Learning the Hierarchical Organization in Brain Network for Brain Disorder Diagnosis

Jingfeng Tang[1,2], Peng Cao[1,2](✉), Guangqi Wen[3], Jinzhu Yang[1,2], Xiaoli Liu[4], and Osmar R. Zaiane[5]

[1] Computer Science and Engineering, Northeastern University, Shenyang, China
[2] Key Laboratory of Intelligent Computing in Medical Image of Ministry of Education, Northeastern University, Shenyang, China
caopeng@cse.neu.edu.cn
[3] School of Computer Science and Artificial Intelligence, Shandong Normal University, Jinan, China
[4] AiShiWeiLai AI Research, Beijing, China
[5] Amii, University of Alberta, Edmonton, Alberta, Canada

Abstract. Brain network analysis based on functional Magnetic Resonance Imaging (fMRI) is pivotal for diagnosing brain disorders. Existing approaches typically rely on predefined functional sub-networks to construct sub-network associations. However, we identified many cross-network interaction patterns with high Pearson correlations that this strict, prior-based organization fails to capture. To overcome this limitation, we propose the Brain Hierarchical Organization Learning (BrainHO) to learn inherently hierarchical brain network dependencies based on their intrinsic features rather than predefined sub-network labels. Specifically, we design a hierarchical attention mechanism that allows the model to aggregate nodes into a hierarchical organization, effectively capturing intricate connectivity patterns at the subgraph level. To ensure diverse, complementary, and stable organizations, we incorporate an orthogonality constraint loss, alongside a hierarchical consistency constraint strategy, to refine node-level features using high-level graph semantics. Extensive experiments on the publicly available ABIDE and REST-meta-MDD datasets demonstrate that BrainHO not only achieves state-of-the-art classification performance but also uncovers interpretable, clinically significant biomarkers by precisely localizing disease-related sub-networks. Our code is available at https://github.com/Jingfeng-Tang/BrainHO.

Keywords: Brain network analysis · Hierarchical brain organization · Attention · Interpretability · Classification.

1 Introduction

Brain disorders, such as Autism Spectrum Disorder (ASD) and Major Depressive Disorder (MDD), are fundamentally underpinned by the dysfunction and atypical interactions of functional brain sub-networks[12,22]. To better analyze sub-networks, recent studies[1,13]have shifted focus from node-level learning

across the whole-brain network to community(subnetwork)-level learning based on existing predefined functional sub-network atlas, such as the Yeo7[19] and Smith10[17] atlas. Com-TF[1] pioneers the use of the Yeo7[19] atlas to independently model intra- and inter-subnetwork interactions via local and global Transformers. Following this separate modeling paradigm, CAGT[13] further integrates spatial and topological information, while DHGFormer[21] leverages it to enhance GNN performance. Existing approaches implicitly assume strict functional boundaries between predefined sub-networks. By modeling them independently, these methods fail to capture latent cross-network interactions. However, it is worth noting that significant correlations often exist between nodes from different sub-networks, shown in Fig.1(a). For example, the research finding [7] reveals that individuals with ASD exhibit significantly altered functional connectivity in key regions, including the left inferior frontal gyrus (Frontal_Inf_Oper_L) and the left superior temporal gyrus (Temporal_Sup_L). However, in fixed atlases like Yeo7, the Frontal_Inf_Oper_L and Temporal_Sup_L are hard-coded into separate communities (e.g., the VAN and the SMN, respectively). Consequently, the vital cross-network interactions between these regions fail to be captured, which obscures critical pathological features and ultimately degrades classification accuracy. Therefore, exploiting the hierarchical brain organization and establishing a flexible subgraph representation for distinct disorders is crucial for better understanding the underlying dysfunction in brain networks.

