Mapping Multi-modal Brain Connectome for Brain Disorder Diagnosis via Cross-modal Mutual Learning

Yanwu Yang, *Member, IEEE*, Chenfei Ye, Xutao Guo, Tao Wu, Yang Xiang, Ting Ma *Member, IEEE*

***Abstract*— Recently, the study of multi-modal brain connectome has recorded a tremendous increase and facilitated the diagnosis of brain disorders. In this paradigm, functional and structural networks, e.g., functional and structural connectivity derived from fMRI and DTI, are in some manner interacted but are not necessarily linearly related. Accordingly, there remains a great challenge to leverage complementary information for brain connectome analysis. Recently, Graph Convolutional Networks (GNN) have been widely applied to the fusion of multi-modal brain connectome. However, most existing GNN methods fail to couple inter-modal relationships. In this regard, we propose a Cross-modal Graph Neural Network (Cross-GNN) that captures inter-modal dependencies through dynamic graph learning and mutual learning. Specifically, the inter-modal representations are attentively coupled into a compositional space for reasoning inter-modal dependencies. Additionally, we investigate mutual learning in explicit and implicit ways: (1) Cross-modal representations are obtained by cross-embedding explicitly based on the inter-modal correspondence matrix. (2) We propose a cross-modal distillation method to implicitly regularize latent representations with cross-modal semantic contexts. We carry out statistical analysis on the attentively learned correspondence matrices to evaluate inter-modal relationships for associating disease biomarkers. Our extensive experiments on three datasets demonstrate the superiority of our proposed method for disease diagnosis with promising prediction performance and multi-modal connectome biomarker location.**

***Index Terms*— Multi-modal learning, graph neural network, brain disorder, self-distillation, cross-modal**

This work is done by Yanwu Yang during his internship at Peng Cheng Laboratory.

The study is supported by grants from the National Natural Science Foundation of China (62106113, 62276081, and 62106115), Innovation Team and Talents Cultivation Program of National Administration of Traditional Chinese Medicine (NO: ZYYCXTD-C-202004), Basic Research Foundation of Shenzhen Science and Technology Stable Support Program (GXWD20201230155427003-20200822115709001), The National Key Research and Development Program of China (2021YFC2501202), and the Major Key Project of PCL. Corresponding authors: Yang Xiang and Ting Ma.

Yanwu Yang and Xutao Guo are with the School of Electronics and Information Engineering, Harbin Institute of Technology at Shenzhen, Shenzhen, China, and the Peng Cheng Laboratory, Shenzhen, Guangdong, China. (e-mail: 20b952019@stu.hit.edu.cn, 18b952052@stu.hit.edu.cn)

Ting Ma is with the School of Electronics and Information Engineering, Harbin Institute of Technology at Shenzhen, Shenzhen, China, the Peng Cheng Laboratory, Shenzhen, Guangdong, China, Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China, and International Research Institute for Artificial Intelligence, Harbin Institute of Technology (Shenzhen), Shenzhen, China. (e-mail: tma@hit.edu.cn)

Chenfei Ye is with the International Research Institute for Artificial Intelligence, Harbin Institute of Technology at Shenzhen, Shenzhen, China, and the Peng Cheng Laboratory, Shenzhen, Guangdong, China. (e-mail: Chenfei.ye@foxmail.com)

Tao Wu is with the Center for Movement Disorders, Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China, and China National Clinical Research Center for Neurological Diseases, Beijing, China. (e-mail: wutao69@163.com)

Yang Xiang is with the Peng Cheng Laboratory, Shenzhen, Guangdong, China. (e-mail: xiangy@pcl.ac.cn)

I. Introduction

Recently, the connectome or brain network has represented an indispensable foundation for brain disorder research, which enables the exploration of the complex relationship between brain dysfunctions and behavioral phenotypes [1], [2]. Conceptualizing the brain as a network provides a more holistic view of relating abnormal discharge of the brain and dysfunction. Recent investigations have successfully examined and identified disconnectome underlying diseases in neurodegenerative and psychiatric brain diseases, such as Alzheimer's disease (AD) and Parkinson's disease (PD) [3], [4].

Various modalities of magnetic resonance imaging (MRI) and neurophysiological data have been reported to constitute brain networks. Functional networks can be constructed from functional MRI (fMRI) and electroencephalography (EEG), while structural networks are constructed from diffusion tensor imaging (DTI) or diffusion spectrum imaging (DSI) [5]–[7]. Combining functional and structural connectome enables the exploration of brain state by neuron activation and connection in vivo, where the complementary between functional and structural networks offers a more constructive scene with distinctive biomarkers.

Among the state-of-the-art approaches, Graph Neural Networks (GNN) are powerful and promising tools for graph-structured brain network learning and have recently gained tremendous interest in the neuroscience field. Typically, a GNN architecture takes each brain as a graph, where each brain region is represented as a node, and the connectivity is implemented to construct the adjacency matrix. With the advantage of capturing the structural information contained within feature interactions in the graph domain, GNNs have achieved inspiring performances in brain disorder classification [8]–[10].

In general, integrative analysis of multiple types of connectomes by GNN could advance the exploration of high-level latent and complicated representations to improve diagnostic performance. However, this feature has been insufficiently investigated, and it remains non-trivial to model multi-modal neuroimages due to the following challenges. First, the sample size of most existing cohorts is relatively small, while the number of nodes can be very large with noisy estimates of connectivity. Feeding multi-modal features into models without consideration of their intrinsic dependencies increases the dimension of input features and might lead to overfitting and decreased performance. Second, the basic physiological mechanism is still not well understood, and the associations between functional and structural pathways remain uncertain [11]. Although the functional and structural pathways are considered to be mediated by each other [12], [13], the complementary information from multi-modal features is not necessarily linearly related. Finally, most existing graph-based approaches are proposed for mono-modal data and are limited to modeling sophisticated associations among multi-modal nodes due to heterogeneity. In addition, there is a lack of joint compositional reasoning of functional and structural connectomes in these graph-based models.

In light of the above-mentioned issues, in this study, we introduce an efficient Cross-modal Graph Neural Network (Cross-GNN) that utilizes a dynamic graph with mutual learning. In detail, to characterize inter-modal intrinsic dependencies and to address the issue of unknown node relationships in multi-modal graphs, we propose to relate multi-modal associations through dynamic graphs, where the adjacent matrix representing inter-modal dependencies is modeled attentively in a data-driven manner. Specifically, the functional and structural connectomes are encoded and further associated into a compositional space for reasoning. A correspondence factor matrix is obtained for each participant that captures the corresponding values of each pair of nodes between modalities.

Moreover, we propose to model multi-modal complementary features in both explicit and implicit ways. On one hand, cross-modal representations are explicitly translated by the inter-modal correspondence matrix. On the other hand, we leverage the self-knowledge distillation mechanism for online mutual learning, where multi-modal representations are cross-distilled to learn with regularized modality-specific semantic context. In this way, cross-modal latent features are regularized with cross-modal semantic contexts implicitly. A Bilateral Graph Convolution (BGC) layer is further proposed to aggregate multi-modal representations for complementary message passing by taking the correspondence matrix as the adjacency matrix.

Extensive experiments on three datasets were carried out to evaluate the superiority of our method in disease prediction. The learned correspondence matrices were further leveraged to examine multi-modal dependencies as well as to relate key biomarkers. The results indicate that our proposed Cross-GNN provides a promising graph framework for multi-modal learning in terms of prediction and explanation.

The rest of our paper is structured as follows. We would like to review related methods in terms of connectome study, multi-modal neuroimaging, and self-knowledge distillation approaches in Section II. The details of the proposed model are introduced in Section III. Section IV describes the experiment settings. The results are provided and discussed in Section V. Section VI draws the conclusions of the work.

II. Related Works

A. Brain connectome network study

Decoding functional or structural neural communication of the brain enables us to understand the associations between brain network organizations and disorders. Promising progress has been made in the past decade using neuroimaging techniques to identify brain network alterations underlying brain disorders. Recent connectome-wide association studies (CWAS) investigations have successfully identified functional dysconnectivity in neurodegenerative and psychiatric brain diseases. Network-based statistical (NBS) analysis [14] and multivariate distance matrix regression (MDMR) [15] are two widely used statistical tools for the systematic quantification of connectome reorganization across the whole-brain network. Based on the returned key biomarkers, machine learning models such as SVM and random forest are leveraged for classification [4], [16], [17]. However, these approaches are often deficient in modeling high-order nonlinear intrinsic attributes and are limited in their performances.

1) Convolution neural network: Deep learning is a different strategy for associating brain disorders with complex brain network variations. Convolutional neural networks (CNN) and graph neural networks (GNN) facilitate end-to-end disease identification with promising performances and have been widely applied for analyzing connectome patterns. BrainNetCNN [18] was first proposed to tackle the brain graph networks as grid-like data, which achieves state-of-the-art performances in brain network studies. [19] built a CNN framework to learn embedded features from both static and dynamic functional connectivity. [20] built a weighted correlation kernel-based convolutional neural network for learning the hierarchical features. [21] proposed a Deep Convolutional Auto-Encoder (DCAE) network for learning multi-scale features with CNN's hierarchical feature abstraction ability.

2) Graph Neural Network: Apart from CNN, GNNs retain a state that can represent information about the neighbors and structures. Two types of GNNs are commonly used for connectome studies.

A) Population graphs. A series of studies built graphs on populations, where each node represents a subject. This type of approach is carried out in the transductive learning framework with semi-supervised training. For instance, Population-GCN [22] represented a population of participants as one graph, where each edge was manifested by inter-participant phenotypic information and feature similarity. InceptionGCN [23] was proposed to capture intra- and inter-graph structural heterogeneity during convolutions. [24], [25] proposed a similarity-aware adaptive calibrated GCN to consider the disease status. HI-GCN [26] and TE-HI-GCN [27] associated graph topology information with the participants' similarities by building kernels of multiple thresholds of connectivity.

MDCN [28] further leveraged hyper-edges to address spatial specificity in population graphs. One challenge of these models is that these models are carried out by transductive learning, and as a result suffers from high computational costs and poor scalability to unseen data.

B) Brain graphs. Another type of GNN architecture is to tackle each brain region as a graph node and high-order message passing or convolution are performed among regions for aggregating neighbor information. The BrainGNN proposed ROI-aware graph convolutional layers and ROI-selection pooling layers for neurological biomarker prediction at the group and individual levels [8]. [29] proposed to learn a graph similarity metric using a siamese graph convolutional neural network. However, applying a graph network directly to the brain connectome is problematic since there remain unknown relations between brain regions [30]. Recently, dynamic graph strategy has been a powerful tool for tackling the unknown structure issue, such as Dynamic Graph Neural Network (DGNN) [31], and Dynamic Hyperedge Graph Neural Network (DHGNN) [32]. [33] proposed a dynamic spectral graph convolution network by building connectivity from time-varying correlations of fMRI signals. [34] trained a dynamic graph network by learning with sparse brain regional connections from dynamically calculated graph features. Based on this, we model multi-modal node associations by an adaptive adjacency matrix, which is obtained by normalizing the correspondence matrix from the compositional space.

