Computers and Artificial Intelligence；Vol.3, No.3; 2026
E-ISSN: 3006 -3728 P-ISSN：3006-371X
DOI: https://doi.org/10.70267/cai.26v3n3.246257
Published by: Zeus Press

Automatic Detection of Epileptic EEG Signals Based on Channel Attention and Bidirectional Temporal Modeling

Xien Gao*

School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

**Corresponding author: Xien Gao.*

Abstract

Epilepsy is a common chronic neurological disease. As an important tool for epilepsy diagnosis and disease assessment, electroencephalogram (EEG) can reflect the abnormal discharge activity of brain neurons. However, traditional EEG interpretation relies heavily on expert manual analysis, which has problems such as time-consuming, strong subjectivity and low efficiency. To improve the automation level of seizure detection, this paper proposes a pyramid-type one-dimensional convolutional neural network model (SE-P1D-LSTM) that integrates a channel attention mechanism and a Bidirectional Long Short-Term Memory (BiLSTM) network. The method first preprocesses the EEG signal through overlapping sliding window and Z-Score standardization, and combines Gaussian noise disturbance and random amplitude scaling for data augmentation; Then use the pyramidal one-dimensional convolutional structure to extract local temporal features, and adaptively strengthen the key channel information through SEBlock; Finally, a BiLSTM is introduced to model the long-range temporal dependencies in EEG signals to realize the automatic classification of normal and seizure EEG signals. The experiment was carried out based on the public epilepsy EEG dataset of the University of Bonn in Germany, and the performance of the model was evaluated using 10-fold cross-verification. The results show that the method in this paper has achieved an accuracy rate of 99.87%, a sensitivity of 99.87% and a specificity of 99.87% in the two-classification task, which is better than a variety of comparative models. The research results show that SE-P1D-LSTM can effectively extract the discriminative characteristics of epilepsy EEG signals, and has good classification performance and application potential in the automatic detection task of seizures.

Keywords

epileptic seizure detection, EEG signals, deep learning, channel attention, bidirectional long short-term memory (BiLSTM), one-dimensional convolutional neural network (1D-CNN)

1. Introduction

Epilepsy is one of the most common neurological diseases in the world, affecting about 50 million people of all ages [1]. As a chronic non-communicable disease, epilepsy is characterized by recurrent and unprovoked seizures caused by abnormal excessive or synchronous neuronal activity in the brain. The impact of seizures is not only limited to the seizure itself, but also leads to cognitive impairment, psychosocial difficulties and a decline in the quality of life. For about one-third of patients with epilepsy who develop drug resistance, the situation is more serious because drugs are difficult to effectively control seizures. After German psychiatrist Hans Berger first discovered electroencephalogram in 1924, electroencephalogram has been the most important tool in the diagnosis and management of epilepsy [2]. Through electrodes placed on the surface of the scalp (scalp electroencephalogram) or electrodes implanted inside the brain (intracranial electroencephalogram), EEG can record the electrical activity of the brain in real time and provide information about neuronal function. In clinical practice, EEG plays an irreplaceable role in detecting epileptic discharge and evaluating the treatment effect.

However, traditional EEG analysis relies heavily on the visual judgment of experienced neurologists. This process is not only extremely time-consuming, but also easily affected by subjective factors, resulting in poor consistency among observers. In clinical practice, a single long-term EEG recording can produce dozens or even hundreds of hours of data, forcing doctors to spend a lot of time on page-by-page review, seriously crowding out the originally limited medical resources. With the continuous growth of the number of epilepsy patients in the world and the serious shortage of professional EEG specialists, the development of an automated, reliable and objective epilepsy detection system has become an urgent clinical need.

The automatic detection of seizures from EEG signals faces challenges. First of all, EEG signals are inherently complex and unstable, and are vulnerable to various artifacts and noise, including eye movement, myoelectric activity and electrode contact noise. Secondly, the seizure pattern shows significant heterogeneity between different patients and different seizures of the same patient, which makes it extremely difficult to define general detection standards. Third, the imbalance between interictal and ictal data poses a challenge to the construction of a balanced classification model. Despite these challenges, the automatic epilepsy detection system has great clinical value. They can continuously monitor patients in long-range EEG records, alert medical personnel to ongoing seizures, reduce the time required for manual review, and provide objective quantitative evaluation to support clinical decision-making. In addition, this kind of system is the basis for the realization of closed-loop treatment equipment, which can provide timely intervention when seizures are detected.

In order to meet the above challenges, this paper proposes an automatic seizure detection system based on deep learning. The system builds a pyramid-type one-dimensional convolutional network (SE-P1D-LSTM) that integrates channel attention and two-way LSTM. It extracts and compresses spatiotemporal features of EEG signals through the decreasing channel structure layer by layer. It combines SEBlock to adaptively strengthen key channel information, employs a BiLSTM to capture long-range temporal dependencies, and reduces the risk of model parameters and over-fitting while ensuring detection accuracy. In response to the limited number of brain electrical data samples, this paper adopts overlapping sliding windows, online Gaussian noise disturbance and random amplitude scaling for data augmentation to improve the robustness and generalization ability of the model. The training stage combines AdamW, weight attenuation, gradient cropping, learning rate adaptive adjustment and early stopping mechanism, and evaluates the model performance through the 10-fold cross-verification system to ensure the stability and reliability of the experimental results. The following part of this article is organized as follows: Chapter II reviews the relevant work of epilepsy detection; Chapter 3 describes in detail the proposed methods, including P-1D-CNN architecture, data augmentation strategy and integrated decision-making integration; Chapter 4 shows the experimental results and conducts comparative analysis; Chapter 5 discusses the significance and limitations of this work; Chapter 6 summarizes the full text and looks forward to the future research direction.

