

Progressive Self-Supervised Learning with Individualized Community Assignment for Brain Network Analysis

Hairui Chen^{1,2}, Yanwu Yang^{3,4}, Jianfeng Cao¹, Hanyang Peng², Chenfei Ye¹, and Ting Ma¹(✉)

¹ Harbin Institute of Technology at Shenzhen, Shenzhen, China
tma@hit.edu.cn

² Peng Cheng Laboratory, Shenzhen, China

³ University of Tübingen, Tübingen, Germany

⁴ Max Planck Institute for Intelligent Systems, Tübingen, Germany

Abstract. Brain networks exhibit a modular community structure that varies across individuals and neurological conditions. However, existing self-supervised learning (SSL) methods often overlook this heterogeneity, relying on generic masking strategies that fail to capture subject-specific functional organization. We propose BrainPICM, a self-supervised framework for brain network analysis via progressive individualized community-aware masking. BrainPICM formulates ROI-to-community mapping as a progressive unbalanced optimal transport process, yielding soft assignments and per-ROI confidence scores. Guided by these confidence estimates, a curriculum-style masking strategy gradually incorporates low-confidence, potentially pathological regions into training, enabling the model to learn both stable modular structures and individual variations. Additionally, a deviation-aware aggregation module quantifies functional reorganization by measuring mass redistribution relative to a population template, enhancing interpretability and downstream prediction. Experiments on three fMRI datasets (ABIDE-I, ADHD-200, ADNI) show that BrainPICM consistently outperforms state-of-the-art supervised and SSL methods in diagnostic accuracy, indicating that explicitly injecting modular community structure into masked modeling yields more functionally consistent and generalizable representations. The source code for this approach will be released at <https://github.com/Hrychen7/BrainPICM>.

Keywords: Self-supervised Learning · Brain Network Analysis · Individualized Community Assignment.

1 Introduction

Brain networks have become a pivotal paradigm in cognitive neuroscience, enabling the characterization of complex relationships between brain abnormalities

H. Chen and Y. Yang contributed equally to this work.

and behavioral changes through network-based modeling [24]. Functional magnetic resonance imaging (fMRI) measures blood-oxygen-level-dependent (BOLD) signals to investigate the functional organization of brain regions and their interactions [23]. A fundamental property of functional brain networks is their modular community organization, which supports coordinated information processing [21, 27]. Importantly, these communities vary across subjects and may be reorganized in disease states, reflecting subject-specific functional heterogeneity [16, 7, 20].

Learning robust representations from functional connectivity is challenging. In addition to the modular community organization that varies across individuals, brain networks are high-dimensional, training samples are often limited, and substantial inter-subject variability exists across cohorts and disorders [22]. These factors can make purely supervised models prone to overfitting and less reliable generalization. Compared with purely supervised learning, self-supervised learning (SSL) provides a promising alternative by exploiting intrinsic structure in unlabeled networks, which can improve generalization while reducing reliance on large annotated datasets.

Among SSL paradigms, masked modeling has recently gained popularity for brain network representation learning, where models are trained to reconstruct functional connectivity from masked ROIs. Existing approaches mainly differ in masking policies, ranging from uniform random masking [13, 25] to attention-guided strategies [29]. While effective in learning generic features, these masking policies may not explicitly incorporate functional community structure. Consequently, the learned representations may be less functionally consistent and may generalize less reliably when community organization varies across individuals.

To address this gap, we propose BrainPICM, a self-supervised framework for **Brain** network analysis using **P**rogressive **I**ndividualized **C**ommunity-aware **M**asking to learn personalized yet consistent representations. Specifically, BrainPICM models ROI-to-community mapping via a progressive unbalanced optimal transport process, yielding soft subject-specific assignments and confidence scores while robustly handling outlier regions. Guided by these scores, a curriculum-inspired masking strategy gradually introduces low-confidence, potentially pathological ROIs into training to capture individual heterogeneity. Additionally, we summarize the soft assignments into compact community-level deviation features and fuse them with ROI-level connectivity features for downstream diagnosis. Our contributions are summarized below:

- (1) We propose BrainPICM, a self-supervised framework with community-guided curriculum masking driven by soft ROI-to-community assignments and ROI confidence.
- (2) We formulate individualized soft community assignment as a progressive unbalanced optimal transport (UOT) process with a virtual community, yielding Q and ROI-level confidence q under functional heterogeneity.
- (3) We derive compact community-level deviation features from Q and fuse them with ROI-level connectivity features in a lightweight dual-branch predictor for improved diagnosis.

