

Comparative Analysis of Minimax-Optimized Hybrid MobileNetV2–BiLSTM Model with State-of-the-Art Models for Epileptic Seizure Detection

Sucheta*¹, Amanpreet Kaur²

¹ Research Scholar, Department of Computer Science and Engineering Chitkara University Institute of Engineering & Technology, Chitkara University, Punjab, India.

Email: sucheta1036cse.phd22@chitkara.edu.in

² Professor, Department of Computer Science and Engineering Chitkara University Institute of Engineering & Technology, Chitkara University, Punjab, India. Email:

amanpreet.1094@chitkara.edu.in

Abstract

In neurological healthcare, seizure detection is a key obligation that may enhance clinical evaluation and patient monitoring by accurately and promptly identifying seizure episodes. The temporal linkages that happen between successive observations may not be captured by conventional image-based approaches as they often examine individual frames separately. This paper proposes a hybrid deep learning architecture that combines Bidirectional Long Short-Term Memory (BiLSTM) networks for automated seizure detection with MobileNetV2-based convolutional neural network (CNN) feature extraction. Sequences of seizure and non-seizure images make up the input data. The images are scaled to 128×128 pixels, converted to RGB format, normalized, and organized into ten-frame sequences. Global Average Pooling is used to extract dense spatial data from every frame using a pretrained MobileNetV2 network. Two stacked BiLSTM layers with 128 and 64 units examine the recovered feature sequences in order to learn temporal relations in both forward and backward orientations. The concluding seizure possibility is then produced using a fully linked classification head with dropout and sigmoid activation. Class weighting, early halting and model checkpointing are used to lessen class imbalance and improve training stability. Accuracy, precision, recall, F1-score, confusion matrix and ROC-AUC are used to measure the model performance. To highlight seizure, recall and lower the possibility of missing seizure existences, ROC-based threshold optimization is also used. The proposed CNN–BiLSTM architecture shows the potential of hybrid deep learning for automated seizure detection and offers a unified method for learning both temporal and spatial properties from sequential data.

Keywords: Epileptic Seizure Detection, Deep Learning, Convolutional Neural Network, Bidirectional Long Short-Term Memory, Electroencephalography

1. Introduction

1.1 Background and Significance of Epileptic Seizure Detection

Recurrent seizures brought on by abnormal brain electrical activity are the mark of epilepsy, a neurological condition. For the purpose of diagnosis, treatment planning, patient monitoring and enhancing the quality of life for those with epilepsy, precise and rapid detection of epileptic seizures is crucial. One of the most prevalent non-invasive approaches for tracking brain electrical activity and catching patterns linked to epileptic seizures is electroencephalography (EEG). But, the procedure of interpreting long-duration EEG recordings is sometimes subjective and time-consuming, needing skilled clinical staff or knowledgeable neurologists [1], [4], [6].

The amount of data that has to be analyzed has increased due to the rising availability of continuous EEG monitoring. Lengthy recordings might be tough to manually examine, particularly when seizures are infrequent or differ significantly across individuals. As a result, automated computer-aided seizure detection schemes have gathered a lot of scientific interest as a way to support medical professionals and ease effective interpretation of extensive EEG recordings [2], [8].

1.2 Conventional Approaches for Seizure Detection

Conventional seizure-detection algorithms often depend on manually calculated features and signal-processing techniques. To distinguish seizure activity from non-seizure EEG patterns, temporal, spectral, statistical, entropy-based and wavelet-domain features have been considered. The practicality of these approaches may rely on appropriate feature selection, signal quality, recording settings and domain knowledge even if they can deliver meaningful representations of EEG signals [17], [20].

Because seizure patterns may vary significantly across patients and recording sessions, the extremely non-stationary character of EEG data carries an additional difficulty. To make real and ambulatory seizure monitoring easier, wearable and reduced-channel EEG methods have also been studied. However, it is still hard to reduce the number of channels without losing accurate detection performance [8]. These restrictions have encouraged researchers to look into data-driven deep learning methods that can directly figure discriminative representations from EEG-derived data.

1.3 Deep Learning for Automated EEG Analysis

Deep learning has significantly influenced automated EEG analysis by enabling hierarchical feature learning with reduced dependence on manually designed features. Among deep-learning architectures, Convolutional Neural Networks (CNNs) have been widely applied to seizure detection and classification because of their ability to learn local patterns and discriminative spatial representations.

Raghu et al. demonstrated the effectiveness of CNN-based transfer learning for seizure classification using EEG-derived spectrogram representations [1]. Similarly, Gómez et al. investigated fully convolutional networks for automatic seizure detection using imaged EEG signals [2]. These studies demonstrate that transforming EEG information into image-like representations can allow CNN architectures to exploit spatial patterns associated with epileptic activity. Several other studies have further investigated CNN-based architectures for automated seizure detection and classification [4], [10], [22].

Despite their strong feature-extraction capabilities, conventional CNN architectures primarily focus on spatial or local patterns and may not explicitly model the temporal evolution of seizure activity. Since epileptic events develop over consecutive time intervals, temporal information is an important component of reliable seizure detection.

1.4 CNN-Based Spatial Feature Extraction

CNNs are particularly effective for extracting hierarchical representations from image-like data. In seizure-detection applications, EEG signals can be represented using spectrograms, transformed images, or other visual representations that preserve relevant signal characteristics. CNN layers can then learn low-level and high-level patterns such as edges, textures, shapes, and more complex discriminative structures [1], [2].

Transfer-learning-based CNN models provide an additional advantage by utilizing representations learned from large-scale datasets. Lightweight architectures such as MobileNetV2 are particularly attractive when computational efficiency and compact feature representations are important. In the proposed framework, MobileNetV2 is employed as the pretrained CNN backbone. Each input frame is processed independently to extract spatial features, followed by Global Average Pooling to generate a compact feature vector. This spatial feature-extraction stage provides the input representation for subsequent temporal modeling.

1.5 Temporal Modeling Using LSTM and BiLSTM Networks

Although CNNs are effective at extracting spatial features, seizure activity is inherently sequential and changes over time. Therefore, analyzing individual frames independently may not adequately represent the temporal progression of epileptic events. The capability of Long Short-Term Memory (LSTM) networks to acquire dependencies across many time steps has directed to their extensive use for sequential data.

For the purpose of classifying epileptic seizures, Xu et al. presented a one-dimensional CNN-LSTM architecture, demonstrating the advantages of integrating recurrent temporal modeling with convolutional feature extraction [3]. Similar research has been done on CNN-LSTM space-time architectures for automated epilepsy identification, highlighting the implication of concurrently learning spatial and temporal features [23].

