Long-Interval Spatio-Temporal Graph Convolution for Brain Disease Diagnosis

Shengrong Li, Qi Zhu, Donghai Guan, Bo Shen, Li Zhang, Yixin Ji, Shile Qi, and Daoqiang Zhang, *Senior Member, IEEE*

***Abstract*— Functional magnetic resonance imaging (fMRI) is a noninvasive neuroimaging technique capable of recording dynamic brain activity, which can be used to construct a dynamic functional brain network (DFBN). DFBN adeptly captures intricate spatio-temporal characteristics of brain connectivities, and this proficiency is increasingly recognized in the field of brain disease diagnosis. However, the existing DFBN analysis methods predominantly focus on capturing the topological information between brain regions within individual time windows, often neglecting the critical long-interval temporal interactions that extend across these windows. Moreover, most of them extract temporal and spatial features independently, which hinders the effective mining of the complex spatio-temporal topological coupling in DFBNs. To tackle these problems, we propose a long-interval interactive graph convolutions method that collaboratively captures dynamic spatio-temporal topological features from DFBNs. Specifically, we first capture the brain networks' long-interval communication patterns through the Hawkes process, which comprehensively exploits the impact of brain connectivity topological associations across different time windows. Then, we develop an interactive graph convolution framework to collaboratively extract spatio-temporal features in DFBNs, which overcomes the shortcomings of existing spatio-temporal models that struggle to characterize complex dynamic topological coupling. Finally, we design a hypergraph-based spatial information augmentation (SIA) module to further extract the high-order dependencies among brain regions. Numerous experimental results demonstrate that the proposed method is superior to several state-of-the-art methods, and can provide effective biomarkers for brain disease diagnosis.**

***Index Terms*— Brain disease diagnosis, dynamic brain network, graph convolution network, Hawkes process, spatio-temporal features.**

Received 22 June 2024; revised 28 November 2024; accepted 12 January 2025. Date of publication 13 March 2025; date of current version 25 March 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62076129, Grant 62371234, Grant 62136004, Grant 62272226, and Grant 62276130; in part by the Natural Science Foundation of Jiangsu Province under Grant BK20231438; in part by the National Key Research and Development Program of China under Grant 2023YFF1204803; and in part by the Key Research and Development Plan of Jiangsu Province under Grant BE2022842. The Associate Editor coordinating the review process was Dr. Zhen Ren. *(Corresponding author: Qi Zhu.)*

Shengrong Li, Qi Zhu, Donghai Guan, Yixin Ji, Shile Qi, and Daoqiang Zhang are with the College of Computer Science and Technology and the Key Laboratory of Brain-Machine Intelligence Technology, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: zhuqinuaa@163.com).

Bo Shen and Li Zhang are with the Department of Geriatrics, Affiliated Brain Hospital, Nanjing Medical University, Nanjing 210029, China.

Digital Object Identifier 10.1109/TIM.2025.3551032

I. Introduction

DYNAMIC functional brain network (DFBN) is essential for gaining a comprehensive understanding for brain function and the interaction pattern changes that occur in various neurological conditions [1]. Unlike static brain network [2], which only provides an average connection overextended periods, DFBN offers a real-time, multifaceted perspective that includes both the temporal and spatial dynamics of brain connectivity patterns. Recent research emphasizes that neurodegenerative diseases, such as mild cognitive impairment (MCI) [3] and Parkinson's disease (PD) [4], can lead to abnormal dynamic properties of brain networks. However, the interrelationships among topological features within these networks are highly complex, especially due to the tightly coupled spatio-temporal information. Therefore, it remains a challenge to develop effective spatio-temporal topological analysis methods to improve the diagnostic performance for brain diseases.

For the construction of DFBN, it typically involves segmenting the functional magnetic resonance imaging (fMRI) signal into multiple subsequences using sliding windows, and employs the similarity function to measure the interregional connectivity. Once the DFBN is established, feature extraction is performed using either feature projection or deep learning methods to facilitate the diagnostic tasks associated with brain diseases. Recently, numerous methods for extracting features from DFBNs have been proposed. Such as, Liu et al. [5] utilized Pearson's correlation coefficient (PC) to construct brain networks for fMRI signals under each window, and designed dual temporal graph learning to capture the dynamic features of each brain network. Zhang et al. [6] stacked DFBNs as tensor, and analyzed DBFNs by tensor decomposition with orthogonality and partial symmetry constraints. Dong et al. [7] added temporal edges between corresponding nodes of adjacent brain networks to describe the evolution trajectory of functional connectivity. However, these methods either treat all functional brain networks across all windows of DFBNs equally during the feature extraction process, or only consider the associations between adjacent window brain networks, failing to adaptively explore the topological relevance of brain connectivity over long intervals.

The integration of long-interval information is potentially related to the brain's higher cognitive functions, such as decision-making and prospective memory [8], [9].

Neuroscientific research indicates that the brain can integrate signals from different sensory and cognitive systems, which may vary across different timescales [10], [11]. Many neurodegenerative diseases are often accompanied by symptoms of cognitive impairment. Therefore, studying changes in long-interval functional patterns is crucial for early brain disease detection and biomarker exploration. For example, Pavlov's memory test showed that the duration of participants' memory for the learned content reflects the fluctuations in cognitive states [12]. Linkenkaer-Hansen et al. [8] studied the temporal correlation of oscillations in normal human brain networks on timescales ranging from seconds to minutes, and found that the correlation decays as a power-law function. However, previous studies only focus on the correlation of brain activity over different time periods, neglecting the exploration of topological connection patterns over long intervals. Although DFBNs contain information about brain topological connections at different times, there are still challenges in modeling topological connection patterns over long intervals from a spatio-temporal analysis perspective. Therefore, there is an urgent need for research on mining long-interval topological connection patterns based on DFBN. This is not only beneficial for the early detection of brain disorders, but also conducive to identifying biomarkers associated with brain diseases that lead to the decline of cognitive and decision-making abilities.

For the extraction of spatio-temporal characteristics from DBFNs, current approaches are predominantly categorized into two types: nongraph-based methods and graph-based methods. The nongraph-based methods convert the brain network at each discrete window into vector form, and, respectively, extract shallow the temporal and spatial information from them. For example, Jie et al. [3] calculated the temporal and spatial variability of DFBN separately and integrated them to diagnose MCI. Ji et al. [13] first adopted nonparametric state estimation to automatically infer the number of brain states and transition times, and then used a generative adversarial network to learn the structure of the DFBN. Graph-based DFBN analysis methods usually utilize graph convolutional networks (GCNs) to extract topological structure features, and employ temporal convolutional networks (TCNs) to capture the dynamic temporal features. For instance, Xu et al. [14] first adopted a multichannel GCN to capture the functional and structural information of DFBNs, and then used a stacked long short-term memory network to extract the temporal information between brain regions. Kong et al. [15] employed the graph attention neural network and temporal fusion module to model spatial and temporal correlation from DFBN, respectively. It is worth noting that the existing spatio-temporal graph analysis models tend to extract spatial and temporal characteristics independently, even though these two types of topological features are highly intertwined.

To tackle the aforementioned issues, we propose a long-interval graph convolution method to cooperatively extract spatio-temporal topological features in DFBNs. Specifically, we first use the Hawkes process to capture the temporal information transport between brain regions across different windows, and comprehensively mine the influence of long-interval cognitive activities relationship from brain networks perspective. Then, we design an interactive graph convolution framework to jointly extract the spatio-temporal features in DFBNs, overcoming the shortcomings of traditional DFBN analysis methods that independently extract temporal and spatial information. Finally, we propose a hypergraph-based spatial information augmentation (SIA) module to further capture the complex spatio-temporal topological dependencies in DFBNs.

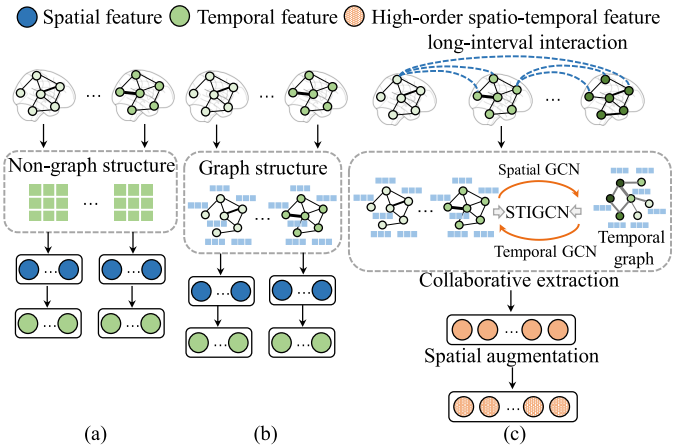

Fig. 1. Comparison of conventional DFBNs analysis strategies. (a) Independently extracting shallow temporal and spatial features with nongraph methods. (b) Independently extracting topological features and temporal information with graph methods. (c) Proposed model can capture the long-interval temporal interactions between brain regions, and collaboratively extract the spatio-temporal graph topological features from DFBNs.

