Disentangled Attention Graph Neural Network for Alzheimer's Disease Diagnosis

,Gurur Gamgam[1], Alkan Kabakcioglu[2], Demet Yüksel Dal[1], and Burak Acar[1]

[1] VAVlab, Department of Electrical And Electronics Engineering, Bogazici University, İstanbul, Turkey
[2] Department of Physics, Koç University, İstanbul, Turkey

Abstract. Neurodegenerative disorders, notably Alzheimer's Disease type Dementia (ADD), are recognized for their imprint on brain connectivity. Recent investigations employing Graph Neural Networks (GNNs) have demonstrated considerable promise in diagnosing ADD. Among the various GNN architectures, attention-based GNNs have gained prominence due to their capacity to emphasize diagnostically significant alterations in neural connectivity while suppressing irrelevant ones. Nevertheless, a notable limitation observed in attention-based GNNs pertains to the homogeneity of attention coefficients across different attention heads, suggesting a tendency for the GNN to overlook spatially localized critical alterations at the subnetwork scale (mesoscale). In response to this challenge, we propose a novel Disentangled Attention GNN (DAGNN) model trained to discern attention coefficients across different heads. We show that DAGNN can generate uncorrelated latent representations across heads, potentially learning localized representations at mesoscale. We empirically show that these latent representations are superior to state-of-the-art GNN based representations in ADD diagnosis while providing insight into spatially localized changes in connectivity.

Keywords: Structural Networks · Graph Neural Networks · Attention Mechanism · Disentanglement.

1 Introduction

The brain is known to undergo significant organizational changes through the course of dementia and numerous techniques have been proposed to examine these changes by studying brain networks [1, 2]. Recently, there has been a growing interest in employing graph neural networks (GNNs) for investigating brain networks [3, 4]. While resting-state fMRI based functional networks (fNETs) have been widely used and proven effective [5–7] , diffusion-weighted imaging (DWI) based structural networks (sNETs) are often neglected, despite the significant role of structural changes during dementia. Besides their importance, the sparsity of structural networks makes them more suitable for GNNs, known to excel in handling sparse graphs.

Lately, attention-based GNNs (AGNNs) have gained prominence in the field of brain networks [8, 9]. Attention mechanism enhances GNNs by enabling them

to identify important connections and filter out unwanted ones achieved through dynamically modeling relationships via computed attention coefficients. AGNNs commonly utilize a multi-head architecture wherein each head operates independently, potentially increasing the model's capacity. However, we observe a key limitation : different attention heads produce similar attention coefficients despite their separate structures. The uniformity of attention coefficients across various attention heads suggests that AGNNs may fail to recognize certain critical subnetworks.

To resolve this issue, we propose a novel model, Disentangled-Attention GNN (DAGNN), that generates spatially distinct attention coefficients across heads, enforced by a disentanglement loss. In addition to rectifying the uniformity of attention coefficients, DAGNN also uncouples representations across heads, resulting in higher-quality outputs. In order to assess its capacity, we have limited our study to discriminate the Mild Cognitive Impairment (MCI) subjects from ADD subjects using sNETs, which is known to be more challenging. Our experimental findings on three diverse datasets demonstrate that DAGNN surpasses existing state-of-the-art methods in AD-MCI classification. Furthermore, we analyzed DAGNN in order to provide explanations regarding spatially localized changes in connectivity.

2 Method

Problem Statement. Consider we have a set of sNETs $\{\mathcal{G}^1, \mathcal{G}^2, ...\mathcal{G}^P\}$ where P is the number of patients. Each sNET is represented by its weighted and undirected adjacency matrix $\mathbf{A}^p \in \mathbb{R}^{N\times N}$ where N is the number of brain regions. Let $\mathbf{X}^p \in \mathbb{R}^{N\times d_{in}}$ be the node feature matrix where d_{in} is the number of features. We seek to find a model g_Θ parameterized with Θ which takes sNET and node features as input and outputs the label of the patient l^p, i.e $g_\Theta(\mathbf{A}^p, \mathbf{X}^p) = l^p$.