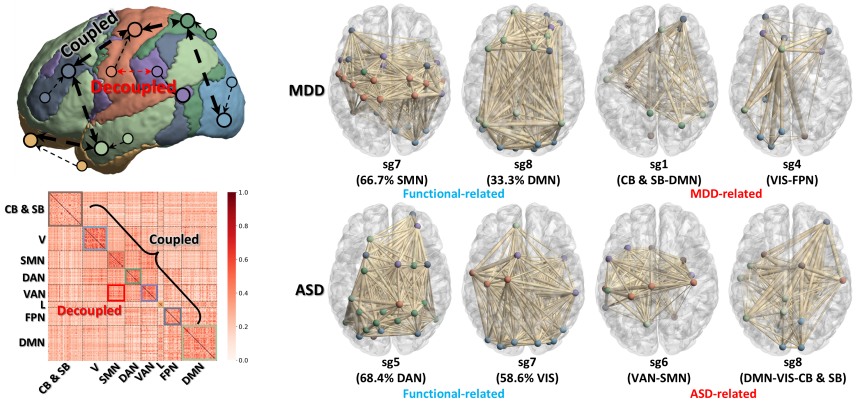


Fig. 1. (a) Average Pearson correlation between brain regions on the ABIDE dataset, revealing strong interactions of brain regions across predefined sub-networks. (b) Learned sub-networks on the ABIDE and REST-meta-MDD datasets. Derived from attention weights, these visualizations demonstrate that BrainHO uncovers disease-related sub-networks that span predefined sub-networks(identified via subgraph to graph attention weights) as well as functionally consistent sub-networks(determined by the proportion of functional sub-networks).

To address this limitation, we propose Brain Hierarchical Organization Learning (BrainHO) for modeling intrinsically hierarchical modular organization, which circumvents the strict functional constraints, thereby emulating the human brain's underlying mechanisms[10]. Different from prior works[1] that rely on the strict parcellation of the brain atlas, we introduce learnable subgraph tokens to dynamically aggregate information from the brain network based on node feature affinity rather than predefined labels. To ensure that the learnable subgraphs capture diverse, complementary and stable organizations, we introduce an orthogonality constraint to encourage diverse connectivity patterns, and a hierarchical consistency strategy to refine node-level features using high-level graph semantics.

We summarize our main contributions as follows:

- We propose BrainHO, a framework that breaks the limits of fixed functional atlases by modeling the hierarchical organization of brain networks in a learnable manner. This approach captures complex cross-network interactions, thereby providing interpretable insights into the underlying mechanisms of brain dysfunctions.
- We design a hierarchical attention mechanism with an orthogonality constraint to learn diverse subgraphs, coupled with a hierarchical consistency constraint mechanism to refine local node features for being highly disease-discriminative.
- Extensive experiments on the ABIDE and REST-meta-MDD datasets demonstrate that BrainHO achieves state-of-the-art performance and validates the clinical significance of the disease-related sub-networks learned by BrainHO.

2 Method

2.1 Overview

As illustrated in Figure 2, BrainHO first projects the fMRI-derived Pearson Correlation Coefficient (PCC) matrix into node tokens $\mathbf{X}_n$. To capture hierarchical brain organization, a hierarchical attention mechanism dynamically aggregates these nodes along with learnable subgraph ($\mathbf{X}_{sg}$) and graph ($\mathbf{X}_g$) tokens based on node feature affinity. Concurrently, an orthogonality constraint ($\mathcal{L}_{oc}$) ensures diverse and non-redundant subgraph patterns. Finally, the final graph token $\hat{\mathbf{X}}_g$ drives disorder classification ($\mathcal{L}_{cls}$), while local node features are refined in parallel via a hierarchical consistency constraint ($\mathcal{L}_{hc}$) and an auxiliary classification loss ($\mathcal{L}_{aux}$).

2.2 Hierarchical Attention Mechanism

To effectively model the brain functional connectivity, we propose a Hierarchical Attention mechanism that organizes information at three distinct levels: node-level, subgraph-level, and graph-level. This module aggregates features in a bottom-up manner using learnable tokens.