2. Related Work

Traditional methods usually follow the serial process of preprocessing, feature extraction, feature selection and classification. Researchers have explored a variety of feature extraction techniques, including time domain characteristics (such as amplitude and curve length), frequency domain characteristics (such as power spectrum density) and time-frequency domain representation (such as wavelet transformation) [3]. The extracted features are then input into classifiers, such as support vector machine, K nearest neighbor, random forest or artificial neural network. A large number of studies have verified the effectiveness of these methods. For example, Subasi and Erçelebi [4] combine the multi-scale frequency band energy characteristics extracted by wavelet transformation with artificial neural networks and logistic regression, which significantly improves the accuracy of the distinction between the seizure phase and the inter-attack period, and the classification accuracy reaches 90%-95%; Guo et al. [5] built a lightweight classifier based on the time domain characteristics of the Line Length of the EEG signal, which verified the high efficiency of simple time domain statistics in low computing power scenarios, and achieved an accuracy rate of 96.3% in the classification of EEG signals; Alotaiby et al. [6] combined the multi-dimensional manual characteristics of the time domain and frequency domain to compare a variety of machine learning and neural network algorithms to complete the electroencephalographic classification of epilepsy. Experimental results show that the integrated model can adapt well to the high-dimensional nonlinear characteristics of EEG, and the sensitivity and specificity of its classification tasks can exceed 90%. However, the traditional method relies heavily on the characteristics of artificial design, and it is difficult for shallow classifiers to fully portray the non-steady and high-dimensional space-time coupling characteristics of EEG signals, prompting researchers to turn to deep learning methods with more expressive ability.

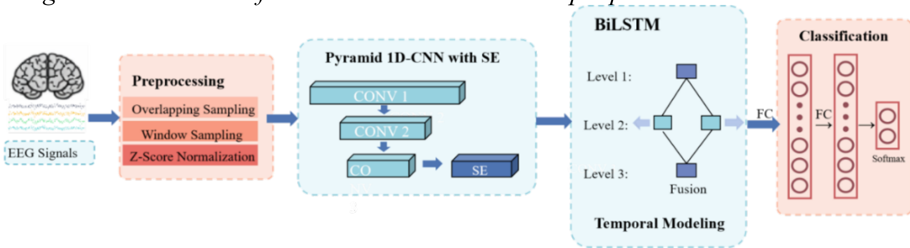
Deep learning has become a research hotspot in the field of epilepsy detection because of its strong automatic feature learning ability. Deep learning can automatically learn hierarchical feature representations directly from the original or minimally preprocessed EEG data. Among them, convolutional neural networks are good at capturing local space-time patterns and have been successfully applied to time series EEG and spectrum images; Recurrent neural networks, especially variants with long-term and short-term memory units, are very suitable for modeling the time dependence inherent in EEG signals; In recent years, the Transformer-based architecture has captured the long-range dependence of the whole sequence through the self-attention mechanism, showing superior performance. For example, the deep one-dimensional convolutional network built by Acharya et al. [7] reached an accuracy rate of 99.03% in the three classification tasks of the University of Bonn dataset; Tsiouris et al. [8] introduced LSTM into long-range electroencephalogram analysis and developed a long-term and short-term memory deep learning network, which effectively improved the robustness of seizure timing modeling and effectively improved the accuracy of EEG signal classification tasks to 92%~96%; In addition, Rukhsar et al. [9] proposed a lightweight convolutional Transformer (LCT) model to carry out cross-patient epilepsy detection based on the CHB-MIT multi-channel scalp EEG public dataset, and achieved an average classification accuracy of 95.1%~96.7% in the multi-test subject scalp electroencephalogram test.

Despite the above progress, there are still some key limitations in the existing research. First, most deep learning models only complete training and testing on a single dataset, and the dataset adopts a standardized acquisition protocol. Such experimental design is difficult to verify the generalization ability of the model across recording devices, cross-electrode lead configuration and cross-patient groups. Research shows that when such models are applied to unseen data from different medical institutions, the detection performance often declines significantly. Second, deep learning models usually rely on large-scale annotation data, and expert-level EEG annotation is time-consuming and requires the support of professional knowledge, so it is difficult to obtain, resulting in the scarcity of available data. At the same time, the natural imbalance in quantity between the seizure sample and the inter-attack sample further aggravates the difficulty of model training.

3. Epileptic Seizure Detection Method Based on SE-P1D-LSTM

In response to the problem that the traditional method has limited feature extraction ability mentioned above and the existing deep learning models are prone to overfitting, this paper proposes a pyramidal one-dimensional convolutional neural network (Squeeze-and-Excitation Pyramidal 1D CNN with LSTM, SE-P1D-LSTM) that integrates a channel attention mechanism and a Bidirectional Long Short-Term Memory (BiLSTM) network. The overall process of this method is shown in Figure 3-1: First, the raw EEG signals are segmented using overlapping sliding windows and standardized preprocessing; Then use the pyramidal convolutional layer to extract local space-time features, and SEBlock is employed to adaptively recalibrate feature weights; Then use bidirectional LSTM to capture the long-term dependence of EEG signals; finally, the probability of seizures is output through the fully connected layer.

