Personalized Functional Connectivity Based Spatio-Temporal Aggregated Attention Network for MCI Identification

Weigang Cui, Yulan Ma, Jianxun Ren, Jingyu Liu, Guolin Ma, Hesheng Liu, and Yang Li, *Senior Member, IEEE*

***Abstract*— Functional connectivity (FC) networks derived from resting-state magnetic resonance image (rs-fMRI) are effective biomarkers for identifying mild cognitive impairment (MCI) patients. However, most FC identification methods simply extract features from group-averaged brain templates, and neglect inter-subject functional variations. Furthermore, the existing methods generally concentrate on spatial correlation among brain regions, resulting in the inefficient capture of the fMRI temporal features. To address these limitations, we propose a novel personalized functional connectivity based dual-branch graph neural network with spatio-temporal aggregated attention (PFC-DBGNN-STAA) for MCI identification. Specifically, a personalized functional connectivity (PFC) template is firstly constructed to align 213 functional regions across samples and generate discriminative individualized FC features. Secondly, a dual-branch graph neural network (DBGNN) is conducted by aggregating features from the individual- and group-level templates with the cross-template FC, which is beneficial to improve the feature discrimination by considering dependency between templates. Finally, a spatio-temporal aggregated attention (STAA) module is investigated to capture the spatial and dynamic relationships between functional regions, which solves the limitation of insufficient temporal information utilization. We evaluate our proposed method on 442 samples from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, and achieve the accuracies of 90.1%, 90.3%, 83.3% for normal control (NC) vs. early MCI (EMCI), EMCI vs. late MCI (LMCI), and NC vs. EMCI vs. LMCI classification tasks, respectively, indicating that our method boosts MCI identification performance and outperforms state-of-the-art methods.**

***Index Terms*— Personalized functional connectivity, MCI, graph neural network, spatio-temporal attention, functional MRI.**

Manuscript received 9 October 2022; revised 18 February 2023 and 23 March 2023; accepted 16 April 2023. Date of publication 27 April 2023; date of current version 12 May 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62201023 and Grant U1809209; in part by the Beijing Natural Science Foundation under Grant Z220017, Grant 4172037, and Grant CRIBJQY202103; in part by the Changping Laboratory under Grant 2021B-01-01; and in part by the Zhejiang Laboratory's International Talent Fund for Young Professionals. *(Weigang Cui and Yulan Ma contributed equally to this work.) (Corresponding authors: Yang Li; Hesheng Liu.)*

Weigang Cui is with the School of Engineering Medicine, Beihang University, Beijing 100191, China (e-mail: cwg1994@buaa.edu.cn).

Yulan Ma and Yang Li are with the Department of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China (e-mail: mayulan@buaa.edu.cn; liyang@buaa.edu.cn).

Jianxun Ren and Hesheng Liu are with the Changping Laboratory, Beijing 100094, China (e-mail: davidren555@outlook.com; heshseng@nmr.mgh.harvard.edu).

Jingyu Liu is with the School of Medical Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: liujingyu@bit.edu.cn).

Guolin Ma is with the Department of Radiology, China-Japan Friendship Hospital, Beijing 100029, China (e-mail: maguolin1007@qq.com).

Digital Object Identifier 10.1109/TNSRE.2023.3271062

I. Introduction

MILD cognitive impairment (MCI), as a prodromal stage of Alzheimer's disease (AD), is a syndrome defined as cognitive decline greater than that expected for an individual's age and education level [1]. In every year, more than 10% MCI patients suffer from irreversible brain atrophy and convert to AD [2]. In order to decease the high risk of progression, there is a clinical demand to accurately diagnosis MCI patients [3], so that early behavioral interventions and pharmacological treatments can be applied to delay the dementia expression and preserve brain cognitive functions. However, due to the mild symptoms and complex disease pathology of MCI, most existing methods based on pathology or anatomical brain images could not distinguish MCI from healthy aging effectively [4], [5].

Recent studies have shown that the functional connectivity (FC) networks based on resting-state functional magnetic imaging (rs-fMRI) is reliable and sensitive in identifying MCI patients [6], [7]. Most of existing methods constructs FC networks for feature extraction from a pre-defined group-averaged anatomical or functional template, such as AAL template [8], the Power Atlas [9] and the Glasser Atlas [10]. However, accumulating evidence from cortical stimulation mapping and neuroimaging studies indicates that individual differences in FC were heterogeneous across cerebral cortex [11], [12]. FC networks derived from a group-level brain template only reflect general principles and may miss important network features that are evident within individuals [13]. Furthermore, abnormal FC strengths that were used to identify MCI patients may attribute to the misaligned functional regions if they are not tightly linked to the macroscopic anatomy [14]. Therefore, substantial inter-subject functional variations may degrade the robustness of MCI identification. To improve the model performance, it is better to identify functional regions in individuals and capture the idiosyncrasies of subjects. Emerging

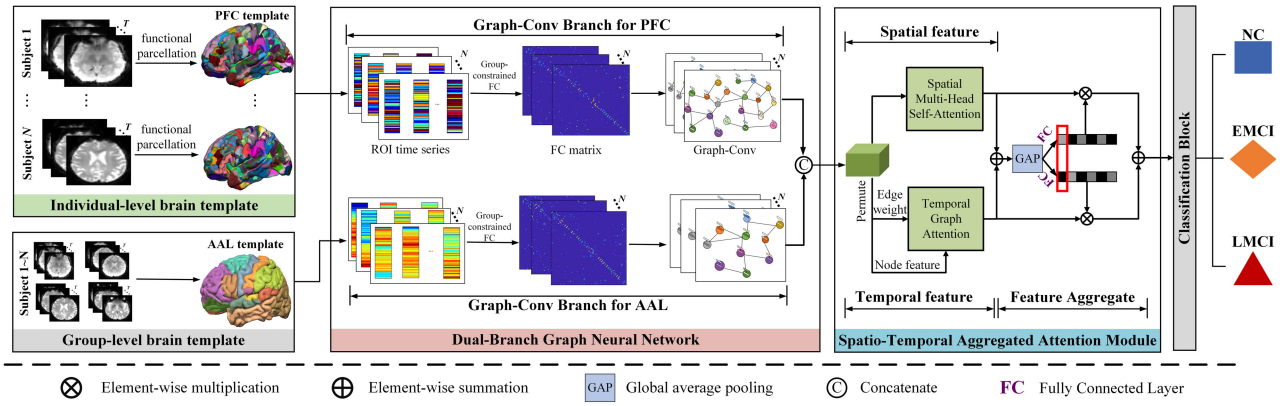

Fig. 1. The schematic diagram of the proposed framework.

individual-specific parcellation approaches provide an effective way to map functional organization at the individual level, but these methods cannot be directly used in MCI identification. *First*, the individual-level cortical templates only identifies 17 large-scale functional networks [15], [16], which may miss important functional features in averaging rs-fMRI time series. *Second*, some individual-level functional studies only concentrate on local brain regions, and cannot construct a FC network for the whole cerebral cortex [13]. Therefore, an individual-level cortical template containing fine-grained functional ROIs is required to facilitate the discovery of meaningful functional biomarkers for MCI identification.

In addition, in order to avoid potential bias associated with one single template, several MCI identification studies attempted to extract multiple feature sets based on different templates. For example, Liu et al. identified MCI patients by using feature representations derived from multiple brain templates determined by the Affinity Propagation clustering algorithm [17]. Lei et al. adopted multiple brain parcellation templates with different sets of ROIs to extract local and global information [18]. Compared with the frameworks based on single template, multi-template methods can provide distinct yet complementary information to identify MCI patients and boost classification accuracy. However, these studies fused the original or regularized feature sets generated from multiple templates by linear concatenation, which neglect the underlying functional relationships among templates. Intuitively, modeling the inter-template interaction information could bring more comprehensive features and thus improve classification performance.

Apart from the brain templates, some deep neural network models are drawing increasing attention in fMRI classification. These methods including graph convolutional network (GCN) [19], graph attention (GAT) model [20], and Transformer [21], can generate discriminative representations for brain FC networks, which provide more powerful classification ability compared to traditional machine learning models. However, fMRI studies using above methods mainly concentrate on spatial correlations between different brain regions. Recent studies reported that abnormal interactions in an MCI patient's brain involve not only spatial connectivity but also temporal dynamics [22], [23]. Therefore, a novel deep learning strategy, which takes advantage of spatial and temporal features in fMRI simultaneously, is highly required for MCI identification.

