Analysis of Epileptic EEG Signals Based on Graph Capsule with Spatio-Temporal Spectral Hierarchical Residual Attention

Sunan Ge, Shuang Wang, Xueqing Zhao, Xinying Wang, Qiong Yang, Song Tian, Jun Chen

***Abstract*—Epilepsy is one of the most common neurological diseases worldwide. However, how to accurately detect epileptic seizure events and their occurrence times remains a challenging problem that is difficult to solve. Therefore, this study proposes an electroencephalogram (EEG) signal analysis method based on a graph capsule network combined with a spatio-temporal spectrum hierarchical residual attention mechanism. The aim is to improve the accuracy of epilepsy detection and prediction by capturing spatio-temporal dependencies and feature correlations. Firstly, a hierarchical spatio-temporal spectrum fusion module is adopted to characterize the dynamic spatio-temporal coupling information features of different rhythms of epileptic EEG signals, thereby obtaining dynamic spatio-temporal features. At the same time, a multiple hybrid attention mechanism is used to effectively extract the graph structure correlation information in the signal spatial interaction space. With the help of these spatial correlation features, a detailed and fine-grained characterization of the spatial information is achieved. Finally, a hierarchical residual attention graph capsule module is established. By fusing multi-information of the graph structure in the time domain, frequency domain, and spatial domain, the most discriminative feature information is obtained. The experimental results show that this method exhibits certain effectiveness in terms of accuracy, sensitivity, specificity, false positive rate per hour (FPR/h), and the area under the curve (AUC).**

***Index Terms*—epilepsy prediction, EEG signals, deep learning, hierarchical spatio-temporal spectral fusion features, hierarchical residual attention graph capsule module**

Manuscript received January 31, 2025; revised June 2, 2025. This work was supported by the Shaanxi Natural Science Youth Foundation (No.2021JQ-693, No.2021JQ-694), the National Social Science Foundation of China Art Project (No.23EH232), the Key Research and Development Program of Shaanxi Province in 2023 (No.2023-YBGY-404, No.2023-ZDLGY-48), the Natural Science Basic Research Program of Shaanxi (No.2023-JC-YB-558), Xi'an Engineering University Fund (GGSZWHFW2025-004), and the Scientific Research Program Funded by the Education Department of Shaanxi Provincial Government (No.23JS027).

Sunan Ge is a lecturer at Shaanxi Key Laboratory of Clothing Intelligence and State-Province Joint Engineering and Research Center of Advanced Networking and Intelligent Information Services, Computer Science School, Xi'an Polytechnic University, Shaanxi, 710048, P. R. China. (corresponding author, e-mail: gesunan@xpu.edu.cn).

Shuang Wang is a postgraduate student at the Computer Science School, Xi'an Polytechnic University, Shaanxi, 710048, P. R. China. (e-mail: 230721103@stu.xpu.edu.cn).

Xueqing Zhao is a professor at the Computer Science School, Xi'an Polytechnic University, 710048, P. R. China. (e-mail: zhaoxueqing@xpu.edu.cn).

Xinying Wang is a director of Big Data Application Research Section, CEPRI (China Electric Power Research Institute), Beijing, 100192, P. R. China. (e-mail: wangxinying@epri.sgcc.com.cn).

Qiong Yang is a lecturer at the Computer Science School, Xi'an Polytechnic University, Shaanxi, 710699, P. R. China. (e-mail: yangqiong@xpu.edu.cn).

Song Tian is a director of Information Technology Department, Shandong Ruyi NOGARA Textile and Garment Co Ltd, Jining, Shandong, 271000, P. R. China. (e-mail: songtian@chinaruyi.com).

Jun Chen is a lecturer at the Computer Science School, Xi'an Polytechnic University, Shaanxi, 710699, P. R. China. (e-mail: chenjun@xpu.edu.cn).

I. Introduction

EPILEPSY is a transient brain dysfunction caused by the sudden abnormal and disordered discharge of neurons in the brain [1]. This discharge is triggered by the local or generalized hypersynchronization of neuronal components. Different from other neurological diseases, epilepsy is a neurological disorder characterized by recurrent seizures, and it is accompanied by various neuropathological changes with increasing age. According to statistics from the World Health Organization (WHO), approximately 2.4 million people worldwide are diagnosed with epilepsy each year, and the affected population exceeds 70 million. The early mortality rate of epilepsy patients is two to three times that of the general population [2][3][4]. In China, the overall prevalence of epilepsy is 7.0 cases per 1,000 people, and the annual incidence rate is 28.8 cases per 100,000 people. The prevalence of active epilepsy (defined as experiencing seizures within one year) is 4.6 cases per 1,000 people. It is estimated that there are approximately 9 million epilepsy patients in China, including 5 to 6 million patients with active epilepsy. In addition, about 400,000 new epilepsy cases are added each year, making it the second most common neurological disease. Among epilepsy patients, about 70% of them can have their conditions controlled by antiepileptic drugs, and about 10% of them can be cured by surgery. However, approximately one-third of patients suffer from drug-resistant epilepsy. Since epilepsy is a syndrome with multiple risk factors and a strong genetic tendency, its pathogenesis is not yet fully understood. This greatly hinders the diagnosis and treatment of epilepsy, leading to increased morbidity and mortality, reduced social participation, and comorbidities such as mental disorders, stigmatization, and other complications. These factors impose a heavy burden on patients, their families, and society [5]. In addition, patients with uncontrolled epileptic seizures are also at risk of permanent memory impairment, depression, anxiety, suicidal ideation, and other mental illnesses, which seriously impair their quality of life [6][7]. Establishing an epileptic seizure warning system to predict seizures and take timely protective measures can significantly reduce the risk of accidental injuries in patients with chronic epilepsy. Meanwhile, accurate early diagnosis and classification of seizure types are crucial for successful treatment. Therefore, studying the detection and prediction of epileptic EEG signals has important clinical and social significance for improving the quality of life of epilepsy patients [8].

Currently, the methods used for diagnosing epilepsy in clinical practice include positron emission tomography

(PET) [9], magnetic resonance imaging (MRI) [10], and EEG [11]. Among them, EEG is widely applied in the detection and prediction of epileptic EEG signals because it can directly capture neuronal signals and continuously monitor the electrical activities and potential characteristics of neurological diseases [12]. Automated epilepsy seizure detection and prediction methods based on EEG can reduce the subjective variability caused by differences in clinicians' experience and alleviate the recording workload of medical staff. In the early stage, the automated detection and prediction of epileptic EEG signals generally follow the pattern of "feature extraction + classifier". The feature extraction methods mainly include wavelet packet transform (WPT), discrete wavelet transform (DWT), statistical time-domain features (STFs), spike discharge rate, and nonlinear feature extraction techniques, etc. The classifiers mainly rely on machine learning algorithms, such as decision tree classifier (DTC), k-nearest neighbors algorithm (KNN), support vector machine (SVM), artificial neural network, and ensemble classifiers, etc.

This comprehensive method that combines multiple feature extraction methods with classifiers demonstrates a certain degree of flexibility. It can obtain features under different frequency transformations and magnify the fine grained features of EEG signals, thereby enhancing the accuracy of feature extraction from coarse to fine in EEG signals and improving the final detection and prediction accuracy. However, due to the highly complex phase space and high dimensional feature information of EEG signals themselves, this low dimensional feature extraction method may lead to a large number of coupling relationships between signals and make it difficult to accurately capture spatiotemporal information. These limitations will ultimately affect the detection and prediction accuracy of epileptic EEG signals.