B. multi-modal connectome learning

A straightforward way for multi-modal learning is to concatenate features and feed them into a classifier such as SVM or MLP for prediction [35], [36]. To mitigate over-fitting, feature selection methods were applied to reduce the feature dimension [37], [38].

Compared with machine learning methods, deep learning methods are feasible to capture high-order representations and achieve better performances. [39] proposed a multimodal fusion network for disease classification based on functional and structural magnetic resonance images. [40] implemented multimodal features into a graph with variational edges. [41] utilized graph hashing learning and converted deep features into hash codes to maintain the original semantic spatial relationships. [42] proposed to perform a two-layer convolution on the fMRI and DTI data simultaneously. [43] regularized convolution on functional connectivity with structural graph Laplacian. [44] proposed a DTI penalty term to fuse functional information and structural information for brain connectivity networks with multi-center and multi-channel mechanisms. A triplet network with a self-attention mechanism was introduced to map high-order multi-modal representations [45]. Besides, hypergraph graph neural networks [31], [32] are proposed to encode hyperstructure. The multi-modal brain networks can be fed as a nature of hyperedges. However, most of these studies potentially ignore the heterogeneity between modalities as a result of a lack of ability to model cross-modality complementary associations and achieve sub-optimal results. In addition, there is a lack of joint compositional reasoning over both functional and structural connectome networks, which plays a key role in phenotype-connectome association studies.

C. Knowledge distillation

Knowledge distillation (KD) was first proposed for model compression by reinforcing the student model learning from the teacher model without a significant drop in performance [46]. Despite its usage on model compression, KD penalizes the predictive distributions with additional knowledge to improve better generalization ability. By learning transferred knowledge from a teacher model, the student can approximate or exceed the teacher model. Based on this mechanism, self-knowledge distillation is proposed to facilitate the model training by learning from itself to mitigate overconfidence and improve performance. [47] implements auxiliary classifiers to utilize the output of intermediate layers, where the knowledge in deeper networks is squeezed into shallow ones. [48] proposed a gating mechanism to filter meaningful knowledge and achieved consistent improvements for brain age prediction. [49] implemented self-distillation for capturing high-level and low-level features to calibrate the neoplasia identification results. [50] devised a distillation module to transfer semantic region information from teacher to student for medical image segmentation. In this study, we utilize the promising self-distillation mechanism for cross-modal mutual learning, where mono-modal representations are regularized with cross-modal contexts.

III. METHOD

Figure 1 illustrates the overall architecture of the proposed Cross-GNN, which consists of functional (orange), structural (blue), and cross-modal (green) pathways. The functional and structural representations are parsed into corresponding pathways and fed to construct the correspondence matrix. Inter-modal representations are translated by cross-embedding in the cross-modal pathway. Moreover, a dynamic multi-modal graph is built by taking multi-modal representations as node features and the normalized correspondence matrix as the adjacency matrix. In addition, a cross-distillation is implemented in the cross-modal pathway to regularize mono-modal representations with cross-modal information. In this section, we introduce the brain network and graph formulation and then the details of the Cross-GNN approach, including dynamic multi-modal graph mapping, bilateral graph convolution, cross-distillation, and optimization.

A. Preliminaries

Brain network: Functional connectivity (FC) and structural connectivity (SC) are derived by mapping processed neuroimages into a template. FC and SC are symmetric positive definite matrices $X^{fc}, X^{sc} \in \mathbb{R}^{M\times M}$, where M denotes the number of brain regions. Each element $x_{i,j}$ denotes a covariance or connectivity strength between the regions i and j. The brain network is usually formulated as an undirected graph $G = (V, E)$, where V is a finite set of vertices with $|V| = M$ and $E \in \mathbb{R}^{M\times M}$ denotes the edges in the graphs. For each

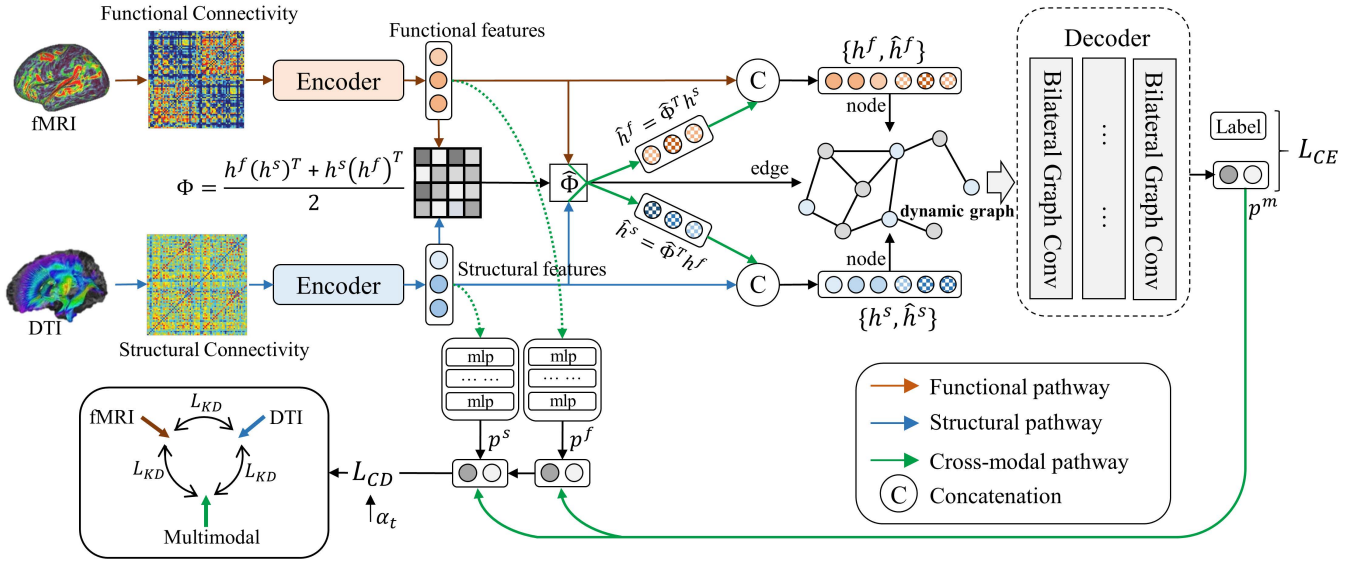


Fig. 1. Illustration of the proposed Cross-GNN method including functional, structural, and cross-modal pathways. The multi-modal brain networks are firstly encoded and then formulated into a correspondence matrix Φ for reasoning. The multi-modal representations are cross-embedded and cross-distilled for cross-modal representation learning.

vertex i, the node feature vector v_i is constructed by the i-th row or column in the matrix $v_i = \{x_{i,k}|k = 1, 2, ..., M\}$. The edges are represented by the matrices directly $e_{i,j} = x_{i,j}$.

Multi-modal brain graph: By taking functional and structural brain networks, the multi-modal graphs are constructed as $\hat{G} = \{G^s, G^f\}$, where G^s and G^f denote the structural and functional brain network graphs, respectively. Formally, given a set of multi-modal graphs $\{\hat{G}_1, \hat{G}_2, ..., \hat{G}_N\}$ with a few labeled graph instances, the aim of the study is to decide the state of the unlabeled graphs as a graph classification task.

Dynamic graph: In this study, our proposed Cross-GNN leverages the dynamic mechanism to model multi-modal node representations and interactions by learning mappings: $f_V : X \to V, f_E : X \to E$. And the learned graphs are denoted as dynamic graphs. In this study, the adjacent matrix in the multi-modal graph is built by the correspondence matrix between functional and structural connectome.

B. Dynamic multi-modal graph mapping

In this section, we leverage the dynamic graph mechanism to build a dynamic multi-modal graph by parsing fMRI and DTI into node representations and building a multi-modal correspondence matrix as the adjacency matrix.

1) Brain network Encoder: Since the multi-modal graphs are heterogeneous, simply embedding features together might decrease the performance. One of the keys is how to parse the brain network of each modality precisely. Formally, we define an encoder operation ψ to learn representations by:

$$h = \psi(G), h \in \mathbb{R}^{M \times H} \tag{1}$$

where M denotes the ROI number and H denotes the hidden size of the encoded features. The choice of brain network encoder functions ψ has a crucial influence on the multi-modal graph representations. For example, a simple way of embedding a connectivity matrix can be a multi-layer perception (MLP):

$$h = \psi_{mlp}(G) \tag{2}$$

The second choice of ψ is by graph embedding. The absolute values of connectivity are normalized by symmetric normalized Laplacian:

$$h = \psi_{graph}(G) = \sigma(\hat{L}HW) \tag{3}$$

$$\hat{L} = I_n - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \tag{4}$$

where D denotes the degree matrix of $A, A = abs(X)$. The graph embedding method facilitates capturing graph topological properties.

The third choice of ψ is to tackle the connectivity matrix as a sequence. Well-known sequence models such as LSTM [51] and GRU [52] are powerful tools with the gating to capture long-range sequential relationships and protect the learning from undesired updates [51]. For example, the GRU layers take the input brain network matrix $X \in \mathbb{R}^{M \times M}$ as a sequence with M nodes and M features. Formally, given an input X, the node features are embedded by:

$$h = \psi_{GRU}(G) = GRU(X) \tag{5}$$

Especially, the GRU operation on each parcel $k \in [1, M]$ is formulated as:

$$z_k = \sigma(W_z \cdot [o_{k-1}, x_k]) \tag{6}$$

$$r_k = \sigma(W_r \cdot [o_{k-1}, x_k]) \tag{7}$$

$$\hat{h}_k = tanh(W \cdot [r_k * o_{k-1}, x_k]) \tag{8}$$

$$o_k = (1 - z_k) * o_{k-1} + z_k * \hat{o}_k \tag{9}$$

where x_k denotes the input vector corresponding to the k-th brain region, and z_k, r_k are the update gate and reset gate respectively. o_k is the output of the GRU.

2) Correspondence matrix: With the embedded h^f and h^s, a soft correspondence matrix Φ is obtained by:

$$\Phi = \frac{h_j^f (h_j^s)^T + h_j^s (h_j^f)^T}{2}, \Phi \in \mathbb{R}^{M \times M} \tag{10}$$

Each row vector in Φ is a probability distribution over potential correspondences to corresponding nodes. The matrix can be regarded as the scores for measuring the goodness of matches between nodes in two modalities. A sinkhorn function is applied to normalize the matrix, which satisfies doubly stochastic, where $\sum_j^M \hat{\Phi}_{i,j} = 1$.