Inspired by the discussions aforementioned, we propose a novel personalized functional connectivity based dual-branch graph neural network with spatio-temporal aggregated attention (PFC-DBGNN-STAA) for MCI classification. Specifically, considering the fact that the group-averaged templates neglect the functional variability across subjects, a personalized functional connectivity (PFC) template is constructed by an iterative strategy, which aligns 213 cortical ROIs based on the FC patterns estimated in personalized fMRI. Then, a dual-branch graph neural network (DBGNN) is employed to fuse rs-fMRI time series and FC features generated from individual- and group-level templates, which utilizes the dependences between templates to boost the robustness of fMRI features. Finally, we characterize the temporal and spatial features by using a spatio-temporal aggregated attention (STAA) module, which employs a multi-head self-attention network to emphasize the spatial correlations among ROIs, and applies a dynamic graph attention mechanism to capture the temporal changes in local regions. The generated spatial and temporal features are aggregated for different MCI identification tasks. We perform classification experiments for our proposed method on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, achieving more satisfactory identification performance against state-of-the-art methods.

The main contributions of this paper are summarized below:

1) A PFC template is developed to align functional ROIs using personalized rs-fMRI data, which offers a complementary connectivity-based functional localizer and provides important individual-specific functional features for MCI identification.

2) We propose a DBGNN to concurrently learn individual-specific and group-level functional features, which contains the correlations between two templates and is beneficial to improve the identification performance.

3) A STAA module is introduced to emphasize spatial and temporal relationships of rs-fMRI features, which aggregates spatial correlations and dynamic changes to provide more powerful feature representations.

II. Methods

The proposed MCI classification framework is outlined in Fig. 1, and summarized as follows: 1) Apply the proposed individual-specific parcellation approach to obtain a PFC template for each sample; 2) employ the DBGNN to fuse the

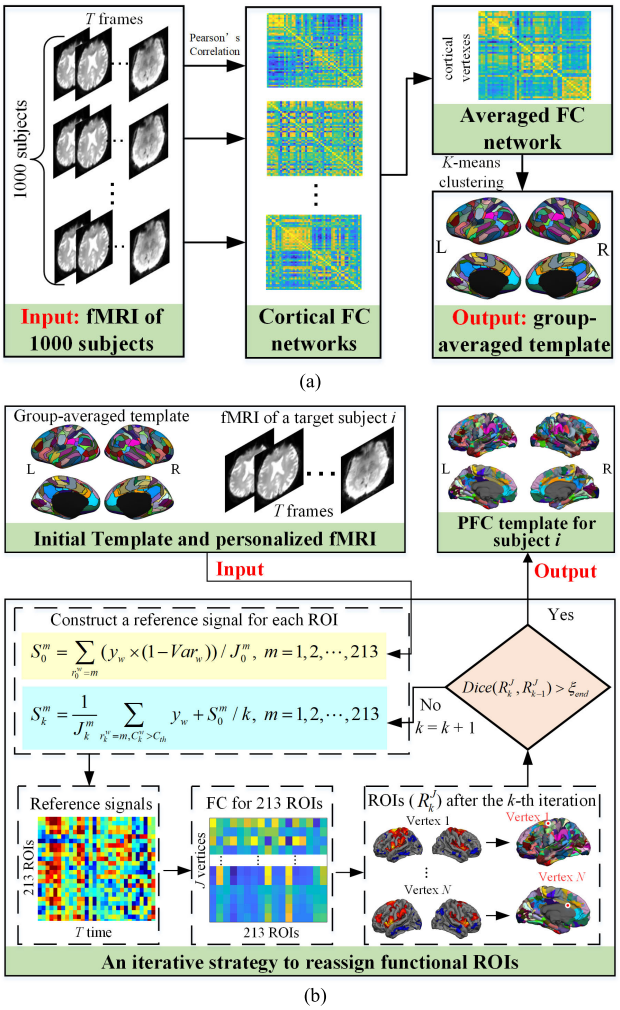

Fig. 2. The construction of (a) the group-averaged functional template and (b) the PFC template.

time series and FC features from the individual- and group-level templates; 3) construct a STAA module to map high-level spatial and temporal features, which is finally used for MCI identification. The details of our method are given in the following subsections.

A. The Construction of Personalized Functional Template

In this study, a novel functional parcellation approach for constructing a PFC template is applied for each subject, and the process of is given in Fig. 2. Specifically, we first construct a group-averaged functional template based on 1,000 healthy subjects, and then individualize the functional template on each subject according to its own fMRI data.

1) The Construction of the Group-Averaged Functional Template: As shown in Fig. 2(a), a group-averaged functional template with 213 fine-grained functional ROIs across the cerebral cortex is created based on the FC profile of 1,000 healthy subjects (https://www.nitrc.org/projects/gspdata) [24]. Specifically, with the fMRI data of 1,000 samples as input, the FC networks of cortical vertexes are calculated by Pearson's correlation for each sample. Then, to get a fine-grained parcellation, we employ a k-means clustering approach on the group-averaged FC networks to parcellate the cortical vertexes into 213 functional ROIs [25]. Specifically, the number of the ROIs in left hemisphere is defined as 108, and the right is 105.

The numbers of the clusters $k = 108$ and $k = 105$ are determined by a stability analysis as follows. First, the 1,000 healthy subjects are randomly and evenly divided into two groups. Second, for each hemisphere, with the number of the clusters k ranging from 20 to 150, the averaged cortical FC networks of each group are inputted into a k-means clustering for a parcellation. Finally, we measure the similarity of the parcellation templates between these two group by calculating the Dice coefficient [26]. We repeat this procedure 100 times by changing the division of subjects, and determine the k-value according to the maximum of the mean Dice coefficients.

2) The Construction of the PFC Template: As shown in Fig. 2(b), an iterative parcellation strategy is conducted based on the obtained group-averaged template to derive a PFC template for each subject. Take the rs-fMRI data of a target subject i as input, the strategy is able to individualize the location of each functional ROI in the group-averaged template according to the personalized FC and produce an PFC template. The details are introduced as follows.

First, we define a reference signal for each functional ROI. Specifically, suppose there are J vertices, the parcellation index from the initial template is $\boldsymbol{R}_0^J = \left[r_0^1, r_0^2, \cdots, r_0^J\right]$, and $r_0^j \in [1, 2, \cdots, 213]$ is the initial label of the corresponding functional ROI for the j-th vertex. Set the rs-fMRI time series for the J vertices as $\left[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_J\right]$, the reference signal $\boldsymbol{S}_0^m$ for the m-th ROI can be defined as the mean signal of all vertices in it:

$$\boldsymbol{S}_0^m = \sum\nolimits_{r_0^w = m} (\boldsymbol{y}_w \times (1 - Var_w))/J_0^m \tag{1}$$

where J_0^m is the number of vertices in the m-th ROI, and Var_w is the inter-subject functional variability for the w-th vertex introduced in [12]. The weight $(1 - Var_w)$ for each vertex is used to reduce the impact of vertex whose FC profile is highly variable in different samples.

Second, we reassign the locations of the 213 functional ROIs. Specifically, the fMRI signal at each vertex is correlated with the reference signals of 213 functional ROIs, and each vertex can be reassigned to one of the 213 functional ROIs according to its maximal correlation to the reference signals. In this way, the new label r_1^j of the j-th vertex is given as:

$$r_1^j = argmax\left(corr\left(\boldsymbol{y}_j, \boldsymbol{S}_0^m\right)\right), m \in [1, 2, \cdots, 213] \tag{2}$$

where $corr(\cdot, \cdot)$ represents the Pearson's correlation between two time series. In addition, we define a confidence value C_1^j for the j-th vertex as the ratio between the largest and the second-largest correlation value. Until now, the first iteration has finished, and we can obtain a temporary template whose ROIs have been reassigned once.

For the subsequent iteration procedure, the reference signals of 213 ROIs need to be recalculated as follows. At the k-th iteration, the functional parcellation is $\boldsymbol{R}_k^J = \left[r_k^1, r_k^2, \cdots, r_k^J\right]$, and the confidence values for all vertices are $C_k^1, C_k^2, \cdots, C_k^J$. In the m-th functional ROI, suppose there are J_k^m core vertices whose confidence value greater than a pre-defined threshold

C_{th}. New reference signal of this ROI can be defined as a weighted average between the signals of J_k^m core vertices and the initial reference signal:

$$S_k^m = \frac{\sum_{r_k^w=m, C_k^w > C_{th}} y_w}{J_k^m} + S_0^m / k \tag{3}$$

Based on the new reference signals, which incorporate the information of both the individualized FC and the initial functional template, all the cortical vertices are further correlated with them and reassigned to one of the 213 functional ROIs. With the increasing number of iterations, the resulting functional ROIs for each subject will converge to be consistent, and functional parcellations will vary substantially across different subjects. The iteration procedure can be stopped if the membership for each ROI remained the same for a predefined ratio ξ_{end} of the vertices. In this way, the PFC template of the target subject i is obtained, where the distribution of the 213 functional ROIs have been individualized.