For comparison purposes, we summarize the differences between the existing DFBNs analysis methods and the proposed method in Fig. 1. Specifically, Fig. 1(a) illustrates the nongraph-based analysis methods, which ignore the valuable topological structure information in DFBNs. Fig. 1(b) shows the graph-based analysis method, which extracts the topological structure and temporal information flows in DFBNs independently, ignoring the property that the spatio-temporal features in brain networks are highly coupled. Different from the above methods, as depicted in Fig. 1(c), the proposed method not only captures the long-interval communication patterns between brain regions across different windows, but also cooperatively extracts the complex high-order spatio-temporal topological characteristics in DFBNs. In brief, the proposed method has the following contributions.

1) In this article, we propose a long-interval DFBN topological analysis method by using the Hawkes process, which comprehensively captures the brain connectivity topological associations across different time windows.
2) We design an interactive graph convolution framework that performs cooperatively with spatial GCN and temporal GCN. Spatial GCN aggregates the topological information of DFBNs, while temporal GCN captures long-interval temporal interactions.
3) A hypergraph-based SIA module is devised to further capture complex high-order topological information from the extracted features of DFBN.

4) Experimental results on ADNI and PD datasets show that the proposed method outperforms various competing baselines, and can provide reliable biomarkers for brain disease diagnosis.

The rest of the article is organized as follows. In Section II, we introduce the related work. Then, Section III presents the proposed method in detail. We describe the ADNI and PD datasets used in our work, and report the experimental results in Section IV. Finally, we summarize this work and draw a conclusion in Section V.

II. Related Work

A. Construction of DFBNs

DFBNs can depict the evolutionary process of brain topology, and hold significant potential in brain disease diagnosis [3], emotion recognition [16], and so on. In general, DFBNs consist of a series of discrete functional networks that capture the dynamic changes of the brain. For example, Lin et al. [17] and Wang et al. [18] first divided the fMRI signal into multiple intervals, and then employed statistical models to estimate the functional connectivity within each interval. With the development of brain research, people realized that the brain is a continuously changing system [19], and a series of discrete networks may not effectively represent the spatio-temporal interactions of the brain. Therefore, several methods for constructing and analyzing DFBNs that interact across time have been developed. Such as, Zhu et al. [20] introduced a stacked DFBN representation method, leveraging temporal variability to quantify the information interaction between brain regions across time domains. Tang et al. [21] segmented fMRI into subsegments and learned asynchronous functional interactions of brain signals.

However, these methods ignore the long-interval and attenuation properties of information interaction between brain regions. Indeed, each brain region is directly or indirectly connected to other brain regions, and the connection strength is related to the proximity in the temporal dimension [22]. Therefore, we propose to use the Hawkes process to model long-interval interactions between brain regions across windows, which may potentially allow us to investigate the effects of long-interval correlation on cognitive activities from a brain network perspective.

B. Spatio-Temporal Graph Convolution Methods

The brain is an evolving nonstationary system, and the dynamic changes of topology can reflect the abnormal connectivity patterns caused by brain diseases. Capturing the spatio-temporal topological features from DFBN is of great significance for improving the diagnostic performance of brain diseases. Existing spatio-temporal graph convolution methods can be divided into two strategies. One of them is to integrate GCN and TCN into a network to extract spatial and temporal information in separate stages

$$H = \mathrm{TCN}(\mathrm{GCN}(A, X, W) \tag{1}$$

where A and X denote the adjacency matrix and node features, respectively, and W is the learnable parameters. For example, Gadgil et al. [23] proposed to construct DFBNs to model the nonstationary properties of the brain, and used GCN and 1-D convolutional neural networks (CNNs) to extract the topology and temporal information flow of DFBNs. Kim et al. [1] first introduced graph isomorphic networks to obtain graph representations of DFBNs, and then designed a spatio-temporal attention mechanism to automatically focus on valuable dynamic topological features. Another spatio-temporal model is the evolutionary graph convolution, which treats the parameters in the GCN as the hidden state, and uses the TCN to update this hidden state

$$W_1 = \mathrm{TCN}(H, W_0), \quad H = \mathrm{GCN}(A, X, W_1) \tag{2}$$

where W_0 and W_1 represent the learnable parameters, H denotes the feature embedding, and A and X are the adjacency matrix and node features, respectively. For example, Dong et al. [7] first employed GCN to obtain the neighboring information, and then used the gated recurrent unit to update the model parameters in GCN, thus capturing the spatio-temporal embedding of functional connectivity.

However, these methods extract temporal and spatial information independently, which may not effectively capture the complex spatio-temporal correlations from DFBNs. In this article, we develop the spatio-temporal interaction GCNs to collaboratively extract dynamic topological features from DFBN.

III. Proposed Method

In this article, we propose an interactive cross-window DFBN analysis method based on the Hawkes process, and develop a spatio-temporal interactive graph convolution network (STIGCN) to cooperatively extract spatio-temporal topological features. Moreover, we introduce a hypergraph-based SIA module to further extract deeper high-order structural features in the brain. The framework of the proposed model is illustrated in Fig. 2.

A. Cross-Window Brain Network Construction

There is continuous communication through neural signals between brain regions with time scales ranging from seconds to minutes, suggesting long-interval temporal interactions between brain regions. In order to comprehensively consider the long-interval temporal brain connectivity, we propose a cross-window brain network incorporating temporal interaction based on the Hawkes process. The cross-window brain network we constructed is shown in Fig. 3.

Specifically, to capture the dynamic evolution process of the brain, we use T nonoverlapping sliding window to divide the fMRI signal into subsignals of length S. Next, we adopt the Pearson's correlation coefficient to calculate the functional brain network under each window. Therefore, the DFBNs can be represented as $G = (V, A, F)$, where V is a set of N nodes in the brain, $A = (A^1, A^2, \ldots, A^{\mathrm{T}}) \in \mathbb{R}^{T\times N\times N}$ and $F = (F^1, F^2, \ldots, F^{\mathrm{T}}) \in R^{T\times N\times S}$ denote the functional brain network and subsignals under all windows, respectively. Naturally, T sliding window yield total of $T * N$ nodes $\{V^t\}_{t\in[1:T]}$. For the information interaction between brain regions across

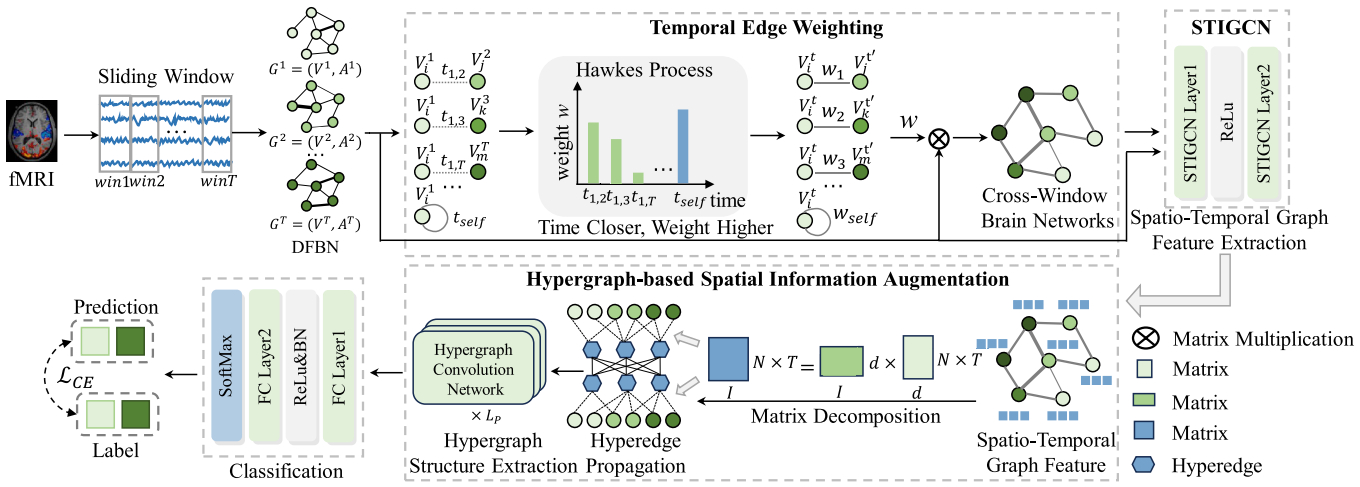

Fig. 2. Framework of the proposed method for brain disease diagnosis.

