DEEP DAG LEARNING OF EFFECTIVE BRAIN CONNECTIVITY FOR FMRI ANALYSIS

Yue Yu[⋆] *Xuan Kan*[†] *Hejie Cui*[†] *Ran Xu*[†] *Yujia Zheng*[‡] *Xiangchen Song*[‡] *Yanqiao Zhu*[§]
Kun Zhang[‡] *Razieh Nabi*[†] *Ying Guo*[†] *Chao Zhang*[⋆] *Carl Yang*[†*]

[⋆] Georgia Tech [†] Emory [‡] CMU [§] UCLA

ABSTRACT

Functional magnetic resonance imaging (fMRI) has become one of the most common imaging modalities for brain function analysis. Recently, graph neural networks (GNN) have been adopted for fMRI analysis with superior performance. Unfortunately, traditional functional brain networks are mainly constructed based on similarities among region of interests (ROIs), which are noisy and can lead to inferior results for GNN models. To better adapt GNNs for fMRI analysis, we propose DABNet, a **D**eep D**A**G learning framework based on **B**rain **Net**works for fMRI analysis. DABNet adopts a brain network generator module, which harnesses the DAG learning approach to transform the raw time-series into effective brain connectivities. Experiments on two fMRI datasets demonstrate the efficacy of DABNet. The generated brain networks also highlight the prediction-related brain regions and thus provide interpretations for predictions.

***Index Terms*—** fMRI Analysis, Brain Network, Direct Acyclic Graph Generation, Graph Neural Network

1. INTRODUCTION

Human brains play a vital role in human neurological systems. Recently, there exists abundant research progresses in neuroscience research showing that neural circuits are key to understanding the differences in brain functioning between populations [1]. Functional magnetic resonance imaging (fMRI) is one of the most commonly used imaging modalities to investigate brain function and organization [2]. There has been a significant increase of interest in utilizing fMRI for brain connectome analysis, which focuses on comprehending the brain organizations and their changes, as well as identifying disease-specific biomarkers.

Current brain network analyses are typically composed of two steps [3]: The first step is generating functional brain networks from individuals' fMRI data. Then, the obtained brain connectivity measures between nodes are used to classify individuals or predict clinical outcomes. However, these brain network generation methods focus on capturing the statistical associations between ROIs. Since correlation does not imply causation, they provide insufficient understandings of the complicated brain organization. Furthermore, the generated brain networks based on correlations can be noisy and inaccurate, which hinders the identifications of biological insights on the structure of brain networks and increase the time complexity of the downstream analysis.

Researchers have proposed a particular type of brain network, effective brain networks [4], which can overcome these two flaws. This type of brain network aims to infer causal relationships among brain regions and produce sparse connections [5, 6]. However, these methods often model the brain connectivity with overly simplistic assumptions, such as the absence of unmeasured confounding and lack of temporal dependencies. In reality, such assumptions are hard to satisfy. Moreover, existing works based on constraint- or score-based methods for brain connectivity generation [5, 7] are usually evaluated on a selected ROI subset (less than 50 regions) for their difficulty on scalability. But in real application scenarios, there exist hundreds of ROIs, and directly adopting these methods could take several hours, or even several days for each instance. Till now, generating effective and interpretable brain networks remains a challenging problem.

Fortunately, there is a recent trend in the machine learning community to view structure learning as a directed acyclic graphs (DAG) generation problem, which can be further converted to a continuous optimization constrained by additional regularizations to ensure acyclicity [8, 9]. Then, it can be solved with gradient-based techniques, which are efficient and can be combined with other deep learning models.

Motivated by these studies, we propose DABNet, a brain network generation approach via modeling the connections among ROIs as DAGs to identify effective brain connectivities and predict the target in an end-to-end fashion. To tackle the inscalability issue, we leverage the recently proposed approach [9] to recast the DAG structure learning task as a gradient-based optimization problem, which benefits from GPU acceleration and scales to hundreds of brain regions.

We evaluate DABNet on two real-world fMRI benchmarks datasets [10, 11] for the important and accessible task of biological sex prediction. The results illustrate that DABNet achieves competitive performance when compared with advanced baselines. Besides, DABNet is able to characterize the most important brain regions for the target tasks, justifying its efficacy on providing clinically useful interpretations.

