Core-Periphery Principle Guided State Space Model for Functional Connectome Classification

Minheng Chen[1], Xiaowei Yu[1], Jing Zhang[1], Tong Chen[1], Chao Cao[1], Yan Zhuang[1], Yanjun Lyu[1], Lu Zhang[2], Tianming Liu[3], and Dajiang Zhu[1(✉)]

[1] Department of Computer Science and Engineering, University of Texas at Arlington, USA
{mxc2442,xxy1302,jxz7537,txc5603,cxc0366,yxz8653,yxl9168}@mavs.uta.edu, dajiang.zhu@uta.edu
[2] Department of Computer Science, Indiana University Indianapolis, USA
lz50@iu.edu
[3] School of Computing, University of Georgia, USA
tliu@uga.edu

Abstract. Understanding the organization of human brain networks has become a central focus in neuroscience, particularly in the study of functional connectivity, which plays a crucial role in diagnosing neurological disorders. Advances in functional magnetic resonance imaging and machine learning techniques have significantly improved brain network analysis. However, traditional machine learning approaches struggle to capture the complex relationships between brain regions, while deep learning methods, particularly Transformer-based models, face computational challenges due to their quadratic complexity in long-sequence modeling. To address these limitations, we propose a Core-Periphery State-Space Model (CP-SSM), an innovative framework for functional connectome classification. Specifically, we introduce Mamba, a selective state-space model with linear complexity, to effectively capture long-range dependencies in functional brain networks. Furthermore, inspired by the core-periphery (CP) organization, a fundamental characteristic of brain networks that enhances efficient information transmission, we design CP-MoE, a CP-guided Mixture-of-Experts that improves the representation learning of brain connectivity patterns. We evaluate CP-SSM on two benchmark fMRI datasets: ABIDE and ADNI. Experimental results demonstrate that CP-SSM surpasses Transformer-based models in classification performance while significantly reducing computational complexity. These findings highlight the effectiveness and efficiency of CP-SSM in modeling brain functional connectivity, offering a promising direction for neuroimaging-based neurological disease diagnosis. Our code is available at https://github.com/m1nhengChen/cpssm

Keywords: Functional connectivity · Core-periphery · State space model.

1 Introduction

The vast assemblage of neurons in human brain forms a complex, interconnected network that achieves a remarkable balance between regional specialization and

global functional integration, enabling diverse cognitive and behavioral processes [9]. Understanding the organization of these neural networks has become a central focus of contemporary neuroscience. By studying and analyzing brain networks, neuroscientists can gain deeper insights into the structural and functional architecture of the human brain, as well as how network dynamics influence the onset, expression, and manifestation of neurological diseases [2,12]. Advancements in neuroimaging, particularly functional magnetic resonance imaging (fMRI), have revolutionized brain network analysis by non-invasively assessing intrinsic functional connectivity through blood oxygen level-dependent (BOLD) signals [10]. Functional connectivity (FC), measured as the correlation of BOLD signals across brain regions, has been widely explored for diagnosing neurological diseases and identifying potential biomarkers [13]. Machine learning techniques are commonly used to classify FC patterns, but traditional approaches struggle to capture the intricate relationships between brain regions. Deep learning has proven more effective in modeling brain functional connectivity [16,29,31]. Notably, Transformer-based architectures, originally designed for natural language processing, have gained traction in medical imaging due to their ability to capture long-range dependencies. Recent studies show that Transformer models significantly outperform other learning-based methods in classifying neurological disorders [15,17,21,30].

Although existing methods can effectively model functional connectivity patterns of the human brain and achieve competitive results in diagnosing neurological diseases such as Alzheimer's disease (AD), they still have the following shortcomings: First, existing works rely heavily on increasingly complex network architectures, which not only lead to overfitting problems due to inductive bias, but also hinder their application on long sequences by computational complexity such as the quadratic computational cost associated with the attention mechanism. These challenges have spurred interest in more computationally efficient alternatives to Transformers, aiming to reduce complexity while preserving the ability to capture long range dependencies and maintain strong representational learning capabilities. Second, current network models often overlook the intrinsic characteristics of brain function during computational modeling, resulting in suboptimal performance in brain network analysis. This gap underscores the need for approaches that align more closely with the unique functional and structural properties of the brain.

