BGL-Net: A brain-inspired global-local information fusion network for Alzheimer's disease based on sMRI

Chen-Chen Fan, Hongjun Yang, Liang Peng, Xiao-Hu Zhou, Zhen-Liang Ni, Yan-Jie Zhou, Sheng Chen, Zeng-Guang Hou, *Fellow*, *IEEE*

***Abstract*—Alzheimer's Disease (AD) is an irreversible neurodegenerative disease, the most common form of dementia, affecting millions worldwide. Neuroimaging-based early AD diagnosis has become an effective approach, especially by using structural Magnetic Resonance Imaging (sMRI). The convolutional neural network (CNN) based method is challenging to learn dependencies between spatially distant positions in the various brain regions due to its local convolution operation. In contrast, the graph convolutional network (GCN) based work can connect the brain regions to capture global information but is not sensitive to the local information in a single brain region. Unlike a separate CNN or GCN-based method, we proposed a brain-inspired global-local information fusion network (BGL-Net) to diagnose AD. It essentially inherits the advantages of both CNN and GCN. The experiments on three public datasets demonstrate the effectiveness and robustness of our BGL-Net. Our method achieved the best performance on three popular public datasets compared with the existing CNN and GCN-based methods. In addition, our visualization results of the learned brain connection on AD and normal people agree with many current AD clinical research.**

***Index Terms*—alzheimer's disease, cognitive assessment, convolutional neural networks, graph neural networks, structural magnetic resonance imaging.**

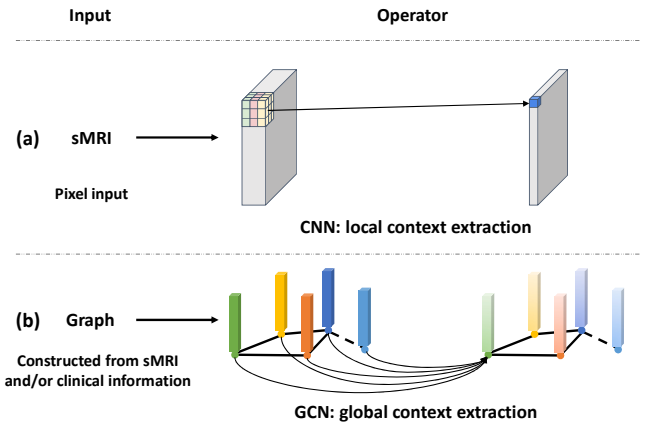


Fig. 1: Comparison of existing methods based on CNN or GCN. The CNN-Based method is challenging to learn dependencies between spatially distant positions in the various brain regions due to its local convolution operation. In contrast, the GCN-Based work can connect the brain regions to capture global information but is not sensitive to the local information in a single brain region.

I. Introduction

Alzheimer's disease (AD) is an irreversible neurodegenerative disease that results in a loss of mental function caused by the deterioration of brain tissue. It is the most common form of dementia, affecting millions of people around the world. While currently there is no cure for AD, early AD detection can contribute to effective interventions to delay the onset of the disease, indicating the importance of accurate AD prediction [1]. Structural Magnetic Resonance Imaging (sMRI), one of the leading techniques in brain abnormality detection, has been extensively used in AD detection [2]. Diagnosis based on sMRI requires a lot of diagnosis experience and is often affected by the clinician's competence [3]. Fortunately, computer-aided diagnosis technology has been applied to AD detection and offers a possible way to detect AD through sMRI quantitatively and effectively.

Motivated by the outstanding performance of CNN in a great number of computer vision tasks, CNN has become a promising method in AD detection. Though CNN-based AD detection methods have achieved relatively satisfactory performance, they have difficulty relating spatially distant regions (Figure 1 (a)). The progression of AD is associated with structural changes over the brain and dysfunction of the brain connectivity network between different brain areas. As feature extraction in CNN is based on receptive fields that operate on constrained local neighborhoods, it is difficult to effectively learn dependencies between spatially

This work was supported in part by the National Key R&D Program of China (Grant 2018YFC2001700), National Natural Science Foundation of China (Grants 61720106012, U1913601, 62073319, 62003343, 62222316), Beijing Natural Science Foundation (Grant L172050), Beijing SciTech Program (Grant Z211100007921021), ANSO Collaborative Research Project (Grant ANSO-CR-PP-2020-03), the Youth Innovation Promotion Association of CAS under Grant 2020140, and by the Strategic Priority Research Program of Chinese Academy of Science (Grant XDB32040000). (Corresponding author: Zeng-Guang Hou.)

C.-C. Fan, H. Yang, L. Peng, X.-H. Zhou, Z.-L. Ni, Y.-J. Zhou, S. Chen, Z.-G. Hou are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. C.-C. Fan, Z.-L. Ni, Y.-J. Zhou, S. Chen, Z.-G. Hou are also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China. Z.-G. Hou is also with the CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing 100190, China, and also with the CASIA-MUST Joint Laboratory of Intelligence Science and Technology, Institute of Systems Engineering, Macau University of Science and Technology, China. (email:{fanchenchen2018, hongjun.yang, liang.peng, xiaohu.zhou, nizhenliang2017, zhouyanjie2017, chensheng2016, zengguang.hou}@ia.ac.cn).

distant positions, making it hard to capture the connectivity disorder caused by AD. Previous approaches have tried to solve this problem by increasing model depth and kernel sizes or adopting new operations like non-local attention layers. These methods increase the computational complexity and only mitigate the problem instead of completely solve it [4], [5].

Existing studies have shown that there is a difference in the connection for areas of the brain between AD and normal people. Graph Convolutional Network (GCN) generalizes the convolution operation from grid data to graph representations [6] and makes it possible to model the relationship between different brain regions (Figure 1 (b)). The intrinsic nature of AD-induced structure and connectivity disorder motivates the exploration of brain network representation based on graph theory [7]. The brain can be regarded as a network graph, where nodes represent different brain regions and edges represent these regions' connections. To date, only a few studies explored AD diagnosis by GCN with sMRI data. Their graph representation construction relies on manually selected features like cortical thickness, which limits the ability of data representation. Also, those manual feature extraction takes plenty of time (about 10 hours/sheet to obtain various cortical thicknesses), making it challenging to meet the needs in real scenes.

