International Journal on Perceptive and Cognitive Computing (IJPCC) Vol 11, Issue 1 (2025)
https://doi.org/10.31436/ijpcc.v11i1.503

ECG Signal Classification Using Hybrid and Non-Hybrid Learning Technologies

Asma Salim Yahya*, Naktal Moaid Edan

Department of Software, College of Computer Science and Mathematics, University of Mosul, Iraq

*Corresponding author asma_alkhairi@uomosul.edu.iq

(Received: 17th September, 2024; Accepted: 29th December, 2024; Published on-line: 30th January, 2025)

Abstract— Most arrhythmias caused by cardiovascular disorders disrupt the electrical activity of the heart, resulting in changes in the morphology of electrocardiogram (ECG) recordings. By analyzing different ECG patterns and comparing machine learning and deep learning techniques, this research aims to accurately identify twenty-nine different cardiac problems and sinus rhythm. The database contains 48 heart rate recordings at a frequency of 360 Hz for about 25 minutes for five classes, namely "N", "S", "V", "F", and "Q". Support Vector Machine (SVM), k-nearest neighbor (k-nearest neighbor) classifier, and random forest (RF) classifier were among the machine learning (ML) techniques used. Experimental results revealed that the random forest classifier achieved the highest classification accuracy, reaching 96.08%, while the support vector machine (SVM) achieved the lowest accuracy, reaching 88.9%. The study included deep learning approaches, namely convolutional neural networks (CNNs), hybrid deep learning models (CNN-LSTM), and recurrent neural networks of the long short-term memory (LSTM) type. Through a comparative analysis of the results of machine learning and deep learning, the best accuracy was achieved by the hybrid deep learning model LSTM-CNN, which achieved 97.25% with a kernel size of 3. Using the Sigmoid and SoftMax activation functions, the model achieved an accuracy of 95.12% and 97.32%, respectively, with the Adam activation function achieving an accuracy of 98.75%, to achieve the highest accuracy of the proposed model and find a balance between accuracy and speed classification. The main objective of this research is to implement a heart rate classification system from adult electrocardiograms using multiple machine learning and deep learning network architectures.

Keywords— Deep Learning, Machine Learning, ECG signals, Classification.

I. Introduction

Heart and blood vessel disorders, including coronary heart disease (CHD), which is characterized by the constriction of blood arteries supplying the heart muscle, are included in the category of cardiovascular illnesses (CVD). Other examples are congenital heart disease, which describes anatomical defects evident from birth, and rheumatic heart disease, which is caused by rheumatic fever induced by streptococcal bacteria and leads to damage to the heart muscle and valves. Globally, cardiovascular disease (CVD) is the primary cause of death. An electrocardiogram, or ECG, is a diagnostic test that captures the electrical activity of the heart and identifies any irregularities. Small electrical impulses generated by the heart travel through spindles to activate the heart's muscles 66.10% [1]. The ECG shows these impulses as a tracing on paper, allowing medical professionals to understand them. Over 17.9 million deaths annually, or 31% of all deaths worldwide, are attributed to it, making it the leading cause of death. 85% of these deaths are linked to cerebrovascular accidents and myocardial infarctions. People with cardiovascular disease or those who have a high risk of getting it because of several risk factors, such as diabetes or hypertension, should be helped to identify the disease's causes and receive treatment with appropriate medication and counselling.

Since its invention, an electrocardiogram, or ECG, has been the main diagnostic tool for determining a variety of heart issues. Recent developments have increased the ECG's significance. It is a commonly used, non-invasive, and reasonably priced diagnostic tool.

Electrocardiography, or ECG, is a diagnostic tool that is widely used to analyse cardiac function. It captures changes in the electrical activity of the heart over time and yields vital physiological data. ECG signals are periodic because they are made up of a wave sequence that repeats throughout time. This pattern consists of a P wave, which is then followed by Q, R, and S waves (which constitute the QRS complex), and a T wave therapy with appropriate medication and counselling comes last. [2], as illustrated in Figure 1.