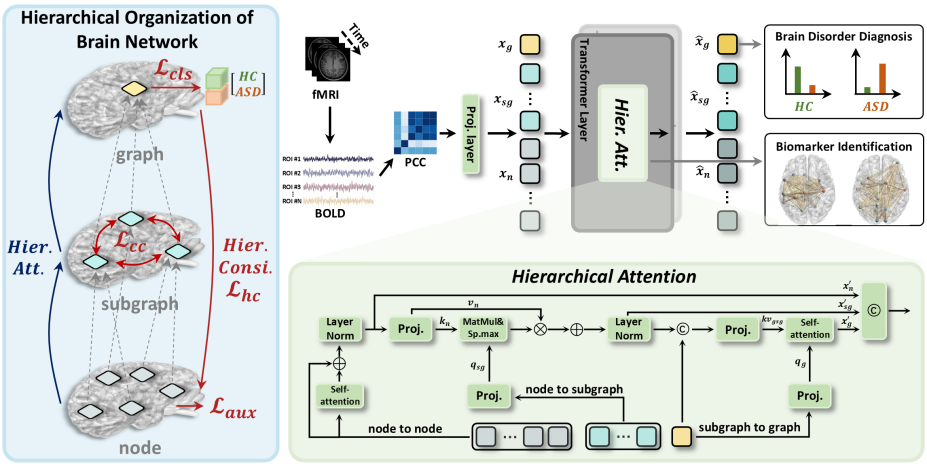


Fig. 2. Framework of the proposed BrainHO. The model constructs the brain network in a bottom-up attention learning manner while enforcing top-down hierarchical consistency. The Hierarchical Attention module comprises three distinct stages: node-to-node, node-to-subgraph, and subgraph-to-graph.

Node to Node: The transition from the node-level to the subgraph-level involves two sequential steps: node interaction and subgraph aggregation.

First, to capture intrinsic dependencies among brain regions, we update the initial node features $\mathbf{X}_n \in \mathbb{R}^{N\times d}$ using a standard self-attention module, ensuring the representations incorporate global network information.

Node to subgraph: We introduce a set of learnable subgraph tokens $\mathbf{X}_{sg} \in \mathbb{R}^{K\times d}$ to represent latent sub-networks. These tokens serve as Queries to aggregate information from the updated node representations $\mathbf{X}'_n$, which act as Keys and Values. To effectively filter noise and learn interpretable sub-networks, we employ the Sparsemax activation function[9]. The updated subgraph features $\mathbf{X}'_{sg}$ are obtained as follows:

$$\mathbf{A}_{(n\to sg)} = \text{Sparsemax}\left(\frac{(\mathbf{X}_{sg}\mathbf{W}_Q)(\mathbf{X}'_n\mathbf{W}_K)^T}{\sqrt{d_h}}\right) \tag{1}$$

$$\mathbf{X}'_{sg} = \text{LN}\left(\mathbf{X}_{sg} + (\mathbf{A}_{(n\to sg)})(\mathbf{X}'_n\mathbf{W}_V)\right) \tag{2}$$

where $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V$ are projection matrices, $\text{LN}(\cdot)$ is layer normalization, and d_h is the dimension of each attention head. Unlike Softmax, Sparsemax induces a sparse probability distribution, ensuring that each subgraph token selectively attends to a relevant subset of brain nodes.

Subgraph to graph: In the final stage, we integrate the learned subgraph-level features into a holistic graph-level representation for disease diagnosis. We

initialize a global graph token $\mathbf{X}_g \in \mathbb{R}^{1\times d}$ to act as the Query. To allow the global token to capture both the subgraph information and its own features, the Keys and Values are projected from the concatenation of the graph token itself and the updated subgraph features $\mathbf{X}''_{sg}$. We employ the standard self attention in this stage to densely aggregate subgraph information:

$$\mathbf{A}_{(sg\to g)} = \text{Softmax}\left(\frac{(\mathbf{X}_g\mathbf{W}'_Q)(\text{Concat}(\mathbf{X}_g, \mathbf{X}''_{sg})\mathbf{W}'_K)^T}{\sqrt{d_h}}\right) \tag{3}$$

This process generates a comprehensive graph-level embedding $\mathbf{X}'_g$ that encapsulates the interactions among different functional sub-networks. Subsequently, $\mathbf{X}'_g$ is concatenated with the updated subgraph tokens $\mathbf{X}'_{sg}$ and node tokens $\mathbf{X}'_n$, and fed into the Feed-Forward Network for classification.

2.3 Subgraph Orthogonality

To encourage diversity and avoid redundancy among the learnable subgraph tokens, we impose a subgraph orthogonality constraint. Specifically, let $\hat{\mathbf{X}}_{sg} \in \mathbb{R}^{K\times d}$ denote the L_2-normalized subgraph tokens from the last transformer layer. We compute the pairwise similarity matrix $\mathbf{S} = \hat{\mathbf{X}}_{sg}\hat{\mathbf{X}}^T_{sg} \in \mathbb{R}^{K\times K}$, where $\mathbf{S}_{i,j}$ represents the cosine similarity between the i-th and j-th subgraph tokens. Ideally, different tokens should be orthogonal ($\mathbf{S}_{i,j} = 0$ for $i \neq j$) while maintaining self-similarity ($\mathbf{S}_{i,i} = 1$). To enforce this, we apply a cross-entropy loss over the similarity matrix, treating the diagonal indices as the ground truth:

$$\mathcal{L}_{\text{OC}} = -\frac{1}{K}\sum_{i=1}^{K}\log\left(\frac{\exp(\mathbf{S}_{i,i})}{\sum_{j=1}^{K}\exp(\mathbf{S}_{i,j})}\right) \tag{4}$$

Minimizing $\mathcal{L}oc$ encourages the model to explore comprehensive and precise subnetworks connectivity patterns.

