BioData Mining

RESEARCH

Open Access

# Deep joint learning diagnosis of Alzheimer's disease based on multimodal feature fusion

Jingru Wang[1], Shipeng Wen[1], Wenjie Liu[2], Xianglian Meng[2,3*] and Zhuqing Jiao[1,3*]

*Correspondence:
mengxl@czust.edu.cn; jzq@cczu.edu.cn

[1] School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou 213164, China
[2] School of Computer Information and Engineering, Changzhou Institute of Technology, Changzhou 213032, China
[3] Wangzheng School of Microelectronics, Changzhou University, Changzhou 213164, China

## Abstract

Alzheimer's disease (AD) is an advanced and incurable neurodegenerative disease. Genetic variations are intrinsic etiological factors contributing to the abnormal expression of brain function and structure in AD patients. A new multimodal feature fusion called "magnetic resonance imaging (MRI)-*p* value" was proposed to construct 3D fusion images by introducing genes as a priori knowledge. Moreover, a new deep joint learning diagnostic model was constructed to fully learn images features. One branch trained a residual network (ResNet) to learn the features of local pathological regions. The other branch learned the position information of brain regions with different changes in the different categories of subjects' brains by introducing attention convolution, and then obtained the discriminative probability information from locations via convolution and global average pooling. The feature and position information of the two branches were linearly interacted to acquire the diagnostic basis for classifying the different categories of subjects. The diagnoses of AD and health control (HC), AD and mild cognitive impairment (MCI), HC and MCI were performed with data from the Alzheimer's Disease Neuroimaging Initiative (ADNI). The results showed that the proposed method achieved optimal results in AD-related diagnosis. The classification accuracy (ACC) and area under the curve (AUC) of the three experimental groups were 93.44% and 96.67%, 89.06% and 92%, and 84% and 81.84%, respectively. Moreover, a total of six novel genes were found to be significantly associated with AD, namely *NTM*, *MAML2*, *NAALADL2*, *FHIT*, *TMEM132D* and *PCSK5*, which provided new targets for the potential treatment of neurodegenerative diseases.

**Keywords:** Alzheimer's disease, Multimodal feature fusion, Deep joint learning diagnosis, Attention mechanism, ResNet

## Introduction

Alzheimer's disease (AD) is a long-term neurodegenerative disorder with complex causes [1]. The main clinical manifestations are memory loss, cognitive impairment, and changes of personality and behavior [2]. Brain tissue morphology manifests as severe atrophy of the hippocampus [3], cerebral cortex [4], and severe enlargement of the ventricles [3]. Since the pathogenesis of AD remains unknown, researchers have often explored AD with the help of its pathognomonic manifestations such as patients' clinical scores and brain images [5]. Despite advances in extensive research and clinical practice,

less than 50% of AD patients are accurately diagnosed via clinical symptoms alone [6]. Patients are not easily detected in the early stages of AD because its pathology occurs several years before the onset of clinical symptoms [7]. Mild cognitive impairment (MCI) is the preclinical stage of AD, which is a transitional state between normal aging and AD. There are no effective treatments to cure or completely halt the progression of AD [8]. Therefore, the early detection and intervention of MCI is of great significance in slowing the patient's condition and preventing future long-term treatment before irreversible brain damage occurs.

Over the last two decades, advances in neuroimaging technology have contributed greatly to the early diagnosis of AD [9]. It has been studied that the temporal lobe, hippocampus, and related regions of the patients' brains have undergone atrophy to varying degrees during the preclinical stage of AD [10]. Magnetic resonance imaging (MRI) can accurately reflect the anatomical structure of the brain of living subjects [11]. Vemuri [12] and Tanveer [13] et al. proposed automated diagnostics with computer assistance from structural MRI. This topic has attracted much attention as a promising area of study. Many studies have demonstrated that MRI technology is an effective adjunctive basis for detecting AD and other brain disorders. However, because single imaging data can only capture obvious differences in brain structure and function, MRI is not effective in facilitating the diagnosis or treatment of patients who do not exhibit these obvious structural changes [14]. Furthermore, MRI-based diagnosis of AD is limited to identifying brain regions associated with the condition, making it difficult to delve deeper into the keys to disease pathogenesis [15]. Thus, the early diagnosis of this disease based on MRI is subject to certain constraints.

Imaging genomics has become an emerging research field with the development of high-throughput histological data and multimodal imaging data [16]. Genome-wide association study (GWAS) is a greatly common research method in imaging genomics studies [17]. Genome-wide high-density genetic marker typing of large-scale population DNA samples is performed to mine genes associated with complex disease phenotypic traits through genome-wide association studies between genes and multiple phenotypes in multicenter, large-sample, and iteratively validated studies [18]. Over the past few decades, several GWASs have identified hundreds of single nucleotide polymorphisms (SNPs) associated with AD [19]. Through genetic linkage analysis, SNPs can help locate biomarkers and genetic risk factors associated with AD. Genetic data can provide new targets for AD diagnosis. The large volume of GWAS results presents a challenge in extracting pertinent genetic variation information and analyzing its association with imaging data to identify biomarkers related to AD [20].

Many studies have explored multimodal feature fusion from imaging data and genetic information. Bi [21] et al. proposed a feature construction method called "brain region-gene pair" to integrate resting-state functional magnetic resonance imaging (rs-fMRI) and SNPs. Hu[22] et al. obtained multimodal fusion features by fusing the voxels of MRIs and $p$-values. Although several feature selection approaches have been proposed, the fusion of imaging and genetic data still faces challenges. Traditional fusion approaches usually select features for each modality separately and then connect the selected features for disease diagnosis, which ignores the connection between the data of different modalities [23]. Nevertheless, not all subjects have complete multimodal data. Samples

with incomplete modes are often discarded in the past [24]. The diagnostic model may not be adequately trained or overfitted if the number of available subjects is too small [25]. Consequently, it is still necessary to further explore the new ways of multimodal feature fusion.

In recent years, deep learning algorithms have been able to adequately analyze the progression of AD by identifying complex changes in brain structure, showing high accuracy and stability in the field of medical diagnosis [26]. Compared with traditional computer-aided diagnostic tools, CNNs do not rely on predefined and handcrafted features and learn local features and diseases classification end-to-end, demonstrating significant effectiveness in medical imaging analysis tasks [27, 28]. Despite the success of CNNs for diagnosis, CNNs with too many deep layers can actually reduce classification accuracy to some extent [29].

Bakkouri [30] et al. introduced a computer-aided diagnosis (CAD) system called DermoNet, which was built upon Multi-Scale Feature Level (MSFL) blocks and Multi-Scale Feature Fusion (MLFF). This system effectively addressed problems related to high interclass similarity and overfitting, thereby improving diagnostic performance. Zhang[31] et al. proposed a 3D residual attention deep neural network (3D ResAttNet), which combined an attention mechanism with a residual neural network (ResNet) to capture the feature of images to obtain better diagnostic results. Bakkouri [32] et al. proposed an innovative multi-level multi-scale gated attentional squeezed network(2MGAS-Net) that effectively captures contextual information across multiple scales. Zhu[33] et al. located the positional information of brain regions to detect regions that showed significant differences between AD and HC by an attention mechanism.

To compensate for the limitations of single-modal MRI information, we introduce genes as prior knowledge and focus on identifying common genetic markers linked to the progression and mechanisms of AD. In addition, we aim to develop a novel feature fusion approach with the valuable information derived from GWAS to tackle the existing challenges in genetic variation and feature fusion, thereby providing a more comprehensive representation. Previous studies have proposed many deep learning methods for AD-related diagnosis, but developing diagnostic models with optimal performance is still the important goal of current research. This study proposes a deep joint learning diagnosis with the attention mechanism and ResNet, which aims to diagnose AD in its early stages with higher accuracy and identify significant biomarkers. The main contributions can be summarized as follows:

(a) A new multimodal feature fusion method called the "MRI-$p$ value" is proposed, which fully considers the intrinsic connection between two modalities. Not only the prior knowledge provided by genetic data is introduced without destroying the 3D MRIs structure, but also sample data are expanded to avoid overfitting problems, which helps the diagnostic model converge more quickly and reduces the need for learning samples.
(b) A new deep joint learning diagnostic model is constructed by combining the attention mechanism with ResNet, which can learn the feature and diagnostic information of the local pathological regions more comprehensively and improve the performance and stability of diagnosis.

**Table 1** The statistical table of subject information

| Subjects | HC | MCI | AD | p |
|---|---|---|---|---|
| Number | 180 | 198 | 125 | - |
| Gender (M/F) | 97/83 | 115/83 | 76/49 | < 0.001 |
| Age (mean ± Sd) | 75.0 ± 5.4 | 71.2 ± 7.1 | 74.7 ± 8.1 | < 0.001 |
| Edu (mean ± Sd) | 16.5 ± 2.6 | 16.0 ± 2.6 | 15.7 ± 2.6 | < 0.001 |

*HC* Healthy control, *MCI* early mild cognitive impairment, *AD* Alzheimer's disease, *M/F* Male/female, *Edu* Education, *sd* standard deviation; $p = p$-value calculated by the t-test

(c) The proposed method is extensively evaluated with several other popular methods in various aspects, including accuracy, precision, sensitivity and area under the curve.

