BrainWeaver: Weaving Dynamic Brain Networks with ROI-Guided Attention for Cognitive Assessment

Tao Liu, Fanwei Zhu, Haijun Han, Zengwei Zheng✉, and Binbin Zhou✉

Hangzhou City University, Hangzhou, China
2240101009@stu.hzcu.edu.cn, {zhufw,hanhj,zhengzw,bbzhou}@hzcu.edu.cn

Abstract. Accurate and interpretable assessment of cognitive functions from neuroimaging data is critical for understanding cognitive health. Existing methods either rely on subjective assessments or struggle to capture the complex, time-varying interplay between brain regions, hindering both predictive accuracy and neurophysiological interpretability. In this paper, we introduce BrainWeaver, a novel framework designed to overcome these limitations by weaving task-relevant neural patterns from dynamic connectomes. Specifically, BrainWeaver uses an ROI-guided dynamic multi-graph fusion module to capture spatial dependencies and temporal fluctuations in fMRI data. To enhance interpretability, we integrate neuroscience priors via a hierarchical graph representation module that reinforces key ROI contributions. Moreover, we employ a contrastive learning paradigm to disentangle cognitive patterns from modality-specific noise and inter-subject variability. Extensive experiments on the large-scale HCP S1200 benchmark show that BrainWeaver outperforms state-of-the-art baselines across three distinct cognitive domains. Ablation and interpretability analyses further demonstrate the substantial contribution of neuroscience-guided hierarchical representation and reveal task-relevant brain network patterns. Code is available at https://github.com/ZJUDataIntelligence/BrainWeaver.

Keywords: Dynamic brain networks · Graph neural networks · Cognitive assessment

1 Introduction

Accurate and objective assessment of cognitive functions is critical for cognitive science and clinical neurology. Cognition is multifaceted requiring a robust predictive framework to generalize across distinct domains. In this work, we focus on three core dimensions, including *working memory* for reasoning and decision-making [1], *inhibitory control* reflecting executive attention [7], and *general cognitive status* for cognitive impairment screening [9]. However, traditional cognitive assessments heavily rely on behavioral paradigms [21]. These methods are fundamentally limited because they rely on indirect behavioral data and are easily

✉ Corresponding authors: Zengwei Zheng and Binbin Zhou.

compromised by subjective factors like participant fatigue [29]. They also lack direct neurophysiological interpretability [35]. Thus, developing objective, robust, and interpretable cognitive assessment methods based on direct neuroimaging data remains an urgent challenge.

The advent of neuroimaging coupled with artificial intelligence has enabled more direct assessment. Early work applied machine learning [24, 29] or deep learning models [27] to functional Magnetic Resonance Imaging (fMRI) and electroencephalography (EEG) signals [2, 12] to decode cognitive states or predict clinical outcomes [28, 29]. Brain connectomes naturally present a high dimensional non-Euclidean graph structure of Regions of Interest (ROIs) and their connections, making Graph Neural Networks (GNNs) [11,16,31] a superior tool. GNNs have achieved remarkable success in clinical diagnosis, such as identifying disease-relevant brain regions [4, 14], enhancing predictive robustness [25], and integrating multimodal data for psychiatric diagnosis [6, 8, 26, 33].

Despite these advances, existing graph-based methods face critical limitations. The first gap is the *static assumption*. Many GNN-based methods treat functional connectivity (FC) as static by averaging brain activity over time [27], ignoring rapid temporal reconfigurations of brain networks driven by spontaneous activity. Dynamic graph neural networks [10, 19, 23, 36] and recent neuroimaging adaptations [15,32] attempt to capture spatial and temporal features but *overlook neuroscience prior knowledge*. The multiple demand (MD) system theory shows that diverse cognitive tasks rely on a shared neural scaffold primarily anchored in the prefrontal and parietal cortex [5]. Current dynamic models fail to hierarchically prioritize this domain knowledge. Existing interpretable models like BrainGNN [17] focus on salient regions but are limited to static graphs. Lastly, fMRI data's *noisy nature*, high inter-subject variability and imaging artifacts, hinders the extraction of task-relevant neural signatures. While contrastive learning [20] enables robust representation learning in neuroscience [3,18], its integration with prior-guided dynamic networks remains unexplored.

