Ada-FCN: Adaptive Frequency-Coupled Network for fMRI-Based Brain Disorder Classification

Yue Xun[1], Jiaxing Xu[2], Wenbo Gao[1], Chen Yang[1], and Shujun Wang[1,3,4⋆]

[1] Department of Biomedical Engineering, Hong Kong Polytechnic University, Hong Kong SAR
shu-jun.wang@polyu.edu.hk
[2] College of Computing and Data Science, Nanyang Technological University, Singapore
[3] Research Institute for Smart Ageing, The Hong Kong Polytechnic University, Hong Kong SAR
[4] Research Institute for Artificial Intelligence of Things, The Hong Kong Polytechnic University, Hong Kong SAR

Abstract. Resting-state fMRI has become a valuable tool for classifying brain disorders and constructing brain functional connectivity networks by tracking BOLD signals across brain regions. However, existing models largely neglect the multi-frequency nature of neuronal oscillations, treating BOLD signals as monolithic time series. This overlooks the crucial fact that neurological disorders often manifest as disruptions within specific frequency bands, limiting diagnostic sensitivity and specificity. While some methods have attempted to incorporate frequency information, they often rely on predefined frequency bands, which may not be optimal for capturing individual variability or disease-specific alterations. To address this, we propose a novel framework featuring Adaptive Cascade Decomposition to learn task-relevant frequency sub-bands for each brain region and Frequency-Coupled Connectivity Learning to capture both intra- and nuanced cross-band interactions in a unified functional network. This unified network informs a novel message-passing mechanism within our Unified-GCN, generating refined node representations for diagnostic prediction. Experimental results on the ADNI and ABIDE datasets demonstrate superior performance over existing methods. The code is available at https://github.com/XXYY20221234/Ada-FCN.

Keywords: fMRI · Functional connectivity network · Disorder classification.

1 Introduction

In the field of neuroscience, a key aim is to derive abnormal patterns in the brain that are linked to neurological disorders such as Alzheimer's, Autism, and Parkinson's. Resting-state state functional magnetic resonance imaging (fMRI)

⋆ Corresponding author.

has been becoming a valuable technique for this objective by tracking Blood Oxygenation Level Dependent (BOLD) signals across paired brain regions and then constructing brain functional connectivity network where nodes represent different brain regions and edges indicate signal correlations between them [2]. The brain functional network helps identify synchronized activity that could serve as dynamic biomarkers for neurological disorders, facilitating early diagnosis and treatment.

Recent advances in graph neural networks (GNNs) have spurred significant progress in brain functional network analysis [6, 11, 16, 17]. Models like BrainGNN [9] employ ROI-selection pooling to highlight disease-relevant regions, while PRGNN [10] enforces group-level consistency through graph pooling with anatomical regularization. Complementary approaches, such as BrainNetCNN [7], leverage edge-to-edge convolutional filters to exploit topological relationships, and Transformer-based architectures [21] utilize attention mechanisms to model global interdependencies. Despite their success, these methods share a critical limitation: they treat BOLD signals as monolithic time series, disregarding the multi-frequency nature of neuronal oscillations [26]. This challenge of identifying and leveraging key frequency components is also a central topic in the broader field of multivariate time series forecasting [22, 23]. It is worth noting that different neurological disorders often manifest as disruptions in specific frequency bands [20]. Therefore, the conventional practice of restricting brain functional connectivity analysis to a single, low-frequency band can obscure critical, band-specific information present in other frequencies, leading to a loss of sensitivity and specificity in disease diagnosis.

Recognizing this, some work has attempted to incorporate frequency information into brain functional network analysis. Hu *et al.* [5] used the discrete wavelet transform (DWT) to decompose BOLD signals into multiple frequency bands and then constructed a sparse functional connectivity network by fusing the information from each band. MFHC [25] constructed both frequency-specific and cross-frequency high-order functional connectivity networks to capture richer interaction patterns. Tewarie *et al.* [12] used a multiplexed graph representation to analyze cross-frequency interactions in magnetoencephalography (MEG) data. While existing frequency-based methods consider multi-band information, they suffer from two critical limitations. First, they rely on preset frequency ranges despite the fact that individual brain functions may work differently, especially when they have neurological disorders. This means that important brain connectivity patterns related to the disease might be missed. Second, they simply combine features from different frequency bands, but signals from distinct frequencies represent different physiological processes. Mixing them directly may fail to reveal interactions between frequencies and lead to misleading conclusions.

