

Dual Attention Graph Convolutional Network Fusing Imaging and Genetic Data for Early Alzheimer's Disease Diagnosis

Jiaqiang Li, Peng Yang, Junlong Qu, Bao Yang, Zhenghua Guan, Xuegang Song, Xiaohua Xiao, Tianfu Wang, Baiying Lei*, *Senior Member, IEEE*

Abstract—Alzheimer's Disease (AD) poses a significant global neurodegenerative challenge, underscoring the urgency of early clinical intervention. Our paper presents a novel approach for early AD diagnosis, focusing on a dual attention graph convolutional network that integrates multi-modal data. This methodology involves constructing image and gene graphs based on the image and genetic information of the subject. Graph convolution networks are then employed to extract embedded information from each graph. Enhanced diagnostic precision is achieved by utilizing self-attention and cross-attention mechanisms, facilitating the fusion of multi-modal state information crucial for early AD identification. Rigorous validation of the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset underscores the model's efficacy. Our method has demonstrated remarkable proficiency in diagnosing early-stage AD through experimental verification, assisting doctors in making accurate diagnoses of AD.

Keywords—Alzheimer's Disease, Multi-Modality, Graph Convolutional Network, Image and Gene, Cross Attention.

I. INTRODUCTION

Alzheimer's disease (AD), a prevalent neurodegenerative condition exhibiting cognitive decline, poor judgment, and memory loss, displays an escalating prevalence with age[1]. Given its rising prevalence and lack of effective therapies, AD poses significant burdens on affected individuals[2]. Consequently, early clinical AD diagnosis and disease progression delay are crucial. Deep learning technology is progressively utilized for early AD diagnosis[3]. However, most existing methods primarily focus on the use of unimodal data, which overlooks the fact that AD is a neurodegenerative disease caused by a multitude of factors. Structural magnetic resonance imaging (sMRI) has illustrated changes in the relevant brain structures of the human body[4]. Single nucleotide polymorphisms (SNP) and clinical data, including age, gender, and body weight, have provided insights into AD pathogenesis[5]. Based on this background, our research will

integrate imaging and genetic data to facilitate the early diagnosis of AD.

Integrating multi-modal information facilitates the acquisition of comprehensive patient data for accurate disease predictions from diverse perspectives, thus enhancing prediction outcomes. Wang *et al.* [6] developed a deep multimodality-disentangled association analysis network for neurodegenerative diseases classification. Wang *et al.* [7] employed hypergraph-regularized multi-modal learning by graph diffusion for AD diagnosis. Therefore, multi-modal information enhances diagnostic precision by extracting patient data from various angles.

As neuroimaging and pattern recognition continue to advance, early diagnosis of AD through graph convolutional learning methods is increasingly emerging as a viable approach[8]. Researchers can train classification models using graph characteristics by constructing a graph where the nodes represent the features of each individual's image and the edges represent the similarity between each pair of nodes. Graph Convolutional Network (GCN) [9] have gained significant recognition due to their exceptional filtering performance, which is invaluable for AD diagnosis. Song *et al.* [3] employed similarity awareness and adaptive calibration in AD diagnosis using GCN. Song *et al.* [10] utilized a multi-center, multi-channel pooling GCN based on dual-modal fusion brain networks for early AD diagnosis. Bi *et al.* [11] developed community GCN for AD classification.

The attention mechanism has demonstrated potential in early AD diagnosis, as recently studied by researchers. Zhu *et al.* [12] used multi-modal triple attention networks for neurological disorder diagnosis. Additionally, Zhu *et al.* [13] further applied sMRI with dual attention multi-instance deep learning for AD diagnosis. Incorporating attention mechanisms enhances the ability to comprehensively extract deep information from MRI images, leading to more effective deep learning models.

This work was supported partly by National Key R&D Program of China (2022YFC2009903/2022YFC2009900), National Natural Science Foundation of China (62271328, 62201360), Guangdong Basic and Applied Basic Research Foundation (2024A1515011950, 2021A1515110746), Shenzhen Key Basic Research Project (KCYFZ20201221173213036, JCYJ20220818095809021).

Jiaqiang Li, Peng Yang, Junlong Qu, Bao Yang, Zhenghua Guan, Xuegang Song, Tianfu Wang, Baiying Lei are with School of Biomedical Engineering, Health Science Center, Shenzhen University, National-

Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Shenzhen, China.