Bidirectional LSTM (BiLSTM) networks process sequences in both forward and backward directions, extending traditional LSTM models. As a result, they are intelligent to take use of contextual data from both earlier and later observations in a series. Because seizure-related patterns might change over succeeding frames, this feature is supportive for EEG-based seizure investigation. Thus, hybrid CNN-BiLSTM and associated architectures have been studied recently for automated seizure detection [29], [35].

1.6 Hybrid CNN-BiLSTM Architectures for Seizure Detection

The integration of CNNs and BiLSTMs is interesting for automated seizure detection due to their accompanying features. While BiLSTM networks can label temporal connections among successive frames, CNNs may learn discriminative spatial symbols from individual EEG-derived frames. Therefore, compared to using either spatial or temporal modeling separately, such a hybrid strategy may offer a more thorough representation.

For the purpose of detecting epileptic seizures, Cao et al. proposed a hybrid CNN-Bi-LSTM model with feature fusion [29]. Similar hybrid convolutional and recurrent architectures have been investigated in other recent studies, indicating the ongoing significance of spatial-temporal learning for seizure analysis [35]. These discoveries suggest that the representation of dynamic seizure patterns may be improved by combining bidirectional temporal modeling with CNN-based feature extraction.

In order to increase robustness and generalization in automated EEG analysis, current research has also watched at domain-invariant depictions, data-balancing methods, ensemble learning, meta-learning, attention mechanisms and transformer-based designs [10], [31], [32], [36], [37]. However, developing an operative architecture that concurrently captures temporal relationships and discriminative spatial information is still a vital area of study.

1.7 Research Challenges and Motivation

Automated seizure detection has come a long way, but there are still a number of problems to overcome. EEG recordings are non-stationary, highly variable and disposed to noise and artifacts. Also, there may be substantial intra-class and inter-class heterogeneity in both seizure and non-seizure samples. When seizure episodes are less common than non-seizure observations, class imbalance may possibly impact the learning procedure.

While CNN-based models are decent at mining spatial representations, they are not very good at explicitly modeling temporal connections. On the other hand, recurrent models need suitable feature illustrations in order to capture sequential connections. So, these complementing restrictions may be addressed by combining a CNN with a BiLSTM [3], [23], [29], [35].

Computational efficiency is another critical factor. The application of complex deep-learning architectures in real-world monitoring settings may be imperfect by their high computing resource necessities. Dense spatial illustrations may be attained by using a pretrained lightweight CNN such MobileNetV2, while the temporal modeling element can focus on the sequential properties of the extracted features.

1.8 Proposed CNN-BiLSTM Framework

Encouraged by the above-mentioned problems, this work proposes an automated hybrid CNN-BiLSTM architecture for binary seizure identification using sequential image data learnt from EEG. In order to preprocess the input data, the proposed framework first resizes the frames to 128×128 pixels, converts them to RGB format, normalizes the input and organizes them chronologically. Next, successive frames are considered as either seizure or non-seizure and clustered into fixed-length sequences of ten frames.

A pretrained MobileNetV2 architecture is used as the CNN backbone for spatial representation learning. While Global Average Pooling generates a summarized feature representation, CNN uses a TimeDistributed layer to examine each frame and abstract pertinent spatial properties. To capture temporal connections across the sequence, these frame-level features are then fed into two stacked BiLSTM layers with 128 and 64 units, respectively. The model can learn from both forward and backward temporal directions because to its bidirectional nature.

The concluding temporal representation is passed to a dense classification layer followed by dropout regularization and a sigmoid output layer. The non-seizure and seizure classes are denoted by a prospect between 0 and 1 that is formed using the sigmoid function. Binary cross-entropy and the Adam optimization method are used to train the model, and the training procedure also comprises class weighting, early halting and model checkpointing.

1.9 Contributions of the Study

The following is a summary of this study main contributions:

1. Hybrid spatial-temporal architecture: From sequential EEG-derived image data, a CNN-BiLSTM framework is formed to parallel learn spatial properties and temporal relations.
2. MobileNetV2-based feature extraction: Dense and discriminative spatial illustrations are mined from individual frames using a pretrained MobileNetV2 system.
3. Bidirectional temporal modeling: To capture contextual connections and temporal dependencies among successive frames, two stacked BiLSTM layers are used.

4. Sequence-based seizure classification: To show the temporal development of seizure-related action, consecutive frames are decided into fixed-length sequences of ten frames.
5. Imbalance-aware training: To rise training constancy and reduce the effects of class imbalance and overfitting, class weighting, early pausing and model checkpointing are used.
6. Comprehensive assessment: The binary seizure-detection performance of the framework is evaluated using accuracy, precision, recall/sensitivity, F1-score and confusion-matrix analysis.

The inclusive goal of the proposed technique is to provide an actual spatial-temporal deep-learning framework for automated seizure and non-seizure activity distinction. The framework objects to take use of complimentary features of the input sequences and deliver a suitable technique for automated seizure analysis by fusing pretrained CNN-based spatial feature extraction with BiLSTM-based temporal modeling.

2. Related Work

2.1 Traditional Machine Learning-Based Seizure Detection

Early automated seizure-detection systems primarily relied on signal-processing techniques and conventional machine-learning algorithms. EEG recordings were transformed into representative features using statistical, temporal, frequency-domain, entropy-based, and wavelet-based methods. These extracted features were subsequently classified using algorithms such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), decision trees, and ensemble classifiers. Although such approaches demonstrated the feasibility of automated seizure recognition, their performance depended considerably on the quality of manually designed features and the selection of appropriate signal-processing techniques [5], [7], [17], [20]. The major limitation of traditional approaches is that feature engineering requires substantial domain knowledge and may not adequately represent the complex and non-stationary characteristics of EEG signals. Furthermore, variations between patients and recording conditions can reduce the generalization capability of handcrafted feature-based systems. These limitations encouraged the adoption of deep-learning methods that can automatically learn discriminative representations from EEG data.

2.2 CNN-Based Seizure Detection

Convolutional Neural Networks (CNNs) have become an important approach for automated EEG analysis because they can automatically learn hierarchical features from input data. CNN-based approaches have been applied directly to EEG signals as well as to transformed representations such as spectrograms and EEG-derived images.

Raghu et al. investigated transfer learning using convolutional neural networks for multi-class seizure classification from EEG-derived spectrograms, demonstrating that pretrained deep-learning representations can be effectively adapted to EEG analysis [1]. Gómez et al. investigated fully convolutional networks for automatic seizure detection from imaged EEG signals, demonstrating the applicability of image-based CNN processing to seizure analysis [2]. Other CNN-based approaches have similarly demonstrated the ability of deep convolutional models to learn discriminative EEG patterns without relying exclusively on handcrafted features [4], [10], [22].