ECGs can now be interpreted by computers in addition to people, thanks to recent developments in computing and improved technology. ECGs are frequently very lengthy, requiring the doctor to read them beat by beat. This is a laborious and timeconsuming procedure that can be challenging for less experienced medical professionals. Replacing it with systems that can accurately classify ECG data and provide counselling and appropriate medications can help this process [4]. Early detection of cardiac arrhythmias is essential for the diagnosis of heart disease and the prompt administration of treatment to patients. Long ECG recordings are difficult for doctors to analyse quickly because the average human eye is not designed to recognize the morphological changes in the ECG signal. For this reason, computer-aided diagnostics are desperately needed. Given the wide range of medical applications where this problem may occur, automated ECG classification is extremely important [5]. The automatic processing of the ECG signal encounters a significant challenge due to the substantial heterogeneity in the morphological and temporal properties of ECG waveforms among different patients, as well as within the same individuals [6]. The primary disadvantage of these ML methods is that they rely on manually extracted features therefore he was numerous machine learning (ML) solutions available for analysing and classifying ECG data at present. The challenge at hand pertains to the potential inability to identify the most appropriate features that will yield optimal classification accuracy. One of the suggested remedies involves the implementation of deep learning architecture in which the initial layers consist of long-memory LSTM layers designed to identify temporal features inherent in the input signal. Furthermore, the ultimate determination of heartbeat classes is entrusted to fully connected FCN layers to determine the pulse class.

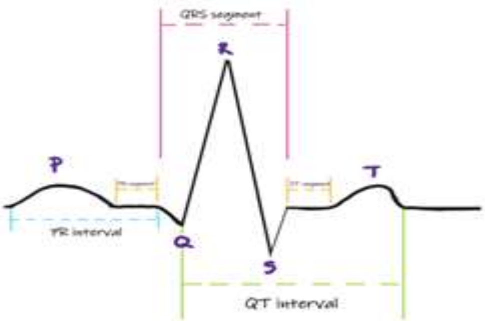

Fig. 1 Normal Electrocardiogram [3]

II. Research Review

Recent studies have used deep learning and machine learning approaches to analyze ECG data to detect cardiovascular diseases.

Abdullah et al was given [7] A thorough overview of IML approaches' applications in healthcare. Unfortunately, there is just one article that discusses using IML for heart disease categorization based on ECG signals. In a similar vein, Rasheed and colleagues [8] surveyed the research on interpreting ECG signals using IML and found only one study. Nonetheless, they justify their choices with a thorough analysis of IML approaches in their explanation of multi-fusion and multi-modal medical image segmentation, Yang et al. [9] demonstrated the advantages of machine learning (ML) interpretable approaches. Stiglic et al. [10] contrasted this with an emphasis on feature importance-based machine learning (ML) explanations.

A thorough overview of IML approaches' applications in healthcare was given by Abdullah et al. [11] and provided a comprehensive theoretical analysis of the popular IML methods now in use. Unfortunately, there is just one article that discusses using IML for heart disease categorization based on ECG signals. In a similar vein, Rasheed et al. [12] reviewed the literature on IML-based ECG signal interpretation and discovered just one report. However, they provide a comprehensive examination of IML techniques to support their decisions.

The DL-based Electrocardiogram arrhythmia classification pipeline relies heavily on the DL model design. When it comes to deep learning models, the architecture is multi-level or multi-layer, with each layer serving as an extractor of features that can improve its ability to summarize signal properties over time. The selected studies' DL classification models can be broadly classified into four categories: convolutional neural networks (CNNs), recurrent neural networks (RNNs), transformer, "hybrid" (combining different DL models), and "others" (representing less popular models like restricted Boltzmann machines and deep-belief networks), all based on the intrinsic properties of the major feature extractor within the neural networks. Here we present a comprehensive review of such DL algorithms for ECG arrhythmia categorization [13]. Classification of CNNs into 1D and 2D CNNs is based on the number of spatial domain filtering directions of the convolutional filters. To be more precise, 1D CNN filters travel in one way (the feature dimension) while 2D CNN filters go in two directions (the filtering axes) [14]. In their explanation of multi-fusion and multi-modal medical image segmentation, Yang et al. [15] demonstrated the advantages of machine learning (ML) interpretable approaches. Stiglic et al. [16] contrasted this with an emphasis on feature importance-based machine learning (ML) explanations. The CNN can work with any sampling rate for ECG signals. To categorize 2, 5, and 20 different kinds of cardiac illnesses, Stiglic et al. [10] use 1D CNNs and take few-shot learning into account due to the short dataset size. Contrarily, 2DCNN primarily considers inputs that resemble images, like the spectrogram in addition to the scalogram of an electrocardiogram (ECG). Electrocardiogram (ECG) signal classification makes use of a traditional 2D convolutional neural network (CNN), namely, AlexNet [17]. The 2D grey-level pictures with dimensions of 15 x 15 are simply converted from the 1D ECG plot in [18] and used as input in the two-dimensional CNN.