Figure 1: Flowchart of the SE-P1D-LSTM-Based Epileptic Seizure Detection Method



3.1 Normalization and Data Augmentation

During acquisition, clinical EEG signals are susceptible to interference from factors such as eye movements, electromyographic (EMG) activity, and environmental noise; additionally, EEG amplitudes vary among individuals. At the same time, the number of epileptic seizure samples is typically limited; using them directly for model training can easily result in poor model generalization. To mitigate the impact of signal amplitude variations on model training and to expand the training samples to some extent, this paper subjects the raw EEG signals to overlapping window sampling, Z-score normalization, and online data augmentation. Let the raw EEG signal sequence be:

$$X = \{x_1, x_2, \dots, x_N\} \tag{1}$$

In the formula, X represents the raw EEG signal sequence, x_i denotes the amplitude at the i-th sampling point, and N represents the total length of the signal. The length of each individual EEG signal used in this paper is 4097. To meet the input requirements of the neural network model while increasing the number of training samples, this paper employs an overlapping sliding window to segment the raw signals. The window length is set to L and the sliding step size to S, meaning the overlap rate between adjacent windows is 50%. The k-th sample obtained after segmentation can be expressed as:

$$X_k = \{x_{kS+1}, x_{kS+2}, \dots, x_{kS+L}\}, k = 0,1, \dots, M-1 \tag{2}$$

In Equation (2), X_k represents the k-th sample obtained after segmentation; L is the window length, S is the sliding step size, and M denotes the total number of samples generated from a single raw signal. The number of samples that can be generated from a single raw signal after segmentation using a sliding window is calculated as follows:

$$M = \left\lfloor \frac{N-L}{S} \right\rfloor + 1 \tag{3}$$

Here, $\lfloor \cdot \rfloor$ denotes the floor operation. Based on the signal length, window length, and sliding step size set in this study, each raw EEG signal can be segmented into multiple fixed-length samples, thereby providing more input data for subsequent model training. This method allows for the expansion of the sample size while preserving the local temporal information of EEG signals, thereby mitigating the problem of model overfitting associated with small sample sizes. Given the potential differences in amplitude scale among different EEG signal samples, this paper further applies Z-score standardization to each sample. For a single sample, the standardization process is defined as:

$$\hat{x}_i = \frac{x_i - \mu}{\sigma + \varepsilon}, \mu = \frac{1}{L}\sum_{i=1}^{L} x_i, \sigma = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(x_i - \mu)^2} \tag{4}$$

In the formula, μ represents the mean amplitude of the current sample, σ denotes the sample standard deviation, L is the sample length, and ε is a small constant introduced to prevent division by zero; in this paper, ε is set to10^{-8}. After this processing, the mean of each sample approaches 0 and the variance approaches 1; this helps mitigate the impact of individual differences and variations in data acquisition conditions on model training, while also enhancing the stability of the network training process.

In addition to standardization, this paper employs online data augmentation during the training phase to enhance the model's robustness against noise and amplitude variations. Specifically, in each training batch, Gaussian noise is added to the input EEG signals, and slight random amplitude scaling is applied. The augmented signals are defined as:

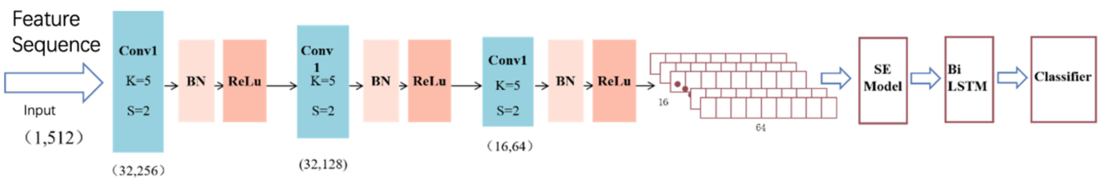
$$\tilde{X} = \alpha(X + \eta), \eta \sim \mathcal{N}(0, \sigma_n^2), \alpha \sim U(a, b) \tag{5}$$

In Equation (5), $\tilde{X}$ represents the augmented EEG signal, X represents the original input signal, η represents the added random noise, and α represents the random amplitude scaling factor. The random noise follows a Gaussian distribution with a mean of 0 and a variance of σ_n^2; in this study, the noise standard deviation is set to 0.10 to simulate the slight equipment noise that may occur during EEG acquisition. The amplitude scaling factor is randomly sampled from a uniform distribution to simulate slight variations in EEG signal amplitude across different individuals or acquisition conditions. It should be noted that online data augmentation is applied only during the training phase; the validation and test sets do not undergo augmentation, thereby ensuring the objectivity and comparability of the model evaluation results.

