Multi-branch fusion graph neural network based on multi-head attention for childhood seizure detection

Yang Li[1], Yang Yang[1]*, Shangling Song[2]*, Hongjun Wang[1], Mengzhou Sun[3], Xiaoyun Liang[4], Penghui Zhao[1], Baiyang Wang[1], Na Wang[1], Qiyue Sun[1] and Zijuan Han[5]

[1]School of Information Science and Engineering, Shandong University, Qingdao, China, [2]Bidding Office, The Second Hospital of Shandong University, Jinan, China, [3]Institute of Research and Clinical Innovations, Neusoft Medical Systems Co., Ltd., Beijing, China, [4]Institute of Research and Clinical Innovations, Neusoft Medical Systems Co., Ltd., Shanghai, China, [5]Center for Optics Research and Engineering, Shandong University, Qingdao, China

The most common manifestation of neurological disorders in children is the occurrence of epileptic seizures. In this study, we propose a multi-branch graph convolutional network (MGCNA) framework with a multi-head attention mechanism for detecting seizures in children. The MGCNA framework extracts effective and reliable features from high-dimensional data, particularly by exploring the relationships between EEG features and electrodes and considering the spatial and temporal dependencies in epileptic brains. This method incorporates three graph learning approaches to systematically assess the connectivity and synchronization of multi-channel EEG signals. The multi-branch graph convolutional network is employed to dynamically learn temporal correlations and spatial topological structures. Utilizing the multi-head attention mechanism to process multi-branch graph features further enhances the capability to handle local features. Experimental results demonstrate that the MGCNA exhibits superior performance on patient-specific and patient-independent experiments. Our end-to-end model for automatic detection of epileptic seizures could be employed to assist in clinical decision-making.

1 Introduction

Epilepsy is a neurological disorder characterized by abnormal synchronous discharges of neurons. Childhood seizures carry a risk for the presence of cognitive impairment and behavioral disorders (Rennie et al., 2004). Therefore, accurate detection of seizures in children is of paramount importance for determining the best treatment plans and preventing adverse conditions. The diagnosis of epilepsy typically relies on the analysis of electroencephalogram (EEG), which is abnormal in the majority of patients. However, this task requires highly experienced experts who must invest a significant amount of time and effort in inspecting lengthy

EEG recordings (Kannathal et al., 2005). This process is also susceptible to the subjective influence of epilepsy experts. The occurrence of epileptic seizures significantly impacts the physical development of children, emphasizing the necessity for early detection and intervention in pediatric epilepsy to mitigate its effects. Therefore, designing a reliable childhood seizure detection model can facilitate the automation of epilepsy diagnosis, holding significant importance in enhancing the quality of life for children.

There is a significant need to expand the utilization of machine learning, particularly within the emerging realm of deep learning, for automating the detection of epilepsy through EEG signal classification. Among the extensively explored techniques for EEG seizure detection, feature extraction rooted in machine learning stands as one of the most prominent approachs. For instance, multiscale entropy were extracted using extreme learning machines (Cui et al., 2017), nonlinear features were extracted using Gradient Boosting Decision Trees (GBDT) (Xu et al., 2022), various features were extracted using Empirical Mode Decomposition (EMD) (Singh et al., 2019), multi-scale features were extracted using wavelet transformation (Zhang et al., 2010), and the Comprehensive Representation of K Nearest Neighbors (CRMKNN) (Na et al., 2021) approach is proposed for epilepsy diagnosis. The traditional machine learning algorithms often struggle to achieve automatic detection of epilepsy, and experimental results are influenced by empirical parameters, making it difficult to stabilize the algorithm's performance. In essence, time series data is nonlinear and dynamic, making it challenging for traditional machine learning algorithms to effectively capture these complex signal characteristics. Moreover, there are large variations between different patients. Therefore, traditional machine learning faces challenges in learning the hidden features of EEG signals and lacks generalization.

Compared to traditional machine learning, neural networks, as a more promising algorithm with greater capacity for learning from complex data, have been applied to various research fields (Zhao P. et al., 2022; Sun and Yang, 2023; Chen J. et al., 2024; Abu and Diamant, 2023). There have been numerous advancements in the detection of epilepsy EEG signals as well (Wang et al., 2023a; Zhao et al., 2023b; He et al., 2022; Xiao et al., 2024). A novel deep network called Two-Stream 3-D Attention Module (TSA3-D) (Cao et al., 2022) was introduced to leverage the multichannel time-frequency and frequency-space features of interictal EEGs for epilepsy classification. In (Feng et al., 2022) a 3D deep network combined with residual attention modules was proposed to explore the spatial and time-frequency features of multi-channel EEG. In (Cui et al., 2022b), a fusion model based on transfer learning and time-frequency features was proposed for the effective detection of childhood epilepsy syndrome. In (Cui et al., 2022a), an analysis of the correlation between time-frequency features and EEG signals was conducted, and a childhood epilepsy syndrome classification model based on transfer networks was proposed.

Attention mechanisms have gained widespread applications in the field of signal recognition (Zhao et al., 2023b; Lian and Xu, 2023; Peh et al., 2023; Qiu et al., 2023; Wang Z. et al., 2023; Liu et al., 2023), emerging as a pivotal technology attracting significant attention and in-depth exploration within the realm of deep learning. Many researchers have integrated attention mechanisms with neural network models, resulting in the creation of a series of innovative models. The emergence of these models has introduced new possibilities for enhancing the accuracy and efficiency of EEG signal recognition, thus steering the direction of development in this field. Ding et al. (2023) proposed a novel seizure prediction model that utilizes a CNN to automatically capture features from EEG signals. This model combines multiple head attention mechanisms to identify relevant information within these features for the recognition of EEG signal segments. Deng et al. (2023) introduced a novel hybrid vision transformer (HViT) model that could enhance the multi-head attention mechanism by augmenting the capability of convolution to process local features, thereby achieving data uncertainty learning. Zhao et al. (Zhao X. et al., 2022) proposed a recommendation detector based on multi-head attention mechanism, utilized for detecting pathological high-frequency oscillations (HFOs) associated with epilepsy to locate the epileptogenic zones. The attention mechanism aids the network in capturing dependencies among features and enhancing the model's sensitivity to local information. In the field of seizure detection, the potential of attention mechanism remains to be further explored.

In the epilepsy detection process using deep learning, EEG signals are represented as two-dimensional signals, considering only channel-based features and disregarding information about the physical distribution of channels. The electrode distribution in EEG exhibits a non-Euclidean topological structure, which can lead to the loss of connectivity information between brain functional regions, neglecting the long-term interdependencies among EEG signals from different channels. The graph convolutional network (GCN) algorithms can effectively leverage the implicit graph representation information within EEG signals. GCN algorithms utilize graph structures and update graph representations through node aggregation. GCN has been widely applied in numerous EEG signals, which have demonstrated excellent performance, such as emotion recognition (Song et al., 2021; Liu et al., 2022; Chen Y. et al., 2024; Li Y. et al., 2022), Alzheimer's disease (Lopez et al., 2023), automatic seizure detection (Wagh and Varatharajah, 2020; Meng et al., 2022; Ho and Armanfard, 2023), driver state monitoring (Kalaganis et al., 2020), motor imagery (Cai et al., 2022), and sleep stage classification (Li M. et al., 2022; Jia et al., 2020; Lee et al., 2024; Ji et al., 2022; Jia et al., 2021). Wang et al. (2022) proposed a spatiotemporal graph attention network (STGAT) based on phase locking value (PLV) to extract spatial and functional connectivity information. Raeisi et al. (2022) constructed graph representations using three different types of spatial information and assessed the performance of neonatal seizure detection. He et al. (2022) utilized the graph attention network (GAT) to extract spatial features and employed a bi-directional long short-term memory network (BiLSTM) to capture temporal relationships before and after the current time frame for epilepsy detection.

The integration of attention mechanisms with GCN has proven effective in enhancing model performance on graph-structured data (Wu et al., 2024; Li et al., 2023; Cheng et al., 2023; Dong et al., 2022; Wang Y. et al., 2023; 2020; Grattarola et al., 2022). Attention helps highlight important nodes or features, improving the ability of GCN to capture both global and local relationships, which

is crucial for EEG classification. The dynamic temporal graph convolutional network (DTGCN) (Wu et al., 2024) is proposed for seizure detection and classification, incorporating a seizure attention layer to capture the distribution patterns of epilepsy and a graph structure learning layer to represent the dynamically evolving graph structure in the data. A spatiotemporal hypergraph convolutional network (STHGCN) (Li et al., 2023) is designed to capture higher-order relationships in EEG recordings, construct feature hypergraphs across the spectral, spatial, and temporal domains to focus on EEG channel correlations and dynamic temporal relationships, and integrate self-attention mechanisms to initialize and update relationships within EEG sequences. A hybrid network (Cheng et al., 2023) is proposed, consisting of a Dynamic Graph Convolution (DGC) module and a Temporal Self-Attention Representation (TSAR) module. This network simultaneously integrates representative knowledge of spatial topology and temporal context into the EEG emotion recognition task. In summary, the integration of attention mechanisms with GCN allows for more effective feature representation by dynamically weighting the importance of spatial and temporal relationships, thereby enhancing the ability of model to capture complex dependencies in EEG data for emotion recognition tasks.

During the training process, traditional GCN describes the dynamic process of epileptic seizures using a single graph representation. During seizures, various complex interactions of neural activities occur, including different types of brainwave changes and alterations in connectivity patterns between brain regions. Therefore, the dynamic process of epileptic seizures cannot be exhaustively represented by a single static graph. To comprehend and describe seizures, it is necessary to consider the changes in the temporal dimension, and the interactions between different brain regions in the spatial dimension. Therefore, this study introduces a multi-branch graph convolutional model with multi-head attention (MGCNA) for childhood seizure detection. Specifically, the MGCNA employs three graph representation approaches to characterize the feature representation of EEG data from multiple dimensions. By incorporating a multi-head attention mechanism, it combines spatial topological information from multi-channel electrodes with dynamic temporal information, enhancing the global contextual awareness and recognition capabilities of model. The major contributions of MGCNA can be summarized as follows:

- A multi-branch GCN model is proposed. It utilizes Euclidean distance to capture spatial information between channels, employs Pearson correlation coefficient to gather functional connectivity information among channels, and supplements latent spatiotemporal correlations through a trainable adjacency matrix. The multi-channel EEG signals are modeled as graph signals, enabling the extraction of synchrony relationships within EEG signals.
- By integrating graph signals with a multi-head attention mechanism, attention weights for graph features are obtained, and the hidden vector representation of graph signals is derived through the summation of weighted values.
- We conduct patient-specific and patient-independent experiments to assist doctors in rapidly identifying the onset period in complex scenarios. The outcomes from these two experiments showcase the effectiveness of our method. Through

comparison with other methods, the sensitivity, which is of utmost clinical concern, is the highest in two experiments.

The organization of the rest of this article is as follows. In Section 2, the proposed seizure detection model is introduced in terms of the overall MGCNA structure, extraction of intrachannels features, multi-branch GCN, multi-head attention mechanism, and classifier module. Section 3 presents the implementation details, performance evaluation metrics and experimental results of the CHB-MIT dataset. Section 4 interprets the results and emphasizes the limitations of the current study, and the conclusion is provided in Section 5.

2 Methods

In this section, we introduce our proposed MGCNA model for epilepsy classification using multi-channel EEG signal.