Xiaohua Xiao is with Affiliated Hospital of Shenzhen University, Health Science Center, Shenzhen University, Shenzhen Second People's Hospital, Shenzhen, China

*Correspondence should be addressed to Baiying Lei (leiby@szu.edu.cn).

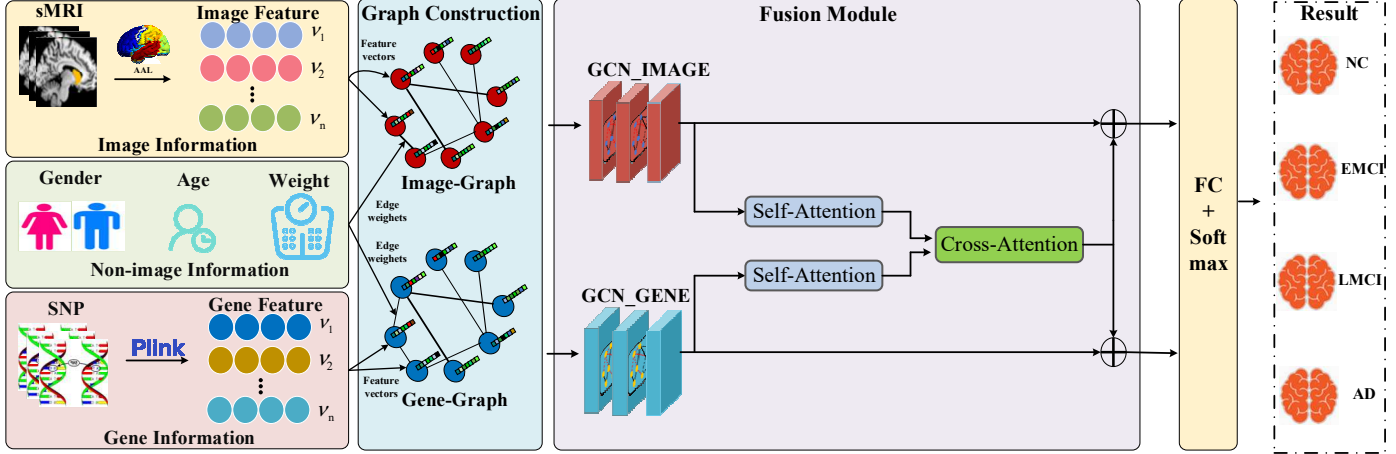


Fig. 1. Overview of the proposed method.

Based on the aforementioned background, this paper proposes a dual-attention GCN that integrates multimodal data for the diagnosis of AD (Fig. 1). For image data and genetic data, we constructed two graphs, namely the Imaging-Graph and Gene-Graph. Subsequently, we employed two GCN modules to extract the respective embedded information from each graph. Then, we utilized self-attention and cross-attention mechanisms to achieve the fusion of the two sets of embedded information. Finally, we employed fully connected layers and a Softmax function to implement disease classification. The model was validated on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset and demonstrated effective performance in disease classification.

II. METHODOLOGY

A. Data Acquisition and Preprocessing

The data utilized was extracted from the ADNI dataset[14]. We used sMRI, clinical information(age, gender, body weight), and genetic information from 726 subjects. The data in this study is categorized into four distinct groups: AD, normal control (NC), early mild cognitive impairment (EMCI), and late mild cognitive impairment (LMCI). The details of the data are shown in Table 1.

For the preprocessing of sMRI, we used Statistical Parametric Mapping (SPM)[15] and Computational Anatomy Toolbox(CAT)[16]. The preprocessing steps are as follows: 1) filtering noise; 2) resampling; 3) bias field correction; 4) bias field correction; 5) skull stripping; 6) linear alignment; 7) tissue segmentation; 8) spatial normalization; 9) cropping bound-aries; 10) gray matter and white matter extraction. After preprocessing, we use the AAL116 template[17] to get the ground truth 116 brain regions of interest (ROIs), and we use the top 90 ROIs[18].

The genetic data were processed using Plink[19] software for quality control operations. It includes: 1) missing value check for each subject and SNP; 2) sex check; 3) minor allele frequency check; 4) Hardy-Weinberg balance test; 5) controlling for heterozygosity rate bias; 6) removal of related samples; 7) population stratification.