B. Dual-Branch Graph Neural Networks

A PFC template is obtained by the proposed individual-specific parcellation method, which generates totally 213 functional ROIs across the cerebral cortex. However, it does not fully consider the anatomical brain parcellation, especially subcortical regions. To supplement structural information, the widely used anatomical parcellation, i.e., AAL template [8], is cooperated with the PFC template to extract FC features for the whole brain. It defines 45 anatomical volumes of interest in each hemisphere, i.e., 39 cortical and 6 subcortical regions.

To consider the dependencies between these two templates, we propose a DBGNN to capture individual- and group-level information simultaneously. It employs two branches with synchronized Graph-Conv and a shared weight matrix to comprehensively map brain FC patterns and fMRI time series from different templates to brain activities. The illustration of DBGNN is given in Fig. 3.

Specifically, the PFC and AAL branches are used to characterize rs-fMRI time series and FC networks generated from the two templates. Inspired by associated works [27], [28], we employ a group-constrained sparsity method to construct FC networks for the two templates, because it can provide a sparse FC network that removes insignificant or spurious connections. Specifically, denote time series of the PFC and AAL templates for a sample as $\boldsymbol{X}_{PFC} \epsilon R^{213\times T}$ and $\boldsymbol{X}_{AAL} \epsilon R^{90\times T}$, to obtain the corresponding FC networks, the sparse brain FC modeling of the m-th ROI can be considered to minimize the loss function defined as follows:

$$f(\boldsymbol{a}_m) = \sum_{n=1}^{N} \left(\frac{1}{2} \left\| \boldsymbol{x}_m^n - \boldsymbol{X}_m^n \boldsymbol{a}_m^n \right\|_2^2 + \lambda \left\| a_m \right\|_{2,1} \right. \tag{4}$$

where $\boldsymbol{a}_m = [\boldsymbol{a}_m^1, \boldsymbol{a}_m^2, \cdots, \boldsymbol{a}_m^N]$ is the FC coefficients for the m-th ROI, $\boldsymbol{x}_m^n$ is the corresponding time series for the n-th sample, and $\boldsymbol{X}_m^n = [\boldsymbol{x}_1^n, \cdots, \boldsymbol{x}_{m-1}^n, \boldsymbol{x}_{m+1}^n, \boldsymbol{x}_M^n]$ is the data matrix which consists of all time series except for the m-th ROI. $\|\cdot\|_{2,1}$ is the summation of l_2-norms, and λ is a regularization hyper-parameter which controls

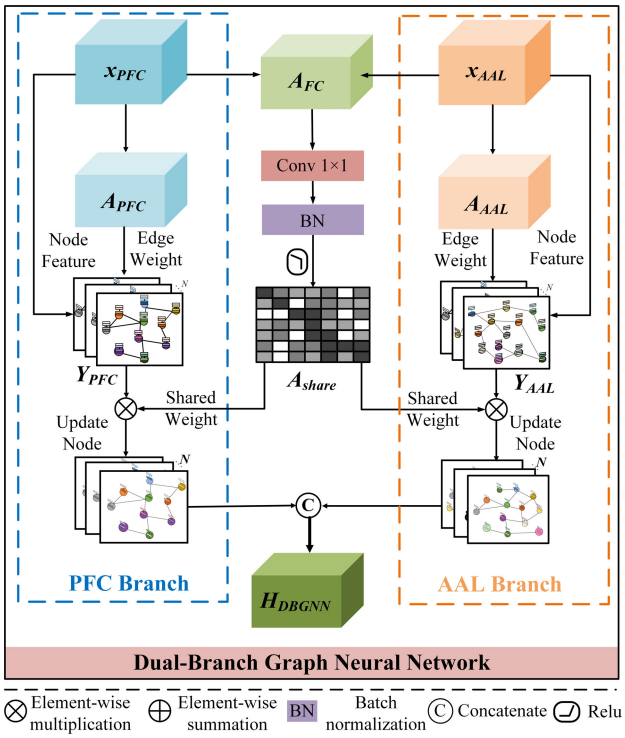

Fig. 3. The illustration of the DBGNN.

the sparsity of the model. We keep the signs of the constructed FC networks for the PFC templates $\boldsymbol{A}_{PFC} = [\boldsymbol{a}_{1,PFC}, \boldsymbol{a}_{2,PFC}, \cdots, \boldsymbol{a}_{213,PFC}] \epsilon R^{213\times213}$ and the AAL templates $\boldsymbol{A}_{AAL} = [\boldsymbol{a}_{1,AAL}, \boldsymbol{a}_{2,AAL}, \cdots, \boldsymbol{a}_{90,AAL}] \epsilon R^{90\times90}$, and then input all the FCs to the following steps.

To fully capture the FC information among brain ROIs while effectively highlighting the local time series, GCN is employed for each branch due to its superiority in learning high-level graph features [29]. In a graph, each node has a corresponding characteristic vector, which is represented as the rs-fMRI time series for a brain ROI. The connection between the nodes in the graph denotes the FC between a pair of ROIs. As an adjacency matrix, the FC network captures the topological correlations among brain ROIs. At this point, a graph data for each branch with 213 or 90 nodes, where each node contains a feature vector with a dimension of $T \times 1$, has been constructed. The results output by the GCNs in two branches are $\boldsymbol{Y}_{PFC} \epsilon R^{213\times D}$ and $\boldsymbol{Y}_{AAL} \epsilon R^{90\times D}$, respectively, where D represents the length of output meaningful time series for each node.

Additionally, in order to mutually map individual-specific and group-level FC features, a FC network $\boldsymbol{A}_{FC} \epsilon R^{213\times90}$ is constructed for each sample by correlating time series $\boldsymbol{X}_{PFC} \epsilon R^{213\times T}$ and $\boldsymbol{X}_{AAL} \epsilon R^{90\times T}$. Then we feed it into a standard convolution with kernel of 1×1, batch normalization, and a rectified linear unit (ReLU), where the convolutional layer can learn the functional weight of brain regions without changing $\boldsymbol{A}_{FC}$ spatial correlation [30]. This strategy not only simulates the dependencies between different brain templates, but also enhances the differences between brain regions. As a result, a non-negative cross-template shared weight matrix $\boldsymbol{A}_{share} \epsilon R^{213\times90}$ is achieved. Next, the $\boldsymbol{A}_{share}$ is multiplied to

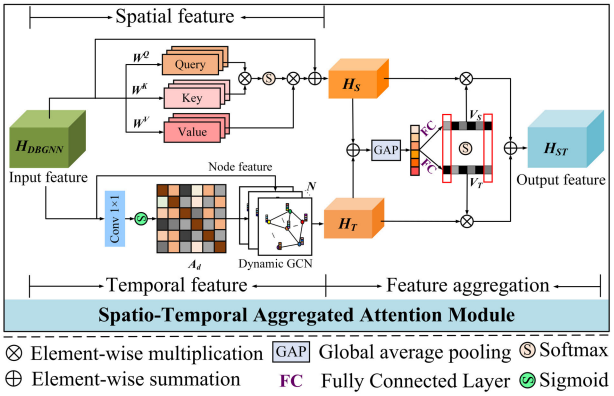

Fig. 4. The illustration of the STAA module.

the outputs of two GCNs (i.e., $\boldsymbol{Y}_{PFC}$ and $\boldsymbol{Y}_{AAL}$) for updating node representations. As a learnable matrix, the $\boldsymbol{A}_{share}$ is updated and optimized according to the classification performance. Finally, the generated representations are concatenated as the high-level feature $\boldsymbol{H}_{DBGNN} \epsilon R^{D\times 303}$, which contains the dependency information between individual- and group-level templates, and can be provided to the following layers to effectively learn spatial and temporal information.

C. Spatio-Temporal Aggregated Attention Module

The rs-fMRI time series contains temporal information reflecting dynamic brain activities. However, most existing methods only concentrate on the spatial connectivity among different brain regions, and ignore the temporal features. To overcome the limitation, a spatio-temporal aggregated attention (STAA) module is proposed to characterize spatial and temporal features simultaneously, and the illustration is outlined in Fig. 4. The high-level representations $\boldsymbol{H}_{DBGNN}$ obtained by DBGNN is regarded as the input feature of STAA. The size of the input is $D\times 303$, where D denotes the length of time series, and 303 denotes the total number of ROIs.

