UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

MAM-GAN: Multimodal association modeling based on generative adversarial networks for Alzheimer's disease diagnosis

Permalink

https://escholarship.org/uc/item/3ds7k9bs

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors

Tang, Binsong
Tian, Yin

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

MAM-GAN: Multimodal association modeling based on generative adversarial networks for Alzheimer's disease diagnosis

Binsong Tang[1], Yin Tian[1,2,*]
1.School of Life Health Information Science and Engineering, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China
2.School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China
* Corresponding author: tianyin@cqupt.edu.cn

Abstract

Alzheimer's disease (AD) is a highly heritable neurodegenerative disease, and brain imaging genetics (BIG) has become a key area for understanding its pathogenesis. However, existing methods often ignore the complex interrelationships between the multiple factors that lead to AD, especially when exploring the intrinsic connection between brain imaging features and gene variation. To address this challenge, we proposed a multimodal association modeling framework (MAM-GAN) based on generative adversarial networks, which aims to deeply reveal the association between genes and brain imaging features and apply it to disease state prediction. To verify the effectiveness of the framework, we conducted experiments using public datasets, and the results showed that MAM-GAN performed well in two classification tasks and successfully identified biomarkers closely related to AD.

Keywords: Alzheimer's disease; Brain Imaging Genetics; Generative Adversarial Networks; Multimodal Association Modeling

Introduction

As a complex neurodegenerative disease, the pathological mechanism of Alzheimer's disease (AD) has always been a core problem in medical research (Liss et al., 2021; Du et al., 2021). Despite the fact that no specific cause or specific treatment for AD has been proven to be effective, many studies have shown that early intervention in its early stages can delay disease progression in clinical practice (Cummings et al., 2022; Hampel et al., 2018). Therefore, elucidating the biological and physiological mechanisms of AD progression is a key factor in effectively controlling AD. However, most current studies focus on revealing damage to specific brain regions through neuroimaging methods (Li et al., 2021; Sarli et al., 2021) or identifying genetic variants through genotyping and sequencing (Choobdar et al., 2019; Wightman et al., 2021). This single-modality study is difficult to fully reveal the complex pathological mechanisms of Alzheimer's disease, because AD involves the interaction of multiple biological levels and factors.

With the development of the field of brain imaging genetics (BIG), the fusion of multimodal data has become an important direction for studying AD (Lorenzi et al., 2018; Ghosal et al., 2022). Du et al. (2020) proposed a multi-task sparse canonical correlation analysis (MTSCCALR) to detect genetic associations with imaging phenotypes in AD. However, they did not consider nonlinear relationships and did not provide intrinsic explanations for imaging genetic biomarkers. Zuo et al. (2022) used a specific autoencoder to identify consistent imaging genomic biomarkers. However, the autoencoder-based method does not consider the information of the interactions between a large number of gene variants that interact through various biological processes. Ghosal et al. (2021) designed an interpretable genetic and imaging graph neural network (GUIDE) based on the gene ontology level for schizophrenia analysis. However, GUIDE ignores the potential ability of genetic information to represent the region of interest in neuroimaging. In addition, most methods use relatively simple ways to jointly analyze phenotypic and genotypic data, such as connecting features, rather than learning the complex distribution relationship of data (Ning et al., 2018; Venugopalan et al., 2021).

Although some studies in recent years have attempted to use generative adversarial networks to explore deep interactions between modalities and data generation, these studies have mostly focused on image generation and data augmentation, with less attention paid to how to incorporate genetic information into the adversarial learning process, resulting in the potential of genetic factors in brain image generation and classification not being fully explored.

In summary, the purpose of most methods is to establish a multimodal deep learning framework rather than specifically obtain new insights into the relationship between them. They still have some limitations, especially in capturing complex associations between modalities and achieving efficient information fusion.

To overcome the limitations of existing methods, this paper proposes an innovative framework MAM-GAN, which combines a data-driven feature selection strategy based on prior knowledge and effectively models the relationship between neuroimaging and genetic data through adversarial training, accurately capturing the complex nonlinear association between the two. Our contributions are as follows:

1) In terms of extracting gene network features, a graph convolution-based module combined with edge enhancement and self-attention mechanism (GCN-EESA) is proposed, which can effectively capture long-range dependencies in gene networks and promote deep interactions between genes;

2) In multiple classification tasks on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, our framework shows significant advantages, further verifying its effectiveness;

