Medical Image Analysis 88 (2023) 102842

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Deep multimodality-disentangled association analysis network for imaging genetics in neurodegenerative diseases

Tao Wang [a], Xiumei Chen [a], Jiawei Zhang [a], Qianjin Feng [a,b,c,*], Meiyan Huang [a,b,c,*]

[a] School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[b] Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou 510515, China
[c] Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Southern Medical University, Guangzhou 510515, China

ARTICLE INFO

Keywords:
Neurodegenerative disease
Imaging genetics
Multimodality-disentangled association analysis
Biomarker detection

ABSTRACT

Imaging genetics is a crucial tool that is applied to explore potentially disease-related biomarkers, particularly for neurodegenerative diseases (NDs). With the development of imaging technology, the association analysis between multimodal imaging data and genetic data is gradually being concerned by a wide range of imaging genetics studies. However, multimodal data are fused first and then correlated with genetic data in traditional methods, which leads to an incomplete exploration of their common and complementary information. In addition, the inaccurate formulation in the complex relationships between imaging and genetic data and information loss caused by missing multimodal data are still open problems in imaging genetics studies. Therefore, in this study, a deep multimodality-disentangled association analysis network (DMAAN) is proposed to solve the aforementioned issues and detect the disease-related biomarkers of NDs simultaneously. First, the imaging data are nonlinearly projected into a latent space and imaging representations can be achieved. The imaging representations are further disentangled into common and specific parts by using a multimodal-disentangled module. Second, the genetic data are encoded to achieve genetic representations, and then, the achieved genetic representations are nonlinearly mapped to the common and specific imaging representations to build nonlinear associations between imaging and genetic data through an association analysis module. Moreover, modality mask vectors are synchronously synthesized to integrate the genetic and imaging data, which helps the following disease diagnosis. Finally, the proposed method achieves reasonable diagnosis performance via a disease diagnosis module and utilizes the label information to detect the disease-related modality-shared and modality-specific biomarkers. Furthermore, the genetic representation can be used to impute the missing multimodal data with our learning strategy. Two publicly available datasets with different NDs are used to demonstrate the effectiveness of the proposed DMAAN. The experimental results show that the proposed DMAAN can identify the disease-related biomarkers, which suggests the proposed DMAAN may provide new insights into the pathological mechanism and early diagnosis of NDs. The codes are publicly available at https://github.com/Meiyan88/DMAAN.

1. Introduction

Neurodegenerative diseases (NDs) are irreversible nervous system diseases that are characterized by progressive degeneration and death of nerve cells (Lei et al., 2020). The most common NDs are Alzheimer's disease (AD) and Parkinson's disease (PD), whose underlying pathogenesis remains unclear and effective cures are still lacking (Chen et al., 2022). In general, before NDs (such as AD and PD) are clinically diagnosed, pathologic processes leading to progressive degeneration begin Mattsson-Carlgren et al. (2020). Therefore, early detection of NDs may largely improve their treatments and slow down their progressions.

Imaging genetics can be applied to potentially explore disease-related biomarkers by analyzing the inherent relationships between neuroimaging and genetic data (e.g., single nucleotide polymorphism [SNP]) (Shen and Thompson, 2019). In addition, previous studies suggested that disease-related biomarkers can be applied to assist in the early detection of NDs. For example, structural magnetic resonance imaging (sMRI) can reveal structural changes in brains and contribute to the understanding of neural changes in NDs (Huang et al., 2021a, 2019). Meanwhile, diffusion tensor imaging (DTI) can investigate white matter changes, which is associated with cognitive, gait, and postural

* Corresponding authors at: School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China.
E-mail addresses: fengqj99@smu.edu.cn (Q. Feng), huangmeiyan16@163.com (M. Huang).

https://doi.org/10.1016/j.media.2023.102842
Received 2 November 2022; Received in revised form 1 March 2023; Accepted 15 May 2023
Available online 19 May 2023
1361-8415/© 2023 Elsevier B.V. All rights reserved.

impairments in PD (Nazmuddin et al., 2021). Moreover, positron emission tomography (PET) can measure Aβ burden, tau aggregation, and neuroinflammation, which can provide remarkable insight directly into the processes underlying the pathophysiology of AD (Chandra et al., 2019). Therefore, multimodal imaging data can measure different yet complementary information, and provide unprecedented opportunities for NDs' study. For the feature extraction of multimodal imaging data, previous studies have given different solutions, such as some deep learning methods for automatic feature extraction based on multimodal whole-brain images (Bonazzola et al., 2021; Wang et al., 2022) and some methods extract ROI-based features according to priori brain regions (Chen et al., 2022; Zhou et al., 2019). However, feature extraction based on whole-brain images is often accompanied by a large number of parameters and training time. Therefore, extensive imaging genetics studies are continuously conducted on ROI-based features for the balance of computational costs and extraction of informative features (Huang et al., 2021b; Ko et al., 2022; Lei et al., 2020). Apart from multimodal imaging biomarkers, NDs are reported to be related to genetic factors (Chen et al., 2022; Stephens et al., 2022), which can be measured accurately from birth. As such, detecting biomarkers of NDs from multimodal imaging and genetic data is a potential way for early prediction, diagnosis, and treatment of NDs (Du et al., 2021; Wen et al., 2020).

However, several challenges exist when performing imaging genetics between multimodal imaging and genetic data for biomarker detection of NDs. First, it is unavoidable to have missing data in multimodal imaging data because of imaging quality and high cost (Chen et al., 2022; Zhou et al., 2020). When discarding subjects with incomplete multimodal data and performing imaging genetics based on the remainder of the subjects with complete multimodal data, a lot of useful information may be lost, and the reliability of the imaging genetic model may be declined. Thus, it is crucial to deal with the missing data issue in multimodal imaging data to construct a reliable imaging genetic model. Second, common and complementary information exists in different modal data. Hence, how to explore crucially modality-shared and modality-specific biomarkers is another key challenge in multimodal imaging genetics. Third, the associations between imaging and genetic data are complex, which includes associations between multi-genetic and multi-imaging data, correlations among genetic data, and correlations among imaging data (Huang et al., 2021b). Such complex associations are difficult to be analyzed with a simple linear model. Finally, considering that genetic data cannot correctly encode pathological behaviors (Bi et al., 2017), such as disease status, traditional imaging genetic methods directly apply imaging data as endophenotypes to detect disease-related biomarkers. However, this step further requires confirming or refuting a suggestive link based on imaging endophenotypes because the detected biomarkers may or may not be related to the diseases (e.g., NDs in this study) (Xu et al., 2017). Therefore, a direct link between imaging genetics and NDs should be built to enhance the power of biomarker detection of NDs.

Thus, an end-to-end deep learning framework, named deep multimodality-disentangled association analysis network (DMAAN), is proposed to analyze the nonlinear associations between incomplete multimodal imaging and genetic data for exploiting disease-related modality-shared and modality-specific biomarkers to overcome the aforementioned challenges. The proposed DMAAN consists of three modules: a multimodality-disentangled module, an association analysis module, and a disease diagnosis module. First, in the multimodality-disentangled module, multimodal imaging data are encoded to achieve latent imaging representations of different modalities, which are further disentangled into common and specific representations. With the achieved common and specific representations of each modality, self- and cross-reconstruction of multimodal imaging data are performed to extract the useful information underly multimodal imaging data. Moreover, a distance ratio loss is applied to boost the power of exploiting modality-shared and modality-specific information. Second, the association analysis module is used to explore potential genetic representations and analyze the associations between imaging and genetic data. On the one hand, the genetic representations are mapped to the common and specific imaging representations for exploring crucially modality-shared and modality-specific information of genetic data. On the other hand, a modality mask vector for each modal data is generated to incorporate genetic information into imaging data for assisting the following disease diagnosis. Finally, the disease diagnosis module is applied to create a link between imaging genetics and disease status of NDs to guide the detection of potentially modality-shared and modality-specific imaging and genetic biomarkers related to NDs. Furthermore, a learning strategy is used to deal with the missing data problem in multimodal imaging data during the whole network training. Based on previous reports, no research has applied a deep learning model to perform imaging genetics to investigate the associations between incomplete multimodal imaging data and genetic data, as well as building a link between imaging genetics and NDs for the detection of modality-shared and modality-specific biomarkers of NDs.