To address these combined challenges, we propose BrainWeaver as a novel framework designed to integrate task relevant neural patterns from dynamic connectomes by explicitly incorporating neuroscientific priors. BrainWeaver systematically overcomes these limitations through three components. To capture the dynamic nature of brain function, we utilize a multi-graph temporal attention fusion guided by ROIs. This accurately models the complex spatial dependencies and temporal fluctuations inherent in the time series data. To incorporate domain knowledge, we design a key ROI-specific hierarchical graph representation method, which leverages neuroscience priors to strategically reinforce the contributions of known functional areas enhancing both predictive performance and model interpretability. We further apply an individual invariance contrastive learning paradigm to disentangle universal cognitive patterns from modality-specific noise and differences among subjects ensuring a robust representation. We conduct extensive experiments on the large-scale HCP S1200 benchmark, demonstrating that BrainWeaver achieves competitive performance against state-of-the-art baselines across diverse cognitive domains, with con-

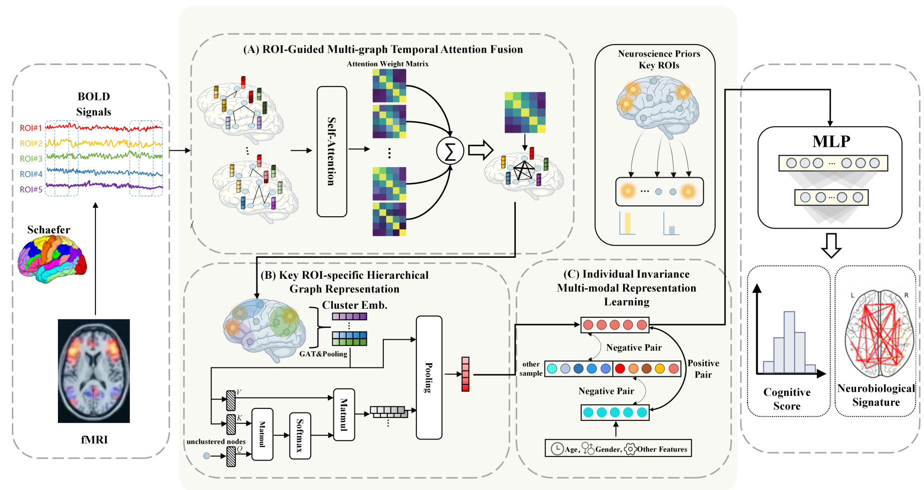

Fig. 1. Overview of BrainWeaver for cognitive assessment, consisting of three key components: (A) ROI-guided multi-graph temporal attention fusion module processing dynamic brain networks through self-attention mechanisms, (B) Key ROI-specific hierarchical graph representation module leveraging neuroscience prior knowledge to prioritize critical brain regions, and (C) Individual invariance multi-modal representation learning module using contrastive learning to reduce modality-specific noise and individual differences.

sistent benefits from the proposed prior-guided representation and contrastive learning modules.

2 Methods

The BrainWeaver framework is an end-to-end framework designed to robustly predict diverse cognitive scores, as shown in Fig. 1. For each subject, we first model the fMRI data as a sequence of dynamic brain graphs: $\{G_1, G_2, ..., G_M\}$, where M is the number of temporal graphs. Each graph $G_i = (V, E_i, X_i)$ is defined over a fixed set of N nodes (ROIs), $V = \{v_1, ..., v_N\}$. However, the edge sets E_i (functional connectivity) and node features $X_i \in \mathbb{R}^{N \times d}$ (e.g., BOLD signals) evolve over time. Our goal is to leverage this dynamic, graph-structured data to predict the subject's cognitive score.

2.1 ROI-Guided Multi-graph Temporal Attention Fusion

To address the failure of static models to capture brain dynamics, an ROI-guided multi-graph temporal attention fusion module is designed to learn a unified representation from the sequence of dynamic graphs. We hypothesize that by modeling the temporal fluctuations of connectivity, we can capture the task-relevant neural patterns missed by static averaging.

