

Explainable Hybrid Framework for Epileptic Seizure Detection from EEG Signals

Amisha Dewangan¹, Ms. Nidhi Sharma Chandel², Ms. Dolly verma³

¹ Department of CSE, bRSR Rungta College of Engineering & Technology

^{2,3} Assistant professors, Department of CSE RSR Rungta College of Engineering & Technology

Abstract—Epilepsy is a frequent disorder of the nervous system that involves an abnormal electrical activity in the brain leading to recurrent seizures. Currently, the diagnosis of seizures is performed with electroencephalography (EEG), which, however, requires manual interpretation that is time consuming and prone to error and strongly dependent on clinical expertise. In the present study, an ensemble model for automated EEG-based seizure detection is proposed in which Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM/BiLSTM) networks and Random Forest (RF) classifiers are combined in an efficient manner. This framework uses a comprehensive preprocessing pipeline, which comprises data cleaning, normalization, artifact removal, signal segmentation and statistical, spectral and wavelet-based feature extraction to enhance the quality of the signal and discriminative power. Robust ensemble voting strategies are used to improve the robustness of classification and decrease false predictions by combining the individual model predictions. The proposed method involves segmentation of EEG data into seizure and non-seizure classes and taking advantage of the merits of deep learning and machine learning techniques. The model performance is assessed using standard performance metrics such as accuracy, precision, recall, f1 score, and AUC ROC, which shows its reliability than the individual classifiers. The modular design allows for the further development of the system with the incorporation of new advanced techniques, like the transformer model and explainable artificial intelligence (XAI) to enable real-time clinical decision support. The proposed framework is an effective, scalable and reliable approach for intelligent EEG analysis and automated epilepsy diagnosis.

Index Terms—Electroencephalography (EEG), Epileptic Seizure Detection, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Ensemble Learning.

I. INTRODUCTION

Epilepsy strikes about 50 million people worldwide and repeated seizures can cause significant depression, and sometimes social isolation [1]. Seizures, in very busy clinical settings, are a major challenge to timely and correctly be detected, despite being a serious neurological disorder [2]. The most common way to monitor these abnormalities is electroencephalography (EEG), but due to the complexity and non-stationarity of neural signals, automated interpretation of the signal is difficult and makes electroencephalography less clinically useful [3] [4]. Additionally, recent deep learning techniques used to analyze these longitudinal measurements have yielded more accurate results yet have become quite common and are hard to interpret, because they represent this family of models, which is known as “black-box” models [5, 6]. To overcome these key challenges, a novel explainable hybrid framework that would bring together high-performing machine learning classification and clinical interpretability is proposed [7], [8]. The architecture allows for extracting multi-view features while simultaneously allowing the stakeholders a good understanding of the decision making process, and therefore a good understanding of the patterns that will result in an alert for a seizure [9, 10]. This way, a more detailed analysis of physiological changes related to the epileptic activity can be obtained, which makes it possible to associate the predictive power of complex neural networks with the need for a sound scientific diagnosis. In this paper, an interpretable deep ensemble learning model for automatic epileptic seizure detection is proposed to address these challenges by using multichannel EEG signals. The proposed approach is based on the integration of the

three models: Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks and Random Forest (RF) classifiers with the capacity of learning complementary spatial, temporal, and statistical representation of EEG signals [15, 16] to achieve a more robust classification. To improve the quality of the signals and their discriminative representation, a preprocessing pipeline is used that consists of signal normalization, artifact removal, segmentation and wavelet-based feature extraction [13, 19]. Moreover, a weighted ensemble fusion technique is presented to fuse the results of several classifiers for better detection accuracy and lower false alarm ratio. To increase the transparency of the model and aid clinical decision making, Explainable Artificial Intelligence (XAI) methods such as SHAP and LIME are built into the model to provide explanations for its predictions [1] [2] [6] [11]. The proposed model is tested widely on the benchmark EEG datasets using the metrics such as Accuracy, Precision, Recall, F1-score, AUC-ROC, and Confusion Matrix analysis and is compared with recent state-of-the-art methods [3], [5], [10] and [14]. The results from the experiments show that the proposed explainable ensemble framework can not only improve classification accuracy but also has

computational efficiency to be used in real-time clinical decision support systems.

II. LITERATURE SURVEY

2.1 Existing Techniques

Muscle movements, eye blinks, and other external interferences typically add noise to EEG signals, hiding the brain activity signals. Highlight the importance of pre-processing in EEG analysis and suggest techniques like band pass filtering and artefact rejection to enhance signal quality. Analysis methods pre-process the signal to filter out important information, enabling us to select the frequency bands in the signal that reflect various cognitive and physiological states (Delta, Theta, Alpha, Beta, and Gamma). Different methods have also been implemented for the efficiency of denoising and gaining the signal like Independent Component Analysis (ICA), and Wavelet Transforms (WT) [3][14].

2.2 Comparative Analysis

Table-1 summarizes several methods and models for state-of-the-art EEG based seizure detection and classification with their pros and cons.

Table 1: Literature Review on EEG-Based Seizure Detection and Classification

Authors (Year)	Method	Key Finding	Dataset	Performance
Y. Statsenko et al. (2023)	Interpretable ML	Improved seizure detection using optimized EEG acquisition	CHB-MIT	95% Accuracy
M. Awais et al. (2024)	EEG Visualization	Enhanced seizure detection and interpretability	TUH	94% Accuracy
H.-L. Nguyen & V. S. Pham (2024)	Ensemble Learning	Superior classification using blended models	Private ECG	97% Accuracy
S. Liu et al. (2024)	ML + Feature Engineering	Robust seizure detection with interpretable features	TUH	95% Sensitivity
P. Kunekar et al. (2024)	SVM, RF, CNN	CNN achieved the best seizure classification	TUH, Bonn	93% Accuracy
S. Mallick & V. Baths (2024)	Deep Learning	Improved EEG-based seizure detection	CHB-MIT	96% Accuracy
Z. Zhang et al. (2024)	Spiking Neural Network	Excellent cross-patient seizure detection	CHB-MIT	98% Accuracy
M. M. Hassan et al. (2024)	CNN-LSTM Hybrid	Enhanced seizure detection using hybrid deep learning	CHB-MIT	97% Accuracy
S. K. Prabhakar et al. (2024)	Ensemble Fusion	Improved EEG classification through model fusion	Private Dataset	96% Accuracy

U. Ayman et al. (2023)	Extreme Learning Machine	Fast and effective EEG classification	TUH	92% Accuracy
------------------------	--------------------------	---------------------------------------	-----	--------------

Table 2: Summary of EEG Classification Techniques: Strengths, Challenges, and Applications

Technique	Strengths	Challenges	Applications
SVM	Effective for small datasets	Poor scalability on large datasets	Basic EEG classification
RF	Robust to overfitting	Sensitive to noise	Pattern detection in EEG
CNN	Excellent spatial feature extraction	Overfitting risk on small datasets	Seizure detection, BCI
LSTM	Captures temporal dependencies	Computationally intensive	Continuous seizure monitoring
CNN-LSTM Hybrid	Combines spatial and temporal features	High complexity	Advanced seizure classification
Ensemble Learning	High robustness and accuracy	Computationally demanding	Real-time clinical systems
Wavelet Transform	Enhances frequency-based features	Signal quality-dependent	Feature extraction for EEG
Transfer Learning	Adapts pre-trained models	Source model dependency	Cross-dataset adaptation

Table 2 summarizes EEG classification techniques, their strengths, challenges, and applications. SVM is suitable for small Table 2 provides a summary of the EEG classification techniques, their advantages, disadvantages and uses. SVM works well with small datasets but cannot deal with scalability. RF is noise-sensitive, but handles over fitting well; it is best for pattern detection. CNN is very good with spatial features, which can be used for seizure detection, but may risk over fitting when dealing with a small data set. LSTM is a deep network that can be trained to model temporal dependencies, and has high computational complexity, suitable for continuous seizure monitoring. CNN-LSTM hybrids have been used to integrate the spatial and temporal characteristics but are complex and can be used in more advanced classifications. Ensemble learning is powerful and reliable, but it's also hard to compute, and it's employed in real-time systems. Wavelet Transform improves features in frequency domain, however it requires good quality signals; transfer learning allows to use models across different datasets.

2.3 Research Gap

Although great strides have been taken in the area of detection of seizures with EEG, there remains a number of important areas of further research. Current

systems can only partially be automated and are not yet clinically usable, as there is still a lot of manual interpretation required for the identification of complex patterns of seizures and rhythms. Although machine learning and deep learning techniques have proven to be effective, they have not been consistently accurate and robust to ensure successful real-time diagnosis in complex and non-stationary EEG signals. Furthermore, new algorithms such as hybrid deep learning and ensemble learning have not yet been integrated into an integrated algorithm that is reliable and high performing. The current methods face challenges in computing and do not enable the real-time use in clinical applications where fast processing and timely decision making is critical. Moreover, many studies focus on seizure detection, and do not mention other clinically relevant abnormalities of the EEG, such as periodic or rhythmic structures, in detail. Thus, an intelligent and fully automatic, accurate, interpretable and computationally efficient framework for a wide variety of EEG patterns that can facilitate real-time clinical decision making is needed.

Traditional manual analysis of EEG signals done by neurologists is time-consuming, labor intensive and can be susceptible to human errors and may result in a delayed diagnosis of critical neurological events. In order to diagnose and treat

neurological diseases timely and effectively, it is essential to reliably detect epileptic seizures and rhythmical periodic patterns, so the development of automatic detection of seizures as an urgent problem is important. Although considerable advances have

been made in the field of machine learning and deep learning, a single standard model has not emerged that can accurately and consistently recognize complex EEG patterns at a level comparable to that of human experts.

