# Deep Fusion of Multi-Template Using Spatio-Temporal Weighted Multi-Hypergraph Convolutional Networks for Brain Disease Analysis

Jingyu Liu, Weigang Cui, Yipeng Chen, Yulan Ma, Qunxi Dong, Ran Cai, Yang Li, *Senior Member, IEEE*, and Bin Hu, *Fellow, IEEE*, for the Alzheimer's Disease Neuroimaging Initiative

***Abstract*—Conventional functional connectivity network (FCN) based on resting-state fMRI (rs-fMRI) can only reflect the relationship between pairwise brain regions. Thus, the hyper-connectivity network (HCN) has been widely used to reveal high-order interactions among multiple brain regions. However, existing HCN models are essentially spatial HCN, which reflect the spatial relevance of multiple brain regions, but ignore the temporal correlation among multiple time points. Furthermore, the majority of HCN construction and learning frameworks are limited to using a single template, while the multi-template carries richer information. To address these issues, we first employ multiple templates to parcellate the rs-fMRI into different brain regions. Then, based on the multi-template data, we propose a spatio-temporal weighted HCN (STW-HCN) to capture more comprehensive high-order temporal and spatial properties of brain activity. Next, a novel deep fusion model of multi-template called spatio-temporal weighted multi-hypergraph convolutional network (STW-MHGCN) is proposed to fuse the STW-HCN of multiple templates, which extracts the deep interrelation information between different templates. Finally, we evaluate our method on the ADNI-2 and ABIDE-I datasets for mild cognitive impairment (MCI) and autism spectrum disorder (ASD) analysis. Experimental results demonstrate that the proposed method is superior to the state-of-the-art approaches in MCI and ASD classification, and the abnormal spatio-temporal hyper-edges discovered by our method have significant significance for the brain abnormalities analysis of MCI and ASD.**

***Index Terms*—Alzheimer's disease (AD), autism spectrum disorder (ASD), deep learning, hyper-connectivity network, multi-template.**

Manuscript received 28 May 2023; revised 10 August 2023; accepted 13 October 2023. Date of publication 17 October 2023; date of current version 2 February 2024. This work was supported in part by the Zhejiang Lab's International Talent Fund for Young Professionals; in part by the Beijing Institute of Technology Research Fund Program for Young Scholars; in part by the National Key Research and Development Program of China under Grant 2019YFA0706200 and Grant 2022YFC3500503; in part by the National Natural Science Foundation of China under Grant 62227807, Grant 61632014, Grant 62302044, Grant 62201023, and Grant 12004034; in part by the National Key Research and Development Program of China under Grant 2021ZD0202000 and Grant 2021ZD0200601; in part by the Beijing Natural Science Foundation under Grant Z220017; in part by the Beijing Municipal Education Commission-Natural Science Foundation under Grant KZ202110025036; and in part by the China Post-Doctoral Science Foundation Funded Project under Grant 2022M720434. *(Corresponding authors: Yang Li; Qunxi Dong; Bin Hu.)*

Jingyu Liu, Qunxi Dong, Ran Cai, and Bin Hu are with the Key Laboratory of Brain Health Intelligent Evaluation and Intervention, Ministry of Education, and the School of Medical Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: liujingyu@bit.edu.cn; dongqx@bit.edu.cn; cairan@bit.edu.cn; bh@bit.edu.cn).

Weigang Cui, Yulan Ma, and Yang Li are with the State Key Laboratory of Virtual Reality Technology and Systems, School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China (e-mail: cwg1994@buaa.edu.cn; mayulan@buaa.edu.cn; liyang@buaa.edu.cn).

Yipeng Chen is with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: z2193896726@yahoo.com).

Digital Object Identifier 10.1109/TMI.2023.3325261

## I. Introduction

FUNCTIONAL connectivity network (FCN) using resting-state fMRI (rs-fMRI) plays a critical role in the detection and analysis of neurological diseases [1], including Alzheimer's disease (AD) and its prodromal stage, mild cognitive impairment (MCI), autism spectrum disorder (ASD), etc. [2], [3], [4]. It measures the dependency of functional activities between any two brain regions [5]. The abnormal changes in functional connectivity are used as a basis for the diagnosis of neurological diseases [6], [7], [8]. However, FCN only measures the pairwise correlation between two brain regions, neglecting high-order interactions among three or more brain regions [9].

Recent studies indicate that high-order interactions are essential to understand activity patterns of the brain nervous system [10]. Therefore, the hyper-connectivity network (HCN) based on hypergraph theory is proposed to measure the high-order relationship among multiple brain regions [9], [11]. For example, Jie et al. [9] adopted a $l_0$-norm sparse regression algorithm to construct a HCN for each subject. Each brain region-of-interesting (ROI) is regarded as the centroid node once connected by other nodes to form a hyper-edge of HCN. Finally, the topological properties of HCN are extracted for MCI identification [9]. Li et al. [12] proposed a functionally-weighted Lasso algorithm to construct multimodal HCN by fusing the blood-oxygen-level-dependent (BOLD) fMRI and arterial spin labeling (ASL) fMRI. The topological features (*i.e.*, three types of unweighted clustering coefficients) of multimodal HCN are calculated for MCI classification [12]. Zhu et al. [11] presented a dynamic hypergraph inference framework based on multimodal neuroimaging for AD and MCI analysis. It should be noted that the "dynamic" in this method refers to the "iterative calculation" of the topological structure in the process of hypergraph learning, without considering the temporal information. Although HCN achieves better performance than FCN for the diagnosis and analysis of neurological diseases, it still has three major limitations. *First*, while the existing HCN models extract high-level spatial information (*i.e.*, high-order interaction among multiple ROIs), they leave the temporal information embedded in rs-fMRI out of consideration. *Second*, most methods only pay attention to the topological properties of HCN, neglecting the different weights of hyper-edges. *Third*, the HCN learning frameworks only employ a single template to partition the brain space, which is difficult to reveal the brain abnormalities of patients comprehensively. To sum up, such loss of temporal, weights and multi-template information, which has a pivotal role in understanding the pathology, may cause performance degradation of the disease analysis model.

According to several studies, compared with single-template approaches, multi-template methods provide a multi-scale perspective of the whole brain, resulting in more comprehensive information and superior performance for detecting and analyzing neurological diseases [13], [14], [15], [16]. For example, Lei et al. [14] partition the T1-weighted MRI using the automated anatomical labeling (AAL-90) template [17] and Craddock 200 (CC-200) template [18], and then concatenate the multi-template features for neurological disease diagnosis, including AD, MCI and Parkinson's disease (PD). Huang et al. [15] construct a FCN for each template, extract a subset of features and develop a classifier for each template, respectively. The voting strategy is adopted to integrate the classification results of different classifiers (corresponding to multi-template) for ASD diagnosis [15]. To fuse multimodal neuroimaging and multi-template information, Long et al. [16] compute the gray matter volumes based on structural MRI and the Hurst exponent based on rs-fMRI data using four types of templates. The multimodal features from different templates are concatenated and then selected via the minimal redundancy maximal relevance method. The optimal features are used for MCI identification [16]. Besides the feature concatenation or classification result voting methods, the multimodal fusion approaches such as canonical correlation analysis [19], multi-task learning [20], and multi-kernel learning [21] can be applied to fuse multi-template. For example, canonical correlation analysis has been employed to integrate the connectivity features of multimodal and multiscale networks for neurobehavioral score prediction [19]. Yao et al. [1] adopt a mutual learning strategy to collaboratively learn feature representation from FCN of different templates for multiple brain disease classification. Although multi-template methods yield better performance, it is still a challenge to achieve multi-template fusion at the FCN or HCN level. This is because different templates divide the brain space into distinct ROIs, leading to disparate nodes within the FCN and HCN of each template. As a result, deep association information between multi-template is difficult to be mined.

To overcome these limitations, we propose a spatio-temporal weighted HCN (STW-HCN) model and further a spatio-temporal weighted multi-hypergraph convolutional network (STW-MHGCN) for the deep fusion of multi-template. Specifically, to construct a more informative analysis framework of neurological diseases, multiple templates are first adopted to divide the brain space into different ROIs. Then, we propose the STW-HCN model based on multi-template data to capture more comprehensive high-level temporal and spatial information. In this model, a spatial weighted HCN is constructed for each template; the rs-fMRI time series of all templates are concatenated along the dimension of ROIs to construct a temporal weighted HCN. During this process, the weights of temporal and spatial hyper-edges are also calculated via a pre-defined function. Next, we propose a novel STW-MHGCN to fuse the temporal weighted HCN and spatial weighted HCN of multi-template by sharing the weight matrix. Finally, the deep fused features generated from STW-MHGCN are used for multiclass classification and abnormal spatio-temporal hyper-edges analysis of brain diseases.

The main contributions of this paper are listed as follows:

1) The proposed STW-HCN model constructs not only the spatial weighted HCN but also the temporal weighted HCN. The former reflects high-order spatial relationships among multiple ROIs, while the latter measures temporal hyper-correlation among several time points, which has been neglected in previous studies. The complementary spatio-temporal correlation representations imply a more powerful brain disease analysis framework.
2) Most HCN construction and learning frameworks only focus on the topology of HCN, ignoring the weight of hyper-edge. Both our proposed HCN construction method (*i.e.*, STW-HCN) and learning model (*i.e.*, STW-MHGCN) consider the topological structure and weight of temporal and spatial HCN, thus providing a more accurate and comprehensive high-level interaction mechanism.
3) Multi-template provides richer characterization for rs-fMRI analysis. In this paper, we proposed the STW-MHGCN to learn deep interrelation representations of multi-template by fusing the STW-HCN of multi-template. To our best knowledge, this is the first time fusing the multi-template information of rs-fMRI at the HCN level.

## II. Methods

The proposed deep fusion framework of multi-template for brain disease analysis is presented in Fig. 1. It is summarized as follows: 1) Construct a spatial weighted HCN for each template; 2) Construct a temporal weighted HCN based on the concatenated rs-fMRI time series of all templates; 3) Apply the STW-MHGCN to fuse the temporal and spatial weighted HCN of multi-template; 4) Employ the deep fused features of

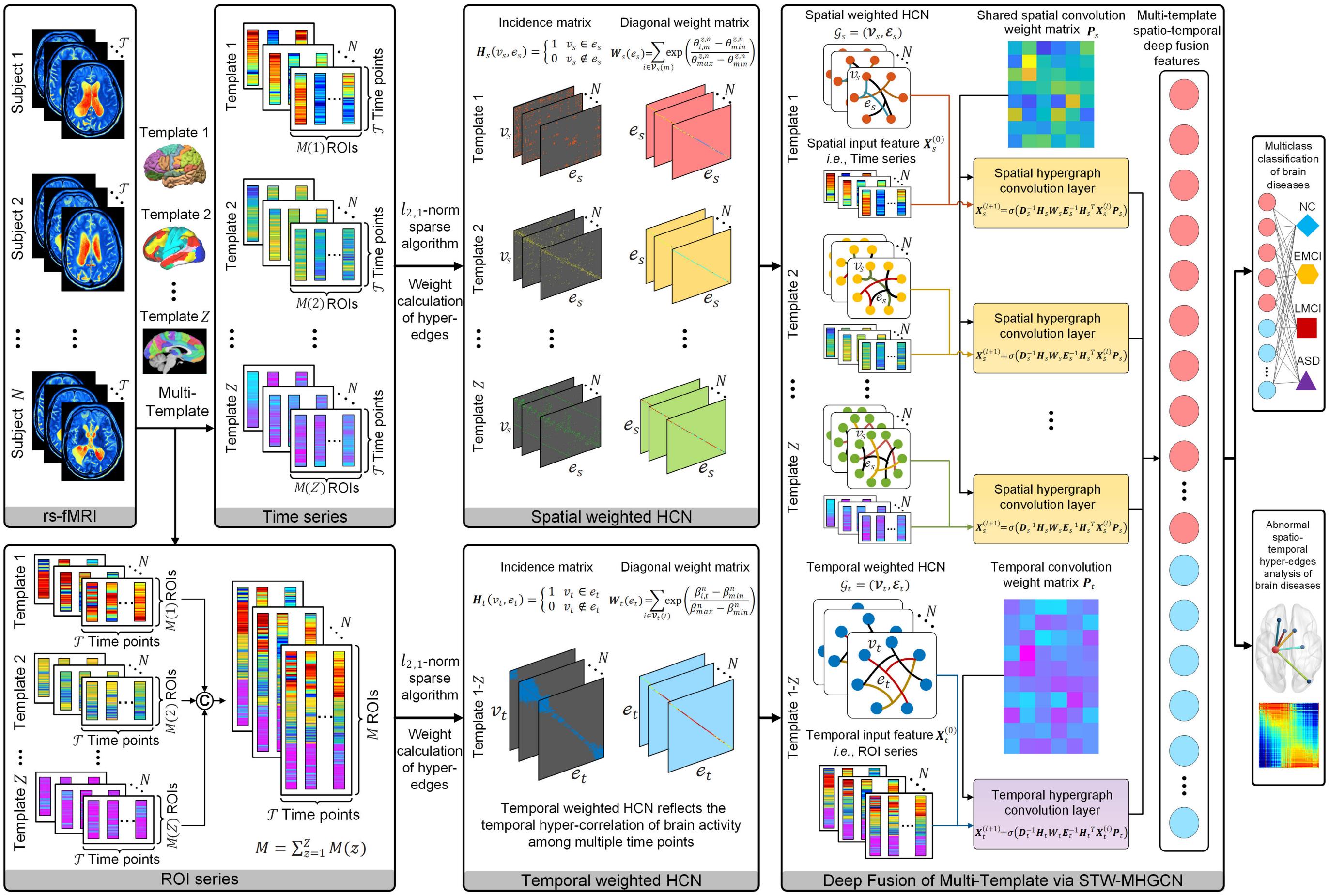

Fig. 1. Illustration of the proposed framework. The multiple templates are first used to partition brain space into different ROIs. Then, the time series and ROI series are extracted to construct spatial and temporal weighted HCN, respectively. Finally, the STW-MHGCN is adopted to fuse multi-template spatio-temporal information, and the deep fused features are employed for classification and analysis of brain diseases.

multi-template for classification and analysis of brain diseases. The preceding steps correspond to the following subsection.

