A Multi-View, Hybrid, Hypergraph Learning Framework with Causal Perturbation for Early Alzheimer's Disease Diagnosis

Yifan Jia[1], Luoyu Wang[1], Yitian Tao[1], Zihao Zhu[1], Yawen Zhang[1], and Han Zhang[1,2(✉)]

[1] School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China
[2] Shanghai Clinical Research and Trial Center, Shanghai 201210, China
zhanghan2@shanghaitech.edu.cn

Abstract. Early diagnosis of Alzheimer's disease (AD) is paramount for effective intervention. Resting-state functional magnetic resonance imaging (rs-fMRI) offers critical insights into how brain connectivity networks transition from cognitively normal (CN) to mild cognitive impairment (MCI, the early stage of AD) status. Most existing methods predominantly rely on *static* functional connectivity (FC), failing to capture *directional (i.e., causal)* and *dynamic* high-order synergistic interactions among brain regions. None of these methods utilize robust features from noisy and heterogeneous multi-view rs-fMRI representations. We propose HyperBrainNet, a causality-aware hybrid hypergraph network for better capturing subtle alterations of such complex interactions. Starting from whole-brain effective connectivity (EC) with causal inferences, we construct six complementary feature views by integrating the EC with 1) time-frequency statistics, 2) raw temporal activities, 3) static FC, 4) graph-theoretic topology, and 5) nodal entropy. We design Adaptive Reliability Fusion (ARF) to mitigate large inter-subject variations in noise levels/distributions across different views, a great challenge in conventional multiview learning. The ARF synthesizes sample-wise quality assessment networks with global learnable weights to achieve dynamic, weighted multiview feature fusion. A hybrid hypergraph then integrates these multiview features through learnable hyperedges. A multi-layer hypergraph convolution mechanism is employed to model high-order, inter-regional dependencies. Finally, dual-path attention pooling on nodes and hyperedges isolates critical hubs to provide interpretability. Our model significantly outperforms other similar methods in CN versus MCI classification, offering a novel solution for comprehensively investigating early pathological mechanisms of AD. The source code is available at `https://github.com/eyeopeners/HyperBrainNet`.

Keywords: Alzheimer's disease · Effective connectivity · Directional brain network · Hypergraph · Multi-view

1 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by continuous cognitive decline, typically evolving along a pathological continuum from cognitively normal (CN) through mild cognitive impairment (MCI) to dementia status [1]. As AD progresses, the disruption of brain functional networks becomes more pronounced. Notably, during the MCI stage, substantial alterations in brain networks can be observed, making it a critical window for early diagnosis and clinical intervention [2]. For this purpose, resting-state functional magnetic resonance imaging (rs-fMRI) is increasingly used as a non-invasive method to identify early functional dysregulations in AD [3]. Recently, deep learning methods have been established as promising individual diagnostic tools using rs-fMRI [4–6]. These methods rely on various types of Graph Neural Networks (GNNs) [7, 8] for graph learning on a simplified, often pairwise and undirected, brain network, despite accumulating evidence indicating that AD involves far more complex brain network changes. This necessitates more sophisticated graph representation paradigms [9].

The current accuracy of rs-fMRI-based early AD diagnosis has reached a bottleneck, which could be partially due to three critical limitations: 1) Lack of a modeling strategy for decoding directional relations among brain regions. The input of mainstream GNNs is overwhelmingly limited to undirected FC, whereas the actual inter-regional relationship is directional, involving both driving and driven regions [10, 11]. Although effective connectivity (EC) can model such directionality [13], its integration and combined learning with FC in deep learning remain largely unexplored. 2) Insufficient modeling of more complex, "high-order" interactions that go beyond pairwise connections: Traditional graphs strictly constrain interactions to node pairs, neglecting synergistic coordination among multiple nodes that coexist in large-scale functional sub-networks [14, 24]. Current hypergraph learning [15] depends on statistical random walks [16] or heuristically determined rigid thresholds, and there is no end-to-end learnable, neuroscientifically meaningful mechanism that can both construct and learn such a hypergraph. 3) Suboptimal multi-view fusion: the most recent sophisticated architectures directly aggregate temporal, spectral, and spatial features [17]. However, the signal and noise distributions across these domains and individual subjects inevitably vary substantially, making naive concatenation or static attention frequently suffer from feature redundancy and noise propagation [18].

