MSTGC: Multi-Channel Spatio-Temporal Graph Convolution Network for Multi-Modal Brain Networks Fusion

Ruting Xu, Qi Zhu[ID], Shengrong Li, Zhenghua Hou, Wei Shao,
and Daoqiang Zhang[ID], *Senior Member, IEEE*

Abstract—**Multi-modal brain networks characterize the complex connectivities among different brain regions from structure and function aspects, which have been widely used in the analysis of brain diseases. Although many multi-modal brain network fusion methods have been proposed, most of them are unable to effectively extract the spatio-temporal topological characteristics of brain network while fusing different modalities. In this paper, we develop an adaptive multi-channel graph convolution network (GCN) fusion framework with graph contrast learning, which not only can effectively mine both the complementary and discriminative features of multi-modal brain networks, but also capture the dynamic characteristics and the topological structure of brain networks. Specifically, we first divide ROI-based series signals into multiple overlapping time windows, and construct the dynamic brain network representation based on these windows. Second, we adopt adaptive multi-channel GCN to extract the spatial features of the multi-modal brain networks with contrastive constraints, including multi-modal fusion Info-Max and inter-channel InfoMin. These two constraints are designed to extract the complementary information among modalities and specific information within a single modality. Moreover, two stacked long short-term memory units are utilized to capture the temporal information transferring across time windows. Finally, the extracted spatio-temporal features are fused, and multilayer perceptron (MLP) is used to realize multi-modal brain network prediction. The experiment on the epilepsy dataset shows that the proposed method outperforms several state-of-the-art methods in the diagnosis of brain diseases.**

Ruting Xu, Qi Zhu, Shengrong Li, Wei Shao, and Daoqiang Zhang are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: zhuqinuaa@163.com; shaowei20022005@nuaa.edu.cn; dqzhang@nuaa.edu.cn).

Zhenghua Hou is with the Department of Psychosomatics and Psychiatry, Affiliated Zhongda Hospital, School of Medicine, Southeast University, Nanjing 210009, China.

I. INTRODUCTION

IN RECENT years, brain network analysis methods based on neuroimaging technology have attracted more and more attention [1], and have been widely used in the brain disease diagnosis [2], [3]. The brain network analysis technology can detect the intrinsic correlation and interaction pattern in brain. A brain network model can be represented by a series of nodes and edges. The node represents brain region defined by physiological templates, and edges measure interactions between brain regions of interest (ROI) [4]. In general, the brain network can be divided into structural connectivity networks, such as diffusion tensor imaging (DTI) [5], and functional connectivity networks (fMRI), such as resting-state functional magnetic resonance imaging (rs-fMRI) [6]. Previous studies have shown that the different modalities of the brain network convey complementary information to each other [7], [8]. However, due to the heterogeneity and the complex topology of different modalities, it is still a challenge to investigate effective fusion of functional connectivity network and structural connectivity networks to improve the performance of feature representation and diagnosis.

There are many multi-modal brain network fusion methods have been proposed for brain disease diagnosis. As shown in **Figure 1**, existing methods for multimodal brain network analysis can be mainly divided into two categories. The first type is based on the non-graph structured data fusion strategy, such as multi-view embedding, multi-kernel learning (MKL) and principal component analysis (PCA). For example, Dai et al. fused fMRI and sMRI image features by MKL for the diagnosis of hyperactivity disorder [9]. Yang et al. used information from one modality to aid the construction of brain network from another modality [10], [11]. Unfortunately, the above fusion methods need to stretch the brain network into vector form, which destroys the topology of the brain network. The second category is based on the conventional graph structured data fusion strategy, which considers the topological characteristics of brain network in fusion. The most
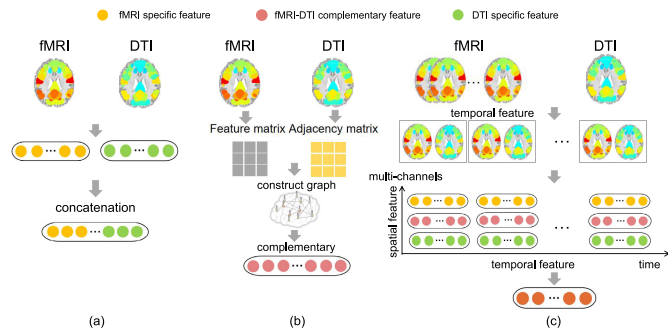
Fig. 1.   The comparison of the multi-modal fusion methods, (a) the non-graph structured data fusion strategy, (b) the conventional graph structured data fusion strategy, and (c) the proposed MSTGC method.

common model for this strategy utilizes a graph convolution network (GCN) to fuse multimodal structural and functional information. As a powerful representation of graph data, GCN can integrate the node features and the internal graph nodes of brain networks by nonlinear mapping. For example, Liu et al. [12] proposed a framework of Siamese community preserving graph convolutional network to learn the structural and functional joint embedding of brain networks.

It is worth noting that although these conventional fusion methods can integrate the structural and functional features of brain network, the topological and discriminative information of brain network over time in each modality is often ignored, which is important in classifying the brain networks. On one hand, the connection pattern of the functional brain pattern changes over time during the scanning period, so capturing temporal dynamic information of the brain network is conducive to improving its feature extraction [13]. Most of the existing multi-modal brain network analysis models are developed to fuse the static brain networks [10], [14], but they cannot take advantage of dynamic topological properties of brain networks for brain disease diagnosis. On the other hand, the specific features in the fMRI modality and DTI modality of the brain networks should also be effectively extracted and fused. With the increase of network layers, it is natural for a deep model to lose the information of the original input due to the information bottleneck problem. But most of the deep models focus on finding the complementary features among modalities, without ensuring how rich the features reflect the original data [15]. Therefore, it is necessary to investigate the spatio-temporal brain network fusion analysis method, which can simultaneously effectively extract the complementary and specific information in multi-modal brain network and exploit the spatial-temporal characteristics of among different modalities.

In order to solve the above challenging problems in multi-modal brain network fusion, in this paper, we propose a multi-channel spatio-temporal graph convolution network to distill both the spatial and temporal topological information from multi-modal brain networks, and apply it to the diagnosis of epilepsy. First, we divide the ROI-based series signals into multiple overlapping time windows, and we construct the dynamic functional brain network based on these windows. Second, an adaptive multi-channel GCN is

employed to obtain the spatial features of the multi-modal brain networks, in which the graph models of functional network and structural network are respectively used as the input of multi-channel GCN. The proposed multi-channel GCN consists of three convolution modules: two specific convolution modules that are utilized to extract unique features of fMRI modality and DTI modality, respectively, and a common convolution module that is used to fuse multiple modalities and extract complementary information. Then we use the attention mechanism to adaptively fuse the features encoded by different channels. Moreover, in order to encourage consistency in multi-modal graph representations and distill discriminative information from each modality, we develop two contrastive objectives: multi-modal fusion InfoMax and inter-channel InfoMin. Among them, the former objective extracts complementary information from different modalities so that the extracted features can reflect the complementary information of the brain networks of different modalities, while the latter distinguishes different graph views for the sake of capturing the specific information in each modality. Next, stacked Long Short-Term Memory (LSTM) units are exploited to capture the temporal information between time windows. Finally, we utilize MLP to classify the spatio-temporal topological features extracted from the multi-modalities. **Figure 2** shows the main difference between the proposed method with the conventional multi-modal brain network fusion methods.