In this paper, we address the aforementioned limitations by introducing a core-periphery principle guided state-space model (CP-SSM) for functional connectome classification. Specifically, 1) we introduce a highly promising long-sequence modeling method with linear complexity based on a selective state space model named Mamba [11] to capture the functional connectivity relationships between different brain regions and identify variations in functional brain patterns across individuals. 2) Core-periphery (CP) organization is a ubiquitous feature of the brain functional network in humans and other mammals. It has been widely demonstrated to enhance the efficiency of information transmission and communication in biological integration processes [7]. Drawing inspiration

from this principle, we propose CP-MoE, which leverages CP organization to guide the redesign of Mixture-of-Experts (MoE) model, thereby effectively enhancing the representation capabilities of the networks. 3) Experimental results on the ABIDE and ADNI datasets demonstrate that the proposed method surpasses existing Transformer-based approaches while achieving lower computational complexity, highlighting its effectiveness and efficiency.

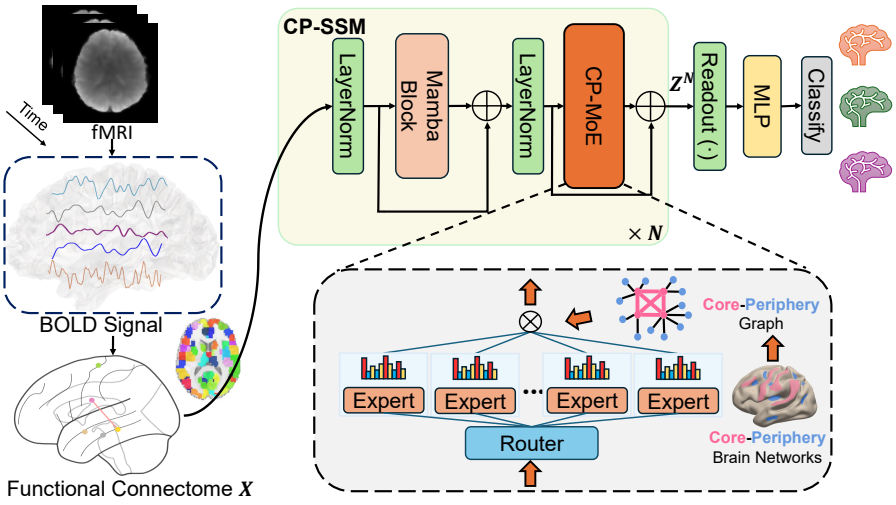


Fig. 1. Overall architecture of the proposed method. Our approach is founded on a state-space model, with its key components comprising a Mamba block and a Core-Periphery principle guided MoE.

2 Material and Method

The overall pipeline of the proposed method is shown in Fig. 1. The FC matrix $X \in \mathbb{R}^{V\times V}$ (V is the number of ROIs), derived by calculating the Pearson cross-correlation of the preprocessed BOLD signals across different brain regions, is first processed through N CP-SSM blocks to generate the node-level embedding representation $Z \in \mathbb{R}^{V\times V}$. Each CP-SSM block comprises two key components: an SSM module and a CP-MoE block. Subsequently, the output is passed through an Orthonormal Clustering Readout function [15] to obtain the graph-level embedding. The embedding is then fed into a fully connected layer, with the final probability predictions generated through a softmax layer. The model is trained using cross-entropy loss.

2.1 Data Acquisition and Preprocessing

In this study, we evaluated the proposed method on two fMRI datasets, assessing its performance in diagnosing two neurological conditions: autism spectrum disorder (ASD) and mild cognitive impairment (MCI)—the prodromal stage of AD. (a) *Autism Brain Imaging Data Exchange* (ABIDE) [5]: This dataset collects resting-state fMRI data from 17 international sites. Based on the given quality control scores, 1009 subjects (516 with ASD and 493 with NC) from 1,112 subjects were selected. The dataset, preprocessed by the Configurable Pipeline for the Analysis of Connectomes tool, underwent band-pass filtering (0.01 - 0.1Hz) without global signal regression. The brain was parcellated using the Craddock 200 atlas [6]. (b) *Alzheimer's Disease Neuroimaging Initiative*(ADNI) [18]: 440 subjects(215 MCI, 225 NC) were selected based on quality control. Each subject's data underwent the same standard preprocessing procedures as detailed in [32]. The Destrieux Atlas was then applied for parcellation.

