



Multi-modal imaging genetics data fusion by deep auto-encoder and self-representation network for Alzheimer's disease diagnosis and biomarkers extraction

Cui-Na Jiao^{a,b}, Ying-Lian Gao^c, Dao-Hui Ge^a, Junliang Shang^a, Jin-Xing Liu^{a,d,*}

^a School of Computer Science, Qufu Normal University, Rizhao, 276826, China

^b School of Engineering, Qufu Normal University, Rizhao, 276826, China

^c Qufu Normal University Library, Qufu Normal University, Rizhao, 276826, China

^d School of Health and Life Sciences, University of Health and Rehabilitation Sciences, Qingdao, 266113, China

ARTICLE INFO

Keywords:

Multi-modal imaging genetics
Auto-encoder
Self-representation
Multi-task learning
Alzheimer's disease

ABSTRACT

Alzheimer's disease (AD) is an incurable neurodegenerative disease, so it is important to intervene in the early stage of the disease. Brain imaging genetics is an effective technique to identify AD-related biomarkers, which can early diagnosis of AD patients once they are clinically verified. With the development of medical imaging and gene sequencing techniques, the association analysis between multi-modal imaging data and genetic data has garnered increasing attention. However, current imaging genetics studies have problem with non-intuitive data fusion. Meanwhile, the characteristics of multi-modal imaging genetics data are high-dimensional, non-linearity, and fewer subjects, so it is necessary to select effective features. In this paper, a multi-modal data fusion framework by deep auto-encoder and self-representation (MFASN) was proposed for early diagnosis of AD. First, a multi-modality brain network was constructed by combining information from the resting-state functional magnetic resonance imaging (fMRI) data and structural magnetic resonance imaging (sMRI) data. Then, we utilized the deep auto-encoder to achieve non-linear transformations and select the informative features. A sparse self-representation module was employed to capture the multi-subspaces structure of the latent representation. At last, a multi-task structured sparse association model was developed to fully mine correlations between the genetic data and multi-modal brain network features. Experiments on AD neuroimaging initiative datasets proved the superiority of the proposed method, while discovering discriminative biomarkers were strongly associated with AD.

1. Introduction

Alzheimer's disease (AD) is a typical progressive neurodegenerative disease with an insidious onset, which lies at a high incidence level in the elderly (Arifoglu et al., 2020; Sharma et al., 2019). AD is characterized by some symptoms such as memory impairment, aphasia, executive dysfunction, and so on. The exact pathogenesis of AD is not yet clear, and its disease course is irreversible. Thus, how to diagnose the disease at an early stage and alleviate the progression of the disease through treatment is a hot research topic (Arifoglu et al., 2020; Erdogmus and Kabakus, 2023).

To achieve early diagnosis and timely treatment of diseases, some disease risk assessment approaches were proposed in recent years. For example, considering the heterogeneity of data, two Einstein

aggregation operators (AOs) including q-rung picture fuzzy dynamic Einstein weighted averaging operator and q-rung picture fuzzy dynamic Einstein weighted geometric operator were proposed to achieve disease risk assessment (Kausar et al., 2022). In 2022, Farid et al. employed the notion of proportional distribution to provide a fair aggregate for liner Diophantine fuzzy numbers and constructed a multi-criteria decision-making (MCDM) approach (Farid et al., 2022). After that, Kausar et al. presented a robust approach for assessing disease risk based on the q-rung orthopair fuzzy set (q-ROFS), meanwhile, the distance-based assessment technique was integrated into the q-ROFS model (Kausar et al., 2023).

As a brain disease, AD mainly manifests in two aspects. One is the degeneration of brain structures and functions in the macrolevel and the other is the potential heritability in the microlevel (Roshchupkin et al.,

* Corresponding author. School of Health and Life Sciences, University of Health and Rehabilitation Sciences, Qingdao, 266113, China.

E-mail address: sdcavell@126.com (J.-X. Liu).

2016). Brain imaging genetics is a rising field of brain research in recent years, which integrated analysis of macroscopic brain imaging data and microscopic genetic data. It provided new insights into the phenotypic, genetic, and molecular characteristics of the brain, as well as their influence on normal and disordered brain function and behavior (Shen and Thompson, 2019). Although these existing multi-omics data have complementary information, it is difficult to perform multi-modal data fusion properly. Traditional multi-modal data fusion only simply spliced the feature matrix of the multi-omics data (Lei et al., 2020) or used the optimization function (Zhang et al., 2022), leading to incomplete information fusion. Thus, it is important to fuse the multi-modal imaging data and genetic data in a reasonable way to perform association analysis.

In this paper, we focused on developing a multi-modal data fusion framework for association analysis between brain imaging data (such as structural magnetic resonance imaging (sMRI) and functional MRI (fMRI)) and genetic data (such as single nucleotide polymorphisms (SNPs)). Our imaging genetics association study can address the correlations between risk SNPs and brain structures or functions, for a better mechanistic understanding of disordered brain functions or behaviors. Moreover, through this association study, our work was also devoted to the identification of AD-related biomarkers. Once they are clinically verified, they can better predict the possibility of a person becoming an AD patient and guide clinical decision-making.

For imaging genetics association studies, multivariate regression analysis and bi-multivariate association analysis were classic methods to explore correlations between SNPs and brain imaging quantitative traits (QTs). In general, regression methods captured correlations by fixing one variable (such as imaging data) and performing regression analysis using the other variable (such as SNP data). For example, Vounou et al. developed the sparse reduced rank regression, a technique for multivariate modeling of high-dimensional imaging data and genetic data, and enforced sparsity in the regression coefficients (Vounou et al., 2010). Later, Hao et al. proposed a diagnosis-guided sparse regression model to identify multi-modal imaging makers that are associated with the risk SNP rs429358 (Hao et al., 2016). In 2019, Zhou et al. presented a joint projection learning and sparse regression model for a brain-wide genome-wide association study (Zhou et al., 2018), which solved issues of data heterogeneity. However, most existing regression methods neglected correlations between multiple imaging QTs data and multiple genetic data.

In contrast, the traditional canonical correlation analysis (CCA) (Yang et al., 2019) as a bi-multivariate model was proposed to detect rich associations between multi-SNPs and multi-QTs. CCA can optimally obtain two correlation weights for linearly transforming features of single modality imaging data and genetic data. Though it ensures the associations between data to be maximal, it may lead to an overfitting issue. As only a limited number of markers are pertinent in association analysis, the sparse CCA (SCCA) with sparsity constraints was proposed for single modality imaging genetics analysis (Du et al., 2018). In real applications, the imaging genetics data are high-dimensional with the non-linear distribution, so some nonlinear CCA methods were presented later. For instance, Kernel CCA (KCCA) (Melzer et al., 2003) which had non-linear mapping ability was presented to learn meaningful representations of data. Deep CCA (DCCA) (Andrew et al., 2013) as a deep learning-based association approach was proposed to perform the potential non-linear transformations of two views of data. Also, (Wang et al., 2022) further developed a graph-based deep learning (DS-SCCA) model to reconstruct the original imaging genetics data as well as identify their correlations. But the approaches mentioned above were two-view association methods that only utilized one modality of imaging data, neglecting the rich information from multi-modal data and restricting the association performance.

Recently, the association analysis between multi-modal imaging data and genetic data has garnered increasing attention with the development of medical imaging and gene sequencing techniques. The multi-

view SCCA (mSCCA) (Hao et al., 2017) as a naive extension of SCCA was presented to capture correlations between multi-modality imaging phenotypes and genotypes. Later, (Du et al., 2019) designed an improved multi-task SCCA (MTSCCA) that jointly learned a series of SCCA tasks to study effective correlations between SNPs and multi-modal imaging QTs. Also, several deep learning methods were also proposed based on multi-modal medical imaging data. Irfan et al. developed a hybrid deep neural network (HDNNs) by using computed tomography and X-ray imaging to predict the risk of diseases (Irfan et al., 2021). Additionally, based on multiple chest X-rays, researchers also introduced an inception-ResNet model to diagnose whether patients are sick automatically (Almalki et al., 2021), which can save time for medical professionals.

Nevertheless, most of the above approaches either concentrated on the correlation analysis between single-modality imaging data and genetic data by optimization algorithms or simply fused multi-modal data by data splicing, resulting in non-intuitive data fusion. In contrast, graph-based data representation can reflect the complex relations between data, further attracting more attention. For example, (Yu et al., 2021) constructed a heterogeneous network and applied deep learning to predict drug-disease associations, which improved the prediction performance. (Bi et al., 2020) designed a genetic evolution of random neural network cluster (GERNNC) method to classify patients with brain disease by constructing brain region-gene networks. Studies (Sporns, 2014) have also shown that the brain network can be seen as a functional structural representation of a brain system. The functional connections of the network reflected the connection strength among brain areas. Hence, for the comprehensive analysis of AD, it is worthy to perform multi-modal data fusion by introducing graph-based data representations into brain imaging genetics research.

Based on the mentioned above, we focused on developing a novel multi-modal data fusion framework by combining rich information from the structural imaging data, functional imaging data, and SNP data. Through the human Brainnetome atlas (Fan et al., 2016), we constructed a multi-modal brain network that took into account both network voxel node information and edge connectivity information. For the connection strength between brain regions, the brain functional connectivity matrices (FCM) were extracted from the resting-state functional magnetic resonance imaging (fMRI) data at first, and then the upper triangle elements of each FCM were collected as edge features for every sample. Also, to combine information from different imaging modalities, structural magnetic resonance imaging (sMRI) data were used to construct node features. Hence, this work aimed to identify the associations between multiple SNPs and multi-modal brain network QTs.