In summary, our contribution is three-fold:

First, a novel deep-learning framework, named DMAAN, is designed for imaging genetics to build a nonlinear relationship between multimodal imaging and genetic data. In addition, a link is created between imaging genetics and NDs to explore potentially modality-shared and modality-specific biomarkers of NDs and improve model interpretability for the underlying pathological mechanisms of the diseases.

Second, a useful learning strategy is applied to deal with the missing data problem in multimodal imaging data where the whole incomplete multimodal imaging data can be fully used to improve the accuracy of biomarker detection.

Third, the proposed DMAAN is performed on Alzheimer's disease neuroimaging initiative (ADNI) and Parkinson's progression markers initiative (PPMI) datasets to evaluate its effectiveness in disease diagnosis and biomarker detection of NDs. Specifically, DMAAN is trained on the ADNI1 dataset, whereas it is tested on the ADNI2 dataset. Moreover, good disease diagnosis performance and potentially disease-related biomarkers are achieved using DMAAN, which further demonstrates the well-generalized ability of DMAAN.

2. Related work

2.1. Imaging genetics with nonlinear model

In general, traditional imaging genetic approaches construct linearly multivariable regression (Wen et al., 2020; Zhu et al., 2018; Zhou et al., 2018) or bimultivariate correlation models (Du et al., 2021, 2020; Kim et al., 2022) combined with several constraints to analyze associations between imaging and genetic data, as well as identify relevant biomarkers. However, the relationship between imaging and genetic data is so complex that a simple linear function cannot completely formulate it. Therefore, the nonlinear function is incorporated into the traditional imaging genetic methods to alleviate this problem. For example, Chen et al. (2022) applied a nonlinear kernel function to project genetic data into imaging data, which can analyze the nonlinear associations between imaging and genetic data to an extent. Nevertheless, only data from a single view (genetic or imaging data) is projected nonlinearly in this kind of method (Chen et al., 2022; Huang et al., 2021a), which may result in the declined performance of association analysis between imaging and genetic data. Kernel canonical correlation analysis (KCCA) projects data from a multiview to a common space and employs the kernel method to maximally capture nonlinear associations between data from a multiview. Although data from multiview are projected nonlinearly in KCCA, it is difficult for KCCA to select important biomarkers, and thus, KCCA is mainly used in feature fusion and classification (Yoshida et al., 2017).

Given the powerful learning ability, deep learning methods can be applied to imaging genetics to project data from multiview in a nonlinear way to accurately model the following relationships: among imaging data, among genetic data, and between imaging and genetic data (Huang et al., 2021b). Hence, Li et al. introduced a deep canonically correlated sparse autoencoder method to extract imaging and genetic features and applied a support vector machine (SVM) classifier to the extracted features for disease diagnosis (Li et al., 2020). Although imaging genetics is linked to disease (Li et al., 2020), the feature extraction and classification steps are separated, which may hamper the biological interpretation related to the disease (Li et al., 2020). Thus, some researchers proposed deep learning methods to analyze the associations between imaging and genetic data to address this problem, as well as perform automated disease diagnosis in a whole network (Huang et al., 2021b; Hu et al., 2021; Ko et al., 2022). The results of these methods can be well interpreted considering that the imaging and genetic data are directly related to the disease. However, these methods are focused on association analysis between single modal imaging and genetic data, and cannot directly handle the missing data issue in incomplete multimodal imaging data.

2.2. Association analysis between multimodal imaging and genetic data

When analyzing multimodal imaging data, data missing is a common issue because of imaging quality and high cost. In previous studies, three types of methods were applied to handle this issue: (a) removing the subjects with missing data (Du et al., 2020; Kim et al., 2022); (b) imputing the missing data (Hastie et al., 2015; Pan et al., 2021); and (c) using some training strategies without removing or imputing data (Chen et al., 2022; Zhou et al., 2019). For the first type of method, imaging genetics can be performed on complete multimodal imaging and genetic data, whose model complexity can be reduced. However, numerous crucial pieces of information may be discarded in this type of methods, in which the performance of biomarker detection and disease diagnosis may be declined (Zhou et al., 2020). Therefore, some studies initially imputed the missing data and then performed the downstream tasks (e.g., imaging genetics and disease diagnosis) to fully use the incomplete multimodal imaging data. However, this type of method may introduce noise into the model, thereby resulting in inaccurate biomarker detection and disease diagnosis (Chen et al., 2022). Different from the first two types of methods, Zhou et al. proposed a three-stage deep feature learning and fusion framework to take full advantage of incomplete multimodal imaging and genetic data without imputing data and learn feature representations of different modal data for accurate disease diagnosis (Zhou et al., 2019). Although disease-related multimodal biomarkers can be explored in Zhou et al. (2019), modality-shared and modality-specific information is lost, which may decline model interpretability for the underlying pathological mechanism of the disease because features extracted from different modal data are concatenated directly for final disease diagnosis.

Disentangled representation learning is designed to encode features into factors with separate meanings, which has shown great promise in the field of multimodal data analysis (Hu et al., 2020; Cheng et al., 2021). Hence, in this study, we propose a novel deep learning framework together with a disentangled representation learning-based module for exploiting the relations between incomplete multimodal imaging and genetic data, which can fully use all available imaging data without imputing data, discover potentially disease-related modality-shared and modality-specific biomarkers, and improve model interpretability for underlying pathological mechanism of the diseases.

3. Method

Fig. 1 presents an overview of the proposed DMAAN that consists of three modules to achieve potentially disease-associated biomarkers. First, a multimodality-disentangled module is constructed to learn latent imaging representations of different modalities nonlinearly, where a multimodality-disentangled strategy is applied to disentangle different representations, as well as learning the complementary information included in different representations. Second, an association analysis module is built to nonlinearly learn genetic representations and modality mask vectors, as well as analyze the associations between imaging and genetic data. With these two modules, the nonlinear associations between data from multiview are established in DMAAN. Furthermore, the learned genetic representations can be used to impute missing imaging data. Third, a disease diagnosis module is adopted to integrate all learned representations and modality mask vectors to achieve disease diagnosis. In this module, the introduction of disease status is useful to detect the potentially disease-related biomarkers, thereby improving model interpretability for the underlying pathological mechanism of the diseases. The details of DMAAN are provided in the following subsections. Moreover, the main notations used in this study are listed in Table 1.

3.1. Multimodality-disentangled module

An important step in achieving multimodality-disentanglement is to force the posterior distribution of data to be close to a pre-defined prior distribution, since the data in this prior distribution can be easy to be disentangled (Chen et al., 2018). Although both variational autoencoder and adversarial autoencoders (AAE) allow the distribution of latent representation obtained by encoder to match the pre-defined prior distribution, AAE may have superior capabilities in capturing data manifolds without the need to obtain the exact functional form of the prior distribution (Hu et al., 2020). Thus, the proposed multimodality-disentangled module consists of M AAEs, where M is modality number and equals 2 in this study (Fig. 1). Moreover, AAE used in this module was composed of an encoder, a decoder, and a shared discriminator. Each of the M modal imaging data $\{\boldsymbol{x}_i\}_{i=1,\ldots,M}$ is firstly fed to a specific encoder $\mathbf{E}_i^{\mathrm{Img}}$ to generate a latent imaging representation $\{\boldsymbol{v}_i\}_{i=1,\ldots,M}$. Then, the generated latent imaging representations for M modal imaging data are forced to approximate a prior distribution (i.e., Gaussian distribution) via a discriminator **Dis**. Furthermore, **Dis** is composed of a shared multi-layer perception (MLP) for different modal imaging data. In particular, **Dis** tries to distinguish if $\boldsymbol{v}_i$ followed the preassigned prior distribution. Let $p(\boldsymbol{v}_i)$, $q(\boldsymbol{v}_i \mid \boldsymbol{x}_i)$, $p(\boldsymbol{x}_i \mid \boldsymbol{v}_i)$, and $p_d(\boldsymbol{x}_i)$ be the prior, encoding, decoding, and imaging data distributions, respectively. Meanwhile, AAE is an autoencoder, which is regularized by matching the aggregated posterior distribution $q(\boldsymbol{v}_i) = \int_{x_i} q(\boldsymbol{v}_i \mid \boldsymbol{x}_i)\, p_d(\boldsymbol{x}_i)\, \mathrm{d}\boldsymbol{x}_i$ to the predefined prior distribution $p(\boldsymbol{v}_i)$. Moreover, the encoder $\mathbf{E}_i^{\mathrm{Img}}$ is the generator of the adversarial network in AAE, so the prior regularization imposed on $\boldsymbol{v}_i$ was an adversarial procedure to optimize the $\mathbf{E}_i^{\mathrm{Img}}$ and **Dis**, which is a minimax game. The encoder $\mathbf{E}_i^{\mathrm{Img}}$ ensures the aggregated posterior distribution can fool the discriminator by minimizing the adversarial loss, which can be defined as:

$$\mathcal{L}_{\mathrm{Adv}}^{i} = \min_{\mathbf{E}_i^{\mathrm{Img}}} \left[\mathbb{E}_{\boldsymbol{x}_i \sim p_d(\boldsymbol{x}_i)} \log \left(1 - \mathbf{Dis}\left(\mathbf{E}_i^{\mathrm{Img}}(\boldsymbol{x}_i) \right) \right) \right] \tag{1}$$