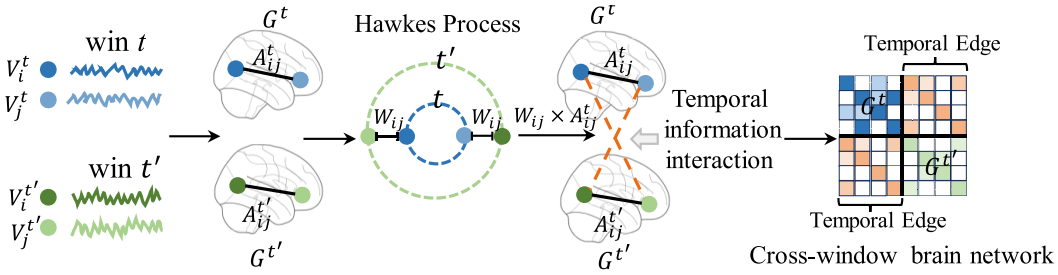

Fig. 3. Schematic of the temporal information interaction between brain regions across windows. The dark and light shades of the same color are used to distinguish the temporal signals of different brain regions within the same window.

windows, we use the Hawkes process [22], [24] to discuss the weighting problem of temporal edges. Hawkes process proved to be beneficial for temporal graph modeling, because it takes into account the effect of temporal changes on the correlation of two brain regions. For the node V_i^t, if its neighboring nodes are closer to V_i^t in time, they tend to have more strong connection. In this way, when a node receives messages from neighboring nodes across windows, these messages should be weighted with temporal information

$$W_{ij} = \exp\left(-\frac{|t'-t|}{T}\right), \quad V_j^t \in N_{V_i^{t'}} \tag{3}$$

where $N_{V_i^{t'}}$ is the neighborhood of node V_i^t, and t' and t denote different windows. This essentially weights the temporal information propagation between brain regions, so that neighboring nodes with smaller time intervals can receive larger weights. In formulation, the adjacency matrix $\hat{A} \in R^{(T*N)\times(T*N)}$ of the cross-window brain network is derived as follows:

$$\hat{A}\left(v_i^t, v_j^{t'}\right) = \begin{cases} A_{ij}^t, & t = t' \\ A_{ij}^t W_{ij}, & \text{otherwise.} \end{cases} \tag{4}$$

Therefore, the cross-window brain network not only contains correlations between brain regions within each time window, but also captures complex interactions between brain regions across different time windows.

B. Spatio-Temporal Interactive Graph Convolution Network

After obtaining the cross-window brain network with valuable dynamic topological information, we develop a novel STIGCN to collaboratively capture the dynamic topology and temporal dependencies in brain networks. The detailed process of STIGCN is shown in Fig. 4. Specifically, for each layer of spatio-temporal interaction graph convolution, we first utilize spatial graph convolution to learn the topology in DFBN. Then, we use the output of the spatial GCN as the node features, and employ the cross-window brain network as the adjacency matrix for performing temporal graph convolution. This step allows the model to capture long-interval interactions between brain regions and effectively fuse spatio-temporal information. In addition, the learned spatio-temporal information serves as input for the next layer of STIGCN. Therefore, the proposed method extracts information in a spatio-temporal interactive manner, allowing it to comprehensively capture the complex dynamic topology of the brain. We take the lth layer of STIGCN as an example

$$H^{(l)} \xrightarrow{f_{\text{spa}}} H_{\text{spa}}^{(l)} \xrightarrow{f_{\text{tem}}} H^{(l+1)} \tag{5}$$

where $H^{(l)}$ is the hidden feature of the lth layer in STIGCN, and $H_{\text{spa}}^{(l)}$ is the output feature after performing spatial graph convolution. f_{spa} and f_{tem} represent the spatial GCN and temporal GCN, respectively. Next, we will introduce two kinds of graph convolution mechanisms in detail.

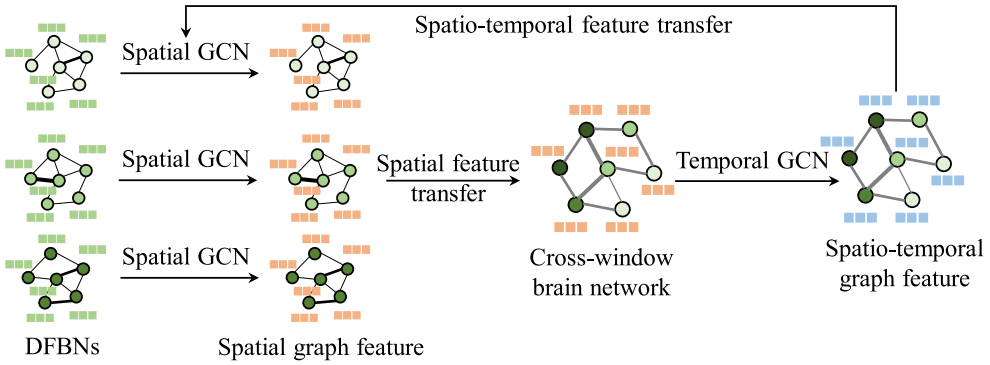

Fig. 4. Framework diagram of STIGCN. The spatial GCN learns the topology from DFBNs, and the temporal GCN uses the output of the spatial GCN as an input feature to learn spatio-temporal information in cross-window brain networks. Moreover, the output of the temporal GCN is used as the input of the next layer in STIGCN.

1) Spatial GCN: GCNs aggregate structural features by propagating information among neighboring nodes. Motivated by this message-passing mechanism, we perform graph convolution for each subnetwork to extract dynamic topological features from DFBNs. The feature of the tth subnetwork in the lth layer is updated by spatial graph convolutional f_{spa} as follows:

$$H_{\text{spa}}^{(l)}(t,:,:) = \sigma\Big(D(t,:,:)^{-\frac{1}{2}}A(t,:,:)D(t,:,:)^{-\frac{1}{2}} \times H^{(l)}(t,:,:)W_{\text{spa}}^{(l,t)}\Big) \tag{6}$$

where $H_{\text{spa}}^{(l+1)}(t,:,:) \in \mathbb{R}^{T\times N\times d}$ is the output of spatial graph convolution, σ denotes the ReLu activation function, $D(t,:,:)$ represents the degree matrix of $A(t,:,:)$, and $W_{\text{spa}}^{(l,t)}$ is the weight matrix of the tth dynamic subnetwork at the lth spatial graph convolution layer. Through the spatial graph convolution operation, we can effectively capture the evolution process of brain topology.

2) Temporal GCN: To effectively characterize the temporal information propagation of brain regions under different windows, we further perform the temporal GCN to aggregate the dynamic correlation across windows. We first transform the output dimension of $H_{\text{spa}}^{(l)}$ into $(T * N) \times d$ to align with the dimensions of cross-window brain network $\hat{A}$. Considering that the temporal edge has been inserted into the cross-window brain network, we use standard graph convolution to aggregate valuable temporal information in the cross-window brain network $\hat{A}$

$$H^{(l+1)} = \sigma\Big(D^{-\frac{1}{2}}\hat{A}D^{-\frac{1}{2}}H_{\text{spa}}^{(l)}W_{\text{tem}}^{(l)}\Big) \tag{7}$$

where $H^{(l+1)} \in \mathbb{R}^{(T*N)\times d}$ is the output of temporal graph convolution, also the input feature of the $(l+1)$th in STIGCN, D denotes the degree matrix of $\hat{A}$, and $W_{\text{tem}}^{(l)}$ represents the weight matrix of the at the lth temporal graph convolutional layer. It is worth noting that STIGCN is quite different from the traditional graph convolution framework. Traditional graph convolutions focus on extracting temporal and spatial information independently, which may not effectively mine the latent spatio-temporal patterns in brain networks. Our proposed STIGCN enables two convolutional operations to be carried out simultaneously in an information-interactive manner, so as to collaboratively extract the inherent spatio-temporal dependencies of brain networks.

