Multi-order Simplex-based Graph Neural Network for Brain Network Analysis

Yechan Hwang[1], Soojin Hwang[1], Guorong Wu[2], and Won Hwa Kim[1]

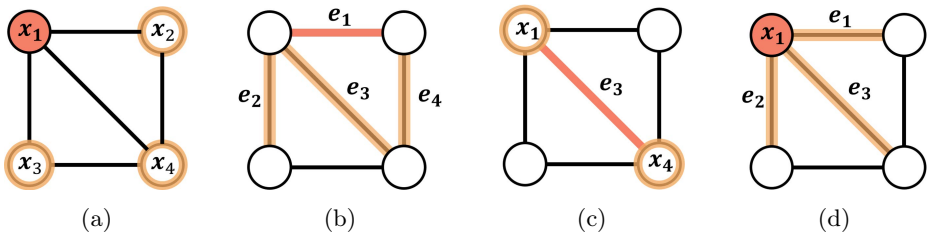
[1] Pohang University of Science and Technology, Pohang, South Korea
{yechan99, soojin0622, wonhwa}@postech.ac.kr
[2] University of North Carolina at Chapel Hill, Chapel Hill, USA

Abstract. A brain network is defined by wiring anatomical regions in the brain with structural and functional relationships. It has an intricate topology with handful early features/biomarkers of neurodegenerative diseases, which emphasize the importance of analyzing connectomic features alongside region-wise assessments. Various graph neural network (GNN) approaches have been developed for brain network analysis, however, they mainly focused on node-centric analyses often treating edge features as an auxiliary information (i.e., adjacency matrix) to enhance node representations. In response, we propose a method that explicitly learns node and edge embeddings for brain network analysis. Introducing a dual aggregation framework, our model incorporates a novel spatial graph convolution layer with an incidence matrix. Enabling concurrent node-wise and edge-wise information aggregation for both nodes and edges, this framework captures the intricate node-edge relationships within the brain. Demonstrating superior performance on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, our model effectively handles the complex topology of brain networks. Furthermore, our model yields interpretable results with Grad-CAM, selectively identifying brain Regions of Interest (ROIs) and connectivities associated with AD, aligning with prior AD literature.

Keywords: Incidence Matrix · Brain Network · Graph Classification.

1 Introduction

Brain Network, a wiring system between different anatomical brain regions of interest (ROIs), is typically viewed via the lens of a graph which comprises nodes and edges [15, 18]. Various approaches utilize Graph Neural Networks (GNNs) [7, 22] which primarily focus on node-centric analyses, emphasizing the importance of understanding individual brain regions through the graph structure [13,21,25]. These methods face critical limitations when analyzing the topology and connectomic features of brain networks, as the graph convolution in GNNs utilizes the graph structure merely as the domain of node-specific measures to derive enhanced node representations [5, 8]. In machine learning, efforts have been spent to analyze the edges directly, but they typically do not consider



(a) (b) (c) (d)

Fig. 1: Comparison between entities passing and receiving messages in conventional spatial convolution (a) and (b), and our proposed incidence matrix-based convolution (c) and (d). The aggregating subject is represented in red, while neighbors being aggregated are depicted in orange. (a): Node feature aggregation for node, (b): Edge feature aggregation for edge, (c) Node feature aggregation for edge, (d): Edge feature aggregation for node.

the analysis of the nodes [9, 10]. Therefore, the limitations above necessitate developing a method for delving into both nodes and edges individually to derive effective biomarkers from the brain network.

In response to these challenges and limitations, we propose a model that captures the intricate connectivity features between ROIs as well as the region-wise characteristics within the brain network from a geometric point of view with simplex. Our model utilizes topological information from both 0-simplices (nodes) and 1-simplices (edges) for graph convolution with Hodge Laplacian, obtaining explicit representations for each. Furthermore, they are mixed to obtain a richer unified representation representing both 0- and 1-simplices. For this, we introduce a dual aggregation framework, which is a novel graph convolution layer that utilizes a boundary matrix, i.e., incidence matrix, for convolutions of 0-simplices and 1-simplices. It allows concurrent aggregation of node-wise and edge-wise information, capturing the inter-simplex relationship, which is distinct from conventional GNN models. Fig. 1 conceptually outlines the description of individual aggregations from which nodes and edges are utilized for the convolution.