For spatial feature extraction, we adopt a self-attention mechanism to emphasize the spatial correlations between brain ROIs. The framework includes a query transform matrix $\boldsymbol{w}^{\boldsymbol{Q}}$, a key transform matrix $\boldsymbol{w}^{\boldsymbol{K}}$ and a value transform matrix $\boldsymbol{w}^{\boldsymbol{V}}$. The outputs of these transforms, i.e., the query $\boldsymbol{Q}$, key $\boldsymbol{K}$, and value $\boldsymbol{V}$, can be defined as:

$$\boldsymbol{Q} = \boldsymbol{H'}_{DBGNN}\boldsymbol{w}^{\boldsymbol{Q}} \tag{5}$$

$$\boldsymbol{K} = \boldsymbol{H'}_{DBGNN}\boldsymbol{w}^{\boldsymbol{K}} \tag{6}$$

$$\boldsymbol{V} = \boldsymbol{H'}_{DBGNN}\boldsymbol{w}^{\boldsymbol{V}} \tag{7}$$

where $\boldsymbol{H'}_{DBGNN}$ is the transposition of $\boldsymbol{H}_{DBGNN}$, and $\boldsymbol{w}^{\boldsymbol{Q}}, \boldsymbol{w}^{\boldsymbol{K}} \in R^{D\times 303}$, $\boldsymbol{w}^{\boldsymbol{V}} \in R^{D\times D}$. The output of the attention layer can be calculated as:

$$\boldsymbol{H}_S = [softmax\left(\frac{\boldsymbol{QK}}{\sqrt{213+90}}\right)\boldsymbol{V} + \boldsymbol{H'}_{DBGNN}]' \tag{8}$$

For temporal feature extraction, we update the time series by using dynamic GCN, where $\boldsymbol{H'}_{DBGNN}$ is set as the node feature. Different from the tradition GCN with a fixed adjacent matrix shared by all the input samples, the adjacent matrix $\boldsymbol{A}_d$ in our dynamic GCN is adaptively adjusted by the input as follows:

$$\boldsymbol{A}_d = \delta(\boldsymbol{W}_A\boldsymbol{H}_{DBGNN}) \tag{9}$$

where δ is sigmoid activation function, $\boldsymbol{W}_A \in R^{303\times D}$ is weight of a convolutional layer to formulate the dynamic adjacent matrix $\boldsymbol{A}_d$. It improves the representational power of the generated features for individual samples, and avoids the overfitting problem induced by a fixed adjacent matrix. In this way, the output $\boldsymbol{H}_T$ of the dynamic GCN can be defined as:

$$\boldsymbol{H}_T = [LRelu(\boldsymbol{A}_d\boldsymbol{H'}_{DBGNN}\boldsymbol{W}_d)]' \tag{10}$$

where $LRelu$ is LeakyReLU activation function, $\boldsymbol{W}_d \in R^{D\times D}$ is the state-update weight.

With the above features, soft attention is applied to adaptively select different spatio-temporal aggregation information, which is guided by the global average pooling layer $\mathcal{F}_{gp}$. Specifically, a softmax operator is applied after feature concatenation of two fully connected layers as follows:

$$[\boldsymbol{V}_S, \boldsymbol{V}_T] = split(\varphi\left(\mathcal{F}^1_{fc}\mathcal{F}_{gp}(\boldsymbol{H}_S + \boldsymbol{H}_T)\right) \times \|\mathcal{F}^2_{fc}\mathcal{F}_{gp}(\boldsymbol{H}_S + \boldsymbol{H}_T))) \tag{11}$$

where $\boldsymbol{V}_S, \boldsymbol{V}_T$ denote the soft attention vector by a softmax operator for $\boldsymbol{H}_S$ and $\boldsymbol{H}_T$, respectively. F^1_{fc} and F^2_{fc} denote the fully connected layers, $||$ is concatenation operation, φ denotes softmax operator, and $split$ means to separate the vectors spliced by the two fully connected layers. Then, the spatio-temporal feature $\boldsymbol{H}_{ST}$ can be obtained by:

$$\boldsymbol{H}_{ST} = \boldsymbol{V}_S \times \boldsymbol{H}_S + \boldsymbol{V}_T \times \boldsymbol{H}_T \tag{12}$$

where $\boldsymbol{V}_S$ and $\boldsymbol{V}_T$ are soft attention vectors defined in Eq. (11), $\boldsymbol{H}_S$ and $\boldsymbol{H}_T$ are high-level spatial and temporal features defined in Eq. (8) and (10), respectively. Finally, as shown in Fig. 1, the classification result is achieved after processing the output of STAA by means of two fully-connected layers and the softmax function.

III. Experiments and Results

A. Data Acquisition and Preprocessing

The rs-fMRI data used in this study are selected from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu/), which is a large-scale collaborative study to explore neuroimaging biomarkers for MCI and early AD. We exclude rs-fMRI scans with excessive head motion (mean relative motion $>$ 0.15 mm). The remained dataset includes 154 NC, 168 early MCI (EMCI), and 120 late MCI (LMCI). Data acquisition was performed using 3 Tesla Philips scanners with the following parameters: TR = 3,000 ms; TE = 30 ms; flip angle = 80 degrees; 140 volumes; matrix size = 64 × 64; 48 slices; voxel thickness = 3.4 mm.

The rs-fMRI data are processed using the same procedures described in the previous studies [16], [31]. In brief, the following steps are performed: slice timing correction (SPM12 software), rigid-body correction for head motion with the FSL package, normalization for global mean signal intensity across runs, and band-pass temporal filtering (0.01-0.08 Hz).

TABLE I
CLASSIFICATION PERFORMANCE COMPARISON OF DIFFERENT TEMPLATES AND FEATURE EXTRACTION MODELS

Template	Feature Extraction	NC vs. EMCI			EMCI vs. LMCI			NC vs. EMCI vs. LMCI				
		ACC (%)	AUC	SEN (%)	ACC (%)	AUC	SEN (%)	ACC (%)	AUC	SEN_N (%)	SEN_E (%)	SEN_L (%)
AAL	GCN	74.8±1.5**	0.82±0.05**	78.0±1.8**	75.0±1.6**	0.79±0.02**	62.5±2.2**	70.4±2.1**	0.81±0.04*	66.2±1.3**	78.6±1.3**	64.2±1.0**
	SSAT	75.2±0.8**	0.83±0.02**	75.6±0.7**	75.7±0.7**	0.77±0.05**	60.8±1.2**	70.8±0.9**	0.77±0.05**	57.8±1.1**	86.9±0.9**	65.0±0.7**
	TGAT	75.8±1.2**	0.80±0.03**	76.2±1.2**	76.4±1.1**	0.77±0.03**	69.2±1.1**	70.8±1.5**	0.83±0.03*	66.2±0.9**	71.4±1.5**	75.8±0.6*
	STAA	78.3±0.5**	0.84±0.04**	79.2±0.7**	77.8±1.0**	0.81±0.04**	67.5±0.6**	72.4±1.3**	0.83±0.04*	65.6±1.2**	82.7±1.1**	66.7±0.9**
PFC	GCN	73.0±2.7**	0.78±0.07**	70.2±1.7**	72.2±2.4**	0.72±0.06**	50.8±3.2**	69.2±1.5**	0.81±0.05**	79.8±1.3*	65.5±1.7**	60.8±1.0**
	SSAT	77.3±1.5**	0.81±0.05**	82.7±1.5**	76.0±1.4**	0.79±0.06**	61.7±2.6**	70.6±1.6**	0.81±0.04**	72.1±1.4**	74.4±2.0**	63.3±1.2**
	TGAT	78.9±1.8**	0.84±0.02**	84.5±2.9**	78.8±1.1**	0.80±0.04**	70.8±2.4**	73.3±1.3**	0.82±0.05*	72.7±2.0**	70.8±2.2**	77.5±1.0*
	STAA	80.8±0.9**	0.85±0.02**	83.9±0.9**	80.2±0.7**	0.82±0.03**	70.0±2.1**	74.2±0.6**	0.84±0.02*	71.4±0.8**	81.6±1.0**	67.5±1.5**
AAL + PFC	GCN	80.4±3.0**	0.86±0.04**	85.7±3.1**	83.0±1.4**	0.85±0.05*	73.3±1.5**	76.7±2.0**	0.83±0.03*	73.4±2.0**	79.8±1.6**	76.7±1.0*
	SSAT	81.7±1.2**	0.87±0.02**	84.5±1.6**	82.3±1.6**	0.86±0.03*	74.2±1.8**	76.2±1.1**	0.81±0.07*	76.6±1.4**	78.0±1.3**	73.3±0.9**
	TGAT	87.3±2.2*	0.90±0.02*	90.5±2.0*	87.2±2.1*	0.88±0.02*	80.0±1.7**	80.5±1.8*	**0.88±0.04**	76.6±1.7**	85.1±1.2**	**79.2±0.5**
	STAA	**90.1±0.3**	**0.90±0.03**	**92.9±1.1**	**90.3±0.3**	**0.90±0.04**	**84.2±1.2**	**83.3±0.3**	**0.88±0.04**	**80.5±0. 9**	**89.3±1.0**	78.3±1.1

where the numbers marked bold indicate the best performance, and the symbols * and ** represent statistically significant difference compared to the proposed method with $p < 0.05$ and $p < 0.001$, respectively.

