Scientific Eminence Group

Journal of Neuroscience Research and Alzheimer's Disease

Research Article

Open Access

jnra@scientificeminencegroup.com

Multi-Scale Sparse Adaptive Fusion Network Predicts Epileptic Seizures

PIAO Yan*, SHEN Xue-ting, ZHAO Hai-tong and YANG Hui-ru

School of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun 130012, China

*Corresponding Author

PIAO Yan, School of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun 130012, China, E-mail: piaoyan@cust.edu.cn

Citation

PIAO Yan, SHEN Xue-ting, ZHAO Hai-tong, YANG Hui-ru (2025) Multi-Scale Sparse Adaptive Fusion Network Predicts Epileptic Seizures. J Neurosci Res and Alzheimer's Dis 4: 1-19

Publication Dates

Received date: December 15, 2024
Accepted date: January 15, 2025
Published date: January 18, 2025

Abstract

The current epilepsy prediction methods have the problems of feature extraction in the fixed frequency domain and redundancy of spatial features, which are difficult to effectively characterize multi-dimensional features, resulting in low prediction accuracy. Therefore, for the effective interactive representation of spatiotemporal spectral features, a new multi-scale sparse adaptive convolutional network model based on attention mechanism (MS-SACN-MM model) was proposed. The model extracts important features by preprocessing EEG data, constructing multi-scale temporal convolutional layers, adaptive spectral convolutional layers, and sparse regularized graph convolutional layers, and fuses multi-domain features through multi-layer perceptron and multi-head attention mechanisms to improve the performance of the model. Experimental results show that the prediction accuracy of the MS-SACN-MM model can reach up to 99.9% 5-30 minutes before seizures on the CHB-MIT dataset, showing superior prediction performance.

Keywords: Computer; Epilepsy Prediction; Multi-Scale Sparse Adaptive Mechanism; Multilayer Perceptron; Multi--Headed Attention; Focus Loss Training Strategy

Introduction

Epilepsy is one of the most common chronic central nervous system diseases worldwide, characterized by suddenness, recurrence, and difficulty in curing, posing a great threat to the quality of life and health of patients. Therefore, accurate prediction of epileptic seizures is of great practical significance.

With the continuous deepening of research on the prediction of epilepsy onset, machine learning methods have been widely applied to epilepsy prediction. Al-Hadeethi et al. [1] selected the covariance matrix to extract statistical features, used non-parametric tests to obtain the most significant feature set, and built an adaptive enhanced least squares support vector machine classification model to achieve better results. Yi Fangji et al. [2] analyzed sample entropy and Pearson correlation coefficients, selected the optimal feature parameter combination, and achieved good results in the two-class classification of interictal and preictal periods using support vector machines. Vicnesh et al. [3] extracted non-linear features of EEG signals, used them as inputs for decision tree classification, and enhanced the inter-pretability of nonlinear features. Zhou Mengni et al. [4] constructed brainwave feature vectors based on per-mutation entropy, used them as inputs for support vector machine learning models, and achieved prediction before the onset of seizures within 50 minutes. However, the complexity of epilepsy EEG features makes manual extraction highly limited, and traditional machine learning algorithms have limited classification capabilities, making it difficult to effectively predict epilepsy onset.

In recent years, with the rapid development of deep learning technology, more and more scholars have started to research epilepsy prediction models based on deep learning. Singh et al. [5] extracted amplitude and power features of different brainwave rhythms as inputs for convolutional neural networks and long short-term memory networks to classify the interictal and preictal periods of epilepsy, achieving high accuracy. Mane et al. [6] proposed the FBCNet model, which aims to con-sider filters with multiple cutoff frequencies to extract information from different frequency bands, thereby enhancing the model's frequency domain representation. Avcu et al. [7] proposed a Seizure-Net method, which converts EEG signals into time-frequency maps through short-time Fourier transform and combines convolutional neural networks to achieve automatic seizure detection. Yang et al. [8] proposed a dual self-attention residual network based on the CHB-MIT dataset to achieve high-precision epilepsy prediction. Jia et al. [9] proposed universal graph convolutional network model architecture for predicting seizures in 18 patients from the CH-B-MIT dataset, with an average prediction accuracy of 92%. Wang Menghao et al. [10] proposed a novel hybrid neural network model that extracts time-frequency domain features through multi-scale convolution, combines channel attention mechanisms to extract spatial features, and achieves a prediction accuracy of 83.53% based on time-frequency-space domain features. Liu et al. [11] used a brainwave channel attention weight learning mechanism to extract discriminative features of continuous multi-channel EEG signals. Kumar et al. [12] proposed a novel classification method based on du-al-stream adaptive multi-attention to extract and integrate spatiotemporal information, revealing hidden emotions. With the development of attention mechanisms, they are also increasingly used in the field of epilepsy prediction. Liu et al. [13] considered using multi-head attention mechanisms to capture the spatial spectral features of brainwave signals, thereby extracting key information related to the spatial frequency domain.

In the field of epilepsy seizure prediction, an in-creasing number of methods are being applied to predict seizures, especially the fusion of neural networks and attention mechanisms, which have achieved good prediction results. Due to the complexity of EEG signals, it is necessary to consider both the time series features and the electrode diversity and channel correlations of EEG signals. However, in existing methods based on multi-domain representation of spatiotemporal spectra, there is often data redundancy, leading to a decrease in the predictive performance of the model. Therefore, methods that can ignore unnecessary features and significantly extract spatiotemporal frequency features are important for enhancing the predictive performance of the model.

Based on the above issues, this article proposes a new multi-scale sparse adaptive convolutional network model (MS-SAC-N). By constructing a triple convolutional layer, including a bidirectional multi-scale temporal convolution block that fully considers temporal information, an adaptive spectral convolution block that dynamically extracts optimal frequency domain features, and a regularization term enforcing sparsity graph convolution block with dimensionality reduction graph connections, effective extraction of spatiotemporal spectral

features is achieved. The extracted information is used as input for the MLP-Multi-Attention (MM) mechanism, which performs multi-domain feature fusion based on attention mechanisms. Considering the issue of imbalanced sample categories, a focal loss training strategy is introduced. Finally, the network is fully trained on the CHB-MIT dataset to classify preictal and interictal stages of epileptic seizures, thereby improving prediction accuracy to a certain extent.

Algorithms

This article constructs the MS-SACN-MM feature extraction model and attention fusion module, as shown in Figure 1.

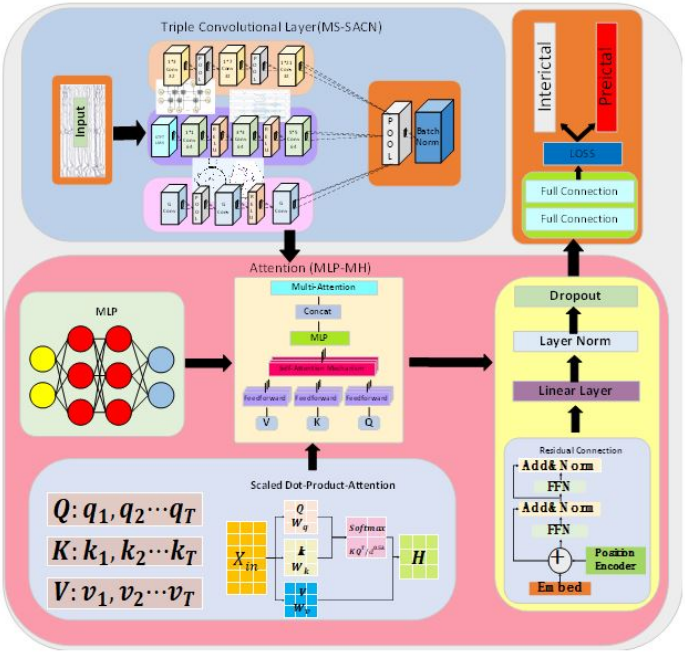


Figure 1: The overall architecture of the network

Triple convolutional layer

To effectively extract the spatiotemporal and frequency-domain features of EEG signals, this paper designs three convolutional blocks. In the time feature extraction convolutional block, a bidirectional convolution is constructed to comprehensively capture key features as important precursors in EEG signals may exist in different time periods. In the frequency domain feature extraction convolutional block, an adaptive mechanism is introduced to dynamically adjust the se-lection of frequency bands based on a recursive updating algorithm to effectively deal with the non-stationarity of EEG signals and reduce false positives. In the spatial feature extraction convolutional block, a sparse regularization mechanism is combined to impose sparsity constraints on the weight matrix and features, removing unimportant edge connections to significantly improve computation time and reduce memory consumption.