This module operates via intra-graph spatial attention and inter-graph temporal fusion. For each G_m in the sequence, we compute a node self-attention matrix A_m to capture the spatial dependencies at that specific time point:

$$A_m = \text{softmax}\left(\frac{Q_m K_m^T}{\sqrt{d_k}}\right) \tag{1}$$

where $Q^m = X^m W^Q$ and $K^m = X^m W^K$. To integrate information across time, we design a cross-graph attention fusion mechanism. This learns to aggregate node features ($\hat{X}$) and attention weights ($\hat{A}$) from all M graphs, assigning higher importance to the most salient time points. This is achieved by calculating temporal attention scores (β_k and γ) to produce a single fused graph representation $\hat{G}$ that encapsulates the most critical spatiotemporal information. For each node v_k, we extract its corresponding attention weight vectors from M graphs to form attention matrices $A_1[:, k], A_2[:, k], ..., A_M[:, k]$. Each attention vector is expanded to dimension $[N, 1]$, and then an individual attention network is used to obtain the inter-graph attention scores $S_k \in \mathbb{R}^{M \times N}$ for node v_k:

$$S_k = \sum_{j=1}^{n} G_k^q[:, :, j] \odot G_k^k[:, :, j] \tag{2}$$

where $G_k^q = W_g^q A_{:,k}$, $G_k^k = W_g^k A_{:,k}$, and $W_g^q \in \mathbb{R}^{d_g \times 1}$, $W_g^k \in \mathbb{R}^{d_g \times 1}$ are the query and key projection matrices for graph attention. d_g denotes the hidden dimension of the graph attention. After normalizing S_k, we obtain the inter-graph attention weights $\beta_k[i]$, and construct the fused attention weight matrix:

$$A_{[:,k]} = \sum_{i=1}^{M} \beta_k[i] \odot A_i[:, k] \tag{3}$$

Similarly, we also fuse node features from multiple graphs. We stack all node features into a tensor and compute feature fusion weights γ through a parallel attention network. The fused node feature matrix $\hat{X}$ is calculated as:

$$\hat{X} = \sum_{i=1}^{m} \gamma[i] \odot X_i \tag{4}$$

Then, we construct a fully connected fused graph $\hat{G} = (\mathcal{V}, \hat{E}, \hat{X})$ where every pair of nodes is connected by an edge with weights derived from the fused attention matrix A. The edge weights in the fused graph are defined as $\hat{E}_{jk} = A_{jk}$ for all node pairs (j, k).

2.2 Key ROI-specific Hierarchical Graph Representation

To enhance the interpretability, we design a hierarchical structure guided by a priori knowledge of the brain's functional architecture. As neuroscience literature [5] confirms, diverse cognitive tasks rely on a shared core network known as the MD system which is primarily anchored in the prefrontal cortex and posterior parietal cortex. We define a set of these key nodes $R = \{r_1, ..., r_K\}$.

We assign an initial importance coefficient to each brain region. Key ROI nodes are initialized with a value of 1, while other nodes are assigned a small value ϵ. A multilayer perceptron then learns to dynamically adjust the importance weight s_i of each node based on the fused input representation z_i:

$$s_i = \begin{cases} \sigma(\mathrm{MLP}(z_i)), & \text{if} \quad i \in R \\ \epsilon \cdot \sigma(\mathrm{MLP}(z_i)), & \text{otherwise} \end{cases} \tag{5}$$

where σ is the sigmoid function. Using key brain regions as core nodes, we form K clusters by selecting neighbors based on edge weights and neighbor importance. The cluster assignment probability for node j to cluster k is defined as:

$$p_{j \to k} = \sigma(w_{jk} \cdot s_j) \tag{6}$$

where w_{jk} is the edge weight between node j and the core node of cluster k.

Cluster representations are then generated using a Graph Attention Network (GAT). Then, each cluster is treated as a single node, and multi-head attention is applied to compute the attention from unclustered nodes to each cluster. During the graph pooling stage, we introduce local and global residual connections. The original unclustered node features and the global average pooled representation of all nodes are added to the final output to preserve fine grained spatial details and global structural context.