3.2 SE-P1D-LSTM Network Architecture

This figure illustrates the overall structure of the SE-P1D-LSTM network proposed in this paper. The model takes an EEG feature sequence of length (1 × 512) as input and passes it sequentially through three 1D pyramidal convolution modules—each comprising a 1D convolution, batch normalization, and a ReLU activation function—to extract local temporal features of the EEG signals layer by layer and reduce feature dimensionality. Subsequently, the extracted multi-channel features are fed into an SEBlock channel attention module to adaptively learn the importance of features across different channels and enhance the representational capacity of key epilepsy-related features. Subsequently, the feature sequences are fed into a two-layer bidirectional LSTM network to further capture forward and backward long-term temporal dependencies within the EEG signals; finally, automated diagnosis of epileptic EEG signals is achieved using a classifier.

Figure 2: Architecture of the SE-P1D-LSTM Network



In the feature extraction stage, traditional CNNs typically employ a "narrow-to-wide" channel design, progressively increasing the number of channels as the network deepens. However, with small-sample epilepsy EEG datasets, excessive channel expansion tends to introduce redundant parameters, increasing the risk of model overfitting. To reduce model complexity and enhance feature representation efficiency, this paper designs a "wide-to-narrow" pyramidal 1D convolutional structure, in which the output channel counts for Conv1, Conv2, and Conv3 are set to 32, 32, and 16, respectively, exhibiting an overall trend of layer-by-layer compression. This structure preserves robust local feature representation capabilities in the earlier layers while reducing redundant channels in the later layers, thereby helping to improve the model's generalization performance in small-sample scenarios. Given an input feature sequence, the 1D convolution operation at layer l can be expressed as:

$$F^l(t) = \sum_{c=1}^{C_{\text{in}}} \sum_{k=1}^{K} W_{c,k} \cdot F^{l-1}(t + k - 1) + b \tag{6}$$

Here, $F^l(t)$ represents the feature value of the l-th convolutional layer's output at time step t, where l is the index of the convolutional layer ($l = 1,2,3$) and t is the time step index; C_{in} denotes the number of input channels; $K = 5$ denotes the convolution kernel size; $W_{c,k}$ represents the weight between the c-th input channel and the k-th position of the convolution kernel; F^{l-1} represents the feature sequence output from the previous layer; and b is the bias term. Through this convolution operation, the model can extract features of waveform variations in EEG signals within local temporal windows, providing a foundational representation for subsequent channel attention modeling and temporal modeling.

Each convolutional layer is followed by a batch normalization layer and a ReLU activation function. Batch normalization is used to standardize the feature distribution of the current mini-batch, thereby mitigating the internal covariate shift problem during training and improving the model's convergence speed and training stability. The calculation process is as follows:

$$\hat{F} = \frac{F-\mu_B}{\sqrt{\sigma_B^2+\epsilon}} \tag{7}$$

$$F_{\text{out}} = \gamma\hat{F} + \beta \tag{8}$$

Here, F represents the raw feature map output by the current convolutional layer; μ_B and σ_B^2 represent the mean and variance of the current mini-batch, respectively; and ϵ is a tiny constant to prevent division by zero. $\hat{F}$ denotes the normalized feature map; γ and β are the learnable scaling and shift parameters, respectively [10]; and F_{out} is the output feature map after batch normalization. Subsequently, the ReLU activation function further enhances the model's nonlinear expressive capability, enabling the network to learn more complex EEG feature patterns.

After passing through the pyramid-structured convolution, the model obtains multi-channel temporal feature representations. Since features associated with epileptic seizures tend to be concentrated in specific frequency bands or channels, the contributions of different channels to the final classification result vary. Therefore, this paper introduces the SEBlock channel attention mechanism after convolutional feature extraction to adaptively recalibrate the importance of each channel. Let the convolutional output feature map be:

$$U = [u_1, u_2, \ldots, u_C] \in \mathbb{R}^{C\times T} \tag{9}$$

Here, C denotes the total number of channels, T represents the number of time steps, and u_c signifies the feature vector of the c-th channel. The SEBlock first compresses the information of each channel along the time dimension via global average pooling [11, 12] to generate a channel descriptor vector:

$$z_c = \frac{1}{T}\sum_{t=1}^{T} u_c\,(t) \tag{10}$$

Here, $u_c(t)$ represents the activation value of the c-th channel at time step t, and z_c represents the global descriptor element corresponding to the c-th channel. This operation compresses the global response of each channel into a scalar, thereby reflecting the channel's average activation level within the overall time series.

After obtaining the channel descriptor vector z, the SEBlock learns the non-linear dependencies between channels through two fully connected layers and generates the corresponding channel excitation weights:

$$s = \sigma(W_2 \cdot \delta(W_1 \cdot z)) \tag{11}$$

Here, $s \in \mathbb{R}^C$ is the channel excitation weight vector; $\sigma(\cdot)$ denotes the sigmoid activation function, used to map the output to the $(0,1)$ interval; $\delta(\cdot)$ denotes the ReLU activation function; $W_1 \in \mathbb{R}^{(C/r)\times C}$ and $W_2 \in \mathbb{R}^{C\times(C/r)}$ denote the weight matrices of the two fully connected layers, respectively; $r = 4$ is the reduction ratio used to decrease the number of parameters and computational complexity. Through this process, the model automatically learns the importance of different channels for the epilepsy recognition task. The learned channel weights are applied channel-wise to the original feature maps to achieve feature recalibration:

$$\tilde{u}_c(t) = s_c \cdot u_c(t) \tag{12}$$

Here, $\tilde{u}_c(t)$ represents the weighted activation value of the c-th channel at time step t, s_c denotes the excitation weight for the c-th channel, and $u_c(t)$ represents the raw activation value before weighting. Through the SEBlock, the network can enhance the response of channels corresponding to epilepsy-related features—such as spikes and sharp waves—while suppressing background noise or interference from irrelevant channels, thereby improving the discriminative power of the feature representation.

