Equity-Aware Self-Supervised Multi-Domain Connectome Representation Learning for Major Depressive Disorder Diagnosis

Jyotismita Barman*[1]✉, Vipul Kumar Singh*[1], Mohammad Yusuf[1], Tapan K. Gandhi[1,2], Jayadeva[1,2], and Sandeep Kumar[1,2]

[1]Department of Electrical Engineering, Indian Institute of Technology, Delhi, India
[2]Yardi School of Artificial Intelligence, Indian Institute of Technology, Delhi, India
{Jyotismita.Barman,Vipul.Kumar.Singh,ksandeep}@ee.iitd.ac.in

Abstract. Graph-based models have shown promise in neuroimaging by identifying abnormal functional brain connectivity in psychiatric disorders. However, current approaches for functional Magnetic Resonance Imaging analysis have several limitations. Most methods estimate connectivity only from time-domain correlations, ignoring frequency dependent phase synchronization that provides complementary information about neural communication. In addition, existing models typically learn representations using diagnostic labels. For disorders such as Major Depressive Disorder, these labels are heterogeneous and site-dependent, which limits generalization across institutions and populations when labeled data are scarce. Finally, reliability and equity are rarely considered during representation learning, leading to demographic bias and inconsistent subgroup performance. To address these challenges, we propose `RESOLVE` (**R**obust and **E**quitable **S**elf supervised learning for functi**O**na**L** connecti**V**ity mod**E**ling), a self-supervised multi-domain connectome representation learning framework. `RESOLVE` employs dual graph encoders to learn from correlation- and coherence-based connectivity and promotes consistency across the two domains. We further introduce a topology-preserving reconstruction objective that maintains the intrinsic organization of the brain network. To improve reliability, we learn demographically invariant representations by aligning subgroup distributions in the latent space using an Optimal Transport objective. Experiments on the REST-meta-MDD dataset demonstrate the effectiveness of the proposed framework in balancing diagnostic performance and demographic bias compared to existing methods. Source code is available at: RESOLVE.

Keywords: Fairness · Functional Connectivity · Graph Neural Network.

1 Introduction

Neurological and psychiatric disorders pose a growing clinical burden, motivating the development of reliable neuroimaging-based diagnostic tools. Resting-state

* These authors contributed equally as first authors.

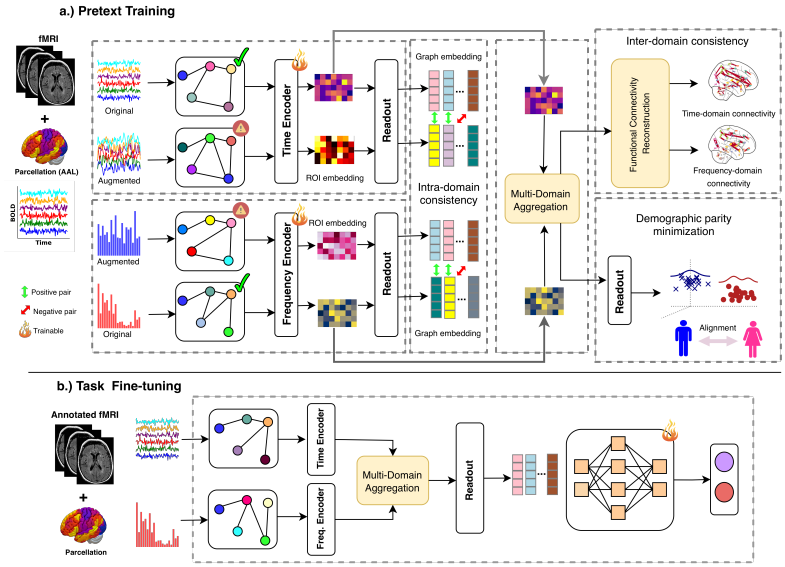

Fig. 1: Overview of the RESOLVE framework, consisting of (a) a self-supervised pretext stage and (b) a downstream diagnostic model. The pretext stage learns representations from correlation- and coherence-based functional connectomes using time and frequency GNN encoders. Intra-domain learning is performed within each connectivity view, followed by a multi-domain aggregator. A functional connectivity reconstruction module preserves network structure, and a latent distribution alignment module reduces demographic discrepancies.

functional magnetic resonance imaging (rs-fMRI) enables non-invasive assessment of brain function through Blood Oxygen Level Dependent (BOLD) signals [11]. Functional connectivity (FC) networks derived from these signals model brain Regions of Interest (ROIs) as nodes and statistical relationships between regions as edges [16]. Graph Neural Networks (GNNs) have therefore become a promising framework for analyzing such connectome data [5,1].