2.1 DAGNN

The Disentangled Attention GNN (DAGNN) model, which we propose here, is depicted in Figure 1. Our model has three key components : encoder, attention-based GNN, and pooling module where each of them can be expressed as the following equations :

$$\mathbf{H} = enc(\mathbf{A}, \mathbf{X}) \qquad \mathbf{Z} = GNN(\mathbf{H}, \mathbf{A}) \qquad \mathbf{y} = pool(\mathbf{Z}) \tag{1}$$

where $\mathbf{H} \in \mathbb{R}^{N\times d_{enc}}$ is the encoded node features, $\mathbf{Z} \in \mathbb{R}^{N\times d_{out}}$ is the node representations generated by GNN and $\mathbf{y} \in \mathbb{R}^{d_{out}}$ is the graph level representation.

Encoder Common practice for AGNNs is to leverage positional embeddings regardless of whether or not node features are present. Eigenvectors and eigenvalues of each node are generally used for such purposes [10] i.e they are used as $\mathbf{H}$ or as complementary to node features. However, eigenvectors suffer from sign

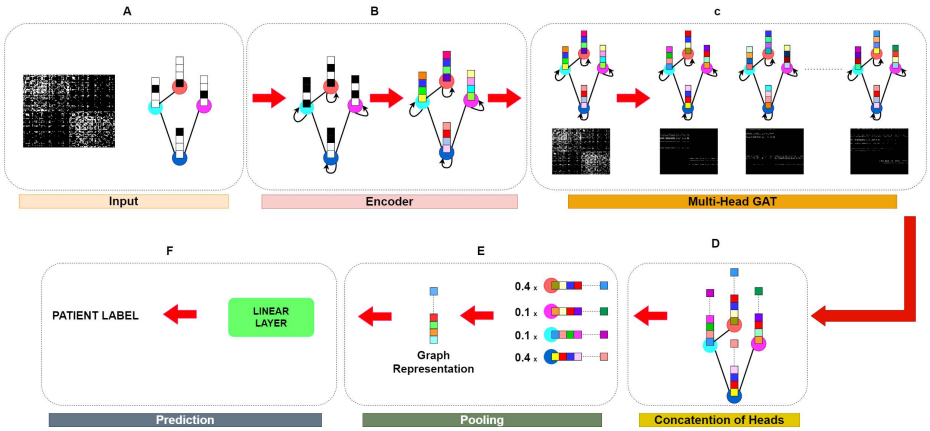

Fig. 1. DAGNN Pipeline. (A) Input sNET. (B) Encoding adjacency matrix with GIN. It is equivalent to apply MLP to each row of binary adjacency matrix with self loops. (C) Multi-head GAT. Each head is independent from each other and compute a set of different attention coefficients along with node features. (D) Node features from heads are concatenated. (E) Attention based pooling is performed to obtain graph-level representation. (F) A linear layer with softmax activation is used to predict label.

ambiguity which can harm the quality of embeddings. To overcome this issue, we propose to use Graph Isomorphism Network(GIN) [11] to encode the graph structure. GIN possesses the capability to uniquely encode different graphs, supported by a well-defined theoretical basis associated with the Weisfeiler-Lehman test [12]. GIN formula is given below :

$$\mathbf{H} = \sigma(f_\Theta((\tilde{\mathbf{A}} + (1+\epsilon)\cdot\mathbf{I})\cdot\mathbf{X})) \tag{2}$$

where σ is ReLU function, f_Θ is a two-layer neural network parameterized by Θ, $\tilde{\mathbf{A}}$ is binary adjacency matrix, ϵ is some irrational number and $\mathbf{I}$ is identity matrix. For practical cases, we can let $\epsilon = 0$ and assign one-hot encoding features to nodes which means to set $X = I \in \mathbb{R}^{N\times N}$. Therefore the equation 2 reduces to :

$$\mathbf{H} = \sigma(f_\theta(\tilde{\mathbf{A}} + \mathbf{I})) \tag{3}$$

which is equivalent to encoding binary adjacency matrix with self loops. This approach could be seen as an alternative way to construct positional encoding in which connectivity values are used directly. It is also consistent with the literature that proposes to use connectivity values as feature vectors in the absence of node features [13, 14].