## A. Construction of Spatial Weighted HCN for Multi-Template

The proposed STW-HCN model consists of the spatial weighted HCN and temporal weighted HCN, which can be denoted as $\mathcal{G}_s = (\mathcal{V}_s, \mathcal{E}_s)$ and $\mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}_t)$, respectively. $\boldsymbol{V}_s = \{v_s\}$ and $\mathcal{V}_t = \{v_t\}$ are the vertex sets of spatial and temporal weighted HCN (*i.e.*, the ROIs and time points, respectively). $\mathcal{E}_s = \{e_s\}$ and $\mathcal{E}_t = \{e_t\}$ are the spatial and temporal hyper-edge sets.

To reveal high-order interactions among multiple ROIs, we first construct the spatial weighted HCN for each template. Specifically, let $\boldsymbol{Y}^z \in \mathbb{R}^{N \times M(z) \times T} (z = 1, 2, \cdots, Z)$ denotes the rs-fMRI time series of $z$-th template, where $N$, $M(z)$, and $T$ are the numbers of training samples, ROIs, and time points of $z$-th template, respectively. $Z$ is the number of templates. It's worth noting that $N$ and $T$ of different templates are the same, while the $M(z)$ are different. For $z$-th template, the $\boldsymbol{y}_m^{z,n} \in \mathbb{R}^{\mathcal{T} \times 1} (n = 1, 2, \cdots, N; m = 1, 2, \cdots, M(z))$ denotes the rs-fMRI time series of $m$-th ROI for $n$-th sample. $\boldsymbol{y}_m^{z,n}$ can be fitted by the rs-fMRI time series of other ROIs:

$$\boldsymbol{y}_m^{z,n} = \boldsymbol{Y}_m^{z,n} \boldsymbol{\theta}_m^{z,n} + \boldsymbol{e}_m^{z,n}, \tag{1}$$

where $\boldsymbol{Y}_m^{z,n} = \left[\boldsymbol{y}_1^{z,n}, \cdots, \boldsymbol{y}_{m-1}^{z,n}, \boldsymbol{y}_{m+1}^{z,n}, \cdots, \boldsymbol{y}_{M(z)}^{z,n}\right] \in \mathbb{R}^{\mathcal{T} \times (M(z)-1)}$ is the matrix consisting of all ROI time series except $m$-th ROI, $\boldsymbol{\theta}_m^{z,n} = [\theta_{1,m}^{z,n}, \cdots, \theta_{m-1,m}^{z,n}, \theta_{m+1,m}^{z,n}, \cdots, \theta_{M(z),m}^{z,n}]^T \in \mathbb{R}^{(M(z)-1) \times 1}$ is the coefficient vector representing the spatial connectivity between $m$-th ROI and other ROIs of $z$-th template for $n$-th sample, and $\boldsymbol{e}_m^{z,n}$ denotes the residual vector. Motivated by the existing research that sparse constraints can remove redundant connections and preserve robust connections [22], [23], we adopt a $l_{2,1}$-norm sparse algorithm to calculate $\boldsymbol{\theta}_m^{z,n}$:

$$\boldsymbol{\Theta}_m^z = \operatorname{argmin} \sum\nolimits_{n=1}^{N} \left\| \boldsymbol{y}_m^{z,n} - \boldsymbol{Y}_m^{z,n} \boldsymbol{\theta}_m^{z,n} \right\|_2^2 + \lambda_s \left\| \boldsymbol{\Theta}_m^z \right\|_{2,1},$$
$$\text{with } m = 1, 2, \cdots, M(z); z = 1, 2, \cdots, Z, \tag{2}$$

where $\boldsymbol{\Theta}_m^z = [\boldsymbol{\theta}_m^{z,1}, \boldsymbol{\theta}_m^{z,2}, \cdots, \boldsymbol{\theta}_m^{z,N}] \in \mathbb{R}^{(M(z)-1) \times N}$ is the coefficient matrix of spatial connectivity between $m$-th ROI and other ROIs of $z$-th template for all samples, $\lambda_s$ denotes the regularization parameter, $\|\cdot\|_{2,1}$ means the $l_{2,1}$-norm of matrix (*i.e.*, the summation of the $l_2$-norm of each row). By using (2), we calculate $\boldsymbol{\theta}_m^{z,n} (m = 1, 2, \cdots, M(z); z = 1, 2, \cdots, Z; n = 1, 2, \cdots, N)$ that reflects the spatial connectivity between any ROI and other ROIs. Therefore, $\boldsymbol{\theta}_m^{z,n}$ can be regarded as a spatial hyper-edge $e_s$ indicating the spatial relationships among a centroid vertex (*i.e.*, $m$-th ROI) and other ROIs with corresponding nonzero elements in $\boldsymbol{\theta}_m^{z,n}$. The spatial weighted HCN $\mathcal{G}_s = (\boldsymbol{\mathcal{V}}_s, \boldsymbol{\mathcal{E}}_s)$ can be represented via the incidence matrix $\boldsymbol{H}_s \in \mathbb{R}^{|\boldsymbol{V}_s|\times|\boldsymbol{E}_s|}$:

$$\boldsymbol{H}_s(v_s, e_s) = \begin{cases} 1 & v_s \in e_s \\ 0 & v_s \notin e_s \end{cases} \tag{3}$$

Compared with the existing HCN treats all hyper-edges equally, we further calculate the weight of each $e_s$. Particularly, the weight of $e_s$ is computed via:

$$\boldsymbol{W}_s(e_s) = \sum\nolimits_{i\in\boldsymbol{\mathcal{V}}_s(m)} \exp\left(\frac{\theta_{i,m}^{z,n} - \theta_{min}^{z,n}}{\theta_{max}^{z,n} - \theta_{min}^{z,n}}\right), \tag{4}$$

where $\boldsymbol{W}_s(e_s)$ is the weight of $e_s$ corresponding to $\boldsymbol{\theta}_m^{z,n}$, $\boldsymbol{\mathcal{V}}_s(m)$ denotes the neighboring ROI sets of $m$-th ROI according to the nonzero elements in $\boldsymbol{\theta}_m^{z,n}$, $\theta_{max}^{z,n}$ and $\theta_{min}^{z,n}$ denote the maximum and minimum spatial connectivity of $z$-th template for $n$-th sample, respectively. Therefore, the weight describes the relative strength of the hyperedge in the entire HCN. Based on (4), the stronger spatial connectivity between centroid ROI and other ROIs, the greater weight of $e_s$. By storing all weights in a diagonal matrix, we obtain the weight matrix $\boldsymbol{W}_s$ of $\mathcal{G}_s$. The incidence matrix $\boldsymbol{H}_s$ and weight matrix $\boldsymbol{W}_s$ represent the topological structure and weight of $G_s$. Therefore, different from the traditional HCN models which only extract topological properties of HCN based on incidence matrix $\boldsymbol{H}_s$ [9], [12], the proposed spatial weighted HCN supplies more comprehensive hyper-connectivity representations.

The degrees of vertexes and hyper-edges in the spatial weighted HCN are defined as:

$$\boldsymbol{D}_s(v_s) = \sum\nolimits_{e_s\in\boldsymbol{E}_s} \boldsymbol{W}_s(e_s)\,\boldsymbol{H}_s(v_s, e_s), \tag{5}$$

$$\boldsymbol{E}_s(e_s) = \sum\nolimits_{v_s\in\boldsymbol{V}_s} \boldsymbol{H}_s(v_s, e_s), \tag{6}$$

where $\boldsymbol{D}_s(v_s)$ and $\boldsymbol{E}_s(e_s)$ denote the degree of vertex $v_s$ and hyper-edge $e_s$, $\boldsymbol{D}_s$ and $\boldsymbol{E}_s$ are the diagonal matrices of vertex and hyper-edge degrees of spatial weighted HCN, respectively.

### B. Construction of Temporal Weighted HCN for Multi-Template

To explore the temporal hyper-correlation among time points, we construct the temporal weighted HCN. In contrast to the previous section, where we constructed one spatial weighted HCN for each template, in this section, we infer one temporal weighted HCN for all templates. Specifically, by concatenating the rs-fMRI time series $\boldsymbol{Y}^z$ of all templates along the dimension of ROIs, we obtain $\boldsymbol{Y}\in\mathbb{R}^{N\times M\times\mathcal{T}}$, where $M = \sum_{z=1}^{Z} M(z)$. For $n$-th sample, let $\boldsymbol{y}_t^n \in \mathbb{R}^{M\times 1}$ denotes the ROI series of $t$-th time point for all templates, and $\boldsymbol{Y}_t^n = \left[\boldsymbol{y}_1^n, \cdots, \boldsymbol{y}_{t-1}^n, \boldsymbol{y}_{t+1}^n, \cdots, \boldsymbol{y}_{\mathcal{T}}^n\right] \in \mathbb{R}^{M\times(\mathcal{T}-1)}$ is the ROI series matrix consisting of all time points except $t$-th time point. The temporal hyper-edge $e_t$ is inferred by using the same $l_{2,1}$-norm sparse algorithm as in the preceding section:

$$\boldsymbol{B}_t = \operatorname{argmin} \sum\nolimits_{n=1}^{N} \left\|\boldsymbol{y}_t^n - \boldsymbol{Y}_t^n\boldsymbol{\beta}_t^n\right\|_2^2 + \lambda_t \left\|\boldsymbol{B}_t\right\|_{2,1}, \quad \text{with } t = 1, 2, \cdots, \mathcal{T}, \tag{7}$$

where $\boldsymbol{\beta}_t^n = [\beta_{1,t}^n, \cdots, \beta_{t-1,t}^n, \beta_{t+1,t}^n, \cdots, \beta_{\mathcal{T},t}^n]^{\mathcal{T}} \in \mathbb{R}^{(\mathcal{T}-1)\times 1}$ is the temporal hyper-edge $e_t$ (*i.e.*, the coefficient vector representing the temporal connectivity between $t$-th time point and other time points of $n$-th sample for all templates), $\boldsymbol{B}_t = [\boldsymbol{\beta}_t^1, \boldsymbol{\beta}_t^2, \cdots, \boldsymbol{\beta}_t^N] \in \mathbb{R}^{(\mathcal{T}-1)\times N}$ is the coefficient matrix for all samples, and $\lambda_t$ denotes the regularization parameter. Nonzero elements in $\boldsymbol{\beta}_t^n$ mean that the corresponding temporal vertexes (*i.e.*, time points) belong to the temporal hyper-edge $e_t$. The incidence matrix $\boldsymbol{H}_t \in \mathbb{R}^{|\boldsymbol{V}_t|\times|\boldsymbol{E}_t|}$ of temporal weighted HCN $\mathcal{G}_t = (\boldsymbol{\mathcal{V}}_t, \boldsymbol{\mathcal{E}}_t)$ is represented as:

$$\boldsymbol{H}_t(v_t, e_t) = \begin{cases} 1 & v_t \in e_t \\ 0 & v_t \notin e_t \end{cases} \tag{8}$$

Similar to the previous section, we compute the weight of $e_t$ by using:

$$\boldsymbol{W}_t(e_t) = \sum\nolimits_{i\in\boldsymbol{V}_t(t)} \exp\left(\frac{\beta_{i,t}^n - \beta_{min}^n}{\beta_{max}^n - \beta_{min}^n}\right), \tag{9}$$

where $\boldsymbol{W}_t(e_t)$ denotes the weight of $e_t$ corresponding to $\boldsymbol{\beta}_t^n$, $\boldsymbol{\mathcal{V}}_t(t)$ is the neighboring time point sets of $t$-th time point according to the nonzero elements in $\boldsymbol{\beta}_t^n$, $\beta_{max}^n$ and $\beta_{min}^n$ are the maximum and minimum temporal connectivity of $n$-th sample, respectively. The weight matrix $\boldsymbol{W}_t$ of $\mathcal{G}_t$ is generated by resetting all weights into a diagonal matrix. By constructing the temporal weighted HCN, we acquire the hyper-connectivity information in the time dimension, which is neglected in the previous studies of HCN.