Bidirectional multi-scale temporal convolutional block

Build bidirectional multi-scale temporal convolutional blocks, using convolutional kernels of different sizes, introducing dilated convolutions to capture temporal features at different time scales.

Due to the high complexity and long duration of epileptic seizures, it is designed to use different sizes of convolutional kernels in bidirectional time convolutional layers to capture time series features at different time scales. This design constructs three time convolutional layers, with two branches (forward and backward) for each convolutional layer. Small convolutional kernels are used to capture detailed features, while large convolutional kernels capture coarse features. The de-sign includes 16 input channels, 32 output channels, and three layers with convolutional kernels of sizes 3, 7, and 21, respectively. A kernel size of 3 captures local features such as rapid changes in peaks or ripples, which can be indicative of epileptic activity. A kernel size of 7 captures features at a medium time scale, corresponding to early signs of epileptic seizures. A kernel size of 21 captures trends and patterns at a

long time scale.

Assuming the input time series is $x = [x_1, x_2, \ldots x_T]$, σ is the activation function, then the formula for multi-scale convolution is (1):

$$y_t^k = \sigma\left(\sum_{(i=0)}^{(K-1)} \omega_i^k \bullet x_{(t-i)} + b^k\right) \qquad (1)$$

There, y_t^k is the output of the k-th convolutional kernel at time t, ω_t^k is the i-th weight of the k-th convolutional kernel, with K being 3,7,21.

Considering the long-term patterns of epileptic seizures, we introduce dilated convolutions: introducing dilation factors in the convolutional layer, which means inserting gaps between the elements of the original convolutional kernel and modifying the dilation rate, thus were increasing the receptive field of the convolutional kernel without adding extra parameters as shown in Figure 2. This allows the model to capture long-term dependencies without sacrificing temporal resolution.

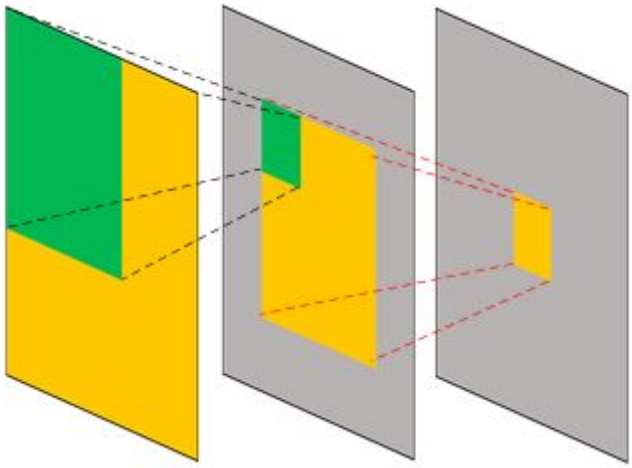


Figure 2: Receptive field

Build bidirectional convolution, convolve in both forward and backward directions of the time series, define the complete time period as *T*, the specific operation is as follows:

Forward convolution: Convolution performed on the input sequence in chronological order from time step *t*=0 to *T*; reverse convolution: perform convolution on the input sequence in reverse time order, from time step *t*=*T* to 0; merge the forward output feature, y_{z1}, and the backward output feature, y_{f1}, by simple concatenation to obtain a multi-scale representation of the time series features. Define the fused feature as *Y*, and solve as shown in equation (2):

$$Y = \frac{1}{2}(y_{z1} + y_{f1}) \qquad (2)$$

Adaptive Spectral Convolution Block

Build an adaptive spectrum convolution block, introduce a fusion adaptive mechanism of discrete short-time Fourier transform (DTFT) to dynamically extract EEG data time-frequency domain features. Traditional DTFT converts time-domain data into fixed frequency domain, while EEG signals are complex and variable in the frequency domain. Therefore, introducing adaptive dynamic feature extraction enriches frequency domain features.

Introduce an adaptive filter composed of band-stop filtering and high-pass filtering based on Butterworth filters. First, perform high-pass filtering to ensure the removal of low-frequency noise, and then perform band-stop filtering to eliminate high-frequency noise and low-frequency drift, retaining its main frequency components and improving the band-stop filtering effect. This can be represented by the formula:

$$\mathrm{DTFT}\left[c,k\right]=\sum_{n=0,k=0}^{N-1}y_{bp}\left[n\right]\omega\left[n-c\right]e^{-j\frac{2\pi kn}{N}} \quad (3)$$

Where, ω[*n*-*c*] is the Hanning smoothing window function, y_{bp} [n] is the result after being processed by an adaptive filter, The process is shown in (4):

$$\left\{\begin{array}{c} y_{hp}\left[n\right]=HpFilter\left(x\left[n\right],cutoff,fs\right) \\ y_{bp}\left[n\right]=BpFilter\left(y_{hp},lc,\mathrm{hc},fs\right) \end{array}\right\} \quad (4)$$

Where, cutoff=1Hz is the cutoff frequency for high-pass filtering,*lc*,*hcare* the cut-off frequencies with band-stop filtering, used to remove frequencies of 60Hz-65Hz and 120Hz-125Hz.

Design an adaptive time-frequency sequence integration mechanism based on the Least Mean Squares (LMS) algorithm with trainable parameters, to adjust the fixed characteristics of the Fourier transform and dynamically extract features in the time-frequency domain. For an input signal *x(n)* after passing through a parameter-adjustable horizontal filter, the output is *y(n)*.The LMS algorithm automatically adjusts the parameters of the filter based on the error between the output signal *y(n)* and the desired signal *d(n)*, allowing the filter to adapt to the time-varying statistical characteristics of the random signal. The implementation process is as follows:

Let the input signal be a vector (5):

$$X\left(n\right)=\left[x\left(n\right),x\left(n-1\right),...x\left(n-L\right)\right]^{T} \quad (5)$$

The output of the LMS adaptive filter is then (6):

$$y\left(n\right)=W^{T}\left(n\right)X\left(n\right) \quad (6)$$

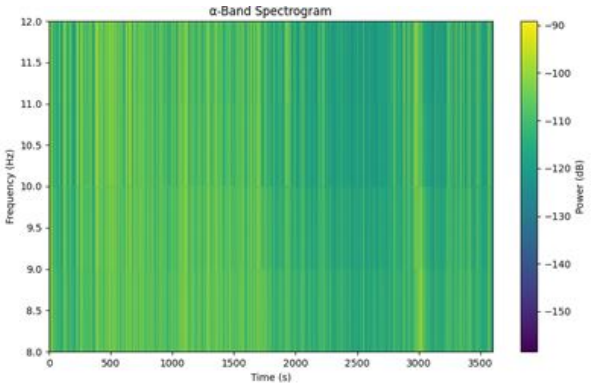
The update formula for the weight vector of the LMS adaptive filter is shown as (7), where, *e(n)* is error signal.

$$W\left(n+1\right)=W\left(n\right)+2\mu e\left(n\right)X\left(n\right) \quad (7)$$