TABLE I. THE DEMOGRAPHIC INFORMATION OF OUR DATASET

Label	Number	Age(years) (mean \pm SD)	Gender (F/M)	Weight (kg) (mean \pm SD)
AD	205	74.86 \pm 5.63	94/111	76.62 \pm 15.94
NC	166	74.73 \pm 8.12	72/94	73.96 \pm 14.78
EMCI	192	70.64 \pm 7.11	87/105	80.51 \pm 16.41
LMCI	163	72.74 \pm 7.54	69/94	77.96 \pm 18.10

B. Graph Construction

Graph construction plays a pivotal role in graph-based classification. A well-designed graph structure effectively guides the input and propagation of features, thereby enhancing classification performance. Our study focuses on constructing two critical graph structures: the Image-Graph $\mathbf{G}_I = (\mathbf{A}, \mathbf{X}_I)$ and Gene-Graph $\mathbf{G}_G = (\mathbf{A}, \mathbf{X}_G)$. For the image graph node feature \mathbf{X}_I , we constructed the graph node feature \mathbf{X}_I based on the ROIs, which have a structural correlation with AD. As for the gene node features \mathbf{X}_G , we used the preprocessed SNP data to construct the gene node features \mathbf{X}_G , which can well reflect the patient's genetic information. n represents the number of patients, d_I represents the image feature dimension, d_G represents the SNP data feature dimension.

$$\mathbf{X}_I \in R^{n \times d_I} \quad (1)$$

$$\mathbf{X}_G \in R^{n \times d_G} \quad (2)$$

In AD diagnosis, gender, age and weight information are valuable because different gender, age and weight information can reflect differences in patient images and genetic information. Therefore, both Image-Graph and Gene-Graph were constructed based on the patient's gender, age and weight information. The adjacency matrix \mathbf{A} is defined as follows:

$$\mathbf{A}(v, w) = r_g(g_v, g_w) + r_a(a_v, a_w) + r_w(w_v, w_w) \quad (3)$$

where g , a , w respectively represent the gender, age and weight information of the subject, $r(\cdot)$ represents the distance between phenotypic characteristics, r_g represents the gender distance, r_a represents the age distance, r_w represents the

weight distance, and we define r as a unit step function as follows:

$$r_g(g_i, g_j) = \begin{cases} 1, & g_i = g_j \\ 0, & g_i \neq g_j \end{cases} \quad (4)$$

$$r_a(a_i, a_j) = \begin{cases} 1, & |a_i - a_j| \leq a_{std} \\ 0, & |a_i - a_j| \geq a_{std} \end{cases} \quad (5)$$

$$r_w(w_i, w_j) = \begin{cases} 1, & |w_i - w_j| \leq w_{std} \\ 0, & |w_i - w_j| \geq w_{std} \end{cases} \quad (6)$$

where a_{std} and w_{std} represent the standard deviation of age and the standard deviation of weight, respectively.

C. Graph Convolutional Network

Building upon the foundation of the graph $\mathbf{G}=(\mathbf{A}, \mathbf{X})$, where $\mathbf{A} \in R^{n \times n}$ is a symmetric adjacency matrix with n nodes, $\mathbf{X} \in R^{n \times d}$ represents the node feature matrix, we employed two specialized graph convolution modules to learn Image-Graph $\mathbf{G}_I=(\mathbf{A}, \mathbf{X}_I)$ and Gene-Graph $\mathbf{G}_G=(\mathbf{A}, \mathbf{X}_G)$ respectively. These modules yield two distinct embedded information \mathbf{Z}_I and \mathbf{Z}_G respectively.

$$\mathbf{Z}_I^l = \text{Relu} \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{Z}_I^{l-1} \mathbf{W}_I^l \right) \quad (7)$$

$$\mathbf{Z}_G^l = \text{Relu} \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{Z}_G^{l-1} \mathbf{W}_G^l \right) \quad (8)$$

where \mathbf{W}_I^l is the weight matrix of the l th layer in GCN_IMAGE, the initial \mathbf{Z}_I^0 is the feature matrix \mathbf{X}_I . \mathbf{W}_G^l is the weight matrix of the l th layer in GCN_GENE, the initial \mathbf{Z}_G^0 is the feature matrix \mathbf{X}_G . $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, \mathbf{A} is the adjacency matrix, and \mathbf{I} is the identity matrix. $\tilde{\mathbf{D}}$ is the diagonal angle matrix of $\tilde{\mathbf{A}}$.