However, the characteristics of multi-modal brain imaging data and genetic data are high-dimensional, non-linear, and low sample size. Most existing methods neglected the non-linear manifolds in the high-dimensional data, meanwhile, they were also difficult to select informative and effective features. Additionally, conventional association methods took the original data as input without considering the potential multi-subspace structure information. Based on the above considerations, a novel multi-modal imaging genetics data fusion framework based on deep auto-encoder and self-representation network (MFASN) was proposed for AD diagnosis and biomarkers identification. The key contributions of this paper are given here.

- To utilize the complementary information from diverse brain imaging modalities, we constructed a multi-modal brain network by using sMRI and fMRI data. Based on the multi-modal brain network and SNP data, we further presented a systematic multi-modal data fusion framework for association analysis.
- A deep auto-encoder was developed to nonlinearly map the input data to a latent space and extract meaningful features. After that, a sparse self-representation module was proposed to perform subspace clustering on the latent representation from the auto-encoder to learn the similarity structures and capture the global information.

- Lastly, based on real sMRI data, fMRI data, and SNP data from AD Neuroimaging Initiative (ADNI) database, a novel multi-task structured sparse association method was developed to explore correlations between SNP data and multi-modal network phenotypes, and identify AD-related risk SNPs and brain connectome for guiding clinical decision-making.

2. Related work

2.1. Enhanced SCCA models

To accommodate various kinds of structures in imaging genetics data, many enhanced SCCA methods with different regularizations have been proposed. Based on three kinds of constraints commonly used in SCCA models: sparsity term, group sparsity, and graph Laplacian-type term, researchers developed the following SCCA variants. To improve the effectiveness of ℓ_1 -norm, Li et al. presented a truncated ℓ_1 -norm penalized SCCA to detect imaging genetic associations (Du et al., 2018). In 2020, Du et al. introduced the improved fused lasso and graph-guided lasso terms and further proposed a novel structured SCCA (FGL-SCCA) (Du et al., 2020a,b,c). On this basis, Wang et al. also developed a structured SCCA with GraphNet regularization in 2022 (Wang et al., 2022) to improve the stability of the graph-guided methods. However, the mentioned SCCA approaches aimed to associate SNP data with a single modality imaging data, neglecting the rich biological information in multi-modality data.

2.2. Multi-modal SCCA models

To combine information from multiple sources, the multi-modal SCCA methods were developed in recent years. For instance, an adaptive sparse multi-view canonical correlation analysis (AdaSMCCA) (Du et al., 2021) was proposed to detect associations among genomic, proteomic, and imaging data in 2021. To combine the merits of both SCCA and logistic regression, a joint multi-task learning model MT-SCCALR was presented to learn associations of multiple tasks (Du et al., 2020a, b, c). Later, Du et al. also developed a dirty MTSCCA (Du et al., 2020a, b, c) to explore modality-specific and shared imaging genetics correlations by using multi-modal imaging data. However, multi-modal imaging genetics data are high-dimensional, non-linear, and low sample size. Most of multi-modal SCCA methods neglected the non-linear distribution of data and could not select effective features. Additionally, they took the original data as input without considering the potential multi-subspace structure information of data.

2.3. Solutions to problems

Therefore, to nonlinearly project the input data to a latent space, we utilized the deep auto-encoder to extract low-dimensional informative features. Moreover, to explore the potential multi-subspace structures of data, our method employed the self-representation network to seek the block diagonal property of data by conducting subspace clustering. Lastly, a multi-task structured sparse association model with fused group lasso and sparsity term was developed to mine the associations between multi-modal brain network and SNP data. Extensive experimental results have proved that our method is effective.

3. Methods

The proposed MFASN framework shown in Fig. 1 was composed of four parts, including A) a multi-modal brain network constructed based on the fMRI and sMRI data, B) a deep auto-encoder employed to map the input data into a non-linear feature space and select the informative features, C) a self-representation network used to perform subspace clustering on the latent representation, and D) a multi-task structured sparse association model utilized to discover disease-relevant brain connectome and risk SNPs. Additionally, a clear block diagram of our proposed model was revealed in Fig. 2.

Next, we introduced the construction of the multi-modal brain network, deep autoencoder, self-representation network, multi-task structured sparse association model, and optimization algorithm.

3.1. Construction of the multi-modal brain network

To identify the potential AD-related brain connectome of the brain network from the fine-grained level, based on the Brainnetome atlas (Fan et al., 2016), the input fMRI and sMRI data were segmented into 246 regions of interest (ROIs), including 210 cortical and 36 subcortical regions. By averaging time series across all voxels in every of the 246 ROIs, the mean regional time series was determined. To further reflect the connection strength between ROI pairs, we calculated the Pearson correlation coefficients between pairs of regional time series. Pearson correlation captures linear relations between variables, which is widely used for quantifying functional correlations between brain regions (Jiang et al., 2020). After that, the obtained functional connectivity matrices (FCM) were normalized to Z scores by Fisher transformation, obtaining a 246×246 symmetric FCM for every sample.

Since a brain network is a depiction of a brain system, each ROI of the brain can be seen as a network vertex and correlation coefficients represent the edge weights. Based on the functional connectivity matrix built for every sample, removing 246 diagonal elements, the upper tri-

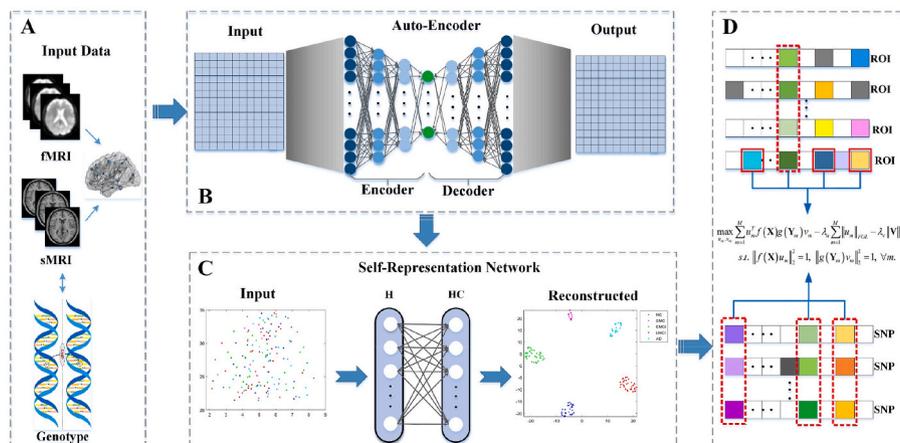


Fig. 1. The pipeline of proposed MFASN.

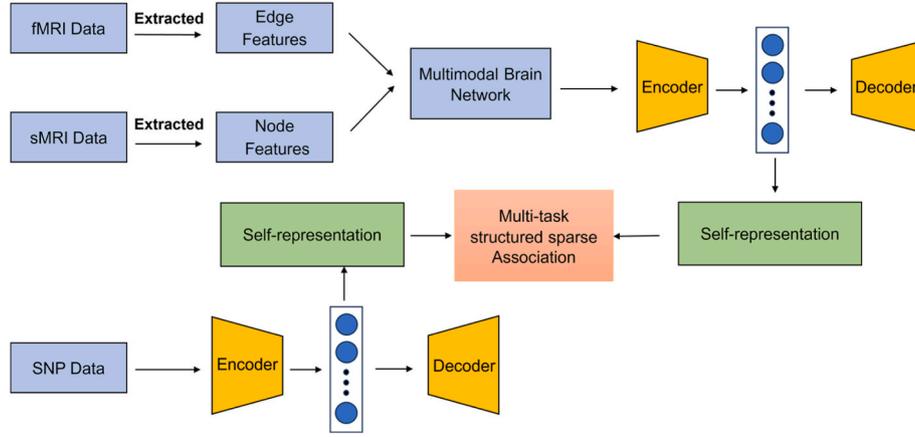


Fig. 2. The block diagram of the proposed MFASN framework.

angle values of the FCM were extracted as edge features for the brain network. At last, each subject has $(246 \times 245)/2 = 30135$ dimensional brain connectivity features.

Moreover, the voxel-based morphometry (VBM) was acquired by preprocessing sMRI data. In detail, the preprocessing steps included head motion correction, alignment, resampling, and image segmentation. Then, in the standard Montreal Neurological Institute (NMI) space, the normalized grey matter density maps were created as $2 \times 2 \times 2 \text{ mm}^3$ voxels and smoothed with an 8 mm FWHM kernel by SPM (Eickhoff et al., 2005). Based on the Brainnetome atlas, the 246 ROIs level measurements of mean grey matter densities were extracted as node features for the brain network. Fig. 3 visualized the association analysis between the multi-modal brain network phenotypes and genotypes.

3.2. Deep auto-encoder

The nonlinear manifold indicates that complex data in high-dimension space can be embedded in a low-dimensional space where it follows a nonlinear structure. To preserve the low-dimensional nonlinear manifold of the high dimensional neuroimaging genetics data, the auto-encoder as an effective deep learning technique was applied to achieve the non-linear mapping and extract the effective features (Fu et al., 2021). The typical architecture of the auto-encoder (AE) is composed of three parts: an encoder, a latent feature representation, and a decoder, which is often used for data reconstruction and non-linear dimensionality reduction (Wang et al., 2016). Moreover, in the typical architecture, the encoder and decoder are neural networks.

In general, the input feature vector $x \in R^p$ of SNP data or multi-modal brain network were encoded by the encoder blocks to obtain the latent feature representation $z \in R^t$,

$$z = \delta_e(\mathbf{W}_e x + b_e), \quad (1)$$

in which $\mathbf{W}_e \in R^{t \times p}$ denoted an encoding weight matrix and $b_e \in R^t$ denoted an encoding bias vector. Here, $z \in R^t$ as the latent feature representation is the output of the encoder blocks. $\delta_e(x)$ referred to an activation function.