The degrees of vertexes and hyper-edges in the temporal weighted HCN are defined as:

$$\boldsymbol{D}_t(v_t) = \sum\nolimits_{e_t\in\boldsymbol{E}_t} \boldsymbol{W}_t(e_t)\,\boldsymbol{H}_t(v_t, e_t), \tag{10}$$

$$\boldsymbol{E}_t(e_t) = \sum\nolimits_{v_t\in\boldsymbol{V}_t} \boldsymbol{H}_t(v_t, e_t), \tag{11}$$

where $\boldsymbol{D}_t(v_t)$and $\boldsymbol{E}_t(e_t)$ are the degree of $v_t$ and $e_t$, $\boldsymbol{D}_t$ and $\boldsymbol{E}_t$ denote the corresponding diagonal matrices of vertex and hyper-edge degrees in the temporal weighted HCN, respectively.

### C. Deep Fusion of Multi-Template via STW-MHGCN

The spatial weighted HCN and temporal weighted HCN of multi-template are constructed via the STW-HCN model. In this section, we proposed a STW-MHGCN to learn more comprehensive multi-template spatio-temporal feature representations from STW-HCN. The STW-MHGCN is composed of spatial and temporal hypergraph convolution layers, which are designed by generalizing graph convolution operations to hypergraph, and following the principles: 1) direct feature propagation only occurs among vertexes connected by the same hyperedge; 2) the influence of feature propagation between vertexes with more shared hyperedges is greater; 3) the hyperedge with larger weight plays a greater role in the feature propagation process. Thus, spatial and temporal hypergraph convolution layers are defined as follows respectively:

$$\boldsymbol{X}_s^{(l+1)} = \sigma\left(\boldsymbol{D}_s^{-1}\boldsymbol{H}_s\boldsymbol{W}_s\boldsymbol{E}_s^{-1}\boldsymbol{H}_s^T\boldsymbol{X}_s^{(l)}\boldsymbol{P}_s\right), \tag{12}$$

$$\boldsymbol{X}_t^{(l+1)} = \sigma\left(\boldsymbol{D}_t^{-1}\boldsymbol{H}_t\boldsymbol{W}_t\boldsymbol{E}_t^{-1}\boldsymbol{H}_t^T\boldsymbol{X}_t^{(l)}\boldsymbol{P}_t\right), \tag{13}$$

where $\boldsymbol{X}_s^{(l)} \in \mathbb{R}^{|\mathcal{V}_s| \times F_s^{(l)}}$ and $\boldsymbol{X}_t^{(l)} \in \mathbb{R}^{|\mathcal{V}_t| \times F_t^{(l)}}$ are spatial and temporal input features, $\boldsymbol{P}_s \in \mathbb{R}^{F_s^{(l)} \times F_s^{(l+1)}}$ and $\boldsymbol{P}_t \in \mathbb{R}^{F_t^{(l)} \times F_t^{(l+1)}}$ denote the spatial and temporal convolution weight matrix, respectively, $\sigma(\cdot)$ denotes a nonlinear activation function. The row-normalization is utilized in (12) and (13).

In this study, the initial spatial and temporal input features $\boldsymbol{X}_s^{(0)} = \left[\boldsymbol{y}_1^{z,n}, \boldsymbol{y}_2^{z,n}, \cdots, \boldsymbol{y}_{M(z)}^{z,n}\right]^{\mathcal{T}} \in \mathbb{R}^{M(z) \times T}$ and $\boldsymbol{X}_t^{(0)} = \left[\boldsymbol{y}_1^n, \boldsymbol{y}_2^n, \cdots, \boldsymbol{y}_{\mathcal{T}}^n\right]^T \in \mathbb{R}^{T \times M}$ are the rs-fMRI time series and ROI series matrix. Considering the limited sample size of medical images, we use single-layer spatial and temporal hypergraph convolution operations to reduce the number of trainable parameters in the STW-MHGCN model. Specifically, we fuse the information of multi-template in the spatial domain and time domain, respectively: 1) For each template, one spatial hypergraph convolution layer is adopted to aggregate the spatial features, and the spatial convolution weight matrix $\boldsymbol{P}_s$ is shared across multi-template. It is inspired by the existing study using a weight-shared convolutional neural network (CNN) to learn multi-scale features for the comprehensive information of images [24]. The weight-shared strategy reduces the number of parameters and fuses multi-scale features effectively [24]. In this study, the time dimensions of spatial input features $\boldsymbol{X}_s^{(l)}$ from different templates are unified and shareable, allowing for the adoption of the shared $\boldsymbol{P}_s$ to compress the time dimension of multi-template features. By aggregating these features along hyperedges in the spatial domain, deep association information across different templates can be extracted; 2) Since a shared temporal weighted HCN is constructed for all templates in the previous section, one temporal hypergraph convolution layer is employed for all templates. Multi-template information is fused in the time domain by constructing the shared temporal weighted HCN and temporal hypergraph convolution for all templates. The outputs of spatial and temporal hypergraph convolution layers are concatenated as the deep fused multi-template spatio-temporal features, and used for subsequent classification and analysis of brain diseases.

### D. Classification and Evaluation

The multi-layer perceptron (MLP) with two fully connected layers is adopted for brain disease classification based on deep fused multi-template spatio-temporal features. The Adam algorithm [25] with a learning rate of 0.001 is used as the optimizer. The cross-entropy loss between predicted labels and actual labels of training set is used as the loss function for model optimization. The trainable parameters including $\boldsymbol{P}_s$ and $\boldsymbol{P}_t$ are initialized via the uniform function. The maximum epoch is set to 1000.

We adopt the nested 10-fold cross-validation to divide the sample for model training and evaluation, which consists of two steps: 1) In the outer cross-validation loop, we move one-fold samples out of the dataset to test the model performance. This process continues until all fold samples have been removed once; 2) The remaining nine-fold samples are then split into the training set and validation set via an inner 10-fold cross-validation loop. In this step, the grid search strategy is employed for hyper-parameters optimization, including $\lambda_s([0.05, 0.1, \cdots, 0.3])$ and $\lambda_t([0.05, 0.1, \cdots, 0.3])$. The details of nested 10-fold cross-validation can be found in our previous studies [7], [26]. Moreover, we perform the nested 10-fold cross-validation 20 times, changing the sample division each time, to acquire more trustworthy evaluation results. The average of these 20 repetitions is regarded as the performance of our model and comparison methods.

It is worth mentioning that, while the proposed method is applicable to the fusion of multi-template HCN, it can also be extended to fuse the multi-template FCN. This is because FCN is a special case of HCN, with each edge connecting two nodes. By utilizing (2) and (7) from *Section* II-A and II-B, we can construct the spatial and temporal FCN of multi-template. Then, by replacing hypergraph convolution layers in *Section* II-C with graph convolution layers, we can obtain the deep fused spatio-temporal features of multi-template FCN for brain disease analysis.

## III. Experiment and Results

### A. Materials and Preprocessing

We obtained the rs-fMRI used in this paper from the ADNI-2 (https://adni.loni.usc.edu/) and ABIDE-I databases [27].

The acquisition parameters of rs-fMRI in ADNI-2 are listed as follows: Field Strength=3.0 Tesla; Manufacturer=Philips; TE=30 ms; TR=3000 ms; flip angle = 80°, imaging matrix size = 64 × 64; Slice Thickness=3.3 mm; 140 volumes; 48 slices. We perform the preprocessing by using SPM12 [28]. The preprocessing contains the removal of the first 4 volumes, slice timing correction, head-motion correction, spatial normalization and smoothing, and band-pass filtering. We exclude the subjects with excessive head motion. More details of preprocessing are available in the previous study [29]. 219 NC, 257 early MCI (EMCI), and 164 late MCI (LMCI) remain for subsequent studies. By using the two-sample $t$-test, we discover no significant differences between NC, EMCI, and LMCI in terms of gender, age, or head motion displacement.

ABIDE-I database contains 573 NC and 539 ASD. The preprocessed rs-fMRI data are publicly accessible from the Preprocessed Connectomes Project (PCP) [27]. The configurable pipeline for the analysis of connectomes (CPAC) is implemented for preprocessing. Details of acquisition and preprocessing process are available at http://preprocessed-connectomes-project.org/abide. Consistent with existing research, due to the requirements of subsequent models, we select 1009 subjects by cropping and fixing the length of rs-fMRI time series, consisting of 516 NC and 493 ASD [30].

In both databases, we partition brain space into ROIs via the AAL-90 template [17], Schaefer (SC-100) template [31], [32], and Brainnetome (BN-246) template [33], respectively. The rs-fMRI time series for each ROI is extracted by averaging gray matter voxels. AAL-90 is the most popular parcellation template, with the manual drawing of ROIs every 2mm on axial slices of a high-resolution T1 volume using anatomical information [17]. SC-100 partitions the brain space using the local gradient and global similarity of fMRI, detecting abrupt transitions in functional connectivity patterns and clustering similar ones [31]. BN-246 is a parcellation template based on both structural and functional connectivity, where voxels within the same ROI exhibit similar connectivity profiles [33]. The multi-template with different parcellation principles offers a comprehensive view of the whole brain at multiple scales. Although we adopt AAL-90, SC-100, and BN-246 templates in this study, other templates can also be used in the proposed framework for richer information.

### B. Overall Performance

We design four classification tasks to evaluate the performance of the model, including 1) NC vs. EMCI, 2) EMCI vs. LMCI, 3) NC vs. EMCI vs. LMCI, and 4) NC vs. ASD. In order to validate the superiority of the deep fusion of multi-template, five comparison frameworks are designed using different templates:

1) **AAL-90**: The rs-fMRI time series based on the AAL-90 template [17] is used for subsequent modeling.
2) **SC-100**: Only the SC-100 template [31] is utilized to parcellate brain space.
3) **BN-246**: Only the BN-246 template [33] is adopted for ROIs parcellation.
4) **AAL-90 + SC-100 + BN-246 (Feature Concatenation)**: This method concatenates the features of different templates directly instead of using the shared convolution weight matrix $\boldsymbol{P}_s$. Moreover, compared with the proposed framework, this method does not construct a shared temporal HCN and temporal hypergraph convolution layer for all templates, but a separate temporal HCN and temporal hypergraph convolution layer for each template.
5) **AAL-90 + SC-100 + BN-246 (Deep Fusion)**: It is the proposed framework that uses the STW-MHGCN to fuse the multi-template.

Each framework with different templates is combined with six HCN construction and learning models:

1) **Spatial HCN** (S-HCN): The S-HCN is represented only by the spatial incidence matrix $\boldsymbol{H}_s$, and the weight matrix $\boldsymbol{W}_s$ is set to the identity matrix.
2) **Temporal HCN** (T-HCN): The T-HCN is represented by $\boldsymbol{H}_t$, and the weight matrix $\boldsymbol{W}_t$ is set to the identity matrix.
3) **Spatio-Temporal HCN**(ST-HCN): The features of S-HCN and T-HCN are combined for brain disease classification.
4) **Spatial Weighted HCN** (SW-HCN): The SW-HCN is constructed based on *Section* II-A, and then inputted into the spatial hypergraph convolution layer.
5) **Temporal Weighted HCN** (TW-HCN): The TW-HCN is constructed via *Section* II-B, and then the temporal hypergraph convolution layer is adopted to learn deep features from TW-HCN.
6) **STW-HCN**: It is the proposed HCN model.