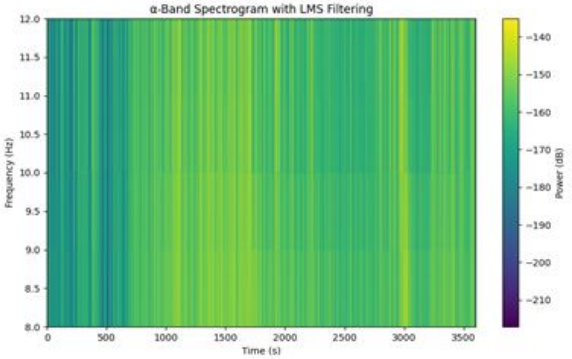
Based on the aforementioned five frequency bands, apply the convolution kernel *H(f)* in the frequency domain. Convolve the components of each frequency band, the formula is (8):

$$Y\left(n\right)=H\left(f\right)\otimes y\left(n\right) \quad (8)$$

The adaptive spectral convolution block con-structed in this article introduces dynamic periodic weights and extended DTFT to enhance the dependency relationships in the processing of time series data using discrete Fourier transform. The EEG signal frequency domain is divided into five common rhythms: *α*, *β*, *θ*, *δ*, and *γ*, each rhythm affecting different brain activity states [14]. By analyzing and extracting significantly relevant information, irrelevant features are ignored to enhance model performance. Figure 3 shows the visual spectrum of the *α* rhythm in patient chb01, where (a) is the regular DTFT transformation and (b) is the transformation after the adaptive mechanism is added.



(a) STFT



(b)STFT+LMS

Figure 3: Chb01 patient α rhythm spectrum

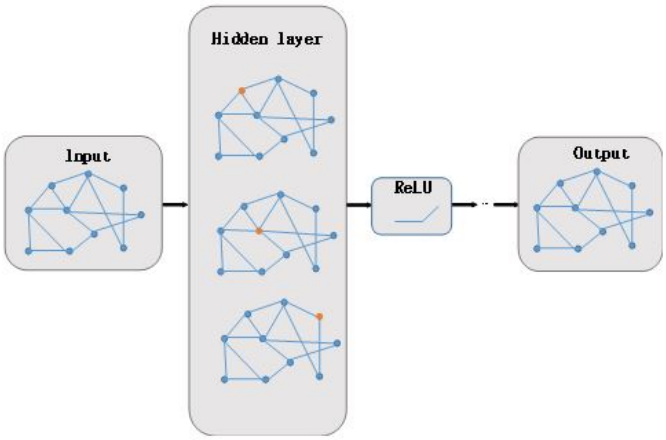


Figure 4: Graph convolution operations

Sparse Graph Convolutional Block

The graph is composed of nodes and edges connecting the nodes, with a complex structure but containing rich potential value. Lian et al. [15] proposed a global-local graph convolutional neural network for predicting epileptic seizures, tested on the Freiburg dataset, and achieved good results. In this paper, we construct sparse graph convolutional blocks, aggregating neighboring and distant nodes in a better way to represent brain structural connections, breaking the common uniform distribution rule. We introduce L_1 regularization strategy to impose sparsity constraints on the model's weight matrix and features, alleviating the oversmoothing issue caused by increasing layers, handling redundancy and avoiding fea-

ture confusion. The graph convolution structure is shown in Figure 4.

Graph convolutional block based on sparse representation, introducing Hilbert transform for extracting signals in different frequency bands, extracting instantaneous phase of signals in different frequency bands for multi-channel correlation analysis. For a given signal *x(t)*, the analytical signal *z(t)* obtained by Hilbert transform, instantaneous phase $\phi(t)$, and instantaneous phase difference $\Delta\phi(t)$ are shown in equation (9):

$$\left\{ \begin{array}{c} z(t) = x(t) + jH(x(t)) \\ \phi(t) = arg(z(t)) \\ \Delta\phi(t) = \phi_i(t) - \phi_j(t) \end{array} \right\} \quad (9)$$

If two signals are phase synchronized, the phase difference $\Delta\phi(t)$ fluctuates within a certain range without significant increase or decrease over time. Introducing the Phase-locked value (PLV) to calculate the average cosine value of the phase difference to evaluate the degree of synchronization, defined as equation (10):

$$\mathrm{PLV} = \left| \frac{1}{N} \sum_{i=1}^{N} e^{j\Delta\phi(t)} \right| \quad (10)$$

Where, $e^{j\Delta\phi(t)}$ is the complex unit representation of phase difference, N is the total number of time points, and PLV ranges from 0 to 1, with values closer to 1 indicating higher phase synchrony.

Design a three-layer graph convolutional layer to capture spatial features of different scales. EEG data from multiple channels typically represent electrical activities in different regions of the brain, which have certain spatial correlations between them. By constructing a three-layer graph convolutional network, as shown in Figure 5, brain graph structures are connected to aggregate features of neighboring nodes, enabling a deeper exploration of the relationships between channels and extraction of spatial features. In the first layer, local connections are considered between adjacent electrode nodes, and the weights defining the adjacency matrix based on physical distances are determined using a Gaussian kernel function, as expressed in Equation (11):

$$e^{-\frac{r(i,j)}{2\tau^2}} \quad (11)$$

Where τ is adjustable parameters, *r(i,j)* represent the physical distance between channels *i* and *j*.

Figure 5: Diagram of brain connectivity

The second layer mainly considers the use of electrodes at a greater distance to obtain more spatial information, while the third layer normalizes the node features to enhance the overall regularity of the graph structure. In addition, combining activation functions (RELU) and pooling (POOL) operations enhances the interpretability of the model.

The graph convolution operation is specifically manifested as follows: Given the adjacency matrix A and node features X, the node feature matrix of the l-th layer is denoted as $H^{(l+1)}$:

$$H^{l+1} = \sigma\left(\tilde{A}H^{l}W^{l}\right) \quad (12)$$

Where, Hl is the node feature matrix of the l-th layer, $\tilde{A}$ is the normalized adjacency matrix used to propagate information between nodes, W^{l} is the learnable weight matrix of the l-th layer, and σ is the ReLU activation function.

Introducing a regularization term enforces sparse representation [16], by adjusting the regularization coefficient to reduce the density of the adjacency matrix and node features, obtaining the sparsest representation for each state, thereby filtering out significant features of different classes to improve the efficiency, generalization ability, and interpretability of the model.

Introducing L_1 regularization, the sparse regularization forms for the adjacency matrix (A) and node features (H) are shown in equation (13):

$$\left\{ \begin{array}{l} L_A = \lambda_A A_1 \\ L_H = \lambda_H H_1 \end{array} \right\} \quad (13)$$

Where, $\|A\|_1$ is the L_1 norm of the adjacency matrix, which is the sum of the absolute values of all elements in the adjacency matrix. λ_A is the regularization coefficient, and the larger the value of λ_A, the sparser the adjacency matrix, ensuring the most important connections. $\|H\|_1$ is the L_1 norm of the node matrix, preserving the most important node features.

MLP-MM Mechanism

After extracting the corresponding features, considering both linear and nonlinear relationships, this paper applies the MLP-MM mechanism for interactive fusion. The extracted time, space, and frequency domain features are defined as a three-dimensional tensor, $X_{in} \in^{R(B\times T\times F\times C)}$, where B is the batch size, T is the time dimension feature, F is the frequency domain feature, and C is the spatial dimension feature. This paper utilizes the powerful feature transformation ability of the multilayer perceptron (MLP) to further process the data fused by the self-attention mechanism. By combining the multi-head mechanism to capture complex linear or nonlinear dependencies between inputs, the model's ability to process information in different dimensions is enhanced, thereby promoting multi-domain feature fusion of the processed EEG signals.