Most existing approaches estimate connectivity only from time-domain correlations between BOLD signals. Correlation captures co-activation patterns, but neural communication also depends on oscillatory synchronization. This information can be reflected in the frequency domain through coherence. Brain networks exhibit frequency-dependent organization, and abnormalities have been reported in Major Depressive Disorder (MDD) [25,20,23]. Relying on a single connectivity measure can therefore lead to incomplete connectome representations, motivating multi-domain connectivity modeling.

In parallel, recent work has highlighted bias in medical imaging models, where predictions vary across demographic attributes [15,26]. Most mitigation strategies operate in supervised classification settings [22]. However, neuroimaging datasets often contain limited annotations due to the cost of reliable clinical la-

beling, and label-dependent models frequently generalize poorly across sites and populations. Self-supervised learning offers an alternative by learning directly from unlabeled connectivity data [24]. Nevertheless, self-supervised models can still encode demographic information present in the data [30].

In supervised settings, fairness can be explicitly enforced at the prediction level. In contrast, self-supervised learning lacks access to task labels, requiring fairness to be imposed directly in the latent representation through bias-invariant embedding learning.

Multi-domain self-supervised learning has been explored for physiological time-series signals [31]. In fMRI, however, the objective is to model interactions among spatially distributed brain regions while preserving network structure. Consequently, methods designed for signal-level representations do not directly account for relational dependencies and network topology. Extending cross-domain consistency to connectome representations, therefore, requires learning embeddings that preserve interactions between regions across connectivity domains. Moreover, prior methods largely overlook bias-invariant representation learning.

This leads to the following research question:

How can complementary connectivity domains be jointly learned to obtain structure-preserving and demographic-invariant brain network representations in a self-supervised setting?

To address the question, we propose `RESOLVE` (Robust and Equitable Self supervised learning for functiOnaL connectiVity modEling), a graph-based self-supervised framework for equity-aware multi-domain functional connectivity learning. As shown in Fig. 1, domain-specific graph encoders learn representations from each connectivity view using intra-domain contrastive learning on augmented connectomes, and the resulting embeddings are fused into a unified representation. To preserve intrinsic network organization, we introduce a topology-preserving reconstruction objective that maintains relational structure across regions and domains. An Optimal Transport–based alignment module further reduces demographic distributional discrepancies in the latent space, producing population-consistent representations. Our contributions are:

❶ **Fairness-aware self-supervised connectome learning.** We propose `RESOLVE`, a graph-based self-supervised framework that learns multi-domain functional connectivity representations without diagnostic labels.

❷ **Structure-preserving multi-domain representation learning.** We introduce a topology-preserving reconstruction objective that retains intrinsic interactions between brain regions across connectivity domains.

❸ **Demographic-invariant representations and improved reliability.** We incorporate an Optimal Transport–based latent alignment to mitigate demographic bias and demonstrate improved cross-site generalization and reduced subgroup disparities.

2 Proposed Method

This section describes the architecture of RESOLVE, a self-supervised framework for functional connectivity modeling. We formalize the learning setting as follows:

Problem Statement. (Fair Self-Supervised Representation Learning for Multi-Domain Functional Connectivity). *Given an unlabeled pre-training fMRI dataset $\mathcal{D}^{pret}$ containing m subjects and a labeled fine-tuning dataset $\mathcal{D}^{tune}$ with p subjects ($p \ll m$), the goal is to learn a mapping function $\mathcal{H}_{\boldsymbol{\Theta}}$ that encodes functional connectivity representations in a self-supervised manner. The learned representations should (i) transfer effectively to downstream tasks when fine-tuned on $\mathcal{D}^{tune}$, and (ii) reduce disparities across demographic subgroups, thereby promoting fair predictions.*

In the following subsections, unless stated otherwise, all data refer to the pre-training phase, and the superscript $^{\text{pret}}$ is omitted for brevity.