The parameter optimization and performance evaluation of all methods are based on the contents of *Section* II-D fairly. The accuracy (ACC), area under curve (AUC), sensitivity (SEN), and specificity (SPE) are used as evaluation metrics for binary classification tasks. The ACC, AUC, SEN of NC, EMCI, and LMCI ($\mathrm{SEN_N}$, $\mathrm{SEN_E}$, and $\mathrm{SEN_L}$) are used to evaluate the three-class classification task.The average performance of 10-fold cross-validation with 20 repetitions is shown in Table I. The proposed framework is superior to all comparison methods in all metrics. It yields the highest ACC of 88.9%, 89.1%, 83.0%, 76.6% and highest AUC of 0.885, 0.880, 0.844, 0.774 for NC vs. EMCI, EMCI vs. LMCI, NC vs. EMCI vs. LMCI, and NC vs. ASD classification tasks, respectively. The effectiveness of the proposed framework is demonstrated from three aspects: 1) Compared with the single template approaches, the proposed framework improves ACCs by at least 6.8%, 6.8%, 7.5%, and 6.1% in four different classification tasks, respectively. In contrast to the feature concatenation strategy for multi-template fusion, the proposed framework yields 5.6%, 4.8%, 4.7%, 4.6% improvement of ACC, and 0.054, 0.055, 0.057, 0.045 improvement of AUC in four different classification tasks, respectively. It confirms the importance of collaborative multi-template information derived from the proposed deep fusion strategy. 2) By considering the weights of hyper-edges in the process of HCN construction and learning, the ACC gains of 4.8%, 5.0%, 4.4%, 3.8% and the AUC gains of 0.060, 0.050, 0.079, 0.044 are obtained in four classification tasks, respectively. It indicates that the more comprehensive hyper-edge representations instead of treating all hyper-edges equally are beneficial to MCI and ASD classification. 3) By incorporating SW-HCN, the proposed framework achieves ACC increase of 4.9%, 4.5%, 5.0%, and 4.9% than TW-HCN in four classification tasks, respectively. Meanwhile, in the proposed framework, TW-HCN contributes to respective enhancements of 6.7%, 5.2%, 5.4%, and 4.3% in ACC across the four classification tasks. SW-HCN reflects the spatial correlation of multiple ROIs, while TW-HCN measures temporal correlation among several time points. The significant performance gains demonstrate the necessity of complementary spatio-temporal hyper-correlation representations for MCI and ASD analysis. Moreover, we apply the two-sample $t$-test between the proposed method and other comparison methods based on the results of 20 repetitions. The $p$-values of all evaluation metrics are less than 0.001, proving the significant superiority of the proposed framework over other methods.

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

In the previous section, we design the relevant HCN comparison model with different templates. To further prove the effectiveness of proposed framework, we compare it with five state-of-the-art methods for brain diseases diagnosis, which include:

1) **Strength and similarity guided group sparse representation** (SSGSR): The strength of individual functional connectivity (FC) and the similarity of inter-subject FC are used to guide the GSR network construction. The elements in GSR network are regarded as features. The Lasso algorithm and SVM are employed for feature selection and classification, respectively [34].
2) **Brain graph neural network** (BrainGNN): The FC network is first constructed by using thresholding partial correlations. The Pearson correlation coefficients

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT HCN MODELS WITH DIFFERENT TEMPLATES

| Template | HCN Method | NC vs. EMCI | | | | EMCI vs. LMCI | | | | NC vs. EMCI vs. LMCI | | | | | NC vs. ASD | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ACC (%) | AUC | SEN (%) | SPE (%) | ACC (%) | AUC | SEN (%) | SPE (%) | ACC (%) | AUC | $SEN_N$ (%) | $SEN_E$ (%) | $SEN_L$ (%) | ACC (%) | AUC | SEN (%) | SPE (%) |
| AAL-90 | S-HCN | 73.7 | 0.743 | 73.6 | 73.7 | 74.4 | 0.749 | 72.7 | 75.5 | 66.0 | 0.683 | 65.8 | 67.0 | 64.6 | 61.1 | 0.626 | 62.7 | 59.7 |
| | T-HCN | 71.8 | 0.739 | 72.6 | 70.8 | 71.8 | 0.729 | 70.0 | 72.9 | 62.4 | 0.679 | 61.5 | 63.7 | 61.6 | 62.7 | 0.653 | 63.3 | 62.2 |
| | ST-HCN | 76.9 | 0.773 | 78.3 | 75.3 | 77.5 | 0.780 | 76.8 | 77.9 | 69.6 | 0.718 | 69.4 | 70.2 | 68.9 | 65.9 | 0.676 | 67.2 | 64.8 |
| | SW-HCN | 77.6 | 0.767 | 79.0 | 76.1 | 78.4 | 0.780 | 77.4 | 79.0 | 71.8 | 0.720 | 71.4 | 73.2 | 70.0 | 65.0 | 0.664 | 66.2 | 63.8 |
| | TW-HCN | 75.4 | 0.766 | 76.7 | 73.9 | 75.8 | 0.767 | 74.5 | 76.6 | 67.9 | 0.709 | 67.2 | 68.9 | 67.4 | 67.3 | 0.691 | 66.9 | 67.6 |
| | STW-HCN | 81.5 | 0.811 | 82.9 | 79.8 | 82.3 | 0.826 | 80.1 | 83.6 | 75.5 | 0.753 | 75.0 | 76.9 | 74.0 | 70.3 | 0.718 | 69.8 | 70.8 |
| SC-100 | S-HCN | 71.6 | 0.719 | 72.9 | 70.2 | 70.2 | 0.740 | 68.5 | 71.4 | 62.9 | 0.657 | 61.7 | 64.4 | 62.0 | 61.8 | 0.632 | 62.3 | 61.4 |
| | T-HCN | 73.2 | 0.752 | 74.4 | 71.9 | 73.1 | 0.739 | 70.9 | 74.5 | 64.8 | 0.695 | 63.6 | 66.8 | 63.2 | 63.1 | 0.651 | 62.7 | 63.4 |
| | ST-HCN | 77.2 | 0.770 | 78.9 | 75.2 | 76.6 | 0.778 | 75.6 | 77.2 | 68.9 | 0.714 | 67.5 | 70.8 | 67.6 | 66.5 | 0.674 | 66.7 | 66.3 |
| | SW-HCN | 75.5 | 0.772 | 76.9 | 73.8 | 77.8 | 0.771 | 76.5 | 78.7 | 67.8 | 0.700 | 66.8 | 68.7 | 67.7 | 64.9 | 0.670 | 65.7 | 64.2 |
| | TW-HCN | 76.4 | 0.780 | 78.1 | 74.3 | 77.2 | 0.782 | 74.8 | 78.8 | 69.4 | 0.727 | 67.6 | 71.6 | 68.4 | 66.4 | 0.696 | 65.5 | 67.3 |
| | STW-HCN | 82.1 | 0.816 | 83.5 | 80.6 | 81.4 | 0.819 | 79.0 | 83.0 | 74.6 | 0.764 | 72.2 | 77.7 | 73.1 | 69.7 | 0.721 | 69.2 | 70.3 |
| BN-246 | S-HCN | 72.2 | 0.714 | 72.0 | 72.5 | 71.6 | 0.708 | 69.4 | 73.0 | 62.6 | 0.671 | 61.8 | 64.0 | 61.4 | 62.9 | 0.648 | 63.1 | 62.6 |
| | T-HCN | 73.9 | 0.753 | 75.2 | 72.4 | 74.0 | 0.740 | 72.0 | 75.2 | 66.3 | 0.707 | 65.6 | 68.1 | 64.6 | 61.6 | 0.634 | 60.2 | 63.0 |
| | ST-HCN | 77.5 | 0.776 | 79.8 | 74.9 | 77.1 | 0.774 | 75.1 | 78.3 | 69.8 | 0.725 | 69.2 | 70.8 | 69.1 | 66.5 | 0.673 | 66.8 | 66.2 |
| | SW-HCN | 75.0 | 0.744 | 75.1 | 75.0 | 76.3 | 0.757 | 74.6 | 77.4 | 67.2 | 0.709 | 66.1 | 69.3 | 65.5 | 66.3 | 0.680 | 66.8 | 65.8 |
| | TW-HCN | 77.7 | 0.783 | 79.0 | 76.2 | 78.1 | 0.782 | 76.0 | 79.4 | 70.5 | 0.733 | 69.2 | 72.0 | 69.8 | 66.1 | 0.686 | 64.2 | 68.0 |
| | STW-HCN | 80.9 | 0.813 | 82.3 | 79.3 | 81.2 | 0.809 | 79.1 | 82.6 | 75.1 | 0.766 | 75.1 | 75.9 | 73.9 | 70.5 | 0.723 | 70.8 | 70.3 |
| AAL-90 + SC-100 + BN-246 (Feature Con.) | S-HCN | 74.9 | 0.744 | 74.1 | 75.9 | 76.3 | 0.758 | 74.4 | 77.6 | 66.2 | 0.701 | 66.2 | 67.8 | 63.8 | 64.7 | 0.655 | 65.2 | 64.1 |
| | T-HCN | 75.8 | 0.765 | 77.3 | 74.0 | 76.7 | 0.769 | 74.5 | 78.1 | 68.4 | 0.710 | 67.2 | 69.8 | 67.6 | 64.4 | 0.670 | 63.0 | 65.9 |
| | ST-HCN | 80.1 | 0.789 | 81.3 | 78.6 | 79.7 | 0.781 | 77.9 | 80.8 | 72.4 | 0.729 | 72.9 | 73.4 | 70.2 | 67.9 | 0.678 | 68.7 | 67.2 |
| | SW-HCN | 80.3 | 0.780 | 80.7 | 79.8 | 79.5 | 0.786 | 78.3 | 80.2 | 73.2 | 0.735 | 73.3 | 75.0 | 70.2 | 67.3 | 0.688 | 67.9 | 66.8 |
| | TW-HCN | 79.7 | 0.786 | 81.5 | 77.5 | 80.0 | 0.794 | 78.2 | 81.1 | 74.1 | 0.750 | 73.4 | 76.0 | 72.1 | 68.0 | 0.693 | 65.9 | 70.1 |
| | STW-HCN | 83.3 | 0.831 | 84.3 | 82.1 | 84.3 | 0.825 | 81.4 | 86.1 | 78.3 | 0.787 | 79.7 | 78.6 | 75.8 | 72.0 | 0.729 | 72.7 | 71.3 |
| **AAL-90 + SC-100 + BN-246 (Deep Fusion)** | S-HCN | 78.9 | 0.772 | 76.9 | 81.1 | 80.9 | 0.782 | 79.4 | 81.9 | 71.3 | 0.730 | 73.2 | 71.3 | 68.7 | 69.5 | 0.689 | 69.9 | 69.1 |
| | T-HCN | 79.7 | 0.799 | 80.3 | 79.1 | 80.5 | 0.799 | 79.9 | 81.0 | 73.8 | 0.752 | 72.8 | 75.0 | 73.4 | 68.3 | 0.690 | 67.6 | 69.0 |
| | ST-HCN | 84.1 | 0.825 | 83.9 | 84.5 | 84.1 | 0.830 | 82.6 | 85.1 | 78.6 | 0.765 | 78.0 | 80.5 | 76.4 | 72.8 | 0.730 | 73.1 | 72.5 |
| | SW-HCN | 82.2 | 0.810 | 83.8 | 80.3 | 83.9 | 0.836 | 82.4 | 84.8 | 77.6 | 0.758 | 77.4 | 79.4 | 75.2 | 72.3 | 0.728 | 72.8 | 71.8 |
| | TW-HCN | 84.0 | 0.845 | 85.0 | 82.9 | 84.6 | 0.837 | 83.6 | 85.3 | 78.0 | 0.798 | 78.2 | 78.2 | 77.3 | 71.7 | 0.725 | 71.0 | 72.4 |
| | **STW-HCN** | **88.9** | **0.885** | **89.9** | **87.8** | **89.1** | **0.880** | **88.1** | **89.8** | **83.0** | **0.844** | **82.8** | **83.9** | **81.8** | **76.6** | **0.774** | **77.1** | **76.1** |

where bold fonts indicate the best performance.

between designated ROI and other ROIs are used as node features. Then, the BrainGNN, which consists of ROI-aware graph convolutional layers and ROI-selection pooling layers, is proposed for feature learning. Finally, the MLP is adopted as the classifier [35].

3) **Mutual multi-scale triplet graph convolutional network**(MMTGCN): The multi-template is used to partition multi-scale ROIs. The FCN of each template is constructed based on the Pearson correlation and $k$-nearest neighbor algorithm. The triplet GCN and mutual learning strategy are employed to learn feature representation from FCN of different templates for multiple brain disease diagnosis. Moreover, it can also be used for structural connectivity networks [1].
4) **Sparse representation with latent temporal dependency** (SRiLT): The SRiLT model uses a latent variable to encode the temporal dependency and sequential order of different rs-fMRI time points during FCN construction. The weights of FC are regarded as features. The $t$-test and SVM are used for feature selection and classification, respectively [36].
5) **High-order connectivity weight-guided graph attention networks** (cwGAT): Dynamic effective connectivity (dEC) is constructed via the group-constrained Kalman filter algorithm. The high-order network is constructed by calculating the correlations between dEC. The features of dEC are aggregated by using the cwGAT based on high-order network topological structure and high-order connectivity weight for classification [26].