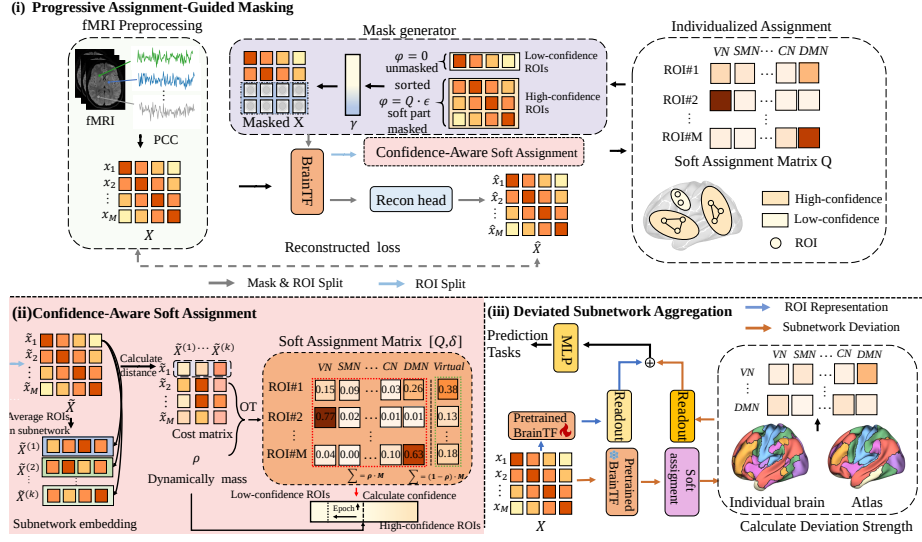


Fig. 1. Overview of BrainPICM. The framework consists of: (i) confidence-aware masking guided by soft assignment; (ii) progressive UOT-based soft assignment deriving ROI-level confidence via a virtual subnetwork; and (iii) deviation-aware subnetwork aggregation capturing functional reorganization for prediction. Recon: Reconstructed. PCC: Pearson Correlation Coefficient. X : Brain network. x_i : the i -th ROI with M features. $\hat{X}$: Reconstructed brain network. ψ : Masking probability. ρ : Dynamic transport mass. ε : Random subnetwork noise vector. γ : Binary mask vector.

2 Method

2.1 Overview

As illustrated in Fig. 1, the proposed BrainPICM is a progressive self-supervised framework that explicitly injects functional community structure into masked modeling of brain networks. In pre-training, the encoder learns ROI-wise representations and simultaneously estimates soft ROI-to-community affiliations with a progressive unbalanced optimal transport (UOT) solver. The resulting soft affiliations provide (i) soft ROI-to-community assignment and (ii) ROI-level confidence scores, which together drive a curriculum-style masking policy: training starts from reliable community-consistent ROIs and gradually incorporates uncertain, potentially pathological ROIs. After pre-training, the encoder is frozen to provide stable community guidance. For downstream diagnosis, we further summarize how each subject deviates from a population-level community reference, and fuse ROI-level connectivity features with community-deviation signals in a lightweight dual-branch predictor. This design enables BrainPICM to learn representations that are both functionally consistent and sensitive to subject-specific heterogeneity.

2.2 Problem Definition

We represent the brain functional network as a symmetric matrix $X \in \mathbb{R}^{M \times M}$ based on an atlas with M Regions of Interest (ROIs). Each ROI is treated as a sequence element $x_i = X_{i,:} \in \mathbb{R}^M$, where the M -dimensional vector reflects its connectivity profile across the whole brain. In this study, we use the Yeo ICN parcellation [27] to define $K = 7$ canonical communities (DMN, VN, DAN, SAN, SMN, CN, LN), and construct a one-hot template assignment $Q_{\text{template}} \in \{0, 1\}^{M \times K}$. Our goal is to learn a Transformer encoder f_θ via masked reconstruction while explicitly leveraging this community organization. Specifically, we estimate soft ROI-to-community assignments $Q \in \mathbb{R}_+^{M \times K}$ and ROI-level confidence scores q , which guide curriculum-style masking; meanwhile, Q is summarized into compact community-level deviation features for downstream diagnosis.