In brief, the proposed multi-modal brain network fusion has the following advantages:

1) To the best of our knowledge, it is the first work that investigates spatio-temporal topological characteristics in multi-modal brain network fusion.

2) A multi-channel graph convolution network with stacked LSTM module is proposed to effectively capture the dynamic graph features embedded in dynamic brain network.

3) We develop mutli-modal graph contrastive learning method to improve the discriminability of the complementary and specific topological information among different modalities.

4) The results on epilepsy dataset show that the proposed method is significantly superior to the state-of-the-art multi-modal fusion diagnosis methods.

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

II. RELATED WORKS

A. Deep Learning Method With the Graph Data

Recently, Graph Neural Network has emerged as a powerful tool for understanding graph-structured data in many domains, such as protein structure [16], and brain network [17]. In essence, theses graph neural network methods in brain disease diagnosis can be mainly divided into two aspects: node classification and graph classification. In node classification, each node represents a subject and is associated with a feature vector extracted from imaging data. The edge weights encode the pairwise similarities between subjects and their

features obtained from auxiliary phenotypic data [18]. Graph classification regards the brain regions of interest as nodes and the constructed brain functional network with the form of the adjacency matrix of the graph [19]. The one-layer of GCN can be described as:

$$G(A, X) = ReLu(AXW) \qquad (1)$$

where A is adjacency matrix, X is feature matrix and W is weight matrix. Therefore, the output of GCN layers considers both brain structural connectome and functional features at the same time. Many multi-modal brain networks using GCNs are often based on guiding strategy. Although the interaction among modalities can be exploited in guiding strategy based fusion methods, they usually tend to ignore the unique characteristics in each modality.

B. Contrastive Learning for Network Embedding

Recent work on self-supervised visual representation learning has shown that comparing congruent and incongruent views allows the encoder to learn the effective features from the data, which is conducive to representing network information [20]. Different from the data in Euclidean space, the features in brain network data have both topological properties and adjacency relations. So it is natural to use graph contrastive learning to improve the feature representation for graph structured brain network. These methods maximize the mutual information between the input and the learned representation for fusion. For example, CREME [21] adopts node-to-node contrastive learning to distill information from embeddings generated from different graph views and capture the complementary information between them. Hassani and Khasahmadi [22] proposed a contrastive multi-view representation learning method on both node and graph levels. The embedding features can reflect the global structure information of the original network. Similarly, in the brain network fusion analysis, we hope that the features obtained can reflect the global and local features of the brain network. Therefore, we introduce graph contrast learning in our model to maximize the mutual information between the features of embedding and the brain network.

III. PROPOSED METHOD

In this work, we propose the multi-channel spatio-temporal graph convolution network for effectively fusing the functional and structural brain networks, and develop graph contrastive learning to enhance feature representation among modalities. Next, this section describes the proposed method in detail.

A. Dynamic Functional Brain Network and Structural Brain Network Construction

We assume that the rs-fMRI time-series data for a subject is $(x_1, \cdots, x_N)^T \in R^{N \times M}$, where each vector $x_n \in R^M (n = 1, \cdots, N)$ contains the blood oxygen level dependent (BOLD) measurement of the nth ROI at M successive time points. N represents the number of ROI. To characterize the temporal variability of the functional architecture associated with a set

of given regions, we segment all rs-fMRI time series into T overlapping windows with the constant length of L. For each subject, we can define a dataset $E = \left\{ \left(\{g_d^i, g_f^i\}, y \right) \right\}_{i=1}^{T}$ consisting of T windows, where g_d^i and g_f^i represent the DTI and fMRI modality of the subject, and y is the label of the corresponding subject. Each modality can be described in a weighted graph $g_*^i = (X, A_*^i)$, where $X \in R^{N \times d}$ is the node feature matrix consisting of time series within each time segment, $A_*^i \in R^{N \times N}$ is the brain network, and $*$ represents specific modality between fMRI and DTI.

B. Adaptive Multi-Channel GCN

The proposed multi-channel GCN is used to extract spatial information from multi-modal brain networks. The overall framework of MSTGC consists of three convolution modules: two specific convolution modules are utilized to extract unique features from fMRI modality and DTI modality respectively, and a common convolution module realizes multi-modal fusion. Then attention mechanism is exploited to adaptively fuse the features encoded by each channel.

1) Specific Convolution Module: The specific convolution module is composed of three stacked GCNs [23], and top-k pooling [24] is used to retain important nodes in the graph. First, in order to obtain the spatial characteristics of the fMRI mode under each window, we regard $g_f^i = (X, A_f^i)$ as the input of the module. Thus, the $l - th$ layer output $Z_f^{i(l)}$ can be represented as:

$$Z_f^{i(l)} = ReLu((D_f^i)^{-\frac{1}{2}} A_f^i (D_f^i)^{-\frac{1}{2}} Z_f^{i(l-1)} W_f^{i(l)}) \qquad (2)$$

where $W_f^{i\,(l)}$ is the weight matrix of the $l-th$ layer in GCN,the initial $Z_f^{i(0)} = X$. Besides, $A_f^i = A_f^i + I_f^i$, where A_f^i and I_f^i denote the adjacent matrix and an identity matrix respectively, and D_f^i is the $N \times N$ degree matrix. We denote the last layer output embedding as Z_f^i. In this way, we can learn the node embedding, which captures the specific information in fMRI modality. Similarly, for the DTI modality, we use $g_d^i = (X, A_d^i)$ as the channel input, and then the specific information encoded in DTI modality can be embedded as Z_d^i. The calculation formula can be expressed as:

$$Z_d^{i(l)} = ReLu((D_d^i)^{-\frac{1}{2}} A_d^i (D_d^i)^{-\frac{1}{2}} Z_d^{i(l-1)} W_d^{i(l)}) \qquad (3)$$

2) Common Convolution Module: Since fMRI reflects the functional characteristics of the brain, DTI reflects the structural characteristics, there is information coupling between the two modalities. Multi-modal fusion can discover complementary information among them, which is beneficial to improve the performance of feature representation and model detection. Here we design a common convolution module with a parameter sharing strategy to fuse multi-modal information. Specifically, we utilize common convolution module to extract the graph embedding Z_{CF}^i from fMRI graph $g_f^i = (X, A_f^i)$ as follows:

$$Z_{CF}^{i(l)} = ReLu((D_f^i)^{-\frac{1}{2}} A_f^i (D_f^i)^{-\frac{1}{2}} Z_{CF}^{i(l-1)} W_c^{i(l)}) \qquad (4)$$