*Corresponding author: Carl Yang <j.carlyang@emory.edu>

2. RELATED WORK

Functional Magnetic Resonance Imaging (fMRI) has been widely used to discover the functional connectivity (FC) between different regions in the brain. Traditional methods usually rely on hand-crafted features to infer the functional connectivity [12, 13], which can often be time-consuming. With the development of graph neural networks (GNNs), much attention has been paid to *GNN models for fMRI-based brain network analysis* [14, 15]. One major advantage of such models is that they can aggregate node features based on graph structures and can be trained in an end-to-end fashion.

However, GNNs usually rely on a pre-defined graph structures, which are often absent for fMRI analysis. To tackle this issue, some works directly use the brain networks constructed based on statistical correlations [16], but such correlations can be noisy and hurt the performance on downstream tasks. There are also some works that jointly generate the brain networks and predict for the downstream tasks. But they generates brain networks based on structure similarity from GNNs [17] or attention weights [18], which still cannot well characterize the complex relationships among ROIs.

On the other hand, learning Directed Acyclic Graphs (DAGs) from data is a fundamental problem in machine learning. Traditional methods include constraint- and score-based techniques [19], but the major drawback is the *computational inefficiency*, as the search problem is NP-hard [20]. To reduce the computational overhead, [8, 9, 21] recasts the DAG search problem as a continuous optimization task, where an additional penalty term is used to enforce the acyclicity of the generated graph. As recent works demonstrate the existence of causal links among ROIs [6], it is possible to leverage SL approaches to capture the relations among different ROIs.

3. METHOD

In this section, we first introduce the task studied in this paper, then introduce DAG learning techniques to generate effective brain networks. Finally, we discuss the overall pipeline.

3.1. Task Definition

In this study, the input $\boldsymbol{X} \in \mathbb{R}^{n\times v\times t}$ is the BOLD time-series for ROIs, where n is the sample size, v is the number of ROIs and t is the length of time-series. The target output for each time-series consists of two terms: (1) a functional brain network $\boldsymbol{A} \in \mathbb{R}^{v\times v}$ (i.e., brain connectivity matrix) for each sample $\boldsymbol{x} \in \mathbb{R}^{v\times t}$, which serves as an intermediate result to the end-to-end framework. (2) the prediction $\boldsymbol{Y} \in \mathbb{R}^{n\times|\mathcal{C}|}$, where $|\mathcal{C}|$ is the number of classes.

3.2. DAG Structure Learning from BOLD Signals

$\diamond$ **Formulation.** For the input BOLD time-series $\boldsymbol{x}$, generating DAGs to capture the effective connectivities is not easy since the current signal of a brain region is related to previous signals from both itself and other ROIs [22]. In this work, we hypothesize that different ROIs influence one another in both contemporaneous and time-lagged manner, and harvest the standard SVAR model [23] for the DAG generation process. In the t-th step, the SVAR model can be expressed as

$$\boldsymbol{x}_t = \boldsymbol{A}^\top \boldsymbol{x}_t + \boldsymbol{A}_1^\top \boldsymbol{x}_{t-1} + \ldots + \boldsymbol{A}_p^\top \boldsymbol{x}_{t-p} + \boldsymbol{\epsilon}_t. \quad (1)$$

To interpret Eq. 1, we note that $\boldsymbol{x}_t$ is an v-dimensional vector that represents the signal of all ROIs at the t-th step, $\boldsymbol{\epsilon}_t$ is a noise vector with independent elements [8], and p is the autoregressive order. Besides, $\boldsymbol{A}_i$ $(i = 1, 2, \ldots, p)$ represent weighted adjacency matrices with nonzero entries corresponding to the time-lagged relations, and $\boldsymbol{A}$ is the directed acyclic graph (DAG) to model the effective relations among ROIs. To effectively model the BOLD signals among different time steps, we further assume the DAG structure is time-invariant and transform Eq. 1 into the matrix form as

$$\boldsymbol{X} = \boldsymbol{A}^\top \boldsymbol{X} + \boldsymbol{A}_1^\top \boldsymbol{X}_1 + \ldots + \boldsymbol{A}_p^\top \boldsymbol{X}_p + \boldsymbol{E}, \quad (2)$$

