 1 to 3 h and 0.5 to 1.5 h, respectively, with center points of 2 h and 1 h. These operating ranges were selected to examine the combined effects of thermal severity and oxidative pretreatment on the formation of condensable liquid products.

According to Table 2, the highest liquid-fuel yield (38.8%) was obtained in Run 19 at 400 °C, a heating rate of 10 °C/min, a holding time of 2 h, and an ozonation time of 0.5 h. In contrast, Run 21 produced the lowest yield (7.4%) at 350 °C, a heating rate of 5 °C/min, a holding time of 2 h, and an ozonation time of 1.5 h. This variation confirms that liquid-fuel production was sensitive to the combined operating conditions. In particular, increasing temperature promotes polymer-chain cleavage and the formation of volatile compounds, which can subsequently condense into liquid products [6, 10, 12, 14]. However, in more severe thermal conditions, secondary cracking may convert condensable vapors into non-condensable gases, thereby reducing liquid-product recovery [6, 10, 12].

Table 2. Experimental design matrix and liquid-fuel yield responses obtained using the Box-Behnken design

Run	Factor 1: Temperature (°C) <i>A</i>	Factor 2: Heating rate (°C/min) <i>B</i>	Factor 3: Holding time (h) <i>C</i>	Factor 4: Ozonation time (h) <i>D</i>	Total mass of oil (g)	Response 1: Yield of oil (%) <i>R1</i>
1.	350	5	1	1	0.51	10.2
2.	300	5	2	1	0.75	15.0
3.	350	15	2	0.5	1.07	21.4
4.	300	10	2	0.5	1.04	20.8
5.	350	15	2	1.5	0.92	18.4
6.	400	10	3	1	1.85	37.0
7.	350	5	2	0.5	1.02	20.4
8.	350	10	3	1.5	1.23	24.6
9.	400	15	2	1	1.87	37.4
10.	400	10	1	1	1.27	25.4
11.	400	5	2	1	1.79	35.8
12.	300	10	3	1	1.03	20.6
13.	400	10	2	1.5	1.65	33.0
14.	300	15	2	1	0.83	16.6
15.	300	10	2	1.5	1.01	20.2
16.	350	5	3	1	1.10	22.0
17.	300	10	1	1	0.82	16.4
18.	350	10	1	0.5	0.67	13.4
19.	400	10	2	0.5	1.94	38.8
20.	350	10	1	1.5	0.59	11.8
21.	350	5	2	1.5	0.37	7.4
22.	350	15	3	1	1.22	24.4
23.	350	10	3	0.5	1.11	22.2
24.	350	15	1	1	0.56	11.2

The screening and optimization results should be interpreted in light of their different experimental conditions. During the screening stage, ozone-treated PP achieved the highest liquid-fuel yield of 46.6% at 400 °C, a heating rate of 5 °C/min, a holding time of 1 h, and an ozonation time of 1 h. In contrast, the Box-Behnken optimization identified a maximum yield of 38.8% at a moderate heating rate of 10 °C/min and a shorter ozonation time of 0.5 h. These findings indicate that liquid-fuel yield depends on the interaction between thermal conditions

and pretreatment duration rather than on any single factor. Previous studies have similarly shown that temperature, heating rate, and residence time collectively influence polymer degradation, vapor-phase cracking, and the distribution of liquid, gas, and solid products during plastic pyrolysis [6, 10, 12, 14, 21]. Therefore, the selected operating conditions should balance sufficient thermal cracking for liquid formation against excessive thermal severity that may promote secondary cracking and gas production.

Moreover, liquid-fuel yield was evaluated as a function of the four process factors: temperature (A), heating rate (B), holding time (C), and ozonation time (D), using Design-Expert 13 software. A model was developed correlating the independent and dependent variables based on the observed responses. The model was statistically significant in affecting the response, as shown by analysis of variance (ANOVA), with an F-value of 8.52 and a p-value of 0.0004. This result implies that there is only a 0.04% probability that the observed response was attributed to noise. Meanwhile, p-values of 0.3606 and 0.2923 showed that the heating-rate (B) and ozonation-time (D) parameters were not significant, whereas temperature (A) and holding time (C) were significant in determining the response variable. This indicates that temperature and holding time are very important in determining the response variable. In contrast, changes in heating rate and ozonation time were less influential under these experimental conditions. From a mass and heat transfer perspective, the insignificance of the heating rate across the 5 to 15 °C/min range suggests that internal thermal gradients within the shredded polymer particles were already minimized at these modest ramp rates, allowing temperature and residence time to dictate the cracking kinetics. Furthermore, the lack of statistical significance for extending ozonation time from 0.5 to 1.5 h implies a rapid saturation of accessible oxidative surface sites on the PP matrix. Once these surface sites are fully saturated with hydroperoxides within the first 30 minutes, prolonged ozone exposure yields diminishing returns, a finding that has beneficial implications for minimizing energy consumption during industrial scale-up scenarios [20, 21].

Table 3. Analysis of variance (ANOVA) for liquid-fuel yield (R1)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	1193.34	4	298.33	8.52	0.0004	significant
A-Temperature	797.07	1	797.07	22.77	0.0001	
B-Heating rate	30.72	1	30.72	0.8775	0.3606	
C-Holding time	324.48	1	324.48	9.27	0.0067	
D-Ozonation time	41.07	1	41.07	1.17	0.2923	
Residual	665.16	19	35.01			
Cor Total	1858.50	23				

Fig. 3 shows the response-surface plots generated using Design-Expert 13 software to illustrate the combined effects of temperature, heating rate, holding time, and ozonation time on liquid-fuel yield. Across the plotted factor combinations, liquid-fuel yield increased primarily with higher temperature and longer holding time, consistent with the ANOVA results showing that these two variables were significant model terms. Higher temperatures promoted polymer-chain cleavage and the formation of volatile products, whereas sufficient holding time allowed the thermal degradation process to proceed more completely and improved the recovery of condensable hydrocarbons [6, 10, 12, 14].

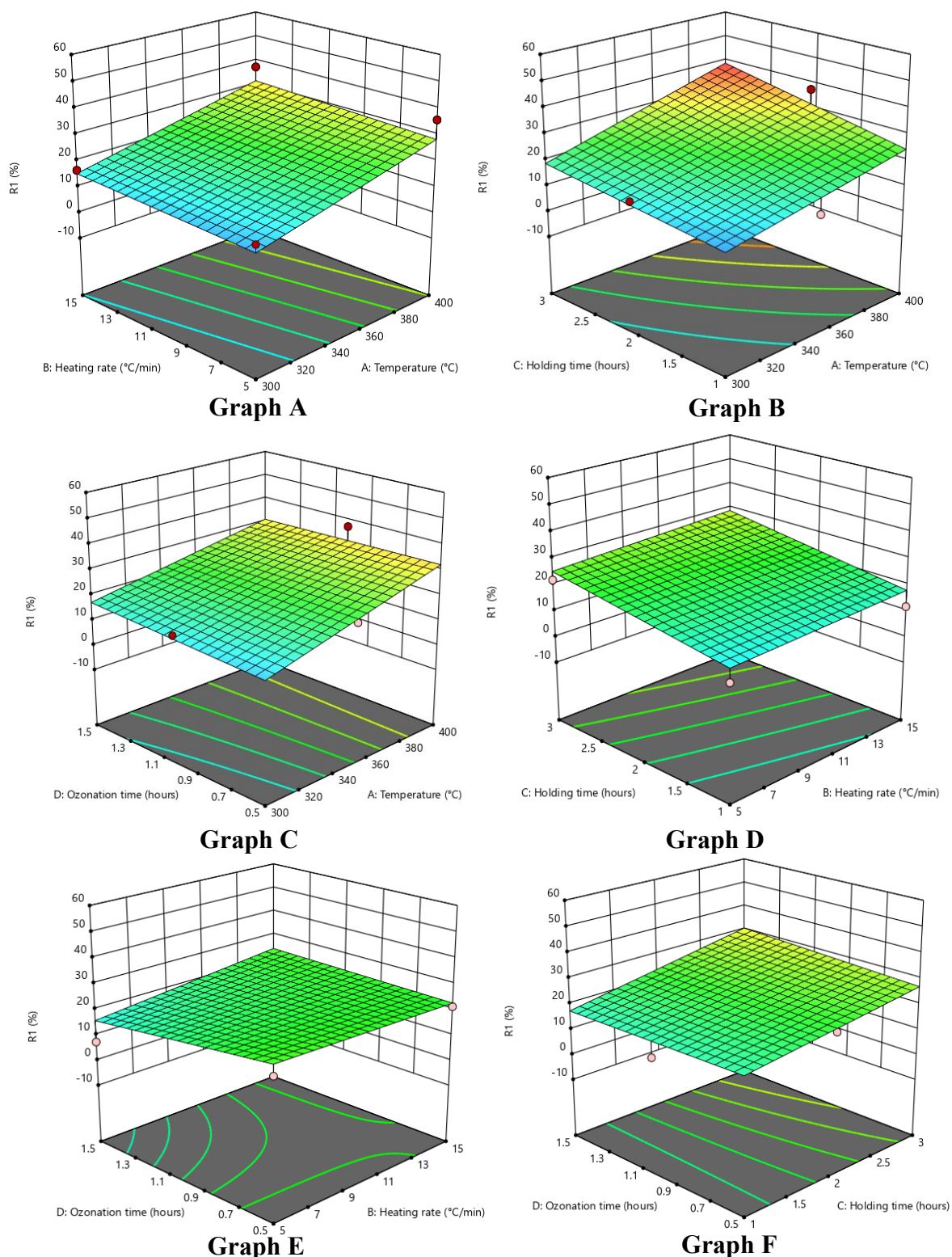


Figure 3. Three-dimensional response surfaces for liquid-fuel yield.

Graphs A–C demonstrate that temperature had the strongest positive influence on liquid-fuel yield within the investigated range. The effect of heating rates was comparatively limited, as reflected by its non-significant ANOVA result. Although moderate heating rates produced favorable yields in some combinations, the response surfaces indicate that temperature remained the dominant factor controlling liquid-fuel formation. This observation is consistent

with previous plastic-pyrolysis studies, which reported that temperature strongly affects the extent of polymer degradation and the distribution of liquid, gas, and solid products [6, 12, 14].

Graphs D–F further illustrate that holding time interacted with the remaining variables to influence liquid-fuel recovery. Longer holding times generally supported higher yields at elevated temperatures by allowing sufficient time for primary polymer cracking and vapor release. In contrast, ozonation time showed a comparatively weak effect within the tested range of 0.5–1.5 h, which agrees with its non-significant ANOVA result. This suggests that the accessible oxidative surface sites on the PP surface were substantially activated during the initial ozonation period, whereas prolonged ozone exposure provided limited additional benefit under the present experimental conditions [20, 21]. Overall, the response surfaces confirm that liquid-fuel yield optimization was governed mainly by temperature and holding time, while moderate heating rates and shorter ozonation durations were adequate for achieving favorable liquid-fuel production.

4. CONCLUSION

This study investigated the effects of pretreatment methods and pyrolysis parameters on liquid-fuel yield from polypropylene (PP), polyethylene terephthalate (PET), and high-density polyethylene (HDPE). Among the tested feedstocks and pretreatment conditions, ozone-treated PP produced the highest liquid-fuel yield (46.6%) at 400 °C, a heating rate of 5 °C/min, a holding time of 1 h, and an ozonation time of 1 h. This result demonstrates the engineering value of catalyst-free ozonation as a chemical pretreatment for improving PP conversion. Ozonation increased the PP liquid-fuel yield from 21.6% for untreated PP to 46.6%, whereas crushing alone produced a lower yield of 31.0%. This improvement is attributed to oxidative modification of the polymer backbone, whereby reactive oxygen intermediates promote chain scission and facilitate subsequent thermal degradation. Additionally, the optimization study identified a maximum liquid-fuel yield of 38.8% at 400 °C, a heating rate of 10 °C/min, a holding time of 2 h, and an ozonation time of 0.5 h. Analysis of variance (ANOVA) confirmed that extending ozonation beyond 0.5 h did not significantly improve the predicted liquid-fuel yield within the investigated operating range. This finding indicates that a shorter ozonation duration may reduce pretreatment time and energy demand while maintaining favorable liquid-fuel production under optimized pyrolysis conditions. Moreover, FTIR spectral analysis showed that the PP-derived oils contained hydrocarbon functional groups comparable with those reported for conventional fuels. However, the presence of oxygenated groups in the ozone-treated products indicates that further upgrading is required before use as a direct fuel substitute.

From an engineering perspective, integrating ozone pretreatment with process optimization provides a practical approach to improving liquid-fuel recovery from PP waste while avoiding catalyst use. Nevertheless, the findings are limited to laboratory-scale experiments using 5 g feed samples. In addition, the present response-surface analysis was based on a linear model, and FTIR provided only qualitative information on the chemical composition of the derived oils. Future work should therefore validate the optimized conditions at pilot scale, develop and validate an appropriate quadratic response-surface model, quantify fuel properties including calorific value, viscosity, density, flash point, sulfur content, oxygen content, and distillation range, and evaluate deoxygenation or other upgrading strategies. Complementary GC–MS analysis and techno-economic assessment should also be undertaken to determine the detailed hydrocarbon distribution, upgrading requirements, and commercial-scale feasibility of the proposed waste-to-fuel process. These outcomes would strengthen the

contribution of ozonation-assisted pyrolysis to sustainable plastic-waste management and circular-economy implementation in Malaysia.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Department of Chemical Engineering and Sustainability (CHES), Kulliyah of Engineering, International Islamic University Malaysia (IIUM), for its assistance and for facilitating this work.

REFERENCES

- [1] OECD. (2022). Global plastics outlook: Economic drivers, environmental impacts and policy options. OECD Publishing. <https://doi.org/10.1787/de747aef-en>
- [2] The Pew Charitable Trusts. (2025). Breaking the plastic wave 2025: An assessment of the global system and strategies for transformative change. <https://www.pew.org/en/research-and-analysis/reports/2025/12/breaking-the-plastic-wave-2025>
- [3] Chen, H. L., Nath, T. K., Chong, S., Foo, V., Gibbins, C., & Lechner, A. M. (2021). The plastic waste problem in Malaysia: Management, recycling and disposal of local and global plastic waste. *SN Applied Sciences*, 3, Article 437. <https://doi.org/10.1007/s42452-021-04234-y>
- [4] Kumar, S., Singh, E., Mishra, R., Kumar, A., & Caucci, S. (2021). Utilization of plastic wastes for sustainable environmental management: A review. *ChemSusChem*, 14(19), 3985–4006. <https://doi.org/10.1002/cssc.202101631>
- [5] Geyer, R., Jambeck, J. R., & Law, K. L. (2017). Production, use, and fate of all plastics ever made. *Science Advances*, 3(7), e1700782. <https://doi.org/10.1126/sciadv.1700782>
- [6] Anuar Sharuddin, S. D., Abnisa, F., Wan Daud, W. M. A., & Aroua, M. K. (2016). A review on pyrolysis of plastic wastes. *Energy Conversion and Management*, 115, 308–326. <https://doi.org/10.1016/j.enconman.2016.02.037>
- [7] Nizami, A. S., Rehan, M., Waqas, M., Naqvi, M., Ouda, O. K. M., Shahzad, K., Miandad, R., Khan, M. Z., Syamsiro, M., Ismail, I. M. I., & Pant, D. (2017). Waste biorefineries: Enabling circular economies in developing countries. *Bioresource Technology*, 241, 1101–1117. <https://doi.org/10.1016/j.biortech.2017.05.097>
- [8] World Bank Group. (2021). Market Study for Malaysia: Plastics circularity opportunities and barriers. <https://www.worldbank.org/en/country/malaysia/publication/market-study-for-malaysia-plastics-circularity-opportunities-and-barriers>
- [9] BERNAMA. (2024). KPKT launches Circular Economy Blueprint, advancing solid waste management. <https://www.bernama.com/tv/news.php?id=2326036>
- [10] Qureshi, M. S., Oasmaa, A., Pihkola, H., Deviatkin, I., Tenhunen, A., Mannila, J., Minkkinen, H., Pohjakallio, M., & Laine-Ylijoki, J. (2020). Pyrolysis of plastic waste: Opportunities and challenges. *Journal of Analytical and Applied Pyrolysis*, 152, 104804. <https://doi.org/10.1016/j.jaap.2020.104804>
- [11] Kunwar, B., Cheng, H. N., Chandrashekar, S. R., & Sharma, B. K. (2016). Plastics to fuel: A review. *Renewable and Sustainable Energy Reviews*, 54, 421–428. <https://doi.org/10.1016/j.rser.2015.10.015>
- [12] Singh, R. K., Balakrishnan, M., & Negi, R. K. (2022). A review on plastic pyrolysis: The current status, challenges, and future perspectives. *Journal of Cleaner Production*, 341, 130862. <https://doi.org/10.1016/j.jclepro.2022.130862>
- [13] Dewangga, P. B., Rochmadi, & Purnomo, C. W. (2019). Pyrolysis of polystyrene plastic waste using bentonite catalyst. *IOP Conference Series: Earth and Environmental Science*, 399(1), 012110. <https://doi.org/10.1088/1755-1315/399/1/012110>
- [14] Cahyono, M. S., & Fenti, U. I. (2017). Influence of heating rate and temperature on the yield and properties of pyrolysis oil obtained from waste plastic bag. *Conserve: Journal of Energy and Environmental Studies*, 1(1), 1–8. <https://doi.org/10.30588/cjees.v1i1.248>

- [15] Belbessai, S., Azara, A., & Abatzoglou, N. (2022). Recent advances in the decontamination and upgrading of waste plastic pyrolysis products: An overview. *Processes*, 10(4), 733. <https://doi.org/10.3390/pr10040733>
- [16] Das, N. C. (2025). Catalytic and noncatalytic pyrolysis of waste plastics: Comparative analysis of liquid products from polyethylene, polypropylene, and mixed plastics with diesel and petrol. *Industrial & Engineering Chemistry Research*, 64(6), 3240–3253. <https://doi.org/10.1021/acs.iecr.4c04155>
- [17] Miandad, R., Barakat, M. A., Aburizaiza, A. S., Rehan, M., Ismail, I. M. I., & Nizami, A. S. (2017). Effect of plastic waste types on pyrolysis liquid oil. *International Biodeterioration & Biodegradation*, 119, 239–252. <https://doi.org/10.1016/j.ibiod.2016.09.017>
- [18] Miandad, R., Rehan, M., Barakat, M. A., Aburizaiza, A. S., Khan, H., Ismail, I. M. I., & Nizami, A. S. (2019). Catalytic pyrolysis of plastic waste: Moving toward pyrolysis based biorefineries. *Frontiers in Energy Research*, 7, 27. <https://doi.org/10.3389/fenrg.2019.00027>
- [19] Lindberg, J., Ljungkvist Nordin, H., & Fråne, A. (2019). Sustainability assessment of pre-treatment methods for plastic waste from hospitals in Sweden (Report No. C 383). IVL Swedish Environmental Research Institute.
- [20] García-López, E. I., Aoun, N., & Marci, G. (2024). An overview of the sustainable depolymerization/degradation of polypropylene microplastics by advanced oxidation technologies. *Molecules*, 29(12), 2816. <https://doi.org/10.3390/molecules29122816>
- [21] Faisal, F., Rasul, M. G., Chowdhury, A. A., & Jahirul, M. I. (2024). Optimisation of process parameters to maximise the oil yield from pyrolysis of mixed waste plastics. *Sustainability*, 16(7), 2619. <https://doi.org/10.3390/su16072619>
- [22] Ahmad, I., Khan, M. I., Khan, H., Ishaq, M., Tariq, R., Gul, K., & Ahmad, W. (2015). Pyrolysis study of polypropylene and polyethylene into premium oil products. *International Journal of Green Energy*, 12(7), 663–671. <https://doi.org/10.1080/15435075.2014.880146>
- [23] Muhammad, C., Onwudili, J. A., & Williams, P. T. (2015). Thermal degradation of real-world waste plastics and simulated mixed plastics in a two-stage pyrolysis–catalysis reactor for fuel production. *Energy & Fuels*, 29(4), 2601–2609. <https://doi.org/10.1021/ef502749h>
- [24] Uebe, J., Kryzevicius, Z., Majauskiene, R., Dulevicius, M., Kosychova, L., & Zukauskaitė, A. (2022). Use of polypropylene pyrolysis oil in alternative fuel production. *Waste Management & Research*, 40(8), 1220–1230. <https://doi.org/10.1177/0734242X211068243>
- [25] Ureel, Y., Chacón-Patiño, M. L., Kusenberg, M., Rodgers, R. P., Sabbe, M. K., & Van Geem, K. M. (2024). Characterization of PP and PE waste pyrolysis oils by ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry. *Energy & Fuels*, 38(12), 11148–11160. <https://doi.org/10.1021/acs.energyfuels.4c01954>
- [26] Khan, M. S., Inamullah, Sohail, M., & Khattak, N. S. (2016). Conversion of mixed low-density polyethylene wastes into liquid fuel by novel CaO/SiO₂ catalyst. *Journal of Polymers and the Environment*, 24, 255–263. <https://doi.org/10.1007/s10924-016-0768-5>

GIS-Based Geospatial Assessment of Net Land Surface Change for Identifying Erosion-Accretion Hotspots Along the Padma River in Bangladesh

MD. TAREQUL ISLAM^{1,2*}, MD. SAMEER HOSSAIN TALUKDER²,
MD. SAKIB HOSSAIN TALUKDER³, MD. SULTANUL ISLAM¹

¹ Department of Civil Engineering, Uttara University, Dhaka-1230, Bangladesh

² Institute of Water and Flood Management, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

³ Department of Water Resources Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

*Corresponding author: md.tarequlislam@uttarauniversity.edu.bd

(Received: 16 April 2026; Accepted: 30 July 2026; Published online: 11 September 2026)

ABSTRACT: Erosion is a continuous natural process that reshapes Earth's surface. It causes geohazards, loss of fertile lands, and has detrimental effects on the environment. This study provides an extensive analysis of erosion and accretion patterns over decades along both banks of the Padma River, the main distributary of the Ganges. Satellite images from Landsat and Sentinel were processed in ArcGIS to extract and compare banklines using a 5 km river-reach framework to identify net land surface change and erosion-accretion hotspots. The temporal scope included both long-term decadal analysis from 1990 to 2020 and recent annual assessment from 2016 to 2022; however, the two periods were interpreted as period-specific estimates rather than statistically tested temporal trends. Results show that erosion dominated accretion across both banks. During 1990-2020, average erosion and accretion rates on the left bank were 464 ha/year and 18 ha/year, respectively, while the corresponding rates on the right bank were 758 ha/year and 285 ha/year, respectively. For the recent 2016-2022 period, the average erosion and accretion rates were 221 ha/year and 177 ha/year on the left bank, and 1035 ha/year and 171 ha/year on the right bank, respectively. These values indicate that the right bank experienced higher erosion activity than the left bank in both the long-term and recent assessment periods. Finally, the 5 km chainage-based hotspots analysis using the net land surface change identified the most critical localized erosion-prone reaches on the right bank within Faridpur, Madaripur, and Shariatpur, while major left-bank erosion hotspots were found within Manikganj, Munshiganj, and Dhaka. These results highlight the importance of priority-based erosion monitoring and riverbank management for industrial site selection, riverbank protection planning, and identifying suitable locations for hydraulic structures along the river.

ABSTRAK: Hakisan adalah proses semula jadi berterusan yang membentuk semula permukaan bumi serta menyebabkan geobahaya, kehilangan tanah subur, dan kesan buruk terhadap alam sekitar. Kajian ini menganalisis corak hakisan dan akresi di kedua-dua tebing Sungai Padma, iaitu cabang utama Sungai Ganges. Imej satelit Landsat dan Sentinel diproses menggunakan ArcGIS bagi mengekstrak serta membandingkan garis tebing berdasarkan rangka kerja rangkaian sungai 5 km bagi mengenal pasti perubahan bersih permukaan tanah dan lokasi tumpuan hakisan-akresi. Skop temporal merangkumi analisis dekad jangka panjang dari 1990 hingga 2020 dan penilaian tahunan terkini dari 2016 hingga 2022; namun, kedua-dua tempoh ini ditafsir sebagai anggaran khusus tempoh dan bukan sebagai trend temporal yang diuji secara statistik. Dapatan menunjukkan bahawa hakisan lebih dominan berbanding akresi di kedua-dua tebing. Bagi tempoh 1990-2020, kadar purata hakisan dan

akresi di tebing kiri masing-masing ialah 464 ha/tahun dan 18 ha/tahun, manakala di tebing kanan masing-masing ialah 758 ha/tahun dan 285 ha/tahun. Bagi tempoh 2016-2022, kadar purata hakisan dan akresi ialah 221 ha/tahun dan 177 ha/tahun di tebing kiri, serta 1035 ha/tahun dan 171 ha/tahun di tebing kanan. Nilai ini menunjukkan bahawa tebing kanan mengalami aktiviti hakisan yang lebih tinggi berbanding tebing kiri pada kedua-dua tempoh penilaian. Analisis lokasi tumpuan berasaskan rangkaian 5 km mengenal pasti kawasan hakisan kritikal di tebing kanan Faridpur, Madaripur, dan Shariatpur, manakala lokasi utama di tebing kiri dikenal pasti Manikganj, Munshiganj, dan Dhaka. Dapatan ini menekankan kepentingan pemantauan hakisan dan pengurusan tebing sungai secara berkeutamaan bagi pemilihan tapak industri, perancangan perlindungan tebing sungai, dan penentuan lokasi struktur hidraulik di sepanjang sungai.

KEYWORDS: *Erosion, Accretion, ArcGIS, Landsat, Sentinel.*

1. INTRODUCTION

Riverbank erosion washes away soil from one side and deposits it on another, eventually reshaping the land. This phenomenon is harmful to the environment in many ways [1]. When the land is eroded, fertile soil and plants are lost, and the quality of the land is reduced [2]. Hence, erosion also destroys habitats for many animals [3]. The eroded soil increases the sediment particles in the water, which increases the turbidity and reduces the light that fish and other aquatic life need to survive [4]. Erosion also forces people to relocate, leaving their homes and livelihoods [5]. This relocation often leads to cutting trees to make new places suitable for living and farming [6]. These changes can create serious pressure on the local ecosystem and natural resources [7]. The occurrence of this riverbank erosion is an unpredictable disaster that occurs frequently within Bangladesh, consequently leading to severe repercussions [8].

The Padma River is one of the major rivers in Bangladesh and is morphologically highly dynamic, which makes it susceptible to erosion of the riverbanks and channel shifting [9]. Owing to its intensive channel shifting and incessant bank erosion, the river is known as ‘Sharbagrashi’, which means ‘disastrous’ in English [10]. There are two main causes of riverbank erosion: natural factors and anthropogenic (human) activities [11]. Natural factors include water discharge, runoff events, active tectonics, sediment supply, and climate change, whereas anthropogenic activities include land cover change, construction of hydraulic structures on the riverbed, and illegal sand mining, all of which accelerate riverbank erosion and channel shifting of the Padma River [11]. Three mechanisms primarily cause riverbank erosion: mass failure, fluvial entrainment, and subaerial weakening and weathering [12].

The left bank, which is located north of the bridge location, shows a certain level of resistance to erosion over considerable stretches owing to the presence of clay soils, whereas significant fluctuations in erosion and deposition patterns are observed along the right bank, which predominantly comprises fine-grained soils with no cohesive arrangement [13]. The outer bank of the flanking meandering channel is typically prone to riverbank erosion, and the erosion rate and erosion pattern are greatly influenced by the characteristics of bank materials [14]. The Padma and the Lower Meghna are alluvial rivers that are predominantly formed of sediments related to floods, while the bank materials mostly consist of sediments that are fine-grained and cohesive in nature [15]. These alluvial rivers are prone to frequent channel changes from reach to reach, owing to the continuous process of erosion and accretion [13]. Planform change has increased the sinuosity ratio and braiding index, clearly indicating differences in erosion along the river [16]. Hence, riverbank erosion in the Padma River is considered a perennial problem, resulting in the loss of a huge amount of landmass, settlements, and

livelihoods [17]. For decades, the Padma River's complex geography has caused it to meander, twist, and weave in zigzag patterns, leading to intense erosion and accretion that displace farmlands, houses, and even lives [10]. As a result, thousands of floodplain dwellers lose their land due to this riverbank erosion, rendering them homeless and landless [6]. The effects of riverbank erosion are particularly severe in the areas of Harirampur, Tongibari, Faridpur Sadar, Char Vodrasion, Dohar, Mawa Ghat, and Shiv Ghat [18].

These landless people are often left with no alternative but to settle on newly accreting char land, which leads to the development of a social and economic char environment [6]. Deforestation and forced habitation by displaced people in reserve forests result in ecological deterioration, as well as shortages of fuelwood and fodder in the area [19]. Therefore, similar to drought risk, erosion-accretion hotspots worsen livelihood and flood hazards by making displaced populations more vulnerable and decreasing ecological resilience through unplanned land-use change [20]. Owing to the socio-economic importance of the river and the development of the Padma Bridge that connects the country's capital and northern parts with its southern parts, the Padma River is very important today [21]. Therefore, timely intervention is necessary to control and manage riverbank erosion in the Padma River [22]. To identify the vulnerable zones, regular erosion-accretion analysis is crucial, and it is also necessary to monitor how the river channel changes over time [23]. Such information is important for planning protective measures such as riverbank protection works, relocation and rehabilitation of local people, risk analysis and mitigation, and land-use management. Since unplanned land-use change increases the Flood Risk Index (FRI), the identification of erosion-accretion hotspots of the Padma River could represent priority zones for flood-risk reduction and resource allocation [24]. Proper monitoring and early action can reduce land loss, protect biodiversity, and support the livelihoods of people living along the river [25].

However, earlier studies did not always quantify erosion at smaller intervals, which would help identify specific locations susceptible to erosion. Furthermore, precise data comparing the net change in land surface area between the left and right banks due to erosion and accretion are lacking. Therefore, the objective of this study is to estimate erosion and accretion along both banks of the river to make a comparison of the left and right banks of the Padma River and then identify erosion- and accretion-prone hotspots within the districts along the river at smaller intervals of 5 km in terms of net land surface change.

2. METHODS

2.1. Study Area

The Padma River gets its name from the Ganges River system, which originates from the Gangotri Glacier of the Himalayas, runs through India, and meets the Brahmaputra River of Bangladesh at Aricha, which then merges with the Upper Meghna River at Chandpur [26]. The length of the Padma River is about 120 km, and its width varies from 4 km to 8 km along its reach [27]. The Padma River is located between latitude 23°53'-23°8'N and longitude 89°38'-90°45'E and flows through the districts of Rajbari, Faridpur, Madaripur, Shariatpur, Chandpur, Munshiganj, Dhaka, and Manikganj, as shown in Fig. 1.

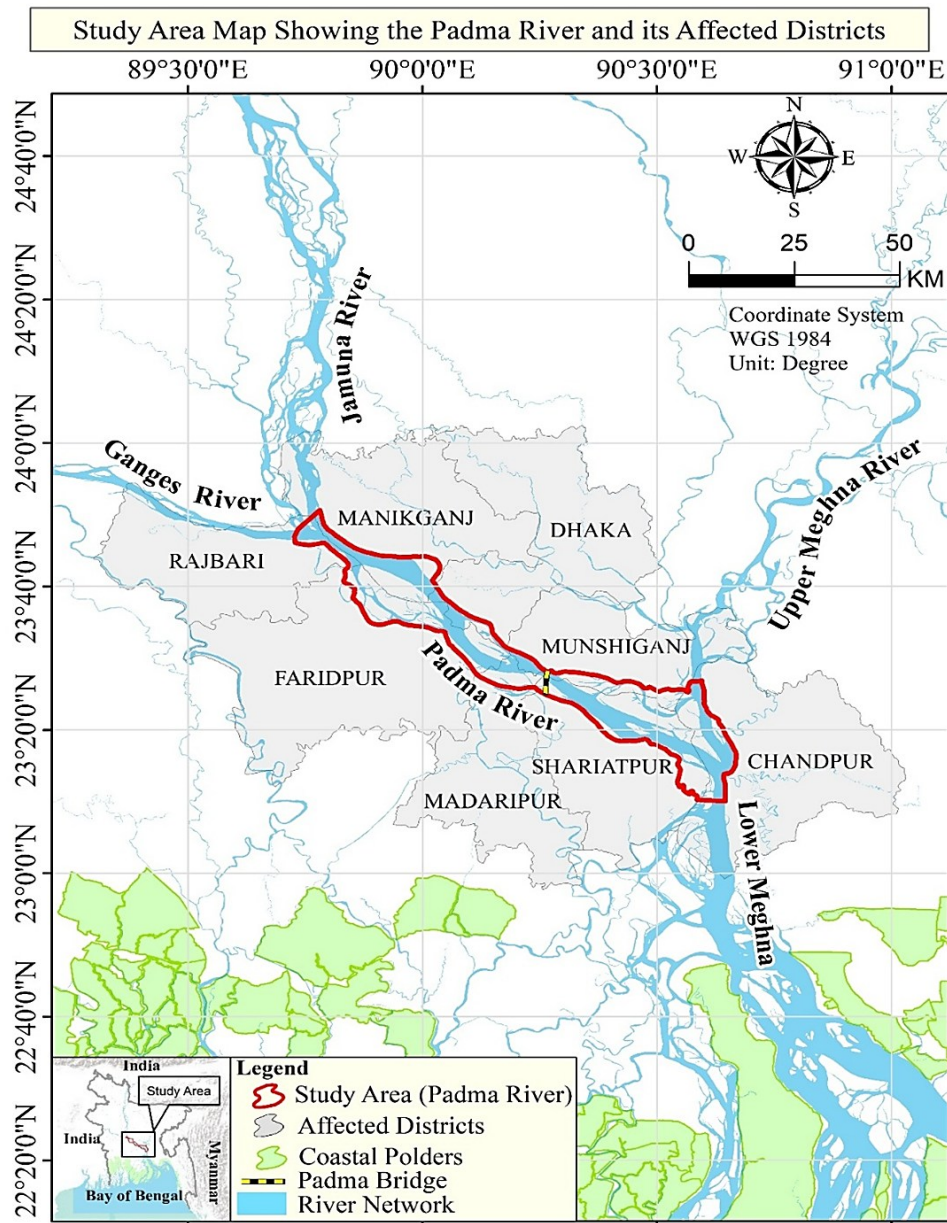


Figure 1. Study area map showing the Padma River.

2.2. Collection of Data and Image Preprocessing

To assess the erosion-accretion pattern along the Padma River, multi-temporal satellite images were collected for the years 1990, 2000, 2010, and 2016-2022. Landsat-5 TM imagery was used for 1990, Landsat-7 ETM+ imagery was used for 2000 and 2010, and Sentinel-2 MSI imagery was used for the annual analysis from 2016 to 2022. All images were selected from the dry season, mainly November to March, to reduce the influence of monsoon flooding, temporary inundation, and seasonal water-level variation on riverbank identification. Images with minimum cloud cover over the Padma River reaches were selected to ensure clear visualization of the river channel and adjacent floodplain areas.

Before bankline delineation, several preprocessing steps were performed in ArcGIS. First, the required spectral bands were selected and combined to prepare false-color composite images for better separation between water and land surfaces. For Landsat-5 TM and Landsat-

7 ETM+ images, bands 4, 3, and 2 were used, while for Sentinel-2 MSI images, bands 8, 4, and 3 were used. These near-infrared false-color combinations enhanced contrast between the river water, exposed bars, vegetation, and surrounding land, improving visual interpretation of the riverbank position.

Second, all satellite images were checked for geometric consistency and projected into a common coordinate system, the Bangladesh Transverse Mercator (BTM) system. This step ensured that images from different years and sensors could be spatially overlaid without major positional mismatch. Using a common projection also helped maintain consistency when calculating erosion, accretion, and net land surface change.

Third, image layers were visually inspected after projection to confirm proper alignment of the river reach and surrounding features. Cloud-affected portions, shadowed areas, and unclear sections were carefully avoided during bankline interpretation. The prepared composite images were then used as the base layers for manual on-screen digitization of the Padma River banklines. Banklines were delineated at a fixed scale of 1:10,000 to maintain consistency among all years. After bankline extraction, the digitized polygons were compared between consecutive time periods using GIS overlay tools to calculate eroded and accreted areas.

Table 1. Specifications of satellite images used in the study (Source: USGS)

Image Type	Bands used	Scale	Spatial resolution (m)	Image Sensor	Date of collection
Landsat-5	4, 3, 2	1:10000	30 x 30	TM	January 1990
Landsat-7	4, 3, 2	1:10000	30 x 30	ETM+	January 2000
Landsat-7	4, 3, 2	1:10000	30 x 30	ETM+	March 2010
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	November 2016
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	December 2017
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	November 2018
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	November 2019
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	November 2020
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	December 2021
Sentinel-2	8, 4, 3	1:10000	10 x 10	MSI	November 2022

After downloading the necessary images from the USGS, they were processed in ArcGIS. Multiple Bands of the Landsat images were then merged to create composite bands. These composite bands offer multiple combinations, producing colors best suited to the work at hand. Finally, the images were ensured to be cloud-free, allowing the topography to be easily perceived.

2.3. Transformation of the Coordinate System

The coordinate system of the images obtained from the USGS was WGS 1984. For convenience, the images were projected to the BTM coordinate system (Bangladesh Transverse Mercator) using ArcGIS 10. The BTM coordinate system is a method used to represent the 3-dimensional Earth's surface on a 2-dimensional plane. The more popular coordinate system, UTM (Universal Transverse Mercator), was avoided because Bangladesh falls between two (2) zones, namely UTM 45N & 46N, which can increase complexity [28].

2.4. Bankline Delineation Accuracy and Uncertainty Assessment

To ensure consistency in bankline delineation, we manually digitized riverbank lines for all selected years using the same on-screen scale of 1:10,000 in ArcGIS. We interpreted the water-land boundary carefully by comparing the spectral contrast between the river channel, exposed bars, and adjacent floodplain areas. During digitization, ambiguous sections such as

mixed pixels, temporary sandbars, and unclear water-land margins were examined cautiously to avoid overestimation or underestimation of bankline movement. The same interpretation approach was followed for all image years to keep the extracted banklines comparable across the study period.

Uncertainty in bankline extraction may arise from sensor resolution differences, georeferencing mismatches, seasonal water-level variation, mixed pixels along the river margin, and manual digitization error. Because Landsat images from earlier years had 30 m spatial resolution and Sentinel-2 images from later years had 10 m spatial resolution, very small bankline shifts near the image-resolution limit were not treated as major changes. Instead, the analysis focused on larger, spatially consistent erosion-accretion patterns that were clearly identifiable across consecutive image years. This helped reduce the influence of image-related and digitization-related uncertainty on the final estimation of erosion, accretion, and hotspot identification.

Using the Polygon feature of ArcGIS 10 software, the banklines of the river were delineated for ten (10) years, as shown in Fig. 2.

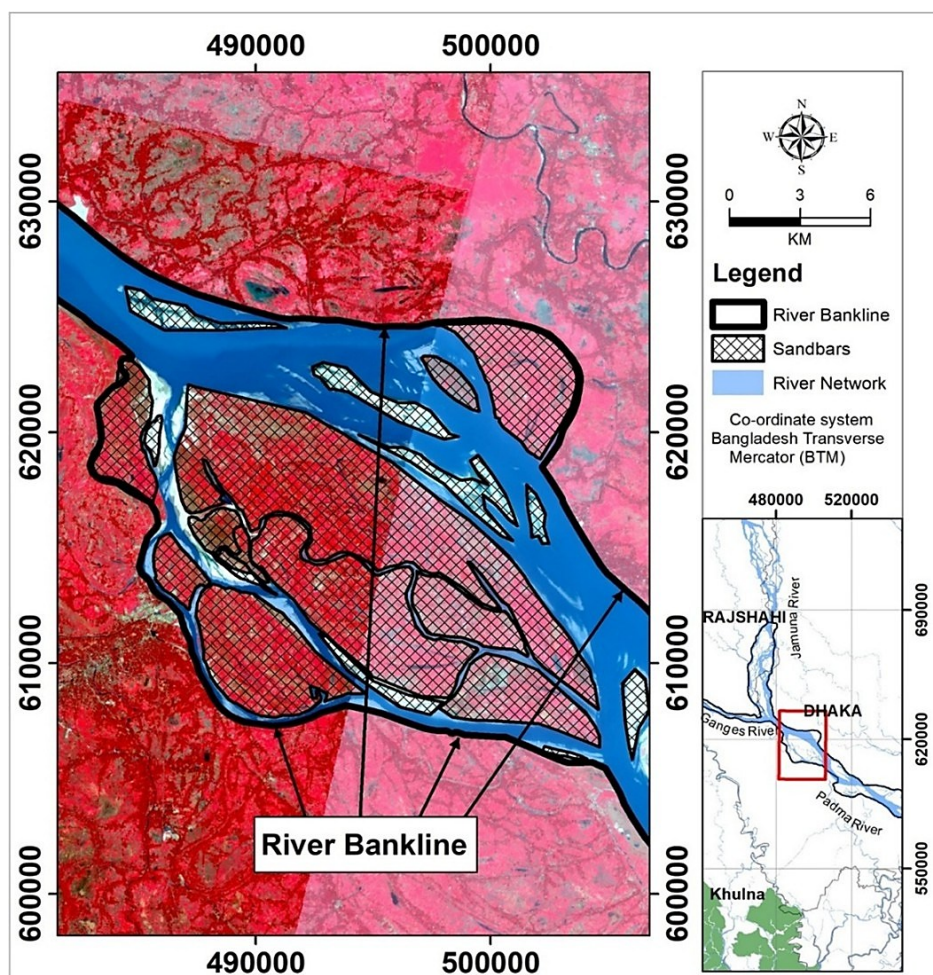


Figure 2. Bankline Delineation.

We performed both long-term decadal analyses for the periods 1990-2000, 2000-2010, and 2010-2020 and short-term year-by-year analyses for the period 2016-2022. Using the vector quantities (Bankline polygon features) as input, erosion and accretion were calculated using the “Erase Tool” in ArcGIS software. Additionally, the net change in land surface area was

calculated from the difference between erosion and accretion (Erosion ~ Accretion) values. Fig. 3 shows the digitized banklines of the Padma River.

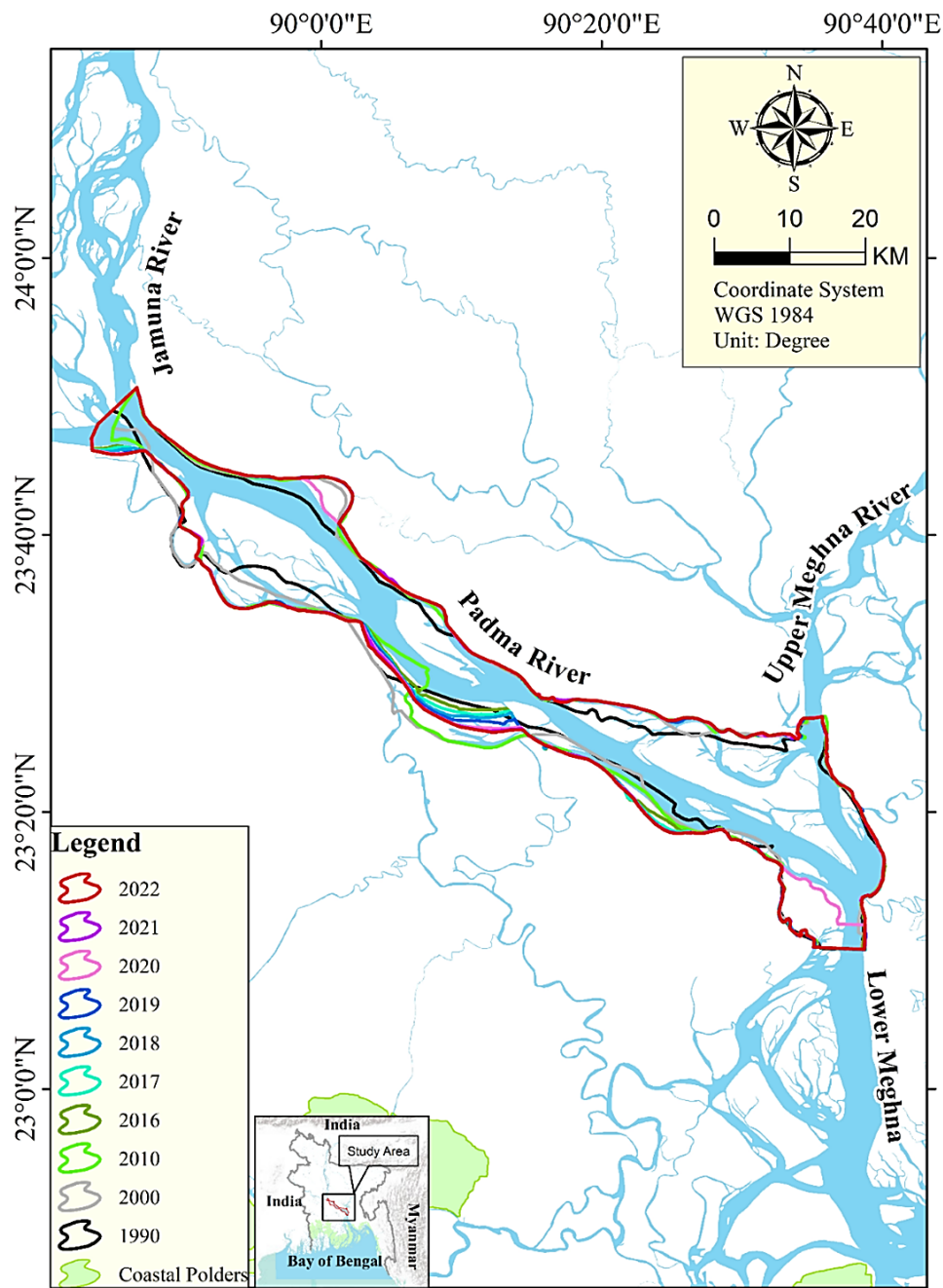


Figure 3. Digitized banklines of the Padma River.

2.5. Partitioning the River Stretch into Small 5 km Intervals

To obtain a more detailed analysis of the erosion-accretion pattern, a 100 km reach of the Padma River was divided into 5 km intervals along its longitudinal section to generate 20 identical parallel sections, as shown in Fig. 4. The starting point (0 chainage) was chosen about 5 km upstream of the Daulatdia Ferry Ghat Terminal, from where 5 km linear intervals were taken, which represented locations along both banks of the river. Considering that chainage along the riverbank would not be feasible in this case, since the right bank chainage would vary

from the left bank chainage. Moreover, the length of the bankline would vary each year, resulting in variation in chainage. This difference in chainage across the banks would not allow for comparison between the two. Thus, using a hypothetical box of equal intervals (5 km in this work) allowed comparison of erosion-accretion patterns over a similar extent along both banks of the Padma River, as well as identifying spots along the river prone to erosion.

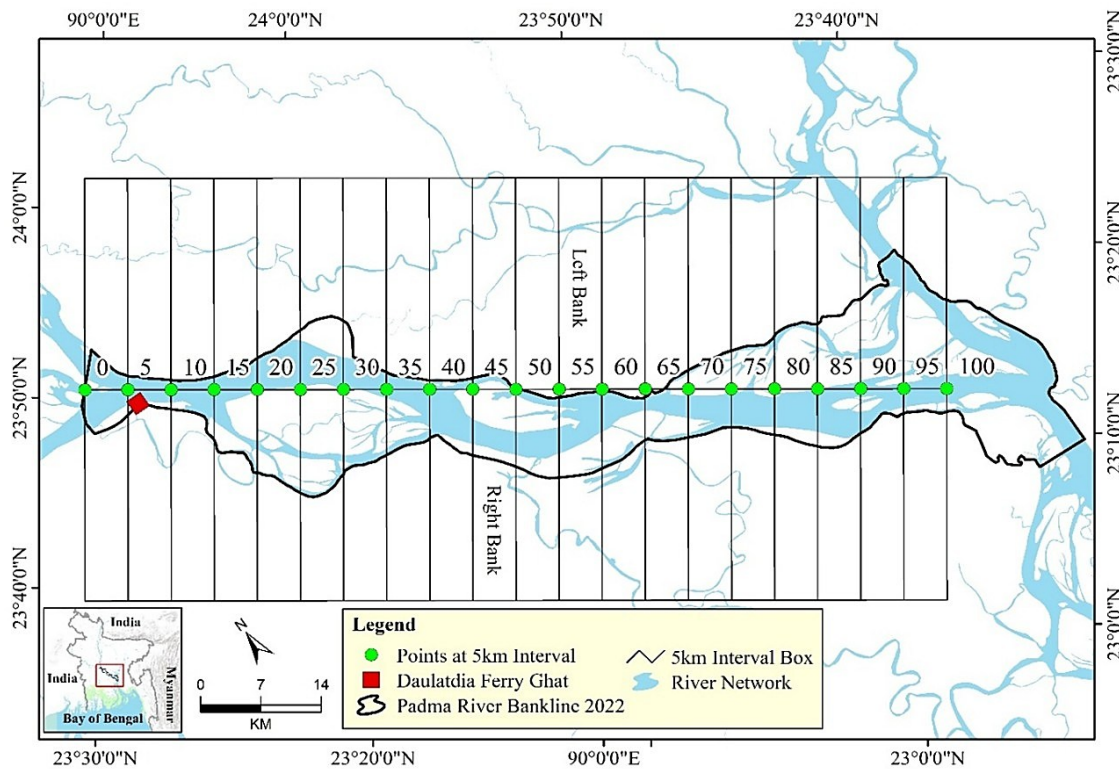


Figure 4. Map showing the division of the Padma River reach into 5 km intervals.

3. RESULTS AND DISCUSSION

Estimates were made of the amount of erosion and accretion along both banks of the Padma River. Dividing the 100 km stretch of the river into smaller 5 km segments allowed visualization and comparison of erosion and accretion patterns within the same extent on either bank. The results were then used to analyze the land surface area loss in the eight affected districts. Finally, the net change in land surface area (Erosion ~ Accretion) was estimated for each bank over the last three decades (1990-2020).

3.1. Erosion and Accretion Rates for the Left and Right Banks of the Padma River

The erosion and accretion of the Padma River were estimated for both the left and right banks to determine total area change and average annual rates. The analysis covered the periods 1990-2000 (10 years), 2000-2010 (10 years), 2010-2020 (10 years), 2016-2022 (6 years), and 1990-2020 (30 years). Tables 2 and 3 present the estimated erosion and accretion rates for these periods. This study included both decadal and annual analyses to describe long-term and recent period-specific patterns; however, these comparisons are not presented as statistically tested temporal trends.

Between 1990 and 2000, erosion of the left and right banks was similar, with estimated erosion of 11130 ha (1113 ha/yr) and 11900 ha (1190 ha/yr), respectively. However, between 2000 and 2010, erosion was significantly higher along the right bank than the left bank, with

estimated erosion of 6230 ha (623 ha/yr) and 1940 ha (194 ha/yr), respectively. Over the 30-year period (1990 to 2020), the average erosion rate on the left bank was 464 ha/yr, and the accretion rate was 18 ha/yr.

Table 2. Erosion and accretion rates for the left bank of the Padma River

Duration	Erosion (ha)	Rate of Erosion (ha/yr)	Accretion (ha)	Rate of Accretion (ha/yr)
1990-2000 (10 years)	11129	1113	140	14
2000-2010 (10 years)	1937	194	235	23
2010-2020 (10 years)	866	87	167	17
2016-2017 (1 Year)	139	139	174	174
2017-2018 (1 Year)	87	87	84	84
2018-2019 (1 Year)	85	85	102	102
2019-2020 (1 Year)	304	304	325	325
2020-2021 (1 Year)	527	527	203	203
2021-2022 (1 Year)	183	183	176	176
2016-2022 (6 years)	1324	221	1065	177
1990-2020 (30 years)	13932	464	541	18

During 2010-2020, the right bank continued to show a higher erosion rate than the left bank, with estimated erosion of 4604 ha (460 ha/year) and 866 ha (87 ha/year), respectively. These corrected values were checked against Tables 2 and 3 to ensure numerical consistency between the text and tabulated results. For the decadal analysis, the erosion and accretion rates during 1990-2020 were estimated by averaging the total eroded and accreted areas over each corresponding 10-year period. Over the 30-year period (1990 to 2020), the average erosion rate on the right bank was 758 ha/yr, and the accretion rate was 285 ha/yr.

Table 3. Erosion and accretion rates for the right bank of the Padma River

Duration	Erosion (ha)	Rate of Erosion (ha/yr)	Accretion (ha)	Rate of Accretion (ha/yr)
1990-2000 (10 years)	11897	1190	1061	106
2000-2010 (10 years)	6227	623	6937	694
2010-2020 (10 years)	4604	460	563	56
2016-2017 (1 Year)	1589	1589	149	149
2017-2018 (1 Year)	992	992	145	145
2018-2019 (1 Year)	1022	1022	113	113
2019-2020 (1 Year)	248	248	486	486
2020-2021 (1 Year)	1880	1880	108	108
2021-2022 (1 Year)	481	481	22	22
2016-2022 (6 years)	6212	1035	1023	171
1990-2020 (30 years)	22728	758	8561	285

Fig. 5 shows the annual erosion and accretion rates of the Padma River averaged over the respective decadal periods. The right bank showed larger erosion and accretion magnitudes than the left bank, indicating greater morphological activity on the right side of the river within the study reach.

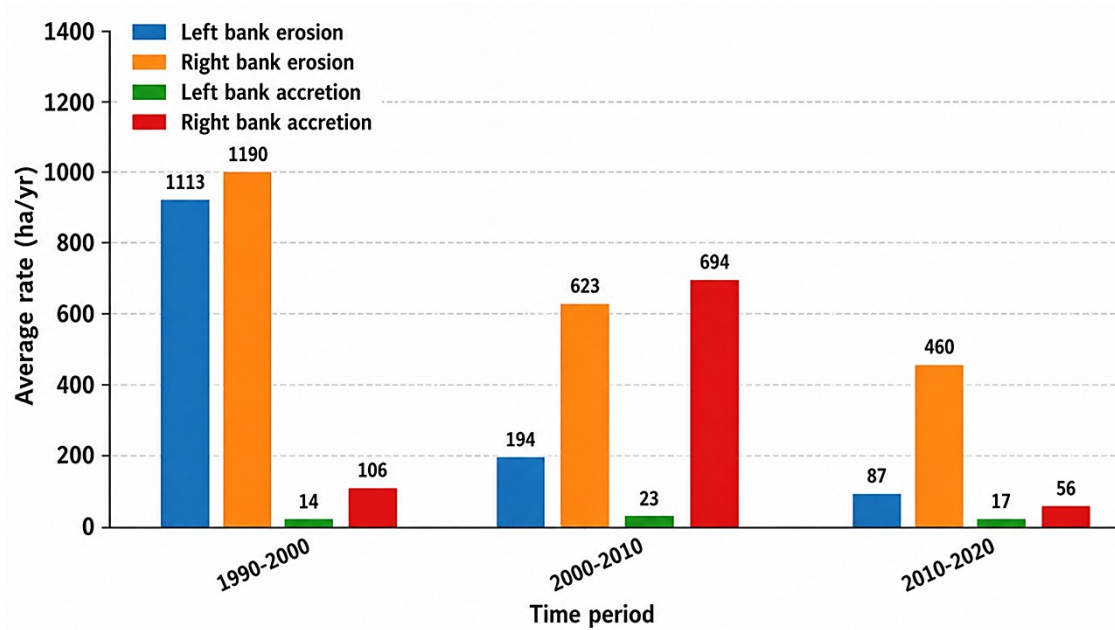


Figure 5. Annual erosion-accretion rates along the Padma River averaged over the respective decades.

Additionally, while analyzing the yearly data between 2016 and 2022, the maximum annual erosion rate was observed during 2020-2021, with approximately 1880 ha/year on the right bank and 527 ha/year on the left bank, as shown in Fig. 6. In each one-year interval except 2019-2020, the annual erosion rate on the right bank was greater than that on the left bank. Considering the 6-year time period, which is 2016 to 2022, the average erosion rate on the left bank was 221 ha/yr, and on the right bank was 1035 ha/yr, and the accretion rate on the left bank was 177 ha/yr and on the right bank was 171 ha/yr, respectively.

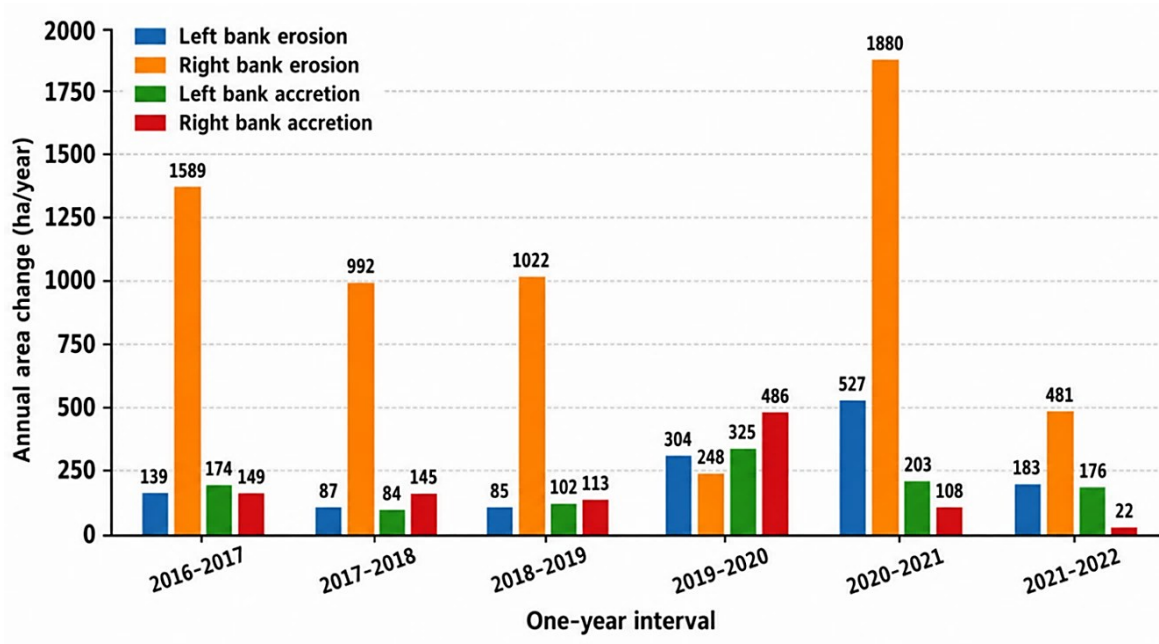


Figure 6. Annual erosion and accretion rates along both banks of the Padma River from 2016 to 2022.

Therefore, the annual observations of recent years (2016-2022) also indicate that the right bank experienced comparatively higher erosion rates than the left bank during most one-year intervals. The higher right-bank values in the recent annual analysis should be interpreted as period-specific observations rather than evidence of a statistically significant increasing trend.

The observed erosion-accretion pattern can be explained by the combined influence of bank material composition, channel curvature, discharge variation, and subaqueous bank failure processes. The right bank's higher erosion susceptibility is consistent with previous studies. The left bank of the Padma River consists of relatively more cohesive, weathered, and consolidated Pleistocene-aged sediments, which act as a natural barrier against erosion [29]. Such resistant materials are comparatively rarer along the right bank; therefore, the right bank is naturally more vulnerable to erosion [29]. This explains why the present study found greater erosion and accretion magnitudes along the right bank than along the left bank. Reports also support the right bank's higher vulnerability, noting that bank failure in the Naria region is aggravated by approximately 200 ft of alternating cross-bedded silt and micaceous fine sand, characterized by high dilatancy and a low angle of friction [30]. They further noted that deep vertical trenches along the Naria curvature increase the possibility of subaqueous slope failure [30].

Hydrological variation also plays an important role in explaining the erosion-accretion dynamics of the Padma River. The river experiences high flow during the monsoon season and low flow during the dry season, which causes repeated erosion and deposition of the riverbed as well as annual flooding [31]. Both the maximum annual discharge and average annual discharge are correlated with total annual erosion in the Padma River [13]. Similarly, a thorough analysis at the Mawa reach of the Padma River reported a strong correlation between annual riverbank erosion and average annual discharge [32]. It was observed that, during and after the 2020 monsoon floods, the largest flood extent occurred in the first week of August, whereas the highest riverbank erosion occurred in July, when both maximum and average river discharge were at their peak [32]. This suggests that bank erosion is more strongly associated with peak river discharge than with maximum flood inundation.

3.2. District-Wise Erosion and Accretion Analysis

The districts directly affected by the Padma River's erosion-accretion process include Dhaka, Manikganj, Rajbari, Faridpur, Munshiganj, Madaripur, Chandpur, and Shariatpur. Among these eight districts, four (Dhaka, Manikganj, Munshiganj, and Chandpur) are on the left bank of the Padma River, while the other four (Rajbari, Faridpur, Madaripur, and Shariatpur) are on the right bank. Tables 4 and 5 provide detailed estimates of erosion, accretion, and net change in land surface area for these districts, along with their respective erosion rates over the past 30 years (1990-2020).

On the left bank, the total amount of eroded land across the districts along the Padma River was approximately 12499 ha over the last 30 years (1990-2020), while the total amount of accreted land was approximately 815 ha. Net land loss during that period was approximately 11684 ha (erosion), highlighting the vulnerability of these districts to erosion compared with accretion (Table 4). Considering the annual average erosion rate calculated over the 30-year period from 1990 to 2020, Munshiganj experienced the highest level of riverbank erosion (174 ha/yr), followed by Manikganj (157 ha/yr), Dhaka (55 ha/yr), and Chandpur (31 ha/yr) (Fig. 7). In terms of the total eroded area in a single decade, it has been observed that the highest erosion occurred during 1990-2000 for all the districts, followed by lower erosion magnitudes during the subsequent decades (Fig. 8).

Table 4. Erosion and accretion analysis for affected districts located along the left bankline

District	Year Interval	Total Eroded Area (ha)	Total Accreted Area (ha)	Net Land Surface Change (ha)	Total Net Land Surface Change over 30 years (ha)	Rate of Net Land Surface Change (ha/yr)	Average Erosion Rate (ha/yr)
Dhaka	1990-2000	1247	0	-1247			
	2000-2010	50	46	-4	-1484	-50	55
	2010-2020	346	112	-234			
Manikganj	1990-2000	3671	10	-3661			
	2000-2010	793	15	-778	-4700	-156	157
	2010-2020	261	0	-261			
Munshiganj	1990-2000	3953	88	-3865			
	2000-2010	1058	142	-916	-4943	-165	174
	2010-2020	202	40	-162			
Chandpur	1990-2000	471	239	-232			
	2000-2010	287	95	-192	-558	-19	31
	2010-2020	161	27	-134			
Total for 30 years (1990-2020)		12499	815	-11684			

Table 5. Erosion and accretion analysis for affected districts located along the right bankline

District	Year Interval	Total Eroded Area (ha)	Total Accreted Area (ha)	Net Land Surface Change (ha)	Total Net Land Surface Change over 30 years (ha)	Rate of Net Land Surface Change (ha/yr)	Average Erosion Rate (ha/yr)
Rajbari	1990-2000	0	404	404			
	2000-2010	643	898	255	369	12	41
	2010-2020	597	307	-290			
Faridpur	1990-2000	4388	136	-4252			
	2000-2010	3022	5853	2831	-3130	-104	318
	2010-2020	2120	411	-1709			
Madaripur	1990-2000	5987	0	-5987			
	2000-2010	0	96	96	-5873	-196	300
	2010-2020	7	25	18			
Shariatpur	1990-2000	1575	650	-925			
	2000-2010	2727	124	-2603	-5368	-179	208
	2010-2020	1918	79	-1840			
Total for 30 years (1990-2020)		22986	8985	-14002			

On the right bank, the total amount of eroded land was approximately 22986 ha over the last 30 years (1990-2020), while the total amount of accreted land was approximately 8985 ha. Net land loss during that time period was approximately 14002 ha (erosion). Hence, districts on the right bank of the river were also more susceptible to erosion, similar to districts on the other bank. Based on the annual average erosion rate calculated over the 30-year period from 1990 to 2020, Faridpur experienced the highest level of riverbank erosion (318 ha/yr), followed by Madaripur (300 ha/yr), Shariatpur (208 ha/yr), and Rajbari (41 ha/yr) (Fig. 7). In terms of total eroded area in a single decade, it has been observed that, among the right-bank districts, Madaripur and Faridpur experienced their highest erosion during 1990-2000, followed by a comparatively lower erosion magnitudes in the subsequent decades. Conversely, Shariatpur and Rajbari showed a different decadal pattern, with the highest erosion recorded during 2000-2010, followed by lower erosion values during 2010-2020 (Fig. 8). Overall, the right-bank

districts showed higher erosion magnitudes than the left-bank districts, indicating greater erosion susceptibility along the right bank of the Padma River.

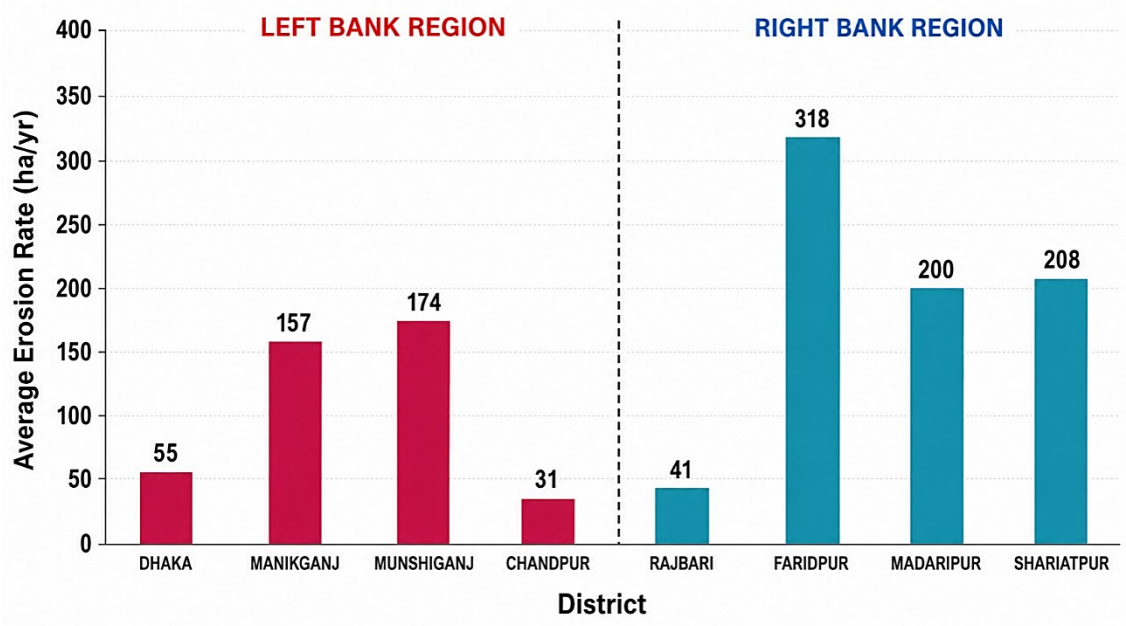


Figure 7. Average erosion rate (ha/yr) over 30 years during 1990-2020

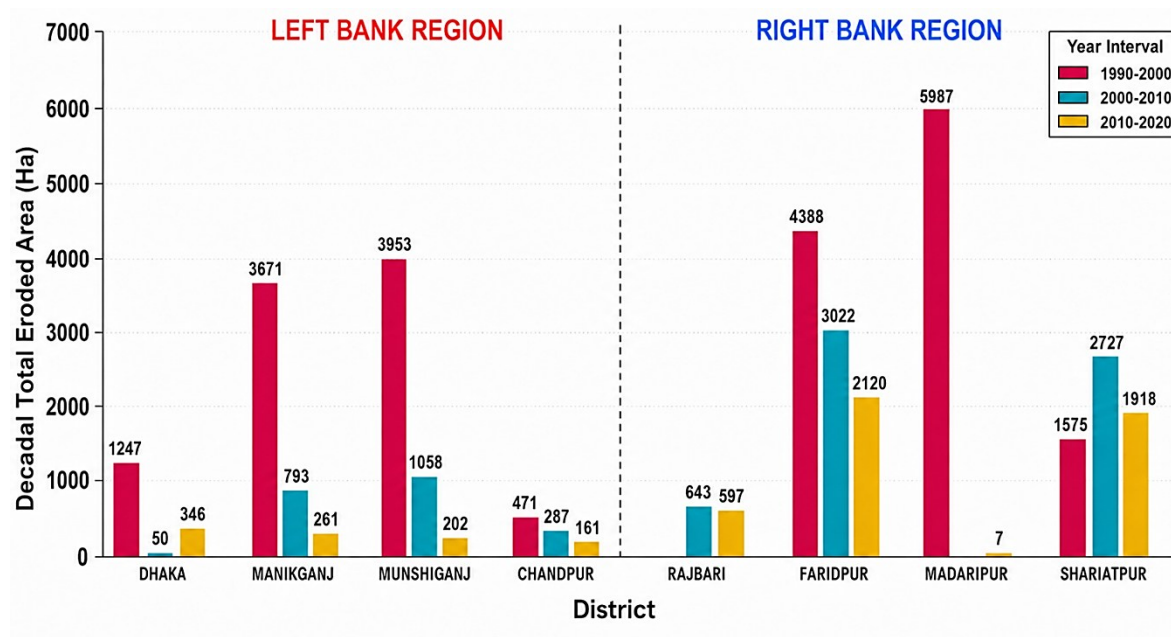


Figure 8. Total eroded area of the districts in a single decade over 1990-2020.

3.3. Net Land Change and Erosion Hotspots (5 km Interval)

Finally, a net change in land surface area from 1990 to 2020 was calculated to present a complete picture of net erosion and accretion along the river and was represented within every 5 km interval along both banks of the river, which would be helpful to visualize the area of the districts that are vulnerable to erosion. It was clear that erosion dominated accretion. Overall, the erosion-accretion process led to net land loss in most areas and net land gain in some areas. These net gains and losses were calculated within the 5 km interval along both banks of the river, deducting erosion from accretion using ArcGIS software. Fig. 9 depicts the net change

in land surface area in the Padma River over the last 30 years (1990-2020). It was observed that both the left bank (solid line) and the right bank (broken line) experienced more net land loss, as the lines spent most of their time below the 0 (zero) line or the x-axis. Positive values indicated net land gain, whereas negative values indicated net land loss.

For the left bank, the chainage-wise net land surface change analysis showed that the maximum total net land loss over the 30-year period occurred at chainage 25-30 km, which lies within Manikganj, amounting to approximately 2933 ha (98 ha/yr). Another prominent erosion peak was observed at chainage 75-80 km within Munshiganj, followed by another loss-prone reach at chainage 45-50 km within Dhaka. In contrast, net land gain along the left bank was insignificant throughout most of the study reach, indicating that the left bank was mainly characterized by localized erosion-prone zones rather than substantial accretion.

This observation was further supported by the district-wise annualized net land surface change analysis, where Munshiganj experienced the highest average net land loss rate of approximately 165 ha/year, followed by Manikganj at 156 ha/year, Dhaka at 50 ha/year, and Chandpur at 19 ha/year during 1990-2020 (Table 4).

As for the right bank, the maximum total net land loss was observed at chainage 25-30 km within Faridpur, amounting to approximately 2877 ha (96 ha/yr) over the 30-year period. Another major net land loss zone was observed between chainage 55-65 km within Madaripur, where the loss reached approximately 2644 ha (88 ha/yr) around chainage 55-60 km, followed by another erosion-prone reach around chainage 80-85 km within Shariatpur. However, unlike the left bank, the right bank also showed noticeable accretion in some reaches. The highest net land gain was observed at chainage 15-20 km within Faridpur, amounting to approximately 1470 ha (49 ha/yr), followed by another accretion-prone reach between chainage 40-50 km, also located within Faridpur. Therefore, the chainage-wise analysis identifies Faridpur as the most dynamic right-bank zone, containing both severe erosion and substantial accretion.

However, when the district-wise annual average rate of net land surface change over the 30-year period was considered, Madaripur showed the highest average net land loss rate on the right bank at approximately 196 ha/year, followed by Shariatpur at 179 ha/year and Faridpur at 104 ha/year, whereas Rajbari showed a slight average net land gain of approximately 12 ha/year (Table 5). Thus, although Faridpur contained the most severe localized erosion and accretion hotspots, Madaripur and Shariatpur experienced greater long-term average net land loss at the district scale.

This comparison indicates that the 5 km chainage-wise analysis is useful for identifying specific erosion-accretion hotspots, while the district-wise annualized rate provides a broader understanding of long-term land loss severity across administrative regions.

The district-wise erosion-accretion findings are consistent with previous studies on the Padma River. Reports show that substantial erosion-accretion activity occurred along the Padma River during 1975-2015, with the right bank showing more pronounced morphological activity [13]. Remote sensing and GIS techniques were used to assess bank erosion and channel shifting of the Padma River in Munshiganj, Madaripur, and Shariatpur districts [10]. The findings confirm that these districts have historically experienced bankline instability, supporting the present study's identification of Munshiganj on the left bank and Madaripur-Shariatpur on the right bank as major erosion-affected districts [10]. Similarly, recent erosion-accretion assessments identified Faridpur, Madaripur, Munshiganj, and Shariatpur as important hotspot districts, which agrees with the chainage-wise hotspots detected in this study.

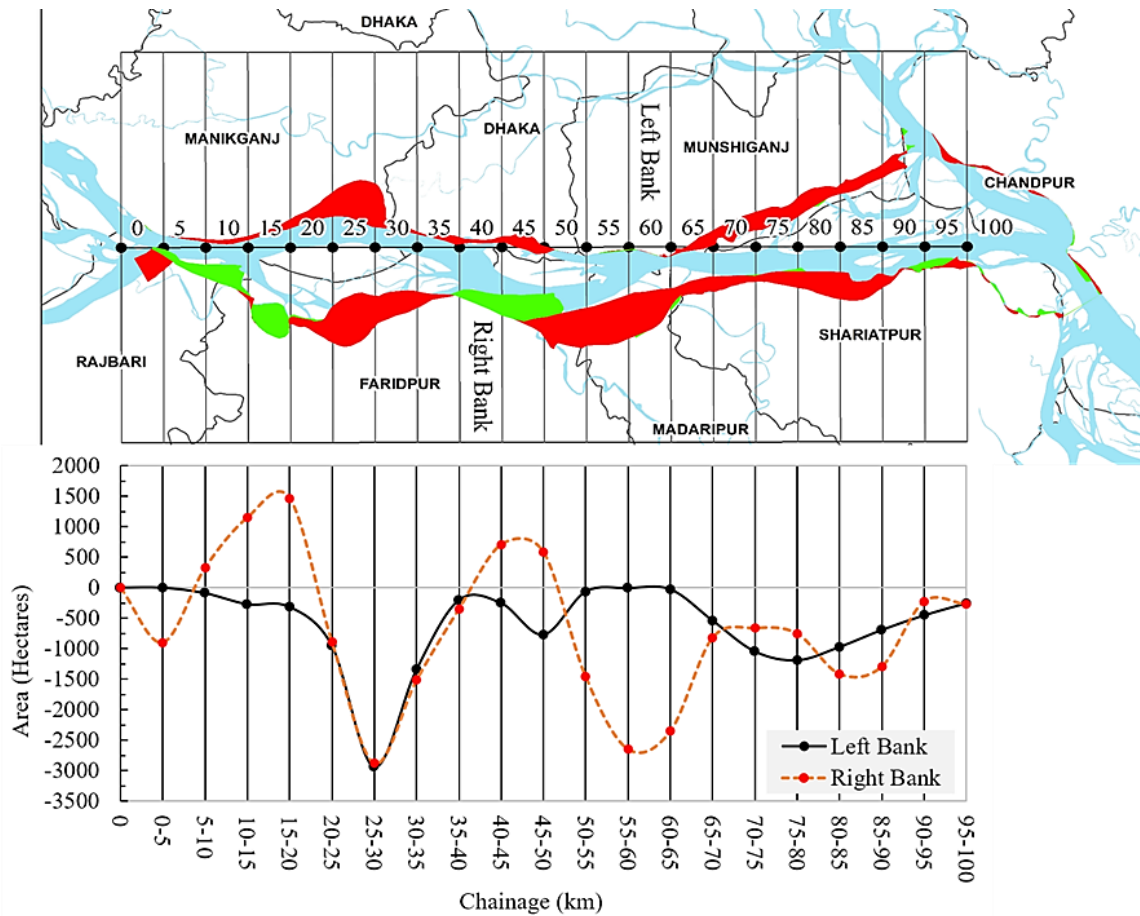


Figure 9. Net change in land surface area across left and right banks (1990-2020).

4. CONCLUSION

The study estimated erosion and accretion rates along both banks of the Padma River. The analysis included a long-term decadal assessment for 1990-2020 and a recent annual assessment for 2016-2022. These two periods were evaluated as separate period-specific estimates and were not interpreted as statistically tested increasing or decreasing trends. The results showed that the right bank experienced higher erosion activity than the left bank in both assessment periods. During 1990-2020, the average erosion rate was 758 ha/year on the right bank and 464 ha/year on the left bank. During 2016-2022, the corresponding average erosion rates were 1035 ha/year and 221 ha/year, respectively. These values indicate greater erosion susceptibility on the right bank, but they should not be interpreted as evidence of a statistically significant temporal increase or decrease because no formal trend test was conducted. The study also shows that, when erosion and accretion rates were considered separately, Faridpur had the highest average erosion rate on the right bank, followed by Madaripur, Shariatpur, and Rajbari. On the left bank, Munshiganj experienced the highest average erosion rate, followed by Manikganj, Dhaka, and Chandpur. However, this comparison does not reflect net land surface change or the specific hotspots within the districts. Accordingly, the analysis of net gain and net loss along the river showed that the most severe localized net land loss on the right bank occurred within Faridpur, followed by major erosion-prone reaches in Madaripur and Shariatpur. On the left bank, the most prominent localized erosion-prone reach was identified within Manikganj, followed by Munshiganj and Dhaka. At the district scale, the annualized net land surface change analysis showed that Madaripur and Shariatpur experienced greater long-

term average net land loss on the right bank, whereas Munshiganj and Manikganj showed the highest average net land loss rates on the left bank.

Combined with the quantitative analysis of eroded land extent, this visual representation may help implement bank protection measures in a priority-based manner. Furthermore, these findings may be useful for predicting riverbank erosion and determining the locations for constructing hydraulic structures, bridges, and industries along the riverbanks. Therefore, such analysis could be useful not only in the context of a riverine country like Bangladesh but also in countries with torrential rivers prone to erosion. In light of the above study, the authors recommend integrating sandbar formation in this study, as sandbar or island formation, locally known as char lands, is an important issue that influences the channel and flow of the Padma River, which will provide more information and a better understanding of the erosion-accretion process of the Padma River.

ACKNOWLEDGMENT

The authors acknowledge and appreciate the generosity of the USGS (United States Geological Survey) for providing the necessary images used for the analysis.

REFERENCES

- [1] Islam A, Guchhait SK (2024). Riverbank erosion: Basic concepts and approaches. In *Riverbank Erosion in the Bengal Delta*. Cham, Springer International Publishing; pp 3-22. doi: 10.1007/978-3-031-47010-3_1
- [2] Rhodes CJ (2014). Soil erosion, climate change and global food security: Challenges and strategies, *Science Progress*, 97(2):97-153. doi: 10.3184/003685014X13994567941465
- [3] Rahman MM, Islam MN (2018). Biodiversity loss by riverbank erosion: A study on the two char unions in Bangladesh. *Journal of Biodiversity & Endangered Species*, 6(1):1-6. doi: 10.4172/2332-2543.1000209
- [4] Rodrigues JN, Ortega JCG, Petsch DK, Padial AA, Moi DA, Figueiredo BRS (2023). A meta-analytical review of turbidity effects on fish mobility. *Rev Fish Biol Fisheries*, 33(4):1113-1127. doi: 10.1007/s11160-023-09785-4
- [5] Kaiser ZRMA (2023). Analysis of the livelihood and health of internally displaced persons due to riverbank erosion in Bangladesh. *Journal of Migration and Health*, 7:100671. doi: 10.1016/j.jmh.2023.100157
- [6] Malak MA, Hossain NJ, Quader MA, Akter T, Islam, MN (2021). Climate change-induced natural hazard: Population displacement, settlement relocation, and livelihood change due to riverbank erosion in Bangladesh. In *Bangladesh II: Climate Change Impacts, Mitigation and Adaptation in Developing Countries*. Cham, Springer International Publishing; pp 193-210. doi: 10.1007/978-3-030-71950-0_6
- [7] Billah MM, Majumdar A, Rahman SMA, Alam MS, Hossain MJ, Talukder J, Islam MM, Khanam T (2023). Riverbank erosion and rural food security in Bangladesh. *World*, 4(3): 528-544. doi: 10.3390/world4030033
- [8] Rahim MA, Akter N, Roy DK, Hasan I, Salman MA, Goswami S, Abir TM, Biswas RK, Rakib MMH (2024). Assessment of socio-economic impacts of river bank erosion: Enhancing resilience in the coastal belt of Bangladesh. *The Journal of Environment & Development*, 33(4):583-612.
- [9] Halder A, Mowla Chowdhury R (2023). Evaluation of the river Padma morphological transition in the central Bangladesh using GIS and remote sensing techniques. *International Journal of River Basin Management*, 21(1):21-35. doi: 10.1080/15715124.2021.1879095
- [10] Ophra SJ, Begum S, Islam R, Islam MN (2018). Assessment of bank erosion and channel shifting of Padma River in Bangladesh using RS and GIS techniques. *Spatial Information Research*, 26(6): 599-605. doi: 10.1007/s41324-018-0202-2

- [11] Berkovich KM, Zlotina LV, Turykin LA (2019). Riverbank erosion: Factors, mechanism, human activity. *Geomorphology RAS*, (2):3-17. doi: 10.31857/S0435-4281201923-17
- [12] Roy S, Barman K, Das VK, Debnath K, Mazumder BS (2020). Experimental investigation of undercut mechanisms of river bank erosion based on 3D turbulence characteristics, *Environmental Processes*, 7(1): 341-366. doi: 10.1007/s40710-019-00417-3
- [13] Billah MM (2018). Mapping and monitoring erosion-accretion in an alluvial river using satellite imagery - The river bank changes of the Padma River in Bangladesh. *Quaestiones Geographicae*, 37(3): 87-95. doi: 10.2478/quageo-2018-0027
- [14] Ielpi A, Lapôtre MGA, Gibling MR, Boyce CK (2022). The impact of vegetation on meandering rivers. *Nature Reviews Earth & Environment*, 3(3):165-178. doi: 10.1038/s43017-021-00249-6
- [15] Miah MD, Chowdhury NA, and Ali Y (2025). Monitoring multi-decadal riverbank dynamics of the Padma River in Bangladesh using geospatial analysis and DSAS model. *Discover Geoscience*, 3(1):238. doi: 10.1007/s44288-025-00353-w
- [16] Okonofua ES, Emeribe CN, Dewingong CL, Butu AW (2023). Concept of channel planforms changes on potential river flooding. *Earth Science Malaysia*, 7(1):20-28. doi: 10.26480/esmy.01.2023.20.28
- [17] Baksh MA, Akter S, Pramanik MS, Bobi F (2022). Disaster affected people's vulnerability assessment through addressing Padma River bank erosion. In *A System Engineering Approach to Disaster Resilience*. Volume 205. Edited by Ghosh C, Kolathayar S, Singapore, Springer Nature Singapore; pp 49-67. doi: 10.1007/978-981-16-7397-9_4
- [18] Khan MFA (2022). Exploring hazard risks and people's adjustment: A study in the downstream of the Padma Multipurpose Bridge Project. PhD thesis. BRAC University. Retrived from <http://hdl.handle.net/10361/17214>
- [19] Hossain B, Shi G, Ajiang C, Sarker MNI, Sohel MS, Sun Z, Yang Q (2022). Climate change induced human displacement in Bangladesh: Implications on the livelihood of displaced riverine island dwellers and their adaptation strategies. *Frontiers in Psychology*, 13:964648. doi: 10.3389/fpsyg.2022.964648
- [20] Tingsanchali T, Piriawong T (2018). Drought risk assessment of irrigation project areas in a river basin. *Engineering Journal*, 22(1):279-287. doi: 10.4186/ej.2018.22.1.279
- [21] Hawlader M, Nazmul Alam S, Akter Epe I, Hossain T (2024). A Study on the impact of the Padma Multipurpose Bridge on the economy of Bangladesh. *International Journal of Innovative Science and Research Technology (IJISRT)*, 9(5):2493-2497. doi: 10.38124/ijisrt/IJISRT24MAY2010
- [22] Rahman MN, Sharifuzzaman SASM, Sarker MNI, Lo K (2025). Socioeconomic impacts and migration dynamics of riverbank erosion. *npj Climate Action*, 4(1):102. doi: 10.1038/s44168-025-00306-6
- [23] Langat PK, Kumar L, Koech R (2019). Monitoring river channel dynamics using remote sensing and GIS techniques. *Geomorphology*, 325:92-102. doi: 10.1016/j.geomorph.2018.10.007
- [24] Kittikhun C, Pilailar S, Chittaladakorn S, Jhonpadit E (2020). Flood Risk Index for land-use changes: Case study of Pakpanang River basin. *Engineering Journal*, 24(5):25-40. doi: 10.4186/ej.2020.24.5.25
- [25] Hemmelder S, Marra W, Markies H, and De Jong SM (2018). Monitoring river morphology & bank erosion using UAV imagery – A case study of the river Buëch, Hautes-Alpes, France. *International Journal of Applied Earth Observation and Geoinformation*, 73:428-437. doi: 10.1016/j.jag.2018.07.016
- [26] De Heer JM, Choudhury GA (2025). The Bangladesh delta and its challenges. In *Strategies for Adaptation to Climate Change in a Transformative Approach*. Cham, Springer Nature Switzerland; pp 71-114. doi: 10.1007/978-3-031-88884-7_5
- [27] Eshita NR, Bhuiyan MAH, Saadat A(2023). Recent morphological shifting of Padma River: Geoenvironmental and socioeconomic implications. *Natural Hazards*, 117(1):447-472.

- [28] Mahato PS(2024). A study of morphological changes in the coastal areas and offshore islands of Sudarban coastline using remote sensing. In *Soil Erosion - Risk Modeling and Management*. Edited by Mahmood S, IntechOpen. doi: 10.5772/intechopen.112243
- [29] Sarker MH (2005). Impact of Upstream Human Interventions on the Morphology of the Ganges-Gorai System. In *The Ganges Water Diversion: Environmental Effects and Implications*. Dordrecht, Kluwer Academic Publishers; pp 49-80. doi: 10.1007/1-4020-2480-0_4
- [30] Karim MF, Khandaker NI, Haque S, Rahman MZ, Hassan MQ, Kamal AMM, Ahmed M, Sayeed BA (2019). An engineering geological anatomy of the Padma River bank failure and erosion, 2018: A case study of Naria bank section, Bangladesh. *GSA Annual Meeting in Phoenix, Arizona, USA*.
- [31] Islam MT (2010). River channel migration: A remote sensing and GIS analysis. In *ESA Living Planet Symposium*, 52:480.
- [32] Aishi AF, Hasan K, Aurib K (2024). Spatiotemporal analysis of 2020 monsoon floods and its impacts along selected reaches of the Padma River, Bangladesh. *International Journal of River Basin Management*, 22(4): 499-513. doi: 10.1080/15715124.2023.2187401

Repeated Loading Behavior of CFRP-Strengthened Continuous RC Beams: Experimental Investigation, Finite Element Analysis, and Predictive Modeling

ZEINAB R. MOUBAREK^{1*}, EHAB M. LOTFY²,
MANAR A. AHMED², ERFAN A. LATEF², EZZAAT A. SALLAM³

¹Dept. of Civil Engineering, Obour High Institute for Engineering and Technology, Cairo, Egypt

²Dept. of Civil Engineering, Faculty of Engineering, Suez Canal University, Egypt

³Dept. of Civil Engineering, Faculty of Engineering, Port Said University, Egypt

*Corresponding author: zredal251@gmail.com

(Received: 11 June 2026; Accepted: 1 July 2026; Published online: 11 September 2026)

ABSTRACT: This study investigates the repeated-loading performance and moment redistribution of continuous reinforced concrete (RC) T-beams strengthened with carbon fiber-reinforced polymer (CFRP) using a combined experimental and numerical approach. The experimental part studies the effect of reinforcement percentage and the CFRP strengthening configuration, while the FE model validates the effect of compressive strength of concrete, ratio of span to depth, CFRP reinforcement percentage, CFRP length percentage, loading frequency, and yield strength of steel on the behavior of the structure under repeated loadings. Five half-scale two-span RC T-beams having different reinforcement percentages and CFRP configurations were experimentally tested under repeated loading. CFRP strengthening increased load capacity by up to 16% and improved ductility, energy dissipation, stiffness degradation, and crack distribution; noticeable benefits were observed with the hybrid-strengthening system. The maximum ductility ratio increased from 2.7 in the control beam to 4.8 in the reinforced beams. In summary, the conflict between the stiffening effect and rotational capacity was found to be the key factor controlling the moment redistribution process. Three-dimensional non-linear finite element (FE) modeling was performed, which showed good agreement with the experimentally obtained load-displacement curve, stiffness degradation, strain evolution, and cracking pattern trends. Utilizing the validated FE database, a predictive model based on multiple regression analysis with a power-law multiplicative formulation was constructed.

ABSTRAK: Kajian ini menyiasat prestasi berterusan rasuk T konkrit bertetulang (RC) yang diperkukuh menggunakan polimer gentian karbon (CFRP) di bawah beban berulang, serta tingkah laku pengagihan semula momen melalui gabungan pendekatan eksperimen dan numerik. Pengaruh nisbah tetulang dan konfigurasi pengukuhan CFRP dinilai melalui eksperimen. Manakala model unsur terhingga (FE) digunakan bagi mengesah kesan kekuatan mampatan konkrit, nisbah rentang kepada kedalaman, nisbah tetulang CFRP, nisbah panjang CFRP, frekuensi pembebanan, dan kekuatan luluh keluli terhadap tindak balas struktur di bawah pembebanan berulang. Lima rasuk T (RC dua bentang) berskala separuh dengan nisbah tetulang dan konfigurasi CFRP berbeza telah diuji secara eksperimen di bawah beban berulang. Penggunaan pengukuhan CFRP meningkatkan kapasiti beban sehingga 16% serta menambah baik kemuluran, pelepasan tenaga, degradasi kekakuan, dan taburan retakan, dengan prestasi terbaik dicapai pada konfigurasi pengukuhan hibrid. Nilai maksimum nisbah kemuluran meningkat daripada 2.7 bagi rasuk kawalan kepada 4.8 bagi rasuk pengukuhan. Secara keseluruhan, dapatan menunjukkan bahawa keseimbangan antara peningkatan kekakuan dan kapasiti putaran merupakan faktor utama mengawal proses pengagihan semula momen. Pemodelan tiga dimensi unsur terhingga (FE) tak linear menunjukkan padanan baik

dengan dapatan eksperimen dari segi lengkung beban–sesaran, degradasi kekakuan, evolusi regangan, dan corak keretakan. Berdasarkan pangkalan data FE yang telah disahkan, satu model ramalan berasaskan analisis regresi berganda menggunakan formulasi hukum kuasa multiplikatif telah dibangunkan.

KEYWORDS: *Repeated loading, Continuous RC beam, CFRP strengthening, Moment redistribution, Finite element modeling.*

1. INTRODUCTION

Continuous RC beams are commonly used because of their structural redundancy, economical use of materials, and ability to redistribute forces [1, 2]. Unlike simple beams, continuous beams allow moment redistribution due to the formation of plastic hinges under repeated or cyclic loading [3, 4]. However, repeated loading induces complex responses associated with stiffness degradation, crack propagation, reinforcement yielding, and progressive damage accumulation, particularly in the negative moment region over intermediate supports where plastic hinges typically develop [5, 6].

Strengthening and retrofitting of the current RC structures have become relevant issues considering the aging of buildings and increased loads on them [7]. One method used for strengthening is carbon fiber reinforced polymer (CFRP) due to its high strength-to-weight ratio, corrosion protection, ease of installation, and lack of geometrical change [8]. Many studies have proved that CFRP strengthening increases the flexural resistance and stiffness of RC beams [9]. Nevertheless, data on the effect of CFRP strengthening on ductility and moment redistribution are lacking, especially for continuous beams, where stiffness compatibility and rotation capacity play a crucial role. Moreover, most previous investigations have focused on simply supported beams under monotonic loading, providing limited insight into the repeated-loading response of continuous systems [10].

Research on CFRP-strengthened continuous reinforced concrete T-beams has been insufficiently investigated because the interaction between the slab and beam significantly influences stiffness distribution, stress transfer, and failure mode [11]. In addition, the interaction between the internal steel reinforcement and externally bonded CFRP, along with the joint effect of structural and strengthening parameters, has not been sufficiently studied [12, 13]. This makes the analysis of strengthened continuous beams more complicated than that of simply supported beams and requires further investigation that accounts for strength enhancement, deformation capacity, and moment redistribution.

In previous investigations of CFRP-strengthened continuous RC beams, most research considered only monotonic loading conditions. Akbarzadeh and Maghsoudi [14] noted that additional CFRP layers increase the load capacity of continuous beams but can decrease their ductility and moment redistribution owing to the increased stiffness. Ashour et al. [15] stated that CFRP strengthening influences the deformation capacity and failure characteristics of continuous beams. El-Ghandour et al. [16] investigated moment redistribution in CFRP-strengthened continuous beams, and Aiello et al. [17] showed that proper strengthening configuration allows good redistribution properties to be maintained. However, this research mainly focused on monotonic loading and did not consider the effects of progressive repeated loading on stiffness, residual deformations, and damage accumulation.

At the same time, most existing investigations under repeated and cyclic loading conditions focused on simply supported beams rather than continuous structures. For example, El-Saikaly and Chaallal [18] and Tanarslan and Altın [19] showed that CFRP strengthening

enhances stiffness retention, crack control, and failure resistance under fatigue and cyclic loading; however, increasing the CFRP strengthening degree decreases ductility due to excessive stiffness. Similar results were obtained in repeated-loading investigations by Fathuldeen and Qissab [20] and Wang et al. [21], where load capacity increased with CFRP strengthening, accompanied by a decrease in deformation capacity due to increased structural stiffness. Nevertheless, these studies mainly considered simply supported beams and did not account for the joint influence of repeated loading, moment redistribution, and CFRP strengthening in continuous RC beams.

To fill this knowledge gap, this study will conduct experimental tests combined with verified non-linear finite element (FE) analysis to investigate the repeated-loading response and moment redistribution of CFRP-strengthened continuous RC T-beams. After verifying the FE model, a series of parametric analyses will be performed to investigate the effect of strengthening arrangement, reinforcement properties, and other factors on load capacity, deformability, stiffness degradation, and redistribution ability in gravity-controlled repeated loading conditions.

The present study provides experimental information on the repeated-loading behavior and moment redistribution of CFRP-strengthened continuous RC T-beams, which have been studied to a relatively lesser extent than simply supported beams. In addition, a non-linear finite element approach will be established and verified to capture the response behavior for conducting parametric studies systematically. Based on the generated database, an empirical equation will be developed to predict load capacity within the tested parameter range. Overall, these contributions will enhance understanding of the interaction between stiffness increase, deformability, and moment redistribution in CFRP-strengthened continuous RC beams under repeated loading conditions.

2. EXPERIMENTAL STUDY

2.1. Specimens and Materials

A series of five half-scale two-span continuous RC T-beams were used to evaluate the influence of the reinforcement ratio and the configuration of the CFRP reinforcement on their behavior. In each specimen, the total length was 4150 mm; the span length was 2000 mm for both spans. The cross section included a web width of 120 mm, an overall depth of 300 mm, and a slab size of 500 mm × 60 mm. The intermediate RC column (150 mm × 200 mm × 150 mm), containing 4 Ø10 longitudinal bars, was utilized to represent the internal support, as shown in Figure 1.

Two reinforcing levels have been examined. In the medium reinforcing series (B-M), 4Ø16 reinforcing bars were used as bottom bars, and 2Ø12 bars were used as top bars. However, the low reinforcing series (B-L) comprised 2Ø12 reinforcing bars as bottom bars and 2Ø10 bars as top bars. All samples had the same transverse reinforcement, consisting of Ø8 plain steel stirrups at 100 mm spacing, and slab reinforcement of 5Ø8/m. All beam specimens were designed following the Egyptian Code of Practice for Design and Construction of Reinforced Concrete Structures (ECP 203-2020) [22] so that flexural failure occurred before shear failure. The yield strength of plain steel stirrups was 240 MPa, which complies with the provisions of ECP 203-2020. The specimen identification includes the beam type (B), reinforcement level (M or L), and strengthening configuration (SL, DL, or DL+D), as summarized in Table 1.

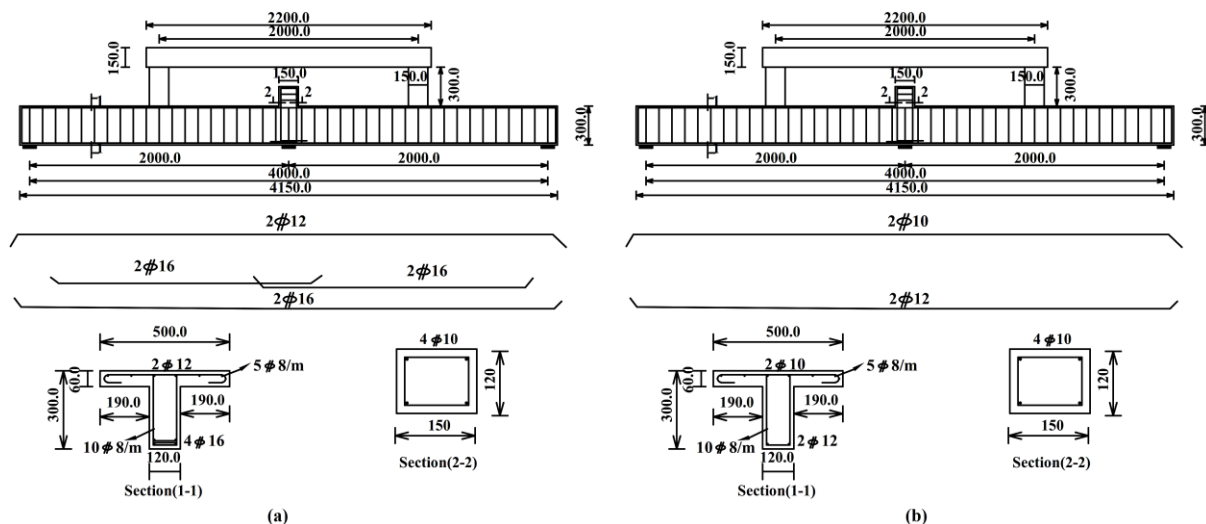


Figure 1. Geometry and reinforcement layout of the tested continuous RC T-beam specimens: (a) medium-reinforcement series (B-M); (b) low-reinforcement series (B-L).

Table 1. Configuration of tested continuous RC T-beam specimens and CFRP strengthening schemes

Specimen ID	Strengthening Scheme	F_{cu} (MPa)	Top Reinforcement (Support)	Bottom Reinforcement (Mid-span)	Stirrups
B-M-CON	Unstrengthened (control)	30	2Ø12	4Ø16	Ø8 @ 100 mm
B-M-SL	SL (single-layer CFRP)	30	2Ø12	4Ø16	Ø8 @ 100 mm
B-L-SL	SL (single-layer CFRP)	30	2Ø10	2Ø12	Ø8 @ 100 mm
B-L-DL	DL (double-layer CFRP)	30	2Ø10	2Ø12	Ø8 @ 100 mm
B-L-(DL+D)	DL+D (hybrid CFRP)	30	2Ø10	2Ø12	Ø8 @ 100 mm

The selected reinforcement layouts were intended to provide different stiffness and deformation behavior and their interaction with CFRP reinforcement under repeated loading. An unstrengthened low-reinforcement control beam was not included because the experimental program was designed to evaluate representative CFRP strengthening configurations rather than to establish a full factorial comparison between reinforcement ratio and strengthening level. Consequently, the individual effects of reinforcement ratio and CFRP strengthening could not be fully isolated experimentally. As a result, direct comparisons between the medium- and low-reinforcement series inherently reflect their combined influence, whereas the low-reinforcement results are interpreted primarily in terms of the incremental effect of increasing the CFRP configuration from single-layer SL to DL and DL+D. To complement the experimental observations, the validated finite element model was subsequently used to vary the reinforcement ratio and CFRP parameters independently, thereby quantifying the individual contribution of each variable.

All specimens were cast with normal-weight concrete with a target compressive strength of 30 MPa. The 28-day strengths of the cube and cylinder specimens were 30 MPa and 25 MPa, respectively, with an elastic modulus of 26 GPa. Mechanical properties of the concrete, steel reinforcement, CFRP sheets, and adhesive are shown in Table 2. Materials testing was carried out according to BS EN 12390-3 [23] for concrete and ASTM A370 [24] for steel.

To maintain uniformity in the properties of the materials used in the test, all beams were cast simultaneously. Instrumentation of the reinforcement cages was done before casting, and

the specimens were demolded 24 hours after casting. The specimens were then cured for 28 days under standard conditions, as shown in Figure 2.

Table 2. Material properties of concrete, steel, CFRP, and adhesive

Material	Property	Value
Concrete	Compressive strength, F_{cu} (MPa)	30
	Cylinder strength (MPa)	25
	Elastic modulus, E_c (GPa)	26
Steel	Yield strength, F_y (MPa)	570
	Elastic modulus, E_s (GPa)	200
	Yield strain, ϵ_y	0.00285
Plain Steel Stirrups	Yield strength, F_y (MPa)	240
	Elastic modulus (GPa)	195
	Yield strain	0.00123
CFRP	Tensile strength (MPa)	4900
	Elastic modulus (GPa)	252
	Ultimate strain	0.012
	Thickness (mm)	0.164
Adhesive	Tensile strength (MPa)	24.8
	Elastic modulus (GPa)	4.5
	Ultimate strain	0.010



Figure 2. Fabrication stages of the tested continuous RC beam specimens.

2.2. Strengthening Procedure

This paper studied three reinforcement configurations for CFRP strengthening: Single Layer, Double Layer, and Longitudinal-Diagonal. Figure 3 shows these configurations. All reinforced specimens consist of externally bonded CFRP strips, 50 mm wide and 0.164 mm thick per layer, applied to the beam soffit in the span areas. The DL configuration consists of two CFRP layers in the same orientation, whereas DL+D is another CFRP configuration with additional diagonal strips at 45° to control cracks in the most vulnerable area of negative moments.

Before applying CFRP strips, the concrete surfaces were mechanically prepared and cleaned to improve bond effectiveness. Bonding was performed with epoxy resin, and consolidation was achieved using rollers. These configurations were used to investigate the effect of CFRP amount and configuration on the flexural behavior and moment redistribution of continuously reinforced concrete beams under repeated loadings.

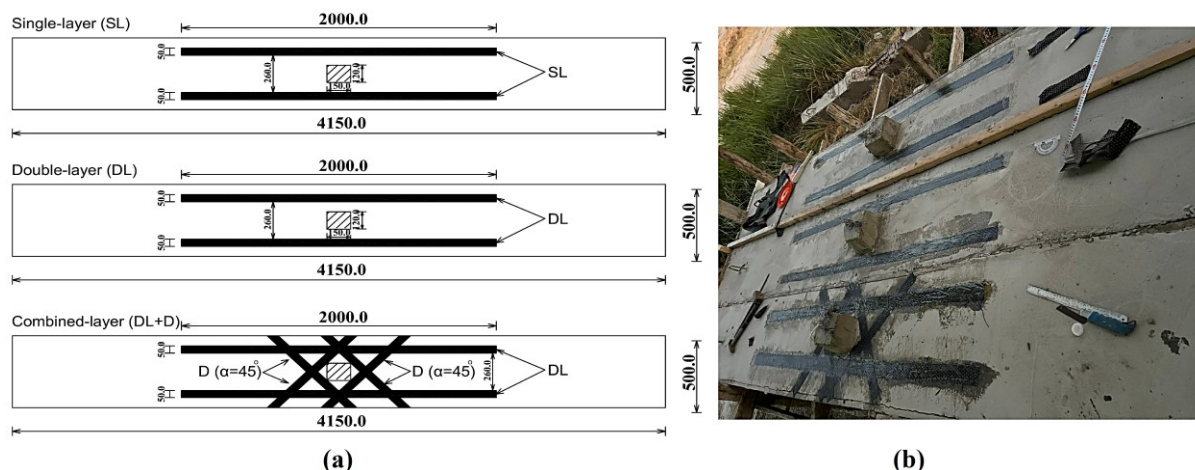


Figure 3. CFRP strengthening details: (a) strengthening layouts and dimensions; (b) representative strengthened specimen after CFRP application.

2.3. Test Setup and Instrumentation

Each specimen was tested using displacement-controlled repeated loading with a servo-controlled hydraulic actuator mounted on a reaction frame. Continuous RC T-beams were tested at both ends and an intermediate RC column to replicate the condition of the continuous beam. Vertical loading was applied using a steel spreader beam placed above the intermediate support to introduce the negative-moment region. Figure 4 schematically shows the test setup.

Vertical displacement was measured using linear variable differential transformers (LVDTs) located at the midspan of each span and recorded continuously using a computerized data acquisition system. Strain measurements focused on the negative-moment region above the intermediate support. Strain gauges were installed in the top longitudinal reinforcement, the concrete surface, and the CFRP sheets (for strengthened specimens) to investigate the strain distribution and the performance of the strengthening process. Figure 5 indicates the strain gauge positions.

The loading procedure, as shown in Figure 6, followed a standard procedure for reinforced concrete structures under repeated loading conditions [25]. A one-way sinusoidal displacement history was introduced at a constant frequency of 0.2 Hz, with unloading to zero displacement after each loading cycle. The displacement amplitude was gradually increased from about 2 mm to 120 mm, allowing analysis of the response from the elastic state to failure. Because no load reversal occurred during the experiment, the loading procedure can be considered one-way repeated loading. This type of loading is typical for reinforced concrete structural elements primarily under repeated gravity loads, such as continuous beam floors, bridge beams, and parking structures under service loads. However, because this test involves no load reversal, the proposed loading procedure cannot capture response mechanisms driven by fully reversed cycles, such as tension-compression alternations, cyclic deterioration from load reversal, and hysteresis during earthquakes. Thus, the results of this research apply only to gravity loading cases and cannot be used for seismic load applications.

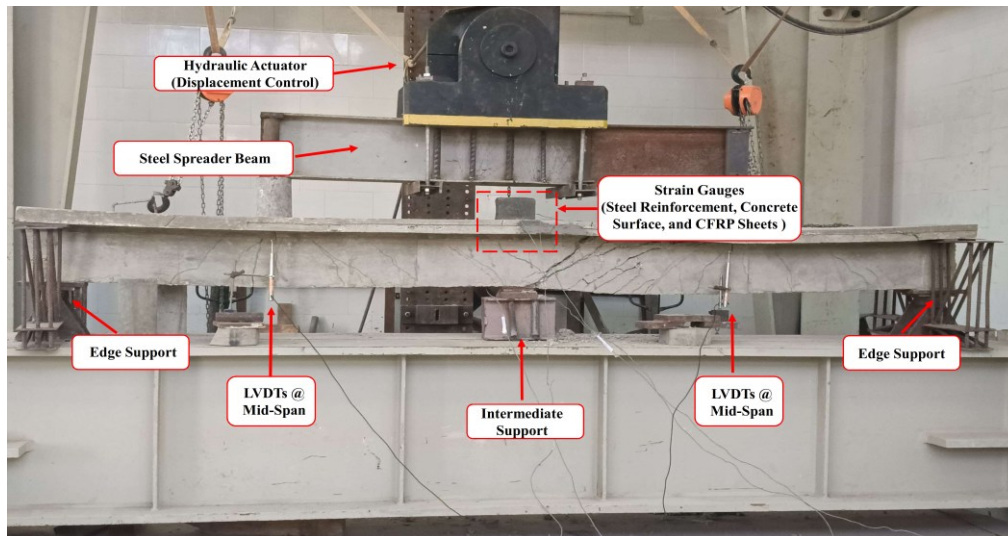


Figure 4. Experimental setup, LVDT layout, and instrumentation of the tested beams.

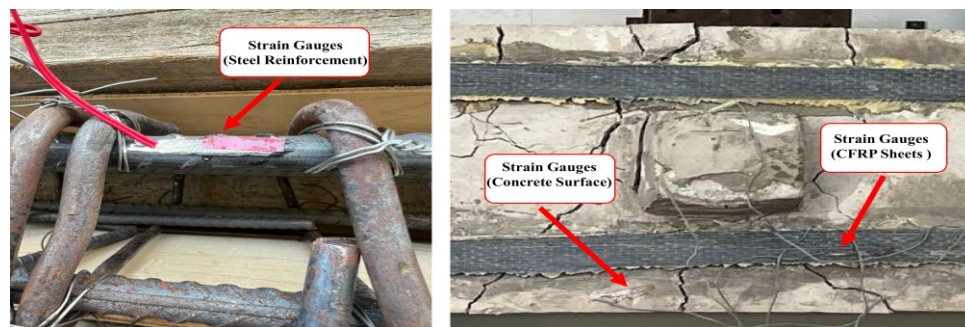


Figure 5. Strain gauge layout in the negative moment region above the intermediate support.

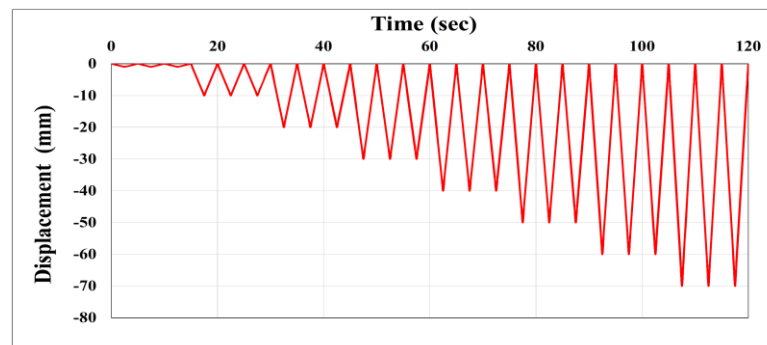


Figure 6. Displacement-controlled repeated loading protocol adopted in this study.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. Global Structural Response under Repeated Loading

The load-displacement responses shown in Figure 7 exhibited nonlinear behavior characterized by progressive cracking, residual deformation accumulation, and widening of the hysteresis loop as the displacement amplitude rises, indicating the gradual shift from elastic to inelastic behavior.

The ductility ratio μ was defined as $\mu = \Delta u / \Delta y$, where the yield displacement Δy refers to the onset of yielding of the top longitudinal reinforcement at the intermediate support ($\epsilon_y \approx$

0.00285). The ultimate displacement Δu was considered as the maximum displacement achieved before any significant strength degradation occurs or when the test is terminated. Although the yield criterion is local, this approach was applied equally to all specimens to ensure consistency.

Parameters that describe the response of the specimens are presented in Table 3. For the control specimen (B-M-CON), the peak force was 448.25 kN, while the ultimate displacement was 32.3 mm and the ductility ratio was 2.7. Adding a CFRP layer increased the peak load by 16% (519 kN), with a marginal ductility increase to 2.9. The low-reinforcement specimens have limited load capacity but high deformation capacity, with ductility ratios of 3.9, 4.6, and 4.8 for the B-L-SL, B-L-DL, and B-L-(DL+D) specimens, respectively. The hybrid-strengthening scheme resulted in the best deformation capacity because of an improved crack pattern and less strain concentration around the intermediate support. Overall, the relationship between strength and ductility is evident: higher reinforcement leads to higher load capacity, whereas lower reinforcement and CFRP strengthening increase deformation capacity.

These findings agree with the repeated-loading results reported by Fathuldeen and Qissab [20] and Wang et al. [21], who also observed that CFRP strengthening increases the load-bearing capacity of RC beams but changes the deformation capacity because of increased structural rigidity. In the current tests on continuous beams, the effect depends not only on structural rigidity but also on the amount of reinforcement and the CFRP configuration pattern.

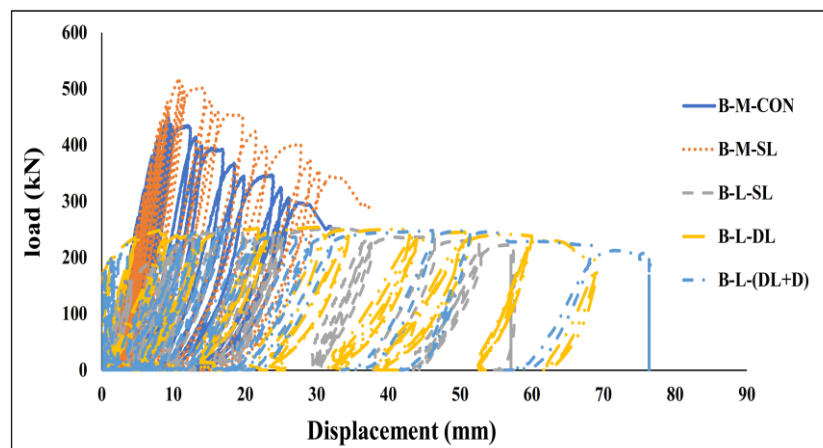


Figure 7. Load-displacement loops of the tested continuous RC beams.

Table 3. Summary of repeated loading response and displacement ductility parameters of tested specimens

Specimen	Peak Load (kN)	Δu (mm)	Δy (mm)	μ	MRI	Failure Mode
B-M-CON	448.25	32.3	12	2.7	0.57	Ductile flexural
B-M-SL	519	37.5	13	2.9	0.55	Flexural (delayed cracking)
B-L-SL	252	54.4	14	3.9	0.53	Flexural (wide cracks)
B-L-DL	254	69.1	15	4.6	0.31	Controlled flexural
B-L-(DL+D)	260	76.7	16	4.8	0.34	Flexural–shear interaction

3.2. Energy Dissipation and Stiffness Degradation

To investigate the behavior under repeated loading, the parameters of cumulative energy dissipation, residual displacement, and normalized stiffness degradation were studied. Cumulative energy dissipation was calculated as the area inside the load-displacement loops, while residual displacement and normalized stiffness were used to quantify deformation

irreversibility and stiffness degradation. Table 4 summarizes the results, and Figure 8 shows the stiffness degradation curves.

Table 4. Quantitative repeated loading performance indicators of tested specimens.

Specimen	Cumulative Energy Dissipation (kN.mm)	Residual Displacement (mm)
B-M-CON	82675	14.15
B-M-SL	93155	16.28
B-L-SL	89943	41.19
B-L-DL	105425	55.10
B-L-(DL+D)	109725	61.25

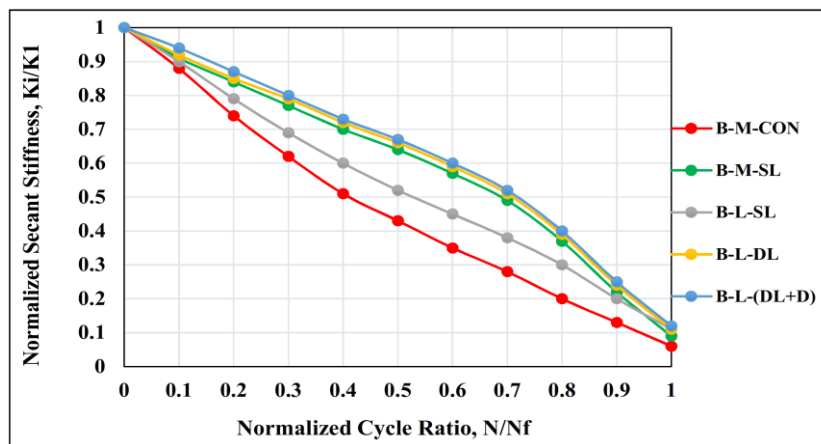


Figure 8. Normalized stiffness degradation behavior of tested continuous RC beams

The results show that CFRP reinforcement improves repeated-loading behavior by increasing energy dissipation capability, deformability, and stiffness retention. Compared with the control sample (B-M-CON), sample B-M-SL had higher cumulative energy dissipation and better stability during loading-unloading cycles. For the low-reinforcement series, significantly larger residual displacements were observed, indicating higher deformability. In this series, as the CFRP reinforcement levels increased progressively, energy dissipation capability and stiffness retention improved; hence, samples B-L-DL and B-L-(DL+D) performed better than sample B-L-SL.

From the stiffness degradation curves, all samples show decreasing stiffness due to cracking and inelastic damage accumulation. However, CFRP-reinforced samples retained a higher percentage of initial stiffness than the control sample. B-L-(DL+D) showed the best overall performance among all samples because it had the highest cumulative energy dissipation, the largest residual displacements, and the most stable stiffness development under repeated loading.

Fathuldeen and Qissab [20] and Wang et al. [21] reported similar results on progressive stiffness degradation under repeated loads. These results reveal that externally applied CFRP prevents cracks from spreading, consistent with results on stiffness maintenance in continuous beams, particularly when reinforced with CFRP.

3.3. Crack Development and Failure Modes

The crack patterns and failure modes shown in Figure 9 were generally consistent with the observed repeated loading response.

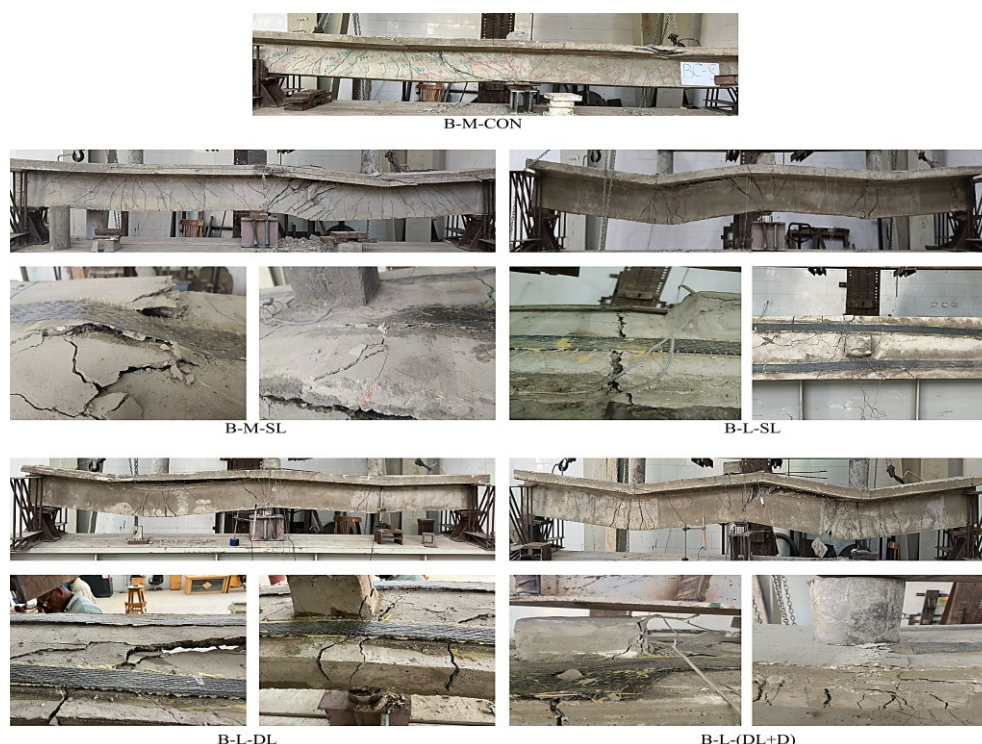


Figure 9. Crack patterns and failure modes of the tested continuous RC beam specimens.

Cracking initiated in the tension zones at locations of maximum bending moments, primarily near the intermediate support region, and progressively propagated with increasing displacement amplitudes. The test specimen B-M-CON had relatively quick crack localization and larger crack openings, indicating concentrated damage development. In contrast, the CFRP-strengthened specimens showed improved crack control, characterized by delayed crack propagation, reduced crack widths, and a more distributed crack pattern.

The influence of CFRP strengthening was particularly evident in the low-reinforcement series. Specimen B-L-SL developed wider and more localized cracks, whereas specimens B-L-DL and B-L-(DL+D) exhibited more uniform crack distribution and slower damage progression. The hybrid-strengthening configuration B-L-(DL+D) showed the most favorable crack pattern, with finer, more evenly distributed cracks throughout the critical region. Overall, CFRP strengthening reduced damage localization and promoted a more distributed cracking response, contributing to improved repeated loading stability and a more gradual failure process.

Crack behaviors found in the present investigation have been confirmed by previous studies by El-Saikaly and Chaallal [18] and Tanarslan and Altin [19], which show that CFRP reinforcement can delay crack growth and minimize crack widths. The present investigation shows that these beneficial effects were more noticeable around the intermediate supports.

3.4. Reinforcement Strain and Inelastic Behavior

The strain response of the longitudinal reinforcement at the intermediate support was used to determine the development of inelastic behavior under repeated loadings. Main parameters of strains are shown in Table 5. Steel strains increased approximately linearly during the initial loading stages and remained below the yield strain ($\epsilon_y \approx 0.00285$). As the displacement amplitude increased, yielding occurred progressively later with increasing CFRP strengthening level. The control specimen (B-M-CON) yielded at approximately 12 mm, whereas the yield

displacement increased to 13, 14, 15, and 16 mm for specimens B-M-SL, B-L-SL, B-L-DL, and B-L-(DL+D), respectively. CFRP-strengthened specimens also developed higher peak and ultimate strain values, indicating improved deformation capacity and a more stable inelastic response.

Specimen B-L-(DL+D) has the most uniform strain development among all tested specimens. This indicates more evenly distributed stresses in the critical negative moment zone. Therefore, one can conclude that CFRP strengthening delays yielding and promotes more uniform inelastic behavior development.

Since strain measurements were limited to the intermediate support region, the observed behavior is interpreted as reduced strain localization rather than direct evidence of plastic hinge relocation. This interpretation is consistent with the crack development and moment redistribution results presented in the following sections.

This delay in reinforcement yielding is consistent with the findings of Akbarzadeh and Maghsoudi [14], who found that CFRP wrapping changes the strain pattern and improves the deformation properties of continuously supported RC beams. Similarly, Fathuldeen and Qissab [20] found the same effect in repeated loading cases, where CFRP wrapping delays the structure's inelastic response. Current findings show that hybrid strengthening provides a more balanced strain distribution around the negative moment zone.

Table 5. Reinforcement strain parameters of tested RC beam specimens

Specimen	Yield Strain (ϵ_y)	Δy , steel (mm)	Peak Strain (ϵ_{max})	Ultimate Strain (ϵ_u)
B-M-CON	0.00285	12	0.011	0.012
B-M-SL	0.00285	13	0.013	0.014
B-L-SL	0.00285	14	0.014	0.015
B-L-DL	0.00285	15	0.015	0.016
B-L-(DL+D)	0.00285	16	0.016	0.017

3.5. Moment Redistribution Behavior

The moment redistribution behavior of the tested continuous RC beams was evaluated using the measured support reactions shown in Figure 10, bending moment distributions shown in Figure 11, and the moment redistribution index (*MRI*) summarized in Table 3. The experimentally measured support moments (M_{exp}) were determined from the recorded reactions using static equilibrium, while the corresponding elastic moments ($M_{elastic}$) were obtained from linear elastic analysis. The *MRI* was calculated using Equation (1).

$$MRI = \frac{M_{elastic} - M_{exp}}{M_{elastic}} \quad (1)$$

where M_{exp} is the experimentally measured support moment, and $M_{elastic}$ is the corresponding elastic support moment. Higher *MRI* values indicate a greater reduction in the negative support moment relative to the elastic prediction.

The load-reaction and moment distribution analyses show that CFRP reinforcement changes the behavior of moment distribution and force redistribution in continuous beams. Specimen B-M-CON produced the maximum support moment value (approximately 224 kN·m), with an *MRI* value of 0.57, indicating a high value of moment redistribution, even though the ductility of this specimen was rather low ($\mu = 2.7$). For the strengthened specimens, the support moments varied approximately within 126- 130 kN·m, while *MRI* values varied between 0.31 and 0.53, reflecting the influence of CFRP-induced stiffness enhancement on support moment development.

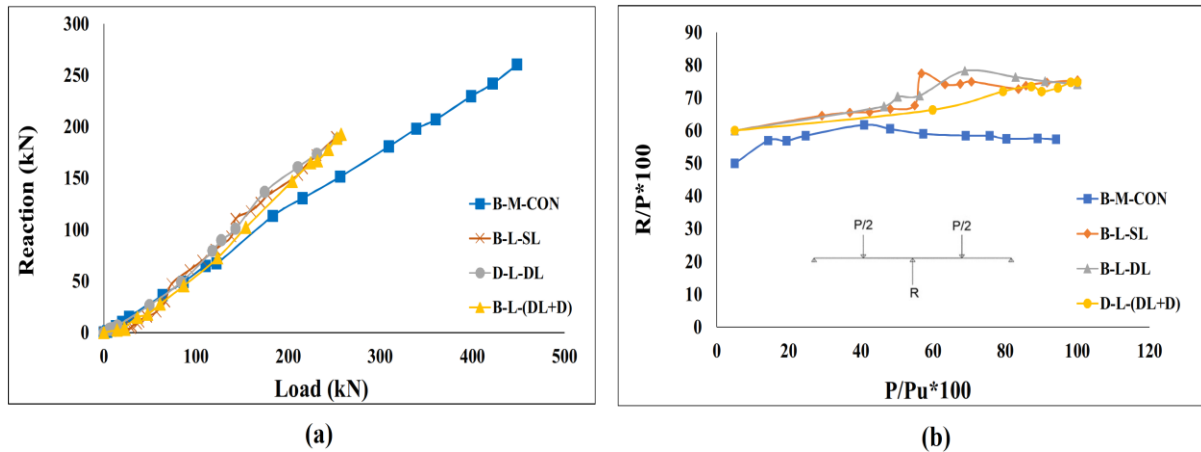


Figure 10. Load-reaction relationships of the tested continuous RC beam specimens: (a) load-reaction curves; (b) normalized load-reaction curves illustrating moment redistribution behaviour.

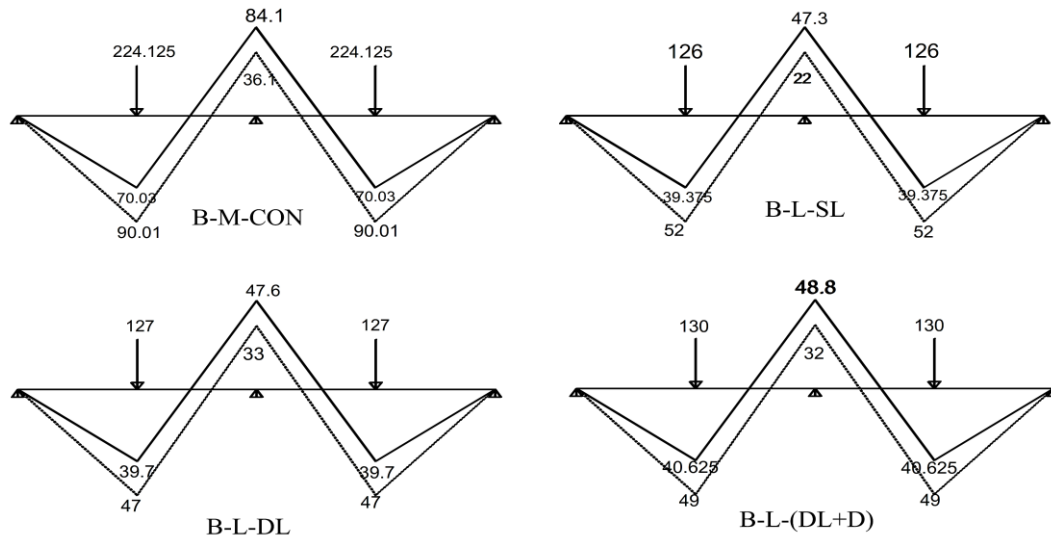


Figure 11. Bending moment distributions along the beam length for the tested beams.

Specimen B-L-SL showed the highest MRI value (0.53) among the strengthened specimens, whereas specimens B-L-DL and B-L-(DL+D) showed smaller MRI values of approximately 0.31 and 0.34, respectively, with increased beam ductility. These findings indicate a balance between rotational capacity and stiffness development, which affects moment redistribution. While CFRP strengthening improved crack control, strain distribution, and repeated-loading performance, higher strengthening levels tended to reduce redistribution efficiency by increasing local stiffness near the intermediate support. Overall, the results indicate that moment redistribution in CFRP-strengthened continuous RC beams depends on the interaction between stiffness enhancement and deformation capacity. Therefore, redistribution behavior must be considered in addition to the increase in strength.

The measured redistribution behavior agrees with previous studies highlighting the sensitivity of continuous RC beams to stiffness variation and rotational capacity. Ashour et al. [15] and El-Ghandour et al. [16] showed that CFRP strengthening affects moment redistribution in continuous beams through changes in flexural stiffness and deformation capacity. The present study observed similar behavior: higher CFRP strengthening improved strength and crack control but reduced redistribution efficiency due to increased local stiffness.

4. NUMERICAL MODELING AND VALIDATION

The experimental program provided valuable insight into the repeated-loading behavior of CFRP-strengthened continuous RC beams; however, its scope was limited by the cost and complexity of full-scale testing. Accordingly, a three-dimensional nonlinear finite element (FE) model was developed in ABAQUS and validated using the experimental results. The model was intended to reproduce the global structural response and the associated damage mechanisms governing stiffness degradation, strain development, and moment redistribution under repeated loading.

4.1. Finite Element Modeling

A three-dimensional nonlinear FE model was developed to simulate the tested continuous RC T-beams under displacement-controlled repeated loading. The model replicated the experimental geometry, reinforcement detailing, support conditions, and CFRP strengthening layouts, as shown in Figure 12.

Concrete was modeled using 8-node reduced-integration solid elements (C3D8R), with mesh refinement in the area of negative moment near the intermediate support. The reinforcing steel was represented with embedded beam elements (B31), assuming full bonding between steel and concrete. This assumption was considered appropriate because the experimental tests showed no evidence of reinforcement slip or bond-related failure, indicating that the global response was governed by flexural behavior rather than bond deterioration. The CFRP strips were modeled using 4-node shell elements (S4R) with orthotropic elasticity and fiber orientation according to the strengthening scheme. The interaction of CFRP with the concrete was simulated using tie constraints, while CFRP debonding was not considered because no premature debonding or rupture was observed experimentally.

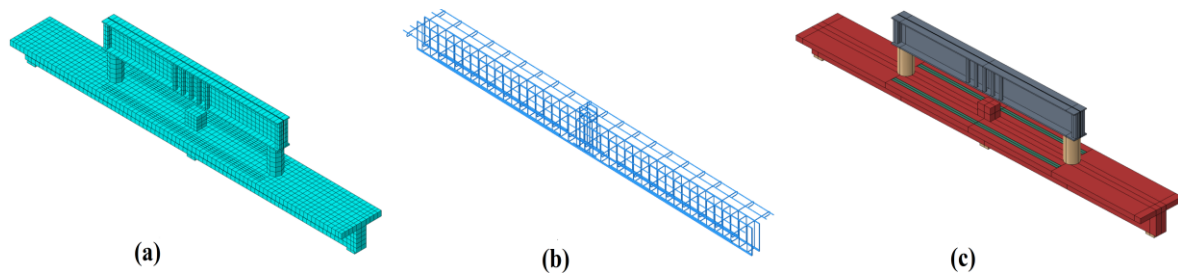


Figure 12. Finite element model of the tested continuous RC beam system: (a) concrete mesh; (b) embedded reinforcement; (c) CFRP strengthening.

Table 6. CDP parameters adopted in the FE model

Parameter	ψ	ε	f_{b0}/f_{c0}	K_c	Viscosity parameter
Value	36°	0.10	1.16	0.667	0.0005

The Concrete Damage Plasticity (CDP) material model in ABAQUS captured the nonlinear response of concrete. The parameters used in the simulation, listed in Table 6, were calibrated based on material characteristics obtained from test results and the literature. The tensile and compressive behavior of concrete included damage accumulation and stiffness reduction, whereas steel reinforcement was modeled using an elastic-plastic constitutive model with isotropic hardening. The simulation loading sequence followed the experimentally

obtained displacement-controlled method, including unloading-reloading cycles with increasing displacement steps. The simulation applied loads using rigid loading plates to ensure uniform load transfer.

4.2. Model Verification

In this work, the FE analysis was validated experimentally based on the structure's global response, strain development in reinforcements, reactions, and progressive damage. Figure 13 compares the numerical and experimental load-displacement responses, while Table 7 compares peak load and ultimate displacement.

The numerical model can predict the entire repeated-loading process, including initial stiffness, yielding, peak load, stiffness reduction, residual displacements, and unloading-reloading behavior. Peak load predictions showed a 2.6%- 11.5% difference compared to experimental peak loads, whereas ultimate displacement showed differences of 7.6%- 11.6%. Mean errors for peak load and ultimate displacement were 7.5% and 9.1%, respectively. Initial stiffness was underpredicted in some experiments, which could be due to the assumption of perfect bonding and no existing microcracks in the numerical simulation.

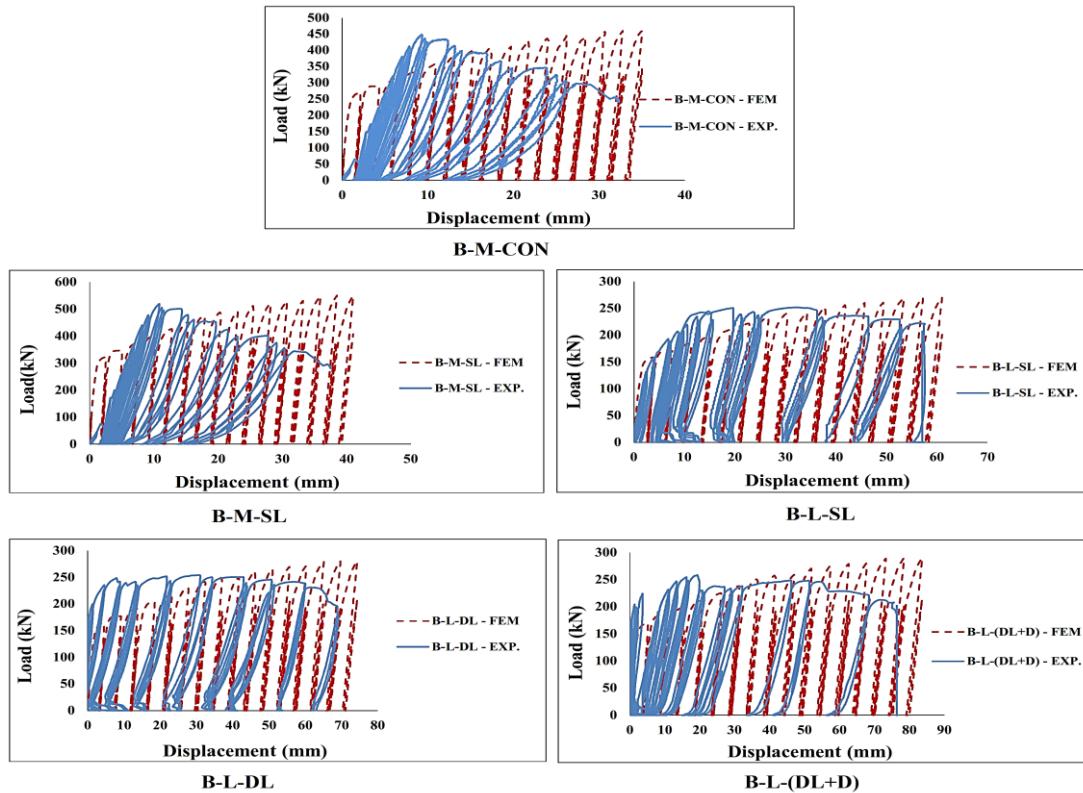


Figure 13. Comparison of experimental and numerical (FEM) load-displacement loops for all beams.

The FE analysis predicted the experimentally observed trends in reinforcement strain evolution and reaction progression, namely delayed yielding and reduced strain localization in CFRP-strengthened beam specimens. Moreover, the tensile damage contours computed based on the CDP model shown in Figure 14 showed good correspondence to the actual cracking, properly detecting regions with increased damage concentrations near the intermediate support point. Although the CDP model damage contours represent material deterioration rather than cracks, the simulation adequately captured the experimentally observed reduction in damage localization and improved cracking behavior caused by CFRP strengthening.

Table 7. Comparison between experimental and FE results

Specimen	$P_{\max}$ -EXP (kN)	$P_{\max}$ -FEM (kN)	Error	Δu -EXP (mm)	Δu -FEM (mm)	Error
B-M-CON	448.25	460	2.6%	32.5	35.0	8.4%
B-M-SL	519.0	550	6.0%	37.5	41.0	9.2%
B-L-SL	252.0	270	7.2%	54.4	60.7	11.6%
B-L-DL	254.0	280	10.2%	69.1	74.3	7.6%
B-L-(DL+D)	260.0	290	11.5%	76.7	83.6	8.9%
Mean Error	—	—	7.5%	—	—	9.1%

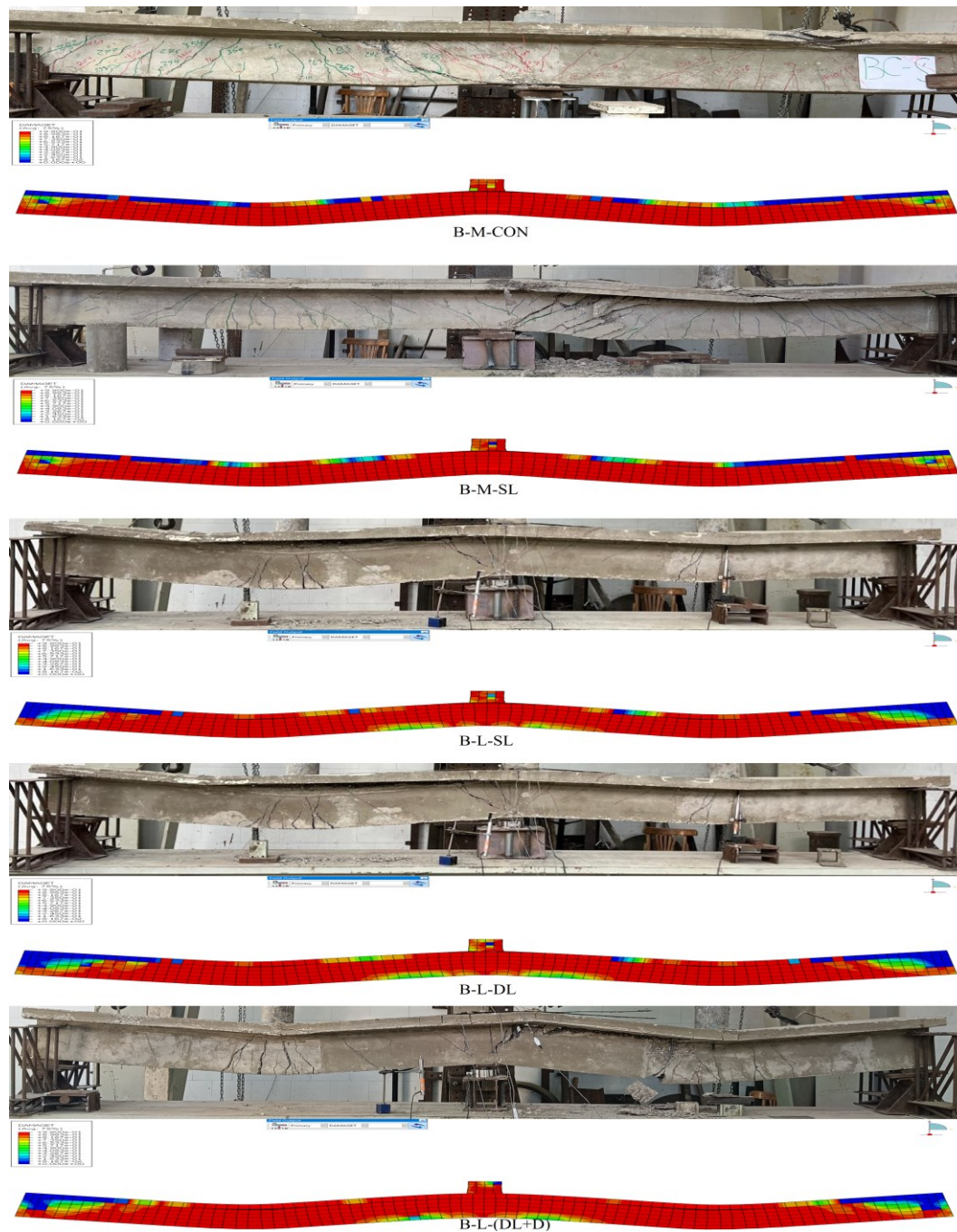


Figure 14. Comparison of numerical damage contours with experimental crack patterns for the tested beams.

Overall, the agreement between the numerical predictions and experimental observations demonstrates that the developed FE model provides a reliable framework for simulating the repeated-loading behavior of CFRP-strengthened continuous RC beams within the investigated parameter range. Nevertheless, the adopted perfect-bond assumption and the smeared-crack representation of concrete damage need to be considered when interpreting localized cracking and damage patterns.

5. PARAMETRIC STUDY AND DISCUSSION

After validating the FE model, a parametric study was performed to examine the effect of various structural, material, and strengthening parameters on the repeated-loading response of CFRP-strengthened RC continuous beams. The validated model enabled the manipulation of the investigated parameters while keeping the same geometric characteristics, boundary conditions, and loading sequence.

Specimen B-L-SL was adopted as the reference model because it exhibited a balanced combination of strength, deformation capacity, and repeated-loading performance. The investigated parameters, summarized in Table 8, included the negative reinforcement ratio (ρ_n), concrete compressive strength (F_{cu}), span-to-depth ratio (L/d), CFRP reinforcement ratio (ρ_f), CFRP length ratio (L_f/L), loading frequency (Δ), and steel yield strength (F_y). The CFRP reinforcement ratio (ρ_f) was calculated as the ratio of the total CFRP cross-sectional area to the concrete web cross-sectional area, considering the number of CFRP layers and the thickness of each layer, as expressed in Equation (2).

$$\rho_f = \frac{n_f \cdot b_f \cdot t_f}{b_w \cdot d} \quad (2)$$

where, n_f is the number of CFRP layers, b_f is the CFRP strip width, t_f is the thickness of one CFRP layer (0.164 mm), b_w is the beam web width, and d is the effective depth of the beam. The CFRP reinforcement ratio was varied by changing the number of CFRP layers (one to five layers), while maintaining a constant CFRP sheet thickness of 0.164 mm per layer.

Table 8. Variables considered in the parametric investigation (baseline: specimen B-L-SL)

Parameter	Levels Considered	Mechanical Role
Negative reinforcement ratio (ρ_n)	0.46, 0.66, 1.20 %	Governs hogging moment capacity, crack control, and stiffness degradation at supports
Concrete compressive strength (F_{cu})	30, 40, 50 MPa	Controls compression zone behavior, ductility, and failure mode
Span-to-depth ratio (L/d)	12, 15, 18	Influences flexural stiffness, deformation capacity, and moment redistribution
CFRP reinforcement ratio (ρ_f)	0.024, 0.048, 0.072, 0.096, 0.12 % (corresponding to one to five CFRP layers with a constant thickness of 0.164 mm per layer)	Enhances flexural capacity, stiffness, and repeated loading performance
CFRP length ratio (L_f/L)	0.25, 0.50, 0.75, 1.00	Affects stress distribution and debonding resistance
Loading frequency (Δ)	0.1, 0.2, 0.3 Hz	Influences rate effects, energy dissipation, and damage accumulation
Steel yield strength (F_y)	350, 570, 650 MPa	Controls yielding behavior, ductility, and energy dissipation

This study used a one-variable-at-a-time approach. In this approach, each parameter is varied individually while the rest remain constant. The study focused on capacity under

repeated loads, displacement capacity, ductility, stiffness degradation, and overall behavior. This will be discussed in the following section.

5.1. Influence of Structural and Material Parameters

Figure 15 presents the influence of the investigated structural and material parameters on the repeated-loading response. Increasing the negative reinforcement ratio (ρ_n) and steel yield strength (F_y) increased load-carrying capacity and stiffness by improving flexural resistance in the negative moment region. However, the associated increase in stiffness reduced deformation capacity and ductility by limiting curvature development and rotational demand.

The span-to-depth ratio (L/d) exhibited an opposite trend. Lower L/d ratios increased stiffness and load resistance, whereas larger ratios promoted greater deformation capacity and ductility at the expense of a moderate strength reduction. Increasing the concrete compressive strength (F_{cu}) also improved load capacity by delaying compressive damage development, although its influence on deformation behavior was less pronounced than that of reinforcement-related parameters.

Overall, the results demonstrate a clear trade-off between strength and deformation capacity. Parameters that enhance stiffness and flexural resistance generally increase load capacity but may reduce ductility and rotational capacity. Therefore, the performance of CFRP-strengthened continuous beams under repeated loading should be assessed by considering both strength and deformation characteristics, not strength enhancement alone.

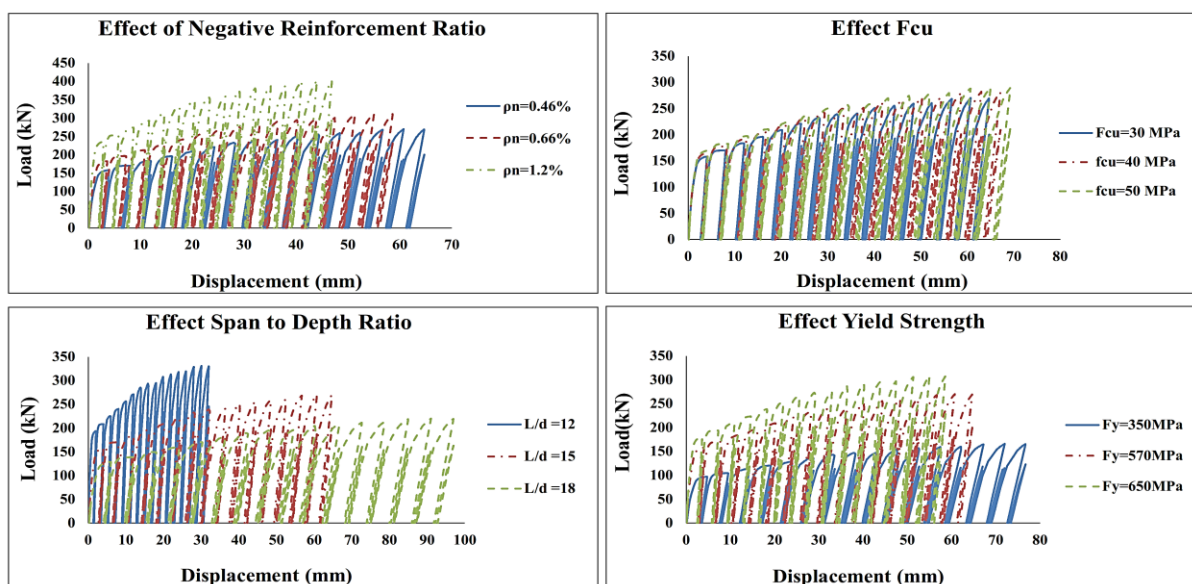


Figure 15. Influence of key structural and material parameters on the load-displacement loops of the tested beams.

5.2. Influence of CFRP Strengthening Parameters

Figure 16 shows the effect of CFRP strengthening parameters on repeated-loading behavior. Increasing the CFRP reinforcement ratio (ρ_f) increased load capacity, crack control, and stiffness. However, excessive strengthening reduced deformation capacity and rotation requirement, suggesting stiffness-controlled behavior.

The CFRP strengthening ratio (L_f/L) affected structural behavior by increasing stress transfer along the reinforced region and reducing stress concentrations at the CFRP ends. Compared with increasing the CFRP ratio alone, increasing the CFRP length provided more

consistent results, with small increases in load capacity and limited decreases in deformation capacity.

Overall, the results indicate that the effectiveness of CFRP strengthening depends not only on the amount of CFRP reinforcement but also on its distribution along the beam. Therefore, strengthening design should balance stiffness enhancement and crack control with adequate deformation capacity to ensure stable behavior under repeated loading.

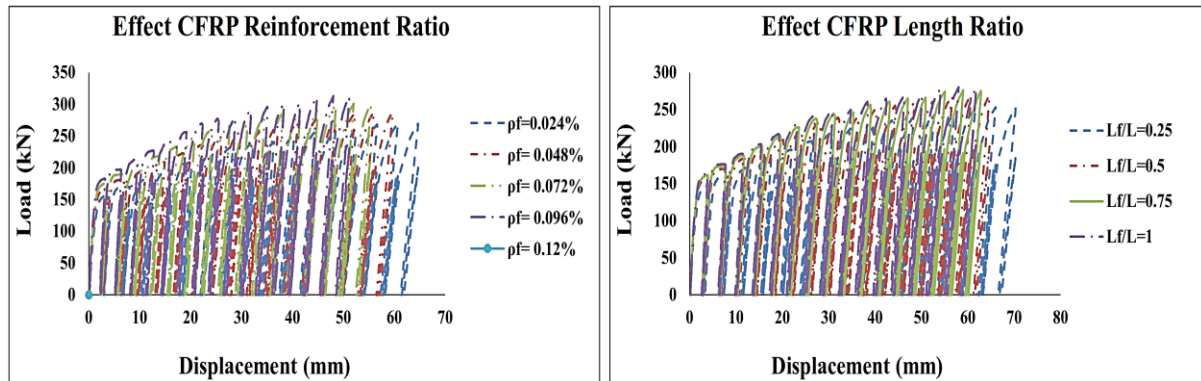


Figure 16. Influence of CFRP strengthening parameters on the load–displacement loops of the tested beams.

5.3. Influence of Loading Conditions

Figure 17 illustrates the influence of loading frequency on repeated-loading behavior. In the investigated quasi-static frequency range (0.1–0.3 Hz), the variation in load-bearing capacity and displacement behavior was minimal, suggesting no sensitivity to loading frequency. This shows that inertia and rate sensitivity are insignificant in the selected frequency range. As a result, the structure's response depends mainly on the beam geometry and the CFRP strengthening parameters.

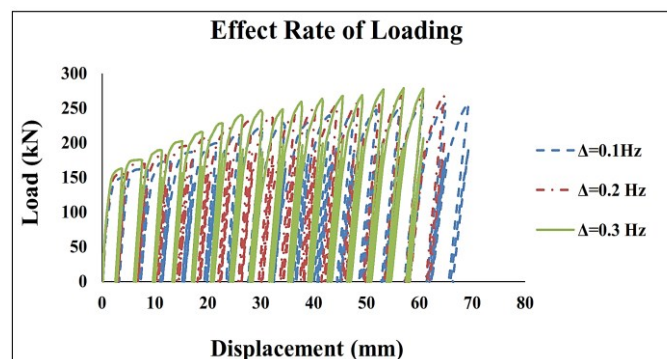


Figure 17. Influence of loading frequency on the load-displacement loops of the tested beams.

5.4. Integrated Discussion of Repeated Loading Behavior and Moment Redistribution

Both the experimental and numerical results suggest that the repeated-loading response of continuously reinforced concrete beams strengthened with CFRP is governed by the interaction among the following factors: stiffness development, damage accumulation, deformation capacity, and rotation compatibility. CFRP strengthening of RC beams typically enhances crack development, delays stiffness reduction, and reduces damage localization, leading to more stable behavior under repeated loads.

The moment redistribution process depends significantly on the relationship between the stiffness and rotation capacity. High values of parameters that increase stiffness, such as high reinforcement ratios, high steel yield strength, and high CFRP ratios, improve load-carrying capacity but reduce deformation capacity and redistribution efficiency. Parameters that increase flexibility and stress uniformity, such as span-to-depth ratio and CFRP layout, improve deformation capacity and redistribution.

The results show that repeated-loading performance and moment redistribution cannot be assessed solely by strength enhancement. Effective strengthening strategies should balance load capacity, crack control, deformation capacity, and redistribution capability to ensure stable performance under repeated loading conditions.

6. SIMPLIFIED LOAD CAPACITY PREDICTION

6.1. Development of the Predictive Model

To provide a practical tool for preliminary estimation of the load capacity of CFRP-strengthened continuous RC beams, a simplified prediction equation was developed based on the validated FE database. The database consisted of 18 numerical cases, including the reference model and 17 parametric variations covering the ranges listed in Table 9. The investigated variables were the negative reinforcement ratio (ρ_n), concrete compressive strength (F_{cu}), span-to-depth ratio (L/d), CFRP reinforcement ratio (ρ_f), CFRP length ratio (L_f/L), loading frequency (Δ), and steel yield strength (F_y). A nonlinear regression analysis based on a multiplicative power-law formulation was performed, resulting in Eq. (3).

$$P_{equ} = 114.27 \rho_n^{0.425} F_{cu}^{0.159} \left(\frac{L}{d}\right)^{-0.977} \rho_f^{0.111} \left(\frac{L_f}{L}\right)^{0.072} \Delta^{0.067} F_y^{0.997} \quad (3)$$

The obtained exponents indicate that steel yield strength and span-to-depth ratio have the greatest influence on load capacity within the investigated parameter ranges, whereas the remaining parameters contribute to a lesser extent. The proposed equation is intended as a simplified engineering estimation tool and should therefore be applied only within the parameter ranges considered in the present study.

6.2. Model Validation

The proposed equation was evaluated first against the FE database used for its development. As shown in Table 9, the model showed excellent agreement, with a mean predicted-to-FE ratio of 0.999 and a coefficient of variation (COV) of 0.8%.

To assess the model's applicability beyond the calibration dataset, we performed independent validation using experimental results reported by El-Ghandour et al. [16] and Ashour et al. [15]. The validation results are summarized in Tables 10 and 11. For the El-Ghandour et al. [16] dataset, the predicted-to-experimental ratios ranged from 0.953 to 1.036, with a mean of 0.988 and a COV of 3.2%. For the Ashour et al. [15] dataset, the ratios ranged from 0.91 to 1.05, with a mean value of 0.95 and a COV of 5.0%. As illustrated in Figure 18, most predictions were located close to the line of equality, indicating satisfactory agreement with the experimental observations.

This proposed formulation can be used for continuous CFRP-strengthened RC beams subjected to repeated loads, where gravity is the main load component, within the parameter range tested in this study. The applicability range includes beams with similar continuous beam configurations, flexural strengthening using CFRP sheets, reinforcement ratios, concrete compressive strengths, CFRP reinforcement ratios, CFRP length ratios, steel grades, and

loading frequencies represented by the developed FE database and validation datasets. Therefore, the equation is intended as a simplified engineering prediction tool for preliminary load-capacity estimation rather than a universal design equation for all CFRP-strengthened RC beams.

Table 9. Validation of the proposed model against FE results

Case	Varied Parameter	Value	P _{FE} (kN)	P _{Eq} (kN)	P _{Eq} / P _{FE}
Ref	—	—	270	267.8	0.992
1	ρ_n	0.66%	312	312.2	1.001
2	ρ_n	1.20%	403	402.5	0.999
3	F _{cu}	40 MPa	282	280.3	0.994
4	F _{cu}	50 MPa	290	290.5	1.002
5	L/d	12	331	333.1	1.006
6	L/d	18	220	224.1	1.019
7	ρ_f	0.048%	288	289.2	1.004
8	ρ_f	0.072%	301	302.6	1.005
9	ρ_f	0.096%	313	312.4	0.998
10	ρ_f	0.12%	322	320.3	0.995
11	L _f /L	0.25	253	254.7	1.007
12	L _f /L	0.75	276	275.7	0.999
13	L _f /L	1.00	280	281.5	1.005
14	Δ	0.1 Hz	258	255.5	0.990
15	Δ	0.3 Hz	279	275.2	0.986
16	F _y	350 MPa	166	164.6	0.992
17	F _y	650 MPa	308	305.3	0.991
Mean	—	—	—	—	0.999
COV	—	—	—	—	0.8%

Table 10. Validation of the proposed model against experimental results reported by El-Ghandour et al. (2004).

Beam	P _{exp} (kN)	P _{eq} (kN)	P _{eq} / P _{exp}
B1	170	162	0.953
B2	235	228	0.970
B3	255	262	1.027
B4	195	188	0.964
B5	250	244	0.976
B6	280	290	1.036
Mean	—	—	0.988
COV	—	—	3.2%

Table 11. Validation of the proposed model against experimental results reported by Ashour et al. (2004).

Beam	P _{exp} (kN)	P _{eq} (kN)	P _{eq} / P _{exp}
H1	138.0	145	1.05
H3	172.9	160	0.93
S4	170.5	155	0.91
E2	178.6	170	0.95
E3	207.0	190	0.92
E4	231.4	215	0.93
E5	174.6	165	0.95
Mean	—	—	0.95
COV	—	—	5.0%

Despite the successful results achieved in this validation process, a few limitations must be considered when using this equation. The formulation was based on a one-factor-at-a-time parametric database, so it does not account for possible interactions between parameters. This equation also does not account for CFRP debonding, bond slip, different failure modes, or fully reversed cyclic loading.

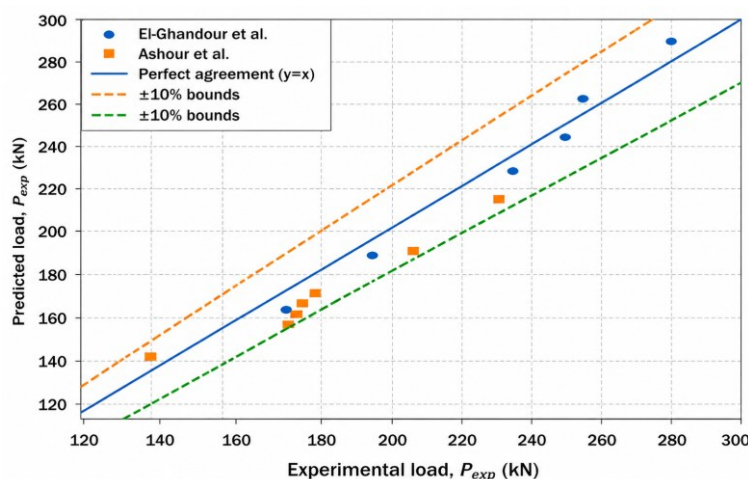


Figure 18. Comparison between predicted and experimental loads for the independent datasets reported by El-Ghandour et al. [16] and Ashour et al. [15]

Overall, the proposed formulation provides a simple and physically interpretable method for estimating the load capacity of CFRP-strengthened continuous RC beams subjected to gravity-dominated repeated loading. Within the identified applicability range, the equation provides reliable predictions; however, further validation using larger databases and broader parameter combinations is recommended to enhance its generalization and applicability.

7. CONCLUSIONS AND FUTURE WORK

This study investigated the repeated-loading response and moment-redistribution capacity of CFRP-strengthened continuous RC T-beams based on experimental tests and nonlinear FE modeling. Results showed that CFRP strengthening increases repeated-loading resistance by providing higher load capacity, better energy absorption, delayed stiffness loss, and more uniformly distributed crack and deformation patterns. The strengthening configuration strongly affects the structure's performance under repeated loading, with the hybrid longitudinal-diagonal CFRP scheme showing the most stable response.

Furthermore, the CFRP strengthening delays yielding of reinforcing bars and reduces strain concentrations in the negative moment zone. However, the influence of CFRP strengthening on moment redistribution is determined by the balance between stiffness improvement and rotational ability of the section. Although CFRP improves crack and damage distribution, excessive stiffening reduces moment-redistribution efficiency, implying that increased structural strength does not always lead to more efficient redistribution.

The parametric study revealed a trade-off between strength and ductility. The reinforcement ratio and yield strength increase the section's load capacity but reduce its deformation capacity, whereas a higher span-to-depth ratio provides ductility and moment redistribution at the cost of reduced load capacity. Lengthening the CFRP improves stress distribution and repeated-loading behavior; however, an excessive CFRP ratio decreases ductility because of increased stiffness. Higher concrete compressive strength increases the

compressive stability, whereas loading frequency has little effect within the analyzed quasi-static range.

The validated FE model reproduced the experimental response with satisfactory accuracy and was used to generate a numerical database for developing a simplified load-capacity prediction model. Overall, the results indicate that effective strengthening design should balance strength enhancement, deformation capacity, rotational compatibility, and moment redistribution rather than maximizing stiffness alone.

Future research should extend the proposed framework to fully reversed cyclic and seismic loading conditions while incorporating bond-slip effects, CFRP debonding behavior, and broader experimental validation covering a wider range of beam geometries and strengthening configurations.

REFERENCES

- [1] Aydogan MS, Aydemir C, Arslan G. (2024) Moment redistribution in continuous reinforced concrete beams: Experimental study and analytical investigation. *Structures*, 63: 106440. <https://doi.org/10.1016/j.istruc.2024.106440>
- [2] Li L, Li B, Zheng W. (2024) Moment Redistribution in UHPC Continuous Beams Reinforced with High-Strength Steel Bars: Numerical Investigation and Prediction Model. *Buildings*, 14(1):149. <https://doi.org/10.3390/buildings14010149>
- [3] Qureshi HJ, Althoey MA, Alqahtani A, Hadi MA, Elhaddad MM, Mohamed MA, et al. (2022) Investigation of CFRP Reinforcement Ratio on the Flexural Capacity and Failure Mode of Plain Concrete Prisms. *Materials*, 15(20):7248. <https://doi.org/10.3390/ma15207248>
- [4] Kim HG, Lee YJ, Kim KH. (2022) Cyclic flexural performance of RC beams with small circular openings in plastic hinge region. *Construction and Building Materials*, 321:126339. <https://doi.org/10.1016/j.conbuildmat.2022.126339>
- [5] Jahami A, Issa CA. (2024) An Updated Review on the Effect of CFRP on Flexural Performance of Reinforced Concrete Beams. *International Journal of Concrete Structures and Materials*, 18(1):14. <https://doi.org/10.1186/s40069-023-00651-y>
- [6] Suliman AKE, Jia Y, Mohammed AAA. (2021) Experimental evaluation of factors affecting the behavior of reinforced concrete beams strengthened by NSM CFRP strips. *Structures*, 32:632–640. <https://doi.org/10.1016/j.istruc.2021.03.058>
- [7] Shaikh ZK, Sohoni VS. (2025) Efficacy of carbon fibre reinforced polymer (CFRP) laminates in retrofitting the damaged moment resisting RC frames. *Discover Civil Engineering*, 2(1):166. <https://doi.org/10.1007/s44290-025-00326-5>
- [8] Alsuhaibani E. (2024) Optimization of Carbon Fiber-Reinforced Polymer (CFRP) Configuration for Enhanced Flexural Performance in Strengthened Concrete Beams. *Buildings*, 14(12):3953. <https://doi.org/10.3390/buildings14123953>
- [9] Lou T, Lopes SMR, Lopes AV. (2014) Factors affecting moment redistribution at ultimate in continuous beams prestressed with external CFRP tendons. *Composites Part B: Engineering*, 66:136–146. <https://doi.org/10.1016/j.compositesb.2014.05.007>
- [10] Kalyani G, Pannirselvam N. (2023) Experimental and numerical investigations on RC beams flexurally strengthened utilizing hybrid FRP sheets. *Results in Engineering*, 19:101337. <https://doi.org/10.1016/j.rineng.2023.101337>
- [11] Yagmur E. (2026) Parametric study on the behavior of CFRP-strengthened reinforced concrete deep beams with cut circular web openings in shear spans. *Scientific Reports*, 16: 9414. <https://doi.org/10.1038/s41598-026-40071-y>
- [12] Li R, Aboutaha RS. (2025) Analytical Investigation of CFRP- and Steel Plate-Strengthened RC Beams with Partially Unbonded Reinforcement. *Buildings*, 15(20):3665. <https://doi.org/10.3390/buildings15203665>

- [13] Szczepański K, Kotynia R. (2024) Bending moments redistribution in two-span reinforced concrete beams reinforced with FRP bars based on collected data research. *Structural Concrete*, 25(2):1076–1091. <https://doi.org/10.1002/suco.202300329>
- [14] Akbarzadeh H, Maghsoudi AA. (2010) Experimental and analytical investigation of reinforced high strength concrete continuous beams strengthened with fiber-reinforced polymer. *Materials and Design*, 31(3):1130–1147. <https://doi.org/10.1016/j.matdes.2009.09.041>
- [15] Ashour AF, El-Refaie SA, Garrity SW. (2004) Flexural strengthening of RC continuous beams using CFRP laminates. *Cement and Concrete Composites*, 26(7):765–775. <https://doi.org/10.1016/j.cemconcomp.2003.07.002>
- [16] El-Ghandour A, Aly H, Nasr E, Abdelrahman A. (2004) Moment redistribution in RC continuous beams strengthened in flexure by CFRP laminates. In: *FRP Composites in Civil Engineering (CICE 2004)*:227–235. <https://doi.org/10.1201/9780203970850.ch22>
- [17] Aiello MA, Valente L, Rizzo A. (2007) Moment redistribution in continuous reinforced concrete beams strengthened with carbon-fiber-reinforced polymer laminates. *Mechanics of Composite Materials*, 43(5):453–466. <https://doi.org/10.1007/s11029-007-0043-x>
- [18] El-Saikaly G, Chaallal O. (2015) Cyclic performance of reinforced concrete T-beams strengthened in shear with fiber-reinforced polymer composites: Sheets versus laminates. *Journal of Reinforced Plastics and Composites*, 34(13):1040–1058. <https://doi.org/10.1177/0731684415587347>
- [19] Tanarslan HM, Altin S. (2010) Behavior of RC T-section beams strengthened with CFRP strips, subjected to cyclic load. *Materials and Structures*, 43(4):529–542. <https://doi.org/10.1617/s11527-009-9509-8>
- [20] Fathuldeen SW, Qissab MA. (2019) Behavior of RC Beams Strengthened with NSM CFRP Strips under Flexural Repeated Loading. *Structural Engineering and Mechanics*, 70(1):67–80. <https://doi.org/10.12989/sem.2019.70.1.067>
- [21] Wang H, Li C, Song S, Wang Y, Meng Q, Li F. (2023) Flexural Performance of Cracked Reinforced Concrete Beams Strengthened with Prestressed CFRP Sheets under Repeated Loads. *Buildings*, 13(8):2115. <https://doi.org/10.3390/buildings13082115>
- [22] ECP 203-2020. (2020) Egyptian Code for Design and Construction of Reinforced Concrete Structures. Housing and Building National Research Center (HBRC), Cairo, Egypt.
- [23] BS EN 12390-3:2019. (2019) Testing Hardened Concrete-Part 3: Compressive Strength of Test Specimens. British Standards Institution.
- [24] ASTM A370-22. (2022) Standard Test Methods and Definitions for Mechanical Testing of Steel Products. ASTM International.
- [25] García JM, Bonett RL, Schultz AE, Carrillo J, Ledezma C. (2019) Flexural behavior of ungrouted post-tensioned concrete masonry beams with unbonded bars. *Construction and Building Materials*, 203:210–221. <https://doi.org/10.1016/j.conbuildmat.2018.12.101>

Machine-Learning Prediction of Marshall Mix-Design Properties and Mechanical Characterization of DCR-Treated Asphaltic Concrete

NUR ATIKAH KAMARUZAMAN¹, WAN NUR AIFA WAN AZAHAR^{1*},
ALTAMASHUDDIN KHAN NADIMALLA^{1,5}, NORHIDAYU KASIM¹,
SITI ALIYYAH MASJUKI¹, ASSAYIDATUL LAILA NOR HAIRIN²,
NORFARAH NADIA ISMAIL³, MASTURA BUJANG⁴

¹ Department of Civil Engineering, Kuliyah of Engineering, International Islamic University of Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

² Department of Manufacturing & Materials Engineering, Kuliyah of Engineering, International Islamic University of Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

³ Faculty of Civil Engineering, Universiti Teknologi MARA, Sarawak Branch, Samarahan 2 Campus, 94300 Kota Samarahan, Sarawak, Malaysia

⁴ Centre of Research for Innovation and Sustainable Development (CRISD), University of Technology Sarawak, 96000 Sibul, Sarawak, Malaysia

⁵ Department of Civil Engineering, Bearys Institute of Technology, Mangalore-574199, India

* Corresponding author: aifa@iium.edu.my

(Received: 19 June 2026; Accepted: 1 August 2026; Published online: 11 September 2026)

ABSTRACT: Marshall mix design remains the dominant empirical framework for proportioning hot-mix asphalt, and coconut-derived by-products such as Desiccated Coconut Residue (DCR) have attracted interest as sustainable bitumen modifiers. Determining the optimum bitumen content (OBC) and characterizing mechanical performance is laboratory-intensive, which motivates data-driven prediction. This study evaluated whether regression models can predict six Marshall responses for DCR-modified asphalt concrete and characterized the mechanical performance of the modified mixtures at their respective OBC. Thirty averaged Marshall records covering bitumen contents of 4–6 %, two gradations (AC14, AC20), and three treatments (Control, untreated DCR, treated DCR) were modeled using eight algorithms and a quadratic response-surface baseline. Models were assessed by pooled out-of-fold prediction under random K-fold, leave-one-mixture-out, and leave-one-bitumen-level-out validation. For the AC14 mixtures, OBC values were 5.30% (Control), 4.95% (untreated), and 5.10% (treated); resilient modulus, indirect tensile strength (ITS), tensile strength ratio (TSR), and dynamic creep were measured for each mixture prepared at its own OBC. Pooled out-of-fold R^2 reached 0.89 for voids filled with bitumen, 0.88 for air voids, 0.83 for bulk density, 0.81 for voids in mineral aggregate, 0.53 for flow, and 0.49 for stability. The quadratic response surface matched or exceeded the machine-learning models on four of six responses. Under leave-one-mixture-out validation, R^2 became negative for most responses, whereas leave-one-bitumen-level-out validation retained R^2 of 0.47–0.91. The models are therefore interpolation tools within the tested design space rather than general predictors. The untreated mixture required the lowest binder content, and the treated mixture recorded the highest dry ITS (1,148.7 kPa) and TSR (50.3%), although all mixtures fell below common TSR acceptance limits despite the hydrated-lime filler.

ABSTRAK: Reka bentuk campuran Marshall, kekal sebagai kerangka empirik utama bagi menentukan kadar campuran asphalt panas, dan hasil sampingan berasaskan kelapa seperti Sisa Kelapa Kering (DCR) telah menarik minat pengubah suai bitumen yang mampan. Penentuan kandungan bitumen optimum (OBC) dan pencirian prestasi mekanikal memerlukan kerja makmal yang intensif, sekaligus mendorong penggunaan ramalan berasaskan data. Kajian ini menilai sama ada model regresi mampu meramal enam tindak balas Marshall bagi konkrit asphalt terubah suai DCR, serta mencirikan prestasi mekanikal campuran tersebut pada OBC masing-masing. Sebanyak 30 rekod Marshall purata merangkumi kandungan bitumen 4–6 %, dua gradasi (AC14, AC20) dan tiga rawatan (Kawalan, DCR tidak dirawat, dan DCR dirawat) dimodelkan menggunakan lapan algoritma serta model permukaan sambutan kuadratik sebagai asas perbandingan. Model dinilai melalui ramalan terkumpul luar-lipatan di bawah pengesahan K-lipatan rawak, tinggal-satu-campuran-keluar dan tinggal-satu-aras-bitumen-keluar. Bagi campuran AC14, nilai OBC ialah 5.30 % (Kawalan), 4.95 % (tidak dirawat) dan 5.10 % (dirawat). Modulus kenyal, kekuatan tegangan tak langsung (ITS), nisbah kekuatan tegangan (TSR) dan rayapan dinamik diukur dengan setiap campuran yang disediakan pada OBC tersendiri. R^2 terkumpul luar-lipatan mencapai 0.89 bagi lompong terisi bitumen, 0.88 bagi lompong udara, 0.83 bagi ketumpatan pukal, 0.81 bagi lompong dalam agregat mineral, 0.53 bagi aliran dan 0.49 bagi kestabilan. Model permukaan sambutan kuadratik menyamai atau mengatasi model pembelajaran mesin bagi empat daripada enam tindak balas. Di bawah pengesahan tinggal-satu-campuran-keluar, R^2 menjadi negatif bagi kebanyakan tindak balas, manakala pengesahan tinggal-satu-aras-bitumen-keluar mengekalkan R^2 0.47–0.91. Justeru model ini merupakan alat interpolasi dalam ruang reka bentuk yang diuji dan bukan peramal umum. Campuran tidak dirawat memerlukan kandungan pengikat terendah, manakala campuran dirawat mencatat ITS kering tertinggi (1,148.7 kPa) dan TSR tertinggi (50.3 %), walaupun semua campuran berada di bawah had penerimaan TSR lazim.

KEYWORDS: *Asphalt mixture; Marshall mix design; Machine learning; Response surface; Desiccated coconut residue; Indirect tensile strength; Cross-validation*

1. INTRODUCTION

Flexible pavements carry most of the world's paved road traffic, and their service life depends on the mechanical performance of the asphalt concrete (AC) wearing course [1, 2]. Three response variables are commonly used to characterize that performance: the resilient modulus (RM), which reflects the stiffness available to distribute traffic stresses; the indirect tensile strength (ITS), which indicates tensile capacity and moisture-damage susceptibility [5, 6]; and the resistance to permanent deformation under repeated loading, usually evaluated through dynamic creep testing [3, 4]. Each test is reliable but slow, and the Marshall mix design that precedes them adds further effort through its sweep of bitumen contents, replicate specimens, and volumetric calculations [9].

Agricultural and biomass residues have been examined as sustainable bitumen modifiers in Malaysia and elsewhere [33]. Coconut-derived materials are of particular interest because of their local availability; coconut shell and husk powder have been reported to increase binder stiffness and rutting resistance [31, 32]. Desiccated Coconut Residue (DCR), a flaky by-product of commercial coconut-milk production, has received less attention, and its effect on Marshall design properties and on the mechanical performance of the resulting mixtures has not been systematically reported.

Data-driven regression offers a complementary route to reducing laboratory effort. Asi et al. [10] compiled 721 hot-mix asphalt observations and reported R^2 of 0.835 and 0.845 for Marshall stability and flow using CatBoost, with Shapley values identifying aggregate proportion and asphalt content as the principal predictors. Awan et al. [11] applied Multi-

Expression Programming to 253 base-course and 343 wearing-course specimens and derived explicit equations for stability and flow with training R^2 above 0.90. Kumar et al. [8] compared conventional and hybrid pipelines for plastic-modified asphalt, where neural networks and bagged support-vector variants exceeded correlation coefficients of 0.95 on 195 records. Al-Ammari et al. [12] evaluated several algorithms on a 60-record Marshall dataset and found the neural-network model strongest overall.

The same approach has been extended to fiber- and waste-modified mixtures. Upadhyaya et al. [24] trained five algorithms on Marshall stability data for glass- and carbon-fiber hybrid asphalt, with Gaussian process regression performing best and fiber content and bitumen percentage jointly governing stability. Jamil et al. [14] applied Random Forest, Gradient Boosting, XGBoost, and a Bagging Regressor to plastic-modified asphalt, obtaining R of 0.95 for training and 0.84 for testing, with SHAP analysis identifying plastic size and bitumen content as dominant. Two findings recur across this literature: bitumen content is the most influential continuous predictor, and non-linear methods outperform linear baselines on volumetric and strength responses.

For the mechanical properties, the resilient modulus measured under repeated indirect tensile loading is a temperature-sensitive stiffness parameter used in mechanistic-empirical design [25]. Shafabakhsh and Tanakizadeh [26] reported marked reductions in resilient modulus between 25 °C and 40 °C, with reduction ratios depending on binder grade and gradation. Indirect tensile strength under the Lottman protocol is the standard moisture-damage screening test, and the wet-to-dry ITS ratio quantifies susceptibility [27, 28]; ratios below roughly 70–80% are conventionally treated as moisture-sensitive [5, 6]. Uniaxial dynamic creep provides cumulative permanent strain and creep stiffness modulus, both of which correlate with wheel-tracking rutting performance [4, 7]. Hussan et al. [3] showed that neural networks reach R^2 up to 0.99 on uniaxial creep data, outperforming non-linear regression on the same records. Reviews of polymer-modified binders report higher resilient modulus at high service temperatures and slower strain accumulation, though the magnitude depends strongly on dosage and base-binder grade [7, 29].

Validation protocol matters more than algorithm choice when the dataset is small [15, 22]. Single train/test splits, still common in pavement machine-learning studies, are sensitive to which records fall in the test partition and yield optimistic estimates that often fail to replicate. Repeated K-fold cross-validation gives a more honest distribution of performance [23], and hyperparameter search must be conducted inside the cross-validation loop to avoid selecting on the test data [15]. Interpretability is commonly addressed through permutation feature importance, which measures the drop in predictive score when a feature is shuffled [19, 22]; its interpretation is affected by correlated or structurally dependent predictors [23], and SHAP values [30] require larger samples than are available here.

Two gaps motivate the present work. First, the published machine-learning literature on asphalt mixtures is dominated by datasets of hundreds to thousands of records, whereas a typical laboratory program produces an order of magnitude fewer observations, and methodological guidance for that regime is sparse. Comparative rankings are also rarely accompanied by a simple physically motivated baseline, so it is seldom clear whether the additional model complexity is justified. Second, mechanical-performance results obtained at the optimum bitumen content are rarely reported alongside the mix-design dataset from which that operating point was derived, even though both inform pavement design [1].

This paper addresses both gaps for DCR-modified asphalt concrete. We model six Marshall responses as functions of bitumen content, treatment condition, and gradation using

eight regression algorithms and a quadratic response-surface baseline. Performance is reported as pooled out-of-fold prediction under three validation designs: random K-fold, leave-one-mixture-combination-out, and leave-one-bitumen-level-out. The optimum bitumen content is determined separately for each AC14 mixture, and the resilient modulus, indirect tensile strength, tensile strength ratio, and dynamic-creep response are measured for each mixture at its optimum, so model behavior and laboratory evidence appear in one document. The objective is to establish what such models can and cannot support for practical mix-design decisions.

2. MATERIALS AND EXPERIMENTAL PROGRAM

2.1. Materials

The additive examined is Desiccated Coconut Residue (DCR), a flaky by-product of commercial coconut-milk production, obtained from Fizah Enterprise, Rawang, Selangor, Malaysia. Two forms were prepared. The untreated form (UT) was oven-dried at 60 °C for 24 hours. The treated form (TR) was oven-dried under the same conditions and then ground for one minute in a high-speed grinder operating at 25,000 rev/min, so that the treated material differs from the untreated material only by the additional grinding step. Particle-size distribution, moisture content, and chemical composition of the DCR were not determined in this program; the study characterizes mixture-level performance rather than additive microstructure, and this restricts mechanistic interpretation of the observed differences.

The base binder was a PEN 60/70 penetration-grade bitumen supplied by K2 Bitumen Sdn. Bhd., selected for the Malaysian climatic range. Measured properties were a penetration of 62.4 dmm at 25 °C (ASTM D5/D5M [34], within the 60–70 dmm range), a softening point of 49 °C (ASTM D36 [35], within the 49–56 °C requirement) and a rotational viscosity of 0.44 Pa·s at 135 °C (ASTM D4402/D4402M [36], below the 3 Pa·s limit). DCR was added at 5 % by mass of the binder for both the UT and TR mixtures, in addition to the binder rather than as a replacement for part of it; the Control mixture used the same binder with no DCR. The 5% dosage was adopted from the preceding phase of the laboratory program, where it produced a workable, homogeneous blend without visible phase separation; the dosage was not varied in the present study. Modified binders were produced by heating 800 g of base binder to 130 °C until pourable, adding the DCR, and mixing for one hour at 160 °C and 1,500 rpm, following a procedure adapted from the biomass-modifier protocol of Mior Sani et al. [33].

Aggregates were supplied by Kajang Rock Quarry, Kajang, Selangor. Coarse fractions of 20, 14, 10, and 5 mm nominal size were blended with fine aggregate to satisfy the JKR gradation envelopes for AC14 (14 mm nominal maximum aggregate size) and AC20 (20 mm) [2]. Hydrated lime was used as the mineral filler throughout, in accordance with JKR requirements, because it improves binder–aggregate adhesion and reduces stripping potential [5, 6]. Aggregate characterization comprised sieve analysis (ASTM C136, AASHTO T27), specific gravity and water absorption (ASTM C127), aggregate crushing value (BS EN 1097-2), Los Angeles abrasion (ASTM C131-81), flakiness index (BS 812; MS 30) and elongation index (ASTM D4791-10), all conducted per the JKR Standard Specification for Road Works [2].

Three treatments were studied for each gradation: a Control mixture without additive, an untreated DCR mixture (UT), and a treated DCR mixture (TR). The full factorial of six gradation-by-treatment mixtures was prepared at five bitumen contents (4, 4.5, 5, 5.5, and 6% by mass of the mixture), with three Marshall replicates per bitumen content, yielding 90 individual specimens and 30 averaged records for volumetric and strength analysis. The optimum bitumen content was determined separately for each AC14 mixture by the Asphalt

Institute MS-2 averaging procedure, in which the bitumen contents satisfying the individual design criteria (peak density, peak stability, target air voids, target VFB, and flow) are averaged. Table 1 presents the resulting values. Each mixture was then prepared at its own optimum for the mechanical-performance tests. The untreated mixture required the lowest binder content of the three, which is of practical interest because binder is the most costly component of the mixture; a full cost assessment would also need to account for DCR processing and the additional mixing energy, which were not quantified here.

Table 1. Optimum bitumen content determined separately for each AC14 mixture by the Asphalt Institute MS-2 averaging procedure. Mechanical-performance tests were conducted with each mixture prepared at its own optimum.

AC14 mixture	OBC (% by mass of mixture)
Control	5.30
Untreated DCR (UT)	4.95
Treated DCR (TR)	5.10

Mechanical performance testing was confined to the AC14 gradation. The AC20 mixtures were included in the Marshall program, and their mix-design records are therefore retained in the modeling dataset, but resilient-modulus, indirect-tensile, and creep testing of AC20 specimens was not undertaken within the scope and resources of this phase of the laboratory program. AC20 mechanical characterization is planned as a follow-up study. Reporting AC20 mechanical results here would have required a separate optimum-bitumen-content determination that was not performed, and the mechanical comparison is therefore restricted to AC14 so that every mixture compared was tested at its own design point.

2.2. Test methods

Marshall mix-design tests followed the Asphalt Institute MS-2 procedure [9]. For each specimen, 1,200 g of proportioned aggregate was heated to 110–130 °C, mixed with the heated binder until uniformly coated, placed in a preheated mould, and compacted with 75 blows per face. After cooling and extrusion, specimens were conditioned in a water bath at 60 °C for 30 minutes before stability and flow measurement. Air voids and VMA were computed from the maximum theoretical specific gravity and the bulk density obtained from saturated surface-dry weighings.

Resilient modulus was measured on a universal testing machine per ASTM D7369-20 [25], using an indirect tensile loading fixture and linear variable differential transformers for horizontal deformation. Repeated haversine pulses were applied along the diametral axis with the load level set to keep the response within the recoverable strain range. One specimen per treatment was tested, and that specimen was measured first at 25 °C and then at 40 °C; the two temperature results are therefore repeated measurements on the same specimen rather than independent replicates. The resilient modulus was taken as the ratio of applied stress to recoverable strain averaged over stable cycles.

Indirect tensile strength followed ASTM D6931 [38] with moisture conditioning per AASHTO T283 [27]. Specimens were compacted to a target air-void level of 7 ± 1 %. The conditioned subset was vacuum saturated to 70–80 % degree of saturation, held in a water bath at 60 °C for 24 hours, and then brought to 25 °C for one hour before testing; no freeze–thaw cycle was applied. Dry and conditioned specimens were loaded diametrically at 50 mm/min to

failure, and the tensile strength ratio was computed as the ratio of the mean conditioned to the mean dry strength. Two replicates were tested in each condition.

Dynamic creep was performed as a uniaxial repeated-load cyclic compression test per EN 12697-25 [37] on cylindrical specimens of nominal length 70 mm and diameter 100 mm, with 3,600 haversine cycles at 40 °C, the temperature at which rutting is most critical [3, 4]. Two specimens per treatment were tested.

Table 2. Variables used in the study and their role in the modeling and experimental analyses.

Variable	Symbol	Unit	Role	Notes
Bitumen content	BC	% by mass of mix	Model input	5 levels: 4.0–6.0 %
Treatment	—	categorical	Model input	Control (reference)
Gradation	—	categorical	Model input	AC14 (reference) /
Bulk density	ρ	g/cm ³	Target	Marshall response
Marshall stability	S	kN	Target	Marshall response
Marshall flow	F	mm	Target	Marshall response
Air voids	Va	%	Target	Marshall response
Voids in mineral	VMA	%	Target	Marshall response
Voids filled with	VFB	%	Target	Marshall response
Resilient modulus	MR	MPa	Reported only	AC14, 3 specimens,
Indirect tensile	ITS	kPa	Reported only	AC14, 12
Tensile strength	TSR	%	Reported only	AC14, derived per
Permanent strain @	$\epsilon(3600)$	%	Reported only	AC14, 4 specimens

Table 3. Composition of the experimental datasets and their role in the analysis. The mechanical datasets are reported descriptively; their sample sizes do not support cross-validated regression.

Dataset	Records	Inputs	Targets	Use
Marshall (bitumen	30 averaged (from	bitumen content,	6 Marshall	Regression
Resilient modulus	3 specimens $\times$ 2	treatment,	MR	Reported
Indirect tensile	12 (3 mixtures $\times$ 2	treatment, condition	ITS, TSR	Reported
Dynamic creep	4 accepted	treatment	$\epsilon(1200)$, $\epsilon(3600)$,	Reported

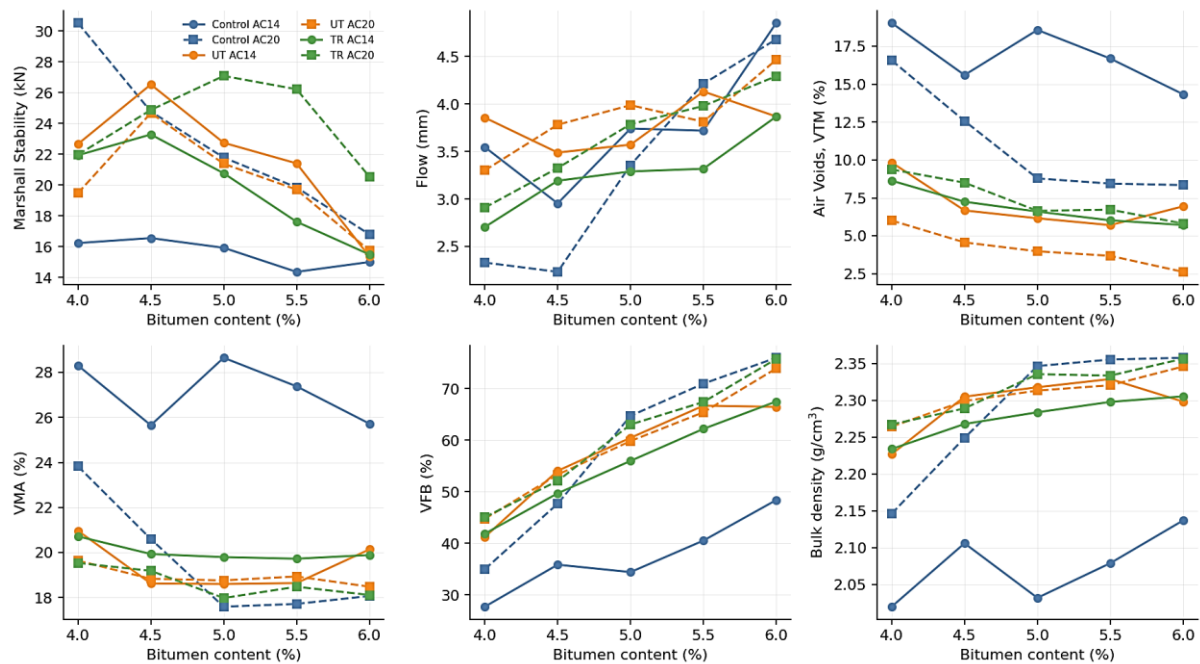


Figure 1. Marshall mix-design responses against bitumen content for the six mixtures (Control, untreated DCR, treated DCR; gradations AC14 and AC20). Each curve traces five bitumen-content levels at the averaged-replicate level. The Control AC14 mixture shows the highest VMA and the lowest density across the bitumen range; both DCR mixtures shift the density curve upwards and the air-voids curve downwards.

3. MODELING METHODOLOGY

The modeling dataset comprises the 30 averaged Marshall records: six mixture combinations (two gradations $\times$ three treatments) at five bitumen contents. Each record is the mean of three replicate specimens, which is the level at which Marshall design curves are conventionally reported. Averaging reduces within-condition experimental scatter, so the models predict mean mixture-condition responses rather than specimen-level variability, and the reported accuracies are correspondingly more favorable than specimen-level modeling would give. The three inputs are bitumen content (continuous), treatment, and gradation. Both categorical inputs are encoded with a reference category dropped: Control for treatment and AC14 for gradation, which avoids exact collinearity with the fitted intercept. Each of the six responses is modeled as a separate regression problem.

Nine models were compared: ordinary least squares, which is unregularized; three regularized linear models, namely Ridge, Lasso, and Elastic Net [16, 17]; support vector regression with a radial basis kernel [18, 39]; three tree ensembles, namely Random Forest, Extra Trees and Gradient Boosting [19, 20, 21]; and a quadratic response-surface model consisting of second-degree polynomial features with ridge regularization. The response surface is included as a physically motivated baseline, because Marshall responses are expected to vary non-linearly with bitumen content, and it provides a reference against which the value of the machine-learning models can be judged. We encapsulate preprocessing that depends on the data distribution in a pipeline so fitting occurs only within training folds [23]. Hyperparameter grids and the selected values are given in the Appendix.

We used three validation designs and compute all metrics from pooled out-of-fold predictions rather than averaging fold-level scores. With only six records in a test fold, fold-

level R^2 is unstable because it depends on the target variance that happens to fall in that fold; pooling the out-of-fold predictions before computing the metric avoids this artifact and is more interpretable. The first design is repeated random 5-fold cross-validation, with five repeats using different random seeds. The second is leave-one-mixture-combination-out validation, in which all five records of one gradation–treatment combination are held out together, so the model must predict a mixture it has never seen. The third is leave-one-bitumen-level-out validation, in which all six records at one bitumen content are held out together, so the model must predict an unobserved binder level. Hyperparameter search is repeated inside the training partition of every outer fold in all three designs. The coefficient of determination, root-mean-square error, and mean absolute error are reported for each.

Random K-fold splitting assigns records from the same mixture curve to both the training and test partitions, so it measures interpolation within the observed factorial rather than generalization to new mixtures. The two grouped designs were added specifically to separate these two questions, and the contrast between them is central to the interpretation offered in Section 4.

4. RESULTS AND DISCUSSION

4.1. Dataset Characteristics

Descriptive statistics for the six responses across the 30 averaged records are given in Table 4. Bulk density ranges from 2.02 g/cm³ for the Control AC14 mixture at 4 % bitumen to 2.358 g/cm³ for the AC20 TR mixture at 6 %. Marshall stability spans 14.4–30.5 kN, and flow spans 2.23–4.85 mm. Air voids vary widely, from 2.62% to 19.03%, and this spread reflects both bitumen-content and treatment effects. VMA spans 17.6–28.6 % and VFB 27.7–76.0 %. The correlation structure in Figure 2 shows near-perfect negative correlations between density and VMA and between density and air voids, which follow from the volumetric definitions rather than from independent observation. Bitumen content correlates positively with flow (+0.74) and VFB (+0.76) and negatively with stability (−0.51), consistent with the expected unimodal stability curve [9].

Table 4. Descriptive statistics for the input variable and the six target responses across the 30 averaged Marshall records.

Variable	n	Mean	SD	Min	Median	Max
Bitumen	30	5.00	0.72	4.00	5.00	6.00
Bulk density	30	2.261	0.097	2.020	2.298	2.358
Stability (kN)	30	20.71	4.20	14.36	21.07	30.52
Flow (mm)	30	3.62	0.62	2.23	3.73	4.85
Air voids (%)	30	8.88	4.53	2.62	7.10	19.03
VMA (%)	30	20.62	3.25	17.60	19.59	28.65
VFB (%)	30	54.94	13.50	27.71	55.03	76.00

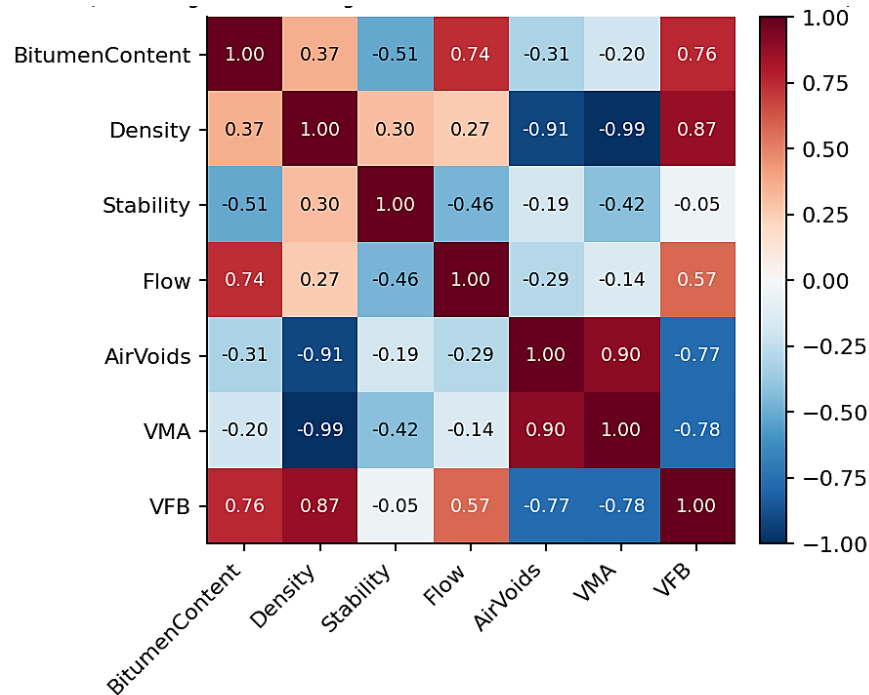


Figure 2. Pearson correlation matrix of the six Marshall design variables and bitumen content. The strong negative correlations between density and air voids and between density and VMA follow from the volumetric definitions rather than from independent measurement.

4.2. Predictive Performance and What It Supports

Table 5 reports pooled out-of-fold performance for the best model of each response under all three validation designs. Under random K-fold validation, the volumetric responses are predicted well: R^2 of 0.89 for VFB, 0.88 for air voids, 0.83 for density, and 0.81 for VMA. Flow reaches 0.53 and stability 0.49. These values are markedly higher than would be obtained by averaging fold-level R^2 , which, for the flow response, yields a near-zero value purely because individual six-record folds contain little flow variance. The pooled estimates are defensible, and the difference illustrates why the choice of metric matters at this sample size.

Table 5. Pooled out-of-fold performance of the best model for each Marshall response under three validation designs. RMSE and MAE correspond to the random K-fold design and are expressed in the units of each response. Results for all nine models are given in the Appendix.

Response	Best model	Random K-	RMSE	MAE	Leave-one-	Leave-one-
Bulk density (g/cm ³)	Response surface	0.826	0.039	0.029	-4.43	0.838
Stability (kN)	SVR (RBF)	0.487	2.896	2.227	-0.15	0.610
Flow (mm)	Response surface	0.528	0.415	0.326	0.00	0.467
Air voids (%)	Response surface	0.879	1.535	1.151	-1.23	0.882
VMA (%)	Response surface	0.812	1.359	1.026	-4.00	0.822
VFB (%)	SVR (RBF)	0.895	4.261	3.572	0.23	0.881

The most consequential result in Table 5 is the contrast between the two grouped designs. Under leave-one-bitumen-level-out validation, where the model must predict an unobserved binder content for mixtures it has otherwise seen, performance is retained across all six responses (R^2 of 0.47–0.88). Under leave-one-mixture-combination-out validation, where an entire gradation–treatment combination is withheld, R^2 becomes negative for four of the six responses, meaning the models perform worse than predicting the training mean. The models interpolate competently along the bitumen-content axis but carry essentially no ability to predict a mixture family they have not encountered. This is a direct consequence of the design: with three treatments and two gradations, withholding one combination removes the only evidence the model has about that category. The practical implication is specific. These models can reduce laboratory effort by predicting intermediate binder contents for a mixture whose Marshall behavior has been partially characterized, as in optimum-bitumen-content determination, but they cannot be used to screen a new additive or gradation without fresh testing.

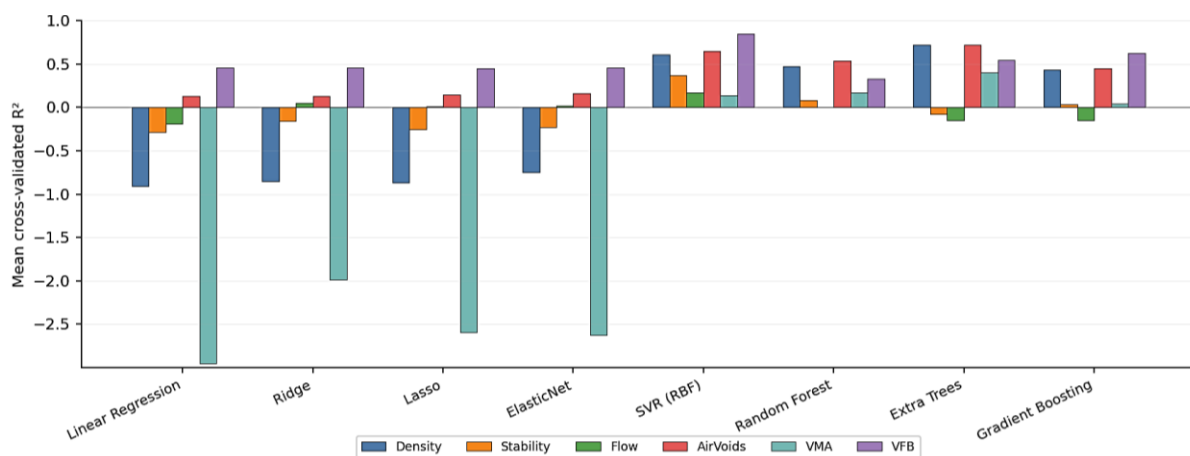


Figure 3. Cross-validated R^2 across the candidate algorithms for each Marshall response under random K-fold validation. Linear baselines trail the kernel, ensemble, and response-surface methods on all six responses.

The quadratic response surface is the best-performing model for density, flow, air voids, and VMA, and it is within 0.01 R^2 of the best model for VFB. Support vector regression retains an advantage only for stability and VFB. This is a useful negative result. The machine-learning algorithms do not deliver a material improvement over a second-degree polynomial in bitumen content combined with treatment and gradation indicators, and for a design space of this size the response surface is preferable for transparency, since its coefficients can be inspected and reported directly. Ordinary least squares, which lacks polynomial terms, performs substantially worse (R^2 of 0.18–0.77), confirming that the curvature rather than the algorithm family is what matters. The classical Marshall stability response peaks near the optimum and declines on either side [9], and any model without curvature in the bitumen content cannot reproduce that shape.

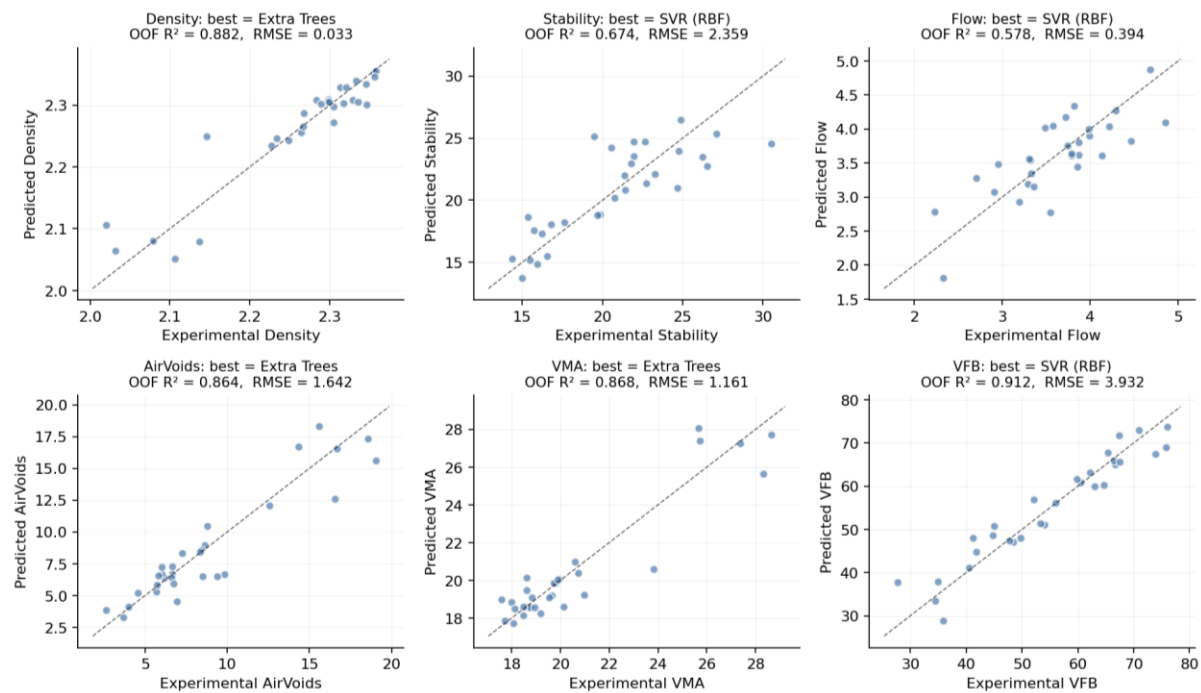


Figure 4. Out-of-fold predictions against experimental values for each Marshall response using the best model per response. Predictions are pooled across one 5-fold partition, giving one point per record.

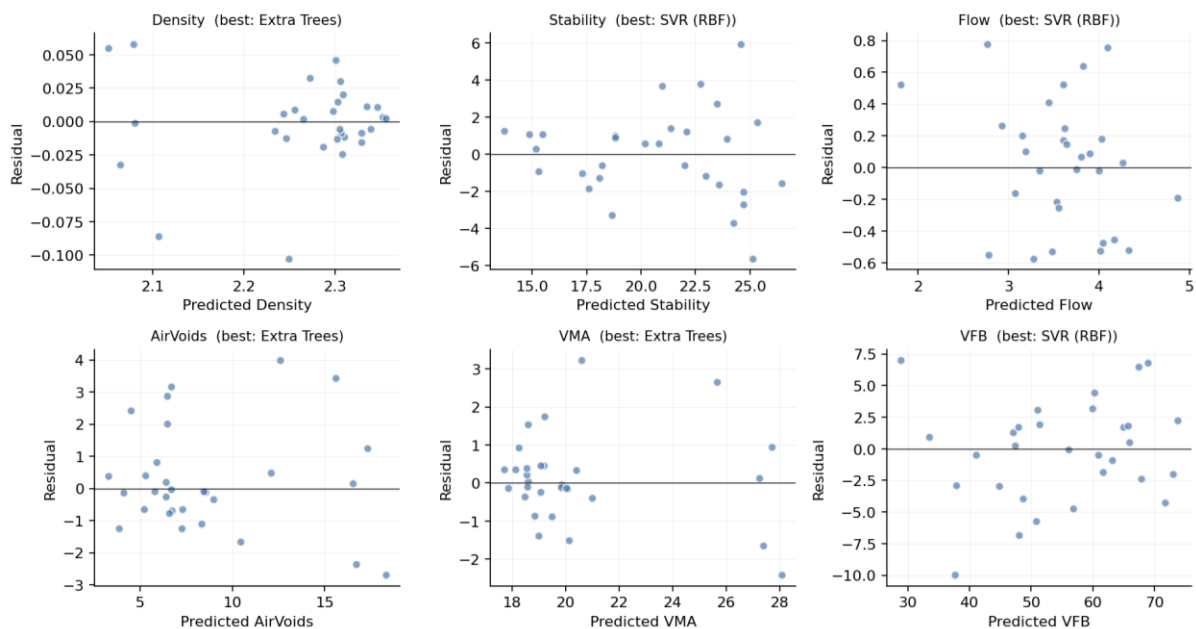


Figure 5. Residuals against predicted values for the best model of each response. Residuals center near zero for all responses; the wider scatter for stability and VFB at the extremes of their ranges reflects the greater experimental variability at the low-bitumen end of the design.

4.3. Feature Importance

Figure 6 shows permutation importance computed on the fitted models. Bitumen content dominates for stability and VFB, consistent with its established mechanistic role [10, 11]. For

density, air voids, and VMA, the treatment indicator separating the Control mixture from the two DCR mixtures dominates. This reflects a categorical separation in the experimental design rather than a dose-response effect: DCR was applied at a single dosage, so the model can only distinguish presence from absence, and the Control mixture is volumetrically distinct from the two DCR mixtures, which cluster together. Because treatment was encoded categorically and dosage did not vary, these results identify predictive separations within the studied design and cannot be read as evidence of a physical mechanism by which DCR alters density. Establishing a dose-response relationship would require multiple DCR dosages [8, 14]. The importance values are also computed from models fitted on the full dataset, which makes them descriptive rather than an unbiased estimate of predictive contribution.

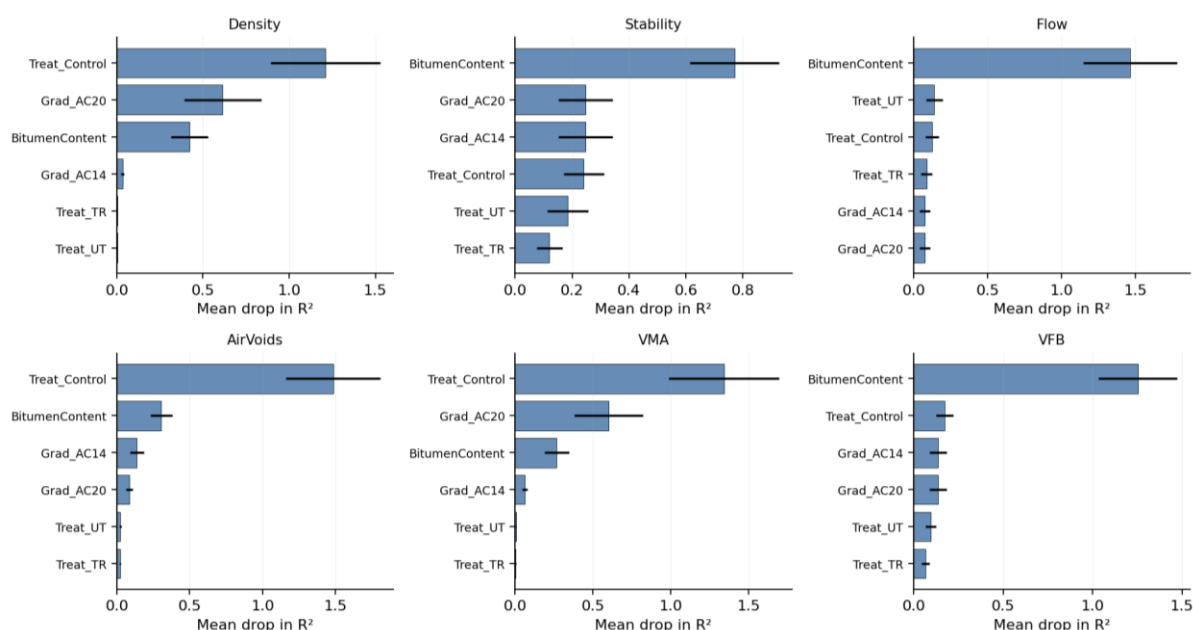


Figure 6. Permutation feature importance for each Marshall response. Bars give the mean reduction in R^2 when a feature is shuffled; error bars are the standard deviation across 50 repeats. Values are descriptive and should not be interpreted as evidence of a physical DCR mechanism.

4.4. Mechanical Performance at the Optimum Bitumen Content

The resilient modulus, indirect tensile strength, and creep results are reported as direct measurements without regression modeling or an associated R^2 . This is deliberate. Each property was measured only at the optimum bitumen content of each mixture, giving three resilient-modulus specimens, twelve indirect-tensile measurements, and four accepted creep specimens. Cross-validated R^2 estimated from samples of this size would be dominated by the chosen partition rather than any underlying response surface, and reporting such values would misrepresent the strength of the evidence. Because only two replicates are available per condition, the comparisons below are descriptive, and the differences reported are not supported by statistical tests.

Table 6 gives the resilient modulus at 25 °C and 40 °C. At 25 °C, the three mixtures fall within a narrow band of 11,815–12,646 MPa, with the untreated mixture highest and the Control 6.6 % below it. At 40 °C, the modulus falls by 62–74 %, which is consistent with the temperature sensitivity of the binder reported in the literature [26, 29], and the Control mixture retains the highest value. The treated mixture shows the largest proportional reduction. With one specimen per treatment and the two temperatures measured on that same specimen, these

values are point measurements, and the ordering between mixtures should not be regarded as established.

Table 6. Resilient modulus of the AC14 mixtures, each tested at its own optimum bitumen content. One specimen per treatment was measured at both temperatures, so the two columns are repeated measurements on the same specimen.

Treatment	MR at 25 °C (MPa)	MR at 40 °C (MPa)	Reduction (%)
Control	11.815	4.488	62.0
Untreated DCR (UT)	12,646	3,878	69.3
Treated DCR (TR)	12,039	3,147	73.9

Indirect tensile strength results are given in Table 7. The treated mixture recorded the highest dry strength (1,148.7 kPa) and the highest conditioned strength (577.4 kPa), both above the corresponding Control values (945.9 and 455.6 kPa). The Control and untreated mixtures recorded closely similar strengths in both conditions, which suggests that grinding rather than the presence of DCR is associated with the strength difference, although with two replicates per condition this observation is descriptive. Tensile strength ratios were 48.2% for the Control, 47.6% for the untreated mixture, and 50.3% for the treated mixture, all well below the 70–80% range commonly adopted as an acceptance criterion [5, 6, 27].

Table 7. Indirect tensile strength of the AC14 mixtures, each tested at its own optimum bitumen content. Means and standard deviations are computed across two replicate specimens per condition.

Treatment	ITS dry (kPa, mean $\pm$ SD)	ITS conditioned (kPa, mean $\pm$ SD)	TSR (%)
Control	945.85 $\pm$ 5.45	455.64 $\pm$ 44.72	48.2
Untreated DCR (UT)	955.54 $\pm$ 82.01	454.87 $\pm$ 34.87	47.6
Treated DCR (TR)	1,148.72 $\pm$ 29.66	577.38 $\pm$ 14.74	50.3

That all three mixtures fell below common acceptance limits despite the use of hydrated lime as filler is the most practically significant result of the experimental program, and it warrants comment. Hydrated lime is specified precisely because it promotes binder–aggregate adhesion and reduces stripping [5, 6], so its presence would ordinarily be expected to lift the tensile strength ratio. Three explanations fit the data. First, the conditioned specimens lost roughly half their dry strength, a reduction characteristic of adhesive failure at the binder–aggregate interface rather than cohesive failure within the binder, pointing to the interface as the controlling weakness. Second, the granite-derived aggregate used here is siliceous and therefore intrinsically hydrophilic, and lime alone may be insufficient to compensate for that in a wearing course subjected to a full 24-hour saturated conditioning cycle. Third, DCR is an organic residue with residual moisture affinity, and its presence at the interface may itself provide pathways for water ingress; the treated form, which was ground more finely, gave the highest ratio of the three, which is at least consistent with better dispersion reducing such pathways. Distinguishing these mechanisms would require interface-level characterization that was not part of this program. For practical purposes, the implication is that DCR-modified mixtures of this type should not be specified for moisture-exposed applications on the basis of lime filler alone; a liquid anti-stripping agent, a higher lime dosage, or a less hydrophilic

aggregate source would need to be evaluated, and drying the DCR to a lower residual moisture content before blending may also be worth testing.