In summary, the CNN-based AD method is challenging to learn dependencies between spatially distant positions in the various brain regions due to its local convolution operation. In contrast, the GCN-Based method can connect the brain regions to capture global information but is not sensitive to the local information in a single brain region. To solve these problems, we propose a brain-inspired global-local information fusion network (BGL-Net) that essentially inherits the advantages of both CNN and GCN. Specifically, it includes global and local information capture: (1) Inspired by the brain partition, region of interest (ROI) level node features are divided by the automated anatomical labeling (AAL) template and extracted by CNN. Then the graph representation is dynamically constructed through the proposed Graph Generation Module (GGM). Finally, global information features are obtained through graph convolution networks, which learn potential links between different brain regions. (2) Whole-brain MRI constructs local information features based on CNN due to the locality of the convolution operation. The global-local information features are fused to realize the classification diagnosis of multiple scale information.

The main contributions of this work can be concluded as follows:

- BGL-Net is proposed to effectively inherit the advantages of both CNN and GCN, achieving the best performance on three popular public datasets.
- An adaptive graph representation building module, GGM, is designed to automatically construct graph representation from sMRI, which avoids manually designing graph node features.
- A brain-inspired graph embedding method is proposed to capture global connection information between different brain regions. The results show that the captured brain area connections are consistent with the existing AD clinical research.

II. Related Work

A. CNN in AD detection

CNNs have achieved great success in various computer vision tasks and have been introduced to sMRI-based AD detection tasks [8]. Researchers [9] tried to convert sMRI to 2D slices, which can be further processed by 2D CNNs. Though this conversion expands the available dataset and facilitates the utility of existing outstanding architectures in a transfer learning fashion, it severely destroys the 3D spatial information. 3D CNNs are also widely used in AD detection [10]. Some studies [11]–[14] selected 3D patches from the whole sMRI and trained independent CNNs for each patch whose results were combined to perform the final classification. While spatial information is preserved, the computation cost and training time overhead are costly, and the selected patches may contain brain regions unrelated to AD progression. Also, some studies [15] focused on disease-related ROIs, like hippocampus, and used ROIs as input of CNN. As they only considered limited brain areas, informative information might have been lost. To better take advantage of the sMRI structure, several studies [16]–[18] performed subject-level methods where the entire sMRI data were used as input.

While CNN-based AD detection models have obtained great attention, they have trouble relating spatially distant information as its feature extraction depends on receptive fields that operate on constrained local neighborhoods. Some studies [19] try to use multi-task learning and multi-view weighted fusion to alleviate this problem, but cannot solve it fundamentally. Therefore, it's difficult for CNNs to effectively capture the potential relationship between different brain regions caused by AD.

B. GCN in AD detection

GCNs have recently gained increasingly more attention in AD diagnosis. GCN-based AD detection models can be divided into two categories according to the type of classification task they perform.

The first type converts the brain to graphs and carries out graph-level prediction tasks to perform AD detection. Wee et al. [20] employed Graph-CNN, a spectral graph convolutional neural network, to process cortical thickness and its underlying geometry information and perform AD detection. Sampathkumar [3] built a cortical thickness-based brain network graph and developed a GCN model, ADiag, to update graph representation and generate classification results. Current GCNs for graph-level AD detection tasks relies on manually selected features for graph representation construction instead of an adaptively data-driven way. Also, the extraction of relevant features (e.g., cortical thickness) takes plenty of time, making it challenging to meet the needs in real scenes.

The second type of model regards each sMRI sample as a node and performs node-level classification to classify unlabeled nodes. Kazi et al. [21], [22] used different clinical

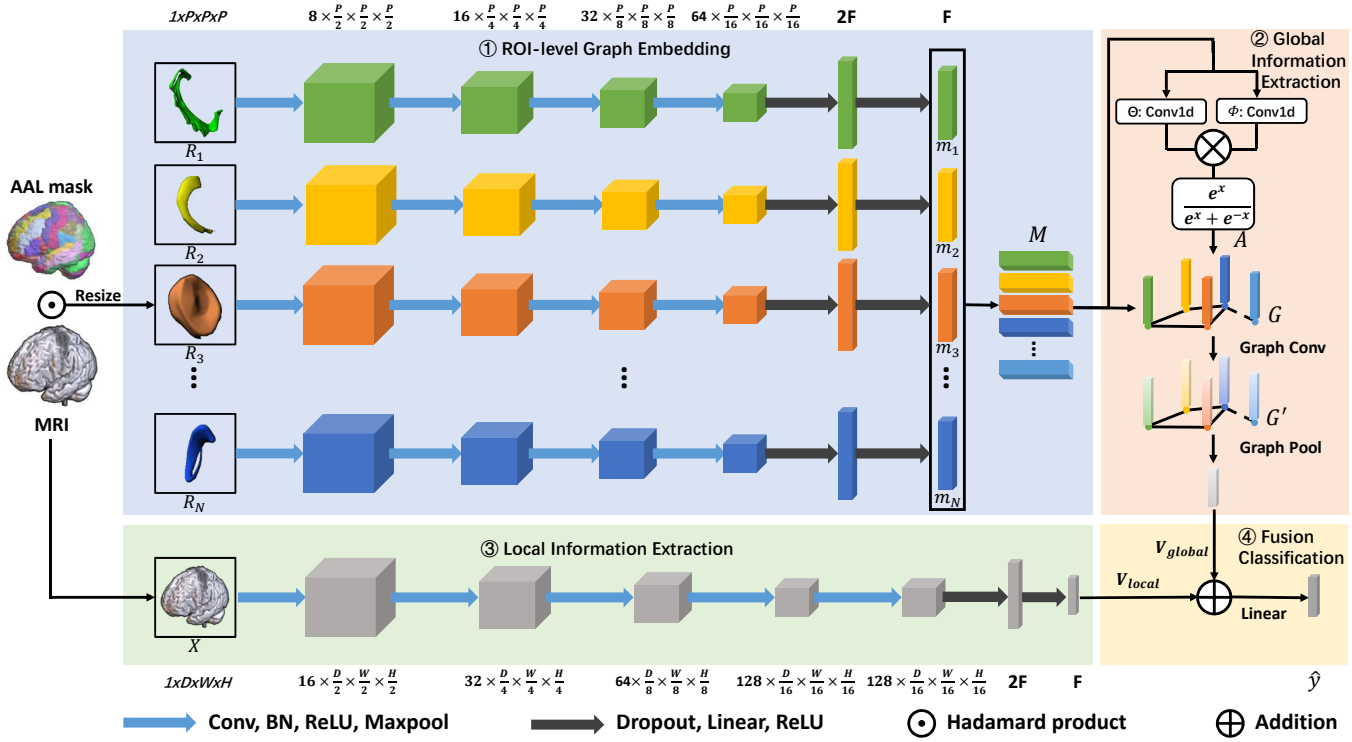


Fig. 2: The framework of the proposed BGL-Net method. It contains four blocks, 1. ROI-level Graph Embedding: AAL atlas is used to segment MRI into different brain regions, and extract the node features of each brain region through a series weight-shared graph embedding network; 2. Global Information Extraction: transform those ROI-level node features into a graph representation through the proposed adaptive graph generation module (GGM), and further capture the global information between different brain regions through graph convolution to obtain advanced classification features; 3. Local Information Extraction: CNN is used to extract features of the entire MRI and obtain advanced classification features focused on local information due to the locality of the convolution operation. 4. Fusion Classification: Fusion classification of the advanced classification features of the global and local information to obtain the diagnosis result.

