

Recognition of Autism Spectrum Disorder via Dual-Channel Feature Fusion and Multi-Resolution Brain Region Segmentation

Lei Chen

College of Information Engineering
Shanghai Maritime University
Shanghai, China
chen_sanshi@163.com

Weiming Zeng

College of Information Engineering
Shanghai Maritime University
Shanghai, China
zengwm86@163.com

Wenhao Dong

College of Information Engineering
Shanghai Maritime University
Shanghai, China
jsxzdwh@163.com

Yueyang Li*

College of Information Engineering
Shanghai Maritime University
Shanghai, China
*lyy20010615@163.com

Abstract—Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by disruptions in brain functional networks. Current classification studies using functional magnetic resonance imaging (fMRI) often struggle with issues of model interpretability and incorporating multi-scale features effectively. This study proposes a multi-resolution brain network model that integrates degree-normalized graph neural networks with fully connected attention mechanisms to boost classification accuracy and identify critical brain networks related to ASD. Experimental results show that the proposed model outperforms traditional methods in accuracy, sensitivity, and other evaluation metrics. Furthermore, analysis of brain network correlations reveals functional networks closely associated with ASD, providing valuable insights and a strong reference for clinical diagnosis.

Keywords—Degree normalization; Multi-resolution partitioning; Full attention; Autism Spectrum Disorder

I. INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by persistent difficulties in social interaction, communication, and behavior, presenting significant challenges for both diagnosis and intervention [1]. Traditional machine learning methods, such as support vector machines (SVM) [2], ridge classifiers [3], and logistic regression [4], have been widely applied to ASD classification using functional magnetic resonance imaging (fMRI) data. These approaches aim to improve diagnostic accuracy by extracting features related to brain function and structure.

With the advancement of deep learning, recent studies leveraging models such as long short-term memory networks (LSTM) [5], graph neural networks (GNN) [6], and Transformers [7] have emerged, employing complex brain connectivity patterns to enhance ASD classification performance. However, these foundational models have certain limitations, such as difficulty in fully capturing the intricate hierarchical structures of brain networks and susceptibility to

overfitting due to their complexity. Fortunately, some recent studies have attempted to address these issues. Recent studies have advanced GNN architectures for fMRI analysis by introducing models like the spatio-temporal hybrid attentive graph network (ST-HAG), which employs Transformer-based attention for temporal modeling [8]. Further advancements include the explainable and generalizable graph neural network (XG-GNN) [9], which employs meta-learning to enhance model generalization. While these methods have achieved improvements in accuracy, they often lack interpretability, particularly in identifying key brain networks associated with ASD. Moreover, they frequently fail to fully leverage the multi-scale structural information of the brain, limiting the comprehensiveness of feature extraction.

To address these limitations, this study proposes a multi-resolution brain region segmentation model based on GNN-DN and Transformer-FA modules. The proposed model demonstrates high accuracy in ASD classification and, through interpretability analysis, identifies brain functional networks closely associated with ASD [10]. This work provides robust support for further elucidating the neural mechanisms underlying ASD.

II. METHODS

Figure 1 illustrates the overall workflow of the model developed in this study, detailing the primary steps from multi-resolution brain region segmentation to functional network construction and feature extraction.

A. Multi-Resolution Brain Region Segmentation

This study employed the Schaefer2018 [11] template to segment the brain into 1000, 400, and 100 regions of interest (ROIs), corresponding to three levels of resolution: microscopic, mesoscale, and macroscopic. This process is illustrated in Figure 1(a).

After obtaining the ROI time series at different resolutions, functional connectivity graphs were subsequently constructed. First, based on the widely used Yeo 17 [12] networks, the time

