

A New Brain Network Construction Paradigm for Brain Disorder via Diffusion-Based Graph Contrastive Learning

Yongcheng Zong, Qiankun Zuo[✉], Michael Kwok-Po Ng[✉], *Senior Member, IEEE*,
Baiying Lei[✉], *Senior Member, IEEE*, and Shuqiang Wang[✉], *Senior Member, IEEE*

Abstract—Brain network analysis plays an increasingly important role in studying brain function and the exploring of disease mechanisms. However, existing brain network construction tools have some limitations, including dependency on empirical users, weak consistency in repeated experiments and time-consuming processes. In this work, a diffusion-based brain network pipeline, DGCL is designed for end-to-end construction of brain networks. Initially, the brain region-aware module (BRAM) precisely determines the spatial locations of brain regions by the diffusion process, avoiding subjective parameter selection. Subsequently, DGCL employs graph contrastive learning to optimize brain connections by eliminating individual differences in redundant connections unrelated to diseases, thereby enhancing the consistency of brain networks within the same group. Finally, the node-graph contrastive loss and classification loss jointly constrain the learning process of the model to obtain the reconstructed brain network, which is then used to analyze important brain connections. Validation on two datasets, ADNI and ABIDE, demonstrates that DGCL surpasses traditional methods and other deep learning models in predicting disease development stages. Significantly, the proposed model improves the efficiency and generalization of brain network construction. In summary, the proposed DGCL can be served as a universal brain network construction scheme, which can effectively identify important brain connections through generative paradigms and has the potential to provide disease interpretability support for neuroscience research.

Index Terms—Alzheimers disease, autism spectrum disorder, brain network, graph contrastive learning, region-aware diffusion.

I. INTRODUCTION

BRAIN diseases represent a broad spectrum of conditions that significantly impact both individuals and society as a whole. These disorders such as Alzheimer's disease (AD) and autism spectrum disorder (ASD) encompass a wide range of brain-related challenges, affecting individuals across different age groups and backgrounds [1]. The complexity of these disorders, coupled with their varied etiologies, makes them a formidable challenge for researchers and clinicians alike. Furthermore, the absence of precise biomarkers and the often subtle nature of early-stage symptoms make early detection a daunting task [2]. Neuroimaging technologies such as magnetic resonance imaging can provide non-invasive and accurate diagnostic methods, and become an important auxiliary means to diagnose brain diseases. Magnetic resonance imaging (MRI) is a non-invasive, repeatable, and high spatial resolution technique widely used in the diagnosis and research of various brain diseases [3]. Among them, diffusion tensor imaging (DTI) is a special form of magnetic resonance imaging that can be used to describe brain structures. According to the DTI reconstruction algorithm, the direction of white matter fiber bundles can be obtained, and finally the structural join of brain networks can be constructed. This technology is widely used for the effective identification of various neurodegenerative diseases and the classification of AD development stages, especially for the recognition of abnormal brain connections [4].

To address these challenges, recent years have seen the rapid development of deep learning techniques in the domain of neuroimaging analysis, which has, in turn, been marked by considerable success [5], [6], [7]. Among them, graph neural networks (GNN) hold promise for modeling intricate, complex networks, such as brain networks [8], [9], [10]. Researchers have increasingly employed GNNs to study brain networks, improving the interpretability of machine learning in brain network analysis [11]. However, these methods typically require preprocessed graph structures as input, which are obtained by medical imaging software such as SPM, resulting in uncertainty of generated brain networks. In these studies, considering the differences in the target dataset and subjective selection of

Manuscript received 24 January 2024; revised 3 July 2024; accepted 8 August 2024. Date of publication 13 August 2024; date of current version 5 November 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFC2506902, in part by the National Natural Science Foundations of China under Grant 62172403, Grant 12326614, and Grant U22A2024, in part by the Distinguished Young Scholars Fund of Guangdong under Grant 2021B1515020019, in part by Hong Kong Research Grants Council (HKRGC) under Grant 17201020 and Grant 17300021, Grant HKRGC CRF C7004-21GF, Grant Joint NSFC-RGC N-HKU769/21, and Grant HLCA/P-211/24. Recommended for acceptance by S. Ji. (Yongcheng Zong and Shuqiang Wang contributed equally to this work.) (Corresponding authors: Baiying Lei; Shuqiang Wang.)

Yongcheng Zong and Shuqiang Wang are with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: sq.wang@siat.ac.cn).

Qiankun Zuo is with the Hubei Key Laboratory of Digital Finance Innovation, School of Information Engineering, Hubei University of Economics, Wuhan, Hubei 430205, China.

Michael Kwok-Po Ng is with the Department of Mathematics, Hong Kong Baptist University, Hong Kong.

Baiying Lei is with the Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen 518060, China (e-mail: leiby@szu.edu.cn).

Digital Object Identifier 10.1109/TPAMI.2024.3442811

various parameters, there will be significant differences in the preprocessing steps. In addition, traditional software heavily relies on setting a large number of parameters, and both optimizing and selecting the correct parameters requires a lot of empirical guidance. Different parameter settings may affect the preprocessing results [12]. When constructing brain networks, traditional research methods often rely on statistical analysis at the group level, mainly using graph attributes and related graph theory indicators for classification. At the same time, traditional software-based methods (such as PANDA) also face some difficulties in non-standard individual space image registration, alignment, and other issues. Overall, due to the complex brain structure and individual differences, research on individual-level brain networks and connectivity changes remains a huge challenge.

Compared with traditional software-based methods, the proposed method is completely data-driven and end-to-end processing flow, avoiding the tedious image data preprocessing steps on traditional software, and improving operational efficiency and accuracy. In this paper, we propose a novel deep learning-based method named DGCL for constructing brain networks, and the main contributions are summarized as follows.

- 1) A new paradigm for constructing brain networks has been proposed, which is completely an end-to-end pipeline for constructing brain networks from brain images. Compared with existing methods that rely on software templates, it can reduce technology dependence, improve the running efficiency, and generate consistent brain networks.
- 2) A Brain Region-Aware Module (BRAM) based on the diffusion mechanism is proposed, which can align and register unprocessed images in non-standard native spaces. After learning more precise information regarding the divisions of brain regions, BRAM demonstrates the capacity to automatically align a wide range of brains, spanning from children to adults.
- 3) The Brain Network Optimization Module (BNOM) utilizes the idea of graph contrastive learning to effectively remove redundant features such as individual differential connections that are unrelated to diseases. By retaining common connections between subjects and enhancing important brain connections, the consistency and robustness of brain network construction have been improved.

II. RELATED WORK

There has been an increasing amount of research on brain network construction and analysis in recent years, particularly for Alzheimer's disease diagnosis and treatment. Considering the lack of research on using diffusion models for the brain network generation, we will review the current research status from three aspects.

A. Brain Network Analysis Methods

For the research of AD diagnosis, the most famous public neuroimaging data set comes from the Alzheimer's Disease Neuroimaging Initiative (ADNI), which is committed to improving the clinical trials of AD prevention and treatment [13].

And the largest dataset for autism research comes from Autism Brain Imaging Data Exchange (ABIDE) [14]. Neuroscientists use neural image data to model the brain as a graph network structure, where graph nodes represent anatomical brain regions connected by edges, while graph edges can be defined according to brain morphology, function or structural connectivity related to paired nodes [15]. Different modalities of brain imaging data are commonly used to construct three types of brain networks: morphological networks, structural networks, and functional networks, respectively, from structural T1 weighted imaging (T1WI), diffusion-weighted imaging (DTI), and resting state functional MRI (rs fMRI). Y. He et al. constructed the brain morphological brain network for the first time in 2007 using the cortical thickness of various brain regions as morphological features [16], and discovered the small world attribute of the brain network. Afterwards, a large number of complex network analysis methods based on graph theory began to study the connection patterns and graph attribute analysis between brain regions, which were used to explore the diagnosis of brain functional disorders such as mental disorders from a structural to functional perspective. To simplify brain network analysis, the threshold method is usually used to sparsize the network. Research has shown that sparsization operations have a significant impact on subsequent analysis [12]. So far, there is still no optimal standard for determining the threshold. In experiments, cross-validation is generally used to traverse all possible value ranges to determine the optimal threshold.

B. Graph Neural Network Methods

In recent years, with the vigorous development of deep learning theory, various advanced neural network models have been used for the analysis and processing of medical images, especially the success of graph neural networks, which provides a new research direction for traditional brain network models based on statistical analysis. Graph neural networks are one of the current research hotspots in deep learning technology, and have become the most widely used method for extracting information from graphs to perform the diagnosis of brain diseases. J. Ma et al. innovatively proposed an attention-guided deep map neural network in 2020 for analyzing AD data from sMRI [17], which can reveal the most relevant brain regions and basic time points of AD progression. J. Gan et al. proposed a multi-graph fusion method in 2021 to explore common and complementary information between functional brain networks [18]. Y. Zhu et al. proposed a dynamic graph convolutional network in 2022 and used feature learning techniques to make it easier to interpret [19]. H. Cui et al. first proposed a benchmark for brain network analysis using message-passing graph convolutional networks in 2022 [20], providing direction for establishing a unified, scalable, and replicable GNN brain network analysis framework. However, combining these advanced graph contrastive learning methods with brain network research still has great potential and exploration space. However, these methods rely on software preprocessing to construct initial brain networks, requiring manual parameter settings on different modal datasets, making it difficult to replicate and fairly compare the