Attention-based GNN We used Graph Attention Network(GAT) [15] for our AGNN module. GAT forward propagation is given by :

$$\mathbf{Z}^k = \sigma(\hat{\mathbf{A}}^k \mathbf{\Theta}^k \mathbf{H}) \tag{4}$$

where $\mathbf{\Theta}^k \in \mathbb{R}^{d_{enc} \times d_{attn}}$ is linear mapping , $\hat{\mathbf{A}}^k \in \mathbb{R}^{N \times N}$ is the computed attention coefficients and $\mathbf{Z}^k \in \mathbb{R}^{N \times d_{attn}}$ is the output for head k. Overall output Z is constructed by concatenating the different heads along the head dimension, i.e if there are K heads, $d_{out} = K \cdot d_{attn}$. In each head, attention coefficients computed as the following :

$$a_{i,j} = \frac{e^{(LeakyReLU(\mathbf{a}^{k\top} \Theta^k \mathbf{h_i} + \mathbf{a_k}^T \Theta^k \mathbf{h_j}))}}{\sum_{k \in \mathcal{N}(i) \cup \{i\}} e^{LeakyReLU(\mathbf{a}^{k\top} \Theta^k \mathbf{h_j} + \mathbf{a_k}^T \Theta^k \mathbf{h_k})}} \tag{5}$$

where $a_{i,j}$ is the attention coefficient between node i and node j, $\mathbf{a}^k \in \mathbb{R}^{2 \cdot d_{attn}}$ is learnable vector, $\mathbf{h_i}$ is the representation of node i, (ith column of $\mathbf{H}$), $\mathcal{N}(i)$ neighbors of node i and Leaky-ReLU is introduced as non-linearity function to prevent cancelling exponentials when $k = i$. GAT learns spatially localized representations since it only computes the relations between neighbors.

Pooling We employed an attention-based pooling method that weights each node with a learnable score computed by a linear layer [16]. Here is the formula for pooling operation :

$$\mathbf{y} = \sum_{k=1}^{k=N} w_i \mathbf{z_i} \tag{6}$$

where w_i is node score of node i calculated by $w_i = softmax(\Theta \mathbf{z_i})$, $\mathbf{z_i}$ is the representation of node i and $\Theta \in \mathbb{R}^{d_{out}}$ is the linear mapping. Our pooling operator can highlight the important nodes which is not achievable by non-parametric pooling functions such as sum, mean or max. Additionally, it is more interpretable compared to Diff-Pool [17] and more expressive than the Top-K pool [18]

2.2 Disentanglement Loss

We noticed a limitation within the multi-head Graph Attention Network (GAT), where the attention coefficients computed by each head, denoted as $\hat{\mathbf{A}}^k$, are strikingly similar to each other. This observation suggests a potential drawback: GAT may not fully exploit its multi-head structure and might overlook significant variations. To address this issue, we introduce a novel loss function, termed as disentanglement loss, aimed at differentiating each head by maximizing the distance between attention coefficient matrices produced by heads. Our disentanglement loss is defined as :

$$\mathcal{L}_{dis} = ReLU(m - \frac{1}{\#pairs} \sum_{b=1}^{B} \sum_{j=1}^{K} \sum_{k=j+1}^{K} dist(\hat{\mathbf{A}}^{b,j}, \hat{\mathbf{A}}^{b,k})) \tag{7}$$

where $\hat{\mathbf{A}}^{b,j}$ is the attention coefficients of bth patient in a batch produced by jth head of GAT, B is the batch size, $\#pairs$ is the number of attention coefficients pairs in the batch, $dist$ is the function for measuring distance and m is the margin hyperparameter.