2.3 Confidence-Aware Soft Assignment via Progressive UOT

We obtain the soft ROI-to-community assignment matrix $Q \in \mathbb{R}_+^{M \times K}$ via a progressive unbalanced optimal transport (UOT) formulation. Given a brain network X , the brain network Transformer (BrainTF) encoder [14, 25] extracts ROI-wise features, and we maintain an exponential moving average (EMA) feature bank $\tilde{X}$ for stable estimation. Based on the Yeo ICN parcellation, we compute the k -th community prototype on EMA features as $\tilde{x}^{(k)} = \frac{1}{|S_k|} \sum_{i \in S_k} \tilde{x}_i$, where S_k denotes the set of ROIs assigned to community k in the template. We then measure ROI-community affinity by cosine similarity $P_{i,k} = \cos(\tilde{x}_i, \tilde{x}^{(k)})$ and construct the transport cost matrix $C = -\log \text{softmax}(P)$. To accommodate uncertain ROIs that do not align well with the K canonical communities, we introduce a virtual community to absorb the remaining assignment mass δ_i . This allows the soft assignment Q to be formulated as a progressive UOT problem, where ν is an atlas-derived prior distribution, λ controls the strength of the prior regularization, and $\rho(t) \in (0, 1]$ specifies the total mass transported to the K real communities. With this formulation, Q is obtained by the following objective:

$$\min_{Q \in \Pi} \langle Q, C \rangle_F + \lambda \text{KL} \left(Q^\top \mathbf{1}_M, \frac{\rho(t)}{K} \nu \right), \quad (1)$$

$$\text{s.t. } \Pi = \{Q \in \mathbb{R}_+^{M \times K} \mid Q \mathbf{1}_K \leq \nu, \mathbf{1}_M^\top Q \mathbf{1}_K = \rho(t)\}. \quad (2)$$

The remaining mass $1 - \rho(t)$ is assigned to the virtual community, yielding an ROI-level confidence $q_i = 1 - \delta_i$. We increase $\rho(t) = \rho_0 + (1 - \rho_0) \exp\left(-5 \left(1 - \frac{t}{T}\right)^2\right)$ progressively, producing a curriculum-style transition from confident to uncertain ROIs and generating Q and q for subsequent masking. To solve this unbalanced OT problem, we adopt an entropy-regularized UOT solver based on the efficient scaling algorithm [3, 28], where the soft assignment matrix is iteratively updated via Sinkhorn-style matrix scaling after introducing the virtual community.

2.4 Progressive Assignment-Guided Masking

Building upon the UOT-based assignments, we design a community-guided masking mechanism to inject functional community structure into masked reconstruction. Given X , we sample a binary mask $\gamma \in \{0, 1\}^M$ (with $\gamma_i = 1$ indicating that ROI i is masked) and form $X^{(t)}$ by replacing masked ROIs with a learnable mask token $x_{\text{mask}} \in \mathbb{R}^M$, followed by adding positional embeddings. The encoder f_θ and a two-layer MLP head g_ϕ reconstruct the full network as:

$$\hat{X} = g_\phi\left(f_\theta\left(X^{(t)}\right)\right), \quad (3)$$

optimized by a masked reconstruction loss:

$$\mathcal{L}_r = \frac{1}{\|\gamma\|_1} \sum_{i=1}^M \gamma_i \cdot \|\hat{X}_i - X_i\|^2. \quad (4)$$

To decide which ROIs to mask at step t , we compute a curriculum-style masking score using the soft assignment vector Q_i and confidence q_i :

$$\psi_i^{(t)} = \begin{cases} Q_i \cdot \epsilon, & \text{if } q_i \geq \text{Quantile}_{1-\rho(t)}(q), \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where $\epsilon \sim \mathcal{U}(0, 1)^K$ introduces community-level stochasticity. We then mask the top- $\lfloor \alpha M \rfloor$ ROIs with the largest $\psi^{(t)}$. As $\rho(t)$ increases, the masking gradually expands from high-confidence ROIs to more uncertain regions, encouraging the encoder to learn community-consistent yet individualized representations.

2.5 Deviated Subnetwork Aggregation

To further account for heterogeneity in functional community organization, we compute an assignment deviation matrix $\Delta Q = Q - Q_{\text{template}}$, where Q_{template} is the one-hot template assignment. We then aggregate ΔQ into a flow matrix $D \in \mathbb{R}^{K \times K}$:

$$D[i, j] = \sum_{n \in S_i} \Delta Q[n, j], \quad i, j = 1, \dots, K, \quad (6)$$

where S_i denotes the set of ROIs assigned to template community i , and we set $D[i, i] = 0$.

