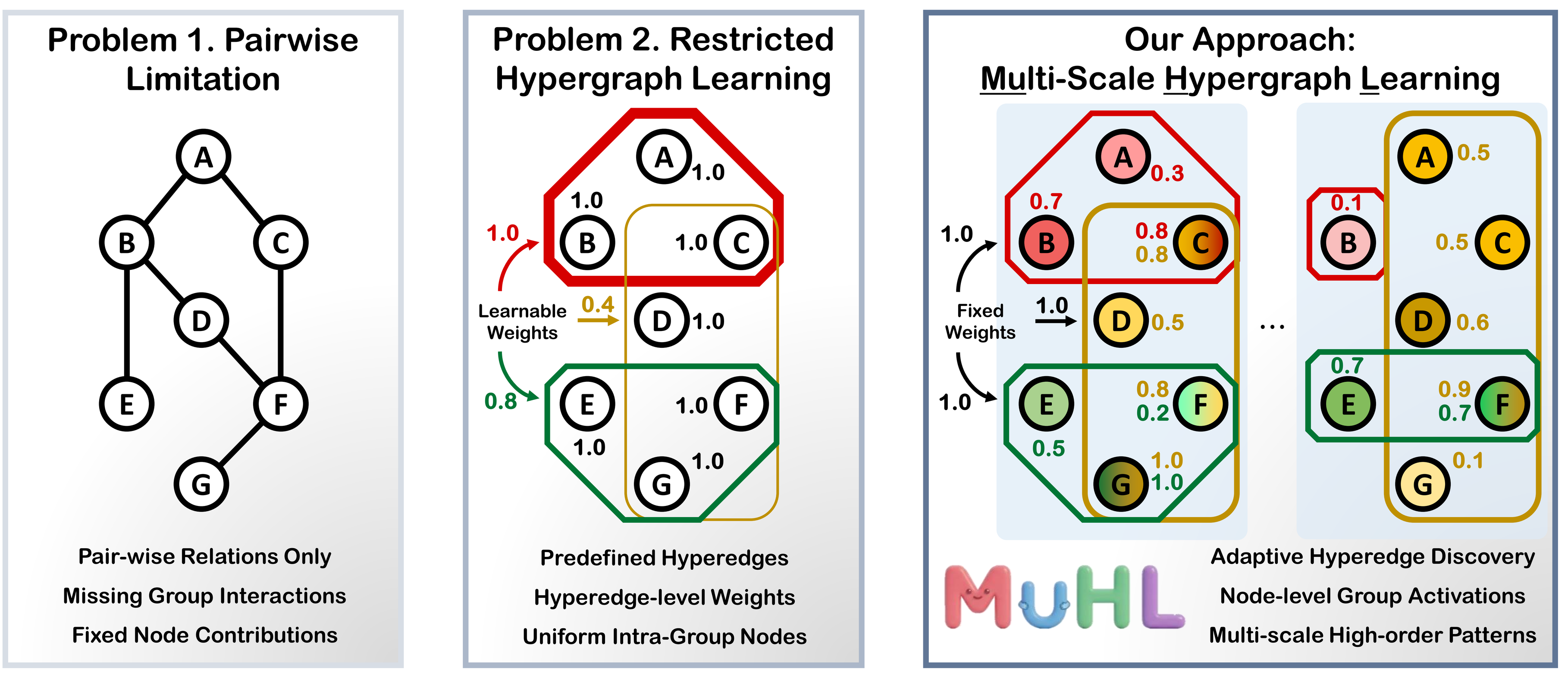
Learning Multi-Scale Hypergraph for High-Order Brain Connectivity Analysis

Jaeyoon Sim[1] Soojin Hwang[1] Seunghun Baek[1] Guorong Wu[2] Won Hwa Kim[1]