where $W_c^{i\,(l)}$ is the $l-th$ layer weight matrix of common-GCN, $Z_{CF}^{i(l-1)}$ is the node embedding in the $(l - 1) - th$ layer, and
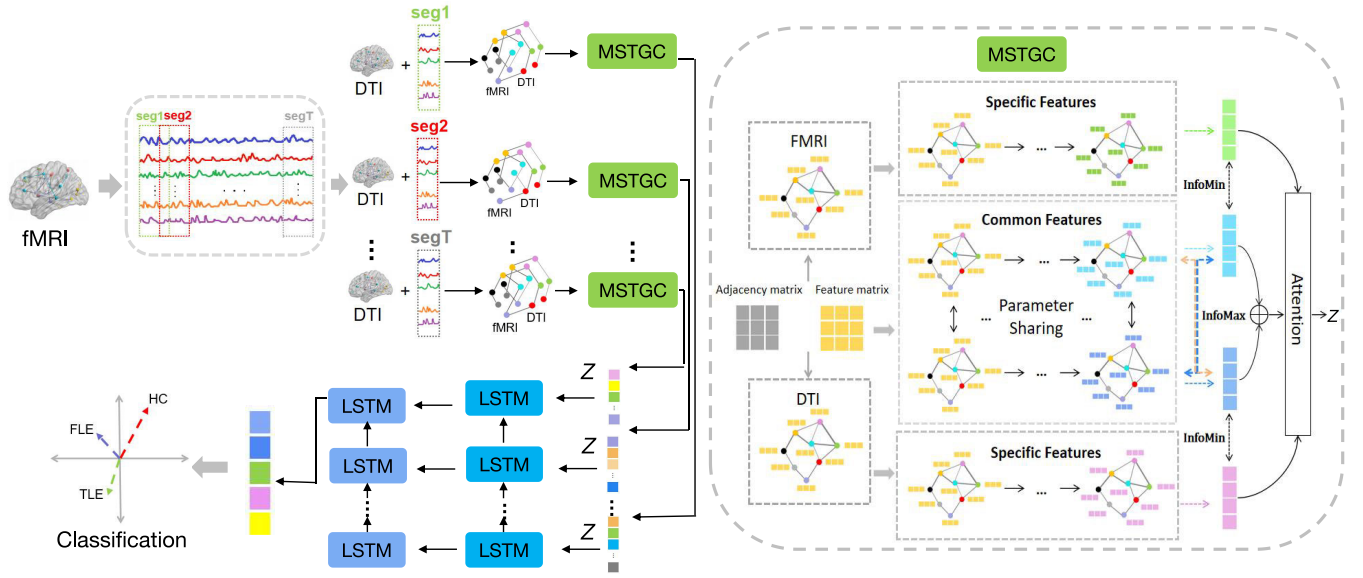
Fig. 2. The framework of the proposed method. The whole framework contains three parts: (1) the construction of dynamic brain network, (2) the extraction of spatial features by MSTGC, (3) the extraction of temporal features. As shown in the right panel, MSTGC consists of three GCN modules, in which the Specific Features module is used to extract the most distinguishing features of each modality separately, and common features module is utilized to extract multi-modal complementary information. In addition, we use attention mechanism and two graph contrastive objectives, InfoMax and InfoMin, to adaptively fuse multi-modal complementary and discriminative information from the learned graph data.

$Z_{CF}^{i(0)} = X$. In order to extract the shared information from DTI modality, we share the same weight matrix $W_c^{i(l)}$. for each layer of common convolution module, so the embedding of DTI graph can be formulated as:

$$Z_{CD}^{i(l)} = ReLu((D_d^i)^{-\frac{1}{2}} A_d^i (D_d^i)^{-\frac{1}{2}} Z_{CD}^{i(l-1)} W_c^{i(l)}) \quad (5)$$

The shared weight matrix can filter out the complementary information from two modalities. Finally, we get two output embeddings Z_{CF}^i and Z_{CD}^i and the common embedding Z_C^i of the two modalities is:

$$Z_C^i = (Z_{CF}^i + Z_{CD}^i)/2 \quad (6)$$

3) Attention Mechanism: Through above multi-channel GCN, we obtain two specific embedding Z_F, Z_D and one common embedding Z_C under each window. In order to extract the most correlated information Z^i to indicate subject, we utilize an attention mechanism $att(Z_F, Z_D, Z_C)$ to adaptively fuse these embeddings with the learned weights [25]. The formula can be described as:

$$(a_f, a_d, a_c) = att(Z_F, Z_D, Z_C) \quad (7)$$

where a_f, a_c, a_d indicate the attention values of Z_F, Z_D, Z_C, respectively. Specifically, take Z_F as example, we firstly transform the embedding through a nonlinear transformation, and use one shared attention vector $q \in R^{h \times 1}$ to get the attention value w_F^i as follows:

$$w_F^i = q^T tanh (W(Z_F^i)^T + b) \quad (8)$$

where W is the weight matrix, and b is the bias vector. Similarly, we can get the attention values of w_D^i and w_C^i. Then we normalize the attention values with the SoftMax function

to get the final weight:

$$a_f^i = softmax (w_F^i) = \frac{exp (w_F^i)}{exp (w_F^i) + exp(w_D^i) + exp(w_C^i)} \quad (9)$$

The value of a_f^i denotes the importance of Z_F. Similarly, we can also attain a_d^i, a_c^i through softmax function. Thus, we have the learned weights $a_f = [a_f^i]$, $a_d = [a_d^i]$, and $a_c = [a_c^i]$, and denote $a_F = diag (a_f)$, $a_D = diag (a_d)$ and $a_C = diag (a_c)$. Then we combine these three embeddings to obtain the final embedding Z of each window:

$$Z = a_F Z_F + a_D Z_F + a_C Z_C \quad (10)$$

Finally, we obtain a low-dimensional and dense representation for each subject.

C. Contrastive Learning of Multi-Modalities

GCN extracts deep features for disease diagnosis by combining the topology of the brain network and the information on the nodes of brain regions. However, the graph representation in graph classification is then transferred to a vector representation, which has the global property by gathering the information from all brain regions. However, both global and local properties of brain networks are equally important for disease diagnosis. Therefore, in order to maximize the mutual information between the global and local representations between fMRI and DTI graphs, we introduce a multi-view fusion InfoMax constraint on the common convolution module, so that the output representation can incorporate the global properties and local attributes of brain network. In addition, in order to make the embedding features under each channel discriminative, we impose the Inter-channel InfoMin constraint among channels.

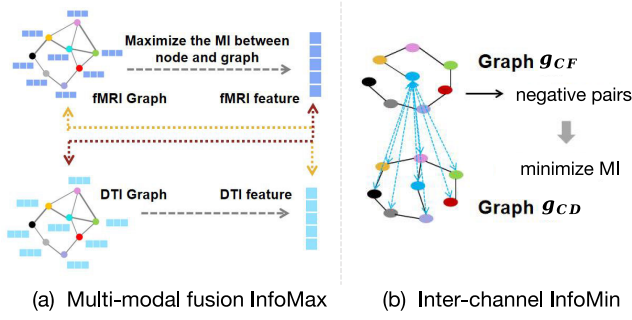(a) Multi-modal fusion InfoMax   (b) Inter-channel InfoMin

Fig. 3.   The schematic of the graph contrastive learning strategies with InfoMax and InfoMin (a) illustrates the contrastive strategies in multi-modal fusion, and (b) indicates the diversity constraints between different channels.