https://doi.org/10.31436/ijpcc.v11i1.503

RNN, or Recurrent Neural Network, is a sort of Deep Learning structure that takes into consideration the temporal relationship of feature sequences [19]. It obtains an advantage in extracting hidden statistical information of ECG features by employing the RNN, which makes it more responsive to the temporal properties of the input sequences [20]. In addition, long short-term memory (LSTM), an enhanced version of the recurrent neural network (RNN), has become more popular than the traditional RNN due to its superior capacity to evaluate time series data [13]. A 6-layer Long Short-Term Memory (LSTM) model is created in [21] to autonomously detect Premature Ventricular Contractions (PVC) using Electrocardiogram (ECG) sequences. Moreover, a bidirectional LSTM (BiLSTM) is a specific kind of LSTM that includes two LSTMsFor ECG classification, the BiLSTM model in [20] was used with retrieved ECG wave statistics in the temporal dimension. The amplitudes of the Q- and R-waves, the RR and QR intervals, and the ST segment starting point are among these statistics. Watson [22], a 2D BiLSTM is employed to detect atrial fibrillation (AF) by analyzing the spectrogram of electrocardiogram (ECG) signals. The input characteristics consist of the frequency components at each time instance. He suggests identifying atrial fibrillation (AF) by use of a Bidirectional Long Short-Term Memory (BiLSTM) model that receives an input of a sequence of RR intervals. To sum up, time-varying pulse statistics, time-frequency representation of the ECG, and raw ECG sequences can all be used as input sequences for RNNs. Numerous studies propose integrating several DL models into a single DL network for the categorization of ECG arrhythmias. For example, [23] combines CNN and RNN to create an encoder-decoder framework for heartbeat classification. CNNs are utilized for feature extraction, while RNNs are employed to convert the extracted features into their appropriate categories. More instances of combining CNN and LSTM with BiLSTM can be found in [24], in which CNNs are stacked in front of LSTM/BiLSTM components to extract features.

Recurrent Neural Networks, or RNNs, are a type of Deep Learning structure that considers feature sequence temporal relationships [25]. It gains an advantage in extracting hidden temporal information of ECG features by applying the RNN, which also makes it more responsive to the temporal properties of the input sequences. Furthermore, because of its greater ability to analyze time series data, long short-term memory (LSTM), an improved recurrent neural network (RNN), has gained popularity over the more conventional RNN [26]. In [27], a 6-layer Long Short-Term Memory (LSTM) model is developed to use Electrocardiogram (ECG) sequences to automatically identify Premature Ventricular Contractions (PVC). Furthermore, a particular sort of LSTM that consists of two LSTMs is called a bidirectional LSTM (BiLSTM).

III. Materials and Methods

A. Applications of Deep Learning in 3D Printing

The data used comes from the MIT-BIH repository (The Massachusetts Institute of Technology - Beth Israel Hospital). This database contains 48 recordings of heartbeats at a frequency rate of 360 Hz for approximately 30 minutes from 47 different individuals and annotated by a minimum of two cardiologists according to the five classes, namely 'N', 'S', 'V', 'F', and 'Q'.

The MIT-BIH cardiac arrhythmia Dataset contains samples that have been resized, down sampled, and padded at zeroes if needed to a fixed dimension of 188, as stated in the Kaggle dataset note. The classes, namely 'N', 'S', 'V', 'F', and 'Q', each represented by a unique numerical value: 0, 1, 2, 3, and 4, respectively. The description is as follows [28]: N - Regular cardiac contraction, S - Supraventricular preterm or as ectopic pulse (atrial or nodal) V- Ventricular premature contraction (VPC) F- The F represents the fusion of a ventricular beat with a regular beat. Q - Uncategorizable beat. Figure 2. shows the sampling of each class.

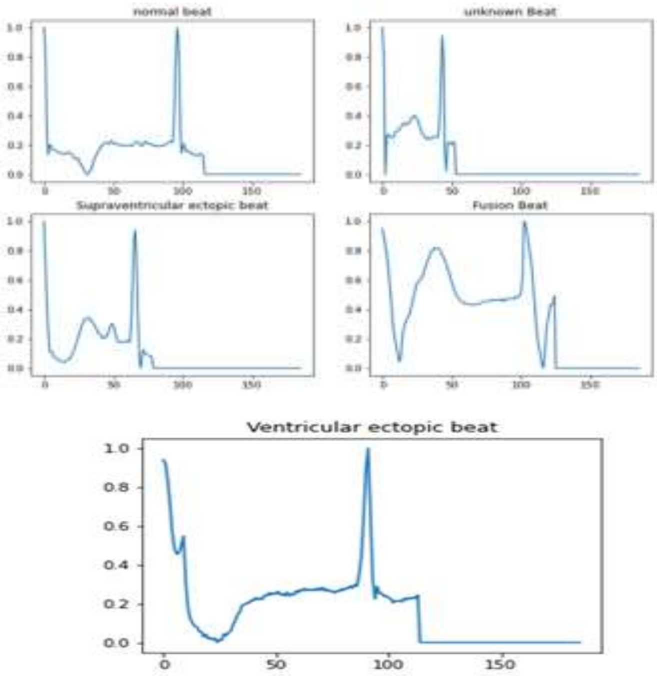

Fig. 2 Sampling of Each ECG Class [28]