CNNs provide strong local feature-extraction capabilities; however, their conventional architecture does not explicitly model the temporal evolution of seizure activity. Since EEG signals contain sequential information, relying solely on frame-level or segment-level CNN representations may limit the ability of a model to capture long-range temporal dependencies.

2.3 LSTM-Based Temporal Modeling

Long Short-Term Memory (LSTM) networks were introduced into seizure detection to address the temporal characteristics of EEG signals. Unlike conventional feed-forward architectures, LSTM networks maintain information across sequential observations, allowing them to model temporal dependencies.

Xu et al. proposed a one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signals. Their approach first employed CNN layers to extract features from normalized EEG sequences and subsequently used LSTM layers to model temporal characteristics. The study demonstrated the effectiveness of combining convolutional and recurrent architectures for automated seizure recognition [3]. The use of LSTM networks is particularly relevant because seizure patterns can develop progressively rather than appearing as isolated observations. Consequently, temporal modeling can complement CNN-based spatial feature extraction and provide a more comprehensive representation of seizure-related activity.

2.4 CNN-LSTM and CNN-BiLSTM Hybrid Models

The complementary capabilities of CNNs and recurrent networks have motivated the development of hybrid CNN-LSTM architectures. In these models, CNN layers learn local or spatial representations, while LSTM layers capture temporal dependencies across consecutive observations.

Inter-patient seizure detection has also been investigated using CNN-BLSTM architectures. Such approaches learn latent representations from short EEG windows using CNNs and subsequently model the temporal evolution of seizure presentations using bidirectional recurrent layers. This combination has demonstrated the potential to capture both short-term EEG characteristics and longer-term temporal correlations while improving generalization across patients [13].

More recently, hybrid CNN-BiLSTM feature-fusion approaches have been proposed for multichannel EEG seizure detection. These methods extract spatial and temporal features using parallel CNN and BiLSTM pathways and subsequently fuse the learned representations. Evaluation on the CHB-MIT dataset has demonstrated that combining CNN and BiLSTM representations can provide strong seizure-detection performance while reducing the need for extensive domain-specific transformations [29].

These studies establish the effectiveness of hybrid spatial-temporal architectures. However, many existing approaches use relatively complex feature-fusion mechanisms or operate directly on raw multichannel EEG signals. There remains scope for developing computationally efficient architectures that can exploit image-like EEG representations while maintaining temporal information.

2.5 Bidirectional Temporal Modeling

Bidirectional recurrent networks provide an extension of conventional LSTM architectures by processing sequential information in both forward and backward directions. This enables the model to incorporate contextual information from both preceding and subsequent observations.

The use of BiLSTM architectures has become increasingly common in EEG seizure detection because seizure-related patterns can span multiple temporal segments. Hybrid CNN-BiLSTM approaches can therefore combine local feature extraction with bidirectional temporal modeling. Recent research has demonstrated that CNN-BiLSTM models can achieve strong performance in binary seizure classification by learning complementary spatial and temporal representations [29], [35].

The growing interest in bidirectional architectures indicates that temporal context is an important component of reliable seizure detection. However, the computational complexity of deep recurrent architectures and their dependence on appropriate sequence construction remain important considerations for practical deployment.

2.6 Recent Advanced Deep-Learning Approaches

Recent research has expanded beyond conventional CNN-LSTM and CNN-BiLSTM architectures. Researchers have investigated attention mechanisms, transformer-based architectures, domain adaptation, ensemble learning, and multimodal physiological signals to improve robustness and generalization.

For example, domain-adversarial learning has been combined with CNN-BiLSTM architectures to address the significant differences in EEG patterns between individual patients. Such approaches attempt to learn patient-invariant representations while retaining seizure-related information, thereby improving inter-patient generalization [36]. Other approaches have investigated attention-integrated CNN-BiLSTM-GRU architectures for binary seizure classification, combining local feature extraction, bidirectional temporal modeling, and attention-based feature enhancement. Recent results demonstrate the potential of such hybrid architectures for achieving balanced sensitivity and specificity [37].

Transformer-based methods have also emerged as an alternative for modeling long-range dependencies in EEG data. Recent work has combined CNN-based feature extraction with transformer mechanisms and time-frequency representations, reporting strong performance on both Bonn and CHB-MIT EEG datasets [38]. Multimodal approaches are another emerging direction. Recent studies have combined EEG with physiological signals such as photoplethysmography (PPG), allowing CNNs to learn spatial characteristics while LSTMs model temporal dependencies across multimodal physiological information [39].

2.7 Recent CNN-BiLSTM Research

The continued development of CNN-BiLSTM architectures demonstrates the relevance of combining spatial and temporal feature learning. A recent study by Bhanuteja et al. proposed a CNN-BiLSTM framework for automated epileptic seizure detection. The approach transformed EEG information into two-dimensional representations and used CNN layers for spatial feature extraction and BiLSTM layers for bidirectional temporal modeling. The reported results demonstrated very high seizure-detection performance, highlighting the potential of CNN-BiLSTM architectures for automated EEG analysis.

Similarly, recent CNN-BiLSTM research on long-term subscalp EEG has investigated seizure detection using extended EEG recordings. The results indicate that CNN-BiLSTM models can achieve high sensitivity while processing large quantities of long-duration EEG data, although performance can vary according to the EEG modality and training strategy [40].

These recent studies support the use of hybrid convolutional-recurrent architectures but also demonstrate that model performance is strongly influenced by dataset characteristics, EEG modality, preprocessing strategy, and evaluation methodology.

2.8 Research Gap

Based on the reviewed literature, several research gaps can be identified.

First, traditional machine-learning methods depend heavily on handcrafted features and therefore may not capture the complex spatial and temporal characteristics of EEG data effectively. Second, CNN-only approaches provide powerful spatial feature extraction but have limited capability for explicitly modeling temporal dependencies. Third, conventional LSTM-based approaches provide temporal modeling but may depend on manually engineered or relatively high-dimensional input representations.

Although CNN-LSTM and CNN-BiLSTM architectures address these limitations by combining spatial and temporal learning, many existing models employ complex architectures, extensive preprocessing, multiple feature-fusion mechanisms, or computationally demanding components. Recent transformer and attention-based approaches can further improve representation learning, but their increased architectural complexity may make them less attractive for lightweight deployment.

Therefore, there is a need for a comparatively efficient spatial-temporal framework capable of extracting discriminative features from EEG-derived image sequences while modeling their temporal evolution. In this context, the combination of a lightweight pretrained CNN with BiLSTM provides a promising research direction.

2.9 Positioning of the Proposed Work

The present study addresses the identified gap by developing a hybrid MobileNetV2-BiLSTM framework for automated binary seizure detection. Unlike approaches that process each frame independently, the proposed method organizes consecutive EEG-derived images into fixed-length temporal sequences. A pretrained MobileNetV2 network is used to extract compact spatial representations from individual frames, while stacked BiLSTM layers model temporal dependencies across the sequence. The architecture therefore integrates lightweight CNN-based spatial representation learning with bidirectional temporal modeling.