In order to further explore the deep and spatiotemporal information of epileptic EEG signals, Gao et al. [13] adopted a spatiotemporal multi-scale convolutional neural network (CNN) with dilated convolutions to expand the receptive field and systematically aggregate global information. They also employed a feature weighting and fusion strategy based on the attention mechanism to better integrate features and eliminate redundant information in the dilated convolution blocks. Ding et al. [14] proposed a signal analysis method based on a convolutional neural network and a multi-head attention mechanism (MHAM). Georgis Yap et al. [15] combined a CNN with a long short-term memory network (LSTM) and a temporal convolutional network (TCN) to develop three unsupervised methods (CNN, CNN-LSTM, TCN autoencoder) to reduce the workload of data annotation. Although the LSTM and the attention mechanism are proficient in extracting semantic relationships and improving the accuracy of epileptic seizure detection and prediction in deep network models combined with EEG signal analysis, the LSTM has difficulties in capturing global information and directly modeling the long-range dependencies in time series. At the same time, the attention mechanism focuses too much on the global context and is likely to overlook time series information. In addition, CNNs that rely on Euclidean structural spectral features to analyze time series signals or spectrograms often ignore key information such as the connectivity and functional connections based on the physical distances between brain regions.

In order to capture the relationships between electrode leads and gain an in-depth understanding of the current state or evolutionary characteristics of the complex brain system, this paper proposes a novel EEG signal analysis method for epilepsy based on a spatio-temporal spectrum hierarchical residual attention graph capsule module. By utilizing the structural information of the EEG signal graph, this method can obtain the physiological and pathological states of individuals. In addition, considering the complex interactions between different structural and functional regions of the brain, this paper employs multiple modular subnetworks to explore the three-dimensional multi channel relationships of multi scale information in the time frequency space. This generates autonomous, multi level, and diverse graph structure features that can fully capture and integrate EEG information with highly complex spatiotemporal interactions. Therefore, this method coordinates the multi rhythm spatio-temporal dependencies in the epileptic brain, which is helpful for improving the prediction of the onset of epileptic seizures. Finally, we introduce quantitative indicators to evaluate the ability of the model to localize the onset of epileptic seizures. To sum up, the main contributions of this study are as follows:

1. By constructing a hierarchical spatio-temporal spectrum attention module, the model can effectively capture the complex spatio-temporal interaction features in EEG signals, as well as the time-frequency information and spatial relationships of multi scale EEG signals.

2. A multiple hybrid attention mechanism is established to extract and integrate the spatio-temporal features in EEG signals, enabling the differentiation of different spatio-temporal states.

3. A hierarchical graph capsule module is proposed to address the problem of missing contextual spatial information. This module converts low-dimensional complex information into geometric graph structure information, thereby enhancing the model's ability to handle complex data.

II. Methods

The graph structure is a data structures composed of nodes and edges, which is used to represent various complex relationships and connections. Generally, a directed graph G can be represented as $G = \{V, U, A\}$, where V is the vertex set with $|V| = n$ nodes; $U \in \mathbb{R}^{n \times c}$ is the node attributes, and its *i-th* row $U_i \in \mathbb{R}^c$ represents the features of the c-channelindexed by node i, and $A \in \mathbb{R}^{n \times n}$ is the adjacent matrix, whose (i, j)th entry indicates connections from node j to i.

To deeply explore the spatial relationships between various attributes (such as frequency, time, and amplitude) in electroencephalogram signals, this paper proposes a seizure detection and prediction framework based on a graph capsule with a hierarchical spatio-temporal spectral residual attention network, as illustrated in Fig.1. This method primarily comprises four subnetworks: (1)Temporal embedding module; (2)Hierarchical spatio-temporal spectral fusion attention module; (3)Multiple hybrid attention mechanism; (4)Hierarchical residual attention graph capsule

module. The following sections provide detailed information for each step.

A. Temporal Embedding Module

To eliminate coupling factors in EEG information without altering the signal size, a temporal embedding module is designed. The dynamic subband EEG processed by this module serves as input data for the hierarchical spatio-temporal spectral fusion attention module.

Specifically, the temporal embedding module is constructed by cascading a temporal convolutional layer and a standard residual block. The temporal convolutional layer integrates a 2D convolutional layer with a kernel size of 1×3, stride of 1, padding of 1, and 4 groups, followed by a batch normalization layer and an Exponential Linear Unit (ELU) activation function. Subsequently, the standard residual block sequentially contains a 2D convolutional layer with a kernel size of 1×3 and 8 groups, a batch normalization layer, another 2D convolutional layer, another batch normalization layer, and a skip connection layer with a kernel size of 1×1. Preprocessed EEG data is fed into this temporal embedding module, and after processing, dynamic subband EEG data is generated as output.

B. Hierarchical Spatio-Temporal Spectral Fusion Attention Module

For further obtaining the multi scale information of the above mentioned dynamic subband EEG, a hierarchical spatio-temporal spectrum fusion attention module is adopted here, which includes three sub modules: 1) hierarchical spectrum block; 2) Multi scale time block; 3) Hierarchical attention block. Each step will be described in detail in the following subsection.

1) Hierarchical Spectrum Block

Considering that epileptic EEG signals exhibit varying frequencies across different brain regions, for instance, during seizure onset, the number of pathological spike waves in EEG signals gradually increases, and there is a significant difference between the average spike rate in the pre ictal phase and the inter-ictal phase [16]. To characterize the dynamic changes in multiple regions of EEG signals, this paper employs Daubechies Wavelet Convolution (Db waveConv) to model spike wave signals and further constructs a hierarchical spectral module based on this approach.

This module includes L-layer wavelet convolutions (waveConv) based on Daubechies order-4 (Db4) wavelet. The dynamic subband EEG embedded in the network is used as input data for the periodic padding wavelet decomposition of consecutive L-layer filled to obtain spectral details at different scales [17]. is determined by the signal sampling rate $L = \lfloor \log_2 f \rfloor - 3$, where f is the sampling rate of the EEG signal data in the dataset, and $\lfloor . \rfloor$ represents the rounding operation.

Due to the fact that the wavelet transform process includes a low-pass coefficient $\circ$ and a high pass coefficient $\bullet$, this module extracts the low frequency $S_w^\circ = (S_\delta^\circ, S_\theta^\circ, S_\alpha^\circ, S_\beta^\circ, S_\gamma^\circ)$ and high frequency information $S_w^\bullet = (S_\delta^\bullet, S_\theta^\bullet, S_\alpha^\bullet, S_\beta^\bullet, S_\gamma^\bullet)$ from the input dynamic subband EEG δ,θ,α,βandγ, separately. The low-frequency $S_w^\circ = (S_\delta^\circ, S_\theta^\circ, S_\alpha^\circ, S_\beta^\circ, S_\gamma^\circ)$ and high frequency information $S_w^\bullet = (S_\delta^\bullet, S_\theta^\bullet, S_\alpha^\bullet, S_\beta^\bullet, S_\gamma^\bullet)$, where w $\in [\delta, \theta, \alpha, \beta, \gamma]$, and $S_\delta^\circ, S_\theta^\circ, S_\alpha^\circ, S_\beta^\circ, S_\gamma^\circ$ are the low frequency information, as well as $S_\delta^\bullet, S_\theta^\bullet, S_\alpha^\bullet, S_\beta^\bullet, S_\gamma^\bullet$ are the high frequency information in the corresponding frequency band. By concatenating low-frequency information and high frequency information output data along dynamic sub bands, the frequency domain features corresponding to the five brain waveforms with shape $S_w = (S_\delta, S_\theta, S_\alpha, S_\beta, S_\gamma)$ are finally obtained.

2) Multi Scale Temporal Block

Meanwhile, this paper adopts a multi scale temporal analysis block(i. e. , five parallel time convolutional layers (TCLs) with corresponding convolutional kernel sizes of $\{k/8, k/4, k/2, k, k\}$) to obtain multi scale temporal information of EEG signals. Each temporal convolutional layer consists of a two dimensional convolutional layer with trainable kernel parameters, a batch normalization layer (BN), and an ELU activation function.

Then, the multi scale temporal analysis module takes the preprocessed dynamic subband EEG as input data, respectively, extracts the relevant features of different frequency bands including δ, θ, α, β and γ, and finally obtains the multi scale temporal domain features as $T_w = (t_\delta, t_\theta, t_\alpha, t_\beta, t_\gamma)$,where $t_\delta, t_\theta, t_\alpha, t_\beta, t_\gamma$ represent the temporal domain features corresponding to the δ, θ, α, β and γ frequency bands.

