The Refining of Brain Connectivity Features on Residual Posterior Patterns

Xinbei Zha[1], Jiaming Zhang[1], and Jin Gu[✉1,2]

[1] Southwest Jiaotong University, Chengdu 611756, CN
[2] Manufacturing Industry Chains Collaboration and Information Support Technology Key Laboratory of Sichuan Province, Chengdu 611756, CN
gujin@swjtu.edu.cn

Abstract. In conjunction with graph neural networks (GNNs), functional connectivity analysis based on fMRI data can provide insights into the interaction and communication patterns in brain network, which has gained increasing attention in the diagnosis of neuropsychiatric disorders. However, traditional GNN based models focus primarily on brain regions, with limited attention given to changes in brain connectivity induced by diseases, and often lack specific methods to address noise and outliers. To accurately preserve and analyze connections in brain networks and retain the structure information in the original graph over message passing, we propose an Residual-Posterior Line Graph Network (RP-LGN). RP-LGN innovatively re-models each edge as a node to highlight functional connectivity information. Subsequently, it integrates residual blocks and a single-pass, low-variance Bayesian variational inference method to approximate the true posterior distribution. Bayesian variational posterior facilitates the quantification of uncertainty in model predictions and enhances model robustness in the presence of noise and anomalous data, ultimately promoting more accurate clinical decision-making. Compared with other models,the performance of RP-LGN was validated on the ABIDE dataset, ADHD-200 dataset,with significant accuracy improvements, and revealed significant site-specific differences and unique connection patterns associated with diseases. Our code is available at: https://github.com/YeDbae/RP-LGN

Keywords: fMRI · Line Graph · Neuropsychiatric Disorders.

1 Introduction

Neuroimaging have revolutionized the ability to explore the intricate functional connectivity of the human brain [6]. Functional magnetic resonance imaging (fMRI) stands out as a powerful tool for capturing dynamics by measuring blood flow changes associated with neural activity across different brain regions [28]. By utilizing deep learning algorithms and computational models, researchers can extract intricate patterns of brain activity from fMRI scans to distinguish between healthy individuals and those affected by various neuropsychiatric conditions [21,9].

Recent advancements in graph neural networks (GNNs), which leverage functional connectivity (FC) matrices derived from fMRI data to construct brain network representations, have successfully simulated inter-regional information transfer and overcome the shortcomings of previous methods. [15,27]. FC networks provide a structured framework to analyze and interpret complex interactions among brain regions, offering insights into both normal cognitive processes and neurological disorders. Existing research on GNNs typically utilizes structured FC frameworks to analyze, simulate, and interpret the complex interactions between brain regions, providing insights into the diagnosis of neuropsychiatric disorders. BrainVGAE, an end-to-end variational graph autoencoder, is to predict partial edges and reduce noise [18]. Multivariate distance-based connectome network (MDCN) integrates populations and maps individual pathological variations [24]. STAGIN employed sliding windows to capture dynamic timestamps and utilized graph convolution and encoding to extract features in the spatial dimension [13]. However, these studies have focused more on functional changes within brain regions and have not explored variations in brain functional connectivity.Brain connectivity can better reflect BOLD signal changes, offering a new perspective for research on neuropsychiatric disorders. Meanwhile, nearly all similar GNNs fMRI classification models place greater emphasis on capturing features through graph structures, utilizing only fully connected layers as classifiers, thereby neglecting the overfitting challenges posed by small sample data.

To fundamentally advance how connectivity patterns are represented and analyzed in neuroimaging, we introduce a novel paradigm shift: directly modeling functional connections as primary entities rather than secondary attributes. This edge-to-node transformation via line graphs, inspired by cross-domain link prediction techniques but uniquely adapted for brain network analysis, explicitly elevates connectivity features to first-class objects in GNN processing. Crucially, this approach incurs minimal computational overhead while adaptively constructing individualized connectomes per subject and enabling direct visualization of clinically significant connections. In this study, we propose a novel Residual-Posterior Line Graph Network (RP-LGN). We address fMRI-specific challenges through task-driven architectural innovations: a residual GraphSAGE backbone preserves structural fidelity against oversmoothing during connectivity-focused message passing, and a Bayesian variational last layer quantifies prediction uncertainty, enhancing robustness against noise and small-sample bias. The incorporation of residual posterior patterns improves the discriminative power and interpretability of functional connectivity matrices, facilitating the identification and characterization of neurodevelopmental disorders like Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD). Our experiments demonstrate robust diagnostic performance on ABIDE I and ADHD-200 datasets, revealing significant effects of neuropsychiatric disorders on brain connectivity, thus supporting the clinical relevance of RP-LGN in early diagnosis and personalized treatment. To our knowledge, this constitutes the first successful integration of line graph transformations with Bayesian poste-

