BrainPrompt: Domain Adaptation with Prompt Learning for Multi-site Brain Network Analysis

Liuzeng Zhang[1,2,†], Lanting Li[2,3,†], Peng Cao[1,2,4(✉)], Jinzhu Yang[1,2,4], and Osmar R. Zaiane[5]

[1] Computer Science and Engineering, Northeastern University, Shenyang, China
[2] Key Laboratory of Intelligent Computing in Medical Image of Ministry of Education, Northeastern University, Shenyang, China
[3] College of Artificial Intelligence, Shenyang Aerospace University, Shenyang, China
[4] National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Shenyang, China
caopeng@cse.neu.edu.cn
[5] Amii, University of Alberta, Edmonton, Alberta, Canada

Abstract. It is challenging to discriminate autism spectrum disorder (ASD) from a highly heterogeneous database, because there is a great deal of uncontrollable variability in the data from different sites. Recently, prompt learning has received considerable attention in domain adaptation as a promising solution. However, its application to graph data like multi-site brain networks has not been fully studied. It faces two major challenges: (1) complex graph structure; and (2) inter-individual variability. To overcome the issues, we propose a novel prompt-tuning paradigm for multi-site brain network analysis (BrainPrompt) using functional magnetic resonance imaging (fMRI). Specifically, we introduce two tunable soft prompts: (1) a mask prompt to prune noisy edges while preserving important connections, and distill it to reduce domain-specific biases; (2) sample prompts to capture inter-individual variations. Our model outperforms other models on the ABIDE dataset, especially at sites with limited samples (*e.g.*, the Stanford site, which has only 39 samples). BrainPrompt achieves a 35.88% improvement in accuracy compared to the state-of-the-art method, highlighting its superiority in small sites. Furthermore, our results demonstrate the interpretability and generalization of the proposed method. Our code is available at https://github.com/zliuzeng/BrainPrompt.

Keywords: Autism Spectrum Disorder · Domain Adaptation · Multi-Site Brain Networks · Prompt Learning.

1 Introduction

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder with significant individual variability influenced by demographic, cognitive,

[1] † Liuzeng Zhang and Lanting Li contributed equally to this work.

and emotional factors [15]. Moreover, the current multi-site fMRI data used for diagnosis brings in heterogeneous distribution among multi-site data due to differing acquisition protocols [9]. These two types of heterogeneity jointly degrade the performance of ASD diagnostic models [6,14]. One common approach to address site-specific heterogeneity is traditional transfer learning. For instance, maLRR [23] utilizes low-rank regression to mitigate domain-specific biases, demonstrating notable performance. These methods typically involve fine-tuning all or part of the model parameters using target domain training data. Despite significant progress, they remain insufficiently effective and efficient. As model complexity increases, fine-tuning the pre-trained model becomes computationally intensive and time-consuming [5]. Particularly for the domains with limited samples, traditional transfer learning is prone to overfitting, leading to performance degradation. Furthermore, these methods fail to account for the inter-individual variability in ASD diagnosis.

Prompt learning [8] is an emerging paradigm in natural language processing (NLP) that adapts pre-trained models to downstream tasks by updating only a small subset of parameters. This approach is particularly advantageous when only limited samples are available, as it effectively guides the model to capture domain-specific knowledge by designing appropriate prompts, which require very few labeled examples [11]. Specifically, prompts are flexible, adjustable instructions or signals that direct the model's focus toward the relevant task or domain, without requiring extensive retraining of the model itself [7]. Fig. 1(c) compares the performance results of traditional transfer learning and our method. Compared to TP-MIDA [12] and LRCDR [17] (Traditional transfer learning), BrainPrompt (Our method) achieves superior performance with lower costs. Inspired