3) Through a series of analyses, we reveal the nonlinear

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

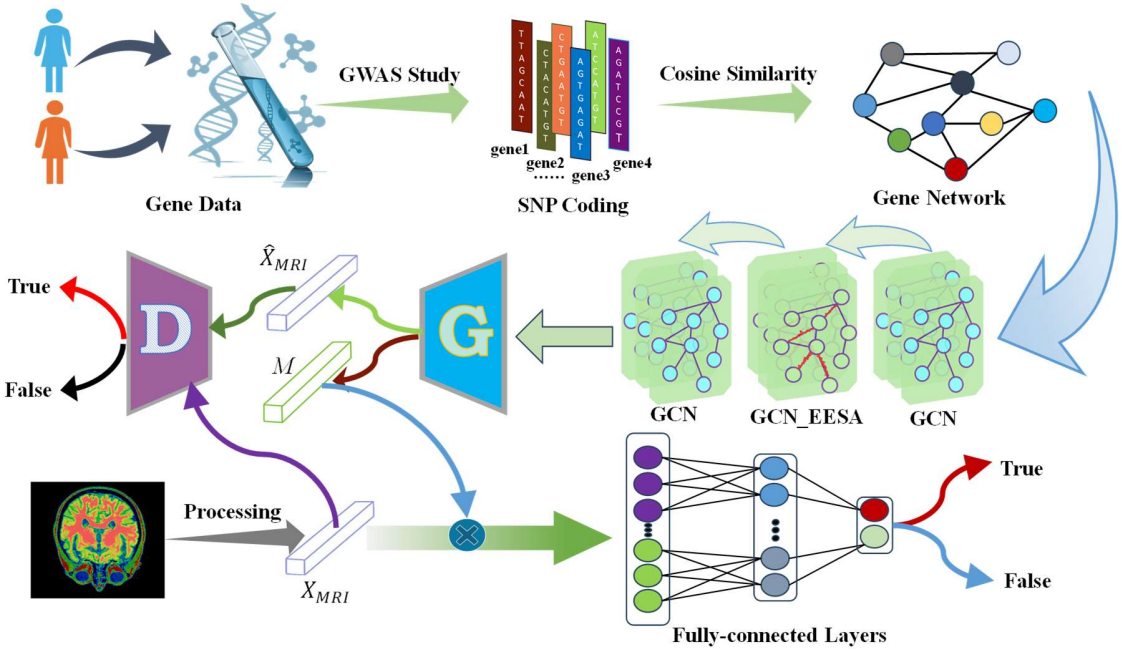


Figure 1: MAM-GAN Framework: A Deep Learning Approach for Multimodal Integration of Genetic and Brain Imaging Data for Disease Diagnosis.

relationship between the neuroimaging features learned by the framework and genetic features, and identify potential imaging biomarkers driven by genetic information;

Method

Model Framework

This study proposes a deep learning-based multimodal association modeling framework (MAM-GAN) to jointly represent the association between genetic information and neuroimaging. The framework includes a gene embedding module, an association module, and a diagnosis module, and its structure is shown in Figure 1.

The gene embedding module effectively captures the complex interactions between nodes in the gene network by combining the GCN with the self-attention mechanism, generating high-quality gene embedding vectors. The generated embedding vector is used as a latent feature input generator to guide the subsequent generation process. The association module uses these latent features to construct an attention mask matrix and continuously optimizes the attention area of the mask through generative adversarial training. To further ensure that the mask focuses on key association areas, the module introduces sparse regularization constraints to strengthen the model's attention to important features. The diagnostic module acts on the neuroimaging data based on the attention matrix output from the correlation module, thus efficiently realizing the classification of diseases and providing accurate disease prediction for clinical decision-making.

The framework adopts a joint loss function to combine the adversarial loss, sparse regularization loss, and classification loss, thereby ensuring the effective fusion of cross-modal information and improving the performance and stability of the model in classification tasks.

Gene Embedding Module The gene network constructed based on genetic information can be represented by a graph structure, where each gene is a node $\boldsymbol{v_p}\,(\boldsymbol{p}=\boldsymbol{1},\ldots,\boldsymbol{n})$, and the edge $\boldsymbol{E}=\{\boldsymbol{a_{p,q}}\mid(\boldsymbol{v_p},\boldsymbol{v_q})\in\boldsymbol{V}\times\boldsymbol{V}\}$ represents the genetic similarity between genes. The adjacency matrix $\boldsymbol{A}\in\boldsymbol{R^{n\times n}}$ of the graph uses the cosine similarity between genes as weight to quantify the strength of the edge. The gene feature matrix $\boldsymbol{H}=[\boldsymbol{h_1},\ldots,\boldsymbol{h_n}]\in\boldsymbol{R^{n\times k}}$ describes the feature vector $\boldsymbol{h_p}$ of each node $\boldsymbol{v_p}$, which consists of SNP codes (0, 1, 2).

In the topological processing of the gene network, we use the graph convolutional network (GCN) to aggregate the features of neighboring nodes through convolution operations, thereby updating the features of each node. The specific formula is:

$$\boldsymbol{H^{(\ell+1)}}=f_e\left(\tilde{\boldsymbol{A}}\boldsymbol{H^{\ell}}\boldsymbol{W^{\ell}}\right) \tag{1}$$

Where $\tilde{\boldsymbol{A}}=\boldsymbol{D^{-\frac{1}{2}}AD^{-\frac{1}{2}}}+\boldsymbol{I}$ is the normalized adjacency matrix, H^{ℓ} is the node feature matrix of the ℓ th layer, $\boldsymbol{W^{(\ell)}}$ is the weight matrix of this layer, f_e is the activation function, and $\boldsymbol{H^{\ell+1}}$ is the updated node feature.