B. Pre-Processing

Preprocessing of the data has been done to see if it needs to be cleaned. Clean data is required for model fitting in the following steps. A dataset's missing values can have an impact on how well a classifier built using that dataset as a training sample performs. Numerous approaches can be used to deal with missing data. In this work, we examine missing values using the (msno) matrix to display the dataset histogram and the is-null and not-null Pandas Data Frame functions. Then, we use the drop-Na (dropna) method to remove rows that contain null values. Unless the in-place option is set to True, the drop-Na method creates a new data frame object; in that case, the drop-Na function removes the previous data

https://doi.org/10.31436/ijpcc.v11i1.503

C. Data Augmentation

To effectively train the model, it is necessary to standardize the level of augmentation for all the inputs. Due to the presence of bias in our data, it is necessary to employ data augmentation techniques to mitigate bias achieve equitable data distributions and improve the accuracy of models in classification. We have created two augmentations along with wrapper classes to facilitate their usage. It is important to note that these augmentations are generated randomly, resulting in two layers of randomization in the development of fresh input.

D. Dataset Balancing

The dataset is rather unbalanced. Therefore, we must up-sample every class. Although class F (fusion of ventricular and normal beat b) may overfit, this will have the same effect on the model as other classes. Consequently, it has been resampled for each group to acquire 20,000 observations for each class.

E. Feature Extraction by Machine Learning

The supervised approach recognizes ECG signals based upon several extracted features, which are then compared to a baseline learnt model. This comparison helps to establish if the ECG signal to normal or arrhythmia. The supervised techniques label a new exemplar as normal if the example baseline learnt model. This comparison helps to establish if the ECG signal to normal or arrhythmia. The supervised techniques label a new exemplar as normal if the example lies within the range of normality; otherwise, the exemplar is marked as abnormal. By using a few of the widely used machine learning approaches for threat detection [29].
The machine-learning techniques utilized in the performance appraisal model through the supervised method of ECG signal detection are based on several extracted features. Algorithms are described as three machine-learning used in the performance appraisal model namely SVM, RF and KNN. The test results are shown in Table I.

TABLE I.
MACHINE LEARNING CLASSIFICATION RESULTS OF ECG SIGNALSFOR MIT-BIH DATASET.

		SVM	RF	KNN
Accuracy		91.27	96.08	93.03
Precision (%)	0	85.16	98.12	86.09
	1	70.01	72.23	77.12
	2	35.82	85.06	45.22
	3	85.06	93.47	81.62
	4	16.56	77.98	55.81
Recall (%)	0	90.23	98.07	90.33
	1	82.01	84.10	77.49
	2	90.99	89.23	92.07
	3	93.21	86.16	82.12
	4	94.08	90.45	92.39
F1-Score (%)	0	90.25	90.47	93.08
	1	56.86	71.03	64.64
	2	88.39	90.15	83.14
	3	37.12	56.03	66.22
	4	87.45	88.09	92.05

Vector Machine (SVM) classifier's highest recall accuracy of classification at 91.27% for 2 classes (Unknown Beat and Fusion Beat).

F. Feature Extraction by Deep Learning

In this part, we analyze the classification performance between the 3 DL models namely CNN, LSTM and CNN-LSTM. However, these models failed to gain good results (or even very bad results by the LSTM model), so modified and redesigned models to be more suitable to ECG classification and gain accuracy and higher results in classification this is done by hybridizing it with a convolutional neural network, and the models are also tested with the same input as in ML test and the test results are shown in Table II.

From Table II the higher accuracy in classification is achieved by the hybrid (CNN-LSTM) model while the second good results have been achieved by the CNN model, while the LSTM model achieved lower classification results. The hybrid (CNN-LSTM) model achieves the highest accuracy rather to other ML and DL models which achieved 97.25% accuracy. On the other hand, CNN achieved the second highest accuracy than other ML and DL models which achieved 96% accuracy LSTM is the worst result compared with others, to show results selected hybrid CNN-STM model is a preferred model for classifying ECG signals. .

G. Optimization Algorithms on the Classification Performance for the CNN-LSTM Model.

Training Parameters and Optimization Algorithms can affect classification accuracy. Hence, it has been studying several parameters and optimization algorithms that can be used in the CNN-LSTM model and find the optimum value that can be used.

TABLE II.
DEEP LEARNING CLASSIFICATION RESULTS OF ECG SIGNALS.