We propose to use the mean of L1 distance as our distance measure since it has two main advantages : The first one is that it is bounded by 2 and helps us to choose meaningful m hyperparameter. The second one is that using L1 distance also promotes the sparsity on the attention coefficients. The proposed distance function is :

$$dist(\hat{\mathbf{A}}^j, \hat{\mathbf{A}}^k) = \frac{1}{N} \sum_{n=1}^{n=N} \sum_{l=1}^{n=N} |C_{n,l}| \tag{8}$$

where $\mathbf{C} = \mathbf{A}^j - \mathbf{A}^k$. L1 distance between two matrices is computed then the average is taken as distance between two matrices.

We trained our model with a combination of classification loss and disentanglement loss. Overall loss function can be written as :

$$\mathcal{L}_{total} = \mathcal{L}_{cls} + \lambda \mathcal{L}_{dis} \tag{9}$$

where λ is hyperparameter that balances classification and disentanglement loss and $\mathcal{L}_{cls}$ is the negative log-likelihood function.

3 Experiments

3.1 Data

We used three datasets : CAPA , Alzheimer's Disease Neuroimaging Initiative (ADNI) and CAPA+ADNI which is the concatenation of the former two datasets. CAPA dataset is a private dataset acquired from 64 volunteers (18 ADD, 46 MCI) at Istanbul University, Istanbul Faculty of Medicine, Neuroimaging Unit of Hulusi Behçet Life Sciences Research Laboratory with written consent and under ethics committee approval. Clinical diagnostic labels are given following the NIA-AA guidelines [20]. We used 148-parcel Destrieux atlas ($N = 148$) [21] to delineate the cortical regions. Our custom pre-processing and network generation pipeline that utilizes FreeSurfer, FSL, and Tortoise can be found here[3].

ADNI dataset is the age-gender matched subset of the publicly available ADNI3 dataset [19] consisting of 42 patients (18 ADD, 24 MCI). Same preprocessing and network construction pipeline steps are repeated as above. This subset was selected to match the size of the CAPA dataset.

CAPA + ADNI dataset is the merging of the CAPA and ADNI datasets which have 106 patients (36 ADD, 70 MCI).

[3] https://vavlab.boun.edu.tr/brainet-structural-and-functional-brain-network-analysis

Model	CAPA		ADNI		CAPA + ADNI	
	F1	AUC	F1	AUC	F1	AUC
BrainGB-GAT	0.63±0.05	0.63±0.05	0.60±0.06	0.67±0.03	0.62±0.05	0.68±0.03
BrainGB-GCN	0.65±0.04	0.68±0.04	0.64±0.05	0.71±0.04	0.66±0.04	0.71±0.02
BrainGNN	0.70±0.06	0.74±0.05	0.72±0.04	0.78±0.02	0.73±0.03	0.77±0.03
DisenGCN	0.65±0.04	0.69±0.03	0.66±0.04	0.72±0.03	0.68±0.03	0.73±0.03
FactorGCN*	0.80±0.03	0.83±0.04	0.78±0.04	0.87±0.02	0.80±0.02	0.84±0.01
AGNN	0.74±0.03	0.79±0.03	0.75±0.03	0.85±0.03	0.76±0.01	0.82±0.01
DAGNN	**0.82±0.03**	**0.86±0.03**	**0.80±0.04**	**0.89±0.02**	**0.82±0.03**	**0.87±0.03**

Table 1. ADD-vs-MCI classification performances computed using cross-validation for end-to-end trained models.

3.2 Experimental Setup

We emprically set $d_{enc} = 128$, $d_{attn} = 32$, $K = 4$ and $d_{out} = 128$ in DAGNN. For disentanglement loss we set m=2 and $\lambda = 0.1$. We used ADAM optimizer with a learning rate of 0.001 , $1e-4$ L2 regularization and the model was trained over 100 epochs. Batch size is set to 6. 8-fold, 7-fold, and 9-fold cross-validation is used to compute test accuracy and area under curve (AUC) scores for CAPA, ADNI, and CAPA+ADNI, respectively. Fold numbers are adjusted to ensure an even distribution of data splitting. Cross-validation experiments are repeated 20 times to eliminate the impact of random seeds. We implemented our DAGNN model with Pytorch Geometric. The source codes are available on Github[4].