2.3 Individual Invariance Multi-modal Representation Learning

To address the high data noise and inter-subject variability, we design the individual invariance multi-modal representation learning module to reduce modality-specific noise (e.g., fMRI scan artifacts, behavioral sensor variations) and inter-individual differences to improve model robustness. Following recent advances in cross modal semantic alignment [3,13,20], we treat fMRI representations and clinical metadata as positive pairs for the same subject. All other combinations within a training batch are considered negative pairs. We project both the brain network representations and the encoded metadata into a shared embedding space as u_i and v_i respectively. We optimize the alignment using a symmetric InfoNCE loss defined as:

$$\mathcal{L}_{CL} = -\frac{1}{N}\sum_{i=1}^{N}\left(\log\frac{\exp(\mathrm{sim}(u_i, v_i)/\tau)}{\sum_{j=1}^{N}\exp(\mathrm{sim}(u_i, v_j)/\tau)} + \log\frac{\exp(\mathrm{sim}(v_i, u_i)/\tau)}{\sum_{j=1}^{N}\exp(\mathrm{sim}(v_j, u_j)/\tau)}\right) \tag{7}$$

where τ is a temperature hyperparameter and *sim* denotes cosine similarity. A prediction head maps the final graph embedding to the estimated cognitive score $\hat{y}_i$. The entire framework is trained end to end by minimizing a joint objective function that combines the mean squared error (MSE) of the prediction and the contrastive alignment loss:

$$\mathcal{L} = \mathcal{L}_{\mathrm{MSE}} + \lambda \mathcal{L}_{\mathrm{CL}} \tag{8}$$

where λ controls the regularization strength of the contrastive objective. To avoid leakage, metadata excludes all cognitive scores.

3 Experiment

3.1 Experimental Settings

Data. We validate our framework on the large-scale HCP S1200 dataset, which provides multimodal neuroimaging data from healthy young adults [30]. Following NeuroGraph [15,22], we use resting-state fMRI dynamic graphs constructed

Table 1. Performance comparison across three cognitive tasks.

Model	Working Memory			Inhibitory Control			General Cognitive Status		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
GCN	14.46 ± 0.77	11.53 ± 0.68	11.54 ± 0.71	12.54 ± 0.56	10.16 ± 0.51	10.15 ± 0.45	2.47 ± 0.08	1.92 ± 0.07	6.65 ± 0.22
GIN	14.90 ± 0.40	11.85 ± 0.33	11.63 ± 0.46	14.17 ± 1.11	11.77 ± 0.97	11.30 ± 0.94	3.79 ± 0.65	3.47 ± 0.67	11.87 ± 2.32
GAT	14.78 ± 0.82	11.69 ± 0.74	11.61 ± 0.82	14.59 ± 1.08	11.58 ± 0.88	11.61 ± 0.82	2.75 ± 0.43	2.12 ± 0.34	7.33 ± 1.19
GraphSAGE	14.99 ± 0.77	11.92 ± 0.50	11.92 ± 0.53	13.15 ± 0.68	10.60 ± 0.61	10.73 ± 0.64	2.44 ± 0.13	1.90 ± 0.09	6.60 ± 0.34
ASTGCN	13.80 ± 0.37	10.95 ± 0.39	11.01 ± 0.45	10.70 ± 0.33	8.76 ± 0.34	8.90 ± 0.33	1.72 ± 0.11	1.33 ± 0.08	4.62 ± 0.32
DySAT	14.84 ± 1.24	11.84 ± 1.05	11.73 ± 1.20	12.07 ± 3.04	10.09 ± 2.80	9.92 ± 2.43	2.46 ± 1.55	2.29 ± 1.60	7.85 ± 5.45
EvolveGCN	13.47 ± 0.34	10.74 ± 0.42	10.69 ± 0.46	10.74 ± 0.45	8.90 ± 0.51	8.80 ± 0.47	1.85 ± 0.40	1.66 ± 0.39	5.66 ± 1.33
TGCN	13.63 ± 0.28	10.89 ± 0.35	10.56 ± 0.40	10.53 ± 0.32	8.70 ± 0.40	8.63 ± 0.38	1.34 ± 0.06	1.17 ± 0.06	4.00 ± 0.22
STAGIN	13.62 ± 0.25	10.88 ± 0.34	11.05 ± 0.42	11.12 ± 0.28	8.92 ± 0.42	9.16 ± 0.36	1.38 ± 0.27	1.12 ± 0.28	3.93 ± 0.95
STNAGNN	13.63 ± 0.52	10.84 ± 0.52	10.63 ± 0.48	12.49 ± 0.88	10.52 ± 0.87	10.12 ± 0.78	3.94 ± 0.76	3.80 ± 0.79	12.98 ± 2.72
Neurograph	13.60 ± 0.49	10.81 ± 0.51	**10.50 ± 0.43**	10.30 ± 0.25	8.47 ± 0.27	8.62 ± 0.24	1.13 ± 0.15	0.93 ± 0.15	3.26 ± 0.52
BrainWeaver	**13.38 ± 0.40**	**10.66 ± 0.53**	10.77 ± 0.66	**10.26 ± 0.28**	**8.40 ± 0.28**	**8.55 ± 0.31**	**1.06 ± 0.08**	**0.81 ± 0.03**	**2.85 ± 0.14**