		CNN	LSTM	CNN-LSTM
Accuracy		96	79.05	97.25
Precision (%)	0	97.15	55.09	97
	1	75.08	71.26	70
	2	95.11	80.14	96
	3	78.88	70.65	81
	4	97.09	82.56	98.09
Recall (%)	0	96.05	70.66	97.01
	1	80.05	62.11	83
	2	94.34	35.00	97
	3	82.16	88.02	85
	4	98.36	75.17	97.09
F1-Score (%)	0	98	55.23	98.25
	1	82.48	67.92	79
	2	96.09	55.84	94
	3	70.05	81.15	92
	4	97	81.09	98.03

https://doi.org/10.31436/ijpcc.v11i1.503

1) *Kernel Size Effects*

This section examines the impact of kernel size variation on the efficacy of training. Four kernel sizes (1), (3), (5), and (7) have been evaluated. Twenty-five epochs are designated as the training epoch, with ten stages comprised of each epoch. The test for two datasets is described previously and the results are shown in Table III.

TABLE III
CLASSIFICATION RESULTS FOR DIFFERENT KERNEL SIZE AND ECG DATASET.

Kernel size	Validation Accuracy	Recall	Precision	F1-score
1	0.8314	0.8145	0.8345	0.8262
3	0.8549	0.9418	0.9355	0.9336
5	0.9449	0.9512	0.8889	0.8735
7	0.8310	0.8415	0.8402	0.8111

The training and validation results for the CNN model for the ECG database are shown in Figure 3.

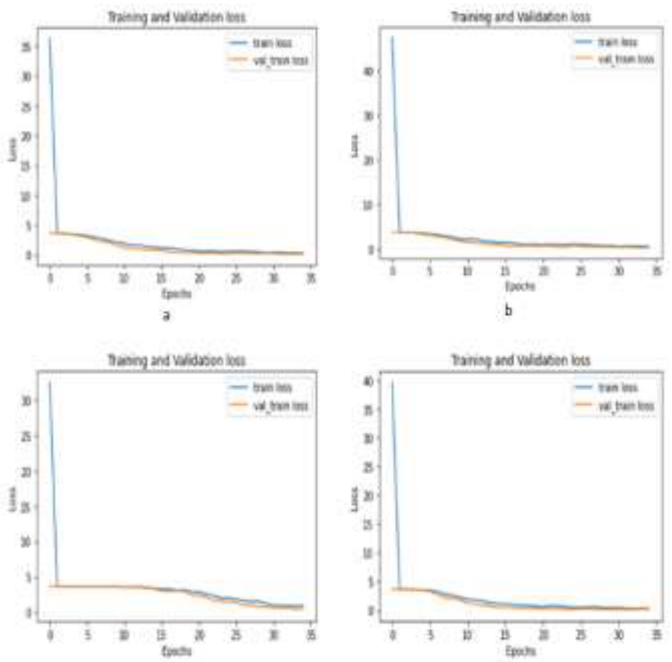

Fig. 3 Training and Test Loss for Various Kernel Sizes for ECG Database. (A)1 Kernel Size, (B) 3 Kernal Size, (C) 5 Kernel Size, (D) 7 Kernel Size

2) Activation Function Effects

In this part, we used the optimal kernel size of 5, we examined the impact of employing various activation functions for the output layer. For this test, four common 4 activation functions from keras have been employed, including SoftMax, sigmoid, Relu, tanh, and Softplus. Table IV and Figure 4 display the results of the tests.

TABLE IV.
THE RESULTS WITH SEVERAL ACTIVATION FUNCTIONS.

Activation Function Type	Test Accuracy	Recall	Precision	f1-score
softmax	0.9732	0.9780	0.9735	0.9756
sigmoid	0.9512	0.9418	0.9355	0.9336
relu	0.0319	0.9631	0.0328	0.0111
tanh	0.0274	0.9741	0.0267	0.0018
soft plus	0.0382	0.9371	0.0383	0.0034

