Dual-Stream Multi-Band Fusion Network for Dynamic Functional Connectivity Analysis in Brain Disorder Classification

Ling Wu[1], Hexi Li[1], Zhengyuan Lyu[1], Zhiwei Song[1], Hu Yu[1], and Xiaojuan Guo[1,2] ✉

[1] School of Artificial Intelligence, Beijing Normal University, Beijing 100875, China
[2] Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing 100875, China
gxj@bnu.edu.cn

Abstract. Dynamic functional connectivity (dFC) derived from fMRI captures the temporal dynamics of brain networks, where cross-frequency features provide complementary characterizations for brain disorder classification. Although existing multi-band approaches incorporate sub-band decomposition, they primarily rely on simplistic averaging or fixed-weight strategies, failing to adaptively fuse information across multiple frequency bands. To handle this limitation, we propose a dual-stream multi-band fusion network (DSMFN): 1) The frequency-domain stream employs a sub-band graph encoding-interaction module, where local graph convolution networks (GCNs) extract band-specific topological features, and lightweight convolutions replace computationally intensive attention mechanisms for data-driven band contribution allocation, followed by a global GCN to aggregate cross-band information; 2) The time-domain stream preserves local dynamic properties via residual multi-layer perceptron networks; 3) A feature-temporal dual-dimension cross-attention mechanism jointly models temporal evolution and cross-domain complementarity to adaptively integrate multi-band features with time-varying characteristics. Experiments on two distinct brain disease datasets demonstrate the effectiveness of DSMFN, achieving accuracies of 91.40% for MCI and 70.18% for ASD classification. This study provides an efficient fusion framework for multi-band dynamic brain network analysis, advancing precise diagnosis of brain disorders. Our code is available at https://github.com/WuLingBNU/DSMFN.

Keywords: Dynamic Functional Connectivity, Multiple Frequency Bands, Graph Neural Network, Brain Disease Classification.

1 Introduction

Functional connectivity (FC) derived from fMRI reflect functional interactions between distinct brain regions [1]. While static functional connectivity (sFC) captures steady-state features through full-time averaging, it may overlook temporal variations [2]. Dynamic functional connectivity (dFC) characterizes temporal evolution of connectivity patterns, not only enhancing sensitivity to abnormal neural coordination in pathological

states through temporal dynamics [3] but also uncovering richer features in the frequency domain. Accordingly, this dynamic feature modeling makes dFC particularly valuable for brain disease classification and diagnostic applications [4].

Recent advances in frequency-domain analysis of fMRI have opened new perspective for brain disorder classification [5]. Studies indicate that spontaneous fluctuations in resting-state BOLD signals predominantly reside within the 0.01-0.08Hz low-frequency range [6], yet this band may contain multiple cognitive function-related sub-bands, such as Slow-4 (0.027-0.073 Hz) and Slow-5 (0.01-0.027 Hz) [7]. Yaesoubi et al. [8] demonstrated that visual network activity predominantly resides in the relatively lower frequency band (0.003–0.037 Hz), while the default mode network shows significant activation in the higher frequency band (0.075–0.113 Hz), with cross-frequency interactions between them. This suggests that it is important to further decompose broad-band signals to capture sub-band-specific information.

Wavelet transform, owing to its time-frequency localization capability, is particularly suited for non-stationary BOLD signals [9] and has become a mainstream approach for multiple frequency bands (abbreviated as multi-band) functional connectivity analysis [10]. For instance, Ding et al. [11] proposed a frequency-adaptive model to dynamically construct functional connectivity by comparing cross-band correlations between brain regions. Hu et al. [12] generated sparse networks through averaging of multi-band connectivity matrices, validating the biological relevance of band fusion. Ding et al. [13] designed independent kernel functions for each band and achieved classification via linear kernel combination. However, existing methods still struggle to fully exploit band-specific information and lack dynamic fusion mechanisms tailored to downstream tasks.