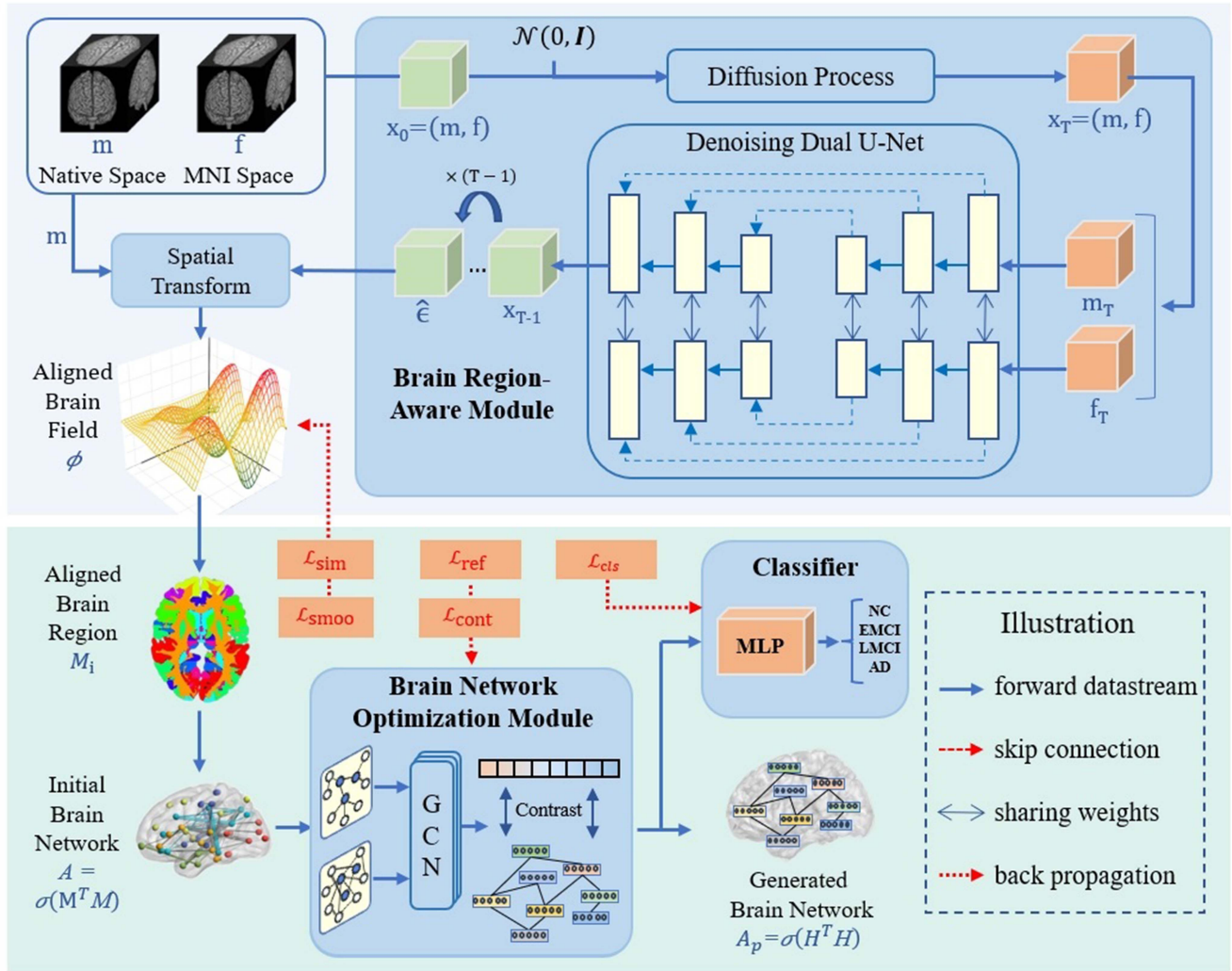


Fig. 1. The network architecture of the proposed DGCL has three sub-modules: (a) Brain Region-Aware Module, (b) Brain Network Optimization Module, and (c) MLP Classifier. For the convenience of description, we only show the process of the second part and the implementation details are illustrated in Fig. 2.

final experimental results under the same conditions. In terms of graph contrastive learning, P. Veličković et al. [21] proposed unsupervised graph contrast learning to learn meaningful node embedding by maximizing mutual information between local node representations and global graph representations. Y. You et al. [22] proposed a novel graph contrast learning framework (GraphCL) for GNN pre-training, which uses four graph enhancement strategies (such as node dropping, edge dropping) to generate different views of graphs, and constructs a general framework. In a recent study, X. Shen et al. [23] proposed Neighbor Contrastive Learning on Learnable Graph Augmentation, which avoids artificially designed graph augmentation strategy and achieves better results when node labels are more scarce. Although the research on graph comparison learning algorithms is very active and has made significant progress in multiple fields, their application in brain network research is still relatively limited. The current research mainly focuses on traditional fields such as social networks, biological networks, and knowledge graphs.

C. Deep Learning Methods

Another end-to-end learning approach attempts to automatically learn features from each voxel in brain images [24], [25]. For example, in the field of natural images, convolutional neural networks (CNN) based methods learn all levels of features from the original pixels, avoiding the process of manual ROI (regions of interest) annotation. Such methods can model or focus on disease-related brain areas such as gray matter or hippocampus at the whole brain level [26]. S. Sarraf et al. used functional magnetic resonance imaging data for deep learning applications for the first time in 2016 [27]. In the preprocessing step, 4D rs-fMRI and 3D MRI data are decomposed into 2D format images, and then these images are received in their input layer based on a CNN architecture. Feng et al. realized a deep learning network based on 3D-CNN and full stack bidirectional short-term memory (FSBi LSTM) in 2019 [28]. Among them, the 3D-CNN network is used to extract features, while FSBi-LSTM extracts advanced semantic and spatial information. W. Kang et al. borrowed from the idea of ensemble learning in 2021

and effectively improved classification accuracy and stability by integrating multiple MRI 2D slices [29]. Loddo et al. conducted a comparative study on the work of deep learning in the field of AD diagnosis in 2022, and then used transfer learning strategies on different heterogeneous data sets to detect AD diseases and distinguish dementia of different degrees. The experimental results show that among the CNN network types considered, ResNet-50 and ResNet-101 models are the most appropriate solutions for transfer learning [30], [31]. Although deep learning methods are convenient and efficient, with continuous breakthroughs in the accuracy of results, most of the existing work focuses on brain segmentation or classification tasks, making it difficult to accurately locate abnormal changes in brain regions and lacking interpretability in neuroscience.

III. METHODOLOGY

As shown in Fig. 1, a two-stage model is proposed for medical brain imaging disease diagnosis, which includes a Brain Region-Aware Module (BRAM) for brain region localization and feature extraction and a Brain Network Optimization Module (BNOM) to obtain better brain network connection. In the first stage, the BRAM uses the standard AAL template [32] to guide brain region localization and alignment to the standard brain space based on diffusion mechanism, automatically learning boundary and position information for brain regions. In the second stage, inspired by the idea of graph contrastive learning [33], the BNOM utilizes the graph or brain network to eliminate individual differences between populations that are unrelated to diseases, generating stable and reliable brain networks.

A. Brain Region-Aware Module

1) *Brain Region Feature Extraction and Alignment*: As is well known, CNNs can extract local features and obtain global information representation through multi-layer stacking. In order to identify targets of different sizes, we need convolutional kernels of different sizes. For the human brain, the cerebral cortex is full of wrinkles, but the boundary of the brain area is not smooth. In addition, different brain regions have different sizes, so refined brain region division is a challenging problem. In order to accurately extract information from various brain regions, we use AAL templates as supervision and use registration models to achieve brain region alignment and capture the features of each brain region.

Image registration methods based on deep learning can estimate the nonrigid voxel correspondence between two images, namely moving and fixed image pairs. The registration methods generally find the optimal voxel deformation relationship between the two images by optimizing the registration field according to the following formula:

$$\psi^* = \arg \min_{\psi} \mathcal{L}_{sim}(f, m \circ \psi) + \mathcal{L}_{reg}(\psi) \quad (1)$$

where ψ^* is the optimal registration field for transforming a moving image into a fixed image, and m is the moving image and f is the fixed image. \mathcal{L}_{sim} is the similarity function, which is used to calculate the similarity between the deformed image

and the fixed image, and \mathcal{L}_{reg} is the regularization penalty for the registration field.

2) *The Structure of Proposed BRAM*: Traditional medical registration models typically consist of two networks, where the registration network receives inputs from f and m to generate a registration field. Then input f and the registration field into the spatial transformation network to obtain the final deformation image. We propose a new unsupervised image registration method based on the diffusion model. The image registration method, by introducing the powerful image generation ability of diffusion, generates the image deformation field in the first stage from noise. RAM consists of two sub-networks: denoising dual U-Net and deformation network, the same spatial transformation network in VoxelMorph [34].

The denoising network estimates the randomly increasing noise in the diffusion forward process, while the deformable network accepts the estimated noise input and the moving image m , and outputs the deformation field. We introduce time embedding into the network to enable the deformable network to learn image features under different noise levels. Since the input DTI image is 3D, the deformation function in the deformation network uses Trilinear interpolation.

The registered image warps the moving image m by using the spatial transformation layer [35]. For 3D image transformation, the transformation function of Trilinear interpolation is used. Obtaining deformed image m' through transformation:

$$m' = m \circ \psi \quad (2)$$

The final output of the first stage model is:

$$M_i = m' \cdot \text{mask}(\text{region}_i) \quad (3)$$

3) *Reverse Alignment Process*: The reverse diffusion process is mainly based on the main structure of the DDPM [36], and the reverse process is used to estimate a conditional score function with a Diffusion network D_{θ} . Specifically, for a moving source image m and a fixed reference image f , the Diffusion network D_{θ} is trained to learn the conditional score function of the deformation between the moving image and the fixed image under the given condition $c = (m, f)$. Therefore, the target latent variable x_t is sampled, and the fixed image is defined as the target, i.e., $x_0 = f$. The loss function of the Diffusion network is as follows:

$$\mathcal{L}_{\text{diffusion}}(c, x_t, t) = \mathbb{E}_{\epsilon, x_t, t} \|D_{\theta}(c, x_t, t) - \epsilon\|_2^2 \quad (4)$$

The optimization objective of the Diffusion model is:

$$\min_{D_{\theta}, M_{\psi}} \mathcal{L}_{\text{diffusion}}(c, x_t, t) + \lambda \mathcal{L}_{\text{regist}}(m, f) \quad (5)$$

where $\mathcal{L}_{\text{regist}}$ is the registration loss, which includes two loss functions of image constraint, \mathcal{L}_{sim} and \mathcal{L}_{smoo} . M is the input moving images, F is the input reference images. Finally, the optimization objective of this module is

$$\psi = \arg \min \mathcal{L}_{sim}(F, M \circ \psi) + \mu \mathcal{L}_{smoo}(\psi) \quad (6)$$

where \mathcal{L}_{sim} is the loss of similarity in image structure, \mathcal{L}_{smoo} is a smoothness constraint on the gradient of the deformation field [37].

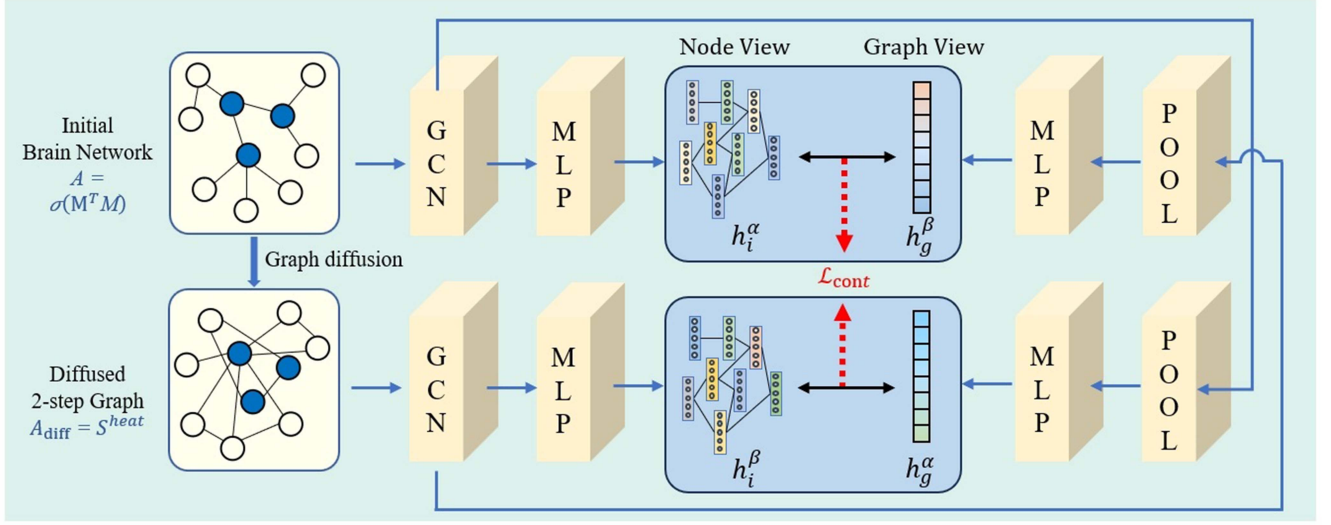


Fig. 2. The detailed architecture of Brain Network Optimization Module. The network takes two graphs as input, one is the original adjacency matrix A and another is the diffusion matrix A_{diff} . The diffusion matrix is an augmentation of the original matrix, providing an additional structural view. This model consists of two parallel GCNs [38], where pooling and MLP share weights to learn node representations in brain regions. The learned features are then fed to a graph pooling layer to obtain the final graph embedding representation.