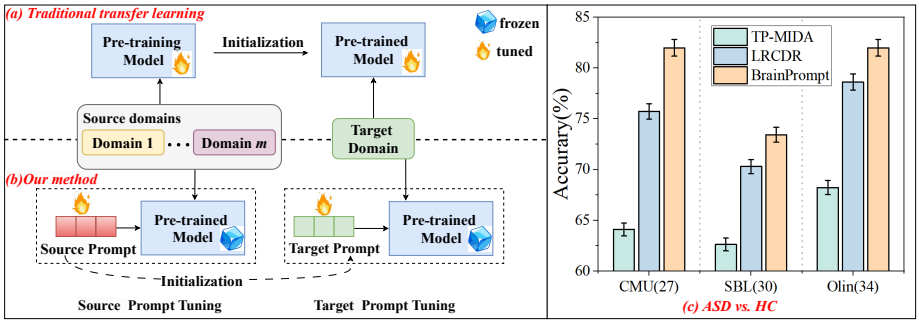


Fig. 1. Comparison of the two methods. Fig. 1(a)-(b): Illustration of traditional transfer learning and our method. Fig. 1(c): Performance of BrainPrompt compared with the state-of-the-art multi-site domain adaptive methods, i.e. TP-MIDA and LRCDR on the sites with the smallest sample sizes of 27,30 and 34, respectively.

by the success of prompt learning in NLP tasks [16], we attempt to apply prompt

learning to multi-site brain network analysis. However, its direct application faces two major challenges:

(1) In contrast to language, brain networks are sparse and contain noisy edges [22]. Therefore, it may cause inaccurate results if the prompt design neglects the inherent structure of brain networks.

(2) Patients with ASD exhibit significant individual variability in the connectivity pattern distribution of functional regions [15]. This variability means that even a well-trained prompt may not generalize across a large population in ASD diagnosis. Designing adaptive sample prompts to capture inter-individual variations in brain networks is critical.

To overcome these limitations, we propose a novel **P**rompt-tuning paradigm for multi-site **B**rain network analysis, named BrainPrompt. First, we pre-train the baseline model in a supervised manner on multiple source domains. Next, we design a tunable source **mask prompt** that adaptively prunes noisy edges while preserving discriminative connections on each source domain. Moreover, we train source domain-specific embeddings for each source domain using Low-Rank Adaptation [10]. Finally, we initialize the target mask prompt by distilling knowledge from multiple source mask prompts, effectively reducing site heterogeneity and enhancing cross-domain knowledge transfer. Meanwhile, we employ a cross-domain attention mechanism to transform source domain-specific embeddings into tunable target **sample prompts** that dynamically adapt to the domain-adaptive distribution biases and capture inter-individual variability. Our contributions are summarized as follows:

- We propose a novel prompt-tuning paradigm for multi-site brain network analysis. To the best of our knowledge, this is the first attempt to leverage prompts for multi-site brain network analysis.
- We introduce the mask prompt and sample prompts to explore the topological structure of brain networks and identify inter-individual heterogeneity.
- We evaluate BrainPrompt on the public Autism Brain Imaging Data Exchange (ABIDE) dataset. The experimental results demonstrate the superior performance of BrainPrompt, especially at sites with limited samples (*e.g.,* Stanford site, which has only 39 samples). BrainPrompt achieves 35.88% and 41.19% accuracy improvements compared to TP-MIDA and LRCDR methods at the Stanford site.

2 Method

The framework of our proposed BrainPrompt is shown in Fig. 2. BrainPrompt operates in three stages: **Stage 1**: We pre-train the baseline model on multiple source domains. **Stage 2**: We adopt a source mask prompt that adaptively prunes noisy edges while preserving discriminative connections for each source domain. Moreover, we train source domain-specific embeddings to capture domain-specific information on each source domain. **Stage 3**: To reduce site heterogeneity and ensure efficient cross-domain knowledge transfer, we initialize the target mask prompt by distilling knowledge from multiple source