For downstream disease diagnosis, the frozen encoder f_θ extracts global connectivity features z_X via a readout operation (e.g., mean pooling over ROIs). Simultaneously, the flow matrix D is flattened and processed by an independent MLP to generate deviation representations z_D . The final diagnostic prediction is:

$$\hat{y} = \text{Softmax}(\text{MLP}(z_X + z_D)). \quad (7)$$

This dual-branch design couples fine-grained ROI connectivity with community-level deviation cues, enabling the predictor to jointly model local patterns and higher-order reorganization.

3 Experiments

3.1 Experimental Setup

Datasets. In this study, we conduct experiments on three real-world fMRI datasets, selected for their coverage of a spectrum of neurodevelopmental and neurodegenerative disorders and their widespread use. (a) Autism Brain Imaging Data Exchange (ABIDE-I) dataset [6]. A total of 1114 subjects are obtained in this dataset, comprising 586 Normal Controls (NC) and 528 subjects diagnosed with Autism spectrum disorder (ASD). (b) Attention Deficit Hyperactivity Disorder (ADHD-200) dataset [1]. This dataset includes 1235 subjects, comprising 711 NC and 524 patients diagnosed with ADHD. (c) Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset [12]. Our analysis focused on 142 NC and 149 patients diagnosed with Alzheimer’s disease (AD). All fMRI images were preprocessed using the Configurable Pipeline for the Analysis of Connectomes (C-PAC) pipeline [4]. This processing includes skull stripping, slice timing correction, motion correction, global mean intensity normalization, regression of nuisance signals using 24 motion parameters, band-pass filtering (0.01-0.08 Hz) and parcellated by Schaefer atlas [19] into 100 ROIs. Pearson Correlation Coefficient was applied to measure functional connectivity.

Baselines. We compare our method with various baselines, including supervised deep learning models (BrainNetCNN [15], BrainNetTF [14], Com-BrainTF [2], CAGT [18], BrainHGT [17], vanillaTF), and self-supervised learning frameworks (MAE [10], EvolvedMask [8], EAG-RS [13], BrainMass [25], TARDRL [29]).

Implementation Details. In this study, all experiments were conducted on the PyTorch 2.4.1 platform using 4 NVIDIA 2080Ti GPUs with 11GB memory. For pretraining, we trained the model for 600 epochs using Adam optimizer with an initial learning rate of $3e-4$ and a weight decay of $5e-5$. For hyperparameters, we set the masking ratio α to 0.1 and λ as 1. The initial value of ρ is dataset-specific: 0.6 for ADNI, 0.8 for ABIDE, and 0.9 for ADHD. In this study, we determined these values by a grid search with a step of 0.1. In the downstream tasks, the model is trained for 50 epochs by using an early stopping strategy. To ensure fairness, we adopt a stratified sampling strategy that accounts for site-specific distributions when splitting the data into training (70%), validation (10%), and testing (20%) sets, following established practice [14]. We evaluated both overall model performance and diagnostic relevance, reporting diagnosis accuracy (ACC) and area under the receiver operating characteristic curve (AUC), averaged over 10 random runs, with both mean and standard deviation on the test dataset.

3.2 Results

Table 1 presents the averaged classification results (ACC and AUC) of our BrainPICM. Our BrainPICM model achieves the best overall accuracy on all three datasets, with relative improvements of 1.71%, 3.13%, and 3.05% over the

Table 1. Classification performance on ADHD-200, ABIDE-I and ADNI (Mean $\pm$ Std). The best results are shown in bold and the second best are underlined.

