Inter-Intra High-Order Brain Network for ASD Diagnosis via Functional MRIs

Xiangmin Han[1], Rundong Xue[2], Shaoyi Du[2], and Yue Gao[1(✉)]

[1] Tsinghua University, 100084, China
{hanxiangmin, gaoyue}@tsinghua.edu.cn
[2] Xi'an Jiaotong University, 710049, China
2438527391@qq.com, dushaoyi@xjtu.edu.cn

Abstract. Currently in the field of computer-aided diagnosis, graph or hypergraph-based methods are widely used in the diagnosis of neurological diseases. However, existing graph-based work primarily focuses on pairwise correlations, neglecting high-order correlations. Additionally, existing hypergraph methods can only explore the commonality of high-order representations at a single scale, resulting in the lack of a framework that can integrate multi-scale high-order correlations. To address the above issues, we propose an Inter-Intra High-order Brain Network (I^2HBN) framework for ASD-assisted diagnosis, which is divided into two parts: intra-hypergraph computation and inter-hypergraph computation. Specifically, the intra-hypergraph computation employs the hypergraph to represent high-order correlations among different brain regions based on fMRI signal, generating intra-embeddings and intra-results. Subsequently, inter-hypergraph computation utilizes these intra-embeddings as features of inter-vertices to model inter-hypergraph that captures the inter-correlations among individuals at the population level. Finally, the intra-results and the inter-results are weighted to perform brain disease diagnosis. We demonstrate the potential of this method on two ABIDE datasets (NYU and UCLA), the results show that the proposed method for ASD diagnosis has superior performance, compared with existing state-of-the-art methods.

Keywords: Brain function network · Inter-intra correlation · High-order correlations · Hypergraph computation

1 Introduction

Autism spectrum disorder (ASD) is a typical neurodevelopmental disorder of the brain. The behavioral and cognitive deficits observed in individuals with ASD are closely linked to abnormal functional brain connectivity [3,12], referred to as intra-correlations. Within ASD cohorts, significant inter-correlations exist among individuals. For example, the social communication difficulties in individuals with ASD were strongly correlated with abnormalities in brain connectivity patterns. Additionally, the repetitive behaviors in ASD patients are often linked to specific neural circuitry disruptions, indicating a consistent neurobiological

basis for these behaviors across different individuals. Thus to accurately diagnose ASD patients, a framework is needed to represent both intra-correlations and inter-correlations simultaneously[6].

With advances in technology, imaging techniques (*e.g.*, fMRI) provide neurological measurements for exploring the interactions among brain regions, enabling a better understanding of the pathological basis of ASD. There has been much effort for ASD diagnosis based on fMRI. For fMRI data classification, graph or hypergraph-based methods can be divided into two types: one is to use the brain region signals of each individual to model the brain functional network for classification tasks [18,8,2], and the other is to use the medical information among individuals to model the population-level graph or hypergraph structure for further node classification tasks [9,13,20]. However, these methods fail to integrate the high-order correlation at individual and population levels into a unified framework.

Recently, some studies have proposed a joint framework that combines the individual graph model with the population graph model. GiG [17,21] is a graph-in-graph neural network that samples intra-graphs as vertices and links each pair of these intra-graphs to generate the population graph. However, GIG is a graph-based method that can only represent pairwise intra- or inter-correlations and ignores their high-order correlations, but such high-order correlations may be significant for understanding the pathological basis of brain diseases.

In this paper, we propose an inter-intra high-order brain network (I^2HBN) framework for ASD diagnosis, which models the intra-hypergraph at the individual level and inter-hypergraph at the population level. For each subject, the intra-module treats brain regions as vertices to model high-order correlations among different brain regions using sparse feature representation. Subsequently, the inter-module uses the generated intra-embeddings as inter-vertex features to construct the inter-hypergraph, capturing the high-order inter-correlations among subjects. Finally, the intra-results and inter-results are fused using weights for ASD diagnosis. The main contributions are summarized as follows:

1) We propose a cross-level hypergraph computation framework that integrates the high-order intra-correlations at the individual level and high-order inter-correlations at the population level into a unified framework.
2) To provide effective individual features for the inter-hypergraph computation module, we propose a learnable representation that fully represents the high-order intra-correlations of an individual and demonstrate the effectiveness of this representation in comparison and ablation experiments.
3) We conduct extensive experiments on the ABIDE dataset and the results indicate that the proposed method has superior classification performance than existing methods.