3) Dynamic graph: By obtaining the normalized correspondence matrix $\hat{\Phi}$, we formulate the normalized correspondence matrix $\hat{\Phi}$ as the dynamic adjacency matrix. Moreover, we can project the representations from one modality field into another (i.e. translate functional/structural representations to structural/functional representations) by:

$$\hat{h}^f = \hat{\Phi}^T h^s, \hat{h}^f \in \mathbb{R}^{M \times H} \tag{11}$$

$$\hat{h}^s = \hat{\Phi}^T h^f, \hat{h}^s \in \mathbb{R}^{M \times H} \tag{12}$$

With the obtained embedded representations $\hat{h}^f$ and $\hat{h}^s$, the dynamic node features can be built by $\{\hat{h}^f, h^f\}$ or $\{\hat{h}^s, h^s\}$, which represents the translated representations of functional/structural field.

C. Bilateral Graph Convolution

In this study, a Bilateral Graph Convolution (BGC) is proposed to perform convolution on the multi-modal graphs. To tackle the heterogeneous features between modalities, the BGC module applies convolutions on each field to aggregate representations of every single modality separately. Moreover, we implement spatial aggregation on the graphs for message passing instead of spectral graph convolution. Since the brain network is fully connected, graph spatial convolution, as well as spectral graph convolutions, can aggregate global information. In this way, the graph spatial convolution is formulated with $\hat{\Phi}$ as:

$$h_{l+1}^f = \sigma(\hat{\Phi} h_l^f W_l^f), \text{where } h_0^f = ||\{\hat{h}^f, h^f\} \tag{13}$$

$$h_{l+1}^s = \sigma(\hat{\Phi} h_l^s W_l^s), \text{where } h_0^s = ||\{\hat{h}^s, h^s\} \tag{14}$$

$||$ denotes a concatenation operation, σ denotes a sigmoid activation function, and W is a learnable matrix for improving node representations. The outputs of the BGC layer are further combined as a feature vector and fed into a multi-layer perception for classification.

D. Cross-distillation

In the proposed Cross-GNN, there exist heterogeneous representations, i.e., functional, structural, and multi-modal compositional features. These features have different yet complementary characteristics in decision-making. In this regard, we proposed to combine the advantages of this heterogeneous information and conduct the training by mutual learning to improve the representations further. In addition, to avoid heterogeneity, we develop a joint optimization strategy based on a self-distillation framework and design an auxiliary branch for cross-distillation.

As is shown in Figure 1, two auxiliary branches are implemented for functional and structural outputs. The auxiliary branches are built by three-layer multi-layer perception with leaky ReLU activation function and dropout. Accordingly, the three outputs, i.e. functional prediction p^f, structural prediction p^s, and multi-modal prediction p^m are leveraged to construct the cross-distillation loss function L_{CD}:

$$L_{CD} = KL(p^m||p^f) + KL(p^m||p^s) + KL(p^s||p^f) \tag{15}$$

where L_{CD} is obtained by a combination of three KL divergences by measuring output distribution similarity.

However, the reliability of the single-modal representations should be considered since the model generally does not have enough knowledge for mutual learning at the early stage of training. Inspired by the progressive self-knowledge distillation [53], we propose to gradually balance the mutual learning loss with the weight α. We apply a linear growth approach:

$$\alpha_t = \alpha_T \times \frac{t}{T} \tag{16}$$

where T is the total epoch for gradual training of α. To summarize, the self-distillation loss function at t-th epoch can be obtained as:

$$L_{CD,t} = \alpha_t \cdot KL(p^m||p^f) + \alpha_t \cdot KL(p^m||p^s) + \alpha_t \cdot KL(p^s||p^f) \tag{17}$$

E. Optimization

In the training process, the objective function is constructed by a combination of a weighted cross-distillation $L_{CD,t}$ and cross-entropy function L_{CE}:

$$L_t = L_{CE}(p^m, y) + L_{CD,t} \tag{18}$$

where y denotes the ground truth. In the inference, the multi-modal outputs p^m are implemented for prediction.

To summarize, the detail of our proposed Cross-GNN is shown in Algorithm 1:

F. Biomarker Interpretation

The relationship between brain structure and function has become an increasingly important study in the clinical neuroscience field [54]–[56]. The correspondence matrices in Cross-GNN attentively model the multi-modal interactions among regions, which are feasible to be applied to pinpoint the key brain biomarkers underlying disease. In this section, we extracted the correspondence matrix for each subject and performed a group statistical analysis for biomarker detection. To associate brain disorders with multi-modal correspondence, we performed the CWAS approach, multi-distance multi-variance regression (MDMR) [15] for group-wise comparison.

In detail, a distance matrix in the subject space was calculated for each region. Within each distance matrix, the distance between multi-modal correspondence patterns Φ for every possible subject pair among all groups related to region i was calculated by

$$d_{uv}^i = dis(\bar{\Phi}_u, \bar{\Phi}_v) \tag{19}$$

Algorithm 1: Cross-modal Graph Neural Network

Input: multi-modal brain networks $\{\hat{G}_1, \hat{G}_2, \hat{G}_3, ..., \hat{G}_n\}$,where $\hat{G} = \{G^s, G^f\}$; For each graph $G = (V, E, X)$

Output: Prediction p^m of the test set

Calculate dynamic node features: $h^s = Encoder(x^s)$, $h^f = Encoder(x^f)$;
Calculate the correspondence matrix: $\Phi = \frac{1}{2}h^f(h^s)^T + \frac{1}{2}h^s(h^f)^T$;
Normalize Φ into $\hat{\Phi}$ to satisfy doubly stochastic;
Obtain the cross-modality mapping representations:
$\hat{h}^f = \hat{\Phi}^T h^s$;
$\hat{h}^s = \hat{\Phi}^T h^f$;
for *i=1:l* **do**
  $h_i^f = \sigma(\hat{\Phi} h_{i-1}^f w_{i-1}^f)$, where $h_0^f = ||\{\hat{h}^f, h^f\}$;
  $h_i^s = \sigma(\hat{\Phi} h_{i-1}^s w_{i-1}^s)$, where $h_0^s = ||\{\hat{h}^s, h^s\}$;
end
Readout $h^m = ||\{h_l^f, h_l^s\}$;
$p^m \leftarrow Readout(h^m)$;
$p^f \leftarrow Readout(h^s)$;
$p^s \leftarrow Readout(h^f)$;
Optimize the objective function Eq. (18)

where dis denotes the distance of multi-modal correspondence vectors, i.e. $\bar{\Phi}_u$, $\bar{\Phi}_v$, of subject u and v. In the implementation, the distance function is obtained by $dis(\bar{\Phi}_u, \bar{\Phi}_v) = ||\bar{\Phi}_u - \bar{\Phi}_v||_2$ by reference to previous works [15], [28]. A pseudo F-statistic analogous to an F-statistic from a standard ANOVA was performed. The total sum of squares for region i was obtained as

$$SS_T^i = \frac{1}{n}\sum_{u=1}^{n}\sum_{v=u+1}^{n} d_{uv}^{i2} \tag{20}$$

where $n = n_1 + n_2$, the total number of subjects. Meanwhile, the within-group sum of squares was formulated By

$$SS_W^i = \frac{1}{n_1}\sum_{u=1}^{n}\sum_{v=u+1}^{n} d_{uv}^{i2}\varepsilon_{uv}^a + \frac{1}{n_2}\sum_{u=1}^{n}\sum_{v=u+1}^{n} d_{uv}^{i2}\varepsilon_{uv}^b \tag{21}$$

where n_1 and n_2 denote the number of each group respectively. ε_{uv}^a represents the belonging of the subject u and v, which equals to one when u and v within the same group. And then the F-statistic score of the region i would be obtained by

$$F^i = (n-1)\frac{SS_T^i - SS_W^i}{SS_W^i} \tag{22}$$

In addition, a random permutation with 2000 times to subjectes was applied to simulate the null distribution, and the pseudo F-statistic score was recomputed for each time. The p-value was finally calculated by counting the pseudo F-statistics from permutated values greater than those derived from the original data. This step was repeated for all ROIs. The Bonferroni correction was applied to control the false positive rate. And the p-value < 0.05 after the correction was determined significant within the experiments.

Finally, to explicitly measure which specific connectivity pattern is primarily driving the association between inter-modal correspondence and disease progression, the δ-statistic was applied here to identify the top five connected network nodes with the greatest effect size from each seed brain region [57].

IV. Experiments

A. Datasets

In this study, two real-world datasets are employed in this study, where functional MRI and DTI are aggregated. All the datasets are enrolled for multi-modal graph classification. The detailed demographic information is listed in Table I with gender (Male/Female), age (Mean±Std), and other scale information (Mean±Std). All three groups were matched for age and sex.

ADNI Dataset[1]: The ADNI dataset is a longitudinal and multi-site multi-modal neuroimaging dataset. In this study, we collected 442 subjects for evaluation, including 142 healthy controls (HC) and 151 with mild cognitive impairment (MCI), and 149 patients with Alzheimer's disease (AD) from ADNI 1, ADNI2, ADNI GO, and ADNI 3. Each involved participant includes fMRI and DTI data. We excluded duplicated scans, and a single scan of fMRI and DTI per subject was collected. Notably, MCI is considered to be a significant stage for the preclinical diagnosis of AD. These subjects are divided into three groups: AD, MCI, and HC, in accordance with the standard clinical criteria, such as Mini-Mental State Examination (MMSE) scores and Clinical Dementia Rating (CDR).

Xuanwu dataset: A total of 138 subjects are included in this dataset, where 53 HCs and 85 subjects with Parkinson's Disease (PD) were recruited from the Movement Disorders Clinic of the Xuanwu Hospital of Capital Medical University. In the dataset, the HCs were all older than 40 years, with no family history of movement disorders and no obvious cerebral lesions found in MR images. The PDs were diagnosed according to the MDS Clinical Diagnostic Criteria for Parkinson's disease. The Movement Disorder Society Unified Parkinson's Disease Rating Scale, part III (UPDRS III) was applied to evaluate the PDs.

PPMI dataset[2]: The PPMI dataset includes 53 HCs and 56 PDs. The patients were diagnosed at baseline, and the HCs were healthy at their first examination. Notably, the PPMI dataset is also a longitudinal database, and each participant has multiple scans. We excluded duplicated scans, and a single scan of fMRI and DTI per subject was collected. The diagnostic criteria of PD followed the inclusion criteria for patients of the PPMI study (`https://www.ppmi-info.org/study-design/research-documents-and-sops`).