After completing channel attention weighting, the model further utilizes a bidirectional long short-term memory (Bi-LSTM) network to model the temporal dependencies of the EEG signals. Epileptic seizures typically manifest as a dynamic process that evolves over time; relying solely on local features extracted by convolutional layers makes it difficult to fully characterize the patterns of their temporal evolution. Therefore, this paper introduces a BiLSTM structure to model EEG sequence information simultaneously in both forward and backward temporal directions, thereby capturing more comprehensive contextual features.

LSTM employs a gating mechanism to control the forgetting, input, and output of information, thereby alleviating the vanishing gradient problem—a common issue in traditional recurrent neural networks when modeling long sequences [13, 14]. At time step t, the forget gate is used to determine what information from the cell state at the previous time step needs to be retained; its calculation formula is:

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \tag{13}$$

Here, f_t represents the output of the forget gate, with a value range of [0, 1]; x_t represents the input vector at the current time step [15]; h_{t-1} represents the hidden state at the previous time step; W_f and U_f represent the weight matrices corresponding to the hidden state and the current input in the forget gate, respectively; and b_f is the bias term for the forget gate. The input gate is used to control the extent to which the current input information is written into the cell state; its calculation formula is:

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{14}$$

Here, i_t represents the output of the input gate; W_i and U_i denote the weight matrices corresponding to the hidden state and the current input for the input gate, respectively; and b_i is the bias term for the input gate. At the same time, the LSTM generates a candidate cell state to provide new information that might be written at the current time step:

$$\tilde{c}_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \tag{15}$$

$\tilde{c}_t$ denotes the candidate cell state, with values in the range $[-1,1]$; $\tanh(\cdot)$ represents the hyperbolic tangent activation function; and W_c, U_c, and b_c are the weight matrices and bias term corresponding to the candidate state, respectively [12, 16]. The LSTM updates the current cell state based on the forget gate, input gate, and candidate cell state:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{16}$$

Here, c_t represents the cell state at the current time step; c_{t-1} represents the cell state at the previous time step; and $\odot$ denotes element-wise multiplication, i.e., the Hadamard product. This process enables the network to incorporate relevant features from the current input while retaining critical historical information. The output gate determines which information from the current cell state is output as the hidden state [17], and its calculation formula is:

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{17}$$

Here, o_t represents the output of the output gate; W_o and U_o denote the weight matrices corresponding to the hidden state and the current input for the output gate, respectively; and b_o is the bias term for the output gate. Ultimately, the hidden state at the current time step is jointly determined by the output gate and the activated cell state [18]:

$$h_t = o_t \odot \tanh(c_t) \tag{18}$$

Here, h_t represents the hidden state at the current time step. The weight matrices and bias terms in the aforementioned equations are learnable parameters of the network, which are automatically updated via the backpropagation algorithm during training.

A BiLSTM consists of two independent branches: a forward LSTM and a backward LSTM. The forward LSTM processes the input sequence in chronological order, while the backward LSTM processes it in reverse chronological order; their outputs are concatenated at each time step.

$$h_t^{\mathrm{bi}} = \overrightarrow{h}_t \oplus \overleftarrow{h}_t \tag{19}$$

Here, h_t^{bi} denotes the output of the BiLSTM at time step t; $\overrightarrow{h}_t$ represents the hidden state of the forward LSTM; $\overleftarrow{h}_t$ represents the hidden state of the backward LSTM; and $\oplus$ denotes the vector concatenation operation. This paper employs a two-layer BiLSTM architecture with a hidden layer dimension of 128. The feature sequence output by the BiLSTM undergoes global average pooling to be aggregated into a fixed-length feature vector, which is then fed into the subsequent classifier.

In the classification stage, the feature vector obtained after global average pooling is mapped to the class space through a fully connected layer, and the probability of the sample belonging to each class is output via the Softmax function:

$$p_k = \frac{\exp(z_k)}{\sum_{j=1}^{2} \exp(z_j)}, k = 0,1 \tag{20}$$

Here, p_k represents the probability that the input sample belongs to class k, with $p_0 + p_1 = 1$; $k = 0$ denotes the normal class, and $k = 1$ denotes the epileptic seizure class. z_k denotes the logit value corresponding to the k-th neuron in the output fully connected layer; $\exp(\cdot)$ denotes the exponential function. The Softmax function converts model outputs into a probability distribution by normalizing the logits of each class. The model is trained end-to-end using the cross-entropy loss function, defined as follows:

$$L = -\frac{1}{B} \sum_{i=1}^{B} [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \tag{21}$$

Here, L denotes the loss function value; B denotes the batch size; i denotes the sample index; $y_i \in \{0,1\}$ represents the true label of the i-th sample, where 0 indicates normal and 1 indicates a seizure; and p_i represents the model's predicted probability that the i-th sample is a seizure. By minimizing the cross-entropy between the predicted probability distribution and the true label distribution, the model can continuously optimize network parameters and improve the performance of epileptic seizure recognition.