where $\boldsymbol{X}$ is a $v \times (T+1-p)$ matrix with each column stands for one timestep, $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_p$ are time-lagged versions of $\boldsymbol{X}$. Denote $\boldsymbol{Z} = [\boldsymbol{X}_1|\cdots|\boldsymbol{X}_p]$ and $\boldsymbol{B} = \left[\mathbf{A}_1^\top|\cdots|\mathbf{A}_p^\top\right]^\top$, we obtain the compact form of Eq. 2 as $\boldsymbol{X} = \boldsymbol{A}^\top \boldsymbol{X} + \boldsymbol{B}^\top \boldsymbol{Z} + \boldsymbol{E}$.

$\diamond$ **Optimization Objective for Vanilla DAG Learning.** To learn the DAG structure for brain networks with $\boldsymbol{X}, \boldsymbol{Z}$, we aim to estimate the matrix $\boldsymbol{A}$ and $\boldsymbol{B}$ to satisfy Eq. 2 as well as the directed acyclic constraint (for $\boldsymbol{A}$ only).

This indeed creates a constraint optimization problem as

$$\min_{\boldsymbol{A},\boldsymbol{B}} \ \ell(\boldsymbol{A}, \boldsymbol{B}) = \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{A}^\top \boldsymbol{X} - \boldsymbol{B}^\top \boldsymbol{Z}\|_F^2 \ \text{ s.t. } \boldsymbol{A} \text{ is acyclic.} \quad (3)$$

Here $\ell(\boldsymbol{A}, \boldsymbol{B})$ is the objective for DAG generation. Directly optimizing Eq. 3 is difficult due to the hard acyclicity constraint [8]. To circumvent this issue, we reformulate the above task to an *unconstrained optimization problem* via applying additional *soft sparsity* and *acyclic* regularization for learning a DAG equivalent to the ground truth DAG [9].

The overall objective is defined as

$$\min_{\boldsymbol{A},\boldsymbol{B}} \mathcal{S}(\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{X}) = \mathcal{L}(\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{X}) + \lambda_1 \mathcal{R}_{\text{DAG}}(\boldsymbol{A}) + \lambda_2 \left(\mathcal{R}_{\text{sparse}}(\boldsymbol{A}) + \mathcal{R}_{\text{sparse}}(\boldsymbol{B})\right) \quad (4)$$

where

$$\mathcal{L}(\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{X}) = -\frac{1}{2}\sum_{i=1}^{v} \log\left(\sum_{j=1}^{T+1-p}\left(\boldsymbol{X}_{i,j} - \boldsymbol{A}_i^\top \boldsymbol{X}_{:,j} - \boldsymbol{B}_i^\top \boldsymbol{Z}_{:,j}\right)^2\right) + \log|\det(I - \boldsymbol{A})| + \text{const.} \quad (5)$$

$\mathcal{L}(\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{X})$ is the learning objective of the maximum log-likelihood estimator of $\boldsymbol{X}, \boldsymbol{Z}$ following multivariate Gaussian distribution. In addition, we use l_1 penalty to approximate the sparse constraint and use the DAG constraint proposed in [8] as the realization of $\mathcal{R}_{\text{DAG}}$ which can be written as

$$\mathcal{R}_{\text{sparse}}(\boldsymbol{A}) = \|\boldsymbol{A}\|_1, \quad \mathcal{R}_{\text{DAG}}(\boldsymbol{A}) = \text{tr}\left(e^{\boldsymbol{A}\circ\boldsymbol{A}}\right) - v. \quad (6)$$

By optimizing $\boldsymbol{A}$ and $\boldsymbol{B}$ with the above equations, we are able to obtain the brain network $\boldsymbol{A}$, which will be also used in the downstream task described as follows.