Algorithm 1: Aligned Image Generation Process.

Input : Conditional images $c = (m, f)$, where m denotes the moving image in native space and f denotes the fixed image in MNI space.

- 1 SET $T \in (0, T_{\text{train}})$;
- 2 SAMPLE $x_T = \sqrt{\bar{\alpha}_T}m + \sqrt{1 - \bar{\alpha}_T}\epsilon$, where $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$, and $\epsilon \sim \mathcal{N}(0, I)$;
- 3 **for** $t = T, T-1, \dots, 1$ **do**
- 4 $z \sim \mathcal{N}(0, I)$;
- 5 $x_{t-1} = \frac{1}{\sqrt{\alpha_t}}(x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}}D_\theta(c, x_t, t)) + \sigma_t z$;
- 6 **end**
- 7 **return** x_0 ;

Result: Deformed image to align.

After training the Diffusion network D_θ , we use the following Algorithm 1 to generate aligned images from the original native spaces.

B. Brain Network Optimization Module

Mining brain network representations using single-modal images is challenging. To address this issue, this section introduces a strategy of graph contrastive learning. Using graph diffusion technology, the extracted original brain network graph and the diffused graph based on hot kernel features are simultaneously input into the graph neural network representation reconstruction module. And based on the similarity of node embedding and graph embedding features, the model is guided to learn the common connections between similar samples and the differences between different categories, improving the performance of the initial single mode constructed brain network.

Reconstructing the embedding features of brain network nodes involves the following three steps. 1) Sample the original brain network separately and construct a set of subgraphs; 2) Aggregating node features of subgraphs and diffused graphs through two layers of GCN; 3) Calculate the similarity of node embedding and graph embedding by contrastive loss function. The specific design of this module can be seen in Fig. 2.

The diffusion map is generated by the thermonuclear features, and the formula is as follows:

$$S^{\text{heat}} = \exp(tAD^{-1} - t) \quad (7)$$

where A represents the adjacency matrix, D represents the corresponding degree matrix, and t is the diffusion time [39]. The diffusion process on the graph can be seen as introducing local information within the neighborhood to the node. The subsampling method involves randomly sampling nodes and their edges from a graph, and selecting precise nodes and edges from another graph, with each graph serving as a different view.

The loss function designed for contrastive learning is as follows:

$$\mathcal{L}_{\text{cont}} = \max_{\theta, \omega, \phi, \psi} \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \sum_{i=1}^{|g|} [\text{MI}(h_i^\alpha, h_g^\beta) + \text{MI}(h_i^\beta, h_g^\alpha)] \quad (8)$$

where MI represents the mutual information function.

Graph contrastive learning of brain networks involves the following processes: 1) Extract graph embedding features from node features through the readout function; 2) Calculate the dot product between node embedding and graph embedding on another view; 3) Update model losses based on similarity, calculate loss gradients, and update network weights.

In order to keep the symmetry of the generated brain network matrix, we conduct matrix inner product operation on the extracted embedding, and limit it to a range of 0 to 1 through the

activation function [40]. The reconstruction loss of edges adopts mean square error:

$$\mathcal{L}_{\text{ref}}(\mathcal{G}) = -\frac{1}{90 \times 90} \sum_{i,j} (A_p[i,j] - A_{\text{ref}}[i,j])^2 \quad (9)$$

where $A_p[i,j] = \sigma(h_i^\top h_j)$.

C. Classifier Design and Model Training

The proposed DGCL ultimately guides brain network generation through disease classification task. The classifier knowledge contains a graphical representation of rich pathological information through multi-layer perceptron aggregation, and is connected with the Softmax function to output the prediction probability of each category of diseases. This module consists of three layers of MLP (Multilayer Perceptron), two layers of ReLU activation function and the Softmax output layer. The Softmax input contains four neurons, mapping the results to the probabilities of four disease categories. The Dropout strategy is used to prevent the model from overfitting [41].

Multi-class cross entropy function is adopted for multi-classification loss [42]. In our study, the number of classes C is 4.

$$\mathcal{L}_{\text{cls}}(p) = -\sum_{i=0}^{C-1} y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \quad (10)$$

At last, the total loss function of DGCL is as follows:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{cont}} + \mathcal{L}_{\text{cls}}(p) + \alpha \mathcal{L}_{\text{ref}}(\mathcal{G}) \quad (11)$$

The overall pipeline of the training process can be summarized in Algorithm 2.

IV. EXPERIMENTS

A. Datasets and Preprocessing

This study utilizes DTI from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [13], an open-source and public dataset to validate the proposed framework. A total of 349 subjects' data are collected, including normal control group (NC), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI), and AD. For the study of ASD, we collect brain imaging data from 160 ASD and NC subjects at San Diego University, NYU Langone Medical Center, and Trinity Centre for Health Sciences State, respectively. Tables I and III provides detailed information about the sample size, gender, and age of all subjects. The PANDA toolbox [43] is used for preprocessing the raw DTI data to obtain the reference structural brain network matrix [44], [45]. The process involves converting the initial DICOM format of the data to NIFTI format, skull stripping, fiber bundle resampling, and head motion correction. Then we calculate the fractional anisotropy (FA) coefficients by fitting the tensor model using the least squares method, and output the DTI data. After resampling, all subjects have $91 \times 109 \times 91$ voxels in DTI, with a voxel size of $2 \text{ mm} \times 2 \text{ mm} \times 2 \text{ mm}$. After that, we register images to the individual brain space, and construct the empirical structural connectivity \hat{A} based on the deterministic fiber tracking method by setting tracking

conditions, network nodes, and tracking stopping conditions. The brain is divided into 90 ROIs based on the AAL atlas [32], with each ROI defined as a node in the brain network. Finally, the structural connectivity of the brain network is determined by fiber tracking between different ROIs. Specifically, the stopping criteria for fiber tracking are defined as follows: (1) the crossing angle between two consecutive directions is greater than 45 degrees, and (2) the anisotropy score value is not within the range of [0.2, 1.0].

B. Experiment Settings

The model is trained and tested using the PyTorch platform, with an NVIDIA RTX 4000 GPU with 20GB memory. During the training process, the optimizer is set to Adam [46], with an initial learning rate of 0.0001, which exponentially decays with the number of training iterations. The number of epochs is set to 300, and the batch size is set to 16. Five-fold cross-validation is used to evaluate the performance of the model. All subjects are randomly divided into five equally sized subsets. One subset is treated as the test set, and the union of the other four subsets is treated as the training set. This process is repeated five times to eliminate bias. The evaluation metrics for the classification performance in this study are accuracy (ACC), sensitivity (SEN), specificity (SPE), and the area under the ROC curve (AUC).

The learnable parameter in BRAM is initialized according to the Xavier scheme. The structure of BNOM is described in detail in Section III-B. We use two GCNs to progressively extract high-order topological features of brain regions, with a brain region feature vector dimension of $d = 128$. The learnable parameter is initialized with the identity matrix. The Discriminator consists of two FC layers, with a LeakyReLU activation function in the hidden layers and no activation function in the output layer. The classification feature $h \in \mathbb{R}^{90 \times 128}$ is inputted into the FC layer for disease prediction, with 128 neurons in the input layer and 4 neurons in the output layer.

C. Ablation Study and Analysis

1) *Hyperparameters Analysis*: In this study, we conduct experiments to investigate the effect of two hyperparameters, namely, λ and μ , in the loss function of the BRAM. λ represents the loss coefficient of the alignment network, while μ represents the constraint coefficient of the smoothness of the deformation field gradient. We also analyze the effect of the α hyperparameter, which is the coefficient of the reference brain network, on the classification performance during the joint training phase. λ is set to 2.0, the optimal value according to the original paper, representing the ratio between the diffusion model and the registration network weights. Then we adjust the parameter μ in the following experiments. In the brain network optimization phase, we evaluate the ACC and AUC values of the binary classification task (AD versus NC) by changing α from 0.1 to 0.9 while keeping other parameters constant.

From the experimental results in Table III above, it can be seen that there is little difference between different methods, but the BRAM module using the diffusion mechanism can get the best effect, in which the Dice coefficient under the hyperparameters

Algorithm 2: The End-to-End Brain Network Construction Pipeline.

Input : X : training dataset of original brain images, and each image is a moving image, and x_{mni} represents the fixed image; A_{ref} : the empirical connection matrix of brain network; Y : the labels of corresponding subjects; T : the maximum timestamp; N : the number of training epochs.

```

1 Randomly initialize the parameters of Dual U-Net model;
2 for  $s \leftarrow 1$  to  $N$  do
3   for  $x \in X$  do
4     Train the Denoising Dual U-Net;
5     Repeat ;
6        $x \sim q(x_0)$  ;
7        $t \sim \text{Uniform}(\{1, \dots, T\})$  ;
8        $\epsilon \sim \mathcal{N}(0, \mathbf{I})$  ;
9       Take gradient descent step on  $\nabla_{\theta} \|\epsilon - D_{\theta}(\sqrt{\alpha_t} \mathbf{x}_0 + \sqrt{1 - \alpha_t} \epsilon, t)\|^2$  ;
10      until converged ;
11       $\epsilon^f = D_{\theta}(c, f, 0)$  ;
12      for  $\eta \in [0, 1]$  do
13         $\epsilon_{\eta}^f \leftarrow \eta \cdot \epsilon^f$  ;
14         $\phi_n \leftarrow M_{\psi}(m, \epsilon_{\eta}^f)$  ;
15      end
16      Obtain: The aligned brain regions  $m' = m \circ \phi_n$  ;
17    end
18    Calculate: ;
19    The feature of each brain region:  $M_{region_i} = m' \cdot \text{mask}(region_i)$  ;
20    The initial brain network:  $A = \sigma(M^T M)$  ;
21     $\mathcal{L}_{\text{BRAM}} = \mathcal{L}_{\text{diffusion}}(c, x_t, t) + \lambda \mathcal{L}_{\text{regist}}(m, f)$  ;
22    for  $x \in X$  do
23       $A_{\text{diff}} = S^{\text{heat}} = \exp(tAD^{-1} - t)$  ;
24       $(h_g^{\alpha}, h_g^{\beta}) = \text{MLP}(\text{GCN}((A, A_{\text{diff}})))$  ;
25       $(h_i^{\alpha}, h_i^{\beta}) = \text{MLP}(\text{POOL}(\text{GCN}((A, A_{\text{diff}}))))$  ;
26    end
27     $A_p = H_i^T H_i$  ;
28     $\mathcal{L}_{\text{cont}} = \max_{\theta, \omega, \phi, \psi} \frac{1}{|\mathcal{G}|} \sum_{g \in G} \sum_{i=1}^{|g|} [\text{MI}(h_i^{\alpha}, h_g^{\beta}) + \text{MI}(h_i^{\beta}, h_g^{\alpha})]$  ;
29     $\mathcal{L}_{\text{ref}}(\mathcal{G}) = -\frac{1}{90 \times 90} \sum_{i,j} (A_p[i, j] - A_{\text{ref}}[i, j])^2$  ;
30     $\mathcal{L}_{\text{cls}}(p) = -\sum_{i=0}^{C-1} y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$  ;
31    Calculate: total loss function  $\mathcal{L}_{\text{total}} = \sum_{t=1}^N \mathcal{L}_{\text{cont}} + \mathcal{L}_{\text{cls}}(p) + \alpha \mathcal{L}_{\text{ref}}(\mathcal{G})$  ;
32    Back propagation and update the parameters ;
33  end

```

Result: The generated brain network A_p .