2.4 Hierarchical Consistency

To refine local node representations under high-level context, we introduce a hierarchical consistency strategy. We construct an auxiliary classifier that predicts disorder labels solely from node-level features. Specifically, the final graph token $\hat{\mathbf{X}}_g$, which aggregates information hierarchically, acts as a teacher to produce logits z_g. Simultaneously, the final node tokens $\hat{\mathbf{X}}_n$ are pooled into a representative vector to generate student logits z_n via an auxiliary head. Additionally, the auxiliary head is supervised by the ground truth labels via a standard cross-entropy loss $\mathcal{L}_{aux}$. To enforce semantic consistency between the global graph context and local node representations, we minimize the Kullback-Leibler divergence:

$$\mathcal{L}_{hc} = \tau^2 \cdot \text{KL}\left(\text{LogSoftmax}\left(\frac{z_n}{\tau}\right) \parallel \text{Softmax}\left(\frac{z_g}{\tau}\right)\right) \tag{5}$$

where τ is the temperature parameter, and a detached copy of z_g is used as the teacher logits.

The total objective function is a weighted sum of the classification loss $\mathcal{L}_{cls}$, auxiliary classification loss $\mathcal{L}_{aux}$, orthogonality loss $\mathcal{L}_{oc}$ and consistency loss $\mathcal{L}_{hc}$.

$$\min_{\Theta} \mathcal{L}_{total} = \mathcal{L}_{cls} + \mathcal{L}_{aux} + \alpha\mathcal{L}_{oc} + \beta\mathcal{L}_{hc} \tag{6}$$

where Θ denotes the learnable parameters of the model, and α, β are the weights to balance the loss terms.

3 Experiments

3.1 Dataset and experimental setup

We utilized rs-fMRI data from ABIDE [3] ($N = 1009$, 516 ASD/493 HCs, C-PAC preprocessed) and REST-meta-MDD [22,2] ($N = 2380$, 1276 MDD/1104 HCs, DPARSF preprocessed). Both datasets were parcellated using the Craddock 200 atlas [4] to extract node features.

We employed 5-fold stratified cross-validation, reserving 25% of training samples as a validation set for early stopping. We report the mean and standard deviation of Accuracy (ACC), AUC, Sensitivity (SEN), and Specificity (SPE) on the test set.

Implemented in PyTorch (A100 GPU), our framework utilized two 8-head transformer layers trained for 200 epochs (batch size 64) using AdamW (LR=1e-4, weight decay=1e-4) and CosineAnnealingLR (min LR=1e-5). We set $\alpha = 1.3$, $K = 8$, $d = 384$, $d_h = d/8$, $\tau = 2.0$, and applied a sigmoid-scheduled β (weight 0.2, center 1/4 steps, slope 0.001). Following prior works[6,21], Mixup[24] augmentation was adopted.

3.2 Results and Discussion

Comparison with state-of-the-art Methods: As shown in Table 1, BrainHO outperformed state-of-the-art baselines on the ABIDE dataset, achieving the highest accuracy (69.68%) and AUC (73.80%). It is noteworthy that while advanced methods such as ALTER[23], DHGFormer[21], and LHDFormer[20] incorporate raw BOLD signals ($\mathcal{B}$) as additional temporal input, BrainHO achieves superior performance relying solely on the static PCC connectivity matrix ($\mathcal{P}$). This result highlights the capability of our framework to extract disease features without complex temporal inputs, validating the superiority of hierarchical subgraph learning over flat node-level modeling approaches(e.g., BNT[6], ALTER[23]). Furthermore, BrainHO attained the highest sensitivity (73.11%), significantly exceeding the second-best Com-BrainTF[1]. Unlike LHDFormer[20], which sacrifices sensitivity for specificity, BrainHO maintained a balanced performance, effectively capturing latent cross-subnetwork interactions. Regarding the larger-scale REST-meta-MDD dataset, BrainHO demonstrated robust generalization, securing the highest Accuracy (64.71%) and Sensitivity (67.43%).