## Materials

### Imaging and gene data

The data were provided by the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset.[1] Launched in 2003, the ADNI was a global, multicenter, long-term tracking research platform comprising four phases (ADNI-1, ADNI-go, ADNI-2 and ADNI-3). The platform provided MRI, positron emission tomography (PET) scans, genetic information, other biomarkers and relevant diagnostic information for a range of participants, with the primary goal of combining these data to actively work towards early detection and follow-up of AD. The samples were complete multimodal data, consisting of MRI and genetic data from 502 subjects. This specifically included 180 HC, 198 patients with MCI, and 125 patients with AD. Table 1 shows the details of these subjects.

### Data processing

Preprocessing operations were performed on the MRI scans with voxel-based morphometry (VBM). Initially, the MRI scans were subjected to operations such as format conversion, the removal of nonbrain tissues (such as neck and skull) from the MRIs, normalization, and segmentation. Subsequently, the acquired images were spatially aligned with the Montreal Neurological Institute (MNI) space, which eliminated the global linear differences of the images and resulted in images with the same spatial resolution (i.e., $1 \times 1 \times 1$mm3). Gray matter density (GMD) images were then extracted and smoothed by an 8 mm full width at half maxima (FWHM) kernel. Additionally, the acquired images were downsampled from $181 \times 218 \times 181$ to $91 \times 109 \times 91$ voxels to reduce the data processing costs. Accordingly, the size of preprocessed 3D brain images was (91,109,91). These images served as inputs of the imaging data for multimodal feature fusion. Moreover, the gray density images obtained after preprocessing were aligned onto the Anatomical Automatic Labeling (AAL) [34] template. Therefore, the voxels from the various local regions of the brain were acquired through the consistent encoding information of the brain regions. The average volumes of the left and right hippocampal regions were extracted as phenotypic data.

[1] http://adni.loni.usc.edu

Genetic data from the target subjects were genotyped by Illumina GWAS arrays (610-Quad v1.0, OmniExpress-24 Kit or HumanOmni2.5-4v1) and blood genomic DNA samples [35]. SNPs and subjects needed to be screened first to achieve complete and effective genetic data. The quality control of genotypic data consisted of six steps. These included screening subjects according to heterozygosity, the deletion rate of the locus, individual independence, and screening loci according to the deletion rate of locus information, Hardy–Weinberg equilibrium, and linkage disequilibrium. SNPs that satisfied the following conditions were extracted by Plink v1.9 [36]: (a) SNPs on chromosomes 1–22 were extracted; (b) the call rate of each SNP was $\geq 95\%$; (c) the estimated value of individual kinship was $\leq 0.2$; (d) the call rate of each participant was $\geq 95\%$; (e) the Hardy–Weinberg equilibrium test was $p \geq 1 \times 10^{-6}$; (f) the minor allele frequency of each SNP was $\geq 5\%$. Ultimately, a total of 563,980 SNPs passed quality control.

Genome-wide association studies were conducted with the genetic and phenotypic data of the target subjects to mine genetic biomarkers that might affect the volumes of brain regions. Principal component analysis was performed on the SNP data after quality control by Plink. The resulting eigenvector matrix provided the top 5 principal components of the population stratification analysis. Association analysis of the SNPs and phenotypic data was conducted by linear regression, in which the population stratification phenomenon was corrected by covariate information. Covariate information specifically included age, sex, education, disease labeling and the first 5 principal components acquired by the population stratification analysis. The results of a GWAS containing 563,980 SNPs were obtained.

The SNPs with $p < 0.05$ were screened. Univariate gene association analysis was then performed on the screened data with an effective chi-square test (ECS). Subsequently, multiple testing and Bonferroni correction were performed to locate genes which were significantly associated with AD. Finally, the corrected genes with $p < 0.001$ were selected data for the subsequent operations of feature fusion.

#### Multimodal feature fusion

The new multimodal feature fusion with the "MRI-$p$ value" was proposed to construct 3D fusion images without destroying the structure of 3D images. First, the preprocessed MRIs were denoted as $\boldsymbol{P_{MRI}} \in \mathrm{R}^{W \times H \times D \times 1}$, while the image features of different categories were defined as $\boldsymbol{P_{AD}}$, $\boldsymbol{P_{MCI}}$ and $\boldsymbol{P_{HC}} \in \mathrm{R}^{W \times H \times D \times 1}$, respectively. Second, the gene tensors were constructed with the genes of $p < 0.001$ after correction. The $-log_{10}(p)$ values were introduced as $\boldsymbol{L_p}$ to integrate the genetic data with significant features. Furthermore, a tensor $\boldsymbol{G_p} \in \mathrm{R}^{W \times H \times D \times 1}$ with the same size as $\boldsymbol{P_{MRI}}$ was constructed by $\boldsymbol{L_p}$, totaling 35 tensors. Each $\boldsymbol{G_p}$ was applied to $\boldsymbol{P_{AD}}$, $\boldsymbol{P_{MCI}}$, and $\boldsymbol{P_{HC}}$, respectively. The fusion features $\boldsymbol{S_1}$, $\boldsymbol{S_2}$ and $\boldsymbol{S_3} \in \mathrm{R}^{W \times H \times D \times 1}$ were calculated using the $\boldsymbol{P_{AD}}$, $\boldsymbol{P_{MCI}}$, $\boldsymbol{P_{HC}}$, and $\boldsymbol{G_p}$. The $\boldsymbol{S_1}$, $\boldsymbol{S_2}$, and $\boldsymbol{S_3}$ are defined as formula (1).

$$\begin{cases} S_1 = [P_{AD} \times G_P,\ P_{HC} \times G_P] \\ S_2 = [P_{AD} \times G_P,\ P_{MCI} \times G_P] \\ S_3 = [P_{HC} \times G_P,\ P_{MCI} \times G_P] \end{cases} \tag{1}$$

where $\boldsymbol{S_1}$, $\boldsymbol{S_2}$ and $\boldsymbol{S_3}$ denote the fusion feature tensors of the AD-HC group, AD-MCI group and MCI-HC group, respectively.

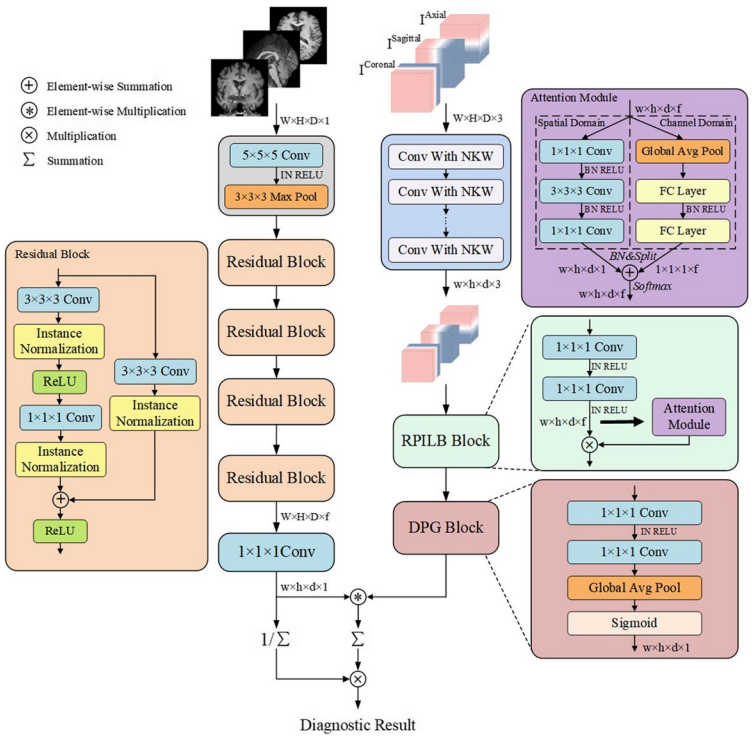


**Fig. 1** The graph of deep joint learning diagnosis

After fusion, the image feature of each subject was annotated with each gene tensor that was significantly associated with AD. The "MRI-*p* value" fusion not only established the link between the imaging data and gene information, but also greatly expanded the amount of training data for the experiment. Specifically, 35 tensors of the fused image were obtained from the image tensor of each subject after fusion. The overall dataset was enlarged by 35 times due to the introduction of genetic data. Therefore, the numbers of AD, HC and MCI fusion images were 4375, 6300 and 6930, respectively. The numbers of $\boldsymbol{S_1}$, $\boldsymbol{S_2}$, and $\boldsymbol{S_3}$ fused tensors were 10,675, 11,305, and 13,230, respectively. These were inputs for the three subsequent experiments. Each experimental dataset was randomly divided into training, validation, and test sets at a ratio of 6:2:2 for subsequent training.

## Methods

An end-to-end deep joint learning diagnosis was constructed with the combination of ResNet and an attention mechanism to automatically learn the feature and location information related to AD. On that basis, the significant features of AD were obtained via the proposed method. The trials included the three classifications of disease diagnosis with respect to AD and HC, AD and MCI, and MCI and HC.

Figure 1 shows the framework of the deep joint learning diagnosis. There were two main branches. The input of one branch was the whole brain image after multimodal fusion and ResNet [37] was trained to learn the features from the local pathological regions of subjects' brain images. The input of another branch was the location information of brain images acquired from a three-dimensional Cartesian space. Attention

convolution training was then applied to learn the location information of brain regions that showed varying changes across different subject categories. Subsequently, discriminant probability information from position was obtained via convolution and global average pooling. Finally, the characteristics of the local pathological regions and the discriminant probability information obtained from the two branches were linearly interacted to provide the diagnostic basis for classifying the different categories of subjects.