B. Experimental Settings and Evaluation

The proposed MCI classification framework is implemented on the Pytorch platform through NVIDIA Tesla V-100 GPU. Our model is trained by the Adam optimizer [32] with a learning rate of 0.001. The maximum epoch is initiated as 1000, and the batch size of the training set is 16. The model is optimized by using the cross-entropy loss between the predicted labels and actual labels of training samples.

In our experiments, we conduct two binary classification tasks and a three-class classification task to evaluate the effectiveness and robustness of the proposed method, including 1) NC vs. EMCI, 2) EMCI vs. LMCI, and 3) NC vs. EMCI vs. LMCI. The generalization performance is obtained via a nested 10-fold cross-validation as our previous study [22], [33], which is proven to be a robust unbiased performance evaluation and hyper-parameter optimization method [34]. Specifically, a nested cross-validation consists of an outer loop and an inner loop. For each outer loop, we set one-fold of samples as testing set; the remaining 9-fold of samples are inputted into the inner loop, and further divided into a training set and a validation set by using another 10-fold cross-validation. In the inner loop, the hyper-parameters, e.g., the confidence threshold C_{th}([1.5, 1.6, $\cdots$, 3.0]), and the pre-defined ratio ξ_{end} ([0.90, 0.92, $\cdots$, 0.98]) to end the iteration strategy of functional parcellation are optimized by using a grid search strategy based on the inner 10-fold cross-validation, which uses the inner training set and validation set for the training and testing process. In order to obtain reliable evaluation results, the nested scheme is repeated 10 times by changing the sample division. Accuracy (ACC), sensitivity (SEN) and area under curve (AUC) are adopted for the performance evaluation [35], where SEN_N, SEN_E and SEN_L represent the sensitivities in identify NC, EMCI and LMCI samples, respectively. In addition, to demonstrate the superior performance of our proposed method, we perform one-way analysis of variance (ANOVA) statistical tests on the 10 times of 10-fold identification results between each comparing method and the PFC-DBGNN-STAA framework.

C. Overall Performance

We evaluate the proposed MCI identification method on the ADNI dataset and compare it with different brain templates and feature extraction models. The group-constrained approach is adopted to construct FC networks for all templates, and the comparison is conducted under the same experimental conditions. The fusion features based on the PFC and AAL templates are compared with the following brain templates: 1) only a group-level anatomical template, i.e., AAL template; 2) only the PFC template introduced in Section II-A. The comparing feature extraction models includes: 1) graph convolutional network (GCN); 2) spatial self-attention (SSAT) network; 3) temporal graph attention (TGAT) network.

From the classification performance of all comparison frameworks given in Table I, we can observe three points. *First*, with the same feature extraction modules, the dual-branch framework generally achieves better performance than single-template methods (only AAL or PFC template). Specifically, the highest ACC achieved by single-template methods is only 80.8%, 80.2% and 74.2% in the three classification tasks, which are noticeably lower than the proposed methods. *Second*, the proposed STAA shows its powerful ability to extract discriminative spatial and temporal features from FC networks and fMRI time series. It can be used with different templates and gain ACC improvements than the single-domain attention model (SSAT and TGAT) in the three classification tasks, which proves the advantage of incorporating the spatial and temporal attention modules for obtaining more comprehensive feature representations. *Finally*, the proposed method yields the highest ACCs of 90.1%, 90.3% and 83.3% for NC vs. EMCI, EMCI vs. LMCI, and NC vs. EMCI vs. LMCI classification tasks, respectively. Compared with other frameworks, our method achieves at least 2.8%, 3.1% and

TABLE II
COMPARISON OF ABLATION RESULTS IN THREE CLASSIFICATION TASKS

Feature Extraction	NC vs. EMCI			EMCI vs. LMCI			NC vs. EMCI vs. LMCI				
	ACC (%)	AUC	SEN (%)	ACC (%)	AUC	SEN (%)	ACC (%)	AUC	SEN_N (%)	SEN_E (%)	SEN_L (%)
AAL+GCN	74.8±1.5**	0.82±0.05**	78.0±1.8**	75.0±1.6**	0.79±0.02**	62.5±2.2**	70.4±2.1**	0.81±0.04*	66.2±1.3**	78.6±1.3**	64.2±1.0**
AAL+PFC+GCN	80.4±3.0**	0.86±0.04*	85.7±3.1**	83.0±1.4**	0.85±0.05*	73.3±1.5**	76.7±2.0**	0.83±0.03*	73.4±2.0**	79.8±1.6**	76.7±1.0*
AAL+PFC+DBGNN+GCN	86.3±1.4**	0.88±0.02*	85.1±2.6**	85.1±1.3**	0.88±0.03*	80.0±2.1**	80.1±1.7**	0.87±0.04*	79.2±1.5*	82.1±1.2**	78.3±0.9*
AAL+PFC+DBGNN+STAA	**90.1±0.3**	**0.90±0.03**	**92.9±1.1**	**90.3±0.3**	**0.90±0.04**	**84.2±1.2**	**83.3±0.3**	**0.88±0.04**	**80.5±0.9**	**89.3±1.0**	**78.3±1.1**

where the numbers marked bold indicate the best performance, and the symbols * and ** represent statistically significant difference compared to the proposed method with $p < 0.05$ and $p < 0.001$, respectively.

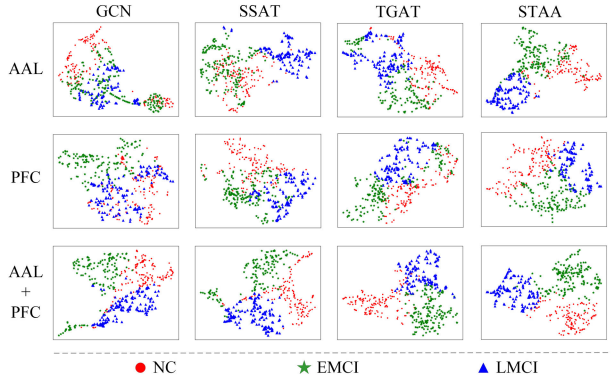

Fig. 5. The t-SNE visualization in 2-D embedding space of features learned by different templates and aggregation modules.

2.8% improvements on accuracy, indicating the promising classification ability. Meanwhile, it outperforms all compared frameworks by yielding the highest AUC and SEN (except SEN_L in NC vs. EMCI vs. LMCI task), proving that the method can robustly identify the positive and negative samples simultaneously.

Additionally, to further validate the superiority of our extracted features, we use the t-SNE visualization to present the features generated by the above methods into a 2D embedding space. From the t-SNE plots in Fig. 5, the feature visualizations from single template are relatively ambiguous in contrast to the dual-branch method. Compared with other aggregation methods, the STAA module generates more separable features. Therefore, our proposed methods using both AAL + PFC and STAA can efficiently distinguish fMRI data from different types of classification tasks.

IV. DISCUSSION

A. Ablation Study

To demonstrate the effectiveness of the components of the proposed MCI classification framework, we perform an ablation study on the ADNI dataset. In this section, the AAL template cooperated with GCN is regarded as the baseline method to explore the contribution of the three proposed modules, i.e., PFC template, DBGNN, and STAA. The competing frameworks in the ablation study includes:

1)**AAL + GCN**: We employ the AAL template to generate ROIs and extract fMRI time series, which are input into group-constrained method to obtain FC networks for each sample. A GCN module is used to aggregate FC features for the classification tasks [36].

2)**AAL + PFC + GCN**: Both the PFC and AAL templates are used to extract fMRI time series and construct FC networks. The features from the two templates are directly concatenated and then inputted into GCN module for MCI classification.

3)**AAL + PFC + DBGNN + GCN**: Different from the method above, the FC networks and fMRI time series from the two templates are fused by the proposed DBGNN, and further mapped to the final classification results by the GCN module.

AAL + PFC + DBGNN + STAA: The method is the proposed framework. Compared with the above method, the proposed STAA described in *Section* II-C is adopted to replace the GCN model for feature aggregation.

From the comparison shown in Table II, the efficacy of each element of the proposed framework is demonstrated as follows:

1) **Efficacy of the PFC**: As shown in Table II, after adding the features of the PFC to the baseline method, the classification performance has increased in terms of three evaluation metrics. For example, compared with the AAL + GCN method, the AAL + PFC + GCN method improves AUCs by 0.04, 0.06 and 0.02 for NC vs. EMCI, EMCI vs. LMCI and NC vs. EMCI vs. LMCI classification tasks. This is due to the fact that the AAL template may misses subtle FC features that are variable across individuals. The PFC template could assess the functional organization of an individual's brain according to its personalized FC, which provides complementary information to boost MCI classification performance.