B. Preprocessing

All the fMRI images were pre-processed by reference to the Configurable Pipeline for the Analysis of Connectomes (C-PAC) pipeline [58], including skull striping, slice timing

[1] `http://www.adni-info.org/`
[2] `https://www.ppmi-info.org/`

TABLE I
CHARACTERISTICS OF PARTICIPANTS OF THREE DATASETS.

Dataset	Type	Sex	Age	CDR	MMSE
ADNI	NC	81/61	75.85±9.7	0.0±0.1	28.9±2.3
	MCI	87/64	77.0±11.7	0.5±0.2	26.6±2.8
	AD	87/62	76.9±8.2	1.0±0.4	21.8±3.6
	F-statistic	0.054	0.825	443.669	219.407
	p-value	0.816	0.364	< 0.001	< 0.001
		Sex	Age	MOCA	UPDRS III
Xuanwu	NC	23/30	60.8±9.4	25.5±3.2	-
	PD	43/42	59.8±8.7	24.0±3.5	25.5±12.8
	F-statistic	0.67	0.162	4.256	-
	p-value	0.414	0.688	0.041	-
PPMI	NC	34/19	64.1±11.3	27.8±1.8	-
	PD	34/22	62.3±9.8	26.3±3.3	23.7±9.7
	F-statistic	0.707	1.002	1.72	-
	p-value	0.401	0.319	0.193	-

correction, motion correction, global mean intensity normalization, nuisance signal regression with 24 motion parameters, and band-pass filtering (0.01-0.08Hz). The functional images were finally registered into standard anatomical space (MNI152). The mean time series for a set of regions were computed and normalized into zero mean and unit variance. Pearson Coefficient Correlation was applied to measure functional connectivity.

The DTI images were pre-processed by image denoising, head motion, eddy-current, susceptibility distortion, and field inhomogeneity correction by MRtrix 3 [59]. We performed the 2-nd order Integration over Fiber Orientation Distributions [60] to reconstruct 10 million streamlines. A Spherical-deconvolution Informed Filtering of Tractograms [61] was applied to reduce the streamline count to 5 million. The number of streamlines connecting each pair of brain regions was used to construct the structural network.

All the pre-processed fMRI and DTI images were mapped by the brain template for parcellations. In this study, the images in ADNI and Xuanwu datasets were segmented by the Schaefer atlas [62], which was parceled by a gradient-weighted Markov random field approach that identified 100 cortical parcels.

C. Implementation details

In our implementation, the number of layers of bilateral graph convolution is decided in a grid search from 1 to 4. The outputs of bilateral graph convolution layers are further fed into a 3-layer multi-layer perception classifier followed by a leaky ReLU activation function and a dropout layer. The learning rate is set as 3e-4, and the weight decay is 5e-5. All the models in this study are trained for 600 epochs and would be stopped early when the loss has not been decreased for 100 epochs. We trained the models with PyTorch on one NVIDIA 2080-Ti GPU. 10-fold cross-validation was applied for evaluation, where 10% samples were randomly selected for testing for each fold. For all experiments, we evaluated the performance in terms of diagnosis accuracy (Acc), sensitivity (Sen), and specificity (Spe). Our code is available at `https://github.com/podismine/Cross-GNN`.

D. Competitive methods

In this study, we compare our proposed Cross-GNN with baseline machine learning approaches and well-estimated graph methods. These methods include:

SVM and MLP. The conventional machine learning methods of the support vector machine (SVM) and multi-layer perception (MLP) are compared as a baseline of classification. The upper matrix of the brain networks is fed into the classifiers to give a score for each subject. The layer number of MLP is searched from 1 to 4.

Population-GCN [22], **BrainNetCNN** [18] and **BrainGNN** [8]. In population-GCN, a population graph is built, where each node is represented by concatenating the vectorized upper matrix of the brain networks of a subject. Key features were selected by recursive feature elimination and vectorized into a set for each vertex and then concatenated. The number of selected features is searched. The adjacency matrix was constructed by the phenotype values as well as the similarity between node features. Moreover, BrainNetCNN was implemented by multiple convolution layers to perform multi-modal data, where functional and structural connectivity matrices are concatenated by channel. In BrainGNN, the original partial Pearson and full correlation coefficients are replaced by functional and structural connectivity. The three models are implemented as baseline deep learning models for comparison.

M-GCN [43], **HGNN** [32], **DHGNN** [31]. These three models are among state-of-the-art multi-modal network approaches. M-GCN aggregates functional representations by regularizing with the structural graph Laplacian. Hypergraph graph neural network (HGNN) encodes multi-modal connectivity by hyperstructure into hyperedges. The dynamic graph hyper-neutral network (DHGNN) extends the HGNN into a dynamic graph. These three models are implemented by reference to the originally proposed architecture.

MMP-GCN [44], Triplet Attention Network (**TAN**) [45] are state-of-the-art networks that are designed specially for brain disorder diagnosis. In MMP-GCN, we discarded the multi-center information and applied the gender information to construct the graph. The layers and heads of self-attention of the Triplet network are decided with a grid search.

V. RESULTS AND DISCUSSION

A. Hyperparameter setting

In our study, we first evaluated the impact of model hyperparameter settings on performance. One key parameter is the hidden size H. As is shown in Fig. 2 A), our model achieved the best accuracy performances with the hidden size $H = 64$ in most cases. The accuracy performances increase when the hidden size H ranges from 16 to 128 and decrease when the hidden size H is larger than 128. Another key parameter is the number of BGC layers l. The layer number decides the multi-hop neighborhood aggregation and message passing among multimodal nodes. Fig. 2 B) demonstrates that $l = 2$ is the optimal setting except for the classification of NC and MCI.

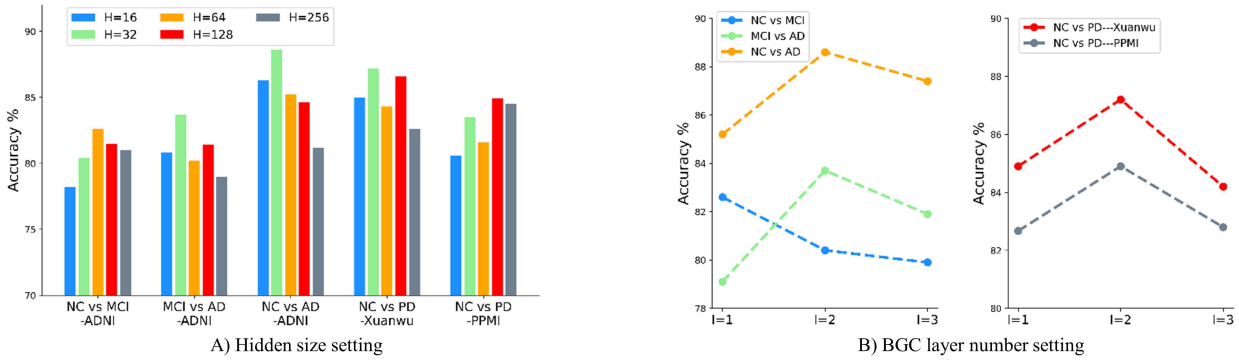


Fig. 2. The effect of the setting of A) the hidden size $\boldsymbol{H}$ on the left and B) the number of BGC layers $\boldsymbol{l}$ on the right. The performances are evaluated on three datasets including the ADNI, the Xuanwu, and the PPMI datasets.

TABLE II
EVALUATIONS OF ENCODERS ON THE ADNI DATASET. THE BEST RESULTS ARE SHOWN IN BOLD.

Type	NC vs MCI			MCI vs AD			NC vs AD		
	Acc	Sen	Spe	Acc	Sen	Spe	Acc	Sen	Spe
ϕ_{MLP}	78.9±5.0	82.6±10.9	78.3±8.0	79.1±7.9	89.7±10.1	75.3±8.3	82.4±5.7	88.8±8.4	79.7±7.3
ϕ_{Graph}	79.5±4.5	83.5±9.5	79.4±8.0	77.4±4.6	85.7±10.4	75.1±8.0	86.2±5.9	92.4±7.4	83.7±7.6
ϕ_{LSTM}	81.3±4.8	**85.0±8.9**	80.0±9.7	82.0±4.1	90.3±7.1	80.4±4.5	**88.6±6.1**	94.5±7.1	85.1±10.3
ϕ_{GRU}	**82.6±4.6**	84.6±5.0	**82.2±8.4**	**83.7±6.7**	**92.6±12.1**	**80.9±8.9**	**88.6±6.3**	**95.9±9.2**	**85.2±7.7**

TABLE III
EVALUATIONS OF ENCODERS ON THE CLASSIFICATION OF NC VS PD ON XUANWU AND PPMI DATASETS. THE BEST RESULTS ARE SHOWN IN BOLD.

Type	NC vs PD (Xuanwu)			NC vs PD (PPMI)		
	Acc	Sen	Spe	Acc	Sen	Spe
ϕ_{MLP}	84.5±5.0	81.6±6.4	**92.4±7.7**	76.7±7.5	84.1±8.4	73.0±11.3
ϕ_{Graph}	83.9±3.2	87.4±9.4	85.5±10.5	81.8±5.2	86.3±6.9	81.5±5.6
ϕ_{LSTM}	83.2±3.2	82.9±7.7	87.4±8.8	84.7±5.7	**88.1±5.2**	81.8±7.5
ϕ_{GRU}	**87.2±8.1**	**88.9±9.9**	89.0±12.2	**84.9±6.1**	86.6±6.5	**82.5±8.5**

As the number of layers is greater than 2, the performances tend to decrease on all five classification tasks, due to the over-smoothing problem [63], [64], even with dynamic graphs.

B. Choice of encoder

We further evaluated the classification performance in different ways of brain network encoders. The results on three datasets are listed in Table II and Table III in terms of accuracy, sensitivity, and specificity, where the best results are shown in bold. From the results, we can see that the approaches of MLP and graph embedding achieve comparable performance in the Xuanwu dataset. While graph embedding outperforms MLP on the ADNI and PPMI datasets. The inconsistent results might be caused by the unclear graph structure, leading to poor generalizability for representation learning. Moreover, the LSTM and GRU methods provide long-range dependencies for modeling graphs as a sequence and outperform graph embedding and MLP. Generally, GRU has better generalization ability, which outperforms LSTM in most cases. The other results of Cross-GNN in this paper will be discussed by using the GRU encoder.