C. Hypergraph-Based SIA

The brain is a dynamic high-order network, and there are spatio-temporal interactions among multiple brain regions. Therefore, we need to mine the nonpairwise structural relationships of brain networks based on spatio-temporal embeddings. Previous methods often capture the complex correlation between brain regions by deepening the number of model layers, which may fail to effectively capture the nonpairwise relationship and is computationally expensive. Therefore, we devise a hypergraph-based SIA module to further extract more complex high-order features. In addition, to reduce the number of parameters and alleviate overfitting, we use matrix factorization to construct a structure matrix. In detail, the incidence matrix of the hypergraph is denoted $\Lambda \in R^{(N*T)\times I}$, where I represents the number of the hyperedge. We decompose the incidence matrix into two low-rank matrices using the extracted spatio-temporal features $H \in R^{(N*T)\times d}$ as follows:

$$\Lambda = HW \tag{8}$$

where $W \in R^{d\times I}$ is the learnable weight matrix, and d denotes the dimension of spatio-temporal feature matrix H, which is a fixed hyperparameter. Thus, learning the incidence matrix merely introduces $\mathcal{O}(I \times d)$ parameters ($d \ll N * T$), which can considerably improve the model efficiency. In addition, low-rank decomposition can alleviate redundant information and noise in the brain [25], [26], promoting the model to focus more on the spatio-temporal dependencies between nonpaired brain regions, thereby effectively capturing the high-order spatio-temporal features in the brain network.

Next, we introduce the hypergraph convolution paradigm to learn hypergraphs with abundant spatio-temporal features. For each layer of hypergraph convolution, we first generate each hyperedge embedding by aggregating the information of all connected nodes. Then, the hypergraph embedding is used to update the node representation, so as to mine the high-order correlations in brain networks. Specifically, the hyperedge embedding matrix $E \in R^{I\times d}$ is derived from the spatio-temporal features H and

the incidence matrix Λ

$$E = \sigma(U\Lambda^{\mathrm{T}}H) + \Lambda^{\mathrm{T}}H \tag{9}$$

where σ denotes the ReLu activation function, and we additionally introduce a learning matrix $U \in R^{I\times I}$ to simulate the implicit dependency among hyperedges. Next, these hyperedge embeddings are aggregated to update the node embedding

$$Z = \Lambda E = \Lambda(\sigma(U\Lambda^{\mathrm{T}}H) + \Lambda^{\mathrm{T}}H). \tag{10}$$

In this way, we can learn nonpairwise spatio-temporal correlation by stacking L_p hypergraph convolution layers, which outputs the node embedding Z.

Finally, we feed Z with complex spatio-temporal features into the MLP, and adopt cross entropy (CE) loss to supervise spatio-temporal features learning by

$$L_{\mathrm{CE}} = -\frac{1}{S}\sum_{s=1}^{S}(y_s \log(\hat{y}_s)\text{-}(1-y_s)\log(1-\hat{y}_s)) \tag{11}$$

where S denotes the number of subjects, $\hat{y}_s$ represents the prediction result of the sth subjects, and y_s is the ground-truth of the sth subjects.

IV. Experiments

A. Materials

1) Data Acquisition: In this article, we conduct experiments on the public Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (http://adni.loni.usc.edu/) and the collected PD dataset. The two datasets are described as follows. ADNI dataset: We downloaded fMRI data for 203 samples from ADNI2 and ADNI3 datasets. The ADNI dataset contains 73 healthy control (HC), 70 significant memory concern (SMC) patients, and 60 early MCI (EMCI) patients. PD dataset: the fMRI data was collected from the Affiliated Brain Hospital of Nanjing Medical University. The dataset contains 162 samples, including 54 HC, 44 Tremor dominant PD (TDPD) patients, and 64 PD with postural instability and gait disorder (PIGD-PD) patients. The equipment for collecting data is the Simens MAGNETOM Verio 3.0 T MRI scanner. The scan parameters of fMRI are as follows: repetition time = 2000 ms, echo time = 25 ms, and flip angle = 90°.

2) Data Preprocessing: We applied the same preprocessing to the ADNI and PD datasets. Specifically, we used the DPARSF toolbox for all fMRI data preprocessing. To correct and realign the initial images, the serialized data was split into several pieces and then adjusted to the EPI template. We utilized detrending to reduce the influence of head motion and the interference from CSF and white matter. Then, we used the automated anatomical labeling (AAL) atlas to divide the fMRI of the ADNI dataset into 90 ROIs with 197 time points. Since the cerebellum plays an important role in the diagnosis of PD, we used AAL to divide the fMRI of the PD dataset into 116 ROIs with 220-time points.

B. Comparison Methods

To verify the effectiveness of the proposed method, we compare it with ten existing brain network analysis methods. Comparison methods can be divided into three categories: temporal feature-based methods, spatial feature-based methods, and spatio-temporal feature-based methods. Temporal feature-based methods include: recurrent neural networks (RNNs) [27], long short-term memory-networks (LSTM) [28], and CNNs [29]. Spatial feature-based methods include: GCN [30], brain network transformer (BNT) [31], and G-Unet [32]. Spatio-temporal feature-based methods are: asynchronous common and individual functional brain network (ACIFBN) [21], spatio-temporal graph convolution network (STGCN) [23], spatio-temporal attention graph isomorphism network (STAGIN) [1], and spatio-temporal graph neural network with dynamic functional and effective connectivity fusion (FE-STGNN) [33].

C. Experimental Settings

The proposed model is implemented using Pytorch, and trained on an NVIDIA GeForce RTX3080 GPU. For optimization, we use the Adam optimizer with a learning rate of $7e^{-3}$ and weight decay is $1e^{-3}$. To mitigate overfitting, we introduce a gradient clipping strategy and set dropout to 0.6. In the experiment, we adopt five nonoverlapping sliding windows to construct DFBNs, and employ one layer of STIGCN to collaboratively extract spatio-temporal features. In the SIA module, we use grid search to determine the number of hyperedges I. Specifically, the number of hyperedges I is located in {4, 8, 16, 32, 64, 128, 256}. After obtaining the high-order spatio-temporal features, we first pull the features into 1-D vectors, and then use a multilayer perceptron to predict the class of each individual. Here, the number of neurons in the two fully connected layers is 90 and 32, respectively.

For the ADNI dataset, we perform four binary classification tasks: HC versus ILL, HC versus SMC, HC versus EMCI, and SMC versus EMCI. For the PD dataset, we also perform four binary classification tasks: HC versus ILL, HC versus TDPD, HC versus PIGD-PD, and TDPD versus PIGD-PD. In the experiments, we use ten-fold cross validation to evaluate the performance of all models. Specifically, the entire dataset is divided into ten approximately equal and mutually exclusive subsets, with nine used for training and the remaining one for testing. To optimize model performance and avoid overfitting, we introduce an early stopping strategy during training. When the training loss does not decrease for 30 consecutive epochs, the model parameters at this time are saved for testing. The experiment is repeated ten times, and the mean and standard deviation (std) of the test results are reported. Experimental performance is measured by accuracy (ACC), sensitivity (SEN), specificity (SPE), and the area under the receiver operation characteristic curve (AUC). To ensure the fairness, we apply the same data processing and experimental settings to all methods.