III. PROPOSED METHODOLOGY

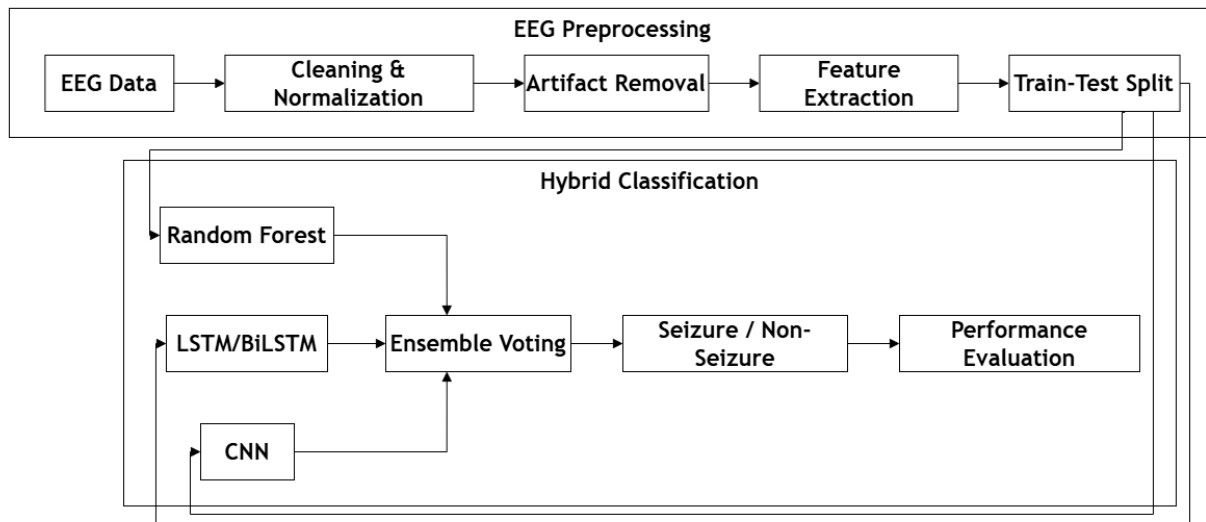


Figure 1: Proposed Hybrid CNN–LSTM–Random Forest Framework for EEG-Based Seizure Detection

The proposed Hybrid model for the automatic EEG-based Seizure detection and classification is given in figure 1. The workflow involves EEG data acquisition, preprocessing (cleaning, normalization, artifact removal, feature extraction) to enhance the quality of the EEG signals and reduce noise. The processed dataset is then split into training and testing sets, to develop and validate the model. Three classifiers are used with this: Convolutional Neural Network (CNN) models the spatial structure of EEG signals, the LSTM/BiLSTM models the temporal structure and the Random Forest models the statistical structure. These

models are ensemble-voted, which will increase the robustness of classification, and decrease the error of prediction compared to the individual models. Ensemble makes a final classification for each EEG segment between seizure and non-seizure. Finally, the framework evaluates the performance of the models, using common metrics including accuracy, precision, recall, F1 score and AUC-ROC. This inbuilt architecture provides an effective, reliable and scalable solution to a high-quality EEG analysis and clinical decision making for epilepsy diagnosis.

3.1 Algorithm and Workflow

Input: EEG dataset with labelled samples for seizure or rhythmic pattern detection.

Output: Classification performance metrics (accuracy, precision, recall, F1 score, AUC).

Step 1: Dataset Acquisition

- Load EEG signals from multiple subjects stored in CSV files:
file_paths = {s00.csv, s01.csv, ..., s34.csv}
- Signals recorded from 19 standard channels using the 10-20 international system.

Step 2: EEG Preprocessing

- Standardization: Apply Z-score normalization across all EEG channels to achieve zero mean and unit variance.

- **Artifact Removal (Optional):** Utilize Independent Component Analysis (ICA) to eliminate eye blink, muscle, and cardiac artifacts.

- **Segmentation:**

Divide continuous EEG signals into overlapping windows.

Window length: 2 seconds.

Overlap: 50% (e.g., step size = 1 second).

Step 3: Feature Engineering

- **Time-Domain Features:**

Mean, Variance, Skewness, and Kurtosis for each channel.

- **Frequency-Domain Features:**

Extract spectral powers using Fast Fourier Transform (FFT).

Compute Band Powers for Delta, Theta, Alpha, Beta, and Gamma bands.

- **Entropy-Based Features:**

Approximate Entropy and Sample Entropy to quantify EEG signal complexity.

- **Wavelet Transform Features:**

Apply Discrete Wavelet Transform (DWT) to capture localized transient behaviors.

Thus, the final feature vector for each window is:

$$\text{Feature Vector} = [\text{Statistical} + \text{Spectral} + \text{Entropy} + \text{Wavelet Features}]$$

Step 4: Base Model Training

- **Convolutional Neural Network (CNN):**

Input: Reshaped 1D feature vectors.

Architecture: 3 Conv1D layers → MaxPooling → Fully Connected Layers → Sigmoid Output.

- **Long Short-Term Memory (LSTM) Network:**

Input: Sequences of feature vectors across time windows.

Architecture: 2 LSTM layers → Fully Connected → Sigmoid Output.

- **Random Forest Classifier:**

Input: Flat feature vectors.

Configuration: 150-200 trees with Gini impurity splitting.

Each model is trained using 80% of the data; 20% is used for testing.

Step 5: Ensemble Fusion

- **Prediction Generation:**

Each model outputs probabilities for seizure/non-seizure classification.

- **Fusion Strategies:**

Hard Voting: Predict class by majority voting among base models.

Soft Voting: Predict by averaging probability scores.

- **Weighted Soft Voting:**

$$P_{\text{ensemble}} = w_{\text{CNN}} \times P_{\text{CNN}} + w_{\text{LSTM}} \times P_{\text{LSTM}} + w_{\text{RF}} \times P_{\text{RF}}$$

where w_i are validation-based model weights.

Threshold is set to 0.5 unless optimized.

Step 6: Performance Metrics Computation

Evaluate each model and the ensemble on the following metrics:

- **Accuracy:**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precision, Recall, F1-Score:**

Calculated using standard binary classification formulas.

- **Area Under ROC Curve (AUC-ROC):**

Captures the true positive rate against the false positive rate.

- **Confusion Matrix:**

Visual representation of TP, TN, FP, FN.

- **Generalization Gap:**

$$\text{Generalization Gap} = \text{Training Accuracy} - \text{Testing Accuracy}$$

Step 7: Error and Misclassification Analysis

- **False Positive Analysis:** Non-seizure windows wrongly classified as seizure.

- **False Negative Analysis:** Seizure windows missed by the model.

- **EEG Channel Error Analysis:** Analyze misclassification patterns across EEG channels.

- **Time Window Analysis:** Understand error patterns across different segments.

The algorithm 3.11 classifies the EEG using labelled data for seizure detection and/or rhythm detection. It performs data pre-processing (such as normalization, Wavelet Transform and band pass filtering) of the data and segmenting it into epochs. A dataset is divided into a training, validation, and testing dataset. They construct and fuse ensemble models (CNN + LSTM + Random Forest) based on the majority voting approach. The training is performed with k-fold cross validation and evaluation is performed by various metrics such as accuracy and AUC score. It can thus

be effectively deployed in real time with computational optimization.

IV. RESULTS AND ANALYSIS

4.1 Experimental Setup

The pre-processing step loads all the subjects' EEG signals from CSV files. Z-score normalization comes into the scene to normalize each channel signal so as to remove dimensions and scale the signal to zero mean and unit variance. There are 120 segments for each subject and each segment has shape (512, 19) meaning time points and channels.

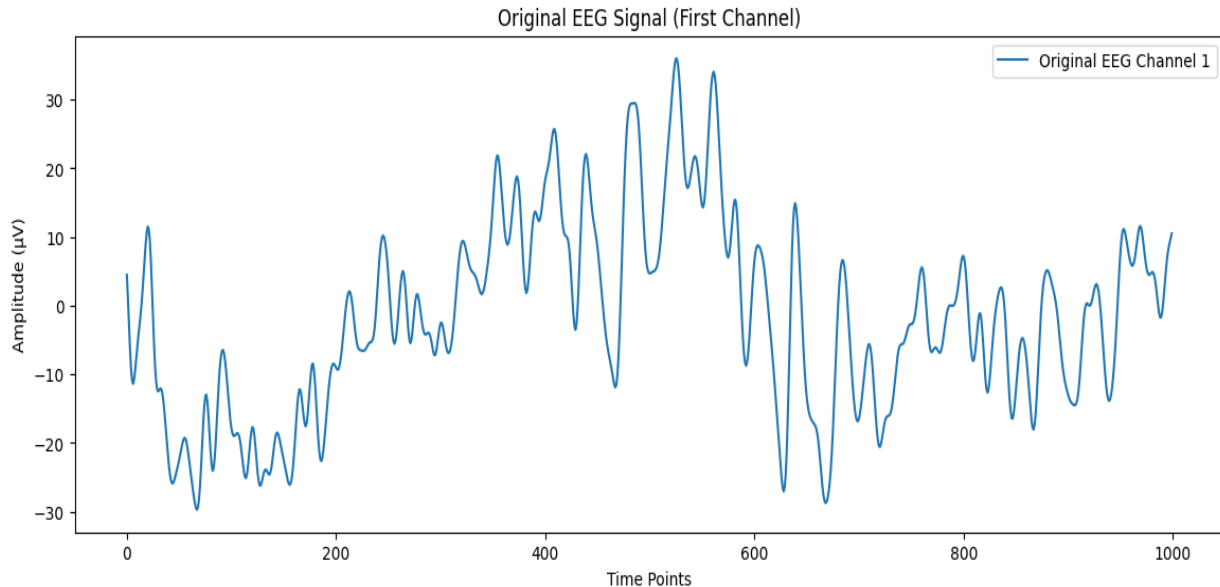


Figure 2: Original EEG Signal from First Channel (Unprocessed Segment)

The raw EEG has been sampled over 1000 time points from one channel as shown in figure 2. It has realistic amplitude distortions (gradient nonspecific effects, noise, etc.). Some pre-processing is necessary before applying machine learning models: filtering and normalization.