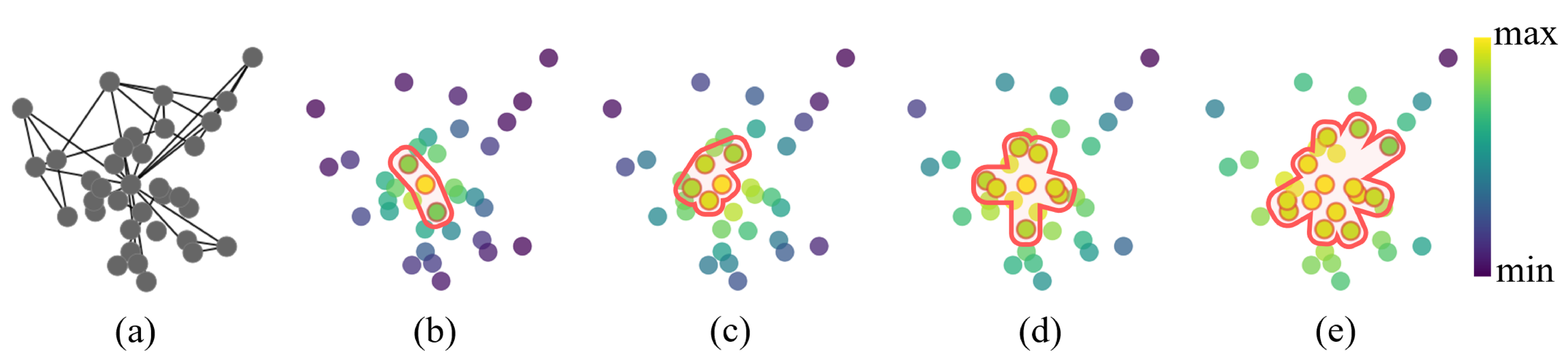
[1]POSTECH, [2]UNC Chapel Hill

Introduction



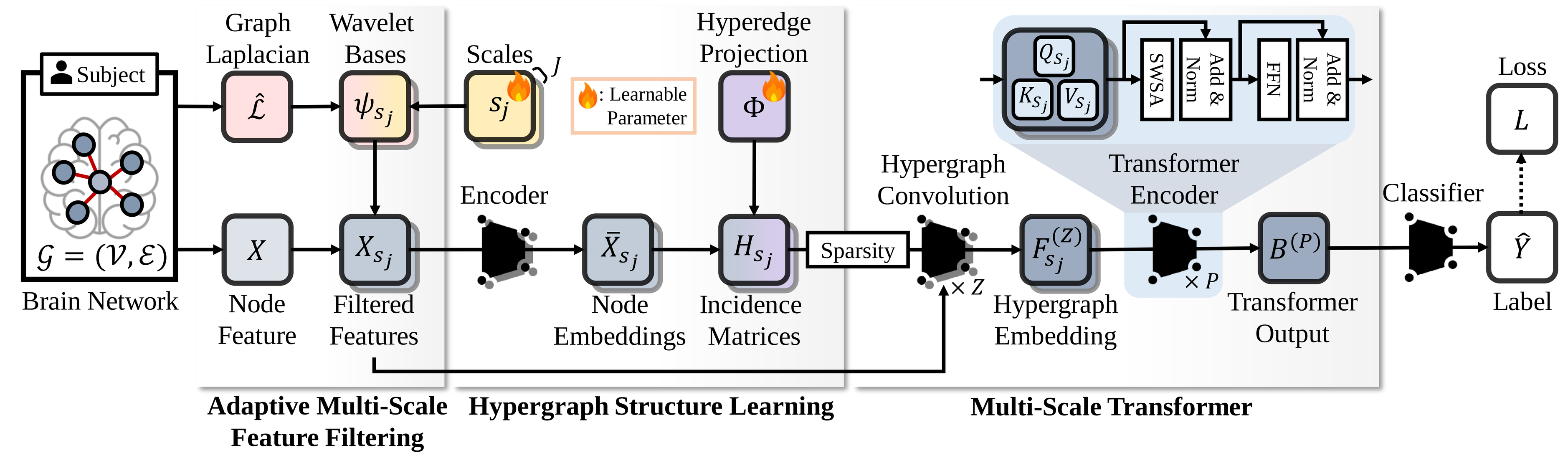
- **Problem:**
  1. Pairwise Limitation: Existing graph-based brain network models mainly capture pairwise ROI connections, limiting their ability to model group-wise interactions among multiple brain regions.
  2. Restricted Hypergraph Learning: Existing hypergraph-based methods often rely on predefined hyperedges or only adjust hyperedge-level weights, without fully learning node-level activation patterns within each hyperedge.
- **Contribution:**
  1. Dynamic Hyperedge Structure Learning: *MuHL* dynamically learns rich and implicit group-wise node dependencies (i.e., hyperedges).
  2. Adaptive Multi-Resolution Brain Representations: *MuHL* adaptively captures structural variations at different scales for hypergraph construction to enhance downstream classification for brain network analysis.
  3. Theoretical Justification: *MuHL* demonstrates theoretical insights that the adaptive hyperedges are derived by modulating node features.