C. Comparison with competitive methods

Table IV and Table V demonstrate the comparison results on three datasets, where the mean and standard deviation across folds are listed. The best results are shown in bold, and the second-best results are underlined. From the results, we can see that SVM performs the worst among all the models. This indicates that the multi-modal data might have nonlinear and heterogeneous structures, and a simple linear classifier fails to distinguish the brain disease state. Moreover, compared with SVM, multi-layer perception improves performances significantly, where nonlinear representations are deeply embedded with multiple layers.

Moreover, BrainNetCNN, Population-GCN, and BrainGNN achieve further improvements, showing that deep learning methods such as CNN and GNN are feasible to capture more meaningful representations. In addition, the multi-modal graph networks, e.g., M-GCN, outperforms BrainNetCNN, Population-GCN, and BrainGNN in most cases, where the complementary information between modalities plays a key role in improving performance. Especially, DHGNN outperforms other baseline approaches in most cases. For example, in distinguishing MCI from NC, DHGNN achieves an accuracy of 79.2%, an sensitivity of 83.0%, and a specificity of 77.6%.

TABLE IV
CLASSIFICATION RESULTS OF DIFFERENT APPROACHES ON THE ADNI DATASET. THE AVERAGE AND STANDARD DEVIATION RESULTS ACROSS FOLDS ARE DISPLAYED.

	NC vs MCI			MCI vs AD			NC vs AD		
	Acc	Sen	Spe	Acc	Sen	Spe	Acc	Sen	Spe
SVM	52.2±9.3	52.4±9.0	52.2±10.6	60.0±9.9	59.8±14.9	60.8±7.8	75.4±9.8	77.5±14.0	75.3±8.8
MLP	70.08±8.8	75.0±15.8	73.1±13.6	69.1±5.4	73.6±9.3	70.1±9.9	77.8±6.5	76.6±6.8	80.4±8.3
BrainNetCNN [18]	72.08±7.1	75.9±9.3	70.4±10.8	71.8±5.0	74.4±6.8	70.8±5.0	81.3±11.3	81.0±11.7	82.1±10.1
Population-GCN [22]	75.5±5.9	82.5±10.0	71.6±8.0	75.5±5.9	80.2±14.4	74.0±5.4	82.0±7.5	85.9±7.5	75.9±8.0
BrainGNN [8]	76.0±5.9	81.1±9.9	71.5±12.6	74.3±4.8	76.6±5.2	73.5±12.8	82.0±2.4	84.6±5.6	76.8±6.1
M-GCN [43]	75.7±6.3	80.1±4.8	72.9±6.1	75.8±6.9	76.1±8.1	75.0±9.0	84.4±5.7	84.1±6.0	85.6±7.1
HGNN [32]	78.1±6.6	80.2±8.4	74.7±9.4	76.8±5.5	83.5±12.8	76.0±7.5	81.4±8.0	82.5±6.1	82.0±12.1
DHGNN [31]	79.2±6.0	83.0±9.9	77.6±7.3	75.8±4.2	85.4±12.9	73.9±8.4	82.5±4.1	86.4±10.7	82.1±11.4
MMP-GCN [44]	80.6±7.6	**86.2±10.2**	79.5±10.1	79.6±11.0	82.52±14.0	**82.2±14.3**	84.4±11.1	88.1±12.4	85.1±13.2
TAN [45]	80.2±5.4	82.0±11.1	79.2±11.6	81.2±3.5	83.7±8.8	79.0±9.7	85.9±4.5	89.3±12.9	**86.3±10.3**
Cross-GNN (Ours)	**82.6±4.6**	84.6±5.0	**82.2±8.4**	**83.7±6.7**	**92.6±12.1**	80.9±8.9	**88.6±6.3**	**95.9±9.2**	85.2±7.7

TABLE V
CLASSIFICATION RESULTS OF DIFFERENT APPROACHES ON THE CLASSIFICATION OF NC AND PD ON THE XUANWU AND PPMI DATASETS. THE AVERAGE AND STANDARD DEVIATION RESULTS ACROSS FOLDS ARE DISPLAYED.

	NC vs PD (Xuanwu)			NC vs PD (PPMI)		
	Acc	Sen	Spe	Acc	Sen	Spe
SVM	64.9±12.7	63.1±9.9	69.7±21.3	63.1±8.0	65.1±8.8	61.4±9.7
MLP	76.8±8.9	76.4±10.7	81.1±13.6	73.7±7.0	76.7±7.6	73.2±8.1
BrainNetCNN [18]	75.6±9.7	77.6±16.5	87.9±9.3	78.7±8.1	83.3±11.7	78.3±11.1
Population-GCN [22]	76.1±7.1	83.3±12.8	75.1±12.3	80.5±12.0	80.2±13.6	**84.8±8.3**
BrainGNN [8]	76.7±8.6	80.0±12.3	74.3±12.5	79.7±6.6	82.0±11.6	76.7±9.1
M-GCN [43]	81.3±4.6	80.0±7.6	86.1±9.9	79.4±7.6	80.4±11.9	84.5±11.5
HGNN [32]	77.7±9.8	75.3±7.5	**89.3±9.8**	83.0±4.1	85.1±9.3	82.7±9.9
DHGNN [31]	78.1±6.8	80.5±10.2	83.6±12.3	83.8±5.7	80.9±8.1	84.7±8.2
MMP-GCN [44]	81.3±9.8	73.7±13.3	83.6±10.5	81.4±9.1	81.7±11.8	80.7±11.1
TAN [45]	82.8±5.8	83.3±13.4	82.9±10.7	82.1±9.2	81.7±16.3	83.6±13.0
Cross-GNN (Ours)	**87.2±8.1**	**88.9±9.9**	89.0±12.2	**84.9±6.1**	**86.6±6.5**	82.5±8.5

Finally, compared with these approaches, our proposed Cross-GNN further improves the accuracy on the ADNI dataset (NC vs. MCI: 82.6%; MCI vs. AD: 83.7%; NC vs. AD: 88.6%), the Xuanwu dataset (NC vs. PD: 87.2%), and the PPMI dataset (NC vs. PD: 84.9%). In particular, compared with the state-of-the-art Transformer model, TAN, 4.4% improvements are achieved in the classification of NC vs. PD on the Xuanwu dataset. We suspect that the improvements might be caused by the powerful cross-modality correspondence reasoning that captures inter-modal dependencies to aggregate multi-modal representations.

D. Ablation Study

As is shown in Eq.18, the objective function is composed of two parts, including the cross-entropy loss and cross-distillation loss. To investigate the effectiveness of cross-distillation loss, we compared it with other combinations in Table VI, where the $KL(p^m||p^{f/s})$ term denotes the cross-distillation between multi-modal and mono-modal knowledge. The function $KL(p^s||p^f)$ represents that between functional and structural knowledge.

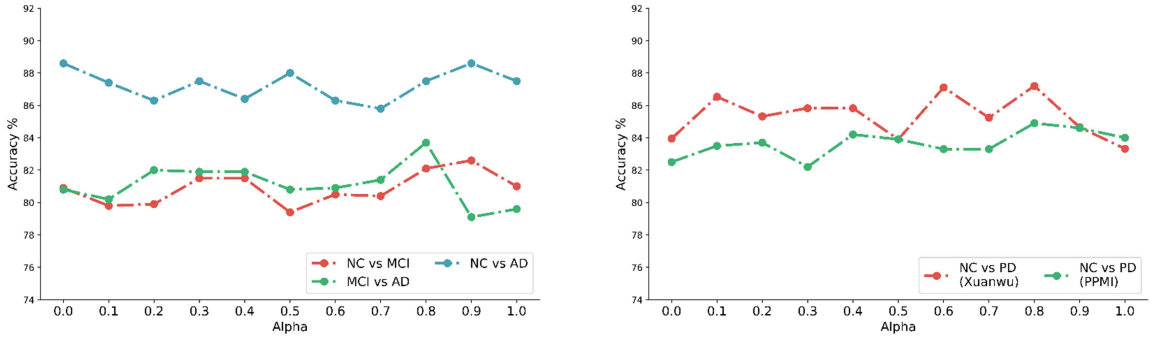
From the results, we can see that the two types could enhance performance in most cases. However, neither of these two components contributes to consistent improvements in the three datasets. For example, there is a drop in the performance in distinguishing AD from NC (-1.7%) with cross-mono-modal learning. However, the best performances are achieved on all tasks when we combine the two components. We suspect that with cross-mono-modal distillation, the mono-modal representations are regularized with a hint of multi-modal and cross-modal information. And in this way, mono-modal learning is equipped with more meaningful knowledge, resulting in improved diversity of representations.

Moreover, we evaluated the relaxation coefficient α_t by adjusting α_T in Eq. (16). The results on three datasets are plotted in Fig. 3. In Fig. 3 A), the red, green, and blue dotted lines indicate the classification results of NC vs. MCI, MCI vs. AD, and NC vs. AD, respectively, and the results of NC vs. PD on the Xuanwu and PPMI datasets are shown in Fig. 3 B). With the value of α_T increasing, the performances are enhanced and then dropped slightly. It is reasonable that the self-knowledge distillation might lead to over-regularization, which could be mitigated by the knowledge distillation coefficient α_T to some extent. In addition, on the ADNI dataset, the best performances are achieved with the α_T value of 0.9 in NC vs. MCI, 0.8 in MCI vs. AD, and 0.8 in NC vs. AD. While on the Xuanwu and PPMI datasets, the value of 0.8 is optimal for the classification of NC and PD.

TABLE VI

ABLATION STUDIES ON THE KNOWLEDGE DISTILLATION IN THE METHOD. THE COMPONENTS EVOLVED ARE LISTED IN THE TABLE WITH ✓.

		ADNI			Xuanwu	PPMI
$KL(p^m\|\|p^{f/s})$	$KL(p^s\|\|p^f)$	NC vs MCI	MCI vs AD	NC vs AD	NC vs PD	NC vs PD
		80.9±5.0	80.8± 4.2	**88.6±6.6**	82.6±6.5	82.5±8.7
	✓	81.4±5.2	80.3±4.4	86.9±8.1	83.9±7.9	83.7±7.3
✓		**82.6±3.6**	81.4±4.1	**88.6±6.3**	85.9±8.1	83.7±4.2
✓	✓	**82.6±4.6**	**83.7±6.7**	**88.6±6.3**	**87.2±8.1**	**84.9±6.1**



A) ADNI B) Xuanwu and PPMI

Fig. 3. The effect of the setting of the weight α_T of cross-distillation loss L_{CD}on three datasets. A) The results on the ADNI dataset are shown in dotted lines, where the accuracy of NC VS MCI is shown in red, MCI VS AD in green, and NC VS AD in blue. B) The results of NC VS PD on the Xuanwu and PPMI datasets are shown in red and green lines.