2.2 State Space Model for Functional Connectome Classification

A State Space Model (SSM) is commonly defined as a linear time-invariant system that maps a one-dimensional input sequence $x(t) \in \mathbb{R}$ to an output response $y(t) \in \mathbb{R}$ through an underlying latent state representation $h(t) \in \mathbb{R}^M$, as formalized in Eq. 1, where $A \in R^{M\times M}$, $B \in \mathbb{R}^{M\times 1}$, $C \in \mathbb{R}^{1\times M}$, and $D \in \mathbb{R}^1$ are the weighting parameters.

$$\begin{aligned} h'(t) &= \mathbf{A}h(t) + \mathbf{B}u(t), \\ y(t) &= \mathbf{C}h(t) + \mathbf{D}u(t), \end{aligned} \tag{1}$$

Mamba [11], a recently proposed variant of SSM, exhibits a remarkable capacity for processing long sequential data with linear computational complexity, making it highly efficient in modeling local interactions between adjacent nodes. In this study, we propose utilizing Mamba modules to characterize brain network connectivity patterns, enabling the capture of long-range dependencies across brain regions that would otherwise be computationally prohibitive to model using self-attention mechanisms. The input functional connectivity matrix X first undergoes Layer Normalization, producing X_n. Subsequently, X_n is processed through a Mamba block, which consists of a linear layer, a SiLU [8] activation function, a one-dimensional convolution layer, and an SSM layer [3]. The output $X' \in \mathbb{R}^{V\times V}$ is then combined with X_n via element-wise addition to generate the latent representation $H \in \mathbb{R}^{V\times V}$. This entire process can be mathematically formulated as follows:

$$\begin{aligned} X' &= SSM(conv(SiLU(linear(X_n))) \\ H &= X' + X_n \end{aligned} \tag{2}$$

In our architecture, the Feedforward Network (FFN) commonly used in Transformers is replaced with the proposed CP-MoE, a widely adopted approach in the field of Sparse MoE (SMoE) [28].The details of CP-MoE will be introduced

later. After normalization, the latent representation H is transformed into H_n. Next, H_n is processed through CP-MoE, yielding H'. The output $H'$is then combined with H_n via element-wise addition to generate the node-level embedding Z. Since our model consists of N cascaded CP-SSM, this process is iteratively repeated N times to obtain the final, more expressive node feature $Z^N \in \mathbb{R}^{V \times V}$.

2.3 Core-periphery Principle Guided Mixture-of-Experts

Mixture-of-Experts improves computational efficiency by integrating sparsity, with SMoE further enhancing scalability by selectively activating relevant experts, reducing computational overhead. Core-periphery is a key structural feature of brain networks, where densely connected core nodes facilitate efficient processing, while sparsely connected peripheral nodes support network integration [7,23,26]. This architecture enhances information transfer and biological integration. Motivated by this principle, we aim to redesign the expert selection mechanism in MoE under the guidance of CP principle. Specifically, we propose leveraging the CP framework to guide the expert assignment process, ensuring a more structured and biologically inspired approach to model the expert selection.

First, we construct a CP graph G to facilitate expert selection based on the CP principle. To achieve this, we introduce a parameter core node rate r, which partitions the graph's nodes into two distinct sets: core nodes $\mathcal{C}$ and peripheral nodes $\mathcal{P}$. Notably, when $r = 1$, the resulting CP graph becomes fully connected. Consistent with the definition in [25,27], G can be represented as:

$$G(i,j) = \begin{cases} 1, & \text{if } (i,j) \in \mathcal{C} \times \mathcal{C} \text{ or } (i,j) \in \mathcal{C} \times \mathcal{P} \text{ or } (i,j) \in \mathcal{P} \times \mathcal{C} \\ 0, & \text{if } (i,j) \in \mathcal{P} \times \mathcal{P} \end{cases}$$