Multi-resolution High-order Relation Modeling



- The original graph with pairwise edges is shown in (a), and from fine scale (b) to coarse scale (e), hyperedges adaptively expand to include more nodes, moving from localized connections to broader groupings.
- Similar colors indicate potentially similar regions, while the red contour indicates the sparsely constructed hyperedge at each scale, determined by corresponding smoothed features.

Multi-Scale Hypergraph Learning (MuHL)



(1) Adaptive Multi-Scale Feature Filtering

- Extract scale-specific ROI features using learnable graph wavelet scales.

$$X_{s_j} = Ug^2(s_j\Lambda)U^\top X, \quad j = 1, \ldots, J$$

- where X is the ROI feature matrix, U and Λ are the eigenvectors and eigenvalues of the graph Laplacian, $g(\cdot)$ is a spectral wavelet kernel, and s_j is a learnable scale parameter.

(2) Hypergraph Structure Learning

- Construct adaptive hyperedges by projecting node embeddings into hyperedge space.

$$\widetilde{H}_{s_j} = \mathrm{Softmax}(\mathrm{ReLU}(\bar{X}_{s_j}\Phi)), \quad \bar{H}_{s_j} = \mathrm{TopK}_\eta(\widetilde{H}_{s_j})$$

- where $\bar{X}_{s_j}$ denotes the embedded node representation, Φ is a learnable hyperedge projector, $\widetilde{H}_{s_j}$ is a soft incidence matrix, and $\bar{H}_{s_j}$ is its sparse version with at most η hyperedges per node.

(3) Multi-Scale Transformer

- Aggregate scale-wise hypergraph features for graph-level disease classification.

$$F_{s_j} = \mathrm{HGConv}(X_{s_j}, \bar{H}_{s_j}), \qquad B = \mathrm{SWSA}\left(\{F_{s_j}\}_{j=1}^{J}\right)$$

- where F_{s_j} is the scale-wise hypergraph embedding from hypergraph convolution, and B is the final multi-scale representation produced by scale-wise self-attention for graph classification.

Experiments Result

(1) Brain Network Classification Performance

Table: Brain network classification performance on the ADNI and PPMI datasets evaluated using 5-fold cross-validation. The best results are shown in **bold**, and the second-best results are underlined. (†: Methods for brain network analysis)