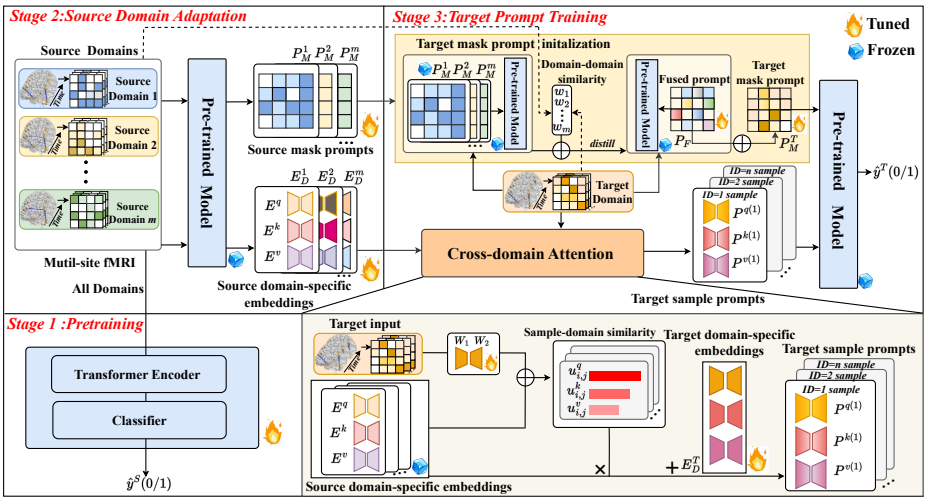


Fig. 2. Illustration of (i) Pretraining, (ii) Source domain adaptation, including constructing source mask prompts and domain-specific embeddings, and (iii) Target prompt training, including training target mask prompt and sample prompts.

mask prompts. Meanwhile, we integrate multiple source domain-specific embeddings through the cross-domain attention mechanism to generate target sample prompts that dynamically adapt to the domain-adaptive distribution biases and capture inter-individual variability.

2.1 Pretraining

We train the baseline model in a supervised manner by mixing data from all source domains. The model comprises a transformer encoder and a multi-layer perceptron (MLP) classifier. Given multiple source training data (X^S) and labels (Y^S), the model is optimized as $\arg\min_\Theta L_{ce}(h_\Theta(X^S), Y^S)$, where L_{ce} denotes the cross-entropy loss and $h_\Theta(\cdot)$ represents the model's prediction, with Θ consisting of two components: the parameters of the encoder and an MLP classifier.

2.2 Source domain adaptation

Source mask prompt: Considering the complexity of brain network structures, we design a mask prompt $\boldsymbol{P}_M \in \mathbb{R}^{M\times M}$ for each domain to identify disease-related edges. We apply $\boldsymbol{L}_1$ regularization to ensure the sparsity of $\boldsymbol{P}_M$. The mask prompt is incorporated into the input as $\hat{x}_i = x_i \odot \sigma(\boldsymbol{P}_M)$, where $\odot$ denotes element-wise multiplication, $\sigma(\cdot)$ represents the sigmoid function, and $x_i = \{x_i^t | x_i^t \in \mathbb{R}^{M\times M}, t \in [1, \mathrm{T}]\}$, with x_i^t denoting the adjacency matrix at time step t for the i-th sample. Note that $\boldsymbol{P}_M$ is consistent across all time steps. We learn the source mask prompt $\boldsymbol{P}_M^S$ for each source domain by minimizing the cross-entropy loss and $\boldsymbol{L}_1$ regularization while freezing the pre-trained model.

Source domain-specific embeddings: To adaptively capture domain-specific information, we train domain-specific embeddings $\boldsymbol{E}_D = \{\boldsymbol{E}^q, \boldsymbol{E}^k, \boldsymbol{E}^v\} \in \mathbb{R}^{D\times d}$ using Low-Rank Adaptation on each domain. These embeddings are incorporated into the linear projections of Q, K, and V in the pre-trained model's multi-head self-attention. The update is defined as $\Theta'_U = \Theta_U + \Delta\Theta_U$, with $U \in \{Q, K, V\}$, $\Delta\Theta_U$ representing the learnable parameters of $\boldsymbol{E}_D$, and $\Theta_U \in \mathbb{R}^{D\times d}$ denoting the original linear projection parameters. The source domain-specific embeddings $\boldsymbol{E}_D^S$ are randomly initialized as low-rank matrices and optimized via the cross-entropy loss while freezing the pre-trained model on each source domain.

2.3 Target prompt Training

Target mask prompt Initialization: To efficiently reduce site heterogeneity and improve cross-domain knowledge transfer, we distill a fused prompt from multiple source domains and use it to initialize the target mask prompt. Specifically, each sample is processed using the pre-trained model with a mask prompt to obtain its hidden representation. Moreover, we compute the cosine distance [19] between the centroids of these representations for each source domain and the target domain, which defines **domain-domain similarity** scores as $w = \{w_1, w_2, \ldots, w_m\}$, where m is the number of source domains. Then, we aggregate the predictions from all source domains by weighting each source prediction with its similarity. This aggregated prediction is defined as:

$$\hat{H}(y_i|\boldsymbol{P}_M^S) = \frac{1}{m}\sum_{i=1}^{m} w_i \cdot H_s(y_i|\boldsymbol{P}_M^i) \tag{1}$$