Then, the decoding process utilized the latent representation z to reconstruct the original data. The reconstructive data $\hat{x} \in R^p$ close to the input x were defined as:

$$\hat{x} = \delta_d(\mathbf{W}_d z + b_d). \quad (2)$$

$\mathbf{W}_d \in R^{p \times t}$ and $b_d \in R^p$ denoted the weight matrix and bias vector in the decoding process, respectively. $\delta_d(x)$ was also an activation function.

For a better understanding, the output functions of every layer in the encoding process were given as:

$$\begin{cases} z_1 = \delta_1(\mathbf{W}_1 x + b_1), \\ z_k = \delta_k(\mathbf{W}_k z_{k-1} + b_k), \quad k = 2, 3, \\ z = \delta_4(\mathbf{W}_4 z_3 + b_4), \end{cases} \quad (3)$$

\mathbf{W}_1 , \mathbf{W}_k and \mathbf{W}_4 of which were related weight matrices. Further, b_1 , b_k and b_4 were bias vectors.

For parameters optimization, the reconstructive error between the input data $x \in R^p$ and the reconstructed data $\hat{x} \in R^p$ was used as the loss function, which was given as:

$$J_{AE} = \frac{1}{2} \sum_{i=1}^p \|\hat{x}_i - x_i\|^2. \quad (4)$$

AE can non-linearly project the input data to a latent space, in the case where the reconstructed data are closer to the input data. Meanwhile, it reflected the characteristics of imaging genetics data and removed redundant features by multiple mapping. This process was realized by optimizing parameters to ensure the minimum loss of the reconstructed data.

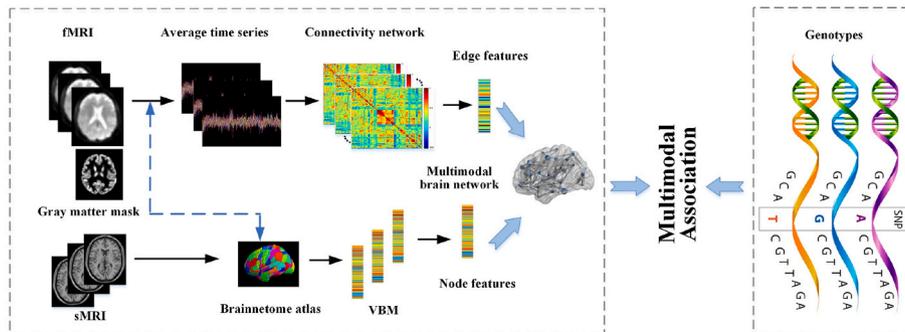


Fig. 3. The multi-modal data fusion-based imaging genetic association model.

AE took the multi-modal brain network phenotype and genotype to the input layer and then encoded data via three hidden layers. At last, the decoder used the learned latent features to reconstruct the original data. For AE, the dimension of the input brain network is 30381, so the hidden layer sizes of the encoder are set to 20480, 10240, and 4000, respectively, and the decoder settings are reversed from the encoder. The bottleneck layer size is 4000, indicating the dimension of the latent representation is 4000. Meanwhile, the learning rate is 0.001 and the training epochs are 500. Since the softplus function $\delta(x) = \log(1 + e^x)$ was similar with the activation models of brain neurons, it was selected as the activation function for both encoding and decoding processes.

3.3. Subspace self-representation network

To reflect the nonlinear distribution of data, a deep auto-encoder was utilized to nonlinearly project the input data to a latent space. As shown in Subsection 3.2, the latent feature matrix $\mathbf{H} = [h_1, h_2, \dots, h_N]^T \in R^{N \times t}$ was generated through this process. Meanwhile, to further reflect the block diagonal property of data between clusters, the self-representation network was employed to preserve the similarity structures of data before the association analysis (Wang et al., 2021). To achieve this idea, we performed subspace clustering on the learned latent feature matrix using the subspace self-representation network.

The learned latent feature matrix $\mathbf{H} = [h_1, h_2, \dots, h_N]^T \in R^{N \times t}$ was a collection of N data samples from multiple linear subspaces, so we can describe the data sample in a specific subspace using a linear combination of other data samples within the same subspace (Meng et al., 2021; Peng et al., 2020). Considering the subspaces are independent, we would obtain a self-representation matrix \mathbf{C}^h with a block-diagonal structure by minimizing certain norms of \mathbf{C}^h . Thus, for the latent representation \mathbf{H} , the self-representation property can be formulated as the following optimization problem:

$$\min_{\mathbf{C}^h} \|\mathbf{C}^h\|, \text{ s.t. } \mathbf{H} = \mathbf{H}\mathbf{C}^h, (\text{diag}(\mathbf{C}^h) = 0), \quad (5)$$

where $\|\cdot\|_l$ was an arbitrary matrix norm. The reconstruction coefficient matrix $\mathbf{C}^h \in R^{N \times N}$ was a symmetric matrix that captured the intrinsic geometric structure of data.

In the work, to further reflect the multi-subspace structures of data by the self-representation property, the diagnostic groups of data samples were used to reconstruct the original data. Given $\mathbf{H} = [h_1^{(1)}, \dots, h_{N_1}^{(1)}, \dots, h_1^{(k)}, \dots, h_{N_k}^{(k)}]^T \in R^{N \times t}$ with k classes, in which $h_i^{(d)}$ indicated the i -th subject in d -th diagnostic group. The subjects with the same diagnostic label were utilized to reconstruct $h_i^{(d)}$. Thus, $h_i^{(d)}$ represented by subjects with the same label was defined as:

$$\begin{aligned} \min_{(\mathbf{C}_i^h)^{(d)} \geq 0} & \frac{1}{2} \|h_i^{(d)} - \mathbf{H}^{(d)} (\mathbf{C}_i^h)^{(d)}\|_2^2 + \alpha \|(\mathbf{C}_i^h)^{(d)}\|_1 \\ & + \beta \sum_{j=1}^{p-1} \left| (\mathbf{C}_{i_{j+1}}^h)^{(d)} - (\mathbf{C}_{i_j}^h)^{(d)} \right|. \end{aligned} \quad (6)$$

Here, $(\mathbf{C}_i^h)^{(d)}$ is a similarity weight vector of i -th subject in d -th diagnosis group, which was constrained by both ℓ_1 -norm and the fused lasso. The first term prevented overfitting and improved the sparsity of the coefficients. Also, the second term encouraged sparsity of their differences, making weights for adjacent features similar, which is especially effective when the number of features is much more than subjects. α and β were sparse penalty parameters. For the weight vector $(\mathbf{C}_i^h)^{(d)}$, the elements in position $h_i^{(d)}$ were zero. The SLEP toolbox was utilized to get the

solution of the optimization problem in (6). When we obtained the $(\mathbf{C}_i^h)^{(d)}$, the sparse self-representation coefficient matrix \mathbf{C}^h was defined as follows:

$$\mathbf{C}^h = \left[(\mathbf{C}_1^h)^{(d)}, \dots, (\mathbf{C}_n^h)^{(d)} \right] + \left[(\mathbf{C}_1^h)^{(d)}, \dots, (\mathbf{C}_n^h)^{(d)} \right]^T. \quad (7)$$

Fig. 1C revealed the distribution of the input SNP data and the reconstructive SNP data through the t-distributed stochastic neighbour embedding (t-SNE) plot, which proved the better subspace clustering ability of the sparse self-representation network.

3.4. Multi-task structured sparse association model

Most of the current imaging genetics studies only considered the correlation analysis between single-modality imaging phenotype and genotype, failing to perform association analysis between the connectome (such as nodes and edges) of multi-modal brain networks and risk SNPs (Wang et al., 2021). Moreover, the multi-modal imaging genetics data exhibit non-linearity, characterized by their high-dimensional features and limited sample sizes, lying around a union of subspaces. It is necessary to develop a novel multi-modal data fusion framework for association analysis and extracting AD-related biomarkers.

Based on the above, sMRI and fMRI data were used to construct node features and edge features for the brain network, which integrated complementary information and helped identify consistent ROI features across modalities. Specifically, SNP data is $\mathbf{X} = [x_1, x_2, \dots, x_N]^T \in R^{N \times q}$ and M modality brain network is $\mathbf{Y}^m = [y_1^m, y_2^m, \dots, y_N^m]^T \in R^{N \times r}$ ($m = 1, \dots, M$), in which N refers to the number of subjects, q and r are the number of SNPs and network QTs features. After building the multi-modal brain network, the network QTs and SNPs were input to the deep AE model for dimension reduction and non-linear transformation. After that, to extract the similarity structure of data, according to Subsection 3.3, the sparse self-representation networks were constructed for latent feature matrices of every modality learned from the deep AE model, which were given as:

$$\mathbf{C}^x = \left[(\mathbf{C}_i^x)^1, \dots, (\mathbf{C}_i^x)^k \right] + \left[(\mathbf{C}_i^x)^1, \dots, (\mathbf{C}_i^x)^k \right]^T, \quad (8)$$

$$\mathbf{C}^m = \left[(\mathbf{C}_i^m)^1, \dots, (\mathbf{C}_i^m)^k \right] + \left[(\mathbf{C}_i^m)^1, \dots, (\mathbf{C}_i^m)^k \right]^T, \quad (9)$$

in which the reconstructed weight matrices \mathbf{C}^x and \mathbf{C}^m denoted $N \times N$ self-representative coefficient matrices for SNP data and the m -th modality brain network data, respectively. The self-representation coefficient matrix also reflected the pairwise affinity between subjects.