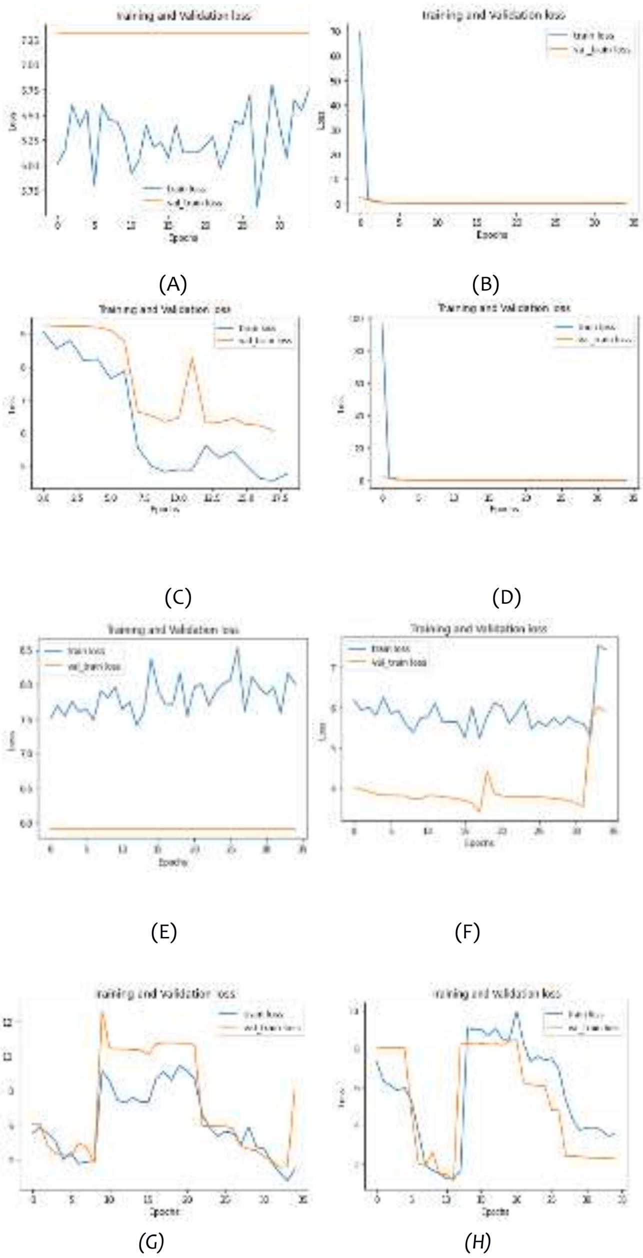

Fig. 4 ECG Classification Results with Various Activation Functions (A) Tanh, (B) Sigmoid, (C) Softmax, (D) Relu, (E) Softsign, (F) Softplus, (G) Elu, (H) Selu.

https://doi.org/10.31436/ijpcc.v11i1.503

Based on the data presented in Table IV the validation and training loss of the ECG database indicate that the model performs well when using the activation function of the sigmoid and the softmax activation function. The validation and training loss consistently reduce and become stable at certain times (at 5 epochs), indicating a good fit. Models with alternative activation functions experience issues with both overfitting and underfitting, as well as a lack of stability. The variation in loss of training is usually ascribed to the presence of disappearing or exploding gradients. It is clear that only sigmoid and SoftMax, two popular activation functions, work well and produce positive results. On the other hand, all other activation functions lead to overfitting and produce notably low accuracy. Even though both perform better in the provided dataset, it is clear that SoftMax performs more accurately than Sigmoid while training the electrocardiogram (ECG) database. It has been concluded that the sigmoid and SoftMax activation functions are appropriate choices for this model.

3) *Optimization Function Effects*

The most efficient activation function, SoftMax, has been utilized to assess how different optimization functions affect the classification accuracy of the suggested model. The seven activation functions that are tested are Adam, Adamax, Adagrad, Nadam, SGD, Ftrl, and RMSprop. TABLE V presents the test results. We set the epoch to 100 in this test so that the accuracy algorithm could compare and optimize them over a longer time frame.

TABLE V.
ECG CLASSIFICATION RESULTS WITH SEVERAL OPTIMIZATION FUNCTIONS.

Activation Function Type	Validation Accuracy	Recall	Precision	F1-score
Adamax	0.9870	0.99	0.9812	0.9813
Adam	0.9875	0.9905	0.9872	0.9885
Nadam	0.9845	0.9815	0.9835	0.9841
Adagrad	0.9715	0.8309	0.9731	0.9721
SGD	0.2105	0.8580	0.1205	0.1524

Based on the results in Table V. the majority of optimization functions perform well and obtain high accuracy for ECG datasets, except for Ftrl and SGD. Specifically, these two approaches produce low accuracy when used with the ECG dataset. Nevertheless, the values of Adam and Adamax are relatively high and they are near to each other. As a result, we decided that Adamax is the best approach for the proposed model.

IV. Experimental Setup and Performance Metrics

The test results were analysed using six widely used assessment techniques, or "performance measures," because they were combined to assess the overall performance of the model. These methods included f1 score, accuracy, memory, and precision. The confusion matrix provides a comprehensive view of your model's performance. The test's results allow for the evaluation of a classification model's performance and the identification of problems it encounters when applied to a binary classification task such as the one under investigation, it would have a 2 x 2 matrix of values, as shown below [30][31]: TP = Positive: Occurrence of positive positivity. TN = True Negative: Positive classes were predicted correctly. FP = incorrectly signed geolocation as came rows. FN = false negative: correctly predicts positive signs.