V. Results

A. Classification Performance

The diagnostic performance of all methods on the ADNI dataset and PD dataset is presented in Tables I and II, respectively. In addition, we perform a t-test on ACC

TABLE I

CLASSIFICATION PERFORMANCE (MEAN/STD) OF THE PROPOSED METHOD AND COMPARISON METHODS ON THE ADNI DATASET (%)

Method	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
	HC vs. ILL				HC vs. SMC			
RNN [27]	83.36/07.42**	90.91/09.48	67.53/13.34	85.25/05.02**	81.05/11.61**	70.08/19.10	90.42/13.15	82.25/12.69**
LSTM [28]	84.79/05.91**	86.27/14.00	72.80/28.95	85.95/05.74**	85.29/08.68**	78.19/22.70	90.11/11.33	85.78/08.60**
CNN [29]	84.69/04.24**	92.13/11.44	70.21/14.00	87.95/06.01**	86.05/09.42*	83.06/13.06	87.32/24.44	88.03/12.54**
GCN [30]	82.76/08.44**	79.50/11.66	87.70/11.71	85.66/10.20**	84.62/07.64**	84.36/23.48	81.38/16.41	86.16/09.72**
BNT [31]	86.00/08.89*	88.26/11.44	80.18/17.74	84.02/10.59**	87.48/03.96*	82.75/14.08	90.15/08.91	87.82/07.74**
G-Unet [32]	84.50/07.23**	79.67/10.72	89.81/15.78	86.37/07.33**	84.67/09.85**	76.27/18.26	92.60/09.65	84.37/16.75**
ACIFBN [21]	89.59/06.55	88.79/08.16	91.07/09.63	91.00/06.88*	88.86/05.48	90.30/10.67	87.66/09.27	91.29/04.61**
ST-GCN [23]	87.17/04.52**	86.06/12.35	86.58/13.40	89.29/05.63**	90.14/05.67	90.06/11.87	88.84/14.19	89.15/10.62**
STAGIN [1]	89.69/05.85	89.72/09.20	88.49/10.62	90.20/06.14*	90.83/05.83	93.48/10.34	84.56/16.71	93.09/05.99**
FE-STGNN [33]	89.10/04.92	89.32/07.57	86.60/10.60	90.36/06.28*	90.86/06.95	89.38/10.92	89.90/12.68	90.31/10.65 *
Ours	**92.14/05.48**	**91.20/07.11**	**93.05/13.16**	**94.93/04.95**	**93.57/08.72**	**95.00/08.03**	**93.50/10.01**	**95.40/08.19**
	HC vs. EMCI				SMC vs. EMCI			
RNN [27]	81.98/06.91**	70.30/14.54	87.94/15.70	75.72/10.94**	73.08/09.88**	73.64/22.84	66.36/19.90	70.51/13.63**
LSTM [28]	83.41/10.03**	73.04/16.85	93.06/11.74	83.19/10.82**	76.15/06.39**	84.52/14.53	65.59/16.06	71.44/11.88**
CNN [29]	84.34/04.77**	70.63/17.84	92.38/07.66	79.52/18.88**	77.69/07.26**	71.63/21.30	78.28/17.86	70.47/14.98**
GCN [30]	85.05/08.04*	74.55/18.94	91.53/11.49	84.82/11.20**	75.38/08.97**	76.08/20.95	65.89/26.06	65.43/10.79**
BNT [31]	87.50/06.72	84.19/12.46	88.31/11.06	86.72/10.77*	76.92/05.96**	87.62/12.92	60.18/17.51	70.92/16.50**
G-Unet [32]	85.82/08.23*	78.71/09.58	91.25/11.85	88.08/08.11*	76.15/10.00**	86.00/09.93	59.21/23.19	72.13/15.84**
ACIFBN [21]	88.24/11.10	80.29/15.15	92.80/13.24	86.83/13.53*	80.77/12.99*	88.89/10.83	74.42/19.69	80.81/10.52**
ST-GCN [23]	87.47/12.09	79.52/14.87	92.90/12.56	85.68/14.04	81.54/07.84**	81.52/20.09	73.81/18.79	72.64/12.57**
STAGIN [1]	90.99/05.68	84.31/19.46	90.90/15.02	88.02/12.78	82.31/07.73**	**89.59/13.18**	72.12/09.02	80.00/14.03
FE-STGNN [33]	88.74/06.78	84.12/15.56	93.29/13.00	87.72/12.28*	83.85/10.00	87.16/10.82	78.14/14.55	83.13/10.81
Ours	**91.81/07.65**	**88.08/10.31**	**95.25/08.10**	**95.08/05.52**	**88.76/04.91**	89.39/10.18	**86.51/13.23**	**87.83/07.28**

ACC means accuracy, SEN means sensitivity, SPE means specificity, and AUC means the area under the receiver operation characteristic curve.
p-value between the comparison methods and the proposed method: * indicating $p \leq 0.15$, ** indicating $p \leq 0.05$.

TABLE II

CLASSIFICATION PERFORMANCE (MEAN/STD) OF THE PROPOSED METHOD AND COMPARISON METHODS ON THE PD DATASET (%)

Method	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
	HC vs. ILL				HC vs.TDPD			
RNN [27]	71.88/06.04**	75.80/16.56	52.57/25.52	58.03/14.18**	75.75/12.55	72.21/28.27	79.79/20.23	73.93/16.50**
LSTM [28]	73.13/06.87**	72.29/16.59	72.51/18.91	71.29/12.83**	73.50/10.44**	49.17/23.69	87.50/16.32	71.97/18.12**
CNN [29]	73.75/09.60**	73.75/16.86	73.71/14.74	71.98/16.18**	78.25/10.25**	75.74/19.88	81.69/18.34	76.98/11.00*
GCN [30]	75.29/10.09**	80.38/09.51	59.59/29.66	68.53/16.53**	78.44/09.79**	63.57/25.92	87.24/14.44	77.58/12.22*
BNT [31]	74.33/06.80**	82.14/30.03	52.38/21.14	59.48/23.11**	82.50/06.12*	65.83/27.24	89.17/13.46	74.83/12.12
G-Unet [32]	72.00/10.56**	72.71/27.34	64.08/28.69	65.10/17.55**	80.78/08.07*	73.75/34.66	77.49/22.40	70.38/14.01**
ACIFBN [21]	79.60/05.64*	83.46/13.31	61.40/29.63	71.21/13.65*	82.89/10.93	81.33/21.66	82.21/14.77	79.80/18.10
ST-GCN [23]	79.67/08.53	86.08/20.38	59.74/28.95	80.81/13.60	82.67/09.02	82.33/16.74	79.48/22.67	80.18/13/04
STAGIN [1]	82.50/09.06	87.00/12.17	71.81/27.77	80.86/14.36**	81.89/09.71*	79.00/29.73	75.42/29.96	76.55/21.18
FE-STGNN [33]	81.47/10.09	85.09/19.31	65.37/28.16	70.27/18.07*	81.33/10.67*	82.67/21.12	77.90/19.75	79.91/16.08
Ours	**85.89/11.10**	**89.00/14.76**	**79.74/22.88**	**85.28/14.40**	**88.75/10.38**	**83.33/21.08**	**91.57/15.47**	**85.65/11.80**
	HC vs. PIGD-PD				TDPD vs. PIGD-PD			
RNN [27]	73.50/06.12**	81.83/17.58	52.31/28.49	66.61/12.41**	71.00/09.43**	76.81/28.63	47.31/33.75	62.14/13.11**
LSTM [28]	75.00/06.83**	73.16/16.49	71.13/20.76	70.18/08.07*	73.18/06.24**	80.75/28.46	50.17/28.92	71.20/06.39
CNN [29]	74.83/08.58**	80.99/25.99	59.17/25.92	65.82/18.13**	74.18/10.37**	81.98/28.41	51.67/27.94	56.40/23.61**
GCN [30]	71.50/13.30**	82.32/11.24	55.80/28.68	68.58/16.52**	75.73/09.07**	72.09/17.05	**77.83/19.69**	69.97/19.59
BNT [31]	79.00/09.60*	81.89/23.02	69.52/20.93	78.62/11.08	76.73/09.82*	82.23/15.43	59.67/35.98	70.20/14.43
G-Unet [32]	73.33/11.88	76.31/30.05	62.33/29.17	66.26/18.90*	73.00/09.11**	81.56/30.71	50.83/28.10	59.74/17.30**
ACIFBN [21]	75.75/12.55*	72.21/28.27	79.79/20.23	73.93/16.50*	74.18/14.29*	79.19/28.61	62.17/32.28	63.79/10.20**
ST-GCN [23]	79.69/10.05	69.49/30.31	80.33/21.47	77.15/13.92	74.45/14.45*	82.08/31.18	59.31/34.31	67.44/18.90*
STAGIN [1]	82.20/06.04	81.23/17.65	75.13/23.67	79.30/09.40*	78.00/13.27	61.55/38.60	70.83/38.60	59.85/21.41**
FE-STGNN [33]	81.36/09.69	79.87/24.06	**83.81/22.05**	80.71/15.26	76.09/08.07**	85.08/17.47	50.83/31.73	62.88/12.50**
Ours	**84.77/06.21**	**84.12/14.19**	81.17/17.57	**85.04/07.01**	**83.36/05.42**	**85.33/17.40**	70.83/23.35	**76.62/10.74**