3.3. GNN for Brain Networks

With the generated brain connectivity $\boldsymbol{A}$, GNNs have been widely used in fMRI-based brain network analysis [14]. In this study, we leverage the *modified* k-layer graph convolutional network architecture [24] to accommodate negative edges in $\boldsymbol{A}$ for clinical prediction tasks. Let $\boldsymbol{\Theta}^{(k)} = (\boldsymbol{\theta}_1^{(k)}, \boldsymbol{\theta}_2^{(k)}, \ldots, \boldsymbol{\theta}_v^{(k)})$ be the matrix of all node embedding vectors at step k, the update rule for node embedding is

$$\boldsymbol{\Theta}^{(k)} = \mathrm{ReLU}\left(\widehat{\boldsymbol{A}}_+ \boldsymbol{\Theta}^{(k-1)} \boldsymbol{W}_+^{(k)}\right), \tag{7}$$

where $\widehat{\boldsymbol{A}}_+ = \boldsymbol{D}_+^{-\frac{1}{2}} \left(I + \mathrm{ReLU}(\boldsymbol{A})\right) \boldsymbol{D}_+^{-\frac{1}{2}}$ stands for the normalized adjacency matrices for positive edges only. $\boldsymbol{D}_+$ is the diagonal matrices representing node degrees. $\boldsymbol{W}^{(k)}$ stands for learnable parameters in convolutional layers and $\boldsymbol{\Theta}^{(0)} = \boldsymbol{F}$ is the *connection vector* that stands for Pearson correlation scores between its time-series with all the nodes contained in the brain network, which is suggested by [15]. We concatenate all the node embedding from the last GNN layer $\boldsymbol{\Theta}^{(k)}$ and use an additional MLP layer for the target prediction as

$$\widehat{\boldsymbol{y}} = \sigma\left(\mathrm{MLP}\left(\mathrm{BatchNorm1D}(\|_{j=1}^{v} \boldsymbol{\theta}_j^{(k)})\right)\right), \tag{8}$$

where $\widehat{\boldsymbol{y}}$ is the probability simplex of the prediction and $\sigma(\cdot)$ is the softmax function. We finally use the cross-entropy loss $\mathcal{L}_{\text{tgt}} = \ell_{\text{CE}}(\widehat{\boldsymbol{y}}, y)$ to train the GNN model.

3.4. End-to-end Training

It is worth noting that the GNN classifier module and DAG structural learning module can be trained in an end-to-end way, as the label y and the task-oriented graphs are learned simultaneously. Moreover, the brain networks of the test set are inferred *separately* from the training set. Thus, we eliminate the issue of information leakage, as *no information* from the test set has been used during the tuning process.

4. EXPERIMENTS

In this section, we conduct extensive experiments to answer the following three questions: **RQ1**: How does DABNet perform as compared with state-of-the-art methods? **RQ2**: How do the key designs in DABNet affect performance? **RQ3**: Does the generated brain connectivity by DABNet offers reasonable interpretability for target prediction tasks?

4.1. Experiment Setup

◇ **Dataset.** We conduct experiments using two real-world fMRI datasets: *(a) Philadelphia Neuroimaging Cohort (PNC)* is curated from the University of Pennsylvania and the Children's Hospital of Philadelphia. It includes a population-based sample of individuals aged 8–21 years [10]. After quality control, we use rs-fMRI data of 503 subjects, with 289 of them being females. Each sample contains 264 nodes collected through 120 time steps. *(b) Adolescent Brain Cognitive Development Study (ABCD)* [11] is one of the largest public fMRI datasets. This study recruited children aged 9–10 years in the U.S. Each child is followed into early adulthood, with repeated imaging scans and cognitive testing. After quality control, ABCD includes fMRI data of 7901 children, and 3961 among them are female. Each sample contains 360 nodes with 512 time points.

◇ **Task.** We choose biological sex prediction as the evaluation task since it is an essential aspect of adolescent development and serves as a critical task for ABCD and PNC. We split the train/valid/test dataset to 7:1:2 for evaluation.

◇ **Evaluation Metrics.** As the label distributions of both PNC and ABCD datasets are balanced, we use both *AUROC* and *accuracy* as the performance metrics. For accuracy, we use 0.5 as the threshold after obtaining the predicted result.