2.1 General structure

Figure 1 provides an overview of the overall process for classifying epileptic EEG signals using the MGCNA proposed in this article. First, intrachannels features are extracted from the raw EEG signals to create graph signals. Second, three different time-series-to-graph representation are employed and graph features are extracted through GCN. Third, a multi-head attention is utilized to learn dependencies of different graph features. Finally, the combined graph features are processed to output classification results.
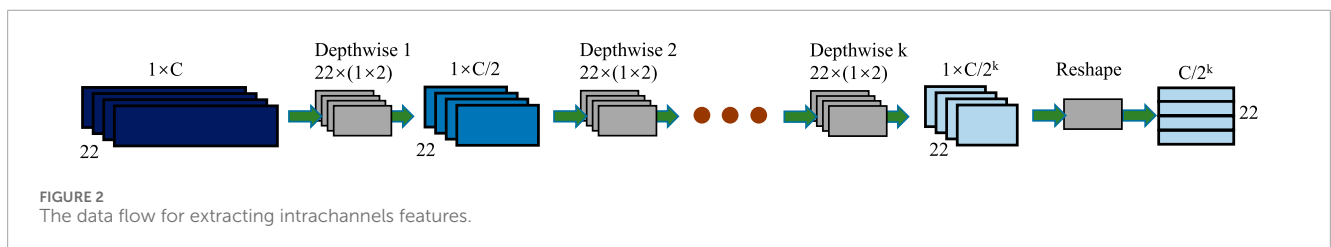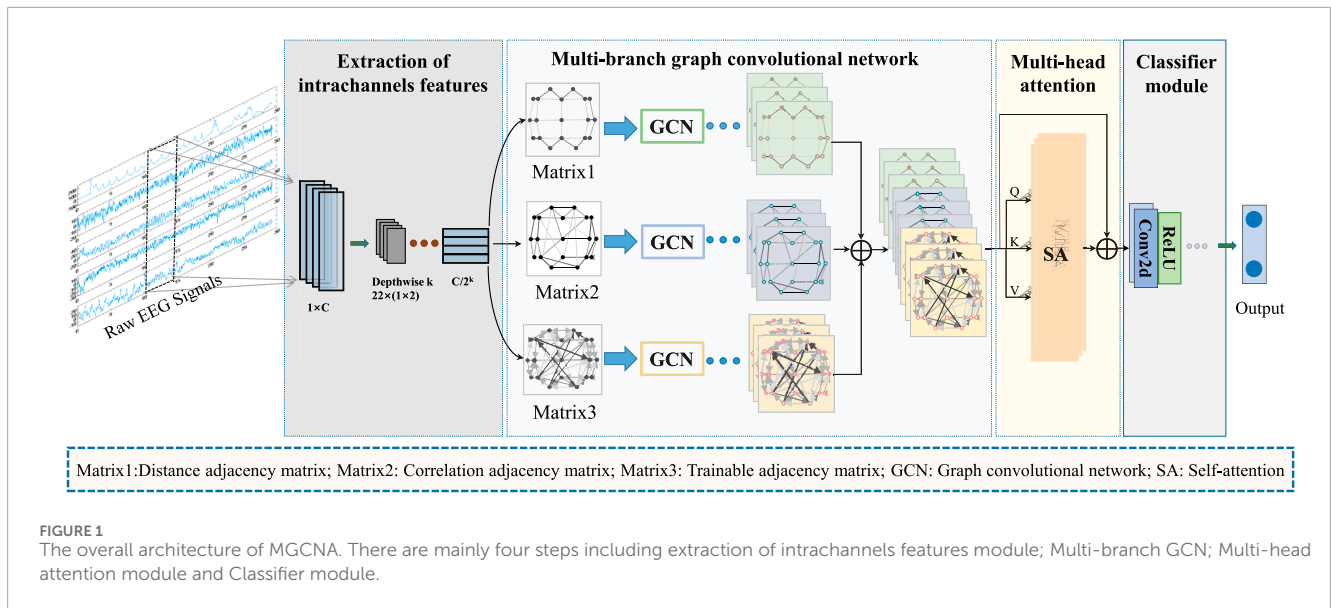
2.2 Extraction of intrachannels features

Convolution has been proven to be highly effective in capturing features of EEG signals. Inspired by the EEGWaveNet (Thuwajit et al., 2021) model, this module utilizes depthwise convolution to compress the input signals' resolution in a channel-wise manner and capture features. Depthwise convolution does not cross information between channels and extracts features independently. This module comprises k consecutive layers, where each layer captures valuable features within channels at half the scale of the resolution of the previous layer.

In this module, the input data X are shaped as $(n, 1, C)$, where n and C represent the number of channels and the number of sampling points, respectively. n is set to 22, and C is defined as $f \times t$, where f denotes the EEG's sampling frequency, and t is the duration of each epoch's segmentation. In each layer, depthwise convolution is applied to each channel with a kernel size of $(1, 2)$, a stride of 2, and no padding. The output size of each kth layer is $(n, 1, C/2^k)$.

The module of extraction of intrachannels features can be represented as:

$$F^k = DWConv\left(F^{k-1}\right) \tag{1}$$

where F^0 is the input signal $X \in \mathbb{R}^{n \times 1 \times C}$, $DWConv(\cdot)$ is the depthwise convolution. After reshape, the features extracted by this module $N \in \mathbb{R}^{n \times (C/2K)}$ will be fed into the next module to obtain multi-scale graph features. Figure 2 depicts the process of extracting intrachannels features.

FIGURE 1
The overall architecture of MGCNA. There are mainly four steps including extraction of intrachannels features module; Multi-branch GCN; Multi-head attention module and Classifier module.

FIGURE 2
The data flow for extracting intrachannels features.

2.3 Multi-branch graph convolutional network

To design an appropriate graph representation for EEG signals, it is necessary to construct a graph based on the physical distribution of channels and the spectral-temporal domain characteristics of the signals. Each channel was considered as a node and the connection of each node is characterized in terms of spatial position information, EEG signal similarity features, and a learnable graph representation to dynamically capture the topological relationship between EEG signal channels. These three types of graph signals are input into GCN, allowing us to leverage the structural characteristics of EEG signals to capture spatial dependencies between nodes. Figure 3 illustrates the overall process of multi-branch GCN. The graph convolution processes at the top, middle, and bottom are based on spatial distance graph representation, functional connectivity graph representation, and adaptive graph representation, respectively.

2.3.1 Graph generator

Graph representation consists of graph signal and graph. Graph signals are extracted from extraction of intrachannels features module. A graph is denoted as $G = \{\mathcal{V}, \epsilon, A\}$, where $\mathcal{V}$ is the set of nodes, ϵ is the set of edges, and A is the adjacency matrix. Graph nodes represent EEG channels. The connections between nodes are controlled by the adjacency matrix A. To construct

the adjacency matrix A, this paper employs three graph learning methods: information based on spatial distance between nodes, measure based on functional connectivity, and adaptive graph learning methods.

2.3.1.1 Spatial distance graph learning

Seizures are produced by abnormal nerve discharges in different areas of the cerebral cortex and cause significant changes in the EEG. And this abnormal discharge spreads and affects neighboring brain regions, making strong correlation between adjacent areas. So the spatial distance between nodes of EEG signals is used for graph representation in EEG signal analysis.

In this approach, we consider the electrode positions of the EEG as a three-dimensional grid model. To represent the spatial distance of EEG electrodes, we compute adjacency matrix $A1_{ij}$ by applying Euclidean distance between pairs of channels in the bipolar montage. The distance between pairs of channels $v_i(x_i, y_i)$ and $v_j(x_j, y_j)$, which can be denoted as:

$$Dist_{ij} = sqrt\left(\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2\right) \tag{2}$$

where $v_i(x_i, y_i)$ and $v_j(x_j, y_j)$ are the centers of the bipolar derivations. The spatial distances obtained from Euclidean distance are transformed into the adjacency matrix while taking into account the strength of connections between nodes. To calculate the adjacency

FIGURE 3
The pipeline of multi-branch graph convolutional network.
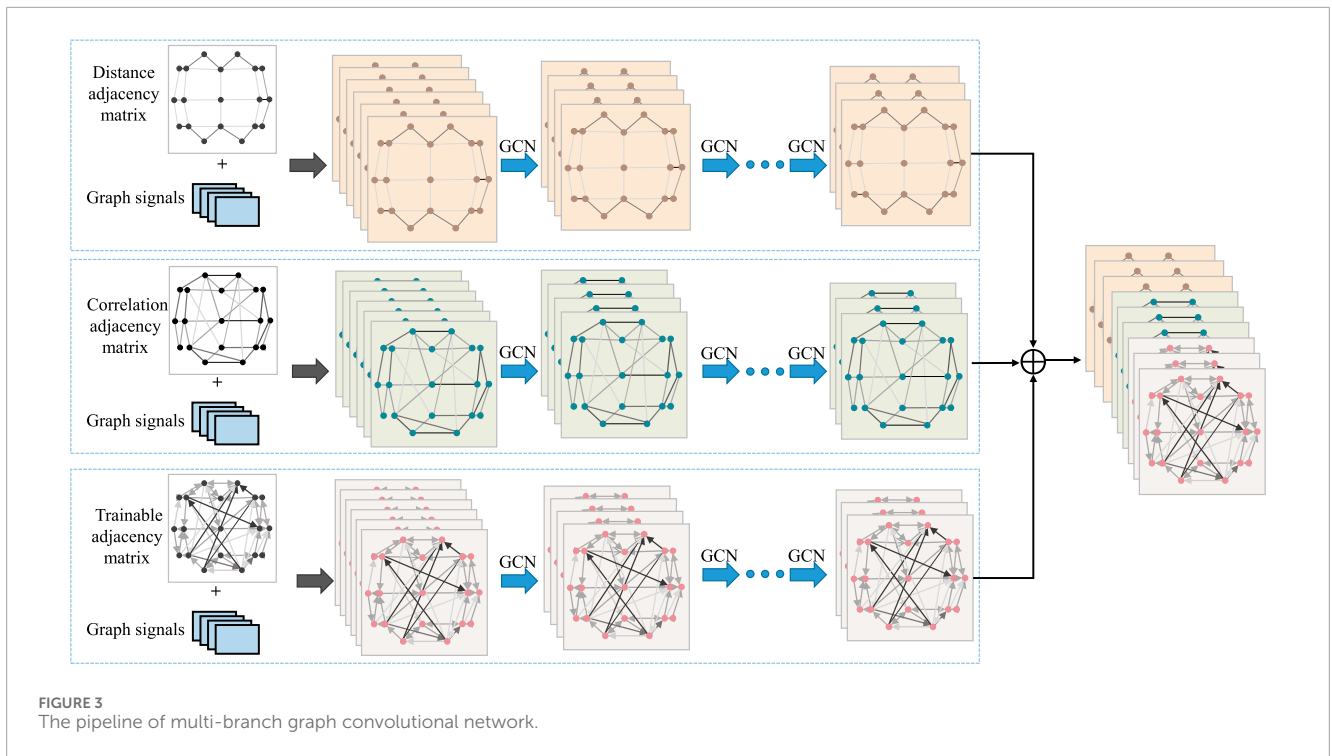
matrix, we apply the following rules:

$$
A1_{ij} = \begin{cases} exp\left(-\dfrac{Dist_{ij}^2}{\delta^2}\right), & if\ Dist_{ij} \le R_1 \\ 0, & otherwise \end{cases} \quad (3)
$$

where R_1 is the threshold for sparsity and δ is the standard deviation of the distances. This results in a universally undirected, weighted graph for all EEG graph representation.

2.3.1.2 Functional connectivity graph learning

The strength of association between channels varies during seizures. and the synchronization between brain regions can be assessed by calculating the connectivity between channels. Therefore, we investigate the relationship of information interaction between channels during seizures.