Table 8 presents the dynamic-creep results for the four accepted AC14 specimens. The Control specimens accumulated a mean permanent strain of 2.57 % at 3,600 cycles against 2.15 % for the treated specimens, and the mean creep stiffness modulus was 79.4 MPa for the Control and 60.9 MPa for the treated mixture. The two indicators therefore do not order the mixtures consistently, and with two specimens per treatment, no conclusion about relative rutting resistance is supported. The individual strain histories in Figure 8 show the divergence directly: specimen TR14-S1 accumulated approximately 25,000 $\mu\epsilon$ by cycle 3,600 while TR14-S2 reached roughly 14,000 $\mu\epsilon$, a spread wider than the difference between the treatment means.

Table 8. Dynamic Creep Results and Untreated Specimen Exclusion

Specimen	Treatment	$\epsilon(1200)$ %	$\epsilon(3600)$ %	CSM (MPa)
C14-S1	Control	2.174	2.467	102.4
C14-S2	Control	2.131	2.663	56.4
TR14-S1	Treated DCR	2.303	2.969	45.0
TR14-S2	Treated DCR	0.930	1.321	76.7

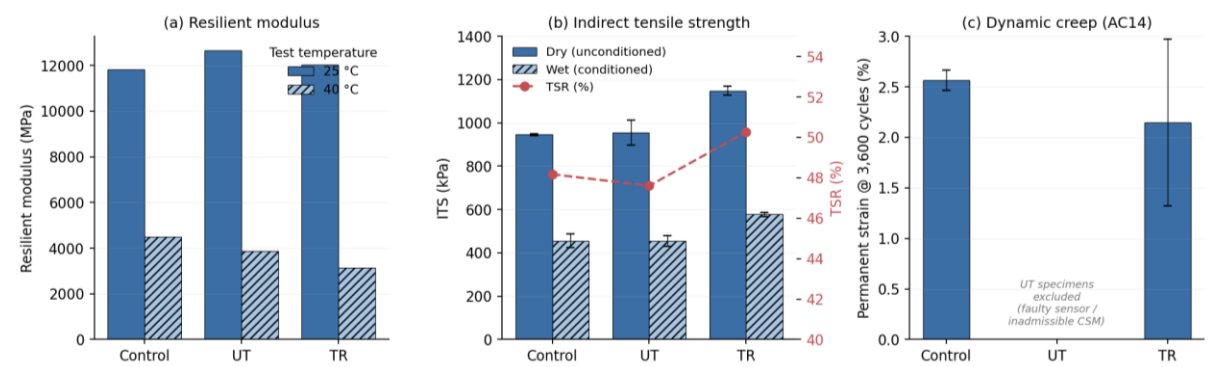


Figure 7. Mechanical performance of AC14 mixtures at optimum bitumen content: (a) resilient modulus at 25 °C and 40 °C; (b) dry and conditioned indirect tensile strength with tensile strength ratio; and (c) permanent strain at 3,600 cycles. Untreated specimens were excluded during data-quality screening (Table 8).

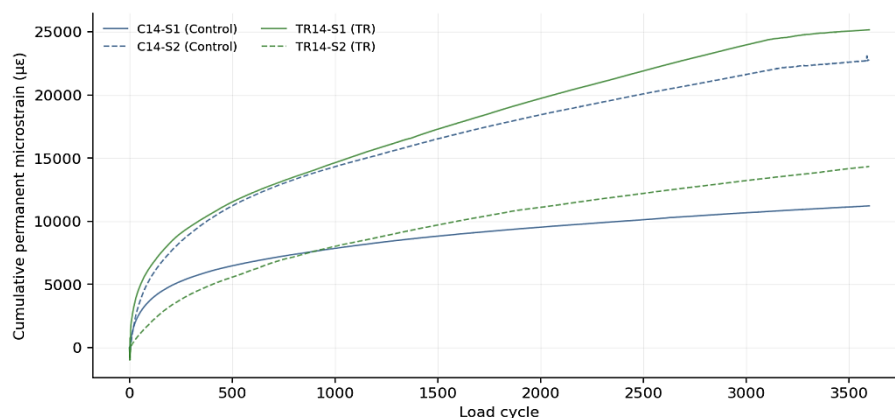


Figure 8. Cumulative permanent microstrain versus load cycles for accepted AC14 creep specimens. Control specimens showed typical primary–secondary creep, while treated specimens diverged, reaching approximately 25,000 and 14,000 $\mu\epsilon$ at 3,600 cycles.

4.5. Synthesis and Scope of the Findings

Taken together, the modeling and experimental results point in the same direction. The regression models reproduce the volumetric responses well because those responses vary smoothly and largely deterministically with binder content within a given mixture; the same smoothness is visible in the Marshall curves of Figure 1. Stability and flow are predicted less well because they depend on mechanical behavior that the three available inputs do not capture, and the mechanical tests show why: at a fixed binder content, the three mixtures differ in tensile strength and creep response in ways that bitumen content, treatment category, and gradation cannot express. The information needed to predict mechanical response is not present in mix-design variables alone.

The scope of what these models support is correspondingly narrow. They are interpolation tools for the tested design space. Within that space, they can reduce the number of trial bitumen contents required during optimum-content determination, which is a real, if modest, saving of laboratory effort. They cannot predict performance for an untested additive, dosage, or gradation, as the leave-one-mixture-out results show plainly. They also cannot substitute for the mechanical tests: no relationship established here would have predicted that all three mixtures fail the tensile strength ratio criterion, which is the finding with the clearest bearing on whether these mixtures could be used in practice. Extending the models beyond interpolation would require a dataset spanning several DCR dosages, multiple aggregate sources, and specimen-level rather than averaged records, together with mechanical testing at more than one binder content so that the mechanical responses themselves become modellable.

5. CONCLUSIONS

This study evaluated regression models for six Marshall mix-design responses of DCR-modified asphalt concrete and characterized the mechanical performance of the modified AC14 mixtures at their respective optimum bitumen contents. The following conclusions are drawn.

The optimum bitumen content differed between mixtures: 5.30% for the Control, 4.95% for the untreated DCR mixture, and 5.10% for the treated DCR mixture. The untreated mixture therefore required the least binder, which is relevant to mixture cost, although a complete assessment would need to include the cost of processing the additive.

Pooled out-of-fold coefficients of determination under random K-fold validation were 0.89 for voids filled with bitumen, 0.88 for air voids, 0.83 for bulk density, 0.81 for voids in mineral aggregate, 0.53 for flow, and 0.49 for stability. Volumetric responses are therefore predictable from bitumen content, treatment, and gradation, whereas the mechanically governed responses are only partially so.

A quadratic response surface matched or exceeded the machine-learning algorithms on four of the six responses. For a design space of this size, the additional complexity of the machine-learning models is not warranted, and the response surface is preferable because its coefficients can be inspected directly.

Validation design determines what the reported accuracy means. When an entire mixture combination was withheld, the coefficient of determination became negative for four of six responses, whereas withholding a bitumen level retained values of 0.47 to 0.88. The models are interpolation tools within the tested design space and cannot predict the behavior of an untested mixture family.

Among the mechanical properties, the treated DCR mixture recorded the highest dry and conditioned indirect tensile strength. All three mixtures returned tensile strength ratios between

47.6% and 50.3%, well below common acceptance limits, despite using hydrated lime as a mineral filler. On this evidence, DCR-modified mixtures of the type tested here require additional anti-stripping measures before they could be considered for moisture-exposed applications.

Further work should extend the experimental matrix across several DCR dosages, characterize the additive itself, repeat the untreated-mixture creep tests that were excluded here, increase replication so that differences in mechanical performance can be tested statistically, and conduct the mechanical tests at more than one binder content so that those responses become amenable to modeling in their own right.

DATA AVAILABILITY

The 30-record modeling dataset, the descriptive experimental results reported in Tables 6 to 8, the complete model-comparison results for all nine algorithms and three validation designs, the analysis code including preprocessing pipeline and hyperparameter grids, and the raw traces of the excluded creep specimens are available from the corresponding author on reasonable request. The analysis was performed in Python 3.12 with scikit-learn 1.8, NumPy 2.4, and pandas 3.0. A fixed random seed of 42 was used for all cross-validation partitioning and model fitting, so that reported results are reproducible.

ACKNOWLEDGEMENTS

The support provided by the Fundamental Research Grant Scheme (FRGS) from the Malaysian Ministry of Higher Education (MOHE), with research grant number FRGS/1/2023/TK02/UIAM/02/3 (Project ID: FRGS23-319-0928), is highly appreciated.

APPENDIX A. MODEL SPECIFICATIONS AND FULL RESULTS

Hyperparameter grids were searched by exhaustive grid search within the training partition of every outer fold. Ridge regularisation strength was searched over $\{0.01, 0.1, 1, 10, 100\}$; Lasso over $\{0.001, 0.01, 0.1, 1\}$; Elastic Net over the same alpha values with l1 ratios of $\{0.2, 0.5, 0.8\}$. Support vector regression used $C \in \{0.1, 1, 10, 100\}$, $\epsilon \in \{0.01, 0.1, 0.5\}$ and $\gamma \in \{\text{scale}, 0.1, 1.0\}$. Random Forest and Extra Trees used 200 estimators with maximum depth $\in \{\text{none}, 4\}$ and minimum samples per leaf $\in \{1, 2\}$. Gradient Boosting used 200 estimators with learning rate $\in \{0.05, 0.1\}$ and maximum depth $\in \{2, 3\}$. The response-surface model used second-degree polynomial features with a ridge penalty searched over $\{0.001, 0.01, 0.1, 1, 10\}$.

Table A1. Pooled out-of-fold coefficient of determination for representative models under the three validation designs. Complete results for all nine models, including root-mean-square and mean absolute errors and the hyperparameters selected in each fold, are available from the corresponding author.

Response	Model	Random K-fold R ²	Leave-one-mixture-out R ²	Leave-one-bitumen-level-out R ²
Density	OLS	0.513	-1.11	0.603
Density	SVR (RBF)	0.823	-0.76	0.754
Density	Extra Trees	0.744	-1.60	0.808
Density	Response surface	0.826	-4.43	0.838
Stability	OLS	0.176	-1.09	0.167
Stability	SVR (RBF)	0.487	-0.15	0.610
Stability	Extra Trees	0.153	-0.81	0.336
Stability	Response surface	0.297	-0.83	0.326
Flow	OLS	0.426	-0.04	0.397
Flow	SVR (RBF)	0.467	0.31	0.255
Flow	Extra Trees	0.293	0.17	0.496
Flow	Response surface	0.528	0.00	0.467
Air voids	OLS	0.773	0.13	0.788
Air voids	SVR (RBF)	0.819	0.09	0.739
Air voids	Extra Trees	0.840	-0.74	0.874
Air voids	Response surface	0.879	-1.23	0.882
VMA	OLS	0.460	-1.34	0.558
VMA	SVR (RBF)	0.764	-0.77	0.704
VMA	Extra Trees	0.780	-1.87	0.840
VMA	Response surface	0.812	-4.00	0.822
VFB	OLS	0.738	-0.03	0.791
VFB	SVR (RBF)	0.895	0.23	0.881
VFB	Extra Trees	0.537	-0.08	0.753
VFB	Response surface	0.885	-1.38	0.913

REFERENCES

- [1] Hunter RN (Ed.) (2015). The Shell Bitumen Handbook, 6th Ed. ICE Publishing, London.
- [2] JKR Malaysia (2008). Standard Specification for Road Works, Section 4: Flexible Pavement (JKR/SPJ/2008-S4). Jabatan Kerja Raya Malaysia, Kuala Lumpur.
- [3] Hussan S, Kamal MA, Hafeez I, Ahmad N (2019). Evaluation and modelling of permanent deformation behaviour of asphalt mixtures using dynamic creep test in uniaxial mode. *International Journal of Pavement Engineering*, 20(9):1026–1043. <https://doi.org/10.1080/10298436.2017.1380805>
- [4] Mashaan NS, Karim MR (2013). Evaluation of permanent deformation of CRM-reinforced SMA and its correlation with dynamic stiffness and dynamic creep. *The Scientific World Journal*, 2013:981637. <https://doi.org/10.1155/2013/981637>
- [5] Solaimanian M, Harvey J, Tahmoressi M, Tandon V (2003). Test methods to predict moisture sensitivity of hot-mix asphalt pavements. In *Moisture Sensitivity of Asphalt Pavements: A National Seminar*, pp. 77–110. Transportation Research Board, Washington, DC.
- [6] Do TC (2024). Engineering properties-based parameters used for moisture damage evaluation of asphalt mixtures: A review. *Canadian Journal of Civil Engineering*, 51(4):390–398. <https://doi.org/10.1139/cjce-2023-0243>
- [7] Mashaan NS, Ali AH, Karim MR, Abdelaziz M (2014). A review on using crumb rubber in reinforcement of asphalt pavement. *The Scientific World Journal*, 2014:214612. <https://doi.org/10.1155/2014/214612>

-
- [8] Kumar B, Alyaseen A, Kumar N (2025). Utilizing conventional and state-of-the-art machine learning algorithms to predict Marshall stability of modified asphalt mixes incorporating PET, HDPE, and PVC plastic waste: Performance evaluation and mix optimization. *International Journal of Pavement Research and Technology*. <https://doi.org/10.1007/s42947-025-00591-8>
 - [9] Asphalt Institute (2014). *Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types (MS-2)*, 7th Ed. Asphalt Institute, Lexington, KY.
 - [10] Asi I, Alhadidi YI, Alhadidi TI (2024). Predicting Marshall stability and flow parameters in asphalt pavements using explainable machine-learning models. *Transportation Engineering*, 18:100282. <https://doi.org/10.1016/j.treng.2024.100282>
 - [11] Awan HH, Hussain A, Javed MF, Qiu Y, Alrowais R, Mohamed AM, Fathi D, Alzahrani AM (2022). Predicting Marshall flow and Marshall stability of asphalt pavements using Multi Expression Programming. *Buildings*, 12(3):314. <https://doi.org/10.3390/buildings12030314>
 - [12] Al-Ammari M, Dong R, Nasser M, Al-Maswari A (2025). Innovative machine learning approaches for predicting the asphalt content during Marshall design of asphalt mixtures. *Materials*, 18(7):1474. <https://doi.org/10.3390/ma18071474>
 - [13] Al Mamun A, Gazder U, Islam MK, Arifuzzaman M, Wahhab HA-A, Rahman MM (2025). Predicting indirect tensile strength of rejuvenated asphalt mixes using machine learning with high reclaimed asphalt pavement content. *Processes*, 13(5):1489. <https://doi.org/10.3390/pr13051489>
 - [14] Jamil MH, Jagirdar R, Kashem A, Ali MN, Deb D (2025). Modeling of Marshall stability of plastic-reinforced asphalt concrete using machine learning algorithms and SHAP. *Hybrid Advances*, 10:100483. <https://doi.org/10.1016/j.hybadv.2025.100483>
 - [15] Cawley GC, Talbot NLC (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11:2079–2107.
 - [16] Tibshirani R (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
 - [17] Zou H, Hastie T (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
 - [18] Cortes C, Vapnik V (1995). Support-vector networks. *Machine Learning*, 20(3):273–297. <https://doi.org/10.1007/BF00994018>
 - [19] Breiman L (2001). Random forests. *Machine Learning*, 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
 - [20] Friedman JH (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232. <https://doi.org/10.1214/aos/1013203451>
 - [21] Geurts P, Ernst D, Wehenkel L (2006). Extremely randomized trees. *Machine Learning*, 63(1):3–42. <https://doi.org/10.1007/s10994-006-6226-1>
 - [22] Hastie T, Tibshirani R, Friedman J (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd Ed. Springer-Verlag, New York.
 - [23] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
 - [24] Upadhyaya A, Thakur MS, Al Ansari MS, Malik MA, Alahmadi AA, Alwetaishi M, Alzaed AN (2022). Marshall stability prediction with glass and carbon fiber modified asphalt mix using machine learning techniques. *Materials*, 15(24):8944. <https://doi.org/10.3390/ma15248944>
 - [25] ASTM (2020). Standard test method for determining the resilient modulus of asphalt mixtures by indirect tension test (ASTM D7369-20). ASTM International, West Conshohocken, PA. <https://doi.org/10.1520/D7369-20>
-

- [26] Shafabakhsh G, Tanakizadeh A (2015). Investigation of loading features effects on resilient modulus of asphalt mixtures using Adaptive Neuro-Fuzzy Inference System. *Construction and Building Materials*, 76:256–263. <https://doi.org/10.1016/j.conbuildmat.2014.11.069>
- [27] Lottman RP (1982). Predicting moisture-induced damage to asphaltic concrete: Field evaluation. NCHRP Report 246. Transportation Research Board, Washington, DC.
- [28] AASHTO (2014). Standard method of test for resistance of compacted asphalt mixtures to moisture-induced damage (AASHTO T283-14). American Association of State Highway and Transportation Officials, Washington, DC.
- [29] Yang Q, Lin J, Wang X, Wang D, Xie N, Shi X (2024). A review of polymer-modified asphalt binder: Modification mechanisms and mechanical properties. *Cleaner Materials*, 12:100255. <https://doi.org/10.1016/j.clema.2024.100255>
- [30] Lundberg SM, Lee S-I (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30:4765–4774.
- [31] Yousif RA, Tayh SA, Jasim AF (2023). The effect of coconut powder on asphalt binder performance under laboratory conditions. *Journal of Engineering and Technological Sciences*, 55(5):577–586. <https://doi.org/10.5614/j.eng.technol.sci.2023.55.5.7>
- [32] Ting TL, Jaya RP, Hassan NA, Yaacob H, Jayanti DS (2015). A review of utilization of coconut shell and coconut fiber in road construction. *Jurnal Teknologi*, 76(14):121–125. <https://doi.org/10.11113/jt.v76.5851>
- [33] Mior Sani WNH, Allias NF, Yaacob H, Al-Saffar ZH, Hashim MH (2024). Utilisation of sawdust and charcoal ash as sustainable modified bitumen. *Smart and Green Materials*, 1(1):13–32. <https://doi.org/10.70028/sgm.v1i1.1>
- [34] ASTM (2019). Standard test method for penetration of bituminous materials (ASTM D5/D5M-20). ASTM International, West Conshohocken, PA. https://doi.org/10.1520/D0005_D0005M-20
- [35] ASTM (2014). Standard test method for softening point of bitumen (ring-and-ball apparatus) (ASTM D36/D36M-14). ASTM International, West Conshohocken, PA. https://doi.org/10.1520/D0036_D0036M-14
- [36] ASTM (2015). Standard test method for viscosity determination of asphalt at elevated temperatures using a rotational viscometer (ASTM D4402/D4402M-15). ASTM International, West Conshohocken, PA. https://doi.org/10.1520/D4402_D4402M-15
- [37] CEN (2016). Bituminous mixtures — Test methods — Part 25: Cyclic compression test (EN 12697-25:2016). European Committee for Standardization, Brussels.
- [38] ASTM (2019). Standard test method for indirect tensile (IDT) strength of bituminous mixtures (ASTM D6931-17). ASTM International, West Conshohocken, PA. <https://doi.org/10.1520/D6931-17>
- [39] Smola AJ, Schölkopf B (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>

Progressive Collapse Analysis of 2D RC Frames Under the Impact of Single Interior and Exterior Columns

TAHARA RAMADZAN MD KASSIM^{1*}, MUHAMMAD FAIZ MOHAMAD NOR¹,
TAKSIAH A. MAJID²

¹*Department of Civil Engineering, Kulliyyah of Engineering,
International Islamic University Malaysia, Kuala Lumpur, Malaysia*
²*School of Civil Engineering, Universiti Sains Malaysia, Penang, Malaysia*

**Corresponding author: tahara@iium.edu.my*

(Received: 16 July 2026; Accepted: 21 August 2026; Published online: 11 September 2026)

ABSTRACT: Structural elements are susceptible to both local and global failures triggered by progressive collapse, making it essential in structural vulnerability assessments. This research investigates and predicts the progressive collapse behavior of two-dimensional (2D) reinforced concrete frames designed for gravity loadings, considering the removal of columns in both interior and exterior locations. Linear static and nonlinear dynamic analyses are conducted to investigate and predict the development of alternate load paths (ALPs) and dynamic structural response under column-loss scenarios using SAP2000 software. Three cases are considered: Case 1 (no column removal), Case 2 (removal of one interior column), and Case 3 (removal of one exterior column). Linear static analysis reveals that the General Services Administration (GSA) yields slightly higher, more conservative axial forces than the British Standard (BS). Regardless of where the column was removed, ALP under progressive collapse significantly altered the structural response in Case 2 and Case 3. Nonlinear dynamic analysis revealed that Case 3 exhibited the highest displacement and interstorey drift ratio (IDR), particularly under Ground Motion (GM5). The findings also demonstrate that exterior column removal triggered the highest drift demands, indicating critical structural vulnerability to progressive collapse.

ABSTRAK: Elemen struktur terdedah kepada kegagalan lokal dan kegagalan menyeluruh boleh mencetuskan keruntuhan progresif, menjadikannya satu aspek penting dalam penilaian struktur. Kajian ini menyiasat dan meramalkan tingkah laku keruntuhan progresif bagi kerangka konkrit bertetulang dua dimensi (2D) yang direkabentuk untuk menampung beban graviti sahaja dengan mengambil kira penyingkiran tiang yang berada di dalam dan tiang luaran. Analisis linear statik linear dan analisis tidak linear dinamik dilakukan untuk mengkaji dan meramalkan laluan "Alternate Load Path" (ALP) serta tindak balas dinamik struktur di bawah senario kehilangan tiang tersebut menggunakan perisian SAP2000. Tiga kes telah dipertimbangkan, iaitu Kes 1 (tiada kegagalan tiang), Kes 2 (penyingkiran satu tiang dalaman) dan Kes 3 (penyingkiran satu tiang luaran). Hasil analisis statik linear menunjukkan bahawa General Services Administration (GSA) menghasilkan daya yang sedikit lebih tinggi dan lebih konservatif berbanding British Standard (BS). Tanpa mengira lokasi tiang yang disingkirkan, ALP mengakibatkan keruntuhan progresif telah mengubah tindak balas struktur dengan ketara, khususnya bagi Kes 2 dan Kes 3. Analisis tidak linear dinamik pula menunjukkan bahawa Kes 3 merekodkan anjakan dan 'interstorey drift ratio' (IDR) yang paling tinggi, terutamanya apabila dikenakan gempa bumi 'Ground Motion' (GM5). Kajian turut menunjukkan bahawa penyingkiran satu tiang luaran menghasilkan IDR yang paling tinggi, sekali gus menunjukkan bahawa keadaan tersebut merupakan senario yang paling kritikal dan menjadikan struktur lebih terdedah kepada keruntuhan progresif.

KEYWORDS: *Progressive Collapse, Column Removal, Alternate Load Path, Nonlinear Dynamic Analysis*

1. INTRODUCTION

Structures are designed to be safe and robust according to design code guidelines. However, extreme events such as earthquakes, explosions, and terrorist attacks usually cause local damage that can lead to total collapse. Initial local damage can be caused by one or more columns (load-bearing) or a portion of the bearing wall failing [1]. These localized structural failures can propagate across adjacent components, potentially triggering a total or disproportionate collapse of the entire system, which is classified as progressive collapse, as shown in Fig. 1. Several progressive collapses have been reported, such as the A. P. Murrah Federal Building (Oklahoma, 1995) and the World Trade Center (New York, 2001), which caused not only economic but also social impacts [2]. Adam et al. [3] stated that the failure of one or more walls or columns, which are vertical load-bearing elements regarded as the most devastating local damage, will set off a series of events that ultimately result in the complete collapse of the building or a significant portion of it. Feng [4] stated that the General Services Administration (GSA), the U.S. Department of Defense (DoD), and several other building codes provide guidelines for progressive collapse analysis. These codes offer detailed information and techniques, such as the alternate path method (ALP), to prevent the gradual collapse of building structures. Nevertheless, the risk of progressive collapse can be assessed using linear static (LS) and nonlinear static (NS) analysis according to the U.S. General Services Administration (GSA) and the U.S. Department of Defense. A threat-independent method can assess a structure's resistance to progressive collapse when high loading leads to the loss of several columns [5]. Fig. 2 shows the most common examples of alternate load paths that may develop during a progressive collapse event. Elkady et al. [6] stated that studying progressive collapse and mitigation has become a major research topic in structural engineering, driven by the increasing number of studies and the development of new design standards. However, comprehensive structural design should focus on the interaction among structural elements to better understand and predict structural behavior under extreme loading conditions [7].

Progressive collapse is a dynamic phenomenon and is reliably evaluated by adopting nonlinear dynamic analysis [8]. Izzuddin et al. [9] studied progressive collapse assessment using nonlinear static response with dynamic effects in multi-story buildings under sudden column-loss scenarios. Arshian et al. [10], Hefez et al. [11], and Kim and Kim [12] studied progressive collapse frames using alternate path methods according to GSA [13] and DoD [14] guidelines. Progressive collapse of steel moment-resisting frames was evaluated using linear static and nonlinear dynamic analysis procedures based on the same design codes. The results indicated that nonlinear dynamic analysis provided better structural performance predictions, while the linear static procedure produced more conservative outcomes [12]. Tsai and Lin [15] performed the progressive collapse analysis using the linear static analysis method recommended by the US General Services Administration (GSA). In progressive collapse, most studies focused on single-column failure scenarios [16-17]. However, multiple-column removal may result in heavy damage or destruction under different time frames [18]. The studies found that removing consecutive corner columns induces more severe damage. Undeniably, corner columns are considered key load-bearing members in a building structure [19-20]. Jin et al. [21] studied the collapse resistance of RC frames, focusing on corner column removal. Compared to interior or edge columns, corner columns may have a higher probability of failure due to their position, especially under vulnerable abnormal load events. Furthermore,

corner column loss would be more significant because there is no column to support it. Buitrago et al. [22] stated that DoD regulations for frame structures require removing corner or exterior middle columns on the long and short sides one by one for collapse analysis.

Using 2D models focuses the analysis on a single structural plane, providing a simplified yet effective way to analyze load redistribution during collapse, as in studies by Ren et al. [23]. Progressive collapse analysis involves geometric and material nonlinearities, making detailed 3D simulations computationally demanding. As a result, 2D models are often used as a practical screening tool to quickly identify the most critical failure modes [37]. This study investigates the progressive collapse behavior of non-seismically designed reinforced concrete (RC) 2D frames under single interior and exterior column removal scenarios. These two column-loss cases are examined because they represent different in-plane load-redistribution mechanisms under both linear and nonlinear dynamic analyses. Sudden column removal is simulated to evaluate the structural vulnerability and robustness of simplified 2D RC frames. A 2D frame provides a computationally efficient approach for conservatively estimating in-plane load redistribution and catenary action, while offering insights into the progressive collapse response that can support structural assessment, especially in analysis and design.



a) Collapse of the A. P. Murrah Building [24]

b) Collapse of Champlain Towers South [25]

Figure 1. Examples of progressive collapse events.

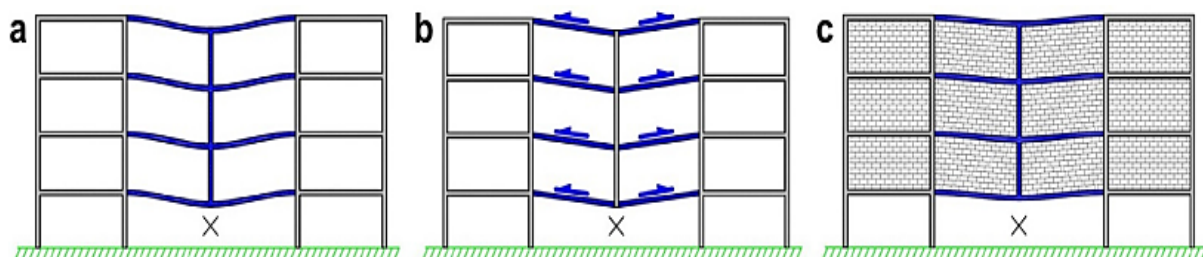


Figure 2. Alternate load paths may occur as (a) Vierendeel action; (b) catenary action; and (c) contribution of non-structural elements [3]

2. MODELING 2D RC FRAMES

This study focuses on the progressive collapse of a non-seismically designed existing building: a 4-story staff quarters at SMK Ranau in Sabah. The building has a symmetrical layout. Meng [26] provides further details on the structural reinforcements. The beam size used is 250 mm × 500 mm, taken for all levels. Each joint is assigned a diaphragm. The column size is taken as 250 mm × 250 mm. The 2D RC frames were modeled using SAP2000 software

[27]. Fig. 3 and Fig. 4 illustrate the front view and layout plan of staff quarters, respectively. Fig. 4 shows the layout of the structure (7-bay frame) selected for the current study. The study assumed the critical section for the 2D RC frame is in the internal section. The total span and height of the 7 frames are 28.06 m and 13.45 m, respectively. The structure height is 13.45 m. The concrete compressive strength, f_{ck} , is 30 N/mm². The structure is assumed to be fixed at the supports. The frames were designed to meet the specifications of British Standard BS 8110 [28], considering gravity loading using a safety factor of 1.4 for dead load (Gk) + 1.6 for live load (Qk). The floor slab thickness is assumed to be 150mm. Table 1 shows the loading details for the frames.



Figure 3. Front view of staff quarters of SMK Ranau.

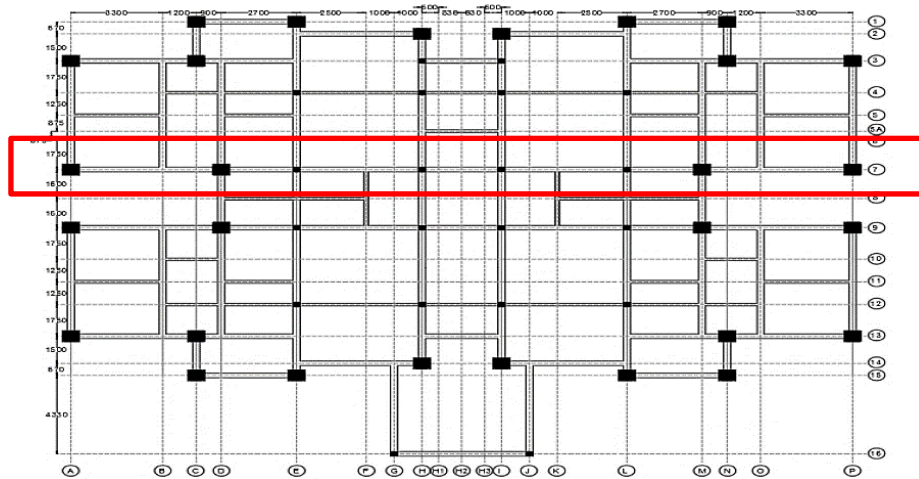


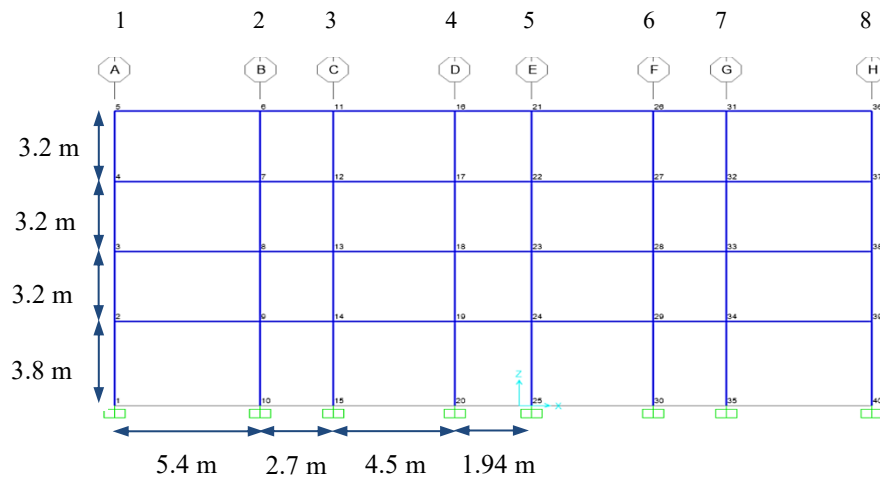
Figure 4. Layout plan of staff quarters of SMK Ranau [26].

Table 1. Detailed loading information on the frames

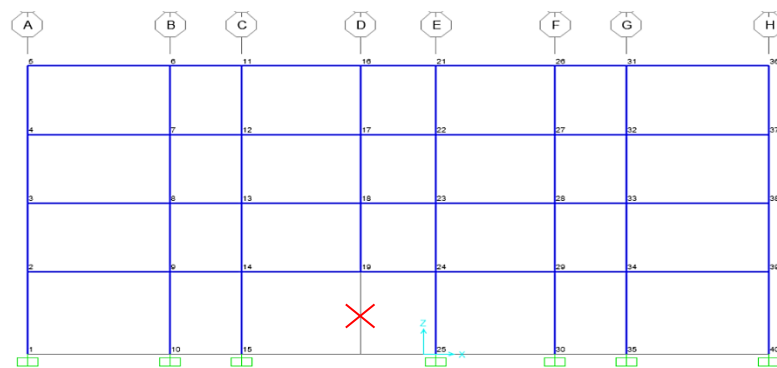
Building Level	Dead Load (kN/m ²)	Live Load (kN/m ²)
1	1.65	1.5
2	1.65	1.5
3	1.65	1.5
4	15 (Water Tank)	1.5

2.1. Progressive Collapse Analysis

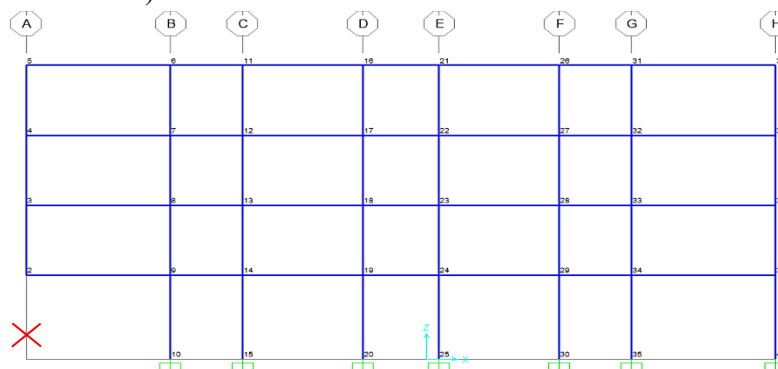
In this study, both interior and exterior columns were deliberately removed to simulate and predict the building's response to progressive collapse using linear and nonlinear analyses. The focus was on removing vertical elements, specifically a single column in the interior and exterior sections of the ground floor.



a) Case 1: No removal of column.



b) Case 2: Removal of one interior column.



c) Case 3: Removal of one exterior column (left).

Figure 5. 2D frames without and with column removal.

Three distinct cases are considered as outlined below:

- Case 1: No removal of column.

- Case 2: Removal of one interior column (middle).
- Case 3: Removal of one exterior column (left).

Fig. 5 shows that three distinct cases are analyzed for progressive collapse in a 2D RC frame. For the progressive collapse analysis, the frames were loaded with a safety factor of 0.2 for dead load (Gk) + 2.0 for live load (Qk), as suggested by GSA [13]. Both GSA [13] and DoD [14] recommend using linear static analysis to address a structure's ability to resist progressive collapse and to observe and assess alternate load path distribution within the existing frame components when columns at different locations are removed. Removing this column denotes a local failure triggered by progressive collapse during earthquakes.

2.2. Nonlinear Dynamic Analysis

This study focused on the progressive collapse of 2D reinforced concrete (RC) frames subjected to seven earthquake datasets. Rock-soil conditions were assumed for all seismic inputs. Fig. 6 illustrates representative earthquake records, with the complete list and details in Table 2. All of these earthquakes are normalized to 0.2g peak ground acceleration (PGA) according to Hatzigeorgiou et al. [29], assuming low- to medium-magnitude earthquakes in Malaysia. The 2D frames are subjected to earthquakes to assess outcomes such as displacements and inter-story drift, mainly under progressive collapse. A damping ratio of 5% was applied, which is typical and appropriate for RC structures. The seven earthquakes are compatible with EC8 [36] for Soil B local conditions with 0.2 PGA, as shown in Fig. 7.

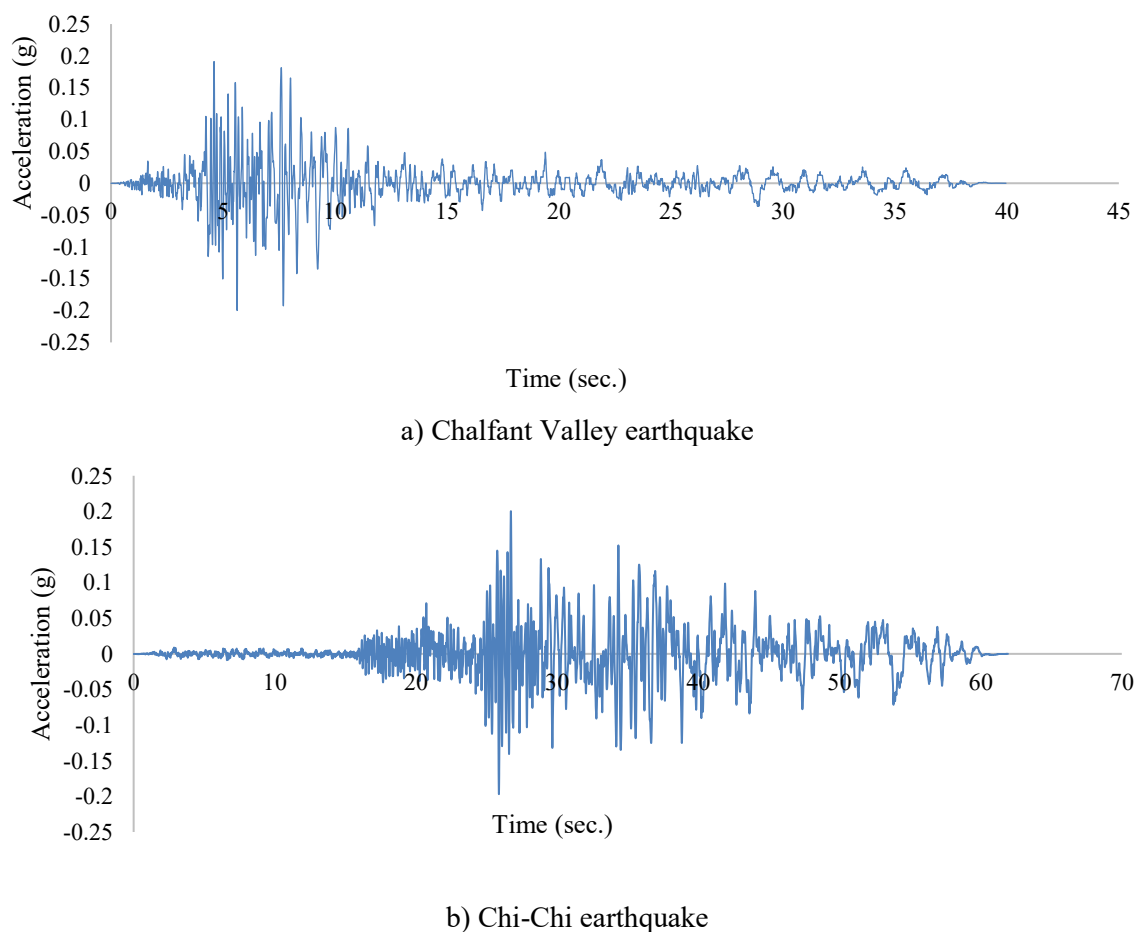


Figure 6. Example of an earthquake used in the current study.

Table 2. List of earthquakes used by Hatzigeorgiou et al., [29]; Abrahamson et al., [30]

GM	Earthquake	Station	Magnitude	PGA (g)
1	Chalfant Valley	Zack Brothers Ranch	6.3	0.0586
2	Chi-Chi	CHY025	6.2	0.0129
3	Coalinga	46T04 CHP	6.0	0.059
4	Imperial Valley	Holtville P.O.	6.6	0.258
5	Irpinia	Bagnoli Irpino	6.2	0.021
6	Mammoth Lakes	Convict Creek	6.1	0.314
7	Whittier Narrow	San Marino	5.3	0.138

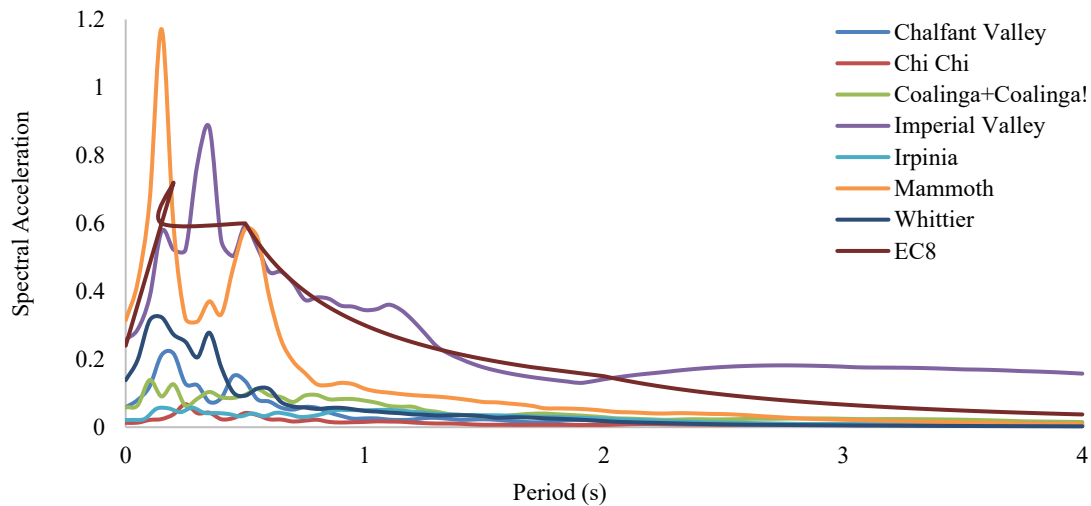


Figure 7. Response spectra of the selected earthquakes with the EC8 design spectrum

3. RESULTS AND DISCUSSION

For validation, this study compared results from Vinay et al. [31], who investigated progressive collapse by removing a single internal column using SAP2000. This study successfully modeled and reproduced the bending moment distribution with the alternate load path, with a 5.5% difference in the second-floor internal beam, as shown in Fig. 8. The difference may be attributed to software implementation, the modeling approach and assumptions, material properties, and numerical discretization, including differences between the modeling approaches adopted. The bending moment patterns were also similar, supporting the adoption of the current study's modeling approach for further analysis.

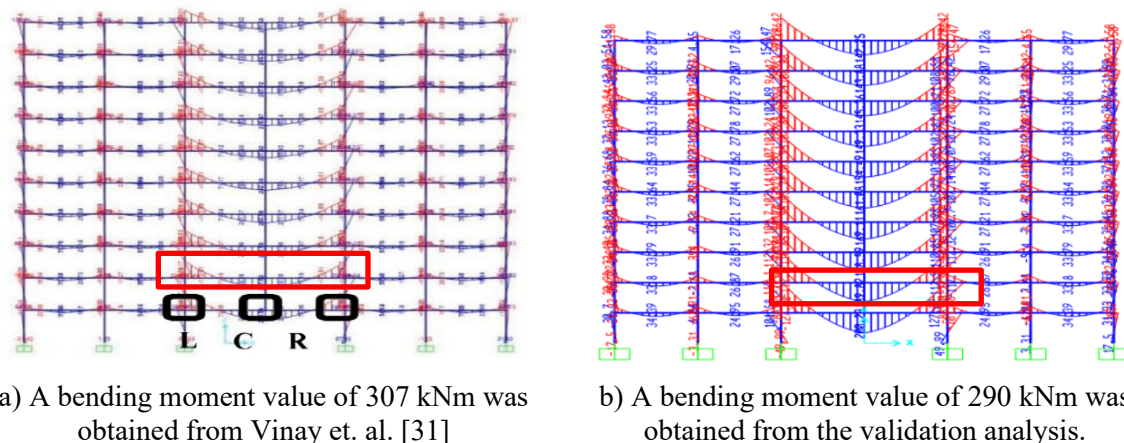


Figure 8. Comparison of bending moment values for validation analysis.

3.1. Alternate Load Path under Progressive Collapse

A linear static analysis was conducted across three distinct scenarios to evaluate the alternative load path mechanism and load redistribution capacity of a 2D reinforced concrete frame following the sudden removal of an interior or exterior column. The analytical findings were validated against the provisions of the British Standard (BS 8110) and the General Services Administration (GSA) guidelines. The structural response indicates that sudden column loss induces abrupt stress concentrations in adjacent columns. If these neighboring structural elements lack the requisite redundant design capacity to sustain these transferred loads, localized overstressing may trigger progressive collapse. Fig. 9 presents a comparative analysis of the resulting bending moments under both BS and GSA standards. An evaluation of the bending moment distributions across the seven spans reveals distinct geometric trends and establishes Case 3 as the most severe scenario for localized flexural demand. Case 1 represents a highly stable baseline across both design frameworks, exhibiting a flat trend line where bending moments remain uniform and minimal, fluctuating strictly between 50 kNm and 100 kNm. Conversely, Case 2 demonstrates a distinct parabolic or bell-shaped trend, wherein flexural demands originate at lower magnitudes at the outer spans, progressively climb to a maximum peak at the central 4th span, and symmetrically taper off toward the terminal spans. Under the British Standard (BS), this central peak reaches approximately 215 kNm, whereas the General Services Administration (GSA) guidelines amplify the entire curve, raising the peak to roughly 245 kNm, a 13.9% increase in peak flexural demand over BS.

Case 3 emerges as a critical vulnerability within the flexural profile, characterized by a sharply downward-sloping trend that originates with a massive, highly concentrated stress concentration at the 1st span. Under BS, this initial peak reaches approximately 300 kNm and increases to around 345 kNm under GSA provisions. This represents a 15.0% increase in localized bending stress compared to BS, constituting the highest absolute flexural magnitude recorded across the entire study and indicating severe localized beam vulnerability at the initial span. The redistribution of axial forces across the eight columns outlines a highly localized load-transfer mechanism, identifying Case 2 as the most severe scenario for critical column overloading. For both Case 1 and Case 3, the axial force profiles follow a symmetric, wave-like trend that peaks near the interior columns (Columns 3, 4, and 5) before gradually tapering down toward the perimeter. For example, the axial load sustained by Column 4 measures approximately 680 kN under BS but scales up to roughly 780 kN under GSA, representing a 14.7% vertical shift in the trend line.

In conclusion, the GSA code consistently projects a trend line that sits between 13.9% and 15.0% higher than the British Standard. This uniform upward shift underscores that the GSA code incorporates a structurally more conservative safety factor or dynamic amplification factor when evaluating alternate load paths (ALP) during progressive collapse scenarios. Case 3 constitutes the most severe condition for beam flexure (1st span peak), while Case 2 represents the most severe condition for column axial capacity (Column 5 peak). These findings are consistent with those reported by Parisi and Scalvenzi [32], Alshaikh et al. [33], Alogla et al. [34], and Jian and Li [35]. The authors observed that the loss of a column causes the surrounding region, which was originally designed to resist tensile forces, to experience significant compressive stresses, thereby increasing the load demand on adjacent structural members. Similarly, Elkady et al. [6] stated that the progressive collapse response varies with the structural system and collapse mechanism, particularly when assessed using the Alternate Load Path (ALP) method. This behavior is also evident in the results of the present study.

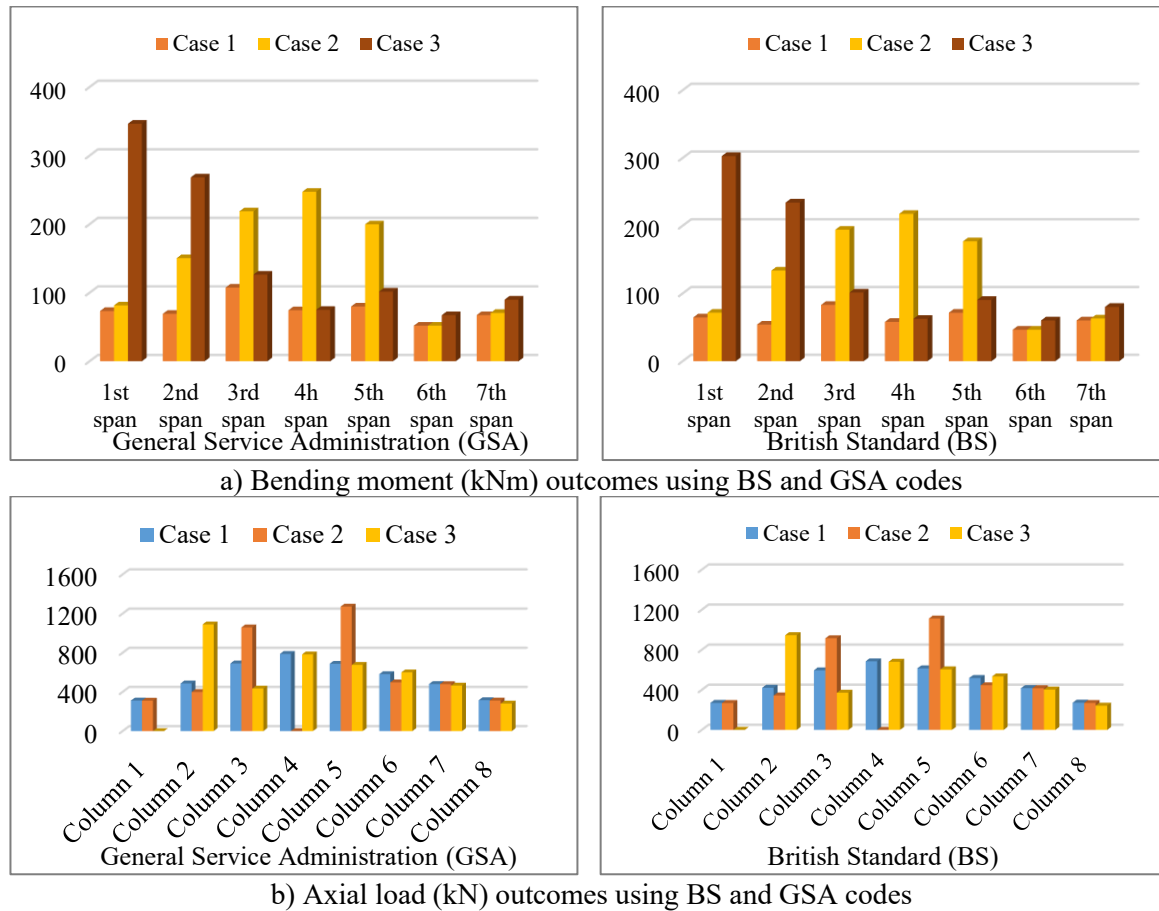


Figure 9. Alternate load path (ALP) under linear static analysis for different cases

3.2. Nonlinear Dynamic Analysis

The seismic response of the two-dimensional reinforced concrete (RC) frame was evaluated under seven earthquake records, with particular emphasis on story displacements and inter-story drift ratio (IDR). The IDR was determined as the ratio of the relative lateral displacement between two consecutive stories to the corresponding story height. This study investigates the impact of different column-removal scenarios on structural response and evaluates the potential for progressive collapse. Fig. 10 presents the displacement and inter-story drift under the GM5 ground motion for three cases: Case 1 (no column removal), removal of one interior (middle) column, and removal of one exterior (left) column. The structural response under progressive collapse induced by Ground Motion 5 (GM5) was evaluated for three design cases. As shown in Fig. 10a, the lateral floor displacement increased with building height, reaching a maximum at the roof (Floor 4). Case 1 exhibited the highest lateral stiffness, with a peak roof displacement of 0.098 m. In contrast, the roof displacement increased to 0.103 m in Case 2 (5.1% higher than Case 1) and to 0.128 m in Case 3, representing a 30.6% increase, indicating a significant reduction in structural stiffness. Fig. 10b presents the interstory drift ratio (IDR), which displays a similar parabolic distribution for all cases, with the maximum drift concentrated at the first floor and gradually decreasing toward the upper stories. The peak IDR was 1.48% for Case 1, increasing to 1.78% in Case 2 (20.3%) and 1.85% in Case 3 (25.0%).

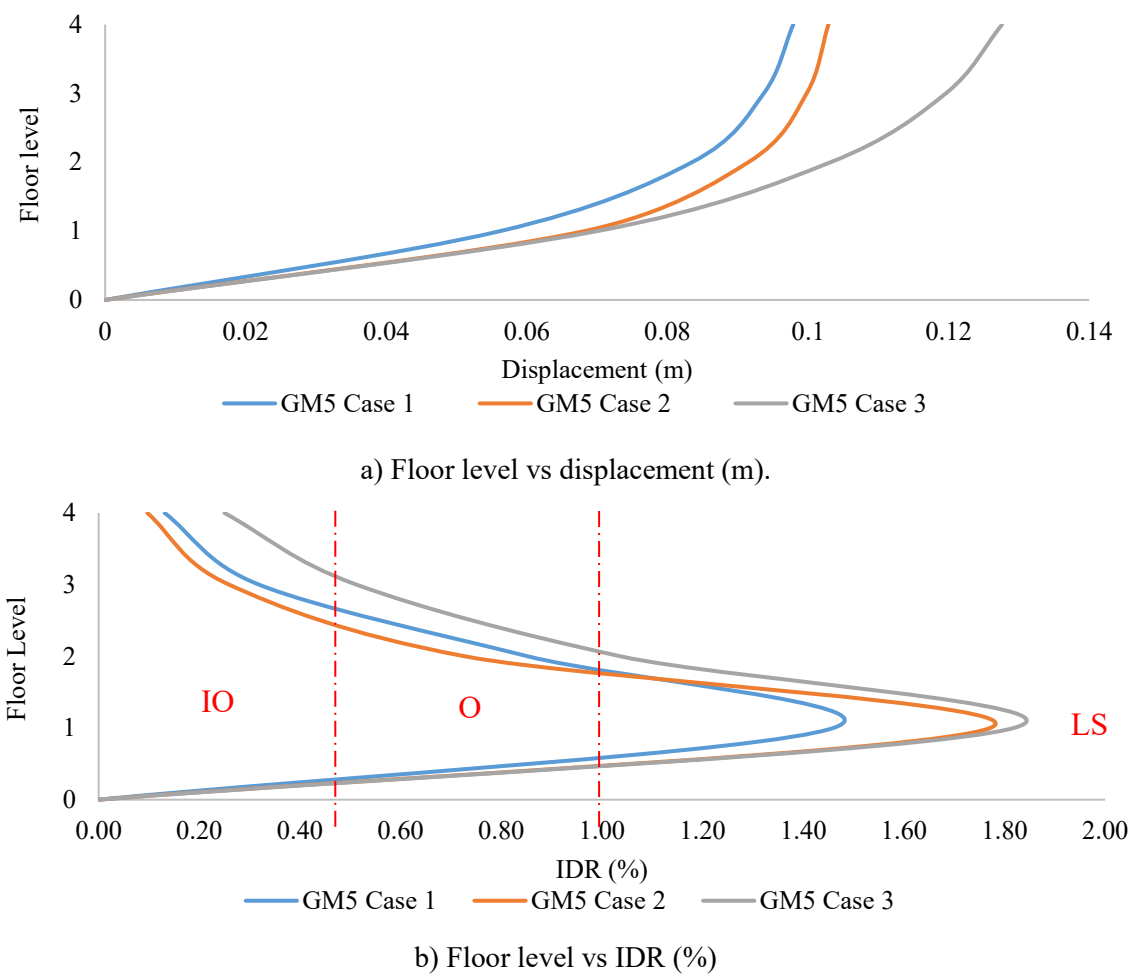


Figure 10. Nonlinear dynamic analysis under progressive collapse for all cases subjected to GM5

Under GM5, comparing the optimized baseline (Case 1) with the most severe scenario (Case 3) shows a 30.6% increase in peak roof displacement and a 25.0% increase in the maximum first-story interstory drift ratio (IDR). These results indicate that drift demand increases as structural robustness decreases. Removing the exterior column significantly reduces structural stability, causing deformation to concentrate at the first story and producing the highest drift demands. In all cases, the first-story IDRs exceed the Immediate Occupancy (IO, 0.50%) and Operational (O, 1.00%) performance limits but remain within the Life Safety (LS) range. This deformation concentration highlights the need to enhance continuity ties and provide redundant load paths to improve progressive collapse resistance. Fig. 11 further confirms that Case 3 exhibits the most critical progressive collapse response under all seven earthquake records. Peak IDRs consistently occur at the first story, although their magnitude depends on the ground-motion frequency characteristics. Among the records, GM5 produces the most critical response, with a peak IDR of 1.85% and a roof displacement of 0.128 m. Within Case 3, the peak IDR under GM5 is 71.6% higher than GM2, 131.3% higher than GM3, and 640.0% higher than GM7 (0.25%), demonstrating the structure's strong sensitivity to seismic excitation and its increased vulnerability to first-story failure following exterior column loss.

According to Ghobarah [37], IDR limits classify progressive collapse performance as IO ($\leq 0.5\%$), O (0.5-1.0%), LS (1.0-2.5%), and collapse ($> 2.5\%$). Most ground motions (GM1-GM4, GM6, and GM7) remained within the O-LS range, indicating controlled damage and

structural stability. However, GM5 produced significantly higher IDR demands at lower stories, exceeding the LS limit and approaching collapse levels, indicating soft-story development and increased vulnerability to progressive collapse.

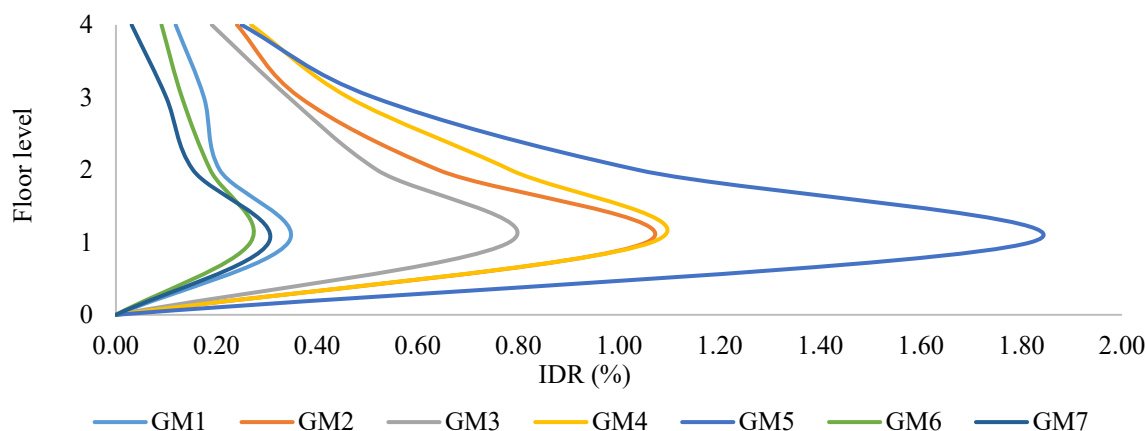


Figure 11. 2D RC frames under seven earthquakes for Case 3

4. CONCLUSIONS

This study used linear static and nonlinear dynamic analyses to predict the progressive collapse resistance of 2D reinforced concrete (RC) frames, accounting for single-column removal in interior and exterior frames. Three cases are considered: Case 1: no column removal, Case 2: removal of one interior column (middle), and Case 3: removal of one exterior column (left). The study's aims and findings can be summarized as:

For alternate load path (ALP), abrupt column removal caused load redistribution, resulting in significant stress increases in adjacent members. Under linear static analysis, the GSA code consistently projects a trend line that sits between 13.9% and 15.0% higher than the British Standard. This uniform upward shift suggests that the GSA code applies a more conservative safety factor or dynamic amplification factor when evaluating alternate load paths (ALP) during progressive collapse scenarios. Case 3 constitutes the most severe condition for beam flexure (1st span peak), while Case 2 represents the most severe condition for column axial capacity (Column 5 peak).

Next, for nonlinear dynamics, Case 3 is the most critical and vulnerable, with the highest IDR of 1.85% under GM5. This triggers a soft-story-like response at level 1, about a 25% increase in lateral sway compared to Case 1. These findings suggest that both internal and exterior column removal should be considered when assessing structural vulnerability, mainly exterior, under progressive collapse.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support provided by the Ministry of Higher Education (MoHE), Malaysia, through the Fundamental Research Grant Scheme-Early Career Researcher (FRGS-EC), under Grant No. FRGS-EC24-014-0014.

REFERENCES

- [1] Ellingwood BR (2006) Mitigating risk from abnormal loads and progressive collapse. *J. Perform. Constr. Facil.*, 20(4):315-323.

-
- [2] Russell JM, Sagaseta J, Cormie D, Jones AEK (2019) Historical review of prescriptive design rules for robustness after the collapse of Ronan Point. *Structures*, 20:365-373. <https://doi.org/10.1016/j.istruc.2019.04.011>
 - [3] Adam JM, Parisi F, Sagaseta J, Lu X (2018) Research and practice on progressive collapse and robustness of building structures in the 21st century. *Eng. Struct.*, 173:122-149. <https://doi.org/10.1016/j.engstruct.2018.06.082>
 - [4] Feng F (2010) 3-D nonlinear dynamic progressive collapse analysis of multi-storey steel composite frame buildings: Parametric study. *Eng. Struct.*, 32:3974-3980. <https://doi.org/10.1016/j.engstruct.2010.09.008>
 - [5] American Society of Civil Engineers (2023) ASCE 76-23: Standard for mitigation of disproportionate collapse potential in buildings and other structures. ASCE.
 - [6] Elkady N, Nelson LA, Weekes L, Makoond N, Buitrago M (2024) Progressive collapse: Past, present, future and beyond. *Structures*, 62:106131. <https://doi.org/10.1016/j.istruc.2024.106131>
 - [7] Li H, El-Tawil S (2014) Three-dimensional effects and collapse resistance mechanisms in steel frame buildings. *J. Struct. Eng.*, 140. [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0000839](https://doi.org/10.1061/(ASCE)ST.1943-541X.0000839)
 - [8] Marjanishvili S, Agnew E (2006) Comparison of various procedures for progressive collapse analysis. *J. Perform. Constr. Facil.*, 20(4):365-374. [https://doi.org/10.1061/\(ASCE\)0887-3828\(2006\)20:4\(365](https://doi.org/10.1061/(ASCE)0887-3828(2006)20:4(365)
 - [9] Izzuddin BA, Vlassis AG, Elghazouli AY, Nethercot DA (2008) Progressive collapse of multi-storey buildings due to sudden column loss—Part I: Simplified assessment framework. *Eng. Struct.*, 30(5):1308-1318. <https://doi.org/10.1016/j.engstruct.2007.07.011>
 - [10] Arshian AH, Morgenthal G, Narayanan S (2016) Influence of modelling strategies on uncertainty propagation in the alternate path mechanism of reinforced concrete framed structures. *Eng. Struct.*, 110:36-47. <https://doi.org/10.1016/j.engstruct.2015.11.019>
 - [11] Hefez I, Khalil A, Mourad S (2013) Alternate path method analysis of RC structures using applied element method. *Int. J. Protect. Struct.*, 4(1):45-64. <https://doi.org/10.1260/2041-4196.4>
 - [12] Kim J, Kim T (2009) Assessment of progressive collapse-resisting capacity of steel moment frames. *J. Constr. Steel Res.*, 65(1):169-179. <https://doi.org/10.1016/j.jcsr.2008.03.020>
 - [13] General Services Administration (2003) Progressive collapse analysis and design guidelines for new federal office buildings and major modernization projects. U.S. General Services Administration.
 - [14] Department of Defense (2009) Unified Facilities Criteria (UFC): Design of structures to resist progressive collapse. U.S. Department of Defense.
 - [15] Tsai MH, Lin BH (2008) Investigation of progressive collapse resistance and inelastic response for an earthquake-resistant RC building subjected to column failure. *Eng. Struct.*, 30(12):3619-3628. <https://doi.org/10.1016/j.engstruct.2008.05.031>
 - [16] Gudmundsson GV, Izzuddin BA (2008) The sudden column loss idealisation for disproportionate collapse assessment. *Struct. Eng.*, 86(6):22-26.
 - [17] Iribarren B, Berke P, Bouillard P, Vantomme J, Massart TJ (2011) Investigation of the influence of design and material parameters in the progressive collapse analysis of RC structures. *Eng. Struct.*, 33:2805-2820. <https://doi.org/10.1016/j.engstruct.2011.06.005>
 - [18] Parisi F, Scalvenzi M (2019) Progressive collapse assessment of gravity-load designed European RC buildings under multi-column loss scenarios. *Eng. Struct.*, 209:110001. <https://doi.org/10.1016/j.engstruct.2019.110001>
 - [19] Yang T, Liu Z, Lian J (2021) Progressive collapse of RC flat slab substructures with unbonded post-tensioning strands after the loss of an exterior column. *Eng. Struct.*, 234:111989. <https://doi.org/10.1016/j.engstruct.2021.111989>
-

-
- [20] Mohammadi A, Pachenari A, Sadeghi B (2021) Numerical study on the behavior and bearing mechanism of flat slabs in column loss events. *Adv. Civ. Eng.*, 2021:1-12. <https://doi.org/10.1155/2021/8834638>
- [21] Jin X, Wang S, Li K, Qian X, Deng F (2021) Study on collapse resistance of RC frame under the corner column removal scenario. *Materials*, 14(23):7157. <https://doi.org/10.3390/ma14237157>
- [22] Buitrago M, Bertolesi E, Sagaseta J, Calderón PA, Adam JM (2021) Robustness of RC building structures with infill masonry walls: Tests on a purpose-built structure. *Eng. Struct.*, 226:111384. <https://doi.org/10.1016/j.engstruct.2020.111384>
- [23] Ren P, Li Y, Lu X, Guan H, Zhou Y (2015) Experimental investigation of progressive collapse resistance of one-story reinforced concrete specimen under penult column removal. *Eng. Struct.*, 87:88-105. <https://doi.org/10.1016/j.engstruct.2015.01.011>
- [24] Jenkins JP (2001) Oklahoma City bombing. *Encyclopaedia Britannica*.
- [25] Winsor M (2021) Surfside building collapse: Death toll rises to 18 after 2 children found. *ABC News*. Available: <https://abcnews.go.com/US/surfside-building-collapse-latest-12-dead-149missing/story?id=78574497>
- [26] Meng HK (2019) Dynamic characteristics and seismic performance of reinforced concrete buildings experiencing 5 June 2015 and 8 March 2018 Sabah earthquakes. Master's thesis, School of Civil Engineering, Universiti Sains Malaysia.
- [27] Computers and Structures, Inc. (2025) SAP2000 (Version 26). Computers and Structures, Inc.
- [28] British Standards Institution (1997) BS 8110-1:1997 Structural use of concrete-Part 1: Code of practice for design and construction. BSI.
- [29] Hatzigeorgiou GD, Liolios AA (2010) Nonlinear behaviour of RC frames under repeated strong ground motions. *Soil Dyn. Earthq. Eng.*, 30(10):1010-1025. <https://doi.org/10.1016/j.soildyn.2010.04.013>
- [30] Abrahamson N, Atkinson G, Boore D, et al. (2008) Comparisons of the NGA ground-motion relations. *Earthq. Spectra*, 24(1):45-66. <https://doi.org/10.1193/1.2924363>
- [31] Vinay M, Rao PKR, Dey S, Swaroop AHL, Sreenivasulu A, Rao KV (2022) Evaluation of progressive collapse behavior in reinforced concrete buildings. *Structures*, 45:1902-1919. <https://doi.org/10.1016/j.istruc.2022.10.001>
- [32] Alshaikh IM, Bakar BA, Alwesabi EA, Akil HM (2020) Experimental investigation of the progressive collapse of reinforced concrete structures: An overview. *Structures*, 25:881-900. <https://doi.org/10.1016/j.istruc.2020.03.018>
- [33] Alogla K, Weekes L, Augusthus-Nelson L (2017) Theoretical assessment of progressive collapse capacity of reinforced concrete structures. *Mag. Concr. Res.*, 69(3):145-162. <https://doi.org/10.1680/jmacr.16.0031>
- [34] Jian H, Li S (2016) Progressive collapse resistance of RC frames under a side column removal scenario: The mechanism explained. *Int. J. Concr. Struct. Mater.*, 10:237-247. <https://doi.org/10.1007/s40069-016-0134-y>
- [35] Ghobarah A (2001) Performance-based design in earthquake engineering: State of development. *Eng. Struct.*, 23:878-884.
- [36] European Committee for Standardization (2004) Eurocode 8: Design of structures for earthquake resistance-Part 1: General rules, seismic actions and rules for buildings (EN 1998-1:2004). CEN.
- [37] El-Ariss B, Elkholy S, Shehada A (2022) Benchmark numerical model for progressive collapse analysis of RC beam-column sub-assemblages. *Buildings*, 12(2):122. <https://doi.org/10.3390/buildings12020122>
-

Numerical Evaluation of Geogrid-Reinforced Slopes under Various Soil and Geometric Configurations

ADIBAH FARHANA HAIRUL HISHAM, NORHIDAYU KASIM*,
WAN NUR AIFA WAN AZAHAR

*Department of Civil Engineering, International Islamic University Malaysia,
Kuala Lumpur, Malaysia*

**Corresponding author: ayukasim@iium.edu.my*

(Received: 31 July 2026; Accepted: 7 September 2026; Published online: 11 September 2026)

ABSTRACT: This study examines the stability of geogrid-reinforced slopes under different soil and geometric conditions using SLOPE/W. The Morgenstern–Price method was used to analyze three homogeneous soil parameter sets: silty clay, clayey sand, and clayey silt. Four slope inclinations (25°, 30°, 45°, and 60°) were considered, along with total slope heights of 18, 24, 30, and 36 m. The reinforced slopes were modeled using UX1100 uniaxial geogrid at vertical spacings of 0.5, 1.0, and 1.5 m. Each reinforced model was compared with its corresponding unreinforced model under the same dry, drained, and static conditions, with pore-water pressure set to zero. The results show that FOS decreased as the slope inclination and total height increased. The calculated FOS also varied with the soil parameters. Silty clay generally produced the highest FOS, while clayey sand produced the lowest despite having a higher friction angle, due to the combined effects of cohesion, friction angle, and unit weight. As expected, the 0.5 m spacing produced the highest FOS among the three spacings because it provided more reinforcement layers within the same slope height. However, closer spacing did not provide adequate stability in every case. For the 36 m slope at 60°, the reinforced FOS values were 1.49, 1.06, and 1.17 for silty clay, clayey sand, and clayey silt, respectively, which were below the adopted requirement of 1.5. The findings show that geogrid spacing should be assessed alongside soil conditions and slope geometry. The results provide comparative trends for the idealized dry slope conditions considered and are not intended as site-specific design recommendations.

ABSTRAK: Kajian ini menilai kestabilan cerun bertetulang geogrid bagi keadaan tanah dan geometri yang berbeza menggunakan SLOPE/W. Kaedah Morgenstern–Price digunakan untuk menganalisis tiga set parameter tanah homogen, iaitu lempung berkelodak, pasir berlempung dan kelodak berlempung. Empat kecerunan, iaitu 25°, 30°, 45° dan 60°, dipertimbangkan bersama jumlah ketinggian cerun 18, 24, 30 dan 36 m. Cerun bertetulang dimodelkan menggunakan geogrid uniaxial UX1100 pada jarak menegak 0.5, 1.0 dan 1.5 m. Setiap model bertetulang dibandingkan dengan model tanpa tetulang yang sepadan dalam keadaan kering, tersalir dan statik yang sama, dengan tekanan air liang ditetapkan kepada sifar. Keputusan menunjukkan bahawa Faktor Keselamatan (FOS) menurun apabila kecerunan dan jumlah ketinggian cerun meningkat. Nilai FOS yang dikira turut berubah mengikut parameter tanah. Lempung berkelodak secara umumnya menghasilkan FOS tertinggi, manakala pasir berlempung menghasilkan FOS terendah walaupun mempunyai sudut geseran yang lebih tinggi, disebabkan oleh pengaruh gabungan kejelekitan, sudut geseran dan berat unit. Seperti yang dijangka, jarak 0.5 m menghasilkan FOS tertinggi antara tiga jarak yang dianalisis kerana lebih banyak lapisan tetulang digunakan pada ketinggian cerun yang sama. Walau bagaimanapun, jarak yang lebih rapat tidak memberikan kestabilan yang mencukupi bagi semua keadaan. Bagi cerun setinggi 36 m pada kecerunan 60°, nilai

FOS bertetulang ialah 1.49, 1.06 dan 1.17 masing-masing untuk lempung berkelodak, pasir berlempung dan kelodak berlempung. Semua nilai tersebut berada di bawah keperluan FOS 1.5 yang digunakan dalam kajian. Dapatan menunjukkan bahawa jarak geogrid perlu dinilai bersama keadaan tanah dan geometri cerun. Keputusan ini memberikan trend perbandingan bagi keadaan cerun kering yang diidealkan dan tidak dimaksudkan sebagai cadangan reka bentuk khusus tapak.

KEYWORDS: *Geogrid reinforcement, Slope stability, SLOPE/W, Slope geometry, Reinforcement spacing*

1. INTRODUCTION

Slope stability is an important consideration in geotechnical engineering, particularly for slopes associated with roads, embankments, and other infrastructure. Slope failure occurs when the driving forces on a slope exceed the available resisting forces, which are mainly governed by the soil's shear strength and the slope's geometry. Slope stability can also be affected by rainfall infiltration, changes in groundwater conditions, and human activities such as slope cutting and inadequate drainage. In tropical regions, prolonged or intense rainfall can increase pore-water pressure and reduce soil shear strength, thereby increasing the possibility of slope failure [1, 2]. Different failure mechanisms and ground conditions may require different slope stabilization measures [3, 4]. Therefore, proper assessment of slope stability and the selection of suitable stabilization measures are important to ensure the safety and long-term performance of slopes. Although rainfall and groundwater are important in field slopes, this study considers only dry, drained, static conditions. Rainfall infiltration and groundwater were not included because the analysis was intended to compare the effects of soil properties, slope geometry, and reinforcement arrangement without an additional hydraulic variable.

Various stabilization methods have been applied to improve slope stability, including soil improvement, soil nailing, retaining structures, ground anchors, and geosynthetic reinforcement [5, 6]. Among these methods, geosynthetics are widely used in geotechnical applications because they provide reinforcement while offering flexibility in construction and application [7]. Geogrids are commonly used as reinforcement due to their tensile strength and ability to interact with the surrounding soil. The reinforcement contributes additional resistance when a geogrid layer crosses a potential slip surface. Its contribution depends on the geogrid's tensile capacity, the spacing and length of the layers, soil–geogrid interaction, and the available anchorage beyond the slip surface [8, 9]. Geosynthetic solutions have also been considered an alternative to conventional stabilization methods because of their versatility and potential environmental benefits [7]. Previous numerical studies have shown that reinforcement measures can increase the Factor of Safety (FOS) of unstable slopes under suitable design conditions [8, 9].

The performance of a reinforced slope, however, depends on several factors and cannot be determined only by the presence of reinforcement. Soil shear strength parameters, particularly cohesion and internal friction angle, have a direct influence on slope stability, while changes in slope height and inclination can considerably affect the FOS [10]. As the slope becomes steeper, the downslope component of soil self-weight increases, resulting in a higher driving force and generally a lower FOS. An increase in total slope height may also increase the size and weight of the potential sliding mass. Reinforcement characteristics and arrangement also determine the stability of reinforced soil systems. Closer vertical spacing increases the number of geogrid layers that may cross the potential slip surface and contribute tensile resistance. However, the resulting improvement also depends on the soil strength and slope geometry. In

addition, berms are commonly introduced in high slopes to modify the slope geometry and improve stability. Previous studies have shown that berms can affect the slope profile and potential sliding mechanism [11, 12]. These findings indicate that soil conditions, slope geometry, and reinforcement arrangement should be considered together when evaluating the performance of reinforced slopes.

Previous studies have investigated different slope stabilization techniques and the parameters affecting their performance. However, many studies have focused on a particular stabilization method or a limited set of parameters. For example, studies on soil nailing have examined the effects of nail length, inclination, and spacing [13], while studies involving geosynthetics have considered reinforcement properties and their effects on stability under specific conditions [7, 9]. The combined effects of soil condition, slope inclination, total slope height, stepped slope configuration, and geogrid spacing have received less attention. In the present analysis, the number of 6 m slope lifts increases together with the total slope height. Therefore, the analyzed cases represent the combined influence of slope height and stepped configuration rather than the independent effect of berm number.

This study numerically evaluates the stability of geogrid-reinforced slopes under different soil and geometric configurations using SLOPE/W. Three soil parameter sets, namely silty clay, clayey sand, and clayey silt, are considered, along with different slope inclinations and combined slope-height and stepped configurations. The reinforced slopes are analyzed using UX1100 uniaxial geogrid at vertical spacings of 0.5, 1.0, and 1.5 m. The calculated FOS values for the reinforced slopes are compared with those of the corresponding unreinforced slopes. The study aims to quantify the changes in FOS produced by the investigated geogrid spacings under different soil conditions, slope inclinations, and combined slope-height and stepped configurations. The analysis also identifies the configurations in which geogrid reinforcement alone is insufficient to achieve the adopted FOS requirement of 1.5.

2. METHODOLOGY

2.1. Numerical Modeling and Slope Geometry

The slope stability analyses were carried out using SLOPE/W to determine the Factor of Safety (FOS) of the unreinforced and geogrid-reinforced slopes. The Morgenstern–Price method was adopted in the limit equilibrium analysis. This method satisfies both force and moment equilibrium and considers the interslice shear and normal forces in the stability calculation [14].

Circular trial slip surfaces were generated using the Entry and Exit search method. The critical slip surface was taken as the trial surface that produced the lowest FOS. The same slip-surface search procedure was used for each unreinforced and reinforced model pair.

All models were analyzed under dry, drained, and static conditions. No groundwater function, rainfall infiltration, surcharge, or seismic coefficient was applied, and the pore-water pressure was set to $u = 0$. These conditions were kept unchanged throughout the analyses to allow direct comparison of the selected soil, slope geometry, and reinforcement variables.

Four local slope inclinations of 25°, 30°, 45°, and 60° were considered. Each slope lift was 6 m. Three, four, five, and six slope lifts produced total heights of 18, 24, 30, and 36 m, respectively. The 6 m slope-lift height and the range of configurations were selected based on the JKR Guidelines for Slope Design [15]. The same geometries were used for the corresponding unreinforced and reinforced models.

Because the number of slope lifts increased with total slope height, the analyzed cases were treated as combined slope-height and stepped configurations. Therefore, the analysis did not isolate the independent effect of berm number.

2.2. Soil and Geogrid Properties

Three homogeneous soil parameter sets were considered: silty clay, clayey sand, and clayey silt. The soil strength was represented using the Mohr–Coulomb model. The required input parameters were unit weight, effective cohesion, and effective friction angle, as listed in Table 1. The soil parameters were obtained from published secondary data [16] rather than laboratory testing for a specific project site. Each parameter set was applied uniformly throughout the corresponding slope model. Clayey sand has a higher friction angle than silty clay but lower cohesion and higher unit weight. Therefore, the FOS results were interpreted based on the combined effects of c' , ϕ' and γ , rather than the friction angle alone.

Table 1. Soil properties used in the numerical analysis

Soil type	Unit weight, γ (kN/m ³)	Cohesion, c' (kPa)	Friction angle, ϕ' (°)
Silty clay	18.0	4	30
Clayey sand	19.0	3	32
Clayey silt	19.5	3	28

For the reinforced slope analyses, Tensar UX1100 high-density polyethylene uniaxial geogrid was used. The geogrid properties were obtained from the manufacturer's product information used in the numerical model. The nominal tensile strength was 58 kN/m. An approximate combined material reduction factor of 2.69 was applied to account collectively for creep, installation damage, and durability. This produced a reduced design tensile capacity of 21.58 kN/m, which was entered into SLOPE/W. Since the material reduction had already been included in the design tensile capacity, the model tensile reduction factor was set to 1.0 to avoid applying the reduction twice.

Table 2. Geogrid properties and modeling settings

Parameter	Value or modeling setting
Product	Tensar UX1100
Material and type	HDPE uniaxial geogrid
Nominal tensile strength	58 kN/m
Combined material reduction factor	2.69
Reduced design tensile capacity	21.58 kN/m
Model tensile reduction factor	1.0
Pullout reduction factor	1.0
Reinforcement force option	FOS-dependent
Force direction	Axial along the horizontal geogrid layer
Vertical spacing	0.5, 1.0 and 1.5 m
Geogrid length for $H = 18\text{m}$	15 m
Geogrid length for $H = 24\text{m}$	20 m
Geogrid length for $H = 30\text{m}$	25 m
Geogrid length for $H = 36\text{m}$	30 m
Approximate L/H ratio	0.83
Layer extent	From the slope face into the stable soil zone
Anchorage	Portion extending beyond the critical slip surface
Tensile stiffness	Not used in the limit equilibrium model
Bond-slip stiffness	Not used in the limit equilibrium model

The geogrid was represented as horizontal axial reinforcement using the FOS-dependent option. It was not applied as a constant prestressing force. The available reinforcement force was controlled by the reduced tensile capacity and the pullout resistance mobilized along the anchorage length beyond the critical slip surface. The pullout reduction factor was set to 1.0.

No tensile stiffness or bond-slip stiffness was used because the limit equilibrium model calculates the global FOS rather than the geogrid deformation response. Table 2 summarises the geogrid properties and modeling settings.

The geogrid layers were placed horizontally from the slope face into the stable soil zone. The total geogrid lengths were 15, 20, 25 and 30 m for the 18, 24, 30 and 36 m slope configurations, respectively. These lengths gave an approximate reinforcement length-to-height ratio, L/H , of 0.83. The geogrid layers intersected by the potential slip surface contributed to the stability calculation, while the portion extending beyond the slip surface provided the available anchorage.

2.3. Analysis Configurations and Stability Assessment

The numerical analysis was conducted in two stages. First, each slope was analyzed without reinforcement to establish its baseline FOS. The same slope model was then reinforced with UX1100 geogrid and reanalyzed. The slope geometry, soil parameters, hydraulic condition, and slip-surface search procedure were kept unchanged within each unreinforced and reinforced model pair.

Three vertical geogrid spacings of 0.5, 1.0, and 1.5 m were investigated for different soil parameter sets, slope inclinations, and combined slope-height and stepped configurations. Table 3 summarizes the numerical analysis configurations.

Table 3. Numerical analysis configurations

Parameter	Conditions analyzed
Soil parameter set	Silty clay, clayey sand and clayey silt
Local slope inclination	25°, 30°, 45° and 60°
Number of 6 m slope lifts	3, 4, 5 and 6
Total slope height	18, 24, 30 and 36 m
Geogrid	Tensar UX1100 uniaxial geogrid
Geogrid vertical spacing	0.5, 1.0 and 1.5 m
Slip-surface search	Entry and Exit
Hydraulic condition	Dry and drained; $u = 0$
External loading	No surcharge or seismic coefficient
Stability parameter	Factor of Safety
Unreinforced-slope requirement	FOS = 1.3
Reinforced-slope requirement	FOS = 1.5

An FOS of 1.3 was adopted as the minimum requirement for unreinforced slopes, while an FOS of 1.5 was used for reinforced slopes, based on the adopted JKR slope design requirements [15]. These values were used to identify the configurations that satisfied the adopted stability requirement and those that remained below the required FOS.

2.4. Comparative Assessment of Reinforcement Performance

The contribution of the geogrid was evaluated by comparing the FOS of each reinforced slope with that of its corresponding unreinforced condition. The comparison considered the

magnitude of the FOS change and whether the reinforced configuration achieved the adopted FOS requirement of 1.5.

The paired comparison provided an internal consistency check because the slope geometry, soil inputs, hydraulic condition, and slip-surface search procedure were kept unchanged. However, this comparison does not constitute independent validation against field monitoring, laboratory testing, or a published benchmark. The findings are therefore interpreted as comparative numerical trends rather than site-specific design values.

3. RESULTS AND DISCUSSION

The numerical results were evaluated by first examining the unreinforced slopes and then comparing the changes after the inclusion of UX1100 geogrid. The discussion focuses on the influence of slope inclination, soil parameter set, combined slope-height and stepped configuration, and geogrid vertical spacing. The FOS was used as the main indicator of slope stability.

3.1. Stability of Unreinforced Slopes

The unreinforced analyses provided the baseline for evaluating the contribution of geogrid reinforcement. Fig. 1 shows a representative SLOPE/W output for the 18 m slope at 30° using the clayey sand parameter set. The critical slip surface produced an FOS of 1.28.

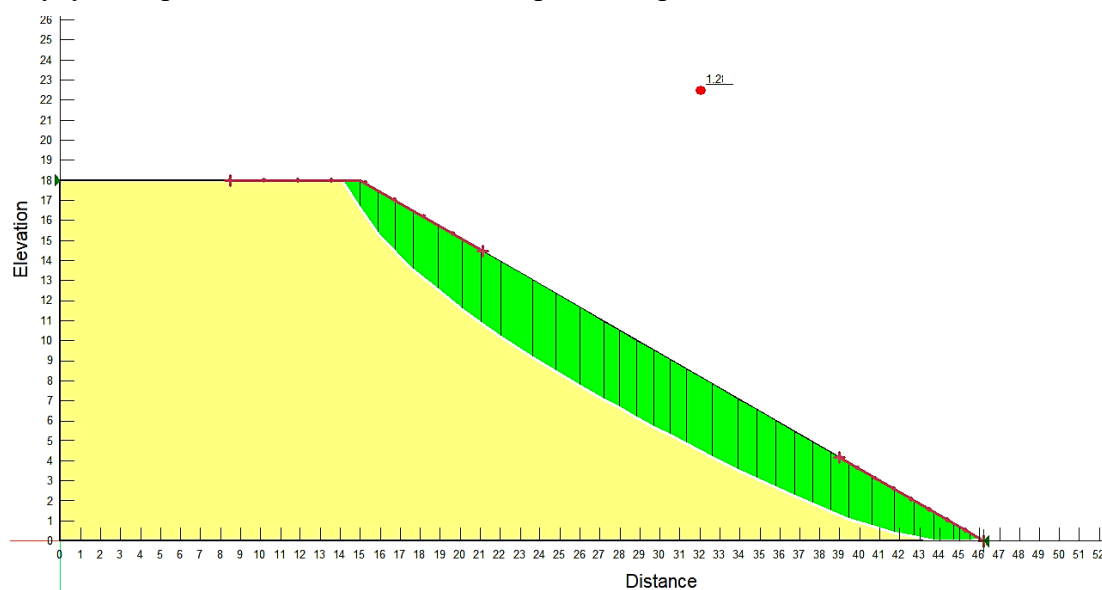


Figure 1. Representative critical slip surface for the unreinforced 18 m slope at 30° using the clayey sand parameter set (FOS = 1.28)

The FOS decreased as the slope inclination increased for all soil parameter sets and total slope heights. For the 18 m slope, the FOS of silty clay decreased from 2.41 at 25° to 1.10 at 60°. Over the same range, the FOS decreased from 1.53 to 0.58 for clayey sand and from 1.72 to 0.67 for clayey silt. A similar trend was observed for the other slope heights. For the 36 m configuration at 60°, the FOS values were 0.85, 0.50 and 0.57 for silty clay, clayey sand and clayey silt, respectively.

As the slope becomes steeper, a larger component of the soil self-weight acts downslope, increasing the driving force relative to the available shear resistance. The change in inclination

may also affect the size and position of the critical sliding mass. Previous numerical studies have reported similar effects of slope inclination on FOS [9, 10, 17].

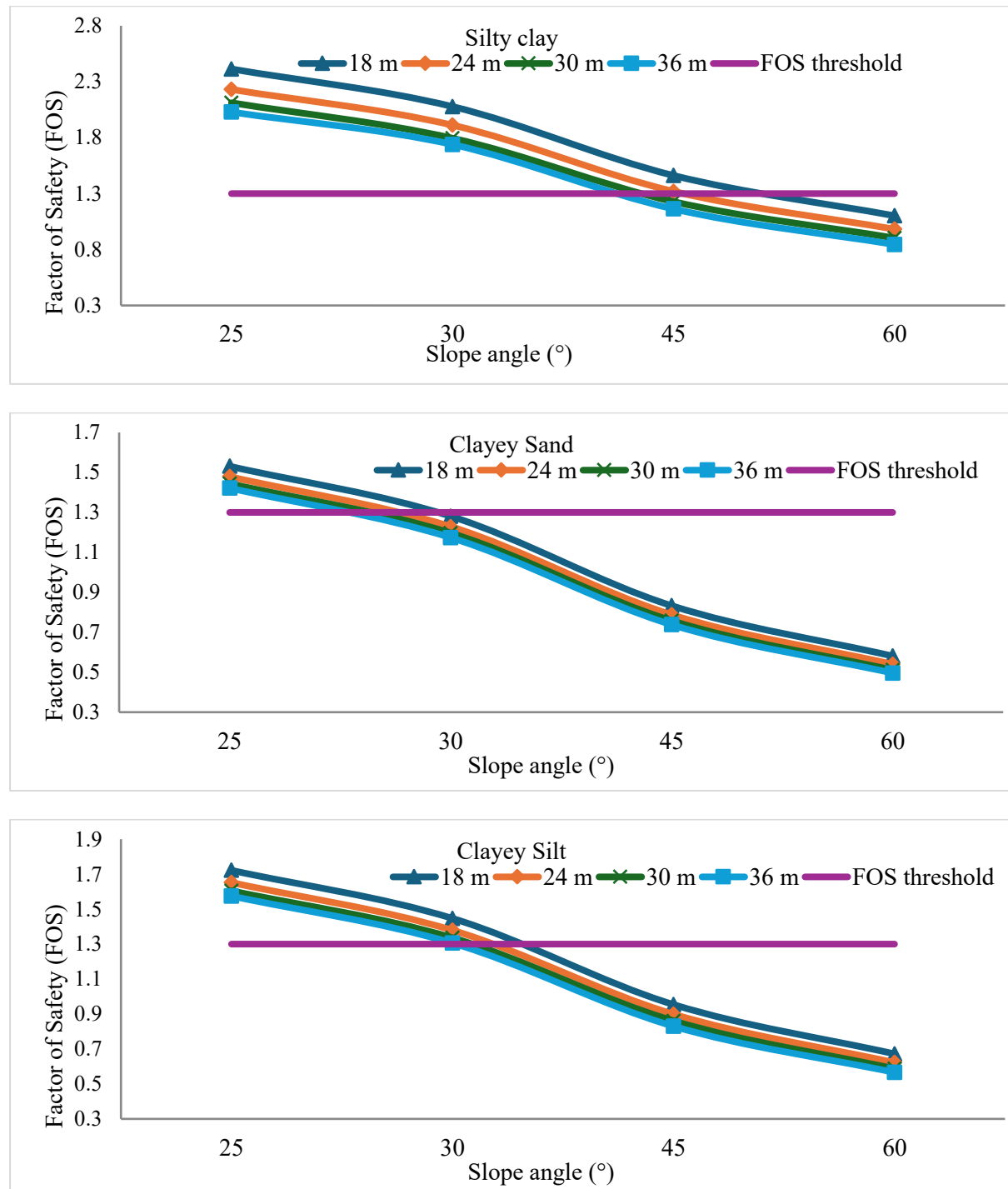


Figure 2. Variation of FOS with slope angle for unreinforced slopes at different slope heights

Fig. 2 compares the unreinforced FOS values with the adopted minimum requirement of 1.3. Several gentler configurations remained above the required value. However, the number of cases below FOS = 1.3 increased as the slope inclination and total height increased. The 60° slopes were the most critical, particularly for clayey sand and clayey silt.

The soil parameter set also influenced the calculated FOS. Silty clay generally produced the highest FOS, followed by clayey silt, while clayey sand produced the lowest. For the 30 m slope at 25°, the FOS values were 2.11, 1.61 and 1.44 for silty clay, clayey silt and clayey sand, respectively. At 45°, the corresponding values decreased to 1.23, 0.86 and 0.76.

The available Mohr–Coulomb shear resistance depends on the combined contribution of effective cohesion and frictional resistance. Although clayey sand had a higher friction angle than silty clay, it also had lower cohesion and higher unit weight. Its lower FOS therefore resulted from the combined soil parameters and stress conditions rather than the friction angle alone.

Figure 3 shows the unreinforced results at 45° for the combined slope-height and stepped configurations. For silty clay, the FOS decreased from 1.46 for the 18 m configuration to 1.16 for the 36 m configuration. The FOS decreased from 0.83 to 0.74 for clayey sand and from 0.95 to 0.83 for clayey silt over the same range.

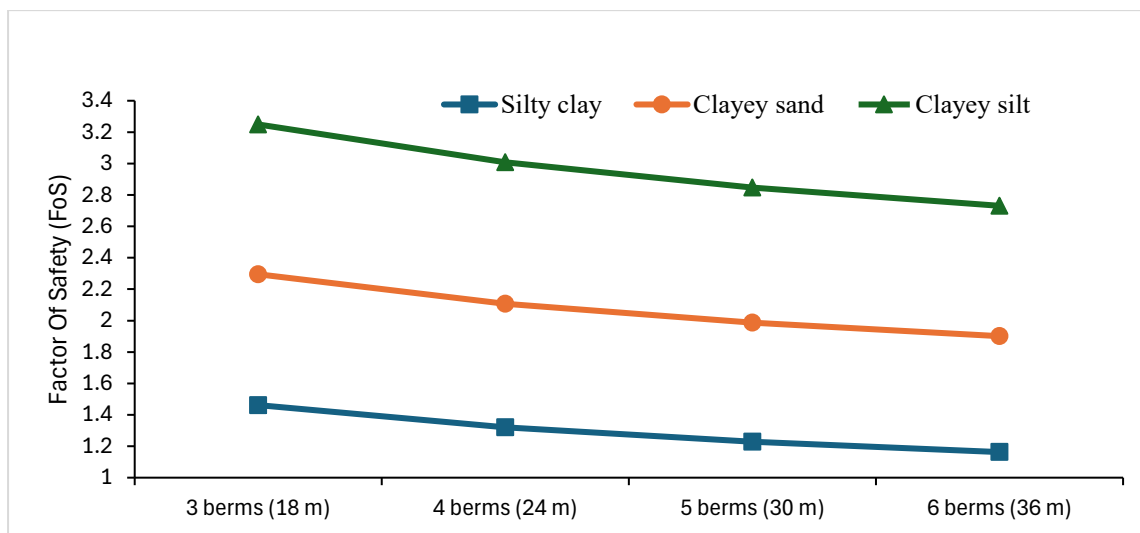


Figure 3. Variation of unreinforced FOS with the combined slope-height and stepped configuration at 45°

The reduction should not be attributed only to the number of berms because the number of slope lifts and total slope height increased together. The observed result represents their combined influence. A constant-height analysis in which the number and width of berms are varied independently would be required to isolate the effect of berm configuration.

3.2. Effect of Geogrid Reinforcement on Slope Stability

Including geogrid increased the FOS relative to the corresponding unreinforced slopes. However, the magnitude of improvement depended on the soil parameter set, slope inclination, total height, and reinforcement spacing. Fig. 4 shows a representative reinforced SLOPE/W model and its critical slip surface.

For the 18 m slope at 45°, the silty clay FOS increased from 1.46 without reinforcement to 2.15 with geogrid at 0.5 m vertical spacing. The clayey sand FOS increased from 0.83 to 1.51, while the clayey silt FOS increased from 0.95 to 1.67.

At 60°, the silty clay FOS increased from 1.10 to 2.12 at the same spacing. Clayey sand increased from 0.58 to 1.43, while clayey silt increased from 0.67 to 1.58. Although clayey sand showed a large percentage improvement, its final FOS remained below 1.5. Therefore,

the percentage increase alone does not indicate whether the reinforced slope meets the adopted stability requirement. The final reinforced FOS must also be considered. Table 4 summarizes selected reinforced and unreinforced results to demonstrate the improvement provided by the geogrid.

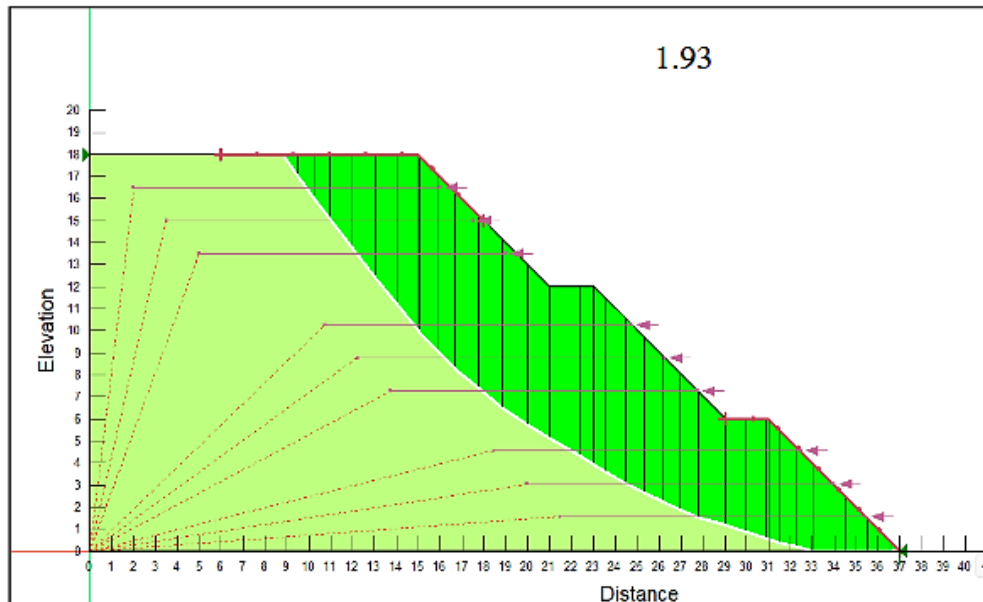


Figure 4. Representative SLOPE/W analysis of a slope reinforced with UX1100 geogrid.

Table 4. Comparison of representative FOS values for unreinforced and reinforced slopes

Height	Angle	Soil type	FOS without geogrid	FOS with geogrid, 0.5 m	Increase in FOS (%)
18 m	45°	Silty clay	1.46	2.15	47.1
18 m	45°	Clayey sand	0.83	1.51	81.8
18 m	45°	Clayey silt	0.95	1.67	75.4
18 m	60°	Silty clay	1.10	2.12	92.2
18 m	60°	Clayey sand	0.58	1.43	145.8
18 m	60°	Clayey silt	0.67	1.58	135.7

The geogrid contributed axial tensile resistance where the reinforcement layers intersected the potential slip surface. This resistance was transferred through soil–geogrid interaction and the available anchorage beyond the sliding mass. The magnitude that could be mobilized was limited by the reduced tensile capacity and pullout resistance. This mechanism is consistent with the improvement observed in previous reinforced-slope studies [8, 9, 18].

The effect differed between the soil parameter sets because the soil controlled both the unreinforced shear resistance and the interaction with the reinforcement. The relatively low initial FOS of clayey sand allowed a high percentage increase after reinforcement, but the resulting FOS was not always sufficient to reach 1.5. This distinction is important when interpreting reinforcement performance.

3.3. Effect of Geogrid Vertical Spacing

The FOS increased as the vertical spacing between the geogrid layers decreased. Among the three spacings investigated, 0.5 m produced the highest FOS, followed by 1.0 and 1.5 m.

This expected trend occurred because the closer spacing provided more reinforcement layers within the same slope height, allowing more layers to intersect the potential slip surface.

Fig. 5 shows the effect of geogrid vertical spacing on FOS for the 18 m slope at 45°. For silty clay, the FOS decreased from 2.15 at 0.5 m spacing to 2.10 and 1.93 at 1.0 and 1.5 m spacing, respectively. For clayey sand, the corresponding values were 1.51, 1.36, and 1.22. Clayey silt produced FOS values of 1.67, 1.52, and 1.37.

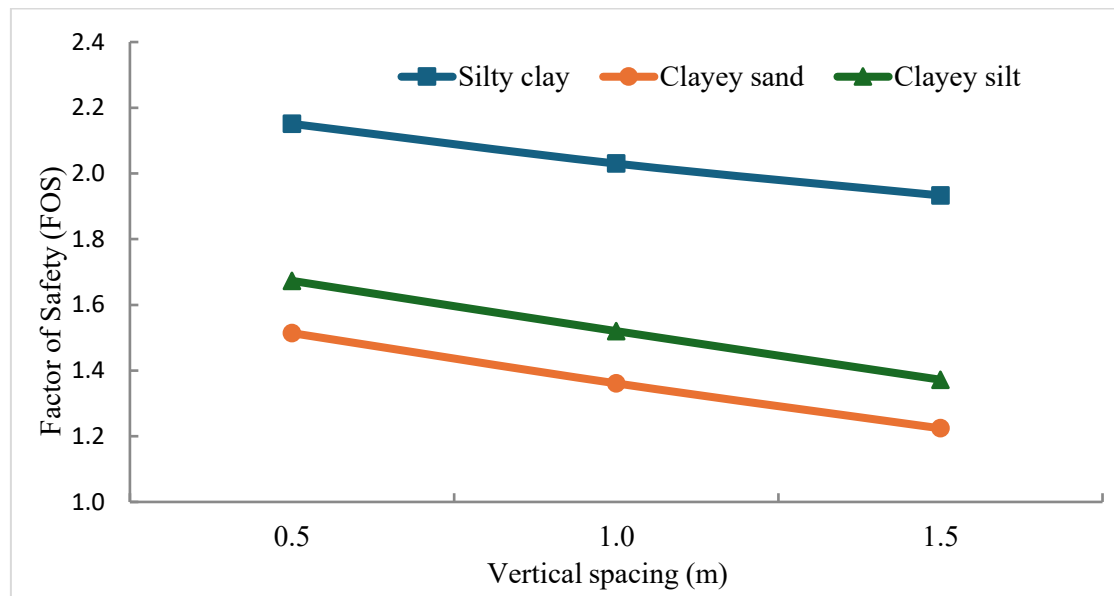


Figure 5. Effect of geogrid vertical spacing on FOS for the 18 m slope at 45°

Spacing also influenced the results at 60°. The silty clay FOS decreased from 2.12 at 0.5 m spacing to 1.68 at 1.0 m and 1.54 at 1.5 m. For clayey sand, the corresponding values were 1.43, 1.06 and 0.94, while clayey silt produced FOS values of 1.58, 1.19 and 1.06.

Although closer spacing increased the FOS, several cases at 60° remained below the reinforced-slope requirement of 1.5. The 0.5 m spacing should therefore be interpreted only as the spacing that produced the highest FOS among the three values investigated. It does not represent an optimum spacing. Determining an optimum would require additional spacings, along with deformation, constructability, and economic considerations.

3.4. Combined Influence of Slope Geometry and Soil Condition

The same reinforcement spacing did not provide the same level of stability for every slope configuration. At 25°, the reinforced slopes generally maintained relatively high FOS values, including the taller configurations. For the 36 m configuration at 0.5 m spacing, the FOS values were 2.46, 1.84, and 2.02 for silty clay, clayey sand, and clayey silt, respectively. All three values exceeded 1.5.

At 60°, substantially lower FOS values were obtained. For the 36 m configuration at the same 0.5 m spacing, the FOS values were 1.49, 1.06, and 1.17 for silty clay, clayey sand, and clayey silt, respectively. None of the three soil parameter sets achieved FOS = 1.5. Thus, closer reinforcement spacing alone was insufficient for this combination of total height and slope inclination. Table 5 compares the reinforced FOS values for the 18 m and 36 m configurations at 60° using the same geogrid spacing of 0.5 m.

Table 5. Reinforced FOS for the 18 m and 36 m configurations at 60° using 0.5 m geogrid spacing

Combined configuration	Soil parameter set	Reinforced FOS
3 lifts; $H = 18\text{m}$	Silty clay	2.12
3 lifts; $H = 18\text{m}$	Clayey sand	1.43
3 lifts; $H = 18\text{m}$	Clayey silt	1.58
6 lifts; $H = 36\text{m}$	Silty clay	1.49
6 lifts; $H = 36\text{m}$	Clayey sand	1.06
6 lifts; $H = 36\text{m}$	Clayey silt	1.17

For silty clay, the reinforced FOS decreased from 2.12 for the 18 m configuration to 1.49 for the 36 m configuration. Clayey sand decreased from 1.43 to 1.06, while clayey silt decreased from 1.58 to 1.17. These comparisons show that reinforcement performance was influenced by the slope geometry and initial soil resistance.

The results should not be interpreted as showing that berms reduce slope stability. The number of slope lifts and total slope height were varied simultaneously. Therefore, the observed reductions cannot be separated into independent berm and height effects. The results only describe the combined slope-height and stepped configurations considered in the study.

3.5. Overall Performance, Limitations and Practical Implications

Fig. 6 presents the variation of reinforced FOS with slope inclination at a vertical spacing of 0.5 m. For the gentler inclinations of 25° and 30°, reinforcement generally produced FOS values above the adopted requirement. As the inclination increased to 45° and 60°, the final FOS became more dependent on the soil parameter set and total slope height. Table 6 summarises the main findings and their implications for assessing geogrid-reinforced slopes.

The study was limited to homogeneous soil parameter sets under dry, drained, and static conditions. The study did not include groundwater, rainfall infiltration, surcharge, seismic loading, drainage performance, or spatial soil variability. The calculated FOS values should therefore not be applied directly to wet, loaded, or site-specific slopes.

The paired reinforced and unreinforced models provided an internal comparison because they maintained the same geometry, soil inputs, and analysis procedure. However, the models were not independently validated using field measurements, laboratory tests, or a published benchmark. The results should therefore be used to describe comparative numerical trends rather than as direct design values.

The results indicate that geogrid spacing should be considered alongside soil conditions and slope geometry. Although the 0.5 m spacing produced higher FOS values than the 1.0 and 1.5 m spacings, it remained insufficient for some of the taller and steeper configurations. This was most evident for the 36 m slope at 60°, where none of the three soil parameter sets achieved $\text{FOS} = 1.5$ at 0.5 m spacing.

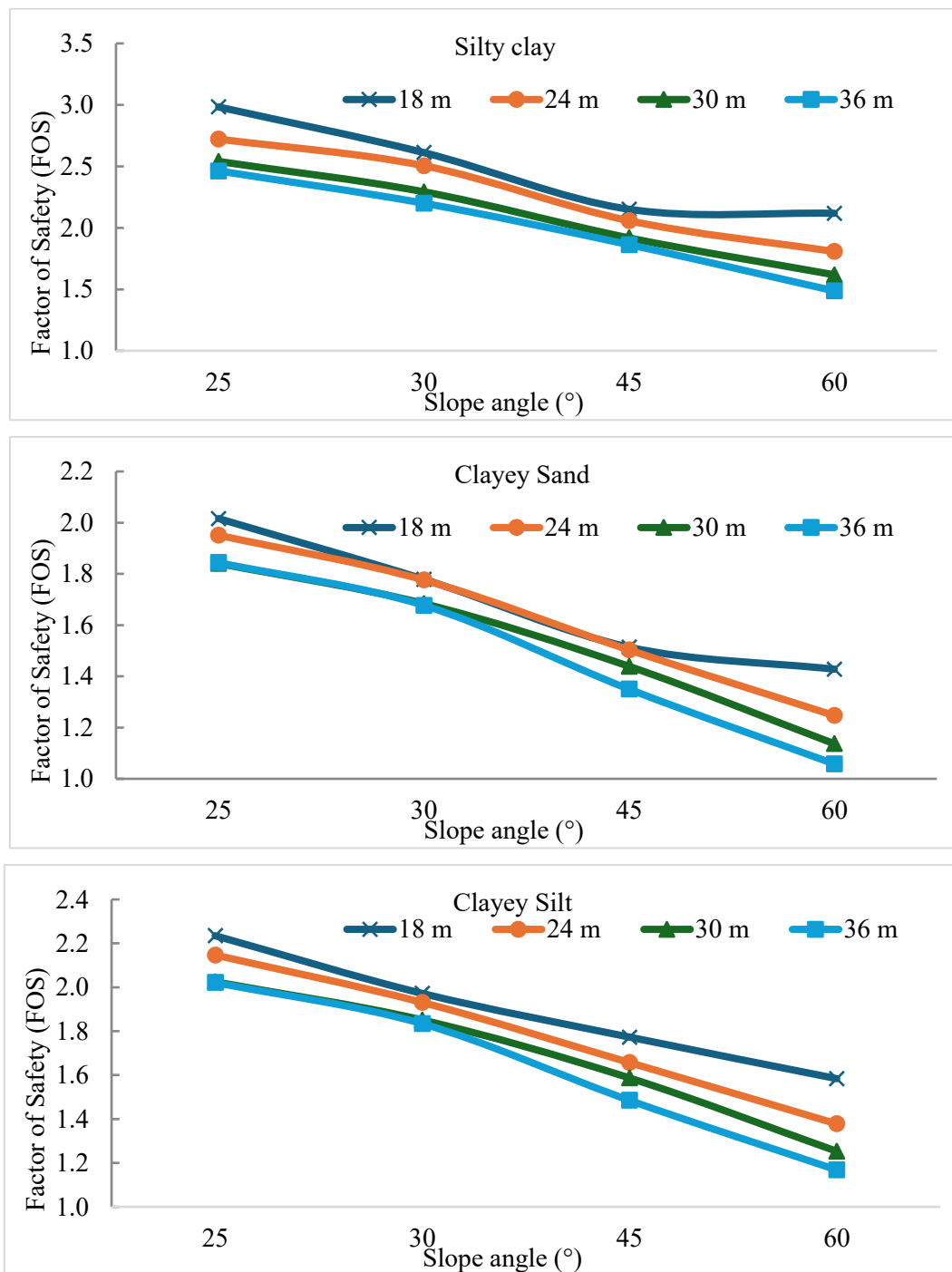


Figure 6. Variation of reinforced FOS with slope inclination at 0.5 m geogrid spacing for different total heights.

Further work should include coupled seepage–stability analysis, surcharge and seismic conditions, soil-property variability, and independent model validation. A constant-height factorial analysis in which berm number and geometry are varied separately is also required to isolate the influence of berm configuration. Additional reinforcement spacings and deformation or economic criteria would be required before an optimum geogrid arrangement could be established.

Table 6. Summary of the main findings and engineering implications

Parameter	Main finding	Engineering implication
Soil parameter set	Silty clay generally produced the highest FOS, while clayey sand produced the lowest	Reinforcement performance depends on the combined soil-strength parameters and stress condition
Slope inclination	FOS decreased as the slope became steeper	Steeper slopes produced a higher driving demand
Geogrid spacing	The 0.5 m spacing produced the highest FOS among the three spacings investigated	The result reflects the larger number of reinforcement layers and is not an optimization result
Combined slope-height and stepped configuration	Taller configurations generally produced lower FOS	The independent effect of berm number cannot be determined from the present analysis
Critical configurations	Reinforcement increased FOS, but some taller and steeper configurations remained below 1.5	An increase in FOS does not necessarily mean that the adopted stability requirement has been achieved

4. CONCLUSIONS

This study evaluated the stability of geogrid-reinforced slopes under different soil and geometric configurations using SLOPE/W. The analyses considered three soil parameter sets, four slope inclinations, four combined slope-height and stepped configurations with total heights ranging from 18 to 36 m, and geogrid vertical spacings of 0.5, 1.0 and 1.5 m. The reinforced and unreinforced slopes were compared under the same dry, drained, and static conditions.

The unreinforced results showed that the FOS generally decreased as the slope inclination and total height increased. Soil condition also influenced the calculated FOS. Silty clay generally produced the highest FOS, while clayey sand produced the lowest. Although clayey sand had a higher friction angle than silty clay, its lower cohesion and higher unit weight contributed to its lower FOS. The results therefore reflected the combined soil parameters rather than the friction angle alone.

Geogrid reinforcement increased the FOS for all the configurations investigated, but the magnitude of improvement varied with the soil parameter set and slope geometry. As expected, the 0.5 m spacing produced the highest FOS among the three spacings because it provided more reinforcement layers within the same slope height. This result should not be interpreted as an optimum spacing. More importantly, higher FOS did not always result in acceptable stability. For the 36 m slope at 60°, none of the three soil parameter sets achieved FOS = 1.5, even at the closest spacing of 0.5 m. This shows that closer reinforcement spacing alone may be insufficient for taller and steeper slope configurations.

The study does not establish the independent influence of berm number because the number of slope lifts and total height increased together. The results represent the combined influence of slope height and stepped configuration. A constant-height analysis with independently varied berm geometry would be required to determine the effect of berm number.

The findings are limited to homogeneous soil parameter sets under dry, drained, and static conditions, with no groundwater, rainfall infiltration, surcharge, or seismic loading. The paired reinforced and unreinforced models provided an internal comparison but were not independently validated using field measurements, laboratory tests, or a published benchmark. Therefore, interpret the findings as comparative numerical trends rather than site-specific

design values. Future work should include independent model validation, wet and loaded slope conditions, constant-height berm configurations, and additional geogrid spacings supported by deformation, constructability, and economic considerations.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Department of Civil Engineering, Kulliyyah of Engineering, International Islamic University Malaysia (IIUM), for providing the facilities and support required for this study.

REFERENCES

- [1] Gajamer V, Kumar A. (2023) A comprehensive review on rainfall-induced slope failures: Mechanism, models, and influencing factors. *Lecture Notes in Civil Engineering*, 303:177–186. https://doi.org/10.1007/978-981-19-7245-4_16
- [2] Lias R, Jais IBM, Lat DC. (2022) Climatic influence on slope failure: A case study at Kem Terendak, Melaka. *International Journal of Sustainable Construction Engineering and Technology*, 13(1):39–49. <https://doi.org/10.30880/IJSCET.2022.13.01.004>
- [3] Kanık M. (2022) Failure mechanism of a soil slope and stabilization method: A case study. *Firat University Journal of Experimental and Computational Engineering*, 1(3):129–138. <https://doi.org/10.5505/FUJECE.2022.35744>
- [4] Shiraliyev N, Khalilova A, Abbasova A. (2024) Geotechnical justification for designing landslide prevention measures. *Nature & Science*, 6(10):28–38. <https://doi.org/10.36719/2707-1146/49/28-38>
- [5] Haque Nayem N. (2023) Enhancement of soil characteristics using different stabilization techniques. *Journal of Civil, Construction and Environmental Engineering*. <https://doi.org/10.11648/J.JCCEE.20230804.12>
- [6] Mizal-Azzmi N, Mohd-Noor N, Jamaludin N. (2011) Geotechnical approaches for slope stabilization in residential area. *Procedia Engineering*, 20:474–482. <https://doi.org/10.1016/J.PROENG.2011.11.190>
- [7] Palmeira EM, Araújo GLS, Santos ECG. (2021) Sustainable solutions with geosynthetics and alternative construction materials—A review. *Sustainability*, 13(22):12756. <https://doi.org/10.3390/SU132212756>
- [8] Idrus J, Hamzah N, Ramli R, Md Nujid M, Sadikon SF. (2023) Enhancing slope stability with different slope stabilization measures: A case study using SLOPE/W software. *Jurnal Kejuruteraan*, 35(6):1427–1434. [https://doi.org/10.17576/JKUKM-2023-35\(6\)-15](https://doi.org/10.17576/JKUKM-2023-35(6)-15)
- [9] Liu H, Zheng J, Zhang R, Xie P. (2022) Probabilistic stability analysis of reinforced soil slope with non-circular RLEM. *Geosynthetics International*, 30(4):1–17. <https://doi.org/10.1680/JGEIN.21.00003>
- [10] Zeng Y, Zhang Y, Hu W, Chen M, Hu Q, Liu X, Zhu X. (2024) A case study on soil slope landslide failure and parameter analysis of influencing factors for safety factor based on strength reduction method and orthogonal experimental design. *PLOS ONE*, 19(5):e0300586. <https://doi.org/10.1371/journal.pone.0300586>
- [11] Basta MM, Rabie MH, Mansour MA, Elbanna WM. (2024) A comparative study between soil nailing and berms for slope stabilization: Performance and analysis. *Engineering Research Journal*, 183(3):65–88. <https://doi.org/10.21608/ERJ.2024.303147.1072>
- [12] Jian-yong S, Sheng L. (2014) Slope stability analysis method considering transfer of sliding failure surface and influence of engineered berm. *Chinese Journal of Geotechnical Engineering*, 36(6):998–1004. <https://doi.org/10.11779/CJGE201406002>

- [13] Elahi TE, Islam MA, Islam MS. (2022) Parametric assessment of soil nailing on the stability of slopes using numerical approach. *Geotechnics*, 2(3):615–634. <https://doi.org/10.3390/GEOTECHNICS2030030>
- [14] Alok A, Burman A, Samui P, Kaloop MR, Eldessouki M. (2024) A generalized limit equilibrium-based platform incorporating simplified Bishop, Janbu and Morgenstern–Price methods for soil slope stability problems. *Advances in Civil Engineering*, 2024:3053923. <https://doi.org/10.1155/2024/3053923>
- [15] Slope Engineering Branch, Public Works Department Malaysia. (2010) Guidelines for Slope Design. JKR 21500-0011-10. Jabatan Kerja Raya Malaysia.
- [16] Marto A, Yusoff SYM. (2017) Major soil type, soil classification, and soil maps. In: *Soils of Malaysia*, pp. 69–101. CRC Press. <https://doi.org/10.1201/B21934-4>.
- [17] Salmasi F, Abraham J, Nourani B. (2022) Determining the analysis of the stability of embankments against sliding and prediction of sliding and critical factor of safety. In: *Novel Perspectives of Engineering Research*, Vol. 9, pp. 98–125. <https://doi.org/10.9734/BPI/NPER/V9/2028A>
- [18] Zhang L, Peng B, Zhou S, Cui P, Liu Y. (2024) Numerical study on stability of geosynthetic-encased stone column-supported embankments based on equivalent methods. *Computers and Geotechnics*, 169:106179. <https://doi.org/10.1016/j.compgeo.2024.106179>

EEG Signal Classification under RF Exposure Using Synthetic Data–Augmented Feedforward Neural Network

MUHAMMAD RASYID ROSLI, ROSHAKIMAH MOHD ISA*,
NUR IRSALINA HUDA NAZRI

*Faculty of Mechanical Engineering, Universiti Teknologi MARA (UiTM),
Shah Alam, 40450 Selangor, Malaysia*

**Corresponding author: roshakimah@uitm.edu.my*

(Received: 04 December 2025; Accepted: 4 June 2026; Published online: 11 September 2026)

ABSTRACT: Electroencephalogram (EEG) signal classification plays a critical role in understanding brain activity under varying radiofrequency (RF) exposure conditions. The demand for improving EEG classification modeling is increasing due to the growing number of gadgets emitting RF exposure. Current Electroencephalogram (EEG) classification methods struggle to process signals accurately because of insufficient data and limited variety in the datasets used to train models. This study aims to improve Electroencephalogram (EEG) signal classification using synthetic data and a Feedforward Neural Network (FNN). The goal is to analyze brain signals under different RF exposure conditions and increase classification accuracy using asymmetry ratio analysis and synthetic data for 3 exposure groups. Conditional Generative Adversarial Networks (CTGAN) are employed to generate 88 synthetic samples from an original dataset of 97 EEG recordings, stratified by exposure type and gender. The two sub-bands from the original data were used to generate synthetic data, divided into 3 groups: Left Exposure (LE), Right Exposure (RE), and Sham Exposure (SE). Each group was further split into male and female categories. Statistically significant variations in alpha- and beta-band activity are observed during RF exposure, validating the discriminative capability of the extracted features. Including synthetic data significantly improves classification performance, particularly during male exposure sessions, where the Feedforward Neural Network (FNN) achieves 98% training accuracy. The findings demonstrate that synthetic data augmentation is an effective strategy for improving Electroencephalogram (EEG) classification reliability and offers strong potential for radiofrequency (RF) exposure analysis and related biomedical applications.

ABSTRAK: Kajian ini bertujuan menambah baik klasifikasi isyarat Elektroensefalogram (EEG) bagi aktiviti otak di bawah pendedahan frekuensi radio (RF) menggunakan data sintetik dan Rangkaian Nural Suap-Depan (FNN). Isu utama adalah kekurangan data dalam set data EEG sedia ada. Kami menggunakan analisis nisbah asimetri bagi mengekstrak ciri dan Analisis Varians (ANOVA) dan membandingkan gelombang alfa dan beta antara lelaki dan wanita dalam tiga kumpulan pendedahan: Kiri (LE), Kanan (RE), dan Palsu (SE). Rangkaian Berseteru Generatif Bersyarat (CTGAN) digunakan bagi menjana 88 data sintetik, meningkatkan set data asal 97 sampel. Penggunaan data sintetik meningkatkan ketepatan klasifikasi model, terutama untuk sesi lelaki, mencapai 98% ketepatan latihan dan 69% ketepatan ujian. Perubahan signifikan dalam gelombang alfa dan beta diperhatikan semasa pendedahan RF. Penemuan ini menunjukkan bahawa data sintetik adalah alat berkesan bagi meningkatkan klasifikasi EEG dalam kajian kesan pendedahan RF.

KEYWORDS: *Neural Network, EEG, CTGAN, Synthetic Data, Classification*

1. INTRODUCTION

Concerns about radiofrequency (RF) exposure devices have raised awareness and sparked interest in investigating their effects on brainwave signal processing. Determining these effects on human brain activity requires researchers to analyze Electroencephalogram (EEG) signal data using advanced analysis techniques. Artificial Neural Networks (ANNs) are large machine learning models developed to study human brain activity and characteristics [1]. Feedforward Neural Networks (FNNs) are among them; they process information in one direction, with no feedback from the output to the input layer, and are used for function approximation, classification, and regression analysis [1, 2]. FNNs, however, lack accuracy with limited data [3-6]. This analysis has led to the introduction of Conditional Generative Adversarial Networks (CTGAN) [7-9]. CTGAN produces more consistent data by eliminating noise and outliers, which enhances statistical analysis by reducing ambiguity and making it easier to identify statistically significant differences [10-12].

EEG signals include frequency bands such as Delta (0–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), and Beta (12–30 Hz). Brain activity can be analyzed in both the Delta and Theta frequency bands, but these bands are generally more susceptible to noise and may be related to pathological conditions. Thus, the alpha and beta frequency bands may be preferred for analyzing brain activity [13-15]. It is crucial to apply significant data pre-processing and feature extraction methods to ensure the accuracy and reliability of the modeling [16]. In this analysis, the study used techniques such as Power Spectral Density (PSD) analysis, normalization, Power Asymmetry Ratio (PAR), Spearman rank correlation, and One-Way ANOVA, among others, to refine and interpret neural activity in the alpha and beta frequency bands. This analysis provides statistical data and includes a data assessment performance prior to classification, allowing for more accurate quantification and interpretation of neural activity.

The frequency content of EEG signals was measured using the PSD technique, which illustrates power distribution across frequencies. PSD is important for identifying how these bands relate to different cognitive states, such as resting, conscious, and deep sleep [17, 18]. Max normalization transforms input data into a fixed range, standardizing the input and improving model consistency, robustness, and performance [19-21]. The PAR correlates left- and right-brain power, which helps explain how the brain lateralizes across different cognitive states [22, 23]. Spearman's rank correlation validated both the original and synthetic EEG data, ensuring statistical equivalence, which is crucial for reliability in clinical applications and for imbalanced brainwave classification and modeling tasks [24-27]. One-way ANOVA helps investigate significant differences in EEG-related features across conditions, supporting classification models by discriminating between groups or sessions based on cognitive tasks or brain stimuli [28, 29].

This study included 97 participants, including males and females, collected through three observation sessions, focusing primarily on higher-frequency bands in left- and right-brainwave data. The EEG signals were obtained from previously collected data, where the original sample consisted of 37 male and 60 female volunteers aged 19 to 30 years from engineering students at Universiti Teknologi MARA, Malaysia [30]. The data were pre-processed; synthetic data were generated using CTGAN, features were extracted, and classification was performed with an FNN. The signal processing and modeling were conducted using Python, MATLAB, and SPSS statistical tests. This could provide a novel feature extraction method using asymmetry, potentially further contributing to a constructed understanding of RF radiation interference on brainwave activity. The study aimed to contribute to science and society discussions, supporting neuroscience and public health solutions.

2. METHODOLOGY

In the present study, data were collected from 97 participants, who were subsequently subdivided into the Left Exposure (LE), Right Exposure (RE), and Sham Exposure (SE) groups. In the LE and RE conditions, participants used a mobile telephone strapped to their left or right ear, respectively, while participants in the SE group served as a control and had the mobile phones turned off. Fig. 1 shows A subject undergoing the EEG experiment with a strapped phone at the right ear. To maintain data validity, all participants were healthy, stable, and not under the influence of medication during data collection. The first preprocessing step was to convert the time-domain signal to the frequency domain using the Fast Fourier Transform (FFT) on the recorded EEG data. Next, the collected data were split by gender across all groups to identify gender-based differences. Boxplots by gender were then compared, with all outliers recorded for each group, and synthetic data were created to enhance the actual dataset.



Figure 1. A subject undergoing the EEG experiment with a strapped phone on the right ear.

Table 1. CTGAN parameter tuning.

Parameter	Value
Epochs number	100 to 1000
Data Batch Size	500
Generator Dimension	(512, 512)
Discriminator Dimension	(512, 512)
Generator Learning Rate	0.0001
Discriminator Learning Rate	0.0001
Discriminator Steps	5
CUDA (GPU Support)	True

The synthetic data were generated by applying CTGAN to the real EEG values in Python using Google Colaboratory. The generated synthetic data were run through controlled conditions to identify negative values and verify that the 'min' and 'max' values varied by ± 0.05 from the original PSD values. To achieve the best data quality, epochs and batch sizes were further tuned. Spearman correlation was subsequently used to determine how closely the synthetic data mimicked the original EEG signals. Values between 0.7 and 1.0 were accepted as strong indicators of similarity under all experimental conditions [31]. The CTGAN parameters were adapted from previous investigations and adjusted for the unique nature of the

PSD dataset across all three groups and for male and female data. The parameters were chosen to introduce complexity while remaining uniform across datasets. Table 1 presents the CTGAN model parameters.

The generated synthetic data were tested using a one-way ANOVA to investigate statistical differences between the original, synthetic, and combined datasets. As with the original data, normalization, feature extraction, and boxplot analysis techniques were applied to the generated dataset to identify unique patterns and similarities. This was significant in further establishing the dataset's reliability, which forms the basis of the results in the following interpretation phase of the study and the later modeling phase. All data, original and generated, were normalized using Eq. (1),

$$X_{new} = \frac{X}{x_{max}} \quad (1)$$

which was used to analyze the entire dataset. The next data analysis step was feature extraction using symmetry, calculated by Eq. (2), where PR is the right-hemisphere signal, and PL is the left-hemisphere signal:

$$PAR = \frac{PR-PL}{PR+PL} \quad (2)$$

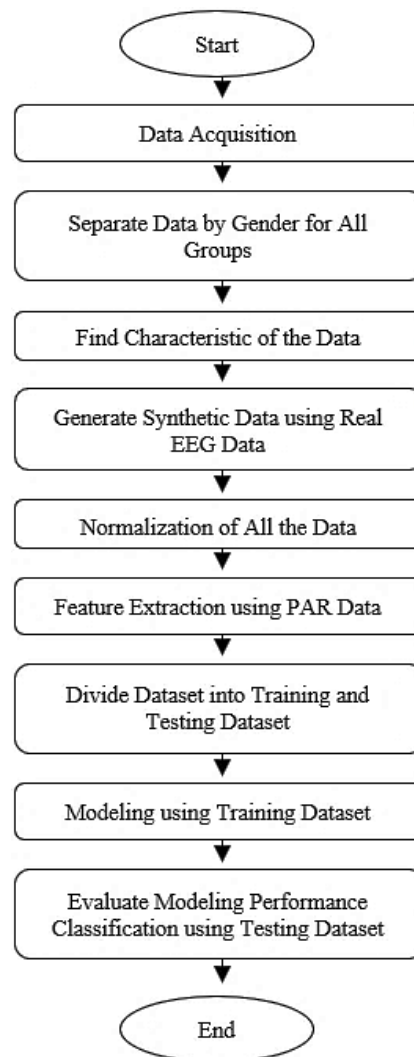


Figure 2. Process flow of the ANN model and classification for original and synthetic EEG datasets.

Boxplots were used to analyze the data, identify patterns and trends, and determine differences between participants and exposure types. Next, the data were separated into training and testing data. The training set was used to build the model in MATLAB 2023a, and the testing set was used to evaluate the model. The model was intended to classify gender from EEG data, focusing on exposure-type differences. Lastly, the model was evaluated on the testing dataset. This evaluation aimed to assess the model's likely accuracy and versatility. To test the model's effectiveness in identifying patterns and making correct classifications, we analyzed the evaluation results. According to the outlined research methodology, Fig. 2 illustrates a clear, simplified process for the ANN model and classification for the original and synthetic EEG datasets.

3. RESULTS AND DISCUSSION

3.1. Normalized PSD EEG Signal Pattern

This section provides a visual comparison of normalized Power Spectral Density (PSD) between original and synthetic EEG data to evaluate the quality of data generated by CTGAN. Radar charts illustrate Alpha and Beta patterns for male and female participants, categorized by exposure type (LE, RE, SE) and session (Before Right, Before Left, During Right, During Left). Solid lines indicate real data, while dashed lines show synthetic data, as shown in Fig. 3. These visualizations help assess the pattern similarity and reliability of the synthetic data.

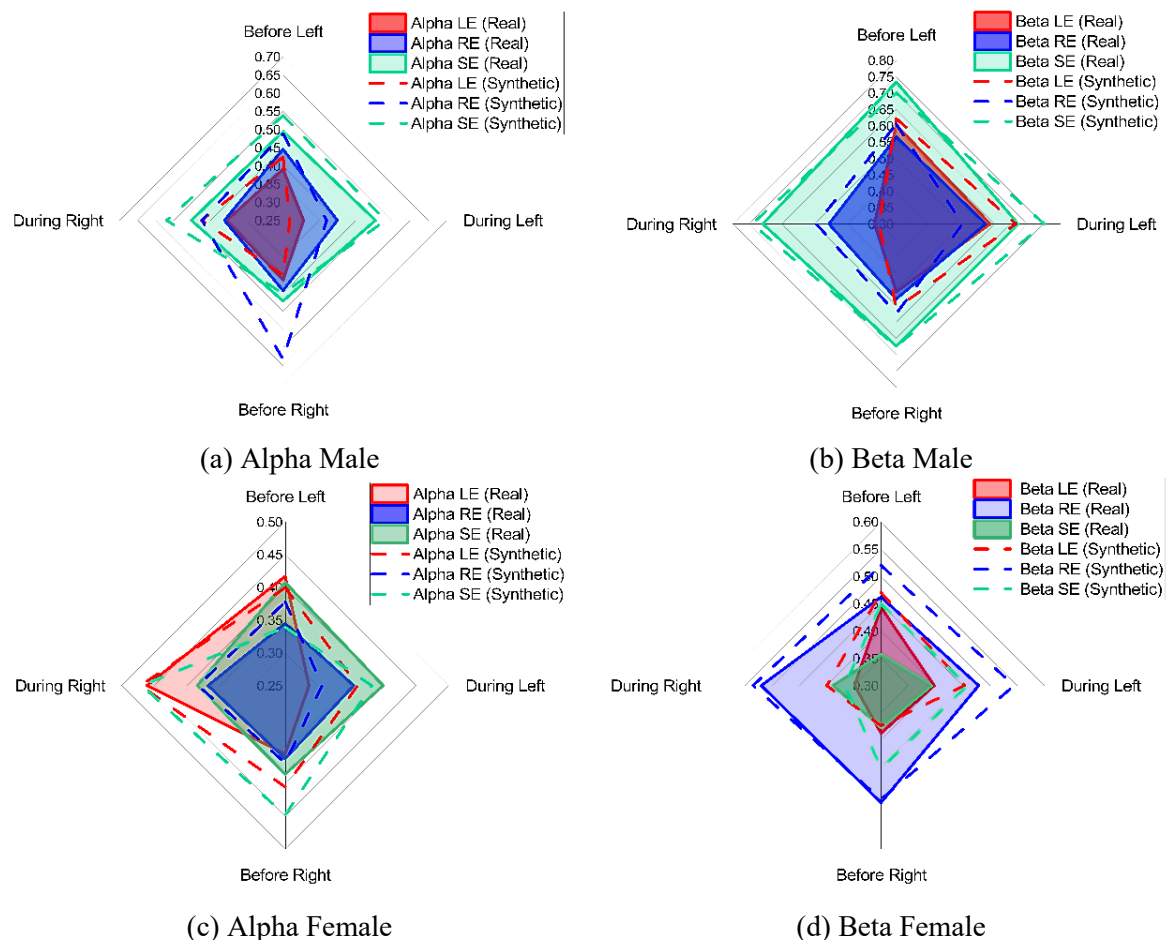


Figure 3. Normalized PSD for original and synthetic EEG data for (a) Alpha Male, (b) Beta Male, (c) Alpha Female, and (d) Beta Female.

Fig. 3 presents the radar chart for the synthetic EEG data, which shows a pattern comparable to the real data across both Alpha and Beta bands. For the male EEG signals, the Beta results show strong similarity, especially in the LE and RE groups during all sessions, indicating that the CTGAN-generated data can follow the trend of the original EEG signals. For Alpha in the male data, the pattern is generally consistent, although minor differences are observed in the LE group in the Before Right session, but still within an acceptable range.

To further validate the similarity between the real and synthetic EEG datasets, we conducted Spearman correlation analysis. Table 2 illustrates the correlation coefficients for each session, frequency band, and gender distribution. This analysis confirms the statistical similarity between the two datasets. The correlation coefficients range from 0.719 to 0.829, indicating strong to very strong relationships and supporting the reliability of the CTGAN model in replicating the statistical behavior of the EEG signals. These correlation ranges are consistent with previous EEG studies involving RF exposure and asymmetry analysis, where such levels were accepted as statistically meaningful indicators of brainwave similarity [31, 32].

Table 2. Spearman correlation coefficient between real and synthetic EEG data.

Exposure session	Brainwave	Gender	Correlation Coefficient	Strength of Correlation
Before RF	Beta	Male	0.768	Strong
		Female	0.791	Strong
	Alpha	Male	0.793	Strong
		Female	0.719	Strong
During RF	Beta	Male	0.778	Strong
		Female	0.829	Very Strong
	Alpha	Male	0.817	Very Strong
		Female	0.794	Strong

3.2. Statistical Evaluation of EEG Signal Distribution Using PAR and ANOVA

EEG datasets comprising original and generated data were analyzed using the PAR to examine the distribution of Alpha and Beta wave activity across different exposure groups. The analysis focused on LE, RE, and SE in the Before and During RF exposure sessions for both male and female EEG signals. Median PAR values were used to generate the bar charts in Fig. 4, illustrating variations in signal asymmetry under different RF exposure conditions.

Beta wave activity in male EEG signals increased significantly in the LE group “During” the exposure session, rising from 0.10 to 0.27, while the RE and SE groups remained relatively stable. A similar pattern was observed in the Alpha wave, where LE values increased to 0.28 “During” exposure. In female EEG signals, Alpha wave activity in the LE group rose sharply from 0.13 to 0.31, indicating stronger hemispheric asymmetry during RF exposure. These variations suggest that RF exposure has a more pronounced impact on EEG asymmetry, particularly under the LE condition.

A one-way ANOVA test was conducted on the combined dataset to evaluate whether Alpha and Beta brainwave activities differed significantly across exposure groups. The hypotheses tested were that RF exposure would not affect EEG asymmetry (null hypothesis) versus the alternative that it would. Table 3 presents the results. Significant differences ($p < 0.05$) were observed in Beta wave activity for male EEG signals during both the before and during exposure sessions. Additionally, Alpha wave activity showed statistically significant changes during exposure for both male and female EEG signals. These findings indicate that RF exposure measurably affects EEG asymmetry, particularly in the Beta and Alpha bands during active exposure. The absence of significant differences in Alpha waves prior to exposure

also suggests that the baseline EEG patterns across the exposure groups were relatively stable and unaffected by external variables.

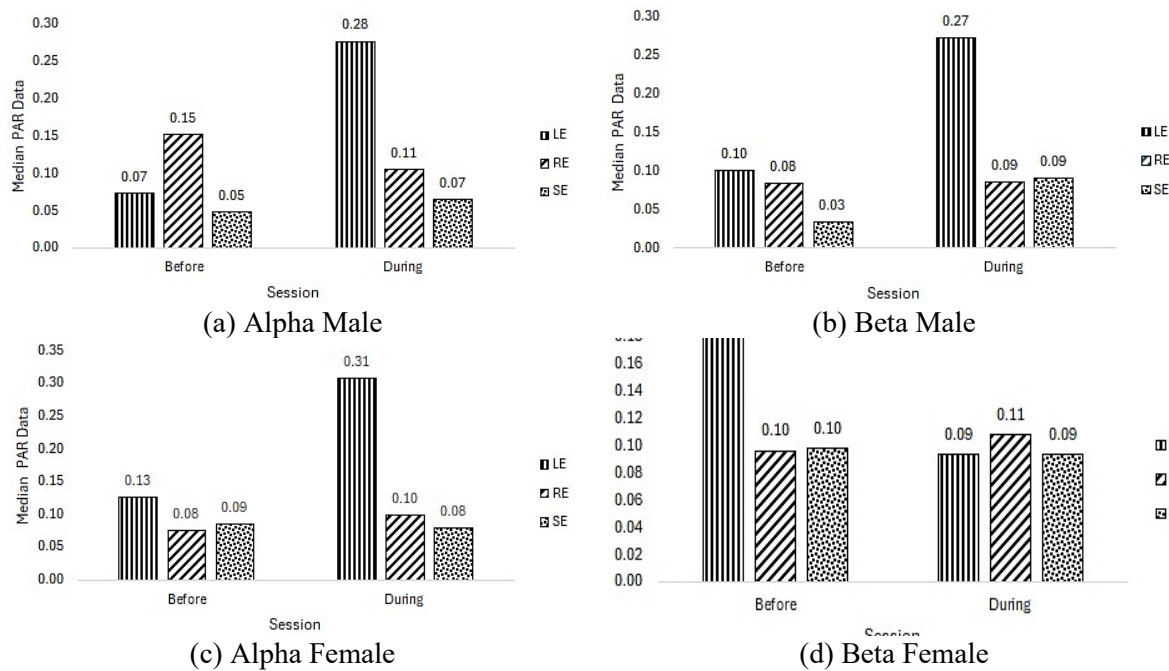


Figure 4. Median Power Asymmetry Ratio (PAR) for Alpha and Beta in LE, RE, and SE exposure groups for (a) Alpha Male, (b) Beta Male, (c) Alpha Female, and (d) Beta Female.

Table 3. One-way ANOVA test using original and synthetic EEG datasets.

Session	Band Wave	Male (p-value)	Female (p-value)
Before	Beta	0.017	0.058
	Alpha	0.115	0.907
During	Beta	0.008	0.455
	Alpha	0.02	0.012

The statistically significant changes in Alpha and Beta asymmetry are favorable signal features that can be used as classification inputs. Because FNNs rely on different and distinctive patterns to make accurate predictions, the significant changes produced by RF exposure will provide a greater opportunity for the model's discriminative ability. Significant and reliable changes in asymmetry, in particular Beta for males and Alpha during the exposure phase, will act as signal features allowing for improvements in classification. This highlights a unique aspect of the overall combined dataset (real and synthetic) for meaningful feature extraction and the generalizing properties of the FNN-based approach.

3.3. FNN Classification Performance for Combined EEG Dataset

Classification results of the Feedforward Neural Network (FNN) are presented using EEG data that combines original and synthetic sources. The model was evaluated based on two RF exposure sessions: Before and During, for both male and female EEG signals. Group classifications were made across LE, RE, and SE. Model performance was assessed through training and testing phases, with accuracy, precision, recall, and F1-score as key performance metrics. Table 4 lists the training parameters used to develop the model for each session and gender, including input nodes, hidden layer size, learning rate, momentum, and epoch count. The parameter tuning was conducted separately for Before male, Before female, During male, and During female.

Table 4. Optimized FNN parameter tuning for each session and gender.

Parameters Tuning	Session			
	Before		During	
	Male	Female	Male	Female
Input Nodes	2.0	2.0	2.0	2.0
Hidden Layer	13.0	18.0	13.0	13.0
Output Nodes	1.0	1.0	1.0	1.0
Learning rate	0.2	0.1	0.3	0.3
Momentum rate	0.2	0.2	0.2	0.6
Epochs	500.0	500.0	500.0	500.0

The parameters listed in Table 4 were obtained through a grid-based tuning process in MATLAB, in which combinations of learning rate, momentum, and hidden layer size were tested to evaluate their impact on model performance. During each run, the network was trained using the Levenberg-Marquardt algorithm, and the mean squared error (MSE) was recorded as the primary evaluation metric. We selected the configuration that produced the lowest MSE during training as the optimal parameter set for each session and gender. This approach ensured that the network generalized well and learned asymmetrical features stably from the EEG data. The selected parameters improved convergence during training and enhanced classification reliability during testing.

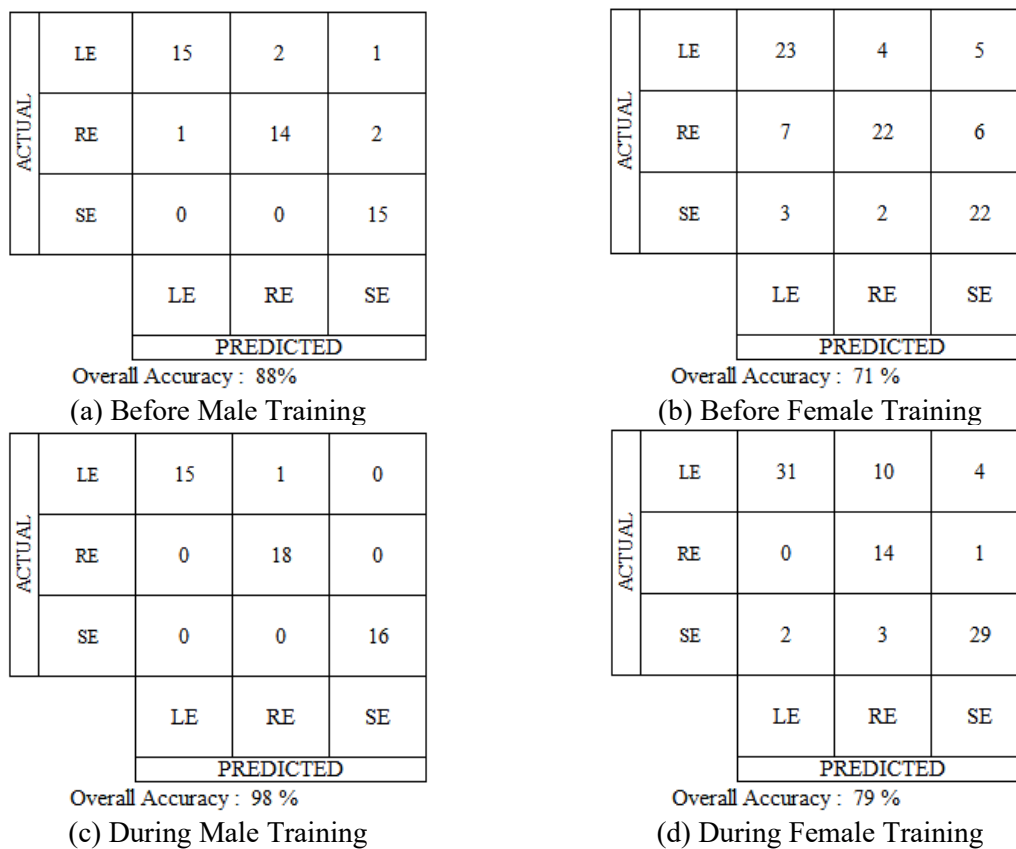


Figure 5. Confusion matrices for FNN training classification: (a) Before Male Training, (b) Before Female Training, (c) During Male Training, (d) During Female Training.

Fig. 5 shows confusion matrices for the training model, with actual and predicted values, while Table 5 summarizes detailed training performance metrics. During the male session, the

FNN model achieved outstanding training results across all exposure groups, with 100% accuracy for LE, 98% for RE, and 98% for SE. This strong performance was supported by equally high precision and recall values, particularly in the SE group, where both reached 100%, resulting in a perfect F1-score. The consistently high metrics across all exposures demonstrate that the model was highly effective at learning and distinguishing male EEG signal patterns during RF exposure.

In contrast, the female session showed slightly lower training accuracy, with 89% for LE, 85% for RE, and 83% for SE. Notably, the RE group had the weakest F1-score (67%), primarily due to a drop in recall (52%), indicating less reliable identification of this group during training. For the Before sessions, male data again outperformed female data, with training accuracy ranging from 90% to 94% and F1-scores between 83% and 88%. Meanwhile, female performance in the Before session ranged from 80% to 83% accuracy, with relatively moderate F1-scores between 70% and 73%.

These results suggest that the FNN learned more effectively from male EEG features, especially during RF exposure, and that the SE group was easier to classify. This was likely due to clearer no-exposure patterns in the SE group, compared with the more complex effects seen in the LE and RE conditions. This pattern supports earlier statistical findings showing stronger group differentiation in male EEG data when comparing all exposure groups.

Table 5. FNN training performance across exposure groups, sessions, and gender.

Session	Gender	Performance Measure	Group Exposure		
			LE	RE	SE
Before	Male	Training Accuracy (%)	94	90	92
		Training Precision (%)	83	82	100
		Training Recall (%)	94	88	83
		Training F1-score (%)	88	85	91
	Female	Training Accuracy (%)	83	80	80
		Training Precision (%)	72	63	81
		Training Recall (%)	70	79	67
		Training F1-score (%)	71	70	73
During	Male	Training Accuracy (%)	100	98	98
		Training Precision (%)	94	100	100
		Training Recall (%)	100	95	100
		Training F1-score (%)	97	97	100
	Female	Training Accuracy (%)	89	85	83
		Training Precision (%)	69	93	85
		Training Recall (%)	94	52	85
		Training F1-score (%)	79	67	85

Fig. 6 presents the confusion matrices, and Table 6 shows the full set of testing metrics. During the male session, the FNN model demonstrated strong testing performance across all exposure groups, with the highest classification accuracy recorded for the Left Exposure (LE) group at 92 percent, followed by Right Exposure (RE) at 77 percent, and Sham Exposure (SE) at 69 percent. These results indicate that the model could effectively differentiate between RF exposure conditions in male EEG signals, particularly in the LE group where separability was most distinct.

For the female session, the classification accuracy was lower across all exposures, with LE and RE both at 74 percent, and SE at 65 percent. While LE showed the strongest results among female exposures, the similarity in performance between RE and SE suggests reduced feature

contrast in the EEG signals. This outcome can be attributed to the stronger signal asymmetry patterns observed in male EEG during RF exposure, as supported by the One-Way ANOVA results, which reported statistically significant differences across exposure groups in the male dataset. In contrast, the female EEG signals showed weaker group differentiation, reflected by higher p-values and less variation in the PAR value.

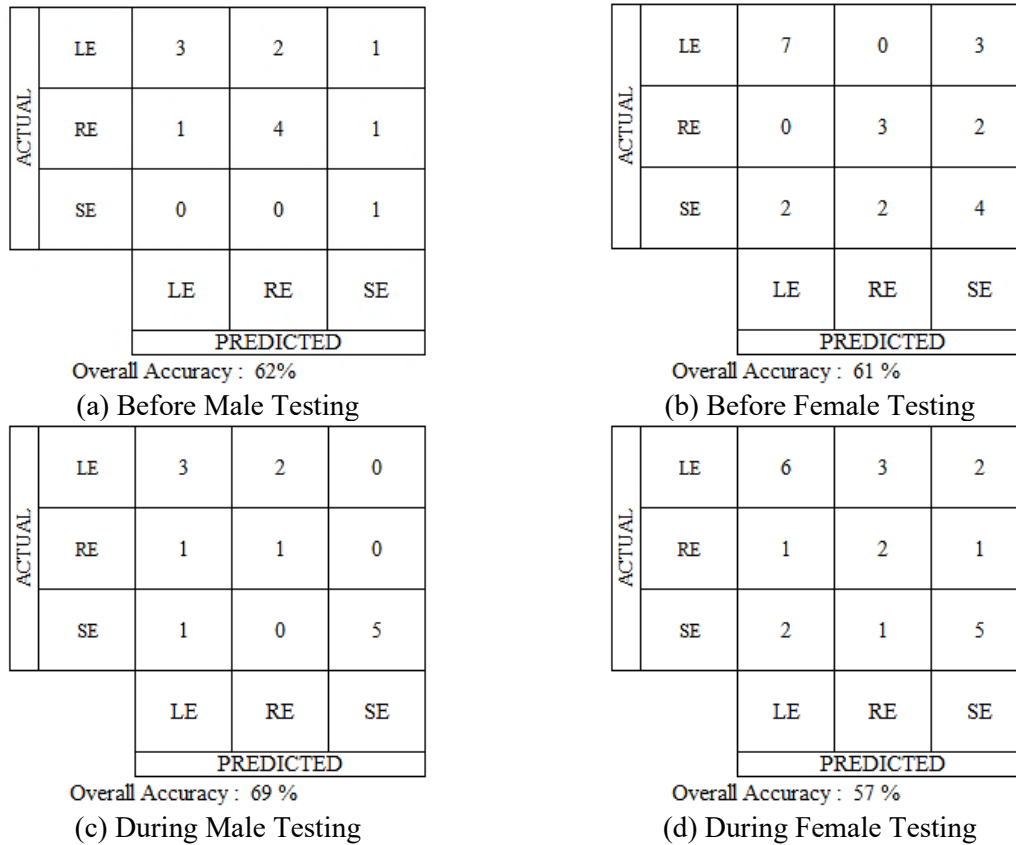


Figure 6. Confusion matrices for FNN training classification: (a) Before Male Testing, (b) Before Female Testing, (c) During Male Testing, (d) During Female Testing.

Table 6. FNN testing performance across exposure groups, sessions, and gender.

Session	Gender	Performance Measure	Group Exposure		
			LE	RE	SE
Before	Male	Testing Accuracy (%)	85	69	69
		Testing Precision (%)	50	67	100
		Testing Recall (%)	75	67	33
		Testing F1-score (%)	60	67	50
	Female	Testing Accuracy (%)	61	83	78
		Testing Precision (%)	70	60	50
		Testing Recall (%)	78	60	44
		Testing F1-score (%)	74	60	47
During	Male	Testing Accuracy (%)	92	77	69
		Testing Precision (%)	60	50	83
		Testing Recall (%)	60	33	100
		Testing F1-score (%)	60	40	91
	Female	Testing Accuracy (%)	74	74	65
		Testing Precision (%)	55	50	63
		Testing Recall (%)	67	33	63
		Testing F1-score (%)	60	40	63

To illustrate the significant impact of combining synthetic and original data on model performance, Table 7 shows training and testing accuracy across different sessions and data types. This table highlights the improvements achieved by incorporating synthetic data into the training process, particularly for male EEG signals. The data clearly demonstrate that combining synthetic data with original EEG signals led to substantial improvements in male classification performance. Specifically, the training accuracy for male EEG increased from 92 percent (Original Data) to 88 percent (Combined Data) in the Before session, and testing accuracy improved from 43 percent to 62 percent. The most significant performance boost was observed in the During session, where testing accuracy for male EEG jumped from 43 percent (Original Data) to 69 percent (Combined Data). Before synthetic data was incorporated, the model exhibited signs of underfitting, with low testing accuracy, especially in the Before exposure session, where it struggled to generalize male EEG data effectively.

Table 7. Comparison of training and testing accuracy using combined and original data.

		Overall Model Accuracy			
Type of Data		Combined Data		Original Data	
		Training	Testing	Training	Testing
Before	Male	88	62	92	43
	Female	71	61	79	25
During	Male	98	69	96	43
	Female	79	57	87	42

These improvements confirm that synthetic data not only enhanced training stability but also helped the model generalize better to previously unseen data. In contrast, female EEG classification showed smaller improvements, consistent with the lower performance observed in the earlier average metrics. This is reflected in the testing accuracy, which remained relatively unchanged (around 57 percent) even with combined data, supporting the conclusion that female EEG signals had less pronounced group separability.

The results align with the One-Way ANOVA, where males showed statistically significant differences ($p < 0.05$) when comparing sessions (Before and During), especially for the During male session, which showed much clearer separability among the exposure groups and contributed to better classification performance. The results for female EEG were far less significant; for Beta, the p-value was 0.057, and for Alpha, it was 0.907 before exposure. This helps to explain the poor classification performance seen in the female data.

The Spearman correlation, on the other hand, showed that the synthetic and real EEG data still had strong similarity, with correlation coefficients ranging from 0.719 to 0.829. This ensured that the synthetic data used to train the classifier were representative of the original EEG signals, which helped contribute to the performance seen in training and testing. The similarity between real and synthetic data showed the model's ability to generalize, as performance was better in testing, particularly for male EEG.

4. CONCLUSION

This research demonstrates the successful application of Feedforward Neural Networks (FNN) to classify EEG signals across various types of radio frequency (RF) exposure conditions using both original and synthetic data. It shows how synthetic EEG signals obtained from a CTGAN improved dataset size and diversity, enabling more stable and reliable training performance. Using PAR as the main feature showed differences in EEG activity between exposure types, most pronounced in the male EEG data during their exposure session. For

classification, the FNN classifier achieved its highest testing accuracy (69%) during the male session, with LE outperforming each exposure condition across all metrics. Female EEG data, especially in the During session, showed poorer classification accuracy and less separability among exposure conditions. This outcome aligns with the One-Way ANOVA results, which also showed statistically significant differences primarily attributed to male EEG signals, strengthening the notion that the model's stronger performance was primarily due to those data. Moreover, the Spearman correlation analysis supported the conclusion that the synthetic data were strongly similar to the original data, which justified their use to improve training and performance.

The training performance also supported that the model learned patterns related to exposures, as it reported the highest training accuracy during the male session (100% LE group). As expected, the SE group across both genders was easier to classify, as there was less RF influence, resulting in clearer, more readable signal patterns. From an engineering perspective, this work demonstrates the advantages of combining deep learning with statistical validation and synthetic data augmentation for biological signal classification challenges. This will not only improve the model's generalization but may also enable this research to impact other engineering problems involving detecting differences in signal patterns, including applications in predictive maintenance, fault detection, and health monitoring systems.

Ultimately, this work does have some limitations. The dataset, although improved by synthesizing data, remains relatively small, and female EEG signals lack feature separability with the methods used in this study. Moreover, although the structural component characteristics were preserved in the synthetic data, they cannot replicate the density and complexity seen in real biological signals. Future work should include obtaining better, larger, and more diverse data, investigating more advanced neural network architectures, and exploring the potential long-term effects of RF exposure. In addition, the ability to use this framework to build on other physiological signals and engineering systems may create opportunities for new innovations in biomedical and industrial disciplines.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the staff and students of Universiti Teknologi MARA, Malaysia, for their valuable contributions and active participation in this research.

REFERENCES

- [1] Liang Y, Jiang J, Chen Y, Zhu R, Lu C, Wang Z (2019). Optimized feedforward neural network training for efficient Brillouin frequency shift retrieval in fiber. *IEEE Access*, 7: 68034–68042. <https://doi.org/10.1109/ACCESS.2019.2919138>
- [2] Ojha VK, Abraham A, Snasel V (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Eng Appl Artif Intell*, 60: 97–116. <https://doi.org/10.1016/j.engappai.2017.01.013>
- [3] Nemzer LR, Cravens GD, Worth RM, Motta F, Placzek A, Castro V, Lou JQ (2021). Critical and ictal phases in simulated EEG signals on a small-world network. *Frontiers in Computational Neuroscience*, 14: 1–13. <https://doi.org/10.3389/fncom.2020.583350>
- [4] Xu Y, Yang J, Sawan M (2022). Multichannel synthetic preictal EEG signals to enhance the prediction of epileptic seizures. *IEEE Transactions on Biomedical Engineering*, 69(12): 3572–3583. <https://doi.org/10.1109/TBME.2022.3171982>
- [5] Leon J, Escobar JJ, Ortiz A, Ortega J, González J, Martín-Smith P, Gan JQ (2020). Deep learning for EEG-based motor imagery classification: Accuracy–cost trade-off. *PLoS One*, 15(6): 1–23. <https://doi.org/10.1371/journal.pone.0234178>

- [6] Ding W, Liu A, Guan L, Chen X (2024). A novel data augmentation approach using mask encoding for deep learning-based asynchronous SSVEP-BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32: 875–886. <https://doi.org/10.1109/TNSRE.2024.3366930>
- [7] Cekic M (2023). Anomaly detection in medical time series with generative adversarial networks: A selective review. In: *Anomaly Detection – Recent Advances, AI and ML Perspectives and Applications*. IntechOpen, 1–18. <https://doi.org/10.5772/intechopen.112582>
- [8] Carrle FP, Hollenbenders Y, Reichenbach A (2023). Generation of synthetic EEG data for training algorithms supporting the diagnosis of major depressive disorder. *Frontiers in Neuroscience*, 17: 1–11. <https://doi.org/10.3389/fnins.2023.1219133>
- [9] Habashi AG, Azab AM, Eldawlatly S, Aly GM (2023). Generative adversarial networks in EEG analysis: An overview. *Journal of NeuroEngineering and Rehabilitation*, 20(1): 1–16. <https://doi.org/10.1186/s12984-023-01169-w>
- [10] Yadav P, Gaur M, Madhukar RK, Verma G, Kumar P, Fatima N, Sarwar S, Dwivedi YR (2024). Rigorous experimental analysis of tabular data generated using TVAE and CTGAN. *International Journal of Advanced Computer Science & Applications*, 15(4): 1250–1262. <https://doi.org/10.14569/IJACSA.2024.01504125>
- [11] Xu L, Skoularidou M, Cuesta-Infante A, Veeramachaneni K (2019). Modeling tabular data using conditional GAN. *Advances in Neural Information Processing Systems*, 32: 7333–7343. <https://doi.org/10.48550/arXiv.1907.00503>
- [12] Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, Kalra MK, Zhang Y, Sun L, Wang G (2018). Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Transactions on Medical Imaging*, 37(6): 1348–1357. <https://doi.org/10.1109/TMI.2018.2827462>
- [13] Zietsch BP, Hansen JL, Hansell NK, Geffen GM, Martin NG, Wright MJ (2007). Common and specific genetic influences on EEG power bands delta, theta, alpha, and beta. *Biol Psychol*, 75(2): 154–164. <https://doi.org/10.1016/j.biopsycho.2007.01.004>
- [14] Michel CM, Lehmann D, Henggeler B, Brandeis D (1992). Localization of the sources of EEG delta, theta, alpha and beta frequency bands using the FFT dipole approximation. *Electroencephalogr Clin Neurophysiol*, 82: 38–44. [https://doi.org/10.1016/0013-4694\(92\)90180-P](https://doi.org/10.1016/0013-4694(92)90180-P)
- [15] Pant A, Kumar A, Verma C, Illés Z (2024). Comparative exploration on EEG signal filtering using window control methods. *Results in Control and Optimization*, 17: 1–16. <https://doi.org/10.1016/j.rico.2024.100485>
- [16] Ahmed MZI, Sinha N, Phadikar S, Ghaderpour E (2022). Automated feature extraction on AsMap for emotion classification using EEG. *Sensors*, 22(6): 1–17. <https://doi.org/10.3390/s22062346>
- [17] Chikhi S, Matton N, Blanchet S (2022). EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology*, 59(6): 1–16. <https://doi.org/10.1111/psyp.14009>
- [18] Ismail LE, Karwowski W (2020). Applications of EEG indices for the quantification of human cognitive performance: A systematic review and bibliometric analysis. *PLoS One*, 15(12): 1–29. <https://doi.org/10.1371/journal.pone.0242857>
- [19] Ambarwari SA, Adrian QJ, Herdiyeni Y (2020). Analysis of the effect of data scaling on the performance of the machine learning algorithm for plant identification. *Jurnal RESTI*, 4(1): 117–122. <https://doi.org/10.29207/resti.v4i1.1565>
- [20] Ahmed HA, Muhammad Ali PJ, Faeq AK, Abdullah SM (2022). An investigation on disparity responds of machine learning algorithms to data normalization method. *Aro-The Scientific Journal of Koya University*, 10(2): 29–37. <https://doi.org/10.14500/aro.10970>
- [21] Singh D, Singh B (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97: 1–15. <https://doi.org/10.1016/j.asoc.2020.105524>

- [22] Gunturkun O, Strockens F, Ocklenburg S (2020). Brain lateralization: A comparative perspective. *Physiol Rev*, 100(3): 1019–1063. <https://doi.org/10.1152/physrev.00006.2019>
- [23] Rogers LJ (2021). Brain lateralization and cognitive capacity. *Animals*, 11(7): 1–11. <https://doi.org/10.3390/ani11071996>
- [24] Szeghalmy S, Fazekas A (2023). A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning. *Sensors*, 23(4): 1–16. <https://doi.org/10.3390/s23042333>
- [25] Liu Q, Li C, Wanga V, Shepherd BE (2018). Covariate-adjusted Spearman’s rank correlation with probability-scale residuals. *Biometrics*, 74(2): 595–605. <https://doi.org/10.1111/biom.12812>
- [26] Arvanitis TN, White S, Harrison S, Chaplin R, Despotou G (2022). A method for machine learning generation of realistic synthetic datasets for validating healthcare applications. *Health Informatics Journal*, 28(2): 1–13. <https://doi.org/10.1177/14604582221077000>
- [27] Li Z, Zhu H, Lu Z, Yin M (2023). Synthetic data generation with large language models for text classification: Potential and limitations. *arXiv preprint*, 1–15. Retrieved from <http://arxiv.org/abs/2310.07849>
- [28] Namazi H, Aghasian E, Ala TS (2020). Complexity-based classification of EEG signal in normal subjects and patients with epilepsy. *Technology and Health Care*, 28(1): 57–66. <https://doi.org/10.3233/THC-191979>
- [29] You Y, Li Y, Yu B, Ying A, Zhou H, Zuo G, Xu J (2024). A study on EEG differences between active counting and focused breathing tasks for more sensitive detection of consciousness. *Frontiers in Neuroscience*, 18: 1–14. <https://doi.org/10.3389/fnins.2024.1341986>
- [30] Isa RM, Pasya I, Taib MN, Jahidin AH, Omar WRW, Fuad N (2012). Assessment of brainwave asymmetry and hemisphere dominance due to RF radiation. *Proceedings – UKSim-AMSS 6th European Modelling Symposium*, 153–157. <https://doi.org/10.1109/EMS.2012.89>
- [31] Ren B, Pan Y (2023). Extracting and supplementing method for EEG signal in manufacturing workshop based on deep learning of time–frequency correlation. *J Intell Manuf*, 34(7): 3179–3196. <https://doi.org/10.1007/s10845-022-01997-y>
- [32] Massar H, Belhoussine Drissi T, Nsiri B, Miyara M (2025). Advancements in blind source separation for EEG artifact removal: A comparative analysis of variational mode decomposition and discrete wavelet transform approaches. *Applied Acoustics*, 228: 1–12. <https://doi.org/10.1016/j.apacoust.2024.110300>

Optimal Operating Cost for a Grid-connected Microgrid Using Quadratic Interpolation Optimization

THACH CHI PHAN¹, THUAN THANH NGUYEN^{1*}, THANG TRUNG NGUYEN²

¹ Faculty of Electrical Engineering Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam.

² Power System Optimization Research Group, Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam.

*Corresponding author: nguyenthanhthuan@iuh.edu.vn

(Received: 3 February 2026; Accepted: 10 May 2026; Published online: 11 September 2026)

ABSTRACT: This study applies the Quadratic Interpolation Optimization Method (QIOM) to minimize the operating cost of a grid-connected MG over a 24-hour horizon. This optimization-based scheduling framework actively controls the diesel generator (DG) and battery energy storage system (BESS). In contrast, renewable sources such as solar and wind are treated as fixed inputs. Three operating scenarios are examined to assess the roles of the controllable generator and the BESS. The findings indicate that cost reduction mainly depends on how energy flows are coordinated over time, rather than on the total energy exchanged. The BESS plays a key role by shifting energy from low-load periods to peak-load periods, which improves overall economic performance. QIOM is also compared with Particle Swarm Optimization (PSO), Gravitational Search Algorithm (GSA), Grey Wolf Optimizer (GWO), and Heap-Based Optimization (HBO) in terms of solution quality, robustness, and computational performance. The comparative results demonstrate that QIOM achieves lower operating costs and more consistent performance across 20 independent runs. The Wilcoxon signed-rank test confirms that these improvements are statistically significant. However, this performance gain comes at the cost of increased computational time. Therefore, QIOM is more suitable for offline scheduling applications.

ABSTRAK: Kajian ini menggunakan Kaedah Pengoptimuman Interpolasi Kuadratik (QIOM) bagi meminimum kos operasi MG yang disambung ke grid selama 24 jam. Kajian ini mencadangkan penjadualan optimum di mana penjana diesel (DG) dan sistem storan tenaga bateri (BESS) dikawal secara aktif. Sebaliknya, sumber boleh diperbaharui seperti solar dan angin dianggap sebagai input tetap. Tiga senario operasi diperiksa bagi menilai peranan penjana boleh kawal dan BESS. Dapatan menunjukkan bahawa pengurangan kos bergantung pada pelarasan aliran tenaga dari semasa ke semasa, bukan pada jumlah penukaran tenaga. BESS memainkan peranan penting dengan mengalihkan tenaga dari tempoh beban rendah kepada tempoh beban puncak, ini meningkatkan prestasi ekonomi keseluruhan. Perbandingan QIOM dengan Pengoptimuman Kawanan Partikel (PSO), Algoritma Carian Graviti (GSA), Pengoptimum Serigala Kelabu (GWO), dan Pengoptimum Berasas Timbunan (HBO) turut dilakukan dari segi kualiti penyelesaian, kekukuhan dan prestasi pengiraan. Dapatan kajian menunjukkan bahawa QIOM mencapai kos operasi lebih rendah dengan prestasi konsisten merentasi 20 larian bebas. Ujian pangkat Wilcoxon mengesahkan bahawa penambahbaikan ini adalah signifikan secara statistik. Walau bagaimanapun, peningkatan prestasi ini datang dengan kos lebih tinggi daripada masa pengiraan. Oleh itu, QIOM lebih sesuai untuk aplikasi penjadualan luar talian.

KEYWORDS: Battery energy storage system, Metaheuristic optimization, Microgrid scheduling, Quadratic Interpolation Optimization, Temporal energy shifting

1. INTRODUCTION

Microgrids (MGs) have become a promising solution for integrating distributed energy resources, including photovoltaic (PV) systems, wind turbines (WT), DG, and BESS, into conventional power systems. By enabling local power generation and flexible operation, MGs improve system resilience, reduce transmission losses, and support the transition to low-carbon energy systems [1-3]. In particular, grid-connected MGs provide an efficient platform for coordinating renewable power generation with conventional sources while maintaining a reliable power supply.

However, the growing penetration of renewable energy sources introduces major operational challenges because of their intermittent characteristics. Changes in solar radiation, wind speed, and load demand complicate operational planning, leading to power imbalances and higher operating costs [4-8]. Recent studies highlight the importance of planning frameworks that account for these uncertainties in practical MG operations, especially under real-time variable electricity tariffs [9, 10]. To address these issues, BESS is widely integrated into MGs to provide energy buffers, reduce peak loads, and improve operational flexibility. However, coordinating multiple energy sources under time-varying conditions remains a challenging optimization task, especially in systems with nonlinear constraints and multidimensional decision spaces. The system's dynamic behavior increases operational complexity, making conventional methods unreliable and ineffective [11].

Metaheuristic methods are widely used in MG power management. PSO is one of the most common options. It is easy to implement and performs well in nonlinear problems. Its applications include MG planning, battery operation under uncertainty [11, 12], and reactive power dispatch with complex constraints. For independent MGs, hybrid PSO–GSA approaches have been used for frequency tuning [13]. GSA alone also performs well in PV-based MG optimization [14]. Even so, PSO has limitations. Performance can vary depending on parameter settings, and early convergence may still occur. This issue becomes more evident in high-dimensional or multimodal problems [15, 16]. These challenges are more pronounced in multi-phase MG optimization. The interaction of variables over time adds complexity. As a result, PSO may lose effectiveness under such conditions. However, most existing methods focus on minimizing total cost, while the temporal coordination of energy flows is often overlooked.

In parallel, nature-inspired artificial intelligence algorithms are emerging. The GWO algorithm demonstrates robust performance in complex MG optimization problems due to its balanced discovery-exploitation mechanism [17-20]. Similarly, the HBO algorithm improves convergence and reduces sensitivity to initial conditions in energy management problems [21-23]. This makes these methods suitable for complex MG optimization problems.

Despite the advantages, several limitations remain. Many methods still suffer from premature convergence, sensitivity to parameter adjustments, and reduced performance in problems with many constraints [24]. For example, PSO-based methods are sensitive to parameter settings such as inertial weights and acceleration coefficients, which can significantly affect convergence behavior. They can also converge prematurely and get stuck at local optimal points in the multimodal search space [15]. Most studies focus on minimizing total operating costs, while the temporal characteristics of energy exchange receive less attention. The timing of energy input and output is crucial because electricity prices and renewable energy generation profiles vary over time. Recent studies indicate that planning performance depends not only on the amount of energy exchanged but also on its timing [25, 26]. To address these limitations, this paper investigates applying QIOM, a derivative-free optimization technique [24], to minimize the operating costs of grid-connected MGs [27].

Unlike conventional metaheuristic algorithms, QIOM uses an interpolation-based search mechanism that balances global exploration and local exploitation without requiring gradient information. This method is evaluated using real-world meteorological data from Ho Chi Minh City, Vietnam, including solar radiation, wind speed, and temperature [28, 29]. This provides a more realistic assessment of its practical performance.

The main contributions of this study are summarized as follows:

- 1) A scenario-based optimization framework is developed to analyze MG operation under three practical configurations. It enables a clearer assessment of the roles of DG and BESS.
- 2) A consistent evaluation protocol is adopted using 20 independent runs. The mean and standard deviation are reported, and the Wilcoxon signed-rank test is used to compare algorithms.
- 3) The findings show that cost reduction is mainly driven by the coordination of energy flows over time, rather than the total energy exchanged.
- 4) The QIOM-based approach, which relies on interpolation without gradient information, is validated using real-world data from Ho Chi Minh City, showing competitive performance under practical operating conditions.

The remainder of this paper is organized as follows. Section 2 presents the system model and problem formulation. Section 3 describes the QIOM-based optimization method and its application to efficient MG scheduling. Section 4 discusses the simulation results, while Section 5 concludes the paper.

2. MICROGRID MODEL AND PROBLEM FORMULATION

Fig. 1 illustrates the structure of the MG system considered in this study. The system comprises four main components: a solar PV farm, a wind power farm, a DG, and a BESS. These energy sources collectively supply electricity to meet the demand profile, which follows a typical 24-hour load curve as depicted in Fig. 2.

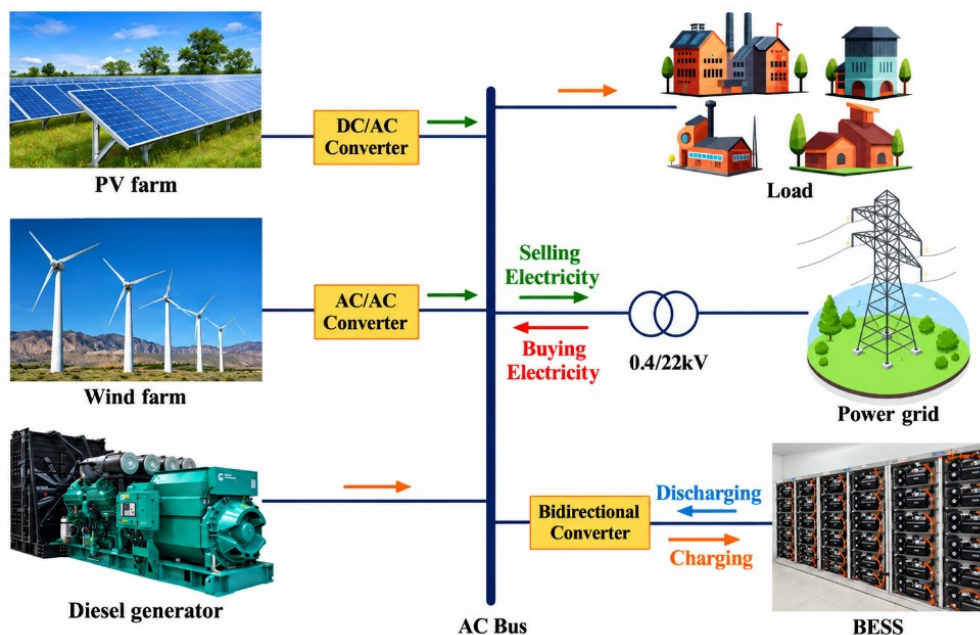


Figure 1. Configuration and energy flow structure of the proposed grid-connected MG.

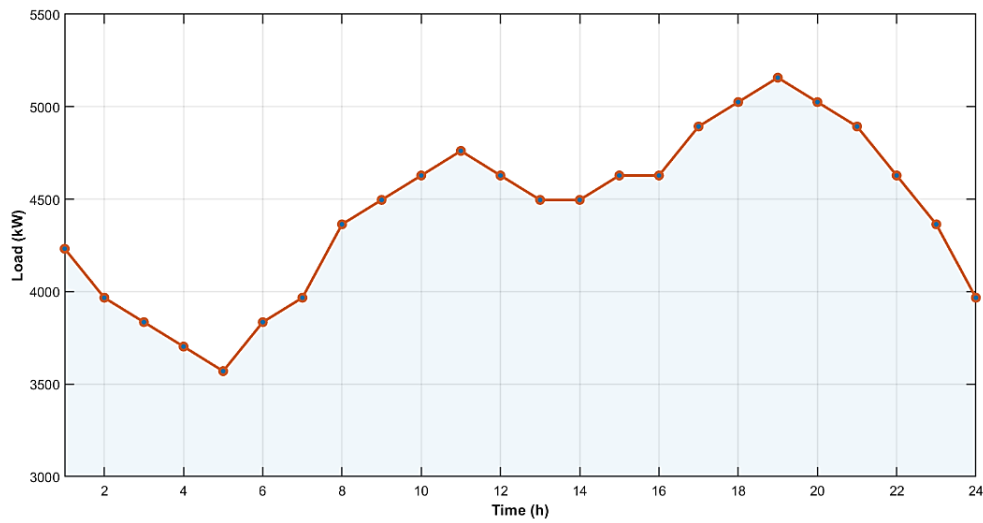


Figure 2. Typical 24-hour load demand profile used in the case study.

2.1. Photovoltaic farm

The PV farm in the MG is designed with a rated capacity of 2500 kW. The direct current (DC) output from the PV is converted to alternating current (AC) via an inverter and then fed into the AC bus. The actual power output at time t , denoted as $P_{PV}(t)$, mainly depends on solar irradiance and the operating temperature of the PV modules. The output power is estimated using the following Eq. (1) [30-32]:

$$P_{PV}(t) = P_{STC} \times (G(t)/1000) \times (1 - \alpha(T_{PV}(t) - 25)) \times \eta_{PV} \quad (1)$$

where: P_{STC} is the rated power under standard test conditions; $G(t)$ and $T_{PV}(t)$ represent the solar irradiance and PV module temperature at time t , respectively, while α denotes the temperature coefficient; and η_{PV} represents the DC/AC converter efficiency.

2.2. Wind farm

The wind farm consists of eight WTs with a combined rated capacity of 4000 kW. The electrical output from the WTs is converted to AC and delivered to the main AC bus via a power electronic converter. The actual power output at time t , denoted as $P_{WT}(t)$, is highly dependent on the wind speed and is calculated using the following expression given by Eq. (2) [33]:

$$P_{WT}(t) = N_{TB} \times \eta_w \times \begin{cases} 0, & V_t(t) < V_{cut_in} \\ P_n \left(\frac{V_t(t) - V_{cut_in}}{V_r - V_{cut_in}} \right)^3, & V_{cut_in} < V_t(t) < V_r \\ P_n, & V_r < V_t(t) < V_{cut_out} \end{cases} \quad (2)$$

where: N_{TB} is the number of WTs in the system; η_w is the AC/AC converter efficiency; $V_t(t)$ is the wind speed at time t ; V_r is the rated wind speed; P_n is the power of one WT; V_{cut_in} and V_{cut_out} are the cut-in and cut-out wind speeds.

The input data used in this study include solar irradiance, ambient temperature [28], and wind speed [29], as shown in Fig. 3. These data represent a typical dry-season day in Ho Chi Minh City, Vietnam.

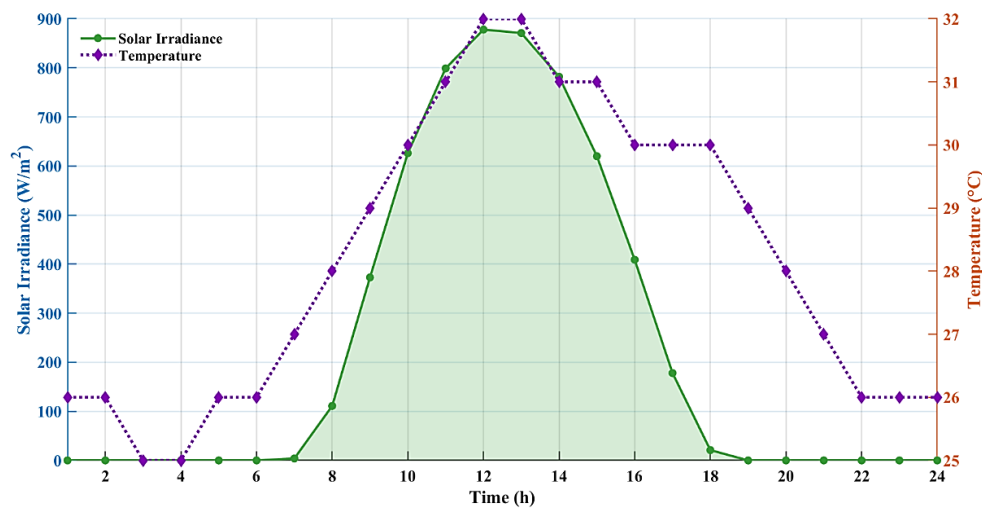


Figure 3. Solar irradiance and ambient temperature profiles for a typical dry-season day.

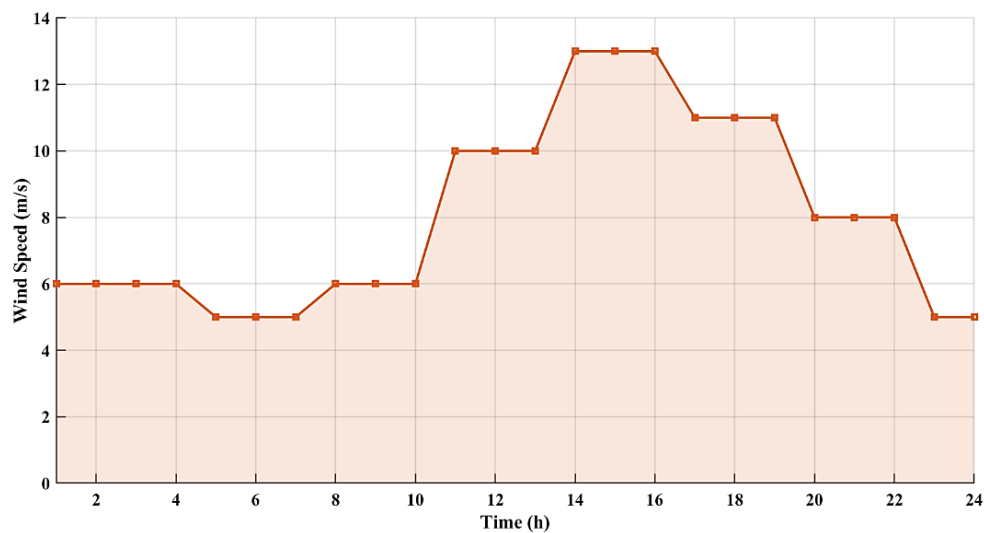


Figure 4. Wind speed profile for a typical dry-season day.

2.3. Diesel generator system

The DG has a capacity of 1750 kW, and its operating cost is determined by fuel consumption and fuel cost. This cost is calculated using Eq. (3) [34]:

$$C_{DG_OP}(t) = C_{diesel} \times (A_{DG} + B_{DG} \times P_{DG}(t) + C_{DG} \times P_{DG}^2(t)) \quad (3)$$

Where: A_{DG} , B_{DG} , and C_{DG} represent the coefficients that influence the fuel consumption of the DG, and C_{diesel} is the diesel fuel cost (\$/gal). The coefficient values are in Table 1.

Table 1. Coefficient values related to the DG [4].

Coefficients	Values
A_{DG}	22.9 (gal/h)
B_{DG}	0.0280 (gal/kWh)
C_{DG}	1.849×10^{-5} (gal/kWh) ²
C_{diesel}	2.258 (\$/gal)

The DG system's power output at time t , denoted as $P_{DG}(t)$, must operate within predefined minimum and maximum limits, as expressed in Eq. (4).

$$P_{DG_min} \leq P_{DG}(t) \leq P_{DG_max} \quad (4)$$

2.4. Battery energy storage system (BESS)

The BESS, with a total capacity of 4600 kWh, operates in both charging and discharging modes. It is connected to the AC bus through a bidirectional converter, enabling two-way energy exchange.

At any time t , the BESS power output/input, denoted as $P_{BESS}(t)$, is subject to operational constraints to ensure safe and efficient operation. These constraints can be expressed as follows:

$$-P_{BESS_max} \leq P_{BESS}(t) \leq P_{BESS_max} \quad (5)$$

where: P_{BESS_max} denotes the maximum discharging power (kW) that the BESS can supply to the MG, while $-P_{BESS_max}$ represents the maximum charging power.

The energy stored in the BESS at time t , represented as $E_{BESS}(t)$, is calculated using Eq. (6) [4].

$$E_{BESS}(t) = \begin{cases} E_{BESS}(t-1) - \frac{\Delta T \times P_{BESS}(t)}{\eta_d}, & P_{BESS}(t) > 0 \\ E_{BESS}(t-1) - \Delta T \times \eta_c \times P_{BESS}(t), & P_{BESS}(t) \leq 0 \end{cases} \quad (6)$$

where: η_c and η_d denote the charging and discharging efficiencies of the BESS, respectively, and ΔT is the sampling time (1 hour).

The BESS state of charge (SOC) is determined using Eq. (7).

$$SOC(t) = \frac{E_{BESS}(t)}{B_{capacity}} \quad (7)$$

where: $B_{capacity}$ denotes the energy capacity of the BESS (kWh).

To ensure battery longevity, the SOC must be maintained within predefined minimum and maximum limits, as expressed in Eq. (8).

$$SOC_{min} \leq SOC(t) \leq SOC_{max} \quad (8)$$

For operational continuity, the battery SOC is constrained to reach 50% at the end of the 24-hour cycle, ensuring energy availability for the subsequent day.

The BESS operating cost function is given by the following expression [35]:

$$C_{BESS_op}(t) = \frac{C_{BESS} \times \eta_c \times \Delta T \times P_{BESS_ch}(t)}{2 \times N_{cycle}} + \frac{C_{BESS} \times \Delta T \times P_{BESS_dis}(t)}{2 \times \eta_d \times N_{cycle}} + C_{deg} \times P_{BESS}(t) \quad (9)$$

where: $P_{BESS}(t)$ denotes the total power of the BESS at time t , $P_{BESS_ch}(t)$ and $P_{BESS_dis}(t)$ denote the charging and discharging power at time t , respectively; C_{BESS} refers to the investment cost of the BESS (\$/kWh), N_{cycle} is the number of cycles corresponding to the battery's lifespan, and C_{deg} is the degradation cost coefficient, representing the penalty associated with battery aging.

2.5. Power grid

A 5 MVA, 0.4/22 kV transformer facilitates power transfer between the MG and the utility grid, subject to transmission and stability constraints as expressed in Eq. (10).

$$P_{Utility_min} \leq P_{Utility}(t) \leq P_{Utility_max} \quad (10)$$

where: $P_{Utility}(t)$ denotes the power exchanged between the main grid and the MG at time t .

A positive value of $P_{Utility}(t) > 0$ indicates that power is imported from the grid to the MG, while a negative value $P_{Utility}(t) < 0$ indicates that power is exported from the MG to the grid.

The power transferred from the grid to the MG at time t is given as:

$$P_{Utility}(t) = P_{Load}(t) - P_{PV}(t) - P_{WT}(t) - P_{BESS}(t) - P_{DG}(t) \quad (11)$$

where: $P_{Load}(t)$ denotes the load power at time t , following the load curve shown in Fig. 2.

2.6. Objective function

The objective of the proposed optimization model is to minimize the MG's total operating cost over a 24-hour period. This objective is formulated through a fitness function, which serves as the performance metric for evaluating candidate solutions. The objective function can be formulated as follows:

$$Min C_{op_MG} = \sum_{t=1}^{24} (C_{MG_buy} + C_{MG_sell} + C_{B_op} + C_{DG_op}) \quad (12)$$

where: C_{MG_buy} denotes the cost of purchasing electricity from the grid (\$), while C_{MG_sell} represents the revenue obtained from selling electricity to the grid (\$). These values are determined by Eq. (13) and Eq. (14):

$$C_{MG_buy} = T_{buy}(t) \times P_{Utility}(t), P_{Utility}(t) > 0 \quad (13)$$

$$C_{MG_sell} = T_{sell}(t) \times P_{Utility}(t), P_{Utility}(t) < 0 \quad (14)$$

where: $T_{buy}(t)$, $T_{sell}(t)$ represent the electricity procurement and sale tariffs at time t , as shown in Table 2.

Table 2. Time-based electricity prices for cost optimization modeling [4].

Type of tariff	Time	Value (\$/kWh)
Purchasing	00:00 – 7:00	0.06
	8:00 – 16:00	0.144
	17:00 – 20:00	0.252
	21:00 – 00:00	0.144
Selling	All day	0.0582

3. OPTIMIZING THE OPERATING COST OF MG SYSTEM USING QUADRATIC INTERPOLATION OPTIMIZATION METHOD

QIOM is a mathematically driven algorithm that leverages quadratic interpolation through three points to locate a function's minimum. It has demonstrated effectiveness in numerous optimization problems, particularly in curve-based search strategies [24, 36-37].

3.1. Quadratic Interpolation Method

Let $q(x)$ be a quadratic function defined as an approximation of the objective function $f(x)$. By identifying the minimum point of $q(x)$, an approximate minimum of $f(x)$ can be obtained. The quadratic function is defined as:

$$q(x) = ax^2 + bx + c, \text{ with } a, b, c \in R \quad (15)$$

Assume that the function $f(x)$ is known at three distinct points $P(x_i, f(x_i))$, $P(x_j, f(x_j))$, and $P(x_k, f(x_k))$, where $a \leq x_i < x_j < x_k \leq b$. The quadratic approximation $q(x)$ satisfies the interpolation conditions at these points:

$$\begin{cases} q(x_i) = ax_i^2 + bx_i + c = f(x_i) \\ q(x_j) = ax_j^2 + bx_j + c = f(x_j) \\ q(x_k) = ax_k^2 + bx_k + c = f(x_k) \end{cases} \quad (16)$$

By differentiating Eq. (15) and setting the derivative to zero, the minimum point of the parabola is found as:

$$x_{\min} = -\frac{b}{2a} \quad (17)$$

Alternatively, the minimum can also be calculated using the following expression derived from Newton's divided differences:

$$x_{\min} = \frac{1}{2}(x_i + x_j) + \frac{1}{2} \frac{(f(x_i) - f(x_j))(f(x_j) - f(x_k))(f(x_k) - f(x_i))}{(x_j - x_k)f(x_i) + (x_k - x_i)f(x_j) + (x_i - x_j)f(x_k)} \quad (18)$$

The interpolation function $q(x)$ can also be constructed explicitly using the Lagrange interpolation formula:

$$q(x) = \frac{(x - x_j)(x - x_k)}{(x_i - x_j)(x_i - x_k)} f(x_i) + \frac{(x - x_i)(x - x_k)}{(x_j - x_i)(x_j - x_k)} f(x_j) + \frac{(x - x_i)(x - x_j)}{(x_k - x_i)(x_k - x_j)} f(x_k) \quad (19)$$

Once $x_{\min}$ is found, a new triplet is formed around it to define a smaller search region. Quadratic interpolation is then repeated iteratively until a convergence criterion or target accuracy is reached, yielding the final approximate minimum.

3.2. Quadratic Interpolation Optimization Method

The QIOM consists of two main phases, namely exploration and exploitation, as illustrated in Fig. 5. The algorithm starts by initializing parameters and generating a random population of candidate solutions. After evaluating the objective function, the best solution is returned if the stopping criterion is met; otherwise, the iterative optimization process continues.

At each iteration, two candidate solutions are randomly selected from the population. Let $rando \sim U(0,1)$ be a control random variable used to probabilistically determine the search phase. If $rando > 0.5$, the Generalized Quadratic Interpolation (GQI) function is computed based on the current solution and the two selected individuals Eq. (20), followed by the exploration phase Eq. (21). Otherwise, the GQI function is constructed using the best-known solution and two randomly selected individuals Eq. (22), leading to the exploitation phase Eq. (23).

After a new candidate solution is generated, the best solution is selected based on the fitness comparison described in Eq. (24). This iterative process continues until the convergence criterion is met.

3.2.1. Exploration phase

This phase explores the full variable space to avoid local optima and premature convergence. Promising regions are identified by randomly selecting individuals that satisfy the GQI conditions. Positions are updated using Eq. (20) and Eq. (21) [24].

$$x_{\min_i, \text{rand}o1, \text{rand}o2}(t) = GQI(x_i(t), x_{\text{rand}o1}(t), x_{\text{rand}o2}(t), \text{fit}(x_i(t)), \text{fit}(x_{\text{rand}o1}(t)), \text{fit}(x_{\text{rand}o2}(t))) \quad (20)$$

$$v_i(t+1) = x_{\min_i, \text{rand}o1, \text{rand}o2}(t) + w_1 \cdot (x_{\text{rand}o3}(t) - x_{\min_i, \text{rand}o1, \text{rand}o2}(t)) + \text{round}(0.5(0.5 + r_1)) \log \frac{r_2}{r_3} \quad (21)$$

where: $x_{\text{rand}o1}$, $x_{\text{rand}o2}$, and $x_{\text{rand}o3}$ represent the positions of selected individuals from the population; r_1 , r_2 , and r_3 are random numbers uniformly distributed in (0, 1), and $\text{fit}(\cdot)$ denotes the fitness function value. The minimizer $x_{\min}$ is estimated from the triplet $(x_i, \text{fit}(x_i))$, $(x_{\text{rand}o2}, \text{fit}(x_{\text{rand}o2}))$ and $(x_{\text{rand}o3}, \text{fit}(x_{\text{rand}o3}))$ via GQI.

3.2.2. Exploitation phase

This phase enhances the best-known solution through a local search strategy to accelerate convergence toward the global optimum. QIOM randomly selects two individuals from the population and uses the current best solution in the GQI process to generate an improved candidate. The updated position is calculated using Eq. (22) within the defined minimum region [24].

$$x_{\min_best, \text{rand}o1, \text{rand}o2}(t) = GQI(x_{\text{best}}(t), x_{\text{rand}o1}(t), x_{\text{rand}o2}(t), \text{fit}(x_{\text{best}}(t)), \text{fit}(x_{\text{rand}o1}(t)), \text{fit}(x_{\text{rand}o2}(t))) \quad (22)$$

$$v_i(t+1) = x_{\min_i, \text{rand}o1, \text{rand}o2}(t) + w_2 \cdot (x_{\text{best}}(t) - \text{round}(1 + \text{rand}o)) \cdot \frac{Ub - Lb}{Ub^{rD} - Lb^{rD}} \cdot x_i^{rD}(t) \quad (23)$$

In the exploitation phase, two individuals $x_{\text{rand}o1}$ and $x_{\text{rand}o2}$, along with the best solution x_{best} identified so far, are used in the GQI process to generate the interpolation minimizer $x_{\min_best, \text{rand}o1, \text{rand}o2}$ as Eq.(22). This point is expected to yield a better solution than those obtained from three randomly selected individuals.

To further enhance local exploitation, the position of the i^{th} individual is updated by searching around $x_{\min_best, \text{rand}o1, \text{rand}o2}$, using a random perturbation x_i^{rD} , as defined in Eq. (23).

Finally, the i^{th} individual's position is updated based on a fitness comparison, as follows:

$$x_i(t+1) = \begin{cases} x_i(t) & \text{fit}(x_i(t)) \leq \text{fit}(v_i(t+1)) \\ v_i(t+1) & \text{fit}(x_i(t)) > \text{fit}(v_i(t+1)) \end{cases} \quad (24)$$

3.3. Application of QIOM to optimize the operating cost of MG system

The QIOM-based framework minimizes the daily operating cost of the MG shown in Fig. 1. Section 2 describes the system configuration, including power components, load demand, and electricity tariff. Renewable power generation is estimated using input data including solar irradiance, ambient temperature, and wind speed, as illustrated in Fig. 3 and Fig. 4. The implementation procedure follows the flowchart in Fig. 5.

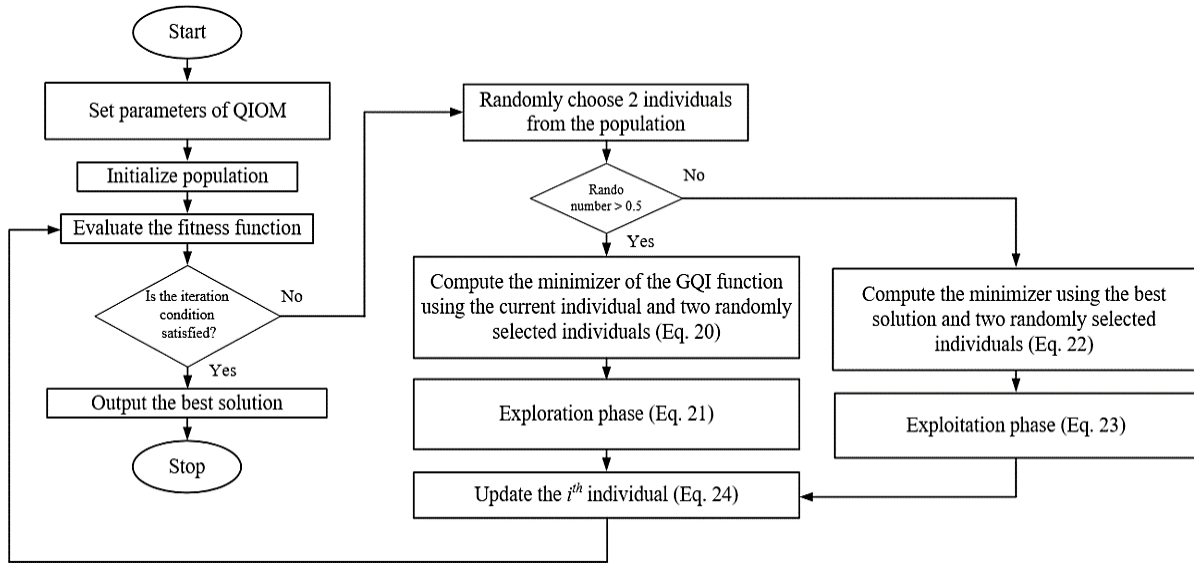


Figure 5. Flowchart of the QIOM algorithm for MG cost optimization.

The main steps are outlined as follows:

- **Step 1:** Define MG parameters

The optimization problem is formulated by specifying decision variables, constraints, and the objective function in Eq. (12). The algorithm is run for up to 500 iterations with a population size of 50. The decision variables are the hourly outputs of the DG, $P_{DG}(t)$, and the BESS, $P_{BESS}(t)$, over a 24-hour horizon. This results in 48 variables.

The PV and WT outputs are treated as fixed inputs and are obtained directly from the data. Other variables are derived through system equations and constraints. In particular, the power exchanged with the grid is computed using the balance equation in Eq. (11). The battery state of charge (SOC) is updated sequentially based on Eq. (7).

- **Step 2:** Initialize population

A randomly generated initial population of 50 individuals is created.

- **Step 3:** Evaluate initial population

The fitness value of each individual x_i (where $i=1 \div 50$) is calculated. The best solution obtained so far, x_{best} , is then identified, representing the selected values of P_{BESS} and P_{DG} .

- **Step 4:** Check the iteration conditions

If the stopping criterion is satisfied, the algorithm terminates. Otherwise, it continues to the next iteration

- **Step 5:** Perform optimization iterations

Each individual is updated based on a probabilistic rule. If $rando > 0.5$, exploration is carried out using three randomly selected individuals Eq. (20) and Eq. (21). Otherwise, the exploitation phase is performed using the current best solution together with two random individuals, as described in Eq. (22) and Eq. (23).

- **Step 6:** Update best individual positions

Update the individuals' positions using Eq. (24). Reevaluate fitness to determine whether a better solution is found.

The procedure repeats until the stopping condition is met. The final output includes the minimum operating cost and the corresponding hourly dispatch of the DG and BESS.

4. RESULTS AND DISCUSSION

All simulations were conducted using MATLAB R2023a on a personal computer equipped with an Intel Core i7 processor (1.9 GHz), 16 GB RAM, and Windows 10 Pro. For consistency, all algorithms used the same parameter settings described in Section 3.3. The compared methods are QIOM, PSO, GSA, GWO, and HBO. Each scenario was simulated over a 24-hour period. Since the algorithms are stochastic, 20 independent runs were carried out for each method. The results are evaluated using standard statistical metrics, including the best, mean, and standard deviation of the objective function.

In addition, the Wilcoxon signed-rank test is applied at a significance level of 0.05 to assess statistical differences between the methods. The numerical results for different operating scenarios are summarized in Tables 3 and 4. The corresponding power profiles are shown in Figs. 6-9.

To further evaluate performance, QIOM is analyzed under three operating scenarios:

- Scenario 1: MG without DG and BESS;
- Scenario 2: MG with DG but without BESS;
- Scenario 3: MG with both DG and BESS.

Table 3. Grid energy exchange under different MG scenarios.

Scenarios	Grid energy import (kWh)	Grid energy export (kWh)
1	66,432	-1,465.80
2	48,088	-2,368.21
3	48,460	-2,815.29

Table 4. Cost components of the MG under different scenarios.

Scenarios	Grid purchase cost (\$)	DG Operating cost (\$)	BESS operating cost (\$)	Grid selling revenue (\$)	Total operating cost (\$)	Cost difference (%) (vs. Scenario 3)
1	8,558.47	0	0	-85.31	8,473.16	5.53
2	5,338.77	3,366.31	0	-137.83	8,567.25	6.70
3	4,705.50	3,322.06	165.58	-163.85	8,029.29	0

4.1. Scenario 1: MG without DG and BESS

Scenario 1 represents the basic configuration without variable power generation or an energy storage system. Under these conditions, the MG relies entirely on renewable energy sources and grid exchange to meet demand. As illustrated in Fig. 6, the electricity output from renewable sources is insufficient most of the time, especially in the early morning and late evening.

In this scenario, electricity imports from the grid become the primary and most consistently high source of supply throughout the day. Around midday, solar power production increases and temporarily exceeds demand, resulting in a short export period. This surplus is limited and cannot be stored because the MG lacks a BESS.

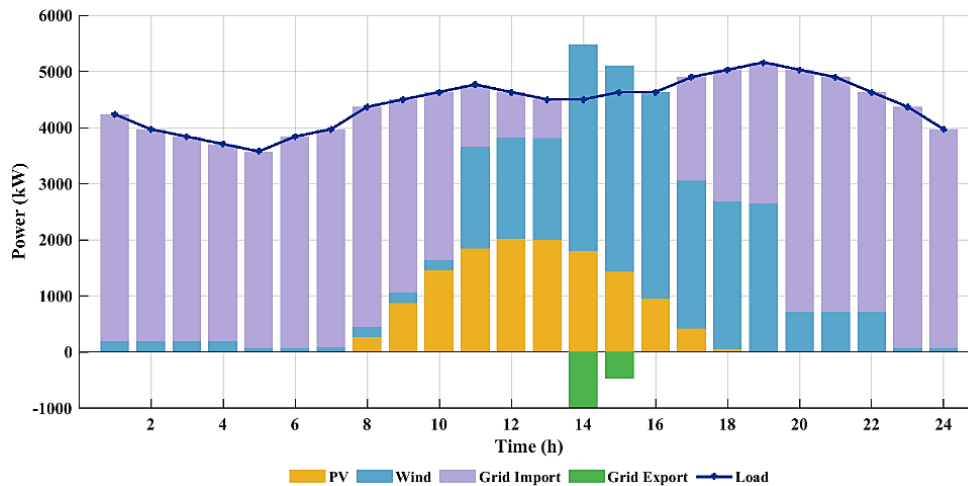


Figure 6. Hourly power distribution of the MG without DG and BESS.

Table 3 reflects this behavior: total energy imports reach 66,432 kWh, while energy exports remain minimal. The MG operates with continuous grid-based compensation. Operationally, the MG cannot regulate power or shift energy over time. The grid acts as the only adjustable supply source. This leads to inefficient use of available renewable energy sources when the time flexibility in energy use is not effectively exploited.

Table 4 clearly shows the economic impact. The cost of purchasing electricity from the utility grid constitutes the largest portion of the total operating cost, resulting in a total cost of \$8,473.16. Although DG and BESS have no costs, the MG remains expensive because it relies on external power sources.

4.2. Scenario 2: MG with DG and without BESS

Compared to scenario 1, adding DG provides a controllable power source, shifting the system from fully reactive operation to partially controllable operation.

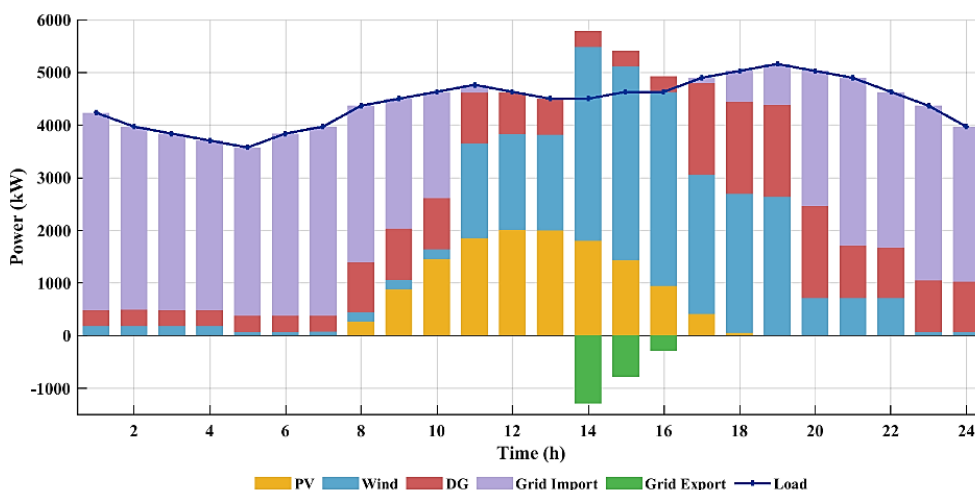


Figure 7. Hourly power distribution of the MG with DG and without BESS.

As illustrated in Fig. 7, DG is deployed during periods of insufficient renewable energy supply. This reduces the amount of electricity purchased from the grid, especially during medium- and peak-load periods. Grid electricity imports decreased to 48,088 kWh, representing a reduction of approximately 27.6% compared to Scenario 1.

From an operational perspective, DG acts as a deployable source that reduces real-time power imbalances but does not provide inter-temporal flexibility, which remains limited due to the lack of BESS. Excess renewable energy continues to be exported instead of stored, and energy cannot be shifted over time. This limitation is reflected in cost outcomes. Although electricity import costs from the grid decreased considerably to \$5,338.77, the DG operating cost of \$3,366.31 offset this benefit. As a result, total operating costs increased slightly to \$8,567.25. Simply reducing electricity imports from the grid does not guarantee lower costs. Total costs depend not only on the amount of energy exchanged but also on the cost of local electricity production and the timing of energy use.

4.3. Scenario 3: MG with DG and BESS

Scenario 3 considers the coordinated operation of both DG and BESS, enabling both controllable power generation and energy storage. Fig. 8 shows a clear shift in operating strategy. BESS is charged during periods of surplus renewable energy generation and discharged during high-demand periods. This reduces the mismatch between power output and load demand.

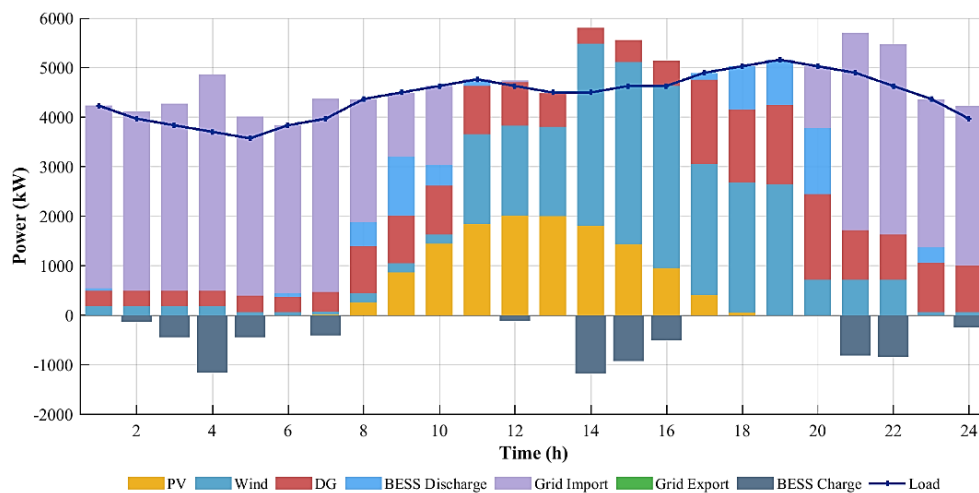


Figure 8. Hourly power distribution of the MG with DG and BESS.

Unlike previous scenarios, surplus renewable energy is not immediately exported but is partially stored for later use. During peak periods, the stored energy is released, reducing both electricity imports from the grid and the DG output.

As shown in Table 3, total energy imports (48,460 kWh) are similar to Scenario 2, but energy exports increase to 2,815.29 kWh, indicating improved utilization of renewable resources. Economically, grid electricity purchases decrease to \$4,705.50. Although additional costs are incurred due to DG and BESS operation, these costs are offset by improved energy management and increased revenue from energy exports. Therefore, total operating costs decreased to \$8,029.29, the lowest among all scenarios. Compared to Scenario 1 and Scenario 2, the cost reductions were approximately 5.53% and 6.70%, respectively. The results indicate that cost reduction is more closely related to the timing of energy exchange than to the total exchanged energy.

4.4. Comparative analysis of scenarios

Fig. 9 shows the power exchange between the MG and the utility grid in three scenarios and highlights the impact of DG and BESS on system operation.

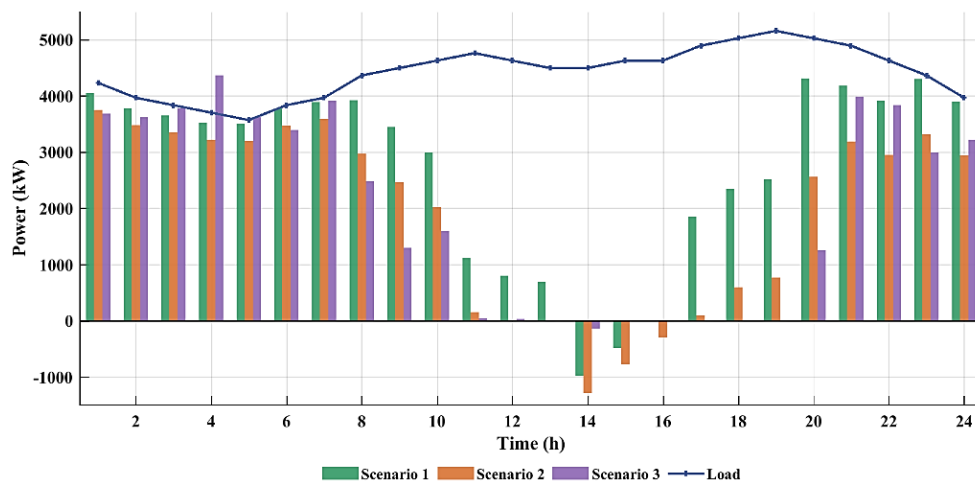


Figure 9. Power exchange between the MG and the utility grid under the three scenarios.

In Scenario 1, electricity imports from the grid predominate, with the grid acting as the primary balancing source. In Scenario 2, DG reduces the amount of electricity drawn from the grid. However, the overall model operates without an energy shift.

In Scenario 3, the grid exchange curve becomes smoother and less correlated with load. Peak-hour electricity imports decrease, and export timing is better controlled. This confirms the effectiveness of the BESS system, allowing for energy storage and redistribution over time.

Although the total amount of energy imported in Scenario 3 is equivalent to Scenario 2, the timing of energy exchange is substantially improved. This demonstrates that cost reduction depends not only on the amount of energy exchanged but also on the timing of the exchange, reflecting a decoupled relationship between energy quantity and economic cost. In this context, DG improves instantaneous power balancing, while BESS enables energy shifting over time; their combined operation enhances both system flexibility and economic efficiency.

4.5. Algorithm performance comparison

QIOM was compared with PSO, GSA, GWO, and HBO based on convergence performance, optimization accuracy, robustness, and computational efficiency. As shown in Fig. 10, QIOM rapidly decreased the objective function in the first few iterations and reached a stable region between iterations 150 and 200. This behavior reflects the algorithm's effectiveness in combining exploration with exploitation.

PSO converges faster initially but shows noticeable fluctuations, reflecting its sensitivity to random updates. GSA shows smoother convergence but at a slower rate, maintaining higher target values over a large number of iterations. GWO and HBO exhibit intermediate behavior, offering improved stability compared to PSO but converging more slowly than QIOM.

The statistical results in Table 5 further confirm these observations. Based on 20 independent runs, QIOM achieves the lowest best cost (\$8,029.29) and the lowest mean cost (\$8,091.60) among all compared algorithms. In addition, it achieves the lowest standard deviation (32.5), demonstrating better stability and robustness against random initialization. These results show that QIOM provides more consistent and reliable performance than the other methods.

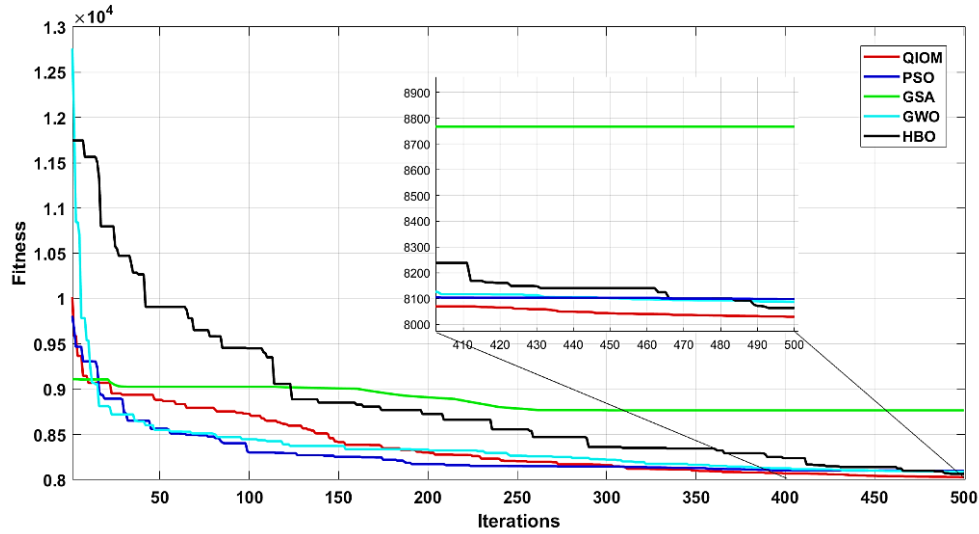


Figure 10. Convergence curves of the algorithms.