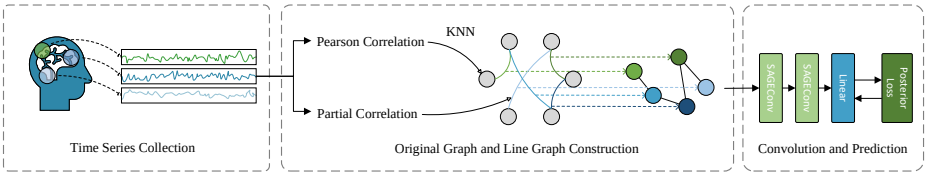

Fig. 1: Overall framework of RP-LGN

rior refinement for brain disorder classification, creating a connection between topological graph theory and clinical neuroscience.

2 Model

RP-LGN transforms the original graph into a line graph via KNN, then obtains logits using SAGEConv and linear layers, and finally refines the prediction results by employing a Bayesian variational posterior. The overall workflow of RP-LGN is illustrated in Fig. 1.

2.1 Initial feature selection

We model the individual brain as a graph structure that simultaneously encompasses a set of nodes $V = \{v_1, v_2, ..., v_n\}$ and a set of edges E based on the original fMRI data. To concretize the abstract nodes and edges into a storage structure, they are mapped to $G = (\mathbf{A}, \mathbf{X})$. In this study, adjacency matrix $\mathbf{A} \in \mathbb{R}^{|V|\times|V|}$ and node feature matrix $\mathbf{X} \in \mathbb{R}^{|V|\times|F|}$ are defined as the partial correlation coefficients and Pearson correlation of pairwise BOLD signals between Regions of Interest (ROIs), where F is the number of features and $|V| = n$ denotes the number of ROIs. Pearson correlation in undirected graphs blurs and confuses the distinction between bidirectional activation/inhibition and unidirectional activation/inhibition among ROIs. To solely represent the strength of correlation itself, we transform all original edge connections into positive values $\mid e_{ij} \mid_{(i,j)\in V} \rightarrow |a_{ij}|_{(i,j)\in n}$. However, constructing graphs in this manner and applying GNNs still primarily focuses on the updating and refinement of region features, rather than treating all brain connection as central features. If each brain connection is modeled as a node in a GNN, the fully connected nature of the initial brain graph results in $(n-1) \times n/2$ edges. Even when converted to an undirected line graph, the scale of the new adjacency matrix dramatically inflates to $(n-1)^2 \times n^2/4$. Given that brain atlases typically classify the brain into hundreds of regions, computing matrices with billions of parameters is a significant challenge for GNNs. To obtain an appropriate number of brain connections, we sparsify the edges in the original graph construction using k-nearest neighbors (KNN). Considering that retaining only the first-order neighbor edges in the graph structure would leave approximately the same number of edges as

the number of nodes, we select $K = 1$ in KNN:

$$d(v_i, v_j) = \|x_i - x_j\|_2 \tag{1}$$

$$\mathcal{N}_i = \text{argmin}_j \{d(v_i, v_j) \mid j \neq i\}_{1 \leq j \leq K, i, j \in V} \tag{2}$$

where node feature x_i denotes the i-th row of node matrix $\mathbf{X}$, $\|\cdot\|_2$ denotes the Euclidean distance formula, d represents the distance between node i and node j, $\mathcal{N}_i$ denotes the K nearest neighbor nodes set of node i. Once the neighbor set is computed, edges not belonging to the set are set to zero, resulting in a new sparse adjacency matrix $A' \in \mathbb{R}^{|V| \times |V|}$ and edge set E'. To highlight the functional alterations of brain connections in psychiatric disorders, it is necessary to map the edges from the original graph G construction into nodes with their own features in a line graph $\tilde{G}$:

$$\tilde{V} = \{\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_{n'}\} = E', e_{ij} \rightarrow \tilde{v}_i \tag{3}$$

where $\tilde{V}$ denotes the node set of line graph, and all $\tilde{v}_i$ are new node constructed by original biran connections. The node features in the line graph node set $\tilde{V}$ are composed of the weights of the corresponding edges in the original graph and the features of the two adjacent nodes connected by these edges:

$$\tilde{x}_i = \text{CONCATE}(\text{SUM}(x_i, x_j), a_{ij}) \tag{4}$$

where $\tilde{x}_i$ represents the features of the i-th node in the line graph, where x_i and x_j denote the endpoints of the edge in the original graph to which the line graph node i corresponds, and a_{ij} indicates the weight of the original edge. After constructing the feature matrix $\tilde{\mathbf{X}} \in \mathbb{R}^{|\tilde{V}| \times |F+1|}$ for the nodes in the line graph, all edges in the original graph that share a common node are treated as connected node pairs in the line graph, resulting in a binarized new adjacency matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{|\tilde{V}| \times |\tilde{V}|}$. The constructed line graph is represented as $\tilde{G} = \{\tilde{\mathbf{A}}, \tilde{\mathbf{X}}\}$. This approach eliminates the need to consider the sequence of node indices, thereby yielding a stable feature representation. Considering that the line graph lacks edge features or edge weights, and aiming to preserve both node and graph features during graph convolution, residual GraphSAGE emerges as a suitable backbone for this purpose. For each SAGEConv layer $\tilde{\mathbf{Z}}^{(l)} = f(\tilde{\mathbf{X}}^{(l)})$, the mapping formula is given by.

$$\tilde{z}_i^{(l)} = \sigma(\theta_i^{(l)} \cdot \text{MEAN}(\{\tilde{x}_i^{(l)}\} || \{\tilde{x}_j^{(l)}, \forall j \in \mathcal{N}_i\}) \tag{5}$$

where $\tilde{x}_i^{(l)}$, $\tilde{z}_i^{(l+1)}$, $\theta_i^{(l)}$ denote the input feature matrix, output feature matrix and learnable parameter of the i-th node in the l-th layer, $||$ represents concatenation operator and σ denotes the activation function. After aggregating neighbor nodes, the model employed residual block to get the input $\tilde{\mathbf{X}}^{(l+1)}$ of the next layer:

$$\tilde{\mathbf{X}}^{(l+1)} = \tilde{\mathbf{X}}^{(l)} + \tilde{\mathbf{Z}}^{(l)} \tag{6}$$

2.2 Classifier

We flatten the node feature matrix $\tilde{\mathbf{X}}^{(L)}$ obtained from the final layer of message passing and use it as the brain map representation $\bar{\mathbf{Z}}$ for each participant:

$$\bar{\mathbf{Z}} = \text{CONCATE}(\tilde{x}_1^{(L)}, \tilde{x}_2^{(L)}, ..., \tilde{x}_{n'}^{(L)}) \tag{7}$$

To get the final predicted label $y = \{y_D, y_{HC}\}$ in $C_y = 2$ classes, the neural network has implemented an input mapping $\phi : \mathbb{R}^F \times \Theta \to \mathbb{R}^{\bar{F}} \sim \phi := \phi(\bar{\mathbf{Z}}, \theta)$, where $\bar{F}$ denotes the output feature dimension of line graph and $\theta \in \Theta$ denotes the parameters in neural network. The unnormalized joint data-label log likelihoods in Bayesian model for classification are:

$$p(y \mid \bar{\mathbf{Z}}, W, \theta) = softmax(\bar{\mathbf{Z}}_L), \bar{\mathbf{Z}}_L = W\phi(\bar{\mathbf{Z}}, \theta) + \varepsilon \tag{8}$$

where $\bar{\mathbf{Z}}_L \in \mathbb{R}^{C_y}$ is logits, $\varepsilon \in \mathbb{R}^{C_y}$ denotes a zero-mean Gaussian noise term with variance Σ, which specify a Gaussian prior for W, and W is the weight of the last linear layer. The weights W of the final layer and the covariance matrix $\Sigma \in \mathbb{R}^{D\times D}$ are treated as random variables, with their variational posterior defined as $q(W, \Sigma|\eta)$, where η represents the variational parameters. It is assumed that the variational posterior factorizes over the weights W for each class:

$$q(W \mid \eta) = \prod_{k=1}^{C_y} q(W_k \mid \eta) \tag{9}$$

where W_k denotes the weight of the k-th class. Meanwhile, the weights for each class possess independent covariance matrices:

$$q(\Sigma) = \prod_{k=1}^{C_y} \mathcal{N}(W_k \mid \mu_k, \Sigma_k) \tag{10}$$

where $\mu_k \in \mathbb{R}^D \sim \mathcal{N}(\mu, \Sigma)$ and Σ_k denote the mean and covariance of variational posterior for the k-th class. To achieve precise marginalization while circumventing the computationally intensive process of evaluating the full marginal likelihood, this study employ stochastic variational inference. This method models the uncertainty of the model's parameters and introduces an approximate posterior $q(\xi \mid \eta)$ to approximate the true posterior distribution. We approximate the posterior distribution as a variational distribution and optimize the model by maximizing the evidence lower bound (ELBO) [8]:

$$S^{-1} \log p(Y|\mathbf{X}, \mathbf{A}, \theta) \geq \mathcal{L}_{VB}(\theta, \eta, \Sigma) - S^{-1}\text{KL}(q(W, \Sigma|\eta) \parallel p(W, \Sigma)) \tag{11}$$

where S denotes the number of subjects, $\mathbf{X}$ and $\mathbf{A}$ denote the original input of this model, $\mathcal{L}_{VB}(\theta, \eta, \Sigma)$ is variational lower bound, KL denotes the Kullback-Leibler divergence. Employing a fully dense variational posterior allows for the

Table 1: Model performance comparison table on ABIDE I and ADHD-200 dataset. SAGE: GraphSAGE, BNT: BrainNetworkTransformer, BNC: BrainNetCNN, FBG: FBNETGEN

Dataset	ABIDE			ADHD-200		
Model	ACC(%)	F1(%)	AUROC(%)	ACC(%)	F1(%)	AUROC(%)
KAN	63.63±3.57	60.10±4.37	72.75±4.38	62.74±9.21	65.22±7.33	63.53±7.90
GAT	58.57±3.36	56.28±3.49	63.26±4.20	51.10±4.31	50.39±4.79	53.92±5.51
GIN	51.93±0.86	42.93±6.39	54.85±2.02	54.17±2.32	53.15±3.04	52.61±5.10
SAGE	55.00±1.11	51.82±2.52	60.67±1.75	60.43±5.90	55.31±9.80	65.06±6.06
BNC	66.21±4.35	69.62±6.14	71.24±3.46	59.18±10.40	74.33±11.59	62.48±12.39
LG-GNN	51.14±0.22	57.12±3.12	62.93±1.35	56.99±0.05	61.91±5.68	51.44±5.24
BrainGNN	55.12±3.26	58.31±2.24	54.79±2.01	62.41±3.43	60.82±2.57	63.24±2.35
BrainGB	64.12±2.90	63.31±3.78	71.23±2.15	59.66±2.09	60.08±3.17	61.94±1.62
FBG	66.55±5.15	66.35±9.58	64.13±11.67	61.22±4.78	**81.77±5.26**	61.95±6.05
BNT	68.74±1.51	68.08±3.68	73.91±4.46	53.05±6.45	72.67±7.92	56.23±6.88
RP-LGN	**73.51±2.86**	**73.16±2.75**	**78.77±0.59**	**66.04±5.25**	74.73±4.30	**71.82±5.77**

derivation of a training objective that encompasses these cross-class relationships. To update the model parameters and minimize prediction error, the loss function is defined as follows:

$$\mathcal{L}_{VB}(\theta, \eta, \Sigma) = \frac{1}{S}\sum_{s=1}^{S}(y_s^\top W\phi_s - log\sum_{k=1}^{C_y}\exp[w_k^\top\phi_s + \frac{1}{2}(\phi_s^\top\Sigma_k\phi_s + \sigma_k^2)]) \quad (12)$$

where $\sigma_i^2 := \Sigma_{ii}$. The goal of the Bayesian variational last layer is to enhance the robustness of the model by incorporating uncertainty factors post-training. The combined training objectives are:

$$\theta^*, \eta^*, \Sigma^* = \arg\max_{\eta,\Sigma}\{\mathcal{L}_{VB}(\theta, \eta, \Sigma) + S^{-1}(\log p(\Sigma) - \mathrm{KL}(q(W \mid \eta)||p(W)))\} \quad (13)$$

where θ^* is trained in SAGEConv layers and η^*, Σ^* are trained through Eq.13. The covariance matrix of the variational posterior Σ are parameterized through Cholesky factorizations, with the means W and η being directly parameterized. Since the variational lower bound already encompasses the objective loss of the classification task, the binary cross entropy loss was not added for the binary classification task in this experiment.