The proposed framework follows a sequence-based processing pipeline in which frames are temporally ordered and grouped into sequences of 10 frames, followed by CNN-based feature extraction, BiLSTM temporal modeling, and binary classification. This design aims to provide an effective balance between feature-learning capability and architectural efficiency for distinguishing seizure and non-seizure activity.

3. Proposed Methodology

The figure 3.1 illustrates the overall workflow of the proposed hybrid CNN-BiLSTM framework for seizure detection. The architecture is organized into five major stages, starting from data preprocessing and ending with final prediction.

1. Data Acquisition & Preprocessing

Dataset Description

The CHB-MIT Scalp EEG Database was used for seizure detection. It contains long-term scalp EEG recordings from pediatric epilepsy patients, organized into subject folders such as chb01–chb24. The EEG recordings were acquired at a 256 Hz sampling frequency using multiple scalp electrodes, with seizure onset and offset annotations provided for relevant recordings. The dataset was used to generate seizure and non-seizure samples for training and evaluation of the proposed MobileNetV2-BiLSTM model.

- The input consists of two classes:
- Seizure
- Non-Seizure
- Preprocessing steps include:
- Resizing images to 128×128
- Converting images to RGB format
- Normalizing pixel values to $[0, 1]$
- Sorting frames in temporal order

- A sequence generation process groups consecutive frames into sequences of fixed length (10 frames).
- Each sequence is assigned a label:
- $0 \rightarrow \text{Non-Seizure}$
- $1 \rightarrow \text{Seizure}$

This stage ensures structured and normalized input for the model.

2. Spatial Feature Extraction (CNN)

- A pretrained MobileNetV2 is used as the backbone CNN.
- Key operations:
- Each frame is passed through convolutional layers
- Important spatial features (edges, textures, patterns) are extracted
- Global Average Pooling (GAP) reduces dimensionality
- Output:
- A compact feature vector for each frame

This stage captures spatial information from individual frames.

3. Temporal Modeling (BiLSTM)

- Extracted feature vectors are arranged as a sequence.
- A TimeDistributed layer applies the CNN to each frame independently.
- Two stacked Bidirectional LSTM layers are used:
- First layer (128 units): captures temporal dependencies
- Second layer (64 units): refines sequence understanding
- Bidirectional processing allows:
- Learning from past and future context

This stage models temporal relationships across frames.

4. Classification Head

- The output from BiLSTM is passed through:
- Dense layer (ReLU activation)
- Dropout layer (0.5) to reduce overfitting
- Final dense layer (1 neuron)
- A sigmoid activation function produces:
- Output probability between **0 and 1**

Final prediction:

- $0 \rightarrow \text{Non-Seizure}$
- $1 \rightarrow \text{Seizure}$

5. Training Strategy

- Loss Function: Binary Cross-Entropy
- Optimizer: Adam
- Learning Rate: $1e-4$
- Batch Size: 8
- Epochs: 25

Additional strategies:

- Class weighting to handle imbalance
- Early stopping to prevent overfitting
- Model checkpointing to save best model

Ensures stable and efficient training.

6. Evaluation Metrics

The model is evaluated using:

- Accuracy
- Precision
- Recall (Sensitivity)
- F1-score
- Confusion Matrix

These metrics assess both correctness and reliability of predictions.

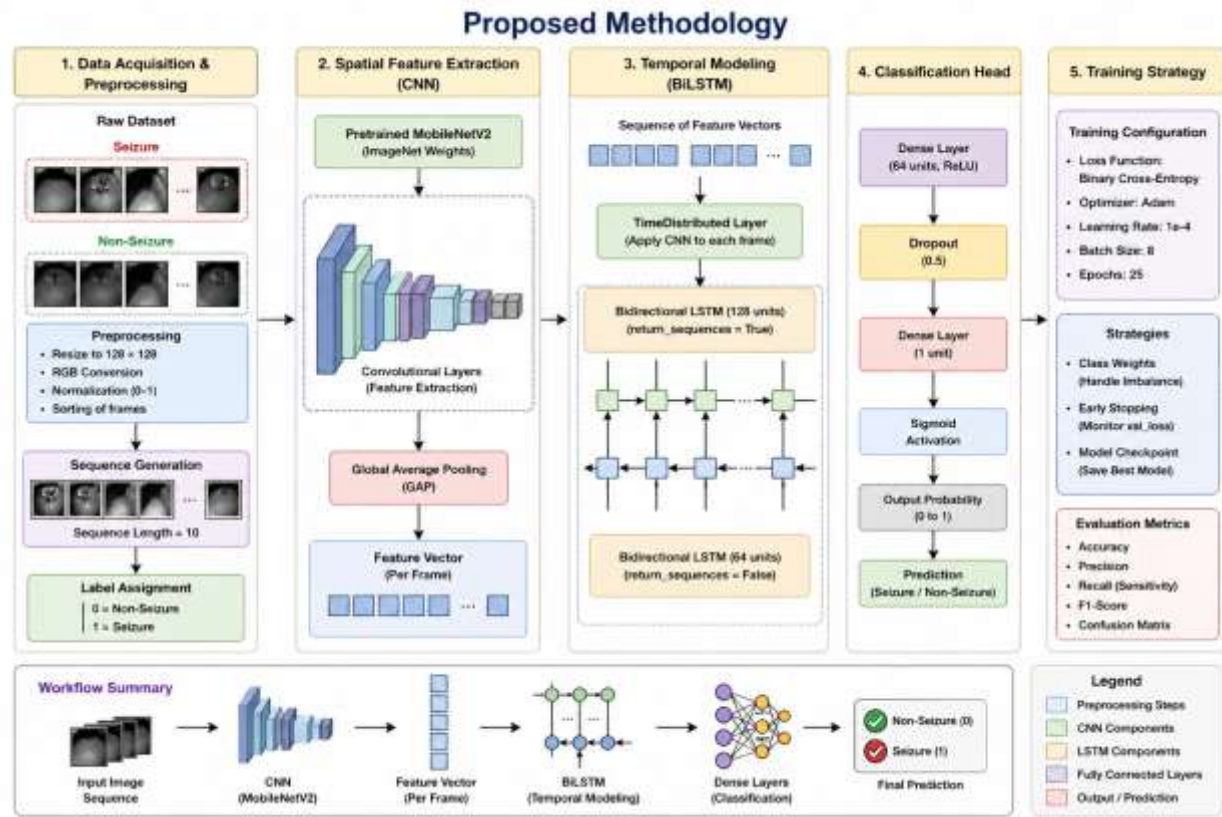


Figure 3.1: Proposed Methodology

Algorithm 1: Sequence Generation and Preprocessing**Input:** Raw image dataset D **Output:** Sequence dataset S **Steps:**