Although LHDFormer exhibited a marginally higher AUC, BrainHO prioritized the identification of positive cases without compromising overall accuracy. This balance makes BrainHO more clinically suitable for effective disease screening compared to baselines that exhibit trade-offs between accuracy and sensitivity.

Table 1. Performance comparison on ABIDE and REST-meta-MDD datasets. The Input column specifies the data type used by each method, where $\mathcal{B}$ represents raw BOLD signals and $\mathcal{P}$ denotes the static PCC connectivity matrix.

Dataset	Methods	Input	ACC(%)	AUC(%)	SEN(%)	SPE(%)
REST-meta-MDD	BNT	$\mathcal{P}$	63.28±1.48	69.08±1.49	65.03±4.83	61.41±6.48
	Com-BrainTF	$\mathcal{P}$	63.24±2.34	68.73±1.76	60.93±7.67	66.21±6.90
	ALTER	$\mathcal{B},\mathcal{P}$	63.95±0.72	68.91±0.88	64.94±6.28	62.75±7.75
	DHGFormer	$\mathcal{B},\mathcal{P}$	61.60±1.36	65.18±1.79	64.73±3.55	58.01±3.84
	LHDFormer	$\mathcal{B},\mathcal{P}$	63.95±1.88	**69.96±1.80**	65.08±5.25	**62.79±8.36**
	BrainHO(Ours)	$\mathcal{P}$	**64.71±2.01**	68.83±1.34	**67.43±4.19**	61.51±3.65
ABIDE	BNT	$\mathcal{P}$	66.61±2.47	73.32±2.48	67.18±6.34	66.08±7.61
	Com-BrainTF	$\mathcal{P}$	67.69±3.09	73.20±3.40	69.28±10.22	65.78±5.56
	ALTER	$\mathcal{B},\mathcal{P}$	66.40±1.73	73.46±2.22	66.17±6.44	67.33±5.46
	DHGFormer	$\mathcal{B},\mathcal{P}$	65.22±3.50	69.15±3.51	67.62±7.03	63.15±7.85
	LHDFormer	$\mathcal{B},\mathcal{P}$	67.50±1.83	73.11±2.10	63.82±6.10	**71.87±7.47**
	BrainHO(Ours)	$\mathcal{P}$	**69.68±2.11**	**73.80±2.52**	**73.11±6.73**	67.08±4.17
ABIDE	Baseline	$\mathcal{P}$	61.55±2.95	68.85±2.91	62.11±13.20	62.05±12.69
	w/o HA	$\mathcal{P}$	61.74±1.78	69.09±1.26	59.57±9.02	64.07±6.99
	w/o $\mathcal{L}_{oc}$	$\mathcal{P}$	65.61±1.46	71.22±1.49	73.10±11.28	59.19±10.74
	w/o $\mathcal{L}_{aux}$	$\mathcal{P}$	67.20±2.13	72.24±3.56	69.32±5.17	65.39±4.98
	w/o $\mathcal{L}_{hc}$	$\mathcal{P}$	68.89±3.05	73.67±2.60	67.69±5.12	70.58±2.71

Ablation study: We validated BrainHO via an ablation study (Table 1). The Baseline (standard Transformer) and w/o HA variant (where HA is replaced by an attention module) yielded only 61.55% and 61.74% accuracy, significantly underperforming the full model and indicating standard attention fails to capture brain hierarchies. Furthermore, removing $\mathcal{L}_{oc}$ led to a sharp accuracy drop to 65.61%, confirming that regularizing subgraph orthogonality is critical for preventing feature collapse and ensuring diverse connectivity patterns. The removal of the auxiliary supervision ($\mathcal{L}_{aux}$) resulted in a performance decline to 67.20%. This suggests that explicit supervision on local representations is indispensable for ensuring the distinctiveness of node-level features. Similarly, excluding $\mathcal{L}_{hc}$ reduced accuracy to 68.89%, highlighting the contribution of the global-to-local (from top to bottom) in refining node features.