3 Experiments

3.1 Datasets

The study focuses on analyzing brain imaging datasets from ABIDE I [5] and ADHD-200 [19], which include diseased samples for ASD and ADHD, respectively. The ABIDE dataset aggregates rs-fMRI and structural MRI (sMRI) data from 17 research centers. From this dataset, we selected 1009 participants with scanning time points durations exceeding 100, including 493 individuals with

Table 2: Model performance (%) comparison table on NYU, Peking and KKI datasets

Dataset	NYU			Peking			KKI		
Model	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC
KAN	**59.74**	56.82	57.80	62.86	52.92	65.68	67.65	44.13	36.38
GAT	51.428	50.07	52.03	59.80	37.42	69.23	73.53	42.36	50.19
GIN	56.99	56.22	55.95	57.73	54.65	56.15	67.43	54.62	50.80
SAGE	56.04	52.43	54.55	56.72	54.05	60.88	69.85	41.09	45.98
BNC	44.76	28.25	60.60	55.78	75.27	54.18	60.00	78.20	45.23
LG-GNN	48.16	54.60	50.95	61.84	60.00	55.43	73.52	62.46	63.75
BrainGNN	52.81	56.12	51.92	53.90	56.25	46.36	66.29	69.83	68.90
BrainGB	53.35	50.24	53.63	54.20	54.43	54.53	70.11	63.97	65.35
FBNETGEN	58.09	37.71	54.75	48.42	82.41	53.25	65.00	85.00	42.84
BNT	55.23	41.80	51.65	59.98	**86.20**	67.52	**85.00**	85.00	57.35
RP-LGN	58.72	**59.57**	**60.88**	**65.55**	70.71	**69.92**	83.88	**85.51**	**80.45**

ASD and 516 healthy controls (HC). To assess the effects of single-site versus multi-site data, we compared the largest three sites from the ADHD-200 dataset—NYU, KKI, and Peking—as well as the combined dataset from these sites. To remove raw noise, we applied a standard preprocessing pipeline to all images, which included motion correction, slice timing correction, spatial normalization, smoothing, artifact removal, spatial registration, and temporal filtering. In addition, given the high resolution and cross-individual consistency of the Craddock 200 brain atlas, we uniformly divided all participants' brain regions into 200 regions of interest (ROIs) in our study.

In each dataset, we performed five-fold cross-validation, where ramdomly allocating 80% of the data for training and 20% for testing in each fold. Given the sample size, we uniformly trained all models applied to the ABIDE dataset for 100 epochs, while models applied to the ADHD-200 and its sub-sites were trained for 50 epochs. The final performance is reported as the mean and standard deviation of the test results across the five folds of cross-validation.

To prevent excessive smoothing of data due to multiple aggregations in the graph convolutional layers, we configured the RP-LGN with two graph convolution layers. Each convolution layer is followed by a linear layer with the same dimensionality, a LeakyReLU, and a BatchNorm layer. The output feature dimensions for the two convolution layers are 256 and 64, respectively. In the experiments, the batch size is uniformly set to 16, the learning rate and weight decay are set to 1e-4, and the dropout rate is 0.5.

3.2 Model performance

To comprehensively evaluate the performance of RP-LGN, we compared it with the single classifier KAN [16], GNN baselines GAT [22], GIN [23], and GraphSAGE [7], the classical brain classification model BrainNetCNN [12], as well as novel deep learning models LG-GNN [26], BrainGNN [14], BrainGB [3], FBNETGEN [10], and BrainNetworkTransformer [11] in Tab.1 and 2. RP-LGN has achieved commendable results on the large-scale datasets ABIDE I and ADHD-

200, as well as on individual sites such as NYU, Peking, and KKI, demonstrating the model's robust generalization performance across datasets of varying scales.

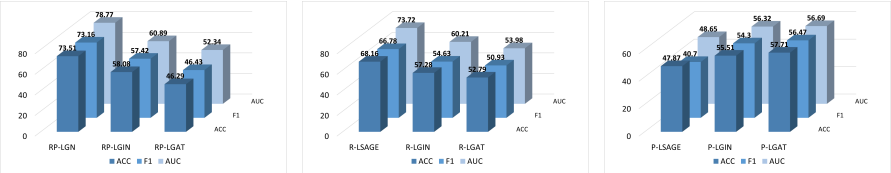

Fig. 2: Resudual and posterior ablation study on ABIDE dataset based on SAGE, GIN and GAT. R: residual, P: posterior, L: line graph.