The functional connectivity between nodes of EEG signals is used for graph representation in EEG signal analysis. To represent the functional connectivity of EEG electrodes, we compute adjacency matrix $A2_{ij}$ by applying Pearson correlation coefficient between the preprocessed signals in v_i and v_j. The Pearson correlation coefficient between pairs of channels v_i and v_j, which can be denoted as:

$$
\rho_{ij} = \frac{\sum \left(X_i - \overline{X_i}\right)\left(X_j - \overline{X_j}\right)}{\sqrt{\sum \left(X_i - \overline{X_i}\right)^2 \sum \left(X_j - \overline{X_j}\right)^2}} \quad (4)
$$

where X_i and X_j are graph signals extracted from extraction of intrachannels features module in v_i and v_j. The computed correlation coefficients are normalized to the range [0, 1]. The functional connectivity obtained from Pearson correlation coefficient are transformed into the adjacency matrix while taking into account the

strength of connections between nodes. To calculate the adjacency matrix, we apply the following rules:

$$
A2_{ij} = \begin{cases} \rho_{ij}, & if\ \rho_{ij} \ge R_2 \\ 0, & otherwise \end{cases} \quad (5)
$$

where R_2 is the threshold. This results in an undirected, weighted graph for all EEG graph representation.

2.3.1.3 Adaptive graph learning

The human brain possesses extremely complex structure and functionality, and analyzing it merely through spatial distance and functional connectivity graph structures does not fully capture its functions and behavioral manifestations. The neural network model aims to simulate the interconnection and information transfer between neurons in the human brain, enabling learning and reasoning of complex tasks. Therefore, we incorporate graph structures as part of the parameters in neural networks, training them to capture the coupling relationships with EEG signals.

The adaptive graph learning method can learn the intrinsic connections between EEG signals. In the adaptive graph learning method we employed, the adjacency matrix $A3 \in \mathbb{R}^{n \times n}$, which characterizes the relationships between individual vertex nodes, is dynamically learned rather than predetermined. During the model training process, the adjacency matrix $A3$ of the adaptive graph representation is updated during training through back propagation.During the training process, $\mathbf{A}_3$ will be constrained by the following formula:

$$
\mathbf{A}3^l = ReLU\left(BN\left(\mathbf{a}3^l\right)\right) \quad (6)
$$

where $\mathbf{a}3^l$ is the adjacency matrix obtained in the l-th layer, using BN and ReLU to prevent overfitting and improve stability. Due to

the fact that the closer spatial connection between channels does not accurately reflect a closer functional relationship between them, and functional connectivity does not fully characterize the intrinsic relationship between channels (Song et al., 2018), we utilize the adaptive graph learning method to capture the inherent connections among EEG channels and enhance EEG recognition.

2.3.2 Graph convolution

A general GCN model (Kipf and Welling, 2016; Defferrard et al., 2016) takes graph signals $N \in \mathbb{R}^{n \times (C/2K)}$ and adjacency matrices $A \in \mathbb{R}^{n \times n}$ as inputs and a graph feature matrix $Z \in \mathbb{R}^{M \times P}$ is generated through node aggregation. P represents the dimension of the output feature vector for each node. GCN updates node features by aggregating features from neighboring nodes. The graph convolutional layer can be represented as follows:

$$H^l = ReLU\left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{l-1} \theta^l\right) \tag{7}$$

where $\tilde{A}$ represents adjacency matrix containing diagonal matrix. $\tilde{D}$ denotes the degree matrix for each node in the graph. H^{l-1} represents the graph feature matrix of the previous output, and $H^0 \in \mathbb{R}^{n \times (C/2K)}$ is the graph signal N. θ^l represents the matrix of learnable parameters.

Notably, the structures of the three GCN branches are identical, but their parameters are not shared. After each GCN, Batch Normalization (BN) is applied to normalize the graph features at each layer, which helps improve the generalization capability of network. And it is presented as the following form:

$$\mathcal{Z}_i^l = BN\left(H_i^l\right) \tag{8}$$

where z_i^l is the input of the next graph convolution layer. i indicated as the ith branch of the GCN. The final output of multi-branch GCN is expressed as the following formula:

$$\mathcal{Z} = [z_1^m, z_2^m, z_3^m] \tag{9}$$

where m represents the number of layers of the GCN, and $[\cdot]$ represents concatenation.

2.4 Multi-head attention mechanism

The attention mechanism, which is a deep learning model, emulates the pivotal information and critical elements that individuals concentrate on during observations. The multi-head attention mechanism makes the output of attention to incorporate encoding from different spatial locations, thereby enhancing the expressive power of the model. In standard multi-head attention with input matrix $Z \in \mathbb{R}^{m \times n \times D}$, the self-attention for each head are computed according to the following formula:

$$Head(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d}}\right) V \tag{10}$$

where Q, K, V represent the query vector, key vector and value vector, respectively. h is the number of heads, $d = m/h$. The original multi-head attention mechanism has a high computational cost, with a quadratic complexity of $\mathcal{O}(L^2 d)$, where L is $n \times D$.

Inspired by MobileViT v2 (Mehta and Rastegari, 2022) and HViT-DUL (Deng et al., 2023), the multi-head design

is incorporated into the separable self-attention mechanism, which is applied to the proposed model for the reduction of computational overhead. The separable multi-head attention mechanism employed in this paper has a linear complexity of $\mathcal{O}(L)$, which is lower compared to the standard multi-head attention mechanism. When multiple self-attention mechanisms are sequentially linked and subsequently subjected to a final projection, it results in the generation of the ultimate values for the multi-head attention mechanism. The spatial topological information and dynamic variation information of multi-channel electrodes are obtained through a multi-head attention mechanism to combine representations of multiple graph signals, thereby enhancing the discrimination ability. Figure 4 illustrates the process of separable multi-head attention. Q, K, V is calculated using the following formula:

$$\begin{cases} Q = Conv_1(\mathcal{Z}) \\ K = Conv_2(\mathcal{Z}) \\ V = Conv_3(\mathcal{Z}) \end{cases} \tag{11}$$

where $\mathcal{Z}$ is the output of multi-branch GCN features. $Conv$ is two-dimensional convolution. We employ three 1×1 convolutional layers, with parameters not shared, to compute the Q K V of the input graph features. After the reshaping process, the outputs Q, K, V are in the following dimensions: $Q \in \mathbb{R}^{h \times 1 \times n \times D}$, $K \in \mathbb{R}^{h \times d \times n \times D}$ and $V \in \mathbb{R}^{h \times d \times n \times D}$.

To calculate self attention scores, we apply the $Softmax$ operation to the Q and taking the element-wise multiplication with K, After that, we perform element-wise multiplication with V after passing it through ReLU. Self attention scores $Atten_i$ are computed based on the following formula:

$$atten_i = ReLU(V)^* (K^* Softmax Q) \tag{12}$$

where $atten$ is the self-attention scores. $ReLU$ is activation function. After concatenating the scores of multiple self-attention mechanisms, we perform a 3×3 convolution on the resulting features, normalize it using BN, and then apply ReLU to enhance generalization ability. The computed multi-head attention scores, matching the dimensions of the input $\mathcal{Z}$, are added to the output of multi-branch GCN, expressed as:

$$Atten = [atten_1, atten_2, \ldots, atten_h] \tag{13}$$

$$MA = ReLU(BN(Conv(Atten))) \tag{14}$$

$$\mathcal{Z}' = BN(\mathcal{Z} + MA) \tag{15}$$

where $[\cdot]$ represents the concatenation.

2.5 Classifier module

A classifier module is placed after the multi-head attention mechanism for the final class inference. Initially, it involves two layers of 2D convolution, with a $ReLU$ activation function applied after each convolution to enhance the robustness of model. This process can be described as:

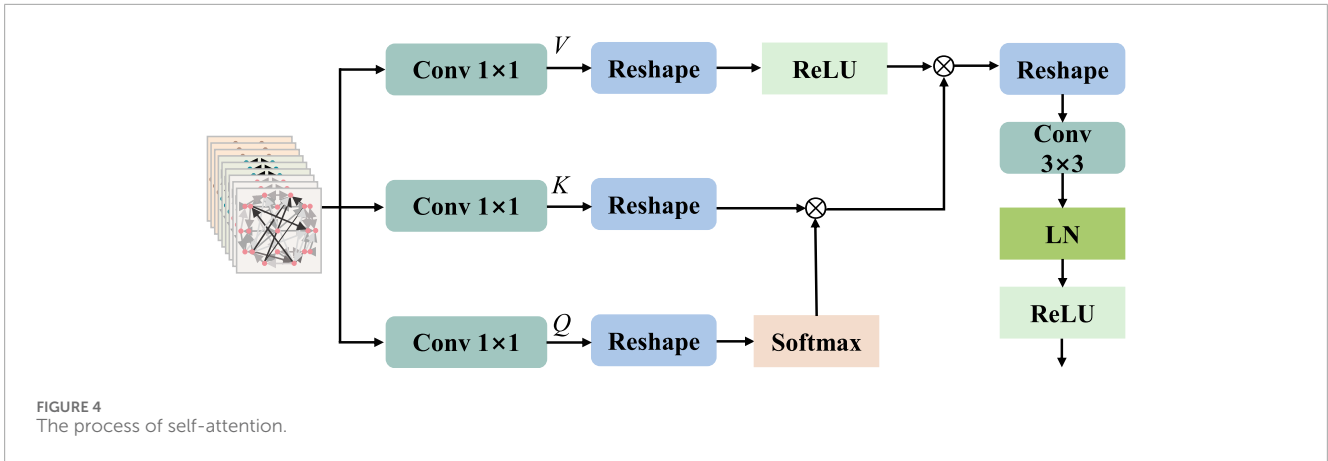$$\mathcal{F}' = ReLU(Conv(ReLU(Conv(\mathcal{Z}')))) \tag{16}$$

FIGURE 4
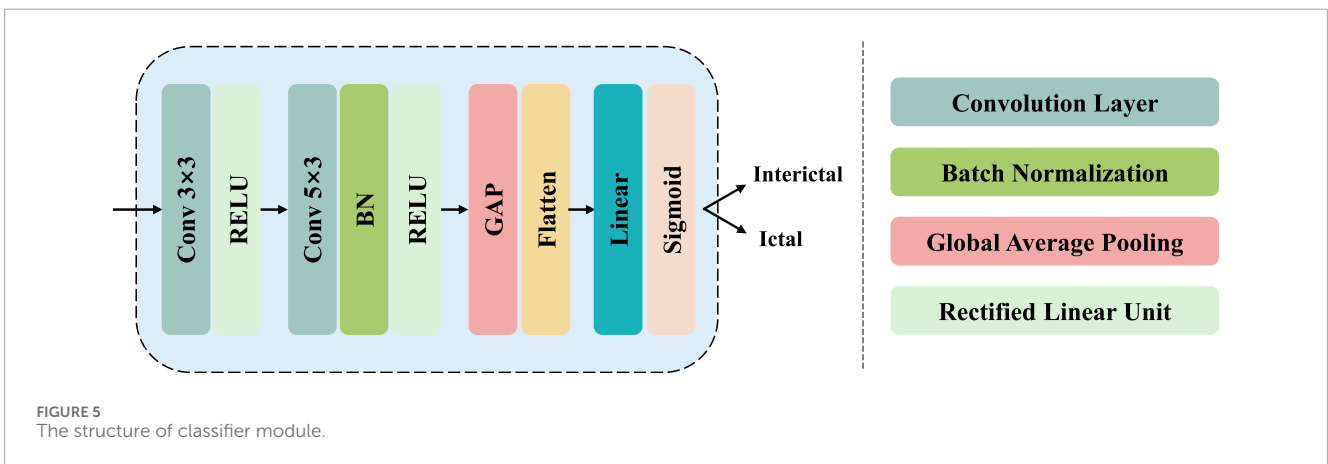The process of self-attention.