Interpretability of the learned sub-networks: Figure 3(a) shows the mapping from the learned sub-networks (rows) to predefined functional networks

(columns). Notably, certain learned sub-networks predominantly align with the defined sub-networks. For example, the No. 7 sub-network for MDD and the No. 5 sub-network for ASD substantially overlap with the SMN and DAN, covering 66.7% and 68.4% of their nodes, respectively. This proves that BrainHO can uncover functionally consistent sub-networks (Shown in Figure 1(b)) consistent with the prior atlas. Moreover, other identified sub-networks flexibly encompass brain regions from multiple predefined functional networks, effectively capture the latent cross-network interactions. This distinct and complementary distribution underscores the orthogonality of the learned sub-networks, precisely enabling precise localization of disease-related cross-network anomalies critical for diagnosis.

Since the CC200 atlas lacks explicit labels, Figure 3(b) visualizes the disease-related sub-networks using the mapped AAL116 names for better interpretation. For ASD, the No. 6 sub-network identifies key nodes like Fusiform_L, Amygdala_R[8], alongside Thalamus_R, Fusiform_R, Vermis, Cerebelum_4-_5_L[11]. Additionally, the No. 8 sub-network highlights Temporal_Sup and Temporal_Mid, which are crucial for ASD-related social perception [14,5]. For MDD, the No. 1 sub-network effectively captures the Hippocampus and Cingulum_Ant, both of which are fundamental to emotion and executive function regulation [16,15]. NO. 4 sub-network includes these nodes (Frontal_Mid_R, Cingulum_Ant_L, Parietal_Inf_L, Insula_R, Frontal_Mid_L, and Precuneus_L), and their functional abnormalities are closely associated with MDD[18].

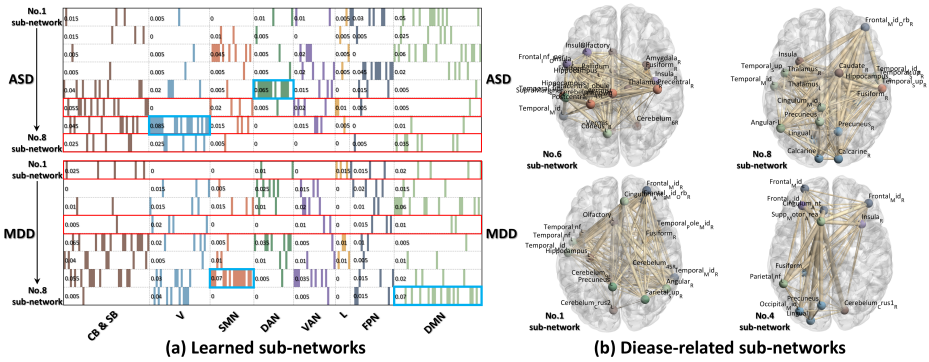


Fig. 3. The interpretability of the sub-networks identified by our model.

4 Conclusion

In this paper, we proposed the Brain Hierarchical Organization Learning framework (BrainHO) to address the limitations of strict, predefined functional atlases in brain network analysis. BrainHO dynamically aggregates brain regions based

on intrinsic functional affinity rather than predefined sub-network labels. We introduced a hierarchical attention mechanism coupled with an orthogonality constraint to ensure diverse sub-network representations, alongside a hierarchical consistency constraint strategy to refine local features using global semantics. Extensive experiments on the ABIDE and REST-meta-MDD datasets demonstrated that BrainHO achieved state-of-the-art classification performance. Furthermore, the visualization of learned organization confirms the model's capacity to uncover interpretable, clinically significant biomarkers, providing new insights into the hierarchical neuropathology of brain disorders.

Acknowledgments. This research was supported by the National Natural Science Foundation of China (62076059), the Science and Technology Joint Project of Liaoning Province (2023JH2/101700367 and ZX20240193) and the Fundamental Research Funds for the Central Universities (N25BJD005).