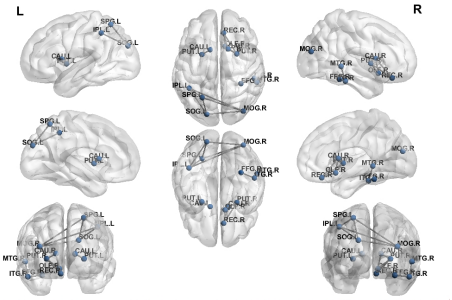

Fig. 3: Significant brain connection of ADHD

Fig. 2 illustrates the average accuracy with standard error on the ABIDE dataset under different module ablations. It can be observed that the combination of GraphSAGE and residual connections preserves more of the original graph information without excessively smoothing the graph, thereby maintaining performance. In contrast, the attention mechanism inherent in GAT and the residual posterior pattern may not integrate well, resulting in a counterproductive effect. In line graph-based models, Grad-CAM can directly analyze the importance of each connection in ADHD disease classification through node gradients. The specific connectivity visualization image is shown in Fig. 3. In this experiment, the RP-LGN model identified connections between Superior Parietal Lobule and Inferior Parietal Lobule in the prefrontal cortex, Caudate Nucleus and Putamen in the basal ganglia, Inferior Temporal Gyrus and Middle Temporal Gyrus in the temporal lobe, and the Fusiform Gyrus. The prefrontal cortex is involved in processing spatial attention and sensory information integration [17]. ADHD patients often exhibit developmental delays and coordination issues in this region. In task execution, reward processing, and impulse control,

ADHD patients commonly show either excessive or insufficient activity in the basal ganglia [4]. The temporal lobe plays a crucial role in auditory information processing, and ADHD patients may have difficulties with auditory attention and information processing tasks [1]. Additionally, alterations in the functional connectivity of the Fusiform Gyrus in ADHD patients impact visual attention, facial expression recognition, and social interactions [20].

4 Conclusion

Based on the findings from the ABIDE, ADHD-200 datasets and subsites of ADHD-200, RP-LGN demonstrated robust performance in analyzing FC in neuropsychiatric disorders. RP-LGN effectively identifies critical brain connections that exhibit disease-related variations, thereby enhancing the understanding of altered brain network dynamics, elucidating complex disease mechanisms, and improving personalized diagnosis and treatment strategies.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (no.62306249), the Sichuan Provincial Natural Science Foundation (no.2025ZNSFSC0484) and the Fundamental Research Funds for the Central Universities (no.2682025ZTPY048).

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

References

1. Chiang, C., Ouyang, C., Yang, R., Wu, R., and Lin, L.: Increased Temporal Lobe Beta Activity in Boys With Attention-Deficit Hyperactivity Disorder by LORETA Analysis. Frontiers in Behavioral Neuroscience, **14**(85) (2020)
2. Constable, R. T.: Challenges in fMRI and Its Limitations. Functional Neuroradiology (2023)
3. Cui, H., Dai, W., Zhu, Y.; Kan, X., Gu, A. A. C., Lukemire, J., Zhan, L., He, L., Guo, Y., and Yang C.: BrainGB: A Benchmark for Brain Network Analysis With Graph Neural Networks. IEEE Transactions on Medical Imaging, **42**(2): 493-506 (2023)
4. Curtin, K., Fleckenstein, A. E., Keeshin, B. R., Yurgelun-Todd, D. A., Renshaw, P. F., Smith, K. R., and Hanson, G. R.: Increased risk of diseases of the basal ganglia and cerebellum in patients with a history of attention-deficit/hyperactivity disorder. Neuropsychopharmacology, **43**: 2548-2555 (2018)
5. Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K.et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry, **19**(6):659-667 (2014)
6. Fu, L., Cai, M., Zhao, Y, Zhang, Z., Qian, Q., Xue, H. et al.: Twenty-five years of research on resting-state fMRI of major depressive disorder: A bibliometric analysis of hotspots, nodes, bursts, and trends. Heliyon, **10**(13) (2024)
7. Hamilton, W. L., Ying, R., and Leskovec, J.: Inductive Representation Learning on Large Graphs. NIPS. Curran Associates, Inc, California, USA (2017)
8. Harrison, J., Willes, J., and Snoek, J.: Variational Bayesian Last Layers. ICLR. PMLR, Vienna, Austria (2024)