#### Feature extraction of pathological regions

Neural networks can effectively learn multilevel features through multilevel deep convolution operations [38]. ResNet is the most widely used CNN architecture and can effectively avoid the problem of gradients vanishing or exploding when the network becomes deeper. In this study, the residual block was optimized by ResNet18 to construct the framework [39], which obtained the features of local pathological regions in the whole brain image. The feature extraction module of pathological regions was mainly composed of the convolution layer, the maximum pooling layer and four residual blocks. They shared learning parameters in spatial dimensions. During training, the convolutional kernels and strides of each convolutional layer were continuously adjusted to capture the characteristics and spatial relationships of each pathological region. The experiments fully considered the effective combination of the information acquired by the two branches, so the parameters of the branch were regulated by controlling the receiving field size of the feature map in the last layer. Therefore, we obtained the features of the local pathological regions in the whole-brain image.

Specifically, the feature representation of the fused image was denoted as $\boldsymbol{X} \in \mathrm{R}^{W \times H \times D \times 1}$. In the first part of the branch, the kernel size and stride of the convolutional layer were set to (5,5,5) and (2,2,2), respectively. The normalization layer and nonlinear activation function were the instance normalization layer and the rectified linear unit (ReLU), respectively. Second, in the maximum pooling layer, the kernel size and stride were set to (3,3,3) and (2,2,2), respectively. After this part, the feature size was (32,19,24,19).

The second part learned richer features hierarchically by four residual blocks. The numbers of output feature mappings were 32, 64, 128, and 256 in the four residuals. At the first convolution layer passed of each branch, the kernel size and stride are (3,3,3,3) and (1,1,1,2), respectively. The normalization layer and activation function were instance normalization and ReLU, respectively. In the subsequent convolutional layers, the kernel size and stride were set to (1,1,1) and (1,1,1), respectively. After this training, the feature size of the images was (256,10,12,10). To obtain the more accurate features of pathological regions, the complexity of the feature was increased while maintaining a fixed feature size. After residual block training, $f$ features with dimensions ($w$, $h$, $d$) were obtained. With a $1 \times 1 \times 1$ convolution, the $f$-dimensional vector was transformed into local features from images, resulting in obtaining the feature $\boldsymbol{X} \in \mathrm{R}^{w \times h \times d \times 1}$ for the local pathological regions. Finally, the feature size was (1, 10, 12, 10).

**Capturing positional information**

The capture of location information in brain regions consisted of three main building blocks. Specifically, it consisted of the centre location extraction block (CLE Block), the relative position information learning block (RPILB Block) and the discriminative probability generation block (DPG Block).

In the transformation of image coordinate information, Liu et al. [40] made additional channels to input different transformed position information, which could facilitate the network to learn the more complete transformational invariance or different degrees of transformation dependence. This approach enhanced the ability of the diagnostic model to learn generalizable functions. Accordingly, the images' coordinate information of different transformations was connected in the form of channels.

Specifically, since all 3D images were aligned to the same MNI space and the position information of the images was not changed during the construction of multimodal fusion, the 3D fused images aligned to the same template shared a 3D Cartesian space. We obtained the coordinate information of the fused image in the coronal, sagittal, and axial planes. This gave three tensors $\boldsymbol{I} \in \mathrm{R}^{W \times H \times D \times 1}$. Second, using different channels to record the coordinates under different transformations, we got a tensor $\boldsymbol{I} \in \mathrm{R}^{W \times H \times D \times 3}$, which was the coordinate information that provided the complete position of the fused image in 3D Cartesian space. It was an input to the CLE block.

In the CLE block, the central location of the pathological regions was extracted hierarchically through multiple convolutional layers. Each layer was performed via deep convolution of non-parametric kernel weights, where the kernel weights were set to 0 and the center position weights were set to 1. The resulting central position information had dimensions of (3, 10, 12, 10) and served as an input to the relative position information learning block.

In the RPILB block, the attention module was introduced to learn the key regions through the weight tensor of the entire image. The attention mechanism was able to assign different attention weights to inputs from different locations to accurately capture the relevant information about each location. Although this avoided interference caused by some nonessential information, it also easily tended to lead to situations where global information was lost. To address this, superposition convolution was employed to detect positional information that differently varied across the different classes of brains. Therefore, the number of feature mappings was increased by superimposed convolutional layers. The most relevant positional information about the disease was acquired from both comprehensive and local perspectives via the combination of superposition convolution and an attention mechanism, providing the support of data for subsequent training. Finally, the feature size was (256,10,12,10).

The attention module of two branches was used to generate attention maps [41] that captured the information about the location of the brain where local changes occurred. Where the information $\boldsymbol{I} \in \mathrm{R}^{W \times H \times D \times f}$ obtained after superimposed convolution was the input to the attention module. One branch was to generate the attention map $W_s(\boldsymbol{I}) \in \mathrm{R}^{W \times H \times D \times 1}$ in the spatial domain. The specific definition is shown in formula (2):

$$W_s\ (I)\ =\ BN(Conv_3(Conv_2(Conv_1(I)))) \tag{2}$$

where $Conv_1$ and $Conv_2$ represent a $1 \times 1 \times 1$ convolutional layer and a $3 \times 3 \times 3$ convolutional layer, respectively. The number of channels was $C/r$, where $r$ was the reduction rate. The feature in the spatial domain was subsequently obtained. $Conv_3$ represents a $1 \times 1 \times 1$ convolutional layer with channel 1, which generates the attention map in the spatial domain. BN represents the batch normalization (BN) layer to ensure that the information of two independent attention branches is at the same scale. Another branch was to generate the attention map $W_c(\boldsymbol{I}) \in \mathrm{R}^{1\times1\times1\times f}$ in the spatial domain. The specific definition is shown in formula (3):

$$W_c\ (I)\ =\ BN(FCLayer_2(FCLayer_1(GlobalAvgPool(I)))) \tag{3}$$

where *GlobalAvgPool* represents the global average pooling (GAP) layer for feature aggregation in the spatial domain; $FCLayer_1$ and $FCLayer_2$ represent the two FC layers, which are feature perception and generation of the attention graph in the channel domain. Their channel numbers were $C/r$ and $C$, respectively. In the spatial domain and channel domain, the normalization layer and activation function were the batch normalization layer and the ReLU, respectively. The reduction rate $r$ was set to 4. In addition, the attention graphs generated by the two branches were combined with the *Softmax* function on the last dimension to obtain the final attention graph $W(\boldsymbol{I}) \in \mathrm{R}^{W\times H\times D\times f}$. The specific definition is shown in formula (4):

$$W(I)\ =\ Softmax\ (W_s(I)\ +\ W_c\ (I)) \tag{4}$$

$W(\boldsymbol{I})$ continued training as an input to the discriminant probability generation module. In the DPG block, firstly, the number of output feature maps was reduced via superposition convolution. The extracted location information was encoded before the selected information was normalized by the *sigmoid* activation function after GAP. Accordingly, the discriminative probability map from the location information was generated. This is expressed as $\boldsymbol{G}=\sum_{i,j,k=1}^{w,h,d} g_{i,j,k}$, where $g_{i,j,k}\in[0,1]$ represents the magnitude of the probability that the regions of the brain image ($i$, $j$, $k$) is associated with AD. Finally, the feature size was (1,10,12,10).

#### Classification of the diagnostic model

The resulting information of two branches were obtained respectively after training. In one branch, the feature about the local pathological regions of the fused images was obtained by trained ResNet, which was denoted as $\boldsymbol{X}\in\mathrm{R}^{w\times h\times d\times 1}$. In the other branch, attentional convolution was trained to capture information about the location of brain regions that were differentially varied across the different classes. The discriminant probability information from the position was denoted as $\boldsymbol{G}\in\mathrm{R}^{w\times h\times d\times 1}$. Inspired by the aggregation method, the diagnostic basis was defined by aggregating the characteristics of local pathological regions and the discriminant probability information [42]. Specifically, the element-by-element multiplication of the feature $\boldsymbol{X}$ with discriminative probability information $\boldsymbol{G}$ yielded the basis $\boldsymbol{E}\in\mathrm{R}^{w\times h\times d\times 1}$ of disease classification. Since the total number of discriminating brain regions was unknown, it was necessary to perform normalization operations according to the amount of discriminating probabilistic information. The specific definition is shown as formula 5:

$$y = \frac{\sum_{i,j,k=1}^{w,h,d} e_{i,j,k}}{\sum_{i,j,k=1}^{w,h,d} g_{i,j,k}} \tag{5}$$

where $e_{i,j,k}=g_{i,j,k}x_{i,\ j,k}$ and $y$ form the basis for identifying the classification of diseases after the normalization operation.

## Experiments and results

### Experimental settings

In this paper, the optimal parameters of the proposed framework were found through the optimization function. Experiments were conducted for five training rounds. The total epoch of each round was set to 200 and the batch was set to 4. Adam was taken as an optimizer [43] and the learning rate was adjusted by learning rate preheating and cosine annealing [44]. To be specific, the learning rate was increased from 0 to 1e-2 within the first 10 epochs of each round by the learning rate preheating scheduler. This value was subsequently reduced to 0 by the cosine annealing learning rate scheduler. The network was also optimized by cross-entropy for the loss function, which alleviated the problem of imbalance between positive and negative samples and learning from samples that were difficult to classify.