features from multi-modal data to form multiple graphs and then fused each graph's classification results for the early diagnosis of AD. Though they achieved satisfactory performance, they suffered from the limitation of lack of flexibility, as they were not able to be used for independent testing. Song et al. [23] proposed an auto-metric graph neural network, AMGNN, which introduced a metric-based meta-learning strategy to improve its performance. As AMGNN was trained on several different tasks, the model complexity was high and the training process was time-consuming.

III. Methods

This section illustrates the proposed BGL-Net method. As shown in Figure 2, the classification of our proposed method relies on the combination of global and local information features. The GM tissue map input is parcellated into N brain areas according to AAL atlas [24], the commonly used brain partition scheme [25]. Those ROIs are resized to the same fixed size, $1 \times P \times P \times P$, and further processed by a series weight-shared graph embedding network. The obtained ROI-level node features are concatenated to construct the node feature matrix, denoted as M. The graph generation module (GGM) adaptively transforms M into a graph representation, then processed by graph convolution for global brain area information capture. Local information features are obtained by the Local Information Extraction block, which takes the original MRI as input. The local information vector V_{local} is further combined with the output (V_{global}) of the Global Information Extraction block to perform the final classification.

A. ROI-level Graph Embedding

As shown in Figure 2, the ROI-level Graph Embedding block receives a set of ROIs data extracted according to the AAL atlas. Then each ROI is processed by a series weight-shared ROI-level graph embedding network. We denoted the input images as $X \in \mathbf{R}^{1\times D\times W\times H}$, the above operation can be formulated as:

$$R_i = \text{Reshape}\left(X \odot Mask_i\right), \quad i = 1, 2, \cdots N \tag{1}$$

$$m_i = \text{Emb}\left(R_i\right), \quad i = 1, 2, \cdots N \tag{2}$$

where $\odot$ represents hadamard product, $Mask_i$ is the binarized mask of the AAL template in the i-th brain area, $R_i \in \mathbf{R}^{1\times P\times P\times P}$, $m_i \in \mathbf{R}^F$, N indicates the number of nodes. $Emb(\cdot)$ is a ROI-level graph embedding network contains four convolution layers and two fully connected layers. The kernel size for each convolution layer is set to three with stride

one and padding one. The convolution operation is followed by batch normalization (BN), rectified linear unit (ReLU) activation, and a maxpooling operation with stride two. As for the two fully connected layers, dropout is adopted, and ReLU activation function is used. As shown in (3), the obtained ROI-level node feature vector m_i is concatenated, denoted as $M \in \mathbf{R}^{N\times F}$, and further processed by graph convolution for global brain area information capture.

$$M = [m_1^T, m_2^T, \cdots m_N^T] \tag{3}$$

B. Global Information Extraction

The Global Information Extraction block transforms the input node feature matrix M to a graph representation and performs graph convolution and pooling operations to aggregate and update the global information feature.

1) Graph Generation Module: We consider the brain as a graph G where nodes feature are denoted by the obtained ROI data and edges represent the relationship between these regions. G is characterized by its node feature matrix $M \in \mathbf{R}^{N\times F}$ and adjacency matrix $A \in \mathbf{R}^{N\times N}$. A describes the latent connection between each node, can be obtained by the following formula:

$$A_{ij} = \frac{e^{[\Theta(M)\Phi(M)^T]_{ij}}}{e^{[\Theta(M)\Phi(M)^T]_{ij}} + e^{-[\Theta(M)\Phi(M)^T]_{ij}}} \tag{4}$$

where $i, j \in [1, 2, \cdots, N]$, $\Theta(\cdot)$ and $\Phi(\cdot)$ are convolution layers with kernel size one and stride one.

2) Graph Convolution: The graph convolution operation is adopted to aggregate and update node representation based on nodes connectivity. For input node feature matrix M and adjacency matrix A, the graph convolutional operator can be denoted as:

$$m_i' = \mathrm{ReLU}\left(\frac{1}{|\mathcal{N}(\mathrm{i})|}\sum_{j\in\mathcal{N}(\mathrm{i})} A_{ij} m_i\right) \tag{5}$$

where $i = 1, 2, \cdots N$, $\mathcal{N}(\mathrm{i})$ represents the set of neighbor nodes of node i.

3) Graph Pooling: Graph pooling operator is adopted to obtain global information features $V_{global} \in \mathbf{R}^F$ by global add pool layers. For input node feature matrix $M' = [{m_1'}^T, {m_2'}^T, \cdots, {m_N'}^T] \in \mathbf{R}^{N\times F}$, the global information features V_{global} can be obtained by the following formula:

$$V_{global} = \sum_{i=1}^{N} m_i', \quad m_i' \in \mathbf{R}^F \tag{6}$$

C. Local Information Extraction

Whole brain MRI constructs local information features based on CNN due to the locality of the convolution operation.

$$V_{local} = f(X) \in \mathbf{R}^F \tag{7}$$

where $f(\cdot)$ contains five convolution layers followed by two fully connected layers and adopts the similar layer setting as the graph embedding network.

D. Fusion Classification

The fusion classification block takes the updated global information feature V_{global} generated by the Global Information Extraction block and the local information feature V_{local} as input. These two feature vectors are added and processed by a fully connected layer to achieve the final classification result. This multi-scale feature fusion operation enables the analysis of alterations of the global and local information simultaneously.

$$\hat{\mathrm{y}} = \mathrm{FC}\left(V_{global} + V_{local}\right) \tag{8}$$

IV. Experiments

A. Dataset

Data used in the preparation of this article are obtained from three public datasets: Alzheimer's Disease Neuroimaging Initiative (ADNI [1]) database, Open Access Series of Imaging Studies (OASIS [2]) [28], and Neuroimaging in Frontotemporal Dementia (NIFD [3]).