Brain networks can be represented as graph structures with nodes denoting brain regions and edges representing FC, thereby forming functional connectivity networks (FCNs). Conventional approaches often use handcrafted network metrics (e.g., clustering coefficient) [4], which may introduce selection bias due to prior assumptions. Graph Neural Networks (GNNs) overcome this by learning topological representations through iterative neighborhood aggregation [14]. The multi-band functional graphs from BOLD signals align with GNNs' ability to model non-Euclidean data, enabling data-driven discovery of connectivity patterns. Therefore, we propose a sub-band graph encoding-interaction (SGEI) module: independent GCN branches first extract topological features from spectrally decomposed sub-bands; lightweight convolutions then perform data-driven weight allocation; finally, a global GCN enables global optimization and integration of cross-band information. To integrate time and frequency domains, we design a feature-temporal dual-dimension cross-attention (FTDC) mechanism: attention weights are independently computed along feature (e.g., different streams) and temporal (e.g., sliding window evolution) dimensions; matrix multiplication combines these weights to generate joint attention maps for feature fusion. We propose a Dual-stream Multi-band Fusion Network (DSMFN), and our main contributions are summarized as follows:

— We design a dual-stream framework that adaptively fuses multi-band connectivity features with time-varying characteristics through attention mechanisms jointly modeling temporal evolution and cross-domain feature interactions.

— We develop a sub-band graph encoding-interaction module that enables both the extraction of band-specific features and adaptive dynamic integration of cross-band information.
— Experimental results demonstrate that our method outperforms existing methods in brain disease diagnosis tasks (such as MCI and ASD classification).

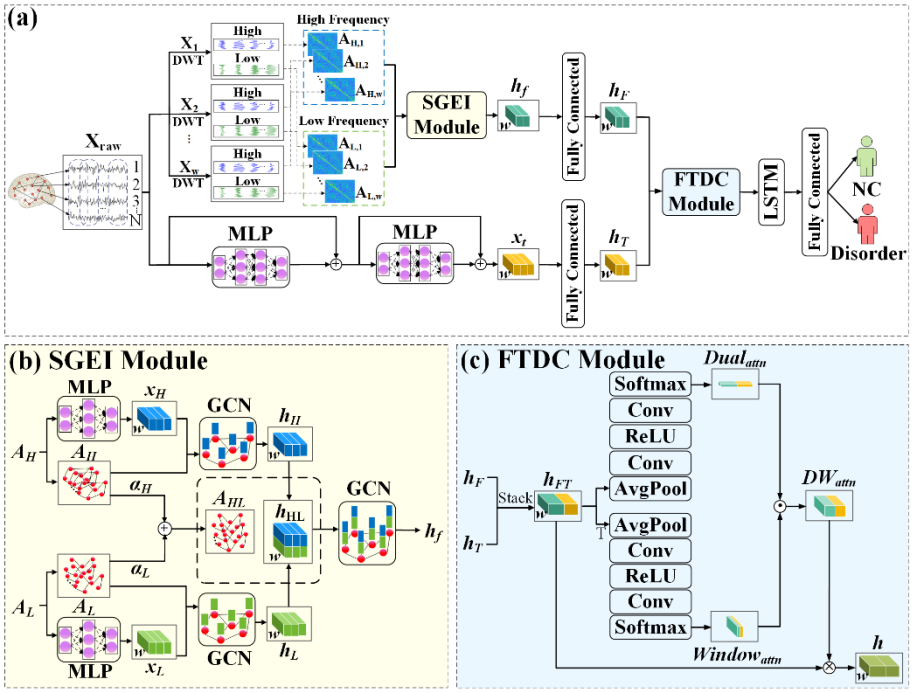

Fig. 1. (a) The overall architecture of DSMFN, (b) The sub-band graph encoding-interaction (SGEI) module, (c) The feature-temporal dual-dimension cross-attention (FTDC) module.

2 Method

We propose a dual-stream dynamic analysis framework for FCNs, termed DSMFN (Dual-stream Multi-band Fusion Network), as illustrated in Fig. 1. Given that brain activity patterns contain both frequency-specific oscillatory features and time-varying dynamic characteristics, we design this dual-stream architecture to separately model and synergistically integrate time-frequency information, achieving cross-domain complementary characterization of neural dynamics. The framework is composed of three key components: 1) frequency-domain feature extraction module, 2) time-domain feature extraction module, and 3) feature fusion and classification module. Given a subject's original BOLD signal $X_{raw} \in \mathbb{R}^{N\times T}$, where N denotes the number of regions of interest (ROIs) and T represents the number of time points. The entire BOLD signal is divided into a series of overlapping time windows $X = \{X_1, X_2, \cdots, X_w\}, X_i \in \mathbb{R}^{N\times l}$

using a sliding window of appropriate size, where w represents the number of windows and l represents the length of the window. Our objective is to develop a comprehensive representation that can effectively capture the temporal dynamics and frequency specificity, thereby enabling a more accurate distinction between patients and healthy controls. The following sections will provide a detailed explanation of the specific components of each module.