FIGURE 5
The structure of classifier module.

where $\mathcal{Z}'$ is the output of multi-head attention mechanism. Global average pooling pools each channel of the convolution result, followed by a linear layer that reduces the feature dimensionality to match the number of classes. After passing through a sigmoid operation, the index with the highest probability is selected as the classification result. This process can be described as:

$$Pre^s = \sigma\big(Linear\big(AvgPool\big(\mathcal{F}'\big)\big)\big) \qquad (17)$$

where $AvgPool(\cdot)$ denotes global average pooling, σ is sigmoid operation, $Linear(\cdot)$ is the Linear operation, and Pre^s is the final classification result. Figure 5 illustrates the structure of the classifier module.

2.6 Training procedure

In order to achieve optimal network parameters during the training process, we employ the backpropagation (BP) algorithm, which iteratively updates the network parameters until an optimal or suboptimal solution is reached. In this context, we introduce a loss function based on binary cross entropy, defined as follows:

$$loss = cross\_entropy\big(Pre, Pre^s\big) + \tau\|\Omega\| \qquad (18)$$

$$cross\_entropy\big(Pre, Pre^s\big) = -\Big[Pre \log Pre^S + (1 - Pre)\log\big(1 - Pre^S\big)\Big] \qquad (19)$$

where Pre and Pre^s represent the actual label and the predicted label, respectively. The binary cross entropy, denoted as $cross\_entropy(Pre, Pre^s)$, quantifies the disparity between the true labels and the predicted labels. τ represents the trade-off parameter, while $\|\cdot\|$ denotes the $l2$-norm to prevent overfitting. Ω represents all the parameters within this model.

For adaptive graph learning, the adjacency matrix A3 is a trainable parameter within the network that optimizes with model optimization. The partial derivatives of the loss function with respect to the optimal adjacency matrix A3 and the loss are expressed as follows:

$$\frac{\partial loss}{\partial A} = \frac{cross\_entropy\big(Pre, Pre^s\big)}{\partial A} + \tau\frac{\|\Omega\|}{\partial A} \qquad (20)$$

2.7 Dataset

The CHB-MIT dataset (Shoeb, 2009) was collected by Boston Children's Hospital and is currently the most widely used public dataset. The multichannel scalp EEG signals consists of EEG signal recordings from 23 children with epilepsy at a sampling rate of

256 Hz. The EEG signals are collected using EEG electrodes placed according to the International 10–20 system. The dataset comprises approximately 983 h of EEG recordings, including 198 seizure onset events, with a total duration of 3 h 15 min.

2.8 Implementation details

In the CHB-MIT dataset, most records have 23 channels, while some records have missing or duplicated channels. In this study, 22 channels are selected in this study to maintain consistency of channels among all patients. The selected 22 channels are FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, FZ-CZ, CZ-PZ, P7-T7, T7-FT9, FT9-FT10, FT10-T8. This dataset is classed into two categories, including interictal and ictal. To reduce the effect of noise, a fifth-order Butterworth band-pass filter ranging from 0.5 Hz to 70 Hz was used. The filtered EEG signals are segmented into 3-s windows. Due to the scarcity of ictal EEG data compared to interictal EEG data, the windows overlap by 50% to obtain more ictal data.

For a comprehensive evaluation of the MGCNA, we train both a patient-independent model and a patient-specific model. In a patient-specific approach, the model is trained, validated, and tested using data from an individual subject. Interictal signals are randomly discarded for each subject, and the ratio of interictal signals to ictal signals is maintained at 5:1. As the number of ictal signals is significantly lower than that of interictal signals, which is detrimental to model training, we employ re-sampling to create a balanced training dataset. In each epoch of training, a random selection of an equal number of interictal signals is made to match the ictal signals in the training set, followed by random shuffling. Consequently, we obtain a balanced training subset for each epoch. We employ a 5-fold cross-validation approach and report the average performance. In a patient-independent approach, the model was trained and validated using data from multiple subjects, and then tested on data from individual subject. We utilized the leave-one-subject-out cross-validation as evaluation method. The data in both categories are kept balanced. The patient-specific approach emphasizes individual differences and personalization, and the patient-independent approach focuses more on overall trends and general patterns.

We employ the Adam optimizer with a learning rate of 3e-4, a weight decay of 1e-3.The dropout rate is experimentally set to 0.5. R_1 is set to 0.4 and R_2 is set to 0.25. k is set to 2. All the above experiments were performed and implemented by Pytorch 1.7.1 in the NVIDIA GTX3090 and CUDA11.0 environment.

2.9 Performance evaluation metrics

We use eight different performance metrics to evaluate the performance of MGCNA including Accuracy (Acc), Sensitivity (Sen), Specificity (Spe), $F1$ Score ($F1$), and Area Under Curve (AUC). These metrics are obtained using true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) values.

Accuracy represents the proportion of samples correctly classified by the model out of the total sample count. Sensitivity and Specificity measure the model's ability to classify positive and negative samples. F1 Score combines precision and recall to provide a comprehensive performance measure. AUC is based on the area under the ROC curve and provides a comprehensive performance metric for unbalanced category distributions.

3 Experimental results

3.1 Patient-specific experiments

Patient-specific experimental results on CHB-MIT dataset using the MGCNA are shown in Table 1. According to the table, the model demonstrates an average accuracy of 99.32%, specificity of 98.4%, sensitivity of 98.74%, an F1 score of 97.76%, and an AUC of 98.51%. For the majority of patients, the recognition accuracy exceeds 97%, with one patient achieving 100% accuracy. The model achieves a specificity of 99% for 12 patients, which represents 50% of all test subjects, and a sensitivity of 99% for 17 patients, covering 70% of all test subjects. F1 scores for all patients are greater than 92%, and the AUC is above 95%. These results indicate the stability and high performance of MGCNA proposed in this study, which can aid medical professionals in diagnosis.

3.2 Patient-independent experiments

In contrast to the patient-specific experiments, patient-independent experiments involve training a general model for all patients. The experiment employed the leave-one-out method. Patient-independent experiments separate one patient's signal as a test set, using the EEG data of other patients as the training set and Validation set. This allows the model to detect epileptic activity in the test set by learning common seizure characteristics from the training set. This experiment is more clinically meaningful but requires to address differences in EEG signals among different patients, which may result from physiological variations, equipment noise, etc. Compared to patient-specific experiments, patient-independent experiments make it more challenging for the algorithm to identify epileptic seizures and lead to poorer detection results.

Patient-independent experimental results on CHB-MIT dataset using the MGCNA are shown in Table 2. As seen in the table, the average values for Accuracy, Sensitivity, Specificity, and F1 are all above 80%. There is a significant individual variation among different patients. Patients ♯4, ♯5, ♯7, ♯9, ♯10, ♯15, ♯17, ♯18, ♯19, ♯22, ♯23 achieved accuracy rates exceeding 90%, whereas patients ♯8, ♯12, ♯13, and ♯24 had comparatively lower accuracy rates. Compared to patient-specific experiments, patient-independent experiments exhibited a 13.95% lower accuracy rate. This is due to the differences in seizures between patients and the presence of different types of symptoms during seizures, which vary in frequency and duration.

3.3 Ablation study

To assess the contributions of different components in our model to classification performance, we conduct comparative experiments using model based only on similarity graph representation, model based only on distance graph representation, model based only

TABLE 1 Patient-specific experimental results on CHB-MIT dataset using the proposed architecture.

Case	Acc(%)	Sens(%)	Spec(%)	F1(%)	AUC(%)
CHB01	99.79	99.12	100.00	99.33	99.52
CHB02	99.56	96.96	100.00	98.01	98.44
CHB03	99.75	99.74	99.74	99.20	99.74
CHB04	99.32	96.58	99.50	94.81	98.05
CHB05	99.48	97.40	99.90	98.34	98.65
CHB06	99.94	100.00	99.94	98.84	98.63
CHB07	99.79	100.00	98.82	99.39	99.42
CHB08	99.62	99.82	98.57	98.83	99.20
CHB09	100.00	100.00	100.00	100.00	100.00
CHB10	99.96	100.00	99.74	99.87	99.87
CHB11	99.83	99.92	99.39	99.51	99.66
CHB12	98.63	99.71	93.77	95.67	96.74
CHB13	97.88	99.26	93.51	95.07	96.38
CHB14	98.11	99.07	93.45	94.18	96.26
CHB15	98.31	98.66	97.91	98.21	98.29
CHB16	99.61	100.00	97.50	98.67	98.75
CHB17	99.29	99.79	96.90	97.82	98.34
CHB18	99.55	99.80	98.28	98.60	99.04
CHB19	99.78	100.00	98.57	99.26	99.29
CHB20	98.91	93.61	99.50	94.32	96.56
CHB21	97.98	92.68	98.82	92.81	95.75
CHB22	99.43	99.53	99.41	98.15	99.47
CHB23	99.58	99.80	98.55	98.78	99.18
CHB24	99.48	98.20	99.77	98.57	98.99
Mean	99.32	98.74	98.40	97.76	98.51

TABLE 2 Patient-independent experimental results on CHB-MIT dataset using the proposed architecture.

Case	Acc(%)	Sens(%)	Spec(%)	F1(%)	AUC(%)
CHB01	91.51	83.73	99.28	90.79	91.51
CHB02	82.10	64.20	100.00	78.20	82.10
CHB03	81.33	62.67	100.00	77.05	81.33
CHB04	99.45	99.18	99.73	99.45	99.45
CHB05	96.49	97.60	95.38	96.53	96.49
CHB06	85.04	92.31	77.78	86.06	85.04
CHB07	98.89	98.41	99.36	98.88	98.89
CHB08	92.34	91.90	92.79	92.31	92.34
CHB09	100.00	100.00	100.00	100.00	100.00
CHB10	99.76	99.76	99.76	99.76	99.76
CHB11	93.89	93.32	94.46	93.86	93.89
CHB12	61.46	46.21	76.70	54.52	61.46
CHB13	70.52	81.09	59.95	73.34	70.52
CHB14	70.44	58.39	82.48	66.39	70.44
CHB15	91.93	82.93	93.33	73.52	91.93
CHB16	86.90	73.81	100.00	84.93	86.90
CHB17	98.05	97.87	98.23	98.05	98.05
CHB18	93.20	86.39	100.00	92.70	93.20
CHB19	94.69	90.27	99.12	94.44	94.69
CHB20	79.25	58.49	100.00	73.81	79.25
CHB21	95.41	91.89	98.92	95.24	95.41
CHB22	98.70	97.93	99.48	98.69	98.70
CHB23	97.61	95.98	99.25	97.57	97.61
CHB24	87.80	88.91	86.70	87.94	87.80
Mean	89.45	84.72	93.86	87.67	89.45

on trainable graph representation, and model without multi-head attention. To ensure the fairness of the experiments, the model settings are kept consistent. The results are shown in Figures 6, 7.

Graph convolutional network uses only spatial distance graph learning, denoted as *h1. Graph convolutional network uses only functional connectivity graph learning, denoted as *h2. Graph convolutional network uses adaptive graph learning, denoted as *h3. Non-attention graph representation models refers to the model that does not employ self-attention mechanisms to obtain self-attention scores for the three graph representations. We directly concatenate

the representations of three graphs into a graph feature, denoted as *self-attention.