Considering the difference in classification tasks of comparison methods (*e.g.*, NC vs. MCI, NC vs. EMCI, etc.) and the larger dataset used in this study, all methods are performed on our dataset for fair comparison. The codes of all comparison methods are from their original authors, which avoids the deviation of method reproduction. The average results of 10-fold cross-validation with 20 repetitions are shown in Table II. The performance of the proposed method exceeds that of other state-of-the-art approaches in all evaluation metrics. It achieves

TABLE II
PERFORMANCE COMPARISON WITH STATE-OF-THE-AET METHODS

| Method | Year | NC vs. EMCI | | | | | EMCI vs. LMCI | | | | | NC vs. EMCI vs. LMCI | | | | | | NC vs. ASD | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ACC (%) | AUC | SEN (%) | SPE (%) | Times (s) | ACC (%) | AUC | SEN (%) | SPE (%) | Times (s) | ACC (%) | AUC | $SEN_N$ (%) | $SEN_E$ (%) | $SEN_L$ (%) | Times (s) | ACC (%) | AUC | SEN (%) | SPE (%) | Times (s) |
| SSGSR [34] | 2019 | 77.9 | 0.805 | 78.0 | 77.8 | 1010.9 | 79.3 | 0.826 | 80.3 | 78.7 | 814.7 | 70.8 | 0.748 | 70.3 | 70.5 | 72.1 | 1267.9 | 66.9 | 0.691 | 65.3 | 68.4 | 2027.2 |
| BrainGNN [35] | 2021 | 82.7 | 0.844 | 82.5 | 82.9 | 730.5 | 77.5 | 0.810 | 76.3 | 78.2 | **538.0** | 74.6 | 0.763 | 75.4 | 75.2 | 72.6 | **925.8** | 68.7 | 0.710 | 66.3 | 71.0 | **1340.2** |
| MMTGCN [1] | 2021 | 80.7 | 0.816 | 81.8 | 79.4 | **705.9** | 80.1 | 0.812 | 78.8 | 81.0 | 631.7 | 76.5 | 0.759 | 75.3 | 77.7 | 76.3 | 1023.1 | 70.7 | 0.705 | 71.4 | 70.0 | 1587.0 |
| SRiLT [36] | 2022 | 82.4 | 0.841 | 82.8 | 81.9 | 1169.5 | 84.4 | 0.853 | 82.6 | 85.5 | 885.9 | 77.1 | 0.793 | 76.6 | 78.9 | 75.0 | 1507.5 | 72.2 | 0.741 | 73.1 | 71.2 | 2335.1 |
| cwGAT [26] | 2022 | 84.1 | 0.848 | 85.1 | 82.8 | 1412.7 | 81.8 | 0.831 | 80.0 | 82.9 | 1238.3 | 78.0 | 0.786 | 77.9 | 78.8 | 76.9 | 1872.8 | 71.1 | 0.729 | 72.4 | 69.8 | 3023.5 |
| **Proposed** | 2022 | **88.9** | **0.885** | **89.9** | **87.8** | 1086.2 | **89.1** | **0.880** | **88.1** | **89.8** | 909.8 | **83.0** | **0.844** | **82.8** | **83.9** | **81.8** | 1450.9 | **76.6** | **0.774** | **77.1** | **76.1** | 2167.6 |

where bold fonts indicate the best performance.

the ACC increase of at least 4.8%, 4.7%, 5.0%, and 4.4% for NC vs. EMCI, EMCI vs. LMCI, NC vs. EMCI vs. LMCI, and NC vs. ASD classification tasks, respectively. Furthermore, compared with the advanced multi-template fusion method MMTGCN [1], our framework obtains the ACC gains of 8.2%, 9.0%, 6.5%, 5.9%, and AUC gains of 0.069, 0.068 0.085, 0.069 in four different classification tasks, respectively. By using the two-sample $t$-test between our method and other competing approaches, our method is proved to be significantly better than all competing approaches in all evaluation metrics under a 99.9% confidence interval ($p$-values < 0.001).

In addition, we compare the computational efficiency of the proposed method and other methods. The computational times are measured on the Intel Core i7-12700F CPU and NVIDIA GeForce RTX 3060 12GB GPU. The software environment includes Matlab R2017b, Python 3.9.12, Pytorch 1.13.1, and CUDA 11.7.1. The time costs of all comparison methods are listed in Table II. We observe that the computational times of these methods are comparable, and the proposed method achieves significant performance improvement at a moderate time cost.

### D. Impact of the Hyper-Parameters

The hyper-parameters $\lambda_s$ and $\lambda_t$ are used to control the sparsity levels of spatial and temporal weighted HCN, respectively. The larger values of $\lambda_s$ and $\lambda_t$ indicate fewer ROIs and time points are connected by the same hyper-edge. To explore the impact of hyper-parameters on the proposed framework, the $\lambda_s$ and $\lambda_t$ are varied from 0.05 to 0.3 with a step size of 0.05 empirically, and the corresponding ACC and AUC in four different classification tasks are shown in Fig. 2. The highest ACC is achieved at $\lambda_s = 0.15$ and $\lambda_t = 0.25$ for four classification tasks, simultaneously. Meanwhile, the optimal AUC is achieved at $\lambda_s = 0.15$ and $\lambda_t = 0.25$ for EMCI vs. LMCI, NC vs. EMCI vs. LMCI, and NC vs. ASD classification tasks, simultaneously. Although in the NC vs. EMCI classification task, the highest AUC of 0.889 is achieved at $\lambda_s = 0.15$ and $\lambda_t = 0.15$, a similar suboptimal AUC of 0.885 is obtained when $\lambda_s = 0.15$ and $\lambda_t = 0.25$. Furthermore, we can observe that the ACC and AUC are robust concerning $\lambda_t$ and relatively sensitive with respect to the changes of $\lambda_s$. The excessively large $\lambda_s$ may result in the hyper-edge not connecting to the critical ROIs. Meanwhile, a too-small $\lambda_s$ may cause redundant ROIs to be reserved in the hyper-edge.

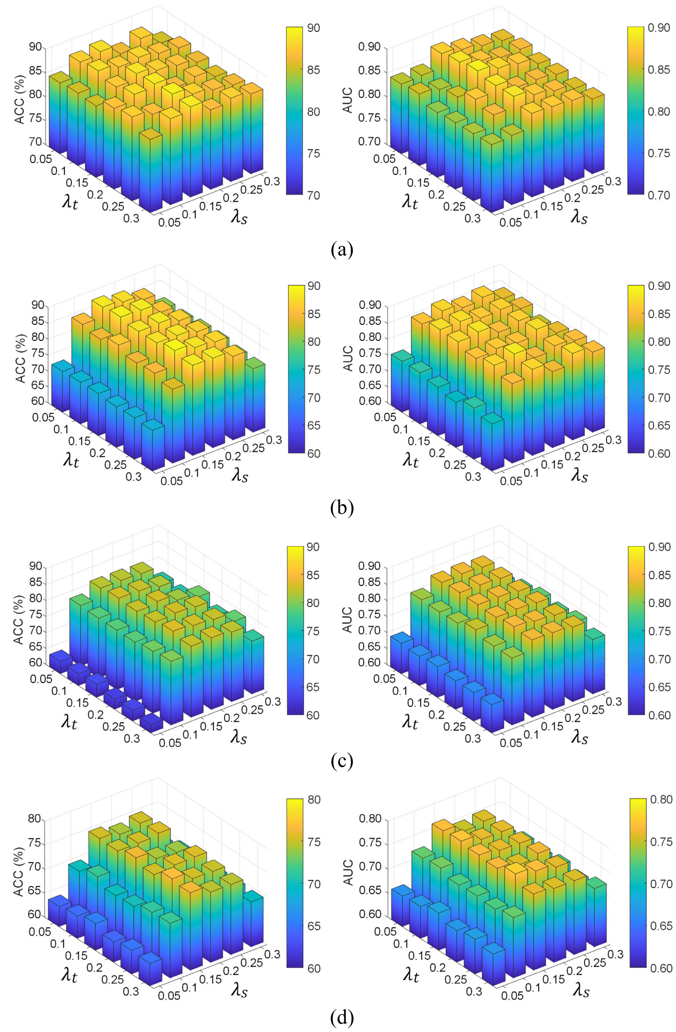


Fig. 2. The classification performance with respect to $\lambda_s$ and $\lambda_t$ in (a) NC vs. EMCI, (b) EMCI vs. LMCI, (c) NC vs. EMCI vs. LMCI, and (d) NC vs. ASD classification tasks.

When $\lambda_s$ varies between 0.15 and 0.25 and $\lambda_t$ varies between 0.05 and 0.3, the high and stable ACC and AUC for four different classification tasks can be achieved. (*i.e.*, ACC > 85% and AUC > 0.84 in NC vs. EMCI and EMCI vs. LMCI classification tasks, ACC > 80% and AUC > 0.83 in NC vs. EMCI vs. LMCI classification task, ACC > 73% and AUC > 0.74 in NC vs. ASD classification task).

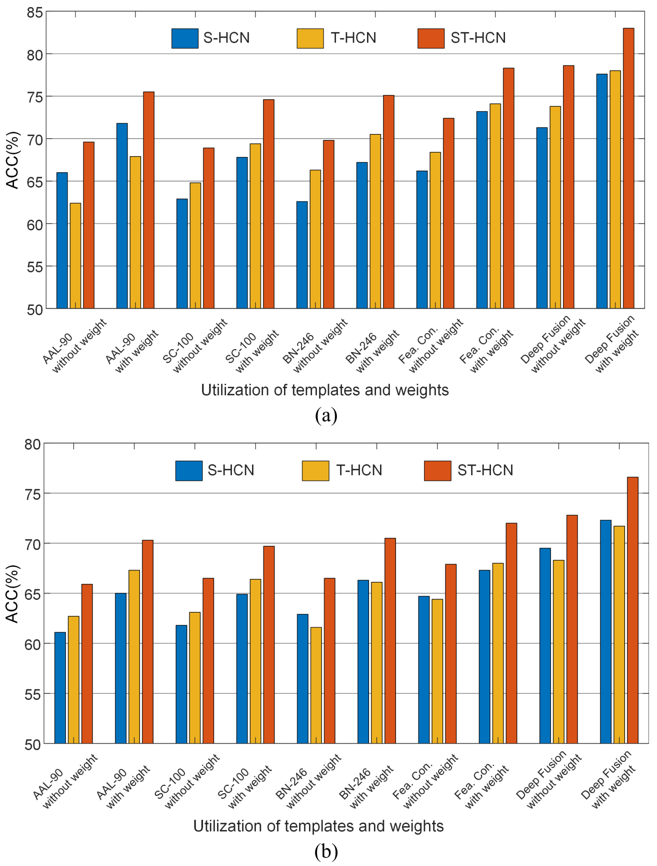


Fig. 3. The comparison of ACC between different HCN models in (a) NC vs. EMCI vs. LMCI and (b) NC vs. ASD classification tasks.

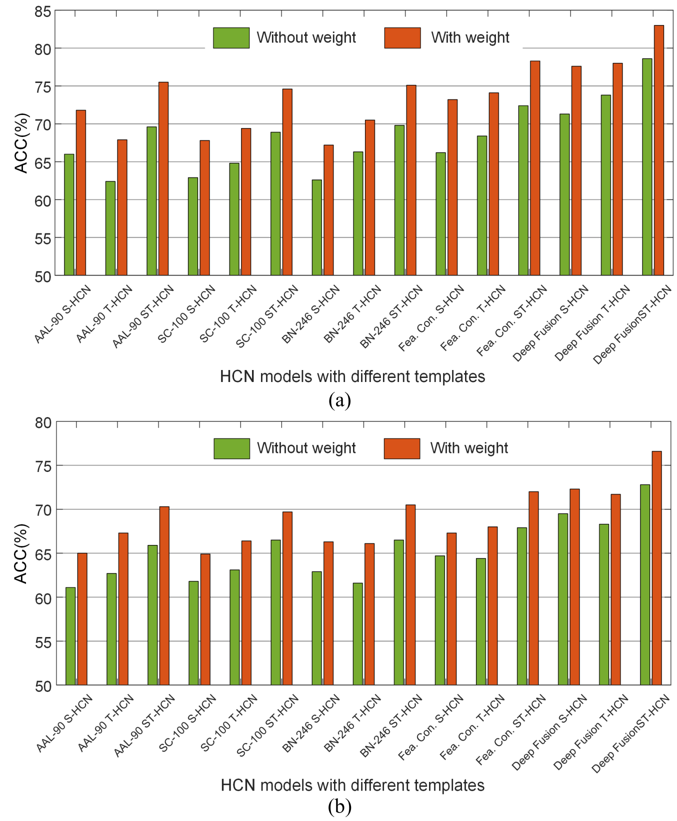


Fig. 4. The effect of the weights of hyper-edges on ACC in (a) NC vs. EMCI vs. LMCI and (b) NC vs. ASD classification tasks.