where $H_s(y_i|\boldsymbol{P}_M^i)$ is the prediction of the i-th sourch domain. We minimize the Kullback-Leibler (KL) divergence [21] between the aggregated source predictions and the fused prompt's predictions. This encourages the fused prompt to align with the aggregated source distribution, facilitating the extraction of consensus knowledge and suppressing domain-specific noise. Finally, we initialize a mask prompt $\boldsymbol{P}_M^T$ using the fused prompt $\boldsymbol{P}_F$ for the target domain as Eq. (2):

$$\boldsymbol{P}_M^T = \boldsymbol{P}_M^T + \alpha\boldsymbol{P}_F \tag{2}$$

where α denotes a weight constrained to lie within the range $[0, 1]$. We fine-tune $\boldsymbol{P}_M^T$ along with other prompts on the target domain.

Target sample prompts: To capture individual variability, we propose sample prompts for each subject on the target domain. Since each domain has a unique contribution to the target samples, we compute **sample-domain similarity** scores between target domain samples and all source domain-specific embeddings and then interpolate those embeddings to generate target sample prompts.

Cross-domain Attention. BrainPrompt controls the influence of the set of source domain-specific embeddings on the final sample prompt by calculating **sample-domain similarity** scores. For the i-th target sample,

$$u_{i,j} = \frac{e^{\boldsymbol{E}_D^j \cdot f(x_i)}}{e^{\boldsymbol{E}_D^T \cdot f(x_i)} + \sum_{k=1}^{m} e^{\boldsymbol{E}_D^k \cdot f(x_i)}} \tag{3}$$

where $u_{i,j} = \{u_{i,j}^q, u_{i,j}^k, u_{i,j}^v\} \in \mathbb{R}^{1\times m}$ denotes the scores of the i-th subject over the j-th source domain and the function $f_w(\cdot) = W_1(\sigma(W_2(\cdot)))$ represents the non-linear mapping, where $W_1 \in \mathbb{R}^{d\times D}$ and $W_2 \in \mathbb{R}^{D\times d}$ are the learnable parameters. $\boldsymbol{E}_D^T$ and $\boldsymbol{E}_D^j$ are the domain-specific embeddings for the target domain and the j-th source domain, respectively.

Target prompt Interpolation. Given the set of source domain-specific embeddings $\boldsymbol{E}_D^S$ and the target sample x_i, the final sample prompts are defined as follows:

$$\boldsymbol{P}_I^{(i)} = F_{\Theta}(x_i, \boldsymbol{E}_D^S) \tag{4}$$

where $\boldsymbol{P}_I^{(i)} = \{\boldsymbol{P}^{q(i)}, \boldsymbol{P}^{k(i)}, \boldsymbol{P}^{v(i)}\} \in \mathbb{R}^{D\times d}$ represents the sample prompts for the i-th subject, $F_{\Theta}(\cdot)$ denotes the target prompt interpolation function with Θ consists of two components: the parameters of $f_w(\cdot)$ and target domain-specific embeddings $\boldsymbol{E}_D^T$. The specific process is described by $\boldsymbol{P}_I^{(i)} = \boldsymbol{E}_D^T + \sum_{j=1}^{m} u_{i,j}\boldsymbol{E}_D^j$, where $u_{i,j}$ are the scores generated by Eq. (3). With the pre-trained model frozen, we fine-tune $\boldsymbol{P}_M^T$, $\boldsymbol{E}_D^T$, and $f_w(\cdot)$ to generate the final prediction.

3 Experiments

3.1 Dataset and experimental setup

The ABIDE dataset [4] contains resting-state fMRI and clinical data from 1,112 subjects across 17 sites. For reproducibility, we used the version from the Preprocessed Connectome Project. After manual artifact removal by three clinicians, we obtained a subset of 1,035 high-quality MRI images with phenotypic data. The rs-fMRI data was processed using the Configurable Pipeline for the Analysis of Connectomes (C-PAC) [2], including slice-timing correction, image realignment, and nuisance regression. Time series of average voxel signals were extracted from each brain region defined by the Anatomical Automatic Labeling (AAL) atlas with 116 ROIs.

We perform 5-fold cross-validation on the ABIDE dataset for ASD diagnosis, treating one site as the target domain and the rest as sources. For each subject's time series, we use a sliding window and masking mechanism for segmentation, followed by Pearson Correlation Coefficients to compute the correlation. Subjects from the OHSU site are excluded due to their short time series.