Our contributions are summarized as follows: **1)** Novel graph convolution with incidence matrix: We introduce a spatial graph convolution layer, leveraging the incidence matrix to model the intricate inter-simplex relationships within the brain network, **2)** Separate transformation of node and edge representations: Unlike conventional models where edges play an auxiliary role, our model directly trains on both node and edge representations, **3)** Superior performance and interpretability: Demonstrating superior performance on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, our model showcases interpretability in classification results by selectively identifying brain Regions of Interest (ROIs) and connectivities associated with Alzheimer's disease using Grad-CAM [20].

2 Related Work

There are models that focus solely on edge features to obtain edge representations. Authors in [10] utilized hypergraph transformation to acquire edge representations, demonstrating its validity on data such as molecules and social datasets, where edge information is crucial. [14, 17] showed efficacy in classification of neurological disorders by learning connectomic features between ROIs without utilizing node-wise features on structural brain datasets. On the other hand, several methods suggested jointly learning node and edge embeddings [5,8,24]. These models iteratively learn node and edge embeddings, obtaining explicit edge representations, which are then used as auxiliary information to enhance the final node representations. Besides, [9] employed a transformer to perform edge representation learning for text information on edges.

3 Preliminary

A simplicial complex, a fundamental concept in topological data analysis (TDA), refers to a collection composed of simplices. Starting from 0-dimension, a simplex extends to higher dimensions, representing basic forms such as a 0-simplex for a point (node), a 1-simplex for a line (edge), a 2-simplex for a triangle, etc. This concept enables a concise representation of complex data, particularly aiding in understanding topological properties of graphs or networks. Typically, a graph can be considered as a simplicial complex made up of 0-simplices (nodes) and 1-simplices (edges) [1, 23].

A boundary matrix $\mathcal{B}_p$ represents the relationship between p-simplices and $(p-1)$-simplices in a simplicial complex. Moreover, a Hodge Laplacian can be defined with $\mathcal{B}_p$, i.e., $\mathcal{L}_p = \mathcal{B}_p^T\mathcal{B}_p + \mathcal{B}_{p+1}\mathcal{B}_{p+1}^T$, which generalizes the conventional graph Laplacian from 0-simplices to p-simplices [19]. A boundary matrix $\mathcal{B}_1 \in \mathbb{R}^{N\times M}$, where N is the number of nodes and M is the number of edges, signifies the association between 0-simplices and 1-simplices. Commonly referred to as an incidence matrix, it provides details on how each edge connects two nodes, as:

$$(\mathcal{B}_1)_{ij} = \begin{cases} -1 & \text{if edge } e_j \text{ leaves vertex } v_i, \\ 1 & \text{if edge } e_j \text{ enters vertex } v_i, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Using this incidence matrix, the Hodge 0-Laplacian $\mathcal{L}_0$, equivalent to graph Laplacian, and Hodge 1-Laplacian $\mathcal{L}_1$ are defined as:

$$\mathcal{L}_0 = \mathcal{B}_1\mathcal{B}_1^T, \tag{2}$$

$$\mathcal{L}_1 = \mathcal{B}_1^T\mathcal{B}_1 + \mathcal{B}_2\mathcal{B}_2^T = \mathcal{B}_1^T\mathcal{B}_1. \tag{3}$$

Here, since a graph is composed of 0-simplices and 1-simplices only, $\mathcal{B}_2 = 0$ resulting in $\mathcal{L}_1 = \mathcal{B}_1^T\mathcal{B}_1$. In this way, we can derive $\mathcal{L}_0 \in \mathbb{R}^{N\times N}$ and $\mathcal{L}_1 \in \mathbb{R}^{M\times M}$ directly from the incidence matrix $\mathcal{B}_1$.