The core of the multi-head attention mechanism is to introduce multiple attention heads on top of the self-attention mechanism, learning input data features in parallel [17], considering each element in the sequence when processing EEG data, so that the model can capture richer contextual information in the sequence data. Based on the three vectors of query (Q), key (K), and value (V), the relationships between inputs are established. Existing methods generate these three vectors from the original input through linear transformations, while the multi-domain representation of EEG signals mostly exhibits nonlinearity. For input features processed after embedding by convolutional modules, $X_{in} \in^{R(N\times d)}$, ($N$ is the sequence length, d is the input feature dimension), they are passed into different "heads" to generate the Q, K, and V vectors.

$$\left\{ \begin{array}{l} Q = X_{in}W_Q \\ K = X_{in}W_K \\ V = X_{in}W_V \end{array} \right\} \quad (14)$$

In the formula, W_Q, W_K, and W_V are the corresponding mapping matrices; the dot product calculation is performed on the query vector Q and the key vector K to obtain the correlation between positions in the sequence, combined with the softmax function for normalization, thus obtaining the influence weight between positions. The formula for calculating attention scores is (15):

$$\text{Attention}(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (15)$$

In the equation, Attention (Q, K, V) are the attention functions for query Q, key K, and value V, where $\sqrt{d_k}$ is the dimension of the key vectors, is the scaling factor to prevent the dot product of vectors from becoming too large, and K^T is the transpose of the set of position vectors to be retrieved.

The core idea of multi-head attention mechanism is to divide the traditional single self-attention mechanism into multiple "heads". By calculating multiple independent "heads" in parallel, different attention distributions are computed. The query, key, and value of each "head" are calculated in different subspaces, generating multiple different attention features, which are then merged. The calculation formula for multi-head concatenation is (16):

$$\left\{\begin{array}{c} MH(Q,K,V) = Concat(head_1,..,head_h)W_o \\ head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right) \end{array}\right\} \quad (16)$$

Where, W_o is the concatenated linear transformation matrix used for the final output.

Focus Loss Training Strategy

Using the focal loss training strategy to increase the weight of hard-to-classify samples, alleviate the problem of sample class imbalance [18], improve AUC (area under the ROC curve), and perform seizure pattern recognition and classification for epilepsy based on the spatiotemporal fusion features of EEG signals over a long time span. By evaluating and optimizing the trained model, the prediction of epilepsy seizures can be achieved. The focal loss, L_p, can be defined as (17):

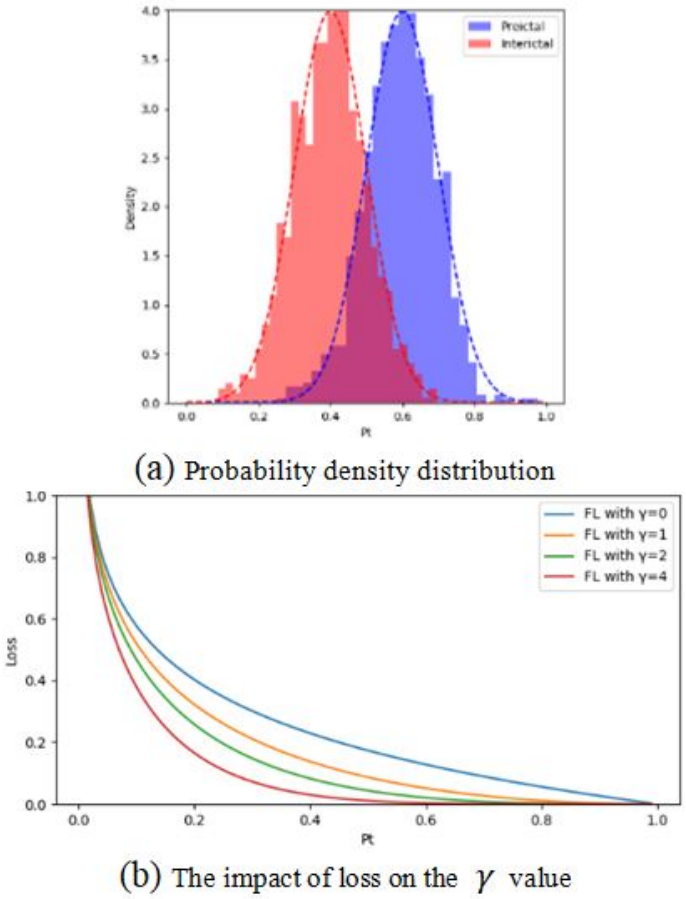
$$L_p = -\alpha_t(1-p)^{\gamma}log(p) \quad (17)$$

Where, p is the probability of correct prediction by the model, α_t is the balancing factor used to adjust the ratio of positive and negative samples, and γ is the adjustment factor. Due to the low proportion of positive samples in the dataset, this paper conducts separate analysis for each patient, dynamically selecting combinations of αt and γ to increase the weight of positive samples, adjust the loss weight, and reduce overfitting. As shown in Figure 6(a), the probability density of the predicted classification for patient chb01 is displayed, with the overlapping area indicating the ambiguous region. By analyzing the probability distribution in Figure 6(b), trends of the loss function under different γ values can be observed. The x-axis and y-axis represent the output probability and corresponding loss value, respectively. For very certain predictions (probability values around 0.75), a larger γ value corresponds to a smaller penalty.

Based on this, this article analyzes the impact of different combinations of α_t and γ on the AUC values of each patient, using polynomial regression to fit the relationship between AUC and α_t and γ, as shown in equation (18):

$$\text{AUC}(\alpha_t,\gamma) = c_0 + c_1\alpha_t + c_2\gamma + c_3\alpha_t^2 + c_4\gamma^2 + c_5\alpha\gamma \quad (18)$$

Where, c_0, c_1, c_2, c_3, c_4, c_5 are the parameters to be fitted, determined by the method of least squares. As shown in Figure 7, for patient chb01, the value of α_t is chosen as 0.75, and γ is chosen as 2.



(a) Probability density distribution

(b) The impact of loss on the γ value

Figure 6: Analysis of the impact of parameters

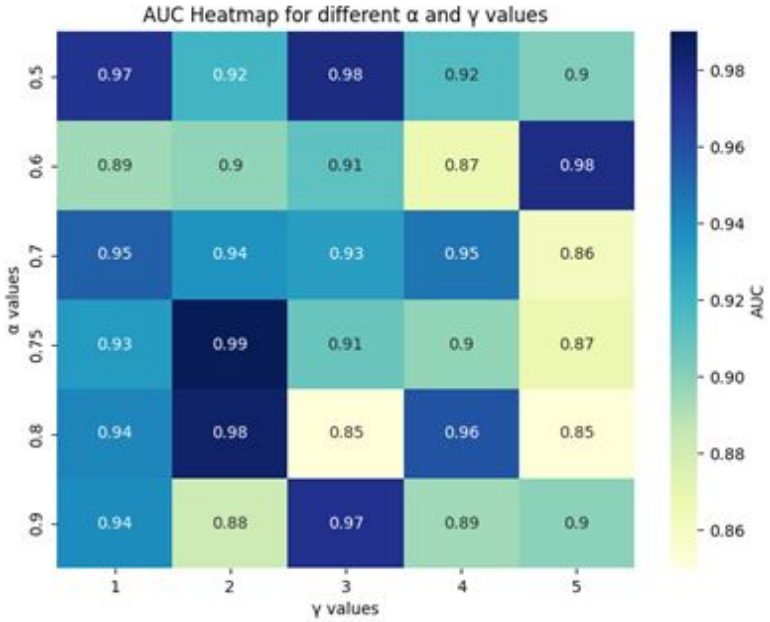


Figure 7: Parameter combinations are associated with AUC

Experiment

CHB-MIT Dataset

The study used the CHB-MIT dataset, which sup-ports epilepsy and EEG research and was jointly created by the Massachusetts Institute of Technology and Boston Children's Hos-

pital[19]. The dataset consists of 23 CHB cases, with 23 case records for 22 patients, including 5 males aged 3 to 22 years old and 17 females aged 1.5 to 19 years old. One of the subjects, CHB21, had their record taken again 1.5 years after CHB01. Each case contains 9-42 continuous .edf files, with EEG signals of patients lasting 1h, 2h, or 4h, sampled at 256 Hz with 16-bit resolution. The records use the international 10-20 system for EEG electrode placement and naming, with 18 channels recording the voltage difference between electrodes, where each channel consists of two vertically adjacent electrodes. Additionally, the CHB-MIT dataset labels whether seizures occurred during the patient's record and records the time of seizure onset.

