



Identification of Autism Spectrum Disorder With Functional Graph Discriminative Network

Jingcong Li^{1,2}, Fei Wang^{1,2*}, Jiahui Pan^{1,2} and Zhenfu Wen³

¹ School of Software, South China Normal University, Guangzhou, China, ² Pazhou Lab, Guangzhou, China, ³ Department of Psychiatry, New York University School of Medicine, New York, NY, United States

Autism spectrum disorder (ASD) is a specific brain disease that causes communication impairments and restricted interests. Functional connectivity analysis methodology is widely used in neuroscience research and shows much potential in discriminating ASD patients from healthy controls. However, due to heterogeneity of ASD patients, the performance of conventional functional connectivity classification methods is relatively poor. Graph neural network is an effective graph representation method to model structured data like functional connectivity. In this paper, we proposed a functional graph discriminative network (FGDN) for ASD classification. On the basis of pre-built graph templates, the proposed FGDN is able to effectively distinguish ASD patient from health controls. Moreover, we studied the size of training set for effective training, inter-site predictions, and discriminative brain regions. Discriminative brain regions were determined by the proposed model to investigate its applicability and biomarkers for ASD identification. For functional connectivity classification and analysis, FGDN is not only an effective tool for ASD identification but also a potential technique in neuroscience research.

Keywords: autism spectrum disorder, ABIDE, graph neural network, functional graph, resting-state functional MRI

OPEN ACCESS

Edited by:

Gopikrishna Deshpande,
Auburn University, United States

Reviewed by:

Niharika D'Souza,
Johns Hopkins University,
United States
Anubha Gupta,
Indraprastha Institute of Information
Technology Delhi, India

*Correspondence:

Fei Wang
fwang@schnu.edu.cn

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 24 June 2021

Accepted: 07 September 2021

Published: 21 October 2021

Citation:

Li J, Wang F, Pan J and Wen Z (2021)
Identification of Autism Spectrum
Disorder With Functional Graph
Discriminative Network.
Front. Neurosci. 15:729937.
doi: 10.3389/fnins.2021.729937

1. INTRODUCTION

Autism spectrum disorder (ASD) is a specific brain disease that causes social and communication impairments, restricted interests, and repetitive behaviors (Lord et al., 2000; Yang et al., 2019). Early diagnosis of ASD is significant for preparing treatment plan and conducting early intervention for ASD. Human brain connectivity network aims to reveal structural and functional interactions between brain regions, which were proven to be potential in identifying predictive biomarkers for neurodevelopmental and neuropsychiatric disorders (Wu et al., 2006; Arslan et al., 2018).

The resting-state functional magnetic resonance imaging (rs-fMRI) is widely used to study functional connectivity between different brain regions with its high spatial resolution (Luca et al., 2006; Aggarwal et al., 2017). Based on rs-fMRI and functional connectivity, it is possible to develop a reliable and objective technique for early diagnosis of ASD (Abraham et al., 2017; Aggarwal and Gupta, 2019b; Dadi et al., 2019). To facilitate techniques for ASD identification with rs-fMRI data, a large multisite dataset termed Autism Brain Imaging Data Exchange (ABIDE) was released (Nielsen et al., 2013). The ABIDE dataset consists of the rs-fMRI data of ASD patients and healthy controls (HC) from different international acquisition sites and using different protocols. In the ABIDE dataset, the challenging problems in identifying ASD patients lie in the individual differences of functional connectivity as well as acquisition protocol differences.

In the past few years, many machine learning approaches were proposed for processing functional connectivity data and identifying ASD. As a popular machine learning method, support vector machine (SVM) was applied for ASD prediction based on Pearson's correlation functional connectivity (Nielsen et al., 2013). In another significant research on ABIDE dataset, the researchers investigated a few classifiers including random forests, Gaussian naive Bayes, support vector classifier, and ridge classifier (Abraham et al., 2017). The major contribution of this research lies in finding best predicting pipelines for ABIDE dataset, which were based on Multi Subject Dictionary Learning (MSDL) atlas, tangent space embedding, and ℓ_2 -regularized classifiers. Also evaluated on the ABIDE dataset, higher performance (e.g., accuracy over 75%) could be obtained by using extraneous information such as "multiple atlases" or "non-physiological" (Karampasi et al., 2020, 2021; Epalle et al., 2021). On the basis of correlation matrices computed from rs-fMRI time-series data, a probabilistic neural network was applied for ASD classification (Iidaka, 2015). A deep neural network based on stacked autoencoder was proposed to identify ASD patients from typical controls with T1-weighted MRI images that it outperformed some state-of-the-art methods (Kong et al., 2018). The deep learning technique also showed potential in identifying ASD based on behavior data like videos (Li et al., 2019). Compared with baseline methods, a mathematical framework based on Riemannian geometry and kernel methods achieved superior performance for functional connectivity graphs classification (Dodero et al., 2015).

Recently, graph representation methodology was proven to be a powerful tool in modeling structured data and achieved significant performance in many applications (Linial et al., 1995; Even, 2011; Aggarwal and Gupta, 2019a). Functional connectivity between brain regions in rs-fMRI data can be considered as typical structured data (Stam et al., 2008). Graph representation approaches also achieved impressive performance in dealing with functional connectivity data (Bullmore and Bassett, 2011; Sporns, 2011; Parisot et al., 2018). Accordingly, graph-based brain network is able to uncover system-level changes of brain regions (Wang et al., 2010). To predict progress of patients with mild cognitive impairment to Alzheimer's disease using rs-fMRI, graph theory, and machine learning approach were utilized (Hojjati et al., 2017). A graph convolutional network termed MTGCN was proposed for learning multi-scale graph representations of brain functional connectivity analysis with rs-fMRI data (Yao et al., 2019). Graph-based network also showed its potentials in ASD diagnosis that a siamese graph convolutional neural network (s-GCN) was proposed and evaluated on ABIDE dataset (Ktena et al., 2018). A graph convolutional neural network was proposed to model phenotypic and demographic information of subjects, and achieved significant improvements in ASD classification accuracy on ABIDE dataset (Parisot et al., 2018). However, due to individual differences and the unknown patterns of functional connectivity, conventional graph-based approaches perform poorly in some functional connectivity classification problems.

In this paper, we propose a modified graph convolution network for ASD classification and conduct a series of

experiments on the ABIDE dataset. The main contributions of this paper can be summarized as follows:

1. An FGDN is proposed for ASD prediction based on resting-state functional MRI.
2. The proposed FGDN is able to achieve a high ASD classification performance on the ABIDE dataset.
3. Inter-site predictions and discriminative brain regions were studied with the proposed model to investigate its applicability and biomarkers for ASD identification.

The remainder of this paper is organized as follows. The proposed FGDN is presented in section 2. In section 3, numerical ASD classification experiments on the ABIDE dataset are carried out. In addition, the performance of the benchmark methods and the proposed methods are presented and compared. A general discussion of the proposed model is presented in section 4. Conclusions of this paper are given in section 5.

2. MATERIALS AND METHODS

The functional connectivity is based on the hypothesis that different regions of interest (ROIs) could capture relevant functional activities within the brain. The ROIs are defined by brain structural atlases, such as Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002), Harvard Oxford (HO) (Desikan et al., 2006), and Massive Online Dictionary Learning (MODL) (Dadi et al., 2019). According to previous research (Ktena et al., 2018; Dadi et al., 2019), functional connectivity is a typical structured data, which is suitable to define on a graph.

2.1. Graph Construction

Generally, a graph model can be constructed as follows:

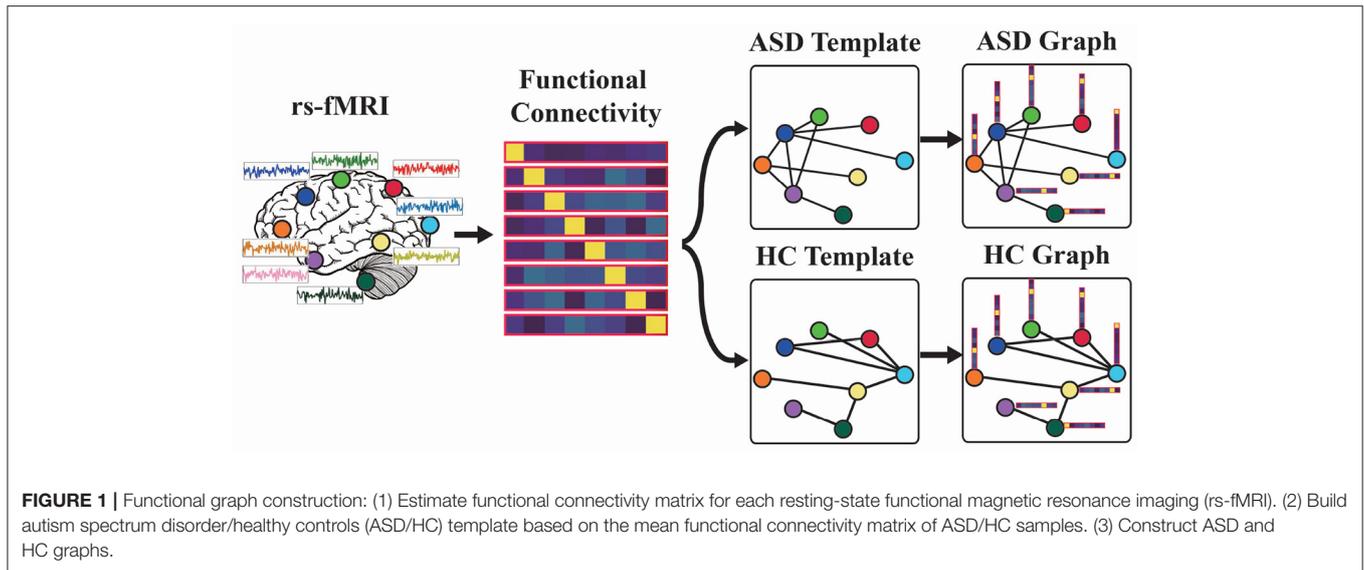
$$\begin{aligned} \mathcal{G} &= (\mathcal{V}, \mathcal{E}, W) \\ \mathcal{V} &= \{v_i | i = 1, \dots, N\} \\ \mathcal{E} &= \{e_{ij} | v_i, v_j \in \mathcal{V}\} \\ W &= \{w_{ij}\} \end{aligned} \quad (1)$$

where \mathcal{V} denotes the set of nodes (totally N nodes) in graph \mathcal{G} , \mathcal{E} are connected edges between different nodes, and $W \in \mathbb{R}^{N \times N}$ is the adjacency matrix whose element w_{ij} denotes the weight of connection between i th node and j th node. The element w_{ij} of adjacency matrix W is usually determined by the distance function and k-nearest neighbor rule. Gaussian kernel function is a commonly used distance function as

$$w_{ij} = \begin{cases} \exp(-\frac{[dist(v_i, v_j)]^2}{2\theta^2}) & dist(v_i, v_j) \leq \tau \\ 0 & otherwise \end{cases} \quad (2)$$

where $dist(v_i, v_j)$ denotes the distance between i th node and j th node, τ and θ are two fixed parameters that determined by K-nearest neighbor (KNN) method for extracting k nearest nodes in graph \mathcal{G} (Ktena et al., 2018; Song et al., 2019).

As mentioned above, the functional connectivity is a typical structured data that could be transformed into a graph (Ktena et al., 2018; Parisot et al., 2018; Dadi et al., 2019). Usually, each row vector of functional connectivity matrix is considered as one



nodes in the graph. To model the connections between different nodes of functional connectivity graph, we utilized a functional graph construction technique (Ktena et al., 2018). Accordingly, such kind of graph structure could reflect functional connection weights between different regions from neuroscientific view.

The mean functional connectivity matrix and the corresponding graph structure are calculated as follows:

$$\begin{aligned} \bar{\mathbf{F}}_c &= \frac{1}{N_c} \sum_{i=1}^{N_c} \mathbf{F}_{ic} = [f_1, \dots, f_i, \dots, f_N]^T \\ \mathcal{V}_{\bar{\mathbf{F}}_c} &= \{f_i | i = 1, \dots, N\} \\ \mathbf{T}_c &= \{\mathcal{E}, W\} = \text{KNN}(\mathcal{V}_{\bar{\mathbf{F}}_c}) \end{aligned} \tag{3}$$

where \mathbf{F}_{ic} denotes functional connectivity matrix of the i th training rs-fMRI data, N_c is the number of training samples, and c denotes the category of ASD ($c = 0$) or HC ($c = 1$); the row vector f_i of $\bar{\mathbf{F}}_c$ is considered as the i th node feature of $\mathcal{V}_{\bar{\mathbf{F}}_c}$; \mathbf{T}_c (termed graph template) denotes the connected edges \mathcal{E} and their weights W between different nodes, \mathcal{E} and W are determined by KNN method and Equations (1) and (2) with node features $\mathcal{V}_{\bar{\mathbf{F}}_c}$. The output graph template could be considered as the typical connections between nodes of ASD or HC samples.

Given the functional connectivity of a test rs-fMRI sample, it could be transformed into two graphs with the ASD and HC templates as shown in **Figure 1**. The interaction between the functional connectivity with the ASD/HC template is as follows:

$$\begin{aligned} \mathbf{F} &= [f_1, \dots, f_i, \dots, f_N]^T \\ \mathcal{V}_{\mathbf{F}} &= \{f_i | i = 1, \dots, N\} \\ \mathbf{T}_c &= \{\mathcal{E}, W\} = \text{KNN}(\mathcal{V}_{\bar{\mathbf{F}}_c}) \\ \mathcal{G}_c &= (\mathcal{V}, \mathcal{E}, W) = (\mathcal{V}_{\mathbf{F}}, \mathbf{T}_c) \end{aligned} \tag{4}$$

where \mathbf{F} is the input functional connectivity, the row vector f_i of \mathbf{F} is the feature of the i th node in $\mathcal{V}_{\mathbf{F}}$, \mathbf{T}_c is graph template mentioned in Equation (3), and \mathcal{G}_c is the output graph of ASD ($c = 0$) or HC ($c = 1$). For a test rs-fMRI sample, its output ASD graph and HC graph will have the same node features while the connections between their nodes are different.

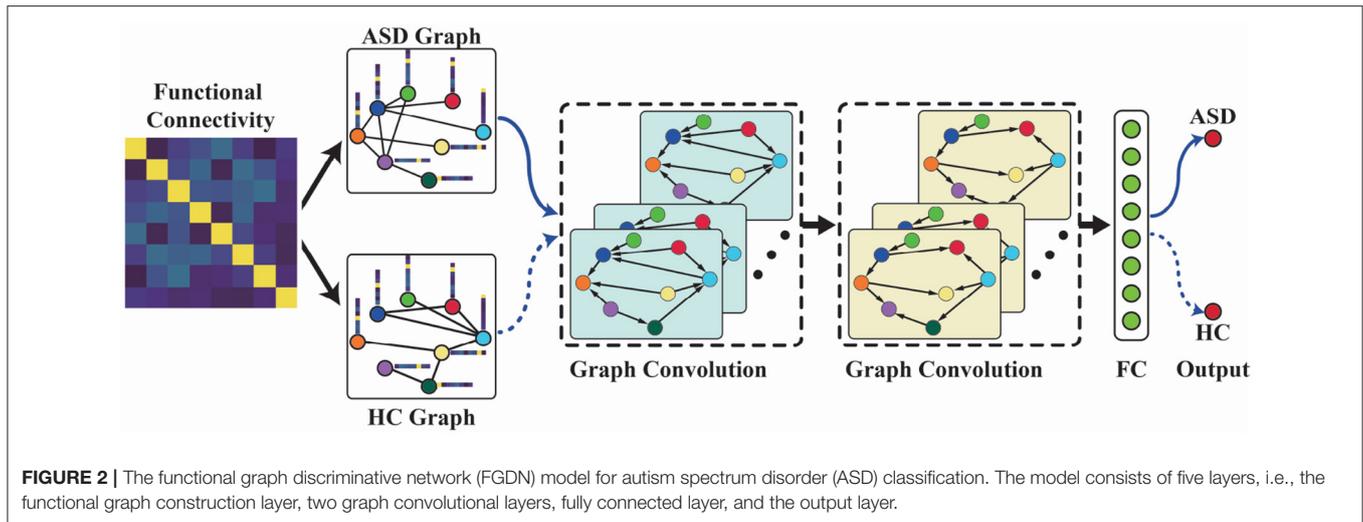
As shown in **Figure 1**, we first estimate the functional connectivity of each rs-fMRI sample. Second, we will build the ASD/HC template based on the mean functional connectivity matrix of ASD/HC data according to Equation (3). Third, the functional connectivity will be combined with the ASD/HC template to construct ASD/HC graph. For a test rs-fMRI sample, we can obtain its node features (i.e., functional connectivity) while the connections between nodes are unknown. So we feed the node features into ASD graph template and HC template to construct ASD graph and HC graph, and then let the model to determine which graph matches the test sample. Therefore, we would like to build a model for ASD graph identification in the next section.

2.2. Proposed Graph Model

Generally, standard convolutional operations for regular data are inappropriate for processing graph data due to its irregular distribution of nodes (Krizhevsky et al., 2012). According to the previous research, two approaches can be utilized for generalizing convolution operations to graph data. First one is to rearrange all of the graph nodes into a regular grid and then conduct standard convolution operation (Niepert et al., 2016). However, the graph structure will be seriously corrupted in this way. The second approach is to apply spectral graph convolution (Bruna et al., 2013). By using convolutions in spectral domains with graph Fourier transform, spectral graph convolution can be feasibly applied on graph data. In addition, there are some following-up studies that the computational complexity of spectral graph convolution can be reduced from $\mathcal{O}(n^2)$ to linear (Defferrard et al., 2016; Kipf and Welling, 2016).

According to previous research (Bruna et al., 2013; Yu et al., 2017; Song et al., 2019), spectral graph convolution is to multiply a signal $x \in \mathbb{R}^n$ with a graph convolution kernel Θ and a graph convolution operator $*_{\mathcal{G}}$ as,

$$\Theta *_{\mathcal{G}} x = \Theta(L)x = \Theta(U\Lambda U^T)x = U\Theta(\Lambda)U^T x \tag{5}$$



where graph Fourier basis $U \in \mathbb{R}^{n \times n}$ is the matrix of eigenvectors of the normalized graph Laplacian $L = I_n - D^{-1/2}WD^{-1/2} = U\Lambda U^T \in \mathbb{R}^{n \times n}$ (I_n is an identity matrix, $D \in \mathbb{R}^{n \times n}$ is the diagonal degree matrix with $D_{ii} = \sum_j W_{ij}$, $W \in \mathbb{R}^{N \times N}$ is the adjacency matrix mentioned in equation (1)); $\Lambda \in \mathbb{R}^{n \times n}$ is the diagonal matrix of eigenvalues of L , and filter $\Theta(\Lambda)$ is also a diagonal matrix. By this definition, a graph signal x is filtered by a kernel Θ with multiplication between Θ and graph Fourier transform $U^T x$ (Shuman et al., 2013). The computational complexity of spectral graph convolution is expensive due to $\mathcal{O}(n^2)$ multiplications with graph Fourier basis.

In order to reduce computational complexity of spectral graph convolution, Chebyshev polynomials approximation technique was proposed (Defferrard et al., 2016). To reduce the number of parameters and localize the graph filter, the graph convolutional kernel Θ is restricted to a Chebyshev polynomial form as

$$\Theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\Lambda}) \quad (6)$$

where $\theta \in \mathbb{R}^K$ is the Chebyshev polynomial coefficients, K is the size of graph convolutional kernel which determines maximum convolutional range from central nodes, Λ is rescaled by $\tilde{\Lambda} = 2\Lambda/\lambda_{\max} - I_n$ (λ_{\max} denotes the largest eigenvalue of L , and I_n is a $N \times N$ identity matrix). The $T_k(x)$ could be recursively calculated as

$$\begin{aligned} T_0(x) &= 1, T_1(x) = x \\ T_k(x) &= 2xT_{k-1}(x) - T_{k-2}(x), k \geq 2 \end{aligned} \quad (7)$$

Then, the spectral graph convolution is rewritten as

$$\Theta *_{\mathcal{G}} x = \Theta(L)x \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{L})x \quad (8)$$

where $T_k(\tilde{L}) \in \mathbb{R}^{n \times n}$ is k -order Chebyshev polynomial estimation for the rescaled Laplacian $\tilde{L} = 2L/\lambda_{\max} -$

I_n . Then the computational complexity of spectral graph convolution by Equation (5) is reduced from $\mathcal{O}(n^2)$ to $\mathcal{O}(K|\varepsilon|)$ (Defferrard et al., 2016).

On the basis of functional graph construction and spectral graph convolution techniques, we proposed an FGDN for ASD classification. The framework of the proposed FGDN model is illustrated in **Figure 2**, which consists of five layers, i.e., the functional graph construction layer, two graph convolutional layers, fully connected layer, and one output layer. Given the ASD graph of an input sample, the corresponding output of the ASD unit (ASD output) is obtained. Likewise, the corresponding HC output is obtained by the HC graph of the same input sample. If the ASD output value is larger than the HC output, the input sample will be categorized as ASD. If not, it will be considered as HC class.

In the first layer of FGDN, each input functional connectivity data is transformed into ASD and HC graphs by Equations (1)–(5) and **Figure 1**. The number of graph nodes is determined by ROIs of the atlas, which were applied to model functional brain regions. In the paper, we will investigate the atlas of AAL (Tzourio-Mazoyer et al., 2002), Harvard Oxford (Desikan et al., 2006), and MODL (Dadi et al., 2019), which consists of 116, 118, or 128 ROIs (nodes). Then, the ASD graph and HC graph will be fed into the following network to obtain ASD and HC output, respectively. In each spectral graph convolution layer, there are 64 graph convolutional kernels. Moreover, we utilized korder Chebyshev spectral graph convolutional operation and applied $k = 3$ following the previous research (Ktena et al., 2018; Parisot et al., 2018). With two concatenated graph convolution layers, the graph features could be extracted hierarchically that is similar to convolution operations on grid data. The output of graph convolution layer is flattened to obtain the full connection (FC) layer. Parametric rectified linear unit (PReLU) activation function is applied in the two graph convolution layers as well as FC layer. Fully connected to the FC layer, the output layer of FGDN model consists of two units, i.e., the ASD and HC output units with sigmoid activation function.

Then, the proposed model can be trained by minimizing cross-entropy error of its predictions and the ground truth. As a result, the loss function is defined as

$$L = - \sum_{i \in \Omega} \sum_c y_{ic} \log(p_{ic}) + (1 - y_{ic}) \log(1 - p_{ic}) \quad (9)$$

where p_{ic} is the value of ASD output unit ($c = 1$) or HC output unit ($c = 2$) of FGDN model with the input of the i th training sample, p_{ic} can be considered as the model's predicted probability of ASD or HC class, y_{ic} is the corresponding ground truth, and Ω denotes all of training samples.

In the next section, a series of experiments will be carried out to evaluate the proposed FGDN model. In addition, the corresponding experimental results of our method will be presented and compared with the other methods.

3. RESULTS

In this section, the experimental procedure and results of the proposed FGDN will be presented. In the experiments, the proposed method and some other popular ASD classification methods are evaluated by the common ABIDE dataset. For ease of reproduction, the training procedure and the parameter settings of the proposed model are illustrated in the paper.

In our experiments, the hardware and software configuration of our system is a platform with Nvidia Titan Xp, Ubuntu 16.04, and Pytorch 1.2.0. The graph convolution operation of our work are achieved under the support of Pytorch-geometric 1.3.2 (Fey and Lenssen, 2019).

3.1. ABIDE Data Preprocessing

The ABIDE dataset was built for investigating the neural basis of autism and facilitating the development of ASD diagnosis techniques (Martino et al., 2014). We use the ABIDE data that were preprocessed by the previous researchers (Dadi et al., 2019) to discriminate individuals with autism spectrum disorder from healthy controls. In the preprocessed ABIDE dataset, there are 402 ASD samples and 464 HC samples. Following the pipeline of previous research (Dadi et al., 2019), we utilized Ledoit-Wolf regularized shrinkage estimator (Ledoit and Wolf, 2004) to efficiently estimate functional connectivity. According to previous research (Dadi et al., 2019), the tangent space embedding technique for estimating functional connectivity could effectively improve ASD classification performance. Likewise, we utilized tangent space embedding method to calculate the functional connectivity of each rs-fMRI samples.

Following the above instructions in Equations (1)–(5), ASD and HC functional graphs could be obtained. Here, the number of neighbors in KNN method was set to 20 for graph construction. Consequently, functional connectivity of each rs-fMRI data is transformed to be an ASD graph and an HC graph with the same node signals, i.e., the row vectors of its functional connectivity matrix. Each rs-fMRI signal will be transformed into ASD and HC graphs, then fed into the FGDN model and obtained the output values of ASD and HC units. Then the category of the input rs-fMRI sample could be identified.

3.2. Experiments

Before the experiment, the ABIDE dataset with 866 samples (402 ASD and 464 HC) is randomly split into five- or ten-folds for cross-validation (CV). Note that five- and ten-fold cross-validations were widely used in the previous studies. To evaluate the performance of different models, we measure the averaged CV accuracy and area under the curve (AUC) from the receiver operating characteristics curve.

In order to train the proposed FGDN model, Adam optimizer (Kingma and Ba, 2014) is applied for minimizing the model's loss function equation (9). The proposed model was trained by Adam optimizer with a learning rate of 0.0001, a weight-decay rate of 0.0005, and mini-batch size of 16. Drop-out operation with a rate of 0.1 was applied in the training procedure for randomly blocking the output units of graph convolution layers. Before the training procedure, we separated 10% training samples as monitoring set. Once the model achieved the highest accuracy on monitoring set, the training procedure was stopped to avoid overfitting. Finally, the trained FGDN model could be applied for ASD prediction.

Once the proposed FGDN model was trained, we could evaluate it and compare with the other models. As shown in **Tables 1, 2**, the experimental results of the proposed FGDN and many other methods on the ABIDE dataset. For evaluation, k-fold cross-validation strategy denoted as CV- n ($n = 5$ and 10) was applied. The averaged accuracy (ACC) of CV-5 and CV-10 along with standard deviation (SD) were presented. The corresponding area under the receive operating characteristic curve (AUC) results were also present. The used atlases, features, and number of subjects of different methods were present for analysis.

Conventional classifiers like logistic regression with L2 regularization (Logistic- ℓ_2) and support vector machine with ℓ_1 regularization (SVM- ℓ_1) were applied as benchmark methods for ASD classification (Rane et al., 2017; Dadi et al., 2019). The corresponding accuracies were 64.1(0.9) and 62.0(–) (– indicated an unknown score), respectively. The Logistic- ℓ_2 was based on tangent feature of MODL atlas, while the SVM- ℓ_1 was based on raw voxel values of standard anatomical space (MNI152). Support vector classifier (SVC) was applied to model tangent feature on MSDL atlas that ACC of 66.9(2.7) was achieved (Abraham et al., 2017). The baseline classifiers such as Gaussian Naive Bayes (GaussianNB), Random Forest (RandomF), Ridge regression (Ridge), and K-nearest neighbors (KNN) were also evaluated in the experiments.

Deep neural networks (DNN) achieved impressive performance in many areas. For ASD classification, a recurrent neural networks with long short-term memory (LSTM) was proposed to model time series on Craddock 200 (CC200) atlas and achieved ACC of 68.5(5.5) (Dvornek et al., 2017). With an ASD classification accuracy of 71.0, the invertible network (InvNet) was proven to be both effective at classification and finding interpretable biomarkers for ASD (Zhuang et al., 2019). As a popular technique in many areas, a DNN model was applied for ASD classification and achieved an ACC of 70.0 (Heinsfeld et al., 2018).

Recently, some studies indicated that graph convolutional network could effectively deal with structured data like functional

TABLE 1 | Five-fold cross-validation performance.

Method	Atlas	Feature	Number of subjects	ACC(SD)%	AUC(SD)%
GaussianNB	MODL	Tangent	866	63.5 (3.1)	65.7 (5.4)
RandomF	MODL	Tangent	866	61.3 (1.8)	66.3 (4.7)
Logistic-L2	MODL	Tangent	866	64.1 (0.9)	69.8 (3.4)
Ridge	MODL	Tangent	866	63.8 (4.4)	63.4 (5.6)
KNN	MODL	Tangent	866	65.9 (4.6)	68.1 (5.2)
SVM-L1	None	Voxel	1,112	62.0 (-)	- (-)
sGCN	HO	Correlation	871	69.0 (-)	- (-)
ASD-SAENet	CC200	Correlation	1,035	70.8 (-)	- (-)
FGDN	HO	Correlation	866	65.7 (2.2)	67.8 (3.5)
FGDN	HO	Tangent	866	71.4 (4.2)	76.9 (3.6)
FGDN	AAL	Correlation	866	66.7 (5.2)	67.5 (3.8)
FGDN	AAL	Tangent	866	72.3 (3.0)	75.1 (4.5)
FGDN	MODL	Correlation	866	68.8 (2.6)	72.2 (1.9)
FGDN	MODL	Tangent	866	72.5 (5.3)	77.8 (4.5)

The bold values indicate the largest values.

TABLE 2 | Ten-fold cross-validation performance.

Method	Atlas	Feature	Number of subjects	ACC(SD)%	AUC(SD)%
GaussianNB	MODL	Tangent	866	64.0 (5.7)	65.0 (6.7)
RandomF	MODL	Tangent	866	62.8 (3.3)	64.9 (5.5)
Logistic-L2	MODL	Tangent	866	64.8 (3.6)	69.0 (5.3)
ridge	MODL	Tangent	866	67.9 (3.2)	70.0 (5.2)
KNN	MODL	Tangent	866	65.8 (6.2)	66.6 (5.9)
SVC	MSDL	Tangent	871	66.9 (2.7)	- (-)
LSTM	CC200	Time Series	1,112	68.5 (5.5)	- (-)
InvNet	CC200	Correlation	1,035	71.0 (-)	- (-)
DNN	CC200	Correlation	1,035	70.0 (-)	- (-)
AttentionET	CC200	Correlation	1,054	72.2 (-)	- (-)
GCN	HO	Correlation	871	70.4 (3.9)	75.0 (4.6)
FGDN	HO	Correlation	866	67.9 (5.5)	69.5 (6.4)
FGDN	HO	Tangent	866	71.4 (4.3)	77.7 (5.3)
FGDN	AAL	Correlation	866	65.9 (3.5)	71.8 (5.1)
FGDN	AAL	Tangent	866	70.8 (5.2)	79.4 (5.0)
FGDN	MODL	Correlation	866	68.0 (3.4)	69.8 (5.2)
FGDN	MODL	Tangent	866	71.8 (4.5)	76.3 (5.6)

The bold values indicate the largest values.

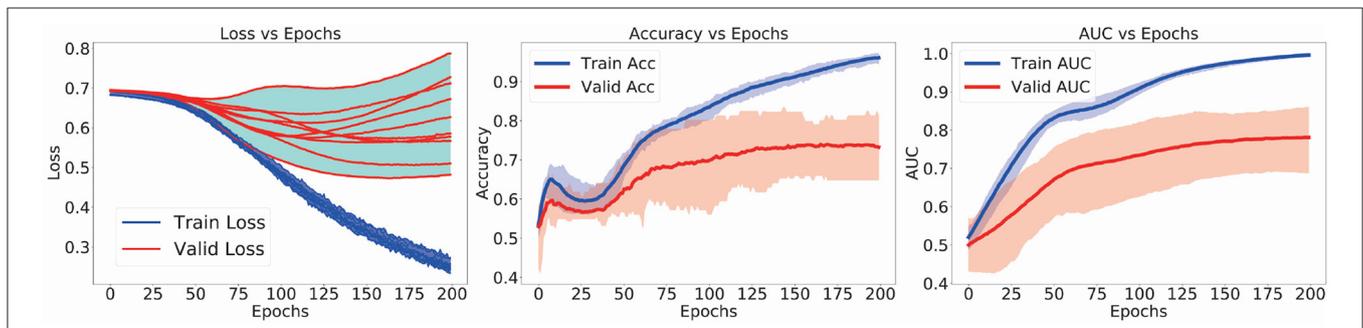


FIGURE 3 | Loss(accuracy or AUC)-vs.-epochs curves in the 10-fold cross-validation (CV) training procedure.

connectivity. A siamese graph convolutional neural network (sGCN) was proposed for functional connectivity classification, which utilized Harvard Oxford (HO) atlas and Pearson correlation features (Ktena et al., 2018). The sGCN was evaluated by CV-5 on 871 ABIDE subjects and achieved an accuracy of 69%. Using the phenotypic information of subjects, a GCN method was proposed to model the connections between subjects that it achieved accuracy of 70.4(3.9) and AUC of 75.0(4.6) in CV-10 experiments (Parisot et al., 2018). An attention mechanism based on Extra-Trees algorithm (AttentionET) achieved an ACC of 72.2, which was state-of-the-art performance on CV-10 of ABIDE dataset (Liu et al., 2020). Based on sparse autoencoder, the ASD-SAENet achieved an accuracy of 70.8 on CV-5 of 1,035 subjects in ABIDE dataset (Almuqhim and Saeed, 2021).

As shown in **Tables 1, 2**, each baseline method had two validation setups, i.e., five-fold CV and 10-fold CV. The FGDN reached a high performance with ACC of 72.5(5.3) and AUC of 77.8(4.5) by five-fold cross-validation (CV-5), and ACC of 71.8(4.5) and AUC of 76.3(5.6) by 10-fold cross-validation (CV-10). The corresponding sensitivity and specificity were 73.9(4.7) and 65.2(5.4) on CV-5, and 74.6(7.0) and 67.8(7.7) on CV-10. Compared to the AttentionET method with sensitivity of 68.8 and specificity of 75.4, the proposed FGDN model had higher sensitivity and lower specificity in CV-10 experiments.

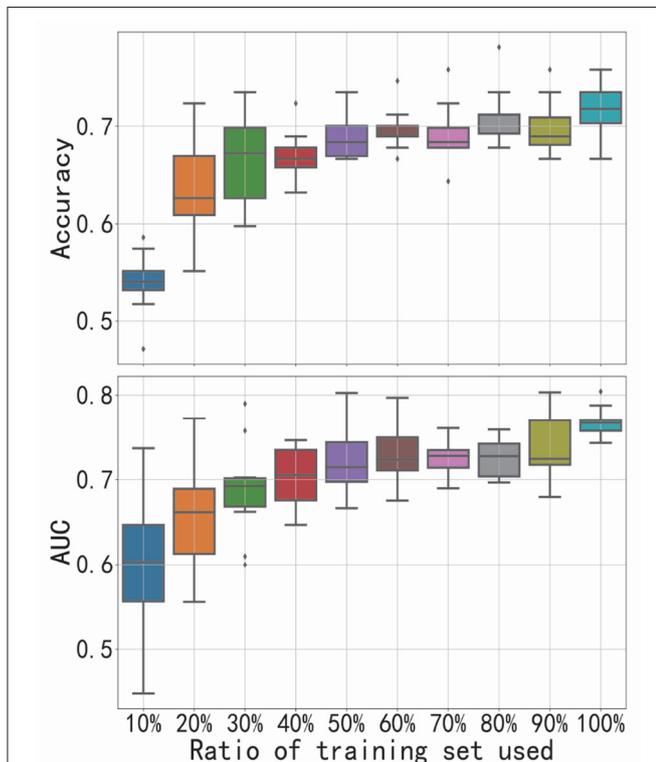


FIGURE 4 | Performance of functional graph discriminative network (FGDN) as the increasing ratio of training set used. Boxplots denote the variant of classification accuracy/area under the curve (AUC) across 10-fold cross-validation.

In **Tables 1, 2**, many baseline methods and the proposed FGDN model were evaluated on the same data subsets (866 samples) and the same input representations (AAL, HO, and MODL atlases with correlation and tangent features). The proposed FGDN achieved higher performance than the GaussianNB, RandomF, Ridge, and KNN. But some of the referenced methods used different atlases, different features, and different number of subjects that we could not compare them in a quantitative way. According to these results, the high performance of the proposed FGDN may depend on the MODL atlas and the tangent features. The combination of MODL and tangent feature was also proven to be most discriminative features on ASD classification according to the previous quantitative experiments (Dadi et al., 2019). Consequently, the cross-validation performance of FGDN demonstrated its effectiveness in ASD classification.

In the next section, we will discuss about the internal properties of the proposed model. It may be significant for further research.

4. DISCUSSION

In this section, the proposed model will be discussed in details. According to the benchmark research on multiple publicly available rest-fMRI datasets (Dadi et al., 2019), the MODL and tangent features were proven to be the best features for rest-fMRI classification. As many previous research (Parisot et al., 2018; Dadi et al., 2019; Zhuang et al., 2019), the 10-fold cross-validation settings was widely used for experiments. Therefore, we all used MODL atlas and tangent features, and evaluated on the CV-10 experiments in the discussion section.

First of all, we present the loss-vs.-epochs curves of the models in the 10-fold CV experiments. In addition, the curves of accuracy and AUC scores were also present. Second, we discuss about the improved performance of FGDN model as the increasing size of training dataset. In the experiments, We varied the number of training samples while keeping a fixed-sized validation set. Third, we analyzed inter-site performance of

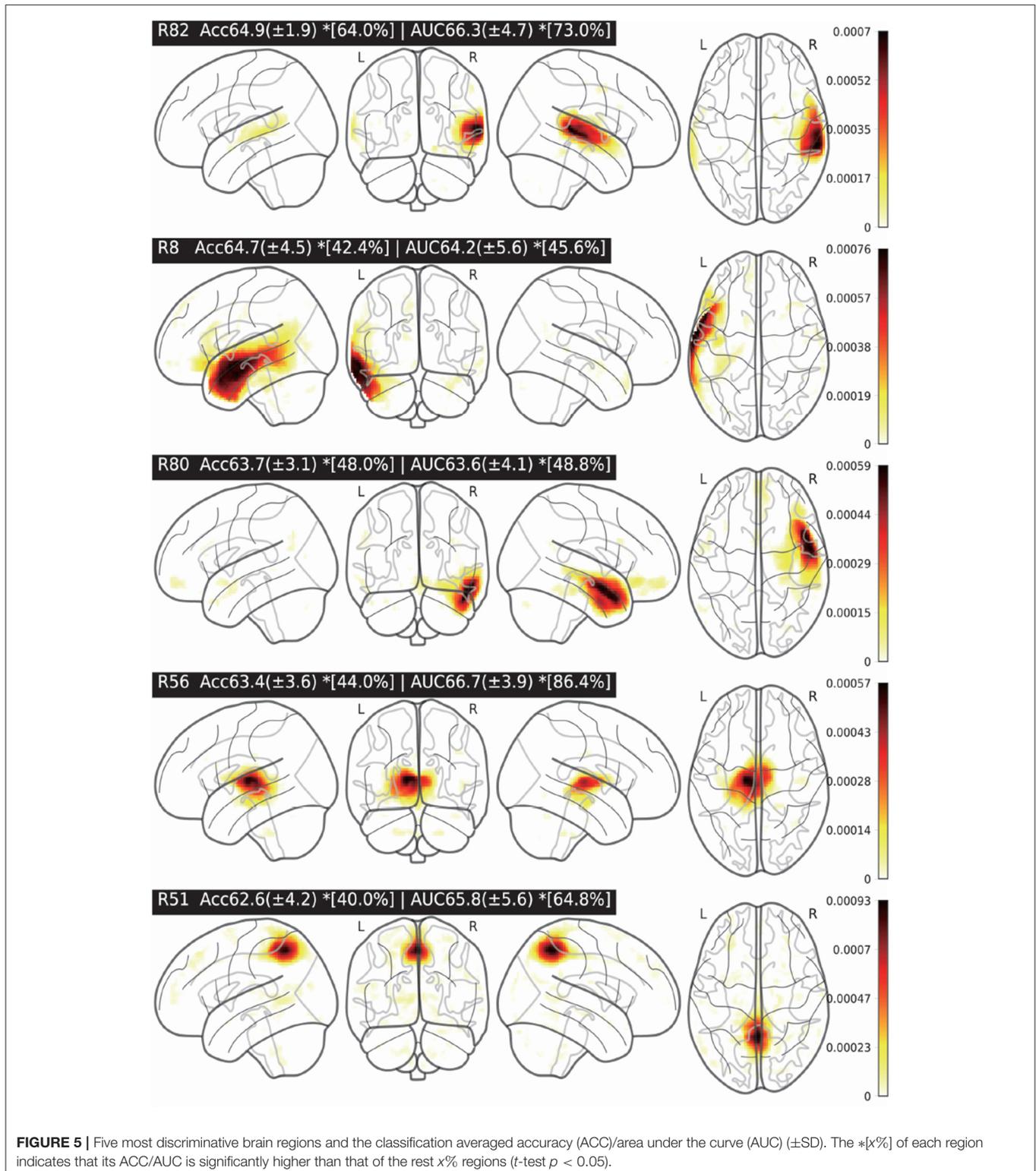
TABLE 3 | Leave-one-site-out autism spectrum disorder (ASD) classification performance of functional graph discriminative network (FGDN) on the Autism Brain Imaging Data Exchange (ABIDE) dataset.

Site	ACC%/AUC%	Site	ACC%/AUC%
NYU	65.8/67.5	CALTECH	73.0/70.3
UM_1	66.9/67.8	SDSU	77.7/81.3
USM	67.3/69.7	OLIN	75.6/71.9
UCLA_1	70.5/80.2	UM_2	67.3/65.0
PITT	70.5/68.5	LEUVEN_2	64.7/68.1
MAX_MUN	63.3/64.3	SBL	65.5/74.7
YALE	75.7/85.7	LEUVEN_1	53.5/60.9
KKI	75.8/84.5	OHSU	64.9/80.6
TRINITY	65.4/63.3	UCLA_2	76.4/83.0
STANFORD	68.6/79.3	CMU	60.5/79.7

FGDN model. Finally, the discriminative brain regions for ASD classification will be extracted for analysis.

In the left subfigure of **Figure 3**, the training(red)/validation(blue) loss-vs.-epochs curves of the

10 models in the CV-10 experiments are presented. During the training procedure, the model was trained to minimize its training loss. In the meanwhile, the validation loss were also decreased. Likewise, the training/validation curves



of accuracy and AUC scores are present in the middle and right subfigures of **Figure 3**. The averaged training curve (red) and validation curve (blue) of accuracy/AUC in 10-fold CV experiments are presented. As the model was optimized by the training samples, the validation accuracy/AUC scores increased. In practice, we will separate 10% training samples as monitoring set to conduct early stopping and avoid overfitting. These curves demonstrated the effectiveness of the proposed method to model and discriminate ASD data.

As the increasing ratio of training set used, the corresponding performance of FGDN is presented in **Figure 4**. We changed the number of samples used for training the model (from 10 to 100% of the training set in CV-10), while the size of validation set (one-fold of CV-10) was unchanged. If the proposed model could work well with a relative small training set, it will be promising for application. With only 20% of training set, the model can achieve a performance better than chance level. With the increase in training samples, the accuracy or AUC also increases. The positive slope indicates that the addition of new training samples substantially improves performance.

Inter-site classification performance is also significant for the techniques in neuroscience (Dvornek et al., 2017; Dadi et al., 2019). Likewise, one of the challenging problems in identifying ASD patients lie in the individual differences of functional connectivity as well as acquisition protocol differences. In the experiment, leave-one-site-out strategy was applied that the samples of the studied site were used for testing while the samples of the rest sites were applied for training. And we used MODL atlas and tangent features for analysis. As shown in **Table 3**, we presented the performance of FGDN for ASD classification of all 20 sites. The proposed FGDN model achieved higher than chance level performance on each site with accuracies from 53.5 to 77.7, AUCs from 60.9 to 85.7. These results demonstrated the effectiveness of the proposed model to overcome different MRI facilities and different MRI settings for inter-site ASD classification.

According to our experiments and the previous research (Dadi et al., 2019), the MODL atlas is promising for ASD classification performance. As a result, we would like to study the most discriminative brain region of MODL for ASD classification. In the MODL atlas, there are totally 128 regions that are represented by $R_i (i = 0 \dots 127)$ in the paper. In the experiments, only the features of the studied region will be kept unchanged while the features of the rest 127 regions were clamped to zeros. Then, we fine-tuned the model which were already trained in the previous CV-10 experiments. We considered that the averaged validation accuracy of each MODL region was its discriminative weight for ASD classification. As shown in **Figure 5**, the visualizations of the five most discriminative MODL regions (R82, R8, R80, R56, and R51) for ASD classification are presented. The classification AUC/ACC and the standard deviation (SD) of each region are presented as well. To study the differences in the discriminative power between the individual MODL regions, we applied *t*-test method to compare the ACC/AUC of different regions in CV-10

experiments. The *t*-test results of the five most discriminative brain regions are also presented in **Figure 5**. The $*[x\%]$ of each region indicates that its ACC/AUC is significantly higher than that of the rest $x\%$ regions with *t*-test $p < 0.05$. For example, the R82 region achieved an AUC of 66.3, which is significantly higher than the rest 73% brain regions. According to the R8, R80, and R82 regions, the superior temporal gyrus is the most discriminative region. This is consistent with findings from the neuroscience field that the cortical volume in several subregions of the superior temporal gyrus are abnormal in individuals with autism (Padmanabhan et al., 2017; Khosla et al., 2018; Zhuang et al., 2019). Therefore, it is a possible technique to determine the discriminative brain regions for ASD identification.

The above experiments and analysis of the proposed FGDN model are significant for studying the ASD. The proposed model can be considered as a new tool for ASD identification as well as the other neuroscience researches based on functional connectivity.

5. CONCLUSION

In this paper, an FGDN was proposed for identifying autism spectrum disorder. The FGDN model was built to discriminate functional graphs of ASD patients and health controls. The proposed model achieved a high ASD classification performance on the ABIDE dataset. In addition, inter-site predictions and discriminative brain regions for ASD prediction were investigated with the proposed model. The experiments and analysis demonstrated that the FGDN model is not only an effective model for identify autism patients but also a potential technique for neuroscience research. In the future, we would like to build more efficient networks to model functional connectivity and identify autism patients. Moreover, some new emerging machine learning techniques can also inspire the methodology in neuroscience like autism identification.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://osf.io/hc4md/download>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of South China Normal University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JL proposed the idea, conducted the experiments, and wrote the manuscript. FW provided advices on the research approaches and revised the manuscript. JP and ZW offered important help

on guiding the experiments and analysis methods. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Key R&D Program of Guangdong Province under Grant 2018B030339001, the National Natural Science Foundation of China (Grant Nos. 62006082, 61836003, and 61906019), the Guangdong Natural Science Foundation (Grant Nos.

2021A1515011600, 2020A1515110294, and 2021A1515011853), and Guangzhou Science and Technology Plan Project (Grant No. 202102020877).

ACKNOWLEDGMENTS

We thank Prof. Zhu Liang Yu, Prof. Yuanqing Li, and their group in South China University of Technology. We would also like to thank the editors, reviewers, and editorial staffs who take part in the publication process of this paper.

REFERENCES

- Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., et al. (2017). Deriving reproducible biomarkers from multi-site resting-state data: an Autism-based example. *Neuroimage* 147, 736–745. doi: 10.1016/j.neuroimage.2016.10.045
- Aggarwal, P., and Gupta, A. (2019a). Group-fused multivariate regression modeling for group-level brain networks. *Neurocomputing* 363, 140–148. doi: 10.1016/j.neucom.2019.06.042
- Aggarwal, P., and Gupta, A. (2019b). Multivariate graph learning for detecting aberrant connectivity of dynamic brain networks in autism. *Med. Image Anal.* 56, 11–25. doi: 10.1016/j.media.2019.05.007
- Aggarwal, P., Gupta, A., and Garg, A. (2017). Multivariate brain network graph identification in functional MRI. *Med. Image Anal.* 42, 228–240. doi: 10.1016/j.media.2017.08.007
- Almuqhim, F., and Saeed, F. (2021). ASD-SANet: a sparse autoencoder, and deep-neural network model for detecting autism spectrum disorder (ASD) using fMRI data. *Front. Comput. Neurosci.* 15:27. doi: 10.3389/fncom.2021.654315
- Arslan, S., Ktena, S. I., Makropoulos, A., Robinson, E. C., Rueckert, D., and Parisot, S. (2018). Human brain mapping: a systematic comparison of parcellation methods for the human cerebral cortex. *Neuroimage* 170, 5–30. doi: 10.1016/j.neuroimage.2017.04.014
- Bruna, J., Zaremba, W., Szlam, A., and LeCun, Y. (2013). Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203*.
- Bullmore, E. T., and Bassett, D. S. (2011). Brain graphs: graphical models of the human brain connectome. *Annu. Rev. Clin. Psychol.* 7, 113–140. doi: 10.1146/annurev-clinpsy-040510-143934
- Dadi, K., Rahim, M., Abraham, A., Chyzyk, D., Milham, M., Thirion, B., et al. (2019). Benchmarking functional connectome-based predictive models for resting-state fMRI. *Neuroimage* 192, 115–134. doi: 10.1016/j.neuroimage.2019.02.062
- Defferrard, M., Bresson, X., Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. in: *Advances in Neural Information Processing Systems*, eds Lee, D., Sugiyama, M., Luxburg, U. Guyon, I. Garnett, R. Barcelona: Curran Associates, Inc.. p. 3844–52.
- Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021
- Dodero, L., Minh, H. Q., San Biagio, M., Murino, V., and Sona, D. (2015). “Kernel-based classification for brain connectivity graphs on the Riemannian manifold of positive definite matrices,” in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)* (Brooklyn, NY: IEEE), 42–45. doi: 10.1109/ISBI.2015.7163812
- Dvornek, N. C., Ventola, P., Pelphrey, K. A., and Duncan, J. S. (2017). “Identifying autism from resting-state fMRI using long short-term memory networks,” in *Machine Learning in Medical Imaging*, eds Q. Wang, Y. Shi, H. I. Suk, and K. Suzuki (Cham: Springer International Publishing), 362–370. doi: 10.1007/978-3-319-67389-9_42
- Epalle, T. M., Song, Y., Liu, Z., and Lu, H. (2021). Multi-atlas classification of autism spectrum disorder with hinge loss trained deep architectures: ABIDE I results. *Appl. Soft Comput.* 107:107375. doi: 10.1016/j.asoc.2021.107375
- Even, S. (2011). *Graph Algorithms*. Cambridge University Press. doi: 10.1017/CBO9781139015165
- Fey, M., and Lenssen, J. E. (2019). “Fast graph representation learning with PyTorch Geometric,” in *ICLR Workshop on Representation Learning on Graphs and Manifolds* Honolulu, HI.
- Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., and Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *Neuroimage Clin.* 17, 16–23. doi: 10.1016/j.nicl.2017.08.017
- Hojjati, S. H., Ebrahimzadeh, A., Khazae, A., Babajani-Feremi, A., Initiative, A. D. N., et al. (2017). Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM. *J. Neurosci. Methods* 282, 69–80. doi: 10.1016/j.jneumeth.2017.03.006
- Iidaka, T. (2015). Resting state functional magnetic resonance imaging and neural network classified autism and control. *Cortex* 63, 55–67. doi: 10.1016/j.cortex.2014.08.011
- Karampasi, A., Kakkos, I., Miloulis, S.-T., Zorzos, I., Dimitrakopoulos, G. N., Gkiatis, K., et al. (2020). “A machine learning fMRI approach in the diagnosis of autism,” in *2020 IEEE International Conference on Big Data (Big Data)* (Atlanta, GA: IEEE), 3628–3631. doi: 10.1109/BigData50022.2020.9378453
- Karampasi, A. S., Savva, A. D., Korfiatis, V. C., Kakkos, I., and Matsopoulos, G. K. (2021). Informative biomarkers for autism spectrum disorder diagnosis in functional magnetic resonance imaging data on the default mode network. *Appl. Sci.* 11:6216. doi: 10.3390/app11136216
- Khosla, M., Jamison, K., Kuceyeski, A., and Sabuncu, M. R. (2018). “3D convolutional neural networks for classification of functional connectomes,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, eds D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, and A. Madabhushi (Cham: Springer International Publishing), 137–145. doi: 10.1007/978-3-030-00889-5_16
- Kingma, D., and Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kipf, T. N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
- Kong, Y., Gao, J., Xu, Y., Pan, Y., Wang, J., and Liu, J. (2018). Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier. *Neurocomputing.* 27:07LT02. doi: 10.1016/j.neucom.2018.04.080
- Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). “ImageNet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems* 25, eds F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Elsevier B.V.).
- Ktena, S. I., Parisot, S., Ferrante, E., Rajchl, M., Lee, M., Glocker, B., et al. (2018). Metric learning with spectral graph convolutions on brain connectivity networks. *Neuroimage* 169, 431–442. doi: 10.1016/j.neuroimage.2017.12.052
- Ledoit, O., and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.* 88, 365–411. doi: 10.1016/S0047-259X(03)00096-4
- Li, J., Zhong, Y., Han, J., Ouyang, G., and Liu, H. (2019). Classifying asd children with lstm based on raw videos. *Neurocomputing.* 390, 226–38. doi: 10.1016/j.neucom.2019.05.106

- Linial, N., London E., and Rabinovich, Y. (1995). The geometry of graphs and some of its algorithmic applications. *Combinatorica* 15, 215–245. doi: 10.1007/BF01200757
- Liu, Y., Xu, L., Li, J., Yu, J., and Yu, X. (2020). Attentional connectivity-based prediction of autism using heterogeneous rs-fMRI data from CC200 atlas. *Exp. Neurobiol.* 29, 27–37. doi: 10.5607/en.2020.29.1.27
- Lord, C., Risi, S., Lambrecht, L., Cook, E., Leventhal, B., DiLavore, P., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* 30, 205–223. doi: 10.1023/A:1005592401947
- Luca, M. D., Beckmann, C. F., Stefano, N. D., Matthews, P. M., and Smith, S. M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage* 29, 1359–1367. doi: 10.1016/j.neuroimage.2005.08.035
- Martino, A. D., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667. doi: 10.1038/mp.2013.78
- Nielsen, J. A., Zielinski, B. A., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., et al. (2013). Multisite functional connectivity MRI classification of autism: ABIDE results. *Front. Hum. Neurosci.* 7:599. doi: 10.3389/fnhum.2013.00599
- Niepert, M., Ahmed, M., and Kutzkov, K. (2016). “Learning convolutional neural networks for graphs,” in *International Conference on Machine Learning* (New York, NY), 2014–2023.
- Padmanabhan, A., Lynch, C., Schaer, M., and Menon, V. (2017). The default mode network in autism. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 2, 476–486. doi: 10.1016/j.bpsc.2017.04.004
- Parisot, S., Ktena, S. I., Ferrante, E., Lee, M., Guerrero, R., Glocker, B., et al. (2018). Disease prediction using graph convolutional networks: application to Autism Spectrum Disorder and Alzheimer’s disease. *Med. Image Anal.* 48, 117–130. doi: 10.1016/j.media.2018.06.001
- Rane, S., Jolly, E., Park, A., Jang, H., and Craddock, C. (2017). Developing predictive imaging biomarkers using whole-brain classifiers: application to the ABIDE I dataset. *Res. Ideas Outcomes* 3:e12733. doi: 10.3897/rio.3.e12733
- Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* 30, 83–98. doi: 10.1109/MSP.2012.2235192
- Song, T., Zheng, W., Song, P., and Cui, Z. (2019). EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.* 11, 532–41. doi: 10.1109/TAFFC.2018.2817622
- Sporns, O. (2011). The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125. doi: 10.1111/j.1749-6632.2010.05888.x
- Stam, C. J., De Haan, W., Daffertshofer, A., Jones, B. F., Manshanden, I., Van Cappellen van Walsum, A. M., et al. (2008). Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer’s disease. *Brain* 132, 213–224. doi: 10.1093/brain/awn262
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Wang, J., Zuo, X., and He, Y. (2010). Graph-based network analysis of resting-state functional MRI. *Front. Syst. Neurosci.* 4:16. doi: 10.3389/fnsys.2010.00016
- Wu, X., Yao, L., Long, Z. Y., Lu, J., and Li, K. C. (2006). “Functional connectivity in the resting brain: an analysis based on ICA,” in *13 International Conference on Neural Information Processing*. Hong Kong. doi: 10.1007/11893028_20
- Yang, X., Islam, M. S., and Khaled, A. M. A. (2019). “Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite ABIDE dataset,” in *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (Chicago, IL: IEEE), 1–4. doi: 10.1109/BHI.2019.8834653
- Yao, D., Liu, M., Wang, M., Lian, C., Wei, J., Sun, L., et al. (2019). “Triplet graph convolutional network for multi-scale analysis of functional connectivity using functional MRI,” in *International Workshop on Graph Learning in Medical Imaging* (Shenzhen: Springer), 70–78. doi: 10.1007/978-3-030-35817-4_9
- Yu, B., Yin, H., and Zhu, Z. (2017). Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*. doi: 10.24963/ijcai.2018/505
- Zhuang, J., Dvornek, N. C., Li, X., Ventola, P., and Duncan, J. S. (2019). Invertible network for classification and biomarker selection for ASD. *arXiv preprint arXiv:1907.09729*. doi: 10.1007/978-3-030-32248-9_78

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Li, Wang, Pan and Wen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.