ACC means accuracy, SEN means sensitivity, SPE means specificity, and AUC means the area under the receiver operation characteristic curve.
p-value between the comparison methods and the proposed method: * indicating $p \leq 0.15$, ** indicating $p \leq 0.05$.

and a DeLong test on the receiver operating characteristic curve to evaluate whether the proposed method achieves statistically significant improvements over the comparison methods. Specifically, on the ADNI dataset, our method

TABLE III
CLASSIFICATION RESULTS OF ABLATION EXPERIMENTS ON ADNI AND PD DATASETS (%)

Datasets	Method	LII	STI GCN	SIA	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
					HC vs. ILL				HC vs. SMC				HC vs. EMCI				SMC vs. EMCI			
ADNI	baseline				82.26	81.03	86.06	82.16	84.62	73.38	85.83	83.79	81.92	63.01	91.71	79.39	72.31	85.65	54.19	67.71
	Ous1	✓			86.71	89.83	79.45	90.08	87.62	81.71	92.08	86.72	83.57	75.94	89.40	85.57	76.15	81.11	57.20	65.44
	Ous2		✓		88.74	88.94	87.67	91.18	89.48	87.99	92.33	89.20	85.82	75.07	87.46	81.96	78.46	77.29	70.92	71.00
	Ous3			✓	87.31	88.43	83.56	91.02	88.74	84.13	90.46	84.71	85.00	69.96	95.24	81.02	77.50	68.40	63.75	67.11
	Ous4	✓	✓		89.21	85.67	87.80	93.20	90.29	83.08	92.92	91.84	90.22	88.20	89.46	86.21	80.83	76.07	70.31	70.23
	Ous5	✓		✓	89.14	90.90	80.53	91.76	89.53	90.89	83.90	91.36	86.54	78.50	92.39	82.19	82.50	83.62	75.18	78.76
	Ous6		✓	✓	90.50	89.41	86.38	91.87	91.52	88.72	92.57	91.70	87.03	76.89	91.56	83.15	81.67	83.67	73.56	79.54
	Ours	✓	✓	✓	**92.14**	**91.20**	**93.05**	**94.93**	**93.57**	**95.00**	**93.50**	**95.40**	**91.81**	**88.08**	**95.25**	**95.08**	**88.76**	**89.39**	**86.51**	**87.83**
					HC vs. ILL				HC vs. TDPD				HC vs. PIGD-PD				TDPD vs. PIGD-PD			
PD	baseline				73.82	72.67	68.17	69.53	73.13	74.42	71.05	74.97	72.83	73.55	51.92	56.72	73.91	75.48	55.00	62.61
	Ous1	✓			74.91	81.58	57.67	60.36	76.56	74.17	69.88	75.20	74.16	82.50	54.02	64.21	74.09	79.05	55.00	65.75
	Ous2		✓		77.64	81.14	54.16	64.73	80.44	68.40	85.00	75.10	76.17	74.37	67.12	66.59	76.73	71.37	73.33	67.84
	Ous3			✓	76.18	74.64	62.50	56.06	76.25	60.33	80.90	70.52	73.50	71.22	62.62	65.72	74.91	81.58	57.67	60.36
	Ous4	✓	✓		82.32	88.79	53.81	71.16	83.67	80.74	78.67	81.52	77.95	83.32	70.82	74.51	79.55	73.29	69.17	66.31
	Ous5	✓		✓	78.91	77.80	62.49	67.41	80.50	63.48	90.17	74.43	76.29	71.14	77.21	70.61	77.00	77.08	53.17	58.67
	Ous6		✓	✓	81.51	84.79	58.38	79.30	81.78	68.67	88.23	77.32	77.05	77.76	72.68	68.50	78.82	80.48	64.50	68.08
	Ours	✓	✓	✓	**85.89**	**89.00**	**79.74**	**85.28**	**88.75**	**83.33**	**91.57**	**85.65**	**84.77**	**84.12**	**81.17**	**85.04**	**83.36**	**85.33**	**70.83**	**76.62**

LII means long-interval interaction.
SIA means spatial information augmentation.
ACC means accuracy, SEN means sensitivity, SPE means specificity, and AUC means the area under the receiver operation characteristic curve.

achieves accuracies surpassing suboptimal results by 2.45%, 2.71%, 0.82%, and 4.91%, respectively. On the PD dataset, our method achieves accuracies that outperform suboptimal results by 3.39%, 5.86%, 2.57%, and 5.36%, respectively. Moreover, we observe that the proposed method achieves significant improvements in ACC and AUC compared to RNN, LSTM, CNN, and GCN ($p \leq 0.05$), while also demonstrating consistent improvements over BNT, ACIFBN, ST-GCN, and other methods in most cases ($p \leq 0.15$).

The advantage of our method lies in its ability to capture long-interval temporal interactions between brain regions, and collaboratively extract spatio-temporal topological features from DFBNs. Furthermore, we introduce a hypergraph-based SIA module to model the interaction among nonpaired brain regions, which further mines the complex high-order features in the brain. Compared with the existing brain network analysis methods, such as RNN, LSTM, GCN, and BNT, which focus solely on temporal or spatial features of brain networks, and cannot effectively extract disease-related spatio-temporal features. While ACIFBN, STGCN, STAGIN, and FE-STGNN extract the spatio-temporal features of brain networks in separate stages, and ignore the long-interval interaction between brain regions. In contrast, the proposed method collaboratively captures the long-interval spatio-temporal dependencies from DFBNs, and effectively simulates the high-order interaction among brain regions.

B. Ablation Study

The proposed method contains three main modules: long-interval interaction, STIGCN, and SIA. To further verify the effectiveness of these modules, we conduct multiple ablation experiments on two datasets, and the experimental results are presented in Table III. In Table III, baseline means using only a multilayer perceptron for brain network classification, while ✓indicates the use of the corresponding module. Comparing Ours1, Ours2, and Ours3 with baseline, we find that these modules can effectively improve the diagnostic performance. Furthermore, comparing Ours4, Ours5, and Ours6 with Ours1, Ours2, and Ours3, we observe that combining any two modules leads to better diagnostic performance than using just one module. This suggests that the three modules in the proposed method synergize with each other and collectively enhance effectiveness. We capture long-interval interactions of the brain, effectively exploring dynamic correlations between brain regions across different windows. STIGCN integrates spatial and temporal GCNs to collaboratively extract spatio-temporal features, overcoming the limitations of traditional graph models to extract these features independently. In addition, the SIA module based on hypergraph can further learn complex high-order topological information in brain networks.

VI. DISCUSSION

A. Visualization of Embedded Features

To measure the feature representation ability of the proposed method, we use t-distributed stochastic neighbor embedding (t-SNE) to visualize the features learned by all methods in Fig. 5. From Fig. 5, the distribution of the original data is quite messy. Compared with the original data, although the methods based on temporal or spatial features, such as RNN, CNN, LSTM, GCN, BNT, and G-Unet aggregate the intraclass data to a certain extent, they struggle to distinguish the common features of different classes. Spatio-temporal feature-based methods, such as ACIFBN, STGCN, STAGIN, and FE-STGNN distance the interclass data to some extent, but cannot effectively aggregate the intraclass data. This is because compared with the temporal or spatial features, the spatio-temporal features contain more abundant information related to brain diseases, thus yielding better visualization

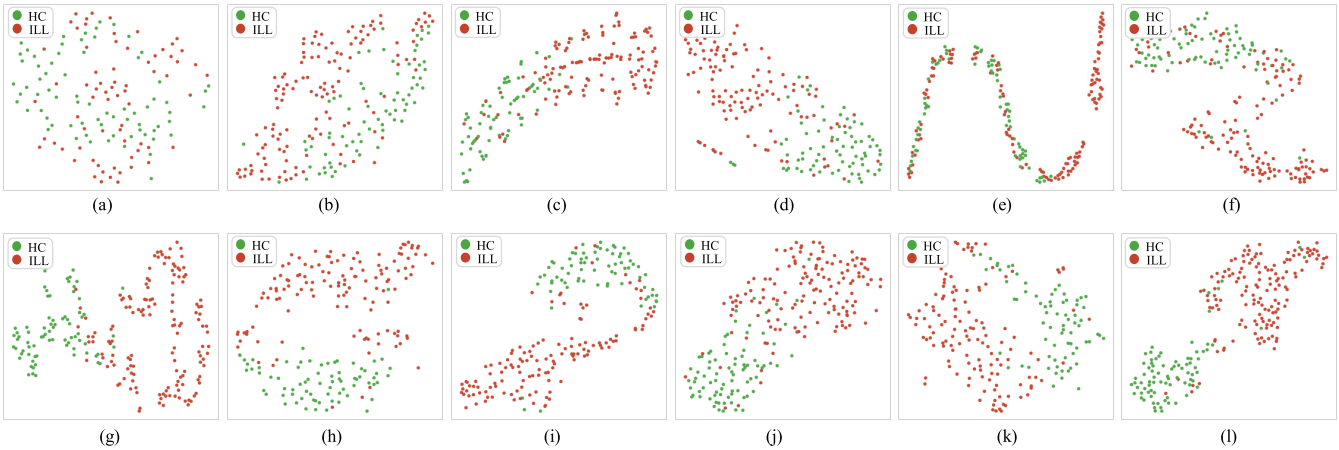

Fig. 5. T-SNE visualization for different methods on the ADNI dataset. (a) Original. (b) RNN. (c) CNN. (d) LSTM. (e) GCN. (f) BNT. (g) G-Unet. (h) ACIFBN. (i) STGCN. (j) STAGIN. (k) FE-STGNN. (l) Ours.