At last, the genotypic and multi-modality phenotypic projection matrix were computed by $f(\mathbf{X}) = \mathbf{C}^x \mathbf{X}$ and $g(\mathbf{Y}_m) = \mathbf{C}^m \mathbf{Y}_m$, respectively. Hence, based on the multi-task learning, the initial multi-modal data fusion-based imaging genetic association model was defined as:

$$\begin{aligned} \max_{u_m, v_m} & \sum_{m=1}^M u_m^T f(\mathbf{X})^T g(\mathbf{Y}_m) v_m, \\ \text{s.t. } & \|f(\mathbf{X}) u_m\|_2^2 = 1, \|g(\mathbf{Y}_m) v_m\|_2^2 = 1, \\ & \Omega(\mathbf{U}) \leq b_1, \Omega(\mathbf{V}) \leq b_2, \forall m, \end{aligned} \quad (10)$$

where $\mathbf{U} \in R^{q \times M}$ was a canonical weight matrix related to SNP data $f(\mathbf{X})$, and $\mathbf{V} \in R^{r \times M}$ denoted the weight matrix related to imaging QTs, in which each v_m was corresponded to $g(\mathbf{Y}_m)$. The constraints $\Omega(\mathbf{U})$ and $\Omega(\mathbf{V})$ were used to ensure sparsity and then identified some significant features. Also, the first two penalties were applied to enhance the covariance structure of data.

Since relations usually exist among SNPs in linkage disequilibrium (LD) blocks in the genome, it is necessary to explore the group-level or graph-level sparsity in the genetic data. To mine the structure information from SNPs, the fused pairwise group lasso (FGL) was applied to every u_m :

$$\|u_m\|_{FGL} = \sum_{i=1}^{q-1} \sqrt{u_{im}^2 + u_{(i+1)m}^2}. \quad (11)$$

Pseudo Code of Algorithm 1. MFASN Algorithm

Pseudo Code of Algorithm 1: MFASN Algorithm

Input: The SNP data $\mathbf{X} \in R^{N \times q}$ and Multi-modal network phenotype $\mathbf{Y}^m \in R^{N \times r}$ ($m = 1, \dots, M$). Subjects with labels (NC, SMC, EMCI, LMCI, and AD).

Output: Canonical weights u_m, v_m .

Optimization:

- 1: Initialize $\mathbf{W}_e, \mathbf{W}_d, b_e$ and b_d ;
 - 2: while not convergence do
 - 3: for each sample x do
 - 4: Calculate the hidden representation z by (1);
 - 5: Calculate the output \hat{x} according to (2);
 - 6: Update weight matrices and bias vectors by backpropagation;
 - 7: end
 - 8: end while
 - 9: Get self-representation matrices \mathbf{C}^x and \mathbf{C}^m by (8) and (9).
 - 10: Initialize $\mathbf{U} \in R^{r \times M}$ and $\mathbf{V} \in R^{r \times M}$;
 - 11: while not convergence do
 - 12: Update diagonal matrix \mathbf{D}_1 and solve u_m by (16);
 - 13: Scale u_m so that $\|f(\mathbf{X})u_m\|_2^2 = 1$;
 - 14: Update diagonal matrix \mathbf{D}_2 and solve v_m by (18);
 - 15: Scale v_m so that $\|g(\mathbf{Y}_m)v_m\|_2^2 = 1$;
 - 16: end while
-

The FGL penalty has merits in both group lasso and fused lasso (Du et al., 2020a,b,c). First, this penalty generated a chain of smoothness over all values of u_m , which can select two adjacent and strongly associated SNPs. Moreover, with the ℓ_2 -norm, the FGL regularized term was sample correlation sign independent. Meanwhile, we also imposed the $\ell_{1,1}$ -norm (ℓ_1 -norm for matrices) penalty on multi-modal imaging QTs to ensure individual-level sparsity.

$$\|\mathbf{V}\|_{1,1} = \sum_{j=1}^r \|v_j\|_1 = \sum_{j=1}^r \sum_{m=1}^M |v_{jm}|. \quad (12)$$

The $\ell_{1,1}$ -norm penalty was effective as it can select important imaging QTs and explore some imaging biomarkers related to a specific imaging modality.

Accordingly, to better perform association analysis between SNP data and multi-modal imaging data, the final multi-task structured sparse association model was defined as:

$$\max_{u_m, v_m} \sum_{m=1}^M u_m^T f(\mathbf{X})^T g(\mathbf{Y}_m) v_m - \lambda_u \sum_{m=1}^M \|u_m\|_{FGL} - \lambda_v \|\mathbf{V}\|_{1,1}, \quad (13)$$

$$s.t. \|f(\mathbf{X})u_m\|_2^2 = 1, \|g(\mathbf{Y}_m)v_m\|_2^2 = 1, \forall m,$$

where λ_u and λ_v were penalty parameters selected by the nested cross-validation on the training set. Meanwhile, u_m and v_m denoted weight vectors that revealed the significance of SNPs and imaging QTs, respectively.

3.5. Optimization Algorithm

For the problem in (13), the \mathbf{C}^x and \mathbf{C}^m for SNPs and multi-modal network QTs were calculated according to (6) and (7). Next, the Lagrange multiplier approach was used to transform (13) into an unconstrained issue and then drop constants. The minimization problem was expressed as follows:

$$\min_{u_m, v_m} \sum_{m=1}^M \left[\|f(\mathbf{X})u_m - g(\mathbf{Y}_m)v_m\|_2^2 + \gamma_u \|f(\mathbf{X})u_m\|_2^2 + \gamma_v \|g(\mathbf{Y}_m)v_m\|_2^2 \right] + \lambda_u \sum_{m=1}^M \|u_m\|_{FGL} + \lambda_v \|\mathbf{V}\|_{1,1}, \quad (14)$$

where $\lambda_u, \lambda_v, \gamma_u$ and γ_v were parameters to control global sparsity. After the experimental verification, results were insensitive to γ_u and γ_v settings. Since existing methods have studied a similar problem, we set γ_u and γ_v to the fixed values for simplicity based on related studies (Wang et al., 2021) here. This problem can be achieved convex by solving one variable with other variables remaining fixed. Hence, we can use the alternative update rule to solve this optimization problem.

Updating u_m : Take the derivative of the Lagrange function in (14) for u_m by fixing v_m and let it be zero, we can obtain:

$$(\gamma_u + 1)f(\mathbf{X})^T f(\mathbf{X})u_m + \lambda_u \mathbf{D}_1 u_m - f(\mathbf{X})^T g(\mathbf{Y}_m)v_m = 0, \quad (15)$$

Table 1

The detail information of the ADNI datasets.

Variables	NC	SMC	EMCI	LMCI	AD
Subject Number	37	14	41	31	26
Gender (male/female)	16/21	5/9	17/24	20/11	11/15
Age (mean \pm std)	75.96 \pm 7.04	73.46 \pm 5.52	71.57 \pm 5.84	71.76 \pm 7.62	72.65 \pm 7.45

where \mathbf{D}_1 was a diagonal matrix, the i -th diagonal value of which was mainly $1/(2\sqrt{u_{(i-1)m}^2 + u_{im}^2}) + 1/(2\sqrt{u_{im}^2 + u_{(i+1)m}^2})$ ($i = 1, \dots, q; m = 1, \dots, M$). By the iterative algorithm, the related updating formula for u_m was given here:

$$u_m = \frac{f(\mathbf{X})^T g(\mathbf{Y}_m) v_m}{\lambda_u \mathbf{D}_1 + (\gamma_u + 1) f(\mathbf{X})^T f(\mathbf{X})}. \quad (16)$$

Updating v_m : Similarly, we can calculate the derivative of v_m by fixing u_m and set it to zero, then we can obtain:

$$-g(\mathbf{Y}_m)^T f(\mathbf{X}) u_m + \lambda_v \mathbf{D}_2 v_m + (\gamma_v + 1) g(\mathbf{Y}_m)^T g(\mathbf{Y}_m) v_m = 0, \quad (17)$$

where \mathbf{D}_2 was a diagonal matrix with its j -th diagonal value being $1/2\|v_{jm}\|_2$ ($j = 1, \dots, r; m = 1, \dots, M$) (Du et al., 2020a,b,c). Hence, the related updating formula for v_m was given as:

$$v_m = \frac{g(\mathbf{Y}_m)^T f(\mathbf{X}) u_m}{\lambda_v \mathbf{D}_2 + (\gamma_v + 1) g(\mathbf{Y}_m)^T g(\mathbf{Y}_m)}. \quad (18)$$

Obviously, \mathbf{D}_1 and \mathbf{D}_2 were dependent on \mathbf{U} and \mathbf{V} that were unknown. First, the initial values of \mathbf{U} and \mathbf{V} were given, and then the diagonal matrices of them were calculated. This process was stopped until the predefined stopping criteria were met. Algorithm 1 gave the pseudo code of the optimization algorithm for MFASN method.

4. Experiment

4.1. ADNI imaging genetic data

The multi-modal neuroimaging data (sMRI, fMRI) and SNP data were from the ADNI database (adni.loni.usc.edu). In 2003, Principal Investigator Michael W. Weiner, MD developed ADNI. It aimed to get clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of AD. For up-to-date information, see www.adni-info.org.

In this study, data samples contained three parts: fMRI data, sMRI data, and SNP data. 149 valid samples were obtained by aligning the SNP data and multi-modal imaging data, and then removing samples with missing values. At last, there were 26 AD, 37 normal control (NC), 14 significant memory concerns (SMC), 41 early mild cognitive impairment (EMCI), and 31 late mild cognitive impairment (LMCI). For the genetic data, 85 SNPs gathered from the *APOE* gene's neighbours were applied in the research, including the famous AD-related risk SNP rs429358. Detailed information of data has been listed in Table 1.

4.2. Experimental settings

In the experimental part, we utilized the Pearson correlation coefficient (CC) to assess the association performance. CC is the unique

Table 2

Association performance on node features and edge features by different methods.