Whereas the discriminator **Dis** tries to distinguish whether $\boldsymbol{v}_i$ comes from distribution $p(\boldsymbol{v}_i)$ or $q(\boldsymbol{v}_i)$ by maximizing the discriminator loss, which can be represented as:

$$\begin{aligned} \mathcal{L}_{\mathrm{Dis}}^{i} = & \max_{\mathbf{Dis}} \left[\mathbb{E}_{\boldsymbol{x}_i \sim p_d(\boldsymbol{x}_i)} \log \left(1 - \mathbf{Dis}\left(\mathbf{E}_i^{\mathrm{Img}}(\boldsymbol{x}_i) \right) \right) \right] \\ & + \max_{\mathbf{Dis}} \left[\mathbb{E}_{\boldsymbol{v}_i \sim p(\boldsymbol{v}_i)} \log \left(\mathbf{Dis}(\boldsymbol{v}_i) \right) \right] \end{aligned} \tag{2}$$

After the adversarial learning, the latent imaging representation that fits the prior distribution is disentangled into a common representation $\{\boldsymbol{v}_i^{\mathrm{c}}\}_{i=1,\ldots,M}$ and a specific representation $\{\boldsymbol{v}_i^{\mathrm{s}}\}_{i=1,\ldots,M}$ by an extra fully connected layer $\mathbf{FC}_i^{\mathrm{Img}}$ for exploiting potential modality-shared and modality-specific biomarkers. Specifically, common representations should be as close as possible, whereas specific representations should be as far as possible. Therefore, a distance ratio loss

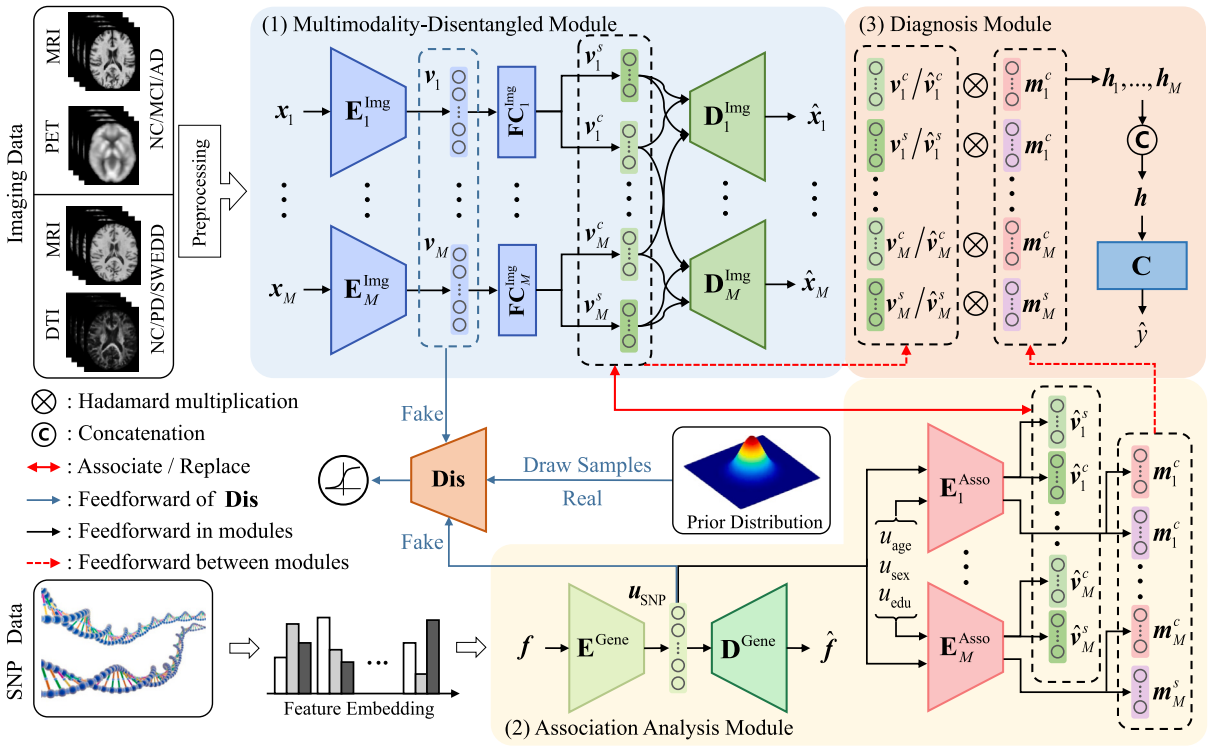

Fig. 1. An overview of the proposed DMAAN.

Table 1
Main notations used in this study.

Symbol	Description	Symbol	Description
Notations in multimodality-disentangled module		Notations in association analysis module	
M	Modality number	$\boldsymbol{f}_n$	Feature embeddings of SNP data for subject n
$\{\boldsymbol{x}_i\}_{i=1,\ldots,M}$	Imaging data	$\boldsymbol{m}_i$	Modality mask vector of ith modality
$\{\hat{\boldsymbol{x}}_i\}_{i=1,\ldots,M}$	Reconstructed imaging data	$\boldsymbol{m}_i^c/\boldsymbol{m}_i^s$	Common/Specific modality mask vector of ith modality
$\{\boldsymbol{v}_i\}_{i=1,\ldots,M}$	Latent imaging representation	$\boldsymbol{u}_{\text{Gene}}$	Latent genetic representations
$p(\boldsymbol{v}_i)$	Prior distribution of ith modality	$u_{\text{age}}/u_{\text{sex}}/u_{\text{edu}}$	Information of age/sex/education years
$q(\boldsymbol{v}_i \mid \boldsymbol{x}_i)$	Encoding distribution of ith modality	$p(\boldsymbol{v}_i)$	Prior distribution of genetic data
$p(\boldsymbol{x}_i \mid \boldsymbol{v}_i)$	Decoding distribution of ith modality	$p_d(\boldsymbol{u}_{\text{Gene}})$	Genetic data distribution
$p_d(\boldsymbol{x}_i)$	Imaging data distribution of ith modality	$\hat{\boldsymbol{v}}_i^c/\hat{\boldsymbol{v}}_i^s$	Mapped common/specific representation for imaging data of ith modality
$q(\boldsymbol{v}_i)$	Aggregated posterior distribution of ith modality	$\mathbf{E}^{\text{Gene}}/\mathbf{D}^{\text{Gene}}$	Encoder/Decoder for genetic data
$\{\boldsymbol{v}_i^c\}_{i=1,\ldots,M}/\{\boldsymbol{v}_i^s\}_{i=1,\ldots,M}$	Common/Specific representation of imaging data	$\mathbf{E}_i^{\text{Asso}}$	Encoder for mapping genetic data to imaging data of ith modality
Dis	Shared discriminator	Notations in disease diagnosis module	
$\mathbf{E}_i^{\text{Img}}/\mathbf{D}_i^{\text{Img}}$	Encoder/Decoder for imaging data of ith modality	$\boldsymbol{h}$	The input of the disease diagnosis module
$\mathbf{FC}_i^{\text{Img}}$	A fully connected layer for disentangling latent imaging representation of ith modality	y	Disease status (disease label)
		$\hat{y}$	The output of the disease diagnosis module
Notations in association analysis module		Notations in disease-related biomarker detection	
N	Subject number	$\bar{\boldsymbol{m}}_i^c/\bar{\boldsymbol{m}}_i^s$	Averaged common/specific modality mask vector of ith modality
N_G	Locus number of SNP data	$\bar{m}_i^{c(k)}/\bar{m}_i^{s(k)}$	kth dimensional value of averaged mask vector
$\boldsymbol{G}$	Raw SNP data	$z_i^{c(k)}, z_i^{s(k)}, z_{\text{Gene}}^{c(k)}, z_{\text{Gene}}^{c(k)}$	Contribution scores of ith modality and genetic data on dimension k
$\boldsymbol{f}$	Feature embeddings of SNP data for all subjects	$z_i^c, z_i^s, z_{\text{Gene}}^c, z_{\text{Gene}}^c$	Final contribution scores of ith modality and genetic data
$\hat{\boldsymbol{f}}$	Reconstructed feature embeddings of SNP data for all subjects		