2.1 FC Graph Construction

Let $\mathbf{B}_i \in \mathbb{R}^{N\times T}$ denote the rs-fMRI time series of the i-th subject, where N is the number of RoIs defined by the brain parcellation atlas and T is the duration of scan. We adopt the AAL atlas, resulting in $N = 116$ RoIs. For each subject, we construct two connectivity graphs corresponding to the time and frequency domains:

$$\mathcal{G}_i^T = (\mathbf{A}_i^T, \mathbf{X}_i^T), \quad \mathcal{G}_i^F = (\mathbf{A}_i^F, \mathbf{X}_i^F), \tag{1}$$

where $\mathbf{A}_i^T, \mathbf{A}_i^F \in \mathbb{R}^{N\times N}$ denote the adjacency matrices and $\mathbf{X}_i^T, \mathbf{X}_i^F \in \mathbb{R}^{N\times D}$ denote the corresponding node feature matrices for the time and frequency domains, respectively.

Time-domain connectivity. The matrix $\mathbf{A}_i^T$ is computed using Pearson correlation between RoI time series. To reduce spurious weak connections and improve numerical stability, we construct a k-nearest-neighbor graph ($k = 10$) by retaining the strongest connections for each node. The node features $\mathbf{X}_i^T$ are derived from correlation-based regional connectivity profiles. This sparse modeling approach is consistent with prior studies [6,21,17].

Frequency-domain connectivity. We transform the fMRI time series $\mathbf{B}_i$ into the frequency domain using the fast Fourier transform, yielding the spectral representation $\mathbf{F}_i$. We follow standard fMRI preprocessing by using the 0.01–0.1 Hz Low-Frequency Fluctuation (LFF) band, which preserves dominant spontaneous BOLD fluctuations while suppressing scanner drift (< 0.01 Hz) and physiological noise [4,2]. Band-limited coherence (0.01–0.1 Hz) between RoIs is computed to form the adjacency matrix $\mathbf{A}_i^F$, and weak connections are removed using a threshold of 0.3, chosen empirically. Node features $\mathbf{X}_i^F$ are constructed from the power spectral density profiles of the RoIs. This modeling strategy is consistent with prior studies on frequency connectome construction[20,22,8].

During self-supervised training, we generate augmented views by perturbing the signals and recomputing the corresponding connectomes. We apply jittering,

scaling, and time shifting to the time-domain signal $\mathbf{B}_i$ to obtain an augmented signal, from which the augmented graph $\tilde{\mathcal{G}}_i^T$ is constructed. In the frequency domain, we modify the spectrum $\mathbf{F}_i$ by selectively adding or removing frequency components to generate the augmented graph $\tilde{\mathcal{G}}_i^F$.

2.2 Methodology

Time and Frequency Encoders. The time- and frequency-domain graphs are processed by domain-specific GNN encoders G_T and G_F, producing node-level embeddings that capture connectivity patterns across RoIs:

$$\mathbf{Z}_i^T = G_T(\mathcal{G}_i^T), \quad \tilde{\mathbf{Z}}_i^T = G_T(\tilde{\mathcal{G}}_i^T), \quad \mathbf{Z}_i^F = G_F(\mathcal{G}_i^F), \quad \tilde{\mathbf{Z}}_i^F = G_F(\tilde{\mathcal{G}}_i^F). \tag{2}$$

Intra-Domain Contrastive Learning. To obtain graph-level representations, node embeddings are aggregated using a READOUT operator:

$$\mathbf{z}_i^T = \mathrm{READOUT}(\mathbf{Z}_i^T), \quad \mathbf{z}_i^F = \mathrm{READOUT}(\mathbf{Z}_i^F). \tag{3}$$