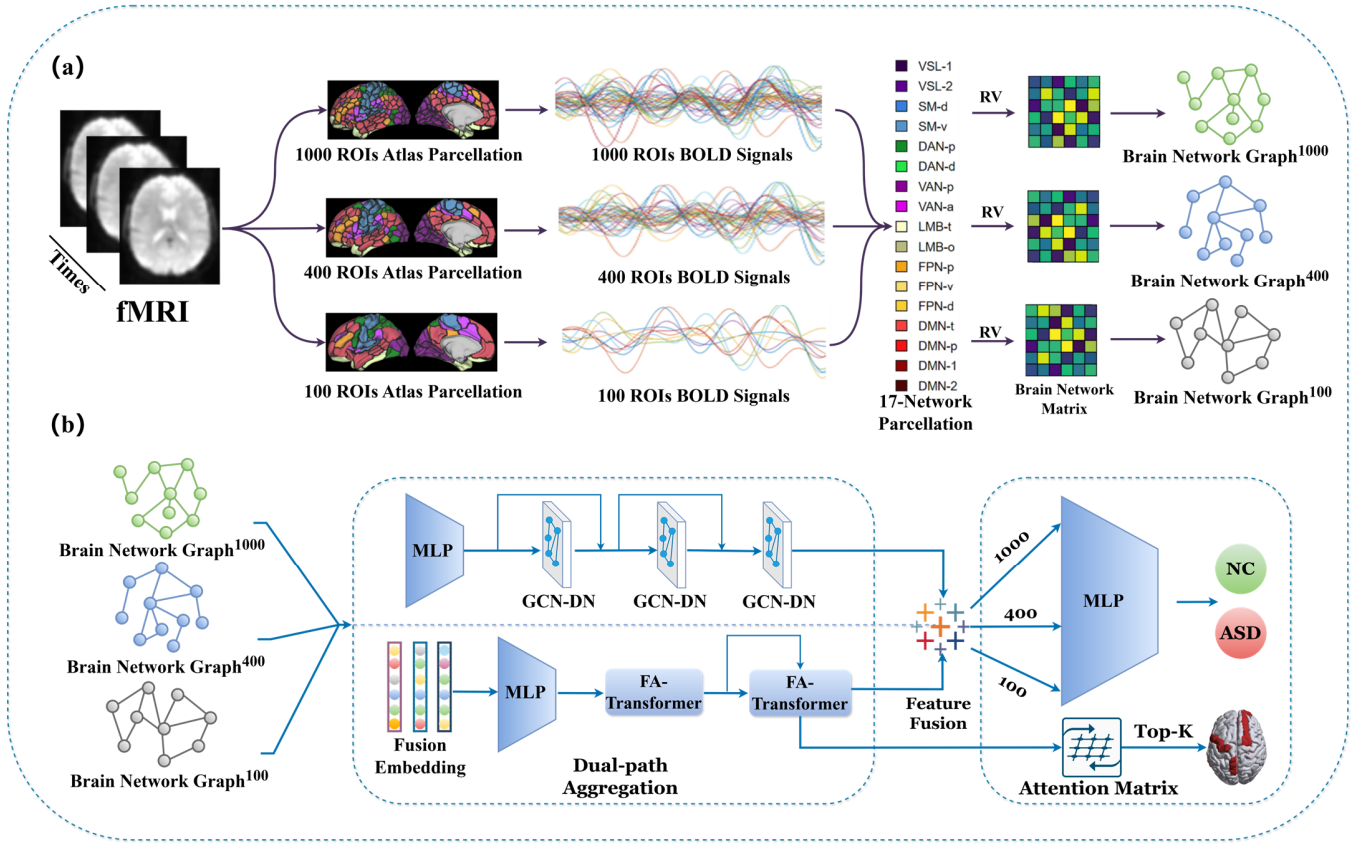


Figure 1. The overall process of the research model, including (a) multi-resolution brain region segmentation and (b) feature extraction and fusion.

series of all ROIs within each of the 17 brain functional networks were extracted and vertically concatenated into matrices to represent the overall activity patterns of the functional networks. Next, the RV coefficient [13] was employed to calculate the similarity between different functional network matrices, defined mathematically as:

$$RV = \frac{\text{tr}(S_i^T S_j)}{\sqrt{\text{tr}(S_i^T S_i) \text{tr}(S_j^T S_j)}}. \quad (1)$$

The numerator represents the structure of the matrix vector spaces, while the denominator serves as a normalization factor to ensure that the RV coefficient value falls within the range $[0, 1]$.

Subsequently, a 17×17 correlation matrix is generated, with each element indicating the strength of similarity between functional networks. Based on this correlation matrix, the graph structure of the functional network is constructed by treating each functional network as a node. To identify the most significant functional connections, a threshold is applied, retaining only edges where the similarity exceeds the specified threshold.