To address these limitations, we propose a novel framework that involves Adaptive Cascade Decomposer and Frequency-Coupled Connectivity Learning to enhance the analysis of brain functional connectivity. We adaptively decompose the raw fMRI time series into task-relevant frequency sub-bands for each brain region. This learned decomposition involves tailored low- and high-frequency sig-

nal pairs. We then construct a comprehensive functional connectivity network, capturing both intra-band and, importantly, nuanced cross-band interactions. These connectivity patterns are unified into a single representation of the brain's functional network. This unified network is the basis for a novel message-passing mechanism within our Unified-GCN. The GCN generates refined node representations, which are aggregated and processed by fully connected layers to produce the final diagnostic prediction. Our contributions can be summarized as follows:

- We propose a learnable decomposition method to adaptively extract task-relevant frequency sub-bands from fMRI time series, overcoming the limitations of fixed or handcrafted frequency band definitions in existing approaches.
- We introduce a Unified-GCN framework, incorporating a novel Dual-Projection Bilinear Attention mechanism for holistic brain network modeling. This end-to-end framework seamlessly integrates adaptive decomposition, frequency-aware message passing, and cross-frequency alignment through attention.
- Extensively evaluated on the ADNI and ABIDE datasets, our model demonstrates superior diagnostic accuracy and AUROC compared to state-of-the-art methods.

2 Method

Neurological disorder classification based on brain functional networks aims to predict the state of the disease of each subject by using fMRI signals. Given a set of fMRI time series $X \in \mathbb{R}^{N\times T}$, which represents the activity of N brain ROIs in T time points. Our objective is to learn a predictive function $f : X \to y \in \{1, 2, \ldots, c\}$, which maps the fMRI signals to a diagnostic label.

As illustrated in Fig. 1, Ada-FCN consists of three key components: an adaptive cascade decomposer, a frequency-coupled connectivity learning network, and classification head and loss function.

2.1 Adaptive Cascade Decomposer

Existing methods for fMRI-based disorder identification often overlook the complex interplay between different frequency components, and preset or full bands might not optimally capture the diverse neural information present across different tasks or individuals. To address this limitation, we introduce an adaptive cascade decomposer that adaptively decomposes the original time series into a hierarchy of frequency sub-bands. Given $X = [x_1, \ldots, x_N]^\top \in \mathbb{R}^{N\times T}$, where $\mathbf{x}_i \in \mathbb{R}^T$ denotes the time series of the i-th ROI, We define $L_i^0 = x_i$ as the initial input. For each ROI i, and at each level $k \in \{1, \ldots, K\}$, the decomposition proceeds in two steps:

$$L_i^k = \mathcal{F}_L^k\big(L_i^{k-1}\big), \quad H_i^k = \mathcal{F}_H^k\Big(L_i^{k-1} - L_i^k\Big), \tag{1}$$

where L_i^k is the low-frequency approximation at level k and H_i^k is the corresponding high-frequency residual. The low-frequency extraction operator $\mathcal{F}_L^k(\cdot)$

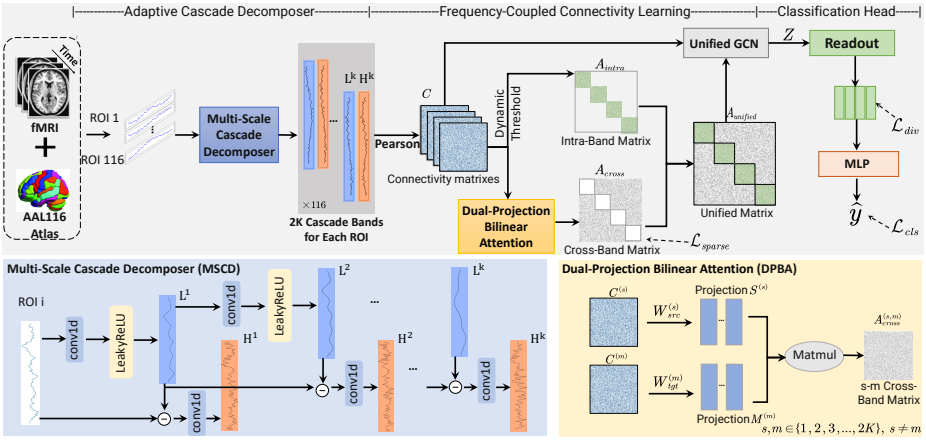

Fig. 1. The framework of Ada-FCN for fMRI-based brain disorder classification.