TABLE I
SUBJECTS INFORMATION IN THIS STUDY

Group	NC(87)	EMCI(135)	LMCI(63)	AD(64)
Male/Female	46M/41F	83M/72F	26M/37F	35M/29F
Age(mean±std)	74.3±5.5	74.9±5.8	75.8±6.1	75.6±5.4

TABLE II
SAMPLE STATISTICAL INFORMATION ON THE ABIDE DATASET

Group	NC(80)	ASD(80)
Male/Female	65M/15F	70M/10F
Age(mean±std)	13.2±3..	12.9±2.7

TABLE III
THE QUALITY OF ALIGNMENT AND REGISTRATION RESULTS UNDER DIFFERENT HYPERPARAMETERS

Method	hyperparameters	learning rate	PSNR	Dice
ANTs	-	-	-	0.694
BRAM	$\lambda = 2.0, \mu = 0.3$	1×10^{-4}	27.3	0.671
BRAM	$\lambda = 2.0, \mu = 0.5$	1×10^{-4}	27.9	0.735
BRAM	$\lambda = 2.0, \mu = 0.8$	1×10^{-4}	28.7	0.798
BRAM	$\lambda = 2.0, \mu = 1.0$	1×10^{-4}	28.3	0.786

$\lambda=2.0$ and $\mu=0.8$ can reach the optimal 0.798, far exceeding the effect of the ANTs software.

The results show that there is no significant difference between the different methods in the brain area localization phase. However, using the diffusion mechanism yield the best results, with a Dice coefficient very close to the optimal value of 0.802 in the original paper, with λ set to 2.0 and μ set to 0.8. To measure

TABLE IV
THE QUALITY OF BRAIN NETWORKS CONSTRUCTED UNDER DIFFERENT HYPERPARAMETERS

α	Epoch	ACC	AUC	Quality	Identifiability
0	36	0.939	0.945	bad	strong
0.1	42	0.953	0.961	bad	strong
0.2	49	0.942	0.958	bad	strong
0.3	53	0.964	0.974	bad	strong
0.4	88	0.966	0.972	bad	weak
0.5	69	0.963	0.971	good	weak
0.6	126	0.968	0.976	good	weak
0.7	147	0.971	0.974	good	strong
0.8	164	0.969	0.977	good	strong
0.9	231	0.965	0.968	superior	strong
1.0	183	0.957	0.962	superior	strong

TABLE V
RESULTS OF ABLATION STUDY ON EACH BLOCK PROPOSED

Method	AUC	ACC	SEN	SPE
w/o BRAM	0.9536	0.9138	0.9625	0.8353
w/o BNOM	0.8914	0.8839	0.9286	0.7867
BRAM+BNOM	0.9781	0.9327	0.9550	0.8625

the difference between the two networks, we choose from several loss functions, including cross-entropy, absolute loss, and square loss (including mean and sum modes). We find that choosing the square loss function yielded better results, with the sum mode being equivalent to punishing every edge, resulting in the slowest convergence rate. The mean mode could converge in 40 epochs, after which they were unable to learn the features of the brain network.

The hyperparameter α mainly affects the brain generation quality. The results are shown in Table IV. The two columns of Quality and Identifiability are judged by experienced experts in the field of brain science and cognitive neuroscience. When α is set to 0.9, the model converged after 231 epochs of training. At this point, the optimized brain networks for different classes have good discriminative power, and exhibit similar patterns to the reference brain network as shown in Fig. 3. Moreover, the brain networks constructed by the proposed DGCL have high consistency on the same group of samples.

We propose the BRAM for brain atlas based on the diffusion model and the BNOM based on graph contrastive learning. We conduct ablation experiments on these two modules, namely BRAM implemented using the DDPM model and BNOM implemented using the GCN model. When the BNOM is not used, two fully connected layers (MLP) are employed for multiclass classification. The results (Table V) show that the brain area localization module has a relatively small effect on performance, while the BNOM for brain network has a more significant impact on performance improvement. Using an unaligned brain area without location information for subsequent classification results in poorer performance compared to the results obtained using the PANDA software. The initial brain network obtained using the brain area localization module is better than the experience-based brain network obtained from the software, even without optimization. The ablation experiments demonstrate the effectiveness of these two modules.

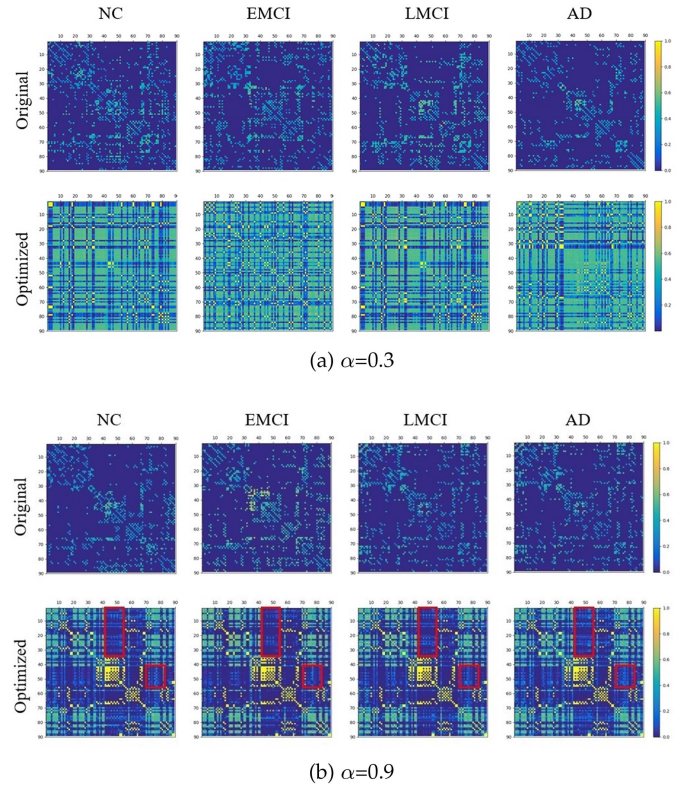


Fig. 3. Visualization of brain network matrix under different hyperparameters. The first row represents the original brain network matrix, and the second row represents the optimized brain network.

TABLE VI
EFFECTS OF LOSS FUNCTIONS ON THE PREDICTION PERFORMANCE

Method	AUC	ACC	SEN	SPE
w/o \mathcal{L}_{cont}	0.6929	0.7105	0.8000	0.6786
w/o \mathcal{L}_{cls}	0.7710	0.8185	0.8518	0.7273
proposed	0.8793	0.8661	0.8999	0.8276

2) *Loss Functions Analysis*: To fully evaluate the performance of the DGCL model, we conduct the ablation experiments of \mathcal{L}_{cont} and \mathcal{L}_{cls} to explore their respective effects on the final model performance. We removed \mathcal{L}_{cont} and \mathcal{L}_{cls} , respectively, and evaluated the model's performance in the brain network classification task. We compared the classification accuracy and other performance indicators of the models in the following three configurations with the same dataset and experimental settings: (1) Remove \mathcal{L}_{cont} : Keep only \mathcal{L}_{cls} and \mathcal{L}_{ref} . (2) Remove \mathcal{L}_{cls} : Keep only \mathcal{L}_{cont} and \mathcal{L}_{ref} . (3) Full loss function: keep \mathcal{L}_{cont} , \mathcal{L}_{cls} , and \mathcal{L}_{ref} . The results are shown in Table VI.

(1) *Effects of removing \mathcal{L}_{cont}* : When removing the \mathcal{L}_{cont} , the classification AUC of the model significantly decreased (from 0.8461 to 0.6929). This decrease indicates that the contrastive learning loss \mathcal{L}_{cont} plays a crucial role in enhancing the model's ability to distinguish features from different brain networks. \mathcal{L}_{cont} effectively improves the model's generalization ability and stability by maximizing mutual information between different samples.

(2) Effects of removing \mathcal{L}_{cls} . When removing \mathcal{L}_{cls} , the classification performance of the model also decreased (from 0.8793 to 0.7710), but the decrease was relatively small. This indicates that the multi-classification loss \mathcal{L}_{cls} is equally important in guiding the model to learn class-distincting features. \mathcal{L}_{cls} improves the classification performance of the model by directly optimizing the classification objectives.

(3) Effects of Full loss function. When using the complete loss function (including \mathcal{L}_{cont} , \mathcal{L}_{cls} , and \mathcal{L}_{ref}), the model performs best on all indicators. The AUC was 0.8461, the classification accuracy reached 0.8793, and the sensitivity and specificity were 0.8999 and 0.8276, respectively. This indicates that combining contrastive learning loss, multiclassification loss, and reconstruction loss can achieve optimal model performance. The combination of these three factors has achieved a good balance in feature extraction, category differentiation, and reconstruction quality, fully verifying the effectiveness and necessity of each loss term.

V. DISCUSSION

A. Comparisons With Baseline

The brain network constructed by the PANDA software plus the dual-layer GCN/four-layer MLP was used as a baseline for experimental comparison. A four-class performance comparison was designed, with each metric result representing the average value of the one versus rest binary classification corresponding indicators for the four classes. Compared with directly using a four-layer fully connected layer for classification, the proposed method performed significantly better in all four indicators, outperforming other methods.

For the MLP method, we first reshaped the brain network into a vector with 8100 neurons. Then, the vector passed through three layers with 1024, 512, and 256 hidden neurons, respectively. Each layer is followed by a ReLU activation layer. The dropout rate is 0.2. Finally, the last layer outputs a vector with 4 or 2 neurons to predict the label. For the GCN method, the graph is a brain network, and the node features are one-hot vectors with a size of 90×90 . Each row in the node feature is a one-hot vector. We designed two GCN layers and three linear layers to predict the label, where each GCN layer is followed by a ReLU activation function. The hidden neurons of the two GCN layers are 128 and 64, respectively. The three linear layers have 64, 32 and 4 or 2 neurons.

The ACC indicator increased by 14.47%, the AUC indicator increased by 11.66%, the sensitivity increased by 9.65%, and the specificity increased by 17.60%. Although the performance slightly improved by replacing the first two fully connected layers with GCN, it still did not reach the performance achieved by the initial brain network constructed based on the region localization method proposed in this study. The results suggest that introducing the region localization module is advantageous for improving the model's disease detection performance.

Figs. 4 and 5 compare the performance indicators of different methods on AD and ASD datasets, respectively. Meanwhile, we further studied our model's classification performance through t-SNE analysis. Fig. 6 compares the visualization results of the

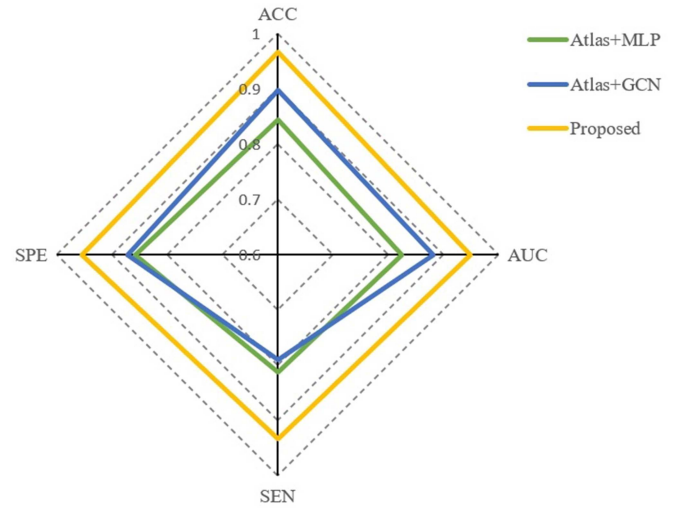


Fig. 4. Comparisons of AD classification performance on different metrics between proposed method with other two basic methods.