1) Multi-View Fusion InfoMax: We design a graph contrastive constraint to further enhance the consistency of two embeddings Z_{CF} and Z_{CD}, so as to promote the module to effectively integrate the information of the two modalities. The contrastive object is designed to maximize the mutual information between node representation of one modality and graph representations of another modality, which is shown in **Figure 3(a)**. Specifically, the brain network graph is first transformed into vector form and then we calculate the mutual information between it and the obtained features by GCN. The consistency constraint L_C between Z_{CF} and Z_{CD} can be formulated as:

$$L_C = \frac{1}{2S}(\frac{1}{N}\sum_{i=1}^{N} I\,(Z_{CF}; g_{CD}^i) + I(Z_{CD}; g_{CF}^i)) \qquad (11)$$

For the convenience of the optimization, we estimate mutual information $I(X; Y)$ in **Eq. (11)** by using Jason-Shannon Divergence (JSD):

$$I\left(Z_{CF}; g_{CD}^i\right) = -sp\left(-d\left(Z_{CF}, g_{CD}^i\right)\right) \\ - \frac{1}{N-1}\sum_{n_j \in N\{n_i\}} sp\left(d\left(Z_{CF}, g_{CD}^j\right)\right) \\ \qquad (12)$$

where $sp\,(x) = \log\,(1 + e^x)$ and d is a discriminator function, which takes the inner product with a sigmoid activation. The resulting multi-modal fusion representation can distill discriminative knowledge from each modality.

2) Inter-Channel InfoMin: Though the embedding Z_F and the embedding Z_{CF} are learned from the same graph structure $g_f = (X, A_f)$, they capture different information from brain networks. In order to make the representation of embeddings under different channels more distinguishable, we consider adding diversity constraints between channels. Our approach is to minimize the mutual information between different channel node representations g_f and g_{cf}. Specifically, we regard all nodes in two graphs as negative pairs and minimize mutual information between node feature vectors on each negative pair, which is shown in **Figure 3(b)**. Thus, the loss function

of (g_{cf}^i, g_f) can be calculated as:

$$l\left(g_{cf}^i, g_f\right) = \log\frac{e^{\theta(g_{cf}^i, g_f)/\tau}}{e^{\theta(g_{cf}^i, g_f)/\tau} + \sum_{j=1}^{k} e^{\theta(g_{cf}^i, g_f^j)/\tau}} \qquad (13)$$

where k is the number of the remaining nodes and τ is temperature parameter. We define $\theta\,(u, v) = s\,(p\,(u)\,, p\,(v))$ as a critic function, $s(\cdot, \cdot)$ is implemented using a simple cosine similarity, and $p(\cdot)$ is a non-linear projection to enhance the expression power of the critic function. In this work, we adopt MLP as above non-linear projection. Since these graphs are symmetric, so we can summarize the L_{SF} as:

$$L_{SF} = \frac{l}{2N}\sum_{i=1}^{N}[l\,(g_{cf}^i, g_f) + l\,(g_{cf}, g_f^i)] \qquad (14)$$

Similarly, the loss L_{SD} between Z_{CD} and Z_D can be calculated as:

$$L_{SD} = \frac{l}{2N}\Sigma_{i=1}^{N}l\,(g_{cd}^i, g_d) + l\,(g_{cd}, g_d^i) \qquad (15)$$

So we set the disparity constrain L_d as:

$$L_d = L_{SF} + L_{SD} \qquad (16)$$

In addition, we use cross-entropy to measure the loss of subject classification and represent it as L_t. In general, combining the subject classification loss and constrains, we can formula the overall objective function as:

$$L = L_t + \alpha\,L_d + \beta L_c \qquad (17)$$

where α and β are parameters of the disparity and consistency constraint terms. Finally, we can optimize the proposed model via backpropagation and learn subjects' embeddings for classification.

D. Temporal Convolutional Layer

The brain is essentially a dynamic system, in which the brain network constantly reconstructing overtime during the scanning period [26], [27]. Brain regions interact dynamically with each other over time. Considering that long short-term memory (LSTM) can properly solve the problem of gradient disappearance and gradient explosion of traditional RNN [28], we use two stacked LSTM units to capture the temporal information transmitted across time windows. Each of these LSTM is followed by batch normalization and tanh activation. Finally, the fully-connected layer is employed to learn a mapping between the dynamic embedding features and the disease progression prediction. The overall framework of the method is shown in **Algorithm 1**.

E. Implementation

The proposed method was implemented using Python based on Pytorch, and the model was trained on a single GPU (NVIDIA GeForce RTX 2080 Ti). We optimize the proposed method via the Adam algorithm, with the learning rate of 0.001, the number of epochs of 300, and the batch-size of 25. In multi-channel GCN module, we used 3 stacked GCN modules for each channel, and the drop-out was set to 0.5.

Algorithm 1 The Algorithm of the Proposed of Our MSTGC Method

---

Require: fMRI data and DTI data; Parameters α and β;
Ensure: ACC, PRE, RECALL, F1, AUC;

1: Initialization: the number of time windows; weight matrices W_f, W_d and W_c;
2: Divide sliding time windows and construct graph $g_f = (X, A_f)$ and $g_d = (X, A_d)$ under each window;
3: **While** epoch

$$<$$

max_epochs **do**
4:     Extract features using multi-channel GCN and then obtain specific feature Z_f, Z_d and common features Z_c;
5:     Fuse multi-channel features by **Eq.(10)**;
6:     Obtain time information by two-stacked LSTM;
7:     Maximize the mutual information between two modalities by **Eq.(11)**;
8:     Minimize the mutual information between different channel L_d by **Eq.(16)**;
9:     $L \leftarrow L_t, L_c, L_d$ by **Eq.(17)**;
10:     Back-propagate L to update model weights;
11:     epoch+=1;
12: **end while**

In the temporal convolutional layer, the number of neurons for each LSTM unit of two layers were 180, 90, respectively. Each temporal convolutional layer was followed by batch normalization, tanh activation. Based on the output of the stacked LSTM, a fully-connected layer with 2 neurons was utilized to predict the category of the subject, where the sigmoid was used as the activation functions of the last fully-connected layer.

IV. Experiment

In this section, we compare the proposed method with the previous brain network analysis algorithms in recent references. In addition, we also carried out ablation experiments, which proved the effectiveness of our proposed method.

A. Data and Pre-Processing

In this work, we used the epilepsy dataset collected from Jinglin Hospital, Nanjing University School of Medicine for experiment. It contains 103 frontal lobe epilepsy (FLE) patients (all right-handed, 50 female, age range: 17-51 years, mean age 24.1), 89 temporal lobe epilepsy (TLE) patients (all right-handed, 45 female, age range: 17-51 years, mean age 25.9) and 114 normal controls (NC) (all right-handed, 56 female, age range: 20-38 years, mean age 26.2). By using Siemens Trio 3T scanner, the raw rs-fMRI data and DTI data of all participants are collected. The scan parameters of rs-fMRI are as follows: TR = 2000ms, TE = 30ms, flip angle = 90°, vorel size = $3.75 \times 3.75 \times 3.75 mm^3$. The scan parameters of DTI are as follows: TR = 6100ms, TE = 93ms, flip angle = 90°, vowel size = $0.94 \times 0.94 \times 3 mm^3$. All rs-fMRI images are preprocessed using SPM8 in the DPARSF

toolbox. The resulting volumes have 240 timepoints and are parcellated into 90 regions of interest (ROIs) using the AAL atlas. These time series reflect information about brain activity. The DTI data are processed by using the PANDA suite. First, we use the FSL toolbox to perform distortion correction on the DTI, remove eddy currents and extract brain masks from the B0 image. Then, based on each subject's co-registered T1 images, the TrackVis is used to obtain fiber images by deterministic tracking method, and anatomical regions were defined using AAL conventions. Finally, through the number of fibers, we can get the structural information of the brain network and the strength of physical connections.