Table 5. Performance comparison of algorithms in terms of operating cost (\$/day) over 20 independent runs.

Algorithm	Best	Time (s)	Mean	Δ Mean (%)	Std	Δ Std (%)
QIOM	8029.29	47.22	8091.6	0	32.5	0
PSO	8098.55	5.23	8266.8	2.16	104.5	221.5
GSA	8766.96	10.32	8949.4	10.6	77.7	139.1
GWO	8087.38	6.32	8204.7	1.4	97.9	201.2
HBO	8063.62	5.94	8170.8	0.98	65.3	100.9

The box plot in Fig. 11 shows that QIOM produces a narrower distribution of results, confirming its stability compared to the other algorithms.

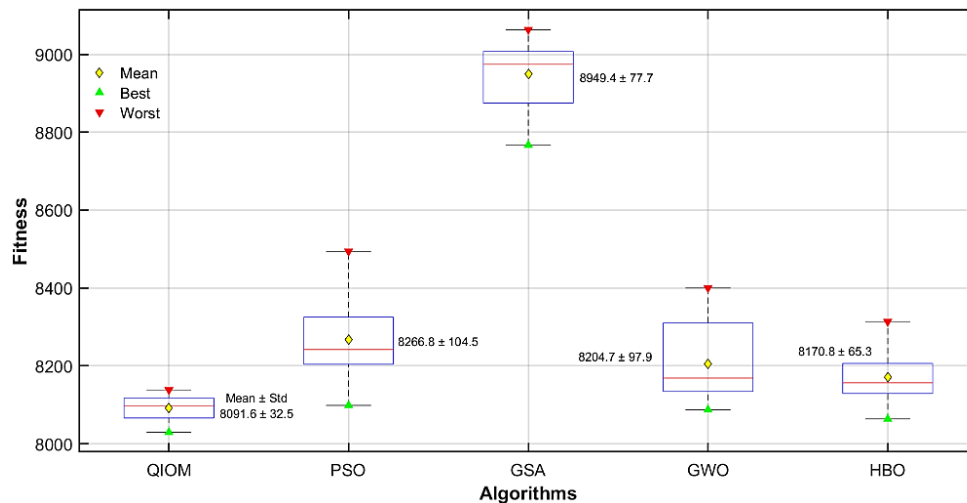


Figure 11. Boxplot comparison of fitness values over 20 independent runs.

To statistically validate QIOM's performance, the Wilcoxon signed-rank test results in Table 6 indicate that all observed differences are significant at the 0.05 level [38-40]. In particular, comparisons with PSO and GSA yield p-values below 0.001, highlighting QIOM's pronounced performance advantage.

Table 6. Wilcoxon signed-rank test results for algorithm comparison.

Comparison	p-value	Significance ($\alpha = 0.05$)
QIOM vs PSO	< 0.001	Yes
QIOM vs GSA	< 0.001	Yes
QIOM vs GWO	0.002	Yes
QIOM vs HBO	0.015	Yes

In terms of computational time, QIOM requires more processing time due to its interpolation-based mechanism in a high-dimensional search space. However, the improved solution quality and stability make it suitable for offline optimization applications, where robustness is more critical than execution speed.

5. CONCLUSIONS

This study applies QIOM to minimize the operating cost of a grid-connected MG over a 24-hour scheduling horizon. The scenario analysis reveals different operational roles. The DG mainly supports real-time power balance. However, when used alone, it does not reduce costs. In contrast, the BESS introduces flexibility by shifting energy across time. This shifting effect improves the overall economic performance. The operating cost depends not only on the total energy exchanged with the grid, but also on when this exchange occurs. With coordinated operation (Scenario 3), the total cost is reduced by approximately 5.53% and 6.70% compared with Scenarios 1 and 2, respectively. In terms of optimization, QIOM yields lower costs than PSO, GSA, GWO, and HBO. The average cost is reduced by up to 2.16% compared to PSO and 10.6% compared to GSA. In addition, QIOM produces more stable results, as indicated by smaller variations across runs. The Wilcoxon signed-rank test further confirms that these improvements are statistically significant ($p < 0.05$). Although QIOM requires more computational time, it provides stable and high-quality solutions. Therefore, it is suitable for offline scheduling problems. Future research will consider uncertainty and real-time operation.

REFERENCES

- [1] Akter A, et al. (2024) A review on microgrid optimization with meta-heuristic techniques: Scopes, trends and recommendation. *Energy Strategy Reviews*, 51:101298.
- [2] Shaker HK, Keshta HE, Mosa MA, Ali AA. (2023) Adaptive nonlinear controllers based approach to improve the frequency control of multi islanded interconnected microgrids. *Energy Reports*, 9:5230-5245.
- [3] Emam AA, Keshta HE, Mosa MA, Ali AA. (2023) Bi-level energy management system for optimal real time operation of grid tied multi-nanogrids. *Electric Power Systems Research*, 214:108957.
- [4] Hassaballah EG, Keshta HE, Abdel-Latif KM, Ali AA. (2024) A novel strategy for real-time optimal scheduling of grid-tied microgrid considering load management and uncertainties. *Energy*, 299:131419.
- [5] Tang Z, Liang T, Mi D, Huang Z, Jing Y, Wu H. (2026) Multi-objective optimization of virtual power plant with mobile storage considering renewable energy uncertainty and multiple flexible loads. *International Journal of Electrical Power & Energy Systems*, 143:108422.
- [6] Halev A, Liu Y, Liu X. (2024) Microgrid control under uncertainty. *Engineering Applications of Artificial Intelligence*, 138:109360.
- [7] Wang C, Zhang Z, Abedinia O, Gholami S. (2020) Modeling and analysis of a microgrid considering the uncertainty in renewable energy resources, energy storage systems and demand management in electrical retail market. *Journal of Energy Storage*, 33:102111.

-
- [8] Shayeghi H, Davoudkhani IF. (2025) Uncertainty aware energy management in microgrids with integrated electric bicycle charging stations and green certificate market. *Scientific Reports*, 15:1-28.
 - [9] Yang M, Wang J, Chen Y, Zeng Y, Su X. (2024) Data-driven robust optimization scheduling for microgrid day-ahead to intra-day operations based on renewable energy interval prediction. *Energy*, 313:134058.
 - [10] Aguilar D, Quinones JJ, Pineda LR, Ostanek J, Castillo L. (2024) Optimal scheduling of renewable energy microgrids: A robust multi-objective approach with machine learning-based probabilistic forecasting. *Applied Energy*, 369:123548.
 - [11] Li Q, Zhang N, Zhu X, Liu J. (2021) Optimization of dynamic reactive power devices' coordinate operation for minimizing HVDC's successive commutation failure based on PSO. *Journal of Engineering Research*, 10(4):234-251.
 - [12] Hwang H, Shi S, Liang X, Zhang D, Dai W, Liu H. (2022) Optimal energy scheduling of grid-connected microgrids with demand side response considering uncertainty. *Applied Energy*, 327:120094.
 - [13] Zishan F, Akbari E, Montoya OD, Giral-Ramírez DA, Molina-Cabrera A. (2022) Efficient PID control design for frequency regulation in an independent microgrid based on the hybrid PSO-GSA algorithm. *Electronics*, 11(23):3886.
 - [14] Pervez I, Sarwar A, Tayyab M, Sarfraz M. (2019) Gravitational Search Algorithm (GSA) based Maximum Power Point Tracking in a Solar PV based Generation System. *Innovations in Power and Advanced Computing Technologies (i-PACT)*, 1-6.
 - [15] Lazreg M, Benamrane N. (2022) Hybrid system for optimizing the robot mobile navigation using ANFIS and PSO. *Robotics and Autonomous Systems*, 153:104114.
 - [16] Amrane F, Gherouat O, Francois B, Itouchene H, Medani KBO, Chaiba A, Belkhier Y. (2025) Robust PSO-Optimized Power Control for Stand-Alone Variable-Speed DFIG Wind-Based High Stability and Power Quality Enhancement: Hardware Investigation-Based dSPACE1103. *International Journal of Energy Research*, 2025:2004125.
 - [17] Kumar G, Guerrero JM, Prakash O. (2021) Optimisation of solar / wind / bio-generator / diesel / battery based microgrids for rural areas: A PSO-GWO approach. *Sustainable Cities and Society*, 67:102723.
 - [18] Biabani F, Shojaei S, Hamzehei-javaran S. (2022) A new insight into metaheuristic optimization method using a hybrid of. *Structures*, 44:1168-1189.
 - [19] Devarapalli R, Rao BV, Al-durra A. (2022) Optimal parameter assessment of Solar Photovoltaic module equivalent circuit using a novel enhanced hybrid GWO-SCA algorithm. *Energy Reports*, 8:12282-12301.
 - [20] Itouchene H, Amrane F, Gherouat O, Boudries Z. (2025) Design and implementation of a grey wolf optimization – Based higher-order sliding mode controller for a DFIG-based wind turbine using dSPACE 1104. *Energy Reports*, 14:4324-4352.
 - [21] Salama D, Houssein EH, Said M, Oliva D, Nabil A. (2022) An Efficient Heap-Based Optimizer for Parameters Identification of Modified Photovoltaic Models. *Ain Shams Engineering Journal*, 13(5):101728.
 - [22] Shaheen AM, El-sehiemy RA, Hasanien HM, Ginidi AR. (2022) An improved heap optimization algorithm for efficient energy management based optimal power flow model. *Energy*, 250:123795.
 - [23] Abd M, El-said EMS, Elsheikh AH, Abdelaziz GB. (2022) Performance prediction of solar still with a high-frequency ultrasound waves atomizer using random vector functional link / heap-based optimizer. *Advances in Engineering Software*, 170:103142.
 - [24] Zhao W, Wang L, Zhang Z, Mirjalili S, Khodadadi N, Ge Q. (2023) Quadratic Interpolation Optimization (QIO): A new optimization algorithm based on generalized quadratic
-

- interpolation and its applications to real-world engineering problems. *Computer Methods in Applied Mechanics and Engineering*, 417:116446.
- [25] Du J, et al. (2024) A multi-objective robust dispatch strategy for renewable energy microgrids considering multiple uncertainties. *Sustainable Cities and Society*, 116:105918.
- [26] Qing K, Du Y, Huang Q, Duan C, Hu W. (2024) Energy scheduling for microgrids with renewable energy sources considering an adjustable convex hull based uncertainty set. *Renewable Energy*, 220:119611.
- [27] Otieno E, Farzaneh H. (2026) Multi-objective optimal sizing and location of the grid-connected renewable microgrids, considering cost-effectiveness, power loss, and network stability approaches. *Energy Conversion and Management: X*, 29:101465.
- [28] Tutiempo Network. (2024) Temperature. <https://en.tutiempo.net/ho-chi-minh-city.html>
- [29] Windfinder Network. (2024) Wind speed. https://www.windfinder.com/forecast/ho_chi_minh_city
- [30] Keshta HE, Saied EM, Malik OP, Bendary FM, Ali AA. (2021) Fuzzy PI controller-based model reference adaptive control for voltage control of two connected microgrids. *IET Generation, Transmission & Distribution*, 15(4):602-618.
- [31] Akter A, et al. (2024) A review on microgrid optimization with meta-heuristic techniques: Scopes, trends and recommendation. *Energy Strategy Reviews*, 51:101298.
- [32] Taha MS, Abdeltawab HH, Mohamed YARI. (2018) An online energy management system for a grid-connected hybrid energy source. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 6(4):2015-2030.
- [33] Babaei M, Azizi E, Beheshti MT, Hadian M. (2020) Data-Driven load management of stand-alone residential buildings including renewable resources, energy storage system, and electric vehicle. *Journal of Energy Storage*, 28:101221.
- [34] Abujarad SY, Mustafa MW, Jamian JJ. (2017) Recent approaches of unit commitment in the presence of intermittent renewable energy resources: A review. *Renewable and Sustainable Energy Reviews*, 70:215-223.
- [35] Garcia-Torres F, Bordons C. (2015) Optimal Economical Schedule of Hydrogen-Based Microgrids With Hybrid Storage Using Model Predictive Control. *IEEE Transactions on Industrial Electronics*, 62(8):5195-5207.
- [36] Dao F, Zeng Y, Zou Y, Qian J. (2024) Fault diagnosis method for hydropower unit via the incorporation of chaotic quadratic interpolation optimized deep learning model. *Measurement*, 237:115199.
- [37] Tong W, Su B, Zhan Y. (2026) An initial value insensitive method for phase equilibrium calculation: Constrained quadratic interpolation optimization algorithm. *Expert Systems with Applications*, 298:129597.
- [38] Derrac J, García S, Molina D, Herrera F. (2011) A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation*, 1(1):3-18.
- [39] Deng J, Zhang X, Jia X. (2025) Cost-efficient PON fault localization utilizing Wilcoxon signed rank test. *Measurement*, 256:118314.
- [40] Zhang X, Feng R, Deng J, Jia X. (2026) An integrated scheme for fault detection and localization in passive optical network monitoring using the Wilcoxon signed-rank test. *Results in Engineering*, 29:109707.

A Hybrid Framework for Real-Time Sentiment Analysis and Adaptive Conversational Outcome Prediction in Domain-Specific Dialogue Systems

MEJDL SAFRAN*

*Department of Computer Science, College of Computer and Information Sciences,
King Saud University, Riyadh, Saudi Arabia*

**Corresponding author: mejdl@ksu.edu.sa*

(Received: 04 February 2026; Accepted: 28 August 2026; Published online: 11 September 2026)

ABSTRACT: The lack of non-verbal signals in text-based communication forces dialogue systems to rely entirely on dynamic textual sentiment changes to infer user intent and emotional shifts. However, substantial class imbalance and the static, post-hoc nature of existing text analysis models prevent dialogue systems from recognizing these emotional dynamics and predicting conversational outcomes in real time. This research presents a hybrid machine learning framework to address these challenges through integrated sentiment-aware analysis and real-time conversation outcome prediction. We first utilize a pre-trained RoBERTa model to assign sentiment scores to individual messages, which is required for two subsequent tasks: (i) conversational outcome classification to classify conversations into Interested in Islam, Wants to Convert, and Accepted Islam, and (ii) message impact prediction to assess the emotional influence of agents' responses before deployment. In the initial task, our Enhanced Bi-LSTM model (enhanced with multi-head attention mechanisms and a boosted input capacity of 2500 tokens) achieves a macro-F1 score of 0.76, a balanced accuracy of 74.3%, and an overall accuracy of 90.0% on the full, unbalanced dataset when addressing the 3-class outcome classification task. By conducting an ablation study, we investigate the main source of error, the extremely ambiguous "Wants to Convert" class, and show the resilience of the model on well-defined classes, revealing that accuracy may exceed 99% by simplifying the task to a binary classification. In the second task, a fine-tuned BERT model achieves 80% accuracy in message-level sentiment prediction. The framework enables real-time feedback to refine adaptive responses and guide conversations toward desired outcomes. We assess our proposed framework using a case study of domain-specific conversations (such as religious outreach), supported by a dataset from a partner organization, to demonstrate applicability. The framework's flexibility makes it applicable to customer service, education, and healthcare, where effective communication depends on real-time sentiment analysis and outcome prediction.

ABSTRAK: Kekurangan isyarat bukan lisan dalam komunikasi berasaskan teks memaksa sistem dialog bergantung sepenuhnya pada perubahan sentimen tekstual dinamik membuat kesimpulan tentang niat pengguna dan perubahan emosi. Walau bagaimanapun, ketidakseimbangan kelas yang ketara dan sifat statik dan pasca-hoc model analisis teks sedia ada menghalang sistem dialog daripada mengenali dinamik emosi ini dan meramalkan hasil perbualan pada masa nyata. Kajian ini membentangi rangka kerja pembelajaran mesin hibrid bagi menangani cabaran melalui analisis kesedaran sentimen bersepadu dan ramalan hasil perbualan masa nyata. Pada permulaan, model RoBERTa dilatih terlebih dahulu bagi memberi skor sentimen kepada mesej individu. Ini diperlukan untuk dua tugas seterusnya: (i) klasifikasi hasil perbualan kepada Berminat tentang Islam, Ingin Beranut Islam dan Menerima Islam, dan (ii) ramalan impak mesej bagi menilai pengaruh emosi respons ejen sebelum penggunaan. Melalui tugas awal, dapatan model Peningkatan Bi-LSTM

(dipertingkatkan dengan mekanisme perhatian berbilang kepala dan peningkatan kapasiti input sebanyak 2500 token) mencapai skor makro-F1 sebanyak 0.76, ketepatan seimbang sebanyak 74.3%, dan ketepatan keseluruhan sebanyak 90.0% pada set data penuh tidak seimbang apabila menangani tugas pengelasan hasil 3 kelas. Dengan menjalankan kajian ablasi, kami menyiasat punca ralat utama, kelas "Ingin Menukar" yang sangat samar-samar, dan menunjukkan daya tahan model pada kelas yang ditakrif baik, mendedahkan bahawa ketepatan mungkin melebihi 99% dengan memudahkan tugas kepada pengelasan binari. Menerusi tugas kedua, model BERT yang ditala halus mencapai ketepatan 80% dalam ramalan sentimen peringkat mesej. Rangka kerja ini memudahkan maklum balas masa nyata bagi memperhalusi respons adaptif dan membimbing perbualan ke arah hasil yang diinginkan. Kami menilai rangka kerja yang dicadangkan menggunakan kajian kes yang mengandungi perbualan khusus domain (seperti jangkauan keagamaan). Ini disokong oleh set data daripada organisasi rakan kongsi, bagi menunjukkan kebolegunaan. Fleksibiliti rangka kerja ini boleh digunakan untuk khidmat pelanggan, pendidikan dan penjagaan kesihatan, yang mana komunikasi berkesan bergantung pada analisis sentimen masa nyata dan ramalan hasil.

KEYWORDS: *Sentiment Analysis; Adaptive Conversational Outcome Prediction; Dialogue Systems; RoBERTa; Bi-LSTM.*

1. INTRODUCTION

Text-based dialogue systems are becoming crucial in areas that demand subtle, immediate adjustments to user emotions and intent, spanning from healthcare advice to customer support. These systems exploit advances in deep learning for sentiment analysis, such as models that analyze medical records to determine treatment quality, with high accuracy in classifying unstructured data [1]. These systems depend entirely on monitoring the temporal evolution of message sentiment to understand emotional state and assess interaction evolution, as text-based communication lacks non-verbal information (e.g., tone, gestures). Although sentiment analysis and conversational classification have been explored in static contexts, these techniques remain largely unintegrated into dynamic, real-time systems [2], especially in challenging situations where emotional shifts affect conversational outcomes.

The effectiveness of text-based conversational agents depends on their ability to understand emotional cues and adapt their responses in real time. Traditional approaches such as lexicon-based sentiment analysis struggle to capture context changes in multi-turn dialogues. Transformer-based models (such as BERT [3] and RoBERTa [4]) and sequential models (such as Bi-LSTM [5]) have their own merits but are seldom integrated end-to-end for dual-task optimization. Other problems, such as class imbalance and preprocessing latency, further plague adaptive dialogue systems, and no special datasets exist for outcome-driven dialogue.

In order to overcome these challenges, we propose a hybrid machine learning framework that combines three components: (1) a pre-trained RoBERTa model to assign sentiment scores to individual messages in dialogue conversations, utilizing its robust contextual embeddings; (2) an enhanced Bi-LSTM to classify conversations into Interested in Islam, Wants to Convert, and Accepted Islam by modeling sequential dependencies; and (3) a fine-tuned BERT model to predict the emotional impact of unsent responses, allowing for proactive adjustments. To ensure resilience in dynamic conversational scenarios, we evaluate this approach using metrics such as accuracy, precision, recall, F1-score, balanced accuracy, and macro F1-score.

The main contributions of this research are outlined as follows:

- We present a dawah conversation dataset of labeled conversations (Interested in Islam, Wants to Convert, and Accepted Islam), in which each conversation's message is assigned a sentiment score (positive/negative).
- We propose a hybrid machine learning framework that combines an Enhanced Bi-LSTM architecture for conversation outcome classification (which is enhanced with multi-head attention layers and an increased token limit) and a fine-tuned BERT for response impact prediction. The Enhanced Bi-LSTM achieves 90% accuracy in a challenging 3-class outcome classification, and BERT reaches 80% accuracy in response impact prediction.
- We comprehensively evaluate the proposed framework using accuracy, precision, recall, F1-score, balanced accuracy, and macro F1-score, along with data validation that includes statistical tests and confusion matrices.
- We validate the framework using a domain-specific use case (persuasive conversations), demonstrating its applicability to scenarios requiring sentiment-aware interaction management.

The paper is structured as follows: Section 2 reviews related work in conversational classification and sentiment analysis. Section 3 describes the framework pipeline, dataset, preparation methods, and the proposed methodology in depth. Section 4 presents the performance of the proposed framework, experimental results, and a real-world implementation prototype. Section 5 concludes the paper with future directions.

2. RELATED WORK

This section summarizes current developments in conversational systems and sentiment analysis, going from static text classification to dynamic, goal-oriented dialogue modeling. We divide the current literature into three main categories: dynamic sentiment analysis in conversational systems, conversational outcome prediction, and hybrid architecture and class imbalance.

2.1. Dynamic Sentiment Analysis in Conversational Systems

Sentiment analysis has progressed from lexicon-based approaches to more advanced deep learning models that can consider context. Early studies focused on static documents, such as social media posts and product reviews. Approaches proven effective in static domains, such as classical Support Vector Machines (SVM) and Naïve Bayes, have been shown to work well for tweets [6] and product review classification [7]. Sharin and Kartiwi [8] used sentiment analysis to determine the public's overall attitude toward Technical and Vocational Education and Training (TVET) on social media and showed the applicability of these methods in supporting policy decisions. Similarly, Buntoro et al. [9] showed that machine learning is reliable for gauging public sentiment in high-stakes political contexts by using SVM and Naïve Bayes algorithms in Indonesia's 2019 presidential election.

However, unlike these static applications, dialogue systems pose special challenges. Dialogues, unlike social media posts, require simulating emotional changes over time. Poria et al. [10] state that dialogue turns, as individual text units, fail to capture the “inter-speaker dependency,” which is crucial to understanding context. While Recurrent Neural Networks (RNNs) have improved at capturing sequential dependencies, they can struggle to retain information over long-term contexts. The advent of Transformer-based models, especially BERT [3] and RoBERTa [4], marked a paradigm shift by enabling bidirectional context and capturing the subtle semantic nuances required for real-time dialogue.

2.2. Conversational Outcome Prediction

Predicting the outcome of a conversation, such as a sale, a resolution, or a conversion, differs from sentiment classification. It requires awareness of the strategic connection between goals and dialogue dynamics. In the business field, researchers have used machine learning to predict stock trends [11] and negotiation outcomes [12], primarily using aggregate sentiment features. More recent studies have explored bidirectional reasoning to track how emotional states change and how they affect resulting states [13]. These deep learning methods are superior to the traditional ones as they capture important conversational turning points. However, a critical issue with adaptive prediction systems is that most existing systems are post hoc, meaning that they are evaluated by analyzing completed dialogues [1]. One important limitation of previous studies is the absence of models designed to forecast what happens during the interaction (at the message level) so that real-time intervention can take place.

2.3. Hybrid Architectures and Class Imbalance

Hybrid solutions have been developed to balance the sequential nature of dialogues with the granular results of sentiment analysis. Pure transformer models (such as BERT) are great for semantic embedding but can be computationally expensive when used for real-time sequence modeling of long conversations. However, Bi-LSTMs are efficient sequence modelers [5] and lack the depth of contextual knowledge of pre-trained models. Recent studies have shown that combining these methods improves performance. Basiri et al. [7] showed that hybrid models, that is, attention models combined with Bi-LSTMs, show a significant improvement over standalone models in complex sentiment tasks. Additionally, the dataset is imbalanced, meaning the class “successful” is not often present, which is a problem that needs to be addressed. Extensive studies [14] show that to achieve good performance with deep neural networks trained on skewed data, dynamic sampling and class weighting are crucial. We build on this foundation by combining the strong embedding power of RoBERTa with our Enhanced Bi-LSTM, which is specifically designed to handle the high class imbalance in persuasive dialogue datasets.

2.4 Research Gaps and Novelty

While there is extensive literature on conversational AI, much work remains to be done on real-time, sentiment-aware dialogue systems. First, interaction is based on text only, without physical non-verbal cues like vocal tone and facial expressions; most high-accuracy sentiment models [2], [6], [8] and [9] are static, working on individual texts, such as reviews or tweets, rather than being dynamic and capturing the turn-taking interactions of live dialogue. Second, current outcome prediction models are mainly used as post-hoc analysis on finished logs [1] and cannot intervene in real time to guide agents turn by turn. Third, hybrid architectures are available [7], but few are optimized or evaluated on highly imbalanced, goal-oriented conversational datasets with rare goal outcomes.

This study directly bridges these gaps by shifting the paradigm from passive post-hoc evaluation to active real-time adaptation. Our novelty lies in integrating granular sentiment scoring and sequential outcome prediction into a unified feedback loop. The framework combines RoBERTa's message-level sentiment analysis with an enhanced Bi-LSTM for trajectory classification and a fine-tuned BERT for pre-send impact analysis, enabling dialogue systems to compensate for the lack of non-verbal cues and proactively guide dialogues toward desired outcomes.

3. METHODOLOGY

This section details the proposed framework, first describing the pipeline and dataset. Then, we discuss the methodology of the main components of our proposed framework: sentiment scoring, conversation outcome classification, and message impact prediction, along with the proposed models for each task. Lastly, we define the evaluation metrics used in this study.

3.1. Framework Overview and Pipeline

Fig. 1 illustrates the proposed framework pipeline, which integrates conversation outcome classification and message impact prediction for adaptive dialogue management. The process starts with a common base, then splits into two interdependent tracks.

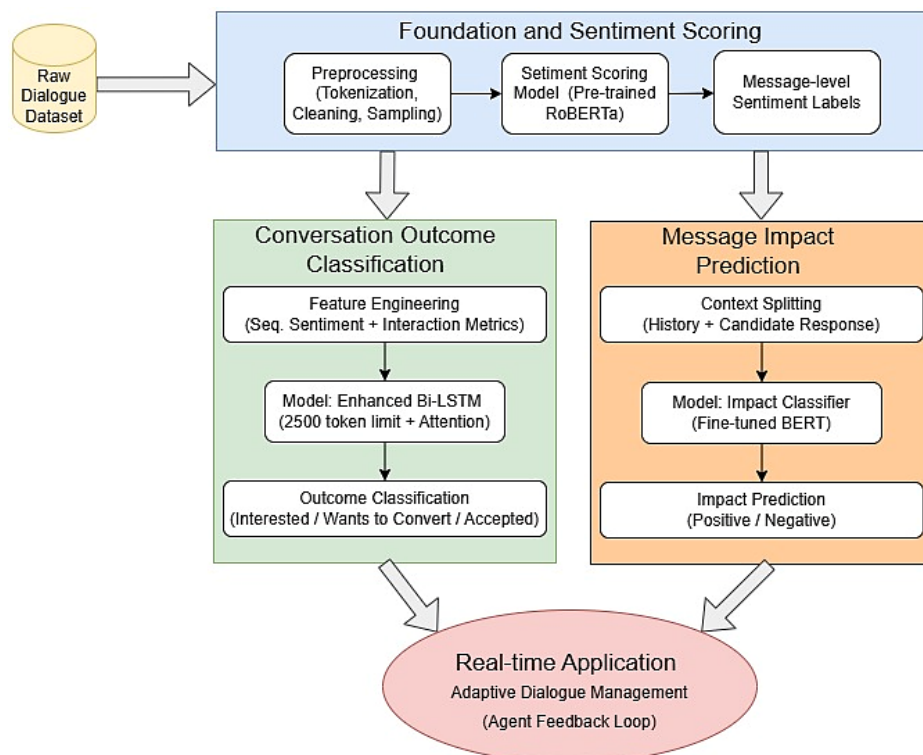


Figure 1. The proposed framework's overall pipeline.

1) Foundation and Sentiment Scoring

- **Preprocessing:** To guarantee data quality, the raw dialogue dataset is first subjected to tokenization, noise reduction, and text standardization.
- **Sentiment Scoring:** A pre-trained RoBERTa model assigns sentiment scores (positive/negative) to individual messages.
- **Message-Level Sentiment:** These granular scores are aggregated to track the emotional trajectory of the conversation turn by turn.

2) Track 1: Conversation Outcome Classification

- Engineered features (such as response latency and keyword density) and sequential sentiment scores are utilized in this track.

- Our enhanced Bi-LSTM model is fed with these inputs to analyze the evolving context and classify the conversation into Interested in Islam, Wants to Convert, or Accepted Islam.

3) Track 2: Message Impact Prediction

- The agent's proposed response, its calculated sentiment score, the visitor's prior conversational context, and the outcome classification of the current ongoing conversation are considered inputs for this track to perform a predictive analysis before a response is sent.
- A fine-tuned BERT model is fed with these inputs to predict whether the proposed response will positively or negatively affect the direction of the conversation.

The pipeline dynamically links these tracks towards real-time adaptive dialogue management. Message-level sentiment tracking guides the conversation outcome classification, while message impact prediction enables agents to iteratively improve their proposed responses to steer that outcome positively.

3.2. Dataset Overview and Preparation

For this study, we use a unique dataset of 25,734 English conversations. The data were gathered from 2019-2022 from Rukn Al-hewar [15], an online platform for conversations about Islam. There are 5,677 (22.1%) conversations about "Accepted Islam", 17,553 (68.2%) conversations about "Interested in Islam", and 2,504 (9.7%) conversations about "Wants to Convert". This indicates class imbalance, as "Interested in Islam" represents 68.2% of conversations, posing challenges for model generalization.

Table 1. Types and numbers of conversations in the dataset

Type of conversation	Number of conversations
Accepted Islam	5,677
Interested in Islam	17,553
Wants to Convert	2,504
Total number of conversations	25,734

Table 2. Key statistical features of the dataset

Type of conversation	Number of conversations
Number of visitors	21,715
Number of agents	120
Number of words identified	26,762,112
Average number of words per conversation	967
Average number of words per question	9

Table 2 provides summary statistics for key features. The dataset contains 26.7 million words, 967 words per conversation, and an average of 9 words per question. The dataset includes 120 agents and 21,715 visitors engaged in conversations. Fig. 2 shows the differences in conversation length for each class. The longest dialogues concern Accepted Islam, while dialogues about Interested in Islam and Wants to Convert are shorter and more consistent. Therefore, pre-processing and sampling methods are needed to avoid training biases caused by this structural imbalance. Because of class imbalance and dialogue length, these factors must be addressed separately to ensure model robustness; however, the dataset size and annotations are domain-specific and can serve as a reliable resource for training a sentiment-aware dialogue system.

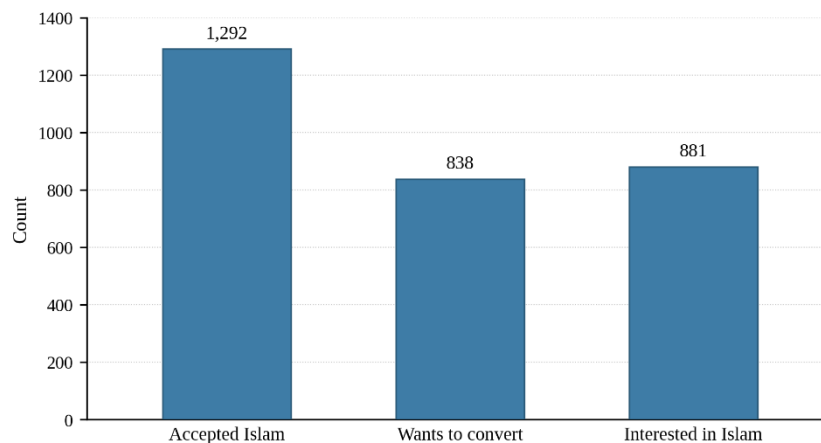


Figure 2. Average word count for each conversation class

To ensure analytical clarity and address potential semantic overlaps, explicit operational boundaries were defined for the three target categories:

- *Interested in Islam*: Exploratory dialogues where the visitor asks general, informational, or theological questions without expressing explicit intent to embrace the faith (e.g., *Can you explain how prayer works?*).
- *Wants to Convert*: Dialogues where the visitor explicitly expresses personal alignment, conviction, or a direct desire to become Muslim, but has not yet pronounced the testimony of faith during the session (e.g., *I believe Islam is the truth and I want to convert; what are the steps?*).
- *Accepted Islam*: Dialogues where the visitor formally completes the conversion process during the live chat by uttering or typing the testimony of faith (e.g., *I declare that there is no God but Allah and Muhammad is His messenger*).

The dataset's ground-truth labels were originally assigned by trained counselors at the partner organization (Rukn Al-hewar) after the session, using standard administrative protocols. The "Accepted Islam" label strictly required an explicit, logged declaration of the Shahada. To rigorously validate the quality of these primary labels and measure human consistency on ambiguous cases, a randomized sample of $N = 1,000$ conversations was independently re-annotated by two domain experts following a formal guidelines codebook. Inter-annotator agreement yielded a Fleiss' Kappa score of 0.82, indicating strong agreement and confirming the reliability of the ground-truth annotations.

To ensure analytical consistency and reduce noise, we prepared the data using standard preprocessing techniques. Text standardization involved lowercasing all characters and removing punctuation, special symbols, and non-alphanumeric elements. Lemmatization was applied to minimize lexical redundancy (e.g., "converting" → "convert"). Domain-specific stop-words (e.g., repetitive greetings such as "Assalamu Alaikum"), URLs, and emojis were removed to concentrate on meaningful dialogue content.

The dataset was partitioned into 70% training, 15% validation, and 15% test splits using a fixed global random seed (seed = 42). To address class imbalance during model development, we evaluated the Synthetic Minority Over-sampling Technique (SMOTE) and undersampling strategies. To strictly prevent data leakage between evaluation subsets, resampling techniques (including SMOTE) were applied exclusively to the 70% training split after partitioning. The

validation (15%) and held-out test split (15%, comprising 3,860 conversations) remained completely unmanipulated to evaluate true real-world operational performance.

3.3. Sentiment Scoring for Individual Messages

In this study, the two main tasks are conversation outcome classification (in Section 3.4) and message impact prediction (in Section 3.5); for both, each unannotated question and answer in the dataset is assigned a sentiment score (positive or negative) as a basis. The two transformer models tested were: (i) RoBERTa [4], an optimized version of BERT [3] known to capture subtle linguistic patterns, and (ii) DistilBERT, a transformer with minimal performance trade-off but drastically reduced computation time. As evaluated in Section 4, these empirical performance metrics determine which of RoBERTa and DistilBERT perform better. In our approach, we fully retain the original conversational context while keeping the pipeline as modular as possible, with sentiment scoring independent of the pre-processing step.

3.4. Conversation Outcome Classification

This subsection outlines the methodology of the conversation outcome classification. We first discuss the domain-specific features and then explain our Enhanced Bi-LSTM architecture. Finally, we highlight the importance of simulating real-time conversation, where classification happens dynamically rather than waiting for the dialogue to conclude.

Table 3. Domain-specific features and their descriptions

Feature	Description
Sentiment Score	Sentiment scores associated with all conversation messages, calculated using the RoBERTa model.
Average Sentiment Score	Average sentiment scores for questions, answers, and the overall conversation, offering insights into the general tone and engagement levels.
Number of Sessions	Tracks the total sessions within a conversation to measure the extent of repeated interactions.
Conversation Length	Represents the total number of words across all messages in a conversation.
Average Message Length	Calculates the average number of words per message, providing insights into the verbosity of participants.
Number of Messages	Counts the total number of messages in each conversation to reflect the level of engagement.
Duration	Measures the total time span of a conversation, from the first to the last message, to understand the temporal dynamics of interactions.

3.4.1. Feature Engineering

Three key aspects – sentiment metrics, interaction patterns, and temporal dynamics – were the focus of domain-specific features engineered to represent conversational dynamics, as shown in Table 3. Sentiment metrics include average sentiment scores for questions, answers, and the overall conversation, as well as per-message sentiment scores produced by the RoBERTa model. Interaction patterns such as conversation length and message frequency quantify engagement. Temporal dynamics such as response latency and keyword density track behavioral shifts.

3.4.2. Full-Conversation Classification Using Enhanced Bi-LSTM

To address the limitations of standard sequential models, we propose an Enhanced Bi-LSTM architecture. Our improved model incorporates multi-head attention mechanisms to dynamically weigh the significance of specific conversation turns, whereas standard Bi-LSTMs struggle with long-term dependencies. Additionally, we increased the input capacity

from the typical 512 or 1024 tokens to 2,500 tokens, enabling the model to analyze long, multi-turn conversations without losing any data. Our proposed model is optimized to classify full conversations into the three target classes using pre-trained GloVe embeddings.

3.4.3. Real-time Classification Simulation

Instead of waiting for the dialogue to end, the system in a live deployment must predict outcomes dynamically as the conversation progresses. When given partial context, models trained only on completed conversations often fall short. Therefore, we truncated the dataset at different completion milestones (50%, 60%, ... 100%) to simulate this real-time scenario and evaluate the robustness of our Enhanced Bi-LSTM. This guarantees that the model can act in advance and provide agents with actionable insights while the conversation is still active.

3.5. Message Impact Prediction

Effective dialogue systems should proactively analyze how individual messages impact interaction trajectories in addition to classifying full conversations into final outcomes. Traditional classification techniques evaluate completed conversations, which restricts their usefulness for real-time adaptation. Predicting whether a possible agent message will positively or negatively affect the conversation flow before it is sent is an essential challenge addressed in this subsection. This framework enables iterative improvement of agent strategies by isolating the causal role of specific responses and ensuring alignment with goals.

3.5.1. Target Label Definition

We created a new target label based on the conversational changes seen before and after a Daei message to understand the impact of these messages. This required examining two versions of the conversations:

- 1) **Pre-Daei Message Context:** this refers to the entire conversation up to the Daei message.
- 2) **Post-Daei Message Context:** this refers to the entire conversation (at most 5 messages from the visitor) after the Daei message.

Table 4. Daei message impact based on the transition of pre- and post-message context classes and their probabilities

Daei Message Impact	Pre-Daei Message Context	Post-Daei Message Context
Positive	Interested in Islam	Interested in Islam (probability increased)
	Interested in Islam	Wants to Convert
	Interested in Islam	Accepted Islam
	Wants to Convert	Wants to Convert (probability increased)
	Wants to Convert	Accepted Islam
	Accepted Islam	Accepted Islam (probability increased)
Negative	Interested in Islam	Interested in Islam (probability decreased)
	Wants to Convert	Interested in Islam
	Wants to Convert	Wants to Convert (probability decreased)
	Accepted Islam	Interested in Islam
	Accepted Islam	Wants to Convert
	Accepted Islam	Accepted Islam (probability decreased)

The predicted classes and their probabilities for both contexts stated above were calculated using the Enhanced Bi-LSTM as the classifier model. The sentiment score of the Daei message—which was computed using RoBERTa—was also included to evaluate how the tone and sentiment of the Daei message affected the classification. The Daei message’s impact was

assigned a positive or negative label based on the transition of the classes and their probabilities from the pre-Daei message context to the post-Daei message context, as illustrated in Table 4.

To ensure that deriving impact labels from model probability transitions does not introduce circularity, we subjected target labels to an independent human reliability audit. Domain experts manually evaluated a sample of dialogue transition segments to verify alignment between automated state shifts and true human-perceived message impact. Section 4.3 reports the detailed reliability findings.

Using the sentiment score of the Daei message and the classes and their probabilities of the pre- and post-Daei messages, we implemented and trained two comparative architectures - a fine-tuned BERT and a standard Bi-LSTM – to predict the Daei message impact. The performance of these two methods is thoroughly evaluated in Section 4.3.

3.5.2. Data Preparation

The following seed-based filtering method was used to prepare the dataset for the message impact prediction task:

- We split conversations into segments in which each segment contained a single Daei message as the focal point. We ensured that each segment had at least five messages from the visitor before the focal point and at least one message after it.
- We combined Daei messages into a single focal point if they appeared consecutively, and we averaged their sentiment scores.

This seed-based filtering method yields a dataset with 122,056 files in the “Accepted Islam” category, 259,720 files in the “Interested in Islam” category, and 32,045 files in the “Wants to Convert” category.

3.6. Evaluation Metrics

To rigorously evaluate model performance across classification and impact prediction tasks, particularly under class imbalance, we use standard quantitative performance metrics [16]. Given True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), the metrics are defined as follows:

- Overall accuracy measures the total proportion of correct predictions:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- Precision and recall measure positive predictive value and sensitivity, respectively [17]:

$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN} \quad (2)$$

- Per-class F1-score is a harmonic mean of precision and recall for class i [17]:

$$F1_i = \frac{2 \times Precision_i \times Recall_i}{Precision_i + Recall_i} \quad (3)$$

- Macro F1-score is the unweighted mean of F1-scores across all C classes, evaluating performance equally across minority and majority categories [16]:

$$Macro - F1 = \frac{1}{C} \sum_{i=1}^C F1_i \quad (4)$$

- Balanced accuracy is the unweighted average of recall across all C classes to account for class imbalance [18]:

$$\text{Balanced Accuracy} = \frac{1}{C} \sum_{i=1}^C \text{Recall}_i \quad (5)$$

3.7. Ethical Considerations and Privacy Protection

Because conversational datasets are typically text-based, they may include personal conversations; we implemented careful data privacy and ethical practices in data collection, preprocessing, and model execution. We thoroughly de-identified every raw conversation log submitted by the partner organization before analysis. We systematically removed Personally Identifiable Information (PII), including personal names, phone numbers, e-mail addresses, IP logs, and specific geographic locations, through automated regex filtering with manual spot-checking. Data collection followed the partner platform's terms of service, which allow interactions to be logged under administrative consent for quality assurance and research.

To address potential concerns regarding manipulative dialogue management, the proposed framework is explicitly architected as a Human-in-the-Loop (HITL) decision-support system rather than a fully autonomous agent. Real-time sentiment scores and message impact predictions serve as advisory feedback within the agent interface, ensuring that human counselors maintain full agency and ethical oversight over all generated responses before deployment.

4. RESULTS AND DISCUSSION

This section presents the experimental findings for sentiment scoring, message impact prediction, and conversation outcome classification. Quantitative performance and comparative analyses validate our framework's ability to improve real-time dialogue.

4.1. Sentiment Scoring of Individual Messages

We analyzed two models, DistilBERT and RoBERTa, to predict the sentiment of specific questions and responses. Table 5 summarizes the performance of both models for each of the three categories based on average sentiment scores. Fig. 3 and Fig. 4 visually compare the average sentiment scores for RoBERTa and DistilBERT, respectively. To determine classification superiority, we manually annotated a randomized gold-standard evaluation subset of $N=1,000$ individual dialogue messages (Positive, Negative, and Neutral) with two domain experts. We used standard classification metrics for both models to assess this benchmark. The quantitative evaluation shows that RoBERTa outperforms DistilBERT in all metrics: accuracy of 88.4%, precision of 0.87, recall of 0.86, and Macro-F1 score of 0.86, compared with DistilBERT's 81.2%, 0.80, 0.79, and 0.79, respectively. The contextual embeddings generated by RoBERTa were especially good at capturing conversational nuances and domain-specific terminology. As a result, we chose RoBERTa as the backbone sentiment-scoring model for all downstream tasks.

Message sentiment scores range continuously between -1.0 (highly negative) and $+1.0$ (highly positive). Scores strictly above $+0.15$ are classified as positive, scores below -0.15 as negative, and scores within the interval as neutral. The neutral boundary was established empirically via grid validation on a representative sample of $N = 1000$ dialogue messages. In conversational contexts, neutral turns predominantly consist of factual inquiries, links, or procedural statements (e.g., asking for website links). A grid search across threshold intervals from to demonstrated that optimally separates purely procedural messages from subtle emotional shifts, preventing minor polarity fluctuations from misclassifying neutral inquiries.

We performed sentiment analysis using the RoBERTa model and computed sentiment scores for each message in the text conversation. Rather than examining the entire conversation, we evaluated the sentiment of individual messages. This gave us a basic understanding of how people convey their views and feelings through messaging. Tables 6 and 7 illustrate some examples of sentiment scores for visitor messages and Daei messages using RoBERTa, respectively.

Table 5. Comparing DistilBERT and RoBERTa for predicting average sentiment scores

Item	RoBERTa	DistilBERT
Accepted Islam Answers	0.3625	0.3031
Accepted Islam Questions	0.2529	0.2535
Wants to Convert to Islam Answers	0.2872	0.2541
Wants to Convert to Islam Questions	0.1526	0.1918
Interested in Islam Answers	0.2131	0.2084
Interested in Islam Questions	0.0853	0.1386

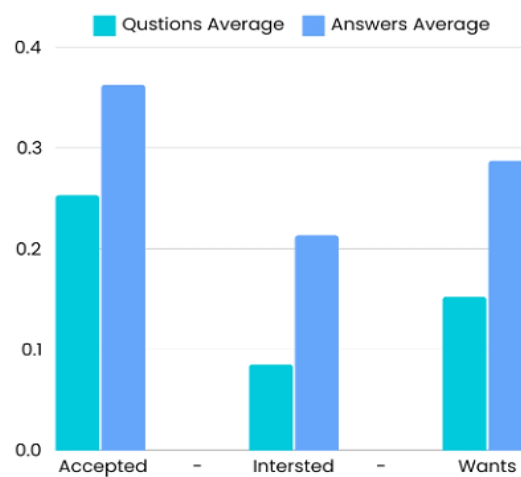


Figure 3. RoBERTa's average sentiment score performance

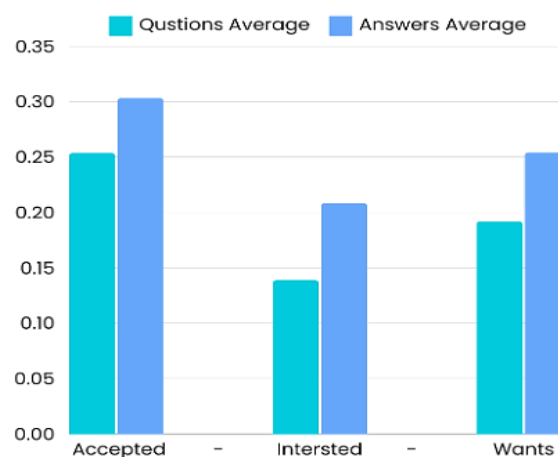


Figure 4. DistilBERT's average sentiment score performance

Table 6. Sentiment Scores for visitor messages using RoBERTa

Visitor Message	RoBERTa Scores
"Yes I would love that."	0.9696
"Yeah, I studied Islam well and understand a lot of it."	0.7180
"So ur logic dont make sense."	-0.7479
" I'm not interested in converting."	-0.4887
"What is the importance of Salah"	0.107

Table 7. Sentiment Scores for Daei messages using RoBERTa

Visitor Message	RoBERTa Scores
"We are happy to help; keep coming to ask, please."	0.9435
"When you take shahadah, all that is forgiven by Allah, and you become new like the time you were born"	0.7180
"The vast universe itself, with its meticulous order and fine-tuning, is proof that God exists"	0.8500
"So would you like for me to share the basics of Islam with you?"	0.0411
"Correct, and that's what the Qur'an says"	0.053

We used scatter charts to display the sentiment scores for the three categories (Accepted Islam, Interested in Islam, and Wants to Convert) in order to better analyze the model's performance. Each scatter chart shows the sentiment scores for each question and response, along with their averages and detected outliers.

With very few outliers, the Accepted Islam scatter chart in Fig. 5 displays the most consistent results. This is expected because conversations in this category are usually clear and direct, as visitors have already accepted Islam, making it simpler for the algorithm to provide accurate sentiment scores.

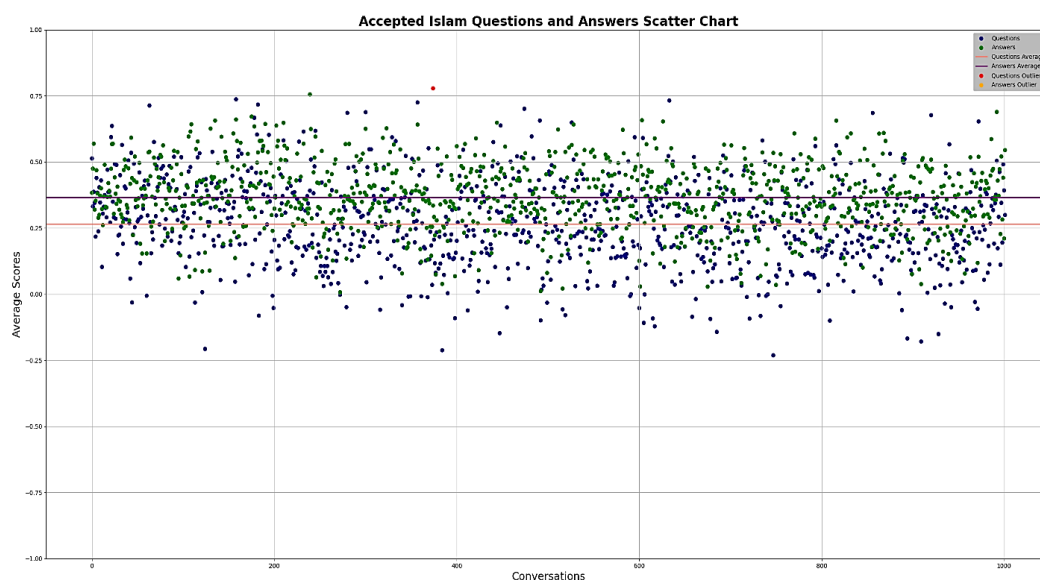


Figure 5. Scatter chart of Accepted Islam questions and answers

Compared to Accepted Islam, the Wants to Convert scatter chart in Fig. 6 shows slightly more variability, but it still performs well with few outliers. This is because, unlike the Accepted Islam category, interactions in this category often feature participants expressing a strong interest or intent to convert, resulting in more varied emotional expressions.

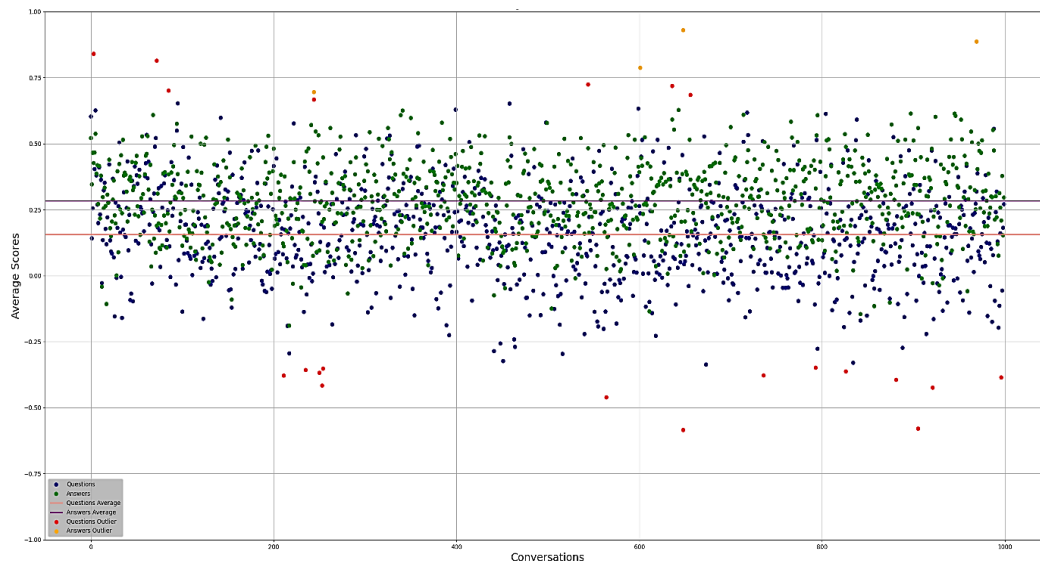


Figure 6. Scatter chart of Wants to Convert questions and answers

The scatter chart for Interested in Islam in Fig. 7 shows the most variability and outliers among the three categories. This is expected, as conversations in this category are more exploratory, with participants asking questions or expressing curiosity without clear intent or decisions. This natural variability in conversations results in a wider range of sentiment scores.

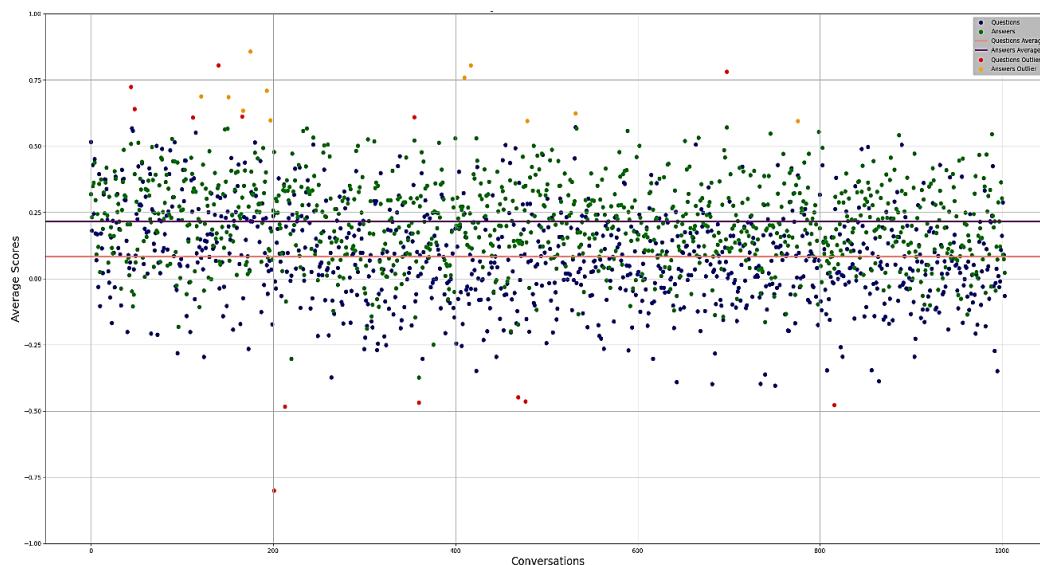


Figure 7. Scatter chart of Interested in Islam questions and answers

4.2. Conversation Outcome Classification

Comprehensive experimental results for conversation classification are shown in this subsection. Using the entire 3-class conversational dataset, we methodically assess the performance of three deep learning algorithms: BERT, Bi-LSTM, and our Enhanced Bi-LSTM. We then thoroughly validate our main model and analyze the effects of several approaches to resolving class imbalance. To confirm the cause of classification difficulty, we also perform an ablation study. Finally, we evaluate the models' adaptability and performance when trained on incrementally truncated conversation lengths, simulating real-time interaction scenarios.

To categorize the conversation into one of the three groups, we initially examined three models on complete conversations without breaking them up into shorter segments.

- **BERT Model:** The accuracy of the pretrained BERT model was 74.8%. BERT's results highlighted its shortcomings for our dataset. In particular, 65% of our conversations were longer than the 512 tokens that BERT can handle. As illustrated in Fig. 8, this limitation reduced its effectiveness. To address this sequence length limitation, we also evaluated Longformer-base (supporting up to 4,096 tokens). However, Longformer exhibited severe training instability, excessive per-epoch latency, and early validation plateauing. Consequently, long-context transformers proved impractical for this task.

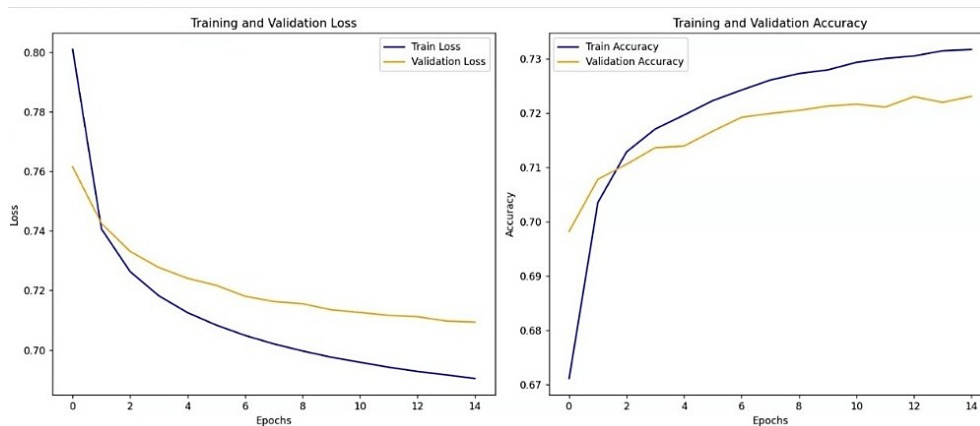


Figure 8. Performance of the pretrained BERT model, showing training and validation loss and accuracy over epochs.

- **Bi-LSTM Model:** A standard Bidirectional LSTM model with a 1024 token limitation demonstrated its ability to handle longer sequences with an accuracy of 80% on the test set (Fig. 9).

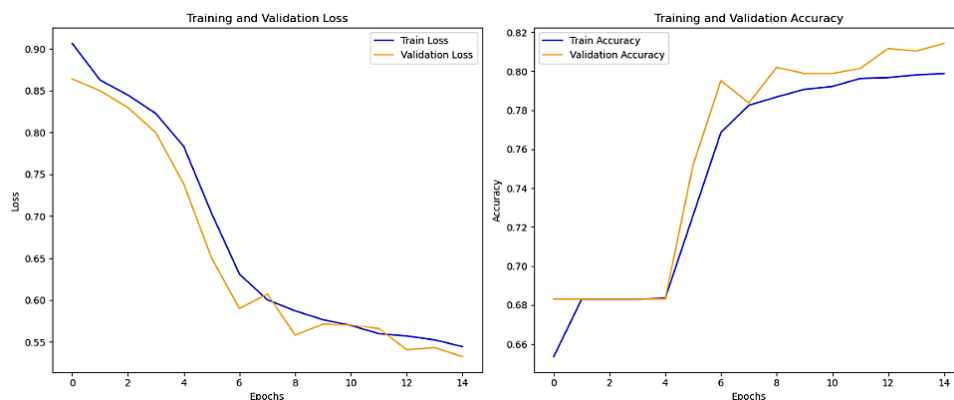


Figure 9. The pretrained Bi-LSTM model's performance on conversations with a token limit of 1024 demonstrates training and validation loss and accuracy over epochs.

- **Enhanced Bi-LSTM Model:** We improved the Bi-LSTM by adding attention layers, increasing the token limit to 2500, and adjusting the learning rate. With a test accuracy of 90%, this model achieved our best 3-class result. These adjustments led to significant improvements, as shown in Fig. 10, with accuracy remaining close to 90% and loss decreasing consistently.

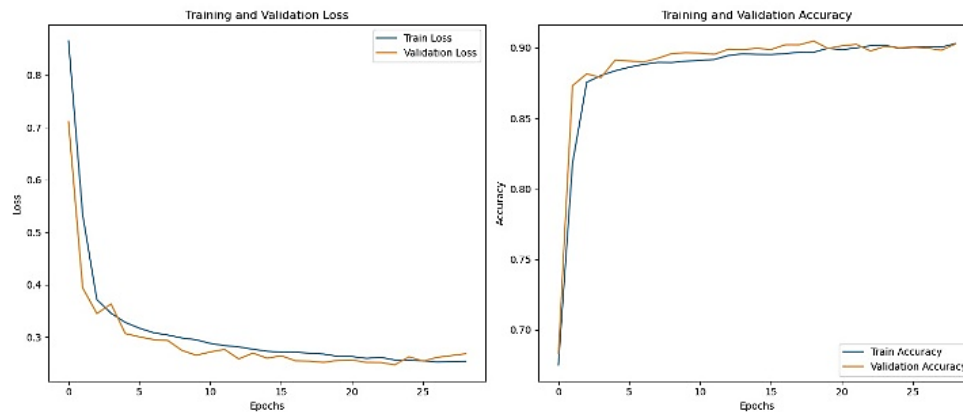


Figure 10. Performance of the Enhanced Bi-LSTM model with training and validation loss and accuracy over epochs after raising the token limit to 2500.

Table 8. Classification results of the Enhanced Bi-LSTM (3-Class Model)

Classes	Precision	Recall	F1-Score
Interested in Islam	0.90	0.96	0.93
Wants to Convert	0.57	0.30	0.39
Accepted Islam	0.96	0.97	0.97
Overall Accuracy	0.900		
Balanced Accuracy	0.743		
Macro-F1 Score	0.763		

The detailed classification results in Table 8 show that while the model achieves an overall accuracy of 90.0%, its balanced accuracy is 74.3% and its macro-F1 score is 0.763. This breakdown shows that overall accuracy is driven largely by strong performance on the majority classes (“Interested in Islam” with F1 = 0.93 and “Accepted Islam” with F1 = 0.97). Conversely, the model struggled with the minority “Wants to Convert” category, achieving a low recall of 0.30 and an F1-score of 0.39, indicating frequent misclassifications due to semantic overlap with “Interested in Islam”. We further analyze this challenge in the following section.

Table 9. Hyperparameters, performance baseline comparison, and execution timing across models

Hyperparameters	BERT	Bi-LSTM	Enhanced Bi-LSTM
Learning rate	5.00e-05	5.00e-05	1.00e-04
Batch size	64	64	128
Number of layers	12 (transformers)	4 layers	4 + Attention layers
Hidden Units / Dimension	768	128 (per direction)	250 (per direction)
Token limit	512	1024	2500
Optimizer	AdamW	Adam	Adam
Number of epochs	15	15	50
Dropout Rate	0.1	0.2	0.3
Early Stopping	Patience = 3	Patience = 5	Patience = 5
Overall Accuracy	74.8%	80%	90%
Training Time/Epoch	~14 min	~3 min	~6 min
Avg. Inference Latency	42 ms	35 ms	85 ms
Additional features	Pretrained embeddings	GloVe embeddings	GloVe embeddings, multi-head attention

Comparing these architectures demonstrates the component-wise contributions of our framework: extending token capacity from 512 (BERT) to 1,024 (Bi-LSTM) yields a 5.2% accuracy gain, while adding multi-head attention and expanding context to 2,500 tokens in the Enhanced Bi-LSTM provides an additional 10.0% accuracy increase (reaching 90.0% accuracy, 74.3% balanced accuracy, and 0.763 macro-F1).

Table 9 summarizes the hyperparameter configurations, architectural dimensions, performance baseline metrics, and execution timing across all evaluated models to establish a consistent experimental setup (with a detailed discussion of real-time latency and computational efficiency in Section 4.5).

4.2.1. Data Validation and Robustness Analysis

We validated the performance of our 90% accuracy 3-class model and transparently analyzed its limitations by presenting and discussing the confusion matrix and the statistical test of agreement (Cohen's Kappa). Then, we applied several techniques to address the imbalance identified in the confusion matrix.

To visualize the challenge presented in Table 7, we reported the confusion matrix (Fig. 11) for the Enhanced Bi-LSTM model's predictions on the 15% test split, which comprises 3860 conversations (2633 "Interested", 376 "Wants to Convert", and 851 "Accepted").

The confusion matrix in Fig. 11 shows that the model is very good at distinguishing between "Interested in Islam" (96.0% recall) and "Accepted Islam" (96.9% recall). Between these two major classes, there is relatively little uncertainty. But the major reason for the model's error is the class "Wants to Convert", where 66.5% of the conversations (250 conversations) are misclassified as "Interested in Islam". This strongly suggests that the language dimensions of "interested" and "wants to convert" are very similar, and this class is inherently ambiguous.

We computed Cohen's Kappa score, which evaluates inter-rater agreement while accounting for chance, to further validate the model's performance beyond simple accuracy. The model received a Cohen's Kappa score of 0.77 on this classification task. This indicates a high degree of agreement, demonstrating that the model's 90% accuracy is reliable and not a result of the class imbalance.

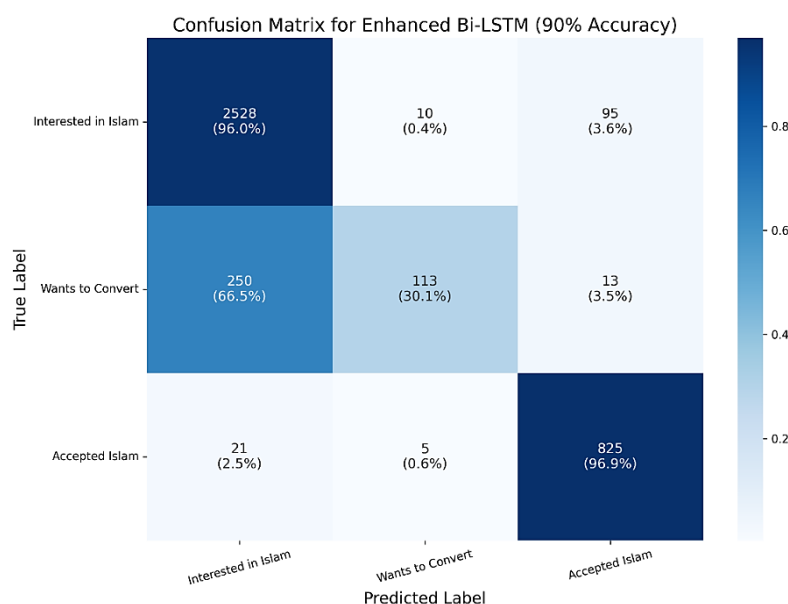


Figure 11. Confusion Matrix for 3-Class Enhanced Bi-LSTM (90% Accuracy)

Additionally, we tested several strategies to address the imbalance found in the confusion matrix, including the following.

a) Using class weights:

During training, we assigned the minority class “Wants to Convert” higher weights. To ensure that the model focused more on the minority class, this adjustment updated the loss function to penalize misclassification of this class more heavily. The performance of this method is shown in Table 10. The model’s recall for “Wants to Convert” improved to 0.71, but its precision for this class stayed low at 0.38, resulting in an F1-score of 0.50. The model’s overall accuracy of 86% shows moderate effectiveness in balancing performance across classes.

Table 10. Enhanced Bi-LSTM’s performance after applying class weights

Classes	Precision	Recall	F1-Score
Interested in Islam	0.96	0.84	0.89
Wants to Convert	0.38	0.71	0.50
Accepted Islam	0.97	0.97	0.97
Accuracy	0.86		

b) Under-sampling the dominant classes:

To balance the dataset, the number of samples in the main categories (“Interested in Islam” and “Accepted Islam”) was reduced. This strategy improved the model’s ability to classify the minority category, “Wants to Convert.” As a result, the model attained an overall accuracy of 84%. Table 11 shows that the precision and recall for “Wants to Convert” were both 0.76, yielding an F1-score of 0.76. F1-scores for “Interested in Islam” and “Accepted Islam” were 0.77 and 0.97, respectively. These findings indicate that while undersampling reduced the model’s bias toward dominant categories, the smaller dataset size marginally limited the model’s learning potential.

Table 11. Performance of the Enhanced Bi-LSTM when the dominant classes were undersampled.

Classes	Precision	Recall	F1-Score
Interested in Islam	0.77	0.78	0.77
Wants to Convert	0.77	0.76	0.76
Accepted Islam	0.97	0.97	0.97
Accuracy	0.84		

c) Oversampling the minority class:

To increase the representation of the “Wants to Convert” category, we employed oversampling by duplicating its data and generating synthetic samples. This method ensured the model was exposed to more minority-class samples during training. According to Table 12, the recall for “Wants to Convert” increased to 0.59, but the precision stayed low at 0.37, yielding an F1-score of 0.45. High F1-scores of 0.90 and 0.96 were attained by the “Interested in Islam” and “Accepted Islam” categories, respectively. Although the precision-recall tradeoff for this category still presented difficulties, these results demonstrate that oversampling improved the model’s capacity to identify the minority class.

Table 12. Performance of Enhanced Bi-LSTM after oversampling the “wants to convert” class.

Classes	Precision	Recall	F1-Score
Interested in Islam	0.94	0.86	0.90
Wants to Convert	0.37	0.59	0.45
Accepted Islam	0.97	0.95	0.96
Accuracy	0.86		

d) Merging classes:

In this technique, we consolidated the “Accepted Islam” and “Wants to Convert” categories into a single class. The model performed well on the entire dataset while focusing more on underrepresented data. The model attained an accuracy of 91%, with detailed results shown in Table 13.

Table 13. Performance of Enhanced Bi-LSTM after merging “accepted Islam” and “wants to convert” classes.

Classes	Precision	Recall	F1-Score
Interested in Islam	0.90	0.97	0.94
Merged (Wants to Convert & accepted Islam)	0.93	0.77	0.85
Accuracy	0.91		

The results of these experiments demonstrate that while typical imbalance strategies can increase minority-class recall, they tend to reduce overall model accuracy and do not address the fundamental problem of linguistic ambiguity.

4.2.2. Ablation Study: Removing the Ambiguous Class

Because of its substantial linguistic overlap with the “Interested” class, the study in the preceding section clearly confirms that the “Wants to Convert” class is the main source of model error. We conducted an ablation study to test this hypothesis by eliminating this one problematic class and reducing the task to a binary classification (Interested vs. Accepted).

Table 14. Performance of Enhanced Bi-LSTM (2-Class Ablation Study)

Classes	Precision	Recall	F1-Score
Interested in Islam	0.99	0.99	0.99
Accepted Islam	0.97	0.98	0.98
Overall Accuracy	0.990		
Balanced Accuracy	0.985		
Macro-F1 Score	0.985		

After removing the “Wants to Convert” class, we retrained our Enhanced Bi-LSTM model on the remaining two categories (2-class ablation study). Table 14 summarizes the findings, and Fig. 12 shows the results. The model attained an overall accuracy of 99%, a balanced accuracy of 98.5%, and a macro-F1 score of 0.985. While this high performance demonstrates the Enhanced Bi-LSTM’s capacity to discriminate between distinct conversational endpoints (Interested vs. Accepted), this binary result must be interpreted cautiously. Simplifying the task to a two-class problem removes the primary source of linguistic overlap but does not resolve the real-world three-class classification task. Thus, this ablation study is not presented as evidence of strong performance for the full system, but rather as a diagnostic experiment confirming that semantic ambiguity in the “Wants to Convert” class is the primary bottleneck of the 3-class architecture. This diagnostic ablation isolates the contribution of architectural

components from dataset characteristics, quantitatively proving that our sequence length expansion and attention layers effectively solved trajectory modeling, leaving semantic boundary overlap as the sole bottleneck.

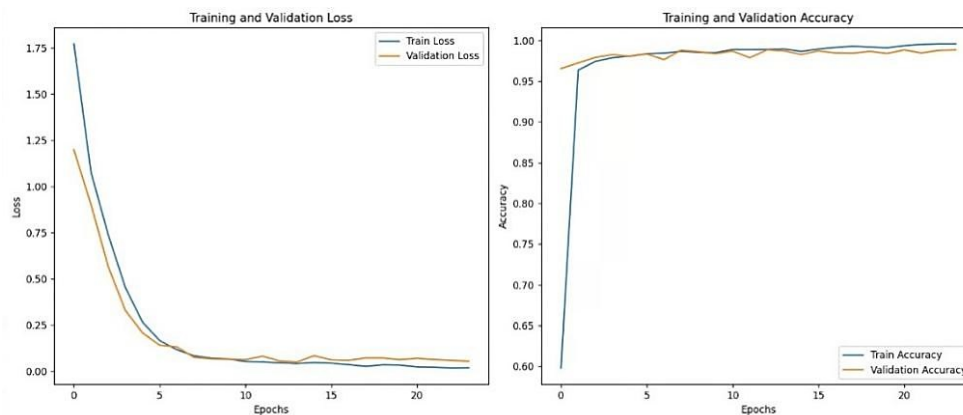


Figure 12. Loss and accuracy curves for training and validation after the removal of the “wants to convert” class (2-Class Ablation Study).

4.2.3. Real-time Classification Simulation: Effect of Conversation Length

To evaluate the model’s performance across conversations of varying lengths (i.e., to more realistically represent real-world chats), we used a second approach to train the model with split data (50%, 60%, up to 100%). We utilized this splitting method exclusively for the Enhanced Bi-LSTM model.

Table 15. Enhanced Bi-LSTM’s performance (training with split conversations in the 3-class classification task).

Classes	Precision	Recall	F1-Score
Interested in Islam	0.90	0.94	0.92
Wants to Convert	0.49	0.41	0.45
Accepted Islam	0.89	0.83	0.86
Overall Accuracy	0.860		
Balanced Accuracy	0.727		
Macro-F1 Score	0.743		



Figure 13. Performance of Enhanced Bi-LSTM (second approach: training on split conversations, 3-class classification task).

When evaluated on split conversation milestones to simulate real-time interaction, the model achieved an overall accuracy of 86%, a balanced accuracy of 72.7%, and a macro-F1 score of 0.743 on the 3-class task. As detailed in Table 15, while the model maintains solid dynamic classification across varying dialogue lengths for the majority classes (Interested F1 = 0.92, Accepted F1 = 0.86), performance on the minority “Wants to Convert” category remains lower (Recall = 0.41, F1 = 0.45) due to mid-conversation semantic ambiguity. Fig. 13’s training and validation plots demonstrate consistent training progress.

As with our full-conversation model, we applied the same ablation study to the split-conversation setup by eliminating the “Wants to Convert” class. As shown in Table 16 and Fig. 14, the model attained an overall accuracy of 99%, a balanced accuracy of 98.5%, and a macro-F1 score of 0.985. This confirms that dynamic truncated sequence modeling remains robust across varying conversation lengths when class definitions are distinct, though addressing the full three-class ambiguity in live settings remains a target for future work.

Table 16. Performance of Enhanced Bi-LSTM (training on split conversations, 2-Class Ablation).

Classes	Precision	Recall	F1-Score
Interested in Islam	0.99	0.99	0.99
Accepted Islam	0.98	0.98	0.98
Overall Accuracy	0.990		
Balanced Accuracy	0.985		
Macro-F1 Score	0.985		

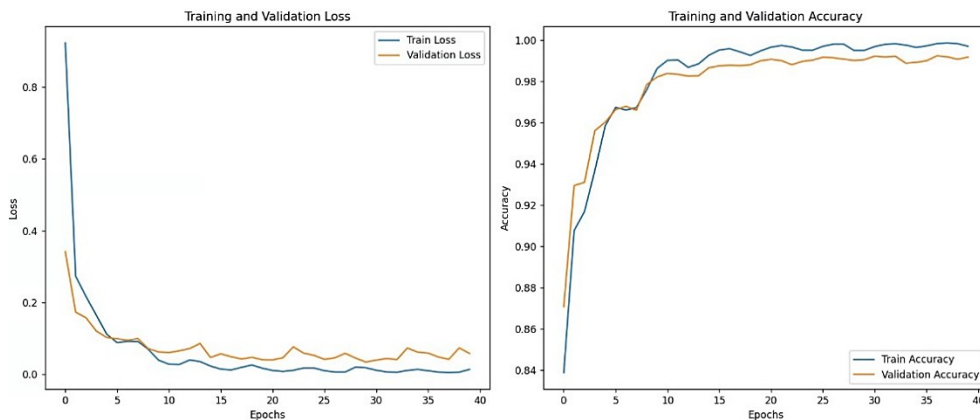


Figure 14. Performance of Enhanced Bi-LSTM (training on split conversations, 2-Class Ablation).

4.3. Message Impact Prediction

In this section, we evaluate the performance of Bi-LSTM and BERT models in predicting the impact of agent (Daei) messages prior to deployment (Positive vs. Negative), as defined in Section 3.5. Before assessing the predictive architectures, we performed a formal reliability analysis to validate the target labels and remove possible circularity in label generation. Two qualified dialogue counselors independently audited a randomized sample of N=500 agent message segments, which included pre-message context, agent message, and visitor turns. Observed agreement between human impact ratings and our automated transition rules was 84.6%, and Cohen's Kappa was 0.71. This large interrater agreement demonstrates the empirical ground-truth validity of the target impact labels.

The class imbalance in the impact dataset was addressed by embedding class weights in the loss function during training to prevent bias toward the dominant negative class.

Table 17 shows that the Bi-LSTM model achieved an accuracy of 78%, which is a good score for the negative class, with an F1-score of 0.85, and a moderate score for the positive class, with an F1-score of 0.59. The weighted average of 0.79 shows balanced overall performance, with weights used to help the model focus on both classes.

The BERT model achieved an accuracy of 80% (as shown in Table 18), which is higher than the Bi-LSTM model. It performed well on the negative class (F1-score of 0.86) and improved on the positive class (F1-score of 0.62). This illustrates BERT's ability to address class imbalance effectively.

Table 17. Performance of Bi-LSTM for predicting message impact.

Classes	Precision	Recall	F1-Score
Negative	0.87	0.83	0.85
Positive	0.56	0.64	0.59
Accuracy	0.78		

Table 18. Performance of BERT for predicting message impact.

Classes	Precision	Recall	F1-Score
Negative	0.87	0.85	0.86
Positive	0.59	0.64	0.62
Accuracy	0.80		

4.4. Prototype Implementation and Real-world Application

We developed a working prototype to validate the practicality and real-time capabilities of our proposed framework. This prototype integrates the best-performing models and techniques from our research, demonstrating their effectiveness in a real-world operational setting. The prototype demonstrates real-time sentiment analysis and conversational outcome prediction. It has been effectively expanded beyond the original domain to a new application where the key classification task is to determine whether a visitor's overall sentiment regarding a conversation is positive or negative.

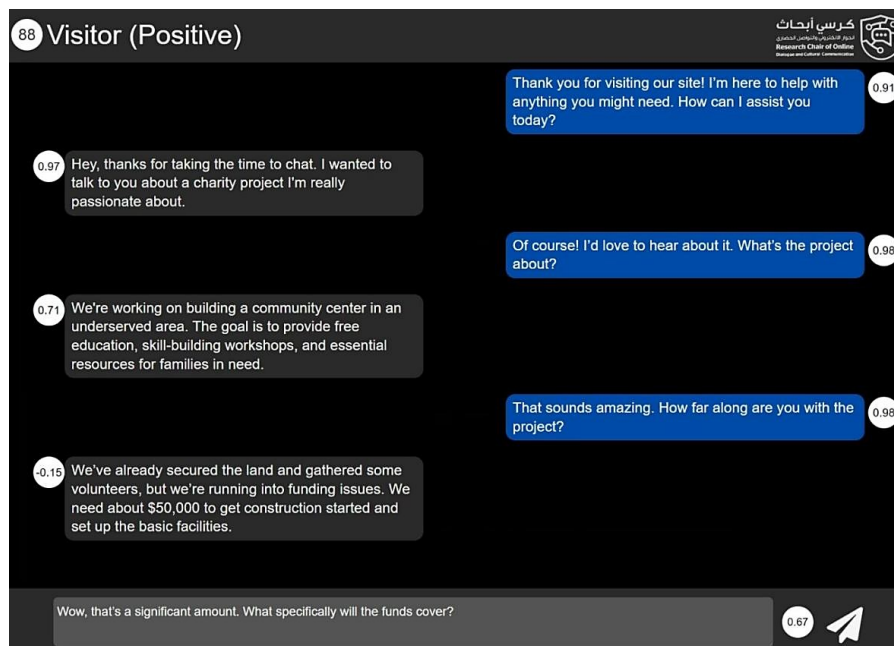


Figure 15. A prototype agent interface demonstrates real-time sentiment scoring, conversation outcome prediction, and pre-send message impact evaluation.

Fig. 15 depicts a screenshot of the agent's UI within this prototype. As shown, a real-time sentiment score is given to each message sent during the conversation, offering detailed emotional insights. Importantly, the expected outcome of the entire conversation (such as "Positive" in the upper-left corner of Fig. 15) is constantly shown, giving the agent an immediate sense of the interaction's trajectory. Additionally, the system predicts and shows the predicted impact of the agent's unsent message right next to the submit button to promote proactive communication. Agents can iteratively improve their responses through this real-time feedback mechanism, ensuring their communication is strategically matched to steer the conversation toward desired outcomes. The successful development and demonstration of these features within the prototype prove the feasibility of real-time deployment and underline the proposed approach's robust adaptability to diverse domains involving customer-agent interactions.

4.5. Real-time Performance and Latency Analysis

We tested the inference latency of RoBERTa sentiment scoring and Enhanced Bi-LSTM outcome prediction on a test set of 1,000 messages to verify the proposed framework's real-time capacity. The experiments were run on a local workstation with an NVIDIA GeForce RTX 4080 GPU. With a 95th percentile latency of 110 milliseconds, the average end-to-end processing time per message was 85 milliseconds. This latency is negligible because turns in human-agent text chat typically occur between seconds and minutes. This confirms that the system can produce sentiment scores and impact predictions instantly without disrupting the conversation, as it falls well within the acceptable threshold for real-time text chat interactions.

5. CONCLUSION

Without non-verbal cues, text-based dialogue systems struggle to infer users' intentions and emotional states; thus, they have limited ability to adapt dynamically and guide the conversation appropriately. This study proposed a machine learning model to address these pressing issues by performing sentiment analysis and predicting conversational outcomes in real time. The main goals of this research were to correctly categorize conversation outcomes and to assess the potential effect of agent messages in real time before deployment.

Our Enhanced Bi-LSTM model performed strongly on the full 3-class challenge for conversational outcome classification, achieving 90% accuracy. The class with the most language ambiguity, and the main source of error, was the "Wants to Convert" class, as shown by the statistical tests and data validation results. We removed this class in an ablation study, and the model's accuracy improved to 99%, showing its effectiveness on well-defined classification problems. Further, our meticulously trained BERT model outperformed the Bi-LSTM (78%), achieving 80% accuracy in real-time evaluation of message impact. This feature ensures high-quality dialog trajectories, allowing agents to proactively improve their answers. We further validated our approach's practical applicability and extensibility to other customer interaction domains by developing and deploying a prototype demonstrated in a real-time agent interface. This framework represents a significant advance for dialogue systems, enabling more intelligent, adaptive, and outcome-driven communication.

REFERENCES

- [1] Waheeb SA, Khan NA, Shang X. (2022) An Efficient Sentiment Analysis Based Deep Learning Classification Model to Evaluate Treatment Quality. *Malaysian Journal of Computer Science*, 35(1):1-20.

-
- [2] Al-Qablan TA, Mohd Noor MH, Al-Betar MA, Khader AT. (2023) A survey on sentiment analysis and its applications. *Neural Comput & Applic*, 35:21567-21601.
 - [3] Devlin J, Chang MW, Lee K, Toutanova K. (2019) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*: 2019; Minneapolis, Minnesota. pp 4171-4186.
 - [4] Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V. (2019) RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692v1*.
 - [5] Siami-Namini S, Tavakoli N, Namin AS. (2019) The Performance of LSTM and Bi-LSTM in Forecasting Time Series Data. In *2019 IEEE International Conference on Big Data (Big Data)*, IEEE, pp. 3285–3292. doi: 10.1109/BigData47090.2019.9005997.
 - [6] Naresh A, Krishna PV. (2021) An efficient approach for sentiment analysis using machine learning algorithm. *Evol Intell*, 14(2):725-731. doi: 10.1007/s12065-020-00429-1.
 - [7] Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., Acharya, U. R. (2021) ABCD: Attention-based bidirectional CNN-RNN-deep model for sentiment analysis. *Future Generation Computer Systems*, 115, 279-294, doi: 10.1016/j.future.2020.08.005.
 - [8] Sharin, N. H., Kartiwi, M. (2025) Identifying Technical and Vocational Education and Training (TVET) Sentiment from Social Media Using a Machine Learning Approach. *IIUM Engineering Journal*, 26(3), 295-303. doi: 10.31436/iiumej.v26i3.3739
 - [9] Buntoro, G., Arifin, R., Syaifuddiin, G., Selamat, A., Krejcar, O., Hamido, F. (2020) The Implementation of the Machine Learning Algorithm for the Sentiment Analysis of Indonesia's 2019 Presidential Election. *IIUM Engineering Journal*, 22(1), 78-92. doi: 10.31436/iiumej.v22i1.1532
 - [10] Poria, S., Majumder, N., Mihalcea, R., Hovy, E. (2019) Emotion recognition in conversation: Research challenges, datasets, and recent advances. *IEEE Access*, 7, 100943-100953, doi: 10.1109/ACCESS.2019.2929050
 - [11] Bouktif S, Fiaz A, Awad M. (2019) Stock Market Movement Prediction using Disparate Text Features with Machine Learning. In *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*. IEEE; pp 1-6. doi: 10.1109/ICDS47004.2019.8942303.
 - [12] Malhotra P, Patil M, Anchan T, Ganatra K, Gaglani J. (2021) Negotiation Closer Bot: Negotiation Outcome Predictor based on Sentiment Analysis. In *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*. IEEE; pp 1-5. doi: 10.1109/ISCON52037.2021.9702347.
 - [13] Li, W., Shao, W., Ji, S., Cambria, E. (2022) BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis. *Neurocomputing*, 467, 73-82, doi: 10.1016/j.neucom.2021.09.057.
 - [14] Johnson, J. M., Khoshgoftaar, T. M. (2019) Survey on deep learning with class imbalance. *Journal of Big Data*, 6, 27, doi: 10.1186/s40537-019-0192-5
 - [15] Rukn Al-hewar non-profit Organization. (2025). Home. [<https://edialogueec.org.sa/>]
 - [16] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437.
 - [17] Van Rijsbergen, C. J. (1979). *Information Retrieval* (2nd ed.). Butterworths.
 - [18] Brodersen, K. H., Ong, C. S., Stephan, K. E., & Buhmann, J. M. (2010). The balanced accuracy and its posterior distribution. In the *2010 20th International Conference on Pattern Recognition (ICPR)*, IEEE, pp. 3121- 3124.
-

Regression-Based Optical Intensity and Image Processing Techniques for Diatom Algae Content Estimation

RAHMAT MAGHROBI ALWAFI¹, AGUS INDRA GUNAWAN^{1*}, SETIAWARDHANA²

¹Department of Electrical Engineering, Politeknik Elektronika Negeri Surabaya, Surabaya, Indonesia

²Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya, Surabaya, Indonesia

*Corresponding author: agus_ig@pens.ac.id

(Received: 13 April 2026; Accepted: 13 August 2026; Published online: 11 September 2026)

ABSTRACT: Microalgae are responsible for oxygen production, pH stabilization, and nitrogen waste absorption in shrimp pond ecosystems, serving as essential biological indicators of water quality. Within the unicellular microalgae group, diatoms such as *Skeletonema* sp. and *Thalassiosira* sp. confer major ecological benefits, excessive growth can disturb pond environments. Despite their significance, monitoring these diatoms, particularly in Indonesian aquaculture systems, remains constrained by time-consuming manual counting. This study proposes a non-invasive integrated framework combining direct raw optical intensity (OI) sensing and digital image processing (IP) techniques to estimate the concentration of *Skeletonema* sp. and *Thalassiosira* sp. across varying laboratory-grown concentration ranges. The proposed regression models demonstrated high predictive performance across all concentration ranges, with test set coefficients of determination exceeding 0.95 ($R^2 > 0.95$). The results indicate that biomass estimations derived from image processing are more accurate than those obtained from direct optical intensity, with data suggesting that species-specific optical properties are modulated by cellular pigment responses. Overall, this non-destructive RGB regression framework offers reliable microalgae biomass estimation for precision aquaculture.

ABSTRAK: Mikroalga bertanggungjawab menghasilkan oksigen, menstabilkan pH, dan menyerap sisa nitrogen dalam ekosistem kolam udang, berfungsi sebagai penunjuk biologi penting kualiti air. Dalam kumpulan mikroalga bersel tunggal, diatom seperti *Skeletonema* sp. dan *Thalassiosira* sp. memberikan manfaat ekologi besar, namun pertumbuhan berlebihan boleh mengganggu persekitaran kolam. Walaupun kepentingannya, pemantauan diatom ini, terutamanya dalam sistem akuakultur Indonesia, masih terhad oleh pengiraan manual yang memakan masa. Kajian ini mencadangkan rangka kerja terintegrasi tidak invasif yang menggabungkan pengesanan intensiti optik mentah secara langsung (OI) dan teknik pemprosesan imej digital (IP) untuk menganggarkan kepekatan *Skeletonema* sp. dan *Thalassiosira* sp. merentasi julat kepekatan yang ditanam di makmal. Model regresi yang dicadangkan menunjukkan prestasi peramalan yang tinggi merentasi semua julat kepekatan, dengan pekali determinan set ujian melebihi 0.95 ($R^2 > 0.95$). Keputusan menunjukkan bahawa anggaran biomassa yang diperoleh daripada pemprosesan imej adalah lebih tepat berbanding yang diperoleh daripada intensiti optik langsung, dengan data mencadangkan bahawa sifat optik spesies tertentu dimodulasi oleh tindak balas pigmen selular. Secara keseluruhannya, rangka kerja regresi RGB bukan merosakkan ini menawarkan anggaran biomassa mikroalga yang boleh dipercayai untuk akuakultur tepat.

KEYWORDS: Microalgae, RGB, Optical Intensity, Image Processing, Regression

1. INTRODUCTION

Shrimp aquaculture is an important part of global aquaculture production. However, maintaining water quality at the level needed for long-term productivity is difficult [1,2]. Water quality parameters, such as dissolved oxygen (DO), pH, temperature, salinity, and nitrogen content, must be maintained within specific ranges to prevent stress and mortality in cultured shrimp [3,4]. Microalgae are important for managing the aquaculture ecosystem because they produce oxygen through photosynthesis, absorb CO₂ to help stabilize pH, and take up nitrogenous compounds from shrimp waste and feed leftovers [5]. The visual characteristics of pond water color indicate microalgal diversity and abundance; green coloration indicates healthy phytoplankton dominance, brown suggests diatom prevalence, and variations in color intensity reflect changes in microalgal biomass concentration, which directly correlates with stability in water quality parameters [6,7,8]. However, uncontrolled microalgal growth can result in harmful blooms that destabilize water quality parameters, particularly when dissolved oxygen levels decrease and pH values fluctuate significantly [9].

That situation requires access to accurate, rapid, and non-invasive methods for measuring microalgal concentration. Currently applied approaches primarily combine optical density (OD) measurements and digital image processing (IP). Traditionally, OD measurements quantify cell density using absorbance at particular wavelengths and standard logarithmic calculations [10]. While these conventional studies establish the core physical principle that light attenuation correlates with biomass concentration, implementing standard logarithmic OD transformations in real-time embedded systems limits direct, efficient implementation on constrained edge devices due to the overhead of floating-point operations. To address this limitation, recent paradigms in multi-sensing monitoring systems suggest that raw transmittance light intensity can serve as a direct, lightweight proxy for biomass, improving resource efficiency on embedded platforms [11]. Concurrently, digital image processing utilizes feature extraction such as average RGB values, color intensity, and pixel distribution to examine the characteristics of visual culture and detect complex, species-specific pigment variations [12]. Therefore, this study adapts the fundamental light-attenuation principle by using a direct, raw RGB optical intensity approach alongside digital image processing to achieve highly accurate, computationally optimized biomass estimation.

Previous studies have demonstrated the effectiveness of the optical method for microalgae monitoring. Cvjetinovic et al. (2023) developed specialized optical systems that use LED light at 505 nm to focus on changes in the concentration of different diatom cultures, achieving coefficient of determination values up to 0.999 with 17.67 mm³ sensor volumes [13]. In a related study, Castaldello et al. (2021) used an image-based method focused on chlorophyll-a, the main pigment involved in photosynthesis associated with microalgal biomes containing diatoms, where the biomass prediction model achieved an average relative error of about 18%, comparable to laboratory measurements [14]. However, research specifically focused on *Skeletonema* sp. and *Thalassiosira* sp. primarily depends on manual counting using hemocytometers, which poses a considerable limitation in aquaculture settings, especially in Indonesia [15]. Although advanced narrow-band optical and multivariate image regression techniques offer high precision, their reliance on specialized hardware or intensive computation limits real-time deployment on low-cost embedded platforms, thereby restricting their application as automated alternatives to manual hemocytometer counting for *Skeletonema* sp. and *Thalassiosira* sp. in Indonesian aquaculture. Therefore, this study aims to develop and validate an integrated, non-invasive methodology that combines direct raw optical intensity readings and image-processing techniques with regression models to accurately and rapidly estimate the biomass concentration of *Skeletonema* sp. and *Thalassiosira* sp.

This study builds upon and expands the methodology introduced in the previous prototype study [16]. While the baseline system utilized optical density to characterize raw sensor outputs, the present work refines this physical framework by formally defining these variables as transmittance-based raw optical intensity features. This terminology adjustment aligns more closely with the system's direct RGB readings, ensuring strict methodological clarity between polychromatic light attenuation and classical logarithmic absorbance metrics.

Beyond this conceptual clarification, this study's principal technical contributions are organized into three main advancements. First, the research focuses exclusively on *Skeletonema* sp. and *Thalassiosira* sp., leveraging their distinct pigment profiles to generate species-specific RGB responses and improve taxonomic specificity over the initial system. Second, expanding the baseline single-modality approach, it evaluates multiple linear regression and second-order polynomial regression models using raw optical data and image-processing features, with concentration-based train-test splits and 5-fold cross-validation. Third, the study provides a comprehensive biological interpretation linking pigment composition to modeling performance. The resulting models reached high predictive accuracy ($R^2 > 0.95$) with minimal errors. The remainder of this paper is organized as follows: Section 2 details the methods and experimental procedures, Section 3 presents the measurement results, and Section 4 offers the study's conclusions.

2. RESEARCH METHODS

This study used a non-destructive measurement method. As shown in Fig. 1(a), the system design describes the measurement workflow, in which a white LED backlight illuminates the specimen. A digital microscope and an RGB sensor recorded data from the specimen simultaneously, producing an image and RGB intensity. The results were then processed using a Raspberry Pi for further analysis. As shown in Fig. 1(b), the prototype consisted of a Class A enclosure equipped with a quartz glass cell unit, a digital microscope, an RGB sensor, and LED lighting integrated into a compact enclosure to minimize external light interference and ensure consistent measurement conditions.

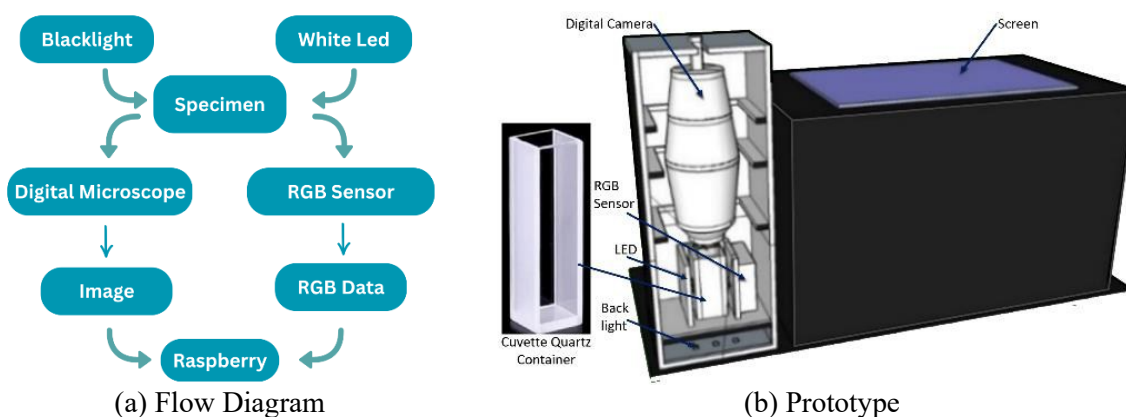


Figure 1. System Design

2.1. Methods and System

To achieve comprehensive and robust biomass estimation, this study employs an integrated dual-modality framework combining direct optical intensity (OI) and image processing (IP). As illustrated in the system architecture in Fig. 1(b), the optical sensing block features a TCS34725 RGB sensor array aligned with a light source. The microalgae sample cuvette was systematically positioned between the light-emitting diode (LED) backlight and

the sensor photodiodes. A white LED was selected to provide a broad, stable polychromatic spectrum, ideal for exciting and analyzing the distinct color and pigment characteristics of the diatom cultures [17]. Short-term exposure during monitoring cycles ensures no impact on microalgae growth [18]. The TCS34725 sensor measured the light intensity incident on the paper in RGB and transmitted the data to the Raspberry Pi via the microcontroller.

Optical measurement relies on polychromatic light attenuation, where microalgae cells absorb and scatter photons, reducing transmitted intensity. In this study, I bypassed classical logarithmic optical density calculations. Because the incident light intensity and the optical path length remained strictly constant, the raw transmitted readings I retained a predictable, monotonic inverse relationship with biomass concentration. Therefore, the direct optical Intensity (OI) framework was operationally formulated as Eq (1):

$$OI(R, G, B) = I(R, G, B) \quad (1)$$

where $I(R, G, B)$ represents the raw digital values extracted directly from the red, green, and blue sensor channels. Bypassing floating-point logarithmic transformations minimizes computational overhead and optimizes real-time processing within the resource-constrained embedded platform. These raw intensity values are referred to as R_{OI} , G_{OI} , and B_{OI} for the red, green, and blue channels, respectively, and represent transmittance-based RGB intensity features rather than classical logarithmic OD values.

IP was also performed by extracting color histograms from digital images of microalgae samples obtained using an integrated camera system [19]. A digital microscope was used as a camera to capture color features rather than for microscopic observation, as shown in Fig. 1(b). The samples were placed between a backlight and a digital microscope. As the measurement space was dark, a white backlight was required to emphasize the intensity captured by the camera from above. To determine the color average parameters for each RGB channel, Eq (2), (3), and (4) were used:

$$R_{ip} = \frac{\sum_{i=1}^n R_i}{n} \quad (2)$$

$$G_{ip} = \frac{\sum_{i=1}^n G_i}{n} \quad (3)$$

$$B_{ip} = \frac{\sum_{i=1}^n B_i}{n} \quad (4)$$

where n is the total number of pixels in the image, and R_i , G_i , and B_i are the intensity values of the i -th pixel in each channel.

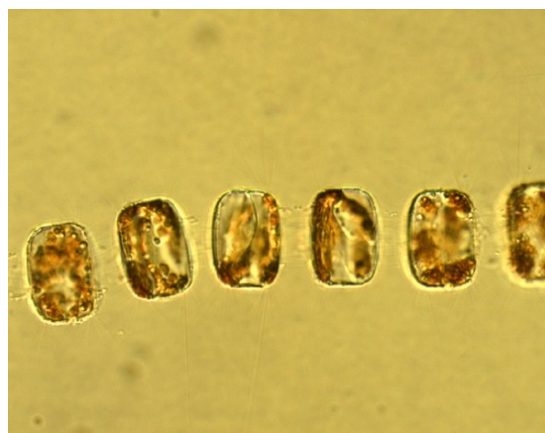
2.2. Microalgae Phycology

Skeletonema sp. is a filamentous diatom that forms long chains of box-shaped cells, as shown in Fig. 2(a) [20]. It does not move by itself, and its color can be yellowish-green or brown because of chlorophyll and carotenoid pigments. The color changes depending on the nutrition, salinity, light intensity, and growth stage. The culture gets darker as biomass increases because cell density and carotenoid levels rise [15]. *Thalassiosira* sp. is a planktonic diatom that comes in cylindrical or box shapes, as shown in Fig. 2(b) [21]. Fish farmers often use it as natural feed. The pigments fucoxanthin and chlorophyll a and c can make it yellow-green or golden brown [22]. This behavior of the three channels suggests that the *Skeletonema* sp. cultures are largely yellowish-green to light brown in color, in agreement with the ever-present, more dominant chlorophyll pigment culture components [23]. *Thalassiosira* sp. and *Skeletonema* sp. have very different amounts of pigments. *Thalassiosira* sp. has more fucoxanthin, which makes it a darker brown color, while *Skeletonema* sp. has more chlorophyll

c, which makes it a lighter or greenish color [23]. This color difference results from a complex mix of pigments, shapes, and cell densities. This is clear from the color and physical traits of each microalga as it grows [24].



(a) *Skeletonema* Sp.[20]



(b) *Thalassiosira* Sp.[21]

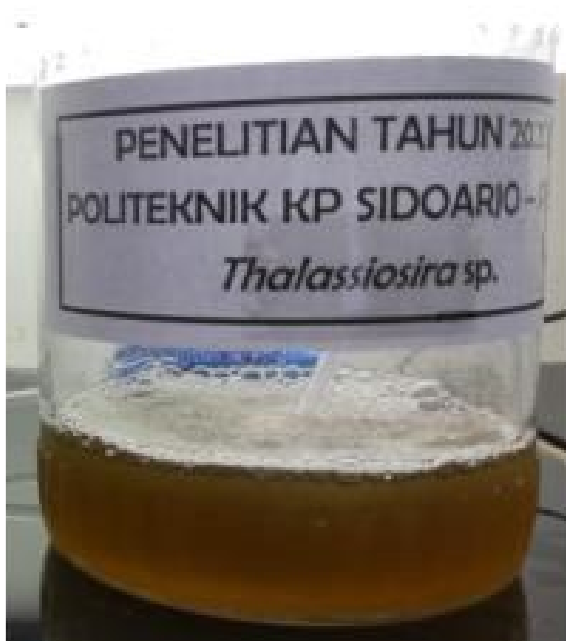
Figure 2. Microalgae

2.3. Sample Preparation

The Department of Fisheries Pathology at Sidoarjo Marine and Fisheries Polytechnic prepared and cultivated samples of *Thalassiosira* Sp. and *Skeletonema* Sp. obtained from the Jepara Brackish Water Aquaculture Research Center (BBPBAP). Fig. 3 shows 43×10^4 cells/mL of *Skeletonema* sp. and 216×10^4 cells/mL of *Thalassiosira* sp. For each sample, a 4 mL cuvette was used to measure concentrations ranging from 0% to 100%. To obtain samples at different concentration levels, each microalgae species was diluted with pure water, producing 10% concentration increments for each species. This cuvette measures $12.5 \times 12.5 \times 45$ mm and has two clear glass sides and two non-clear glass sides. The materials and measuring tools were cleaned in an autoclave to kill any germs that could slow down growth.



(a) *Skeletonema* Sp.



(b) *Thalassiosira* Sp

Figure 3. Alga Sample

The samples were placed in cuvette containers and placed in the measuring chamber, as shown in Fig. 1(b). For the OI method, white LEDs and RGB sensors were placed on the left and right sides, facing the clear glass. When the sample was exposed to white light, the RGB sensor on the other side measured the optical intensity response, which revealed the sample's properties. The IP technique is similar, using a digital sensor to capture the reaction and a backlight to illuminate the sample from below.

2.4. Data Collection and Model Validation

For dataset eleven, we prepared concentrations for each microalga from 0% to 100% in 10% increments. At each concentration, we obtained 100 replicate measurements using both the OI and IP methods. We averaged 100 measurements at each concentration to obtain a single representative value for each of the red, green, and blue channels.

We implemented a concentration-based data split to evaluate model performance on previously unseen concentrations. The training set comprised averaged values from nine concentration levels (0%, 20%, 30%, 40%, 60%, 70%, 80%, 90%, and 100%), while the testing set included the remaining two levels (10% and 50%). This partitioning enabled the assessment of the models' interpolation capability at both low (10%) and mid-range (50%) concentrations. Tables 1 and 2 explicitly report the test concentrations, showing actual concentration (AC) versus predicted concentration (PC). All regression modeling and performance evaluation, including coefficient of determination (R^2) and mean absolute error (MAE), were conducted using Python's scikit-learn library.

To more rigorously assess model robustness and address potential overfitting concerns, we performed 5-fold cross-validation on the training dataset with nine concentration levels. For each fold, we trained the model on four subsets and evaluated it on the remaining subset. To capture both predictive stability and variance, we report the cross-validation metrics as the mean $\pm$ standard deviation (SD) of R^2 and MAE across the five validation folds. This approach provides a more reliable estimate of generalization performance than a single train-test split, particularly given the limited number of available concentration levels.

3. RESULTS AND DISCUSSION

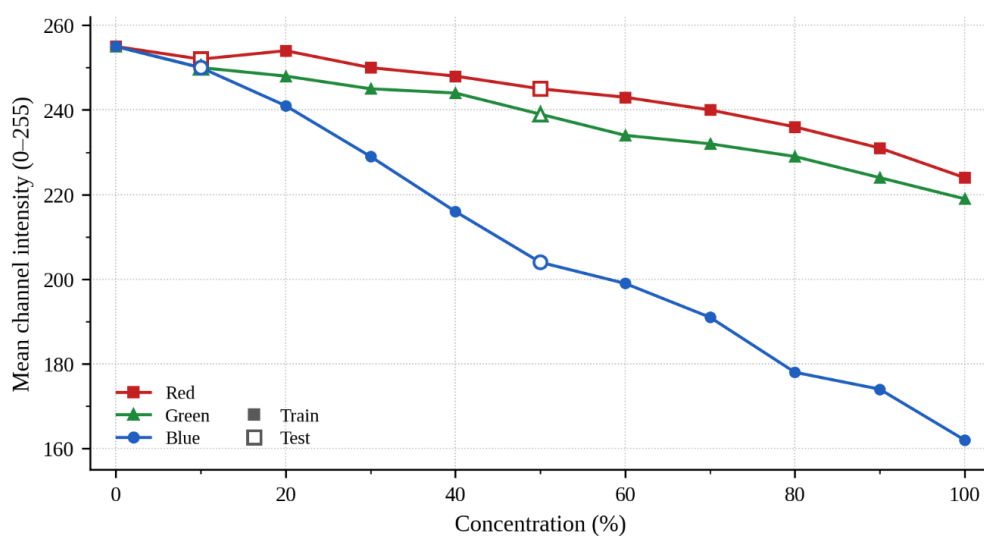
This section presents the results of measurements of *Skeletonema* sp. and *Thalassiosira* sp. using OI and IP techniques. We estimated biomass concentration using both simple and multiple regression models. We used both methods to examine how including multiple factors affects biomass-estimation accuracy, because conventional regression and multiple linear regression differ in the number of predictors considered. Individual plots illustrate the relationship between each predictor and biomass concentration to improve visualization. We evaluated the performance of each regression method for the OI and IP approaches using R^2 and MAE, where R^2 indicates the model's ability to explain variation in the data, and MAE reflects the average prediction error.

Both the image-based average intensity channels (R_{IP} , G_{IP} , and B_{IP}) and channels derived from optical intensity (R_{OI} , G_{OI} , and B_{OI}) served as predictors in a multiple linear regression (MLR) model formulated as $\text{Concentration} = \beta_0 + \beta_1 R + \beta_2 G + \beta_3 B + \varepsilon$. Additionally, a second-order polynomial regression (PR) approach was implemented, in which individual polynomial models were fitted for each feature, and the final concentration was determined by averaging the predictions from all features. A multivariate polynomial regression model was examined, but it did not yield satisfactory results. The model's complexity proved excessive for the dataset's limited size, leading to unstable solutions. Strong correlations among the RGB

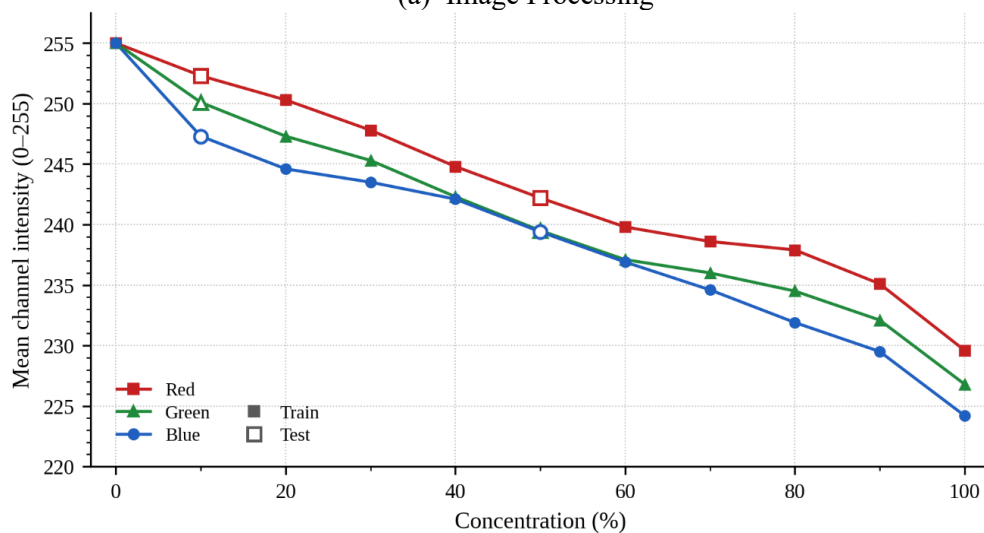
channels led to overfitting, in which the model learned the training data too well but failed to generalize to the test data. As a result, because of complementary patterns in the training data, the multivariate polynomial regression model did not generalize well to unseen data, even compared with individual polynomials and MLR methods.

A custom Python program using the scikit-learn library estimated parameters via least-squares fitting on the training dataset to produce all regression coefficients. The distribution of RGB intensities, changes in OI, and the chosen concentration levels all affect the final coefficients. With the same data and preprocessing circumstances, this formulation allows a direct comparison of linear and nonlinear modeling approaches.

3.1. Skeletonema Sp.



(a) Image Processing



(b) Optical Intensity

Figure 4. *Skeletonema* Sp. measurement results (a) IP and (b) OI

As shown in Fig. 4, the red and green channels were the dominant color channels during both the IP and OI analyses for *Skeletonema* sp. In general, the blue channel was the most diminished as biomass concentration increased. The image data-based method showed greater separation among the three (red, green, and blue) color channels, indicating that mean RGB values from the image data could potentially be used with greater sensitivity to detect changes

in the culture's pigment presence. The curves for each OI channel showed a steady, almost linear trend. This behavior across the three channels suggests that *Skeletonema* sp. cultures are largely yellowish-green to light brown, consistent with the ever-present, more dominant chlorophyll pigment components [23]. The increased coverage of the green color by the fucoxanthin pigment could explain the increase in blue-channel capture. The variation in response between the three channels was much sharper in the IP, implying that this method preserves more color information than the OI method, which provides a more direct optical response to biomass change. Based on the relationships between RGB channel responses and biomass concentration shown in these trends, MLR and second-order PR were acquired for *Skeletonema* sp. for both OI and IP. Eq (6) and (7) provide the descriptive equations for the relationship of the RGB channels to concentration (Y), and Eq (8) to (13) give the individual fit equations for each RGB channel with the average response, which are then used to estimate the overall concentration.

$$Y_{IP} = 473.871827 + 74.817793 \cdot R_{IP} - 398.203066 \cdot G_{IP} - 150.075396 \cdot B_{IP} \quad (5)$$

$$Y_{OI} = 1053.113088 - 1296.950445 \cdot R_{OI} + 640.025413 \cdot G_{OI} - 396.188056 \cdot B_{OI} \quad (6)$$

$$Y_{R_{IP}} = -3504.9130 + 8391.1507 \cdot R_{IP} - 4882.7901 \cdot R_{IP}^2 \quad (7)$$

$$Y_{G_{IP}} = -66.0257 + 1004.2964 \cdot G_{IP} - 940.5092 \cdot G_{IP}^2 \quad (8)$$

$$Y_{B_{IP}} = 291.2321 - 320.0015 \cdot B_{IP} + 31.6565 \cdot B_{IP}^2 \quad (9)$$

$$Y_{R_{OI}} = 143.8409 + 875.3660 \cdot R_{OI} - 1021.4505 \cdot R_{OI}^2 \quad (10)$$

$$Y_{G_{OI}} = 2137.1253 - 3427.0015 \cdot G_{OI} + 1286.5958 \cdot G_{OI}^2 \quad (11)$$

$$Y_{B_{OI}} = 1706.5261 - 2609.8157 \cdot B_{OI} + 899.0882 \cdot B_{OI}^2 \quad (12)$$

Table 1. *Skeletonema* Sp. Evaluation

Methods	Type	R^2		AC	PC	MAE	
		Train	Test				
Multiple Linear Regression	Optical Intensity	0.9892	0.9947	10 50	12.05464 50.21728	2.054638 0.217281	1.136
	Image Processing	0.9992	0.9954	10 50	11.37673 51.32571	1.376728 1.32571	1.3512
2 nd Polynomial Regression	Optical Intensity	0.9854	0.9683	10 50	14.86567 51.29772	4.865673 1.297719	3.0817
	Image Processing	0.9993	0.9969	10 50	8.80026 51.02743	1.199742 1.027428	1.1136

Based on Table 1, IP models outperformed OI models for *Skeletonema* sp. Among all approaches, the second-order PR using IP features delivered the highest predictive accuracy, achieving a test R^2 of 0.9969 and an MAE of 1.1136. The MLR model on IP data also performed strongly (test $R^2 = 0.9954$, $MAE = 1.3512$). OI models were comparatively less accurate, especially the second-order PR variant ($MAE = 3.0817$). These results indicate that RGB features from images better reflect the pigment variations in *Skeletonema* sp. cultures than transmittance-based measurements, and that second-order PR is more suitable for modeling the nonlinear relationships involved. Therefore, the second-order PR model based on IP is recommended as the most reliable method for estimating this species' concentration.

The MLR's strength lies in its ability to use all data in a single regression equation, allowing predictions to be much closer to the actual values, providing consistent performance across the dataset, and reducing the risk of overfitting. Second-order PR is better suited for data

with complex nonlinear patterns because it can capture the data variation more completely. Using R^2 and MAE together provides a more complete analysis of model performance. While R^2 measures the model's overall explanation of variation in the data, MAE assesses individual prediction accuracy.

3.2. Thalassiosira Sp.

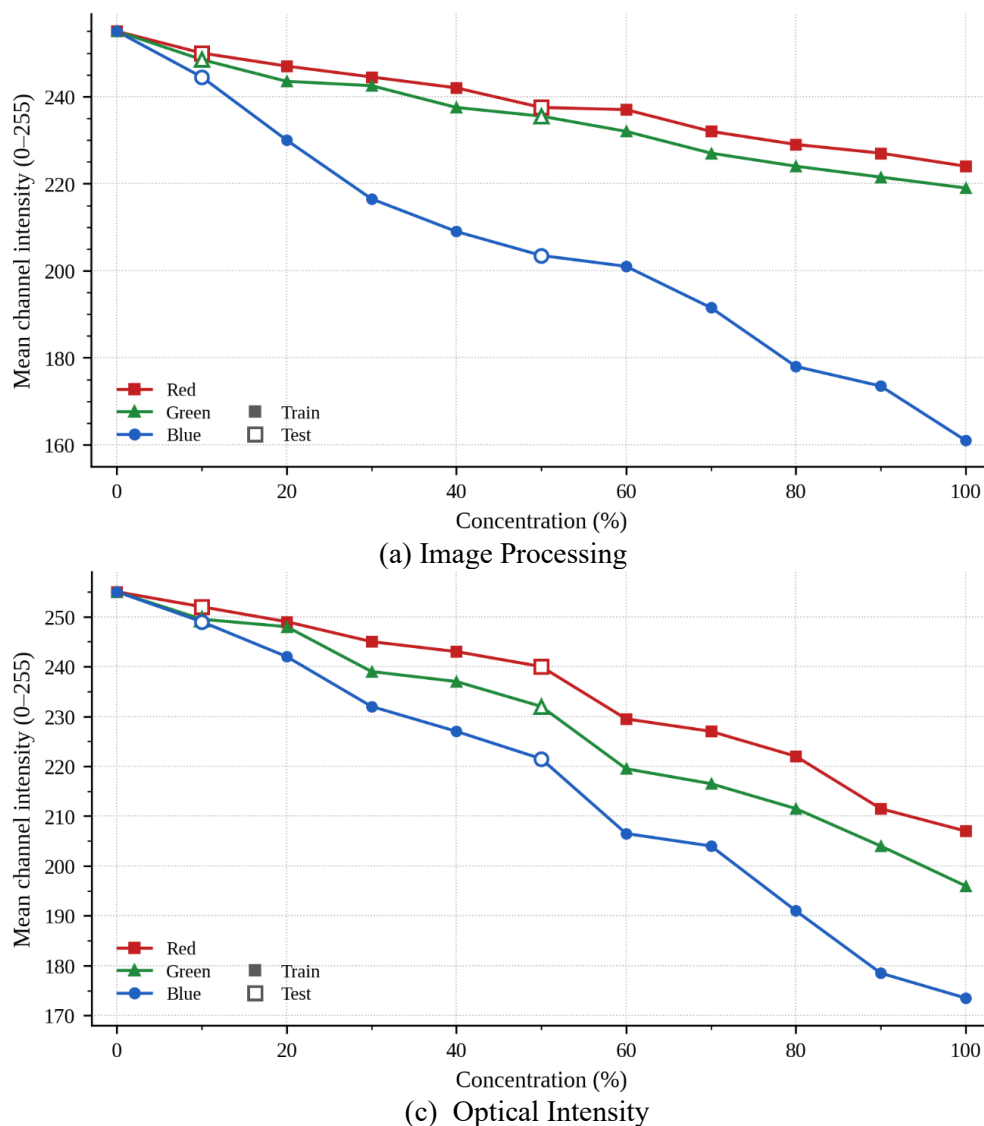


Figure 5. *Thalassiosira* Sp. measurement result

For *Thalassiosira* sp., an all-channel decrease with increasing concentration can be observed in Fig. 5(a) for IP and Fig. 5(b) for OI, with the blue channel changing at a faster rate than the red or green channels. This led to greater channel differentiation with increasing biomass, as indicated by the shapes of the curves on the optical intensity plot, providing a strong optical response to concentration. The extracted mean RGBs for the cultures in the IP method showed a similar decreasing behavior, although there was less distinction among the different channels. Compared with *Skeletonema* sp., *Thalassiosira* sp. generally exhibited lower RGB values, indicating a darker sample appearance. This behavior results from the specific pigment profile of *Thalassiosira* sp., where fucoxanthin is the dominant pigment, giving the cultures a rich brown to golden-brown hue [23]. The different responses of the channels indicate that the pigments may influence the color channels unequally, providing a

basis for quantitative concentration modeling. MLR equations, Eq (14) and (15), and second-order PR Eq (16) - (21) were created using OI and IP data to predict the biomass concentration (Y) from the three channel values, with a mean over the channels equaling the estimated concentration:

$$Y_{IP} = 765.643697 - 1047.030808 \cdot R_{IP} + 299.449323 \cdot G_{IP} - 23.515404 \cdot B_{IP} \quad (14)$$

$$Y_{OI} = 265.198939 + 281.705902 \cdot R_{OI} - 262.164290 \cdot G_{OI} - 283.011439 \cdot B_{OI} \quad (15)$$

$$Y_{RIP} = 2160.9818 - 3752.3202 \cdot R_{IP} + 1589.9766 \cdot R_{IP}^2 \quad (16)$$

$$Y_{GIP} = 2381.4556 - 4317.3729 \cdot G_{IP} + 1935.2007 \cdot G_{IP}^2 \quad (17)$$

$$Y_{BIP} = 373.5593 - 516.4265 \cdot B_{IP} + 140.2109 \cdot B_{IP}^2 \quad (18)$$

$$Y_{ROI} = -534.2934 + 1819.5588 \cdot R_{OI} - 1282.4392 \cdot R_{OI}^2 \quad (19)$$

$$Y_{GOI} = 171.3751 + 161.4931 \cdot G_{OI} - 330.2908 \cdot G_{OI}^2 \quad (20)$$

$$Y_{BOI} = 166.3622 + 37.1178 \cdot B_{OI} - 202.9260 \cdot B_{OI}^2 \quad (21)$$

Table 2. *Thalassiosira* Sp evaluation

Methods	Type	R^2		AC	PC	MAE	
		Train	Test				
Multiple Linear Regression	Optical Intensity	0.9949	0.9713	10	9.81557	0.184433	2.4857
				50	45.21303	4.786969	
	Image Processing	0.989	0.9893	10	7.44456	2.55544	1.9847
				50	48.5861	1.413901	
2 nd Polynomial Regression	Optical Intensity	0.9947	0.9589	10	10.14191	0.14191	2.9362
				50	44.26954	5.730462	
	Image Processing	0.9936	0.9945	10	10.72131	0.721315	1.3483
				50	48.02474	1.975263	

Based on Table 2, IP models outperformed OI models for *Thalassiosira* sp. The second-order PR model trained on IP data achieved the best overall performance, with the highest coefficient of determination on the test set ($R^2 = 0.9945$) and the lowest mean absolute error ($MAE = 1.3483$). The MLR model using IP data also performed very well (test $R^2 = 0.9893$, $MAE = 1.9847$). In contrast, OI-based models were less accurate, particularly the second-order PR using OI features ($MAE = 2.9362$). These findings suggest that RGB values extracted from images more effectively capture the pigment variations of *Thalassiosira* sp. than transmittance measurements. Furthermore, the second-order PR better captured the nonlinear relationship between image features and biomass concentration than MLR.

To evaluate model durability and generalization more systematically beyond a single train–test split, we conducted 5-fold cross-validation on the training dataset consisting of nine concentration levels, as presented in Table 3. For *Skeletonema* sp., the second-order PR combined with IP features delivered the highest predictive accuracy and stability, achieving a mean R^2 of 0.9932 ± 0.0081 and a mean MAE of 1.1738 ± 0.6896 . The MLR model trained on IP data also performed well, whereas the OI models showed lower accuracy and higher variance across validation folds. For *Thalassiosira* sp., the second-order PR applied to the OI data performed best, with a mean R^2 of 0.9440 ± 0.0746 and a mean MAE of 3.3455 ± 1.1081 . The mean R^2 values for the MLR models were similar between the OI and IP approaches, averaging approximately 0.93. However, the OI framework showed greater variability, with an R^2 standard deviation of 0.1038 for the OI model compared to 0.0273 for the IP model. In contrast, the second-order PR on IP data for *Thalassiosira* sp. produced a lower mean R^2

(0.8891 ± 0.1136) and higher variability, indicating less stable cross-validation performance for this specific configuration.

Table 3. 5-fold cross-validation performance (mean $\pm$ standard deviation) for *Skeletonema* sp. and *Thalassiosira* sp.

Microalgae	Methods	Type	R^2 (mean $\pm$ SD)	MAE (mean $\pm$ SD)
<i>Skeletonema</i> Sp.	Multiple Linear Regression	Optical Intensity	0.9575 ± 0.0208	5.3555 ± 2.0538
		Image Processing	0.9841 ± 0.0167	2.1859 ± 0.6299
	2 nd Polynomial Regression	Optical Intensity	0.8529 ± 0.1173	8.0305 ± 4.2258
		Image Processing	0.9932 ± 0.0081	1.1738 ± 0.6896
<i>Thalassiosira</i> Sp.	Multiple Linear Regression	Optical Intensity	0.9283 ± 0.1038	3.3033 ± 1.5831
		Image Processing	0.9309 ± 0.0273	4.6568 ± 2.6311
	2 nd Polynomial Regression	Optical Intensity	0.9440 ± 0.0746	3.3455 ± 1.1081
		Image Processing	0.8891 ± 0.1136	4.1481 ± 3.2165

These cross-validation dynamics indicate that optimal modeling depends heavily on the microalgae species and modality. Direct OI favors MLR because of the nearly linear inverse relationship between transmittance and cell concentration under stable illumination. Conversely, IP captures complex color shifts from pigments by aggregating spatial RGB values, making second-order PR more suitable. However, these nonlinearities can induce isolated prediction outliers across validation folds. Because R^2 relies on squared errors, it heavily penalizes these anomalies, leading to a lower mean R^2 with increased variance (e.g., *Thalassiosira*'s second-order PR model on IP features at 0.8891 ± 0.1136 , whereas MAE remains more stable due to its linearity. This localized instability aligns with a profile dominated by fucoxanthin, which induces darker, nonlinear absorption. Conversely, the low SD in *Skeletonema* sp. models confirms excellent structural stability, as this chlorophyll-rich species produces a brighter culture with a more straightforward concentration gradient.

Cross-referencing Table 3 with Tables 1 and 2 confirms a highly consistent performance hierarchy. The top models, namely second-order PR on OI data for *Thalassiosira* sp. (mean $R^2 = 0.9440$) and second-order PR on IP features for *Skeletonema* sp. (mean $R^2 = 0.9932$), closely align with the baseline tests. Minor numerical variations between the single-split and k-fold validation are statistically expected, representing the transition from a single random partition to an iterated mean. The narrow SD (± 0.0746 and ± 0.0081 , respectively) proves that these slight fluctuations reflect realistic generalization boundaries rather than model instability. This close convergence demonstrates that integrating raw optical intensities and digital image features provides a highly stable, reproducible biomass estimation framework free from partition-dependent bias.

4. CONCLUSION

This study demonstrates that integrating direct raw OI sensing and digital IP provides a reliable, non-destructive framework for estimating the biomass of *Skeletonema* sp. and *Thalassiosira* sp. Bypassing classical logarithmic transformations successfully eliminates unnecessary computational steps, ensuring that the system is highly optimized for real-time applications. The evaluation confirms high accuracy, where the optimal model selection depends on the microalgae species and the analytical method used. The second-order PR achieved the best overall stability and performance, yielding a mean R^2 of 0.9440 on direct OI data for *Thalassiosira* sp. and 0.9932 on IP features for *Skeletonema* sp. The direct OI method

aligns effectively with MLR because of the near-linear relationship established under stable illumination. Conversely, the IP method aggregates color information across numerous visual coordinates to capture complex pigment variations, making the second-order PR more suitable. The narrow standard deviations across the validation subsets prove that these models are stable and free from partition-dependent bias. Ultimately, this dual-modality framework provides a robust, reproducible technological solution for automated microalgae biomass monitoring in precision aquaculture systems.

ACKNOWLEDGMENT

This research was supported by the Directorate of Academic Affairs for Vocational Higher Education (DAPTV) through the 2025 Master's Thesis Research Grant scheme. The authors also thank the Electrical Engineering for Aquaculture Technology Research Group, Politeknik Elektronika Negeri Surabaya, for providing the necessary laboratory facilities and equipment.

REFERENCES

- [1] Yusoff FM, Umi WAD, Ramli NM, Harun R (2024) Water quality management in aquaculture. *Cambridge Prism. Water*, 2:e8. <https://doi.org/10.1017/wat.2024.6>
- [2] Terkula B, Azman N (2021) Recent advances in shrimp aquaculture wastewater management. *Heliyon*, 7:e08283. <https://doi.org/10.1016/j.heliyon.2021.e08283>
- [3] Moayed A, Yargholi B, Pazira E, Babazadeh H (2021) Investigation of bio-desalination potential algae and their effect on water quality. *Desalin. Water Treat.*, 212:78-86. <https://doi.org/10.5004/dwt.2021.26638>
- [4] Yin S, Jin W, Xi T, Zhou X, He Z, Meng X, Naushad M, Jiang G, Li X (2024) Factors affect the oxygen production of *Chlorella pyrenoidosa* in a bacterial-algal symbiotic system: Light intensity, temperature, pH and static magnetic field. *Process Saf. Environ. Prot.*, 184:492-501. <https://doi.org/10.1016/j.psep.2024.02.004>
- [5] Ende S, Henjes J, Spiller M, Elshobary M, Hanelt D (2024) Recent advances in recirculating aquaculture systems and role of microalgae to close system loop. *Bioresour. Technol.*, 407:131107. <https://doi.org/10.1016/j.biortech.2024.131107>
- [6] Winata HN, Nasution MA, Ahamed T, Noguchi R (2021) Prediction of concentration for microalgae using image analysis. *Multimed. Tools Appl.*, 80(6):8541-8561. <https://doi.org/10.1007/s11042-020-10052-y>
- [7] Yang W, Zheng Z, Lu K, Zheng C, Du Y, Wang J, Zhu J (2020) Manipulating the phytoplankton community has the potential to create a stable bacterioplankton community in a shrimp rearing environment. *Aquaculture*, 520:734789. <https://doi.org/10.1016/j.aquaculture.2019.734789>
- [8] M (1997) Phytoplankton dynamics in shrimp ponds. *Aquac. Res.*, 28(5):351-360. <https://doi.org/10.1046/j.1365-2109.1997.00865.x>
- [9] Xiao X, Peng Y, Zhang W, Yang X, Zhang Z, Ren B, Zhu G, Zhou S (2024) Current status and prospects of algal bloom early warning technologies: A review. *J. Environ. Manage.*, 349:119510. <https://doi.org/10.1016/j.jenvman.2023.119510>
- [10] Hotos GN, Avramidou D, Bekiari V (2020) Calibration curves of culture density assessed by spectrophotometer for three microalgae (*Nephroselmis* sp., *Amphidinium carterae* and *Phormidium* sp.). *Eur. J. Biol. Biotechnol.*, 1(6):e132. <https://doi.org/10.24018/ejbio.2020.1.6.132>
- [11] Dębowski M, Kazimierowicz J, Zieliński M (2025) Multi-sensing monitoring of the microalgae biomass cultivation systems for biofuels and added value products synthesis—challenges and opportunities. *Appl. Sci.*, 15(13):7324. <https://doi.org/10.3390/app15137324>

-
- [12] Sunoj S, Hammed A, Igathinathane C, Eshkabilov S, Simsek H (2021) Identification, quantification, and growth profiling of eight different microalgae species using image analysis. *Algal Res.*, 60:102487. <https://doi.org/10.1016/j.algal.2021.102487>
- [13] Cvjetinovic J, Perkovic SA, Kurochkin MA, Sergeev IS, German SV, Bedoshvili YD, Davidovich NA, Korsunsky AM, Gorin DA (2023) Concentration dependence of optical transmission and extinction of different diatom cultures. *J. Biomed. Photonics Eng.*, 9(1):010303. <https://doi.org/10.18287/JBPE23.09.010303>
- [14] Castaldello C, Gubert A, Sforza E, Facco P, Bezzo F (2021) Microalgae monitoring in microscale photobioreactors via multivariate image analysis. *ChemEngineering*, 5(3):49. <https://doi.org/10.3390/chemengineering5030049>
- [15] Badraeni, Trijuno DD, Eriswandi I (2021) Use of bioassay CG (Colour Graduation) to determine density of *Skeletonema* sp. at hatchery. *IOP Conf. Ser. Earth Environ. Sci.*, 763(1):012039. <https://doi.org/10.1088/1755-1315/763/1/012039>
- [16] Kamaluddin MW, Gunawan AI, Setiawardhana, Dewantara BSB, Insivitawati E, Asmarany A, Pratama AE (2024) Algae content estimation utilizing optical density and image processing method. *Int. J. Electr. Comput. Eng.*, 14(6):6248-6257. <https://doi.org/10.11591/ijece.v14i6.pp6248-6257>
- [17] Kondzior P, Tyniecki D, Butarewicz A (2019) Influence of color temperature of white LED diodes and illumination intensity on the content of photosynthetic pigments in *Chlorella vulgaris* algae cells. MDPI AG, pp. 46.
- [18] Al-amshawee S, Bin Mohd Yunus MY (2020) Influence of light emitting diode (LED) on microalgae. *J. Chem. Eng. Ind. Biotechnol.*, 5(2):9-16. <https://doi.org/10.15282/jceib.v5i2.3771>
- [19] Wood NJ, Baker A, Quinnell RJ, Camargo-Valero MA (2020) A simple and non-destructive method for chlorophyll quantification of *Chlamydomonas* cultures using digital image analysis. *Front. Bioeng. Biotechnol.*, 8:00746. <https://doi.org/10.3389/fbioe.2020.00746>
- [20] Phycokey - *Skeletonema* images. Available: https://cfb.unh.edu/phycokey/Choices/Bacillariophyceae/Centric/Centric_Filament/s/SKELETONEMA/Skeletonema_Image_page.html
- [21] Phycokey - *Thalassiosira*. Available: https://cfb.unh.edu/phycokey/Choices/Bacillariophyceae/Centric/Centric_Filament/s/THALASSIOSIRA/Thalassiosira_key.html
- [22] Stenger-Kovács C, Béres VB, Buczkó K, Al-Imari JT, Lázár D, Padisák J, Lengyel E (2023) Review of phenotypic response of diatoms to salinization with biotechnological relevance. *Hydrobiologia*, 850(20):4665-4688. <https://doi.org/10.1007/s10750-023-05194-7>
- [23] Mishra B, Tiwari A (2021) *Spirulina platensis* for the production of C-Phycocyanin, C-Phycocerythrin and *Thalassiosira*, *Skeletonema*, *Chaetoceros* for fucoxanthin. *Syst. Microbiol. Biomanufacturing*. <https://doi.org/10.1007/s43393-020-00020-w>
- [24] Sharma A, Singh P, Srivastava P (2023) Photosynthetic pigments in diatoms. In: *Insights into the World of Diatoms: From Essentials to Applications*. Srivastava P, Khan AS, Verma J, Dhyani S (Eds.). Springer Nature Singapore, Singapore; pp. 1-20.
-

Efficient Edge-Based Object Detection via Multi-Signal Teacher–Student Knowledge Distillation

RICHA SHARMA^{1,2}, ASHUTOSH GUPTA¹,
PRASUN CHAKRABARTI¹, ARVIND SHARMA¹

¹*Faculty of Computing and Informatics, Sir Padampat Singhanian University, Udaipur, India*

²*Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering,
Mumbai, India*

Corresponding author: richa.gairola@gmail.com

(Received: 16 April 2026; Accepted: 11 August 2026; Published online: 11 September 2026)

ABSTRACT: A key obstacle to bringing computer vision everywhere, from robots and drones to smart cities and wearable tech, is the difficulty of deploying real-time object detectors on resource-limited edge hardware such as the NVIDIA Jetson Nano, Raspberry Pi 5, and ARM-based mobile processors. State-of-the-art detectors such as YOLOv8-L achieve outstanding accuracy on server GPUs but often deliver only single-digit or low FPS on embedded hardware. This paper presents EdgeKD-Net, a multi-signal Knowledge Distillation (KD) framework that transfers the representational knowledge of a high-capacity teacher to a compact edge-deployable student. The student employs a MobileNetV3-Large backbone augmented with a quantization- and distillation-aware Dual-Path Feature Pyramid (DPFP). Three complementary distillation signals are introduced: (i) temperature-scaled response-level KL divergence ($\tau=4$); (ii) Channel-Affinity Feature Alignment (CAFA), transferring structural channel co-activation patterns via cosine-affinity matrices; and (iii) Detection Fidelity Loss (DFL), applied exclusively on the formally defined consistent anchor set, eliminating noisy gradients from anchor-assignment mismatch. Across three independent runs, EdgeKD-Net achieves a mAP of 38.6 ± 0.14 mAP@[0.50:0.95] on MS-COCO 2017, only 1.4 mAP below its YOLOv8-L teacher while operating at 61 FPS (FP16) and 89 FPS (INT8) on NVIDIA Jetson Nano with 8.77 mAP/W energy efficiency. The framework is also evaluated on the VisDrone 2023 dataset, which contains dense aerial small objects. It achieves 14.1 ± 0.22 mAP@[0.50:0.95], compared with 12.4 for the best distillation baseline. This improvement is larger than the 0.9-point gain on MS-COCO, showing that the proposed method is effective on both datasets. The evaluation is limited to these two datasets, and we do not claim broader generalization.

ABSTRAK: Halangan utama dalam memperluas penggunaan penglihatan komputer, daripada robot dan dron kepada bandar pintar serta teknologi boleh pakai, ialah kesukaran melaksanakan pengesanan objek masa nyata pada perkakasan pengkomputeran pinggir yang mempunyai sumber terhad, seperti NVIDIA Jetson Nano, Raspberry Pi 5 dan pemproses mudah alih berasaskan ARM. Pengesanan tercanggih seperti YOLOv8-L mencapai ketepatan yang sangat baik pada unit pemprosesan grafik (GPU) pelayan, tetapi lazimnya hanya menghasilkan kadar bingkai sesaat (FPS) satu digit atau rendah apabila dilaksanakan pada perkakasan terbenam. Makalah ini memperkenalkan EdgeKD-Net, iaitu kerangka Penyulingan Pengetahuan (Knowledge Distillation, KD) berbilang isyarat yang memindahkan pengetahuan representasi daripada model pengajar berkapasiti tinggi kepada model pelajar kompak yang sesuai dilaksanakan pada peranti pinggir. Model pelajar menggunakan tulang belakang MobileNetV3-Large yang dipertingkatkan dengan Piramid Ciri Dwialiran (Dual-Path Feature Pyramid, DPFP) yang peka terhadap pengkuantuman dan penyulingan. Tiga isyarat penyulingan yang saling melengkapi diperkenalkan: (i) pencapaian Kullback–Leibler (KL) pada aras respons dengan penskalaan suhu ($\tau = 4$); (ii) Penjajaran Ciri

Afiniti Saluran (Channel-Affinity Feature Alignment, CAFA), yang memindahkan pola struktur pengaktifan bersama saluran melalui matriks afiniti kosinus; dan (iii) Kehilangan Kesetiaan Pengesanan (Detection Fidelity Loss, DFL), yang digunakan secara eksklusif pada set sauh tekal yang ditakrifkan secara formal bagi menghapuskan kecerunan hingar akibat ketidakpadanan penetapan sauh. Berdasarkan tiga larian bebas, EdgeKD-Net mencapai mAP sebanyak 38.6 ± 0.14 bagi mAP@[0.50:0.95] pada set data MS-COCO 2017, iaitu hanya 1.4 mata mAP lebih rendah daripada model pengajarnya, YOLOv8-L. Pada NVIDIA Jetson Nano, model ini beroperasi pada 61 FPS menggunakan FP16 dan 89 FPS menggunakan INT8, dengan kecekapan tenaga sebanyak 8.77 mAP/W. Kerangka ini turut dinilai menggunakan set data VisDrone 2023 yang mengandungi objek udara bersaiz kecil dan padat. EdgeKD-Net mencapai 14.1 ± 0.22 mAP@[0.50:0.95], berbanding 12.4 yang dicatatkan oleh kaedah asas penyulingan terbaik. Peningkatan ini lebih besar daripada peningkatan 0.9 mata pada MS-COCO, sekali gus menunjukkan bahawa kaedah yang dicadangkan berkesan pada kedua-dua set data. Walau bagaimanapun, penilaian kajian ini terhad kepada kedua-dua set data tersebut dan tiada dakwaan dibuat mengenai keupayaan generalisasi yang lebih luas.

KEYWORDS: *Knowledge distillation, object detection, edge computing, anchor-assignment consistency, FPN, aerial object detection*

1. INTRODUCTION

Deploying vision systems on robots, drones, and embedded cameras is, in practice, more challenging than it appears. These devices need to detect objects in real time, typically at 20–30 frames per second or faster, while drawing only a few watts of power. Models like YOLOv8 [1], YOLOv9 [30], RT-DETR [2], and DINO [3] hit impressive accuracy numbers on server benchmarks, but their inference speed collapses on embedded hardware. A YOLOv8-L running on an NVIDIA Jetson Nano, for instance, manages around 7 FPS. This falls well short of the requirements of a delivery robot that needs a new perception update every 50 ms.

Knowledge distillation (KD) [4] is one of the more practical approaches to bridging this gap. The central idea is to train a lightweight student model to replicate the behavior of a heavier teacher, transferring learned knowledge without carrying over the computational cost. For image classification, this performs well and has been studied extensively [5–7]. Detection is a harder case, however, and this work identifies a specific issue that, to our knowledge, has not been formally addressed before.

The problem is anchor-assignment mismatch. In anchor-based detectors, each spatial position in the feature map is associated with one or more anchor boxes, and each anchor is assigned responsibility for detecting a particular object. When the teacher and student use different backbones and therefore develop different spatial feature representations, they often disagree on which anchor should handle which object. Existing knowledge distillation methods, including FGD [11], MGD [12], DeFeat [13], and PKD [9], apply distillation uniformly across all anchors, regardless of whether the teacher and student agree at that position. When they disagree, the student receives conflicting supervisory signals. Our experiments show this happens at roughly 22.9% of matched anchor positions across the COCO training set, and that naively distilling through those positions costs approximately 0.5 mAP.

We built EdgeKD-Net to address this, along with two other issues: the capacity gap between teacher and student feature representations (handled by CAFA) and the student neck architecture (handled by DFPF). We settled on MobileNetV3-Large [14] as the backbone after running latency measurements on an NVIDIA Jetson Nano at 416×416 resolution. It delivered 44 FPS at 4.0 W, which beat both EfficientViT-M0 (37 FPS / 4.8 W) and FastViT-T8 (41 FPS

/ 4.3 W) under the same conditions. The complete framework, EdgeKD-Net, makes the following contributions:

- A quantization- and distillation-aware dual-path pyramid (DPFP) that integrates top-down semantic context with bottom-up spatial detail at a cost of 1.8 M additional parameters. We do not claim the dual-path topology itself as novel; it follows PANet [16] and BiFPN [17], but rather the specific operator and width choices that make it suited to INT8 edge deployment and to channel-aligned distillation from the teacher neck.
- CAFA transfers structural relationships between feature channels rather than duplicating raw values. It is applicable to teacher–student networks of different capacities.
- Detection Fidelity Loss (DFL) is only trained with the examples on which there is agreement between teacher and student (boxes overlap $\geq 50\%$ and same predicted class). Approximately 77.1% of matched anchors pass this criterion. We evaluate the classification and localization components separately.
- Statistical validation was performed using three independent runs, with results reported as mean $\pm$ standard deviation. The proposed method showed statistically significant improvement over the Probabilistic Knowledge Distillation (PKD) baseline ($p < 0.01$) on 5,000 evaluation images.

2. LITERATURE REVIEW

The field of efficient object detection for edge devices has developed along several overlapping directions over the past few years. We review the most relevant threads here and situate EdgeKD-Net within them.

2.1. Lightweight Object Detectors and Edge-Efficient Architectures.

MobileNetV3 [14] achieved good accuracy, latency, and power consumption by using Neural Architecture Search (NAS) with depth-wise separable convolution. EfficientDet [17] improvises the detection neck by using BiFPN and a compound scaling rule that works in conjunction with backbone depth, width, and input resolution. GhostNet [22] improves computational efficiency by using linear transformations to create feature maps, a principle adopted in our DPFP module. More recently, EfficientViT [28] showed that linear attention with multi-scale feature aggregation can result in high accuracy at much lower FLOPs. Recent real-time detectors like RTMDet [31] further emphasize the trade-off between accuracy and efficiency, particularly for deployment on resource-constrained edge devices. MobileViT-v2 [29] is a hybrid model that combines depthwise convolutions with separable self-attention, making it suitable for edge devices. We tested MobileNetV3-Large, EfficientViT-M0, and FastViT-T8 directly on the NVIDIA Jetson Nano to select the student backbone; we discuss the results in Section III.B. This persistent 4–10 mAP gap between compact and large models on COCO motivates distillation-based approaches.

Taken together, these works treat architectural efficiency and knowledge transfer as separate problems: each optimizes a backbone or neck in isolation, with no mechanism for borrowing representational strength from a larger network. EdgeKD-Net treats the two jointly, pairing an empirically selected backbone with a distillation objective (CAFA) designed around the specific capacity gap that choice creates.

2.2. Knowledge Distillation

The original KD paper by Hinton et al. [4] showed that training a student on the teacher's softened output probabilities (controlled by a temperature τ), rather than hard one-hot labels, provides a richer learning signal; the soft outputs encode information about which classes look similar to each other. Romero et al. [5] extended this to intermediate layers with FitNets, showing the student can benefit more from mimicking hidden representations, not just final outputs. Zagoruyko and Komodakis [6] found that matching spatial attention patterns transfers more useful information than matching raw activations, while Tian et al. [7] used contrastive learning, though the memory cost made it impractical for edge-oriented training. Park et al. [8] shifted the focus entirely away from individual feature vectors toward relational structure, the pairwise distances and angles between samples. This relational distillation idea is what directly inspired our CAFA module, though we apply it at the channel level instead of the sample level. Passalis and Tefas [9] used Pearson correlation kernels to match the statistical structure of feature distributions, producing smoother gradients than direct activation matching when the capacity gap is large.

Across this line of work, the trend is toward progressively richer distillation targets, soft logits, hidden activations, attention maps, relational structure, distributional statistics, but every one of these targets is defined over a fixed, shared input representation: a classifier's output is anchored to one image, not to a spatial grid the teacher and student may partition differently. None of these formulations has to contend with a structural disagreement about which output corresponds to which target, which is exactly the complication that anchor-based detection introduces and that we take up next.

2.3. Knowledge Distillation for Object Detection

Using KD for detection raises complications that classification does not have. Many anchor locations are actually background, so applying distillation across all of them mostly captures irrelevant information. Object detection also uses multiple scales, and each anchor corresponds to a specific location, so differences between teacher and student can cause position errors instead of just confidence differences. Chen et al. [10] were pioneers in tackling this issue. They introduced L2 loss between teacher and student feature maps. It performed well when the two networks had similar capacity but degraded when they did not. Yang et al. [11] presented FGD, where the distillation signal was guided by the teacher's attention maps, enabling the student to focus on relevant areas. Yang et al. [12] proposed MGD, in which student features were randomly masked and trained to recover the complete teacher features, encouraging better feature learning. Guo et al. [13] suggested that jointly training classification and regression heads created conflicting gradients and separated them in DeFeat. Zheng et al. [24] later introduced Localization Distillation (LD), which helped in decoupling the spatial and categorical knowledge streams by matching only the teacher's localization distribution. In our DFL, LD still applies its matching globally across all anchors and does not address the fact that the teacher and student may disagree on which anchor is responsible for which object. Zhu et al. [25] proposed ScaleKD as a cross-architecture distillation bridge; feature matching degrades badly when the two networks differ by more than about $2\times$ in depth, which motivated our CAFA.

Jang et al. [26] proposed GLAMD, using gradient-weighted class activation maps to identify which pyramid levels the teacher finds most informative for each object, then focusing distillation there. GLAMD addresses which levels receive distillation but not whether anchor assignments are consistent within those levels. Our analysis shows that anchor-assignment agreement breaks down at 22.9% of matched anchor positions across the COCO training set.

To our knowledge, this anchor-assignment mismatch problem has not been formally defined, measured, or corrected in any detection-distillation literature surveyed here.

A closely related family of methods does select where to distill rather than distilling uniformly, and it is worth separating these from DFL precisely. Region- and position-selection approaches restrict the distillation signal to areas judged informative by teacher attention (FGD), pyramid level (GLAMD), or IoU-guided selection of candidate regions. Separately, the semi-supervised and pseudo-labeling literature filters teacher outputs by estimated quality, typically thresholding on the teacher's own confidence or on prediction stability under augmentation, so that only reliable teacher signals propagate. Both families share DFL's premise that not all supervision is equally trustworthy. They differ in what they condition on: region-selection methods ask where the teacher is informative, and pseudo-label filtering asks whether the teacher is likely to be correct, both of which are properties of the teacher alone. DFL instead conditions on a relational property of the pair, whether teacher and student currently assign the same anchor to the same object, which is invisible to any criterion computed from the teacher in isolation. A teacher prediction may be confident, high-quality, and in an informative region, yet still attached to an anchor the student uses for something else. However, we do not evaluate teacher-confidence filtering or IoU-guided region selection as empirical baselines against DFL; establishing how much of DFL's gain is attributable to the relational criterion specifically, instead of selective distillation in general, would require such a comparison, and we identify it as future work.

Table 1. Comparison of representative detection-distillation methods on anchor-assignment consistency handling

Method	Distillation Signal	Operates At	Conditions on Teacher–Student Agreement?	Key Limitation Relative to DFL
Chen et al. [10]	Feature L2 distance	Whole feature map	No	Uniform loss regardless of agreement; degrades as capacity gap widens
FGD [11]	Attention-guided feature loss	Spatial region (focal/global masks)	No	Re-weights by teacher attention, not by anchor-level assignment agreement
MGD [12]	Masked feature generation	Randomly masked positions	No	Masking is stochastic, independent of whether teacher and student agree at that position
DeFeat [13]	Decoupled cls./reg. distillation	Per detection head	No	Separates gradient sources but not agreeing vs. disagreeing anchors
PKD [9]	Pearson-correlation feature matching	Whole feature-map statistics	No	Aggregate correlation averages away the minority of disagreeing anchors
LD [24]	Localization-distribution matching	Per-anchor, but applied globally	Partial (decouples streams, not agreement)	Matches teacher's localization everywhere; no gate on assignment consistency
ScaleKD [25]	Cross-architecture feature bridging	Whole feature map	No	Degrades when teacher/student differ by $>2\times$ in depth; not assignment-aware
GLAMD [26]	Grad-CAM pyramid-level gating	Pyramid level	Partial (level, not anchor)	Selects informative levels but not consistent anchors within a level
DFL (Ours)	KL divergence gated by Ω	Individual anchor position	Yes, explicit gate ($\text{IoU} \geq 0.5 + \text{class match}$)	Directly measures and removes contradictory gradients from disagreeing anchors

This limitation arises because these methods were not designed to model anchor-assignment consistency. FGD reweights by region, so a region with both agreeing and disagreeing anchors still gets one uniform weight. MGD masks features at random, without checking whether the masked anchor is one the two networks agree on. DeFeat splits classification from regression but does not check anchor-level agreement either way. And PKD matches whole feature-map statistics, so a small pocket of disagreement gets averaged out by everything that does agree. Table 1 compares these methods, along with LD, ScaleKD, and GLAMD, on exactly this point.

2.4. Multi-Scale Feature Representations and Pyramid Architectures

Lin et al. [15] introduced Feature Pyramid Networks (FPN), which connect backbone feature maps top-down via lateral connections to enable multi-scale detection. Liu et al. [16] proposed PANet on top of FPN, adding a second bottom-up path so that fine spatial detail from shallow layers also reaches deeper feature levels. Tan et al. [17] introduced BiFPN in EfficientDet, replacing fixed fusion with learned scalar weights and removing single-input nodes. DPFP is our contribution to this line of work.

The architectural differences from PANet and BiFPN are not just cosmetic. PANet fuses by concatenation, then full convolution, which doubles the channel count before projecting back and adds $O(C^2)$ parameters per scale. BiFPN uses softmax-weighted fusion, which requires division; division is a known source of rounding error under INT8 quantization in general. DPFP instead uses unweighted element-wise addition followed by a single depthwise convolution, with no learned scalars and no normalization overhead, which motivated our design choice on quantization-friendliness grounds. Since we have not quantized PANet- or BiFPN-neck variants under identical INT8 settings, this is a design motivation rather than an experimentally verified comparison; we leave such a comparison for future work. On the bottom-up path, PANet uses full 3×3 convolutions with stride 2, while DPFP uses depthwise separable convolutions, cutting the parameter count by a factor of $C/9$ per stride step. This allows DPFP to achieve essentially the same dual-path information flow as PANet with 1.8 M vs. 2.4 M extra parameters (Table 10). The third difference is that DPFP's output width of 256 channels was specifically chosen to match the YOLOv8-L teacher neck, eliminating any projection mismatch in the CAFA distillation layers. This width was selected deliberately for this reason, rather than adopted as a default configuration.

2.5. Edge Deployment

2.5.1. Quantization and Hardware Optimization

INT8 quantization cuts memory bandwidth by $4 \times$ on embedded hardware and is a primary factor in achieving high FPS on devices like the NVIDIA Jetson Nano. TensorRT, ONNX Runtime, and NCNN are the main inference frameworks that support INT8 on these platforms. We use TensorRT minmax calibration with a 512-image COCO validation subset for our INT8 deployment.

For reproducibility, the complete hardware benchmarking protocol used to collect all latency, power, and INT8 accuracy figures reported in Table 5 is as follows. Edge device: NVIDIA Jetson Nano (4 GB). Operating system: Ubuntu 18.04 LTS. JetPack SDK 4.6.2 (CUDA 10.2, cuDNN 8.2, TensorRT 8.2). Power mode: 5 W (nvpmodel mode 1). Batch size: 1 (single-image inference, matching the deployment scenario this work targets). Precision modes evaluated: FP16 and INT8. INT8 calibration: TensorRT post-training calibration using the same 512-image COCO validation subset noted above. Each reported measurement is

preceded by 10 warm-up iterations (excluded from timing) and averaged over 500 measurement runs.

EdgeKD-Net shows an INT8 accuracy drop of only 0.7 mAP, compared to 1.9 for YOLOv8-N and 1.8 for NanoDet-Plus. We measured the kurtosis of weight distributions in the final classification sub-head and found $r = -0.94$ Pearson correlation with the INT8 accuracy drop, consistent with the hypothesis that temperature-scaled distillation produces smoother weight distributions. This evidence is correlational rather than causal; a controlled ablation isolating this factor would be required to establish a causal link.

2.5.2. Anchor-Free Detectors and Future Extensions

CenterNet [18], FCOS [19], and TOOD represent objects via heatmaps or set-prediction instead of anchor grids, so DFL as defined here cannot be applied directly. However, the underlying consistency principle generalizes: for heatmap-based methods, agreement could be measured by the proximity of predicted object centers; for DETR-style methods, bipartite matching cost could serve the same role. We plan to extend DFL to anchor-free architectures.

3. METHODOLOGY

This section presents the complete EdgeKD-Net architecture and training procedure. We describe the teacher and student network designs, the Dual-Path Feature Pyramid neck, the SSD detection head, the three complementary distillation objectives, the combined training loss, and the full set of training hyperparameters

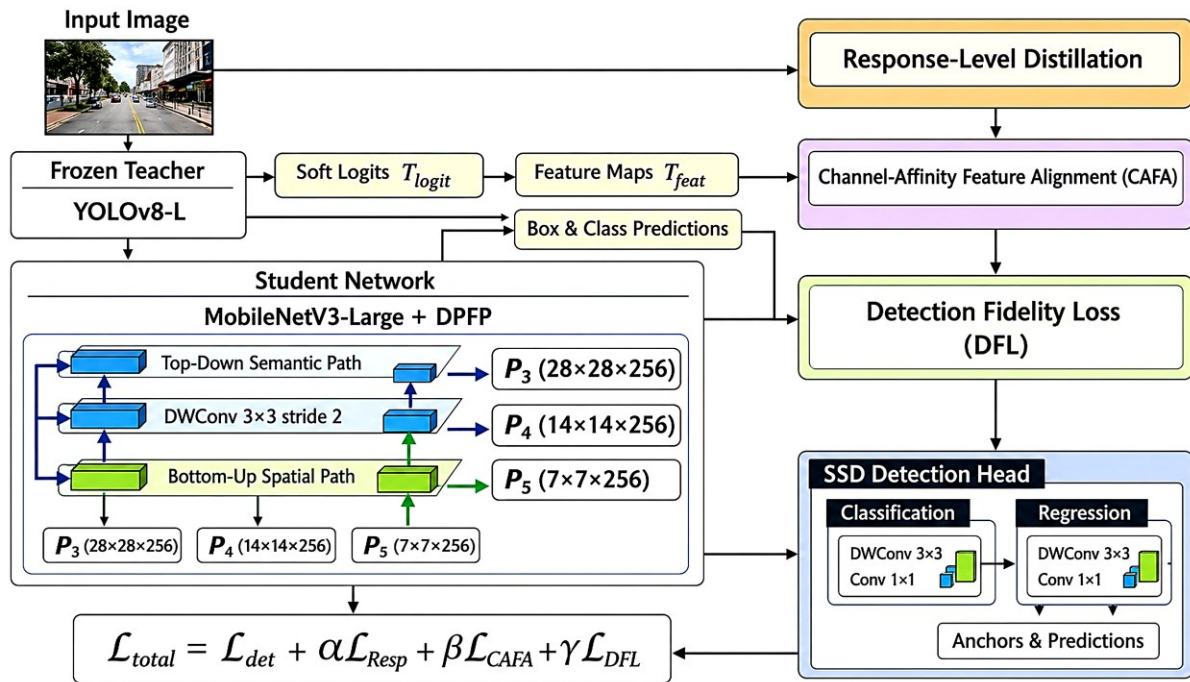


Figure 1. System Architecture of EdgeKD-Net

3.1. Teacher: YOLOv8-L

YOLOv8-L is used as the teacher network. The teacher was trained independently on the same train/validation split used for all student experiments, at the same 224×224 input resolution used throughout this work, and achieved 40.0 mAP@[0.50:0.95] on that validation set. This figure is below commonly reported YOLOv8-L results because those are obtained at substantially larger input resolutions (typically 640×640); the value quoted here reflects our

teacher's accuracy under our evaluation protocol and serves as the correct reference point for the teacher–student gaps reported in this paper. We held this single trained model fixed as the reference teacher across every distillation experiment reported in this paper. Once loaded with these weights, all teacher parameters are set to eval() mode and frozen for the entire student training (`requires_grad=False`). The teacher serves only as a fixed source of soft supervision signals and is discarded at inference time, adding no computational cost to the deployed model.

At each forward pass, the teacher produces a logit tensor over all anchor positions and classes, which serves as the shared reference for all three distillation losses:

$$T(x) \in \mathbb{R}^{A \times C} \quad (1)$$

where A is the total number of anchor positions, and C is the number of classes.

3.2. Student: MobileNetV3-Large + DPFP

The student backbone is MobileNetV3-Large [14], initialized from ImageNet pre-trained weights. We selected this backbone based on direct latency measurements on the deployment hardware (NVIDIA Jetson Nano) at 416×416 input resolution. MobileNetV3-Large delivered 44 FPS at 4.0 W average power, outperforming both EfficientViT-M0 (37 FPS / 4.8 W) and FastViT-T8 (41 FPS / 4.3 W) under identical conditions. Among the three backbones screened, its combination of throughput and energy efficiency made it the most suitable choice for our edge deployment scenario; we did not survey backbones beyond these three. This backbone-screening benchmark was run at 416×416 to match the deployment-time inference resolution used in the edge measurements of Table 5. It is independent of the 224×224 resolution at which all detector training and accuracy evaluation in this paper were performed.

The backbone extracts three hierarchical feature maps at progressively coarser spatial resolutions: C_3 (28×28, 40 channels), C_4 (14×14, 112 channels), and C_5 (7×7, 160 channels). These three maps are passed to the Dual-Path Feature Pyramid module described in the following section.

3.3. Dual-Path Feature Pyramid (DPFP)

Conventional single-path feature pyramids propagate information in one direction only, either top-down to distribute semantic context, or bottom-up to consolidate spatial detail. Each direction alone has limitations: the top-down path loses fine localization precision as it propagates downward, while a bottom-up path alone lacks the category-level context needed for reliable classification of small objects. As shown in Fig. 2, DPFP addresses this by running both paths simultaneously and fusing their outputs at each scale level, at a cost of only 1.8 M additional parameters and 1.3 GFLOPs at 224×224 resolution.

3.3.1. Top-Down Semantic Path

The top-down path begins at C_5 , which carries rich semantic content but coarse spatial resolution. A 1×1 convolution converts it to 256 channels, and 2× bilinear upsampling then transfers this semantic context toward the shallower feature maps:

$$TD_4 = \text{Upsample}(\text{Conv}_1 \times_1(C_5; 160 \rightarrow 256 \text{ ch})) \quad (2)$$

$$TD_3 = \text{Upsample}(\text{Conv}_1 \times_1(TD_4; 256 \rightarrow 256 \text{ ch})) \quad (3)$$

This allows every pyramid level to exploit the high-level category information encoded in the deepest layers, which is particularly important for detecting small objects that would otherwise lack the semantic context needed for confident classification.

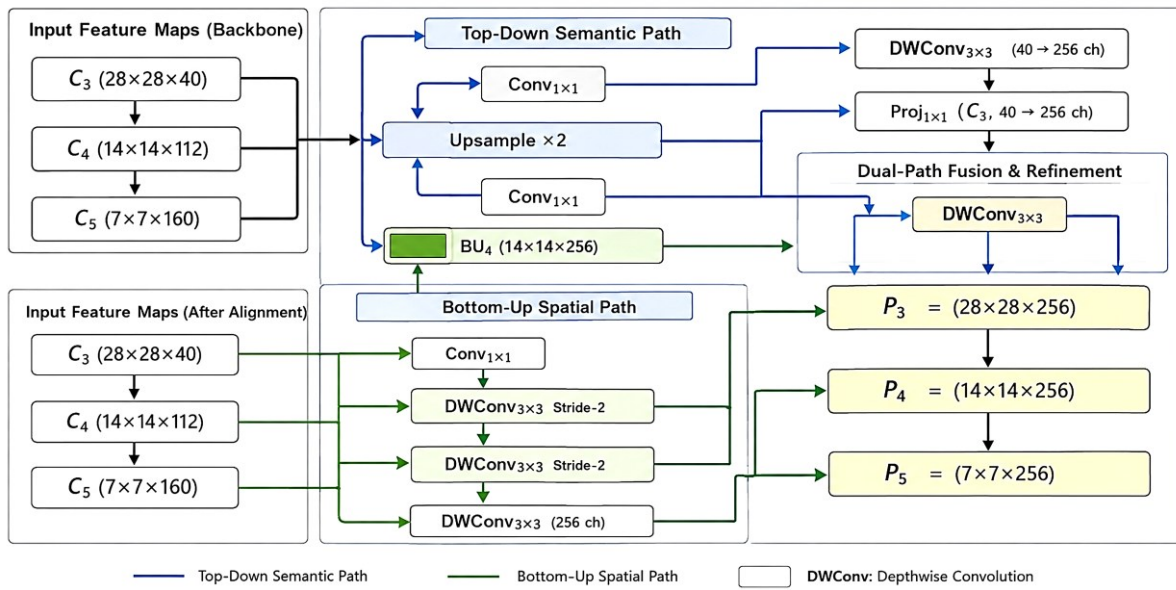


Figure 2. Dual-Path Feature Pyramid (DPFP) with complete channel dimension and stride annotations. Top-down semantic path (blue) and Bottom-up spatial path (green)

3.3.2. Bottom-Up Spatial Path

The bottom-up path runs in the opposite direction, propagating fine-grained spatial detail from the shallowest feature map C_3 upward using stride depthwise separable convolutions:

$$BU_4 = DWConv_3 \times_3 (C_3; 40 \rightarrow 256 \text{ ch}, \text{stride} = 2) \quad (4)$$

$$BU_5 = DWConv_3 \times_3 (BU_4; 256 \rightarrow 256 \text{ ch}, \text{stride} = 2) \quad (5)$$

Depthwise separable convolutions are used here in place of standard convolutions to keep the parameter cost low while still achieving the necessary stride-2 spatial downsampling.

3.3.3. Dual-Path Fusion and Refinement

At each pyramid scale $s \in \{3, 4, 5\}$, the corresponding backbone feature is first projected to 256 channels via a 1×1 convolution, and all three signal sources are then merged by element-wise addition and refined with a depth-wise convolution:

$$P_s = DWConv_{3 \times 3} (TD_s \oplus BU_s \oplus Proj_{1 \times 1}(C_s)), s \in \{3, 4, 5\} \quad (6)$$

This yields three output feature maps: P_3 (28x28x256), P_4 (14x14x256), and P_5 (7x7x256), each combining semantic depth from the top-down path, spatial precision from the bottom-up path, and the backbone's original local features. The 256-channel width was selected for three reasons: it matches the YOLOv8-L teacher's neck output width, eliminating any projection mismatch in the CAFA distillation layers; it is consistent with the standard FPN output width used by FPN [15], PANet [16], and BiFPN [17], keeping comparisons fair; and a controlled width ablation identified it as the best-performing of the three widths tested (128, 256, 512), 128 channels reduced mAP by 0.8 points due to insufficient fusion capacity, while 512 channels added 3.1 M parameters for only a 0.2 mAP gain, an unfavourable trade-off for edge deployment.

3.4. SSD Detection Head Architecture

A decoupled SSD-style prediction head operates on the three DPFP outputs P_3 , P_4 , and P_5 . Three anchor boxes are assigned at each spatial location within each map, with aspect ratios $\{0.5, 1.0, 2.0\}$ and base scales of 32, 64, and 128 pixels, respectively. Separate sub-heads

handle classification and regression, each consisting of a DWConv $_{3 \times 3}$ layer followed by a Conv $_{1 \times 1}$ projection. Decoupling the two tasks prevents the conflicting gradient interactions that arise when they share a common prediction branch [13]. The total number of anchor positions per image is fixed at: $A = 3 \times (28^2 + 14^2 + 7^2) = 3,087$ anchors per image.

3.5. Multi-Signal Distillation

Three distillation losses are applied simultaneously, each targeting a distinct type of knowledge in the teacher.

1. Response-level distillation operates at the output layer and transfers inter-class similarity structure.
2. Channel-Affinity Feature Alignment (CAFA) operates at the feature pyramid level and transfers structural co-activation patterns between channels.
3. Detection Fidelity Loss (DFL) operates at the anchor level and transfers detection quality only at positions where teacher and student are in mutual agreement.

The proposed loss functions provide complementary supervision at the output, feature, and anchor levels, ensuring that each contributes distinct information during training.

3.5.1. Response-Level Distillation

The response-level loss applies temperature-scaled KL divergence between the teacher and student output distributions. The temperature parameter τ softens both distributions before comparison, revealing the relative probability mass the teacher places on non-target classes. This inter-class similarity structure is invisible in hard one-hot labels and provides the student with richer gradient information per training sample [4]. The τ^2 prefactor compensates for the gradient magnitude reduction that temperature scaling introduces:

$$\mathcal{L}_{resp} = \tau^2 \cdot KL\left(\text{softmax}\left(\frac{T_{logit}}{\tau}\right) \parallel \text{softmax}\left(\frac{S_{logit}}{\tau}\right)\right) \quad (7)$$

We chose $\tau = 4$ via Bayesian optimization over [1, 8] on the COCO validation set. At this value, the distillation loss and ground-truth detection loss produce gradients of comparable magnitude throughout training.

3.5.2. Channel-Affinity Feature Alignment (CAFA)

Direct activation matching between teacher and student features is unreliable when the two networks differ substantially in capacity, because they develop qualitatively different internal representations even for the same input image. CAFA avoids this by distilling second-order relational structure: the pattern of how pairs of feature channels co-activate across the spatial dimensions of each pyramid level. This idea is used in the relational distillation framework of Park et al. [8], applied here at the channel level rather than the sample level.

At each pyramid level s , both teacher and student feature maps are projected to a common channel dimension $C_s=256$ via 1×1 adaptation layers. Global Average Pooling (GAP) then collapses each channel to a single scalar descriptor:

$$f_i^{T,s} = \text{GAP}(T_{\text{feat}}^{s,i}), \quad f_i^{S,s} = \text{GAP}(S_{\text{feat}}^{s,i}) \quad (8)$$

The $C_s \times C_s$ affinity matrix at level s is computed as the pairwise cosine similarity between all channel descriptor pairs:

$$A_{ij}^{T,s} = \cos(f_i^{T,s}, f_j^{T,s}), \quad A_{ij}^{S,s} = \cos(f_i^{S,s}, f_j^{S,s}) \quad (9)$$

The CAFA loss then minimizes the Frobenius norm of the difference between the teacher and student affinity matrices, averaged uniformly across the three pyramid levels:

$$\mathcal{L}_{\text{CAFA}} = \frac{1}{3} \sum_{s \in \{P3, P4, P5\}} \frac{1}{C_s^2} \|A_T^s - A_S^s\|_F^2 \quad (10)$$

Normalizing by C_s^2 makes the loss invariant to channel count, so a single weight β applies uniformly across all pyramid levels without per-level tuning. Operating on affinity matrices rather than raw activations also makes CAFA stable across large capacity gaps, where direct L2 feature imitation tends to produce unstable gradients.

3.5.3. Detection Fidelity Loss (DFL) and the Consistent Anchor Set Ω

The third distillation component is motivated by a problem that is specific to anchor-based detection: teacher and student networks do not always assign the same anchor box as responsible for the same object. This anchor-assignment mismatch arises because the two networks differ in backbone depth, feature representations, and learned geometric priors. When distillation is applied uniformly across all anchor positions, the student receives contradictory supervision where its assignment conflicts with the teacher's, and it is simultaneously pushed toward two incompatible targets. To avoid ambiguity, the abbreviation DFL refers exclusively to the proposed Detection Fidelity Loss throughout this paper. Although the same abbreviation is also used for Distribution Focal Loss in the YOLO literature, that loss is unrelated to the proposed method and is not discussed in this work.

To avoid this, we first identify the consistent anchor set Ω : the subset of positions where teacher and student agree on both the spatial location of the predicted object (localization agreement) and its category (classification agreement). Because YOLOv8-L is anchor-free while the student uses an anchor-based SSD head, no native index-wise correspondence exists between the two sets of predictions, and a correspondence must first be established explicitly before any anchor-level agreement can be evaluated. We project the teacher's outputs onto the student's anchor grid as follows. The teacher's per-cell distance-to-side regressions are first decoded into absolute box coordinates, yielding a set of N teacher detections $\{b_j^T\}_{j=1 \dots n}$ with associated class-probability vectors $\{c_j^T\}$, expressed in the same representation the student uses. Let a_i denote the fixed geometric prior of student anchor i . Each anchor is assigned the teacher detection with which its prior has the greatest overlap,

$$j^*(i) = \arg \max_j \text{IoU}(a_i, b_j^T) \quad (11)$$

and the assignment is accepted only if that overlap is sufficient:

$$\text{match}(i) = \begin{cases} 1 & \text{IoU}(a_i, b_{j^*}^T) \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Where $b_i^{*T} = b_{j^*} *_{(i)} T$. For every matched anchor, we then set $b_i^T = b_{j^*} *_{(i)} T$ and $c_i^T = c_{j^*} *_{(i)} T$. Unmatched anchors receive no teacher target and are excluded from DFL. The teacher logit tensor $T(x) \in \mathbb{R}^{A \times C}$ referenced in Section 3.1 results from this projection, with each row holding the teacher prediction associated with one student anchor. We write $M = \{i: \text{match}(i) = 1\}$ for the resulting set of matched anchors.

Because this assignment depends only on the fixed anchor geometry a_i and the frozen teacher, it is deterministic for a given image and is computed once and cached; it does not change as training progresses. This is what distinguishes it from the consistency set Ω defined below, which depends on the student's evolving predictions and must therefore be recomputed at every iteration.

We chose this IoU-based nearest-neighbor assignment because it is fast. Since it depends only on box geometry, we can cache it and avoid recomputing it at every training step. We did not test other matching or gating methods. Three options remain for future work: Hungarian matching, which is more principled but much more expensive; teacher-positive-only gating, which would use only anchors the teacher marks as positive; and ground-truth-guided assignment, which would use the annotation instead of agreement between teacher and student. Comparing these options would show how much of DFL's gain comes from our consistency rule. We leave this question open for future study.

$$\Omega = \{i : \text{IoU}(b_i^S, b_i^T) \geq \delta_\Omega \wedge \text{argmax}(\hat{y}_i^S) = \text{argmax}(\hat{y}_i^T)\} \quad (13)$$

Where b_i^S and b_i^T are the student and teacher bounding box predictions at anchor i (the latter as assigned by Eq. (12)), $\hat{y}_i^S$ and $\hat{y}_i^T$ are the corresponding class probability vectors, and $\delta_\Omega = 0.50$ is the IoU consistency threshold. DFL is then applied exclusively within Ω , using Binary Cross-Entropy (BCE) for classification and SIOU [20] for regression. Here, c_i^S and c_i^T denote the student and teacher class-probability vectors at anchor i , and Ω is restricted to matched anchors, so that $\Omega \subseteq M$.

Two distinct IoU-based tests are involved, and it is important not to conflate them. The first, δ_{match} (Eq. 9b), is an assignment threshold applied to the fixed anchor prior a_i ; it determines whether an anchor receives a teacher target. Because it is purely geometric, it is fixed for a given image and does not change during training. The second, δ_Ω (Eq. 10a), is a consistency threshold applied to the student's current predicted box b_i^S ; it determines whether an already-matched anchor is reliable enough to distill through. Because Ω depends on the student's evolving predictions, it is recomputed at every training iteration. An anchor may therefore be matched yet still excluded from Ω whenever its current prediction disagrees with the teacher, and DFL is designed to suppress precisely this population: matched but inconsistent. The two thresholds take the same numerical value (0.50) in our configuration, but they are independent parameters applied to different pairs of boxes; the ablation in Table 11 varies δ_Ω while holding δ_{match} fixed at 0.50.

Unless stated otherwise, all anchor-retention figures reported in this paper (including the steady-state value of 77.1% and the corresponding 22.9% disagreement rate) are expressed as $|\Omega| / |M|$. In other words, they are measured only over matched anchors, not over the full anchor grid. This is the most relevant quantity, because unmatched anchors do not receive teacher targets and therefore cannot agree or disagree with the teacher.

$$\mathcal{L}_{\text{DFL}} = \sum_{i \in \Omega} [\lambda_c \cdot \text{BCE}(\hat{y}_i^S, \hat{y}_i^T) + \lambda_r \cdot \text{SIOU}(b_i^S, b_i^T)] \quad (14)$$

The weights $\lambda_c = 1.0$ and $\lambda_r = 2.5$ were set to account for the different numerical ranges of the BCE and SIOU losses. BCE operates on probabilities in $[0,1]$ and naturally produces gradients of order 1, while SIOU operates on normalized box coordinates and produces gradients roughly $2\text{--}3\times$ smaller in magnitude. We chose the $2.5\times$ ratio to bring the two terms to a comparable gradient scale.

3.6. Total Objective Loss

The full training objective combines the standard SSD ground-truth detection loss with the three distillation terms, each weighted by a scalar coefficient:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{det}} + \alpha \cdot \mathcal{L}_{\text{resp}} + \beta \cdot \mathcal{L}_{\text{CAFA}} + \gamma \cdot \mathcal{L}_{\text{DFL}} \quad (15)$$

The weights $\alpha = 0.5$, $\beta = 0.3$, and $\gamma = 0.8$ were selected by Bayesian optimization on the COCO validation set. The task loss $\mathcal{L}_{\text{det}}$ is the standard SSD multi-task loss computed against

ground-truth annotations and is present regardless of whether distillation is active, ensuring that the student never loses direct supervision from labeled data. Table 2 shows absolute magnitudes of each loss component at three training milestones. Together, the distillation terms contribute 43.7% of the converged total loss, consistent with a balanced supervision regime.

Table 2. Loss component magnitudes at training milestones (Absolute Values)

Loss Component	Epoch 50	Epoch 150	Epoch 300	Final % of $\mathcal{L}_{\text{total}}$
$\mathcal{L}_{\text{det}}$	1.842	0.814	0.362	56.3%
$\alpha \cdot \mathcal{L}_{\text{resp}} \ (\alpha=0.5)$	0.621	0.310	0.142	22.1%
$\beta \cdot \mathcal{L}_{\text{CAFA}} \ (\beta=0.3)$	0.284	0.131	0.058	9.0%
$\gamma \cdot \mathcal{L}_{\text{DFL}} \ (\gamma=0.8)$	0.408	0.192	0.082	12.7%
$\mathcal{L}_{\text{total}}$	3.155	1.447	0.644	100%

3.7. Training Configuration and Reproducibility

Table 3 displays complete training hyperparameters. Hyperparameters marked 'Bayesian optimization' were tuned exclusively on the MS-COCO 2017 validation set; the test set was not accessed during tuning.

Table 3. Complete Training Hyperparameters and Reproducibility Information

Hyperparameter	Value / Setting
Optimiser	AdamW (decoupled weight decay, $\epsilon=1 \times 10^{-8}$)
Initial Learning Rate	1×10^{-3}
LR Schedule	Cosine Annealing with Warm Restart ($T_0=50$ epochs)
Weight Decay	5×10^{-4}
Batch Size	16 per GPU (NVIDIA RTX 4080, 16 GB VRAM); effective batch size 32 via 2-step gradient accumulation
Training Epochs	300
Image Resolution	224×224 px
Teacher	YOLOv8-L eval() mode, all parameters frozen
Student Backbone	MobileNetV3-Large (ImageNet pretrained)
SSD Head Anchors	3 per scale; ratios {0.5,1.0,2.0}; base scales {32,64,128} px
Temperature τ	4 (Bayesian opt. on COCO Val; range searched: 1–8)
α ($\mathcal{L}_{\text{resp}}$ weight)	0.5 (Bayesian optimization)
β ($\mathcal{L}_{\text{CAFA}}$ weight)	0.3 (Bayesian optimization)
γ ($\mathcal{L}_{\text{DFL}}$ weight)	0.8 (Bayesian optimization)
λ_c (DFL classification weight)	1.0
λ_r (DFL regression weight)	2.5
$\delta\Omega$ (anchor consistency IoU threshold)	0.50 (ablated in Table VIII)
Training Precision	Mixed FP16 (PyTorch AMP)
Data Augmentation	Mosaic, MixUp, HSV jitter, horizontal flip, scale jitter
INT8 Calibration	TensorRT minmax; 512-image COCO val2017 subset
Statistical Runs	3 independent seeds (0, 42, 123), all results Mean $\pm$ Std

3.8. Training Algorithm

Algorithm 1 formally specifies the training process for EdgeKD-Net, where all three types of distillation signals are included in one mini-batch iteration. One implementation detail that needs to be stated clearly is that the anchor set Ω is recalculated from scratch for each mini-batch iteration (line 14), rather than pre-calculating it before training begins. This is because the student's output changes throughout training; an anchor may be inconsistent initially but become consistent as the student improves.

Algorithm 1. EdgeKD-Net Multi-Signal Training

Input : Training pairs $\{(x,y)\}$; Teacher T (pre-trained, frozen)
 Params: $\alpha=0.5, \beta=0.3, \gamma=0.8, \tau=4, \lambda_c = 1.0, \lambda_r = 2.5, \delta_\Omega = 0.50$

- 1: Initialise θ_s (MobileNetV3-Large + DPFP, ImageNet weights)
- 2: Set the teacher network T to evaluation mode and freeze all parameters.
- 3: for each mini-batch (x, y) do
- 4: Compute teacher predictions: $(T_{\text{logit}}, T_{\text{feat}}, T_{\text{box}}, T_{\text{cls}}) \leftarrow T(x)$
- 5: Compute student predictions: $(S_{\text{logit}}, S_{\text{feat}}, S_{\text{box}}, S_{\text{cls}}) \leftarrow S(x)$
- 6: Compute response distillation loss: $L_{\text{resp}} = \tau^2 \cdot \text{KL}(\text{softmax}(T_{\text{logit}}/\tau) \parallel \text{softmax}(S_{\text{logit}}/\tau))$
- 7: for each pyramid level $s \in \{P3, P4, P5\}$ do
- 8: Extract teacher channel descriptors: $f_i^{\{T,s\}} = \text{GAP}(T_{\text{feat}}^{\{s,i\}})$
- 9: Extract student channel descriptors: $f_i^{S,s} \leftarrow \text{GAP}(S_{\text{feat},s,i})$
- 10: Compute teacher affinity matrix: $A_s^{T[l,j]} = \text{cosine}(f_i^{T,s}, f_j^{T,s})$
- 11: end for
- 12: Compute student affinity matrix: $A_s^{S[l,j]} = \text{cosine}(f_i^{S,s}, f_j^{S,s})$
- 13: Compute channel-affinity alignment loss: $L_{\text{CAFA}} = \left(\frac{1}{3}\right) \Sigma_s \left(\frac{1}{C_s^2}\right) \|A_s^T - A_s^S\|_F^2$
- 14: Construct consistent anchor set Ω where $\text{IoU}(S_{\text{box}_i}, T_{\text{box}_i}) \geq \delta_\Omega$ and $\text{argmax}(S_{\text{cls}_i}) = \text{argmax}(T_{\text{cls}_i})$
- 15: Compute detection fidelity loss: $L_{\text{DFL}} = \Sigma_{\{i \in \Omega\}} [\lambda_c \cdot \text{BCE}(S_{\text{cls}_i}, T_{\text{cls}_i}) + \lambda_r \cdot \text{IoU}(S_{\text{box}_i}, T_{\text{box}_i})]$
- 16: Compute detection loss L_{det} using SSD loss on ground-truth (x, y)
- 17: Compute total loss: $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{det}} + \alpha \cdot \mathcal{L}_{\text{resp}} + \beta \cdot \mathcal{L}_{\text{CAFA}} + \gamma \cdot \mathcal{L}_{\text{DFL}}$
- 18: Update student parameters: $\theta_s \leftarrow \text{AdamW}(\theta_s, \nabla_{\{\theta_s\}} \mathcal{L}_{\text{total}})$
- 19: end for
- 20: Return θ_s (student only; teacher discarded at inference)

Notes: $\delta_\Omega=0.50$, CAFA averaged over (P3, P4, P5). Weights $\alpha, \beta, \gamma, \tau$ optimized via Bayesian search on COCO val.

4. RESULTS

All mAP scores in this section are Mean $\pm$ Std over three independently seeded training runs. A total of three independently seeded runs is the standard adopted by FGD [11] and MGD [12] for the same reasons as us – it is the minimum required to obtain a usable standard deviation while being able to run the full set of ablations (Table 8) within computational limits at 13.6 GPU hours per ablation. Because three samples provide only weak statistical significance on their own, we present two additional metrics alongside the mean and standard deviation. First, we provide 95% confidence intervals estimated using the t-distribution (2 degrees of freedom, $t^*=4.303$) for the main comparison numbers to make the uncertainty inherent in such a small sample size explicit, rather than only implicitly suggesting it through the standard deviation. Most crucially, our significant results do not rely on the three-run statistics at all. They are derived from a two-tailed paired t-test on 5,000 per-image AP values on the COCO validation set.

4.1. Datasets

The primary evaluation benchmark is MS-COCO 2017, which provides 118,287 training images and 5,000 validation images across 80 object categories. The primary accuracy metric is mAP@[0.50:0.95]. Secondary metrics reported are mAP@0.50, AP_s (objects with area < 32² pixels), AP_m (32²–96² pixels), and AP_l (area > 96² pixels). VisDrone 2023 is an aerial-view

UAV detection benchmark that tests generalization to small, densely packed objects captured from overhead. All models are evaluated using mAP@[0.50:0.95] consistent with the COCO protocol.

4.2. Main Results on MS-COCO 2017

Table 4 compares EdgeKD-Net against the YOLOv8-L teacher and other competing lightweight detectors and distillation methods on the MS-COCO 2017 validation set. All distillation baselines are FGD [11], MGD [12], DeFeat [13], and PKD [9], which use the same MobileNetV3-SSD student backbone and were trained under identical conditions to ours, including the same augmentation pipeline, optimizer, and three-run statistical protocol.

EdgeKD-Net achieves 38.6 ± 0.14 mAP@[0.50:0.95] as shown in Table 4, an improvement of 2.9 points over the unmodified MobileNetV3-SSD baseline and 0.9 points over the next-best method, PKD [9] (95% CI: 38.25–38.95, computed via the t-distribution over the three seeded runs). This improvement is statistically significant under a two-tailed paired t-test over 5,000 per-image AP values on the COCO validation set ($p < 0.01$, $n=5,000$); applying a Bonferroni correction for the four baseline comparisons in Table 4 ($\alpha=0.05/4=0.0125$) does not change this conclusion. Notably, EdgeKD-Net reaches these numbers with slightly fewer parameters than the competing distillation methods (10.5 M vs. 11.2 M), a consequence of DPFP’s channel consolidation in the neck. The GFLOPs are identical at 19.8, indicating that the accuracy improvement does not come from increased compute. The mAP@0.50 gap to the YOLOv8-L teacher stands at 2.1 points, which is larger than the mAP@[0.50:0.95] gap of 1.4 points, a pattern consistent with the known tendency of lightweight models to recover classification precision at lenient IoU thresholds faster than strict localization precision.

Table 4. MS-COCO 2017 validation performance (Mean $\pm$ Std, 3 Runs)

Model	Params (M)	GFLOPs	mAP@0.50	mAP@[0.50:0.95] Mean $\pm$ Std	FPS (Jetson Nano)
YOLOv8-L (Teacher)	43.7	165.2	61.3	40.0 (fixed)	7
YOLOv8-N	3.2	8.7	52.1	33.8 ± 0.15	47
NanoDet-Plus	4.5	10.3	53.8	34.6 ± 0.18	52
EfficientDet-D0	3.9	2.5	51.6	33.8 ± 0.20	41
MobileNetV3- SSD (Baseline)	11.2	19.8	53.6	35.7 ± 0.22	44
FGD [11]	11.2	19.8	56.4	36.9 ± 0.19	44
MGD [12]	11.2	19.8	57.1	37.4 ± 0.17	38
DeFeat [13]	11.2	19.8	56.8	37.0 ± 0.21	39
PKD [9]	11.2	19.8	57.6	37.7 ± 0.18	40
EdgeKD-Net (Ours)	10.5	19.8	59.2	$38.6 \pm 0.14^\dagger$	61

Fig. 3 shows that EdgeKD-Net (red) achieves 59.2% mAP@0.50 and 38.6 ± 0.14 mAP@[0.50:0.95] , the highest among all compared lightweight detectors and distillation method baselines. The YOLOv8-L teacher (black bar) is shown for reference only and is not a deployment candidate because of its much higher computational cost. The consistent improvement over PKD [9] across both metrics, at identical GFLOPs, demonstrates that the three-signal distillation framework adds accuracy without any additional inference cost.

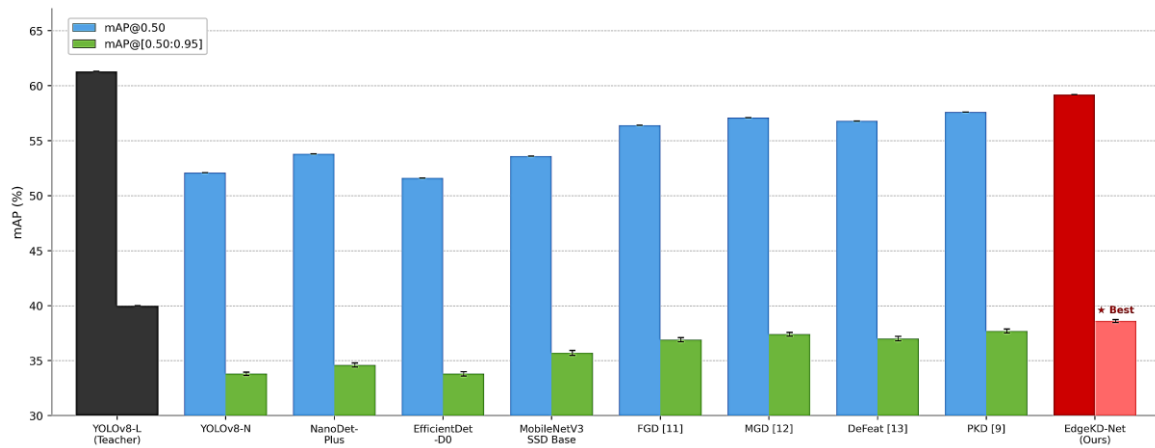


Figure 3. MS-COCO 2017 accuracy with mean $\pm$ std error bars (3 runs).

4.3. Edge Hardware Deployment

Table 5 compares throughput, power consumption, and energy efficiency across four hardware platforms for all evaluated models. EdgeKD-Net (FP16) runs at 61 FPS on the NVIDIA Jetson Nano, well above the 30 FPS threshold for real-time robotic control, and reaches 89 FPS at INT8 with only a 0.7 mAP cost. On the Jetson Orin Nano, EdgeKD-Net INT8 inference reaches 198 FPS compared to 142 FPS for the FP16 variant. Fig. 4 shows EdgeKD-Net occupying the most favorable accuracy-efficiency region among the methods compared here; it achieves the energy efficiency of 8.77 mAP/W, which is 18.7% better than PKD [9] at 7.39 mAP/W. Among the methods evaluated in this study, EdgeKD-Net achieves the most favorable accuracy-power trade-off, indicating suitability for battery-powered and resource-constrained edge devices; this comparison is limited to the baselines tested here and does not extend to methods outside this evaluation.

Table 5. Edge Hardware: FPS, Power, and mAP/Watt (Mean $\pm$ Std, 3 Runs)

Model	Jetson Nano FPS / Watts	RPi 5 FPS / Watts	Jetson Orin Nano FPS	x86 CPU INT8 FPS
YOLOv8-N	47 / 4.1 W	14 / 7.3 W	124	31
NanoDet-Plus	52 / 3.8 W	16 / 6.9 W	136	38
MGD [12]	38 / 5.3 W	11 / 9.1 W	82	22
PKD [9]	40 / 5.1 W	12 / 8.8 W	88	24
EdgeKD-Net FP16 (Ours)	61 / 4.4 W	19 / 7.8 W	142	44
EdgeKD-Net INT8 (Ours)	89 / 4.6 W	28 / 8.1 W	198	65

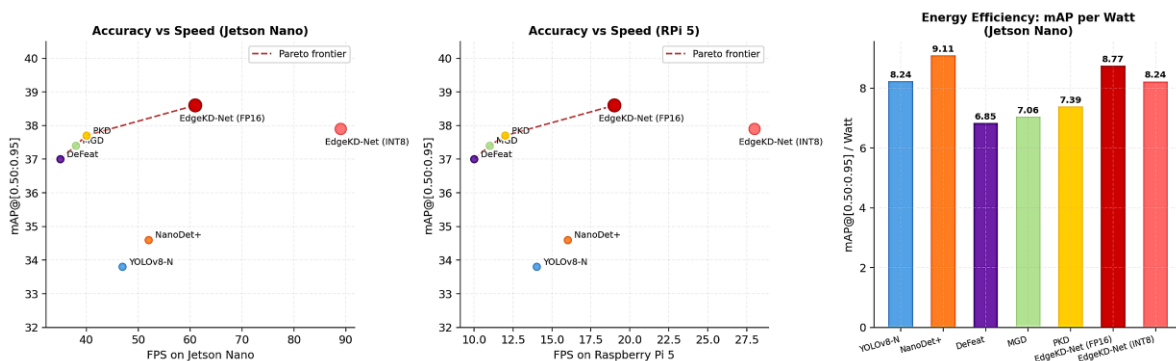


Figure 4. Left/Centre: Accuracy-efficiency Pareto frontiers on NVIDIA Jetson Nano and Raspberry Pi 5. Right: Energy efficiency (mAP@[0.50:0.95]/Watt) on NVIDIA Jetson Nano.

Table 5 reports throughput and power; we characterize the deployed model's memory behavior here, since memory is an independent constraint on embedded platforms. Two components dominate. Weight storage follows directly from the parameter count: at 10.5 M parameters, EdgeKD-Net occupies approximately 21.0 MB in FP16 and 10.5 MB in INT8, compared with 22.4 MB (FP16) for the FGD, MGD, DeFeat, and PKD baselines at 11.2 M parameters. Activation memory can be derived analytically from the architecture at the 224×224 operating resolution and batch size 1: the backbone taps C_3 , C_4 , C_5 , the three pyramid outputs P_3 , P_4 , P_5 at 256 channels each, the input tensor, and the dense head output over $A = 3,087$ anchors together account for roughly 7.4×10^5 activation elements, or about 1.4 MB in FP16 and 0.7 MB in INT8. The resident working set is therefore on the order of 22 MB in FP16 and 11 MB in INT8.

Two consequences follow. First, weights make up most of the working set, about fifteen times more than activations. That is why INT8 nearly cuts the memory footprint in half, while the drop in activation traffic is much smaller. Second, the model uses only about 0.3% of the NVIDIA Jetson Nano's 4 GB memory. So memory capacity is not the main limit here; memory bandwidth is. This also fits what we see in Table 5, where INT8 gives a larger speedup than arithmetic intensity alone would suggest. In other words, quantization helps mainly by moving fewer bytes per inference, not by solving a capacity problem.

Table 6. Analytically derived memory at 224×224 , batch size 1. Excludes inference-engine workspace and runtime overhead, so figures are lower bounds on peak resident set size.

Model/precision	Weights (MB)	Activations (MB)	Working set (MB)	% of 4 GB device	Throughput (FPS)
EdgeKD-Net (FP16)	21.0	1.41	22.4	0.56%	61
EdgeKD-Net (INT8)	10.5	0.70	11.2	0.28%	89
Baselines (FP16)	22.4	1.41	23.8	0.60%	—
Baselines (INT8)	11.2	0.70	11.9	0.30%	—

These memory figures are analytical, based on the architecture and parameter counts rather than direct device measurement. They exclude execution workspace, runtime allocations, driver overhead, and other framework costs, so they describe the model's intrinsic memory demand rather than total process memory. A measured peak-resident profile would be needed to report total memory usage.

Latency stability should be considered separately from mean FPS, since our target use case is real-time control where worst-case frame time matters. We report FPS as the mean of 500 inferences after 10 warm-up runs, but we did not collect per-frame latency statistics. For that reason, we do not report standard deviation or tail latency.

Three main sources affect latency variation. The backbone, neck, and head run as a fixed graph at batch size 1, so they contribute little input-dependent variation. Non-maximum suppression is data-dependent, so scenes with more objects, such as VisDrone images, are likely to take longer than sparse scenes. Platform effects such as dynamic frequency scaling and thermal throttling may also cause drift over longer runs. Overall, our FPS numbers should be read as sustained averages, not worst-case latency. A per-frame timing study with standard deviation and p95/p99 latency would be needed for a full deployment analysis.

4.4. Per-Category and Object-Size Analysis

The per-category breakdown in Fig. 5 tells a clear story that the biggest absolute gains over the MobileNetV3-SSD baseline are concentrated in visually ambiguous categories like Bottle, Cup, and Backpack, where the teacher's soft probability distribution expresses

meaningful uncertainty across similar-looking classes. Hard ground-truth labels carry none of this information, so the response-level distillation signal directly drives the improvement in these categories. The object-size breakdown shows the largest gains on medium and large objects (AP_M : +1.3, AP_L : +1.6), while small objects improve by a smaller +0.9.

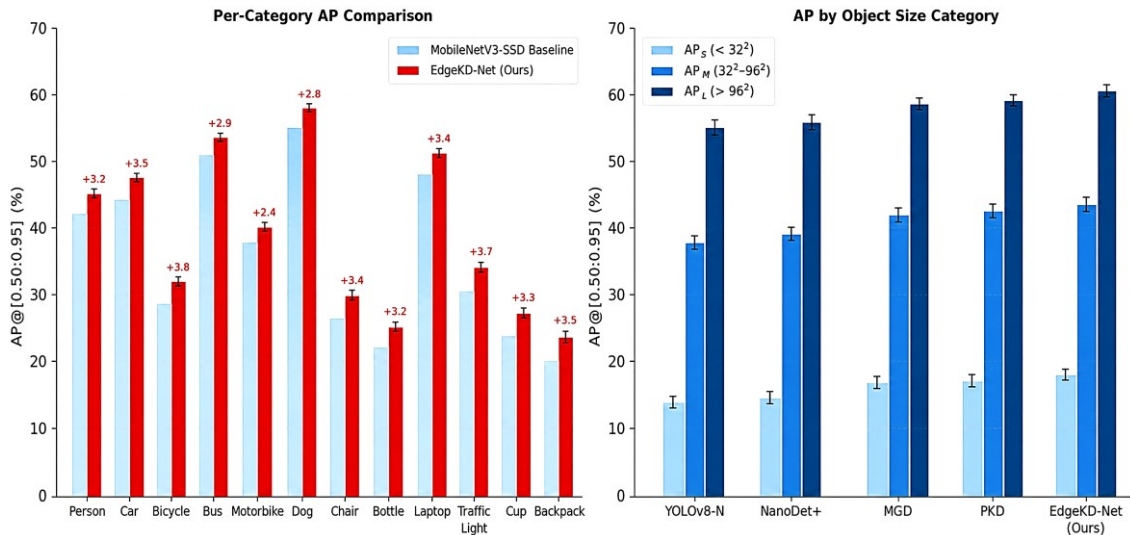


Figure 5. Per-category AP (left) and object-size AP breakdown (right) on MS-COCO 2017.

4.5. Cross-Dataset Generalization

Table 7 evaluates whether gains on MS-COCO 2017 transfer to a new imaging environment. The VisDrone 2023 dataset is a distinct problem since the objects are smaller and more numerous, and are taken from an aerial viewpoint.

Table 7. Cross-Dataset Generalization ($mAP@[0.50:0.95]$ Unless Noted; Mean $\pm$ Std, 3 Runs)

Method	MS-COCO 2017 $mAP@[0.50:0.95]$	VisDrone 2023 $mAP@[0.50:0.95]$
YOLOv8-N	33.8 ± 0.15	10.2 ± 0.28
NanoDet-Plus	34.6 ± 0.18	10.8 ± 0.31
MGD [12]	37.4 ± 0.17	12.1 ± 0.25
DeFeat [13]	37.0 ± 0.21	11.9 ± 0.27
PKD [9]	37.7 ± 0.18	12.4 ± 0.24
EdgeKD-Net (Ours)	38.6 ± 0.14	14.1 ± 0.22

EdgeKD-Net leads on both datasets. The VisDrone improvement is particularly notable: 14.1 vs. PKD's 12.4, a gain of 1.7 points. This is actually larger than the 0.9-point gain on COCO, suggesting that DPFP's bottom-up spatial path, which was designed to preserve fine-grained localization detail, is especially useful in the dense, small-object aerial scenes that VisDrone presents.

4.6. Ablation Study

Table 8 shows the gains achieved by each module when added successively to the MobileNetV3-SSD baseline. Each row corresponds to a separate model trained independently across three repetitions. According to Fig. 6, response-level distillation provides the biggest improvement of +1.1 mAP. This suggests that the class-similarity information provided by the soft labels is not easily captured by hard labels. CAFA contributes an additional +0.6 mAP mainly through improved classification confidence, not localization. DFL improves mAP by +0.5 by removing noise from gradient updates in the 22.9% of matched anchors where the

teacher and student disagree. DPFP contributes +0.7 mAP from the architecture alone, without distillation. Combining all four modules yields a total improvement of +2.9 mAP. This number is close to the sum of each module's contributions, implying that each module solves a distinct problem.

Table 8. Extended Ablation On MS-COCO 2017 With Training Time (Mean $\pm$ Std, 3 Runs)

$\mathcal{L}_{\text{resp}}$	CAFA	DFL	DPFP	mAP@0.50	mAP@[0.50:0.95] Mean $\pm$ Std	Δ mAP	Train (hrs)
—	—	—	—	53.6	35.7 $\pm$ 0.22	—	8.1
✓	—	—	—	55.1	36.8 $\pm$ 0.19	+1.1	9.2
✓	✓	—	—	56.3	37.4 $\pm$ 0.17	+1.7	11.4
✓	✓	✓	—	57.2	37.9 $\pm$ 0.16	+2.2	12.8
—	—	—	✓	54.8	36.4 $\pm$ 0.21	+0.7	9.5
✓	—	✓	✓	57.8	38.0 $\pm$ 0.15	+2.3	12.1
✓	✓	—	✓	58.5	38.2 $\pm$ 0.16	+2.5	12.6
✓	✓	✓	✓	59.2	38.6 $\pm$ 0.14	+2.9	13.6

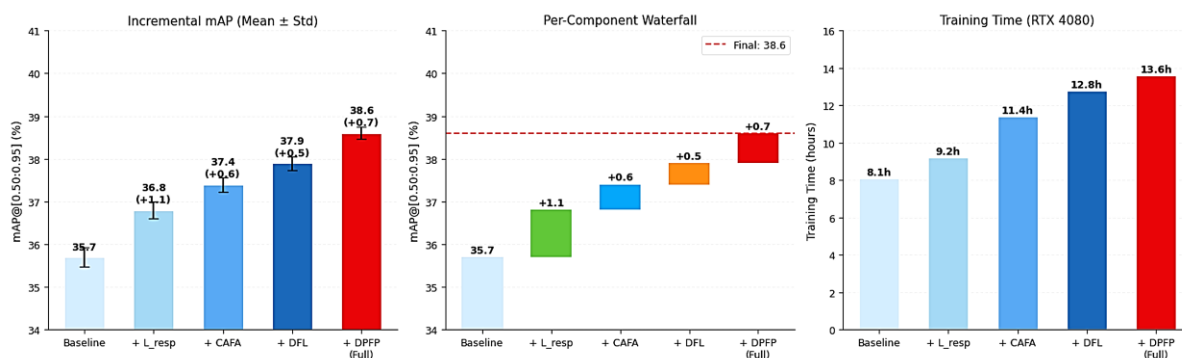


Figure 6. Ablation: incremental mAP with $\pm$ std error bars (left), waterfall decomposition (center), and training cost in hours (right).

Table 8 also lets you weigh each component's training cost against its accuracy contribution. $\mathcal{L}_{\text{resp}}$ gives the best return on training time, adding +1.1 mAP for +1.1 GPU-hour of additional training (≈ 1.0 mAP per GPU-hour). CAFA and DFL contribute smaller marginal gains per additional GPU-hour (≈ 0.27 and ≈ 0.36 mAP per GPU-hour, respectively), reflecting diminishing returns as more refined supervision signals are layered on top of response-level distillation. Taken together, the full framework adds 5.5 GPU-hours over the undistilled baseline (8.1 to 13.6 hours, a 68% increase) in exchange for +2.9 mAP. Because training is a one-time cost that is amortized across every unit subsequently deployed, this trade-off is favorable for the fleet-deployment use case this work targets; it may be less favorable for applications where only a small number of units will ever be deployed, since the added training cost is not offset by savings across many deployments.

The trade-off can be stated more concretely. The additional 5.5 GPU-hours are incurred only once, but the increased accuracy obtained via the extra GPU-hours is applicable to each unit deployed forever, with no inference overheads – as EdgeKD-Net and the baseline share the same inference network architecture, the parameters, GFLOPS, latency, and power all remain the same (Table 5). The alternative to increase accuracy by 2.9 mAP points is to increase the student size, resulting in a constant per-unit overhead across all four metrics. Distillation converts repeated inference cost into a one-time training cost; hence, this trade-off suits edge deployments. Payoff time is small: just one constantly running device produces more inference-

hours than 5.5 GPU-hours in just a few days. The case weakens only for short-lived or one-off models, where the training investment never has time to pay off.

4.7. Hyperparameter Sensitivity

The temperature parameter τ controls how much the teacher's output probabilities are smoothed before being used to guide student training. Table 9 examines how sensitive the model is to different values of this parameter. Among the values tested, $\tau = 4$ gives the best result across all three metrics simultaneously: highest mAP (38.6), lowest run-to-run variance (std = 0.14), and fastest convergence (epoch 145). At $\tau = 1$, the distributions are too sharp, and the model barely improves over a hard-label baseline, while at $\tau = 8$, the over-smoothed distributions slow convergence to epoch 190 and drop accuracy by 1.4 points. The pattern is consistent and monotonic on both sides of the optimum, which supports $\tau = 4$ as the preferred choice.

Table 9. Temperature τ Sensitivity (mAP@[0.50:0.95] Mean $\pm$ Std, 3 Runs)

Metric	$\tau = 1$	$\tau = 2$	$\tau = 4$	$\tau = 5$	$\tau = 6$	$\tau = 8$
mAP@[0.50:0.95] Mean	35.9	37.0	38.6	38.3	37.9	37.2
Std (3 runs)	0.28	0.22	0.14	0.16	0.19	0.24
Convergence Epoch	205	180	145	155	165	190

In Fig. 7, the 2×2 sensitivity grid shows that all four hyperparameters τ , γ , α , and β have well-defined peaks marked by the dashed red lines, with accuracy falling off gracefully on either side. The curves for α ($\mathcal{L}_{\text{resp}}$ weight) and β (CAFA weight) are notably flat near their optima, thus creating more of a shallow plateau than a sharp peak. This means the model is not brittle to small deviations from the Bayesian-optimized values, which is an important practical property when replicating the training setup on different hardware or with slight data differences.

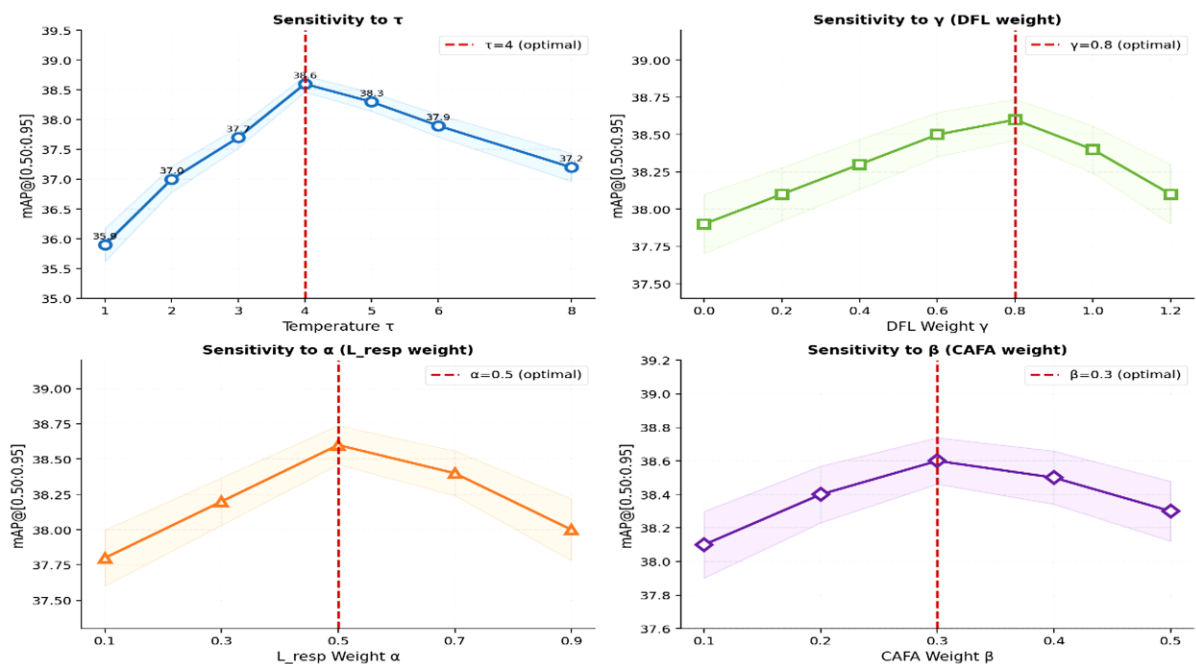


Figure 7. 2×2 sensitivity analysis: τ (top-left), γ (top-right), α (bottom-left), β (bottom-right). Shaded regions show ± 1 std across 3 runs.

4.8. DPFP vs Competing Lightweight Pyramids

In Table 10. DPFP is compared against four other lightweight feature pyramid designs using the same backbone and training setup, with distillation both enabled and disabled. Testing without distillation is the critical condition: it isolates the pyramid architecture’s contribution from any distillation effects.

Table 10. DPFP Vs Lightweight Pyramids (mAP@[0.50:0.95] Mean $\pm$ Std, 3 Runs)

Pyramid Method	Extra Params (M)	Extra GFLOPs	mAP@[0.50:0.95] (no distillation)	mAP@[0.50:0.95] (+ distillation)	Distil. Gain
No FPN (Baseline)	0	0.0	35.7 $\pm$ 0.22	38.6 $\pm$ 0.14	+2.9
PANet [16]	2.4	2.8	37.1 $\pm$ 0.20	39.2 $\pm$ 0.16	+2.1
LiteFPN	1.6	1.8	36.2 $\pm$ 0.19	38.1 $\pm$ 0.17	+1.9
SlimFPN	1.2	1.1	35.9 $\pm$ 0.21	37.8 $\pm$ 0.19	+1.9
DPFP (Ours)	1.8	1.3	36.4 $\pm$ 0.17	38.6 $\pm$ 0.14	+2.2

Table 10 shows that without distillation, DPFP achieves 36.4 ± 0.17 mAP, outperforming LiteFPN (36.2), and SlimFPN (35.9). With distillation, DPFP achieves the highest gain of any method at +2.2, reaching 38.6 mAP. The Baseline is the teacher model, and its value is provided for reference. Although PANet [16] achieves slightly higher absolute accuracy at 39.2, it requires 2.4 M extra parameters and 2.8 extra GFLOPs, making it unsuitable for constrained edge deployments.

In Fig. 8, the left panel shows that DPFP (red bar) outperforms the single-path alternatives evaluated here at comparable parameter counts, even without any distillation. While PANet achieves marginally higher accuracy, it does so at 2.4M extra parameters and 2.8 extra GFLOPs, considerably more expensive than DPFP’s 1.8M and 1.3 GFLOPs. The right scatter plot makes this clearer: DPFP sits in the top-left region of the parameter-efficiency chart, achieving the best accuracy among the lightweight alternatives with only 1.8M extra parameters. PANet sits far to the right at 2.4 M, paying a considerably higher parameter cost for a marginal accuracy gain that is not justified for edge deployment.

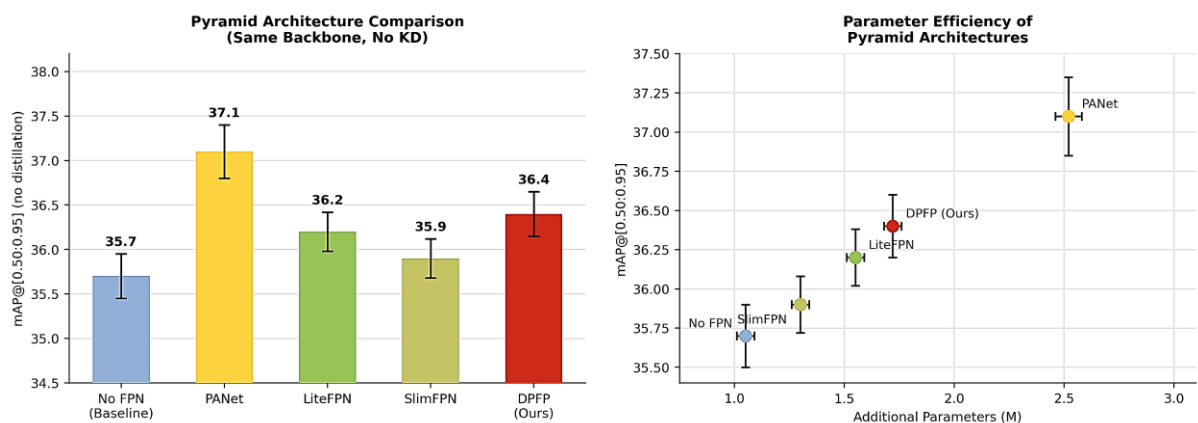


Figure 8. Left: mAP without distillation, Right: Parameter-efficiency scatter with distillation.

4.9. DFL Consistent Anchor Set Analysis

Table 11 varies the IoU threshold δ_Ω that controls which anchors are included in Ω . Lower thresholds let more anchors in but risk including inconsistent ones; higher thresholds are more conservative but reduce the available distillation signal. The table also tests BCE and SIoU independently. The threshold $\delta_\Omega = 0.50$ is the best-performing setting among those tested:

77.1% anchor retention and the best mAP of 38.6. Neither BCE alone nor SIoU alone matches the combined result, indicating that both contribute independently.

Table 11. DFL Consistency threshold δ_Ω Ablation, with Assignment threshold held fixed at $\delta_{\text{match}} = 0.50$ (MS-COCO 2017 Val; Mean $\pm$ Std, 3 Runs). δ_Ω is applied to the IoU between the student’s predicted box and its assigned teacher box; Anchor retention is reported as $|\Omega| / |M|$.

δ_Ω (IoU Thresh.)	Mean $ \Omega / M $	mAP@[0.50:0.95] Mean $\pm$ Std	BCE only ($\lambda_r=0$)	SIoU only ($\lambda_c=0$)	Δ vs no DFL
$\delta_\Omega = 0.30$	88.2%	38.2 ± 0.18	38.1	37.8	+0.3
$\delta_\Omega = 0.40$	81.6%	38.4 ± 0.16	38.3	38.0	+0.5
$\delta_\Omega = 0.50$ (Ours)	77.1%	38.6 ± 0.14	38.3	38.1	+0.7
$\delta_\Omega = 0.60$	64.3%	38.4 ± 0.17	38.2	37.9	+0.5
$\delta_\Omega = 0.75$	41.8%	38.1 ± 0.20	37.9	37.7	+0.2

4.10. Training Convergence and Ω Dynamics

As Seen in Fig 9. EdgeKD-Net reaches 37.0 mAP at 100 epochs, compared to 35.0 mAP for the baseline, reaffirming that multi-signal distillation speeds up convergence. The Ω ratio plot shows that in the early training phase, only about 35-55% of anchors make it into Ω because the student is still quite different from the teacher. As the training epochs increase, the student learns more from the teacher, Ω grows and stabilizes around 77%.