2) **Efficacy of DBGNN**: To demonstrate the effectiveness of the DBGNN module, we add this module to the above method. Table II shows that this framework outperforms the method that concatenates individual- and group-level features directly. It improves ACCs by 5.9%, 2.1%, 3.4% in the three classification tasks, respectively. The advantage of the proposed DBGNN module is that it considers the correlations between individual- and group-level features, and thus extracts more enriched and comprehensive dual-branch information, which can effectively improve the feature discrimination of the MCI classification framework.

3) **Efficacy of STAA module**: Compared with the GCN module, the STAA module not only uses a self-attention mechanism to explore the spatial correlations, but also applies dynamic graph attention to capture the temporal relationship of fMRI time series. The spatial and temporal features are fused to generate high-level representations containing the

TABLE III
CLASSIFICATION PERFORMANCE COMPARISON WITH STATE-OF-THE-ART METHODS

Method	Year	NC vs. EMCI			EMCI vs. LMCI			NC vs. EMCI vs. LMCI				
		ACC (%)	AUC	SEN (%)	ACC (%)	AUC	SEN (%)	ACC (%)	AUC	SEN_N (%)	SEN_E (%)	SEN_L (%)
SSGSR	2020	80.1±1.7**	0.82±0.04**	82.5±1.4**	75.4±1.9**	0.79±0.09*	66.1±3.5**	71.5±2.3**	0.79±0.10*	77.9±0.3**	67.3±2.6**	69.2±2.7**
ADB-NN	2020	81.4±2.0**	0.84±0.07*	80.4±1.9**	82.6±2.3**	0.84±0.06*	74.2±2.4**	73.5±1.7**	0.83±0.03*	76.6±1.0**	70.8±2.1**	73.3±1.4**
ST-GCN	2020	82.3±1.1**	0.86±0.05*	83.3±1.7**	80.6±1.5**	0.86±0.03*	71.7±2.1**	75.6±0.7**	0.84±0.04*	**82.5±0.6**	72.0±0.8**	71.7±1.2**
VAT-CNN	2021	83.2±1.6**	0.83±0.05**	84.5±1.3**	80.9±2.7**	0.81±0.08**	67.5±3.3**	72.2±2.5**	0.81±0.07*	79.2±2.0*	69.1±2.9**	67.5±3.0**
SRiLT	2022	80.1±4.2**	0.85±0.06*	81.6±2.3**	84.4±2.4**	0.88±0.03*	75.0±2.9**	76.0±1.2**	0.85±0.05*	74.0±2.1**	77.4±1.5**	76.7±1.3*
Proposed	2023	**90.1±0.3**	**0.90±0.03**	**92.9±1.1**	**90.3±0.3**	**0.90±0.04**	**84.2±1.2**	**83.3±0.3**	**0.88±0.04**	80.5±0.9	**89.3±1.0**	**78.3±1.1**

where the numbers marked bold indicate the best performance, and the symbols * and ** represent statistically significant difference compared to the proposed method with $p < 0.05$ and $p < 0.001$, respectively.

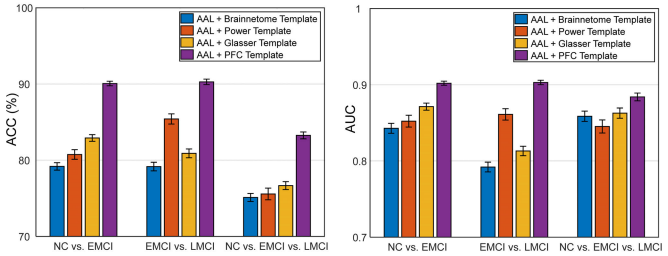

Fig. 6. The comparison of (a) ACC and (b) AUC with standard deviation between different dual-branch methods.

information of the fMRI time series and the relevance of FC networks, simultaneously. The classification results shown in Table II verifies the effectiveness of STAA module, which improves ACC (86.3%/85.1%/80.1% to 90.1%/90.3%/83.3%) and AUC (0.88/0.88/0.87 to 0.90/0.90/0.88) in the three MCI classification tasks.

B. Efficacy of Feature Fusion of Individual- and Group-Level Templates

To prove the importance of the feature fusion of our individual- and group-level brain templates, we replace the proposed PFC template with some commonly used group-averaged templates, including the Brainnetome template (246 regions) [37], the Power template (264 regions) [9] and the Glasser template (360 regions) [10]. These three templates are fused with the AAL template by our DBGNN and used for MCI identification, respectively. The classification results and the corresponding standard deviations are compared in Fig. 6. Compared with these group-level templates, our proposed PFC template could detect functional ROIs according to personalized fMRI and provide complementary individual-specific features for classification. Therefore, although less brain regions are used for feature extraction, our proposed framework based on PFC template improves ACCs by 7.1%, 4.9% and 6.6% in NC vs. EMCI, EMCI vs. LMCI and NC vs. EMCI vs. LMCI classification tasks, respectively, indicating that spatio-temporal features from the PFC template is beneficial to the classification performance.

C. Performance Comparison With State-of-the-Art Methods

To demonstrate the superiority of our proposed method, we compare it with several state-of-the-art rs-fMRI classification methods, which include:

Strength and similarity guided group sparse representation(SSGSR): Both fMRI signal temporal correlation and inter-subject similarity of FC are used to guide the GSR-based network modeling. The connection coefficients of FC networks are inputted into a SVM classifier for classification [38].

Attention-Diffusion-Bilinear Neural Network(ADB-NN): An ADB-NN framework including a data preprocessing module, attention-diffusion-bilinear neural network module and the decision model is used for brain network analysis and classification [39].

Spatio-temporal graph convolutional network (ST-GCN): A ST-GCN is trained on rs-fMRI time series to model the non-stationary nature of FC, and the importance of graph edges within ST-GCN is learned for classification [40].

Dynamic effective connectivity with a virtual adversarial training convolutional neural network(VAT-CNN): The local features of dynamic effective connectivity extracted by a VAT-CNN is inputted into a weight-guided graph attention networks for feature aggregation and classification [20].

Sparse representation with latent temporal dependency (SRiLT): The brain functional networks are constructed by the SRiLT estimation method. The connection weights between different ROIs are used as features for MCI classification tasks [7].

Because the sample sizes of the existing research are quite different from this article, we reproduce the state-of-the-art classification methods on our processed dataset. For a fair comparison, all of the comparison results shown in Table III are based on a nested 10-fold cross-validation. The experimental environment is the same as our method. Compared with these existing studies, one advantage of our proposed method is that the PFC template injects an individual-specific FC network into feature sets, and another advantage is the utilization of high-level spatial and temporal representations. It can be observed that the proposed method achieves more excellent performance than the state-of-the-art methods. The classification ACC is 6.9%, 5.9%, and 7.3% higher than the

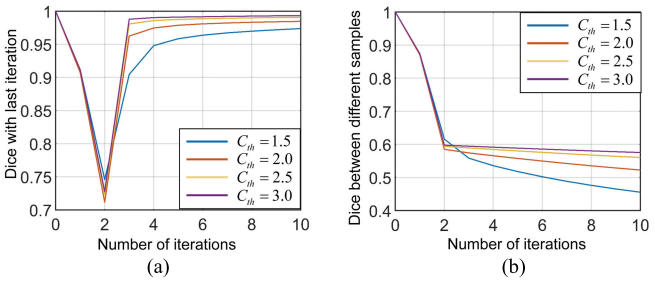

Fig. 7. The convergence of the parcellation algorithm with different C_{th} is evaluated by (a) averaged Dice with last iteration for each sample and (b) averaged Dice between different samples.

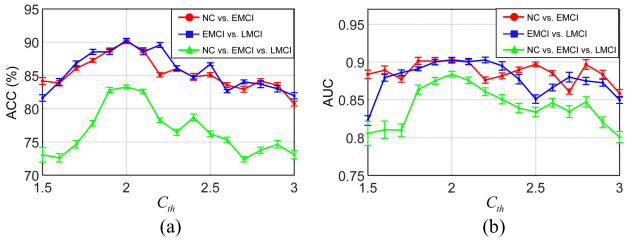

Fig. 8. The influence of the confidence threshold C_{th} on (a) ACC and (b) AUC with standard deviation for the three MCI classification tasks.

best results from other methods for the three classification tasks, respectively. The significant improvements on AUC and SEN also demonstrate the effectiveness of our proposed method for MCI classification.