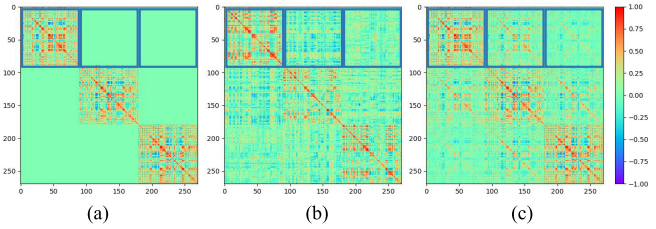

Fig. 6. Dynamic brain networks constructed by different methods. (a) DFBNs. (b) ACIFBN. (c) Ours.

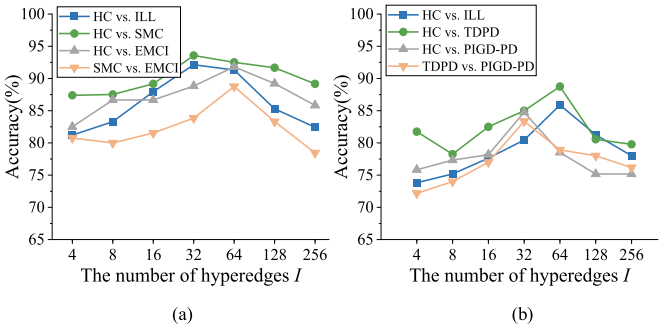

Fig. 7. Classification accuracy under different hyperedges on (a) ADNI and (b) PD datasets.

results. This also proves the importance of extracting spatio-temporal features. Notably, the embedded features learned by the proposed method have clear boundaries between different classes, because our method captures the long-interval interactions between brain regions across windows, and cooperatively captures the complex spatio-temporal dependencies in brain networks.

B. Visualization of Dynamic Brain Networks

To visually demonstrate the ability of DFBN analysis methods to describe temporal interactions between brain regions, we visualize the networks learned by our method and the other two dynamic brain networks in Fig. 6. As depicted in Fig. 6, DFBNs directly constructed by sliding window correlation only focus on the topological changes within each window, and ignore the interactions across windows. Although ACIFBN can extract asynchronous interactions between brain regions, it ignores the influence of temporal proximity on cross-window brain region interactions, resulting in an inconspicuous evolution process. Different from the above two methods, it can be seen from the circled blue box that the brain network constructed by the proposed method has an obvious evolution trajectory. This is because the cross-window brain network constructed based on Hawkes process considers the influence of temporal distance on interaction, and effectively captures long-interval interactions.

C. Analysis of Parameter Sensitivity

In the experiments, we determine the number I of hyperedges by grid search. Specifically, we vary the number of hyperedges I in the range of {4, 8, 16, 32, 64, 128, 256}, and discuss the effect of different numbers of hyperedges on classification accuracy, as shown in Fig. 7. From Fig. 7, we observe that the classification outcomes are superior on two datasets with the utilization of 32 or 64 hyperedges, outperforming those achieved with other values. The possible reason for this phenomenon is that a small number of hyperedges only capture the coarse features in the brain network, which hinders the accuracy of disease diagnosis. Conversely, too many hyperedges may introduce extra connections and redundant information. Therefore, an intermediate number of hyperedges is advantageous for extracting high-order information related to brain diseases.

D. Computational Efficiency Analysis

To evaluate the computational efficiency of models, we calculate the number of model parameters (#Param) and the number of floating-point operations (FLOPs) of the proposed method and several spatio-temporal methods, such as ACIFBN, STGCN, STAGIN, and FE-STGNN, shown in Table IV. As can be seen from Table IV, the proposed method achieves the lowest #Param: 0.81 M, and the lowest FLOPs: 0.81 M. In particular, compared with ACIFBN, STGCN, STAGIN, and FE-STGNN, our method significantly reduces FLOPs. Therefore, our method achieves superior experimental

TABLE IV

COMPUTATIONAL EFFICIENCY COMPARISON BETWEEN THE PROPOSED METHOD AND OTHER SPATIO-TEMPORAL ANALYSIS METHODS

Method	#Param(M)	FLOPs(M)
ACIFBN	1.18	14.24
ST-GCN	2.40	13.30
STAGIN	1.06	96.50
FE-STGNN	1.22	1.26
Ours	**0.81**	**0.81**

#Param means the number of model parameters.
FLOPs means the number of floating-point operations.

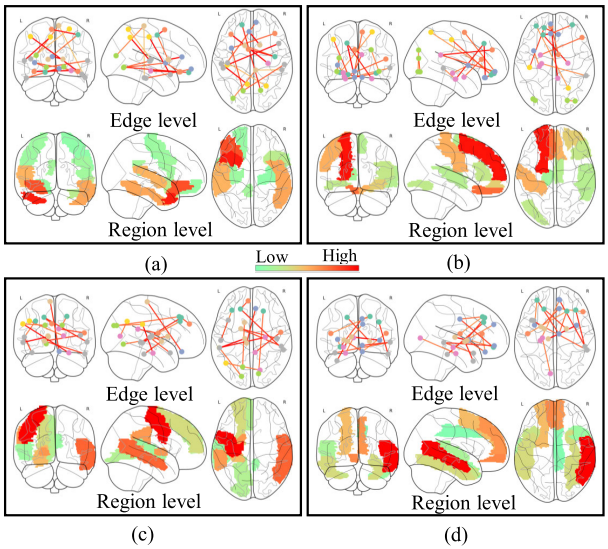

Fig. 8. Top 15 most important brain connectivities and the top 10 most key brain regions on different diagnosis tasks. (a) HC versus ILL. (b) HC versus SMC. (c) HC versus EMCI. (d) SMC versus EMCI.

performance with the least computational resources, which further verifies the effectiveness of the proposed method.

E. Discriminative Brain Connectivity and Regions

We investigate the effectiveness of the proposed method in identifying discriminative brain connectivity and regions. Specifically, we use the nonnegative elastic network [34] to find the most important 15 brain connections and 10 brain regions, and the experimental results are shown in Fig. 8. For HC versus ILL, the important brain connections and brain regions are mainly concentrated in Temporal.Pole.Mid.L, Temporal.Sup.L, Precentral and Frontal.Inf.Oper.R, which are responsible for cognitive function and closely related to cognitive impairment. For HC versus SMC, the discriminative brain regions are concentrated in the Frontal.sup.L, Gyrus rectus, Frontal.mid.orb.R, Hippocampus and other regions. These regions control the memory and emotions of the brain, so the abnormality of these regions may lead to SMC. For NC versus EMCI, key brain regions include the Middle temporal gyrus, Superior frontal gyrus, Hippocampus, and Temporal.sup.L. These regions are responsible for processing visual and memory information in the brain. Therefore, these regions are crucial for distinguishing HC from EMCI patients. For SMC versus EMCI, the discriminative brain regions are the Middle temporal gyrus, Superior frontal gyrus, and Cingulum.Mid.R. We find that the identified biomarkers corresponded to relevant regions with functional deficits in SMC and EMCI patients [35], [36], [37], [38]. Therefore, our method can provide effective biomarkers for the diagnosis of cognitive impairment.

