https://doi.org/10.31436/iiumej.v27i3.4079

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 Nueral 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),

$$X_{new} = \frac{X}{X_{max}} \tag{1}$$

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:

$$PAR = \frac{PR-PL}{PR+PL} \tag{2}$$

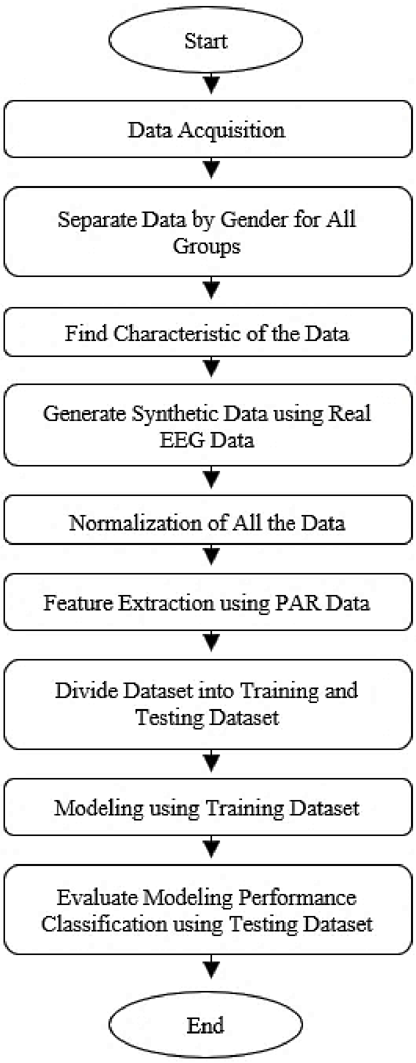


Figure 2. Process flow of the ANN model and classification for original and synthetic EEG datasets.

https://doi.org/10.31436/iiumej.v27i3.4079

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.

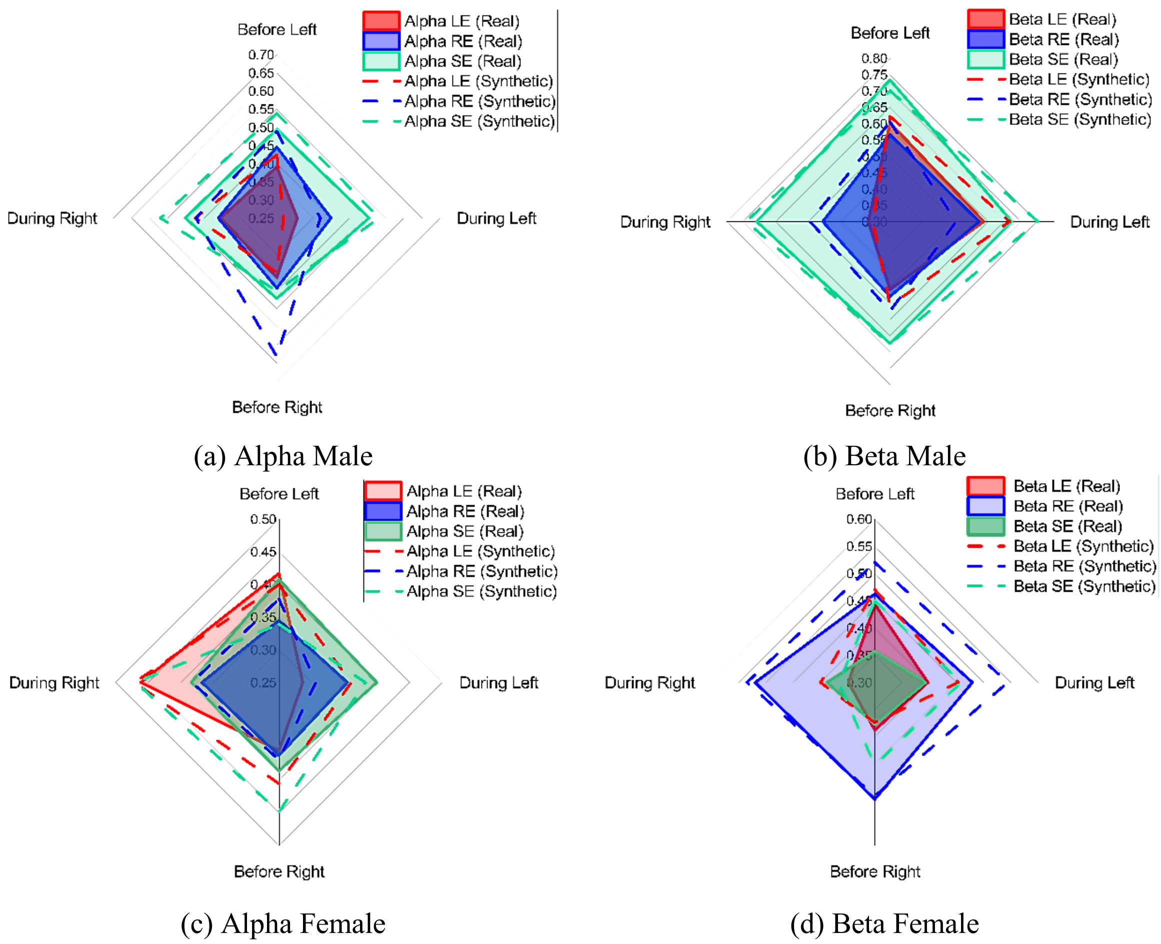


Figure 3. Normalized PSD for original and synthetic EEG data for (a) Alpha Male, (b) Beta Male, (c) Alpha Female, and (d) Beta Female.