Thus, the Top-k expert selection of the router $\mathcal{R}(\cdot)$ in MoE, determined by the highest scores from softmax function with the learnable gating function $g(\cdot)$, can be formally expressed as follows:

$$\mathbf{y} = G \cdot \sum_{i=1}^{E} \mathcal{R}(\mathbf{x})_i \cdot f_i(\mathbf{x}) \tag{3}$$

$$\mathcal{R}(\mathbf{x}) = \text{Top-}k(\text{softmax}(g(\mathbf{x})), k) \tag{4}$$

$$\text{Top-}k(\mathbf{v}, k) = \begin{cases} \mathbf{v}, & \text{if } \mathbf{v} \text{ is in the top } k, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

where $\mathbf{x}$ and $\mathbf{y}$ denote the input and output of MoE, respectively. E represents the total number of experts, and i denotes the index of a specific expert.

3 Experiments

3.1 Experiment Settings

We partition each dataset into 70% for training, 10% for validation, and 20% for testing. For evaluation on the test set, we select the epoch that achieves

the highest AUROC score on the validation set. We evaluate the performance of our proposed method against several baseline approaches, including 3 traditional machine learning methods: support vector machine (SVM), random forest (RF), and XGBoost; 2 CNN/GNN-based methods: FBNETGEN [14] and BrainNetCNN [16]; and 4 Transformer-based methods: VanillaTF [15], BrainNetTF [15], Com-BrainTF [1], and GBT [21].

Implementation details. Our method is configured with a state-space dimension of 16 in the Mamba block, an expansion factor of 2, two CP-SSM blocks ($N = 2$), and a convolution kernel dimension of 4. The CP-MoE module consists of 8 experts, utilizing a Top-4 selection strategy. For dataset-specific settings, we set the core node rate to 0.2 for the ABIDE and 0.8 for ADNI. The CP-SSM model is trained using the Adam optimizer with an initial learning rate of 10e-4 and a weight decay of 10e-4. A cosine annealing schedule is applied, gradually reducing the learning rate from 10e-4 to 10e-5 without a warm-up phase. Training is conducted over 200 epochs with a batch size of 64. All experiments were performed on a PC equipped with an NVIDIA RTX 6000 Ada GPU and a 3.6-GHz Intel Core i7 processor.

Table 1. Performance comparison with different baselines on ADNI and ABIDE.

Methods	**Dataset: ABIDE**				**Dataset: ADNI**			
	AUROC	ACC	SEN	SPE	AUROC	ACC	SEN	SPE
SVM	70.4±5.2	63.3±5.2	64.8±7.1	61.6±7.0	65.1±8.2	61.5±4.4	51.2±7.9	69.7±8.2
RF	69.2±4.3	63.8±3.4	71.0±5.2	56.1±5.2	67.9±3.8	63.9±1.2	55.9±4.7	71.4±4.5
XGBoost	71.2±4.4	63.4±5.1	68.6±4.9	57.8±9.1	65.4±5.0	62.7±1.1	61.5±6.0	63.8±6.2
FBNETGNN	72.9±5.1	65.7±5.6	64.3±10.6	66.6±8.2	69.1±7.9	66.3±3.9	66.7±8.1	65.7±8.4
BrainNetCNN	73.2±3.0	66.6±4.0	64.6±6.2	68.7±4.8	65.8±1.0	65.4±5.2	60.7±1.3	68.7±4.6
VanillaTF	79.6±4.6	69.8±6.0	64.1±8.1	**76.4±9.1**	73.1±6.4	69.3±3.8	68.7±8.1	70.3±8.8
BrainNetTF	79.1±4.8	70.1±4.9	67.9±5.0	72.2±6.6	73.0±7.4	70.3±4.7	69.7±7.6	69.9±7.7
Com-BrainTF	77.3±4.1	71.6±4.5	75.1±11.9	67.4±9.3	-	-	-	-
GBT	78.3±4.1	71.5±5.8	75.5±14.7	68.2±12.8	74.5±5.5	71.1±3.2	69.0±7.8	73.6±5.1
CP-SSM	**82.4±2.2**	**76.9±1.4**	**79.3±8.8**	74.5±8.4	**80.7±2.6**	**74.9±2.6**	**75.4±5.7**	**74.2±2.6**

