Causal Brain Connectivity: Integrating Granger Directed Graphs in fMRI Analysis

Tianyi Zhang[1], Keqi Han[1], Jiawei Nie[1], Chenyu You[2], Sanne van Rooij[3], Jennifer Stevens[3], Boadie Dunlop[3], Charles Gillespie[3], and Carl Yang[1]

[1] Department of Computer Science, Emory University, Atlanta, GA, USA
{tzha325, keqi.han, jiawei.nie, j.carlyang}@emory.edu
[2] Department of Computer Science, Stony Brook University, Stony Brook, NY, USA
chenyu.you@stonybrook.edu
[3] Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, USA
{sanne.van.rooij, jennifer.stevens, bdunlop, cgilles}@emory.edu

Abstract. Functional Magnetic Resonance Imaging (fMRI) has significantly advanced our understanding of human brains by capturing dynamic neural activities, providing basis for causal analysis between brain regions. However, conventional correlation-based analyses often fail to account for the directionality and complexity of neural interactions. We propose an approach that integrates Granger causality with graph-based deep learning to better capture *effective connectivity* between brain regions. Specifically, we compare three methods: Multilayer Perceptron (MLP)-based approaches on flattened time series, Graph Convolutional Networks (GCNs) using undirected connectivity, and a GCN framework incorporating directed Granger-causal influences into brain graph construction. Through the optimization of Granger parameters such as the lag order via Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), we investigate the impact of different graph construction methods on connectome-based outcome prediction. The directed graph framework demonstrates robustness to hyperparameter variations, while also providing biologically plausible insights into brain functionalities that complement undirected correlation-based graphs. Evaluations on classification and regression tasks using large-scale fMRI datasets reveal that directionality preserves predictive performance while offering additional understanding of information flow within brain networks. These findings emphasize the potential of Granger-causality-informed graphs for robust, nuanced, and causality-aware fMRI analyses.

1 Introduction

Functional magnetic resonance imaging (fMRI) has significantly advanced our understanding of brain function by non-invasively measuring blood oxygen level-dependent signals (BOLD), thus capturing *dynamic* aspects of neural activity. However, many existing analyses continue to rely on *static, undirected pairwise correlations* between regions of interest (ROIs), overlooking the *directional*

and *causal* nature of neural interactions [7]. This omission limits our ability to determine how brain regions *causally influence* each other and restricts our understanding of the underlying mechanisms of neural communication.

Recent advances in *effective connectivity* emphasize modeling how activity causally propagates between brain regions, rather than merely identifying co-activation patterns [1, 2]. *Granger causality (GC)* is a well-established technique for estimating effective connectivity that provides a robust statistical framework to infer directed influences from time-series data, offering deeper insights into the flow and directionality of neural interactions than correlation-based methods [3–5]. In recent years, Graph Neural Networks (GNNs) have attracted broad interest for modeling graph-structured interactions between brain regions. Several pioneering methods have been developed for cognition assessment and neural disorder detection [6–8]. However, to the best of our knowledge, most existing neuroimaging applications of GNNs are built on correlation-based brain connectivity, which overlooks the causal directionality inherent in brain activity.

In this work, we propose a novel framework that integrates Granger-causal directed graphs with Graph Convolutional Networks (GCNs) to rigorously model *effective connectivity* in fMRI data. We validate our approach on the Adolescent Brain Cognitive Development Study (ABCD) and the Philadelphia Neurodevelopmental Cohort (PNC) datasets, which are large neurodevelopmental cohorts well-suited for studying how functional networks evolve over development. Our approach centers on four insights. First, while correlation-based connectivity captures co-activation patterns, it overlooks the *causal directionality* essential for understanding how neural signal propagate. We address this by constructing *directed adjacency matrices* via Granger causality, capturing asymmetrical information flow across brain regions. Second, we systematically tune Granger parameters (e.g., lag order) using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), ensuring both statistical validity and scalability to large fMRI datasets. Third, integrating these Granger-causal graphs enables GCNs to explicitly model directed causal influences in brain networks, leveraging both local and global connectivity patterns for more accurate clinical outcome prediction. Finally, our Granger-causality-informed GCN directly models *directed, causal influences*, offering a complementary perspective to traditional correlation-based methods. Comprehensive validation shows that incorporating causality-guided connections delivers robust, competitive performance across diverse clinical tasks, while revealing deeper insights into neural information flow. Overall, our contributions bridge the gap between statistical causal inference and graph-based deep learning, addressing limitations of static functional connectivity analysis and paving the way for a new generation of *causality-aware* fMRI studies.