Category	Method	ADHD-200		ABIDE-I		ADNI	
		ACC $\uparrow$	AUC $\uparrow$	ACC $\uparrow$	AUC $\uparrow$	ACC $\uparrow$	AUC $\uparrow$
Supervised	BrainNetCNN	<u>64.24$\pm$2.36</u>	66.79 $\pm$ 2.31	67.76 $\pm$ 2.09	73.52 $\pm$ 2.87	74.30 $\pm$ 2.06	81.97 $\pm$ 1.78
	vanillaTF	62.65 $\pm$ 2.51	66.19 $\pm$ 2.46	67.51 $\pm$ 2.27	72.58 $\pm$ 2.27	74.57 $\pm$ 6.91	81.72 $\pm$ 3.94
	BrainNetTF	63.86 $\pm$ 2.71	67.77 $\pm$ 2.32	70.03 $\pm$ 1.96	74.38 $\pm$ 2.17	77.12 $\pm$ 4.10	83.50 $\pm$ 4.20
	Com-BrainTF	63.51 $\pm$ 2.09	66.69 $\pm$ 2.09	68.86 $\pm$ 0.55	75.43 $\pm$ 0.86	76.13 $\pm$ 5.72	83.12 $\pm$ 5.06
	CAGT	62.93 $\pm$ 2.78	66.95 $\pm$ 3.53	68.27 $\pm$ 2.42	74.04 $\pm$ 2.69	76.61 $\pm$ 4.02	<u>84.17$\pm$4.41</u>
	BrainHGT	63.47 $\pm$ 0.88	<u>67.82$\pm$2.94</u>	68.81 $\pm$ 1.29	74.76 $\pm$ 2.64	<u>78.98$\pm$4.10</u>	83.96 $\pm$ 2.84
SSL	MAE	62.01 $\pm$ 2.18	63.79 $\pm$ 2.60	67.87 $\pm$ 1.41	74.17 $\pm$ 2.58	75.49 $\pm$ 2.49	81.63 $\pm$ 2.08
	EvolvedMask	62.92 $\pm$ 1.13	67.11 $\pm$ 1.71	66.96 $\pm$ 1.28	72.70 $\pm$ 1.59	73.72 $\pm$ 8.45	83.95 $\pm$ 0.95
	EAG-RS	62.97 $\pm$ 0.78	66.47 $\pm$ 2.98	69.90 $\pm$ 3.46	74.51 $\pm$ 3.05	73.56 $\pm$ 1.65	79.60 $\pm$ 2.87
	TARDRL	61.37 $\pm$ 1.37	61.33 $\pm$ 1.67	68.36 $\pm$ 0.19	72.45 $\pm$ 1.13	74.41 $\pm$ 3.51	80.86 $\pm$ 4.32
	BrainMass	64.12 $\pm$ 2.08	66.41 $\pm$ 1.78	<u>70.32$\pm$1.91</u>	<u>76.21$\pm$1.13</u>	76.78 $\pm$ 2.74	81.22 $\pm$ 3.01
Ours	BrainPICM	67.37$\pm$0.83	68.37$\pm$1.31	72.03$\pm$0.82	77.06$\pm$1.25	82.03$\pm$2.75	84.33$\pm$3.62

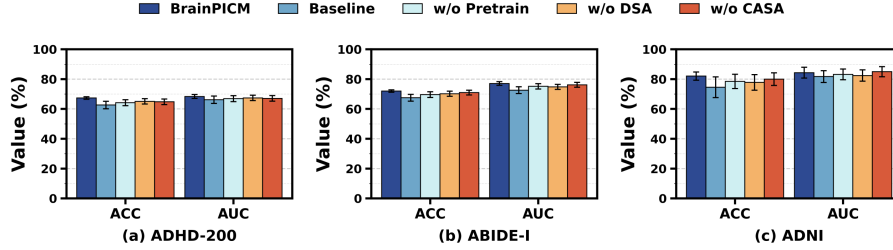


Fig. 2. Ablation studies on the elements of Our BrainPICM.

second-best methods on ABIDE-I, ADHD-200, and ADNI, respectively. In addition to accuracy, BrainPICM also demonstrates consistently high AUC across the datasets. While some SSL methods do not show significant advantages over BrainNetTF, our method outperforms existing SSL baselines. This performance gain is attributed to the integration of individualized subnetwork assignment, a confidence-aware masking strategy, and deviation-aware feature aggregation, which together enable more precise and robust representation learning of subject-specific brain network structures.

3.3 Model Analysis

Ablation Studies. Figure 2 evaluates BrainPICM’s core components: self-supervised pretraining, Confidence-Aware Soft Assignment (CASA), and Deviated Subnetwork Aggregation (DSA). Comparisons with a vanilla Transformer and a non-pretrained variant (using direct OT assignment) show that pretraining provides essential stability for individualized priors. Specifically, removing CASA increases assignment noise, while omitting DSA neglects critical group-to-individual reorganization signals. The consistent performance drop across

ρ_0	Sigmoid		Linear		Fixed	
	ACC $\uparrow$	AUC $\uparrow$	ACC $\uparrow$	AUC $\uparrow$	ACC $\uparrow$	AUC $\uparrow$
0.6	82.03	84.33	78.14	83.58	75.42	81.17
0.7	79.83	83.09	76.61	81.87	73.56	81.96
0.8	81.53	84.01	79.15	83.43	75.25	82.09
0.9	78.64	84.12	78.21	84.10	74.58	80.62

Table 2. Analysis of ρ_0 with different ramp-up strategies on the ADNI dataset. Fixed means ρ is held constant.