2 Method

Fig. 1 summarizes the pipeline of our method. It consists of two components: 1) Intra-hypergraph computation models the associations among functional areas

and generates high-order embeddings for each individual. 2) Inter-hypergraph computation models the associations among individuals based on intra-embeddings. The final result is calculated by weighting the intra- and inter-result.

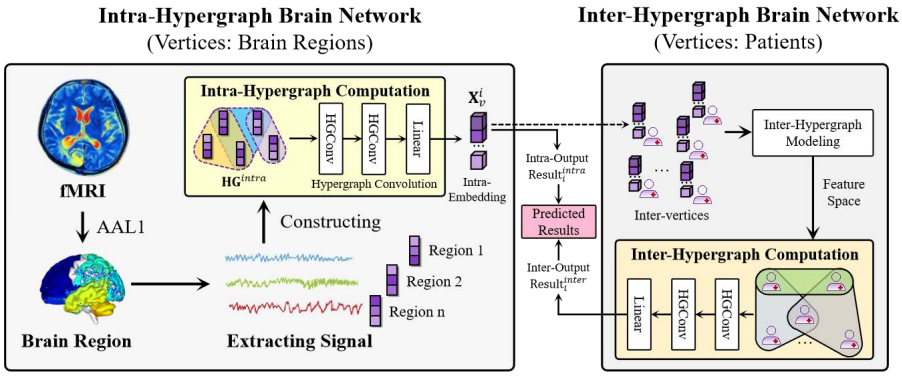


Fig. 1. The pipeline of the proposed I^{2}HBN

2.1 Data Pre-Processing

Data pre-processing is performed using Data Processing Assistant for Resting-State fMRI (DPARSF[3]). Image preprocessing consists of: (1) slice timing correction; (2) head motion correction; (3) T1 alignment to functional image space; (4) nuisance covariate regression; (5) spatial normalization to the MNI template; (6) spatial smoothing with a 6 mm Gaussian kernel; (7) band-pass filtering (0.01 - 0.1 Hz). Individuals with maximum absolute head motion larger than 2 mm or 2° in any of the three translation planes or rotational axes, respectively, need to be eliminated before signal extraction. The scanned brain space of each individual is then parcellated into 116 regions of interest (ROI) based on the Automated Anatomical Labeling (AAL) atlas. Finally, mean time series are obtained by calculating the processed fMRI data voxels within each ROI.

2.2 Intra-Hypergraph Computation

Intra-Hypergraph Modeling In our study, we construct the intra-hypergraph based on fMRI time series using sparse representation [19]. Specifically, for individual R-fMRI images, the data preprocessing is first implemented to generate M time series, where M is the number of ROIs. Each ROI is considered as an intra-vertex and the i^{th} ROI time series x_i is treated as a vertex feature. We denote $\mathbf{X} = [x_1, \ldots, x_i, \ldots, x_M]^\top \in \mathbb{R}^{M \times P}$ as an individual, where P is the length

[3] http://rfmri.org/DPARSF

of time series. Then, each ROI time series (*i.e.*, x_i) is regarded as a response vector and can be linearly represented by other M - 1 ROIs time series as follows:

$$x_i = A_i\boldsymbol{\alpha}_i + \boldsymbol{\tau}_i, \quad i = 1, 2, \ldots, M \tag{1}$$

where $A_i = [x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_M]^\top$ represents the time series of all the ROIs except the i^{th} ROI. $\boldsymbol{\alpha}_i$ is the weight vector quantifying the impact of other ROIs on the i^{th} ROI, and $\boldsymbol{\tau}_i$ is a noise term. Please note that zero elements in $\boldsymbol{\alpha}_i$ indicate that the corresponding ROIs do not contribute to the accurate estimation of the i^{th} ROI time series.