is implemented using a dilated 1D convolution with an appropriately chosen kernel, padding, and dilation factor, followed by a LeakyReLU activation. In contrast, the high-frequency extraction operator $\mathcal{F}_H^k(\cdot)$ employs a separate 1D convolution with a smaller kernel and standard padding, without an intervening activation function to refine the residual signal. In this manner, for each ROI i, we obtain a cascade of sub-band signals:

$$\widetilde{X} = \{(L_i^1, H_i^1), \ldots, (L_i^K, H_i^K)\}, \tag{2}$$

where each pair (L_i^k, H_i^k) represents the low-frequency approximation and high-frequency residual at the k-th scale. By stacking these $2K$ sub-band signals along a new dimension, the original fMRI time series $X \in \mathbb{R}^{N\times T}$ is transformed into a multi-band representation $\widetilde{X} \in \mathbb{R}^{(2K)\times N\times T}$.

2.2 Frequency-Coupled Connectivity Learning

Intra-Band Connectivity via Dynamic Thresholding. Intra-band connectivity is crucial because each frequency band offers distinct insights into brain function. We use an adaptive dynamic threshold to retain robust connections tailored to each band's statistics, effectively filtering out noise. For each subject's decomposed sub-bands $\widetilde{X} \in \mathbb{R}^{(2K)\times N\times T}$, we compute the Pearson correlation to generate a connectivity matrix $C^{(k)} = \text{Pearson}\left(\widetilde{X}^{(k)}\right) \in \mathbb{R}^{N\times N}$, $k = 1, \ldots, 2K$, then apply an dynamic threshold to eliminate weak connections:

$$A_{\text{intra}}^{(k)} = T_{\tau k}\left(C^{(k)}\right), \quad T_{\tau k}(\cdot) : \text{mask}_{ij} = \begin{cases} 1, & \text{if } |C_{ij}^{(k)}| > \tau_k, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

These $2K$ sparsified intra-band matrices are then integrated using the Kronecker direct sum $A_{\text{intra}} = \bigoplus_{k=1}^{2K} A_{\text{intra}}^{(k)} \in \mathbb{R}^{(2KN)\times(2KN)}$ which is a block-diagonal matrix capturing intra-band transitions within each sub-band.

Cross-Band Coupling with Dual-Projection Bilinear Attention. To model interactions across frequency bands, we adopt an attention-based coupling mechanism that learns off-diagonal adjacency blocks. Consider two distinct frequency bands $s, m \in \{1, 2, 3, \ldots, 2K\}, s \neq m$. Let $C^{(s)},\ C^{(m)} \in \mathbb{R}^{N\times N}$ be their correlation matrices. First, each matrix is transformed into source or target spaces via learnable weights $W_{src}^{(s)},\ W_{tgt}^{(m)} \in \mathbb{R}^{d\times d}$, where d is the hidden dimension. Concretely, we obtain $S^{(s)} = C^{(s)} W_{src}^{(s)} \in \mathbb{R}^{N\times d}$ and $M^{(m)} = C^{(m)} W_{tgt}^{(m)} \in \mathbb{R}^{N\times d}$. This dual-projection allows asymmetric and band-specific interactions. Next, we compute the bilinear interaction $S^{(s)}\big(M^{(m)}\big)^\top$, where each entry in the resulting $\mathbb{R}^{N\times N}$ matrix captures the pairwise similarity between the i-th row of $S^{(s)}$ and the i-th row of $M^{(m)}$:

$$A_{\text{cross}}^{(s,m)} = S^{(s)}\big(M^{(m)}\big)^\top. \tag{4}$$

Finally, we construct $A_{\text{cross}} \in \mathbb{R}^{(2KN)\times(2KN)}$ by arranging all cross-band matrices $\mathbf{A}_{\text{cross}}^{(s,m)}$ at off-diagonal blocks corresponding to band pairs, while setting the diagonal blocks to zero.