To reduce the difference in the fused images, the experiment enhanced the training data with operations such as image translation, random cropping, and random generation of the starting point of cropping for each subject image. The size of the processed images was (75, 93, 75). Image translation and center cropping were conducted on the test data. Moreover, the position information derived from the 3D Cartesian space was also transformed in the same way as the fused images. The obtained position information of the cropped images was taken for subsequent training.

### Evaluation metrics

In this study, we evaluated three binary classification tasks: AD and HC, AD and MCI, and HC and MCI by four common standard indicators. Specifically, these metrics included the classification accuracy (ACC), specificity (SPE), sensitivity (SEN), and area under the curve (AUC) of the receiver operating characteristic (ROC) curve. These metrics are defined as formula (6), (7) and (8):

$$ACC = \frac{T_P+T_N}{T_P+T_N+F_P+F_N} \tag{6}$$

$$SPE = \frac{T_N}{T_N+F_P} \tag{7}$$

$$SEN = \frac{T_P}{T_P+F_N} \tag{8}$$

where $T_P$, $F_P$, $T_N$ and $F_N$ are true positive, false positive, true negative and false negative, respectively. Constantly changing the execution threshold enabled to obtain all possible true positive rates (TPR = SEN) and false positive rates (FPR = 1-SPE). The AUC was calculated by the true and false positive rates under different thresholds and was taken to measure the classifier's ability to distinguish categories.

### Comparison with other methods

Previous studies mostly adopted deep learning frameworks with imaging features for disease diagnosis. Hence, this study made an extensive comparison between the most advanced methods and the proposed framework. These specifically included 3D CNN, XGA-ResNet18, Hybrid-MACN, and PG-BrainBagNet.

3D CNN [45]: An optimized neural network with 3D CNNs was developed to enable differentiation between subjects with normal cognition, MCI and AD.

XGA-ResNet18 [46]: This method proposed a new learning-explanation-reinforcement (LEAR) framework. It applied a three-level ResNet18 as the backbone diagnostic for disease classification and injected an explanation-guided attention (XGA) module that adaptively modulated the layer output to improve the performance and interpretability of the diagnostic.

Hybrid-MACN [41]: These authors proposed an AD diagnostic network with hybrid multiscale attentional convolution and aging transformers. First, the feature mappings learned from the multiscale kernel were adaptively aggregated by an attention module. Second, the correlations between features were modelled by pyramidal nonlocal blocks. Third, an aging transformer subnetwork embedded age into image features and learned potential dependencies between subjects of different ages. An end-to-end hybrid network was constructed for the diagnosis of AD.

PG-BrainBagNet [47]: They proposed an end-to-end joint learning framework for pathological regions localization and AD diagnosis. A total of two branches were included. The plaque-level prediction branch extracted local features from specific receptive field sizes and generated plaque-level responses based on MRIs. The other part identified brain regions associated with disease salience via position-gated BrainBagNet (PG-BrainBagNet). Ultimately, the results of the two branches were aggregated to generate categorical diagnostic evidence.

This study made the best effort to implement codes and frameworks provided in the literature [41, 45–47] to ensure a fairer comparison. All methods were trained, validated and tested with the same dataset, i.e., expanded the 3D fused images constructed after multimodal fusion. We calculated their ACC, SEN, SPE and AUC, respectively. Table 2 shows the results for three sets of diagnostic tasks.

The results of various evaluation parameters indicated that regardless of the group of classification tasks, the proposed framework performed much better than the other methods did. Additionally, owing to the differences in the subjects selected for the dataset, the results in Table 3 may differ from the results demonstrated in references. Through the results of comparative experiments, more powerful characterization was obtained via the combined diagnosis of the attention mechanism and the residual module.

### Effectiveness of multimodal feature fusion

The validity of multimodal fusion images was also verified with single-modal MRIs. Table 3 shows the performance of the three diagnostic tasks. The results showed that the diagnosis with multimodal fusion features was superior to the diagnosis with only MRI as input in all aspects. As a consequence, the proposed "MRI-$p$ value" fusion can

**Table 2** The performance of comparative experiments

| Method | AD vs. HC (%) | | | | AD vs. MCI (%) | | | | MCI vs. HC (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ACC | SEN | SPE | AUC | ACC | SEN | SPE | AUC | ACC | SEN | SPE | AUC |
| 3D CNN[45] | 85.25 | 76 | 91.67 | 91.67 | 71.88 | 64 | 76.92 | 79.59 | 73.33 | 71.79 | 75 | 78.28 |
| XGA-ResNet18[46] | 83.61 | 80 | 86.11 | 90.67 | 75 | 68 | 79.49 | 83.18 | 70.67 | 69.23 | 72.22 | 76.85 |
| Hybrid-MACN[41] | 78.69 | 72 | 83.33 | 81.78 | 70.31 | 64 | 74.36 | 77.54 | 64 | 61.54 | 66.67 | 66.45 |
| PG-BrainBagNet[47] | 85.25 | 84 | 87.18 | 89.03 | 81.25 | 80 | 82.05 | 88.92 | 76 | 74.36 | 77.78 | 80.34 |
| Ours | 93.44 | 92 | 94.44 | 96.67 | 89.06 | 88 | 89.74 | 92 | 84 | 87.18 | 80.56 | 81.84 |

**Table 3** The performance of single-mode experiments

| Target | Modality | Performance (%) | | | |
|---|---|---|---|---|---|
| | | ACC | SEN | SPE | AUC |
| AD vs. HC | MRI | 86.89 | 88 | 86.11 | 92.22 |
| AD vs. MCI | | 82.81 | 80 | 84.62 | 88.62 |
| MCI vs. HC | | 78.67 | 80.56 | 76.92 | 80.77 |

make full use of imaging data and genetic information to provide more comprehensive features for diagnosis.

#### Effectiveness of the attention module

The validity of the attention module was verified in the attentional convolutional subnetwork (ACS) in terms of diagnostic performance. The attention module consisted of spatial attention and channel attention branches. The test was divided into four parts: model diagnosis without attention, model diagnosis with only channel attention, only spatial attention, and combined channel and spatial attention. These were denoted as Baseline, ACS-Channel, ACS-Spatial, and ACS-Combined, respectively. Table 4 shows the performance of the four groups of experiments in the diagnosis of AD and HC, AD and MCI, and MCI and HC. The results showed that the diagnostic performance of the model was further improved with spatial attention and channel attention. Spatial attention had better classification performance than channel attention because spatial attention aggregated more feature information. Nevertheless, ACS-Combined further improved classification performance by combining channel and spatial attention.

#### SNPs results

Figure 2 shows SNPs and corresponding *p*-values extracted from the GWAS via the Manhattan plot. The X-axis represents the coding information of the chromosome where SNPs are located, and the Y-axis represents the $-log_{10}(p)$ of the SNPs. SNP sites on different chromosomes are distinguished by different colors. The smaller the *p*-value, the higher the position of the corresponding point of SNPs, indicating that this SNP is more significantly associated with AD. ($-log\ (p < 0.05) = 1.301$).

### Results for significantly correlated genes

After GWAS analysis, the gene information was correspondingly calculated with SNP locus. Figure 3 demonstrates the visualization results of genes by the Manhattan plot. The X-axis represents the coding information of chromosome where the genes are located, and the Y-axis represents the $-log_{10}(p)$ of the genes. A smaller $p$-value means a higher position of the corresponding point of the gene, which represents the more significant association of genes with AD. This analysis provides information on significant genes associated with AD.

After Bonferroni correction, the genes and the corresponding $p$-values of the target subjects were obtained. Among them, there were 119 genes with $p < 0.05$ and 35 genes with $p < 0.001$. Table 5 lists the significant gene information. It was found that disease-causing genes with $p < 0.001$ such as *CSMD1*, *PTPRD*, *RBFOX1*, *CDH13* and *WWOX* were significantly correlated with AD. More importantly, many previous studies provided evidence for the significant correlation between these genes and AD, verifying the validity of this study. In addition, *NTM*, *MAML2*, *NAALADL2*, *FHIT*, *TMEM132D* and *PCSK5* were also significantly associated with AD. As potential therapeutic targets for neurodegenerative diseases, these genes would warrant further functional studies in the future.

### Visualization of experimental performance

The performance of the experiments was also demonstrated by ROC curves and confusion matrices. Figures 4 and 5 show the ROC curves and confusion matrices for the three groups of experiments, respectively. Each ROC curve and confusion matrix represented the performance of the framework on its task and dataset. The results revealed that the AUC values for the AD-HC, AD-MCI and MCI-HC groups were 96.67%, 92% and 81.84%, respectively. The true positive and true negative values for diagnosis in the AD-HC group were 1190 and 805. In the AD-MCI group, the true positive and true negative values for diagnosis were 1,225 and 770. In the AD-HC group, the true positive and true negative values for diagnosis were 1015 and 1190.