D. Convergence of the Individual-Specific Functional Parcellation Approach

To objectively evaluate the convergence of the proposed functional parcellation approach, the similarity of PFC templates with different iterations is measured with the Dice coefficient [15]. Fig. 7(a) shows the average Dice between the template with k iterations and with $k-1$ iterations for each sample, where the initial template is set as iteration $k=0$ in the evaluation. It is observed that after 3 or 4 iterations, the template is highly consistent with it in the last iteration ($Dice > 0.95$). As the iterative procedure progresses, it converges to a stable functional template which can map the unique spatial distribution of the functional ROIs at the individual level. The convergence speed is correlated with the confidence threshold C_{th}. Furthermore, we compare the similarity between PFC templates of different samples. It is shown in Fig. 7(b) that the inter-subject variability increases and stabilizes after several iterations. Between any two individuals, the Dice coefficient is about 0.5 to 0.6 for the final PFC templates. It indicates that the functional architecture varies across different individuals. The proposed parcellation algorithm is able to obtain individual-specific functional ROIs and can reflect the network distribution differences. In addition, we find that the confidence threshold has an influence on the inter-subject variability of templates. A larger value could lead to a more unique functional template, but the reliability is decreased. It is a requirement for a powerful MCI classification performance to detect the optimal C_{th} in the parcellation strategy.

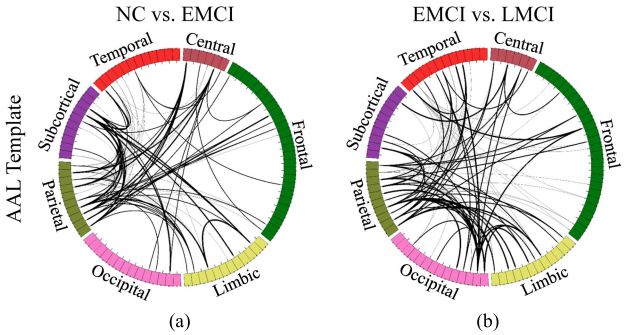

Fig. 9. The top 100 most discriminative FC on AAL template for (a) NC vs. EMCI, (b) EMCI vs. LMCI.

E. Influence of the Confidence Threshold C_{th} on Classification Performance

In the construction of PFC template, the confidence threshold C_{th} is used to determine core vertexes and further generate a reference signal for each ROI. To explore the influence of C_{th} on the MCI classification performance, we vary C_{th} from 1.5 to 3.0 in steps of 0.1 and report the corresponding classification ACC and AUC of the proposed framework with their standard deviations in Fig. 8. It is shown that $C_{th}=0.2$ is an optimal parameter and achieves the best classification performance for NC vs. EMCI, EMCI vs. LMCI, and NC vs. EMCI vs. LMCI classification tasks simultaneously. Meanwhile, we note that a too-large or too-small value of C_{th} degrades the generalization performance of the model. An excessively high value of C_{th} indicates that few vertexes are selected as core signals, and they tend to stay within the original ROI area. The generated functional template thus includes less individualized FC information and is more similar to the initial template, which cannot provide enough complementary information for the classification model. On the contrary, a too-small C_{th} represents that a wide range of core vertexes are selected by the individualized FC. The reference signals completely dominated by the individualized information may reduce the reliability of the generated functional template due to the possible noise and limited sample size of personalized fMRI data. It may lead to the overfitting problem and performance degradation of the MCI identification framework.

F. Most Discriminative Brain Correlations

Inspired by related works [22], [41], we perform a significance test to investigate the most discriminative FCs in the AAL and PFC templates, which selects the features by the following steps. First, in each fold of cross-validation, we perform a Pearson's correlation test between the FCs and the labels, and select the features with FDR corrected p-value < 0.01. Second, since the selected features are different in each fold, and 10 times of 10-fold cross-validation are performed in our experiment, we rank the FCs according to the numbers of times they are selected. In this way, the FC features that always occur in the cross-validations are defined as the most discriminative brain correlations.

Fig. 9 and Fig. 10 shows the top 100 most frequently selected features in the AAL and PFC templates across

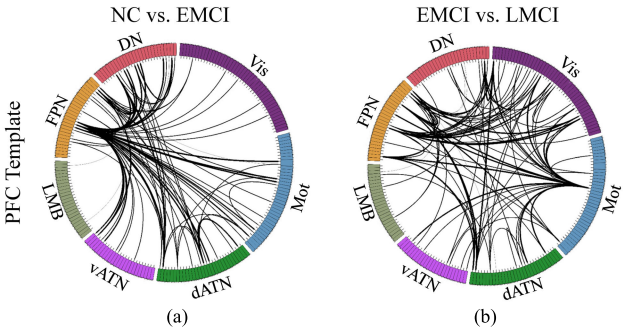

Fig. 10. The top 100 most discriminative FC on PFC template for (a) NC vs. EMCI, (b) EMCI vs. LMCI.

10 times of classifications, where the line thickness represents the importance of functional connection. The ROIs on the AAL anatomical template are classified by their structural locations, i.e., temporal, central, frontal, limbic, occipital, parietal lobes and subcortical regions. The ROIs on the PFC template are classified into 7 functional networks defined in Yeo's functional atlas [42], i.e. visual network (Vis), motor network (Mot), dorsal attention network (dATN), ventral attention network (vATN), limbic (LMB), frontoparietal control network (FPN) and default network (DN).

From the AAL results in Fig. 9(a) and (b), it can be observed that the discriminative FC features used in the two classification tasks (NC vs. EMCI, EMCI vs. LMCI) are relatively similar. The connections from parietal lobe are most frequently as discriminative features, which is consistent with previous MCI studies [43]. From the perspective of brain function, we find that different FC features from the PFC template are used in the two tasks as shown in Fig. 10. In the early stage of MCI, the discriminative FC features are mainly involved in the high-order association networks, including FPN, DN, dATN and vATN, which is correlated with the cognitive and memory functions. In the late stage of MCI, the primary cortex including Vis and Mot networks also show function degenerations. More motor-cognitive and visual-cognitive interactions are chosen to discriminate EMCI and LMCI patients, which can be regarded as the biomarkers for late stage of MCI. These regions is also reported in the previous MCI/AD studies [23], [44]. Note that the proposed PFC template is sensitive to different stages of MCI, and it demonstrates the atrophy of cortical FC is not consistent in early and late MCI. This phenomenon is similar with recent studies that report a more widespread reduction of cortical thickness over time [45], [46].

G. Limitations and Future Directions

Although the proposed framework achieves promising MCI classification performance, there are still two limitations in this work. One limitation is the computational complexity of our framework. Compared with existing MCI identification methods which only using group-averaged templates, our proposed framework spends time on constructing a PFC template before inputting fMRI data into a deep learning model. According to our experiments on Intel CPU with 8GB RAM, it takes about 20.4s for a 12-min rs-fMRI sample. However, in practical applications, multiple samples can be processed simultaneously when a high-RAM CPU is available, and the computational issue becomes less critical. Another limitation is that the genetics data is not utilized in this work. Several studies demonstrated that multimodal data analysis based on rs-fMRI and gene data facilitates the study of brain functions [47], [48], which provides a new solution to identify MCI patients more accurately. Our future work will try to reveal the correlations between individual-specific FC and gene of MCI patients by multimodal data fusion.

V. Conclusion

In this paper, we propose a novel MCI classification framework called PFC-DBGNN-STAA, which fuses individual- and group-level functional features and is effective in identifying both EMCI and LMCI. First, we introduce an individual-level parcellation method to build a PFC template for each sample. This method can automatically project 213 functional ROIs onto each individual's cerebral cortex according to its FC patterns, and supplement important individual-specific functional features. Then, with FC networks detected by a group-constrained algorithm, the individual- and group-level functional features are fused by a DBGNN. This feature extraction approach captures the relationships between two templates and achieves more promising performance compared with the linear concatenation of features. Finally, we feed these fused features into a novel STAA module to generate high-level spatio-temporal FC representations, which emphasizes both spatial correlations and temporal dynamics in brain ROIs. We perform experiments on the ADNI database to evaluate our proposed method, which verifies the efficacy of the PFC template, DBGNN, and STAA. The detailed discussion shows that the proposed method captures individual-specific functional features, which can be regarded as the critical MCI-related discriminative biomarkers. Furthermore, experimental results show that our method achieves better performance against existing MCI classification methods, which has a potential application prospect in clinical MCI diagnosis.

References

[1] S. Gauthier et al., "Mild cognitive impairment," *Lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.

[2] K. Ma, S. Huang, and D. Zhang, "Diagnosis of mild cognitive impairment with ordinal pattern kernel," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1030–1040, 2022.

[3] S. Siuly et al., "A new framework for automatic detection of patients with mild cognitive impairment using resting-state EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 9, pp. 1966–1976, Jul. 2020.

[4] J. P. Chhatwal et al., "Plasma N-terminal tau fragment levels predict future cognitive decline and neurodegeneration in healthy elderly individuals," *Nature Commun.*, vol. 11, no. 1, pp. 1–10, Nov. 2020.

[5] R. Cabeza et al., "Maintenance, reserve and compensation: The cognitive neuroscience of healthy ageing," *Nature Rev. Neurosci.*, vol. 19, no. 11, pp. 701–710, Nov. 2018.

[6] 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.