D. Multi-modal Fusion

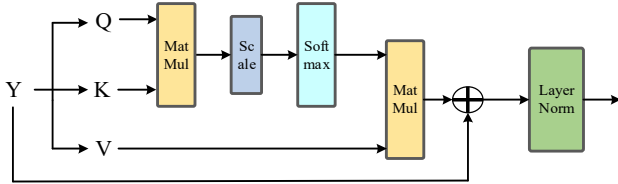


Fig. 2. Principle diagram of attention mechanism.

We set up a dual attention mechanism to fully mine the above-mentioned embedded information. This mechanism adaptively fuses the embedded information with the learned weights, enhancing the accuracy of early AD diagnosis. The dual attention mechanism encompasses both the self-attention mechanism and the cross-attention mechanism. The self-attention module is implemented as follows (Fig.2):

$$\mathbf{Q} = \mathbf{Y}\mathbf{W}_Q \quad \mathbf{K} = \mathbf{Y}\mathbf{W}_K \quad \mathbf{V} = \mathbf{Y}\mathbf{W}_V \quad (9)$$

$$\mathbf{f} = \text{Softmax} \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}} \right) \mathbf{V} \quad (10)$$

where \mathbf{Y} is the input, \mathbf{W}_Q , \mathbf{W}_K and \mathbf{W}_V are the parameter matrices used to generate queries, keys, and values respectively, which are updated through backpropagation of the network during the model training process. We capture the attentional coefficients between brain regions by computing the scaled dot product correlation between \mathbf{Q} and \mathbf{K} , which we then feed into the Softmax function. Finally, the feature

matrix \mathbf{f} with self-attention is calculated using the product of the attention weight and \mathbf{V} .

To fully explore the relationship between the two embedded information, we set up a cross-attention module in this work. The specific implementation details are as follows:

$$\mathbf{f}_{CA} = \text{Softmax} \left(\frac{\mathbf{f}_G \mathbf{f}_I^T}{\sqrt{d}} \right) \mathbf{f}_I \quad (11)$$

where \mathbf{f}_I and \mathbf{f}_G are the results obtained by \mathbf{Z}_I and \mathbf{Z}_G respectively after passing through the self-attention module, \mathbf{f}_{CA} is the result of the cross-attention module.

III. EXPERIMENTAL AND RESULTS

A. Experimental Setup

We used PyTorch for distributed parallel training of algorithmic models. We utilized a five-fold cross-validation method on the ADNI dataset. To ensure the fairness of the experimental results, all settings in the experiment are consistent with different comparison methods. In this article, we evaluate the network performance through some standard evaluation metrics such as accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the ROC curve (AUC).

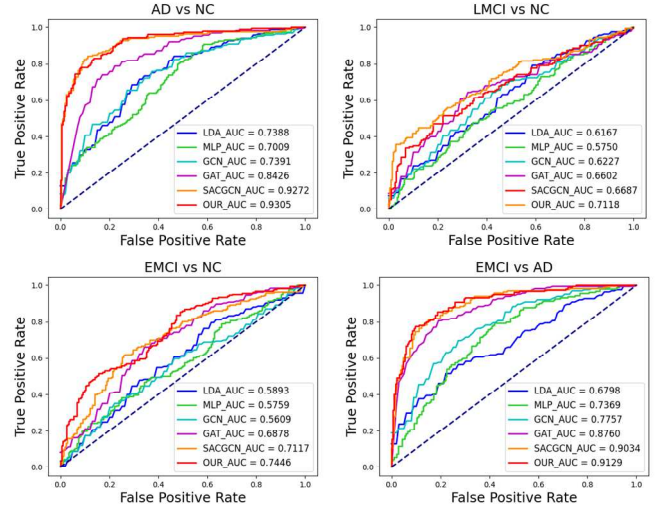


Fig. 3. The ROC curves of four classification tasks.