**Table 4** The performance of the individual attention module

| Target | Method | Performance (%) | | | |
|---|---|---|---|---|---|
| | | ACC | SEN | SPE | AUC |
| AD vs. HC | Baseline | 88.52 | 88 | 88.89 | 93.78 |
| | ACS-Channel | 90.17 | 88 | 91.67 | 94.33 |
| | ACS-Spatial | 91.8 | 96 | 88.89 | 95.22 |
| | ACS-Combined | 93.44 | 92 | 94.44 | 96.67 |
| AD vs. MCI | Baseline | 84.38 | 80 | 87.18 | 90.46 |
| | ACS-Channel | 85.94 | 88 | 84.62 | 91.18 |
| | ACS-Spatial | 87.5 | 84 | 89.74 | 91.38 |
| | ACS-Combined | 89.06 | 88 | 89.74 | 92 |
| MCI vs. HC | Baseline | 80 | 80.56 | 79.49 | 80.34 |
| | ACS-Channel | 81.33 | 84.62 | 77.78 | 80.56 |
| | ACS-Spatial | 82.67 | 84.62 | 80.56 | 80.7 |
| | ACS-Combined | 84 | 87.18 | 80.56 | 81.84 |

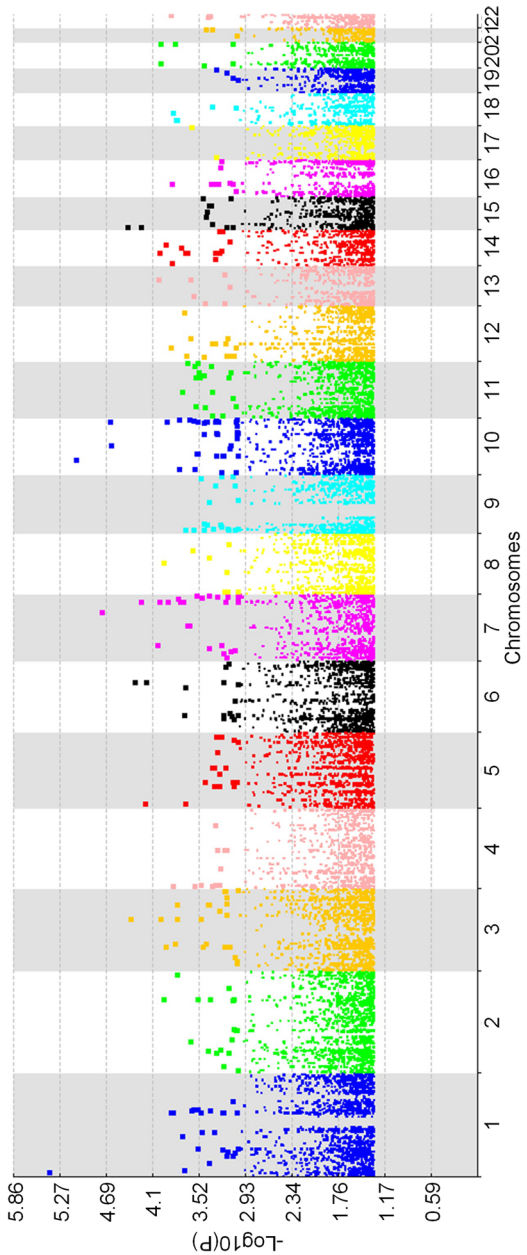


**Fig. 2** Graph of the SNP Manhattan plot

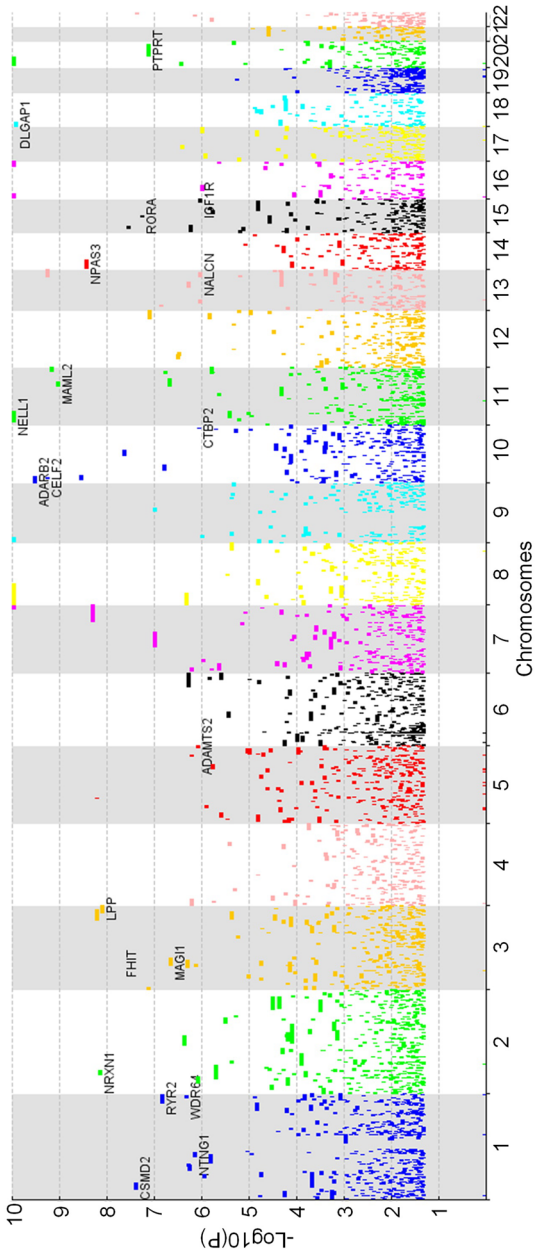


**Fig. 3** Manhattan plot of the genes

**Table 5** Information on genes significantly associated with AD

| Gene | *p*-Value | References |
|---|---|---|
| CSMD1 | 3.774126 × 10^-22 | Parcerisas et al. [48] |
| PTPRD | 1.178651 × 10^-16 | Uhl et al. [49] |
| RBFOX1 | 1.416824 × 10^-12 | Raghavan et al. [50] |
| CDH13 | 2.943574 × 10^-10 | Huang et al. [51] |
| WWOX | 4.055048 × 10^-9 | Hsu et al. [52] |
| DPP6 | 1.062618 × 10^-8 | Lin et al. [53] |
| GALNT18 | 1.032083 × 10^-7 | Li et al. [54] |
| MACROD2 | 4.574143 × 10^-7 | Kohannim et al. [55] |
| NELL1 | 6.65663 × 10^-7 | James et al. [56] |
| DLGAP1 | 7.63375 × 10^-7 | Katsumata et al. [57] |
| ADARB2 | 1.874849 × 10^-6 | Lee et al. [58] |
| FAM155A | 3.401599 × 10^-6 | Horimoto et al. [59] |
| CELF2 | 3.548167 × 10^-6 | Tran et al. [60] |
| NTM | 4.226044 × 10^-6 | - |
| MAML2 | 5.917683 × 10^-6 | - |
| CAMK1D | 1.758816 × 10^-5 | Floudas et al. [61] |
| NPAS3 | 2.314553 × 10^-5 | Nucifora et al. [62] |
| CNTNAP2 | 3.120677 × 10^-5 | Hirano et al. [63] |
| PDE4D | 3.829089 × 10^-5 | Xiang et al. [64] |
| NAALADL2 | 3.829089 × 10^-5 | - |
| NRXN1 | 4.531394 × 10^-5 | Hu et al. [65] |
| LPP | 4.995526 × 10^-5 | Patel et al. [66] |
| FHIT | 1.34354 × 10^-4 | - |
| CTNNA3 | 1.447359 × 10^-4 | Kondo et al. [67] |
| MEIS2 | 1.8321 × 10^-4 | Huang et al. [68] |
| CSMD2 | 2.577154 × 10^-4 | Stein et al. [69] |
| TAFA5 | 2.668759 × 10^-4 | Seong et al. [70] |
| RORA | 3.432134 × 10^-4 | Acquaah-Mensah et al. [71] |
| PTPRT | 4.64132 × 10^-4 | Ben-Avraham et al. [72] |
| CNTN4 | 4.781781 × 10^-4 | Bamford et al. [73] |
| TMEM132D | 4.910028 × 10^-4 | - |
| MAGI2 | 6.41235 × 10^-4 | Kim et al. [74] |
| PCSK5 | 6.53449 × 10^-4 | - |
| MTUS2 | 8.73301 × 10^-4 | Xicota et al. [75] |
| RYR2 | 9.09943 × 10^-4 | Stutzmann et al. [76] |

## Discussion

In this research, the new feature fusion method called the "MRI-*p* value" was proposed to construct 3D multimodal images and an attention mechanism and the residual module were applied to construct the diagnostic model. Traditional multimodal fusion faces two main problems. On the one hand, only the features selected for each modality are simply connected, ignoring the connection between the data of different modalities [23]. On the other hand, choosing to discard modally incomplete samples to construct their own dataset leads to the problem of insufficient data and unbalanced data from different categories [25], which is an important reason for not obtaining the best performance for diagnosis. As shown in Tables 2 and 3, we used the same model to train the data

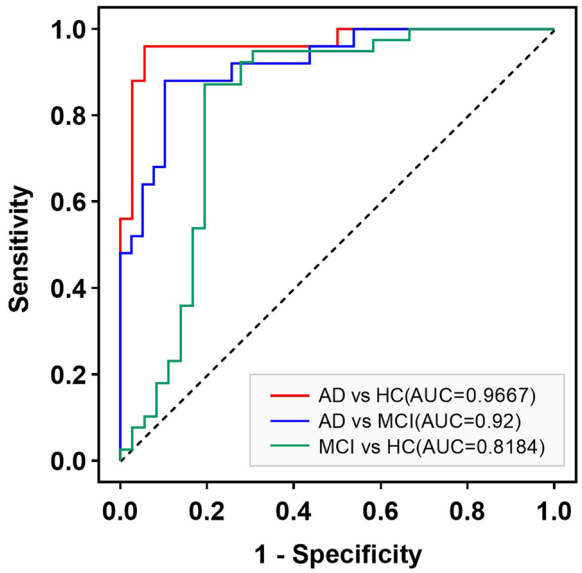


**Fig. 4** The graph of ROC curves

before and after fusion, and the resulting accuracy rates presented significant differences (6.55% in the AD-HC group, 6.25% in the AD-MCI group, and 5.33% in the MCI-HC group). This finding indicates that our proposed data fusion method fully considers the intrinsic relationship between imaging data and genetic information, effectively enhancing the differences between different groups of data and obtaining significant gene information related to AD through model screening. Moreover, our data fusion approach has enriched the dataset, providing a diverse source of data and a more comprehensive feature representation for the training of diagnostic models.