From the Figures 6, 7, it is evident that MGCNA outperforms the other four ablation experiments. Compared to our model in this paper, the accuracy of the four ablation experiment models is lower by 5.22%, 5.30%, 3.37%, and 3.99%, respectively. As illustrated in Figure 6, the trainable graph representation model demonstrates superior performance compared to the other two graph representation learning methods, particularly for patients ♯4, ♯12, ♯13, ♯14, ♯17, and ♯21. Between spatial distance graph
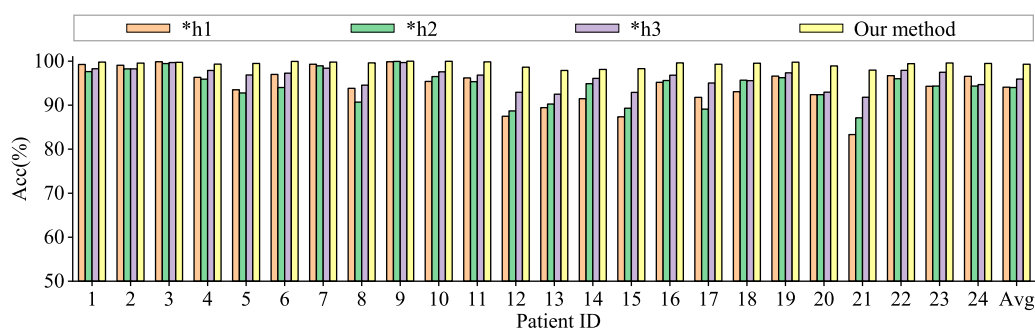
FIGURE 6
Ablation study for epilepsy recognition CHB-MIT dataset of different graph learning.*h1 represents the use of spatial distance graph learning only. *h2 indicates the use of functional connectivity graph learning only. *h3 denotes the use of adaptive graph learning only.
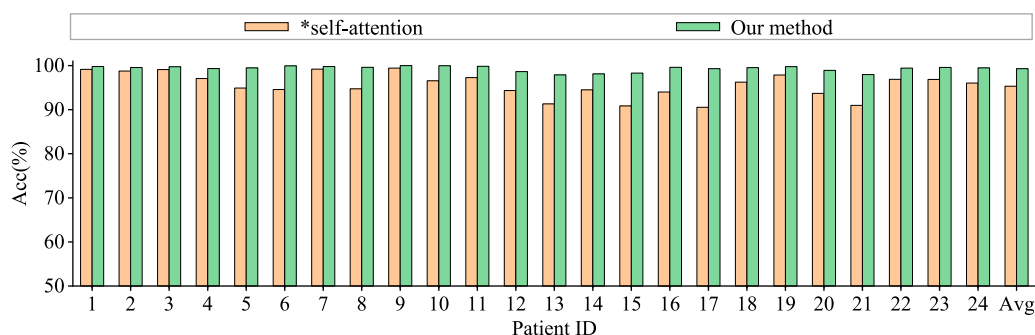

FIGURE 7
Ablation study for epilepsy recognition CHB-MIT dataset of multi-head attention.

learning and functional connectivity graph learning, the latter more effectively captures the interrelationships among EEG signals.

The effect of the self-attention mechanism on epilepsy detection approximates that of the model using only distance graph representation. Figure 7 illustrates that allocating different attention weights to graph features through self-attention mechanism helps to capture dependencies between graph features and enhances the model's ability to model relationships between graph features, which can achieve good performance in epilepsy detection. The model can effectively utilize the spatiotemporal relationships among EEG signals, supplement information through learnable graph representations, and obtain attention scores through self-attention mechanisms, and therefore the model's learning capability has been significantly enhanced.

To further validate the superiority of MGCNA, t-SNE is applied to visualize and analyze the features extracted from the ablation study. As depicted in Figure 8, the t-SNE visualization in two-dimensional embedding space illustrates the interictal and ictal features for both patient-specific and patient-independent experiments. We can see that in both patient-specific and patient-independent experiments, our approach exhibits superior recognition capabilities compared to the ablated models. Particularly, models utilizing only one graph construction method tend to confuse some interictal and ictal features. In contrast, better discriminative features were obtained using MGCNA, mainly
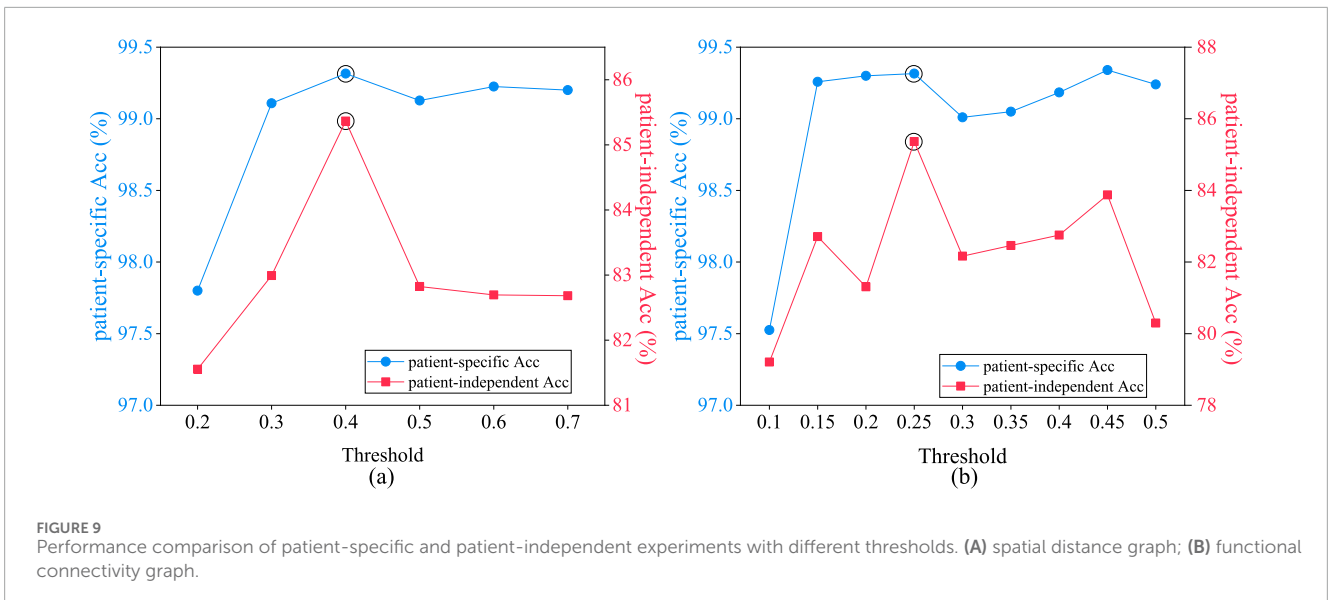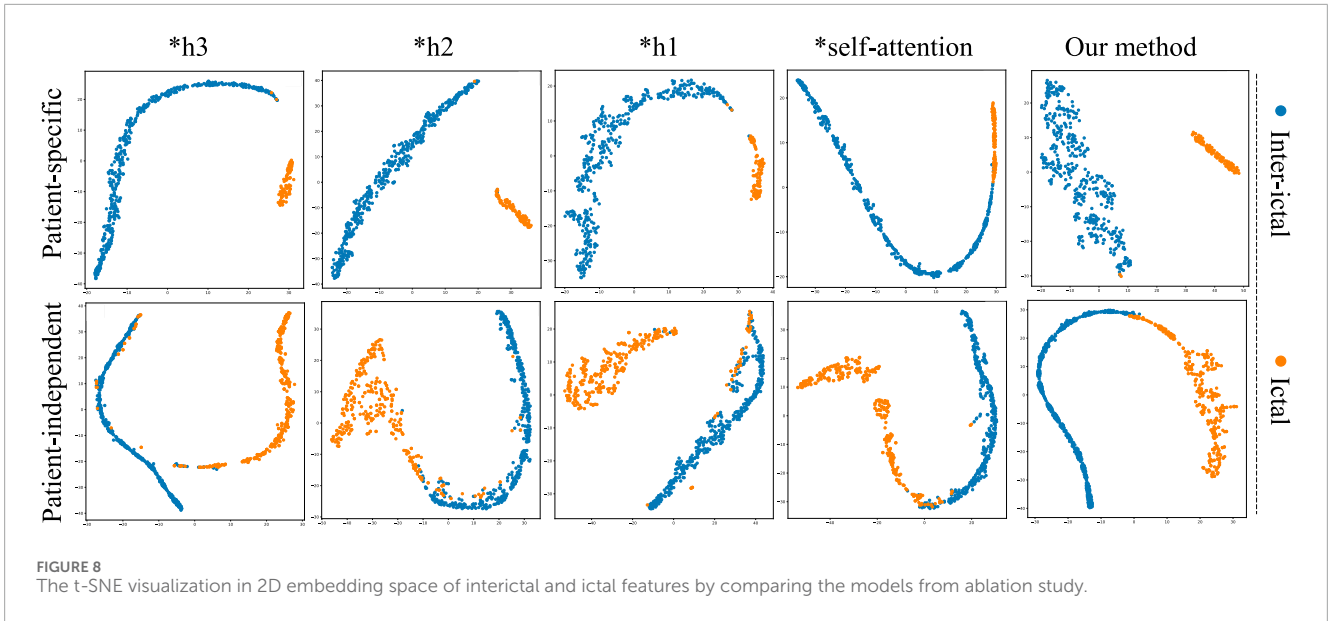
in terms of significant inter-ictal distances and dense intra-ictal distributions. These observations indicate that combining multi-branch GCN with self-attention mechanisms can yield the optimal performance for epileptic seizure classification.

4 Discussion

4.1 The influence of thresholds on graph representations

In both functional connectivity graph learning and spatial distance graph learning methods, it is necessary to set a threshold for constructing an adjacency matrix. The constructed adjacency matrix must not only ensure the sparsity of the graph but also be capable of distinguishing temporal and spatial characteristics of different types of EEG data to enhance the accuracy of the model in identifying epileptic seizures.

Figure 9A depicts the average epileptic seizure detection results of the CHB-MIT dataset in patient-specific and patient-independent experiments, under varying thresholds R_1 for spatial distance graph learning. The threshold range selected spans from 0.2 to 0.7. Notably, at a threshold of 0.2, the patient-specific epileptic seizure detection accuracy is the lowest, as lower thresholds tend to introduce excessive irrelevant physical connections between unrelated nodes

FIGURE 8
The t-SNE visualization in 2D embedding space of interictal and ictal features by comparing the models from ablation study.

FIGURE 9
Performance comparison of patient-specific and patient-independent experiments with different thresholds. **(A)** spatial distance graph; **(B)** functional connectivity graph.

in the spatial distance graph. Therefore, to maintain the sparsity of the adjacency matrix and achieve optimal seizure detection results, configuring the threshold R_1 for spatial distance graph learning to 0.4 is advocated.

Figure 9B illustrates the average epileptic seizure detection results in both patient-specific and patient-independent experiments under various threshold values R_2 for functional connectivity graph learning. The threshold range selected spans from 0.1 to 0.5, with a stride of 0.05. From Figure 9, it is evident that R_2 has a more pronounced impact on patient-independent epileptic seizure detection results. In the patient-specific experiment, where the training and testing datasets originate from the same patient, they exhibit similar data distributions. However, in the patient-independent experiment, differences exist between various patients, leading to inconsistent effects of different thresholds on the patients. Appropriate thresholds are applied to eliminate unrelated channel signals, utilizing functional

connectivity to capture neuronal synchronized discharge during epileptic seizures (Abbas et al., 2021). To maximize the spatial discriminative power of the functional connectivity graph and achieve optimal results for seizure detection, the threshold R_2 for functional connectivity graph learning is set at 0.25.