E. Evaluation on inter-center generalization

To evaluate the generalization ability and reproducibility, we further apply the Xuanwu and the PPMI datasets for inter-center testing, where the models are trained on the Xuanwu datasets and validated on the PPMI dataset. The experiments are repeated 10 times and the averaged results as well as deviations are displayed. Table VII shows the performances of our proposed Cross-GNN model, compared with other baseline models and state-of-the-art networks. The best results are shown in bold, and the second best are underlined. Our proposed method achieves the best performance among all the models, with an accuracy of 78.1%, a sensitivity of 80.5%, and a specificity of 79.6%. Compared with the results that are obtained by trained and validated on the same PPMI dataset, there is a 6%± drop in performances. One main reason is that the criteria for including participants in training Xuanwu dataset are different from the test PPMI database. For example, the average MoCA score of PD in the Xuanwu dataset is 24.0, while the average in the PPMI dataset is 26.3. Besides, the various scan parameters and MRI protocol settings also affect the performance. Overall, our Cross-GNN could achieve the best inter-center generalization ability among all the state-of-the-art models.

TABLE VII

EVALUATIONS ON THE INTER-CENTER GENERALIZATION OF VARIOUS APPROACHES. ALL MODELS ARE TRAINED ON THE XUANWU DATASET AND EVALUATED ON THE PPMI DATASET. THE EXPERIMENTS ARE REPEATED 10 TIMES AND THE AVERAGED RESULTS AS WELL AS DEVIATIONS ARE SHOWN IN THE TABLE.

	NC vs PD (inter-center)		
	Acc	Sen	Spe
SVM	60.1±3.1	61.9±3.3	58.3±3.2
MLP	70.5±4.8	75.4±7.8	68.8±7.5
BrainNetCNN [18]	72.6±5.2	78.0±10.8	74.9±13.8
Population-GCN [22]	75.0±6.1	79.0±7.7	73.0±6.7
BrainGNN [8]	<u>77.1±6.9</u>	**82.0±11.5**	74.7±9.1
M-GCN [43]	72.5±4.8	74.8±5.8	72.4±10.7
HGNN [32]	73.0±4.2	72.8±7.0	75.6±11.3
DHGNN [31]	74.2±3.2	71.1±6.3	78.4±9.4
MMP-GCN [44]	76.9±12.7	73.9±6.6	**81.5±8.1**
TAN [45]	74.1±7.3	72.5±11.3	76.8±12.6
Cross-GNN (Ours)	**78.1±4.6**	<u>80.5±8.3</u>	<u>79.6±4.1</u>

F. Biomarker detection

The correspondence matrices learned by Cross-GNN attentively model the multi-modal interactions among regions. In this section, we extracted the correspondence matrix for each subject and performed a group statistical analysis for biomarker detection.

Fig. 4 and Fig. 5 demonstrate the statistical analysis of the ADNI, the Xuanwu, and the PPMI datasets on the left side, where the regions with significant differences (p-value < 0.05) are displayed with the statistical values (cutoff from 2 to 5). As is shown in the figures, the number of significant regions is correlated to the classification performances. For example, in the ADNI dataset, the number of regions with significant differences in the classification of NC and PD is the largest, which corresponds to the results in Table IV, i.e., NC VS AD outperforms the tasks of NC VS MCI and MCI VS AD.

In detail, the right posterior cingulate gyrus in the default mode network was found with significant changes in the ADNI dataset among all tasks. Previous studies demonstrate that the AD process has been hypothetically explained by PCC hypofunction due to the effect of the degeneration of cingulum

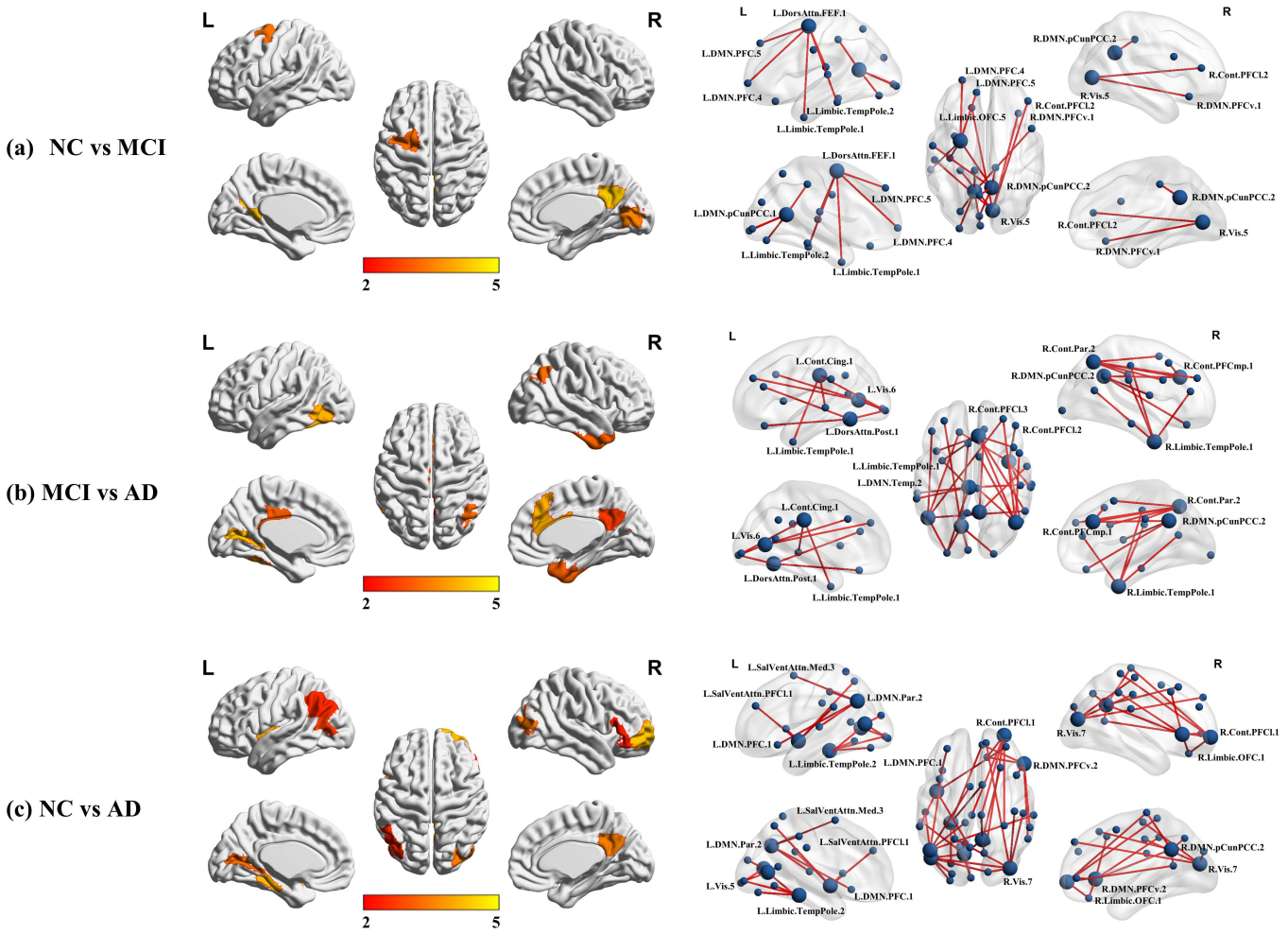


Fig. 4. Group-wise comparison on (a) NC vs. MCI, (b) MCI vs. AD. (c) NC vs. AD. Left: statistical results (F-statistic value) on the correspondence matrices on the Xuanwu dataset, where only significant regions (adjusted p-value $< \mathbf{0.05}$) are displayed. Right: key connectivity pairs returned from the significant regions. L: the left hemisphere. R: the right hemisphere. DMN: the default mode network. DorsAttn: the dorse attention network. Vis: the visual network. Cont: the frontoparietal network. PFC: the prefrontal cortex. FEF: the frontal eye fields. OFC: the orbital frontal cortex. Temp: the temporal. Par: the parietal. Cing: the cingulate.

fibers [65]. And PCC hypofunction could be caused by early PCC atrophy, which has been proven to be a useful biomarker [66]. Moreover, the right prefrontal cortex is found with significant changes in the Xuanwu and the PPMI datasets (Xuanwu: R.DMN.PFCv.2; PPMI: R.DMN.PFCv.1). PD is shown to be associated with impaired frontal lobe functions, which include defective use of memory stores and a dysexecutive syndrome of processes that are commonly considered to be controlled by the prefrontal cortex [67], [68]. Our finding is in agreement with previous studies that showed patients with PD to have changes in the prefrontal cortex, which plays a critical role in postural control in older adults [69], [70]. In addition, key biomarker regions such as the somatomotor network were also found in the PPMI dataset.

We further empirically selected five pair-wise connectivity with the greatest effect size for each seed region as the key structural connectivity. These connections are displayed on the right side of Fig. 4 and Fig. 5. Notably, during the AD progress, brain function changes are found in connectivity pairs, i.e. pairs between the right posterior cingulate gyrus in the default mode network and visual network, limbic network, and dorsal and ventral attention network. This indicates that the right posterior cingulate is highly functionally heterogeneous, but also has a central role in supporting internally-directed cognition [71], [72]. Moreover, reduced metabolism in the PCC is an early feature of Alzheimer's disease and is often present before a definitive clinical diagnosis [73], [74]. Our work has consistently shown abnormal PCC function in the early stage of AD. In addition, connections between the right prefrontal cortex in the default network and other networks (e.g. the somatomotor network, the frontal-parietal control network, and the limbic network) are found in the PD progress, which is consistent with previous studies [75]–[77]. These findings indicate that our proposed Cross-GNN could locate meaningful and interpretative key biomarkers.