B. GNN-DN Module

The GNN-DN module is a core component developed in this study for processing brain functional networks. Building upon the traditional graph neural network, this module incorporates degree normalization and adaptive residual

connections [14].

1) The Weighting Mechanism for Degree Normalization

Significant differences exist in the degree distribution of nodes within brain functional networks. Traditional graph convolution methods predominantly employ Laplacian regularization [15] but often fail to account for the heterogptive weighting mechanism based on node degrees during feature propagation between high-degree and low-degree nodes. To address this issue, the GNN-DN module incorporates an adaptive propagation, ensuring balanced feature distribution. Specifically, for the edge between nodes i and j , the degree-normalized weight W_{ij} is defined as follows:

$$W_{ij} = \frac{1}{\sqrt{d_i d_j}}, \quad (2)$$

where d_i and d_j denote the degrees of nodes i and j , respectively. Based on this weighting mechanism, the feature propagation formula is defined as follows:

$$H^{(l+1)} = \sigma(D^{-1/2} A D^{-1/2} H^{(l)} W^{(l)}), \quad (3)$$

where A represents the adjacency matrix, and D is the degree matrix with diagonal elements $D_i = d_i$, H denotes the node features of the l -th layer, $W^{(l)}$ is the weight matrix of the current layer, and σ is the nonlinear activation function, ensuring consistent feature propagation between high- and low-degree nodes.

2) Residual Connection

To enhance the effectiveness of feature propagation, the GNN-DN module incorporates an adaptive residual connection to mitigate the gradient vanishing problem in multi-layer networks by retaining the initial input feature $H^{(0)}$. After each convolutional layer, the initial feature $H^{(0)}$ is directly added to the output of the current layer, forming an adaptive residual connection. The feature update formula is defined as follows:

$$H^{(l+1)} = H^{(l+1)} + H^{(0)}, \quad (4)$$

$H^{(0)}$ is the initial feature of the input layer.

C. Transformer-FA Module

The Transformer-FA module is designed to capture global feature correlations across regions within the brain functional network. This module integrates a fully connected attention mechanism with an attention-based brain network selection method.

1) Fully Connected Attention Convolution

Cross-regional dependencies [16] exist between nodes in the brain functional network, and interactions between distant brain regions play a critical role in classification. The Transformer-FA module incorporates a fully connected attention mechanism, enabling each node to globally interact with all other nodes, thereby comprehensively capturing cross-regional features.

Given the input feature vectors for query, key, and value, they are denoted as $Q \in R^{N \times H \times M}$, $K \in R^{L \times H \times M}$ and $V \in R^{L \times H \times D}$, respectively, where N represents the number of query nodes, L denotes the number of source nodes, H is the number of attention heads, and M and D are the feature dimensions.

To ensure feature scale consistency, Q and K are first subjected to $L2$ normalization:

$$Q = \frac{Q}{\|Q\|_2}, K = \frac{K}{\|K\|_2}, \quad (5)$$

Next, the weighted sum of the source nodes is computed by multiplying K and V , resulting in an intermediate aggregation tensor KV_a . The computation is defined as follows:

$$KV_a = \sum_{l=1}^L K_l \otimes V_l, \quad (6)$$

The operator $\otimes$ denotes the inner product operation along dimension l . The KV_a generated in this step captures the weighted feature information of the source nodes.

Subsequently, the query vector Q is multiplied by the intermediate aggregation tensor KV_a to produce the attention-weighted aggregation result A_{num} , defined as:

$$A_{\text{num}} = \sum_{n=1}^H Q_n \cdot KV_a, \quad (7)$$

The result A_{num} represents the similarity weight between each query node and the source nodes, which is utilized for subsequent feature aggregation.

To normalize the attention weights, a normalization factor A_{denom} is defined. First, the sum of the key vectors, K_{sum} , is calculated. The query vector is then multiplied by this sum to derive the normalization factor:

$$A_{\text{denom}} = \sum_{n=1}^H Q_n \cdot K_{\text{sum}}, \quad (8)$$

by dividing the attention value A_{num} by the normalization factor A_{denom} , the weighted feature aggregation output A_{out} is obtained. This output represents the weighted feature representation of each query node across different attention heads.