The spatial distance graph learning based on Equation 3 with the threshold R_1 set to 0.4 is illustrated in Figure 10. The distance between two bipolar derivations was determined by measuring the separation between the centers of the bipolar derivations. For instance, the coordinates of FP1-F3 are the centers of the FP1 coordinates and the F3 coordinates. After calculating the Euclidean distances, these distances are normalized to create the distance graph representation, and self-connections are added, i.e., the graph representation of the Euclidean distances plus the diagonal matrix.

Figure 11 depicts the functional connectivity graph learning of interictal and ictal signals under thresholds set at 0.25. In the
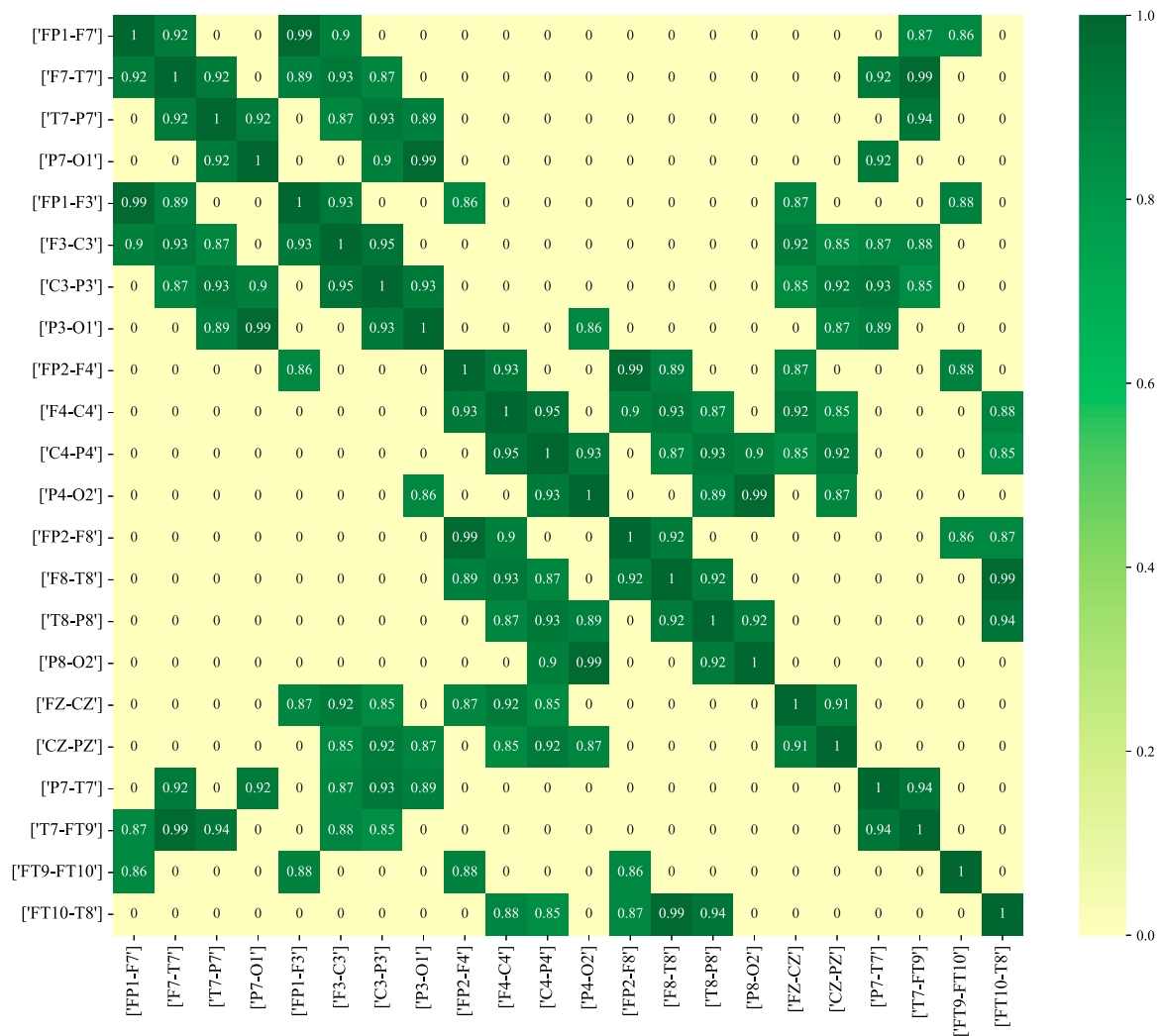
FIGURE 10
The spatial distance graph learning of EEG signals under thresholds R_1 set as 0.4.

figure, we have selected and described the functional connectivity graph learning that represents interictal and ictal signals with the highest degree of similarity. The values in the graph are calculated as the Pearson correlation coefficients, which are then normalized into a similarity adjacency matrix ranging from 0 to 1. From Figure 11, it can be observed that functional connectivity graph learning for interictal signals exhibits a strong degree of similarity, while the ictal signal displays weaker inter-correlations across many bipolar derivations. Graph representations based on functional connectivity can depict the functional connections between brain regions as interdependencies among EEG signals.

4.2 Comparisons with state-of-the-art methods

Table 3 presents a performance comparison between the proposed MGCNA and state-of-the-art epilepsy detection algorithms in the patient-specific experimental setting. Most of these methods involve feature extraction and deep learning algorithms for epilepsy detection. It is evident that the advancement of deep learning methods holds significant importance for signal recognition and detection.

Machine learning methods are classical approaches for epilepsy detection. Ein Shoka et al. (2021) conducted feature extraction after channel selection and evaluated the performance of different classifiers for epilepsy detection, with KNN exhibiting the best classification performance. Sukriti et al. (Sukriti et al., 2021) extracted multiscale spectral features (MSSFs) and employed a random forest classifier for epilepsy classification exhibits a higher specificity compared to MGCNA (99.17% *versus* 98.4%). However, the MGCNA achieves higher accuracy and sensitivity than that method. Cimr et al. (2023) achieved EEG classification by normalizing input signals and an 8-layer depth CNN model. The MGCNA is superior to that method in terms of accuracy, specificity and sensitivity. Zhao et al. (2023a) proposed using CNNs to extract local features and transformers to capture global information.
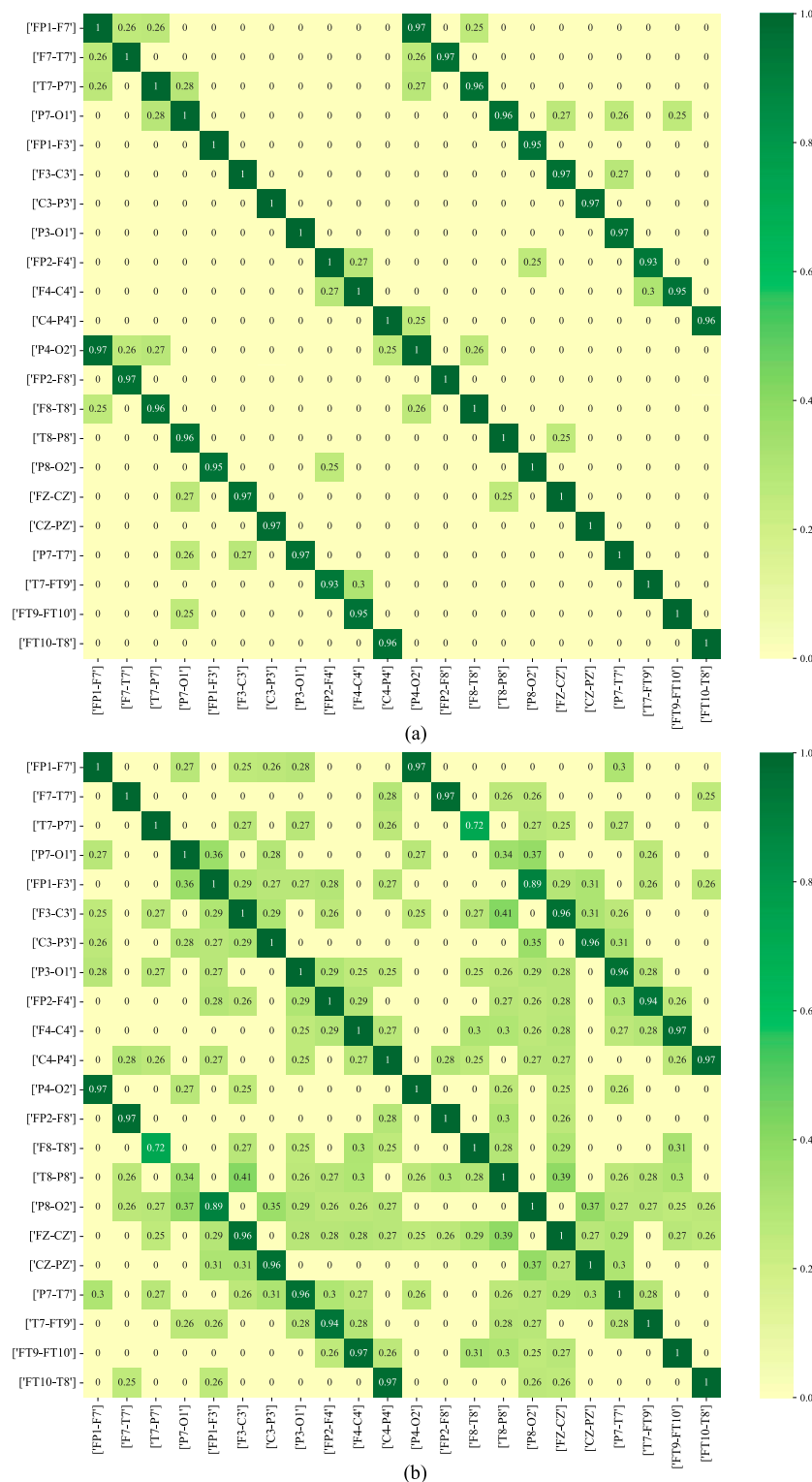
FIGURE 11
The functional connectivity graph learning of interictal and ictal signals under thresholds R_2 set as 0.25. **(A)** ictal signal; **(B)** interictal signal.

Although MGCNA outperforms this approach, the differences in preprocessing methods also affect the results of epilepsy detection.

GCN can analyze signals by considering the three-dimensional spatial positions of EEG electrodes, thereby compensating for

potential spatial information loss in deep learning methods such as CNN. The introduction of GCN can be utilized to analyze the spatiotemporal correlations among channels. Methods based on GCN have already demonstrated excellent

TABLE 3 Comparisons of performance between proposed method and recent works (patient-specific experimental setting).

Method	Acc(%)	Sens(%)	Spec(%)
Channel selection + KNN (Ein Shoka et al., 2021)	85	86.04	83.78
Multiscale spectral features + RF (Sukriti et al., 2021)	98.9	98.12	**99.17**
normalization + CNN (Cimr et al., 2023)	96.99	97.06	96.89
WNG-TS-1DCNN (Wang et al., 2023b)	93.1	91.8	96.3
HAN (Zhao et al., 2023b)	98.3	97.34	96.07
GAT + BiLSTM (He et al., 2022)	98.52	97.75	94.34
CNN + Transformer (Zhao et al., 2023a)	98.76	97.70	97.60
proposed MGCNA	**99.32**	**98.74**	98.40

The bold values represent the highest values in each column.

performance for the detection of epileptic signals. Wang et al. (2023b) introduced a two-stream graph-based framework for learning the Weighted Neighbour Graph (WNG) representation in both the frequency and time domains. However, this approach achieved inferior accuracy, sensitivity, and specificity, with values of 93.1%, 91.8%, and 96.3%, respectively, in comparison to the method proposed in the current study. Zhao et al. (2023b) introduced a hybrid Attention Network that utilizes the GAT to extract spatial features and the Transformer to extract temporal features addresses the issue of imbalanced data. However, in the patient-specific experiment, it achieved slightly lower accuracy, specificity, and sensitivity compared to MGCNA. He et al. (2022) utilized the GAT to extract spatial features and employed a BiLSTM network to capture temporal features, achieving an accuracy of 98.52% on the CHB-MIT dataset, slightly lower than MGCNA.