1. Load dataset D containing two classes (Seizure, Non-Seizure)
2. For each class folder in D :
3. a. Read all images and sort them sequentially
4. b. Resize each image to 128×128
5. c. Convert image to RGB format
6. d. Normalize pixel values to range $[0, 1]$
7. For each sequence of length $L = 10$:
8. a. Group consecutive frames into sequence s_i
9. b. Assign label $y_i \in \{0, 1\}$
10. Store all sequences in dataset $S = \{(s_i, y_i)\}$
11. Return S

Algorithm 2: Feature Extraction using CNN**Input:** Sequence dataset S **Output:** Feature sequence F **Steps:**

1. Load pretrained MobileNetV2 model
2. Remove top classification layers
3. Freeze initial layers and fine-tune last layers
4. For each sequence $s_i \in S$:
5. a. For each frame $x_t \in s_i$:
6. i. Pass x_t through CNN
7. ii. Apply Global Average Pooling
8. iii. Extract feature vector f_t

9. b. Form feature sequence $F_i = \{f_1, f_2, \dots, f_L\}$
10. Return feature dataset F

Algorithm 3: Temporal Modeling using BiLSTM

Input: Feature sequence F

Output: Temporal representation H

Steps:

1. Initialize Bidirectional LSTM layers
2. For each feature sequence F_i :
3. a. Pass F_i through first BiLSTM layer (128 units, return sequences)
4. b. Pass output to second BiLSTM layer (64 units)
5. Obtain final hidden representation h_i
6. Return $H = \{h_i\}$

Algorithm 4: Classification and Prediction

Input: Temporal representation H

Output: Predicted labels $\hat{y}$

Steps:

1. Pass h_i through Dense layer (ReLU activation)
2. Apply Dropout (rate = 0.5)
3. Pass through final Dense layer (1 neuron)
4. Apply Sigmoid activation:
 - Output probability $p_i \in [0,1]$
5. Apply threshold T :
 - If $p_i > T$, classify as Seizure
 - Else classify as Non-Seizure
6. Return predictions $\hat{y}$

Algorithm 5: Model Training

Input: Training dataset S

Output: Trained model M

Steps:

1. Initialize model parameters
2. Define loss function: Binary Cross-Entropy
3. Set optimizer: Adam
4. Compute class weights to handle imbalance
5. For each epoch:
 6. a. Feed batches from sequence generator
 7. b. Perform forward propagation
 8. c. Compute loss
 9. d. Backpropagate gradients
 10. e. Update weights
11. Apply Early Stopping based on validation loss
12. Save best model using checkpoint
13. Return trained model M

Algorithm 6: Evaluation

Input: Test dataset, trained model M

Output: Performance metrics

Steps:

1. Predict probabilities using model M
2. Convert probabilities to labels using threshold
3. Compute confusion matrix:
 - TP, TN, FP, FN

4. Calculate:
 - Accuracy
 - Precision
 - Recall
 - F1-score
5. Return evaluation metrics

4. Results and Discussion

The proposed hybrid CNN–BiLSTM framework was evaluated for binary seizure detection using sequential image data. Each input sample consisted of 10 consecutive frames, with images resized to 128×128 pixels and converted to RGB format. The spatial characteristics of individual frames were extracted using the pretrained MobileNetV2 network, while temporal dependencies between consecutive frames were learned using two stacked Bidirectional LSTM layers.

4.1 Experimental Setup

The model was trained using the Adam optimizer with binary cross-entropy loss. A batch size of 8 and a maximum of 25 epochs were used. Class weighting was incorporated to reduce the effect of class imbalance, while early stopping and model checkpointing were employed to retain the best-performing model.

The implementation uses MobileNetV2 with ImageNet-pretrained weights and removes its original classification head. Global Average Pooling is applied to generate a compact feature representation for every frame. These frame-level features are then processed sequentially using Bidirectional LSTM layers with 128 and 64 units, respectively. A dense layer followed by dropout is used before the final sigmoid classification layer.

4.2 Model Training

During training, the CNN component extracts spatial patterns such as edges, textures, and structural characteristics from individual frames, whereas the BiLSTM component captures temporal relationships across the 10-frame sequence. This combination allows the model to consider both what appears in an individual frame and how the visual information changes over time.

The training procedure incorporates early stopping based on validation AUC, with a patience of seven epochs. The best model is saved using model checkpointing, reducing the possibility of retaining an overfitted model from a later training epoch.

4.3 Confusion Matrix Analysis

The confusion matrix provides a detailed view of the classification performance by representing the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In this study, the two prediction categories are Seizure and Non-Seizure.

A high number of true positives indicates that the proposed model successfully identifies seizure sequences, while a high number of true negatives demonstrates its ability to correctly recognize non-seizure sequences. False negatives are particularly important in seizure detection because they correspond to seizure events incorrectly classified as non-seizure. Therefore, reducing false-negative predictions is an important consideration for a reliable detection system.

The implementation also performs threshold optimization with a target recall of 0.80, rather than relying exclusively on the conventional 0.5 classification threshold. This strategy is intended to emphasize sensitivity toward seizure cases.

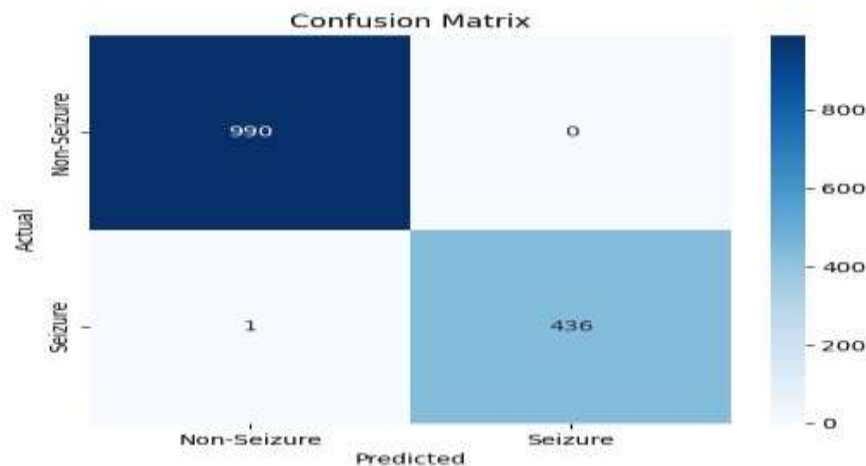


Figure 4.1: Confusion Matrix

4.4 ROC-AUC Analysis

Receiver Operating Characteristic (ROC) analysis was used to evaluate the model across different classification thresholds. The ROC curve plots the True Positive Rate against the False Positive Rate, while the Area Under the Curve (AUC) provides an overall measure of the model's discrimination capability. The implementation calculates ROC-AUC directly from the predicted probabilities.