Unified Graph Convolution. The unified graph convolution integrates both intra-band and cross-band connectivity patterns to enable holistic message passing across all frequency-specific nodes, by:

$$A_{\text{unified}} = A_{\text{intra}} + \lambda A_{\text{cross}} \in \mathbb{R}^{(2KN)\times(2KN)}, \tag{5}$$

$$H_0 = \text{Stack}\Big(C^{(1)}, C^{(2)}, \ldots, C^{(2K)}\Big) \in \mathbb{R}^{(2KN)\times N}, \tag{6}$$

where λ is a learnable scaling factor that balances the contribution of cross-band interactions. The initial node features H_0 are derived from the raw Pearson correlation matrices. These are used for unified graph convolution:

$$H^{(i)} = \sigma\Big(D^{-\frac{1}{2}} A_{\text{unified}} D^{-\frac{1}{2}} H^{(i-1)} W^{(i)}\Big), \tag{7}$$

where $W^{(i)}$ is the learnable weight matrix for the i-th layer, D is the degree matrix computed from the unified adjacency matrix A_{unified} and $\sigma(\cdot)$ is the activation function(e.g., ReLU).

2.3 Classification Head and Loss Function

Classification Head. Let $Z \in \mathbb{R}^{(2KN)\times d}$ denote the node representations output by the final unified graph convolution layer, where i is the hidden dimension.

To obtain graph-level classification features, we first split the multi-band representations as $Z = [Z^{(1)}; \ldots; Z^{(2K)}]$ to obtain the embedding of each band, where $Z^{(k)} \in \mathbb{R}^{N \times d}$ for $k = 1, \ldots, 2K$. For each sub-band representation $Z^{(k)}$, we apply mean readout to obtain graph-level embeddings and the final prediction is computed by concatenating all sub-band embeddings and feeding them into an MLP:

$$h_k = \frac{1}{N} \sum_{i=1}^{N} Z_i^{(k)} \in \mathbb{R}^d, \quad \hat{y} = \mathrm{MLP}\Big([h_1 \parallel \ldots \parallel h_{2K}]\Big) \in \mathbb{R}^c, \tag{8}$$

where $\parallel$ denotes concatenation and c is the number of classes.

Loss Function. In order to make model optimization easier to converge, we utilize three loss functions to guide the end-to-end training: (1) A commonly-used cross-entropy loss $\mathcal{L}_{ce}$ for graph classification; (2) a band diversity loss $\mathcal{L}_{div}$ to encourage distinct frequency bands to capture complementary patterns; (3) a sparsity loss $\mathcal{L}_{sparse}$ for cross-band matrices that highlights only the most salient cross-band connections.

$$\mathcal{L}_{\text{div}} = \frac{1}{2K\,(2K-1)} \sum_{i=1}^{2K} \sum_{\substack{j=1 \\ j \neq i}}^{2K} \cos\big(h_i, h_j\big), \quad \mathcal{L}_{\text{sparse}} = \left\| \mathbf{S}^{(s)} \mathbf{M}^{(m)\top} \right\|_1, \tag{9}$$

where $\cos(\cdot)$ computes cosine similarity. $\mathcal{L}_{div}$ operates on graph-level embeddings $\{h_k\}_{k=1}^{2K}$ and $\|\cdot\|_1$ is L_1 penalty. The total loss is computed by:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{cls}} + \lambda_1 \mathcal{L}_{\text{div}} + \lambda_2 \mathcal{L}_{\text{sparse}}, \tag{10}$$

where λ_1 and λ_2 are trade-off hyperparameters for balancing different losses.

3 Experiments and Results

3.1 Experiment Settings

Datasets. We use two brain network datasets constructed by Xu *et al.* [19] for evaluation, which are ADNI [1] for Alzheimer's Disease (AD) and ABIDE [3] for Autism (ASD). ADNI is categorized into four classes based on the progression of cognitive impairment: cognitive normal (CN), significant memory concern (SMC), mild cognitive impairment (MCI) and Alzheimer's disease (AD). The dataset comprises 914 CN, 73 SMC, 264 MCI, and 65 AD samples. The time series for each sample of AD has a length of 197. For ABIDE dataset, we focus on the binary classification task of TC vs. ASD, with 61 (44.2%) samples from patients with ASD. We truncate the first 300 time points for samples of ABIDE. For brain region parcellation, we employed the AAL116 atlas [13] on both datasets.