Table 4 presents comparisons of performance between the proposed method and state-of-the-art epilepsy detection algorithms in the patient-independent experimental setting. Due to the substantial variations in signals among different patients, the overall performance of the patient-independent experiment is lower than that of the patient-specific experiment. The accuracy of the patient-independent experiment based on SVM (Shoeb, 2009) only reaches 58.32%, indicating a significant performance gap compared to MGCNA. Wei et al. (2019) proposed an epilepsy detection algorithm by combining CNN and Wasserstein Generative Adversarial Nets (WGANs), achieving lower sensitivity than MGCNA, i.e. 72.11% vs. 83.54%, respectively. Zhao et al. (2023b) introduced the HAN model, which performs epilepsy detection not only in patient-specific experiments but also in patient-independent experiments. In both types of experiments, the three performance metrics were lower than MGCNA. Additionally, the ablation studies conducted in the patient-independent experimental setting are showcased herein, demonstrating that the MGCNA method exhibits superior performance compared to each ablation study. Since the three graph learning models can better extract the structural information of EEG and perform graph feature fusion,

TABLE 4 Comparisons of performance between proposed method and recent works (patient-independent experimental setting).

Method	Acc(%)	Sens(%)	Spec(%)
SVM (Shoeb, 2009)	58.32	74.6	42.19
WGANs (Wei et al., 2019)	84.00	72.11	**95.89**
HAN (Zhao et al., 2023b)	73.15	72.75	75.7
*h1	82.49	80.33	85.74
*h2	85.21	79.83	91.27
*h3	84.72	80.83	89.16
*self-attention	84.19	79.49	89.56
proposed MGCNA	**85.45**	**84.72**	93.86

The bold values represent the highest values in each column.

the overall performance is better than other epilepsy detection algorithms.

Through the above analysis, the MGCNA framework offers several key advantages. First, it employs a multi-branch graph convolutional network structure that dynamically learns temporal correlations and spatial topological information, enhancing the ability to process complex EEG signals, particularly in capturing spatial and temporal dependencies in epileptic brains. Second, the use of three graph learning approaches allows for a comprehensive evaluation of connectivity and synchronization across multiple channels, improving adaptability in different patient scenarios. Additionally, the multi-head attention mechanism further strengthens the framework's ability to handle local features and complex EEG patterns. Experimental results demonstrate that MGCNA outperforms other methods in both patient-specific and patient-independent tasks, highlighting its strong generalization capabilities. Lastly, as an end-to-end automatic seizure detection

model, MGCNA can be applied in clinical decision-making, helping clinicians diagnose childhood epileptic seizures more quickly and accurately, providing significant practical value.

4.3 Limitations and future work

While the MGCNA achieved satisfactory results in both patient-specific and patient-independent experiments, the MGCNA has several limitations. Firstly, spatial distance graph learning requires the prior determination of channel locations and is not robust to variations in EEG channel count. If channels are missing or if channel locations change, it can adversely affect recognition performance. Secondly, our model has a relatively long computation time because it involves calculating Pearson correlation coefficients for each sample to extract features for constructing EEG graph representations. Finally, the performance of model may depend on the quality and diversity of the data it has been trained on, and it is crucial to validate it on larger, more diverse datasets to ensure its generalizability. Specifically, both functional connectivity graph learning and spatial distance graph learning employ thresholds.

In future work, it is essential to explore graph representation methods that are more suitable for epileptic signals. While many scholars have already employed various graph representations (Raeisi et al., 2022), these often involve manual feature engineering and have long computation times. In the future, it's possible to explore alternative composition methods or utilize clustering algorithms to identify the most suitable adjacency matrix. In the future, we will explore alternative graph generation techniques, such as imposing appropriate constraints on trainable adjacency matrix or using clustering algorithms to identify the most suitable graph generator. Identifying the most suitable graph representation method for epileptic signal recognition is of paramount importance. In this study, patient-independent experiments hold more clinical relevance, and there is significant room for improvement in accuracy. To capture common seizure characteristics among different patients, techniques such as transfer learning will be considered to enhance the accuracy of patient-independent experiments. In current research, considerable attention has been devoted to the occurrence of seizures, yet there exist variations in seizure types among individuals. Consequently, in future investigations, emphasis will be placed on the analysis of seizure types in epilepsy research.

5 Conclusion

This study proposes a children epilepsy detection model named as MGCNA that combines a multi-branch GCN with multi-head attention. The MGCNA leverages three graph structures to learn spatiotemporal features among channels. It uses spatial graph representations to capture spatial distances between channels, functional connections between channels to learn spatial dependencies in the signals, and employs learnable graph representations to complement spatiotemporal features. The model employs a multi-head attention to assign importance weights to graph signals, learning relationships between graph representations. The model's performance in classifying epileptic

EEG signals is validated on the CHB-MIT dataset through patient-specific and patient-independent experiments. The experimental results indicate that the MGCNA shows excellent performance of childhood seizure detection surpassing other existing methods. This method can be used to assist in the childhood seizure detection and effectively reduce the workload of physicians. The EEG classification algorithm introduced in this research provides the potential to establish an EEG monitoring system for children with epilepsy.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

YL: Data curation, Investigation, Methodology, Writing–original draft, Writing–review and editing, Validation. YY: Funding acquisition, Methodology, Resources, Writing–original draft. SS: Investigation, Resources, Supervision, Writing–review and editing. HW: Funding acquisition, Methodology, Writing–review and editing. MS: Investigation, Methodology, Writing–review and editing, Validation. XL: Investigation, Methodology, Writing–review and editing. PZ: Methodology, Writing–review and editing. BW: Methodology, Writing–review and editing. NW: Methodology, Writing–review and editing. QS: Methodology, Writing–review and editing. ZH: Methodology, Writing–review and editing.

Funding

Conflict of interest

Authors MS and XL were employed by Neusoft Medical Systems Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbas, A. K., Azemi, G., Ravanshadi, S., and Omidvarnia, A. (2021). An eeg-based methodology for the estimation of functional brain connectivity networks: application to the analysis of newborn eeg seizure. *Biomed. Signal Process. Control* 63, 102229. doi:10.1016/j.bspc.2020.102229

Abu, A., and Diamant, R. (2023). Underwater object classification combining sas and transferred optical-to-sas imagery. *arXiv Prepr. arXiv:2304.11875* 144, 109868. doi:10.1016/j.patcog.2023.109868

Cai, S., Li, H., Wu, Q., Liu, J., and Zhang, Y. (2022). Motor imagery decoding in the presence of distraction using graph sequence neural networks. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 30, 1716–1726. doi:10.1109/TNSRE.2022.3183023

Cao, J., Feng, Y., Zheng, R., Cui, X., Zhao, W., Jiang, T., et al. (2022). Two-stream attention 3-d deep network-based childhood epilepsy syndrome classification. *IEEE Trans. Instrum. Meas.* 72, 1–12. doi:10.1109/tim.2022.3220287

Chen, J., Chen, C., Huang, W., Zhang, J., Debattista, K., and Han, J. (2024a). Dynamic contrastive learning guided by class confidence and confusion degree for medical image segmentation. *Pattern Recognit.* 145, 109881. doi:10.1016/j.patcog.2023.109881

Chen, Y., Xu, X., Bian, X., and Qin, X. (2024b). Eeg emotion recognition based on ordinary differential equation graph convolutional networks and dynamic time wrapping. *Appl. Soft Comput.* 152, 111181. doi:10.1016/j.asoc.2023.111181

Cheng, C., Yu, Z., Zhang, Y., and Feng, L. (2023). Hybrid network using dynamic graph convolution and temporal self-attention for eeg-based emotion recognition. *IEEE Trans. Neural Netw. Learn. Syst.*, 1–11. doi:10.1109/tnnls.2023.3319315

Cimr, D., Fujita, H., Tomaskova, H., Cimler, R., and Selamat, A. (2023). Automatic seizure detection by convolutional neural networks with computational complexity analysis. *Comput. Methods Programs Biomed.* 229, 107277. doi:10.1016/j.cmpb.2022.107277

Cui, G., Xia, L., Tu, M., and Liang, J. (2017). Automatic classification of epileptic electroencephalogram based on multiscale entropy and extreme learning machine. *J. Med. Imaging Health Inf.* 7, 949–955. doi:10.1166/jmihi.2017.2121

Cui, X., Cao, J., Hu, D., Wang, T., Jiang, T., and Gao, F. (2022a). Regional scalp eegs analysis and classification on typical childhood epilepsy syndromes. *IEEE Trans. Cognitive Dev. Syst.* 15, 662–674. doi:10.1109/tcds.2022.3175636

Cui, X., Hu, D., Lin, P., Cao, J., Lai, X., Wang, T., et al. (2022b). Deep feature fusion based childhood epilepsy syndrome classification from electroencephalogram. *Neural Netw.* 150, 313–325. doi:10.1016/j.neunet.2022.03.014

Defferrard, M., Bresson, X., and Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. *Adv. neural Inf. Process. Syst.* 29.

Deng, Z., Li, C., Song, R., Liu, X., Qian, R., and Chen, X. (2023). Eeg-based seizure prediction via hybrid vision transformer and data uncertainty learning. *Eng. Appl. Artif. Intell.* 123, 106401. doi:10.1016/j.engappai.2023.106401

Ding, X., Nie, W., Liu, X., Wang, X., and Yuan, Q. (2023). Compact convolutional neural network with multi-headed attention mechanism for seizure prediction. *Int. J. Neural Syst.* 33, 2350014. doi:10.1142/S0129065723500144

Dong, C., Zhao, Y., Zhang, G., Xue, M., Chu, D., He, J., et al. (2022). Attention-based graph resnet with focal loss for epileptic seizure detection. *J. Ambient Intell. Smart Environ.* 14, 61–73. doi:10.3233/ais-210086

Ein Shoka, A. A., Alkinani, M. H., El-Sherbeny, A., El-Sayed, A., and Dessouky, M. M. (2021). Automated seizure diagnosis system based on feature extraction and channel selection using eeg signals. *Brain Inf.* 8, 1–16. doi:10.1186/s40708-021-00123-7

Feng, Y., Zheng, R., Cui, X., Wang, T., Jiang, T., Gao, F., et al. (2022). 3d residual-attention-deep-network-based childhood epilepsy syndrome classification. *Knowledge-Based Syst.* 248, 108856. doi:10.1016/j.knosys.2022.108856

Grattarola, D., Livi, L., Alippi, C., Wennberg, R., and Valiante, T. A. (2022). Seizure localisation with attention-based graph neural networks. *Expert Syst. Appl.* 203, 117330. doi:10.1016/j.eswa.2022.117330

He, J., Cui, J., Zhang, G., Xue, M., Chu, D., and Zhao, Y. (2022). Spatial–temporal seizure detection with graph attention network and bi-directional lsm architecture. *Biomed. Signal Process. Control* 78, 103908. doi:10.1016/j.bspc.2022.103908