3.3 Model Training and Evaluation Strategy

During model training, the AdamW optimizer is employed to update the network parameters. Compared to the traditional Adam optimizer, AdamW decouples weight decay from the gradient update process, enabling more effective suppression of model overfitting and improving generalization performance on the test set. The parameter update process can be expressed as:

$$\theta_{t+1} = \theta_t - \eta \left(\frac{\widehat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} + \lambda \theta_t \right) \tag{22}$$

Here, θ_t represents the network parameter vector at the t-th iteration; θ_{t+1} represents the updated parameter vector; and $\eta = 2 \times 10^{-3}$ is the initial learning rate; $\widehat{m}_t$ denotes the bias-corrected estimate of the first moment of the gradient; $\hat{v}_t$ denotes the bias-corrected estimate of the second moment of the gradient; ϵ is a small constant to prevent division by zero; $\lambda = 1 \times 10^{-4}$ is the weight decay coefficient. This optimization strategy reduces the risk of overfitting caused by excessively large parameters while maintaining a rapid convergence rate.

Furthermore, to enhance the stability of the recurrent network training process, this paper employs a gradient clipping strategy to limit the gradient norm. When the gradient norm exceeds a preset threshold, the gradient is scaled proportionally; the calculation is as follows:

$$g \leftarrow \min\left(1, \frac{1.0}{\|g\|}\right) \cdot g \tag{23}$$

Here, $\boldsymbol{g}$ denotes the current accumulated gradient vector; $\| \boldsymbol{g} \|$ denotes the $\boldsymbol{L_2}$-norm of the gradient vector; and $\mathbf{min}(\cdot)$ denotes the function that returns the minimum value. When the gradient norm is less than or equal to 1.0, the gradient remains unchanged; when the gradient norm exceeds 1.0, the gradient is scaled to the threshold range. This method effectively mitigates the potential gradient explosion problem during BiLSTM training, thereby enhancing the stability and convergence reliability of model training.

4. Experimental Results and Analysis

4.1 Dataset and Experimental Setup

This experiment uses the University of Bonn EEG dataset to verify the proposed model. This dataset is a public benchmark dataset commonly used in seizure detection research. It contains a total of 5 groups of single-channel EEG signals from Set A to Set E, each group contains 100 samples, with a sampling frequency of 173.61 Hz, and a signal length of 4097 sampling points [19]. According to the research purpose of this paper, a binary classification task is constructed using Set A and Set E as the two classes. Among them, Set A collected scalp EEG signals from healthy subjects with eyes open and Set E collected the intracranial EEG signal from the ictal state of epilepsy patients. This task is used to evaluate the ability of the model to learn the distinguishing characteristics between normal and seizure EEG signals.

The experiments were conducted on a platform featuring an Intel Core i9 processor and an NVIDIA RTX 3080 GPU (10 GB VRAM), using Python 3.9 and PyTorch 1.10. The model was trained using the AdamW optimizer, with the initial learning rate set to 1×10^{-3} and the weight decay coefficient set to 1×10^{-4}. During the training process, the ReduceLROnPlateau learning rate scheduling strategy is used. When the accuracy rate of the validation accuracy has not been improved for 5 consecutive epochs, the learning rate decays to 0.5 of the current value [20,21]; At the same time, the early stop mechanism is introduced to terminate the training when the accuracy rate of the validation accuracy does not exceed the historical optimal result for 20 consecutive epochs. To enhance the model's robustness against noise and amplitude variations, Gaussian noise with a standard deviation of 0.10 is added to the input signals during the training phase, and random amplitude scaling is applied, with the scaling factor α sampled uniformly from the interval $[0.9, 1.1]$. All experimental results are obtained based on 10-fold stratified cross-validation to reduce the chances caused by data division and improve the reliability of the evaluation results.

4.2 Evaluation Metrics

In order to comprehensively evaluate the performance of the model in the seizure detection task, this paper adopts accuracy, sensitivity and specificity as the core evaluation indicators. These indicators are calculated based on the four basic variables in the confusion matrix: true positive (TP) refers to the number of samples that have been correctly detected, true negative (TN) is the number of normal samples correctly identified, false positive (FP) indicates the number of normal samples that have been incorrectly classified as seizures, and false negative (FN) represents the number of seizure samples that were missed. Based on the above variables, the accuracy rate measures the overall classification accuracy rate of the model; Sensitivity (also known as recall rate) reflects the model’s ability to detect seizures. The higher the value, the lower the missed detection rate; the specificity reflects the model‘s ability to identify the normal state. The higher the value, the lower the misdetection rate. Each indicator is defined as follows:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \tag{24}$$

$$Sen = \frac{TP}{TP+FN} \tag{25}$$

$$Spe = \frac{TN}{TN+FP} \tag{26}$$

Among them, Acc stands for accuracy, measuring the model's overall classification correctness; Sen stands for sensitivity (also known as recall), reflecting the model's ability to detect epileptic seizures, where a higher value indicates a lower rate of missed detections. Spe stands for specificity, reflecting the model's ability to identify normal states; a higher value indicates a lower false-positive rate. TP, TN, FP, and FN are all count variables derived by comparing model predictions with ground-truth labels.