Accuracy= (TP+TN)/(TP+TN+FP+FN).

Precision= TP/(TP+TN).

Recall= TP/(TP+FN).

F1 Score= (2(precision*recall))/((precision recall))*

V. CONCLUSION

After conducting an inquiry on the impact of various training settings on the accuracy as well as the stability of the DL models, and can summarize the findings as follows: The stability and precision of the findings can be greatly impacted by choosing an appropriate kernel size. Our results suggest that for these kinds of networks, a kernel size of 5 is more appropriate. By contrasting the outcomes of deep learning and machine learning, Using Sigmoid and SoftMax activation functions, the hybrid deep learning model LSTM-CNN achieved the highest accuracy in ECG classification, reaching 98.10% with a kernel size of 5. Adam's activation function achieved 98.75%, the highest accuracy for the proposed model, and a balance between accuracy and speed was achieved. The choice of activation function greatly impacts the training and testing results, it has been selected Adam as the preferred method for training the electrocardiogram (ECG) database. The use of state-of-the-art fast computational ML algorithms, such as SVM, DT, and KNN and compared with high complexity DL algorithms in ECG classification applications to determine the most preferred method that can give a balance between performance and speed. Figure 1 shows that the training and test losses of the ECG database and the dataset that is provided for all kernel sizes have a good match, with both training and testing losses decreasing as they stabilize at a given time (about one epoch). Following that, it decreased steadily until the 35 designated epochs came to an end. As a result, it is evident that the training and testing curve fittings, which begin at 25 epochs and remain steady until the conclusion, were similar to one another when the model employs a kernel size of the outcome goal of this research, that intend to implement a heartbeat classification system from adult ECGs by using several machine learning and deep learning network architectures.

ACKNOWLEDGMENT

The authors hereby acknowledge the review support offered by the IJPCC reviewers who took their time to study the manuscript and find it acceptable for publishing.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest

REFERENCES

[1] T. Gaziano, K. S. Reddy, F. Paccaud, and S. Horton, "Cardiovascular Disease," *Dis. Control Priorities Dev. Ctries. (2nd Ed.*, no. Cvd, pp. 645–662, 2006, doi 10.1596/978-0-8213-6179-5/chpt-33.

[2] H. C. Kim, "Epidemiology of cardiovascular disease and its risk factors in Korea," *Glob. Heal. Med.*, vol. 3, no. 3, pp. 134–141, 2021, doi: 10.35772/ghm.2021.01008.

[3] Mørk, M. S. (2022). *Cilostazol a medical treatment for bradycardia of cardiac origin in dogs* (Doctoral dissertation).

[4] X. Yang et al., "The highly-cited Electrocardiogram-related articles in science citation index expanded: characteristics and hotspots," J. Electrocardiol., vol. 47, no. 5, pp. 738–744, 2014, doi: https://doi.org/10.1016/j.jelectrocard.2014.03.005.

[5] E. A. Ashley and J. Niebauer, *Conquering the ECG*. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK2214/

[6] K. Prashant, P. Choudhary, T. Agrawal, and E. Kaushik, "OWAE-Net: COVID-19 detection from ECG images using deep learning and optimized weighted average ensemble technique," *Intell. Syst. with Appl.*, vol. 16, no. November, p. 200154, 2022, doi: 10.1016/j.iswa.2022.200154.

[7] T. A. A. Abdullah, M. S. M. Zahid, and W. Ali, "A review of interpretable ml in healthcare: Taxonomy, applications, challenges, and future directions," *Symmetry (Basel).*, vol. 13, no. 12, pp. 1–28, 2021, doi: 10.3390/sym13122439.

[8] K. Rasheed, A. Qayyum, M. Ghaly, A. Al-Fuqaha, A. Razi, and J. Qadir, "Explainable, trustworthy, and ethical machine learning for healthcare: A survey," *Comput. Biol. Med.*, vol. 149, no. September, p. 106043, 2022, doi 10.1016/j.compbiomed.2022.106043.

[9] G. Yang, Q. Ye, and J. Xia, "Unbox the black box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond," *Inf. Fusion*, vol. 77, no. May 2021, pp. 29–52, 2022, doi: 10.1016/j.inffus.2021.07.016.

[10] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 5, pp. 1–13, 2020, doi: 10.1002/widm.1379.

[11] T. A. A. Abdullah, M. S. M. Zahid, and W. Ali, "A Review of Interpretable ML in Healthcare: Taxonomy, Applications, Challenges, and Future Directions," Symmetry 2021, Vol. 13, Page 2439, vol. 13, no. 12, p. 2439, Dec. 2021, doi: 10.3390/SYM13122439.