To tackle these challenges, we propose *HyperBrainNet*, a directionality-aware hybrid hypergraph network for unified learning from directed, dynamic, high-order functional topologies in an end-to-end manner, fortified by adaptive cross-domain noise suppression. HyperBrainNet is equipped with three core innovations: 1) *Adaptive Reliability Fusion (ARF)*: computes sample-specific quality assessments combined with global priors to dynamically re-weight multi-view rs-fMRI representations. 2) *Hybrid Hypergraph Learning*: a dual-pathway hypergraph that seamlessly integrates deterministic anatomical priors (neuroscientifically meaningful Schaefer seven networks [19, 25]) with orthogonally regular-

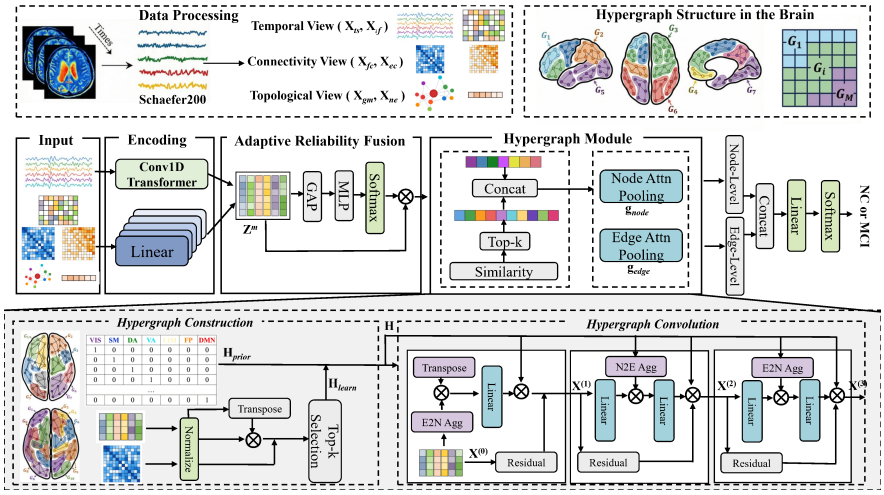

Fig. 1. The overall framework of our proposed HyperBrainNet.

ized learnable hyperedges, which can capture both established network modularity and latent pathological reorganization. 3) *Hierarchical Attention Pooling*: a node-hyperedge dual attention mechanism, which extracts highly discriminative and interpretable brain functional sub-network biomarkers for improving diagnostic transparency.

2 Methods

2.1 Overall Framework

Fig. 1 details the architecture of HyperBrainNet. Given preprocessed rs-fMRI time series $\mathbf{X}_{\text{ts}} \in \mathbb{R}^{T \times N}$ from N ROIs over T time points, the analytical pipeline is designed to progressively distill pathological biomarkers. The framework proceeds through three stages: multi-view encoding of six complementary representations, Adaptive Reliability Fusion (ARF) for subject-specific noise suppression, and directionality-aware hybrid hypergraph learning with dual-path attention pooling for final classification.

2.2 Causality-Enhanced Encoding and Adaptive Reliability Fusion

Effective Connectivity (EC) via Neural Perturbational Inference (NPI). To map directed connectivity disruptions associated with AD [9], we extract EC via the Neural Perturbational Inference (NPI) framework [13]. NPI treats the brain as a dynamical system where the state of a region depends on the historical states of its associated network.

Specifically, an MLP surrogate f_θ is trained on CN subjects as a healthy reference to predict the next rs-fMRI state by one-step mean-squared error. The

trained surrogate is then applied to all CN/MCI subjects. Once f_θ captures the underlying dynamics, we apply systematic *virtual perturbations* [12] to individual source ROIs. The directed influence (the EC measurement in our method) from ROI i to ROI j is quantified by observing the deviation in the predicted future state of ROI j in response to the perturbation at ROI i:

$$\mathrm{EC}_{ij} = \|\mathbf{x}_j^{\text{perturbed}} - \mathbf{x}_j^{\text{baseline}}\|_2 \tag{1}$$

The deviation is averaged over time and normalized to obtain directed, model-inferred EC that approximates the surrogate Jacobian. The *causal perturbation* in our title refers to this virtual perturbation of a learned surrogate, rather than interventionist causality. EC is encoded row-wise, so ROI i's outgoing profile and the asymmetry $\mathrm{EC}_{ij} \neq \mathrm{EC}_{ji}$ remain in the node features; the hypergraph operators themselves are not directed.