## IV. Discussion

### A. Efficacy of the Spatio-Temporal Hyper-Correlation Learning

In view of that most existing HCN models are spatial HCN losing sight of the temporal dependency among multiple time points, we propose the ST-HCN to provide complementary spatio-temporal hyper-correlation representations. To validate the effectiveness of spatial and temporal hyper-edges in improving performance, we compare the ST-HCN with methods that solely use either spatial or temporal hyper-edges (*i.e.*, S-HCN and T-HCN) in ten frameworks. The ten frameworks include AAL-90, SC-100, BN-246, AAL-90 + SC-100 + BN-246 (Feature Concatenation), and AAL-90 + SC-100 + BN-246 (Deep Fusion) frameworks, and whether they use the weights of hyper-edges respectively. The comparison results of NC vs. EMCI vs. LMCI classification tasks are shown in Fig. 3(a). In ten frameworks, ST-HCN improves ACCs by 3.6%, 3.7%, 6.0%, 6.8%, 7.2%, 7.9%, 6.2%, 5.1%, 7.3% and 5.4% than S-HCN respectively, proving the importance of temporal hyper-correlation information. In addition, compared with T-HCN, the ST-HCN gains ACC improvement of 7.2%, 7.6%, 4.1%, 5.2%, 3.5%, 4.6%, 4.0%, 4.2%, 4.8% and 5.0% in ten frameworks, respectively. Analogously, Fig. 3(b) shows the comparison results of NC vs. ASD classification tasks. S-HCN contributes to improvements of 3.2%, 3.0%, 3.4%, 3.3%, 4.9%, 4.4%, 3.5%, 4.0%, 4.5% and 4.9% in ACC in ten frameworks, respectively. Meanwhile, the ACC increase achieved by incorporating T-HCN in ten frameworks is 4.8%, 5.3%, 4.7%, 4.8%, 3.6%, 4.2%, 3.2%, 4.7%, 3.3% and 4.3%, respectively. It indicates that the more comprehensive spatio-temporal hyper-correlation learning strategy is superior to using single-domain representations.

### B. Efficacy of the Weights of Hyper-Edges

The proposed method calculates the weights of spatial and temporal hyper-edges in the STW-HCN, and incorporates them into the HCN learning process via the STW-MHGCN. By changing the condition that whether incorporate the weights of hyper-edges into HCN construction and learning frameworks, we design the comparison experiment to verify the efficacy of the weights. As shown in Fig. 4(a), we explore the effect of the weights in fifteen HCN frameworks in NC vs. EMCI vs. LMCI classification tasks. The fifteen frameworks include S-HCN, T-HCN, and ST-HCN models with five different templates respectively. In AAL-90 framework, the three models (*i.e.*, S-HCN, T-HCN, and ST-HCN) with weights yield 5.8%, 5.5%, and 5.9% improvement of ACC than those without weights, respectively. In SC-100 framework, the ACC increase of 4.9%, 4.6%, and 5.7% is achieved by adopting the weights. In BN-246 framework, the ACC increase of 4.6%, 4.2%, and 5.3% is achieved. Meanwhile, the ACC gains of 7.0%, 5.7%, 5.9%, and gains of 6.3%, 4.2%, 4.4% are obtained in AAL-90 + BN-246 (Feature Concatenation) and AAL-90 + BN-246 (Deep Fusion) frameworks, respectively.

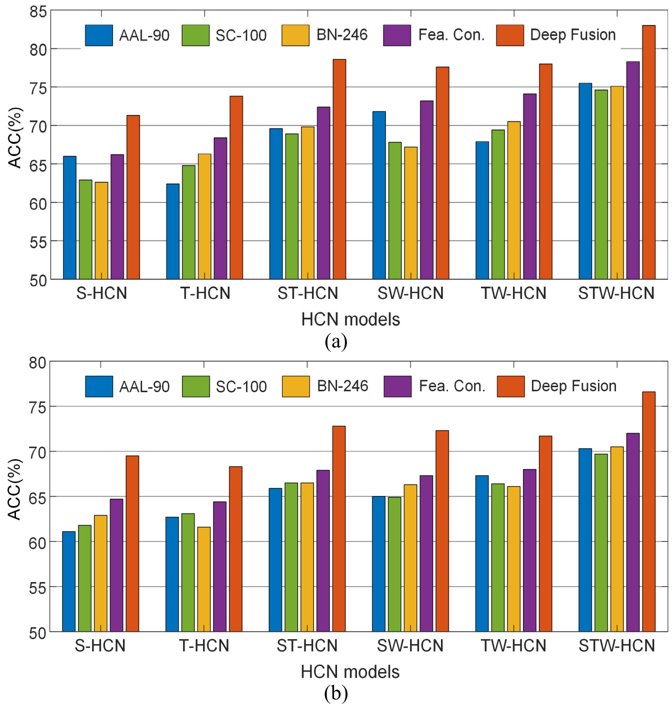


Fig. 5. The comparison of ACC between frameworks using different templates in (a) NC vs. EMCI vs. LMCI and (b) NC vs. ASD classification tasks.

Analogously, Fig. 4(b) shows the comparison results of NC vs. ASD classification tasks. By calculating the weights of hyper-edges, the ACC gains of at least 3.9%, 3.1%, 3.4%, 2.6%, and 2.8% with five different templates, respectively. The significant improvement of ACC confirms the necessity of incorporating the weights of hyper-edges into HCN construction and learning framework.

## C. Efficacy of the Deep Fusion of Multi-Template

To investigate the efficacy of multi-template deep fusion framework, we compare it with AAL-90, SC-100, BN-246, and AAL-90 + SC-100 + BN-246 (Feature Concatenation) methods. Fig. 5(a) shows the NC vs. EMCI vs. LMCI classification performance of different templates in S-HCN, T-HCN, ST-HCN, SW-HCN, TW-HCN, and STW-HCN models. Compared with the single template methods, the deep fusion framework yields at least 5.3%, 7.5%, 8.8%, 5.8%, 7.5%, and 7.5% improvement of ACC in six HCN models respectively, proving the benefit of multi-template carrying richer information. Furthermore, the deep fusion framework gains ACC improvement of 5.1%, 5.4%, 6.2%, 4.4%, 3.9%, and 4.7% compared to the corresponding AAL-90 + SC-100 + BN-246 (Feature Concatenation) methods. Analogously, Fig. 5(b) shows that the deep fusion method improves ACCs by 6.4% and 4.5% on average than single template methods and feature concatenation methods respectively. As stated in *Section* II-C, our proposed framework deeply fuses multi-template information in the spatial and/or time domain, extracting more enriched and powerful feature representation than the feature concatenation approach.

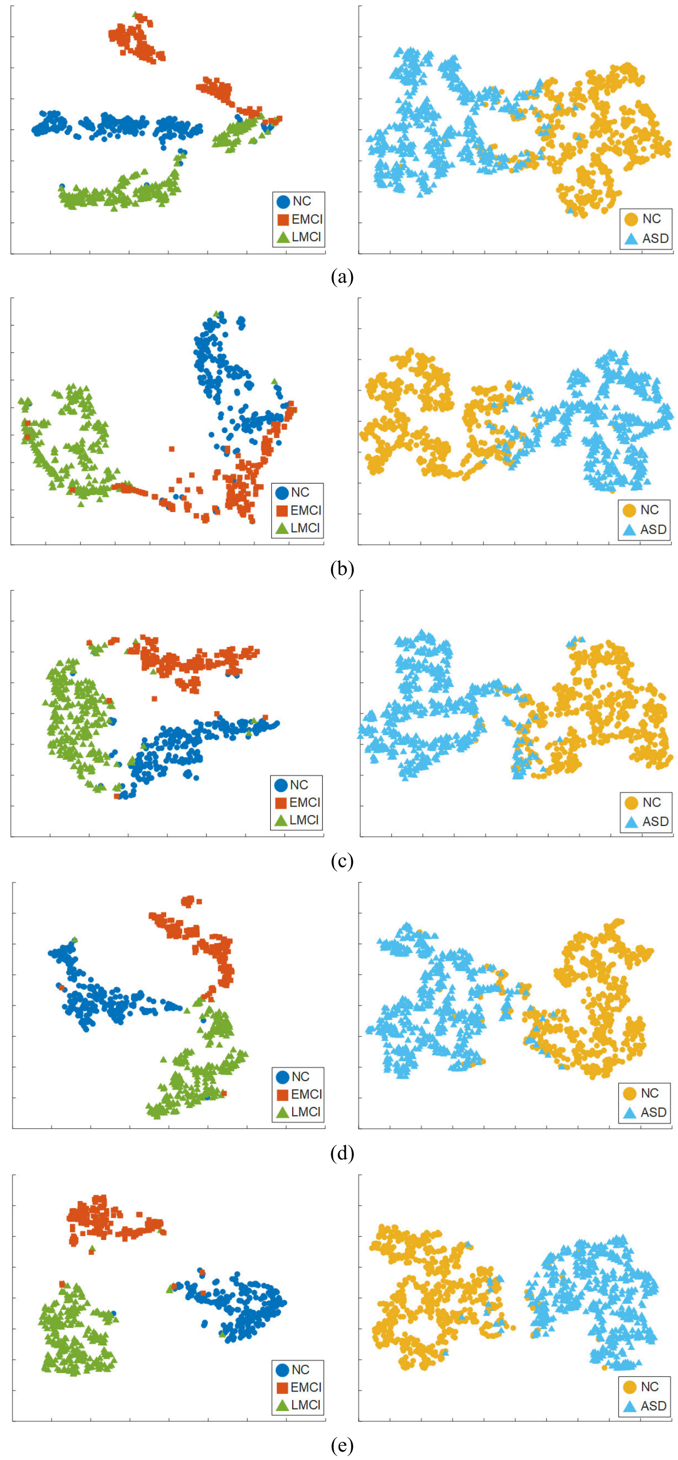


Fig. 6. Visualizations of spatio-temporal features of STW-HCN models with (a) AAL-90, (b) SC-100, (c) BN-246, (d) AAL-90 + SC-100 + BN-246 (Feature Concatenation), and (e) AAL-90 + SC-100 + BN-246 (Deep Fusion).

To further validate the deep fusion framework, we adopt the t-SNE to embed spatio-temporal features of STW-HCN models with different templates into two-dimensional space [37]. The visualization results of NC vs. EMCI vs. LMCI and NC vs. ASD classification tasks are shown in Fig. 6. Compared with the single template methods and the feature concatenation approach, the deep fusion model provides more discriminative

TABLE III
PERFORMANCE COMPARISON USING SEPARATE WEIGHTS AND SHARED WEIGHT

| Task | Metric | Weight-separated | Weight-shared | $p$-values |
|---|---|---|---|---|
| NC vs. EMCI | ACC(%) | 86.5 | 88.9 | 0.022 |
| | AUC | 0.852 | 0.885 | 0.006 |
| | SEN(%) | 88.0 | 89.9 | 0.095 |
| | SPE(%) | 84.6 | 87.8 | 0.023 |
| EMCI vs. LMCI | ACC(%) | 87.8 | 89.1 | 0.002 |
| | AUC | 0.861 | 0.880 | 0.003 |
| | SEN(%) | 86.4 | 88.1 | 0.025 |
| | SPE(%) | 88.6 | 89.8 | 0.006 |
| NC vs. EMCI vs. LMCI | ACC(%) | 81.6 | 83.0 | 0.018 |
| | AUC | 0.824 | 0.844 | 0.001 |
| | $SEN_N$(%) | 81.6 | 82.8 | 0.135 |
| | $SEN_E$(%) | 82.7 | 83.9 | 0.096 |
| | $SEN_L$(%) | 79.8 | 81.8 | 0.032 |
| NC vs. ASD | ACC(%) | 74.9 | 76.6 | 0.001 |
| | AUC | 0.766 | 0.774 | 0.008 |
| | SEN(%) | 75.4 | 77.1 | 0.001 |
| | SPE(%) | 74.5 | 76.1 | 0.001 |

spatio-temporal features of multi-template for MCI and ASD classification.

## D. Effectiveness of the Weight-Shared Strategy

To demonstrate the effectiveness of the weight-shared strategy adopted in *Section* II-C, we compare the proposed method with the approach that utilizes separate weights for each template. We perform the experiment 20 times, changing the sample division of 10-fold cross-validation each time. As shown in Table III, the proposed method is superior to the weight-separated method on all metrics for four classification tasks. The weight-shared strategy achieves an average improvement in ACC and AUC of 1.7% and 0.020, respectively. Furthermore, as shown in Table III, the weight-shared strategy is proven to be significantly better than the weight-separated method in terms of ACC and AUC for four classification tasks under a 95% confidence interval ($p$-values $< 0.05$) using the two-sample $t$-test.

## E. Most Discriminative Spatio-Temporal Hyper-Edges

We investigate the discriminating ability of spatial and temporal hyper-edges for brain disease identification. The matrix multiplication of $\boldsymbol{H}_s$ and $\boldsymbol{W}_s$ is used to represent spatial weighted hyper-edges. Each column in $\boldsymbol{H}_s\boldsymbol{W}_s$ denotes a spatial weighted hyper-edge. The significance test of correlation between the elements in $\boldsymbol{H}_s\boldsymbol{W}_s$ and sample labels are tested. The spatial weighted hyper-edges containing the elements whose $p$-value $< 0.001$ with FDR correction are considered as the most discriminative hyper-edges. Fig. 7 shows the most discriminative spatial weighted hyper-edges in different templates for MCI classification. The details of ROIs defined in AAL-90, SC-100, and BN-246 templates are listed in [17], [32], and [33], respectively. Fig. 7(a) compares the average hyper-edges with right olfactory cortex (OLF.R) as the centroid ROI among NC, EMCI, and LMCI. It can be observed that the topology of hyper-edge in different populations is consistent, and weights decrease sequentially as the disease progresses.

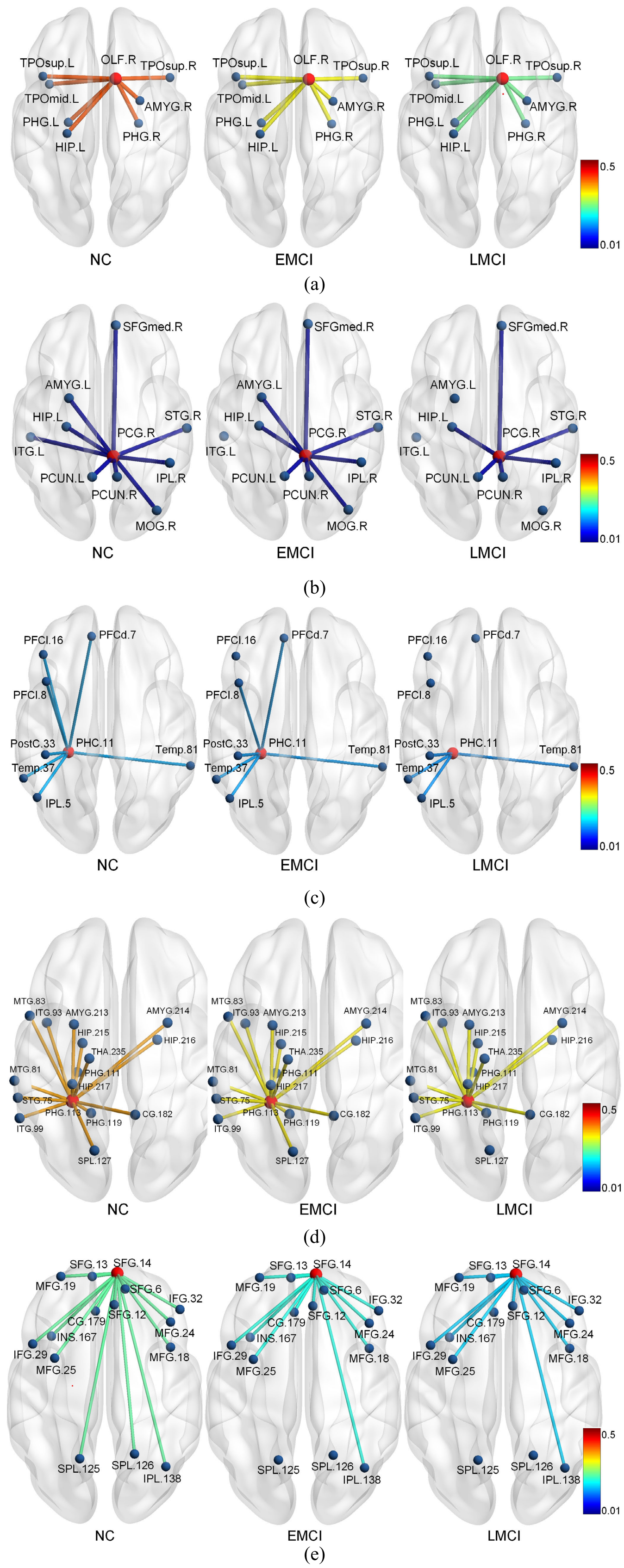


Fig. 7. The most discriminative spatial weighted hyper-edges for MCI classification in (a)-(b) AAL-90, (c) SC-100, and (d)-(e) BN-246 templates, where the red node denotes the centroid ROI and the color of connecting line indicates the weight of hyper-edge.

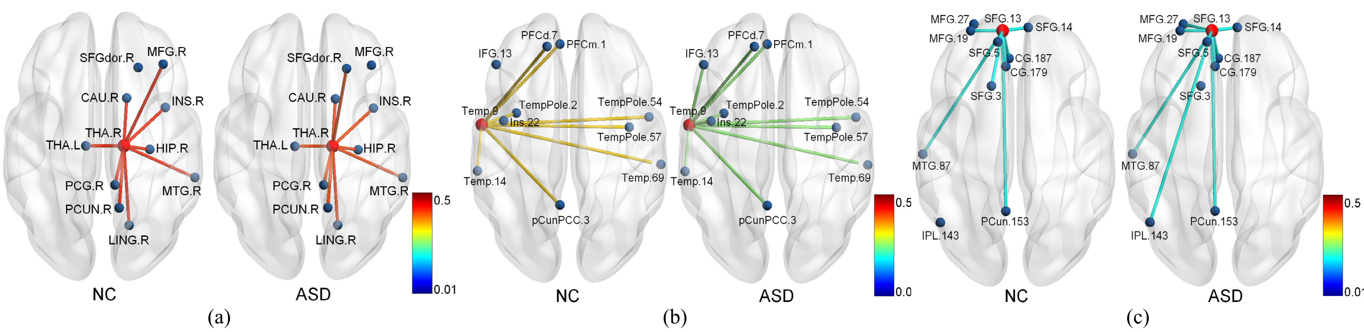


Fig. 8. The most discriminative spatial weighted hyper-edges for ASD classification in (a) AAL-90, (b) SC-100, and (c) BN-246 templates, where the red node denotes the centroid ROI and the color of connecting line indicates the weight of hyper-edge.

The centroid ROI OLF.R and other connected ROIs such as the left hippocampus (HIP.L), bilateral parahippocampal gyri (PHG.L and PHG.R), amygdala (AMYG.R), etc., are reported as the critical ROIs for MCI pathological mechanisms [38], [39], [40]. For example, olfactory defects are common in MCI, which appear earlier than cognitive and memory deficits, and the degeneration of OLF is a biomarker for MCI progression [38]. In addition, the previous study has shown that the HIP, PHG, and AMYG experienced the most salient decreases in gray matter volume for MCI compared to NC [39]. Therefore, the decrease in the weight of hyper-edge composed of these ROIs indicates the progression of MCI. Fig. 7(b) shows the hyper-edge with right posterior cingulate gyrus (PCG.R) as the centroid ROI. There is no significant difference in the weight of this hyper-edge among NC, EMCI, and LMCI. However, compared with NC, the left inferior temporal gyrus (ITG.L) is not connected to the hyper-edge in EMCI and LMCI, which supports previous findings that ITG.L is highly associated with the risk of disease progression in MCI [41]. Moreover, compared with NC and EMCI, the AMYG.L and right middle occipital gyrus (MOG.R) are not connected to PCG.R in the hyper-edge of LMCI. The previous studies report that the gray matter density of PCG, AMYG, and MOG significantly decreases ($p$-value$<$0.0001) simultaneously in AD/MCI patients [42], [43]. Thus, the effectiveness of the most discriminative hyper-edge is proved. Fig. 7(c) shows the hyper-edge with parahippocampal cortex (PHC.11) as the centroid ROI, where 11 denotes the index of ROI in the SC-100 template. The weight of this hyper-edge is similar across different populations. Nevertheless, relative to NC, the lateral prefrontal cortex (PFCl.16) is not connected to the centroid ROI in EMCI and LMCI, and dorsal prefrontal cortex (PFCd.7) and PFCl.8 are not connected to the centroid ROI in LMCI. The aberrant brain activity of PHC and PFC is highly correlated to MCI [44]. Fig. 7(d) displays the hyper-edge of PHG.113 as the centroid ROI, where 113 denotes the index of ROI in the BN-246 template. The index of odd numbers indicates that the ROI is in the left hemisphere, while the even number indicates the right hemisphere. We observed that the weight of hyper-edge is greater in NC than in EMCI and LMCI, while EMCI and LMCI have similar weights. In addition, compared with NC and EMCI, the hyper-edge of PHG.113 is not connected to superior parietal lobule (SPL.127) in LMCI. The prior study finds that both decreased gray matter density in PHG and hypometabolism in SPL occur during the conversion from MCI to AD [45]. Thus, the abnormal hyper-edge of PHG is a biomarker for EMCI and LMCI identification. As shown in Fig. 7(e), the weight of hyper-edge of superior frontal gyrus (SFG.14) decreases gradually in NC, EMCI, and LMCI. The SPL.125 and SPL.126 disconnect to the hyper-edge of SFG.14 in EMCI and LMCI, and the cingulate gyrus (CG.179) disconnect to the hyper-edge in LMCI. As a critical ROI for higher-order cognitive function, the hyper-edge abnormality of SFG can be regarded as a biomarker for MCI analysis. Fig. 8 displays the most discriminative spatial weighted hyper-edges in AAL-90, SC-100, and BN-246 templates for ASD classification. As shown in Fig. 8(a), the weights of the hyper-edges of right thalamus (THA.R) are similar in NC and ASD. Relative to NC, we observe the presence of an atypical connection between THA.R and right dorsolateral superior frontal gyrus (SFGdor.R) in ASD. Conversely, the connection between THA.R and right middle frontal gyrus (MFG.R) disappears. In Fig. 8(b), the weight of the hyper-edge of temporal (Temp.9) in ASD decreases compared with NC. Moreover, the inferior frontal gyrus (IFG.13) and insula (Ins.22) are atypically connected to the hyper-edge of Temp.9 in ASD. The over-connectivity in ASD has been reported in previous studies [46], [47]. Fig. 8(c) shows the hyper-edge with SFG.13 as the centroid ROI. Compared with NC, ASD has increased connections of MFG.27 and inferior parietal lobule (IPL.143) to the centroid ROI, and lost the connection of SFG.3 to the centroid ROI. The aberrant hyper-edges might shed new light on ASD analysis.

In this experiment, we adopt the *Section* II-B method to construct the temporal weighted HCN for each population, respectively. The temporal weighted HCN is represented by the matrix multiplication of $\boldsymbol{H}_t$ and $\boldsymbol{W}_t$, with each column in $\boldsymbol{H}_t\boldsymbol{W}_t$ denotes a temporal weighted hyper-edge. The hyper-edge for each population is denoted by the average result across subjects. The properties of temporal weighted hyper-edges for each population are presented in Table IV, while the results of two-sample $t$-test comparing these properties between different populations are provided in Table V. For the ADNI-2 database, NC exhibits a significantly larger degree of temporal hyper-edge compared to EMCI ($p$-value$<$0.01) and LMCI ($p$-value$<$0.05). Additionally, the temporal hyper-edge in NC is significantly stronger than that in EMCI ($p$-value$<$0.05) and LMCI ($p$-value$<$0.01). The weight

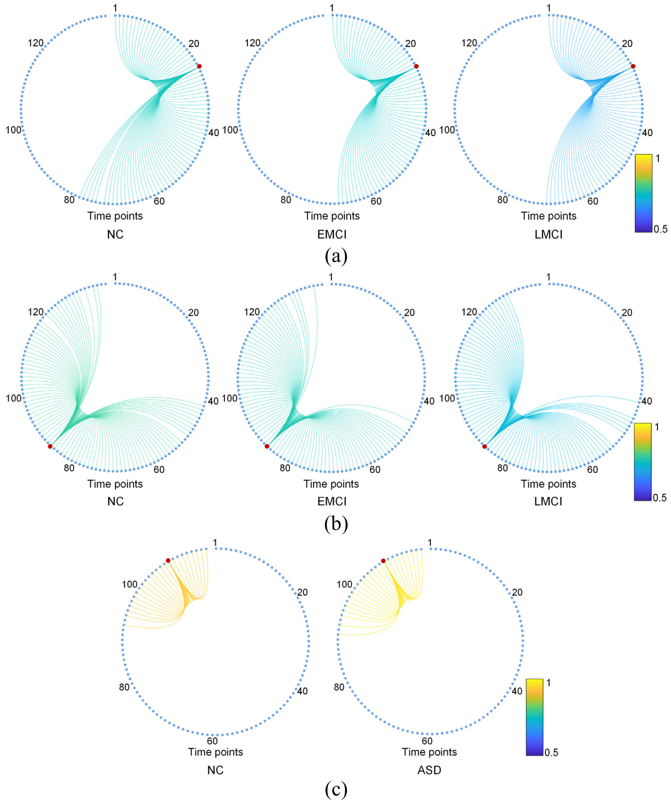


Fig. 9. The most discriminative temporal weighted hyper-edges for (a)-(b) MCI classification in ADNI-2 database and (c) ASD classification in ABIDE-I database, where the red node denotes the centroid node and the color of connecting line indicates the weight of hyper-edge.

of hyper-edge in EMCI is also significantly larger than that in LMCI ($p$-value$<$0.01). On the other hand, for the ABIDE-I database, there is no significant difference observed in degree of hyper-edge between NC and ASD ($p$-value=0.325); however, the weight of hyper-edge in NC is significantly smaller than that in ASD ($p$-value$<$0.01). Furthermore, we conduct a significance test of correlation between the elements in $\boldsymbol{H}_t \boldsymbol{W}_t$ and labels. Fig. 9 shows the most discriminative temporal weighted hyper-edges ($p$-value $<$ 0.001 with FDR correction). In Fig. 9(a), it is evident that the temporal hyper-edge connects a greater number of time points in NC compared to EMCI and LMCI. And the weight of hyper-edge in NC and EMCI is larger than that in LMCI. In Fig. 9(b), a larger degree of hyper-edge is observed in NC than that in EMCI and LMCI, with the weight gradually decreasing from NC to EMCI then LMCI. These results indicate that the reduction of the degrees and weights of the temporal hyper-edges may imply MCI progression. Fig.9 (c) displays the most discriminative temporal hyper-edges for ASD identification. Although the topology of hyper-edge is similar across different populations, ASD possesses the stronger temporal weighted hyper-edge than NC.