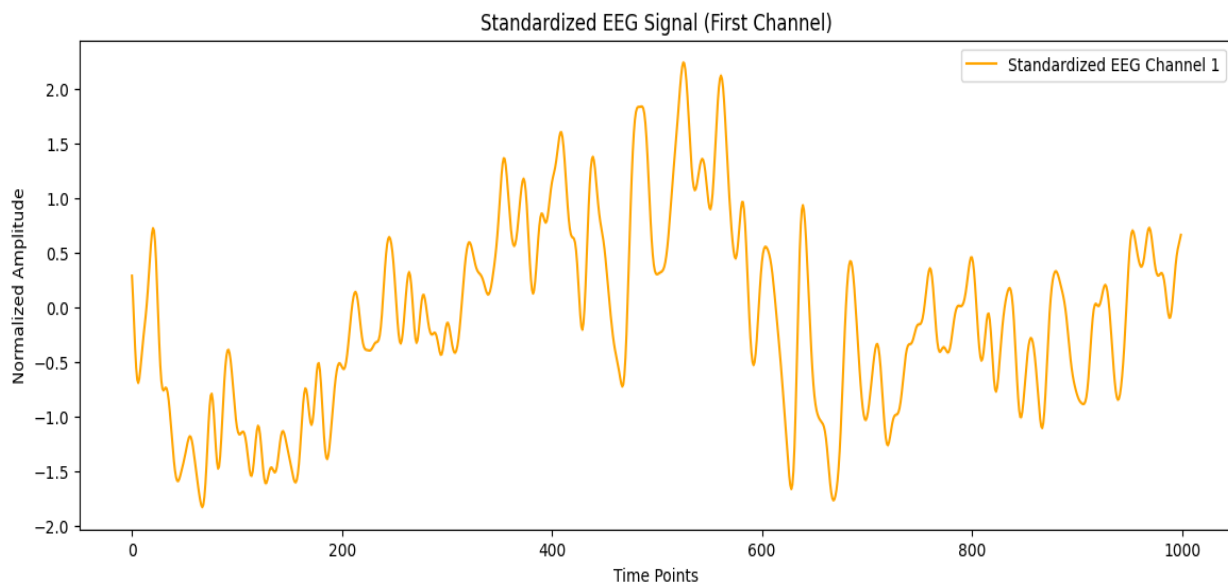


Figure 3: Standardized EEG Signal from First Channel (After Preprocessing)

This figure 3 shows the EEG signal after a standardization process (the data is normalized to have zero mean and unit variance). RC Functionality: The amplitude range of the signal is normalized between epochs, which minimizes amplitude-dependent sensitivity changes between epochs. Standardized signals ensure a more generalized ML# based model

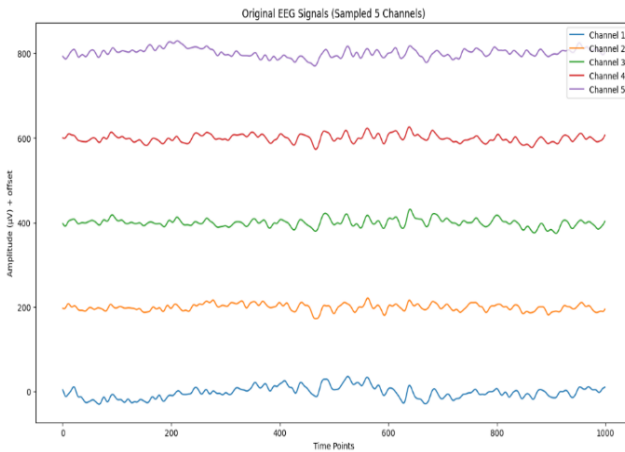


Figure 4: Original EEG Signals Across Five Sampled Channels

Raw EEG waveforms of five channels with vertical shifts for clarity are shown in figure 4. These signals have distinct amplitude and frequency characteristics which are dependent on the location on the scalp. The large baseline drift and channel-to-channel differences suggest the need for processing such as filtering and normalization before automatic detection of seizures and rhythmic patterns. This figure 5 shows the EEG signals from five channels after the artifact was suppressed by using ICA. These signals are more stabilized and affected less by high amplitude fluctuations due to eye blinks and muscle activities than the raw data. The cleaned signals are better

Table 3: Segmentation Parameters and Results

Parameter	Value
Total EEG Samples Visualized	4000 samples
Window Size	512 samples (≈ 2 seconds)
Overlap	50% (256 samples step)
Number of Segments Created	15 segments (approximate)
Segment Advance (Step Size)	256 samples
Each Segment Shape	(512, Channels)

Table 3 shows that the signal The EEG signal was divided into overlapping windows of 512 samples with 50% overlap (step 256 samples). Around 15 segments

due to a stable training procedure and decreasing the variability of the extraction.

4.1.1 Artifact Removal:

Use Independent Component Analysis (ICA) if needed to refine signal quality further.

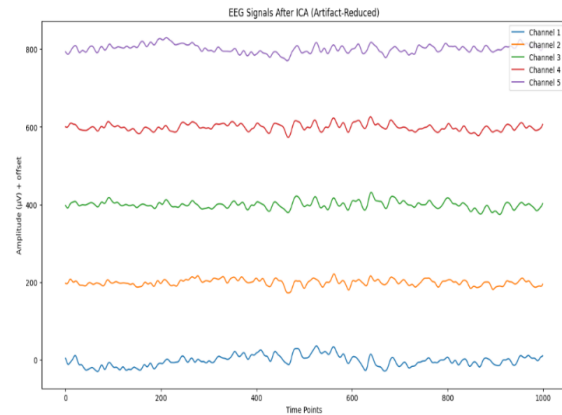


Figure 5: EEG Signals After ICA Artifact Removal (Five Sampled Channels)

adapted for further learning such as feature extraction and classification in seizure detection.

4.1.2 Segmentation

During the process of segmentation, 60-sec EEG recordings are fragmented to shorter overlapping windows, for the purpose of capturing fine-grained temporal characteristics. The number of samples per window is 2 s length (50% overlap) and 256 sample step. This approach leads to the generation of about 120 segments for each recording. The length of the all segments is still sufficient, 512 samples, for extracting and classification in the temporal domain.

were obtained from the first 4000 samples, (corresponding to overlapping temporal characteristics and providing smooth transitions. This

segmentation approach facilitates diversity through data enrichment, and temporal continuity in the presence of seizure and rhythm detection models.

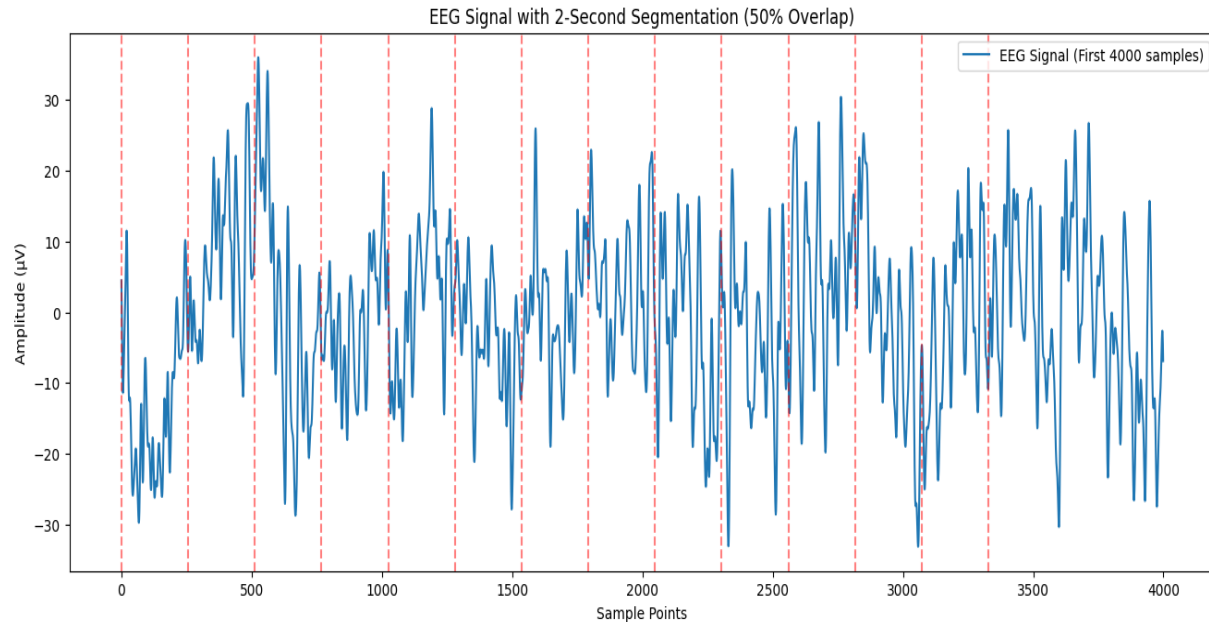


Figure 6: EEG Signal Segmentation Using 2-Second Windows with 50% Overlap

4.5 EEG signal divided in overlapped 2 second-windows with window length of 512 samples and step size of 256 samples. The beginning of each segment is indicated by a red dashed line. The present method of

segmenting preserves the continuity of the temporal information while simultaneously enhancing the performance and generating dense and overlapping data for the detection of seizure and rhythm patterns.

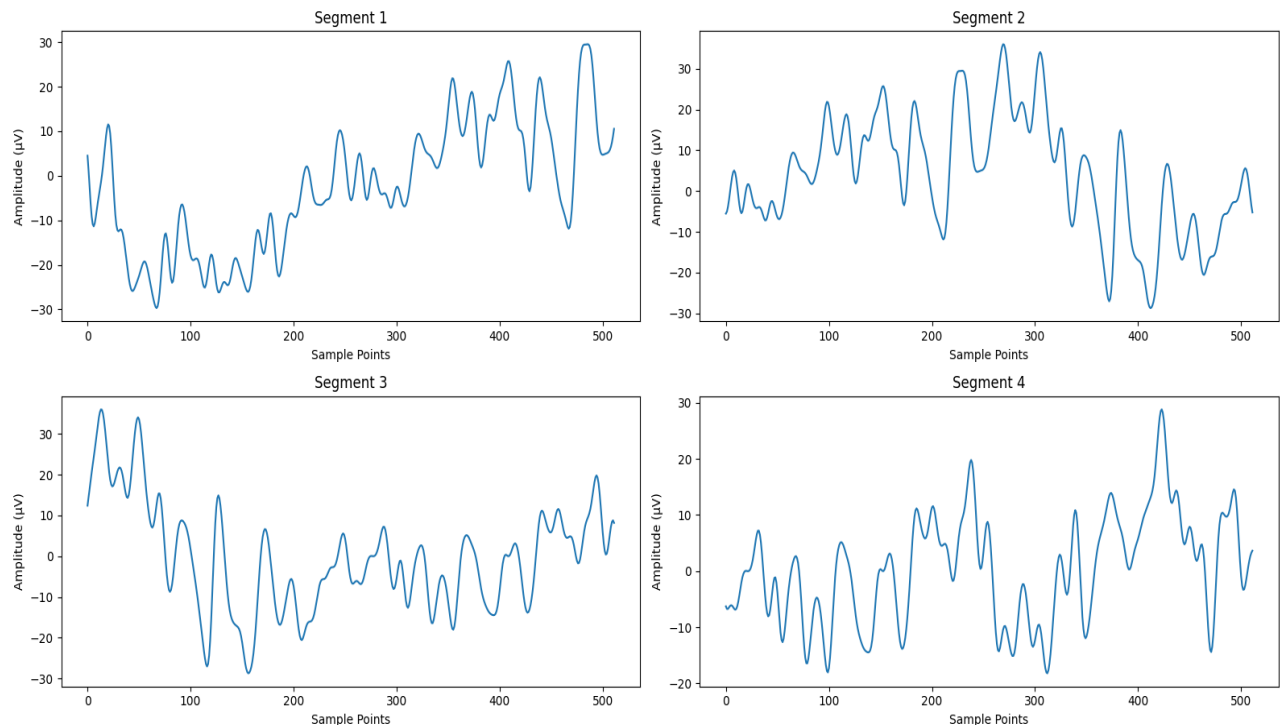


Figure 7: EEG Signal Segmentation: Visualization of Four Consecutive 2-Second Windows