A new deep joint learning framework was proposed for the assisted diagnosis of AD with multimodal feature fusion images. On the basis of ResNet, the addition of an attention module enhances the extraction of visual features through integrated residual attention learning. This effectively transmits and accumulates attention information from queries and keys, strengthening the model's generalization capabilities. In previous diagnostics, the location branch often constructed gated branches from positions and obtained patch-level evidence by the traditional convolutional layer [47]. In this study, we optimized this approach and introduced a new attention mechanism to accurately capture the location information of brain regions. First, attention maps were generated in the spatial and channel domains respectively by the attention modules of the two branches. Second, the relative position information of local brain regions was generated by combining the two branches. The combination of attention module and superposition convolution not only accurately learns the relative location information of pathological regions, but also globally controls the location information of the different categories of brain variations, further improving the final classification effect. As shown in Tables 2 and 4, there was a significant difference in the diagnostic results between the model without attention and those with mixed channel and spatial attention (4.92% in the AD-HC group, 4.68% in the AD-MCI group, and 4% in the MICI-HC group). The validity of the diagnostic model was verified.

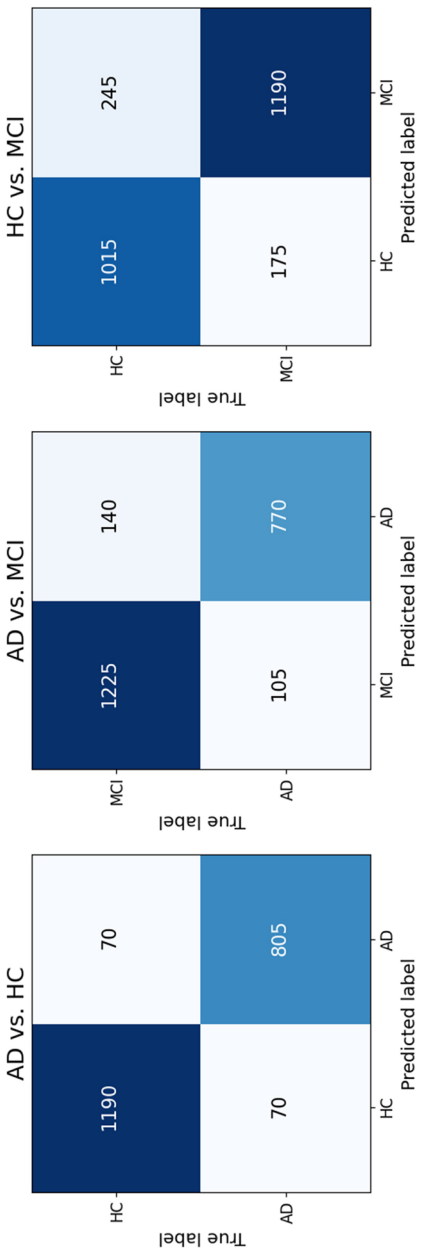


**Fig. 5** The graph of confusion matrices

Through data fusion and model optimization, we identified significant genes related to AD, including *CSMD1, PTPRD, RBFOX1* and *WWOX*. Among these genes, *CSMD1* might play a regulatory and modifying role by participating in various regulatory and signaling pathways involved in disease development [77]. *PTPRD* plays a role in neurofibrillary pathology, neurodevelopmental disorders, and cognitive impairment of AD [49]. The involvement of BACE1-mediated cleavage of *PTPRD* represents a potentially important new mechanism for heart metabolic diseases and AD risk [78]. The decrease of glucose metabolism in the posterior cingulate cortex (PCC) is one of the earliest biomarkers of AD [79]. Genome-wide significant SNPs (rs12444565) in *RBFOX1* and four suggestive loci (rs235141, rs79037, rs12526331, and rs12529764) are associated with fluorodeoxyglucose f18 [80]. *WWOX* is a gene that has been implicated in the pathology of hippocampal sclerosis (HS) [81], and the largest AD GWAS to date has also revealed an association between AD and *MAF*, which is a downstream gene of *WW*OX [82]. Moreover, we also identified several potential genes related to AD, such as *NTM*, *MAML2*, *NAALADL2*, *FHIT*, *TMEM132D*, and *PCSK5*, which awaited further experimental validation.

In addition, there is room for further improvement. In future research, other pretrained architectures will be introduced and the proposed method will be fine-tuned to obtain more desirable results. At the same time, we will also work on the following aspects: (a) we will conduct in-depth research on the fusion of imaging data and genetic information to further explore the influence of feature correlation in more models on the disease and play the role of genetic information in the diagnosis of AD; (b)a subnetwork will be proposed to encode age into image features and learn the potential dependencies between the subjects of different ages to reduce the effects of normal aging on classification; (c) more challenging multiclassification tasks will be studied by a large number of experiments to build a multiclassification diagnosis with the best performance.

## Conclusion

In this paper, we proposed a new feature fusion method named the "MRI-*p* value" which made full use of the imaging and genetic information from the participants. The fused 3D images did not damage the basic structure of the original 3D images. Additionally, we introduced prior knowledge provided by genetic data to provide a more comprehensive feature representation for subsequent diagnostic models. Moreover, we combined the attention mechanism and ResNet network to construct a new deep joint learning diagnostic model. By fully learning the feature information and location information of local pathological regions, we obtained the basis for disease diagnosis through aggregation. We evaluated this method using subjects downloaded from the ADNI dataset. The proposed model achieved an accuracy (ACC) and area under the curve (AUC) of 93.44% and 96.67% for HC and AD diagnosis, 89.06% and 92% for AD and MCI diagnosis, and 84% and 81.84% for HC and MCI diagnosis tasks, respectively. Through our experiments, we discovered significant genes related to AD, such as *CSMD1*, *PTPRD*, *RBFOX1*, and *WWOX*, and identified several potential genes related to AD, specifically including *NTM*, *MAML2*, *NAALADL2*, *FHIT*, *TMEM132D*, and *PCSK5*.

**Acknowledgements**

This work was supported by the National Natural Science Foundation of China under Grant No. 62371074, Jiangsu Provincial Key Research and Development Program under Grant No. BE2021636, Natural Science Foundation of the Jiangsu

Higher Education Institutions of China under Grant No. 23KJB520002, No.24KJA520002 and Changzhou Science and Technology Plan under Grant No. CZ20230028. This work was also sponsored by Qing Lan Project of Jiangsu Province.

**Authors' Contributions**

Jingru Wang: Conceptualization, methodology, formal analysis, data curation, Writing – original draft, validation, writing – review and editing. Shipeng Wen: Conceptualization, methodology, writing – review and editing. Wenjie Liu: Conceptualization, methodology, validation, writing – review & editing. Xianglian Meng: Conceptualization, methodology, formal analysis, data curation, validation, writing – review and editing. Zhuqing Jiao: Conceptualization, methodology, formal analysis, writing – original draft, validation, writing – review and editing.

**Data availability**

The datasets analyzed during the current study are available in the Alzheimer's Disease Neuroimaging Initiative (ADNI) public dataset which is publicly available on the website (http://adni.loni.usc.edu).

## Declarations

**Competing interest**

The authors declare no competing interests.