between common and specific representations (Hu et al., 2020) can be formulated as:

$$\mathcal{L}_{\text{Dist}} = \left\|\boldsymbol{v}_1^c - \boldsymbol{v}_2^c\right\|_2 \Big/ \left\|\boldsymbol{v}_1^s - \boldsymbol{v}_2^s\right\|_2 \tag{3}$$

where $\|\cdot\|_2$ represents ℓ2-norm. Subsequently, the latent representations are used as inputs of their corresponding decoders $\mathbf{D}_i^{\text{Img}}$ for reconstruction of different modal imaging data $\hat{\boldsymbol{x}}_i$. In addition, given that share information among different modalities exists in common representations, $\boldsymbol{v}_i^c$ generated from the ith modality imaging data should provide useful information for the reconstruction of another modality. Thus, the reconstruction loss can be defined as:

$$\mathcal{L}_{\text{Recon}}^{\text{Img}} = \sum_{i=1}^{M}\sum_{j=1}^{M}\left\|\boldsymbol{x}_i - \mathbf{D}_i^{\text{Img}}\left(\boldsymbol{v}_j^c, \boldsymbol{v}_i^s\right)\right\|_2 \tag{4}$$

3.2. Association analysis module

The association analysis module contains an AAE and two association networks (Fig. 1). Herein, the AAE is applied to generate potential genetic representations which are also constrained within a prior distribution. Moreover, the two association networks are used to map the genetic representations to the modality-shared and modality-specific imaging representations for analyzing associations between imaging and genetic data. For a given SNP, the value of each locus is categorical (i.e., 0, 1, and 2 for normal, homogeneous mutant, and heterogeneous mutant, respectively), which may influence the performance of the deep neural network (Ko et al., 2022). Therefore, feature embeddings $\boldsymbol{f} = \{\boldsymbol{f}_n\}_{n=1}^N$ are used as inputs instead of using raw SNP data $\boldsymbol{G} \in \mathbb{R}^{N\times N_G}$ for the AAE (Huang et al., 2021b), where N and N_G are the

subject number and locus number, respectively. Specifically, the proportion of the three values (0/1/2) of all the subjects can be calculated to replace the three values. Therefore, each feature embedding $\boldsymbol{f}_n$ is a matrix with a size of $3\times N_G$. With the calculated feature embeddings, an AAE is built, which consists of an encoder $\mathbf{E}^{\text{Gene}}$, a decoder $\mathbf{D}^{\text{Gene}}$, and a discriminator **Dis**. Given that the Gaussian prior distribution is chosen for the potential genetic representations, the AAE of this module shares the same discriminator **Dis** as the multimodality-disentangled module. The genetic representations are learned from the feature embeddings by using the encoder, that is, $\boldsymbol{u}_{\text{Gene}} = \mathbf{E}^{\text{Gene}}(\boldsymbol{f})$. Furthermore, the feature embeddings can be reconstructed using the genetic representations and the decoder, that is, $\hat{\boldsymbol{f}} = \mathbf{D}^{\text{Gene}}(\boldsymbol{u}_{\text{Gene}})$. Similarly, the adversarial and discriminator losses of AAE in the association analysis module can be represented as:

$$\mathcal{L}_{\text{Adv}}^{\text{Gene}} = \min_{\mathbf{E}^{\text{Gene}}} \left[\mathbb{E}_{\boldsymbol{f}\sim p_d(\boldsymbol{f})} \log\left(1 - \mathbf{Dis}\left(\mathbf{E}^{\text{Gene}}(\boldsymbol{f})\right)\right)\right] \tag{5}$$

$$\begin{aligned}\mathcal{L}_{\text{Dis}}^{\text{Gene}} =& \max_{\mathbf{Dis}} \left[\mathbb{E}_{\boldsymbol{f}\sim p_d(\boldsymbol{f})} \log\left(1 - \mathbf{Dis}\left(\mathbf{E}^{\text{Gene}}(\boldsymbol{f})\right)\right)\right] \\ &+ \max_{\mathbf{Dis}} \left[\mathbb{E}_{\boldsymbol{u}_{\text{Gene}}\sim p(\boldsymbol{u}_{\text{Gene}})} \log\left(\mathbf{Dis}\left(\boldsymbol{u}_{\text{Gene}}\right)\right)\right]\end{aligned} \tag{6}$$

where $p_d(\boldsymbol{f})$ and $p(\boldsymbol{u}_{\text{Gene}})$ are SNP data and prior distributions, respectively. Meanwhile, the reconstruction loss is imposed between $\boldsymbol{f}$ and $\hat{\boldsymbol{f}}$, which can be formulated as:

$$\mathcal{L}_{\text{Recon}}^{\text{Gene}} = \left\|\boldsymbol{f} - \mathbf{D}^{\text{Gene}}\left(\boldsymbol{u}_{\text{Gene}}\right)\right\|_2 \tag{7}$$

This study aims to detect modality-shared and modality-specific imaging and genetic biomarkers. Hence, two association networks are constructed for different modalities to analyze the associations between imaging and genetic data to achieve this goal. For each association network, an encoder $\mathbf{E}_i^{\text{Asso}}$ is applied to map genetic representations $\boldsymbol{u}_{\text{Gene}}$ to the common and specific representations $\left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right]$, as well as generate a modality mask vector $\boldsymbol{m}_i = \left[\boldsymbol{m}_i^c, \boldsymbol{m}_i^s\right]$ that contains crucial genetic information for assisting the following disease diagnosis. In addition, NDs are diseases with progressive degeneration. The brain structure, brain function, and disease diagnosis of a subject change over time, whereas his/her genetic information is never changed. Therefore, the age of a subject is integrated into the corresponding genetic representation for analyzing the associations between genetic and longitudinal imaging data to fully use this characteristic of NDs. Moreover, some clinical covariates, such as sex and education years, are included in the genetic representation to exploit statistically express socioeconomic information (Ko et al., 2022). Therefore, the input of the association network is $\boldsymbol{u} = \left[\boldsymbol{u}_{\text{Gene}}, u_{\text{age}}, u_{\text{sex}}, u_{\text{edu}}\right]$. Furthermore, the ith association network can be represented as:

$$\left[\hat{\boldsymbol{v}}_i^c, \hat{\boldsymbol{v}}_i^s, \boldsymbol{m}_i^c, \boldsymbol{m}_i^s\right] = \mathbf{E}_i^{\text{Asso}}(\boldsymbol{u}) \tag{8}$$

Herein, the reconstruction loss can be formulated as:

$$\mathcal{L}_{\text{Recon}}^{\text{Asso}} = \sum_{i=1}^{M} \left(\lambda_{i1} \left\|\boldsymbol{v}_i^c - \hat{\boldsymbol{v}}_i^c\right\|_2 + \lambda_{i2} \left\|\boldsymbol{v}_i^s - \hat{\boldsymbol{v}}_i^s\right\|_2\right) \tag{9}$$

where λ_{i1} and λ_{i2} are hyperparameters that controls the tradeoff among different reconstruction errors.