3) Hierarchical Attention Block

After obtaining multi scale spatio-temporal information, a hierarchical attention module is designed to capture the heterogeneity of spatio-temporal multi domain features. This module sequentially connects a 2D convolutional layer, a batch normalization layer, a variant of the Squeeze and Excitation (SE) module, an Exponential Linear Unit (ELU) activation function, another 2D convolutional layer, another batch normalization layer, another variant of the SE module, another ELU activation function, and an adaptive average pooling layer. It enables the parallel and in-depth capture of spectral and temporal features while jointly enhancing cross-domain channel responses.

The specific process is as follows. First, the multi scale temporal frequency fusion features obtained from the multi scale temporal block and the hierarchical spectrum module are concatenated along the dimension of the dynamic subband to obtain cross domain fusion features $U_{ST} = (s_\delta \oplus t_\delta, s_\theta \oplus t_\theta, s_\alpha \oplus t_\alpha, s_\beta \oplus t_\beta, s_\gamma \oplus t_\gamma)$, where $\oplus$ represents the element-wise addition method. Secondly, the obtained cross domain fusion features U_{ST} are convolved and batch normalized, and then output to the variant SE block. The variant SE block is a computational unit that includes two steps: squeezing and excitation. The squeezing is achieved through the adaptive average pooling of the variant SE block, which compresses the global spatial information of the data after cross-domain fusion features U_{ST} are convolved and batch normalized into a channel descriptor $A \in R^{1\times1\times(V+V)}$, fully capturing channel dependencies. Here, is the number of channels corresponding to the EEG signal data, and A is calculated by a descriptive statistic. Therefore, the statistic of the Vth

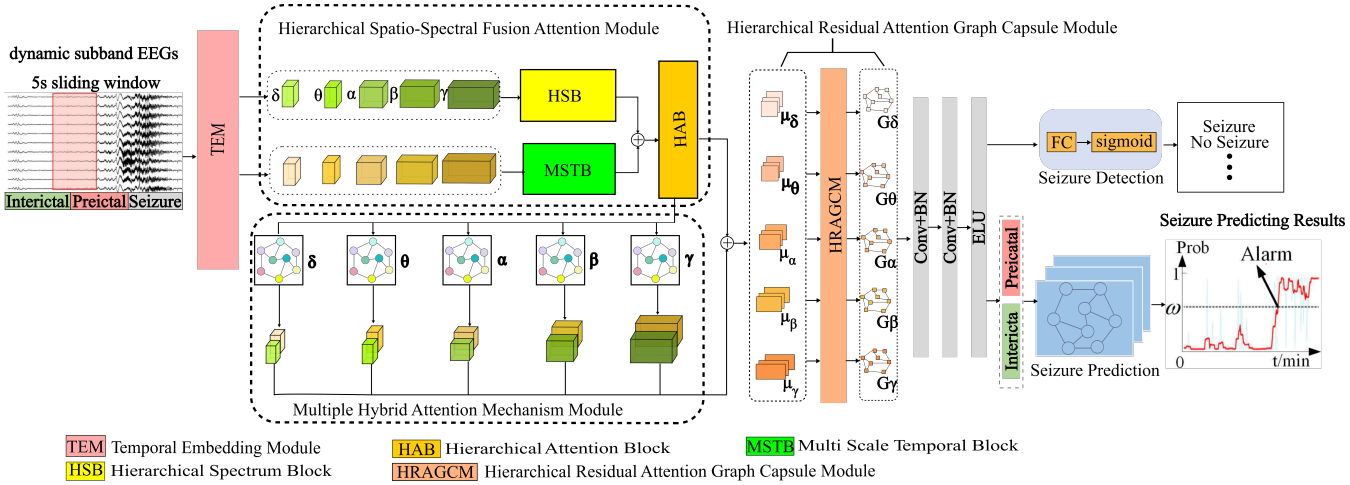

Fig. 1. Overall network framework.

channel of A is defined as:

$$A_V = \frac{1}{\rho}\sum_{l=1}^{\rho} H(S_w^V t_w^V)(V,1,l), \tag{1}$$

in which S_w^V is the frequency-domain feature of the Vth channel inside the brainwave w, and w is δ, θ, α, β and γ. According to $t_w = [t_w^1, t_w^2, \ldots, t_w^V]$, is the temporal-domain feature of the Vth channel inside the brainwave w. is a channel cascade function and represents the spatial dimension ρ.

Here, a simple self-gating mechanism is adopted to model the channel dependencies in order to obtain the channel weight $\hat{A}$ that are adapted to the specific channel descriptor A. The weight parameter $\hat{A}$ of the Vth channel is defined as follows:

$$\hat{A}_V = \alpha(W_2\beta(W_1 A_V)), \tag{2}$$

where, and $W_2 \in R^{(V+V)\times((V+V)/r)}$ represent the weights of two fully connected layers, and $\beta(\cdot)$ respectively represent the ELU activation function and the sigmoid activation function, is the number of EEG channels, and r is the compression ratio parameter, which represents the bottleneck of the self-gating mechanism.

Perform channel-by-channel operations on the captured relevant features, adaptively recalibrate the importance of the cross domain aggregation features U_{ST} of the five brain waveforms in different channels, and jointly improve the distinguishability of the features. It is defined as follows:

$$\begin{aligned}\hat{S}_w &= [\hat{A}_1 S_w^1, \hat{A}_2 S_w^2, \ldots, \hat{A}_V S_w^V] \\ \hat{T}_w &= [\hat{A}_{V+1} t_w^1, \hat{A}_{V+2} t_w^2, \ldots, \hat{A}_{V+V} t_w^V]\end{aligned} \tag{3}$$

Finally, through adaptive average pooling operation, the distinguishable cross domain fusion features corresponding to the five brain waveforms are obtained $U_{\hat{S}\hat{T}} = (\hat{S}_\delta \oplus \hat{T}_\delta, \hat{S}_\theta \oplus \hat{T}_\theta, \hat{S}_\alpha \oplus \hat{T}_\alpha, \hat{S}_\beta \oplus \hat{T}_\beta, \hat{S}_\gamma \oplus \hat{T}_\gamma)$.

C. Multiple Hybrid Attention Mechanism

To eliminate redundant information in spatio-temporal cross domain fusion features, the concept of cross dimensional interaction is introduced into the channel based attention module (CBAM) [18]. By integrating an improved dual attention module (DAM) [19] and a tripartite attention module (TAM) [20], a multi hybrid attention mechanism is designed. This mechanism captures interaction information between spatial and channel dimensions of the input tensor, fully exploits the spatio-temporal spectral fusion features of EEG signals, enhances the expression ability of multi scale features in complex EEG signals, and improves the model's learning ability and robustness.

The multi-hybrid attention mechanism takes the output of the hierarchical attention module as input. The data is processed through a 2D convolutional layer and a batch normalization layer before being fed into the hybrid attention module. This hybrid attention module is composed of three parallel spatial attention modules and channel attention modules, each with three distinct branches. Finally, the multi-hybrid fusion features $U = (\mu_\delta, \mu_\theta, \mu_\alpha, \mu_\beta, \mu_\gamma)$ corresponding to different brain wave patterns with frequencies of δ, θ, α, β and γ are obtained [21].

D. Hierarchical Residual Attention Graph Capsule Module

To obtain the fusion feature graph information of EEG signals, as shown in Fig.2. This paper proposes a hierarchical residual attention graph capsule module based on a multi layer neural network. This module is mainly composed of three parts: the Residual Attention Graph Capsule (RAGC) block, a fully connected layer, and a capsule attention block.