3.2 Performance Comparison with Baseline Methods

We evaluate the proposed CP-SSM against baseline methods on ASD classification using the ABIDE dataset and MCI classification using the ADNI dataset. Each experiment is repeated ten times per dataset, and the disease classification results are reported. Notably, the brain network community division in Com-TF follows the Yeo 7-network template [24], which is based on brain network analysis in young individuals. This makes it well-suited for the ABIDE dataset (ages 7–64, median 14.7). However, since the ADNI dataset consists of older participants (ages 50.5–82.8), this template is less appropriate. Therefore, we do not report Com-TF results for ADNI. The results in Table 1 demonstrate that CP-SSM consistently outperforms all baselines on both ABIDE and ADNI, highlighting its superior capability in modeling functional connectivity for neurological disease diagnosis compared to both traditional machine learning and deep learning-based approaches.

3.3 Ablation Study

Table 2. Ablation study of CP-SSM on ABIDE with the best and second-best values in **boldface** and underline, respectively.

	AUC	ACC	SEN	SPE
CP-SSM	**82.4±2.2**	**76.9±1.4**	79.3±8.8	**74.5±8.4**
w/o CP	79.9±3.3	75.9± 2.3	**81.9±7.3**	69.2±8.9
w/o MoE	81.1±3.1	74.1±1.6	75.7±5.7	72.5±8.0
w/o CP-MoE	80.1±2.3	73.0±1.3	72.4±4.5	73.0±1.3
w/o SSM	81.1±2.2	74.7±0.8	77.3±5.0	72.3±5.3
learnable mask	79.4±2.3	74.0±1.7	73.5±6.2	**74.5±4.9**

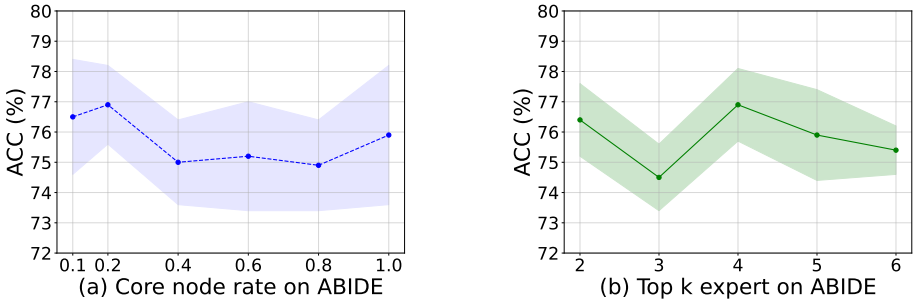
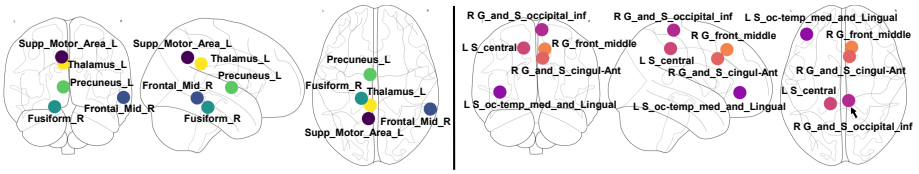


Fig. 2. Sensitivity analysis of CP-SSM on ABIDE. The hyperparameters include core node rate r and top-k expert selection.

We perform ablation studies to assess the effectiveness of each component in the design of the proposed CP-SSM, with the corresponding experimental results presented in Table 2. 1) w/o CP: the CP mask in CP-MoE is removed, making the MoE structure equivalent to other SMoE methods [22,28]; 2) w/o MoE: directly replace all MoE layers in the network with MLPs with the same number of neurons; 3) w/o CP-MoE: CP-MoE is replaced by FFN in the common Transformer architecture; 4) w/o SSM: replace SSM with the same transformer block as implemented in [15]; 5) learnable mask: substitute the core-periphery guided mask with a weighted learnable mask. Furthermore, replacing the SSM module with a Transformer block led to an increase in the number of model parameters from 11.6M to 12.2M, while classification accuracy decreased by 2%. This finding highlights the computational efficiency of the SSM module in maintaining model performance while reducing parameter complexity. The

sensitivity analysis of the proposed method on ABIDE, presented in Fig. 2, explores the impact of the core node rate r and the top-k selection.