2 Problem Formulation

This study aims to perform downstream prediction using resting-state fMRI (rs-fMRI) data by modeling the brain as a network of interacting regions. To enhance predictive accuracy and interpretability, we incorporate directionality and causal inference into graph-based connectivity representations [12].

Using rs-fMRI data from multiple cohorts, BOLD signals are recorded across N Regions of Interest (ROIs) over T time steps. For a given subject m, the data is represented as a matrix $\mathbf{X}^{(m)} \in \mathbb{R}^{N \times T}$, where $x_{i,t}^{(m)}$ denotes the BOLD intensity at ROI i and time step t. Each subject is associated with a label $y^{(m)}$, which may correspond to a continuous cognitive score or a categorical behavioral classification, depending on the prediction task. The objective is to construct graph-based connectivity representations from $\mathbf{X}^{(m)}$ and use them to predict $\hat{y}^{(m)}$ while capturing causal relationships within brain networks to improve the understanding of functional brain organization and its connection to cognition and behavior.

3 Method

In this section, we first introduce two existing fMRI-based brain analysis paradigms: MLP-Based approaches and graph-based models utilizing undirected functional connectivity. We then present the proposed method based on Granger causality, which can capture directional interactions between brain regions, providing insights into causal relationships and temporal dependencies that are not detectable in undirected correlation-based networks.

Multilayer Perceptron-Based Methods

A baseline approach to fMRI-based prediction is to apply a Multilayer Perceptron (MLP) directly to the BOLD signals, treating the data as a high-dimensional input. Instead of explicitly modeling connectivity between brain regions, this method learns feature representations through fully connected layers.

For a given subject m, the fMRI data $\mathbf{X}^{(m)} \in \mathbb{R}^{N \times T}$ is flattened into a feature vector $\mathbf{z}^{(m)} \in \mathbb{R}^{N \cdot T}$, which serves as the input to an MLP:

$$\hat{y}^{(m)} = f_\theta(\mathbf{z}^{(m)}),$$

where f_θ represents the MLP model with learnable parameters θ. The MLP consists of multiple fully connected layers, each followed by a non-linear activation function.

Graph-Based Models

For each subject m, the node feature matrix $\mathbf{X}^{(m)} \in \mathbb{R}^{N \times T}$ is derived from time-series BOLD signals, where N represents the number of ROIs and T represents the number of time steps. Each row of $\mathbf{X}^{(m)}$ encapsulates the temporal activity of an ROI. Adjacency matrices encode ROI relationships in two forms. Undirected graphs represent functional connectivity based on Pearson correlation coefficients, producing symmetric adjacency matrices $\mathbf{C}^{(m)} \in \mathbb{R}^{N \times N}$. Directed graphs capture effective connectivity using Granger causality, resulting in directed adjacency matrices $\mathbf{A}^{(m)} \in \mathbb{R}^{N \times N}$.

Undirected Graphs: Functional Connectivity. Functional connectivity between ROIs is computed using Pearson correlation coefficients. For each subject m, the adjacency matrix $\mathbf{C}^{(m)}$ is defined as:

$$C_{ij}^{(m)} = \frac{\sum_t \left(x_{i,t}^{(m)} - \mu_i^{(m)}\right)\left(x_{j,t}^{(m)} - \mu_j^{(m)}\right)}{\sqrt{\sum_t \left(x_{i,t}^{(m)} - \mu_i^{(m)}\right)^2}\sqrt{\sum_t \left(x_{j,t}^{(m)} - \mu_j^{(m)}\right)^2}},$$

where $\mu_i^{(m)}$ and $\mu_j^{(m)}$ are the mean values of ROI i and ROI j time-series, respectively. The resulting graph $\mathcal{G}^{(m)}$ models symmetric relationships between ROIs, with edge weights reflecting the strength of functional connectivity.