Complementary Feature Views. Five views complement EC: 1) raw time series encoded by a 1D-CNN and a two-layer Transformer over time [17]; 2) 16-D time-frequency statistics per ROI, including mean/std/skewness and Welch-PSD features; 3) static FC; 4) degree, betweenness, clustering, local efficiency, participation coefficient, and within-module z-score on the top-20% |FC| graph; and 5) nodal Shannon entropy $H_i = -\sum_j p_{ij} \log p_{ij}$ from sliding-window FC variability. Each view is mapped independently to a shared latent space before fusion.

Adaptive Reliability Fusion (ARF). Because the noise profiles of rs-fMRI (e.g., motion artifacts, scanner variations) differ significantly across subjects and distinct feature views, naive concatenation frequently propagates redundant or corrupted signals. To achieve robust feature integration, the ARF module is designed to dynamically evaluate the feature reliability of each view. For a given subject, sample-specific quality scores s^m are generated via lightweight MLPs. These dynamic scores are then combined with learnable global baseline weights w^m (acting as population-level priors) and normalized via softmax:

$$\alpha^m = \frac{\exp(s^m + w^m)}{\sum_{m'=1}^{6} \exp(s^{m'} + w^{m'})} \tag{2}$$

The final unified node representation is obtained through a weighted summation $\mathbf{Z} = \sum_m \alpha^m \mathbf{Z}^m$. This strategy not only suppresses noisy inputs on a per-sample basis but also provides interpretability regarding which views contribute most effectively to the diagnosis.

2.3 Hybrid Hypergraph Construction and Message Passing

To account for complex functional modules where multiple ROIs co-activate synergistically [14], we project the fused node features into a hybrid hypergraph space governed by an incidence matrix $\mathbf{H} = [\mathbf{H}_{\text{prior}}, \mathbf{H}_{\text{learn}}]$, seamlessly merging fixed neuroscientific knowledge with adaptive topological discovery. In total, the hypergraph comprises 17 hyperedges (seven prior plus ten learnable), defined as follows:

(1) Prior Hyperedges: Anatomical constraints are enforced via the well-established Schaefer seven-network parcellation scheme [19], yielding seven static, binary, non-overlapping hyperedges that explicitly group functionally relevant ROIs (e.g., Default Mode, Somatomotor). This anchors the learning process in biological reality.
(2) Learnable Hyperedges: While prior parcellations capture normal neurofunctional anatomy, AD induces complex topological shifts that may transcend standard sub-network boundaries. To capture this pattern, we introduce a data-driven mechanism alongside the hypothesis-derived one. We initialize K orthogonal query vectors $\{\mathbf{q}_e\}_{e=1}^{K}$. The membership of node i to learnable hyperedge e is determined via a temperature-scaled similarity attention mechanism:

$$H_{\text{learn}}(i,e) = \sigma\left(\tau^{-1} \cdot \frac{\mathbf{z}_i^\top \mathbf{q}_e}{\|\mathbf{z}_i\|\|\mathbf{q}_e\|}\right) \tag{3}$$

This incidence is soft during training; at inference, top-k node selection yields sparse binary incidence. We use $K = 10$ to bridge the granularity of the Schaefer 7- and 17-network parcellations while allowing data-adaptive reorganization.
Hypergraph Convolution Layers. Following the construction of the hybrid incidence matrix, feature refinement operates through an alternating message-passing protocol between nodes and hyperedges [15]. At layer ℓ, the feature of hyperedge e is first aggregated from its incident nodes and normalized by the hyperedge degree $\mathbf{D}_e = \sum_i H(i,e)$:

$$\mathbf{h}_e^{(\ell)} = \mathbf{D}_e^{-1} \sum_{i \in \mathcal{N}(e)} H(i,e) \cdot \phi_{\text{n2e}}(\mathbf{z}_i^{(\ell)}) \tag{4}$$

where ϕ_{n2e} represents a learnable transformation. Subsequently, the node features are symmetrically updated by aggregating information from all connected hyperedges. A residual connection is incorporated to preserve local node semantics and mitigate over-smoothing in deeper layers:

$$\mathbf{z}_i^{(\ell+1)} = \mathbf{D}_i^{-1} \sum_{e \in \mathcal{E}(i)} H(i,e) \cdot \phi_{\text{e2n}}(\mathbf{h}_e^{(\ell)}) + \mathbf{z}_i^{(\ell)} \tag{5}$$