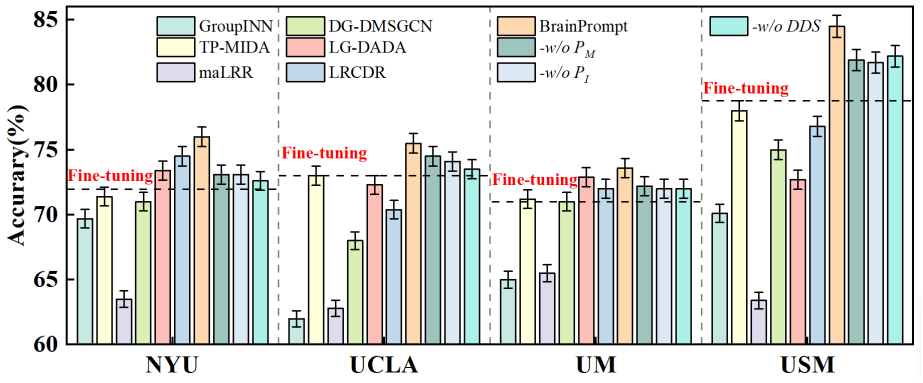


Fig. 3. Comparison of BrainPrompt with the state-of-the-art methods on the four largest sites. We regard fine-tuning as the baseline, which indicates that the baseline model, consisting of a transformer-encoder and MLP, is pre-trained on the source domain and then fine-tuned on the target domain.

3.2 Results and Discussion

Comparison with state-of-the-art Methods: We evaluate BrainPrompt on the four largest sites (NYU, UM, UCLA, USM). Except GroupINN [24] which is non-domain adaptation, we focus on the comparison with five state-of-the-art domain adaptation methods: TP-MIDA [12], maLRR [23], DG-DMSGCN [3], LG-DADA [1], and LRCDR [17]. From Fig. 3, we observe that BrainPrompt outperforms all the comparable methods. Specifically, BrainPrompt significantly outperforms the previous best methods, *i.e.*, TP-MIDA and LRCDR, achieving improvements of 6.44%, 3.42%, 3.37%, and 8.33%, and 2.01%, 7.10%, 2.22%, and 10.03% in terms of accuracy on NYU, UCLA, UM, and USM sites, respectively. This demonstrates the effectiveness of our method in ASD diagnosis. Additionally, our method updates only the prompt (1.95M parameters),which is 80.52% fewer than native fine-tuning method that updates all parameters (10.01M) in the target domain. Compared to other methods (*e.g.*, maLRR: 8.09M; LRCDR: 12.51M), our model achieves competitive efficiency.

Ablation study: We also individually removed the proposed modules (-*w/o* $\boldsymbol{P}_M$: without the mask prompt, -*w/o* DDS: without the domain-domain similarity, i.e. assuming equal similarity across domains, -*w/o* $\boldsymbol{P}_I$: without the sample prompts) from BrainPrompt to verify the effectiveness of each component. From Fig. 3, we conclude that the proposed components are necessary and complementary, as they collaborate effectively and contribute positively to ASD diagnosis.

Evaluation on sites with limited samples: In addition to the evaluation of the four major sites, we also assessed our model on multiple sites with limited samples. In each fold of 5-fold cross-validation, each site is individually split into

Table 1. Performance of BrainPrompt compared with four state-of-the-art methods on the sites with limited samples. Avg: the average result across all sites. Num: the number of samples at different sites. The best results are boldfaced and the second-best results are underlined.