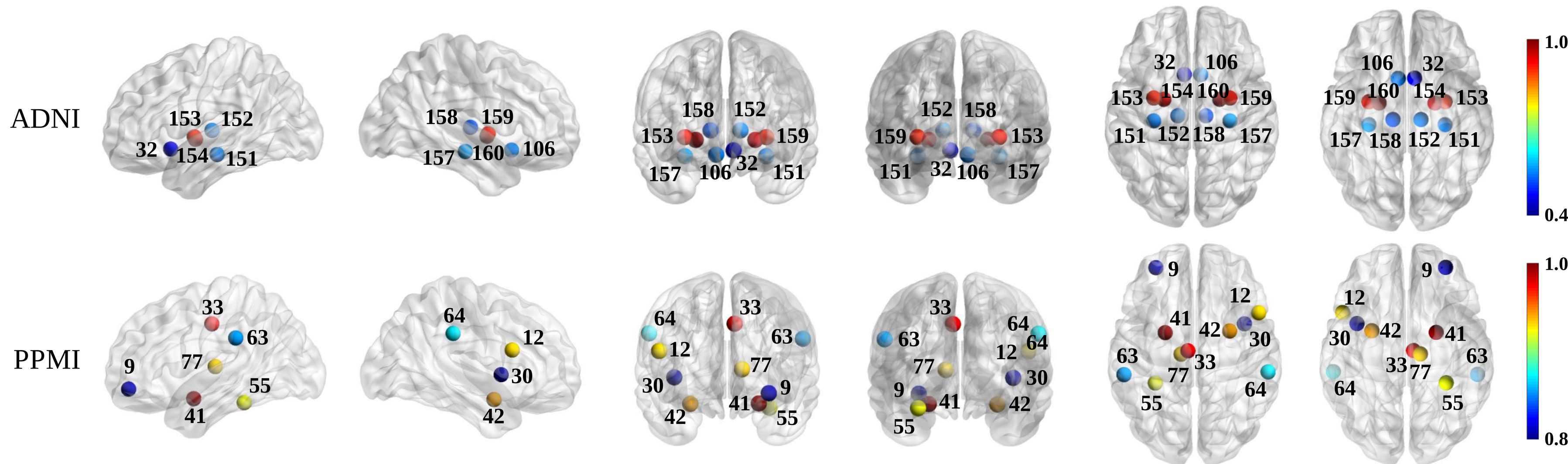
Category	Method	ADNI (CN vs SMC vs EMCI vs LMCI vs AD)				PPMI (CN vs Prodromal vs PD)			
		Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
Graph based method	GCN	85.7 ± 1.7	88.1 ± 3.0	82.7 ± 5.9	83.7 ± 3.4	66.3 ± 2.4	46.2 ± 5.4	41.5 ± 3.7	38.9 ± 4.4
	GAT	89.2 ± 2.1	92.7 ± 1.8	86.6 ± 4.6	88.9 ± 3.6	<u>72.9 ± 2.5</u>	55.1 ± 7.8	52.0 ± 3.3	52.3 ± 4.3
	GCNII	86.3 ± 4.0	91.2 ± 1.5	85.8 ± 6.6	87.4 ± 4.8	68.5 ± 2.3	52.5 ± 6.0	42.8 ± 4.5	41.4 ± 6.3
	BrainGNN†	43.4 ± 4.7	31.8 ± 3.4	28.3 ± 1.7	27.0 ± 1.3	68.1 ± 2.0	48.8 ± 5.3	40.8 ± 4.4	38.2 ± 5.3
	IBGNN†	81.4 ± 2.4	85.8 ± 0.7	82.3 ± 3.4	83.3 ± 2.3	68.1 ± 2.4	51.1 ± 6.0	42.1 ± 5.2	38.7 ± 5.4
	BrainGB†	83.5 ± 1.6	87.1 ± 2.6	80.7 ± 3.2	82.2 ± 2.3	68.8 ± 3.3	48.8 ± 6.0	42.8 ± 3.2	41.6 ± 4.5
	BrainNetTF†	89.4 ± 2.5	88.9 ± 5.7	84.6 ± 7.3	85.8 ± 6.0	67.9 ± 2.9	55.6 ± 8.8	45.7 ± 5.1	45.0 ± 7.2
	SGCN†	84.0 ± 2.1	87.6 ± 3.3	82.1 ± 4.6	83.8 ± 3.5	67.6 ± 3.2	39.0 ± 9.3	40.6 ± 4.4	37.6 ± 6.8
	ALTER†	<u>90.8 ± 2.6</u>	<u>93.3 ± 2.4</u>	<u>89.6 ± 5.6</u>	<u>90.9 ± 4.6</u>	69.6 ± 2.3	55.2 ± 8.5	<u>55.8 ± 5.3</u>	54.5 ± 8.3