after brain parcellation, scanner-drift removal, motion correction, and signal normalization. Each graph uses 100 Schaefer ROIs as nodes, with edges defined by the top 10% positive Pearson correlations within each sliding window. The ROI time series is cropped to 150 time points and segmented with a window length of 50 and stride of 3, yielding $M = 34$ graph snapshots per subject.

Implementation Details. Our framework is implemented using PyTorch and PyTorch-Geometric. We employ the Adam optimizer with a learning rate of 0.008 and a weight decay of 1e-5. Training proceeds for 100 epochs with a batch size of 16 and an early stopping strategy based on the validation MSE. Model performance is evaluated using three standard regression metrics, including root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

Baselines. We compare our BrainWeaver against two categories of established graph models. Static graph models include GCN [16], GAT [31], GIN [34], and GraphSAGE [11]. Dynamic models include EvolveGCN [19], DySAT [23], T-GCN [36], ASTGCN [10], STAGIN [15], and STNAGNN [32]. We also include the standardized Neurograph evaluation baseline [22].

3.2 Results

Baselines Comparison. Table 1 presents the predictive performance across the three cognitive assessment tasks. Our BrainWeaver achieves the best overall performance, with the lowest prediction errors across all evaluated cognitive tasks and metrics. We also observe three insights. First, dynamic graph models consistently outperform static configurations. Models incorporating temporal processing significantly reduce prediction errors, confirming the necessity of capturing temporal fluctuations for cognitive decoding. Second, BrainWeaver further reduces prediction errors, while advanced dynamic models like EvolveGCN and TGCN achieve relatively low error margins by modeling spatiotemporal dynamics, demonstrating the distinct advantage of the incorporation of hierar-

Table 2. Ablation study on component contributions across three cognitive tasks.

Modules			Working Memory			Inhibitory Control			General Cognitive Status		
Fusion	Prior	CL	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
×	✓	✓	13.53 ± 0.34	10.84 ± 0.40	10.92 ± 0.46	10.61 ± 0.20	8.63 ± 0.24	8.81 ± 0.26	1.08 ± 0.10	0.84 ± 0.06	2.95 ± 0.23
✓	✓	×	14.13 ± 0.48	11.30 ± 0.47	11.47 ± 0.59	11.11 ± 0.78	8.97 ± 0.64	9.19 ± 0.65	1.19 ± 0.18	0.94 ± 0.12	3.30 ± 0.44
✓	×	✓	13.98 ± 0.79	11.22 ± 0.66	11.03 ± 0.31	12.02 ± 0.56	9.88 ± 0.45	9.98 ± 0.57	1.36 ± 0.41	1.14 ± 0.37	3.97 ± 1.32
✓	✓	✓	**13.38± 0.40**	**10.66± 0.53**	**10.77± 0.66**	**10.26± 0.28**	**8.40± 0.28**	**8.55± 0.31**	**1.06± 0.08**	**0.81± 0.03**	**2.85± 0.14**