B. Method for Comparison

To validate the effectiveness of our proposed method, we compare it with several state-of-the-art brain network representation and classification methods. Among them, the first 4 methods are based on single-modal learning, including G-Unet [24], dynamic functional connectivity network (dFCN-LSTM) [29], dynamic functional connectivity (DFC) [4], BrainNetCNN [30], and the following 5 methods are multi-modal fusing approaches, including self-attention graph pooling (SAGPool) [31], siamese community preserving graph convolutional network (SCP-GCN) [12], multi-graph multi-view graph embedding(M2E) [32], multilinear principal component analysis(MPCA) [33], and GCNeuro [34].

(1) G-Unet: In this work, the brain functional connectivity network constructed by fMRI of each subject is considered as the adjacency matrix of the graph, and the time series is considered as the feature matrix of each node.

(2) dFCN-LSTM: This model fuses past and future information from fMRI to effectively learn time-series changes in signals from brain regions by using LSTM.

(3) DFC: In DFC, Pearson correlation coefficient of the whole time series is used for static FC calculation. Dynamic FC algorithm considers the moving window of time series. After that, the main repetitive FC matrix is found by clustering algorithm.

(4) BrainNetCNN: Three specialized convolutional layer types for DTI datasets are proposed, aiming to exploit the inherent structure of weighted brain networks.

(5) SAGPool: In SAGPool, a graph pooling method based on self-attention is used to extract the information of graph structure, which can take into account both node features and graph topology.

(6) SCP-GCN: In SCP-GCN, a framework of Siamese community preserving graph convolutional network are utilized to learn the structural and functional joint embedding of brain networks.

(7) M2E: In this method, a multi-view brain network processing framework is used to extract the high-order representation features in multi-modal data. We treat fMRI network and DTI network as different views and use tensor techniques to exploit the correlations between the multi-view brain networks.

(8) MPCA: The MPCA method mainly uses the multilinear principal component analysis method to analyze the brain network. We first concatenate the fMRI and DTI data into

TABLE I
PERFORMANCE COMPARISON OF THE PROPOSED AND COMPETING METHODS (%)

Method	ACC	F1	AUC	ACC	F1	AUC
	NC vs. TLE & FLE			NC vs. TLE		
G-Unet	73.96±0.85	81.19±0.91	74.24±1.78	68.18±0.60	69.57±1.56	65.88±1.01
dFCN_LSTM	66.72±1.11	56.90±1.61	66.35±1.07	62.64±1.19	65.02±2.43	62.63±2.28
DFC	75.78±2.13	66.56±2.01	69.23±2.78	70.93±1.94	76.30±2.30	68.11±2.44
BrainNetCNN	73.77±2.38	78.95±2.95	71.07±2.09	72.73±1.67	66.67±1.95	68.34±1.64
SAGPool	78.13±1.50	82.93±1.86	76.52±1.24	69.70±1.53	70.26±1.81	66.32±1.89
SCP-GCN	77.41±1.70	74.07±1.74	73.11±1.54	75.75±1.01	78.64±1.10	68.33±1.60
M2E	65.69±1.09	55.37±2.02	69.78±2.00	61.29±1.38	62.52±1.01	55.35±1.27
MPCA	62.45±3.48	70.30±3.65	67.78±3.23	62.52±2.84	57.66±2.78	56.04±2.72
GCNeuro	67.98±1.04	72.36±0.91	56.09±0.71	67.40±1.04	55.86±0.90	66.53±0.71
MSTGC	**81.58±0.55**	**85.37±0.62**	**80.47±1.50**	**83.30±0.70**	**79.20±0.75**	**77.64±0.54**
	NC vs. FLE			TLE vs. FLE		
G-Unet	70.00±0.63	72.73±1.06	70.83±0.81	63.33±0.46	61.11±1.47	58.42±0.88
dFCN_LSTM	64.01±1.12	70.97±2.20	68.40±2.99	60.47±1.40	66.98±3.51	59.50±2.87
DFC	67.65±0.90	67.75±1.38	66.01±1.77	69.72±1.74	2.69±2.10	68.59±1.98
BrainNetCNN	67.50±1.73	71.11±2.10	67.04±1.45	66.67±1.04	57.14±0.91	65.00±0.71
SAGPool	71.43±0.72	66.07±1.57	70.84±0.94	67.50±0.72	66.04±2.73	65.94±1.16
SCP-GCN	78.57±0.74	71.43±1.34	74.55±0.80	71.05±0.55	66.67±2.01	69.34±1.51
M2E	64.04±1.28	68.12±1.76	57.56±1.55	69.27±0.97	70.95±1.31	63.89±1.22
MPCA	54.98±2.10	55.57±2.33	52.73±1.79	54.13±1.09	53.30±1.17	60.23±1.01
GCNeuro	70.87±1.04	43.33±0.91	59.54±0.71	66.79±3.19	41.04±3.65	58.79±3.11
MSTGC	**83.27±1.01**	**78.57±1.60**	**77.23±1.60**	**75.59±0.73**	**74.57±1.52**	**71.07±0.80**

a three-dimensional tensor, and then utilize MPCA to obtain the features of the tensor.

(9) GCNeuro: In the GCNeuro method, we trained two Graph Attention Networks (GATs) for the fMRI modality and the DTI modality separately by sharing weight. Then we aggregated the features of each node in two modalities and regard it as feature matrix of the composite graph.

C. Experimental Setup

In the experiment, we evaluate the proposed method and other comparison methods on epilepsy dataset based on a 10-fold cross-validation strategy. Specifically, we divided the data set into 10 parts. Due to the slight imbalance of the subject categories, we utilize stratified 10-fold cross validation, that is, each compromise maintains the same proportion as the original subject category. After that, eight folds are taken as the training dataset each time, one fold is used as the validation dataset to determine the optimal value of the objective function parameters, and the remaining fold is used as the test data. In order to compare the performance of each method fairly, grid search is adopted to find the optimal model parameters. For the deep learning based methods, the training will not stop until the loss converges to the threshold. Four binary classification tasks are utilized to evaluate our model's diagnostics capacity for epilepsy (NC vs. TLE & FLE), diagnostic capacity for each subtype (NC vs. TLE, and NC vs. FLE), and diagnostic capacity between the two subtypes (TLE vs. FLE). The performance of each task are measured by five

metrics, i.e., accuracy (ACC), precision (Pre), recall (Rec), F-Measure (F1), and area under the ROC curve (AUC). For all the methods, we perform 10-fold cross-validation 10 times, and report the average results.

D. Classification Performance on Epilepsy Dataset

We divided the comparison methods into two groups for analysis: single-modal group and multi-modal group. The result can be seen in **Table I**. It shows that our proposed method is effective under all 4 tasks on the epilepsy dataset and outperforms other algorithms. To be specific, compared with the static methods such as G-Unet and BrainNetCNN, our proposed method achieves better performance by integrating the spatio-temporal information of brain network. Compared with DFC and dFCN-LSTM, our method has greatly improvement on each task, which proves multi-modal classification is indeed an effective method for diagnosis of brain diseases. The other multi-modal fusion methods are also compared in the experiment, including machine learning methods, including M2E, MPCA and some deep learning methods, including GCNeuro, SCP-GCN, SAGPool. We find that deep learning models generally achieve better classification results than other comparison methods. This is because GCN can preserve direct and indirect relationships between nodes, so that it can capture more potential high-order features. In addition, although SAGPool, SCP-GCN and GCNeuro utilize GCN model and introduce fMRI and DTI to jointly learn brain network, they only retain the common feature between modalities and do not

retain the unique features of fMRI and DTI. In general, our method can achieve best classification results among all the methods on all the four diagnosis tasks.