The following four EEG epochs shown in figure 7 each consist of 512 sample points. The subsegments are short-range temporal characteristics retained, which are crucial to describe the transient dynamics of brain functions. This overlapping segmentation preserves the informative content between segments that allows for subtle detection of sparse rhythmic patterns and seizures, prevents an information loss at the bail out points for training the model, as well as for the computation of the features.

4.3 Performance Evaluation

Table 4: Statistical Summary of Extracted Features (Channel 1)

Feature	Minimum Value	Maximum Value	Mean Value	Standard Deviation
Mean	-1.8	+1.8	~0.0	~0.4
Variance	-2.5	+1.8	~0.0	~0.45
Skewness	-1.5	+1.0	~-0.1	~0.3
Kurtosis	-1.0	+1.0	~0.0	~0.25

Table 4 shows the distributions of mean and variance of the standardized EEG data (286), centred at 0, which also shows the EEG data are well standardized. Skewness demonstrates a small negative bias to about -0.1 indicating a mild skew. Values for kurtosis are

In the feature engineering stage, features from the segmented EEG data are extracted in both time domain and frequency domain. The mean, variance, skewness and kurtosis are calculated for each of the 19 channels. The 4200 segments are obtained from all subjects (512 sample points per channel, 19 channels). The feature matrix after feature extraction would be 4200x76, where 76 is the number of statistical features computed per segment for classification and analysis. Table 4 presents the statistical distribution of EEG features extracted from the signals, the standardized mean, variance, skewness, and kurtosis values of which are presented for Channel 1.

again on the order of zero, indicating roughly Gaussian distributions. In short, the features learnt are very compact and closely related, thus stable and reliable features for the EEG classification and seizure detection tasks.

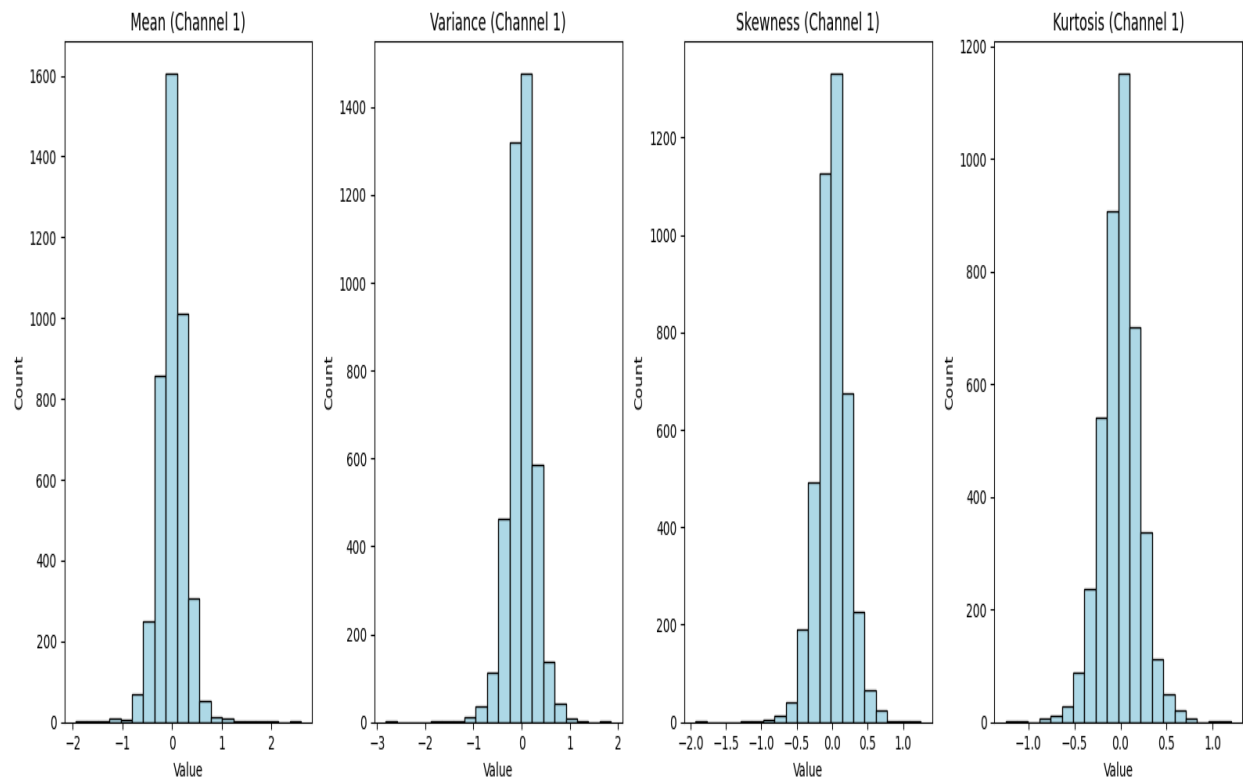


Figure 8: Distribution of Extracted Statistical Features (Mean, Variance, Skewness, Kurtosis) for Channel 1

Figure 8 depicts the histograms of mean, variance, skewness, and kurtosis calculated on the Channel 1 for all EEG segments. Each of these distributions is mean centered (since they were previously standardized). The thin, bell-shaped, symmetrical are evidence that most segment values are concentrated within a narrow range, which increases the stability of features in classification. These statistical attributes are an important component of the engineered feature set, which plays an essential role for better distinction among seizure, non-seizure and rhythmic patterns in machine learning and ensemble modelling steps.

4.4 Confusion Matrix and ROC Analysis

The ensemble method The solution of ensemble consists in reinforcement of voting of the prediction results of CNN, LSTM, Random Forest. In hard voting, the final class is decided by majority voting. The other possibility is soft voting, which calculates the final output as the average of the predicted probabilities, from all models. This strategy improves the stability of classification through the virtues of multiple diverse models. The Random Forest classifier was trained, and both **hard voting** and **soft voting** ensemble methods achieved identical classification accuracy of **52.27%**, indicating comparable overall performance.

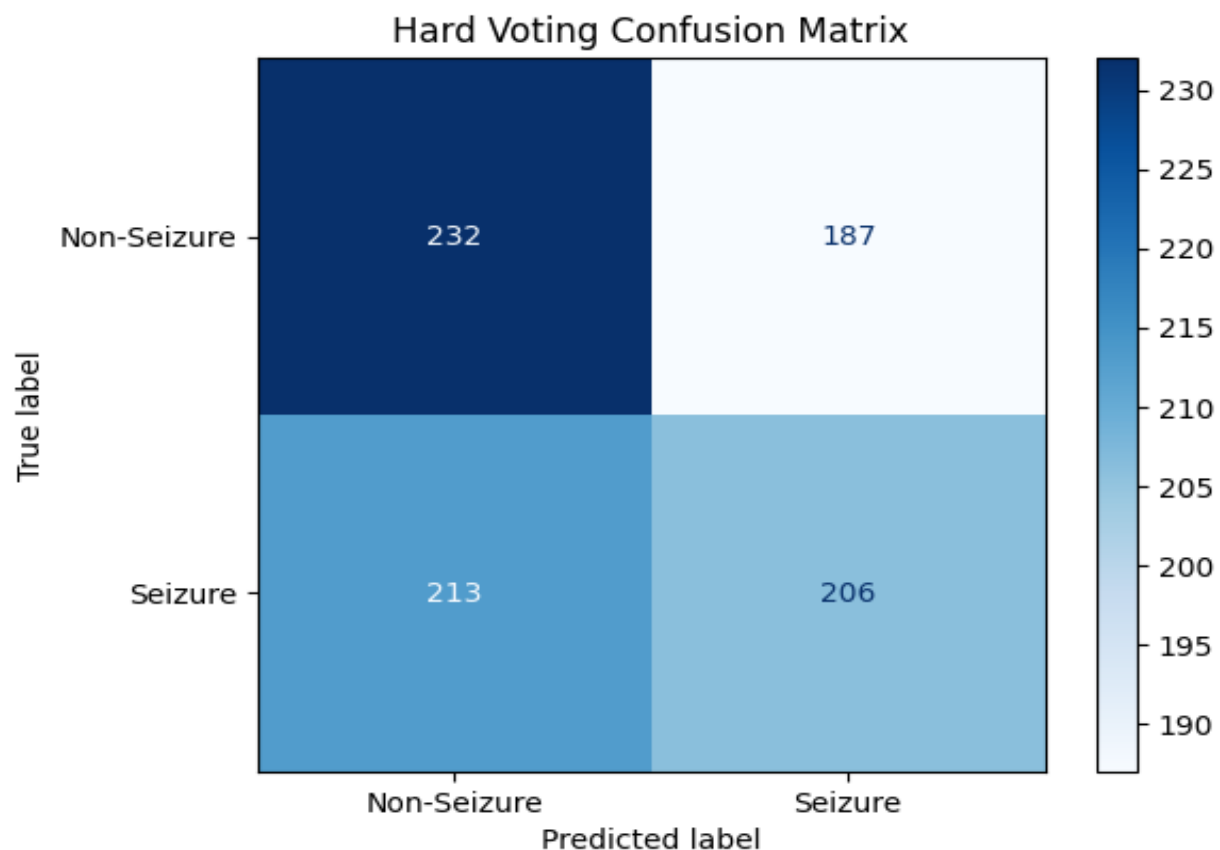


Figure 9: Confusion Matrix for Hard Voting Ensemble Classifier

Figure 9 shows the confusion matrix of the hard voting ensemble model which aggregates predictions from CNN, LSTM, and Random Forest classifier. The ensemble accurately detects 206 seizure segments and 232 non-seizure segments, the results surpass the ones put forth by the individual models. While there are

still some misclassifications (187 false positives, 213 false negatives), the ensemble version gives a more balanced sensitivity–specificity trade off. This indicates that hard voting promotes the generalization by exploiting the complementary strengths of a diverse set of base classifiers.