2) Attention-based Brain Network Selection

Through the attention mechanism, source nodes with the most similar features to the query node are assigned higher attention weights. Higher weights indicate that the features of these nodes dominate the aggregated output.

After constructing the attention matrix, the global node importance is derived by summing the columns of the attention matrix. Node pairs with high attention weights typically represent brain region connections with strong correlations in the functional network, thereby identifying brain networks associated with ASD.

III. EXPERIMENT AND DISCUSSION

A. Data Preprocessing

This study utilized fMRI data from the Autism Brain Imaging Data Exchange (ABIDE) dataset [17], which included data from ASD patients and healthy controls (HC). Due to the multi-institutional nature of the dataset, some data contained image artifacts and head motion interference, necessitating a screening process prior to analysis [18]. Samples with severe motion artifacts were excluded, resulting in a final dataset of 871 subjects.

For these 871 subjects, a standardized preprocessing pipeline was applied, which included time slice correction, head motion correction, spatial normalization, and temporal smoothing:

1) *Time slice correction*: Interpolation was used to align the acquisition times of different slices, eliminating temporal differences between slices.

2) *Head motion correction*: Spatial deviations caused by motion were reduced by calculating translations and rotations at each time point and applying inverse transformations.

3) *Spatial normalization*: Image data were registered to the standard MNI152 space to ensure spatial consistency.

4) *Temporal smoothing*: A Gaussian kernel was applied to smooth the time series of each voxel, reducing noise and enhancing signal quality.

B. Experimental Setup and Evaluation Metrics

In the experimental setup, the data of 871 subjects were randomly divided into training, validation, and test sets, with proportions of 70%, 15%, and 15%, respectively. A 10-fold cross-validation method was employed to mitigate result bias

caused by the randomness of data splitting.

During model training and evaluation, standard performance metrics were selected. Accuracy (Acc) reflects overall classification correctness, Precision (Prec) and Recall (Rec) assess the model's performance on the positive class, and AUC measures overall discriminative ability. Collectively, these metrics provide a comprehensive evaluation of the model's effectiveness in ASD classification.

C. Result

1) Module performance comparison

To validate the effectiveness of the GNN-DN and Transformer-FA modules, multiple experiments were conducted using the ABIDE dataset and the Schaefer2018 template. The experimental results are presented in Table 1. As shown, the combined model proposed in this study outperforms all baseline models across various metrics, demonstrating the superior discriminative ability of these modules in identifying ASD patients.

Table 1. Experimental results on ABIDE dataset

Method	Acc	Prec	Rec	AUC
SVM	65.72 ±3.11	54.68 ±6.88	66.38 ±6.79	71.33 ±4.56
LSTM	66.45 ±3.25	65.35 ±8.92	66.71 ±7.52	70.15 ±3.55
SwinT	68.45 ±4.64	60.32 ±7.58	68.36 ±6.78	70.26 ±4.26
BrainGNN	73.86 ±0.57	68.30 ±0.92	73.89 ±1.37	69.94 ±0.31
Ours	78.96 ±0.13	75.46 ±0.14	78.34 ±0.51	74.56 ±0.23

2) Ablation Experiment

To evaluate the contributions of each module and hierarchical brain region division, an ablation experiment was conducted. The GNN-DN module, Transformer-FA module, and hierarchical brain region division were removed individually, and the changes in model performance were observed. The results are presented in Table 2.

Table 2. Ablation Experimental

A	B	C	Acc	Prec	Rec	AUC
√	×	√	74.72 ±1.13	68.48 ±1.85	73.18 ±2.58	69.23 ±1.56
√	√	×	75.45 ±2.25	70.35 ±1.52	74.71 ±1.02	73.65 ±0.55
×	√	√	72.45 ±3.54	65.32 ±5.68	71.53 ±3.28	70.86 ±2.86

The columns labeled A, B, and C in the table represent the following model components: A: Multi-level Parcellation. B: GNN-DN Model. C: Transformer-FA Model.

D. Brain network correlation

This study leverages the attention matrix from the Transformer-FA module to analyze correlations within brain functional networks. Specifically, the fully connected attention mechanism identifies the top six brain network nodes with the highest attention weights. The correlations among these nodes were visualized, as shown in Figure 2, to underscore their importance in ASD recognition.