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

References

1. Bannadabhavi, A., Lee, S., Deng, W., Ying, R., Li, X.: Community-aware transformer for autism prediction in fmri connectome. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. pp. 287–297. Springer Nature Switzerland, Cham (2023)
2. Chen, X., Lu, B., Li, H.X., et al.: The direct consortium and the rest-meta-mdd project: towards neuroimaging biomarkers of major depressive disorder. Psychoradiology **2**(1), 32–42 (06 2022)
3. Craddock, C., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakab, A., Khundrakpam, B.S., Lewis, J.D., Li, Q., Milham, M., et al.: The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. Frontiers in Neuroinformatics **7**(27), 5 (2013)
4. Craddock, R.C., James, G.A., Holtzheimer III, P.E., Hu, X.P., Mayberg, H.S.: A whole brain fmri atlas generated via spatially constrained spectral clustering. Human brain mapping **33**(8), 1914–1928 (2012)
5. Jou, R.J., Minshew, N.J., Keshavan, M.S., Vitale, M.P., Hardan, A.Y.: Enlarged right superior temporal gyrus in children and adolescents with autism. Brain Research **1360**, 205–212 (2010)
6. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain network transformer. In: Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A. (eds.) Advances in Neural Information Processing Systems. vol. 35, pp. 25586–25599. Curran Associates, Inc. (2022)
7. Lee, Y., Park, B.y., James, O., Kim, S.G., Park, H.: Autism spectrum disorder related functional connectivity changes in the language network in children, adolescents and adults. Frontiers in Human Neuroscience **Volume 11 - 2017** (2017)
8. Ma, H., Cao, Y., Li, M., Zhan, L., Xie, Z., Huang, L., Gao, Y., Jia, X.: Abnormal amygdala functional connectivity and deep learning classification in multifrequency bands in autism spectrum disorder: A multisite functional magnetic resonance imaging study. Human Brain Mapping **44**(3), 1094–1104 (2023)

9. Martins, A., Astudillo, R.: From softmax to sparsemax: A sparse model of attention and multi-label classification. In: Balcan, M.F., Weinberger, K.Q. (eds.) Proceedings of The 33rd International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 48, pp. 1614–1623. PMLR, New York, New York, USA (20–22 Jun 2016)
10. Meunier, D., Lambiotte, R., Fornito, A., Ersche, K., Bullmore, E.T.: Hierarchical modularity in human brain functional networks. Frontiers in Neuroinformatics **Volume 3 - 2009** (2009)
11. Nagai, Y., Kirino, E., Tanaka, S., Usui, C., Inami, R., Inoue, R., Hattori, A., Uchida, W., Kamagata, K., Aoki, S.: Functional connectivity in autism spectrum disorder evaluated using rs-fmri and dki. Cerebral Cortex **34**(13), 129–145 (11 2023)
12. Padmanabhan, A., Lynch, C.J., Schaer, M., Menon, V.: The default mode network in autism. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging **2**(6), 476–486 (2017)
13. Pei, S., Ma, J., Lv, Z., Zhang, C., Guan, J.: Community-aware graph transformer for brain disorder identification. In: Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence. pp. 4191–4199 (2025)
14. Pelphrey, K.A., Morris, J.P., McCarthy, G.: Neural basis of eye gaze processing deficits in autism. Brain **128**(5), 1038–1048 (03 2005)
15. Pizzagalli, D.A.: Frontocingulate dysfunction in depression: toward biomarkers of treatment response. Neuropsychopharmacology **36**(1), 183–206 (2011)
16. Schmaal, L., Veltman, D.J., van Erp, T.G., Sämann, P., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W., et al.: Subcortical brain alterations in major depressive disorder: findings from the enigma major depressive disorder working group. Molecular psychiatry **21**(6), 806–812 (2016)
17. Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F.: Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences **106**(31), 13040–13045 (2009)
18. Sundermann, B., Olde lütke Beverborg, M., Pfleiderer, B.: Toward literature-based feature selection for diagnostic classification: a meta-analysis of resting-state fmri in depression. Frontiers in Human Neuroscience **Volume 8 - 2014** (2014)
19. Thomas Yeo, B.T., Krienen, F.M., Sepulcre, J., et al.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology **106**(3), 1125–1165 (2011), pMID: 21653723
20. Xue, R., Han, X., Hu, H., Zhang, Z., Du, S., Gao, Y.: Adaptive Embedding for Long-Range High-Order Dependencies via Time-Varying Transformer on fMRI . In: proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. vol. LNCS 15971. Springer Nature Switzerland (September 2025)
21. Xue, R., Hu, H., Zhang, Z., Han, X., Wang, J., Gao, Y., Du, S.: Dhgformer: Dynamic hierarchical graph transformer for disorder brain disease diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 268–278. Springer (2025)
22. Yan, C.G., Chen, X., Li, L., et al.: Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. Proceedings of the National Academy of Sciences **116**(18), 9078–9083 (2019)
23. Yu, S., Jin, S., Li, M., Sarwar, T., Xia, F.: Long-range brain graph transformer. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (eds.) Advances in Neural Information Processing Systems. vol. 37, pp. 24472–24495. Curran Associates, Inc. (2024)

24. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. In: International Conference on Learning Representations (2018)