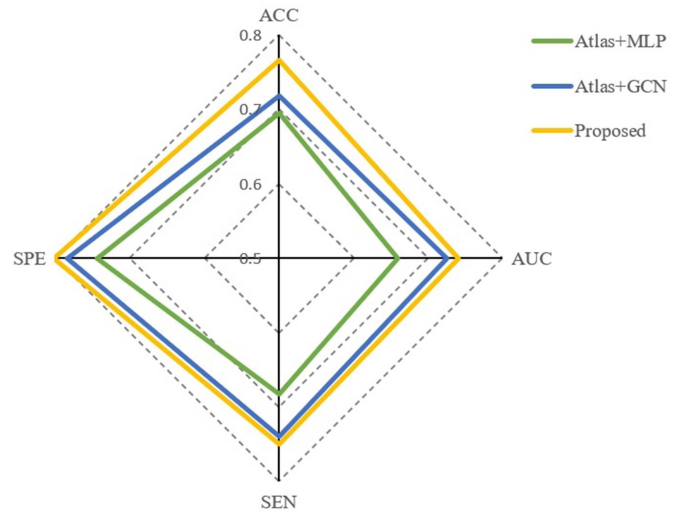


Fig. 5. Comparisons of ASD classification performance on different metrics between proposed method with other two basic methods. The classification performance of the proposed model on two datasets is comprehensively better than that of the comparison method in all indicators.

brain network embedding features of different methods, showing the two-dimensional projection of the learned brain network features of three methods.

Compared with directly adopting the brain network constructed by the PANDA software, our model can generate more easily segmented planes, indicating that the features obtained by our method are more discriminative than those of PANDA. The visualization results also intuitively explain why our model performs better than other models in disease classification tasks as show in Fig. 6.

To evaluate the quality of generating brain networks, multiple comprehensive indicators are used to compare the differences between the proposed method and the software. Among them, ICC is the most commonly used measure in assessing the reliability [47]. The experimental platform was completed on Dell

TABLE VII
RESULTS OF 5-FOLD CROSS VALIDATION ON EACH CATEGORY OF AD, AND VALUES ARE REPORTED AS MEAN \pm STANDARD DEVIATION

category	Validation Set				Test Set			
	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE
NC	0.9583 \pm 0.0563	0.8957 \pm 0.0884	0.9850 \pm 0.0200	0.8641 \pm 0.1201	0.8641 \pm 0.0381	0.8489 \pm 0.0311	0.8600 \pm 0.0490	0.8451 \pm 0.0362
EMCI	0.9648 \pm 0.0516	0.9126 \pm 0.1004	0.9400 \pm 0.0344	0.9019 \pm 0.1331	0.8699 \pm 0.0213	0.8279 \pm 0.0535	0.8309 \pm 0.0961	0.8259 \pm 0.0839
LMCI	0.9670 \pm 0.0268	0.8722 \pm 0.0816	0.9886 \pm 0.0229	0.8377 \pm 0.1048	0.9308 \pm 0.0506	0.8912 \pm 0.0751	0.9111 \pm 0.0831	0.8855 \pm 0.0933
AD	0.9706 \pm 0.0208	0.9036 \pm 0.0663	0.9719 \pm 0.0354	0.8835 \pm 0.0911	0.8384 \pm 0.0931	0.8016 \pm 0.0625	0.8000 \pm 0.2475	0.8028 \pm 0.1381

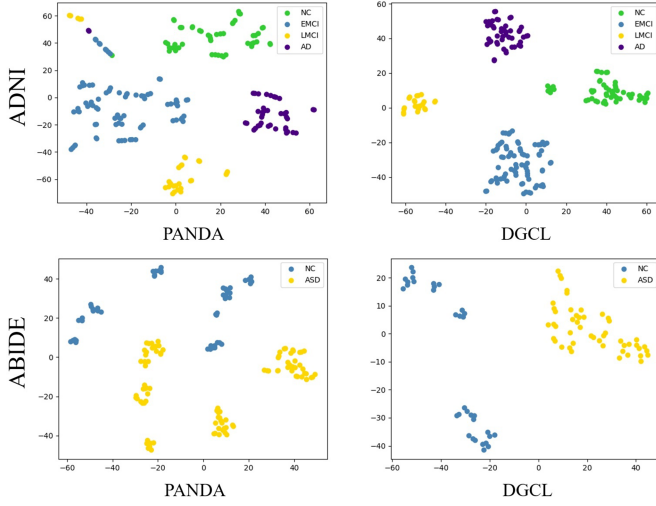


Fig. 6. Visualization of the features of different disease categories of the comparative and proposed methods under the t-SNE tool. The above two figures show the dimensionality reduction visualization results of brain networks constructed by DGCL and PANDA on the dataset ADNI. The below figures show the feature visualization on the dataset ABIDE under the same conditions. Compared to the brain network features generated by PANDA, the proposed DGCL can generate planes that are easier to segment.

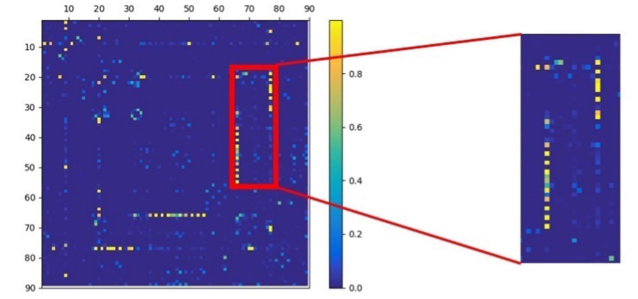
TABLE VIII
THE COMPREHENSIVE COMPARISON OF BRAIN NETWORK CONSTRUCTION USING PANDA AND DGCL

Datasets	Methods	Time(s)	Reliability	Parameter Dependency
ADNI	PANDA	$7.32 \times 10^2 \pm 27$	0.73	technician-dependent
	Ours	$2.23 \times 10^0 \pm 0.43$	0.79	highly-automated
ABIDE	PANDA	$7.25 \times 10^2 \pm 23$	0.64	technician-dependent
	Ours	$2.19 \times 10^0 \pm 0.36$	0.71	highly-automated

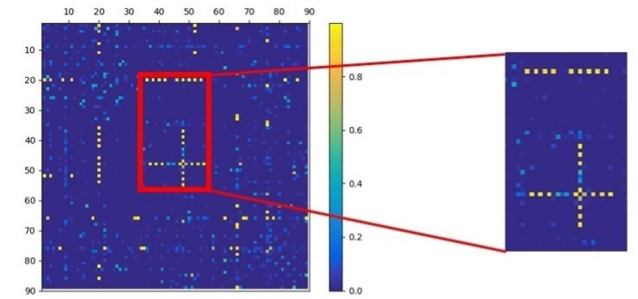
Precision T3630, with an Intel (R) Core (TM) i7-9700 CPU @ 3.00 GHz and 32 GB of memory. The comparative experimental results of brain network construction on two datasets are shown in Table VIII, indicating that the efficiency of our proposed deep learning based method far exceeds that of traditional software.

B. The Optimized Brain Networks

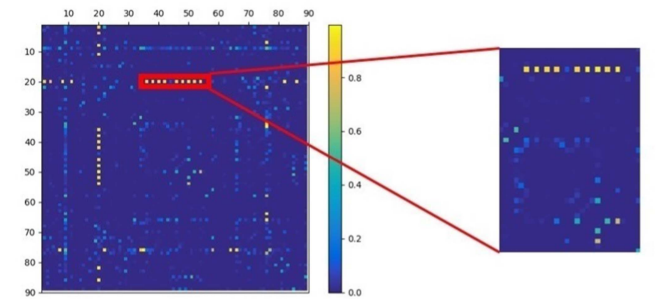
The brain network optimization module learns the differences in brain networks between different disease categories through graph comparison, enabling the classifier to better distinguish between different disease categories. To verify the expression ability of AD-related features in the optimized brain network, a classification experiment is designed to compare the predictive performance of the reference brain network and the optimized brain network at different stages of AD. This



(a) The increased connection from NC to EMCI compared to PANDA.



(b) The increased connection from NC to LMCI compared to PANDA.



(c) The increased connection from NC to AD compared to PANDA.

Fig. 7. The difference of brain network connection. Figures (a)–(c) show the changes in brain connectivity from NC to EMCI, LMCI, and AD, respectively.

experiment achieves four-class classification for different diseases, using five-fold cross-validation, and calculated the corresponding classification indicators using one-vs-rest for each class. The performance of this model on the test set is shown in Table VII. The best accuracy rates in the four binary classification tasks were 84.89%, 82.79%, 89.12%, and 80.16%, respectively.

To quantitatively analyze the changes in brain connectivity features in different cognitive disease stages, based on the output

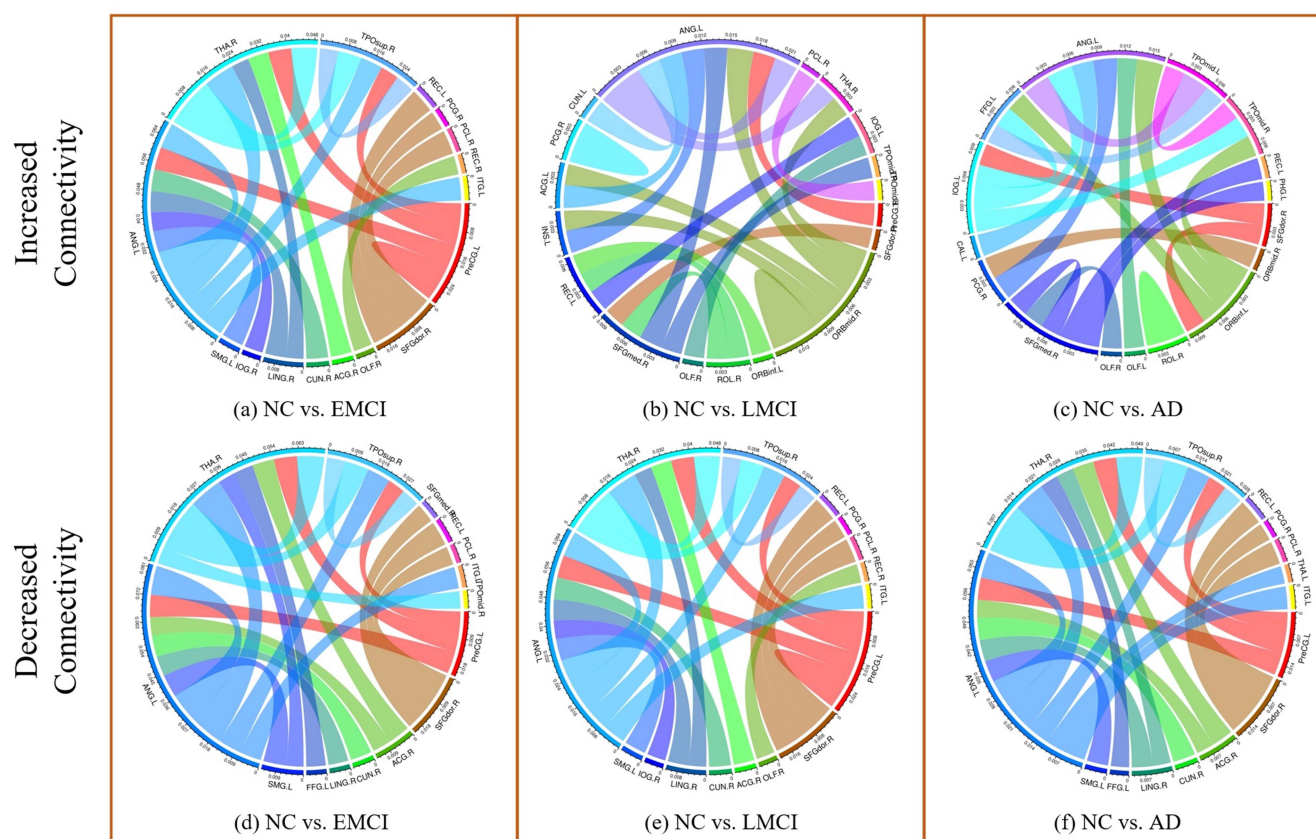


Fig. 8. (a)–(c) represents the brain connections with increased connectivity, and (d)–(f) represents the brain connections with decreased connectivity. There are significant differences in the brain regions with increased abnormal connections from EMCI to LMCI and AD stages. For the decreased connections, all three stages of the disease are concentrated in the connections between the frontal lobe and other brain regions.