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.

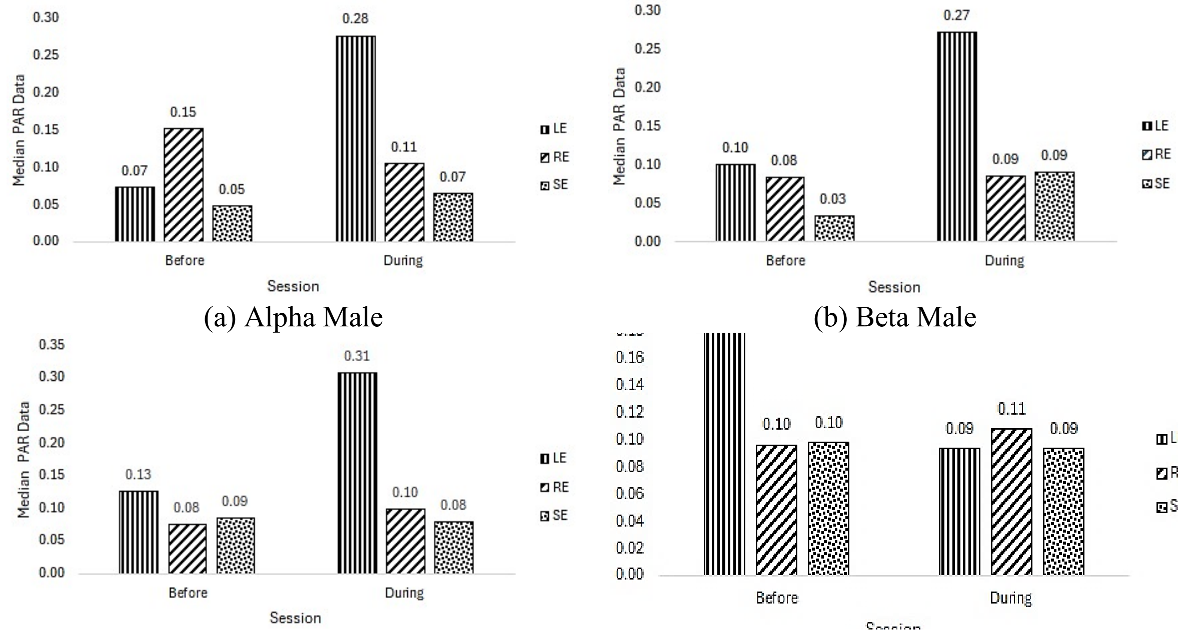
Exposure session	Brainwave	Gender	Correlation Coefficient	Strength of Correlation
Before RF	Beta	Male	0.768	Strong
		Female	0.791	Strong
	Alpha	Male	0.793	Strong
		Female	0.719	Strong
During RF	Beta	Male	0.778	Strong
		Female	0.829	Very Strong
	Alpha	Male	0.817	Very Strong
		Female	0.794	Strong

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.



(a) Alpha Male (b) Beta Male

(c) Alpha Female (d) Beta Female

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.

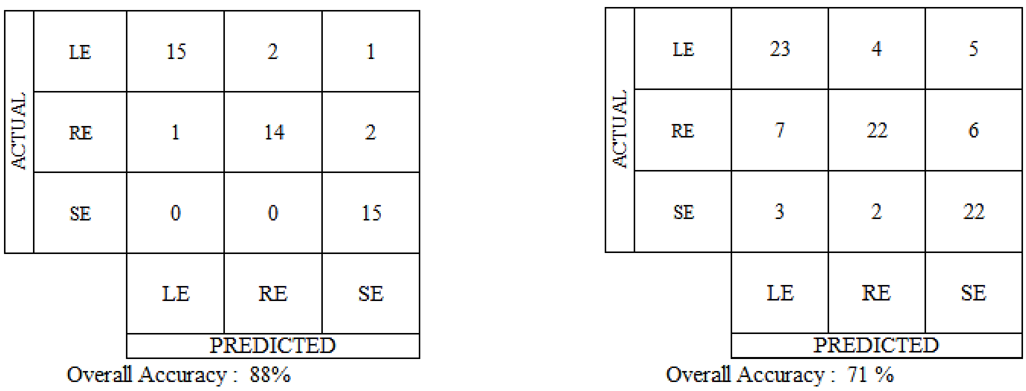
3.3. FNN Classification Performance for Combined EEG Dataset

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**	
Gender	**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.



(a) Before Male Training (b) Before Female Training

(c) During Male Training (d) During Female Training

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.

ACTUAL \ PREDICTED	LE	RE	SE
LE	3	2	1
RE	1	4	1
SE	0	0	1

Overall Accuracy : 62%

(a) Before Male Testing

ACTUAL \ PREDICTED	LE	RE	SE
LE	7	0	3
RE	0	3	2
SE	2	2	4

Overall Accuracy : 61 %

(b) Before Female Testing

ACTUAL \ PREDICTED	LE	RE	SE
LE	3	2	0
RE	1	1	0
SE	1	0	5

Overall Accuracy : 69 %

(c) During Male Testing

ACTUAL \ PREDICTED	LE	RE	SE
LE	6	3	2
RE	1	2	1
SE	2	1	5

Overall Accuracy : 57 %

(d) During Female Testing

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

https://doi.org/10.31436/iiumej.v27i3.4079

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

[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