Table 1. Results over 10-fold-CV (Average ± Standard Deviation). The best result is highlighted in **bold** while the second-best result is in underline.

Methods	Year	ADNI		ABIDE	
		Accuracy(%)	AUROC(%)	Accuracy(%)	AUROC(%)
GCN [8]	2017	62.05±4.71	63.41±3.25	67.35±4.16	67.93±3.19
GraphSAGE [4]	2017	69.55±4.86	71.79±3.32	69.92±4.12	71.28±2.49
GAT [14]	2018	64.12±2.18	66.68±4.23	63.64±3.73	62.82±4.54
BrainNetCNN [7]	2017	73.27±4.59	72.46±3.66	72.71±2.32	73.37±1.22
PRGNN [10]	2020	62.63±2.29	60.32±1.71	67.34±2.80	68.52±1.85
BrainGNN [9]	2021	74.31±1.62	69.73±2.39	74.92±1.68	74.20±2.12
Contrasformer [18]	2024	73.32±4.04	69.26±1.03	73.06±1.35	73.53±2.86
Tewarie et al. [12]	2016	74.18±2.34	72.58±1.52	66.32±4.14	67.58±4.45
MFHC [25]	2017	68.26±1.64	71.94±2.41	71.27±2.49	71.05±2.73
Hu et al. [5]	2021	67.24±3.85	72.31±2.96	66.28±1.27	67.06±2.58
Autoformer [15]	2021	65.29±3.24	63.85±2.16	68.49±4.78	67.35±3.52
Leddam [24]	2024	70.35±3.77	65.31±3.41	70.22±1.25	69.42±1.75
Ada-FCN (Ours)	-	**79.68±2.65**	**75.30±1.24**	**77.89±1.52**	**77.62±1.68**

Implementation details. All experiments were conducted on a Linux server with a NVIDIA GeForce RTX 4090 with 24GB memory. The whole network is trained in an end-to-end manner using the Adam optimizer with an initial learning rate of 1×10^{-3}, a weight decay of 1×10^{-4} and a batch size of 32. We use early stopping based on AUROC metric, terminating training if AUROC does not improve for 15 consecutive epochs. The data is split to 8:1:1 for training, validation, and testing with 10-fold cross-validation. In our experiments, we found that setting the decomposition level $K = 2$ in the Adaptive Cascade Decomposer acheived the best performance on both the ADNI and ABIDE datasets.

3.2 Results

Baseline Models. We carefully choose eleven well-acknowledged neural network models as our baseline methods, including : (1) General-Purpose GNNs: GCN [8], GAT [14] and GraphSAGE [4]. (2) Advanced Brain Connectivity Networks: BrainNetCNN [7], BrainGNN [9], PRGNN [10] and Contrasformer [18]. (3) Frequency Domain Methods: Hu et al. [5], MFHC [25] and the work of Tewarie et al. [12]. Although these approaches exploit multi-band representations, they generally lack a unified mechanism for adaptive cross-band connectivity learning. (4) Time Series Domain Methods: We adapt Leddam [24] and Autoformer [15], originally for long-term series forecasting, due to their learnable decomposition strategies. We report the classification Accuracy and AUROC over 10-fold cross-validation in Table 1.

Ablation Studies. To demonstrate the effectiveness of our proposed Ada-FCN model, we conducted ablation studies focusing on key components. Table 2

Table 2. Ablation study on important components in Ada-FCN on ADNI and ABIDE datasets. The best result is highlighted in **bold**.

			ADNI	ABIDE
DT	$\mathcal{L}_{sparse}$	$\mathcal{L}_{div}$	**Accuracy (%)**	**Accuracy (%)**
✓	✓	✓	**79.68±1.65**	**77.89±1.52**
✓	✓		77.35±1.22	76.53±1.95
✓		✓	76.14±1.66	73.29±1.61
	✓	✓	79.33±1.29	76.26±1.14