3.4 Neuroscientific Analysis



(a) Top 5 discriminative brain regions on ABIDE in ASD diagnosis

(b) Top 5 discriminative brain regions on ADNI in MCI diagnosis

Fig. 3. Top 5 discriminative brain regions derived from the learnable weight in the last CP-SSM layer on (a) ABIDE and (b) ADNI, with different colormap intensities reflecting relative significance.

As shown in Fig 3, we visualize the learnable weights of the last layer of CP-SSM block to show the top 5 rated brain regions for ASD diagnosis and MCI diagnosis, respectively. In the figure, we present the brain region names from the Automated Anatomical Labeling (AAL) atlas that exhibit the highest overlap with each corresponding region.

- **ASD analysis**: the identified regions, including the Supplementary Motor Area (SMA), Thalamus, Precuneus, Fusiform Gyrus, and Middle Frontal Gyrus, are associated with social cognition and motor functions. Studies indicate reduced connectivity in the Precuneus and Middle frontal Gyrus in individuals with ASD, which affects social behavior and theory of mind [4].
- **MCI Analysis**: the key regions identified, including the Inferior Occipital Gyrus, Middle Frontal Gyrus, Anterior Cingulate Gyrus, Central Sulcus, and Occipito Temporal and Lingual Gyrus, show strong associations with neurodegeneration and cognitive decline. The findings of the discrimination region are consistent with existing studies [19,20].

These findings corroborate existing research, highlighting the involvement of these brain regions in ASD and MCI.

4 Conclusion

In this paper, we propose a core-periphery principle guided state-space model for functional connectome classification. We utilize the Mamba block to model long-range dependencies in brain connectivity and incorporate the core-periphery principle to guide the MoE. To the best of our knowledge, this work represents

the first adaptation of the Mamba and MoE architectures for functional brain network analysis. Extensive experimental results demonstrate that our method not only outperforms other deep learning approaches but also offers strong interpretability, making it a robust and insightful framework for functional brain network analysis. Future work will focus on validating the effectiveness of our method across a broader range of datasets and assessing its performance on tasks beyond classification.

Acknowledgments. This work was supported by the National Institutes of Health (R01AG075582 and RF1NS128534)

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: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 287–297. Springer (2023)
2. Chen, M., Cao, C., Chen, T., Zhuang, Y., Zhang, J., Lyu, Y., Yu, X., Zhang, L., Liu, T., Zhu, D.: Using structural similarity and kolmogorov-arnold networks for anatomical embedding of cortical folding patterns. In: 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). pp. 1–5 (2025)
3. Chen, S., Atapour-Abarghouei, A., Zhang, H., Shum, H.P.H.: Mxt: Mamba x transformer for image inpainting. In: 35th British Machine Vision Conference 2024, BMVC 2024, Glasgow, UK, November 25-28, 2024. BMVA (2024)
4. Cheng, W., Rolls, E.T., Gu, H., Zhang, J., Feng, J.: Autism: reduced connectivity between cortical areas involved in face expression, theory of mind, and the sense of self. Brain **138**(5), 1382–1393 (2015)
5. 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)
6. 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)
7. Csermely, P., London, A., Wu, L.Y., Uzzi, B.: Structure and dynamics of core/periphery networks. Journal of Complex Networks **1**(2), 93–123 (2013)
8. Elfwing, S., Uchibe, E., Doya, K.: Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural networks **107**, 3–11 (2018)
9. Fornito, A., Zalesky, A., Breakspear, M.: Graph analysis of the human connectome: promise, progress, and pitfalls. Neuroimage **80**, 426–444 (2013)
10. Glover, G.H.: Overview of functional magnetic resonance imaging. Neurosurgery Clinics **22**(2), 133–139 (2011)
11. Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752 (2023)