Received: 28 August 2024 Accepted: 8 October 2024
Published online: 05 November 2024

**References**

1. Porsteinsson A P, Isaacson R, Knox S, et al. Diagnosis of early Alzheimer's disease: clinical practice in 2021. The journal of prevention of Alzheimer's disease. 2021; 8: 371–86. https://doi.org/10.14283/jpad.2021.23.
2. Srivastava S, Ahmad R, Khare SK. Alzheimer's disease and its treatment by different approaches: A review. Eur J Med Chem. 2021;216: 113320. https://doi.org/10.1016/j.ejmech.2021.113320.
3. Rao Y L, Ganaraja B, Murlimanju B, et al. Hippocampus and its involvement in Alzheimer's disease: a review. 3 Biotech. 2022; 12(2): 55. https://doi.org/10.1007/s13205-022-03123-4.
4. Frizzell TO, Glashutter M, Liu CC, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: A systematic review. Ageing Res Rev. 2022;77: 101614. https://doi.org/10.1016/j.arr.2022.101614.
5. Teng Z. Novel Development and Prospects in Pathogenesis, Diagnosis, and Therapy of Alzheimer's Disease. Journal of Alzheimer's Disease Reports. 2024;8(1):345–54. https://doi.org/10.3233/ADR-230130.
6. Golovanevsky M, Eickhoff C, Singh R. Multimodal attention-based deep learning for Alzheimer's disease diagnosis. J Am Med Inform Assoc. 2022;29(12):2014–22. https://doi.org/10.1093/jamia/ocac168.
7. Sheng J, Xin Y, Zhang Q, et al. Predictive classification of Alzheimer's disease using brain imaging and genetic data. Sci Rep. 2022;12(1):2405. https://doi.org/10.1038/s41598-022-06444-9.
8. Fareed MMS, Zikria S, Ahmed G, et al. ADD-Net: an effective deep learning model for early detection of Alzheimer disease in MRI scans. IEEE Access. 2022;10:96930–51. https://doi.org/10.1109/ACCESS.2022.3204395.
9. Zhang F, Pan B, Shao P, et al. A single model deep learning approach for Alzheimer's disease diagnosis. Neuroscience. 2022;491:200–14. https://doi.org/10.1016/j.neuroscience.2022.03.026.
10. Zhang P, Lin S, Qiao J, et al. Diagnosis of Alzheimer's Disease with ensemble learning classifier and 3D convolutional neural network. Sensors. 2021;21(22):7634. https://doi.org/10.3390/s21227634.
11. Teipel SJ, Grothe M, Lista S, et al. Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease. Medical Clinics. 2013;97(3):399–424. https://doi.org/10.1016/j.mcna.2012.12.013.
12. Vemuri P, Jack CR. Role of structural MRI in Alzheimer's disease. Alzheimer's research & therapy. 2010;2:1–10. https://doi.org/10.1186/alzrt47.
13. Tanveer M, Richhariya B, Khan R U, et al. Machine learning techniques for the diagnosis of Alzheimer's disease: A review. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM). 2020; 16(1s): 1–35. https://doi.org/10.1145/3344998.
14. Tu Y, Lin S, Qiao J, et al. Alzheimer's disease diagnosis via multimodal feature fusion. Comput Biol Med. 2022;148: 105901. https://doi.org/10.1016/j.compbiomed.2022.105901.
15. Jiao C-N, Gao Y-L, Ge D-H, et al. Multi-modal imaging genetics data fusion by deep auto-encoder and self-representation network for Alzheimer's disease diagnosis and biomarkers extraction. Eng Appl Artif Intell. 2024;130: 107782. https://doi.org/10.1016/j.engappai.2023.107782.
16. Li L, Yu X, Sheng C, et al. A review of brain imaging biomarker genomics in Alzheimer's disease: implementation and perspectives. Translational Neurodegeneration. 2022;11(1):42. https://doi.org/10.1186/s40035-022-00315-z.
17. Ying Q, Xing X, Liu L, et al. Multi-modal data analysis for alzheimer's disease diagnosis: An ensemble model using imagery and genetic features; proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), F, 2021 [C]. IEEE. https://doi.org/10.1109/EMBC46164.2021.9630174.
18. Novikova G, Kapoor M, Tcw J, et al. Integration of Alzheimer's disease genetics and myeloid genomics identifies disease risk regulatory elements and genes. Nat Commun. 2021;12(1):1610. https://doi.org/10.1038/s41467-021-21823-y.
19. Ridge P G, Hoyt K B, Boehme K, et al. Assessment of the genetic variance of late-onset Alzheimer's disease. Neurobiology of aging. 2016; 41: 200. e13-. e20. https://doi.org/10.1016/j.neurobiolaging.2016.02.024.
20. Meng X, Wei Q, Meng L, et al. Feature fusion and detection in Alzheimer's disease using a novel genetic multi-kernel SVM based on MRI imaging and gene data. Genes. 2022;13(5):837. https://doi.org/10.3390/genes13050837.

21. Bi X-A, Hu X, Wu H, et al. Multimodal data analysis of Alzheimer's disease based on clustering evolutionary random forest. IEEE J Biomed Health Inform. 2020;24(10):2973–83. https://doi.org/10.1109/JBHI.2020.2973324.
22. Hu Z, Wang X, Meng L, et al. Detection of Association Features Based on Gene Eigenvalues and MRI Imaging Using Genetic Weighted Random Forest. Genes. 2022;13(12):2344. https://doi.org/10.3390/genes13122344.
23. Zhou T, Thung K-H, Zhu X, et al. Feature learning and fusion of multimodality neuroimaging and genetic data for multi-status dementia diagnosis; proceedings of the Machine Learning in Medical Imaging: 8th International Workshop, MLMI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 10, 2017, Proceedings 8, F, 2017 [C]. Springer. https://doi.org/10.1007/978-3-319-67389-9_16.
24. Pan Y, Liu M, Lian C, et al. Synthesizing missing PET from MRI with cycle-consistent generative adversarial networks for Alzheimer's disease diagnosis; proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16–20, 2018, Proceedings, Part III 11, F, 2018 [C]. Springer. https://doi.org/10.1007/978-3-030-00931-1_52.
25. Dolci G, Rahaman M A, Chen J, et al. A deep generative multimodal imaging genomics framework for Alzheimer's disease prediction; proceedings of the 2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE), F, 2022 [C]. IEEE. https://doi.org/10.1109/BIBE55377.2022.00017.
26. Zeng N, Li H, Peng Y. A new deep belief network-based multi-task learning for diagnosis of Alzheimer's disease. Neural Comput Appl. 2023;35(16):11599–610. https://doi.org/10.1007/s00521-021-06149-6.
27. Adedigba AP, Adeshina SA, Aibinu AM. Performance evaluation of deep learning models on mammogram classification using small dataset. Bioengineering. 2022;9(4):161. https://doi.org/10.3390/bioengineering9040161.
28. Chen L, Qiao H, Zhu F. Alzheimer's disease diagnosis with brain structural mri using multiview-slice attention and 3D convolution neural network. Frontiers in Aging Neuroscience. 2022;14: 871706. https://doi.org/10.3389/fnagi.2022.871706.
29. Amini M, Pedram M, Moradi A, et al. Diagnosis of Alzheimer's disease severity with fMRI images using robust multitask feature extraction method and convolutional neural network (CNN). Comput Math Methods Med. 2021;2021:1–15. https://doi.org/10.1155/2021/5514839.
30. Bakkouri I, Afdel K. DermoNet: A computer-aided diagnosis system for dermoscopic disease recognition; proceedings of the Image and Signal Processing: 9th International Conference, ICISP. Marrakesh, Morocco, June 4–6, 2020, Proceedings 9, F, 2020 [C]. Springer. 2020. https://doi.org/10.1007/978-3-030-51935-3_18.
31. Zhang X, Han L, Zhu W, et al. An explainable 3D residual self-attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE J Biomed Health Inform. 2021;26(11):5289–97. https://doi.org/10.1109/JBHI.2021.3066832.
32. Bakkouri I, Bakkouri S. 2MGAS-Net: multi-level multi-scale gated attentional squeezed network for polyp segmentation. Signal, Image and Video Processing. 2024: 1–10. https://doi.org/10.1007/s11760-024-03240-y.
33. Zhu W, Sun L, Huang J, et al. Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI. IEEE Trans Med Imaging. 2021;40(9):2354–66. https://doi.org/10.1109/TMI.2021.3077079.
34. Amunts K, Mohlberg H, Bludau S, et al. Julich-Brain: A 3D probabilistic atlas of the human brain's cytoarchitecture. Science. 2020;369(6506):988–92. https://doi.org/10.1126/science.abb4588.
35. Saykin AJ, Shen L, Foroud TM, et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. Alzheimer's & Dementia. 2010;6(3):265–73. https://doi.org/10.1016/j.jalz.2010.03.013.
36. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics. 2007;81(3):559–75. https://doi.org/10.1086/519795.
37. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2016 [C]. https://doi.org/10.1109/CVPR.2016.90.
38. Lecun Y, Bengio Y, Hinton G. Deep learning nature. 2015;521(7553):436–44. https://doi.org/10.1038/nature14539.
39. Ebrahimi A, Luo S, Chiong R. Introducing transfer learning to 3D ResNet-18 for Alzheimer's disease detection on MRI images; proceedings of the. 35th international conference on image and vision computing New Zealand (IVCNZ), F, 2020 [C]. IEEE. 2020. https://doi.org/10.1109/IVCNZ51579.2020.9290616.
40. Liu R, Lehman J, Molino P, et al. An intriguing failing of convolutional neural networks and the coordconv solution. Advances in neural information processing systems. 2018; 31. https://doi.org/10.48550/arXiv.1807.03247.
41. Gao X, Cai H, Liu M. A hybrid multi-scale attention convolution and aging transformer network for Alzheimer's disease diagnosis. IEEE J Biomed Health Inform. 2023. https://doi.org/10.1109/JBHI.2023.3270937.
42. Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning; proceedings of the International conference on machine learning, F, 2018 [C]. PMLR.
43. Kingma D P, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980. 2014. https://doi.org/10.48550/arXiv.1412.6980.
44. He T, Zhang Z, Zhang H, et al. Bag of tricks for image classification with convolutional neural networks; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2019 [C]. https://doi.org/10.48550/arXiv.1812.01187.
45. Liu S, Masurkar AV, Rusinek H, et al. Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs. Sci Rep. 2022;12(1):17106. https://doi.org/10.1038/s41598-022-20674-x.
46. Oh K, Yoon JS, Suk H-I. Learn-explain-reinforce: counterfactual reasoning and its guidance to reinforce an Alzheimer's Disease diagnosis model. IEEE Trans Pattern Anal Mach Intell. 2022;45(4):4843–57. https://doi.org/10.1109/TPAMI.2022.3197845.
47. Park C, Jung W, Suk H-I. Deep joint learning of pathological region localization and Alzheimer's disease diagnosis. Sci Rep. 2023;13(1):11664. https://doi.org/10.1038/s41598-023-38240-4.
48. Parcerisas A, Rubio SE, Muhaisen A, et al. Somatic signature of brain-specific single nucleotide variations in sporadic Alzheimer's disease. Journal of Alzheimer's Disease. 2014;42(4):1357–82. https://doi.org/10.3233/JAD-140891.
49. Uhl GR, Martinez MJ. PTPRD: neurobiology, genetics, and initial pharmacology of a pleiotropic contributor to brain phenotypes. Ann N Y Acad Sci. 2019;1451(1):112–29. https://doi.org/10.1111/nyas.14002.