B. Experimental Results

In this study, we perform four binary classification tasks: AD vs. NC, LMCI vs. NC, EMCI vs. NC, EMCI vs. AD. To demonstrate the advantages of our proposed method, we compare it with Linear Discriminant Analysis(LDA)[20], Multilayer Perceptron (MLP)[21], GCN[9], Graph Attention Networks(GAT)[22], similarity-aware adaptive calibrated GCN (SACGCN)[3]. Table 2 presents the performance of different methods for these tasks. Our model exhibits superior performance, achieving accuracy rates of 85.95%, 67.99%, 67.51%, and 83.20% for the four binary classifications. The ROC curves of different models and tasks are shown in Fig. 3. Generally, our proposed method performs best among four binary classification tasks. These results show that our method can effectively identify early AD subjects from NC or several subcategories.

TABLE 2. COMPARISON OF DIFFERENT METHODS WITH FOUR CLASSIFICATION TASKS(%).

Task	Method	ACC	SEN	SPE	AUC
AD vs. NC	LDA	68.92±4.98	66.67±7.90	64.71±5.19	73.88±5.90
	MLP	61.89±2.62	53.90±6.28	68.02±4.84	70.09±2.66
	GCN	69.46±5.96	62.42±11.27	67.47±8.87	73.91±4.90
	GAT	78.38±5.27	75.76±9.77	75.81±5.11	84.26±4.75
	SACGCN	85.68±4.95	82.14±9.88	89.27±2.97	92.72±3.59
	Ours	85.95±3.15	83.03±11.59	85.81±4.76	93.05±3.06
LMCI vs. NC	LDA	57.07±2.67	52.77±6.19	51.41±2.72	61.67±4.21
	MLP	57.31±4.33	45.08±8.83	66.80±8.11	57.50±6.37
	GCN	59.74±6.65	52.59±8.71	54.00±8.11	62.27±4.35
	GAT	64.28±5.20	63.82±6.28	65.21±7.00	66.02±5.08
	SACGCN	65.14±4.14	61.29±6.80	68.57±5.30	66.87±2.52
	Ours	67.99±4.14	70.80±5.52	65.10±2.49	71.18±6.26
EMCI vs. NC	LDA	57.54±5.15	58.62±7.44	55.20±5.28	58.93±3.25
	MLP	53.46±5.71	52.50±8.60	54.54±7.21	57.59±5.87
	GCN	55.50±6.01	51.07±7.94	53.94±6.79	56.09±5.23
	GAT	64.18±4.25	67.70±5.33	61.36±4.27	68.78±2.58
	SACGCN	63.45±5.55	60.75±6.74	65.98±8.78	71.17±4.26
	Ours	67.51±4.59	67.20±4.01	66.02±5.49	74.46±4.93
EMCI vs. AD	LDA	60.43±7.25	56.36±7.81	58.65±8.66	67.98±5.21
	MLP	65.81±2.02	59.70±2.87	72.67±4.18	73.69±4.09
	GCN	71.80±4.96	65.45±6.53	71.91±5.29	77.57±4.47
	GAT	79.50±4.69	75.76±9.58	79.63±3.71	87.60±3.86
	SACGCN	81.21±5.46	76.40±3.16	86.25±9.24	90.34±2.33
	Ours	83.20±5.17	81.82±10.67	82.82±6.05	91.29±3.61

IV. CONCLUSION

This study constructs two graphs to use imaging data, genetic information, and primary patient demographic data. We employ two GCN to extract the respective embedded information from each graph. Subsequently, we adopt an attention mechanism to integrate multimodal information to accomplish the classification task for early AD. The model is validated on the ADNI dataset, demonstrating the effectiveness of our approach. However, there are limitations to this work: we should have considered specific modalities of data, such as proteins and scales from AD patients. Therefore, in future work, we aim to integrate a broader range of modalities for more accurate AD diagnosis.

REFERENCES

[1] P. Scheltens *et al.*, "Alzheimer's Disease," *The Lancet*, vol. 397, no. 10284, pp. 1577-1590, 2021.
 [2] G. Livingston *et al.*, "Dementia Prevention, Intervention, and Care: 2020 Report of The Lancet Commission," *The Lancet*, vol. 396, no. 10248, pp. 413-446, 2020.