Directed Graphs: Effective Connectivity. Effective connectivity is captured using Granger causality [3–5]. First, each ROI time-series is segmented into overlapping sliding windows with a specified window size W and step size S. For each window w, the mean signal is computed as

$$\bar{x}_{i,w}^{(m)} = \frac{1}{W}\sum_{t \in w} x_{i,t}^{(m)}$$

which produces a window-averaged time-series $\bar{\mathbf{X}}^{(m)} \in \mathbb{R}^{N \times W_{\text{windows}}}$. Next, for every window and each ROI pair (i, j), a Vector Autoregressive (VAR) model of order L is fitted to predict the time-series of ROI j from its own past values and those of ROI i, mathematically expressed as

$$x_j(t) = \alpha_{j,0} + \sum_{k=1}^{L} \alpha_{j,k} x_j(t-k) + \sum_{k=1}^{L} \beta_{j,k} x_i(t-k) + \epsilon_j(t),$$

where $\alpha_{j,0}$ is the intercept, $\alpha_{j,k}$ and $\beta_{j,k}$ are lag coefficients, and $\epsilon_j(t)$ is the error term [9]. Hypothesis testing then determines whether the coefficients $\beta_{j,k}$ are significantly non-zero; if so, a directed edge $(i \to j)$ is added to the adjacency matrix $\mathbf{A}^{(m)}$. Finally, causal relationships detected over multiple windows are aggregated to construct the final binary matrix:

$$A_{ij}^{(m)} = \begin{cases} 1 & \text{if ROI } i \text{ Granger-causes ROI } j, \\ 0 & \text{otherwise.} \end{cases}$$

The directed graph $\mathcal{G}_{\text{dir}}^{(m)} = (\mathcal{V}, \mathcal{E}_{\text{dir}})$ encodes causal influences, capturing the flow of information among ROIs.

We evaluate the constructed graphs using a Graph Convolutional Network (GCN) [7, 10]. Specifically, GCNs are applied separately to undirected $\mathbf{C}^{(m)}$ and directed $\mathbf{A}^{(m)}$. In undirected GCNs, $\mathbf{C}^{(m)}$ serves as the adjacency matrix for symmetric feature propagation, while directed GCNs use $\mathbf{A}^{(m)}$ to incorporate directional dependencies. The model generates embeddings, which are evaluated on downstream classification and regression tasks.

4 Experiments

Our experimental design is structured to address three central research questions. **RQ1**: Does model performance vary across different graph construction methods? To answer this, we compare models based on flattened time-series

data, undirected correlation graphs, and directed Granger-causal graphs. **RQ2**: What parameters best optimize Granger-causal graphs? We systematically tune hyperparameters such as window size, step size, and lag order using model selection criteria (e.g., AIC and BIC) to identify the optimal configurations for constructing directed graphs. **RQ3**: How do directed connectivity insights complement undirected graphs? We compare the interpretability and neurobiological insights derived from directed Granger-causal graphs with those from undirected correlation-based graphs, focusing on their respective contributions to understanding brain connectivity.

4.1 Experimental Settings

Datasets. We use two large-scale neuroimaging datasets: the Adolescent Brain Cognitive Development Study (ABCD) and the Philadelphia Neurodevelopmental Cohort (PNC).

Table 1: Summary of Datasets and Tasks

Dataset	Task	Atlas	# Subjects	# Time Steps	# Nodes	Response	# Classes
ABCD [12]	Classification	HCP 360	7,901	512	360	Gender	2
ABCD [12]	Regression	HCP 360	4,613	1,024	360	Cognitive Score	–
PNC [13]	Classification	Power 264	503	120	264	Gender	2

Nodes: The number of brain Regions of Interest (ROIs) used for constructing connectivity matrices (e.g., the HCP 360 atlas divides the brain into 360 ROIs).

ABCD Dataset. The ABCD dataset tracks 9-10-year-olds through early adulthood with repeated MRI scans [12]. It includes 7,901 subjects, parcellated using the HCP 360 atlas [7, 12], with a balanced gender distribution. For gender prediction tasks, 512 time steps are used, while regression tasks predicting the Cognition Summary Score utilize 1,024 time steps. Samples with fewer than 1,024 time steps were excluded, resulting in 4,613 samples for regression.