◇ **Implementations.** We implement our model in PyTorch[1]. We use Adam as the optimizer with the learning rate 1e-4. The key hyperparameters in DABNet include regularization weight λ_1, λ_2 and the number of steps p in SVAR model. Following common practice [9], we set $\lambda_1 = 10, \lambda_2 = 1$ without further tuning. Following [7], we set $p = 3$ as further increasing p will introduce more learnable parameters which causes OOM error for large datasets (ABCD).

4.2. Baselines

We compare DABNet with the following baselines:

(a) Time-series Model. We use *bi-GRU* as the baseline to encode BOLD time-series *without* modeling brain networks.

(b) Statistical Methods for Brain Networks. It sets the weight in adjacency matrices as Pearson correlation of BOLD signals, named as *GNN-Pearson*.

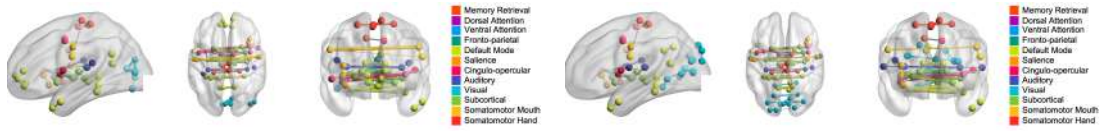
(c) Deep Learning Models for Brain Networks. *BrainnetCNN* [25], *BrainGNN* [14], and *BrainGB* [15] are three representative methods that develops advanced neural networks with the fixed correlation-based functional brain networks to model the relations among different ROIs.

(d) Models with Learnable Brain Network Generation. *LDS* [26], *FBNetGen* [17] are two baselines which generate brain networks and perform predictions jointly. LDS learns graph structures and model parameters through bilevel optimization. FBNetGen learns task-aware brain networks based on embedding similarities. Note that we use the same GNN classifier to adapt these methods for target tasks.

4.3. Main Experiment (RQ1)

From the experimental results shown in Table 1, we observe that DABNet outperforms all the baselines for both PNC and ABCD dataset with 3% and 2% performance gain in terms of accuracy. This is because ABCD is a much larger dataset (16 times larger than PNC) with plenty of labels, and directly using embedding similarity is sufficient to learn the relations

[1] https://pytorch.org/

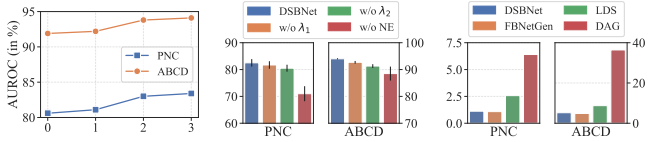


(a) Male (Label 0) (b) Female (Label 1)

Fig. 1: Visualizations of predominant connectivities generated via DABNet on different biological sexes. Edges spanning multiple neural systems are colored gray, whereas those linking nodes within the same neural system are colored appropriately.

Type	Method	Dataset: PNC		Dataset: ABCD	
		AUROC	Accuracy	AUROC	Accuracy
Time-series	bi-GRU	65.1 ± 3.5	58.1 ± 2.4	51.2 ± 1.0	49.9 ± 0.8
Statistical	GNN-Pearson	76.5 ± 2.7	69.2 ± 1.8	91.0 ± 0.5	82.4 ± 0.5
Deep Model	BrainnetCNN [25]	78.5 ± 3.2	71.9 ± 4.9	93.5 ± 0.3	85.7 ± 0.8
	BrainGNN [14]	77.5 ± 3.2	70.6 ± 4.8	OOM	OOM
	BrainGB [15]	76.6 ± 5.0	69.8 ± 4.2	91.8 ± 0.3	83.1 ± 0.9
Learnable Graph	LDS [26]	78.2 ± 3.8	70.8 ± 6.2	90.7 ± 0.3	82.5 ± 0.9
	FBNetGen [17]	80.1 ± 3.3	72.5 ± 2.3	93.8 ± 0.7	86.2 ± 1.2
Ours	DABNet	**82.5 ± 1.4**	**76.0 ± 1.6**	**94.0 ± 0.2**	**87.9 ± 0.4**

Table 1: Performance (in %) comparison with different types of baselines. Note that **OOM** means out-of-memory error.