2.1 Frequency-domain Feature Extraction

To align with the discrete structure from the division of time windows, we apply Discrete Wavelet Transform (DWT) [15] to decompose BOLD signals X_i within each time window into various frequency sub-bands. In practical applications, the DWT is commonly implemented using a filter bank, where each decomposition separates the signal into low- and high-frequency components. The low-frequency components undergo iterative decomposition until a specified resolution threshold is met [7].

Next, we calculate the Pearson correlation coefficient between each pair of brain regional BOLD signals, yielding a set of functional connectivity matrices $\{A_{L,1}, \cdots, A_{L,w}, A_{H,1}, \cdots, A_{H,w}\}$. We adopt single-level decomposition, and these matrices capture the interaction patterns of brain regions at different frequency scales, where L and H denote the low-frequency and high-frequency sub-bands.

Then, we introduce a sub-band graph encoding-interaction (SGEI) module to extract and hierarchically fuse multi-band features. For the *i*-th time window and *j*-th frequency band, we first encode its functional connectivity matrix $A^{j,i}$ to capture the matrix's features:

$$x^{j,i} = \varphi_{MLP}(A^{j,i}) \tag{1}$$

We apply independent graph convolutional branch to learn topological features for each frequency band, where adjacency matrices undergo symmetric normalization and node representations are iteratively updated layer-wise. We then integrate multi-band local topological patterns via feature concatenation operations.

$$\hat{A}^{j,i} = D^{-\frac{1}{2}} A^{j,i} D^{-\frac{1}{2}} \tag{2}$$

$$h_{l+1}^{j,i} = \sigma\left(\hat{A}^{j,i} h_l^{j,i} W_l^{j,i}\right), \qquad h_0^{j,i} = x^{j,i} \tag{3}$$

$$h^{multi,i} = Concat\_Fuse(h_{l+1}^{1,i}, h_{l+1}^{2,i}, \cdots, h_{l+1}^{f,i}) \tag{4}$$

Here, $h_l^{j,i}$ represents the hidden features at the *l*-th layer, and $W_l^{j,i}$ is a learnable transformation matrix that maps node features to a new feature space. To overcome limitations of fixed band-weighting schemes, we use lightweight 1×1 convolutional kernels to dynamically learn the importance weights α_j of each band, generating a cross-band adjacency matrix $A^{multi,i}$. This enables data-driven adaptive allocation of band-specific contributions.

$$A^{multi,i} = \sum_{j=1}^{f} \alpha_j A^{j,i} \tag{5}$$

Finally, we construct a global graph convolutional layer to capture nonlinear cross-band interaction patterns through multi-layer feature propagation.

$$h_{l+1}^{cross,i} = \sigma\left(\hat{A}^{multi,i} h_l^{cross,i} W_l^{cross,i}\right), \qquad h_0^{cross,i} = h^{multi,i} \tag{6}$$

Through this dual mechanism of local band feature extraction and global cross-band information aggregation, the resulting feature representation $h_{l+1}^{cross,i} \in \mathbb{R}^{w \times D}$ preserves both band-specific discriminability and cross-band synergy.

2.2 Time-domain Feature Extraction

To effectively preserve the local dynamic properties of raw BOLD signals and compensate for potential information loss during frequency-domain feature extraction, we design a parallel stream for time-domain feature extraction. This stream employs a residual multi-layer perceptron (MLP) network to encode and extract features, where residual connections [16] are introduced to mitigate gradient vanishing and maintain the integrity of original temporal patterns. And a dimensional alignment layer projects the temporal features into a latent space compatible with frequency-domain representations.

2.3 Feature Fusion and Classification

To improve the integration of dual-stream features, we design a feature-temporal dual-dimension cross-attention (FTDC) module. This module performs adaptive fusion by constructing an interaction between the dual-stream and temporal features.