The sparse solution $\boldsymbol{\alpha}_i$ can be approximately recovered by solving a standard l_1-norm regularized optimization problem with the following objective function:

$$\min_{\boldsymbol{\alpha}_i} \parallel x_i - A_i\boldsymbol{\alpha}_i \parallel_2 + \boldsymbol{\lambda} \parallel \boldsymbol{\alpha}_i \parallel_1, \tag{2}$$

where $\boldsymbol{\lambda} > 0$ is a regularization parameter controlling the sparsity of the solution. Larger values of $\boldsymbol{\lambda}$ indicate more regularization, *i.e.*, more elements in α_i are zero. Upon obtaining $\boldsymbol{\alpha}_i$, we eliminate the ROIs with corresponding elements in $\boldsymbol{\alpha}_i$ that are either negative or zero since they have a negative or no effect on the i^{th} ROI. The hyperedge e_i is composed of the i^{th} ROI and the remaining ROIs. In this way, the interactions of the i^{th} ROI with a few other ROIs within the same hyperedge can be obtained, while insignificant and spurious interactions are filtered out. We then apply sparse linear regression to each ROI, ultimately generating M hyperedges to construct the intra-hypergraph $\mathcal{H}^{intra}$ for that individual.

Intra-Hypergraph Learning Given the training individuals set **intra**, we input $\{(\mathbf{X}^{(0)}, \mathcal{H}^{intra}\}$ into the hypergraph convolution layers for training. Hypergraph convolution operation in each layer consists of four stages[4]: 1) Vertex Feature Reweighting. 2) Message Passing from $\mathcal{V}$ to $\mathcal{E}$. 3) Hyperedge Feature Mask. 4) Message Passing from $\mathcal{E}$ to $\mathcal{V}$,

which can be formulated as:

$$\mathbf{X}^{(l+1)} = \delta(\mathbf{H}(\text{Mask}(\mathbf{H}^\top\mathbf{X}^{(l)}\Theta^{(l)}))), \tag{3}$$

where $\delta(\cdot)$ denotes a nonlinear activation function, $\mathbf{H}$ represents the incidence matrix of the hypergraph $\mathcal{H}$, Mask$(\cdot)$ denotes the Hyperedge Feature Mask operation and $\Theta^{(l)}$ denotes a learnable parameter.

After passing several hypergraph convolution layers, the learned vertex features $\mathbf{X}^{(l+1)}$ are obtained. The embedding(pooling) layer converts the learned features into intra-embeddings $e \in \mathbb{R}^{1\times C}$ that represent the corresponding intra-hypergraph features, *i.e.*, individual vectorization, where C denotes the embedding scale. Finally, the result Result^{intra} via MLP is used for ASD diagnosis.

2.3 Inter-Hypergraph Computation

Inter-Hypergraph Modeling. Inter-hypergraph computation treats individuals as inter-vertices to model the high-order correlations among individuals,

and the intra-embeddings $e \in \mathbb{R}^{1\times C}$ learned through intra-hypergraph as initial features of the corresponding vertices. We combine the embedding of each vertex into $\mathbf{E} = [e_1, \ldots, e_N] \in \mathbb{R}^{N\times C}$, where N denotes the number of individuals. An inter-hyperedge is composed of a vertex and its neighbors. The neighbors are determined by calculating the Euclidean distance between initial features (*i.e.* intra-embeddings) of different vertices. The Euclidean distance between any two vertices v_i, v_j can be calculated by Eq. (4).

$$d(v_i, v_j) = (\sum_{k=1}^{C}(e_{v_i}[k] - e_{v_j}[k])^2)^{\frac{1}{2}}, \tag{4}$$

where E_{v_i} and E_{v_j} denote the initial features of vertices v_i and v_j respectively. Then, we scale the $d(v_i, v_j)$ to $[0, 1]$ according to Eq. (5)

$$d_{\text{norm}}(v_i, v_j) = d(v_i, v_j) / \max_{j=1}^{N}(d(v_i, v_j)), \tag{5}$$

where $d(v_i)_{\max}$ denotes the maximum Euclidean distance from the vertex v_i to other vertices. After that, the set of neighbors of v_i can be calculated as Eq. (6).

$$\text{neigh}_{\text{th}}(v_i) = \{v_j | d_{\text{norm}}(v_i, v_j) < th\}, \tag{6}$$

where $\text{th} \in [0, 1]$ is the threshold. We follow the above steps to obtain the set of neighbors of each vertex, combine a vertex and its neighbors set as a hyperedge and finally generate N hyperedges to construct the inter-hypergraph $\mathcal{H}^{inter}$.