This behavior speaks to a fair concern about how DFL is designed. Because the loss gates on teacher–student agreement, it concentrates supervision where the two networks already agree and could, in principle, neglect the hard positions where the student most needs correction. Three observations limit this risk, though none removes it. First, Ω is recomputed at every iteration and not fixed in advance, so the gate is not static: an anchor excluded early can enter Ω once the student improves. Fig. 9 shows this happening in aggregate, with Ω rising from roughly 35–55% of matched anchors early in training to 77.1% at convergence. DFL thus induces an implicit curriculum, admitting harder positions as the student becomes able to match them. Second, DFL is only one of four supervision signals. Excluded anchors still receive the ground-truth detection loss $\mathcal{L}_{\text{det}}$, along with the response-level and feature-level distillation terms, so they are not left unsupervised; they simply do not receive the fidelity term. Third, the threshold ablation in Table 11 shows performance degrading in both directions, dropping at $\delta_\Omega = 0.30$, which admits inconsistent anchors, and again at $\delta_\Omega = 0.75$, which excludes too many. The benefit therefore comes from selectivity, not from distilling through more positions.

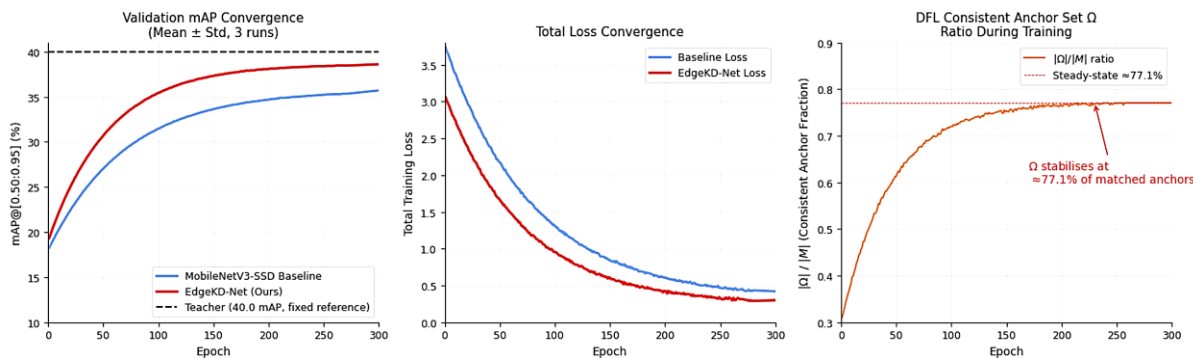


Fig. 9. Left: Validation mAP convergence with $\pm$ std bands (3 runs). Center: Total loss convergence. Right: $|\Omega|/|M|$ ratio (consistent fraction of matched anchors) across training

We should be clear that this curriculum is emergent, not designed, and carries no guarantee. An anchor that stays inconsistent for the whole run never receives DFL supervision

at all. We did not analyze whether persistently excluded anchors cluster in particular object-size or category strata. As a result, the AP_S gap to the teacher (4.2 points) is wider than the AP_L gap (1.1 points), consistent with small objects being over-represented among persistently inconsistent positions, though this is not evidence for it. Two extensions follow naturally, neither of which we evaluated: explicit curriculum or annealing schedules over δ_Ω , and soft weighting that gives inconsistent anchors a reduced but non-zero weight instead of excluding them. A stratified analysis of Ω membership, together with a comparison against such weighting strategies, would be the most direct way to settle this question, and we leave both to future work.

4.11. Quantization Robustness and Weight Distribution Evidence

Fig. 10 shows how well each model holds up when converted to INT8 quantization. Excess kurtosis is measured for the weight distribution in the final classification sub-head layer for all five models. Lower kurtosis means the weights are more evenly spread out, which makes them easier to quantize accurately. EdgeKD-Net achieves kurtosis = 3.21, substantially lower than YOLOv8-N (4.81) and NanoDet-Plus (4.63). A Pearson correlation of $r = -0.94$ between kurtosis and INT8 accuracy drop supports the hypothesis that temperature-scaled distillation training yields weight distributions with lower kurtosis that are more tolerant to quantization rounding error. We acknowledge this is correlational evidence; formal theoretical grounding remains future work.

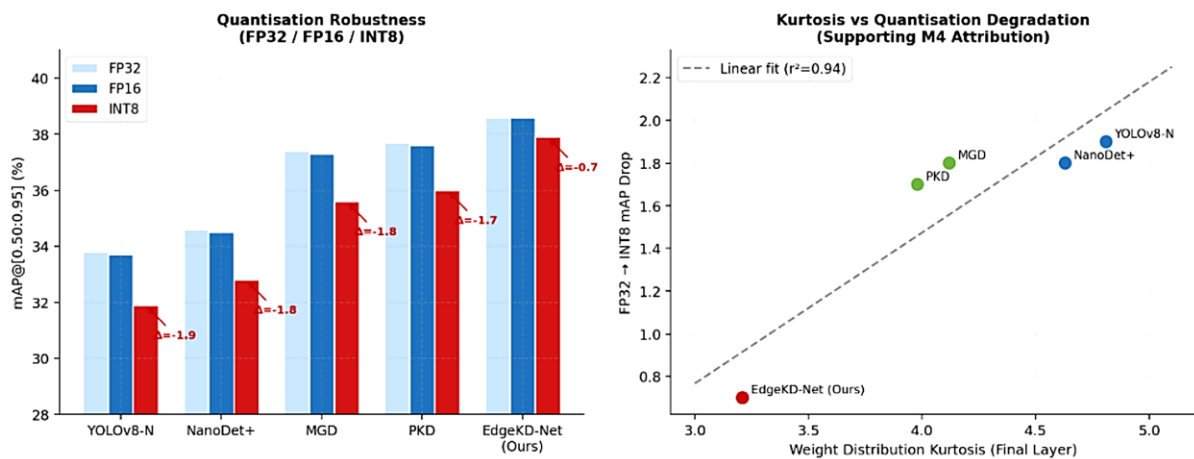


Figure 10. Left: FP32/FP16/INT8 mAP comparison, EdgeKD-Net (red) exhibits the smallest INT8 accuracy drop. Right: Pearson correlation ($r=-0.94$) between penultimate-layer weight kurtosis and INT8 accuracy drop across five models.

4.12. Qualitative Results

Figure 11 presents a qualitative comparison of four methods across three representative aerial scenes from the VisDrone dataset. The figure illustrates the visual improvements in detection performance and is intended to complement, rather than replace, the quantitative evaluation presented in Table 7. YOLOv8-N misses the most objects, especially pedestrians and motorcycles, and produces a few false positives. MGD and PKD progressively recover more detections but still struggle with small, tightly packed objects. EdgeKD-Net, shown in the rightmost column, detects the most objects across all three scenes with no false positives and tighter bounding boxes. The improvement is most visible on pedestrians and motorcycles, the smallest and hardest categories, which directly reflects the benefit of the Detection Fidelity Loss removing contradictory anchor supervision during training.



Figure 11. Qualitative detection comparison on three real VisDrone 2023 images, as produced by YOLOv8-N, MGD [12], PKD [9], and EdgeKD-Net (Ours).

In summary, we evaluated EdgeKD-Net on MS-COCO 2017 and VisDrone 2023 against seven competing methods using three independent training runs for statistical reliability. On COCO, it reaches 38.6 ± 0.14 mAP@[0.50:0.95], 2.9 points above the MobileNetV3-SSD baseline and 0.9 points above the best prior distillation method PKD [9], at identical GFLOPs and slightly fewer parameters. On the NVIDIA Jetson Nano, it runs at 61 FPS (FP16) and 89 FPS (INT8) with an energy efficiency of 8.77 mAP/W, outperforming all compared methods on the accuracy-per-watt metric. The ablation indicates that all four components contribute independently, with response-level distillation adding the most (+1.1 mAP), followed by DFPF (+0.7), CAFA (+0.6), and DFL (+0.5). The INT8 accuracy drop of just 0.7 mAP is well within acceptable deployment limits, and results hold across both evaluated benchmarks, supporting generalization to the domains tested here; generalization beyond COCO and VisDrone has not been evaluated and is not claimed.

5. DISCUSSION

Taken together, the results from both benchmarks give us reasonable confidence in the core claim: anchor-assignment mismatch is a real and measurable source of training noise in knowledge distillation, and fixing it produces gains that stack on top of existing approaches. DFL complements them; it does not replace them. A 0.9 mAP improvement over PKD [9] at the same GFLOPs sounds modest in isolation, but PKD itself already incorporates several years of distillation research, and reducing the teacher–student gap from 2.3 mAP (PKD) to 1.4 mAP (EdgeKD-Net) is a 39% reduction. The 18.7% energy-efficiency gain over PKD (8.77 vs

7.39 mAP/W) may be more relevant than the accuracy gain for the battery-powered robotics and drone use cases that motivated this work.

The ablation deserves a bit more interpretation. DFL has the smallest individual contribution (+0.5 mAP) partly because it only operates on a filtered anchor subset, so it has a smaller ‘surface area’ of supervision to improve compared to response-level distillation, which operates across the full output distribution. The near-additive behavior of all four components (+1.1 + 0.6 + 0.5 + 0.7 $\approx$ +2.9 observed) is the more meaningful finding; it means each component addresses a distinct failure mode and doesn't just duplicate the others.

Regarding limitations, small-object detection is the most obvious gap: AP_S for EdgeKD-Net is about 4.2 points below the YOLOv8-L teacher, versus only 1.1 points for AP_L. This is expected, since small objects depend most on fine-grained spatial resolution, which degrades progressively through downsampling in both the backbone and the feature pyramid. DPFP's bottom-up path was designed to partially counteract this, and the VisDrone results suggest it helps. But P3 at 28×28 spatial resolution is still coarser than the teacher's internal feature representations. The attention-based DPFP enhancements at P3 that we list in the future work section directly address this observed gap, not just a speculative extension.

On whether DPFP is simply a recombination of PANet [16] and BiFPN [17]: all three use dual paths, yes, but the differences matter. PANet fuses by concatenation, then full convolution, adding $O(C^2)$ parameters per scale and not being designed with quantization in mind. BiFPN uses learned softmax-weighted fusion, whose division operation is a recognized source of rounding error under INT8 quantization. DPFP uses element-wise addition with no learned scalars and no normalization, which we adopted as a quantization-friendly design choice. We did not run a controlled INT8 comparison against PANet- or BiFPN-neck variants, and therefore make no claim of experimentally verified INT8 superiority; that comparison remains future work. On the bottom-up path, DPFP uses depthwise separable convolutions rather than PANet's full 3×3 convolutions, which is what allows it to match PANet's dual-path information flow at 1.8 M vs. 2.4 M parameters. The 256-channel output width is deliberately chosen to match the YOLOv8-L teacher neck, which eliminates a projection mismatch that would otherwise affect the CAFA distillation layers. This is an intentional consequence of a distillation-aware design decision.

This study has several limitations that provide opportunities for future work. First, evaluation covers MS-COCO 2017 and VisDrone 2023 only; extending to autonomous driving, indoor inspection, industrial, or thermal and low-light imagery would give a broader assessment, and our conclusions are scoped to the two domains tested. Second, we characterized deployment on Jetson Nano and Jetson Orin Nano, with Raspberry Pi 5 reported for reference; the Coral Edge TPU and Intel Movidius impose different quantization constraints than the TensorRT pipeline used here and would test broader versatility.

Third, the reported FPS values are averages. The backbone, neck, and head run as a static graph at batch size 1 and contribute little input-dependent variance; non-maximum suppression, whose cost scales with surviving candidate boxes, is the main data-dependent term, with frequency scaling and thermal throttling adding drift over longer periods. A per-frame profile reporting tail latency would therefore matter for applications with strict worst-case timing.

Finally, all experiments use an anchor-based MobileNetV3-SSD student. CAFA and DPFP are head-independent and transfer to anchor-free detectors unchanged, but DFL is not: Ω is defined over discrete anchor slots, and its motivating problem, that teacher and student may assign different anchors to the same object, presupposes the student has anchors to assign. An

anchor-free student would need Ω redefined over spatial cells rather than the existing loss re-run, and formulating this for YOLOv8-N/S students is the most direct route to establishing generality. Comparisons with certain recent distillation methods surveyed in Table 1 are likewise absent, as they either target anchor-free detectors or require backbone-specific projection layers incompatible with MobileNetV3-SSD without substantial architectural change. Head-to-head comparison under common conditions remains future work.

6. CONCLUSION AND FUTURE WORK

We presented EdgeKD-Net, a knowledge distillation framework for object detection on edge hardware. The three main components are DPFP, CAFA, and DFL, each addressing a different weakness of the compact student model. The central contribution is DFL: formally identifying the anchor-assignment mismatch problem, measuring it (22.9% of matched anchors), and correcting it by restricting distillation to the consistent anchor set Ω . This single fix contributes +0.7 mAP on its own. The complete system reaches 38.6 ± 0.14 mAP@[0.50:0.95] on MS-COCO 2017, 1.4 mAP short of the YOLOv8-L teacher, at 61 FPS (FP16) and 89 FPS (INT8) on NVIDIA Jetson Nano with 8.77 mAP/W efficiency. The improvement over PKD [9] is statistically significant ($p < 0.01$) and holds across both COCO and VisDrone. For future work, the priority is small-object detection, specifically attention-based DPFP enhancements at the P3 scale to improve AP_s. Another major direction is generalizing DFL to anchor-free models (CenterNet [18], FCOS [19], TOOD) by redefining the consistent set in terms of heatmap proximity or bipartite-matching cost. We also plan to test EdgeKD-Net on other embedded platforms beyond the Jetson series and in challenging real-world conditions, including nighttime and adverse weather.

ACKNOWLEDGMENT

The authors acknowledge all the help received from Sir Padampat Singhania university to carry out this research work. No grant or external funding was received for this research work. The code and trained models used in this study are available from the authors upon reasonable request.

REFERENCES

- [1] Jocher G, Chaurasia A, Qiu J (2023). Ultralytics YOLOv8. GitHub. Available: <https://github.com/ultralytics/ultralytics>
- [2] Lv Y, Xu Y, Zhao W, Wang G, Wei J, Cui Y. (2024) DETRs Beat YOLOs on Real-time Object Detection. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR); Seattle, USA, 16965–16974. doi: 10.1109/CVPR52733.2024.01605
- [3] Zhang H, Li F, Liu S, Zhang L, Su H, Zhu J, Ni LM, Shum H (2022). DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection. arXiv preprint arXiv:2203.03605. Available: <https://arxiv.org/abs/2203.03605>
- [4] Hinton G, Vinyals O, Dean J (2014) Distilling the Knowledge in a Neural Network. Proc. NeurIPS Workshop on Deep Learning; Montreal, Canada.
- [5] Romero A, Ballas N, Kahou SE, Chassang A, Gatta C, Bengio Y. (2015) FitNets: Hints for Thin Deep Nets. Proc. Int. Conf. Learn. Represent. (ICLR); San Diego, USA.
- [6] Zagoruyko S, Komodakis N. (2017) Paying More Attention to Attention: Improving the Performance of Convolutional Neural Networks via Attention Transfer. Proc. Int. Conf. Learn. Represent. (ICLR); Toulon, France.
- [7] Tian Y, Krishnan D, Isola P. (2020) Contrastive representation distillation. In Proc. Int. Conf. Learning Representations (ICLR); Addis Ababa, Ethiopia. <https://doi.org/10.48550/arXiv.1910.10699>

- [8] Park W, Kim D, Lu Y, Cho M. (2019) Relational knowledge distillation. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); Long Beach, USA. pp 3962-3971. <https://doi.org/10.1109/CVPR.2019.00409>
- [9] Passalis N, Tefas A. (2018) Learning deep representations with probabilistic knowledge transfer. In Proc. European Conf. Computer Vision (ECCV); Munich, Germany. Lecture Notes in Computer Science, vol 11215, pp 268-284. https://doi.org/10.1007/978-3-030-01252-6_17
- [10] Chen G, Choi W, Yu X, Han T, Chandraker M. (2017) Learning efficient object detection models with knowledge distillation. In Proc. 31st Int. Conf. Neural Information Processing Systems (NIPS); Long Beach, USA. pp 742-751.
- [11] Yang Z, Li Z, Jiang X, Gong Y, Yuan Z, Zhao D, Yuan C. (2022) Focal and global knowledge distillation for detectors. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); New Orleans, USA. pp 4643-4652. <https://doi.org/10.1109/CVPR52688.2022.00460>
- [12] Yang Z, Li Z, Shao M, Shi D, Yuan Z, Yuan C. (2022) Masked generative distillation. In Proc. European Conf. Computer Vision (ECCV); Tel Aviv, Israel. Lecture Notes in Computer Science, vol 13671, pp 53-69. https://doi.org/10.1007/978-3-031-20083-0_4
- [13] Guo J, Han K, Wang Y, Wu H, Chen X, Xu C, Xu C. (2021) Distilling object detectors via decoupled features. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); Nashville, USA. pp 2154-2164. <https://doi.org/10.1109/CVPR46437.2021.00219>
- [14] Howard A, Sandler M, Chu G, Chen LC, Chen B, Tan M, Wang W, Zhu Y, Pang R, Vasudevan V, Le QV, Adam H. (2019) Searching for MobileNetV3. In Proc. IEEE Int. Conf. Computer Vision (ICCV); Seoul, Korea. pp 1314-1324. <https://doi.org/10.1109/ICCV.2019.00140>
- [15] Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. (2017) Feature pyramid networks for object detection. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); Honolulu, USA. pp 936-944. <https://doi.org/10.1109/CVPR.2017.106>
- [16] Liu S, Qi L, Qin H, Shi J, Jia J. (2018) Path aggregation network for instance segmentation. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); Salt Lake City, USA. <https://doi.org/10.1109/CVPR.2018.00913>
- [17] Tan M, Pang R, Le QV. (2020) EfficientDet: scalable and efficient object detection. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); Seattle, USA. pp 10778-10787. <https://doi.org/10.1109/CVPR42600.2020.01079>
- [18] Zhou X, Wang D, Krähenbühl P. (2019) Objects as points. arXiv:1904.07850. <https://doi.org/10.48550/arXiv.1904.07850>
- [19] Tian Z, Shen C, Chen H, He T. (2019) FCOS: fully convolutional one-stage object detection. In Proc. IEEE Int. Conf. Computer Vision (ICCV); Seoul, Korea. pp 9626-9635. <https://doi.org/10.1109/ICCV.2019.00972>
- [20] Gevorgyan Z. (2022) SIOU loss: more powerful learning for bounding box regression. arXiv:2205.12740. <https://doi.org/10.48550/arXiv.2205.12740>
- [21] Xu S, Wang L, Wen G (2023). LiteFPN: A Lightweight Feature Pyramid Network for Real-Time Object Detection on Edge Devices. IEEE Trans. Circuits Syst. Video Technol., 33(8), 4021–4033.
- [22] Han K, Wang Y, Tian Q, Guo J, Xu C, Xu C. (2020) GhostNet: more features from cheap operations. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR); Seattle, USA. pp 1577-1586. <https://doi.org/10.1109/CVPR42600.2020.00165>
- [23] Dong Y, Gao P, Liu C, Zhao X. (2023) SlimFPN: Slimming Feature Pyramid Networks via Structured Pruning for Efficient Object Detection. Proc. Int. Joint Conf. Neural Networks (IJCNN); Gold Coast, Australia.
- [24] Zheng G, Ding X, Huang X, Feng J, Wei Y, Zhang Y, Liu J, Yuille A, Kong T. (2022) Localisation Distillation for Dense Object Detection. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR); New Orleans, USA, 9407–9416.

- [25] Zhu Y, Zhou Q, Liu N, Xu Z, Ou Z, Mou X, Tang J. (2023) ScaleKD: Distilling Scale-Aware Knowledge in Small Object Detector. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR); Vancouver, Canada, 19723–19733.
- [26] Jang Y, Shin W, Kim J, Woo S, Bae SH. (2022) GLAMD: global and local attention mask distillation for object detectors. In Proc. European Conf. Computer Vision (ECCV); Tel Aviv, Israel. Lecture Notes in Computer Science, vol 13670, pp 460-476. https://doi.org/10.1007/978-3-031-20080-9_27
- [27] Li Y, Li X, Li W, Hou Q, Liu L, Cheng MM, Yang J. (2024) SARDet-100K: towards open-source benchmark and toolkit for large-scale SAR object detection. arXiv:2403.06534. <https://doi.org/10.48550/arXiv.2403.06534>
- [28] Cai H, Li J, Hu M, Gan C, Han S. (2023) EfficientViT: lightweight multi-scale attention for high-resolution dense prediction. In Proc. IEEE Int. Conf. Computer Vision (ICCV); Paris, France. pp 17256-17267. <https://doi.org/10.1109/ICCV51070.2023.01587>
- [29] Mehta S, Rastegari M. (2022) Separable self-attention for mobile vision transformers. arXiv:2206.02680. <https://doi.org/10.48550/arXiv.2206.02680>.
- [30] Wang CY, Yeh IH, Liao HYM. (2024) YOLOv9: learning what you want to learn using programmable gradient information. In Proc. European Conf. Computer Vision (ECCV); Milan, Italy. Lecture Notes in Computer Science, vol 15089. https://doi.org/10.1007/978-3-031-72751-1_1
- [31] Lyu C, Zhang W, Huang H, Zhou Y, Wang Y, Liu Y, Zhang S, Chen K. (2022) RTMDet: an empirical study of designing real-time object detectors. arXiv:2212.07784. <https://doi.org/10.48550/arXiv.2212.07784>.

Pipe-to-Soil Potential State Classification For Equatorial Gas Pipeline ICCP Systems Under Geomagnetically Induced Current Influence Using XGBoost Algorithm

MAIZURA MOHD SANI¹, AHMAD IHSAN MOHD YASSIN¹,
MOHAMAD HUZAIMY JUSOH^{1,3}, NURHANI KASUAN²,
NUR DALILA KHIRUL ASHAR², MUHAMMAD ASRAF HAIRUDDIN²

¹ Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor Malaysia

² Faculty of Electrical Engineering, Universiti Teknologi MARA, Pasir Gudang Campus, Masai, Johor, Malaysia

³ MAGDAS/CPMN Research Group, International Center for Space and Planetary Environmental Science, Kyushu University, Japan

*Corresponding author: maizura23@uitm.edu.my

(Received: 17 April 2026; Accepted: 15 July 2026; Published online: 12 September 2026)

ABSTRACT: Pipeline corrosion poses critical challenges for the oil and gas industry, resulting in economic losses and safety risks to humans, assets, and the environment. To mitigate corrosion, the Impressed Current Cathodic Protection (ICCP) system is employed, which typically maintains the pipe-to-soil potential (PSP) within the optimal protected range of -0.85 V to -1.2 V. However, geomagnetically induced currents (GICs) associated with geomagnetic storms can reduce the system's efficiency, causing the pipeline to be either underprotected or overprotected by the ICCP. In equatorial regions, this interference is quite challenging, as the Equatorial Electrojet (EEJ) can enhance daily daytime geomagnetic variability and may increase GIC-related interference in susceptible pipelines. This study proposes an Extreme Gradient Boosting (XGBoost) classification model to classify three gas pipeline ICCP protection states, i.e., protected, underprotected, or overprotected, at two equatorial pipeline locations, which is trained using ICCP and GIC data. The proposed model achieves robust classification performance, with accuracies of 94.22% at Location 1 and 86.50% at Location 2, and weighted F1-scores of 0.9432 and 0.8662, respectively. The model also demonstrates strong discrimination, with ROC-AUC values of 0.9922 and 0.9738 across sites.

ABSTRAK: Kakisan paip merupakan cabaran kritikal industri minyak dan gas. Ini mengakibatkan kerugian ekonomi serta risiko keselamatan kepada manusia, aset, dan alam sekitar. Bagi mengurangkan kakisan, sistem Perlindungan Katodik Arus Paksa (ICCP) digunakan, lazimnya mengekalkan potensi paip-ke-tanah (PSP) dalam julat perlindungan optimum antara -0.85V hingga -1.2V. Walau bagaimanapun, kecekapan sistem ini dipengaruhi oleh Arus Teraruh Geomagnet (GIC), yang berkait dengan ribut geomagnet dan boleh menyebabkan saluran paip mengalami perlindungan tidak mencukupi atau perlindungan berlebihan oleh ICCP. Di kawasan khatulistiwa, gangguan ini menjadi lebih mencabar kerana Elektrojet Khatulistiwa (EEJ) meningkatkan variasi geomagnet harian pada waktu siang dan berpotensi menambah gangguan berkaitan GIC pada saluran paip yang terdedah. Kajian ini mencadangkan model pengelasan berasaskan Extreme Gradient Boosting (XGBoost) bagi klasifikasi tiga status perlindungan ICCP saluran paip gas, iaitu dilindungi, kurang dilindungi, dan terlebih dilindungi, di dua lokasi saluran paip kawasan khatulistiwa menggunakan data ICCP dan GIC. Model yang dicadangkan mencapai prestasi pengelasan tinggi, dengan ketepatan sebanyak 94.22% di Lokasi 1 dan 86.50% di Lokasi 2, serta skor F1 berlawanan masing-masing 0.9432 dan 0.8662. Model ini juga menunjukkan keupayaan

diskriminasi kukuh, dengan nilai ROC-AUC sebanyak 0.9922 dan 0.9738 bagi kedua-dua lokasi. Model ini boleh diguna sebagai sistem sokongan keputusan bagi pengendali saluran paip, membolehkan strategi penyelenggaraan proaktif sistem perlindungan katodik di kawasan khatulistiwa.

KEYWORDS: *Corrosion, Impressed current cathodic protection (ICCP), Geomagnetically induced current (GIC), Equatorial gas pipeline, XGBoost classification*

1. INTRODUCTION

Underground steel pipelines are the primary infrastructure for transporting oil and gas over long distances and serve as a main energy supply system for industrial, commercial, and residential sectors worldwide. One major challenge for these pipelines is corrosion, which can lead to pipe leakage, economic losses, safety risks, and environmental damage. To control corrosion, pipeline coatings and Impressed Current Cathodic Protection (ICCP) systems are commonly used together. Mimicking the electrochemical process, the ICCP system aims to ensure the pipeline acts as a cathode, reducing anodic oxidation reactions that cause corrosion [1]. As shown in Fig. 1, the ICCP system consists of anode beds of inert materials installed near the pipeline. Direct current (DC) supplied by a Transformer Rectifier Unit (TRU) is discharged through the anode bed into the surrounding soil, providing protective current to the pipeline and cathodically polarising the steel surface, thereby suppressing anodic metal dissolution and mitigating external corrosion. Pipeline protection is typically adequate when the pipe-to-soil potential (PSP), a monitoring ICCP parameter, is maintained between -850 mV and -1200 mV [2]. However, if the PSP is excessively negative, the ICCP overprotects the pipeline, which can lead to detrimental effects such as hydrogen embrittlement and damage to the pipeline coating [3].

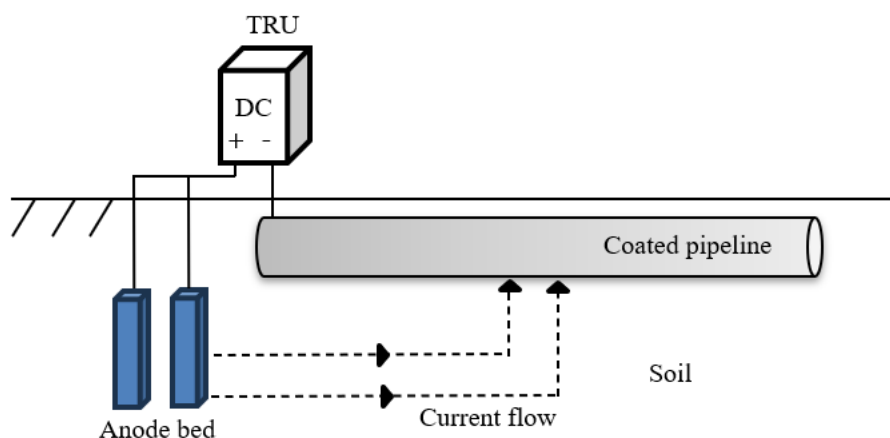


Figure 1. Schematic of the impressed current cathodic protection (ICCP) system for a buried pipeline

One factor that leads to overprotection and underprotection of the ICCP systems is the presence of stray telluric currents, or geomagnetically induced currents (GICs), which can be observed as fluctuations in PSP. Originating from space weather activity, the GIC hazard to the ICCP system has historically been associated with high-latitude regions due to the influence of the auroral electrojet near the pole [4]. However, the risk is now recognized as a global concern.. Pipelines in mid- and low-latitude regions have experienced out-of-safe-range PSP conditions during geomagnetic storm events [5] [6]. In Equatorial regions, pipelines experience more dynamic GIC interference. Equatorial pipelines experience interruptions from

geomagnetic storms and interference from the Equatorial Electrojet (EEJ), a regular diurnal phenomenon caused by daily variations in Earth's magnetic field during solar-quiet (Sq) time. [7]. This indicates that pipelines in this region are subjected to slowly varying GICs every day. Although the magnitude of this daily interference may be less extreme than during a major storm, its constant and prolonged nature challenges the ICCP system's ability to operate efficiently. This is because the EEJ continuously fluctuates the PSP, affecting the ICCP system and potentially accelerating long-term corrosion at coating defects [7].

Machine learning has been increasingly applied to pipeline integrity monitoring and corrosion protection management, with early efforts focusing on physics-based and statistical models before transitioning to data-driven approaches. Physics-based modeling approaches, such as the Distributed-Source Transmission Line Theory (DSTL) model and equivalent circuit models [8], have been adopted to identify pipeline locations affected by GIC and have been applied in continuous monitoring systems, such as the nowcasting service in Finland [9]. While DSTL models simulate GIC propagation along pipeline networks, their computational complexity and reliance on geological conductivity profiles pose challenges for extensive pipeline systems. As an alternative, regression-based approaches using a Multilayer Perceptron (MLP) and a Nonlinear Autoregressive with Exogenous Inputs (NARX) model have been demonstrated as a reliable method for predicting PSP values in equatorial pipeline systems under GIC influence [10], showing that geomagnetic parameters such as dH/dt and the EEJ index carry significant predictive signal for PSP behavior. However, regression output provides continuous potential estimates rather than direct categorical protection status, which limits its utility as an actionable decision-support tool for pipeline operators. Rossouw and Doorsamy [11] presented one of the first attempts to classify PSP protection states, specifically the three operational categories of protected, underprotected, and overprotected for a pipeline near the South African region using historical ICCP system data. While their framework demonstrated the feasibility of data-driven PSP classification, the model's performance reportedly degraded significantly in the presence of stray current interference because the feature set was limited to ICCP system parameters and did not incorporate stray currents.

Despite this progress, a gap persists in the literature. No prior study has classified PSP protection states using a data-driven approach that explicitly incorporates GIC-related geomagnetic features as model inputs. This gap is particularly significant in the equatorial context, where the EEJ introduces a persistent, diurnal source of GIC-related PSP interference that operates independently of geomagnetic storm activity. Studies on GIC effects in mid- and low-latitude pipeline infrastructure have been reported [5][6]. However, these works documented telluric current interference without a machine learning framework for automated protection-state classification. Furthermore, no existing classification framework has been validated specifically for equatorial pipeline systems where the combined influence of EEJ-driven quiet-time variability and storm-time GIC disturbances creates a uniquely challenging operational environment for ICCP systems.

This study addresses the identified gap by proposing the first GIC-PSP classification framework for equatorial gas pipelines, implemented using the Extreme Gradient Boosting (XGBoost) algorithm [12], which is efficient at handling missing values and imbalanced data, both of which are characteristics of this study's ICCP pipeline data [13]. The framework incorporates dH/dt and the EEJ index (EUEL) as geomagnetic feature inputs alongside ICCP system data, addressing the limitation of ICCP-only approaches that resulted in reduced accuracy under stray-current conditions [11]. The following section details the methodology, including XGBoost model development and hyperparameter tuning, followed by the

experimental results, performance evaluation, discussion, conclusions, and future research directions.

2. METHODOLOGY

To classify the PSP, this work proposes a four-stage methodology, as illustrated in Fig. 2. First is the data collection stage, which involves acquiring raw time-series sensor data from the pipeline's ICCP system, the PSP, and the TRU output current (I_{out}). Other data include the telluric current, or GIC, with indices of dH/dt , and the EEJ index (EUEL). The next stage is data preprocessing, which prepares the raw data by imputing missing values, resampling to ensure consistency, and extracting temporal features with specific time lags. The next stage is model training and optimization, where the data is split into training and test sets using a walk-forward validation scheme. The XGBoost model is then trained, and its hyperparameters are optimized using random search to find the best parameters. Finally, the performance metrics evaluate the model on the test data using accuracy, F1-score, the Receiver Operating Characteristic (ROC) curve, and the confusion matrix to confirm its reliability and effectiveness.

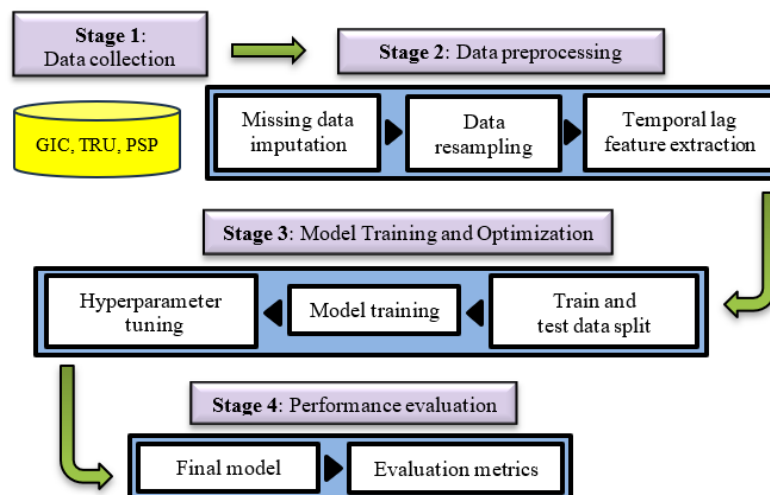


Figure 2. Block diagram of the ICCP classification system status using the XGBoost model

2.1. Data collection

The ICCP system PSP measurements were captured using the Close Interval Potential Survey (CIPS) [14] method. The pipeline's true potential, or off-potential, is measured over a 10-day period, from October 2, 2022, to October 12, 2022, using a data acquisition unit (DAQ) on two test posts at each location. These DAQs consist of voltage and current sensors integrated with a microcontroller, a real-time clock (RTC), and an SD memory card to record PSP and TRU output current. Both datasets have a sampling interval of 1 millisecond. Figure 3 shows the ICCP system and the CIPS measurement implemented in this study.

For this study, the PSP measurements were categorized into a three-class variable, 'underprotected', 'protected', and 'overprotected', based on the voltage thresholds as tabulated in Table 1. For the underprotected class, the dataset comprised 5,527 samples (27.42%) at Location 1 and 7,937 samples (39.38%) at Location 2. For the protected class, the corresponding numbers were 12,175 samples (60.39%) and 9,334 samples (46.30%), respectively. For overprotected samples, 2,458 (12.19%) were observed at Location 1 and 2,889 (14.32%) at Location 2. Figure 4 shows the plot of the normalized PSP values for this

study. This distribution shows the operational behavior of the ICCP system, where the protected status corresponds to the dominant operational state during normal cathodic protection performance [11].

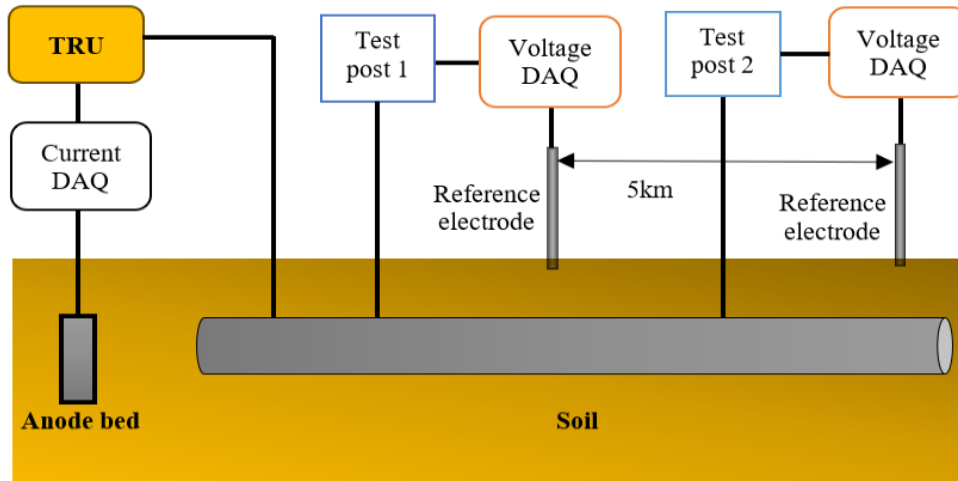


Figure 3. Data collection for PSP, TRU Iout at two pipeline locations at test posts 1 and 2

Table 1. PSP threshold of protected, underprotected and overprotected ICCP status

Characteristic	PSP value (V)
Protected	-0.85 to -1.2
Underprotected	Greater than -0.85
Overprotected	Less than -1.2

GICs originate from space weather activity, such as geomagnetic storms from solar disturbances, that perturb the Earth's magnetosphere and cause variations in the Earth's magnetic field [15]. Based on Faraday's law, these rapid temporal variations in the Earth's magnetic field give rise to geoelectric fields at the Earth's surface, which then drive GICs. In this study, the time derivative of the horizontal Earth's magnetic field, i.e., dH/dt represents a proxy for GIC activity, specifically in the low-latitude region [16]. According to established space weather studies, dH/dt is strongly correlated with the intensity of induced electric fields and has been adopted as a severity level for GICs in ground-based technological systems [17]. To compute dH/dt , the H-component was obtained from the magnetic observatory station at Universiti Teknologi MARA (UiTM), Pasir Gudang, and then numerically differentiated. In equatorial regions, the EEJ is a concentrated eastward electric current flowing in the ionospheric E-layer along the magnetic dip equator. We used the EEJ index (EUEL) from Kyushu University, Japan, to quantify EEJ strength [18]. Fig. 4 shows plots of PSP at Locations 1 and 2, TRU Iout, dH/dt , and EUEL. All variables are min-max normalized to the range [0, 1] to synchronize the data. Fluctuations in PSP at both locations indicate GIC influence on pipeline protection states.

2.2. Data Preprocessing

2.2.1. Missing data imputation

The PSP and TRU datasets contained occasional short-term gaps due to sensor interruptions, with missing values ranging between 7.5% and 18.6%. We imputed these gaps using linear interpolation, which assumes a straight-line relationship between adjacent

observations and preserves signal continuity. Given two consecutive observations, x_i and x_{i+1} a missing value, $\hat{x}(t)$ at time t , was estimated using Eq. (1):

$$\hat{x}(t) = x_i(t) + \frac{(t-t_i)}{(t_{i+1}-t_i)}(x_{i+1} - x_i) \quad (1)$$

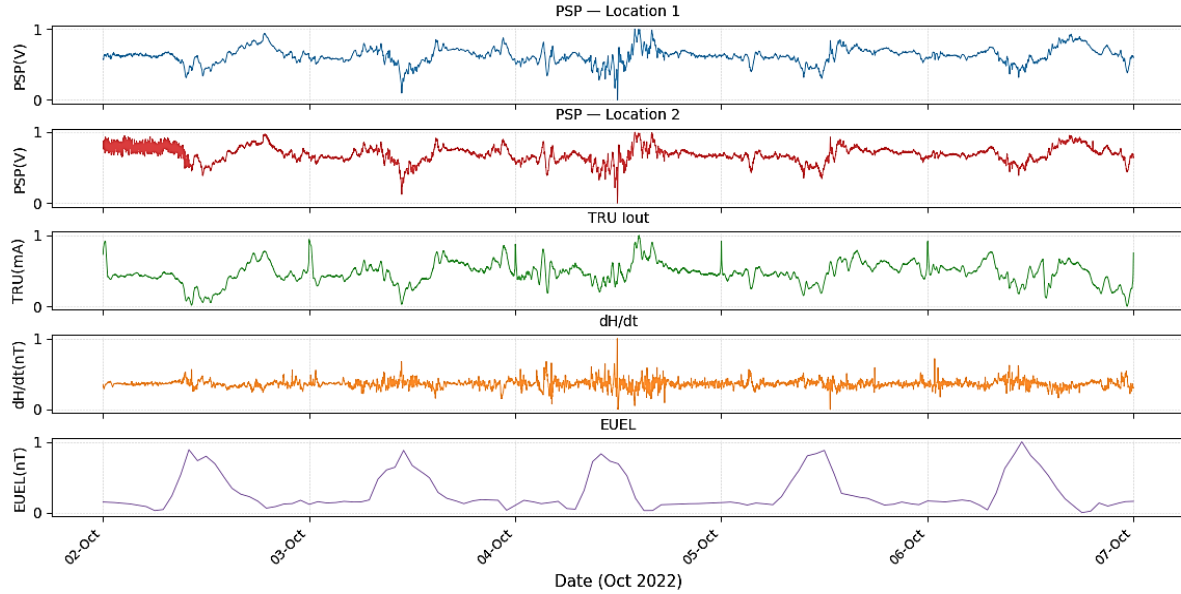


Figure 4. Normalized PSP for Locations 1 and 2, TRU lout, dH/dt, and EUEL from October 2, 2022, to October 7, 2022

2.2.2. Data resampling

The PSP, TRU output current, and dH/dt signals were originally sampled at millisecond resolution. To harmonize the dataset sampling rate, all variables were resampled to a one-minute resolution (1/60 Hz) to synchronize with the EUEL data. Down-sampling was performed using a two-step procedure. First, a digital low-pass filter was applied to attenuate frequency components above the new Nyquist limit, followed by decimation by a factor of 60. The filtering stage employed an anti-aliasing FIR filter (MATLAB resample with the default Kaiser window design) to keep the signals band-limited before decimation.

2.2.3. Temporal lag feature extraction

Temporal lag features were extracted to capture variations and historical patterns across all input data, which improve prediction performance. For EUEL, dH/dt, and TRU, lagged features were constructed at 1, 5, 10, and 30 minutes. The lag intervals were selected to capture data at different timescales for short-term transients (1-5 minutes), medium-term trends (10 minutes), and extended variations (30 minutes). The temporal lag set can be expressed as in Eq. (2):

$$X = [v(t - \tau)] \text{ for } v \in \{EUEL, \frac{dH}{dt}, TRU\} \text{ and } \tau \in \{0, 1, 5, 10, 30\} \quad (2)$$

where t represents the current time step of 1-minute resolution.

2.3. Model training and optimization

2.3.1. Train and test data split

The dataset was split into training and test sets, labeled as blue and red boxes, respectively, as shown in Figure 5. We split the data using a day-grouped walk-forward validation scheme

[19] to reduce overfitting in time-series data. In this figure, chronological order prevents information leakage from future samples into the training set, providing an independent validation set for performance assessment. Five folds were constructed so that the training window expanded progressively over time, and each test fold consisted of the subsequent day's observations.

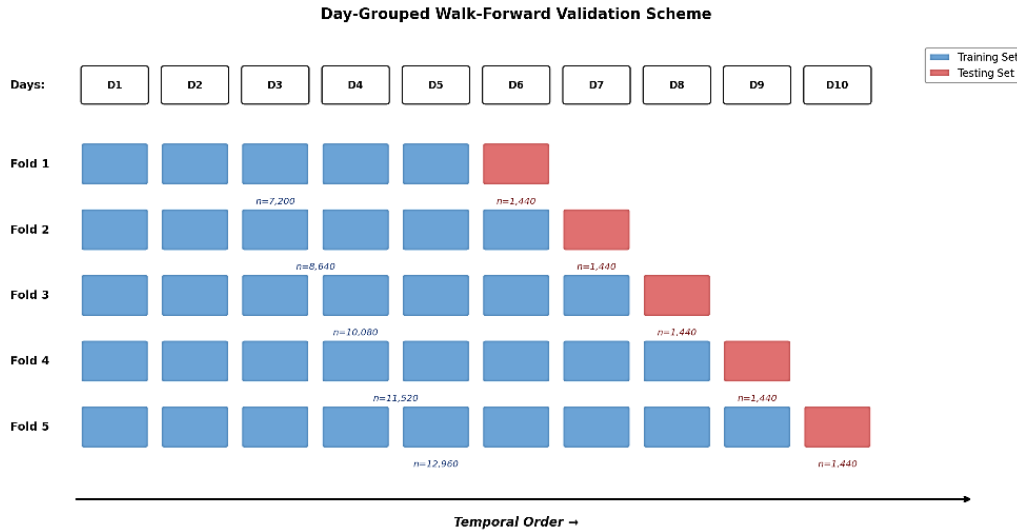


Figure 5. Day-grouped walk-forward validation

2.3.2. Model training

The XGBoost classifier was trained on the training partition using the selected hyperparameters. XGBoost learns an additive ensemble of decision trees by minimizing a regularised objective function, $\mathcal{L}(\phi)$ which balances data fit and model complexity. The overall objective function is defined as in Eq. (3):

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^t \Omega(f_k) \quad (3)$$

where $l(\hat{y}_i, y_i)$ denotes the loss between the predicted output and the true label, and $\Omega(f_k)$ is the regularisation term that penalizes the complexity of the tree f_k . At iteration t , XGBoost adds a new tree, f_t by minimising $\mathcal{L}^{(t)}$ as in Eq. (4):

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \sum_{k=1}^t \Omega(f_t) \quad (4)$$

The loss term, given by $\sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i))$, measures the discrepancy between the true label, y_i and the updated prediction obtained by adding the output of the new tree $f_t(x_i)$, to the prediction from the previous $t - 1$ trees, $\hat{y}_i^{(t-1)}$. The regularization term, $\Omega(f_t)$ penalizes the complexity of the newly added tree at the iteration t to control model complexity and reduce overfitting.

Multiclass log-loss was selected as the evaluation metric because it penalizes overconfident incorrect predictions more severely than uncertain incorrect predictions. The histogram-based tree construction algorithm was implemented to reduce computational complexity by discretizing continuous features, enabling efficient training in large feature spaces while maintaining model accuracy.

2.3.3. Hyperparameter tuning

Hyperparameter tuning was implemented using a random search over 100 iterations, with each candidate configuration evaluated using five-fold time-series cross-validation and scored

on weighted F1. The hyperparameter search space comprised the following values; maximum depth of 4, 6, 8, 10 and 12, minimum child weight with the value of 1, 3, 5, and 7, subsample ratio; 0.7, 0.8, 0.9, and 1.0, column sampling ratio; 0.7, 0.8, 0.9, and 1.0, gamma; 0.0, 0.1, 0.2, 0.3, L1 regularisation; 0.0, 0.001, 0.01, 0.1, L2 regularisation; 0.1, 0.5, 1.0, 5.0, and 10.0, learning rate with 0.01, 0.02, 0.03, and 0.05, and number of estimators of 1500 and 2000. Early stopping with a patience of 50 rounds was applied during each fold to prevent overfitting. The combination with the highest mean weighted F1 across folds was then used to train the final model.

2.4. Performance evaluation

Classification performance was evaluated using standard metrics, i.e., accuracy, macro and weighted F1-scores, and Receiver Operating Characteristic (ROC) analysis [20]. The weighted F1-score served as the main metric, weighted by each class's number of true instances, making it suitable for imbalanced multiclass data. ROC analysis using a one-vs-rest approach was conducted, and the Area Under the ROC Curve (AUC-ROC) was calculated for each class as a threshold-independent measure of performance.

3. RESULTS AND DISCUSSION

3.1. Model performance for 5-fold cross-validation

Figure 6 shows the results for accuracy, weighted F1-score, macro F1-score, and ROC-AUC from 5-fold cross-validation for both locations using a day-grouped walk-forward validation scheme. Location 1 outperformed Location 2 across all metrics, with higher mean values observed in each subplot. Performance at Location 1 was generally more consistent across folds, with smaller fluctuations. In contrast, Location 2 showed greater variability, particularly in fold 2, with lower values observed across all metrics. Despite this variation, ROC-AUC remained high across all folds at both locations, suggesting that the classifier maintained good class separability even when classification scores were lower. The fifth fold achieved the highest performance across all metrics at both locations, indicating that more training data progressively improves model generalization.

3.2. Performance per class

The one-vs-rest ROC curves are depicted in Fig. 7(left) for Location 1 and Fig. 7(right) for Location 2. Location 1 achieved high performance, with an AUC ranging from 0.986 to 0.997, whereas Location 2 achieved an AUC ranging from 0.940 to 0.995. The over-protected class consistently achieved the highest AUC at both locations (0.997 and 0.995 for Locations 1 and 2, respectively). Conversely, for both locations, the protected class demonstrated slightly lower discrimination performance. Despite these minor variations, the sustained high ROC-AUC performance across both locations validates the model's robustness for ICCP pipeline protection monitoring applications.

Fig.8 shows the confusion matrices for both monitoring sites. At Location 1, the model correctly identifies 94.57% (209/221) of under-protected samples, 95.25% (702/737) of protected samples, and 90.79% (138/152) of over-protected samples. At Location 2, the model correctly identifies 92.76% (346/373) of under-protected samples, 79.61% (449/564) of protected samples, and 92.65% (189/204) of over-protected samples. The protected class at Location 2 has the highest number of misclassifications, with 105 samples classified as under-protected and 10 as over-protected. This higher misclassification rate is attributed to the narrower margin between the protected (46.30%) and under-protected (39.38%) class

distributions at Location 2 compared to Location 1 (60.39% vs 27.42%), which reduces inter-class separability during training.

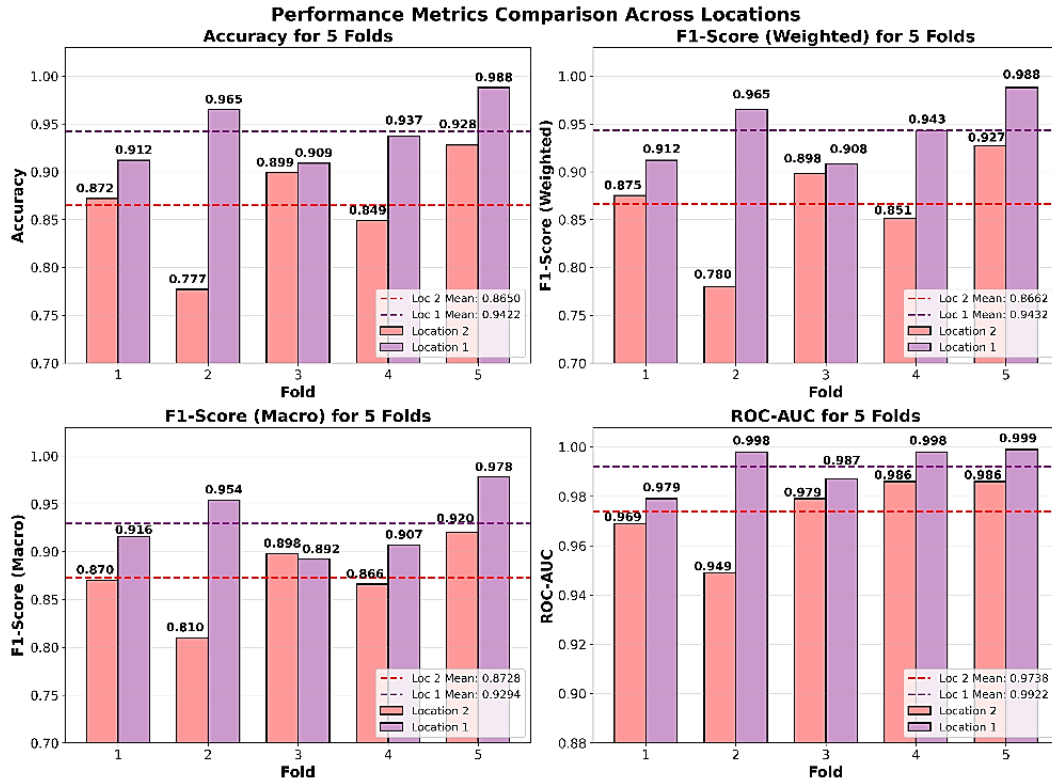


Figure 6. Model performance for five-fold cross-validation

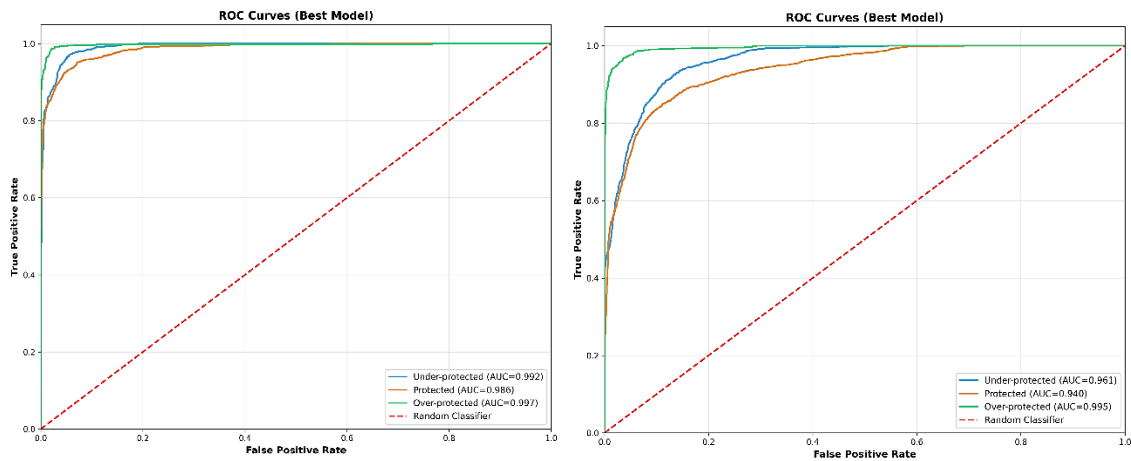


Figure 7. One-vs-rest ROC curves for three-class pipeline protection status classification for Location 1(left) and Location 2(right)

No under-protected sample is misclassified as over-protected, or vice versa, at both locations. This is attributed to the distinct separation of their PSP voltage ranges, with the entire protected zone (-0.85 V to -1.2 V) residing between them. As the protected samples lie in the intermediate range of the PSP spectrum, this class is more susceptible to boundary effects when readings fluctuate near either threshold, where PSP values near -0.85 V and -1.2 V introduce ambiguity in class assignment between adjacent classes.

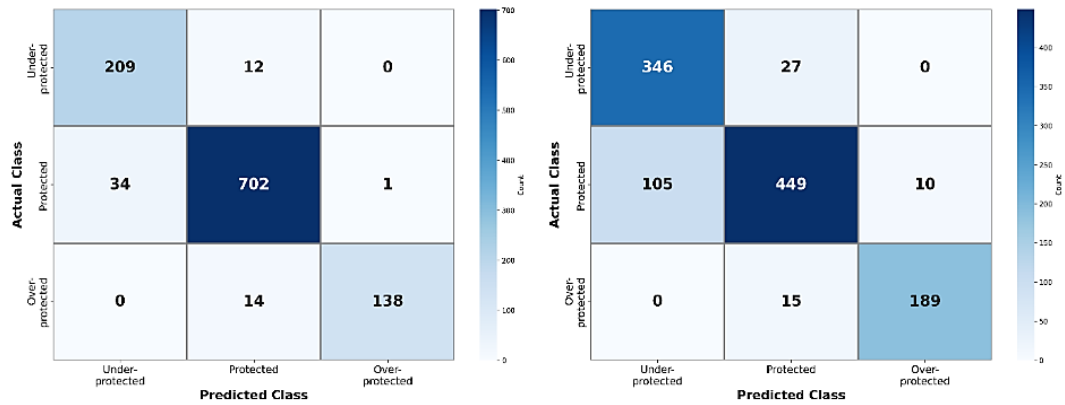


Figure 8. Confusion matrix of the ICCP system status classification (under-protected, protected, and over-protected) for Location 1(left) and Location 2(right)

3.3. Feature importance analysis

Fig. 9 presents the Shapley Additive explanations (SHAP) [21] global feature importance ranking for Locations 1 and 2. At both sites, the current (un-lagged, $\tau=0$) values of EUCL and TRU dominate the ranking, occupying the top two positions: EUCL_lag0m ranks first at Location 1, while TRU_lag0m ranks first at Location 2, with the other swapping to second place. This indicates that the instantaneous EEJ intensity and TRU output current carry the strongest direct relationship with the current PSP protection state, consistent with their role as the most immediate drivers of ICCP system behavior.

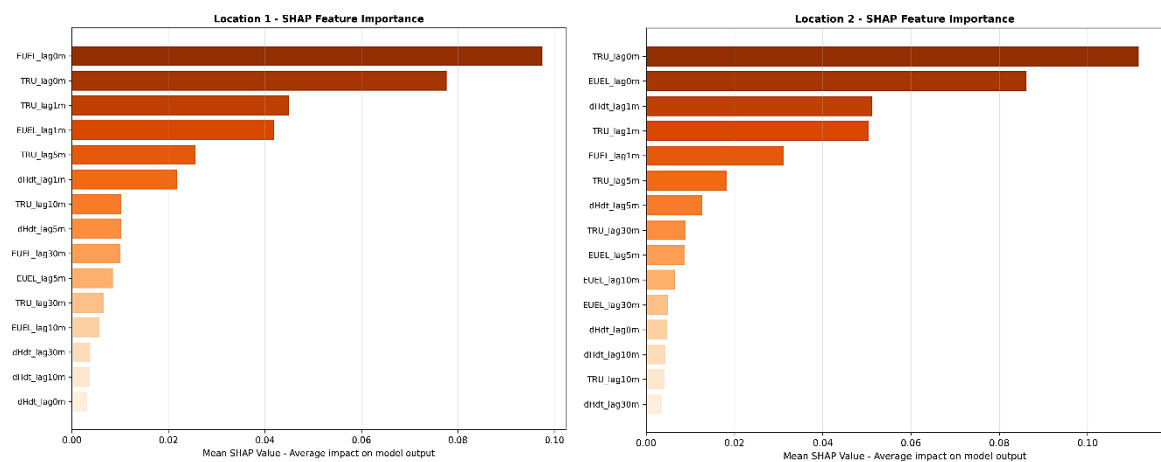


Figure 9. Global SHAP feature ranking for Location 1 and Location 2

A consistent pattern across both locations concerns dH/dt. Its current value (dHdt_lag0m) ranks lowest among all fifteen features at both sites, while its 1-minute lag (dHdt_lag1m) ranks substantially higher, sixth at Location 1 and third at Location 2. This contrast suggests that the instantaneous dH/dt reading carries comparatively little direct classification value, whereas the value recorded one minute prior is considerably more informative. The consistently lower SHAP importance of the instantaneous dH/dt value (dHdt_lag0m) relative to its one-minute lag (dHdt_lag1m) at both locations suggests a finite temporal delay between geomagnetic field variation and its manifestation in the PSP signal. This may be because the magnetic observatory at UiTM Pasir Gudang is not co-located with the pipeline monitoring sites, meaning the recorded dH/dt is regional rather than a direct local measurement of the geoelectric field at the pipeline [22]. This spatial offset may introduce a temporal lagged discrepancy between the observed dH/dt and the resulting PSP response. Beyond a 1-minute lag, the reduced

importance of 10-minute and 30-minute lagged variables has been observed, which indicates that the system response is mainly affected by short-term dynamics within the first few minutes.

3.4. Contribution of GIC features on model performance

Table 2 shows XGBoost model performance using three input feature sets: TRU only, TRU with dH/dt, or TRU with dH/dt and EUEL. Using TRU current output alone as input features has exhibited limited classification performance at both monitoring locations. At Location 1, the accuracy, F1-macro, F1-weighted, and ROC-AUC were 0.6769, 0.6092, 0.6664, and 0.8572, respectively, while the corresponding values at Location 2 were 0.5360, 0.5451, 0.5207, and 0.7716. Although TRU current reflects the impressed current supplied to the pipeline, it does not capture transient geomagnetic disturbances that influence the PSP. This modest performance indicates that TRU current data alone is insufficient for predicting PSP protection states under GIC-disturbed conditions.

Among the evaluated feature combinations, including dH/dt yields the greatest performance improvement, with consistent gains across all metrics. The F1-macro increased by 0.2031 and 0.2146 at Locations 1 and 2, respectively, indicating improved class balance in prediction performance. The ROC-AUC values also increased to 0.9422 and 0.9194, demonstrating enhanced class separability. This improvement is physically justified, as dH/dt represents rapid geomagnetic fluctuations associated with GIC-induced disturbances that directly affect PSP behavior.

Table 2. Performance of the XGBoost model based on different input combinations.

Loc	Features	Accuracy	F1-scores (macro)	F1-scores (weighted)	ROC-AUC
1	TRU	0.6769	0.6092	0.6664	0.8572
	TRU, dH/dt	0.8346	0.8123	0.8381	0.9422
	TRU, dH/dt, EUEL	0.9422	0.9294	0.9432	0.9922
2	TRU	0.5360	0.5451	0.5207	0.7716
	TRU, dH/dt	0.7493	0.7597	0.7509	0.9194
	TRU, dH/dt, EUEL	0.8650	0.8728	0.8662	0.9738

Adding EUEL to form the complete feature set produced further improvements across all evaluation metrics, although the magnitude of improvement was smaller than that associated with dH/dt. At Location 1, the model achieved an accuracy of 0.9422, F1-macro of 0.9294, F1-weighted of 0.9432, and ROC-AUC of 0.9922. At Location 2, the corresponding values were 0.8650, 0.8728, 0.8662, and 0.9738. The consistent improvement for all metrics indicates that EUEL provides important information. As an indicator of EEJ strength, EUEL captures geomagnetic variations that dH/dt does not, therefore improving characterization of GIC-induced PSP perturbations.

3.5. Comparison of different classification model performance

Table 3 presents a comparative evaluation of the XGBoost classifier, which was performed using five commonly deployed machine learning algorithms, i.e., Random Forest (RF), LightGBM, CatBoost, Multilayer Perceptron (MLP) Artificial Neural Network (ANN), and Support Vector Machine (SVM), evaluated under identical day-grouped walk-forward cross-validation conditions. All comparison classifiers were hyperparameter-optimized using random search over a single chronological train-validation split (75/25), with 100 search iterations for RF, LightGBM, CatBoost, and MLP, and 15 iterations for SVM due to its higher computational cost. The XGBoost model achieved the highest classification accuracy at both

monitoring locations, with mean accuracies of 94.22% and 86.50% at Locations 1 and 2, respectively, outperforming the next-best classifier (LightGBM) by approximately 6 percentage points at both sites. Among the comparison classifiers, LightGBM performed most competitively, achieving mean accuracies of 86.03% and 79.76% at Locations 1 and 2, respectively. RF and CatBoost performed moderately, while MLP-ANN showed greater variability. SVM performed worse than all methods at both locations, with mean accuracies of 49.22% and 45.42%, marginally above the theoretical random baseline for a three-class problem. XGBoost's consistent superiority across all metrics at both monitoring locations indicates that this algorithm is well suited to the GIC-influenced PSP classification task.

Table 3. Performance of the XGBoost model compared with other classification models.

Loc	Algorithm	Accuracy	F1-scores (macro)	F1-scores (weighted)	ROC- AUC
1	Random Forest	0.8390	0.8200	0.8405	0.9581
	Light GBM	0.8603	0.8443	0.8642	0.9613
	CatBoost	0.8306	0.8224	0.8377	0.9625
	MLP(ANN)	0.8147	0.7860	0.8182	0.9304
	SVM	0.4922	0.5374	0.4733	0.8699
	XGBoost	0.9422	0.9294	0.9432	0.9922
2	Random Forest	0.7312	0.8121	0.7312	0.9364
	Light GBM	0.7976	0.7546	0.7976	0.9509
	CatBoost	0.7143	0.7364	0.7143	0.9440
	MLP(ANN)	0.7835	0.7814	0.7835	0.9423
	SVM	0.4542	0.4390	0.4542	0.8512
	XGBoost	0.8650	0.8728	0.8662	0.9738

3.6. Computational Performance and Operational Considerations

Computational performance shows the speed at which this study's computer system can execute the XGBoost classification model, demonstrating efficiency suitable for near-real-time pipeline protection monitoring. All timing measurements were obtained using Python's high-resolution performance counter (`time.perf_counter()`). Timing measurements were recorded on a system equipped with an Intel Core i7-14700HX processor and 16 GB of RAM. Model training across five walk-forward cross-validation folds took approximately 11.12 seconds, with a mean training time of 2.22 seconds per fold. For the inference stage, in which the trained model classifies the instantaneous PSP protection state from incoming sensor features, it is computationally inexpensive, with a mean classification time of 0.032 milliseconds per sample, equivalent to approximately 31,000 classifications per second. This indicates that the model can process the input sensor data resolution within the available sampling window.

Table 4 compares mean training and inference times across all six classifiers evaluated in this study. All classifiers achieve inference times well below the sensor data's one-minute sampling resolution, indicating that real-time classification is computationally feasible for all methods evaluated. Among the tree-based ensemble methods, XGBoost and RF demonstrated comparable training times of 2.22 and 1.29 seconds per fold, respectively, while LightGBM was marginally slower at 2.41 seconds per fold. CatBoost exhibited the highest training time among tree-based methods at 19.19 seconds per fold. The MLP achieved the fastest inference time at 0.0014 ms per sample, followed closely by CatBoost at 0.0026 ms per sample. XGBoost's inference time of 0.032 ms per sample, while slightly higher than MLP and CatBoost, remains well within real-time operational requirements. SVM has the highest inference time at 0.221 ms per sample. XGBoost is therefore selected not for computational efficiency alone; rather, its high classification accuracy at both monitoring locations, combined

with acceptable inference latency and low training time, justifies its adoption as the primary classifier.

Table 4. Mean training time and mean inference time for different classifier models

Classifier	Mean Training Time	Mean Inference Time
MLP (ANN)	1.909	0.0014
CatBoost	19.194	0.0026
LightGBM	2.406	0.0099
XGBoost	2.224	0.0321
Random Forest	1.285	0.0432
SVM	8.963	0.2207

With low computational overhead, this model's classification output could map directly to actionable decisions within the pipeline operator's maintenance workflow. Pipeline engineers monitoring the ICCP system can use the PSP output state as a continuous decision-support signal alongside conventional PSP readings. A protected classification indicates that the system is operating within the acceptable potential range of -0.85 V to -1.2 V and requires no immediate intervention. An underprotected classification alerts the engineer that the pipeline is at risk of accelerated corrosion and prompts actions such as increasing the TRU output current, inspecting for coating defects, or verifying the integrity of ICCP system components. An overprotected classification indicates excessive cathodic current, which may lead to hydrogen embrittlement or coating disbondment, and requires reducing TRU output current. In persistent or severe cases, particularly where recurrent underprotection is detected despite TRU adjustment, the classification output may serve as an early indicator warranting physical excavation and direct pipeline inspection. Integrating this classification framework into a complete real-time deployment architecture, including communication protocols and SCADA system interfacing, represents a direction for future work.

3.7. Limitations and future work

The present study spans approximately ten days under geomagnetically quiet conditions and a moderate storm event from 3 to 5 Oct 2022 (minimum EDst index value of -59 nT) [23]. As a result, the training data largely captures EEJ-driven diurnal variability but does not encompass strong or severe GIC behavior or seasonal variations in EEJ intensity. Because the observation window contains no major geomagnetic storm events, the model's performance under stronger geomagnetic forcing has not been evaluated. These constraints reflect a broader challenge in equatorial GIC research, where observational data from equatorial and low-latitude regions remain limited [24]. A similar challenge has been reported in machine-learning-based space weather prediction, where data scarcity and class imbalance in rare-event datasets have led researchers to employ strategies such as transfer learning, data augmentation, and resampling methods to improve model robustness and generalization [25]. In pipeline corrosion monitoring, collecting continuous, high-resolution field data is also constrained by cost, sensor access, and operational considerations, as interrupting the ICCP system to obtain PSP data is not feasible. Future work should prioritize acquiring longer-term, multi-season datasets covering varying levels of geomagnetic activity, including storm-time events, at different geographical sites to improve model robustness and validate generalization across a broader range of ICCP operational conditions. Additionally, data augmentation strategies such as conditional generative adversarial networks (cGAN) or synthetic minority oversampling may help address the inter-class separability and temporal coverage challenges identified in this study.

4. CONCLUSION

This study developed an XGBoost classification model incorporating temporal lag features derived from GIC-influenced geomagnetic field variations to address the limited availability of data-driven frameworks for classifying PSP protection states in an ICCP system for an equatorial-region gas pipeline under GIC disturbances. The proposed model achieved classification accuracies of 94.22% and 86.50% at Locations 1 and 2, respectively, with ROC-AUC values of 0.9922 and 0.9738, indicating strong discriminative capability across all three protection states at both sites. A comparative evaluation of five classifiers, i.e., Random Forest, LightGBM, CatBoost, MLP, and SVM, showed consistent superiority of XGBoost at both monitoring locations. XGBoost's consistent superiority across all metrics and both locations, combined with an inference time of 0.032 milliseconds per sample, indicates the feasibility of the proposed framework for real-time operational deployment within the sensor time-sampling window. Among the input features, dH/dt contributed the largest performance increment as the most informative variable for capturing GIC-induced PSP disturbances, while EUCL provided additional discriminative information that further refined the classification boundaries. The present study has certain limitations. The ten-day observation window predominantly captures EEJ-driven diurnal variability without encompassing strong or severe storm-time GIC events (which depend on space weather geomagnetic storm events) or seasonal variations in EEJ intensity. Furthermore, the dH/dt and EUCL features should be specific to locally acquired geomagnetic data near the pipeline. Future work should prioritize acquiring longer-term, multi-season datasets spanning varied geomagnetic activity levels, exploring data augmentation strategies to address inter-class separability challenges, and cross-regional validation of the proposed methodology for equatorial pipeline systems. The proposed approach highlights its potential as a reliable support tool for pipeline engineers, particularly for ICCP protection state monitoring in equatorial pipeline systems under the influence of geomagnetic storms and the EEJ phenomenon.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support provided by Universiti Teknologi MARA and the Ministry of Higher Education Malaysia.

REFERENCES

- [1] Peabody AW. (2018) Peabody's Control of Pipeline Corrosion (3rd ed.). Houston, NACE International. <https://doi.org/10.5006/37617>
- [2] Gummow RA. (2018) Cathodic protection criteria-A critical review of NACE standard RP-01-69. *Mater. Perform.*, 57(1):28-35. https://doi.org/10.5006/MP2018_57_1-28
- [3] Hosokawa Y, Kajiya F, Nakamura Y. (2004) New cathodic protection criteria based on direct and alternating current densities measured using coupons and their application to modern steel pipelines. *Corrosion*, 60(3):304-312. doi:10.5006/1.3287735
- [4] Gummow RA, Eng P. (2002) GIC effects on pipeline corrosion and corrosion control systems. *J. Atmos. Sol. Terr. Phys.*, 64(16):1755-1764. doi:10.1016/S1364-6826(02)00125-6
- [5] Yu Z, Hao J, Liu L, Wang Z. (2019) Monitoring Experiment of Electromagnetic Interference Effects Caused by Geomagnetic Storms on Buried Pipelines in China. *IEEE Access*, 7:14603-14610. doi:10.1109/ACCESS.2019.2893963
- [6] Fernandes De Moraes J, Paulino I, Alves LR, Marcos Denardini C. (2020) Evaluation of possible corrosion enhancement due to telluric currents: Case study of the Bolivia-Brazil pipeline. *Ann. Geophys.*, 38(4):881-888. doi:10.5194/angeo-38-881-2020

-
- [7] Sani MM, Kasuan N, Jusoh MH, Mohd Yassin AI, Hussain Z, Hairuddin MA. (2023) Preliminary Study on the Impact of GIC on Pipe-to-Soil Potential of Buried Pipeline Near Equatorial Region. In Proceedings of the IEEE Symposium on Industrial Electronics and Applications: 15-16 July 2023; Malaysia; pp 1-6. doi:10.1109/ISIEA58478.2023.10212318
 - [8] Boteler DH. (2013) A new versatile method for modelling geomagnetic induction in pipelines. *Geophys. J. Int.*, 193(1):98-109. doi:10.1093/gji/ggs113
 - [9] Viljanen A, Pulkkinen A, Pirjola R, Pajunpää K, Posio P, Koistinen A. (2006) Recordings of geomagnetically induced currents and a nowcasting service of the Finnish natural gas pipeline system. *Space Weather*, 4(10). doi:10.1029/2006SW000234
 - [10] Sani MM, Yassin AIM, Kasuan N, Ashar NDK, Hairuddin MA, Jusoh MH. (2025) Prediction of the Pipe-to-Soil Potential Based on GIC for Buried Gas Pipeline Using Machine Learning Approach. *Journal of Physics: Conference Series*, 2998(1):012021. doi:10.1088/1742-6596/2998/1/012021
 - [11] Rossouw E, Doorsamy W. (2021) Predictive maintenance framework for cathodic protection systems using data analytics. *Energies*, 14(18):5805. doi:10.3390/en14185805
 - [12] Yang Y, Liu JJ, Feng XS, Chen PF, Zhang B. (2023) Prediction of the Transit Time of Coronal Mass Ejections with an Ensemble Machine-learning Method. *Astrophys. J. Suppl. Ser.*, 268(2):69. doi:10.3847/1538-4365/acf218
 - [13] Tanha J, Abdi Y, Samadi N, Razzaghi N, Asadpour M. (2020) Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*, 7(1):70. doi:10.1186/s40537-020-00349-y
 - [14] Dzakyprasetyo, Ferdian D. (2023) Analysis of external corrosion protection performance on buried gas pipeline using CIPS and DCVG methods. *ITM Web of Conferences*, 61:01022. doi:10.1051/itmconf/20246101022
 - [15] Patel S, Chamadia PK. (2025) Effect of Solar and Interplanetary Disturbances on Geomagnetic Field: A Comprehensive Review. *Journal of Advances in Science and Technology*, 22(2):583-596. doi:10.29070/426v7408
 - [16] Zainuddin A, Hairuddin MA, Yoshikawa A, Helmy M, Abd Latiff ZI, Yassin I, Jusoh H. (2024) Comparative Analysis of dB/dt Response Towards dH/dt as Proxies for Geomagnetically Induced Currents (GICs) in Low-Latitude Region. In Proceedings of the IEEE Symposium on Wireless Technology and Applications: 20-21 July 2024; pp 235-240. doi:10.1109/ISWTA62130.2024.10651826
 - [17] Viljanen A, Wintoft P, Wik M. (2015) Regional estimation of geomagnetically induced currents based on the local magnetic or electric field. *Journal of Space Weather and Space Climate*, 5:A24. doi:10.1051/swsc/2015022
 - [18] Fujimoto A, Uozumi T, Abe S, Matsushita H, Ishitsuka JK, Yoshikawa A. (2016) Long-term EEJ variations by using the improved EE-index. *Sun and Geosphere*, 11(1):37-47.
 - [19] Wahyuddin EP, Caraka RE, Kurniawan R, Caesarendra W, Gio PU, Pardamean B. (2025) Improved LSTM hyperparameters alongside sentiment walk-forward validation for time series prediction. *Journal of Open Innovation: Technology, Market, and Complexity*, 11(1). doi:10.1016/j.joitmc.2024.100458
 - [20] Chukwura Obi J. (2023) A comparative study of several classification metrics and their performances on data. *World Journal of Advanced Engineering Technology and Sciences*, 8(1):308-314. doi:10.30574/wjaets.2023.8.1.0054
 - [21] Sheng L, Ruan X, Wang Y. (2025) Interpretable XGBoost-SHAP machine learning model for identifying scientific breakthroughs. *Scientometrics*, 130:6801-6832. doi:10.1007/s11192-025-05497-7
 - [22] Ingham M, Divett T, Rodger CJ, Sigley M. (2025) Observed Effects of the May 2024 Gannon Storm on the New Zealand Gas Pipeline Network-Toward Predicting the Effects of an Extreme Storm. *Space Weather*, 23(5). doi:10.1029/2024SW004258
-

- [23] Uozumi T, Yumoto K, Kitamura K, et al. (2008) A new index to monitor temporal and long-term variations of the equatorial electrojet by MAGDAS/CPMN real-time data: EE-index. *Earth, Planets and Space*, 60(7):785-790. doi:10.1186/BF03352828
- [24] Caraballo R, González-Esparza JA, Pacheco CR, Corona-Romero P, Arzate-Flores JA, Castellanos-Velazco CI. (2025) The Impact of Geomagnetically Induced Currents (GIC) on the Mexican Power Grid: Numerical Modeling and Observations From the 10 May 2024, Geomagnetic Storm. *Geophys. Res. Lett.*, 52(4). doi:10.1029/2024GL112749
- [25] Thaker P, Desai A, Panchal P, Sutaria D, Shah M, Prajapati M. (2025) Evaluating AI approaches for space weather prediction: Strengthening satellite resilience and technology systems. *Space Habitation*, 1(3).doi:10.1016/j.spaceh.2025.100032

A CatBoost-Based Framework for Autism Spectrum Disorder Classification from EEG Signals with Blackman–Harris FIR Filtering

BULEUN WALADAH¹, MELINDA MELINDA^{1*}, YUNIDAR YUNIDAR¹,
AUFA RAFIKI¹, SITI RUSDIANA², WEI KITT WONG³

¹*Department of Electrical and Computer Engineering, Faculty of Engineering,
Universitas Syiah Kuala, Banda Aceh, Indonesia*

²*Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Syiah Kuala, Banda Aceh, Indonesia*

³*Department of Electrical and Computer Engineering, Faculty of Engineering and Science,
Curtin University Malaysia, Miri, Malaysia*

**Corresponding author: melinda@usk.ac.id*

(Received: 15 June 2026; Accepted: 24 August 2026; Published online: 11 September 2026)

ABSTRACT: Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder marked by impairments in social interaction, communication, and repetitive behavior. Early identification remains difficult given its clinical heterogeneity and the subjectivity of conventional diagnosis. This study proposes an EEG-based ASD classification framework combining Blackman–Harris FIR filtering, feature extraction and selection, and machine learning. EEG data were collected from 16 participants (8 ASD, 8 neurotypical) using a 16-channel system, band-pass filtered at 4–40 Hz, segmented into 4-second epochs with 50% overlap, and represented by 176 temporal and spectral features per epoch (11 per channel), followed by fold-specific Mutual Information feature selection. SVM-RBF, CatBoost, LDA, and Gaussian Naive Bayes were evaluated using nested subject-wise 8-fold cross-validation. Based on 7,145 pooled out-of-fold epoch predictions, CatBoost achieved the highest accuracy (99.58%), followed by LDA (97.47%), SVM-RBF (91.60%), and Gaussian Naive Bayes (57.49%). An exploratory paired t-test on fold-wise accuracy showed a nominally significant CatBoost advantage over SVM-RBF (mean difference 8.00 percentage points, $p = 0.042$; Cohen's $d_z = 0.88$). At the participant level, majority voting correctly classified 16/16 participants with CatBoost and LDA, 15/16 with SVM-RBF, and 9/16 with Gaussian Naive Bayes; the 95% CI for 100% accuracy was 79.41–100.00%. These findings support preliminary within-cohort feasibility but not clinical or screening readiness.

ABSTRAK: Gangguan Spektrum Autisme (ASD) adalah gangguan perkembangan neuro yang dicirikan oleh kekurangan interaksi sosial, komunikasi, dan tingkah laku berulang. Pengesanan awal masih sukar disebabkan kepelbagaian klinikal dan sifat subjektif diagnosis konvensional. Kajian ini mencadangkan rangka kerja pengelasan ASD berasaskan EEG yang menggabungkan penapisan Blackman–Harris FIR, pengekstrakan dan pemilihan ciri, serta pembelajaran mesin. Data EEG diperoleh daripada 16 peserta (8 ASD, 8 neurotipikal) menggunakan sistem 16 saluran, ditapis jalur 4–40 Hz, dibahagikan kepada epok 4 saat dengan pertindihan 50%, dan diwakili oleh 176 ciri temporal dan spektral bagi setiap epok (11 setiap saluran), diikuti pemilihan ciri Informasi Bersama khusus bagi setiap lipatan. SVM-RBF, CatBoost, LDA, dan Gaussian Naive Bayes dinilai menggunakan validasi silang 8-lipatan berasaskan subjek. Berdasarkan 7,145 ramalan epok terkumpul luar lipatan, CatBoost mencapai ketepatan tertinggi (99.58%), diikuti LDA (97.47%), SVM-RBF (91.60%), dan Gaussian Naive Bayes (57.49%). Kajian pada ketepatan setiap lipatan bagi ujian-t

berpasangan menunjukkan kelebihan nominal yang signifikan bagi CatBoost berbanding SVM-RBF (perbezaan min 8.00 mata peratus, $p = 0.042$; Cohen's $d_z = 0.88$). Pada peringkat peserta, pengelasan undian majoriti adalah betul iaitu 16/16 peserta menggunakan CatBoost dan LDA, 15/16 menggunakan SVM-RBF, dan 9/16 menggunakan Gaussian Naive Bayes; 95% CI bagi ketepatan 100% ialah 79.41–100.00%. Dapatan ini menyokong kebolehlaksanaan awal kohort ini, tetapi bukan kesediaan klinikal atau saringan.

KEYWORDS: ASD; Electroencephalography; Blackman–Harris FIR; SVM-RBF; CatBoost

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a heterogeneous neurodevelopmental disorder characterized by persistent impairments in social communication and restricted or repetitive behaviors [1], [2]. Neurophysiological studies have reported atypical functional connectivity and neuronal synchronization in individuals with ASD, including altered local and long-range interactions during resting conditions [3], [4]. Electroencephalography (EEG) provides a non-invasive means of examining these neural dynamics and has increasingly been used for ASD classification through conventional feature-based and more recent representation-learning approaches [5], [6], [7], [8].

EEG-based classification remains challenging because of signal nonstationarity, artifact sensitivity, and variability in acquisition protocols and participant characteristics [8], [9]. Preprocessing is therefore important for reducing out-of-band components while preserving information within the frequency range of interest. Window-based FIR filtering provides a finite, non-recursive response with controllable spectral characteristics, while the Blackman–Harris window is distinguished by strong sidelobe suppression that can reduce spectral leakage associated with finite-length signal processing [10], [11], [12]. In this study, a 4–40 Hz Blackman–Harris FIR filter was therefore evaluated against raw EEG and a conventional Butterworth preprocessing alternative [13].

Following preprocessing, temporal and spectral EEG features were classified using SVM-RBF and CatBoost as the primary models, with LDA and Gaussian Naive Bayes (GNB) included as additional baselines. SVM-RBF provides nonlinear decision boundaries through the radial basis function kernel [14], whereas CatBoost uses gradient-boosted decision trees to model complex nonlinear relationships [15], [16]. Recent studies have also demonstrated CatBoost's applicability to EEG and other biomedical classification problems [17], [18], supporting its evaluation as a complementary tree-based approach.

Previous EEG-based ASD studies have primarily emphasized feature representation and classifier development, whereas preprocessing choices have received comparatively less systematic evaluation [6], [8]. Moreover, direct comparison of nonlinear kernel-based and gradient-boosting pipelines under participant-disjoint validation remains limited. Accordingly, this study evaluates Blackman–Harris FIR preprocessing by comparing it with raw EEG and Butterworth filtering, and compares SVM-RBF, CatBoost, LDA, and GNB using nested subject-wise 8-fold cross-validation. Performance is assessed at both epoch and participant levels using classification metrics, confidence intervals, aggregation sensitivity analysis, and statistical testing to determine the within-cohort feasibility and stability of the evaluated EEG-based ASD classification pipelines.

2. MATERIALS AND METHODS

This study developed a machine-learning framework distinguishing individuals with ASD from typically developing (TD) controls using multichannel EEG. The workflow comprised participant recruitment, EEG acquisition, electrode configuration, preprocessing, epoch segmentation, feature extraction and selection, classification, and statistical evaluation. EEG signals were acquired via a 16-channel OpenBCI Cyton system, preprocessed with a Blackman–Harris FIR band-pass filter, then used to extract temporal and spectral features for classification, as illustrated in Fig. 1.



Figure 1. Workflow of the proposed EEG classification system.

2.1. EEG Dataset and Participants

The EEG dataset was collected as a primary dataset at Universitas Syiah Kuala, Indonesia, under a controlled acquisition protocol and environment. The dataset comprised recordings from 16 participants: eight individuals with ASD aged 15–25 years and eight TD controls aged 17–25 years. ASD participants were recruited from a special education school in Banda Aceh, whereas TD controls were recruited from the surrounding university community. The acquisition and participant-group framework are consistent with previous EEG-based ASD datasets and comparative studies [7], [8].

Participant information reported in this study reflects the demographic and diagnostic-group metadata available in the original acquisition records, including group designation, age range, and recruitment source. Additional structured clinical variables, such as standardized ASD severity scores, medication status, detailed comorbidity profiles, and formal group-matching variables, were not part of the analytical metadata and were therefore not retrospectively inferred or used during model development.

Following participant preparation and electrode placement, each EEG recording lasted approximately 15 minutes. The Ethics Committee of the Faculty of Medicine, Universitas Syiah Kuala approved the study protocol (approval no. 117/EA/FK/2024).

2.2. EEG Acquisition Device and Channel Configuration

EEG signals were acquired using an OpenBCI Cyton-based biosensing system at a sampling frequency of 250 Hz. Sixteen EEG channels were positioned according to the international 10–20 system at Fp1, Fp2, F3, F4, F7, F8, C3, C4, T3, T4, T5, T6, P3, P4, O1, and O2, as illustrated in Fig. 2. This configuration is consistent with the acquisition protocol reported in previous studies using the same EEG setup [8]. The ground electrode was positioned at AFz and the common reference electrode at CPz; these auxiliary electrodes were used only for signal acquisition and were not treated as EEG analysis channels.

Recordings were performed under an eyes-open resting condition in a quiet, low-distraction environment between 09:00 and 12:00. Participants were seated comfortably and instructed to remain calm, minimize head and body movements, and face a blank wall. Conductive gel was applied to improve scalp-electrode contact, and signals were monitored in real time using the OpenBCI software. The acquisition hardware incorporated an ADS1299 analog front-end with 24-bit resolution. No dedicated ocular, myogenic, or movement-artifact correction was applied; subsequent preprocessing was limited to the filtering procedures described in Section 2.3.

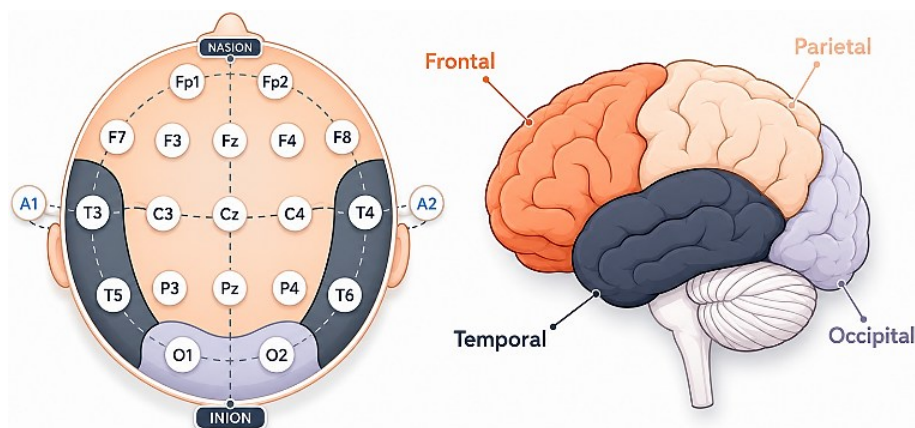


Figure 2. Electrode placement of the 16-channel EEG recording.

2.3. EEG Preprocessing Using Blackman–Harris FIR Filter

Raw continuous EEG recordings were preprocessed before feature extraction to attenuate baseline drift and out-of-band frequency components that could interfere with subsequent analysis. No dedicated ocular, myogenic, or movement-artifact correction procedure was applied; therefore, the filtering stage was not interpreted as removing these artifacts.

Preprocessing was performed independently on each EEG channel using a finite impulse response (FIR) band-pass filter designed with a four-term Blackman–Harris window. This window was selected because its low sidelobe levels and strong stopband attenuation reduce spectral leakage associated with truncation of the ideal impulse response [10], [11], [12]. Filtering was applied to the continuous EEG recordings before epoch segmentation.

The FIR filter was implemented using forward–backward filtering (filtfilt), producing a zero-phase response and preventing phase shifts in the filtered signals. Signal boundaries were handled using the default odd-symmetric extension implemented by SciPy's filtfilt function. For the 101-coefficient FIR filter, a padding length of 303 samples, corresponding to approximately 1.21 s at 250 Hz, was applied at each end of the continuous recording. Filtering the continuous signals before segmentation also avoided introducing separate filtering transients at the boundaries of individual 4-second epochs.

An FIR filter produces each output sample from a finite weighted combination of the current and preceding input samples without recursive feedback. The filtering operation is expressed as:

$$y(n) = \sum_{k=0}^{N-1} h(k)x(n-k) \quad (1)$$

where $y(n)$ is the filtered EEG signal, $x(n-k)$ denotes the input signal delayed by k samples, $h(k)$ represents the FIR coefficient, and N is the number of filter coefficients.

In a window-based FIR design, the final coefficients are obtained by multiplying the ideal impulse response by the selected window function:

$$h(n) = h_d(n)w(n) \quad (2)$$

where $h(n)$ is the resulting FIR coefficient, $h_d(n)$ is the ideal impulse response, and $w(n)$ is the window function. Applying a tapered window smooths the truncated impulse response and reduces spectral leakage associated with abrupt truncation [11].

The Blackman–Harris window has a smooth and symmetric profile with strong sidelobe suppression, as illustrated in Fig. 3.

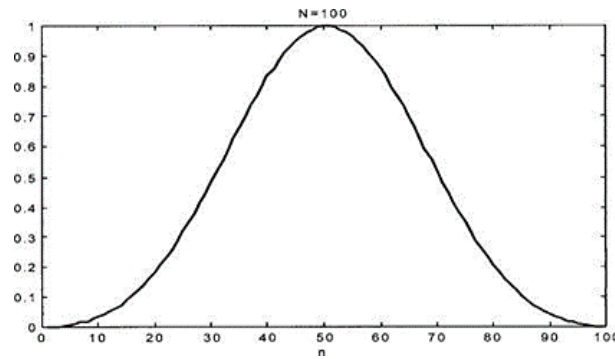


Figure 3. Blackman–Harris window.

The four-term Blackman–Harris window is defined as [12]:

$$w(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{L-1}\right) + a_2 \cos\left(\frac{4\pi n}{L-1}\right) - a_3 \cos\left(\frac{6\pi n}{L-1}\right), 0 \leq n \leq L-1 \quad (3)$$

where L is the window length, n is the sample index, and a_0 , a_1 , a_2 , and a_3 are the Blackman–Harris coefficients. The coefficients used were $a_0 = 0.35875$, $a_1 = 0.48829$, $a_2 = 0.14128$, and $a_3 = 0.01168$. Table 1 summarizes the complete FIR configuration.

Table 1. Configuration of the Blackman–Harris FIR band-pass filter

Parameter	Description / Value
Sampling frequency	250 Hz
Filter type	FIR band-pass filter
Window function	Blackman–Harris window
Passband frequency range	4–40 Hz
Filter order/length	100th-order FIR filter / 101 coefficients
Filtering mode	Zero-phase forward–backward filtering

For the preprocessing ablation reported in Section 3.1, a conventional Butterworth IIR band-pass filter was included as a comparator. The comparator used the same 4–40 Hz frequency range and 250 Hz sampling frequency as the conventional Butterworth

preprocessing approach reported by Mulyadi *et al.* [13]. In the present study, the Butterworth filter was implemented as a fourth-order IIR band-pass filter with lower and upper cut-off frequencies of 4 and 40 Hz, respectively. The filter was applied independently to each EEG channel before epoch segmentation using forward–backward filtering to obtain a zero-phase response. Thus, the Butterworth and Blackman–Harris FIR pipelines used the same 4–40 Hz passband, sampling frequency, filtering direction, and subsequent epoch-generation procedure, while differing in filter structure and design order. Accordingly, the ablation was intended to compare two consistently implemented preprocessing alternatives rather than filters matched by numerical order.

The 4–40 Hz passband retained theta, alpha, beta, and low-gamma activity while attenuating lower-frequency drift and higher-frequency out-of-band components. The effects of the preprocessing alternatives were subsequently assessed using power spectral density analysis and the classification ablation reported in Section 3.1.

The filtered recordings were segmented into 4-second epochs, corresponding to 1,000 samples at a sampling frequency of 250 Hz. Consecutive epochs overlapped by 50%, resulting in a 500-sample (2-second) step; successive epochs therefore covered 0–4 s, 2–6 s, 4–8 s, and so forth.

A total of 7,145 epochs were included across the 16 participants. Fifteen participants contributed 449 epochs each from complete 15-minute recordings, whereas one ASD participant contributed 410 epochs because the recording ended before the planned duration, corresponding to approximately 13.7 minutes of usable EEG data. Each epoch contained all 16 EEG channels and was treated as one sample for epoch-level feature extraction. All epochs from the same participant remained grouped within the same outer cross-validation fold, as described in Section 2.5. The resulting epochs were subsequently processed using the feature-extraction and feature-selection procedures described in Section 2.4.

2.4. Feature Extraction and Feature Selection

After preprocessing, temporal and spectral features were extracted independently from each channel of every 4-second EEG epoch. Eleven features were calculated per channel: log-variance, root mean square (RMS), Hjorth activity, mobility, and complexity, spectral centroid, spectral bandwidth, and absolute band power in the theta, alpha, beta, and gamma bands. These features were selected to represent complementary amplitude, temporal-dynamic, and spectral characteristics of the EEG signals, consistent with previous EEG-based ASD studies [7]. With 16 EEG channels, the resulting representation comprised 176 features per epoch.

For an EEG epoch $x[n]$, $n = 0, \dots, N - 1$, log-variance and RMS were calculated as

$$\text{LogVar} = \ln[\text{Var}(x) + \varepsilon], \quad \text{RMS} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} x^2[n]}, \quad (4)$$

where N is the number of samples and $\varepsilon = 10^{-12}$ is a small constant introduced for numerical stability.

The Hjorth parameters were calculated as

$$\text{Activity} = \text{Var}(x) + \varepsilon, \quad (5)$$

$$\text{Mobility} = \sqrt{\frac{\text{Var}(\Delta x) + \varepsilon}{\text{Var}(x) + \varepsilon}}, \quad (6)$$

$$\text{Complexity} = \frac{\sqrt{\frac{\text{Var}(\Delta^2 x) + \varepsilon}{\text{Var}(\Delta x) + \varepsilon}}}{\text{Mobility} + \varepsilon}, \quad (7)$$

where Δx and $\Delta^2 x$ denote the first and second discrete differences, respectively. Hjorth parameters were included because they characterize EEG signal activity and temporal complexity and have previously been applied to the analysis of ASD-related EEG patterns [7].

Spectral features were derived from a Hann-windowed one-sided real Fourier transform. The power spectral density (PSD) was estimated as

$$X[k] = rFFT x[n]w[n], \quad P[k] = \frac{|X[k]|^2}{f_s \sum_{n=0}^{N-1} w^2[n]}, \quad (8)$$

where $w[n]$ denotes the Hann window and f_s is the sampling frequency. Each 4-second epoch contained $N = 1,000$ samples at $f_s = 250$ Hz; therefore, a 1,000-point rFFT was applied, providing a frequency resolution of 0.25 Hz. No Welch segmentation or spectral averaging was used for feature extraction.

Using the resulting PSD, the spectral centroid f_c and spectral bandwidth B_w were calculated as

$$f_c = \frac{\sum_k f_k P[k]}{\sum_k P[k]}, \quad B_w = \sqrt{\frac{\sum_k P k^2}{\sum_k P[k]}}, \quad (9)$$

where f_k represents the frequency associated with the k -th spectral bin.

Absolute band power was calculated by integrating the PSD within each predefined frequency range:

$$P_B = \int_{f_l}^{f_h} P(f), df, \quad (10)$$

where f_l and f_h denote the lower and upper limits of the corresponding frequency band. Numerical integration was performed using the trapezoidal rule. The frequency ranges were theta, 4–8 Hz; alpha, 8–13 Hz; beta, 13–30 Hz; and gamma, 30–40 Hz. The same feature definitions and spectral parameters were applied consistently to all channels, epochs, and participants.

Mutual Information (MI) was subsequently used for model-specific feature selection within the nested subject-wise cross-validation procedure described in Section 2.5. MI ranking and feature selection were performed exclusively within the corresponding training partitions, without access to the outer-test participants. Any subsequent feature scaling was performed according to the classifier-specific pipeline described in Section 2.6.

2.5. Nested Subject-Wise 8-Fold Cross-Validation

A nested subject-wise cross-validation procedure was used to minimize participant leakage during model development and performance estimation. In the outer loop, the 16 participants (8 ASD and 8 TD) were divided into eight stratified folds, each containing one ASD and one TD participant. In each outer iteration, one fold comprising two participants was reserved exclusively for testing, while the remaining 14 participants were used for model development. All epochs belonging to the same participant remained within the same partition throughout the procedure.

Model selection was performed exclusively within the 14 outer-training participants using an inner subject-wise cross-validation loop. The inner folds were participant-disjoint, such that epochs from an inner-validation participant were never included in the corresponding inner-

training set. Classifier-specific preprocessing, MI estimation, feature ranking and selection, model hyperparameter selection, and, where applicable, CatBoost early stopping were performed using only the inner-training and inner-validation data. Z-score normalization was applied only to the SVM-RBF pipeline and was fitted exclusively on the corresponding training partition.

After the optimal configuration was determined from the inner cross-validation results, the selected preprocessing and feature-selection parameters were refitted using all 14 outer-training participants, and the resulting model was evaluated once on the two previously unseen outer-test participants. This procedure was repeated for all eight outer folds so that each participant contributed exactly once to the out-of-fold test predictions.

For model selection, candidate MI feature-retention fractions of 10%, 20%, ..., 100% were evaluated within the inner cross-validation loop. For SVM-RBF, the candidate search space comprised $C \in \{0.001, 0.01, 0.1, 1, 10, 100\}$ and $\gamma \in \{\text{scale}, 0.01, 0.1\}$. For CatBoost, candidate configurations combined learning rates of $\{0.01, 0.03\}$, tree depths of $\{2, 4\}$, and L2 leaf-regularization values of $\{10, 50\}$. Candidate configurations were ranked primarily using mean inner-validation balanced accuracy, with lower inner-fold standard deviation and the predefined tie-breaking criteria used when necessary.

2.6. Classification Models

The 176-dimensional EEG feature vectors were classified into ASD and TD classes using SVM-RBF and CatBoost as the primary classifiers, with Linear Discriminant Analysis (LDA) and Gaussian Naive Bayes (GNB) included as additional baseline models. All classifiers received the same initial feature representation. Model-specific feature selection and model configuration were performed exclusively within the nested subject-wise validation procedure described in Section 2.5, with the outer-test participants remaining completely unseen during feature selection and model development.

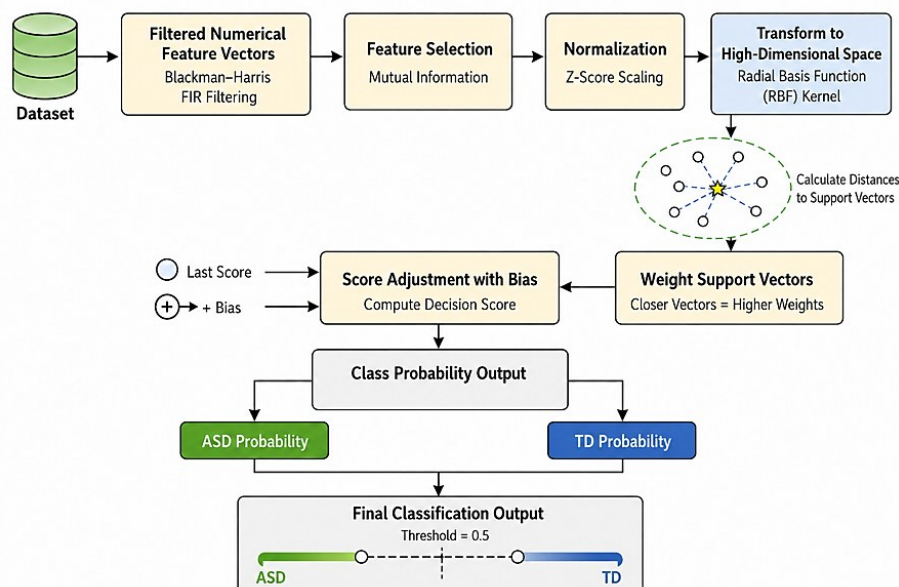


Figure 4. SVM-RBF classification workflow.

2.6.1. Support Vector Machine with Radial Basis Function Kernel

SVM-RBF classified the selected EEG feature vectors after Blackman-Harris FIR preprocessing, feature extraction, MI-based feature selection, and Z-score normalization, as

illustrated in Fig. 4. The RBF kernel enables nonlinear separation of EEG feature patterns by mapping pairwise similarity according to [14]:

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (11)$$

where x and x' are EEG feature vectors and γ controls the kernel width. The SVM decision function is expressed as:

$$f(x) = \text{sign}(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b) \quad (12)$$

where α_i denotes the support-vector coefficient, y_i is the corresponding class label, $K(x_i, x)$ is the kernel similarity, and b is the bias term.

SVM-RBF model selection was performed exclusively within the inner subject-wise cross-validation loop. For each inner split, Z-score normalization and MI ranking were fitted using only the inner-training participants and subsequently applied to the participant-disjoint inner-validation set. Candidate feature-retention fractions and SVM-RBF configurations were evaluated using mean inner-validation balanced accuracy. After the optimal configuration was selected, MI ranking and normalization were refitted using all participants in the corresponding outer-training set, and the final SVM-RBF model was trained without access to the outer-test participants. The resulting configuration used $C = 0.01$, $\gamma = \text{scale}$, and the top 30% of MI-ranked features. The trained model was then evaluated once on the previously unseen outer-test participants.

2.6.2. CatBoost

CatBoost is a gradient-boosted decision-tree classifier that iteratively combines weak learners while using ordered boosting to reduce prediction shift [16]. As illustrated in Fig. 5, the selected EEG features were used to estimate class probabilities for ASD and TD classification.

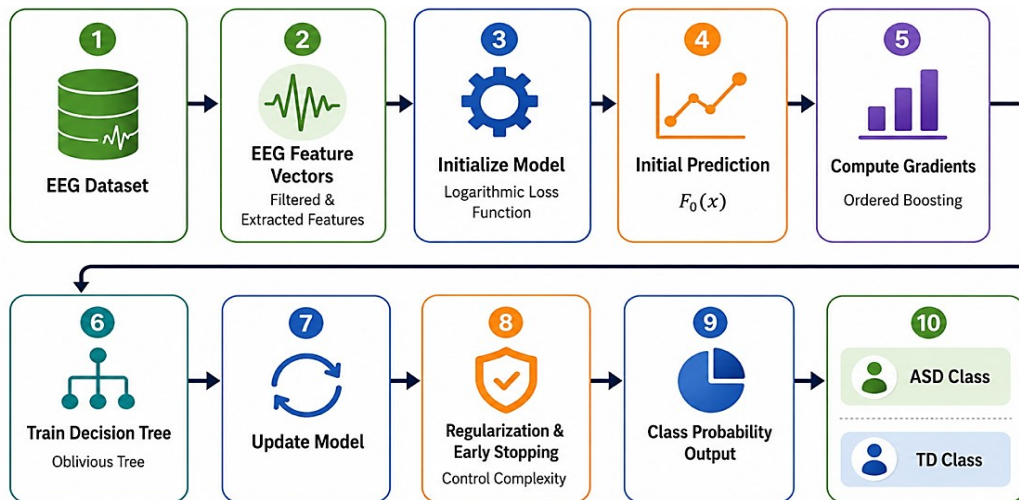


Figure 5. CatBoost classification workflow.

The binary classification objective can be represented as:

$$\mathcal{L}(\theta) = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (13)$$

where $L(y_i, \hat{y}_i)$ denotes the classification loss, $\Omega(f_k)$ represents the regularization term for the k -th tree, and $\hat{y}_i$ is the predicted output. The ensemble is updated iteratively according to:

$$F_{t+1}(x) = F_t(x) + \eta h_t(x) \quad (14)$$

where $F_t(x)$ is the ensemble at iteration t , η is the learning rate, and $h_t(x)$ is the newly added decision tree.

CatBoost feature selection, hyperparameter selection, and early-stopping assessment were performed exclusively within the inner subject-wise cross-validation loop. For each inner split, MI ranking was fitted using only the inner-training participants and subsequently applied to the participant-disjoint inner-validation set. No feature scaling was applied to CatBoost because of its tree-based structure. Candidate CatBoost configurations were evaluated using mean inner-validation balanced accuracy, while early stopping was monitored exclusively on the inner-validation participants. No epochs, labels, or predictions from the outer-test participants were used for feature selection, hyperparameter selection, or early stopping. After the optimal configuration was selected, MI ranking was refitted using all participants in the corresponding outer-training set, and the final CatBoost model was retrained using the selected hyperparameters and boosting-iteration setting determined from the inner-validation procedure. The model was then evaluated once on the previously unseen outer-test participants. The final CatBoost configuration used a maximum of 800 boosting iterations, a learning rate of 0.01, tree depth of 2, L2 leaf regularization of 50, bagging temperature of 2.0, an early-stopping patience of 30 iterations, and the top 40% of MI-ranked features.

2.6.3. Additional Baseline Models

LDA and GNB were included as additional baseline classifiers to determine whether the observed class separation was specific to the primary nonlinear models. Both baselines were evaluated using the same outer subject-wise folds and the same initial 176-dimensional feature representation. All data-dependent preprocessing and feature-selection operations were restricted to the corresponding training partitions.

LDA used the singular-value decomposition (SVD) solver without shrinkage, whereas GNB used a variance-smoothing parameter of 1×10^{-9} . Both baseline models used the top 40% of MI-ranked features within the nested validation procedure. These models were reported as baseline pipelines rather than as evidence of direct classifier superiority because SVM-RBF and the other models did not necessarily retain identical final feature subsets.

2.6.4. Participant-Level Aggregation

Participant-level predictions were generated exclusively from the out-of-fold epoch predictions of each held-out participant. Majority voting was used as the primary aggregation rule. For participant s , the proportion of epochs predicted as ASD was calculated as:

$$V_s = \frac{1}{N_s} \sum_{j=1}^{N_s} \mathbb{I}(\hat{y}_{sj} = 1), \quad (15)$$

where N_s is the number of epochs belonging to participant s , $\hat{y}_{sj}$ is the predicted label for epoch j , with ASD coded as 1 and TD as 0, and $\mathbb{I}(\cdot)$ is the indicator function. The final participant-level prediction was defined as:

$$\hat{Y}_s = \begin{cases} 1 \text{ (ASD)}, & V_s > 0.5, \\ 0 \text{ (TD)}, & V_s < 0.5, \\ \mathbb{I}(\bar{p}_s \geq 0.5), & V_s = 0.5, \end{cases} \quad (16)$$

where $\bar{p}_s$ denotes the mean out-of-fold ASD probability across all epochs of participant s . Thus, when ASD and TD votes were equal, the participant was classified using a prespecified probability threshold of 0.5; no voting ties occurred in the present analysis.

Each participant contributed exactly one final classification decision, whether 449 or 410 epochs were available, preventing participants with more epochs from receiving greater weight

in the participant-level metrics. Accuracy, sensitivity, specificity, precision, and F1-score were calculated from the 16 participant-level decisions. Exact two-sided Clopper–Pearson 95% confidence intervals were calculated for accuracy, sensitivity, and specificity.

As a sensitivity analysis, participant-level predictions were also obtained by averaging the out-of-fold ASD probabilities across all epochs of each participant and applying the same fixed threshold of 0.5. Neither the majority-voting rule nor the probability threshold was optimized using the outer-test predictions.

2.6.5. Performance Evaluation

All four classifiers were evaluated using the outer subject-wise folds of the nested cross-validation procedure described in Section 2.5. Performance was assessed using accuracy, precision, sensitivity (recall), specificity, F1-score, ROC-AUC, and confusion matrices, with ASD treated as the positive class and TD as the negative class. Participant-level accuracy, sensitivity, and specificity were additionally reported with exact two-sided Clopper–Pearson 95% confidence intervals [19].

Accuracy was designated as the primary overall metric for the inferential comparison between SVM-RBF and CatBoost, whereas precision, sensitivity, specificity, and F1-score were retained as complementary descriptive measures. An exploratory paired-samples t-test was performed using the accuracy values obtained from the same eight outer subject-wise test folds [20], [21]. For fold i , the paired difference was defined as:

$$d_i = A_{\text{CatBoost},i} - A_{\text{SVM-RBF},i}, \quad (17)$$

such that a positive d_i indicated higher accuracy for CatBoost. The paired t -statistic was calculated as

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}, \quad (18)$$

where $\bar{d}$ is the mean paired difference, s_d is the standard deviation of the paired differences, and $n = 8$ is the number of outer folds. Statistical significance was assessed at a nominal level of $\alpha = 0.05$. The two-sided 95% confidence interval for the mean difference was calculated as

$$\bar{d} \pm t_{0.975,n-1} \frac{s_d}{\sqrt{n}}, \quad (19)$$

and the standardized paired effect size was calculated using Cohen's d_z :

$$d_z = \frac{\bar{d}}{s_d}. \quad (20)$$

Because the eight outer folds contained substantially overlapping training sets, their accuracy estimates were not fully independent [22]. Therefore, the paired t-test, confidence interval, and effect size were interpreted as exploratory within-cohort summaries rather than confirmatory participant- or population-level inference. No multiplicity correction was applied because accuracy was the only metric subjected to this inferential comparison.

To complement the exploratory fold-wise analysis, an exact participant-level permutation test was performed using the final out-of-fold CatBoost participant predictions. Participant identity was treated as the exchangeable unit, while the original class composition of eight ASD and eight TD participants was preserved. Because the cohort contained only 16 participants, all possible class-balanced diagnostic-label assignments were evaluated exhaustively rather than approximated using Monte Carlo sampling. The total number of unique assignments was

$$M = \binom{16}{8} = 12,870. \quad (21)$$

For each assignment, the diagnostic labels were reassigned to the fixed out-of-fold participant predictions, and participant-level accuracy was recalculated. The exact one-sided permutation p -value was defined as

$$p_{\text{perm}} = \frac{1}{M} \sum_{b=1}^M I(A_b \geq A_{\text{obs}}), \quad (22)$$

where A_{obs} denotes the observed participant-level accuracy, A_b is the accuracy obtained under the b -th class-balanced label assignment, $I(\cdot)$ is the indicator function, and M is the total number of unique class-balanced assignments defined in Eq. (21).

This analysis tested whether the association between the held-out participant predictions and the diagnostic labels exceeded that expected under random class-balanced label assignment. Because all possible assignments were enumerated, no Monte Carlo approximation or correction was required. The test was conditional on the observed out-of-fold predictions and was therefore interpreted as a within-cohort robustness analysis rather than as an assessment of external generalizability [23].

3. RESULTS AND DISCUSSION

This section presents the preprocessing ablation results, out-of-fold epoch-level classifier performance, participant-level classification using majority voting, exploratory fold-wise statistical comparisons, and the study's principal findings and limitations.

3.1. EEG Signal Preprocessing Results

Blackman–Harris FIR filtering attenuated baseline drift and frequency components outside the 4–40 Hz passband and was not interpreted as an artifact-removal procedure, as no dedicated ocular, myogenic, or movement-artifact correction was applied. Representative ASD and TD signals before and after filtering are shown in Figs. 6 and 7, while the corresponding PSD comparison in Fig. 8 confirms attenuation below 4 Hz and above 40 Hz with preservation of activity within the selected theta, alpha, beta, and low-gamma ranges.

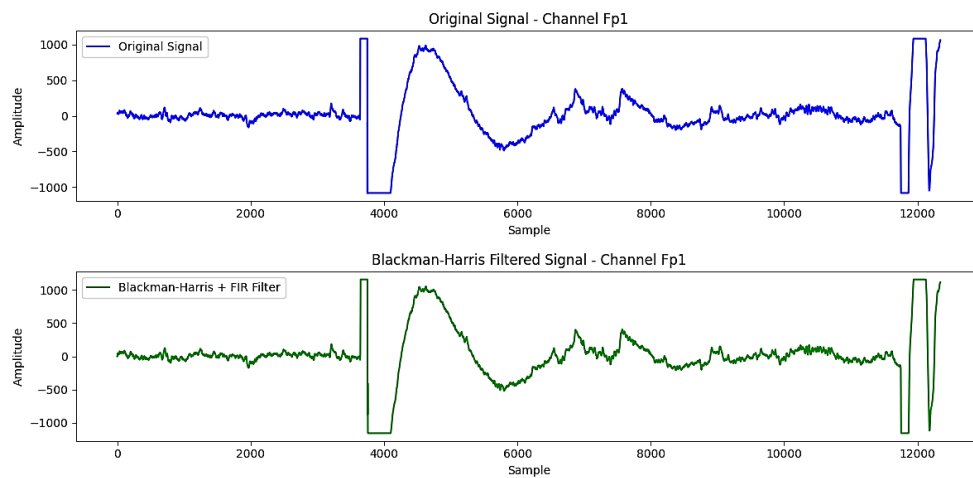


Figure 6. ASD EEG signals before and after Blackman–Harris FIR filtering.

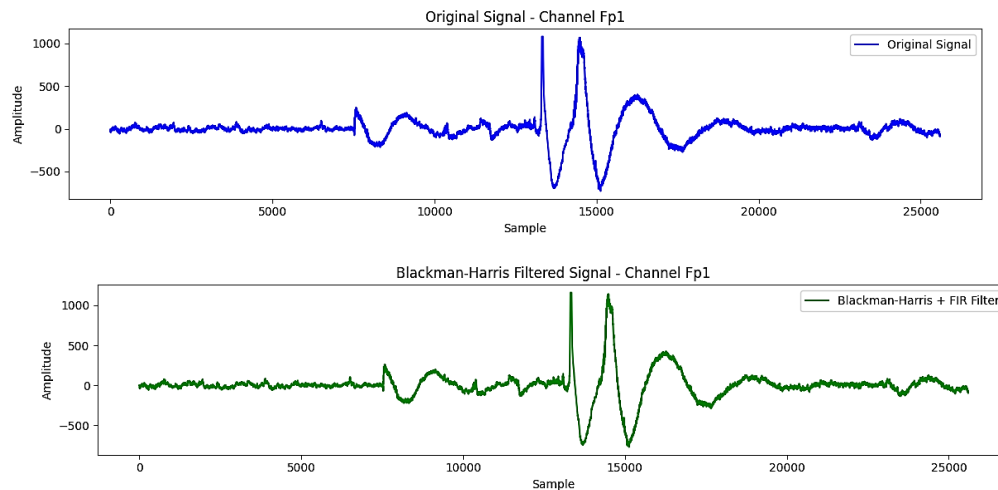


Figure 7. TD EEG signals before and after Blackman–Harris FIR filtering.

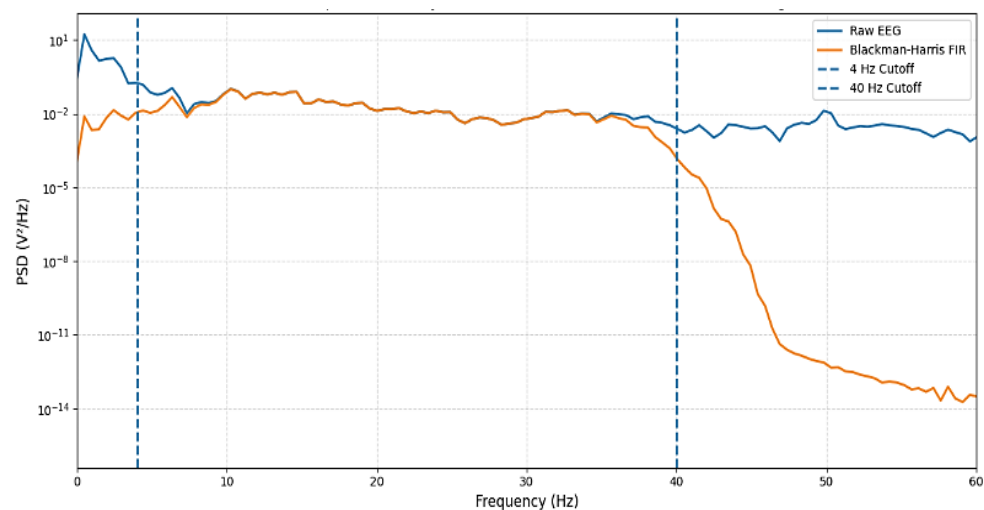


Figure 8. PSD of EEG signals before and after Blackman–Harris FIR filtering.

To quantify the effect of preprocessing on classification, raw EEG, Butterworth band-pass filtering, and Blackman–Harris FIR filtering were compared, as summarized in Table 2.

Table 2. Performance comparison of EEG preprocessing approaches

Preprocessing pipeline	Mean accuracy (%)	SD (%)
Raw EEG + SVM-RBF	80.01	1.14
Raw EEG + CatBoost	88.92	3.88
Butterworth band-pass + SVM-RBF	88.53	1.16
Butterworth band-pass + CatBoost	93.97	0.86
Blackman–Harris FIR + SVM-RBF	91.59	9.15
Blackman–Harris FIR + CatBoost	99.58	0.23

Both filtering approaches improved accuracy relative to raw EEG, with Blackman–Harris FIR producing the highest accuracy for both classifiers. The Blackman–Harris FIR–CatBoost pipeline achieved the best performance at 99.58%, indicating that preprocessing influenced feature discriminability within the present cohort.

3.2. Epoch-Level Performance of SVM-RBF

SVM-RBF was evaluated using out-of-fold predictions from the outer folds of the nested subject-wise eight-fold cross-validation. In each fold, one ASD and one TD participant were held out for testing, with all epochs from each participant remaining exclusively within the corresponding partition. Table 3 summarizes the fold-wise epoch-level results.

Table 3. Out-of-fold epoch-level performance of SVM-RBF across eight outer folds

Fold	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-score (%)	Specificity (%)
Fold 1	93.21	100.00	86.41	92.71	100.00
Fold 2	71.71	64.03	99.11	77.80	44.32
Fold 3	94.88	100.00	89.76	94.60	100.00
Fold 4	99.44	100.00	98.89	99.44	100.00
Fold 5	88.31	81.73	98.66	89.40	77.95
Fold 6	88.47	88.40	87.32	87.85	89.53
Fold 7	99.89	100.00	99.78	99.89	100.00
Fold 8	96.77	100.00	93.54	96.66	100.00
Mean	91.59	91.77	94.18	92.29	88.98
SD	9.15	13.20	5.67	7.29	19.73

SVM-RBF achieved a mean fold-wise accuracy of 91.59%, with notable variability across held-out participants. Fold 2 showed the lowest performance, where specificity decreased to 44.32% because 250 of 449 TD epochs were classified as ASD, whereas specificity reached 100% in five folds. Despite this variability, sensitivity remained generally high across folds. The corresponding out-of-fold confusion matrices are shown in Fig. 9.

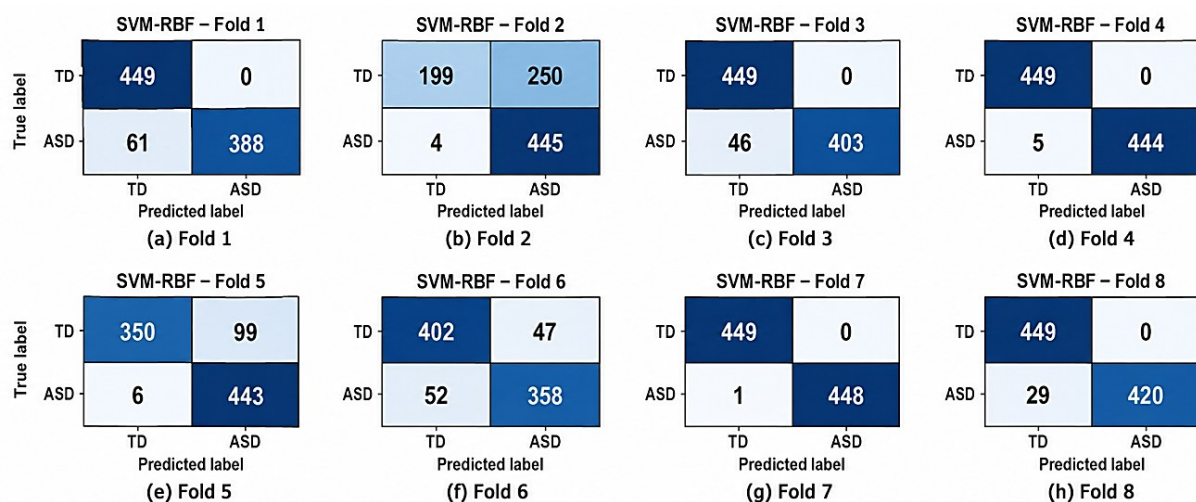


Figure 9. Confusion matrices of SVM-RBF across eight outer folds.

3.3. Epoch-Level Performance of CatBoost

CatBoost was evaluated using the same eight outer subject-wise folds, with fold-wise out-of-fold epoch-level performance summarized in Table 4. CatBoost achieved consistently high performance across all outer folds, with a mean accuracy of 99.58% and low fold-to-fold variability (SD = 0.23%). Accuracy ranged from 99.33% to 99.88%. Across the 7,145 out-of-fold epoch predictions, CatBoost produced eight false-positive and 22 false-negative classifications. Fig. 10 shows the corresponding confusion matrices.

Table 4. Out-of-fold epoch-level performance of CatBoost across eight outer folds

Fold	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-score (%)	Specificity (%)
Fold 1	99.33	100.00	98.66	99.33	100.00
Fold 2	99.33	98.90	99.78	99.33	98.89
Fold 3	99.78	99.78	99.78	99.78	99.78
Fold 4	99.67	100.00	99.33	99.66	100.00
Fold 5	99.78	100.00	99.55	99.78	100.00
Fold 6	99.88	99.76	100.00	99.88	99.78
Fold 7	99.55	99.78	99.33	99.55	99.78
Fold 8	99.33	100.00	98.66	99.33	100.00
Mean	99.58	99.78	99.39	99.58	99.78
SD	0.23	0.37	0.50	0.23	0.38

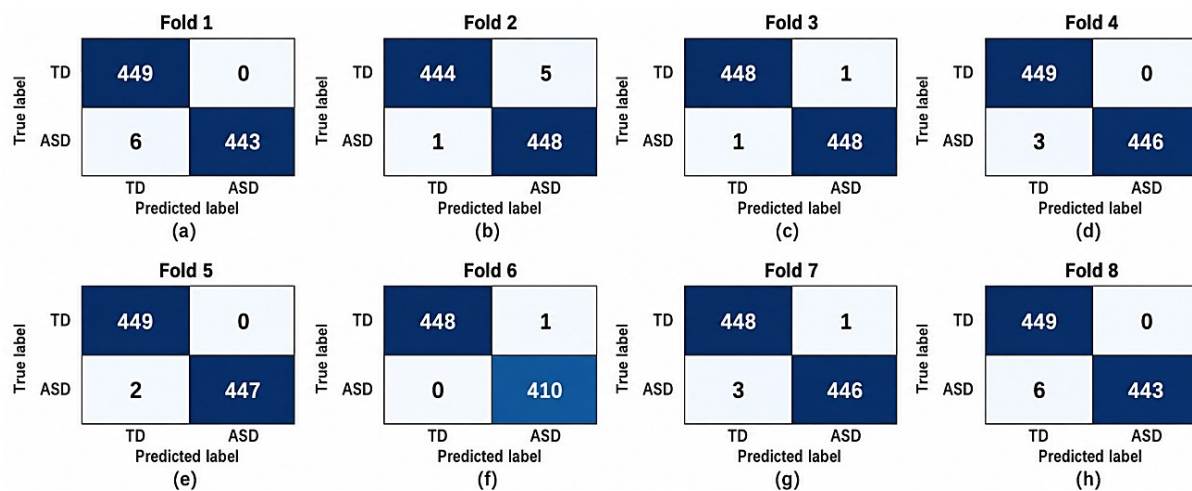


Figure 10. Confusion matrices of CatBoost across eight outer folds.

3.4. Pooled Out-of-Fold Epoch-Level Comparison

The pooled out-of-fold predictions from all 7,145 epochs were used to compare the four classification pipelines. SVM-RBF retained the top 30% of MI-ranked features, whereas CatBoost, LDA, and Gaussian Naive Bayes retained the top 40%. Thus, Table 5 compares separately configured pipelines rather than classifiers using identical final feature subsets.

Table 5. Pooled out-of-fold epoch-level performance of the four classification pipelines

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	ROC-AUC (%)
CatBoost	99.58	99.77	99.38	99.78	99.58	99.99
LDA	97.47	96.65	98.31	96.63	97.47	99.58
SVM-RBF	91.60	89.43	94.26	88.98	91.78	98.21
Gaussian Naive Bayes	57.49	53.94	99.49	15.95	69.95	63.41

CatBoost achieved the highest pooled accuracy, precision, specificity, F1-score, and ROC-AUC, followed by LDA and SVM-RBF. Gaussian Naive Bayes showed high sensitivity but very low specificity (15.95%), indicating a strong tendency to classify TD epochs as ASD. These results compare the performance of the configured pipelines and do not, by themselves, exclude potential cohort-related confounding.

3.5. Participant-Level Classification Using Majority Voting

Participant-level performance was obtained by majority voting across the out-of-fold epoch predictions of each held-out participant. Each participant contributed one final decision regardless of the number of available epochs, thereby preventing participants with more epochs from receiving greater weight. Table 6 summarizes the results.

Table 6. Participant-level performance using majority voting

Model	Accuracy, % (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	F1-score (%)
CatBoost	100.00 (79.41–100.00)	100.00 (63.06–100.00)	100.00 (63.06–100.00)	100.00
LDA	100.00 (79.41–100.00)	100.00 (63.06–100.00)	100.00 (63.06–100.00)	100.00
SVM-RBF	93.75 (69.77–99.84)	100.00 (63.06–100.00)	87.50 (47.35–99.68)	94.12
Gaussian Naive Bayes	56.25 (29.88–80.25)	100.00 (63.06–100.00)	12.50 (0.32–52.65)	69.57

CatBoost and LDA correctly classified all 16 participants, whereas SVM-RBF classified 15/16 and Gaussian Naive Bayes 9/16. Despite the perfect CatBoost and LDA point estimates, the 95% CI for accuracy was 79.41–100.00%, reflecting uncertainty due to the small participant sample.

As a sensitivity analysis, participant labels were also obtained by averaging out-of-fold ASD probabilities and applying the fixed threshold of 0.5. CatBoost, LDA, and Gaussian Naive Bayes produced unchanged decisions, whereas SVM-RBF improved from 15/16 under majority voting to 16/16 under probability averaging, indicating some sensitivity to the aggregation rule.

For CatBoost, an exact participant-level permutation test evaluated all $M = \binom{16}{8} = 12,870$ class-balanced diagnostic-label assignments. Only one assignment achieved an accuracy equal to the observed 100.00%, yielding $p_{perm} = 7.77 \times 10^{-5}$ ($p < 0.001$). This supports an association between the held-out CatBoost predictions and diagnostic labels beyond random class-balanced assignment, while remaining conditional on the observed out-of-fold predictions and therefore representing within-cohort rather than external or clinical evidence.

3.6. Exploratory Fold-Wise Statistical Comparison of Accuracy

Accuracy was treated as the primary inferential metric, while precision, sensitivity, specificity, and F1-score were reported descriptively. An exploratory paired t-test compared CatBoost and SVM-RBF accuracy across the same eight outer subject-wise folds. Because these folds shared substantially overlapping training sets, the resulting inference was interpreted as within-cohort exploratory evidence rather than confirmatory population-level evidence.

Table 7. Exploratory paired t-test results for CatBoost minus SVM-RBF fold-wise differences

Comparison	Mean difference (pp)	95% CI (pp)	t- statistic	p- value	Cohen's d_z	Conclusion ($\alpha = 0.05$)
CatBoost – SVM-RBF	8.00	0.38–15.61	2.48	0.042	0.88	Nominally significant; exploratory

CatBoost achieved a mean fold-wise accuracy 8.00 percentage points higher than SVM-RBF (95% CI: 0.38–15.61; $t = 2.48$, $p = 0.042$, Cohen's $d_z = 0.88$). Although the difference

was nominally significant with a large standardized effect, it should be interpreted cautiously because the eight folds were not fully independent and the analysis was based on only 16 participants.

3.7. Reproducibility and Implementation Details

All analyses were implemented in Python 3.12.12 using NumPy 2.0.2, Pandas 2.3.3, SciPy 1.16.3, scikit-learn 1.6.1, and CatBoost 1.2.10. Experiments were conducted on Kaggle Notebook under Linux 6.6.122+ x86_64 with glibc 2.35, an Intel Xeon 2.00 GHz CPU (four cores), 31.35 GB RAM, and two NVIDIA Tesla T4 GPUs with 15,360 MiB memory each. A fixed random seed of 1234 was used for data splitting, MI estimation, and algorithms supporting random-state control. All data-dependent preprocessing, feature selection, model tuning, and early-stopping procedures followed the nested subject-wise protocol described in Section 2.5.

SVM-RBF used $C = 0.01$, $\gamma = \text{scale}$, and the top 30% of MI-ranked features. CatBoost used up to 800 iterations, a learning rate of 0.01, depth 2, L2 leaf regularization of 50, bagging temperature of 2.0, early-stopping patience of 30, and the top 40% of MI-ranked features. LDA used the SVD solver without shrinkage, while Gaussian Naive Bayes used a variance-smoothing parameter of 1×10^{-9} ; both baselines retained the top 40% of MI-ranked features. All results in Sections 3.2–3.6 were derived from the same 7,145 out-of-fold predictions for each classifier.

3.8. Discussion

The preprocessing ablation showed that filtering influenced the discriminability of the extracted EEG features. Both Butterworth and Blackman–Harris FIR filtering improved performance relative to raw EEG, with Blackman–Harris FIR producing the highest accuracy for the evaluated classifiers. However, the 4–40 Hz filtering stage only attenuated baseline drift and out-of-band components and did not specifically remove ocular, myogenic, or movement artifacts.

At the epoch level, CatBoost achieved the highest pooled performance and the lowest fold-to-fold variability, followed by LDA and SVM-RBF. SVM-RBF was more sensitive to the held-out participants, particularly in Fold 2, where specificity decreased to 44.32%. The exploratory CatBoost–SVM-RBF comparison yielded an 8.00-percentage-point difference in fold-wise accuracy ($p = 0.042$, 95% CI: 0.38–15.61; Cohen’s $d_z = 0.88$). Because the outer folds shared substantially overlapping training sets, interpret this result as exploratory within-cohort evidence rather than confirmatory population-level inference.

Participant-level evaluation provides the more clinically relevant assessment because each participant contributes only one final decision. CatBoost and LDA correctly classified all 16 participants, whereas SVM-RBF classified 15/16 under majority voting and 16/16 under probability averaging. Thus, participant-level superiority of CatBoost over LDA was not established, and the SVM-RBF result showed some sensitivity to the aggregation rule. Moreover, despite 100% accuracy for CatBoost and LDA, the 95% confidence interval remained wide, reflecting the uncertainty associated with only eight participants per group.

Using 50% overlapping 4-second epochs increased the number of segments but introduced dependence between adjacent observations. Subject-wise cross-validation prevented participant leakage between outer training and test sets, but epoch-level results still represent correlated repeated observations from the same individuals and should therefore be regarded as secondary measures of segment-level discriminability. In addition, the present ablation did

not separately quantify the effects of MI-based feature selection, temporal versus spectral feature families, or epoch overlap; these component-wise effects require further evaluation under the same nested subject-wise framework.

The single-site design also leaves open potential recruitment, demographic, recording-session, and hardware-related confounding. High classification accuracy across several models does not exclude these alternative explanations. Accordingly, the findings should be regarded as preliminary within-cohort evidence, and larger, demographically matched, independent, and multi-site datasets are required before considering the framework for ASD screening or clinical decision support.

Finally, the exact participant-level permutation test for CatBoost yielded $p_{\text{perm}} = 7.77 \times 10^{-5}$, indicating that the observed association between held-out predictions and diagnostic labels was unlikely under random class-balanced assignment. Nevertheless, this test remained conditional on the observed out-of-fold predictions and does not establish external validity or eliminate potential cohort- and acquisition-related confounding.

4. CONCLUSIONS

This study developed an EEG-based ASD classification framework combining Blackman–Harris FIR preprocessing, temporal and spectral feature extraction, MI-based feature selection, and machine-learning classification. Blackman–Harris FIR preprocessing yielded higher classification accuracy than raw EEG and Butterworth filtering. Among the evaluated pipelines, CatBoost achieved the highest pooled out-of-fold epoch-level accuracy (99.58%), followed by LDA (97.47%), SVM-RBF (91.60%), and Gaussian Naive Bayes (57.49%). The exploratory fold-wise comparison favored CatBoost over SVM-RBF ($p = 0.042$, Cohen's $d_z = 0.88$), although this result should be interpreted cautiously because the outer folds were not fully independent. At the participant level, CatBoost and LDA correctly classified all 16 participants, while SVM-RBF and Gaussian Naive Bayes classified 15/16 and 9/16, respectively. Thus, participant-level superiority of CatBoost over LDA was not established. The exact CatBoost permutation test also exceeded random class-balanced label assignment ($p < 0.001$). Overall, the findings support preliminary within-cohort feasibility, but the small sample size, overlapping epochs, single-site acquisition, and lack of external validation preclude claims of clinical or screening readiness.

ACKNOWLEDGEMENT

The authors thank Universitas Syiah Kuala for institutional and technical support.

REFERENCES

- [1] Rogala J, et al. (2023) Enhancing autism spectrum disorder classification in children through the integration of traditional statistics and classical machine learning techniques in EEG analysis. *Sci. Rep.*, 13:21748. <https://doi.org/10.1038/s41598-023-49048-7>
- [2] Romero M, et al. (2016) Psychiatric comorbidities in autism spectrum disorder: A comparative study between DSM-IV-TR and DSM-5 diagnosis. *Int. J. Clin. Health Psychol.*, 16(3):266-275. <https://doi.org/10.1016/j.ijchp.2016.03.001>
- [3] Wantzen P, et al. (2022) EEG resting-state functional connectivity: evidence for an imbalance of external/internal information integration in autism. *J. Neurodev. Disord.*, 14:47. <https://doi.org/10.1186/s11689-022-09456-8>

-
- [4] Kang J, Xie H, Mao W, Wu J, Li X, Geng X (2023) EEG connectivity diversity differences between children with autism and typically developing children: A comparative study. *Bioengineering*, 10(9):1030. <https://doi.org/10.3390/bioengineering10091030>
 - [5] Dissanayake T, Fernando T, Denman S, Sridharan S, Fookes C (2021) Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals. *IEEE Sens. J.*, 21(7):9377-9388. <https://doi.org/10.1109/JSEN.2021.3057076>
 - [6] Melinda M, Gazali S, Away Y, Rafiki A, Wong WK, Mulyadi M, Rusdiana S (2026) Heavy–light soft-vote fusion of EEG heatmaps for autism spectrum disorder detection. *J. Electron. Electromed. Eng. Med. Inform.*, 8(1):409-429. <https://doi.org/10.35882/jeeemi.v8i1.1377>
 - [7] Mulyadi M, Melinda M, Away Y, Gazali S (2025) Comparative analysis of normal and ASD EEG data using Hjorth parameters. 2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA), pp. 1-6. <https://doi.org/10.1109/ICHORA65333.2025.11017066>
 - [8] Melinda M, Purnamasari PD, Fahmi F, Sinulingga EP, Mulyadi M, Away Y, Yunidar Y, Juwono FH (2025) A comprehensive EEG dataset and performance assessment for autism spectrum disorder. *Sci. Rep.*, 15:34981. <https://doi.org/10.1038/s41598-025-18934-7>
 - [9] Ding Y, Robinson N, Zhang S, Zeng Q, Guan C (2023) TSception: Capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition. *IEEE Trans. Affect. Comput.*, 14(3):2238-2250. <https://doi.org/10.1109/TAFFC.2022.3169001>
 - [10] Pant A, Kumar A (2024) Hanning FIR window filtering analysis for EEG signals. *Biomedical Analysis*, 1(2):111-123. <https://doi.org/10.1016/j.bioana.2024.05.003>
 - [11] Pant A, Kumar A, Verma C, Illés Z (2024) Comparative exploration on EEG signal filtering using window control methods. *Results Control Optim.*, 17:100485. <https://doi.org/10.1016/j.rico.2024.100485>
 - [12] Harris FJ (1978) On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE*, 66(1):51-83. <https://doi.org/10.1109/PROC.1978.10837>
 - [13] Mulyadi M, Melinda M, Away Y, Gazali S, Rafiki A, Wong WK (2026) Bayesian-optimized Butterworth-PID preprocessing for EEG-based autism spectrum disorder classification. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2026.3709120>
 - [14] Mastropietro A, Feldmann C, Bajorath J (2023) Calculation of exact Shapley values for explaining support vector machine models using the radial basis function kernel. *Sci. Rep.*, 13:19561. <https://doi.org/10.1038/s41598-023-46930-2>
 - [15] Geeitha S, Ravishankar K, Cho J, Easwaramoorthy SV (2024) Integrating cat boost algorithm with triangulating feature importance to predict survival outcome in recurrent cervical cancer. *Sci. Rep.*, 14:19828. <https://doi.org/10.1038/s41598-024-67562-0>
 - [16] Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A (2018) CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pp. 6638-6648.
 - [17] Nassir LM, Ramadhan AJ, Al-Sharify NT, Khalaf MI, Ogaili AAF, Jaber AA, Al-Sharify ZT (2025) Robust multi-state EEG cognitive classification via optimized time-domain features and CatBoost. *Int. J. Robot. Control Syst.*, 5(2):968-989. <https://doi.org/10.31763/ijres.v5i2.1799>
 - [18] Bhaskar N, Borhade RR, Barekar S, Bachute M, Bairagi V (2024) CNN-CatBoost ensemble deep learning model for enhanced disease detection and classification of kidney disease. *Indonesian J. Electr. Eng. Comput. Sci.*, 34(1):144-151. <https://doi.org/10.11591/ijeecs.v34.i1.pp144-151>
 - [19] Garthwaite PH, Moustafa MW, Elfadaly FG (2024) Locally correct confidence intervals for a binomial proportion: A new criteria for an interval estimator. *Scand. J. Stat.*, 51(1):220-244. <https://doi.org/10.1111/sjos.12672>
-

- [20] Chicco D, Sichenze A, Jurman G (2025) A simple guide to the use of Student's t-test, Mann-Whitney U test, Chi-squared test, and Kruskal-Wallis test in biostatistics. *BioData Min.*, 18:56. <https://doi.org/10.1186/s13040-025-00465-6>
- [21] Nariya MK, Mills CE, Sorger PK, Sokolov A (2023) Paired evaluation of machine-learning models characterizes effects of confounders and outliers. *Patterns*, 4(8):100791. <https://doi.org/10.1016/j.patter.2023.100791>
- [22] Jafrasteh B, Adeli E, Pohl KM, Kuceyeski A, Sabuncu MR, Zhao Q (2025) Statistical variability in comparing accuracy of neuroimaging based classification models via cross validation. *Sci. Rep.*, 15:28745. <https://doi.org/10.1038/s41598-025-12026-2>
- [23] Steinert S, Ruf V, Dzsotjan D, Großmann N, Schmidt A, Kuhn J, Küchemann S (2024) A refined approach for evaluating small datasets via binary classification using machine learning. *PLoS ONE*, 19(5):e0301276. <https://doi.org/10.1371/journal.pone.0301276>

Performance Evaluation of Boosting Algorithms on Intrusion Detection System Based on Real-Time CICIOT2023 Dataset

MUHAMMAD HAFIZ AMRULLAH¹, FAVIAN DEWANTA^{1*},
MUHAMAD ERZA AMINANTO²

¹*School of Electrical Engineering, Telkom University, Bandung, Indonesia*

²*Cybersecurity, Monash University, Jakarta, Indonesia*

*Corresponding author: favian@telkomuniversity.ac.id

(Received: 20 June 2026; Accepted: 15 August 2026; Published online: 11 September 2026)

ABSTRACT: Currently, the utilization of machine learning methods for identifying intrusions in Internet of Things (IoT) networks demonstrates intriguing prospects. Choosing the right machine learning technique for an Intrusion Detection System (IDS) is challenging, as an unsuitable algorithm can reduce threat-detection accuracy and compromise network security. This research comprehensively assesses various boosting algorithms—including GBM, XGB, LGBM, CatBoost, and AdaBoost—using a stratified sample of 100,000 records from the latest CICIOT2023 dataset. The evaluation encompasses both binary (2-Class) and multi-class (8-Class) classification across four ablation sampling scenarios: Pure Sample, Up-Sample, Down-Sample, and Hybrid-Sample. To simulate real-world deployment and avoid unsubstantiated real-time claims, we formally deployed and evaluated the trained models on a lightweight Lubuntu Virtual Machine (VM) with a single 1-core CPU in resource-constrained environments (2 GB and 4 GB RAM) to measure empirical inference rates and latency. The empirical results demonstrate that GBM achieves the highest performance during VM deployment, reaching 99.9% accuracy for binary tasks and 99.0% for multi-class tasks, while maintaining a high inference throughput of 680,341 samples per second. Conversely, XGB provides the best trade-off during the offline phase by significantly reducing training time while maintaining near-peak accuracy. These findings provide critical guidelines for deploying lightweight, high-performance boosting frameworks for securing resource-constrained IoT infrastructures.

ABSTRAK: Pada masa ini, penggunaan kaedah pembelajaran mesin untuk mengenal pasti pencerobohan dalam rangkaian *Internet of Things* (IoT) menunjukkan prospek yang menarik. Memilih teknik pembelajaran mesin yang sesuai untuk *Intrusion Detection System* (IDS) menimbulkan cabaran yang ketara, kerana algoritma yang tidak sesuai boleh mengakibatkan pengurangan ketepatan dalam mengesan ancaman dan menjejaskan keselamatan rangkaian. Kajian ini menjalankan penilaian komprehensif terhadap pelbagai algoritma penggalak (*boosting*)—merangkumi GBM, XGB, LGBM, CatBoost, dan AdaBoost—menggunakan sampel terstratifikasi sebanyak 100,000 rekod daripada set data CICIOT2023 terkini. Pendekatan ini merangkumi penilaian algoritma tersebut dalam senario pengelasan binari (2-Class) dan berbilang kelas (8-Class) merentasi empat senario pensampelan ablasi. Bagi mensimulasikan penggunaan dunia nyata, model yang telah dilatih turut dihantar dan dinilai secara formal pada Mesin Maya (VM) ringan Lubuntu berspesifikasi 1 teras CPU di bawah persekitaran kekangan sumber dengan konfigurasi RAM 2 GB dan 4 GB untuk mengukur kadar inferensi serta kependaman masa nyata secara empirikal. Keputusan empirikal menunjukkan bahawa GBM menghasilkan tahap ketepatan tertinggi semasa simulasi VM, iaitu 99.9% untuk tugas binari dan 99.0% untuk tugas berbilang kelas, dengan kadar inferensi mencapai 680,341 sampel sesaat. Sebaliknya, XGB memberikan imbangan optimum

(trade-off) pada fasa luar talian dengan memangkas masa latihan secara signifikan sambil mengekalkan ketepatan yang tinggi. Temuan ini memberikan garis panduan krusial dalam melaksanakan kerangka kerja *boosting* yang ringan dan berprestasi tinggi untuk mengamankan infrastruktur IoT yang mempunyai keterbatasan sumber daya.

KEYWORDS: *Internet of Things, Intrusion Detection System, Boosting Algorithms, Machine Learning, Lightweight Virtual Machine, CICIOT2023.*

1. INTRODUCTION

The Internet of Things (IoT) has expanded substantially in recent years, with millions of devices connecting to the internet daily [1]. These IoT devices generate large volumes of data suitable for various purposes such as environmental monitoring, automated control systems, and data analytics [2]. As a result, IoT has been adopted across many sectors, including industry, manufacturing, healthcare, and daily life. The rapid advancement of IoT technology has transformed businesses by enabling seamless connectivity, information exchange, and automation [3].

Nevertheless, the extensive implementation of IoT also introduces new security challenges, particularly in identifying network breaches. The IoT comprises a wide variety of devices with diverse capabilities but often limited in functionality. This diverse IoT landscape is prone to malicious attacks and vulnerable to various security issues [4]. Ongoing advances in IoT standards and protocols have led to inconsistent security practices across IoT devices and environments [5]. Poor implementation of security protocols can create vulnerabilities that attackers exploit to take advantage of network gaps. IoT devices are often used in physically exposed and unmanaged environments, such as manufacturing facilities or public spaces [6]. Greater access to a device's physical components increases the risk of tampering, unauthorized access, and overall security compromise.

Ensuring the security of IoT systems is crucial for mitigating and overcoming these security risks [7]. Protecting IoT networks is essential to safeguard sensitive data, maintain privacy, and uphold the reliability of IoT devices. An Intrusion Detection System (IDS) is one strategy that prioritizes IoT security. The IDS protects the system's Confidentiality, Integrity, and Availability (CIA) from a variety of attacks. Put simply, security tools like an IDS detect activity that violates security policies. Therefore, modern network architectures strongly emphasize using an effective IDS [8].

Effective network intrusion detection is vital for providing immediate protection in IoT environments. This method centralizes network data to distinguish benign from malicious network behavior. Effective strategies are essential to detect and respond to network breaches in IoT devices, given their extensive deployment, diversity, and severe resource constraints. Conversely, as attackers become more advanced, new risks and weaknesses emerge rapidly. The likelihood of cyberattacks penetrating critical infrastructure increases substantially over short periods. Consequently, there is a growing need for an IDS that can identify and respond to novel IoT-specific cyber threats.

Researchers have developed various methodologies to improve IDS accuracy and effectiveness. Machine Learning (ML) is considered one of the leading approaches to address these challenges [8]. An IDS primarily differentiates between legitimate and harmful network traffic flowing through major network nodes. Moreover, it must determine the exact classification of an attack within the protected system by examining specific network fingerprints. Therefore, developing highly accurate anomaly detection models is essential.

Machine Learning (ML) can be employed to develop Anomaly Detection Systems (ADS) capable of accurately detecting and identifying attacks on IoT networks. Among ML approaches, boosting methods are widely used to improve classification precision. Boosting works by iteratively building a series of weak models, where each subsequent estimator is trained on the remaining residuals from the previous model, focusing predominantly on correcting significant classification errors [9]. These weak models are then combined into a stronger ensemble structure.

However, to bridge the gap between theoretical, cloud-based evaluations and practical deployment, it is essential to benchmark these algorithms in resource-constrained environments that mathematically and physically mirror actual IoT edge components.

The primary contributions of this paper can be summarized as follows:

- We evaluate and contrast multiple boosting classification algorithms using a statistically justified, stratified sample of 100,000 records from the latest large-scale IoT dataset, CICIoT2023.
- We provide a rigorous performance comparison of five state-of-the-art boosting frameworks using standardized international acronyms: Gradient Boosting Machine (GBM), eXtreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM), Categorical Boosting (CatBoost), and Adaptive Boosting (AdaBoost).
- We perform an empirical deployment simulation by evaluating the trained models on a resource-constrained Lubuntu Virtual Machine (VM) configured with a 1-core CPU and restricted RAM, replacing generic "real-time" claims with formal latency and throughput metrics.
- We conduct an extensive ablation analysis of preprocessing procedures (No balancing vs. SMOTE vs. Undersampling vs. Hybrid) applied strictly after train-test splitting to completely eliminate data leakage.

The paper is structured into five main sections. Section 1 offers an overview of security challenges in IoT. Section 2 presents the critical synthesis of related works. Section 3 outlines the methodology employed. Section 4 presents the analysis derived from simulation outcomes and statistical tests. Finally, Section 5 interprets the findings, limitations, and future directions.

2. RELATED WORK

Boosting algorithms have existed for a considerable time, but only recently have they gained popularity and acceptance within the machine learning community. This section provides an overview and analysis of existing research on network intrusion detection in IoT contexts, focusing specifically on boosting methods.

Joffrey et al. [10] conducted a study consisting of two primary components: ensemble feature selection and model comparison. The study used the CSE-CIC-IDS2018 dataset. In the initial phase, the researchers selected specific attributes from the dataset as input for two classifiers. In the subsequent phase, the researchers used these classifiers for classification. In the second phase, the researchers assessed the classifier's performance using the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) and F1-score. The results show that the LGBM algorithm performs well, with AUC and F1-score values of 0.9689 and 0.96134.

Eman et al. [11] introduced the concept of a Federated Intrusion Detection Blockchain (FIDC) that uses lightweight Artificial Neural Networks (ANN) with Federated Learning (FL) to safeguard the privacy of healthcare data. This is achieved by leveraging blockchain technology, which provides a decentralized ledger. Researchers evaluated the performance of the ANN and XGB models using the BoT-IoT dataset. The results indicate that the ANN model shows superior accuracy and performance when handling data heterogeneity on IoT devices. The researchers tested FIDChain on various datasets, including Bot-Net-IoT, CSE-CIC-IDS2018, and KDDCup99. The findings indicated that the BoT-IoT dataset yielded the most dependable and accurate results when evaluating IoT applications, particularly those utilized in Healthcare systems. The accuracy improvements achieved were 99.99% with ANN and 98.40% with XGB.

Garg et al. [12] conducted a study to identify DDoS and IoT attacks using three datasets: IoT-23, BoT-IoT, and CIC-DDoS2019. The research included training and testing with two gradient boosting strategies: XGB and Light Gradient Boosting Machine (LGBM). The study then evaluated these methods against Cascaded Deep Forest (CDF). The feature extraction and selection (FES) process uses two widely recognized techniques: analysis of variance (ANOVA) and principal component analysis (PCA). The boosting approaches achieve at least 16% higher accuracy than CDF. Boosting algorithms show a minimum speed advantage of 240 times over CDF. Between the two boosting algorithms, LGBM has the shortest execution time, taking 54 seconds or less, and achieves an accuracy of up to 94.79%.

Maryam et al. [13] employed the CICIoT2023 dataset to develop machine learning methods for effectively identifying abnormal network traffic. In their study, the dataset undergoes pre-processing, and they use a balanced class representation to build unbiased supervised machine learning models with several algorithms, including Random Forest, Adaptive Boosting, Logistic Regression, Perceptron, and Deep Neural Network. These models undergo further analysis by removing highly correlated features, reducing dimensionality, mitigating overfitting, and accelerating training. The Random Forest algorithm demonstrated excellent performance in classifying IoT Attacks, with an estimated accuracy of 99.55% in both binary and multiclass classification tasks. This high accuracy was observed in both restricted and full feature space scenarios. Adaptive Boosting method achieved multiclass accuracy of 48.38% and binary class accuracy of 99.40%. This enhancement was praised for its concurrent reduction in response time, a critical factor for promptly identifying and addressing attacks in real time.

In a recent pivotal work, Saied et al. [18] provided a foundational comparative analysis of ensemble architectures for botnet detection in the N-BaIoT dataset. They thoroughly demonstrated the robustness of tree-based boosting structures in discerning complex network anomalies across large traffic captures. While their study offered significant theoretical insights into the hierarchical superiority of XGB and GBM over traditional classifiers, their evaluation was confined to standard, high-resource workstation environments without imposing hardware latency restrictions.

Collectively, these studies leave a substantial "deployment gap." By providing descriptive summaries of high classification accuracy on cloud-tier hardware, previous literature neglects the severe hardware limitations—such as restricted RAM and low-power 1-core CPUs—that strictly characterize actual IoT edge devices. Our research explicitly distinguishes itself from Saied et al. [18] and others by shifting the paradigm from theoretical performance to formal, deployment-level efficiency. By imposing strict hardware constraints through a Ubuntu 1-core VM emulation, this work enriches the discourse with physical inference metrics and provides a realistic synthesis of how algorithmic complexity interacts with processing bottlenecks at the

IoT edge. Table 1 summarizes the key research objectives and constraints across the discussed literature.

Table 1. Comparative analysis of related works

Ref	Year	Algorithm	Dataset	Objective	Accuracy(%)
[10]	2020	XGB, LGBM	CSE-CIC-IDS2018	Ensemble feature selection and model comparison	96.89 & 96.13
[11]	2022	XGB	N-BaIoT, CIC-IDS2018, KDDCup99	FIDC using Federated Learning (FL) to safeguard the privacy of healthcare data	98.4
[12]	2022	XGB, LGBM	IoT-2023, BoT-IoT, CIC-DDoS-19	Identifying IoT attacks and DDoS attacks	94.49 & 94.79
[13]	2024	AdaBoost	CICIoT2023	Develop machine learning methods for effectively identifying abnormal network traffic	48.38 (Multi-class) 99.4 (Binary class)
[18]	2023	AdaBoost, GBM, XGB, CatBoost, HGB, LGBM	N-BaIoT	Investigating the potential of boosting-based methods for detecting IoT botnet attacks	99.99 (HGB)
Ours	2026	GBM, XGB, LGBM, CatBoost, AdaBoost	CICIoT2023	Empirical deployment benchmarking on resource-constrained (1-core CPU, 2&4 GB RAM) IoT edge environments	99.00 (Multi) & 99.99 (Binary)

3. METHODOLOGY

This section outlines the research methodology used to design, evaluate, and benchmark the machine learning framework. Figure 1 illustrates the macro-level operational architecture of the proposed evaluation pipeline.

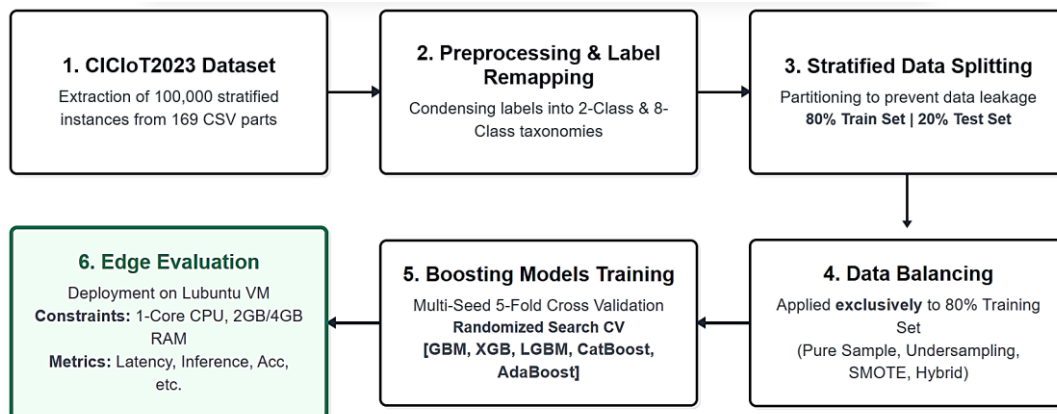


Figure 1. Machine learning pipeline model

3.1. Overview and Strategic Selection of the CICIoT2023 Dataset

Developed by the Canadian Institute for Cybersecurity, CICIoT2023 is a state-of-the-art benchmark dataset designed to advance security analytics in large-scale IoT infrastructures [14]. Figure 2 details the multi-layered IoT network topology deployed during data collection. The architecture models a smart home ecosystem comprising 105 physical IoT gadgets targeted by 33 distinct cyber-attack profiles, summarized in Table 2.

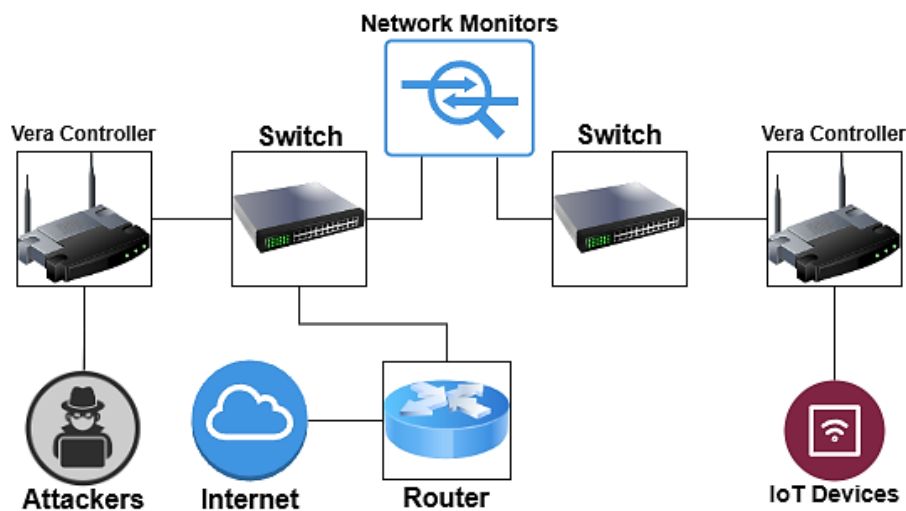


Figure 2. Network topology for generating the CICIOT2023 dataset

To ensure the efficiency and reproducibility of the experimental pipeline in a restricted Google Colab environment, we implemented a systematic data-reduction strategy. The original CICIOT2023 dataset consists of 169 fragmented CSV files, which collectively create significant memory overhead that exceeds the standard RAM allocation of cloud-based runtime environments and often leads to system crashes or kernel termination.

To address this challenge, we performed a global stratified sampling approach. Instead of selecting subsets from a limited number of files, we aggregated the class distribution profile across all 169 dataset parts. We then applied a stratified sampling technique to extract a total of 100,000 records. This sampling methodology ensures that the proportional representation of both benign traffic and various intrusion categories is preserved, maintaining the statistical integrity of the original dataset while reducing the memory footprint to a manageable size and preventing out-of-memory (OOM) failures.

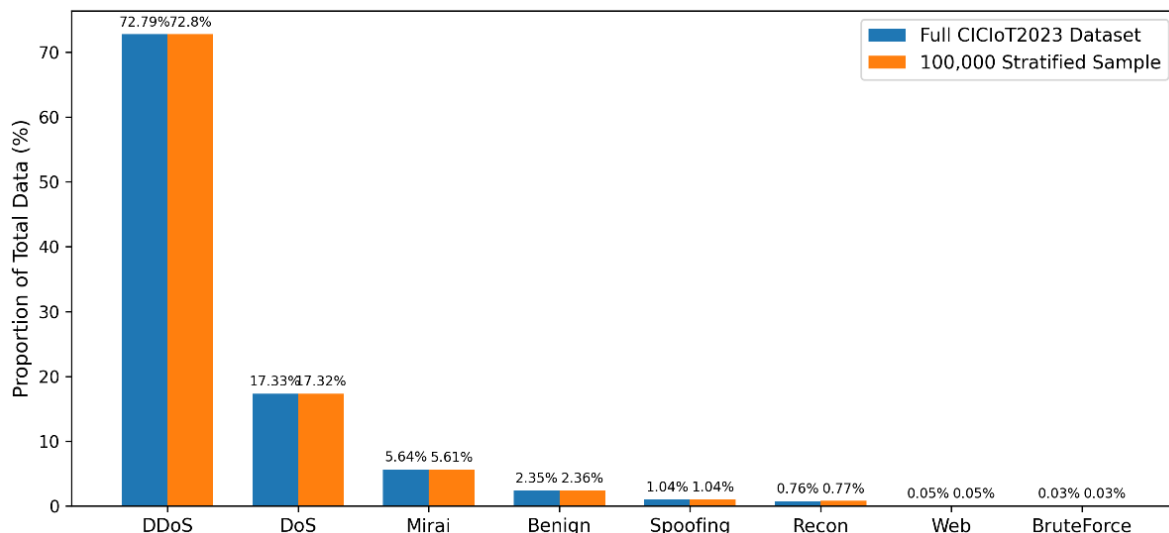


Figure 3. Empirical Comparison of Class Distribution (Full Dataset vs. 100,000 Samples)

To empirically demonstrate the representativeness of the extracted 100,000-record sample against the full-scale CICIOT2023 dataset, we conducted a comparative distribution analysis. As shown in Figure 3, the stratified sampling methodology mirrors the exact proportional ratios

of the full population across all 8 class taxonomies. For instance, the absolute majority class (DDoS) constitutes 72.79% of the original 46 million records and is perfectly retained at 72.80% within the extracted sample. Similarly, the extreme minority class (BruteForce) represents 0.03% in both the full dataset and the localized subset. This empirical alignment confirms that the 100,000-record subset used in this research is statistically unbiased and fully representative of the real-world IoT traffic patterns provided by the original CICIoT2023 benchmark.

Table 2. Attack Traffic in CICIoT-2023 dataset

Attack	Number	Attack	Number
DDoS-ICMP_Flood	7,200,504	DoS-TCP_Flood	2,671,445
DDoS-UDP_Flood	5,412,287	DoS-SYN_Flood	2,028,834
DDoS-TCP_Flood	4,497,667	BenignTraffic	1,098,195
DDoS-PSHACK_Flood	4,094,755	Mirai-greeth_flood	991,866
DDoS-SYN_Flood	4,059,190	Mirai-udpplain	890,576
DDoS-RSTFINFlood	4,045,285	Mirai-greip_flood	751,682
DDoS-SynonymousIP_Flood	3,598,138	DDoS-ICMP_Fragmentation	452,489
DoS-UDP_Flood	3,318,595	MITM-ArpSpoofing	307,593
Recon-PingSweep	2,262	Uploading_Attack	1,252
DDoS-UDP_Fragmentation	286,925	DDoS-HTTP_Flood	28,790
DDoS-ACK_Fragmentation	285,104	DDoS-SlowLoris	23,426
DNS_Spoofing	178,911	DictionaryBruteForce	13,064
Recon-HostDiscovery	134,378	BrowserHijacking	5,859
Recon-OSScan	98,259	CommandInjection	5,409
Recon-PortScan	82,284	SqlInjection	5,245
DoS-HTTP_Flood	71,864	XSS	3,846
VulnerabilityScan	37,382	Backdoor_Malware	3,218

3.2. Model Construction, Ablation, and Data Leakage Prevention

The model development pipeline begins with uniform class mapping, compressing the 34 native labels into 2-Class (Binary) and 8-Class (Multi-class) taxonomies. After label harmonization, we intentionally omitted feature scaling (e.g., Min-Max or Z-score) because boosting architectures are intrinsically scale-invariant [15]. Hastie et al. [19] and Breiman et al. [20] show that tree node splits rely strictly on monotonic rank-ordering; avoiding unnecessary scaling preserves identical decision boundaries while significantly reducing computational overhead during real-time edge inference.

To strictly prevent data leakage—a critical vulnerability in which testing data inadvertently influences model training—we first extracted a stratified subset of 100,000 instances to preserve original class proportions. We then partitioned this subset into isolated training (80.0%) and testing (20.0%) subarrays via a stratified split before any data balancing. An extensive ablation analysis was then conducted exclusively on the training set to evaluate preprocessing resilience:

1. **Pure Sample (Baseline):** No balancing applied.
2. **Down-Sample:** RandomUnderSampler [16] was applied to compress majority classes to the level of the absolute minority.
3. **Up-Sample:** Synthetic Minority Over-sampling Technique (SMOTE) [17] was applied strictly to the training set to mathematically synthesize minority points based on k-nearest neighbors.

4. **Hybrid-Sample:** A combined approach utilizing SMOTETomek (oversampling followed by Tomek Link cleaning) to smooth decision boundaries.

Crucially, the 20% testing partition remained entirely unmodified throughout all ablation stages, guaranteeing that model evaluation reflects purely unseen, real-world data distributions. Table 3 details the final training set dimensions across the ablation scenarios.

Table 3. Training Set Dimensions Across Ablation Sampling Scenarios

Classification Task	Sampling Scenario	Instances (Rows)	Features
8-Class (Multi)	Pure Sample (Baseline)	80,000	46
	Down Sample (Under)	216	46
	Up Sample (SMOTE)	465,936	46
	Hybrid Sample	464,222	46
2-Class (Binary)	Pure Sample (Baseline)	80,000	46
	Down Sample (Under)	3,780	46
	Up Sample (SMOTE)	156,220	46
	Hybrid Sample	156,058	46

The isolated, balanced training arrays are subsequently fed into five state-of-the-art tree-based boosting architectures: GBM, XGB, LGBM, CatBoost, and AdaBoost. Hyperparameter configurations—including learning rates, maximum tree depths, and estimator counts—are optimized using a standardized Randomized Search CV. To ensure statistical robustness and mitigate random-initialization bias, the evaluation phase uses a rigorous multi-seed 5-fold cross-validation strategy. Specifically, we ran the cross-validation pipeline independently across three random seeds (42, 100, and 2026) to ensure fairness across all five frameworks. We report the final classification metrics and training latencies as the mean, with the standard deviation across these multi-seed folds, ensuring that all comparative claims about model efficacy are statistically validated and immune to data-split variation.

3.3. Hardware Emulation Scenario in a Resource-Constrained Environment

To substantiate the deployment readiness of the compiled boosting estimators without relying on unsubstantiated "real-time" claims, this study integrates a localized hardware emulation phase. The trained models are serialized into .joblib files and transferred onto an isolated virtual machine running Lubuntu OS, a lightweight Linux distribution optimal for IoT gateway emulation.

To evaluate performance resilience under hardware degradation and strict edge computing limits, the evaluation is systematically executed across two independent memory profiles:

1. A highly restricted 2.0 GB RAM memory constraint.
2. A moderate 4.0 GB RAM memory constraint.

During this implementation phase, we evaluate the models using 1,000 untouched samples from the isolated testing set withheld during training. Beyond capturing standard performance indicators (Accuracy, Precision, Recall, and Macro F1-score), the emulation infrastructure tracks physical computation metrics directly from the OS environment, specifically recording:

- **Inference Rate (samples/second):** The total volume of network packets successfully classified per second.
- **Latency per Sample (ms/sample):** The exact duration required to compute a prediction for a single network packet.

This operational profiling shows how the structural complexity and memory footprint of different boosting frameworks interact with real physical bottlenecks, providing critical insights for deploying machine learning-based IDS on memory-restricted IoT edge devices.

4. RESULTS AND DISCUSSION

The evaluation was executed systematically to dissect the critical trade-offs between accuracy and computational overhead, utilizing the environments detailed in Table 4.

Table 4. Simulation parameter list: Experimental and Hardware Emulation Environments

Parameter / Component	Phase 1: Model Training Environment	Phase 2: Hardware Emulation (Edge Deployment)
Platform	Google Colaboratory (Cloud)	Oracle VirtualBox (Local VM)
Operating System	Ubuntu 22.04 LTS	Lubuntu 22.04 LTS (Lightweight OS)
CPU / Processor	Intel Xeon	1 vCPU core
RAM (Memory)	~53 GB (Colab Pro L4)	Tested across two profiles: 2.0 GB and 4.0 GB
Programming Language	Python 3.x	Python 3.x
Core Libraries	Scikit-learn, Imbalanced-learn, Pandas	Joblib, Pandas, Scikit-learn (for inference)
Task Executed	Data Preprocessing, Model Training	Loading .joblib models, Measuring Inference Latency

4.1. Model Training Performance and Hyperparameter Optimization

To ensure complete algorithmic fairness, we determined the optimal hyperparameters for each of the five boosting architectures via Randomized Search CV under identical cross-validation split seeds. The finalized optimal hyperparameter configurations were settled as follows:

- **Gradient Boosting Machine (GBM):** {'n_estimators': 25, 'max_depth': 3, 'learning_rate': 0.1}
- **AdaBoost:** {'n_estimators': 30, 'learning_rate': 0.1}
- **XGBoost (XGB):** {'n_estimators': 50, 'max_depth': 3, 'learning_rate': 0.1}
- **LightGBM (LGBM):** {'n_estimators': 50, 'max_depth': 3, 'learning_rate': 0.1}
- **CatBoost:** {'iterations': 50, 'depth': 3, 'learning_rate': 0.1}

Tables 5 and 6 summarize empirical validation accuracy (Mean $\pm$ Standard Deviation) and offline training latency (ms) across the multi-seed cross-validation splits. To validate the comparative claims formally, a Wilcoxon signed-rank test ($\alpha = 0.05$) was conducted across the folds. The test confirmed that the performance superiority of GBM and XGB over AdaBoost in the multi-class task under Pure Sample conditions is statistically significant ($p < 0.01$), proving that this dominance is not a byproduct of random initialization. However, it is crucial to state explicitly that this study does not claim marginal accuracy differences among closely performing top-tier models (e.g., XGBoost versus LightGBM) to be statistically significant. For these closely grouped architectures, the primary deployment differentiator is empirical computational latency and edge memory efficiency, rather than absolute classification precision.

Table 5. Empirical Performance and Training Latencies for 8-Class (Multi-Class) Classification

Sampling Scenario	Metric / Latency	GBM	AdaBoost	XGB	LGBM	CatBoost
Pure Sample (80,000 rows)	Mean Accuracy	0.9930 ± 0.0008	0.9005 ± 0.0002	0.9906 ± 0.0004	0.9883 ± 0.0040	0.9860 ± 0.0009
	Train Time (ms)	66,248.01	4,115.21	328.23	659.14	689.20
Down Sample (216 rows)	Mean Accuracy	0.6959 ± 0.0683	0.2671 ± 0.0413	0.6773 ± 0.0536	0.7051 ± 0.0691	0.6743 ± 0.0506
	Train Time (ms)	492.84	69.98	194.87	50.00	639.87
Up Sample (465,936 rows)	Mean Accuracy	0.9212 ± 0.0013	0.3089 ± 0.0591	0.9188 ± 0.0013	0.9421 ± 0.0008	0.8596 ± 0.0031
	Train Time (ms)	1,150,132.43	65,099.00	1,131.26	3,729.25	939.16
Hybrid Sample (464,222 rows)	Mean Accuracy	0.9221 ± 0.0020	0.2618 ± 0.0042	0.9198 ± 0.0016	0.9430 ± 0.0012	0.8625 ± 0.0031
	Train Time (ms)	1,091,629.04	62,996.50	1,118.65	5,651.37	939.99

Table 6. Empirical Performance and Training Latencies for 2-Class (Binary) Classification

Sampling Scenario	Metric / Latency	GBM	AdaBoost	XGB	LGBM	CatBoost
Pure Sample (80,000 rows)	Mean Accuracy	0.9956 ± 0.0004	0.9861 ± 0.0009	0.9943 ± 0.0008	0.9942 ± 0.0008	0.9936 ± 0.0006
	Train Time (ms)	8,130.05	3,928.51	114.62	208.67	700.66
Down Sample (3,780 rows)	Mean Accuracy	0.9904 ± 0.0027	0.9845 ± 0.0036	0.9904 ± 0.0038	0.9915 ± 0.0035	0.9900 ± 0.0030
	Train Time (ms)	590.25	311.99	50.16	23.48	655.54
Up Sample (156,220 rows)	Mean Accuracy	0.9927 ± 0.0006	0.9877 ± 0.0008	0.9945 ± 0.0006	0.9945 ± 0.0007	0.9937 ± 0.0008
	Train Time (ms)	33,145.60	14,979.07	217.56	565.23	743.23
Hybrid Sample (156,058 rows)	Mean Accuracy	0.9932 ± 0.0005	0.9882 ± 0.0005	0.9949 ± 0.0003	0.9949 ± 0.0003	0.9939 ± 0.0005
	Train Time (ms)	32,984.17	15,233.62	222.68	795.03	739.45

4.2. Behavioral Analysis of the Preprocessing Ablation

The empirical results reveal key behavioral patterns. Under the multi-class Pure Sample scenario, GBM achieves peak accuracy (99.30%) but incurs significant latency (66,248.01 ms) due to single-core sequential processing. Conversely, XGB demonstrates its algorithmic superiority by achieving comparable accuracy (99.06%) in merely 328.23 ms.

When preprocessing shifted to the Down-Sample ablation (compressing the data to 216 rows), model capacity degraded aggressively. Accuracies plummeted to ~67%, and AdaBoost collapsed to 26.71%. Without sufficient spatial volume, the models overfitted on tabular noise. However, under extensive data scaling via SMOTE (Up-Sample, 465,936 rows), GBM becomes an immense computational bottleneck (1,150,132.43 ms). Under this ablation, LGBM

demonstrates superior resilience. Using leaf-wise tree growth, it processed the artificially inflated synthetic data rapidly (3,729.25 ms), mitigating the SMOTE latency penalty.

4.3. Edge-VM Deployment Emulation and Per-Class Profiling

To substantiate real-world deployment viability, we executed the serialized boosting estimators within the resource-constrained Lubuntu Virtual Machine configured with a single 1-core CPU. The testing set comprised exactly 1,000 uncorrupted, unbalanced test packets to prevent synthetic inflation of accuracy metrics. Table 7 details the actual physical performance metrics recorded across the 2.0 GB and 4.0 GB RAM profiles for the 2-Class (Binary) classification models under the baseline Pure Sample scenario, while Table 8 reports the physical performance of the 8-Class (Multi-Class) counterparts. The hardware emulation benchmarks reported in Tables 7 and 8 provide crucial system-level insights into deploying boosting classifiers at the IoT network edge.

Table 7. Real-Time Hardware Emulation Performance for 2-Class (Binary) Classification

Algorithm	Latency (ms/sample) @ 2GB	Latency (ms/sample) @ 4GB	Inference Rate (samples/sec) @ 2GB	Inference Rate (samples/sec) @ 4GB	Accuracy (VM)	Precision (VM)	Recall (VM)	F1-Score (VM)
AdaBoost	0.00638	0.01222	156,831.70	81,810.71	0.989	0.733	0.880	0.800
CatBoost	0.00406	0.00415	246,291.65	241,127.47	0.997	0.958	0.920	0.939
GBM	0.00147	0.00127	680,341.07	788,468.18	0.999	1.000	0.960	0.980
LGBM	0.00222	0.00205	451,373.12	486,905.17	0.994	0.880	0.880	0.880
XGB	0.00441	0.01221	226,971.27	81,914.43	0.997	0.923	0.960	0.941

Table 8. Real-Time Hardware Emulation Performance for 8-Class (Multi-Class) Classification

Algorithm	Latency (ms/sample) @ 2GB	Latency (ms/sample) @ 4GB	Inference Rate (samples/sec) @ 2GB	Inference Rate (samples/sec) @ 4GB	Accuracy (VM)	Precision (VM)	Recall (VM)	F1-Score (VM)
AdaBoost	0.00832	0.01527	120,259.19	65,499.16	0.911	0.831	0.911	0.869
CatBoost	0.00285	0.03507	350,698.52	233,879.89	0.986	0.984	0.986	0.983
GBM	0.00278	0.00784	359,661.13	127,605.90	0.990	0.987	0.990	0.988
LGBM	0.01012	0.01331	98,862.06	75,117.20	0.988	0.984	0.988	0.986
XGB	0.00662	0.01962	151,137.13	50,967.51	0.986	0.982	0.986	0.983

In the 2-Class (Binary) task (Table 7), GBM emerges as the absolute performance winner. GBM achieves a latency of only 0.00147 ms under the strict 2 GB RAM constraint, which further drops to an exceptional 0.00127 ms per sample when RAM is expanded to 4 GB. This translates to inference throughputs of 680,341.07 samples/sec and 788,468.18 samples/sec, respectively, while still yielding the highest accuracy of 0.999. This indicates that while GBM suffers from massive sequential offline training latency (66,248.01 ms, as shown in Table 5), its serialized execution path during active deployment is mathematically lightweight and highly optimized for single-threaded CPUs. An intriguing finding is identified when examining XGB and AdaBoost models. In the 2-Class scenario, expanding VM memory from 2 GB to 4 GB paradoxically increases latency and reduces throughput for both algorithms (e.g., XGB's throughput drops from 226,971.27 to 81,914.43 samples/sec, and AdaBoost drops from 156,831.70 to 81,810.71 samples/sec). This paradoxical drop in performance when RAM is increased cannot be explained solely by basic virtualization overhead. Instead, it suggests a deeper issue with how the operating system manages larger memory spaces on a severely restricted 1-core CPU. With more RAM, the virtual system might spend more time searching,

allocating, or cleaning up memory spaces than executing model predictions, creating a bottleneck.

For the 8-Class task (Table 8), the complexity of multiclass decision boundaries (which process multiple class trees per packet) heavily constrains the single-core CPU bandwidth. CatBoost proves highly resilient under 2 GB of RAM, achieving a latency of 0.00285 ms (350,698.52 samples/sec throughput) and an F1-Score of 0.983. Similarly, under the 4 GB RAM configuration, almost all models experience this same memory-scaling paradox. For example, GBM's throughput drops to 127,605.90 samples/sec, and XGB drops to 50,967.51 samples/sec. This confirms that inside a limited 1-core environment, blindly adding larger memory allocation hinders rather than helps the processing speed. These results emphasize that memory bandwidth limits and cache management are highly critical factors in single-core IoT edge routers, making the selection of lightweight tree structures a primary design priority.

Table 9. Per-Class Performance Profiling of the Leading GBM Model (8-Class Task under 2 GB RAM Constraint)

Attack Class	Precision	Recall	F1-Score	Support (Instances)
Benign	0.9986	1.0000	0.9993	728
DDoS	1.0000	1.0000	1.0000	49
DoS	0.8077	0.9130	0.8571	23
Mirai	1.0000	1.0000	1.0000	184
Recon	0.7143	0.8333	0.7692	6
Spoofing	0.6000	0.5000	0.5455	6
Web	0.0000	0.0000	0.0000	1
BruteForce	0.0000	0.0000	0.0000	3
Weighted Average	0.9865	0.9900	0.9881	1000

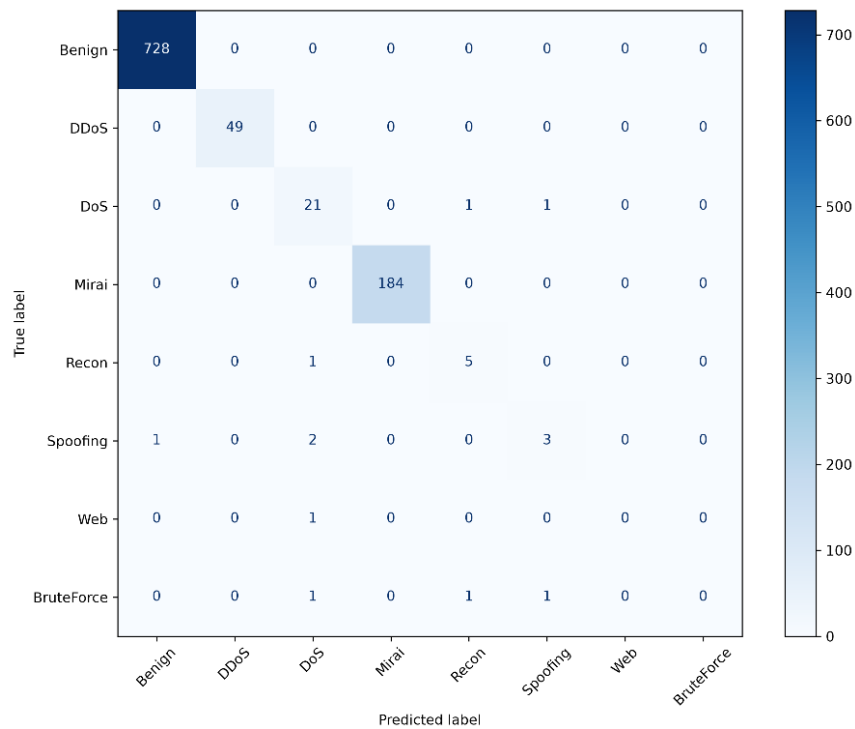


Figure 4. Confusion Matrix of the GBM Model (8-Class Task) under Edge VM Emulation, 2 GB RAM

To address classification robustness and provide a granular view of the leading estimator, we extracted comprehensive per-class profiling for the GBM model on the 8-class task during active VM deployment with the highly restricted 2 GB RAM profile, as detailed in Table 9. While the aggregate weighted F1-score remains exceptionally high at 0.988, per-category metrics reveal the model's behavior under severe minority-class constraints.

As observed, GBM achieves perfect and near-perfect classification resilience for high-volume network events, securing an F1-score of 1.0000 for DDoS and Mirai attacks, and 0.9993 for Benign traffic. However, a notable degradation is evident in the extreme minority classes, specifically Web and BruteForce attacks, which yielded an F1-score of 0.000. This drop is mathematically tied to the critically low support sizes in the unbalanced testing subset (only 1 and 3 instances, respectively). This profiling is crucial for edge deployment: it demonstrates that while tree-based boosting handles macro-level network anomalies flawlessly under 1-core constraints, detecting micro-anomalies in unsampled, real-world traffic requires localized threshold tuning. To visually corroborate these classification boundaries and empirical misclassifications, Figure 4 presents the high-resolution Confusion Matrix captured directly from the 2 GB RAM Ubuntu VM inference phase.

5. CONCLUSION

This research conducted a rigorous, deployment-level empirical evaluation of tree-based boosting frameworks (GBM, XGB, LGBM, CatBoost, and AdaBoost) to secure resource-constrained IoT environments. By deploying an ablation methodology comprising Pure, Up, Down, and Hybrid sampling strictly after the train-test split, the study eliminated data-leakage vulnerabilities. Statistical validation ($p < 0.05$) confirmed that GBM yields peak performance under restricted edge-deployment conditions, achieving 99.9% accuracy in binary classification and 99.0% in multi-class classification during VM emulation. Edge-emulation profiles identified that sequentially structured frameworks like GBM provide the highest, most stable inference throughput inside severely memory-restricted (2 GB RAM) single-core architectures, handling over 680,000 samples per second.

A primary limitation of this study is the reliance on virtualized software hardware emulation (Ubuntu VM). While VM profiles accurately restrict memory and CPU allocation, they abstract away specific thermal throttling and physical instruction-set bottlenecks present in bare-metal microcontrollers (e.g., Raspberry Pi or ESP32). Future work will prioritize deploying these joblib artifacts directly onto physical Edge TPUs and examining power-consumption profiles. Furthermore, the memory-scaling paradox observed during the VM emulation—where increasing RAM from 2 GB to 4 GB unexpectedly reduced inference speed—remains an open issue. Because this study used a software-based VM, this anomaly may stem from how the virtual OS manages larger RAM allocations. Future research should deploy these models directly on physical, bare-metal IoT devices to confirm whether this memory bottleneck stems from the machine learning algorithm itself or is merely a side effect of the virtualization software.

DATA AVAILABILITY STATEMENT

The CICIoT2023 dataset supporting this research is publicly accessible via the Canadian Institute for Cybersecurity (CIC) repository. Analyzed subsets are available upon reasonable request.

ACKNOWLEDGEMENT

This research is funded by the Ministry of Education, Culture, Research, and Technology of Indonesia under grant numbers 003/SP2H/RT-MONO/LL4/2023 and 350/PNLT2/PPM/2023. The authors thank several anonymous reviewers and colleagues who shared their thoughts to improve the quality of this paper.

REFERENCES

- [1] J. C. Cano, V. Berrios, B. Garcia and C. K. Toh, "Evolution of IoT: An Industry Perspective," in *IEEE Internet of Things Magazine*, vol. 1, no. 2, pp. 12-17, December 2018, doi: 10.1109/IOTM.2019.1900002.
- [2] A. Imteaj, U. Thakker, S. Wang, J. Li and M. H. Amini, "A Survey on Federated Learning for Resource-Constrained IoT Devices," in *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 1-24, 1 Jan. 2022, doi: 10.1109/JIOT.2021.3095077.
- [3] R. Ande, B. Adebisi, M. Hammoudeh, and J. Saleem, "Internet of Things: Evolution and technologies from a security perspective", *Sustainable Cities and Society*, vol. 54, p. 101728, 2020. doi: 10.1016/j.scs.2019.101728.
- [4] M. Almiani, A. AbuGhazleh, A. Al-Rahayfeh, S. Atiewi, and A. Razaque, "Deep recurrent neural network for IoT intrusion detection system", *Simulation Modelling Practice and Theory*, vol. 101, p. 102031, 2020. doi: 10.1016/j.simpat.2019.102031.
- [5] K. Ponnusamy and N. Rajagopalan, "Internet of Things: A Survey on IoT Protocol Standards", in *Progress in Advanced Computing and Intelligent Engineering*, 2018, pp. 651–663. doi: 10.1007/978-981-10-6875-1_64
- [6] S. Sathyadevan, K. Achuthan, R. Doss and L. Pan, "Protean Authentication Scheme – A Time-Bound Dynamic KeyGen Authentication Technique for IoT Edge Nodes in Outdoor Deployments," in *IEEE Access*, vol. 7, pp. 92419-92435, 2019, doi: 10.1109/ACCESS.2019.2927818.
- [7] W. E. Mohammed Aziz Al Kabir and M. S. Sharif, "Securing IoT Devices Against Emerging Security Threats: Challenges and Mitigation Techniques", *Journal of Cyber Security Technology*, vol. 7, no. 4, pp. 199–223, 2023. doi: 10.1080/23742917.2023.2228053.
- [8] P. Mishra, V. Varadharajan, U. Tupakula and E. S. Pilli, "A Detailed Investigation and Analysis of Using Machine Learning Techniques for Intrusion Detection," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 686-728, Firstquarter 2019, doi:10.1109/COMST.2018.2847722.
- [9] H. Binder, O. Gefeller, M. Schmid, and A. Mayr, "The evolution of boosting algorithms", *Methods of Information in Medicine*, vol. 53, no. 06, pp. 419–427, 2014. doi: 10.3414/ME13-01-0122.
- [10] J. L. Leevy, J. Hancock, R. Zuech and T. M. Khoshgoftaar, "Detecting Cybersecurity Attacks Using Different Network Features with LightGBM and XGBoost Learners," 2020 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI), Atlanta, GA, USA, 2020, pp. 190-197, doi: 10.1109/CogMI50398.2020.00032.
- [11] E. Ashraf, N. F. F. Areed, H. Salem, E. H. Abdelhay, and A. Farouk, "FIDChain: Federated Intrusion Detection System for Blockchain-Enabled IoT Healthcare Applications", *Healthcare*, vol. 10, no. 6, 2022. doi: 10.3390/healthcare10061110.
- [12] S. Garg, V. Kumar, and S. R. Payyavula, "Identification of internet of things (IoT) attacks using gradient boosting: a cross dataset approach," *Telematique*, vol. 21, no. 1, pp. 6982–7012, 2022. Available: <https://www.provinciajournal.com/index.php/telematique/article/view/965>
- [13] M. M. Khan and M. Alkhathami, "Anomaly detection in IoT-based healthcare: machine learning for enhanced security", *Scientific Reports*, vol. 14, no. 1, p. 5872, Mar. 2024. doi: 10.1038/s41598-024-56126-x.

- [14] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment", *Sensors*, vol. 23, no. 13, 2023. doi: 10.3390/s23135941.
- [15] M. M. Ahsan, M. A. P. Mahmud, P. K. Saha, K. D. Gupta, and Z. Siddique, "Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance", *Technologies*, vol. 9, no. 3, 2021. doi: 10.3390/technologies9030052.
- [16] M. Bach, A. Werner, and M. Palt, "The Proposal of Undersampling Method for Learning from Imbalanced Datasets", *Procedia Computer Science*, vol. 159, pp. 125–134, 2019. doi: 10.1016/j.procs.2019.09.167.
- [17] A. R. Hakim, K. Ramli, M. Salman, and E. R. Agustina, "Improving Model Performance for Predicting Exfiltration Attacks Through Resampling Strategies", *IIUM EJ*, vol. 26, no. 1, pp. 420–436, Jan. 2025.
- [18] M. Saied, S. Guirguis, and M. Madbouly, "A Comparative Study of Using Boosting-Based Machine Learning Algorithms for IoT Network Intrusion Detection", *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, p. 177, Nov. 2023. doi:10.1007/s44196-023-00355-x.
- [19] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Belmont, CA: Wadsworth, 1984.
- [20] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY: Springer, 2009.

A Predictive Energy Management Framework for Malaysian Residential Microgrid Using LSTM-Driven Adaptive Rule-Based Control

KAMAL AKASHAH CHE KAMARUDDIN, SITI HAJAR YUSOFF*,
TEDDY SURYA GUNAWAN, SARAH YASMIN MOHAMAD,
AISHA HASSAN ABDALLA HASHIM, NUR SHAHIDA MIDI

*Department of Electrical and Computer Engineering, International Islamic University Malaysia,
Kuala Lumpur, Malaysia*

**Corresponding author: siti Yusoff@iium.edu.my*

(Received: 17 July 2026; Accepted: 19 August 2026; Published online: 11 September 2026)

ABSTRACT: The rapid increase in global energy demand and the growing emphasis on environmental sustainability have accelerated the integration of renewable energy sources (RES) into modern microgrids. However, the stochastic nature of RES, the unpredictable behavior of load demand, and the variability of utility electricity tariffs under the Time of Use (ToU) pricing mechanism increase the operational complexity of microgrid energy management systems (EMS). To address these challenges, this research focuses on developing and implementing a hybrid EMS that integrates machine learning (ML)-based Long Short-Term Memory (LSTM) predictive models with conventional Rule-based Control (RBC) strategies for tertiary-level control of a centralized, grid-connected residential microgrid. The proposed EMS approach combines one-hour-ahead forecasts of solar Photovoltaic (PV) energy generation and residential load demand with predefined Load Following (LF) and Cycle Charging (CC) control strategies, enabling the EMS to anticipate future operating conditions while preserving the simplicity of the conventional RBC framework. The main objective of this research is to evaluate whether integrating LSTM predictive models into conventional RBC strategies can reduce the total operating cost of a Malaysian residential microgrid under the applicable ToU electricity tariff and Solar Accelerated Transition Action Program (Solar ATAP) scheme. The microgrid is modeled in MATLAB Simulink as a grid-connected system comprising a solar PV system, battery energy storage system (BESS), and three residential load profiles representing terrace, apartment, and condominium households in Kuala Lumpur, Malaysia. The data used to develop the proposed hybrid EMS are sourced from various trusted platforms, including Tenaga Nasional Berhad (TNB) energy smart meters, Huawei FusionSolar, and Solcast. The simulation results showed that the developed LSTM predictive models achieved higher forecasting accuracy for both solar PV energy generation and residential load demand than the baseline Feedforward Neural Network (FFNN) models. Leveraging these forecasts, the proposed LSTM-RBC-based EMS effectively coordinated the distributed energy resources (DERs) under the ToU electricity tariff and Solar ATAP scheme by strategically scheduling BESS charging and discharging operations and managing grid energy import and export based on economic operating conditions. The proposed hybrid EMS achieved a significant reduction in total operating cost by approximately 30% to 40% compared with conventional RBC strategies under different initial BESS State of Charge (SoC) scenarios. These results demonstrate that integrating ML-based LSTM predictive models into conventional RBC strategies enables more economical microgrid operation, highlighting the potential of predictive control to improve the operational performance of residential microgrids.

ABSTRAK: Peningkatan pesat dalam permintaan tenaga global serta penekanan yang semakin meningkat terhadap kelestarian alam sekitar telah mempercepat integrasi sumber tenaga boleh baharu (RES) ke dalam mikrogrid moden. Walau bagaimanapun, sifat stokastik RES, ketidakpastian tingkah laku permintaan beban, serta perubahan tarif elektrik utiliti di bawah mekanisme harga Masa Penggunaan (ToU) telah meningkatkan kerumitan operasi sistem pengurusan tenaga (EMS) mikrogrid. Bagi menangani cabaran ini, kajian ini memberi tumpuan kepada pembangunan dan pelaksanaan EMS hibrid yang mengintegrasikan model ramalan Ingatan Jangka Pendek Berpanjangan (LSTM) berasaskan pembelajaran mesin (ML) dengan strategi konvensional Kawalan Berasaskan Peraturan (RBC) bagi kawalan peringkat tertier untuk mikrogrid kediaman berpusat yang disambungkan kepada grid. Pendekatan EMS ini menggabungkan ramalan satu jam ke hadapan bagi penjanaan tenaga solar foto volta (PV) dan permintaan beban kediaman dengan strategi kawalan Pengikutan Beban (LF) dan Kitar Pengecasan (CC) yang telah ditetapkan, bagi membolehkan EMS menjangka keadaan operasi pada masa hadapan sambil mengekalkan kesederhanaan kerangka RBC konvensional. Objektif utama kajian ini adalah bagi menilai sama ada integrasi model ramalan LSTM ke dalam strategi RBC konvensional dapat mengurangkan jumlah kos operasi mikrogrid kediaman di Malaysia di bawah tarif elektrik ToU yang berkuat kuasa dan skim Program Solar Accelerated Transition Action Programme (Solar ATAP). Mikrogrid tersebut dimodelkan dalam MATLAB Simulink sebagai sistem yang disambung kepada grid dan terdiri daripada sistem solar PV, sistem penyimpanan tenaga bateri (BESS), serta tiga profil beban kediaman yang mewakili rumah teres, apartmen, dan kondominium di Kuala Lumpur, Malaysia. Data yang digunakan dalam pembangunan EMS hibrid ini diperoleh dari beberapa platform yang dipercayai, termasuk meter pintar tenaga Tenaga Nasional Berhad (TNB), Huawei FusionSolar, dan Solcast. Hasil simulasi menunjukkan bahawa model ramalan LSTM yang dibangunkan mencapai ketepatan ramalan lebih tinggi bagi kedua-dua penjanaan tenaga solar PV dan permintaan beban kediaman berbanding model asas Rangkaian Neural Suap ke Depan (FFNN). Melalui ramalan ini, EMS berasaskan LSTM-RBC berjaya menyelaras sumber tenaga teragih (DERs) dengan berkesan di bawah tarif elektrik ToU dan skim Program Solar ATAP melalui penjadualan strategik operasi pengecasan dan nyahcas BESS, serta pengurusan import dan eksport tenaga grid berdasarkan keadaan operasi yang lebih ekonomik. EMS hibrid ini mencapai pengurangan ketara dalam jumlah kos operasi, iaitu kira-kira 30% hingga 40% berbanding strategi RBC konvensional di bawah senario keadaan cas awal (SoC) BESS berbeza. Dapatan ini menunjukkan bahawa pengintegrasian model ramalan LSTM berasaskan ML ke dalam strategi RBC konvensional membolehkan operasi mikrogrid yang lebih ekonomik, sekaligus menonjolkan potensi kawalan ramalan dalam meningkatkan prestasi operasi mikrogrid kediaman.

KEYWORDS: *Energy Management Systems (EMS), Long Short-Term Memory (LSTM), Machine Learning (ML), Rule-Based Control (RBC), Tertiary-Level Control*

1. INTRODUCTION

A microgrid, as the term suggests, is a localized or autonomous electrical network that can produce, consume, store, and control energy flow within its boundaries. It typically involves multiple distributed energy resources (DERs), including both dispatchable and non-dispatchable units such as solar Photovoltaic (PV), wind turbines, battery energy storage systems (BESS), diesel generators (DG), and fuel cells. Integrating these energy resources improves the microgrid's flexibility to adapt quickly to changes in generation, storage, and demand, while also increasing reliability in maintaining power supply during grid outages. However, high penetration of renewable energy sources (RES) in modern microgrids introduces significant operational challenges, primarily because of their inherent stochastic and intermittent nature. RES depend heavily on weather conditions, which can result in inconsistent

power generation and system instability. For instance, solar PV systems rely heavily on sunlight availability, and wind turbine output fluctuates with varying wind speed.

Microgrids are dynamic in nature since they have the capability of operating in multiple modes, including islanded, grid-connected, or shut down [1]. Microgrids can seamlessly connect and disconnect from the main grid at any time. In islanded mode, the microgrid supplies power to the local consumers without being connected to the main grid, whereas in grid-connected mode, it operates in synchronization with the main grid, enabling power exchange and enhanced reliability. Shut-down mode is normally used for maintenance, fault conditions, or system protection purposes. Microgrids can operate across a wide range of voltage levels, from low-voltage systems (230V single-phase and 400V three-phase) to medium voltage levels such as 11kV and 33kV. In Malaysia, microgrids operate using the same regulated voltage standards applied to the national distribution network to ensure compliance with the Malaysian Grid Code, Distribution Code, and MS IEC60038 voltage standard [2].

Solar PV systems are currently the fastest-growing RES in Malaysia, with a growth rate expectation of 13.46% by 2035 [3]. The rapid expansion of solar PV is due to the country's strategic equatorial location and consistent year-round climate, which provide abundant solar irradiance. Moreover, government-led initiatives such as the Feed-in Tariff (FiT), Net Energy Metering (NEM), and the Solar Accelerated Transition Action Program (Solar ATAP) have further encouraged consumers to adopt solar PV systems. However, the increasing integration of solar PV systems in Malaysia may lead to system instability due to intermittency and uncertainty of energy generation, thereby necessitating more advanced energy management systems (EMS).

In modern microgrid development, EMS is regarded as the core component, responsible for coordinating energy resources and making optimal control decisions. Optimal control decisions typically focus on minimizing operating costs, maximizing power generation, extending the lifespan of the energy storage system (ESS), and ensuring a consistent electricity supply to consumers. According to [4], the essential aspects of microgrid control that must be addressed are: load tracking, real and reactive power control, voltage regulation, frequency regulation, load sharing, and protection.

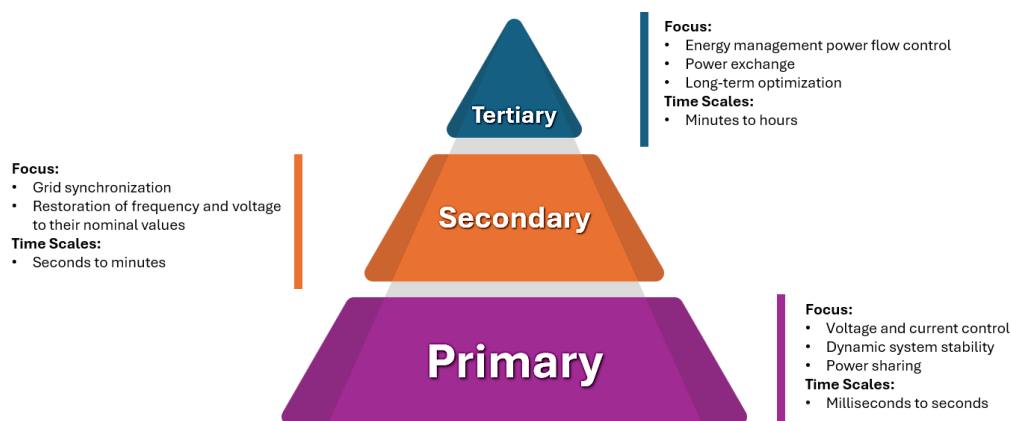


Figure 1. Hierarchical control levels of microgrids

Microgrid EMS has three hierarchical control levels: primary, secondary, and tertiary control, as shown in **Figure 1**. Each control level features different functions, objectives, and time scales [5]. Primary-level control serves as the basis of hierarchical control in microgrid EMS and is important for fast voltage and frequency stabilization, as well as effective power sharing among the DERs in the microgrid. This control layer is implemented locally at the

inverter or DER level and operates on time scales of milliseconds to seconds. Secondary-level control, on the other hand, focuses on restoring voltage and frequency to their nominal values and improves grid synchronization during transitions between microgrid operation modes. This control layer operates at slower time scales, between seconds and minutes, depending on the communication infrastructure. Finally, tertiary-level control is the highest level in hierarchical microgrid control. This control layer focuses on the long-term optimization of the microgrid's EMS in terms of economic dispatch, power scheduling, system constraints, or power exchanges between microgrids and external networks. Tertiary-level control generally works on time scales ranging from several minutes to hours.

While hierarchical control levels describe the focus and time-scale separation of microgrid EMS, practical implementation can use different control structures, such as centralized, decentralized, or distributed approaches. In a centralized control structure, a single central EMS controller collects all system information from the DERs and makes operational decisions for the entire microgrid. This structure provides strong system-wide coordination and can deliver high optimization performance, as the controller has full visibility of the microgrid. It can be advantageous for microgrids with small resources and relatively low demand. However, this structure scales poorly as the system expands and is highly vulnerable to a single point of failure, where a central controller failure may disrupt the entire EMS operation. In contrast, a decentralized control structure allows each subsystem to operate with its own local controller, thereby enabling independent decision-making based on local objectives. Compared to the centralized control structure, this structure offers improved reliability and scalability. However, it often results in suboptimal performance due to the lack of global coordination. Meanwhile, the distributed control structure allows multiple controllers to communicate and cooperate to achieve common control objectives without relying entirely on a single central controller. Compared to the centralized control structure, this structure offers greater scalability, while also providing better coordination than the decentralized approach, despite being more complex to implement [6]. **Figure 2** summarizes the microgrids classification based on operational modes, control levels, control structures, and configurations.

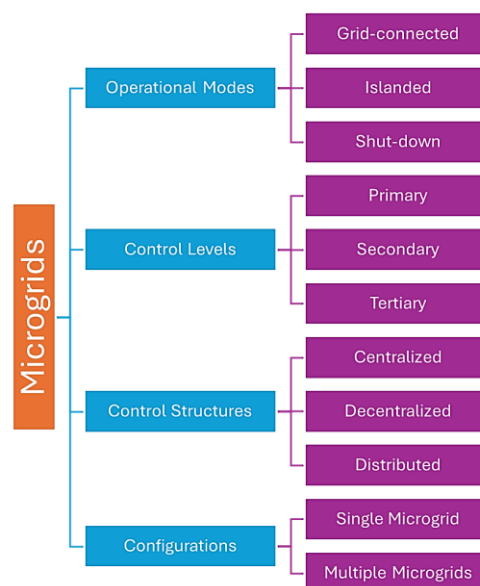


Figure 2. Microgrids classification based on operational modes, control levels, control structures, and configurations

Numerous methods have been proposed for tertiary-level microgrid EMS in the literature. Among these, heuristic Rule-based Control (RBC) is one of the most straightforward and

widely adopted approaches, which operates based on predefined “if-then” rules that make it simple to design, implement, and interpret [7]. The study in [8] implemented a centralized RBC-based EMS strategy to examine the techno-economic and environmental performance of an isolated microgrid comprised of a solar PV system, DG, and BESS. The study compared and evaluated the performance of three distinct RBC control strategies known as Load Following (LF), Cycle Charging (CC), and Combined Dispatch (CD), with their performance metrics assessed based on the Net Present Cost (NPC), Cost of Energy (COE), and Carbon Dioxide (CO₂) emissions. The results showed that CD achieved the lowest NPC, COE, and CO₂ emissions compared to LF and CC. On the other hand, the research in [9] focused on optimizing the operational cost of a grid-connected microgrid consisting of a solar PV system and BESS using a centralized RBC-based EMS strategy. The proposed EMS considered Time of Use (ToU) electricity pricing and the maximum sanctioned limits for power imported from the grid. This research successfully optimized overall energy consumption costs and effectively managed peak demand by controlling the BESS based on real system states and ToU. The study in [10] implemented a centralized RBC-based EMS strategy for an off-grid electric vehicle (EV) charging station integrated with a solar PV system and BESS, where the EV load demand was dynamically modeled using home charging data from metropolitan Melbourne, Australia. The simulations considered three scenarios: sunny, cloudy, and nighttime conditions. The proposed EMS demonstrated stable direct-current (DC) bus operation within $\pm 5\%$, the ability to operate independently of the grid, and effective BESS operation within safe State of Charge (SoC) levels of 25% to 95%. Despite the encouraging results shown in [8]-[10], the EMS based on the RBC strategy does not guarantee optimal microgrid operation, particularly under dynamic and uncertain operating conditions, because it depends on predefined control rules.

To enhance the responsiveness of EMS, some studies have implemented a Model Predictive Control (MPC) strategy, which uses mathematical models to forecast future system states and optimize operation at each timestep [11]. The work in [12] focused on developing a robust and practical centralized EMS based on the integrated MPC and RBC strategy for a grid-connected microgrid comprised of solar PV and BESS. MPC was employed as the first control layer to determine the optimal BESS SoC, aiming to keep the microgrid’s operation within the white zone, where system constraints are comfortably satisfied. Meanwhile, RBC served as the second control layer to ensure operational safety and constraint satisfaction. However, this strategy resulted in a slight increase in daily operating cost compared to conventional MPC. The study in [11] further improved the centralized EMS based on MPC by integrating it with an Artificial Neural Network (ANN), which is a subset of Machine Learning (ML). ANN was used to predict solar PV energy generation, whereas MPC was used to manage peak demand and optimize energy usage through load curtailment and load shifting. This hybrid EMS approach reduced grid energy usage and associated costs by up to 48% and 34% on sunny and cloudy days, respectively.

Another study in [13] combined ANN with the metaheuristic Artificial Bee Colony algorithm into a centralized EMS to achieve optimal power flow control for a home EMS in a residential microgrid connected to the main grid, solar PV, and BESS. The proposed EMS achieved significant reductions in electricity consumption cost of 9.05% during non-summer months and 29.7% during summer months compared to the RBC-only strategy. The work in [14] also employed a metaheuristic algorithm, namely Particle Swarm Optimization (PSO), integrated with the RBC strategy into a distributed EMS to optimize energy management within community-based grid-connected residential microgrids comprising solar PV and BESS. PSO was used in the first control layer to determine optimal load scheduling, while RBC was applied in the second control layer to perform peak demand shaving by regulating the BESS charging and discharging schedule. This hybrid EMS approach achieved an average improvement of

8.5% in Percentage Peak Shaving (PPS) across various PV power and load demand patterns for both winter and summer conditions.

To maintain the simplicity of RBC while enhancing its decision-making capability and adaptability, many studies have integrated ML techniques with the RBC strategy. The work in [15] compared the performance of four different ML techniques, known as Decision Tree, Random Forest, Deep ANN, and Long Short-Term Memory (LSTM), for forecasting RES generation and load consumption. These forecasting models were then integrated into a centralized RBC-based EMS framework for an islanded microgrid comprising solar PV, wind turbines, BESS, DG, and a fuel cell, focusing on improving scheduling performance through peak shaving and valley filling. The results indicated that the LSTM model performed best in forecasting RES generation and load demand, while the ML-based EMS demonstrated greater robustness to uncertainties in RES generation and load demand. Similar to [15], the work in [16] implemented a Temporal Convolutional Neural Network predictive model, which was subsequently integrated into a centralized RBC-based EMS framework. However, the proposed hybrid EMS considered a power flow algorithm that coordinated RES utilization and ToU pricing to improve the economic operation of the BESS. The study in [17] adopted a slightly different approach, in which the centralized EMS used a dynamic decision-making mechanism capable of switching between RBC and deep learning models. RBC served as a foundational method by providing a decision-making framework based on predefined logic and rules to manage various energy resources, demand responses, and microgrid components. Conversely, the deep learning models were only used for adaptive control actions, particularly when rules might be inadequate or under unexpected conditions.

This study proposes a hybrid approach that integrates an ML-based LSTM predictive model with conventional RBC strategies for tertiary-level control of a centralized, grid-connected residential microgrid, with specific consideration of the Malaysian operating context. The microgrid is configured as a grid-connected system consisting of a solar PV system, BESS, and three types of Malaysian residential loads, all modeled in MATLAB Simulink. The core objective of this study is to investigate whether the proposed hybrid EMS can reduce the total operating cost of the residential microgrid through more economical DER coordination under the Malaysian ToU electricity tariff and Solar ATAP scheme. The main contributions of this study are the development, implementation, and evaluation of the proposed LSTM-RBC-based EMS, with the details as follows:

- Development of ML-based LSTM predictive models for one-hour-ahead forecasting of solar PV energy generation and residential load demand.
- Development of forecast-informed adaptive RBC strategies that integrate one-hour-ahead LSTM predictions and upcoming ToU electricity tariff information into the decision-making process of the conventional LF and CC strategies.
- Implementation and evaluation of the proposed LSTM-RBC-based EMS using Malaysian residential load profiles, solar PV generation, weather data, and electricity tariff conditions to analyze its applicability and economic performance under the Malaysian ToU tariff and Solar ATAP scheme.
- Evaluation of the economic and operational trade-offs introduced by the proposed hybrid EMS in terms of total operating cost, grid energy exchange, BESS degradation cost, self-sufficiency ratio (SSR), PV utilization, and BESS SoC behavior under different initial BESS SoC conditions.

The rest of this paper is structured as follows: methodology, followed by results and discussion, and finally, the conclusion.

2. METHODOLOGY

2.1. General Framework of the Proposed Microgrid and Hybrid EMS

In this study, the proposed microgrid system operates as a grid-connected system comprising a solar PV system, BESS, and three Malaysian residential load profiles, as shown in Figure 3.

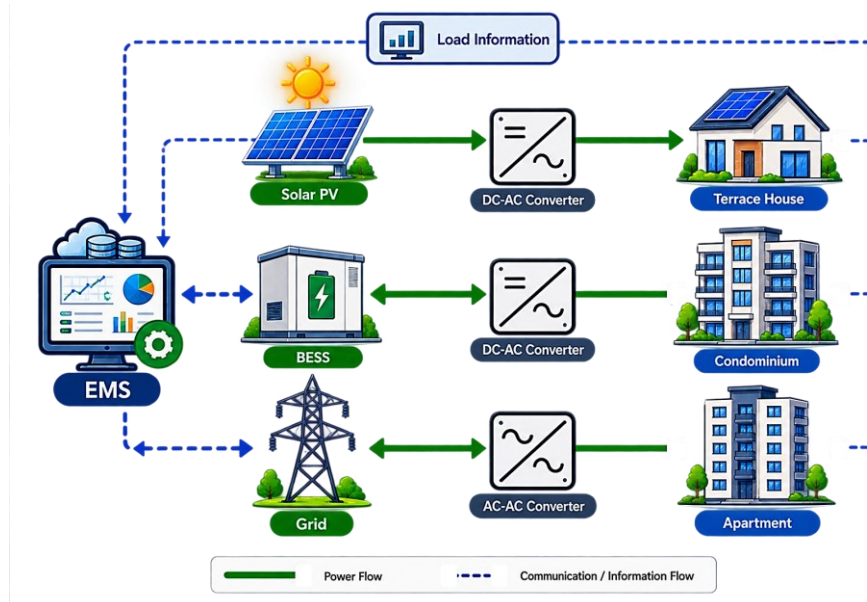


Figure 3. Proposed grid-connected microgrid system

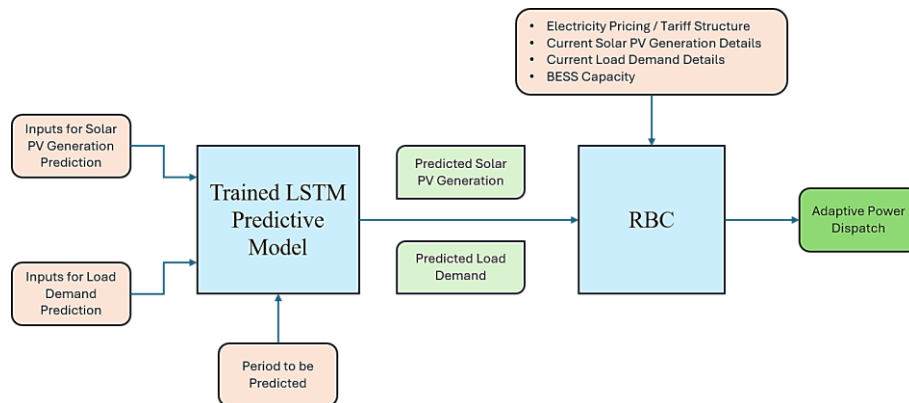


Figure 4. Proposed LSTM-RBC-based EMS framework

The proposed microgrid system prioritizes energy supply from the solar PV system, followed by the BESS, and finally the main grid. The solar PV system operates at its maximum power point to maximize solar energy generation, while the BESS stabilizes microgrid operation. For this study, the BESS is assumed to be charged only using surplus energy generated by the solar PV system, while grid charging is not considered. Moreover, the BESS is not intentionally discharged to export stored energy to the main grid and capture economic benefits from periods of higher electricity prices. These BESS operating assumptions emphasize self-consumption and the effective utilization of locally generated renewable energy. The main grid serves as a backup energy source for the microgrid only when the combined available power from the solar PV system and BESS is insufficient to meet the

residential load demand. **Figure 3** also illustrates the communication flow between the centralized EMS and the solar PV system, BESS, main grid, and residential loads.