Ho, T. K. K., and Armanfard, N. (2023). Self-supervised learning for anomalous channel detection in eeg graphs: application to seizure analysis. *Proc. AAAI Conf. Artif. Intell.* 37, 7866–7874. doi:10.1609/aaai.v37i7.25952

Ji, X., Li, Y., and Wen, P. (2022). Jumping knowledge based spatial-temporal graph convolutional networks for automatic sleep stage classification. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 30, 1464–1472. doi:10.1109/tnsre.2022.3176004

Jia, Z., Lin, Y., Wang, J., Ning, X., He, Y., Zhou, R., et al. (2021). Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 29, 1977–1986. doi:10.1109/TNSRE.2021.3110665

Jia, Z., Lin, Y., Wang, J., Zhou, R., Ning, X., He, Y., et al. (2020). Graphsleepnet: adaptive spatial-temporal graph convolutional networks for sleep stage classification. *IJCAI* 2021, 1324–1330. doi:10.24963/ijcai.2020/184

Kalaganis, F. P., Laskaris, N. A., Chatzilari, E., Nikolopoulos, S., and Kompatsiaris, I. (2020). A data augmentation scheme for geometric deep learning in personalized brain–computer interfaces. *IEEE access* 8, 162218–162229. doi:10.1109/access.2020.3021580

Kannathal, N., Choo, M. L., Acharya, U. R., and Sadasivan, P. (2005). Entropies for detection of epilepsy in eeg. *Comput. methods programs Biomed.* 80, 187–194. doi:10.1016/j.cmpb.2005.06.012

Kipf, T. N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks

Lee, S., Yu, Y., Back, S., Seo, H., and Lee, K. (2024). Sleepy: automatic sleep scoring with feature pyramid and contrastive learning. *Expert Syst. Appl.* 240, 122551. doi:10.1016/j.eswa.2023.122551

Li, M., Chen, H., and Cheng, Z. (2022a). An attention-guided spatiotemporal graph convolutional network for sleep stage classification. *Life* 12, 622. doi:10.3390/life12050622

Li, M., Qiu, M., Zhu, L., and Kong, W. (2023). Feature hypergraph representation learning on spatial-temporal correlations for eeg emotion recognition. *Cogn. Neurodynamics* 17, 1271–1281. doi:10.1007/s11571-022-09890-3

Li, Y., Chen, J., Li, F., Fu, B., Wu, H., Ji, Y., et al. (2022b). Gmss: graph-based multi-task self-supervised learning for eeg emotion recognition. *IEEE Trans. Affect. Comput.* 14, 2512–2525. doi:10.1109/taffc.2022.3170428

Lian, J., and Xu, F. (2023). Epileptic eeg classification via graph transformer network. *Int. J. neural Syst.* 33, 2350042. doi:10.1142/S0129065723500429

Liu, D., Dong, X., Bian, D., and Zhou, W. (2023). Epileptic seizure prediction using attention augmented convolutional network. *Int. J. Neural Syst.* 33, 2350054. doi:10.1142/S0129065723500545

Liu, H., Zhang, J., Liu, Q., and Cao, J. (2022). Minimum spanning tree based graph neural network for emotion classification using eeg. *Neural Netw.* 145, 308–318. doi:10.1016/j.neunet.2021.10.023

Lopez, S., Del Percio, C., Lizio, R., Noce, G., Padovani, A., Nobili, F., et al. (2023). Patients with alzheimer's disease dementia show partially preserved parietal 'hubs' modeled from resting-state alpha electroencephalographic rhythms. *Front. Aging Neurosci.* 15, 780014. doi:10.3389/fnagi.2023.780014

Mehta, S., and Rastegari, M. (2022). Separable self-attention for mobile vision transformers. *arXiv Prepr. arXiv:2206.02680*. doi:10.48550/arXiv.2206.02680

Meng, L., Hu, J., Deng, Y., and Hu, Y. (2022). Electrical status epilepticus during sleep electroencephalogram waveform identification and analysis based on a graph convolutional network. *Biomed. Signal Process. Control* 77, 103788. doi:10.1016/j.bspc.2022.103788

Na, J., Wang, Z., Lv, S., and Xu, Z. (2021). An extended k nearest neighbors-based classifier for epilepsy diagnosis. *IEEE Access* 9, 73910–73923. doi:10.1109/access.2021.3081767

Peh, W. Y., Thangavel, P., Yao, Y., Thomas, J., Tan, Y.-L., and Dauwels, J. (2023). Six-center assessment of conn-transformer with belief matching loss for patient-independent seizure detection in eeg. *Int. J. Neural Syst.* 33, 2350012. doi:10.1142/S0129065723500120

Qiu, X., Yan, F., and Liu, H. (2023). A difference attention resnet-lstm network for epileptic seizure detection using eeg signal. *Biomed. Signal Process. Control* 83, 104652. doi:10.1016/j.bspc.2023.104652

Raeisi, K., Khazaei, M., Croce, P., Tamburro, G., Comani, S., and Zappasodi, F. (2022). A graph convolutional neural network for the automated detection of seizures in the neonatal eeg. *Comput. Methods Programs Biomed.* 222, 106950. doi:10.1016/j.cmpb.2022.106950

Rennie, J., Chorley, G., Boylan, G., Pressler, R., Nguyen, Y., and Hooper, R. (2004). Non-expert use of the cerebral function monitor for neonatal seizure detection. *Archives Dis. Childhood-Fetal Neonatal Ed.* 89, 37–40. doi:10.1136/fn.89.1.f37

Shoeb, A. H. (2009). Application of machine learning to epileptic seizure onset detection and treatment. *Ph.D. thesis, Mass. Inst. Technol.*

Singh, G., Singh, B., and Kaur, M. (2019). Grasshopper optimization algorithm–based approach for the optimization of ensemble classifier and feature selection to classify epileptic eeg signals. *Med. and Biol. Eng. and Comput.* 57, 1323–1339. doi:10.1007/s11517-019-01951-w

Song, T., Liu, S., Zheng, W., Zong, Y., Cui, Z., Li, Y., et al. (2021). Variational instance-adaptive graph for eeg emotion recognition. *IEEE Trans. Affect. Comput.* 14, 343–356. doi:10.1109/taffc.2021.3064940

Song, T., Zheng, W., Song, P., and Cui, Z. (2018). Eeg emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.* 11, 532–541. doi:10.1109/taffc.2018.2817622

Sukriti, Chakraborty, M., and Mitra, D. (2021). A computationally efficient automated seizure detection method based on the novel idea of multiscale spectral features. *Biomed. Signal Process. Control* 70, 102990. doi:10.1016/j.bspc.2021.102990

Sun, Q., and Yang, Y. (2023). Unsupervised video anomaly detection based on multi-timescale trajectory prediction. *Comput. Vis. Image Underst.* 227, 103615. doi:10.1016/j.cviu.2022.103615

Thuwajit, P., Rangpong, P., Sawangjai, P., Autthasan, P., Chaisaen, R., Banluesombatkul, N., et al. (2021). Eegwavenet: multiscale conn-based spatiotemporal feature extraction for eeg seizure detection. *IEEE Trans. Industrial Inf.* 18, 5547–5557. doi:10.1109/tii.2021.3133307

Wagh, N., and Varatharajah, Y. (2020). "Eeg-gcnn: augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network," in *Machine Learning for health (PMLR)*, 367–378.

Wang, H., Xu, L., Bezerianos, A., Chen, C., and Zhang, Z. (2020). Linking attention-based multiscale conn with dynamical gcn for driving fatigue detection. *IEEE Trans. Instrum. Meas.* 70, 1–11. doi:10.1109/tim.2020.3047502

Wang, J., Gao, R., Zheng, H., Zhu, H., and Shi, C.-J. R. (2023a). Ssgcnet: a sparse spectra graph convolutional network for epileptic eeg signal classification. *IEEE Trans. Neural Netw. Learn. Syst.* 35, 12157–12171. doi:10.1109/tnnls.2023.3252569

Wang, J., Liang, S., Zhang, J., Wu, Y., Zhang, L., Gao, R., et al. (2023b). EEG signal epilepsy detection with a weighted neighbor graph representation and two-stream graph-based framework. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 31, 3176–3187. doi:10.1109/tnsre.2023.3299839

Wang, Y., Cui, W., Yu, T., Li, X., Liao, X., and Li, Y. (2023c). Dynamic multi-graph convolution based channel-weighted transformer feature fusion network for epileptic seizure prediction. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 31, 4266–4277. doi:10.1109/tnsre.2023.3321414

Wang, Y., Shi, Y., Cheng, Y., He, Z., Wei, X., Chen, Z., et al. (2022). A spatiotemporal graph attention network based on synchronization for epileptic seizure prediction. *IEEE J. Biomed. Health Inf.* 27, 900–911. doi:10.1109/JBHI.2022.3221211

Wang, Z., Hou, S., Xiao, T., Zhang, Y., Lv, H., Li, J., et al. (2023d). Lightweight seizure detection based on multi-scale channel attention. *Int. J. Neural Syst.* 33, 2350061. doi:10.1142/S0129065723500612

Wei, Z., Zou, J., Zhang, J., and Xu, J. (2019). Automatic epileptic eeg detection using convolutional neural network with improvements in time-domain. *Biomed. Signal Process. Control* 53, 101551. doi:10.1016/j.bspc.2019.04.028

Wu, G., Yu, K., Zhou, H., Wu, X., and Su, S. (2024). Time-series anomaly detection based on dynamic temporal graph convolutional network for epilepsy diagnosis. *Bioengineering* 11, 53. doi:10.3390/bioengineering11010053

Xiao, T., Wang, Z., Zhang, Y., Wang, S., Feng, H., Zhao, Y., et al. (2024). Self-supervised learning with attention mechanism for eeg-based seizure detection. *Biomed. Signal Process. Control* 87, 105464. doi:10.1016/j.bspc.2023.105464

Xu, X., Lin, M., and Xu, T. (2022). Epilepsy seizures prediction based on nonlinear features of eeg signal and gradient boosting decision tree. *Int. J. Environ. Res. Public Health* 19, 11326. doi:10.3390/ijerph191811326

Zhang, M., Zhang, B., Wang, F., Chen, Y., and Jiang, N. (2010). Multi-scale phase average waveform of electroencephalogram signals in childhood absence epilepsy using wavelet transformation. *NEURAL Regen. Res.* 5, 774–780. doi:10.3969/j.issn.1673-5374.2010.10.010

Zhao, P., Zheng, Q., Ding, Z., Zhang, Y., Wang, H., and Yang, Y. (2022a). A high-dimensional and small-sample submersible fault detection method based on feature selection and data augmentation. *Sensors* 22, 204. doi:10.3390/s22010204

Zhao, X., Peng, X., Niu, K., Li, H., He, L., Yang, F., et al. (2022b). A multi-head self-attention deep learning approach for detection and recommendation of neuromagnetic high frequency oscillations in epilepsy. *Front. Neuroinformatics* 16, 771965. doi:10.3389/fninf.2022.771965

Zhao, Y., Chu, D., He, J., Xue, M., Jia, W., Xu, F., et al. (2023a). Interactive local and global feature coupling for eeg-based epileptic seizure detection. *Biomed. Signal Process. Control* 81, 104441. doi:10.1016/j.bspc.2022.104441

Zhao, Y., He, J., Zhu, F., Xiao, T., Zhang, Y., Wang, Z., et al. (2023b). Hybrid attention network for epileptic eeg classification. *Int. J. Neural Syst.* 33, 2350031. doi:10.1142/S0129065723500314