Inter-Hypergraph Learning. $\{\mathbf{E}^{(0)}, \mathcal{H}^{inter}\}$ is fed into inter-hypergraph models. Finally, the MLP outputs the inter-results Result_i^{inter} for all individuals. In order to optimize the utilization of intra- and inter-correlations learned through the inter-intra hypergraph computation, we weight the i^{th} individual's inter-result Result_i^{inter} and intra-result Result_i^{intra} as its final result Result_i^{final} for ASD diagnosis, shown as follows:

$$\text{Result}_i^{final} = \beta \text{Result}_i^{intra} + (1 - \beta)\text{Result}_i^{inter}, \tag{7}$$

where the parameter β is the weight used to measure the equilibrium relationship between Result_i^{intra} and Result_i^{inter}. After testing, the desired result can be obtained when $\beta = 0.5$.

3 Experiments

3.1 Experiment Settings

Datasets and Compared Methods. We conduct experiments on five data centers (*i.e.*, ABIDEI-NYU, ABIDEI-UCLA, ABIDEII-NYU1, ABIDEII-UCLA1 and ABIDEII-UCLA2) in the Autism Brain Imaging Data Exchange (ABIDE) datasets. Considering the relatively small amount of data in one data center, we

combine the data from the NYU and UCLA data centers separately into two new datasets NYU, and UCLA. The NYU dataset contains 127 ASD patients and 135 Typical Controls (TC) and the UCLA includes 78 ASD and 63 TC [4]. Based on the same dataset, we compare our method with 4 representative state-of-the-art methods, including GCN [10], GAT [16], HGNN+ [5] and GiGCN [7].

Implementation Details. We conduct experiments on PyTorch with a 3090 GPU. In our experiments, the training/test data is randomly split by five-fold cross-validation, and the experiment is repeated 5 times using random seeds. Note that the folds of the data split used in our proposed method and compared methods are the same. For training, we adopt a two-stage training approach, i.e., we train the intra module separately and then train the inter module, the two modules are independently supervised with their respective loss. The proposed method utilizes cross-entropy as the loss function in both stages.

The involved hyperparameters in our I^{2}HBN are optimally set as follows. The intra module consists of two hypergraph convolutional layers, followed by concat pooling and two fully connected layers with 16, 2 neurons in each layer. Dropout is set to 0.5, the optimizer is Adam, with a learning rate of 1e-5, 5e-4 weight decay and 100 epochs. In the inter module, the learning rate is set to 1e-2. The methods compared in this paper follow the training protocols mentioned in the cited literature. The graph methods use validated features: graph edge weight matrices as node features. The input features for hypergraph methods and our method are the same, all being fMRI signal information.

Evaluation Metrics. Classification performance is tested using five metrics: accuracy, the area under the subject operating characteristic curve (AUC), F1_score, sensitivity, and specificity. The final results are the combination of mean values and standard deviation of all methods through the five-fold cross-validation.

3.2 Classification Performance Results and Analysis

Table 1 and 2 show the performance metrics of all methods on the NYU and UCLA datasets, our method significantly outperforms the compared methods. Particularly, I^{2}HBN improves at least 5.27%, 4.08%, 5.70% and 3.10% for the compared methods with regard to ACC, AUC, F1, and Spe on all datasets. Although I^{2}HBN does not reach the highest precision in Sen metrics on the UCLA dataset, its Spe and Sen are relatively average and high, indicating that I^{2}HBN has a high precision in diagnosing diseased and non-diseased respectively.

I^{2}HBN *vs.* Graph-based Methods. By comparing the classification performance of I^{2}HBN, GCN, and GAT, we can conclude that I^{2}HBN effectively captures the high-order correlations of individual fMRI signals compared to graph-based methods, which ignore these important high-order correlations.

[4] http://fcon_1000.projects.nitrc.org/indi/abide/

Table 1. Experimental results for different methods on NYU datasets.

Methods	ACC	AUC	F1	Spe	Sen
GCN	$63.40_{\pm 5.73}$	$67.60_{\pm 7.52}$	$63.13_{\pm 5.51}$	$59.02_{\pm 3.80}$	$67.35_{\pm 9.98}$
GAT	$60.16_{\pm 4.56}$	$64.04_{\pm 4.13}$	$59.69_{\pm 4.68}$	$62.83_{\pm 4.30}$	$57.21_{\pm 7.43}$
HGNN+	$61.79_{\pm 2.99}$	$64.45_{\pm 5.00}$	$61.33_{\pm 3.10}$	$69.02_{\pm 3.22}$	$53.91_{\pm 4.97}$
GIGCN	$69.84_{\pm 1.44}$	$79.70_{\pm 1.72}$	$69.06_{\pm 1.38}$	$81.40_{\pm 3.67}$	$57.09_{\pm 3.08}$
I^{2}HBN	$\mathbf{76.75}_{\pm 1.96}$	$\mathbf{83.78}_{\pm 1.11}$	$\mathbf{76.41}_{\pm 1.87}$	$\mathbf{84.50}_{\pm 4.25}$	$\mathbf{68.25}_{\pm 1.47}$

Table 2. Experimental results for different methods on UCLA datasets.