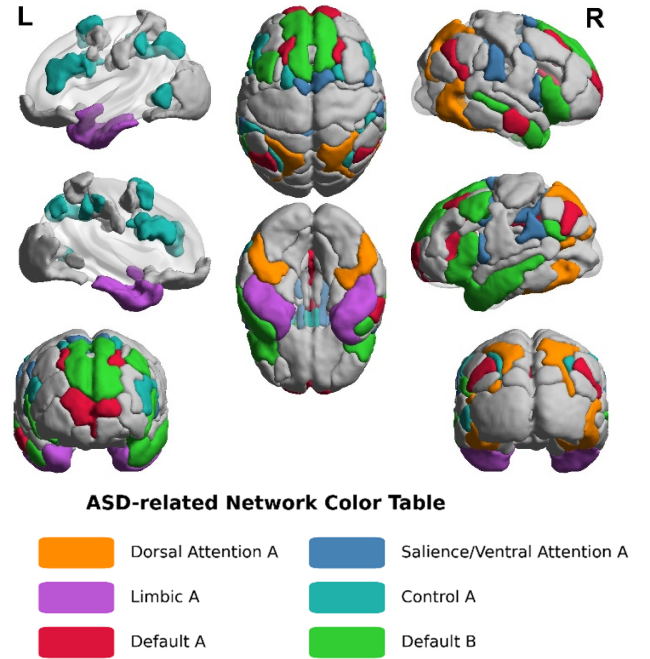


Figure 2. The six most important brain networks for ASD classification

E. Discussion

Experimental results demonstrate that the proposed model achieves excellent performance in ASD classification. The GNN-DN and Transformer-FA modules leverage degree normalization and fully connected attention mechanisms, respectively, to balance feature propagation across nodes, ensuring consistency between high and low degree nodes. These modules enhance global information capture and exhibit strong performance in multi-scale feature integration, effectively capturing both local and global characteristics of brain functional networks. The hierarchical brain region division further supports multi-scale information integration, enabling the model to extract multi-level brain region features at different resolutions and improving ASD recognition accuracy. Ablation experiments validate the independent contributions of each module and the hierarchical brain division, highlighting their synergistic effect on classification performance.

Existing brain partition templates face a trade-off in resolution: high-resolution divisions capture detailed local information but increase computational complexity, while low-resolution divisions may lack sufficient detail for characterizing fine-scale neural activity. The hierarchical brain region division used in this study (1000, 400, and 100) enables multi-resolution feature extraction: high resolution (1000) captures fine-grained connections, medium resolution (400) balances information retention and efficiency, and low resolution (100) facilitates global network analysis. This multi-scale approach provides diverse feature inputs, allowing the GNN-DN and Transformer-FA modules to learn more representative cross-regional features, thereby enhancing classification performance.

However, hierarchical brain region division also incurs additional computational costs, especially at high resolution

(1000). To mitigate this, the degree normalization in the GNN-DN module and the fully connected attention in the Transformer-FA module play key roles in feature extraction and information integration, partially alleviating computational overhead while improving model effectiveness.

In the visualization of brain network correlations, the attention weights extracted by the Transformer-FA module highlight brain regions closely associated with ASD, such as the default mode network (Default), control network (Cont), and salient attention network (SalVentAttn). These subnetworks encompass core functions relevant to ASD, including self-referential processing, social cognition, emotional regulation, attentional control, and executive function. These findings align with existing medical research [19], indicating that the model not only achieves high classification accuracy but also possesses biological interpretability. By identifying brain regions most critical for ASD recognition, this study provides a potential basis for applications in early ASD diagnosis.

IV. CONCLUSIONS

This study proposes a multi-resolution brain region partitioning model based on GNN-DN and Transformer-FA modules for identifying ASD. By leveraging the degree normalization of the GNN-DN module and the fully connected attention mechanism of the Transformer-FA module, the model effectively integrates local and global features within brain functional networks. Additionally, the model successfully identifies key brain networks associated with ASD, aligning with existing medical research and supporting its biological interpretability.

Despite current limitations in the spatial resolution and granularity of brain region analyses, future research could enhance model interpretability by incorporating high-precision brain region coordinates or finer partitioning methods. Overall, the multi-resolution model developed here demonstrates high accuracy and reliability in ASD identification, offering a valuable tool for early ASD screening and diagnosis. Future work may further optimize feature extraction or extend the model's application to other neuropsychological disorders, providing more biologically meaningful diagnostic support for clinical practice.