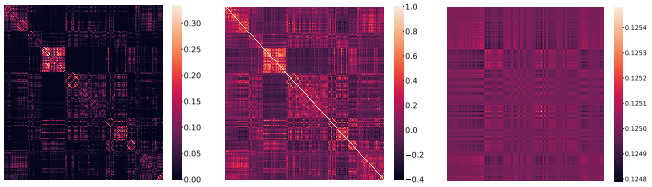
(a) p (b) Ablation Study (c) Running Time

Fig. 2: Effect of different components (a, b) and running time (in hours) (c) of DABNet.

among ROIs. Conversely, when labeled data is limited, FBNetGen cannot perform as well as DABNet. On the other hand, directly using statistical correlations as brain connectivity is insufficient to capture the relationships among ROIs, as they achieve suboptimal results on two datasets. Designing more complicated GNN architectures does not address this challenge — baselines with carefully designed GNN models [15, 25] still underperform DABNet on both datasets. These results justify the advantage of DABNet for generating more effective brain networks to benefit downstream tasks.

4.4. Ablation Studies and Efficiency Analysis (RQ2)

We further examine the effect of key modules in our graph generation module (Section 3.2). From Fig 2a, we find that when the time-lagged length p is larger, the performance gain is more significant. However, this will also lead to larger memory overhead. As a result, we keep $p = 3$ to balance between the performance and efficiency. From Fig 2b, we demonstrate that the regularization of DAG (λ_2) and sparsity (λ_2), as well as the strategy for handling negative weight (Eq. 7) are beneficial for target tasks. Removing any of them hurts the performance. From Fig 2c, we illustrates that DABNet is efficient when compared with other structure learning approaches, thanks to the GPU acceleration. On the contrary, Conventional DAG approaches (shown as DAG in Fig 2c) is inefficient and takes more than 1 day for large datasets.



(a) DABNet (b) Pearson (c) FBNetGen

Fig. 3: Generated brain networks of DABNet and baselines.

4.5. Case Studies (RQ3)

We visualize the connections of the learned brain networks between two biological sex groups in Fig. 1. Specifically, we divide the learned graphs based on their class labels and calculate an average network by averaging weights of each edge within the same class. The top 100 predominant edges are visualized using the BrainNet Viewer [27]. By comparing Fig. 1(a) and Fig. 1(b), we observe that the main difference in connections between two groups lies in the Default Mode Network (DMN), the Auditory Network (Aud) and the Visual Network (Vis), which is in line with previous studies [28] that ROIs with significant biological sex differences are located in the DMN and Aud systems.

Besides, we plot the generated brain network of DABNet and baselines in Fig 3. From the result, we find that both Pearson correlation graph and FBNetGen are suboptimal — the Pearson correlation graph (Fig. 3b) is too dense and contains many negative edges. The network generated by FBNetGen (Fig. 3c) is almost fully-connected, as all the edges have similar weight. In contrast, DABNet produces sparse connections among different ROIs and preserves the edge strength patterns observed in Pearson graphs, indicating that DABNet can highlight meaningful signals in the brain connectivities.

5. CONCLUSION

In this paper, we propose DABNet, an effective brain connectivity generation approach to support clinical predictive tasks. In particular, we leverage the DAG learning techniques to encode the relations among ROIs. Our end-to-end framework is efficient with GPU acceleration. Experiments on two real-world datasets illustrate the better performance of DABNet when compared with advanced baselines, and demonstrate valuable neurological interpretations on downstream tasks.

6. ETHICAL STATEMENTS

The research work is related to human studies. The PNC and ABCD dataset used in this study are de-identified and owned by a third-party organization, where informed consent was obtained for all subjects. All studies are conducted according to the Good Clinical Practice guidelines and U.S. 21 CFR Part 50 (Protection of Human Subjects), and under the approval of Institutional Review Boards.

7. REFERENCES

[1] L. Williams, "Precision psychiatry: a neural circuit taxonomy for depression and anxiety," *The Lancet*, 2016.

[2] G. Ganis and S. Kosslyn, "Neuroimaging," in *Encyclopedia of the Human Brain*. 2002.

[3] S. Smith, "The future of fmri connectivity," *Neuroimage*, 2012.