Site	Num	ACC(AUC)				
		maLRR	LG-DADA	TP-MIDA	LRCDR	BrainPrompt
CMU	27	58.45(60.67)	65.54(68.32)	64.11(65.11)	75.74(76.58)	**82.00(80.66)**
SBL	30	65.23(62.45)	61.01(63.12)	62.62(60.12)	70.36(68.67)	**73.43(73.33)**
Olin	34	57.23(60.50)	66.25(67.11)	68.24(69.23)	78.65(74.28)	**82.00(76.83)**
SDSU	36	63.78(62.12)	65.65(64.80)	68.13(70.51)	74.58(70.39)	**81.10(86.10)**
Caltech	37	61.01(60.90)	62.42(64.00)	61.11(66.78)	70.86(65.28)	**72.27(70.66)**
Stanford	39	60.25(62.28)	64.22(65.11)	68.11(66.22)	65.55(64.47)	**92.55(92.50)**
Trinity	47	62.78(64.21)	68.33(71.23)	63.32(60.11)	72.69(69.65)	**76.04(72.50)**
KKI	48	65.21(64.10)	67.74(68.45)	65.95(66.02)	70.54(72.48)	**74.70(75.04)**
MaxMum	52	62.02(60.18)	63.08(66.91)	60.25(61.35)	68.14(62.71)	**70.93(74.40)**
Yale	56	54.22(58.55)	60.64(61.02)	61.45(61.12)	75.25(75.96)	**80.33(81.21)**
Pitt	56	64.12(62.12)	66.15(62.65)	65.87(66.11)	70.27(68.25)	**73.47(73.77)**
Leuven	63	61.21(62.25)	65.72(69.10)	63.83(64.30)	72.11(70.25)	**79.15(78.21)**
Avg	-	61.29(61.69)	64.73(65.65)	64.41(64.62)	72.31(70.49)	**78.16(77.93)**

training and test subsets. The results in Table 1 clearly indicate that our method exhibits superior performance across all sites compared to other methods. Specifically, compared with LRCDR, our model achieves an average increase of 8.09% (ACC) and 10.55% (AUC) on Avg. Furthermore, the model achieves favorable results on different small sites not limited to the four major sites, demonstrating its capability of generalization.

Prompt Interpretation: We investigate the interpretability of ASD through the mask prompt in our model. Specifically, we extracted the weights of the functional connectivity (FCs) obtained from the mask prompts of all the sites and identified the common top 30 FCs across all the sites (shown in Fig. 4.). From Fig. 4(b), we observe that the critical FCs are primarily distributed across the frontal, temporal, and cerebellar regions, consistent with the findings of previous studies [20,13]. In addition, 10 of the top 30 discriminative FCs are associated with the temporal regions, which emerge as a key area for ASD diagnosis. These findings are consistent with the current interpretation of the pathological pathways of ASD [25,18].

4 Conclusion

We propose a novel prompt-tuning framework for multi-site brain network analysis and improve ASD diagnostic performance. Our framework effectively captures the topological structure of brain networks and mitigates inter-individual heterogeneity by leveraging the mask prompt and sample prompts. Additionally, we distill the mask prompt to reduce domain-specific biases and enhance efficient

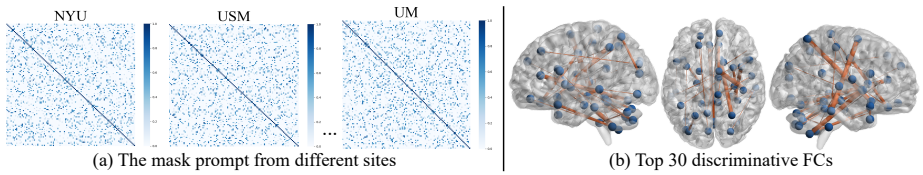


Fig. 4. Exploring the interpretability of mask prompts. Fig. 4(a): Visualization of mask prompts of different sites. Fig. 4(b): Top 30 discriminative FCs identified by BrainPrompt.

cross-domain knowledge transfer. Experimental results on the ABIDE dataset show that the classification performance of the proposed method outperforms the state-of-the-art methods for ASD diagnosis, especially at sites with limited samples. Furthermore, our framework holds promise for applications in other multi-site databases.

Acknowledgments. This research was supported by the National Natural Science Foundation of China (No.62076059), the Science and Technology Joint Project of Liaoning province (2023JH2/101700367, ZX20240193). Osmar Zaiane acknowledges the funding from NSERC and the Canada CIFAR AI Chairs Program.