PNC Dataset. The PNC dataset comprises 503 subjects from the University of Pennsylvania and Children's Hospital of Philadelphia, also with a balanced gender distribution [13]. Each subject provides 120 time steps of rs-fMRI data from 264 ROIs, with preprocessing steps including motion correction, normalization, and bandpass filtering [13].

Metrics. For binary gender classification, we use the Area Under the Receiver Operating Characteristic (AUC), Accuracy, and F1 Score, reflecting the model's ability to distinguish between classes, overall correctness, and balance of precision and recall. The classification threshold is set at 0.5 as a standard choice for binary classification. For cognitive score regression, Mean Squared Error (MSE) is employed to measure the average squared difference between predicted and actual scores.

Implementation Details. All models were optimized using Adam with a learning rate and weight decay of 1×10^{-4} (fine-tuned as in BrainGB), with data split 80:20 (train:test) and the training set further partitioned via k-fold cross-validation to keep subjects exclusive to one set [7, 11]. We used binary cross-entropy for classification, MSE for regression, training up to 100 epochs with early stopping (patience = 10), a batch size of 16 (validation checked every 5 epochs), and mixup regularization (`mixup`=1). Experiments were repeated 5

times with different seeds. We evaluated two architectures: the **MLP**, which flattens time-series data and passes it through two fully connected (FC) layers (512 and 256 units, ReLU, dropout = 0.5) with sigmoid (classification) or linear (regression) output; and the **GCNs**, which utilize two graph convolution layers (256 units, ReLU), followed by a 256-unit FC layer (dropout = 0.5) and an output layer with sigmoid or linear activation.

Baselines We compare our proposed model with a range of baselines: (i) *MLP Approach*—a time-series MLP that directly encodes BOLD data without network modeling; (ii) *Direction-Free Deep Learning Approaches*—including BrainNetCNN, FCNet, and BrainGB with functional connectivity (BrainGB w/FC), which exploit correlation-based connectivity features via BrainGB's GCN architecture to model ROI relationships without directionality [7, 17, 18]; (iii) *Effective Connectivity Method*—which estimates directed interactions using the Noise-Diffusion Network (NDNetwork) followed by a GCN module [7, 10, 16].

4.2 Model Performance (RQ1)

Table 2: Performance Comparison Across Models and Tasks

Type	**Method**	**Dataset: PNC**			**Dataset: ABCD**			
		AUC ↑	**Acc** ↑	**F1** ↑	**AUC** ↑	**Acc** ↑	**F1** ↑	**MSE** ↓
Time-series	MLP	53.80	53.28	52.48	51.29	50.38	49.17	106.43
Direction-Free	BrainNetCNN [18]	52.73	54.16	53.33	54.36	53.29	54.10	85.81
	FCNet [17]	52.95	51.64	50.82	50.36	51.29	50.10	103.65
	BrainGB w/FC [7, 10]	<u>56.39</u>	**56.74**	**55.32**	<u>63.32</u>	**59.74**	<u>59.33</u>	**79.93**
Effective Connectivity	NDNetwork [10, 16]	54.46	53.58	50.95	TLE	TLE	TLE	TLE
Directed (Ours)	BrainGB w/GC [4, 5, 10, 7]	**56.74**	<u>56.35</u>	<u>54.93</u>	**63.76**	<u>59.65</u>	**59.56**	<u>81.32</u>

Bold values denote the best performance, **underlined** values the second-best. **Arrows** indicate the desired direction (↑ = higher is better; ↓ = lower is better), and **TLE** means training exceeds 72 hours.

Graph-Based Models Outperform MLP. Table 2 shows that graph-based models outperform the baseline MLP. By incorporating functional connectivity, these models better capture the spatial and temporal organization of neural systems than the MLP, which treats each feature independently.

Compatibility of Directed and Undirected Graphs. Our findings indicate that incorporating directionality does not degrade performance—BrainGB w/GC attains comparable results to the undirected BrainGB w/FC—while also providing additional insight into causal interactions. In the next section, we further demonstrate that the dispersed connectivity patterns revealed by the directed model uncover interesting neurobiological patterns, offering complementary and plausible insights into how the connectome and brain regions interact.