[12] K. Rasheed, A. Qayyum, M. Ghaly, A. Al-Fuqaha, A. Razi, and J. Qadir, "Explainable, trustworthy, and ethical machine learning for healthcare: A survey," Comput Biol Med, vol. 149, p. 106043, Oct. 2022, doi: 10.1016/J.COMPBIOMED.2022.106043.

[13] A. A. Ahmed, W. Ali, T. A. A. Abdullah, and S. J. Malebary, "Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model," *Mathematics*, vol. 11, no. 3, pp. 1–16, 2023, doi: 10.3390/math11030562.

[14] Á. T. Escottá, W. Beccaro, and M. A. Ramírez, "Evaluation of 1D and 2D Deep Convolutional Neural Networks for Driving Event Recognition," *Sensors*, vol. 22, no. 11, pp. 1–21, 2022, doi: 10.3390/s22114226.

[15] G. Yang, Q. Ye, and J. Xia, "Unbox the black box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond," Information Fusion, vol. 77, pp. 29–52, Jan. 2022, doi: 10.1016/J.INFFUS.2021.07.016.

[16] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," Wiley Interdiscip Rev Data Min Knowl Discov, vol. 10, no. 5, p. e1379, Sep. 2020, doi: 10.1002/WIDM.1379.

[17] A. Y. Hannun *et al.*, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nat. Med.*, vol. 25, no. 1, pp. 65–69, 2019, doi: 10.1038/s41591-018-0268-3.

[18] E. A. Ashley and J. Niebauer, "explained Euan A Ashley and Josef Niebauer explained".

[19] C. Antzelevitch and A. Burashnikov, "Overview of Basic Mechanisms of Cardiac Arrhythmia," *Card. Electrophysiol. Clin.*, vol. 3, no. 1, pp. 23–45, 2011, doi: 10.1016/j.ccep.2010.10.012.

[20] H. Liu *et al.*, "A large-scale multi-label 12-lead electrocardiogram database with standardized diagnostic statements," *Sci. Data*, vol. 9, no. 1, pp. 1–10, 2022, doi: 10.1038/s41597-022-01403-5.

[21] P. Wagner *et al.*, "PTB-XL, a large publicly available electrocardiography dataset," *Sci. Data*, vol. 7, no. 1, pp. 1–16, 2020, doi: 10.1038/s41597-020-0495-6.

[22] D. Watson, "Conceptual Challenges for Interpretable Machine Learning," *SSRN Electron. J.*, no. June, 2020, doi: 10.2139/ssrn.3668444.

[23] G. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, 2009, doi: 10.4249/scholarpedia.5947.

[24] P. Sharma, M. Kumar, H. K. Sharma, and S. M. Biju, *Generative adversarial networks (GANs): Introduction, Taxonomy, Variants, Limitations, and Applications*, no. 0123456789. Springer US, 2024. doi: 10.1007/s11042-024-18767-y.

[25] C. Antzelevitch and A. Burashnikov, "Overview of Basic Mechanisms of Cardiac Arrhythmia," Card Electrophysiol Clin, vol. 3, no. 1, p. 23, Mar. 2011, doi: 10.1016/J.CCEP.2010.10.012.

[26] B. B. Sahoo, R. Jha, A. Singh, and D. Kumar, "Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting," Acta Geophysica, vol. 67, no. 5, pp. 1471–1481, Oct. 2019, doi: 10.1007/S11600-019-00330-1/METRICS.

[27] P. Wagner et al., "PTB-XL, a large publicly available electrocardiography dataset," Scientific Data 2020 7:1, vol. 7, no. 1, pp. 1–15, May 2020, doi: 10.1038/s41597-020-0495-6.

[28] S. Mukhopadhyay, S. Biswas, A. B. Roy, and N. Dey, "Wavelet-Based QRS Complex Detection of ECG Signal," vol. 2, no. 3, pp. 2361–2365,

2012, [Online]. Available: http://arxiv.org/abs/1209.1563

[29] H. A. Marzog and H. J. Abd, "Machine Learning ECG Classification Using Wavelet Scattering of Feature Extraction," *Appl. Comput. Intell. Soft Comput.*, vol. 2022, 2022, doi: 10.1155/2022/9884076.

[30] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 01–11, 2015, doi: 10.5121/ijdkp.2015.5201.

[31] R. Mathur, C. P. Gupta, V. Katewa, D. S. Jat, and N. Yadav, Eds., "Emerging Trends in Data-Driven Computing and Communications," 2021, doi: 10.1007/978-981-16-3915-9.