Figure 4 shows the proposed LSTM-RBC-based EMS framework in this study, featuring an LSTM predictive model and RBC in a cascaded configuration. The proposed EMS begins by forecasting solar PV energy generation and load demand using the trained LSTM predictive models. The RBC then uses these predictions, along with required local information, to provide adaptive power-dispatch decisions for the dispatchable DERs.

2.2. Data Collection

This study considers four types of data: weather data, load demand data, solar PV energy generation data, and energy price data. All datasets were sourced from reputable sources to ensure data quality, reliability, and accuracy, which are essential for developing a robust LSTM predictive model and an efficient EMS.

2.2.1. Weather Data

Weather data is an essential component of the dataset, providing valuable insights into the environmental factors that influence solar PV energy generation. This dataset was obtained from Solcast, a data service platform that provides up to 20 years of historical data. In this study, the weather dataset comprises 5415 hourly observations recorded between August 2025 and March 2026. Each observation contains five weather variables: air temperature, global horizontal irradiance (GHI), water precipitation, water precipitation rate, and relative humidity, resulting in a total of 27075 samples. The weather dataset corresponds to a specific geographical location at coordinates (3.1864, 101.7059), near Kuala Lumpur, Malaysia, as shown in **Figure 5**.



Figure 5. Location coordinates (3.1864, 101.7059) in Kuala Lumpur, Malaysia

2.2.2. Solar PV Energy Generation Data

Solar PV energy generation data was collected via the Huawei FusionSolar monitoring platform from an in-house solar PV system installed within the same geographical area, as illustrated in **Figure 5**. The solar PV system consists of nine 620 W Trina TSM-NEG19R.20-620 solar panels with a total capacity of 5.58 kWp and a 5 kWac Huawei SUN2000-5KTL-L-1 inverter, with detailed specifications provided in **Table 1**.

A total of 5415 hourly solar PV energy generation samples were obtained for the period spanning from August 2025 to March 2026, as depicted in **Figure 6**. **Figure 7** presents the average hourly solar PV energy generation, with peak sun hours typically occurring for 5 hours between 10 am and 3 pm at the selected location.

Table 1. Solar PV systems parameters

PV Panel TSM-NEG19R.20-620 @ Standard Test Condition (STC)	
Peak Power (P_{max})	620 W
Maximum Power Voltage (V_{mpp})	41.4 V
Maximum Power Current (I_{mpp})	14.99 A
Open Circuit Voltage (V_{oc})	49.6 V
Short Circuit Current (I_{sc})	15.91 A
Module Efficiency (η_{pv})	23.0 %
Inverter SUN2000-5KTL-L-1	
Maximum PV power	7500 Wp
Maximum Input Voltage	600 V
Startup Voltage	100 V
Maximum Power Point Tracking (MPPT) Operating Range	90 V – 560 V
Rated Input Voltage	360 V
Maximum Input Current per MPPT	12.5 A
Maximum I_{sc}	18 A
Inverter Efficiency (η_{inv})	98.4 %

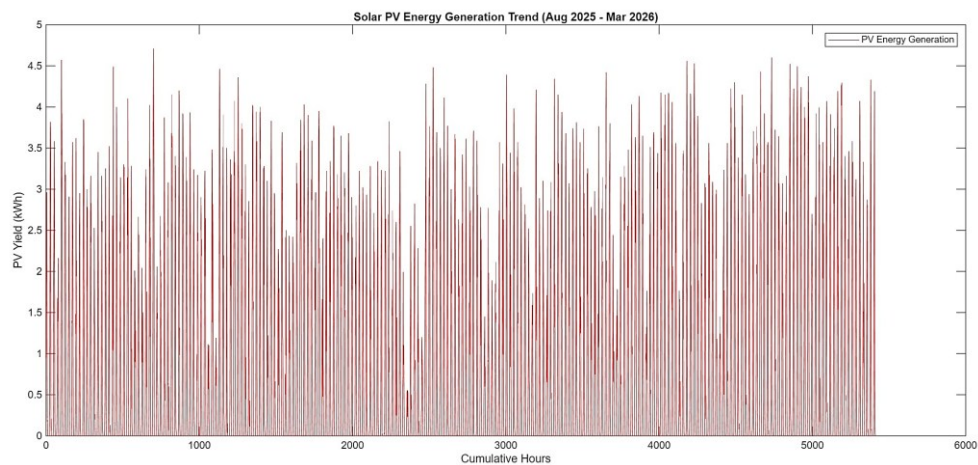


Figure 6. Solar PV energy generation trend between August 2025 and March 2026

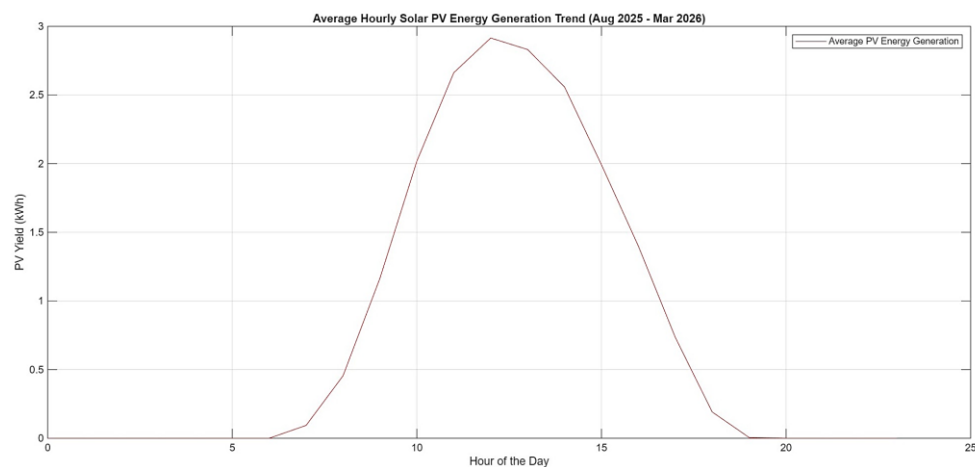


Figure 7. Average hourly solar PV energy generation trend between August 2025 and March 2026

2.2.3. Load Demand Data

This study considers three types of Malaysian residential load profiles located within the same geographical area, as shown in **Figure 5**, namely terrace house, apartment, and condominium, using the data sourced from Tenaga Nasional Berhad (TNB)’s energy smart meter platform. This platform facilitates the analysis of the historical hourly load demand patterns. A total of 32760 samples were collected, spanning from January 2025 to March 2026, comprising 10920 samples for each residential type. Considering different residential load types is important to improve the LSTM predictive model's generalization across different housing scenarios. **Figure 8** shows the Malaysian residential load demand trends, with average electricity consumption of 0.67 kWh for the terrace house, 0.43 kWh for the apartment, and 0.37 kWh for the condominium.

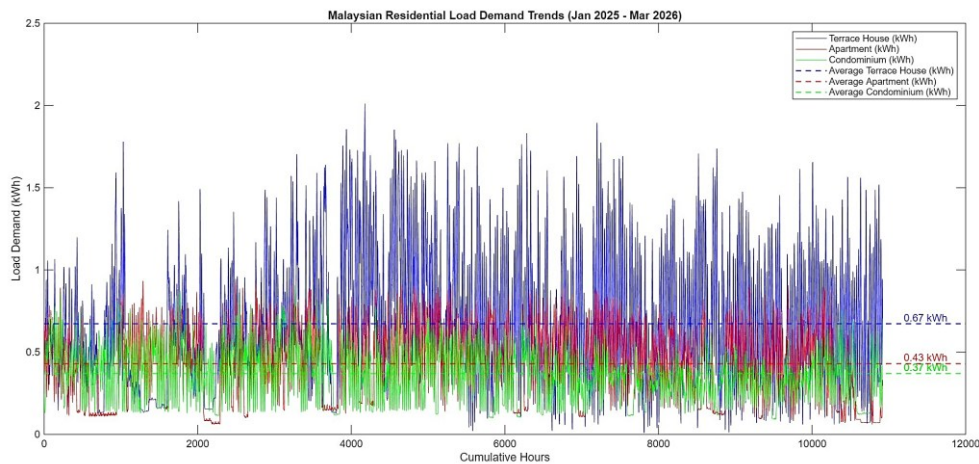


Figure 8. Malaysian residential load profiles for three different types of Malaysian residential houses between January 2025 and March 2026

2.2.4. Grid Energy Pricing based on ToU λ_{ToU}

In Malaysia, TNB introduced the ToU electricity tariff starting in July 2025 as part of the government’s electricity tariff restructuring, which provides consumers with differentiated rates based on peak and off-peak hours. The peak tariff applies only on weekdays from 2 pm to 10 pm, while the off-peak tariff applies from 10 pm to 2 pm on weekdays and throughout the day on weekends and public holidays. **Table 2** shows the ToU electricity tariff applied to domestic consumers in Malaysia.

Table 2. Malaysian domestic ToU electricity tariff as of July 2025 [18]

Domestic ToU Electricity Tariff		
For total consumption of 1500 kWh and below per month		
Peak Energy Charge, C_{peak}	RM/kWh	0.2852
Off-peak Energy Charge, $C_{offpeak}$	RM/kWh	0.2443
Capacity Charge, $C_{capacity}$	RM/kWh	0.0455
Network Charge, $C_{network}$	RM/kWh	0.1285
Retail Charge, C_{retail}	RM/month	10.00
For total consumption of more than 1500 kWh per month		
Peak Energy Charge, C_{peak}	RM/kWh	0.3852
Off-peak Energy Charge, $C_{offpeak}$	RM/kWh	0.3443
Capacity Charge, $C_{capacity}$	RM/kWh	0.0455
Network Charge, $C_{network}$	RM/kWh	0.1285
Retail Charge, C_{retail}	RM/month	10.00

2.2.5. Solar ATAP Export Energy Offset Rate λ_{ATAP}

Solar ATAP is a rooftop solar scheme introduced by the Malaysian government in January 2026 to further accelerate the adoption of solar PV systems. The scheme serves as a successor to the NEM while preserving its core principles. Under the NEM scheme, excess energy generated by the solar PV system is credited on a one-to-one energy-offset basis, whereas Solar ATAP implements an energy-offset mechanism based on the applicable electricity tariff. According to [19], the energy exported to the main grid under Solar ATAP is offset against the bill at a rate ranging from RM 0.27 to RM 0.37 per kWh for the residential consumers. In this study, the export energy offset rate is assumed to be RM 0.37 per kWh.

2.3. Data Preprocessing

All previously collected datasets were preprocessed in MATLAB before developing LSTM predictive models. The data preprocessing procedures consisted of feature selection, data cleaning, chronological data splitting, and data normalization. Feature selection was performed first to identify the input variables that significantly influence the LSTM models' forecasting performance. For the solar PV energy generation prediction model, the selected input features include:

- Hour of the day
- Month of the year
- Air temperature
- GHI
- Water precipitation
- Water precipitation rate
- Relative humidity
- Current PV yield
- PV yield recorded at 1-hour intervals over the previous 24 hours

For the residential load demand prediction model, the selected input features comprise:

- Hour of the day
- Month of the year
- Day of the week
- Type of day
- Type of house
- Current load demand
- Load demand recorded at 1-hour intervals over the previous 24 hours

Subsequently, the datasets were cleaned using the MATLAB Data cleaner, which involved outlier detection and removal through Z-score filtering, followed by outlier replacement and missing-value handling using linear interpolation, and finally a causal data-smoothing procedure to reduce measurement noise. Next, the datasets were arranged chronologically and split into training, validation, and testing subsets using a 70:20:10 ratio. This chronological splitting strategy ensured that the training, validation, and testing datasets followed the natural temporal sequence of the data, thereby preventing temporal data leakage from future observations into the model training process. Finally, the data normalization process was performed using the min-max technique, as expressed in Eq. (1), to ensure consistent input conditions for the LSTM predictive model.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where x_{norm} is the normalized data, x is the data value, x_{min} is the minimum data value, and x_{max} is the maximum data value.

2.4. LSTM Predictive Modeling

Among the various ANN-based predictive models, the LSTM predictive model has been selected for forecasting solar PV energy generation and load demand due to its effectiveness in capturing both short- and long-term temporal dependencies, as demonstrated in the previous study [15]. LSTM, a variant of Recurrent Neural Network (RNN), can model sequential data while mitigating the vanishing and exploding gradient problems that commonly occur in conventional RNNs. **Figure 9** shows the structure of a single LSTM cell that is controlled by three main gates known as the Forget gate, Input gate, and Output gate. These gates will determine whether the information should be retained, updated, or discarded.

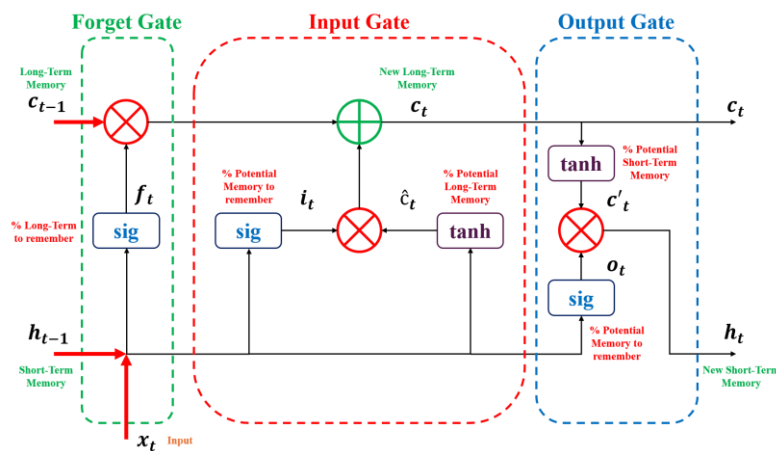


Figure 9. Structure of a single LSTM cell

According to **Figure 9**, the forget gate contains a sigmoid activation function that determines the percentage of long-term memory from the previous cell state to be remembered. This percentage can be calculated using Eq. (2).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

where σ is a sigmoid function, W_f is the weight matrix associated with the forget gate, h_{t-1} is the previous hidden state, x_t is the current input, and b_f is the forget gate bias.

Next, the input gate consists of two activation functions, sigmoid and hyperbolic tangent, used to determine the percentage of new information to be remembered and to create candidate values representing potential updates to memory, respectively. The new memory percentage can be calculated using Eq. (3), whereas the memory potential update percentage is determined according to Eq. (4). The current cell state, c_t , can be calculated as in Eq. (5) by adding the new candidate values calculated from the multiplication of i_t and $\hat{c}_t$, with the multiplication of the previous state c_{t-1} with f_t .

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\hat{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \hat{c}_t \quad (5)$$

where $\tanh$ is a hyperbolic tangent function, W_i and W_c are the weight matrices associated with the input gate for the sigmoid and hyperbolic tangent functions, respectively, and b_i and

b_c are the biases of the input gate for the sigmoid and hyperbolic tangent functions, respectively.

Finally, the output gate contains two activation functions: the sigmoid and hyperbolic tangent functions. The sigmoid function is used to determine the percentage of potential memory to be the output, whereas the hyperbolic tangent function is used to scale the cell state value c_t to a normalized range between -1 and 1 before it is filtered by the sigmoid function of the output gate. The calculation of the potential memory percentage to be the output is as in Eq. (6), while the scaling of the cell state value can be done using Eq. (7). By multiplying Eq. (6) and Eq. (7), the hidden state, h_t , can be calculated as in Eq. (8).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$c'_t = \tanh(c_t) \quad (7)$$

$$h_t = o_t \cdot c'_t \quad (8)$$

where W_o and b_o are the weight matrix and bias associated with the output gate, respectively.

2.4.1. LSTM Training Settings and Performance Metrics

In this research, the LSTM predictive model for solar PV energy generation consisted of a single input layer of 32 neurons (corresponding to the 32 selected input features), a single hidden layer of 50 neurons, and a single output neuron. Similarly, for load demand prediction, the LSTM had a single input layer of 30 neurons (corresponding to the 30 selected input features), a single hidden layer of 50 neurons, and a single output neuron. For both predictive models, a hidden layer with 50 neurons was selected to provide adequate prediction capability while maintaining a simple model structure. The final output layer in both predictive models represents the predicted value for the next time step (one-hour ahead). This short-term prediction horizon is implemented to balance forecasting accuracy and model complexity, as prediction performance typically declines for longer horizons [20]. The number of training epochs was set to 2000 based on the observation that training and validation losses approached convergence within this range. Both models were optimized using the Adam optimizer with a constant learning rate of 0.01. This learning rate was selected based on the observed training behavior, as it provided stable convergence without excessive fluctuations in the training loss. A 50% dropout rate and L2 regularization were also applied to both models to strengthen regularization and reduce overfitting. Table 3 summarizes the training settings of LSTM models.

To evaluate the performance of the LSTM predictive models, regression-based error metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R-Squared) were used to quantify the discrepancy between predicted and actual values. Smaller MAE, MSE, and RMSE values indicate higher forecasting accuracy, whereas an R-Squared value closer to 100% shows a stronger correlation between predicted and actual values.

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_{actual} - Y'_{predicted}| \quad (9)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_{actual} - Y'_{predicted})^2 \quad (10)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{actual} - Y'_{predicted})^2} \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_{actual} - Y'_{predicted})^2}{\sum_{i=1}^N (Y_{actual} - Y_{mean})^2} \quad (12)$$

Table 3. Summary of LSTM training settings

	LSTM Solar PV Energy Generation Prediction	LSTM Load Demand Prediction
Layer	1 input (32 neurons) 1 hidden (50 neurons) 1 output (1 neuron)	1 input (30 neurons) 1 hidden (50 neurons) 1 output (1 neuron)
No. of Epochs	2000	2000
Optimizer	Adam	Adam
Learning Rate	0.01	0.01
Type of Learning Rate	Constant	Constant
L2 Regularization	0.01	0.0001

2.5. Problem Formulation

2.5.1. Performance Evaluation Metric

The economic performance of the proposed LSTM-RBC-based EMS is formulated by coordinating the solar PV system, BESS, and the main grid to meet residential load demand. The primary objective is to minimize the total operating cost associated with the power exchange with the main grid, BESS degradation, and the operation and maintenance (O&M) of the solar PV system under the applicable Malaysian ToU electricity tariff and Solar ATAP scheme. Therefore, the total operating cost over the considered simulation period is formulated as the sum of the main grid cost, BESS degradation cost, and solar PV O&M cost, as given in Eq. (13).

$$C_{total} = \sum_{t=0}^T C_{grid,t} + C_{bess,t} + C_{pv(O\&M),t} \quad (13)$$

where C_{total} , $C_{grid,t}$, $C_{bess,t}$, and $C_{pv(O\&M),t}$ represent the total operation cost, main grid, BESS, and solar PV O&M, respectively.

The main grid cost, $C_{grid,t}$, is determined based on the energy imported from and exported to the main grid and their corresponding electricity rates during each operating interval Δt , as expressed in Eq. (14). The first term of Eq. (14) represents the cost of electricity imported from the main grid according to the applicable ToU electricity tariff, whereas the second term represents the offset obtained from electricity exported to the main grid under the Solar ATAP scheme.

$$C_{grid,t} = (\lambda_{ToU,t} \cdot P_{GIm,t} \cdot \Delta t) \mu_{GIm,t} - (\lambda_{ATAP,t} \cdot P_{GEx,t} \cdot \Delta t) \mu_{GEx,t} \quad (14)$$

where $P_{GIm,t}$ and $P_{GEx,t}$ denote the power imported from and exported to the main grid, respectively, while $\mu_{GIm,t}$ and $\mu_{GEx,t}$ represent the binary variables associated with the grid import and export operating states. The role of these binary variables is further described in Section 2.5.2.

In this study, the applicable ToU electricity tariff, $\lambda_{ToU,t}$, is represented by the combination of the energy, capacity, and network charges, as expressed in (15). The retail charge is excluded from the calculation since it is generally imposed as a fixed monthly charge and does not vary with the hourly electricity consumption. Among the considered tariff components, only the energy charge varies according to the applicable ToU period, with the corresponding rate determined based on whether the operating interval Δt falls within the peak or off-peak period, as summarized in **Table 2**.

$$\lambda_{ToU,t} = \begin{cases} C_{peak} + C_{capacity} + C_{network}, & t \in \tau_{peak} \\ C_{offpeak} + C_{capacity} + C_{network}, & t \in \tau_{offpeak} \end{cases} \quad (15)$$

where C_{peak} and $C_{offpeak}$ denote the applicable peak and off-peak energy charges, respectively, while $C_{capacity}$ and $C_{network}$ represent the capacity and network charges. The sets τ_{peak} and $\tau_{offpeak}$ represent the operating intervals corresponding to the peak and off-peak periods, respectively.

In contrast to the ToU electricity tariff applied to the grid-imported energy, the Solar ATAP export energy offset rate, $\lambda_{ATAP,t}$, is assumed to remain constant throughout the operating period. Therefore, the same export offset rate applies to electricity exported to the main grid, regardless of whether the interval falls within the peak or off-peak period.

Following the formulation of the main grid cost function, the operating cost associated with the BESS, $C_{bess,t}$, is determined by considering the BESS degradation cost incurred during BESS discharge. As shown in Eq. (16), the BESS degradation cost is calculated as the product of the BESS degradation cost coefficient, $\lambda_{deg,t}$, and the BESS discharge power, $P_{dch,t}$, at the operating interval Δt .

$$C_{bess,t} = \lambda_{deg,t} \cdot P_{dch,t} \cdot \Delta t \quad (16)$$

The BESS degradation cost coefficient, $\lambda_{deg,t}$, represents the equivalent degradation cost incurred per unit of energy discharged from the BESS and is expressed in RM/kWh. In this study, the coefficient is estimated based on the BESS capital cost, expected cycle life, battery capacity, and battery round-trip efficiency, as formulated in Eq. (17).

$$\lambda_{deg,t} = \frac{C_{bess\ capital}}{(N_{cycle})(BESS_{capacity})(DoD)(\eta_{rte})} \quad (17)$$

where $C_{bess\ capital}$ denotes the capital cost of the BESS, N_{cycle} is the battery life cycle, $BESS_{capacity}$ denotes the battery capacity, DoD is the allowable depth of discharge, and η_{rte} represents the battery round-trip efficiency.

In addition to the main grid and BESS costs, the O&M cost of the solar PV system, $C_{pv(O\&M),t}$, is also included in the overall performance evaluation metric to account for the recurring costs required to ensure reliable operation of the solar PV system. As expressed in Eq. (18), the solar PV O&M cost is calculated as the product of the PV O&M cost coefficient, $\lambda_{pv(O\&M),t}$, and the solar PV power output, $P_{pv,t}$, over the operating interval Δt .

$$C_{pv(O\&M),t} = \lambda_{pv(O\&M),t} \cdot P_{pv,t} \cdot \Delta t \quad (18)$$

The PV O&M cost coefficient, $\lambda_{pv(O\&M),t}$, is expressed in RM/kWh and is determined by allocating the annual solar PV O&M cost over its corresponding annual solar PV energy generation, as expressed in Eq. (19).

$$\lambda_{pv(O\&M),t} = \frac{C_{O\&M\ annual}}{E_{pv\ annual}} \quad (19)$$

where $C_{O\&M\ annual}$ represents the annual solar PV O&M cost, while $E_{pv\ annual}$ denotes the corresponding annual solar PV energy generation.

2.5.2. System Constraints

Several system constraints have been considered in the EMS optimization framework to ensure reliable and stable system operation. One primary constraint is the system power balance constraint. The total power consumption must always satisfy the power balance constraint to ensure that the total power generated is sufficient to meet the residential load

demand throughout the scheduling horizon, as defined in Eq. (20). The net power contributions of the BESS and main grid are further defined in Eq. (21) and Eq. (22), respectively.

$$P_{load,t} = P_{pv,t} + P_{bess,t} + P_{grid,t} \quad (20)$$

$$P_{bess,t} = P_{dch,t}\mu_{dch,t} - P_{ch,t}\mu_{ch,t} \quad (21)$$

$$P_{grid,t} = P_{GIm,t}\mu_{GIm,t} - P_{GEx,t}\mu_{GEx,t} \quad (22)$$

where $P_{load,t}$ denotes the residential load demand, $P_{bess,t}$ is the net BESS power, $P_{grid,t}$ is the net main grid power, $P_{ch,t}$ is the BESS charging power, while $\mu_{ch,t}$ and $\mu_{dch,t}$ are binary variables representing the BESS charging and discharging operating states, respectively.

Similarly, the power exchange between the microgrid and the main grid is constrained to prevent simultaneous electricity import and export within the same operating interval. The binary variables associated with the grid import and export operating states are introduced as defined in Eq. (23).

$$\mu_{GIm,t} + \mu_{GEx,t} \leq 1, \text{ where } \mu_{GIm,t}, \mu_{GEx,t} \in \{0,1\} \quad (23)$$

Another important constraint is related to the charging and discharging operations of the BESS, which must be regulated within the allowable charging and discharging power limits to prevent battery degradation, as expressed in Eq. (24) and Eq. (25). In addition, the charging and discharging operations of the BESS must not occur simultaneously, as stated in Eq. (26).

$$0 \leq P_{ch,t} \leq P_{chMAX} \quad (24)$$

$$0 \leq P_{dch,t} \leq P_{dchMAX} \quad (25)$$

$$\mu_{ch,t} + \mu_{dch,t} \leq 1, \text{ where } \mu_{ch,t}, \mu_{dch,t} \in \{0,1\} \quad (26)$$

where P_{chMAX} and P_{dchMAX} denote the maximum allowable BESS charging and discharging power, respectively.

The SoC, which represents the percentage of available energy stored in the BESS relative to its rated capacity, is an important parameter for determining the BESS operational state. The SoC dynamics can be estimated using Eq. (27). The BESS SoC is commonly regulated within predefined minimum and maximum operating limits, as in Eq. (28), to protect the battery and preserve lifetime.

$$SoC_{t+1} = SoC_t + \frac{(P_{ch,t})(\eta_{bess})(\Delta t)}{BESS_{capacity}} \mu_{ch,t} - \frac{(P_{dch,t})(\Delta t)}{(BESS_{capacity})(\eta_{bess})} \mu_{dch,t} \quad (27)$$

$$SoC_{MIN} \leq SoC_t \leq SoC_{MAX} \quad (28)$$

where SoC_{t+1} and SoC_t denote the BESS SoC at the next and current operating intervals, respectively, η_{bess} is the BESS efficiency, while SoC_{MIN} and SoC_{MAX} are the minimum and maximum allowable BESS SoC limits, respectively.

2.5.3. Sizing of Solar PV System

The sizing of the solar PV system is estimated based on the daily load demand, available peak sun hours, and the overall solar PV system efficiency, as shown in Eq. (29).

$$PV_{capacity} = \frac{E_{daily\ load}}{(PSH)(\eta_{system})} \quad (29)$$

where $PV_{capacity}$ is the required rated capacity of the solar PV system in kWp, $E_{daily\ load}$ denotes the total daily load consumption in kWh, PSH represents the average peak sun hours, and η_{system} is the overall efficiency of the solar PV system.

Based on the average hourly solar PV energy generation trend in **Figure 7**, the selected location experiences approximately 5 peak sun hours per day. The overall solar PV system efficiency, η_{system} , is assumed to be 0.7, accounting for various system losses such as temperature effects, wiring, and inverter losses.

Once the required installed PV capacity is obtained, the number of solar PV panels can be determined using Eq. (30).

$$Quantity\ of\ PV\ panels = \frac{PV_{capacity}}{P_{max}} \quad (30)$$

where P_{max} is the peak power of the solar PV panel, based on the specifications provided in **Table 1**.

2.5.4. Sizing of BESS

The sizing of BESS involves the estimation of the required battery capacity capable of storing excess energy generated by RES and supporting the load demand during periods of insufficient generation. Proper BESS sizing is essential to ensure reliable microgrid operation, improve RES utilization, and maintain power-balance stability within the microgrid. In this study, the battery capacity is estimated based on the total daily load consumption, required number of autonomy days, battery depth of discharge, and battery efficiency, as expressed in Eq. (31).

$$BESS_{capacity} = \frac{(E_{daily\ load})(Day_{auto})}{(DoD)(\eta_{bess})} \quad (31)$$

where Day_{auto} denotes the specified number of autonomy days and η_{bess} represents the battery efficiency.

2.6. Control Strategies

2.6.1. Load Following

Figure 10 shows the flowchart of the LF control strategy for the grid-connected microgrid with a solar PV system and BESS. This strategy is categorized into four predefined control rules.

- Rule 1: If $P_{pv,t} > P_{load,t}$, $P_{pv,t}$ supplies the load demand while the excess power is used to charge the BESS. If the BESS reaches full charge, any remaining excess power is exported to the main grid.
- Rule 2: If $P_{pv,t} = P_{load,t}$, $P_{pv,t}$ is fully used to meet the load demand.
- Rule 3: If $P_{pv,t} \leq P_{load,t}$ and $SoC_t \geq SoC_{MIN}$, BESS discharges to serve the net load. In the case of $P_{pv,t} + P_{bess,t} < P_{load,t}$, the main grid will serve the net load.
- Rule 4: If $P_{pv,t} \leq P_{load,t}$ and $SoC_t < SoC_{MIN}$, the main grid will serve the net load.

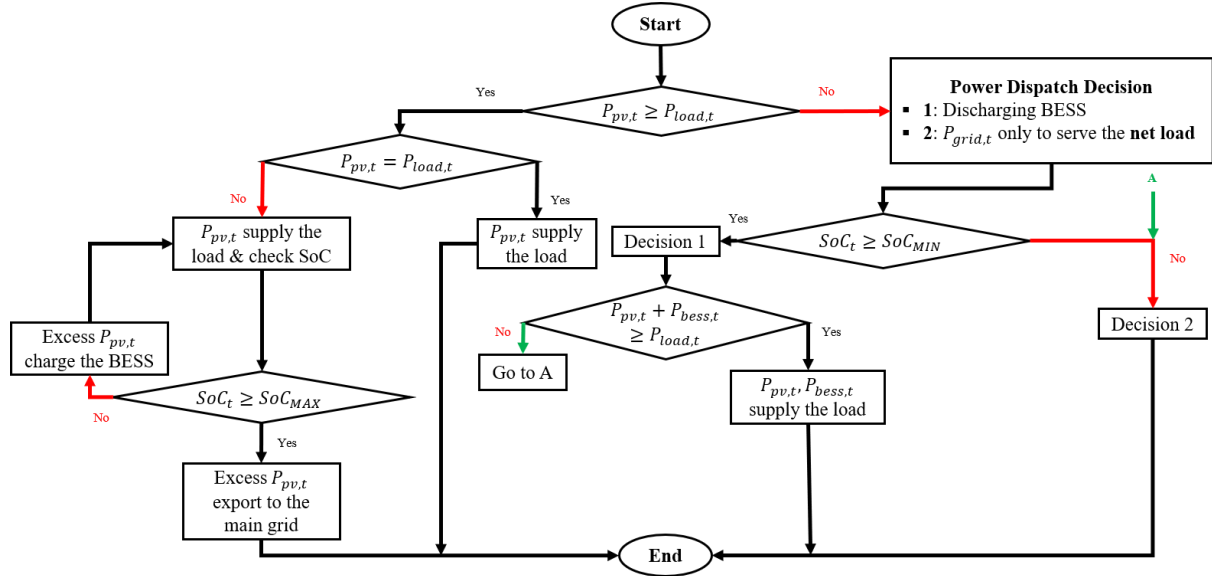


Figure 10. Flowchart of LF control strategy for the solar PV, BESS, and main grid

2.6.2. Cycle Charging

Figure 11 shows the flowchart of the CC control strategy for the same grid-connected microgrid system configuration. This strategy is categorized into four predefined control rules.

- Rule 1: The operating condition is similar to Rule 1 of the LF strategy.
- Rule 2: The operating condition is similar to Rule 2 of the LF strategy.
- Rule 3: If $P_{pv,t} \leq P_{load,t}$ and $SoC_t \geq SoC_{MIN}$, BESS discharges to serve the net load. However, in the case of $P_{pv,t} + P_{bess,t} < P_{load,t}$, the load demand will be fully served by the main grid.
- Rule 4: If $P_{pv,t} \leq P_{load,t}$ and $SoC_t < SoC_{MIN}$, the load demand will be fully served by the main grid. Any excess $P_{pv,t}$ will follow the operating condition in Rule 1, but following flow B.

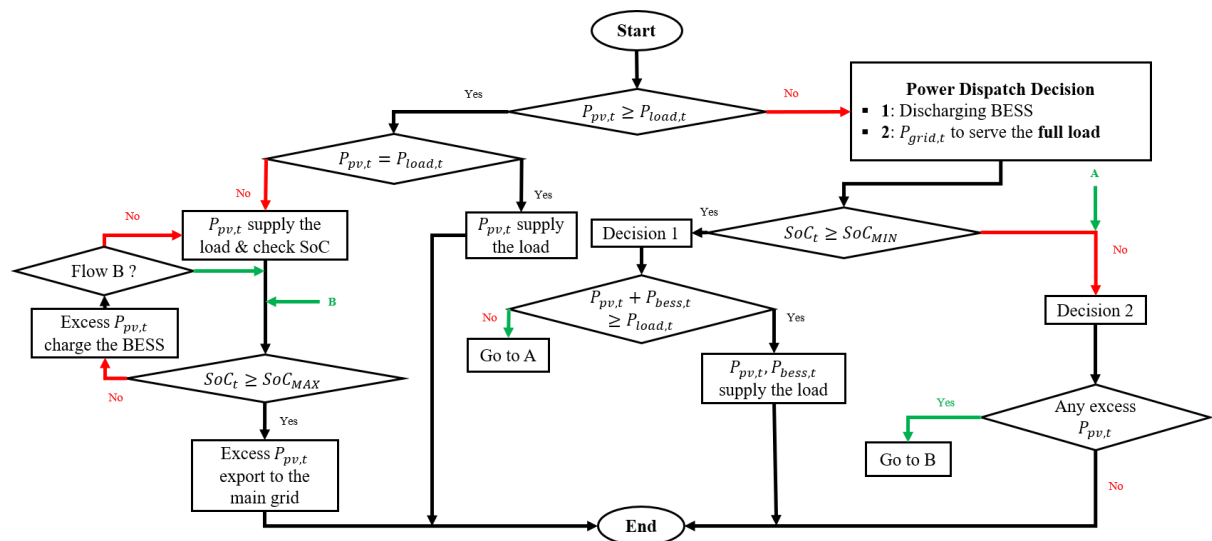


Figure 11. Flowchart of CC control strategy for the solar PV, BESS, and main grid

2.6.3. LSTM-Load Following

Figure 12 shows the flowchart of the LSTM-LF control strategy for the same grid-connected microgrid system configuration. In contrast with the conventional LF control strategy, the proposed LSTM-LF strategy involves forecast information on future solar PV energy generation, $P'_{pv,t}$, and residential load demand, $P'_{load,t}$, together with the corresponding future ToU electricity tariff, $\lambda'_{ToU,t}$, into the control decision-making process. The operation of this strategy is categorized into seven different predefined control rules.

- Rule 1: If $P_{pv,t} > P_{load,t}$, $P'_{pv,t} < P'_{load,t}$ and low $\lambda'_{ToU,t}$, $P_{pv,t}$ supplies the load demand while the BESS charges partially. If the BESS reaches its maximum SoC, it exports any remaining excess power to the main grid.
- Rule 2: If $P_{pv,t} > P_{load,t}$, $P'_{pv,t} < P'_{load,t}$ and high $\lambda'_{ToU,t}$, $P_{pv,t}$ supplies the load demand while the BESS charges aggressively. If the BESS reaches its maximum SoC, it exports any remaining excess power to the main grid.
- Rule 3: If $P_{pv,t} > P_{load,t}$, $P'_{pv,t} > P'_{load,t}$ and low $\lambda'_{ToU,t}$, $P_{pv,t}$ supplies the load demand while the BESS charges moderately. If the BESS reaches its maximum SoC, it exports any remaining excess power to the main grid.
- Rule 4: If $P_{pv,t} > P_{load,t}$, $P'_{pv,t} > P'_{load,t}$ and high $\lambda'_{ToU,t}$, $P_{pv,t}$ supplies the load demand while the BESS charges aggressively. If the BESS reaches its maximum SoC, any remaining excess power is exported to the main grid.
- Rule 5: The operating condition is similar to Rule 2 of the LF strategy.
- Rule 6: The operating condition is similar to Rule 3 of the LF strategy.
- Rule 7: The operating condition is similar to Rule 4 of the LF strategy.

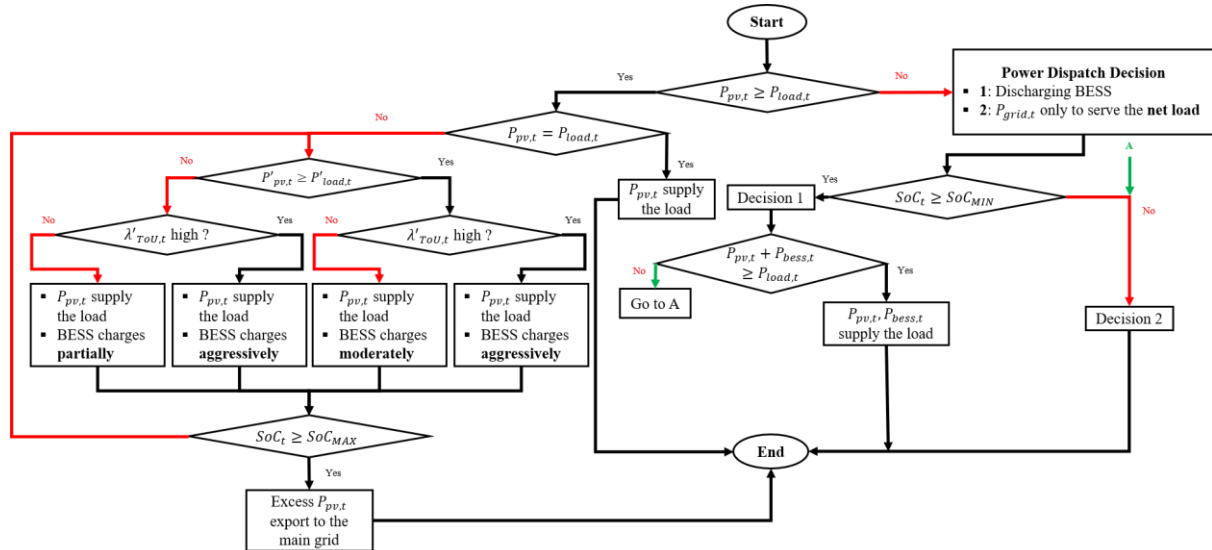


Figure 12. Flowchart of LSTM-LF control strategy for the solar PV, BESS, and main grid

2.6.4. LSTM-Cycle Charging

Figure 13 shows the flowchart of the LSTM-CC control strategy for the same grid-connected microgrid system configuration. Similar to the LSTM-LF strategy, the proposed LSTM-CC strategy involves forecast information on future solar PV energy generation, $P'_{pv,t}$, and residential load demand, $P'_{load,t}$, together with the corresponding future ToU electricity tariff, $\lambda'_{ToU,t}$, into the control decision-making process. This strategy is categorized into seven predefined control rules.

- Rule 1: The operating condition is similar to Rule 1 of the LSTM-LF strategy.
- Rule 2: The operating condition is similar to Rule 2 of the LSTM-LF strategy.
- Rule 3: The operating condition is similar to Rule 3 of the LSTM-LF strategy.
- Rule 4: The operating condition is similar to that in Rule 4 of the LSTM-LF strategy.
- Rule 5: The operating condition is similar to that in Rule 2 of the LF strategy.
- Rule 6: The operating condition is similar to that in Rule 3 of the CC strategy.
- Rule 7: The operating condition is similar to that in Rule 4 of the CC strategy.

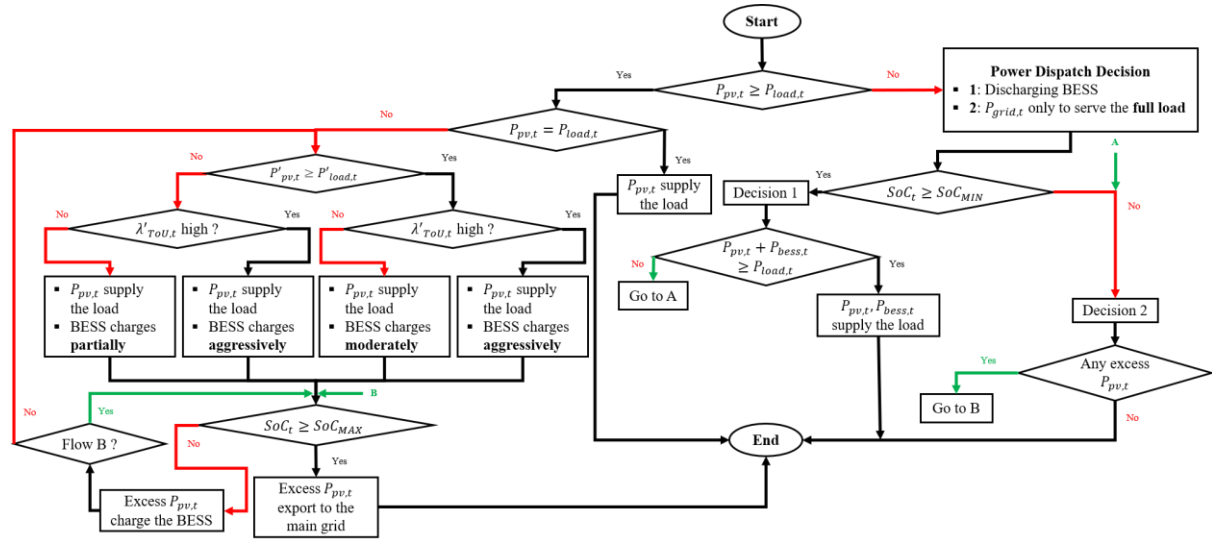


Figure 13. Flowchart of LSTM-CC control strategy for the solar PV, BESS, and main grid

3. RESULTS AND DISCUSSION

3.1. Analysis of LSTM Predictive Models Performance

Table 4 summarizes the performance comparison between the developed LSTM and baseline Feedforward Neural Network (FFNN) predictive models for solar PV energy generation and load demand forecasting using the training, validation, and test datasets. To ensure a fair comparison, the FFNN predictive models were developed using the same input variables, data preprocessing procedures, and training settings as the LSTM predictive models. The only difference is that the FFNN predictive models replace the LSTM layer with a feedforward fully connected layer.

For the solar PV energy generation forecasting, the developed LSTM model achieved MAE values of 0.113933 kWh, 0.079352 kWh, and 0.081602 kWh for the training, validation, and test datasets, respectively. Similarly, the MSE values were 0.045747 kWh², 0.019729 kWh², and 0.019591 kWh², while the RMSE values were 0.213836 kWh, 0.139233 kWh, and 0.139149 kWh. The model also achieved high R-Squared values of 94.93%, 97.83%, and 97.88% for the training, validation, and test datasets, respectively. Compared with the FFNN model, the LSTM consistently achieved lower prediction errors and higher R-Squared values across all datasets, demonstrating its better ability to model and capture the nonlinear relationship between input weather variables and solar PV energy generation. Furthermore, the comparable performance on the validation and test datasets shows that the LSTM model generalizes well to unseen data without significant performance degradation, indicating minimal overfitting.

Figure 14(a) shows the actual and predicted solar PV energy generation profiles for a single solar PV system from 11th to 21st September 2025. In the figure, the red solid line denotes the actual energy generation, while the blue dashed line represents the predicted values. As shown in the figure, the predicted profile closely follows the actual solar PV energy generation pattern, with only minor deviations during the spikes, demonstrating the developed LSTM model's ability to accurately forecast short-term solar PV output.

Similarly, Table 4 shows that the developed LSTM model also outperformed the FFNN model in residential load demand forecasting. The developed LSTM model achieved MAE values of 0.040838 kWh, 0.016715 kWh, and 0.016092 kWh for the training, validation, and test datasets, respectively. The corresponding MSE values were 0.004079 kWh², 0.000557 kWh², and 0.000521 kWh², while the RMSE values were 0.063749 kWh, 0.023357 kWh, and 0.022831 kWh. In addition, the model achieved R-Squared values of 95.24%, 99.21%, and 99.19%, indicating excellent predictive capability across all datasets. Compared with the FFNN model, the LSTM consistently produced lower prediction errors and higher coefficients of determination, indicating that its memory-based architecture is more effective in capturing the sequential patterns in residential load demand data. Furthermore, the comparable validation and test results indicate the robustness and generalization capability of the developed LSTM model for residential load demand prediction.

Table 4. Performance comparison of the developed LSTM and the baseline FFNN

Performance of LSTM Predictive Models						
Solar PV Energy Generation			Load Demand			
	Training	Validation	Test	Training	Validation	Test
MAE (kWh)	0.113933	0.079352	0.081602	0.040838	0.016715	0.016092
MSE (kWh ²)	0.045747	0.019729	0.019591	0.004079	0.000557	0.000521
RMSE (kWh)	0.213836	0.139233	0.139149	0.063749	0.023357	0.022831
R-Squared (%)	94.93	97.83	97.88	95.24	99.21	99.19
Performance of FFNN Predictive Models						
MAE (kWh)	0.146536	0.081079	0.086748	0.046429	0.023673	0.021795
MSE (kWh ²)	0.079981	0.021226	0.022103	0.005286	0.001071	0.000804
RMSE (kWh)	0.282509	0.144911	0.147929	0.072555	0.032290	0.028354
R-Squared (%)	91.16	97.59	97.61	93.53	98.69	98.86

Figure 14(b), (c) and (d) present the predicted load demand profiles of the three selected Malaysian residential load types, namely terrace house, apartment, and condominium, from 11th to 21st September 2025. The red solid line denotes the actual values, whereas the blue dashed line represents the predicted values. Based on these figures, the predicted load demand closely follows the actual load demand pattern, with only minor deviations observed during the spikes.

The comparative results demonstrate that the developed LSTM models consistently outperformed the baseline FFNN models for both solar PV energy generation and residential load demand forecasting. The LSTM achieved lower MAE, MSE, and RMSE values, along with higher R-Squared values, across the training, validation, and test datasets. The LSTM's improved performance can be attributed to its ability to retain and use information from previous time steps, enabling it to capture sequential patterns that are difficult to represent with the conventional feedforward architecture of the FFNN models. In addition, the training and validation loss values remained closely aligned throughout the 2000 training epochs, suggesting stable convergence with minimal overfitting. The comparable validation and test performance further demonstrates the developed LSTM models' generalization capability on unseen data. These findings justify selecting the developed LSTM predictive models for

integration into the proposed hybrid EMS, where accurate short-term forecasting is essential to support effective energy management and control decisions.

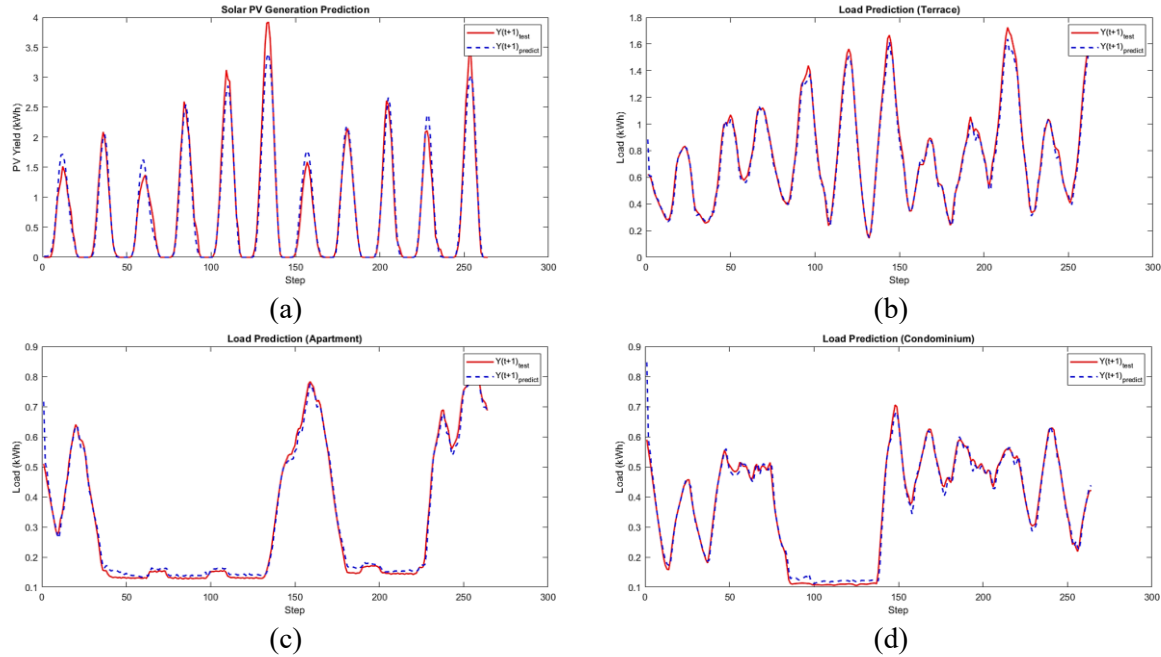


Figure 14. Predicted patterns of (a) solar PV energy generation, (b) load demand for a terrace house, (c) load demand for an apartment house, and (d) load demand for a condominium house, between 11th and 21st September 2025

3.2. Evaluation of the Proposed EMS Performance

Following the successful development of LSTM models for solar PV energy generation and load demand forecasting, this section presents the performance evaluation of the proposed LSTM-RBC-based EMS (LSTM-LF and LSTM-CC). It compares performance with the conventional RBC-based EMS (LF and CC) in managing the operation of the grid-connected residential microgrid under initial BESS SoC scenarios of 80%, 50%, and 30%. The evaluation considers both the operational and economic performance of the EMS, including the power flow analysis, the BESS SoC behavior under ToU electricity pricing, the grid energy exchange, RES utilization, and the associated operating costs, such as grid import and export costs, BESS degradation cost, PV O&M cost, and the overall operating cost of the residential microgrid system.

In this research, all EMS control strategies were evaluated over an 11-day period, from 11th to 21st September 2025, covering 264 operating hours. The evaluation used data for the selected location in Kuala Lumpur, Malaysia (3.1864, 101.7059), as described in the methodology section. **Table 5** summarizes the simulation period, including the day type and the applicable ToU electricity tariff category in Malaysia. The evaluation covers weekdays, weekends, and a holiday (Malaysia Day), allowing the EMS to operate under varying electricity tariff conditions.

The performance evaluation begins with the analysis of the power flow profiles of the four EMS control strategies over the 11-day simulation period, as illustrated in **Figure 15**, **Figure 16**, and **Figure 17**. **Figure 15** compares the power flow profiles under an initial BESS SoC of 80%, while Figures 16 and 17 show the corresponding comparisons for initial BESS SoC scenarios of 50% and 30%. In these figures, the green, blue, red, and purple solid lines represent

the power flow profiles of the solar PV system, BESS, main grid, and residential load demand, respectively.

Table 5. Classification of the simulation days according to the day type and the applicable ToU electricity tariff category based on the Malaysian context

Date	Type of Day	Applicable ToU Electricity Tariff
11 th September 2025	Weekday	Peak & Off-peak pricing
12 th September 2025	Weekday	Peak & Off-peak pricing
13 th September 2025	Weekend	Off-peak pricing only
14 th September 2025	Weekend	Off-peak pricing only
15 th September 2025	Weekday	Peak & Off-peak pricing
16 th September 2025	Holiday (Malaysia Day)	Off-peak pricing only
17 th September 2025	Weekday	Peak & Off-peak pricing
18 th September 2025	Weekday	Peak & Off-peak pricing
19 th September 2025	Weekday	Peak & Off-peak pricing
20 th September 2025	Weekend	Off-peak pricing only
21 st September 2025	Weekend	Off-peak pricing only

As shown in Figure 15(a, c), Figure 16(a, c), and Figure 17(a, c), the conventional RBC-based EMS controls dispatchable DERs solely based on the predefined control rules presented earlier in Figure 10 and Figure 11. The dispatch decisions for the BESS and the main grid were determined only by the instantaneous operating conditions of the residential microgrid, including the net load condition, BESS SoC, and current ToU electricity pricing. As a result, the BESS responded reactively to the variations in solar PV energy generation and residential load demand. In addition, the main grid participated in grid energy exchange only when surplus solar PV energy remained after the BESS reached its maximum allowable SoC, or when the available power from the solar PV system and BESS was insufficient to meet residential load demand.

In contrast, the proposed LSTM-RBC-based EMS retains the predefined RBC structure while incorporating LSTM-based forecasting to anticipate future net load conditions and upcoming ToU electricity tariff periods. Using this predictive information, the EMS can schedule BESS charging and discharging decisions in advance, enabling more economical coordination between the BESS and the main grid. As observed in **Figure 15(b, d)**, **Figure 16(b, d)**, and **Figure 17(b, d)**, this predictive capability results in more active grid energy exchange across all initial BESS SoC scenarios. Unlike the conventional RBC-based EMS, where the main grid serves only as a backup energy source, the proposed hybrid EMS incorporates grid energy exchange into its overall dispatch strategy by importing and exporting electricity based on future net load and the ToU electricity tariff. Although grid energy exchange increases, the proposed hybrid EMS consistently maintains sufficient BESS energy to support residential load demand during high-tariff periods, as shown in **Figure 18**. This coordinated dispatch strategy prioritizes electricity import during lower-tariff periods while minimizing grid import during higher-tariff periods, thereby reducing the overall operating cost of the residential microgrid.

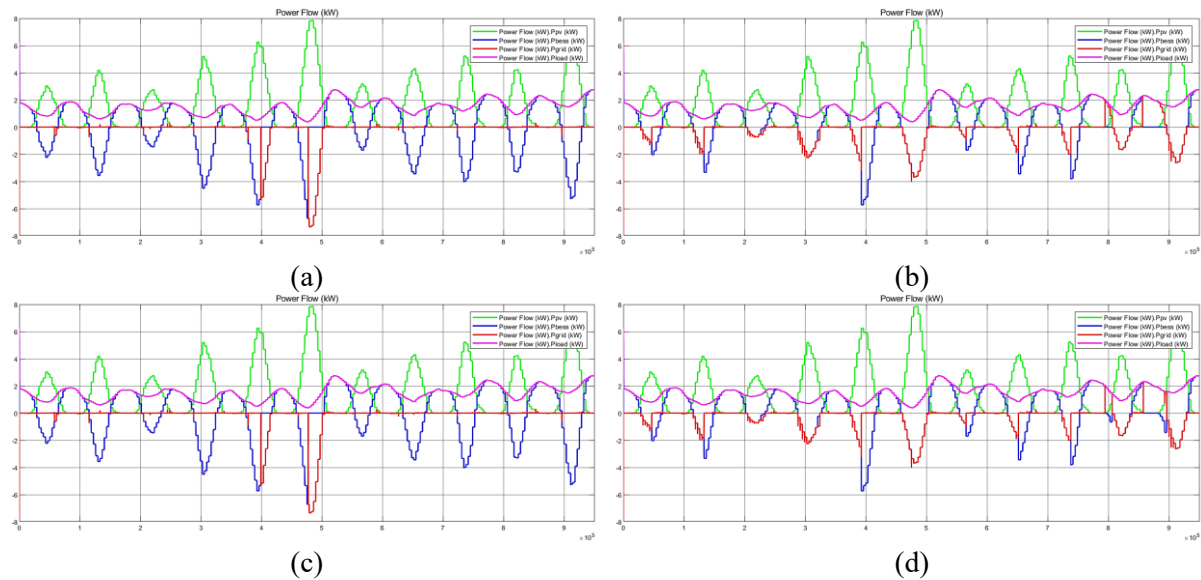


Figure 15. Comparison of power flow profiles for different EMS control strategies at an initial BESS SoC of 80%: (a) LF, (b) LSTM-LF, (c) CC, and (d) LSTM-CC

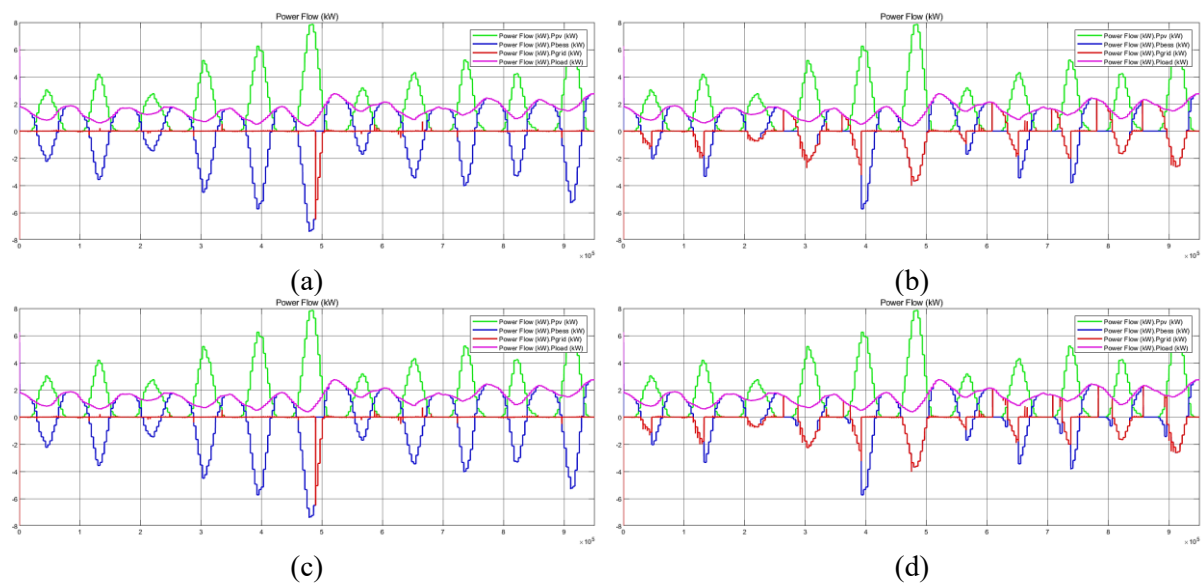


Figure 16. Comparison of power flow profiles for different EMS control strategies at an initial BESS SoC of 50%: (a) LF, (b) LSTM-LF, (c) CC, and (d) LSTM-CC

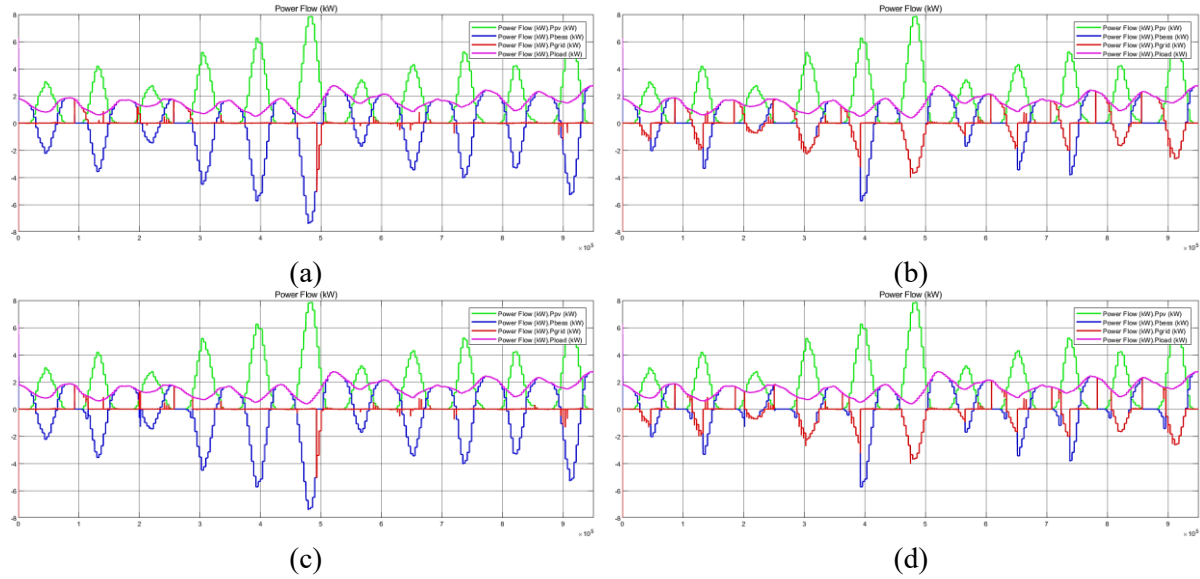


Figure 17. Comparison of power flow profiles for different EMS control strategies at an initial BESS SoC of 30%: (a) LF, (b) LSTM-LF, (c) CC, and (d) LSTM-CC

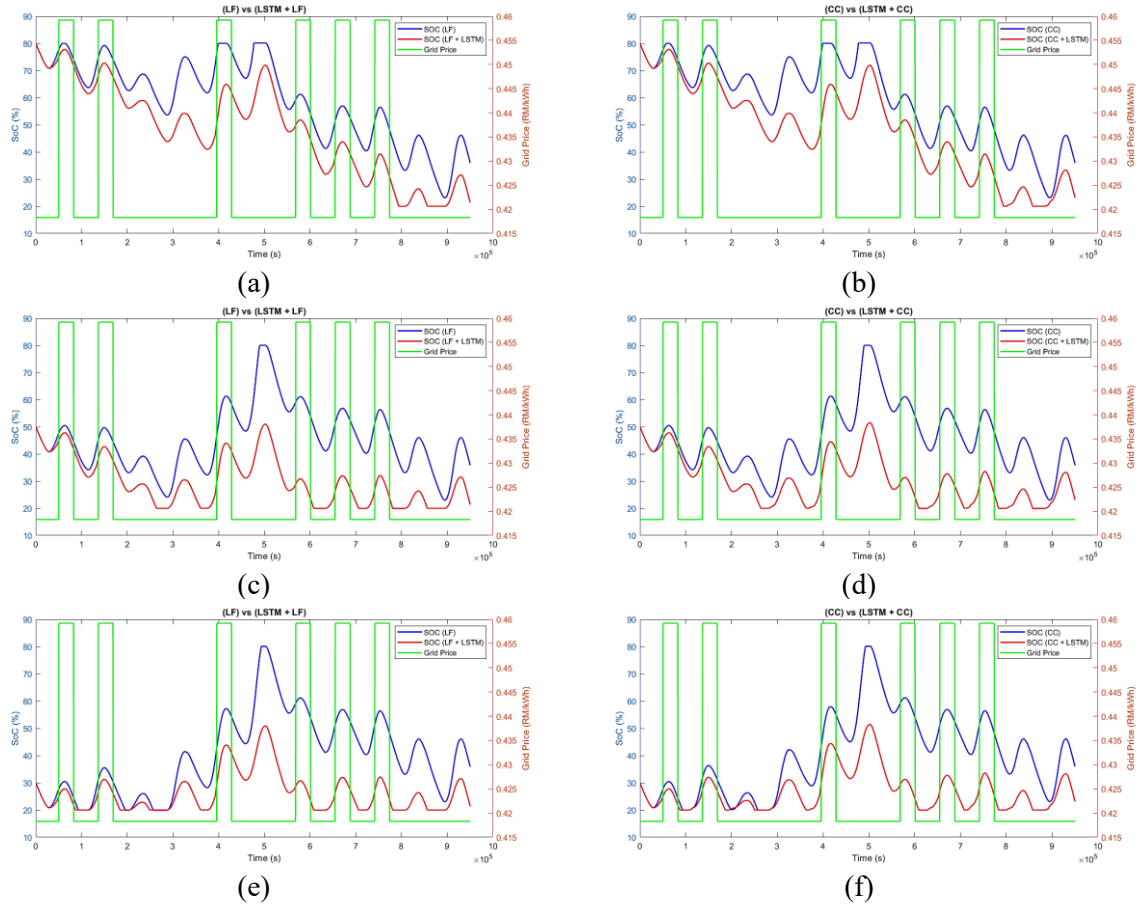


Figure 18. Comparison of BESS SoC profiles under grid ToU electricity prices for the LF and LSTM-LF control strategies at initial SoC of (a) 80%, (c) 50%, and (e) 30%, and for the CC and LSTM-CC control strategies at initial SoC of (b) 80%, (d) 50%, and (f) 30%

Table 6 summarizes the operational and economic performance of the evaluated EMS control strategies under different initial BESS SoC scenarios. The comparison includes key performance indicators such as grid energy import and export, SSR, PV utilization, and associated operating costs, including grid import and export costs, BESS degradation cost, PV O&M cost, and the residential microgrid's total operating cost.

As summarized in **Table 6**, the proposed LSTM-RBC-based EMS consistently achieved a lower total operating cost than the conventional RBC-based EMS across all initial BESS SoC scenarios. This reduction demonstrates the economic benefit of integrating LSTM-based forecasting into the conventional RBC framework, as predicted future net load conditions and upcoming ToU electricity tariff periods enable the EMS to make more informed dispatch decisions rather than relying solely on the microgrid's instantaneous operating conditions. Thus, the BESS and main grid can be coordinated based on both anticipated net load conditions and the corresponding ToU electricity tariff periods.

Table 6. Comparison of the economic and operational performance of different EMS control strategies under varying initial BESS SoC scenarios

EMS Control Strategy					
Initial SoC of 80%					
Parameter	Unit	LF	CC	LSTM-LF	LSTM-CC
Grid Import	kWh	0.06	0.06	24.47	26.82
Grid Export	kWh	48.56	48.56	93.77	93.77
SSR	%	99.98	99.98	93.72	93.12
PV Utilization	%	87.93	87.93	76.70	76.70
Grid Import Cost	RM	0.03	0.03	10.24	11.22
Grid Export Offset	RM	17.97	17.97	34.69	34.69
BESS Degradation Cost	RM	29.33	29.33	26.30	26.40
PV O&M Cost	RM	13.66	13.66	13.66	13.66
Total Operation Cost	RM	25.05	25.05	15.51	16.60
Initial SoC of 50%					
Grid Import	kWh	0.08	0.08	59.72	62.36
Grid Export	kWh	10.46	10.46	93.81	93.75
SSR	%	99.98	99.98	84.68	84.00
PV Utilization	%	97.40	97.40	76.68	76.70
Grid Import Cost	RM	0.03	0.03	24.98	26.09
Grid Export Offset	RM	3.87	3.87	34.71	34.69
BESS Degradation Cost	RM	29.33	29.33	21.94	22.54
PV O&M Cost	RM	13.66	13.66	13.66	13.66
Total Operation Cost	RM	39.15	39.15	25.88	27.60
Initial SoC of 30%					
Grid Import	kWh	18.81	19.80	83.43	86.15
Grid Export	kWh	5.18	6.03	93.96	93.85
SSR	%	95.17	94.92	78.59	77.89
PV Utilization	%	98.71	98.50	76.64	76.67
Grid Import Cost	RM	7.87	8.23	34.90	36.04
Grid Export Offset	RM	1.92	2.23	34.77	34.72
BESS Degradation Cost	RM	27.01	27.15	19.00	19.84
PV O&M Cost	RM	13.66	13.66	13.66	13.66
Total Operation Cost	RM	46.63	46.87	32.80	34.81

Compared with the conventional RBC-based EMS, the proposed hybrid EMS showed higher grid energy import and export values. However, this does not mean that the residential microgrid relies heavily on the main grid; it reflects a more strategic use of grid energy exchange. By using future net load and the ToU electricity tariff schedule, the proposed hybrid

EMS prioritizes grid electricity import during lower ToU tariff periods while minimizing grid import during higher tariff periods. Meanwhile, surplus solar PV energy that cannot be used by the residential load or stored in the BESS is exported to the main grid, providing an energy offset under the Solar ATAP scheme. Although the proposed hybrid EMS strategy involves more grid energy exchange, it does not necessarily increase operating costs, as costs depend on both the amount and timing of energy exchanged with the grid.

However, the proposed hybrid EMS introduces trade-offs between microgrid SSR and local RES utilization. As shown in Table 6, the proposed hybrid EMS achieved lower SSR and PV utilization than the conventional RBC-based EMS. The reduction in SSR is mainly due to greater grid imports during lower ToU tariff periods, allowing the BESS to preserve its SoC for later use during higher-tariff periods. Similarly, the reduction in PV utilization mainly results from exporting more surplus PV energy to the main grid rather than using it locally within the microgrid.

The proposed hybrid EMS also resulted in lower BESS degradation costs than the conventional RBC-based EMS across all evaluated initial SoC scenarios. This shows that predictive BESS scheduling enables more effective coordination of charging and discharging operations, reducing unnecessary battery cycling while maintaining reliable support for the residential load.

Overall, Table 6 demonstrates that integrating LSTM-based predictive models into the conventional RBC framework improves the EMS's operational and economic performance, while introducing trade-offs in SSR and PV utilization. The proposed hybrid EMS achieved a significant reduction in total operating cost by approximately 30% to 40% compared with the RBC-based EMS across different initial BESS SoC scenarios. Although the proposed hybrid EMS results in more active grid energy exchange and noticeably lower SSR and PV utilization, this reflects a trade-off between economic performance and energy self-sufficiency. Active grid energy exchange during favorable ToU electricity tariff periods reduces total operating cost. However, this economic benefit also increases the dependency on the main grid and reduces the utilization of locally generated solar PV energy.

4. CONCLUSION

This research proposed a hybrid EMS that integrates ML-based LSTM predictive models with the conventional RBC strategies, such as LF and CC, with a particular focus on the Malaysian residential microgrid scenario. The proposed hybrid EMS was designed to reduce the residential microgrid's total operating cost through more economical DER coordination under the Malaysian ToU electricity tariff and Solar ATAP scheme. The developed grid-connected microgrid consists of a solar PV system, BESS, and three residential load profiles representing terrace, apartment, and condominium households in Kuala Lumpur, Malaysia. Within the proposed framework, the LSTM predictive models provide one-hour-ahead predictions of solar PV energy generation and load demand, while the RBC strategies use these predictions, along with instantaneous operating conditions, to strategically coordinate DER dispatch.

The simulation results demonstrate that the developed LSTM predictive models consistently outperformed the baseline FFNN models in forecasting both solar PV energy generation and residential load demand. For solar PV energy generation, the LSTM model achieved a test MAE of 0.081602 kWh, MSE of 0.019591 kWh², RMSE of 0.139149 kWh, and R-Squared of 97.88%. Similarly, the residential load demand model achieved a test MAE of 0.016092 kWh, MSE of 0.000521 kWh², RMSE of 0.022831 kWh, and R-Squared of

99.19%. These results demonstrate that the developed LSTM models can accurately capture short-term variations in solar PV energy generation and residential load demand, providing reliable predictive information for the proposed hybrid EMS.

Leveraging these forecasts, the proposed hybrid EMS achieved a significant reduction in total operating cost by around 30% to 40%, along with lower BESS degradation costs, compared with the RBC-only strategies across different initial BESS SoC scenarios. The predictive information enabled the proposed hybrid EMS to schedule BESS charging and discharging in advance and to coordinate the dispatchable DERs more economically. As a result, the system strategically imported electricity during lower ToU tariff periods, while exporting excess solar PV energy to the main grid when conditions were favorable under the Solar ATAP scheme. At the same time, the BESS SoC was maintained at a level sufficient to support the residential load during high-tariff periods. Although the proposed hybrid EMS increased grid energy exchange and reduced SSR and PV utilization, these trade-offs lowered total operating costs.

Despite the encouraging results, several limitations remain. The performance evaluation was conducted over a relatively short simulation period to maintain computational feasibility while still capturing short-term variations in solar PV energy generation and residential load demand. Although Malaysia's tropical climate generally experiences less seasonal variation than regions with four distinct seasons, a longer evaluation period would still provide a more comprehensive analysis of seasonal variations throughout the year. In addition, the proposed hybrid EMS was validated only through MATLAB Simulink simulation and has not yet been evaluated under real operating conditions. The study is also geographically limited to the weather, residential load, and tariff conditions of Kuala Lumpur, Malaysia, which may not generalize to other locations. Furthermore, the comparative analysis was limited to the baseline FFNN model and conventional RBC strategies; therefore, it does not establish superiority over other forecasting models or advanced EMS techniques. Finally, the study also did not include a statistical uncertainty analysis of the forecasting and EMS performance.

Future research may address these limitations by extending the evaluation to longer and more diverse operating periods and analyzing the proposed hybrid EMS across different geographical locations. Its applicability may also be investigated for islanded microgrid operation, particularly in rural and remote areas with limited access to the main grid. In addition, future studies may explore more advanced EMS strategies, such as Reinforcement Learning, Multi-Criteria Decision Making techniques, and other intelligent optimization methods, to improve the DERs dispatch under dynamic operating conditions. Future studies may also include statistical uncertainty and sensitivity analyses to assess the robustness of forecasting and EMS performance under variations in input data, prediction errors, and system parameters. Last but not least, integrating intraday LSTM prediction may further improve the microgrid's operational and economic performance.

ACKNOWLEDGEMENT

This work is supported by the Ministry of Higher Education (MoHE) Malaysia under the Fundamental Research Grant Scheme FRGS (Ref: FRGS/1/2023/TK07/UIAM/02/5 – Project ID: FRGS23-320-0929).

REFERENCES

- [1] L. Fusheng, L. Ruisheng, and Z. Fengquan, "Composition and classification of the microgrid," *Microgrid technology and engineering application*, pp. 11–27, 2016.

-
- [2] “Energy Commission - Voltan Nominal.” Accessed: Mar. 20, 2026. [Online]. Available: <https://www.st.gov.my/en/web/general/details/144>
 - [3] “Energy transition in Malaysia.” Accessed: Mar. 19, 2026. [Online]. Available: <https://www.energymonitor.ai/data-insights/energy-transition-in-malaysia/?cf-view>
 - [4] T. B. Lopez-Garcia, A. Coronado-Mendoza, and J. A. Domínguez-Navarro, “Artificial neural networks in microgrids: A review,” *Eng. Appl. Artif. Intell.*, vol. 95, p. 103894, Oct. 2020, doi: 10.1016/J.ENGAPPAI.2020.103894.
 - [5] P. A. Cárdenas, M. Martínez, M. G. Molina, and P. E. Mercado, “Development of Control Techniques for AC Microgrids: A Critical Assessment,” *Sustainability* 2023, Vol. 15, Page 15195, vol. 15, no. 21, p. 15195, Oct. 2023, doi: 10.3390/SU152115195.
 - [6] A. Elmouatamid, R. Ouladsine, M. Bakhouya, N. El Kamoun, M. Khaidar, and K. Zine-Dine, “Review of Control and Energy Management Approaches in Micro-Grid Systems,” *Energies* 2021, Vol. 14, Page 168, vol. 14, no. 1, p. 168, Dec. 2020, doi: 10.3390/en14010168.
 - [7] K. Hassan, T. Trireksani, A. Hussain, and H.-M. Kim, “A Rule-Based Modular Energy Management System for AC/DC Hybrid Microgrids,” *Sustainability* 2025, Vol. 17, Page 867, vol. 17, no. 3, p. 867, Jan. 2025, doi: 10.3390/SU17030867.
 - [8] A. S. Aziz, M. F. N. Tajuddin, M. R. Adzman, M. A. M. Ramli, and S. Mekhilef, “Energy Management and Optimization of a PV/Diesel/Battery Hybrid Energy System Using a Combined Dispatch Strategy,” *Sustainability* 2019, Vol. 11, Page 683, vol. 11, no. 3, p. 683, Jan. 2019, doi: 10.3390/SU11030683.
 - [9] S. Chakraborty, G. Modi, and B. Singh, “A Cost Optimized-Reliable-Resilient-Realtime-Rule-Based Energy Management Scheme for a SPV-BES-Based Microgrid for Smart Building Applications,” *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 2572–2581, Jul. 2023, doi: 10.1109/TSG.2022.3232283.
 - [10] M. M. Awad, A. Bhattacharjee, and K. N. Hasan, “Design and Energy Management of Solar PV-BESS Integrated Off-Grid Electric Vehicle Charging Station,” *2025 12th National Power Electronics Conference (NPEC)*, pp. 1–6, Dec. 2025, doi: 10.1109/NPEC66512.2025.11450150.
 - [11] K. Nassereddine, M. Turzynski, H. Biellokha, and R. Strzelecki, “Simulation of energy management system using model predictive control in AC/DC microgrid,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–12, Dec. 2025, doi: 10.1038/S41598-025-89036-7.
 - [12] M. Kaheni, J. Fu, and A. V. Papadopoulos, “Rule-Based Predictive Control for Battery Scheduling in Microgrids Under Power Generation and Load Uncertainties,” *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 9782–9792, 2025, doi: 10.1109/TASE.2024.3512882.
 - [13] S. Y. Chen and C. H. Chang, “Optimal Power Flows Control for Home Energy Management With Renewable Energy and Energy Storage Systems,” *IEEE Transactions on Energy Conversion*, vol. 38, no. 1, pp. 218–229, Mar. 2023, doi: 10.1109/TEC.2022.3198883.
 - [14] A. Abbasi, H. A. Khalid, H. Rehman, and A. U. Khan, “A Novel Dynamic Load Scheduling and Peak Shaving Control Scheme in Community Home Energy Management System Based Microgrids,” *IEEE Access*, vol. 11, pp. 32508–32522, 2023, doi: 10.1109/ACCESS.2023.3255542.
 - [15] M. M. Shibl, L. S. Ismail, and A. M. Massoud, “An Intelligent Two-Stage Energy Dispatch Management System for Hybrid Power Plants: Impact of Machine Learning Deployment,” *IEEE Access*, vol. 11, pp. 13091–13102, 2023, doi: 10.1109/ACCESS.2023.3243097.
 - [16] E. Giglio, G. Luzzani, V. Terranova, G. Trivigno, A. Niccolai, and F. Grimaccia, “An Efficient Artificial Intelligence Energy Management System for Urban Building Integrating
-

- Photovoltaic and Storage,” *IEEE Access*, vol. 11, pp. 18673–18688, 2023, doi: 10.1109/ACCESS.2023.3247636.
- [17] O. Akbulut, M. Cavus, M. Cengiz, A. Allahham, D. Giaouris, and M. Forshaw, “Hybrid Intelligent Control System for Adaptive Microgrid Optimization: Integration of Rule-Based Control and Deep Learning Techniques,” *Energies* 2024, Vol. 17, Page 2260, vol. 17, no. 10, p. 2260, May 2024, doi: 10.3390/EN17102260.
- [18] “Fahami Tarif Elektrik Anda | TNB Malaysia.” Accessed: Aug. 13, 2025. [Online]. Available: <https://www.mytnb.com.my/tariff/index.html?v=1.1.34>
- [19] “Solar ATAP Malaysia 2026 | Cut TNB Bills 90% | No Quota, Apply Anytime.” Accessed: Aug. 12, 2026. [Online]. Available: <https://solaratap.my/>
- [20] R. R. Onteru, V. Sandeep, Rupesh, and R. Onteru, “An intelligent model for efficient load forecasting and sustainable energy management in sustainable microgrids,” *Discover Sustainability* 2024 5:1, vol. 5, no. 1, pp. 170–, Jul. 2024, doi: 10.1007/s43621-024-00356-6.

DesnowOracle: Quantifying Mask-Oracle Gap Recovery by a Training-Free Snow Prior in Compact Image Desnowing

TEDDY SURYA GUNAWAN^{1*}, MIRA KARTIWI²

¹*Department of Electrical and Computer Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia*

²*Department of Information Systems, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia*

*Corresponding author: tsgunawan@iium.edu.my

(Received: 21 July 2026; Accepted: 19 August 2026; Published online: 11 September 2026)

ABSTRACT: Snow occludes visual information needed by outdoor perception systems. This study evaluates whether a deterministic, training-free spatial snow-response map improves compact image desnowing beyond the effect of its conditioning branch. All conditions use a fixed approximately 2-million-parameter backbone and matched controls. Under the original 2,500-image, 100-epoch protocol, estimated-map SSM improves mean PSNR by 1.004 dB over Base and by 0.490 and 0.517 dB over zero- and shuffled-input SSM. A paired analysis averaging predictions across three seeds gives median gains of 0.387 and 0.420 dB, respectively. Under an extended protocol that also changes training-set size and learning-rate scheduling, this ordering reverses: estimated-map SSM trails zero and shuffled controls by 0.109 and 0.160 dB in mean PSNR. Thus, the experiments demonstrate protocol sensitivity, not an isolated convergence effect. A mask oracle remains an architecture-specific, synthesis-privileged diagnostic reference; the estimated map recovers 12.4% of its short-protocol PSNR headroom. Results on 1,329 unpaired real images are mixed and indirect: no-reference indices favor the estimated map, whereas unlabelled detector counts do not distinguish its content from the controls. These findings show why spatial-prior claims require matched controls, correct repeated-measures analysis, and protocol checks.

ABSTRAK: Salji menghalang maklumat visual yang diperlukan oleh sistem persepsi luar bangunan. Kajian ini menilai sama ada peta tindak balas salji spatial yang deterministik dan tanpa latihan meningkatkan penyahsaljian imej padat melebihi kesan cabang pengkondisian. Semua keadaan menggunakan rangkaian asas tetap dengan kira-kira 2 juta parameter dan kawalan yang sepadan. Di bawah protokol asal 2,500 imej dan 100 epok, SSM dengan peta anggaran meningkatkan purata PSNR sebanyak 1.004 dB berbanding Base dan sebanyak 0.490 serta 0.517 dB berbanding SSM input sifar dan teracak. Purata analisis ramalan berpasangan merentas tiga benih memberikan peningkatan median masing-masing sebanyak 0.387 dan 0.420 dB. Di bawah protokol lanjutan yang turut mengubah saiz set latihan dan penjadualan kadar pembelajaran, susunan ini berbalik kepada SSM dengan peta anggaran ketinggalan berbanding kawalan sifar dan teracak sebanyak 0.109 dan 0.160 dB dalam purata PSNR. Oleh itu, eksperimen menunjukkan kepekaan terhadap protokol, bukan kesan penumpuan terasing. Topeng Oracle kekal sebagai rujukan diagnostik khusus seni bina dan keistimewaan sintesis; peta anggaran memulihkan 12.4% ruang PSNR protokol pendek. Dapatan daripada 1,329 imej sebenar tanpa pasangan adalah bercampur dan tidak langsung: indeks tanpa rujukan memihak kepada peta anggaran, manakala kiraan pengesan tanpa label tidak membezakan kandungannya daripada kawalan. Dapatan ini menunjukkan bahawa tuntutan sebelum spatial memerlukan kawalan sepadan, analisis ukuran berulang yang betul, dan semakan protokol.

KEYWORDS: *Image Desnowing; Adverse Weather Perception; Spatial Conditioning; Snow Severity Prior; Compact Restoration*

1. INTRODUCTION

Winter storms create a compound threat to transportation and machine perception in countries with recurrent seasonal snowfall. In the United States alone, an estimated 101,390 police-reported crashes occurred under snow or sleet conditions in 2023, including 22,293 injury crashes and 320 fatal crashes, as reported by the National Highway Traffic Safety Administration (NHTSA) in 2025. During a snowstorm, airborne particles reduce visibility and contrast, obscure pedestrians, vehicles, traffic signals, and road markings, and produce veiling and reflected-light artifacts. Accumulated snow can further conceal lane boundaries and alter the appearance of familiar objects. These effects degrade the visual evidence required by driver-assistance systems, autonomous vehicles, aerial platforms, outdoor robots, and surveillance systems precisely when reliable perception is most critical. Desnowing is therefore not merely a visual-enhancement task. Restoration-aware snowy-weather detection has produced mean average precision gains of 3.80, 2.10, and 1.86 percentage points across three real-world datasets, demonstrating that recovering useful image structure can improve downstream perception [1].

Mask-guided desnowing can improve restoration, but attribution remains difficult: a conditioned model may benefit from extra capacity or optimization even when its spatial input is uninformative. The relevant comparison is therefore not only against an unconditioned network, but also against the same conditioning branch supplied with zero and shuffled maps. This study measures that distinction and tests whether the resulting ordering changes under a substantially extended protocol.

Snow is especially difficult to remove because opaque, irregular particles can replace local scene structure rather than merely attenuating it. DesnowNet introduced Snow100K and used a learned intermediate mask to guide restoration [2]. Subsequent methods modeled contradict-channel information, particle size, transparency, semantic context, depth, and scale-aware interactions [3-6]. Lightweight architectures have also addressed the tension between restoration quality and network capacity [7-9]. More recently, a snow-shape prior was introduced by [10] as surrogate supervision when real snow masks are unavailable. Their prior replaces mask supervision during training, whereas the present prior replaces a mask input at inference. Neither purpose is interchangeable, and the earlier study does not quantify recovery relative to a ground-truth-mask oracle.

This study introduces DesnowOracle, a controlled framework for measuring spatial-prior effects within a fixed compact backbone. An identity-initialized spatial modulation head receives an estimated map, an all-zero map, a map from another image, a dilated ground-truth mask, or the exact ground-truth mask. The design separates the conditioning-branch effect from the effect of prior content. A second protocol tests whether the ordering is stable when training-set size, training duration, and learning-rate scheduling change together.

The study makes four contributions. First, it specifies a deterministic, training-free spatial snow-response map and evaluates its mask agreement by particle scale. Second, it compares estimated, zero, shuffled, dilated-mask, and exact-mask inputs through the same spatial modulation branch, with the exact mask treated as a synthesis-privileged diagnostic reference. Third, it reports scale-stratified restoration and an exploratory degradation-ladder analysis showing that mask AUC correlates with gain within that ladder but does not predict the estimated map out of family. Fourth, it shows that the estimated map ranks above the

uninformative controls under the original protocol but below them under an extended protocol that changes data volume, optimizer steps, and learning-rate scheduling. Real-image no-reference indices, unlabelled detector responses, and measured computational costs provide supporting evidence with explicit limitations.

2. RELATED WORK

The literature has established that snow masks, spatial priors, multiscale representations, and degradation-aware conditioning can improve image restoration, but it has not yet determined how much of that benefit arises from valid spatial information rather than additional model capacity.

2.1. Desnowing and Snow Priors

DesnowNet established Snow100K with paired synthetic snowy and clean images, snow masks, and an unpaired real subset, and used a learned intermediate mask with translucency recovery [2]. JSTASR modeled particle size and transparency through modified partial convolution [4], while HDCW-Net used a handcrafted contrast-channel signal as a training loss rather than an explicit network input [5]. DDMSNet introduced semantic and depth priors [6], and SnowFormer modeled scale-dependent context using a transformer [11]. These methods show that snow location, scale, and opacity are useful, but they do not determine how much of the improvement comes from accurate spatial guidance rather than architectural capacity.

Compact desnowing introduces a second challenge: preserving sufficient contextual reasoning without relying on a large model. The trade-off between network capacity and restoration performance was explicitly studied by [8]. LMQFormer combined a learned coarse Laplace-prior mask with a 2.181-million-parameter query transformer [9], while an efficient multiscale context-aware architecture was developed by [7]. These methods establish that compact desnowing is feasible, but baseline comparisons alone cannot establish whether a prior-guided model actually uses the prior spatially.

Recent prior-based methods sharpen this distinction. A snow-shape prior was embedded in a deep unfolding framework to replace impractical real-mask supervision [10]. By contrast, the present study supplies a deterministic prior directly to the restoration network and tests whether its spatial correspondence is causally useful. BAT-Net jointly predicts a snow mask and clean image through coupled transformer decoders, allowing the two tasks to refine each other [12]. Its mask remains a learned internal representation rather than a controllable input that can be removed, shuffled, spatially degraded, or replaced by ground truth. SnowMamba reports PSNR, SSIM, LPIPS, and DISTS separately across Snow100K-S, Snow100K-M, and Snow100K-L [13], confirming the value of scale-stratified evaluation. The distinction in the present study is therefore not stratification itself, but using stratification to measure how much mask-oracle headroom an estimated prior recovers.

Real-world evaluation has also advanced beyond isolated qualitative examples. SnowMaster uses multimodal large language model feedback and multiple restoration models to address real snowy scenes [14], while WeatherBench provides aligned real degraded and clean pairs for rain, snow, and haze [15]. These developments expose the limitation of the unpaired Snow100K real subset. Accordingly, the present study uses no-reference metrics on that subset as conservative supporting evidence and does not treat them as a substitute for paired real-world evaluation.

2.2. General Adverse-Weather Restoration and Conditioning Mechanisms

All-in-one restoration seeks to handle multiple degradations within a shared model, with recent work organized around degradation representations, prompts, adaptive pathways, and modulation [16]. AirNet learns contrastive degradation representations [17], TransWeather uses weather-type queries [18], and TKL distills degradation-specific teachers into a unified student [3]. More recent architectures use self-estimated degradation prompts, efficient convolutional prompts, and diffusion guidance [19-21]. These approaches primarily condition restoration through latent degradation representations or prompts rather than through an externally testable per-pixel prior.

AdaIR is the closest recent conceptual comparator because it mines frequency components and uses them to modulate restoration features [22]. The present study instead conditions features using an explicit spatial severity map. This design follows the broader principle of feature-wise linear modulation [23] and spatial feature transformation [24], but differs in two respects: the conditioning map is deterministic and training-free, and its value is evaluated through matched falsification and mask-oracle controls.

Taken together, prior work has established the effectiveness of learned masks, surrogate priors, multiscale representations, degradation prompts, and feature modulation. What remains unmeasured is how much mask-oracle headroom a training-free spatial prior recovers under a fixed backbone, and whether its gain survives zero-prior and shuffled-prior controls. This is the specific gap this study addresses. This unresolved attribution problem motivates DesnowOracle, which evaluates spatial-prior utility under matched falsification and mask-oracle controls. Throughout this paper, the zero- and shuffled-prior conditions are negative (placebo) controls in the standard experimental-design sense; the term falsification control is retained for continuity with the framework name and implies no stronger epistemic status.

2.3. Research Positioning

Table 1 positions DesnowOracle against representative desnowing methods according to the methodological evidence required to assess a spatial prior: the prior's role, matched falsification controls, comparison with a ground-truth-mask diagnostic, particle-scale stratification, real-image evaluation, and model size. This comparison is not intended as an absolute performance ranking because the methods use different architectures, training sets, resolutions, and evaluation protocols. Instead, it highlights the difference between showing that a prior-guided architecture performs well and establishing that the prior's spatial content drives the improvement.

Earlier methods employ learned snow masks, semantic or depth information, surrogate shape priors, and multiscale representations, demonstrating that degradation-aware guidance is valuable. Compact models such as LMQFormer further show that prior guidance can be incorporated with relatively few parameters, while SnowMamba establishes the importance of reporting performance separately across Snow100K-S, -M, and -L. However, the selected studies do not report matched zero- and shuffled-prior controls or quantify improvement relative to the same architecture supplied with a ground-truth mask. Consequently, their reported gains cannot isolate valid spatial correspondence from additional capacity, regularization, or optimization effects. Real-image evaluation is also uneven and, where unpaired data are used, cannot directly establish restoration fidelity.

DesnowOracle addresses this attribution gap within a fixed 1.966-million-parameter architecture. Its contribution is therefore not a claim of universal superiority, but a controlled evaluation framework combining matched negative controls, an architecture-specific mask

diagnostic, scale-stratified analysis, real-image supporting evidence, and an extended-protocol sensitivity check. The mask condition remains a privileged diagnostic reference, not a theoretical upper bound on desnowing performance.

Table 1. Research positioning relative to selected desnowing methods. The falsification, mask-oracle, stratification, and real-image columns indicate what each cited study explicitly reports. NR indicates not reported.

Method	Prior or mask role	Falsification controls	Mask-diagnostic reference	Scale-stratified S/M/L	Real-image evaluation	Parameters (M)
DesnowNet [2]	Learned intermediate snow mask	NR	NR	Aggregate only	Yes	26.151
JSTASR [4]	Learned particle-size and transparency representation	NR	NR	Aggregate only	NR	83.347
HDCW-Net [5]	Handcrafted contradict-channel training signal	NR	NR	Aggregate only	NR	18.958
DDMSNet [6]	Learned semantic and depth priors	NR	NR	Aggregate only	NR	229.454
SnowFormer [11]	Learned scale-aware contextual representation	NR	NR	Aggregate only	NR	NR
LMQFormer [9]	Learned coarse Laplace-prior mask	NR	NR	Aggregate only	NR	2.181
Deep unfolding network [10]	Snow-shape prior used as surrogate supervision	NR	NR	Aggregate only	NR	NR
SnowMamba [13]	Multiscale state-space representation	NR	NR	Yes	Reported	NR
BAT-Net [12]	Jointly learned snow mask and restored image	NR	NR	NR	Yes	NR
This study	Deterministic, training-free snow-severity prior	Zero, shuffled, and dilated controls	Yes; architecture-specific mask-gap recovery	Yes, diagnostic analysis	Yes, no-reference evaluation	1.966

3. DESNOWORACLE FRAMEWORK AND EXPERIMENTAL PROTOCOL

DesnowOracle combines deterministic prior construction, a fixed compact backbone, matched falsification controls, and versioned experimental manifests to isolate the contribution of spatial snow information and enable exact replication of every reported experiment. Figure

1 summarizes the complete experimental pipeline end-to-end, from prior construction through conditioning mechanisms, falsification and oracle controls, training protocol, and the two evaluation tracks, and makes explicit which conditions receive the prior (Concat, FiLM, SSM) and which do not (Base).

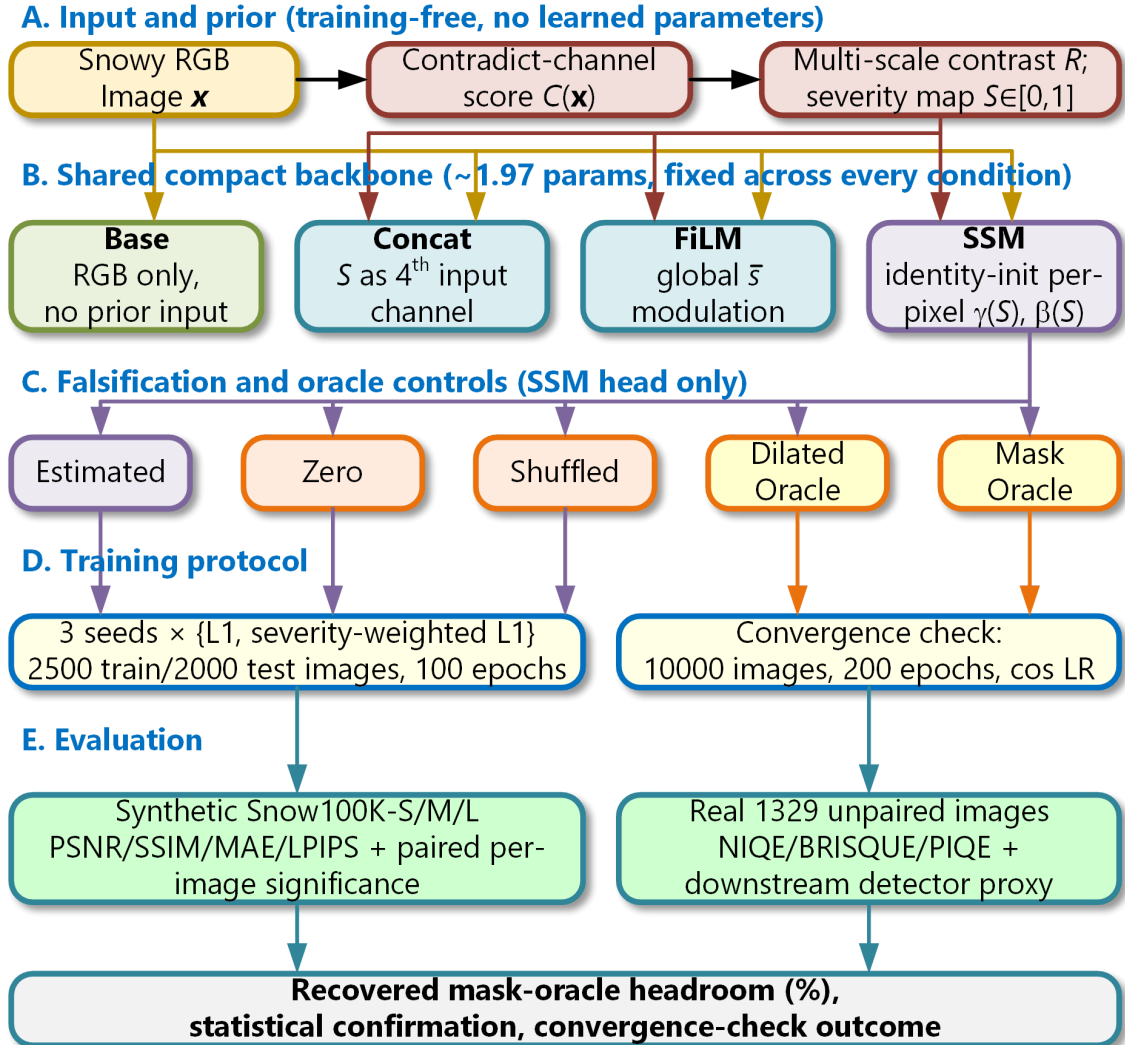


Figure 1. Overall experimental procedure. Base receives no prior input; Concat, FiLM, and SSM all receive the deterministic severity map $S(\mathbf{x})$; the SSM head is additionally evaluated under zero, shuffled, dilated-oracle, and mask-oracle controls (Sec. 3.3) and under both the original and convergence-check training protocols (Sec. 3.4, 4.7).

3.1. Snow-Severity Prior

Let $I \in [0,1]^{H \times W \times 3}$ denote an RGB snowy image and p a pixel location. Bright, nearly achromatic snow tends to have high channel intensity but low interchannel variation. Accordingly, define the value and channel-difference terms as $V(p) = \max_{c \in [R,G,B]} I_c(p)$, $D(p) = \max_c I_c(p) - \min_c I_c(p)$. Bright, nearly achromatic snow generally produces high V and low D . The contradict score is therefore:

$$C(p) = \text{clip}(V(p) - \lambda_c D(p), 0, 1), \quad (1)$$

where $\lambda_c = 0.75$. Positive local contrast is extracted using Gaussian kernels G_σ :

$$R_\sigma(p) = \max(C(p) - (G_\sigma * C)(p), 0), \quad R(p) = \max_{\sigma \in [1,2,3,4]} R_\sigma(p). \quad (2)$$

Each Gaussian kernel has radius 4σ and width $8\sigma + 1$, producing kernel sizes of 9, 17, 25, and 33 pixels for $\sigma = 1, 2, 3, 4$, respectively. Convolution uses OpenCV `BORDER_REFLECT_101` padding. Retaining only positive residuals suppresses the dark side of ordinary edges and concentrates the response on bright, locally contrasting structures.

The normalized severity prior and its spatial mean are

$$S(p) = \frac{R(p) - \min_q R(q)}{\max_q R(q) - \min_q R(q) + \varepsilon}, \quad \bar{S} = \frac{1}{HW} \sum_p S(p), \quad (3)$$

where $\varepsilon = 10^{-6}$. A constant response therefore produces an all-zero prior. The estimator contains no learned parameters and uses no ground-truth mask. In `Snow100KManifest.__getitem__`, the snowy image is first resized to 256×256 using `cv2.INTER_AREA`; the prior is then computed from that resized image. Clean targets are resized identically, while masks use `cv2.INTER_NEAREST`. No cropping or spatial augmentation is applied. The prior is recomputed deterministically during data loading rather than retrieved from a learned model or stored cache.

Algorithm 1. Snow-Severity Prior $S(\mathbf{x})$

Input: RGB image $I \in [0,1]^{H \times W \times 3}$; $\lambda_c = 0.75$; $\sigma \in [1,2,3,4]$; $\varepsilon = 10^{-6}$
Output: Severity map $S \in [0,1]^{H \times W}$; spatial mean $\bar{S}$

- 1: $V \leftarrow$ per-pixel $\max(R, G, B)$ channel of $\mathbf{I}$
- 2: $D \leftarrow$ per-pixel $[\max \text{ channel} - \min \text{ channel}]$ of $\mathbf{I}$
- 3: $C \leftarrow \text{clip}(V - \lambda_c D, 0, 1)$ $\triangleright$ Eq. (1)
- 4: **for** $\sigma \in [1,2,3,4]$ **do**
- 5: $R_\sigma \leftarrow \max(C - \text{GaussianBlur}(C, \sigma), 0)$
- 6: **end for**
- 7: $R \leftarrow \max(R_1, R_2, R_3, R_4)$ elementwise $\triangleright$ Eq. (2)
- 8: $S \leftarrow \frac{R - \min(R)}{\max(R) - \min(R) + \varepsilon}$ $\triangleright$ Eq. (3)
- 9: $\bar{S} \leftarrow \text{mean}(S)$
- 10: **return** $S, \bar{S}$

Algorithm 1 summarizes the deterministic construction of the spatial snow-response map S from the resized RGB input using Eqs. (1)–(3). Equation (3) normalizes each image independently. The output therefore describes relative within-image response, not absolute severity calibrated across images. In a low-snow image, normalization can stretch the residual noise floor across $[0,1]$, limiting interpretation of the FiLM global mean and preventing one threshold from being assumed valid across particle scales.

3.2. Shared Compact Backbone and Conditioning Mechanisms

Figure 2 presents the complete DesnowOracle pathway, from deterministic prior construction to the shared compact encoder-decoder and identity-initialized spatial snow modulation. Every condition uses the same three-stage U-Net-style encoder-decoder with residual RGB prediction [25]. With `base_ch=32`, the encoder widths are 32, 64, and 128 channels, followed by a 256-channel bottleneck. The decoder widths are 128, 64, and 32

channels. Each stage contains one **ConvBlock** comprising two 3×3 convolutions, each followed by GELU. No batch, instance, or layer normalization is used.

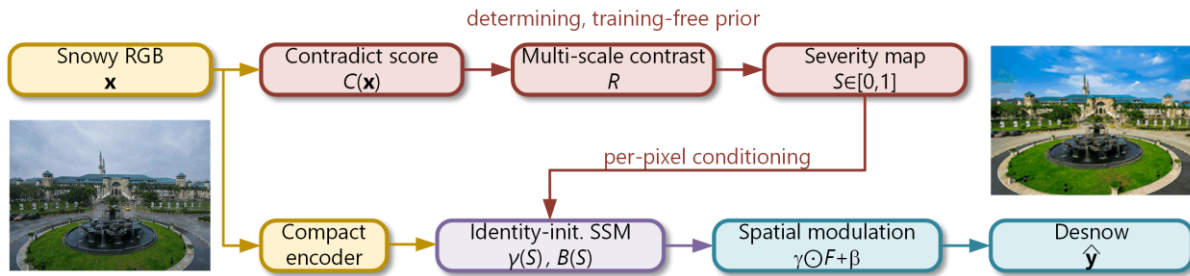


Figure 2. DesnowOracle architecture showing deterministic snow-prior construction and identity-initialized spatial snow modulation. The training-free severity map is converted into per-pixel scale and bias fields that modulate the final full-resolution decoder features.

Downsampling uses 2×2 max pooling. Decoder features are bilinearly interpolated to the corresponding skip-feature size with `align_corners=False`, then concatenated channel-wise with the encoder feature. The final 1×1 convolution maps 32 channels to RGB. The restored image is

$$\hat{Y} = \text{clamp}(I + \text{out}(d_1), 0, 1), \quad (4)$$

where d_1 is the final full-resolution decoder feature.

The Base condition receives RGB only. Concat appends S as a fourth input channel. This condition is retained as a capacity-matched baseline because an earlier prior-swap ablation in lightweight image dehazing found that replacing a physically meaningful prior with an all-zero map produced comparable performance, leaving the source of the apparent gain unattributable [26]. FiLM maps $\bar{s}$ through `Linear(1, 64)`, GELU, and `Linear(64, 64)`, then divides the output into 32-channel scale and bias vectors. Its final linear layer is zero-initialized.

SSM maps S through `Conv2d(1, 32, 3, padding=1)`, GELU, and `Conv2d(32, 64, 3, padding=1)`. The output is divided into 32-channel spatial fields $\gamma(S)$ and $\beta(S)$, which modulate the final decoder feature F as

$$F' = [1 + \gamma(S)] \odot F + \beta(S). \quad (5)$$

The final SSM convolution has zero-initialized weights and biases, giving $F' = F$ at initialization. Before modulation, the prior is bilinearly resized with `align_corners=False` to the feature-map dimensions; this is effectively a no-op in the present pipeline because both are 256×256 . SSM adds 18,816 parameters, producing a total model size of 1.966 million parameters. Because the backbone contains no batch, instance, or layer normalization, an SSM head driven by a spatially constant input can still learn a global per-channel affine transform of the final decoder feature. The zero-prior control defined in Section 3.3 therefore bounds this effect as well as the generic benefit of an auxiliary branch.

3.3. Mask-Oracle and Falsification Controls

DesnowOracle evaluates five inputs through the same SSM head. This holds the SSM architecture fixed for the estimated, zero, shuffled, dilated-mask, and exact-mask conditions. Concat, FiLM, and SSM share the backbone but have different conditioning-head capacities and are therefore compared as complete conditioning mechanisms, not as perfectly parameter-matched tests of spatial versus global information. The five SSM inputs are:

1. **Estimated prior:** The deterministic severity map S defined by Eqs. (1)–(3).
2. **Zero prior:** An all-zero map with the same spatial dimensions as S , used to determine whether the conditioning path provides a benefit without informative spatial guidance.
3. **Shuffled prior:** Within each manifest, row idx receives the prior from $rows[(idx+7919) \bmod N]$. The mapping is formed separately for training and testing. For the reported manifest sizes, the fixed offset cannot produce a self-match. This control tests a model trained and evaluated with mismatched maps; it is distinct from substituting a map at test time into a model trained with the estimated map.
4. **Dilated oracle:** The ground-truth snow mask dilated once using a 31×31 square structuring element, producing an effective dilation radius of 15 pixels. This condition tests the restoration's sensitivity to localization precision.
5. **Mask oracle:** The normalized Snow100K mask. The source files contain grayscale compositing amplitude; values are scaled to $[0,1]$ before SSM. This target-derived input is used only as an architecture-specific, synthesis-privileged diagnostic reference.

The mask-oracle and dilated-oracle inputs are resized using nearest-neighbor interpolation (`cv2.INTER_NEAREST`) to preserve their discrete spatial support. The estimated prior is computed directly from the snowy image after it has been resized. The shuffled prior is resized to the target dimensions using bilinear interpolation. The mask oracle is an architecture-specific diagnostic reference for the fixed DesnowOracle backbone, not a theoretical upper bound applicable to all desnowing methods. It is also privileged in a stronger sense than the word oracle usually implies: in Snow100K the mask is the compositing stencil used to synthesize each snowy image from its clean counterpart, so supplying it discloses part of the data-generating process rather than merely indicating where snow lies. The resulting headroom is consequently inflated relative to what any estimator operating on the snowy image could attain, and no equivalent quantity exists for real snowfall. Section 5.3 discusses the consequences for interpreting the recovered fraction and proposes a non-leaking alternative.

For PSNR, the recovered mask-oracle fraction is defined as

$$\rho = \frac{Q_{est} - Q_{base}}{Q_{oracle} - Q_{base}}, \quad (6)$$

where Q_{est} , Q_{base} , and Q_{oracle} are mean PSNR values from the same test manifest and matched seeds. A value of zero indicates no recovery beyond Base, and one indicates a gain equal to that of the architecture-specific mask diagnostic. The reported uncertainty is approximate because the available seed summary does not support covariance-aware propagation of this correlated ratio; the component PSNR differences are therefore emphasized alongside the percentage.

The standard reconstruction objective is

$$\mathcal{L}_1 = \frac{1}{3HW} \sum_{p,c} |\hat{Y}_c(p) - Y_c(p)| \quad (7)$$

and the severity-weighted alternative is

$$\mathcal{L}_{SW} = \frac{1}{3HW} \sum_{p,c} [1 + S(p)] |\hat{Y}_c(p) - Y_c(p)|, \quad (8)$$

where $\hat{Y}_c(p)$ and $Y_c(p)$ are the predicted and ground-truth intensities, respectively, for color channel c at pixel p , and H and W are the image height and width. All mask-oracle-gap comparisons use the standard $\mathcal{L}_1$ objective so that the spatial conditioning input is the only experimental factor that changes. The severity-weighted objective is evaluated separately to

determine whether the estimated prior is more effective as an explicit conditioning input or as a pixelwise loss weight.

3.4. Dataset and Training Protocol

The audited Snow100K tree contains 50,000 aligned synthetic training triples comprising a snowy image, clean image, and snow mask; 16,611 small-particle (S), 16,588 medium-particle (M), and 16,801 large-particle (L) test triples; and 1,329 unpaired real snowy images without clean references [2]. The fixed manifests select 2,500 training triples and 2,000 test triples, comprising 666 S, 667 M, and 667 L images. The official training directory is labeled only as “all” and provides no S/M/L assignments; therefore, only the locked test evaluation is stratified by particle size. This subsampling makes the ten-condition, three-seed experimental matrix computationally tractable, and all comparisons use the same manifests.

Every condition is trained independently using seeds 42, 1337, and 2024 for 100 epochs with a batch size of 88, eight data-loading workers, and an input resolution of 256×256 pixels. RGB images are resized using `cv2.INTER_AREA`, masks are resized using `cv2.INTER_NEAREST`, and no cropping or data augmentation is applied. Training uses AdamW with a constant learning rate of 2×10^{-4} , weight decay of 0.01, no learning-rate scheduler, and automatic mixed precision [27]. All experiments run on one NVIDIA GeForce RTX 4080 GPU with 16 GB of memory using CUDA 12.1, PyTorch 2.5.1, and Python 3.10.20. Each run stores its resolved configuration, training log, final checkpoint, and evaluation metrics in a condition-specific and seed-specific results directory. The 2,500-image manifest yields 29 batches per epoch and 2,900 optimizer steps. No validation set or early stopping is used; final-epoch checkpoints are evaluated, and the test manifest is not used for checkpoint selection. The extended protocol uses a 10,000-image manifest containing the original 2,500 rows, 200 epochs, 114 batches per epoch, and 22,800 optimizer steps (approximately 7.9 times the short protocol), with warmup and cosine decay. Because data volume, duration, and scheduling all change, this is a protocol-sensitivity check rather than a convergence-isolation experiment.

3.5. Evaluation Metrics

Synthetic test results use RGB PSNR, SSIM [28], MAE, and LPIPS [29]. PSNR, SSIM, and MAE are computed on RGB images in $[0, 1]$. SSIM uses `scikit-image` 0.25.2 with `data_range=1.0`, a uniform 7×7 window, $K_1 = 0.01$, $K_2 = 0.03$, sample-covariance normalization, and no Gaussian weighting. LPIPS uses package version 0.1.4 with the AlexNet backbone and inputs rescaled to $[-1, 1]$. Results are reported as the mean $\pm$ sample standard deviation across the three training seeds; lower values are better for MAE and LPIPS.

Because the 1,329 real images have no clean references, they are evaluated only using NIQE [30], BRISQUE [31], and PIQE [32], implemented with `pyiqa=0.1.16`; lower values indicate better predicted quality. Complexity is reported as parameter count and GFLOPs, with THOP multiply-accumulate counts converted using two floating-point operations per multiply-accumulate. Latency at 256×256 uses 30 warm-up iterations, 100 timed iterations, and five repeats. Conditioned-model latency includes CPU-side prior construction and host-to-device transfer. These no-reference indices estimate perceptual quality, not verified snow-removal fidelity. Detector-response counts are reported only as an unlabelled proxy and are not detection accuracy. Network GFLOPs exclude prior construction; conditioned latency includes it.

4. RESULTS AND ANALYSIS

The results first establish the short-protocol conditioning and prior-content effects, then test their stability under the extended protocol. Additional analyses examine scale, diagnostic oracles, an exploratory mask-degradation ladder, unpaired real images, and computational cost.

4.1. Prior Agreement and Scale-Stratified Quality

Table 2 shows that the positive local-contrast residual supplies most of the map's agreement with the Snow100K mask. Relative to the contradict channel alone, the single-scale response raises AUC from 0.7883 to 0.8678 and best IoU from 0.1086 to 0.1883. Four-scale aggregation changes AUC by +0.0016 and IoU by −0.0003. AUC evaluates pixel ranking across thresholds, whereas best IoU evaluates overlap at an optimized threshold; they are related but not interchangeable.

For Table 2, the IoU threshold is optimized separately for each image against its ground-truth mask, and the resulting best IoU and threshold are averaged. These are diagnostic oracle thresholds, not a deployable calibration procedure. On the locked test set, M has the highest AUC and IoU. The mean oracle threshold decreases from 0.4213 for S to 0.2576 for L, consistent with the map's lack of cross-image calibration.

Table 2. Prior agreement. The component ablation uses the training split; the S/M/L rows use the locked estimator evaluated on the corresponding test subset.

Split/component	AUC	Best IoU	Mean best threshold
Train, contradict channel only	0.7883	0.1086	0.7030
Train, single scale ($\sigma = 3$)	0.8678	0.1883	0.3331
Train, proposed multi-scale	0.8694	0.1880	0.3485
Test S, locked estimator	0.8745	0.1431	0.4213
Test M, locked estimator	0.8903	0.2122	0.3743
Test L, locked estimator	0.8380	0.1992	0.2576

Table 3 confirms that particle scale affects restoration difficulty: Base PSNR decreases by 5.307 dB from S to L. Severity-weighted SSM gains 0.857, 1.675, and 0.956 dB on S, M, and L. The largest gain and strongest mask agreement both occur on M. This is an association within the experiment, not evidence that localization alone causes the restoration difference.

The mask oracle improves average PSNR from 26.425 to 34.539 dB, exposing 8.114 dB of architecture-specific headroom. However, this is a diagnostic reference for the fixed backbone, not a universal upper bound. Because the best deployable model in Table 3 uses severity-weighted loss while the oracle uses ordinary L1, the matched-L1 results in Table 5 provide the valid oracle-gap calculation.

Table 3. PSNR / SSIM by particle-size subset, mean $\pm$ standard deviation over three seeds.

Metric	Particle scale	Base, L1	SSM, severity-weighted L1	SSM, ground-truth mask oracle
PSNR	S	28.972 $\pm$ 0.251	29.829 $\pm$ 0.038	36.735 $\pm$ 0.121
	M	26.640 $\pm$ 0.233	28.315 $\pm$ 0.063	35.157 $\pm$ 0.133
	L	23.665 $\pm$ 0.302	24.621 $\pm$ 0.023	31.729 $\pm$ 0.139
	Average	26.425 $\pm$ 0.250	27.587 $\pm$ 0.022	34.539 $\pm$ 0.130
SSIM	S	0.8679 $\pm$ 0.0077	0.8918 $\pm$ 0.0010	0.9427 $\pm$ 0.0008
	M	0.8428 $\pm$ 0.0072	0.8754 $\pm$ 0.0017	0.9307 $\pm$ 0.0010
	L	0.7600 $\pm$ 0.0095	0.7950 $\pm$ 0.0023	0.8845 $\pm$ 0.0019
	Average	0.8235 $\pm$ 0.0081	0.8541 $\pm$ 0.0017	0.9193 $\pm$ 0.0012

4.2. Factorial and Falsification Tests

Table 4 separates conditioning from loss weighting. Applying severity-weighted L1 to Base changes PSNR by only -0.009 dB and produces mixed changes in SSIM, MAE, and LPIPS. In contrast, estimated SSM with ordinary L1 improves Base by 1.004 dB PSNR, 0.0273 SSIM, 11.4% MAE, and 22.5% LPIPS. Spatial conditioning therefore provides the principal improvement.

Adding severity-weighted L1 to SSM yields a further 0.158 dB PSNR gain and improves every reported metric under the short protocol. Estimated SSM with L1 also exceeds Concat and FiLM by 0.529 and 0.618 dB. Because the conditioning heads differ in capacity and information flow, these results compare complete mechanisms rather than isolating spatial versus global information.

Table 4. Factorial comparison of conditioning mechanism and loss, reported as mean $\pm$ standard deviation over three seeds

Architecture/loss	PSNR (dB)	SSIM	MAE	LPIPS
Base, L1	26.425 $\pm$ 0.250	0.8235 $\pm$ 0.0081	0.0307 $\pm$ 0.0011	0.1652 $\pm$ 0.0106
Base, severity-weighted L1	26.416 $\pm$ 0.563	0.8239 $\pm$ 0.0179	0.0313 $\pm$ 0.0028	0.1636 $\pm$ 0.0212
SSM, L1	27.429 $\pm$ 0.082	0.8508 $\pm$ 0.0036	0.0272 $\pm$ 0.0004	0.1281 $\pm$ 0.0038
SSM, severity-weighted L1	27.587 $\pm$ 0.022	0.8541 $\pm$ 0.0017	0.0269 $\pm$ 0.0003	0.1239 $\pm$ 0.0027
Concat, L1	26.900 $\pm$ 0.151	0.8366 $\pm$ 0.0074	0.0291 $\pm$ 0.0007	0.1488 $\pm$ 0.0093
FiLM, L1	26.811 $\pm$ 0.185	0.8359 $\pm$ 0.0053	0.0294 $\pm$ 0.0006	0.1480 $\pm$ 0.0066

Table 5 separates the conditioning-branch and prior-content effects under the short protocol. Zero and shuffled inputs improve Base by 0.514 and 0.487 dB, showing that the SSM branch can help without informative spatial input. The estimated SSM adds 0.490 dB over zero and 0.517 dB over shuffled in mean PSNR. Thus, only part of its 1.004 dB gain over Base is associated with the estimated map's content.

Table 5. Short-protocol matched-L1 mask-diagnostic controls, mean $\pm$ sample SD over three seeds. Recovered headroom uncertainty is approximate because Eq. (6) contains correlated quantities.

Prior supplied to SSM	PSNR (dB)	SSIM	LPIPS	Oracle headroom recovered
Estimated	27.429 $\pm$ 0.082	0.8508 $\pm$ 0.0036	0.1281 $\pm$ 0.0038	12.4 $\pm$ 1.7%
Zero	26.939 $\pm$ 0.384	0.8417 $\pm$ 0.0092	0.1376 $\pm$ 0.0083	6.3 $\pm$ 3.2%
Shuffled	26.912 $\pm$ 0.310	0.8413 $\pm$ 0.0074	0.1400 $\pm$ 0.0083	6.0 $\pm$ 2.8%
Dilated oracle	27.905 $\pm$ 0.135	0.8519 $\pm$ 0.0013	0.1253 $\pm$ 0.0062	18.2 $\pm$ 1.8%
Mask oracle	34.539 $\pm$ 0.130	0.9193 $\pm$ 0.0012	0.0775 $\pm$ 0.0009	100% (by definition)

Estimated SSM recovers 12.4% of the mask-diagnostic headroom under the short protocol, compared with 6.3% for zero and 6.0% for shuffled input. The corresponding absolute gains over Base are 1.004, 0.514, and 0.487 dB. Recovery is highest on M, but the percentage depends on both the gain numerator and the subset-specific oracle denominator.

The same three seeds are paired across conditions. Their mean PSNR differences favor estimated SSM over zero and shuffled on all three seeds, but $n=3$ provides weak seed-level inference. Because aligned per-image outputs are available, each condition's prediction was averaged over the three seeds before calculating one paired difference per unique test image. Conditional on these trained models, estimated SSM has a median PSNR advantage of 0.387 dB over zero (95% BCa CI 0.354–0.415) and 0.420 dB over shuffled (0.390–0.457), with two-sided Wilcoxon $p<10^{-130}$ for both comparisons. A separate dependence check uses the three

estimated-map SSM checkpoints and changes only their test-time input. After averaging each image's drop across seeds, median PSNR losses are 4.247 dB for zero input (95% BCa CI 4.114–4.409), 4.754 dB for a different-image map (4.596–4.919), 4.647 dB for a different image from the same particle stratum (4.517–4.749), and 3.626 dB for a horizontally flipped map (3.546–3.724). Halving and doubling the aligned map give smaller median losses of 1.375 and 2.659 dB. These substitution results show that the trained checkpoints depend on their expected input; they do not establish that training with estimated content is superior across protocols. All image-level intervals are conditional on the three trained models and do not replace uncertainty over independently trained models.

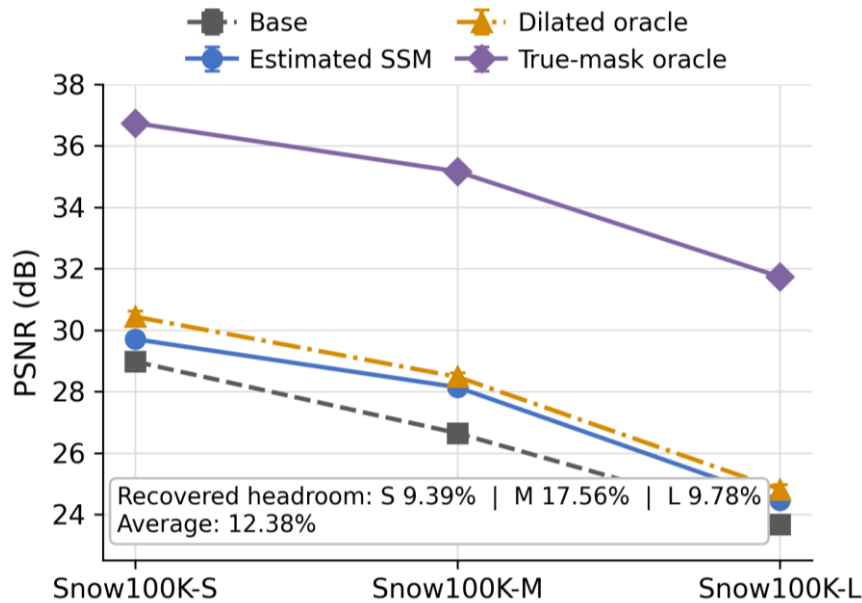


Figure 3. Scale-stratified mask-oracle gap, reported as mean $\pm$ standard deviation over three 100-epoch seeds. Estimated SSM benefits from image-aligned spatial information, while the exact mask oracle exposes substantial unrecovered headroom.

4.3. Real-World Images and Computational Complexity

Table 6 summarizes the original real-image comparison between Base and the best short-protocol variant; Table 8 extends it to matched zero and shuffled controls. Lower no-reference scores indicate improved predicted naturalness, but NIQE, BRISQUE, and PIQE do not measure snow-removal fidelity. Verified raw snowy-input values were not produced by the no-reference pipeline, so that comparison is unavailable.

Table 6. No-reference quality metrics on 1,329 real images, mean $\pm$ standard deviation over three seeds.

Model	NIQE	BRISQUE	PIQE
Base	5.5284 $\pm$ 0.0864	18.7490 $\pm$ 0.4902	32.6003 $\pm$ 0.2292
Estimated SSM, severity-weighted $\mathcal{L}_1$	5.1588 $\pm$ 0.0759	15.8150 $\pm$ 0.4094	31.8118 $\pm$ 0.3511

Figure 4 provides fixed-seed qualitative examples selected before inference. The examples illustrate the quantitative pattern, with estimated SSM remaining visually closer to Base than to the mask oracle. They support interpretation but do not replace the complete test-set evaluation.

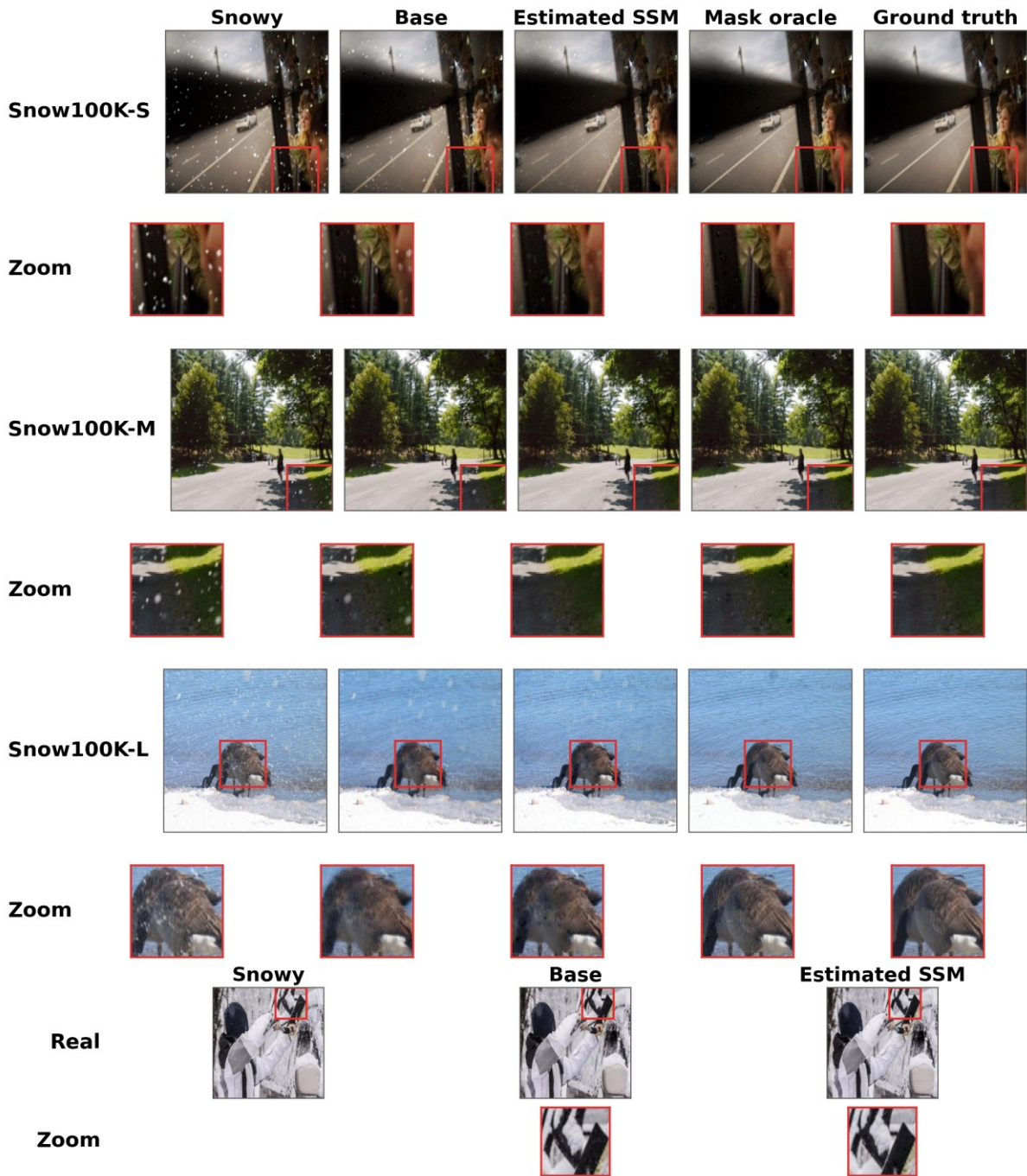


Figure 4. Fixed-seed qualitative comparison. Synthetic rows show the snowy input, Base output, estimated-SSM output, mask-oracle output, and clean target. The real-image row has neither a mask nor a clean reference. Each synthetic and real row is followed by a red-boxed zoomed crop of the same region (chosen as the highest-difference window against the clean target, or against Base for the real row) so the restoration difference between conditions is legible at print size.

Table 7 shows that SSM adds 18,816 parameters (0.97% over Base) and 2.45 network GFLOPs (11.5%). The GFLOP values exclude prior construction. Complete conditioned latency rises from 0.750 to 4.488 ms, mainly because the CPU multiscale prior costs about 3.5 ms. The SSM modulation head alone is not responsible for this increase. SSM adds approximately 0.017 GB of batch-1 inference memory and 1.423 GB of batch-88 training memory relative to Base.

Each training run, including periodic evaluation, requires 740 to 836 seconds at 310 to 352 images per second. The complete 30-run matrix requires approximately 6.6 GPU-hours on the RTX 4080. These results establish experimental tractability, but the sub-5-ms latency measured on a desktop GPU should not be interpreted as evidence of embedded-device performance.

Table 7. Computational complexity at 256×256 resolution. Conditioned-model latency includes prior construction and transfer; Base latency excludes them.

Model	Parameters (M)	GFLOPs	Latency (ms)	Inference memory (GB)	Training memory (GB)
Base	1.947	21.27	0.750±0.032	0.077	11.151
Concat	1.947	21.30	4.381±0.017	0.078	11.269
FiLM	1.951	21.27	4.462±0.015	0.085	11.524
SSM	1.966	23.72	4.488±0.028	0.094	12.574
SSM, GPU-fused single-scale (deployment)	1.966	23.72	0.941±0.001	0.094	12.574

The real-image evaluation was extended in two ways after the original protocol. First, `scripts/evaluate_real.py` was generalized to respect each checkpoint's own training-time prior mode, in which it had defaulted every real-subset evaluation to the estimated prior regardless of what a checkpoint was actually trained with, which would have silently performed a test-time swap on the zero- and shuffled-prior models instead of evaluating their matched control condition. Table 8 reports the corrected comparison.

Table 8. Real-image no-reference metrics extended to the zero- and shuffled-prior controls, mean ± standard deviation over three seeds, 1,329 images; lower is better for all three indices.

Model	NIQE	BRISQUE	PIQE
Base	5.528±0.086	18.749±0.490	32.600±0.229
SSM, zero prior	5.308±0.045	17.073±0.312	32.694±0.201
SSM, shuffled prior	5.302±0.038	16.858±0.276	32.688±0.187
SSM, estimated (matched L1)	5.237±0.041	15.973±0.298	32.003±0.215
SSM, estimated, severity-weighted L1 (best variant)	5.159±0.076	15.815±0.409	31.812±0.351

The estimated map scores are better than zero and shuffled controls on all three real-image no-reference indices. A fixed pretrained detector was also applied to raw and restored images, and high-confidence responses were counted for person, car, truck, and traffic light. With no ground-truth annotations, these counts are a detector-response proxy: additional responses may be true or false positives and do not establish improved perception.

The detector-response totals differ between Base and the conditioned models, but estimated, zero, and shuffled inputs are not clearly separated. Because the same images and seeds are repeated, the earlier pooled image-seed p-values are not used. The detector proxy is reported descriptively and does not support a prior-content advantage.

A GPU-native PyTorch implementation of the same multiscale estimator matches the OpenCV reference within 1.73×10^{-6} over 200 test images. A single-scale variant is faster, but restoration performance was not evaluated after replacing the multiscale map. Table 7 therefore reports its latency as an implementation benchmark, not a recommended deployment default. The 79% reduction from 4.488 to 0.941 ms concerns the complete conditioned inference path and is driven mainly by removing CPU prior construction and transfer.

Table 9. Unlabelled detector-response counts across 1,329 real images. Counts are descriptive proxies, not detection accuracy; no independent-observation p-values are reported.

Condition	Person	Car	Truck	Comparison with Base	Comparison within SSM	vs. zero/shuffled, Wilcoxon p
Raw snowy input	2223	1882	247	Descriptive only	Descriptive only	--
Base	2218.0	1848.7	241.3	Descriptive only	Descriptive only	--
SSM, zero prior	2223.3	1858.3	257.0	Descriptive only	Descriptive only	--
SSM, shuffled prior	2229.3	1871.3	256.7	Descriptive only	Descriptive only	--
SSM, estimated	2229.7	1846.7	275.3	Descriptive only	Descriptive only	0.30 / 0.40 (not significant)

4.4. Scale-Stratified Falsification Controls

Table 10 shows that the estimated prior consistently exceeds zero and shuffled controls across S, M, and L, supporting a spatial-alignment contribution under the short-budget protocol. However, the absolute margins are modest: the estimate exceeds zero by 0.349, 0.764, and 0.357 dB, respectively. The larger M recovery fraction partly reflects its subset-specific oracle denominator and cannot be interpreted as stronger causal sensitivity to medium particles. Zero and shuffled conditions reproduce the same S/M/L ordering, indicating that baseline restoration difficulty and conditioning-branch behavior also shape recovery. Dilating the oracle improves all subsets but remains far below masking, showing that localization precision matters.

Table 10. Scale-stratified matched-L1 falsification controls, mean over three seeds. Headroom recovered uses Eq. (6) per subset, with Q_{base} and Q_{oracle} taken from the same subset.

Prior supplied to SSM	S PSNR (dB)	M PSNR (dB)	L PSNR (dB)	S recovered	M recovered	L recovered
Estimated	29.701±0.144	28.136±0.043	24.453±0.073	9.39%	17.56%	9.78%
Zero	29.352±0.408	27.372±0.468	24.096±0.297	4.89%	8.59%	5.35%
Shuffled	29.340±0.277	27.311±0.418	24.087±0.239	4.73%	7.88%	5.23%
Dilated oracle	30.433±0.192	28.476±0.132	24.810±0.169	18.81%	21.55%	14.20%
Mask oracle	36.735±0.121	35.157±0.133	31.729±0.139	100%	100%	100%

4.5. Target-Referenced Diagnostic Oracles and Exploratory Support/Amplitude Ablations

Table 11 reveals that restoration benefit depends not only on locating snow but also on the numerical meaning and distribution of the conditioning signal. Binary support yields a 4.610 dB gain, far above the estimated map's 1.004 dB, whereas pairing correct support with wrong-image amplitude reduces the gain to 1.067 dB. This contrast suggests that inaccurate amplitude can actively misdirect modulation; however, because each condition is trained independently, the difference also includes optimization variability and cannot causally isolate an amplitude effect. Severity on dilated support provides only 1.480 dB, indicating that approximate coverage does not substitute for precise support.

The clean-referenced residual exceeding the exact mask by 1.273 dB does not establish a superior oracle: it encodes target discrepancy closely aligned with the reconstruction objective and therefore leaks richer information. Accordingly, interpret these conditions as separate stress tests, not a ranked continuum of deployability or additive evidence. The 12.4% mask-gap recovery remains descriptive because its uncertainty depends on correlated Base,

estimated, and mask results. Also, remove the duplicated mask-oracle row and verify the uncertainty labels: the displayed values appear to be standard errors rather than sample standard deviations.

Table 11. Target-referenced diagnostic oracles and exploratory support/amplitude ablations, matched L1, mean $\pm$ sample SD over three seeds.

Condition	PSNR (dB)	Δ vs. Base (dB)	Interpretation	Mask-gap recovery
Base	26.425 $\pm$ 0.145	--	Unconditioned reference	--
Estimated map	27.429 $\pm$ 0.047	+1.004	Snowy-image-derived input	12.4%
Binary mask support	31.035 $\pm$ 0.056	+4.610	Amplitude discarded	--
Severity on dilated support	27.905 $\pm$ 0.078	+1.480	Imprecise support	--
True support, wrong-image amplitude	27.492 $\pm$ 0.094	+1.067	Mismatched amplitude	--
Clean-referenced residual	35.812 $\pm$ 0.060	+9.388	Target-leaking diagnostic	--
Exact Snow100K mask	34.539 $\pm$ 0.075	+8.115	Synthesis-privileged diagnostic	100%
Mask oracle (privileged reference)	34.539 $\pm$ 0.075	+8.115	100%	denominator (100% by construction)

4.6. Exploratory Association Between Mask Agreement and Restoration Gain

A controlled ladder evaluates whether restoration gain changes with mask agreement within a single target-derived corruption family. The 15 fitted perturbations comprise dilation at 2, 5, 10, and 25 pixels, erosion at 2 and 5 pixels, Gaussian blur at $\sigma=1, 2$, and 4, false-positive rates of 0.05, 0.10, and 0.20, and false-negative rates of 0.10, 0.25, and 0.50. The unperturbed mask is retained only as a consistency check and is not fitted. Dilation at 15 pixels is held out. Each rung trains separate SSM models across three seeds under the short protocol. Restoration is modeled as

$$\Delta PSNR(Q) = a \cdot Q + b \quad (9)$$

where Q is AUC or best IoU. Regression on the 15-rung means avoids treating three initializations as independent observations. However, the rungs remain statistically related because they are generated from the same masks, images, and nested corruption levels. The reported ordinary least-squares intervals therefore depend on independence, linearity, and constant-variance assumptions that have not been verified. Within this restricted ladder, AUC is the stronger linear correlate, with a slope of 13.95 dB per unit and $R^2 = 0.817$, compared with 5.79 dB and $R^2 = 0.413$ for IoU. These results do not establish a transferable prediction rule.

The held-out dilation point falls within one residual SD of the AUC fit, but the estimated map falls 4.52 SD below it. Its scene-correlated errors differ from the geometric and randomized perturbations used to construct the ladder. The clean-referenced residual falls 2.54 SD above the fit because it contains target information that mask agreement cannot represent. These opposing deviations identify corruption type and information content as omitted factors. Table 12 and Figure 5 therefore support only an exploratory within-family association, not a general transfer function.

Table 12. Exploratory mask-agreement association and residuals from the AUC fit. The estimated and clean-referenced maps are out-of-family observations, not validation points.

Prior	AUC	Best IoU	Δ PSNR (dB)	Residual vs. AUC fit	Residual vs. IoU fit
Estimated (deployable)	0.868	0.185	1.004	-4.80 dB (-4.52 SD)	-3.25 dB (-1.71 SD)
Clean-image ceiling	0.931	0.317	9.388	+2.70 dB (+2.54 SD)	+4.37 dB (+2.30 SD)
Mask oracle (consistency check, $r = 0$)	1.000	1.000	8.115	+0.46 dB (+0.43 SD)	-0.86 dB (-0.45 SD)
Held-out ladder point (dilation $r = 15$)	0.621	0.035	1.480	-0.88 dB (-0.83 SD)	-1.91 dB (-1.00 SD)

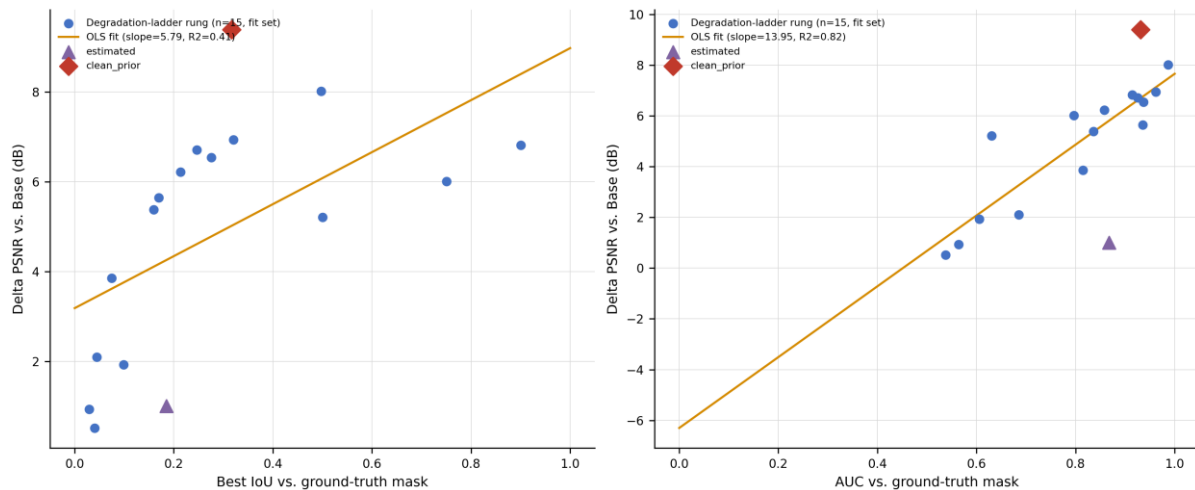


Figure 5. Fitted prior-quality transfer function (restoration gain vs. mask AUC) over the 15-rung degradation ladder, with the estimated prior and the clean-image ceiling located relative to the fitted line and its confidence band.

4.7. Extended-Training Protocol Sensitivity Check

The core comparison was repeated with 10,000 training images, 200 epochs, approximately 22,800 optimizer steps, and warmup followed by cosine learning-rate decay. The short protocol used 2,500 images, 100 epochs, approximately 2,900 steps, and a constant learning rate. Because data volume, training duration, and optimization schedule changed simultaneously, this experiment evaluates sensitivity to the combined protocol rather than convergence alone.

Table 13. Extended-protocol PSNR (10,000 training images, 200 epochs, warmup and cosine decay), mean $\pm$ sample SD over three seeds.

Condition	Extended-protocol PSNR (dB)	Δ vs. Base (dB)
Base	28.976 $\pm$ 0.021	--
Base + learnable per-channel affine (no conditioning)	28.986 $\pm$ 0.029	+0.010
SSM, zero prior	29.412 $\pm$ 0.123	+0.436
SSM, shuffled prior	29.463 $\pm$ 0.018	+0.487
SSM, estimated (deployable)	29.303 $\pm$ 0.078	+0.327
SSM, clean-image target-leaking diagnostic	38.900 $\pm$ 0.022	+9.924
SSM, mask oracle	37.187 $\pm$ 0.206	+8.211

Base plus a learned output affine transform reaches 28.986 dB, only 0.010 dB above Base. This near-equivalence indicates that a simple global affine adjustment does not explain the

SSM advantage. All SSM conditions exceed Base, which supports a benefit from the spatial conditioning architecture. However, the estimated input reaches 29.303 dB and falls below both zero at 29.412 dB and shuffled at 29.463 dB. The ordering observed under the short protocol therefore does not persist.

After averaging per-image results across seeds, estimated-minus-zero has a median difference of -0.111 dB, while estimated-minus-shuffled reaches -0.141 dB. Their bootstrap intervals remain below zero, conditional on the three trained models. These results show that the estimated map's content offers no advantage over uninformative controls under the extended protocol. They do not identify which protocol change caused the reversal. Isolating duration would require keeping the original 2,500-image manifest, optimizer, and learning-rate schedule fixed while training continues to a prespecified plateau criterion.

5. DISCUSSION

DesnowOracle's central result is protocol sensitivity. The conditioning branch improves on Base under both reported protocols, but the estimated map's ranking relative to zero and shuffled inputs reverses when data volume, optimizer steps, and learning-rate scheduling change together.

5.1. Particle Scale, Mask Agreement, and Conditioning Utility

Under the short protocol, the M subset combines the strongest mask agreement with the largest estimated-map improvement. However, zero and shuffled inputs show the same S, M, and L recovery ordering. The M advantage therefore cannot be attributed solely to better localization. It may also reflect subset difficulty, optimization behavior, or the smaller relationship between Base performance and the subset-specific mask-diagnostic denominator.

The S and L subsets expose different potential limitations. For S, relatively high AUC but lower IoU indicates that the map ranks many relevant pixels correctly without producing precise thresholded support around fine particles and bright scene structures. For L, large and opaque particles remove more underlying information, increasing dependence on contextual reconstruction rather than localization alone. These explanations are plausible but remain untested because particle scale, opacity, scene content, and restoration difficulty are not independently manipulated.

The 12.4% recovery estimate must also be interpreted cautiously. It is defined relative to a synthesis-privileged, architecture-specific mask diagnostic, not universally attainable performance. The absolute 1.004 dB gain over Base is more transparent, while the additional gains over zero and shuffled controls are only 0.490 and 0.517 dB. Moreover, the estimated map falls below both controls under the extended protocol. The evidence therefore supports protocol-dependent conditioning utility, not a general conclusion that accurate spatial guidance is consistently most valuable on medium-scale snow.

5.2. Contribution Beyond Prior-Guided Desnowing

Prior-guided desnowing is well established. DesnowNet predicts an intermediate snow mask [2], LMQFormer converts a learned coarse mask into restoration queries [9], and the deep unfolding approach uses a snow-shape prior as surrogate supervision [10]. These studies show that prior information can support restoration, but their architectures do not isolate spatial-content utility by supplying the same conditioning head with estimated, zero, shuffled, degraded, and exact-mask inputs. DesnowOracle contributes to this controlled attribution structure rather than another prior-guided architecture alone.

Under the short protocol, estimated SSM exceeds Base, zero, and shuffled conditions. The corrected unique-image analysis supports these differences for the three trained models. Test-time substitutions also produce substantial degradation, showing that each trained checkpoint depends on the input distribution it encountered during training. This dependence does not necessarily reflect useful prior content because replacing the expected map creates a distribution shift. The separately trained controls address a different question by testing whether estimated content improves the final trained condition. Under the extended protocol, estimated input falls below zero and shuffled inputs, although every SSM condition remains above Base. The combined evidence supports an architectural conditioning benefit, but it does not support a protocol-invariant benefit from the estimated map.

The severity-weighted L1 objective adds 0.158 dB to the estimated SSM and improves all reported metrics under the short protocol. Because this interaction was not repeated under the extended protocol, it should be treated as protocol-specific. The same caution applies to the scale analysis. SnowMamba establishes the value of S, M, and L reporting [13], while DesnowOracle relates these subsets to mask agreement and diagnostic headroom. However, shared scale ordering across estimated, zero, and shuffled controls shows that localization quality alone does not explain the pattern.

Absolute PSNR comparisons with LMQFormer are also inappropriate because its learned prior, architecture, training data, and evaluation protocol differ [9]. TDesnowOracle's defensible contribution is a framework that separates conditioning-branch effects, training-time input dependence, prior-content utility, privileged diagnostic headroom, and protocol sensitivity.

Real-image findings further limit broad claims. No-reference indices favor estimated SSM, but they measure predicted naturalness rather than paired restoration fidelity. Unlabeled detector counts do not distinguish estimated input from the controls and cannot separate correct detections from false positives. Synthetic fidelity, no-reference quality, and detector response therefore provide complementary but non-equivalent evidence. Together, they indicate that the value of prior content depends on the training protocol and evaluation criterion.

5.3. Scope, Limitations, and Research Direction

The evidence is bounded by three training seeds and restricted Snow100K subsets, which limit inference about initialization variability and full-data performance. The absence of a validation set, early stopping, and a prespecified plateau criterion also prevents a rigorous claim of convergence. Moreover, the extended protocol changes data volume, training duration, and learning-rate scheduling together. The observed reversal therefore demonstrates protocol sensitivity but cannot identify its cause. A controlled follow-up should retain the original manifest, optimizer, and scheduler; vary only training duration; define convergence in advance; and use additional seeds. The Base-plus-affine control excludes one simple architectural explanation under the extended protocol, but it does not explain the short-protocol ordering.

Real Snow100K images provide neither paired clean targets nor detector annotations. Consequently, no-reference indices and detection counts remain indirect proxies rather than evidence of restoration fidelity or perception accuracy. The exact mask and clean-referenced residual are target-derived diagnostics that cannot represent deployable performance. The regression ladder covers only geometric and randomized mask perturbations, while the estimated map contains scene-correlated errors, so external validity remains unknown. Finally, RTX 4080 latency does not establish embedded feasibility. The faster single-scale

implementation requires restoration testing and device-specific benchmarking before deployment claims are justified.

6. CONCLUSIONS

DesnowOracle demonstrates that compact spatial conditioning can improve image desnowing, but the benefit attributable to estimated prior content depends on the training protocol. The SSM head adds only 18,816 parameters to a 1.966-million-parameter model. Under the short protocol, estimated SSM improves PSNR by 1.004 dB over Base and by 0.490 and 0.517 dB over zero and shuffled controls, respectively, recovering 12.4% of the architecture-specific mask-diagnostic gap. Under the extended protocol, estimated SSM remains 0.327 dB above Base but falls 0.109 dB below zero and 0.160 dB below shuffled input. This reversal establishes sensitivity to the combined changes in data volume, training duration, and learning-rate scheduling, not a convergence mechanism. On 1,329 unpaired real images, no-reference indices favor the estimated map, while unlabeled detector responses show no advantage over the controls, leaving real-world fidelity unresolved. A GPU-native single-scale implementation reduces conditioned latency from 4.488 to 0.941 ms, although restoration equivalence and embedded performance remain unverified. For prior-guided restoration research, these findings show that conditioning-branch gains must be separated from prior-content gains using matched controls, unique-image inference, and protocol-robust evaluation. Future work should isolate training duration, expand seeds and training data, develop a non-leaking diagnostic reference, validate the faster prior, and evaluate paired real images.

CODE AVAILABILITY

The implementation of the snow-response map, compact backbone, conditioning heads, control conditions, and GPU-native prior is available at <https://github.com/tsgunawan/DesnowOracle> under the MIT License. The repository provides training, evaluation, statistical analysis, and aggregation code, fixed manifests, pinned dependencies, protocol configurations, and figure-regeneration scripts. Snow100K dataset is not redistributed and must be obtained from its original source.

ACKNOWLEDGEMENT

This work is supported by and also conceptually aligned with ongoing research under the AOARD project FA2386-25-1-4061 and recorded at IIUM SPI25-295-0295. The authors used generative artificial intelligence (GenAI) solely to improve language clarity and proofreading. The authors reviewed and verified all GenAI-assisted content and take full responsibility for the accuracy, originality, and integrity of the published work.

REFERENCES

- [1] J. Wang, M. Xu, H. Xue, Z. Huo, and F. Luo, "Joint image restoration for object detection in snowy weather," *IET Computer Vision*, vol. 18, no. 6, pp. 759-771, 2024.
- [2] Y.-F. Liu, D.-W. Jaw, S.-C. Huang, and J.-N. Hwang, "DesnowNet: Context-aware deep network for snow removal," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3064-3073, 2018.
- [3] W.-T. Chen, Z.-K. Huang, C.-C. Tsai, H.-H. Yang, J.-J. Ding, and S.-Y. Kuo, "Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: Toward a unified model," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022: IEEE, pp. 17632-17641.

- [4] W.-T. Chen, H.-Y. Fang, J.-J. Ding, C.-C. Tsai, and S.-Y. Kuo, "JSTASR: Joint size and transparency-aware snow removal algorithm based on modified partial convolution and veiling effect removal," in *European Conference on Computer Vision*, 2020: Springer, pp. 754-770.
- [5] W.-T. Chen *et al.*, "All snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 4196-4205.
- [6] K. Zhang, R. Li, Y. Yu, W. Luo, and C. Li, "Deep dense multi-scale network for snow removal using semantic and depth priors," *IEEE Transactions on Image Processing*, vol. 30, pp. 7419-7431, 2021.
- [7] S. A. Agyemang, H. Shi, X. Nie, and N. Y. Asabere, "An integrated multi-scale context-aware network for efficient desnowing," *Engineering Applications of Artificial Intelligence*, vol. 151, p. 110769, 2025.
- [8] Z. Chen, Y. Sun, X. Bi, and J. Yue, "Lightweight image de-snowing: A better trade-off between network capacity and performance," *Neural Networks*, vol. 165, pp. 896-908, 2023.
- [9] J. Lin, N. Jiang, Z. Zhang, W. Chen, and T. Zhao, "LMQFormer: A laplace-prior-guided mask query transformer for lightweight snow removal," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 11, pp. 6225-6235, 2023.
- [10] X. Guo, X. Wang, X. Fu, and Z.-J. Zha, "Deep unfolding network for image desnowing with snow shape prior," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 35, no. 5, pp. 4740-4752, 2025.
- [11] S. Chen, T. Ye, Y. Liu, E. Chen, J. Shi, and J. Zhou, "SnowFormer: Scale-aware transformer via context interaction for single image desnowing," *arXiv preprint arXiv:2208.09703*, 2022.
- [12] Y. Zhang, "BAT-Net: Bidirectional Attention Transformer Network for Joint Single-Image Desnowing and Snow Mask Prediction," *Information*, vol. 16, no. 11, p. 966, 2025.
- [13] G. Wang, Y. Zhou, F. Shi, and Z. Jia, "SnowMamba: Achieving more precise snow removal with Mamba," *Applied Sciences*, vol. 15, no. 10, p. 5404, 2025.
- [14] J. Lai *et al.*, "Snowmaster: Comprehensive real-world image desnowing via mllm with multi-model feedback optimization," in *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025: IEEE, pp. 4302-4312.
- [15] Q. Guan, Q. Yang, X. Chen, T. Song, G. Jin, and J. Jin, "Weatherbench: A real-world benchmark dataset for all-in-one adverse weather image restoration," in *Proceedings of the 33rd ACM international conference on multimedia*, 2025, pp. 12607-12613.
- [16] J. Jiang, Z. Zuo, G. Wu, K. Jiang, and X. Liu, "A survey on all-in-one image restoration: Taxonomy, evaluation and future trends," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.
- [17] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, "All-in-one image restoration for unknown corruption," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022: IEEE, pp. 17431-17441.
- [18] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022: IEEE, pp. 2343-2353.
- [19] C. Li, F. Sun, H. Zhou, Y. Xie, Z. Li, and L. Zhu, "Multi-weather restoration: An efficient prompt-guided convolution architecture," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 35, no. 2, pp. 1436-1450, 2024.
- [20] Y. Wen, T. Gao, Z. Li, J. Zhang, K. Zhang, and T. Chen, "All-in-one weather-degraded image restoration via adaptive degradation-aware self-prompting model," *IEEE Transactions on Multimedia*, vol. 27, pp. 3343-3355, 2025.
- [21] D. Xie, X. Hu, Y. Zhou, and S. Duan, "All-in-one adverse weather removal via dual state space-based diffusion model with degradation-aware guidance," *Pattern Recognition*, vol. 171, p. 112081, 2026.
- [22] Y. Cui, S. W. Zamir, S. Khan, A. Knoll, M. Shah, and F. Khan, "Adair: Adaptive all-in-one image restoration via frequency mining and modulation," in *Proceedings of the 13th International Conference on Learning Representations (ICLR)*, 2025.

- [23] E. Perez, F. Strub, H. De Vries, V. Dumoulin, and A. Courville, "FiLM: Visual reasoning with a general conditioning layer," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, vol. 32, no. 1.
- [24] X. Wang, K. Yu, C. Dong, and C. C. Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: IEEE, pp. 606-615.
- [25] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015: Springer, pp. 234-241.
- [26] T. S. Gunawan and M. Kartiwi, "Stability-Aware Physics-Prior Conditioning for Lightweight Image Dehazing," presented at the 2026 IEEE 12th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA 2026), Kuching, Sarawak, Malaysia, 28-30 July 2026, 2026.
- [27] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," presented at the Proceedings of the 13th International Conference on Learning Representations (ICLR), 2019.
- [28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [29] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: IEEE, pp. 586-595.
- [30] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209-212, 2013.
- [31] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, 2012.
- [32] N. Venkatanath, D. Praneeth, S. C. Sumohana, and S. M. Swarup, "Blind image quality evaluation using perception based features," in *2015 Twenty First National Conference on Communications (NCC)*, 2015: IEEE, pp. 1-6.

Compute-Aware Evaluation of Severity Conditioning and Lightweight Architectures for Real-World Image Deraining

TEDDY SURYA GUNAWAN^{1*}, MIRA KARTIWI²

¹*Department of Electrical and Computer Engineering, Kuliyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia*

²*Department of Information Systems, Kuliyah of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia*

**Corresponding author: tsgunawan@iium.edu.my*

(Received: 28 July 2026; Accepted: 29 August 2026; Published online: 12 September 2026)

ABSTRACT: Image deraining for embedded perception must balance restoration quality against complete inference cost. Yet density-conditioned methods require annotations unavailable for captured rain, while efficiency studies often omit conditioning-signal cost. This study examines whether label-free severity conditioning justifies its end-to-end cost and whether complexity measures predict realized latency. Under a common L_1 protocol, we compare a PReNet baseline with scalar feature-wise linear modulation, spatial severity modulation (SSM), a matched-capacity constant-map control, and two lightweight architectures. Zero-, shuffled-, and constant-prior interventions test SSM's dependence on prior presence, image correspondence, and spatial structure. Six models were trained for 100 epochs on 10,000 LHP-Rain patches using three seeds and evaluated on the 1,000-image test split. SSM achieved 31.775 ± 0.109 dB, compared with 31.741 ± 0.095 dB for Base and 31.754 ± 0.102 dB under zero-prior replacement. A seed-level equivalence analysis based on paired image differences supported practical equivalence between SSM and Base (95% CI: -0.026 to 0.094 dB; margin fixed before primary analysis: ± 0.1 dB), consistent with the half-scale experiment. The two-stage model reduced computation from 265.0 to 44.6 GFLOPs and model latency from 20.43 to 4.50 ms, but the prior generation increased total latency to 24.74 ms. CBAM-light required 14.1 GFLOPs yet recorded 33.68 ms. The label-free statistic achieved a cross-scene AUC of 0.772 on RealRain-1k, indicating moderate agreement with density labels. Thus, severity-conditioned and lightweight deraining systems should be assessed through verified prior use and measured end-to-end latency rather than accuracy, parameters, or GFLOPs alone.

ABSTRAK: Penyahhujan untuk sistem terbenam perlu mengimbangi kualiti pemulihan dengan kos inferens menyeluruh. Namun, kaedah yang bergantung pada kekerapan memerlukan anotasi yang tidak tersedia untuk hujan yang dirakam, manakala kajian kecekapan sering kali mengabaikan kos isyarat pengkondisian. Kajian ini meneliti sama ada penentu tahap tanpa label membenarkan kos hujung-ke-hujung dan sama ada ukuran kerumitan meramalkan kelewatan sebenar. Di bawah protokol L_1 , garis dasar PReNet dibandingkan dengan modulasi linear mengikut ciri skalar, modulasi tahap spatial (SSM), kawalan peta tetap berkapasiti sepadan, dan dua seni bina ringan. Intervensi sifar-, diacak-, dan malar menguji kebergantungan SSM pada kewujudan, kesepadanan imej, dan struktur spatial. Enam model dilatih untuk 100 epok menggunakan 10,000 tampalan LHP-Rain dengan tiga benih dan dinilai pada 1,000 imej ujian. SSM mencapai 31.775 ± 0.109 dB, berbanding 31.741 ± 0.095 dB bagi Base dan 31.754 ± 0.102 dB di bawah penggantian sebelum sifar. Analisis kesetaraan tahap biji berdasarkan perbezaan imej berpasangan menyokong kesetaraan praktikal antara SSM dan Base (95% CI: -0.026 hingga 0.094 dB; margin ditetapkan sebelum analisis utama: ± 0.1 dB), ianya konsisten dengan eksperimen separuh skala. Model

dua peringkat mengurangi pengiraan daripada 265.0 kepada 44.6 GFLOP dan model pendaman daripada 20.43 kepada 4.50 ms, tetapi bagi penjanaian sebelum, meningkatkan jumlah pendaman kepada 24.74 ms. CBAM-light memerlukan 14.1 GFLOPs tetapi merekodkan 33.68 ms. Statistik tanpa label mencapai AUC rentas senarai 0.772 pada RealRain-1k, menunjukkan persetujuan sederhana pada label ketumpatan. Jadi, sistem penyahhujan ringan dan bersyarat teruk perlu dinilai melalui penggunaan yang disahkan sebelumnya dan ukuran kelewatan hujung-ke-hujung, bukan pada ketepatan, parameter, atau GFLOPs.

KEYWORDS: *Image Deraining; Severity Conditioning; Lightweight Architectures; End-to-End Latency; Real-World Rain*

1. INTRODUCTION

Rain introduces streaks, veiling, splash, and contrast loss into images acquired by mobile robots, unmanned aerial vehicles, and advanced driver-assistance systems. These degradations can obscure scene structure and compromise downstream perception precisely when reliable visual information is most needed. Because restoration may need to run on the sensing platform, practical deraining requires more than high peak signal-to-noise ratio (PSNR) or structural similarity (SSIM). Moreover, it must also satisfy constraints on latency, computation, and memory [1]. Nevertheless, deraining studies commonly use parameter count and floating-point operations (FLOPs) as efficiency surrogates without measuring the full inference path. This convention is particularly problematic for conditional models, whose reported cost may omit generation of the auxiliary signal that drives the restoration network.

Deep image deraining has progressed from convolutional and recurrent networks to transformer- and attention-based restoration systems. PReNet established a compact recurrent formulation that progressively refines the restored image through repeated stages [2]. Density-aware deraining further showed that restoration can be adapted to different rain regimes by predicting a density category and routing features through specialized streams [3]. More recent high-capacity architectures, including Restormer and frequency-enhanced state-space models, have expanded the accuracy-oriented design space through long-range and frequency-aware modeling [4, 5]. In parallel, lightweight approaches have pursued recursive pyramids, recursive window transformers, and saliency-guided convolutional attention to reduce model size and arithmetic cost [6-8]. These developments demonstrate that both conditional adaptation and architectural compression can improve the accuracy-complexity tradeoff, but they do not establish whether low parameter counts or FLOPs translate into low realized latency under a common implementation and hardware configuration.

The shift from synthetic benchmarks to captured rain creates a second challenge. Real-rain datasets better represent the interaction of streaks, accumulated rain, splash, illumination, and scene content, but constructing aligned clean references is intrinsically difficult. GT-RAIN acquires real rainy and clean observations through carefully controlled capture, whereas LHP-Rain derives pseudo-ground-truth references from video through robust low-rank tensor recovery [9, 10]. These strategies improve realism but can retain alignment errors or residual degradation in the reference images. Moreover, LHP-Rain does not provide rain-density annotations. Such annotations are generally unavailable in uncontrolled capture because density is neither prescribed during acquisition nor readily recoverable afterward as an objective ground-truth quantity. Consequently, a density-conditioned model developed for synthetic or controlled data cannot be transferred directly to LHP-Rain without constructing a label-free conditioning signal.

Conditioning mechanisms also differ in the information they communicate to a restoration network. Feature-wise linear modulation (FiLM) converts a global conditioning vector into channel-wise affine transformations, whereas spatial feature transformation produces location-dependent modulation from a spatial prior [11, 12]. Global conditioning can represent image-level degradation, but it cannot indicate where individual degradation structures occur. Spatial conditioning preserves this location information, but its larger conditioning head and map-generation procedure add computational cost. Related work in adverse-weather restoration has applied color-imbalance conditioning to a compact transformer for sandstorm restoration and stability-aware physics-prior conditioning to lightweight image dehazing [13, 14]. Although these studies support the broader use of degradation-aware priors, they do not establish whether a label-free spatial rain prior is informative for real-world image deraining or whether any resulting improvement justifies its end-to-end cost. The present study addresses this question by coupling a deterministic, training-free rain-severity estimator with an identity-initialized spatial severity modulation (SSM) head and evaluating the resulting system using matched-capacity conditioning alternatives and prior-replacement controls.

These limitations define the research gap this study addresses. Existing work has not jointly isolated the contribution of conditioning granularity, controlled for conditioning-head capacity, challenged a trained spatially conditioned model using falsification controls, and charged prior generation to the end-to-end inference cost. Likewise, lightweight deraining architectures are often compared through published parameter or FLOP values obtained under different protocols rather than through measured latency under a common evaluation. The present study is therefore positioned as a controlled evaluation rather than a claim of a universally superior architecture. Its objective is to determine whether a label-free severity prior provides a repeatable restoration advantage on real captured rain commensurate with its full computational cost, and whether conventional complexity measures reliably predict realized latency.

To achieve this objective, the study uses PReNet as a fixed backbone and compares an unconditioned baseline with global FiLM, spatial SSM, and a matched-capacity SSM supplied with a constant map. We further evaluate each trained SSM checkpoint after replacing its estimated prior with zero, shuffled, and constant alternatives, testing dependence on the prior's presence, image correspondence, and spatial structure. We then compare the matched-backbone family with a prior-gated two-stage network and an image-only CBAM-light network under the same reconstruction objective and test protocol [7, 15]. A deterministic estimator supplies the label-free severity map, and we assess its cross-scene agreement with low- and high-density labels independently on RealRain-1k rather than assuming it represents ground-truth density [16]. Evaluation combines PSNR, SSIM, and learned perceptual image patch similarity with parameter count, GFLOPs, model latency, prior-generation latency, and total sequential latency. Severity-stratified and upper-versus-lower regional analyses examine whether aggregate results conceal condition- or region-specific behavior.

This study makes four contributions. First, it establishes a label-free difficulty-stratification protocol and quantifies its agreement with labeled real-rain data while explicitly limiting its interpretation as a density measure. Second, it provides a matched-backbone comparison that separates global conditioning, spatial conditioning, and added conditioning-head capacity, supplemented by three prior-replacement controls. Third, it evaluates two lightweight alternatives against the PReNet family while accounting for the cost of every auxiliary signal consumed during inference. Fourth, it introduces severity-stratified and regional analyses to test whether spatial conditioning reduces performance disparities associated with difficult and ground-dominant image content. Together, these contributions

provide an evidence-based assessment of when severity conditioning and lightweight architectural design translate into practical end-to-end efficiency for real-world image deraining.

2. LABEL-FREE SEVERITY CONDITIONING AND LIGHTWEIGHT DERAISING ARCHITECTURES

This section develops the label-free conditioning framework and comparison architectures needed to determine whether spatial rain information improves restoration without confounding signal granularity with model capacity.

2.1. Comparative Framework and PReNet Backbone

The comparison is organized around four conditions that share a common PReNet backbone: an unconditioned baseline (Base), global feature-wise linear modulation (FiLM), spatial severity modulation (SSM), and a matched-capacity constant-map condition (SSM-const). PReNet provides a compact recurrent architecture that progressively refines the restored image through repeated stages [2]. Its recurrent depth, feature width, and main restoration pathway are held fixed across the four conditions. Consequently, differences within this family reflect the conditioning mechanism rather than changes in backbone capacity.

This comparison distinguishes three factors often conflated in conditional restoration: the presence of a conditioning signal, the spatial information in that signal, and the additional capacity introduced by the conditioning head. Base establishes the unconditioned reference. FiLM tests whether an image-level severity statistic is sufficient. SSM retains the spatial distribution of the estimated rain response, whereas SSM-const preserves the SSM head but removes spatial variation from its input. Two additional lightweight models, including a prior-gated two-stage network and an image-only CBAM-light network, are evaluated as independent architectures rather than PReNet variants.

2.2. Label-Free Rain-Severity Estimator

A deterministic, training-free procedure derives a per-pixel response map $S(x)$ and an image-level summary statistic $\bar{s}$ directly from the rainy input, as shown in Fig. 1. The procedure is designed for real-rain datasets that do not provide density annotations. It identifies bright, narrow, predominantly near-vertical structures consistent with rain streaks, but it is not intended to recover a physically calibrated rain density.

Let $Y(x)$ denote the luminance of the rainy image at spatial location x . A guided self-filter $G(\cdot)$, with radius 4 and regularization parameter $\varepsilon_g = 0.08^2$, estimates the locally smooth component of the luminance image [17]. The positive high-frequency residual is

$$R(x) = \max(Y(x) - G(Y)(x), 0). \quad (1)$$

Retaining only the positive residual suppresses slowly varying illumination and dark edges while emphasizing locally bright structures. Because scene textures can produce similar responses, $R(x)$ is treated as an intermediate streak-sensitive representation rather than a rain mask.

The residual is then convolved with a bank of zero-mean line kernels $K_{\theta_k, l}$, indexed by orientation θ_k and length l . The strongest absolute response across the bank is

$$E(x) = \max_{\theta_k, l} |(K_{\theta_k, l} * R)(x)|. \quad (2)$$

The kernel bank spans five near-vertical orientations from 60° to 120° in 15° increments and uses lengths of 7 and 11 pixels. Taking the maximum across orientations and lengths permits the estimator to respond to streaks with modest variations in direction and scale without introducing learned parameters.

Severity estimator — deterministic, training-free, no learned parameters

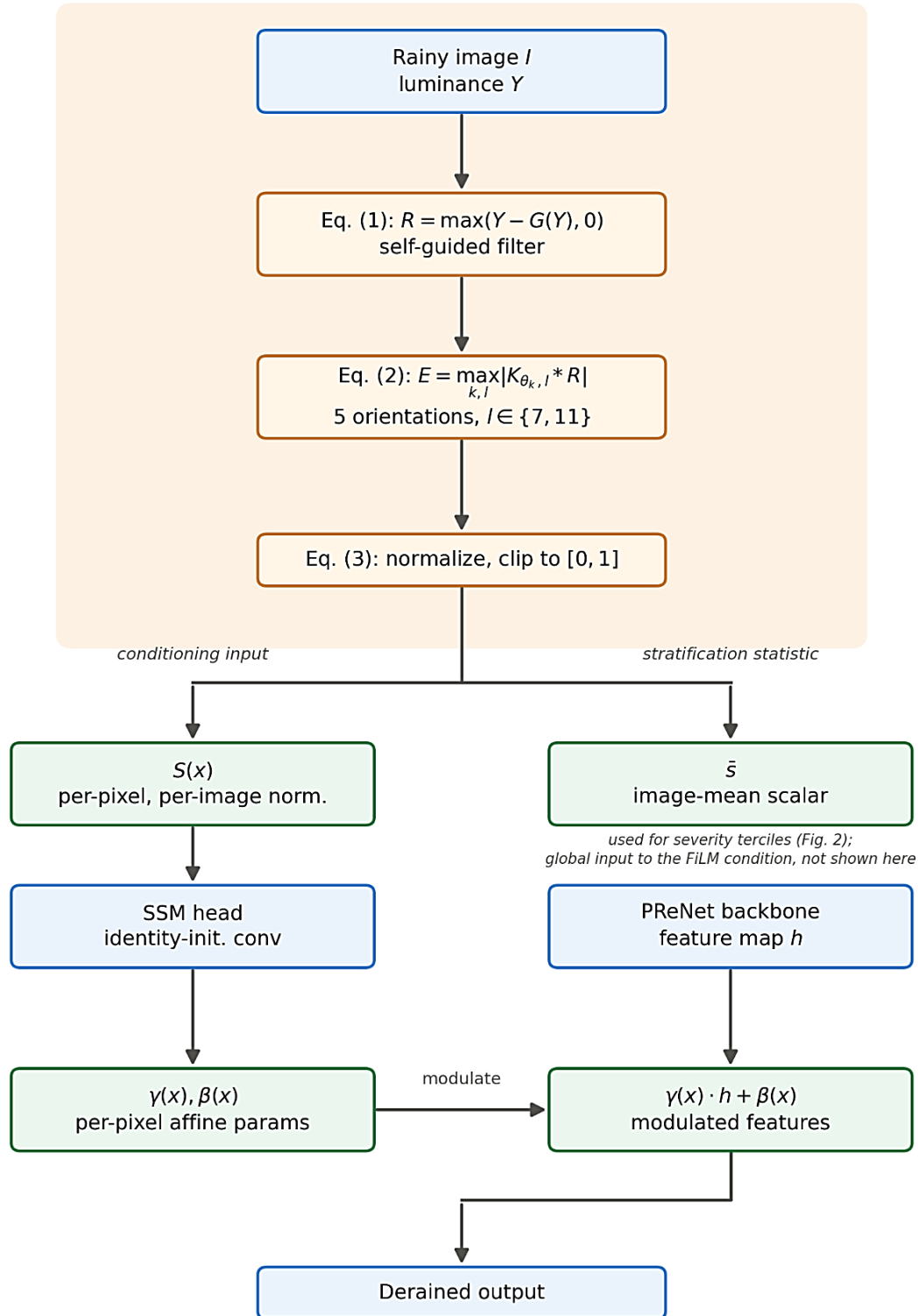


Figure 1. Construction of the label-free severity map through Eqs. (1)–(3) and its integration into the SSM conditioning head.

The response is normalized within each image using its 5th and 95th percentiles:

$$S(x) = \text{clip}\left(\frac{E(x) - q_{0.05}(E)}{q_{0.95}(E) - q_{0.05}(E) + \varepsilon}, 0, 1\right), \quad \bar{s} = \frac{1}{HW} \sum_x S(x), \quad (3)$$

where H and W are the image height and width, respectively, and ε prevents division by zero. The normalized map $S(x)$ serves as the spatial conditioning input, while $\bar{s}$ summarizes the distribution of the normalized response at the image level. Because the normalization is performed independently for each image, $\bar{s}$ is not interpreted as an absolute rain-density measurement. It is used only as a reproducible difficulty statistic, and its agreement with externally provided low- and high-density labels is evaluated separately.

2.3. Global and Spatial Severity Conditioning

The four PReNet-family conditions differ only in the signal supplied to the backbone and the mechanism used to transform that signal, as shown in Table 1.

Table 1. Comparison of Conditioning Mechanisms.

Condition	Conditioning signal	Purpose
Base	None	Establishes the unconditioned reference
FiLM	Image-level statistic $\bar{s}$	Tests whether global severity information is sufficient
SSM	Spatial map $S(x)$	Tests whether location-dependent severity information is beneficial
SSM-const	Constant map with value $\bar{s}$	Controls for SSM-head capacity while removing spatial variation

FiLM and SSM apply the same general affine transformation to an intermediate recurrent feature tensor h :

$$h' = h \odot [1 + \gamma(c) + \beta(c)], \quad c = \begin{cases} \bar{s}, & \text{FiLM,} \\ S(x), & \text{SSM.} \end{cases} \quad (4)$$

FiLM converts the scalar $\bar{s}$ into channel-wise modulation parameters γ and β , which are broadcast across spatial positions. This design follows the principle of feature-wise affine conditioning [11]. SSM instead derives channel- and position-dependent modulation maps from $S(x)$, following the broader principle of spatial feature transformation [12]. This modulation is applied once within each of PReNet's six recurrent iterations, immediately after the convolutional-LSTM gate update and before the five residual blocks, using the same conditioning head (i.e., shared, not iteration-indexed, parameters) at every iteration.

The final affine layer of each conditioning head is initialized to zero. Therefore, $\gamma(x) = 0$ and $\beta(c) = 0$ at initialization, and Eq. (4) initially reduces to $h' = h$. This identity initialization lets each conditioned model start from the unmodulated feature pathway and learn only the degree of conditioning supported by the training data. SSM-const uses the same architecture and parameter count as SSM but replaces $S(x)$ with a spatially constant map whose value is $\bar{s}$. It therefore preserves the conditioning head and image-level response magnitude while removing location-dependent structure. The contrast between SSM and SSM-const isolates the contribution of spatial variation more directly than a comparison between SSM and Base alone.

2.4. Lightweight Comparison Architectures

Two independent lightweight architectures broaden the comparison beyond alternative conditioning heads on PReNet. They use the same reconstruction objective but are not exact reproductions of previously published models.

The first is a prior-gated two-stage architecture organized as local repair followed by global correction. A depthwise-separable repair network receives the concatenation of the rainy image I and severity map S . Its residual output produces an intermediate image I_o :

$$I_o = I + g_\phi([I, S]). \quad (5)$$

An AOD-Net-style module then estimates a spatial correction factor $K_\psi(I_o)$ and produces the restored image [18]:

$$\hat{J} = \text{clip}(K_\psi(I_o) \odot I_o - K_\psi(I_o) + b, 0, 1). \quad (6)$$

The first stage is termed de-occlusion because its local prior emphasizes bright, oriented streak structures that may obscure scene details. This term describes the intended computational function rather than asserting that the stage identifies every form of rain occlusion or splash. The second stage is termed de-veiling because it applies the AOD-style reformulation to correct broader contrast and visibility degradation without separately estimating transmission and atmospheric light. The de-occlusion-first ordering was fixed before the main comparison based on preliminary architectural screening; that screening motivates the configuration but is not treated as causal evidence for its physical interpretation.

The second architecture, CBAM-light, is an image-only model that combines dilated convolution with channel and spatial attention. It follows the general lightweight visual-saliency pattern used in LDVS and implements attention through convolutional block attention modules [7, 15]. CBAM-light is a study-specific reimplementation, not an exact reproduction of LDVS. Its default width yields 54,998 trainable parameters, close to the approximately 50,703 parameters reported for LDVS. Unlike FiLM, SSM, SSM-const, and the two-stage architecture, CBAM-light receives no externally computed severity signal. It therefore provides a lightweight, image-only comparison to determine whether reduced arithmetic complexity translates into reduced realized latency.

3. EXPERIMENTAL DESIGN

This section establishes a common and reproducible protocol for testing restoration quality, dependence on the conditioning prior, and realized end-to-end efficiency under controlled training and evaluation conditions.

3.1. Datasets and Data Splits

LHP-Rain is the primary training and restoration-evaluation dataset [10]. It contains rain captured under real outdoor conditions across multiple cities and provides pseudo-ground-truth clean references recovered from video through robust low-rank tensor modeling. Although this construction offers more realistic degradation than synthetic rain generation, residual rain and imperfect alignment may remain in the recovered references. Reference-based metrics are therefore interpreted as measures of agreement with the released pseudo-ground truth rather than as complete measures of perceptual restoration quality. LHP-Rain does not provide rain-density annotations.

The experiments use the released train-simple and local test configurations rather than the full-resolution benchmark protocol. The train-simple split contains 75,838 pre-cropped 256×256 patches, representing approximately 10% of the complete training material. The primary configuration uses a fixed-seed sample of 10,000 patches generated with seed 20260814, together with 1,000 additional patches reserved for validation. The test set contains 1,000 paired 512×512 crops, all of which are included in the reported evaluation. By contrast,

the published LHP-Rain benchmark evaluates full-resolution 1920×1080 frames. Published benchmark values are therefore included only as contextual anchors and not as like-for-like comparisons.

The 10,000-patch sample represents 13.2% of train-simple. However, because train-simple consists of cropped patches rather than independently captured scenes, the nominal patch count may overstate the diversity of the training data. The initial experiment screened all six conditions using 5,000 patches for 50 epochs. To determine whether its principal finding persisted under a materially larger training budget, we subsequently doubled the number of unique training patches to 10,000 and increased the training duration to 100 epochs. This change increased the total number of sample presentations and optimizer updates by approximately fourfold. We adopted the larger configuration as the primary training and reporting scale, while retaining the original 5,000-patch experiment as a training-budget sensitivity check. Because we increased patch count and training duration simultaneously, this comparison evaluates robustness to a larger overall training budget rather than isolating the effect of data quantity alone.

A filename-level audit of the released manifests showed that the train-simple subsample and local test split are not scene-disjoint. Using the `<scene>_<frame>` naming convention of the LHP-Rain release, 446 of the test set's 463 distinct source scenes (96.3%) also contributed patches to the 10,000-patch training sample. The corresponding overlap was 399 of 463 scenes (86.2%) for the earlier 5,000-patch sample and 454 of 463 scenes (98.1%) for the complete 75,838-patch train-simple release. By contrast, the training and validation subsamples shared no scene identifiers under the same audit procedure.

Matching filename identifiers across the training and test sets were comparatively uncommon: 7 of 10,000 training images at the primary scale, 5 of 5,000 at the screening scale, and 101 of 75,838 in the complete train-simple release. Because the released training and test images can differ in resolution and crop offset, a matching identifier does not necessarily indicate pixel-identical content. Nevertheless, the extensive scene-level overlap shows that most test scenes originate from source videos also represented in the training material. This overlap is a property of the officially released train-simple and local test configurations, not an artifact introduced by the present subsampling procedure.

All evaluated conditions use identical training, validation, and test lists; therefore, the overlap does not create an explicit condition-specific procedural advantage. It may nevertheless elevate absolute PSNR and SSIM values by exposing models to scene-related content and could interact differently with model capacity or conditioning. Accordingly, interpret both the absolute restoration scores and comparisons with published anchors as results on the official overlapping configuration rather than as evidence of generalization to unseen scenes.

RealRain-1k is used only to assess whether the label-free statistic agrees with externally assigned low- and high-density conditions [16]. The dataset contains the same 1,120 scenes represented at two rain densities through density-controllable temporal filtering. The present validation uses the released 224-scene test split from RealRain-1k-L and RealRain-1k-H, producing 448 images. RealRain-1k is not used to train or evaluate the restoration architectures.

3.2. Training Objective and Protocol

All six trained conditions, including Base, FiLM, SSM, SSM-const, two-stage, and CBAM-light, are optimized using the same reconstruction objective and data protocol. This common training design isolates the effects of conditioning and architecture from differences introduced by model-specific loss functions.

Let $I \in [0,1]^{3 \times H \times W}$ denote the rainy input, J its pseudo-ground-truth reference, and z the model-specific conditioning input. The restored image is expressed as $\hat{J} = f_{\theta}(I, z)$, where θ denotes the trainable model parameters. The conditioning input is defined as

$$z = \begin{cases} \emptyset & \text{Base and CBAM-light} \\ \bar{s} & \text{FiLM} \\ S(x) & \text{SSM and two-stage} \\ \bar{s}\mathbf{1}_{H \times W} & \text{SSM-const} \end{cases} \quad (7)$$

where $\emptyset$ denotes the absence of an auxiliary conditioning input, $\bar{s}$ is the image-level severity statistic, $S(x)$ is the estimated spatial severity map, and $\mathbf{1}_{H \times W}$ is an all-one map. Thus, $\bar{s}\mathbf{1}_{H \times W}$ supplies SSM-const with the same image-level statistic as FiLM while removing spatial variation.

Every condition minimizes the mean absolute reconstruction error:

$$\mathcal{L}_1(\theta) = \frac{1}{3HW} \sum_{c=1}^3 \sum_x |f_{\theta}(I, z)_c(x) - J_c(x)|. \quad (8)$$

Using the same $\mathcal{L}_1$ objective prevents the architecture comparison from being confounded by differences in loss formulation. The retrained Base condition is not intended to reproduce the originally published PReNet score because the official PReNet implementation uses mean squared error or negative SSIM, whereas the present study applies $\mathcal{L}_1$ uniformly to all conditions [2]. Severity-weighted and detector-based auxiliary losses are excluded from the validated comparison because they would introduce additional objectives requiring independent weight selection and controlled ablation.

Each condition is trained for 100 epochs with a batch size of 16 using random seeds 42, 1337, and 2024. The same training, validation, and test partitions are maintained across conditions. Reported results are calculated as the arithmetic mean and sample standard deviation of the three seed-level measurements. For reproducibility, each run retains its resolved configuration, training log, metric output, final checkpoint, latency measurements, and source-revision identifier. The three prior-replacement controls are not separately trained models; they are post-training interventions applied to each retained SSM checkpoint, as described in Section 3.3.

All conditions were optimized with Adam (learning rate 1×10^{-4} , no learning-rate schedule) and a batch size of 16. The only data augmentation applied during training was random 256×256 cropping of each rainy/clean pair; no flipping, rotation, or color jittering was used. Evaluation uses each run's final-epoch checkpoint for every condition, with no best-on-validation selection, so no condition benefits from selective early stopping. All experiments used PyTorch 2.5.1 (CUDA 12.1, cuDNN 9.1.0) on a single NVIDIA GeForce RTX 4090 GPU under Python 3.10.20; LPIPS was computed with the lpips package (v0.1) using its AlexNet backbone.

3.3. Prior-Replacement Falsification Controls

Three post-training interventions test whether the trained SSM model depends on the information carried by its estimated prior. These interventions are applied to every trained SSM checkpoint without retraining or changing any model weights.

In the **zero-prior control**, $S(x)$ is replaced with an all-zero map. This intervention tests whether the model depends on a nonzero conditioning signal. In the **shuffled-prior control**, $S(x)$ is replaced with a severity map obtained from another image in the same evaluation batch. This preserves the broad distribution of prior values while disrupting correspondence between

the rainy image and its prior. In the **constant-prior control**, $S(x)$ is replaced with a spatially constant map whose value equals the image's $\bar{s}$. This retains its image-level statistic while removing spatial variation.

The input tensors used for the estimated, shuffled, zero, and constant conditions are verified to be distinct where required. Together, these controls test three increasingly specific forms of prior use: sensitivity to the prior's presence, dependence on image–prior correspondence, and dependence on its spatial structure. They complement SSM-const, which tests the effect of training a matched-capacity model from the outset with a spatially invariant input.

3.4. Restoration-Quality Metrics

PSNR and SSIM are computed on the luminance channel of YCbCr, whereas learned perceptual image patch similarity (LPIPS) is computed on RGB. This combination captures pixel fidelity, structural similarity, and learned perceptual distance.

For images normalized to $[0,1]$, luminance mean squared error and PSNR are

$$MSE_Y = \frac{1}{HW} \sum_x [\hat{Y}(x) - Y_f(x)]^2, \quad PSNR_Y = 10 \log_{10} \left(\frac{1}{MSE_Y} \right) \quad (9)$$

where $\hat{Y}$ and Y_f denote the luminance channels of the restored image and reference image, respectively.

SSIM measures agreement in local luminance, contrast, and structure [19]:

$$SSIM(a, b) = \frac{(2\mu_a\mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)}. \quad (10)$$

LPIPS measures the distance between normalized deep representations extracted at multiple feature levels [20]:

$$d(a, b) = \sum_l \frac{1}{H_l W_l} \sum_x \|w_l \odot [\hat{\phi}_l(a)_x - \hat{\phi}_l(b)_x]\|_2^2. \quad (11)$$

Higher PSNR and SSIM indicate better agreement with the reference, whereas lower LPIPS indicates greater perceptual similarity. We first compute metrics for each test image and checkpoint, then aggregate across the full test split. The three seed-level results are summarized using their mean and sample standard deviation. Because three seeds provide limited power for formal significance testing, the experiment emphasizes the magnitude and consistency of observed differences rather than claiming statistical significance from seed counts alone.

To determine whether the small mean difference reported in Section 4.2 represents practical equivalence rather than merely an absence of statistical significance, we compute paired per-image PSNR differences between SSM and Base on the shared 1,000-image test split for each seed. Each seed's mean paired difference is treated as one independent training replicate ($n=3$), and the mean across replicates is reported with its 95% confidence interval. We evaluate equivalence using a project-defined margin of ± 0.1 dB, selected after the 5,000-patch screening experiment but fixed before analysis of the primary 10,000-patch configuration as the smallest PSNR difference considered operationally meaningful. The margin was not adjusted after observing the primary results. Equivalence is assessed using two one-sided tests (TOST) at $\alpha = 0.05$ [21]. Because the margin is project-defined and only three independent training replicates are available, the result is interpreted as equivalence under the stated experimental criterion rather than as evidence of exact equality.

3.5. Severity Validation and Stratified Analyses

The label-free statistic $\bar{s}$ is validated independently on RealRain-1k through within-scene and cross-scene analyses. Within-scene ordering examines whether the high-density version of a scene receives a larger value than its corresponding low-density version. Cross-scene discrimination is assessed using the area under the receiver operating characteristic curve (AUC) and point-biserial correlation between $\bar{s}$ and the binary low/high label. These analyses distinguish paired sensitivity to changes in rain conditions from general discrimination across scenes with different textures and content.

For LHP-Rain, tercile boundaries are computed from the $\bar{s}$ distribution of the 1,000-image test split. The resulting low-, medium-, and high-statistic groups provide a deterministic difficulty partition for evaluating whether aggregate performance conceals condition-specific behavior. Because $\bar{s}$ is affected by both streak-like responses and scene texture, the groups are described as severity-statistic or difficulty strata rather than ground-truth rain-density classes.

A complementary regional analysis divides each 512×512 test image into an upper half (rows 0–255) and a lower half (rows 256–511). We compute regional PSNR and SSIM for Base and SSM using the same checkpoints as the whole-image evaluation. The division is motivated by the ground-dominant degradation reported for LHP-Rain, but the two regions are geometric proxies rather than semantic segmentations of sky, road, splash, or other scene components. The analysis therefore tests whether SSM changes a reproducible regional performance gap without assigning that gap to a specific physical degradation.

3.6. Complexity and Latency Evaluation

Computational efficiency is assessed using trainable parameter count, GFLOPs, model latency, prior-generation latency, and total sequential latency. Parameter count measures model storage requirements, while GFLOPs estimate the arithmetic operations required for a single 512×512 forward pass. Neither measure substitutes for realized latency because execution time also depends on memory traffic, kernel-launch overhead, operator implementation, and hardware utilization.

Latency is measured at batch size 1 on an NVIDIA GeForce RTX 4090. Each model is evaluated over 50 timed repetitions following five warm-up passes, with CUDA synchronization before and after timing to prevent asynchronous execution from biasing the measurements. Model latency covers the restoration-network forward pass. The deterministic prior is computed on the host CPU, and its measured generation time is charged only to conditions that consume $\bar{s}$ or $S(x)$.

Total latency is defined as the sum of prior-generation latency and model latency under sequential execution:

$$T_{total} = T_{prior} + T_{model}. \quad (12)$$

This definition does not assume asynchronous overlap between CPU prior generation and GPU inference. It therefore provides a transparent end-to-end estimate for the tested implementation rather than a hardware-independent claim about deployment speed.

4. RESULTS

This section evaluates the severity statistic, restoration accuracy, dependence on the estimated prior, regional behavior, and the relationship between theoretical complexity and measured end-to-end latency.

4.1. Agreement of the Label-Free Statistic With Rain-Density Labels

The label-free statistic $\bar{s}$ was first evaluated on the 448-image RealRain-1k test set, comprising 224 scenes under each low- and high-density condition. Terciles calculated from the RealRain-1k distribution had boundaries of 0.287 and 0.329. Among the low-density images, 48.2% were assigned to the low tercile, while 48.7% of the high-density images were assigned to the high tercile. Both proportions exceeded the 33.3% expected under equal random assignment, but neither indicated clean separation.

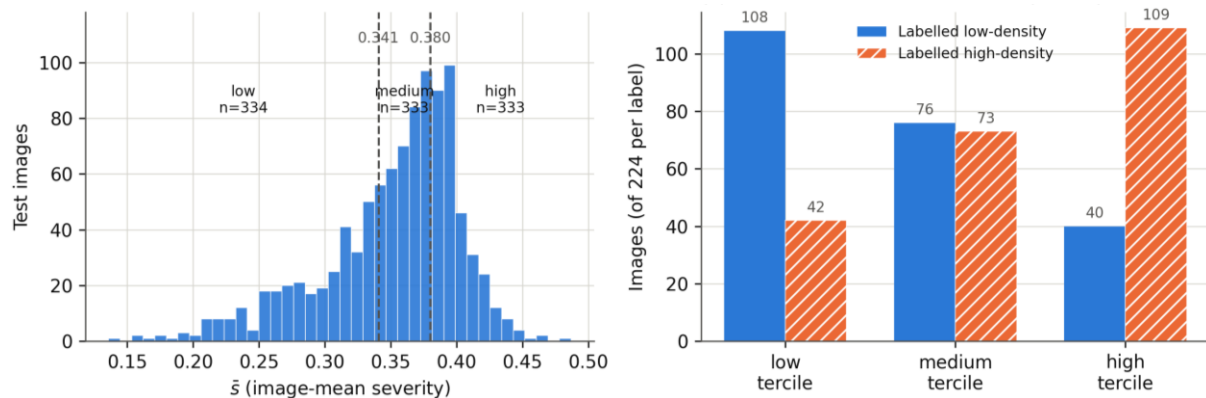


Figure 2. Distribution and external validation of the label-free statistic. (a) Distribution of $\bar{s}$ on the LHP-Rain test set ($n = 1,000$), with tercile boundaries at 0.341 and 0.380. (b) RealRain-1k tercile assignments by low- and high-density label ($n = 448$).

Across scenes, $\bar{s}$ achieved an area under the receiver operating characteristic curve of 0.772 and a point-biserial correlation of $r=0.455$ with the binary density label. These results indicate moderate agreement, not a direct measurement of rain density. Accordingly, $\bar{s}$ is used as a reproducible difficulty statistic, not as a substitute for a validated density annotation.

On the 1,000-image LHP-Rain test split, $\bar{s}$ had a mean of 0.351 and a standard deviation of 0.053. Tercile boundaries of 0.341 and 0.380 divided the test set into 334 low-, 333 medium-, and 333 high-statistic images. Figure 2 summarizes the LHP-Rain distribution and the relationship between RealRain-1k tercile assignments and the externally assigned density labels.

4.2. Restoration Accuracy and Dependence on the Prior

Table 2 examines whether the principal finding changes under a larger training budget. Increasing the configuration from 5,000 patches and 50 epochs to 10,000 patches and 100 epochs improved Base by 0.099 dB in PSNR and 0.0054 in SSIM, while SSM improved by 0.116 dB and 0.0059, respectively. However, the SSM–Base PSNR difference increased only from 0.017 to 0.034 dB. Because the number of patches and training epochs were increased simultaneously, this comparison cannot separate the effects of additional data from those of additional optimization. It nevertheless shows that the small advantage of SSM persisted but did not become practically substantial under the larger training budget.

Table 3 reports restoration quality on the complete LHP-Rain test split under the primary configuration. Among the trained PReNet-family conditions, SSM attained the highest mean PSNR and SSIM and the lowest LPIPS. Its improvements over Base were nevertheless limited to 0.034 dB in PSNR, 0.0011 in SSIM, and 0.0016 in LPIPS. FiLM and SSM-const also differed only marginally from Base. The numerical ranking therefore does not establish a practically meaningful advantage for spatial severity conditioning.

Table 2. Training-budget sensitivity of Base and SSM under the 5,000-patch/50-epoch screening configuration and the primary 10,000-patch/100-epoch configuration; values are mean $\pm$ sample standard deviation across three seeds.

Condition	Training configuration	PSNR $\uparrow$ (dB)	SSIM $\uparrow$
Base	5,000 patches/50 epochs	31.642 $\pm$ 0.047	0.8975 $\pm$ 0.0012
Base	10,000 patches/100 epochs (primary)	31.741 $\pm$ 0.095	0.9029 $\pm$ 0.0030
SSM	5,000 patches/50 epochs	31.659 $\pm$ 0.034	0.8981 $\pm$ 0.0011
SSM	10,000 patches/100 epochs (primary)	31.775 $\pm$ 0.109	0.9040 $\pm$ 0.0043

Table 3. Restoration quality on the LHP-Rain test split ($n = 1,000$). Values are mean $\pm$ sample standard deviation across three seeds. The zero-, shuffled-, and constant-prior rows are post hoc interventions applied to the trained SSM checkpoints without retraining.

Condition	PSNR $\uparrow$ (dB)	SSIM $\uparrow$	LPIPS $\downarrow$
Base	31.741 $\pm$ 0.095	0.9029 $\pm$ 0.0030	0.1112 $\pm$ 0.0026
FiLM	31.758 $\pm$ 0.104	0.9034 $\pm$ 0.0030	0.1108 $\pm$ 0.0033
SSM with estimated $S(x)$	31.775 $\pm$ 0.109	0.9040 $\pm$ 0.0043	0.1096 $\pm$ 0.0036
SSM with zero prior	31.754 $\pm$ 0.102	0.9014 $\pm$ 0.0037	0.1107 $\pm$ 0.0038
SSM with shuffled prior	31.739 $\pm$ 0.110	0.9028 $\pm$ 0.0044	0.1108 $\pm$ 0.0037
SSM with constant prior	31.753 $\pm$ 0.121	0.9038 $\pm$ 0.0044	0.1093 $\pm$ 0.0034
SSM-const, trained	31.765 $\pm$ 0.100	0.9030 $\pm$ 0.0031	0.1106 $\pm$ 0.0027
Two-stage	31.626 $\pm$ 0.032	0.8964 $\pm$ 0.0003	0.1180 $\pm$ 0.0010
CBAM-light	31.652 $\pm$ 0.025	0.8974 $\pm$ 0.0002	0.1148 $\pm$ 0.0003

The prior-replacement interventions test whether the trained SSM model uses the information contained in the estimated spatial map. Relative to the estimated prior, replacing $S(x)$ with zeros reduced mean PSNR by 0.021 dB, shuffling priors across images reduced it by 0.036 dB, and replacing each prior with its corresponding constant map reduced it by 0.022 dB. These effects were small and were not consistently ordered across PSNR, SSIM, and LPIPS. In particular, the constant-prior intervention produced marginally better LPIPS than the estimated prior. The separately trained SSM-const condition was also only 0.010 dB below SSM in PSNR. The controls therefore provide limited evidence that prior correspondence or spatial variation materially contributes to restoration performance.

The seed-level equivalence analysis supports this interpretation. Based on the mean paired image difference calculated independently for each seed, the SSM–Base PSNR difference was +0.034 dB, with a 95% confidence interval of $[-0.026, +0.094]$ dB. The interval lies within the project-defined practical-equivalence margin of ± 0.1 dB, which was fixed after the screening experiment but before analysis of the primary configuration. The two one-sided tests support equivalence under this margin ($p=0.021$). This result indicates practical equivalence within the stated experimental framework rather than exact equality between the models. Its interpretation remains bounded by the project-defined margin and the availability of only three independent seed-level replicates.

Figure 3 tests whether aggregate averaging conceals a greater benefit on more difficult images. PSNR decreases from the low- to high-statistic tercile for every condition, confirming that the label-free statistic separates images by restoration difficulty. However, the differences among Base, FiLM, SSM, and SSM-const remain below 0.1 dB across the three terciles. Replacing the estimated SSM prior with zero, shuffled, or constant maps likewise produces no consistent deterioration, including within the high-statistic tercile where severity conditioning would be expected to offer its greatest benefit. The statistic therefore identifies difficult images more clearly than conditioning improves their restoration.

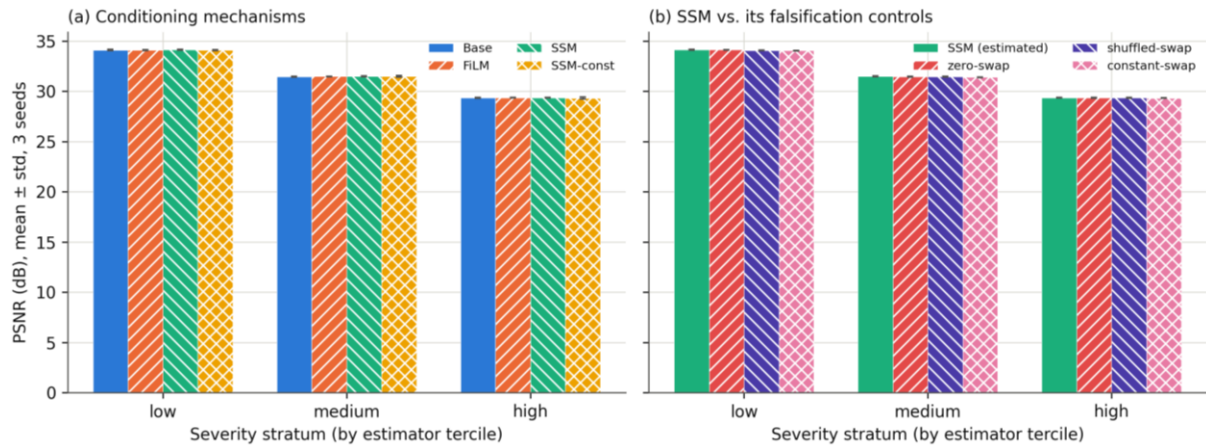


Figure 3. Restoration PSNR across LHP-Rain severity-statistic terciles under the primary training configuration. (a) Comparison of Base, FiLM, SSM, and SSM-const. (b) Comparison of the estimated SSM prior with zero-, shuffled-, and constant-prior interventions.

The lightweight architectures incurred modest accuracy penalties relative to Base. Two-stage was lower by 0.115 dB in PSNR, 0.0065 in SSIM, and 0.0068 in LPIPS, whereas CBAM-light was lower by 0.089 dB, 0.0055, and 0.0036, respectively. Relative to the mean PSNR of the four trained PReNet-family conditions, the deficits were 0.134 dB for Two-stage and 0.108 dB for CBAM-light. These accuracy reductions must be considered together with the computational and latency results rather than interpreted independently.

4.3. Computational Complexity and Realized Latency

Table 4 compares model size, arithmetic cost, and measured latency at 512×512 resolution. SSM increased the PReNet parameter count and GFLOPs by approximately 11.1%. Its model-only latency increased from 20.43 to 23.79 ms, while prior generation added another 19.73 ms. The resulting sequential total of 43.52 ms was more than twice Base's latency.

Table 4. Model complexity and latency at batch size 1 and 512×512 resolution on an NVIDIA GeForce RTX 4090.

Condition	Parameters	GFLOPs	Model (ms)	Prior (ms)	Total (ms)
Base	168,963	265.0	20.43	0	20.43
FiLM	171,139	265.0	20.90	20.11	41.01
SSM	187,779	294.4	23.79	19.73	43.52
SSM-const	187,779	294.4	23.79	19.73	43.52
Two-stage	170,446	44.6	4.50	20.23	24.74
CBAM-light	54,998	14.1	33.68	0	33.68

The two-stage model reduced restoration-network computation from 265.0 to 44.6 GFLOPs and model-only latency from 20.43 to 4.50 ms. However, generating its spatial prior increased total latency to 24.74 ms, which was 21% higher than Base. The approximately 20 ms prior-generation time therefore became a dominant fixed cost once the restoration network itself was made inexpensive.

CBAM-light exhibited the opposite discrepancy. It used approximately 3.1 times fewer parameters and 18.8 times fewer GFLOPs than Base, yet its measured latency was 33.68 ms, or 1.6 times higher. Because we did not conduct operator-level profiling, Table 4 shows an implementation- and hardware-specific mismatch between arithmetic complexity and realized latency rather than a universal limitation of CBAM-based architectures.

4.4. Regional Performance

Table 5 presents the upper- and lower-half results for Base and SSM. Both models performed less favorably in the lower half. For Base, the upper-to-lower difference was 0.179 dB in PSNR and 0.0156 in SSIM. For SSM, the corresponding differences were 0.165 dB and 0.0152. SSM improved lower-half PSNR over Base by only 0.041 dB and therefore did not materially reduce the regional disparity.

Table 5. Upper- and lower-half restoration quality on LHP-Rain; values are mean $\pm$ sample standard deviation across three seeds.

Condition	Region	PSNR $\uparrow$ (dB)	SSIM $\uparrow$
Base	Upper half	32.074 $\pm$ 0.074	0.9106 $\pm$ 0.0022
Base	Lower half	31.895 $\pm$ 0.109	0.8950 $\pm$ 0.0038
SSM	Upper half	32.101 $\pm$ 0.073	0.9116 $\pm$ 0.0032
SSM	Lower half	31.936 $\pm$ 0.128	0.8964 $\pm$ 0.0055

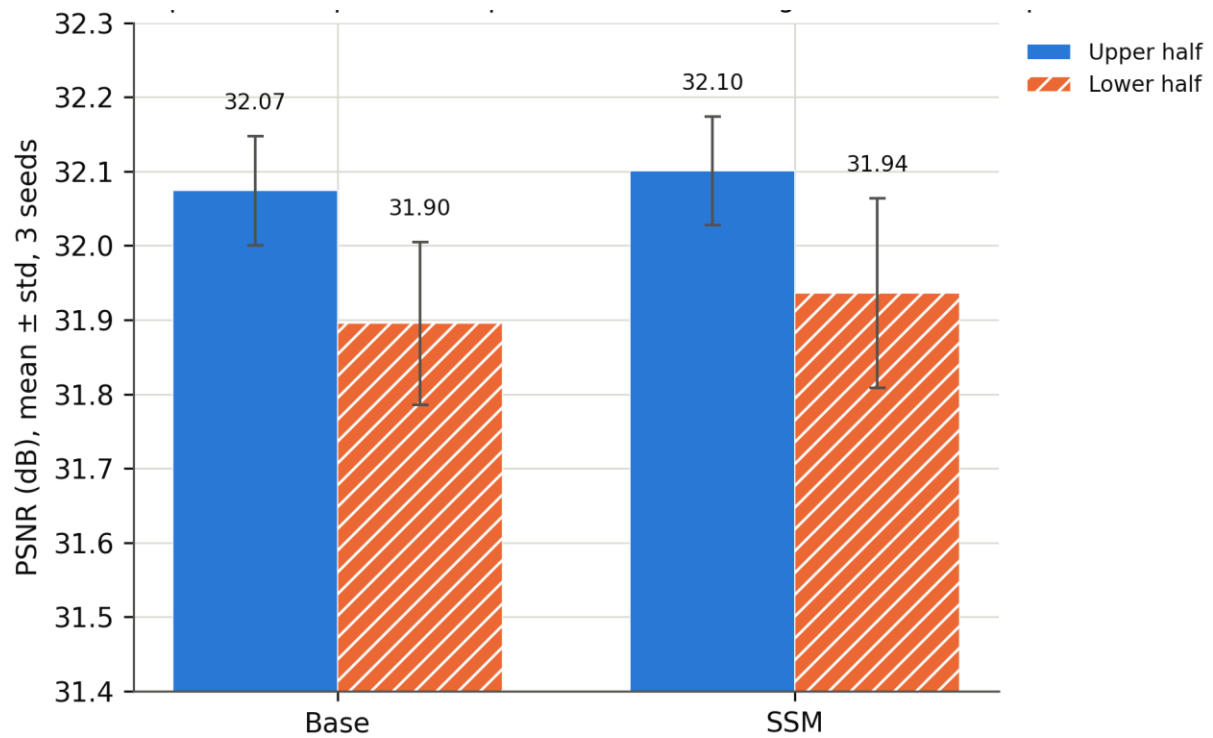


Figure 4. Upper- and lower-half PSNR for Base and SSM on LHP-Rain. Bars show means across three seeds, and error bars show sample standard deviations. The truncated vertical axis highlights a regional difference below 0.3 dB; the regions are geometric proxies rather than semantic masks.

The arithmetic mean of upper- and lower-half PSNR is not expected to equal whole-image PSNR in Table 3 because PSNR is a logarithmic transformation of MSE. Averaging separately calculated log-domain values differs from calculating PSNR from the combined image-level MSE. The corresponding SSIM averages do not show this offset.

Figure 4 visualizes the regional difference. The geometric division does not semantically identify sky, ground, road, or splash; therefore, the lower performance cannot be attributed to a specific rain phenomenon. It nevertheless shows that the tested spatial prior did not eliminate the reproducible lower-region deficit.

4.5. Contextual Comparison With Published LHP-Rain Results

Table 6 provides published LHP-Rain values only to contextualize the scale of the present results [10]. The protocols are not directly comparable. The published values use the full-resolution configuration, different training losses, and 1920×1080 test frames, whereas the present study uses 10,000 training patches, a common $\mathcal{L}_1$ objective, and 512×512 test crops. The published PReNet value exceeds the present Base result by 0.389 dB, which is plausible given its larger training configuration and different loss design. Table 6 should therefore not be interpreted as a ranking of the evaluated methods.

Table 6. Contextual LHP-Rain results reported under different training and evaluation protocols.

Method	PSNR (dB)	SSIM
Rainy input, as reported	29.97	0.8497
PReNet, as reported	32.13	0.9177
SCD-Former, as reported	34.33	0.9468
Base, present study	31.741 ± 0.095	0.9029 ± 0.0030

4.6. Qualitative Comparison

The qualitative examples in Figure 5 reinforce the small quantitative differences reported in Table 3. Base and SSM produce nearly indistinguishable outputs across the low-, medium-, and high-severity-statistic terciles. Both models generally preserve major edges, colors, and scene structures, but SSM shows no obvious advantage in suppressing rain or recovering detail. This similarity is especially important in the high-severity example: although restoration becomes more difficult as the severity statistic increases, the spatial prior does not visibly redirect the model toward more effective correction of the prominent rain-affected regions. This observation is consistent with Figure 3, where the SSM advantage remains below 0.1 dB across all terciles.

The final row presents a more demanding regional case, selected because Base produced its largest upper-minus-lower PSNR difference among the high-tercile images. Nevertheless, Base and SSM again produce visually similar lower-region reconstructions. This observation supports the regional results in Table 5: SSM improves lower-half PSNR by only 0.041 dB and does not materially reduce the upper-to-lower performance gap. However, the image cannot establish that the remaining degradation is caused by splash. The geometric regional criterion does not provide a semantic label, and the estimator primarily responds to bright, near-vertical, high-frequency structures. The example therefore indicates a limitation in coverage rather than demonstrating a specific physical failure mechanism.

The practical implication is that the spatial prior does not deliver an observable improvement commensurate with its computational cost. When Base and SSM produce comparable quantitative and visual outcomes, Base is the more defensible deployment choice because it avoids the conditioning head and approximately 20 ms of prior-generation latency reported in Table 4. This negative result is operationally meaningful: adding a physically plausible map is insufficient unless the model demonstrably uses its spatial information to improve difficult regions.

The qualitative evidence must nevertheless be interpreted conservatively. Figure 5 contains only four examples, uses one seed, and was not evaluated through blinded human assessment, magnified residual views, or downstream perception tasks. Moreover, the reference images are video-derived pseudo-ground truth and may contain residual degradation or smoothing, particularly in difficult regions. Figure 5 therefore corroborates the aggregate

findings but does not independently establish perceptual equivalence. Stronger perceptual evaluation would require enlarged crops, Base–SSM difference maps, multiple checkpoints, human ratings, or task-level measures such as detection accuracy.

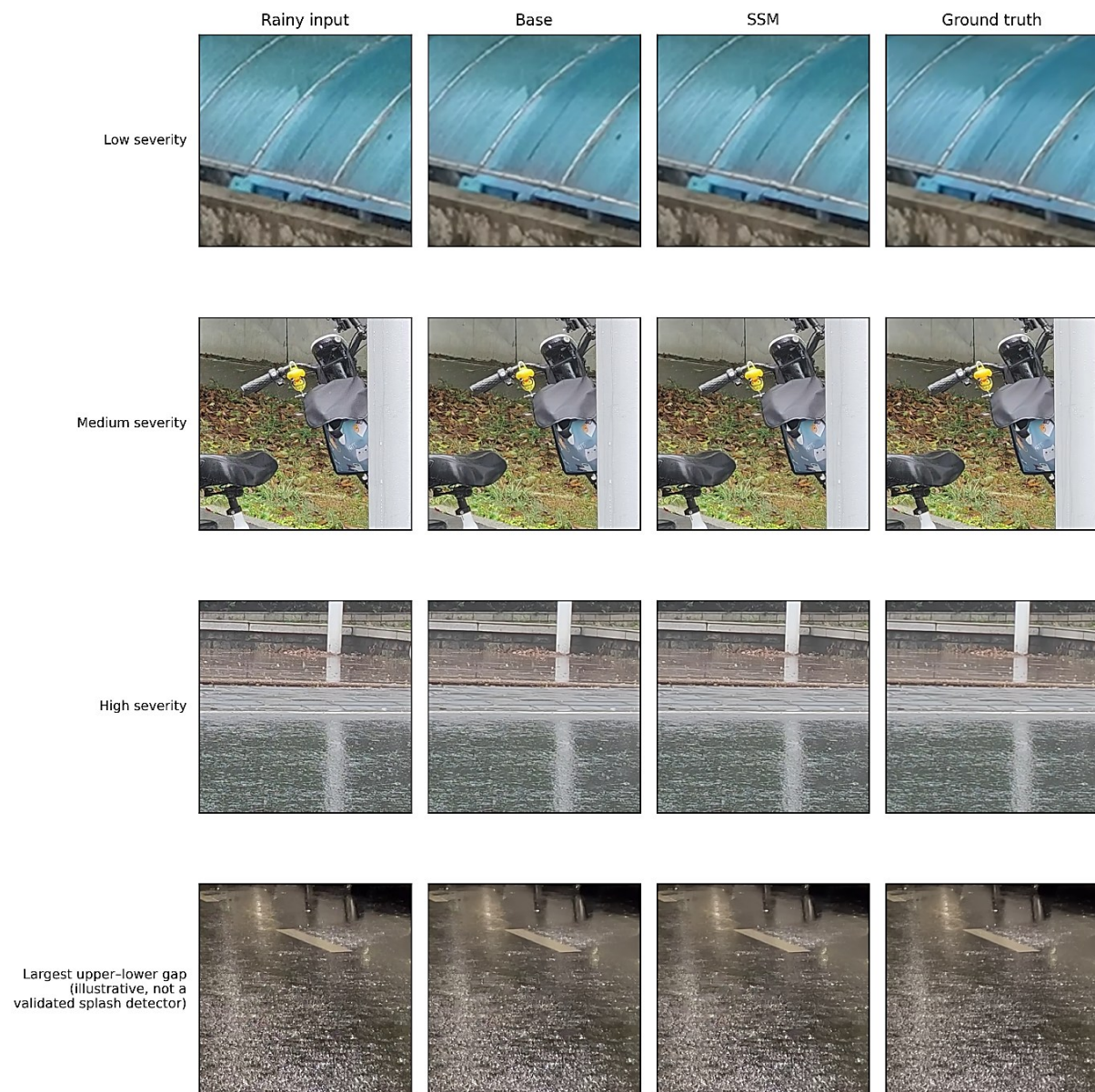


Figure 5. Qualitative comparison of rainy inputs, Base outputs, SSM outputs, and pseudo-ground-truth references. The first three rows represent the low-, medium-, and high-severity-statistic terciles.