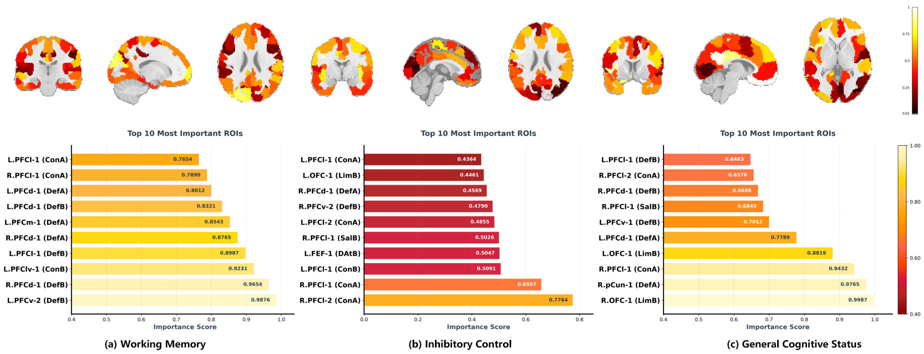

Fig. 2. Spatial distribution of top contributing anatomical regions identified by learned attention weights. The upper panels display normalized importance maps projected onto anatomical templates, while the lower panels list ranked top contributing ROIs.

chical neuroscience priors. Finally, BrainWeaver shows relatively low standard deviations across metrics, suggesting stable prediction performance across repeated runs and indicating that individual-invariance contrastive learning may contribute to more robust cognitive representations.

Ablation Study. We evaluate the individual contributions of the BrainWeaver components by systematically removing the ROI-guided multi-graph temporal attention fusion module (Fusion), the key ROI-specific hierarchical graph representation module (Prior), and the individual invariance multi-modal representation learning module (CL). Table 2 demonstrates that excluding the Prior module causes the most severe performance degradation, particularly for the inhibitory control and general cognitive status assessments. This highlights the importance of explicitly incorporating prior knowledge from neuroscience into the model. Removing the CL module results in the second largest error increase, which confirms the critical function of contrastive learning in mitigating modality-specific noise and high inter-subject variability inherent in fMRI data.

Brain Network Interpretability Analysis. To verify the neurobiological specificity of learned representations, we visualize regional importance weights on three tasks in Fig. 2. Distinct spatial patterns demonstrate that BrainWeaver captures functionally meaningful and behaviorally relevant neural configurations.

For working memory, importance weights are mainly distributed across dorsolateral, ventrolateral prefrontal cortices (PFC), posterior parietal cortex, and dorsal attention regions, together with default mode network (DMN), forming a strong fronto-parietal integration pattern consistent with sustained executive control and active information manipulation. For inhibitory control, the model highlights frontoparietal control and salience-related regions, including lateral PFC and frontal eye fields, showing a more distributed and balanced interaction pattern that reflects large-scale coordination during interference resolution. In contrast, general cognitive status shows a shift toward widespread default mode and associative cortical regions, particularly medial prefrontal and posterior cingulate areas, suggesting dense large-scale network coordination that supports global cognitive integrity rather than domain-specific executive processing.

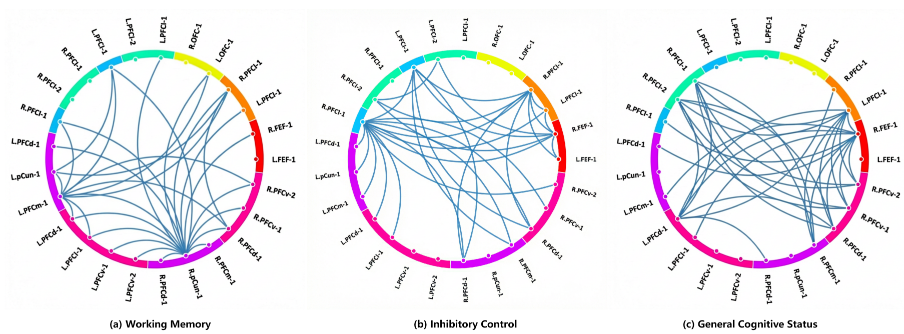

Fig. 3. Topological comparison of learned brain network connectivity across three cognitive domains.