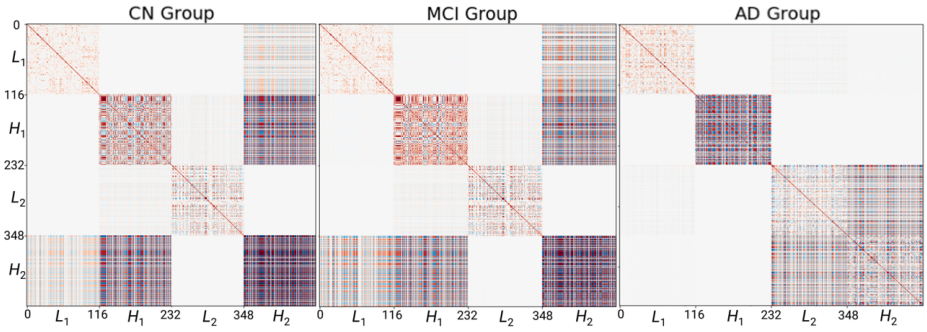

Fig. 2. Distinct frequency-coupled connectivity patterns revealed by group-averaged A_{unified} matrices for CN, MCI, and AD.

presents the results. DT denotes the dynamic threshold approach, where ablation experiments are performed using the first 25% as the threshold value. On both datasets, removing the DT significantly reduces accuracy, indicating that DT is crucial for model performance. Removing $\mathcal{L}_{\text{spasity}}$ decreases Accuracy on both datasets. This suggests that the sparsity constraint helps the model to learn more representative brain network features. Removing $\mathcal{L}_{\text{div}}$ also decreases Accuracy on both datasets. Ablation experiments show that each component of the Ada-FCN model contributes to the final performance.

Interpretability. We visualize A_{unified}, which integrated intra- and cross-band connectivity across four sub-bands, revealed distinct group differences. As shown in Fig. 2. The AD group exhibited significantly stronger H_1 intra-band connectivity than CN and MCI, suggesting a potential role for altered high-frequency oscillations in AD. While some cross-band connections (e.g., L_1-H_1, L_2-H_2) were similar across groups, AD generally showed weaker cross-frequency interactions. While the L_1-L_2 cross-band connectivity was found to be uniformly weak across all groups, the MCI group uniquely demonstrated a prominent enhancement of intra-band connectivity within the H_1 frequency band. Furthermore, MCI often presented connectivity patterns intermediate between AD and CN, consistent with its potential as a prodromal AD stage.

4 Conclusion

In this paper, we present Ada-FCN, a novel framework for neurological disorder classification using fMRI data, designed to overcome the limitations of prior methods that disregard inherent multi-frequency characteristics of fMRI. Ada-FCN adaptively extracts task-relevant frequency sub-bands, moving beyond the constraints of pre-defined frequency ranges, and constructs a unified brain functional network by capturing both intra-band and complex cross-band interactions, leading to refined node representations for enhanced classification. Future work will focus on examining the interpretability of the learned frequency-specific features.

Acknowledgments. This work was partially supported by RGC Collaborative Research Fund (No. C5055-24G), the Start-up Fund of The Hong Kong Polytechnic University (No. P0045999), the Seed Fund of the Research Institute for Smart Ageing (No. P0050946), and Tsinghua-PolyU Joint Research Initiative Fund (No. P0056509). The authors also acknowledge the use of AI tools for assistance in grammar enhancement and spelling checks during the preparation of this manuscript.

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

References

1. 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 (2013)
2. Cui, H., Dai, W., Zhu, Y., Li, X., He, L., Yang, C.: Interpretable graph neural networks for connectome-based brain disorder analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 375–385. Springer (2022)
3. Dadi, K., Rahim, M., Abraham, A., Chyzhyk, D., Milham, M., Thirion, B., Varoquaux, G., Initiative, A.D.N., et al.: Benchmarking functional connectome-based predictive models for resting-state fmri. NeuroImage **192**, 115–134 (2019)
4. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. Advances in neural information processing systems **30** (2017)
5. Hu, R., Peng, Z., Zhu, X., Gan, J., Zhu, Y., Ma, J., Wu, G.: Multi-band brain network analysis for functional neuroimaging biomarker identification. IEEE transactions on medical imaging **40**(12), 3843–3855 (2021)
6. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain network transformer. Advances in Neural Information Processing Systems **35**, 25586–25599 (2022)
7. Kawahara, J., Brown, C.J., Miller, S.P., Booth, B.G., Chau, V., Grunau, R.E., Zwicker, J.G., Hamarneh, G.: BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage **146**, 1038–1049 (2017)
8. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations (ICLR) (2017)