Methods	ACC	AUC	F1	Spe	Sen
GCN	$63.59_{\pm 6.68}$	$59.78_{\pm 4.96}$	$62.15_{\pm 7.71}$	$58.26_{\pm 29.96}$	$75.44_{\pm 18.68}$
GAT	$61.73_{\pm 7.23}$	$62.22_{\pm 7.71}$	$61.56_{\pm 7.27}$	$62.70_{\pm 20.99}$	$65.44_{\pm 13.34}$
HGNN+	$67.83_{\pm 7.06}$	$72.12_{\pm 12.6}$	$67.50_{\pm 6.96}$	$74.55_{\pm 15.64}$	$61.67_{\pm 8.50}$
GIGCN	$71.59_{\pm 3.58}$	$77.36_{\pm 2.77}$	$71.05_{\pm 3.80}$	$66.15_{\pm 13.02}$	$\mathbf{76.73}_{\pm 9.85}$
I^{2}HBN	$\mathbf{76.82}_{\pm 3.42}$	$\mathbf{82.37}_{\pm 1.70}$	$\mathbf{76.75}_{\pm 3.50}$	$\mathbf{80.38}_{\pm 4.93}$	$73.45_{\pm 6.86}$

I^{2}HBN *vs.* Hypergraph-based Methods. While hypergraph-based methods (*e.g.*, HGNN+) only represent high-order inter- or intra-correlations individually, I^{2}HBN achieves the integration of the two within a unified framework that considers multi-scale associations to generate more accurate results.

I^{2}HBN *vs.* GiG. In contrast, I^{2}HBN exhibits superior performance. Although the GiG considers the unique individual information and common information across individuals, it focuses only on pairwise associations, ignoring high-order associations. Additionally, GIG is an end-to-end model, and the embeddings generated by the IGC module in the early stages of training cannot effectively display the characteristics of the individual. Conversely, I^{2}HBN adopts a two-phase training strategy to ensure the effectiveness of the set of high-order embeddings of the inter module at the early stage of training, which can better improve the structural modeling of the inter-hypergraph and the node information transfer. The results also prove the effectiveness of this training strategy.

3.3 Ablation Study

We perform ablation experiments on the NYU and UCLA datasets to evaluate the respective contribution of the I^{2}HBN method in the inter and intra stages. The experiments are conducted in three ways: ASD diagnosis using only the results of intra-Hypergraph Computation (namely Intra), ASD diagnosis using only the intra-Hypergraph Computation, and combining the two stages (namely Inter-intra). Note that for a comprehensive and fair evaluation, we conduct two experiments in inter single stage: using intra-embedding (namely $\text{Inter}_{embedding}$) and degrees (namely Inter_{degree}) that reflect the signal characteristics of the brain regions as the vertex features respectively. The results of the ablation experiments on NYU and UCLA datasets are shown in Table 3 and Table 4.

Table 3. Ablation study results on NYU datasets.

Stage	ACC	AUC	F1	Spe	Sen
Intra	$61.79_{\pm 2.99}$	$64.45_{\pm 5.00}$	$61.33_{\pm 3.10}$	$69.02_{\pm 3.22}$	$53.01_{\pm 4.97}$
Inter_{degree}	$54.87_{\pm 2.23}$	$58.08_{\pm 6.11}$	$48.15_{\pm 3.98}$	$\mathbf{85.32}_{\pm 8.16}$	$21.34_{\pm 8.82}$
$\text{Inter}_{embedding}$	$71.79_{\pm 3.14}$	$81.38_{\pm 2.86}$	$71.56_{\pm 3.46}$	$68.06_{\pm 10.96}$	$\mathbf{79.50}_{\pm 5.50}$
Inter-Intra	$\mathbf{76.75}_{\pm 1.96}$	$\mathbf{83.78}_{\pm 1.11}$	$\mathbf{76.41}_{\pm 1.87}$	$84.50_{\pm 4.25}$	$68.25_{\pm 1.47}$

Table 4. Ablation study results on UCLA datasets.