However, although GCN can expand its ability to capture long-range dependencies in the network by stacking more layers, GCN may encounter the problem of over-smoothing when stacking deep layers, causing the representation of nodes to tend to be similar, thereby reducing the model's discrimination. Therefore, in order to better capture long-

range dependencies in gene networks, we propose the GCN-EESA module. Specifically, the feature update formula of node p in the ℓ layer is:

$$\boldsymbol{h}_p^{\ell} = O_h^{\ell}||_{h=1}^{H}\left(\sum_{q\in\mathcal{N}_p} \boldsymbol{w}_{pq}^{h,\ell}\left(\boldsymbol{V}^{h,\ell}\boldsymbol{h}_q^{\ell}\right)\right) + \boldsymbol{h}_p^{\ell} \quad (2)$$

where $\boldsymbol{h}_p^{\ell}$ represents the feature of node p in the ℓ layer, $\boldsymbol{H}$ represents the number of attention heads, $||_{h=1}^{H}$ represents the concatenation operation of multi-head attention output, $\mathcal{N}_p$ is the neighbor node set of node p, $\boldsymbol{w}_{pq}^{h,\ell}$ is the attention weight between nodes p and q, $\boldsymbol{V}^{h,\ell}$ is the value matrix, and O_h^{ℓ} is the linear projection operation of the ℓ layer.

The edge enhancement mechanism further adjusts the edge weight by calculating the edge embedding feature $\boldsymbol{E}_{pq}^{(l)}$, the formula is:

$$\boldsymbol{w}_{pq}^{h,\ell} = softmax\left(\frac{\boldsymbol{Q}_p^{h,\ell}\cdot(\boldsymbol{K}_q^{h,\ell})^T}{\sqrt{d_\ell}} + \alpha\cdot\boldsymbol{E}_{pq}^{h,\ell}\right) \quad (3)$$

$$\boldsymbol{E}_{pq}^{h,\ell} = \sigma\left(concat\left(\boldsymbol{h}_p^{(l)};\boldsymbol{h}_q^{(l)}\right)\boldsymbol{W}_e\right) \quad (4)$$

where $\boldsymbol{Q}_p^{h,\ell}$ and $\boldsymbol{K}_q^{h,\ell}$ are the query vector and key vector of node p and node q respectively, α is the hyperparameter of the edge feature, $concat\left(\boldsymbol{h}_p^{(l)};\boldsymbol{h}_q^{(l)}\right)$ is the feature concatenation of node p and node q, $\boldsymbol{W}_e$ is the linear transformation matrix, and $\boldsymbol{E}_{pq}^{h,\ell}$ is the edge embedding feature.

Then, we introduce multi-layer GCN in the last layer h_p^L of GCN-EESA to further capture the local topological characteristics of the gene network.

$$\boldsymbol{h}_p^{g} = \sigma\left(\sum_{q\in\mathcal{N}_p}\frac{1}{\sqrt{deg(p)deg(q)}}\boldsymbol{W}\boldsymbol{h}_p^{L}\right) \quad (5)$$

Where, $\boldsymbol{h}_p^L$ is the node feature output by the last layer of GCN-EESA. q is the neighbor of node p, $deg(p)$ is the degree of node p, and $\boldsymbol{W}$ is the learnable weight matrix.

Finally, we aggregate the node features into a global feature vector via global pooling:

$$gene_emb = \frac{1}{|\boldsymbol{V}|}\sum_{p\in V}\boldsymbol{h}_p^{g} \quad (6)$$

Where: $\boldsymbol{V}$ is the node set.

Compared with traditional graph convolutional networks, this method introduces edge enhancement and self-attention mechanisms, which effectively improves the influence of edge features on node updates and enhances the representation ability of the graph, especially in capturing the spatial associations and complex relationships of gene networks. In this way, the model provides effective prior constraints for downstream tasks and prepares the generator for reasoning about the attention matrix.

Association Module In this module, our framework generates latent distribution data through gene network embedding vectors to guide the generation of mask matrices and further learn the association between neuroimaging data and gene data. Given the high complexity of neuroimaging data, it is still challenging to generate neuroimaging-assisted attention matrices relying solely on genetic information. In order to better learn the joint representation of neuroimaging and genetic data, we design a strategy that enables the generator to implicitly learn the distribution information of neuroimaging data. Specifically, we introduce the Generative Adversarial Network (GAN) (Goodfellow et al., 2014), which is widely used in generation tasks because it can effectively estimate the true distribution of data.