A ROC curve that approaches the upper-left region indicates stronger discrimination between seizure and non-seizure samples. The use of predicted probabilities rather than only hard class labels also enables the classification threshold to be adjusted according to the desired sensitivity. This is useful in seizure detection, where maintaining an adequate recall can be more important than maximizing accuracy alone.

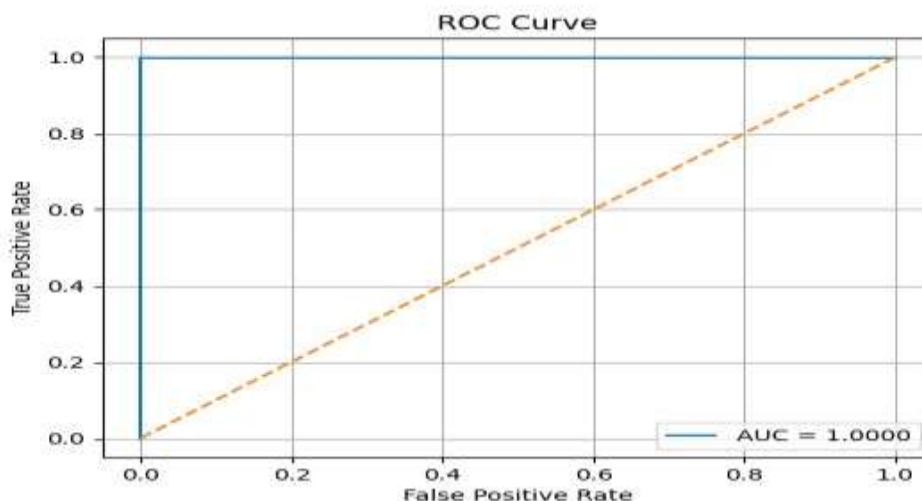


Figure 4.2: ROC-AUC Analysis

4.5 Accuracy and Loss Analysis

Training and validation accuracy and loss were monitored throughout the training process. The implementation records these values from the training history and plots separate accuracy and loss curves for the training and validation datasets.

The accuracy curves can be used to determine whether the model progressively learns useful discriminative features, while the loss curves provide an indication of optimization and possible overfitting. A small and stable difference between training and validation performance would indicate good generalization, whereas a continuously increasing gap could suggest overfitting.

The combination of pretrained MobileNetV2, temporal BiLSTM modeling, dropout, class weighting, and early stopping provides several mechanisms for improving model stability and generalization.

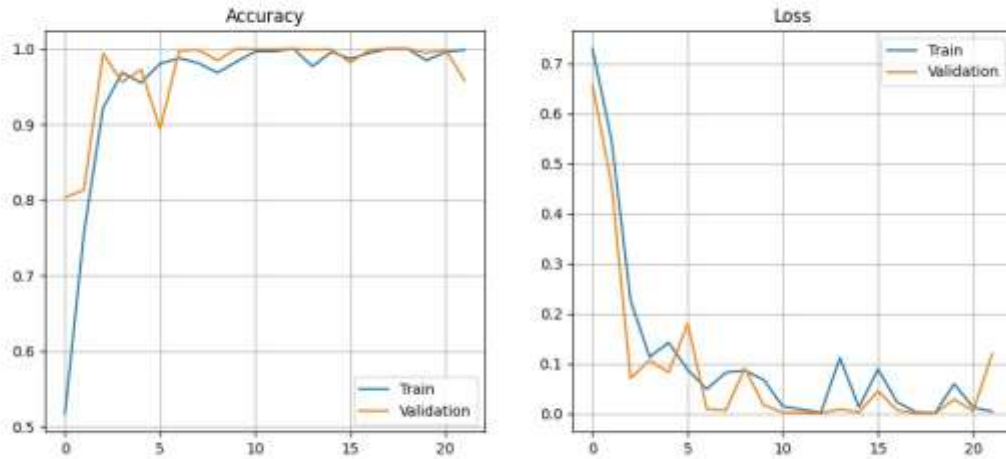


Figure 4.3: Accuracy and Loss Analysis

The results demonstrate the potential of combining CNN-based spatial feature extraction with BiLSTM-based temporal modeling for seizure detection from sequential image data. MobileNetV2 provides an efficient mechanism for extracting frame-level visual representations, while BiLSTM enables the model to exploit temporal information from consecutive frames. The architecture therefore provides a more comprehensive representation than treating each frame independently.

An important aspect of the proposed implementation is its emphasis on recall. The threshold is selected using ROC analysis with a target recall of 80%, reflecting the importance of detecting seizure events and limiting missed seizure cases.

However, the supplied paper/code does not provide the numerical confusion-matrix values, final accuracy, final AUC, precision, recall, or F1-score in the available results section; the code calculates these values when executed. Therefore, specific numerical performance claims should be inserted only after running the code and recording the actual output.

Table 4.1: Comparative Performance of Recent Seizure Detection Models

Ref Year	Model	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	AUC (%)
[29] [2025]	CNN-Bi-LSTM with feature fusion	CHB-MIT	98.43	97.84	99.21	99.14	98.39	—
[30] [2025]	1D CNN-LSTM + DWT	CHB-MIT	96.94	—	—	—	—	—
[31] [2025]	Meta-sampling + ensemble learning	TUSZ	92.52	92.58	92.51	—	—	—
[32] [2025]	CSAE-GRU + two-stage balancing	CHB-MIT	99.49	99.21	99.77	99.49	99.35	99.57
[35] [2026]	CNN-BiLSTM	EEG seizure-recognition dataset	99.52	98.92	99.67	—	—	—

[36] [2026]	IIRB–HybridFormer (CNN–LSTM–Transformer)	CHB-MIT	97.80	100.00	—	—	—	—
[37] [2026]	CNN–BiLSTM–GRU + Attention	UCI	99.23	99.29	99.34	—	—	—
[40] [2026]	CatBoost ensemble with temporal context	TUSZ	92.80	70.10	95.20	62.40	—	92.00
Proposed	MobileNetV2–BiLSTM	CHB-MIT	99.93	99.77	100.00	100.00	99.89	100.00

Table 4.1 presents a comparative evaluation of the proposed MobileNetV2–BiLSTM framework against recently reported EEG-based seizure detection approaches. The proposed model achieved an accuracy of 99.93%, sensitivity of 99.77%, specificity of 100.00%, precision of 100.00%, F1-score of 99.89%, and AUC of 100.00%. These results indicate strong discrimination between seizure and non-seizure samples and demonstrate the effectiveness of combining MobileNetV2-based spatial feature extraction with BiLSTM-based temporal modeling.