Stage	ACC	AUC	F1	Spe	Sen
Intra	$61.69_{\pm 6.20}$	$61.55_{\pm 9,66}$	$61.44_{\pm 6.49}$	$57.64_{\pm 10.15}$	$65.45_{\pm 3.64}$
Inter_{degree}	$54.11_{\pm 2.23}$	$53.12_{\pm 11.80}$	$53.26_{\pm 10.14}$	$48.18_{\pm 12.57}$	$60.00_{\pm 12.58}$
$\text{Inter}_{embedding}$	$68.41_{\pm 9.56}$	$69.44_{\pm 14.81}$	$65.59_{\pm 13.52}$	$53.85_{\pm 26.26}$	$\mathbf{82.18}_{\pm 8.32}$
Inter-Intra	$\mathbf{76.82}_{\pm 3.42}$	$\mathbf{82.37}_{\pm 1.70}$	$\mathbf{76.75}_{\pm 3.50}$	$\mathbf{80.38}_{\pm 4.93}$	$73.45_{\pm 6.86}$

Intra *vs.* Inter_{degree}: The superior classification performance demonstrated by the Intra stage suggests that Intra generates more accurate representations of ASD than Inter_{degree} by analyzing high-order associations among brain regions. However, Inter_{degree} makes it difficult to capture associations among individuals using only fMRI signals, which further illustrates the necessity of Intra stage.

$\text{Inter}_{embedding}$ *vs.* Intra: The performance of $\text{Inter}_{embedding}$ is further improved based on the intra-embedding, which shows that the complex interactions among individuals discovered by $\text{Inter}_{embedding}$ stage are helpful for ASD diagnosis. This demonstrates the superiority of the Inter-Hypergraph Computation.

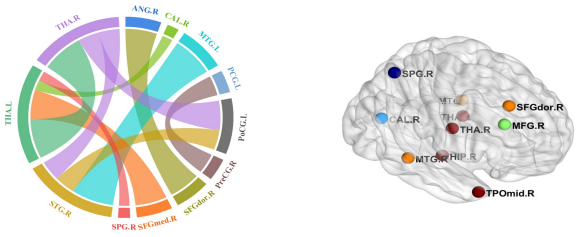
$\text{Inter}_{embedding}$ *vs.* Inter_{degree}: The accuracy of $\text{Inter}_{embedding}$ is higher than Inter_{degree} in the disease diagnosis task, which indicates that the learnable embedding is better at capturing and representing high-order intra-correlations within individuals compared to the initial fMRI signal features.

Inter-intra *vs.* $\text{Inter}_{embedding}$: The evaluation metrics of the Inter-intra are superior to $\text{Inter}_{embedding}$, due to the fact that the $\text{Inter}_{embedding}$ stage only represents high-order inter-correlations based on intra-embedding. In contrast, Inter-intra integrates intra-correlations among brain regions and inter-correlations among individuals, thereby obtaining a reliable result.

3.4 Most Discriminative Functional Connectivity (FC) and ROIs.

Revealing the most discriminative FC and ROIs facilitates pathological analysis. Specifically, functional connectivity networks are constructed by calculating the corresponding adjacency matrix based on the intra-hypergraph of each individual, followed by the use of standard t-test and Lasso to assess the significance of each FC. Moreover, we identify the top 10 ROIs with the highest occurrence frequency in those important FCs with p-value less than 0.05. Figure 2 shows the 10 most discriminative FCs and ROIs selected, where the thickness of the line

represents the discriminative power of the FC. We find that these FCs and ROIs are primarily located in the frontal, temporal, parietal, and Thalamus, which is consistent with the conclusions reached in previous studies. [14,1,15,11].



(a) Discriminative FCs (b) Discriminative ROIs

Fig. 2. Most Discriminative Functional Connectivity (FC) and ROIs.