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.: Brain graph synthesis by dual adversarial domain alignment and target graph prediction from a source graph. Medical Image Analysis **68**, 101902 (2021)
2. 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). Front Neuroinform **42**(10.3389) (2013)
3. Cui, W., Du, J., Sun, M., Zhu, S., Zhao, S., Peng, Z., Tan, L., Li, Y.: Dynamic multi-site graph convolutional network for autism spectrum disorder identification. Computers in Biology and Medicine **157**, 106749 (2023)
4. 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)
5. Dong, Z., Wu, Y., Chen, Z., Zhang, Y., Jin, Y., Zhou, J.H.: Prompt your brain: Scaffold prompt tuning for efficient adaptation of fmri pre-trained model. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 512–521. Springer (2024)
6. Farahani, F.V., Karwowski, W., Lighthall, N.R.: Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review. frontiers in Neuroscience **13**, 585 (2019)

7. Feng, X., Li, X., Hu, S., Zhang, D., Zhang, J., Chen, X., Huang, K., et al.: Memvlt: Vision-language tracking with adaptive memory-based prompts. Advances in Neural Information Processing Systems **37**, 14903–14933 (2025)
8. Ge, C., Huang, R., Xie, M., Lai, Z., Song, S., Li, S., Huang, G.: Domain adaptation via prompt learning. IEEE Transactions on Neural Networks and Learning Systems (2023)
9. Han, X., Xue, R., Du, S., Gao, Y.: Inter-intra high-order brain network for asd diagnosis via functional mris. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 216–226. Springer (2024)
10. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al.: Lora: Low-rank adaptation of large language models. ICLR **1**(2), 3 (2022)
11. Jin, C., Huang, E., Chang, H., Wang, Y., Cao, P., Zaiane, O.R.: Reusing transferable weight increments for low-resource style generation. In: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. pp. 2470–2488 (2024)
12. Kunda, M., Zhou, S., Gong, G., Lu, H.: Improving multi-site autism classification based on site-dependence minimisation and second-order functional connectivity. bioRxiv pp. 2020–02 (2020)
13. Li, X., Dvornek, N.C., Zhuang, J., Ventola, P., Duncan, J.S.: Brain biomarker interpretation in asd using deep learning and fmri. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part III 11. pp. 206–214. Springer (2018)
14. 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)
15. Liu, L., Wen, G., Cao, P., Hong, T., Yang, J., Zhang, X., Zaiane, O.R.: Braintgl: A dynamic graph representation learning model for brain network analysis. Computers in Biology and Medicine **153**, 106521 (2023)
16. Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., Neubig, G.: Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM Computing Surveys **55**(9), 1–35 (2023)
17. Liu, X., Wu, J., Li, W., Liu, Q., Tian, L., Huang, H.: Domain adaptation via low rank and class discriminative representation for autism spectrum disorder identification: A multi-site fmri study. IEEE Transactions on Neural Systems and Rehabilitation Engineering **31**, 806–817 (2023)
18. Mandell, D.S., Novak, M.: The role of culture in families' treatment decisions for children with autism spectrum disorders. Mental retardation and developmental disabilities research reviews **11**(2), 110–115 (2005)
19. Senoussaoui, M., Kenny, P., Stafylakis, T., Dumouchel, P.: A study of the cosine distance-based mean shift for telephone speech diarization. IEEE/ACM Transactions on Audio, Speech, and Language Processing **22**(1), 217–227 (2013)
20. Song, R., Cao, P., Wen, G., Zhao, P., Huang, Z., Zhang, X., Yang, J., Zaiane, O.R.: Braindas: Structure-aware domain adaptation network for multi-site brain network analysis. Medical Image Analysis **96**, 103211 (2024)
21. Van Erven, T., Harremos, P.: Rényi divergence and kullback-leibler divergence. IEEE Transactions on Information Theory **60**(7), 3797–3820 (2014)
22. Wang, B., Pourshafeie, A., Zitnik, M., Zhu, J., Bustamante, C.D., Batzoglou, S., Leskovec, J.: Network enhancement as a general method to denoise weighted biological networks. Nature communications **9**(1), 3108 (2018)

23. Wang, M., Zhang, D., Huang, J., Yap, P.T., Shen, D., Liu, M.: Identifying autism spectrum disorder with multi-site fmri via low-rank domain adaptation. IEEE transactions on medical imaging **39**(3), 644–655 (2019)
24. Yan, Y., Zhu, J., Duda, M., Solarz, E., Sripada, C., Koutra, D.: Groupinn: Grouping-based interpretable neural network for classification of limited, noisy brain data. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. pp. 772–782 (2019)
25. Zhang, P., Wen, G., Cao, P., Yang, J., Zhang, J., Zhang, X., Zhu, X., Zaiane, O.R., Wang, F.: Brainusl: U nsupervised graph s tructure l earning for functional brain network analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 205–214. Springer (2023)