4.3 Hyperparameter Study for Directed Graphs (RQ2)

The construction of directed Granger-causal graphs involves optimizing three key hyperparameters: window size (W), step size (S), and lag order (L) for Vector Autoregressive (VAR) models. These parameters are critical for accurately capturing temporal dependencies and causal relationships in fMRI time-series data.

The **window size (W)** defines the number of consecutive time steps in each window, where larger values capture longer temporal dependencies. The **step size (S)** determines the shift between windows; smaller values improve temporal resolution by increasing overlap, while larger steps reduce computational load but may miss finer dynamics. The **lag order (L)** specifies the number of past time steps considered in the VAR model, balancing complexity and predictive accuracy. Higher lag orders risk overfitting, whereas lower ones may omit significant temporal dependencies.

To select the optimal L, we employ the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). AIC is calculated as $\text{AIC} = 2k - 2\ln(L)$, where k is the number of model parameters and $\ln(L)$ is the log-likelihood [14]. BIC applies a stricter penalty for complexity, defined as $\text{BIC} = k\ln(n) - 2\ln(L)$, where n is the number of observations [15]. The lag order minimizing both AIC and BIC is chosen as optimal.

We conducted a grid search over combinations of W, S, and L, summarized in Table 3.

Table 3: Grid Search Results for Granger Causality Hyperparameters

Window Size (W)	**Step Size (S)**	**Lag Order (L)**	**AIC** [14]	**BIC** [15]	**AUC**	**Accuracy**	**F1**
32	16	1	79.59	<u>83.90</u>	63.32	59.44	59.33
32	16	2	78.10	85.11	62.82	58.93	58.75
32	16	3	<u>77.34</u>	86.92	63.03	59.19	59.08
32	32	1	78.40	**82.70**	<u>63.62</u>	<u>59.61</u>	**59.71**
32	32	2	77.72	84.73	63.12	59.21	59.01
32	32	3	**76.96**	86.53	**63.76**	**59.65**	<u>59.56</u>
64	32	1	167.37	173.79	62.53	58.02	58.07
64	32	2	166.26	176.90	63.04	58.58	58.56
64	32	3	165.43	180.21	63.58	59.05	59.07
64	64	1	161.80	168.23	62.89	58.20	58.24
64	64	2	160.97	171.61	63.20	58.64	58.62
64	64	3	160.44	175.21	63.55	59.00	59.03
128	64	1	352.23	360.76	61.82	57.94	57.83
128	64	2	350.05	364.24	62.19	58.21	58.19
128	64	3	348.88	368.68	62.40	58.54	58.32
128	128	1	339.90	348.43	62.71	58.86	58.60
128	128	2	337.63	351.81	63.05	59.11	58.99
128	128	3	336.39	356.18	63.36	59.33	59.22

Bold values denote the best performance, **underlined** values the second-best.

Findings *Despite the assumption that lower AIC and BIC would yield better results, performance remained robust under various hyperparameter settings.* Table 3 shows that although ($W = 32, S = 32, L = 3$) achieves the lowest AIC (76.96) and highest AUC (63.76%) and Accuracy (59.65%), ($W = 32, S = 32, L = 1$) attains the lowest BIC (82.70) and the highest F1 score (59.71%). These variations underscore a trade-off when balancing AIC versus BIC, yet both settings deliver competitive performance. More broadly, the grid search reveals that small adjustments to window size, step size, and lag order generally yield stable outcomes, suggesting that meaningful temporal dependencies are captured across a spectrum of configurations rather than relying solely on minimizing AIC or BIC.