4.3 Results Analysis

To evaluate the contributions of the SE attention module and the BiLSTM temporal modeling module to model performance, this paper designed ablation experiments comparing the complete SE-P1D-LSTM model against three variants: CNN-only, CNN + LSTM, and CNN + SE. Experimental results show that the complete model achieved the best performance across all three metrics—Accuracy, Sensitivity, and Specificity—reaching 99.87%, 99.87%, and 99.87%, respectively. Compared to “CNN Only,” the complete model shows an increase of 0.43 percentage points in accuracy, 1.00 percentage points in sensitivity, and 0.17 percentage points in specificity. Compared with the CNN + LSTM model, accuracy increased by 0.43 percentage points, sensitivity by 0.75 percentage points, and specificity by 0.12 percentage points; Compared with CNN + SE, accuracy increased by 1.43 percentage points, sensitivity by 2.12 percentage points, and specificity by 0.75 percentage points. The above results demonstrate that the combined use of SE and BiLSTM modules further enhances the model's ability to identify epileptic seizure samples—particularly regarding sensitivity—thereby helping to reduce the risk of missing such samples.

In terms of the functions of the various modules, the pyramid-structured 1D CNN is primarily used to extract local temporal features from EEG signals and to reduce redundant parameters by progressively compressing the number of channels layer by layer. The SEBlock channel attention module adaptively learns the importance of different feature channels, enhancing the response of key channels associated with epileptic seizures while suppressing irrelevant noise features. The BiLSTM module is used to model long-range temporal dependencies in EEG signals in both forward and backward directions, addressing the limitations of purely convolutional architectures in capturing long-term dynamic changes. Ablation study results demonstrate that while using CNN, CNN + LSTM, or CNN + SE in isolation yields high performance, the complete model achieves superior classification results by combining local feature extraction, channel attention enhancement, and bidirectional temporal modeling, thereby realizing a more stable and comprehensive feature representation.

Figure 1: Ablation Study

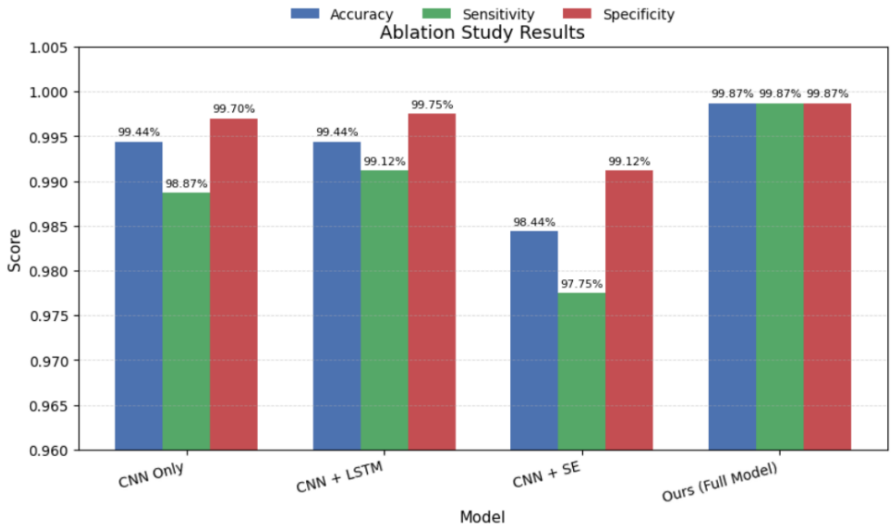


Table 1 compares the performance of different models on the task of epileptic EEG signal classification. It can be seen that the model proposed in this paper achieves 99.87% across all three metrics—Accuracy, Sensitivity, and Specificity—demonstrating the best overall performance. Compared with the LSSVM method, the model proposed in this paper achieved improvements of 5.17 percentage points in accuracy, 3.37 percentage points in sensitivity, and 1.57 percentage points in specificity. Compared with the TCN-SA method, the model in this paper achieved a 2.50 percentage point increase in accuracy and a 4.99 percentage point increase in sensitivity, though its specificity was slightly lower by 0.04 percentage points. Compared with the Conv1D+LSTM+Bayesian method, the model proposed in this paper achieved improvements of 0.40 percentage points in accuracy, 0.42 percentage points in sensitivity, and 0.30 percentage points in specificity. Although the FT-VGG16+CWT method achieves a sensitivity of 100%, its accuracy and specificity are 0.49 and 1.12 percentage points lower, respectively, than those of the model proposed in this paper. Overall, while maintaining high sensitivity, the model proposed in this paper achieved a more balanced accuracy and specificity, demonstrating its effective ability to distinguish between epileptic and normal samples.

Table 1: Performance Comparison of Different Models

Method	Year	Accuracy	Sensitivity	Specificity
LSSVM [22]	2021	94.7%	96.5%	98.3%
FT-VGG16+ CWT [23]	2021	99.38%	100%	98.75%
VMD+LECM+ DF [24]	2022	99.2%	99.31%	99.26%
TCN-SA [25]	2024	97.37%	94.88%	99.91%
Conv1D+LSTM+Bayesian [26]	2024	99.47%	99.45%	99.57%
This work	2026	99.87%	99.87%	99.87%

5. Conclusion and Future Work

This research focuses on signal preprocessing, model construction, parameter setting and experimental result analysis through the classification and identification of EEG electroencephalogram signals as the research object. By establishing a deep learning classification model and conducting feature learning and classification judgment of input signals, the experimental results show that the method can identify the signal characteristics of different categories to a certain extent, which has a good classification effect. On the whole, this article completes the basic process from data processing to model training and result analysis, which provides a certain reference for the subsequent further optimization of the intelligent diagnosis method of EEG signals.