ACKNOWLEDGMENT

This research was supported by the National Natural Science Foundation of China with grant number 31870979.

REFERENCES

- [1] Li, Y., Zeng, W., Dong, W., Cai, L., Wang, L., Chen, H., ... & Wang, N. (2024). MHNet: Multi-view High-order Network for Diagnosing Neurodevelopmental Disorders Using Resting-state fMRI. arXiv preprint arXiv:2407.03217.
- [2] Morra, J. H., Tu, Z., Apostolova, L. G., Green, A. E., Toga, A. W., & Thompson, P. M. (2009). Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. *IEEE transactions on medical imaging*, 29(1): 30-43.
- [3] Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., & Varoquaux, G. (2017). Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. *NeuroImage*, 147: 736-745.
- [4] Lee, S. H., Yu, D., Bachman, A. H., Lim, J., & Ardekani, B. A. (2014). Application of fused lasso logistic regression to the study of corpus callosum thickness in early Alzheimer's disease. *Journal of neuroscience methods*, 221: 78-84.
- [5] Oota, S. R., Rowtula, V., Gupta, M., & Bapi, R. S. (2019). StepEncog: A convolutional LSTM autoencoder for near-perfect fMRI encoding. In: 2019 International Joint Conference on Neural Networks. Budapest. pp. 1-8.
- [6] Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., ... & Duncan, J. S. (2021). Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74: 102233.
- [7] Kim, P., Kwon, J., Joo, S., Bae, S., Lee, D., Jung, Y., ... & Moon, T. (2023). Swift: Swin 4d fmri transformer. *Advances in Neural Information Processing Systems*, 36: 42015-42037.
- [8] Liu, R., Huang, Z. A., Hu, Y., Huang, L., Wong, K. C., & Tan, K. C. (2024). Spatio-Temporal Hybrid Attentive Graph Network for Diagnosis of Mental Disorders on fMRI Time-Series Data. *IEEE Transactions on Emerging Topics in Computational Intelligence*.
- [9] Qiu, X., Wang, F., Sun, Y., Lian, C., & Ma, J. (2024). Towards Graph Neural Networks with Domain-Generalizable Explainability for fMRI-Based Brain Disorder Diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer Nature Switzerland. pp. 454-464.
- [10] Hernandez, L. M., Rudie, J. D., Green, S. A., Bookheimer, S., & Dapretto, M. (2015). Neural signatures of autism spectrum disorders: insights into brain network dynamics. *Neuropsychopharmacology*, 40(1): 171-189.
- [11] Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., ... & Yeo, B. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral cortex*, 28(9): 3095-3114.
- [12] Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 106: 1125-1165.
- [13] Smilde, A. K., Kiers, H. A., Bijlsma, S., Rubingh, C. M., & Van Erk, M. J. (2009). Matrix correlations for high-dimensional data: the modified RV-coefficient. *Bioinformatics*, 25(3): 401-405.
- [14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas. pp. 770-778.
- [15] Qu, G., Hu, W., Xiao, L., Wang, J., Bai, Y., Patel, B., ... & Wang, Y. P. (2021). Brain functional connectivity analysis via graphical deep learning. *IEEE Transactions on Biomedical Engineering*, 69(5): 1696-1706.
- [16] Stieger, J. R., Pinheiro-Chagas, P., Fang, Y., Li, J., Lusk, Z., Perry, C. M., ... & Parvizi, J. (2024). Cross-regional coordination of activity in the human brain during autobiographical self-referential processing. *Proceedings of the National Academy of Sciences*, 121(32): e2316021121.
- [17] Liu, M., Li, B., & Hu, D. (2021). Autism spectrum disorder studies using fMRI data and machine learning: a review. *Frontiers in Neuroscience*, 15: 697870.
- [18] Yan, C. G., Wang, X. D., Zuo, X. N., & Zang, Y. F. (2016). DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*, 14: 339-351.
- [19] Bicks, L. K., & Geschwind, D. H. (2024). Functional neurogenomics in autism spectrum disorders: A decade of progress. *Current Opinion in Neurobiology*, 86: 102858.