	BioBGT†	80.3 ± 1.7	79.7 ± 7.1	77.3 ± 6.7	77.2 ± 6.1	68.0 ± 2.8	54.3 ± 8.2	48.7 ± 4.2	45.7 ± 3.2
	BQN†	81.5 ± 4.7	74.3 ± 5.6	72.3 ± 5.5	72.6 ± 5.2	71.9 ± 2.9	<u>63.2 ± 6.2</u>	51.0 ± 8.4	<u>56.4 ± 7.3</u>
Hypergraph based method	HGNN	83.2 ± 2.7	84.8 ± 7.1	80.7 ± 6.8	82.1 ± 6.6	65.2 ± 4.8	43.3 ± 9.0	41.9 ± 6.1	40.3 ± 7.1
	HNHN	87.3 ± 1.8	89.6 ± 1.0	85.6 ± 4.2	87.0 ± 3.0	69.1 ± 2.9	52.6 ± 5.1	48.5 ± 8.4	48.0 ± 9.4
	UniGCNII	89.6 ± 1.6	91.1 ± 1.8	88.0 ± 3.8	89.9 ± 2.0	70.8 ± 3.3	59.6 ± 9.4	54.3 ± 7.2	55.2 ± 7.3
	HyperDrop	72.3 ± 0.8	74.4 ± 3.1	65.9 ± 0.2	68.6 ± 2.2	67.5 ± 4.0	41.5 ± 8.9	39.2 ± 3.8	36.0 ± 6.4
	dwHGCN†	90.2 ± 1.3	87.2 ± 7.1	86.3 ± 6.8	86.2 ± 6.4	69.0 ± 2.9	51.5 ± 7.9	44.2 ± 6.1	43.3 ± 8.0
	HyperGT	81.5 ± 1.6	92.1 ± 1.8	89.0 ± 3.8	89.9 ± 2.0	68.5 ± 1.5	51.7 ± 4.4	42.0 ± 4.8	39.6 ± 6.1
	HyBRiD†	86.6 ± 4.2	87.5 ± 3.9	84.3 ± 6.4	84.6 ± 4.7	67.2 ± 6.5	54.8 ± 6.8	52.0 ± 6.2	52.8 ± 6.5
	DHHNN	81.1 ± 2.5	89.0 ± 4.5	84.5 ± 3.2	86.7 ± 3.4	60.6 ± 3.8	62.3 ± 9.8	52.7 ± 4.3	54.0 ± 6.3
	MuHL (Ours)	**93.2 ± 2.4**	**95.4 ± 1.3**	**94.2 ± 1.6**	**94.7 ± 1.5**	**76.8 ± 3.7**	**66.6 ± 7.6**	**60.7 ± 6.0**	**62.4 ± 6.6**