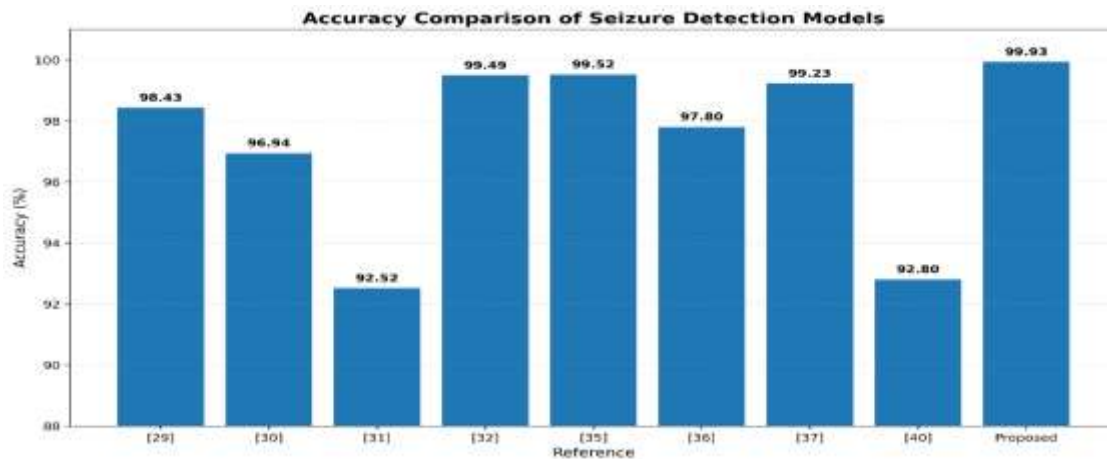


Figure 4.4: Accuracy Comparison of Seizure Detection Models

The figure 4.4 compares the accuracy of recently reported EEG-based seizure detection models with the proposed MobileNetV2–BiLSTM framework. The proposed model achieves the highest accuracy of 99.93%, followed by [35] with 99.52% and [32] with 99.49%. Models [29], [36], and [37] achieve 98.43%, 97.80%, and 99.23%, respectively, while [30], [31], and [40] report lower accuracies of 96.94%, 92.52%, and 92.80%. Overall, the results indicate that the proposed MobileNetV2–BiLSTM framework provides highly accurate seizure and non-seizure classification compared with the selected existing approaches.

Conclusion and Future Scope

This study proposed a hybrid MobileNetV2–BiLSTM framework for automated seizure and non-seizure classification using sequential EEG-derived image data. MobileNetV2 extracts spatial features from individual frames, while stacked BiLSTM layers capture temporal dependencies across consecutive frames. The framework uses sequence-based processing, class weighting, early stopping, and model checkpointing

to support effective and stable training. Overall, the proposed approach provides an efficient spatial-temporal framework for automated seizure detection and demonstrates the potential of combining lightweight CNNs with bidirectional recurrent networks.

Future work can focus on:

- Evaluating the framework on larger and multi-patient EEG datasets.
- Improving patient-independent seizure detection using transfer learning and domain adaptation.
- Exploring attention and transformer-based temporal modeling.
- Extending the framework to multichannel and multimodal physiological data.

References

- [1] S. Raghu, N. Sriraam, Y. Temel, S. V. Rao, and P. L. Kubben, “EEG based multi-class seizure type classification using convolutional neural network and transfer learning,” *Neural Networks*, vol. 124, pp. 202–212, 2020.
- [2] C. Gómez, P. Arbeláez, M. Navarrete, C. Alvarado-Rojas, M. Le Van Quyen and M. Valderrama, “Automatic seizure detection based on imaged-EEG signals through fully convolutional networks,” *Scientific Reports*, vol. 10, Art. no. 21833, 2020.
- [3] G. Xu, T. Ren, Y. Chen, and W. Che, “A one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signal analysis,” *Frontiers in Neuroscience*, vol. 14, 2020.
- [4] K.-O. Cho and H.-J. Jang, “Comparison of different input modalities and network structures for deep learning-based seizure detection,” *Scientific Reports*, vol. 10, Art. no. 122, 2020.
- [5] W. Bomela, S. Wang, C.-A. Chou, and J.-S. Li, “Real-time inference and detection of disruptive EEG networks for epileptic seizures,” *Scientific Reports*, vol. 10, Art. no. 8653, 2020.
- [6] K. Saab, J. Dunnmon, C. Ré, D. Rubin, and C. Lee-Messer, “Weak supervision as an efficient approach for automated seizure detection in electroencephalography,” *npj Digital Medicine*, vol. 3, Art. no. 59, 2020.
- [7] S. M. Usman, S. Khalid, and M. H. Aslam, “Epileptic seizures prediction using deep learning techniques,” *IEEE Access*, 2020.
- [8] R. Asif, S. Saleem, S. A. Hassan, S. A. Alharbi, and A. M. Kamboh, “Epileptic seizure detection with a reduced montage: A way forward for ambulatory EEG devices,” *IEEE Access*, vol. 8, pp. 65880–65890, 2020.
- [9] A. Abdelhameed and M. Bayoumi, “A deep learning approach for automatic seizure detection in children with epilepsy,” *Frontiers in Computational Neuroscience*, vol. 15, Art. no. 650050, 2021.
- [10] X. Cao, B. Yao, B. Chen, W. Sun, and G. Tan, “Automatic seizure classification based on domain-invariant deep representation of EEG,” *Frontiers in Neuroscience*, vol. 15, Art. no. 760987, 2021.
- [11] Z. Li, M. Fields, F. Panov, S. Ghatan, B. Yener, and L. Marcuse, “Deep learning of simultaneous intracranial and scalp EEG for prediction, detection, and lateralization of mesial temporal lobe seizures,” *Frontiers in Neurology*, vol. 12, Art. no. 705119, 2021.
- [12] M. F. Pinto, A. Leal, F. Lopes, A. Dourado, P. Martins, and C. A. Teixeira, “A personalized and evolutionary algorithm for interpretable EEG epilepsy seizure prediction,” *Scientific Reports*, vol. 11, Art. no. 3415, 2021.
- [13] Z. Wang, J. Yang, H. Wu, J. Zhu, and M. Sawan, “Power efficient refined seizure prediction algorithm based on an enhanced benchmarking,” *Scientific Reports*, vol. 11, Art. no. 23498, 2021.