VII. CONCLUSION

In this article, we propose a brain disease diagnosis method for collaboratively extracting long-interval spatio-temporal features from DFBNs. The proposed method not only effectively simulates the long-interval and attenuation properties of brain connectivities, but also can collaboratively extract the spatio-temporal dependencies from brain networks. Specifically, we utilize the Hawkes process to capture communication patterns across window brain regions, potentially allowing us to investigate the effects of long-interval correlation on cognitive activities from a brain network perspective. Second, we design a novel graph interaction model to collaboratively extract spatio-temporal topological features in brain networks, which overcomes the shortcomings of traditional graph methods that are difficult to effectively aggregate complex dynamic topological structures. In addition, we introduce a hypergraph-based SIA module to further extract deeper high-order structural information in brain networks. Extensive experiments show that the proposed method not only helps medical professionals explore the association between long-interval interactions and cognitive function impairment, but also assists clinicians in diagnosing and provides reliable biomarkers. In future work, we plan to integrate structural and DFBNs to investigate the impact of brain diseases on function-structure spatio-temporal coupling.

REFERENCES

[1] B.-H. Kim, J. C. Ye, and J.-J. Kim, "Learning dynamic graph representation of brain connectome with spatio-temporal attention," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 4314–4327.
[2] S. Pei, C. Wang, S. Cao, and Z. Lv, "Data augmentation for fMRI-based functional connectivity and its application to cross-site ADHD classification," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–15, 2023.
[3] B. Jie, M. Liu, and D. Shen, "Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease," *Med. Image Anal.*, vol. 47, pp. 81–94, Jul. 2018.
[4] M. Filippi et al., "Longitudinal brain connectivity changes and clinical evolution in parkinson's disease," *Mol. Psychiatry*, vol. 26, no. 9, pp. 5429–5440, 2021.
[5] L. Liu et al., "BrainTGL: A dynamic graph representation learning model for brain network analysis," *Comput. Biol. Med.*, vol. 153, Feb. 2023, Art. no. 106521.
[6] Q. Zhang, B. Guo, W. Kong, X. Xi, Y. Zhou, and F. Gao, "Tensor-based dynamic brain functional network for motor imagery classification," *Biomed. Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102940.
[7] Z. Dong, Y. Wu, X. Yu, J. S. X. Chong, Y. Jin, and J. Zhou, "Beyond the snapshot: Brain tokenized graph transformer for longitudinal brain functional connectome embedding," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Cham, Swizerland. Springer, Jan. 2023, pp. 348–357.
[8] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, "Long-range temporal correlations and scaling behavior in human brain oscillations," *J. Neurosci.*, vol. 21, no. 4, pp. 1370–1377, Feb. 2001.
[9] Y. Xu, X. Long, J. Feng, and P. Gong, "Interacting spiral wave patterns underlie complex brain dynamics and are related to cognitive processing," *Nature Hum. Behav.*, vol. 7, no. 7, pp. 1196–1215, Jun. 2023.

[10] R. Liégeois et al., "Resting brain dynamics at different timescales capture distinct aspects of human behavior," *Nature Commun.*, vol. 10, no. 1, p. 2317, May 2019.

[11] C. Favaretto et al., "Subcortical-cortical dynamical states of the human brain and their breakdown in stroke," *Nature Commun.*, vol. 13, no. 1, p. 5069, Aug. 2022.

[12] J. Zeng et al., "Local 5-HT signaling bi-directionally regulates the coincidence time window for associative learning," *Neuron*, vol. 111, no. 7, pp. 1118–1135, Apr. 2023.

[13] J. Ji, L. Han, F. Wang, and J. Liu, "Dynamic effective connectivity learning based on nonparametric state estimation and GAN," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–12, 2024.

[14] R. Xu, Q. Zhu, S. Li, Z. Hou, W. Shao, and D. Zhang, "MSTGC: Multi-channel spatio-temporal graph convolution network for multi-modal brain networks fusion," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 2359–2369, May 2023, doi: 10.1109/TNSRE.2023.3275608.

[15] Y. Kong et al., "Spatio-temporal graph convolutional network for diagnosis and treatment response prediction of major depressive disorder from functional connectivity," *Hum. Brain Map.*, vol. 42, no. 12, pp. 3922–3933, 2021.

[16] T. B. Tang, J. S. Chong, M. Kiguchi, T. Funane, and C.-K. Lu, "Detection of emotional sensitivity using fNIRS based dynamic functional connectivity," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 894–904, 2021.

[17] Y. Lin et al., "Learning dynamic graph embeddings for accurate detection of cognitive state changes in functional brain networks," *NeuroImage*, vol. 230, Apr. 2021, Art. no. 117791.

[18] M. Wang, J. Huang, M. Liu, and D. Zhang, "Modeling dynamic characteristics of brain functional connectivity networks using resting-state functional MRI," *Med. Image Anal.*, vol. 71, Jul. 2021, Art. no. 102063.

[19] J. Bijsterbosch et al., "Challenges and future directions for representations of functional brain organization," *Nature Neurosci.*, vol. 23, no. 12, pp. 1484–1495, Dec. 2020.

[20] Q. Zhu, R. Xu, R. Wang, X. Xu, Z. Zhang, and D. Zhang, "Stacked topological preserving dynamic brain networks representation and classification," *IEEE Trans. Med. Imag.*, vol. 41, no. 11, pp. 3473–3484, Nov. 2022.

[21] X. Tang, X. Zhang, M. Liu, and J. Zhang, "Learning asynchronous common and individual functional brain network for AD diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Cham, Swizerland. Springer, 2023, pp. 215–225.

[22] A. G. Hawkes, "Point spectra of some mutually exciting point processes," *J. Roy. Stat. Soc., Ser. B Methodol.*, vol. 33, no. 3, pp. 438–443, Oct. 1971.

[23] S. Gadgil et al., "Spatio-temporal graph convolution for resting-state fMRI analysis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Lima, Peru. Cham, Swizerland: Springer, Oct. 2020, pp. 528–538.

[24] M. Liu et al., "TMac: Temporal multi-modal graph learning for acoustic event classification," in *Proc. 31st ACM Int. Conf. Multimedia*, Oct. 2023, pp. 3365–3374.

[25] J. Tan, X. Zhang, C. Qing, and X. Xu, "Fourier domain robust denoising decomposition and adaptive patch MRI reconstruction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 6, pp. 7299–7311, Jun. 2022.

[26] M. Wan, M. Cai, Z. Yang, H. Tan, G. Yang, and M. Tang, "Robust latent nonnegative matrix factorization with automatic sparse reconstruction for unsupervised feature extraction," *Inf. Sci.*, vol. 648, Nov. 2023, Art. no. 119517.

[27] S. Grossberg, "Recurrent neural networks," *Scholarpedia*, vol. 8, no. 2, p. 1888, 2013.

[28] J. Cheng, L. Dong, and M. Lapata, "Long short-term memory-networks for machine reading," 2016, *arXiv:1601.06733*.

[29] J. Gu et al., "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018.

[30] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[31] X. Kan, W. Dai, H. Cui, Z. Zhang, Y. Guo, and C. Yang, "Brain network transformer," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 25586–25599.

[32] H. Gao and S. Ji, "Graph U-Nets," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 2083–2092.

[33] D. Chen and L. Zhang, "FE-STGNN: Spatio-temporal graph neural network with functional and effective connectivity fusion for MCI diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Cham, Switzerland. Springer, 2023, pp. 67–76.

[34] Q. Zhu, J. Huang, and X. Xu, "Non-negative discriminative brain functional connectivity for identifying schizophrenia on resting-state fMRI," *Biomed. Eng. Online*, vol. 17, no. 1, pp. 1–15, 2018.

[35] D. Mousa, N. Zayed, and I. A. Yassine, "Correlation transfer function analysis as a biomarker for Alzheimer brain plasticity using longitudinal resting-state fMRI data," *Sci. Rep.*, vol. 13, no. 1, p. 21559, Dec. 2023.

[36] X. Song et al., "Multicenter and multichannel pooling GCN for early AD diagnosis based on dual-modality fused brain network," *IEEE Trans. Med. Imag.*, vol. 42, no. 2, pp. 354–367, Feb. 2023.

[37] W. Cui et al., "Personalized functional connectivity based spatio-temporal aggregated attention network for MCI identification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 2257–2267, 2023, doi: 10.1109/TNSRE.2023.3271062.

[38] J. Li, Y. Wei, C. Wang, Q. Hu, Y. Liu, and L. Xu, "3-D CNN-based multichannel contrastive learning for Alzheimer's disease automatic diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.