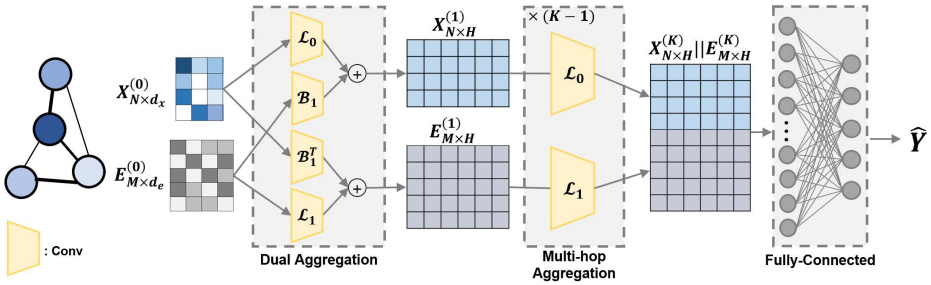


Fig. 2: The overall process of the proposed method. Both node and edge features, i.e., X and E, are convolved with incidence matrix and Laplacians through Dual and Multi-hop aggregations for a unified representation.

4 Method

Consider an undirected weighted graph $G = (A, X, E)$ composed of N nodes and M edges, along with a multi-variate feature matrix $X = \{x_1, ..., x_N\} \in \mathbb{R}^{N\times d_x}$ for the nodes and a $E = \{e_1, ..., e_M\} \in \mathbb{R}^{M\times d_e}$ for the edges in d_x- and d_e-dimension, respectively. $A \in \mathbb{R}^{N\times N}$ represents a binary adjacency matrix indicating node connections. From A, connectivities between nodes and edges are determined and it can be expressed as the incidence matrix $\mathcal{B}_1$. Utilizing $\mathcal{B}_1$, we obtain $\mathcal{L}_0$ and $\mathcal{L}_1$ through Eq. (2) and Eq. (3).

The complete configuration of our model is illustrated in Fig. 2. To separately learn node and edge representations, we conduct graph convolution on nodes and edges using $\mathcal{L}_0$ and $\mathcal{L}_1$, respectively. Simultaneously, we perform graph convolution using $\mathcal{B}_1$ to obtain embeddings based on the node-edge relationships for each node and edge. The resulting embeddings are then individually combined for each node and edge using element-wise averaging to achieve topologically enhanced embeddings. Subsequently, additional convolution layers utilizing $\mathcal{L}_0$ and $\mathcal{L}_1$ facilitate the exploration of a broader range of graph structures. The obtained embeddings for nodes and edges are concatenated and input into a multi-layer perceptron (MLP), yielding class predictions. Layer normalization [2] is applied to X and E before model input to prevent biased learning, considering the diverse distribution of features. Through this comprehensive process, our model effectively captures the complex topology of the brain network based on node-edge connectivity.

4.1 Dual Aggregation for Nodes

Let $f_1(\cdot)$ and $f_2(\cdot)$ be functions designed to generate embeddings for nodes, mapping $\mathbb{R}^{d_x} \to \mathbb{R}^H$ and $\mathbb{R}^{d_e} \to \mathbb{R}^H$, respectively, where H represents the hidden dimension. In the context of conventional node-wise spatial graph convolution, we utilize $\mathcal{L}_0$ to aggregate information about neighboring nodes for each node,

resulting in the following embedding:

$$f_1(X^{(0)}) = \sigma(\mathcal{L}_0 X^{(0)} \theta^{(0)}), \tag{4}$$

where $\theta^{(0)}$ denotes learnable parameters and $X^{(0)} = X$. Here, $\sigma(\cdot)$ signifies a rectified linear unit (ReLU) nonlinearity, consistently used with the same meaning throughout. Subsequently, for edge-wise feature aggregation for each node, we employ $\mathcal{B}_1$ as:

$$f_2(E^{(0)}) = \sigma(\mathcal{B}_1 E^{(0)} \theta^{(1)}), \tag{5}$$

where $\theta^{(1)}$ denotes learnable parameters and $E^{(0)} = E$. Through this process, each node captures information about its neighboring edges. The obtained embeddings $f_1(X^{(0)})$ and $f_2(E^{(0)})$ are then combined to derive $X^{(1)}$:

$$X^{(1)} = (f_1(X^{(0)}) + f_2(E^{(0)}))/2. \tag{6}$$

Thus, $X^{(1)}$ encompasses information about both the adjacent nodes and edges.