Specifically, we first stack the frequency-domain and time-domain features along the feature dimension to form the input $h^{ft} \in \mathbb{R}^{2 \times w \times D}$. Attention weights are then extracted separately from both the feature streams $Dual_{attn} \in \mathbb{R}^{2 \times 1 \times 1}$ and the temporal dimension $Window_{attn} \in \mathbb{R}^{w \times 1 \times 1}$:

$$Dual_{attn} = \sigma(W_2\left(ReLU\left(W_1\left(AAP(h^{ft})\right)\right)\right)) \tag{7}$$

$$Window_{attn} = \sigma(V_2\left(ReLU\left(V_1\left(AAP((h^{ft})^T)\right)\right)\right)) \tag{8}$$

Here, AAP refers to the adaptive average pooling operation and σ represents the softmax activation function. Next, we combine the attention weights from both dimensions through matrix multiplication to weight the input features:

$$h = h^{ft} \otimes (Dual_{attn} \cdot Window_{attn}^T) \tag{9}$$

Here, $\otimes$ denotes element-wise multiplication. This module not only adaptively adjusts the importance of dual-stream features but also effectively captures the critical information via a window-level attention mechanism.

The fused features are standardized using layer normalization. Then, we use a Long Short-Term Memory (LSTM) [17] network to model the fine-grained temporal dynamics and long-range dependencies in the sequence features, with its internal gating mechanisms implicitly regulating the information flow. The global feature representation is derived from LSTM's terminal hidden state h_w. The final classification layer maps these features to the target class space y, while we use cross-entropy loss to compute the classification loss, and apply label smoothing to mitigate model overfitting and enhance generalization performance.

$$L_{CE} = -\sum_{c=1}^{C} y_c log(\widehat{p_c}) \tag{10}$$

3 Experiments

3.1 Dataset and Experimental Settings

Dataset. We evaluated the proposed method on two neurological conditions: MCI and ASD. The rs-fMRI datasets used were from ADNI-2 (https://adni.loni.usc.edu/) and ABIDE-II (https://fcon_1000.projects.nitrc.org/). For ADNI-2, we selected 135 healthy controls and 274 MCI patients (age range: 56-95 years). For ABIDE-II, we similarly selected 557 healthy controls and 496 ASD patients (age range: 5-64 years). Due to site-specific scanning parameters, each site's dataset underwent independent preprocessing. These datasets were preprocessed using the DPARSFA [18] toolbox, with steps including slice timing correction, realignment, co-registration of T1-weighted images to functional images, spatial normalization by using EPI templates, spatial smoothing with a 6mm×6mm×6mm Gaussian kernel [19], nuisance covariates regression, band-pass filtering (0.01-0.08 Hz), and so on. The AAL atlas was used to extract the average time series from 90 brain regions in the gray matter.

Implementation Details. The model was implemented on an NVIDIA RTX3090 GPU using the PyTorch 2.1.0 framework. During training, we set the batch size to 16, the learning rate to 0.001, used the db4 wavelet function, and the Adam optimizer. To ensure the reliability of the evaluation, we employed a nested 10-fold cross-validation strategy [20]. The evaluation metrics included accuracy (ACC), sensitivity (SEN), specificity (SPE), area under the curve (AUC), and F1 score (F1).

3.2 Performance Comparison

To validate the effectiveness of our proposed method, we conducted comparative experiments using several state-of-the-art deep learning methods on the MCI and ASD diagnostic tasks: 1) LSTM [17] and CNN-LSTM, 2) BrainNetCNN [21], 3) BrainGNN [22] and ST-GCN [23], and 4) Transformer [24] and ACI-FBN [25]. All comparison methods used identical datasets and a nested 10-fold cross-validation strategy to ensure the comparability of the experiments.

The experimental results in Table 1 showed that our method outperformed others in both MCI and ASD classifications, achieving the highest accuracy of 91.40% for MCI and 70.18% for ASD. On the ACC metric, we performed paired samples t-tests and our model showed statistically significant differences ($p < 0.05$) when compared with each baseline method listed in Table 1. This demonstrated the effectiveness of our multi-band fusion and time-frequency dual-stream complementary brain network analysis framework.

Table 1. Classification results of MCI and ASD.