[7] Y. Xue, Y. Zhang, L. Zhang, S.-W. Lee, L. Qiao, and D. Shen, "Learning brain functional networks with latent temporal dependency for MCI identification," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 590–601, Feb. 2022.

[8] N. Tzourio-Mazoyer et al., "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, 2002.

[9] J. D. Power, D. A. Fair, B. L. Schlaggar, and S. E. Petersen, "The development of human functional brain networks," *Neuron*, vol. 67, no. 5, pp. 735–748, 2010.

[10] M. F. Glasser et al., "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171–178, Aug. 2016.

[11] C. Schifani et al., "Repetitive transcranial magnetic stimulation (rTMS) alters variability in brain function in schizophrenia: Data from a double-blind, randomized, sham-controlled trial," *Biol. Psychiatry*, vol. 89, no. 9, pp. S83–S84, May 2021.

[12] S. Mueller et al., "Individual variability in functional connectivity architecture of the human brain," *Neuron*, vol. 77, no. 3, pp. 586–595, 2013.

[13] R. M. Braga and R. L. Buckner, "Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity," *Neuron*, vol. 95, no. 2, pp. 457.e5–471.e5, Jul. 2017.

[14] M. Li et al., "Performing group-level functional image analyses based on homologous functional regions mapped in individuals," *PLOS Biol.*, vol. 17, no. 3, Mar. 2019, Art. no. e2007032.

[15] D. H. Wang et al., "Parcellating cortical functional networks in individuals," *Nature Neurosci.*, vol. 18, no. 12, pp. 1853–1860, 2015.

[16] W. Cui et al., "Personalized fMRI delineates functional regions preserved within brain tumors," *Ann. Neurol.*, vol. 91, no. 3, pp. 353–366, Mar. 2022.

[17] M. Liu, D. Zhang, and D. Shen, "Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1463–1474, Jun. 2016.

[18] B. Lei et al., "Adaptive sparse learning using multi-template for neurodegenerative disease diagnosis," *Med. Image Anal.*, vol. 61, Apr. 2020, Art. no. 101632.

[19] S. Parisot et al., "Disease prediction using graph convolutional networks: Application to autism spectrum disorder and Alzheimer's disease," *Med. Image Anal.*, vol. 48, pp. 117–130, Aug. 2018.

[20] Y. Li, J. Liu, Y. Jiang, Y. Liu, and B. Lei, "Virtual adversarial training-based deep feature aggregation network from dynamic effective connectivity for MCI identification," *IEEE Trans. Med. Imag.*, vol. 41, no. 1, pp. 237–251, Jan. 2022.

[21] J. Zhang, L. Zhou, L. Wang, M. Liu, and D. Shen, "Diffusion kernel attention network for brain disorder classification," *IEEE Trans. Med. Imag.*, vol. 41, no. 10, pp. 2814–2827, Oct. 2022.

[22] Y. Li, J. Liu, Z. Tang, and B. Lei, "Deep spatial–temporal feature fusion from adaptive dynamic functional connectivity for MCI identification," *IEEE Trans. Med. Imag.*, vol. 39, no. 9, pp. 2818–2830, Sep. 2020.

[23] J. Alty et al., "The TAS test project: A prospective longitudinal validation of new online motor-cognitive tests to detect preclinical Alzheimer's disease and estimate 5-year risks of cognitive decline and dementia," *BMC Neurol.*, vol. 22, no. 1, pp. 1–13, Dec. 2022.

[24] L. A. M. Lebois et al., "Large-scale functional brain network architecture changes associated with trauma-related dissociation," *Amer. J. Psychiatry*, vol. 178, no. 2, pp. 165–173, Feb. 2021.

[25] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electronics*, vol. 9, no. 8, p. 1295, Aug. 2020.

[26] X. Shen, F. Tokoglu, X. Papademetris, and R. T. Constable, "Groupwise whole-brain parcellation from resting-state fMRI data for network node identification," *NeuroImage*, vol. 82, pp. 403–415, Nov. 2013.

[27] Y. Li, H. Yang, B. Lei, J. Liu, and C.-Y. Wee, "Novel effective connectivity inference using ultra-group constrained orthogonal forward regression and elastic multilayer perceptron classifier for MCI identification," *IEEE Trans. Med. Imag.*, vol. 38, no. 5, pp. 1227–1239, May 2019.

[28] Y. Jiang, H. Huang, J. Liu, C.-Y. Wee, and Y. Li, "Adaptive functional connectivity network using parallel hierarchical BiLSTM for MCI diagnosis," in *Proc. 10th Int. Workshop Mach. Learn. Med. Imag. (MLMI)*, Shenzhen, China. Cham, Switzerland: Springer, Oct. 2019, pp. 507–515.

[29] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: A comprehensive review," *Comput. Social Netw.*, vol. 6, no. 1, pp. 1–23, Dec. 2019.

[30] Y. Li, Y. Liu, Y.-Z. Guo, X.-F. Liao, B. Hu, and T. Yu, "Spatio-temporal-spectral hierarchical graph convolutional network with semisupervised active learning for patient-specific seizure prediction," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 12189–12204, Nov. 2022.

[31] J. Ren et al., "Dissociable auditory cortico-cerebellar pathways in the human brain estimated by intrinsic functional connectivity," *Cerebral Cortex*, vol. 31, no. 6, pp. 2898–2912, May 2021.

[32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[33] Y. Li et al., "Fusion of ULS group constrained high- and low-order sparse functional connectivity networks for MCI classification," *Neuroinformatics*, vol. 18, no. 1, pp. 1–24, Jan. 2020.

[34] L. Dora, S. Agrawal, R. Panda, and A. Abraham, "Nested cross-validation based adaptive sparse representation algorithm and its application to pathological brain classification," *Expert Syst. Appl.*, vol. 114, pp. 313–321, Dec. 2018.

[35] Y. Li, Y. Zhang, W. Cui, B. Lei, X. Kuang, and T. Zhang, "Dual encoder-based dynamic-channel graph convolutional network with edge enhancement for retinal vessel segmentation," *IEEE Trans. Med. Imag.*, vol. 41, no. 8, pp. 1975–1989, Aug. 2022.

[36] D. Yao et al., "A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity," *IEEE Trans. Med. Imag.*, vol. 40, no. 4, pp. 1279–1289, Jan. 2021.

[37] L. Fan et al., "The human brainnetome atlas: A new brain atlas based on connectional architecture," *Cerebral Cortex*, vol. 26, no. 8, pp. 3508–3526, 2016.

[38] Z. Zhou et al., "A toolbox for brain network construction and classification (BrainNetClass)," *Hum. Brain Mapping*, vol. 41, no. 10, pp. 2808–2826, Jul. 2020.

[39] 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.

[40] S. Gadgil, Q. Zhao, A. Pfefferbaum, E. V. Sullivan, E. Adeli, and K. M. Pohl, "Spatio-temporal graph convolution for resting-state fMRI analysis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 528–538.

[41] H. Zhang et al., "Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment," *J. Alzheimer's Disease*, vol. 54, no. 3, pp. 1095–1112, Oct. 2016.

[42] B. T. T. Yeo, F. M. Krienen, J. Sepulcre, and M. R. Sabuncu, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *J. Neurophysiol.*, vol. 106, pp. 1125–1165, Sep. 2011.

[43] C. R. Ilardi, S. Chieffi, T. Iachini, and A. Iavarone, "Neuropsychology of posteromedial parietal cortex and conversion factors from mild cognitive impairment to Alzheimer's disease: Systematic search and state-of-the-art review," *Aging Clin. Exp. Res.*, vol. 34, pp. 1–19, Jul. 2021.

[44] G. G. Yener, D. D. Emek-Savaş, B. Güntekin, and E. Başar, "The visual cognitive network, but not the visual sensory network, is affected in amnestic mild cognitive impairment: A study of brain oscillatory responses," *Brain Res.*, vol. 1585, pp. 141–149, Oct. 2014.

[45] E. C. Edmonds et al., "Patterns of longitudinal cortical atrophy over 3 years in empirically derived MCI subtypes," *Neurology*, vol. 94, no. 24, pp. e2532–e2544, Jun. 2020.

[46] H. Cho et al., "Longitudinal changes of cortical thickness in early-versus late-onset Alzheimer's disease," *Neurobiol. Aging*, vol. 34, no. 7, pp. 1921.e9–1921.e15, 2013.

[47] X.-A. Bi, W. Zhou, L. Li, and Z. Xing, "Detecting risk gene and pathogenic brain region in EMCI using a novel GERF algorithm based on brain imaging and genetic data," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 8, pp. 3019–3028, Aug. 2021.

[48] W. Hu et al., "Interpretable multimodal fusion networks reveal mechanisms of brain cognition," *IEEE Trans. Med. Imag.*, vol. 40, no. 5, pp. 1474–1483, May 2021.