structural brain connectivity matrix of the model, we study the changes in the structural connectivity of the brain network (brain network weighted adjacency matrix) in different populations. The visualization is shown in Fig. 7, which shows the brain connectivity differences between the NC versus EMCI, NC versus LMCI, and NC versus AD groups, respectively. For each binary classification task, we represented the connectivity matrix of each group, subtracted the connectivity matrix of the NC group from the connectivity matrix of the other groups, and obtained the changes in brain network connectivity.

In order to intuitively understand the differences in brain network connectivity between subjects in different disease stages and normal populations, we did not threshold the results and selected the top 2% of brain connections. From (a) to (c), it can be seen that with the deterioration of the disease, the increased connectivity initially rises in the early stage and then decreases to a low level when it worsens to AD. This result is consistent with neuroscience research, indicating that the brain may first produce compensatory mechanisms during the MCI to AD progression and weaken in the later stages of the disease.

Fig. 8 shows the visualization results of important brain connections in different stages of Alzheimer's disease. Fig. (a)–(c) represent the increased changes in brain connections, and (d)–(f) represent the decreased changes in brain connections between

different stages of the disease (NC, EMCI, LMCI, or AD). These abnormal brain connections can reflect the damage to important brain connections during the development of Alzheimer's disease. The brain regions involved in the abnormal increase of connections in EMCI and LMCI include the cingulate gyrus and occipital lobe, while in the AD stage, there is also an increase in connectivity between the frontal and temporal lobes, with the abnormal brain connections concentrated in the left cingulate gyrus and left occipital lobe. For the decreased connections, all three stages of the disease are concentrated in the connections between the frontal lobe and other brain regions. The regions with increased connectivity are consistent with most research, but the frontal lobe region is not reflected in this paper. According to [48], the development of the AD brain network leads to a transformation towards a more regular network topology and a loss of diffusion in local structural connections. Compared with healthy elderly people, MCI and AD patients have reduced structural connections in the posterior cingulate gyrus of the DMN, while the decrease in the DMN posterior cingulate cortex and anterior cingulate cortex can be revealed from the above figure.

For ASD, the ROIs of the dataset can be divided into eight functional subnetworks: cerebellum and subcortical structure (CB), visual network (VN), somatic motor network (SMN), dorsal attention network (DAN), ventral attention network (VAN),

TABLE IX
DETECTION OF DISEASE-RELATED BRAIN REGIONS/SUB-NETWORKS

Comparison	Software	Proposed	# Common
NC vs. EMCI	SFGdor.R, ORBmid.L, ORBmid.R, IFGoperc.L, IFGoperc.R, IFGtriang.R, ROL.L, SMA.R, OLF.L, SFGmed.R, ORBsupmed.R, REC.L, REC.R, INS.R, ACG.L, PHG.L, PHG.R, IOG.L, IOG.R, IPL.R, SMG.R, CAU.R	SFGdor.R, ORBmid.L, ORBmid.R, IFGoperc.R, IFGtriang.L, ORBinf.L, SMA.L, MA.R, OLF.L, ORBsupmed.L, ORBsupmed.R, REC.L, PCG.L, HIP.R, IOG.R, FFG.R, PoCG.L, SPG.R, SMG.R, CAU.R, THA.L, HES.L	10
NC vs. LMCI	SFGdor.R, ORBsup.R, FGoperc.L, ORBinf.L, ROL.L, SMA.L, OLF.L, SFGmed.R, ORBsupmed.R, REC.R, HIP.L, LING.L, SOG.L, MOG.R, FFG.R, PoCG.R, IPL.R, SMG.R, ANG.L, ANG.R, PCUN.L, STG.R	ORBsup.R, ROL.L, ROL.R, SMA.L, SMA.R, OLF.L, SFGmed.R, REC.R, CUN.L, LING.L, SOG.L, MOG.L, MOG.R, FFG.R, PoCG.L, PoCG.R, SPG.L, SMG.L, SMG.R, ANG.L, ANG.R, PCUN.L, STG.L, STG.R, ITG.L, ITG.R	13
NC vs. AD	SFGdor.R, ORBsup.R, ORBmid.L, IFGtriang.L, ROL.R, SMA.R, LING.L, SOG.L, MOG.R, PoCG.L, PoCG.R, SMG.L, SMG.R, ANG.R, PCL.L, PCL.R, CAU.R, PUT.L, HES.R, STG.L, ITG.L, ITG.R	SFGdor.R, ORBsup.R, ROL.R, LING.L, SOG.L, MOG.R, PoCG.L, PoCG.R, SMG.L, SMG.R, ANG.L, ANG.R, PCUN.L, PCL.L, PCL.R, CAU.L, CAU.R, PUT.L, PUT.R, HES.R, TPOsup.L, MTG.L, ITG.R	17
HC vs. ASD	SMN, DAN, VAN, DMN	DAN, VAN, FPN, DMN	3

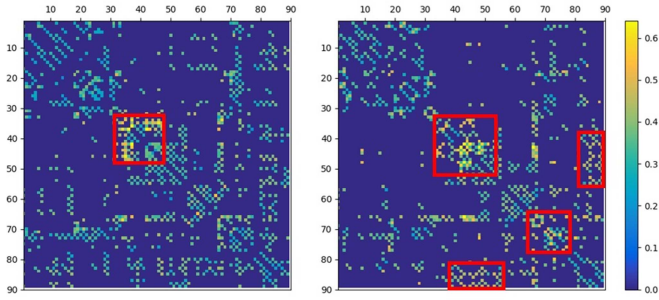


Fig. 9. Subnetwork patterns of autism spectrum disorder. The red box clearly indicates the difference between HC (left) and ASD (right). In these subnetworks, there have been significant changes in brain connectivity.

edge network (LN), frontoparietal network (FPN), and default mode network (DMN).

Fig. 9 shows the differences between the proposed method and the brain template (left), as well as the average matrix of the autism population with reference template (right). These matrices are obtained by averaging weight scores on correctly classified test data. Compared to the brain template, our model's learning score clearly highlights the important sub-networks for predicting ASD. Specifically, the attention score in the SMN region is high, indicating that connectivity plays a crucial role in ASD prediction in SMN.

According to relevant literature on autism, the reduction of FC in SMN reflects changes in sensory and motor processing. In addition, attention scores were higher in the DMN, DAN, and FPN regions (highlighted in the red box), indicating that functional connectivity between DMN and DAN, as well as between DMN and FPN, is also crucial for predicting ASD, consistent with previous studies reporting DMN abnormalities. Therefore, the proposed model can correctly identify functional networks related to ASD, and the results are consistent with the findings of neuroscience.

C. Detection of Disease-Related Brain Regions

The degree centrality of a brain region node is defined as the importance of the brain region, which is calculated using the

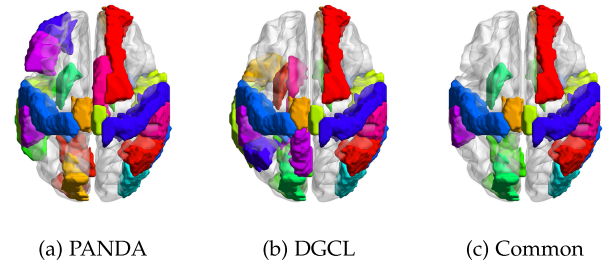


Fig. 10. Visualization of the Alzheimer's disease-related brain regions detected by (a) PANDA, (b) the proposed DGCL and (c) the common regions. Each brain region is represented by a different color, and the same color in the image represents the jointly detected area. From (c), it can be seen that the proposed model has high consistency with the software.

following formula:

$$C_{deg}(v_i) = \frac{\sum_j A[i, j]}{\sum_j \mathbf{1}(A[i, j] > 0)} \quad (12)$$

Based on the importance of brain regions, we visualized the brain regions associated with AD generated by PANDA and proposed method, as shown in Fig. 10. In our four-class classification scenario, 17 ROIs overlapped with the brain regions identified and they show strong consistency in the common regions (Table IX). These important ROIs include the frontal lobe, the precuneus, the paracentral lobule, the superior frontal gyrus, and the caudate nucleus.

Most of these regions are consistent with previous research, demonstrating that our model has captured important brain regions associated with AD. According to research in neuroscience and cognitive brain diseases, the widely recognized and highly correlated ROIs in AD/MCI diseases include the hippocampus (HIP. R), inferior frontal gyrus (IFGtriang. L), occipital gyrus (MOG. R), rectus gyrus (REC. L and REC. R), etc [49], [50], [51], [52]. These ROIs are mainly distributed in the frontal, temporal, and occipital lobes. The frontal lobe is primarily associated with higher-order mental activities, including motor control, language expression, self-awareness, and emotional expression. The temporal lobe is responsible for storing visual and

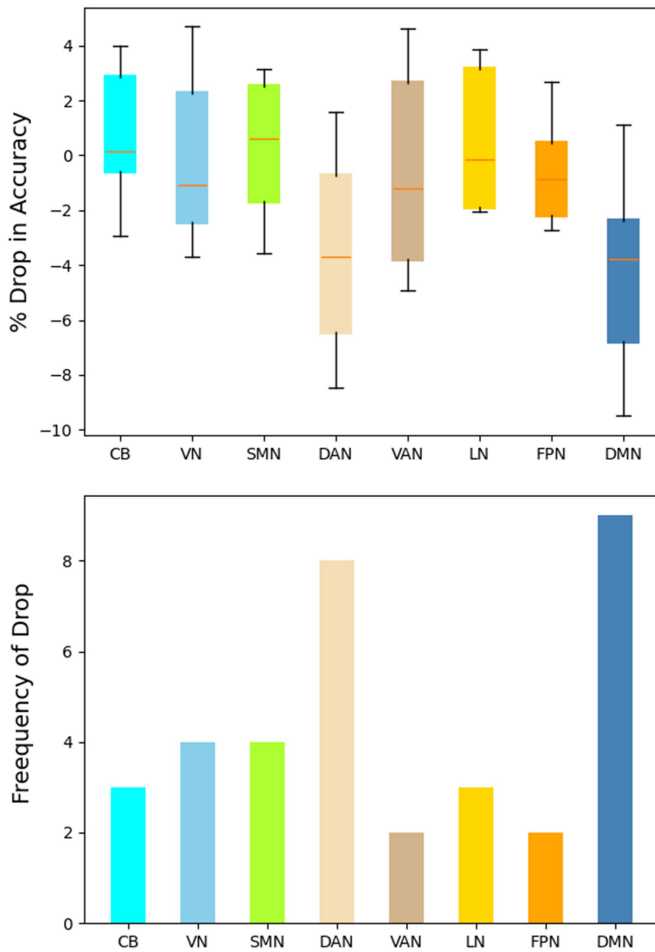


Fig. 11. The percentage drop in accuracy (including the median and range) across all sites when each of the sub-networks is occluded in the ablation analysis (top). The frequency of drop in accuracy, i.e. the number of sites where a drop in accuracy is observed, for occlusion of each of the sub-networks in ablation analysis (bottom).

auditory information in memory, and in AD patients, abnormal levels of tau protein have been found in the inferior temporal gyrus [53], [54]. Overall, the important ROIs and abnormal connections inferred by our model can reflect the major features associated with AD.