1) Residual Attention Graph Capsule Block

Transform the fused features of the above five frequency bands into a graph representation as $G_w = \{V_w, U_w, \hat{A}_{vs}\}$, and simultaneously depict the spatiotemporal correlation of the epileptogenic regions under different rhythms, where $w \in [\delta, \theta, \alpha, \beta, \gamma]$. V_w is a vertex set with $|V_w| = \lambda$ nodes, and the node attributes are the spatiotemporal spectrum fused features under U_w. The dependencies between nodes are determined by the dynamic adjacency matrix $\hat{A}_{vs}$. The feature on node a of G_w in the l-th residual graph capsule is represented as $h_a^{(m,l)}$, where $a = 1, 2, \ldots, \lambda$. $H^{(m,l)}$ is the dynamic feature of the l-th layer of the residual graph capsule network, and $H^{(w,0)} = U_w$. Therefore, the aggregator in the l-th residual graph capsule network for G_w is defined as

$$\varpi_{res}^{(m,l)} = \frac{\sum_{b=1}^{\lambda} \hat{A}_{a,b} \times \alpha_1(h_a^{(m,l)})\beta_1^{(m,l)}\beta_2^{(m,l)}}{\sum_{b=1}^{\lambda} \hat{A}_{a,b}}, \tag{4}$$

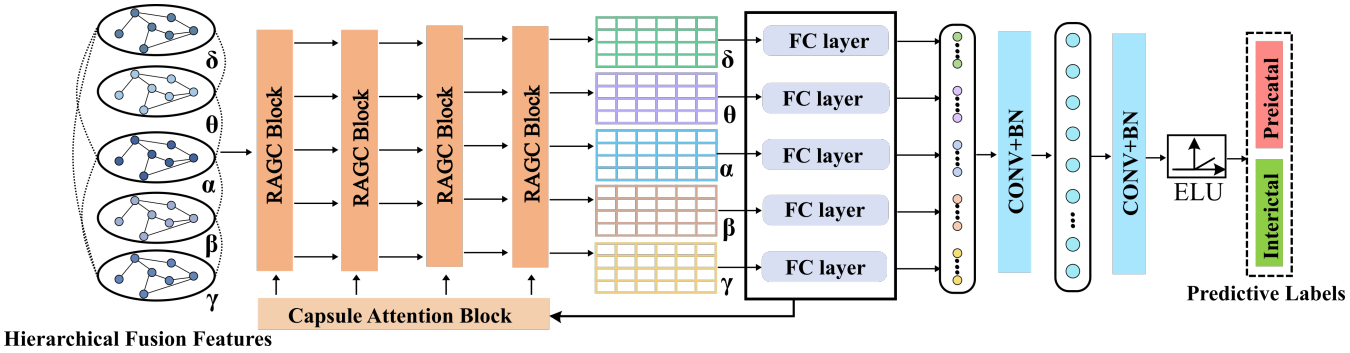

Fig. 2. Hierarchical residual attention graph capsule module.

in which $\hat{A}_{a,b}$ is the $\hat{A}_{vs}$ of the (a,b) channel, and m corresponds to five different brain frequency bands respectively. $l = 1, 2, \ldots, L-1$ represents the number of layers. $\beta_1^{(m,l)}$ and $\beta_2^{(m,l)}$ are the weight matrices in the l-th layer, and $\alpha_1(\cdot)$ is the ELU activation function.

Since the coupling intensities $\hat{A}_{a,b}/\sum_{b=1}^{\lambda}\hat{A}_{a,b}$ obtained from different dynamic subband EEG are different under the gating response $\alpha_1(h_b^{(m,l)}\beta_1^{(m,l)})\beta_2^{(m,l)}$ at different times. In order to describe the dynamic information and local features in the EEG signals, a residual updater $\nu_{res}^{(m,l)}$ is set after the aggregator $\varpi_{res}^{(m,l)}$ at the l-th layer, and its definition is as follows

$$\nu_{res}^{(m,l)} = \beta_2^{(m,l)}\left(h_{Nb}^{(m,l)} + \sum_{n=0}^{l-1}\beta_1^{(m,l)}\right), \tag{5}$$

where, $h_{Nb}^{(m,l)}$ is the neighborhood attribute obtained from $\varpi_{res}^{(m,l)}$.

According to equations (4) and (5), the formula of the residual graph capsule network on the intracortical graph G_w under different rhythms is obtained as follows to generate fine spatiotemporal response feature

$$H^{(m,l+1)} = \alpha_2\left(D^{(-1)}\hat{A}_{a,b}\alpha_1\left(H^{(m,l)}\beta_1^{(m,l)}\right)\beta_2^{(m,l)} + \sum_{n=0}^{l-1}H^{(m,l)}\right), \tag{6}$$

in which $l = 1, 2, \ldots, L-1$ represents the layer number. m corresponds to five different brain frequency bands respectively. $D = \sum_m \hat{A}_{vs}^{mn}$ is the degree matrix of $\hat{A}_{vs}$.$\beta_1^{(m,l)}$ and $\beta_2^{(m,l)}$ are the weight matrices in the l-th layer. α_1 and α_2 are ELU activation functions.

2) Capsule Attention Block

To further capture the intrinsic relationships between local and global aspects of EEG information, a capsule attention module is employed. Specifically, an attention mechanism is introduced into the dynamic routing between capsule layers. By utilizing attention scores and transformation matrices, this module distinguishes different temporal states of EEG signals. The length of each capsule represents the probability of occurrence of each state, as shown in Fig.3.

The $\mathrm{caps}_i(i = 1, 2, \ldots, \mathrm{caps}_{\sigma_1})$ is defined to represent the i-th node capsule. After multiplying caps_i by the weight matrix $W_{ij}(j = 1, 2, \ldots, \mathrm{caps}_{\sigma_2})$,the predicted feature $\mathrm{caps}_{j|i}$ containing high-level features is obtained as follows

$$\mathrm{caps}_{j|i} = W_{ij} \otimes \mathrm{caps}_i, \tag{7}$$

where, the matrix W_{ij} represents the internal spatial relationship between local features and global features, and $\otimes$ represents the element wise multiplication operation.

Multiply the predicted feature $\mathrm{caps}_{j|i}$ by different coupling coefficients, and then obtain the state capsule caps_j by a summation operation, which is used to represent the correlation between the key information, as follows

$$\mathrm{caps}_j = \sum_i h_{ij} \otimes \mathrm{caps}_{j|i}, \tag{8}$$

in which, h_{ij} is the coupling coefficie nt with an equal-probability distribution obtained after Softmax routing, which is expressed as

$$h_{ij} = \frac{e^{con_{ij}}}{\sum_{\sigma_2} e^{con_{i\sigma}}}, \tag{9}$$

and, where con_{ij} is the logarithmic prior probability initialized to 0. The log prior probability can be discriminatively learned simultaneously with all other weights. This process depends on the positions and types of the two capsules, rather than the current input content.

Subsequently, in order to make the length of the state capsule represent the probability of each state of the input EEG signal, a squash operation is used to normalize the vectors. This is achieved through a nonlinear activation function to ensure that the vector length of $cons_j$ falls between 0 and 1. The formula is as follows

$$act_j = \frac{\|\mathrm{caps}_j\|_2^2}{1 + \|\mathrm{caps}_j\|_2^2}\frac{\mathrm{caps}_j}{\|\mathrm{caps}_j\|_2}, \tag{10}$$

in which act_j is the output of the activation vector of the state capsule j.

The consistency between act_j and $caps_{j|i}$ is evaluated by calculating their scalar product, and this process is expressed as

$$con_{ij} = con_{ij} + act_{ij} \otimes \mathrm{caps}_{j|i}, \tag{11}$$

where, the scalar product between act_j and $caps_{j|i}$ is used to update their respective coupling coefficients. If the scalar product between act_j and $caps_{j|i}$ is large, it proves that there is a strong correlation between the feature $caps_{j|i}$ and the activation vector act_j, which indicates that this capsule is more suitable for describing the features during the inter-ictal or pre-ictal period. Therefore, the weight of this capsule vector can be increased in the next iteration to dynamically adjust the weights.