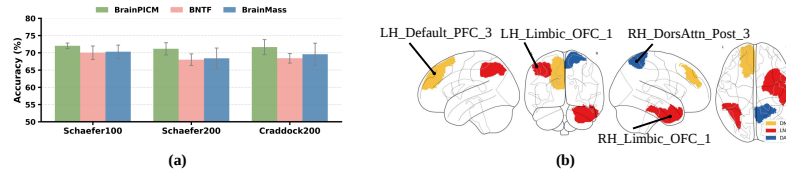


Fig. 3. (a) Classification results on ABIDE-I dataset using different brain atlases; (b) Statistically significant ROIs ($p < 0.05$) identified by multivariate analysis. LH: left hemisphere; RH: right hemisphere; PFC: prefrontal cortex; OFC: orbital frontal cortex.

all variants validates that these modules synergistically enhance representation learning.

Analysis of ρ We compared fixed, linear, and sigmoid scheduling for the transport mass ρ . Table 2 shows that dynamic scheduling consistently outperforms fixed settings. The sigmoid ramp achieves the best results, confirming that progressively increasing transport mass effectively balances initial structural stability with late-stage individual refinement.

Influence of Different Atlases We evaluated BrainPICM on the ABIDE-I dataset using different brain atlases (Schaefer 100 atlas, Schaefer 200 atlas, Craddock 200 atlas [5]). As shown in Fig. 3(a), BrainPICM achieves stable diagnostic accuracy across all atlases, demonstrating robustness to different parcellations.

3.4 Biological Explanation

Fig. 3(b) illustrates the ROIs identified through multivariate analysis [26], indicating the significant involvement of the Default Mode Network (DMN), Dorsal Attention Network (DAN), and Limbic Network (LN). Notably, these systems are closely related to social cognition, attentional control, and affective processing, respectively, suggesting that the detected alterations may reflect core functional disruptions in autism. Consistent with these findings, prior studies have also reported connectivity abnormalities within these specific networks in individuals with autism [9, 11].

4 Conclusion

In this paper, we introduced BrainPICM, a self-supervised framework that addresses the limitations of generic masking by injecting functional community structure into masked brain network modeling. By formulating ROI-to-community mapping as a progressive unbalanced optimal transport process, BrainPICM derives soft assignments and ROI-level confidence scores to drive curriculum-style community-guided masking from reliable to uncertain regions. We further summarize the soft assignments into compact community-level deviation features and fuse them with ROI-level connectivity features in a lightweight dual-branch predictor. Experiments on three fMRI datasets demonstrate that BrainPICM consistently outperforms state-of-the-art supervised and self-supervised baselines in diagnostic accuracy.

Acknowledgments. This study is supported by grants from the National Key Research and Development Program of P.R. China (2025YFF0517803), the National Natural Science Foundation of P.R. China (62276081), Guangdong S&T Programme (2025B0101130004), and the Shenzhen Science and Technology Program (CJGJZD 20230724093959002, JCYJ20250604145427037). The funders only provided financial support for this study, and did not participate in the study design, data collection and analysis, manuscript writing, or submission decision-making. YY gratefully acknowledge support from the Alexander von Humboldt Foundation for his research fellowship.

Disclosure of Interests. The authors declare no competing financial or personal interests that could have influenced this work.

References

1. ADHD-200 Consortium: The adhd-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience* **6**, 62 (2012)
2. 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)
3. Chizat, L., Peyré, G., Schmitzer, B., Vialard, F.X.: Scaling algorithms for unbalanced optimal transport problems. *Mathematics of computation* **87**(314), 2563–2609 (2018)
4. Craddock, C., Sikka, S., Cheung, B., Khanuja, R., Ghosh, S.S., Yan, C., Li, Q., Lurie, D., Vogelstein, J., Burns, R., et al.: Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac). *Frontiers in Neuroinformatics* **42**(10.3389) (2013)
5. 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)
6. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry* **19**(6), 659–667 (2014)