We also present the topological comparison of learned brain network connectivity across three tasks as shown in Fig. 3. The working memory task exhibits a hub-centric topology centered on left prefrontal regions with strong fronto-parietal connections, reflecting the dependence on prefrontal executive control. The inhibitory control task shows a balanced mesh topology with bidirectional connections between frontal eye fields, prefrontal cortex, and parietal regions, embodying the collaborative nature of attention networks. The general cognitive status task presents a highly interconnected dense network with strong intra-DMN connections, reflecting the multi-domain characteristics of comprehensive cognitive assessment.

4 Conclusion

In this study, we introduce BrainWeaver, a novel framework designed for comprehensive and interpretable cognitive assessment. To overcome the limitations of static connectivity assumptions and generic graph structures, our model leverages an ROI-guided temporal attention module to capture the spatiotemporal dynamics inherent in functional brain networks. By explicitly integrating neuroscience priors into a hierarchical graph representation, the model strategically reinforces task-relevant neural pathways, thereby enhancing both predictive accuracy and neurophysiological interpretability. Furthermore, the incorporation of an individual invariance contrastive learning paradigm effectively disentangles core cognitive signatures from modality-specific noise and inter-subject variability. Extensive experiments demonstrate that BrainWeaver establishes a new SOTA performance across three distinct cognitive domains, i.e. working memory, inhibitory control, and general cognitive status. Future investigations will extend this framework to clinical populations to evaluate its diagnostic utility for early cognitive impairment screening and personalized psychiatric interventions.

Acknowledgments. This work was supported by the National Key R&D Program of China (No.2025YFG0100700), the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (No.2025C02068), the Natural Science Foundation of Zhejiang Province (No.LTGG24F020002, LY24F020013), the National Natural Science Foundation of China (No.62576304, 62102349), the Zhejiang Provincial Natural Science Foundation of China (No.LD26H300001), the Scientific Research Cultivation Fund of Hangzhou City University (No.J-202404) and the Scientific Research Cultivation Fund of Hangzhou City University: Special Project for Affiliated Hospitals (No.F-202408). The authors would like to acknowledge the Supercomputing Center of Hangzhou City University, for the support of advanced computing resources.

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

References

1. Baddeley, A.: Working memory. Memory pp. 71–111 (2020)
2. Balaji, S.S., Parhi, K.K.: Classifying subjects with pfc lesions from healthy controls during working memory encoding via graph convolutional networks. In: International IEEE/EMBS Conference on Neural Engineering (NER). IEEE (2023)
3. Chen, J., et al.: Bridging the semantic latent space between brain and machine: Similarity is all you need. In: AAAI Conference on Artificial Intelligence. vol. 38 (2024)
4. Chen, Z., Beeler-Duden, S., Lawson, S., Jacokes, Z., Van Horn, J.D., Pelphrey, K.A., Venkataraman, A.: Biscot: Behavior-informed subgroup-consistent connectome template for interpretable brain network analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 109–119. Springer (2025)
5. Duncan, J.: The multiple-demand (md) system of the primate brain: mental programs for intelligent behaviour. Trends in cognitive sciences **14**(4), 172–179 (2010)
6. D'Souza, N.S., Wang, H., Giovannini, A., Foncubierta-Rodriguez, A., Beck, K.L., Boyko, O., Syeda-Mahmood, T.F.: Fusing modalities by multiplexed graph neural networks for outcome prediction from medical data and beyond. Medical Image Analysis **93**, 103064 (2024)
7. Eriksen, B.A., Eriksen, C.W.: Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & psychophysics **16**(1), 143–149 (1974)

8. Feng, J., Xu, H., Cai, J., Chang, Y., Zhang, D., Du, S., Wang, J.: Cross-modal brain graph transformer via function-structure connectivity network for brain disease diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 247–256. Springer (2025)
9. Folstein, M.F., Folstein, S.E., McHugh, P.R.: "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research **12**(3), 189–198 (1975)
10. Guo, S., et al.: Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In: AAAI Conference on Artificial Intelligence. vol. 33 (2019)
11. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. Advances in Neural Information Processing Systems **30** (2017)
12. Jiang, W., et al.: Multimodal brain signal analysis with state-space modeling: A study on working memory. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE (2024)
13. Jin, J., Chai, H., Huang, X., Guo, X., Zheng, Z., Zhou, Z., Wang, J., Wang, X., Liu, J., Zhou, B.: Ultrasound-clip: Semantic-aware contrastive pre-training for ultrasound image-text understanding. arXiv preprint arXiv:2604.01749 (2026)
14. Kazi, A., Shekarforoush, S., Kortuem, K., Albarqouni, S., Navab, N., et al.: Self-attention equipped graph convolutions for disease prediction. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). pp. 1896–1899. IEEE (2019)
15. Kim, B.H., Ye, J.C., Kim, J.J.: Learning dynamic graph representation of brain connectome with spatio-temporal attention. Advances in Neural Information Processing Systems **34**, 4314–4327 (2021)
16. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
17. 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)
18. Lin, A., Wang, W., Han, H., Zhu, F., Ma, Q., Zheng, Z., Zhou, B.: Knowmdd: Knowledge-guided cross contrastive learning for major depressive disorder diagnosis. In: Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence. pp. 7536–7544 (2025)
19. Pareja, A., et al.: Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In: AAAI Conference on Artificial Intelligence. vol. 34 (2020)
20. Radford, A., et al.: Learning transferable visual models from natural language supervision. In: International Conference on Machine Learning. PMLR (2021)
21. Rodriguez, J.S., Paule, M.G.: Working memory delayed response tasks in monkeys. PubMed (2009)
22. Said, A., et al.: Neurograph: Benchmarks for graph machine learning in brain connectomics. In: Advances in Neural Information Processing Systems. vol. 36, pp. 6509–6531 (2023)
23. Sankar, A., et al.: Dysat: Deep neural representation learning on dynamic graphs via self-attention networks. In: International Conference on Web Search and Data Mining (2020)
24. Singh, P., et al.: Analysis of fmri signals from working memory tasks and resting-state of brain: Neutrosophic-entropy-based clustering algorithm. International Journal of Neural Systems **32**(04), 2250012 (2022)

25. Song, X., Frangi, A., Xiao, X., Cao, J., Wang, T., Lei, B.: Integrating similarity awareness and adaptive calibration in graph convolution network to predict disease. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 124–133. Springer (2020)
26. Song, X., Zhou, F., Frangi, A.F., Cao, J., Xiao, X., Lei, Y., Wang, T., Lei, B.: Multicenter and multichannel pooling gcn for early ad diagnosis based on dual-modality fused brain network. IEEE Transactions on Medical Imaging **42**(2), 354–367 (2022)
27. Sridhar, S., Romney, A., Manian, V.: A deep neural network for working memory load prediction from eeg ensemble empirical mode decomposition. Information **14**(9), 473 (2023)
28. Talarowska, M., Zboralski, K., Gałecki, P.: Correlations between working memory effectiveness and depression levels after pharmacological therapy. Psychiatria Polska **47**(2), 255–267 (2013)
29. Tran, T.M., Tran, T.T., Khanmohammadi, S.: High-order resting-state functional connectivity is predictive of working memory decline after brain tumor resection. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE (2024)
30. Van Essen, D.C., et al.: The wu-minn human connectome project: an overview. NeuroImage **80**, 62–79 (2013)
31. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)
32. Wang, J., Dvornek, N.C., Duan, P., Staib, L.H., Ventola, P., Duncan, J.S.: Stnagnn: Spatiotemporal node attention graph neural network for task-based fmri analysis. arXiv e-prints pp. arXiv–2406 (2024)
33. Wang, M., Guo, J., Wang, Y., Yu, M., Guo, J.: Multimodal autism spectrum disorder diagnosis method based on deepgcn. IEEE Transactions on Neural Systems and Rehabilitation Engineering **31**, 3664–3674 (2023)
34. Xu, K., et al.: How powerful are graph neural networks? arXiv preprint arXiv:1810.00826 (2018)
35. Zhang, Y., et al.: Prediction of working memory ability based on eeg by functional data analysis. Journal of Neuroscience Methods **333**, 108552 (2020)
36. Zhao, L., et al.: T-gcn: A temporal graph convolutional network for traffic prediction. IEEE Transactions on Intelligent Transportation Systems **21**(9), 3848–3858 (2019)