3.3. Disease diagnosis module

Herein, a link between imaging genetics and disease status of NDs is created to guide the detection of potential modality-shared and modality-specific imaging and genetic biomarkers related to NDs (Huang et al., 2021a). Specifically, the input $\boldsymbol{h}$ of the disease diagnosis module can be formulated as:

$$\begin{aligned}\boldsymbol{h}_i &= \left\{\left[\boldsymbol{m}_i^c, \boldsymbol{m}_i^s\right] \otimes \left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right] + \left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right]\right\}_{i=1,\ldots,M} \\ \boldsymbol{h} &= \text{Concat}\left(\boldsymbol{h}_1, \ldots, \boldsymbol{h}_i, \ldots, \boldsymbol{h}_M\right)\end{aligned} \tag{10}$$

where $\text{Concat}(\cdot)$ and $\otimes$ represent feature concatenation and Hadamard multiplication operations, respectively. Then, a classification network $\mathbf{C}$ is built for the disease diagnosis module to link imaging genetics and disease status of NDs:

$$\hat{y} = \mathbf{C}(\boldsymbol{h}) \tag{11}$$

where $\hat{y}$ is the output of the disease diagnosis module. In this study, cross-entropy loss is used in the classification network:

$$\mathcal{L}_{\text{Cla}} = \text{CE}(y, \hat{y}) \tag{12}$$

where y is the disease status (disease labels) of the subjects.

3.4. Optimization and learning strategy

The proposed framework is trained in an end-to-end manner. The objective functions to optimize **Dis** can be written as:

$$\mathcal{L}_{\text{Dis}} = \sum_{i=1}^{M} \mathcal{L}_{\text{Dis}}^i + \mathcal{L}_{\text{Dis}}^{\text{Gene}} \tag{13}$$

Moreover, the total objective function for the optimization of $\mathbf{E}_i^{\text{Img}}$, $\mathbf{D}_i^{\text{Img}}$, $\mathbf{FC}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{D}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$, and $\mathbf{C}$ can be defined as:

$$\begin{aligned}\mathcal{L}_{\text{Total}} =& \gamma_1 \mathcal{L}_{\text{Dist}} + \gamma_2 \left(\mathcal{L}_{\text{Recon}}^{\text{Img}} + \mathcal{L}_{\text{Recon}}^{\text{Gene}}\right) + \mathcal{L}_{\text{Recon}}^{\text{Asso}} \\ &+ \gamma_3 \mathcal{L}_{\text{Cla}} + \gamma_4 \left(\sum_{i=1}^{M} \mathcal{L}_{\text{Adv}}^i + \mathcal{L}_{\text{Adv}}^{\text{Gene}}\right)\end{aligned} \tag{14}$$

where γ_1, γ_2, γ_3, and γ_4 are hyperparameters that control the ratio of different loss values. Please note that the hyperparameters λ_{i1} and λ_{i2} are contained in $\mathcal{L}_{\text{Recon}}^{\text{Asso}}$ as listed in Eq. (9). In summary, DMAAN first updates the discriminator **Dis** with Eq. (13), and then updates the parameters of other architecture in DMAAN with Eq. (14) when **Dis** is fixed.

Meanwhile, for multimodal imaging data, data missing problems exist because of imaging quality and high cost (Chen et al., 2022). Therefore, the following learning strategy is used to deal with the missing data problem by employing more samples to train a more reliable network. When the ith modal data is missed, the modality-shared and modality-specific imaging representations $\left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right]$ are imputed with the corresponding reconstruction values $\left[\hat{\boldsymbol{v}}_i^c, \hat{\boldsymbol{v}}_i^s\right]$ generated from the ith association network. With this strategy, multimodal imaging data with missing data can be well handled. The process of optimization and imputation is detailed in Algorithm 1.

Algorithm 1: Pseudo-code for DMAAN

Input: Training dataset $\boldsymbol{x}_i$, $\boldsymbol{f}$, u_{age}, u_{sex}, u_{edu}, and y;
network architectures **Dis**, $\mathbf{E}_i^{\text{Img}}$, $\mathbf{D}_i^{\text{Img}}$, $\mathbf{FC}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{D}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$, and $\mathbf{C}$

Initialize the parameters of **Dis**, $\mathbf{E}_i^{\text{Img}}$, $\mathbf{D}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{D}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$, and $\mathbf{C}$
for number of training iterations **do**
  $\boldsymbol{u} \leftarrow \mathbf{E}^{\text{Gene}}(\boldsymbol{f})$
  $\left[\hat{\boldsymbol{v}}_i^c, \hat{\boldsymbol{v}}_i^s, \boldsymbol{m}_i^c, \boldsymbol{m}_i^s\right] \leftarrow \mathbf{E}_i^{\text{Asso}}\left[\boldsymbol{u}_{\text{Gene}}, u_{\text{age}}, u_{\text{sex}}, u_{\text{edu}}\right]$
  if modality i is complete **then**
    $\boldsymbol{v}_i \leftarrow \mathbf{E}_i^{\text{Img}}(\boldsymbol{x}_i)$
    $\left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right] \leftarrow \mathbf{FC}_i^{\text{Img}}(\boldsymbol{v}_i)$
  else
    $\left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right] \leftarrow \left[\hat{\boldsymbol{v}}_i^c, \hat{\boldsymbol{v}}_i^s\right]$
  end if
  $\boldsymbol{h}_i \leftarrow \left\{\left[\boldsymbol{m}_i^c, \boldsymbol{m}_i^s\right] \otimes \left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right] + \left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right]\right\}_{i=1,\ldots,M}$
  $\boldsymbol{h} \leftarrow \text{Concat}\left(\boldsymbol{h}_1, \ldots, \boldsymbol{h}_i, \ldots, \boldsymbol{h}_M\right)$
  $\hat{y} \leftarrow \mathbf{C}(\boldsymbol{h})$
  (1) **Optimization of Dis**
  Train **Dis** using stochastic gradient descent with momentum (SGDM)
  (2) **Optimization of** $\mathbf{E}_i^{\text{Img}}$, $\mathbf{D}_i^{\text{Img}}$, $\mathbf{FC}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{D}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$, and $\mathbf{C}$
  Fix the parameters of **Dis** and update the parameters of $\mathbf{E}_i^{\text{Img}}$, $\mathbf{D}_i^{\text{Img}}$, $\mathbf{FC}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{D}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$, and $\mathbf{C}$ by SGDM.
end for

3.5. Disease-related biomarker detection

The genetic representations are used to generate the common and specific imaging representations in the association analysis module, and the modality mask vectors are synchronously synthesized. During the representation generated process, the imaging and genetic data are mapped into the same latent space, and the distance between them is reduced. Thus, the associations between imaging and genetic representations can be learned inherently. Simultaneously, the representations are connected to the disease labels, hence, these learned associations are highly disease-related. Therefore, a hypothesis is established that these associations are also included in the modality mask vectors. Particularly, the modality mask vectors act as attention vectors in the classification task, and the disease-related information is explicitly reflected by the weights in the vectors. However, these associations are built between genetic and imaging representations, rather than between the raw imaging and SNP data. Therefore, disease-related biomarkers cannot be obtained directly via the weights of the modality mask vectors. To address this problem, a gradient-based method is designed to search the disease-related biomarkers.

First, the absolute value of the weights in modality mask vectors are averaged among all subjects, which are defined as $\bar{\boldsymbol{m}}_i^c$ and $\bar{\boldsymbol{m}}_i^s$. Second, partial derivatives of each dimension of different representations $\left[\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right]$ or $\left[\hat{\boldsymbol{v}}_i^c, \hat{\boldsymbol{v}}_i^s\right]$ are calculated with regard to the raw input data $\boldsymbol{x}_i$ or $\boldsymbol{f}$, which can reflect the importance of different biomarkers regarding the representations (Ning et al., 2018). Specifically, the kth dimensional data $\left[\boldsymbol{v}_i^{c(k)}, \boldsymbol{v}_i^{s(k)}\right]$ and $\left[\hat{\boldsymbol{v}}_i^{c(k)}, \hat{\boldsymbol{v}}_i^{s(k)}\right]$ from different representations are multiplied with the corresponding kth dimensional mask values $\bar{m}_i^{c(k)}$ and $\bar{m}_i^{s(k)}$ from different mask vectors $\bar{\boldsymbol{m}}_i^c$ and $\bar{\boldsymbol{m}}_i^s$. And the multiplication results are taken as absolute values and then averaged among all subjects, which are defined as contribution scores $\begin{bmatrix} \boldsymbol{z}_i^{c(k)}, \boldsymbol{z}_i^{s(k)} \\ \boldsymbol{z}_i^{c(k)}, \boldsymbol{z}_i^{s(k)} \end{bmatrix}$ and $\left[\boldsymbol{z}_{\text{Gene}}^{c(k)}, \boldsymbol{z}_{\text{Gene}}^{s(k)}\right]$ of ith modality and genetic data. Thus, $\left[\boldsymbol{z}_i^{c(k)}, \boldsymbol{z}_i^{s(k)}\right]$ and $\left[\boldsymbol{z}_{\text{Gene}}^{c(k)}, \boldsymbol{z}_{\text{Gene}}^{s(k)}\right]$ can be formulated as:

$$\left[\boldsymbol{z}_i^{c(k)}, \boldsymbol{z}_i^{s(k)}\right] = \left[E\left(A\left(\frac{\partial \boldsymbol{v}_i^{c(k)} \bar{m}_i^{c(k)}}{\partial \boldsymbol{x}_i}\right)\right), E\left(A\left(\frac{\partial \boldsymbol{v}_i^{s(k)} \bar{m}_i^{c(k)}}{\partial \boldsymbol{x}_i}\right)\right)\right] \tag{15}$$

$$\left[\boldsymbol{z}_{\text{Gene}}^{c(k)}, \boldsymbol{z}_{\text{Gene}}^{s(k)}\right] = \left[E\left(A\left(\frac{\partial \hat{\boldsymbol{v}}_i^{c(k)} \bar{m}_i^{c(k)}}{\partial \boldsymbol{f}}\right)\right), E\left(A\left(\frac{\partial \hat{\boldsymbol{v}}_i^{s(k)} \bar{m}_i^{c(k)}}{\partial \boldsymbol{f}}\right)\right)\right] \tag{16}$$

where $E(\cdot)$ and $A(\cdot)$ represent the functions of average and absolute operations, respectively. Third, final contribution scores $\left[\boldsymbol{z}_i^c, \boldsymbol{z}_i^s\right]$ and $\left[\boldsymbol{z}_{\text{Gene}}^c, \boldsymbol{z}_{\text{Gene}}^s\right]$ are obtained by accumulating the results of Eqs. (15)–(16) for all dimensions, which can be represented as:

$$\left[\boldsymbol{z}_i^c, \boldsymbol{z}_i^s\right] = \sum_k^{K_1} \left[\boldsymbol{z}_i^{c(k)}, \boldsymbol{z}_i^{s(k)}\right] \tag{17}$$

$$\left[\boldsymbol{z}_{\text{Gene}}^c, \boldsymbol{z}_{\text{Gene}}^s\right] = \sum_k^{K_2} \left[\boldsymbol{z}_{\text{Gene}}^{c(k)}, \boldsymbol{z}_{\text{Gene}}^{s(k)}\right] \tag{18}$$

where K_1 and K_2 denote the dimensions of imaging representation and genetic representation, respectively. Finally, the top modality-shared and modality-specific imaging biomarkers and top genetic biomarkers can be obtained by the scores of different biomarkers in $\left[\boldsymbol{z}_i^c, \boldsymbol{z}_i^s\right]$ and $\left[\boldsymbol{z}_{\text{Gene}}^c, \boldsymbol{z}_{\text{Gene}}^s\right]$.

4. Experiments and discussions

4.1. Simulation study

4.1.1. Generation of simulated data

The simulated data were generated by following steps in a previous study (Huang et al., 2021b). First, two sparse vectors: $\boldsymbol{Q} \in \mathbb{R}^{a\times 1}$ and $\boldsymbol{T} \in \mathbb{R}^{b\times 1}$, where nonzero values in $\boldsymbol{Q}$ and $\boldsymbol{T}$ represent disease-related loci and ROIs in genetic data and imaging data, respectively; and a and b indicate the dimension number for SNPs and imaging data with values of 3,000 and 100, respectively. Herein, the numbers of nonzero values in $\boldsymbol{Q}$ and $\boldsymbol{T}$ were fixed to $n_a = 300$ and $n_b = 10$, respectively. Second, a latent vector of r^{th} class $\boldsymbol{Z} \in \mathbb{R}^{n_r\times 1}$ was generated randomly, where r can be set to 0 or 1, n_r was set to 1200 and denotes the number of subjects. Third, simulated imaging data $\hat{\boldsymbol{X}} = \left\{\hat{\boldsymbol{X}}_r\right\}_{r=0,1}$ were generated based on the latent vector $\boldsymbol{Z}_r$ where $\hat{\boldsymbol{X}}_r \sim \mathrm{N}\left(\boldsymbol{Z}_r\boldsymbol{T}, \mathbf{I}_{b\times b}\right)$ and $\mathbf{I}_{b\times b}$ is an identity matrix. Similarly, simulated genetic data $\hat{\boldsymbol{G}} = \left\{\hat{\boldsymbol{G}}_r\right\}_{r=0,1}$ were generated by $\hat{\boldsymbol{X}}_r$ with linear or nonlinear association to evaluate the classification and detection performance of proposed DMAAN. The linear and nonlinear associations can be formulated as:

(a) Linear association: $\hat{\boldsymbol{G}}_r \sim \mathrm{N}\left(\boldsymbol{Z}_r\boldsymbol{Q}, \mathbf{I}_{a\times a}\right)$

(b) Nonlinear association: $\hat{\boldsymbol{G}}_r \sim \mathrm{N}\left(L\left(\boldsymbol{Z}_r\right)\boldsymbol{Q}, \mathbf{I}_{a\times a}\right)$, where $L(\boldsymbol{Z}_r) = \boldsymbol{Z}_r^{1/5} + \boldsymbol{Z}_r^{1/4} + \boldsymbol{Z}_r^{1/2}$

Fourth, a binomial distribution $\mathrm{B}\left(2, \mathrm{logit}^{-1}\left(\boldsymbol{X} + \mathrm{logit}(\omega)\right)\right)$ was utilized to convert $\hat{\boldsymbol{G}}$ into discrete values 0, 1, and 2 (Kim et al., 2019), where $\mathrm{logit}(\omega) = \log(\omega/(1-\omega))$, and ω was drawn from a uniform distribution $\mathrm{U}(0.1, 0.7)$.

4.1.2. Results of simulated data

In the experiment of the simulated data, $\boldsymbol{Q}$ and $\boldsymbol{T}$ were set to multiple sets of different values. In particular, the values of $\boldsymbol{Q}$ and $\boldsymbol{T}$ were sampled from different uniform distributions, where uniform distributions were set to $\mathrm{U}(0.9, 1.0)$, $\mathrm{U}(1.0, 1.1)$, and $\mathrm{U}(1.1, 1.2)$. Moreover, ACC, AUC, and causal biomarker rate (CBR) were used to evaluate the classification and detection performance of DMAAN on simulated data, where CBR was defined as the ratio of the number of detected loci from genetic data (or ROIs from imaging data) in top n_a (or n_b) over the n_a (or n_b). A hold-out method was used on the simulated data and repeated five times. Moreover, all simulated data were divided into six non-overlapping subsets, that is, four as the training set, one as the validation set, and the remaining one as the testing set.

To consider the association between imaging and genetic data separately without demographic information (i.e., age, gender, and education years), DMAAN was modified, and the corresponding variant is denoted as DMAAN-oCov. Specifically, DMAAN-oCov projected SNP representations into the space where the image representations reside with the removal of demographic information, and genetic mask vectors were generated from the concatenation of common and specific representations of imaging data. The classification results of AUC and the detection results of CBR on different associations (linear or nonlinear) are listed in Table 2. The results in Table 2 show that DMAAN outperforms DMAAN-oCov in classification and biomarker detection, thereby demonstrating that the demographic information is helpful in classification and biomarker detection. In previous study, it was necessary to remove demographic information in advance to eliminate its impact on the input data when using demographic information as covariates (Huang et al., 2015; Pirinen et al., 2012). In this study, when demographic information was included as one of the inputs, DMAAN may be able to utilize it to remove the impact of covariates in the input data. Additionally, the greater the non-zero values in $\boldsymbol{Q}$ and $\boldsymbol{T}$, the greater the degree of association between disease-related genetic/imaging biomarkers and disease. According to the results in Table 2, high CBR is obtained by DMAAN when large non-zero values are set in $\boldsymbol{Q}$ and $\boldsymbol{T}$, implying that DMAAN can capture disease-related biomarkers when the association is strong enough. Furthermore, a similar performance was achieved by DMAAN on linear and nonlinear associations between genetic and imaging data (non-zero values > 1.0), which suggested that DMAAN is robust to linear and nonlinear data.