4.2 Dual Aggregation for Edges

Let $g_1(\cdot)$ and $g_2(\cdot)$ be functions designed to obtain embeddings for edges, each mapping $\mathbb{R}^{d_e} \to \mathbb{R}^H$ and $\mathbb{R}^{d_x} \to \mathbb{R}^H$. As done in [17], we utilize $\mathcal{L}_1$ for edge-wise spatial graph convolution to capture information about neighboring edges for each edge:

$$g_1(E^{(0)}) = \sigma(\mathcal{L}_1 E^{(0)} \theta^{(2)}), \tag{7}$$

where $\theta^{(2)}$ denotes learnable parameters. Next, for node-wise feature aggregation per edge, we employ $\mathcal{B}_1^T$ as follows, with $\theta^{(3)}$ representing learnable parameters:

$$g_2(X^{(0)}) = \sigma(\mathcal{B}_1^T X^{(0)} \theta^{(3)}). \tag{8}$$

Through this, each edge takes information about its neighboring nodes, i.e., the two nodes connected by the edge. The obtained embeddings $g_1(E^{(0)})$ and $g_2(X^{(0)})$ are then average-pooled to derive $E^{(1)}$:

$$E^{(1)} = (g_1(E^{(0)}) + g_2(X^{(0)}))/2. \tag{9}$$

As a result, $E^{(1)}$ involves information regarding both adjacent edges and nodes.

4.3 Multi-hop Aggregation

The embeddings $X^{(1)}$ and $E^{(1)}$ obtained through dual aggregation encapsulate information exclusively from directly connected nodes and edges. To achieve a broader scope of information, additional layers utilizing $\mathcal{L}_0$ and $\mathcal{L}_1$ are consecutively stacked K-1 times, facilitating multi-hop aggregation as done in [22]. For this purpose, we define $h_1(\cdot), h_2(\cdot) : \mathbb{R}^H \to \mathbb{R}^H$ as:

$$X^{(k+1)} = h_1(X^{(k)}) = \sigma(\mathcal{L}_0 X^{(k)} \theta^{(2k+2)}) \quad k = 1, \ldots, K-1, \tag{10}$$

$$E^{(k+1)} = h_2(E^{(k)}) = \sigma(\mathcal{L}_1 E^{(k)} \theta^{(2k+3)}) \quad k = 1, \ldots, K-1. \tag{11}$$

Since $X^{(1)}$ and $E^{(1)}$ already contain information about neighboring nodes and edges through convolution with the incidence matrix, subsequent convolution using $\mathcal{L}_0$ and $\mathcal{L}_1$ alone ensures comprehensive information propagation for both nodes and edges.

The resulting embeddings $X^{(K)} \in \mathbb{R}^{N\times H}$ and $E^{(K)} \in \mathbb{R}^{M\times H}$ are concatenated ($||$), flattened, and fed as input to an MLP. Following a softmax layer, the model produces its final output $\hat{Y}$. Specifically, the class prediction $\hat{Y}^c$ for each class c is expressed as:

$$\hat{Y}^c = \frac{\mathrm{MLP}(X^{(K)}||E^{(K)})^c}{\sum_{c'\in C}\mathrm{MLP}(X^{(K)}||E^{(K)})^{c'}}. \tag{12}$$

Finally, the model is trained on T samples using the cross-entropy loss:

$$L_{ce} = -\sum_{t=1}^{T}\sum_{c\in C} Y_t^c \log \hat{Y}_t^c, \tag{13}$$

where, for the t-th sample, $Y_t^c = 1$ if the class is c, and $Y_t^c = 0$ otherwise.

5 Experiments

5.1 Dataset and Setup

Dataset. Diffusion Tensor Images (DTI) for a total of N=2020 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) are used. The images were registered on the Destrieux atlas [4], which includes 148 cortical regions and 12 subcortical regions. Applying tractography yields structural brain networks, representing 160 regions as nodes and the white matter fiber tracts between these nodes as edges, and the number of tracts between the regions serves as edge weight. For each subject, three types of node features were measured region-wise: cortical thickness from MRI, and Standardized Uptake Value Ratios (SUVR) of β-amyloid and FDG from PET, where SUVR was normalized with the Cerebellum as the reference. The dataset comprises five diagnostic labels: 843 Control (CN), 197 Significant Memory Concern (SMC), 490 Early Mild Cognitive Impairment (EMCI), 250 Late Mild Cognitive Impairment (LMCI), and 240 Alzheimer's Disease (AD) subjects. Details regarding the number of subjects available for each feature can be found in the supplementary materials.
Setup. Experiments are conducted for four scenarios with node measures, i.e., using 1) cortical thickness, 2) Amyloid-PET, 3) FDG-PET and 4) all three together. The structural brain networks for all subjects are averaged and edges below a specific threshold (=10000) are dropped. Additional details on the effect of thresholds on number of edges and model performance can be found in the supplementary. As baselines, the following models, utilizing both node and edge features, are used: Support Vector Machine (SVM), GCN [22], CensNet [8], and EGNN [5]. SVM takes the flattened concatenation of node and edge features as input, and the others, including ours, are evaluated by stacking two layers. For our model, additional evaluations are conducted under two distinct conditions: when the number of hops ($= K$) is 1 and in the absence of dual aggregation.