3.3 Results

We performed ADD-MCI classification task. For comparison, we also report the performance of different state-of-the-art (SOTA) models (best performing GAT and Graph Convolution Network (GCN) [22] models from BrainGB [13] and BrainGNN [23]) and other disentanglement based methods (FactorGCN [24] and DisenGCN [25]) in the table 1. FactorGCN factorizes input graph into factor graphs where each factor graph is assumed to be distinguishable by a classifier. To compare DAGNN with FactorGCN in a fair setup, which we denote as FactorGCN*, we replaced our disentanglement loss in Equation 7 with a classifier-based loss :

$$\mathcal{L}_{factor} = L_{cls}(k, G(\hat{\mathbf{A}}^{b,j})) \tag{10}$$

where G is a GNN based graph encoder and k is the label of factor graph $\hat{\mathbf{A}}^{b,j}$. On the other hand, DisenGCN aims to disentangle node representations by partitioning the neighbors. In addition, to justify the disentanglement we also add the performance of AGNN which is the same as DAGNN but does not utilize disentanglement loss.

[4] https://github.com/gururgg/DAGNN

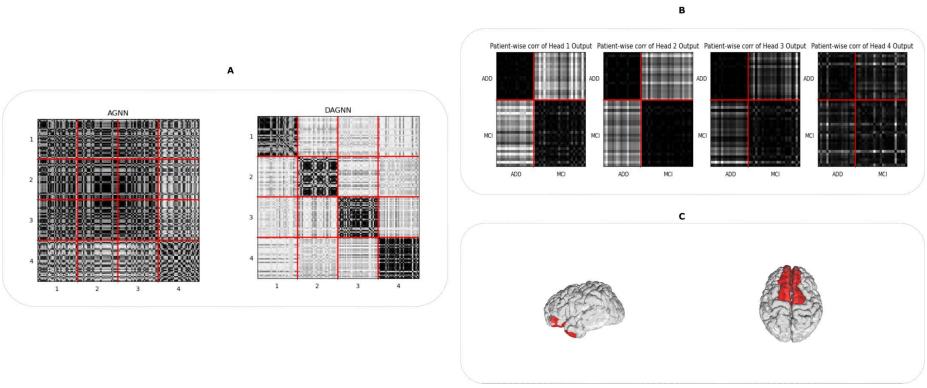

Fig. 2. DAGNN Analysis. The model is trained on the ADNI+CAPA dataset. (A) Correlation matrices of features among latent representations generated by AGNN (left) and DAGNN (right) is illustrated. The numerical annotations on the figures shows that which head contributes to which part of the representations. (B) Correlation of latent representations between patients for each head output is shown. (C) Highlighted brain regions by Head 1 (left) and Head 2 (right) are given.

4 Discussion

Results in table 1 show that DAGNN significantly outperforms other SOTA models and disentanglement based models. DAGNN' ability to generate decorrelated latent representations, each localized at mesoscale, might play a key role in its performance. Compared to AGNN, DAGNN is able to learn distinct latent representations which is indicated by reduced correlation values between off-diagonal elements as depicted in Figure 2. in Figure 2. Our proposed distance based disentanglement improves AGNN more than the classification based loss, since it directly maximizes the distances and hence produces more distinct representations across heads, while classification based loss softens the distinctness condition by working on the latent space generated by G.

In addition, we investigate which attention heads play a significant role in distinguishing patients. To find it, we take each $\mathbf{Z}^p$, GAT output of patient p, for all patients and compute the correlation coefficient matrix across all patients. As depicted in Figure 2, outputs of heads 1 and 2 exhibit lower inter-class correlations, indicating that the representations they generate are discriminatory. To determine which parts of the brain are highlighted by these heads, we calculated the mean attention map denoted as $\ddot{\mathbf{A}}^k$ which is equal to the average of $\hat{\mathbf{A}}^{p,k}$ over all patients. We treated the row sum of $\ddot{\mathbf{A}}^k$ as an importance score of brain regions. For each of the two heads, we identified two nodes displaying positive deviations from the mean node score that exceeds 3 standard deviations. We showed these regions in Figure 2 within the brain images where these regions are colored as red. Our analysis showed that head 1 focuses on left Orbital gyri and left Temporal pole while head 2 focuses on left Superior frontal gyrus and right

Superior frontal gyrus. Among these four regions, three of them (left Orbital gyri, left and right Superior frontal gyrus) are parts of default mode network (DMN) [26], which is known to be affected by ADD progression [27].