Methods	Modality	CC Results	
		Train (mean \pm SD)	Test (mean \pm SD)
SCCA	node	0.4669 \pm 0.0001	0.2812 \pm 0.0001
	edge	0.2547 \pm 0.0001	0.1585 \pm 0.0003
KSCCA	node	0.3021 \pm 0.0001	0.3002 \pm 0.0001
	edge	0.5567 \pm 0.0001	0.4145 \pm 0.0006
DCCA	node	0.7433 \pm 0.0001	0.6703 \pm 0.0003
	edge	0.8900 \pm 0.0001	0.8131 \pm 0.0002
DS-SCCA	node	0.7426 \pm 0.0001	0.6132 \pm 0.0008
	edge	0.7280 \pm 0.0001	0.7218 \pm 0.0001
mSCCA	node	0.3346 \pm 0.0002	0.1987 \pm 0.0047
	edge	0.2812 \pm 0.0002	0.1544 \pm 0.0043
MTSCCA	node	0.3867 \pm 0.0001	0.2629 \pm 0.0017
	edge	0.3906 \pm 0.0007	0.2484 \pm 0.0013
AdaSMCCA	node	0.4825 \pm 0.0002	0.2927 \pm 0.0026
	edge	0.4058 \pm 0.0006	0.1574 \pm 0.0019
FGL-SCCA	node	0.8318 \pm 0.0003	0.6529 \pm 0.0002
	edge	0.8448 \pm 0.0002	0.6748 \pm 0.0003
MFASN	node	0.7433 \pm 0.0001	0.7088 \pm 0.0003
	edge	0.8863 \pm 0.0001	0.8529 \pm 0.0004

measure to predict the genotype-phenotype association performance in almost all imaging genetics studies (Du et al., 2019; Wang et al., 2019; Zhang et al., 2020).

To improve the significance of the performance, 20 times independent 5-fold cross-validation (5-CV) was carried out to evaluate the effectiveness of the proposed approach. For the 5-CV each time, the original dataset was divided into five independent subsets, in each cross-validation, one of which was used as the testing set, while the remaining four subsets were used as the training set. To control the sparsity of the coefficient matrix in (6), the grid search method was utilized to obtain the optimal combination of two self-representation parameters α and β within a specified range [0.1:0.1:0.9]. As revealed in Fig. 4, from the experimental results on node features and edge features, we found that different parameter combinations have little effect on the final prediction performance. Based on the parameter selection process, we set $\alpha = 0.6$ and $\beta = 0.1$ in the association tasks on node features and set $\alpha = 0.8$ and $\beta = 0.2$ on edge features, respectively. Meanwhile, the regularization parameters (including λ_u and λ_v) in the multi-task structured sparse association model were optimally tuned using a grid search from the range of {0.001, 0.01, 0.1, 1, 10} by another nested 5-CV on the training data. In the nested 5-CV, the current training set was further divided into five subsets. One of these subsets was used as the validation set, while the remaining four subsets were used as training sets. Different parameter combinations were used to train the model in the nested cross-validation, and the performance was evaluated on the internal validation set.

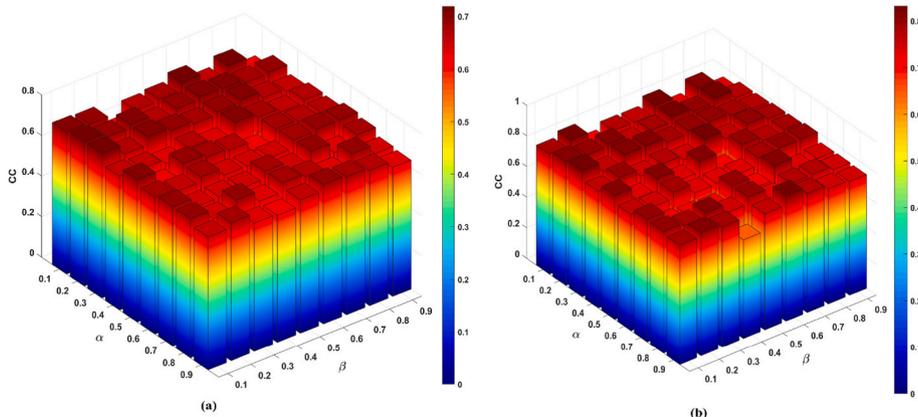


Fig. 4. Experimental results with different self-representation parameters α and β on (a) node features and (b) edge features.

Meanwhile, eight advanced methods were compared in this experiment, including SCCA (Chi et al., 2013), DCCA (Andrew et al., 2013), KSCCA (Melzer et al., 2003), DS-SCCA (Wang et al., 2022), mSCCA (Hao et al., 2017), MTSCCA (Du et al., 2019), FGL-SCCA (Du et al., 2020a,b, c), and AdaSMCCA (Du et al., 2021). The hyper-parameters of all methods were adjusted according to the related papers and the best results were obtained.

4.3. Results of comparative experiments

The association results of the proposed MFASN were compared with eight advanced methods. The average training CC and testing CC on node and edge features were listed in Table 2. It is noticed that both training and testing CC results on edge features of most approaches were better than results on node features, which proved the importance of functional connectivity information among brain areas. That is, edge features provided rich prior knowledge for early diagnosis of AD.

As revealed in Table 2, the traditional SCCA and mSCCA methods are difficult to handle with high-dimensional multi-modal data, which shows poor association performance. FGL-SCCA recovered graphical smoothness and structure information by introducing improved fused penalty and graph-guided penalty, which achieved higher and more robust association performance. Compared with two-view SCCA, multi-modal SCCA models have obtained better CC results, especially MTSCCA and AdaSMCCA. These results indicated that multi-task learning can learn complementary information from distinct imaging modalities, which performed better than approaches learning each task separately. Moreover, deep non-linear methods including DS-SCCA, KSCCA, and DCCA outperformed other linear SCCA-based approaches significantly. The reason may be that real data may not be linearly represented by each other in the space thereby reducing performance. The deep extension methods can nonlinearly project the original data to the latent space and capture the effective features of data, thus having superior results. More importantly, whether for the node modality or edge modality, our proposed MFASN method obtained the best training and testing CC results, which proved that our method was an effective attempt to achieve multi-modal data fusion. The self-representation network was applied for the latent representation from the auto-encoder, which can make the subjects of the same subspace with nonlinear structures more clustered in the low-dimensional space. Through deep subspace clustering, we reconstructed the multi-modal imaging genetics data by using a new data similarity matrix with sparsity and nonlinearity principles.

In the following biological application experiment, we also proved that the proposed MFASN model not only obtained the best association performance but also identified the consistent imaging connectome and genetic biomarkers for AD diagnosis.

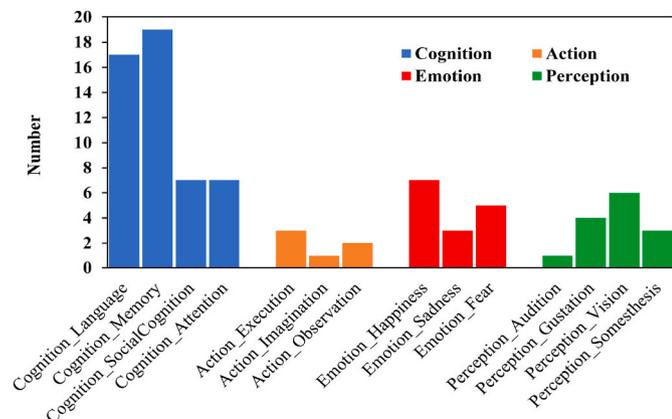


Fig. 5. A summary of functional annotations of top 100 ROIs.

Table 3

Top 10 ROIs selected by the MFASN. L = Left; R=Right.

ID	ROI	Gyrus	Studies
212	mAmyg. R	Amygdala	Poulin et al. (2011)
211	mAmyg. L	Amygdala	Poulin et al. (2011)
218	cHipp. R	Hippocampus	Pennanen et al. (2004)
213	lAmyg. L	Amygdala	Guo et al. (2018)
110	A35/36r. R	Parahippocampal	Echavarrri et al. (2011)
99	A20cl. L	Inferior Temporal	Scheff et al. (2011)
117	TI. L	Parahippocampal	Mitchell et al. (2002)
109	A35/36r. L	Parahippocampal	Echavarrri et al. (2011)
215	rHipp. L	Hippocampus	Pasquini et al. (2015)
116	A28/34. R	Parahippocampal	McLachlan et al. (2018)

4.4. Identification of risk SNP loci and ROIs

The important goal of this research is to identify some risk SNPs and crucial brain regions related to AD. Specifically, the sparse correlation weights by 5-CV of MFASN on the node modality were averaged to select significant ROIs. First, Fig. 5 showed the functional annotation of the top 100 important ROIs. We found that these ROIs exhibited higher correlation levels in functions related to cognition, memory, and language, which were consistent with the symptoms of AD. Second, details of the top 10 crucial brain imaging markers were listed in Table 3. These significant ROIs were also visualized by the BrainNet Viewer shown in Fig. 6A. The identified fine-grained subregions included *cHipp*, *rHipp*, *mAmyg*, and *lAmyg* that belong to the *Hippocampus* gyrus (Pennanen et al., 2004) and *Amygdala* gyrus (Poulin et al., 2011), which were strongly associated with the memory function of AD. Pennanen et al. proved that the entorhinal cortex atrophy precedes hippocampal atrophy in AD (Pennanen et al., 2004). Poulin et al. also found that the magnitude of amygdala atrophy was comparable to that of the hippocampus in the earliest clinical stages of AD, and was associated with global illness severity (Poulin et al., 2011). Meanwhile, researchers have found that the *Parahippocampal* volume was more discriminative than the *Hippocampus* volume in the early stage of the disease, which was proven as an early diagnostic marker of AD (van Hoesen et al., 2000). Its fine-grained subregions contained *A35/36r*, *TI*, and *A28/34*. Further, the discovered subdivided brain region *A20cl* from the *Inferior Temporal* gyrus (Scheff et al., 2011) also played a major role in verbal fluency, a cognitive function affected early in the onset of AD.