VI. Conclusion

In this study, we propose a cross-modal graph neural network for multi-modal connectome analysis, which exploits a

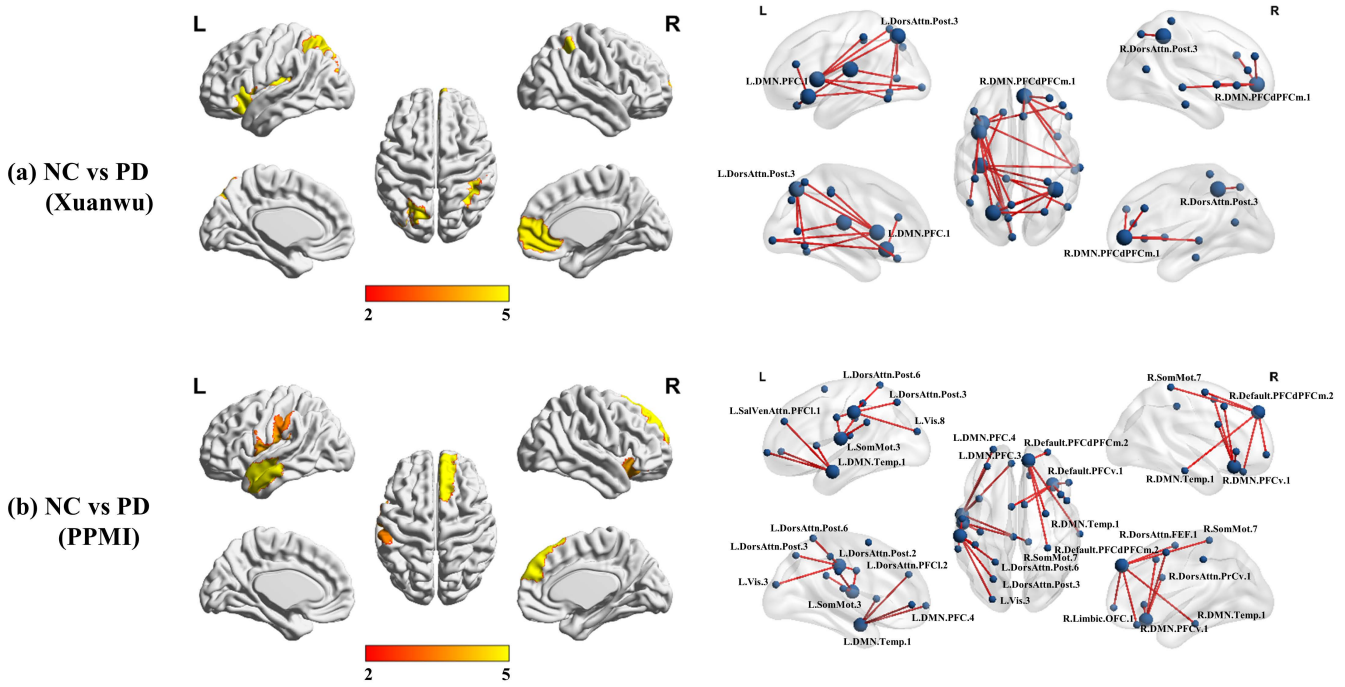


Fig. 5. Group-wise comparison on NC vs. PD on (a) the Xuanwu dataset, and (b) the PPMI dataset. Left: statistical results (F-statistic value) on the correspondence matrices, where only significant regions (adjusted p-value $< \mathbf{0.05}$) are displayed. Right: key connectivity pairs returned from the significant regions. L: the left hemisphere. R: the right hemisphere. DMN: the default mode network. DorsAttn: the dorse attention network. SomMot: the somotomoto network. PFC: the prefrontal cortex. Temp: the temporal. OFC: the orbital frontal cortex.

dynamic graph for tighter coupling of multi-modal representations with regularized knowledge. Cross-modal embedding and knowledge distillation are leveraged for mutual learning to capture inter-modal dependencies. Moreover, the derived correspondence matrix provides a compositional space for reasoning multi-modal dependencies. Experimental results on two datasets demonstrate that our proposed method is feasible to model multi-modal graphs and outperforms other state-of-the-art methods. In addition, our Cross-GNN provides a way of locating the discriminative multi-modal biomarkers. The proposed method has great potential for multi-modal connectome-based brain disorder diagnosis as well as multi-modal biomarker detection.

REFERENCES

[1] J. Pu, J. Wang, W. Yu, Z. Shen, Q. Lv, K. Zeljic, C. Zhang, B. Sun, G. Liu, and Z. Wang, "Discriminative structured feature engineering for macroscale brain connectomes," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, pp. 2333–2342, 2015.

[2] J. D. Gabrieli, S. S. Ghosh, and S. Whitfield-Gabrieli, "Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience," *Neuron*, vol. 85, no. 1, pp. 11–26, 2015.

[3] C. Ye, S. Mori, P. Chan, and T. Ma, "Connectome-wide network analysis of white matter connectivity in alzheimer's disease," *NeuroImage: Clinical*, vol. 22, p. 101690, 2019.

[4] Y. Yang, C. Ye, J. Sun, L. Liang, H. Lv, L. Gao, J. Fang, T. Ma, and T. Wu, "Alteration of brain structural connectivity in progression of parkinson's disease: a connectome-wide network analysis," *NeuroImage: Clinical*, vol. 31, p. 102715, 2021.

[5] E. T. Bullmore and D. S. Bassett, "Brain graphs: graphical models of the human brain connectome," *Annual review of clinical psychology*, vol. 7, pp. 113–140, 2011.

[6] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, 2006.

[7] A. Zalesky and A. Fornito, "A dti-derived measure of cortico-cortical connectivity," *IEEE transactions on medical imaging*, vol. 28, no. 7, pp. 1023–1036, 2009.

[8] X. Li, Y. Zhou, N. Dvornek, M. Zhang, S. Gao, J. Zhuang, D. Scheinost, L. H. Staib, P. Ventola, and J. S. Duncan, "Braingnn: Interpretable brain graph neural network for fmri analysis," *Medical Image Analysis*, vol. 74, p. 102233, 2021.

[9] H. Zhang, R. Song, L. Wang, L. Zhang, D. Wang, C. Wang, and W. Zhang, "Classification of brain disorders in rs-fmri via local-to-global graph neural networks," *IEEE Transactions on Medical Imaging*, 2022.

[10] Y. Yang, X. Guo, Z. Chang, C. Ye, Y. Xiang, and T. Ma, "Multi-modal dynamic graph network: Coupling structural and functional connectome for disease diagnosis and classification," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2022, pp. 1343–1349.

[11] P. Skudlarski, K. Jagannathan, V. D. Calhoun, M. Hampson, B. A. Skudlarska, and G. Pearlson, "Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations," *Neuroimage*, vol. 43, no. 3, pp. 554–561, 2008.

[12] M. Fukushima, R. F. Betzel, Y. He, M. P. van den Heuvel, X.-N. Zuo, and O. Sporns, "Structure–function relationships during segregated and integrated network states of human brain functional connectivity," *Brain Structure and Function*, vol. 223, no. 3, pp. 1091–1106, 2018.

[13] S. Atasoy, I. Donnelly, and J. Pearson, "Human brain networks function in connectome-specific harmonic waves," *Nature communications*, vol. 7, no. 1, pp. 1–10, 2016.

[14] A. Zalesky, A. Fornito, and E. T. Bullmore, "Network-based statistic: identifying differences in brain networks," *Neuroimage*, vol. 53, no. 4, pp. 1197–1207, 2010.

[15] Z. Shehzad, C. Kelly, P. T. Reiss, R. C. Craddock, J. W. Emerson, K. McMahon, D. A. Copland, F. X. Castellanos, and M. P. Milham, "A multivariate distance-based analytic framework for connectome-wide association studies," *Neuroimage*, vol. 93, pp. 74–94, 2014.

[16] C.-Y. Wee, P.-T. Yap, D. Zhang, K. Denny, J. N. Browndyke, G. G. Potter, K. A. Welsh-Bohmer, L. Wang, and D. Shen, "Identification of mci individuals using structural and functional connectivity networks," *Neuroimage*, vol. 59, no. 3, pp. 2045–2056, 2012.

[17] M. Plitt, K. A. Barnes, and A. Martin, "Functional connectivity classification of autism identifies highly predictive brain features but falls

short of biomarker standards," *NeuroImage: Clinical*, vol. 7, pp. 359–366, 2015.

[18] J. Kawahara, C. J. Brown, S. P. Miller, B. G. Booth, V. Chau, R. E. Grunau, J. G. Zwicker, and G. Hamarneh, "Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.

[19] T.-E. Kam, H. Zhang, Z. Jiao, and D. Shen, "Deep learning of static and dynamic brain functional networks for early mci detection," *IEEE transactions on medical imaging*, vol. 39, no. 2, pp. 478–487, 2019.

[20] B. Jie, M. Liu, C. Lian, F. Shi, and D. Shen, "Designing weighted correlation kernels in convolutional neural networks for functional connectivity based brain disease diagnosis," *Medical image analysis*, vol. 63, p. 101709, 2020.

[21] H. Huang, X. Hu, Y. Zhao, M. Makkie, Q. Dong, S. Zhao, L. Guo, and T. Liu, "Modeling task fmri data via deep convolutional autoencoder," *IEEE transactions on medical imaging*, vol. 37, no. 7, pp. 1551–1561, 2017.

[22] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, and D. Rueckert, "Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease," *Medical image analysis*, vol. 48, pp. 117–130, 2018.

[23] A. Kazi, S. Shekarforoush, S. Arvind Krishna, H. Burwinkel, G. Vivar, K. Kortüm, S.-A. Ahmadi, S. Albarqouni, and N. Navab, "Inceptiongcn: receptive field aware graph convolutional network for disease prediction," in *Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26*. Springer, 2019, pp. 73–85.

[24] X. Song, F. Zhou, A. F. Frangi, J. Cao, X. Xiao, Y. Lei, T. Wang, and B. Lei, "Graph convolution network with similarity awareness and adaptive calibration for disease-induced deterioration prediction," *Medical Image Analysis*, vol. 69, p. 101947, 2021.

[25] X. Song, A. Frangi, X. Xiao, J. Cao, T. Wang, and B. Lei, "Integrating similarity awareness and adaptive calibration in graph convolution network to predict disease," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23*. Springer, 2020, pp. 124–133.

[26] H. Jiang, P. Cao, M. Xu, J. Yang, and O. Zaiane, "Hi-gcn: a hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction," *Computers in Biology and Medicine*, vol. 127, p. 104096, 2020.

[27] L. Li, H. Jiang, G. Wen, P. Cao, M. Xu, X. Liu, J. Yang, and O. Zaiane, "Te-hi-gcn: An ensemble of transfer hierarchical graph convolutional networks for disorder diagnosis," *Neuroinformatics*, pp. 1–23, 2021.

[28] Y. Yang, C. Ye, and T. Ma, "A deep connectome learning network using graph convolution for connectome-disease association study," *Neural Networks*, 2023.

[29] S. I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. Lee, B. Glocker, and D. Rueckert, "Metric learning with spectral graph convolutions on brain connectivity networks," *NeuroImage*, vol. 169, pp. 431–442, 2018.

[30] Y. Zhang and H. Huang, "New graph-blind convolutional network for brain connectome data analysis," in *International Conference on Information Processing in Medical Imaging*. Springer, 2019, pp. 669–681.

[31] J. Jiang, Y. Wei, Y. Feng, J. Cao, and Y. Gao, "Dynamic hypergraph neural networks." in *IJCAI*, 2019, pp. 2635–2641.