9. 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)
10. Li, X., Zhou, Y., Dvornek, N.C., Zhang, M., Zhuang, J., Ventola, P., Duncan, J.S.: Pooling regularized graph neural network for fmri biomarker analysis. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23. pp. 625–635. Springer (2020)
11. Peng, C., Huang, Y., Dong, Q., Yu, S., Xia, F., Zhang, C., Jin, Y.: Biologically plausible brain graph transformer. arXiv preprint arXiv:2502.08958 (2025)
12. Tewarie, P., Hillebrand, A., van Dijk, B.W., Stam, C.J., O'Neill, G.C., Van Mieghem, P., Meier, J.M., Woolrich, M.W., Morris, P.G., Brookes, M.J.: Integrating cross-frequency and within band functional networks in resting-state meg: A multi-layer network approach. Neuroimage **142**, 324–336 (2016)
13. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. Neuroimage **15**(1), 273–289 (2002)
14. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)
15. Wu, H., Xu, J., Wang, J., Long, M.: Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. Advances in neural information processing systems **34**, 22419–22430 (2021)
16. Xu, J., Bian, Q., Li, X., Zhang, A., Ke, Y., Qiao, M., Zhang, W., Sim, W.K.J., Gulyás, B.: Contrastive graph pooling for explainable classification of brain networks. IEEE Transactions on Medical Imaging (2024)
17. Xu, J., Chen, Y., Dong, X., Lan, M., HUANG, T., Bian, Q., Cheng, J., Ke, Y.: BrainOOD: Out-of-distribution generalizable brain network analysis. In: The Thirteenth International Conference on Learning Representations (2025), https://openreview.net/forum?id=3xqqYOKILp
18. Xu, J., He, K., Lan, M., Bian, Q., Li, W., Li, T., Ke, Y., Qiao, M.: Contrasformer: A brain network contrastive transformer for neurodegenerative condition identification. In: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. pp. 2671–2681 (2024)
19. Xu, J., Yang, Y., Huang, D.T.J., Gururajapathy, S.S., Ke, Y., Qiao, M., Wang, A., Kumar, H., McGeown, J., Kwon, E.: Data-driven network neuroscience: On data collection and benchmark. In: Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track (2023)
20. Yang, L., Yan, Y., Li, Y., Hu, X., Lu, J., Chan, P., Yan, T., Han, Y.: Frequency-dependent changes in fractional amplitude of low-frequency oscillations in alzheimer's disease: a resting-state fmri study. Brain imaging and behavior **14**, 2187–2201 (2020)
21. Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D., Shen, Y., Liu, T.Y.: Do transformers really perform badly for graph representation? Advances in neural information processing systems **34**, 28877–28888 (2021)
22. Yu, G., Li, Y., Guo, X., Wang, D., Liu, Z., Wang, S., Yang, T.: Lino: Advancing recursive residual decomposition of linear and nonlinear patterns for robust time series forecasting (2025), https://arxiv.org/abs/2410.17159
23. Yu, G., Li, Y., Wang, J., Guo, X., Aviles-Rivero, A.I., Yang, T., Wang, S.: Refocus: Reinforcing mid-frequency and key-frequency modeling for multivariate time series forecasting (2025), https://arxiv.org/abs/2502.16890

24. Yu, G., Zou, J., Hu, X., Aviles-Rivero, A.I., Qin, J., Wang, S.: Revitalizing multivariate time series forecasting: Learnable decomposition with inter-series dependencies and intra-series variations modeling. arXiv preprint arXiv:2402.12694 (2024)
25. Zhang, Y., Zhang, H., Chen, X., Shen, D.: Constructing multi-frequency high-order functional connectivity network for diagnosis of mild cognitive impairment. In: Connectomics in NeuroImaging: First International Workshop, CNI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Proceedings 1. pp. 9–16. Springer (2017)
26. Zuo, X.N., Di Martino, A., Kelly, C., Shehzad, Z.E., Gee, D.G., Klein, D.F., Castellanos, F.X., Biswal, B.B., Milham, M.P.: The oscillating brain: complex and reliable. Neuroimage **49**(2), 1432–1445 (2010)