50. Raghavan NS, Dumitrescu L, Mormino E, et al. Association between common variants in RBFOX1, an RNA-binding protein, and brain amyloidosis in early and preclinical Alzheimer disease. JAMA Neurol. 2020;77(10):1288–98. https://doi.org/10.1001/jamaneurol.2020.1760.
51. Huang J, Lu D, Meng G. Module analysis using single-patient differential expression signatures improve the power of association study for Alzheimer's disease. bioRxiv. 2020: 2020.01. 05.894931. https://doi.org/10.1101/2020.01.05.894931.
52. Hsu C-Y, Lee K-T, Sun T-Y, et al. WWOX and its binding proteins in neurodegeneration. Cells. 2021;10(7):1781. https://doi.org/10.3390/cells10071781.
53. Lin L, Petralia RS, Holtzclaw L, et al. Alzheimer's disease/dementia-associated brain pathology in aging DPP6-KO mice. Neurobiol Dis. 2022;174: 105887. https://doi.org/10.1016/j.nbd.2022.105887.
54. Li X, Chu S-G, Shen X-N, et al. Genome-wide association study identifies SIAH3 locus influencing the rate of ventricular enlargement in non-demented elders. Aging (Albany NY). 2019; 11(21): 9862. https://doi.org/10.18632/aging.102435.
55. Kohannim O, Hibar DP, Stein JL, et al. Discovery and replication of gene influences on brain structure using LASSO regression. Front Neurosci. 2012;6:115. https://doi.org/10.3389/fnins.2012.00115.
56. James AW, Shen J, Zhang X, et al. NELL-1 in the treatment of osteoporotic bone loss. Nat Commun. 2015;6(1):7362. https://doi.org/10.1038/ncomms8362.
57. Katsumata Y, Nelson PT, Estus S, et al. Translating Alzheimer's disease–associated polymorphisms into functional candidates: a survey of IGAP genes and SNPs. Neurobiol Aging. 2019;74:135–46. https://doi.org/10.1016/j.neurobiolaging.2018.10.017.
58. Lee E, Giovanello KS, Saykin AJ, et al. Single-nucleotide polymorphisms are associated with cognitive decline at Alzheimer's disease conversion within mild cognitive impairment patients. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2017;8:86–95. https://doi.org/10.1016/j.dadm.2017.04.004.
59. Horimoto A R, Boyken L A, Blue E E, et al. Admixture mapping implicates 13q33. 3 as ancestry-of-origin locus for Alzheimer disease in Hispanic and Latino populations. Human Genetics and Genomics Advances. 2023; 4(3). https://doi.org/10.1016/j.xhgg.2023.100207.
60. Peng S, Cai X, Chen J, et al. The role of CELF family in neurodevelopment and neurodevelopmental disorders. Neurobiology of Disease. 2024:106525. https://doi.org/10.1016/j.nbd.2024.106525.
61. Floudas CS, Um N, Kamboh MI, et al. Identifying genetic interactions associated with late-onset Alzheimer's disease. BioData mining. 2014;7:1–19. https://doi.org/10.1186/s13040-014-0035-z.
62. Nucifora LG, Wu YC, Lee BJ, et al. A mutation in NPAS3 that segregates with schizophrenia in a small family leads to protein aggregation. Molecular Neuropsychiatry. 2016;2(3):133–44. https://doi.org/10.1159/000447358.
63. Hirano A, Ohara T, Takahashi A, et al. A genome-wide association study of late-onset Alzheimer's disease in a Japanese population. Psychiatr Genet. 2015;25(4):139–46. https://doi.org/10.1097/YPG.0000000000000090.
64. Xiang J, Wang X, Gao Y, et al. Phosphodiesterase 4D gene modifies the functional network of patients with mild cognitive impairment and Alzheimer's disease. Front Genet. 2020;11:890. https://doi.org/10.3389/fgene.2020.00890.
65. Hu Z, Xiao X, Zhang Z, et al. Genetic insights and neurobiological implications from NRXN1 in neuropsychiatric disorders. Mol Psychiatry. 2019;24(10):1400–14. https://doi.org/10.1038/s41380-019-0438-9.
66. Patel H, Iniesta R, Stahl D, et al. Working towards a blood-derived gene expression biomarker specific for Alzheimer's disease. Journal of Alzheimer's Disease. 2020;74(2):545–61. https://doi.org/10.3233/JAD-191163.
67. Kondo T, Hara N, Koyama S, et al. Dissection of the polygenic architecture of neuronal Aβ production using a large sample of individual iPSC lines derived from Alzheimer's disease patients. Nature aging. 2022;2(2):125–39. https://doi.org/10.1038/s43587-021-00158-9.
68. Huang M, Deng C, Yu Y, et al. Spatial correlations exploitation based on nonlocal voxel-wise GWAS for biomarker detection of AD. NeuroImage: Clinical. 2019; 21: 101642. https://doi.org/10.1016/j.nicl.2018.101642.
69. Stein J L, Hua X, Lee S, et al. Voxelwise genome-wide association study (vGWAS). neuroimage. 2010; 53(3): 1160–74. https://doi.org/10.1016/j.neuroimage.2010.02.032.
70. Seong J Y, Kim H-B, Yoo S, et al. Antibody-based immunotherapy targeting FAM19A5 reverses synaptic loss and improves cognitive function in Alzheimer's disease. bioRxiv. 2023: 2023.11. 22.568357. https://doi.org/10.1101/2023.11.22.568357.
71. Acquaah-Mensah GK, Agu N, Khan T, et al. A regulatory role for the insulin-and BDNF-linked RORA in the hippocampus: implications for Alzheimer's disease. Journal of Alzheimer's Disease. 2015;44(3):827–38. https://doi.org/10.3233/JAD-141731.
72. Ben-Avraham D, Karasik D, Verghese J, et al. The complex genetics of gait speed: genome-wide meta-analysis approach. Aging (Albany NY). 2017; 9(1): 209. https://doi.org/10.18632/aging.101151.
73. Bamford RA, Widagdo J, Takamura N, et al. The interaction between contactin and amyloid precursor protein and its role in Alzheimer's disease. Neuroscience. 2020;424:184–202. https://doi.org/10.1016/j.neuroscience.2019.10.006.
74. Kim H-R, Lee T, Choi JK, et al. Polymorphism in the MAGI2 gene modifies the effect of Amyloid β on neurodegeneration. Alzheimer Dis Assoc Disord. 2021;35(2):114–20. https://doi.org/10.1097/WAD.0000000000000422.
75. Xicota L, Cosentino S, Vardarajan B, et al. Whole genome-wide sequence analysis of long-lived families (Long-Life Family Study) identifies MTUS2 gene associated with late-onset Alzheimer's disease. Alzheimer's & Dementia. 2024. https://doi.org/10.1002/alz.13718.
76. Stutzmann G E. RyR2 calcium channels in the spotlight—I'm ready for my close up, Dr. Alzheimer! Cell Calcium. 2021; 94: 102342. https://doi.org/10.1016/j.ceca.2020.102342.
77. Stepanov VA, Bocharova AV, Marusin AV, et al. Replicative association analysis of genetic markers of cognitive traits with Alzheimer's disease in the Russian population. Mol Biol. 2014;48(6):835–44. https://doi.org/10.1134/S0026893314060168.
78. Taylor H A, Simmons K J, Clavane E M, et al. PTPRD and DCC Are Novel BACE1 Substrates Differentially Expressed in Alzheimer's Disease: A Data Mining and Bioinformatics Study [J/OL] 2022, 23(9):4568. https://doi.org/10.3390/ijms23094568.

79. Minoshima S, Giordani B, Berent S, et al. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Annals of Neurology. 1997; 42(1): 85–94. https://doi.org/10.1002/ana.410420114.
80. Kong L-L, Miao D, Tan L, et al. Genome-wide association study identifies RBFOX1 locus influencing brain glucose metabolism. Annals of Translational Medicine. 2018; 6(22): 436. https://doi.org/10.21037/atm.2018.07.05.
81. Bellenguez C, Küçükali F, Jansen I E, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nature genetics. 2022;54(4):412-36. https://doi.org/10.1038/s41588-022-01024-z.
82. Nelson PT, Estus S, Abner EL, et al. ABCC9 gene polymorphism is associated with hippocampal sclerosis of aging pathology. Acta Neuropathol. 2014;127(6):825–43. https://doi.org/10.1007/s00401-014-1282-2.

## Publisher's Note