The final row shows the preselected high-tercile image with the largest Base upper-minus-lower PSNR difference. All outputs use the seed-42 checkpoint.

5. DISCUSSION

This section interprets the results by determining what the limited conditioning gains and mismatched complexity–latency relationships imply for designing and deploying severity-aware, lightweight real-world image deraining systems.

5.1. Utility of Spatial Severity Conditioning

The matched-backbone experiment provides no evidence that the tested spatial severity prior produces a practically meaningful restoration advantage at the present scale. As shown in Table 3, SSM improved mean PSNR over Base by only 0.034 dB and over its zero-prior intervention by 0.021 dB. Figure 3 further shows that this limited aggregate improvement does not conceal a substantial benefit in any severity-statistic tercile. The seed-level equivalence analysis described in Section 3.4 and reported in Section 4.2 strengthens this interpretation: the SSM–Base difference had a 95% confidence interval of $[-0.026, +0.094]$ dB, which lies within the project-defined ± 0.1 -dB equivalence margin (TOST $p=0.021$). This result supports practical equivalence under the stated criterion rather than exact equality between the models. Its interpretation remains bounded by the project-defined margin and the availability of only three independent training replicates.

Two mechanisms may account for the limited effect. First, $S(x)$ is a deterministic transformation of the rainy image already available to PReNet. It therefore supplies an inductive bias rather than independent information, and the recurrent convolutional backbone may learn comparable local evidence directly from I . Second, the per-image percentile normalization in Eq. (3) suppresses absolute response magnitude. Although this stabilizes the spatial map, it also makes $\bar{s}$ sensitive to the distribution of oriented scene textures such as foliage, brickwork, and railings. The moderate agreement observed in Figure 2 is consistent with this interpretation.

The prior-replacement controls nevertheless suggest that the trained SSM is not entirely insensitive to its conditioning input. Shuffling or spatially flattening the map reduced performance more than replacing it with zeros, but the ordering was not consistent across PSNR, SSIM, and LPIPS. The current evidence therefore supports a narrow conclusion: this particular estimator–head combination does not provide sufficient benefit to justify its cost. It does not establish that all physical or degradation-aware priors are ineffective.

5.2. Arithmetic Efficiency Versus End-to-End Efficiency

Table 4 demonstrates that model-only complexity can misrepresent operational efficiency in two distinct ways. For the two-stage architecture, network compression was successful: the restoration model required 5.9 times fewer GFLOPs and ran 2.7 times faster than Base. However, prior generation eliminated much of that advantage and raised total latency above the unconditioned baseline. The prior therefore becomes the principal optimization target when the restoration network is already lightweight.

CBAM-light presents the converse result. Despite having the fewest parameters and GFLOPs, it produced the highest image-only latency. Parameters and GFLOPs do not account for kernel-launch overhead, memory access, operator fusion, or hardware utilization. Consequently, the paper’s compute-aware contribution lies not in declaring one architecture universally faster, but in showing that efficiency claims should include measured latency for the complete inference pathway.

These latency rankings remain specific to the tested implementation and RTX 4090 platform. Different results may arise with TensorRT, operator fusion, mixed precision, CPU vectorization, asynchronous prior generation, or deployment-class accelerators. Nevertheless, the comparison establishes that theoretical arithmetic savings cannot be assumed to translate directly into end-to-end speed.

5.3. Interpretation of Severity and Regional Behavior

The severity statistic is useful primarily as a difficulty partition. Figure 2 shows moderate agreement with the RealRain-1k density labels, while Figure 3 shows a substantial decline in restoration quality from the low to high tercile. However, these two observations do not imply that $\bar{s}$ measures rain density alone. Scene texture, illumination, residual degradation in the pseudo-ground truth, and other sources of restoration difficulty may also influence the statistic.

The regional findings reinforce this cautious interpretation. Table 5 and Figure 4 identify a consistent lower-half performance deficit, but the geometric split cannot determine whether the difference arises from splash, road texture, reflections, accumulated rain, or reference quality. SSM does not materially reduce the gap, which is compatible with an estimator designed around bright, near-vertical streaks rather than low-frequency veiling or isotropic splash. Figure 5 similarly shows little visible separation between Base and SSM, but the examples cannot establish the physical cause of the regional result.

5.4. Limitations and Scope of Inference

The primary limitation is the training scale. The models use 10,000 of the 75,838 released train-simple patches, and these cropped patches may not represent independent scenes. The numerical ranking could change with a larger training sample or longer optimization. The difference between the present results and the published values in Table 6 further indicates that the findings should not be generalized to the full-resolution LHP-Rain benchmark.

The scene-overlap audit in Section 3.1 further shows that the training samples and test split are not scene-disjoint, with 86–98% of test-scene identifiers shared with the training data depending on the training scale. The training and validation subsamples, by contrast, share no scene identifiers. Because every condition uses the same training and test lists, the overlap does not create an explicit condition-specific split advantage; however, it could interact differently with model capacity or conditioning. Consequently, the absolute PSNR and SSIM values, including the contextual comparison with published LHP-Rain results in Table 6, cannot be assumed to generalize to unseen scenes. A released scene-disjoint partition was not available for the simple splits, and constructing one remains necessary for stronger generalization claims.

The equivalence analysis is also conditional on its specified criterion. The ± 0.1 -dB margin used in Sections 3.4 and 4.2 was selected after the screening experiment but fixed before analysis of the primary configuration as the smallest PSNR difference considered operationally meaningful. Because this threshold is project-defined rather than externally established or preregistered, the equivalence result should be interpreted as evidence under the stated margin, not as an unconditional demonstration that SSM and Base are identical. The availability of only three independent training replicates further limits the precision and generalizability of the equivalence estimate.

Reference quality also bounds the interpretation. LHP-Rain uses video-derived pseudo-ground truth; consequently, residual rain and alignment errors can affect L_1 training and all reference-based metrics. This limitation may be most consequential for difficult images, where reference recovery itself may be less reliable.

The estimator is intentionally limited to bright, near-vertical, high-frequency structures. It does not explicitly model low-frequency veiling, oblique rain outside the kernel bank, or ground splash. Likewise, the upper/lower analysis uses fixed geometric regions rather than semantic labels. The reported regional gap is therefore descriptive rather than causal.

Finally, conditioning is evaluated on one common backbone, external validation is limited to RealRain-1k, and latency is measured using one software–hardware stack. Three seeds characterize training variability but limit the precision of the estimated effects. The findings should therefore be interpreted as a controlled assessment of the tested configurations rather than a universal rejection of severity-aware deraining.

6. CONCLUSIONS

This study establishes that severity conditioning must be justified by verified information use and complete inference cost, not architectural plausibility alone. Under a common 10,000-patch, 100-epoch, three-seed protocol, spatial severity modulation improved mean PSNR over the unconditioned PReNet baseline by only 0.034 dB and over its zero-prior intervention by 0.021 dB, while increasing parameters, computation, model latency, and prior-generation time. The matched-capacity control, prior-replacement interventions, seed-level equivalence analysis, and half-scale sensitivity experiment consistently showed no practically meaningful restoration advantage from the tested spatial prior. The lightweight architectures exposed a related deployment limitation: the two-stage network reduced model computation from 265.0 to 44.6 GFLOPs but became slower than Base when prior generation was included, whereas CBAM-light had the fewest parameters and GFLOPs yet the highest image-only latency. Model size and arithmetic complexity are therefore insufficient indicators of operational efficiency; the entire inference pathway must be measured on the intended platform. The label-free statistic remains useful for reproducible difficulty stratification, but its cross-scene AUC of 0.772 indicates only moderate agreement with rain-density labels and precludes its interpretation as ground-truth density. A severity prior is operationally valuable only when its restoration benefit survives controlled falsification and outweighs its end-to-end computational cost. Future work should evaluate larger, scene-disjoint training configurations, introduce semantically annotated weather regions, and benchmark optimized end-to-end inference on deployment-class accelerators.

ACKNOWLEDGEMENT

This work is supported by the AOARD project FA2386-25-1-4061 and recorded at IIUM SPI25-295-0295. The authors used an AI coding assistant for clarity and proofreading, reviewed and verified all AI-assisted content, and take full responsibility for the accuracy, originality, and integrity of the published work.

REFERENCES

- [1] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, "Perception and sensing for autonomous vehicles under adverse weather conditions: A survey," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 196, pp. 146-177, 2023.
- [2] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019: IEEE, pp. 3937-3946.
- [3] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: IEEE, pp. 695-704.
- [4] S. Yamashita and M. Ikehara, "Image deraining with frequency-enhanced state space model," in *Asian Conference on Computer Vision*, 2024: Springer, pp. 318-334.

- [5] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5728-5739.
- [6] B. Xiaojun, C. Zheng, Y. Jianyu, and W. Haibo, "LRP-net: A lightweight recursive pyramid network for single image deraining," *Neurocomputing*, vol. 490, pp. 181-192, 2022.
- [7] M. Hu, J. Yang, N. Ling, Y. Liu, and J. Fan, "Lightweight single image deraining algorithm incorporating visual saliency," *IET Image Processing*, vol. 16, no. 12, pp. 3190-3200, 2022.
- [8] Y. Liang, S. Anwar, and Y. Liu, "Drt: A lightweight single image deraining recursive transformer," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022: IEEE, pp. 589-598.
- [9] Y. Ba *et al.*, "Not just streaks: Towards ground truth for single image deraining," in *European Conference on Computer Vision*, 2022: Springer, pp. 723-740.
- [10] Y. Guo, X. Xiao, Y. Chang, S. Deng, and L. Yan, "From sky to the ground: A large-scale benchmark and simple baseline towards real rain removal," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023: IEEE, pp. 12097-12107.
- [11] E. Perez, F. Strub, H. De Vries, V. Dumoulin, and A. Courville, "FiLM: Visual reasoning with a general conditioning layer," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, vol. 32, no. 1.
- [12] X. Wang, K. Yu, C. Dong, and C. C. Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: IEEE, pp. 606-615.
- [13] T. S. Gunawan and M. Kartiwi, "Stability-aware physics-prior conditioning for lightweight image dehazing," in *The 12th IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA) 2026*, Kuching, Sarawak, 29-30 July 2026, pp. 210-215.
- [14] T. S. Gunawan and M. Kartiwi, "CIART: A compact color-imbalance-conditioned transformer for sandstorm image restoration," in *The 12th IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA) 2026*, Kuching, Sarawak, 29-30 July 2026, pp. 221-226.
- [15] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *European Conference on Computer Vision*, 2018: Springer, pp. 3-19.
- [16] W. Li, Q. Zhang, J. Zhang, Z. Huang, X. Tian, and D. Tao, "Toward real-world single image deraining: A new benchmark and beyond," *arXiv preprint arXiv:2206.05514*, 2022.
- [17] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2013.
- [18] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-net: All-in-one dehazing network," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017: IEEE, pp. 4770-4778.
- [19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [20] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: IEEE, pp. 586-595.
- [21] D. Lakens, "Equivalence tests: A practical primer for t tests, correlations, and meta-analyses," *Social psychological and personality science*, vol. 8, no. 4, pp. 355-362, 2017.

Similar Accuracy, Different Explanations in YOLO-Based Palm Fruit Ripeness Detection

QUEENTA PARADISSA RAMADHANI¹, FAVIAN DEWANTA¹,
TEDDY SURYA GUNAWAN^{2*}, MIRA KARTIWI³, AND BAMBANG SETIA NUGROHO¹

¹ School of Electrical Engineering, Telkom University, Bandung, Indonesia.

² Department of Electrical and Computer Engineering, Kulliyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia

³ Department of Information Systems, Kulliyah of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia

*Corresponding author: tsgunawan@iium.edu.my

(Received: 31 July 2026; Accepted: 30 August 2026; Published online: 11 September 2026)

ABSTRACT: Modern palm-fruit detectors can achieve similarly high benchmark accuracy, yet class activation mapping (CAM) methods are often applied without establishing whether their explanatory behavior transfers across detector configurations. This study is a controlled exploratory within-dataset comparison: model scale and target layer were selected on the same 416-image test partition used for final CAM evaluation, so the reported values are not independent estimates of explanation generalization. We evaluated 15 YOLOv8, YOLO11, and YOLO26 variants under a matched training protocol and assessed representative medium models across three random seeds. Backbone representations were then audited on the 416 test images using class-agnostic EigenCAM, backbone-adapted Grad-CAM++ (bGrad-CAM++), and backbone-adapted Score-CAM (bScore-CAM). Explanation behavior was quantified with CAM-box intersection over union (IoU) and Pointing Game for spatial localization and with Mean Absolute Confidence Drop (MACD) and Increase in Confidence as image-level confidence-retention diagnostics. Repeated-run mean mAP_{50–95} was narrowly separated at 0.903 ± 0.002 , 0.899 ± 0.002 , and 0.901 ± 0.003 for YOLOv8m, YOLO11m, and YOLO26m, respectively. In contrast, Pointing Game varied by 68.03 percentage points and MACD by 48.62 percentage points across model–explainer configurations. EigenCAM showed the strongest observed profile on YOLOv8m, whereas bScore-CAM achieved the highest localization on YOLO11m; no explainer dominated across all evaluated backbones. The ranking reversal was also present within the difficult underripe, ripe, and overripe subset, and the recorded layer search showed that layer 8 was the strongest compatible backbone target across the layer 6–10 window in all three detectors. Similar detection accuracy therefore did not imply transferable explanation behavior. CAM explanations should be validated jointly with the explanatory target, layer, and detector rather than treated as interchangeable post-hoc visualizations.

ABSTRAK: Pengesanan buah sawit moden boleh mencapai ketepatan penanda aras tinggi dan hampir setara, namun kaedah pemetaan pengaktifan kelas (CAM) sering diguna tanpa memastikan sama ada tingkah laku penjelasannya boleh dipindah merentas konfigurasi pengesanan. Kajian ini merupakan perbandingan penerokaan terkawal dalam set data: skala model dan lapisan sasaran dipilih menggunakan pecahan ujian 416 imej sama seperti penilaian CAM akhir, di mana dapatan yang dilaporkan bukan anggaran bebas bagi generalisasi penjelasan. Sebanyak 15 varian YOLOv8, YOLO11 dan YOLO26 dinilai menggunakan protokol latihan sepadan, manakala model sederhana yang mewakili setiap keluarga dinilai menggunakan tiga benih rawak. Perwakilan tulang belakang kemudiannya diaudit pada 416 imej ujian menggunakan EigenCAM yang bebas kelas, Grad-CAM++

teradaptasi tulang belakang (bGrad-CAM++), dan Score-CAM teradaptasi tulang belakang (bScore-CAM). Tingkah laku penjelasan dinilai menggunakan CAM-box IoU dan Pengarahan Perburuan bagi menempat spatial, serta Purata Mutlak Keyakinan Menurun (MACD) dan Keyakinan Bertambah sebagai diagnostik pengekal keyakinan pada aras imej. Purata mAP50–95 bagi larian berulang masing-masing adalah 0.903 ± 0.002 , 0.899 ± 0.002 dan 0.901 ± 0.003 untuk YOLOv8m, YOLO11m dan YOLO26m. Sebaliknya, Pengarahan Pemburuan berbeza sebanyak 68.03 mata peratusan dan MACD sebanyak 48.62 mata merentas konfigurasi model–penerang. EigenCAM menunjukkan profil pemerhatian terkuat pada YOLOv8m, manakala bScore-CAM mencapai penyetempatan tertinggi pada YOLO11m; tiada penerang yang mendominasi semua tulang belakang yang dinilai. Pembalikan kedudukan ini turut hadir dalam subset sukar seperti kurang masak, masak dan terlebih masak, manakala carian lapisan yang direkod menunjukkan bahawa lapisan 8 ialah sasaran tulang belakang serasi terkuat dalam tettingkap lapisan 6–10 bagi ketiga-tiga pengesan. Dapatan ini menunjukkan bahawa ketepatan pengesanan yang hampir setara tidak semestinya menghasilkan tingkah laku penjelasan yang boleh dipindahkan. Oleh itu, penjelasan CAM perlu disahkan secara bersama dengan sasaran penjelasan, lapisan dan pengesan.

KEYWORDS: *Explainable Artificial Intelligence; Class Activation Mapping; Object Detection; YOLO; Oil Palm Fruit.*

1. INTRODUCTION

Oil palm fresh fruit bunch (FFB) ripeness is an operational quality determinant, not merely a visual classification problem. Harvesting bunches outside the appropriate maturity stage can reduce recoverable oil yield or compromise crude palm oil quality, while inconsistent grading can propagate economic losses from harvesting through mill processing. This problem is substantial. Malaysia's oil palm sector generated RM109.39 billion in export revenue in 2024, yet the national oil extraction rate declined from 19.86% in 2023 to 19.67% in 2024 despite a 5.8% increase in FFB yield [1]. This decline was partly attributed to lower-quality FFB processed following prolonged rainfall and flooding [2], illustrating that increased fruit yield does not necessarily translate into greater extraction efficiency. Reliable and auditable ripeness assessment therefore remains important for both operational performance and economic value.

Computer vision provides a scalable approach to improving grading consistency, and one-stage YOLO detectors are particularly attractive because they integrate localization, classification, and real-time inference. Previous work identified YOLOv8m as offering a favorable accuracy–computational cost trade-off for six-class FFB detection [3], while subsequent studies have continued to compare detector generations primarily using precision, recall, inference speed, and mean average precision [4]. These metrics quantify predictive performance but provide limited evidence about the visual representations underlying a model's decisions. Two detectors can achieve nearly identical benchmark accuracy while relying on markedly different spatial features, and a visually plausible heatmap does not necessarily indicate that an explanation is informative or transferable across models.

Class activation mapping (CAM) methods, including EigenCAM, Grad-CAM++, and Score-CAM, are widely used to visualize salient image regions. However, these methods were developed largely for classifier-style networks with tensor outputs and scalar class scores. Modern object detectors present a more complex explanatory setting: they produce multiple bounding boxes, class labels, and confidence scores; their necks combine skip-connected feature maps through multi-path operations; and recent architectures increasingly modify the detection and post-processing pathway. Consequently, applying a classifier-oriented CAM

method to an object detector requires explicit choices about the explanatory target and target layer. Without these choices, a plausible saliency map may be incorrectly interpreted as attribution of a specific final box–class prediction.

The resulting research gap is therefore not another comparison of YOLO generations, but whether explanation behavior transfers across detector configurations with closely matched predictive performance. Differences in detector configuration can coincide with differences in learned representations, while CAM methods themselves interrogate those representations through distinct activation-, gradient-, or perturbation-based mechanisms. Whether their relative behavior remains consistent across closely performing YOLO detectors has received limited quantitative examination in FFB ripeness detection.

Accordingly, this study asks whether YOLO detectors with practically similar ripeness-detection accuracy exhibit similarly consistent backbone-level CAM explanations across detector–explainer configurations. The study makes three contributions. First, it provides a controlled exploratory comparison showing that explainer rankings can change across closely performing YOLOv8m, YOLO11m, and YOLO26m configurations. Second, it establishes a matched benchmark of 15 YOLO variants, supported by three-seed evaluations of the representative medium models, to place explanation differences alongside small variations in predictive accuracy. Third, it presents a backbone-level compatibility procedure for EigenCAM, backbone-adapted Grad-CAM++ (bGrad-CAM++), and backbone-adapted Score-CAM (bScore-CAM) that preserves trained weights while explicitly distinguishing feature-representation auditing from attribution of a final box–class prediction.

Two scope conditions apply throughout. First, all explanation results are strictly within-dataset because model-scale and target-layer selection reused the 416-image test partition used for final CAM evaluation. The reported metrics therefore describe relative behavior on this dataset rather than out-of-sample explanation generalization. Second, the explanatory target is the pre-neck backbone representation rather than a final box–class decision. The maps audit spatial selectivity in the selected representation and do not provide causal attributions for individual detections.

2. RELATED WORK

Research on automated fresh fruit bunch (FFB) grading has advanced rapidly in predictive performance, particularly through recent YOLO-based detectors. However, the methodological basis for interpreting and auditing those detections remains less mature. This gap matters because high benchmark accuracy does not ensure that the underlying visual evidence is spatially meaningful, stable across detector configurations, or suitable for post-hoc explanation.

2.1. FFB Ripeness Detection and YOLO Benchmarking

FFB ripeness assessment has progressed from manual color inspection and specialized optical sensing toward RGB-based deep learning. Camera-based systems are non-destructive, comparatively inexpensive, and readily integrated into harvesting and mill workflows [5, 6]. Nevertheless, reliable ripeness discrimination remains difficult because biologically adjacent stages, particularly underripe, ripe, and overripe, exhibit gradual rather than discrete changes in color, texture, and fruitlet detachment. Field imagery further introduces substantial variability in illumination, viewing angle, distance, occlusion, and plantation background [7]. Consequently, near-ceiling mAP50 on curated datasets does not imply that ripeness boundaries are fully resolved or that comparable performance will generalize to field conditions.

YOLO-based approaches have established strong FFB detection baselines. Comparisons across YOLOv8 scales have shown favorable accuracy–cost trade-offs for medium-sized models [3], while cross-generation studies have evaluated YOLOv8, YOLOv9, and optimized YOLOv7 variants primarily in terms of precision, recall, mean average precision, and inference speed [4, 8]. YOLOv8 has also achieved near-saturated performance on curated ripeness datasets [9]. However, reported results are difficult to compare directly because studies vary in class definitions, dataset composition, augmentation, training configuration, hardware, and model-selection criteria. More broadly, controlled agricultural benchmarks indicate that the relative ranking of recent YOLO generations is sensitive to both dataset and evaluation protocol [10]. Cross-generation comparisons are therefore most informative when data, training, model selection, and evaluation conditions are held constant.

2.2. Evolving Detector Architectures

Successive YOLO generations differ not only in computational cost but also in the internal representations exposed to explanation methods. YOLOv8 uses an anchor-free design with C2f-based feature extraction and decoupled multiscale prediction [11]. YOLO11 replaces key C2f stages with C3k2 blocks and incorporates C2PSA attention while retaining the conventional backbone–neck–head organization [10]. YOLO26 extends the design toward deployment-oriented end-to-end detection through an NMS-free inference pathway and lighter regression components, reflecting a broader movement toward reducing dependence on post-processing [12].

These changes do not guarantee monotonic gains in predictive accuracy. Controlled architectural analyses instead indicate shifting accuracy–latency trade-offs across model generations [13]. This distinction matters for explainability because CAM methods operate on intermediate activations, gradients, or perturbation responses rather than mAP itself. Two detectors can therefore produce similar final predictions while exposing materially different internal representations to the same explanatory procedure. The present study evaluates that empirical possibility without attributing the observed differences to any single architectural block.

2.3. CAM Methods and the Detector Barrier

CAM methods interrogate learned representations through different mechanisms. Grad-CAM and Grad-CAM++ derive saliency from gradient information, with Grad-CAM++ using higher-order weighting to improve localization in complex scenes [14, 15]. EigenCAM is class-agnostic and derives saliency from the dominant principal component of the activation tensor [16], whereas Score-CAM estimates channel importance through masked forward responses rather than gradients [17]. Because these methods probe different properties of the learned representation, their relative performance may not transfer unchanged across detector configurations.

A further difficulty is that conventional CAM methods were developed primarily for classifier-style networks with tensor outputs and scalar class scores. Object detectors instead produce multiple boxes, classes, and confidence values, making the explanatory target inherently more complex. Detector-specific methods such as D-RISE address this issue by combining category confidence with bounding-box similarity to explain complete detection outputs [18]. Quantitative studies of detector explanations also distinguish localization, faithfulness, and computational cost, emphasizing that visually plausible heatmaps may correspond only weakly to actual detector behavior [19, 20]. In agricultural imaging, CAM-

family explanations are increasingly common, but quantitative validation remains comparatively limited [21].

2.4. Research Positioning

Detector-output attribution and backbone-level feature auditing address different explanatory questions. The former asks why a particular box–class prediction was produced; the latter asks whether the representation entering subsequent neck fusion and detection stages is spatially selective for the object of interest. This distinction matters because predictive accuracy can conceal substantial differences in internal feature organization.

Table 1. Comparison of Prior Research and the Positioning of the Present Study

Research focus	Typical evidence	Explanatory target	Key limitation relative to this study
FFB detector benchmarking	mAP, precision, recall, latency	Final detector output	Does not quantify explanation behavior
Agricultural CAM visualization	Qualitative or limited quantitative saliency	Intermediate feature maps	Often evaluates one detector–explainer pairing
Detector-specific XAI, e.g., D-RISE	Box- and class-conditioned attribution	Complete detection decision	Addresses final prediction attribution rather than backbone representation
This study	<i>Matched detection benchmark, CAM-box IoU, Pointing Game, confidence-retention diagnostics</i>	<i>Pre-neck backbone representation</i>	<i>Tests whether explanation behavior remains consistent across closely performing YOLO configurations</i>

Accordingly, this study treats backbone saliency as an explanatory target in its own right rather than as a surrogate for complete detector attribution. EigenCAM is applied in its original class-agnostic form, whereas Grad-CAM++ and Score-CAM are adapted to explicit backbone targets as bGrad-CAM++ and bScore-CAM. The central question is whether the relative localization and confidence-retention behavior of these feature-auditing methods remains stable across YOLOv8m, YOLO11m, and YOLO26m when predictive accuracy is practically close.

3. METHODOLOGY

This section describes the matched detector benchmark, backbone-level CAM compatibility procedure, saliency evaluation, and statistical analysis used to determine whether closely performing YOLO detectors exhibit consistent explanation behavior. The experimental design separates predictive performance from backbone-level feature auditing so that differences in CAM behavior are not conflated with differences in detector accuracy.

3.1. Problem Formulation and Experimental Overview

Let an input image be denoted by $x \in [0,1]^{3 \times H_0 \times W_0}$, where $H_0 = W_0 = 640$. A detector D_θ , parameterized by θ , predicts a set of J detections,

$$D_\theta(x) = [(\hat{b}_j, \hat{c}_j, \hat{p}_j)]_{j=1}^J \quad (1)$$

where $\hat{b}_j$ is a predicted bounding box, $\hat{c}_j \in \mathcal{C}$ is its predicted class, and $\hat{p}_j \in [0,1]$ is its confidence score. The class set is $\mathcal{C} = [\text{unripe, underripe, ripe, overripe, empty bunch, abnormal}]$. The detection objective is to localize each FFB and assign it to the appropriate Malaysian Palm Oil Board (MPOB) ripeness class. The explanatory objective is deliberately

narrower. Given a trained detector D_θ , an explainer E_m , and a target backbone layer L , the objective is to generate a normalized spatial saliency map

$$S_{m,L}(x) \in [0,1]^{H_0 \times W_0}, \quad (2)$$

and evaluate its spatial correspondence with annotated FFB regions and its ability to preserve image-level detector confidence after saliency-based masking.

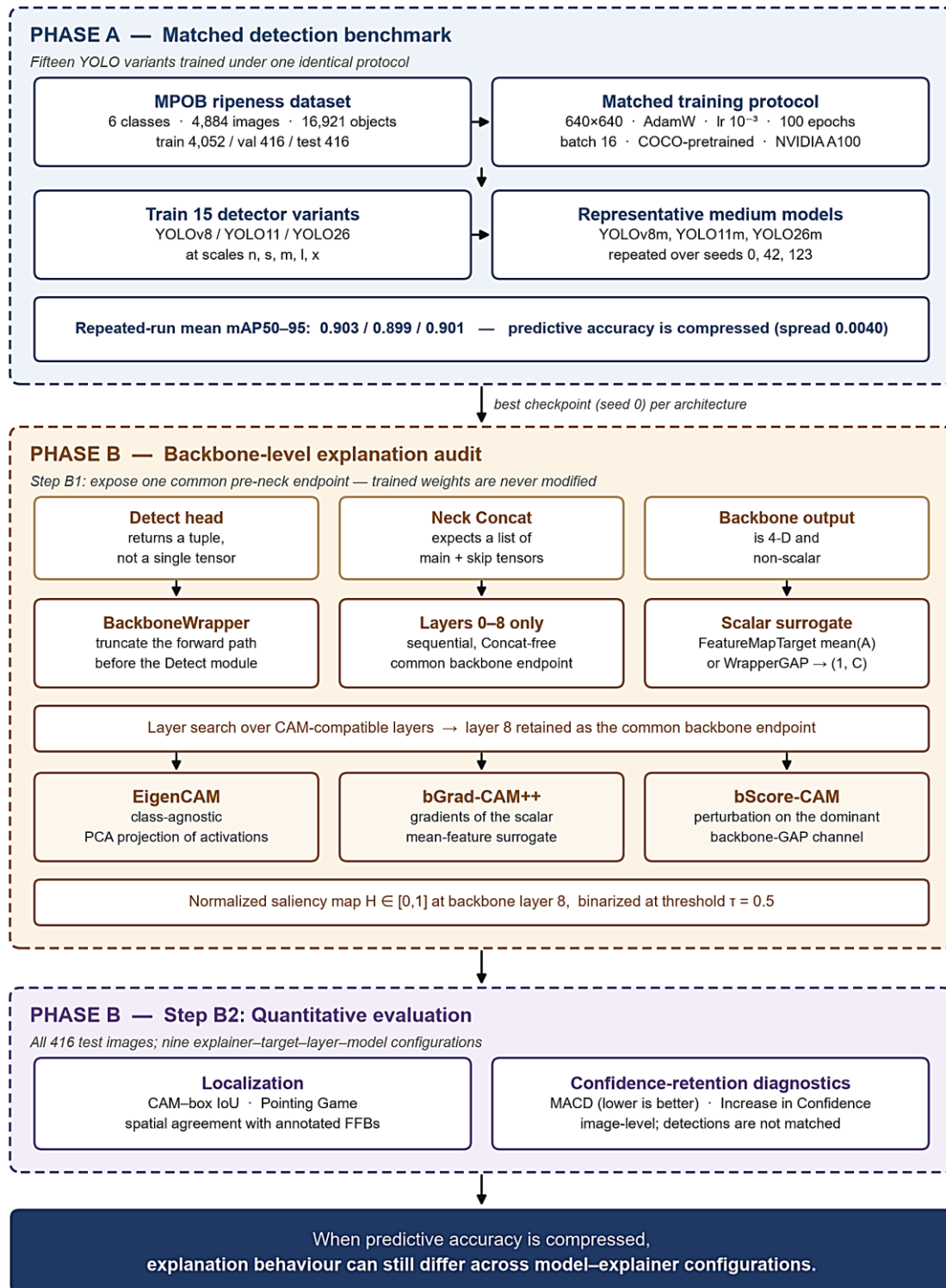


Figure 1. Two-phase experimental framework.

The explanatory target is defined at the *backbone representation*, rather than at the final detector output. Accordingly, $S_{m,L}(x)$ characterizes whether the selected pre-neck feature representation is spatially concentrated on annotated FFB regions before multiscale feature fusion and final box–class prediction. The resulting saliency maps therefore support backbone-level feature auditing and should not be interpreted as causal attribution of a specific final detection. This objective complements box- and class-conditioned detector attribution rather than replacing it.

The experimental pipeline comprises two phases, as illustrated in Figure 1. Phase A benchmarks 15 detector configurations formed by three YOLO families, including YOLOv8, YOLO11, and YOLO26, at five model scales: nano, small, medium, large, and extra-large. We then evaluate the medium variant from each family across three random seeds to characterize variability in predictive performance. Phase B exposes a common compatible pre-neck backbone endpoint from the three medium detectors to EigenCAM, backbone-adapted Grad-CAM++ (bGrad-CAM++), and backbone-adapted Score-CAM (bScore-CAM), after which the resulting saliency maps are evaluated on the test set.

Because YOLOv8m, YOLO11m, and YOLO26m achieve practically close mAP50–95 under the matched training protocol, Phase B addresses the central methodological question of the study: whether similarly accurate detectors also exhibit similarly consistent explanation behavior across model–explainer configurations.

3.2. Dataset and Preprocessing

The experiments used a publicly available oil palm fresh fruit bunch (FFB) dataset developed for six-class ripeness grading [22]. The original dataset comprised 4,160 images with predefined training, validation, and test partitions in an approximately 80:10:10 ratio. The six classes were abnormal, empty bunch, overripe, ripe, underripe, and unripe.

Because the empty-bunch class was underrepresented in the training partition, with 266 dominant-class images, offline augmentation was applied exclusively to training images dominated by this class. The augmentation operations comprised horizontal and vertical flipping, 90° rotation, hue–saturation–value adjustment, brightness and contrast modification, blur, additive noise, synthetic shadow, and contrast-limited adaptive histogram equalization. This procedure increased the empty-bunch subset to 990 dominant-class images and yielded a final dataset of 4,884 unique images containing 16,921 annotated object instances. Neither the validation nor the test partition was modified by offline augmentation.

Table 2. Dataset distribution after training-only offline augmentation

Class	Dominant-class images	Images containing class	Object instances
Abnormal	869	1,237	2,599
Empty bunch	990	1,197	2,574
Overripe	658	1,366	2,845
Ripe	606	2,080	3,173
Underripe	819	1,332	2,618
Unripe	942	1,274	3,112
Total	4,884	8,486	16,921

Table 2 reports the resulting class distribution using three complementary counts. Dominant-class images assign each image to the class with the most annotated objects. Images containing class count an image once for each class represented in that image; consequently, these counts do not sum to the total number of unique images. Object instances represent

individual bounding-box annotations. This distinction matters because a single image may contain multiple FFBs from different ripeness classes.

During detector training, images were resized to 640×640 pixels. The Ultralytics training pipeline additionally applied online mosaic, mixup, HSV transformation, scaling, translation, and RandAugment, with mosaic augmentation disabled during the final ten epochs. For CAM analysis, each test image was resized to 640×640, converted to RGB, normalized to [0,1], and represented as a 1×3×640×640 tensor. The unchanged test partition contained 416 images and 1,493 annotated FFB instances.

3.3. Detection Models and Training Protocol

The detector benchmark comprised 15 configurations obtained by combining three YOLO families, including YOLOv8, YOLO11, and YOLO26, with five model scales: nano, small, medium, large, and extra-large. All configurations were trained under the same externally controlled protocol to minimize confounding from model-specific hyperparameter tuning. Each model was initialized from COCO-pretrained weights and trained for 100 epochs using AdamW with an initial learning rate of 1×10^{-3} , a batch size of 16, and 640×640 inputs. Experiments were conducted on an NVIDIA A100-SXM4-40GB GPU using Ultralytics 8.4.41 and PyTorch 2.11.0. Table 3 summarizes the training configuration.

Table 3. Externally controlled training configuration applied to all 15 detector variants.

Parameter	Value
Optimizer	AdamW
Initial learning rate	0.001
Epochs	100
Batch size	16
Input resolution	640×640
Initialization	COCO-pretrained weights
Initial benchmark seed	0
Repeated-training seeds	0, 42, and 123
Hardware	NVIDIA A100-SXM4-40GB

The initial benchmark used seed 0. Within each YOLO family, we ranked the five scales by test-set mAP50–95, and carried forward the highest-scoring scale to the explanation analysis. We selected the medium variant for all three families. To assess whether the small performance differences among these representative models were stable across training runs, we additionally trained YOLOv8m, YOLO11m, and YOLO26m using seeds 42 and 123. The subsequent CAM analysis used the primary seed-0 checkpoint, while the additional checkpoints were used only to characterize variation in detection performance.

Because model-scale selection used test-set mAP50–95, the test partition was not fully independent of model selection. The same partition was then used for target-layer selection and CAM evaluation. These choices may introduce selection optimism; therefore, the explanation results are interpreted as a controlled exploratory within-dataset comparison rather than an independent estimate of explanation generalization. An independent study should select model scale and explanatory layer using validation data before evaluating CAM behavior on a separate test set.

3.4. Backbone-Level Compatibility Procedure

Conventional CAM implementations could not be applied directly to the complete YOLO computation graph because of differences in detector outputs, internal feature routing, and

explanatory target requirements. Rather than modifying the trained weights, we used a backbone-level compatibility procedure to expose a common intermediate representation to each CAM method. The procedure addresses the three implementation barriers summarized in Table 4 while explicitly restricting interpretation to pre-neck backbone feature auditing.

Let B_L denote the YOLO backbone truncated after layer L . For the common compatible endpoint $L = 8$,

$$A = B_8(x) \in \mathbb{R}^{1 \times K \times h \times w}, \quad (3)$$

where K is the number of channels and $h \times w$ is the spatial resolution of the activation tensor. In YOLOv8m, YOLO11m, and YOLO26m, layers 0–8 form a sequential backbone path before the multi-input concatenation operations used in the neck. Layer 8 therefore provides a common evaluable representation while avoiding both structured detection-head outputs and neck operations that a simple sequential wrapper cannot reproduce.

Table 4. CAM–YOLO incompatibilities and corresponding backbone-level compatibility procedures.

Incompatibility	Observed error	Cause	Procedure response
Detection-head output	tuple object has no attribute CPU	The complete detector returns structured detection outputs rather than a single tensor compatible with the CAM interface.	Terminate the explanatory forward path before the detection head.
Neck concatenation	Invalid cat() argument combination	Neck concatenation requires multiple tensors from main and skip paths, whereas a sequential wrapper propagates a single tensor.	Restrict the common explanatory path to the sequential backbone through layer 8.
Non-scalar backbone output	Scalar-gradient or tensor-to-scalar error	The truncated backbone produces a four-dimensional activation tensor rather than the scalar or vector target required by the adapted CAM methods.	Use the scalar mean target in Eq. (4) for bGrad-CAM++ and global average pooling in Eq. (5) for bScore-CAM; EigenCAM requires no target adaptation.

The three CAM methods require different interfaces to this backbone representation. EigenCAM operates directly on A because it is class-agnostic and requires neither a scalar class score nor a gradient-defined target. By contrast, conventional Grad-CAM++ and Score-CAM are formulated around class-discriminative targets and therefore require explicit adaptation when the explanatory target is a backbone representation rather than a detector class output. To distinguish these implementations from their original formulations, we denote them as backbone-adapted Grad-CAM++ (bGrad-CAM++) and backbone-adapted Score-CAM (bScore-CAM).

For bGrad-CAM++, the four-dimensional backbone activation is reduced to the scalar surrogate

$$q_G(A) = \frac{1}{Khw} \sum_{k=1}^K \sum_{i=1}^h \sum_{j=1}^w A_{ij}^k. \quad (4)$$

which provides the scalar target required for gradient computation. The resulting gradients therefore characterize the sensitivity of the aggregated backbone representation rather than the sensitivity of a particular FFB class score.

For bScore-CAM, the truncated backbone is followed by global average pooling,

$$z(x) = \text{GAP}(B_8(x)) \in \mathbb{R}^K, \quad z_k = \frac{1}{hw} \sum_{i=1}^h \sum_{j=1}^w A_{ij}^k \quad (5)$$

producing a channel-response vector that supports perturbation-based weighting. As described in Section 3.5, the dominant pooled backbone channel is then used as the feature-level target for each image. This target is a learned backbone feature response, not one of the six FFB ripeness classes.

These adaptations preserve the trained detector parameters and the core weighting principles of the three CAM methods while placing them on a common explanatory interface. EigenCAM retains its original class-agnostic formulation, whereas bGrad-CAM++ and bScore-CAM replace their conventional class-oriented targets with explicitly defined backbone targets.

The resulting saliency maps therefore characterize the spatial selectivity of the selected *pre-neck backbone representation*. They should not be interpreted as causal explanations of a specific final bounding box or ripeness-class prediction. Saliency generation is defined mathematically in Section 3.5, while target-layer selection and quantitative evaluation are described in Sections 3.6 and 3.7, respectively.

3.5. Backbone-Level CAM Formulations

Let $A = [A_k]_{k=1}^K$, $A_k \in \mathbb{R}^{h \times w}$, denote the activation tensor at the selected backbone layer. Let $\mathcal{U}(\cdot)$ denote bilinear upsampling to the input resolution and $\mathcal{N}(\cdot)$ denote min-max normalization to $[0,1]$.

3.5.1. EigenCAM

EigenCAM derives a class-agnostic saliency map from the dominant variation in the backbone activation tensor. The tensor is reshaped as $Z \in \mathbb{R}^{n \times K}$, $n = hw$, where each row represents a spatial location and each column an activation channel. After channel-wise centering,

$$\tilde{Z} = Z - \mathbf{1}\mu^T, \mu = \frac{1}{n} Z^T \mathbf{1}. \quad (6)$$

The channel covariance matrix and its dominant eigenvector are

$$C = \frac{1}{n-1} \tilde{Z}^T \tilde{Z}, C v_1 = \lambda_1 v_1, \quad (7)$$

where v_1 corresponds to the largest eigenvalue. The raw EigenCAM map is

$$R^E = \text{reshape}(\tilde{Z} v_1) \quad (8)$$

and the final normalized heatmap is

$$S^E = \mathcal{N}[\mathcal{U}(\text{ReLU}(R^E))]. \quad (9)$$

Because EigenCAM does not require a class score, S^E represents dominant spatial variation in the selected backbone representation rather than evidence for a specific ripeness class.

3.5.2. Backbone-Adapted Grad-CAM++ (bGrad-CAM++)

bGrad-CAM++ retains the Grad-CAM++ weighting mechanism but replaces the conventional class score with the backbone-derived scalar target $q_G(A)$ defined in Eq. (4). The gradient at each spatial location is

$$g_{i,j}^k = \frac{\partial q_G}{\partial A_{i,j}^k}. \quad (10)$$

The implementation-equivalent pixel coefficient is

$$\alpha_{i,j}^k = \frac{(g_{i,j}^k)^2}{2(g_{i,j}^k)^2 + (\sum_{a=1}^h \sum_{b=1}^w A_{ab}^k)(g_{i,j}^k)^3 + \varepsilon}, \quad (11)$$

where $\varepsilon > 0$ prevents division by zero. The channel weight is

$$w_k^G = \sum_{i=1}^h \sum_{j=1}^w \alpha_{i,j}^k \text{ReLU}(g_{i,j}^k), \quad (12)$$

yielding

$$S^G = \mathcal{N} \left[\mathcal{U} \left(\text{ReLU}(\sum_{k=1}^K w_k^G A_k) \right) \right]. \quad (13)$$

Because $q_G(A)$ aggregates the backbone representation rather than a detector class score, S^G is interpreted as gradient-weighted backbone saliency and is not class-discriminative with respect to the six FFB labels.

3.5.3. Backbone-Adapted Score-CAM (bScore-CAM)

For bScore-CAM, each activation channel is first upsampled and normalized to form a spatial mask,

$$M_k = \mathcal{N}[\mathcal{U}(A_k)]. \quad (14)$$

The corresponding masked input is

$$x_k = x \odot M_k, \quad (15)$$

where M_k is broadcast across the RGB channels. Using the pooled backbone response $z(x)$ from Eq. (5), the dominant feature channel is selected as,

$$k^* = \arg \max_k z_k(x). \quad (16)$$

Each masked input x_k is then forwarded through the same backbone-GAP wrapper, and its response at the fixed target channel k^* determines the Score-CAM weight w_k^S . The resulting heatmap is

$$S^S = \mathcal{N} \left[\mathcal{U} \left(\text{ReLU}(\sum_{k=1}^K w_k^S A_k) \right) \right]. \quad (17)$$

Importantly, k^* denotes a backbone feature channel rather than an FFB ripeness class. Thus, bScore-CAM preserves the perturbation-based principle of Score-CAM while explaining the dominant pooled backbone response rather than a specific final box–class prediction.

3.6. Target-Layer Selection

CAM behavior depends on the depth of the selected representation. Shallower layers generally preserve finer spatial detail, whereas deeper layers encode increasingly abstract semantic information [23]. We therefore conducted a target-layer search separately for each representative medium detector. For candidate layer L , localization quality was summarized as

$$Q_L = \frac{1}{2} \left[\frac{1}{N} \sum_{i=1}^N \text{IoU}_{i,L} + \frac{1}{N} \sum_{i=1}^N \text{PG}_{i,L} \right], \quad (18)$$

where CAM-box IoU and Pointing Game (PG) are both expressed on the [0,1] scale. We excluded layers associated with the final Detect module.

The highest observed localization scores occurred in the neck: layer 18 for YOLOv8m and layer 19 for YOLO11m and YOLO26m. These layers require multi-input, list-based concatenation and therefore cannot be exposed through the common sequential wrapper described in Section 3.4. Selecting them would have required a different explanatory pathway for each detector and would have confounded the cross-model comparison with implementation differences, which is precisely the confound this design is intended to avoid; Table 4 documents the specific failures encountered. Layer 8 was consequently selected as the strongest compatible backbone layer and the common evaluable endpoint across all three detectors. This choice should not be interpreted as evidence that layer 8 is universally optimal for CAM analysis. Rather, it reflects a trade-off between observed localization performance and implementation compatibility. To show that this is not a knife-edge choice within its neighborhood, we also report the layer search over the full compatible window from layer 6 to layer 10 in Section 4.3. The target layer is therefore part of the explanatory design and must be reported explicitly.

The layer search used all 416 test images, which were subsequently reused for final CAM evaluation. This reuse may favor the selected layer and introduces selection optimism. Accordingly, the target-layer analysis and CAM results are interpreted as exploratory within-dataset evidence rather than independent estimates of explanation generalization.

3.7. Evaluation Metrics

3.7.1. Detection Performance

We assessed detection performance using precision, recall, F1-score, mAP50, and mAP50–95. The primary predictive metric was mAP50–95, which averages class-level average precision over the IoU thresholds $\mathcal{T} = [0.50, 0.55, \dots, 0.95]$. We reported computational efficiency using end-to-end inference time, frames per second, and serialized model size. Because latency depends on both hardware and software configuration, these values characterize execution on the stated NVIDIA A100 platform and should not be interpreted as hardware-independent measures.

Latency benchmarking used the native PyTorch 2.11.0 Ultralytics 8.4.41 validation runtime on the NVIDIA A100-SXM4-40GB GPU recorded in the Colab notebooks, with batch size one and 640×640 inputs. No TensorRT engine, ONNX Runtime backend, explicit FP16 mode, INT8 quantization, pruning, or batched-throughput optimization was used; the reported values are therefore native PyTorch full-precision timings rather than optimized deployment measurements. Total time is the sum of the recorded per-image preprocessing, inference, and post-processing times, and FPS is its reciprocal. Because all 15 variants were evaluated under the same runtime conditions, the values support relative accuracy–efficiency comparison on this platform but should not be transferred directly to TensorRT, quantized, or different-hardware deployments.

3.7.2. Saliency Localization

Let S_i denote the normalized saliency map for the test image i . Using the common threshold $\tau = 0.5$, the binary saliency mask is

$$P_i(u, v) = \mathbf{1}[S_i(u, v) \geq \tau]. \quad (19)$$

For an image containing R_i annotated FFBs, each bounding box is rasterized into a binary mask G_{ir} . The image-level ground-truth mask is their union,

$$G_i = \bigvee_{r=1}^{R_i} G_{ir}. \quad (20)$$

CAM-box IoU is then

$$\text{IoU}_i^{\text{CAM}} = \frac{|P_i \cap G_i|}{|P_i \cup G_i|}. \quad (21)$$

The reported CAM-box IoU is the mean image-level IoU across the test set. Because all annotated boxes in an image are combined, this metric measures spatial overlap with the overall annotated FFB region rather than instance-specific attribution.

Pointing Game evaluates whether the most salient spatial location falls within any annotated FFB. Let

$$(u_i^*, v_i^*) = \arg \max_{u,v} S_i(u, v), \quad (22)$$

and define

$$PG_i = \mathbf{1}[(u_i^*, v_i^*) \in \cup_{r=1}^{R_i} b_{ir}]. \quad (23)$$

The dataset-level Pointing Game score is

$$PG = \frac{100}{N} \sum_{i=1}^N PG_i. \quad (24)$$

Unlike CAM-box IoU, Pointing Game is independent of τ because it depends only on the location of maximum saliency.

3.7.3. Image-Level Confidence Retention

Confidence retention was evaluated using the complete YOLO detector rather than the truncated backbone. For the original image,

$$p_i = \begin{cases} \max_j \hat{p}_{ij}, & J_i > 0, \\ 0, & J_i = 0, \end{cases} \quad (25)$$

where J_i is the number of detections. The keep-bright masked image is

$$\tilde{x}_i = x_i \odot P_i, \quad (26)$$

with P_i broadcast across the three RGB channels. The corresponding masked-image confidence is

$$\tilde{p}_i = \begin{cases} \max_j \tilde{p}_{ij}, & \tilde{J}_i > 0, \\ 0, & \tilde{J}_i = 0, \end{cases} \quad (27)$$

Mean Absolute Confidence Drop (MACD) is defined as

$$\text{MACD} = \frac{100}{N} \sum_{i=1}^N \max(0, p_i - \tilde{p}_i). \quad (28)$$

MACD is a study-specific descriptive measure of the absolute reduction in maximum image-level confidence after masking. Unlike conventional normalized Average Drop, it is not divided by the original confidence. Lower MACD therefore indicates greater confidence retention.

Increase in Confidence (IC) is

$$\text{IC} = \frac{100}{N} \sum_{i=1}^N \mathbf{1}[\tilde{p}_i > p_i]. \quad (29)$$

Because detections before and after masking are not matched by object, box, or class, MACD and IC are interpreted strictly as image-level confidence-retention diagnostics, not as object-specific or class-specific faithfulness measures.

All CAM methods used the same fixed threshold $\tau = 0.5$, with no model- or explainer-specific tuning. This value is the implementation default and was fixed before any model-explainer comparison: the identical threshold constant was already in place during the target-layer search of Section 3.6, which preceded the CAM evaluation, and it was not adjusted afterward for any configuration. It is therefore prespecified for the comparison reported here, although it was not registered in advance of the study as a whole. Because CAM-box IoU nevertheless depends on this threshold, every threshold-dependent conclusion is cross-checked against Pointing Game, which is threshold-free by construction, and threshold-dependent metrics are interpreted comparatively across configurations rather than as absolute measures of explanation validity.

3.8. Class-Stratified Analysis

The primary EigenCAM-bScore-CAM comparison was stratified post hoc using the per-image CAM records already produced in the original Colab evaluation. Each image was assigned to the first annotated class recorded by the evaluation code. Overripe, ripe, and underripe images were also pooled as a difficult subset, while abnormal, empty-bunch, and unripe images formed a visually distinctive subset. This aggregation introduces no new model, checkpoint, image, or CAM run; it describes the existing 416-image results at a finer class resolution.

3.9. Statistical Analysis

Uncertainty in CAM localization was quantified using percentile bootstrap 95% confidence intervals based on 10,000 resamples of the 416 test images with a fixed random seed of 0. For the primary localization comparison, paired Wilcoxon signed-rank tests evaluated image-level CAM-box IoU differences between EigenCAM and bScore-CAM within each detector. Holm adjustment across the three detector-specific tests controlled the family-wise error rate; all tests were two-sided. Because model-scale and target-layer selection reused the test partition, these analyses are interpreted as exploratory estimates of effect magnitude and consistency rather than confirmatory estimates of out-of-sample explanation generalization. The class-stratified results are reported descriptively.

4. RESULTS

The results are organized to test the study's central proposition in two stages. First, the 15 YOLO variants are compared to determine whether conventional detection accuracy and computational efficiency provide a clear basis for model selection. Second, repeated runs and class-level performance are examined to establish whether the three representative medium models are sufficiently close in predictive performance to justify a separate analysis of explanation behavior.

4.1. Detection Accuracy and Efficiency

All 15 detector variants achieved mAP50 values above 0.993, while mAP50-95 ranged from 0.8860 to 0.9048 (Table 5). Within each YOLO family, the medium-scale model achieved the highest mAP50-95: 0.9048 for YOLOv8m, 0.8998 for YOLO11m, and 0.9043 for YOLO26m. Increasing model scale beyond medium did not yield consistent accuracy gains. Instead, the large and extra-large variants generally incurred higher model size and latency while remaining within the same narrow mAP50-95 range. The benchmark therefore indicates a compressed accuracy regime rather than a clear accuracy-driven ranking.

Table 5. Detection accuracy and end-to-end inference performance of the 15 YOLO variants on the NVidia A100.

Model	Size (MB)	mAP50–95	Total (ms)	FPS
YOLOv8n	5.97	0.8901	1.61	620.25
YOLOv8s	21.49	0.8997	2.60	384.13
YOLOv8m	49.63	0.9048	5.05	197.97
YOLOv8l	83.60	0.8970	7.79	128.40
YOLOv8x	130.40	0.9027	12.12	82.48
YOLO11n	5.23	0.8921	1.79	559.51
YOLO11s	18.30	0.8984	3.10	322.28
YOLO11m	38.65	0.8998	5.20	192.44
YOLO11l	48.84	0.8982	6.24	160.35
YOLO11x	109.10	0.8965	11.45	87.36
YOLO26n	5.15	0.8860	1.29	773.49
YOLO26s	19.38	0.8998	2.73	365.70
YOLO26m	42.00	0.9043	4.76	210.19
YOLO26l	50.54	0.8969	5.77	173.39
YOLO26x	112.86	0.8995	10.93	91.49

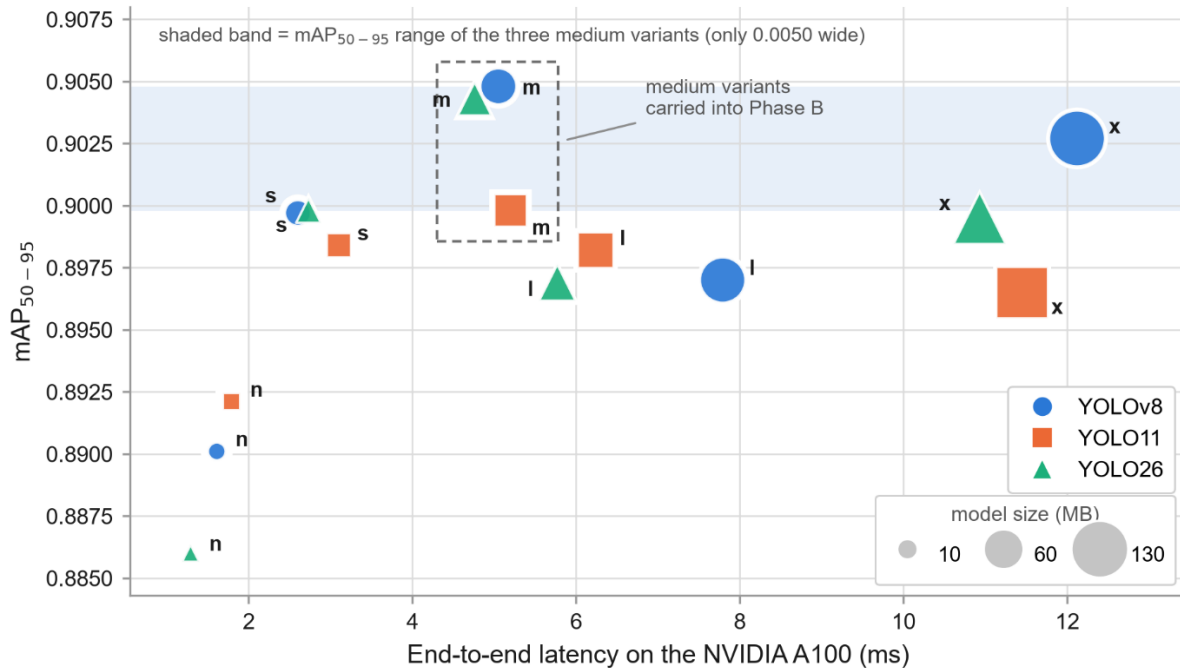


Figure 2. Accuracy–efficiency trade-off of the 15 YOLO variants on the NVidia A100. Marker area represents serialized model size; *n*, *s*, *m*, *l*, and *x* denote model scale. The three medium variants occupy a narrow high-accuracy region, while latency and model size more clearly differentiate their deployment profiles.

Figure 2 further illustrates this compressed accuracy–efficiency region. The three medium models achieved single-run mAP50–95 values of 0.9048, 0.8998, and 0.9043 for YOLOv8m, YOLO11m, and YOLO26m, respectively, corresponding to a maximum spread of only 0.0050. However, their deployment characteristics were more clearly separated. YOLO26m required 4.76 ms per image, compared with 5.05 ms for YOLOv8m and 5.20 ms for YOLO11m, corresponding to 210.19, 197.97, and 192.44 FPS, respectively. Thus, YOLO26m offered the lowest measured latency among the medium variants, although the absolute difference remained modest and is specific to the reported hardware and software environment.

Two observations follow from these results. First, near-ceiling mAP50 should not be interpreted as evidence that the detection task is saturated. The lower mAP50–95 values, centered near 0.90, indicate that stricter localization criteria continue to expose performance differences that are largely hidden at IoU = 0.50. Second, increasing detector size does not produce monotonic accuracy improvements. The medium variants provide the strongest overall accuracy–efficiency balance, whereas larger variants incur substantial computational cost without consistent gains in localization performance.

Accordingly, the benchmark does not identify a dominant detector solely on predictive accuracy. Instead, it establishes a setting in which several models achieve closely matched accuracy while differing in computational cost. This compressed performance regime provides the basis for examining whether explanation behavior remains similarly close.

4.2. Repeated Runs and Class-Level Behavior

The single-run ranking of the three medium detectors was examined across three random seeds to determine whether their small accuracy differences were stable or comparable to ordinary training variation.

Table 6. Repeated-training mAP50–95 of the three medium YOLO models.

Seed	YOLOv8m	YOLO11m	YOLO26m
0	0.905	0.900	0.904
42	0.904	0.900	0.901
123	0.901	0.897	0.899
Mean ± SD	0.903 ± 0.002	0.899 ± 0.002	0.901 ± 0.003

As shown in Table 6, across the three runs, mean mAP50–95 was 0.903±0.002 for YOLOv8m, 0.899±0.002 for YOLO11m, and 0.901±0.003 for YOLO26m. The maximum separation between the model means was only 0.004, which is of the same order as the observed run-to-run standard deviations. The small differences in mean performance therefore provide limited evidence for a practically meaningful ranking among the three medium detectors. Importantly, these results support practical closeness, not statistical equivalence: three seeds are sufficient to show that a single favorable run should not determine model selection, but they are insufficient for a formal equivalence claim or robust between-model significance analysis.

Table 7. Class-level mAP50–95 of the three representative medium detectors on the test partition.

Class	YOLOv8m	YOLO11m	YOLO26m
Abnormal	0.854	0.852	0.851
Empty bunch	0.910	0.907	0.902
Overripe	0.893	0.885	0.876
Ripe	0.915	0.908	0.919
Underripe	0.914	0.907	0.909
Unripe	0.931	0.929	0.935
All classes	0.9048	0.8998	0.9043

Note: Values are seed-0 checkpoints evaluated on the unmodified test partition of 416 images and 1,493 annotated instances. The all-classes row reports the aggregate mAP50–95 of Table 5.

Class-level performance further clarifies where detection remains difficult. As Table 7 shows, abnormal bunches consistently produced the lowest mAP50–95, at 0.854, 0.852, and 0.851 for YOLOv8m, YOLO11m, and YOLO26m, whereas unripe bunches achieved 0.931,

0.929, and 0.935. The most operationally consequential ambiguity occurred among underripe, ripe, and overripe FFBs, where visual characteristics evolve gradually rather than at sharply defined boundaries [6]. Overripe was the weakest of the three stages for every detector, at 0.893, 0.885, and 0.876. Aggregate mAP therefore conceals meaningful class asymmetry: visually distinctive classes are detected more reliably, whereas the biologically adjacent ripeness stages remain comparatively difficult. The class ordering is nearly identical across the three detectors, further indicating that they have converged on similar predictive behavior and motivating the class-stratified explanation analysis in Section 4.5.

These findings change the role of predictive accuracy in the subsequent analysis. Detection performance alone does not establish that YOLOv8m, YOLO11m, and YOLO26m are interchangeable, but the observed differences are too small to provide a compelling basis for selection on this dataset. Once predictive accuracy becomes compressed, *computational efficiency and explanation behavior become complementary evaluation dimensions*, motivating the backbone-level CAM analysis that follows.

4.3. Compatible Target-Layer Selection

Because CAM behavior depends on the representation being interrogated, target-layer selection was examined before comparing explainers. Table 8 summarizes the exploratory layer search using the combined localization score defined in Eq. (18).

Table 8. Exploratory target-layer search.

Metric	YOLOv8m	YOLO11m	YOLO26m
Overall best layer	18 (C2f)	19 (C3k2)	19 (C3k2)
Overall combined score	53.40%	54.78%	51.27%
Compatible backbone layer	8 (C2f)	8 (C3k2)	8 (C3k2)
Backbone combined score	51.23%	39.08%	46.70%

Note. Combined score is the mean of CAM-box IoU and Pointing Game, both expressed on the same scale. The formal cross-model comparison uses the common compatible backbone layer 8.

The highest observed localization scores occurred in the neck: layer 18 for YOLOv8m and layer 19 for YOLO11m and YOLO26m. However, these representations depend on multi-input list-based concatenation and could not be exposed through the common sequential CAM wrapper. The formal comparison therefore used layer 8, the strongest compatible backbone layer shared across all three detectors.

This constraint affected the explanatory target differently across models. For YOLOv8m, the combined score decreased only slightly from 53.40% at the overall-best layer to 51.23% at layer 8. The reduction was more pronounced for YOLO11m, from 54.78% to 39.08%, and more moderate for YOLO26m, from 51.27% to 46.70%. Thus, implementation compatibility is not methodologically neutral: the strongest common evaluable layer does not necessarily coincide with the representation showing the strongest observed localization.

Because the choice of layer 8 governs every subsequent comparison, its sensitivity was examined across the neighboring layers rather than asserted from a single value. Table 9 reports the full layer 6 to layer 10 window for all three detectors using the same combined localization score.

Layer 8 attains the highest combined score in all three detectors within this window, exceeding the next-best candidate by 26.10 percentage points in YOLOv8m, 4.96 points in YOLO11m, and 12.82 points in YOLO26m. The selection is therefore not a knife-edge choice among neighboring representations, although the margin is far smaller for YOLO11m than for

the other two detectors, and a study adopting a different compatible endpoint in that model would obtain a correspondingly different explanatory target.

Table 9. Target-layer sensitivity across the compatible window, layers 6 to 10. Each cell reports Pointing Game (%) / CAM-box IoU (%) / combined score (%).

Layer	Block type (v8m / 11m / 26m)	YOLOv8m	YOLO11m	YOLO26m
6	C2f / C3k2 / C3k2	36.54 / 5.02 / 20.78	60.10 / 8.02 / 34.06	58.89 / 8.87 / 33.88
7	Conv / Conv / Conv	35.10 / 15.16 / 25.13	54.87 / 13.37 / 34.12	7.69 / 11.14 / 9.42
8	C2f / C3k2 / C3k2	85.82 / 16.64 / 51.23	63.94 / 14.22 / 39.08	79.33 / 14.06 / 46.70
9	SPPF / SPPF / SPPF	0.96 / 6.26 / 3.61	43.79 / 12.57 / 28.18	20.91 / 13.72 / 17.32
10	Upsample / C2PSA / C2PSA	1.20 / 6.47 / 3.83	38.46 / 15.18 / 26.82	1.92 / 3.12 / 2.52

Note: Scores are computed from the unweighted spatial mean of each layer's activation tensor over all 416 test images, so they characterize the representation rather than any particular CAM method. Layer 10 lies outside the backbone in YOLOv8m, where the SPPF stage at layer 9 terminates the sequential path.

The window also shows that localization degrades sharply and unevenly once the sequential backbone terminates. Pointing Game at the post-backbone stages falls to 0.96% and 1.20% in YOLOv8m and to 20.91% and 1.92% in YOLO26m, while YOLO11m retains 43.79% and 38.46% at the same positions. Layer selection is thus a model-specific decision even within a narrow window, which reinforces the study's central point that the explanatory configuration, rather than the explainer alone, is the unit that must be reported.

Layer 8 should therefore be interpreted as a common compatible backbone endpoint, not as an optimal CAM layer. Because both layer selection and final CAM evaluation reused the same 416-image test set, these results remain exploratory and may reflect selection optimism; the sensitivity window of Table 9 shows only that the choice is stable among its neighbors, not that it is independent of the evaluation data.

4.4. Model–Explainer Explanation Behavior

Table 10 presents the central empirical result. Predictive accuracy among the three medium detectors was tightly compressed, whereas explanation behavior varied substantially across model–explainer configurations. No explainer dominated across all three backbones.

For YOLOv8m, EigenCAM produced the highest Pointing Game score (71.39%) and lowest MACD (32.68%), while CAM-box IoU differed only marginally among the three methods (14.85–15.70%). The EigenCAM–bScore-CAM IoU contrast was not significant after Holm adjustment ($p=0.650$). EigenCAM therefore showed the strongest observed peak-localization and confidence-retention profile for YOLOv8m, but did not statistically establish overall superiority.

The ranking changed markedly for YOLO11m. EigenCAM yielded 7.00% CAM-box IoU, 35.10% Pointing Game, and 81.30% MACD, whereas bScore-CAM achieved 22.83%, 75.72%, and 41.89%, respectively. bGrad-CAM++ produced similarly strong localization, with 21.38% IoU and 67.07% in Pointing Game. The EigenCAM–bScore-CAM IoU difference was 15.82 percentage points and remained significant after Holm adjustment ($p<0.001$). This rank reversal is the primary empirical finding: an explainer that performed strongly on one detector did not retain the same relative behavior on another detector with similar predictive accuracy.

Table 10. Backbone-CAM evaluation on all 416 test images at layer 8.

Model	Method	IoU (%)	PG (%)	MACD (%) ↓	IC (%)
YOLOv8m	EigenCAM	15.70	71.39	32.68	0.24
YOLOv8m	bGrad-CAM++	15.27	35.10	40.14	0.72
YOLOv8m	bScore-CAM	14.85	66.35	48.62	0.00
YOLO11m	EigenCAM	7.00	35.10	81.30	0.00
YOLO11m	bGrad-CAM++	21.38	67.07	43.66	0.00
YOLO11m	bScore-CAM	22.83	75.72	41.89	0.00
YOLO26m	EigenCAM	10.68	19.71	66.51	0.24
YOLO26m	bGrad-CAM++	11.16	7.69	74.65	0.00
YOLO26m	bScore-CAM	10.50	9.86	78.03	0.00

Note: Higher IoU and PG are better; lower MACD is better. MACD is the study-specific, non-normalized reduction in maximum image-level detector confidence. bGrad-CAM++ and bScore-CAM are the backbone-adapted formulations defined in Section 3.4. Shading, if retained, should indicate the strongest observed value within each architecture and should not imply statistical superiority.

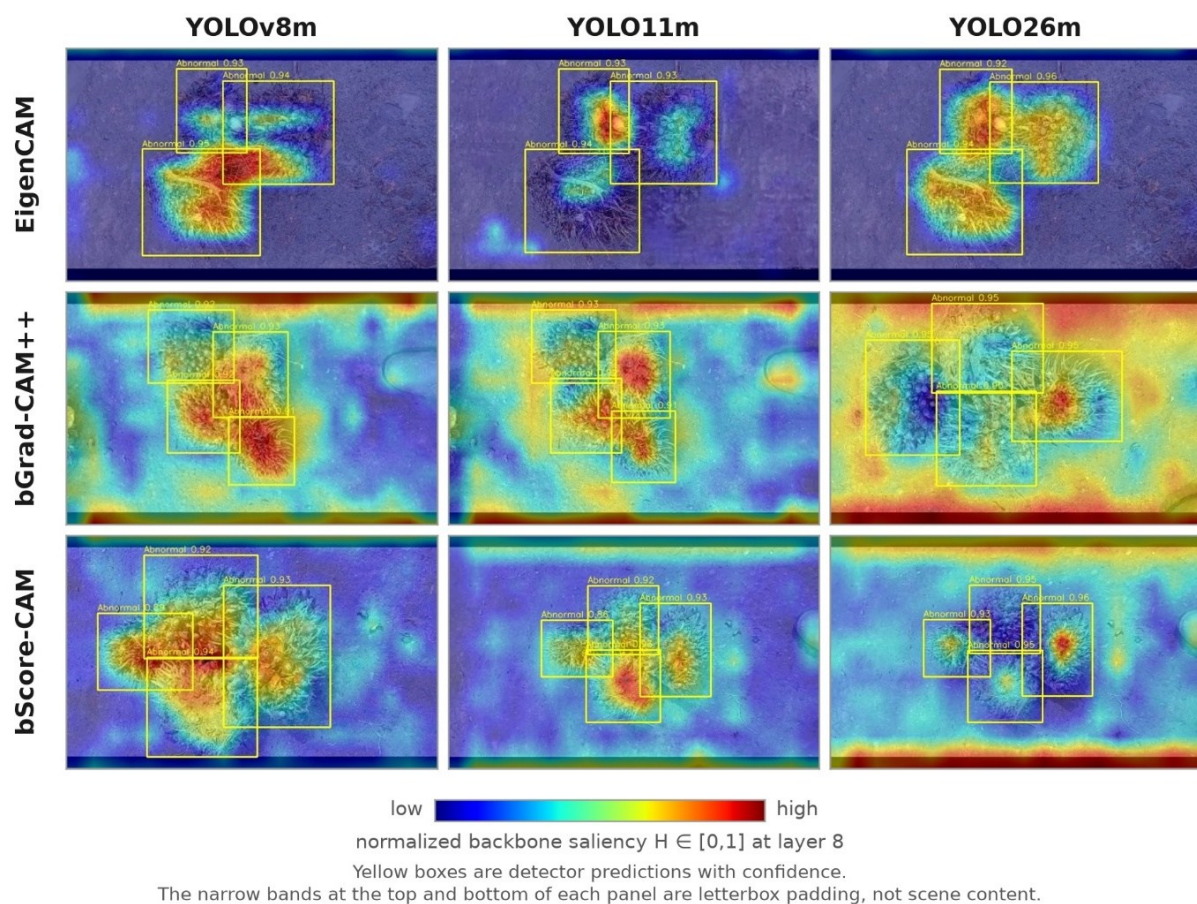


Figure 3. Representative layer-8 backbone heatmaps for an abnormal-bunch scene. EigenCAM is applied in its original class-agnostic form, whereas bGrad-CAM++ and bScore-CAM use the backbone-derived targets defined in Section 3.4. The examples illustrate the quantitative trends but do not establish statistical superiority. All overlays represent backbone feature saliency rather than attribution to a specific final box–class prediction.

For YOLO26m, all evaluated layer-8 configurations showed comparatively weak localization. Pointing Game ranged from 7.69% to 19.71%, while MACD ranged from 66.51% to 78.03%. EigenCAM and bScore-CAM produced nearly identical IoU values (10.68% and

10.50%; Holm-adjusted $p=0.650$), with overlapping bootstrap 95% confidence intervals of 9.60–11.78% and 10.01–11.00%, respectively. This result applies only to the selected backbone target and layer and should not be generalized to YOLO26 as a whole.

Across all nine configurations, Pointing Game spanned 68.03 percentage points, and MACD spanned 48.62 points, whereas repeated-run mean mAP50–95 differed by only 0.0040 across the three medium detectors. Thus, models that are nearly indistinguishable by predictive accuracy need not be interchangeable when explanation behavior is included as an evaluation criterion. Figure 3 shows representative heatmaps illustrating these contrasting saliency patterns.

The absolute CAM-box IoU values remained modest, with a maximum of 22.83%. This is consistent with the study design: the maps originate from a deep, lower-resolution backbone representation, are not conditioned on individual detections, and are evaluated against the union of annotated boxes in multi-object images. They are therefore best interpreted as relative measures of backbone spatial selectivity rather than precise pixel-level attribution. Section 5.4 assesses whether this level of overlap is practically sufficient and why direct comparison with COCO-style detection benchmarks is not meaningful. A class-stratified analysis next examines whether the aggregate pattern is concentrated in visually distinctive classes or also occurs in the biologically adjacent underripe, ripe, and overripe stages.

Increase in Confidence was only 0–0.72% across all configurations. Because detections were not matched before and after masking, IC and MACD remain coarse image-level confidence-retention diagnostics rather than object-specific faithfulness measures.

4.5. Class-Stratified Explanation Behavior

Aggregate metrics can conceal class-specific behavior, and the operationally consequential decisions in FFB grading occur among the biologically adjacent underripe, ripe, and overripe stages. Table 11 therefore stratifies the primary contrast, EigenCAM against bScore-CAM, by the dominant annotated class of each test image, and pools the three adjacent stages into a difficult subset and the visually distinctive classes into a comparison subset.

Table 11. Class-stratified localization for the primary EigenCAM–bScore-CAM contrast at layer 8. Each cell reports CAM-box IoU (%) / Pointing Game (%).

Class (n)	YOLOv8m EigenCAM	YOLOv8m bScore- CAM	YOLO11m EigenCAM	YOLO11m bScore- CAM	YOLO26m EigenCAM	YOLO26m bScore- CAM
Abnormal (56)	12.71 / 62.50	19.91 / 92.86	5.39 / 35.71	22.95 / 80.36	11.43 / 12.50	8.44 / 10.71
Empty bunch (27)	12.21 / 70.37	14.60 / 66.67	4.45 / 25.93	25.30 / 70.37	9.21 / 22.22	8.86 / 11.11
Overripe (72)	14.51 / 63.89	15.64 / 52.78	4.80 / 38.89	24.05 / 93.06	11.88 / 29.17	12.99 / 8.33
Ripe (103)	14.46 / 72.82	16.04 / 66.02	5.03 / 24.27	22.11 / 84.47	12.36 / 19.42	11.05 / 16.50
Underripe (77)	15.11 / 70.13	14.27 / 80.52	5.46 / 31.17	19.27 / 63.64	12.18 / 16.88	10.23 / 2.60
Unripe (81)	22.12 / 83.95	9.78 / 46.91	14.89 / 51.85	25.12 / 59.26	6.04 / 18.52	9.80 / 8.64
Difficult subset (252)	14.67 / 69.44	15.38 / 66.67	5.10 / 30.56	21.80 / 80.56	12.17 / 21.43	11.35 / 9.92
Distinctive subset (164)	17.27 / 74.39	14.03 / 65.85	9.93 / 42.07	24.41 / 68.29	8.40 / 17.07	9.18 / 9.76

Note: n is the number of test images whose first annotated instance carries the listed class. The difficult subset pools overripe, ripe, and underripe; the distinctive subset pools abnormal, empty bunch, and unripe

The descriptive rank reversal is present within the difficult subset and is not produced by the visually distinctive classes. On YOLO11m, bScore-CAM exceeds EigenCAM by 16.70 percentage points of IoU on these 252 images, at 21.80% versus 5.10%, and by 50.00 points of Pointing Game, at 80.56% versus 30.56%. On YOLOv8m, the corresponding IoU values are close, at 15.38% for bScore-CAM and 14.67% for EigenCAM; on YOLO26m, they are 11.35% and 12.17%, respectively. The class-stratified evidence therefore supports the large YOLO11m contrast but does not justify superiority claims where the methods remain closely matched.

Two class-specific effects nevertheless qualify any single-number summary. First, the aggregate EigenCAM advantage on YOLOv8m is concentrated in the unripe class, where it reaches 22.12% IoU and 83.95% Pointing Game against 9.78% and 46.91% for bScore-CAM; restricted to the difficult subset, bScore-CAM is marginally ahead on IoU. Second, on YOLO26m, every configuration is weak, but the failure is not uniform: bScore-CAM Pointing Game falls to 2.60% on underripe images while reaching 16.50% on ripe images. Explanation quality therefore depends on the class, the model, the explainer, and the configuration; aggregate metrics may be weakest on the classes that matter most for grading.

5. DISCUSSION

The results show that the principal distinction among the evaluated detectors is not simply predictive accuracy, but how much explanation behavior depends on the specific model, explanatory target, selected layer, and CAM procedure. The three medium detectors achieved practically close detection performance, yet their backbone-level saliency profiles differed substantially. This contrast is central to interpreting the study: explanation behavior cannot be inferred from mAP alone and should not be treated as a model-independent property of the explainer.

5.1. Detection Accuracy Has Limited Discriminative Power

The three medium detectors occupy a narrow accuracy range. Their repeated-run mean mAP_{50–95} values differ by no more than 0.004, with standard deviations of 0.002–0.003. These results do not establish statistical equivalence, but they provide little basis for claiming meaningful superiority from the observed ordering alone. For example, selecting YOLOv8m over YOLO26m because its mean mAP_{50–95} is higher by 0.002 would assign disproportionate importance to a difference that is comparable to ordinary run-to-run variation.

This compressed accuracy regime also limits the scientific value of ranking newer YOLO generations exclusively by mAP. Once predictive performance converges, other criteria matter more. YOLO26m illustrates this point: it retained nearly the same detection accuracy as YOLOv8m while achieving lower measured end-to-end latency on the NVidia A100. Computational efficiency therefore constitutes a legitimate additional selection criterion when deployment resources matter.

The latency and CAM findings, however, must remain analytically separate. The CAM procedure terminates at the backbone and does not interrogate the NMS-free detection pathway. Consequently, the weaker layer-8 saliency observed for YOLO26m cannot be attributed to its NMS-free design or other downstream head characteristics. Such an explanation would be causal speculation beyond the experimental evidence. The reference-baseline diagnostics of Section 4.7 support the same restriction: the weaker layer-8 behavior of YOLO26m is visible relative to its own unweighted channel mean, which is a property of how the explainers interact with that representation rather than of the detection head.

5.2. Explanation Behavior Is Conditional on the Model–Explainer Configuration

The central finding is not that one CAM method is universally superior, but that explainer performance changes across model–target–layer configurations. EigenCAM captures dominant activation variation, bGrad-CAM++ weights features using gradients of a backbone-derived scalar target, and bScore-CAM uses perturbation responses associated with a pooled backbone feature. Because these mechanisms interrogate different aspects of the learned representation, their relative rankings need not remain stable across detectors even when predictive accuracy is nearly identical.

The clearest evidence is the rank reversal between YOLOv8m and YOLO11m. EigenCAM produced the strongest observed peak-localization and confidence-retention profile on YOLOv8m, whereas bScore-CAM produced substantially stronger localization on YOLO11m. This reversal is more informative than the absolute ranking of any single method because it directly challenges the assumption that a CAM method validated on one detector can be transferred unchanged to another.

The present evidence establishes an empirical association but does not identify its mechanism. EigenCAM, bGrad-CAM++, and bScore-CAM interrogate the representation through principal-component, gradient-weighted, and perturbation-weighted operations, respectively, so differences in activation covariance, gradient stability, or channel-response distributions plausibly explain the reversal. These quantities were not recorded in the Colab runs and cannot be reconstructed reliably from aggregate CAM metrics. The observed differences therefore cannot be attributed to C2f, C3k2, C2PSA, or any other architectural component in isolation. Instrumented analysis of the existing checkpoints is reserved for future work.

The appropriate unit of evaluation is therefore the complete *explainer–target–layer–model configuration*, not the explainer alone. This distinction matters for post-hoc XAI, where visually plausible saliency maps can encourage overgeneralization if their validity is assumed to transfer across models.

5.3. Implications for Model Selection and Deployment

The results suggest three complementary model-selection dimensions: *predictive accuracy*, *computational efficiency*, and *explanation auditability*. In this study, accuracy alone provided little separation among the three medium detectors, while latency and explanation behavior produced materially different comparative profiles.

For the evaluated layer-8 backbone target, YOLOv8m–EigenCAM yielded the highest Pointing Game score and lowest MACD, whereas YOLO11m–bScore-CAM achieved the strongest localization. YOLO26m provided the lowest latency among the medium variants but weaker localization across the evaluated layer-8 CAM configurations. These rankings are descriptive and configuration-specific; they should not be interpreted as universal recommendations for either a detector or an explainer.

This distinction matters for FFB grading. The most difficult classes are the biologically adjacent underripe, ripe, and overripe stages, where visual differences are gradual and grading decisions are consequential, and Section 4.5 shows that the explainer ranking reverses within exactly this subset rather than being averaged into existence across easier classes. It also shows that failure is class-specific: on YOLO26m, bScore-CAM localizes correctly on 16.50% of ripe images but only 2.60% of underripe images. In such settings, a high aggregate mAP may be insufficient if the accompanying explanation consistently emphasizes regions that are difficult to reconcile with annotated FFB locations, and an aggregate explanation score may itself be

insufficient if it conceals a class on which the explanation fails. Explanation auditing therefore provides an additional diagnostic dimension rather than a substitute for predictive accuracy, and it should be reported per class.

More broadly, CAM should not be treated as an interchangeable visualization layer attached after model training. The explanatory target and layer should be specified explicitly, and the selected explainer–target–layer–model configuration should be validated before its saliency output is used for interpretation or decision support.

5.4. Is the Observed Localization Practically Sufficient?

The largest CAM-box IoU observed in this study, 22.83%, is modest in absolute terms and not sufficient on its own to establish deployment-ready explanation quality. A direct numerical comparison with COCO detection IoU or ImageNet localization accuracy would be misleading: those benchmarks evaluate supervised predicted regions or class-conditioned localization, whereas the present metric compares a thresholded, class-agnostic saliency mask from one intermediate tensor with the union of all annotated boxes. Threshold choice, feature-map resolution, and multi-object box union therefore affect the value in ways that do not apply to standard detector mAP [24, 25]. The result is best interpreted comparatively within the present protocol.

Within that protocol, bScore-CAM on YOLO11m combined the highest observed CAM-box IoU (22.83%) with a Pointing Game score of 75.72%, indicating that its peak saliency often fell on an annotated bunch even though the thresholded map did not closely reproduce box extent. This supports its use as a backbone-auditing signal, not as a precise segmentation or as evidence that the explanation is sufficient for autonomous deployment.

Pointing Game and CAM-box IoU should therefore be read together: the former tests whether peak saliency reaches an annotated FFB and the latter tests coarse spatial extent at the fixed threshold. Neither metric validates a specific class decision, and this study did not establish an acceptance threshold for operational deployment.

Instance-level IoU was not added because the evaluated saliency map is class-agnostic and is not conditioned on an individual detection. In images containing multiple bunches, assigning one shared saliency map to a particular instance would require an arbitrary matching rule and would not measure the stated backbone target. A future box- and class-conditioned detector explanation, such as D-RISE [18], would make instance-level overlap well defined.

5.5. Limitations and Validity Threats

Four limitations bound the interpretation of these findings. First, model-scale and target-layer selection reused the test partition. The layer 6–10 sensitivity analysis shows that layer 8 is the strongest compatible candidate within the recorded window, but it does not remove selection dependence. The study therefore remains a controlled exploratory within-dataset analysis rather than a confirmatory estimate of explanation generalization. Second, the explanatory scope is restricted to pre-neck backbone saliency and does not attribute individual box–class decisions; CAM analysis was also limited to the primary seed-0 checkpoints.

Third, localization uses the union of annotated boxes rather than instance-level assignment, together with a single saliency threshold. The threshold was defined in the shared Colab evaluation functions before the model–explainer runs and was not tuned by configuration. However, it was not preregistered, and no threshold sweep was performed. Threshold-dependent IoU is therefore reported alongside the threshold-free Pointing Game. MACD and IC operate on unmatched image-level maximum confidence rather than object- or class-specific

detector responses. Fourth, external validity is limited to one public FFB dataset without independent field validation, repeated CAM evaluation across seeds, explanation-runtime measurement, or operator-based assessment; latency reflects an unoptimized native PyTorch run on the recorded NVIDIA A100 environment rather than TensorRT or a quantized edge deployment.

These constraints limit the generalizability of the absolute CAM values and rankings. Future work should prespecify model and layer selection on validation data, evaluate a genuinely independent test set, sweep the saliency threshold, repeat CAM analysis across seeds, record activation, gradient, and perturbation statistics, and evaluate box- and class-conditioned detector attribution with instance-level overlap on field imagery.

6. CONCLUSION

This study shows that similar detection accuracy does not imply similar explanation behavior in YOLO-based palm fruit ripeness detection. Under a matched protocol, YOLOv8m, YOLO11m, and YOLO26m achieved repeated-run mean mAP50–95 values of 0.903, 0.899, and 0.901, respectively, with a maximum separation of 0.004. In contrast, Pointing Game and MACD varied by 68.03 and 48.62 percentage points across model–explainer configurations, and the explainer ranking changed across detectors. EigenCAM produced the strongest observed peak-localization and confidence-retention profile for YOLOv8m, whereas bScore-CAM achieved the strongest localization for YOLO11m; no method dominated all evaluated backbones. The descriptive reversal remained evident within the difficult underripe, ripe, and overripe subset, while the recorded layer 6–10 search confirmed layer 8 as the strongest common compatible backbone endpoint. These findings support evaluating the complete explainer–target–layer–model configuration rather than treating a CAM method as an interchangeable visualization step. Because the study audits pre-neck representations and reuses the test partition for selection, its results are exploratory, within-dataset evidence rather than attribution of individual box–class decisions or independent estimates of generalization. Future work should address those boundaries with validation-based selection, independent field data, instrumented representation analysis, and box- and class-conditioned explanations.

ACKNOWLEDGEMENT

This work was supported by the International Islamic University Malaysia through the Recognition in Scholarly Excellence 2025 (RISE) Grant (RISE25-007-0007). The first and second authors also acknowledge support from the Ministry of Higher Education, Science, and Technology of Indonesia (Kemendikristek) under grant numbers 1609/LL4/PG/2026 and 069/LIT07/PPM-LIT/2026. Generative AI tools were used solely to assist with language refinement and proofreading. All AI-assisted content was reviewed, verified, and revised by the authors, who retain full responsibility for the accuracy, originality, and integrity of the work.

REFERENCES

- [1] T. S. Gunawan, M. Kartiwi, and A. Jamali, "Drone and deep learning-based instrumentation for palm fruit detection and yield measurement: A Southeast Asian roadmap," *IEEE Instrumentation & Measurement Magazine*, pp. 1-17, 2026, doi: 10.1109/MIM.2026.11224873.
- [2] G. K. A. Parveez et al., "Oil palm economic performance in Malaysia and R&D progress in 2024," *Journal of Oil Palm Research*, vol. 37, no. 2, pp. 187-208, 2025, doi: <https://doi.org/10.21894/jopr.2025.0031>.

-
- [3] T. S. Gunawan, M. Kartiwi, H. Mansor, and N. M. Yusoff, "Palm fruit ripeness detection and classification using various yolov8 models," in *2023 IEEE 9th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, 2023: IEEE, pp. 193-198.
 - [4] S. J. Purba, W. F. Tandion, and E. Irwansyah, "Comparison of the latest version of deep learning YOLO model in automatically detecting the ripeness level of oil palm fruit in PTPN IV, North Sumatra, Indonesia," in *2024 International Conference on Information Technology and Computing (ICITCOM)*, 2024: IEEE, pp. 295-300.
 - [5] J. Y. Goh, Y. Md Yunos, and M. S. Mohamed Ali, "Fresh fruit bunch ripeness classification methods: A review," *Food and Bioprocess Technology*, vol. 18, no. 1, pp. 183-206, 2025.
 - [6] J. W. Lai, H. R. Ramli, L. I. Ismail, and W. Z. Wan Hasan, "Oil palm fresh fruit bunch ripeness detection methods: A systematic review," *Agriculture*, vol. 13, no. 1, p. 156, 2023.
 - [7] J. Y. Goh, M. S. Mohamed Ali, Y. Md Yunos, U. U. Sheikh, and M. S. Khan, "Outdoor RGB and Point Cloud Depth Dataset for Palm Oil Fresh Fruit Bunch Ripeness Classification and Localization," *Scientific Data*, vol. 12, no. 1, p. 687, 2025.
 - [8] P. Chotikawanid, P. Saeleung, Y. Pianroj, S. Jumrat, T. Punvichai, and J. Muangprathub, "Optimizing an Object Detection Algorithm for Detecting Oil Palm Fruit Bunches and Their Ripeness," *Applied Computational Intelligence and Soft Computing*, vol. 2025, no. 1, p. 6263757, 2025.
 - [9] J. Josdaan, V. C. Tamsil, J. Harefa, and K. Jingga, "Revolutionizing palm oil ripeness classification: Utilizing YOLOv8 for ultra-precise ripeness detection," *Procedia Computer Science*, vol. 245, pp. 700-709, 2024.
 - [10] L. T. Ramos and A. D. Sappa, "A comprehensive analysis of YOLO architectures for tomato leaf disease identification," *Scientific Reports*, vol. 15, no. 1, p. 26890, 2025.
 - [11] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680-1716, 2023.
 - [12] Y. Zhao et al., "DETRs beat YOLOs on real-time object detection," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024: IEEE, pp. 16965-16974.
 - [13] M. Tornero-Soria, A.-J. Sánchez-Salmerón, and E. Vendrell Vidal, "Toward a Deeper Understanding of YOLO26: Block-Level Architectural Analysis and Ablation Studies," *Applied Sciences*, vol. 16, no. 13, p. 6758, 2026.
 - [14] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in *2018 IEEE winter conference on applications of computer vision (WACV)*, 2018: IEEE, pp. 839-847.
 - [15] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 618-626.
 - [16] M. Bany Muhammad and M. Yeasin, "Eigen-CAM: Visual explanations for deep convolutional neural networks," *SN Computer Science*, vol. 2, no. 1, p. 47, 2021.
 - [17] H. Wang et al., "Score-CAM: Score-weighted visual explanations for convolutional neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 24-25.
 - [18] V. Petsiuk et al., "Black-box explanation of object detectors via saliency maps," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 11443-11452.
-

- [19] J. Adebayo, J. Gilmer, M. Muelly, I. Goodfellow, M. Hardt, and B. Kim, "Sanity checks for saliency maps," *Advances in Neural Information Processing Systems*, vol. 31, 2018.
- [20] A. Andres, A. Martinez-Seras, I. Laña, and J. Del Ser, "On the black-box explainability of object detection models for safe and trustworthy industrial applications," *Results in Engineering*, vol. 24, p. 103498, 2024.
- [21] D. G. Pai, M. Balachandra, and R. Kamath, "Explainable AI in agriculture: Review of applications, methodologies, and future directions," *Engineering Research Express*, vol. 7, no. 3, p. 032202, 2025.
- [22] Suharjito *et al.*, "Annotated datasets of oil palm fruit bunch piles for ripeness grading using deep learning," *Scientific data*, vol. 10, no. 1, p. 72, 2023.
- [23] P.-T. Jiang, C.-B. Zhang, Q. Hou, M.-M. Cheng, and Y. Wei, "LayerCAM: Exploring hierarchical class activation maps for localization," *IEEE Transactions on Image Processing*, vol. 30, pp. 5875-5888, 2021.
- [24] J. Choe, S. J. Oh, S. Lee, S. Chun, Z. Akata, and H. Shim, "Evaluating weakly supervised object localization methods right," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020: IEEE, pp. 3133-3142.
- [25] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *European Conference on Computer Vision*, 2014: Springer, pp. 740-755.

CFD Analysis of the Effect of a Ball Throttle on the Hydraulic Drive Stability

GULSARA ANNAKULOVA, SARVAR SAIDOV*, BEKZOD ASTANOV

Department of Mechanics of Fluids, Gas and Drive Systems, Institute of Mechanics and Seismic Stability of Structures named after M.T. Urazbaev, Uzbekistan Academy of Sciences, Tashkent, Uzbekistan

**Corresponding author: sarvarsaid1987@gmail.com*

(Received: 4 June 2026; Accepted: 13 August 2026; Published online: 11 September 2026)

ABSTRACT: This paper investigates the impact of a ball throttle on the stability of a vehicle's hydraulic drive system. To do so, a simulation model of the hydraulic system was developed in MATLAB/Simulink, and a 3D computational fluid dynamics (CFD) analysis of the working fluid flow was conducted in SolidWorks Flow Simulation. Additionally, experimental studies were performed on a KI-4815M hydraulic test bench, utilizing SS302 pressure sensors and an Arduino-based data acquisition system. The findings indicate that implementing a ball throttle helps reduce pressure fluctuations and dynamic loads within the hydraulic system. The differences between the CFD results, Simulink simulations, and experimental data do not exceed 5.67% at a load of 700 N. These results confirm the effectiveness of the proposed method for enhancing the stability of hydraulic drives.

ABSTRAK: Kajian ini menyelidiki kesan pendikit bola terhadap kestabilan sistem pemacu hidraulik kenderaan. Bagi mencapai tujuan ini, satu model simulasi sistem hidraulik telah dibangun menggunakan MATLAB/Simulink, manakala analisis dinamik bendalir pengiraan (CFD) 3D bagi aliran bendalir kerja telah dilakukan mengguna pakai simulator SolidWorks Flow. Selain itu, kajian eksperimen telah dilaksanakan pada bangku ujian hidraulik KI-4815M dengan penderia tekanan SS302 serta sistem pemerolehan data berasaskan Arduino. Dapatan kajian menunjukkan bahawa pelaksanaan pendikit bola membantu mengurangkan turun naik tekanan dan beban dinamik dalam sistem hidraulik. Perbezaan antara keputusan CFD, simulasi Simulink, dan data eksperimen tidak melebihi 5.67% pada beban 700 N. Dapatan ini mengesahkan keberkesanan kaedah yang dicadangkan dalam meningkatkan kestabilan pemacu hidraulik.

KEYWORDS: *hydraulic systems, ball throttle, flow control, CFD (computational fluid dynamics), CAD (computer-aided design).*

1. INTRODUCTION

A machine-controlled hydraulic drive is a volumetric hydraulic drive in which a variable-speed pump or variable-speed hydraulic motor regulates the output link motion parameters. In a drive-motor-controlled hydraulic drive, the drive motor's speed controls the output link motion parameters. Furthermore, a throttle-controlled hydraulic drive operates with output link motion parameters governed by a regulating hydraulic device. A machine-throttle-controlled hydraulic drive combines these two approaches, controlling the output link motion parameters using both a regulating hydraulic device and a volumetric hydraulic machine [1].

Determining pressure drops and resistance losses is essential for ensuring the stability and robustness of hydraulic drives. In hydraulic systems, pressure fluctuations or flow irregularities

can cause actuator instability, reduce hydraulic cylinder positioning accuracy, and shorten the service life of pumps and valves.

Classic analytical methods for calculating hydraulic drives do not always account for the complex effects of turbulence, swirl, and throttle geometry, which directly affect the system's stability and stable operation. Computational fluid dynamics (CFD) methods are currently widely used to analyze the internal flow structure, pressure distribution, and hydrodynamic forces in hydraulic valves. In recent years, CFD modeling has become a key tool for improving the reliability and efficiency of servo valves and control valves by more accurately describing complex transient processes. It enables numerical calculation of velocity and pressure drop (ΔP) in complex flows, and prediction of pulsation and swirl zones that can cause instability in hydraulic drives [2-4].

Hydraulic systems in vehicles – pumps, hydraulic motors, hydraulic cylinders, valves, and associated lines – must meet requirements for high efficiency, reliability, and cost-effectiveness. CFD can analyze pressure losses, cavitation, and processes dependent on external conditions. For example, CFD can more accurately calculate flow forces around a moving valve spool [5]. One advantage of CFD is virtual prototyping, which can significantly reduce the time and cost of physical testing (e.g., of a pump or cylinder) [6].

In [7], the internal leakage processes in the main cascade of a servo valve were investigated using 2D CFD modeling. The study found that leakage occurs predominantly through the gap and increases nonlinearly with increasing pressure drop. In [8], a three-dimensional unsteady CFD analysis of the hydrodynamic forces acting on the spool of a control valve was carried out. It was shown that 3D modeling allows more accurate consideration of flow asymmetry and dynamic oscillations than the 2D case. In [9], approaches to reducing hydrodynamic forces on the spool through geometric optimization of the flow path were proposed, improving valve-operation stability. In [10], the authors analyzed flow-pressure characteristics and showed the influence of geometry and flow regimes on valve-operation stability. In [11], a dynamic model of a direct-drive rotary servo valve was developed, confirming its high response speed and control accuracy. A study in [12] analyzed relief-valve stability and showed that hydrodynamic vibrations can lead to self-oscillating modes. In [13], they investigated the effect of a ball valve's opening degree on flow structure and hydraulic losses, and proposed approaches to its design optimization. Furthermore, vehicle hydraulic systems often operate under various operating conditions - changing load, speed, and temperature. CFD models allow analysis of these conditions in a virtual environment and system optimization by changing geometry and operating parameters [14].

Thus, in developing vehicle hydraulic systems, CFD modeling is an important technology that speeds up design, reduces testing, enables early detection of flow problems, and improves system efficiency. Unlike other studies, this article provides a comprehensive comparison of CFD modeling results, MATLAB/Simulink simulations, and experimental data for a hydraulic ball throttle. The schematic diagram of the KI 4815M hydraulic test bench with the specified throttle control methods (Figure 1) includes: pump P with drive motor M, hydraulic cylinder C, hydraulic distributor D (four-way, three-position with manual control), hydraulic throttle DR, relief valve KP, filter F, pressure gauges MN1...MN2, and hydraulic tank T.

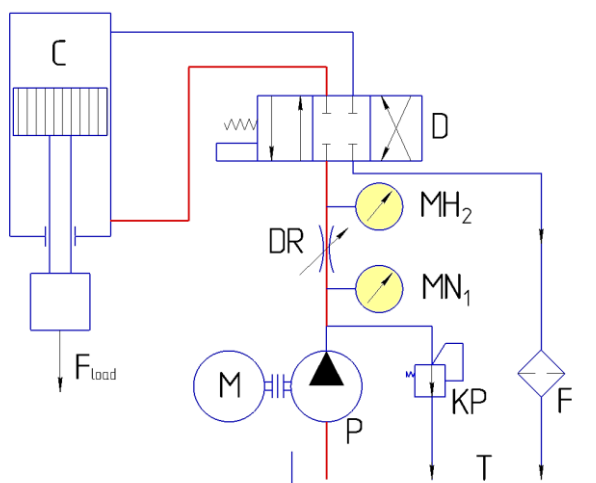


Figure 1. Schematic diagram of the KI 4815M hydraulic test bench with a DR throttle installed at the C hydraulic cylinder inlet.

The ball throttle creates a complex turbulent flow, which directly affects the stability and operation of the hydraulic drive. The presence of a sharp narrowing of the flow, turbulence, and localized pressure drops leads to:

- pressure and flow fluctuations, which can cause unstable operation of the hydraulic cylinders;
- increased load on the pump and valves, which reduces the service life of hydraulic system components;
- difficulties in accurately adjusting the speed and force of actuators. Classic analytical methods do not allow for a detailed consideration of these effects.

CFD modeling provides:

- accurate prediction of pressure drop and velocity distribution, which is important for assessing the stability of the hydraulic drive;
- visualization of turbulence zones and possible turbulent pulsations affecting actuator vibrations;
- analysis of the effect of throttle geometry on flow stability and the possibility of design optimization to reduce vibrations. According to the literature, CFD-based prediction of local resistances achieves 85–95% accuracy and provides a reliable assessment of the throttle's impact on hydraulic system stability [15, 16].

2. FORMULATION OF THE PROBLEM

The system describing the dynamics of a hydraulic drive consists of differential equations for the motion of the moving parts, an equation for the flow of the working fluid in the hydraulic drive elements, and a flow equation. The equations of motion for the moving elements of a hydraulic cylinder are based on the equilibrium of the moving element under applied forces and moments. We assume that the working fluid's properties (temperature, density, viscosity, and undissolved air content) remain constant during the transient process, and that there are no leaks or cavitation. We ignore compressibility in the return pipe.

The fluid flow rate through the throttle is determined according to [17]:

$$Q = c_0 A \sqrt{\frac{2\Delta p}{\rho}}, \quad (1)$$

where c_0 is the flow rate coefficient, Δp - is the pressure drop, ρ is the working fluid density, A is the flow channel area for a ball throttle, and $A(\theta) = A_{max} \sin(\theta)$ for a spool throttle; $A = b y_s$ A_{max} is the section in the fully open state, θ is the valve opening angle. y_s is the spool displacement, b is the spool valve window.

The pressures in the cavities of the hydraulic cylinder are determined by the following expressions:

$$\begin{aligned} \dot{p}_a &= \frac{E'}{V_A(y_{p0}+y_p)A_p} [Q_a - A_p \dot{y}_p], \\ \dot{p}_b &= \frac{E'}{V_B(y_{p0}+y_p)A_s} [Q_b - A_s \dot{y}_p], \end{aligned} \quad (2)$$

where V_A, V_B are the volumes of the pipelines on the rodless A and rod B cavities of the hydraulic cylinder, respectively; E' is the reduced modulus of elasticity of the working fluid, which can be determined according to [18]

$$E' = 0,5 E_{max} l g \left(c_1 \frac{p}{p_{max}} + c_2 \right). \quad (3)$$

We write the flow rates Q_a and Q_b of the spool valve in the following form [19]:

$$\begin{aligned} Q_a &= c_0 \left[sg(y_s) sign(p_p - p_2) \sqrt{|p_p - p_2|} - sg(-y_s) sign(p_2 - p_c) \sqrt{|p_2 - p_c|} \right] \\ Q_b &= c_0 \left[sg(y_s) sign(p_3 - p_c) \sqrt{|p_3 - p_c|} - sg(-y_s) sign(p_p - p_3) \sqrt{|p_p - p_3|} \right], \\ sign(y_s) &= \begin{cases} 1, & \text{for } y_s > 0, \\ 0, & \text{for } y_s = 0, \\ -1, & \text{for } y_s < 0, \end{cases} \end{aligned} \quad (4)$$

The equation of motion of the spool valve is:

$$\ddot{y}_s + 2\zeta \omega_n \dot{y}_s + \omega_n^2 y_s = \omega_n^2 u, \quad (5)$$

where u is the control action from the electromagnet; ω_n - is the natural frequency.

The equations of motion of moving elements of the hydraulic cylinder are based on the equilibrium of the moving element under applied forces and moments. For a hydraulic cylinder, it is written according to [20, 21, 22] in the following form:

$$m_p \frac{d^2 y_p}{dt^2} = p_a A_p - p_b A_s - \left(F_{fr} sign \frac{dy_p}{dt} + k_v \frac{dy_p}{dt} \right) - F_{load}, \quad (6)$$

where m_p is the mass of the moving parts reduced to the piston; y_p is the piston displacement, F_{load} is the external load force, A_p, A_s are the active areas on the rodless and rod ends of the hydraulic cylinder, F_{fr} is the modulus of the dry friction force; k_v is the coefficient of viscous friction of the piston against the cylinder walls.

3. CONSTRUCTION OF A SIMULATION MODEL

Based on the above relations (1)-(6), a simulation model of the hydraulic system with the following parameter values was developed in the Matlab-Simulink package (Table 1). The

Simulink model of the hydraulic system consists of the following subsystems: “Ball throttle,” “Hydraulic distributor,” and “Hydraulic cylinder” (Figure 2). In the “Hydraulic distributor” subsystem, the flow rates Q_a and Q_b are determined based on the input pressure and the tank pressure using Eq. (4).

Table 1. The basic simulation parameters

Parameter	Value
Hydraulic cylinder piston diameter, mm	75
Hydraulic cylinder rod diameter, mm	30
Maximum hydraulic cylinder rod stroke, mm	200
Weight of piston rod, kg	3,5
Spool valve inlet pressure, MPa	16
External load, N	700
Distributor spring natural frequency, s^{-1}	(300÷500)
Damping coefficient	(0,7÷1)
Reduced modulus of elasticity of the working fluid, MPa	1.8
Fluid flow rate coefficient	0.7
Viscous friction coefficient, N	376
Dry friction coefficient, N	247

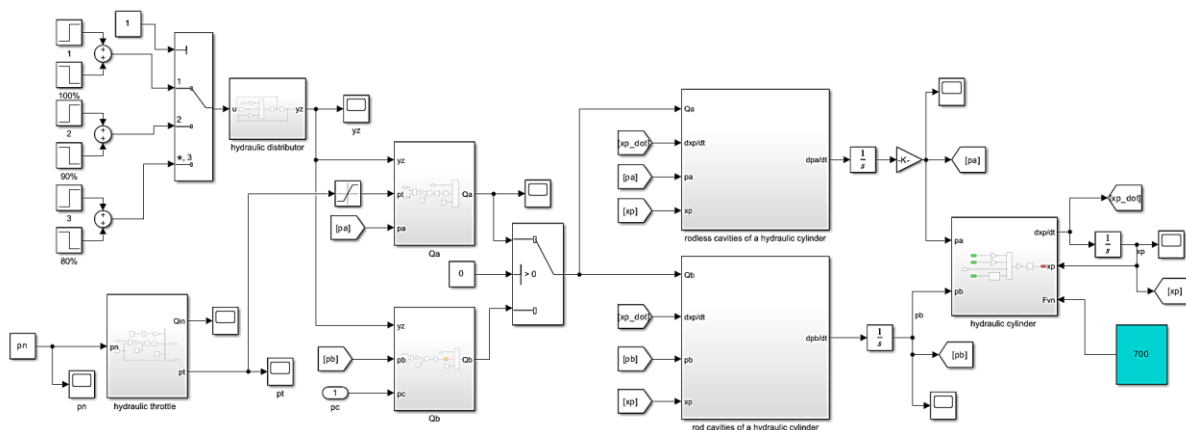


Figure 2. General view of the Simulink model of the hydraulic system

4. CAD MODEL CONSTRUCTION METHOD

The 3D model of the throttle and piping sections of the hydraulic system was created in SolidWorks [23] based on the actual dimensions of the experimental setup (Figure 3).

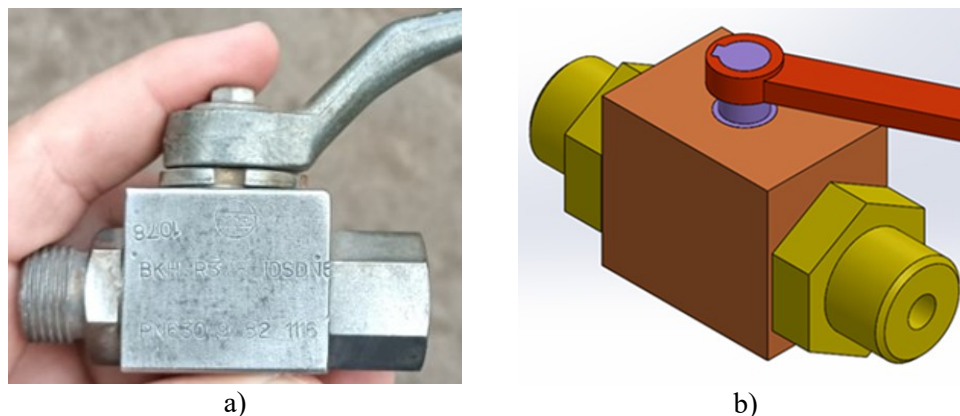


Figure 3. Ball throttle: a) test bench throttle; b) 3D model of the throttle

CFD model setup. The 3D flow domain was discretized in SolidWorks Flow Simulation using an adaptive Cartesian mesh with local refinement near the ball throttle orifice and solid walls to resolve the high velocity gradients in that region. Turbulence was modeled using the standard k- ϵ model with standard wall functions. The following boundary conditions were applied: a specified inlet pressure/velocity corresponding to the pump flow rate at the inlet, atmospheric static pressure at the outlet, and no-slip conditions on all solid walls. The working fluid was treated as incompressible with constant density and viscosity, consistent with the assumptions adopted in the simulation model. The solution was considered converged when the residuals for mass, momentum, and turbulence quantities dropped below $10^{-3} \div 10^{-4}$, together with stabilization of the monitored outlet pressure and mass flow rate.

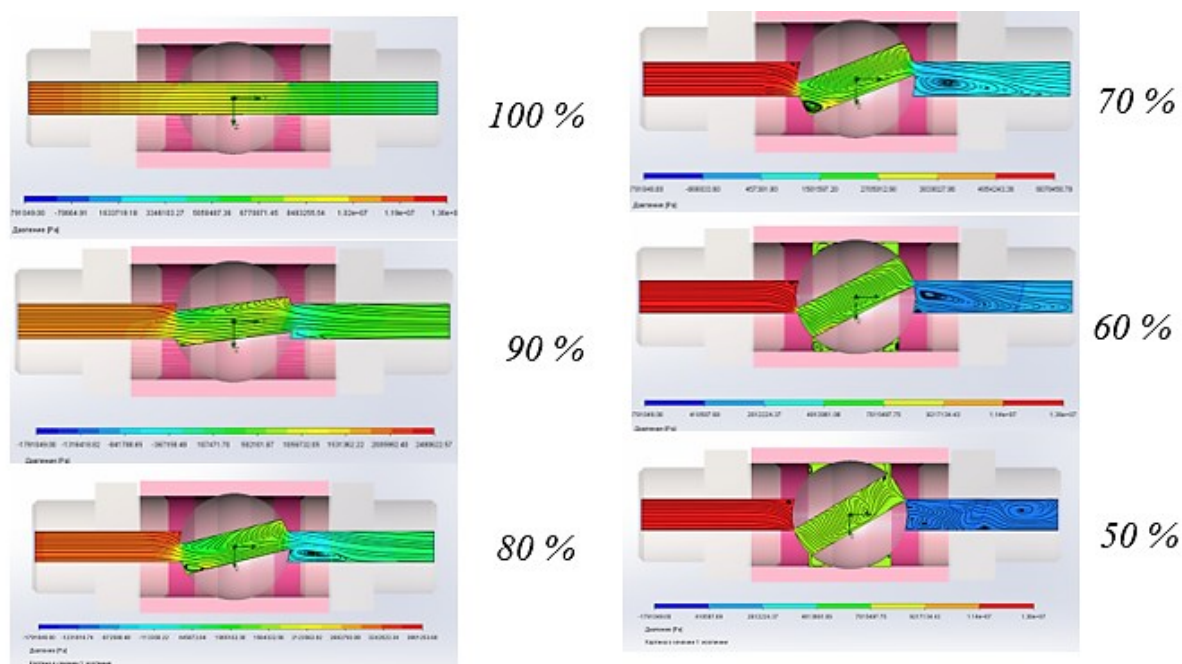


Figure 4. Channel pressure isolines for various ball throttle openings

Figure 4 shows the channel pressure isolines for various ball throttle openings. The CFD results for flow through the ball throttle, considering its geometry, are compared with experimental data measured on the KI 4815M hydraulic test bench to verify the accuracy of the calculations and analyze the throttle's impact on hydraulic system stability.

5. EXPERIMENTAL STUDY

The study focused on the KI 4815M test rig's hydraulic circuit. A pump supplies the working fluid, which passes through a ball throttle and enters a hydraulic cylinder. SS302 diaphragm sensors connected to an Arduino analog-to-digital converter (ADC) measured pressure. The SS302 piezoresistive pressure sensor has an accuracy of $\pm 0.5\%$ F.S. (including non-linearity, hysteresis, and repeatability), a response time of less than 4 ms, and a temperature coefficient of $\pm 0.03\%$ F.S./ $^{\circ}\text{C}$; the pressure range used for the present test bench corresponds to the working pressure of the KI-4815M hydraulic circuit. The Arduino-based data acquisition system employs a 10-bit ADC ($0 \div 1023$ counts), providing a voltage resolution of approximately 4.88 mV at a 5 V reference.

The preparation of the computer-controlled measuring systems of the KI-4815M hydraulic test bench consists of the following stages: selection and placement of sensors (mechanical-to-

electrical converters or vice versa); preliminary processing of analog signals; conversion of analog signals to digital; classification and storage of measured values or their preliminary processing; software for mathematical processing and analysis of the obtained results. Figure 5 shows the functional diagram of the measuring system connections to the computer.

During distributor operation on the KI-4815M test bench a hydraulic pump of appropriate capacity is installed. A control check of the distributor consists of determining the actuation pressure of the relief valve and the spool return mechanism. The tightness of the housing components and oil leakage in the valves are checked.

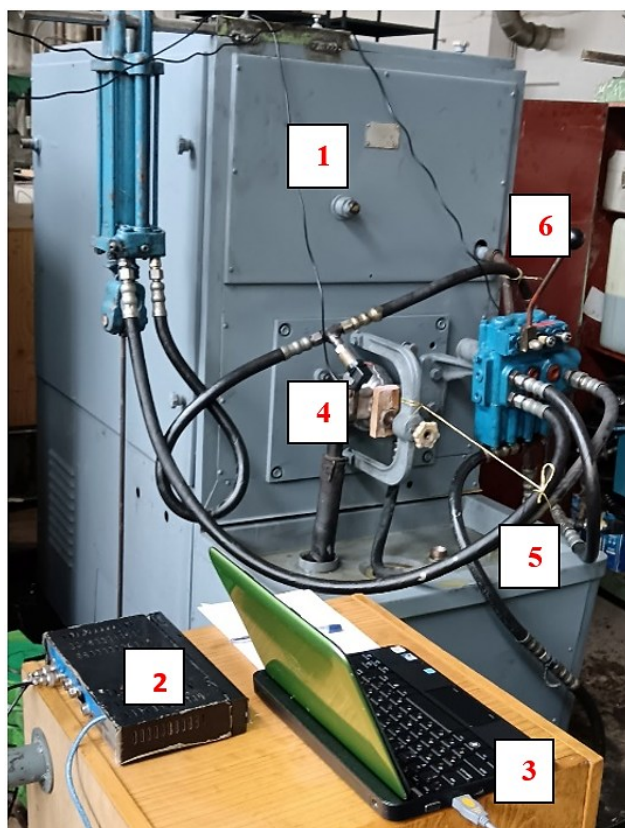


Figure 5. Functional diagram of the connections of the measuring system to the computer of the KI 4815M bench. 1 – KI-4815M bench; 2 – Arduino analog-to-digital converter (ADC); 3 – laptop computer; 4, 5 – pressure sensors; 6 – ball throttle

To quantify the measurement uncertainty, each pressure reading was repeated three times ($n = 3$) at every throttle opening. For each measurement point, the sample mean and sample standard deviation were calculated, and the 95% confidence interval was obtained using the Student's t-distribution ($t_{0,9752} = 4,303$) as $\Delta = \frac{ts}{\sqrt{n}}$. The resulting uncertainty does not exceed ± 0.16 MPa at the throttle inlet and ± 0.02 MPa at the throttle outlet (relative uncertainty below 7,1% at all measurement points, and below 2% for the majority of them), confirming a high level of measurement repeatability. Figures 6 and 7 show the corresponding confidence intervals as error bars on the experimental points.

6. RESULTS ANALYSIS

Figure 4 shows pressure isolines in the channel for different ball-throttle openings. A noticeable narrowing of the jet, followed by an expansion, occurs in the area of the throttle's flow opening. Maximum velocity gradients occur at the minimum-flow section, accompanied

by a significant pressure drop. This pressure drop is the primary source of hydraulic losses in the throttle. The pressure fields illustrate a distinct difference between the inlet and outlet cavities. A zone of high pressure is formed at the inlet, while a sharp decrease in pressure is observed downstream of the throttle. At small openings (50÷60%), the pressure drop is greatest because of increased local resistance and higher flow turbulence. Moreover, the velocity-field analysis reveals local vortex structures behind the edge of the spherical opening. These vortices contribute to additional energy losses and can generate pressure pulsations within the hydraulic system. As the throttle opening increases, vortex formation weakens, and the flow structure becomes more orderly.

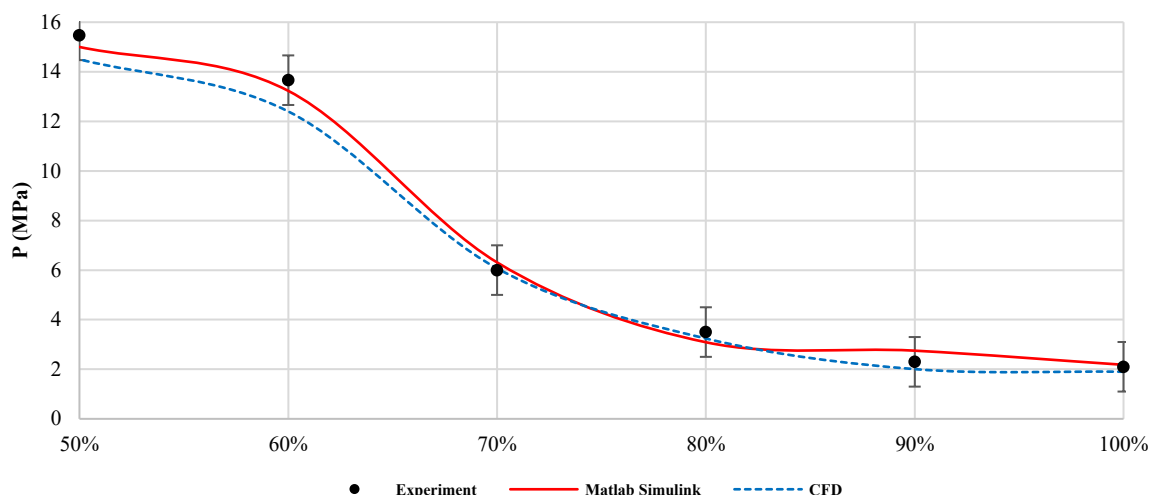


Figure 6. Temporary changes in pressure at the throttle inlet at a load of Flood=700 N

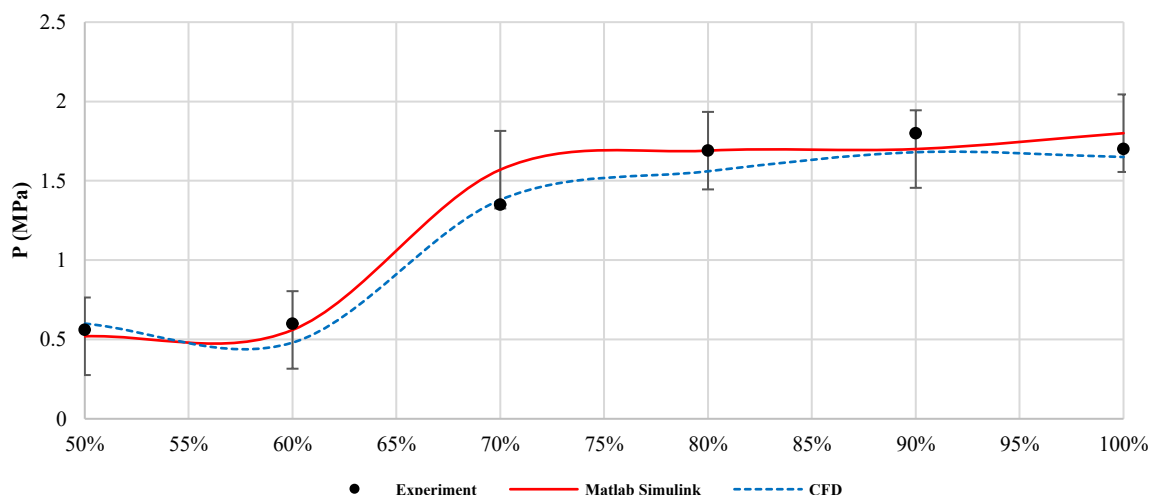


Figure 7. Temporary changes in pressure at the throttle outlet at a load of Flood=700 N

Figures 6 and 7 show the pressure dependence at the ball throttle inlet and outlet, respectively, under a load of Flood=700 N. Comparing the graphs allows us to evaluate the impact of throttling on the system's hydrodynamic processes. Analysis revealed that the pressure at the throttle inlet (Figure 6) is higher and characterized by noticeable pulsations, especially at small openings (50÷60%). This is due to the throttle's increased hydraulic resistance, which causes pressure buildup upstream of the flow area. Meanwhile, the pressure at the throttle outlet (Figure 7) is significantly lower and smoother. The observed pressure drop

confirms a pressure drop across the throttle, which regulates the working-fluid flow. As the throttle opening increases from 70% to 100%, the pressure drop decreases, accompanied by a reduction in pulsation amplitude at both the inlet and outlet of approximately 31÷35% (rising to approximately 13÷15% when the full 50÷100% opening range is considered). This indicates reduced local hydraulic losses and more stable flow.

A comparison of the results from MATLAB/Simulink, CFD, and experiment showed good agreement at the vibronic measurement points. The discrepancy does not exceed 5,67% at a load of $F_{load}=700\text{ N}$, confirming the adequacy of the developed model and the reliability of the CFD calculations. The ball throttle creates a controlled pressure differential between the inlet and outlet, facilitating flow regulation and reducing dynamic oscillations in the hydraulic drive as the throttle opening increases.

7. CONCLUSION

The ball throttle establishes a stable pressure differential and effectively regulates the working-fluid flow rate. As the throttle opening increases, the amplitude of pressure pulsations decreases, stabilizing the flow. Specifically, based on the inlet pressure data, the pulsation amplitude decreases by approximately 13÷15% as the throttle opening increases from 50% to 100%, and by approximately 31÷35% between 70% and 100% opening, consistently across the experimental, MATLAB/Simulink, and CFD results, confirming a substantial improvement in hydraulic drive stability. A comparison of CFD modeling, simulation, and experimental results showed a strong correlation, with the maximum discrepancy not exceeding 5,67% at a load of 700 N. This confirms the accuracy of the developed models and the reliability of the obtained results.

ACKNOWLEDGMENTS

We express our sincere gratitude to the Academy of Sciences of the Republic of Uzbekistan for their ongoing support.

REFERENCES

- [1] Avrunin GA, Gritsai IV, Kirichenko IG, Moroz II, Shcherbak OV (2008) Volumetric hydraulic drive and hydropneumatic automation. KhNADU, Kharkov.
- [2] Anderson JD (2017) Computational Fluid Dynamics: The Basics with Applications. McGraw-Hill.
- [3] Versteeg H, Malalasekera W (2007) An Introduction to Computational Fluid Dynamics. Pearson.
- [4] Domagala M, Fabis-Domagala J (2023) A review of the CFD method in the modeling of flow forces. *Energies*, 16:6059.
- [5] Musa AA, et al. (2021) Advances in CFD-driven design for fluid-particle separation and filtration systems in engineering applications. *Iconic Research and Engineering Journals*, 5(3):347-365.
- [6] Ahmed R, Motin A, Srinivasan C, Zhang Y (2022) CFD analysis of a centrifugal pump controlling a vehicle coolant hydraulic system: A comparison between MRF and transient approaches. *SAE Technical Paper*.
- [7] Tamburrano P, et al. (2019) 2D CFD analysis of servovalve main stage internal leakage. *Proceedings of the ASME/BATH 2019 Symposium on Fluid Power and Motion Control*, Longboat Key, FL, USA.

- [8] Frosina E, Senatore A, Buono D (2015) 3D CFD transient analysis of the forces acting on the spool of a directional valve. *Energy Procedia*, 81:1090-1101.
- [9] Heraković N, Duhovnik J, Šimic M (2015) CFD simulation of flow force reduction in hydraulic valves. *Tehnički Vjesnik*, 22(2):453-463.
- [10] Wu D, Li S, Wu P (2015) CFD simulation of flow-pressure characteristics of a pressure control valve for automotive fuel supply system. *Energy Conversion and Management*, 101:658-665.
- [11] Yu J, Zhuang J, Yu D (2014) Modeling and analysis of a rotary direct drive servovalve. *Chinese Journal of Mechanical Engineering*, 27(5):1064-1074.
- [12] Bazsó C, Hős C (2012) A CFD study on the stability of a hydraulic pressure relief valve. *Conference on Modeling Fluid Flow (CMFF'12)*, Budapest, Hungary.
- [13] Wang JX, Li FX (2020) Flow analysis and structural optimization of a ball valve focusing on the effects of opening degree. *Academic Journal of Manufacturing Engineering*, 18(2):76-86.
- [14] White FM (2016) *Fluid Mechanics*. McGraw-Hill.
- [15] Çengel Y, Cimbala J (2020) *Fluid Mechanics: Fundamentals and Applications*. McGraw-Hill.
- [16] Menter F (1994) Two-equation turbulence models for engineering applications. *AIAA Journal*, 32(8):1598-1605.
- [17] Avtushko VP, et al. (2015) Theory and design of hydraulic-pneumatic drives. Part 1. Two-position hydraulic-pneumatic drives with relay control. BNTU, Minsk.
- [18] Šulc B, Jan JA (2002) Non linear modelling and control of hydraulic actuators. *Acta Polytech.*, 42(3):41-47.
- [19] Jelali M, Kroll A (2012) *Hydraulic Servo-Systems: Modelling, Identification and Control*. Springer Science & Business Media.
- [20] Petrov VV (1951) On self-oscillations of two-cascade servo mechanisms with relay control. *Automation and Telemechanics*, 12:1-28.
- [21] Khokhlov VA, et al. (1971) *Electrohydraulic servo systems*. Mashinostroenie, Moscow.
- [22] Metlyuk NF, Avtushko VP (1980) *Dynamics of pneumatic and hydraulic drives of automobiles*. Mashinostroenie, Moscow.
- [23] SolidWorks Corporation (2022) *Flow Simulation Technical Reference*. Dassault Systèmes.

A TRIZ-Based Conceptual Design of an Automatic Footrest Stand to Enhance Lateral Stability of the Flexible Unicycle Prototype During Mounting and Dismounting

NOORHIDAYAH RAMLI, MUHAMMAD AZIM AZIZI*,
MUHAMMAD AMIN AZMAN, NURAINI AB AZIZ

*Department of Mechanical and Manufacturing Engineering, Faculty of Engineering,
Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia*

**Corresponding author: muhdazim@upm.edu.my*

(Received: 21 July 2026; Accepted: 18 August 2026; Published online: 11 September 2026)

ABSTRACT: This paper outlines the conceptual design of a support mechanism developed to improve the lateral stability of electric unicycles (EUCs), specifically the existing Flexi-UC design, during mounting and dismounting without compromising its compact and lightweight features. The study highlights the application of the Theory of Inventive Problem Solving (TRIZ), which enables the mechanism concept to be developed systematically according to the specified design requirements. Several TRIZ tools were applied, namely Functional Analysis, Cause-and-Effect Chain Analysis, Engineering Contradiction, and Physical Contradiction. As a result, Inventive Principle 7, the Nested Doll Principle, was selected and applied in developing the automatic stand design. Finite Element Analysis (FEA) was conducted to verify the automatic stand's structural reliability. The results showed that the structure could safely support a maximum rider mass of 100 kg. The corresponding load produced a maximum von Mises stress of 66.723 MPa and a maximum deformation of 0.363 mm, both of which are below the allowable limits for mild steel. In addition, theoretical stability analysis indicates that the critical angle increases from -6.01° without the stand to 10.9° with the automatic stand deployed, providing a significantly wider stability margin. Overall, the results demonstrate that the TRIZ-based approach is an effective method for generating conceptual designs that enhance the lateral stability of the Flexi-UC during mounting and dismounting.

ABSTRAK: Kajian ini menekankan aplikasi Teori Penyelesaian Masalah Inventif (TRIZ), membolehkan konsep mekanisme dibangunkan secara sistematik dalam memenuhi keperluan reka bentuk yang ditetapkan. Beberapa kaedah TRIZ telah digunakan iaitu Analisis Fungsi, Analisis Rantaian Sebab dan Akibat, Percanggahan Kejuruteraan dan Percanggahan Fizikal. Hasilnya, Prinsip Rekacipta 7 iaitu Nested Doll telah terpilih dan rekabentuk tongkat automatik telah digunakan. Analisis Unsur Terhingga (FEA) telah dijalankan bagi mengesahkan kebolehpercayaan struktur tongkat automatik tersebut, dan hasil analisis menunjukkan bahawa struktur tersebut mampu menampung beban penunggang maksimum seberat 100 kg dengan selamat. Beban tersebut menghasilkan tegasan Von Mises maksimum sebanyak 66.723 MPa dan ubah bentuk maksimum sebanyak 0.363 mm, kedua-duanya berada di bawah had yang dibenarkan bagi bahan keluli ringan. Selain itu, analisis kestabilan secara teori menunjukkan bahawa sudut kecondongan kritikal meningkat daripada -6.01° tanpa sokongan tambahan kepada 10.9° dengan penggunaan tongkat automatik, sekaligus memberikan margin kestabilan yang lebih besar. Secara keseluruhan, dapatan kajian menunjukkan bahawa pendekatan berasaskan TRIZ merupakan kaedah berkesan dalam menghasilkan reka bentuk konseptual yang dapat meningkatkan kestabilan lateral Flexi-UC semasa proses naik dan turun.

KEYWORDS: TRIZ, Lateral Stability, Micromobility, Electric Unicycle, Finite Element Analysis.

1. INTRODUCTION

Conceptual design is a process in which possible design options are defined early in the design phase after identifying product requirements and specifications. This step is crucial as it establishes the foundation of the final product and has a significant bearing on the overall performance and functionality [1]. Traditional conceptual design methods are often based on brainstorming to generate intuitive ideas or on logical methods that apply structured procedures [2]. TRIZ offers a more analytical and systematic method compared to these approaches, enabling the creation of innovative yet realistic solutions [3]. Unlike conventional brainstorming methods, TRIZ uses structured tools, including the contradiction matrix, separation principles, and functional analysis. Building on these strengths, this study uses TRIZ to conceptually design an automatic footrest stand system. This system aims to enhance the lateral stability of electric unicycles (EUCs) during the critical phases of mounting and dismounting.

Recent studies have demonstrated TRIZ's applicability in micromobility design. For instance, in 2023, a systematic review noted the effective use of TRIZ in designing e-scooters and e-bikes to enhance safety while reducing weight and size [4]. Moreover, case studies in 2023 and 2024 have shown how the contradiction matrix and functional analysis can be used to improve the design of foldable scooters and their ergonomics [5]. More advanced methods have also been adopted in the development of lightweight vehicles, where efforts have focused on computer-aided TRIZ techniques and systematic prototyping to achieve a balance between structural integrity and lightness [6]. While TRIZ has proven effective for multi-wheel micromobility systems, its application to self-balancing single-wheel vehicles remains largely unexplored in the literature. This conceptual stage provides the basis for subsequent structural analysis and prototype development, contributing to a more balanced and user-friendly electric unicycle. Among micromobility systems, the EUC is one of the most compact yet most challenging platforms to balance.

An EUC is a small personal electric motorized vehicle with a single wheel, which is driven by an electric motor and regulated by an inbuilt gyroscope that reacts to changes in the rider's body angle and position [6]. When the rider leans forward, the EUC moves forward, and when the rider leans backward, it reverses. EUCs have integrated control mechanisms that include both gyroscopes and accelerometers, which enable them to maintain balance in both the forward and backward directions. The EUC consists of a frame with the battery, motor, electronic controller, and a central wheel.

Despite continuous technological advances, most commercial EUCs continue to experience poor lateral balance, particularly at low speeds or when at rest [6]. Similar to a bicycle, when the device's forward speed is reduced, the stabilizing effect of forward momentum and gyroscopic action diminishes, increasing susceptibility to lateral instability at low speeds. Novice riders may experience difficulty placing their feet on the footrests, and any minor imbalance may cause the EUC to lean over. The rider's weight transfer during dismounting may disrupt the gyroscopic system's response, resulting in unintended motion that may cause imbalance or a fall [7]. This sudden action may reduce novice riders' confidence and perception of safety. Lateral stability is therefore critically important, especially during starting and stopping, when riders are at the greatest risk of losing balance [7]. Previous studies have investigated rider balancing and EUC dynamics during different riding maneuvers [8]. However, stability during mounting and dismounting remains an important issue that requires further consideration.

The development of the Flexible Electric Unicycle, known as Flexi-UC [9] focused on improving lateral stability during riding, cornering, and low-speed operation. As illustrated in

Figure 1, both sides of the main wheel were equipped with small auxiliary wheels to provide lateral support, helping riders stay balanced during riding. This prototype showed that lateral instability could be minimized by adding compact support elements without sacrificing the EUC's portability.



Figure 1. Flexi-UC with 2 auxiliary wheels on both sides

However, the Flexi-UC does not assist the rider during mounting and dismounting, when lateral instability still exists. This gap motivated the present study to develop an additional lateral-support mechanism through a TRIZ-based conceptual design. This systematic approach is intended to generate a mechanism that provides lateral support without compromising the compactness and portability of the Flexi-UC, specifically during mounting and dismounting. The findings are limited to the geometry, configuration, and operating conditions of the Flexi-UC used in this study and should not be generalized to other electric unicycle configurations without further validation.

Table 1. The comparison of the Flexi-UC with 2 existing electric unicycles on the market

Parameter	Original Flexi-UC	InMotion V11Y [10]	Airwheel Q3 [11]
Dimensions (mm)	690 x 500 x 450	500 × 180 × 744	518 × 408 × 200
Wheel/tire configuration	1 main + 2 side wheels	Single 18 × 3 in tire	Two coaxial 14 in wheels
Stabilization mechanism	Electronic self-balancing + Flexible auxiliary wheels	Electronic self-balancing + suspension	Electronic self-balancing + passive lateral support from twin wheels
Lateral support during mounting/dismounting	Side-wheels ground support	Rider dependent	Twin-wheel ground contact
Advantage	Improved lateral stability during riding/cornering	Suspension and ride comfort	Wider contact area provides improved lateral stability
Limitation	Insufficient support during mounting/dismounting	Requires rider lateral balance	Fixed twin-wheel configuration increases width and cannot retract during riding

Table 1 shows that the three designs use different approaches to improve stability and riding performance. The Flexi-UC incorporates side wheels to improve lateral stability during riding, while the commercial models mainly rely on electronic self-balancing and suspension systems. However, lateral stability during mounting and dismounting remains dependent on adequate ground support or rider control. This limitation provides the basis for developing an additional support mechanism for the existing Flexi-UC. Therefore, this study proposes a

TRIZ-based automatic footrest stand to improve the Flexi-UC's lateral stability during mounting and dismounting.

The objective of this paper is to develop and evaluate a conceptual support mechanism to improve the lateral stability of a Flexi-UC, particularly during mounting and dismounting. The study applies a TRIZ-based approach to identify the root causes of instability and to guide the development of the proposed design. In addition, analytical stability analysis and finite element analysis are carried out to assess the effectiveness and structural reliability of the proposed mechanism.

2. METHODOLOGY

The methodology will be divided into four stages to develop and evaluate the proposed design. Each stage will address a specific aspect of the research, where Stage 1 will focus on problem identification and conceptual design using the TRIZ method, Stage 2 will involve theoretical stability analysis using analytical equations, Stage 3 will cover Finite Element Analysis (FEA), focusing on structural analysis and optimization, and Stage 4 will involve the fabrication and integration of the proposed conceptual design into the Flexi-UC.

2.1. Problem Analysis and Conceptual Design Using the TRIZ Method.

TRIZ is a systematic method that helps identify design conflicts and generate suitable solution ideas without restricting how the mechanism should be developed. However, TRIZ does not consider the mechanical feasibility of the proposed conceptual solutions. FEA, by contrast, enables numerical evaluation of structural behavior under applied loading conditions. Thus, TRIZ and FEA are complementary: TRIZ drives conceptual innovation and guides functional and mechanism-level design, while FEA verifies structural integrity, quantifies performance, and supports further refinement of the design parameters. In this work, TRIZ was applied to develop the deployable lateral support concept, which was subsequently evaluated using FEA to assess and optimize the design geometry and ensure that the proposed inventive solution has sufficient strength and stiffness during operation.

The TRIZ problem-solving map shown in Figure 2 is adapted from the approach described in [12]. It outlines a sequence of steps to investigate a technical problem and develop innovative solutions. The diagram indicates that the process begins with defining the original problem, followed by functional analysis and Cause and Effect Chain Analysis (CECA), which help identify the root causes of the harmful function, and then proceeds through systematic analytical and creative steps toward identifying an appropriate solution. Rather than following a fixed linear sequence, TRIZ allows selective use of each step, depending on its relevance to the engineering problem [13]. Some tools may also be unnecessary or redundant in certain projects [13].

The current design of Flexi-UC is compact, with a small support base defined by the two auxiliary wheels, making it laterally unstable during mounting and dismounting. Specifically, the footrest lacks structural support beneath it to counteract the moment generated by the rider's weight, causing the unicycle to topple. This mechanical limitation was evident in the current Flexi-UC prototype, which, although effective in enhancing lateral stability during motion via auxiliary wheels, still requires improvement during mounting and dismounting. This functional weakness defined the critical design problem of the present work.

To address this, the problem was initially identified and analyzed through functional analysis and Cause and Effect Chain Analysis. These analyses revealed the primary weakness in lateral stability and the competing needs of the footrest system. Based on these findings, the

TRIZ tools, namely Engineering Contradiction (EC), Physical Contradiction (PC), and Function Modeling (FM), were applied to identify the most feasible solution. Through this process, a paradox was identified: the stand must be deployed to provide support but retracted during riding to avoid compromising the vehicle's compactness. This contradiction was addressed using the TRIZ separation principles to ensure that both conditions could be achieved without impeding performance.

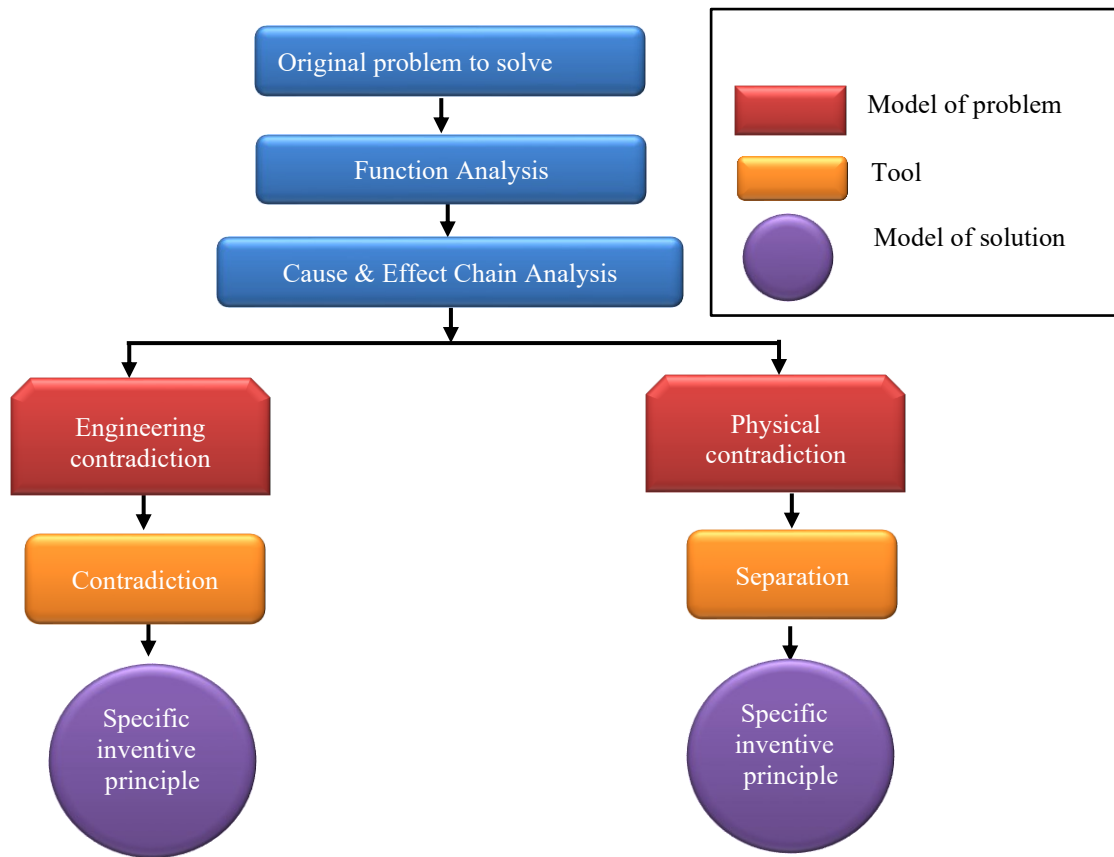


Figure 2. TRIZ problem-solving map adapted from [12].

2.1.1. Function Analysis

The Function Analysis (FA) process has three distinct steps: component analysis, interaction analysis, and functional analysis. Component analysis identifies the system's components and their corresponding functions. Interaction analysis focuses on how the system components affect each other, including both beneficial and harmful interactions. Functional analysis further assesses these interactions to determine whether the functions support or hinder overall system performance [13]. This analysis used the existing Flexi-UC prototype as the reference model and identified the sources of lateral instability.

In the first step, known as the component analysis phase, the Flexi-UC system is defined by identifying its system boundary and the elements that interact with it. As summarized in Table 2, the system consists of the left and right footrests, the unicycle frame, main wheel, the support wheel, the motor, and the electronic board. The supersystem includes the rider, the ground surface, and the surrounding environment, which provide the external loads and contact conditions that affect the system's stability and operation. By considering both the system and supersystem components, the functional analysis ensures that all relevant interactions influencing stability are identified, reducing the likelihood of overlooking critical factors during the design process.

Table 2. The component analysis of the engineering system and its supersystem

Description of engineering system	Components of the system	Components of the supersystem
Engineering system for Flexi-UC	Left footrest Right footrest Unicycle frame Main wheel 2 support wheels Motor Electronic board	Rider's foot, ground, weather, road condition, weight of the rider.

Meanwhile, Table 3 presents the second step: component interaction analysis. The plus sign (+) is used to describe the existence of a functional interaction between two components, whereas the minus sign (-) is used to describe the non-existence of direct interaction.

Table 3. Component interaction analysis

Components	x-axis								
	Rider	Right Footrest	Left Footrest	Unicycle Frame	Electronic Board	Motor	Main wheel	2 support wheels	Ground
Rider	-								
Right footrest	+	-							
Left Footrest	+		-						
Unicycle frame		+	+	-					
Electronic board				+	-				
Motor					+	-			
Main wheel				+		+	-		
2 support wheels				+				-	
Ground							+	+	-

The rider contacts the footrests, which are the primary interfaces for transmitting weight and balance input to the unicycle. The left and right footrests are mechanically connected to the unicycle frame, the primary load-bearing structure that integrates the motor, main wheel, electronic board, and support wheels. The main wheel provides direct contact with the ground, enabling propulsion and transferring vertical loads. The support wheels connect the unicycle frame to the ground and provide additional lateral contact in the Flexi-UC baseline design. Elements of the supersystem, including the rider's foot, ground, road conditions, weather, and the rider's weight, interact with the engineering system and influence its stability behavior. Overall, the component analysis identifies the principal load-transfer paths and control interactions within the system, thereby facilitating the identification of missing or weak interactions that contribute to lateral instability during mounting and dismounting.

The earlier design of Flexi-UC [9] lacks a support mechanism to assist with lateral balance during mounting and dismounting, and this aspect has received limited attention in previous studies because riders often rely on external support, such as a wall or support handle, when mounting and dismounting the unicycle. However, Nicholas [14] stated that mastering the EUC, particularly for beginners, is challenging because it requires precise control and balance to mount and dismount successfully.

The final step in the function analysis process is functional analysis, which evaluates the functions resulting from interactions between system and supersystem components to identify

effective functions and potential problem areas. Figure 3 shows the functional analysis of the Flexi-UC system, illustrating the footrest's interaction with the internal subsystems and the external rider–environment interface. The diagrams use different arrow and line types to categorize the system's main functions, auxiliary functions, interrelations among system elements, and possible negative effects. The right footrest and the rider's foot are positioned lower during lateral imbalance, as shown in the baseline scenario in Figure 3, causing an uneven distribution of the rider's weight. As stability decreases, the load shifts increasingly to one side, reducing the footrest's ability to act as an effective stabilizing surface. This increases the likelihood that the system will tip sideways.

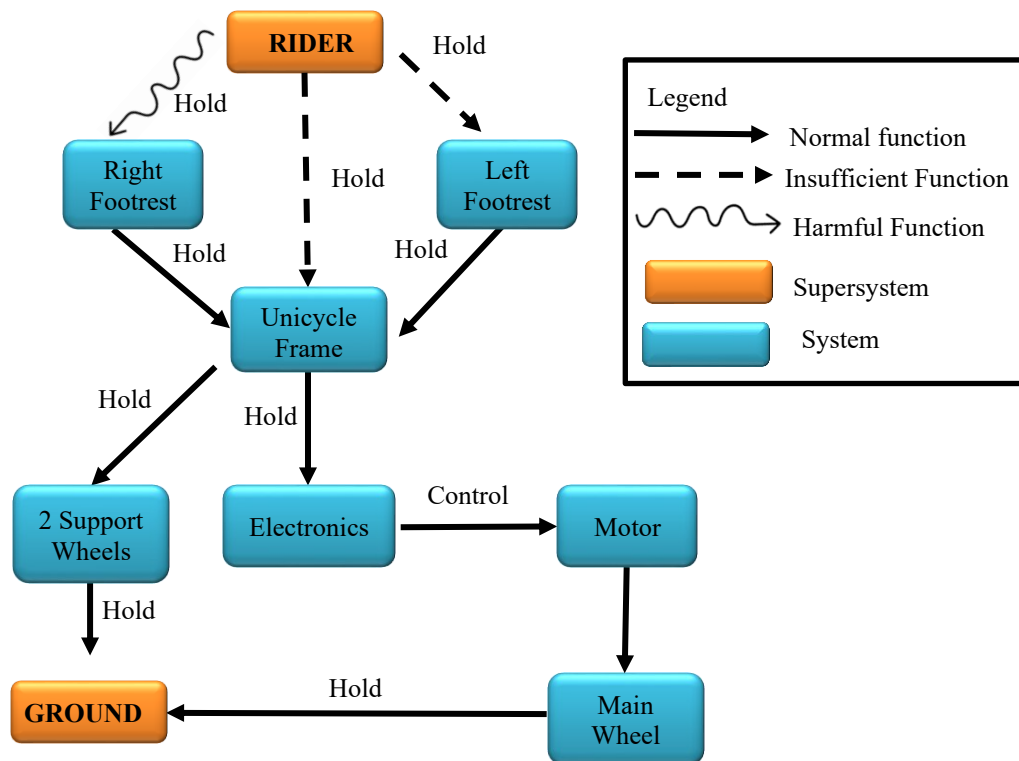


Figure 3. Function analysis of the system before improvement

Focusing on the mounting and dismounting phase, Figure 3 shows the functional model of the Flexi-UC before implementing the stability enhancement mechanism. In this configuration, the rider's balance acts as the primary stabilizing factor, and the left and right footrests serve only as platforms for the rider's feet. In the diagram, the rider's right foot is loosely placed on the right footrest, while the left foot remains on the ground. This instability worsens in the next step, when the rider attempts to quickly place the left foot on the left footrest. At this stage, a novice rider may experience significant lateral instability, which can result in unsuccessful mounting or even a fall. The electronic board, gyroscope, motor, and main wheel are integrated into the unicycle frame, enabling forward and backward movement through the control system; however, these subsystems do not directly contribute to lateral stabilization during mounting.

2.1.2. Cause and Effect Chain Analysis (CECA)

One major TRIZ tool is Cause and Effect Chain Analysis (CECA), which identifies and visualizes the root causes of an identified problem. It starts with defining the system's undesired function and then repeatedly asks “why” to identify further causal connections [15]. This method helps uncover underlying contradictions and root causes. CECA is commonly applied

alongside functional and system analyses, serving as a transition between identifying the problem and developing creative TRIZ solutions.

2.1.3. Engineering contradiction

Engineering Contradiction (EC) was applied after CECA to identify conflicting design requirements, where improving one parameter may worsen another. The identified parameters were mapped into the TRIZ contradiction matrix, which consists of 39 engineering parameters and 40 inventive principles, to generate relevant inventive principles [13]. The generated principles were then evaluated against the design requirements to identify the most suitable principle for the proposed mechanism.

2.1.4. Physical Contradiction

Physical Contradiction (PC) was applied to address opposite requirements of the same design parameter, particularly the need for a large ground-contact area for stability and a small area for compactness. The four TRIZ separation principles, namely Separation in Time, Separation in Space, Separation upon Condition, and Separation between Whole and Parts, were evaluated based on their relevance to the design problem [13]. The most applicable separation principles were then used to determine when the proposed support mechanism should be deployed or retracted.

2.2. Stability analysis based on analytical equations.

The second stage will use analytical equations to evaluate the support stand's impact on stability. Two main indicators were considered: critical angle and stability margin.

2.2.1. Critical angle

The critical angle is the inclination at which a ground-contact system loses static stability. This occurs when the stability margin becomes zero, and the vertical projection of the center of gravity reaches the boundary of the support area. Beyond this angle, the system will tip over, as reported in [16]. The critical tilt angle is calculated based on geometric considerations of static stability, where the vertical projection of the combined center of mass reaches the edge of the support base. In this study, we define the critical angle using the same geometric principle as the rider tilt angle reported by Mungan [17], where the angle is obtained from the inverse tangent of the ratio of the horizontal offset of the center of mass to its vertical height.

$$\theta_{\text{critical}} = \tan^{-1} \left(\frac{\text{Horizontal offset}}{\text{Vertical offset}} \right) \quad (1)$$

For clarity, the horizontal offset term in the present work corresponds to the horizontal coordinate difference reported in prior unicycle tilt studies, in which b is half of the effective width of support, x is the horizontal deflection of the center of gravity (CoG), while h_{eff} representing the vertical offset, and both describe the horizontal distance between the center of mass and the reference support point; the formula is simplified to be:

$$\theta_{\text{critical}} = \tan^{-1} \left(\frac{b - x_{\text{eff}}}{h_{\text{eff}}} \right) \quad (2)$$

The system requires accurate geometric parameters for computation; thus, the unicycle–rider system was modeled in CATIA. The simulation provided CoG coordinates under static conditions, which were then applied in the stability equations to compare the existing Flexi-UC with the proposed design incorporating the stand.

2.2.2. Stability margin

In this study, stability is determined by evaluating how the object's center of gravity (CoG) relates to the support area on the ground. The system is considered stable when the projected CoG lies within the support polygon, defined by the points of contact between the ground and the system. When the projection falls outside the support limit, the system becomes unstable [16]. Thus, the stability margin is the distance between the projected center of gravity (CoG) and the nearest support edge. The critical tipping point occurs when this distance is zero [16].

The stability margin provides a convenient measure of how far the center of gravity (CoG) projection remains within the support polygon during a lateral lean. While Morales et al. [18] formulated the stability margin for a general polygon geometry, the present study adopts an expression adapted to the Flexi-UC configuration, where lateral stability is governed primarily by a single effective support boundary. The lean angle (φ) refers to the instantaneous tilt of the unicycle–rider system away from the vertical, representing the degree of leaning at a given moment. In its general form, the margin is expressed as $SM = b - d_{CoG}$, where b is the horizontal distance from the system's centerline to the edge of the support base, and d_{CoG} is the lateral displacement of the projected CoG. A positive value indicates stable equilibrium, whereas a negative value indicates tipping.

For the present unicycle application, the CoG projection displacement can be approximated as:

$$d_{CoGProjection} = \chi_{eff} + h \cdot \tan(\varphi) \quad (3)$$

where h is the unicycle CoG height and φ represents the effective instantaneous lean angle. Substituting this relation yields the simplified expression:

$$SM = b - (\chi_{eff} + h \cdot \tan \varphi) \quad (4)$$

The stability margin is directly related to the critical angle. If the critical angle $\theta_{critical}$ is substituted into Eq.(3), replacing φ , the SM value equals zero.

$$SM = b - (\chi_{eff} + h \cdot \tan \theta_{critical}) = 0 \quad (5)$$

2.3. Structural Analysis based on Finite Element Analysis (FEA)

The final stage involved structural analysis using FEA and optimization analysis to determine the design parameters of the proposed automatic support mechanism. The 3D CAD model created in CATIA was prepared for analysis by removing minor features that did not affect the main load-carrying paths. This reduced the computational load while maintaining the geometry needed for stress prediction. The stand was hinged at the end of the left footrest, and the rider load during mounting and dismounting was applied vertically at the main load-transfer area of the footrest. The von Mises stress distribution and total deformation were used to evaluate the mechanism's structural performance.

A static structural analysis was conducted in ANSYS Workbench to evaluate the automatic stand's strength. Mild steel was modeled as a linear elastic and isotropic material, with an elastic modulus of 200 GPa, Poisson's ratio of 0.30, yield strength of 250 MPa, and density of 7,850 kg/m³. A safety factor of 3 was selected to provide a conservative design margin for uncertainties in loading, material properties, and operating conditions, resulting in an allowable stress of 83.33 MPa.

A load of 981 N, representing the maximum rider mass of 100 kg, was applied at the main load-transfer area of the footrest. Fixed supports were assigned to the stand's ground-contact

surfaces. The model was meshed using SOLID187 elements. Mesh convergence was checked by comparing maximum equivalent stress across different mesh sizes. The mesh was considered converged when further refinement produced only a small change in stress. Therefore, an element size of 5 mm was selected, resulting in 6,985 nodes and 3,540 elements. No contact definition was required because the stand was modeled as a single continuous solid body.

2.3.1. Boundary condition

The support and load application during the analysis. In this study, two boundary conditions were defined for the FEA, as shown in Figure 4. A fixed support was applied to the surfaces where the stand connects to the cylinders in contact with the ground. These surfaces were assumed to be fully constrained because they represent the parts of the stand that are held in place and do not move when a load is applied.

The second boundary condition was defined by applying a maximum load to the stand's upper surface to represent the maximum rider weight of 100 kg. By applying both the support constraints and the load in this manner, the stand's structural behavior can be evaluated under the specified loading conditions.

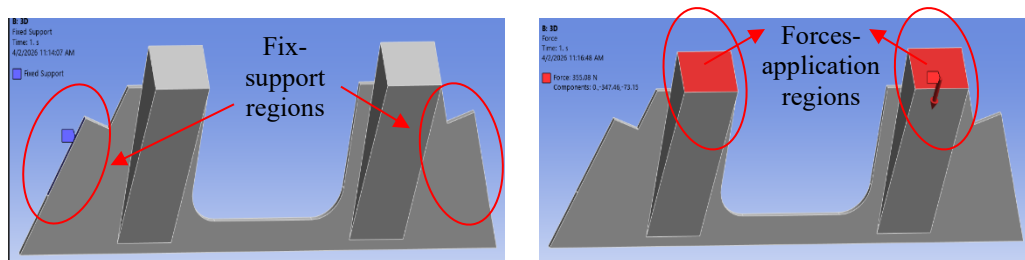


Figure 4. The identification of the boundary conditions: (a) Fixed supports (b) Applied forces

2.3.2. Mesh Sensitivity

Mesh sensitivity analysis examines how mesh size influences the stress results. The model is analyzed with different mesh sizes, producing different numbers of elements and corresponding stress values. These results are plotted as a graph of the number of elements versus the equivalent stress. When the curve begins to flatten, the results are considered to have converged. The mesh size at the start of this convergence region is selected as the optimal mesh size, as it provides stable stress results without unnecessary refinement. This mesh size is then used in the optimization analysis to determine the optimal design parameters.

2.3.3. Optimization

A parameter-based optimization was performed using ANSYS DesignXplorer. The center width and stand-plate thickness were defined as input parameters, while the maximum equivalent stress and stand mass were selected as output parameters. The center width was varied from 40 to 48 mm, and the stand-plate thickness from 2 to 6 mm. The analysis generated 8 design points within these ranges. All other parameters, including the material properties, boundary conditions, and applied load, were kept constant. The optimization aimed to reduce the maximum equivalent stress and stand mass while keeping the stress below the allowable limit of 83.33 MPa.

Among the eight design points, a center width of 43 mm and a plate thickness of 2 mm were selected because they best balanced lower mass with acceptable stress. The final maximum stress was 66.723 MPa, which remained below the allowable limit of 83.33 MPa.

Therefore, these dimensions represent the best feasible design among the evaluated cases rather than a formally determined global optimum.

2.4. Conceptual Design Development and System Integration

Following the TRIZ analysis and concept selection, the selected Nested Doll principle was translated into a foldable automatic stand mechanism. The stand was designed to remain compact beneath the existing Flexi-UC footrest and deploy only when additional lateral support was required. The conceptual design was developed using CATIA V5, where the mechanical components were modeled, assembled, and integrated with the existing Flexi-UC.

Figure 5 shows the main components of the proposed automatic stand, including the footrest stand, support plate, hinge, actuator link, servo motor, and servo motor casing. The components were modeled and assembled using CATIA V5 and integrated beneath the existing Flexi-UC footrest. This arrangement lets the stand provide extra ground support during mounting and dismounting while remaining compact during riding.

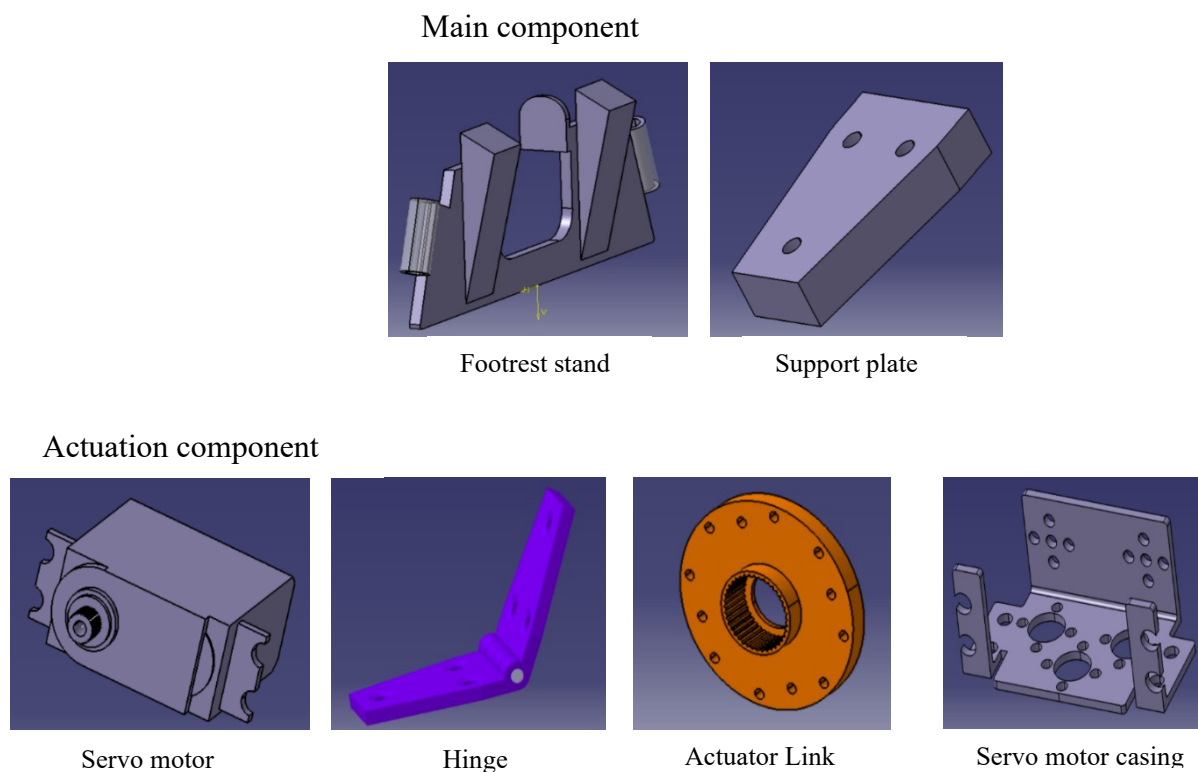


Figure 5. Main mechanical and actuation components of the proposed automatic stand

2.4.1. Automatic Actuation and Control System

The automatic actuation system was developed using an infrared (IR) reflective sensor, an Arduino Uno, and an MG996R servo motor. The IR sensor detected wheel rotation through black-and-white markings on the rotating wheel. The Arduino processed the sensor signal to determine the wheel speed and compared it with the predefined threshold of 0.5 m/s. At speeds of 0.5 m/s or below, the servo deployed the stand to provide lateral support during mounting and dismounting. At speeds above 0.5 m/s, the stand was retracted to prevent ground contact during riding. Separate power supplies were used for the Arduino and the servo motor to minimize voltage drop and electrical noise. Table 4 summarizes the main electrical and electronic components used in the automatic stand system and their respective functions. The

components support the sensing, control, actuation, and power supply required for automatic stand operation.

Table 4. Electrical and electronic components of the automatic stand system

No	Component	Function
1	Arduino Uno	Microcontroller that controls the entire system
2	Breadboard	Provides a platform for building circuits without soldering
3	MG996R Servo Motor	Produces rotational motion to actuate the mechanism
4	IR Sensor	Detects objects or movement using infrared signals
5	Step-Down Converter	Converts higher voltage to stable 5V supply
6	Rechargeable Batteries	Supplies electrical power to the system
7	Jumper Wires	Connect electrical components together

3. RESULTS AND DISCUSSION

This section discusses the findings and results in relation to the study's objectives. The findings are organized into four sections. The first section presents the TRIZ results, where the identified contradictions and inventive principles lead to the development of the proposed mechanism. The second section presents the theoretical stability analysis using analytical equations to evaluate the improvement in lateral stability. The third section presents the FEA results to evaluate the structural strength and reliability of the proposed design under the applied loading conditions. Finally, the fourth section presents the fabricated conceptual design, in which the automatic stand is integrated at the end of the Flexi-UC footrest to provide additional lateral support during mounting and dismounting. Collectively, these findings provide an overall evaluation of the proposed solution and its capability to improve the lateral stability of the Flexi-UC.

3.1. TRIZ-Based Problem Analysis Result

3.1.1. Function Analysis

The functional analysis after improvement complements the FEA results and supports the conclusion that the additional mechanism can withstand the applied load with adequate structural strength.

After developing the conceptual design, the functional analysis was updated, as shown in Figure 6. The automatic stand introduces an additional ground-support interaction, widening the support polygon and improving lateral stability. It also addresses the previously harmful rider-footrest interaction while retaining the original structure with only one additional subsystem. This minimal modification is consistent with the TRIZ concept of increasing system ideality.

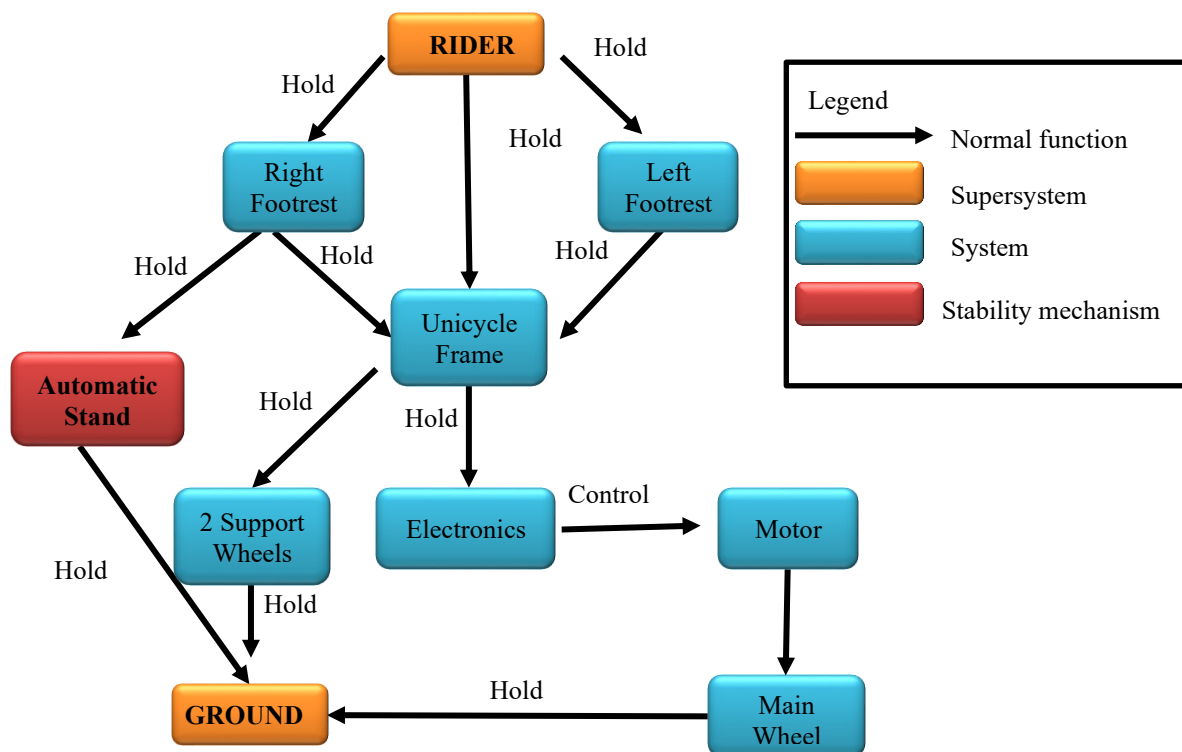


Figure 6. Function Analysis for the research after the improvement

3.1.2. Cause and Effect Chain Analysis (CECA)

The chains used to identify the root causes of Flexi-UC's tendency to topple during mounting and dismounting are shown in Figure 7. The major problem is insufficient contact area between the unicycle and the ground, which reduces available lateral support. Several factors further influence this issue: the limited contact footprint restricts the reaction forces needed for stabilization; the lightweight design with fewer structural components reduces support between the unicycle and the ground; and together, these effects increase the likelihood of toppling when the unicycle is stationary or moving slowly. Collectively, these root causes highlight the need for an auxiliary mechanism that increases ground contact and provides additional support to overcome Flexi-UC instability during mounting, dismounting, and low-speed phases.

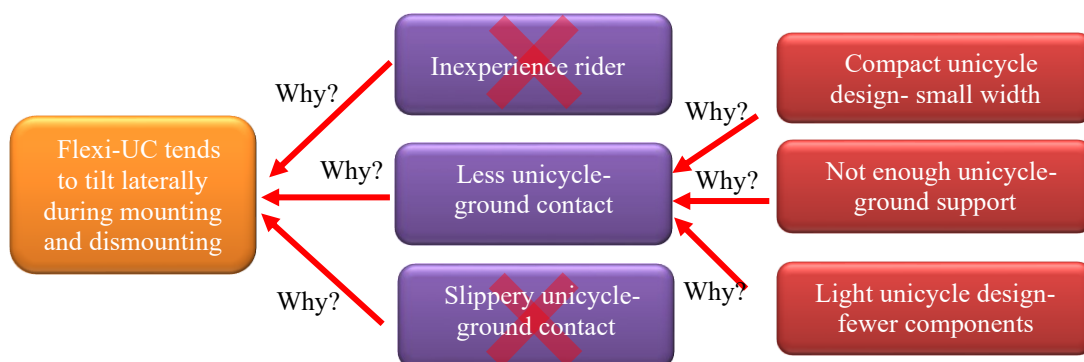


Figure 7. The cause-and-effect chain analysis (CECA)

3.1.3. Engineering Contradiction

Engineering contradiction (EC), also known as a technical contradiction, refers to a situation in which improving one aspect of an engineering system often leads to the degradation of another [19]. In this study, we considered two pairs of engineering contradictions (ECs): the area of the stationary object vs. its length, and the volume of the stationary object vs. its length. Table 5 presents the inventive principles obtained from the contradiction matrix based on the selected contradiction parameters.

Table 5. Extracted from the contradiction matrix [12]

Improving features	Worsening features	Solution from contradiction matrix	
Area of stationary object	Length of stationary object	7 - Nested 9 - Preliminary Anti-Action	26-Copying 39 – Inert atmosphere
Volume of stationary object	Length of stationary object	2 - Taking out 8 - Anti-Weight	14 - Curvature 35 – Parameter changes

Table 6. Relevance analysis of TRIZ inventive principles for Flexi-UC stability enhancement

Principle	Application idea	Relevance	Reasons
26– Copying	Use a simple mock-up or passive stabilizer instead of complex gyroscopic/sensor systems.	×	Doesn't address the core need for additional support during static phases; unsuitable for the project goal.
7 - Nested Doll	Design a mechanism that folds and stores compactly inside the unicycle's body/frame.	✓	Supports the challenge; mechanisms can fold/unfold neatly inside without affecting portability or aesthetics.
9 - Preliminary Anti-Action	Pre-deploy a stabilizer or footrest before the unicycle stops to prevent toppling.	×	This refers to internal counteractions (like stress) that prevent harm, not a structural/mechanical anti-toppling system.
39 - Inert Atmosphere	Use passive damping materials or inert mechanical elements to reduce vibrations/movement.	×	Focuses on chemical or environmental stability, not on a structural or mechanical stability issue.
35- Parameter Changes	Use materials that change stiffness or shape based on temperature or electric current for a deployable footrest.	×	Material transformation doesn't relate to adding a structural extension to enhance stability.
8. Anti-Weight	Use a lightweight material to keep the unicycle from feeling heavy.	×	Stability is a structural issue, not a weight-based one.
14. Curvature	Design the footrest to curve around the unicycle frame.	×	This principle does not directly address that specific challenge.
2. Taking Out	Eliminate the rider's reliance on body balancing by making the support function a new, separate mechanism.	×	The rider's balance control is self-regulated; it's not a system component that can be removed/isolated by design

Meanwhile, Table 6 shows the assessment of inventive principles based on the TRIZ contradiction matrix to enhance the Flexi-UC's lateral stability. The principles were evaluated separately for relevance, practicability, and appropriateness to the project objectives, focusing on increased stability during mounting and dismounting while maintaining a compact installation size. Most of the proposed principles were less suitable because they did not directly address the need for structural stability during mounting and dismounting. Among the

evaluated principles, only Principle 7, Nested Doll, was identified as directly applicable. The principle led to the idea of a foldable and retractable footrest stand that can be stored within the unicycle frame when not in use and deployed when needed, addressing the need for greater stability while overcoming space limitations. The other principles, although novel in other contexts, were less suitable because they primarily involve material modification, environmental effects, redistribution of internal forces, or weight minimization, none of which directly enhances mechanical stability in this case. One principle was also considered impractical because the rider controls balance and cannot regulate it externally through mechanical means. As a result, the Nested Doll principle emerged as the most appropriate conceptual basis for the automatic footrest stand, as it enables a compact yet deployable stabilization system while retaining the unicycle's overall functionality and design integrity.

3.1.4. Physical Contradictions

Physical contradictions (PCs) refer to opposite requirements for a design parameter within the same system [13]. The paradox of the existing Flexi-UC design is that a large ground-contact area is needed during mounting and dismounting to maintain lateral stability, while a small contact area is required during riding to reduce interference.. The improved Flexi-UC addresses this contradiction by applying TRIZ Separation in Time and Separation upon Condition, so the stand deploys only during mounting and dismounting and remains compact during riding. This technical paradox can be resolved using appropriate TRIZ tools, such as the separation principles for physical contradiction.

Table 7. The physical contradiction involved

Element Involved	Conflicting Requirements
Length and area of contact between unicycle and ground	Needs to be large for stable ground contact and small for compactness

The first step in a physical contradiction analysis is to identify the element involved and to explain the conflicting requirements. In this case, as shown in Table 7, the conflict is between the unicycle's size and its ground contact area. A larger contact area improves stability when the unicycle is static, but keeping the unicycle small and compact reduces that area. In other words, the contact area or length must be both large and small at the same time. To resolve these conflicting requirements, TRIZ provides four separation principles: Separation in Time, Separation in Space, Separation upon Condition, and Separation between Whole and Parts [13].

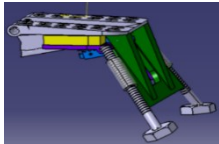
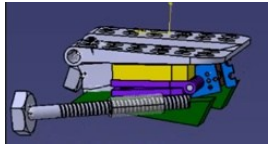
A relevance analysis, shown in Table 8, revealed that the most applicable principles are Separation in Time and Separation upon Condition, meaning the footrest stand can be deployed when needed and retracted when not. This approach increases stability without compromising the unicycle's compactness. However, Separation in Space and Separation between Whole and Parts were less applicable because these principles increase design complexity.

Subsequently, Table 9 presents the TRIZ-based resolution strategies derived from these separations. Two key problems were identified: (1) insufficient ground contact during mounting and dismounting, and (2) a bulky stand design that disrupts dynamic motion. To address these, we applied two separation strategies. In time separation, the stand deploys when stationary and retracts during motion, ensuring both stability and compactness. In condition separation, the stand activates only below a threshold speed, making it responsive to real-time riding conditions. Together, these strategies keep the footrest compact when not in use and deploy it automatically when needed, improving stability without compromising maneuverability.

Table 8. Relevance of TRIZ separation concepts in addressing the physical contradiction

Type of separation	Contradiction Element	Application Idea in the project	Relevant to Project?	Reasons
Separation in Time	Contact ground area must be small for compactness and large for stability	Keep unicycle compact; footrest unfolds for lateral support	✓ Relevant	Unfolds only when needed
Separation in Condition		Compact design: parts extend temporarily for support	✓ Relevant	Depends on speed/rider input
Separation in Space		Not applicable	✗ Irrelevant	Increase design complexity
Separation between whole and parts		Not applicable	✗ Irrelevant	Increase design complexity

Table 9. The physical contradiction and TRIZ-based solution

Types of separation	TRIZ-based solution	
<u>In time</u>		
i. The stand is deployed only when stationary, and retracted during motion.		
<u>In condition</u>		
i. The stand activates only when the speed falls below a threshold of 0.5 m/s.	Deployed during mounting and dismounting phase.	Neatly installed under footrest.

Although the control system prevents commanded deployment above 0.5 m/s, an operational safety response was considered. If the stand fails to deploy, the rider can move slowly towards a level stopping area or nearby fixed support before dismounting. If unintended deployment occurs during riding, the stand remains above the ground and does not immediately interfere with wheel movement. The rider can then reduce the speed, stop on a level surface, and dismount safely. Therefore, these failure conditions do not immediately compromise rider safety.

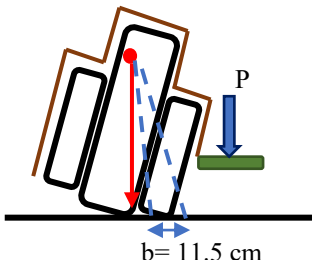
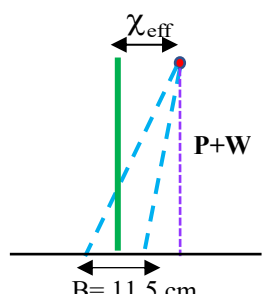
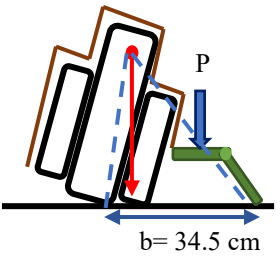
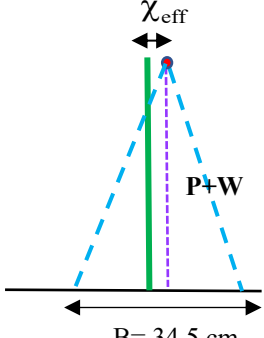
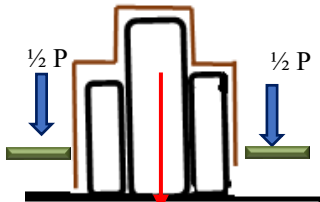
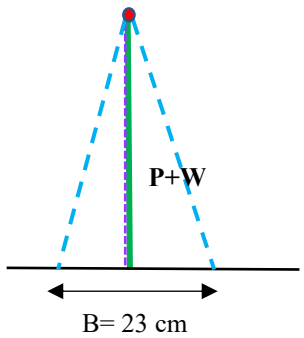
3.2. Stability Analysis Using Analytical Equations

A preliminary static stability analysis was conducted to verify improved lateral stability. This evaluation provides a simple yet quantitative measure of how the proposed foldable footrest stand influences the unicycle's resistance to toppling. A comparison of the support polygon width and center of gravity (CoG) projection under three conditions, namely mounting the Flexi-UC with the stand, mounting without the stand, and riding the Flexi-UC, is presented in Table 8. P denotes the weight of the rider, W is the weight of the Flexi-UC, b is the distance at which the Flexi-UC contacts the ground, and e_1 to be denoted as the distance between the center of the b and P while e_2 is the distance between the center of b and W .

To clarify the diagram in Table 10, the purple line represents the Center of Gravity (CoG) of the rider, P , while the red dashed line represents the CoG of the Flexi-UC, W . The green line indicates the center of the base width, and the blue dashed lines indicate the stability boundaries of the base. During mounting and dismounting with the stand deployed, the base width increases to 34.5 cm, keeping the CoG projection within the stability boundary. This provides greater lateral stability and makes mounting and dismounting easier and safer. Without stand

deployment, the base width is only 11.5 cm, causing the CoG projection to fall outside the stability boundary. This creates an unstable condition and increases the likelihood that the Flexi-UC will tip sideways. During riding, when both feet are in full contact with the footrests, the base width is 23 cm, and the CoG projection remains within the stability boundary, which is sufficient to maintain lateral stability.

Table 10. Comparison of support polygon and CoG projection in three conditions

Condition	Support polygon and CoG projection (P+W)	
Mounting and dismounting Without stand		 The base width is narrow, the projected CoG is already outside the boundary.
Mounting and dismounting with stand		 The base width is wider, and the center of the width shifted to the right
Normal riding		 The system's CoG is aligned with the centre of the support base.

Indicators:

-  Centre of Gravity of Flexi UC and rider
-  Centre of the base width
-  Stability boundary (length of the base)

3.2.1. Lateral CoG displacement.

Lateral CoG displacement is the effective lateral displacement of the center of gravity (CoG) of the unicycle. All computations used a typical rider weight of 70 kg and a unicycle weight of 25 kg, standard values in the stability literature representing an average adult and a mid-weight EUC. The effective lateral CoG displacement and the vertical CoG displacement h_{eff} of the combined rider–unicycle system are approximated using a mass-weighted average of the individual lateral CoG positions, assuming quasi-static conditions and a centrally aligned unicycle CoG. When $g = 9.81 \text{ m/s}^2$, the weight of the rider is $P = 686.7 \text{ N}$, and the weight of the unicycle is $W = 245.25 \text{ N}$. e_1 is the distance between the CoG of the Flexi UC and the center of the width base, while e_2 is the distance between the CoG of the rider and the center of the width base. For the without-stand condition, e_1 is -1.305 cm, while e_2 is 19.05 cm, and for the with-stand condition, e_1 is 12.27 cm, while e_2 is 8.1 cm. Meanwhile h_1 and h_2 are the individual CoG heights for the Flexi-UC and the rider, respectively. When the stand is deployed, the support base is not displaced; it is only spread out. Therefore, the collective offset in the lateral CoG is solely dependent on P , and W will remain identical with or without the stand. Consequently, the formula for lateral CoG displacement. Equation (4) and vertical CoG displacement, h_{eff} ; using the adaptation of the formula introduced by Rahman [20] will give:

$$\chi_{eff} = \frac{W.e_1 + P.e_2}{W + P} \quad (6)$$

$$h_{eff} = \frac{W.h_1 + P.h_2}{W + P} \quad (7)$$

χ_{eff} are calculated separately in both cases: without stand and with stand, where the value of e is different for both cases.

χ_{eff} without stand:

$$\chi_{eff1} = \frac{245.25(-0.01305) + 686.7(0.1905)}{245.25 + 686.7} = 0.1368 \text{ m} \quad (8)$$

χ_{eff} with stand:

$$\chi_{eff2} = \frac{245.25(-0.1227) + 686.7(0.081)}{245.25 + 686.7} = 0.0274 \text{ m} \quad (9)$$

While for h_{eff} is the same for both cases.

$$h_{eff} = \frac{W.h_1 + P.h_2}{W + P} = \frac{245.25(0.2) + 686.7(0.95)}{245.25 + 686.7} = 0.753 \text{ m} \quad (10)$$

The values of the and h_{eff} above are used to calculate the critical angle and SM.

a. Critical angle during mounting and dismounting with stand:

The value from equation (2) will be substituted into equation (5), which will give

$$\theta_{critical} = \tan^{-1} \left(\frac{b - \chi_{eff2}}{h_{eff2}} \right) \Rightarrow \tan^{-1} \left(\frac{0.1725 - 0.0274}{0.7527} \right) = 10.9^\circ \quad (11)$$

When the stand is deployed, the base width of the unicycle increases substantially, and the critical toppling angle increases to 10.9° . This shows that the system can tolerate a larger sideways lean before tipping over, representing a substantial improvement compared with the condition without stand deployment. Although the unicycle reaches its critical angle at 10.9° , the rider can still maintain balance because the critical lean angle for a standing human is about 17.7° [21]. This is comparable to standing on a stationary four-wheel skateboard or climbing a V-shaped ladder, where the toppling angle is around 10° . These situations suggest that a person can maintain balance under similar static conditions. As a result, the stand helps keep the unicycle upright and stable during mounting and dismounting, similar to the situations described above.

b. Critical angle during mounting and dismounting without stand:

$$\theta_{\text{critical}} = \tan^{-1} \left(\frac{b - \chi_{\text{effl}}}{h_{\text{effl}}} \right) \Rightarrow \tan^{-1} \left(\frac{0.0575 - 0.1368}{0.7527} \right) = -6.01^\circ \quad (12)$$

The small critical angle indicates that the human–unicycle system is near the limit of static stability at rest, even in the absence of additional external disturbances, since the effective center of gravity lies very close to the edge of the base. The sign of the critical angle indicates the direction of lean, while its magnitude of 6.01° represents the onset of lateral instability. Under this condition, the unicycle cannot remain steady on its own without external support, as even minor disturbances may cause the center of gravity projection to move beyond the support boundary. This finding indicates that toppling can be induced by very slight lateral movement without a stand or other external support. It highlights the inherent instability first-time riders encounter during mounting and dismounting, when maintaining balance is difficult unless additional structural support is provided.

c. Critical angle during riding:

$$\theta_{\text{critical}} = \tan^{-1} \left(\frac{b - \chi_{\text{eff}}}{h_{\text{eff}}} \right) \Rightarrow \tan^{-1} \left(\frac{0.115}{0.753} \right) = 8.7^\circ \quad (13)$$

The lateral load of the rider is symmetrically distributed between both footrests during riding while the stand remains fully retracted; therefore, there is no distinction between riding with or without the stand, and the effective lateral CoG offset is nearly zero $\chi_{\text{eff}} \approx 0^\circ$. The width of the unicycle with the footrests unfolded defines the support base, while the stand width does not contribute to stability when it is folded. This positive angle suggests that, when riding straight with the CoG centered, the CoG projection remains well within the support base over a relatively large lean range, with the theoretical stability limit reached when the lean exceeds 8.7° . In practice, the achievable lean before loss of control may be lower due to additional dynamic constraints such as main wheel grip, surface irregularities, and rider control. During riding, small steering corrections and the gyroscopic effects of the moving wheel also support balance. Thus, this value provides a reference point for why riding is more stable than mounting and dismounting.

Overall, the critical angle analysis shows a clear difference among the three operating conditions. Without the stand, the system's critical angle during mounting and dismounting is very low at -6.01° , indicating that the system is near the stability limit and can be easily disturbed by minor lateral movements. When the stand is deployed, the support base widens, and the tipping threshold shifts outward, with the critical angle increasing to approximately 10.9° , providing a greater stability margin during weight transfer in mounting and dismounting. During riding, the critical angle is 8.7° , indicating sufficient stability for forward motion, with

the rider and dynamic effects largely maintaining balance. These findings demonstrate that the stand enhances lateral stability in slow or near-static conditions, particularly during mounting and dismounting, without adversely affecting stability during riding.

Table 11. Comparison of critical angle for unicycle with stand and without stand while mounting and dismounting, and in addition during riding

Condition	Critical angle (Theoretical)
Mounting and dismounting with stand	10.9 ⁰
Mounting and dismounting without stand	-6.01 ⁰
During riding	8.7 ⁰

Table 11 shows that the stand plays an important role in improving lateral stability during mounting and dismounting. The stand's enlarged support polygon keeps the CoG projection well within the base, providing greater rider stability. Conversely, without the stand, the support base is very narrow, and the CoG projection is near the edge of the support polygon. This condition makes the unicycle highly sensitive to minor lateral disturbances, which explains why balancing becomes difficult during mounting and dismounting. Once the rider has both feet on the footrests during riding, however, the effective base width increases. In this riding condition, the CoG projection remains within the support polygon, and adequate stability is maintained; thus, the stand becomes less critical.

3.2.2. Stability Margin (SM)

A positive SM indicates that the CoG projection lies within the support base, signifying a stable condition, while a negative SM indicates that the CoG projection has exceeded the boundary, leading to instability. In this way, χ_{eff} from Equation (4) represents the shifted position of the CoG due to mounting, while $h \tan \theta$ accounts for further displacement due to leaning, and SM quantifies the remaining margin before tipping occurs. Selected lean angles are presented as representative low-angle conditions to illustrate the influence of lean angle on the stability margin. We calculated the stability margins by substituting the base width b from Table 10 into Equation (4), and present the results in Table 12.

From Table 12 and Figure 8, the three conditions differ in stability margins. In the upright posture, the system is already unstable with $SM = -0.079$ m at 0° , and when leaning, the system becomes more unstable with $SM = -0.185$ m at 8° , indicating that the projection of the CoG is outside the support base and that the unicycle is susceptible to toppling during mounting or dismounting.

Table 12. Comparison of stability margin (SM) at selected lean angles for Flexi-UC with stand and without stand

Lean Angle φ	SM without stand	SM with stand	SM during riding
0 ⁰	-0.079	0.145	0.088
2 ⁰	-0.106	0.119	0.061
5 ⁰	-0.145	0.079	0.022
8	-0.185	0.039	-0.018

The tipping boundary shifts outward when the stand is deployed, resulting in a large positive margin: $SM = 0.145$ m at 0° and $SM = 0.039$ m at 8° . This indicates that the CoG projection remains within the support polygon over the tested lean range, providing greater stability during mounting and dismounting. The system is therefore unstable during unsupported mounting and dismounting unless additional support is provided. However, after

deploying the automatic stand, the support base widens, shifting the tipping boundary further outward and creating a consistently positive stability margin, thereby providing greater stability during mounting and dismounting. This enhancement shows that the optimized Flexi-UC provides higher static stability than the baseline configuration without the stand.

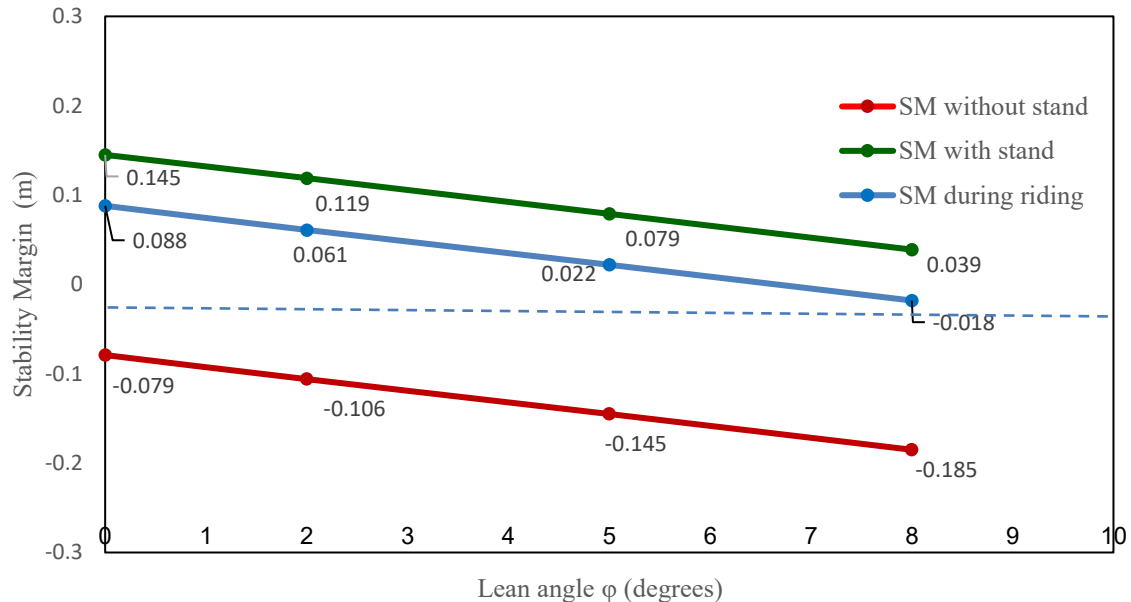


Figure 8. Comparison of stability margin (SM) versus lean angle for three operating phases

During riding, when the stand is folded, the base configuration is positively stable at the lower lean angles tested, where $SM = 0.088$ m at 0° and decreases to 0.061 m at 2° . However, the stability margin decreases as the lean angle increases and becomes negative at higher angles, with $SM = -0.018$ m at 8° . This tendency shows that active rider control and dynamic balance become increasingly important in maintaining balance at higher lean angles or under external disturbances. Overall, the results showed that the stand played a major role in improving static stability during mounting and dismounting.

The derived analytical formula verified the relationship between the stability margin and the critical angle. From Equation (3), substituting the critical angle for the lean angle results in $SM = 0$, corresponding to the zero-crossing points of the stability margin curves. The lean angles at which $SM = 0$ are very similar to the calculated critical angles for all configurations, demonstrating the internal consistency of the static stability model and supporting the validity of the adopted formulation.

3.3. Structural Analysis Using Finite Element Analysis (FEA)

FEA was undertaken to determine the structural behavior of the automatic stand mechanism. Because the stand must bear the applied load during mounting and dismounting, a design load of 100 kg, representing the maximum rider mass, was considered. Static loading was considered sufficient because mounting and dismounting occur under stationary or very low-speed conditions and do not require a complex dynamic simulation. It is necessary to ensure that the added structure does not deform beyond acceptable limits and that stress concentrations remain within allowable limits. The analysis was based on three criteria: total deformation to assess the stiffness and geometric stability of the stand; von Mises stress to verify that the maximum stress remains below the allowable stress of mild steel after applying a safety factor of 3, which is 83.33 MPa; and design optimization to reduce the overall mass of

the component without adversely affecting performance. A safety factor of 3 was used in this design. This value is supported in the literature, where a safety factor of 3 has been used in tubular space-frame chassis designs to ensure structural safety under high loading conditions. Therefore, a safety factor of 3 was considered suitable for the proposed structural component [22].

3.3.1. Mesh sensitivity

Before optimization, a mesh sensitivity analysis was conducted to determine a suitable mesh size for the subsequent optimization analysis. Figure 9 shows a sensitivity graph of maximum stress convergence versus the number of elements. From the graph, the stress values begin to converge at 3,522 elements, corresponding to a mesh size of 5 mm. Therefore, this mesh size was selected for the optimization analysis.

A mesh convergence study was conducted using element sizes of 20, 13, 10, 6, 5, 4, 3, and 3.5 mm. The maximum equivalent stress values obtained for each mesh were 7.021, 10.24, and 10.63 MPa, respectively. Mesh convergence was considered achieved when the difference in maximum equivalent stress between two successive mesh refinements was below 5%. Therefore, the final analysis used an element size of 5 mm.

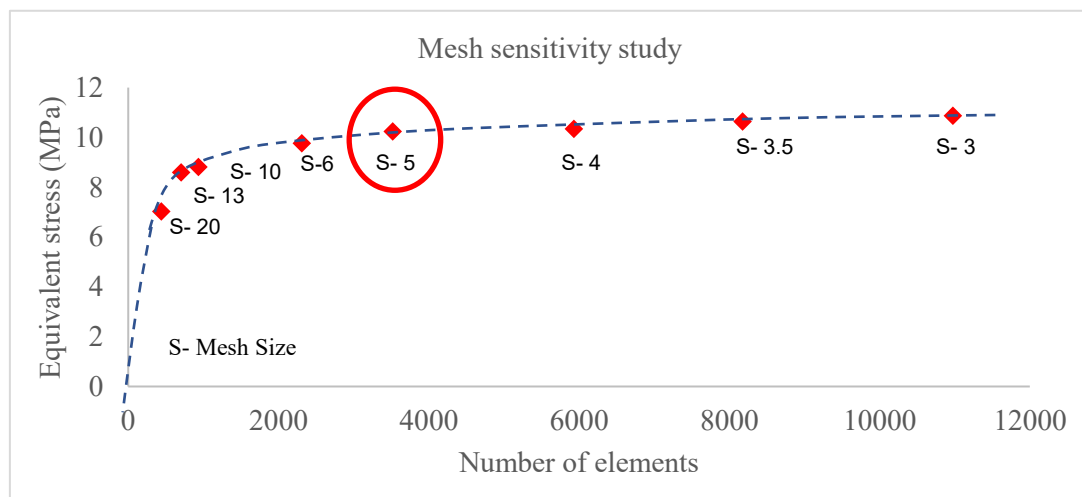


Figure 9. Mesh sensitivity graph of Equivalent Stress vs. Number of elements

3.3.2. Optimization

For this design, two parameters were considered as independent variables: the center width and the thickness of the stand plate, as shown in Figure 10. The dependent variables are the maximum stress in the stand under the maximum applied load and the stand weight. Other parameters were treated as controlled variables.

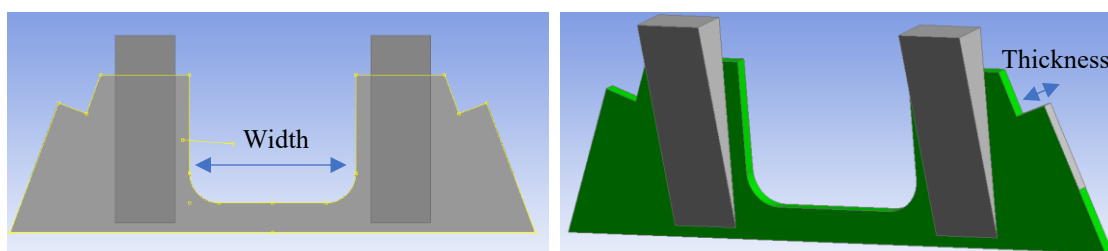


Figure 10. Input parameters used in the ANSYS DesignXplorer optimization

Based on the evaluated design points, the analysis identified a center width of 43 mm and a plate thickness of 2 mm as the most suitable combination within the specified parameter ranges. This configuration produced a stand mass of 0.6 kg and a maximum equivalent stress of 66.723 MPa under a static load equivalent to a 100 kg rider.

Since the maximum stress remained below the allowable stress of 83.33 MPa, the selected geometry was considered structurally safe. The configuration also maintained a compact and lightweight design suitable for integration with the Flexi-UC footrest. The selected dimensions represent the most suitable design identified within the evaluated parameter ranges rather than a global optimum. Overall, the selected geometry can safely support the maximum design load of 100 kg, representing the maximum rider mass, while minimizing unnecessary material use. The final design is structurally adequate, compact, and suitable as a lateral stabilizing mechanism for the improved Flexi-UC system.

3.3.3. Total deformation

The maximum total deformation of the automatic stand was 0.363 mm at the lower part of the stand, as shown in Figure 11. This displacement is very small compared with the component's overall size, indicating that the structure maintains geometric stability under the applied load. The minimal bending also confirms that the stand is sufficiently stiff to remain dimensionally stable during rider mounting and dismounting, with only minimal changes to the support geometry. These results confirm that the stand maintains sufficient structural rigidity and is structurally viable under the intended operating conditions.

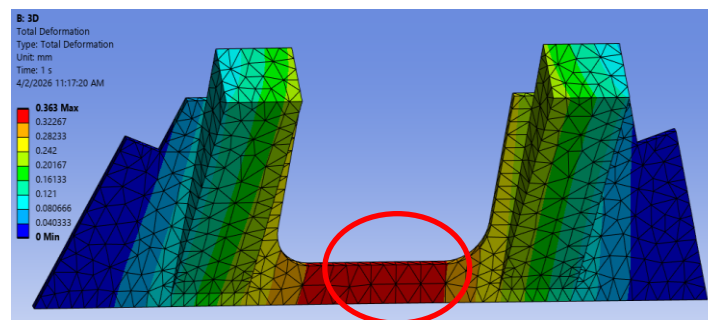


Figure 11. Total deformation contour of the optimized design

3.3.4. Equivalent Stress

Figure 12 shows the von Mises stress distribution of the automatic stand, with a maximum stress of 66.723 MPa, which is approximately 80.1% of the allowable stress of 83.33 MPa. These results confirm that the automatic stand does not exhibit critical stress concentrations exceeding the allowable limit and is structurally suitable for mounting and dismounting under the specified static loading conditions.

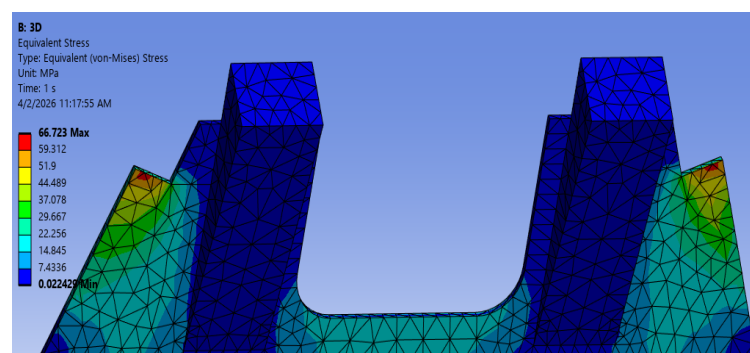


Figure 12. Equivalent Von Mises stress distribution of the optimized component

3.4. Conceptual Design and Prototype Integration

The final conceptual design was developed based on the selected TRIZ solution. Figure 13 shows the final CAD assembly and fabricated prototype of the automatic footrest stand. The mechanism was then integrated with the existing Flexi-UC, as shown in Figure 14. The stand was positioned beneath the footrest to provide additional lateral ground support during mounting and dismounting while maintaining compactness when retracted.

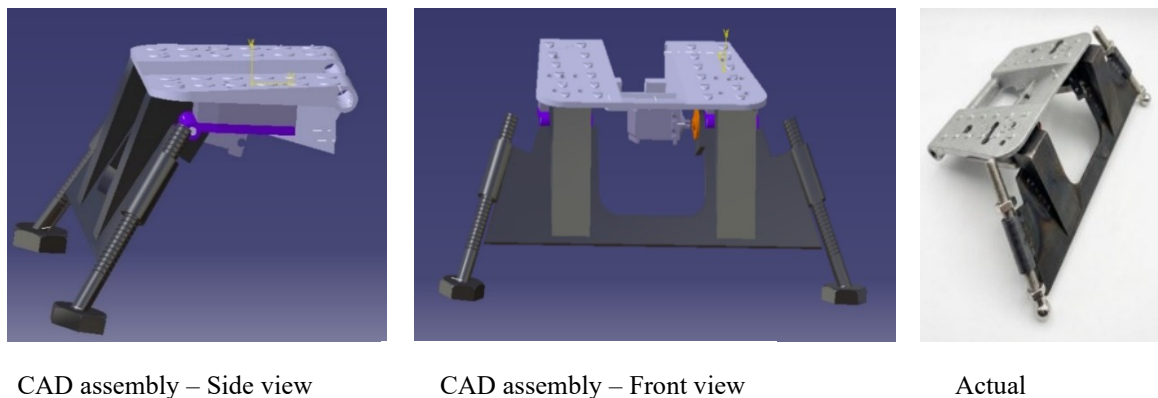


Figure 13. Final automatic footrest stand: CAD model and fabricated prototype

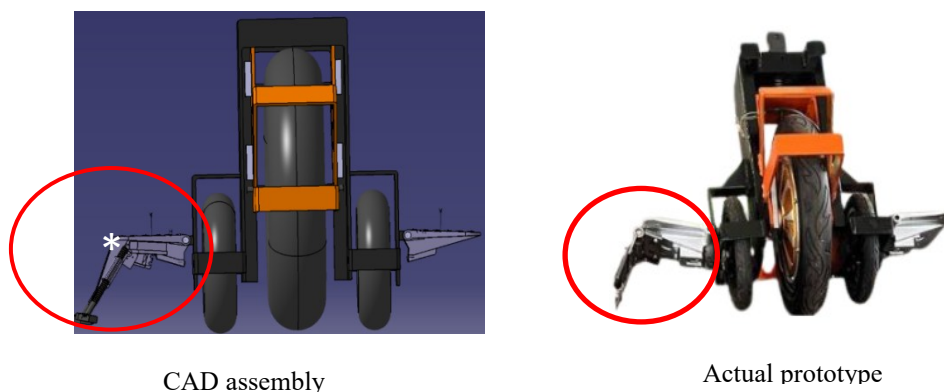


Figure 14. Integration of the automatic footrest stand with the Flexi-UC

The automatic stand added to the Flexi-UC, as shown in Figure 14, provides additional lateral support during mounting and dismounting and retracts during riding to keep the unicycle compact and prevent interference with the rider's movement. The FEA results confirm that the maximum stress remains below the allowable limit and that the design's total deformation is minimal. The improved functional analysis was subsequently refined to reflect the Engineering Contradiction (EC), Physical Contradiction (PC), and the validated performance of the optimized design. The mechanical integrity demonstrated by the simulations supports including the automatic stand as a stabilizing subsystem, indicating that the proposed mechanism can provide the intended ground-contact function without compromising the design's compactness. With the contradictions resolved, the updated functional analysis shows how the automatic stand integrates into the system as a functional stabilizing element.

Overall, Table 13 summarizes the key features of the Flexi-UC before and after the improvement. This comparison clearly shows how the improved Flexi-UC addresses the earlier design problem, with the added support function improving stability and overall system performance.

Table 13. The comparison table for before improvement and after improvement

Feature	Before Improvement	After Improvement
Stability mechanism	Entirely on the rider	Rider + mechanical stand support
Footrest role	Passive platform	Interface for stand activation (indirect)
Ground interaction	2 points (main wheel and one support wheel)	3 points (main wheel + 1 support wheel + stand)
Sensitivity to supersystem	High	Reduced
System ideality	Lower	Higher due to added passive function

4. CONCLUSION

The lateral stability issue during mounting and dismounting of EUCs has received little attention, and this paper addresses it systematically through the TRIZ method. This issue was addressed through an inventive solution developed using functional analysis, cause-and-effect chain analysis (CECA), engineering contradiction, and physical contradiction to conceptualize an innovative mechanism to overcome the main problem. This added mechanism does not compromise the Flexi-UC's compactness or riding dynamics. The TRIZ-guided design process helped resolve major physical contradictions and led to the development of an additional mechanism, namely the automatic stand, thereby improving the existing Flexi-UC design. The stability analysis indicated that the stability margin improved substantially and that the theoretical critical angle increased from -6.01° to 10.9° , while the stability margin increased from a negative value of approximately -0.08 m to a positive value of about $+0.15$ m under the stand-deployed condition. The finite element analysis confirmed that the structure operates safely within the elastic range with minimal deformation under the applied load, supporting the proposed mechanism's structural integrity.

The results indicate that TRIZ offers a systematic approach to converting design contradictions into practical engineering solutions, leading to a compact and reliable Flexi-UC mechanism that increases lateral stability, particularly during mounting and dismounting, while also contributing to innovation in micromobility design. Further experiments can be conducted with users under riding conditions to evaluate the mechanism's robustness, ergonomic comfort, and long-term durability. Future work should focus on improving the control logic and actuation responsiveness while minimizing the impact on battery performance. Dynamic, fatigue, and impact testing under cyclic loading conditions are recommended to evaluate the automatic stand's long-term structural durability. Future work should include dynamic and impact loading, along with fatigue and cyclic testing, to assess the automatic stand's long-term durability. The actuation system can also be improved in response and reliability. Consider an independent mechanical fail-safe or interlocking mechanism to prevent unintended stand deployment during riding. Further improvement should also focus on reducing the overall system mass. Testing with a larger and more diverse group of participants is recommended to provide a broader evaluation of the proposed design's usability and ergonomic performance. Overall, this research shows that TRIZ can be used as a systematic approach to produce practical design improvements and solve engineering problems in future personal mobility systems.

ACKNOWLEDGMENT

The authors gratefully acknowledge their respective institutions for their strong support in this study. All authors consent to the publication of this manuscript.

REFERENCES

- [1] M. Gonçalves and P. Cash, "The life cycle of creative ideas: Towards a dual-process theory of ideation," *Des. Stud.*, vol. 72, Jan. 2021, <https://doi.org/10.1016/j.destud.2020.100988>.
- [2] L. Chechurin, "TRIZ in Science. Reviewing Indexed Publications," in *Procedia CIRP*, Elsevier B.V., 2016, pp. 156–165. <https://doi.org/10.1016/j.procir.2016.01.182>.
- [3] P. Cuffe, "Flexible Mobility in the Smart City: The Role of Small Personal Electric Vehicles Flexible mobility in the smart city," *DIT – eseia Conference on Smart Energy Systems in Cities and Regions, Dublin, Ireland*, pp. 10–12, Feb. 2018, [Online]. Available: <http://hdl.handle.net/10197/9938>
- [4] D. Reyes-Huerta, H. Mitre-Hernandez, and U. Jaramillo-Avila, "Teaching and Learning TRIZ as an Innovative Educational Technology: A Systematic Literature Review," 2023.
- [5] O. Şen, "An Implementation of TRIZ in Innovative Product Development: An Oil Cap for Propeller Shaft," *Sakarya University Journal of Science*, vol. 28, no. 6, pp. 1296–1306, Dec. 2024, <https://doi.org/10.16984/saufenbilder.1514093>
- [6] N. Ramli, M. A. Azizi, N. A. Aziz, and M. A. Azman, "Advances and challenges in electric unicycle stability: a review of structural design and control mechanisms," *Journal of Engineering and Applied Science*, vol. 72, no. 1, p. 197, Dec. 2025, <https://doi.org/10.1186/s44147-025-00778-7>
- [7] Travers et al., "Moral Panic and Electric Micromobilities: Seeking Space for Mobility Justice," *Sociological Perspectives*, vol. 67, no. 1–3, pp. 83–108, Jun. 2024, doi: <https://doi.org/10.1177/07311214231193355>
- [8] L. Mihályi, X. A. Ji, G. Orosz, and D. Takács, "Data Collection for Learning the Dynamics and Control of an Electric Unicycle," *The Evolving Scholar*, 2023, <https://doi.org/10.24404/63fdc92b57f7d2ada3a24289>
- [9] Kang FK, "Stabilized electric unicycle, Flexi UC fabrication and experimental,," Universiti Putra Malaysia, 2024.
- [10] INMOTION USA, "InMotion V11 Electric Unicycle," INMOTION USA. [Online]. Available: <https://www.myninmotion.com/products/inmotion-v11-electric-unicycle>. [Accessed: Aug. 6, 2026].
- [11] Zuo, G. (2018). Intelligent somatic full-balance electric vehicle (U.S. Patent No. 9,908,579 B2). United States Patent and Trademark Office. <https://patents.google.com/patent/US9908579B2/en>
- [12] T. S. Yeoh, T. J. Yeoh, and C. L. Song, *TRIZ Systematic Innovation in Manufacturing*, 11th ed., vol. 1. Malaysia: Firstfruits Sdn Bhd, 2009.
- [13] K. Szczepanik and J. A. Chudziak, "TRIZ Agents: A Multi-Agent LLM Approach for TRIZ-Based Innovation," in *International Conference on Agents and Artificial Intelligence*, Science and Technology Publications, Lda, 2025, pp. 196–207. <https://doi.org/10.5220/0013321900003890>.
- [14] Scott, N., & Travers, N. (2023). E-micromobility, cycling, and 'good' active travel. *Active Travel Studies*, 3(1). <https://doi.org/10.16997/ats.1206>
- [15] L. Chen, Y. Song, S. Ding, L. Sun, P. Childs, and H. Zuo, "TRIZ-GPT: An LLM-augmented method for problem-solving," Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2408.05897>
- [16] Su, Y., Ding, L., Xu, P., Wang, G., Wei, Q., Liu, T., Wang, H., & Gao, H. (2025). Dimensional design method based on stability analysis to enhance traversability of heavy-duty hexapod robots over challenging terrains. *Mechanism and Machine Theory*, 214, 106141. <https://doi.org/10.1016/j.mechmachtheory.2025.106141>
- [17] Mungan, C. E., & Kabutz, H. (2024). The driven unicycle. *Physics Education*, 59(3), 035026. <https://doi.org/10.1088/1361-6552/ad3cc5>
- [18] Morales, J., Martínez, J. L., Mandow, A., Serón, J., & García-Cerezo, A. J. (2013). Static tip-over stability analysis for a robotic vehicle with a single-axle trailer on slopes based on altered

- supporting polygons. IEEE/ASME Transactions on Mechatronics, 18(2), 697–705.
<https://doi.org/10.1109/TMECH.2011.2181955>
- [19] Santoso, G., Ammarullah, M. I., Sugiharto, S., Hidayat, T., Khoeron, S., Bayuseno, A. P., & Jamari, J. (2024). TRIZ-based method for developing a conceptual laparoscopic surgeon's chair. Cogent Engineering, 11(1), 2298786. <https://doi.org/10.1080/23311916.2023.2298786>
- [20] Rahman, M. T. A., Ahmad, S., Akmeliawati, R., Altalmas, T., & Aula, A. (2013). Centre of gravity (C.O.G.)-based analysis on the dynamics of the extendable double-link two-wheeled mobile robot. IOP Conference Series: Materials Science and Engineering, 53(1), 012079. <https://doi.org/10.1088/1757-899X/53/1/012079>
- [21] Egoyan, A., & Moistsrapishvili, K. (2013). Equilibrium and stability of the upright human body. General Science Journal, 1–10.
[https://www.gsjournal.net/Science Journals/Research%20Papers/View/4581](https://www.gsjournal.net/Science%20Journals/Research%20Papers/View/4581)
- [22] C. K. Anil, S. Srinivas, M. Manjunath, M. Satish, and M. B. Khot, “Selection of Materials, Design and Analysis of FSAE Chassis for Different Engineering Application- A Comprehensive Review,” Mar. 31, 2025, *Society of Automotive Engineers Turkey*. <https://doi.org/10.30939/ijastech..1574611> .

A Modular, 3D-Printed, and Cost-Effective Near-Infrared Spectrometer with Machine Learning for Commodity Plastic Identification

WEI JIM TAN, ZULKIFLI MOHAMAD ARIFF*,
ARJULIZAN RUSLI, MUHAMAD SHARAN MUSA

*School of Materials and Mineral Resources Engineering, Universiti Sains Malaysia,
Nibong Tebal, Malaysia.*

**Corresponding author: zulariff@usm.my*

(Received: 6 May 2026; Accepted: 14 July 2026; Published online: 11 September 2026)

ABSTRACT: Efficient segregation of post-consumer plastics remains a critical challenge in achieving a sustainable circular economy because current sorting technologies are limited in cost, adaptability, and accessibility for community-level use. This study presents the development of a modular, cost-effective, and 3D-printed transmission near-infrared (NIR) spectrometer that operates in the 850 nm to 1100 nm range. The spectrometer was developed using open-source tools and resources such as Theremino and Scikit-learn, as well as commercially available optical components such as a desktop webcam, a diffraction grating, and an incandescent penlight. Validation of the spectrometer revealed minor wavelength deviations between theoretical and measured values. Measured values deviated by 5.1% with respect to the 546 nm mercury characteristic peak and 7.6% with respect to the 436 nm characteristic mercury peak emitted from a fluorescent lamp. This deviation is due to the low-cost optical components, but it did not impair the classification performance of the polymers. 504 plastic samples from seven material classes, namely acrylonitrile butadiene styrene (ABS), low-density polyethylene (LDPE), polyethylene terephthalate (PET), polymethyl methacrylate (PMMA), polypropylene (PP), green polystyrene (GPS), and red polystyrene (RPS), together with an empty baseline reference (MT) sample, were collected and processed using Principal Component Analysis (PCA), where 12 principal components were retained to capture 95% of the cumulative spectral variance. Classification performance was subsequently evaluated using Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and K-Nearest Neighbors (KNN) models. After hyperparameter tuning, both the SVM and KNN models achieved the highest held-out test accuracy (82.24%). SVM was the most reliable overall classifier, with slightly better class-balanced performance (F1-score of 82.36%) than KNN (81.98%) and LDA (73.71%). The empty baseline reference spectrum was classified perfectly with a precision, recall, and F1-score of 100%, confirming the consistency of the baseline penlight source spectrum. RPS showed greater feature overlap with GPS, and LDPE with PP. This study demonstrated that a low-cost, transmission NIR spectrometer operating between 850 nm and 1100 nm, combined with machine learning, can accurately identify commodity plastics, offering a promising community-level solution.