Also, the top 10 risk SNPs related to AD have been listed in Table 4. It can be noticed that MFASN identified meaningful and consistent AD-relevant SNPs including *rs429358* (*APOE*) (Kulminski et al., 2020), *rs118170342* (*TOMM40*) (Roses et al., 2016), and other significant SNPs *rs12691088*, *rs72654473* and *rs114536010* from the disease-related genes *APOE*, *APOC1* and *TOMM40*. More importantly, existing studies confirmed that *APOE* and *APOC1* genes played important parts in the AD development process (Kulminski et al., 2018). Especially, *APOE* was especially related to increased risk of AD, and *rs429358* remained the most notable risk factor for AD development. Roses et al. proved that *TOMM40* is active in AD by regulating mitochondrial biogenesis. *NECTIN2* was located on both sides of *TOMM40*. The pathogen hypothesis thought that *NECTIN2* was a pathogenic factor of AD (Curtis and Initiative, 2021). To further validate associations between the top 10 brain ROIs and SNPs, Fig. 6B showed the pairwise correlation of each SNP-ROI pair. As expected, most SNP-ROI associations obtained considerable correlation coefficients. Fig. 6B also revealed that the effects of the same genetic loci on distinct brain areas have similar positive or negative relations on the whole. Some significant SNP-ROI pairs were plotted in Fig. 6D. Among them, different colour represented brain regions with different correlation levels.

4.5. Connectivity analysis

Since a brain network is a depiction of a brain system, the functional

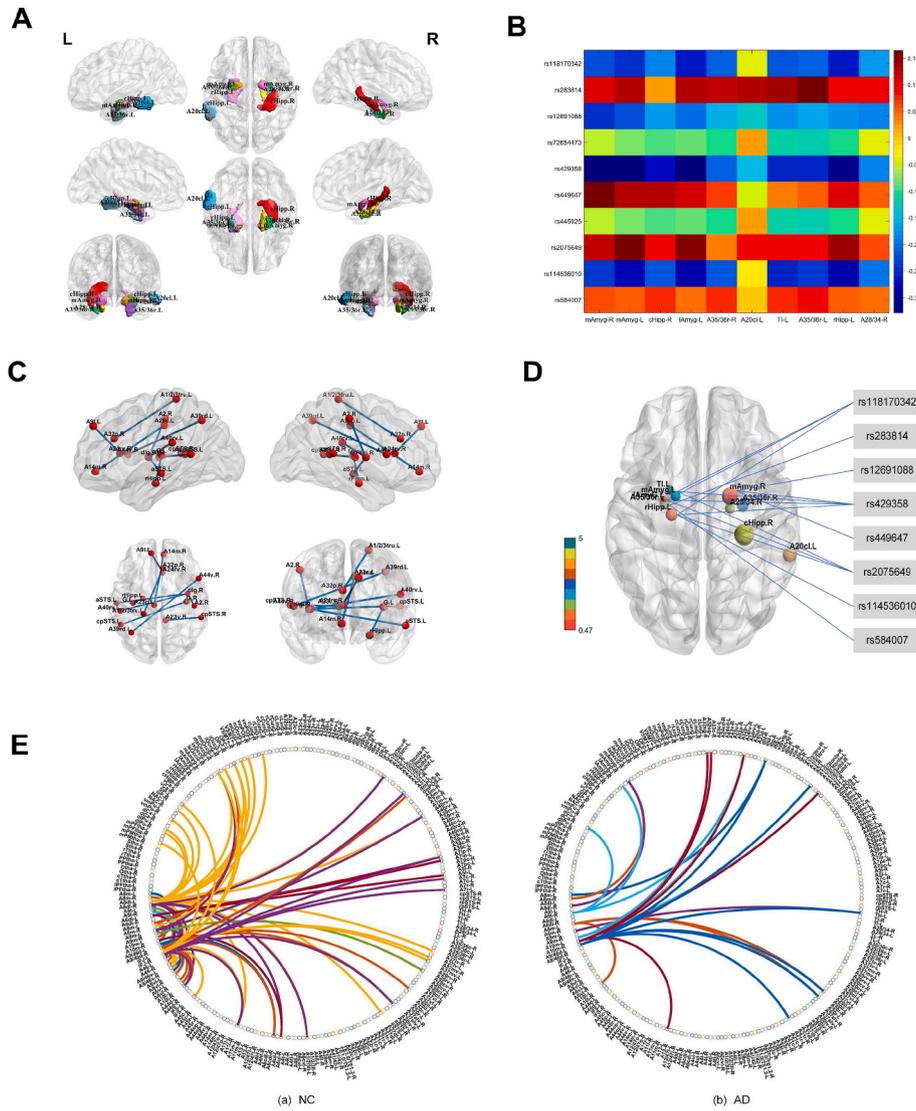


Fig. 6. The comprehensive biological analysis of the proposed MFASN.

Table 4
Top 10 SNPs selected by the proposed method.

ID	SNP	Gene	Studies
65	rs118170342	TOMM40	Roses et al. (2016)
50	rs283814	NECTIN2	Mizutani et al. (2022)
78	rs12691088	APOC1	Liu et al. (2021)
28	rs72654473	APOE	Shi and Holtzman (2018)
27	rs429358	APOE	Kulminski et al. (2020)
71	rs449647	APOE	Shi and Holtzman (2018)
29	rs445925	APOC1	Ki et al. (2002)
51	rs2075649	TOMM40	Lin et al. (2016)
64	rs114536010	TOMM40	Roses et al. (2010)
77	rs584007	APOC1	Kulminski et al. (2022)

connectivity values denote the edge weights of the brain network, which also reflects the importance of edges for association analysis. According to the obtained average weight coefficients by 5-CV on the edge modality, our MFASN selected the top 10 edges with maximum weight, which were listed in Table 5 and visualized in Fig. 6C.

The popular default mode network (DMN) mainly included the posterior cingulate cortex, precuneus, inferior parietal lobule, bilateral temporal cortex, and medial prefrontal cortex. Previous literature concluded that local abnormal activities of DMN may result in AD,

Table 5
Top 10 edges selected by the MFASN. L. = Left; R. = Right.

ID	ROI 1	ID	ROI 2
172	R. Insular Gyrus (dlg. R)	145	L. Inferior Parietal Lobule (A40rv. L)
137	L. Inferior Parietal Lobule (A39rd. L)	40	R. Inferior Frontal Gyrus (A44v. R)
178	R. Cingulate Gyrus (A24rv. R)	5	L. Superior Frontal Gyrus (A9l. L)
172	R. Insular Gyrus (dlg. R)	87	L. Middle Temporal Gyrus (aSTS. L)
164	R. Insular Gyrus (G. R)	123	L. posterior Superior Temporal Sulcus (cpSTS. L)
215	L. Hippocampus (rHipp. L)	163	L. Insular Gyrus (G. L)
185	L. Cingulate Gyrus (A23c. L)	42	R. Orbital Gyrus (A14m. R)
172	R. Insular Gyrus (G. R)	160	R. Postcentral Gyrus (A2. R)
180	L. Cingulate Gyrus (A23c. L)	161	L. Postcentral Gyrus (A1/2/3tru. L)
182	L. Cingulate Gyrus (A23c. L)	124	L. posterior Superior Temporal Sulcus (rpSTS. L)

depression, and even schizophrenia (Koch et al., 2012). As expected, our biological experimental results proved that the connections between DMN and other brain areas (Hippocampus, Insular gyrus, Postcentral, and Frontal gyrus) are closely associated with AD. Related results also revealed that DMN was a momentous functional system in the brain. Apart from this, some brain regions such as the Superior Frontal gyrus and

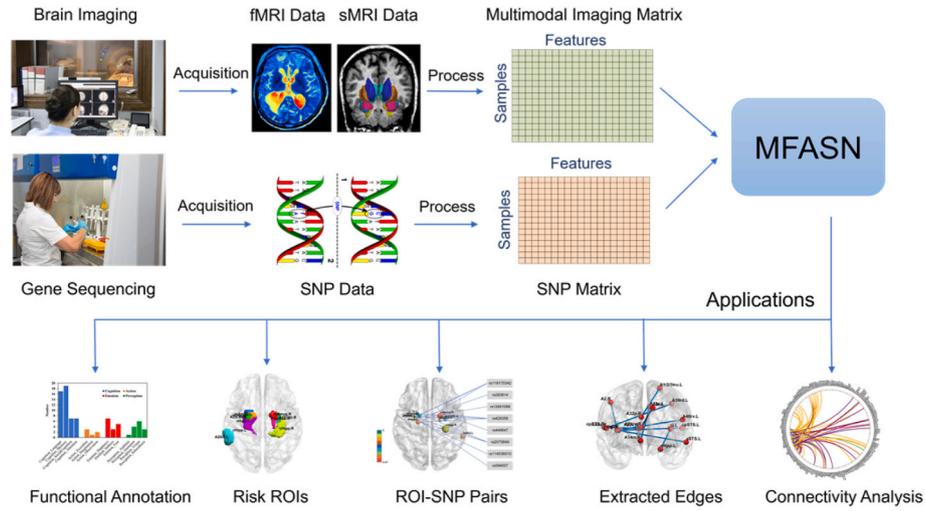


Fig. 7. The real applications of the proposed MFASN.

Table 6

CC results of different correlation methods for brain network construction on node features and edge features.