(2) Zero-Shot Classification Performance

Table: Zero-shot classification performance across datasets and acquisition phases. MuHL and baselines are trained on a source dataset or phase and directly evaluated on unseen target datasets or phases without fine-tuning.

Method	ADNI-2→ADNI-1/3/GO				PPMI→TaoWu				PPMI→Neurocon			
	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑	Acc ↑	Pre ↑	Rec ↑	F1s ↑
BrainGB	42.9 ± 3.2	48.4 ± 8.1	36.5 ± 2.4	34.4 ± 4.3	53.0 ± 1.2	64.1 ± 6.8	53.0 ± 1.2	43.5 ± 4.0	62.9 ± 6.4	52.5 ± 2.8	51.6 ± 7.3	45.1 ± 5.3
BrainNetTF	43.5 ± 8.4	<u>48.9 ± 2.4</u>	50.8 ± 2.8	44.8 ± 1.6	50.0 ± 1.6	46.0 ± 3.9	50.0 ± 1.6	46.2 ± 4.6	62.9 ± 4.4	62.7 ± 8.1	55.0 ± 5.6	50.3 ± 8.7
ALTER	43.8 ± 3.0	46.8 ± 2.0	54.6 ± 6.4	44.3 ± 2.3	54.5 ± 1.9	<u>65.1 ± 6.0</u>	52.5 ± 1.9	45.3 ± 4.4	<u>63.8 ± 3.2</u>	<u>63.8 ± 7.6</u>	<u>55.3 ± 4.4</u>	<u>51.7 ± 8.0</u>
BQN	<u>52.3 ± 2.9</u>	39.5 ± 6.3	32.8 ± 7.7	39.0 ± 7.6	<u>57.0 ± 4.0</u>	56.3 ± 4.6	<u>55.0 ± 4.0</u>	<u>49.5 ± 3.0</u>	55.1 ± 3.7	47.8 ± 3.2	48.3 ± 2.9	46.8 ± 2.9
dwHGCN	52.8 ± 2.9	48.0 ± 2.7	<u>57.0 ± 2.8</u>	<u>48.8 ± 3.1</u>	50.5 ± 2.9	45.1 ± 5.2	50.5 ± 2.9	36.0 ± 4.5	60.9 ± 2.7	59.3 ± 3.4	53.6 ± 3.7	46.1 ± 7.0
HyBRiD	47.6 ± 5.1	46.7 ± 4.6	53.3 ± 2.9	44.0 ± 5.2	56.2 ± 3.1	53.8 ± 3.5	49.8 ± 4.7	44.7 ± 3.8	63.4 ± 2.4	51.7 ± 2.5	50.0 ± 4.2	48.8 ± 3.6
MuHL	**56.0 ± 1.6**	**52.1 ± 6.1**	**58.4 ± 8.3**	**54.8 ± 7.7**	**60.5 ± 1.6**	**72.8 ± 2.4**	**60.5 ± 1.6**	**58.9 ± 3.0**	**65.9 ± 3.4**	**68.1 ± 3.3**	**57.3 ± 3.2**	**53.7 ± 3.6**

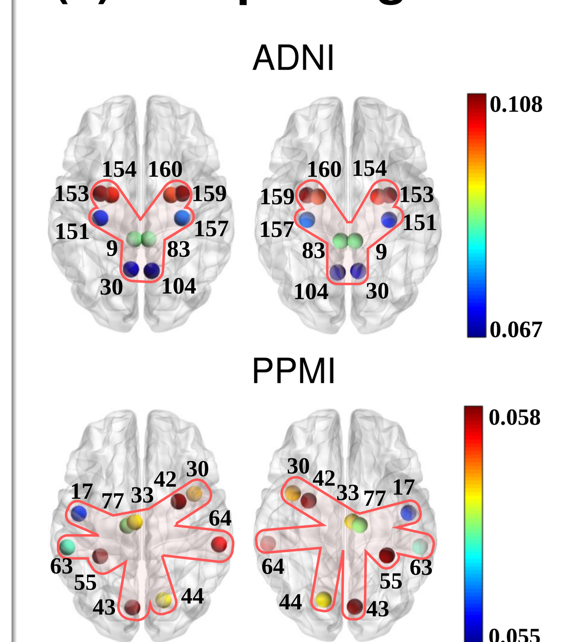
(3) Identifying Key ROIs via Hyperedge Activations



ADNI						PPMI					
Idx	ROI	Act.	Idx	ROI	Act.	Idx	ROI	Act.	Idx	ROI	Act.
154	(L) Globus.Pallidus	1.000	160	(R) Globus.Pallidus	0.963	41	(L) Amygdala	1.000	33	(L) G.Cingulum.Mid	0.974
159	(R) Putamen	0.944	153	(L) Putamen	0.941	42	(R) Amygdala	0.948	77	(L) Thalamus	0.939
157	(R) Hippocampus	0.551	151	(L) Hippocampus	0.532	12	(R) G.Frontal.Inf.Oper	0.939	55	(L) G.Fusiform	0.929
106	(R) G.Subcallosal.Area	0.529	152	(L) Thalamus	0.528	64	(R) G.Supramarginal	0.888	63	(L) G.Supramarginal	0.870
158	(R) Thalamus	0.499	32	(L) G.Subcallosal.Area	0.432	9	(L) G.Frontal.Mid.Orb	0.829	30	(R) Insula	0.829