1) ADNI: The ADNI (adni.loni.usc.edu) is a longitudinal multi-site observational study of controls (CN), mild cognitive impairment (MCI), and AD. A total number of 1253 subjects, including 330 CN subjects, 587 MCI subjects, and 336 AD patients, are used in our experiments. In MCI subjects, we divide it to 296 stable MCI subjects (sMCI: MCI patients who did not progress to AD in 36 months) and 248 progressive MCI (pMCI: MCI patients who progress to AD in 36 months). As data for each subject might be collected repeatedly, there are 5132 available images for the 1253 subjects. The testing set consists of the baseline data for 100 randomly chosen subjects in each diagnostic class (i.e., 100 CN subjects, 100 MCI subjects, 100 AD patients, 100 sMCI subjects, 100 pMCI subjects). The remaining data is used as the training set where we perform 5-fold cross-validation.

2) OASIS: The OASIS is a series of magnetic resonance imaging datasets that is publicly available for study and analysis. Data used in this work consists of 193 images from 100 AD patients and 93 CN participants in OASIS-1. We randomly select 20% of patients for each diagnostic class as the testing set, and the rest of the data is used as the training set and performed 5-fold cross-validation.

3) NIFD: NIFD is aimed to characterize longitudinal clinical and imaging changes in frontotemporal lobar degeneration (FTLD). We adopt 213 AD patients with 543 images and 114 CN participants with 312 images. Like OASIS, 20% of the

[1]As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

[2]Data were provided by OASIS: Cross-Sectional: Principal Investigators: D. Marcus, R, Buckner, J, Csernansky J. Morris; P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 RR021382

[3]NIFD is the nickname for the frontotemporal lobar degeneration neuroimaging initiative (FTLDNI, AG032306), which was funded by the NIA and NINDS to characterize longitudinal clinical and imaging changes in FTLD. The imaging and clinical methods are the same for NIFD and for the 4-Repeat Tauopathy Neuroimaging Initiative (4RTNI), which is also available for download from LONI. Controls for NIFD are the same controls as those collected for 4RTNI.

TABLE I: Experimental results of AD vs. CN classification with 5-fold CV on three large public datasets: ADNI, OASIA, and NIFD. BGL-Net achieves the best test BA on the three datasets.

Dataset	Model	Para. (M)	Valid			Test		
			BA↑	SEN↑	SPE↑	BA↑	SEN↑	SPE↑
ADNI	DeepCNN [16]	0.12	89.13 ± 1.96	85.60 ± 2.41	92.66 ± 2.65	87.60 ± 1.24	83.80 ± 3.19	91.40 ± 4.84
	ConvNet3D [17]	8.38	86.26 ± 3.55	81.36 ± 2.82	91.17 ± 4.54	86.30 ± 1.03	83.40 ± 3.98	89.20 ± 5.27
	VoxCNN [18]	2.14	87.91 ± 2.22	85.59 ± 6.37	90.22 ± 6.71	87.10 ± 1.36	86.80 ± 5.60	87.40 ± 7.47
	ResNet101 [26]	85.21	90.07 ± 2.39	89.41 ± 2.28	90.72 ± 6.78	86.30 ± 2.82	87.80 ± 4.53	84.80 ± 7.08
	DenseNet121 [27]	11.24	**91.09 ± 2.76**	**91.06 ± 5.45**	91.12 ± 6.07	88.20 ± 0.98	**88.00 ± 2.61**	88.40 ± 2.25
	BGL-Net(ours)	3.12	90.18 ± 3.79	84.32 ± 5.49	**96.05 ± 3.73**	**89.00 ± 0.89**	86.00 ± 2.61	**92.00 ± 1.67**
OASIS	DeepCNN [16]	0.12	72.50 ± 2.34	70.00 ± 13.35	75.00 ± 11.18	65.00 ± 5.13	60.00 ± 16.66	70.00 ± 8.32
	ConvNet3D [17]	8.38	56.87 ± 9.76	71.25 ± 37.42	42.50 ± 38.41	54.92 ± 6.53	64.29 ± 37.25	45.56 ± 39.50
	VoxCNN [18]	2.14	71.88 ± 1.98	61.25 ± 10.75	**82.50 ± 10.00**	65.63 ± 4.90	55.71 ± 22.77	**75.56 ± 16.33**
	ResNet101 [26]	85.21	**75.62 ± 4.15**	76.25 ± 8.29	75.00 ± 11.18	63.65 ± 4.14	62.86 ± 14.57	64.44 ± 13.43
	DenseNet121 [27]	11.24	69.37 ± 5.73	56.25 ± 24.37	56.25 ± 18.71	64.52 ± 4.55	55.71 ± 27.99	73.33 ± 24.44
	BGL-Net(ours)	3.12	75.00 ± 5.59	**80.00 ± 12.75**	70.00 ± 6.12	**67.22 ± 5.76**	**70.00 ± 12.29**	64.44 ± 11.44
NIFD	DeepCNN [16]	0.12	85.55 ± 3.99	79.30 ± 11.79	91.79 ± 7.46	79.64 ± 3.24	72.27 ± 5.26	87.00 ± 11.22
	ConvNet3D [17]	8.38	76.56 ± 8.01	**86.06 ± 9.43**	67.05 ± 22.52	72.05 ± 7.45	**79.09 ± 11.17**	65.00 ± 25.88
	VoxCNN [18]	2.14	83.98 ± 4.55	76.03 ± 8.97	91.92 ± 10.35	79.45 ± 2.73	70.91 ± 11.45	88.00 ± 7.48
	ResNet101 [26]	85.21	87.54 ± 2.55	79.95 ± 6.35	95.13 ± 3.99	77.09 ± 2.42	73.18 ± 6.49	81.00 ± 10.68
	DenseNet121 [27]	11.24	87.73 ± 2.90	77.13 ± 5.84	**98.33 ± 3.33**	81.64 ± 4.26	72.27 ± 2.23	91.00 ± 6.63
	BGL-Net(ours)	3.12	**88.37 ± 2.63**	79.95 ± 8.25	96.79 ± 3.93	**83.27 ± 2.35**	74.55 ± 6.65	**92.00 ± 5.10**

BA: Balanced accuracy; SEN: Sensitivity; SPE: Specificity; Para.: Parameter.

TABLE II: Experimental results of sMCI vs pMCI classification with 5-fold CV on ADNI datasets. BGL-Net achieves the best BA on both validation and test sets.

Model	Valid			Test		
	BA↑	SEN↑	SPE↑	BA↑	SEN↑	SPE↑
DeepCNN [16]	74.89 ± 4.94	76.52 ± 3.32	73.26 ± 9.89	69.90 ± 1.02	77.20 ± 3.87	62.60 ± 2.33
ConvNet3D [17]	61.07 ± 3.41	43.26 ± 13.67	**78.88 ± 11.71**	54.60 ± 4.31	36.00 ± 30.74	**73.20 ± 23.37**
VoxCNN [18]	60.60 ± 4.57	50.83 ± 21.02	70.37 ± 16.92	53.30 ± 3.11	73.00 ± 18.30	33.60 ± 20.81
ResNet101 [26]	73.98 ± 4.87	84.66 ± 8.27	63.31 ± 17.07	66.90 ± 4.44	80.20 ± 5.78	53.60 ± 13.66
DenseNet121 [27]	76.45 ± 4.27	**85.74 ± 10.05**	67.17 ± 11.61	70.60 ± 2.52	**82.40 ± 7.17**	58.80 ± 10.68
BGL-Net(ours)	**77.27 ± 1.70**	79.72 ± 9.13	74.83 ± 10.36	**71.90 ± 3.46**	77.40 ± 4.13	66.40 ± 9.67

BA: Balanced accuracy; SEN: Sensitivity; SPE: Specificity.

patients for each diagnostic class are randomly selected to compose the testing set, leaving the remaining data as training set where 5-fold cross-validation is performed.

4) Data Preprocessing: The original data have been curated and converted to the Brain Imaging Data Structure (BIDS) format [29] using Clinica [30]. Then the acquired data passes through the t1-volume pipeline of Clinica [31]. The Unified Segmentation procedure [32] is used to simultaneously perform tissue segmentation, bias correction, and spatial normalization. After that, t1-weighted volumetric images are segmented into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Only GM tissue maps are used in the following experiments as they are more related to the AD diagnosis. The size of GM tissue maps are $128 \times 128 \times 128$. During training, we perform 5-fold cross-validation. For each fold, the model with the highest balanced accuracy (BA) on the validation set is saved and further tested on the testing set.

B. Implementation Details

The models are implemented using Python 3.7.9, PyTorch on a workstation with Nvidia Tesla V100. For all experiments, we set the batch size to 8 and carry out an exhaustive grid search for learning rate and weight decay parameter combination. For models on ADNI, OASIS, and NIFD, their learning rate is set to 7e-4, 3e-4, and 1e-3, respectively, with weight decay of 5e-4, 1e-5, and 5e-3, respectively. All models' dropout rate is set to 0.5 and use AdamW as optimizer. The parameter of P is set to 32.

We adopt balanced accuracy (BA) to evaluate the classification performance. Also, sensitivity (SEN) and specificity (SPE) are used. The average and standard deviation of each metric are recorded to compare the performance of different models. The formula of BA, SEN, and SPE is as follows:

$$\text{BA} = \frac{1}{2}\left(\frac{TP}{P} + \frac{TN}{N}\right) \tag{9}$$

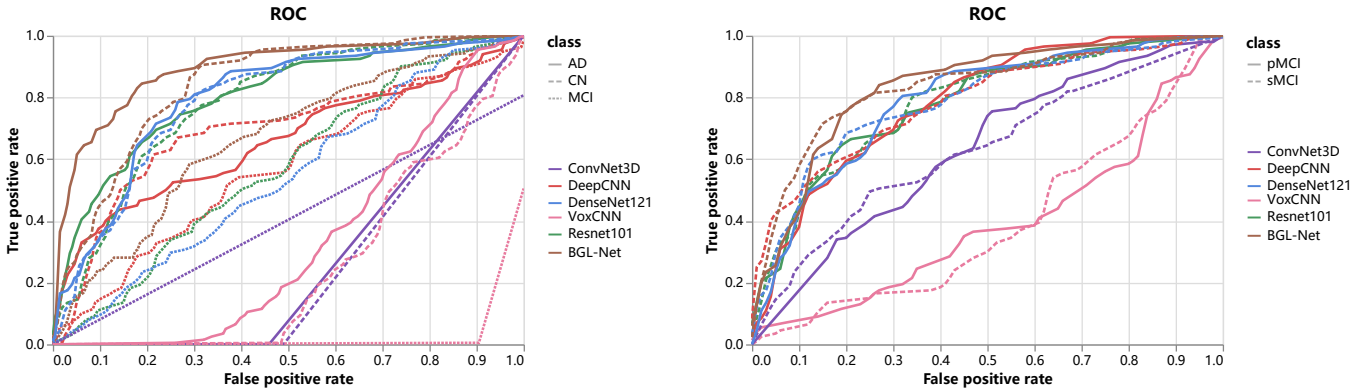
$$\text{SEN} = \frac{TP}{TP + FN} \tag{10}$$

$$\text{SPE} = \frac{TN}{FP + TN} \tag{11}$$

Where TP, TN, FP, and FN are the number of true positive, true negative, false positive, and false negative terms, respectively, in the confusion matrix, P and N is the number of all actual Positive and negative terms in confusion matrix respectively.

C. Results

For comparison, we adopt three existing state-of-the-art methods, DeepCNN [16], ConvNet3D [17], VoxCNN [18]



(a). The ROC curve of AD vs. MCI vs. CN task.

(b). The ROC curve of sMCI vs. pMCI task.

Fig. 3: The test ROC curve of AD vs. MCI vs. CN and sMCI vs. pMCI task in ADNI dataset.

TABLE III: Experimental results of AD vs CN vs MCI classification with 5-fold CV on ADNI datasets. Sensitivity and specificity were not available because the number of classes is greater than two. We use BA and F1 score to evaluate the model. BGL-Net achieves the best test BA.

Model	Valid		Test	
	BA↑	F1↑	BA↑	F1↑
DeepCNN [16]	46.81 ± 2.88	49.21 ± 3.17	47.54 ± 3.69	47.43 ± 3.83
ConvNet3D [17]	70.41 ± 8.96	72.90 ± 6.65	46.55 ± 8.50	46.96 ± 8.51
VoxCNN [18]	**72.60 ± 7.86**	**74.88 ± 6.29**	50.64 ± 6.71	51.08 ± 6.87
ResNet101 [26]	54.27 ± 4.98	52.79 ± 6.79	52.02 ± 5.92	52.03 ± 6.17
DenseNet121 [27]	56.60 ± 1.81	44.43 ± 4.15	53.17 ± 3.01	**53.78 ± 3.04**
BGL-Net(ours)	54.30 ± 4.76	56.78 ± 2.31	**53.74 ± 2.27**	53.58 ± 2.26

BA: Balanced accuracy; F1: F1 Score.

TABLE IV: Cross-site evaluation results. The AD vs. CN models trained on the ADNI dataset is used to test all data in NIFD and OASIS. BGL-Net maintains the best BA under the two untrained datasets.

Model	NIFD			OASIS		
	BA↑	SEN↑	SPE↑	BA↑	SEN↑	SPE↑
DeepCNN [16]	73.96 ± 2.76	59.03 ± 7.95	86.87 ± 16.97	72.70 ± 1.78	60.40 ± 3.93	**85.00 ± 2.95**
ConvNet3D [17]	61.42 ± 9.22	72.20 ± 17.44	50.64 ± 35.03	59.07 ± 5.38	68.80 ± 16.25	49.35 ± 24.19
VoxCNN [18]	49.83 ± 0.22	**78.96 ± 39.51**	20.71 ± 39.67	50.10 ± 0.48	**77.80 ± 39.08**	22.39 ± 39.08
ResNet101 [26]	75.63 ± 2.01	57.28 ± 3.55	93.98 ± 5.06	71.12 ± 2.40	66.60 ± 10.40	75.38 ± 12.52
DenseNet121 [27]	75.88 ± 1.00	56.12 ± 3.48	**95.64 ± 2.17**	72.79 ± 2.45	70.80 ± 6.11	74.78 ± 8.20
BGL-Net(ours)	**76.63 ± 1.02**	57.61 ± 2.85	**95.64 ± 2.39**	**73.08 ± 1.44**	71.60 ± 1.74	74.57 ± 2.71

BA: Balanced accuracy; SEN: Sensitivity; SPE: Specificity.

and convert classic deep learning models (esNet101 [26], DenseNet121 [27]) to 3d versions to perform classification on ADNI, OASIS, and NIFD dataset. Specifically, we construct AD vs. CN experiments on three datasets, AD vs. CN vs. MCI and sMCI vs. pMCI experiments on ADNI dataset.

1) Results on ADNI dataset: As shown in TABLE I, BGL-Net achieves the best test BA and SPE while maintaining the lowest variance. For baseline models, 3d version deep learning models outperform the three popular methods. While DenseNet121 obtains the highest validation BA, its BA on the testing set is lower than our proposed BGL-Net, indicating that there may be an overfitting problem. At the same time, BGL-Net's parameter size (3.12M) is much smaller than that of two deep learning models. Specifically, it is about 1/3 of DenseNet121 and about 1/27 of ResNet101.

As shown in Table II, in the sMCI vs. pMCI experiment, BGL-Net achieves the best BA on both validation and test sets. Combined with Figure. 3-(b), it can be found that the ROC curve of our method is at the top, which shows that our model outperforms other comparative methods.

In the AD vs. CN vs. MCI experiment, sensitivity and specificity were not available because the number of classes is greater than two. We use BA and F1 score to evaluate the model. As shown in Table III, BGL-Net achieves the best BA on the test set, and DenseNet121 achieves the best F1 score on the test set. ConvNet3D and VoxCNN have high BA on the validation set and poor performance on the test set. This shows that the generalization ability of these models is poor.

TABLE V: Comparison with GCN-based methods for AD and CN classification. As different data and preprocessing pipelines are adopted in these studies, it is difficult to reproduce them. Therefore, the results reported in their papers are used here for a naive direct comparison.

Method	Modality	Splits	Train and valid samples (AD/CN)	Test samples (AD/CN)	Valid				Test			
					ACC↑	BA↑	SEN↑	SPE↑	ACC↑	BA↑	SEN↑	SPE↑
Graph-CNN [20]	sMRI	10-fold	960/592	NA	85.80	NA	83.5	87.5	NA	NA	NA	NA
ADiag [3]	sMRI	train-valid	60/61	NA	83.44	NA	NA	NA	NA	NA	NA	NA
AMGNN [23]	sMRI, Clinical features*	5-fold	337/413	NA	88.33	NA	NA	NA	NA	NA	NA	NA
BGL-Net (ours)	sMRI	5-fold	236/230	100/100	**89.77**	**90.18**	**84.32**	**96.05**	**89.00**	**89.00**	**86.00**	**92.00**

* Clinical features include age, gender, year of education, APOe4 gene information, cognitive test score.

TABLE VI: Ablation study on ADNI datasets.

Feature	Valid			Test		
	BA↑	SEN↑	SPE↑	BA↑	SEN↑	SPE↑
Brain	87.83 ± 2.44	83.04 ± 5.73	92.62 ± 6.77	87.20 ± 1.57	84.00 ± 5.25	90.40 ± 5.08
ROI	85.21 ± 4.55	82.64 ± 8.30	87.77 ± 9.36	85.10 ± 1.02	85.00 ± 7.64	85.20± 6.46
Brain + ROI	**90.18 ± 3.79**	**84.32 ± 5.49**	**96.05 ± 3.73**	**89.00 ± 0.89**	**86.00 ± 2.61**	**92.00 ± 1.67**

Combined with Figure. 3-(a), the ROC curves of different classes of BGL-Net are all at the top, which demonstrates the excellent performance of our model.

2) Results on OASIS dataset: According to the results on the OASIS dataset in TABLE I, BGL-Net achieves the top BA which is 1.59% higher than the second-place VoxCNN obtained. Meanwhile, BGL-Net has the highest SPE on both validation and testing sets. Compared with VoxCNN, BGL-Net's SPE is 18.75% and 14.29% higher on validation and testing set, respectively.

3) Results on NIFD dataset: According to TABLE I, BGL-Net performs the best on the NIFD dataset, obtaining the top BA on both validation and testing set. Generally, for baseline models, while deep learning models perform better on the validation set, they suffer from the problem of overfitting as their testing performances are not as good as the three popular methods.

4) Cross-site evaluation: To verify the performance of the proposed method on the multi-site dataset, we test the AD vs. CN model trained on the ADNI dataset on the full dataset of OASIS and NIFD, respectively. As shown in Table IV, BGL-Net achieves the highest BA in both datasets, which further illustrates the generalization of the proposed method is better than other contrasting methods.

5) Comparison with other GCN methods: We compared the performance of BGL-Net with recent other GCN-based methods that focus on AD and CN classification. The results are illustrated in TABLE V. Various data and preprocessing pipelines are adopted in these studies. In addition, the lack of implementation details makes these methods can not to reproduce. Therefore, the results reported in their papers are used here for a naive direct comparison.

The construction of graph representations in these GCN-based methods relies on manual intervention. For example, the extraction of node features might depend on the cortical thickness given by software like Freesurfer and the adjacency matrix may be built based on the correlation between features from different nodes. Unlike these methods, BGL-Net adaptively generates the graph representation with the help of a learnable GGM module to build more robust feature space mapping. It can be seen from TABLE V that compared with manually defined graph representation, BGL-Net is able to achieve better performance while using fewer data.

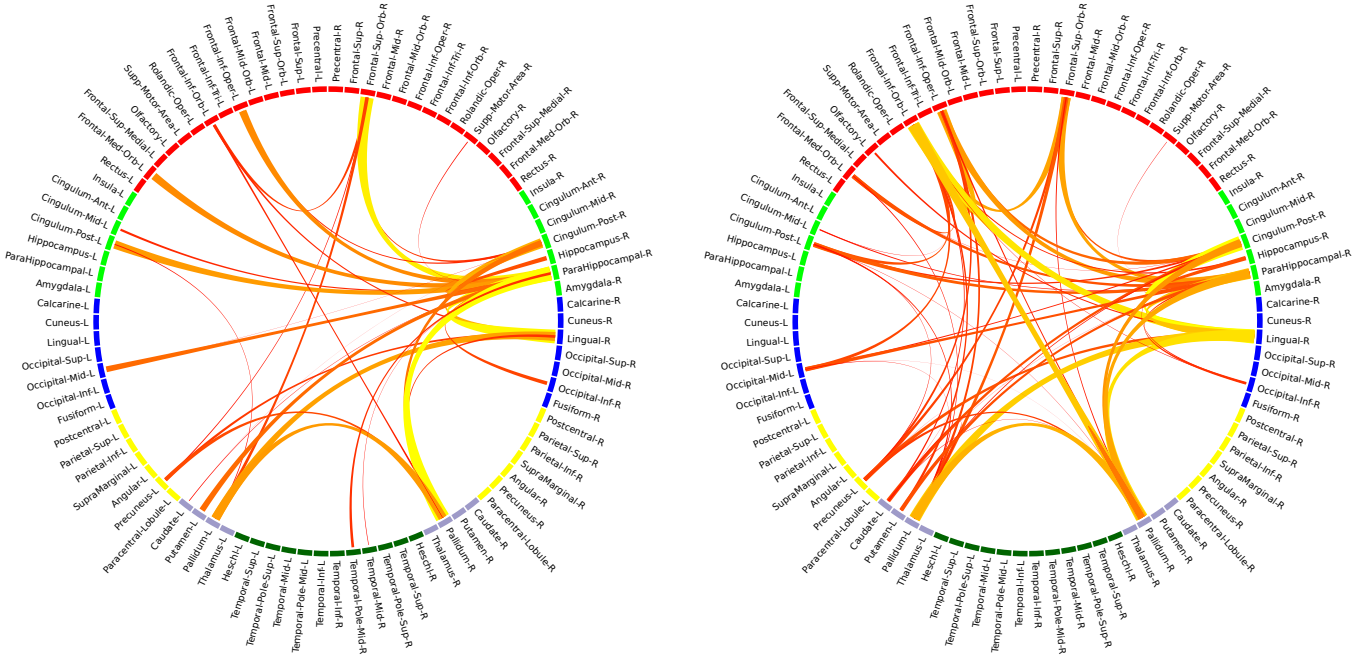
TABLE VII: Comparison of computing resources.

Model	Para.(M)	Infer time(ms)	FLOPs(G)	Memory(M)
DeepCNN [16]	0.12	1.29	0.87	1431
ConvNe3D [17]	8.38	105.84	548.66	7445
VoxCNN [18]	2.14	35.34	50.49	3439
ResNet101 [26]	85.21	111.79	409.65	8131
DenseNet121 [27]	11.24	120.17	345.02	15859
BGL-Net(ours)	3.12	70.59	84.68	17137

6) Ablation Study: An ablation study is conducted on the ADNI dataset to evaluate the effectiveness of the global and local information features. As shown in Table VI, the local information features achieved a BA of 87.83% and 87.20% in the validation set and testing set, respectively, which is similar to other CNN-based methods. The BA of global information features on the validation set and testing set is 85.21% and 85.10%, respectively, which is comparable to other GCN-based methods. After integrating the features of the two type levels, the BA of the validation set is increased by 2.35% and 4.97%, respectively. Compared with the Brian-level and global information features, and the testing set is increased by 1.8% and 3.9%, respectively.

D. Computing resources compare

The computation and inference time of BGL-Net are 84.68G flops and 70.59ms, respectively, which are better than ConvNet3D, DenseNet121, and ResNet101. In terms of parameters, BGL-Net's parameter size (3.12M) is much smaller than that of two deep learning models. Specifically, it is about 1/3 of DenseNet121 and about 1/27 of ResNet101. In terms of memory occupation, since BGL-Net has two branches for global and local information extraction, it requires more memory during training. Overall, BGL-Net achieves the best performance on the three datasets at the cost of a limited memory increase.



(a). Brain connection of AD. (b). Brain connection of CN.

Fig. 4: The connectograms for areas of the brain with different diagnosis groups which learned by BGL-Net. The left and right halves in the figure represent the brain's left and right brain regions, and curves represent the connections between different brain regions. The thicker the curve and the yellower the color, the stronger the connection between brain regions. Adjacency matrices of all subjects in AD and CN are averaged and then normalized, respectively. Here we demonstrate the results of connections stronger than 0.6.

E. Interpretability study

To better understand how BGL-Net arrives at its predictions, as shown in Figure 4, we visualize the connection between different brain regions by further processing the adjacency matrix A. Specifically, the adjacency matrices of all subjects in each category are averaged and then normalized. Only those connections that are stronger than 0.6 are demonstrated.

As AD causes brain disorders among different brain regions, it is interesting to see that the connection between different brain ROIs of AD patients is weaker than that of CN subjects. Specifically, the weakened connection mainly occurs in the left hemisphere, and such biased detection agrees with many current AD clinical research [33], [34].

V. Conclusion

This paper proposes a brain-inspired global-local information fusion network BGL-Net for AD diagnosis. It essentially inherits the advantages of both CNN and GCN to explore global and local information capture. The experiments on three public datasets demonstrate the effectiveness and robustness of our BGL-Net. Our method achieved the best performance on three popular public datasets compared with the existing CNN and GCN-based methods. We believe the proposed method takes a significant step in constructing an sMRI based AD diagnosis medical diagnostic system.

References

[1] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward et al., "The alzheimer's disease neuroimaging initiative ADNI: MRI methods," Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, vol. 27, no. 4, pp. 685–691, 2008.