4 Conclusion

This study proposes a new hypergraph brain functional network for ASD diagnosis, which achieves the integration of information at multiple scales. The framework consists of two parts: the intra-hypergraph represents the intra-correlations of the individual scale, and the inter-hypergraph represents the inter-correlations of the population scale. The intra-correlations of an individual are efficiently represented by proposing a learnable embedding. Experimental results on the ABIDE dataset show that the classification performance of the proposed method outperforms the state-of-the-art methods for ASD diagnosis.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abraham, A., Milham, M.P., Di Martino, A., Craddock, R.C., Samaras, D., Thirion, B., Varoquaux, G.: Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. NeuroImage **147**, 736–745 (2017)
2. Cui, H., Dai, W., Zhu, Y., Kan, X., Gu, A.A.C., Lukemire, J., Zhan, L., He, L., Guo, Y., Yang, C.: Braingb: A benchmark for brain network analysis with graph neural networks. IEEE transactions on medical imaging **42**(2), 493–506 (2022)
3. Farahani, F.V., Karwowski, W., Lighthall, N.R.: Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review. frontiers in Neuroscience **13**, 585 (2019)
4. Feng, Y., You, H., Zhang, Z., Ji, R., Gao, Y.: Hypergraph neural networks. In: Proceedings of the AAAI conference on artificial intelligence. pp. 3558–3565 (2019)

5. Gao, Y., Feng, Y., Ji, S., Ji, R.: HGNN+: General hypergraph neural networks. IEEE Trans. Pattern Anal. Mach. Intell. **45**(3), 3181–3199 (2022)
6. Gao, Y., Ji, S., Han, X., Dai, Q.: Hypergraph computation. Engineering (2024)
7. Jia, S., Jiang, S., Zhang, S., Xu, M., Jia, X.: Graph-in-graph convolutional network for hyperspectral image classification. IEEE Trans. Neural Networks Learn. Syst. (2022)
8. Jie, B., Wee, C.Y., Shen, D., Zhang, D.: Hyper-connectivity of functional networks for brain disease diagnosis. Medical image analysis **32**, 84–100 (2016)
9. Kazi, A., Shekarforoush, S., Arvind Krishna, S., Burwinkel, H., Vivar, G., Kortüm, K., Ahmadi, S.A., Albarqouni, S., Navab, N.: Inceptiongcn: receptive field aware graph convolutional network for disease prediction. In: Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26. pp. 73–85. Springer (2019)
10. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: ICLR (2017)
11. Li, X., Gu, Y., Dvornek, N., Staib, L.H., Ventola, P., Duncan, J.S.: Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: Abide results. Medical Image Analysis **65**, 101765 (2020)
12. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L.H., Ventola, P., Duncan, J.S.: Braingnn: Interpretable brain graph neural network for fmri analysis. Medical Image Analysis **74**, 102233 (2021)
13. Parisot, S., Ktena, S.I., Ferrante, E., Lee, M., Moreno, R.G., Glocker, B., Rueckert, D.: Spectral graph convolutions for population-based disease prediction. In: Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20. pp. 177–185. Springer (2017)
14. Shahamat, H., Abadeh, M.S.: Brain mri analysis using a deep learning based evolutionary approach. Neural Networks **126**, 218–234 (2020)
15. Turner, A.H., Greenspan, K.S., van Erp, T.G.: Pallidum and lateral ventricle volume enlargement in autism spectrum disorder. Psychiatry Research: Neuroimaging **252**, 40–45 (2016)
16. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. In: ICLR (2017)
17. Wang, J., Song, S., Chen, L., Yang, J., Deng, J., Gunes, H.: Graph in Graph neural network. In: ICLR (2023)
18. Wei, L., Liu, B., He, J., Zhang, M., Huang, Y.: Autistic spectrum disorders diagnose with graph neural networks. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 8819–8827 (2023)
19. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. IEEE transactions on pattern analysis and machine intelligence **31**(2), 210–227 (2008)
20. Xiao, L., Wang, J., Kassani, P.H., Zhang, Y., Bai, Y., Stephen, J.M., Wilson, T.W., Calhoun, V.D., Wang, Y.P.: Multi-hypergraph learning-based brain functional connectivity analysis in fmri data. IEEE transactions on medical imaging **39**(5), 1746–1758 (2019)
21. Zaripova, K., Cosmo, L., Kazi, A., Ahmadi, S.A., Bronstein, M.M., Navab, N.: Graph-in-Graph (GiG): Learning interpretable latent graphs in non-euclidean domain for biological and healthcare applications. Med. Image Anal. **88**, 102839 (2023)