5 Conclusion

We proposed a novel model, DAGNN, to alleviate the homogeneity of attention coefficients across different attention heads. Utilizing disentanglement loss, DAGNN effectively distinguishes attention coefficients across various heads. Our experimental results on three sNET based datasets show that DAGNN' ability to learn distinct latent representations for each head improves the ADD-MCI classification. Additionally we analysed DAGNN by examining attention coefficients produced by heads and found that DAGNN highlights brain regions that are neurologically relevant to ADD. Future work will involve conducting additional experiments with the complete ADNI3 set to demonstrate the robustness of DAGNN.

Acknowledgments. This work was in part supported by the Turkish Directorate of Strategy and Budget [Grant: TAM Project number 2007K12-873], TÜBİTAK-ARDEB 1003 Programme [Grant #114E053], Bogazici University Research Fund [Grant #16862]. We would specifically like to thank H. Gurvit, Z. Yıldırım, T. Demiralp, Ç. Ulaşoğlu-Yıldız, E. Kurt, G. Bayır, E. Özarslan, E. Özacar, B. Bilgiç, A. Demirtaş-Tatlıdede for their contributions in data collection, preprocessing, and clinical evaluation.

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

References

1. Durusoy, G., Yıldırım, Z., Dal, D.Y.,Ulasoglu-Yildiz, Cigdem.,Kurt, Elif., Bayır, Güneş., Özacar, Erhan., Özarslan, Evren., Demirtaş-Tatlıdede, Aslı., Bilgiç, Başar., Demiralp, Tamer., Gurvit, Hakan., Kabakçıoğlu, Alkan.,Acar, Burak.: B-Tensor: Brain Connectome Tensor Factorization for Alzheimer's Disease. IEEE Journal of Biomedical and Health Informatics **25**(5), 1591–1600 (2021)
2. Sarraf, S., Tofighi, G.: Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data. In: 2016 Future Technologies Conference (FTC), pp. 816–820. IEEE (2016)
3. Ktena, S.I., et al.: Distance metric learning using graph convolutional networks: application to functional brain networks. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10433, pp. 469–477. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_54
4. Parisot, S., et al.: Spectral graph convolutions for population-based disease prediction. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 177–185. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_21