E. Ablation Studies

In order to verify the effectiveness of the three components of multi-channel GCN, graph contrastive learning, and LSTM in our proposed MSTGC method, we conduct the ablation experiments and present experiment results in **Table II**. In the table, MSTGC-1,..., MSTGC-7 represent the variants methods of our proposed MSTGC method. We use "√" to indicate that the corresponding component is used. Otherwise, it is not used. For example, MSTGC-1 indicates that it neither uses graph contrastive learning to enhance the topology feature representation of brain network nor utilizes LSTM to obtain the temporal information of dynamic brain network. Compared with the accuracies of MSTGC-4, MSTGC-5, and MSTGC-6, we can find that abandoning any components of multi-channel GCN, graph contrastive learning, and LSTM will reduce the experiment accuracies. Compared with group of MSTGC-1, MSTGC-2, and MSTGC-3 and the group of MSTGC-4, MSTGC-5, and MSTGC-6, it can be concluded that if any two components are discarded, the accuracies will be lower than one of them is discarded, which further demonstrates that these components can improve disease diagnostic accuracy.

Multi-channel GCN obtains specific information and complementary features from fMRI modality and DTI modality by multiples channels, and combines the attention mechanism to adaptively fuse these features. Graph contrastive learning maximizes the mutual information between graph representation and the extracted features, thereby enhancing the feature representation of brain networks from both global and local perspectives. LSTM module obtains the temporal information of dynamic brain network. These three modules make our model have promising disease diagnosis performance.

V. DISCUSSION

In this section, we discuss the sensitivity of parameters in the proposed model, and then analyze the discriminative connectivities and brain regions discovered by our method. In addition, we also visualize the embedding features.

A. Analysis of Parameter Sensitivity

According to **Eq. (17)**, there are two parameters in the total loss function. We use the grid search method to show the influence of parameters on our experimental results and find the optimal parameters. Specifically, we set parameter α in the range of [1, 0.1] and parameter β in the range of [1, 0.1]. The results are shown in **Figure 4**. In general, it can be seen from the figure that no matter what the values of parameters α and β are, our model can obtain good results on the four tasks of epilepsy, which proves that our proposed method is robust to these parameters.

In addition, we also discuss the influence of the number of windows dividing on the experimental results. Specifically, we constructed the DCNs with different window numbers located in [5] and [15] and conducte experiments under
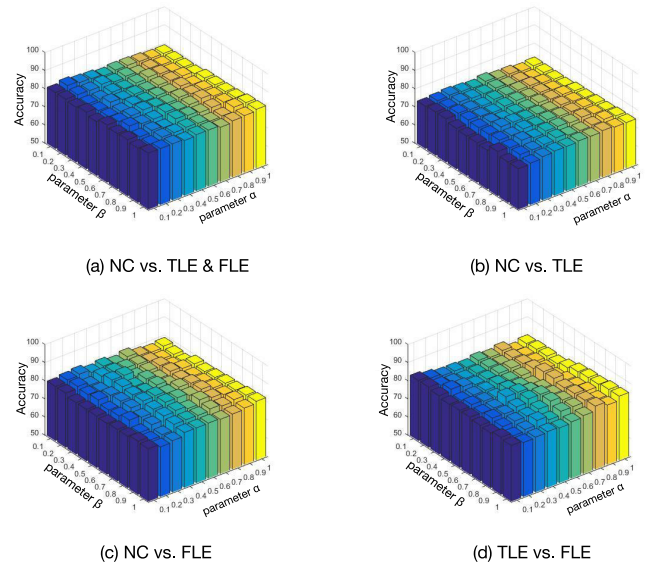

Fig. 4. The accuracies (%) of our method with different parameter combinations under the four diagnosis tasks.

four classification tasks respectively. The result is shown in **Figure 5**. As can be seen from **Figure 5**, for NC vs. TLE & FLE task, the classification accuracy reaches highest point, when the number of time windows equal to 6. For NC vs. TLE, the accuracy of classification reaches a peak, when the number of time window is 5. For the NC vs. FLE task and TLE vs FLE task, the model performs best, when the number of sliding windows was 7. In addition, we find that the change of accuracies with different sliding time windows number has the similar trend under four tasks. With the increasing of number of windows, the accuracy of classification first increases and then decreases. This trend is reasonable, because the number of windows is too small, information interaction among ROIs under different time windows will be lost. When the number of windows is small, it may contain too much redundant information in each time window, which will decrease the classification accuracy as well. In addition, we also discuss the effect of the dropout on the model results. The experimental results are shown in **Figure 5**, where the blue line represents dropout value equal to 0.5, the orange line is 0.4, and the green line is 0.3. From the figure, we can see that the model achieves the best classification performance when the dropout value is 0.5 on all the four tasks. We believe this phenomenon is reasonable, because the lower dropout is, the fewer nodes are retained, which is easy to cause information loss, resulting in the decline of classification accuracy.

B. Discriminative Functional Connectivity and Regions

Because not all ROIs are strongly associated with epilepsy, and we attempt to utilize our proposed method to figure out the most discriminative ROIs for understanding brain abnormalities. Specifically, we respectively visualized the 10 most relevant ROIs and the connections between them under the three task states including NC vs. TLE, NC vs. FLE, and TLE vs. FLE. For each classification task, since the selected features are different in each 10-fold cross-validation, we choose

TABLE II
THE ACCURACIES OF ABLATED MODELS ON DIFFERENT DIAGNOSIS TASKS (%)