We define a generator network ($\boldsymbol{G}$) and a discriminator network ($\boldsymbol{D}$). The generator G accepts the gene vector $\boldsymbol{C}_{gene}$ as input and generates an attention mask $\boldsymbol{M}\in R^{F_{MRI}}$ and fake brain image data $\widehat{\boldsymbol{X}}_{MRI}\in R^{F_{MRI}}$. Where F_{MRI} represents the feature dimension of the brain image data. The discriminator $\boldsymbol{D}$ accepts real and fake brain image data and distinguishes them.

In order to make the generated attention mask focus on key brain areas, we introduce a sparse regularization term to constrain the sparsity of the mask by applying a $\ell 1$ norm penalty to the elements of the mask $\boldsymbol{M}$. This constraint forces the model to focus on areas that have a greater impact on brain image features and reduce interference from irrelevant areas, thereby improving biological significance. The introduction of sparse regularization enables the generator to more accurately generate brain image features that reflect the impact of genetic variation.

In the association module, the training objective loss function combines the GAN loss and the $\ell 1$ regularization term as follows:

$$L_{GAN} = L_{\boldsymbol{D}} + L_{\boldsymbol{G}} + \lambda\left|\left|\boldsymbol{M}\right|\right|_1 \quad (7)$$

Where λ is the weight of the sparse regularization. The loss term is defined as:

$$L_{\boldsymbol{D}} = -E[logD(\boldsymbol{X}_{MRI})] - E\left[log\left(1-\boldsymbol{D}\left(\widehat{\boldsymbol{X}}_{MRI}\right)\right)\right] \quad (8)$$

$$L_{\boldsymbol{G}} = -E\left[logD\left(\widehat{\boldsymbol{X}}_{MRI}\right)\right] \quad (9)$$

During the training process, the generator generates brain imaging data that is consistent with gene embedding consistency by minimizing $L_{\boldsymbol{G}}$ and $\lambda\left|\left|\boldsymbol{M}\right|\right|_1$, while focusing on key areas; the discriminator distinguishes between real and fake images by minimizing $L_{\boldsymbol{D}}$. This adversarial training can enhance the model's ability to learn the biological significance of brain imaging features, ensure that the mask focuses on brain areas that contribute to disease classification, and be used for downstream diagnostic tasks.

Diagnostic Module The diagnosis module is used for downstream diagnosis tasks of brain diseases. The input neuroimaging data $\boldsymbol{X}_{\boldsymbol{MRI}}$ is modulated by the gene-guided attention mask $\boldsymbol{M}$. It is sent to the classification network $\boldsymbol{C}$ for disease diagnosis. Specifically, we perform Hamiltonian

product on the generated attention mask $\boldsymbol{M}$ and the real brain imaging data $\boldsymbol{X_{MRI}}$ to obtain the representation $\hat{\boldsymbol{y}}$ for brain disease diagnosis:

$$\hat{\boldsymbol{y}} = \boldsymbol{C}\big((\boldsymbol{X}_{MRI} \odot \boldsymbol{M})\big) \quad (10)$$

Where $\odot$ represents Hamiltonian product.

The diagnosis network consists of multiple independent feedforward layers. First, the intermediate feature vector $f(i)$ is extracted by the following formula:

$$\boldsymbol{f}(i) = \sigma\big((\boldsymbol{X}_{MRI} \odot \boldsymbol{M})\boldsymbol{W} + b\big) \quad (11)$$

Where σ is a nonlinear activation function, $\boldsymbol{W} \in R^{F_{MRI} \times df}$ is a weight matrix, and b is a bias term.

In order to perform the classification task, we define a feedforward layer to output the prediction results of brain diseases:

$$\hat{\boldsymbol{y}} = \sigma_c(f(i)\boldsymbol{W_c} + b_c) \quad (12)$$

Where $\boldsymbol{W_c} \in R^{df \times 1}$ and b_c are the classification weight matrix and bias term respectively.

In order to deal with the problem of class imbalance, we use focal cross entropy (FCE) as the classification loss function. FCE reduces the influence of easy-to-classify samples by adjusting the weights of difficult-to-classify samples. For a given model's predicted result $\hat{y}$ and true label y, the FCE loss function is defined as:

$$L_{\boldsymbol{C}} = -E[\alpha(1 - P_t)^{\gamma} log P_t]$$

Where,

$$P_t = P(y \mid \boldsymbol{X}_{MRI} \odot \boldsymbol{M}) \quad (13)$$

FCE loss evaluates the model training effect by comparing the predicted labels and the true labels.

Experiment

Dataset

All subjects in this paper were from the ADNI database (Mueller et al., 2005), including 179 CN, 191 MCI and 96 AD, a total of 466. All subjects had matched structural magnetic resonance imaging (SMRI) and genetic data, and detailed statistical information is shown in Table I.