Data Preprocessing

According to the status of epileptic seizures, four time periods are divided into pre-seizure, seizure, post-seizure, and interictal periods [20], as shown in Figure 8. The pre-seizure period generally lasts from a few minutes to tens of minutes before the seizure; the seizure period typically lasts from a few seconds to a few minutes from the beginning to the end; the post-seizure period refers to the time from the end of the seizure to the recovery of normalcy in a few minutes; the interictal period refers to the long time between the post-seizure period and the patient's next seizure, during which the patient's condition is no different from that of a normal person. To predict epileptic seizures in advance, it is necessary to analyze the patient's brain electrical signals and effectively identify whether the patient is in a pre-seizure or interictal state. Therefore, accurately dividing epileptic brain electrical signals is a prerequisite for predicting seizures.

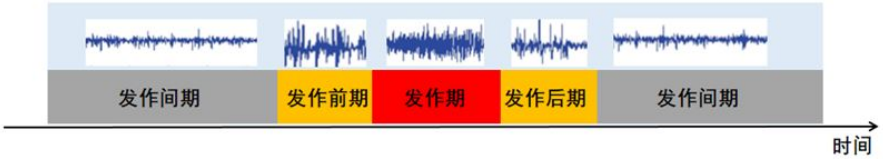


Figure 8: Seizure state

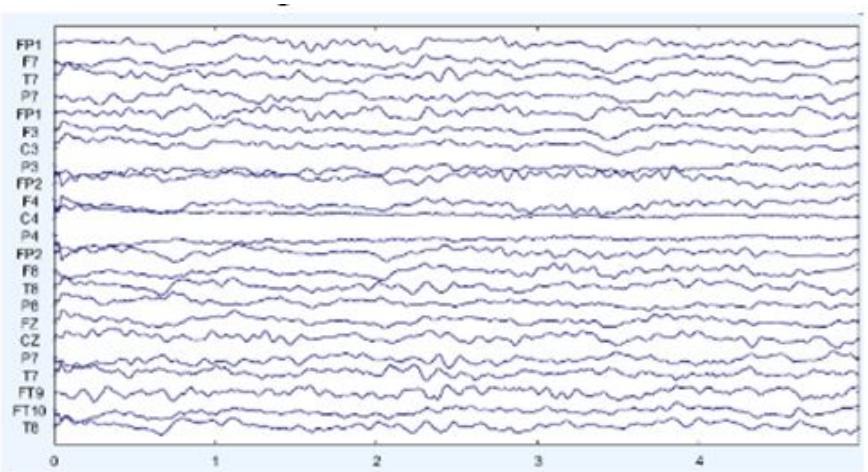


Figure 9: EEG signal after processing

During the data collection process, due to the subjective behavior of patients and external influences, EEG signals may be contaminated with eye movement and muscle artifacts. Therefore, data cleaning, filtering, noise reduction, baseline correction, artifact removal, and data enhancement are necessary. In this study, EEG data was resampled and divided into fixed-length segments. An 8-second window size was used, with moving distances of 4s, 2s, and 1s between two consecutive windows, resulting in 8-second data segments. Independent Component Analysis (ICA) was employed at a sampling rate of 256Hz to remove artifacts and enhance useful components while suppressing noise. Epileptic patients exhibit specific waveforms in their EEG signals during seizures, such as spike waves, sharp waves, slow waves, spike-slow complex

waves, and high-frequency oscillations [21]. Bandpass filters ranging from 0.5Hz to 60Hz and a notch filter at 50Hz were used to eliminate noise and power line interference. Additionally, baseline correction was applied to remove the DC component from the signal, ensuring that the average value within a selected time period is zero. The preprocessing work was mainly carried out using the EEGLAB software, and Figure 9 shows the EEG signals before and after preprocessing (4-second segments).

Model Implementation

This simulation experiment was conducted on the Windows 10 operating system. The experiment was carried out on a computing platform with a GPU of 16vCPU Intel(R) Xeon(R) Platinum 8352V CPU @ 2.10GHz and 3 NVIDIA GeForce RTX 4090 24G graphics cards. The relevant code was written in Python 3.8 and implemented under the Pytorch 1.13.1+cu118 framework.

This article uses the publicly available CHB-MIT dataset and employs leave-one-out cross-validation ten-fold cross-training for each patient. 70% of the data is used as the training set, while 30% is used as the test set. The training is conducted for 200 batches with 100 epochs, totaling 20,000 iterations. The initial learning rate is set at 0.001 and dynamically adjusted during training to find the optimal solution. The focal loss function is used along with dropout to prevent overfitting, with a dropout parameter value of 0.5. By combining the focal loss training strategy and selecting appropriate values for α_t and γ, the issue of class imbalance is addressed, reducing training loss and improving the reliability of the model. The results obtained are shown in Figure 10.

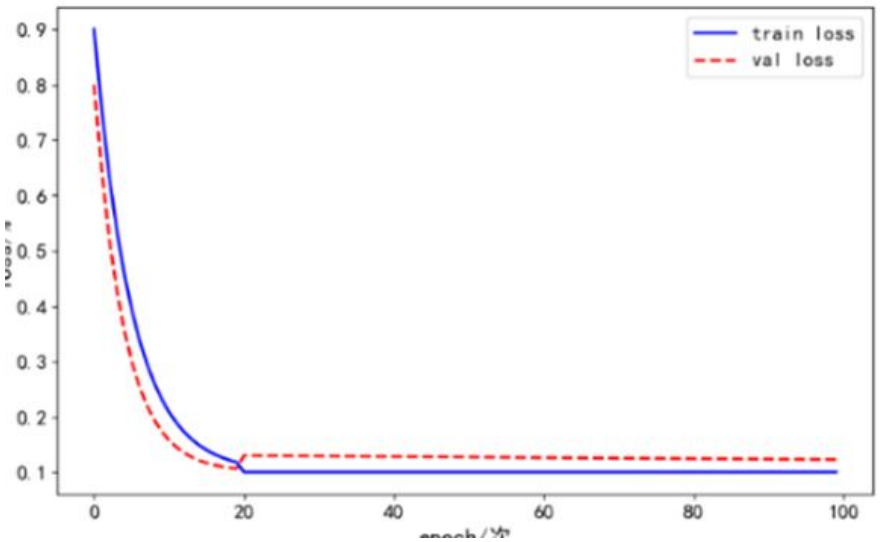


Figure 10: Loss curve graph

Result

In this section, in order to evaluate the performance of the proposed MS-SACN-MM model for predicting epileptic seizures, we analyzed the experimental results and assessed the timeliness and novelty of our method. In the experiments, we selected recent years' learning-based methods for comparison. It is worth noting that existing networks often overlook the dynamic frequency domain characteristics and multi--channel redundancy properties, making it difficult to fully characterize the multidimensional features of EEG signals. Therefore, we constructed an adaptive spectral convolution to dynamically extract frequency domain characteristics, effectively extracting useful α, β, θ, δ and γ five common rhythms, and quantitatively analyzed accuracy (ACC), specificity (SP), recall rate (RE), and F1 parameters. In response to the large amount of data redundancy in channel interactions, our experiments introduced a sparse regularization mechanism to eliminate meaningless node connections, reduce the dimension of the topological structure, and significantly retain strongly correlated features. In this experiment, leave-one-out cross-validation was used to ensure that each patient's data was sufficiently trained and tested. To evaluate quantitative performance, we calculated the values of the area under the ROC curve (AUC), sensitivity (Sn), and false positive rate (FPR) to measure the accuracy of predictions, the sensitivity of

the model, and the false positive rate. The higher the corresponding values, the better the performance of the proposed model. Compared to other methods, our proposed model has significantly improved performance and is more versatile.