Method	MCI/NC(%)					ASD/NC(%)				
	ACC	SEN	SPE	AUC	F1	ACC	SEN	SPE	AUC	F1
LSTM	71.41	63.55	40.11	73.33	62.17	62.42	61.73	60.52	61.20	60.23
CNN-LSTM	82.52	78.37	65.30	85.98	79.18	64.22	63.53	70.80	63.58	62.78
BrainNetCNN	74.12	67.81	49.74	73.30	66.91	53.41	51.04	48.46	54.01	42.83
BrainGNN	84.73	82.77	76.23	91.38	82.64	65.93	64.55	**78.52**	63.96	63.92
ST-GCN	83.78	79.16	65.56	79.51	79.32	65.93	65.17	71.05	63.67	64.96
Transformer	74.39	68.60	53.45	77.14	67.95	54.11	51.10	64.97	56.89	37.70
ACI-FBN	88.88	**95.96**	73.34	93.67	**91.75**	59.72	58.07	56.92	58.84	48.61
DSMFN (Ours)	**91.40**	89.64	**84.34**	**94.06**	89.96	**70.18**	**69.90**	73.67	**70.76**	**69.75**

3.3 Ablation Study

Effectiveness of the Proposed Dual-Stream Architecture. We designed five sets of experimental comparisons: a single frequency-domain stream (without frequency band division), a single frequency-domain stream (with multi-frequency bands), a single time-domain stream, a dual-stream architecture (without frequency band division), and a dual-stream architecture (with multi-frequency bands). The results were shown in Table 2.

It was evident that the fusion of multi-band features in the single frequency-domain stream did not significantly improve performance, likely due to the inevitable information loss during wavelet decomposition. In the two diagnostic tasks, the dual-stream architecture did not always outperform the single time-domain stream, excluding the performance improvement solely due to added parameters. The best performance was achieved with the combination of multi-band fusion and the dual-stream architecture.

This demonstrated that detailed frequency-domain analysis and dual-domain complementarity enhanced functional network feature extraction across multiple dimensions and scales.

Table 2. Ablation study results of the Dual-Stream architecture.

Ablation Method	MCI/NC(%)					ASD/NC(%)				
	ACC	SEN	SPE	AUC	F1	ACC	SEN	SPE	AUC	F1
Base_Freq	70.68	64.22	47.15	71.56	63.32	65.53	64.62	69.01	64.07	64.00
Multi_Freq	68.98	62.46	42.98	68.90	62.01	63.52	63.08	68.69	61.88	62.43
Only_Time	89.16	87.13	80.53	91.92	87.46	65.53	65.00	**76.29**	65.00	64.33
Dual_Stream	87.90	86.13	80.43	91.45	86.07	66.33	65.64	73.82	65.61	65.50
Dual_Multi_Freq	**91.40**	**89.64**	**84.34**	**94.06**	**89.96**	**70.18**	**69.90**	73.67	**70.76**	**69.75**

Effectiveness of the Frequency Band Integration Strategy. Using MCI as an example, we compared three different frequency band integration strategies: Concat-Fusion, Attention-Fusion, and MGCN fusion proposed in this paper. The results were shown in Fig. 2 (a). Fig. 2 (b) showed that only splitting the signal into high-frequency and low-frequency components (i.e., the number of frequency bands=2) yielded the best performance. Over-partitioning of frequency bands may introduce unnecessary complexity.

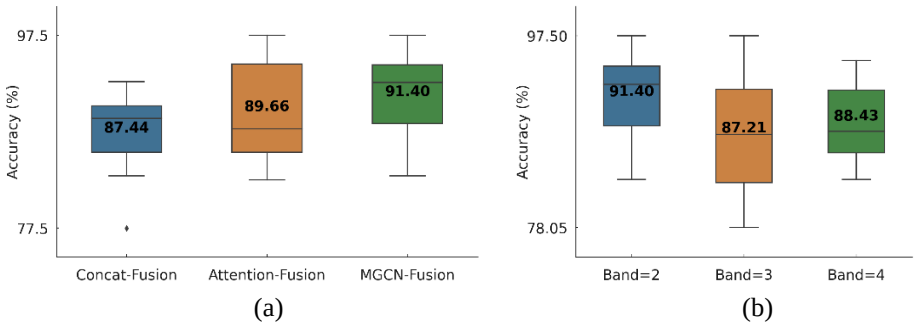

Fig. 2. (a) Performance of three multi-band integration strategies. (b) Experimental selection of optimal number of frequency bands.

4 Conclusion

In this study, we introduce an innovative framework for multi-band dynamic brain network analysis—DSMFN. By fusing multi-band features and complementing time-frequency dual-streams, this method highlights two crucial challenges in traditional research: 1) the specific oscillatory patterns of the brain across different frequency bands, and 2) the complementary nature of the time-frequency features. Experimental results

on two brain disorder diagnostic datasets validate the effectiveness of our method. They show that DSMFN significantly enhances disease diagnosis accuracy by learning brain network features in a multi-dimensional and multi-scale manner.