Table 1: Subjects information

Disease	CN	MCI	AD
Number	179	191	96
Age(mean)	74.02	71.82	75.25
Gender(M/F)	91/88	106/85	53/43
Education(mean)	16.42	15.88	15.25

Data Preprocessing and Feature Extraction

After preprocessing the structural MRI (sMRI) scans of all subjects using the recon-all command of FreeSurfer (version 7.2.0) (Fischl, 2012), 308 cortical regions were segmented based on the Desikan-Killiany atlas (Romero-Garcia et al . , 2012). Next, we extracted gray matter volume and cortical thickness from the brain of each subject. These morphological indicators were used as brain region features for the input of brain imaging data.

In the gene data processing step, we used the PLINK tool (Purcell et al., 2007) for quality control, and the screening conditions included missing rate, minimum allele frequency, Hardy-Weinberg equilibrium test significance, and linkage disequilibrium. To reduce the high-dimensional redundancy of SNPs, we screened out significant SNPs through genome-wide association analysis (GWAS) and mapped them to the corresponding gene ontology. After screening, 132 related genes were identified.

In addition, to further enrich the gene ontology information, we also included other SNPs nearby, giving priority to sites located in introns and coding regions. These SNPs jointly characterize the gene ontology characteristics represented by the significant SNPs. All SNPs are coded with {0, 1, 2}, representing homozygous common genotypes, heterozygous genotypes, and homozygous minor genotypes, respectively, to more comprehensively characterize the gene ontology characteristics.

Experiment Settings

The experiment uses five-fold cross-validation to ensure the reliability of the performance evaluation. During each iteration, four subsets are used for training and the remaining one subset is used for validation and testing. All experiments were conducted under the same conditions and with the same hyper-parameter settings to guarantee the fairness of the comparison. The Adam optimizer was chosen for model training in the experiments, specifically, the learning rate was set to 0.0001 and the weight decay parameter was set to 0.001. The optimal model was ultimately determined by minimizing the classification error in the validation set.

Evaluation Metrics

In order to verify the effectiveness of the model, the experiment adopted the evaluation indicators commonly used in classification tasks. We combined multiple indicators to comprehensively evaluate the performance of the model, including accuracy (ACC), area under the curve (AUC), sensitivity, and specificity, and the F1-Score.

Results

To verify the effectiveness of the model, we conducted comprehensive classification experiments on the ADNI dataset, including comparative experiments and ablation experiments.

Comparative Experiments

In this study, we compared the proposed framework with several classic representation methods and current mainstream deep learning representation methods (see Table 2). Baseline methods include linear discriminant analysis (LDA) (Sharma & Dey, 2021), canonical correlation analysis (CCA) (Hardoon et al., 2004), and multi-layer perceptron

Table2: Classification performance (%) of comparative experiments

Methods	CN.VS AD					CN.VS MCI				
	Accuracy	AUC	SEN	SPE	F1-score	Accuracy	AUC	SEN	SPE	F1-score
LDA	78.96	82.60	61.52	82.24	69.81	59.24	49.85	67.09	62.70	64.82
CCA	81.28	78.12	62.57	75.48	68.42	62.22	58.25	73.45	68.14	70.69
MLP-Concat	82.93	85.19	68.14	83.83	73.66	64.18	66.80	68.64	70.22	70.70
MLP-Fusion	85.14	87.89	74.27	84.46	77.99	72.10	69.35	72.84	73.52	73.72
SWDNN	86.15	88.68	73.84	83.69	76.55	72.25	76.30	69.68	72.24	72.20
G-MIND	82.74	85.58	78.24	86.68	80.55	73.60	73.58	74.35	77.75	75.11
IGNET	85.98	89.32	82.68	90.26	84.64	75.26	74.58	75.82	78.82	76.56
Ours	**92.59**	**94.93**	**87.50**	**95.24**	**91.20**	**77.32**	**79.52**	**77.82**	**80.64**	**78.98**

directly cascaded after feature extraction (MLP-Concat) (Zhou et al., 2019) and multi-layer perceptron with more advanced fusion strategies (MLP-Fusion) (Venugopalan et al., 2021). Deep learning methods include staged deep neural network (SWDNN) (Zhou, Thung, Zhu, & Shen, 2019), neural network framework for genetic and multimodal imaging data (G-MIND) (Ghosal et al., 2021), and imaging genetic deep neural network system (IGNET) (Wang et al., 2022).

Experimental results show that our proposed framework outperforms other methods in two classification tasks when inputting the same modal data. This highlights the effectiveness and robustness of the MAM-GAN methods in disease classification using multimodal data, and strongly demonstrates its effectiveness in modeling the association between neuroimaging and genetics.

Ablation Experiments

In ablation experiments, we remove different modules in the framework separately to evaluate their respective impact on the overall performance.

First, the gene embedding and gene-neuroimaging association modules are removed, and only neuroimaging data is used for diagnosis (Case I).