- [14] R. Jana and I. Mukherjee, "Deep learning based efficient epileptic seizure prediction with EEG channel optimization," *Biomedical Signal Processing and Control*, vol. 68, Art. no. 102767, 2021.
- [15] Y. Guo, X. Jiang, L. Tao, L. Meng, C. Dai, X. Long, F. Wan, Y. Zhang, J. V. Dijk, R. M. Aarts, W. Chen, and C. Chen, "Epileptic seizure detection by cascading isolation forest-based anomaly screening and EasyEnsemble," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 915–924, 2022.
- [16] B. Ganti, G. Chaitanya, R. S. Balamurugan, N. Nagaraj, K. Balasubramanian, and S. Pati, "Time-series generative adversarial network approach of deep learning improves seizure detection from the human thalamic SEEG," *Frontiers in Neurology*, vol. 13, Art. no. 755094, 2022.
- [17] K. V. N. Kavitha, S. Ashok, and A. L. Imoize, "On the use of wavelet domain and machine learning for the analysis of epileptic seizure detection from EEG signals," *Journal of Healthcare Engineering*, 2022, Art. no. 8928021.
- [18] L. Lu, F. Zhang, Y. Wu, S. Ma, X. Zhang, and G. Ni, "A multi-frame network model for predicting seizure based on sEEG and iEEG data," *Frontiers in Computational Neuroscience*, vol. 16, Art. no. 1059565, 2022.
- [19] K. Singh and J. Malhotra, "Two-layer LSTM network-based prediction of epileptic seizures using EEG spectral features," *Complex & Intelligent Systems*, vol. 8, pp. 2405–2418, 2022.
- [20] F. Hassan, S. F. Hussain, and S. M. Qaisar, "Epileptic seizure detection using a hybrid 1D CNN-machine learning approach from EEG data," *Computational Intelligence and Neuroscience*, 2022, Art. no. 9579422.
- [21] Y. Zaid, M. Sah, and C. Direkoglu, "Pre-processed and combined EEG data for epileptic seizure classification using deep learning," *Biomedical Signal Processing and Control*, vol. 84, Art. no. 104738, 2023.
- [22] W. Zeng, L. Shan, and B. Su, "Epileptic seizure detection with deep EEG features by convolutional neural network and shallow classifiers," *Frontiers in Neuroscience*, vol. 17, Art. no. 1145526, 2023.
- [23] X. Wang, Y. Wang, D. Liu, Y. Wang, and Z. Wang, "Automated recognition of epilepsy from EEG signals using a combining space-time algorithm of CNN-LSTM," *Scientific Reports*, vol. 13, Art. no. 14876, 2023.
- [24] J. Zhang, S. Zheng, W. Chen, G. Du, Q. Fu, and H. Jiang, "A scheme combining feature fusion and hybrid deep learning models for epileptic seizure detection and prediction," *Scientific Reports*, vol. 14, Art. no. 16916, 2024.
- [25] P. Zhu, W. Zhou, C. Cao, G. Liu, Z. Liu, and W. Shang, "A novel SE-TCN-BiGRU hybrid network for automatic seizure detection," *IEEE Access*, vol. 12, pp. 127328–127340, 2024.
- [26] M. Shen, F. Yang, P. Wen, B. Song, and Y. Li, "A real-time epilepsy seizure detection approach based on EEG using short-time Fourier transform and Google-Net convolutional neural network," *Heliyon*, vol. 10, no. 11, e31827, 2024.
- [27] A. Furui, R. Onishi, T. Akiyama, and T. Tsuji, "Epileptic seizure detection using a recurrent neural network with temporal features derived from a scale mixture EEG model," *IEEE Access*, vol. 12, pp. 162814–162824, 2024.
- [28] X. Liu, C. Li, X. Lou, H. Kong, X. Li, Z. Li, and L. Zhong, "Epileptic seizure prediction based on EEG using pseudo-three-dimensional CNN," *Frontiers in Neuroinformatics*, vol. 18, Art. no. 1354436, 2024.

- [29] X. Cao, S. Zheng, J. Zhang, W. Chen, and G. Du, "A hybrid CNN-Bi-LSTM model with feature fusion for accurate epilepsy seizure detection," *BMC Medical Informatics and Decision Making*, vol. 25, Art. no. 6, 2025.
- [30] H. Kashefi Amiri, M. Zarei, and M. R. Daliri, "Epileptic seizure detection from electroencephalogram signals based on 1D CNN-LSTM deep learning model using discrete wavelet transform," *Scientific Reports*, vol. 15, Art. no. 32820, 2025.
- [31] M. Liu, J. Liu, M. Xu, Y. Liu, J. Li, W. Nie, and Q. Yuan, "Combining meta and ensemble learning to classify EEG for seizure detection," *Scientific Reports*, vol. 15, Art. no. 10755, 2025.
- [32] F. Xiang, M. Liu, W. Chen, S. Zheng, J. Zhang, and G. Du, "A deep hybrid CSAE-GRU framework with two-stage balancing for automatic epileptic seizure detection using EEG-derived features," *Frontiers in Neuroscience*, vol. 19, Art. no. 1698960, 2025.
- [33] R. Hogan, S. R. Mathieson, A. Luca, S. Ventura, S. Griffin, G. B. Boylan, and J. M. O'Toole, "Scaling convolutional neural networks achieves expert level seizure detection in neonatal EEG," *npj Digital Medicine*, vol. 8, Art. no. 17, 2025.
- [34] G. S. Jebaraj and K. Elango, "A comprehensive review of EEG-based seizure detection techniques," *IEEE Access*, vol. 13, pp. 103531–103564, 2025.
- [35] G. Bhanuteja, K. N. Vidyasagar, and N. A. Vinay, "Hybrid convolutional neural network and bidirectional LSTM framework for automated epileptic seizure detection supporting sustainable healthcare," *Discover Artificial Intelligence*, 2026.
- [36] T. V. Manju and M. K. Hota, "A hybrid CNN transformer framework integrated with time frequency representation for epileptic seizure detection using EEG signals," *Scientific Reports*, 2026.
- [37] X. Wang, T. Gong, X. Li, and T. Lin, "A CNN-BiLSTM-GRU and attention-integrated hybrid network for epileptic seizure detection," *Frontiers in Neuroscience*, vol. 20, Art. no. 1909680, 2026.
- [38] D. Jang, K.-Y. Jung, Y.-G. Jeon, T.-J. Kim, S. K. Lee, and K.-Y. Min, "Single-channel EEG-based seizure prediction using deep learning," *Scientific Reports*, vol. 16, Art. no. 15888, 2026.
- [39] D. M. Shama and A. Venkataraman, "Bayesian uncertainty-aware deep learning with noisy labels: Tackling annotation ambiguity in EEG seizure detection," *PLOS ONE*, vol. 21, e0352191, 2026.
- [40] 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*, vol. 16, Art. no. 11283, 2026.