7. Dujardin, S., Commins, C., Lathuiliere, A., Beerepoot, P., Fernandes, A.R., Kamath, T.V., De Los Santos, M.B., Klickstein, N., Corjuc, D.L., Corjuc, B.T., et al.: Tau molecular diversity contributes to clinical heterogeneity in alzheimer's disease. *Nature medicine* **26**(8), 1256–1263 (2020)
8. Feng, Z., Zhang, S.: Evolved part masking for self-supervised learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10386–10395 (2023)
9. He, C., Chen, Y., Jian, T., Chen, H., Guo, X., Wang, J., Wu, L., Chen, H., Duan, X.: Dynamic functional connectivity analysis reveals decreased variability of the default-mode network in developing autistic brain. *Autism Research* **11**(11), 1479–1493 (2018)
10. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16000–16009 (2022)
11. Hong, S.J., Vos de Wael, R., Bethlehem, R.A., Larivière, S., Paquola, C., Valk, S.L., Milham, M.P., Di Martino, A., Margulies, D.S., Smallwood, J., et al.: Atypical functional connectome hierarchy in autism. *Nature communications* **10**(1), 1022 (2019)
12. Jack, Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., et al.: The alzheimer's disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging* **27**(4), 685–691 (2008)
13. Jung, W., Jeon, E., Kang, E., Suk, H.I.: Eag-rs: a novel explainability-guided roi-selection framework for asd diagnosis via inter-regional relation learning. *IEEE Transactions on Medical Imaging* **43**(4), 1400–1411 (2023)
14. 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)
15. 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)
16. Keller, A.S., Pines, A.R., Shanmugan, S., Sydnor, V.J., Cui, Z., Bertolero, M.A., Barzilay, R., Alexander-Bloch, A.F., Byington, N., Chen, A., et al.: Personalized functional brain network topography is associated with individual differences in youth cognition. *Nature communications* **14**(1), 8411 (2023)
17. Ma, J., Zhang, Y., Zhang, C., Lv, Z., Pei, S.: Brainhgt: A hierarchical graph transformer for interpretable brain network analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 40, pp. 17617–17625 (2026)
18. Pei, S., Ma, J., Lv, Z., Zhang, C., Guan, J.: Community-aware graph transformer for brain disorder identification. In: Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence. pp. 4191–4199 (2025)
19. Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.: Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral Cortex* **28**(9), 3095–3114 (2018)
20. Segal, A., Parkes, L., Aquino, K., Kia, S.M., Wolfers, T., Franke, B., Hoogman, M., Beckmann, C.F., Westlye, L.T., Andreassen, O.A., et al.: Regional, circuit and network heterogeneity of brain abnormalities in psychiatric disorders. *Nature Neuroscience* **26**(9), 1613–1629 (2023)
21. Sporns, O., Betzel, R.F.: Modular brain networks. *Annual review of psychology* **67**(1), 613–640 (2016)

22. Yang, Y., Chen, H., Hu, J., Guo, X., Ma, T.: Advancing brain imaging analysis step-by-step via progressive self-paced learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 58–68. Springer (2024)
23. Yang, Y., Wolfers, T.: Hierarchical characterization of brain dynamics via state space-based vector quantization. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 394–404. Springer (2025)
24. Yang, Y., Ye, C., Ma, T.: A deep connectome learning network using graph convolution for connectome-disease association study. *Neural Networks* **164**, 91–104 (2023)
25. Yang, Y., Ye, C., Su, G., Zhang, Z., Chang, Z., Chen, H., Chan, P., Yu, Y., Ma, T.: Brainmass: Advancing brain network analysis for diagnosis with large-scale self-supervised learning. *IEEE transactions on medical imaging* **43**(11), 4004–4016 (2024)
26. Yang, Y., Ye, C., Sun, J., Liang, L., Lv, H., Gao, L., Fang, J., Ma, T., Wu, T.: Alteration of brain structural connectivity in progression of parkinson’s disease: a connectome-wide network analysis. *NeuroImage: Clinical* **31**, 102715 (2021)
27. 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)
28. Zhang, C., Ren, H., He, X.: P²OT: Progressive partial optimal transport for deep imbalanced clustering. *arXiv preprint arXiv:2401.09266* (2024)
29. Zhao, Y., Nie, D., Chen, G., Wu, X., Zhang, D., Wen, X.: Tardrl: Task-aware reconstruction for dynamic representation learning of fmri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 700–710. Springer (2024)