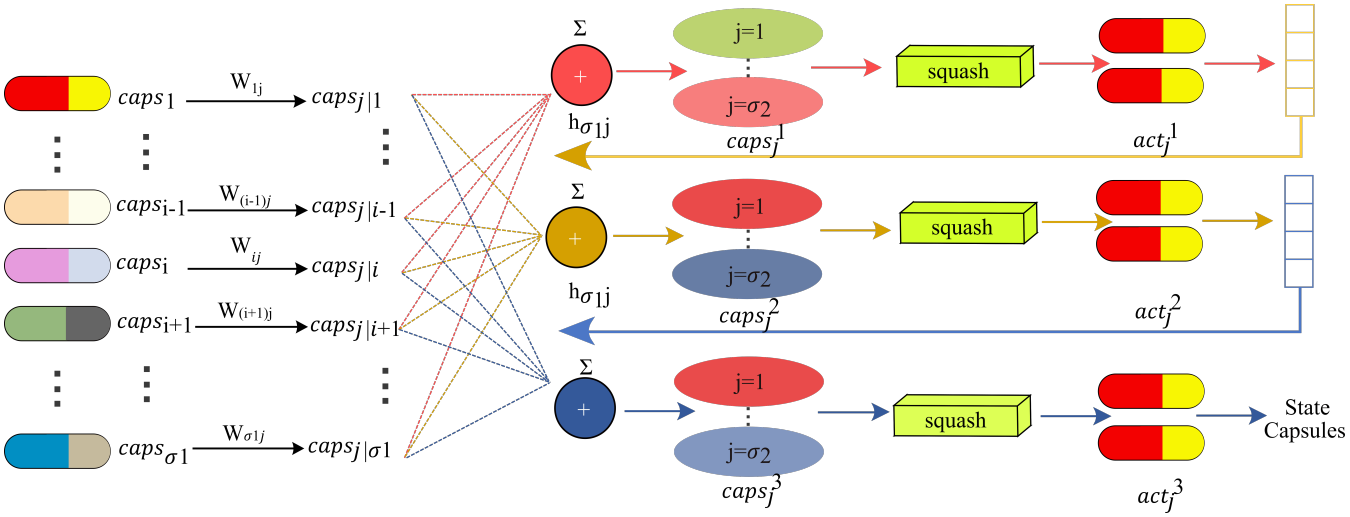

Fig. 3. Capsule attention block.

III. Experimentation and Results

A. Datasets

1) CHB-MIT dataset

The CHB-MIT dataset is one of the publicly available continuous long-term epileptic seizure datasets, created by researchers from the Massachusetts Institute of Technology (MIT) and Boston Children's Hospital (CHB). This dataset can be downloaded from the following website: https://physionet.org/content/chbmit/1.0.0/. This dataset uses the bipolar lead technology of the international 10-20 system to collect EEG signals from 21 electrodes, with a sampling rate of 256 hz and a resolution of 16 bits.

This dataset was collected from 23 patients with refractory epilepsy, totaling 24 cases (among them, cases chb21 and chb01 were from the same patient, with an interval of 1.5 years between the two recordings. Case chb24 was added to the dataset in December 2010 and is currently not included in the "Subject Information"). There are 5 male patients (aged 3 to 22 years) and 17 female patients (aged 1.5 to 19 years) in the dataset, and the gender and age data of one patient are missing. The dataset contains a total of 967.55 hours of continuous EEG recordings, capturing 198 epileptic seizures. The annotation files of this dataset provide EEG channel information as well as the start and end times of epileptic seizures.

The objective of the epileptic seizure prediction algorithm is to determine that a seizure is imminent within a specific time frame before the epileptic seizure occurs, that is, to intervene within the Seizure Prediction Horizon (SPH) [22]. In this process, the interictal period of epilepsy is regarded as the normal period of the patient, while the Seizure Onset Period (SOP) is the period when the patient exhibits epileptic symptoms. By accurately predicting the time of an epileptic seizure, it can provide an opportunity for medical staff to intervene, thus achieving better management and treatment of epilepsy.

To ensure the effectiveness and rationality of the SPH and the SOP in the epileptic seizure prediction algorithm, the following criteria need to be met:

a). An appropriate time interval should be set between the SPH and the SOP, so as to provide doctors, patients and their families with sufficient time to take necessary measures to deal with the impending epileptic seizure.

b). Premature or excessively long time intervals should be avoided, because this will not only cause unnecessary anxiety to patients, but also increase the difficulty of prediction.

Based on the above considerations, this paper selects records from the CHB-MIT dataset that contain at least two epileptic seizures and have an inter-ictal period of three hours for the evaluation of epileptic seizure prediction. Specifically, the pre-ictal period is defined as at least 15 minutes. For the SPH, the signal data from 15 minutes to 1 hour before the seizure are selected. For the inter-ictal period, the signal data from 2 hours before to 2 hours after the seizure are used. This setting aims to ensure the prediction accuracy while taking into account the practical needs and response strategies of patients and the medical team.

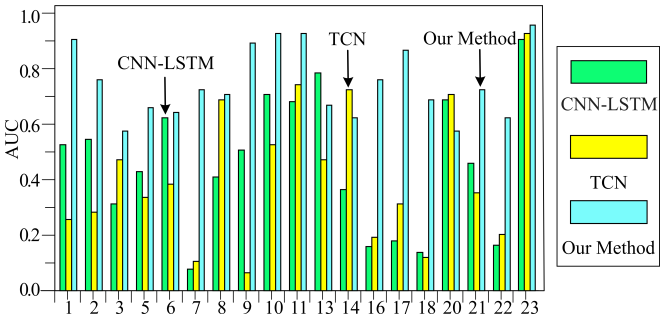

Fig. 4. Comparison of AUC between CNN-LSTM, TCN, and the method used in this paper.

2) TUSZ dataset

The Temple University Hospital (TUH) EEG Seizure Corpus (TUSZ) dataset is the only open-source EEG dataset containing annotations for multiple types of epileptic seizures. Considering the superiority of the TUSZ dataset in long-term maintenance and updates, we adopted the TUSZ v2.0.1 version updated on October 4, 2023, as the experimental data source for epilepsy detection. Since the sampling rates of different patients in the TUSZ dataset vary, we resampled the EEG recordings to 256 hz.