2.4 Hierarchical Attention Pooling and Training Objectives

Dual-Path Attention Pooling. Not all brain regions or functional modules contribute equally to the pathology of AD. Therefore, instead of standard mean pooling, we implement an attention-based pooling mechanism to derive the graph-level embedding. We calculate attention scores for both nodes and hyperedges to dynamically emphasize diagnostically critical hubs:

$$\mathbf{g}_{\text{node}} = \sum_{i=1}^{N} \alpha_i \cdot \mathbf{z}_i^{(L)}, \quad \alpha_i = \frac{\exp(\text{MLP}_{\text{attn}}(\mathbf{z}_i^{(L)}))}{\sum_{j=1}^{N} \exp(\text{MLP}_{\text{attn}}(\mathbf{z}_j^{(L)}))} \tag{6}$$

Hyperedge-level pooling symmetrically yields a compact representation $\mathbf{g}_{\text{edge}}$. The concatenated vector $\mathbf{g} = [\mathbf{g}_{\text{node}}, \mathbf{g}_{\text{edge}}]$ forms the final embedding passed to the classifier.

Multi-Objective Training. The entire network is jointly optimized using a composite loss function to balance classification accuracy, topological diversity, and feature discriminability:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{CE}} + \lambda \mathcal{L}_{\text{div}} + \beta \mathcal{L}_{\text{SCL}} \tag{7}$$

Here, $\mathcal{L}_{\text{CE}}$ is a class-weighted focal loss [20] with label smoothing ($\epsilon = 0.1$). $\mathcal{L}_{\text{div}}$ acts as a critical diversity regularizer, penalizing non-orthogonality among the learnable query vectors in $\mathbf{Q}$ to maintain distinct topological modules ($\lambda = 0.1$). $\mathcal{L}_{\text{SCL}}$ [21] denotes an auxiliary supervised contrastive loss included to encourage tighter intra-class representation clustering ($\beta = 0.01$).

3 Experimental Results

Datasets and Evaluation metrics. The performance of the proposed HyperBrainNet framework is evaluated on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [3]. The cohort consists of 462 CN subjects and 242 MCI subjects. All BOLD fMRI data underwent standardized preprocessing and quality control via the QuNex pipeline [22]. To ensure the reliability of the findings and prevent data leakage, a five-fold cross-validation scheme is implemented using subject-level stratified splitting.

Implementation Details. The framework is implemented in PyTorch and trained on an NVIDIA RTX 4090 GPU. The optimization process employs the AdamW optimizer with an initial learning rate of 1×10^{-4} and a weight decay coefficient of 1×10^{-5}. The training utilizes a batch size of 16 and continues for up to 150 epochs. To enhance stability, a warmup strategy of 10 epochs is applied. HyperBrainNet contains approximately 0.51M parameters.

3.1 Comparison with State-of-the-Art Methods

Table 1 benchmarks HyperBrainNet against leading architectural baselines, including graph convolutional networks (GCN [4], GAT [5]), specialized brain architectures (BrainNetCNN [7], BrainGNN [8]), and contemporary transformer models (BrainNetTransformer [17], DHGFormer [23]). All baseline models were implemented using their official code and standard settings reported in the respective literature.

HyperBrainNet achieves superior performance across primary metrics, yielding an ACC of 74.4% and an AUC of 74.9%. The >11% ACC margin over standard pairwise graphs (GCN/GAT) substantiates that AD-induced network disruptions occur mostly in high-order networks. Standard GNNs are typically designed for static, undirected graphs, limiting their ability to incorporate directed dynamics (through EC) and high-order topological interactions. Consequently, they often underperform in scenarios requiring the integration of rich,

Table 1. Classification performance comparison (CN vs. MCI).

Method	CN vs. MCI				
	ACC(%)	AUC(%)	SEN(%)	SPE(%)	F1(%)
GCN [4]	60.9±8.0	59.9±6.3	59.8±5.7	66.9±6.1	61.1±5.8
GAT [5]	62.5±4.0	63.6±2.9	63.0±5.7	67.5±2.9	64.3±3.0
BrainNetCNN [7]	73.1±2.8	73.2±4.5	64.1±8.3	**86.2±6.2**	74.2±2.7
HGNN [15]	67.2±3.5	65.6±3.2	62.9±5.3	74.7±4.8	69.2±3.3
BrainGNN [8]	69.0±2.0	69.6±1.8	65.4±5.2	76.2±4.5	70.6±1.8
BrainNetTrans. [17]	67.8±1.3	65.7±0.8	62.1±1.5	76.0±2.1	70.7±1.2
DHGFormer [23]	71.0±2.0	73.4±3.5	71.8±2.3	70.1±2.3	71.5±2.2
Ours	**74.4±3.5**	**74.9±4.9**	**74.9±3.2**	78.8±5.9	**74.8±3.1**

multi-view information. In contrast, our framework maintains a rigorously balanced diagnostic profile (SEN 74.9%, SPE 78.8%), which is clinically paramount for MCI screening.