Methods	Multi-channel GCN	graph contrastive learning	LSTM	NC vs. TLE&FLE	NC vs. TLE	NC vs. FLE	TLE vs. FLE
MSTGC-1	✓			77.79	75.58	78.46	70.43
MSTGC-2		✓		74.19	70.43	70.00	68.42
MSTGC-3			✓	67.74	61.90	65.00	60.47
MSTGC-4	✓	✓		79.79	76.01	76.01	70.96
MSTGC-5		✓	✓	75.49	71.37	77.28	69.38
MSTGC-6	✓		✓	70.96	81.24	79.87	74.93
MSTGC	✓	✓	✓	**81.58**	**83.30**	**83.27**	**75.59**
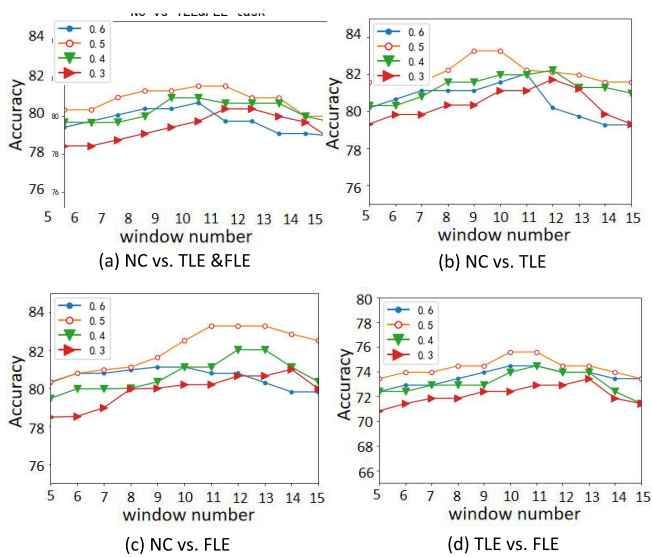
Fig. 5. Effects of different values of window number and dropout on model classification results.
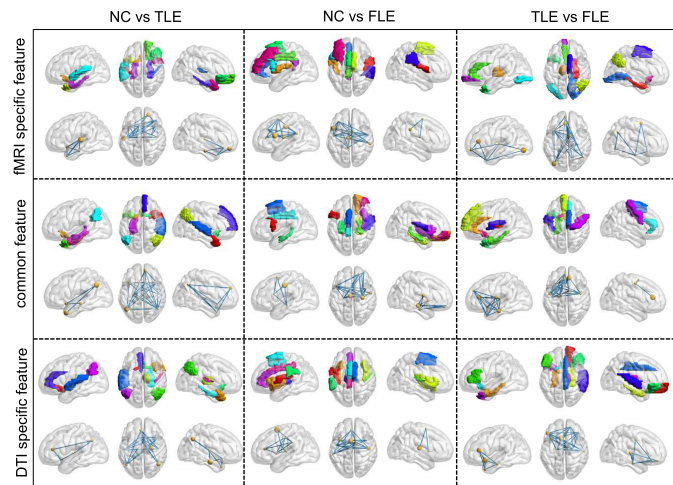

Fig. 6. The spatial location of the top 10 most differentiated brain regions and connectivities among them under three task states, including NC vs. TLE, NC vs. FLE, and TLE vs. FLE.

features that always occur in all folds of cross-validation as the most important features. On the task of the NC vs. TLE, the most relevant ROIs are concentrated on Parahippocampal gyrus, Olfactory cortex and middle temporal gyrus. On the task of NC vs. FLE, the key regions include Superior frontal gyrus, middle temporal gyrus and Supplementary motor area. On the task of TLE vs. FLE, the most essential region includes middle temporal gyrus and Parahippocampal gyrus. These brain areas have also been suggested to be related to epilepsy in previous studies [34], [35], [36], [37]. In addition, according to **Figure 6**, we find that although the most relevant brain regions searched under fMRI-based, DTI-based and fMRI-DTI-based are not all the same, they are all related to epilepsy and complement each other, which further proves that mining specific and complementary features from different modes can better describe brain networks.

C. The Effectiveness of Multi-Modal Fusion

In the experiment, we utilize multi-channel GCN to extract specific and common features in fMRI modalities and DTI modalities and adaptively fuse them by attention mechanism. MSTGC learns specific features in each modality and the

multi-modal complementary features among different modalities. The significance of each feature is evaluated by attention mechanism. We analyzed these attention values under 4 tasks on the epilepsy dataset, and the results are shown in **Figure 7**. From **Figure 7**, we can see that for these four tasks, the attention value of common embeddings is larger than that in DTI specific features, and the attention value of specific features contained in fMRI modality is between them. It means that complementary information between multi-modalities is more important than specific information under a single modality. This phenomenon is also consistent with the results of our ablation experiments in **Table II**. In addition, it can be seen from **Figure 7** that, under the four tasks, the specific features extracted from fMRI modality, the specific features extracted from DTI modality, and the common features among them always have relatively high attention values, which demonstrates the necessity of fusing modal specific and complementary features. In addition, in **Table II**, the diagnosis accuracy of MSTGC-5, which only uses the complementary features of fMRI and DTI, is obviously lower than our method. Therefore, the adaptive fusion of specific features and complementary features from multi-modal brain networks can effectively improve the diagnosis performance.
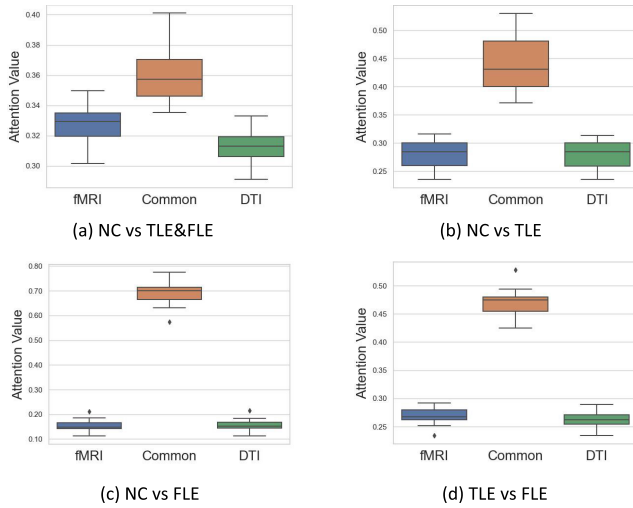
Fig. 7.   The analysis of attention distribution under different diagnosis tasks.
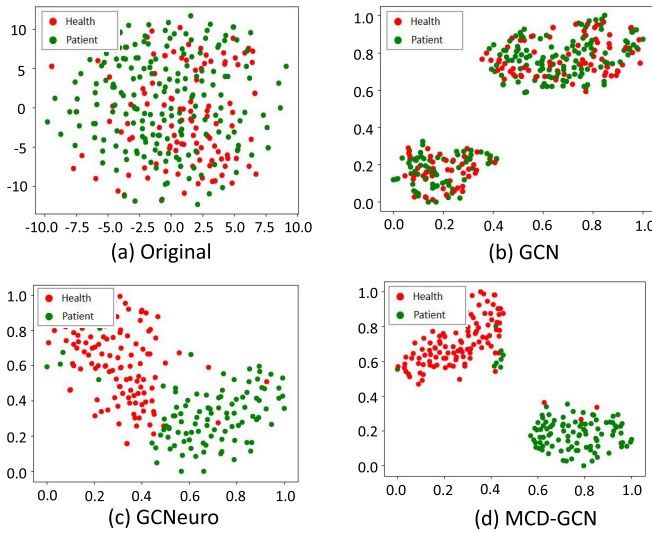

Fig. 8.   The t-SNE visualization of original data and the graph embeddings learned by Original, GCN, GCNeuro, MSTGC, respectively, on the epilepsy dataset.

D. Analysis of Embedding Features

What's more, in order to compare more intuitively and further demonstrate the effectiveness of our proposed method, we conduct the task of visualization on epilepsy dataset. We use the output embedding on the last layer of MSTGC before SoftMax and plot the embedding by t-SEN. The result is illustrated in **Figure 8**, in which the red dots represent healthy subjects and the green dots represent patients. According to **Figure 8**, it can be seen that the result of GCN and GCNeuro are not satisfactory, because the nodes with different labels are mixed together. Our method performs best, where the embedding features we learned have the clearest distinct boundaries among different classes. The reason why our method can clearly distinguish healthy subjects and patients is that we introduce graph contrast learning, which makes the embedding features more discriminative, reflecting the global and local features of original brain network.