Second, GCN-EESA is removed, and a single GCN is used to process the original gene graph data to generate the attention mask (Case II). Finally, the gene-neuroimaging association module is not used, that is, after obtaining the guidance of gene prior information, the generator directly generates the attention mask matrix (Case III).

The results are shown in Table 3. Our proposed framework performs better in all ablation scenarios.

Table3: Classification accuracy (%) of ablation experiments

Methods	Metrics	CN.VS AD	CN.VS MCI
Case I	ACC	85.17	72.12
Case II	ACC	90.14	74.68
Case III	ACC	90.25	75.02
Proposed	**ACC**	**92.59**	**77.32**

By comparison, it can be seen that the performance of Case I is significantly lower than that of the complete model, indicating that genetic information is crucial in revealing pathological factors and providing rich information. For Case II, after losing GCN-EESA, the model cannot fully capture the spatial distribution and complex relationships of the gene network, and the prior guidance provided to the generator is significantly weakened. For Case III, in the absence of a discriminator, genetic information is difficult to effectively guide neuroimaging representation, resulting in the mask matrix being unable to fully learn deep cross-modal associations.

Based on the above results, we conclude that the proposed overall architecture design and learning strategy are reasonable and effective, and each module in the framework plays a key role in parameter optimization under limited sample conditions.

Discussion

Identification of Risk Gene Associated With AD

This section aims to evaluate the effectiveness of our framework in identifying genes with significant representational power. To this end, we systematically analyzed the global node attention weighted average gene scores generated after processing by the GCN-EESA module at the population level. In order to verify whether the key genes identified by our framework are closely related to the progression of neurodegenerative diseases, we took the CN and AD classification task on the test set as an example, normalized the gene scores, and displayed the top 25 genes, as shown in Figure 2.

Among them, APOE, TOMM40, and APOC1 have been widely identified as core risk genes for AD (Corder et al., 1993; Roses et al., 2010; Zhou et al., 2014), while genetic variation in CSMD1 suggests that it may play an important role in neuronal synaptic plasticity and immune regulation（Jansen et al., 2022）.

In the study of the molecular mechanism of AD, PAK2 and AKT3 have been proposed as potential regulatory genes. PAK2 promotes abnormal phosphorylation of tau protein,

exacerbating the formation of neurofibrillary tangles (Chen et al., 2023), while AKT3 may affect the balance between neuronal survival and apoptosis by regulating the PI3K/AKT signaling pathway (Li et al., 2022). Notably, epigenetic studies have shown that dysregulated expression of the imprinted gene SNRPN may be associated with neurodevelopmental defects in early-onset AD（Smith et al., 2022）. These AD-associated risk genes exemplify the ability of the proposed framework to enhance the diagnosis of brain diseases by effectively representing data.

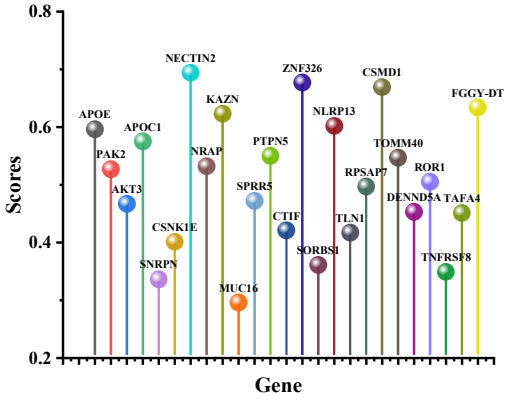


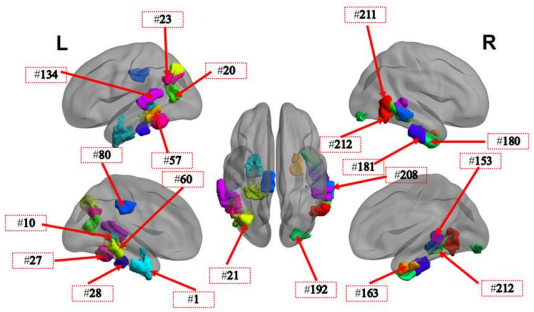
Figure 2: Importance scores of the top 25 risk genes.

Discriminative Brain Regions

In order to reveal which ROIs the proposed framework focuses on, we extracted the attention matrices of correctly classified samples in the CN and AD classification experiments and averaged the attention matrices of each type of sample at the group level. Subsequently, a statistical multivariate t-test (p-value < 0.01) was used to distinguish the key brain regions of CN and AD. We projected the top 20 significant brain regions onto the cortical surface (as shown in Figure 3(a)), and presented detailed information in descending order of brain region importance in Figure 3(b).

We found that regions such as the hippocampus, temporal lobe, cingulate gyrus, and entorhinal cortex showed higher attention levels in the CN-AD group (marked in red). These regions have been shown to be morphological biomarkers of Alzheimer's disease (AD) progression. In addition, the lateral occipital lobe and inferior parietal lobule also showed relevant changes. Studies have shown that the lateral occipital lobe is closely related to visual symptoms such as visual agnosia in AD patients and may be significantly affected in the late stage of the disease（Johnson et al., 2022）; damage to the inferior parietal lobule will aggravate AD patients' problems in spatial perception and sensory integration, which is usually associated with spatial disorientation and abnormal sensory information processing (Smith et al., 2023).