9. He, J., Kurita, K., Yoshida, T., Matsumoto, K., Shimizu, E., and Hirano, Y.: Comparisons of the amplitude of low-frequency fluctuation and functional connectivity in major depressive disorder and social anxiety disorder: A resting-state fMRI study. Journal of Affective Disorders, 362: 425-436 (2024)
10. Kan, X., Cui, H., Lukemire, J., Guo, Y., and Yang, C.: FBNETGEN: Task-aware GNN-based fMRI Analysis via Functional Brain Network Generation. Proceedings of The 5th International Conference on Medical Imaging with Deep Learning, 618-637. PMLR, Montreal, Canada (2022a)
11. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., and Yang, C.: Brain Network Transformer. NeurIPS. NeurIPS Foundation, Vancouver, Canada (2022b)
12. Kawahara, J., Brown, C. J., Miller, S. P., Booth, B. G., Chau, V., Grunau, R. E., Zwicker, J. G., and Hamarneh, G.: BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage, **146**: 1038-1049 (2017)
13. Kim, B., Ye, J. C., and Kim, J.: Learning Dynamic Graph Representation of Brain Connectome with Spatio-Temporal Attention. NeurIPS, Neural Information Processing Systems Foundation (2021)
14. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L. H., Ventola, P., and Duncan, J. S.: BrainGNN: Interpretable Brain Graph Neural Network for fMRI Analysis. Medical Image Analysis, **74** (2021)
15. Liu, S, and Gui, R.: Fusing multi-scale fMRI features using a brain-inspired multi-channel graph neural network for major depressive disorder diagnosis. Biomedical Signal Processing and Control, **90** (2024)
16. Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., and Tegmark, M.: KAN: Kolmogorov-Arnold Networks. arXiv:2404.19756 (2024)
17. Long, Y., Pan, N., Ji, S., Qin, K., Chen, Y., Zhang, X., He, M., Suo, X., Yu, Y., Wang, S., and Gong, Q.: Distinct brain structural abnormalities in attention-deficit/hyperactivity disorder and substance use disorders: A comparative meta-analysis. Translational Psychiatry, **12**(368) (2022)
18. Mai, Q., Nakarmi, U., and Huang, M.: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, Las Vegas, USA (2022)
19. Milham, M. P., Fair, D., Mennes, M., and Mostofsky, S. H., The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. Frontiers in Systems Neuroscience, **6**(62) (2012)
20. Parlatini, V., Andrews, D. S., Pretzssch, C. M., Arenella, M., Daly, E., Ecker, C., and Murphy, D. G.: Cortical alterations associated with lower response to methylphenidate in adults with ADHD. Nature Mental Health, **2**: 514-524 (2024)
21. Shahzad, M. N., and Ali, H.: Deep learning based diagnosis of PTSD using 3D-CNN and resting-state fMRI data. Psychiatry Research: Neuroimaging, **343** (2024)
22. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y.: Graph Attention Networks. ICLR. ICLR, Vancouver, Canada (2018)
23. Xu, K., Hu, W., Leskovec, J., and Jegelka, S.: How Powerful are Graph Neural Networks? ICLR. ICLR, New Orleans, USA (2019)
24. Yang, Y., Ye, C., and Ma, T.: A deep connectome learning network using graph convolution for connectome-disease association study. Neural Networks, **164**: 91-104 (2023)
25. Yin, W., Li, L., and Wu, F.: Deep learning for brain disorder diagnosis based on fMRI images. Neurocomputing, **469**: 332-345 (2022)

26. Zhang, H., Song, R., Wang, L., Zhang, L., Wang, D., Wang, C., and Zhang, W.: Classification of Brain Disorders in rs-fMRI via Local-to-Global Graph Neural Networks. IEEE Transactions on Medical Imaging, **42**(2): 444-455 (2023)
27. Zhang, R., Murray, S. B., Duval, C. J., Wang, D. J. J., and Jann, K.: Functional connectivity and complexity analyses of resting-state fMRI in pre-adolescents demonstrating the behavioral symptoms of ADHD. Psychiatry Research, **334** (2024)
28. Zhen, Y., Yang, Y., Zheng, Y., Wang, X., Liu, L., Zheng, Z., Zheng, H., and Tang, S.: The heritability and structural correlates of resting-state fMRI complexity. NeuroImage, **296** (2024)