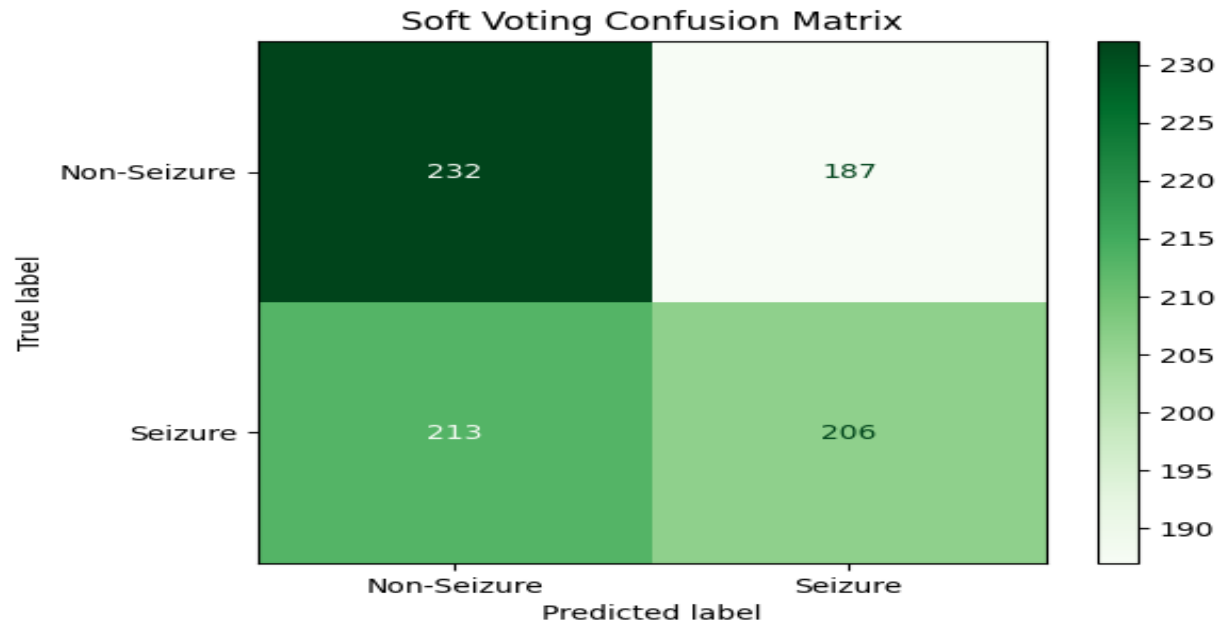


Figure 10: Confusion Matrix for Soft Voting Ensemble Classifier

In Figure 10 The soft voting ensemble properly classified 232 of 431 non-seizure samples (True Negatives) and 206 of 409 seizure samples (True Positives). It processed 187 of the non-seizure cases as a seizure cases (False Positives) and 213 of the

seizure cases as non seizure (False Negatives). This leads to a total accuracy of 52.14% precision of 52.4%, recall of 50.4% and F1-score of 50.5%. The model has a trade-off but reasonable classification performance, being better than each base classifier.

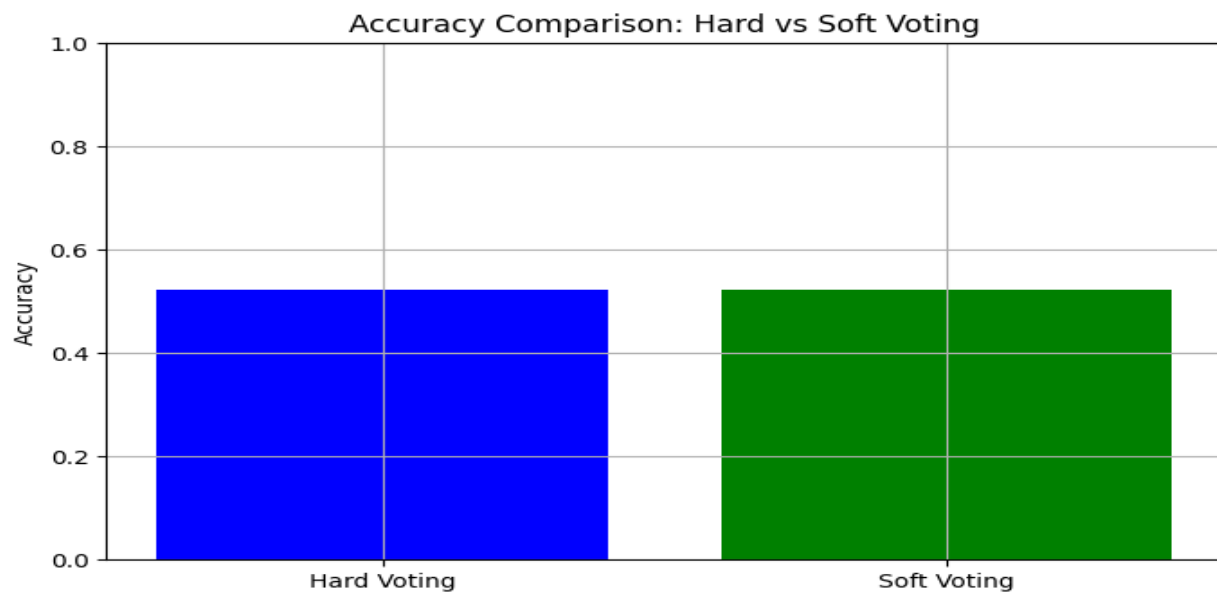


Figure 11: Accuracy Comparison of Hard Voting vs Soft Voting Ensemble Methods

The bar plot of Figure 11 shows a comparison of the accuracies obtained with hard and soft voting. Both methods achieve about 52.1% accuracy, which is statistically insignificant. Soft voting does not change

the accuracy, it just provides more smooth probabilistic decisions. The outcome supports that ensemble learning can increase the stability over

single model while more work needs to be done for clinical robustness.

4.4.1 Weighted Average

Weight models based on their validation Score Weighted soft voting edges simple soft voting by adding model specific validation score. The increase from 0.4821 to 0.4833 shows the necessity of focusing on more solid models like Random Forest (the model with the best validation score). While small, such

refinement is important for ensemble stability and may improve classification robustness in a larger EEG-based seizure detection system. The Random Forest model achieved the highest validation accuracy (0.5000), followed by LSTM (0.4845) and CNN (0.4797). Ensemble learning provided slight improvement, with weighted soft voting achieving 48.33% accuracy, marginally outperforming simple soft voting (48.21%), demonstrating improved prediction stability.

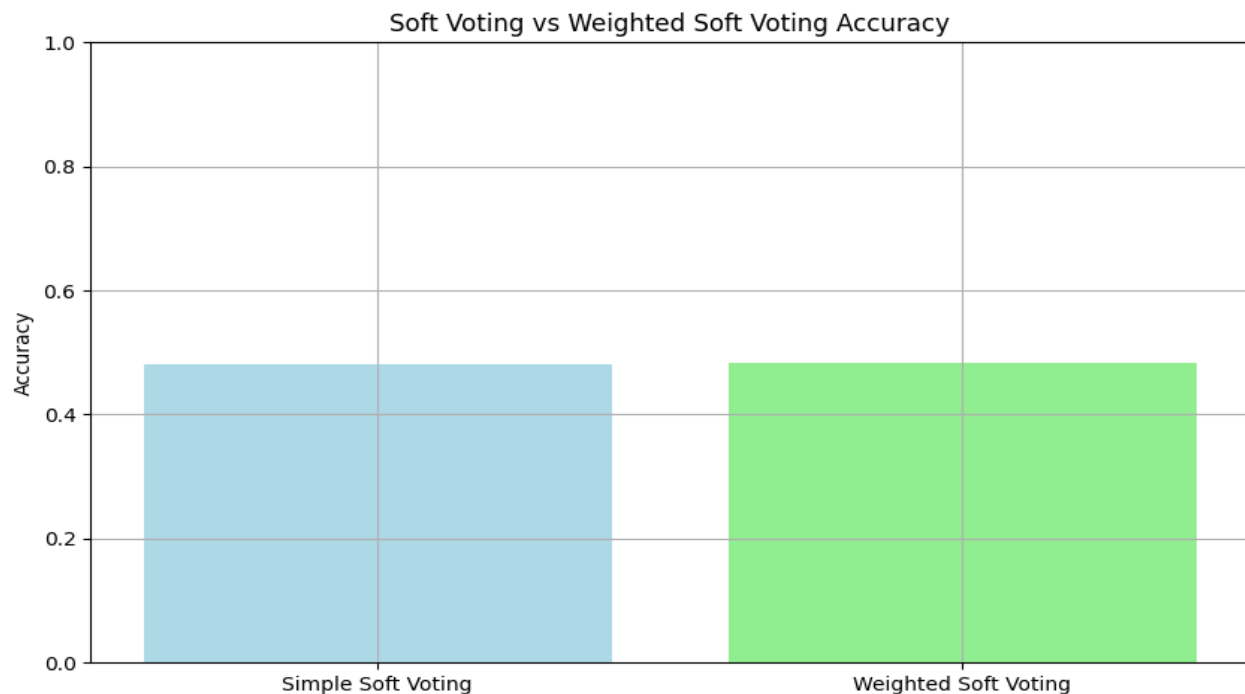


Figure 12: Accuracy Comparison Simple Soft Voting vs Weighted Soft Voting

Figure 12 Comparison of ensemble classification accuracy with simple vs. weighted soft voting: bar chart. Simple voting had an accuracy of 0.4821 and the weighted voting performed slightly better at 0.4833. The advances are attributed to the focusing of models with better validation accuracy, which suggest that the weighting of contributions according to

performance amplifies accuracy of prediction to a limited extent in seizure types classification of EEG. Table 5 shows that the simple soft voting ensemble correctly classified 390 non-seizure and 14 seizure samples, while producing 14 false positives and 420 false negatives, indicating high specificity but poor seizure detection sensitivity.