3.2 Ablation Studies

Table 2 dissects the isolated contributions of our multi-view semantics and structural modules.

Multiple Feature View and Their Importance. The removal of the connectivity view (containing EC and FC) precipitates a severe 5.0% ACC decline. This unequivocally demonstrates that temporal and topological markers alone are insufficient; the directed influence embedded within the EC encapsulates unique pathological trajectories. Similarly, excluding the temporal feature view heavily penalizes performance (−6.2% ACC), reinforcing the necessity of capturing low-frequency oscillatory dynamics alongside the brain graphs.

Architectural Module Impact. Omitting the hybrid hyperedge module (reverting to either solely pairwise graphs or prior-only modeling) drops ACC to 70.3%. This validates our core hypothesis: pure neuroscientific priors lack individual adaptability, while pure data-driven edges risk biological implausibility; their hybrid combination optimally models AD-induced compensatory subnetworks. Furthermore, substituting the ARF module with naive concatenation degrades accuracy to 69.7%, confirming the imperative nature of dynamically suppressing modality-specific inter-subject noise. Finally, reverting from attention pooling to standard mean pooling leads to an F1 score drop to 60.6%, proving that explicitly weighting critical functional hubs is essential for robust classification.

3.3 Interpretability Analysis

We visualized the top-10 discriminative ROIs extracted from the spatial attention weights generated by the attention pooling module (Fig. 2). These ROIs

Table 2. Ablation studies. Top: feature-view ablation. Bottom: architectural-module ablation (CN vs. MCI).

Feature-View Ablation							
Temporal	Conn.	Topology	ACC(%)	AUC(%)	SEN(%)	SPE(%)	F1(%)
	✓	✓	68.2±3.5	67.5±2.4	56.2±3.8	74.5±3.4	54.1±5.5
✓		✓	69.4±2.7	72.8±3.8	64.0±8.5	74.9±6.2	60.2±3.9
✓	✓		71.4±3.8	74.1±5.2	69.8±5.6	73.0±6.0	64.3±3.8
✓	✓	✓	**74.4±3.5**	**74.9±4.9**	**74.9±3.2**	**78.8±5.9**	**74.8±3.1**
Architectural-Module Ablation							
Hyperedge	ARF	Pooling	ACC(%)	AUC(%)	SEN(%)	SPE(%)	F1(%)
	✓	✓	70.3±2.5	73.2±2.7	69.0±7.8	71.6±5.1	61.7±3.4
✓		✓	69.7±2.8	74.1±4.1	69.0±4.5	70.3±4.8	61.2±3.5
✓	✓		69.4±5.7	72.1±6.3	67.3±5.7	71.5±4.9	60.6±5.2
✓	✓	✓	**74.4±3.5**	**74.9±4.9**	**74.9±3.2**	**78.8±5.9**	**74.8±3.1**

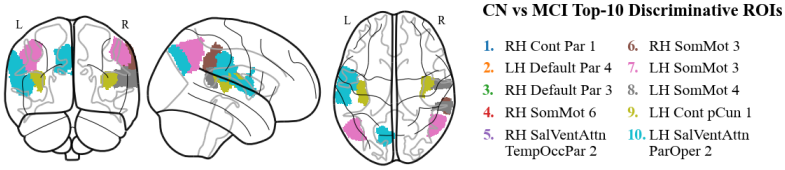

Fig. 2. Spatial mapping of the top-10 discriminative ROIs for CN vs. MCI classification.

exhibit striking anatomical concordance with established neurodegenerative etiology.

Notably, three distinct hubs within the default mode network emerged prominently. Given that this network's hypometabolism and early amyloid-beta accumulation are hallmarks of prodromal AD, their high ranking validates the biological soundness of the feature extraction pipeline. Interestingly, four regions map to the somatomotor network. While AD is conventionally viewed as an exclusively cognitive disorder, emerging longitudinal literature underscores that subtle sensorimotor dysfunction substantially precedes overt memory loss, rendering this network a highly sensitive early biomarker. The involvement of the salience/ventral attention network further indicates that executive control and attention-shifting mechanisms are fundamentally compromised in MCI.