Figure: Top: Visualization of top-10 ROIs with the highest relative hyperedge importance on the ADNI (top) and PPMI (bottom) datasets. (a)/(b): outer views of left/right hemisphere, (c)/(d): front/rear views, (e)/(f): top/bottom views. Node color reflects the relative contribution of each ROI based on aggregated hyperedge weights.
Bottom: ROIs with 10 highest relative importance of aggregated hyperedge activation (Act.) on for classification. (L) and (R) denote the left and right hemispheres, respectively.

(4) Interpreting Most Salient Hyperedge



Idx	ROI
9	(L) G.Cingul.Post.Dorsal
83	(R) G.Cingul.Post.Dorsal
30	(L) G.Precuneus
104	(R) G.Precuneus
151	(L) Hippocampus
157	(R) Hippocampus
153	(L) Putamen
159	(R) Putamen
154	(L) Pallidum
160	(R) Pallidum

Idx	ROI
17	(L) Rolandic.Oper
77	(L) Thalamus
33	(L) Cingulum.Mid
30	(R) Insula
43	(L) Calcarine
42	(R) Amygdala
55	(L) Fusiform
44	(R) Calcarine
63	(L) SupraMarginal
64	(R) SupraMarginal

Figure: Left: Visualization of the top-10 ROIs associated with the most important hyperedge for top/bottom views, based on the sum of node activations for each hyperedge. Node color indicates activation value (i.e., weight) per node within the hyperedge. Right: Corresponding ROI labels.

(5) Hyper-parameter Studies

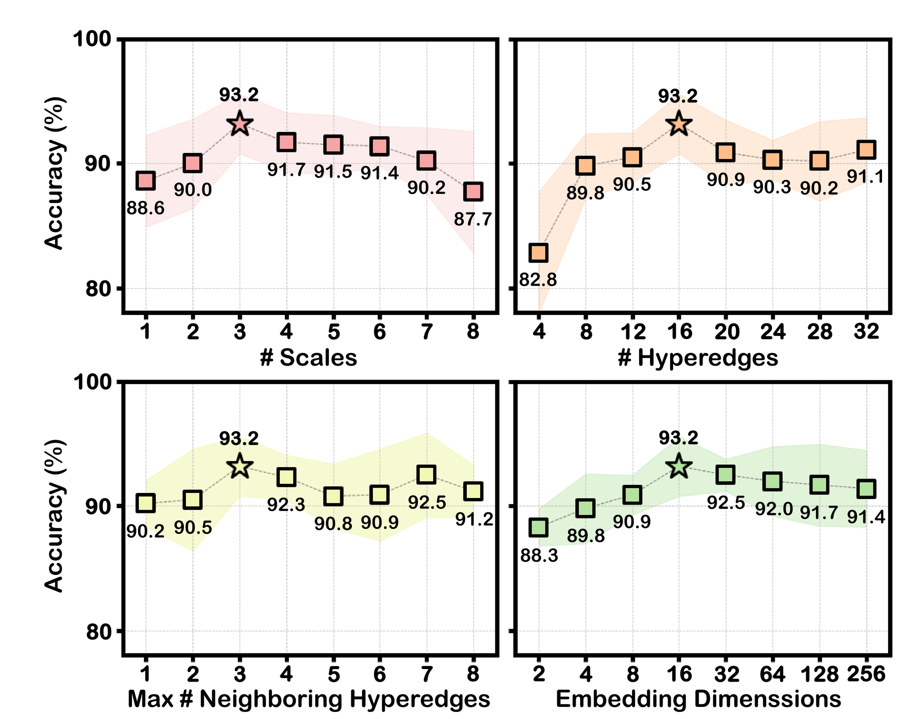


Figure: Sensitivity studies of crucial hyperparameters in MuHL on the ADNI dataset. Boxes denote the mean, with shading indicating standard deviation across folds.