ABSTRAK: Pengasingan plastik pasca guna secara berkesan masih merupakan cabaran utama dalam mencapai ekonomi kitaran lestari. Kekurangan teknologi pengasingan plastik sedia ada dari segi kos, kebolehsuaian, dan kebolehcapaiangunaan di peringkat komuniti telah menjadi penyumbang kepada cabaran tersebut. Kajian ini menumpukan kepada pembangunan spektrometer inframerah dekat (NIR) jenis transmisi melalui pencetakan 3D bersifat modular, kos efektif, dan mempunyai julat operasi dari 850 nm hingga 1100 nm. Spektrometer ini dibina menggunakan platform sumber terbuka seperti Theremino dan Scikit-learn. Selain itu, komponen optik yang mudah diperoleh seperti kamera web desktop yang diubah suai, kisi

pembelauan, dan lampu pijar jenis pen turut digunakan. Semasa proses pengesanan fungsi spektrometer ini, sisihan panjang gelombang kecil antara nilai teori dan nilai yang diukur telah dikesan. Berdasarkan dapatan pengukuran, terdapat sisihan sebanyak 5.1% bagi puncak ciri merkuri pada panjang gelombang 546 nm dan 7.6% bagi puncak ciri merkuri pada panjang gelombang 436 nm yang dipancarkan oleh lampu pendarfluor. Penggunaan komponen optik kos rendah mengakibatkan berlakunya sisihan tersebut, tetapi keupayaan spektrometer ini dalam pengelasan polimer tidak terjejas. Sebanyak 504 sampel plastik daripada tujuh kelas bahan iaitu akrilonitril butadiena stirena (ABS), polietilena berketumpatan rendah (LDPE), polietilena tereftalat (PET), polimetil metakrilat (PMMA), polipropilena (PP), polistirena hijau (GPS), dan polistirena merah (RPS), bersama satu garis dasar rujukan kosong (MT) telah dikumpulkan dan diproses menggunakan Analisis Komponen Utama (PCA), di mana 12 komponen utama dikekalkan bagi mewakili 95% varians spektrum kumulatif. Prestasi pengelasan kemudiannya dinilai menggunakan model Mesin Vektor Sokongan (SVM), Analisis Diskriminan Linear (LDA), dan K-Jiran Terdekat (KNN). Selepas penalaan hiperparameter, model SVM dan KNN mencapai ketepatan ujian ketahanan tertinggi sebanyak 82.24%. SVM dikenal pasti sebagai pengelas keseluruhan yang paling boleh percaya dengan skor F1 sebanyak 82.36%, berbanding KNN (81.98%) dan LDA (73.71%). Spektrum sampel MT berjaya diklasifikasi dengan sempurna dengan nilai ketepatan, penarikan semula, dan skor F1 sebanyak 100%, sekaligus mengesahkan konsistensi spektrum asas yang dihasilkan daripada cahaya lampu pijar. Namun, bahan seperti RPS menunjukkan pertindihan ciri spektrum yang lebih tinggi dengan GPS, manakala spektrum LDPE pula bertindih dengan spektrum PP. Kajian ini berjaya membuktikan bahawa spektrometer NIR transmisi kos rendah yang dihasilkan melalui pencetakan 3D dan pembelajaran mesin dapat beroperasi pada julat 850 nm hingga 1100 nm dan mampu mengenal pasti plastik komoditi dengan tepat di samping menawarkan penyelesaian terbaik pada peringkat komuniti.

KEYWORDS: *Near-Infrared Spectroscopy (NIR), Plastic Identification, Low-Cost Spectrometer, Machine Learning, Principal Component Analysis (PCA).*

1. INTRODUCTION

Global plastic consumption continues to rise because of its remarkable combination of properties and characteristics, making it essential to daily life; yet only 9% of all waste plastic generated worldwide is recycled. The most preferred forms of plastic waste management remain landfilling, which accounts for almost 50% of plastic waste, followed by incineration at approximately 19% [1]. Furthermore, the complexity of plastic polymers and their additives further complicates recycling because plastic products are often compounded, with polymer molecules that are no longer uniform and include copolymers, blends, and composites [2]. Most recycling systems rely on centralized post-sorting technologies such as Material Recovery Facilities (MRFs), where contamination is already significant and sorting accuracy depends heavily on expensive optical systems or trained personnel [1], [3]. Fourier transform infrared spectroscopy (FTIR) and other laboratory-grade instruments provide excellent identification performance but remain inaccessible for community or household applications because of their cost, required expertise, and limited adaptability. A key advantage of FTIR is its non-destructive measurement mode, which preserves material integrity and supports material recovery during recycling processes.

In this context, a promising alternative, near-infrared (NIR) spectroscopy, offers similarly non-destructive and rapid identification capabilities. NIR spectroscopy is widely used for polymer analysis because of characteristic absorption features associated with molecular vibrations in the wavelength range of approximately 800–2500 nm [4]. However, despite these

advantages, commercially available NIR spectrometers remain relatively expensive and are not readily accessible for community or household applications.

Industrial-grade NIR and hyperspectral imaging (HSI) systems are effective for high-speed material recovery, but their high investment costs range from \$15,000 to \$50,000, which renders them inaccessible for community-level applications. This study demonstrates that by utilizing 3D printing and consumer-grade silicon sensors, the cost of NIR technology can be reduced to a few hundred dollars, making it a viable tool for decentralized recycling initiatives, fablabs, and small-scale workshops [5].

Nonetheless, recent advancements in open-source hardware, 3D printing, and low-cost optical components could extend access to commodity plastic segregation capabilities for the general public. Open-source devices such as the RM90 Theremino spectrometer show that affordable designs can achieve competitive performance in the visible region comparable to a commercial USB4000 spectrometer costing over RM5,000, suggesting similar approaches may apply in the NIR domain [6]. NIR spectrometer instrumentation requires appropriate diffraction optics, guided by the grating equation and long-pass filters to minimize order overlap.

Several DIY spectrometers are available online that employ different spectroscopic methods. Two projects designed specifically for plastic identification are the “Plastic Scanner” by Straller & Gessler [7] and “Ozirma” by Vivien Henry (<https://hackaday.io/project/179969-ozirma-near-infrared-spectrometer>). The Plastic Scanner is highly compact and features an array of IR LEDs spanning 850 to 1650 nm paired with an InGaAs sensor, but these components are expensive and difficult to source. Ozirma is larger and utilizes a concave mirror, a rotatable reflection diffraction grating controlled by a stepper motor, and a single-pixel InGaAs photodiode. Additionally, DIY spectrometers not specifically developed for plastic identification offer useful design insights, such as the OSMS visible-range spectrometer for liquid analysis and a 3D-printable modular smartphone visible spectrophotometer [8].

Despite existing advancements in DIY spectroscopy, a significant gap remains between high-cost industrial systems and current open-source tools. Innovative projects such as the 'Plastic Scanner' have demonstrated the efficacy of using discrete LED wavelengths for material identification; however, extending these systems to longer NIR wavelengths typically requires InGaAs sensors, which can be expensive and difficult for the general public to source. In this study, we propose a complementary approach that utilizes more accessible silicon-based sensors. Although these sensors restrict the detection range to 850 nm to 1100 nm, our prototype uses continuous-spectrum measurements to extract a higher density of meaningful classification patterns from this narrower range. This provides a cost-effective alternative for decentralized, community-level plastic segregation without expensive longer-wavelength components.

The prototype features a modular, 3D-printed spectrometer that utilizes unconventional, low-cost materials such as razor blades to create sharp aperture slit edges and modified incandescent penlights to achieve continuous NIR acquisition. This hardware innovation improves accessibility by showing that standard silicon-based webcam sensors, traditionally limited to the visible spectrum, can be repurposed for NIR polymer identification up to 1100 nm through simple filter modification. These hardware capabilities are supported by an optimized chemometric workflow featuring a rigorous machine learning pipeline that uses Parallel Analysis (PA) to objectively retain principal components. This approach ensures a high classification accuracy of 82.24% for seven commodity plastic classes, demonstrating the system's robustness even when accounting for inherent hardware-induced spectral deviations.

2. MATERIALS AND METHODS

2.1. Spectrometer Prototype Fundamental Components

The developed spectrometer, as shown in Fig. 1, was designed based on the principle of transmission spectroscopy, in which the plastic sample was positioned between the illumination source and the detector. The spectrometer design is governed by the grating equation Eq. (1), which defines the relationship between the groove spacing (d), the angle of incidence (θ), and the angle of diffraction (φ) for a given wavelength:

$$n\lambda = d(\sin\theta + \sin\varphi) \quad (1)$$

where n is the diffraction order. By utilizing a transmission grating with 1000 lines/mm ($d=1000$ nm), the system was optimized to disperse incident light across the webcam sensor. Spectral acquisition was performed using a desktop webcam (CW834M) as the detector after removal of its built-in infrared (IR) blocking filter. Because the system uses a silicon-based image sensor, the practical detection range was limited to approximately 1100 nm. An educational-grade transmission diffraction grating with a groove density of 1000 lines/mm was used to disperse incident light, while a Zomei 850 nm long-pass optical filter was incorporated to minimize overlap between diffraction orders in the visible and NIR regions. A modified Fable incandescent penlight served as the broadband light source and was powered using a regulated 3.3 V supply from an Arduino to reduce baseline spectral drift caused by decreasing battery discharge. The aperture slit was constructed using razor blades to produce sharp and straight slit edges for improved resolution.

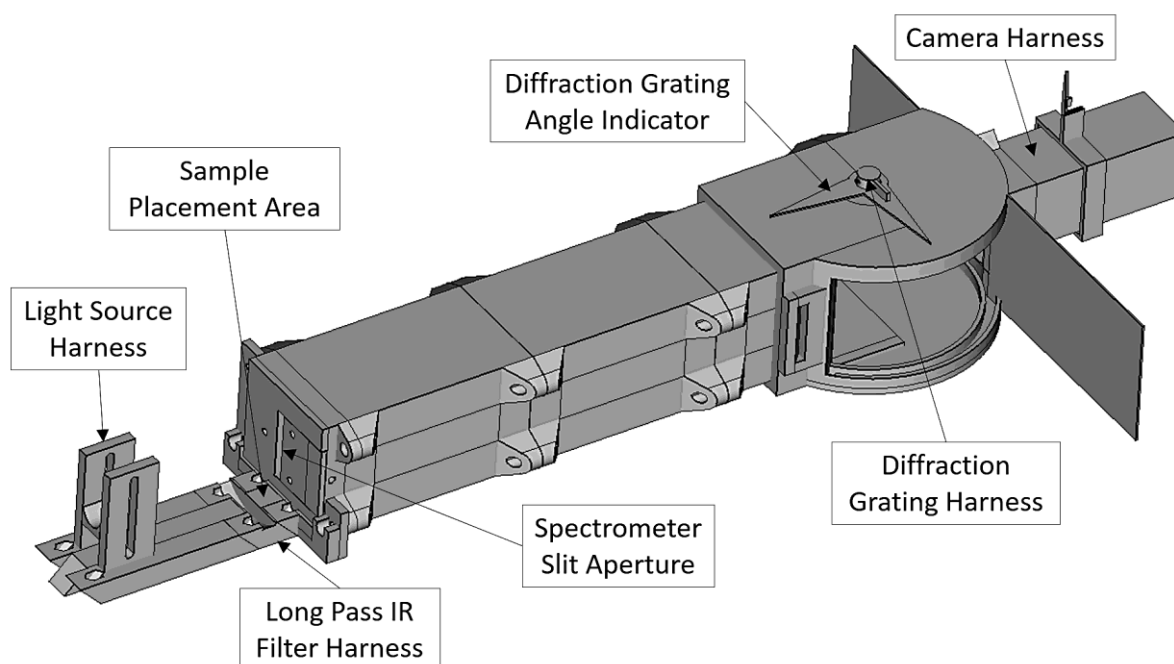


Figure 1. CAD 3D model of spectrometer

A total of seven polymer samples and one empty baseline reference sample were used in this study, namely acrylonitrile butadiene styrene (ABS), low-density polyethylene (LDPE), polyethylene terephthalate (PET), polymethyl methacrylate (PMMA), polypropylene (PP), green polystyrene (GPS), and red polystyrene (RPS), together with an empty baseline reference (MT) representing the absence of a sample. The GPS and RPS samples used the same base polystyrene polymer but differed in colorant. To ensure spectral consistency during model

development, all laboratory-prepared samples were compressed to a standardized thickness of 1.0 mm prior to measurement.

2.2. Spectrometer Operation and Data Analysis

2.2.1. Spectrometer Prototype Setup and Software Settings Standardization

The open-source Theremino Spectrometer Software was utilized. To optimize signal quality, the contrast setting was standardized at the maximum value to eliminate peripheral background noise inherent to the low-cost camera. The aperture width was standardized at 2.5 mm. The angle of incidence was set to 42.5° and the angle of diffraction to 22° to maximize NIR signal detection within the camera's sensor limits. These angles were determined by iteratively adjusting the diffraction grating and camera position relative to the optical axis. This configuration was finalized once the continuous spectral distribution was fully projected across the active area of the silicon sensor, validated through real-time visualization in the Theremino software interface.

2.2.2. Machine Learning Workflow

The machine learning workflow starts with a dataset of 504 samples and 460 spectral wavelength features, partitioned into a 70% training set and a 30% held-out test set using the `train_test_split` function. A stratified split, implemented via the `stratify` parameter, was applied to preserve class balance across both subsets. To avoid data leakage and overly optimistic performance estimates, all preprocessing steps, including feature standardization using `StandardScaler` and Principal Component Analysis (PCA) from Scikit-learn, were encapsulated within a Pipeline and fitted exclusively on the training data within each cross-validation loop before being applied to the corresponding validation folds, as shown in Fig. 2.

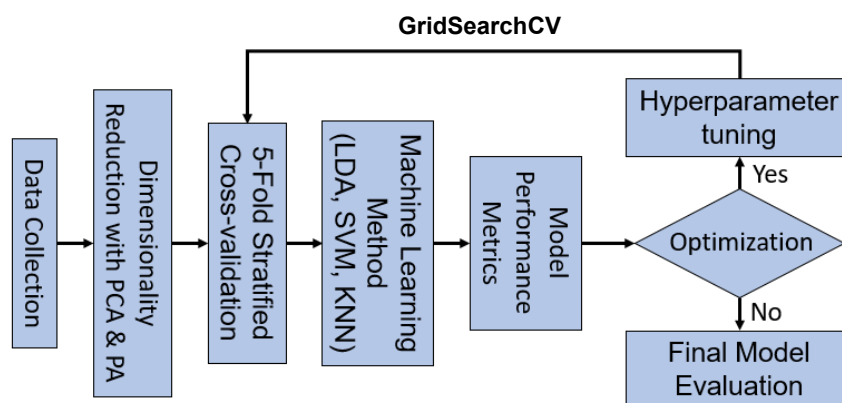


Figure 2. Workflow diagram of the proposed classification framework.

Principal component retention was determined using a combination of parallel analysis (PA) and a cumulative explained variance threshold of 95%. Parallel analysis compared eigenvalues from the training data with the 95th percentile of eigenvalues from randomly generated datasets using `rng.normal` noise. Based on this diagnostic analysis, the maximum value among these criteria was used, resulting in 12 principal components being retained for the final machine learning pipeline.

Classification performance of three machine learning algorithms was compared using 5-fold stratified cross-validation (`StratifiedKFold`), namely Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and K-Neighbors Classifier (KNN). Baseline performance was evaluated using models instantiated with their default parameters, followed by hyperparameter optimization using `GridSearchCV` to identify the best-performing model

configuration. The optimized models were then evaluated on the held-out test dataset, and confusion matrices were generated. Fig. 2 illustrates the process flowchart.

The classification performance of the machine learning models implemented for the polymer sample identification process was quantitatively evaluated using a set of performance metrics. These metrics include identification accuracy, precision, recall, and the F1-score, which are commonly used to assess machine learning algorithms in chemometric and material identification tasks. The derivation of these metrics utilizes data from the four fundamental outcomes of the classification task: True Positive (TP) and True Negative (TN), representing correct recognition of target and non-target classes, respectively, alongside False Positive (FP) and False Negative (FN), which denote instances of incorrect recognition. Machine learning model accuracy, precision, recall, and the F1-score are calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

3. RESULTS

3.1. Spectrometer Prototype Performance Discussion

The camera's maximum operating wavelength was approximately 1100 nm due to limitations of the silicon-based photovoltaic sensor in the modified webcam [9]. Spectrometer accuracy was verified by comparing the detected value to the actual value of characteristic mercury emission peaks (436 nm and 546 nm) emitted from a fluorescent lamp [10] using the diffraction grating equation Eq. (1). Wavelengths at the measured angles, namely 28° and 35° as shown in Table 1, were compared with the actual wavelength of characteristic mercury emission peaks. Deviations of 27.6 nm (5.1%) for the 546 nm peak and 33.5 nm (7.6%) for the 436 nm peak were observed, as shown in Table 1.

Table 1. Deviation between theoretical and experimental values

Measured angle of diffraction	Experimental Value	Actual value	Wavelength difference	Error
28°	469.5 nm	436 nm	33.5 nm	7.6%
35°	573.6 nm	546 nm	27.6 nm	5.1%

These deviations were attributed to the low-cost imaging system, including radial lens distortion [11] and surface non-planarity in the educational-grade transmission grating, as shown in Fig. 3. Fig. 3 visualizes the extent of this surface irregularity by displaying the diffraction grating's warp profile. This profile was qualitatively assessed by capturing the reflection of a digital grid from a computer screen off the grating's surface. A perfectly planar surface would reflect the grid in straight lines; the observed reflected grid curved in certain areas, indicating a warped surface and contributing to the observed deviations. The resulting accuracy was nevertheless sufficient for pattern-based polymer classification provided that the warp profile remains constant.

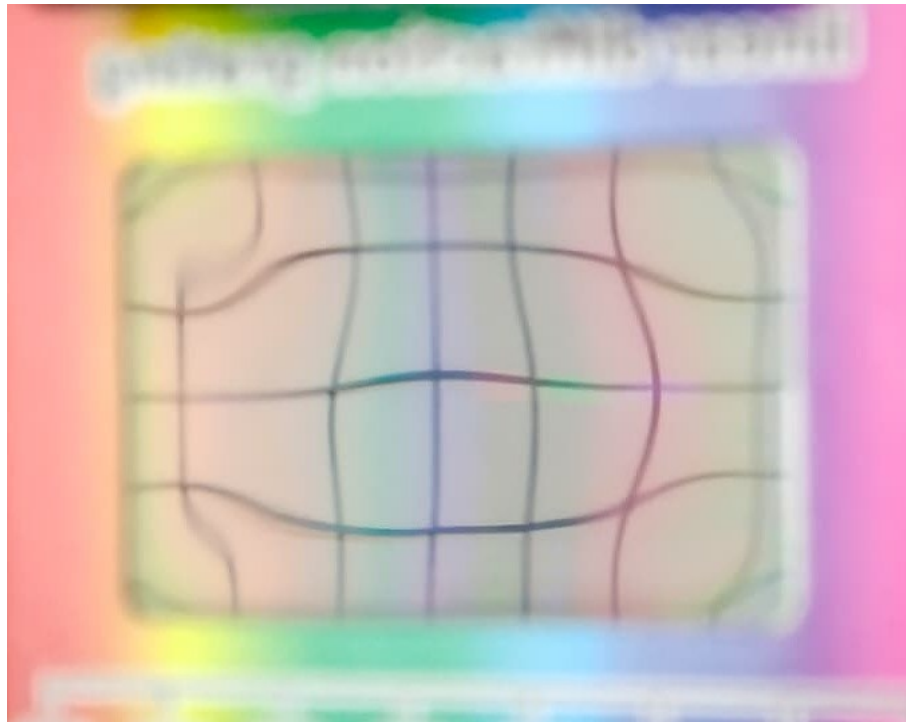


Figure 3. Reflection of grid with right angles across the diffraction grating.

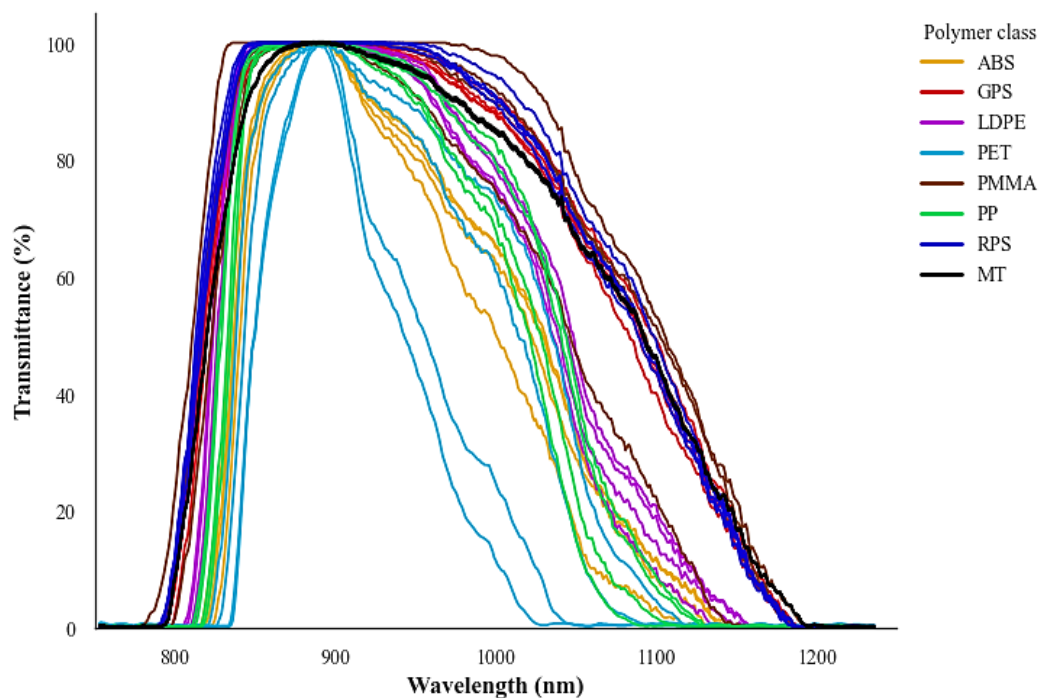


Figure 4. Different spectra of the seven lab-manufactured plastic samples and an empty baseline reference.

Although the low-cost optical components introduced minor deviations in absolute wavelength precision, classification performance was unaffected. The machine learning algorithms prioritize overall relative spectral features. Fig. 4 shows the different spectra of the seven lab-manufactured plastic samples and an empty baseline reference.

Transparent samples have higher transmittance compared to semi-transparent samples. As a result, transparent samples such as GPS, RPS, and PMMA exhibit a wider range of high transmittance between 900 nm and 1200 nm, overlap with one another, and closely resemble the empty baseline reference (MT) spectrum.

3.2. Dimensionality Reduction Results Using PCA

Principal Component Analysis (PCA) was conducted using Parallel Analysis (PA), integrating a scree plot and a line indicating the 95th percentile eigenvalues obtained in Monte Carlo PA. Eigenvalues above this percentile indicate the optimal number of components to retain. As shown in Fig. 5, the PA results suggest a cutoff of 4 principal components (PCs) [12]. However, 12 PCs were required to explain 95% of the dataset's cumulative variance. Over-extraction is a less severe risk than under-extraction because the variation captured by successive components decreases sequentially [13]. In this context, retaining 12 components safeguarded against the negative effects of under-extraction. Therefore, 12 PCs were selected for subsequent machine learning development using SVM, LDA, and KNN classification models, which were commonly employed in plastic material identification [14]. For clarity, only the leading principal components are presented in the scree plot in Fig. 5.

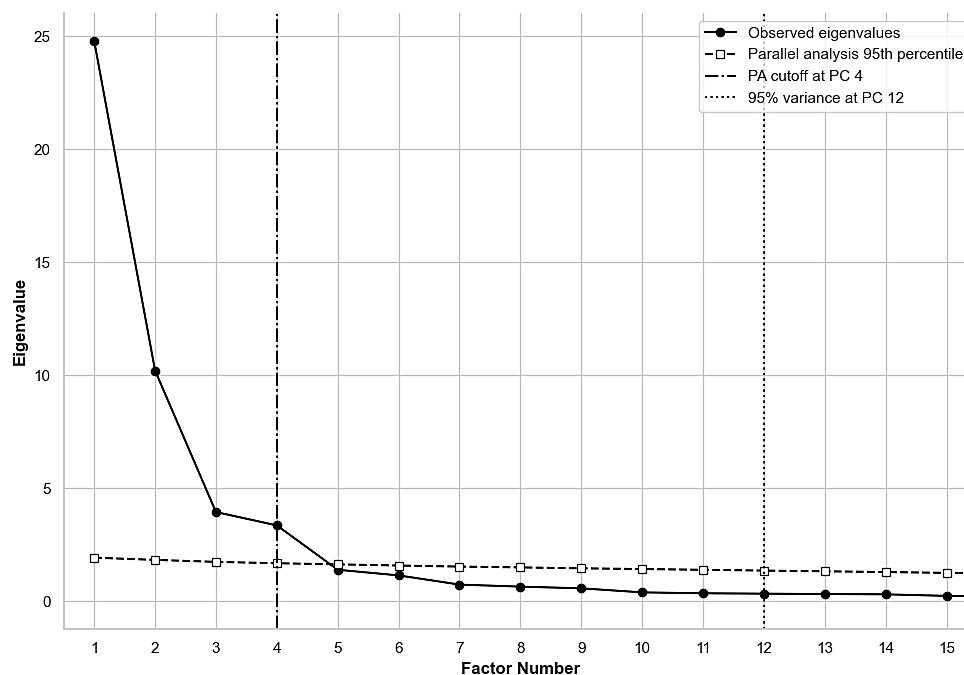


Figure 5. Parallel Analysis showing cut-off at 4 PCs

A 2D PCA projection, as shown in Fig. 6, capturing 69.88% of the variance from PC1 and PC2, was generated for visualization purposes to illustrate clustering patterns across classes. The empty baseline reference (MT) class displayed the tightest distribution. Similarly, Red Polystyrene (RPS) and Green Polystyrene (GPS) clustered very closely together. Meanwhile, PET exhibited the largest within-class spread, indicating higher intra-class variability. Partial class overlap was observed between the LDPE, PP, and ABS classes.

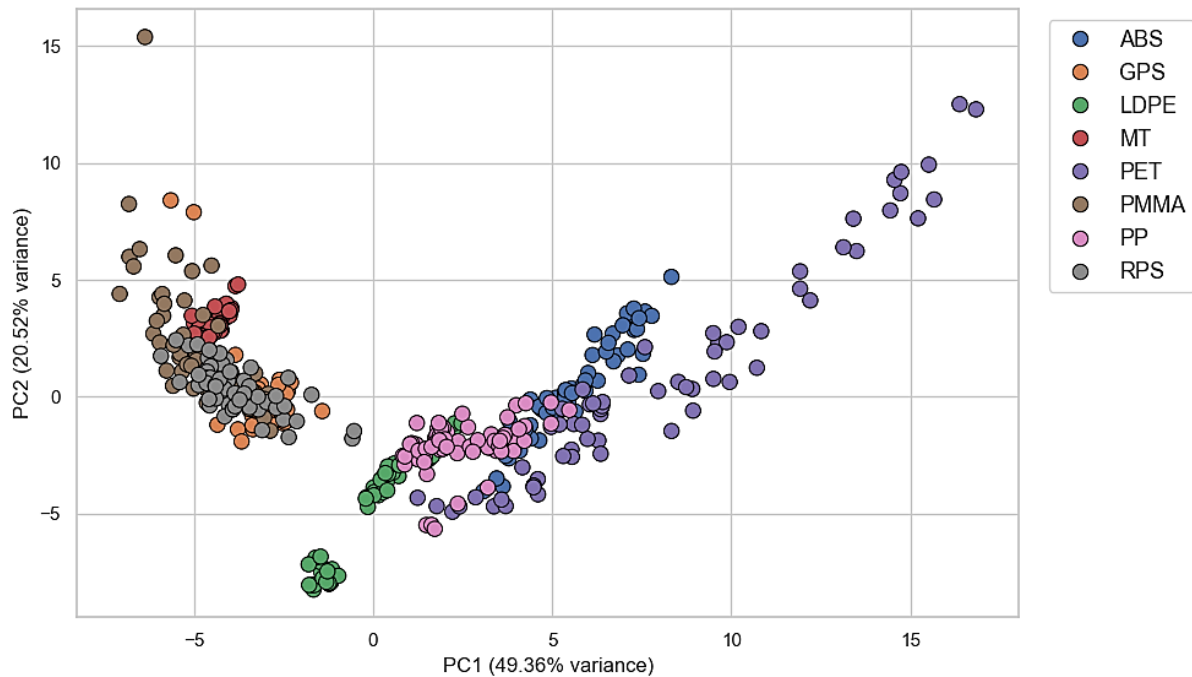


Figure 6. 2D PCA projection capturing 69.88% of the total variance.

3.3. Model Performance Results: Cross-Validation and Hyperparameter Tuning

Baseline classification performance was evaluated using 5-fold stratified cross-validation on the training set. The Support Vector Machine (SVM) achieved the highest baseline accuracy of 87.79%, followed by Linear Discriminant Analysis (LDA) at 84.66% and K-Nearest Neighbors (KNN) at 83.24%.

Following hyperparameter tuning, SVM and KNN achieved the highest held-out test accuracy, both reaching 82.24%. SVM used a linear kernel with a regularization parameter of $C = 0.1$, while KNN performed best with 7 neighbors and distance weighting. Although both models obtained the same test accuracy, SVM showed slightly better overall class-balanced performance, with a higher F1-score (82.36%) than KNN (81.98%), and was therefore selected as the preferred final model. In contrast, LDA showed lower held-out performance, with a test accuracy of 73.68%. Table 2 summarizes the results.

Table 2. Model performances following hyperparameter tuning

Model	PC (n)	Prediction (%)				
		Baseline Accuracy	Held-out test Accuracy	Precision	Recall	F-1 Score
SVM	12	87.79	82.24	83.94	82.24	82.36
LDA	12	84.66	73.68	75.32	73.68	73.71
KNN	12	83.24	82.24	83.94	82.24	81.98

3.4. Classification Behavior Across Prediction Classes

Table 3 further illustrates each prediction class's classification behavior in terms of precision, recall, and F1-score. LDA remains the weakest model, except that it has a higher precision score for the PMMA class than the KNN model. The SVM and KNN models showed varying strengths in predicting different classes. For example, the SVM model predicted ABS, PET, and PMMA better than KNN, while KNN performed better on GPS, LDPE, and PP.

Nonetheless, all models classified the MT class perfectly, with precision, recall, and F1-score of 100%, confirming the consistency of the empty baseline reference spectrum.

For the SVM model, LDPE and RPS both showed low recall values of 63.16%, indicating that a substantial fraction of true samples from these classes were misclassified. This is understandable for the case of RPS due to its base polystyrene polymer, which is similar to that of GPS, as shown in the confusion matrix in Fig. 7. For the case of LDPE, the low recall value of 63.16%, indicating misclassification with the PP class as shown in Fig. 7, was largely due to the chemical similarity, being in the same polyolefin class [15] Furthermore, the SVM performance results highlight that PP exhibited a high recall of 89.47% but the lowest overall precision of 60.71%, meaning the SVM model also frequently overpredicted the PP class.

For KNN, RPS remained particularly problematic, with recall dropping even further to 52.63%. Like SVM, KNN also struggled with PP, showing high recall (89.47%) but relatively low precision (62.96%), further suggesting a tendency to overpredict the PP class.

Table 3. Precision, recall, and F1-score of the prediction classes by respective classification models.

Model	Metric	Scores (%)							
		ABS	GPS	LDPE	MT	PET	PMMA	PP	RPS
SVM	Precision	100.00	75.00	85.71	100.00	94.12	80.95	60.71	75.00
	Recall	89.47	78.95	63.16	100.00	84.21	89.47	89.47	63.16
	F1-score	94.44	76.92	72.73	100.00	88.89	85.00	72.34	68.57
LDA	Precision	80.95	62.50	66.67	100.00	100.00	76.47	56.00	60.00
	Recall	89.47	78.95	63.16	100.00	68.42	68.42	73.68	47.37
	F1-score	85.00	69.77	64.86	100.00	81.25	72.22	63.64	52.94
KNN	Precision	88.89	80.95	86.67	100.00	93.75	75.00	62.96	83.33
	Recall	84.21	89.47	68.42	100.00	78.95	94.74	89.47	52.63
	F1-score	86.49	85.00	76.47	100.00	85.71	83.72	73.91	64.52

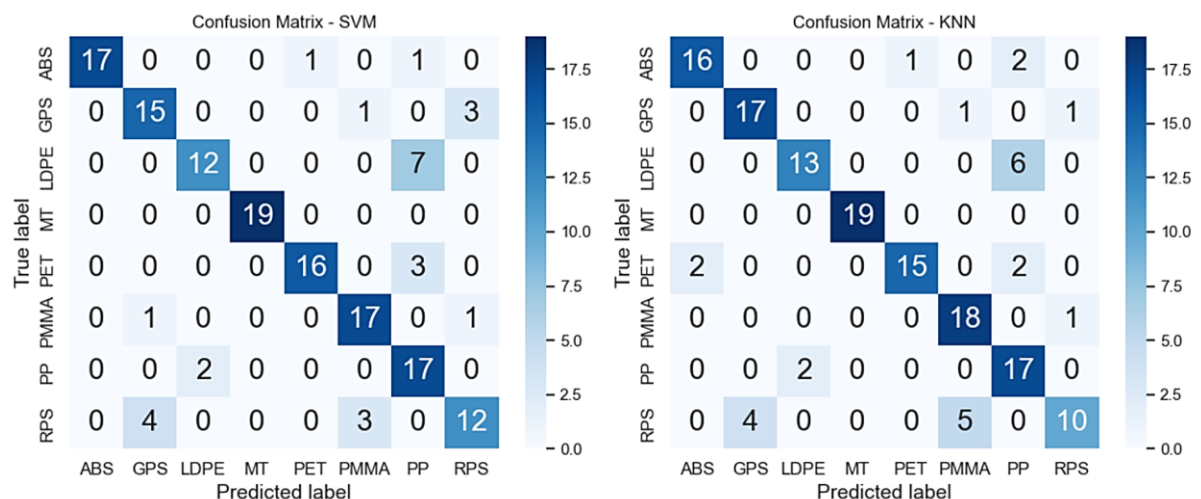


Figure 7. Confusion matrices of SVM and KNN from the held-out test set.

4. DISCUSSION

4.1. Discussion of Model Performance and Polymer Identification

Previous research reported 72.50% accuracy for waste plastics when operating between 410 nm and 940 nm using Convolutional Neural Networks [16]. This demonstrated that low-cost near-infrared sensors combined with machine learning methods can recognize plastics effectively. This performance is comparable to the overall identification accuracy of 82.24% achieved in the current study. While high-end laboratory prototypes using Principal Component Analysis (PCA)-optimized filters have achieved accuracies ranging from 94.44% to 100%, those systems often utilize expensive InGaAs sensors to access wavelengths above 1100 nm [14]. The proposed prototype's efficacy suggests that continuous-spectrum acquisition, combined with robust chemometric models, can extract high-density information even from the restricted range of a silicon-based sensor.

4.2. Machine Learning Models Comparison

Selecting SVM, KNN, and LDA as classification models aligns with established benchmarks in the plastic recycling industry. These three algorithms are fundamental chemometric tools for handling the high dimensionality and overlapping spectral features of polymers [14].

In this study, Support Vector Machine (SVM) emerged as the most reliable classifier, achieving an F1-score of 82.36%. This performance is consistent with research showing that SVM offers superior classification efficacy for highly affordable short-wave near-infrared (SW-NIR) miniaturized spectrometers compared with Linear Discriminant Analysis (LDA) or K-Nearest Neighbors (KNN) [4]. These results also align with findings that identify SVM as an optimal model for rapid, on-site material classification using a handheld near-infrared (NIR) spectrometer [17].

SVM's superiority likely stems from its ability to use a linear kernel that models the linear relationship between absorbance and chemical composition described by the Beer-Lambert law while remaining robust to hardware-induced noise inherent in low-cost sensors. Furthermore, SVM is particularly effective at preventing underfitting and overfitting in high-dimensional spectral data, making it well suited for classification tasks with limited training samples [17].

4.3. Analysis of Polymer Feature Overlap

The observed confusion between LDPE and PP reflects a documented challenge in vibrational spectroscopy. Near-infrared spectroscopy often struggles to differentiate polyolefins such as LDPE and PP because of their extreme spectral similarity and chemical overlap. These polymers have nearly identical C-H stretching overtones—specifically the second overtone near 1200 nm—and differentiating them often requires high-precision instrumentation to resolve subtle differences in methyl group density and short-chain branching (SCB) [15].

Similarly, the spectral overlap between green and red polystyrene (GPS and RPS) is anticipated, as both variants share an identical molecular fingerprint associated with the polystyrene base polymer. While specific colorants can shift absolute transmittance, these additives do not fundamentally alter the polymer backbone's molecular vibrations.

4.4. Study Limitations and Socio-Economic Implications

The primary physical limitation of this study is the silicon bandgap of 1.12 eV, which restricts the prototype's detection capability to wavelengths below 1100 nm [9]. Beyond this threshold, standard silicon sensors become "blind" to the more prominent second overtones of C-H vibrations found in industrial NIR applications [14].

Despite these technical trade-offs, the system provides a necessary technological bridge for decentralized recycling. Industrial units like Titech's PolySort can process up to 5 tons per hour [18], but their high investment costs, such as hyperspectral imaging (HSI) technology of up to \$ 50,000, are prohibitive for community workshops [5]. The proposed 3D-printed modular framework reduces these costs to a few hundred dollars, democratizing access to scientifically sound material segregation tools for fablabs and small-scale recovery initiatives.

5. CONCLUSION

This study successfully demonstrated the development of a modular, cost-effective, and 3D-printed transmission NIR spectrometer capable of operating in the 850 nm to 1100 nm range using open-source tools and commercially available optical components. By integrating machine learning with Principal Component Analysis (PCA) for dimensionality reduction, the system effectively extracted classification patterns from a restricted spectral range. Among the evaluated models, both the Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) achieved a held-out test accuracy of 82.24%, with SVM proving to be the most reliable overall classifier due to its superior class-balanced performance (F1-score of 82.36%).

While the hardware setup introduced wavelength deviations, specifically a 7.6% error at the 436 nm peak and 5.1% at the 546 nm peak, these were attributed to inherent optical limitations such as radial lens distortion and surface non-planarity in the educational-grade transmission diffraction grating. Despite these absolute precision errors, the machine learning algorithms maintained high performance by prioritizing relative spectral features, successfully identifying the empty baseline (MT) with 100% accuracy. However, challenges remain in distinguishing polymers with significant chemical similarity or feature overlap, such as LDPE with PP and RPS with GPS.

In conclusion, this research provides a viable framework for creating low-cost identification tools that bridge the gap between expensive laboratory-grade instruments and the practical needs of decentralized, community-level plastic segregation. Such accessible technology is essential to reduce contamination in mixed plastic waste streams and accelerate the transition toward a sustainable circular economy.

ACKNOWLEDGMENT

The authors acknowledge the Ministry of Higher Education, Malaysia, for the financial support provided through the Fundamental Research Grant Scheme (Ref. No. FRGS/1/2022/TK09/USM/01/4).

REFERENCES

- [1] Macheca AD, Mutuma B, Adalima JL, Midheme E, Lucas LHM, Ochanda VK, Mhlanga SD. (2024) Perspectives on plastic waste management: challenges and possible solutions to ensure its sustainable use. *Recycling*, 9(5): 77. <https://doi.org/10.3390/recycling9050077>
- [2] Lahl U, Zeschmar-Lahl B. (2024) Material recycling of plastics—a challenge for sustainability. *Sustainability*, 16(15): 6630. <https://doi.org/10.3390/su16156630>

-
- [3] Bernat K. (2023) Post-consumer plastic waste management: from collection and sortation to mechanical recycling. *Energies*, 16(8): 3504. <https://doi.org/10.3390/en16083504>
 - [4] Grabska J, Huck CW, Bec KB. (2021) Principles and applications of miniaturized near-infrared (NIR) spectrometers. *Chem. Eur. J.*, 27(5): 1514–1532. <https://doi.org/10.1002/chem.202002838>
 - [5] Werner T, Dawoud M, Aschenbrenner D, Taha I. (2025) Cost-efficient detection of plastics from post-consumer packaging waste using selected bands in the near-infrared spectrum. *Macromol. Mater. Eng.*, 2025: 2500143. <https://doi.org/10.1002/mame.202500143>
 - [6] Silva WRF, Suarez WT, Reis C, dos Santos VB, Carvalho FA, Reis EL, Vicentini FC. (2021) Multifunctional webcam spectrophotometer for performing analytical determination and measurements of emission, absorption, and fluorescence spectra. *J. Chem. Educ.*, 98(4): 1442–1447. <https://doi.org/10.1021/acs.jchemed.0c01085>
 - [7] Straller A, Gessler B. (2019) Identification of plastic types using discrete near infrared reflectance spectroscopy. ResearchGate. https://www.researchgate.net/publication/337868860_Identification_of_Plastic_Types_Using_Discrete_Near_Infrared_Reflectance_Spectroscopy
 - [8] Laganovska K et al. (2020) Portable low-cost open-source wireless spectrophotometer for fast and reliable measurements. *HardwareX*, 7: e00108. <https://doi.org/10.1016/j.ohx.2020.e00108>
 - [9] Huang S et al. (2024) Broadband-spectral-responsivity of black silicon photodetector with high gain and sub-bandgap sensitivity by titanium hyperdoping. *Opt. Laser Technol.*, 171: 110399. <https://doi.org/10.1016/j.optlastec.2023.110399>
 - [10] Rodrigues M, Marques MB, Carvalho PS. (2016) How to build a low cost spectrometer with Tracker for teaching light spectra. *Phys. Educ.*, 51(1): 014002. <https://doi.org/10.1088/0031-9120/51/1/014002>
 - [11] Liebold F, Mader D, Sardemann H, Eltner A, Maas HG. (2023) A bi-radial model for lens distortion correction of low-cost UAV cameras. *Remote Sens.*, 15(22): 5283. <https://doi.org/10.3390/rs15225283>
 - [12] Ledesma RD, Valero-Mora P, Macbeth G. (2015) The scree test and the number of factors: a dynamic graphics approach. *Span. J. Psychol.*, 18: E11. <https://doi.org/10.1017/sjp.2015.13>
 - [13] Peres-Neto PR, Jackson DA, Somers KM. (2005) How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Comput. Stat. Data Anal.*, 49(4): 974–997. <https://doi.org/10.1016/j.csda.2004.06.015>
 - [14] Neo ERK, Yeo Z, Low JSC, Goodship V, Debattista K. (2022) A review on chemometric techniques with infrared, Raman and laser-induced breakdown spectroscopy for sorting plastic waste in the recycling industry. *Resour. Conserv. Recycl.*, 180: 106217. <https://doi.org/10.1016/j.resconrec.2022.106217>
 - [15] Li S, Ivancic RJS, Sutliff BP, Huang D, Blázquez-Blázquez E, Martin TB, Migler KB, Audus DJ, Orski SV. (2026) Predicting properties from near-infrared spectra with machine learning for improved polyolefin differentiation. *ACS Polym. Au*, 6(4): 246–255. <https://doi.org/10.1021/acspolymersau.5c00131>
 - [16] Martinez-Hernandez U, West G, Assaf T. (2024) Low-cost recognition of plastic waste using deep learning and a multi-spectral near-infrared sensor. *Sensors*, 24(9): 2821. <https://doi.org/10.3390/s24092821>
 - [17] Yang Y, Zhang X, Yin J, Yu X. (2020) Rapid and nondestructive on-site classification method for consumer-grade plastics based on portable NIR spectrometer and machine learning. *J. Spectrosc.*, 2020: 6631234. <https://doi.org/10.1155/2020/6631234>
 - [18] Brunner S, Fomin P, Kargel C. (2015) Automated sorting of polymer flakes: fluorescence labeling and development of a measurement system prototype. *Waste Manag.*, 49: 600–607. <https://doi.org/10.1016/j.wasman.2014.12.006>
-

Kansei-Based AHP Framework for Feeding Mechanism Concept Selection in Agroindustrial Coconut Deshelling Machine Design

HETY HANDAYANI HIDAYAT^{1,2}, BAMBANG PURWANTANA^{1*},
RADI¹, MIRWAN USHADA³

¹Department of Agricultural and Biosystem Engineering, Faculty of Agricultural Technology,
Universitas Gadjah Mada, Yogyakarta, Indonesia

²Department of Agricultural Engineering, Faculty of Agriculture,
Universitas Jenderal Soedirman, Banyumas, Indonesia

³Department of Agroindustrial Technology, Faculty of Agricultural Technology,
Universitas Gadjah Mada, Yogyakarta, Indonesia

*Corresponding author: bambang_pw@ugm.ac.id

(Received: 7 July 2026; Accepted: 15 August 2026; Published online: 11 September 2026)

ABSTRACT: Indonesia is one of the world's biggest producers, with an estimated US\$2.05 billion in export value for the coconut industry in 2023. However, mechanized deshelling is limited by the non-uniformity of coconut geometry and intensive operator interaction with sharp rotating parts, which may lead to coconut meat breakage, unstable feeding, fatigue, and safety risks. While human-centered agricultural machinery design and multicriteria equipment selection have received increasing attention, integrating user-oriented Kansei criteria with AHP at the pre-prototype concept-selection stage remains limited. This paper develops a Kansei-based Analytic Hierarchy Process (AHP) framework to select the most feasible feeding mechanism concept before detailed design and prototyping. The baseline requirements identified for guidance, clamping, positioning, pushing, rotation control, operator protection, and implementation feasibility were developed into three alternative concepts: V-Guided Mechanical Feeding, Box-Guided Mechanical Feeding, and Assisted Controlled Feeding. Nine experts evaluated seven Kansei criteria, namely Performance, Practicality, Functionality, Durability, Usability, Pleasurability, and Efficiency. All the aggregated pairwise-comparison matrices were consistent ($CR \leq 0.10$). Box-Guided Mechanical Feeding performed best (0.411) with the baseline weights, and Assisted Controlled Feeding was close behind (0.401). This result indicates a near-tie rather than an outright win. Furthermore, the sensitivity analysis showed that the preferred concept changes when much higher emphasis is placed on functionality or pleasurability. The preferred concept is box-guided for both baseline and performance-oriented conditions. The framework shows a context-dependent trade-off between mechanical control, implementation practicality, and human-centered requirements, providing a transparent basis for concept selection in small- and medium-scale agro-industrial machinery.

ABSTRAK: Indonesia merupakan antara pengeluar kelapa terbesar di dunia, dengan nilai eksport industri kelapa dianggarkan sebanyak AS\$2.05 bilion pada tahun 2023. Walau bagaimanapun, proses pengupasan tempurung secara mekanikal masih terhad disebabkan oleh ketidakseragaman geometri kelapa serta keperluan interaksi intensif antara pengendali dengan komponen tajam yang berputar. Keadaan ini boleh mengakibatkan isi kelapa pecah, proses suapan tidak stabil, keletihan pengendali dan risiko keselamatan. Walaupun reka bentuk jentera pertanian berpusatkan manusia dan pemilihan peralatan berasaskan pelbagai kriteria semakin mendapat perhatian, pengintegrasian kriteria Kansei yang berorientasikan pengguna dengan Proses Hierarki Analitik (Analytic Hierarchy Process, AHP) pada peringkat

pemilihan konsep sebelum pembangunan prototaip masih terhad. Makalah ini membangunkan kerangka AHP berasaskan Kansei bagi memilih konsep mekanisme suapan yang paling sesuai sebelum peringkat reka bentuk terperinci dan pembangunan prototaip. Keperluan asas yang dikenal pasti merangkumi aspek pemanduan, pengapitan, peletakan, penolakan, kawalan putaran, perlindungan pengendali dan kebolehlaksanaan pelaksanaan. Keperluan ini digunakan untuk membangunkan tiga konsep alternatif, iaitu Suapan Mekanikal Berpandu-V, Suapan Mekanikal Berpandu-Kotak dan Suapan Terkawal Berbantu. Sembilan orang pakar menilai tujuh kriteria Kansei, iaitu Prestasi, Kepraktisan, Kefungsian, Ketahanan, Kebolegunaan, Keseronokan dan Kecekapan. Semua matriks perbandingan berpasangan yang diagregatkan menunjukkan tahap ketekalan yang boleh diterima ($CR \leq 0.10$). Berdasarkan pemberat asas, Suapan Mekanikal Berpandu-Kotak mencatatkan prestasi terbaik (0.411), diikuti rapat oleh Suapan Terkawal Berbantu (0.401). Dapatan ini menunjukkan keputusan yang hampir setara dan bukannya kemenangan mutlak bagi satu konsep. Selain itu, analisis sensitiviti menunjukkan bahawa konsep pilihan berubah apabila penekanan yang jauh lebih tinggi diberikan kepada aspek kefungsian atau keseronokan. Konsep berpandu kotak kekal sebagai pilihan utama dalam keadaan asas dan keadaan yang berorientasikan prestasi. Kerangka ini menunjukkan pertukaran keutamaan yang bergantung pada konteks antara kawalan mekanikal, kepraktisan pelaksanaan dan keperluan berpusatkan manusia, sekaligus menyediakan asas yang telus bagi pemilihan konsep jentera agroindustri berskala kecil dan sederhana.

KEYWORDS: *Coconut Deshelling Machine, Human-oriented, Multicriteria, Optimum solution*

1. INTRODUCTION

The coconut processing industry is a developing sector in Indonesia, supported by abundant raw materials and Indonesia's status as the world's largest coconut producer [1]. In addition, coconuts also have a wide range of economically valuable products in the global market, such as coconut milk and coconut oil [2], [3]. In 2023, Indonesia produced 14.967 million tonnes of coconut in whole-nut equivalent, and the value of coconut exports reached about US\$2.047 billion, the International Coconut Community reported [1]. In the industrial context, transforming coconuts into various products involves a long, complex series of processes. One crucial process is deshelling. In this process, the coconut kernel is separated from its shell. The expected result of deshelling is an intact coconut kernel, which facilitates the next process: removing the testa, i.e., the thin brown skin covering the white coconut flesh. The clean coconut kernel will subsequently serve as the main input for various coconut-derived products [4].

In various coconut industry centers in Indonesia, the deshelling process has been using a deshelling machine [5]. However, the current use of machines still faces several issues, especially in the feeding section, such as the risk of workplace accidents and broken coconuts. This is because coconut sizes are not uniform [6], [7]. In some industries in Indonesia, the main raw material is limited to tall coconut varieties, but their sizes vary widely with age and soil conditions [8]. Moreover, the variability in coconut shell thickness and hardness complicates the mechanical deshelling process for coconut feeding [9].

In addition to issues arising from the coconut material, deshelling problems also stem from the current high level of operator involvement. Operators interact closely with the material and the sharp, moving components of the deshelling machine. The operator must direct, press, and maintain the coconut in the correct position so that it continuously touches the rotary cutter area and is correctly positioned between the rotary and the blade [5]. Therefore, the operator

must have high skill and concentration during the deshelling process. Consequently, operator fatigue may cause the process to become unstable and inconsistent. This condition may increase the risk of kernel breakage and workplace accidents. Problems arising from both the coconut material and the operator's side necessitate a better mechanism, especially in feeding the coconut during the deshelling process with this machine. Hence, the feeding problem stems from a combination of the coconut's non-uniform geometry and the need for intense manual control near sharp, moving components, which ultimately affects process stability, kernel integrity, and operator safety.

When designing this feeding mechanism concept, many aspects must be considered. Specifically, this mechanism must ensure the main functions of the existing process, namely guiding, holding, pushing, and maintaining the coconut position relative to the rotary cutter and fixed blade [5]. This requirement is necessitated by the operating sequence, in which the coconut must be continuously positioned and maintained in contact with the cutting zone during deshelling [5]. The selection of the concept should be assessed not only on its technical capabilities but also on its feasibility of implementation. The concept of this feeding mechanism must be applicable and realistic, especially given the condition of the coconut industry in Indonesia, which is mostly small- and medium-scale enterprises [10]. In addition, the new mechanism needs to place the operator at the center of the design process, including safety and comfort. Previous studies on agricultural processing machinery have revealed that human factors and user experience should be considered early in the design rather than only after prototype development [11], [12]. Therefore, in this study, user expectations and impressions are translated into technical design requirements using Kansei Engineering [13].

The decision-making for selecting the feeding mechanism concept must consider multidimensional factors simultaneously. Each alternative concept may have strengths in some factors, but weaknesses in others. A concept may be easier to fabricate and maintain, but it might not sufficiently accommodate the coconut variability in size and operator safety. On the other hand, another concept may improve performance and safety, but it may be difficult to implement because of its mechanical complexity and high investment requirements. Therefore, Kansei Engineering is used to translate the expectations of operators, such as confidence, perceived safety, and ease of use, into explicit evaluation criteria, and AHP structures the trade-offs between these criteria and tests the logical consistency of expert judgments [13], [14], [15]. In current research in agricultural machinery, multicriteria methods are used to compare alternatives. However, such studies are mainly dedicated to the selection of existing machines or developed systems, rather than to human-oriented concept selection before the prototyping phase [16], [17]. Therefore, there is a lack of evidence for a structured concept-stage framework that integrates technical feasibility with operator-oriented Kansei criteria for coconut deshelling feeding mechanisms. Hence, this study aims to develop a Kansei-based AHP framework to select the most feasible feeding mechanism concept for the coconut deshelling machine design as a basis for decision-making before detailed design and prototype development.

2. METHODOLOGY

This research uses the Analytical Hierarchy Process (AHP) integrated with Kansei engineering to select the most feasible feeding mechanism concept for the coconut deshelling machine design. In this case, KE serves as a bridge, translating the operators' expectations into design attributes that, in the context of AHP, serve as criteria. In general, the research proceeds as shown in Fig. 1, which consists of eight stages: identification of design needs through literature review and expert interviews, formulation of alternative concepts, determination of

AHP criteria based on Kansei factors, collection of expert assessments through pairwise comparison, calculation of priority weights and Consistency Ratio (CR), determination of global rankings of alternative concepts, sensitivity analysis, and selection of the best concept.



Figure 1. Research framework

2.1. Feeding Mechanism Concept Alternatives

The existing feeding process has five limitations, which form the basis for the alternatives: non-uniform coconuts, unstable positioning, continuous manual force near the cutting zone, absence of self-centering support, and limited control of pushing and rotation. The existing arrangement was considered the baseline functional condition rather than an alternative. Three concepts were considered: simple mechanical guidance, self-centering box guidance, and assisted controlled feeding. These are progressively higher levels of guidance and control.

Table 1. Main configuration of the alternative feeding mechanism concept

Task Unit	Alternative Concepts		
	V-guided Mechanical Feeding	Box-guided Mechanical Feeding	Assisted Controlled Feeding
Guiding unit	V-shaped guide rails	Box guide	Funnel guide or box guide with centering block
Clamping unit	Dual-side screw clamp with rubber-lined pads	3-jaw/self-centering holder with free-wheel ball support	Spring-loaded clamp with assisted holding
Adjustment system	Lead screw adjustment	Sliding jaw and locking pin	Telescopic clamp arm
Pushing unit	Rack-and-pinion pusher and linear pusher arm	Rack-and-pinion pusher and linear pusher arm	Screw-drive actuator or assisted pusher
Speed/thrust control	Gear ratio control	Pulley-belt speed reduction	Servo-based or controlled thrust system
Positioning and stopping	V-guide centering and simple adjustable pin stop	Box-guide centering and mechanical end-stop	Limit switch and controlled positioning
Rotation control	Spur-gear drive and pin-lock mechanism	Spur-gear drive and pin-lock mechanism	Chain-sprocket drive with brake or ratchet lock
Frame and safety	Hollow steel frame and full metal guard	Hollow steel frame, adjustable brackets, full metal guard	Integrated frame with sliding or transparent safety shield

This study evaluated three alternative concepts: the V-Guided Mechanical Feeding Concept, Box-Guided Mechanical Feeding Concept, and Assisted Controlled Feeding Concept.

The V-Guided Mechanical Feeding Concept is the simplest mechanical concept featuring V-shaped guide rails, a dual-sided screw clamp, and a rack-and-pinion pusher. This concept emphasizes ease of fabrication and maintenance. The Box-Guided Mechanical Feeding Concept uses a box guide, a self-centering holder, a free-wheel ball support, a sliding jaw with a locking pin, and a rack-and-pinion pusher. The advantage of this alternative is that it maintains the coconut's positional stability and adapts to variations in material size. Meanwhile, the Assisted Controlled Feeding Concept is a more controlled approach that uses actuators, assisted clamps, controlled-rotation units, and safety guards. This concept emphasizes improved motion control, efficiency, ease of use, and operator protection. Table 1 presents the configurations of the three alternative concepts, and Fig. 2 shows the conceptual designs used in the questionnaire.

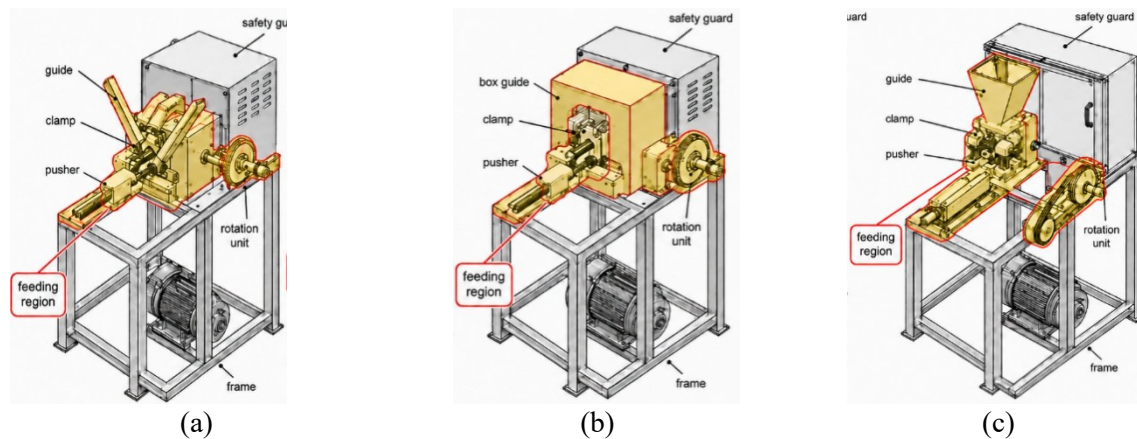


Figure 2. Feeding mechanism concept alternatives: (a) V-guided Mechanical feeding, (b) Box-Guided Mechanical Feeding, and (c) Assisted Controlled Feeding

2.2 Kansei-Based Evaluation Criteria and AHP Hierarchy

Kansei is used to accommodate both the technical aspects and the customer experience in a product [13], [18]. In this case, the product is the feeding mechanism. Kansei factors serve as the main criteria in the AHP structure [15]. Several operational indicators clarify each criterion, but the AHP calculation does not treat these indicators as separate sub-criteria. This approach reduces the burden on respondents and maintains the consistency of expert assessments, without altering the meaning of the operational indicators within each Kansei factor. The criteria used in this study consist of performance, practicality, functionality, durability, usability, pleasurability, and efficiency.

Table 2. Kansei-based Criteria

Criterion	Operational definition
Performance	Ability of the concept to support stable, smooth, and consistent coconut feeding.
Practicality	Feasibility of implementation, fabrication, adjustment, and use in small- and medium-scale industry.
Functionality	Ability to perform essential feeding functions, including guiding, clamping, positioning, pushing, and rotation control.
Durability	Potential to withstand repeated use, load, vibration, coconut impact, and friction.
Usability	Ease with which operators can understand, operate, and control the mechanism.
Pleasurability	Positive user perception, including confidence, perceived safety, neatness, and comfort.
Efficiency	Ability to reduce feeding time, unnecessary operator movement, and interruption in the process flow.

Table 2 presents the operational definitions of each criterion. The seven criteria were adapted from Kansei literature [13], [15] and recent human-factor design guidance for agricultural processing machinery [11], and refined through literature review and expert interviews to correspond to the identified feeding functions. Before pairwise comparison, experts were presented with their operational definitions to ensure a shared understanding.

Based on the Kansei factor descriptions, Fig. 3 illustrates the AHP hierarchy. The first level is the goal of selecting the most suitable feeding mechanism concept for the coconut deshelling machine. The second level comprises seven Kansei-based criteria, while the third level comprises three alternative feeding mechanisms.

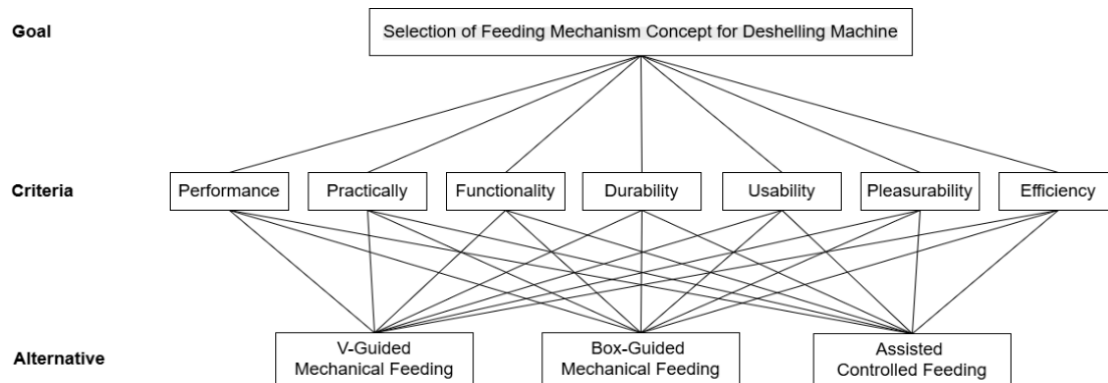


Figure 3. Hierarchy of feeding machine concept for deshelling machine design

2.3 Expert judgment and pairwise comparison

In this study, purposive sampling was used to select 9 experts to evaluate the AHP case. Expert selection was based on their involvement and experience [19] in coconut deshelling, machine maintenance, machine design, production process management, and occupational safety. Experts were contacted individually and invited according to these predefined professional roles. Individuals without relevant direct experience or with incomplete questionnaires were not included. In this study, experts were deliberately selected from practitioners and academics with more than 5 years of experience in their fields. Practitioners include deshelling operators, process supervisors, machine technicians, and GMs, who are managerial personnel. Academics are selected for their expertise in machine design and occupational health and safety. For further details, the expert profiles are available in Table 3.

Table 3. Expert profiles

Expert background	Number	Main contribution
Coconut deshelling operators	2	Practical feeding operation and usability
Deshelling process supervisor	1	Process flow and efficiency
Machine technicians	2	Mechanical feasibility and maintenance
Official Occupational Safety and Health (OSH)	1	General occupational safety risk identification
General manager	1	Feasibility and productivity perspective
Academic expert in machinery design	1	Design logic and technical evaluation
Academic expert in Occupational Safety and Health (OSH)	1	Occupational safety and health perspective

This study involved professional expert judgments on engineering design alternatives. No vulnerable participants or sensitive personal information were involved. All experts were given written information about the purpose of the study, the use of their responses for research and AHP analysis, and the confidentiality of their identities and individual assessments before

completing the questionnaire. Participation was voluntary, and providing the information in the questionnaire was considered informed consent to participate. We analyzed and reported individual responses only in the aggregate. Formal institutional ethics/IRB approval was not obtained for this study.

The experts completed pairwise-comparison questionnaires using Saaty's 1-9 scale, as shown in Table 4 [14]. The pairwise comparison assessment is divided into 2 stages: comparisons among the 7 Kansei criteria and comparisons among the 3 concept alternatives for each Kansei criterion. Consequently, the experts completed 42 pairwise comparisons. An example item for criteria comparison was: "In the selection of the feeding mechanism, how important is Functionality compared to Pleasurability?" A sample alternative-comparison-item was: "Which of the two feeding concepts is preferred and with what intensity regarding functionality?" All expert assessments were then combined using the geometric mean to obtain the combined comparison matrix, as this method is suitable for grouping AHP and maintains the matrix's reciprocal property.

Table 4. Saaty scale on pairwise comparison questionnaires

Scale	Interpretation
1	Equal importance
3	Moderate importance
5	Strong importance
7	Very strong importance
9	Extreme importance
2, 4, 6, 8	Intermediate values

2.4 Priority Weight, Consistency Ratio, and Global Priority

The pairwise comparison data from experts were converted into a numerical matrix to calculate both the criterion priority weights and the local weights of each alternative [14]. The criteria weights were obtained by normalizing the matrix and calculating the row average Eq. (1):

$$W_j = \frac{1}{n} \sum_{j=1}^n \frac{a_{ij}}{\sum_{k=1}^n a_{kj}} \quad (1)$$

where W_j is the priority weight of criterion j , a_{ij} is the element in the pairwise comparison matrix, and n is the number of criteria.

After knowing the criteria weights, the next step is to calculate the Consistency Ratio (CR) [16] to assess the logical consistency of expert judgments. First, the Consistency Index (CI) was computed using Eq. (2), and then CR was calculated using Eq. (3).

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (2)$$

$$CR = \frac{CI}{RI} \quad (3)$$

where $\lambda_{\max}$ is the maximum eigenvalue of the matrix, n is the number of elements compared, RI is the Random Index, based on the number of elements. A $CR \leq 0.10$ indicates acceptable consistency, while a higher CR suggests that the judgments need to be revised.

The global priority of each alternative concept was calculated by combining the criteria weights with the local weights of alternatives under each criterion in Eq. (4):

$$GP_i = \sum_{j=1}^n W_j \times LP_{ij} \quad (4)$$

where GP_i is the global priority of alternative i , W_j is the weight of criterion j , LP_{ij} is the local weight of alternative i under criterion j , and n is the total number of criteria. The global priority values were used to rank the alternatives.

2.5 Sensitivity Analysis

Sensitivity analysis was conducted to evaluate the stability of alternative rankings when the weights of the dominant criteria were varied. Following the robustness-testing approach of Manaf et al. [20], the three criteria with the highest baseline weights were varied individually from 0% to 100% in 10% increments, while the remaining criteria were proportionally adjusted to maintain a total weight of 1. The global priority of each alternative was then recalculated to examine whether the selected concept remained stable under different weighting conditions.

3. RESULT AND DISCUSSION

This research aims to select an alternative feeding mechanism concept for the coconut deshelling machine before entering the prototyping stage. The Analytic Hierarchy Process (AHP) method is used and combined with Kansei Engineering. Both approaches are necessary because they involve multidimensional factors that must be considered in selecting the appropriate mechanism. Variability in coconut size and operator involvement can affect process stability, operational efficiency, and workplace safety. With this approach, 9 experts evaluate each alternative through pairwise comparisons, enabling the determination of criterion weights and the global ranking of alternatives in an objective and systematic manner.

Table 5. Consistency Ratio Summary

Matrix	λ_{max}	CI	CR	Status
Criteria	7.7281	0.1213	0.0919	Consistent
Performance	3.0066	0.0033	0.0057	Consistent
Practicality	3.0522	0.0261	0.0450	Consistent
Functionality	3.0038	0.0019	0.0033	Consistent
Durability	3.0106	0.0053	0.0092	Consistent
Usability	3.0017	0.0008	0.0014	Consistent
Pleasurability	3.0649	0.0324	0.0559	Consistent
Efficiency	3.0386	0.0193	0.0333	Consistent

Table 5 reports the maximum eigenvalue (λ_{max}), Consistency Index (CI), and Consistency Ratio (CR) for the criteria matrix and each alternative-comparison matrix. $RI = 1.32$ was used for the seven-criterion matrix, while $RI = 0.58$ was used for the three-alternative matrices. The Consistency Ratio (CR) for each criterion's comparison of criteria and alternatives is below the threshold ($CR \leq 0.10$). This indicates that the experts' judgments are logical and consistent. This proves that the obtained priority weights can be reliably used for further analysis [14].

After ensuring an acceptable Consistency Ratio (CR), the next step is to analyze the criterion weights, which reflect the experts' main focus and considerations in selecting the optimal feeding mechanism concept. Fig. 4 shows the distribution of weights for each Kansei-based criterion, with the total value of all criteria set to 1 (100%). From the histogram, the Functionality criterion has the highest weight (0.293), followed by Pleasurability (0.274) and Performance (0.210), indicating the experts' main priorities are the mechanism's ability to perform important functions, operator comfort, and operational stability. These three factors

have relatively close values. This indicates that these three factors are considered almost equally important. Together, the weights of these three factors exceed 75%, indicating that they play a dominant role in this AHP model. This also aligns with research by [15], where safety (part of pleasurability in this study) and performance are the two dominant factors in selecting machine design concepts. Other criteria, namely usability, durability, practicality, and efficiency, have lower weights. This phenomenon indicates that, although important, they are considered less decisive than the three dominant criteria. These weight values subsequently serve as the basis for calculating the global priority of alternatives and as a reference for sensitivity analysis, since criteria with higher weights have a greater impact on the ranking of alternatives.

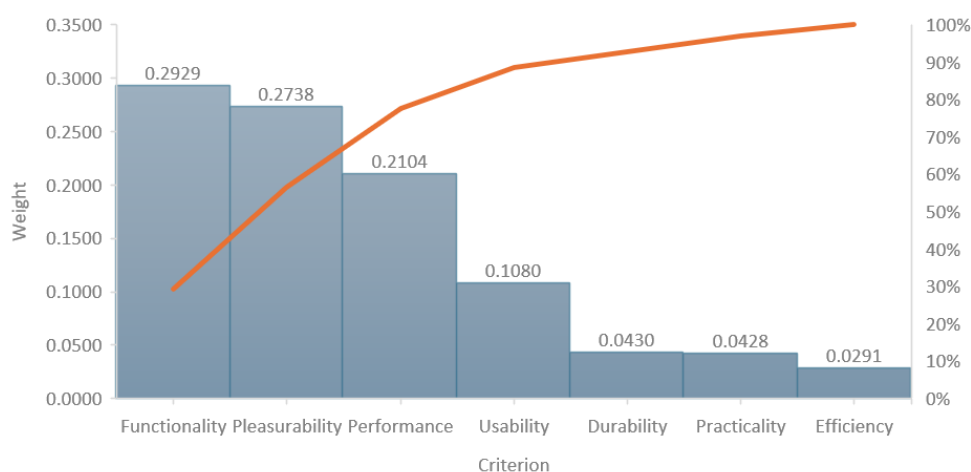


Figure 4. Criteria Weight Histogram

Fig. 5 illustrates the local priority weight values for each feeding mechanism concept across each Kansei criterion. The darker the criterion's color, the greater its weight, and vice versa. A lighter color indicates a lower local priority value. From the image, the three concepts each have advantages, distributing the dominant weights across different criteria. V-Guided Mechanical Feeding has the highest weight in the Practicality criterion (0.614), indicating that this concept excels in ease of implementation and maintenance, but has low weights in Functionality, Durability, and Efficiency (0.09–0.12), showing that its performance and efficiency are quite limited. Conversely, in the Box-Guided Mechanical Feeding concept, the 3 dominant criteria are Usability (0.492), Durability (0.451), and Functionality (0.430). This indicates a balance between mechanical stability and operator comfort. The Assisted Controlled Feeding concept is slightly different, with Pleasurability (0.546) and Efficiency (0.550) as the dominant criteria. This indicates that this concept can provide the operator with safety and comfort at high process speeds. On the other hand, this concept has a low Practicality weight (0.078), indicating higher implementation complexity, especially for small- to medium-sized coconut industries.

These findings are consistent with prior studies on agricultural machinery selection using AHP, which indicate a trade-off between ease of implementation and optimal operational performance. Similarly, recent studies on machinery selection show that multicriteria rankings depend on trade-offs among technical effectiveness, cost, usability, and operational constraints [16], [17]. In addition, [17] found that tractors with greater practical ease or lower mechanical complexity often scored lower on efficiency or sustainability metrics. V-guided, as the simplest concept, is easier to implement but less optimal in terms of function and efficiency. Higher-control concepts, such as Assisted Controlled, are more efficient and provide operator comfort

but are mechanically complex. The Box-Guided concept is proposed as the best compromise, balancing mechanical stability, functionality, and operator comfort, in line with recommendations from previous studies that emphasize integrating technical criteria and human factors in the design of agricultural and post-harvest machinery.

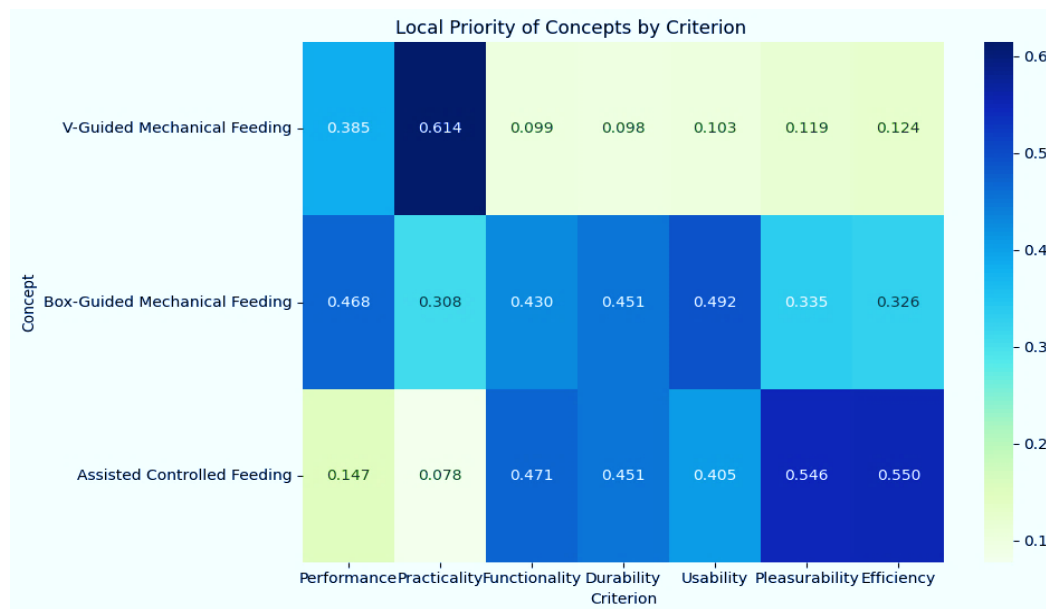


Figure 5. Concepts Local Priority by Criterion Heatmap

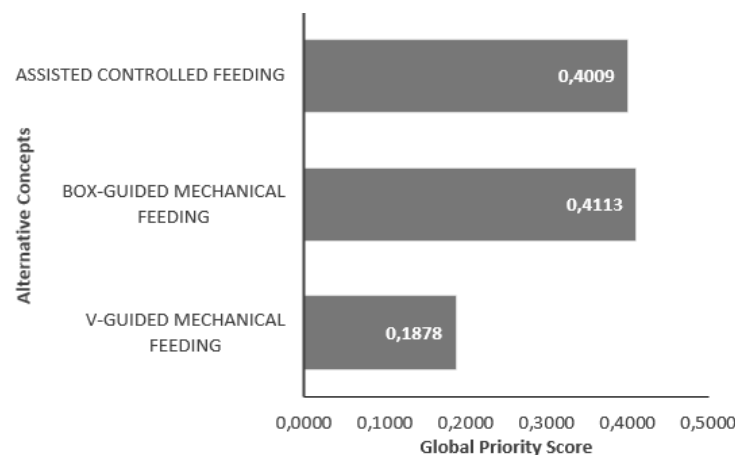


Figure 6. Global Priority Alternatives Diagram

After determining the local priority for each concept in each criterion, the next step is to analyze the global priority to determine the global ranking of the feeding mechanism concept alternatives. Global Priority integrates criteria and local weights for each alternative, providing a comprehensive picture of which concept is most worthy of selection. Fig. 6 shows the global priority values. The Box-Guided Mechanical Feeding concept receives the highest global priority (0.411), followed by Assisted Controlled Feeding (0.401) and V-Guided Mechanical Feeding (0.188). The 0.010 difference between the two highest-ranked concepts is very small and should not be interpreted as an absolute winner. Box-Guided is the baseline preferred concept under the original criterion weights, while the Assisted Controlled concept remains practically competitive and may be selected when the target industry has more financial, technical, and maintenance capability. Meanwhile, V-Guided has the lowest value because it is simpler but less optimal overall. V-Guided remains the simplest alternative but is less

competitive overall because of its lower functionality, durability, and efficiency priorities. Some prior literature has also reported similar results. The agricultural and post-harvest machines chosen balance performance and ease of use, avoiding options that are too simple or too complex [15].

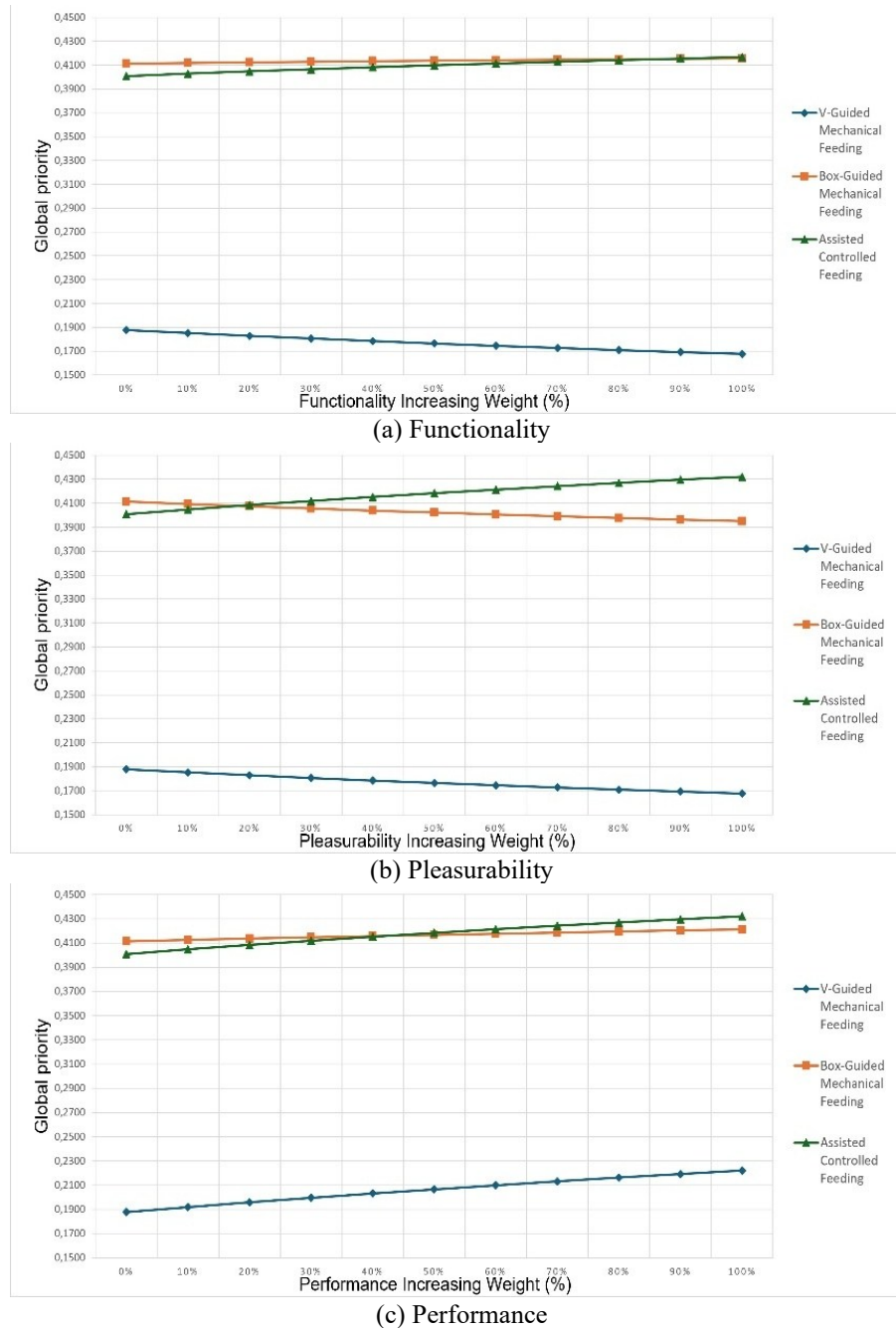


Figure 7. Sensitivity analysis of global priorities under varying weights of the three dominant Kansei criteria

A sensitivity study assessed the robustness of the selected feeding mechanism concept to variations in the main Kansei criteria. The study focused on the three most heavily weighted criteria, which together account for more than 75% of the overall criteria weight: Functionality, Pleasurability, and Performance, as shown in Fig 4 above. This technique is similar to robustness testing in AHP-based decision models, where sensitivity analysis is performed to assess whether the final ranking is stable under variations in the relative importance of criteria

[20]. In this study, each dominant criterion was varied from 0% to 100%, and the other criteria were proportionally adjusted to ensure the total weight remained 1. Figure 7 shows the results (a, b, and c).

The sensitivity study indicated that the ranking of alternatives was fairly robust, while some changes occurred under certain weighting conditions. In the Functionality sensitivity case in Fig 7a, Box-Guided Mechanical Feeding remained the top-ranked concept across Functionality weights from 0% to 40%. However, when the Functionality weight reached 50% or higher, Assisted Controlled Feeding became the first-ranked concept. The highest-ranked concept is context-dependent. This crossover confirms, as shown in the graph, that Assisted Controlled becomes more attractive as decision makers attach very high importance to controlled pushing, clamping, positioning, and rotation. Box-Guided remains the baseline preferred option at the original Functionality weight of 0.293.

The same tendency was observed in the Pleasurability sensitivity scenario in Fig 7b. Box-Guided Mechanical Feeding was the best-ranked alternative when Pleasurability was weighted between 0% and 30%. However, Assisted Controlled Feeding became dominant when the Pleasurability weight was 40% or higher. Consequently, Assisted Controlled may be preferred when perceived safety, operator confidence, comfort, and control dominate the decision, although its higher mechanical complexity and lower practicality remain an important implementation constraint for small- and medium-scale enterprises.

However, the Performance sensitivity case in Fig. 7c provided greater support for Box-Guided Mechanical Feeding. Box-Guided Mechanical Feeding was still the preferred concept at a Performance weight of 20% and above. Assisted Controlled Feeding was dominant only when the Performance weight was very low (0–10%). This indicates that Box-Guided Mechanical Feeding is more strongly correlated with reliable, smooth, and consistent coconut feeding performance. The result supports Box-Guided for performance-oriented and baseline decision conditions, but does not rule out Assisted Controlled as a viable option for different strategic priorities.

Table 6. Sensitivity Analysis Scenario and Selected Concept Result

Scenario	Weight adjustment	Selected concept	Interpretation
Baseline	Original weights	Box-Guided Mechanical Feeding (baseline-preferred)	Top-ranked under baseline weights, but near-tied with Assisted Controlled.
Functionality sensitivity	Functionality varied from 0% to 100% at 10% intervals	Box-Guided at 0–40%; Assisted Controlled at $\geq 50\%$	Ranking changes when functionality receives very high emphasis.
Pleasurability sensitivity	Pleasurability varied from 0% to 100% at 10% intervals	Box-Guided at 0–30%; Assisted Controlled at $\geq 40\%$	Ranking is sensitive to operator comfort and perceived safety.
Performance sensitivity	Performance varied from 0% to 100% at 10% intervals	Assisted Controlled at 0–10%; Box-Guided at $\geq 20\%$	Box-Guided is supported under baseline and performance-oriented conditions.

The sensitivity analysis indicates that Box-Guided Mechanical Feeding is the preferred compromise under the baseline weights, but the choice between Box-Guided and Assisted Controlled is not absolute. Box-Guided is preferred with the original weights and moderate changes in the dominant criteria, while Assisted Controlled becomes top-ranked with greater weight on Functionality or Pleasurability. On the other hand, V-Guided Mechanical Feeding was not selected in any sensitivity case because its simplicity and practicality did not outweigh

its poor functionality, performance, and efficiency. For medium-scale enterprises with conventional fabrication and limited maintenance capability, Box-Guided offers a viable path to more stable feeding and less manual intervention. However, higher-level enterprises with larger automation budgets and stronger control-system capability may choose Assisted Controlled to prioritize operator comfort, perceived safety, and motion control.

4. CONCLUSIONS

The research shows that a Kansei-based AHP framework can be used in the conceptual design of the feeding mechanism before detailed design by integrating technical requirements with operator-oriented considerations. In the baseline expert priorities, the preferred compromise feeding mechanism for implementation in medium-scale coconut agroindustry is the Box-Guided Mechanical Feeding, which combines mechanical guidance and self-centering capability with lower implementation complexity than the assisted-control alternative. However, the tight baseline ranking and the sensitivity analysis suggest that Assisted Controlled Feeding is also a feasible option and can become more attractive when more weight is given to overall functionality, operator confidence, perceived safety, or process efficiency. Therefore, the decision is more about context than a hunt for the universally best mechanism. Importantly, the favored Box-Guided concept directly addresses the original design challenges by reducing dependence on continuous manual holding and repositioning, improving repeatability in accommodating non-uniform coconut sizes, and increasing separation between the operator and the cutting interface. These anticipated benefits may reduce conditions associated with fatigue, kernel breakage, and accident exposure. This can be demonstrated following experimental verification on the physical prototype in the subsequent stage. The framework can be extended to other decisions in agro-industrial machinery where technical performance, implementation feasibility, and human factors must be considered simultaneously. The next step should be to develop and test a validated design through prototype development, mechanical performance testing, ergonomic assessment, and safety evaluation under actual coconut deshelling conditions.

ACKNOWLEDGEMENT

The authors would like to thank Universitas Gadjah Mada for partially funding this research through the Final Assignment Recognition Program (Program Rekognisi Tugas Akhir, RTA), under Assignment Letter No. 5286/UN1.P1/PT.01.03/2024, and the Center for Higher Education Funding of Indonesia for providing a scholarship to the first author. The authors also extend their sincere gratitude to all experts who participated in this study for their willingness, time, professional judgment, and valuable insights during the AHP assessment. Their contributions provided essential input to this research. The authors also acknowledge using generative AI solely to assist in creating visually engaging representations and polishing the manuscript's language. The authors reviewed, revised, and validated all final visualizations, scientific content, data analyses, interpretations, and conclusions.

REFERENCES

- [1] International Coconut Community, Statistical yearbook 2023. Jakarta: International Coconut Community, 2023. [Online]. Available: <https://coconutcommunity.org/files-/document/20250516104911pub-coconut-statistical-year-book.pdf>
- [2] P. Rethinam, "International scenario of coconut sector," in *The Coconut Palm (Cocos nucifera L.) - research and development perspectives*, K. U. K. Nampoothiri, Ed., Singapore: Springer Nature Singapore Pte Ltd, 2018, pp. 21–56. doi: 10.1007/978-981-13-2754-4_2.

-
- [3] F. A. Zainol et al., “Coconut value chain analysis: A systematic review,” *Agriculture*, vol. 13, no. 7, p. 1379, 2023, doi: 10.3390/agriculture13071379.
 - [4] R. Nirmalaruban, R. Suvitha, and S. P. Sharan, “Unlocking the potential of Coconut : Nutritional value, processing technologies, and product development,” *Annu. Res. Rev. Biol.*, vol. 40, no. 9, pp. 39–49, 2025.
 - [5] H. H. Hidayat, B. Purwantana, Radi, and M. Ushada, “Mechanization of Coconut deshelling : Ergonomic and productivity impacts in an integrated coconut industry,” in *IOP Conf. Series: Earth and Environmental Science*, Yogyakarta: IOP Publishing, 2026. doi: 10.1088/1755-1315/1584/1/012105.
 - [6] G. K. Ekanayake, S. A. C. N. Perera, P. N. Dassanayake, and J. M. D. T. Everard, “Varietal classification of New Coconut (*Cocos nucifera* L.) forms identified from Southern Sri Lanka,” *Cocos*, vol. 19, no. 1, pp. 41–50, 2012, doi: 10.4038/cocos.v19i1.4749.
 - [7] R. Pandiselvam et al., “Engineering properties of five varieties of coconuts (*Cocos nucifera* L.) for efficient husk separation,” *J. Nat. Fibers*, vol. 17, no. 4, pp. 589–597, 2020, doi: 10.1080/15440478.2018.1507863.
 - [8] S. Ghribi et al., “Characterization and evaluation of the diversity of local and imported coconut varieties collected from the Sultanate of Oman,” vol. 60, no. 1, pp. 43–51, 2023, doi: 10.21162/PAKJAS/23.266.
 - [9] A. F. Alonge, “Properties of Coconut (*Cocos nucifera* L.) relevant to its dehusking p of Coconut (*Cocos nucifera* L.),” *J. Agric. Sci. Technol.*, no. January, pp. 1089–1094, 2016, doi: 10.17265/2161-6256/2011.11A.020.
 - [10] M. Elfahmi, L. Sutiarto, D. Purwadi, and M. M. Machfoedz, “Development of integrated Coconut agroindustry from a circular economy perspective: A literature review,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1364, no. 1, 2024, doi: 10.1088/1755-1315/1364/1/012001.
 - [11] T. Rahman, Yassierli, and A. Widyanti, “Design guidelines for sustainable utilization of agricultural appropriate technology : Enhancing human factors and user experience,” *Open Agric.*, vol. 8, no. 1, 2023, doi: <https://doi.org/10.1515/opag-2022-0232>.
 - [12] B. Qiu, Y. Zhang, H. Shen, J. Zhou, and L. Chu, “Ergonomic researches in agricultural machinery- a systematic review using the PRISMA method,” *Int. J. Ind. Ergon.*, vol. 95, no. March, p. 103446, 2023, doi: 10.1016/j.ergon.2023.103446.
 - [13] M. Nagamachi and A. M. Lokman, *Kansei innovation: Practical design applications for product and service development*. Boca Raton: CRC Press, 2015. doi: 10.1201/b18054.
 - [14] T. L. Saaty and Luis G. Vargas, *Model, methods, concepts & application of the Analytic Hierarchy Process*, 2nd ed., no. July. New York: Springer, 2012. doi: 10.1016/b978-0-08-032599-6.50008-8.
 - [15] R. Ramanathan, L. Abdullah, M. H. F. M. Fauduadi, M. S. S. Mohamed, and K. N. Kamaludin, “A hybrid of Kansei Engineering (KE) and Analytical Hierarchy Process (AHP) to develop conceptual design of portable oil spill skimmer,” *IIUM Eng. J.*, vol. 24, no. 1, 2023, doi: <https://doi.org/10.31436/iiumej.v24i1.2426>.
 - [16] Z. Ünal, “Plant protection machine selection with Analytical Hierarchy Process : Maize plant example,” *Turkish J. Agric. - Food Sci. Technol.*, vol. 12, no. 3, pp. 453–461, 2024, doi: <https://doi.org/10.24925/turjaf.v12i3.453-461.6722>.
 - [17] H. A. A. Sayed, M. A. Abdelhamid, T. K. Abdelkader, Q. Lai, A. M. Mousa, and M. Refai, “Machine learning and analytic hierarchy process integration for selecting a sustainable tractor,” *Sci. Rep.*, vol. 14, p. 26735, 2024, doi: <https://doi.org/10.1038/s41598-024-78023-z>.
 - [18] F. Fadhillah, N. P. Sari, and W. Prastiwinarti, “Determining packaging design concepts using K-Means cluster Genetic Algorithm methods through a Kansei Engineering approach,” in *Proceedings of the 8th Mechanical and Industrial Engineering Symposium (MIE 2025)*, Yogyakarta: Atlantis Press International BV, 2026, pp. 286–304. [Online]. Available: http://dx.doi.org/10.2991/978-94-6239-687-6_21
-

- [19] V. V Tsyganok, S. V Kadenko, and O. V Andriichuk, “Significance of expert competence consideration in group decision making using AHP,” *Int. J. Prod. Res.*, no. August 2013, pp. 37–41, 2012, doi: 10.1080/00207543.2012.657967.
- [20] L. A. Manaf et al., “A consensus-driven Hybrid Delphi and Analytic Hierarchy Process framework for evaluating sustainable food waste treatment alternatives,” *Process Integr. Optim. Sustain.*, 2026, doi: <https://doi.org/10.1007/s41660-026-00776-z>.

Surface Treatment for Improved Electrochemical Performance of Activated Carbon Derived from Palm Empty Fruit Bunches

HASAN MARZUKI¹, ALYA NAILI ROZHAN^{1*},
MOHD HANAFI ANI¹, HADI PURWANTO²

¹Department of Manufacturing and Materials Engineering, Kulliyyah of Engineering,
International Islamic University Malaysia, 53100, Kuala Lumpur, Malaysia

²Department of Engineering Management, University of International Cement Indonesia UIIS,
Gresik, West Java 61122, Indonesia

*Corresponding author: alyanaili@iium.edu.my

(Received: 23 July 2026; Accepted: 29 August 2026; Published online: 11 September 2026)

ABSTRACT: Recent interests in sustainable and green technologies for energy storage systems require extensive research to develop sustainable electrode materials. This includes developing biomass-derived activated carbons. In this research, empty fruit bunch (EFB) has been selected as the precursor material to produce activated carbon by the physical activation method, performed at 500 °C to 800 °C. Subsequently, a surface treatment process was performed using nitric acid to introduce oxygen functional groups. This step further improved the electrochemical properties of the EFB-based activated carbon material. The resulting samples were analyzed using Fourier transform infrared spectroscopy (FTIR), Raman spectroscopy, and X-ray photoelectron spectroscopy (XPS). The electrochemical performance was examined through cyclic voltammetry (CV), galvanostatic charge–discharge, and electrochemical impedance spectroscopy (EIS). FTIR analysis confirmed successful incorporation of oxygen functional groups, enhanced surface wettability, and improved electrolyte penetration. CV measurements showed a significant increase in the specific capacitance after treatment. Surface treatment of the activated carbon positively impacted all samples. The largest percentage increase was observed in sample AC ST 500, with a 71% increase in specific capacitance from 24 F g⁻¹ to 71 F g⁻¹. AC ST 800 displayed the highest overall capacitance after surface treatment, at 183 F g⁻¹, a 9% increase from 166 F g⁻¹ before surface treatment. This result shows that adding oxygen functional groups improves these properties by increasing surface wettability, which increases ion penetration within the activated carbon.

ABSTRAK: Minat yang semakin meningkat terhadap teknologi mampan dan hijau bagi sistem penyimpanan tenaga memerlukan penyelidikan menyeluruh bagi membangunkan bahan elektrod mampan, termasuk karbon teraktif berasaskan biojisim. Dalam kajian ini, tandan buah kosong (TBK) dipilih sebagai bahan pelopor bagi menghasilkan karbon teraktif melalui kaedah pengaktifan fizikal pada suhu antara 500 °C hingga 800 °C. Seterusnya, rawatan permukaan menggunakan asid nitrik dijalankan bagi menerapkan kumpulan berfungsi yang mengandungi oksigen dan seterusnya meningkatkan sifat elektrokimia bahan karbon teraktif berasaskan TBK. Sampel terhasil dianalisis menggunakan spektroskopi inframerah transformasi Fourier (FTIR), spektroskopi Raman dan spektroskopi fotoelektron sinar-X (XPS), manakala prestasi elektrokimia ini dinilai melalui voltametri berkitar (CV), pengecasan–nyahcas galvanostatik dan spektroskopi impedans elektrokimia (EIS). Analisis FTIR mengesahkan kejayaan penerapan kumpulan berfungsi yang mengandungi oksigen, peningkatan kebolehbasaan permukaan dan penembusan elektrolit yang lebih baik. Pengukuran CV menunjukkan peningkatan ketara dalam kemuatan tentu selepas rawatan

permukaan, dengan kesan positif diperhatikan pada semua sampel. Peningkatan peratusan terbesar dicatat pada sampel AC ST 500, dengan kekuatan tentu meningkat kira-kira 71%, daripada 24 F g⁻¹ kepada 71 F g⁻¹. Sampel AC ST 800 pula mencatatkan kekuatan tentu keseluruhan tertinggi selepas rawatan permukaan, iaitu 183 F g⁻¹, bersamaan peningkatan kira-kira 9% daripada 166 F g⁻¹ sebelum rawatan. Dapatan ini menunjukkan bahawa penerapan kumpulan berfungsi yang mengandungi oksigen dapat meningkatkan prestasi elektrokimia karbon teraktif dengan memperbaiki kebolehasahan permukaannya, sekaligus meningkatkan penembusan ion ke dalam struktur karbon teraktif.

KEYWORDS: *Empty fruit bunch, activated carbon, physical activation, acid surface treatment, supercapacitor*

1. INTRODUCTION

Interest in developing sustainable, green energy storage system technologies has grown, as these systems are expected to have long life cycles, high performance, and be driven by low-carbon life cycle materials. This approach ensures that both high performance and minimal adverse environmental impact can be achieved [1]. Supercapacitors are promising energy storage systems that can deliver high performance due to their high cycling stability and power density. However, most commercially available supercapacitors use conventional carbon-based electrode materials derived from non-renewable fossil fuels.

In energy storage systems, demand for sustainable materials has increased, making renewable biomass materials an attractive alternative because of their strong potential as raw materials. Biomass-derived activated carbon has generated significant interest as a more economical and environmentally friendly substitute for conventional activated carbon [2]. This type of carbon, derived from natural resources, reduces dependence on non-renewable carbon sources. However, when meeting the performance demands of energy storage systems such as supercapacitors and batteries, these biomass-based materials often lack the required properties [3].

One abundantly available biomass source with strong potential for energy storage systems is palm empty fruit bunch (EFB) [4, 5]. Numerous studies have focused on optimizing synthesis parameters to produce high-performance activated carbon [6, 7]. However, many still produce activated carbon with relatively low electrochemical performance. Therefore, other techniques to improve electrochemical properties, such as surface treatment and element doping, need further exploration [8, 9].

Acid treatment is one feasible method to improve the efficiency of these biomass-derived activated carbons. This surface treatment adds functional groups, improving the surface characteristics of biomass-carbon materials. This improves ionic transport and the interface between the carbon materials and the electrolyte, which helps increase energy storage capacity [10]. Table 1 shows the effect of the acid treatment on the electrochemical properties of activated carbon.

Despite this development, we still need to fully understand how acid treatment influences the electrochemical performance of activated carbon, especially from agricultural biomass such as EFB. This is because few studies have demonstrated the effect of acid surface treatment on the electrochemical performance of activated carbon derived from agricultural biomass such as EFB. Table 1 shows that surface treatment greatly improved conventional materials such as commercial activated carbon, tire, and polymer, while the performance of palm kernel shell after surface treatment was relatively poor. A critical observation is that the surface-treatment

methodology is similar for these biomass and conventional materials, which employ acid surface treatment at elevated temperatures. This is because the structure of biomass materials differs from that of conventional materials; therefore, a similar approach may not be practical. To address this gap, this study focuses on acid surface treatment of biomass material at room temperature.

Table 1. Literature data on the electrochemical properties of activated carbon before and after surface treatment

Raw Material	Surface Treatment Method	Specific Capacitance Before Surface Treatment (Fg ⁻¹)	Specific Capacitance After Surface Treatment (Fg ⁻¹)	Ref
Scrap Tire	30% nitric acid at 50 °C	188	190	[11]
Polyaniline	Concentrated nitric acid at 60 °C	50	350	[12]
Palm kernel shell	30% nitric acid at 60 – 80 °C	45	Less than 45	[13]
Commercial Activated Carbon	Ammonium persulfate and nitric acid at 60 °C	35	140	[14]
Commercial Activated Carbon	Nitric acid at 120 °C	142.47	229.17	[15]

In our previous study [16], we successfully converted EFB into activated carbon via physical and chemical activation methods. Their electrochemical performances have been examined. Therefore, this present study further explores how acid surface treatment affects the electrochemical properties of EFB. The effect of acid treatment is demonstrated by comparing the properties of activated carbon derived from EFB before and after acid surface treatment.

2. EXPERIMENTAL METHODS

2.1. Experiment

EFB was used as the raw material and was first prepared by pulverizing and drying it in an oven at 105°C for 24 hours to remove moisture. After drying, a two-step process converted the EFB into activated carbon. Initially, EFB was pyrolyzed at 500°C with a heating rate of 10°C/min in an argon atmosphere, with a residence time of 1 hour. This process yields solid carbon in the form of biochar. Next, the resulting biochar was activated to produce high-quality activated carbon via physical activation at 500°C, 600°C, 700°C, and 800°C, with a heating rate of 10°C/min and a flowing CO₂ gas flow rate of 0.3 L/min. The residence time was one hour. CO₂ acted as the oxidizing agent in this process. This physical activation of the biochar to produce activated carbon was followed by a surface modification treatment using acid treatment.

Surface modification treatment enhances activated carbon by adding oxygen-containing functional groups to its surface. In this process, the activated carbon was soaked in 40 mL of 65% nitric acid for one hour at room temperature. Afterward, it was thoroughly rinsed with distilled water to remove any remaining acid, then dried at 105 °C for 24 hours in an oven to complete the treatment. The samples were categorized by treatment. Samples without surface modification were labeled as AC 500, AC 600, AC 700, and AC 800. Samples with surface modification were identified as AC ST 500, AC ST 600, AC ST 700, and AC ST 800.

2.2. Physical Characterization

To evaluate the physical properties before and after surface modification, activated carbon was characterized using Fourier transform infrared spectroscopy (FTIR), X-ray photoelectron spectroscopy (XPS), and Raman spectroscopy. The defect structure of the activated carbon was evaluated by the ID/IG ratio in the Raman spectra. FTIR analysis was performed to examine the surface functional groups of the activated carbon derived from EFB, while XPS was used to identify and quantify these groups, complementing the FTIR analysis.

2.3. Preparation of Electrode

For slurry preparation, activated carbon was mixed with graphite as a conductive additive and polytetrafluoroethylene (PTFE) as a binder. A ratio of 8:1:1 was used, and the mixture was stirred well to achieve good homogeneity. Stainless steel mesh was selected as the current collector. It was cut into 1 cm x 1 cm pieces. Finally, 0.01 g of the slurry was hot-pressed onto the mesh at 10 MPa and 80°C.

2.4. Electrochemical Analysis

For electrochemical measurements, a Metrohm Autolab electrochemical workstation was used with a three-electrode setup in a 6 M KOH aqueous solution. In this setup, the activated carbon electrode, graphite, and an Ag/AgCl electrode served as the working, counter, and reference electrodes, respectively. The cyclic voltammetry (CV) potential window was tested from -0.8 to 0.4 V at a scan rate of 2 mV/s. The specific capacitance was calculated from the CV test using the equation below [17].

$$C_p = \frac{A}{2mk(\Delta V)} \quad (1)$$

where C_p is the specific capacitance, A is the area inside the CV curve, m is the mass of activated carbon, k is the scan rate, and ΔV is the potential window.

The voltage range for the galvanostatic charge and discharge (GDC) test was -0.8 to 0.1 V with a current density of 0.1 A/g. Electrochemical impedance spectroscopy (EIS) test was carried out with a frequency range of 100 kHz to 10 mHz with an amplitude of 10 mV.

In addition, this research uses a similar approach to electrode preparation and electrochemical test setup. This shows that the methodology adopted in this research is consistent with established published literature [18].

3. RESULTS AND DISCUSSION

The acid surface treatment introduced oxygen-containing functional groups such as carbonyl, carboxyl, and hydroxyl on the activated carbon surface. Figure 1 illustrates the mechanism of the activated carbon synthesis via physical activation and the surface treatment of the activated carbon [19]. This illustration also explains the mechanism of each process, including how surface treatment enhances the surface wettability of activated carbon to facilitate electrolyte movement and increase the efficiency of ion penetration within the activated carbon networks [20].

FTIR analysis showed functional groups on the activated carbons. The wavenumbers 2400 to 3600 cm^{-1} , 1710 cm^{-1} , and 1500 cm^{-1} indicated the presence of O-H, C=O, and C=C functional groups, respectively, while the wavenumbers 1205 cm^{-1} and 1103 cm^{-1} indicated the presence of the C-O functional group [21]. Figure 2 shows the FTIR spectra of the activated carbon before and after surface treatment. The peak at the representative wavenumber of the

surface functional group increased. This observation indicated that the surface treatment had successfully introduced oxygen functional groups on the surface of the activated carbon. The amount of functional groups developed on the surface of activated carbon varied among the samples, although constant parameters of the surface treatment process were being applied. This could be due to the original physical properties of activated carbon during the physical activation process.

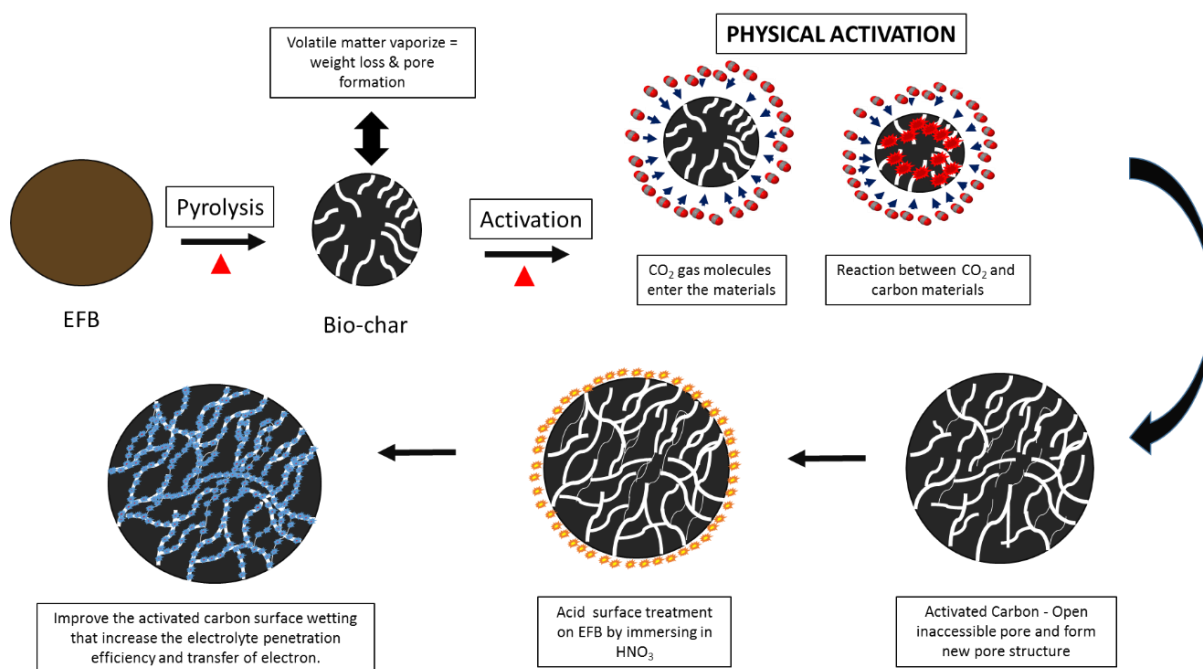


Figure 1. Synthesis pathways of the activated carbon from biomass

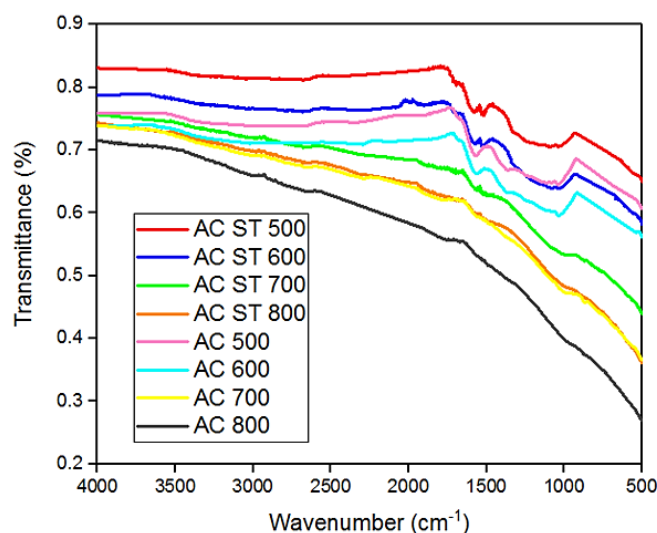


Figure 2. FTIR spectra of activated carbon before and after surface treatment

XPS analysis was conducted on AC ST 800 to support the FTIR findings, since AC ST 800 showed fewer FTIR peaks than the other surface-treated samples. Figure 3 shows a peak in the XPS spectra representing C-O, C=O, and O-C=O. From the XPS analysis, the surface functionality present on the activated carbon sample can be quantified as shown in Table 2. This quantitative analysis of functional-group percentages better verifies surface functional-group presence after surface treatment [22].

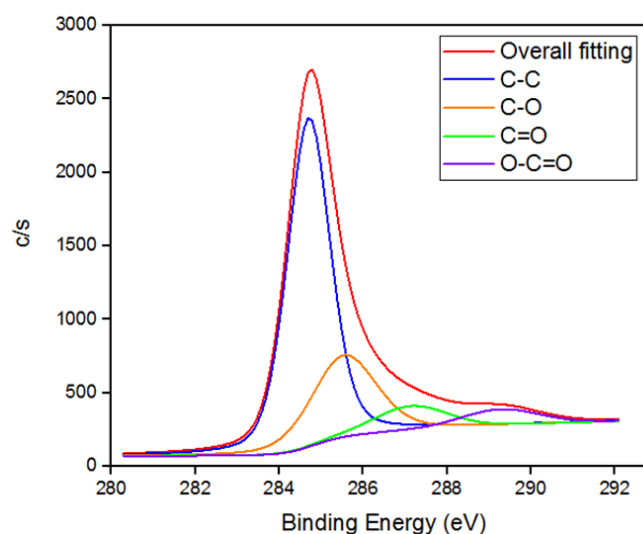


Figure 3. XPS spectra of AC ST 800

Table 2. Surface functional groups on AC ST 800

Sample	AC ST 800
C-C (%)	63.10
C-O (%)	23.41 – 63.15
C=O (%)	8.03 – 33.55
O-C=O (%)	3.31 – 5.55

Structural defects in activated carbon materials play an important role in determining their electrochemical performance. Raman spectroscopy analysis was used to analyze the structural defects in the samples. Characterizing defects is important for understanding the relationship between structural defects within the activated carbon network and its electrochemical performance. Although a higher defect density in carbon materials increases the density of electronic states and can lead to higher capacitance, structural defects in activated carbon are not the only factor affecting electrochemical performance. Other factors also contribute to the electrochemical properties of activated carbon, such as surface area, pore connectivity, and surface wettability [23].

Defects in carbon materials can be analyzed from the I_D and I_G of the Raman spectra, where a higher I_D/I_G indicates a higher defect structure. The D and G bands corresponding to activated carbon are 1348 cm^{-1} and 1600 cm^{-1} , respectively [24]. Figure 4 illustrates the Raman spectra of AC ST 500 and AC ST 800, the samples with the highest and lowest functional group peaks in the FTIR spectra, along with the untreated samples, AC 500 and AC 800.

Table 3 compares the I_D/I_G ratios of the Raman spectra of the activated carbon before and after surface treatment. The I_D/I_G ratios for the samples showed no major differences. This suggests that the surface treatment in this experiment did not significantly affect the activated carbon's defect structure. The I_D/I_G ratio increased slightly for AC ST 500, while it remained unchanged for AC ST 800. The I_D/I_G values of AC ST 500 and AC ST 800 were also in a similar range, indicating that structural defects may not be a major factor in determining the electrochemical properties of activated carbon after surface treatment.

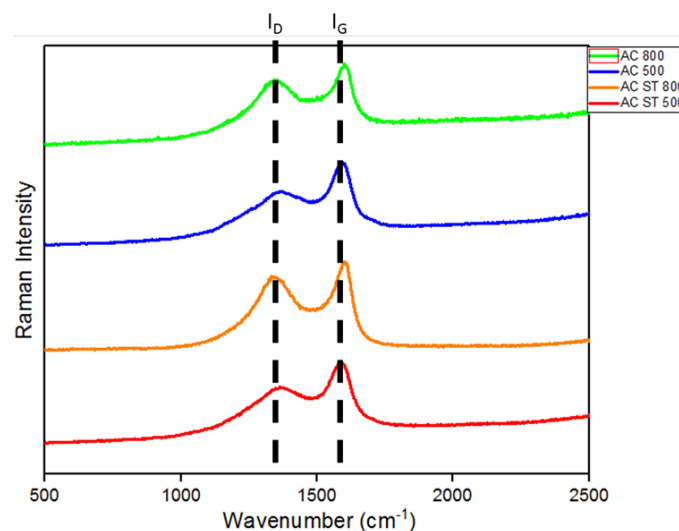


Figure 4. Raman spectra of activated carbon before and after surface treatment

Table 3. ID/IG of the activated carbon before and after surface treatment

Sample	I _D /I _G
AC 500	0.70
AC ST 500	0.73
AC 800	0.84
AC ST 800	0.84

Figure 5 shows the CV curve for activated carbon before and after surface treatment. The area under a CV curve represents the specific capacitance of the activated carbon; a larger area indicates higher specific capacitance. The CV curve shape can also be used to analyze the electrochemical energy storage mechanism. A perfect square CV curve indicates a purely electrochemical double-layer mechanism, without any pseudocapacitive reaction. All samples, before and after surface treatment, did not show a perfect square curve, indicating that some pseudocapacitive reactions occurred [25]. This observation may be related to the FTIR spectra in Figure 2, which showed that all samples had a peak indicating an oxygen functional group, typical of biomass carbon-based materials with traces of functional-group impurities. The surface treatment on activated carbon was applied to determine whether the electrochemical properties could be improved or negatively affected by adding functional groups from the treatment. Figure 5 shows that all CV curves improved after surface treatment, indicating that the surface treatment process used in this research is compatible with activated carbon derived from EFB via physical activation. Table 4 compares the specific capacitance of activated carbon before and after surface treatment.

As shown in Table 4, all samples increased in specific capacitance after surface treatment, indicating that introducing oxygen functional groups using the method applied in this study did not collapse the activated carbon structure, which could otherwise negatively affect capacitance [13]. The specific capacitance of activated carbon increased from AC 500 to AC 800 before surface treatment, which was attributed to the activation temperature, as higher temperatures produced activated carbon with greater specific capacitance. AC 800, activated at 800 °C, produced the highest capacitance. After surface treatment, the same trend held: AC ST 800 produced the highest capacitance. However, the percentage change in specific capacitance before and after treatment followed an inverse pattern, with AC ST 500 showing the highest percentage increase, which gradually decreased toward AC ST 800. FTIR analysis supported

this pattern: AC ST 500 showed the highest peak for oxygen functional groups, which led to greater wettability improvement. This resulted in a higher percentage increase in capacitance. While wettability enhances electrolyte penetration, it does not change the activated carbon structure, such as pore connectivity, surface area, and structural defects, that play major roles in electrochemical performance [26]. Although AC ST 800 has the lowest oxygen functional group content after surface treatment, it still showed the highest overall specific capacitance, which was contributed by the original properties of the activated carbon prior to surface treatment. Introducing oxygen functional groups enhanced its capacitance. This indicates that the overall effect depends heavily on the activated carbon's structural properties [27].

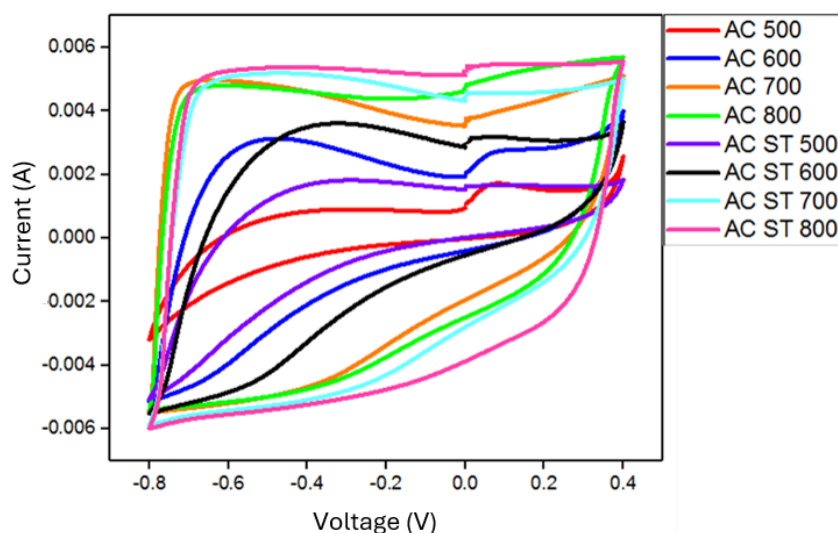


Figure 5. CV curve for activated carbon before and after surface treatment

Table 4. Specific capacitance of activated carbon before and after surface treatment

Sample	Specific Capacitance, C_p (Fg ⁻¹)	Sample	Specific Capacitance, C_p (Fg ⁻¹)	Percentage Difference (%)
AC 500	24.72	AC ST 500	42.35	71.32
AC 600	76.00	AC ST 600	88.59	16.57
AC 700	146.60	AC ST 700	161.84	10.40
AC 800	166.68	AC ST 800	183.03	9.73

Figure 6 shows the galvanostatic charge-discharge (GCD) curve, which illustrates the charging and discharging behavior of the activated carbon electrode during energy storage. This curve also shows ion diffusion mobility, represented by the voltage drop after the peak. The discharging time of AC ST 800 was higher than that of AC 800. A longer discharge time indicates greater electrolyte-ion involvement during discharge [28]. Greater electrolyte-ion involvement indicates ion movement within the activated carbon pore network. This can be attributed to surface functional groups, which enhance the wettability and hydrophilicity of the activated carbon surface. Improved surface wettability and hydrophilicity can enhance ion movement, as shown by the charging and discharging times in the GCD analysis of activated carbon [29, 30].

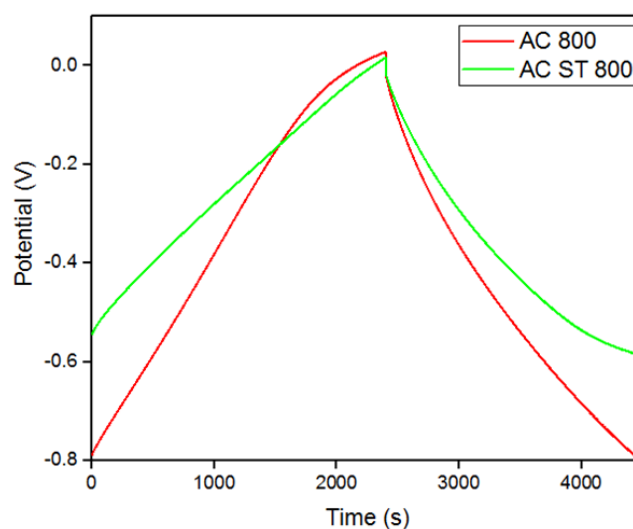


Figure 6. Galvanostatic charge-discharge (GCD) curve of AC 800 and AC ST 800

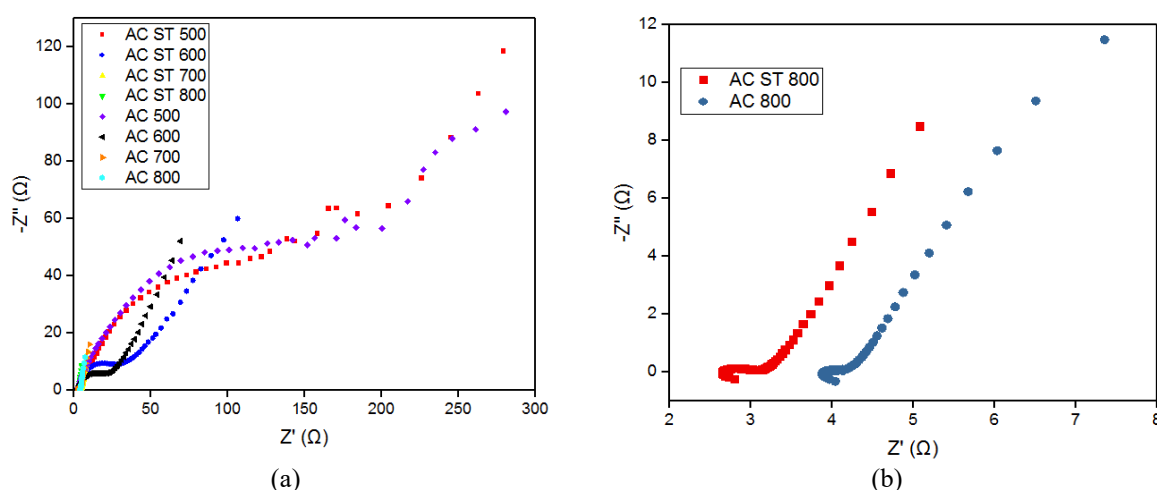


Figure 7. Nyquist plots of a) activated carbon before and after surface treatment; b) AC 800 and AC ST 800

Figure 7 shows the Nyquist plots of the EIS analysis. These Nyquist plots allow determination of the activated carbon resistance. These values were determined by evaluating the semicircle in the plots, which represents the electrolyte resistance at the electrode interface (R_s), the internal resistance of the activated carbon electrode (R_p), and the equivalent series resistance (ESR) of the activated carbon electrode. The values of R_s , R_p , and ESR were represented by the start of the semicircle, the end of the semicircle, and the diameter of the semicircle ($R_p - R_s$), respectively [31]. From the Nyquist plot, activated carbon samples with higher specific capacitance show lower resistance values.

Table 5. ESR value of activated carbon before and after surface treatment

Sample	ESR (Ω)	Sample	ESR (Ω)	Percentage Difference (%)
AC 500	153.96	AC ST 500	155.21	0.81
AC 600	18.96	AC ST 600	25.39	22.91
AC 700	1.25	AC ST 700	1.02	-18.40
AC 800	0.10	AC ST 800	0.31	244.44

Several factors influence the resistance value, including pore characteristics and surface area, defect structure, and the presence of functional groups. In addition, these ESR values can also indicate the behavior of the carbon-based supercapacitor. Typical ESR values for carbon electrode supercapacitors range from 0.134 to 1.149 [32]. Based on Table 5, some activated carbon samples showed decreased ESR, while others increased. Notably, the ESR values for samples AC ST 700 and AC ST 800 remained below 1.149 Ω after surface treatment. Although AC ST 800 showed an increase in ESR, this can be attributed to its initially low ESR, where even a slight increase resulted in a relatively larger percentage increment. The AC ST 700 and AC ST 800, which exhibited higher specific capacitance and lower resistance, proved that with the right combination of the physical activation parameter, a better electrochemical performance could be achieved, and the addition of the oxygen functional group would enhance its specific capacitance while not compromising the internal resistance for the electrolyte and ionic movement within the activated carbon network [22].

Meanwhile, the other segment of the Nyquist plot, the straight line after the semicircle, represents the resistive and capacitive behavior of ions penetrating the activated carbon pores. Therefore, this steep slope indicates efficient ion penetration into the pore networks of activated carbon [33]. For AC ST 600, AC ST 700, and AC ST 800, the slopes were similar, while AC ST 500 showed a significant difference after $Z' = 200 \Omega$, which may also be related to its large change in specific capacitance. This suggests better ion penetration, which led to improved capacitance performance [25].

4. CONCLUSION

This research compares the physical properties and electrochemical performance of activated carbon before and after surface treatment prepared by physical activation. The main conclusions drawn from this research are as follows:

1. The amount of the oxygen functional group developed during surface treatment varied among the samples, although the same treatment parameters were applied. This subsequently led to different electrochemical performance rates.
2. The Raman spectroscopy analysis showed that the surface treatment did not have major impacts on the structural defects of the activated carbon.
3. All samples exhibited a positive increment in specific capacitance after the surface treatment. This indicated that the surface treatment process and parameters used in this study did not cause collapse of the activated carbon structure and, consequently, lower electrochemical performance.
4. Although AC ST 500 recorded the highest amount of oxygen functional groups developed, and translated to the highest specific capacitance increase, it still has the lowest specific capacitance value as compared to the other samples. This is because the overall structure of the activated carbon, such as surface area, porosity, and structural defects developed during the physical activation process, still plays a major role in shaping the electrochemical properties of the activated carbon, while the addition of oxygen functional groups contributed mainly to enhancing those properties.
5. While there were samples that exhibited an increase in the resistance value, this did not have major implications for the electrochemical performance, as all the samples showed a positive increase in the specific capacitance.

REFERENCES

- [1] Rustamaji H, Prakoso T, Devianto H, Widiatmoko P, Kurnia KA (2023) Facile synthesis of N, S-modified activated carbon from biomass residue for promising supercapacitor electrode applications. *Bioresour. Technol. Rep.*, 21:101301.
- [2] Ferdous AR, Shah SS, Shaikh MN, Barai HR, Marwat MA, Oyama M, Aziz MA (2024) Advancements in biomass-derived activated carbon for sustainable hydrogen storage: A comprehensive review. *Chem. Asian J.*, 19(16):e202300780.
- [3] Lobato-Peralta DR, Okoye PU, Alegre C (2024) A review on carbon materials for electrochemical energy storage applications: State of the art, implementation, and synergy with metallic compounds for supercapacitor and battery electrodes. *J. Power Sources*, 617:235140.
- [4] Rozhan AN, Ani MH, Salleh HM, Akiyama T, Purwanto H (2015) Development of carbon-infiltrated bio-char from oil palm empty fruit bunch. *ISIJ Int.*, 55(2):436-440.
- [5] Purwanto H, Salleh HM, Rozhan AN, Mohamad AS, Zakiyuddin A (2018) Phase change of iron ore reduction process using EFB as reducing agent at 900-1200C. *IOP Conf. Ser. Mater. Sci. Eng.*, 342:012054.
- [6] Choi MS, Park S, Lee H, Park HS (2018) Hierarchically nanoporous carbons derived from empty fruit bunches for high performance supercapacitors. *Carbon Lett.*, 25:103-112.
- [7] Nasir S, Hussein MZ, Zainal Z, Yusof NA (2018) Carbon-based nanomaterials/allotropes: A glimpse of their synthesis, properties and some applications. *Materials*, 11(2):1-24.
- [8] Jha MK, Joshi S, Sharma RK, Kim AA, Pant B, Park M, Pant HR (2021) Surface modified activated carbons: Sustainable bio-based materials for environmental remediation. *Nanomaterials*, 11:3140.
- [9] Malini K, Selvakumar D, Kumar NS (2023) Activated carbon from biomass: Preparation, factors improving basicity and surface properties for enhanced CO₂ capture capacity – A review. *J. CO₂ Util.*, 67:102318.
- [10] Petrovic B, Gorbounov M, Soltani SM (2022) Impact of surface functional groups and their introduction methods on the mechanisms of CO₂ adsorption on porous carbonaceous adsorbents. *Carbon Capture Sci. Technol.*, 3:100045.
- [11] Han Y, Zhao PP, Dong XT, Zhang C, Liu SX (2014) Improvement in electrochemical capacitance of AC from scrap tires by nitric acid treatment. *Front. Mater. Sci.*, 8(4):391-398.
- [12] Wei X, Jiang X, Wei J, Gao S (2015) Functional groups and pore size distribution do matter to hierarchically porous carbons as high-rate-performance supercapacitor. *Chem. Mater.*, 28(2):445-458.
- [13] Arango DI, Zapata-Benabith Z, Arenas EC, Perez-Osorno JC (2018) Influence of surface modification with nitric acid on electrochemical performance of agroindustrial waste-based AC. *J. Mater. Sci. Mater. Electron.*, 29(18):15557-15569.
- [14] Liu C, Koyyalamudi BB, Li L, Emani S, Wang C, Shaw LL (2016) Improved capacitive energy storage via surface functionalization of AC as cathodes for lithium ion capacitors. *Carbon*, 109:163-172.
- [15] Zhang C, Wei YL, Cao PF, Lin MC (2018) Energy storage system: Current studies on batteries and power condition system. *Renew. Sustain. Energy Rev.*, 82:3091-3106.
- [16] Marzuki H, Rozhan AN, Purwanto H (2024) Electrochemical performance of activated carbon derived from empty fruit bunch via chemical and physical activation methods. *J. Kejuruteraan*, 36(4):1437-1450.
- [17] Darkwa MK, Zquine C, Kahol PK, Gupta RK (2019) Novel biomass-derived smoke-like carbon as a supercapacitor electrode material. *R. Soc. Open Sci.*, 6(7):190132.
- [18] Haitao S, Qiang Q, Zhuangzhuang Y, Yanrong Z, Ling Q, Yibo Z, Chenguang L, Mingqiang Z, Xuanmin Y (2025) Supercapacitor performance of activated carbon from *Eucommia ulmoides* Oliver wood optimized by the activation method. *ACS Omega*, 10:15368-15380.

- [19] Figueiredo JL (2018) Nanostructured porous carbons for electrochemical energy conversion and storage. *Surf. Coat. Technol.*, 350:307-312.
- [20] Sarwar A, Ali M, Khoja AH, Nawar A, Waqas A, Liaquat R, Asjid M (2021) Synthesis and characterization of biomass-derived surface-modified AC for enhanced CO₂ adsorption. *J. CO₂ Util.*, 46:101476.
- [21] Shi G, Liu C, Wang G, Chen X, Li L, Jiang X, Zhang H (2018) Preparation and electrochemical performance of biomass-based AC nanofibers. *Ionics*, 25(4):1805-1812.
- [22] Yang W, He W, Zhang F, Hickner MA, Logan BE (2014) Single-step fabrication using a phase inversion method of poly(vinylidene fluoride) (PVDF) AC air cathodes for microbial fuel cells. *Environ. Sci. Technol. Lett.*, 1(10):416-420.
- [23] Zheng Y, Deng T, Yue N, Zhang W, Zhu X, Yang H, Zheng W (2021) Raman spectroscopy and correlative-Raman technology excel as an optimal stage for carbon-based electrode materials in electrochemical energy storage. *J. Raman Spectrosc.*, 52(12):2119-2130.
- [24] Singh A, Ghosh K, Kumar S, Agarwal AK, Jassal M, Goswami P, Chaturvedi H (2019) Interdigitated flexible supercapacitor using AC synthesized from biomass for wearable energy storage. *arXiv:1903.02384*.
- [25] Frackowiak E, Béguin F (2001) Carbon materials for the electrochemical storage of energy in capacitors. *Carbon*, 39(6):937-950.
- [26] Olivares-Marín M, Fernández-González C, MacÍas-García A, Gómez-Serrano V (2012) Preparation of AC from cherry stones by physical activation in air. Influence of the chemical carbonisation with H₂SO₄. *J. Anal. Appl. Pyrolysis*, 94:131-137.
- [27] Dieu HTTD, Charoensook K, Tai HC, Lin YT, Li YY (2021) Preparation of activated carbon derived from oil palm empty fruit bunches and its modification by nitrogen doping for supercapacitors. *J. Porous Mater.*, 28:9-18.
- [28] Yu H, Wu J, Fan L, Xu K, Zhong X, Lin Y, Lin J (2011) Improvement of the performance for quasi-solid-state supercapacitor by using PVA-KOH-KI polymer gel electrolyte. *Electrochim. Acta*, 56(20):6881-6886.
- [29] Taer E, Melisa M, Agustino A, Taslim R, Sinta MW, Apriwandi A (2021) Biomass-based AC monolith from *Tectona grandis* leaf as supercapacitor electrode materials. *Energy Sources Part A Recovery Util. Environ. Eff.*, 1-12.
- [30] Ishak MM, Deraman M, Dolah BNM, Othman MAR, Omar R, Basri NH, Aziz AA (2015) Effects of activation time on the performance of supercapacitor binderless AC electrodes derived from fibers of oil palm empty fruit bunches. *Adv. Mater. Res.*, 1112:308-312.
- [31] Farma R, Deraman M, Awitdrus, Talib IA, Omar R, Manjunatha JG, Dolah BNM (2013) Physical and electrochemical properties of supercapacitor electrodes derived from carbon nanotube and biomass carbon. *Int. J. Electrochem. Sci.*, 8(1):257-273.
- [32] Kim YJ, Lee BJ, Suezaki H, Chino T, Abe Y, Yanagiura T, Endo M (2006) Preparation and characterization of bamboo-based ACs as electrode materials for electric double layer capacitors. *Carbon*, 44(8):1592-1595.
- [33] Chen WC, Wen TC (2003) Electrochemical and capacitive properties of polyaniline-implanted porous carbon electrode for supercapacitors. *J. Power Sources*, 117:273-282.

Comparison of Fatigue Crack Growth Models for Predicting Residual Life

MUHAMMAD HUSNAIN MOHD NOH¹, MOHD AKRAMIN MOHD ROMLAY^{1*},
MOHD SHAMIL SHAARI¹, AKIYUKI TAKAHASHI²

¹Faculty of Mechanical & Automotive Engineering Technology,
Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang Darul Makmur, Malaysia

²Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

*Corresponding author: akramin@umpsa.edu.my

(Received: 19 August 2024; Accepted: 10 February 2026; Published online: 11 September 2026)

ABSTRACT: Fatigue is one of the most critical failure mechanisms in structural materials. Accurate prediction of fatigue crack growth and residual fatigue life is essential to ensure structural safety and reliability during design and service-life assessment. This study aims to predict the residual fatigue life of a four-point bending specimen under fatigue loading conditions using the S-version finite element model (S-version FEM). Four fatigue crack growth models are implemented within the S-version FEM framework, namely Paris' Law, Frost and Pook, Walker, and Huang Moan models, to simulate crack propagation and estimate residual fatigue life. Experimental data from Al7075-T6 are used for validation purposes. The results indicate that the Frost and Pook model achieves an error below 20%, demonstrating acceptable engineering accuracy. In contrast, the Paris' Law, Walker, and Huang Moan models show significantly higher deviations from the experimental results. Overall, the Frost and Pook model exhibits the best agreement with the experimental findings and provides improved predictive capability compared to the other empirical models, confirming the effectiveness of the proposed S-version FEM approach for fatigue crack growth and residual fatigue life prediction in four-point bending structures.

ABSTRAK: Retak-lesu merupakan salah satu mekanisme kegagalan paling kritikal dalam bahan struktur. Ramalan yang tepat terhadap pertumbuhan retak-lesu dan baki jangka hayat adalah penting bagi memastikan keselamatan dan kebolehpercayaan struktur dalam reka bentuk serta penilaian hayat perkhidmatan. Kajian ini bertujuan bagi meramalkan baki jangka hayat retak-lesu spesimen lenturan empat titik menggunakan model unsur terhingga S-version (S-version FEM) di bawah beban keletihan. Empat model pertumbuhan retak-lesu dijalankan dalam rangka kerja S-version FEM, iaitu Hukum Paris, Frost dan Pook, Walker, dan Huang Moan bagi mensimulasikan penyebaran retak serta menganggar baki jangka hayat retak-lesu. Data eksperimen daripada Al7075-T6 digunakan untuk tujuan pengesahan model. Keputusan menunjukkan bahawa model Frost dan Pook memberikan ralat kurang daripada 20%, sekaligus menunjukkan ketepatan kejuruteraan yang boleh diterima. Sebaliknya, model Hukum Paris, Walker, dan Huang Moan menunjukkan sisihan lebih tinggi berbanding keputusan eksperimen. Secara keseluruhan, model Frost dan Pook menunjukkan kesepadanan terbaik dengan keputusan eksperimen serta memberikan keupayaan ramalan yang lebih baik berbanding model empirik yang lain, sekaligus mengesahkan keberkesanan pendekatan S-version FEM yang dicadangkan dalam meramal pertumbuhan retak-lesu dan baki jangka hayat bagi struktur lenturan empat titik.

KEYWORDS: *Residual life; S-version finite element model; Fatigue Crack Growth Models*

1. INTRODUCTION

Typically, more than 90% of engineering component failures are attributed to fatigue. The primary objective of fatigue crack growth (FCG) analysis is to identify the parameters governing crack initiation and propagation under cyclic loading conditions. Therefore, understanding fatigue crack growth behavior is essential because it significantly influences structural integrity and material performance. Fatigue crack growth problems are complex to simulate because of the nonlinear nature of crack initiation and propagation mechanisms. Although several numerical methods, including the hybrid method, the Extended Finite Element Method (XFEM), and the conventional Finite Element Method (FEM), have been widely applied, each presents notable limitations. The hybrid method is associated with high computational cost; XFEM requires complex formulation and implementation; while conventional FEM struggles to accurately capture evolving crack paths without remeshing.

In fracture mechanics, fatigue life estimation has been extensively developed based on crack propagation analysis. One of the most important parameters in this field is the stress intensity factor (SIF), which characterizes the stress field near the crack tip and governs crack growth behavior [1]. However, accurately predicting fatigue crack growth in complex geometries remains challenging. Conventional FEM is widely used for such analyses, but it requires continuous remeshing as cracks propagate, making it computationally expensive and inefficient, particularly for three-dimensional problems [2].

To overcome these limitations, the S-version FEM has been developed. This method significantly reduces remeshing requirements and enables efficient simulation of complex crack propagation paths [2]. The S-version FEM is commonly integrated with linear elastic fracture mechanics (LEFM) for fatigue crack growth and residual life prediction. In LEFM, the stress concentration near the crack tip, expressed through the SIF, is a key parameter in describing crack propagation behavior.

Furthermore, fatigue crack growth rates are typically predicted using empirical crack growth laws combined with LEFM principles [3]. FEM is widely applied to evaluate fracture mechanics parameters such as SIF and energy release rate, which quantify crack driving forces and stress concentration at the crack tip [4]. Fatigue is defined as the progressive and localized structural degradation of materials subjected to repeated cyclic loading [5]. Therefore, accurate prediction of FCG is essential for estimating the remaining useful life of engineering components and remains a major focus in both academic and industrial research [6, 7].

The fatigue crack growth rate is commonly described using material constants and the SIF range, which govern crack propagation behavior [8]. Crack growth significantly reduces the fatigue life of engineering structures. From a fracture mechanics perspective, the energy release rate represents the rate of energy transformation during crack propagation. Numerical approaches such as the virtual crack closure technique (VCCT) [9], [10] and the J-integral method [11, 12] are widely implemented within FEM frameworks to evaluate the energy release rate. Both energy release rate and SIF are strongly correlated and widely used in fatigue analysis.

The VCCT determines the energy required to close a crack by multiplying nodal reaction forces with crack opening displacements [4]. Its advantages include computational simplicity and the ability to decompose energy release rates into mode I, mode II, and mode III components. VCCT has been successfully applied to various three-dimensional fracture mechanics problems [4, 13, 14].

FCG models are mathematical formulations used to predict crack propagation rates under cyclic loading conditions [3]. Common models include Paris' Law [15], the Wheeler model [16], the Forman model [17], the Walker model [18], the Frost & Pook model [19], and the modified Paris-Erdogan model [20], among others [21-23]. These models aim to describe crack growth behavior over loading cycles [20]. Although numerous models have been proposed, their comparative performance within advanced numerical frameworks remains insufficiently investigated.

Among these models, Paris' Law is one of the most widely used empirical formulations for predicting fatigue crack growth under cyclic loading. Paris, Gomez, and Anderson initially introduced the relationship in 1961, and Paris and Erdogan subsequently formalized it in 1963 [15]. It relates the fatigue crack growth rate, (da/dN) , to the SIF range, (ΔK) , using a power-law relationship and has been extensively validated and refined for various engineering materials [24]. However, the original Paris formulation is mainly applicable to the stable crack growth region under constant-amplitude loading and a fixed stress ratio. It does not explicitly account for stress-ratio effects, crack-growth thresholds, load-sequence interactions, or retardation caused by overloads. The Frost and Pook model provides an alternative formulation in which the fatigue crack growth rate is related to the square of the SIF range and the material's elastic properties [25]. This formulation introduces greater sensitivity to material behavior and can be useful when comparing crack growth characteristics across different materials. Nevertheless, the model does not explicitly incorporate load-history or load-interaction effects, which may limit its accuracy under complex variable-amplitude loading.

The Walker model extends Paris' Law by introducing the stress ratio (R) through an equivalent SIF range and a material-dependent exponent, allowing fatigue crack growth data obtained under different stress ratios to be unified into a single relationship and improving prediction accuracy when mean stress varies; however, it remains an empirical steady-state model and does not capture load-sequence effects such as overload-induced retardation or underload acceleration [26-28]. In contrast, the Huang Moan model provides a more comprehensive approach for variable-amplitude loading by combining an equivalent SIF (normalized to $R = 0$) with a modified Wheeler retardation concept [29, 30], while accounting for both stress-ratio effects and plastic zone size ahead of the crack tip, enabling it to describe key load-interaction phenomena such as crack growth retardation or temporary arrest after overloads and acceleration following underloads, and has been validated for materials such as aluminium alloys and structural steel under complex loading spectra; therefore, it is generally more suitable than the Paris, Frost & Pook, and Walker models for fatigue crack growth analyses involving variable-amplitude loading histories.

Fatigue crack initiation and propagation under cyclic loading are among the primary causes of structural failure. However, damage-tolerant structures are designed to sustain applied loads even in the presence of cracks or partial damage. Therefore, an understanding of fatigue crack growth laws is essential for damage-tolerant design and reliable fatigue life prediction [31]. This knowledge enables engineers to make informed decisions regarding structural design, maintenance scheduling, and performance optimization.

Fatigue and fracture failure modes significantly influence the integrity and reliability of metallic materials. Investigating these mechanisms is essential to ensure structural safety, enhance design reliability, and improve material performance across engineering applications. However, challenges such as multiscale material behavior, environmental effects, and experimental limitations make fatigue and fracture analysis highly complex. Therefore, advanced modeling techniques and accurate predictive approaches are essential for improving fatigue life estimation and structural integrity assessment.

2. METHODOLOGY

This study implements FCG models, namely Paris' Law, Frost & Pook, Walker, and Huang-Moan models, within the S-version FEM framework for fatigue life prediction. Integrating these models enables accurate simulation of crack propagation behavior and residual life estimation under cyclic loading conditions. Modern engineering structures require more precise and reliable assessment techniques due to increasing complexity in design, material properties, geometry, loading conditions, operating environments, manufacturing processes, and structural applications. Therefore, an advanced numerical approach is necessary to evaluate fatigue behavior using computational modeling. In this study, the VCCT is employed to compute fatigue crack growth parameters. VCCT is used to determine the energy required for crack propagation by evaluating nodal forces and crack opening displacements. Within the LEFM framework, the primary governing parameters for crack growth are the SIF and the energy release rate, which describe the stress state and energy driving crack propagation at the crack tip. The fatigue crack growth process is modeled using the computed SIF to predict crack extension per loading cycle. The number of cycles to failure is then determined by incrementally updating the crack length using the selected fatigue crack growth models until the critical crack size is reached.

2.1. S-version Finite Element Model (S-version FEM)

This section explains crack growth at the crack front within the S-version FEM framework. Important variables such as the SIF and energy release rate are employed to describe crack growth behavior in LEFM. The VCCT, which has been further developed to be compatible with the S-version FEM framework, is utilized to define the energy release rate [32]. The crack propagation direction at the crack front can be determined with the support of this extension [33]. Assuming that the faces of the finite element mesh are positioned orthogonally at the crack front, Fig. 1 illustrates the three-dimensional VCCT in its original configuration. As shown in Fig. 1, the nodal opening displacement P_i , where $(i = 1, 2, 3, 4, 5, 6)$, is measured near the crack front within the VCCT framework. The five nodes at the crack front boundary represent the opening displacement, v_i . S_1 and S_2 denote the areas of the finite element faces at the crack front segment.

As shown in Fig. 1, the displacement of the upper crack surface is represented by v_{iU} , while the displacement of the lower crack surface is denoted by v_{iL} . Therefore, the definition of the opening displacement v_i is given as follows [6]:

$$v_i = v_{iU} - v_{iL} \quad (1)$$

If the analysis incorporates a combination of global and local meshes, the displacement is calculated as the sum of the global and local displacement components for both the upper and lower crack surfaces. Accordingly, the displacement for both surfaces can be expressed as:

$$v_{iU} = v_{iU}^G - v_{iU}^L, v_{iL} = v_{iL}^G - v_{iL}^L \quad (2)$$

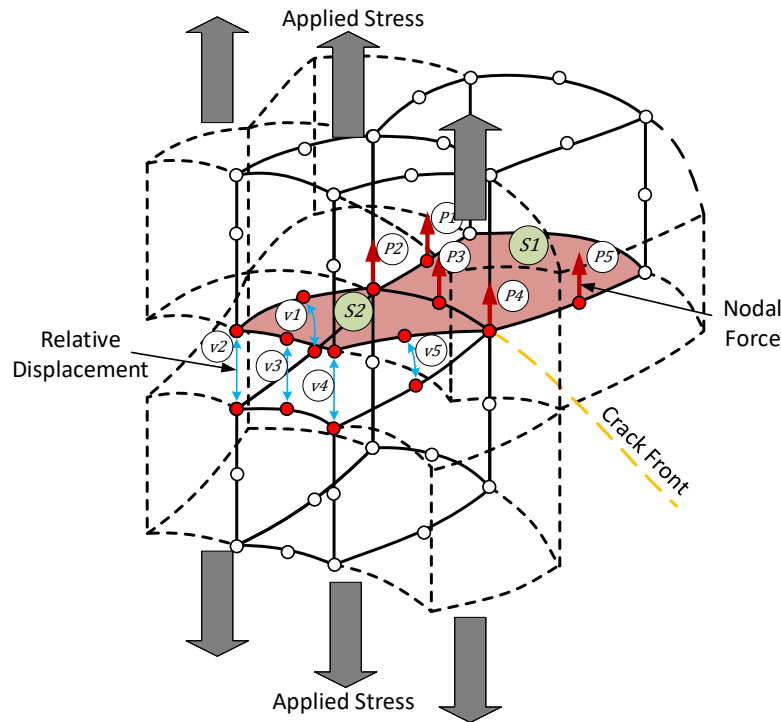


Figure 1. VCCT for local mesh at the crack front in three-dimensional analysis.

For this method, the opening displacement exists exclusively within the local mesh. In contrast, within the global region, the opening displacement is assumed to be zero. Therefore,

$$v_{iU}^G = -v_{iL}^G \quad (3)$$

Eqs. (1), (2), and (3) are combined to determine the cumulative opening displacement v_i at the nodes within the local region. The resulting expression for v_i is given as follows [34]:

$$v_i = (v_{iU}^G + v_{iU}^L) - (v_{iL}^G - v_{iL}^L) = v_{iU}^L - v_{iL}^L \quad (4)$$

The opening displacement around the crack front is computed first, followed by the energy release rate. Okada et al. [33] first introduced the energy release rate (start equation cap G) for the non-symmetrical configuration of finite element faces at the crack front. They proposed, as shown in Fig. 2, that the energy release, denoted by ΔG_I , is evaluated at the instant the crack front opens within the region S_1 . This computation, which occurs during fracture propagation under Mode I failure, is expressed as follows:

$$\Delta G_I = \frac{1}{2} \int_{S_1} \sigma_{33}(r) v_3 (\Delta - r) dS \quad (5)$$

The element configuration at the crack front is shown in Fig. 2, where r and Δ represent the element width and length, respectively, in directions perpendicular and parallel to the fracture front. Three fracture modes contribute to crack propagation: Mode I, Mode II, and Mode III. The cohesive stress, σ_{33} , within the crack front plane is indicated by the first subscript 3, which represents the face of the crack front in the vertical direction, and by the second subscript 3, which corresponds to the crack front coordinate system, x_1 , x_2 , and x_3 , as seen in Fig. 2. The displacement v_i represents the crack opening displacement function in the Mode I direction. Broek [34] provides further details on stress and displacement at the crack face. The formulation for the cohesive stress and displacement function is given as follows:

$$\sigma_{33}(\bar{r}) = \frac{K_I}{\sqrt{2\pi\bar{r}}}, v_3(\bar{r}) = 4\sqrt{\frac{2\bar{r}}{\pi}} \frac{K_I}{E^*} \quad (6)$$

$E' = E/(1 - \nu^2)$ represents the plane strain condition, while $E^* = E$ represents the plane stress condition. The Poisson's ratio and Young's modulus of the material are denoted by ν and E , respectively. The angle between the loading direction and the average crack front direction is denoted by θ , whereas $\bar{r}$ represents the distance from the crack front. By substituting the above expressions into Eq. (4), the stress and displacement fields can be expressed as follows:

$$\Delta G_I = \int_{\theta_1}^{\theta_2} \int_0^\Delta \frac{K_I}{\sqrt{2\pi r \cos\theta}} 2\sqrt{\frac{2(\Delta-r)\cos\theta}{\pi}} \frac{K_I}{E'} (R+r) dr d\theta \quad (7)$$

Therefore, the following equation can be used to represent Eq. (7), which describes the SIFs for the S_1 and S_2 regions, as illustrated in Fig. 2:

$$G_I = \frac{K_I^2}{E'} = \frac{1}{2[S_1 - \frac{1}{4}(S_1 - S_2)]} \sum_{i=1}^5 v_i^3 P_i^3 \quad (8)$$

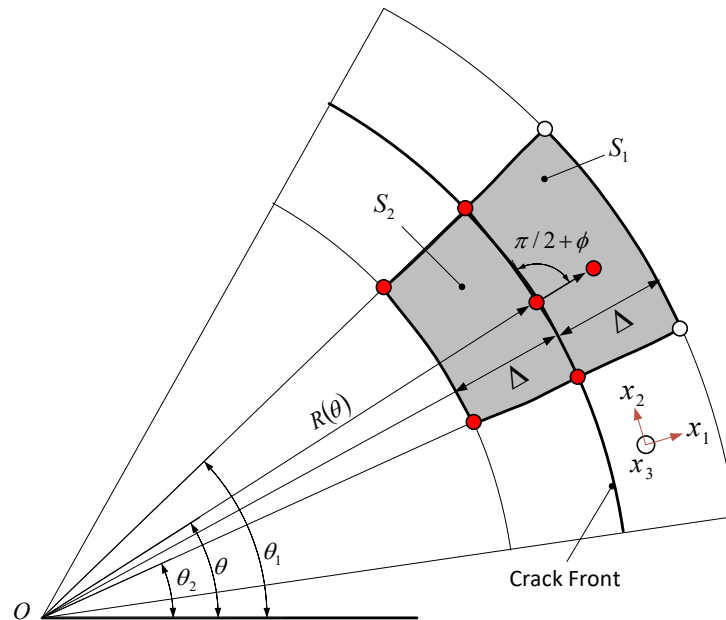


Figure 2. Element arrangement at the crack front of the material.

The energy release rate for the remaining failure modes is stated as follows:

$$G_{II} = \frac{K_{II}^2}{E'} = \frac{1}{2[S_1 - \frac{1}{4}(S_1 - S_2)]} \sum_{i=1}^5 v_i^1 P_i^1 \quad (9)$$

$$G_{III} = \frac{K_{III}^2}{2\mu} = \frac{1}{2[S_1 - \frac{1}{4}(S_1 - S_2)]} \sum_{i=1}^5 v_i^2 P_i^2 \quad (10)$$

Here, the shear modulus is denoted by μ . In the energy release rate G , each component is represented by a subscript. The total energy release rate, G_{Total} , is obtained by summing G_I , G_{II} , and G_{III} . The conversion of the energy release rate to the SIF is possible, as shown in Eqs. (8), (9), and (10). For further details on the derivation of the element arrangement near the fracture front, refer to Okada et al. [33].

2.1.1. Paris Law Model

The Paris law, often expressed as a power law and originally proposed by Paris and Erdogan [15], is a well-known and straightforward method for predicting fatigue crack growth. This relationship provides the governing equation and marks one of the earliest applications of fracture mechanics to fatigue analysis.

$$\frac{da}{dN} = C_p(\Delta K)^n \quad (11)$$

where n is the slope and C_p is the intercept on the log–log plot of da/dN versus ΔK . Eq. (11) describes Region II of the fatigue crack growth rate curve, which exhibits a linear relationship on a log–log scale between da/dN and ΔK . The Paris law is simple to apply, requiring only two curve-fitting parameters that can be readily determined from experimental data. It is widely used in fatigue analysis, particularly when the experimental data exhibits a linear relationship in Region II.

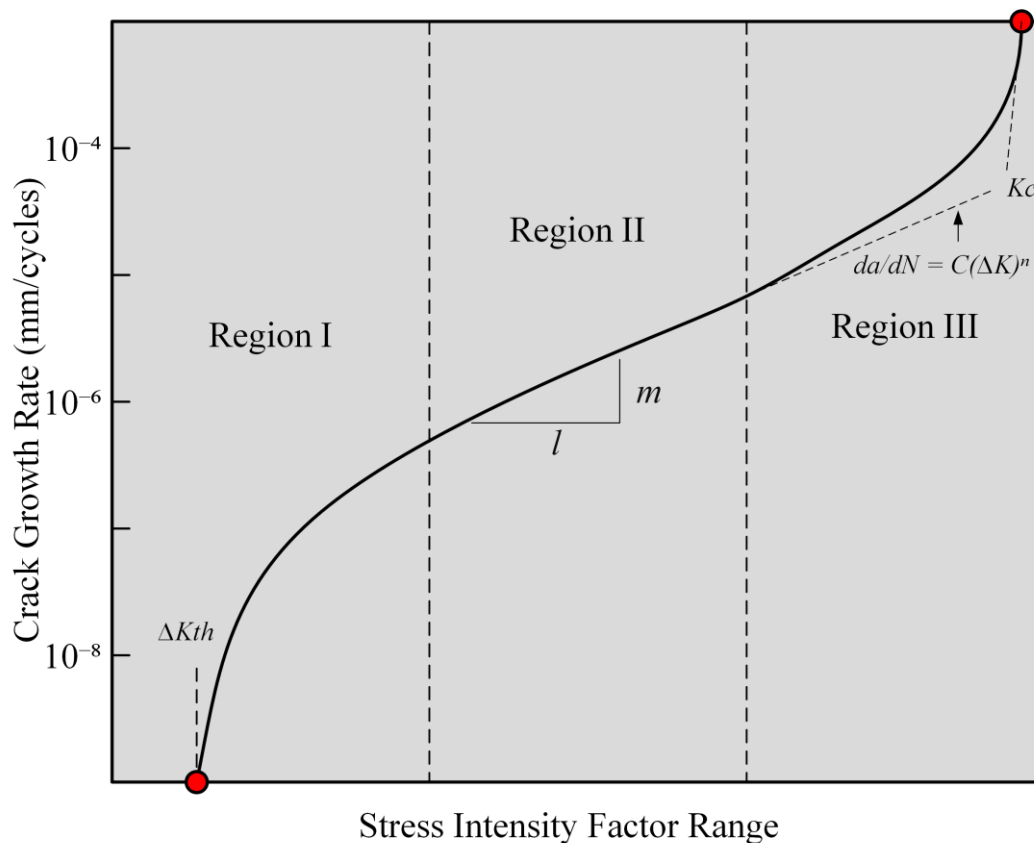


Figure 3. Typical da/dN versus ΔK curve $\text{Log } \Delta K$

The Paris law is limited to describing fatigue crack growth data in Region II (see Fig. 3). It cannot adequately characterize Regions I and III, where threshold behavior and rapid crack growth occur, respectively. Therefore, additional modeling approaches may be required depending on the nature of the analysis. Furthermore, the Paris law is material-dependent and does not account for the influence of stress ratio. For steels tested under different stress ratios, a family of parallel straight lines is typically obtained. This indicates that, while the value of C_p varies with stress ratio, the exponent n remains constant for a given material. As a result, ensure that fatigue crack growth data is applied using the appropriate stress ratio.

2.1.2. Frost & Pook Model

In 1971, Frost and Pook [19] developed an equation for fatigue crack growth that incorporates physical significance. According to their theory, cyclic loading causes cracks to propagate not through gradual structural damage, but because unloading reconfigures the crack tip after each cycle, which in turn forms striations [35]. Consequently, a simple model based on incremental crack growth per cycle can be related to the evolving crack tip geometry during opening and closure. The relationship is given as follows:

$$\frac{da}{dN} = \frac{9}{\pi} \left(\frac{\Delta K}{E} \right)^2 \quad (\text{plane stress}) \quad (12)$$

$$\frac{da}{dN} = \frac{7}{\pi} \left(\frac{\Delta K}{E} \right)^2 \quad (\text{plane strain}) \quad (13)$$

At crack growth rates of approximately 3×10^{-5} mm/cycle, a strong correlation was observed between experimental results plotted in terms of $\Delta K/E$ versus da/dN and the predictions of the plane stress formulation. However, at higher crack growth rates, the model tends to overestimate crack growth, while at lower rates it tends to underestimate it. Since ΔK is derived from the relationship between E' and the energy release rate G , the model implicitly incorporates the initial crack length and mean stress.

2.1.3. Huang Moan Model

Unlike the basic Paris law model, which does not explicitly incorporate stress ratio effects, the Huang Moan model accounts for the influence of stress ratio in the fatigue crack growth formulation [36]. The Huang and Moan model is mathematically expressed as:

$$\frac{da}{dN} = C_m (\Delta K_{eq})^{n_m} \quad (14)$$

$$\text{where, } M = \begin{cases} (1-R)^{-1.2\beta} & (-5 \leq R < 0) \\ (1-R)^{-\beta} & (0 \leq R < 0.5) \\ (1.05 - 1.4R + 0.6R^2)^{-\beta} & (0.5 \leq R < 1) \end{cases} \quad (15)$$

where C_m and n_m are material constants, M is a correction factor used to account for the mean stress effect over the range $-5 \leq R < 1$, and ΔK_{eq} is the equivalent SIF range. The stress ratio range is divided into three intervals, and the correction factor M is used to shift fatigue crack growth curves at different stress ratios to the corresponding curve at $R = 0$. The parameter β is determined by fitting experimental fatigue crack growth data and is dependent on the material type [37].

2.1.4. Walker Model

The Walker model is commonly regarded as a traditional approach for estimating fatigue crack growth in different materials, especially in the stable crack propagation region [38]. The Walker model is expressed as:

$$\frac{da}{dN} = C_w \left[\frac{\Delta K_{eq}}{(1-R)^{1-\gamma}} \right]^{n_w} \quad (16)$$

The variables include a , which represents crack length, and N , which is the number of load cycles. C_w and n_w are material constants, while γ is the Walker coefficient specific to the material. ΔK denotes the SIF range, and R represents the stress ratio.

3. RESULTS AND DISCUSSION

Fig. 4 shows the geometry of the four-point bending specimen subjected to fatigue loading. Based on the experimental configuration, the S-version FEM was applied to simulate the four-point bending setup. The model incorporates a semi-elliptical crack located at the center of the specimen [39, 40]. The initial crack length and depth were defined using experimental data from Akramin [41] were used to establish the initial crack's length and depth. The experimental setup was deliberately developed as a non-standard test rather than strictly adhering to ASTM standards. This approach better captures the behavior of real engineering structures under service-like conditions, where idealized ASTM loading and boundary assumptions may not fully represent actual structural responses. Although the ASTM four-point bending configuration serves as a useful benchmark for controlled laboratory testing, this study focuses on a more realistic scenario that reflects practical engineering applications. As a result, the non-standard experimental design improves the applicability of the findings to real-world engineering problems. Table 1 summarizes the material properties of Aluminum Alloy 7075-T6, which was selected for this study. This alloy was chosen due to its widespread use in older aircraft manufactured in the 1970s [20].

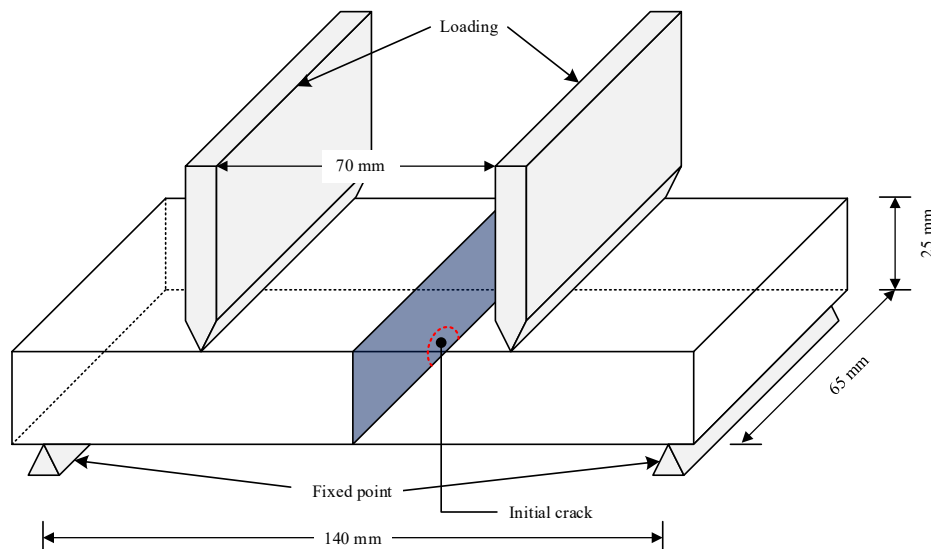


Figure 4. Geometry of four-point bending model.

The four-point bending model was subjected to fatigue loading with a stress ratio of 0.1, corresponding to a maximum load of 45 kN and a minimum load of 4.5 kN.

Table 1. Material properties of Aluminum Alloy 7075-T6

Variables	Values
Fracture toughness, K_{IC}	$29 \text{ MPa} \cdot \sqrt{\text{m}}$
Young's modulus, E	71.7 GPa
Tensile strength, Yield	503 MPa
Poisson's ratio, ν	0.33
Fatigue power parameter, n	2.88
Paris coefficient, C	2.29×10^{-10}
The threshold value, K_{th}	$3.66 \text{ MPa} \cdot \sqrt{\text{m}}$

Fig. 5 compares fatigue-life predictions for the four-point bending model using Frost & Pook, Walker, Huang Moan, and Paris' Law approaches against the experimental data reported

by Akramin et al. [41]. The results show a typical nonlinear crack-growth response, where the crack length increases gradually in early loading cycles, followed by an accelerated growth stage before catastrophic failure. The experimental results indicate an initial stable crack propagation stage, where the crack length increases slowly from approximately 4 mm to around 6 mm up to about 250×10^3 cycles. This is followed by a transition region where the crack growth rate begins to accelerate, leading to a rapid increase in crack length beyond approximately $350\text{--}380 \times 10^3$ cycles. Final failure occurs at approximately $450\text{--}460 \times 10^3$ cycles, where the crack length approaches 18–19 mm.

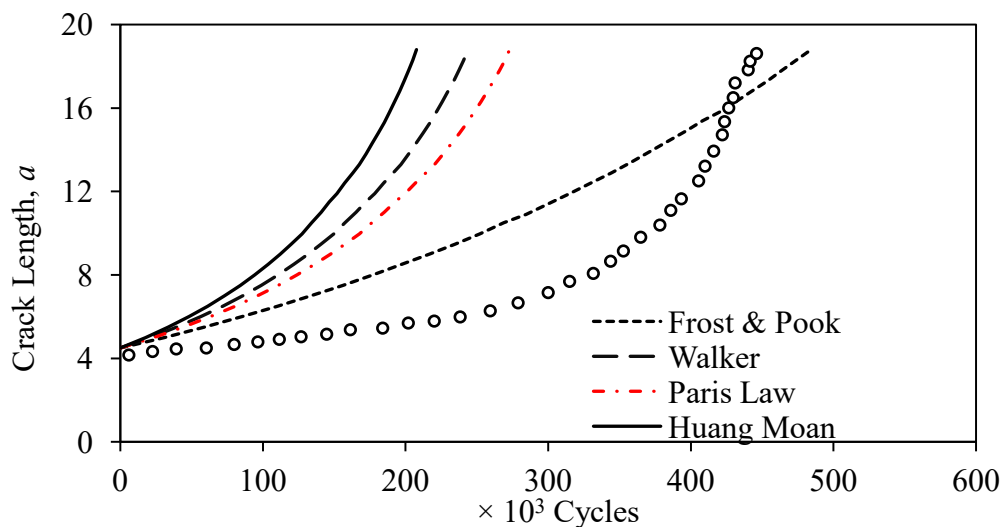


Figure 5. Comparison of experimental, Paris law model, Walker model, Huang Moan model and Frost & Pook model according to the four-point bending specimen's fatigue life prediction.

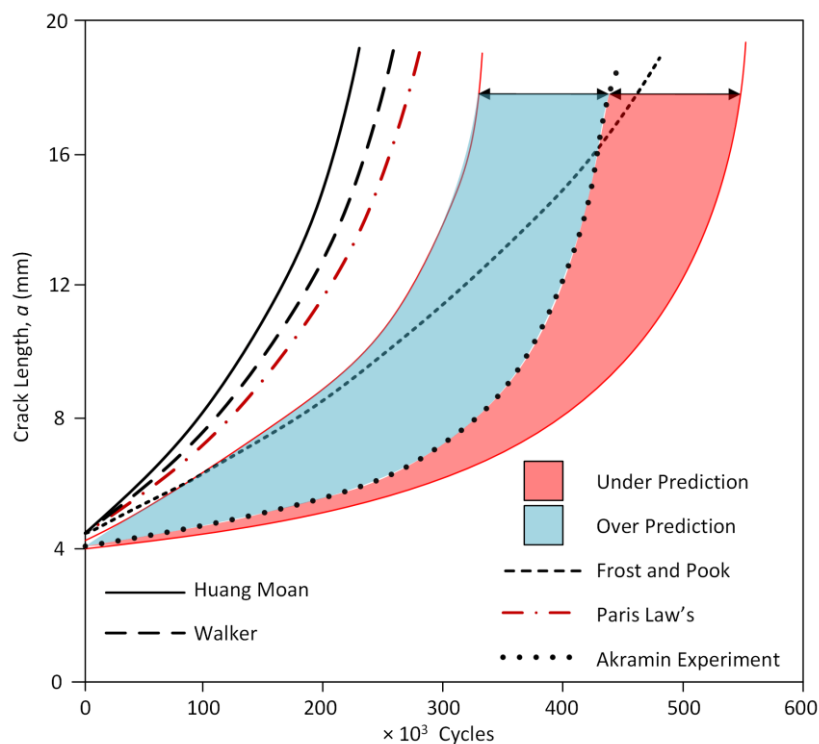


Figure 6. Schematic illustration of FCG models for fatigue life prediction.

Among the predictive models, the Huang-Moan model shows the fastest crack growth rate, significantly underestimating fatigue life by predicting failure at around 210×10^3 cycles. Similarly, both Walker and Paris' Law models underestimate fatigue life, predicting failure earlier than the experimental data at approximately $240\text{--}270 \times 10^3$ cycles. This indicates that both models are overly conservative for the current four-point bending configuration. In contrast, the Frost & Pook model best matches the experimental trend. It captures the gradual transition from stable to unstable crack growth more realistically and predicts fatigue failure at approximately $470\text{--}490 \times 10^3$ cycles, which is slightly higher but much closer to the experimental fatigue life compared to the other models. Although minor deviations exist in the early crack growth stage, with slight overprediction, the overall trend alignment is satisfactory. Overall, the comparison indicates that the Frost & Pook model offers the most reliable prediction of fatigue crack growth behavior under four-point bending loading conditions, particularly in capturing the long stable crack-growth region before final rapid failure.

Fig. 6 illustrates the under- and over-life predictions based on the experimental outcome, which may lead to excessive risk-taking or maintenance costs. Excessive risk could lead to early failure, resulting in equipment and life loss. Because parts would need to be replaced too soon, excessive maintenance would shorten the life. As such, a precise forecast of the damage's actual response is necessary. Thus, fatigue-life forecasts based on FCG models can help avoid catastrophic failure. To achieve these aims, damage-life projections must be well defined.

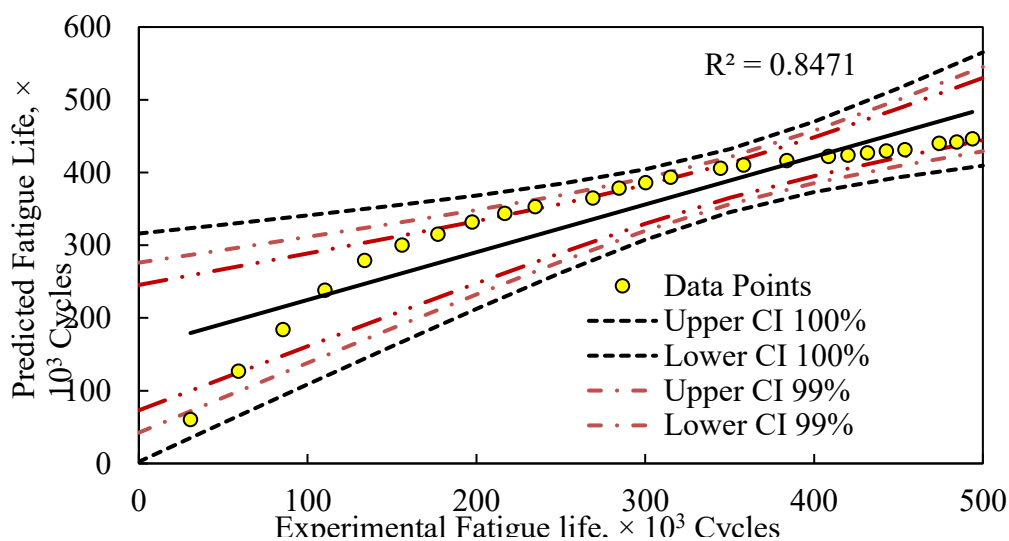


Figure 7. Frost & Pook fatigue crack growth model with regression line and confidence interval.

Fig. 7 and Fig. 8 show the correlation analysis of the fitted line plots of the Frost & Pook model and Paris Law model cycles against experimental cycles. The Walker and Huang Moan models were excluded from the graph because they ranked third and fourth in prediction performance, and because including them would clutter the graph. This technique was crucial in assessing fatigue life estimates that showed a higher degree of goodness of fit with the experimental data. The Paris law model shows a higher determination coefficient value, R^2 with 0.9076, against the Frost & Pook model, R^2 value of 0.8471. Thus, the Paris law model provides a better fit to the observations. Moreover, Fig. 7 and Fig. 8 show fatigue life predictions compared with experimental results, with 100%, 99%, and 95% confidence intervals within the life factor of 2 bands. The fatigue life prediction performs excellently compared with the experimental results within the 100% confidence interval. The predictions agree, as they fall within the 100% confidence interval. Most predictions fall within a life factor of two for the 95% confidence interval. The trend of the prediction interval among the two

models is different. The Frost & Pook model trend is broader at the beginning and narrower in the middle, starting within the range of 200,000 to 400,000 cycles. Furthermore, the Frost & Pook model deviated from the experimental results [41] from 0 to 350,000 cycles, but after 350,000 cycles, it approached the experimental results. Meanwhile, the deviation of the Paris Law model is always constant across the number of experimental cycles, according to Fig. 8. Thus, the Frost & Pook model is considered to perform better because it converges more closely to the experimental data in the final stage of fatigue life, despite its lower R^2 value. In contrast, the Paris law model shows a relatively constant deviation across the overall fatigue life range.

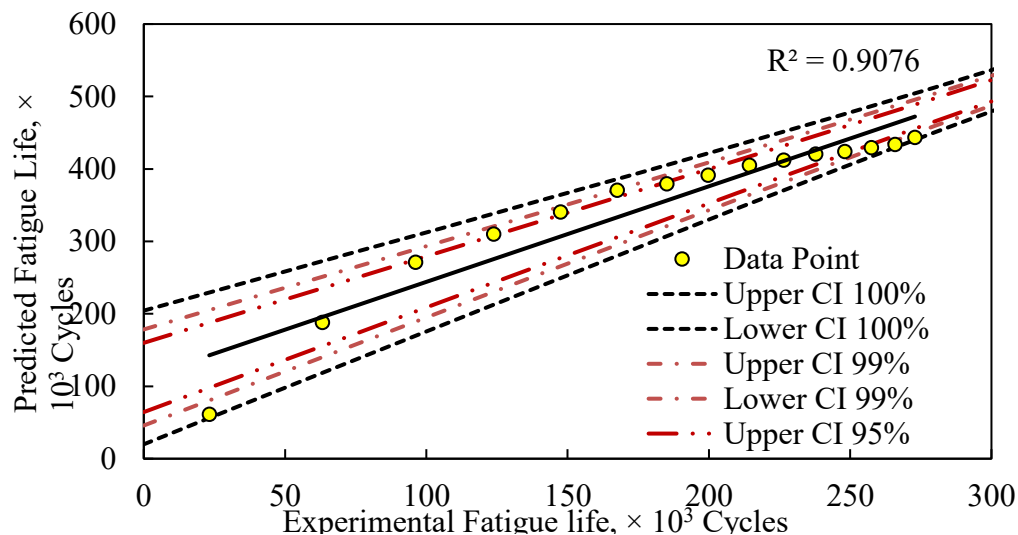


Figure 8. Paris Law fatigue crack growth model with regression line and confidence interval.

The prediction accuracy of the fatigue life models was evaluated using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE), as shown in Figs. 9, 10, and 11. These statistical indicators provide a comprehensive assessment of the deviation between predicted and experimental fatigue life results for the four-point bending configuration. From Fig. 11, the Frost and Pook model has the lowest NRMSE (0.1723), indicating the smallest normalized deviation from the experimental data among the models. In contrast, the Paris Law, Walker, and Huang Moan models show significantly higher NRMSE values of 0.4574, 0.5102, and 0.5759, respectively, demonstrating poorer predictive performance and larger deviations from the experimental trend.

A similar pattern is observed in Fig. 10, where the Frost and Pook model records the lowest RMSE value (65.85×10^3 cycles), whereas the Paris Law (174.77×10^3 cycles), Walker (194.94×10^3 cycles), and Huang Moan (220.06×10^3 cycles) models exhibit substantially higher error magnitudes. This indicates that the Frost and Pook formulation more accurately represents crack growth behavior under four-point bending loading conditions. Furthermore, Fig. 9 shows that the Frost and Pook model achieves the lowest MAPE value (16.87%), while the Paris Law, Walker, and Huang Moan models record higher errors of 50.64%, 55.93%, and 62.45%, respectively. Notably, all competing models exceed 50% error, whereas the Frost and Pook model remains below 20%, which is generally considered an acceptable range for engineering prediction accuracy. Overall, the combined error analysis confirms that the Frost and Pook model best agrees with experimental fatigue life data [41] under four-point bending conditions. Its consistently lower MAPE, RMSE, and NRMSE values indicate higher predictive reliability and better ability to capture fatigue crack growth behavior than the Paris Law, Walker, and Huang Moan models.

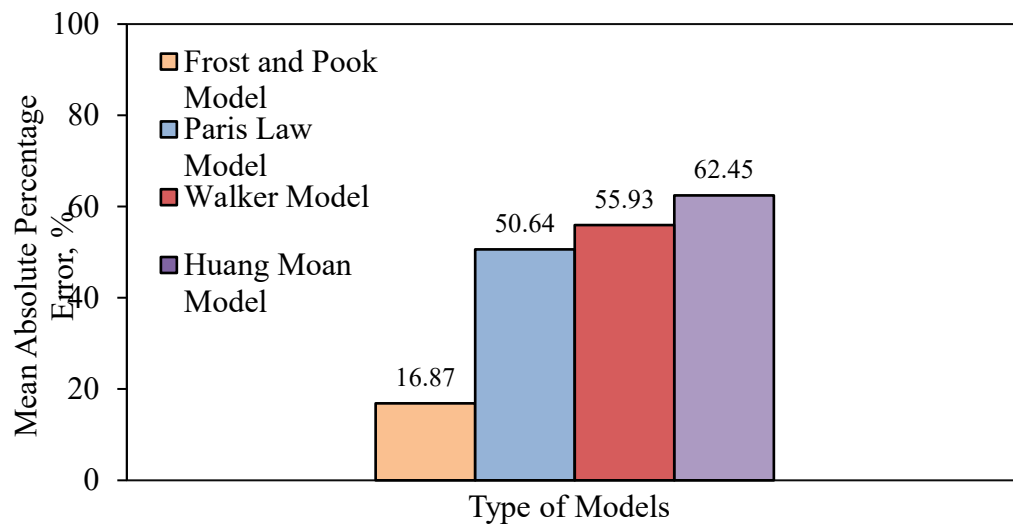


Figure 9. Bar chart of MAPE for fatigue life prediction.

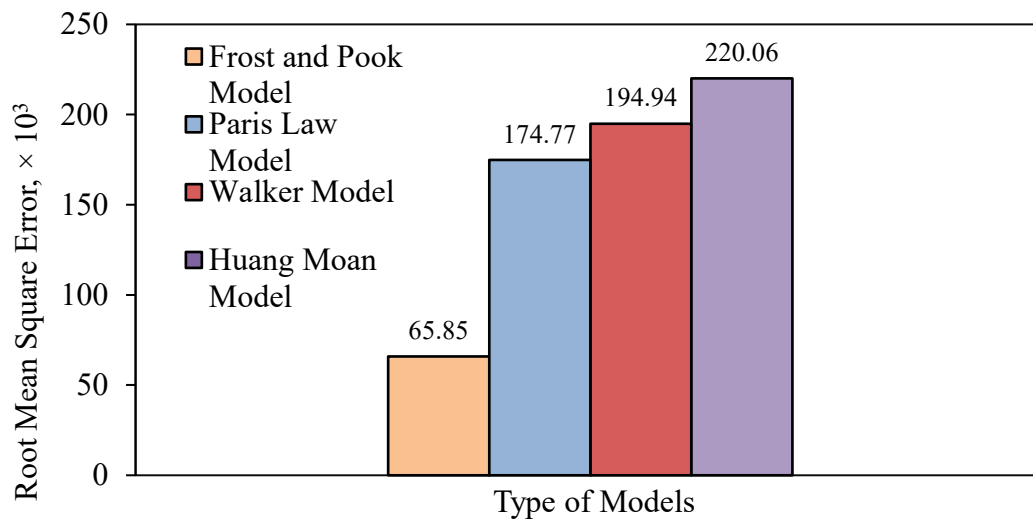


Figure 10. Bar chart of RMSE for fatigue life prediction.

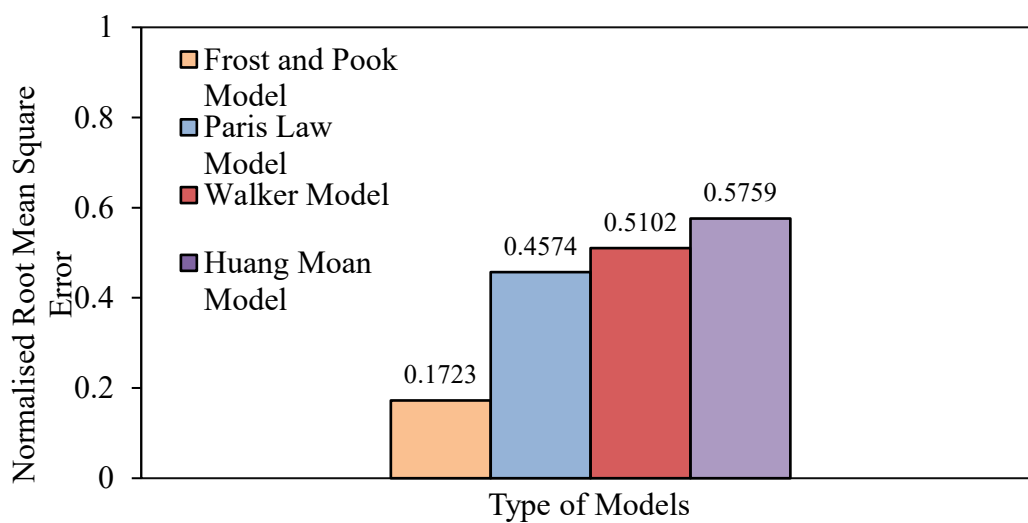


Figure 11. Bar chart of normalized RMSE for fatigue life prediction.

Although the Paris Law model exhibited a higher R^2 value in the correlation analysis, this indicator primarily reflects the overall linear relationship between the predicted and experimental fatigue lives rather than the prediction accuracy across the full fatigue crack growth life. Under four-point bending, the semi-elliptical surface crack did not propagate uniformly along the crack front. As the crack propagated, the remaining uncracked ligament became smaller, causing stresses near the crack front to redistribute and the crack growth rate to accelerate. Because the Paris Law model is based on a single power-law relationship, it shows a relatively constant deviation across the entire fatigue life range. In contrast, the Frost & Pook model captured the accelerated crack growth near the final stage of failure, which explains its closer agreement with the experimental fatigue life after approximately 300,000 cycles. This behavior is further supported by the lower RMSE, NRMSE, and MAPE values obtained for the Frost & Pook model, indicating superior predictive accuracy, particularly near failure.

4. CONCLUSIONS

Fatigue surface cracking is a principal cause of failure in metallic components and structures across various industries. Concerns about fatigue crack growth and structural reliability motivated this investigation into surface crack propagation under cyclic loading. From a practical engineering perspective, accurate fatigue life prediction supports damage-tolerant design and structural integrity management by enabling more reliable maintenance planning and reducing the risk of catastrophic failure. Four fatigue crack growth models Paris law, Frost and Pook, Walker, and Huang Moan were implemented in the fatigue crack analysis of a four-point bending specimen. The predicted fatigue lives were compared with experimental results to assess the performance of each model. The prediction accuracy was evaluated using the MAPE, RMSE, and NRMSE, as presented in Figs. 9–11. The results show that the Frost and Pook model provides the closest agreement with the experimental data [41], producing the lowest MAPE of 16.87%, RMSE of (65.85×10^3) cycles, and NRMSE of 0.1723. In comparison, the Paris law, Walker, and Huang Moan models produce substantially higher errors, with MAPE values exceeding 50% and NRMSE values ranging from 0.4574 to 0.5759. These results indicate lower predictive accuracy and greater deviation from the experimental fatigue life. Overall, the three error metrics consistently demonstrate that the Frost and Pook model provides the most accurate and reliable fatigue life prediction under the investigated four-point bending conditions. Its MAPE remains within the acceptable engineering error range of less than 20%, whereas the other models considerably overestimate the fatigue life. Therefore, fatigue life prediction using the Frost and Pook model may help prevent catastrophic failure by supporting appropriate inspection and maintenance intervals. The findings may also help prioritize critical components for monitoring and improve remaining useful life assessments for structural components subjected to cyclic bending loads.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Ministry of Higher Education, Malaysia, for their invaluable financial support through the Fundamental Research Grant Scheme (FRGS), Grant No. FRGS/1/2023/TK10/UMP/02/26 (RDU230134). Their support has been instrumental in the successful execution of this research project.

REFERENCES

- [1] A. Grbović, G. Kastratović, A. Sedmak, I. Balać, and M. D. Popović, “Fatigue crack paths in light aircraft wing spars,” *Int. J. Fatigue*, vol. 123, pp. 96–104, 2019, doi: <https://doi.org/10.1016/j.ijfatigue.2019.02.013>.
- [2] M. Kikuchi, Y. Wada, and Y. Li, “Crack growth simulation in heterogeneous material by S-FEM and comparison with experiments,” *Eng. Fract. Mech.*, vol. 167, pp. 239–247, 2016, doi: <https://doi.org/10.1016/j.engfracmech.2016.03.038>.
- [3] M. Aursand, P. J. Haagenzen, and B. H. Skallerud, “Remaining fatigue life assessment of corroded mooring chains using crack growth modelling,” *Marine Structures*, vol. 90, p. 103446, 2023, doi: <https://doi.org/10.1016/j.marstruc.2023.103446>.
- [4] H. Okada, H. Kawai, and K. Araki, “A virtual crack closure-integral method (VCCM) to compute the energy release rates and stress intensity factors based on quadratic tetrahedral finite elements,” *Eng. Fract. Mech.*, vol. 75, no. 15, pp. 4466–4485, 2008, doi: <https://doi.org/10.1016/j.engfracmech.2008.04.014>.
- [5] R. Masoudi Nejad and A. Gholami, “Fatigue crack growth and microstructural analysis of rail steel specimens under periodic overloads,” *Mater. Today Commun.*, vol. 39, p. 108592, 2024, doi: <https://doi.org/10.1016/j.mtcomm.2024.108592>.
- [6] T. Wang, Z. Bahrami, G. Renaud, C. Yang, M. Liao, and Z. Liu, “A probabilistic model for fatigue crack growth prediction based on closed-form solution,” *Structures*, vol. 44, pp. 1583–1596, 2022, doi: <https://doi.org/10.1016/j.istruc.2022.08.066>.
- [7] S. O. Afolabi, B. I. Oladapo, C. O. Ijagbemi, A. O. M. Adeoye, and J. F. Kayode, “Design and finite element analysis of a fatigue life prediction for safe and economical machine shaft,” *Journal of Materials Research and Technology*, vol. 8, no. 1, pp. 105–111, 2019, doi: <https://doi.org/10.1016/j.jmrt.2017.10.007>.
- [8] W. Liang, M. Lou, C. Zhang, D. Zhao, D. Yang, and Y. Wang, “Experimental investigation and phenomenological modeling of fatigue crack growth in X80 pipeline steel under random loading,” *Int. J. Fatigue*, vol. 182, p. 108169, 2024, doi: <https://doi.org/10.1016/j.ijfatigue.2024.108169>.
- [9] C. H. Wang and C. N. Duong, “Chapter 6 - Damage tolerance and fatigue durability of scarf joints,” in *Bonded Joints and Repairs to Composite Airframe Structures*, C. H. Wang and C. N. Duong, Eds., Oxford: Academic Press, 2016, pp. 141–172. doi: <https://doi.org/10.1016/B978-0-12-417153-4.00006-2>.
- [10] H. Zhang and P. Qiao, “Virtual crack closure technique in peridynamic theory,” *Comput. Methods Appl. Mech. Eng.*, vol. 372, p. 113318, 2020, doi: <https://doi.org/10.1016/j.cma.2020.113318>.
- [11] M. Goswami, S. Sharma, G. Subbarayan, S. P. A. Bordas, and S. Chattopadhyay, “Chapter Two - Historical purview and recent advances in fracture mechanics of elastomeric matrix composites,” in *Advances in Applied Mechanics*, vol. 56, S. P. A. Bordas, Ed., Elsevier, 2023, pp. 139–187. doi: <https://doi.org/10.1016/bs.aams.2022.09.002>.
- [12] B. Karihaloo and Q. Z. Xiao, “2.03 - Linear and Nonlinear Fracture Mechanics,” in *Comprehensive Structural Integrity*, I. Milne, R. O. Ritchie, and B. Karihaloo, Eds., Oxford: Pergamon, 2003, pp. 81–212. doi: <https://doi.org/10.1016/B0-08-043749-4/02128-5>.
- [13] H. Okada, H. Koya, H. Kawai, Y. Li, and K. Osakabe, “Computations of stress intensity factors for semi-elliptical cracks with high aspect ratios by using the tetrahedral finite element (fully automated parametric study),” *Eng. Fract. Mech.*, vol. 158, pp. 144–166, 2016, doi: <https://doi.org/10.1016/j.engfracmech.2016.02.049>.
- [14] W. Xiang, M. Chen, Y. Tian, and B. Liu, “Application of a hierarchical quadrature element method to interface crack analysis in bi-material systems,” *Eng. Fract. Mech.*, vol. 300, p. 109983, 2024, doi: <https://doi.org/10.1016/j.engfracmech.2024.109983>.
- [15] P. Paris and F. Erdogan, “A critical analysis of crack propagation laws,” 1963.

-
- [16] O. E. Wheeler, "Spectrum loading and crack growth," 1972.
- [17] R. G. Forman, "Study of fatigue crack initiation from flaws using fracture mechanics theory," *Eng. Fract. Mech.*, vol. 4, no. 2, pp. 333–345, 1972, doi: [https://doi.org/10.1016/0013-7944\(72\)90048-3](https://doi.org/10.1016/0013-7944(72)90048-3).
- [18] K. Walker, "The effect of stress ratio during crack propagation and fatigue for 2024-T3 and 7075-T6 aluminum," 1970.
- [19] N. E. Frost, L. P. Pook, and K. Denton, "A fracture mechanics analysis of fatigue crack growth data for various materials," *Eng. Fract. Mech.*, vol. 3, no. 2, pp. 109–126, 1971, doi: [https://doi.org/10.1016/0013-7944\(71\)90003-8](https://doi.org/10.1016/0013-7944(71)90003-8).
- [20] T. Wang, Z. Bahrami, G. Renaud, C. Yang, M. Liao, and Z. Liu, "A probabilistic model for fatigue crack growth prediction based on closed-form solution," *Structures*, vol. 44, pp. 1583–1596, 2022, doi: <https://doi.org/10.1016/j.istruc.2022.08.066>.
- [21] M. R. M. Akramin, M. S. Marizi, M. N. M. Husnain, and M. Shamil Shaari, "Analysis of Surface Crack using Various Crack Growth Models," *J. Phys. Conf. Ser.*, vol. 1529, no. 4, p. 042074, 2020, doi: 10.1088/1742-6596/1529/4/042074.
- [22] W. J. Wang and M.-S. Yim, "A fatigue crack growth prediction model for cracked specimen under variable amplitude loading," *Int. J. Fatigue*, vol. 168, p. 107387, 2023, doi: <https://doi.org/10.1016/j.ijfatigue.2022.107387>.
- [23] V. Ribeiro et al., "Fatigue crack growth modelling by means of the strain energy density-based Huffman model considering the residual stress effect," *Eng. Fail. Anal.*, vol. 140, p. 106543, 2022, doi: <https://doi.org/10.1016/j.engfailanal.2022.106543>.
- [24] N. Pugno, M. Ciavarella, P. Cornetti, and A. Carpinteri, "A generalized Paris' law for fatigue crack growth," *J. Mech. Phys. Solids*, vol. 54, no. 7, pp. 1333–1349, 2006, doi: <https://doi.org/10.1016/j.jmps.2006.01.007>.
- [25] V. Ribeiro, J. Correia, G. Lesiuk, A. Gonçalves, A. De Jesus, and F. Berto, "Application and discussion of various crack closure models to predict fatigue crack growth in 6061-T651 aluminium alloy," *Int. J. Fatigue*, vol. 153, p. 106472, 2021, doi: <https://doi.org/10.1016/j.ijfatigue.2021.106472>.
- [26] M. F. V Montezuma, E. P. De Deus, C. O. F. T. Rüdert, M. C. De Carvalho, and M. A. E Silva Filho, "Numerical-experimental procedure for predicting fatigue life in SAE AMS 7475-T7351 aluminum alloy considering the effect of stress ratio," *An. Acad. Bras. Cienc.*, vol. 96, 2024, doi: 10.1590/0001-3765202420231400.
- [27] Y. Li, H. Wang, and D. Gong, "The interrelation of the parameters in the Paris equation of fatigue crack growth," *Eng. Fract. Mech.*, vol. 96, pp. 500–509, 2012, doi: 10.1016/j.engfracmech.2012.08.016.
- [28] Y. Li and J. Jia, "A modified walker's equation in fatigue crack growth," *Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University*, vol. 26, no. 2, pp. 221–224, 2008, [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-44849123145&partnerID=40&md5=a11aa8b7e0638602de90c8576fa4c368>
- [29] P.-T. Guan, C.-Y. Lü, X.-T. Tang, J.-W. Zhu, Z.-Y. Ding, and Z.-L. Gao, "Study on fatigue crack growth behavior of q345r steel under multiple overloads," *Gongcheng Lixue/Engineering Mechanics*, vol. 34, no. 7, pp. 224–231 and 240, 2017, doi: 10.6052/j.issn.1000-4750.2016.02.0113.
- [30] B. K. C. Yuen and F. Taheri, "Proposed modifications to the Wheeler retardation model for multiple overloading fatigue life prediction," *Int. J. Fatigue*, vol. 28, no. 12, pp. 1803–1819, 2006, doi: 10.1016/j.ijfatigue.2005.12.007.
- [31] K. Tanaka, "4.04 - Fatigue Crack Propagation," in *Comprehensive Structural Integrity*, I. Milne, R. O. Ritchie, and B. Karihaloo, Eds., Oxford: Pergamon, 2003, pp. 95–127. doi: <https://doi.org/10.1016/B0-08-043749-4/04089-1>.
-

- [32] M. Shaari, A. K. Ariffin, A. Takahashi, S. Abdullah, M. Kikuchi, and M. R. M. Akramin, "Fatigue crack growth analysis on square prismatic with embedded cracks under tension loading," *JOURNAL OF MECHANICAL ENGINEERING AND SCIENCES*, vol. 14, pp. 2511–2525, Mar. 2017, doi: 10.15282/jmes.11.1.2017.10.0231.
- [33] H. Okada, M. Higashi, M. Kikuchi, Y. Fukui, and N. Kumazawa, "Three dimensional virtual crack closure-integral method (VCCM) with skewed and non-symmetric mesh arrangement at the crack front," *Eng. Fract. Mech.*, vol. 72, no. 11, pp. 1717–1737, 2005, doi: <https://doi.org/10.1016/j.engfracmech.2004.12.005>.
- [34] D. Broek, *Elementary engineering fracture mechanics*. Springer Netherlands, 1982. [Online]. Available: <https://books.google.com.my/books?id=LC4E7OjnhuEC>
- [35] G. Chell, "Developments in fracture mechanics," 1979. [Online]. Available: <https://api.semanticscholar.org/CorpusID:137575003>
- [36] M. Xinting, L. Guoxu, P. Jian, T. Ping, Z. Jinbo, and L. Jian, "Study on the I-II mixed mode fatigue crack growth behavior of TC4 titanium alloy," *Theoretical and Applied Fracture Mechanics*, vol. 138, p. 104985, 2025, doi: <https://doi.org/10.1016/j.tafmec.2025.104985>.
- [37] M.-S. Lee and M.-H. Kim, "A new fatigue crack growth model considering underloads and overloads history together with stress ratio," *International Journal of Naval Architecture and Ocean Engineering*, vol. 14, p. 100481, 2022, doi: <https://doi.org/10.1016/j.ijnaoe.2022.100481>.
- [38] J. Cong, Z. Liu, S. Zhou, S. Gao, and X. Zhu, "Predicting fatigue life and crack growth rate of TC4 titanium alloy based on PINN before and after ultrasonic impact treatment," *Eng. Fail. Anal.*, vol. 166, p. 108875, 2024, doi: <https://doi.org/10.1016/j.engfailanal.2024.108875>.
- [39] M. S. Shaari, M. R. M. Akramin, A. K. Ariffin, S. Abdullah, and M. Kikuchi, "Fatigue crack growth behaviour of semi-elliptical surface cracks for an API 5L X65 gas pipeline under tension," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 308, no. 1, p. 012041, 2018, doi: 10.1088/1757-899X/308/1/012041.
- [40] F. Y. Chin, M. S. Shaari, A. Takahashi, M. R. M. Akramin, and S. Urai, "Fatigue Crack Growth Behavior of AlSi10Mg Material Produced by Direct Metal Laser Sintering Using S-Version Finite Element Method," *Journal of Failure Analysis and Prevention*, vol. 23, no. 2, pp. 601–608, 2023, doi: 10.1007/s11668-023-01643-6.
- [41] M. R. M. Akramin, A. K. Ariffin, M. Kikuchi, and S. Abdullah, "Sampling method in probabilistic S-version finite element analysis for initial flaw size," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 39, no. 1, pp. 357–365, 2017, doi: 10.1007/s40430-016-0549-z.

Digital Twin Framework for Integrated Sequential Laser Beam Micromachining (LBMM) and Micro-EDM

WAN AHMAD BIN WAN AZHAR¹, TANVEER SALEH^{1*},
KHAIRUL AFFENDY BIN MD NOR², MUHAMMAD MUKHTAR BIN NOOR
AWALLUDIN³, ABD HALIM BIN EMBONG¹, ALI SOPHIAN¹,
MOHD ASYRAF BIN MOHD RAZIB¹, AZHAR BIN MOHD IBRAHIM²

¹*Autonomous Systems and Robotics Research Unit, Department of Mechatronics Engineering, International Islamic University Malaysia, 53100, Jalan Gombak, Kuala Lumpur, Malaysia*

²*Department of Mechatronics Engineering, International Islamic University Malaysia, 53100, Jalan Gombak, Kuala Lumpur, Malaysia*

³*Department of Manufacturing and Materials Engineering, International Islamic University Malaysia, 53100, Jalan Gombak, Kuala Lumpur, Malaysia*

*Corresponding author: tanveers@iium.edu.my

(Received: 17 March 2026; Accepted: 30 July 2026; Published online: 11 September 2026)

ABSTRACT: Micro-hole arrays are widely used in aerospace, biomedical devices, and filtration applications. While many methods exist for producing these arrays, achieving high precision, efficiency, and low cost in a single process remains challenging. An integrated sequential method combining Laser Beam Micromachining (LBMM) for fast material removal and micro Electrical Discharge Machining (μ EDM) for precise finishing offers a promising solution. However, the sequential operation of these two processes presents challenges such as collision risk, coordination difficulties, and monitoring limitations. To address these issues, this paper proposes a Digital Twin (DT) framework integrated with the Internet of Things (IoT) for real-time analysis and decision-making, which improves process coordination, collision-aware operation, and virtual validation of sequential LBMM- μ EDM machining. The DT enables real-time monitoring by leveraging accurate 3D geometries of machine components to enhance the realism and precision of collision detection. The virtual 3D representation was developed using Unity3D software, while IoT connectivity for real-time data exchange was established through Firebase, resulting in an average transmission latency of 16 ms. The DT eliminates the need for live camera feeds by enabling flexible, low-bandwidth, multi-angle visualization without extra hardware, while also supporting offline virtual simulations for motion coordination, realistic cutting effects, and machining time estimation with a 79.8% prediction accuracy.

ABSTRAK: Susunan lubang mikro digunakan secara meluas dalam pelbagai aplikasi seperti aeroangkasa, peranti bioperubatan dan sistem penapisan. Walaupun pelbagai kaedah pemesinan telah dibangunkan untuk menghasilkan susunan lubang ini, pencapaian ketepatan tinggi, kecekapan proses serta kos pengeluaran yang rendah secara serentak masih menjadi satu cabaran. Kaedah pemesinan berjujukan yang menggabungkan pemesinan mikro pancaran laser bagi menyingkirkan bahan secara pantas dan pemesinan nyahcas elektrik mikro bagi membuat kemasan berketepatan tinggi, berpotensi menjadi satu penyelesaian berkesan. Namun, pelaksanaan kedua-dua proses secara berjujukan menimbulkan beberapa cabaran seperti risiko perlanggaran antara komponen mesin, kesukaran dalam penyelarasan proses serta keterbatasan pemantauan operasi. Sehubungan itu, kajian ini mencadangkan satu rangka kerja kembar digital yang diintegrasikan bersama teknologi Internet Kebendaan bagi menyokong analisis masa nyata dan pembuatan keputusan, membolehkan penyelarasan proses dengan lebih baik, operasi berasaskan pengesanan perlanggaran, serta pengesanan maya bagi

pemesinan berjujukan LBMM- μ EDM. Rangka kerja kembar digital yang dicadangkan membolehkan pemantauan masa nyata melalui penggunaan model geometri tiga dimensi komponen mesin yang tepat, sekaligus meningkatkan tahap realisme dan ketepatan dalam pengesanan perlanggaran. Persekitaran maya tiga dimensi dibangun menggunakan perisian pembangunan grafik tiga dimensi, manakala sambungan Internet Kebendaan bagi pertukaran data masa nyata dilaksanakan melalui pangkalan data awan, dengan purata kelewatan penghantaran data direkodkan sebanyak 16 milisaat. Selain itu, sistem kembar digital yang dibangunkan dapat menggantikan keperluan penggunaan kamera langsung dengan menyediakan visualisasi berbilang sudut yang fleksibel serta penggunaan jalur lebar yang rendah tanpa memerlukan perkakasan tambahan. Sistem ini juga menyokong simulasi maya secara luar talian bagi tujuan penyelarasan pergerakan mesin, visualisasi kesan pemotongan yang realistik serta penganggaran masa pemesinan dengan ketepatan ramalan sebanyak 79.8%.

KEYWORDS: *Digital Twin; Sequential Hybrid Machining, Laser Beam Micromachining (LBMM); Micro Electrical Discharge Machining (μ EDM); Real-time Monitoring; Collision Detection.*

1. INTRODUCTION

In advanced engineering applications, a micro hole array is a patterned grid or arrangement of repetitive micro-scale holes created on a material surface to serve structural, fluidic, optical, or functional purposes. They play a vital role in numerous industries because of their versatile functional advantages. In the aerospace sector, they are used to cool turbine blades under high temperatures [1]. In water filtration systems, these arrays improve the detection of trace heavy metals [2]. In the field of microfluidics, ceramic-based arrays facilitate accurate fluid control [3].

Despite their importance, producing high-precision, repeatable micro-hole arrays remains a manufacturing challenge. Various fabrication methods are used to create these structures, including electrochemical machining (ECM), micro-EDM, laser drilling, and mechanical punching. ECM provides high surface quality without tool wear but requires precise parameter control [4]. Micro-EDM (μ EDM) is suitable for hard, conductive materials and achieves excellent accuracy, though it is slow and limited in material types [5]. Laser drilling offers fast, contactless processing but can cause thermal damage [6]. Mechanical punching is productive but struggles with miniaturization and alignment [7]. The choice of method depends on the material, feature size, and intended application.

Researchers have investigated various approaches to enhance micromachining efficiency by integrating different machining techniques. The core concept of hybrid machining involves combining two or more machining methods or principles to overcome the limitations of standalone processes and achieve superior component quality and performance. For instance, in laser-assisted micro-milling, laser preheating softens the workpiece, thereby improving material removal rates up to six times and extending tool lifespan [8]. One of the most promising hybrid approaches combines laser beam micromachining (LBMM) with micro-EDM (μ EDM) in sequence. In this method, LBMM is used for rapid roughing to reduce material volume, minimize thermal damage, and limit recast-layer formation, followed by μ EDM for fine surface quality and precise dimensions. Studies have shown that this combined method reduces processing time for fuel injector nozzles by up to 70% while maintaining high hole quality at diameters below 145 μm [9]. In some cases, hole drilling times dropped by 90% for 125–130 μm diameter holes compared to conventional μ EDM alone [10].

Despite the advantages of hybrid or sequential machining processes, several challenges remain, with collision risk during setup and operation being a major concern. Hybrid setups often involve multiple machining heads, increasing the potential for collisions. This risk is especially significant in computer-aided manufacturing (CAM), where simulations may not account for all physical components such as clamps or vices [11]. Some CAM software simulates only toolpaths or material removal, ignoring fixture geometry, which can result in undetected collisions during actual machining. Chen et al. (2017) also noted that in hybrid additive–subtractive processes, improperly planned material heights can prevent cutting tools from reaching certain areas, causing collisions [12]. This issue is particularly critical in sequential LBMM– μ EDM systems, where multiple heads operate in a confined space alongside components such as the workpiece holder and dielectric tank.

Another critical challenge in hybrid machining is coordinating and sequencing different machining stages. Basinger et al. (2019) highlighted this issue in their development of the Feature-based Advanced Hybrid Process Planning System (FAH-PS), which automates process planning using feature-specific geometric and material data [13]. Their study showed that manual process planning often takes weeks because of these coordination complexities, whereas the FAH-PS achieved a 35% reduction in machining time in a patient-specific bone plate case study. Furthermore, the complexity of coordination and sequencing in hybrid machining also affects operator training. Operators need to understand both additive and subtractive processes, which makes the learning process longer and more difficult [14]. In the sequential LBMM– μ EDM system, operators must be well trained in both processes, as any mistake in sequencing, tool alignment, or parameter setup can affect machining accuracy and overall process efficiency.

Monitoring machining processes using live camera feeds presents several limitations. Although they provide real-time visuals, cameras are restricted to fixed viewpoints. They are often obstructed by cutting fluids or debris, making it challenging to inspect hidden or critical areas during operation [15]. Expanding camera coverage requires additional hardware, increasing both system complexity and cost. Furthermore, live video streaming demands high bandwidth and may experience latency or degraded quality, especially under high-resolution or remote monitoring conditions [16]. These challenges highlight the need for advanced solutions such as Digital Twin (DT) technology to enhance monitoring, coordination, sequencing, and collision detection in sequential LBMM– μ EDM processes.

A Digital Twin (DT) is a virtual representation of a physical system that continuously mirrors its real-time behavior using 3D models, positional data, sensor information, and control parameters. Unlike a digital model, which is static and requires manual updates, or a digital shadow, which supports only one-way data flow from the physical to the virtual system, a DT enables bi-directional communication, allowing both physical and virtual entities to exchange information in real time for predictive analysis, optimization, and autonomous decision-making [17]–[19]. In manufacturing, DT provides a comprehensive 3D visualization of the machining environment, including machine components, tool heads, workpieces, and fixtures, enabling process planning, collision detection, operation sequencing, and performance optimization. When integrated with the Internet of Things (IoT), DT also supports remote monitoring and virtual operator training without requiring access to the physical machine. DT platforms are commonly developed using tools such as Siemens TIA Portal, Unity, Game4Automation, MATLAB/Simulink, and ANSYS Twin Builder for real-time visualization, control integration, and physics-based simulation.

A recent DT application in machining is the prediction and compensation of thermal-induced time-varying errors in CNC machine tool axes, where a time-varying error model

based on heat transfer theory and a visual model was used [20]. This approach effectively predicted and compensated for errors, reducing hole-pitch fluctuation by 69.19%, demonstrating the potential of DT in enhancing machining accuracy and performance. In contrast, a novel DT framework was developed to support offline robotic EDM cutting-path programming by simulating tool kinematics and addressing complex issues such as collisions and singularities. This approach enhances the integration of industrial robots with EDM for the precise machining of hard-to-cut materials [21].

This paper proposes a Digital Twin (DT) framework integrated with the Internet of Things (IoT) for real-time analysis and decision-making. This study aims to develop a DT framework integrated with IoT technology to support sequential machining processes, improving process coordination, collision-aware operation, and virtual validation of sequential LBMM- μ EDM machining. The framework addresses critical challenges in hybrid and sequential machining, particularly collision risk, process coordination, and real-time monitoring. By leveraging 3D geometric models of the LBMM head, μ EDM head, and surrounding fixtures, the DT enables real-time monitoring and enhances collision detection throughout the machining operations, without relying on live camera systems. In addition, the framework simulates coordination and sequencing between different machining stages, enabling users to plan and monitor operations more efficiently. This work also contributes to operator training by offering a virtual platform for observing and understanding process flows and interactions, eliminating the need for physical machines or on-site presence.

2. DIGITAL TWIN SYSTEM ARCHITECTURE

This section details the system architecture and framework of the digital twin for the sequential LBMM- μ EDM process, including the physical machine, related software, IoT implementation, DT virtual representation, and the Unity-based user interface.

2.1. Integrated LBMM- μ EDM Physical Setup

The physical setup integrates LBMM and μ EDM processes into a single system, as illustrated in Figure 1. In this configuration, the LBMM module is mounted horizontally offset from the μ EDM unit, while the workpiece remains fixed within the μ EDM workspace and is accessible to both processes. The LBMM system employs a 50W fiber laser (1060 nm wavelength, 50 μ m spot size, 200 mm focal length) driven by an X/Y galvanometer scanner with a $\pm 15^\circ$ scan range and 12 μ rad resolution. Laser output was calibrated using a Gentec Pronto-250 power meter, which records the average output power. Calibration was performed from 10% to 100% laser setting. Although rated at 50 W, the measured output at 100% was approximately 38.6 W, indicating a power loss of about 22.8%. This reduction is consistent with the laser supplier's specifications and is attributed to transmission losses through multiple optical components, such as collimators, beam expanders, mirrors, and the galvanometric scanner. Each optical component contributes to reflection, scattering, or absorption losses, reducing the power delivered to the workpiece surface. A calibration curve was generated to correlate the percentage setting with the actual delivered power. For this experiment, the laser was operated at 90%, corresponding to a calibrated output of 35.4 W. The laser head is height-adjustable via a rotary mechanism and traverses horizontally on a motorized linear stage, while EZCAD Lite software controls the process parameters.

The μ EDM subsystem consists of a three-axis CNC machine with a repeatability of 15 μ m, equipped with a μ EDM spindle and a microscopic camera mounted on the Z-axis plate. Custom control software provides camera visualization, parameter configuration, G-code

execution, and real-time axis monitoring. The process flow and sequence are illustrated in Figure 2. Initially, as shown in Figure 2(a), the LBMM head extends over the workpiece to perform laser machining while the μ EDM head remains retracted in a safe position. Once the LBMM process is completed, the laser head retracts to allow for the subsequent μ EDM operation.

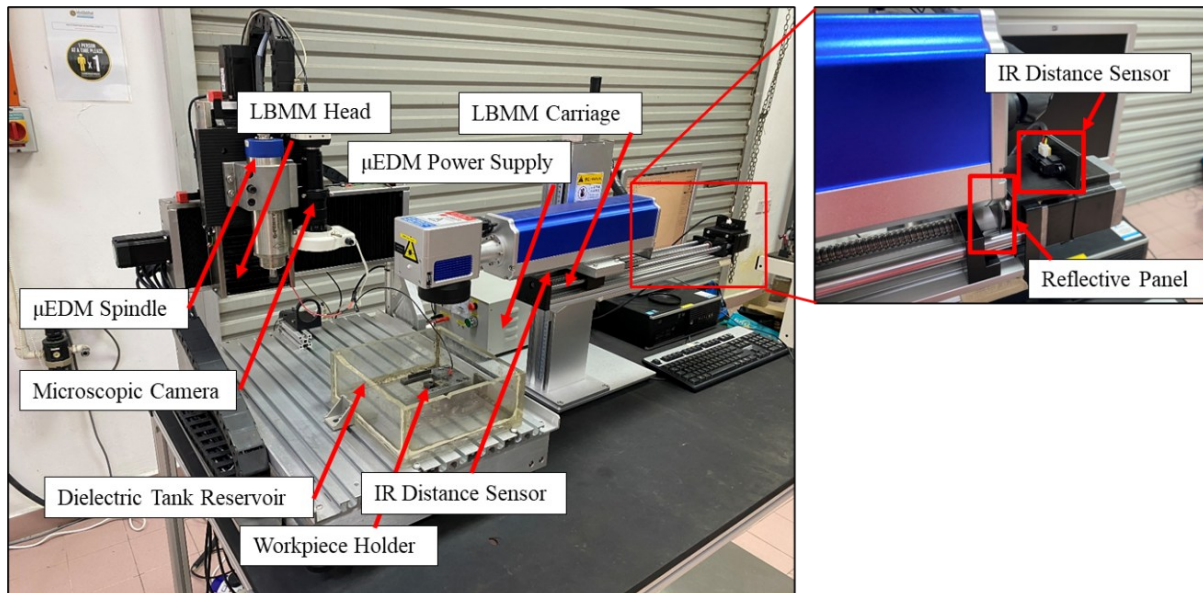


Figure 1. Physical setup of the hybrid Laser Beam Micromachining (LBMM) and micro-EDM (μ EDM) system, showing the integrated configuration used for sequential machining operations on a shared workpiece.

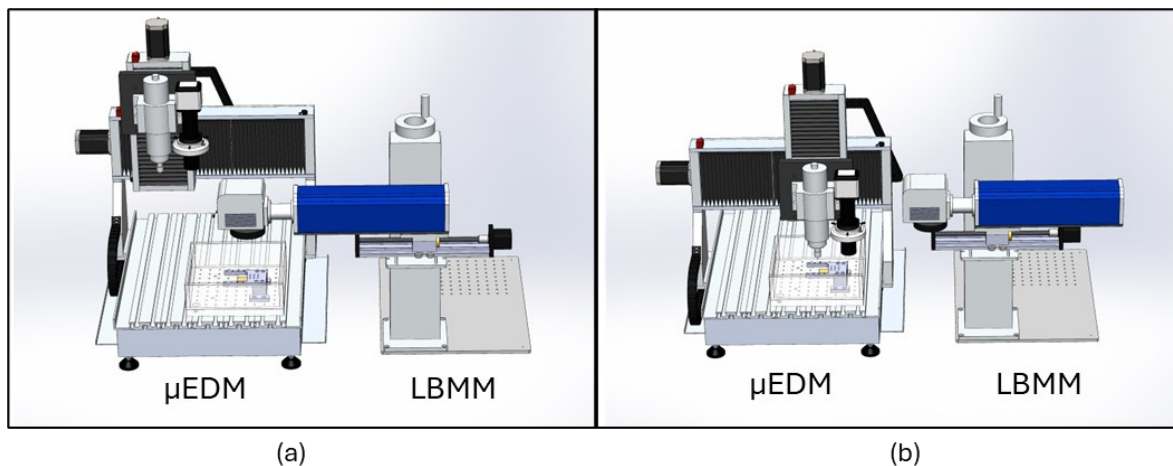


Figure 2. Machine Configuration (3D model):
(a) During the LBMM process; (b) During the μ EDM process.