VI. CONCLUSION

In this work, we propose a multi-channel spatio-temporal GCN that mines the dynamic typological features from rs-fMRI and DTI brain networks. Specifically, we employ an adaptively multi-channel GCN to learn the specific and complementary features of rs-fMRI and DTI. In order to reveal the global and local structures from the original brain networks, two graph contrast learning constraints, i.e. multi-modal fusion InfoMax and inter-channel InfoMin, are developed and imposed on the model. After obtaining the spatial characteristics of multi-modal brain networks, LSTM units are leveraged to capture the temporal dynamic pattern of functional connectivities along multiple time windows. Finally, MLP is used to classify the spatio-temporal features. The proposed MSTGC can effectively fuse the specific and complementary typological features of multi-modal and the dynamic characteristics of brain networks simultaneously. Experimental results on the epilepsy dataset show that our proposed MSTGC has achieved better diagnosis performance than the state-of-the-art methods in identifying epilepsy patients from healthy controls.

REFERENCES

[1] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009.

[2] L. Qiao, H. Zhang, M. Kim, S. Teng, L. Zhang, and D. Shen, "Estimating functional brain networks by incorporating a modularity prior," *NeuroImage*, vol. 141, pp. 399–407, Nov. 2016.

[3] R. Yu, H. Zhang, L. An, X. Chen, Z. Wei, and D. Shen, "Connectivity strength-weighted sparse group representation-based brain network construction for MCI classification," *Hum. Brain Mapping*, vol. 38, no. 5, pp. 2370–2383, May 2017.

[4] S. Yu et al., "Multi-scale enhanced graph convolutional network for early mild cognitive impairment detection," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*. Cham, Switzerland: Springer, 2020, pp. 228–237.

[5] S. Mori and J. Zhang, "Principles of diffusion tensor imaging and its applications to basic neuroscience research," *Neuron*, vol. 51, no. 5, pp. 527–539, Sep. 2006.

[6] P. M. Matthews and P. Jezzard, "Functional magnetic resonance imaging," *J. Neurol., Neurosurgery Psychiatry*, vol. 75, no. 1, pp. 6–12, 2004.

[7] X. Chen, H. Zhang, L. Zhang, C. Shen, S. Lee, and D. Shen, "Extraction of dynamic functional connectivity from brain grey matter and white matter for MCI classification," *Hum. Brain Mapping*, vol. 38, no. 10, pp. 5019–5034, Oct. 2017.

[8] D. Zhu et al., "Fusing DTI and fMRI data: A survey of methods and applications," *NeuroImage*, vol. 102, pp. 184–191, Nov. 2014.

[9] D. Dai, J. Wang, J. Hua, and H. He, "Classification of ADHD children through multimodal magnetic resonance imaging," *Frontiers Syst. Neurosci.*, vol. 6, p. 63, Sep. 2012.

[10] J. Yang, Q. Zhu, R. Zhang, J. Huang, and D. Zhang, "Unified brain network with functional and structural data," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*. Cham, Switzerland: Springer, 2020, pp. 114–123.

[11] M. S. Oechslin, M. Gschwind, and C. E. James, "Tracking training-related plasticity by combining fMRI and DTI: The right hemisphere ventral stream mediates musical syntax processing," *Cerebral Cortex*, vol. 28, no. 4, pp. 1209–1218, Apr. 2018.

[12] J. Liu, G. Ma, F. Jiang, C. Lu, P. S. Yu, and A. B. Ragin, "Community-preserving graph convolutions for structural and functional joint embedding of brain networks," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 1163–1168.

[13] Z. Jiao et al., "Dynamic routing capsule networks for mild cognitive impairment diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*. Cham, Switzerland: Springer, 2019, pp. 620–628.

[14] A. Kazi et al., "Graph convolution based attention model for personalized disease prediction," in *Proc. Int. Conf. Med. Image Comput. Computer-Assist. Intervent.* Cham, Switzerland: Springer, 2019, pp. 122–130.

[15] E. Jeon, E. Kang, J. Lee, J. Lee, T.-E. Kam, and H.-I. Suk, "Enriched representation learning in resting-state fMRI for early MCI diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 397–406.

[16] S. Sanyal, I. Anishchenko, A. Dagar, D. Baker, and P. Talukdar, "ProteinGCN: Protein model quality assessment using graph convolutional networks," *bioRxiv*, Apr. 2020, doi: 10.1101/2020.04.06.028266.

[17] Y. Zhu, H. Cui, L. He, L. Sun, and C. Yang, "Joint embedding of structural and functional brain networks with graph neural networks for mental illness diagnosis," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 272–276.

[18] S. I. Ktena et al., "Metric learning with spectral graph convolutions on brain connectivity networks," *NeuroImage*, vol. 169, pp. 431–442, Apr. 2018.

[19] L. Zhang, L. Wang, and D. Zhu, "Jointly analyzing Alzheimer's disease related structure-function using deep cross-model attention network," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 563–567.

[20] Y. Tian, D. Krishnan, and P. Isola, "Contrastive multiview coding," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2020, pp. 776–794.

[21] M. Zhang, Y. Zhu, Q. Liu, S. Wu, and L. Wang, "Deep contrastive multiview network embedding," 2021, *arXiv:2108.08296*.

[22] K. Hassani and A. H. Khasahmadi, "Contrastive multi-view representation learning on graphs," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 4116–4126.

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

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

[25] X. Wang, M. Zhu, D. Bo, P. Cui, C. Shi, and J. Pei, "AM-GCN: Adaptive multi-channel graph convolutional networks," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 1243–1253.

[26] U. Braun et al., "Dynamic reconfiguration of frontal brain networks during executive cognition in humans," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 37, pp. 11678–11683, Sep. 2015.

[27] H. Ju and D. S. Bassett, "Dynamic representations in networked neural systems," *Nature Neurosci.*, vol. 23, no. 8, pp. 908–917, Aug. 2020.

[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[29] W. Yan, H. Zhang, J. Sui, and D. Shen, "Deep chronnectome learning via full bidirectional long short-term memory networks for MCI diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2018, pp. 249–257.

[30] J. Kawahara et al., "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017.

[31] J. Lee, I. Lee, and J. Kang, "Self-attention graph pooling," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 3734–3743.

[32] Y. Liu, L. He, B. Cao, P. Yu, A. Ragin, and A. Leow, "Multiview multi-graph embedding for brain network clustering analysis," in *Proc. AAAI Conf. Artif. Intell.*, vol. 32, no. 1, 2018, pp. 1–8.

[33] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," *IEEE Trans. Neural Netw.*, vol. 19, no. 1, pp. 18–39, Jan. 2008.

[34] J. Huang, L. Zhou, L. Wang, and D. Zhang, "Attention-diffusion-bilinear neural network for brain network analysis," *IEEE Trans. Med. Imag.*, vol. 39, no. 7, pp. 2541–2552, Jul. 2020.

[35] X. Wu et al., "Altered intrinsic brain activity associated with outcome in frontal lobe epilepsy," *Sci. Rep.*, vol. 9, no. 1, pp. 1–8, Jun. 2019.

[36] A. B. Waites, R. S. Briellmann, M. M. Saling, D. F. Abbott, and G. D. Jackson, "Functional connectivity networks are disrupted in left temporal lobe epilepsy," *Ann. Neurol.*, vol. 59, no. 2, pp. 335–343, Feb. 2006.

[37] H. M. H. Braakman et al., "Frontal lobe connectivity and cognitive impairment in pediatric frontal lobe epilepsy," *Epilepsia*, vol. 54, no. 3, pp. 446–454, Mar. 2013.