In summary, the framework we proposed can effectively identify brain regions closely related to AD pathology, providing strong support for the early diagnosis of AD and the study of pathological mechanisms.



(a)

Index	ROI	Index	ROI
#60	lh_parahippocampal_part2	#180	rh_inferiortemporal_part2
#163	rh_entorhinal_part1	#57	lh_middletemporal_part4
#21	lh_inferiorparietal_part6	#211	rh_middletemporal_part4
#20	lh_inferiorparietal_part5	#212	rh_middletemporal_part5
#192	rh_lateraloccipital_part7	#179	rh_inferiortemporal_part1
#1	lh_bankssts_part1	#10	lh_entorhinal_part1
#208	rh_middletemporal_part1	#80	lh_posteriorcingulate_part1
#28	lh_inferiortemporal_part5	#181	rh_inferiortemporal_part3
#153	rh_bankssts_part1	#134	lh_superiortemporal_part3
#27	lh_inferiortemporal_part4	#23	lh_inferiorparietal_part8

(b)

Figure 3: Spatial locations of the top-20 significant brain regions. (a) CN-AD significant regions. (b) List of top-20 significant brain regions from (a).

Conclusion

In this paper, we proposed a multimodal association framework (MAM-GAN) based on deep learning. This framework fully explores the potential of genetic factors in characterizing morphological indicators and associated lesion areas by promoting deep interactions between genes and neuroimaging features, thereby providing a new idea for achieving accurate diagnosis and identification of key pathogenic factors. Experimental results show that MAM-GAN has achieved significant performance improvements compared with existing models. In addition, considering the important information contained in other neuroimaging modalities (such as positron emission tomography, PET), future work can further integrate multimodal neuroimaging data to further improve the performance and application value of the model.

Acknowledgements

This work was supported by the Chongqing Natural Science Fund Innovation and Development Joint Fund (Municipal Education Commission) project under Grant 2024NSCQ-LZX0058; the National Natural Science Foundation of China under Grant 62171074, W2411084; and the funding of Institute for Advanced Sciences of Chongqing University of Posts and Telecommunications under Grant E011A2022327.

References

Liss, J. L., Seleri Assunção, S., Cummings, J., Atri, A., Geldmacher, D. S., Candela, S. F., ... & Sabbagh, M. N. (2021). Practical recommendations for timely, accurate diagnosis of symptomatic Alzheimer's disease (MCI and dementia) in primary care: a review and synthesis. *Journal of internal medicine*, *290*(2), 310-334.

Du, L., Zhang, J., Liu, F., Wang, H., Guo, L., Han, J., & Alzheimer's Disease Neuroimaging Initiative. (2021). Identifying associations among genomic, proteomic and imaging biomarkers via adaptive sparse multi-view canonical correlation analysis. *Medical Image Analysis*, *70*, 102003.

Cummings, J., Lee, G., Nahed, P., Kambar, M. E. Z. N., Zhong, K., Fonseca, J., & Taghva, K. (2022). Alzheimer's disease drug development pipeline: 2022. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *8*(1), e12295.

Hampel, H., Mesulam, M. M., Cuello, A. C., Farlow, M. R., Giacobini, E., Grossberg, G. T., ... & Khachaturian, Z. S. (2018). The cholinergic system in the pathophysiology and treatment of Alzheimer's disease. *Brain*, *141*(7), 1917-1933.

Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., ... & Duncan, J. S. (2021). Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, *74*, 102233.

Sarli, G., De Marco, M., Hallikainen, M., Soininen, H., Bruno, G., & Venneri, A. (2021). Regional strength of large-scale functional brain networks is associated with regional volumes in older adults and in Alzheimer's disease. *Brain Connectivity*, *11*(3), 201-212.

Choobdar, S., Ahsen, M. E., Crawford, J., Tomasoni, M., Fang, T., Lamparter, D., ... & Marbach, D. (2019). Assessment of network module identification across complex diseases. *Nature methods*, *16*(9), 843-852.

Wightman, D. P., Jansen, I. E., Savage, J. E., Shadrin, A. A., Bahrami, S., Holland, D., ... & Posthuma, D. (2021). A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. Nature genetics, 53(9), 1276-1282.

Lorenzi, M., Altmann, A., Gutman, B., Wray, S., Arber, C., Hibar, D. P., ... & Alzheimer's Disease Neuroimaging Initiative. (2018). Susceptibility of brain atrophy to TRIB3 in Alzheimer's disease, evidence from functional prioritization in imaging genetics. *Proceedings of the National Academy of Sciences*, *115*(12), 3162-3167.