Table 1: Comparison of model performance with 5-fold cross validation. Two-hop configurations ($K = 2$) are used across models, SVM excepted. Our model is further evaluated at $K = 1$ and without dual aggregation (DA), respectively.

Method	Cortical Thickness				β-Amyloid			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
SVM	66.29	77.22	55.69	61.18	71.17	78.13	67.98	71.12
GCN [22]	80.79	83.27	77.59	79.82	82.79	85.16	82.67	83.58
EGNN [5]	82.48	84.48	79.12	81.30	82.26	83.74	81.84	82.43
CensNet [8]	82.18	85.27	79.20	81.73	85.25	87.65	84.70	85.92
Ours	**86.53 (±1.85)**	**87.34 (±1.11)**	84.65 (±2.35)	**85.85 (±1.74)**	89.34 (±1.77)	90.11 (±1.52)	**89.81 (±1.91)**	**89.91 (±1.70)**
Ours ($K = 1$)	86.09 (±0.90)	86.55 (±0.54)	**84.66 (±1.94)**	85.42 (±1.01)	**89.49 (±1.67)**	**90.59 (±1.53)**	89.24 (±1.72)	89.80 (±1.61)
Ours (w/o DA)	82.97 (±1.82)	83.72 (±1.52)	81.51 (±1.67)	82.48 (±1.47)	85.62 (±1.61)	87.33 (±1.71)	85.50 (±1.30)	86.30 (±1.39)
	FDG				All Features			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
SVM	71.06	75.89	69.36	71.63	71.32	76.43	68.94	71.32
GCN [22]	81.56	83.59	82.14	82.44	82.42	84.06	82.84	83.19
EGNN [5]	90.32	91.21	90.70	90.79	89.38	90.56	89.65	89.91
CensNet [8]	82.83	84.50	82.93	83.55	84.28	86.28	84.07	84.98
Ours	91.22 (±2.45)	92.34 (±2.23)	91.74 (±1.87)	91.94 (±1.96)	91.81 (±0.77)	92.95 (±1.05)	92.03 (±0.76)	92.38 (±0.75)
Ours ($K = 1$)	**91.68 (±1.82)**	**92.52 (±1.81)**	**92.19 (±1.42)**	**92.26 (±1.56)**	**93.03 (±1.06)**	**93.90 (±1.17)**	**93.35 (±1.02)**	**93.53 (±0.93)**
Ours (w/o DA)	87.47 (±2.34)	88.97 (±2.56)	88.13 (±1.55)	88.41 (±1.90)	86.87 (±1.55)	87.40 (±2.22)	87.18 (±0.98)	87.14 (±1.48)

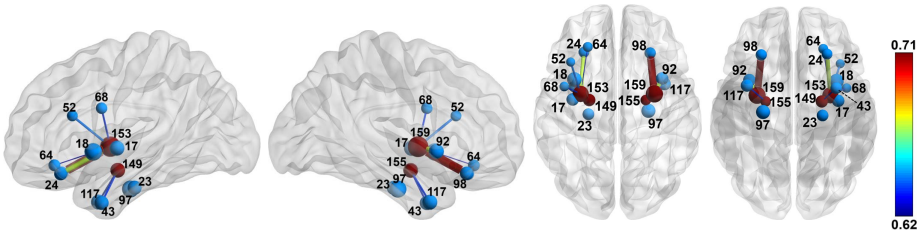
Table 2: The Grad-CAM outcomes depicting the top-10 ROIs (Left) and connectomes (Right) with the highest activation to classify AD. The indices align with the index values in the Destrieux atlas [4].