Follow-up research can be further verified by combining real clinical experimental data. By introducing more sources and larger-scale clinical electroencephalogram data, the stability and generalization ability of the model in actual medical scenarios can be tested. At the same time, it can also be compared and analyzed with the doctor‘s diagnosis results, evaluate the reliability of the model in auxiliary diagnosis, and provide support for the application of intelligent medical auxiliary diagnosis system.

References

[1] GBD 2021 Epilepsy Collaborators. (2025). Global, regional, and national burden of epilepsy, 1990–2021: A systematic analysis for the Global Burden of Disease Study 2021. *The Lancet Neurology*.

[2] Guo, J., & Wei, X. N. (2007). Research progress in the application of electroencephalography. *Journal of Continuing Education of Shaanxi Normal University, 24*(2), 122–124.

[3] Li, Y. (2025). *Research on epileptic EEG signal analysis methods based on graph neural networks* (Master's thesis, Shandong University, Jinan, China).

[4] Subasi, A., & Ercelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine, 78*(2), 87–99.

[5] Guo, L., Rivero, D., Dorado, J., et al. (2010). Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. *Journal of Neuroscience Methods, 191*(1), 101–109.

[6] Alotaiby, T. N., Alshebeili, S. A., Alshawi, T., et al. (2014). EEG seizure detection based on linear and nonlinear features using artificial neural networks. *Clinical Neurophysiology, 125*(7), 1405–1414.

[7] Acharya, U. R., Fujita, H., Lih, O. S., et al. (2017). Automated detection of seizures using a deep convolutional neural network. *International Journal of Environmental Research and Public Health, 14*(10), 1182.

[8] Tsiouris, K. M., Tzallas, A. T., Tsipouras, M. G., et al. (2018). A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. *Computers in Biology and Medicine, 99*, 24–37.

[9] Rukhsar, S., & Tiwari, A. K. (2024). Lightweight convolution transformer for cross-patient seizure detection in multi-channel EEG signals. *Biomedical Signal Processing and Control, 95*, 106452.

[10] Li, X. S., Hu, C. X., Tian, Z. J., et al. (2026). Lightweight attention residual-enhanced YOLO architecture for remote sensing object detection. *Journal of Geo-Information Science*.

[11] Wang, W. H. (2024). *Semi-supervised ship engine state recognition method based on efficient self-attention* (Master's thesis, Dalian Maritime University, Dalian, China).

[12] Zhang, W. S. (2025). *Deep learning-based network intrusion detection methods* (Master's thesis, Guilin University of Electronic Technology, Guilin, China).

[13] Luo, J. (2025). *Chinese text classification based on syntactic analysis and abstract semantic representation* (Master's thesis, Wuhan Institute of Technology, Wuhan, China).

[14] Li, J. H., Gao, L. F., & Li, X. W. (2026). Application of surface acoustic wave wireless temperature sensors in fault prediction of power equipment. *Electric Power Equipment Management*.

[15] Chen, N. (2025). *Research on molecular property prediction based on multi-view graph learning* (Master's thesis, Nanning Normal University, Nanning, China).

[16] Ruan, Z. J. (2025). *Research on sequence recommendation algorithms for bi-objective optimization* (Master's thesis, Guangdong University of Foreign Studies, Guangzhou, China).

[17] Li, X. L., Zhang, Y., Zhang, W., et al. (2024). State-of-health prediction for lithium-ion batteries based on health factors and CNN-BiLSTM. *Journal of Nanjing University of Information Science & Technology*.

[18] Huang, Q. M. (2024). River water level prediction method based on long short-term memory networks. *Information Recording Materials*.

[19] Andrzejak, R. G., Lehnertz, K., Morrison, C., et al. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E, 64*(6), 061907.

[20] Wang, J., & Li, L. Y. (2025). Breast ultrasound lesion segmentation network based on channel-wise spatial adaptive selective kernel convolution and bidirectional boundary perception mechanism. *Journal of South China University of Technology (Natural Science Edition)*.

[21] Zhang, X. T. (2025). *Application of deep learning algorithms integrating Transformer and mixture-of-experts in sleep staging* (Master's thesis, Shanghai Polytechnic University, Shanghai, China).

[22] Torse, D. A., & Khanai, R. (2021). Classification of epileptic seizures using ensemble empirical mode decomposition and least squares support vector machine. In *2021 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1–5). IEEE.

[23] Rashed-Al-Mahfuz, M., et al. (2021). A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (EEG) data. *IEEE Access, 9*, 1–12.

[24] Liu, X., Wang, J., Shang, J., et al. (2022). Epileptic seizure detection based on variational mode decomposition and deep forest using EEG signals. *Biosensors, 12*(10), 1275.

[25] Huang, L., Zhou, K., Chen, S., et al. (2024). Automatic detection of epilepsy from EEGs using a temporal convolutional network with a self-attention layer. *Bioengineering, 23*(1), 50.

[26] Jain, M., Bhardwaj, H., & Srivastav, A. (2024). Bayesian-enhanced EEG signal analysis for epileptic seizure recognition: A 1D-CNN LSTM approach. In *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (pp. 1–6). IEEE.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights