Discovering Brain Network Dysfunction in Alzheimer's Disease Using Brain Hypergraph Neural Network

Hongmin Cai[1], Zhixuan Zhou[1], Defu Yang[2], Guorong Wu[2], and Jiazhou Chen[1(✉)]

[1] School of Computer Science and Engineering, South China University of Technology, Guangzhou, China
csjzchen@scut.edu.cn

[2] Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, USA

Abstract. Previous studies have shown that neurodegenerative diseases, specifically Alzheimer's disease (AD), primarily affect brain network function due to neuropathological burdens that spread throughout the network, similar to prion-like propagation. Therefore, identifying brain network alterations is crucial in understanding the pathophysiological mechanism of AD progression. Although recent graph neural network (GNN) analyses have provided promising results for early AD diagnosis, current methods do not account for the unique topological properties and high-order information in complex brain networks. To address this, we propose a brain network-tailored hypergraph neural network (BrainHGNN) to identify the propagation patterns of neuropathological events in AD. Our BrainHGNN approach constructs a hypergraph using region of interest (ROI) identity encoding and random-walk-based sampling strategy, preserving the unique identities of brain regions and characterizing the intrinsic properties of the brain-network organization. We then propose a self-learned weighted hypergraph convolution to iteratively update node and hyperedge messages and identify AD-related propagation patterns. We conducted extensive experiments on ADNI data, demonstrating that our BrainHGNN outperforms other state-of-the-art methods in classification performance and identifies significant propagation patterns with discriminative differences in group comparisons.

Keywords: Hypergraph Neural Network · Alzheimer's Disease · Brain Network · Propagation Patterns

1 Introduction

Alzheimer's disease (AD) is an irreversible and progressive neurodegenerative disorder, and it is the most common cause of dementia. Early diagnosis and appropriate interventions play a vital role in managing AD [22]. Existing studies suggest that AD is a disconnection syndrome manifesting the brain

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14224, pp. 230–240, 2023.
https://doi.org/10.1007/978-3-031-43904-9_23

network alterations caused by neuropathological processes before the onset of clinical symptoms [14,15]. Therefore, identifying the propagation patterns of neuropathological burdens provides a new window to comprehend the pathophysiological mechanism of AD and predict the early stage of AD. Recent advances in neuroimaging techniques, such as diffusion-weighted imaging (DWI), allow us to observe the fiber bundles between two anatomical regions *in-vivo*, thereby encoding the structural brain network into a graph data structure [13,17].

In recent years, brain network analyses via graph neural networks have been widely used in brain disease diagnosis [10,12,21]. The advantage of graph convolution is that it takes the topological information of the graph into account, so that the information of neighbor nodes can be integrated into the target nodes via a message-passing scheme to upgrade their discriminative performance. However, these GNN-based approaches assume that the relationship between nodes is pairwise, which ignores the high-order relationships widely existing in brain networks (e.g., functional interactions between multiple brain regions) [8]. To address this limitation, hypergraph, a special graph structure where one hyperedge can contain multiple nodes, is proposed to capture the high-order information in graph [24]. Moreover, Feng et al. [2] extended the hypergraph learning method to hypergraph neural network (HGNN) to learn an optimal data representation. In neuroscience, Ji et al. [7] proposed a hypergraph attention network (FC-HAT) consisting of a dynamic hypergraph generation phase and a hypergraph attention aggregation phase to classify the functional brain networks.

Although HGNN methods have achieved promising progress in many fields, such approaches have three major limitations for brain network analyses. (1)*Lack of consideration of the unique topological properties in the brain network.* The brain networks are cost-efficient small-world networks, which contain a series of organization patterns such as hub nodes and hierarchical modularity. However, current hypergraph constructed using K-nearest neighbors (KNN) exhibit a lack of power to extract specific high-order topological information from the brain network. (2)*Lack of a robust ROI-aware encoding to overcome the anonymity of nodes in HGNN.* Current GNN or HGNN methods are permutation invariant, indicating that the node order in a graph is insensitive to the performance of the graph neural network models. However, every brain region in the brain network has its specific brain function and location; such permutation invariance and anonymity are problematic for brain network analyses. (3)*Lack of an appropriate mechanism to identify brain network dysfunction.* Compared to identifying the disease-related brain regions, less attention has been paid to detecting the propagation patterns of neuropathological burdens using the HGNN methods, despite the brain network alterations with strong interpretation being more attractive to neuroimaging researchers.

To overcome these limitations, we propose a hypergraph neural network explicitly tailored for brain networks to perform (1) identification of propagation patterns of neuropathology and (2) early diagnosis of AD. In this study, we adopt a second-order random walk on a brain network reference to generate two groups

of hyperedges to depict the topology of the brain network. Additionally, an ROI identity encoding module is incorporated into our model to avoid the influence of anonymity of nodes during HGNN convolution on brain network analysis. Furthermore, we design a self-learned weighted hypergraph convolution to extract the discriminative hyperedges, which can be used to characterize the spreading pathways of neuropathological events in AD. The framework of our BrainHGNN is shown in Fig. 1. We have evaluated the effectiveness and robustness of our proposed BrainHGNN on neuroimaging data from the ADNI database. Compared to other methods, the BrainHGNN achieves enhanced discriminative capacity for CN, EMCI, and LMCI classification, as well as illustrates its potential for discovering the putative propagation patterns of neuropathology in AD.

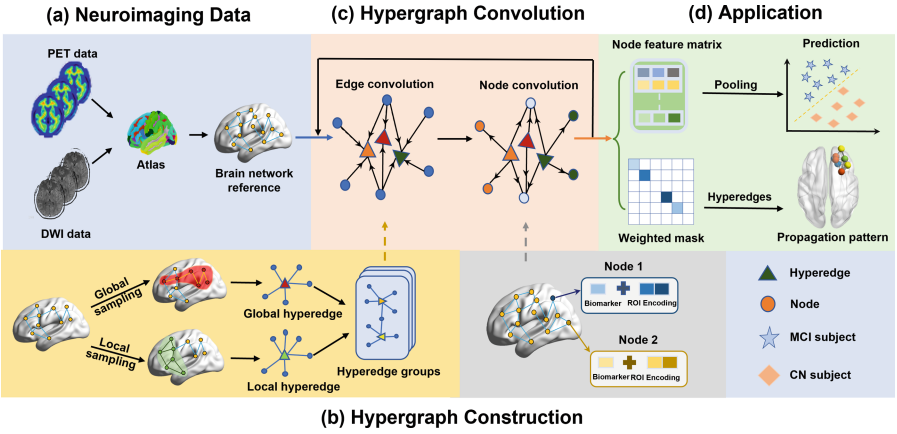

Fig. 1. Framework of our BrainHGNN for brain network analyses. (a) Utilizing DWI image and PET image data to construct the common brain network reference and its initial node feature respectively. (b) For hypergraph generation, applying biased random walk to form two hyperedge groups and then combining them together. For node feature, ROI encoding is concatenated with initial feature to obtain augmented feature. (c) Updating node feature and edge weight through two-stage edge-node convolution. (d) Classification and interpretation according to output of convolution.

2 Methodology

2.1 Hypergraph Construction

A brain network is often encapsulated into an adjacency matrix $\mathbf{A}_s \in \mathbb{R}^{N\times N}(s = 1, \cdots, S)$, where N stands for the number of brain regions in the brain network and S denotes the number of samples. To characterize the spreading pathway of neuropathological events in AD, we first calculate a group-mean adjacency matrix $\bar{\mathbf{A}} = \frac{1}{S}\sum_{s=1}^{S}\mathbf{A}_s$ as a brain network reference from a population of brain networks. Since complex brain functions are usually carried out

by the interaction of multiple brain regions, a hypergraph neural network is introduced as the backbone of our proposed model to capture high-order relationships. A hyperedge in hypergraph connects more than two nodes, characterizing the high-order relationships. Let $G = \{\mathcal{V}, \mathcal{E}, \mathbf{W}\}$ denote a hypergraph with node set $\mathcal{V}$ and hyperedge set $\mathcal{E}$, where each hyperedge is actually a subset of $\mathcal{V}$ randomly sampled based on $\bar{\mathbf{A}}$. Each diagonal element in the diagonal matrix $\mathbf{W}$ represents a weight of hyperedge. The hypergraph structure is represented by an incident matrix $\mathbf{H} \in \mathbb{R}^{|\mathcal{V}|\times|\mathcal{E}|}$, where $\mathbf{H}(v, e) = 1$ means v node is in the e hyperedge. Moreover, the diagonal matrices of the hyperedge degrees $\mathbf{D}_e$ and the vertex degrees $\mathbf{D}_v$ can be calculated by $\delta(e) = \sum_{v\in\mathcal{V}} \mathbf{H}(v, e)$ and $d(v) = \sum_{e\in\mathcal{E}} w(e)\mathbf{H}(v, e)$, respectively.

Node Representation. Empirical biomarkers (such as deposition of amyloid plaques) calculated on PET neuroimaging data are the hallmarks of AD. We obtain the standard uptake value ratio for each brain region and represent them as a column vector $\mathbf{f}_s \in \mathbb{R}^N$, with each element corresponding to the feature of the node. However, the structural brain network has specific regions with distinct functions and messages, making the permutation invariance of hypergraph neural networks less suitable. We address this by generating a region-of-interest (ROI) identity encoding $\mathbf{p} \in \mathbb{R}^{N\times m}$ for each brain region and concatenating it with the empirical biomarkers, thereby reducing permutation invariance and anonymity of nodes in graph convolution. The parameter m denotes the dimension of identity encoding and is set to 2 in our experiments. The final node representation is formulated as,

$$\mathbf{x}_s^0 = Concat(\mathbf{f}_s^0, \mathbf{p}) \tag{1}$$

where $\mathbf{f}_s^0$ is a vector of the initial input biomarkers and $\mathbf{p}$ is a learnable matrix to encode the regional information.

Hyperedge Generation. The brain, a complex network, exhibits information processing with maximum efficiency and minimum costs. Such brain organization requires the brain network to be equipped with both high local clustering (hierarchical modularity) and global efficiency (shortest path length) [18]. In this context, inspired by the random walk sampling in [6], we propose a second-order biased random walk sampling strategy to each node of the weighted graph $\bar{\mathbf{A}}$ to generate a group of local hyperedges $\mathbf{H}_{local} \in \mathbb{R}^{N\times N}$ and a group of global hyperedges $\mathbf{H}_{global} \in \mathbb{R}^{N\times N}$, respectively. Specifically, the transition probability π_{ux} on edge (u, x) with source node u is decided by edge weight $\bar{\mathbf{A}}(u, x)$ and a bias term $\alpha_p(v, x)$ related to source node v of previous step, that is,

$$\pi_{ux} = \alpha_p(v, x) \cdot \bar{\mathbf{A}}(u, x) \tag{2}$$

where $\alpha_p(v, x) = p^{d_{vx}-1}, d_{vx} \in \{0, 1, 2\}$ and d_{vx} is the shortest path distance between node v and candidate target x. In Eq. (2), we can observe that when p sets to a small value $(0 < p < 1)$, node u tends to transfer to the neighbors of node v, thus forming a local network community. However, when p takes a large value $(p > 1)$, node u has a high probability of moving to a node far away

from node v, thereby generating a global network community. Therefore, the hyperedge generation can be described mainly in three steps. (1) Calculate the transition probability matrix with the original edge weight matrix and parameter p according to Eq. (2). (2) For each node, set it as the beginning of the walk, and calculate which neighbor to be traversed according to transition probability. Then we sample this neighbor to the hyperedge and set this neighbor as the current node. Repeat this procedure l times to generate a hyperedge. Iterate all nodes to construct a hypergraph. (3) Execute step 1–2 for local sampling and global sampling to generate local and global hypergraphs respectively. Fuse two hypergraphs by concatenation to form the final hypergraph. In the following experiments, we set walk length $l_{local} = 14$ which is 1/4 average node degree and parameter $p = 1/4$ to generate local hyperedge group, as well as apply walk length $l_{global} = 23$ which is 1/4 of the average of shortest path length and parameter $p = 4$ to construct global hyperedge group. Finally, the incidence matrix $\mathbf{H} \in \mathbb{R}^{N \times 2N}$ can be obtained by concatenating these two hyperedge groups, i.e., $\mathbf{H} = Concat(\mathbf{H}_{local}, \mathbf{H}_{global})$.

2.2 Hypergraph Convolution

As stated in [4], the hypergraph convolution can be decoupled as a two-stage message passing procedure, which can be formulated as:

$$\begin{aligned} \mathbf{Z}^t &= \mathbf{W}\mathbf{H}^\top \mathbf{D}_e^{-1} \mathbf{X}^t \\ \mathbf{X}^{t+1} &= \sigma(\mathbf{D}_v^{-1} \mathbf{H} \mathbf{Z}^t \mathbf{\Theta}^{t+1}) \end{aligned} \tag{3}$$

where $\mathbf{X}^t \in \mathbb{R}^{|\mathcal{V}| \times M_t}$ is the node feature matrix in t^{th} layer and $\mathbf{Z}^t \in \mathbb{R}^{|\mathcal{E}| \times M_t}$ is the feature matrix of hyperedges in t^{th} layer. $\mathbf{\Theta}^{t+1} \in \mathbb{R}^{M_t \times M_{t+1}}$ is the learning parameter in $(t+1)^{th}$ layer. The intuition behind Eq. (3) is that we first use the incidence matrix $\mathbf{H}$ to guide each node to aggregate and generate the hyperedge feature matrix $\mathbf{Z}^t$, then the updated node feature matrix $\mathbf{X}^{t+1}$ can be obtained by aggregating node-relevant hyperedge features with a learning parameter $\mathbf{\Theta}^{t+1}$ and a nonlinear activation function $\sigma(\cdot)$.

In most cases, the hyperedge weight matrix $\mathbf{W}$ is pre-defined as either an identical matrix $\mathbf{I}$ or a diagonal matrix with specific elements determined by prior knowledge, for ease of use. Nevertheless, these two types of hyperedge weight matrices pose difficulties in identifying the discriminant hyperedges associated with AD-related brain network dysfunction. To address this issue, we propose using a learnable weight matrix mask to guide the learning of hyperedge weights, enabling us to identify propagation patterns of neuropathological burden in AD. The hyperedge convolution can be rewritten as,

$$\mathbf{Z}^t = \tilde{\mathbf{W}}\mathbf{H}^\top \mathbf{D}_e^{-1} \mathbf{X}^t \tag{4}$$

where $\tilde{\mathbf{W}} = sigmoid(\tilde{\mathbf{M}} \circ \mathbf{W})$ is the learned hyperedge weight matrix, $\tilde{\mathbf{M}}$ is the learnable weight mask and $\circ$ means Hadamard product.

2.3 Optimization

In this study, we use the cross entropy loss to train our model and predict disease status. To avoid overfitting and discover AD-related propagation patterns, we first impose an l_1-norm constraint to the learnable weight mask $\tilde{\mathbf{M}}$ to overshadow the irrelevant hyperedges and preserve the discriminative hyperedges that are highly related to the prediction of labels. Then, we require $\mathbf{\Theta}$ to be smooth with an l_2-norm constraint to void its divergence. Finally, the loss function can be formulated as follows:

$$loss_{ce} = \frac{1}{N}\sum_{i=1}^{N} -[y_i \cdot \log(z_i) + (1-y_i) \cdot \log(1-z_i)] + \lambda_1 ||\tilde{\mathbf{M}}||_1 + \lambda_2 ||\mathbf{\Theta}||_2 \quad (5)$$

where z_i is the predictive label of i^{th} sample output from the fully connected layer by our BrainHGNN and y_i is the ground truth of i^{th} sample.

3 Experiments

3.1 Dataset Description and Experimental Settings

Data Processing. We collect 94 structural brain networks from the ADNI database (https://adni.loni.usc.edu) to calculate a group-mean adjacency matrix $\bar{\mathbf{A}}$ to construct the incident matrix $\mathbf{H}$ of the common hypergraph. For each subject, we construct the structural brain network in the following steps. First, we apply Destrieux atlas [1] to the T1-weighted MR to obtain 148 ROIs. Then we apply surface seed-based probabilistic fiber tractography [3] to the DWI image to generate a adjacency matrix $\mathbf{A}_s$. Regarding the feature representation of each ROI, we adopt a similar image processing pipeline to pracellate the ROIs and calculate the standard update value ratio (SUVR) score of each ROI for the amyloid-PET and FDG-PET data and represent them as a column vector $\mathbf{f}_s$ of whole-brain pathological events. The detailed information on multi-modal neuroimaging data is shown in Table 1.

Parameter Settings. In our experiments, we set parameters $num_{layer} = 1$, $num_{epoch} = 250$, $batchsize = 256$, $learning\,rate = 0.05$. LeakyReLu activation function $\sigma(\cdot)$ is used for node convolution. The l_1-norm is adopted on the learnable weight mask $\tilde{\mathbf{M}}$ with a hyper-parameter $\lambda_1 = 1e-3$ to extract the discriminative hyperedges. The l_2-norm is applied on the learning parameter matrix $\mathbf{\Theta}$ with a hyper-parameter $\lambda_2 = 1e-5$.

Baselines. The comparison baselines incorporate two traditional methods, including support vector machines (SVM) and random forest (RF), as well as five graph-based methods, including graph convolution neural network (GCN) [9], graph attention network (GAT) [19], hypergraph neural network (HGNN) [4], Dynamic Graph CNN (DGCNN) [20], and BrainGNN [11].

Table 1. Demographic Information of Amyloid-PET and FDG-PET Data.

PET	Gender	Number	Range of Age	Average Age	CN	EMCI	LMCI
Amyloid	Male	450	55.0–91.4	73.4	136	184	130
	Female	389	55.0–89.6	71.7	148	145	96
	Total	**839**	**55.0–91.4**	**72.6**	**284**	**329**	**226**
FDG	Male	592	55.0–91.4	73.9	169	182	241
	Female	472	55.0–89.6	72.2	166	148	158
	Total	**1064**	**55.0–91.4**	**73.1**	**335**	**330**	**399**

3.2 Evaluating Diagnostic Capability on Amyloid-PET and FDG-PET Data

In this section, our objective is to evaluate the diagnostic performance of our BrainHGNN model. We first construct three group comparisons, including CN/EMCI, EMCI/LMCI, and CN/LMCI, for amyloid-PET or FDG-PET data. For each comparison, we apply our BrainHGNN on empirical biomarkers to evaluate the classification metrics, including accuracy, F1-score, sensitivity, and specificity, according to 10-fold cross validation. From Table 2, it is obvious that (1) our proposed method achieves the highest classification performance in most of the metrics on both amyloid-PET and FDG-PET data, compared to the other methods; (2) the graph-based deep learning methods (GCN, GAT, DGCNN, BrainGNN, HGNN, and BrainHGNN) are generally superior to traditional machine learning methods (SVM and RF), mainly because graph-based methods fully incorporate the topology of the brain network, thus enhancing the classification performance; (3) the BrainHGNN constructs the hypergraph customized to the brain network's unique topology, which is more suitable for brain network analyses than the traditional hypergraph neural networks.

3.3 Evaluating the Statistical Power of Identifying Brain Network Dysfunction in AD

In this section, we investigate the capability of identifying AD-related discriminative hyperedges by our BrainHGNN. Benefiting from the sparse weight mask, we can extract four discriminative hyperedges that most frequently occurred in the top-10 highest weight hyperedge list. Here, we take CN/LMCI group comparison as an example and map the discriminative hyperedges on the cortical surface for amyloid-PET data (first row in Fig. 2(a)) and FDG-PET data (second row in Fig. 2(a)). We can observe that most of nodes in these discriminative hyperedges are located in the Default Mode network, Cingulo-Opercular network, FrontoParietal network, DorsalAttention network, and VentralAttention network. Numerous studies have demonstrated a significant association between these subnetworks and the progress of AD [5,16,23].

Table 2. The Classification Results on Amyloid-PET and FDG-PET Data

Data	GROUP	MODEL	ACCURACY	F1-SCORE	SENSITIVITY	SPECIFICITY
Amyloid	CN/LMCI	SVM	0.691 ± 0.057	0.572 ± 0.086	0.476 ± 0.092	0.861 ± 0.064
		RF	0.690 ± 0.064	0.572 ± 0.086	0.555 ± 0.090	0.861 ± 0.064
		GCN	0.709 ± 0.052	0.587 ± 0.119	0.498 ± 0.154	0.863 ± 0.132
		GAT	0.718 ± 0.060	0.620 ± 0.113	0.553 ± 0.167	0.828 ± 0.161
		DGCNN	0.735 ± 0.039	0.578 ± 0.058	0.491 ± 0.077	0.841 ± 0.065
		BrainGNN	0.745 ± 0.079	0.615 ± 0.115	0.577 ± 0.122	**0.898 ± 0.088**
		HGNN	0.706 ± 0.056	0.567 ± 0.114	0.460 ± 0.143	0.891 ± 0.112
		BrainHGNN	**0.751 ± 0.040**	**0.651 ± 0.089**	**0.583 ± 0.070**	0.873 ± 0.075
	EMCI/LMCI	SVM	0.628 ± 0.067	0.353 ± 0.109	0.257 ± 0.094	0.886 ± 0.059
		RF	0.607 ± 0.059	0.396 ± 0.105	0.333 ± 0.060	0.803 ± 0.070
		GCN	0.692 ± 0.059	0.498 ± 0.160	0.422 ± 0.194	0.859 ± 0.134
		GAT	0.692 ± 0.048	0.417 ± 0.083	0.332 ± 0.023	**0.912 ± 0.089**
		DGCNN	0.672 ± 0.060	0.451 ± 0.070	0.341 ± 0.074	0.893 ± 0.054
		BrainGNN	0.673 ± 0.072	0.486 ± 0.060	0.412 ± 0.092	0.828 ± 0.162
		HGNN	0.693 ± 0.057	0.499 ± 0.161	0.423 ± 0.098	0.862 ± 0.110
		BrainHGNN	**0.695 ± 0.050**	**0.507 ± 0.139**	**0.455 ± 0.095**	0.865 ± 0.070
	CN/EMCI	SVM	0.550 ± 0.061	0.629 ± 0.060	0.731 ± 0.145	0.373 ± 0.156
		RF	0.581 ± 0.059	0.619 ± 0.066	0.642 ± 0.085	**0.512 ± 0.091**
		GCN	0.621 ± 0.046	0.655 ± 0.126	0.732 ± 0.218	0.461 ± 0.263
		GAT	0.632 ± 0.036	0.648 ± 0.127	0.701 ± 0.023	0.510 ± 0.027
		DGCNN	0.613 ± 0.025	**0.669 ± 0.033**	0.732 ± 0.072	0.470 ± 0.08
		BrainGNN	0.639 ± 0.080	0.629 ± 0.113	0.706 ± 0.166	0.473 ± 0.025
		HGNN	0.626 ± 0.045	0.648 ± 0.133	0.715 ± 0.024	0.485 ± 0.028
		BrainHGNN	**0.647 ± 0.053**	0.668 ± 0.067	**0.733 ± 0.024**	0.504 ± 0.031
FDG	CN/LMCI	SVM	0.553 ± 0.062	0.688 ± 0.047	**0.913 ± 0.061**	0.136 ± 0.083
		RF	0.617 ± 0.049	0.647 ± 0.053	0.655 ± 0.080	0.575 ± 0.086
		GCN	0.658 ± 0.050	0.673 ± 0.097	0.695 ± 0.193	0.585 ± 0.025
		GAT	0.655 ± 0.049	0.657 ± 0.100	0.651 ± 0.197	**0.638 ± 0.259**
		DGCNN	0.639 ± 0.013	0.678 ± 0.050	0.715 ± 0.115	0.539 ± 0.131
		BrainGNN	0.635 ± 0.064	0.624 ± 0.041	0.624 ± 0.133	0.593 ± 0.021
		HGNN	0.648 ± 0.046	0.673 ± 0.111	0.726 ± 0.215	0.516 ± 0.027
		BrainHGNN	**0.670 ± 0.053**	**0.707 ± 0.055**	0.741 ± 0.103	0.577 ± 0.061
	EMCI/LMCI	SVM	0.562 ± 0.064	0.691 ± 0.047	**0.902 ± 0.068**	0.166 ± 0.099
		RF	0.665 ± 0.051	0.696 ± 0.054	0.712 ± 0.070	**0.614 ± 0.081**
		GCN	0.649 ± 0.047	0.678 ± 0.089	0.718 ± 0.192	0.544 ± 0.260
		GAT	0.660 ± 0.029	0.691 ± 0.065	0.734 ± 0.020	0.555 ± 0.029
		DGCNN	0.650 ± 0.033	0.683 ± 0.043	0.723 ± 0.169	0.544 ± 0.193
		BrainGNN	0.639 ± 0.057	0.683 ± 0.070	0.637 ± 0.143	0.560 ± 0.165
		HGNN	0.651 ± 0.045	0.664 ± 0.104	0.684 ± 0.021	0.576 ± 0.028
		BrainHGNN	**0.675 ± 0.037**	**0.713 ± 0.060**	0.764 ± 0.014	0.547 ± 0.081
	CN/EMCI	SVM	0.557 ± 0.071	0.595 ± 0.129	0.712 ± 0.207	0.440 ± 0.213
		RF	0.620 ± 0.058	0.608 ± 0.066	0.605 ± 0.091	0.641 ± 0.082
		GCN	0.629 ± 0.038	0.587 ± 0.149	0.613 ± 0.254	0.599 ± 0.026
		GAT	0.631 ± 0.047	0.638 ± 0.087	0.679 ± 0.186	0.564 ± 0.117
		DGCNN	0.636 ± 0.034	0.559 ± 0.118	0.560 ± 0.160	**0.655 ± 0.106**
		BrainGNN	0.621 ± 0.069	0.557 ± 0.109	0.616 ± 0.187	0.567 ± 0.161
		HGNN	0.616 ± 0.037	0.560 ± 0.017	0.585 ± 0.175	0.596 ± 0.175
		BrainHGNN	**0.639 ± 0.042**	**0.653 ± 0.067**	**0.717 ± 0.155**	0.544 ± 0.179

Discussion. Here, we design an ablation experiment to evaluate the power of each module used in BrainHGNN. We divide the models into five different counterparts, including the HGNN, model without random walk sampling, model without self-learned edge weight, model without ROI encoding, and our BrainHGNN. We apply these five methods on three group comparisons using empirical amyloid SUVR and FDG SUVR to calculate the classification accuracy, as shown

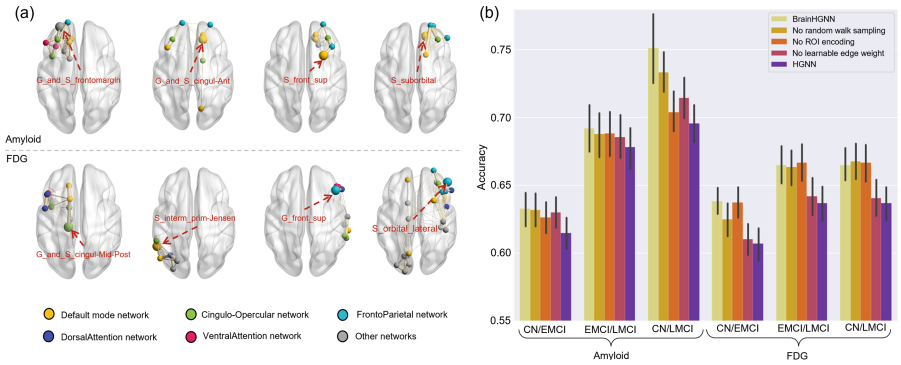

Fig. 2. (a) The brain mappings of four discriminative hyperedges on CN/LMCI group comparison. (b) The classification accuracy of ablation study.

in Fig. 2(b). It is clear that (1) compared to baseline model HGNN, all proposed modules used in HGNN show enhanced classification performance, indicating these modules are effective; (2) the BrainHGNN with full settings achieves best classification performance than other counterparts.

4 Conclusion

In this paper, we propose a random-walk-based hypergraph neural network by integrating the topological nature of the brain network to predict the early stage of AD and discover the propagation patterns of neuropathological events in AD. Compared with other methods, our proposed BrainHGNN achieve enhanced classification performance and statistical power in group comparisons on ADNI neuroimaging dataset. In the future, we plan to apply our BrainHGNN to other neurodegenerative disorders that manifest brain network dysfunction.

Acknowledgements. This work was supported in part by the National Key Research and Development Program of China (2022YFE0112200), the National Natural Science Foundation of China (U21A20520,62102153), the Science and Technology Project of Guangdong Province (2022A0505050014), the Guangdong Key Laboratory of Human Digital Twin Technology (2022B1212010004), Natural Science Foundation of Guangdong Province of China (2022A1515011162), Key-Area Research and Development Program of Guangzhou City (202206030009), and the China Postdoctoral Science Foundation (2021M691062, 2023T160226). The neuroimaging datasets used in this study were supported by the Alzheimer's Disease Neuroimaging Initiative (ADNI).

References

1. Destrieux, C., Fischl, B., Dale, A., Halgren, E.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage **53**(1), 1–15 (2010)
2. Feng, Y., You, H., Zhang, Z., Ji, R., Gao, Y.: Hypergraph neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 3558–3565 (2019)
3. Fillard, P., et al.: Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. Neuroimage **56**(1), 220–234 (2011)
4. Gao, Y., Feng, Y., Ji, S., Ji, R.: HGNN$^+$: general hypergraph neural networks. IEEE Trans. Pattern Anal. Mach. Intell. (2022). https://doi.org/10.1109/tpami.2022.3182052
5. Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V.: Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc. Natl. Acad. Sci. **101**(13), 4637–4642 (2004)
6. Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 855–864 (2016)
7. Ji, J., Ren, Y., Lei, M.: Fc-hat: Hypergraph attention network for functional brain network classification. Inf. Sci. **608**, 1301–1316 (2022)
8. Jie, B., Wee, C.Y., Shen, D., Zhang, D.: Hyper-connectivity of functional networks for brain disease diagnosis. Med. Image Anal. **32**, 84–100 (2016)
9. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
10. Li, X., Dvornek, N.C., Zhou, Y., Zhuang, J., Ventola, P., Duncan, J.S.: Graph neural network for interpreting task-fMRI biomarkers. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11768, pp. 485–493. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32254-0_54
11. Li, X., et al.: BrainGNN: Interpretable brain graph neural network for fMRI analysis. Med. Image Anal. **74**, 102233 (2021)
12. Li, X., et al.: Pooling regularized graph neural network for fMRI biomarker analysis. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12267, pp. 625–635. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59728-3_61
13. Liu, J., et al.: Complex brain network analysis and its applications to brain disorders: a survey. Complexity **2017** (2017)
14. Pievani, M., Filippini, N., Van Den Heuvel, M.P., Cappa, S.F., Frisoni, G.B.: Brain connectivity in neurodegenerative diseases-from phenotype to proteinopathy. Nat. Rev. Neurol. **10**(11), 620–633 (2014)
15. Sepulcre, J., et al.: Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. Nat. Med. **24**(12), 1910–1918 (2018)
16. Sorg, C., et al.: Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. Proc. Natl. Acad. Sci. **104**(47), 18760–18765 (2007)
17. Sporns, O.: Structure and function of complex brain networks. Dialogues Clin. Neurosci. **15**, 247–262 (2013)
18. Stam, C.J.: Modern network science of neurological disorders. Nat. Rev. Neurosci. **15**(10), 683–695 (2014)
19. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., et al.: Graph attention networks. Stat **1050**(20), 10–48550 (2017)

20. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph CNN for learning on point clouds. ACM Trans. Graph. (TOG) **38**(5), 1–12 (2019)
21. Yang, H., et al.: Interpretable multimodality embedding of cerebral cortex using attention graph network for identifying bipolar disorder. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11766, pp. 799–807. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32248-9_89
22. Zhang, L., Wang, M., Liu, M., Zhang, D.: A survey on deep learning for neuroimaging-based brain disorder analysis. Front. Neurosci. **14**, 779 (2020)
23. Zhang, Y., et al.: Joint assessment of structural, perfusion, and diffusion MRI in Alzheimer's disease and frontotemporal dementia. Int. J. Alzheimer's Dis. **2011** (2011)
24. Zhou, D., Huang, J., Schölkopf, B.: Learning with hypergraphs: clustering, classification, and embedding. In: Advances in Neural Information Processing Systems, vol. 19 (2006)