4.4 Neurological Insights (RQ3)

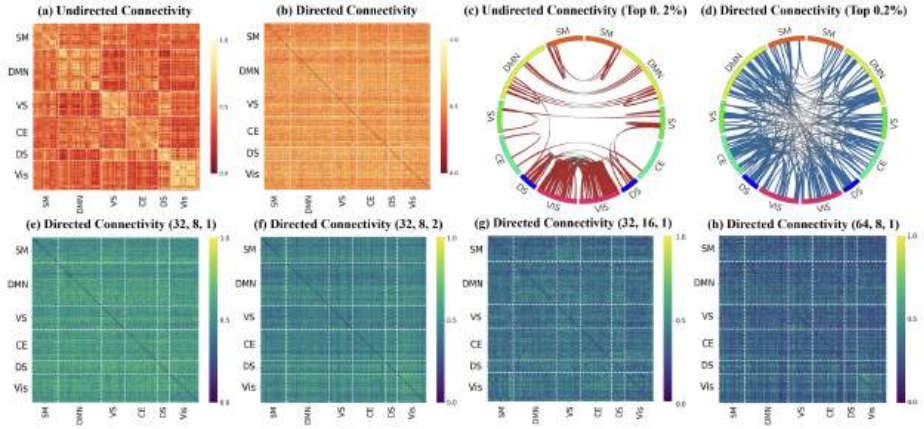

Fig. 1: Plots are from the ABCD dataset using the HCP 360 atlas, with nodes reordered into six functional network regions. Connectivity strength is rescaled to [0,1] for direct comparison. The top row compares undirected and directed connectivity matrices and chord plots (top 0.2% strongest connections), while the bottom row presents directed connectivity matrices under various modeling parameters, illustrating how these choices impact inferred connectivity.

In Figure 1, the undirected connectivity in *panel a* reveals discernible neurological structure, with intra-regional clustering aligned with key networks such as the Sensorimotor (SM), Default Mode (DMN), and Visual (Vis) regions [19, 20]. In contrast, the directed connectivity in *panel b* emphasizes asymmetric, hierarchical interactions that disrupt such clustering. Thus, while undirected connectivity captures reciprocal co-activation from shared inputs, directed connectivity exposes broader, causal inter-regional influences. Moreover, because undirected connectivity is symmetric—with each connection represented twice—the chord plot in *panel c* shows roughly half as many unique edges as that from the directed matrix in *panel d*, reinforcing that undirected approaches reflect primarily intra-network co-activation.

The *bottom row* presents directed adjacency matrices (*panels e–h*) derived from Granger causality analyses under various hyperparameter settings ("W32 S8 L1," "W32 S8 L2," "W32 S16 L1," and "W64 S8 L1"), which modulate sparsity and network structure. These variations offer plausible neurological insights, such as elucidating how SM activation may precede and potentially drive changes in other networks. However, as evidenced in Table 2, while tuning parameters like window size, step size, and lag order can optimize model fit in terms of lower AIC and BIC, they do not necessarily translate into enhanced predictive performance. This highlights the complexity of interpreting effective connectivity and underscores the need for careful calibration when inferring causal neural dynamics.

Even in a resting state, the brain exhibits spontaneous, organized activity that reflects its intrinsic functional architecture. For example, the sensorimotor

cortex may remain active in a baseline readiness to process sensory inputs or prepare motor responses. Recent studies suggest that resting-state connectivity is not mere noise but rather reveals stable network patterns, even in the absence of overt tasks [21]. In this view, specific activations in regions like the sensorimotor cortex arise because they must remain functionally prepared, preserving a baseline of interaction that can quickly adapt to external stimuli or initiate motor actions when necessary.

5 Conclusion

This study integrates directed graph structures into brain connectivity analysis for gender classification, addressing three key research questions.
RQ1: Graph-based models outperform MLPs by capturing complex neural interactions via functional connectivity, with the directed model (BrainGB w/ GC) delivering comparable performance to its undirected counterpart (BrainGB w/ FC) while providing additional causal insights.
RQ2: Our experiments show that the directed graph construction using Granger causality remains stable across varying hyperparameters such as window size and lag order.
RQ3: While undirected methods capture broad co-activation patterns, directed graphs reveal dispersed, asymmetrical causal pathways, offering complementary insights into inter-regional interactions.

Despite promising results, challenges remain, including the computational demands of Granger causality at higher lag orders and the need for specialized architectures for directed graphs. Future work should focus on more efficient causal inference and models tailored to directed connectivity, ultimately advancing our understanding of neural information flow in cognitive and clinical settings.