For each subject, the original and augmented graphs form a positive pair, while graphs from other subjects serve as negatives. This objective encourages invariance to signal perturbations while preserving subject-specific structure. For each domain $w \in \{T, F\}$, the intra-domain contrastive loss for subject i is defined as

$$\mathcal{L}_{\mathrm{intra},i;w} = -\log \frac{\exp\left(\mathrm{sim}(\mathbf{z}_i^w, \tilde{\mathbf{z}}_i^w)/\tau\right)}{\sum_{j \neq i} \exp\left(\mathrm{sim}(\mathbf{z}_i^w, \mathbf{z}_j^w)/\tau\right)}. \tag{4}$$

where $\mathrm{sim}(\cdot,\cdot)$ denotes cosine similarity and τ is a temperature parameter. The total intra-domain loss is:

$$\mathcal{L}_{\mathrm{intra},i} = \mathcal{L}_{\mathrm{intra},i;T} + \mathcal{L}_{\mathrm{intra},i;F}. \tag{5}$$

Multi-Domain Aggregation. We combine the node-level embeddings from the two connectivity domains to obtain a unified representation. In this work, fusion is performed via concatenation, $\mathbf{Z}_i^{\mathrm{fuse}} = \mathrm{concat}\left(\mathbf{Z}_i^T, \mathbf{Z}_i^F\right)$.

Structure-Preserving Reconstruction. Following multi-domain feature aggregation, we enforce structural consistency by ensuring that RoI embeddings preserve the input functional connectivity patterns across both time and frequency domains. To achieve this, we introduce a joint reconstruction loss that aligns the fused embeddings with their respective connectivity structures. Specifically, we minimize the Frobenius norm error between the reconstructed and original functional connectivity matrices:

$$\mathcal{L}_{\mathrm{inter},i} = \left\|\sigma(\mathbf{Z}_i^{\mathrm{fuse}}\mathbf{Z}_i^{\mathrm{fuse}\top}) - \mathbf{A}_i^T\right\|_F^2 + \left\|\sigma(\mathbf{Z}_i^{\mathrm{fuse}}\mathbf{Z}_i^{\mathrm{fuse}\top}) - \mathbf{A}_i^F\right\|_F^2, \tag{6}$$

where $\sigma(\cdot)$ is the sigmoid function and $\top$ denotes matrix transpose.

Demographic-Invariant Representation Learning. To promote fairness, we enforce demographic invariance at the representation level. Specifically, the

learned graph-level embedding $\mathbf{z}_i^{\text{fuse}} = \text{READOUT}(\mathbf{Z}_i^{\text{fuse}})$ should not encode information related to sensitive attributes. To enforce fairness at the representation level, we impose demographic parity by aligning the conditional distributions of learned representations across demographic subgroups [18]. The disparity is defined as:

$$\text{DP} = \left|\mathbb{P}(\mathbf{z}^{\text{fuse}} \mid s_1) - \mathbb{P}(\mathbf{z}^{\text{fuse}} \mid s_2)\right|, \tag{7}$$

where $s \in \{s_1, s_2\}$ denotes the binary subgroup attribute. Directly matching high-dimensional distributions is challenging. We therefore model fairness as a distribution alignment problem and measure the discrepancy between subgroup representations using Optimal Transport (OT). Let P and Q denote the empirical distributions of embeddings for subgroups s_1 and s_2. We minimize the entropy-regularized OT distance, implemented via the Sinkhorn divergence:

$$\mathcal{L}_{\text{DP}} = \mathcal{S}(P, Q). \tag{8}$$

where $\mathcal{S}$ is the Sinkhorn divergence estimator with entropy regularization [9]. Minimizing $\mathcal{L}_{\text{DP}}$ aligns subgroup representation distributions in the latent space, encouraging demographic-invariant embeddings while preserving connectivity structure. The total pre-training objective will then be:

$$\mathcal{L}_{\texttt{RESOLVE}} = \lambda_1 \, \mathbb{E}_i[\mathcal{L}_{\text{intra},i}] + \lambda_2 \, \mathbb{E}_i[\mathcal{L}_{\text{inter},i}] + \lambda_3 \, \mathcal{L}_{\text{DP}}, \tag{9}$$

where λ_1, λ_2, and λ_3 control the contribution of each component.
Upon completion of pre-training, we store the model parameters and represent the trained model as $\mathcal{H}(\cdot, \boldsymbol{\Theta})$, where $\boldsymbol{\Theta}$ denotes the set of all trainable parameters. During fine-tuning, fused multi-domain embeddings from $\mathcal{D}^{\text{tune}}$ are input to a multilayer perceptron for classification with cross-entropy loss.