4 Conclusions

We presented HyperBrainNet, a directionality-aware hybrid hypergraph neural network engineered for early AD detection. By synergizing perturbation-derived EC with a combined theory- and data-driven feature-fusion mechanism, we rigorously address the limitations of undirected functional-connectivity-based mod-

els and cross-domain noise-distribution gaps. Integrating structural brain priors with orthogonally regularized learnable hyperedges establishes a highly robust method for tracking high-order network reorganization, delivering both accurate diagnosis and transparent clinical interpretability. Future work will explore explicitly directed, asymmetry-preserving hypergraph message passing.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Nos. 62571330, 62131015, and U23A20295) and the Shanghai Pilot Program for Basic Research – Chinese Academy of Sciences, Shanghai Branch (No. JCYJ-SHFY-2022-014). The computation in this work is also supported by the HPC Platform of ShanghaiTech University.

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

References

1. Jack, C.R., et al.: NIA-AA research framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia **14**(4), 535–562 (2018)
2. Sperling, R.A., et al.: Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines. Alzheimer's & Dementia **7**(3), 280–292 (2011)
3. Weiner, M.W., et al.: The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. Alzheimer's & Dementia **13**(5), 561–571 (2017)
4. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: ICLR (2017)
5. Veličković, P., et al.: Graph attention networks. In: ICLR (2018)
6. Cui, H., et al.: BrainGB: A benchmark for brain network analysis with graph neural networks. IEEE Trans. Med. Imaging **42**(2), 493–506 (2023)
7. Kawahara, J., et al.: BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage **146**, 1038–1049 (2017)
8. Li, X., et al.: BrainGNN: Interpretable brain graph neural network for fMRI analysis. Med. Image Anal. **74**, 102233 (2021)
9. Biswas, R., Sripada, S.N.: Causal functional connectivity in Alzheimer's disease computed from time series fMRI data. Front. Comput. Neurosci. **17**, 1251301 (2023)
10. Schippers, M.B., et al.: Mapping the information flow from one brain to another during gestural communication. Proc. Natl. Acad. Sci. USA **107**(21), 9388–9393 (2010)
11. Manjunatha, K.K.H., et al.: Controlling target brain regions by optimal selection of input nodes. PLoS Comput. Biol. **20**, e1011274 (2024)
12. Ivanovs, M., Kadikis, R., Ozols, K.: Perturbation-based methods for explaining deep neural networks: A survey. Pattern Recognit. Lett. **150**, 228–234 (2021)
13. Luo, Z., et al.: Mapping effective connectivity by virtually perturbing a surrogate brain. Nature Methods (2025)
14. Gao, Y., et al.: Hypergraph learning: Methods and practices. IEEE Trans. Pattern Anal. Mach. Intell. **44**(5), 2548–2566 (2022)

15. Feng, Y., et al.: Hypergraph neural networks. In: AAAI, vol. 33, pp. 3558–3565 (2019)
16. Guo, J., Wang, M., Guo, X.: Hierarchical hyperedge graph transformer: Toward dynamic interactions of brain networks for neurodevelopmental disease diagnosis. IEEE Access **12**, 162145–162156 (2024)
17. Kan, X., et al.: Brain network transformer. In: NeurIPS, vol. 35, pp. 25586–25598 (2022)
18. Zhang, Y., et al.: Transformer-based multimodal fusion for early diagnosis of Alzheimer's disease using structural MRI and PET. In: ISBI, pp. 1–5 (2023)
19. Schaefer, A., et al.: Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb. Cortex **28**(9), 3095–3114 (2018)
20. Lin, T.Y., et al.: Focal loss for dense object detection. In: ICCV, pp. 2980–2988 (2017)
21. Khosla, P., et al.: Supervised contrastive learning. In: NeurIPS, vol. 33, pp. 18661–18673 (2020)
22. Ji, J.L., et al.: QuNex—An integrative platform for reproducible neuroimaging analytics. Front. Neuroinform. **17**, 1104508 (2023)
23. Xue, R., et al.: DHGFormer: Dynamic hierarchical graph transformer for disorder brain disease diagnosis. In: MICCAI 2025, LNCS, vol. 15971, pp. 268–278. Springer, Cham (2025)
24. Friston, K.: Hierarchical models in the brain. PLoS Comput. Biol. **4**(11), e1000211 (2008)
25. Akiki, T.J., Abdallah, C.G.: Determining the hierarchical architecture of the human brain using subject-level clustering of functional networks. Sci. Rep. **9**, 1–15 (2019)