### F. Limitations and Future Works

Although the proposed framework is effective for brain disease analysis, there are some limitations and potential enhancements that could be explored in the future. First, we develop the spatial and temporal weighted HCN to capture the high-order relationships among multiple ROIs and

TABLE IV
DEGREES AND WEIGHTS (MEAN ± STANDARD DEVIATION) OF TEMPORAL HYPER-EDGES IN DIFFERENT POPULATIONS

| Properties | ADNI-2 | | | ABIDE-I | |
|---|---|---|---|---|---|
| | NC | EMCI | LMCI | NC | ASD |
| Degree | 74.6±19.5 | 68.8±17.4 | 70.5±19.1 | 31.4±9.4 | 31.9±9.7 |
| Weight | 0.828±0.062 | 0.810±0.066 | 0.766±0.063 | 0.929±0.017 | 0.940±0.016 |

TABLE V
TWO-SAMPLE T-TEST RESULTS ($p$-VALUES) OF DEGREES AND WEIGHTS OF TEMPORAL HYPER-EDGES BETWEEN DIFFERENT POPULATIONS

| Properties | ADNI-2 | | | ABIDE-I |
|---|---|---|---|---|
| | NC vs. EMCI | NC vs. LMCI | EMCI vs. LMCI | NC vs. ASD |
| Degree | $5.81\times10^{-3}$ | 0.043 | 0.227 | 0.325 |
| Weight | 0.012 | $6.21\times10^{-15}$ | $2.12\times10^{-8}$ | $2.26\times10^{-7}$ |

time points, respectively. However, the spatial and temporal hyper-edges reflect the connectivity of the entire scanning period and the whole brain space, respectively. How the spatial hyper-edges change over time points and how the temporal hyper-edges change over ROIs have not been studied yet. In future studies, we will investigate the dynamics of spatial/temporal hyper-edges across time points/ROIs. Second, the predefined function is employed to compute the weights of spatial and temporal hyper-edges, which follows the principle that stronger connectivity between the centroid node and other nodes results in a greater weight of the hyper-edge. Therefore, the benefit of using the predefined function is to fully exploit the connection information between the centroid node and other nodes. Nevertheless, the deep learning framework can provide an end-to-end learning strategy and automatically learn the weight of hyper-edge. However, this will increase the number of parameters that require training. We will conduct a comparative study between the predefined method and the adaptive learning method in the future.

## V. CONCLUSION

In this paper, we propose the STW-HCN to capture the complementary spatio-temporal hyper-correlation among ROIs and time points. Then, the novel STW-MHGCN model is presented to fuse the STW-HCN of multiple templates deeply to obtain collaborative multi-template features. In comparison with state-of-the-art methods, the promising performance in four different classification tasks proves the effectiveness of proposed framework. The abnormal spatial and temporal hyper-edges discovered in this study are meaningful for the MCI and ASD identification and analysis. The deep fusion framework can also be extended to more templates or multi-modal data.

## REFERENCES

[1] 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, Apr. 2021.
[2] P. Yang et al., “Fused sparse network learning for longitudinal analysis of mild cognitive impairment,” *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 233–246, Jan. 2021.

[3] A. K. Malik, M. A. Ganaie, M. Tanveer, P. N. Suganthan, and Alzheimer's Disease Neuroimaging Initiative Initiative, "Alzheimer's disease diagnosis via intuitionistic fuzzy random vector functional link network," *IEEE Trans. Computat. Social Syst.*, early access, Feb. 17, 2022, doi: 10.1109/TCSS.2022.3146974.

[4] X. Song et al., "Graph convolution network with similarity awareness and adaptive calibration for disease-induced deterioration prediction," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101947.

[5] 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.

[6] J. F. Strain et al., "Covariance-based vs. correlation-based functional connectivity dissociates healthy aging from Alzheimer disease," *NeuroImage*, vol. 261, Nov. 2022, Art. no. 119511.

[7] 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.

[8] S. Dautricourt et al., "Dynamic functional connectivity patterns associated with dementia risk," *Alzheimer's Res. Therapy*, vol. 14, no. 1, p. 72, May 2022.

[9] B. Jie, C.-Y. Wee, D. Shen, and D. Zhang, "Hyper-connectivity of functional networks for brain disease diagnosis," *Med. Image Anal.*, vol. 32, pp. 84–100, Aug. 2016.

[10] L. Xiao et al., "Multi-hypergraph learning-based brain functional connectivity analysis in fMRI data," *IEEE Trans. Med. Imag.*, vol. 39, no. 5, pp. 1746–1758, May 2020.

[11] Y. Zhu, X. Zhu, M. Kim, J. Yan, D. Kaufer, and G. Wu, "Dynamic hyper-graph inference framework for computer-assisted diagnosis of neurodegenerative diseases," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 608–616, Feb. 2019.

[12] Y. Li et al., "Multimodal hyper-connectivity of functional networks using functionally-weighted LASSO for MCI classification," *Med. Image Anal.*, vol. 52, pp. 80–96, Feb. 2019.

[13] 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.

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

[15] F. Huang et al., "Self-weighted adaptive structure learning for ASD diagnosis via multi-template multi-center representation," *Med. Image Anal.*, vol. 63, Jul. 2020, Art. no. 101662.

[16] Z. Long et al., "A multi-modal and multi-atlas integrated framework for identification of mild cognitive impairment," *Brain Sci.*, vol. 12, no. 6, p. 751, Jun. 2022.

[17] 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, Jan. 2002.

[18] R. C. Craddock, G. A. James, P. E. Holtzheimer, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Hum. Brain Mapping*, vol. 33, no. 8, pp. 1914–1928, Aug. 2012.

[19] I. Fernandez-Iriondo, A. Jimenez-Marin, B. Sierra, N. Aginako, P. Bonifazi, and J. M. Cortes, "Brain mapping of behavioral domains using multi-scale networks and canonical correlation analysis," *Frontiers Neurosci.*, vol. 16, Jun. 2022, Art. no. 889725.

[20] L. Xiao, J. M. Stephen, T. W. Wilson, V. D. Calhoun, and Y.-P. Wang, "A manifold regularized multi-task learning model for IQ prediction from two fMRI paradigms," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 3, pp. 796–806, Mar. 2020.

[21] B. Lei et al., "Self-calibrated brain network estimation and joint non-convex multi-task learning for identification of early Alzheimer's disease," *Med. Image Anal.*, vol. 61, Apr. 2020, Art. no. 101652.

[22] 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.

[23] Y. Li, J. Liu, J. Huang, Z. Li, and P. Liang, "Learning brain connectivity sub-networks by group—Constrained sparse inverse covariance estimation for Alzheimer's disease classification," *Frontiers Neuroinform.*, vol. 12, p. 58, Sep. 2018.

[24] X. Wang, A. Bao, Y. Cheng, and Q. Yu, "Weight-sharing multi-stage multi-scale ensemble convolutional neural network," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 7, pp. 1631–1642, Jul. 2019.

[25] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, San Diego. CA, USA, 2015, pp. 1–15.

[26] 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.

[27] C. Craddock et al., "The neuro bureau preprocessing initiative: Open sharing of preprocessed neuroimaging data and derivatives," *Frontiers Neuroinformatics*, vol. 7, no. 27, p. 5, 2013.

[28] J. Ashburner et al., "SPM12 manual," Funct. Imag. Lab., Wellcome Trust Centre Neuroimag., Inst. Neurology, UCL, London, U.K., Tech. Rep., 2014.

[29] 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.

[30] X. Deng, J. Zhang, R. Liu, and K. Liu, "Classifying ASD based on time-series fMRI using spatial–temporal transformer," *Comput. Biol. Med.*, vol. 151, Dec. 2022, Art. no. 106320.

[31] A. Schaefer et al., "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI," *Cerebral Cortex*, vol. 28, no. 9, pp. 3095–3114, Sep. 2018.

[32] R. Kong et al., "Individual-specific areal-level parcellations improve functional connectivity prediction of behavior," *Cerebral Cortex*, vol. 31, no. 10, pp. 4477–4500, Aug. 2021.

[33] 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, Aug. 2016.

[34] Y. Zhang et al., "Strength and similarity guided group-level brain functional network construction for MCI diagnosis," *Pattern Recognit.*, vol. 88, pp. 421–430, Apr. 2019.

[35] X. Li et al., "BrainGNN: Interpretable brain graph neural network for fMRI analysis," *Med. Image Anal.*, vol. 74, Dec. 2021.

[36] 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.

[37] A. C. Belkina, C. O. Ciccolella, R. Anno, R. Halpert, J. Spidlen, and J. E. Snyder-Cappione, "Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets," *Nature Commun.*, vol. 10, no. 1, Nov. 2019.

[38] M. M. Vasavada et al., "Olfactory cortex degeneration in Alzheimer's disease and mild cognitive impairment," *J. Alzheimer's Disease*, vol. 45, no. 3, pp. 947–958, Mar. 2015.

[39] P. J. Raine and H. Rao, "Volume, density, and thickness brain abnormalities in mild cognitive impairment: An ALE meta-analysis controlling for age and education," *Brain Imag. Behav.*, vol. 16, no. 5, pp. 2335–2352, Oct. 2022.

[40] B. Chen, "Abnormal cortical regions and subsystems in whole brain functional connectivity of mild cognitive impairment and Alzheimer's disease: A preliminary study," *Aging Clin. Experim. Res.*, vol. 33, no. 2, pp. 367–381, Feb. 2021.

[41] K. Ota, N. Oishi, K. Ito, and H. Fukuyama, "Prediction of Alzheimer's disease in amnestic mild cognitive impairment subtypes: Stratification based on imaging biomarkers," *J. Alzheimer's Disease*, vol. 52, no. 4, pp. 1385–1401, Jun. 2016.

[42] Z. Dai et al., "Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3)," *NeuroImage*, vol. 59, no. 3, pp. 2187–2195, Feb. 2012.

[43] Z. Long et al., "A support vector machine-based method to identify mild cognitive impairment with multi-level characteristics of magnetic resonance imaging," *Neuroscience*, vol. 331, pp. 169–176, Sep. 2016.

[44] Y. J. Lin et al., "Temporal dynamic changes of intrinsic brain activity associated with cognitive reserve in prodromal Alzheimer's disease," *J. Alzheimers Dis.*, vol. 81, no. 3, pp. 1295–1304, Jun. 2021.

[45] S. Morbelli et al., "Mapping brain morphological and functional conversion patterns in amnestic MCI: A voxel-based MRI and FDG-PET study," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 37, no. 1, pp. 36–45, Jan. 2010.

[46] M. D. Shen et al., "Atypical lexicosemantic function of extrastriate cortex in autism spectrum disorder: Evidence from functional and effective connectivity," *NeuroImage*, vol. 62, no. 3, pp. 1780–1791, Sep. 2012.

[47] A. Di Martino et al., "Aberrant striatal functional connectivity in children with autism," *Biol. Psychiatry*, vol. 69, no. 9, pp. 847–856, May 2011.