Ghosal, S., Chen, Q., Pergola, G., Goldman, A. L., Ulrich, W., Weinberger, D. R., & Venkataraman, A. (2021). A biologically interpretable graph convolutional network to link genetic risk pathways and neuroimaging markers of disease. *bioRxiv*, 2021-05.

Bi, X. A., Chen, K., Jiang, S., Luo, S., Zhou, W., Xing, Z., ... & Liu, T. (2023). Community graph convolution neural network for alzheimer's disease classification and pathogenetic factors identification. *IEEE Transactions on Neural Networks and Learning Systems*.

Du, L., Liu, K., Yao, X., Risacher, S. L., Han, J., Saykin, A. J., ... & Shen, L. (2020). Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach. *Medical image analysis*, *61*, 101656.

Zuo, Y., Wu, Y., Lu, Z., Zhu, Q., Huang, K., Zhang, D., & Shao, W. (2022, September). Identify consistent imaging genomic biomarkers for characterizing the survival-associated interactions between tumor-infiltrating lymphocytes and tumors. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 222-231). Cham: Springer Nature Switzerland.

Ning, K., Chen, B., Sun, F., Hobel, Z., Zhao, L., Matloff, W., ... & Alzheimer's Disease Neuroimaging Initiative. (2018). Classifying Alzheimer's disease with brain imaging and genetic data using a neural network framework. *Neurobiology of aging*, *68*, 151-158.

Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific reports*, *11*(1), 3254.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems, 27*.

Fischl, B. (2012). FreeSurfer. *NeuroImage, 62*(2), 774–781.

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C., Jagust, W., ... & Beckett, L. (2005). The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America, 15*(4), 869–877.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics, 81*(3), 559–575.

Romero-Garcia, R., Atienza, M., Clemmensen, L. H., & Cantero, J. L. (2012). Effects of network resolution on topological properties of human neocortex. *NeuroImage, 59*(4), 3522–3532.

Hardoon, D. R., Szedmak, S., & Shawe-Taylor, J. (2004). Canonical correlation analysis: An overview with application to learning methods. *Neural Computation, 16*(12), 2639–2664.

Sharma, A., & Dey, P. (2021). A machine learning approach to unmask novel gene signatures and prediction of Alzheimer's disease within different brain regions. *Genomics, 113*(4), 1778–1789.

Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports, 11*(1), 3254.

Wang, J. X., Li, Y., Li, X., & Lu, Z. H. (2022). Alzheimer's disease classification through imaging genetic data with IGnet. *Frontiers in Neuroscience, 16*, 846638.

Zhou, T., Thung, K. H., Zhu, X., & Shen, D. (2019). Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Human Brain Mapping, 40*(3), 1001–1016.

Ghosal, S., Chen, Q., Pergola, G., Goldman, A. L., Ulrich, W., Berman, K. F., ... & Venkataraman, A. (2021, February). G-MIND: An end-to-end multimodal imaging-genetics framework for biomarker identification and disease classification. In *Medical Imaging 2021: Image Processing* (Vol. 11596, pp. 63–72). SPIE.

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G., ... & Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science, 261*(5123), 921–923.

Chen, X., Li, Y., & Zhang, J. (2023). PAK2-mediated tau phosphorylation in Alzheimer's disease: Mechanisms and therapeutic implications. *Journal of Neuroscience*, 43(5), 789-801.

Li, H., Wang, Z., & Liu, Y. (2022). AKT3 signaling in neuronal survival and death: Implications for Alzheimer's disease. *Cell Death & Disease*, 13(1), 1-12.

Jansen, I. E., Savage, J. E., Watanabe, K., Bryois, J., Williams, D. M., & Steinberg, S. (2022). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics*, 54(3), 412-436.

Roses, A. D., Lutz, M. W., Amrine-Madsen, H., Saunders, A. M., Crenshaw, D. G., Sundseth, S. S., ... & Reiman, E. M. (2010). A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *The Pharmacogenomics Journal, 10*(5), 375–384.

Zhou, Q., Zhao, F., Lv, Z. P., Zheng, C. G., Zheng, W. D., Sun, L., ... & Yang, Z. (2014). Association between APOC1 polymorphism and Alzheimer's disease: A case-control study and meta-analysis. *PLoS ONE, 9*(1), e87017.

Smith, A. R., Smith, R. G., & Lunnon, K. (2022). Epigenetic regulation of SNRPN in early-onset Alzheimer's disease: A potential link to neurodevelopmental deficits. *Epigenetics*, 17(3), 245-260.

Johnson, K. A., Fox, N. C., Sperling, R. A., & Klunk, W. E. (2022). Brain imaging in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine, 12*(3), a037218.

Smith, R., Wibom, M., Olsson, E., Hägerström, D., & Hansson, O. (2023). Spatial navigation deficits in preclinical Alzheimer's disease: A review of neuroimaging findings. *Neurobiology of Aging, 121*, 1-12.