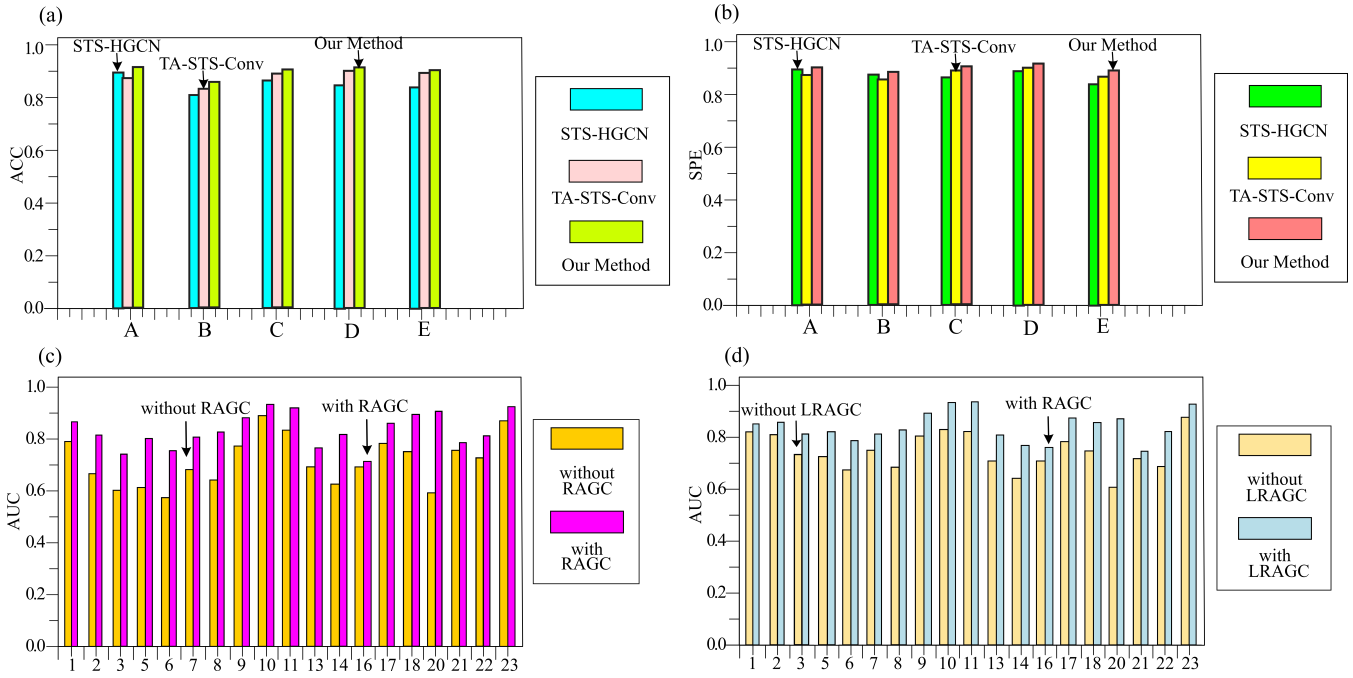

Fig. 5. (a)ACC comparison among STS-HGCN, TA-STS-Conv and our method.(b)SPE comparison among STS-HGCN, TA-STS-Conv and our method.(c)AUC comparison of with and without RAGC.(d)AUC comparison of with and without LRAGC.

B. Evaluation metrics

To compare with other existing methods, this paper adopts five of the most commonly used evaluation indicators in epilepsy-related research to assess the overall performance of different models on the test dataset. These five indicators are Accuracy (ACC), Specificity (SPE), Sensitivity (SEN), False Positive Rate (FPR), and Area Under the Curve (AUC).

Among these indicators, ACC represents the proportion of correctly predicted samples among all samples; the higher this value is, the stronger the model's ability to distinguish between different categories. SPE represents the proportion of correctly predicted negative results among all actual negative results. The higher this value is, the better the model's ability to predict and identify the inter-ictal period of epilepsy. SEN reflects the proportion of correctly predicted epileptic seizures. The FPR represents the number of incorrect predictions, and the FPR/h specifically represents the number of incorrect predictions per hour. Finally, the AUC is used to evaluate the classification performance of the model, and its definition formula is as follows:

$$ACC = \frac{TP + TN}{TP + TN + FN + FP}, \tag{12}$$

$$SPE = \frac{TN}{FP + TN}, \tag{13}$$

$$SEN = \frac{TP}{TP + FN}, \tag{14}$$

$$FPR = \frac{FP}{FP + TN}, \tag{15}$$

$$AUC = \frac{1}{2}\sum_{\tau=1}^{\xi-1}(x_{\tau+1} - x_\tau)(y_\tau - x_{\tau+1}), \tag{16}$$

where, TP, FP, TN, and FN in Formulas (12) to (16) represent true positive, false positive, true negative, and false negative, respectively. In Formula (16), x and y are the continuous coordinate points on the ROC curve, which are represented by $\{(x_1 = 0, x_\xi = 1)|(x_1, y_1), (x_2, y_2), \ldots, (x_\xi, y_\xi)\}$.

C. Single Subject Experiment

To demonstrate the advantages of the proposed method over traditional structured Capsule Networks (CapsNet) and Graph Convolutional Networks (GCN), this paper compares the performance of three different models for seizure detection on the TUSZ dataset. As shown in the table I, for the training set, 1D-CapsNet [23] and the STS-HGCN [24] achieved average ACC of 75.21% and 81.44%, respectively, and average SPE of 82.64% and 87.34%, respectively. In contrast, the proposed model achieved an average accuracy of 86.28% and an average specificity of 93.47%. Additionally, the average SEN of the comparison methods was 59.88% and 66.32%, respectively, while the proposed method achieved an average sensitivity of 70.12%.

To demonstrate the advantages of the method proposed in this paper over traditional structured CapsNet and GCN, we compared the performance of three different models for seizure detection on the TUSZ test set, with the results shown in the test set section of the table I. The findings indicate that the 1D-CapsNet and the STS-HGCN achieved average ACC of 73.48% and 77.31%, respectively, and average SPE of 80.43% and 83.10%, respectively. In contrast, the proposed model achieved an average accuracy of 84.11% and an average SPE of 91.85%. Additionally, the average SEN of the comparison methods was 58.35% and 61.57%, respectively, while the proposed method achieved an average SEN of 67.74%.

To validate the effectiveness of the method proposed in this paper for seizure classification, single-subject experiments were conducted on the TUSZ dataset, comparing the proposed method with LSTM, CNN-LSTM, Diffusion Convolutional Recurrent Neural Networks (DCRNN), and

TABLE I
RESULTS OF EPILEPTIC SEIZURE DETECTION PERFORMANCE ON THE TUSZ TRAINING SET AND TEST SET

Method	Training set			Test set		
	ACC(%)	SPE(%)	SEN(%)	ACC(%)	SPE(%)	SEN(%)
1D-CapsNet	75.21	82.64	59.88	73.48	80.43	58.35
STS-HGCN	81.44	87.34	66.32	77.31	83.10	61.57
Our method	86.28	93.47	70.12	84.11	91.85	67.74

TABLE II
EXPERIMENTAL RESULTS OF EPILEPTIC SEIZURE CLASSIFICATION ON THE TUSZ DATASET

Method	ACC(%)	SPE(%)	F1(%)	Precision(%)
LSTM	71.1	71.8	65.7	65.7
CNN-LSTM	68.5	75.4	69.5	68.4
DCRNN	77.1	79.8	72.4	61.6
Mamba	72.6	80.3	69.3	70.2
Our method	78.8	81.6	74.8	75.2

the Mamba model. As shown in the table II, LSTM, CNN-LSTM, DCRNN, and Mamba achieved average ACC of 71.1%, 68.5%, 77.1%, and 72.6%, respectively; average SPE of 71.8%, 75.4%, 79.8%, and 80.3%, respectively; F1 scores of 65.7%, 69.5%, 72.4%, and 69.3%, respectively; and precision values of 65.7%, 68.4%, 61.6%, and 70.2%, respectively. In contrast, the proposed model demonstrated superior performance with an average ACC of 78.8%, an average SPE of 81.6%, an F1 score of 74.8%, and a precision of 75.2%.

In seizure prediction experiments, the effectiveness of the method proposed in this paper was validated by comparing it with various advanced methods on the CHB-MIT dataset. As shown in Fig. 4, comparisons were made with advanced traditional deep neural networks, including CNN-LSTM [15] and TCN [15]. The results demonstrate that, for the vast majority of patients, the AUC of the proposed method was significantly higher than that of CNN-LSTM and TCN, further proving the superiority of the proposed method in predicting seizures.

To demonstrate the advantages of the method adopted in this paper over the traditional structured CapsNet and GCN, we compared the performance of three different models in predicting epileptic seizures on the CHB-MIT dataset, as shown in Table III. The results show that, on the CHB-MIT dataset, the average ACC of 1D-CapsNet and the Spiking-GCNN [25]reaches 85.21% and 90.71% respectively, and the average SPE is 85.70% and 91.18% respectively. In contrast, the model implemented in this paper has an average ACC of 96.04% and an average SPE of 99.16%. Meanwhile, the average FPR of the comparative methods is 0.295/h and 0.091/h respectively, while the average FPR of the method in this paper is only 0.026/h.