To further validate the importance of brain regions on the autism dataset, we used ablation analysis, specifically, to mask the brain regions of the 8 sub-networks of the imaging data to evaluate the degree of performance degradation. The upper part of Fig. 11 shows the decrease in accuracy when any of the given 8 sub-networks is occluded from the analysis. It can be seen that there are some differences between different brain regions. Therefore, we evaluated the frequency of accuracy decrease, i.e. the number of locations where accuracy decrease is observed in ablation analysis. When each sub network is occluded in the ablation analysis, the percentage of accuracy decrease (median and range) for all stations is also shown in Fig. 11. In terms of the absolute percentage decrease in frequency and accuracy, both the Dorsal Attention Network (decrease of 8 sites) and the Default Mode Network (decrease of 9 sites) performed outstandingly.

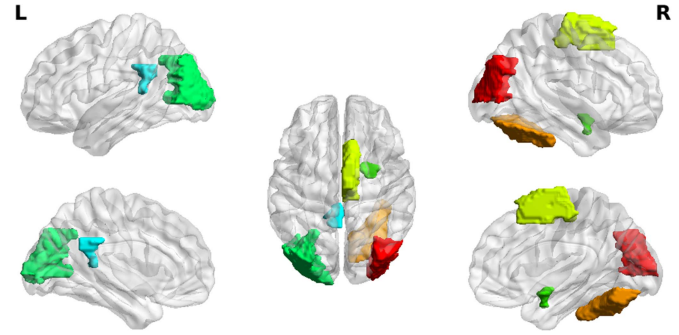


Fig. 12. Visualization of the top 5 highest importance of brain region vectors in ASD groups.

Distinguish the differences between ASD and HC based on the attention vectors generated by the matrix of the brain network. We study the first row of the learned brain network matrix, which is the correlation coefficient with the corresponding ROI. Use the average cue vector to generate ROI standardized attention scores on correctly classified test data, and we visualized the brain region where TOP5 is located in Fig. 12. It can be seen that there are significant differences in important brain regions between ASD and HC subjects. In addition, most of the differences in generating vector embeddings occur in DMN and SMN, which is consistent with previous neuroscience literature.

VI. CONCLUSION

In this paper, we proposed a new brain network reconstruction model based on the diffusion mechanism for aligning brain images and the graph contrastive learning for brain network generation. By using a diffusion-guided BRAM, the model can accurately locate the spatial position of brain regions partitioned by the template and extract important features with spatial information semantics on each brain region. In addition, we use the idea of graph contrastive learning to eliminate individual differences in connectivity unrelated to the disease among populations, and can well retain and strengthen common connections among similar samples, highlighting abnormal brain connections between different categories and improving the efficiency and stability of brain network construction. Although the disease studied in this paper is targeted at AD or ASD, the proposed model can easily be extended to other neurodegenerative diseases. Unlike traditional template-based methods, the proposed model is expected to provide a new deep learning analysis framework for abnormal brain connection detection. Overall, our approach outperforms traditional methods in terms of running efficiency, and accuracy. However, our work has two main limitations: one is that the proposed model still cannot explain the pathogenesis of the detected abnormal brain connections; the other is that the dataset used in the current study is relatively small. In the future, we plan to use other deep learning methods to analyze fMRI data and further study changes in brain function through functional networks. Meanwhile, we may validate the effectiveness of the proposed model on larger medical image datasets such as U.K. Biobank.

REFERENCES

- [1] L. Minati, T. Edginton, M. Grazia Bruzzone, and G. Giaccone, "Reviews: Current concepts in alzheimer's disease: A multidisciplinary review," *Amer. J. Alzheimer's Dis. Other Dementias*, vol. 24, no. 2, pp. 95–121, 2009, doi: [10.1177/1533317508328602](https://doi.org/10.1177/1533317508328602).
- [2] G. Mahesh and S. Tasneem, "Biomarker controversies and diagnostic difficulties in Alzheimers disease," *Amer. J. Phytomedicine Clin. Therapeutics*, vol. 2, no. 4, pp. 463–8, 2014, [Online]. Available: <https://www.academia.edu/download/34031088/PA-400318--1.pdf>
- [3] S. Wang et al., "An ensemble-based densely-connected deep learning system for assessment of skeletal maturity," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 1, pp. 426–437, Jan. 2022, doi: [10.1109/TSMC.2020.2997852](https://doi.org/10.1109/TSMC.2020.2997852).
- [4] R. Balachandar et al., "Resting-state functional connectivity changes associated with visuospatial cognitive deficits in patients with mild Alzheimer disease," *Dement. Geriatr. Cogn. Disord.*, vol. 43, no. 5/6, pp. 229–236, 2017, doi: [10.1159/000457118](https://doi.org/10.1159/000457118).
- [5] S. Wang et al., "Skeletal maturity recognition using a fully automated system with convolutional neural networks," *IEEE Access*, vol. 6, pp. 29979–29993, 2018, doi: [10.1109/ACCESS.2018.2843392](https://doi.org/10.1109/ACCESS.2018.2843392).
- [6] S. Wang, Y. Hu, Y. Shen, and H. Li, "Classification of diffusion tensor metrics for the diagnosis of a myelopathic cord using machine learning," *Int. J. Neural Syst.*, vol. 28, no. 02, 2018, Art. no. 1750036, doi: [10.1142/S0129065717500368](https://doi.org/10.1142/S0129065717500368).
- [7] B. Lei et al., "Predicting clinical scores for Alzheimer's disease based on joint and deep learning," *Expert Syst. Appl.*, vol. 187, 2022, Art. no. 115966, doi: [10.1016/j.eswa.2021.115966](https://doi.org/10.1016/j.eswa.2021.115966).
- [8] Y. Li et al., "Brain anatomical network and intelligence," *PLoS Comput. Biol.*, vol. 5, no. 5, 2009, Art. no. e1000395, doi: [10.1371/journal.pcbi.1000395](https://doi.org/10.1371/journal.pcbi.1000395).
- [9] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, 2009, doi: [10.1038/nrn2575](https://doi.org/10.1038/nrn2575).
- [10] Z. Yao, B. Hu, Y. Xie, P. Moore, and J. Zheng, "A review of structural and functional brain networks: Small world and atlas," *Brain Inform.*, vol. 2, pp. 45–52, 2015, doi: [10.1007/s40708-015-0009-z](https://doi.org/10.1007/s40708-015-0009-z).
- [11] A. Bessadok, M. A. Mahjoub, and I. Rekik, "Graph neural networks in network neuroscience," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 5, pp. 5833–5848, May 2023, doi: [10.1109/TPAMI.2022.3209686](https://doi.org/10.1109/TPAMI.2022.3209686).
- [12] K. A. Garrison, D. Scheinost, E. S. Finn, X. Shen, and R. T. Constable, "The (in) stability of functional brain network measures across thresholds," *Neuroimage*, vol. 118, pp. 651–661, 2015, doi: [10.1016/j.neuroimage.2015.05.046](https://doi.org/10.1016/j.neuroimage.2015.05.046).
- [13] S. G. Mueller et al., "Ways toward an early diagnosis in Alzheimer's disease: The alzheimers disease neuroimaging initiative (ADNI)," *Alzheimer's Dis. Dement.*, vol. 1, pp. 55–66, 2005, doi: [10.1016/j.jalz.2005.06.003](https://doi.org/10.1016/j.jalz.2005.06.003).
- [14] D. Martino et al., "The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism," *Mol. Psychiatry*, vol. 19, no. 6, pp. 659–667, 2014, doi: [10.1038/mp.2013.78](https://doi.org/10.1038/mp.2013.78).
- [15] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010, doi: [10.1016/j.neuroimage.2009.10.003](https://doi.org/10.1016/j.neuroimage.2009.10.003).
- [16] Y. He, Z. J. Chen, and A. C. Evans, "Small-world anatomical networks in the human brain revealed by cortical thickness from MRI," *Cereb. Cortex*, vol. 17, no. 10, pp. 2407–2419, 2007, doi: [10.1093/cercor/bhl149](https://doi.org/10.1093/cercor/bhl149).
- [17] J. Ma, X. Zhu, D. Yang, J. Chen, and G. Wu, "Attention-guided deep graph neural network for longitudinal Alzheimer's disease analysis," in *Proc. 23rd Int. Conf. Med. Image Comput. Comput. Assist. Interv.*, Lima, Peru, Springer, 2020, pp. 387–396, doi: [10.1007/978-3-030-59728-3_38](https://doi.org/10.1007/978-3-030-59728-3_38).
- [18] J. Gan, Z. Peng, X. Zhu, R. Hu, J. Ma, and G. Wu, "Brain functional connectivity analysis based on multi-graph fusion," *Med. Image Anal.*, vol. 71, 2021, Art. no. 102057, doi: [10.1016/j.media.2021.102057](https://doi.org/10.1016/j.media.2021.102057).
- [19] Y. Zhu, J. Ma, C. Yuan, and X. Zhu, "Interpretable learning based dynamic graph convolutional networks for Alzheimer's disease analysis," *Inf. Fusion*, vol. 77, pp. 53–61, 2022, doi: [10.1016/j.inffus.2021.07.013](https://doi.org/10.1016/j.inffus.2021.07.013).
- [20] H. Cui et al., "BrainGB: A benchmark for brain network analysis with graph neural networks," *IEEE Trans. Med. Imag.*, vol. 42, no. 2, pp. 493–506, Feb. 2023, doi: [10.1109/TMI.2022.3218745](https://doi.org/10.1109/TMI.2022.3218745).
- [21] P. Velićković, W. Fedus, W. L. Hamilton, P. Liò, Y. Bengio, and R. D. Hjelm, "Deep graph infomax," 2018, *arXiv: 1809.10341*. [Online]. Available: <https://arxiv.org/abs/1809.10341>
- [22] Y. You, T. Chen, Y. Sui, T. Chen, Z. Wang, and Y. Shen, "Graph contrastive learning with augmentations," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 5812–5823, doi: [10.2139/ssrn.4331194](https://doi.org/10.2139/ssrn.4331194).
- [23] X. Shen, D. Sun, S. Pan, X. Zhou, and L. T. Yang, "Neighbor contrastive learning on learnable graph augmentation," in *Proc. AAAI Conf. Artif. Intell.*, 2023, pp. 9782–9791, doi: [10.1609/aaai.v37i8.26168](https://doi.org/10.1609/aaai.v37i8.26168).
- [24] Y. Cui et al., "Identifying brain networks at multiple time scales via deep recurrent neural network," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 6, pp. 2515–2525, Nov. 2019, doi: [10.1109/JBHI.2018.2882885](https://doi.org/10.1109/JBHI.2018.2882885).
- [25] H. Wang et al., "Recognizing brain states using deep sparse recurrent neural network," *IEEE Trans. Med. Imag.*, vol. 38, no. 4, pp. 1058–1068, Apr. 2019, doi: [10.1109/TMI.2018.2877576](https://doi.org/10.1109/TMI.2018.2877576).
- [26] C. Lian, M. Liu, J. Zhang, and D. Shen, "Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 880–893, Apr. 2020, doi: [10.1109/TPAMI.2018.2889096](https://doi.org/10.1109/TPAMI.2018.2889096).
- [27] S. Sarraf, D. D. DeSouza, J. Anderson, G. Tofighi, and A. D. N. Initiative, "DeepAD: Alzheimers disease classification via deep convolutional neural networks using MRI and fMRI," *BioRxiv*, 2016, Art. no. 070441, doi: [10.1101/070441](https://doi.org/10.1101/070441).
- [28] C. Feng et al., "Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBI-LSTM," *IEEE Access*, vol. 7, pp. 63605–63618, 2019, doi: [10.1109/ACCESS.2019.2913847](https://doi.org/10.1109/ACCESS.2019.2913847).
- [29] W. Kang et al., "Multi-model and multi-slice ensemble learning architecture based on 2D convolutional neural networks for alzheimer's disease diagnosis," *Comput. Biol. Med.*, vol. 136, 2021, Art. no. 104678, doi: [10.1016/j.compbiomed.2021.104678](https://doi.org/10.1016/j.compbiomed.2021.104678).
- [30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
- [31] A. Loddo, S. Butta, and C. Di Ruberto, "Deep learning based pipelines for Alzheimer's disease diagnosis: A comparative study and a novel deep-ensemble method," *Comput. Biol. Med.*, vol. 141, 2022, Art. no. 105032, doi: [10.1016/j.compbiomed.2021.105032](https://doi.org/10.1016/j.compbiomed.2021.105032).
- [32] E. T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng, and M. Joliet, "Automated anatomical labelling atlas 3," *Neuroimage*, vol. 206, 2020, Art. no. 116189, doi: [10.1016/j.neuroimage.2019.116189](https://doi.org/10.1016/j.neuroimage.2019.116189).
- [33] K. Hassani and A. H. Khasahmadi, "Contrastive multi-view representation learning on graphs," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2020, pp. 4116–4126. [Online]. Available: <http://proceedings.mlr.press/v119/hassani20a/hassani20a.pdf>
- [34] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, 2019, doi: [10.1109/TMI.2019.2897538](https://doi.org/10.1109/TMI.2019.2897538).
- [35] M. Jaderberg et al., "Spatial transformer networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 2017–2025.
- [36] J. Ho, A. Jain, and P. Abbeel, "Denosing diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 6840–6851.
- [37] M. Meng, L. Bi, M. J. Fulham, D. D. Feng, and J. Kim, "Enhancing medical image registration via appearance adjustment networks," *NeuroImage*, vol. 259, 2021, Art. no. 119444, doi: [10.1016/j.neuroimage.2022.119444](https://doi.org/10.1016/j.neuroimage.2022.119444).
- [38] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*. [Online]. Available: <https://arxiv.org/pdf/1609.02907.pdf>
- [39] J. Klicpera, S. Weißenberger, and S. Günnemann, "Diffusion improves graph learning," in *Proc. Neural Inf. Process. Syst.*, 2019, pp. 13333–13345. [Online]. Available: <https://arxiv.org/pdf/1911.05485.pdf>
- [40] S. Wang, L. He, B. Cao, C.-T. Lu, P. S. Yu, and A. B. Ragin, "Structural deep brain network mining," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2017, pp. 475–484, doi: [10.1145/3097983.3097988](https://doi.org/10.1145/3097983.3097988).
- [41] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [42] C. Lian, M. Liu, L. Wang, and D. Shen, "End-to-end dementia status prediction from brain MRI using multi-task weakly-supervised attention network," in *Proc. 22nd Int. Conf. Med. Image Comput. Comput. Assist. Interv.*, Shenzhen, China, Springer, 2019, pp. 158–167, doi: [10.1007/978-3-030-32251-9_18](https://doi.org/10.1007/978-3-030-32251-9_18).
- [43] Z. Cui, S. Zhong, P. Xu, Y. He, and G. Gong, "PANDA: A pipeline toolbox for analyzing brain diffusion images," *Front. Hum. Neurosci.*, vol. 7, 2013, Art. no. 42, doi: [10.3389/fnhum.2013.00042](https://doi.org/10.3389/fnhum.2013.00042).