idx	ROI	Activation
159	right-putamen	0.9025
153	left-putamen	0.8224
18	left-g.insular.short	0.7131
97	right-g.oc-temp.med-parahip	0.6752
92	right-g.insular.short	0.5580
17	left-g.ins.lg.and.s.cent.ins	0.5559
117	right-pole.temporal	0.5552
149	left-amygdala	0.5445
43	left-pole.temporal	0.5060
23	left-g.oc-temp.med-parahip	0.5014

idx1	ROI1	idx2	ROI2	Activation
98	right-g.orbital	159	right-putamen	0.7101
18	left-g.insular.short	153	left-putamen	0.6980
92	right-g.insular.short	159	right-putamen	0.6708
24	left-g.orbital	153	left-putamen	0.6707
52	left-s.front.inf	153	left-putamen	0.6410
43	left-pole.temporal	149	left-amygdala	0.6347
117	right-pole.temporal	155	right-amygdala	0.6346
68	left-s.precentral-inf-part	153	left-putamen	0.6345
64	left-s.orbital-h.shaped	153	left-putamen	0.6252
17	left-g.ins.lg.and.s.cent.ins	153	left-putamen	0.6206

5.2 Results

Quantitative Results. As shown in Tab. 1, recent models, i.e., CensNet and EGNN, outperform traditional baselines such as SVM and GCN, but our model for both $K = 1, 2$ consistently demonstrates the highest performance across all datasets when compared to all baselines. The performance ranking of our model



(a) (b) (c) (d)

Fig. 3: Visualization of the top-10 ROIs and connectomes with the highest activation for classifying AD. (a)/(b): outer view of left/right hemisphere, (c)/(d): top/bottom. Blue nodes belong to the cortical region, and red nodes belong to the subcortical region. Node size and edge color/thickness represent the activation.

in $K = 1, 2$ varies depending on the node features and evaluation metric used. Furthermore, it can be observed that our model without dual aggregation shows inferior performance in all cases compared to when it is present, indicating the efficacy of dual aggregation.

Analysis on the Grad-CAM Results. Using class-averaged Grad-CAM [20], we obtained clinically interpretable results on the impact of each ROI and connectome in the classification of the brain network. In Tab. 2, we specified the top 10 nodes and edges displaying the highest gradient activation for classifying AD, while Fig. 3 visually depicts them.

Regarding ROIs from both nodes and edges, our model selected four ROIs (left/right putamen, left/right amygdala) within the subcortical region. Interestingly, putamen exhibited the highest activation among the nodes, concentrated on 8 out of 10 connectomes. These regions are known to exhibit significant atrophy in AD subjects compared to normal individuals [3, 11]. In the temporal lobe, the parahippocampal gyrus and temporal pole are detected. Notably, the parahippocampal gyrus plays a crucial role in memory encoding and retrieval [12, 16], with patients diagnosed with AD showing noteworthy reductions in its gray matter volume [6].

Additionally, our model selected the same 6 regions (putamen, short insular gyri, parahippocampal gyrus, temporal pole, amygdala, orbital gyri) in both the left and right hemispheres, indicating its ability to identify consistent features on both hemispheres of the brain. There are also 8 common ROIs (left/right putamen, left amygdala, left/right short insular gyri, left long insular gyrus and central sulcus of the insula, left/right temporal pole) selected for both nodes and edges. It can be inferred that, via dual aggregation, each node and edge adeptly incorporate information from one another.

6 Conclusion

We proposed a model that explicitly learns both node and edge features of a graph, effectively capturing complex node-edge relationships. Utilizing the incidence matrix, our model employs a dual aggregation framework to enable both 0-simplices (nodes) and 1-simplices (edges) to gather information from their surrounding 0- and 1-simplices. This holds the potential to be extended to higher-dimensional simplices. With the interpretablity identifying key ROIs and connectomes related to the progression of AD, the superior performance on the ADNI dataset demonstrates our model's capability in analyzing both region-wise and connectomic features of a brain.

Acknowledgments. This research was supported by NRF-2022R1A2C2092336 (50%), RS-2022-II2202290 (20%), RS-2019-II191906 (AI Graduate Program at POSTECH, 10%) funded by MSIT, RS-2022-KH127855 (10%), and RS-2022-KH128705 (10%) funded by MOHW in South Korea.