To demonstrate the superior performance of the method proposed in this paper compared to the state-of-the-art techniques, we conducted comparative experiments with other multi layer deep neural networks. Table IV shows the performance of seven different methods in predicting epileptic seizures on the CHB-MIT dataset, including comparisons of ACC, SPE, SEN, and FPR/h. It can be seen that the proposed method performs excellently in epileptic seizure prediction, achieving higher accuracy, sensitivity, and specificity, while also having a lower FPR/h.

D. Cross Subject Experiment

In this paper, the advanced nature of the proposed method in predicting seizures was demonstrated by comparing its performance with two models, the STS-HGCN and the TA-STS-Conv [26], on the CHB-MIT dataset. Table V presents the performance results of these three models on the CHB-MIT dataset, including three metrics: SEN, FPR/h, and AUC.

In terms of average SEN and AUC, the method proposed in this paper achieved 89.86% and 0.889, respectively, which are superior to the 86.87% and 0.850 of the STS-HGCN, as well as the 87.67% and 0.870 of the TA-STS-Conv. Furthermore, the average FPR/h of the proposed method was 0.103/h , which is significantly lower than the 0.126/h of STS-HGCN and the 0.112/h of TA-STS-Conv.These results indicate that, after improvement and optimization, the proposed method performs better in distinguishing seizure states and can more accurately differentiate between inter-ictal and pre-ictal states.

As shown in Fig.5(a), it visually presents the performance results of the three models in terms of average ACC across five datasets in cross-subject experiments. It can be seen that the method proposed in this paper achieved higher prediction accuracy, indicating that the model used in this paper has stronger classification ability.

As shown in Fig.5(b), it visually presents the performance results of the three models in terms of average SPE across five datasets in cross-subject experiments. It can be seen that the method proposed in this paper has higher prediction specificity, indicating better prediction and recognition of inter-ictal periods.

E. Influence of the Networkt

Overall, compared to the traditional CapsNet structure, the Residual Attention Graph Capsule Network (RAGC) proposed in this paper demonstrates stronger generalization ability. The advantage of this network lies in its ability to effectively represent complex temporal information and precise spatial relationships in EEG signals. By fusing channels, it captures and integrates spatio-temporal features, enabling more effective discrimination of seizure states. Fig.5(c) shows the AUC results of single subject experiments, while Table VI presents the average performance results of the other four metrics. When processing EEG signals, the RAGC Network has stronger representation ability and prediction accuracy, providing a more reliable solution for distinguishing seizure states.

After a comprehensive analysis, the Hierarchical Residual Attention Graph Capsule Module introduced in this paper demonstrates more significant generalization ability

TABLE III
PERFORMANCE COMPARISON OF 1D CAPSNET, SPIKING GCNN, AND OUR METHOD ON THE CHB-MIT DATASET

Patient	1D-CapsNet			Spiking-GCNN			Our method		
	ACC(%)	SPE(%)	FPR/h	ACC(%)	SPE(%)	FPR/h	ACC(%)	SPE(%)	FPR/h
1	88.92	89.22	0.249	96.28	96.14	0.032	96.47	99.66	0.000
2	89.07	90.78	0.281	92.18	98.31	0.025	97.89	99.70	0.000
3	84.53	83.99	0.358	96.09	93.89	0.071	96.35	99.64	0.000
5	84.84	82.73	0.345	91.58	88.84	0.125	95.76	99.58	0.027
6	81.03	80.56	0.337	87.61	89.28	0.112	94.89	99.03	0.000
7	90.27	91.25	0.202	89.69	91.25	0.086	98.38	99.27	0.000
8	84.35	85.89	0.317	86.49	90.21	0.098	96.48	99.85	0.000
9	90.74	89.87	0.223	90.51	88.89	0.112	98.86	99.71	0.000
10	88.56	88.23	0.220	89.78	89.26	0.113	97.02	99.50	0.000
11	89.40	87.56	0.201	91.39	90.34	0.095	98.46	99.22	0.114
13	80.73	83.05	0.397	89.89	90.27	0.098	94.96	98.52	0.000
14	78.68	80.62	0.257	89.01	85.96	0.142	92.75	98.79	0.000
16	76.44	79.73	0.534	89.98	89.97	0.112	89.63	98.81	0.065
17	88.62	87.30	0.201	89.12	90.28	0.098	96.89	99.44	0.049
18	84.45	85.39	0.298	92.05	91.64	0.085	96.98	98.56	0.000
20	90.09	91.52	0.176	92.85	95.61	0.044	98.77	98.76	0.247
21	79.33	85.21	0.488	91.34	93.14	0.069	92.34	97.90	0.000
22	88.34	83.19	0.223	89.85	89.04	0.114	97.86	99.20	0.000
23	80.51	82.11	0.302	87.89	90.16	0.099	93.96	98.87	0.000
mean	85.21	85.70	0.295	90.71	91.18	0.091	96.04	99.16	0.026

TABLE IV
COMPARISON OF SINGLE SUBJECT EXPERIMENTAL PERFORMANCE OF THE METHOD PROPOSED IN THIS ARTICLE ON THE CHB-MIT DATASET

author	method	Number of patients	ACC(%)	SPE(%)	SEN(%)	FPR/h
Xu et. al	RF+GBDT	-	91.76	-	91.87	0.083
Kapoor et. al	AdaBoost+DT+RF	-	93.81	88.57	91.68	-
Hussein et. al	PIL-SEG+SA	10	80.61	80.71	79.32	0.239
Lu et. al	CNN-LSTM	19	80.96	68.80	78.04	-
Alizadeh et. al	KP-Ki-KD+CNN	16	97.1	96.3	97.5	-
Gao et. al	MSPPNet	16	-	-	93.8	0.054
Our method	SSHRAGC	19	96.04	99.16	94.81	0.026

TABLE V
COMPARISON OF CROSS SUBJECT EXPERIMENTAL PERFORMANCE OF THE METHOD PROPOSED IN THIS ARTICLE ON THE CHB-MIT DATASET

Grop	STS-HGCN			TA-STS-Conv			Our method		
	SEN(%)	FPR/h	AUC	SEN(%)	FPR/h	AUC	SEN(%)	FPR/h	AUC
A	91.22	0.077	0.851	85.71	0.081	0.849	91.37	0.078	0.881
B	85.21	0.162	0.818	83.33	0.198	0.822	87.39	0.141	0.853
C	88.19	0.156	0.842	91.34	0.058	0.875	92.01	0.128	0.918
D	86.11	0.041	0.902	90.51	0.044	0.915	91.30	0.096	0.914
E	83.64	0.192	0.837	87.45	0.177	0.887	89.24	0.072	0.901
mean	86.87	0.126	0.850	87.67	0.112	0.870	89.86	0.103	0.889

TABLE VI
PERFORMANCE COMPARISON OF THE EFFECTS OF RESIDUAL ATTENTION GRAPH CAPSULE NETWORK AND HIERARCHICAL RESIDUAL ATTENTION GRAPH CAPSULE MODULE ON CHB-MIT DATASET

method	ACC(%)	SPE(%)	SEN(%)	FPR/h
Without RAGC	90.32	94.57	89.23	0.159
With RAMG	93.75	95.31	90.48	0.126
Without LRAGC	93.75	95.31	90.48	0.126
With LRAGC	96.01	98.92	93.14	0.038

compared to the Single layer Residual Attention Graph Capsule Network (LRAGC). Its advantage lies in the ability to effectively represent complex temporal information and precise spatial relationships in EEG signals. By fusing channels, it captures and integrates spatio-temporal features, enabling more effective discrimination of seizure states. Fig.5(d) shows the AUC results of single-subject experiments, while Table VI lists the average performance results of the other four metrics.