To demonstrate the ability of DMAAN to handle missing data on simulated dataset, different data missing rates were set for the training sets of the simulation data. Under the premise that non-zero values were sampled in $\mathrm{U}(1.0, 1.1)$, the data from one of the modalities in the multimodal imaging simulated data were randomly discarded at the data missing rates of 10%, 20%, and 30%. Table 3 lists the classification and detection performance of the simulated data with different data missing rates. The results from Table 3 suggest that a slight fluctuation is observed in all metrics when different amounts of data are missing, thereby proving the stability of DMAAN in handling missing data.

Table 2
Experimental results on simulated data, where "Asso", "Distribution", "Lin", and "Nolin" denote association types, sampling distribution of non-zero values, linear association, and nonlinear association, respectively. The results (%) are reported as mean ± std.

Asso	Distribution	Method	CBR(IMG)	CBR(Gene)	ACC	AUC
Lin	U(0.9,1.0)	DMAAN	100.00 ± 0.00	79.47 ± 2.06	99.90 ± 0.22	100.00 ± 0.00
	U(1.0,1.1)	DMAAN	100.00 ± 0.00	89.80 ± 1.12	100.00 ± 0.00	100.00 ± 0.00
	U(1.1,1.2)	DMAAN	100.00 ± 0.00	98.08 ± 0.32	100.00 ± 0.00	100.00 ± 0.00
	U(0.9,1.0)	DMAAN-oCov	100.00 ± 0.00	75.40 ± 2.10	98.60 ± 0.22	100.00 ± 0.00
	U(1.0,1.1)	DMAAN-oCov	100.00 ± 0.00	85.60 ± 2.13	100.00 ± 0.00	100.00 ± 0.00
	U(1.1,1.2)	DMAAN-oCov	100.00 ± 0.00	95.27 ± 2.50	100.00 ± 0.00	100.00 ± 0.00
Nolin	U(0.9,1.0)	DMAAN	100.00 ± 0.00	73.27 ± 1.85	98.70 ± 0.27	99.75 ± 0.29
	U(1.0,1.1)	DMAAN	100.00 ± 0.00	86.40 ± 0.72	99.50 ± 0.35	100.00 ± 0.00
	U(1.1,1.2)	DMAAN	100.00 ± 0.00	95.84 ± 0.96	100.00 ± 0.00	100.00 ± 0.00
	U(0.9,1.0)	DMAAN-oCov	96.00 ± 5.48	68.00 ± 1.78	95.10 ± 0.89	96.53 ± 0.39
	U(1.0,1.1)	DMAAN-oCov	98.00 ± 4.47	82.53 ± 2.01	97.50 ± 0.50	98.47 ± 0.00
	U(1.1,1.2)	DMAAN-oCov	100.00 ± 0.00	91.53 ± 2.07	100.00 ± 0.00	100.00 ± 0.00

Table 3
Experimental results of handling missing data on simulated dataset, where "Asso", "Rate", "Lin", and "Nolin" represent association types, the data missing rate, linear association, and nonlinear association, respectively. The results (%) are reported as mean ± standard deviation.

Asso	Rate	CBR(IMG)	CBR(Gene)	ACC	AUC
Lin	10%	100.00 ± 0.00	79.08 ± 1.42	99.80 ± 0.27	100.00 ± 0.00
	20%	100.00 ± 0.00	79.00 ± 1.18	99.70 ± 0.27	100.00 ± 0.00
	30%	100.00 ± 0.00	78.87 ± 1.26	99.50 ± 0.35	99.95 ± 0.10
Nolin	10%	100.00 ± 0.00	72.93 ± 1.55	98.70 ± 0.45	99.27 ± 0.32
	20%	100.00 ± 0.00	72.87 ± 1.59	98.60 ± 0.42	99.25 ± 0.28
	30%	100.00 ± 0.00	72.73 ± 1.89	98.60 ± 0.55	99.13 ± 0.27

4.2. Analysis on real data

4.2.1. Dataset and preprocessing

Data used in the preparation of this study were obtained from the ADNI (https://adni.loni.usc.edu/) and PPMI databases (https://www.ppmi-info.org). ADNI was launched in 2003 as a public–private partnership. The primary goal of ADNI has been to test whether or not serial MRI, PET, and other biological markers are useful in clinical trials of mild cognitive impairment and early AD. Moreover, PPMI is the first internationally recognized observational study created to identify and validate biomarkers for the prediction of PD progression.

For the ADNI database, 3166 subjects obtained from ADNI-1 and ADNI-2 were included in this study. Furthermore, T1-weighted MRI images, fluorodeoxyglucose PET (FDG-PET) images, and corresponding SNP data of all subjects were downloaded and preprocessed. Based on mini-mental state examination (MMSE) scores and clinical dementia rating, the subjects were classified into four categories, namely, normal control (NC), stable mild cognitive impairment (sMCI), progressive mild cognitive impairment (pMCI), and AD. Notably, those subjects whose disease status were reverted were excluded and the subjects that appeared in both ADNI-1 and ADNI-2 were removed from ADNI-2. For the PPMI database, 960 subjects with their corresponding T1-weighted MRI images, DTI images, and SNP data were collected. In addition, the 960 subjects were classified into three categories, namely, NC, PD, and scans without evidence for dopaminergic deficit (SWEDD). The details of the subjects obtained from different databases are listed in Table 4.

For MRI data, region of interest (ROI) based features were extracted under the following preprocessing steps: (a) anterior commissure-posterior commissure correction with MIPAV software (https://mipav.cit.nih.gov/); (b) image intensity inhomogeneity correction by applying N3 algorithm (Sled et al., 1998); (c) skull stripping via HD-BET (Isensee et al., 2019); (d) registering images to Montreal Neurological Institute (MNI) space via advanced normalization tools (ANTs) (Tustison et al., 2014); (e) three main tissues (i.e., gray matter (GM), white matter, and cerebrospinal fluid) segmentation using Atropos algorithm in ANTs; (f) applying the automated anatomical label (AAL) atlas of MNI space to project 90 brain ROIs; and (g) computing GM tissue volume of each ROI in the MNI space. For PET images, the co-registration strategy was utilized to align them to the corresponding MRI images via ANTs and the average intensity of each ROI was calculated as its feature. Finally, 90-dimensional ROI-based were separately extracted from the MRI and PET data for each subject.

For DTI data, 65 original format images where the b0 images do not activate the diffusion gradient were contained in each subject, whereas the other 64 images have different gradient directions. The DTI data were preprocessed by the following steps: (a) generating a b-vector file and a b-value file indicating each gradient direction and its scalar value via dcm2niix tool; (b) eddy correction with FMRIB Software Library (FSL) (Jenkinson et al., 2012); (c) skull stripping on b0 image using BET algorithm of FSL; (d) using generated files in step (a), the results of skull stripping and DTI images to calculate fractional anisotropy (FA) via difiti command in FSL; (e) aligning b0 image to MNI space via affine registration; (f) applying the transformation matrix in FA and labeling 90 ROIs with AAL; and (f) calculating the mean tissue density of each ROI of FA and then the corresponding 90-dimensional ROI-based features could be obtained.

For SNP data, the first line quality control steps include (a) call rate check per subject and SNP marker, (b) gender check, (c) sibling pair identification, (d) the Hardy-Weinberg equilibrium test, (e) marker removal by the minor allele frequency, and (f) population stratification. The second line preprocessing steps include the removal of SNPs with (a) more than 2% missing values, (b) minor allele frequencies of below 5%, and (c) Hardy-Weinberg equilibrium p-value $< 10^{--6}$. The Michigan Imputation Server (https://imputationserver.readthedocs.io/en/latest/) (Das et al., 2016) with Minimac4 was applied on all subjects to perform genotype imputation, where 1000G Phase I Integrated Release Version 3 haplotypes (http://www.1000genomes.org) (The 1000 Genomes Project Consortium et al., 2012) was used as reference panel. Additionally, a global sure independence screening procedure presented in our previous study (Huang et al., 2015) was applied to select the candidate SNPs. Herein, the selection of p-values in the SNP screening procedure is based on the amount of data in different datasets. In the ADNI database, the number of SNP data after PLINK (Purcell et al., 2007) is over 500,000, compared with just over 30,000 in the PPMI database. On the premise of preserving SNP information as much as possible and based on the consideration of data quantity, 0.01 and 0.1 were chosen as the p-value of SNP screening procedure for ADNI and PPMI database, respectively. Finally, 2960 and 3373 candidate SNPs for ADNI and PPMI datasets were preselected, respectively.