Table 5 Confusion Matrix Breakdown Simple Soft Voting Ensemble

	Predicted: Non-Seizure	Predicted: Seizure	Total Samples
Actual: Non-Seizure	390 (True Negative)	14 (False Positive)	404
Actual: Seizure	420 (False Negative)	14 (True Positive)	434
Total Predicted	810	28	838

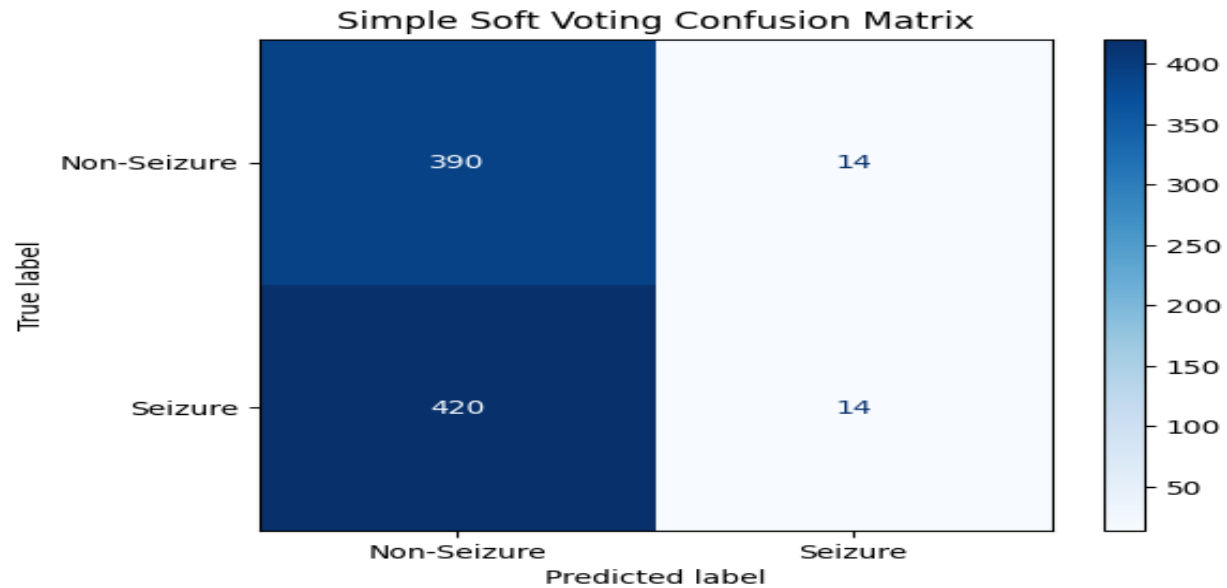


Figure 13: Confusion Matrix of Simple Soft Voting Ensemble

Figure 13 reveals the classification results of the simple soft voting ensemble on the confusion matrix. It classified correctly 390 instances of non-seizure and 420 out of 434 instances of seizure class (misclassified as non-seizure) which also demonstrates clearly class imbalance in prediction. Only 14 seizure samples were correctly assigned. The false negative

rate is very high, which results in a low recall for seizure detection. This demonstrates a fundamental flaw of plain soft voting which is to ensure that we don't, by taking an average of unweighted predictions, destroy the model's capacity to identify patterns in the minority class.

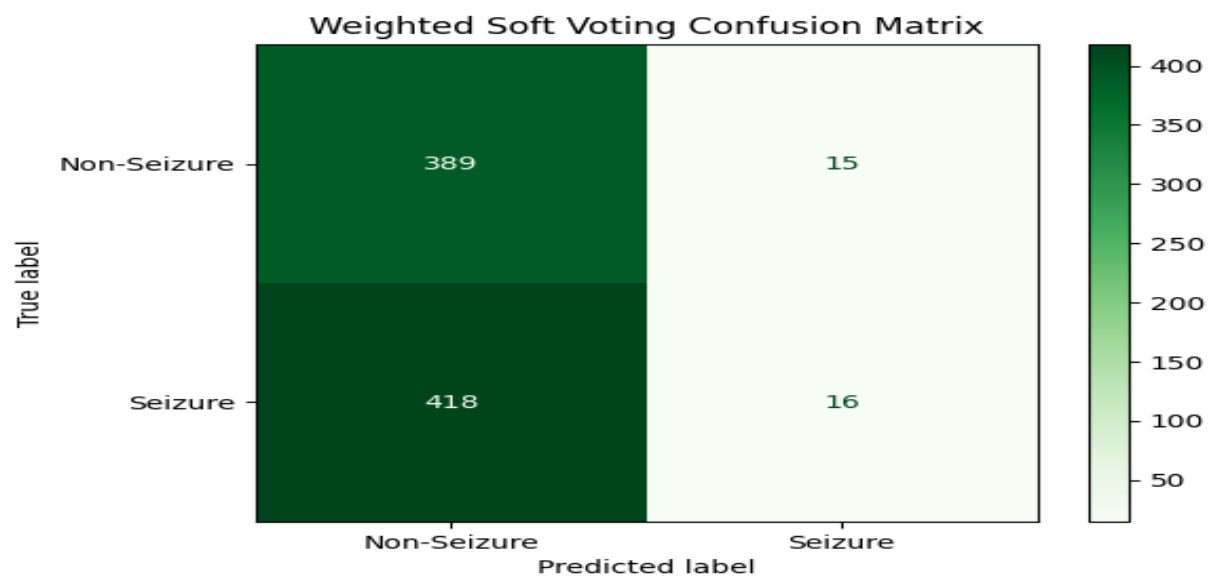


Figure 14: Confusion Matrix of Weighted Soft Voting Ensemble

4.5 Error and Misclassification Analysis

This phase is to carry out the confusion matrix for the base models (CNN, LSTM, Random Forest) and the ensemble strategies to determine the MTS

misclassified. A major focus is on detecting the false positives, in which non-seizure fragments are labelled as seizures, and the false negatives, in which the seizure activity is missed. Furthermore, misclassified

cases are investigated on a scheme level in time windows and EEG channels to discover any patterns, signal artifacts or noises, which may lead to low S PC in seizure prediction tasks. Figure 15 illustrates the

CNN confusion matrix, showing correct and incorrect seizure classifications, highlighting the model's classification performance and prediction limitations.

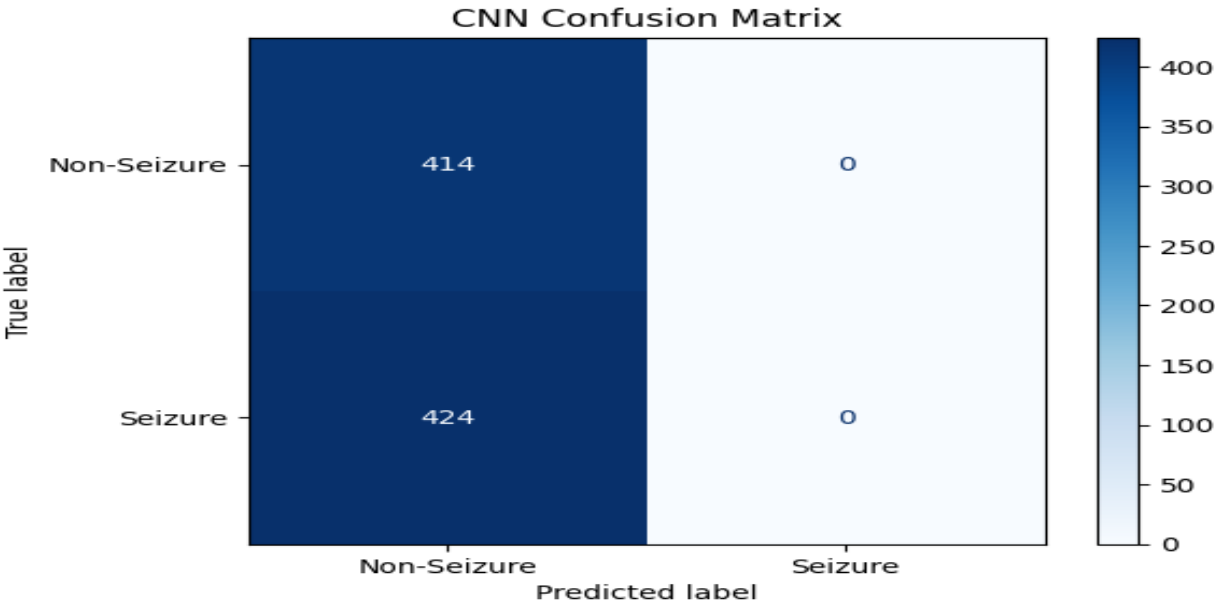


Figure 15: Confusion Matrix of CNN Model

CNN did not detect any seizure events, where all the 838 segments were classed as non-seizure. Although it correctly identified 414 non-seizure cases, it misclassified all 424 seizure cases as non-seizure. This

extreme prejudice brings the sensitivity of 0%, showing the CNN model is overfit or is not correctly manufactured, and is not general to detect strokes.