3 Experiments

Dataset and Experimental Protocol. We evaluate the proposed framework on the REST-meta-MDD dataset [28], a multi-site rs-fMRI cohort for MDD. To construct the brain graphs, fMRI data were preprocessed using DPARSF [29]. The pipeline included slice-timing correction, head motion realignment, spatial normalization, band pass filtering (0.01-0.1 Hz), and global signal regression. Regional time series were then extracted for each ROI using the AAL atlas.
To assess generalization in a self-supervised setting, we follow the cross-site evaluation protocol proposed in [24]. For pre-training, we use the largest available site (Site 20), containing 533 subjects (282 MDD, 251 healthy controls). The site also exhibits demographic imbalance, comprising 186 males and 347 females. To evaluate cross-site generalization, the pretrained model is tested on two independent sites. Site 21 includes 156 subjects (86 MDD, 70 healthy controls; 69 males, 87 females), and Site 14 includes 93 subjects (32 MDD, 61 healthy controls; 34 males, 59 females). This cross-site setting assesses generalization under distribution shift and is substantially more challenging than random subject-level splits within a single site. After pre-training, the model is fine-tuned on each target

Table 1: Performance on REST-meta-MDD. $S \rightarrow T$ denotes source and target sites. FATE is computed relative to the **GCN** baseline. Best and second-best results are in red and blue, respectively.

	Site 20 → Site 14			Site 20 → Site 21		
Methods	**AUC** (↑)	**Acc** (↑)	**FATE** (↑)	**AUC** (↑)	**Acc** (↑)	**FATE** (↑)
GCN [14]	$51.35_{\pm 12.36}$	$64.61_{\pm 2.11}$	0.000	$53.80_{\pm 7.92}$	$53.16_{\pm 2.97}$	0.000
BrainGNN [16]	$60.21_{\pm 10.85}$	$64.56_{\pm 6.88}$	−0.77	$52.28_{\pm 18.00}$	$54.24_{\pm 8.53}$	−0.53
BrainNetCNN [13]	$58.14_{\pm 9.75}$	$63.67_{\pm 1.92}$	−0.70	$58.28_{\pm 10.08}$	$57.64_{\pm 6.68}$	−0.58
BrainNetTF [12]	$65.01_{\pm 8.95}$	$67.01_{\pm 6.56}$	−0.88	$60.07_{\pm 11.32}$	$57.94_{\pm 10.82}$	−1.57
CoRE-BOLD [22]	$60.24_{\pm 6.66}$	$65.61_{\pm 3.67}$	0.17	$58.92_{\pm 4.70}$	$57.35_{\pm 4.70}$	−0.14
GATE [19]	$61.28_{\pm 2.75}$	$66.14_{\pm 2.74}$	−0.51	$52.08_{\pm 5.81}$	$52.76_{\pm 6.20}$	−0.45
GCDA [24]	$64.80_{\pm 7.53}$	$62.51_{\pm 6.80}$	−0.26	$59.49_{\pm 10.03}$	$59.11_{\pm 10.08}$	−0.11
R-Supervised	$61.18_{\pm 7.98}$	$67.63_{\pm 10.27}$	−0.41	$57.16_{\pm 7.77}$	$58.78_{\pm 9.38}$	−0.01
No Pre-train	$56.31_{\pm 7.76}$	$62.28_{\pm 6.46}$	0.08	$55.09_{\pm 4.05}$	$53.47_{\pm 2.67}$	−0.21
`RESOLVE`	$67.38_{\pm 9.66}$	$70.93_{\pm 6.63}$	0.33	$61.15_{\pm 10.65}$	$62.06_{\pm 9.58}$	0.50

site using 5-fold cross-validation, with 72% of the data for training, 8% for validation, and 20% for testing.