[3] X. Song *et al.*, "Graph Convolution Network with Similarity Awareness and Adaptive Calibration for Disease-Induced Deterioration Prediction," *Medical Image Analysis*, vol. 69, p. 101947, 2021.
 [4] Y. Kim *et al.*, "Multimodal Phenotyping of Alzheimer's Disease with Longitudinal Magnetic Resonance Imaging and Cognitive Function Data," *Scientific Reports*, vol. 10, no. 1, p. 5527, 2020.
 [5] M. Yu, O. Sporns, and A. J. Saykin, "The Human Connectome in Alzheimer Disease —Relationship to Biomarkers and Genetics," *Nature Reviews Neurology*, vol. 17, no. 9, pp. 545-563, 2021.
 [6] T. Wang, X. Chen, J. Zhang, Q. Feng, and M. Huang, "Deep Multimodality-Disentangled Association analysis Network for Imaging Genetics in Neurodegenerative Diseases," *Medical Image Analysis*, vol. 88, p. 102842, 2023.
 [7] M. Wang, W. Shao, S. Huang, and D. Zhang, "Hypergraph-Regularized Multimodal Learning by Graph Diffusion for Imaging Genetics Based Alzheimer's Disease Diagnosis," *Medical Image Analysis*, vol. 89, p. 102883, 2023.
 [8] T. Tong, R. Wolz, Q. Gao, R. Guerrero, J. V. Hajnal, and D. Rueckert, "Multiple Instance Learning for Classification of Dementia in Brain MRI," *Medical Image Analysis*, vol. 18, no. 5, pp. 808-818, 2014.
 [9] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," in *IEEE Conference on Computer Vision & Pattern Recognition (CVPR 2016)*, 2016, pp. 1-12.
 [10] X. Song *et al.*, "Multicenter and Multichannel Pooling GCN for Early AD Diagnosis Based on Dual-Modality Fused Brain Network," *IEEE Transactions on Medical Imaging*, vol. 42, no. 2, pp. 354-367, 2023.
 [11] X.-A. Bi *et al.*, "Community Graph Convolution Neural Network for Alzheimer's Disease Classification and Pathogenetic Factors Identification," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
 [12] Q. Zhu, H. Wang, B. Xu, Z. Zhang, W. Shao, and D. Zhang, "Multimodal Triplet Attention Network for Brain Disease Diagnosis," *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3884-3894, 2022.
 [13] W. Zhu, L. Sun, J. Huang, L. Han, and D. Zhang, "Dual Attention Multi-Instance Deep Learning for Alzheimer's Disease Diagnosis With Structural MRI," *IEEE Transactions on Medical Imaging*, vol. 40, no. 9, pp. 2354-2366, 2021.
 [14] S. G. Mueller *et al.*, "The Alzheimer's Disease Neuroimaging Initiative," *Neuroimaging Clinics*, vol. 15, no. 4, pp. 869-877, 2005.
 [15] G. Flandin and K. J. Friston, "Statistical Parametric Mapping (SPM)," *Scholarpedia*, vol. 3, no. 4, 2008.
 [16] R. Seiger, S. Ganger, G. S. Kranz, A. Hahn, and R. Lanzenberger, "Cortical Thickness Estimations of FreeSurfer and The CAT12 Toolbox in Patients with Alzheimer's Disease and Healthy Controls," *Journal of Neuroimaging*, vol. 28, no. 5, pp. 515-523, 2018.
 [17] E. T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng, and M. Joliot, "Automated Anatomical Labelling Atlas 3," *NeuroImage*, vol. 206, p. 116189, 2020.
 [18] Z. Qiu *et al.*, "3D Multimodal Fusion Network with Disease-induced Joint Learning for Early Alzheimer's Disease Diagnosis," *IEEE Transactions on Medical Imaging*, 2024.
 [19] S. Purcell *et al.*, "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *The American journal of human genetics*, vol. 81, no. 3, pp. 559-575, 2007.
 [20] P. Xanthopoulos, P. M. Pardalos, T. B. Trafalis, P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, "Linear Discriminant Analysis," *Robust data mining*, pp. 27-33, 2013.
 [21] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer Perceptron and Neural Networks," *WSEAS Transactions on Circuits and Systems*, vol. 8, no. 7, pp. 579-588, 2009.
 [22] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph Attention Networks," *stat*, vol. 1050, no. 20, pp. 10-48550, 2017.