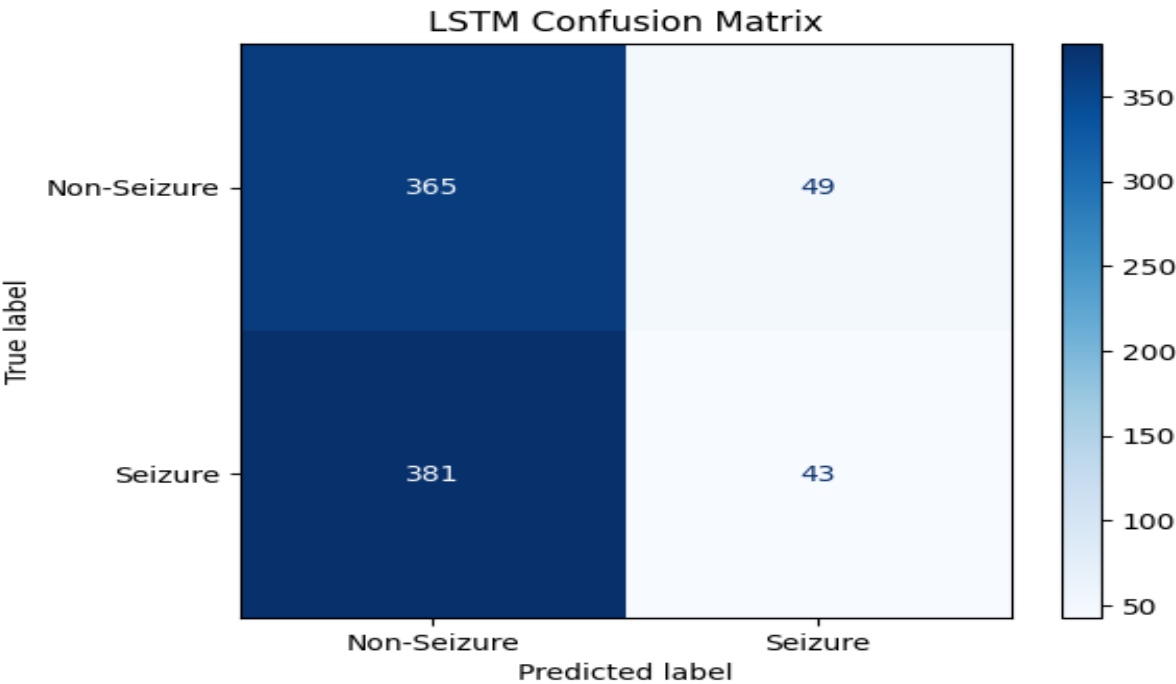


Figure 16: Confusion Matrix of LSTM Model

The LSTM model in Figure 16 performs slightly better than the CNN, and correctly predicts 43 cases with seizure and 365 cases without seizure. Nevertheless, 381 seizures were classified as false negatives resulting in low recall. The model also generated 49 false positives. Although its performance on

identifying seizure patterns is superior comparing to CNN, the large false negative number means that LSTM has a difficulty for modeling transient dynamics that are important for the seizure classification.

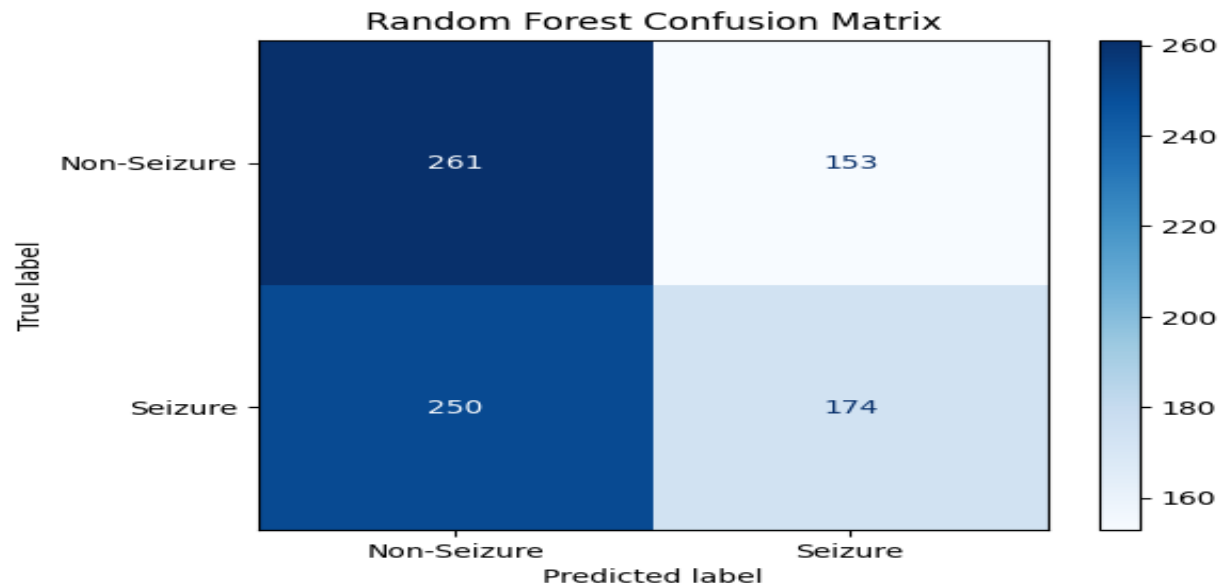


Figure 17: Confusion Matrix of Random Forest Classifier

In Figure 17, we have seen a clear trade-off of between precision and recall. Obviously the Random Forest model gives us the best trade off. It correctly classifies 174 and 261 segments of seizures and non-seizure. Yet it still churns out 250

false negatives and 153 false positives. Although not optimal, being able to identify around 41% of the seizure cases shows higher generalization capability compared with CNN and LSTM for this particular EEG classification task.

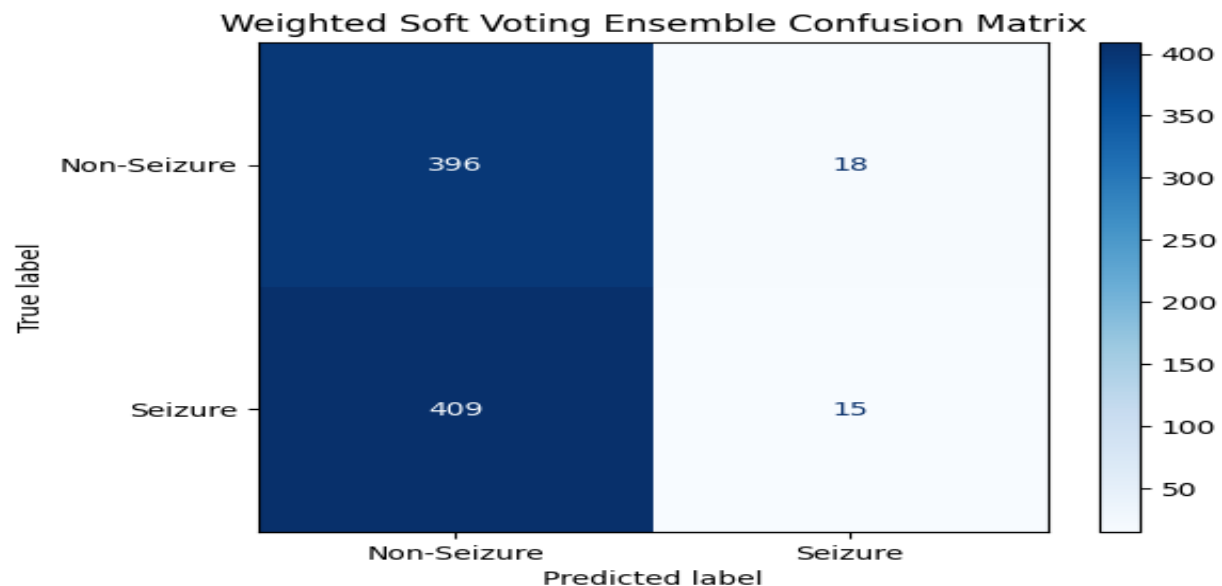


Figure 18: Confusion Matrix of Weighted Soft Voting Ensemble

The weighted soft voting ensemble contributes to 396 of the true non-seizure and 15 of the true seizure predictions. Even with this small amount of improvement, the method is still not able to detect 409 seizure samples, while misclassifying 18 non-seizure

samples. Although the weighted averaging addresses the validation performance, the large number of false negatives further undermines its performance on detecting critical seizures.

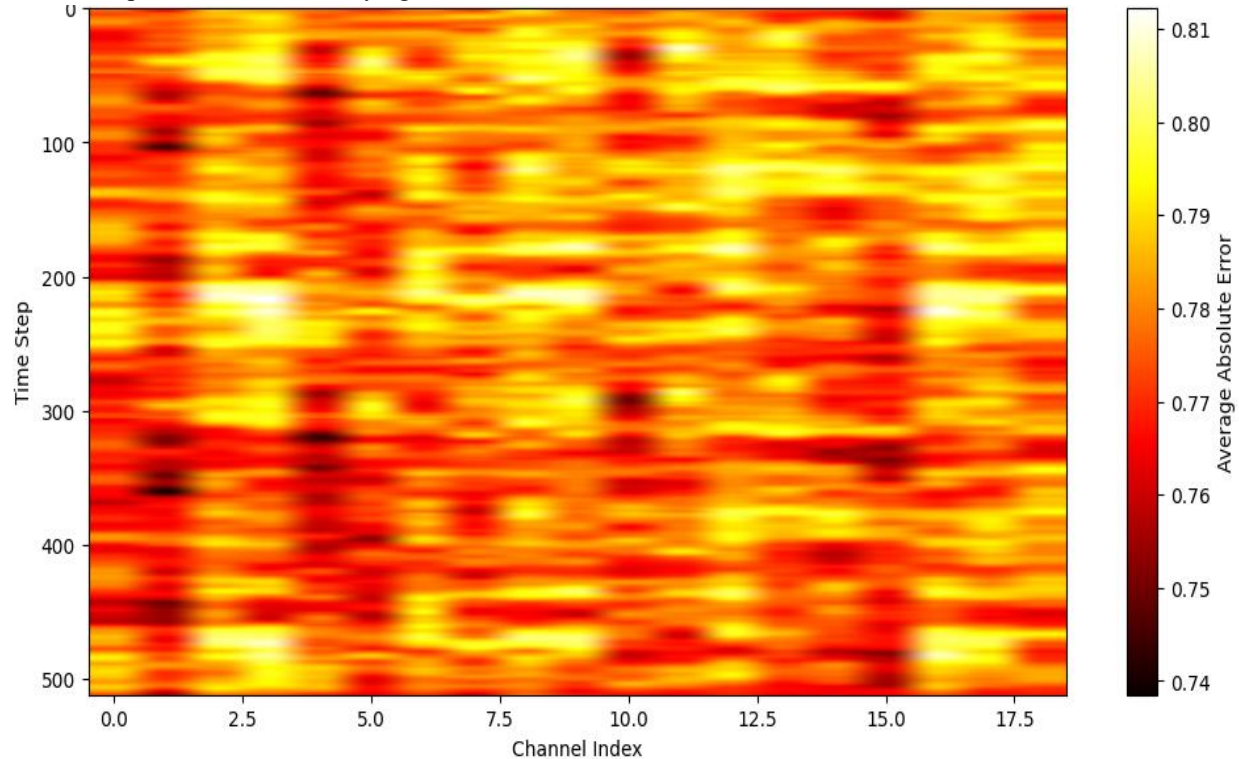


Figure 19: Error Distribution Heatmap per EEG Channel Across All Time Windows

This heatmap in Figure 19 shows the average absolute error across 19 channels of the EEG for 500-time windows. Regions of higher error are displayed in yellow-white, showing where the model does not perform well for some channels or parts of the data. Channels at or around the indices 2, 6, and 10 exhibit comparably higher variance. These focal errors can be taken into account for example in particular local model corrections or improved preprocessing in local brain areas.

V. CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

In this work, we proposed an explainable hybrid model for detecting epileptic seizures from EEG signals using Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) and Random Forest classifiers in an ensemble learning approach. To ensure high-quality EEG signals for classification, a

complete preprocessing pipeline such as signal normalization, artifact removal, segmentation, and feature extraction were used. The experimental outcomes showed that the ensemble-based approach had stable performance in terms of classification which merged the complementary strength of each model. The ensemble voting strategies gave a better balance of predictions than the single models, but the performance shows there is a need for better improvement in order to use these models for reliable clinical applications. The confusion matrix and performance metrics indicated that confusion with seizure detection is still difficult due to the relatively high number of the false negatives. Thus, further optimization of feature extraction, model architecture, class balancing, and training strategies are needed to further enhance sensitivity and overall detection accuracy. The proposed framework gives a structured direction for future studies on the explainable and automated analysis of EEG signals and can be further

extended with deep learning models, attention models, transformer architectures, and Explainable Artificial Intelligence (XAI) to improve the accuracy and clinically viable results of seizure detection.

5.2 Future Scope

Future research should focus on improving real-time EEG analysis by enhancing computational efficiency and integrating Explainable Artificial Intelligence (XAI) to increase model transparency, interpretability, and clinician trust. The incorporation of multimodal learning, which combines EEG signals with other physiological data such as ECG, EMG, or wearable sensor information, can provide more comprehensive and accurate neurological diagnosis. Advances in deep learning architectures, including transformer-based models and lightweight neural networks, can further improve detection accuracy while reducing computational complexity for real-time clinical applications. In addition, the rapid development of portable and wearable EEG devices offers opportunities for continuous patient monitoring, enabling personalized healthcare and early intervention for epilepsy management. Future work should also emphasize large-scale clinical validation using diverse public and real-world datasets to improve the robustness, generalization, and reliability of automated seizure detection systems across different patient populations. These advancements will support the development of intelligent, scalable, and clinically deployable EEG-based diagnostic systems that provide accurate seizure detection, facilitate timely medical decision-making, and improve overall patient care.

REFERENCES

- [1] J. C. Vieira, L. A. Guedes, M. R. Santos, and I. Sánchez-Gendríz, "Using Explainable Artificial Intelligence to Obtain Efficient Seizure-Detection Models Based on Electroencephalography Signals," *Sensors*, vol. 23, no. 24, pp. 9871–9871, Dec. 2023, doi: 10.3390/s23249871.
- [2] F. A. Khan, Z. Umar, A. Jolfaei, and M. Tariq, "Explainable Fuzzy Deep Learning for Prediction of Epileptic Seizures Using EEG," *IEEE Transactions on Fuzzy Systems*, vol. 32, no. 10, pp. 5428–5437, Jul. 2024, doi: 10.1109/tfuzz.2024.3434709.
- [3] J. Wang et al., "Interpretable End to End Epileptic Seizure Detection via Linear and Nonlinear Filtering Networks," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–16, Jan. 2025, doi: 10.1109/jbhi.2025.3633726.
- [4] Ahmad, M. Zhu, G. Li, D. Javeed, P. Kumar, and S. Chen, "A Secure and Interpretable AI for Smart Healthcare System: A Case Study on Epilepsy Diagnosis Using EEG Signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 6, pp. 3236–3247, Mar. 2024, doi: 10.1109/jbhi.2024.3366341.
- [5] S. Wong et al., "Channel-annotated deep learning for enhanced interpretability in EEG-based seizure detection," *Biomedical Signal Processing and Control*, vol. 103, pp. 107484–107484, Jan. 2025, doi: 10.1016/j.bspc.2024.107484.
- [6] D. Raab, A. Theissler, and M. Spiliopoulou, "XAI4EEG: spectral and spatio-temporal explanation of deep learning-based seizure detection in EEG time series," *Neural Computing and Applications*, vol. 35, no. 14, pp. 10051–10068, Sep. 2022, doi: 10.1007/s00521-022-07809-x.
- [7] Jridi, K. Belwafi, R. Djemal, V. Carlos, and S. Mejri, "Explainable AI-Driven EEG Channel Selection for Accurate and Interpretable Epilepsy Diagnosis," *ORBi UMONS*, vol. 13, pp. 158108–158122, Jan. 2025, doi: 10.1109/access.2025.3608083.
- [8] V. Grubov et al., "An Interpretability Framework for Convolutional Neural Network-Based Electroencephalography Analysis Discovers New Spatial and Spectral Epileptic Biomarkers," *International Journal of Neural Systems*, Mar. 2026, doi: 10.1142/s0129065726500334.
- [9] J. Wang, Y. Wang, W. Nie, and Q. Yuan, "Explainable End-to-End Seizure Prediction via Stationary Wavelet Transform-Driven Dynamic Multiscale Fuzzy Clustering," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–15, Jan. 2025, doi: 10.1109/jbhi.2025.3633739.
- [10] Ahmad et al., "An AI-Driven Interpretable Multiview Feature Learning Approach for EEG Based Epileptic Seizure Detection," *IEEE Journal of Biomedical and Health Informatics*,

- pp. 1–14, Jan. 2025, doi: 10.1109/jbhi.2025.3604638.
- [11] P. Mishra and H. Das, “An Explainable AI Framework Integrating Variational Sparse Autoencoder and Random Forest for EEG-Based Epilepsy Detection,” *IEEE Access*, vol. 13, pp. 179568–179583, Jan. 2025, doi: 10.1109/access.2025.3620762.
- [12] J. Wang, Y. Wang, W. Nie, and Q. Yuan, “Explainable End-to-End Seizure Prediction via Dynamic Multiscale Cross-Band Fusion Filter Network,” *International Journal of Neural Systems*, vol. 36, no. 5, pp. 2650010–2650010, Dec. 2025, doi: 10.1142/s0129065726500103.
- [13] M. K. Siddiqui, R. Morales-Menéndez, X. Huang, and N. Hussain, “A review of epileptic seizure detection using machine learning classifiers,” *Brain Informatics*, vol. 7, no. 1, pp. 5–5, May 2020, doi: 10.1186/s40708-020-00105-1.
- [14] Ahmad et al., “Robust Epileptic Seizure Detection Based on Biomedical Signals Using an Advanced Multi-View Deep Feature Learning Approach,” *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 10, pp. 5742–5754, May 2024, doi: 10.1109/jbhi.2024.3396130.
- [15] H. Kode, K. Elleithy, and L. Almazaydeh, “Epileptic Seizure Detection in EEG Signals Using Machine Learning and Deep Learning Techniques,” *IEEE Access*, vol. 12, pp. 80657–80668, Jan. 2024, doi: 10.1109/access.2024.3409581.
- [16] X. Tian et al., “Deep Multi-View Feature Learning for EEG-Based Epileptic Seizure Detection,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 10, pp. 1962–1972, Sep. 2019, doi: 10.1109/tnsre.2019.2940485.
- [17] Y. Gao, A. Liu, L. Wang, R. Qian, and X. Chen, “A Self-Interpretable Deep Learning Model for Seizure Prediction Using a Multi-Scale Prototypical Part Network,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1847–1856, Jan. 2023, doi: 10.1109/tnsre.2023.3260845.
- [18] R. J. Venkatesh, P. Sahay, and S. Shridevi, “Investigating Intra-Patient Seizure Heterogeneity via EEG Fingerprinting and Explainable AI,” *IEEE Access*, vol. 13, pp. 199170–199186, Jan. 2025, doi: 10.1109/access.2025.3636046.
- [19] S. Rukhsar and A. K. Tiwari, “Lightweight convolution transformer for cross-patient seizure detection in multi-channel EEG signals,” *Computer Methods and Programs in Biomedicine*, vol. 242, pp. 107856–107856, Oct. 2023, doi: 10.1016/j.cmpb.2023.107856.
- [20] M. F. Pinto, A. Leal, F. Lopes, A. Dourado, P. Martins, and C. Teixeira, “A personalized and evolutionary algorithm for interpretable EEG epilepsy seizure prediction,” *Scientific Reports*, vol. 11, no. 1, pp. 3415–3415, Feb. 2021, doi: 10.1038/s41598-021-82828-7.
- [21] Jemal, N. Mezghani, L. Abou-Abbas, and A. Mitiche, “An Interpretable Deep Learning Classifier for Epileptic Seizure Prediction Using EEG Data,” *IEEE Access*, vol. 10, pp. 60141–60150, Jan. 2022, doi: 10.1109/access.2022.3176367.
- [22] J. Germino et al., “Mixture of checkpoint experts for explainable seizure detection using wearable devices,” *Scientific Reports*, vol. 15, no. 1, pp. 43976–43976, Dec. 2025, doi: 10.1038/s41598-025-27610-9.
- [23] W. U. R. Qamar, M. Lee, and B. Abibullaev, “Deep learning in intracranial EEG for seizure detection: advances, challenges, and clinical applications,” *Frontiers in Neuroscience*, vol. 19, pp. 1677898–1677898, Oct. 2025, doi: 10.3389/fnins.2025.1677898.
- [24] S. Tan et al., “Automatic detection and prediction of epileptic EEG signals based on nonlinear dynamics and deep learning: a review,” *Frontiers in Neuroscience*, vol. 19, pp. 1630664–1630664, Aug. 2025, doi: 10.3389/fnins.2025.1630664.
- [25] M. Zabihi, E. J. Gilmore, K. Ding, and E. S. Rosenthal, “A transparent AI assurance and benchmarking framework for EEG seizure detection on TUSZ seeded with a reproducible gradient-boosting ensemble,” *Scientific Reports*, Feb. 2026, doi: 10.1038/s41598-026-41358-w.