Acknowledgments. This study was supported by the National Natural Science Foundation of China under Grant 62071051.

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

References

1. Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T.: Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nature Neuroscience **18**(11), 1664-1671(2015).
2. Preti, M.G., Bolton, T.A.W., Van De Ville D.: The dynamic functional connectome: State-of-the-art and perspectives. NeuroImage **160**, 41-54 (2017).
3. Liu, J., Liao, X., Xia, M., He, Y.: Chronnectome fingerprinting: Identifying individuals and predicting higher cognitive functions using dynamic brain connectivity patterns. Human Brain Mapping **39**(2), 902-915 (2018).
4. Li, Y., Liu, J., Tang, Z., Lei, B.: Deep spatial-temporal feature fusion from adaptive dynamic functional connectivity for MCI identification. IEEE Transactions on Medical Imaging **39**(9), 2818-2830 (2020).
5. Wang, B.: Exploring intricate connectivity patterns for cognitive functioning and neurological disorders: incorporating frequency-domain NC method into fMRI analysis. Cerebral Cortex **34**(5), bhae195 (2024).
6. Fransson, P.: Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. Human Brain Mapping **26**(1), 15-29 (2005).
7. Kajimura, S., Margulies, D., Smallwood, J.: Frequency-specific brain network architecture in resting-state fMRI. Scientific Reports **13**(1), 2964 (2023).
8. Yaesoubi, M., Silva, R.F., Iraji, A., Calhoun, V.D.: Frequency-aware summarization of resting-state fMRI data. Frontiers in Systems Neuroscience **14**, 16 (2020).
9. Grossmann, A., Morlet, J.: Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM Journal on Mathematical Analysis **15**(4), 723-736 (1984).
10. Bullmore, E., Fadili, J., Maxim, V., Sendur, L., Whitcher, B., Suckling, J., Brammer, M., Breakspear, M.: Wavelets and functional magnetic resonance imaging of the human brain. NeuroImage **23**, S234-S249 (2004).
11. Ding, Y., Xu, X., Peng, L., Zhang, L., Li, W., Cao, W., Gao, X.: Wavelet transform-based frequency self-adaptive model for functional brain network. Cerebral Cortex **33**(22), 11181-11194 (2023).
12. Hu, R., Peng, Z., Zhu, X., Gan, J., Zhu, Y., Ma, J.: Multi-band brain network analysis for functional neuroimaging biomarker identification. IEEE Transactions on Medical Imaging **40**(12), 3843-3855 (2021).
13. Ding, Y., Zhang, T., Cao, W., Zhang, L., Xu, X.: A multi-frequency approach of the altered functional connectome for autism spectrum disorder identification. Cerebral Cortex **34**(8), bhae341 (2024).

14. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P.S.: A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems **32**(1), 4-24 (2020).
15. Guo, T., Zhang, T., Lim, E., López-Benítez, M., Ma, F., Yu, L.: A review of wavelet analysis and its applications: Challenges and opportunities. IEEE Access **10**, 58869-58903 (2022).
16. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778 (2016).
17. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation, **9**(8), 1735-1780 (1997).
18. Yan, C., Zang, Y.: DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. Frontiers in Systems Neuroscience, **4**, 1377 (2010).
19. Li, L., Zhang, L., Cao, P., Yang, J., Wang, F., Zaiane, O.R.: Exploring spatio-temporal interpretable dynamic brain function with transformer for brain disorder diagnosis. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 195-205. Springer (2024).
20. Liu, J., Cui, W., Chen, Y., Ma, Y., Dong, Q., Cai, R.: Deep fusion of multi-template using spatio-temporal weighted multi-hypergraph convolutional networks for brain disease analysis. IEEE Transactions on Medical Imaging **43**(2), 860-873 (2023).
21. 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).
22. 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).
23. Gadgil, S., Zhao, Q., Pfefferbaum, A., Sullivan, E.V., Adeli, E., Pohl, K.M.: Spatio-temporal graph convolution for resting-state fMRI analysis. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 528-538. Springer (2020).
24. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. In: Conference on Neural Information Processing Systems, pp. 5998-6008 (2017).
25. Zhang, J., Wu, X., Tang, X., Zhou, L., Wang, L., Wu, W.: Asynchronous functional brain network construction with spatiotemporal transformer for MCI classification. IEEE Transactions on Medical Imaging **44**(3), 1168-1180 (2024).