Methods	Modality	CC Results	
		Train (mean \pm SD)	Test (mean \pm SD)
KCC	node	0.8145 \pm 0.0001	0.7898 \pm 0.0001
	edge	0.8353 \pm 0.0001	0.7954 \pm 0.0001
SCC	node	0.7091 \pm 0.0001	0.6593 \pm 0.0003
	edge	0.7493 \pm 0.0001	0.6753 \pm 0.0005
PCC	node	0.7433 \pm 0.0001	0.7088 \pm 0.0003
	edge	0.8863 \pm 0.0001	0.8529 \pm 0.0004

Inferior Frontal gyrus were usually related to mild cognitive impairment. Studies (Guo et al., 2016) also indicated that the *Postcentral gyrus* was commonly related to the depressive symptom severity of AD. And the unawareness of deficits in early AD patients was related to the decreased function of the *Cingulate gyrus*. Furthermore, the *Insular gyrus* played an important role in AD subjects who have behavior out of control and visceral dysfunction (Kitamura et al., 2020). Research (Wang et al., 2011) has found decreased connectivity of the *Insular-Hippocampus gyrus* in MCI patients. A short-term memory task researched in AD samples was also related to a brain subnetwork including *bilateral Inferior Frontal, Insular, and Postcentral* regions (Peters et al., 2009). All in all, these potential associations between genotype and phenotype proved that our MFASN was stable and effective in the multi-modal imaging genetics association study.

Lastly, to explore the development pattern of AD based on the functional connectivity information, the reconstructed average functional connections of AD and NC patients were shown in Fig. 6E. As can be found, AD patients revealed a loss of functional connectivity compared to NC individuals, which reflected further degeneration of brain functions. More importantly, most brain areas connected by functional connectivity were also verified by the extracted significant brain regions. This further demonstrated the effectiveness of the developed multi-modal data fusion model. Meanwhile, the overall applications framework of our proposed research model was shown in Fig. 7.

4.6. Correlation methods for brain network construction

Moreover, to verify the effectiveness of the Pearson correlation coefficient (PCC) for multimodal brain network construction, the comparison experiments between PCC and two common distance correlation measurements were done in this work. The different distance correlation

Table 7

Ablation experimental results with five-fold cross-validation on node features and edge features.

Method		Train (mean \pm SD)		Test (mean \pm SD)	
		CC	P-value	CC	P-value
Degrade 1	node	0.4121 \pm 0.0001	1.19e-28	0.2490 \pm 0.0055	1.85e-16
	edge	0.3582 \pm 0.0002	1.51e-30	0.1414 \pm 0.0027	5.21e-23
Degrade 2	node	0.4421 \pm 0.0001	6.49e-30	0.2795 \pm 0.0039	2.47e-17
	edge	0.4243 \pm 0.0002	5.58e-29	0.2195 \pm 0.0018	1.38e-26
Degrade 3	node	0.4631 \pm 0.0001	1.43e-32	0.4378 \pm 0.0003	5.08e-21
	edge	0.6502 \pm 0.0002	8.19e-24	0.6219 \pm 0.0011	3.78e-17
MFASN	node	0.7433 \pm 0.0001	–	0.7088 \pm 0.0003	–
	edge	0.8863 \pm 0.0001	–	0.8529 \pm 0.0004	–

measuring methods included the Kendall correlation coefficient (KCC), Spearman correlation coefficient (SCC), and Pearson correlation coefficient (PCC). The detailed results were shown in Table 6. As revealed in Table 6, we observed that the PCC obtained the best association performance in comparison with the KCC and SCC. In detail, the above three distance correlation measures have achieved better experimental results, especially for KCC and PCC. However, compared to PCC, the CC results of KCC in the node modality are better than PCC, but in edge modality is lower than PCC. Considering the overall level, we chose PCC to build the multimodal brain network.

4.7. Ablation experiments

There were three key components in our proposed method, including the deep auto-encoder network, the self-representation subspace learning, and the multi-task structured sparse association model. Here, to evaluate the effect of different components in MFASN, we conducted the ablative experiment.

These three unique variants were summarized below. *First*, to verify the performance of the multi-task structured sparse association model, that is, removing the deep auto-encoder and self-representation network, we obtained Degrade 1. *Second*, to learn the multi-subspace mapping ability of the self-representation module, the loss function of MFASN without self-representation was called Degrade 2. *Third*, to

prove the non-linear projection ability of the auto-encoder, the loss function of MFASN without a deep auto-encoder was named Degrade 3. Table 7 listed the association results of the three cases. Also, the pairwise *t*-test based on the 20 times 5-CV results was performed and the *p*-values were also listed in Table 7.

From these results, the following conclusions are given here. (1) Degrade 1 was a simple multi-task structured sparse association method, which seemed a basic multi-modal SCCA model that had poor association results, but was also effective on multi-modal data. (2) Degrade 2 contained deep auto-encoder, as an effective neural network, which nonlinearly mapped the input data to a latent space and selected informative features, helping for later association analysis. (3) Degrade 3 contained the sparse self-representation module, which was used to perform subspace clustering and further learn similarity structures of data. The Degrade 3 with higher results confirmed that self-representation module played a key role in our work. (4) Our presented MFASN performed better than the three degraded models and the *p*-values ($p < 0.05$) displayed that all results were statistically significant, which revealed the significance and effectiveness of MFASN.

5. Conclusion

In this study, we developed a multi-modal imaging genetics data fusion framework for AD diagnosis and biomarkers identification. This work mainly utilized brain imaging technology to evaluate the genetic influence on individuals by using brain structure and function as phenotypes, and explored how SNPs affect the neural structure and function of the brain. Specifically, we integrated sMRI and fMRI data to construct a multi-modality brain network. Then, this model employed the deep auto-encoder to select the informative features. The self-representation network was developed to explore the complex multi-subspace structures of data. Lastly, a novel multi-task structured sparse association model was proposed to discover correlations between the multimodal brain network and SNP data. In the experiments, first, on the real ADNI datasets with a total of 149 valid subjects, MFASN achieved the correlation coefficient 0.7088 of SNPs-ROIs and 0.8529 of SNPs-Brain connectivity, which were higher than other advanced methods. This also proved that our method was an effective attempt to achieve multi-modal data fusion. More importantly, MFASN identified significant ROIs *Amygdala* and *Hippocampal*, and risk SNPs *rs429358* and *rs118170342*, which were crucial for early diagnosis of AD. Third, by revealing the pairwise correlations of SNP-ROI pairs, our study proved that the effects of the same SNP on different brain regions were basically related in the same direction. Finally, since alteration in brain functional connectivity is expected to provide potential biomarkers for AD prediction, our study also identified the popular DMN, the abnormal activities of which may result in AD. The above intuitive biological analysis can make it easier for us to interpret AD pathology-related problems.

Although MFASN can improve the association performance and identify risk biomarkers, the deep learning module can be further improved. In future work, first, we will focus on optimizing the iteration process of the presented method to reduce the computational time and further form an end-to-end deep learning framework to perform brain imaging genetics analysis. Second, restricted by the current experimental conditions, some unproven pathogenic biomarkers need to be studied in subsequent work. Thus, we will utilize more data modalities for more comprehensive analysis, and cooperate with clinicians to facilitate new biomarkers and data updates in our future schedule.

CRedit authorship contribution statement

Cui-Na Jiao: Methodology, Data curation, Software, Investigation, Writing – original draft, Visualization. **Ying-Lian Gao:** Conceptualization, Writing – review & editing, Validation, Supervision. **Dao-Hui Ge:** Conceptualization, Writing – review & editing, Validation, Supervision. **Junliang Shang:** Writing – review & editing, Validation, Supervision.

Jin-Xing Liu: Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 62172254). The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Almalki, Y.E., et al., 2021. A novel method for COVID-19 diagnosis using artificial intelligence in chest X-ray images. *Healthcare*. MDPI 9 (5), 522.
- Andrew, G., et al., 2013. Deep canonical correlation analysis. In: *International Conference on Machine Learning*. PMLR, pp. 1247–1255.
- Arifoglu, D., et al., 2020. Detecting indicators of cognitive impairment via graph convolutional networks. *Eng. Appl. Artif. Intell.* 89, 103401.
- Bi, X.-a., et al., 2020. Morbigenous brain region and gene detection with a genetically evolved random neural network cluster approach in late mild cognitive impairment. *Bioinformatics* 36 (8), 2561–2568.
- Chi, E.C., et al., 2013. Imaging genetics via sparse canonical correlation analysis. In: *2013 IEEE 10th International Symposium on Biomedical Imaging*. IEEE, pp. 740–743.
- Curtis, D., Initiative, A.S.D.N., 2021. Analysis of whole genome sequenced cases and controls shows that the association of variants in TOMM40, BCAM, NECTIN2 and APOC1 with late onset Alzheimer's disease is driven by linkage disequilibrium with APOE $\epsilon 2/\epsilon 3/\epsilon 4$ alleles. *J. Neurogenet.* 35 (2), 59–66.
- Du, L., et al., 2020a. Identifying diagnosis-specific genotype-phenotype associations via joint multitask sparse canonical correlation analysis and classification. *Bioinformatics* 36 (Suppl. ment 1), i371–i379.
- Du, L., et al., 2020b. Associating multi-modal brain imaging phenotypes and genetic risk factors via a dirty multi-task learning method. *IEEE Trans. Med. Imag.* 39 (11), 3416–3428.
- Du, L., et al., 2019. Multi-task sparse canonical correlation analysis with application to multi-modal brain imaging genetics. *IEEE ACM Trans. Comput. Biol. Bioinf* 18 (1), 227–239.
- Du, L., et al., 2020c. Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach. *Med. Image Anal.* 61, 101656.
- Du, L., et al., 2018. A novel SCCA approach via truncated ℓ_1 -norm and truncated group lasso for brain imaging genetics. *Bioinformatics* 34 (2), 278–285.
- Du, L., et al., 2021. Identifying associations among genomic, proteomic and imaging biomarkers via adaptive sparse multi-view canonical correlation analysis. *Med. Image Anal.* 70, 102003.
- Echavarrri, C., et al., 2011. Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease. *Brain Struct. Funct.* 215 (3), 265–271.
- Eickhoff, S.B., et al., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25 (4), 1325–1335.
- Erdogmus, P., Kabakus, A.T., 2023. The promise of convolutional neural networks for the early diagnosis of the Alzheimer's disease. *Eng. Appl. Artif. Intell.* 123, 106254.
- Fan, L., et al., 2016. The human brainnetome atlas: a new brain atlas based on connective architecture. *Cerebr. Cortex* 26 (8), 3508–3526.
- Farid, H.M.A., et al., 2022. Linear diophantine fuzzy fairly averaging operator for suitable biomedical material selection. *Axioms* 11 (12), 735.
- Fu, S., et al., 2021. A re-optimized deep auto-encoder for gas turbine unsupervised anomaly detection. *Eng. Appl. Artif. Intell.* 101, 104199.
- Guo, Z., et al., 2016. Abnormal degree centrality in Alzheimer's disease patients with depression: a resting-state functional magnetic resonance imaging study. *Exp. Gerontol.* 79, 61–66.
- Guo, Z., et al., 2018. Abnormal changes in functional connectivity between the amygdala and frontal regions are associated with depression in Alzheimer's disease. *Neuroradiology* 60 (12), 1315–1322.
- Hao, X., et al., 2017. Mining outcome-relevant brain imaging genetic associations via three-way sparse canonical correlation analysis in Alzheimer's disease. *Sci. Rep.* 7 (1), 1–12.
- Hao, X., et al., 2016. Identifying multimodal intermediate phenotypes between genetic risk factors and disease status in Alzheimer's disease. *Neuroinformatics* 14, 439–452.