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

References

1. Anand, D.V., Chung, M.K.: Hodge Laplacian of brain networks. IEEE Transactions on Medical Imaging (2023)
2. Ba, J.L., Kiros, J.R., Hinton, G.E.: Layer normalization. arXiv preprint arXiv:1607.06450 (2016)
3. Cuénod, C.A., Denys, A., et al.: Amygdala atrophy in Alzheimer's disease: an in vivo magnetic resonance imaging study. Archives of Neurology **50**(9), 941–945 (1993)
4. Destrieux, C., Fischl, B., et al.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage **53**(1), 1–15 (2010)
5. Gong, L., Cheng, Q.: Exploiting edge features for graph neural networks. In: Computer Vision and Pattern Recognition. pp. 9211–9219 (2019)
6. Guo, X., Wang, Z., et al.: Voxel-based assessment of gray and white matter volumes in Alzheimer's disease. Neuroscience Letters **468**(2), 146–150 (2010)
7. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. Advances in Neural Information Processing Systems **30** (2017)
8. Jiang, X., Ji, P., Li, S.: Censnet: Convolution with Edge-Node Switching in Graph Neural Networks. In: International Joint Conference on Artificial Intelligence. pp. 2656–2662 (2019)
9. Jin, B., Zhang, Y., et al.: Edgeformers: Graph-empowered transformers for representation learning on textual-edge networks. International Conference on Learning Representations (2023)
10. Jo, J., Baek, J., et al.: Edge representation learning with hypergraphs. Advances in Neural Information Processing Systems **34**, 7534–7546 (2021)
11. de Jong, L.W., van der Hiele, K., et al.: Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. Brain **131**(12), 3277–3285 (2008)

12. Köhler, S., Black, S., et al.: Memory impairments associated with hippocampal versus parahippocampal-gyrus atrophy: an MR volumetry study in Alzheimer's disease. Neuropsychologia **36**(9), 901–914 (1998)
13. Li, X., Zhou, Y., et al.: Braingnn: Interpretable brain graph neural network for fMRI analysis. Medical Image Analysis **74**, 102233 (2021)
14. Ma, X., Wu, G., et al.: Learning multi-resolution graph edge embedding for discovering brain network dysfunction in neurological disorders. In: Information Processing in Medical Imaging. pp. 253–266. Springer (2021)
15. Miraglia, F., Vecchio, F., Rossini, P.M.: Searching for signs of aging and dementia in EEG through network analysis. Behavioural Brain Research **317**, 292–300 (2017)
16. Owen, A.M., Milner, B., et al.: A specific role for the right parahippocampal gyrus in the retrieval of object-location: A positron emission tomography study. Journal of Cognitive Neuroscience **8**(6), 588–602 (1996)
17. Park, J., Hwang, Y., et al.: Convolving directed graph edges via Hodge Laplacian for brain network analysis. In: Medical Image Computing and Computer Assisted Intervention. pp. 789–799. Springer (2023)
18. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. Neuroimage **52**(3), 1059–1069 (2010)
19. Schaub, M.T., Zhu, Y., et al.: Signal processing on higher-order networks: Livin'on the edge... and beyond. Signal Processing **187**, 108149 (2021)
20. Selvaraju, R.R., Cogswell, M., et al.: Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: International Conference on Computer Vision. pp. 618–626 (2017)
21. Sim, J., Jeon, S., et al.: Learning to Approximate Adaptive Kernel Convolution on Graphs. In: AAAI Conference on Artificial Intelligence. vol. 38, pp. 4882–4890 (2024)
22. Welling, M., Kipf, T.N.: Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations (2016)
23. Yang, M., Isufi, E., Leus, G.: Simplicial convolutional neural networks. In: International Conference on Acoustics, Speech and Signal Processing. pp. 8847–8851. IEEE (2022)
24. Yang, Y., Li, D.: Nenn: Incorporate node and edge features in graph neural networks. In: Asian Conference on Machine Learning. pp. 593–608. PMLR (2020)
25. Zhang, W., Zhan, L., et al.: Deep representation learning for multimodal brain networks. In: Medical Image Computing and Computer Assisted Intervention. pp. 613–624. Springer (2020)