The position of the LBMM carriage is monitored using an infrared (IR) distance sensor (model: GP2D120XJ00F), as illustrated in Figure 1, which operates effectively within a measurement range of 0–30 cm. The sensor has a typical response time of approximately 39 ms, with output sensitivity decreasing from ~ 0.12 V/mm at shorter distances to ~ 0.025 V/mm at longer distances. Next, the μ EDM head, equipped with a spindle, a microscopic camera, and a ring light, advances toward the workpiece. This operation is illustrated in Fig.2(b). The microscopic camera that is mounted on the Z-axis is used to localize the LBMM-machined holes and determine their center coordinates, as described in a previous study [22]. After

verifying the coordinates, μ EDM power supply parameters such as gap voltage, capacitance, discharge and short thresholds, and polarity are configured. Subsequently, the software automatically generates the G-code for the μ EDM fine finishing process, enabling a fully coordinated and automated machining sequence. The changeover between the two processes takes approximately 36 seconds. The LBMM and μ EDM Parameters, and IR sensor specifications are tabulated in Table 1.

Table 1. LBMM and μ EDM Parameters, and IR Sensor specifications.

LBMM Parameters	
Loop Count	75
Programmed Diameter (μm)	180
Scanning Speed (mm/s)	10
Laser Rated Power (W)	50
Actual Measured Laser Power (W)	38.6
Applied Laser Power (W)	35.4 (or 90%)
Frequency (Khz)	50
Hatch Angle ($^{\circ}$)	0
Line Space (mm)	0.01
Focal Spot Size (μm)	50
Hatch Type	Two-way Hatch 
μ EDM Parameters	
Power Supply Type	Resistor- Capacitor (RC)
Resistance ($\text{K}\Omega$)	1
Electrode Material	Tungsten
Electrode Diameter (μm)	300
Workpiece Material	Stainless Steel (SUS304)
• Thermal Conductivity (W/m-K)	16.2
• Melting Point ($^{\circ}\text{C}$)	1400 - 1455
• Electrical Resistivity ($\mu\Omega\cdot\text{m}$)	0.72
Workpiece Thickness (μm)	200, 400 and 600
Gap Voltage (V)	90
Capacitance (nF)	10 (1.8 Stray Cap)
Discharge Energy (μJ)	40.5
Spindle Speed (RPM)	1200
Feedrate ($\mu\text{m/s}$)	5
Flushing Method	Atmospheric Pressure
Dielectric Fluid	Mineral-based EDM oil
Polarity	Workpiece: Positive Tool: Negative

For validation purposes, a stainless steel (SS304) plate with dimensions of 20 mm $\times$ 20 mm and thicknesses of 200 μm and 600 μm was used. A cylindrical tungsten electrode with a 300 μm diameter was chosen because of its high thermal wear resistance and low tool wear rate. The machining process involved creating an array of three LBMM pilot micro-holes in a single row, evenly spaced 1 mm apart, followed by fine finishing of the through-holes using μ EDM with a 300 μm diameter tungsten electrode (as available in the lab). The LBMM pilot-hole size plays an important role in achieving efficient and high-quality machining. A hole that is too small restricts tool access during μ EDM, increasing machining time, while an overly large hole increases the extent of HAZ, recast layer, and spatter, which can degrade final feature

quality. To balance these factors, preliminary trials were conducted to identify an appropriate pilot-hole size that provides both effective LBMM performance and efficient μ EDM finishing. The LBMM process parameters were selected based on the optimization performed in our previous study [23]. Each LBMM pilot hole had an average entry and exit diameter of approximately 200 μm and 100 μm , respectively, residual spatter of 280 μm , and a HAZ extending up to 470 μm .

2.2. Integrated LBMM- μ EDM Physical Setup

Figure 3 illustrates the Digital Twin (DT) framework developed for the unified Laser Beam Micromachining (LBMM) and Micro Electrical Discharge Machining (μ EDM) setup. In this integrated system, the LBMM process is managed through EZCAD software, while the μ EDM process is controlled via custom-developed μ EDM CNC controller software. Both systems communicate through serial interfaces, enabling coordinated operation and seamless data exchange.

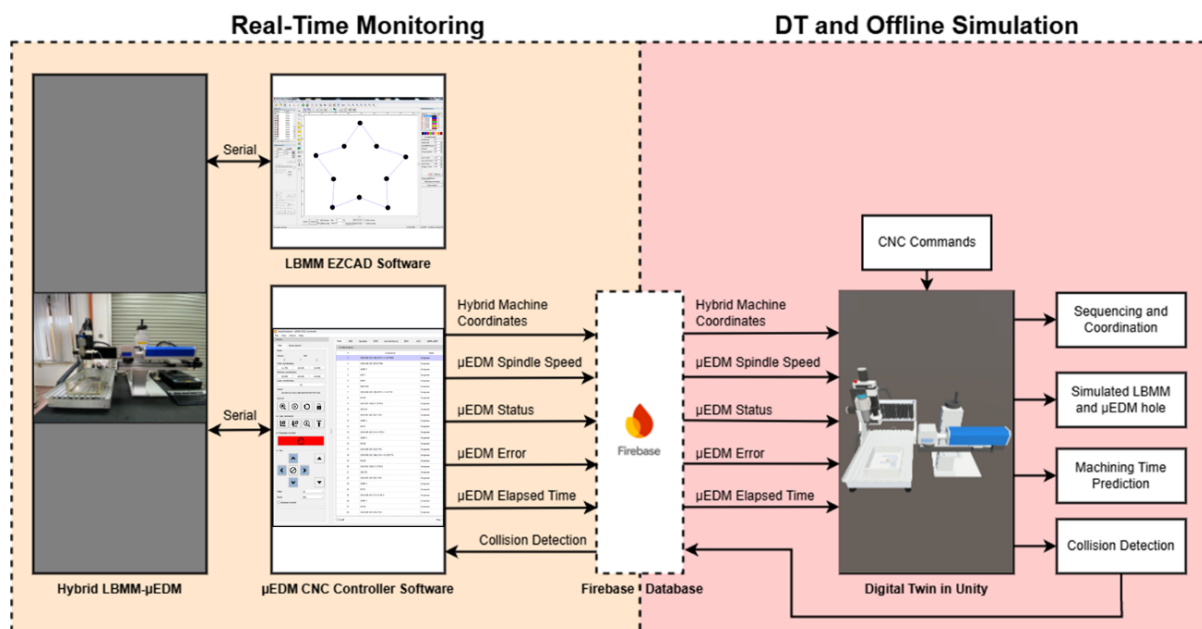


Figure 3. Digital Twin (DT) framework for the unified LBMM- μ EDM setup, illustrating both real-time monitoring and offline simulation modes.

For the LBMM process, EZCAD software controls key laser parameters such as scanning speed, frequency, loop count, LBMM carriage position, and other settings crucial to achieving the desired machining accuracy. In addition, an infrared (IR) distance sensor tracks the position of the LBMM carriage. This positional data is sent to the μ EDM CNC Controller, enabling synchronization between the LBMM and μ EDM stages for improved coordination in the integrated setup.

For the μ EDM CNC machine, a customized control software named *the μ EDM CNC Controller* was developed by modifying the open-source *CANDLE* software, written in C++ using the Qt5 framework. This software has been extensively enhanced to manage μ EDM power supply parameters, including gap voltage, RC circuit capacitance, discharge and short-circuit thresholds, and voltage polarity. In addition, it controls the CNC motion system and implements an automatic retraction mechanism that holds the position upon discharge and retracts to clear shorts when detected. The μ EDM CNC controller executes G-code toolpaths and includes real-time X, Y, and Z plotting for process monitoring. Additionally, on-machine

measurement (OMM) and automation features were integrated to streamline the transition between LBMM and μ EDM processes, enabling seamless and efficient sequential operation.

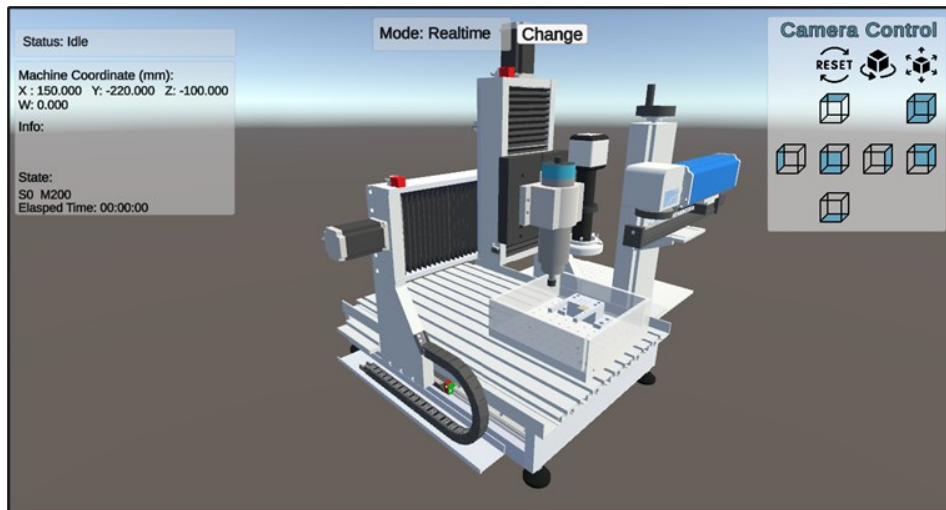
Furthermore, Firebase Realtime Database integration was implemented to support bidirectional data exchange, enabling real-time remote monitoring and seamless synchronization with the digital twin environment in Unity. This data includes the machine coordinates, μ EDM spindle speed, μ EDM status and error messages, and μ EDM elapsed time.

Unity3D, functioning as the DT platform, continuously receives and processes real-time data from the μ EDM CNC system. One critical DT function is collision detection, which is available in both real-time and simulation modes. When a potential collision is detected based on position and the 3D model, the system sends a warning signal back to Firebase. The μ EDM CNC controller monitors this signal and promptly halts machine motion to prevent any damage. For safety, no control commands are sent through Firebase, and only monitoring and warning functions are enabled. This ensures that the μ EDM CNC controller software remains securely isolated from unauthorized or accidental external operations. A subsequent section explains the detailed logic for collision detection, collider setup, and Unity implementation.

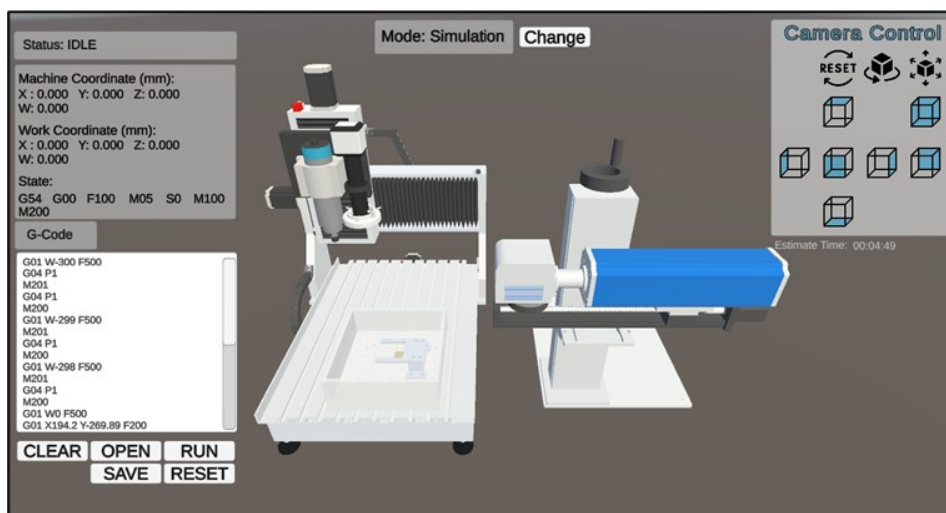
The Digital Twin for the sequential LBMM- μ EDM system operates in two modes: real-time mode for live monitoring and collision detection, and simulation mode for offline process planning and visualization. In real-time mode, Figure 4(a) depicts the Unity3D interface with a live 3D visualization of the machine. The machine's position is dynamically updated based on data received from Firebase, allowing users to accurately monitor tool and workpiece movement. A status panel shows the current process state, error messages, and elapsed machining time. A camera control panel enables users to switch between predefined views, zoom in and out, and rotate the scene to observe operations from different perspectives. This flexibility makes it useful for monitoring and offers advantages over live camera feeds, which are limited by fixed viewpoints and potential obstructions. Real-time collision warnings are indicated visually when a collision is detected.

The system also includes a simulation mode, as illustrated in Figure 4(b), for offline testing and process planning. In this mode, users can input CNC commands using a dedicated G-code panel, with options to write, load, and save code files. Once the simulation runs, the virtual machine follows the programmed toolpath, showing animated motion and hole-generation sequences for both LBMM and μ EDM. The visual logic for the hole-cutting simulation is detailed in the following section. A reset button allows users to restart the simulation. The system calculates and displays estimated machining time based on feedrate and delay values defined in the G-code. However, for a more realistic estimation, the simulation incorporates actual μ EDM Z-displacement data to account for the additional μ EDM processing time.

The DT is entirely software-based and requires no additional hardware. All machines have existing 3D models that can be directly imported into Unity3D for simulation and visualization. Unity3D is the environment used to develop the digital twin (DT). While the Personal plan is free for individual use, the Pro plan is available at USD 200 per month per seat for professional development. The system also requires an internet connection and uses Firebase for data storage and communication. Firebase pricing depends on the selected plan and usage level. Under typical operating conditions, expenses are minimal, as the free Spark plan supports moderate data storage and real-time updates, while the Blaze plan offers a pay-as-you-go option for higher usage. Overall, this approach is highly cost-effective compared to hardware-based solutions.



(a)



(b)

Figure 4. GUI Interfaces of the Digital Twin System. (a) Real-time mode GUI displaying live machine position updates from Firebase, μ EDM process status, error alerts, elapsed time, and camera control panel for view navigation. (b) Simulation mode GUI featuring G-code editor, run/reset controls, estimated machining time display, and animated toolpath with hole cutting simulation for both LBMM and μ EDM processes.

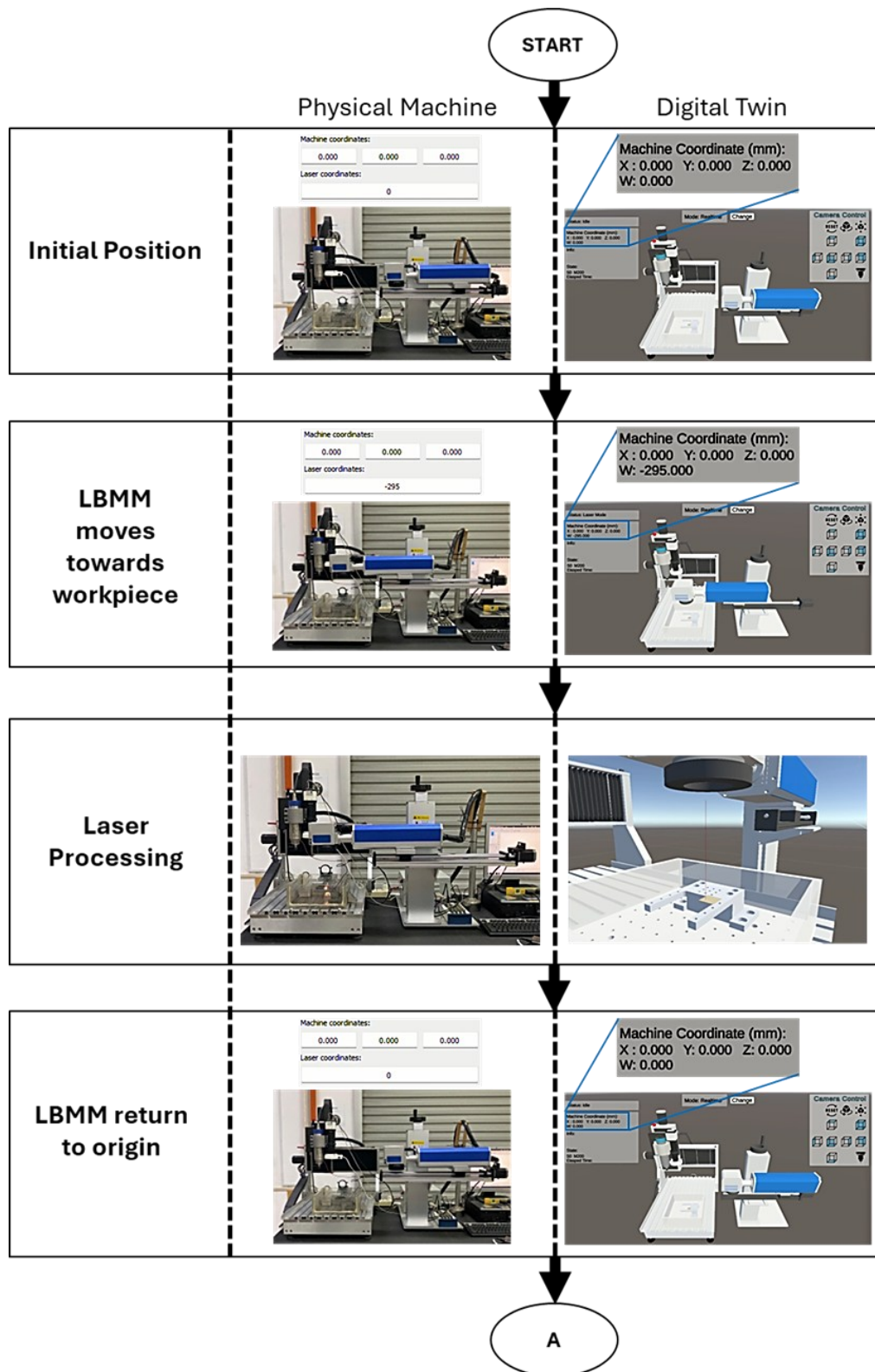
3. DIGITAL TWIN FOR REAL-TIME MONITORING WITH COLLISION DETECTION

This section presents the implementation of a Digital Twin for real-time monitoring, including online simulation and integrated collision detection. It is divided into three parts: Section 3.1 describes real-time monitoring, Section 3.2 describes the setup and components of the collision detection system, and Section 3.3 discusses latency results and system response, evaluating the effectiveness and responsiveness of the safety mechanism during operation.

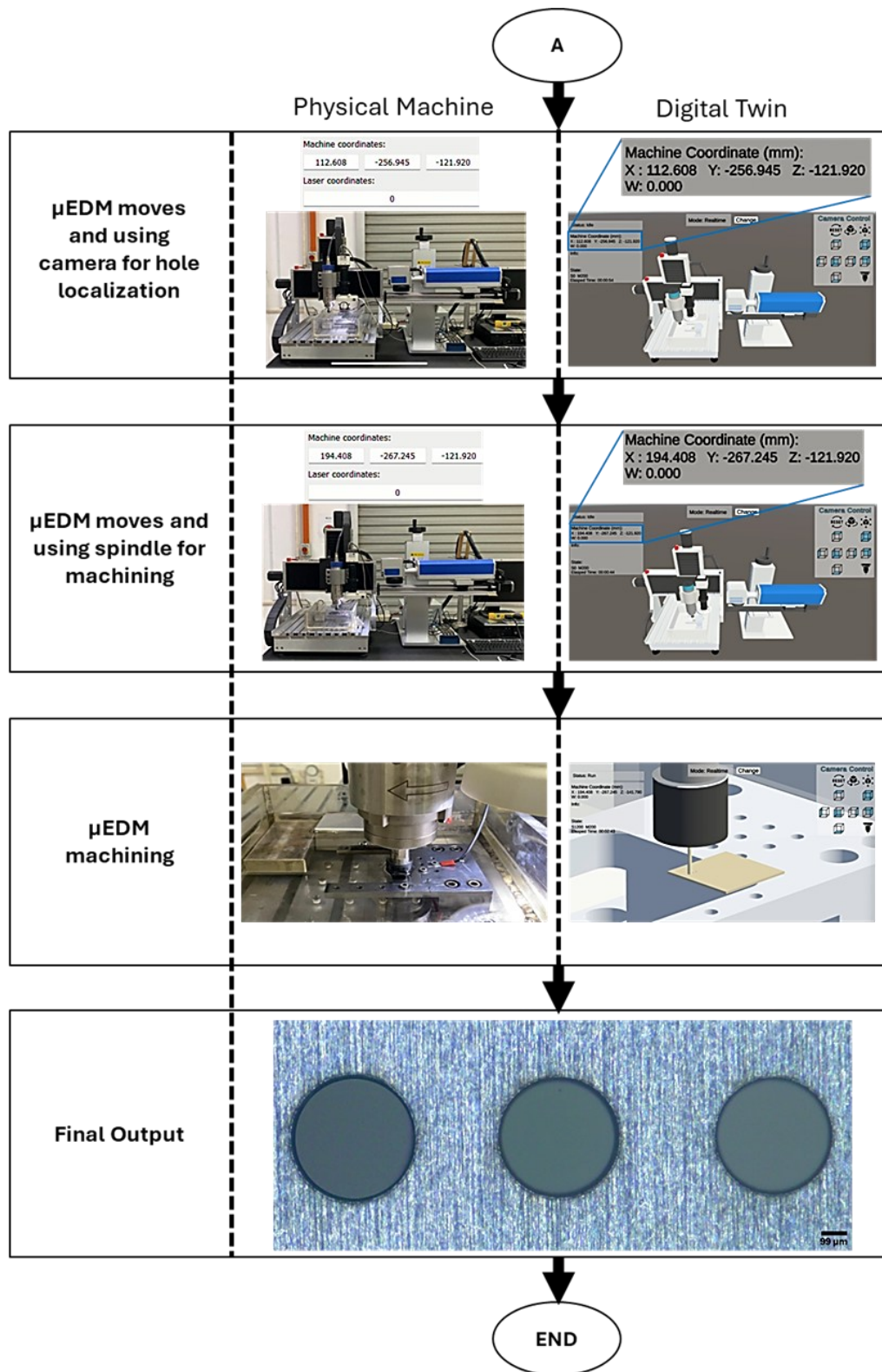
3.1. Real-time Monitoring

To demonstrate the effectiveness of the DT during actual LBMM- μ EDM operation, a machining experiment was performed. In this test, the LBMM created 200 μ m pilot holes in a 200 μ m-thick stainless steel (SS304) workpiece, which were subsequently enlarged and fine-

finished to 300 μm using the μEDM process. A total of three holes were machined. The complete sequence of the physical machine workflow and its corresponding DT operations is illustrated in Figure 5.



(a)



(b)

Figure 5. Flowchart of the integrated sequential LBMM- μ EDM machining process, showing both the physical machine operations and the corresponding Digital Twin (DT) updates: (a) Initial position and LBMM operations, (b) μ EDM operations, including hole localization, spindle positioning, and final output.

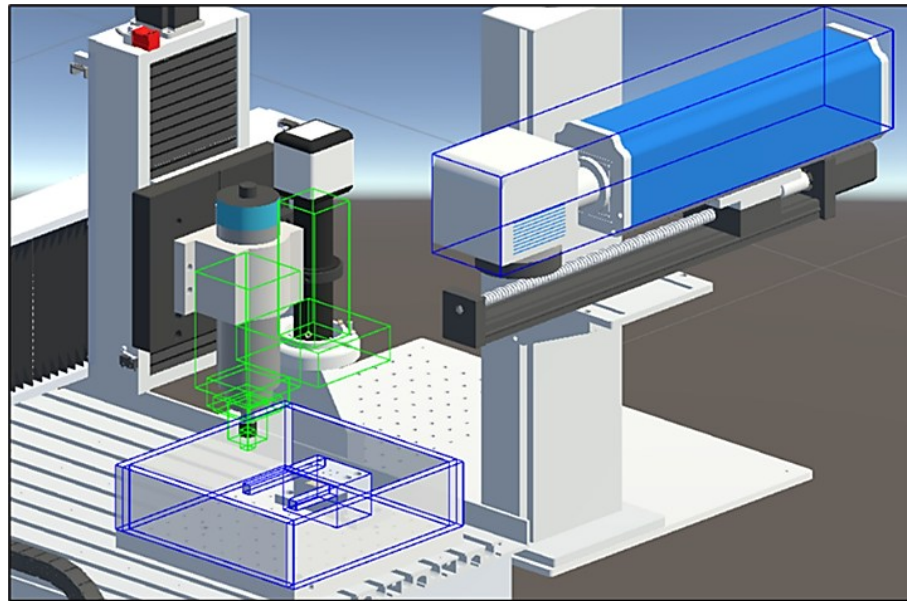
The process begins with both modules at their initial positions, as illustrated in Figure 5(a). The LBMM head then moves toward the workpiece and performs the laser machining operation. After completing laser processing, the LBMM returns to its origin to allow the μ EDM head to proceed. The μ EDM head unit subsequently moves forward, where the camera system is utilized to localize the LBMM-drilled pilot hole and determine its precise coordinates for the fine-finishing stage. The μ EDM Z-axis is equipped with both a camera for LBMM hole localization and a spindle for the actual μ EDM machining, each serving a distinct function. Once localization is completed, the μ EDM head moves to the identified position, and the μ EDM process is carried out, as shown in Figure 5(b). The DT continuously tracks machine motions and matches the coordinate values between the μ EDM CNC controller software (physical system) and the DT throughout the sequence.

3.2. Collision Detection System Setup

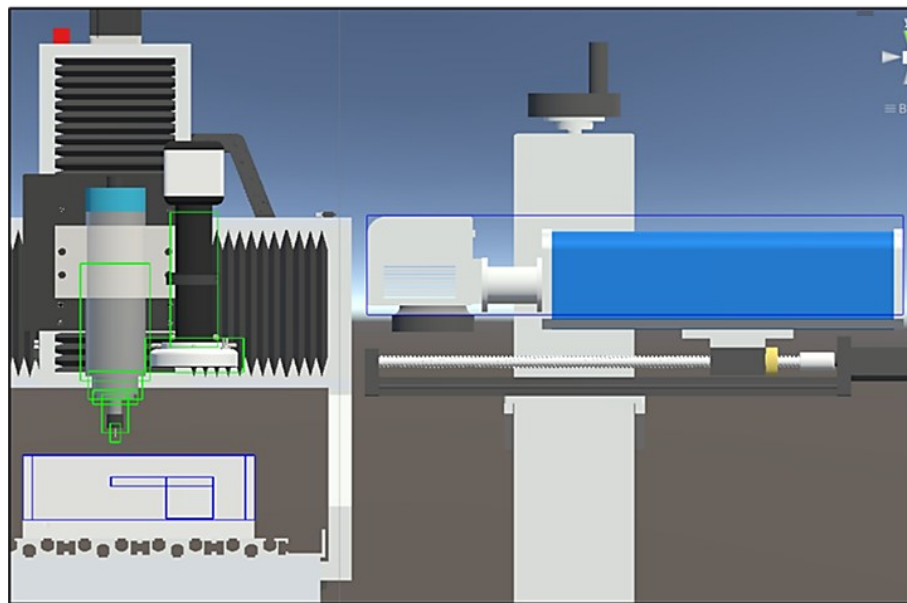
This section outlines the setup and components of the collision detection system. The primary components being monitored are the μ EDM head (comprising the spindle, microscopic camera, and ring light), the LBMM head, the dielectric tank reservoir, and the workpiece holder. The collision detection mechanism relies on continuous monitoring of the 3D geometries and distances between these components within a virtual environment. To achieve accurate real-time collision detection, the system employs Unity3D's Box Collider components, which are assigned to the respective monitoring elements. In Unity3D, colliders detect physical interactions, with common types including box, cylinder, sphere, and capsule colliders. The Box Collider was selected because its simple geometric structure reduces computational load and enables faster, more efficient collision detection in real-time simulations. Each Box Collider defines the geometric boundary of its corresponding physical component. In addition, a dynamic safety offset, ranging from 5 mm to 10 mm depending on the feed rate or motion speed, is applied to facilitate early collision detection before physical contact occurs. As shown in Figure 6 (a) and (b), green box colliders represent the μ EDM head, and blue represents the other monitoring components. Unlike conventional point-based or bounding-box proximity checks, this method uses well-positioned 3D box colliders that reflect the actual tool and machine geometry, enhancing the realism and accuracy of collision detection. When the tool's collider intersects with another component's collider, the DT sends a signal to the μ EDM CNC controller through Firebase, prompting an immediate halt in machine movement, effectively preventing mechanical damage.

This shape-based awareness is particularly useful in complex regions of the machine, such as the dielectric tank reservoir, where confined space and irregular geometry present higher risks. For instance, when the μ EDM tool approaches from outside the dielectric tank reservoir, the collision detection system triggers as the tool's box collider intersects with the tank's box collider, as shown in Figure 7 (a) and (b). Additionally, within the confined space of the tank, the system effectively detects potential collisions between the tool and the workpiece holder, as illustrated in Figure 8 (a) and (b), showcasing the system's ability to ensure safe operations in both open and confined environments.

Proper machine architecture and motion planning should minimize collision risk without heavy reliance on real-time detection. The proposed Digital Twin-based collision detection provides additional safety during manual jogging and setup, which is especially beneficial during alignment operations where tool movements are unpredictable and require live feedback. The same monitoring system also runs during automated G-code machining, providing continuous protection throughout the full operation.



(a)



(b)

Figure 6. Box colliders with safety offset used for collision detection. Green box colliders represent the μ EDM head, while blue box colliders represent other monitoring components. (a) Isometric view, (b) Side view.

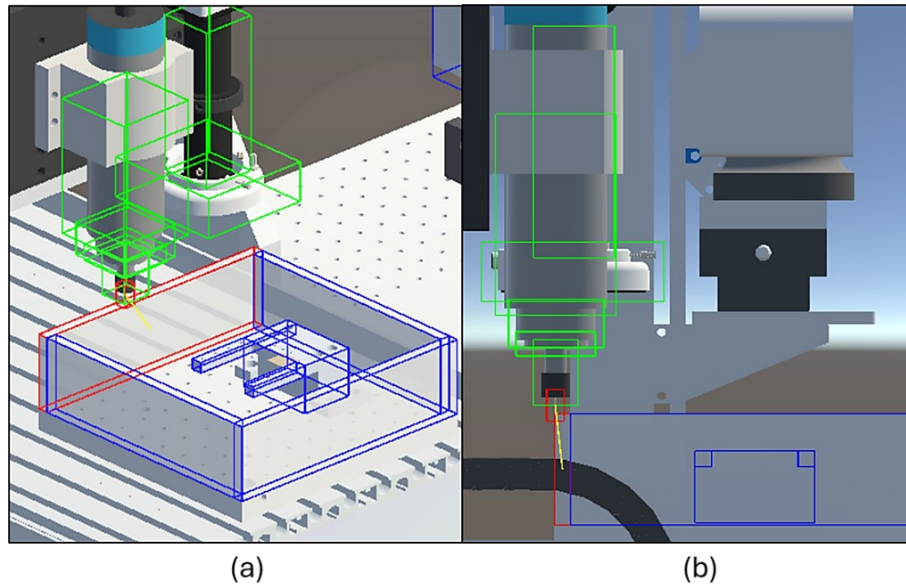


Figure 7. Collision detection (with safety offset) between the tool and the dielectric tank reservoir. (a) Isometric view, (b) Side view. Green box colliders represent the tool, blue box colliders represent the dielectric tank and workpiece holder, and red areas indicate detected collisions.

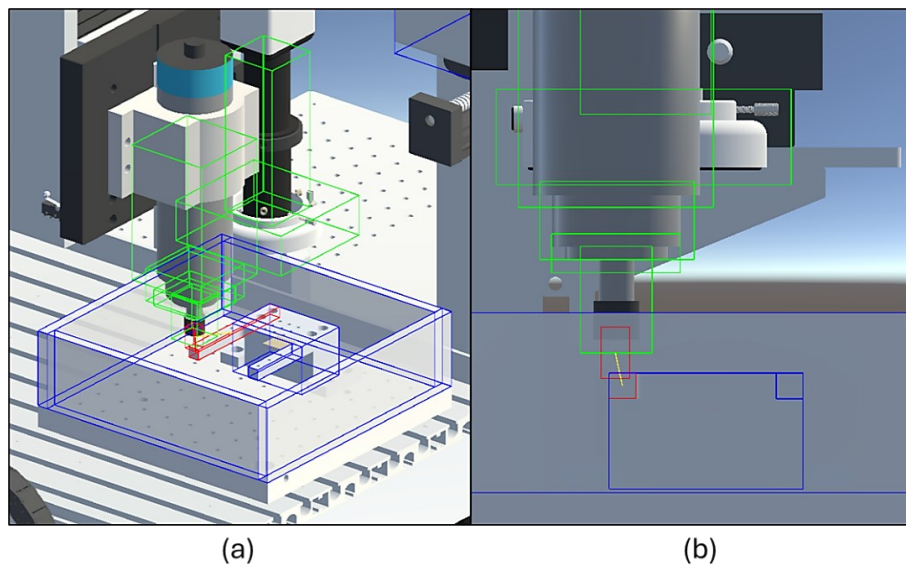


Figure 8. Collision detection (with safety offset) between the tool and the workpiece holder. (a) Isometric view, (b) Side view. Green box colliders represent the tool, blue box colliders represent the workpiece holder and dielectric tank, and red areas indicate detected collisions.

3.3. Latency and System Response

To evaluate the real-time responsiveness of the Digital Twin system, a latency test measured the time delay between issuing a movement command from the μ EDM CNC control software and the corresponding position update in the Unity-based visualization interface. Two critical timestamps were recorded: (1) when the command was sent from the control software, and (2) when the updated tool position was visually reflected in Unity. The test was repeated ten times to ensure consistency and reliability. The results, illustrated in Figure 9, showed an average latency of approximately 16 milliseconds, with the maximum recorded latency being 19 milliseconds. This low delay confirms the system's capability to provide near-instantaneous feedback, which is essential for effective real-time monitoring and collision detection.

Compared to other Digital Twin (DT)-based monitoring systems that achieve synchronization rates of around one second [37], the proposed system demonstrates a significantly faster response time, making it better suited to time-sensitive industrial applications. Latency remained stable across all test iterations, indicating the robustness of the communication pipeline. The DT eliminates the need for live camera feeds by enabling flexible, low-bandwidth, multi-angle visualization without added hardware costs.

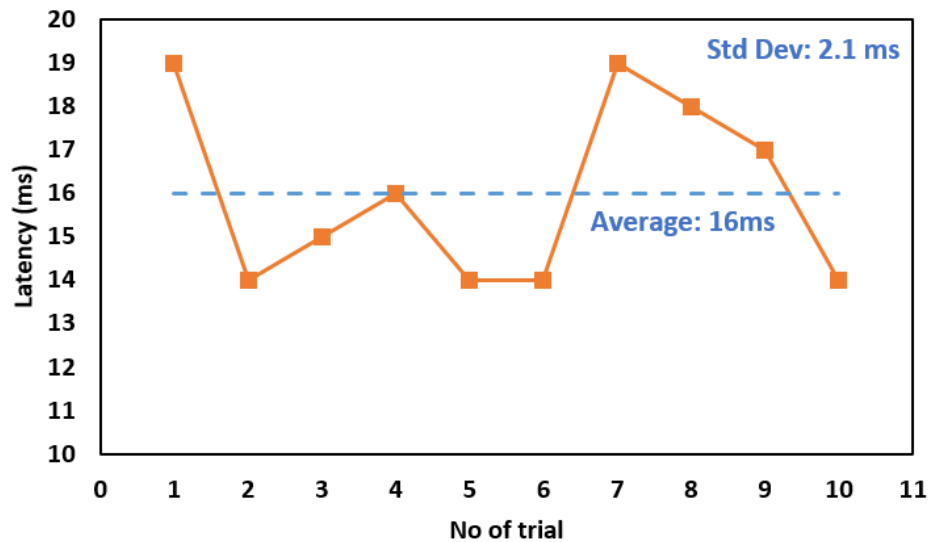


Figure 9. Latency Measurement for Real-Time Digital Twin System

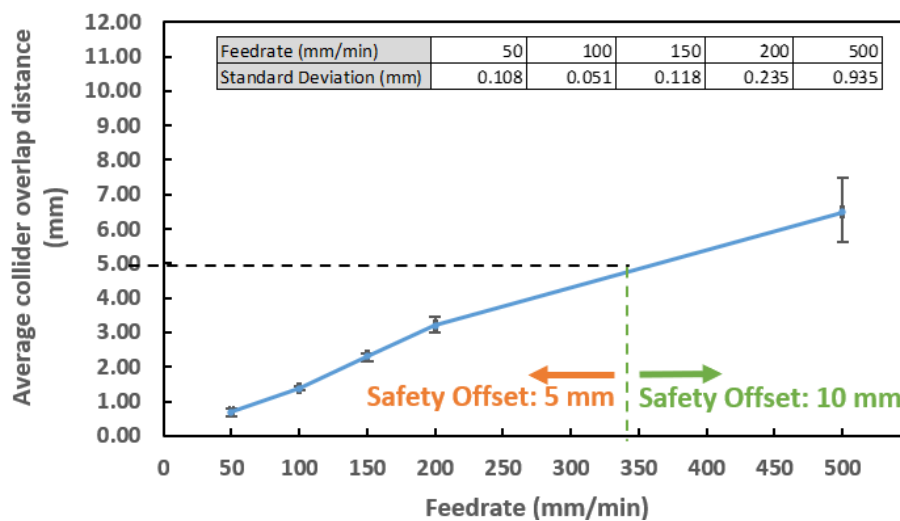


Figure 10. Average collider overlap distance at different feed rates ($n = 3$) with error bars representing the minimum and maximum values observed for each feed rate.

In the context of machine, workpiece, and tool motion, latency refers to the time delay between issuing a command and the system's corresponding physical response. This delay must be accounted for in the control strategy so that, when a halt signal is issued, the tool decelerates appropriately and does not overshoot or collide with the workpiece. In this work, a dynamic safety offset, ranging from 5 mm to 10 mm depending on the feed rate or motion speed, is applied to facilitate early collision detection before physical contact occurs. To further validate this approach, experiments were conducted at different feed rates, with a 5 mm safety offset. The experiment involves moving the μ EDM tool toward the dielectric tank along the Y-axis (as in Figure 10) at feed rates of 50, 100, 150, 200, and 500 mm/min. The resulting collider

overlap distances were recorded three times for each feed rate and are presented in Figure 10. The collider overlap distance increases with feed rate, as higher feed rates result in greater momentum, requiring a longer distance for the tool to decelerate safely. At a feed rate of 500 mm/min, the measured overlap of 6.5 mm exceeds the default safety offset of 5 mm. Based on linear interpolation, feed rates of 350 mm/min and above also surpass the 5 mm safety offset. Therefore, for feed rates above 350 mm/min, the DT system will automatically adjust the safety offset to 10 mm to ensure the tool stops in time and avoids contact with the dielectric tank.

4. DIGITAL TWIN FOR OFFLINE PROCESS COORDINATION

This section discusses the use of the Digital Twin system for offline simulation and planning of the sequential LBMM- μ EDM machining process. It highlights the system's ability to interpret G-code and M-code commands and simulate tool movements and process coordination. Section 4.1 demonstrates how tool and workpiece interactions, such as hole formation during LBMM or μ EDM, are visually represented in the Unity environment. Meanwhile, Section 4.2 focuses on the system's ability to estimate machining time based on G-code feed rates and interpolated real μ EDM process data, providing predictive insights to support process planning.

4.1. CNC Command Simulation with Visualized Cutting Effects

In simulation mode, the Digital Twin was tested for its ability to accurately interpret and execute a variety of standard CNC commands commonly used in CNC machining operations. It utilizes the same set of commands as the actual μ EDM CNC controller, ensuring consistency between the virtual simulation and real-world execution. The supported commands are tabulated in Table 2. These M-codes are not standard ISO or FANUC codes and apply only to this system, where they are custom-defined and may vary between manufacturers.

Table 2. List of supported CNC Commands on the DT for offline simulation

CNC Command	Description
G0, G01	Rapid and linear interpolation movement (X,Y,Z,W)
G90, G91	Absolute and incremental positioning
G38.2	Probing command for surface detection
G04	Dwell command for pause simulation
G10 L20	Work offsets G54 to G59
M03, M04, M05	Spindle control (clockwise, counterclockwise, stop)
M200, M201	Laser ON/OFF commands
M101, M100	μ EDM simulation ON/OFF commands

In addition, in the Digital Twin environment, the interaction between the machining tools and the workpiece is simulated using visual cues that mimic material removal. When the LBMM laser beam or the μ EDM electrode engages with the workpiece, the system creates see-through or hollow areas on the simulated workpiece mesh to illustrate the cutting or drilling effect. For instance, when the LBMM beam interacts with the surface, a 300 μ m diameter hole is visualized, while μ EDM operations produce a 500 μ m diameter hole, reflecting the size of the tungsten electrode used. Figure 11(a) shows an example where a 300 μ m LBMM hole is first created, followed by larger 500 μ m μ EDM holes at the same position, as shown in Figure 11(b), illustrating the sequential machining process. These visualized effects allow users to intuitively verify the cutting sequence and assess potential errors or collisions in an offline environment. The workpiece is modeled using a ProBuilder mesh in Unity, and hole generation is implemented through mesh subtraction library [38]. In this process, the volume of the

LBMM beam or the μ EDM tool is subtracted from the workpiece mesh, effectively simulating material removal. This approach creates realistic visual effects, with holes appearing at the exact locations where the tool interacts with the workpiece.

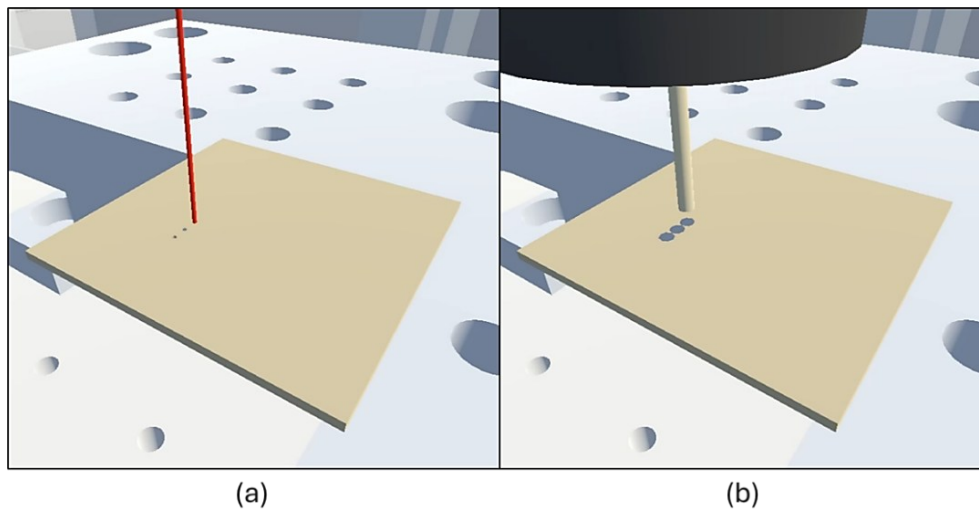


Figure 11. Visualization of sequential machining effects in the Digital Twin environment: (a) 300 μ m diameter LBMM holes, (b) 500 μ m diameter μ EDM holes at the same locations.

4.2. Machining Time Estimation

The DT system can calculate machining times based on the G-code feed rate, as tool movement is governed by the feed rate specified in the G-code. However, calculating the machining time for the LBMM process is not feasible, as the actual processing time is recorded internally by the EZCAD software (closed system), which does not share data with the μ EDM CNC controller software. Similarly, estimating machining time for the μ EDM process is not possible using G-code feed rate alone, as μ EDM machining time is stochastic and influenced by various factors such as material removal rate, discharge frequency, and spark stability. To address this, we include Z-displacement data from actual μ EDM experiments to improve machining-time prediction.

Two sets of experiments were conducted to obtain the datasets: one for a 0.2 mm-thick and another for a 0.6 mm-thick stainless steel (SS304) workpiece, both machined at 90V and 10nF capacitance. These datasets enable the prediction of material removal for workpiece thicknesses between 0.2 and 0.6 mm. For this study, a thickness of 0.4 mm was selected for prediction, where a linear interpolation method was used to estimate its Z-depth profile and subsequently validate it against the corresponding 0.4 mm experimental data. Linear interpolation was used because the Z-depth data from the 0.2 mm and 0.6 mm workpieces show a similar and steady trend under the same machining conditions (90 V and 10 nF). Since no experimental data exist for intermediate thicknesses, interpolation provides a simple, practical way to estimate the Z-depth for the 0.4 mm workpiece. This allows a direct comparison with the actual experimental result for validation.

Figure 12 illustrates the comparison between the predicted and experimental Z-depth profiles for the 0.4 mm workpiece, with experimental Z-depth profiles in orange (0.2 mm), green (0.6 mm), and yellow (0.4 mm), and the predicted 0.4 mm profile shown in blue. It should be noted that, since the machining depth of the 0.2 mm workpiece is only 0.2 mm, the 0.2 mm data are linearly extrapolated to 0.04 mm prior to performing the 0.4 mm prediction. Then, the machining time was calculated by interpolating data from the 0.2 mm and 0.6 mm thickness

profiles. To validate the prediction, an additional experiment was conducted using a 0.4 mm thick workpiece.

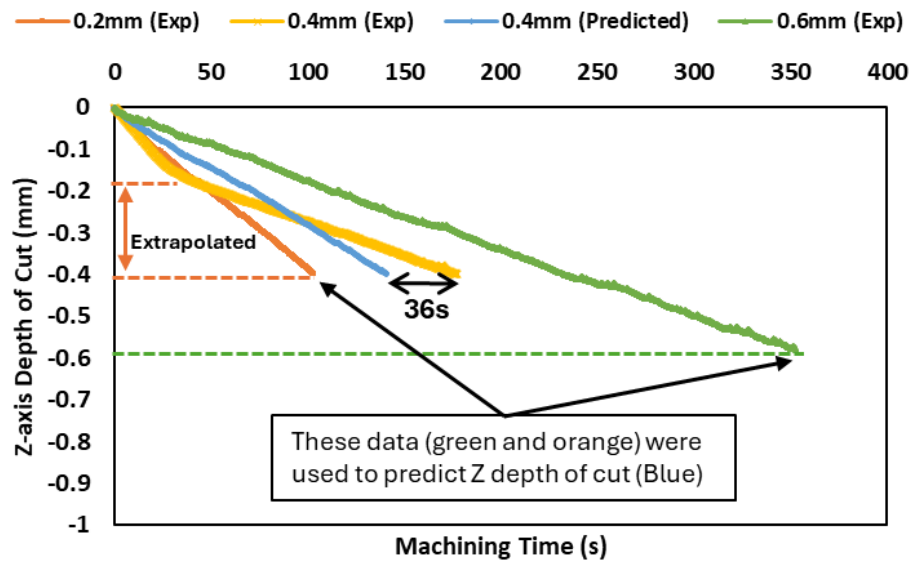


Figure 12. Comparison of predicted and actual Z-depth profiles for 0.4 mm workpiece thickness, predicted based on interpolation from 0.2 mm and 0.6 mm real machining data.

The results show a close agreement between the predicted and experimental machining times, with a difference of approximately 36 s, corresponding to a 20.2% error and an overall prediction accuracy of 79.8%. Although the percentage error appears relatively high, μ EDM is inherently a nonlinear and stochastic process. Unlike conventional machining processes, material removal in μ EDM is governed by a sequence of discrete electrical discharges, each exhibiting random variations in discharge energy, plasma channel formation, ignition delay, and crater size. Consequently, the material removal rate (MRR) fluctuates continuously throughout machining and does not scale linearly with workpiece thickness or machining depth.

As machining progresses, several depth-dependent phenomena further contribute to the stochastic behavior. Electrode wear gradually tapers the tool, altering the effective tool geometry and spark gap. At greater depths, debris evacuation becomes increasingly difficult due to restricted flushing conditions, leading to unstable discharge conditions, secondary discharges, arcing, and occasional short circuits. These phenomena reduce machining efficiency and introduce additional variability in the instantaneous MRR. Furthermore, the servo control system continuously adjusts the tool position to maintain a stable discharge gap, resulting in variable feed rates and intermittent machining pauses that are difficult to predict using deterministic models.

Given these characteristics, achieving 79.8% prediction accuracy with an interpolation-based approach shows that the proposed method captures the overall machining trend despite the process's randomness. The predicted Z-depth profile closely follows the experimental profile, indicating that interpolation provides a practical and computationally efficient way to estimate machining time. Nevertheless, further improvements may be achieved by incorporating nonlinear depth-dependent correction factors, adaptive electrode wear compensation, or data-driven approaches such as artificial intelligence and machine learning models capable of learning the complex stochastic relationships between machining parameters, electrode wear, flushing conditions, and material removal behavior. Such

enhancements could further reduce prediction error and improve the robustness of machining time estimation for μ EDM applications.

5. CONCLUSION

In conclusion, a comprehensive Digital Twin (DT) framework for the sequential Laser Beam Micromachining (LBMM) and micro-EDM (μ EDM) process is proposed to address critical challenges in collision detection and process coordination. Integrating Internet of Things (IoT) technology with a Unity3D-based virtual environment enables real-time monitoring and interactive visualization of the machining system. Through IoT connectivity established via Firebase, the system achieves a low average latency of 16 ms, supporting responsive collision detection.

Unlike conventional point-based or bounding-box proximity checks, this method uses well-positioned 3D box colliders that reflect the actual tool and machine geometry, enhancing the realism and accuracy of collision detection. In addition to real-time monitoring, the DT framework supports offline process coordination through virtual simulation. Users can preview and validate the whole sequence of machining operations using G-code and M-code commands, visually assess tool-path execution, and estimate machining time. The results show a close match, with only a slight difference in overall machining time of approximately 36 seconds, corresponding to a 79.8% prediction accuracy. Mitigation: Simulated cutting effects, such as through-holes corresponding to LBMM and μ EDM operations, enhance process interpretability. By incorporating real experimental Z-displacement data, the DT also demonstrates predictive capability for machining outcomes, illustrating its potential for future integration with AI-driven models. The DT eliminates the need for live camera feeds by enabling flexible, low-bandwidth, multi-angle visualization without added hardware costs.

The main limitation of this study for real-time collision detection is its reliance on geometry-based box colliders, which may not fully capture complex dynamic effects such as vibration, tool deflection, or rapid transient movements during high-speed machining. For machining time prediction, a key limitation is the use of a simplified interpolation-based approach, which does not fully account for the nonlinear and stochastic nature of μ EDM material removal, reducing accuracy under varying machining conditions. For future improvements, more advanced physics-based collision models or sensor-assisted detection can be introduced to improve realism and accuracy, while data-driven or machine learning models can be developed to better capture nonlinear machining behavior and enhance the prediction of machining time under different process conditions.

The findings support the main objective by demonstrating that the developed Digital Twin framework successfully enables real-time analysis and decision-making for sequential machining. The low communication latency (16 ms) confirms its capability for real-time data exchange, while the synchronized 3D simulation and collision detection validate improved process coordination and safety. In addition, the online virtual environment and 79.8% machining time prediction accuracy show that the framework can perform effective virtual validation of LBMM– μ EDM operations, fulfilling the intended objectives of monitoring, coordination, and predictive assessment.

ACKNOWLEDGMENT

This research is also funded under the grant PRGS/1/2022/TK03/UIAM/02/1 provided by the Ministry of Higher Education, Malaysia. Wan Ahmad Bin Wan Azhar is supported by the

KOE IIUM Engineering Merit Scholarship (KOEiems 2023). The data that support the findings of this study are available from the corresponding author upon request.

REFERENCES

- [1] Li ZY, Wei XT, Guo YB, Sealy MP. (2015) State-of-art, challenges, and outlook on manufacturing of cooling holes for turbine blades. *Mach. Sci. Technol.*, 19(3):361-399. <https://doi.org/10.1080/10910344.2015.1051543>
- [2] Shang J, et al. (2023) Micro-hole array sprayer combined with an organic membrane to assist LIBS (MASOM-LIBS): A novel highly sensitive detection method for dissolved trace heavy metals in water. *Talanta*, 264:124780. <https://doi.org/10.1016/j.talanta.2023.124780>
- [3] Wang C, Lu Z, Zhang K. (2018) Fabrication of micro-parts with high-aspect ratio micro-hole array by micro-powder injection molding. *Materials*, 11(10):1864. <https://doi.org/10.3390/ma11101864>
- [4] Xu Z, Wang Y. (2021) Electrochemical machining of complex components of aero-engines: Developments, trends, and technological advances. *Chinese J. Aeronaut.*, 34(2):28-53. <https://doi.org/10.1016/j.cja.2019.09.016>
- [5] Davim JP. (2013) *Nontraditional Machining Processes*, 1st Ed. Springer London. <https://doi.org/10.1007/978-1-4471-5179-1>
- [6] Zhang N, Wang M, Ban M, Guo L, Liu W. (2024) Femtosecond laser drilling 100 μm diameter micro holes with aspect ratios in a Nickel based superalloy. *J. Mater. Res. Technol.*, 28:1415-1422. <https://doi.org/10.1016/j.jmrt.2023.12.049>
- [7] Yung-Chou H. (2024) Applying punching without die to micro-hole array processing. *J. Manuf. Process.*, 116:284-292. <https://doi.org/10.1016/j.jmapro.2024.02.064>
- [8] Chavoshi SZ, Luo X. (2015) Hybrid micro-machining processes: A review. *Precis. Eng.*, 41:1-23. <https://doi.org/10.1016/j.precisioneng.2015.03.001>
- [9] Li L, Diver C, Atkinson J, Giedl-Wagner R, Helml HJ. (2006) Sequential laser and EDM micro-drilling for next generation fuel injection nozzle manufacture. *CIRP Ann.*, 55(1):179-182. [https://doi.org/10.1016/S0007-8506\(07\)60393-X](https://doi.org/10.1016/S0007-8506(07)60393-X)
- [10] Kim S, Kim BH, Chung DK, Shin HS, Chu CN. (2010) Hybrid micromachining using a nanosecond pulsed laser and micro EDM. *J. Micromechanics Microengineering*, 20(1):015037. <https://doi.org/10.1088/0960-1317/20/1/015037>
- [11] Racz SG, Breaz RE, Cioca LI. (2020) Hazards that can affect CNC machine tools during operation—An AHP approach. *Safety*, 6(1):10. <https://doi.org/10.3390/safety6010010>
- [12] Chen L, Xu K, Tang K. (2018) Optimized sequence planning for multi-axis hybrid machining of complex geometries. *Comput. Graph.*, 70:176-187. <https://doi.org/10.1016/j.cag.2017.07.018>
- [13] Basinger K, Webster C, Keough C, Wysk R, Harrysson O. (2020) Advanced manufacturing using linked processes: Hybrid manufacturing. In: *Mass Production Processes*, IntechOpen. <https://doi.org/10.5772/intechopen.88560>
- [14] Cortina M, Arrizubieta JI, Ruiz JE, Ukar E, Lamikiz A. (2018) Latest developments in industrial hybrid machine tools that combine additive and subtractive operations. *Materials*, 11(12):2583. <https://doi.org/10.3390/ma11122583>
- [15] Christiand, Kiswanto G. (2020) Digital twin approach for tool wear monitoring of micro-milling. *Procedia CIRP*, 93:1532-1537. <https://doi.org/10.1016/j.procir.2020.03.140>
- [16] He Y, Guo J, Zheng X. (2018) From surveillance to digital twin: Challenges and recent advances of signal processing for industrial Internet of Things. *IEEE Signal Process. Mag.*, 35(5):120-129. <https://doi.org/10.1109/MSP.2018.2842228>
- [17] Wright L, Davidson S. (2020) How to tell the difference between a model and a digital twin. *Adv. Model. Simul. Eng. Sci.*, 7(1):13. <https://doi.org/10.1186/s40323-020-00147-4>

- [18] Semeraro C, Lezoche M, Panetto H, Dassisti M. (2021) Digital twin paradigm: A systematic literature review. *Comput. Ind.*, 130:103469. <https://doi.org/10.1016/j.compind.2021.103469>
- [19] Liu M, Fang S, Dong H, Xu C. (2021) Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.*, 58:346-361. <https://doi.org/10.1016/j.jmsy.2020.06.017>
- [20] Liu K, Song L, Han W, Cui Y, Wang Y. (2022) Time-varying error prediction and compensation for movement axis of CNC machine tool based on digital twin. *IEEE Trans. Ind. Informatics*, 18(1):109-118. <https://doi.org/10.1109/TII.2021.3073649>
- [21] Almeida S, Mo JPT, Bil C, Ding S, Cheng CT. (2023) A novel robotic EDM digital twin for offline cutting-path programming. <https://doi.org/10.3233/ATDE230624>
- [22] Wan Azhar WAB, Saleh T. (2025) An integrated setup for sequential hybrid laser beam micromachining (LBMM) and micro-EDM for machining micro-hole arrays with automated tool localization. *Int. J. Precis. Eng. Manuf.*, 26(8):1889-1906. <https://doi.org/10.1007/s12541-025-01251-4>
- [23] Wan Azhar WAB, Saleh T. (2025) Radial tool wear analysis and compensation strategy for microhole arrays in sequential hybrid laser beam micromachining (LBMM) and micro-EDM. *Int. J. Adv. Manuf. Technol.*, 141(5-6):3411-3430. <https://doi.org/10.1007/s00170-025-16717-y>

Budget-Efficient PID Tuning of a Line-Following Differential-Drive AGV Using Extra Trees Surrogate-Assisted CMA-ES

MASTANG¹, BAKR AHMED TAHA¹, RETNA APSARI^{2,3}, MUHAMMAD MUKHLISIN⁴,
ENI DWI WARDIHAN⁵, NORHANA ARSAD^{1,2*}

¹Dept. of Electrical, Electronic & Systems Engineering, Universiti Kebangsaan Malaysia,
43600, Selangor, Malaysia

²Dept. of Engineering, Faculty of Advanced Technology and Multidiscipline, Airlangga University,
60115, Surabaya, Indonesia

³Dept. of Physics, Faculty of Science and Technology, Airlangga University,
60115, Surabaya, Indonesia

⁴Dept. of Civil Engineering, Faculty of Engineering, Semarang State Polytechnic,
50275, Semarang, Indonesia

⁵Dept. of Electrical Engineering, Faculty of Engineering, Semarang State Polytechnic,
50275, Semarang, Indonesia

*Corresponding author: noa@ukm.edu.my

(Received: 7 June 2026; Accepted: 22 August 2026; Published online: 11 September 2026)

ABSTRACT: Physical PID tuning for line-following automated guided vehicles (AGVs) is sample-inefficient because the gain-to-performance relationship is nonlinear and noisy, while every candidate must be executed on the real platform. Conventional population optimizers can therefore spend a limited experimental budget on unproductive trials, and evidence for surrogate-assisted CMA-ES under explicit true-evaluation accounting on embedded AGVs remains limited. This study evaluates an Extra Trees surrogate-assisted CMA-ES supervisory tuner on a single line-following differential-drive AGV. CMA-ES generates a virtual gain population in normalized space, and Extra Trees ranks it using the predicted objective and inter-tree dispersion; only a small, ranked subset undergoes physical evaluation. The objective combines integral absolute error, root mean square error, overshoot, actuation smoothness, and a lost-line penalty. We compared the method with pure CMA-ES, particle swarm optimization, and a genetic algorithm on Circle and 8Shape trajectories, using an identical budget of 57 true evaluations for every method, seed, and trajectory. The proposed method achieved the lowest mean final-best objective on both trajectories and improved on the strongest baseline by 30.8% on Circle and 20.8% on 8Shape. Cross-run distributions, empirical cumulative distribution functions, and time-domain responses support the same ranking. These results show that surrogate screening can improve sample efficiency for PID tuning on the tested embedded platform.

ABSTRAK: Penalaan PID secara fizikal bagi kenderaan berpandu automatik (AGV) pengikut garisan tidak cekap dari segi sampel kerana hubungan antara gandaan dan prestasi adalah tak linear serta bising, sedangkan setiap calon perlu diuji pada platform sebenar. Pengoptimum berasaskan populasi konvensional boleh menggunakan bajet eksperimen yang terhad pada percubaan yang kurang produktif, dan bukti bagi CMA-ES berbantuan surrogat dengan perakaunan penilaian sebenar yang jelas pada AGV terbenam masih terhad. Kajian ini menilai penala penyeliaan CMA-ES berbantuan surrogat Extra Trees pada sebuah prototaip AGV pacuan kebezaan pengikut garisan. CMA-ES menjana populasi gandaan maya dalam ruang ternormal dan Extra Trees menyusunnya menggunakan objektif ramalan dan sebaran antara pokok, manakala hanya subset kecil dengan skor terbaik menjalani penilaian fizikal. Fungsi

objektif menggabungkan ralat mutlak bersepadu, ralat punca min kuasa dua, limpahan, kelancaran aktuasi, dan penalti kehilangan garisan. Kaedah ini dibandingkan dengan CMA-ES tulen, pengoptimuman kawanan zarah, dan algoritma genetik pada trajektori Circle dan 8Shape menggunakan bajet yang sama, iaitu 57 penilaian sebenar bagi setiap kaedah, benih rawak, dan trajektori. Kaedah yang dicadangkan mencapai purata objektif terbaik akhir terendah pada kedua-dua trajektori serta mengatasi garis dasar terbaik sebanyak 30.8% pada Circle dan 20.8% pada 8Shape. Taburan antara larian, fungsi taburan kumulatif empirik, dan respons domain masa menyokong kedudukan yang sama. Dapatan ini menunjukkan bahawa penapisan surrogat dapat meningkatkan kecekapan sampel bagi penalaan PID pada platform terbenam yang diuji.

KEYWORDS: *Automated Guided Vehicle, PID tuning, CMA-ES, Extra Trees Surrogate, Fixed-Budget Optimization.*

1. INTRODUCTION

Automated guided vehicles (AGVs) are a core technology in modern manufacturing and warehousing intralogistics. On production lines, they transport raw materials, work-in-process, components, tools, and finished products among storage locations, workstations, assembly cells, inspection points, and shipping areas. Their role extends beyond replacing manually pushed trolleys, as AGVs can coordinate material flow, reduce dependence on repetitive handling tasks, and connect production planning with physical logistics. Recent reviews therefore identify AGVs as an important element of smart manufacturing systems that integrate equipment, storage, and human work areas [1, 2].

AGVs are most effective when they can follow a planned route repeatedly with minimal supervision. In manufacturing, many vehicles operate on fixed or semi-fixed paths defined by magnetic tape, painted lines, QR codes, or virtual routes. Even when the global route is predetermined, the vehicle must maintain accurate local tracking because small lateral deviations can develop into oscillation or line-loss events. In this study, a line-following differential-drive AGV is used as a controlled experimental platform for local tracking, not as a proxy for every AGV architecture. A 15-sensor array estimates the lateral position of the line, and the controller maps the resulting error to differential left-right motor commands. Accordingly, the conclusions are restricted to this platform and the tested operating conditions.

Despite significant advances in sensors, mapping, reinforcement learning, and predictive control models, PID controllers are still relevant in AGV applications. PIDs are simple, computationally lightweight, and can be implemented on low-cost embedded hardware. The proportional component reacts to current errors, the integral component compensates for persistent bias, and the derivative component reduces rapid changes. For small AGVs operating on Raspberry Pi or microcontroller-class devices, PIDs provide a practical control law with low memory requirements and latency [3]. In real-world manufacturing, this simplicity matters because maintenance technicians must understand controller behavior, diagnose instabilities, and retune parameters when loads, floor friction, motor response, or sensor heights change. Therefore, PIDs often serve as inner-loop controllers, while more advanced algorithms act as supervisory tuners or trajectory planners [4].

Fig. 1 illustrates the classical closed-loop PID structure. The reference signal $r(t)$ defines the desired centerline, while the measured output $y(t)$ is obtained from the AGV sensor system. The tracking error $e(t)$ is processed by the proportional, integral, and derivative terms, whose sum produces the motor-control signal $u(t)$. A high proportional gain can speed up correction but can also induce oscillation. A large integral gain can remove persistent bias but increase

windup, whereas a large derivative gain can damp rapid error changes but amplify sensor noise. PID tuning therefore requires a balance among tracking accuracy, transient behavior, command smoothness, and recovery from line-loss events.

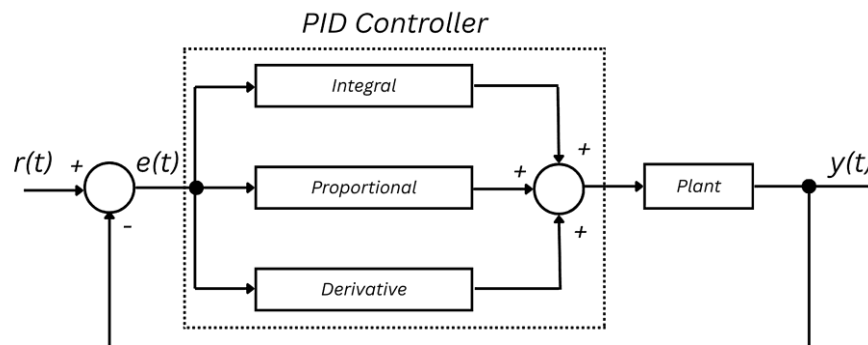


Figure 1. Classical closed-loop PID structure used as the conceptual basis for AGV line tracking.

The challenge of PID tuning for AGVs is the relationship between PID gain and tracking performance. It is nonlinear, noisy, and highly dependent on operating conditions. A gain set that works well on a smooth circle track may not remain optimal on an 8Shape track because the vehicle experiences alternating curvatures, sharper transitions, and different motor saturation patterns. Tracking sensors can also produce uncertainty. When a line is only partially detected, the estimated error can spike. When the line is lost, the controller must recover position without generating aggressive steering commands. Furthermore, AGVs are constrained by their onboard computing power and limited experiment budget. Each true evaluation requires operating the vehicle for a specified duration, recording sensor responses, calculating performance metrics, and resetting the experiment. Such evaluations are more expensive than surrogate predictions.

Classic empirical tuning methods such as manual tuning, Ziegler-Nichols, and relay feedback are attractive due to their simplicity. However, their performance is often limited when the objective combines multiple criteria such as error magnitude, overshoot, smoothness, and line-loss stability. Population-based optimizers, including genetic algorithm (GA), particle swarm optimization (PSO), differential evolution (DE), and beetle antennae search (BAS), can search nonlinear PID space without requiring gradients [5]. However, these methods generally require many true evaluations. In tight budget settings, candidate exploration can be inefficient because many gains are tested. CMA-ES is robust for continuous black-box optimization because it learns the covariance distribution over promising regions [6]. However, pure CMA-ES also consumes true evaluations for each candidate, making it expensive for AGV experiments.

Surrogate-assisted optimization offers a suitable compromise. A surrogate approximates the objective from evaluated samples, allowing an optimizer to screen many candidates before submitting a selected subset to high-fidelity evaluation [7, 8]. Extra Trees is attractive in this study because its randomized regression-tree ensemble can represent nonlinear interactions among K_p , K_i , and K_d , does not require feature scaling, trains quickly on small tabular datasets, and provides inter-tree prediction dispersion. The ensemble mean estimates candidate quality, while the dispersion is used as a heuristic exploration indicator. Combining Extra Trees with CMA-ES therefore couples covariance-based search with model-based filtering of expensive physical trials.

This study's contribution lies in the system-level integration and experimental protocol rather than in any individual algorithmic component. CMA-ES, Extra Trees, PID control, and composite tracking metrics are established methods. Within the scope of a single line-following differential-drive AGV, this study contributes: (i) an explicit accounting scheme that separates virtual surrogate screening from costly true AGV episodes, (ii) uncertainty-informed Extra Trees ranking embedded in a CMA-ES supervisory loop so that only selected candidates consume the physical evaluation budget, (iii) integration of a 15-sensor error model, low-pass filtering, and a lost-line-aware composite objective, and (iv) a controlled fixed-budget physical comparison with pure CMA-ES, particle swarm optimization, and a genetic algorithm on Circle and 8Shape tracks. These contributions constitute a specific configuration and evaluation protocol for the tested platform. They are not presented as a new CMA-ES or Extra Trees algorithm, and broader transfer requires further validation.

Table 1. Representative PID tuning methods and their relevance to AGV line tracking

No	Tuning Method	Principle	Advantages	Disadvantages	Ref
1	Manual/empirical tuning	Adjusting K_p , K_i , K_d based on trial response.	Easy, transparent, no computation.	Subjective and often fails on multi-curvature trajectories.	[9]
2	Ziegler-Nichols / relay feedback	Using the oscillation gain and oscillation period to derive the gain.	Quick as a baseline for simple plants.	Aggressive tuning can cause oscillations and overshoot.	[9]
3	Cohen-Coon/model-based rules	Using the identified first-order-plus-delay response.	Good for known process models.	Requires a reliable model because AGV dynamics change by floor and load.	[9]
4	IMC-based PID	Lowering the gain of the internal-model-control filter.	Robust and easy to interpret.	Requires a plant model and can be suboptimal for nonlinear line tracking.	[9]
5	Fuzzy self-tuning PID	Fuzzy rules adapt the gain from error and the change in error.	Dealing with uncertainty and expert knowledge.	Rule design becomes complex and difficult to generalize.	[10]
6	Neural-network PID	Studying nonlinear mapping from state/error to gain.	Adaptive and expressive.	Requires training data and safety validation.	[11]
7	GA-PID	Evolutionary selection, crossover, and mutation optimize gain.	Global search and simple constraints.	Many evaluations and risk of premature convergence.	[12]
8	PSO-PID	Particles move based on personal best and global best.	Initial convergence is fast, and parameters are relatively few.	Can be stagnant, sensitive to inertia and swarm diversity.	[3, 13]
9	BAS-PID	Beetle antennae search using objective left-right probing.	Light and simple.	Sensitive to step size and noisy objectives.	[5]
10	DE-PID	The vector difference mutates the gain candidate.	Robust continuous optimization baseline.	Can consume many true trials in embedded experiments.	[9]
11	GWO-PID	The grey-wolf hierarchy directs exploration and exploitation.	Competitive metaheuristics for PID benchmarks.	Noisy AGV objectives can mislead the leadership hierarchy.	[9, 14]

No	Tuning Method	Principle	Advantages	Disadvantages	Ref
12	WOA-PID	The search is based on encircling and spiraling like a whale. Employed, onlooker, and scout bee colony exploration.	A simple population algorithm.	Limited adaptation of local covariance among interdependent PID gains	[15]
13	ABC-PID		Good exploration.	It's inefficient when every trial is expensive.	[9]
14	Simulated annealing PID	Probabilistically accept worse solutions during the cooling process.	Can get out of local minima.	Sequential and sensitive to cooling schedule.	[9]
15	Bayesian optimization PID	Probabilistic surrogate and acquisition function select trial.	Very efficient sampling.	GP cost and kernel selection are difficult on noisy data.	[16]
16	Gaussian-process CMA-ES	CMA-ES uses a GP surrogate to reduce evaluations.	Uncertainty awareness and efficient sampling.	GP scalability and hyperparameter fitting can be expensive.	[17]
17	RBF surrogate CMA-ES	RBF approximates the objective in the CMA-ES loop.	Fast interpolation for smooth functions.	Less robust to discontinuities or line-loss penalties.	[18]
18	Reinforcement-learning PID	Agents learn the gain adjustment policy of rewards.	Potential for online adaptation.	Requires safe training and lots of interaction.	[19, 20]
19	Pure CMA-ES PID	Studying the covariance distribution directly from real objectives.	Powerful continuous black-box optimizer.	Each candidate consumes a real evaluation budget.	[6]
20	Proposed CMA-ES Tree	Extra Trees ranks CMA-ES candidates; only the top $q = 2$ candidates undergo true evaluation.	Budget-friendly, nonlinear, embedded-friendly, and uncertainty-aware.	Requires careful initial dataset and surrogate updates.	This paper

Table 1 summarizes the evolution of PID tuning methods from classical approaches to modern surrogate-assisted optimization approaches for AGV applications. Early methods such as manual tuning, Ziegler-Nichols, Cohen-Coon, and IMC have the main advantages of simplicity, transparency, and ease of implementation. However, these methods generally assume a relatively simple, stable, or well-modeled plant. In line-following AGVs, these assumptions often do not hold because vehicle response is affected by trajectory changes, floor friction, load, sensor noise, motor saturation, and line-loss conditions. Therefore, while classical methods are useful as an initial baseline, they cannot perform optimally when AGVs encounter multi-curvature trajectories such as Circle and 8Shape. Adaptive methods such as fuzzy self-tuning PID and neural-network PID can handle nonlinearity and uncertainty but require more complex rule design, training data, and safety validation before being applied to embedded systems. Several methods, such as GA, PSO, BAS, DE, GWO, WOA, ABC, simulated annealing, reinforcement learning, and CMA-ES, can perform a global search for combinations of K_p , K_i , and K_d without requiring analytical gradients. This suits AGV PID tuning because the objective function is black-box and obtained from tracking evaluation results. However, the main drawback of most of these methods is the need for many evaluations, whereas each PID trial on an AGV can be time-consuming, requires a trajectory reset, and risks producing unstable behavior. This is where surrogate-assisted methods become

important. Bayesian optimization, GP-CMA-ES, RBF surrogate CMA-ES, and CMA-ES Tree aim to reduce real-world evaluations by predicting candidate quality before testing. The proposed method, CMA-ES Tree, is more budget-efficient because Extra Trees filter CMA-ES candidates, so only the most promising candidates are sent for true evaluation.

2. METHODOLOGY

The methodology proposed in this paper treats PID tuning as a bounded black-box optimization problem. The objective is not assumed to be differentiable because it results from interactions among AGV dynamics, sensor readings, actuator saturation, and lost-line events. In this paper, the term true evaluation refers to a single high-fidelity episode execution using a prototype AGV. Surrogate prediction, internal candidate ranking, and CMA-ES population generation are not counted as true evaluations. This definition is used consistently to ensure that comparison with the baseline measures sampling efficiency, not the amount of unseen internal computation.

Fig. 2 summarizes the proposed Extra Trees surrogate-assisted CMA-ES supervisory architecture. On the optimizer side, CMA-ES generates $\lambda=32$ virtual PID candidates in normalized space. Extra Trees predicts the objective and inter-tree dispersion for each candidate, and the score in Eq. (18) ranks the population. Only the selected subset of $q=2$ candidates is executed on the physical AGV. On the control side, the PID controller receives tracking errors, generates the correction signal, and drives the differential motors while sensor feedback closes the loop. This architecture separates the fast, deterministic control loop from the slower supervisory optimization loop, which updates between episodes.

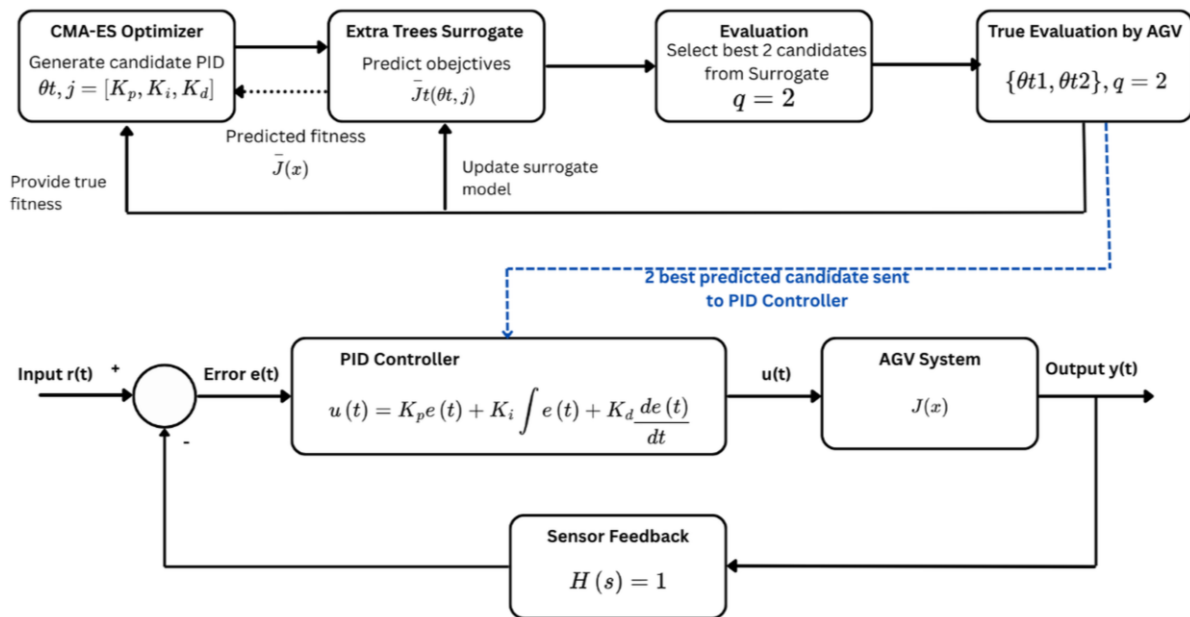


Figure 2. Proposed Extra Trees surrogate-assisted CMA-ES supervisory architecture and its relationship with the AGV PID feedback loop.

2.1. Problem Formulation

Eq. (1) defines the decision vector θ , where K_p , K_i , and K_d denote the proportional, integral, and derivative gains, respectively, and L denotes the feasible search domain. The gains are bounded by $0 \leq K_p \leq 30$, $0 \leq K_i \leq 3$, and $0 \leq K_d \leq 10$.

$$\theta = [K_p \ K_i \ K_d]^T, \ L = [0,30] \times [0,3] \times [0,10] \quad (1)$$

These bounds are platform-specific safety and search limits established during preliminary bench testing with the same base speed, actuator limits, sensor arrangement, control direction, and track surface used in the experiments. The lower bounds are zero because negative gains reverse the intended corrective action. Above the selected upper limits, larger proportional gains produced sustained actuator saturation and lateral oscillation, larger integral gains increased windup and recovery time, and larger derivative gains amplified sensor noise and command chatter. The same domain was applied to every optimizer for fairness where these values are not proposed as universal AGV limits.

The optimization seeks a feasible PID gain vector that minimizes, as defined in Eq. (2) where θ^* is optimal PID gain. Because the J value is obtained from AGV experiments, the optimization process is considered black box. This means that the optimizer can only try a PID combination and see the resulting J value but does not know the derivative or gradient formula directly.

$$\theta^* = \arg \min_{\theta \in L} J(\theta) \quad (2)$$

CMA-ES is more stable when all variables operate on comparable numerical scales. Therefore, the physical PID gain is transformed into a normalized vector z as shown in Eq. (3). Each z component is on a unit interval, so that K_p , K_i , K_d contribute equally to the covariance adaptation.

$$z = \frac{\theta - \theta_{min}}{\theta_{max} - \theta_{min}}, \ z \in [0,1]^3 \quad (3)$$

Before the candidate is executed on the AGV, the normalized vector is converted back to the original PID domain as formulated in Eq. (4). This equation ensures that each selected candidate corresponds to a physically interpretable PID gain.

$$\theta = \theta_{min} + z \odot (\theta_{max} - \theta_{min}) \quad (4)$$

2.2. AGV Error Model

The AGV uses 15-line sensors. The relative positions of the sensors are modeled linearly from the left to the right side of the AGV body. With $n_s=15$ sensors, the first sensor has a position of -1, the middle sensor is around 0, and the last sensor has a position of +1. This normalized coordinate system is shown in Eq. (5). It can convert raw sensor readings into lateral errors that are independent of the physical distance between the sensors.

$$p_i = -1 + \frac{2(i-1)}{n_s-1} \quad (5)$$

Eq. (6) converts the raw reading $s_i(k)$ into a nonnegative activation weight $\omega_i(k)$ by subtracting the detection threshold $\tau = 400$. The value $\tau = 400$ was obtained from preliminary static calibration of the 15-channel infrared array under the sensor height, track surface, and ambient lighting used in the experiments. It provided a common separation between line and background responses while retaining the response magnitude for raw weighting. The threshold was fixed for all methods and runs and is therefore a hardware and condition-specific setting, not a universal value.

$$\omega_i(k) = \max(0, s_i(k) - \tau) \quad (6)$$

Eq. (7) shows the logic of the sensor detection threshold. The robot is considered to have detected a line when at least one sensor has a positive activation. This logic is important because the objective function must penalize line-missing events. Without this component, the

optimizer might choose an aggressive gain that reduces short-term error but often causes the robot to deviate from the path.

$$\sum_{i=1}^{n_s} \omega_i(k) > 0 \quad (7)$$

Eq. (8) shows the measured lateral error calculated as the weighted centroid of the active sensors $e_m(k)$. If the left-hand sensor detects a stronger activation, the centroid shifts to the left. If the right-hand sensor is dominant, the centroid shifts to the right. This formulation yields a continuous error estimate from discrete sensor readings.

$$e_m(k) = \frac{\sum_{i=1}^{n_s} \omega_i(k) p_i}{\sum_{i=1}^{n_s} \omega_i(k)} \quad (8)$$

Sensor readings can fluctuate due to illumination, band reflectance, small vibrations, and sampling jitter. Therefore, a first-order low-pass filter as shown in Eq. (9) is used to smooth the measured error before it enters the PID controller. The parameter α_e determines the trade-off between responsiveness and noise suppression. $e(k)$ is filtered error.

$$e(k) = \alpha_e e_m(k) + (1 - \alpha_e) e(k-1) \quad (9)$$

2.3. Objective Function

The objective function is constructed from several complementary performance components. Eq. (10) represents the integral absolute error *IAE*, which measures the accumulated deviation from the center line. This component directly represents the tracking quality over the entire evaluation interval.

$$IAE = \sum_{k=1}^N |e(k)| \Delta t \quad (10)$$

Eq. (11) defines the RMSE, which gives greater weight to large and persistent errors. This metric is useful because two controllers can have similar *IAE* but different error variability.

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N e^2(k)} \quad (11)$$

Eq. (12) shows the overshoot *OS* value calculated as the maximum absolute tracking error. This metric captures extreme deviations and is particularly relevant when the AGV is navigating curves or track transitions.

$$OS = \max_{1 \leq k \leq N} |e(k)| \quad (12)$$

Eq. (13) shows the actuation smoothness *SM* calculated from the change in the PID correction signal $u(k)$. A low value means the controller avoids unnecessary command oscillations, thereby reducing motor stress and improving energy behavior.

$$SM = \sum_{k=2}^N [u(k) - u(k-1)]^2 \quad (13)$$

Eq. (14) defines the line-loss penalty *PEN*, which increases the objective value when the robot no longer detects the path. This component converts qualitative failures into quantitative optimization signals.

$$PEN = 100N_{lost} \quad (14)$$

Eq. (15) defines the scalar objective used by every optimizer. The fixed weights in Table 2 place the greatest emphasis on accumulated tracking error while retaining smaller contributions from *RMSE*, overshoot, actuation smoothness, and line-loss events. Applying the same weights to every method ensures the comparison reflects how each optimizer allocates the evaluation budget, rather than differences in the performance criterion.

$$J(\theta) = 1.0 \times IAE + 0.10 \times RMSE + 0.20 \times OS + 0.0005 \times SM + 0.01 \times PEN \quad (15)$$

2.4. Surrogate Extra Trees Model

After t true evaluations, the algorithm stores a dataset $\mathcal{D}_t$ as shown in Eq. (16) that contains the normalized PID vector and the observed objective values. This dataset is small but informative because it records regions where the real AGV performed well or poorly.

$$\mathcal{D}_t = \{(z_i, J_i)\}_{i=1}^t \quad (16)$$

Eq. (17) shows an Extra Trees Regressor F consisting of M random regression trees. Each tree f_m learns a different partition of the PID search space. Randomness is beneficial in this application because the dataset is small, nonlinear, and potentially noisy.

$$F = \{f_1, f_2 \dots f_M\} \quad (17)$$

Eq. (18) explains the candidate score function in surrogate model-based optimization. $S(z)$ is the candidate's score. The smaller the value of $S(z)$, the more attractive the candidate is to be selected. $\hat{\mu}(z)$ is the average predicted value of the surrogate model for candidate z . $\hat{\sigma}(z)$ is the level of uncertainty of the model's prediction on candidate z . If $\hat{\sigma}(z)$ is large, the model is not very confident in its prediction in that area. κ is an exploration control parameter. A large κ value makes the optimizer more daring to try uncertain areas. A small κ value makes the optimizer more focused on candidates that have been predicted well.

$$S(z) = \hat{\mu}(z) - \kappa \hat{\sigma}(z) \quad (18)$$

Eq. (19) shows the average prediction $\hat{\mu}(z)$ obtained by averaging the predictions across all trees in the ensemble. This value estimates the actual expected objective of the candidate without running an expensive AGV experiment.

$$\hat{\mu}(z) = \frac{1}{M} \sum_{m=1}^M f_m(z) \quad (19)$$

Eq. (20) defines $\hat{\sigma}(z)$, which represents the level of uncertainty in the model prediction for candidate z .

$$\hat{\sigma}(z) = \sqrt{\frac{1}{M-1} \sum_{m=1}^M [f_m(z) - \hat{\mu}(z)]^2} \quad (20)$$

2.5. CMA-ES Candidate Generation

Eq. (21) shows the CMA-ES mechanism in sampling candidate vectors from a multivariate Gaussian distribution in normalized space. The mean m_g represents the current search center, σ_g represents the global step size, and C_g represents the covariance matrix that captures the correlation between PID gains.

$$z_{g,j} \sim N(m_g, \sigma_g^2 C_g) \quad (21)$$

Eq. (22) shows a framework for preventing generated candidates from leaving the feasible interval. It does this by clipping each component to $[0,1]$. This repair rule is simple, deterministic, and consistent across all comparison algorithms.

$$z_{g,j} = \begin{cases} 0, & \text{if } z_{g,j} < 0 \\ z_{g,j}, & \text{if } 0 \leq z_{g,j} \leq 1 \\ 1, & \text{if } z_{g,j} > 1 \end{cases} \quad (22)$$

Eq. (23) defines Q_g as the subset of virtual candidates with the smallest surrogate scores. In each generation, $q = 2$ of the $\lambda = 32$ virtual candidates (6.25%) are sent to true evaluation. After the nine initial evaluations, this choice preserves frequent surrogate retraining within the 57-evaluation budget while providing two observed candidates for each distribution update. A larger q would reduce the screening benefit and the number of update cycles, whereas $q = 1$ would make an update more sensitive to a single noisy episode. Thus, $q = 2$ is an experiment-specific budget-noise trade-off rather than a universally optimal value.

$$Q_g = \text{Top}_q^{\min} \left(\{z_{g,j}\}_{j=1}^{\lambda}; S \right), \quad q = 2 \quad (23)$$

2.6. CMA-ES Distribution Update

After true evaluation is performed, candidates are ranked based on the observed objective J . If the implementation uses the standard CMA-ES library, the mean, step-size, and covariance updates follow the CMA-ES procedure. Eqs. (24)-(26) are presented as a conceptual summary of the distribution update to explain how elite candidates shift the search center towards more promising PID regions.

$$m_{g+1} = \sum_{i=1}^{\mu} \omega_i z_{g,i}^e \quad (24)$$

The covariance estimate measures the distribution of elite candidates around the new mean. A small regularization term εI as shown in Eq. (25), is added to maintain numerical stability and avoid singular covariances.

$$C_{\text{new}} = \sum_{i=1}^{\mu} \omega_i \frac{(z_{g,i}^e - m_{g+1})(z_{g,i}^e - m_{g+1})^T}{\sigma_g^2} + \varepsilon I \quad (25)$$

Eq. (26) shows how the covariance matrix is updated by smoothing the previous covariance and the new estimate. This step avoids abrupt distribution changes and preserves useful search-direction information.

$$C_{g+1} = (1 - c_{\text{cov}})C_g + c_{\text{cov}}C_{\text{new}} \quad (26)$$

2.7. Algorithm Design

The CMA-ES Tree algorithm is designed for embedded deployment as a supervisory tuning module on the Raspberry Pi. This AGV prototype uses two controllers: an Arduino Uno microcontroller as the PID controller and a Raspberry Pi to coordinate the experiment, read performance logs, train the Extra Trees surrogate, generate gain candidates, and write the selected PID gain to the controller. This separation is crucial for real-time safety. The inner PID loop must remain deterministic, while the optimizer can operate at a slower timescale after each evaluation episode.

In a practical implementation, each evaluation episode begins by loading a vector of candidate PIDs. The AGV follows a defined trajectory for the duration of the evaluation, while the Raspberry Pi records timestamps, raw and filtered errors, line detection status, correction commands, and wheel commands. At the end of the episode, the objective component is calculated and stored. Extra Trees retrains on the updated dataset, and CMA-ES proposes the next candidate population. The surrogate filters the candidates so that only a small number of promising candidates are physically evaluated. This architecture reduces risky or unproductive real-world trials and suits low-cost AGV platforms.

Algorithm 1 shows how Fixed-Budget CMA-ES Tree PID Tuning optimizes PID parameters (K_p, K_i, K_d) on an AGV with a limited evaluation budget $B = 57$. The process

begins with normalization of the search space, initial evaluation of $N_0 = 9$ deterministic candidates, and then the best candidate is used to initialize CMA-ES. At each generation, CMA-ES generates virtual candidates; the Extra Trees surrogate then predicts their performance and uncertainty, so only the most promising candidates are evaluated directly on the AGV. The evaluation results are added to the history H , while CMA-ES adapts its mean, covariance, and sampling scale to improve the search. A restart mechanism is used when stagnation occurs, while micro-polishing will utilize the remaining budget to refine the best solution. With this strategy, the algorithm can obtain the best-observed PID parameters and the minimum objective value efficiently in a limited number of real evaluations.

Algorithm 1. Embedded supervisory tuning process for CMA-ES Tree

Input	Random seed s , PID bounds $K_p \in [0,30], K_i \in [0,3], K_d \in [0,10]$, evaluation budget $B = 57$, deterministic initial evaluations $N_0 = 9$, CMA-ES virtual population size $\lambda = 32$, number of true evaluations per generation $q = 2$, elite size $\mu=6$, minimum surrogate training size $N_{min} = 8$, surrogate exploration factor $\kappa = 0.35$, micro-polish threshold $B_m = 12$, episode duration $T = 20s$, and objective weights $\omega_{IAE}, \omega_{RMS}, \omega_{OS}, \omega_{SM}, \omega_{PEN}$.
Step 1	Define the PID search space using the global bounds of K_p, K_i , and K_d
Step 2	Normalize the PID search space into $z \in [0,1]^3$. Initialize the evaluation history $H \leftarrow \emptyset$, evaluation counter $n \leftarrow 0$, and generation counter $g \leftarrow 0$
Step 3	Generate $N_0 = 9$ deterministic initial PID candidates.
Step 4	For each initial candidate K_j , while $n < B$, run one AGV episode, calculate J_j , append (K_j, J_j) to H , and update $n \leftarrow n + 1$.
Step 5	Select $(K^*, J^*) = \arg \min_{(K,J) \in H} J$. Initialize the CMA-ES mean using the normalized K^* , with $\sigma = 0.20$ and $C = I$
Step 6	While $n < B$, set $g \leftarrow g + 1$, calculate the remaining budget $r \leftarrow B - n$, and set the number of true evaluations to $q_g \leftarrow \min(2, r)$.
Step 7	If the number of evaluated candidates in H is at least $N_{min} = 8$, fit the Extra Trees surrogate using all evaluated pairs in H . Otherwise, skip surrogate fitting and use the trust-region-based candidate-selection fallback.
Step 8	If $r \leq B_{micro} = 12$, activate micro-polish: center the CMA-ES mean on the current K^* , set $\lambda_g = 2\lambda = 64$, and use $\sigma_{sample} = 0.40\sigma$. Otherwise, set $\lambda_g = 32$ and $\sigma_{sample} = \sigma$.
Step 9	Generate λ_g virtual candidates from the CMA-ES distribution in the normalized search space. Virtual candidates do not increase n .
Step 10	Remove candidates that duplicate or nearly duplicate previously evaluated candidates in H
Step 11	Use the Extra Trees ensemble to predict the mean objective $\hat{\mu}(z)$ and inter-tree standard deviation $\hat{\sigma}(z)$ of each remaining virtual candidate.
Step 12	Rank the virtual candidates using $S(z) = \hat{\mu}(z) - \kappa\hat{\sigma}(z) + 10[\max(0, d_{TR}(z)/R - 1)]^2$.
Step 13	Select the q_g candidates with the lowest scores. For each selected candidate, if $n < B$, run one AGV episode, calculate its objective J , append (K, J) to H , and update $n \leftarrow n + 1$.
Step 14	Update (K^*, J^*) using the lowest observed objective value in H .
Step 15	Select the $\mu = 6$ best unique elites from H , and update the CMA-ES mean and covariance using the selected elites.
Step 16	Update the CMA-ES sampling scale and trust-region radius according to whether J^* improved.
Step 17	If the failure streak reaches the predefined stagnation threshold, restart the CMA-ES around K^* , reset $C \leftarrow I$, and set $\sigma \leftarrow 0.35$
Step 18	End while. If $n < 57$, return to Step 6. When $n=57$, terminate the optimization without evaluating additional candidates.
Output	Best observed PID vector $K^* = (K_p^*, K_i^*, K_d^*)$ and its objective value J^* .

2.8. Experimental Design, Fairness, and Validation

Fig. 3 shows an experimental AGV prototype used as a hardware validation platform to test the control system and optimize PID parameters. Fig. 3(a) shows a side view of the prototype, where several key components are integrated on a single acrylic frame, including a

Raspberry Pi, Arduino Uno, motor driver module, battery power supply, DC motor, drive wheels, and a sensor array on the front. This arrangement indicates that the prototype is designed as a compact mobile mechatronics system, with electronic components and actuators placed centrally to maintain mechanical balance during movement. The drive wheels and motor control module allow the AGV to perform differential maneuvers, while the front sensor module serves as the main input for detecting trajectory or environmental changes during navigation.

Fig. 3(b) shows the front view of the prototype, emphasizing the sensor position at the front of the AGV. Front-mounted sensor placement is important because it allows the system to obtain trajectory information early, before the vehicle body passes through the path, enabling real-time control corrections. This hardware architecture supports a closed-loop experimental scenario in which the control system reads sensor data, processes it into correction signals based on PID parameters, and then sends them to the left and right motors to adjust the AGV's direction. With this configuration, this prototype provides an objective evaluation platform to compare the performance of PID tuning methods, particularly based on indicators such as trajectory tracking accuracy, motion stability, overshoot, actuation smoothness, and trajectory maintenance ability under real-world experimental conditions.

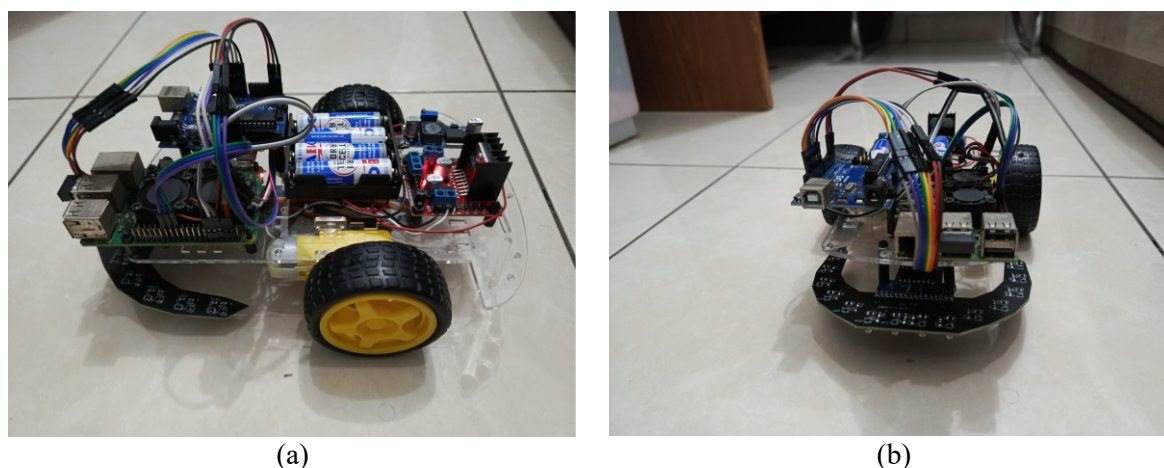


Figure 3. (a) Side view of AGV prototype, (b) Front view of AGV prototype

Fig. 4 shows the experimental track layout used to evaluate the AGV's path-tracking capabilities at two levels of geometric complexity. Fig. 4(a) shows an 8Shape track, which is more challenging because it combines straight segments, sharp turns, repeated direction changes, and a central intersection. This configuration tests the control system's dynamic response to rapid curvature changes, particularly when the AGV must make consecutive course corrections in a short period. Thus, the 8Shape track provides a more complex test scenario for assessing control stability, line-following ability, overshoot tendencies on turns, and the system's resilience to tracking errors as the vehicle moves from one track segment to another.

Fig. 4(b) shows the Circle trajectory, a closed pattern dominated by continuous turns and smoother directional transitions than the 8Shape trajectory. This trajectory was used to evaluate the AGV's ability to maintain stable motion on a circular path, particularly in balancing speed, steering correction, and consistent line sensor readings. Compared to the 8Shape trajectory, the Circle trajectory provides a more regular control load, making it suitable for observing the PID system's basic performance under relatively consistent curvature conditions. Using these two trajectories allows for a more comprehensive evaluation because the PID tuning method is tested not only under simple, stable conditions but also under track conditions that require more

aggressive control adaptation. Therefore, the experimental setup in Fig. 4 plays a crucial role in assessing the AGV's objective performance based on metrics such as tracking error, actuation smoothness, overshoot, lost-line count, and aggregate objective value.

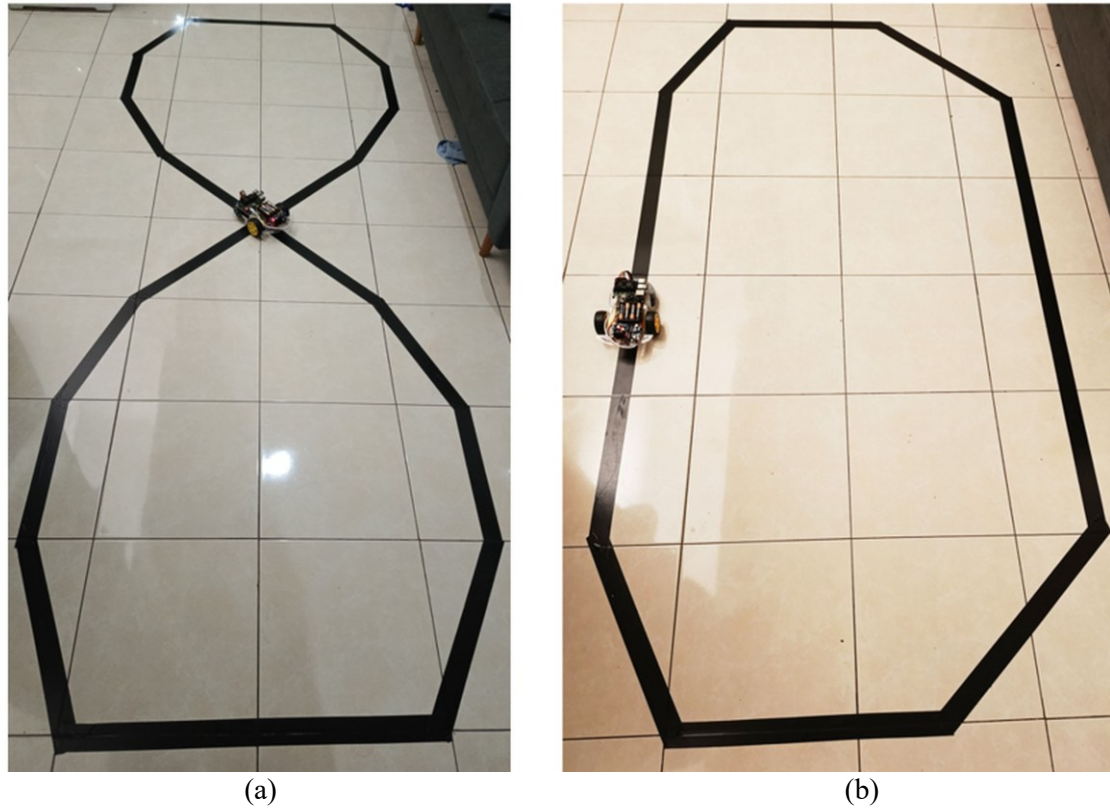


Figure 4. (a) 8Shape track, (b) Circle track

Table 2 lists the parameter settings used for the comparison. We repeated each method-track combination with 20 random seeds, and limited each run to 57 true evaluations. We held the PID domain, objective function, line threshold, evaluation duration, base speed, anti-windup policy, error filter, and initial conditions constant. We did not count surrogate predictions as true evaluations because they were not executed on the AGV. The proposed method could generate and score many candidates internally, but only physically executed candidates consumed the budget. Pure CMA-ES, PSO, and GA were subject to the same 57-call limit, so the experiment compares sample efficiency under an identical physical budget.

Table 2. Reproducibility and fairness parameters for all methods

Group	Parameter	Value	Fairness Objective
PID Domain	PID Search Limit: Kp_Bounds, Ki_Bounds, Kd_Bounds	$K_p \in [0, 30]$ $K_i \in [0, 3]$ $K_d \in [0, 10]$	The platform-specific bounds were set by preliminary safety tests and applied identically to all methods where they are not universal AGV limits.
Line sensor	Num_Sensors, Line_Threshold, Use_Raw_Weight, sensor_positions	15 IR sensors, threshold 400, raw weighting True, sensor position $[-1, 1]$	The threshold was calibrated once for the tested sensor height, surface, and lighting and then fixed for all methods and runs.

Group	Parameter	Value	Fairness Objective
Control and Actuator	Sampling interval/control frequency, base speed, actuator limits, integral clamp, derivative filter, initial/reset condition, and lost-line recovery.	$Dt=0.032s$, $fc=31.25Hz$, $vbase=3$ rad/s, clipped to ± 6.28 rad/s, $\alpha_e = 0.30$, clip $I(k) = -2$ to 2 , $\alpha_D = 0.20$, $e_f = 0$, $e_{last} = 0$, $e_m(k) = 0$	All control and actuator parameters are applied identically to all methods.
Budget	Eval_Budget, episode duration, objective weight	57 real evaluations, Evaluation_Duration_Ms = 20000, $W_Iac=1.0$, $W_Rms=0.10$, $W_Os=0.20$, $W_Sm=0.0005$, $W_Pen=0.01$	All optimizers receive the same number of evaluations and use the same objective function, so the method's advantage comes from search efficiency, not a larger budget.
CMA-ES Tree	Sa_Init_Evals, Sa_Cma_Lambda, Sa_True_Evals_Per_Gen, Sa_Cma_Mu, Sa_Sigma_Init, Sa_Sigma_Min, Sa_Sigma_Max, Sa_Cov_Learning_Rate, Sa_Stagnation_Restart, Sa_Micro_Polish_Last number of trees M, max_features, min_samples_leaf, maximum depth, Random state	Initial true eval 9, virtual candidates 32, true eval/generation 2, elite 6, initial sigma 0.20, sigma min-max 0.015–0.55, covariance learning rate 0.28, stagnation restart default 8, but sweep seeds 2–10, micro-polish last 12 eval, M=1000, max_features=1.0, min_samples_leaf=1, max-depth=none, rand_state=s+1000+g	CMA-ES Tree uses virtual candidates for exploration, but only candidates actually tested in the field are counted toward the budget. This maintains fairness compared to other methods, which are also limited to 57 true evaluations.
Pure CMA-ES	Cma_Lambda, Cma_Mu, Cma_Sigma_Init, Cma_Sigma_Min, Cma_Sigma_Max, Cma_Cov_Learning_Rate, Cma_Stagnation_Restart, Cma_Micro_Polish_Last	Population $\lambda=6$, elite $\mu=3$, initial sigma 0.20, sigma min-max 0.015–0.55, covariance learning rate 0.28, stagnation restart default 8, sweep seeds 2–10, micro-polish last 12 eval	All Pure CMA-ES candidates are fully evaluated directly in AGV, with no Extra Trees surrogates. Therefore, each candidate consumes a real budget directly.
PSO	Swarm_Size, Pso_Iters, W_Max, W_Min, C1, C2, Vmax_Frac	Initial Default Swarm_Size=12, Pso_Iters=8, But Automatically Selected to Have Swarm_Size $\times$ Pso_Iters = 57, Effective= $19 \times 3 = 57$, Inertia $0.90 \rightarrow 0.40$, C1=1.80, C2=1.80, Vmax_Frac=0.20	PSO is automatically adjusted so that the number of particle evaluations exactly matches the budget of 57, so it does not get more evaluations than other methods.
GA	Fair_Auto_Params, Pop_Size, Generations, Elite_N, Tournament_K, Crossover_Rate, Mutation_Rate_Gene, Mut_Sigma_Frac	Fair_Auto_Params=True, Initial Default Is Pop_Size=14, Generations=7, But Is Automatically Selected to Have Pop_Size $\times$ Generations = 57, Effective: $19 \times 3 = 57$, Elite 2, Tournament 3, Crossover 0.90, Mutation Gene 0.25, Mutation Sigma Fraction 0.12	The GA is made fair by adjusting the population size and number of generations so that total individual evaluations equal the budget of 57, rather than following the default configuration, which can exceed the budget.

3. RESULTS

This section interprets the convergence, distribution, and time-response results from CMA-ES Tree and the three baselines. It focuses on the final objective J, convergence efficiency, cross-seed distribution, and time-domain trade-offs. Fig. 5(a) shows the

convergence curves of the mean best-so-far objective J on the Circle path for four methods: CMA-ES Tree, CMA-ES Pure, PSO, and GA. In general, all methods decrease J as the evaluation budget increases. This indicates that the optimization process has successfully found a better combination of PID parameters. However, CMA-ES Tree shows the lowest observed mean best-so-far objective across all methods at the evaluated budgets. From the start of the evaluation, this method reduces the objective value sharply and achieves the lowest final value compared to the other methods. CMA-ES Pure requires a larger budget to decrease due to its very high initial value, while PSO and GA decrease more steadily but stop at a higher final value than CMA-ES Tree. This shows that on the Circle trajectory, the Tree-based surrogate-assisted approach uses the evaluation budget more efficiently to find the best PID parameters.

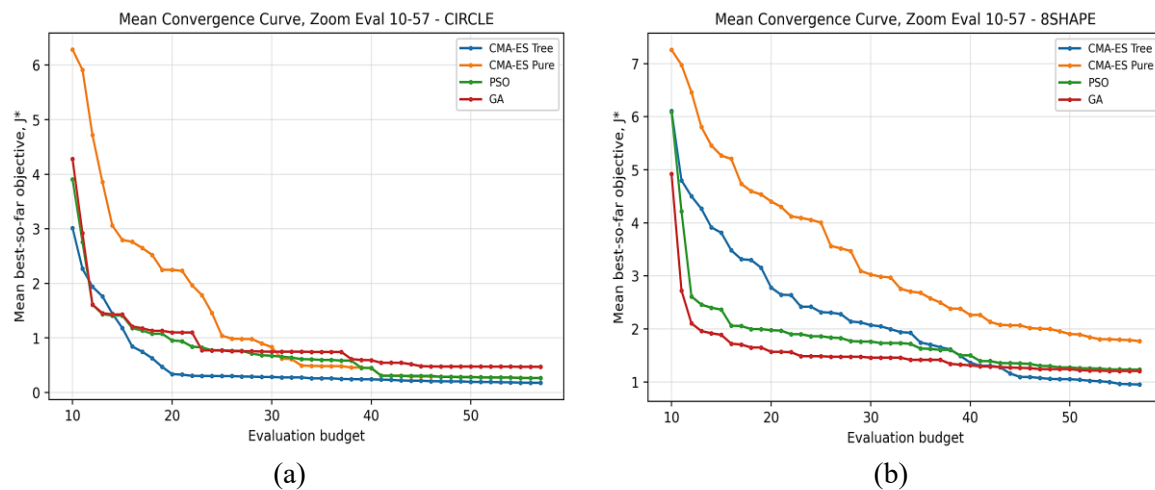


Figure 5. (a) Average convergence in circle track (b) Average convergence in 8Shape track

Fig. 5(b) shows a more challenging pattern on the 8Shape trajectory, as the initial objective value is higher and the decrease is slighter than on the Circle trajectory. This is understandable because the 8Shape trajectory has more complex directional changes and curvature, so PID tuning requires stronger adaptation. On this trajectory, CMA-ES Tree still shows the best performance at the end of the evaluation, with the lowest final J value and a curve that continues to decrease until it approaches the maximum budget. GA is also competitive because it decreases rapidly at the beginning of the evaluation but then tends to stagnate. PSO shows a gradual decrease, while CMA-ES Pure is the slowest method because its objective value remains the highest until the end of the evaluation. Thus, these two graphs confirm that CMA-ES Tree has the best convergence efficiency and solution quality on both simple and complex trajectories, especially in fixed-budget scenarios with a limited number of evaluations.

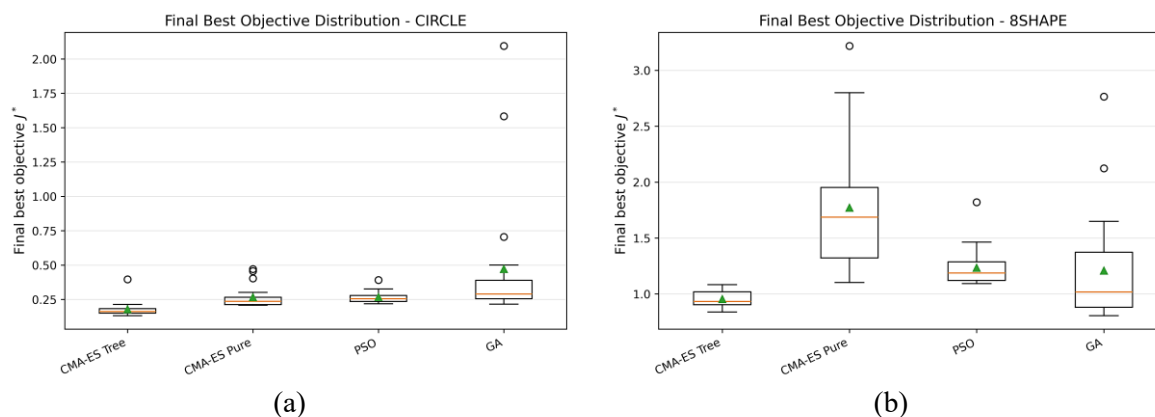


Figure 6. (a) Final best objective in circle track (b) Final best objective in 8Shape track

Fig. 6(a) shows the final distribution of the best objective J on the Circle path after the entire evaluation budget has been used. A lower J value indicates better PID tuning quality. On this path, CMA-ES Tree performs best because it has the lowest median, the narrowest data distribution, and few outliers. This indicates that CMA-ES Tree not only finds solutions with small objective values but is also more consistent across runs. CMA-ES Pure and PSO have slightly higher medians, with relatively controlled variation, while GA shows the widest distribution and several large outliers. The outliers in GA indicate that in some runs, this method failed to find the best-observed combination of PID parameters within the same evaluation budget.

Fig. 6(b) shows that the 8Shape trajectory is more challenging than the Circle trajectory, as evidenced by the generally higher final J values across all methods. On this trajectory, CMA-ES Tree again performs best, with the lowest median and the most stable distribution compared to other methods. GA has a fairly competitive median, but its spread is larger and includes high outliers, resulting in less consistent performance. PSO shows intermediate results with relatively small variations, but its objective median is still higher than CMA-ES Tree. Meanwhile, CMA-ES Pure has the highest final objective value and the largest distribution, indicating that without the assistance of a surrogate Tree, the search process requires more real-world evaluations to reach a competitive solution. Thus, this boxplot confirms that CMA-ES Tree is more consistent across the tested runs and more budget-efficient at producing the best PID parameters, on both simple and complex trajectories.

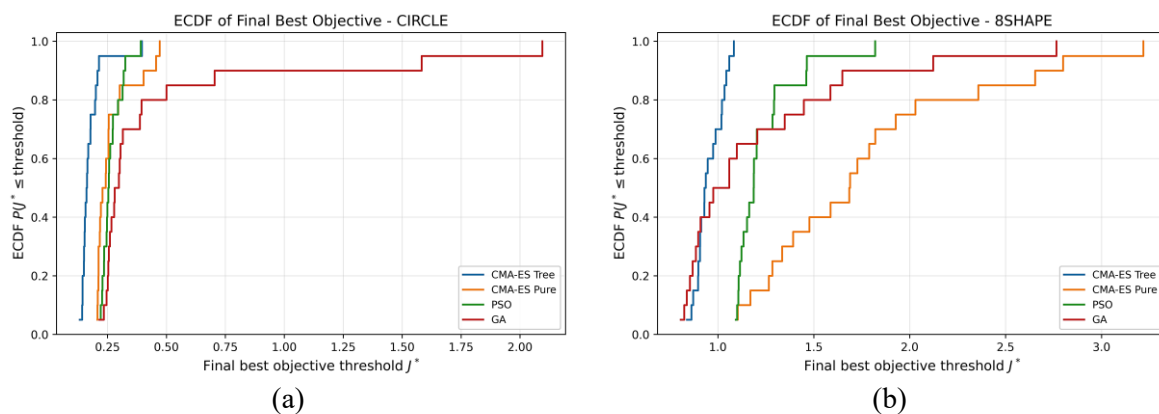


Figure 7. (a) ECDF in circle track (b) ECDF in 8Shape track

Fig. 7(a) shows the ECDF of the final best objective J on the Circle path. The ECDF measures the probability that a method produces a final objective value less than or equal to a given threshold. Therefore, a curve farther to the left and rising more quickly to 1 indicates a better and more consistent method. On the Circle path, CMA-ES Tree is the best method because its curve is farthest to the left and rises very sharply. This means that most CMA-ES Tree runs have already achieved low J values at a small threshold. PSO and CMA-ES Pure are in the middle, while GA has a very long distribution tail up to large J values. The long tail for GA indicates that while some runs can produce reasonably good solutions, others fail to achieve optimal performance within the same evaluation budget.

Fig. 7(b) shows that the 8Shape path is more difficult than the Circle path because the entire curve shifts to the right, meaning the final objective values are generally higher. On this path, CMA-ES Tree remains superior because its curve is farthest to the left and reaches a cumulative probability of 1 at the lowest threshold. This means all CMA-ES Tree runs produced relatively good and stable final solutions. PSO came in second because its curve still rose quite quickly but required a larger threshold J than CMA-ES Tree. GA showed quite

competitive initial performance in some runs, but had a long right tail, indicating performance variation and the risk of bad runs. Meanwhile, CMA-ES Pure was the weakest method on 8Shape because its curve was farthest to the right and required the highest objective threshold to cover all runs. Overall, this ECDF graph confirms that CMA-ES Tree has the best reliability in fixed-budget scenarios, especially because it produced consistently low objective values on both simple and complex paths.

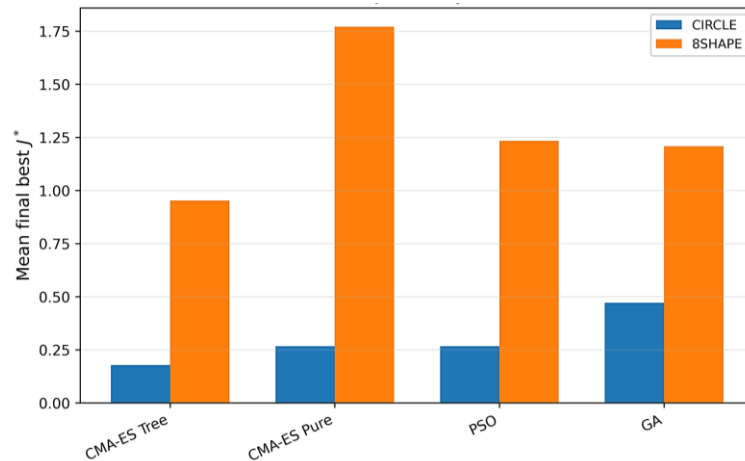


Figure 8. Average final best objective value by method and trajectory.

Fig. 8 compares the mean final best objective J by optimization method and trajectory type. Because J represents the objective function to be minimized, a smaller J value indicates better-quality PID parameters. On the Circle trajectory, all methods produced relatively low final objective values compared to the 8Shape trajectory. This indicates that the Circle trajectory is easier to optimize due to its simpler motion pattern and more stable curvature changes. Among all methods on the Circle trajectory, CMA-ES Tree obtained the lowest average J value, followed by CMA-ES Pure and PSO, which were nearly equal. Meanwhile, GA produced the highest average J value on the Circle trajectory, making its performance less efficient than the other methods under simple trajectory conditions.

On the 8Shape trajectory, all methods experienced an increase in J values. Nevertheless, CMA-ES Tree still shows the best performance with the lowest average J value compared to CMA-ES Pure, PSO, and GA. PSO and GA are at an intermediate level with relatively close values, while CMA-ES Pure produces the highest objective value on 8Shape, indicating that pure CMA-ES search without surrogate assistance requires more evaluations to reach a competitive solution. Overall, this graph confirms that CMA-ES Tree is more effective and budget-efficient at producing the best PID parameters on both simple and complex trajectories.

Fig. 9 shows the RMSE distribution of the best candidates for each run for the four PID optimization methods on two trajectory types, Circle and 8Shape. In general, the Circle trajectory exhibits lower RMSE than the 8Shape trajectory. On the Circle trajectory, all methods fall within a narrow error range, but CMA-ES Tree Circle appears to perform best due to its lowest median RMSE and very narrow distribution. PSO and CMA-ES Pure are also quite stable, while GA Circle has a slightly higher median and several large outliers, making its performance less consistent than the other methods.

On the 8Shape trajectory, the differences between the methods become more pronounced. CMA-ES Tree 8Shape performs best, with the lowest median RMSE among the methods and a relatively narrow distribution indicating more stable tuning results across runs. GA 8Shape has a fairly competitive median, but its spread is wider, and outliers reduce consistency.

Meanwhile, PSO 8Shape and especially CMA-ES Pure 8Shape show higher RMSE, with large variations and outliers. Thus, this graph supports that the CMA-ES Tree approach is more effective and more consistent across the tested runs, especially on complex trajectories, as it produces PID candidates with smaller and more stable tracking errors.

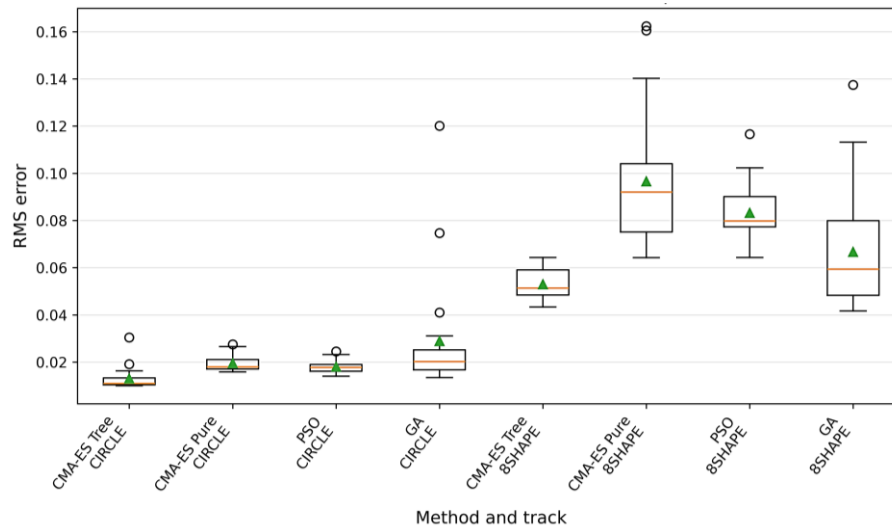


Figure 9. Distribution of RMSE of the best candidate.

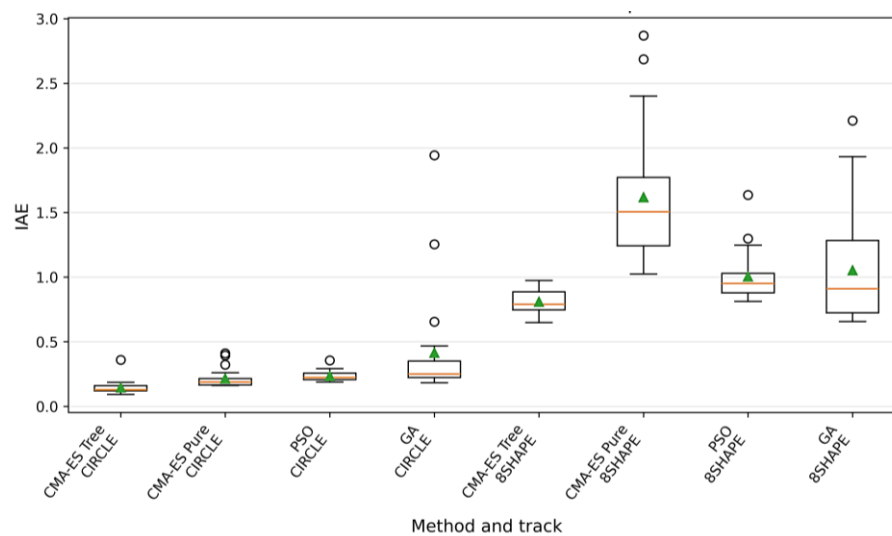


Figure 10. Distribution of the best candidate IAEs.

Fig. 10 shows the distribution of the Integral Absolute Error (*IAE*) of the best PID candidate for each run for the four optimization methods on the Circle and 8Shape trajectories. *IAE* represents the accumulated total tracking error over the course of an episode. The lower the *IAE*, the better the *AGV* maintains its position relative to the track line. On the Circle trajectories, all methods produced relatively low *IAE*. Among the methods on the Circle trajectories, CMA-ES Tree performed best, with the lowest median *IAE* and the narrowest data distribution. CMA-ES Pure and PSO remained relatively stable but had slightly higher medians. Meanwhile, the Circle GA showed greater variation and several high outliers, indicating that the GA method produced suboptimal PID parameters in some runs.

On the 8Shape trajectories, *IAE* values increased for all methods. CMA-ES Tree 8Shape still produces the lowest *IAE* and a relatively narrow distribution compared to other methods,

so it can be said to be more consistent in minimizing error accumulation. PSO 8Shape is in the middle, with a higher median IAE than CMA-ES Tree but more stable than GA. GA 8Shape has quite wide variations and several high outliers, while CMA-ES Pure 8Shape shows the highest IAE and the largest distribution, indicating less consistent performance within a limited evaluation budget. Overall, this graph confirms that CMA-ES Tree is more effective and more consistent across the tested runs in producing PID parameters with low accumulated tracking error on both simple and complex trajectories.

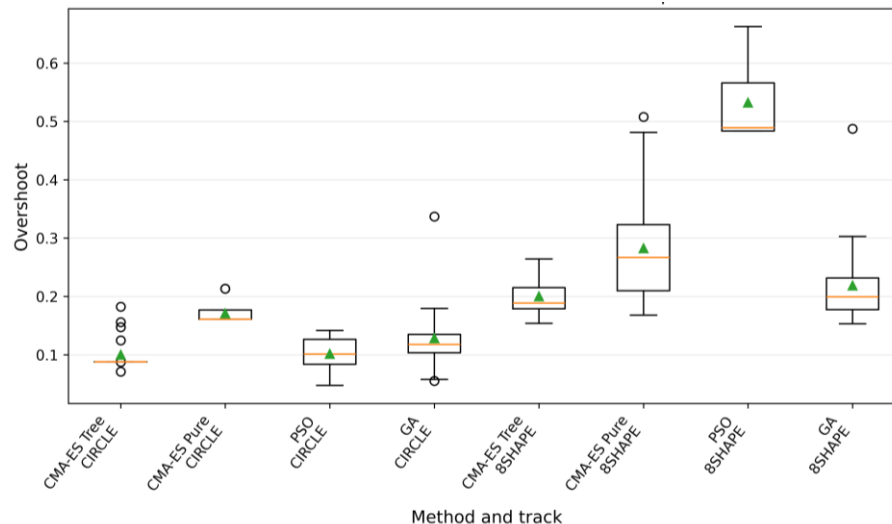


Figure 11. Distribution of overshoot of the best candidates.

Fig. 11 shows the overshoot distribution of the best PID candidates for each run for the CMA-ES Tree, CMA-ES Pure, PSO, and GA methods on the Circle and 8Shape trajectories. Overshoot represents the maximum deviation of the AGV response when it crosses or overshoots the reference position of the trajectories. The smaller the overshoot value, the better the stability of the PID control. On the Circle trajectories, the overshoot values for all methods were relatively low compared to the 8Shape trajectories. CMA-ES Tree Circle and PSO Circle performed best due to their low median overshoot. However, CMA-ES Tree had several outliers, indicating that higher overshoot occurred in a small portion of the runs. CMA-ES Pure Circle had a very narrow distribution, but its median overshoot was higher than those of CMA-ES Tree and PSO. Meanwhile, GA Circle was intermediate, with one relatively high outlier.

On the 8Shape trajectory, the overshoot value increases across all methods. CMA-ES Tree 8Shape produces relatively low and stable overshoot compared to CMA-ES Pure and PSO. This indicates a better ability to prevent overly aggressive AGV responses. GA 8Shape also has a relatively low median overshoot, but high outliers indicate potential instability in some runs. In contrast, PSO 8Shape has the highest overshoot, with a large median and spread, indicating that the PSO-derived PID parameters tend to produce more aggressive control responses on complex trajectories. Overall, this graph confirms that CMA-ES Tree is more consistent across the tested runs in suppressing overshoot, especially on complex trajectories, making it more suitable for AGV PID tuning in fixed-budget scenarios.

Fig. 12(a) shows the filtered tracking error on the Circle trajectory for the best candidate for each method. The main comparison is based on the oscillation amplitude, stability, and number of spikes on each curve. On the Circle trajectory, CMA-ES Tree showed the most stable response because, after the initial transient, the error rapidly decreased and the curve remained relatively flat. CMA-ES Pure still exhibited fairly consistent periodic oscillations,

PSO had some fluctuations and spikes, while GA showed coarser error variations and was not as smooth as CMA-ES Tree. This indicates that the PID parameters from CMA-ES Tree kept the AGV closer to the trajectory line more smoothly on simple trajectories.

Fig. 12(b) shows conditions on the 8Shape trajectory. On this trajectory, all methods experienced more noticeable fluctuations in tracking error than on the circle track, particularly at the beginning and at several trajectory transition points. Nevertheless, CMA-ES Tree still produces the most controlled error pattern, with relatively small oscillation amplitudes and no prolonged large deviations. CMA-ES Pure shows larger ripples and error changes; PSO experiences a fairly strong initial transient and fluctuations at several time intervals, while GA shows higher error variations and tends to be less smooth. Overall, this graph confirms that CMA-ES Tree produces a more stable and smoother tracking response on both Circle and 8Shape paths, making it more effective for AGV PID tuning in maintaining path-tracking accuracy.

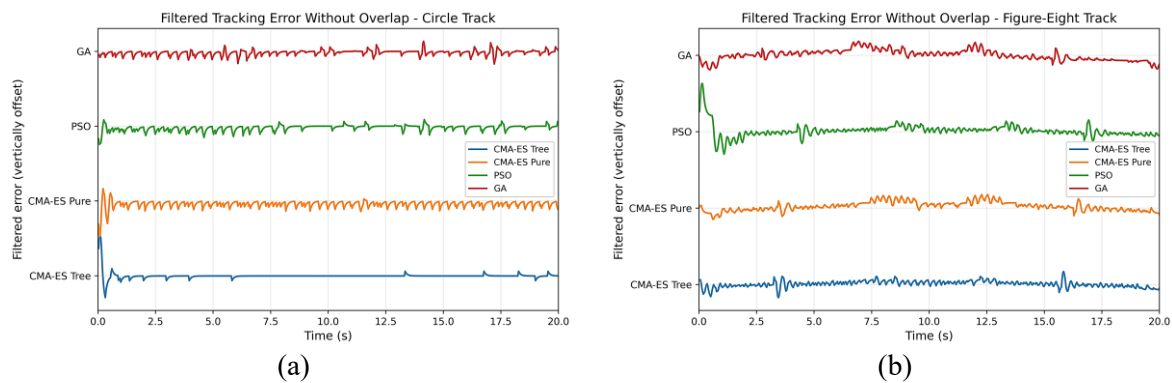


Figure 12. (a) Filtered Tracking Error in Circle track (b) Filtered Tracking Error in 8Shape track

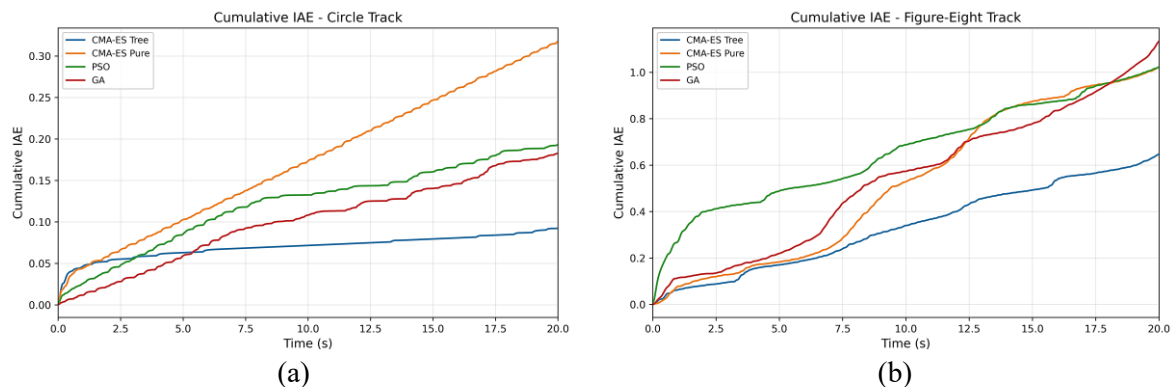


Figure 13. (a) Cumulative IAE in circle track (b) Cumulative IAE in 8Shape track

Fig. 13(a) shows the cumulative IAE for the Circle trajectory, which is the accumulation of absolute error during one AGV operating cycle. A lower curve indicates a smaller total tracking error accumulated over the entire episode. On the Circle trajectory, CMA-ES Tree produced the lowest cumulative IAE until the end, around 0.09, and the curve flattened relatively gradually after the initial phase. This indicates that the CMA-ES Tree PID candidate maintained a consistently small error throughout the trajectory. GA and PSO ranked in the middle with larger final error accumulations, while CMA-ES Pure had the highest cumulative IAE, around 0.32, indicating that errors were small but repeated over time, resulting in the largest accumulation.

Fig. 13(b) shows the cumulative IAE on the 8Shape trajectory. On this trajectory, CMA-ES Tree remains the best method because its curve is the lowest until the end, around 0.65, resulting in the smallest total tracking error. PSO experiences a very rapid increase in error at the beginning, while CMA-ES Pure and GA show a sharp increase in the middle to the end of the trajectory. GA even reaches the highest cumulative IAE at the end. Overall, this graph confirms that CMA-ES Tree is more effective at suppressing the accumulation of tracking error on both simple and complex trajectories.

4. DISCUSSION

Figs. 5-8 indicate that the principal advantage of CMA-ES Tree was how it allocated the fixed true-evaluation budget. CMA-ES retained a broad virtual search over the normalized PID domain, whereas Extra Trees screened the virtual population before physical execution. The real platform was therefore used mainly to evaluate candidates predicted to be promising or informative under the inter-tree dispersion term. Pure CMA-ES, PSO, and GA had no comparable screening stage, so every candidate immediately consumed a true evaluation. The observed convergence advantage should consequently be interpreted as improved sample allocation under an equal physical budget, not as an advantage obtained from additional AGV trials.

Trajectory geometry changed the difficulty of the search and exposed different optimizer behavior. The Circle track imposed relatively regular curvature, allowing all methods to enter a lower-objective region. The alternating curvature, direction changes, and central transition of the 8Shape track generated a more heterogeneous mapping between PID gains and tracking performance. Under that condition, the surrogate-assisted method continued to improve late in the budget, whereas the baselines more often plateaued or retained broad final distributions. This pattern supports the proposed screening mechanism for the two tested trajectories, but it does not establish that the same ranking will hold for every path geometry.

The boxplots and empirical cumulative distribution functions provide information beyond the mean final objective. Their left-shifted and comparatively concentrated patterns show that CMA-ES Tree was more likely to return a low-objective solution across repeated seeds. This cross-run reliability matters in experiment-based tuning because a single favorable run does not show that an optimizer can repeatedly find a usable controller within a restricted budget. The longer tails observed for some baselines indicate a greater risk of ending the experiment with a poor gain set, even when their best runs were competitive.

Figs. 9-13 summarize the distributions of time-domain-derived performance metrics and reveal trade-offs rather than uniform dominance across all metrics. CMA-ES Tree produced the most consistent reduction in RMSE, accumulated error, and cumulative IAE, particularly on the 8Shape track. Its overshoot, however, was not the minimum in every condition. PSO was competitive on Circle, and GA showed a competitive median on 8Shape but with less reliable tails. This distinction matters because the optimizer minimizes a composite objective that also includes actuation smoothness and a lost-line penalty. The selected controller should therefore be viewed as a budget-constrained compromise among tracking accuracy, extreme deviation, command variation, and recovery reliability rather than a solution optimized for one metric alone.

From an implementation perspective, the two-level architecture keeps the time-critical PID loop on the Arduino, while training and candidate selection occur between episodes on the Raspberry Pi. This separation preserves deterministic inner-loop execution and permits the supervisory tuner to use more virtual candidates than could be tested physically. Extra Trees is

suitable for the small nonlinear tabular dataset because it retrains quickly and does not assume a smooth objective. Nevertheless, inter-tree dispersion is a heuristic indicator rather than a calibrated error probability. Its value depends on the diversity and coverage of the samples accumulated.

The methodological contribution is therefore the integration of explicit true-evaluation accounting, uncertainty-informed surrogate ranking, and embedded physical validation under a common budget. We do not claim that CMA-ES, Extra Trees, PID control, or the individual performance metrics are new. The evidence instead shows that their specific integration can improve the probability of finding a useful PID setting when physical evaluations are selective and expensive. This positioning aligns the contribution with the scope of the experiment and avoids treating two-track validation as a universal AGV framework.

Several limitations constrain the interpretation of these findings. The experiments used one line-following differential-drive prototype, sensor arrangement, two trajectories, and fixed hardware and operating settings. The PID bounds and the threshold $\tau = 400$ were calibrated for the tested actuator limits, sensor height, surface, and lighting, so they should be recalibrated for another platform. The study did not vary steering architecture, payload, base speed, floor friction, battery state, sensor layout, or environmental illumination. In addition, surrogate quality depends on the coverage and noise level of the initial and sequential data, while inter-tree dispersion is not a formally calibrated uncertainty measure. Future work should validate the method on longer industrial routes and multiple AGV platforms, include systematic load and speed changes, and conduct ablation and sensitivity studies for the surrogate score, q , initialization, and other supervisory design choices.

5. CONCLUSION

This study evaluated an Extra Trees surrogate-assisted CMA-ES supervisory tuner for PID gain selection on a line-following differential-drive AGV under an identical budget of 57 true evaluations. By ranking a larger virtual CMA-ES population and physically executing only selected candidates, the method achieved the best end-of-budget objective and the most favorable cross-run reliability on both tested trajectories. The time-domain evidence further showed that the selected gains reduced tracking-error accumulation while maintaining a practical compromise with overshoot, command smoothness, and line-loss protection. The results support surrogate screening as a sample-efficient supervisory strategy for the tested embedded platform when physical trials are costly and limited. They are not a claim of universal performance across AGV architectures. The study therefore provides a focused experimental basis for broader multi-platform and variable-operating-condition validation.

ACKNOWLEDGEMENT

This research was funded by the Ministry of Higher Education (MoHE) Malaysia through the Fundamental Research Grant Scheme (FRGS), grant number FRGS/1/2021/TK0/UKM/02/17. We also acknowledge support from the National Research and Innovation Agency (BRIN) Indonesia and the Indonesia Endowment Fund for Education Agency (LPDP) under the Perjanjian Pendanaan Program Riset dan Inovasi untuk Indonesia Maju Gelombang 4 tahun 2023.

REFERENCES

- [1] Oyekanlu EA, et al. (2020) A review of recent advances in automated guided vehicle technologies: Integration challenges and research areas for 5G-based smart manufacturing applications. *IEEE Access*, 8:202312-202353.
- [2] Dos Reis WPN, Couto GE, Junior OM. (2023) Automated guided vehicles position control: A systematic literature review. *Journal of Intelligent Manufacturing*, 34(4):1483-1545.
- [3] Alwan MMA, Green AA, Noori AS, Aldair AA. (2021) Design and implementation of line follower Arduino mobile robot using Matlab Simulink toolbox. *Iraqi Journal for Electrical and Electronic Engineering*, 17(2):11-16.
- [4] Moshayedi AJ, Abbasi A, Liao L, Li S. (2019) Path planning and trajectory tracking of a mobile robot using bio-inspired optimization algorithms and PID control. 2019 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 1-6.
- [5] Moshayedi AJ, et al. (2022) Simulation and validation of optimized PID controller in AGV (Automated Guided Vehicles) model using PSO and BAS algorithms. *Computational Intelligence and Neuroscience*, 2022(1):7799654.
- [6] Henclová K. (2019) Using CMA-ES for tuning coupled PID controllers within models of combustion engines. *Neural Network World*, 29(5):325-344.
- [7] Chen JC, Chen TL, Teng YC. (2021) Meta-model based simulation optimization for automated guided vehicle system under different charging mechanisms. *Simulation Modelling Practice and Theory*, 106:102208.
- [8] Krawczyk K, Arabas J. (2025) Single-objective surrogate models for continuous metaheuristics: An overview. *Applied Sciences*, 15(16):9068.
- [9] Patil RS, Jadhav SP, Patil MD. (2024) Review of intelligent and nature-inspired algorithms-based methods for tuning PID controllers in industrial applications. *Journal of Robotics and Control (JRC)*, 5(2):336-358.
- [10] Aliskan I. (2025) The optimization-based fuzzy logic controllers for autonomous ground vehicle path tracking. *Engineering Applications of Artificial Intelligence*, 151:110642.
- [11] Yu J, Liang M, Peng W, Wu T, Rong C, Zhang D. (2023) Speed control based on an improved PID controller with BP neural network for two wheel differential AGV system. 2023 IEEE 6th Student Conference on Electric Machines and Systems (SCEMS), 1-5.
- [12] Ha VT, Thuong TT, Truc LN. (2024) Trajectory tracking control based on genetic algorithm and proportional integral derivative controller for two-wheel mobile robot. *Bulletin of Electrical Engineering and Informatics*, 13(4):2348-2357.
- [13] Demir MH, Demirok M. (2023) Designs of particle-swarm-optimization-based intelligent PID controllers and DC/DC buck converters for PEM fuel-cell-powered four-wheeled automated guided vehicle. *Applied Sciences*, 13(5):2919.
- [14] Jangid P, Nagar V, Sharma A, Nagar S, Palwalia DK. (2025) Design a PID controller based on grey wolf optimization algorithm for single-area load frequency control. *Journal of Emerging Science and Engineering*, 3(1):e36-e36.
- [15] Mosaad AM, Attia MA, Abdelaziz AY. (2019) Whale optimization algorithm to tune PID and PIDA controllers on AVR system. *Ain Shams Engineering Journal*, 10(4):755-767.
- [16] Van Niekerk JA, le Roux JD, Craig IK. (2023) On-line automatic controller tuning of a multivariable grinding mill circuit using Bayesian optimisation. *Journal of Process Control*, 128:103008.
- [17] Tumpach J, Koza J, Holeňa M. (2023) Improving optimization with Gaussian processes in the covariance matrix adaptation evolution strategy. *ITAT'23: Information Technologies – Applications and Theory*, 3498(10).

- [18] Khouzani FF, Mirzaei A, Plante PL, Gewali L. (2025) CMA-ES with radial basis function surrogate for black-box optimization. ITNG 2025. Advances in Intelligent Systems and Computing. 1463:367-379.
- [19] Bujgoi G, Sendrescu D. (2025) Tuning of PID controllers using reinforcement learning for nonlinear system control. Processes, 13(3):735.
- [20] Van Niekerk JA, le Roux JD, Craig IK. (2025) Reinforcement learning based automatic tuning of PID controllers in multivariable grinding mill circuits. Control Engineering Practice, 165:106522.

Proof-of-Concept Validation in Healthy Participants of an sEMG-Controlled 4-DOF Upper Limb Exoskeleton for Post-Stroke Rehabilitation

ANG LING WEAY* AND TANG MING

School of Artificial Intelligence and Computing, Malaysia University of Science and Technology (MUST), Block B, Encorp Strand Garden Office, 12, Jalan PJU 5/5, Kota Damansara, 47810 Petaling Jaya, Selangor, Malaysia

**Corresponding author: ang.weay@must.edu.my*

(Received: 7 July 2026; Accepted: 2 September 2026; Published online: 11 September 2026)

ABSTRACT: Two improvements are needed in upper-limb exoskeletons (ULEs) for rehabilitation after stroke: a mechanism for the shoulder joint that will ensure interaction forces are minimized throughout the entire joint's range of motion and an easy-to-use interface that detects neural intent. These have previously been addressed separately; here, we validate them together in one integrated device, the first closed-loop 4-DOF ULE with a surface electromyography (sEMG)-based classification interface. The objective is to test, in healthy participants, whether an anatomically aligned shoulder mechanism preserves sEMG signal quality during physical coupling while a lightweight classifier recognizes movement intent in real time. The mechanism uses three revolute glenohumeral joints whose axes intersect at the shoulder center of rotation (COR), giving the basic shoulder degrees of freedom (DOF) with minimal acromion interference, plus a proximal joint emulating the sternoclavicular joint to expand the workspace; an elbow joint completes the kinematic chain. The controller identifies four activities of daily living (ADL), namely shoulder abduction (SAB), pick-up cup (PUC), push-forward cup (PFC), and rest arm (RA). These activities are recognized by classifying a 72-dimensional feature vector (6 features $\times$ 12 channels) at 16 Hz, using a 125 ms sliding time window and a linear support vector machine (SVM). In a closed-loop study with 15 healthy volunteers, recognition accuracy was 92.4%, 89.7%, and 85.1% for SVM, artificial neural network (ANN), and logistic regression, respectively; SVM recall was 96% (SAB) and 98% (RA). End-to-end latency was 148 ± 19 ms, below the accepted 200 ms threshold. Participants rated the system "Good" (SUS 71.3 ± 8.4) with only slight discomfort (NRS 1.8 ± 0.7). Misclassification was dominated by the kinematically adjacent PUC–PFC pair. As validation used healthy participants, these findings establish technical feasibility as a precursor to future clinical trials in stroke patients.

ABSTRAK: Eksoskeleton anggota atas (upper limb exoskeletons, ULE) untuk pemulihan pasca-strok memerlukan dua kemajuan: mekanisme bahu yang mengekalkan daya interaksi rendah sepanjang julat pergerakan, dan antara muka intuitif yang mengenali niat neuro-otot tanpa beban mental tambahan. Kedua-duanya sebelum ini ditangani secara berasingan; di sini kami mengesahkan gabungannya dalam satu peranti bersepadu — ULE 4-DOF gelang tertutup pertama dengan antara muka pengelasan sEMG. Objektif kajian ialah menguji, dengan peserta sihat, sama ada mekanisme bahu yang sejajar secara anatomi dapat mengekalkan kualiti isyarat sEMG semasa gandingan fizikal, sementara pengelas ringan mengenali niat pergerakan secara masa nyata. Mekanisme ini menggunakan tiga sendi glenohumeral berputar yang paksinya bersilang pada pusat putaran (center of rotation, COR) bahu dengan gangguan minimum pada akromion, serta satu sendi proksimal yang meniru sendi sternoklavikular untuk memperluas ruang kerja; satu sendi siku melengkapkan rantaian kinematik. Pengawal tersebut mengenal pasti empat jenis aktiviti kehidupan harian (Activities

of Daily Living, ADL), iaitu abduksi bahu (Shoulder Abduction, SAB), mengambil cawan (Pick-Up Cup, PUC), menolak cawan ke hadapan (Push Forward Cup, PFC), dan merehatkan lengan (Rest Arm, RA). Aktiviti-aktiviti ini dikelaskan melalui vektor ciri berdimensi 72 (6 ciri $\times$ 12 saluran) pada kadar 16 Hz, menggunakan tettingkap masa gelongsor 125 ms dan Mesin Vektor Sokongan linear (linear SVM). Dalam eksperimen gelung tertutup dengan 15 subjek sihat, ketepatan pengecaman ialah 92.4%, 89.7%, dan 85.1% masing-masing bagi SVM, rangkaian neural buatan (ANN), dan regresi logistik; kepekaan (recall) SVM ialah 96% (SAB) dan 98% (RA). Kependaman hujung-ke-hujung ialah 148 ± 19 ms, di bawah ambang 200 ms yang boleh diterima. Peserta menilai sistem ini “Baik” (SUS 71.3 ± 8.4) dengan sedikit ketidakselesaan sahaja (NRS 1.8 ± 0.7). Salah pengelasan didominasi oleh pasangan PUC–PFC yang berdekatan secara kinematik. Oleh kerana pengesanan dijalankan ke atas peserta sihat, dapatan ini merupakan langkah awal kepada ujian klinikal dengan pesakit strok.

KEYWORDS: *Upper Limb Exoskeleton, sEMG Pattern Recognition, Rehabilitation Robotics, 4-DOF Shoulder Mechanism, Closed-Loop Control.*

1. INTRODUCTION

Stroke is the second leading global cause of death and third leading cause of disability, leaving more than 70% of survivors with chronic upper limb hemiparesis [1]. Restoration of upper limb function is the primary goal of post-acute rehabilitation, yet the intensive, high-dosage training required to drive neuroplasticity is difficult to deliver consistently through manual therapy alone due to clinician shortages and therapist fatigue [2, 3]. Robotic-assisted upper limb exoskeletons (ULEs) offer high-repetition movement assistance that can be delivered, quantified, and adapted to individualized motor intentions [3].

Adoption for clinical use is restricted by the mechanical problems that persist. In the shoulder complex, there is a complicated structure, and the movement of several joints results in translation of the COR of the glenohumeral joint [3]. Exoskeletons that constrain motion to a fixed mechanical pivot generate parasitic interaction forces that lead to shear, localized pressure, and compensatory movements, ultimately resulting in device rejection or joint damage [4, 5]. Traditional configurations are also susceptible to kinematic singularities and potential collisions with the operator’s head or body as a result of complex motion tasks [5]. By designing redundant kinematic chains like 4-DOF active configurations and redundant 2R-3R-2R configurations, collision-free zones could be created and the workspace would be expanded to more closely approximate biological reachability [4].

The second problem is non-intuitive control, as most current platforms drive the upper-limb exoskeleton (ULE) along pre-programmed trajectories triggered by button presses or simple amplitude-threshold signals [6, 7]. This approach positions the patient as a passive recipient of movement rather than an active agent, which is contrary to modern neurorehabilitation theory [7]. Active volitional engagement is a prerequisite for cortical reorganization and durable functional recovery: actively generated movement produces greater performance gains and more pronounced motor-cortical reorganization than passive movement matched for kinematics [8], so the patient must intentionally initiate movement in order to drive neuroplasticity [2, 3]. Surface electromyography (sEMG) pattern recognition—the classification of spatial and temporal muscle activation into discrete intent classes—was established as a viable multifunction control strategy in the prosthetics literature [9], [10] and has since been proposed as the key enabling technology for achieving active, intent-driven rehabilitation control [3, 6].

Though both mechanical design and classifiers have been studied separately, the two have not yet been validated together as part of a single system [3, 6]. Classifiers have been evaluated based on signal obtained without any form of physical connection, whereas the design of the shoulder mechanisms has often lacked a proper intent-recognition pipeline [4, 11]. This is a critical issue, as physical connection introduces vibration and joint loading effects as well as perturbation of the electrode contacts which could affect the sEMG signal quality and thus the classification performance [4, 11]. On the other hand, the alignment of the kinematics would determine how large the perturbations could become due to the misalignment causing unnecessary interaction forces and hence the compensatory motions and rejection of the devices [3-5].

In this paper, we will be presenting a closed-loop validation of three systems – a four degree of freedom mechanical shoulder mechanism, real-time sEMG pipeline and the pattern recognition classifier – together as a single system. The validation is conducted in healthy participants as a proof-of-concept precursor to trials in stroke patients; clinical validation in the target population remains future work [4, 6].

This paper makes three scientific contributions, each addressing a distinct aspect of the integrated validation. First, the contribution to validating the mechanical design: we experimentally verify, on participants physically wearing the device, that the anatomically aligned 4-DOF shoulder mechanism supports functional activities of daily living while keeping compensatory movement low [3, 4]. Second, the contribution to validating the sEMG intent-recognition system: we quantify online, closed-loop classification accuracy and end-to-end latency for four daily-living movement classes under real device use, together with participant-reported usability and comfort [5, 6]. Third, the contribution to demonstrating the interaction between both systems through physical coupling: by comparing signal-to-noise ratio with and without device coupling on the same participants, we quantify how the mechanical design affects sEMG signal quality, and we frame this within a structured comparison of pattern-recognition versus proportional control [4, 6].

2. BACKGROUND AND FRAMING

2.1. The Glenohumeral Misalignment Problem

As shown in Fig. 1, the human shoulder complex involves four articulations that move in a coordinated fashion during arm elevation [2, 3]. The glenohumeral (GH) joint provides the dominant 3-DOF rotational motion, but its center of rotation (COR) is not stationary in the thoracic frame; specifically, the humeral head migrates dynamically during abduction as it must clear the acromion process [4]. Recent biomechanical research confirms that this translation—which can involve displacements of approximately 10 mm is essential for achieving a full range of motion [5]. The ball joint used as an exoskeleton joint does not provide a possibility for such translation and always causes kinematic incompatibility and the appearance of parasitic forces causing shear, localization of loads, and compensatory actions [4, 5]. Such interaction forces increase as the angle of abduction grows, and can cause rejection of the device or injuries to the joints themselves [4, 5].

Conventional 3R shoulder configurations, in which three revolute joints are connected in series to approximate the 3-DOF glenohumeral articulation, offer advantages in simplicity and control but suffer from three fundamental limitations (Fig. 2(a)). First, a singularity occurs when the arm reaches vertical elevation, causing the abduction and internal/external rotation axes to align and rendering independent rotation impossible. Second, the physical bulk of motors and linkages above and behind the shoulder creates collision risk with the user's head

and torso at abduction angles beyond approximately 90° , severely restricting the usable workspace. Third, even modified 3R designs with a tilted base axis (Fig. 2(b)) — which improve the abduction range — cannot resolve the singularity problem or provide the scapular mobility needed for cross-body and overhead movements.

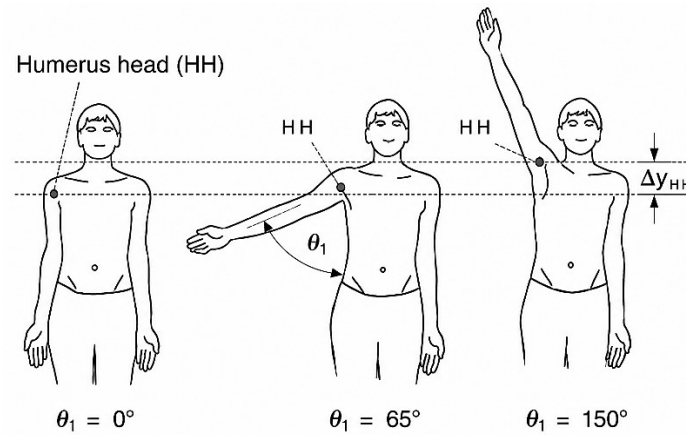


Figure 1. Location of the humeral head at different arm abduction angles in the frontal plane

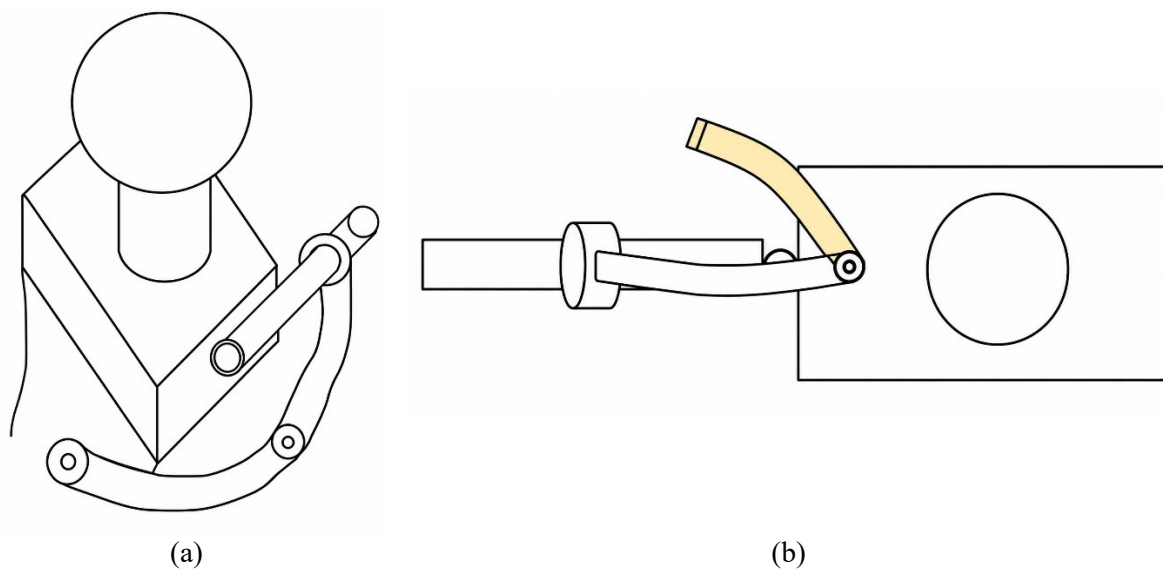


Figure 2. Comparison of serial 3R shoulder configurations: (a) conventional configuration and (b) modified configuration with a tilted base axis.

2.2. The 4R Redundant Shoulder Configuration

Adding a fourth revolute joint to create a 4R configuration introduces kinematic redundancy: the mechanism has more DOFs than are strictly required to position and orient the upper arm, providing multiple joint-configuration paths to the same end-effector pose. This redundancy can be exploited to continuously avoid singularities and to navigate around collision obstacles by selecting configurations that maintain safe clearance from the user's head and torso (Fig. 3). The fourth joint in the present design is placed proximally to replicate the protraction/retraction and elevation/depression of the sternoclavicular joint and scapular girdle, which is the primary mechanism by which the shoulder complex expands its reach for cross-body, overhead, and backward extension movements.

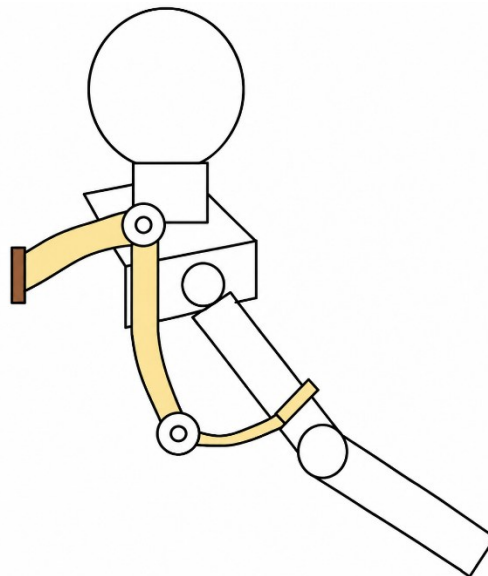


Figure 3. Four-revolute-joint (4R) redundant shoulder configuration.

2.3. Pattern Recognition vs. Proportional Control for Rehabilitation

In proportional myoelectric control, the amplitude envelope of the sEMG signal is mapped directly to a joint torque or velocity command, and it remains the most commonly used approach in commercial myoelectric controllers and EMG-driven rehabilitation robots [7, 11]. Although this characteristic (i.e., continuous modulation of the command through continuous muscle activation) was its main defining feature, it also drove its major drawbacks in a rehabilitation context because it requires repeated demands on an already impaired neuromuscular system [3, 7].

Two consequences follow. Sustained submaximal contraction accelerates the onset of neuromuscular fatigue, expressed as a decline in force-generating capacity together with a compression of the sEMG power spectrum toward lower frequencies [12]. The early stage of stroke rehabilitation faces a major obstacle because patients have weak voluntary muscle contractions and their bodies cannot activate enough motor units [2, 13]. Calibration contraction is no longer useful because fatigue results in non-stationarity; therefore, repeated calibrations are needed within the same session [3, 6, 11]. The system becomes unstable because the electrodes change position with the individual's limb, and contact resistance between the electrodes and the skin changes [14].

Pattern recognition methodology stems from an altogether different perspective. The system acquires feature vectors from small sliding windows and then classifies the motion patterns into one of a few predefined categories. The early stages of contraction exhibit a deterministic pattern which enables classification into categories using few channels of data [9, 10]. The user is therefore required only to generate an activation burst distinguishable from the background signal over a brief interval—135 ms in the implementation reported by Chen et al. [6]—with analysis windows of roughly 150–250 ms generally identified as the best compromise between class separability and perceived controller delay [15]. Once a class is detected, the device executes a pre-programmed trajectory without further volitional modulation [2, 3]. Three lines of evidence support the adoption of pattern recognition specifically for rehabilitation, as distinct from prosthetic applications:

- It preserves voluntary drive, consistent with the AAN paradigm. Assist-as-needed control requires that robotic assistance be scaled to the patient's residual voluntary effort, so that

the device does not substitute for the neural drive that training is intended to elicit [7]. The rationale is empirical: actively generated movement produces greater cortical reorganization than passive movement matched for kinematics [8], and active practice is a precondition for the experience-dependent plasticity on which motor recovery depends. Research showed that EMG-driven robotic wrist training for patients with chronic stroke produced better results than continuous passive motion, which provided an equal amount of treatment [13]. The system requires activation bursts at the start of each movement because pattern recognition enforces this requirement directly rather than relying on automatic detection [3].

- The system establishes a motion path that follows biomechanical principles to achieve natural movement patterns. The clinical team sets predetermined movement routes for classified trajectories that maintain patient safety through correct anatomical positions during movement [2, 3]. This matters because hemiparetic reaching is characterized by stereotyped abnormal synergies: under continuous proportional control, the patient directly authors the trajectory and may therefore reinforce compensatory patterns, whereas a guided path constrains practice to the intended movement [7].
- It provides an explicit safety gate through reliable rest detection. Robust discrimination of a no-motion class is regarded as a prerequisite for clinically viable myoelectric control, since spurious actuation during intended rest is the failure mode most damaging to user trust [14]. Reported rest-detection accuracies approach 99% [6], yielding a discrete gate between intention and actuation. Continuous proportional control offers no equivalent gate, since any suprathreshold activity—including involuntary spasm or co-contraction—is translated into device motion [11].

Intent-based pattern recognition therefore offers a more natural and safer mode of human–device interaction for rehabilitation. However, much of this evidence derives from prosthetics research and healthy cohorts, and classification performance degrades in stroke populations: subject-specific classifiers applied to hemiparetic limbs have achieved approximately 71% accuracy in moderately impaired participants but below 40% in severely impaired participants [16]. The present study accordingly validates the integrated system in healthy participants as a necessary precursor to evaluation in the target clinical group.

3. INTEGRATED SYSTEM DESCRIPTION

3.1. Mechanical Subsystem

The exoskeleton offers five actuated revolute joints, which include Joint 0 (proximal/sternoclavicular approximation; range of motion $\pm 5^\circ$), Joint 1 through Joint 3 (glenohumeral; range of motion 0 – 150° , $\pm 90^\circ$, and -50° to $+150^\circ$ respectively), and Joint 4 (elbow flexion; range of motion 0 – 85°). The glenohumeral joints Joint 1, Joint 2, and Joint 3 have a common intersection point of the axes at the anatomical center of rotation, which allows the kinematic equivalence to a spherical joint of the glenohumeral joint. Importantly, all motors and links are located away from the point of intersection of the axes of the joints, thus shifting the mechanical mass laterally and inferiorly from the acromion point of view instead of being right on top of the shoulder joint. This is the main approach used to avoid the problem of head-and-torso collisions encountered by 3R mechanisms discussed in Section 2.1. The proximal joint J0 provides the sternoclavicular mobility that expands the functional workspace for cross-body and forward-reach motions and accommodates the scapulothoracic coupling observed during arm elevation. Fig. 4 shows the proposed shoulder mechanism. Three revolute joints

(J1, J2, J3), whose axes intersect at the glenohumeral COR, provide 3-DOF spherical motion. A fourth proximal joint (J0) approximates sternoclavicular/scapular mobility. The elbow joint (J4) completes the 5-joint chain. Motor and linkage offset from the COR reduce mechanical bulk over the acromion.

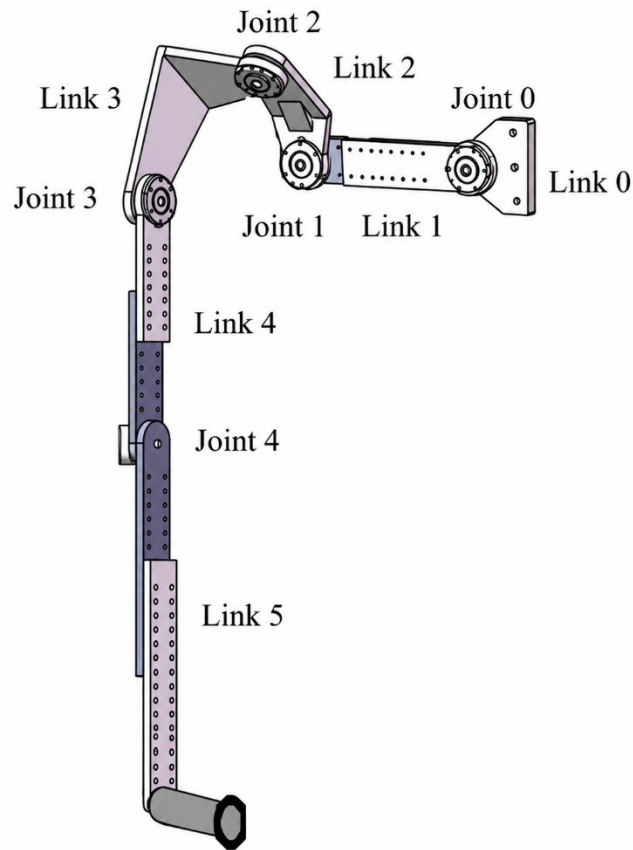


Figure 4. Proposed 4-DOF shoulder mechanism.

Each joint is driven by a Kollmorgen KH4248-B95020 closed-loop stepper motor (200 steps/rev, integrated encoder 1,000 counts/rev) coupled to a 1:100 planetary gearbox, yielding a peak joint torque of approximately 180 N·m and a maximum effective joint velocity of $\leq 60^\circ/\text{s}$. Closed-loop stepper motors were selected over conventional BLDC servo motors because they integrate the motor, encoder, and driver electronics into a compact unit and provide reliable position hold without active current at rest — eliminating the drift and low-frequency vibration that would degrade sEMG signals under continuous servo torque. A gear reduction of 1:100 was chosen to ensure the speed does not exceed safe limits for rehabilitation (shoulder: up to $60^\circ/\text{s}$; elbow: up to $30^\circ/\text{s}$).

Fig. 5 shows the physical prototype of the 4-DOF upper limb exoskeleton. Twelve Delsys Trigno Avanti wireless sEMG sensors (not shown) are placed at the shoulder, arm, and forearm positions. All structural components (links, brackets, motor mounts, and cuffs) were fabricated via fused deposition modeling (FDM) 3D printing using engineering-grade PETG polymer, with reinforced infill geometry in load-bearing zones. Telescoping aluminum segments in the upper arm and forearm provide anthropometric adjustability across the 5th–95th adult male percentile range. Complete device mass is 3.8 kg. The device is mounted on a height-adjustable aluminum stand for seated use; participants do not bear device weight on their own skeletal

structure. Joint hard stops and software torque limits at 80% of stall torque implement dual independent safety layers.



Figure 5. Physical prototype of the 4-DOF upper limb exoskeleton.

3.2. sEMG Acquisition and Processing Subsystem

Twelve Delsys Trigno Avanti wireless sensors sample at 1,000 Hz per channel and stream data via TCP/IP over the Trigno SDK. Sensor placement follows SENIAM guidelines and covers the primary movers for all four ADL classes: middle and anterior deltoid, posterior deltoid, supraspinatus, pectoralis major, upper trapezius, infraspinatus, teres major, biceps brachii, triceps brachii, and flexor and extensor carpi radialis. Per-session baseline normalization (z-score using a 10-second rest recording prior to each block) compensates for inter-session skin impedance variation. Fig. 6 illustrates the data flow from the Trigno system to the LabVIEW processing pipeline. The Trigno SDK streams 12-channel sEMG data at 1,000 Hz on port 50043. A 500 ms circular buffer absorbs wireless packet jitter up to 18 ms peaks.

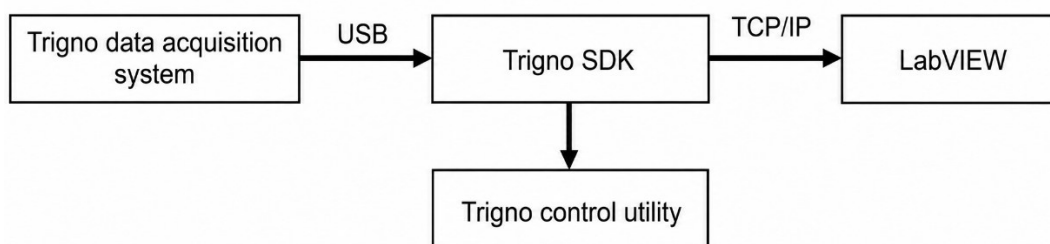


Figure 6. Data flow from the Delsys Trigno Avanti wireless sensor system to the LabVIEW acquisition loop via TCP/IP.

The LabVIEW processing pipeline runs in four parallel dataflow loops at tiered priorities, enabling concurrent acquisition, feature extraction, classification, and UI refresh without blocking. A 500 ms circular buffer absorbs Trigno wireless packet jitter of up to 18 ms peak while maintaining a continuous sample stream to the feature extraction loop. Feature extraction updates every 62.5 ms by computing six features per channel per 125 ms overlapping window: mean absolute value (MAV), root mean square (RMS), slope sign changes (SSC), Willison

amplitude (WAMP), mean frequency (MNF), and median frequency (MDF), yielding a 72-dimensional feature vector. MAV and RMS characterize signal amplitude and correlate with muscle force; SSC and WAMP capture the frequency complexity of the signal burst; MNF and MDF characterize the spectral centroid and are sensitive to fatigue-induced spectral compression [12]. The time-domain subset follows the feature ensemble established by Hudgins et al. for multifunction myoelectric control [9], and the use of overlapping rather than disjoint windows with majority-vote post-processing follows the continuous-decision scheme of Englehart and Hudgins [10]. We deliberately chose the 125 ms window below the 150–250 ms range reported to minimize classification error [15], trading a small accuracy margin for reduced controller delay [15]. We selected the six features using offline optimization experiments, which showed better discriminative performance among the four ADL classes than models with fewer features. Predictions are stabilized by taking a majority vote of three consecutive predictions (within 187.5 ms). The predicted intent is sent to the Arduino through the USB serial connection (115,200 baud rate). Fig. 7 illustrates the entire process flow of online pattern recognition of sEMG signals. The twelve-channel sEMG signals are captured and processed with a window size of 125 ms and an interval of 62.5 ms to generate a 72-dimensional vector at each 62.5 ms interval.

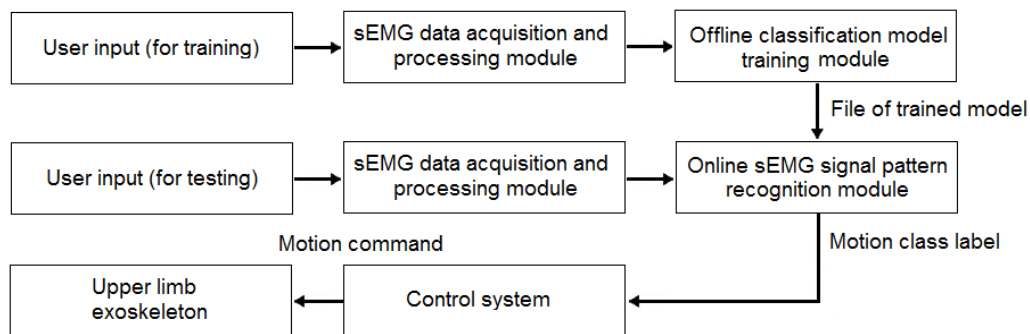


Figure 7. Online sEMG pattern recognition workflow implemented in LabVIEW.

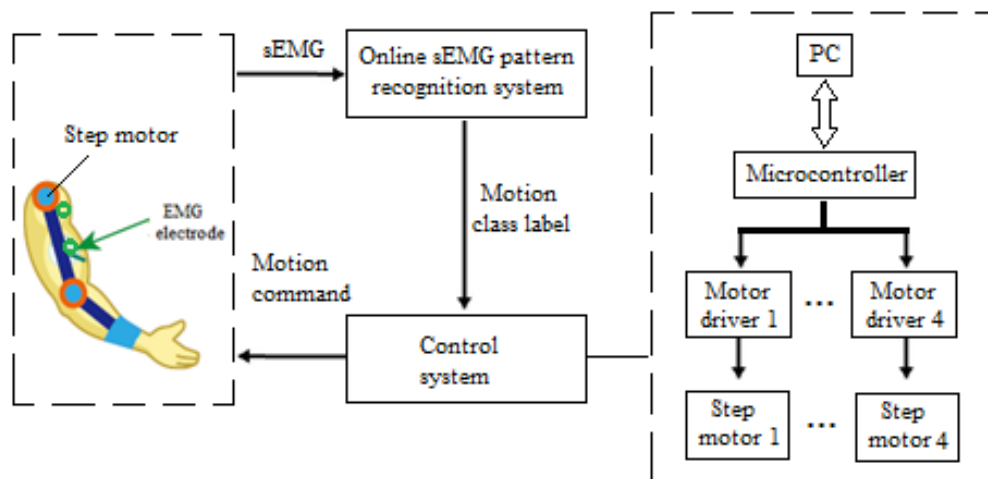


Figure 8. Integrated system architecture.

Figure 8 illustrates the system architecture. Three separate loops in LabVIEW run in parallel (acquisition loop, feature extraction/classification loop, and command generation loop) and exchange data through internal queues. Communication is enabled through external interfaces such as TCP/IP (Delsys Trigno SDK) and USB serial communication (Arduino Mega 2560).

3.3. Kinematic Modeling and Simulation

The 5-joint kinematic chain was modeled mathematically using the Denavit–Hartenberg (DH) convention, where coordinate frames were defined for each joint. Since J1, J2, and J3 have their centers of rotation (COR) coinciding with that of the glenohumeral joint, their coordinate frames have their origin at the same point and hence have link offsets of zero in the DH parameter table. The mathematical modeling of the 5-joint chain was done using MATLAB's Robotics Toolbox, where the Link and SerialLink objects were used with parameters derived from the SolidWorks CAD model. Joint range of motion limits were derived based on population-based anthropometric data, which are: J0 ($\pm 5^\circ$), J1 (0° - 150°), J2 ($\pm 90^\circ$), J3 (-50° - 150°), J4 (0° - 85°). Figure 9 shows the kinematic model and its workspace, respectively.

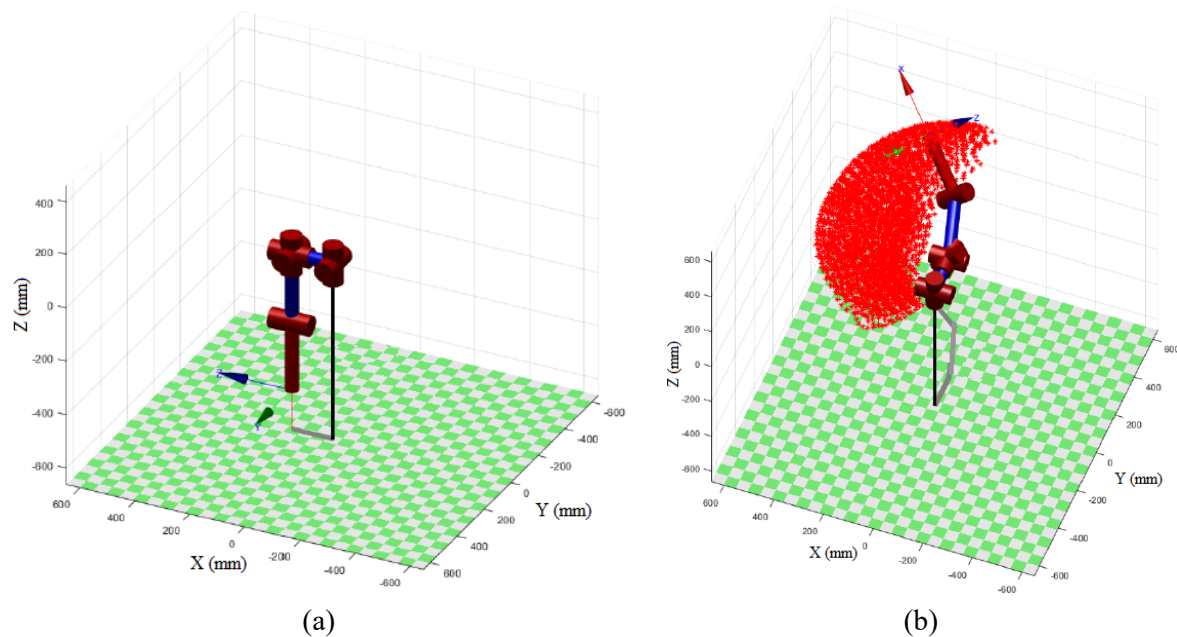


Figure 9. Kinematic modeling and workspace simulation of the five-joint upper-limb exoskeleton: (a) DH-based kinematic model implemented in MATLAB Robotics Toolbox and (b) simulated end-effector workspace.

3.4. Low-Level Actuation Subsystem

Four Summator MB 450A stepper motor drivers receive step and direction pulses from the Arduino Mega 2560. The Arduino converts each incoming ASCII intention command from LabVIEW into a preset joint trajectory programmed into the firmware, as a sequence of angular waypoints at $24^\circ/\text{s}$. Trajectories were designed collaboratively with a physiotherapist to match normative upper limb kinematics for each ADL: SAB guides J1 from 0° to 90° in the frontal plane with accompanying scapular elevation compensation through J0; PUC executes combined J1 shoulder flexion to 45° and J4 elbow flexion to 70° ; PFC executes J1 shoulder flexion to 60° while maintaining elbow near-extension; RA issues a STOP command and holds the current position with motor holding torque. Trapezoidal velocity profiles ($30^\circ/\text{s}^2$ ramp-up and ramp-down) prevent jerk at movement initiation and termination.

4. EXPERIMENTAL METHODS

The validation followed a five-step workflow (see Fig. 10) in which each subsystem was assessed both on its own and in combination. The mechanical subsystem was first characterized

in isolation through kinematic modeling and workspace simulation (Section 3.3); the sEMG intent-recognition pipeline was first evaluated as a standalone classifier on previously recorded offline data from the same volunteers; the two were then brought together in the integrated, physically coupled closed-loop setup described below; and finally the standalone and integrated conditions were cross-validated using the identical set of participants. Because the same individuals were tested offline, in free space, and under physical coupling, each participant serves as their own control, so any change in classification accuracy or signal quality can be attributed to the physical coupling itself rather than to between-subject variability. This within-subject cross-validation provides tighter control than validating the mechanism and the classifier on separate cohorts, and it underpins the signal-quality comparison reported in Section 5.2.

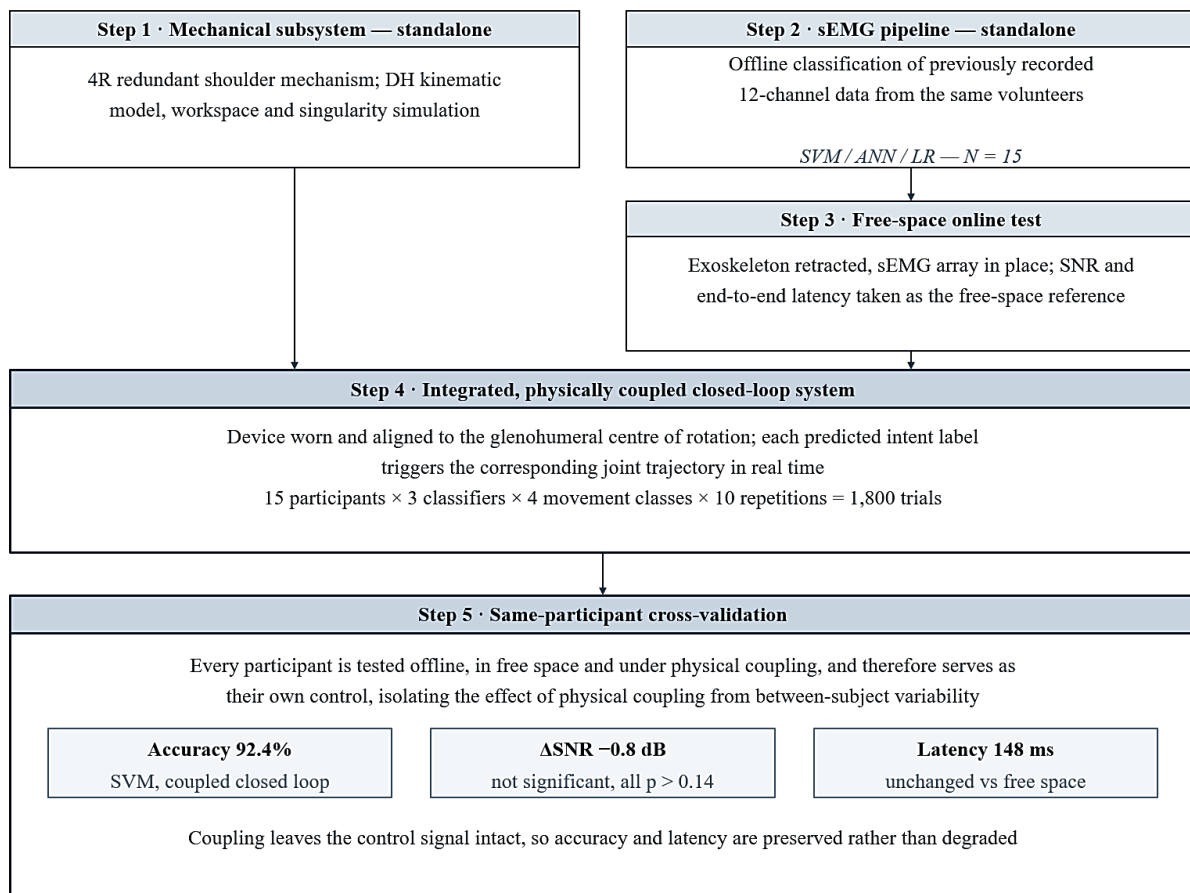


Figure 10. Staged validation workflow.

4.1. Participants

A total of fifteen healthy volunteers were included (10 males, 5 females; aged 23 ± 3.7 years; stature 1.71 ± 0.14 m; body weight 67.2 ± 4.8 kg). All had taken part in the earlier offline dataset study, so that the offline, free-space, and physically coupled conditions were assessed within the same individuals. Subjects with neuromuscular diseases, arm injuries within the past 12 months, and/or those with conditions influencing the voluntary control of movements and sEMG signals were excluded. The Institutional Research Ethics Committee reviewed and approved all experimental procedures involving human participants (approval no. MREC/AS/25/26/006), and the study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from every participant prior to enrolment.

This was a non-clinical pilot study conducted with healthy volunteers under the appropriate institutional clearance.

4.2. Movement Set and Clinical Rationale

Four movement classes were evaluated and chosen for their clinical importance in early post-stroke upper limb rehabilitation. Pilot testing also confirmed they were feasible within the device's range of motion. Shoulder abduction (SAB, 0–90° frontal plane) is an isolated proximal movement closely tied to regaining anti-gravity control after a stroke. “Pick up a cup” (PUC) – shoulder flexion and elbow flexion movements for reaching, grasping, and lifting a 50 g foam cup that is placed 30 cm anteriorly – is a basic task of reach and grasp activity of daily living used in a number of standardized tests (ARAT, WMFT). Pushing a cup forward (PFC) includes pushing the cup 30 cm anteriorly on the table using a flat palm of the hand. It uses similar proximal musculature as PUC but different distal musculature. Resting arm (RA) is a relaxed arm hanging down at the side. This is a crucial safety point; good recall of RA will protect against accidental device activation.

4.3. Experimental Protocol

Each participant underwent one session of integration, which lasted about 125 minutes, in four steps:

- Step 1 – Device fitting and familiarization (20 min). The alignment of the glenohumeral COR was checked by aligning the point of crossing of the J1- J3 axes with respect to a calibration marker placed over the acromion process of the subject. The link length of the device was adjusted such that it was equal to the length of the subject's upper arm and forearm segments, respectively. Each subject performed each action three times without the sEMG pipeline to become accustomed to the movement mechanism.
- Stage 2 — Free-space reference block (15 min). Before coupling the exoskeleton to the limb, participants performed 10 repetitions of each of the four movement classes with the twelve-channel sEMG array in place, the device retracted, and the arm moving unconstrained in free space. Acquisition, feature extraction, and classification ran with settings identical to the coupled condition, but no motor commands were issued. This block provided the free-space reference for the per-channel SNR comparison reported in Section 5.2 and the end-to-end latency comparison reported in Section 5.3.
- Stage 3 — Evaluation of classifiers (80 min). Evaluation of classifiers was done in three blocks (SVM, ANN, and LR, respectively) in counterbalanced order across subjects, with a rest of 10 minutes for every two successive blocks, totaling 80 min in Stage 3, i.e., $(3 \times 20 \text{ min}) + (2 \times 10 \text{ min}) = 80 \text{ min}$. Within each block, participants performed 10 repetitions of each of the four movement classes (40 trials per block, 120 trials per participant, 1,800 trials in total across the 15 participants) in a randomized block design cued by a visual display. This was a fully closed-loop system, meaning that every prediction of an intent label was associated with the predefined motor trajectory that the device followed. The repetitions were spaced with a 5-second break, and there were 10-minute breaks after each classifier block.
- Stage 4 — Usability and comfort questionnaires (10 min).

4.4. Outcome Measures

Intent Recognition Accuracy: Accuracy, Precision, Recall, and F1 Scores for each classifier were calculated across all subjects and runs. An intent was accurately recognized if

the majority vote label in the cue movement window was equal to the instructed class label. End-to-end latency was defined as the interval from EMG burst onset — detected using a double-threshold algorithm (RMS envelope >3 SD above baseline, sustained ≥ 20 ms) — to the Arduino serial command timestamp, measured using LabVIEW microsecond-resolution Tick Count primitives. Signal quality: Per-channel SNR was computed as the ratio of mean RMS amplitude during active movement windows to mean RMS during RA windows (dB scale), compared between physically coupled and free-space conditions (the free-space reference being the Stage 2 block described in Section 4.3) using 12 paired t-tests with Bonferroni correction (adjusted $\alpha = 0.004$). Usability: System Usability Scale (SUS, 10 items, 0–100; scores of approximately 68 and above correspond to a "Good" rating on the Bangor et al. adjective-rating scale [17]). Discomfort: Numeric Rating Scale (NRS, 0–10) at three attachment sites. Compensatory movement: A physiotherapist observer recorded trunk lean or shoulder shrug (blinded to classifier block order).

4.5. Statistical Analysis

Classifier effects on accuracy and latency were evaluated using one-way repeated-measures ANOVA with Greenhouse–Geisser correction for sphericity violations and Bonferroni post-hoc pairwise comparisons. Free-space-to-coupled accuracy degradation was tested with paired t-tests. Effect sizes were reported as partial eta-squared (η^2_p) for ANOVA and Cohen's d for pairwise tests. The significance threshold was $\alpha = 0.05$. All analyses were performed in MATLAB R2023b.

5. RESULTS

5.1. Intent-Recognition Accuracy

Table 1 situates the present system within a representative set of recently reported integrated sEMG-controlled upper limb exoskeletons. The comparison is restricted to recent systems that pair a physically worn mechanism with an sEMG-driven control model, because classifier accuracies reported on disembodied benchmark recordings are not directly comparable to values obtained during active, physically coupled device use. The set is nevertheless heterogeneous: not every system actuates the shoulder, and not everyone classifies movement intent online, as noted below. Across this set, the present system is the only one to report an end-to-end online latency measured under full physical coupling (148 ms), and the only one to quantify the effect of that coupling on sEMG signal quality directly. Its online accuracy of 92.4% is obtained with the mechanism worn and moving.

A few caveats apply in making this comparison and are highlighted here to avoid overstating the outcome. First, the present study classifies four movement classes, fewer than the five [20] and six [18] used elsewhere. Because class separability generally decreases as classes are added, part of the accuracy difference reflects the smaller decision space rather than the system itself. Similarly, the nominal number of DOF in the shoulder joints in this study differs from those of the other approaches, as this study uses a redundant 4-DOF configuration that utilizes the redundancy to perform singularity and collision avoidance for the shoulder (see Section 2.2), while the number of DOFs in the other approaches describes the kinematic configuration of the system. Zhao et al. [20] report higher accuracy than the present system, but on offline recordings and over activity and load states rather than discrete ADL movement classes. Regarding response time, the present study is the only system in Table 1 to report end-to-end latency (148 ms, below the 200 ms threshold); the remaining systems did not report latency (listed as N/R in Table 1), so a direct latency comparison is not possible. Taken

together, the present system is the only entry in Table 1 that combines a redundant 4-DOF shoulder mechanism, full physical coupling during evaluation, a reported sub-150 ms end-to-end latency, and the highest online accuracy among the systems evaluated under physical coupling.

Table 1. Representative integrated sEMG-controlled ULE systems compared with the present study

System	DOFs (shoulder)	Classifier	Classes	Reported acc. %	Latency (ms)	Physical coupling
Guo et al. [18]	N/R	SVM	6	79	N/R	Yes — robotic arm
Trigili et al. [19]	4	GMM	N/R	N/R	N/R	Yes — full
Zhao et al. [20]	N/R	CNN-BiLSTM-attention	5	97.3 / 88.2	N/R	Yes — full (offline)
Present study	4	SVM/ANN/LR	4	92.4	148	Yes — full, physical

SVM: support vector machine; ANN: artificial neural network; CNN: convolutional neural network; LSTM: long short-term memory network. Accuracies are as reported by the original authors. Zhao et al. [20] evaluated offline; their two values are the within-subject test-set and leave-one-subject-out cross-validation accuracies, respectively.

Table 2 reports overall and per-class classifier accuracy under physically coupled closed-loop conditions. One-way repeated measures ANOVA indicated a significant effect of classifier on accuracy ($F(2,28) = 19.3$, $p < 0.001$, $\eta^2p = 0.58$). Post hoc tests: SVM > ANN (mean difference 2.7%, $p = 0.003$, $d = 0.54$); SVM > LR (7.3%, $p < 0.001$, $d = 1.14$); ANN > LR (4.6%, $p < 0.001$, $d = 0.71$). SVM accuracy exceeded 80% for all 15 participants.

Table 2. Closed-loop intent-recognition performance (mean $\pm$ SD, N = 15). Best values per metric in bold

Metric	SVM (linear)	ANN (MLP)	LR
Overall accuracy (%)	92.4 $\pm$ 4.1	89.7 $\pm$ 5.8	85.1 $\pm$ 6.9
SAB recall (%)	96.0 $\pm$ 3.8	95.0 $\pm$ 4.1	86.0 $\pm$ 6.2
PUC recall (%)	86.0 $\pm$ 6.7	78.0 $\pm$ 8.4	70.0 $\pm$ 9.3
PFC recall (%)	82.0 $\pm$ 7.4	76.0 $\pm$ 8.1	72.0 $\pm$ 8.8
RA recall (%)	98.0 $\pm$ 2.1	97.0 $\pm$ 2.6	96.0 $\pm$ 3.0
Mean precision (%)	88.1 $\pm$ 5.2	85.6 $\pm$ 6.0	80.6 $\pm$ 7.5
Mean F1 score (%)	90.3 $\pm$ 4.6	87.4 $\pm$ 5.9	82.8 $\pm$ 7.1
End-to-end latency (ms)	148 $\pm$ 19	158 $\pm$ 23	155 $\pm$ 22
SUS score (0–100)	71.3 $\pm$ 8.4	68.9 $\pm$ 9.1	66.2 $\pm$ 10.7

SVM: support vector machine (linear kernel, one-vs-one); ANN: multilayer perceptron (20 neurons, ReLU, softmax); LR: logistic regression (one-vs-rest).

Classifier performance in terms of precision and recall is shown in Fig. 11. The main issue across all classifiers is confusion between PUC and PFC tasks (for SVM, 9% PUC classified as PFC and 10% PFC classified as PUC), which relates to the common forward-reach part of the task in which both muscles are engaged. All three classifiers perform best for the RA and SAB classes.

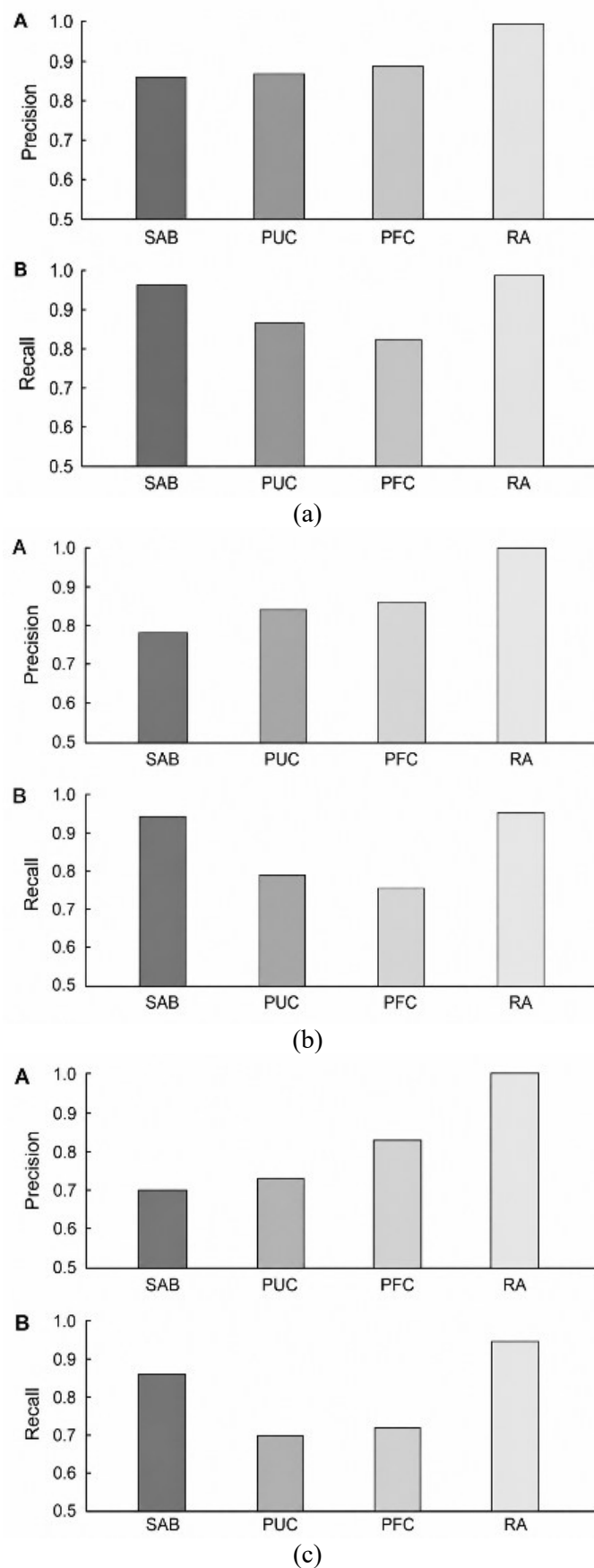


Figure 11. Class-wise precision and recall of the three classifiers: (a) support vector machine (SVM), (b) artificial neural network (ANN), and (c) logistic regression (LR).

5.2. Effect of Physical Coupling on sEMG Signal Quality

Table 3 presents per-channel SNR under physically coupled (device worn) and free-space (device absent) conditions. The mean SNR decrease of -0.8 dB under physical coupling was not statistically significant after Bonferroni correction (corrected threshold $p < 0.004$; all $p > 0.14$). This confirms that the 4-DOF mechanism's anatomical alignment prevents the vibration-induced and shear-induced electrode displacement that would substantially degrade sEMG quality in a mechanically misaligned device.

Table 3. Per-channel SNR (dB, mean $\pm$ SD, N = 15) under free-space and physically coupled conditions

Ch.	Muscle	Free-space SNR (dB)	Coupled SNR (dB)	Δ SNR (dB)	p-value
1	Middle deltoid	18.4 ± 3.2	17.6 ± 3.8	-0.8	0.19
2	Anterior deltoid	17.1 ± 2.9	16.2 ± 3.4	-0.9	0.14
3	Posterior deltoid	16.8 ± 3.1	15.9 ± 3.6	-0.9	0.17
4	Supraspinatus	14.2 ± 3.7	13.4 ± 4.1	-0.8	0.22
5	Pectoralis major	19.3 ± 2.8	18.7 ± 3.1	-0.6	0.31
6	Trapezius (upper)	15.7 ± 3.4	14.9 ± 3.9	-0.8	0.21
7	Infraspinatus	13.9 ± 4.0	13.1 ± 4.4	-0.8	0.24
8	Teres major	14.6 ± 3.5	13.8 ± 3.9	-0.8	0.20
9	Biceps brachii	20.1 ± 2.6	19.4 ± 2.9	-0.7	0.26
10	Triceps brachii	17.8 ± 3.0	17.0 ± 3.4	-0.8	0.18
11	Wrist flexor (FCR)	16.5 ± 3.2	15.8 ± 3.6	-0.7	0.27
12	Wrist extensor (ECR)	15.9 ± 3.3	15.2 ± 3.7	-0.7	0.25
Mean	All channels	16.7 ± 3.2	15.9 ± 3.6	-0.8	0.21

Δ SNR = coupled minus free-space. No comparison reached significance after Bonferroni correction (corrected $p < 0.004$; all $p > 0.14$).

5.3. End-to-End Latency

SVM latency under physical coupling was 148 ± 19 ms — identical within measurement uncertainty to free-space conditions. ANN (158 ± 23 ms) and LR (155 ± 22 ms) were similarly unchanged. A paired t-test confirmed no significant difference between coupled and free-space latency for SVM ($t(14) = 0.41$, $p = 0.69$). All three classifiers remained below the 200 ms therapeutic latency threshold in every session and every participant [15]. The dominant latency contributor was the windowing stage (~ 60 ms, accounting for $>40\%$ of total latency), confirming that reducing window length — rather than faster classifiers — is the primary lever for further latency reduction.

5.4. Usability, Comfort, and Compensatory Movement

SUS mean scores when using the SVM control were 71.3 ± 8.4 , classified as "Good" on the adjective rating scale [17]. The mean scores for ANN and LR were 68.9 ± 9.1 and 66.2 ± 10.7 , respectively. The effect of the classifier on the SUS score was statistically significant ($F(2,28) = 4.87$, $p = 0.015$, $\eta^2 p = 0.26$). The SVM classifier achieved a significantly higher SUS score than LR ($p = 0.008$, $d = 0.55$), as participants indicated that it provided faster and more reliable classification of their intent, minimizing uncertainty about the device's recognition success. The mean discomfort NRS score was 1.8 ± 0.7 . No participant got an NRS score of 4. Three participants reported a mild, temporary tingling sensation within the first five minutes of use. The compensation rate during the SAB task was 4.1% (in some cases above 80° of abduction) and 2.8% during the PUC task, resulting in a mean rate of 3.5%.

6. DISCUSSION

6.1. Integration Findings: Mechanical–Signal Interaction

The primary integration result is that physical coupling had little effect on sEMG signal quality: the mean Δ SNR was -0.8 dB for all twelve channels (Table 3). Hence, mechanical coupling introduced no noticeable electrical noise [4]. Importantly, the preservation of the quality of the signal is the basis by which the results of the mechanical and classification experiments can be linked to each other: because the coupling has not led to SNR degradation, the classifier has processed the signals in much the same way as it does in free space, which is the reason why the accuracy of the closed loop remained at the level of 92.4%, and the latency of the loop was 148 ms. Thus, the results demonstrate the effectiveness of the anatomically matched shoulder design in reducing the electrode displacement that conventional upper-limb exoskeletons experience [4, 5, 14]. Conventional designs confine shoulder movement to a fixed mechanical pivot and do not allow the glenohumeral joint COR to shift [5]. This kinematic mismatch generates interaction forces that lead to shear force, increased skin pressure, electrode displacement, and ultimately device rejection or joint injury [4, 5].

The lack of signal degradation shows that matching the mechanism with the COR eliminates the displacement mechanism and preserves signal quality [4]. This supports perceptual transparency: mechanical coupling has not destroyed the physiological signal on which intent recognition relies [4]. According to the results published by Pan et al. [4], self-aligning designs with passive DOFs allow keeping the glenohumeral rotation deviations within approximately 20 mm and having the same muscle activation patterns as in the case of no load, while fixed-axis designs cause the trajectory deviation up to 50 mm and greatly increase peak activation of the deltoid and upper trapezius muscles [8].

The practical implication of integrating the two subsystems in the experiment is that each subsystem's results would be meaningless on its own. Classifier accuracy measured in free space provides no guarantee that it persists when electrodes are loaded and the device vibrates, while the kinematic compatibility test results provide no guarantee that the remaining misalignment is small enough to allow use of the control signal.

6.2. System Responsiveness and Clinical Deployment

The mean latency of 148 ms matters for safe and intuitive physical human–robot interaction, since assistive responses must fall within biologically natural timing [3, 15]. This latency remains significantly below the accepted 200 ms threshold for implementing natural assistive responses in real-time control [3]. The stability of latency in free-space and coupled conditions (Section 5.3) shows that the device's mechanical structure does not introduce vibration noise or computational delay in the real-time loop [4, 6].

This responsiveness follows from using computationally inexpensive time-domain features with a classical classifier: linear SVMs have been reported to reach online accuracies of approximately 90% for upper limb movement classification, comparable to more complex models [3, 6, 14]. Deep learning architectures achieve high classification accuracy but impose substantial computational demands and end-to-end latencies of 200–500 ms [3]. Avoiding architectures that require GPU acceleration therefore allows stable operation on standard laboratory hardware, which matters for eventual clinical deployment [3, 6].

6.3. Implications for Intent-Driven Rehabilitation

The findings show that a 4-DOF mechanism can be combined with real-time intent recognition for functional ADL tasks [3, 6]. Preserving sEMG integrity under mechanical loading allows the user to remain an active participant in the movement, which is a precondition for regaining motor control [7, 8]. These findings align with contemporary neurorehabilitation theory, which states that voluntary patient involvement helps facilitate neuroplasticity [13]. Because this validation was limited to a single seated laboratory session, robustness across different environments and repeated daily sessions has not yet been established; domain-adaptation methods are therefore a promising direction for future work to address inter-session and inter-environment variability [14].

6.4. Biomechanical Interpretation of the PUC–PFC Confusion

The 9–10% mutual misclassification between picking up a cup (PUC) and pushing a cup forward (PFC) under SVM is consistent with the biomechanical literature on upper limb task initiation [6]. PUC and PFC both involve an initial reaching phase in an identical arc of shoulder flexion, powered mainly by the anterior deltoid and pectoralis major muscles [2, 3]. During this initial phase, the two tasks are kinematically indistinguishable, producing substantial overlap in activation patterns that no classifier can resolve from proximal channels alone [3, 6, 14].

The discriminating features are distal: flexor carpi radialis activation in preparation for grasping in PUC, and extensor carpi radialis activation for flat-palm pushing in PFC [6]. Because these muscles activate late in both movements, they fall outside the 125–135 ms analysis window anchored at movement onset [6, 9]. A short window is necessary to keep latency below the 200 ms therapeutic thresholds, but it cannot capture the late activation needed to separate functionally similar reach-to-grasp and reach-to-push intentions [3, 6, 15].

Solving this issue will require more sensing modalities that provide contextual information [3]. IMUs or force sensors can monitor joint coordination and interaction forces before distal muscle activation [3, 6]. Moreover, a hierarchy-based classification process may place more importance on distal cues as motion progresses, leading to better intention estimation [3]. This is one way perceptual transparency can be achieved in such a system [4].

6.5. Pattern Recognition vs. Proportional Control: Evidence Analysis

Experimental findings provide empirical evidence to compare PR and proportional control through five dimensions. The first dimension is effort demands. Participants experienced no difficulty (NRS 1.8) during 20 minutes of continuous task performance, with compensatory movements occurring in just 3.5% of all trials, indicating low effort demands. Proportional control depends on calibrated muscle contraction for every 2- to 4-second trial period, requiring much more physical effort from the participant. The low effort demand and near absence of compensatory movement align with the clinical requirement for low-effort practice in early post-stroke rehabilitation [2, 3]. In contrast, proportional control relies on sustained, calibrated muscle contractions to change the torque level, whereas PR requires only a burst of volitionally initiated muscular contractions [6, 11]. This difference matters because physical fatigue is a major obstacle during stroke rehabilitation, often changing the characteristics of the bioelectrical signal and requiring continuous recalibration in proportional control systems [3], [12]. By shortening the required contraction, pattern recognition mitigates the fatigue effects that otherwise degrade task performance and contribute to device rejection [3].

The second dimension is trajectory quality. Each SAB repetition follows an identical physiotherapist-designed frontal plane arc at 24°/s irrespective of the amplitude modulation. Each repetition therefore follows a physiotherapist-designed path whose biomechanical correctness does not depend on the participant's instantaneous EMG amplitude. A stable, guided path provides the repeatability required for both training and quantitative assessment. According to the literature, such guidance protects against reinforcing spastic synergies or motor maladaptation that could occur when a patient with limited coordination guides the proportional system [2, 3].

Third, calibration stability. Session-based baseline normalization maintained accuracy for all fifteen participants across 20-minute blocks without recalibration. In contrast, the proportional control requires MVC calibration for each channel with recalibration when amplitude drift occurs. The PR framework's stability over 20 minutes without recalibration solves one of the biggest technical problems in wearable robotics [3]. Traditional proportional myoelectric control is extremely sensitive to non-stationarity, drift, and bioelectrical signal changes caused by fatigue, requiring repeated Maximum Voluntary Contraction (MVC) tests that take valuable clinical time [3, 6, 11]. Using session-based normalization and computationally efficient time-domain features, the PR systems maintain higher long-term accuracy, making them deployable in resource-limited clinical environments [6].

Fourth, safety. Recall of 98% for the resting arm (RA) class corresponds to unintended actuation in only 2% of rest periods, and constitutes the safety gate on which clinically viable myoelectric control depends [6]. According to research, reliable intention recognition requires robust rest detection, with benchmarks up to 99%, to avoid the unintended actuation inherent to continuous proportional controllers [6, 14]. This way, the exoskeleton stays inactive during involuntary muscle spasms or other unintended movements, improving interaction safety during rehabilitation.

Fifth, active engagement. Requiring a volitional burst before each movement satisfies the assist-as-needed principle, under which robotic assistance is provided in proportion to the user's own effort [3]. Contemporary neurorehabilitation theory states that active volitional engagement is the prerequisite for cortical reorganization and durable functional recovery [2, 8]. By positioning the patient as an active agent rather than a passive recipient of movement, pattern-recognition control engages the neuroplastic processes on which motor recovery depends [2, 3, 13].

6.6. Limitations and Future Work

Four key restrictions need to be recognized. First, the study subjects were healthy young adults. Patients who have suffered stroke, with spasticity, co-contraction, and abnormal muscle synergies show a very different set of sEMG characteristics, especially in more complex movements such as PUC and PFC, where distal selectivity is the bottleneck; subject-specific classifiers applied to hemiparetic limbs have reported approximately 71% accuracy in moderately impaired participants and below 40% in severely impaired participants [16]. Extending to post-stroke subjects is the main next step and will require an adaptive or transfer-learning classifier approach due to large inter-subject and inter-session variability in the population. Second, each classifier block lasted 20 minutes, and all blocks were completed on a single day; inter-day variability in electrode placement and classifier performance in multi-session protocols has not been characterized. Third, the stand-mounted system does not have subjects bear the weight of the device through their own skeleton; back-mounted or orthotic mounting will change the weight-bearing properties of the shoulder mechanism and the

electrode array. Fourth, the wrist was not actuated, so it cannot produce the coordinated wrist and hand motion required for many ADLs.

7. CONCLUSION

This paper presented a fully integrated, sEMG-controlled 4-DOF upper limb exoskeleton intended for post-stroke rehabilitation, and validated it experimentally in healthy participants as a proof of concept. The system combines a novel shoulder mechanism — three glenohumeral revolute joints with co-located axes at the anatomical COR and a proximal sternoclavicular-replicating joint — with a LabVIEW-based real-time sEMG pipeline classifying four functional ADL classes using a linear SVM on 72-dimensional feature vectors extracted every 62.5 ms.

In a closed-loop experiment with 15 healthy participants physically wearing and using the device, the system achieved 92.4% overall intent-recognition accuracy, with 96% recall for SAB and 98% recall for RA—the highest online accuracy among comparable systems in Table 1 evaluated under physical coupling. End-to-end latency was 148 ± 19 ms, below the 200 ms therapeutic threshold. Physical coupling produced no statistically significant sEMG degradation (mean Δ SNR -0.8 dB, all $p > 0.14$ after Bonferroni correction), demonstrating that anatomical alignment directly preserves signal quality. Participants rated usability as “Good” (SUS 71.3) and reported only mild discomfort (NRS 1.8). Compensatory movement was observed in 3.5% of trials, substantially below rates reported for misaligned designs.

In summary, anatomically aligned mechanical design and lightweight sEMG pattern recognition are complementary: the same joint alignment that reduces parasitic interaction forces also limits electrode displacement, so the classifier receives cleaner signals than a mechanically coupled but misaligned device would deliver. The next phase will involve chronic stroke patients, expanding the range of movements from four to seven classes or more, and incorporating instrumented cuffs.

ACKNOWLEDGEMENT

The authors thank the 15 participants for their time and commitment at Xi'an Peihua University for professional guidance in trajectory design and compensatory movement assessment.

REFERENCES

- [1] Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, Fisher M, Pandian J, Lindsay P (2022) World Stroke Organization (WSO): Global stroke fact sheet 2022. *Int. J. Stroke*, 17(1):18-29. <https://doi.org/10.1177/17474930211065917>
- [2] Hasan SK, Bhujel SB, Niemiec GS (2024) Emerging trends in human upper extremity rehabilitation robot. *Cognit. Robot.*, 4:174-190. <https://doi.org/10.1016/j.cogr.2024.09.001>
- [3] Karasheva M, Saudanbekova A, Utepbergen A, Akkulova S, Niyetkaliyev A, Ozhikenov K, Ozhiken A, Alimbayev C, Shylmyrza U, Aimukhanbetov Y (2025) Sensor-driven control strategies for post-stroke shoulder rehabilitation exoskeletons: A systematic review. *MethodsX*, 15:103648. <https://doi.org/10.1016/j.mex.2025.103648>
- [4] Pan J, Astarita D, Baldoni A, Dell'Agnello F, Crea S, Vitiello N, Trigili E (2023) A self-aligning upper-limb exoskeleton preserving natural shoulder movements: Kinematic compatibility analysis. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 31:4954-4964. <https://doi.org/10.1109/TNSRE.2023.3341219>

- [5] Pramod AS, Mohan S, Thondiyath A (2025) A modified human shoulder model considering the migration of the humeral head from the scapular glenoid fossa for the design of shoulder exoskeletons. *Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci.*, 239(5):1441-1456. <https://doi.org/10.1177/09544062241295631>
- [6] Chen B, Zhou Y, Chen C, Sayeed Z, Hu J, Qi J, Frush T, Goitz H, Hovorka J, Cheng M, Palacio C (2023) Volitional control of upper-limb exoskeleton empowered by EMG sensors and machine learning computing. *Array*, 17:100277. <https://doi.org/10.1016/j.array.2023.100277>
- [7] Marchal-Crespo L, Reinkensmeyer DJ (2009) Review of control strategies for robotic movement training after neurologic injury. *J. NeuroEngineering Rehabil.*, 6:20. <https://doi.org/10.1186/1743-0003-6-2>
- [8] Lotze M, Braun C, Birbaumer N, Anders S, Cohen LG (2003) Motor learning elicited by voluntary drive. *Brain*, 126(4):866-872. <https://doi.org/10.1093/brain/awg079>
- [9] Hudgins B, Parker P, Scott RN (1993) A new strategy for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.*, 40(1):82-94. <https://doi.org/10.1109/10.20477>
- [10] Englehart K, Hudgins B (2003) A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.*, 50(7):848-854. <https://doi.org/10.1109/TBME.2003.813539>
- [11] Fougner A, Stavdahl Ø, Kyberd PJ, Losier YG, Parker PA (2012) Control of upper limb prostheses: Terminology and proportional myoelectric control—A review. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 20(5):663-677. <https://doi.org/10.1109/TNSRE.2012.2196711>
- [12] Enoka RM, Duchateau J (2008) Muscle fatigue: What, why and how it influences muscle function. *J. Physiol.*, 586(1):11-23. <https://doi.org/10.1113/jphysiol.2007.139477>
- [13] Hu XL, Tong KY, Song R, Zheng XJ, Leung WWF (2009) A comparison between electromyography-driven robot and passive motion device on wrist rehabilitation for chronic stroke. *Neurorehabil. Neural Repair*, 23(8):837-846. <https://doi.org/10.1177/1545968309338191>
- [14] Scheme E, Englehart K (2011) Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. *J. Rehabil. Res. Dev.*, 48(6):643-659. <https://doi.org/10.1682/JRRD.2010.09.0177>
- [15] Smith LH, Hargrove LJ, Lock BA, Kuiken TA (2011) Determining the optimal window length for pattern recognition-based myoelectric control: Balancing the competing effects of classification error and controller delay. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 19(2):186-192. <https://doi.org/10.1109/TNSRE.2010.2100828>
- [16] Lee SW, Wilson KM, Lock BA, Kamper DG (2011) Subject-specific myoelectric pattern classification of functional hand movements for stroke survivors. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 19(5):558-566. <https://doi.org/10.1109/TNSRE.2010.2079334>
- [17] Bangor A, Kortum P, Miller J (2009) Determining what individual SUS scores mean: Adding an adjective rating scale. *J. Usability Stud.*, 4(3):114-123.
- [18] Guo B, Ma Y, Yang J, Wang Z, Zhang X (2020) Lw-CNN-based myoelectric signal recognition and real-time control of robotic arm for upper-limb rehabilitation. *Comput. Intell. Neurosci.*, 2020:8846021. <https://doi.org/10.1155/2020/884602>
- [19] Trigili E, Grazi L, Crea S, Accogli A, Carpaneto J, Micera S, Vitiello N, Panarese A (2019) Detection of movement onset using EMG signals for upper-limb exoskeletons in reaching tasks. *J. NeuroEngineering Rehabil.*, 16:45. <https://doi.org/10.1186/s12984-019-0512-1>
- [20] Zhao D, Ye X, Wang S, Zhang C, Sun S, Zhang X, Cheng R (2025) Upper limb human-exoskeleton system motion state classification based on sEMG: Application of CNN-BiLSTM-attention model. *Sci. Rep.*, 15:18969. <https://doi.org/10.1038/s41598-025-02864-5>