Utility is evaluated using AUC and accuracy, while fairness is assessed by the demographic parity difference. Predictive utility and fairness often exhibit a trade-off in clinical models. To jointly evaluate both aspects, we report the Fairness-Accuracy Trade-off Efficiency (FATE) metric, commonly used in fair medical imaging studies [27,3]. A higher FATE score indicates better balance between utility and fairness. For a model m and baseline b, FATE $= \frac{ACC_m - ACC_b}{ACC_b} - \lambda \frac{\Delta_{DP,m} - \Delta_{DP,b}}{\Delta_{DP,b}}$, where Δ_{DP} denotes demographic parity difference and $\lambda = 1$.

We compare `RESOLVE` with nine baselines: GCN [14], BrainNetCNN [13], BrainGNN [16], BrainNetTF [12], GATE [19], GCDA [24], and CoRE-BOLD [22], along with two ablations: **R-Supervised** (Task-specific `RESOLVE` trained with labeled data during both pre-training and fine-tuning.) and **No Pre-train**.

The graph encoder uses a three-layer GCN with ReLU activations. Training is performed using Adam (learning rate 10^{-2}) for 50 epochs. Global mean pooling is used as the READOUT function. Additional implementation details and hyperparameter configurations are detailed in the code repository.

3.1 Results

Utility and Fairness Comparison. Table 1 summarizes the classification performance for MDD vs. HC across cross-site evaluations. `RESOLVE` consistently achieves the best performance in both utility and fairness metrics. Compared with existing baselines, the proposed method improves AUC and accuracy while substantially reducing the demographic parity gap, indicating more reliable predictions across subgroups. The *No Pre-train* model shows degraded performance, demonstrating the importance of self-supervised pretraining, while the

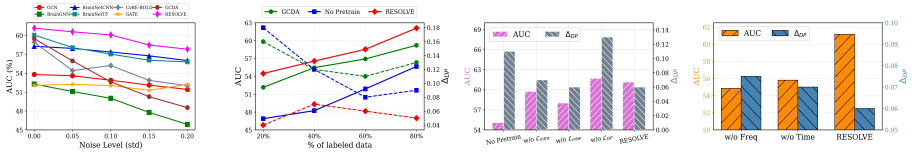

Fig. 2: Additional experiments: (a) robustness analysis under Gaussian noise, (b) performance across varying labeled target data (solid lines: AUC, dotted lines: fairness), (c) ablation study of key modules in RESOLVE, and (d) ablation study of using dual-domain connectivity information.

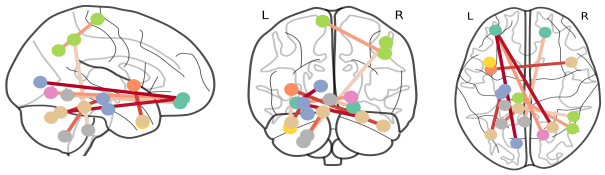

Fig. 3: Top 10 significant functional connectivity alterations between MDD and HCs ($p < 0.05$). Darker edges indicate smaller p-values (greater significance).

R-Supervised variant exhibits reduced cross-site generalization, indicating that label-free pre-training produces more generalizable representations before fine-tuning. For supervised baselines, training is performed only on the fine-tuning dataset, leading to lower performance than typically reported in the literature.

Robustness Analysis. Fig. 2(a) illustrates the performance of competing methods under increasing levels of zero-mean Gaussian noise on Site 21 for MDD classification. RESOLVE consistently demonstrates superior robustness.

Impact of Labeled Target Data. Fig. 2(b) presents performance of GCDA, No Pre-train, and RESOLVE under varying proportions of labeled fine-tuning BOLD signals from Site 21 in MDD classification. The results highlight improved generalization and fairness of RESOLVE with fewer labeled target samples.

Ablation Study. To evaluate the contribution of each component in RESOLVE, we conduct an ablation study for MDD classification on Site 21. We sequentially remove each objective from (9) and present the results in Fig. 2(c). The findings demonstrate the positive impact of each module on the overall performance of RESOLVE in terms of both utility and fairness. To further assess the impact of multi-domain information, we evaluate the model using each domain individually. As shown in Fig. 2(d), the dual-domain model outperforms the single-domain variants.