5. Xing, X. et al. (2019). Dynamic Spectral Graph Convolution Networks with Assistant Task Training for Early MCI Diagnosis. In: Shen, D., et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019. LNCS, vol 11767. Springer, Cham. https://doi.org/10.1007/978-3-030-32251-9_70
6. Li X, Dvornek NC, Zhou Y, Zhuang J, Ventola P, Duncan JS. Graph Neural Network for Interpreting Task-fMRI Biomarkers. Med Image Comput Comput Assist Interv. pp. 485-493 (2019) https://doi.org/10.1007/978-3-030-32254-0_54.
7. Zhang, Y., Huang, H. (2019). New Graph-Blind Convolutional Network for Brain Connectome Data Analysis. In: Chung, A., Gee, J., Yushkevich, P., Bao, S. (eds) Information Processing in Medical Imaging. IPMI 2019. LCNS, vol 11492. Springer, Cham. https://doi.org//10.1007/978-3-030-20351-1_52
8. Ma, J., Zhu, X., Yang, D., Chen, J., Wu, G. Attention-Guided Deep Graph Neural Network for Longitudinal Alzheimer's Disease Analysis. In: Martel, A.L., et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. MICCAI 2020. LCNS, vol 12267. Springer, Cham. https://doi.org/doi.org/10.1007/978-3-030-59728-3_38
9. Nerrise, F., Zhao, Q., Poston, K.L., Pohl, K.M., Adeli, E. An Explainable Geometric-Weighted Graph Attention Network for Identifying Functional Networks Associated with Gait Impairment. In: Greenspan, H., et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. MICCAI 2023. LCNS, vol 14221. Springer, Cham. https://doi.org/10.1007/978-3-031-43895-0_68
10. Rampášek, L., Galkin, M., Dwivedi, V.P., Luu, A.T., Wolf, G., Beaini, D . : Recipe for a General, Powerful, Scalable Graph Transformer. Advances in Neural Information Processing Systems (2022)
11. Xu, K. , Hu,W. , Leskovec J., Jegelka, S. : How Powerful are Graph Neural Networks? International Conference on Learning Representations (2019)
12. Weisfeiler, B., Lehman, A.A. : A reduction of a graph to a canonical form and an algebra arising during this reduction. Nauchno-Technicheskaya Informatsia, **2**(9):12–16, (1968)
13. H. Cui, W. Dai, Y. Zhu, X. Kan, A. A. C. Gu, J. Lukemire, L. Zhan, L. He, Y. Guo, C. Yang, Braingb: A benchmark for brain network analysis with graph neural networks, IEEE Transactions on Medical Imaging 42 (2) pp. 493–506 (2023)
14. Kan, X. , Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C. : Brain Network Transformer.Advances in Neural Information Processing Systems (2022)
15. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. In: ICLR, Vancouver, Canada (2018)
16. Li, Y., Tarlow, D., Brockschmidt, M., Zemel, R. : Gated Graph Sequence Neural Networks. ICLR (2016)
17. Ying, R., You, J., Morris, C., Ren, X., Hamilton, W.L., Leskovec, J. : Hierarchical graph representation learning with differentiable pooling. In Proceedings of the 32nd International Conference on Neural Information Processing Systems. Curran Associates Inc., Red Hook, NY, USA, pp. 4805–4815 (2018)
18. Knyazev, B., Taylor, G.W., Amer, M.R. : Understanding Attention and Generalization in Graph Neural Networks. NIPS (2019)
19. Jack Jr., C.R., et al.: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging Off. J. Int. Soc. Magn. Reson. Med. **27**(4), 685–691 (2008)
20. M. S. Albert, S. T. DeKosky, D. Dickson, B. Dubois, H. H. Feldman, N. C. Fox, A. Gamst, D. M. Holtzman, W. J. Jagust, R. C. Petersen, P. J. Snyder, M. C. Carrillo, B. Thies, C. H. Phelps. : The diagnosis of mild cognitive impairment due

to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's Disease, Alzheimer's & Dementia **7**(3) (2011) 270–279.
21. Destrieux, C., Fischl, B., Dale, A., Halgren, E.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage **27**(1), 1–15 (2010)
22. Kipf, T. N. and Welling, Max. : Semi-Supervised Classification with Graph Convolutional Networks. ICLR (2017)
23. X. Li, Y. Zhou, N. Dvornek, M. Zhang, S. Gao, J. Zhuang, D. Scheinost, L. H. Staib, P. Ventola, J. S. Duncan. : Braingnn: Interpretable brain graph neural network for fmri analysis, Medical Image Analysis 74 (2021)
24. Yang, Y., Feng, Z., Song, M., Wang, X. : Factorizable Graph Convolutional Networks. NIPS (2020)
25. Ma, J., Cui, P., Kuang, K., Wang, X., Zhu, W. : Disentangled Graph Convolutional Networks. Proceedings of the 36th International Conference on Machine Learning, in Proceedings of Machine Learning (2019)
26. Thomas Yeo, B. T. , Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., Buckner, R. L. : The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology **106**(3) pp. 1125–1165 (2011)
27. Mevel, K., Chételat, G., Eustache, F., Desgranges, B : The default mode network in healthy aging and Alzheimer's disease. Int J Alzheimers Dis. (2011) https://doi.org/10.4061/2011/535816