IV. CONCLUSIONS

This paper proposes a novel epileptic seizure predictor based on EEG and implements a spatio-temporal spectrum analysis method based on hierarchical residual attention graph capsules for analyzing epileptic EEG signals. We adopt a sub-network, the hierarchical spatio-temporal spectrum fusion attention network, to capture the temporal states of the epileptic cortex under different rhythms, and use a multiple hybrid attention convolutional network to capture spatial couplings and eliminate false couplings. Finally, by introducing the hierarchical residual attention graph capsule module, we overcome the limitations of convolutional neural networks in perceiving the relative positional relationships between local features and address the deficiencies of traditional capsule networks in handling EEG biosignal features. When tested on the publicly available TUSZ

epileptic EEG dataset, for the epileptic seizure prediction experiment of a single subject, the average accuracy is 84.11%, the average specificity is 91.85%, and the sensitivity is 67.74%. For the epileptic seizure classification experiment of a single subject, the average accuracy is 78.8%, the average specificity is 81.6%, the F1 score is 74.8%, and the precision is 75.2%. When tested on the publicly available CHB-MIT epileptic EEG dataset, for the single subject experiment, the average accuracy is 96.04%, the average specificity is 99.16%, and the average FPR/h is 0.026/h. In the cross subject test, the average sensitivity is 89.86%, the average AUC is 0.889, and the average FPR/h is 0.103/h, which further improves the accuracy of the epileptic seizure prediction method.

REFERENCES

[1] J. Moini, A. LoGalbo, and R. Ahangari, *Foundations of the mind, brain, and behavioral relationships: understanding physiological psychology*. Elsevier, 2023.

[2] E. Gonzalez-Viana, A. Sen, A. Bonnon, and J. H. Cross, "Epilepsies in children, young people, and adults: summary of updated nice guidance," *British Medical Journal*, vol. 10, no. 1446, pp. 378–391, 2022.

[3] M. Mula, "Suicidality and antiepileptic drugs in people with epilepsy: an update," *Expert Review of Neurotherapeutics*, vol. 22, no. 5, pp. 405–410, 2022.

[4] R. Whitney, S. Sharma, and R. Ramachandrannair, "Sudden unexpected death in epilepsy in children," *Developmental Medicine & Child Neurology*, vol. 65, no. 9, pp. 1150–1156, 2023.

[5] I. Jemal, A. Mitiche, and N. Mezghani, "A study of eeg feature complexity in epileptic seizure prediction," *Applied Sciences*, vol. 11, no. 4, pp. 1579–1594, 2021.

[6] A. P. d. A. Boleti, P. H. d. O. Cardoso, B. E. F. Frihling, L. F. R. N. de Moraes, E. A. C. Nunes, L. T. H. Mukoyama, E. A. C. Nunes, C. M. E. Carvalho, M. L. R. Macedo, and L. Migliolo, "Pathophysiology to risk factor and therapeutics to treatment strategies on epilepsy," *Brain Sciences*, vol. 14, no. 1, pp. 71–101, 2024.

[7] H. O. Heyne, F.-D. Pajuste, J. Wanner, J. I. Daniel Onwuchekwa, R. Mägi, A. Palotie, FinnGen, E. B. research team, R. Kälviainen, and M. J. Daly, "Polygenic risk scores as a marker for epilepsy risk across lifetime and after unspecified seizure events," *Nature Communications*, vol. 15, no. 1, pp. 6277–6289, 2024.

[8] S. Chaibi, C. Mahjoub, W. Ayadi, and A. Kachouri, "Epileptic eeg patterns recognition through machine learning techniques and relevant time–frequency features," *Biomedical Engineering/Biomedizinische Technik*, vol. 69, no. 2, pp. 111–123, 2024.

[9] H. Anwar, Q. U. Khan, N. Nadeem, I. Pervaiz, M. Ali, and F. F. Cheema, "Epileptic seizures," *Discoveries*, vol. 8, no. 2, pp. e110–e129, 2020.

[10] P. Krkoska, V. Kokosova, M. Dostal, D. Vlazna, M. Kerkovsky, M. Straka, R. Gerstberger, K. Matulova, P. Ovesna, and B. Adamova, "Assessment of lumbar paraspinal muscle morphology using mdixon quant magnetic resonance imaging (mri): a cross-sectional study in healthy subjects," *Quantitative Imaging in Medicine and Surgery*, vol. 14, no. 8, pp. 6015–6036, 2024.

[11] M. Safari, R. Shalbaf, S. Bagherzadeh, and A. Shalbaf, "Classification of mental workload using brain connectivity and machine learning on electroencephalogram data," *Scientific Reports*, vol. 14, no. 1, pp. 9153–9168, 2024.

[12] J. Batista, M. F. Pinto, M. Tavares, F. Lopes, A. Oliveira, and C. Teixeira, "Eeg epilepsy seizure prediction: the post-processing stage as a chronology," *Scientific Reports*, vol. 14, no. 1, pp. 407–420, 2024.

[13] Y. Gao, X. Chen, A. Liu, D. Liang, L. Wu, R. Qian, H. Xie, and Y. Zhang, "Pediatric seizure prediction in scalp eeg using a multi-scale neural network with dilated convolutions," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, no. 1, pp. 1–9, 2022.

[14] X. Ding, W. Nie, X. Liu, X. Wang, and Q. Yuan, "Compact convolutional neural network with multi-headed attention mechanism for seizure prediction," *International Journal of Neural Systems*, vol. 33, no. 3, pp. 2 350 014–2 350 032, 2023.

[15] Z. Georgis-Yap, M. R. Popovic, and S. S. Khan, "Supervised and unsupervised deep learning approaches for eeg seizure prediction," *Journal of Healthcare Informatics Research*, vol. 8, no. 2, pp. 286–312, 2024.

[16] I. B. Slimen, L. Boubchir, and H. Seddik, "Epileptic seizure prediction based on eeg spikes detection of ictal-preictal states," *Journal of Biomedical Research*, vol. 34, no. 3, pp. 162–170, 2020.

[17] J. Maan and A. Prasad, "Wave packet transform and wavelet convolution product involving the index whittaker transform," *The Ramanujan Journal*, vol. 64, no. 1, pp. 19–36, 2024.

[18] Y. Wang, W. Wang, Y. Li, Y. Jia, Y. Xu, Y. Ling, and J. Ma, "An attention mechanism module with spatial perception and channel information interaction," *Complex & Intelligent Systems*, vol. 10, no. 4, pp. 5427–5444, 2024.

[19] Y. Xue, X. Han, and Z. Wang, "Self-adaptive weight based on dual-attention for differentiable neural architecture search," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 4, pp. 6394–6403, 2024.

[20] Y. Tang, L. Zhang, Q. Teng, F. Min, and A. Song, "Triple cross-domain attention on human activity recognition using wearable sensors," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 5, pp. 1167–1176, 2022.

[21] S. Ge, W. Wang, X. Shi, and M. Wang, "Hierarchical spatio-temporal spectral multiple hybrid attention mechanisms for seizure prediction," *International Journal of Computer Science*, vol. 51, no. 3, pp. 317–328, 2024.

[22] J. Batista, M. F. Pinto, M. Tavares, F. Lopes, A. Oliveira, and C. Teixeira, "Eeg epilepsy seizure prediction: the post-processing stage as a chronology," *Scientific Reports*, vol. 14, no. 1, pp. 407–420, 2024.

[23] S. Toraman, "Automatic recognition of preictal and interictal eeg signals using 1d-capsule networks," *Computers & Electrical Engineering*, vol. 91, no. 1, pp. 107 033–107 346, 2021.

[24] Y. Li, Y. Liu, Y. Z. Guo *et al.*, "Spatio-temporal-spectral hierarchical graph convolutional network with semisupervised active learning for patient-specific seizure prediction," *IEEE Transactions on Cybernetics*, vol. 52, no. 11, pp. 12 189–12 204, 2021.

[25] Q. Lian, Y. Qi, G. Pan *et al.*, "Learning graph in graph convolutional neural networks for robust seizure prediction," *Journal of Neural Engineering*, vol. 17, no. 3, pp. 035 004–035 028, 2020.

[26] T. Yu, "Triple-attention-based spatio-temporal-spectral convolutional network for epileptic seizure prediction," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, no. 1, pp. 3915–3926, 2023.