Discriminative FC for MDD. To identify the significant FC differences, edge-wise two-sided Welch's t-tests were performed across subjects. We use embeddings derived from RESOLVE to compute Z-transformed Pearson correlation matrices, enabling detection of biologically meaningful effects. From Fig. 3, significant pairwise functional connectivity differences were observed between regions involved in memory and emotion processing (e.g., hippocampus–insula), higher-

order cognition (inferior parietal gyrus, angular gyrus, and medial orbitofrontal cortex), and visual processing (lingual and calcarine cortices). Several cerebellar subregions (Crus I and lobules VI, VIII, and X, including vermis IV–V) were also involved in altered FC, suggesting disrupted fronto-limbic and cerebello-cortical interactions in MDD compared with HCs, consistent with prior studies [7,10].

4 Conclusion

In this work, we presented `RESOLVE`, a self-supervised graph representation learning framework that leverages functional connectivity in both the time and frequency domains. The model further promotes demographic invariance by aligning subgroup representations using a Optimal Transport–based objective. Experiments on multi-site rs-fMRI data demonstrate improved generalization and reduced subgroup disparity compared with competing methods.

Acknowledgments. This work was supported by the Science and Engineering Research Board (SERB) under Grant No. RP04820G. Jyotismita Barman is supported by Tata Consultancy Services Research Fellowship. Vipul Kumar Singh is supported by Prime Minister's Research Fellowship. The authors would like to sincerely thank the DIRECT Consortium for giving the access to the REST-meta-MDD data.

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

References

1. Bessadok, A., Mahjoub, M.A., Rekik, I.: Graph neural networks in network neuroscience. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(5), 5833–5848 (2022)
2. Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magnetic resonance in medicine **34**(4), 537–541 (1995)
3. Chiu, C.H., Chen, Y.J., Wu, Y., Shi, Y., Ho, T.Y.: Achieve fairness without demographics for dermatological disease diagnosis. Medical Image Analysis **95**, 103188 (2024)
4. Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E.: Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. American journal of neuroradiology **22**(7), 1326–1333 (2001)
5. Cui, H., Dai, W., Zhu, Y., Kan, X., Gu, A.A.C., Lukemire, J., Zhan, L., He, L., Guo, Y., Yang, C.: Braingb: a benchmark for brain network analysis with graph neural networks. IEEE transactions on medical imaging **42**(2), 493–506 (2022)
6. Dai, P., Lu, D., Shi, Y., Zhou, Y., Xiong, T., Zhou, X., Chen, Z., Zou, B., Tang, H., Huang, Z., et al.: Classification of recurrent major depressive disorder using a new time series feature extraction method through multisite rs-fmri data. Journal of Affective Disorders **339**, 511–519 (2023)