12. van den Heuvel, M.P., Sporns, O.: A cross-disorder connectome landscape of brain dysconnectivity. Nature reviews neuroscience **20**(7), 435–446 (2019)
13. Hohenfeld, C., Werner, C.J., Reetz, K.: Resting-state connectivity in neurodegenerative disorders: Is there potential for an imaging biomarker? NeuroImage: Clinical **18**, 849–870 (2018)
14. Kan, X., Cui, H., Lukemire, J., Guo, Y., Yang, C.: Fbnetgen: Task-aware gnn-based fmri analysis via functional brain network generation. In: International Conference on Medical Imaging with Deep Learning. pp. 618–637. PMLR (2022)
15. 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)
16. 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)
17. Kong, Y., Zhang, X., Wang, W., Zhou, Y., Li, Y., Yuan, Y.: Multi-scale spatial-temporal attention networks for functional connectome classification. IEEE Transactions on Medical Imaging (2024)
18. Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L.: The alzheimer's disease neuroimaging initiative. Neuroimaging Clinics of North America **15**(4), 869 (2005)
19. Palejwala, A.H., O'Connor, K.P., Milton, C.K., Anderson, C., Pelargos, P., Briggs, R.G., Conner, A.K., O'Donoghue, D.L., Glenn, C.A., Sughrue, M.E.: Anatomy and white matter connections of the fusiform gyrus. Scientific reports **10**(1), 13489 (2020)
20. Peng, G., Wang, J., Feng, Z., Liu, P., Zhang, Y., He, F., Chen, Z., Zhao, K., Luo, B.: Clinical and neuroimaging differences between posterior cortical atrophy and typical amnestic alzheimer's disease patients at an early disease stage. Scientific reports **6**(1), 29372 (2016)
21. Peng, Z., He, Z., Jiang, Y., Wang, P., Yuan, Y.: Gbt: Geometric-oriented brain transformer for autism diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 142–152. Springer (2024)
22. Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., Dean, J.: Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. arXiv preprint arXiv:1701.06538 (2017)
23. Shu, Z., Yu, X., Wu, Z., Jia, W., Shi, Y., Yin, M., Liu, T., Zhu, D., Niu, W.: Real-time core-periphery guided vit with smart data layout selection on mobile devices. Advances in Neural Information Processing Systems **37**, 95744–95763 (2024)
24. Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology (2011)
25. Yu, X., Wu, Z., Zhang, L., Zhang, J., Lyu, Y., Zhu, D.: Cp-clip: Core-periphery feature alignment clip for zero-shot medical image analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 88–97. Springer (2024)
26. Yu, X., Zhang, L., Cao, C., Chen, T., Lyu, Y., Zhang, J., Liu, T., Zhu, D.: Gyri vs. sulci: Core-periphery organization in functional brain networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 164–174. Springer (2024)

27. Yu, X., Zhang, L., Wu, Z., Zhu, D.: Core-periphery multi-modality feature alignment for zero-shot medical image analysis. IEEE Transactions on Medical Imaging (2024)
28. Yun, S., Choi, I., Peng, J., Wu, Y., Bao, J., Zhang, Q., Xin, J., Long, Q., Chen, T.: Flex-moe: Modeling arbitrary modality combination via the flexible mixture-of-experts. In: The Thirty-eighth Annual Conference on Neural Information Processing Systems (2024)
29. Zhang, J., Lyu, Y., Yu, X., Zhang, L., Cao, C., Chen, T., Chen, M., Zhuang, Y., Liu, T., Zhu, D.: Classiffication of mild cognitive impairment based on dynamic functional connectivity using spatio-temporal transformer. In: 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). pp. 1–5 (2025)
30. Zhang, J., Yu, X., Lyu, Y., Zhang, L., Chen, T., Cao, C., Zhuang, Y., Chen, M., Liu, T., Zhu, D.: Brain-adapter: Enhancing neurological disorder analysis with adapter-tuning multimodal large language models. In: 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). pp. 1–5 (2025)
31. Zhou, H., He, L., Chen, B.Y., Shen, L., Zhang, Y.: Multi-modal diagnosis of alzheimer's disease using interpretable graph convolutional networks. IEEE Transactions on Medical Imaging (2024)
32. Zhu, D., Li, K., Terry, D.P., Puente, A.N., Wang, L., Shen, D., Miller, L.S., Liu, T.: Connectome-scale assessments of structural and functional connectivity in mci. Human brain mapping **35**(7), 2911–2923 (2014)