6 Acknowledgment

This research was partially supported by the US National Science Foundation under Award Number 2319449 and Award Number 2312502.

References

1. Stephan, K.E. et al.: Analyzing effective connectivity with functional magnetic resonance imaging. Wiley Interdisciplinary Reviews: Cognitive Science. 1, 3, 446–459 (2010). https://doi.org/10.1002/WCS.58.
2. Chuang, K.C. et al.: Brain effective connectivity and functional connectivity as markers of lifespan vascular exposures in middle-aged adults: The Bogalusa Heart Study. Frontiers in Aging Neuroscience. 15, (2023). https://doi.org/10.3389/fnagi.2023.1110434.
3. Granger, C.W.J.: Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. Econometrica. 37, 3, 424–438 (1969). https://doi.org/10.2307/1912791.
4. Deshpande, G. et al.: Effective connectivity during haptic perception: A study using Granger causality analysis of functional magnetic resonance imaging data. NeuroImage. 40, 4, 1807–1814 (2008). https://doi.org/10.1016/J.NEUROIMAGE.2008.01.044.

5. Shojaie, A., Fox, E.B.: Granger Causality: A Review and Recent Advances. Annual Review of Statistics and Its Application. 9, 1, 289–319 (2022). https://doi.org/10.1146/annurev-statistics-040120-010930.
6. Li, X. et al.: BrainGNN: Interpretable Brain Graph Neural Network for fMRI Analysis. Medical Image Analysis. 74, 102233 (2021). https://doi.org/10.1016/J.MEDIA.2021.102233.
7. Cui, H. et al.: BrainGB: A Benchmark for Brain Network Analysis with Graph Neural Networks (Extended Abstract). 4968–4969 (2022). https://doi.org/10.1109/TMI.2022.3218745.
8. Kan, Xuan, et al. "Dynamic brain transformer with multi-level attention for functional brain network analysis." In 2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE, 2023.
9. Sims, C.A.: Macroeconomics and reality. Econometrica. 48, 1, 1–48 (1980). https://doi.org/10.2307/1912017.
10. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv:1609.02907 (2016)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv:1412.6980 (2014)
12. Casey, B.J. et al.: The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Developmental Cognitive Neuroscience. 32, 43–54 (2018). https://doi.org/10.1016/J.DCN.2018.03.001.
13. Satterthwaite, T.D. et al.: Neuroimaging of the Philadelphia neurodevelopmental cohort. NeuroImage. 86, 544–553 (2014). https://doi.org/10.1016/J.NEUROIMAGE.2013.07.064.
14. Akaike, H.: Akaike's information criterion. In: *International Encyclopedia of Statistical Science*, 25–25. Springer, Berlin, Heidelberg (2011)
15. Neath, A.A., Cavanaugh, J.E.: The Bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**(2), 199–203 (2012)
16. Gilson, M. et al.: Estimation of Directed Effective Connectivity from fMRI Functional Connectivity Hints at Asymmetries of Cortical Connectome. PLOS Computational Biology. 12, 3, (2016). https://doi.org/10.1371/JOURNAL.PCBI.1004762.
17. Riaz, A. et al.: Fcnet: a convolutional neural network for calculating functional connectivity from functional mri. Connectomics in NeuroImaging: First International Workshop, CNI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Proceedings 1. Proceedings 1, 70–78 (2017).
18. Kawahara, J. et al.: BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage. 146, 146, 1038–1049 (2017). https://doi.org/10.1016/J.NEUROIMAGE.2016.09.046.
19. Smallwood, J. et al.: The default mode network in cognition: a topographical perspective. Nature Reviews Neuroscience. 22, 8, 503–513 (2021). https://doi.org/10.1038/S41583-021-00474-4.
20. Hale, T.S. et al.: Visual network asymmetry and default mode network function in ADHD: An fMRI study. Frontiers in Psychiatry. 5, 81, 81 (2014). https://doi.org/10.3389/FPSYT.2014.00081.
21. Preti, M.G., Bolton, T.A.W., Van De Ville, D.: The dynamic functional connectome: State-of-the-art and perspectives. NeuroImage. 160, 41–54 (2017). https://doi.org/10.1016/j.neuroimage.2016.12.061.