7. Depping, M.S., Schmitgen, M.M., Kubera, K.M., Wolf, R.C.: Cerebellar contributions to major depression. Frontiers in psychiatry **9**, 418513 (2018)
8. Duan, P., Dvornek, N.C., Wang, J., Eilbott, J., Du, Y., Sukhodolsky, D.G., Duncan, J.S.: Spectral brain graph neural network for prediction of anxiety in children with autism spectrum disorder. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2024)
9. Feydy, J., Séjourné, T., Vialard, F.X., Amari, S.i., Trouvé, A., Peyré, G.: Interpolating between optimal transport and mmd using sinkhorn divergences. In: The 22nd international conference on artificial intelligence and statistics. pp. 2681–2690. PMLR (2019)
10. Gao, J., Li, Y., Wei, Q., Li, X., Wang, K., Tian, Y., Wang, J.: Habenula and left angular gyrus circuit contributes to response of electroconvulsive therapy in major depressive disorder. Brain Imaging and Behavior **15**(5), 2246–2253 (2021)
11. Hull, J.V., Dokovna, L.B., Jacokes, Z.J., Torgerson, C.M., Irimia, A., Van Horn, J.D.: Resting-state functional connectivity in autism spectrum disorders: a review. Frontiers in psychiatry **7**, 205 (2017)
12. 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)
13. 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)
14. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
15. Larrazabal, A.J., Nieto, N., Peterson, V., Milone, D.H., Ferrante, E.: Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proceedings of the National Academy of Sciences **117**(23), 12592–12594 (2020)
16. 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)
17. Noman, F., Ting, C.M., Kang, H., Phan, R.C.W., Ombao, H.: Graph autoencoders for embedding learning in brain networks and major depressive disorder identification. IEEE Journal of Biomedical and Health Informatics **28**(3), 1644–1655 (2024)
18. Oneto, L., Donini, M., Luise, G., Ciliberto, C., Maurer, A., Pontil, M.: Exploiting mmd and sinkhorn divergences for fair and transferable representation learning. Advances in Neural Information Processing Systems **33**, 15360–15370 (2020)
19. Peng, L., Wang, N., Xu, J., Zhu, X., Li, X.: Gate: Graph cca for temporal self-supervised learning for label-efficient fmri analysis. IEEE Transactions on Medical Imaging **42**(2), 391–402 (2022)
20. Peng, L., Su, J., Hu, D., Yu, Y., Wei, H., Li, M.: Measuring functional connectivity in frequency-domain helps to better characterize brain function. Human Brain Mapping **45**(10), e26726 (2024)
21. Qin, K., Lei, D., Pinaya, W.H., Pan, N., Li, W., Zhu, Z., Sweeney, J.A., Mechelli, A., Gong, Q.: Using graph convolutional network to characterize individuals with major depressive disorder across multiple imaging sites. EBioMedicine **78** (2022)
22. Singh, V.K., Barman, J., Kumar, S., Jayadeva, J.: Core-bold: Cross-domain robust and equitable ensemble for bold signal analysis. In: Machine Learning for Health (ML4H). pp. 961–975. PMLR (2025)

23. Wang, L., Ma, Q., Sun, X., Xu, Z., Zhang, J., Liao, X., Wang, X., Wei, D., Chen, Y., Liu, B., et al.: Frequency-resolved connectome alterations in major depressive disorder: A multisite resting fmri study. Journal of Affective Disorders **328**, 47–57 (2023)
24. Wang, X., Fang, Y., Wang, Q., Yap, P.T., Zhu, H., Liu, M.: Self-supervised graph contrastive learning with diffusion augmentation for functional mri analysis and brain disorder detection. Medical image analysis **101**, 103403 (2025)
25. Wu, C.W., Gu, H., Lu, H., Stein, E.A., Chen, J.H., Yang, Y.: Frequency specificity of functional connectivity in brain networks. Neuroimage **42**(3), 1047–1055 (2008)
26. Xu, Z., Li, J., Yao, Q., Li, H., Zhao, M., Zhou, S.K.: Addressing fairness issues in deep learning-based medical image analysis: a systematic review. npj Digital Medicine **7**(1), 286 (2024)
27. Xu, Z., Zhao, S., Quan, Q., Yao, Q., Zhou, S.K.: Fairadabn: Mitigating unfairness with adaptive batch normalization and its application to dermatological disease classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 307–317. Springer (2023)
28. Yan, C.G., Chen, X., Li, L., Castellanos, F.X., Bai, T.J., Bo, Q.J., Cao, J., Chen, G.M., Chen, N.X., Chen, W., et al.: Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. Proceedings of the National Academy of Sciences **116**(18), 9078–9083 (2019)
29. Yan, C., Zang, Y.: Dparsf: a matlab toolbox for" pipeline" data analysis of resting-state fmri. Frontiers in systems neuroscience **4**, 1377 (2010)
30. Yfantidou, S., Spathis, D., Constantinides, M., Vakali, A., Quercia, D., Kawsar, F.: Using self-supervised learning can improve model fairness. In: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. pp. 3942–3953 (2024)
31. Zhang, X., Zhao, Z., Tsiligkaridis, T., Zitnik, M.: Self-supervised contrastive pre-training for time series via time-frequency consistency. Advances in neural information processing systems **35**, 3988–4003 (2022)