- Irfan, M., et al., 2021. Role of hybrid deep neural networks (HDNNs), computed tomography, and chest X-rays for the detection of COVID-19. *Int. J. Environ. Res. Publ. Health* 18 (6), 3056.
- Jiang, R., et al., 2020. Gender differences in connectome-based predictions of individualized intelligence quotient and sub-domain scores. *Cerebr. Cortex* 30 (3), 888–900.
- Kausar, R., et al., 2022. Cancer therapy assessment accounting for heterogeneity using q-rung picture fuzzy dynamic aggregation approach. *Symmetry* 14 (12), 2538.
- Kausar, R., et al., 2023. Innovative CODAS algorithm for q-rung orthopair fuzzy information and cancer risk assessment. *Symmetry* 15 (1), 205.
- Ki, C.-S., et al., 2002. Genetic association of an apolipoprotein C1 (APOC1) gene polymorphism with late-onset Alzheimer's disease. *Neurosci. Lett.* 319 (2), 75–78.
- Kitamura, J., et al., 2020. The insular cortex, Alzheimer disease pathology, and their effects on blood pressure variability. *Alzheimer Dis. Assoc. Disord.* 34 (3), 282–291.
- Koch, W., et al., 2012. Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease. *Neurobiol. Aging* 33 (3), 466–478.
- Kulminski, A.M., et al., 2018. Apolipoprotein E region molecular signatures of Alzheimer's disease. *Aging Cell* 17 (4), e12779.
- Kulminski, A.M., et al., 2022. Definitive roles of TOMM40-APOE-APOC1 variants in the Alzheimer's risk. *Neurobiol. Aging* 110, 122–131.
- Kulminski, A.M., et al., 2020. Genetic and regulatory architecture of Alzheimer's disease in the APOE region. *Alzheimer's Dementia: Diagnosis, Assessment & Disease Monitoring* 12 (1), e12008.
- Lei, B., et al., 2020. Adaptive sparse learning using multi-template for neurodegenerative disease diagnosis. *Med. Image Anal.* 61, 101632.
- Lin, R., et al., 2016. Association of common variants in TOMM40/APOE/APOC1 region with human longevity in a Chinese population. *J. Hum. Genet.* 61 (4), 323–328.
- Liu, H., et al., 2021. Association between polygenic risk score and the progression from mild cognitive impairment to Alzheimer's disease. *J. Alzheim. Dis.* 84 (3), 1323–1335.
- McLachlan, E., et al., 2018. Reduced parahippocampal volume and psychosis symptoms in Alzheimer's disease. *Int. J. Geriatr. Psychiatr.* 33 (2), 389–395.
- Melzer, T., et al., 2003. Appearance models based on kernel canonical correlation analysis. *Pattern Recogn.* 36 (9), 1961–1971.
- Meng, X., et al., 2021. Multi-view Low-rank Preserving Embedding: a novel method for multi-view representation. *Eng. Appl. Artif. Intell.* 99, 104140.
- Mitchell, T.W., et al., 2002. Parahippocampal tau pathology in healthy aging, mild cognitive impairment, and early Alzheimer's disease. *Ann. Neurol.: Official Journal of the American Neurological Association and the Child Neurology Society* 51 (2), 182–189.
- Mizutani, K., et al., 2022. Nectin-2 in general and in the brain. *Mol. Cell. Biochem.* 477 (1), 167–180.
- Pasquini, L., et al., 2015. Link between hippocampus' raised local and eased global intrinsic connectivity in AD. *Alzheimer's Dementia* 11 (5), 475–484.
- Peng, X., et al., 2020. Deep subspace clustering. *IEEE Transact. Neural Networks Learn. Syst.* 31 (12), 5509–5521.
- Pennanen, C., et al., 2004. Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. *Neurobiol. Aging* 25 (3), 303–310.
- Peters, F., et al., 2009. The neural correlates of verbal short-term memory in Alzheimer's disease: an fMRI study. *Brain* 132 (7), 1833–1846.
- Poulin, S.P., et al., 2011. Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatr. Res. Neuroimaging* 194 (1), 7–13.
- Roses, A., et al., 2016. Understanding the genetics of APOE and TOMM40 and role of mitochondrial structure and function in clinical pharmacology of Alzheimer's disease. *Alzheimer's Dementia* 12 (6), 687–694.
- Roses, A.D., et al., 2010. A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *Pharmacogenomics J.* 10 (5), 375–384.
- Roshchupkin, G.V., et al., 2016. Heritability of the shape of subcortical brain structures in the general population. *Nat. Commun.* 7 (1), 13738.
- Scheff, S.W., et al., 2011. Synaptic loss in the inferior temporal gyrus in mild cognitive impairment and Alzheimer's disease. *J. Alzheim. Dis.* 24 (3), 547–557.
- Sharma, P., et al., 2019. Comprehensive review of mechanisms of pathogenesis involved in Alzheimer's disease and potential therapeutic strategies. *Prog. Neurobiol.* 174, 53–89.
- Shen, L., Thompson, P.M., 2019. Brain imaging genomics: integrated analysis and machine learning. *Proc. IEEE* 108 (1), 125–162.
- Shi, Y., Holtzman, D.M., 2018. Interplay between innate immunity and Alzheimer disease: APOE and TREM2 in the spotlight. *Nat. Rev. Immunol.* 18 (12), 759–772.
- Sporns, O., 2014. Contributions and challenges for network models in cognitive neuroscience. *Nat. Neurosci.* 17 (5), 652–660.
- van Hoesen, G.W., et al., 2000. The parahippocampal gyrus in Alzheimer's disease: clinical and preclinical neuroanatomical correlates. *Ann. N. Y. Acad. Sci.* 911 (1), 254–274.
- Vounou, M., et al., 2010. Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. *Neuroimage* 53 (3), 1147–1159.
- Wang, M., et al., 2022. Identify connectome between genotypes and brain network phenotypes via deep self-reconstruction sparse canonical correlation analysis. *Bioinformatics* 38 (8), 2323–2332.
- Wang, M., et al., 2019. Identify consistent cross-modality imaging genetic patterns via discriminant sparse canonical correlation analysis. *IEEE ACM Trans. Comput. Biol. Bioinf* 18 (4), 1549–1561.
- Wang, M., et al., 2021. Identify complex imaging genetic patterns via fusion self-expressive network analysis. *IEEE Trans. Med. Imag.* 40 (6), 1673–1686.
- Wang, S., et al., 2022. An improved fusion paired group lasso structured sparse canonical correlation analysis based on brain imaging genetics to identify biomarkers of alzheimer's disease. *Front. Aging Neurosci.* 13, 817520.
- Wang, Y., et al., 2016. Auto-encoder based dimensionality reduction. *Neurocomputing* 184, 232–242.
- Wang, Z., et al., 2011. Baseline and longitudinal patterns of hippocampal connectivity in mild cognitive impairment: evidence from resting state fMRI. *J. Neurol. Sci.* 309 (1), 79–85.
- Yang, X., et al., 2019. A survey on canonical correlation analysis. *IEEE Trans. Knowl. Data Eng.* 33 (6), 2349–2368.
- Yu, Z., et al., 2021. Predicting drug-disease associations through layer attention graph convolutional network. *Briefings Bioinf.* 22 (4), bbaa243.
- Zhang, Y., et al., 2020. Canonical correlation analysis of imaging genetics data based on statistical independence and structural sparsity. *IEEE journal of biomedical and health informatics* 24 (9), 2621–2629.
- Zhang, Y., et al., 2022. Multi-modal imaging genetics data fusion via a hypergraph-based manifold regularization: application to schizophrenia study. *IEEE Trans. Med. Imag.* 41 (9), 2263–2272.
- Zhou, T., et al., 2018. Brain-wide genome-wide association study for Alzheimer's disease via joint projection learning and sparse regression model. *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 66 (1), 165–175.