[2] B. Murugesan, S. V. Raghavan, K. Sarveswaran, K. Ram, and M. Sivaprakasam, "Recon-GLGAN: A global-local context based generative adversarial network for MRI reconstruction," in International Workshop on Machine Learning for Medical Image Reconstruction. Springer, 2019, pp. 3–15.

[3] V. R. Sampathkumar, "Adiag: Graph neural network based diagnosis of alzheimer's disease," arXiv preprint arXiv:2101.02870, 2021.

[4] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7794–7803.

[5] B. Wu, C. Xu, X. Dai, A. Wan, P. Zhang, M. Tomizuka, K. Keutzer, and P. Vajda, "Visual transformers: Token-based image representation and processing for computer vision," arXiv preprint arXiv:2006.03677, 2020.

[6] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2020.

[7] J. delEtoile and H. Adeli, "Graph theory and brain connectivity in alzheimer's disease," The Neuroscientist, vol. 23, no. 6, pp. 616–626, 2017.

[8] S. Rathore, A. Abdulkadir, and C. Davatzikos, "Analysis of MRI data in diagnostic neuroradiology," Annual Review of Biomedical Data Science, vol. 3, pp. 365–390, 2020.

[9] A. Valliani and A. Soni, "Deep residual nets for improved alzheimer's diagnosis," in Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, 2017, pp. 615–615.

[10] D. Jin, B. Zhou, Y. Han, J. Ren, T. Han, B. Liu, J. Lu, C. Song, P. Wang, D. Wang et al., "Generalizable, reproducible, and neuroscientifically interpretable imaging biomarkers for alzheimer's disease," Advanced Science, vol. 7, no. 14, p. 2000675, 2020.

[11] D. Cheng, M. Liu, J. Fu, and Y. Wang, "Classification of MR brain images by combination of multi-CNNs for ad diagnosis," in Ninth International Conference on Digital Image Processing (ICDIP 2017), vol. 10420. International Society for Optics and Photonics, 2017, p. 1042042.

[12] M. Liu, D. Cheng, K. Wang, and Y. Wang, "Multi-modality cascaded convolutional neural networks for alzheimer's disease diagnosis," Neuroinformatics, vol. 16, no. 3, pp. 295–308, 2018.

[13] F. Li, M. Liu, A. D. N. Initiative et al., "Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks," Computerized Medical Imaging and Graphics, vol. 70, pp. 101–110, 2018.

[14] 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 Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 4, pp. 880–893, 2018.

[15] K. Aderghal, A. Khvostikov, A. Krylov, J. Benois-Pineau, K. Afdel, and G. Catheline, "Classification of alzheimer disease on imaging modalities with deep CNNs using cross-modal transfer learning," in 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS). IEEE, 2018, pp. 345–350.

[16] D. Cheng and M. Liu, "CNNs based multi-modality classification for ad diagnosis," in 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017, pp. 1–5.

[17] K. Bäckström, M. Nazari, I. Y.-H. Gu, and A. S. Jakola, "An efficient 3D deep convolutional network for alzheimer's disease diagnosis using mr images," in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 149–153.

[18] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, "Residual and plain convolutional neural networks for 3D brain MRI classification," in 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 2017, pp. 835–838.

[19] W. Zhang, G. Yang, N. Zhang, L. Xu, and V. Albuquerque, "Multi-task learning with multi-view weighted fusion attention for artery-specific calcification analysis," Information Fusion, vol. 71, no. 6, 2021.

[20] C.-Y. Wee, C. Liu, A. Lee, J. S. Poh, H. Ji, A. Qiu, A. D. N. Initiative et al., "Cortical graph neural network for ad and MCI diagnosis and transfer learning across populations," NeuroImage: Clinical, vol. 23, p. 101929, 2019.

[21] A. Kazi, S. Shekarforoush, K. Kortuem, S. Albarqouni, N. Navab et al., "Self-attention equipped graph convolutions for disease prediction," in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019, pp. 1896–1899.

[22] A. Kazi, S. Shekarforoush, S. A. Krishna, H. Burwinkel, G. Vivar, B. Wiestler, K. Kortüm, S.-A. Ahmadi, S. Albarqouni, and N. Navab, "Graph convolution based attention model for personalized disease prediction," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2019, pp. 122–130.

[23] X. Song, M. Mao, and X. Qian, "Auto-metric graph neural network based on a meta-learning strategy for the diagnosis of alzheimer's disease," IEEE Journal of Biomedical and Health Informatics, 2021.

[24] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," Neuroimage, vol. 15, no. 1, pp. 273–289, 2002.

[25] Z. Long, J. Huang, B. Li, Z. Li, Z. Li, H. Chen, and B. Jing, "A comparative atlas-based recognition of mild cognitive impairment with voxel-based morphometry," Frontiers in Neuroscience, vol. 12, p. 916, 2018.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[27] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

[28] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies OASIS: cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," Journal of Cognitive Neuroscience, vol. 19, no. 9, pp. 1498–1507, 2007.

[29] K. J. Gorgolewski, T. Auer, V. D. Calhoun, R. C. Craddock, S. Das, E. P. Duff, G. Flandin, S. S. Ghosh, T. Glatard, Y. O. Halchenko et al., "The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments," Scientific Data, vol. 3, no. 1, pp. 1–9, 2016.

[30] A. Routier, N. Burgos, M. Díaz, M. Bacci, S. Bottani, O. El-Rifai, S. Fontanella, P. Gori, J. Guillon, A. Guyot et al., "Clinica: An open-source software platform for reproducible clinical neuroscience studies," Frontiers in Neuroinformatics, vol. 15, 2021.

[31] J. Samper-Gonzalez, N. Burgos, S. Bottani, S. Fontanella, P. Lu, A. Marcoux, A. Routier, J. Guillon, M. Bacci, J. Wen et al., "Reproducible evaluation of classification methods in alzheimer's disease: Framework and application to MRI and PET data," NeuroImage, vol. 183, pp. 504–521, 2018.

[32] J. Ashburner and K. J. Friston, "Unified segmentation," Neuroimage, vol. 26, no. 3, pp. 839–851, 2005.

[33] J. Hutsler and R. A. Galuske, "Hemispheric asymmetries in cerebral cortical networks," Trends in Neurosciences, vol. 26, no. 8, pp. 429–435, 2003.

[34] X. Ma, G. Wu, and W. H. Kim, "Enriching statistical inferences on brain connectivity for alzheimer's disease analysis via latent space graph embedding," in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE, 2020, pp. 1685–1689.