[32] Y. Feng, H. You, Z. Zhang, R. Ji, and Y. Gao, "Hypergraph neural networks," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 3558–3565.

[33] X. Xing, Q. Li, H. Wei, M. Zhang, Y. Zhan, X. S. Zhou, Z. Xue, and F. Shi, "Dynamic spectral graph convolution networks with assistant task training for early mci diagnosis," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV*. Springer, 2019, pp. 639–646.

[34] K. Zhao, B. Duka, H. Xie, D. J. Oathes, V. Calhoun, and Y. Zhang, "A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in adhd," *Neuroimage*, vol. 246, p. 118774, 2022.

[35] Y. Li, J. Liu, Z. Tang, and B. Lei, "Deep spatial-temporal feature fusion from adaptive dynamic functional connectivity for mci identification," *IEEE Transactions on Medical Imaging*, vol. 39, no. 9, pp. 2818–2830, 2020.

[36] D. Salas-Gonzalez, J. Górriz, J. Ramírez, I. Illán, M. López, F. Segovia, R. Chaves, P. Padilla, C. Puntonet, and A. D. N. Initiative, "Feature selection using factor analysis for alzheimer's diagnosis using pet images," *Medical physics*, vol. 37, no. 11, pp. 6084–6095, 2010.

[37] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, "Comparison of adaboost and support vector machines for detecting alzheimer's disease through automated hippocampal segmentation," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 30–43, 2009.

[38] E. Westman, J.-S. Muehlboeck, and A. Simmons, "Combining mri and csf measures for classification of alzheimer's disease and prediction of mild cognitive impairment conversion," *Neuroimage*, vol. 62, no. 1, pp. 229–238, 2012.

[39] R. Liu, Z.-A. Huang, Y. Hu, Z. Zhu, K.-C. Wong, and K. C. Tan, "Attention-like multimodality fusion with data augmentation for diagnosis of mental disorders using mri," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[40] Y. Huang and A. C. Chung, "Disease prediction with edge-variational graph convolutional networks," *Medical Image Analysis*, vol. 77, p. 102375, 2022.

[41] J. Ji and Y. Zhang, "Functional brain network classification based on deep graph hashing learning," *IEEE Transactions on Medical Imaging*, vol. 41, no. 10, pp. 2891–2902, 2022.

[42] Y. Wang, Y. Yang, X. Guo, C. Ye, N. Gao, Y. Fang, and H. T. Ma, "A novel multimodal mri analysis for alzheimer's disease based on convolutional neural network," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 754–757.

[43] N. S. Dsouza, M. B. Nebel, D. Crocetti, J. Robinson, S. Mostofsky, and A. Venkataraman, "M-gcn: A multimodal graph convolutional network to integrate functional and structural connectomics data to predict multidimensional phenotypic characterizations," in *Medical Imaging with Deep Learning*. PMLR, 2021, pp. 119–130.

[44] X. Song, F. Zhou, A. F. Frangi, J. Cao, X. Xiao, Y. Lei, T. Wang, and B. Lei, "Multi-center and multi-channel pooling gcn for early ad diagnosis based on dual-modality fused brain network," *IEEE Transactions on Medical Imaging*, 2022.

[45] Q. Zhu, H. Wang, B. Xu, Z. Zhang, W. Shao, and D. Zhang, "Multimodal triplet attention network for brain disease diagnosis," *IEEE Transactions on Medical Imaging*, 2022.

[46] G. Hinton, O. Vinyals, J. Dean *et al.*, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, vol. 2, no. 7, 2015.

[47] L. Zhang, J. Song, A. Gao, J. Chen, C. Bao, and K. Ma, "Be your own teacher: Improve the performance of convolutional neural networks via self distillation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 3713–3722.

[48] Y. Yang, G. Xutao, C. Ye, Y. Xiang, and T. Ma, "Regularizing brain age prediction via gated knowledge distillation," in *Medical Imaging with Deep Learning*, 2021.

[49] W. Hou, L. Wang, S. Cai, Z. Lin, R. Yu, and J. Qin, "Early neoplasia identification in barrett's esophagus via attentive hierarchical aggregation and self-distillation," *Medical image analysis*, vol. 72, p. 102092, 2021.

[50] D. Qin, J.-J. Bu, Z. Liu, X. Shen, S. Zhou, J.-J. Gu, Z.-H. Wang, L. Wu, and H.-F. Dai, "Efficient medical image segmentation based on knowledge distillation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3820–3831, 2021.

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

[52] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[53] K. Kim, B. Ji, D. Yoon, and S. Hwang, "Self-knowledge distillation with progressive refinement of targets," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 6567–6576.

[54] D. Zhu, T. Zhang, X. Jiang, X. Hu, H. Chen, N. Yang, J. Lv, J. Han, L. Guo, and T. Liu, "Fusing dti and fmri data: a survey of methods and applications," *NeuroImage*, vol. 102, pp. 184–191, 2014.

[55] A. A. Baird, M. K. Colvin, J. D. VanHorn, S. Inati, and M. S. Gazzaniga, "Functional connectivity: integrating behavioral, diffusion tensor imaging, and functional magnetic resonance imaging data sets," *Journal of cognitive neuroscience*, vol. 17, no. 4, pp. 687–693, 2005.

[56] M. Ystad, E. Hodneland, S. Adolfsdottir, J. Haász, A. J. Lundervold, T. Eichele, and A. Lundervold, "Cortico-striatal connectivity and cognition in normal aging: a combined dti and resting state fmri study," *Neuroimage*, vol. 55, no. 1, pp. 24–31, 2011.

[57] D. B. McArtor, G. H. Lubke, and C. Bergeman, "Extending multivariate distance matrix regression with an effect size measure and the asymptotic null distribution of the test statistic," *Psychometrika*, vol. 82, pp. 1052–1077, 2017.

[58] C. Craddock, S. Sikka, B. Cheung, R. Khanuja, S. S. Ghosh, C. Yan, Q. Li, D. Lurie, J. Vogelstein, R. Burns *et al.*, "Towards automated

analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac)," *Front Neuroinform*, vol. 42, pp. 10–3389, 2013.
[59] J.-D. Tournier, F. Calamante, and A. Connelly, "Mrtrix: diffusion tractography in crossing fiber regions," *International journal of imaging systems and technology*, vol. 22, no. 1, pp. 53–66, 2012.
[60] A. W. Anderson, "Measurement of fiber orientation distributions using high angular resolution diffusion imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 54, no. 5, pp. 1194–1206, 2005.
[61] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly, "Sift: Spherical-deconvolution informed filtering of tractograms," *Neuroimage*, vol. 67, pp. 298–312, 2013.
[62] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. Yeo, "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri," *Cerebral cortex*, vol. 28, no. 9, pp. 3095–3114, 2018.
[63] G. Li, M. Muller, A. Thabet, and B. Ghanem, "Deepgcns: Can gcns go as deep as cnns?" in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 9267–9276.
[64] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" *arXiv preprint arXiv:1810.00826*, 2018.
[65] I. H. Choo, D. Y. Lee, J. S. Oh, J. S. Lee, D. S. Lee, I. C. Song, J. C. Youn, S. G. Kim, K. W. Kim, J. H. Jhoo *et al.*, "Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and alzheimer's disease," *Neurobiology of aging*, vol. 31, no. 5, pp. 772–779, 2010.
[66] T. C. Chua, W. Wen, X. Chen, N. Kochan, M. J. Slavin, J. N. Trollor, H. Brodaty, and P. S. Sachdev, "Diffusion tensor imaging of the posterior cingulate is a useful biomarker of mild cognitive impairment," *The American Journal of Geriatric Psychiatry*, vol. 17, no. 7, pp. 602–613, 2009.
[67] A. Brück, T. Kurki, V. Kaasinen, T. Vahlberg, and J. Rinne, "Hippocampal and prefrontal atrophy in patients with early non-demented parkinson's disease is related to cognitive impairment," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 75, no. 10, pp. 1467–1469, 2004.
[68] B. Dubois and B. Pillon, "Cognitive deficits in parkinson's disease," *Journal of neurology*, vol. 244, pp. 2–8, 1996.
[69] J. R. Mahoney, R. Holtzer, M. Izzetoglu, V. Zemon, J. Verghese, and G. Allali, "The role of prefrontal cortex during postural control in parkinsonian syndromes a functional near-infrared spectroscopy study," *Brain research*, vol. 1633, pp. 126–138, 2016.
[70] S. Stuart and M. Mancini, "Prefrontal cortical activation with open and closed-loop tactile cueing when walking and turning in parkinson disease: a pilot study," *Journal of Neurologic Physical Therapy*, vol. 44, no. 2, pp. 121–131, 2020.
[71] R. Leech and D. J. Sharp, "The role of the posterior cingulate cortex in cognition and disease," *Brain*, vol. 137, no. 1, pp. 12–32, 2014.
[72] J. M. Pearson, S. R. Heilbronner, D. L. Barack, B. Y. Hayden, and M. L. Platt, "Posterior cingulate cortex: adapting behavior to a changing world," *Trends in cognitive sciences*, vol. 15, no. 4, pp. 143–151, 2011.
[73] S. Minoshima, B. Giordani, S. Berent, K. A. Frey, N. L. Foster, and D. E. Kuhl, "Metabolic reduction in the posterior cingulate cortex in very early alzheimer's disease," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 42, no. 1, pp. 85–94, 1997.
[74] K. A. Johnson, K. Jones, B. L. Holman, J. A. Becker, P. A. Spiers, A. Satlin, and M. S. Albert, "Preclinical prediction of alzheimer's disease using spect," *Neurology*, vol. 50, no. 6, pp. 1563–1571, 1998.
[75] R. Randver, "Repetitive transcranial magnetic stimulation of the dorsolateral prefrontal cortex to alleviate depression and cognitive impairment associated with parkinson's disease: A review and clinical implications," *Journal of the Neurological Sciences*, vol. 393, pp. 88–99, 2018.
[76] A. Nagano-Saito, Y. Washimi, Y. Arahata, T. Kachi, J. Lerch, A. Evans, A. Dagher, and K. Ito, "Cerebral atrophy and its relation to cognitive impairment in parkinson disease," *Neurology*, vol. 64, no. 2, pp. 224–229, 2005.
[77] J.-S. Chu, T.-H. Liu, K.-L. Wang, C.-L. Han, Y.-P. Liu, S. Michitomo, J.-G. Zhang, T. Fang, and F.-G. Meng, "The metabolic activity of caudate and prefrontal cortex negatively correlates with the severity of idiopathic parkinson's disease," *Aging and disease*, vol. 10, no. 4, p. 847, 2019.