[4] G. Deshpande and X. Hu, "Investigating effective brain connectivity from fmri data: past findings and current issues with granger causality analysis," *Brain connectivity*, 2012.

[5] J. Runge, "Causal network reconstruction from time series: From theoretical assumptions to practical estimation," *Chaos*, 2018.

[6] S. Siddiqi, K. Kording, J. Parvizi, and M. Fox, "Causal mapping of human brain function," *Nature reviews neuroscience*, 2022.

[7] S. Saetia, N. Yoshimura, and Y. Koike, "Constructing brain connectivity model using causal network reconstruction approach," *Frontiers in Neuroinformatics*, 2021.

[8] Xun Zheng, Bryon Aragam, Pradeep K Ravikumar, and Eric P Xing, "Dags with no tears: Continuous optimization for structure learning," *NeurIPS*, 2018.

[9] Ignavier Ng, AmirEmad Ghassami, and Kun Zhang, "On the role of sparsity and dag constraints for learning linear dags," *NeurIPS*, 2020.

[10] T. Satterthwaite, M. Elliott, K. Ruparel, J. Loughead, K. Prabhakaran, et al., "Neuroimaging of the Philadelphia Neurodevelopmental Cohort," *NeuroImage*, 2014.

[11] B. Casey, T. Cannonier, and M. Conley et al., "The adolescent brain cognitive development (abcd) study: Imaging acquisition across 21 sites," *Developmental Cognitive Neuroscience*, 2018.

[12] T. Nichols and A. Holmes, "Nonparametric permutation tests for functional neuroimaging: a primer with examples," *Human brain mapping*, 2002.

[13] J. Lukemire, S. Kundu, G. Pagnoni, and Y. Guo, "Bayesian joint modeling of multiple brain functional networks," *JASA*, 2021.

[14] X. Li, Y. Zhou, S. Gao, N. Dvornek, M. Zhang, et al., "Braingnn: Interpretable brain graph neural network for fmri analysis," *Medical Image Analysis*, 2021.

[15] H. Cui, W. Dai, Y. Zhu, X. Kan, A. Gu, et al., "BrainGB: A Benchmark for Brain Network Analysis with Graph Neural Networks," *ArXiv.org*, 2022.

[16] S. Smith, K. Miller, G. Salimi-Khorshidi, M. Webster, et al., "Network modelling methods for FMRI," *NeuroImage*, 2011.

[17] X. Kan, H. Cui, J. Lukemire, Y. Guo, and C. Yang, "FBNETGEN: Task-aware GNN-based fMRI analysis via functional brain network generation," in *MIDL*, 2022.

[18] U. Mahmood, Z. Fu, V. Calhoun, and S. Plis, "A deep learning model for data-driven discovery of functional connectivity," *Algorithms*, 2021.

[19] D. Chickering and D. Heckerman, "Efficient approximations for the marginal likelihood of bayesian networks with hidden variables," *Machine learning*, 1997.

[20] Gerhard J. Woeginger, *Exact Algorithms for NP-Hard Problems: A Survey*, 2003.

[21] R. Pamfil, N. Sriwattanaworachai, S. Desai, P. Pilgerstorfer, et al., "Dynotears: Structure learning from time-series data," in *AISTATS*, 2020.

[22] Anish Mitra, Abraham Z Snyder, Carl D Hacker, and Marcus E Raichle, "Lag structure in resting-state fmri," *Journal of neurophysiology*, 2014.

[23] N. Swanson and C. Granger, "Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions," *JASA*, 1997.

[24] T. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *ICLR*, 2017.

[25] J. Kawahara, C. Brown, S. Miller, B. Booth, et al., "BrainNetCNN: Convolutional neural networks for brain networks," *NeuroImage*, 2017.

[26] L. Franceschi, M. Niepert, M. Pontil, and X. He, "Learning discrete structures for graph neural networks," in *ICML*, 2019.

[27] M. Xia, J. Wang, and Y. He, "Brainnet viewer: a tool for brain connectomics," *PloS one*, 2013.

[28] D. Satterthwaite, H. Wolf, R. Roalf, K Ruparel, G. Erus, et al., "Linked Sex Differences in Cognition and Functional Connectivity in Youth," *Cerebral Cortex*, 2015.