- [44] Q. Zuo, H. Wu, C. L. P. Chen, B. Lei, and S. Wang, "Prior-guided adversarial learning with hypergraph for predicting abnormal connections in Alzheimer's disease," *IEEE Trans. Cybern.*, vol. 54, no. 6, pp. 3652–3665, Jun. 2024, doi: [10.1109/TCYB.2023.3344641](https://doi.org/10.1109/TCYB.2023.3344641).
- [45] Q. Zuo, N. Zhong, Y. Pan, H. Wu, B. Lei, and S. Wang, "Brain structure-function fusing representation learning using adversarial Decomposed-VAE for analyzing MCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 4017–4028, 2023, doi: [10.1109/TNSRE.2023.3323432](https://doi.org/10.1109/TNSRE.2023.3323432).
- [46] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <https://arxiv.org/pdf/1412.6980.pdf>
- [47] Y. Hu, Q. Li, K. Qiao, X. Zhang, B. Chen, and Z. Yang, "PhiPipe: A multi-modal MRI data processing pipeline with test–retest reliability and predicative validity assessments," *Hum. Brain Mapping*, vol. 44, no. 5, pp. 2062–2084, 2023, doi: [10.1002/hbm.26194](https://doi.org/10.1002/hbm.26194).
- [48] L. Fan et al., "The human brainnetome atlas: A new brain atlas based on connectonal architecture," *Cereb. Cortex*, vol. 26, no. 8, pp. 3508–3526, 2016, doi: [10.1093/cercor/bhw157](https://doi.org/10.1093/cercor/bhw157).
- [49] X. Song et al., "Multi-center and multi-channel pooling GCN for early AD diagnosis based on dual-modality fused brain network," *IEEE Trans. Med. Imag.*, vol. 42, no. 2, pp. 354–367, Feb. 2023, doi: [10.1109/TMI.2022.3187141](https://doi.org/10.1109/TMI.2022.3187141).
- [50] C.-Y. Wee, S. Yang, P.-T. Yap, D. Shen, and A. D. N. Initiative, "Sparse temporally dynamic resting-state functional connectivity networks for early mci identification," *Brain Imag. Behav.*, vol. 10, pp. 342–356, 2016, doi: [10.1007/s11682-015-9408-2](https://doi.org/10.1007/s11682-015-9408-2).
- [51] D. Yao et al., "A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity," *IEEE Trans. Med. Imag.*, vol. 40, no. 4, pp. 1279–1289, Apr. 2021, doi: [10.1109/TMI.2021.3051604](https://doi.org/10.1109/TMI.2021.3051604).
- [52] Y. Zhang, H. Zhang, E. Adeli, X. Chen, M. Liu, and D. Shen, "Multiview feature learning with multiatlas-based functional connectivity networks for MCI diagnosis," *IEEE Trans. Cybern.*, vol. 52, no. 7, pp. 6822–6833, Jul. 2022, doi: [10.1109/TCYB.2020.3016953](https://doi.org/10.1109/TCYB.2020.3016953).
- [53] J. W. Prescott, A. Guidon, P. M. Doraiswamy, K. R. Choudhury, C. Liu, and J. R. Petrella, "The Alzheimer structural connectome: Changes in cortical network topology with increased amyloid plaque burden," *Radiology*, vol. 273, no. 1, pp. 175–184, 2014, doi: [10.1148/radiol.14132593](https://doi.org/10.1148/radiol.14132593).
- [54] Z. Yao, Y. Zhang, L. Lin, Y. Zhou, C. Xu, and T. Jiang, "Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease," *PLoS Comput. Biol.*, vol. 6, 2010, Art. no. e1001006, doi: [10.1371/journal.pcbi.1001006](https://doi.org/10.1371/journal.pcbi.1001006).



Yongcheng Zong received the BE degree in software engineering from Xidian University, Xi'an, China, in 2021. He is currently working toward the PhD degree with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Science, Shenzhen, China. His research interest includes medical image computing.



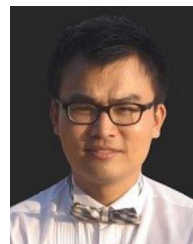
Qiankun Zuo received the PhD degree from the China University of Petroleum (Beijing), China, in 2018. He joined as a postdoctoral researcher with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. He is currently a lecturer with the School of Information Engineering, Hubei University of Economics, Wuhan. His research interests include deep learning, medical image analysis, and brain network computation.



Michael Kwok-Po Ng (Senior Member, IEEE) received the BSc and MPhil degrees from the University of Hong Kong, in 1990 and 1992, respectively, and the PhD degree from the Chinese University of Hong Kong, in 1995. He was a chair professor with the Research Division of Mathematical and Statistical Science, The University of Hong Kong from 2019 to 2023. He is currently a chair professor in mathematics and a chair professor in data science with Hong Kong Baptist University. His research interests include bioinformatics, image processing, scientific computing, and data mining. He is selected for the 2017 Class of Fellows of the Society for Industrial and Applied Mathematics. He obtained the Feng Kang Prize for his significant contributions in scientific computing. He serves on the editorial board members of several international journals.



Baiying Lei (Senior Member, IEEE) received the PhD degree from Nanyang Technological University (NTU), Singapore, in 2013. She is currently a distinguished professor with the School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen, China. She has coauthored more than 220 scientific articles. Her research interests include medical image analysis, machine learning, and pattern recognition. She serves as an associate editor for *IEEE Transactions on Medical Imaging*, *IEEE Transactions on Cybernetics* and *IEEE Transactions on Neural Networks and Learning Systems*, *Journal of Biomedical and Health Informatics*, and an editorial board member of *Medical Image Analysis*, *Neural Computing and Application*.



Shuqiang Wang (Senior Member, IEEE) received the PhD degree from the City University of Hong Kong, in 2012. He was a research scientist of Huawei Technologies Noahs Ark Lab. He held a postdoctoral fellowship with the University of Hong Kong from 2013 until 2014. He is currently an professor with the Shenzhen Institutes of Advanced Technology (SIAT), Chinese Academy of Science. He has published more than 150 papers on journals, such as *IEEE Transactions on Neural Networks and Learning Systems*/*IEEE Transactions on Image Processing*/*IEEE Transactions on Signal Processing*/*IEEE Transactions on Medical Imaging* etc. His current research interests include machine learning, brain image computing and optimization theory.