The Influence of Different Rhythms

The EEG signal is cropped into segments using a sliding window of 1 to 5 seconds. Based on the five common rhythms of EEG signals – α, β, θ, δ, and γ the performance of each rhythm in different stages and states of epilepsy is analyzed as shown in Table 1. δ waves show local enhancement in the pre-seizure period, significantly enhanced during the seizure period and post-seizure period. θ waves show signs of enhancement except during the interictal period. α waves weaken from the interictal period to the post-seizure period. β waves show no significant changes throughout the process, while γ waves show abnormal discharge during the seizure period.

By classifying the interictal and pre-ictal periods in the CHB-MIT dataset, it can be considered as a binary classification problem. Evaluation metrics such as ACC, SP, RE, F1 score are used to predict the classification performance of the model, and the results of epilepsy prediction under different rhythms are shown in Figure 11.

状态	δ波	θ波	α波	β波	γ波
发作间期	弱	弱或正常	强	正常	正常
发作前期	局部增强	增强	减弱或阻断	增强（不显著）	增强
发作期	显著增强	增强（局灶性显著）	显著减弱或不规则、消失	增强（不显著）	高频异常放电
发作后期	显著增强	增强	恢复或较弱	减弱	减弱

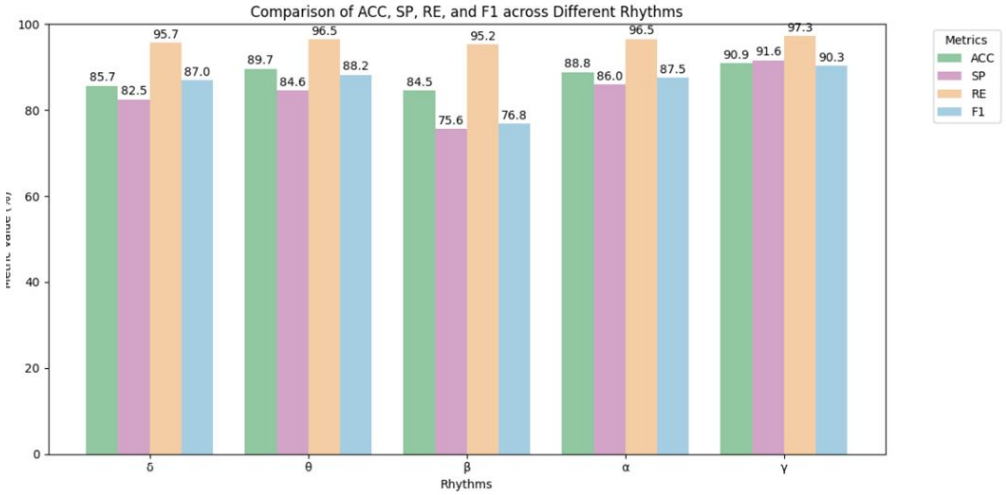


Figure 11: Seizure prediction performance at different rhythms

Correlation of Each Channel

Based on the above prediction of epileptic seizures by dividing different frequency bands, we conducted a phase synchronization analysis to study the functional connections and relationships between nodes in different channels of frequency band interactions, constructing a phase synchronization matrix of multi-channel brain areas to study the phase coupling relationships between different brain areas in EEG, in order to more comprehensively analyze the impact of multi-channel leads on predicting epileptic seizures. Each electrode serves as a node in the network, with the correlation between channels serving as connections between nodes. The results of the significant lead correlation analysis for the chb01 patient recorded in the CHB-MIT dataset are shown in Figure 12.

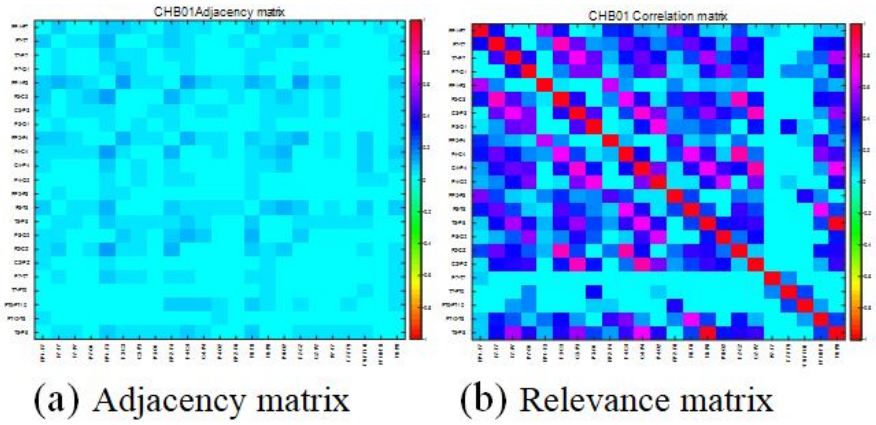


Figure 12: Significance analysis of chb01 patients

The adjacency matrix is used to connect different nodes of various channels. The numbers on the right indicate the heterogeneity between nodes. In the correlation matrix, different colors on the right represent different correlations. The higher the overall correlation, the stronger the overall coordinated activity of the brain in the current state. Epileptic seizures are often accompanied by abnormal high synchrony, especially showing high correlation between channels in the seizure focus and surrounding areas. This can help identify the propagation pathway of epileptic activity and locate the epileptic focus.

Result Analysis

Model Performance Validation

We calculate true positive (TP), false positive (FP), true negative (TN), and false negative (FN), and test the epilepsy EEG data samples using the above network model. The confusion matrix for the predicted classification of chb02 patients is shown in Figure 13, from which it can be visually seen that the values of TP and TN are large, indicating good performance in correct classification.

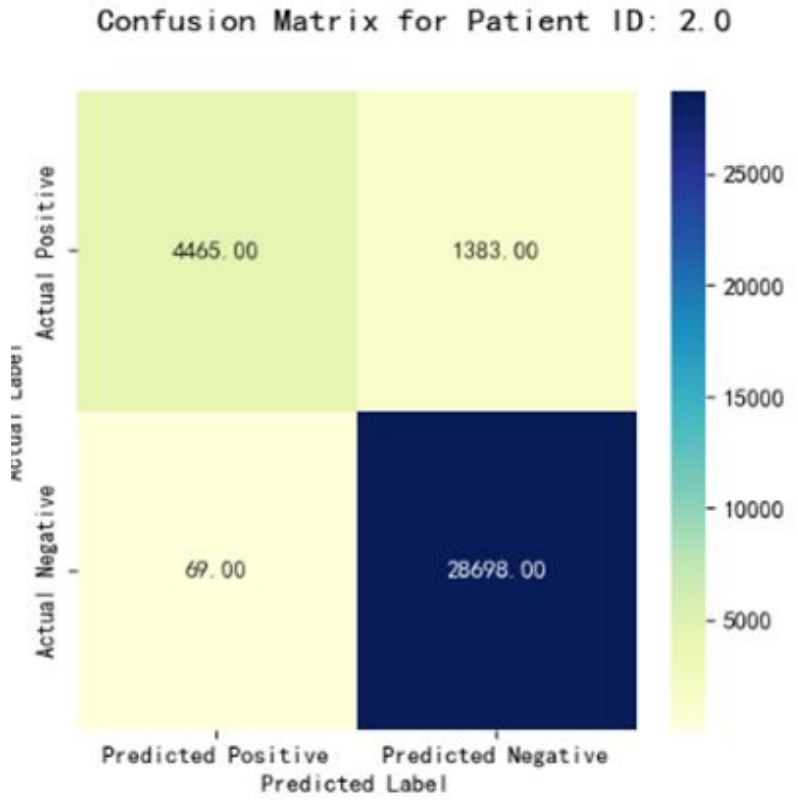


Figure 13: Confusion matrix for patient chb02

We divided the CHB-MIT data into three parts: seizure, seizure enhancement, and non-seizure, and input the preprocessed signal segments into the proposed MS-SACN network. In addition, we defined the 15 minutes before a seizure as the preictal period, and at least 2 hours after the end of a seizure as the interictal period. We evaluated the model performance by plotting ROC curves, with the false positive rate (FPR) and true positive rate (TPR) as the *x* and *y* axes, respectively.

Figure 14 (a) Shows the predicted ROC curve for patient chb08, with a larger corresponding AUC value indicating better model performance. From the graph, it can be seen that the AUC value is 0.9781, indicating a good predictive result. (b) Can be used as a warning for epileptic seizures. When the red curve shows abnormal enhancement, it can serve as a warning for an impending seizure. The bottom line corresponds to the corresponding seizure state, and a significant and sustained increase in the red curve indicates an ongoing seizure.

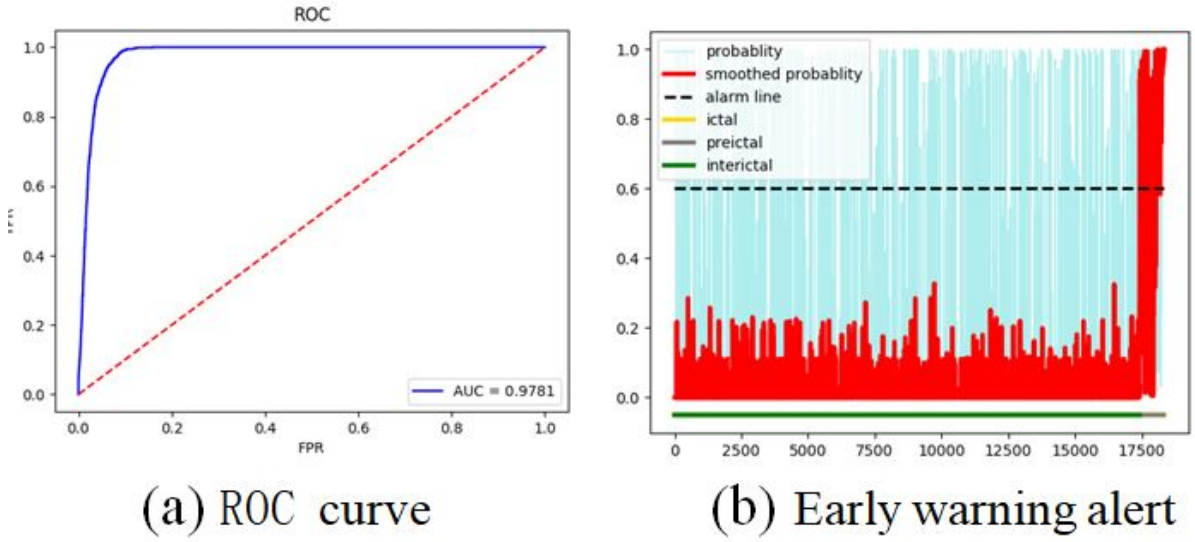


Figure 14: Predicted results of patient chb08

The proposed MS-SACN network constructs time, space, and frequency multidimensional feature extraction convolutional layers and introduces an attention mechanism for interactive fusion, which is a competitive method. We use leave-one-out cross-validation method to divide patients into training and test sets, and the results obtained are shown in Table 2, with AUC values of 0.999 for chb01, chb13, chb16, and chb23. The p-values in the table are used to measure the significance of the results, with values less than 0.005 indicating high confidence and significant results. Among them, chb22 may have a p-value less than 0.005 due to uneven sample distribution or complex noise. Comprehensive analysis shows that our model has strong predictive and robust performance.

In addition, existing methods such as TA-STS- ConvNet [18] and STFT+CNN [22] models based on convolutional neural networks, DCNN+Bi-LSTMM [23], GAM-RNN [24], MT-CRNN [25], BN-LSTM-CASA [26] models based on recurrent neural networks and their variants, CE-stSENet [27] based on attention mechanism, and BiConvLST-Attention3D[28]models have achieved good results in EEG feature extraction, classification, and other aspects in the time, frequency, and spatial domains. Table 3 shows the advantages of our method compared to the above methods.

Ablation Experiment

To better demonstrate the impact of the proposed innovative points on the prediction results of epileptic seizures, we conducted ablation experiments by isolating the Bidirectional Multi-Scale Time Convolutional Layer (Bi-Mul-CNN), Adaptive Spectral Convolutional Layer (DF-SFTF), Sparse Graph Convolutional Layer (SR-GNN), and Frequency Domain Feature Extraction Convolutional Layer for seizure prediction. The results of the experiment isolating the graph convolutional layer are shown in Figure 15, indicating that extracting graph structure information is crucial for predicting epileptic seizures. The results of the isolated experiments are shown in Table 4, and upon analysis, we found that effectively extracting spectral information is crucial for predicting epileptic seizures, while also constraining spatial characteristics.

Table 2: Prediction performance of CHB-MIT under the MS-SACN-MM model

Patient ID	AUC	Sn(%)	FPR/h	p
1	0.999	100.0	0.000	<0.001
2	0.958	98.2	0.003	<0.001
3	0.991	100.0	0.000	<0.001
5	0.963	98.8	0.001	<0.001
6	0.982	100.0	0.005	<0.001
7	0.895	89.4	0.102	0.001
8	0.978	100.0	0.000	<0.001
9	0.966	99.2	0.014	<0.001
10	0.992	100.0	0.000	<0.001
11	0.984	100.0	0.000	<0.001
13	0.999	100.0	0.106	<0.001
14	0.896	89.4	0.103	<0.001
16	0.999	100.0	0.000	<0.001
17	0.956	97.2	0.012	<0.001
18	0.962	100.0	0.000	<0.001
20	0.982	100.0	0.000	<0.001
21	0.926	90.2	0.100	<0.001
22	0.911	96.8	0.200	0.010
23	0.999	100.0	0.000	<0.001
aver	0.965	97.9	0.034	-

Table 3: Comparison of the performance of each model

method	AUC	Sn(%)	FPR/h
DCNN+Bi-LSTM	0.865	90.7	0.592
STFT+CNN	0.886	82.2	1.237
CE-stSENet	0.857	86.0	0.369
GAMRNN	0.917	88.1	0.053
TA-STS-ConvNet	0.918	96.7	0.072
BiCovLST-Att3D	0.920	97.8	0.107
MT-CRNN	0.948	91.7	-
BNLSTM-CASA	0.956	96.2	-
Ours	0.965	97.9	0.034

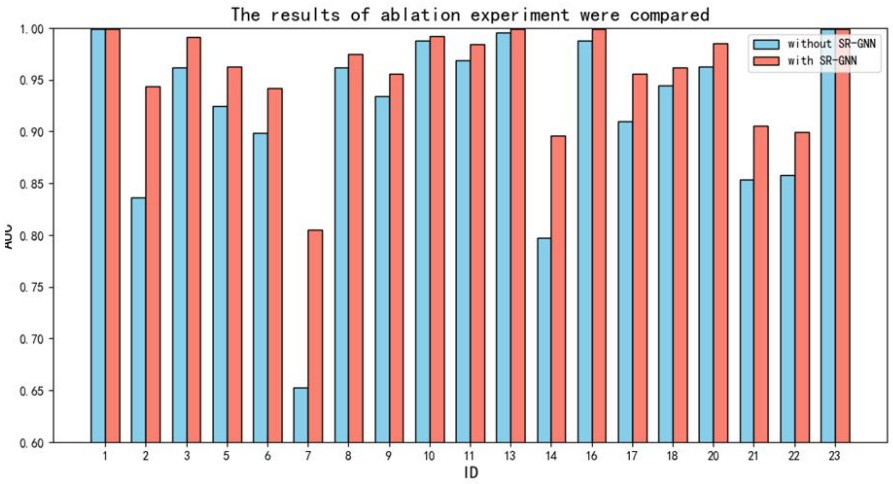


Figure 15: Isolated SR-GNN convolutional layer results

Table 4: Comparison of ablation experimental results

Dataset	Methods	AUC	Sn(%)	FPR/h
CHB-MIT	Without DF-SFTF	0.926	96.56	0.120
	Without SR-GNN	0.935	96.68	0.087
	Without Bi-Mul-CNN	0.943	97.95	0.052
	Ours	0.956	97.97	0.034

Conclusion

To address the problem of low prediction accuracy and high false alarm rate caused by the lack of significant multi-dimensional representation of EEG signals, this paper proposes a multi-domain feature interaction model based on multiple convolutional blocks and attention mechanism for accurate prediction of epileptic seizures. In the feature extraction part, different convolutional blocks are designed to meet the feature requirements in time, space, and frequency domains, obtaining significantly correlated features and reducing feature redundancy. In the feature fusion part, considering the representation of data in space and channels, the MLP-MultiAttention mechanism is proposed to comprehensively consider linear and non-linear relationship interactions to fuse multi-dimensional features. In the model implementation part, the focal loss training strategy is introduced to improve the model's generalization ability. Compared with other methods, our model shows significant improvement, and experimental results demonstrate that our method is highly competitive.

References

1. Al-Hadeethi H, Abdulla S, Diykh M, et al. (2020) Adaptive boost LS-SVM classification approach for time-series signal classification in epileptic seizure diagnosis applications[J]. Expert Systems with Applications, 161: 113676

2. Yi F, Zhong L, Li Z. Research on the Extraction Method of Spatiotemporal Features of Epileptic EEG Based on SVM Classifier. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 34: 444-50.

3. Vicnesh J, Hagiwara Y (2019) Accurate detection of seizure using nonlinear parameters extracted from EEG signals [J]. Journal of Mechanics in Medicine and Biology, 19: 1940004.

4. Zhou M, Cui H, Cao R, et al. (2017) Epileptic seizure prediction research based on permutation entropy and support vector machine. Application Research of Computers, 36: 1696-9.

5. SINGH K, MALHOTRA J (2022) Prediction of Epileptic Sei zures from Spectral Features of Intracranial EEG Record ings Using Deep Learning Approach[J]. Multimed Tools Appl, 81: 28875-98.

6. MANE R, CHEW E, CHUA K, et al. (2021) FBCNet: A multi-view convolutional neural network for brain-computer in--terface [EB/0L].

7. Avcu MT, Zhang Z, Chan D WS (2019) Seizure detection using least eeg channels by deep convolutional neural network[C]. ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP), 1120-4.

8. ANG X, ZHAO J, SUN, et al. (2021) An Effective Dual Self--attention Residual Network for Seizure Predic tion[J]. IEEE Trans Neural Syst Rehabil Eng, 29: 1604-13.

9. JIA MH, LIU WJ, DUAN JW, et al. (2022) Efficient Graph Convolutional Networks for Seizure Prediction Using Scalp EEG[J]. Front Neurosci, 16: 967116.

10. Wang M, Fang H, Gong, H, et al. Feature Extraction and Recognition of EEG Signals Using Multi-scale Hybrid Convolutional Networks. Journal of Huaqiao University (Natural Science), 44: 628-35.

11. Liu S, Wang X, Zhao L, et al. (2021) 3DCANN:a spatio-temporal convolution attention neural network for EEG emotion recognition [J]. IEEE Journal of Biomedical and Health Informatics, 26: 5321-31.

12. Kumar AJR, Bhanu B (2024) Uncovering hidden emotions with adaptive Multi-Attention Graph Networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 4822-31.

13. Liu J, Wu H, Zhang L, et al. (2022) Spatial-temporal transformers for EEG emotion recognition[C]//Proceedings of the 6th International Conference on Advances in Artificial Intelligence, 2022: 116-20.

14. RM, DU S, CHEN Z, et al. (2014) Study on Nonlinear Dynamic Characteristic Indexes of Epileptic Electroencephalography and Electroencephalography Subbands [J]. J Bio-medEng, 31: 18-22.

15. LIAN Q, QI Y, PAN G, et al. (2020) Learning Graph in Graph Convolutional Neural Networks for Robust Seizure Prediction [J]. J Neural Eng, 17: 035004.

16. Sheykhivand S, Rezaii TY, Mousavi Z, et al. (2020) Automatic identification of epileptic seizures from EEG signals using sparse representation-based classification[J]. IEEE Access, 8: 138834-45.

17. Chen J, Liu B, Lin W, et al. (2024) A review of time series pre-diction methods based on Transformer. Computer Science, 1-17.

18. Yu T (2023) Triple attention based spatio temporal spectral convolutional network for epileptic seizure prediction[J]. Authorea Preprints.

19. SHOEB A (2009) Application of Machine Learning to Epilep tic Seizure Onset Detection and Treatment[D]. Cam bridge: Massachusetts Institute of Technology, 2009: 27-156.

20. Ma L, Li S, Guo S, et al. (2024) Research on Epilepsy Prediction and Classification Optimized by PSO-LSTM. Journal of Tianjin University of Technology and Education, 34: 21-6.

21. PENG RM, JIANG J, KUANG GT, et al. (2022) EEG-based Automatic Epilepsy Detection: Review and Outlook [J]. Acta Autom Sin, 48: 335-50.

22. ND Truong et al. (2018) "Convolutional neural networks for seizure prediction using intracranial and scalp electroen--cephalogram," Neural Netw., 105: 104-11.

23. Daoud and MA Bayoumi (2019) "Efficient Epileptic Seizure Prediction Based on Deep Learning," IEEE Trans. Biomed. Circuits Syst., 13: 804-13.

24. Ji H, Xu T, Xue T, Xu T, Yan Z, Liu Y, Chen B, Jiang W (2023) An effective fusion model for seizure prediction: GAM-RNN. Front Neurosci. 17: 1246995.

25. L Duan, J Hou, Y Qiao, J Miao (2019) "Epileptic seizure prediction based on convolutional recurrent neural network with multi-timescale," in Proc. Int. Conf. Intell. Sci. Big Data Eng, 139-50.

26. M Maetaz (2021) Early prediction of epileptic seizure based on the BNLSTM CASAmodel," IEEE Access, 9: 79600-10.

27. Y Li, Y Liu, WG Cui, YZ Guo, H Huang, ZY Hu (2020) "Epileptic Seizure Detection in EEG Signals Using a Unified Temporal-Spectral Squeeze-and-Excitation Net-work," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 28: 4.

28. X Liu, J Li, M Shu (2020) "Epileptic seizure prediction based on region correlation of EEG signal,"in Proc. IEEE 33rd Int. Symp. Comput. Based Med. Syst. (CBMS), 120-5.