4.2.2. Experimental setup

The proposed DMAAN was implemented using Tensorflow on a workstation with one NVIDIA GeForce 2080Ti GPU. Moreover, four different classification tasks about ND diagnosis, namely, NC vs. AD, sMCI vs. pMCI, NC vs. PD, and PD vs. SWEDD, were concerned to verify the effectiveness of DMAAN. Among them, the disease-related information can be introduced by classification tasks of NC vs. AD and NC vs. PD

Table 4
Demographic information of different datasets. The age, education years (Edu), and MMSE scores are presented by mean ± standard deviation.

Dataset	Total	Category	Female/Male	Age	Edu	MMSE	PET
ADNI-1	2698	AD	135/218	76.74 ± 6.77	14.86 ± 3.08	22.34 ± 3.86	Not missing
		pMCI	75/144	75.79 ± 6.86	15.95 ± 2.78	26.60 ± 2.08	Not missing
		sMCI	96/240	76.94 ± 7.37	15.79 ± 2.92	27.43 ± 2.25	Not missing
		NC	122/246	76.39 ± 4.73	15.98 ± 2.98	29.13 ± 1.07	Not missing
		AD	208/231	75.58 ± 7.58	14.95 ± 3.14	21.72 ± 4.38	Missing
		pMCI	78/132	74.95 ± 6.67	15.99 ± 2.97	25.90 ± 2.31	Missing
		sMCI	115/178	75.85 ± 6.97	15.60 ± 3.18	27.33 ± 2.75	Missing
		NC	261/219	77.04 ± 5.19	16.30 ± 2.37	29.14 ± 1.09	Missing
ADNI-2	468	AD	26/49	75.33 ± 8.12	15.85 ± 2.55	23.05 ± 2.01	Not missing
		pMCI	33/35	72.60 ± 7.06	16.26 ± 2.64	27.35 ± 1.68	Not missing
		sMCI	68/93	71.31 ± 7.48	16.25 ± 2.69	28.24 ± 1.68	Not missing
		NC	78/86	73.46 ± 5.96	16.43 ± 2.54	29.01 ± 1.26	Not missing
PPMI	960	PD	141/298	62.32 ± 9.58	15.21 ± 2.88	–	Not missing
		SWEDD	25/43	61.94 ± 9.91	14.54 ± 4.92	–	Not missing
		NC	36/82	60.51 ± 11.39	16.01 ± 2.83	–	Not missing
		PD	79/135	62.06 ± 9.93	15.79 ± 2.72	–	Missing
		SWEDD	6/15	63.11 ± 10.75	15.71 ± 2.90	–	Missing
		NC	38/62	60.08 ± 12.22	16.48 ± 2.81	–	Missing

to guide biomarker detection. The classification tasks of sMCI vs. pMCI and PD vs. SWEDD are also interesting topics in clinical practice (Ko et al., 2022; Pan et al., 2021; Zhou et al., 2020). For the classification tasks of NC vs. AD and sMCI vs. pMCI, the data from ADNI-1 were selected as a training set, whereas the data from ADNI-2 were selected as independent testing set to demonstrate the robustness of DMAAN. Furthermore, the success of the deep learning model greatly relies on an adequate number of training data. Thus, the longitudinal data of ADNI-1 were applied as a training set to increase the data amount in the training stage (Zhou et al., 2019). Specifically, MRI and PET images at the baseline visit (BL) and 6, 12, 24, and 36 months were included when available. Due to the enormous computational costs associated with the dual internal and external loops required to use the cross-validation strategy (Lu et al., 2018; Zhou et al., 2019), and a hold-out strategy was selected for hyperparameter selection and data partitioning. The training set was divided based on the subjects' unique Roster IDs and disease status at BL because longitudinal data were used. All subjects within the training set were divided into five non-overlapping subsets with the same proportion of subjects' disease status at BL. One of the five subsets was successively selected as a validation set, and the four remaining subsets were combined for model training at each run. Thus, the results of the validation set can be obtained on each run, and the best hyperparameter combination can be chosen based on the highest averaged results of the validation set. Additionally, the model was retrained on the ADNI-1 dataset using the chosen hyperparameters and then tested on the ADNI-2 dataset. Similarly, for the classification tasks of NC vs. PD and PD vs. SWEDD, a hold-out method was also applied to the longitudinal data from PPMI, and all subjects with complete multimodal data were partitioned into 10 non-overlapping subsets with the same proportion of each class. Among these, eight subsets were applied for training, one was utilized for validation, and another one was used for testing. The data partitioning process was repeated five times. Given that the data missing problem cannot be solved in some compared methods used in this study, the subjects with incomplete multimodal data were included in the training set, whereas complete multimodal data were included in the validation and testing sets to ensure all methods make a fair comparison. The results of the validation and testing sets were achieved in each process. Specifically, the validation sets were utilized to select the best hyperparameters and models, and the final results of the PD classification were obtained from the average of five results in the testing set. Herein, for a fair and reasonable comparison, the same training settings were adopted for all experiments unless otherwise specified. Moreover, the area under the receiver operating characteristic curve (AUC), accuracy (ACC), sensitivity (SEN), specificity (SPE), and balanced class accuracy (BCA) were applied to evaluate the methods' performance. Furthermore, a pair t-test (at 95% significance level) was conducted on AUC for a statistical significance test among the compared methods and DMAAN.

To compensate for the computational cost of tuning the hyperparameters, the best hyperparameter combinations were selected from a small and predefined search range via a grid search (Zhou et al., 2019). Herein, the hyperparameters used in DMAAN were tuned carefully in three separate steps, and the hyperparameters of each step were fixed after being selected. (a) The architectures of different modules in DMAAN were mainly composed of MLPs, including the encoders, decoders, discriminators, and classification network. Thus, the hyperparameters associated with the architectures were selected in the first step. Considering that the node number of the input layer depended on the input data, and the structure of the decoder in different modules was reversed according to the structure of the encoder, the layer and node numbers of hidden and output layers in encoders and classification network attracted our major attention. The predefined search range in the first step are presented as follows: the layer number was selected from 2, 3, and 4; the node number of the hidden layers in $\mathbf{E}_i^{\text{Img}}$ was selected from 100, 200, 300, and 400; the node number of the hidden layers in $\mathbf{E}^{\text{Gene}}$ ranged from 200, 500, 1000, 2000, and 4000; the node number of the hidden layers in $\mathbf{E}_i^{\text{Asso}}$ was selected from 50, 100, 200, and 300; the node number of the output layers in $\mathbf{E}_i^{\text{Img}}$ and $\mathbf{E}^{\text{Gene}}$ was chosen from 30, 50, 100, and 200; the node number of the hidden layers in $\mathbf{C}$ ranged from 25, 50, 100, and 200; and the node number of the hidden layers in **Dis** was selected from 50, 100, 200, and 300; other hyperparameters with regard to the architectures, including the node number of input layers in encoders (i.e., $\mathbf{E}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$) and **Dis**, the node numbers of different layers in the decoders (i.e., $\mathbf{D}_i^{\text{Img}}$, $\mathbf{D}^{\text{Gene}}$), and the node number of $\mathbf{FC}_i^{\text{Img}}$, can be determined according to the aforementioned hyperparameters. The parameters of each module are determined independently because of the large number of MLPs. For the multimodality-disentangled module, an additional FC layer was added after the feature concatenation of $\boldsymbol{v}_i^c$ and $\boldsymbol{v}_i^s$ to classify the disease status according to the Eqs. (1)–(4) and extra cross-entropy loss, and the corresponding hyperparameters can be determined by AUC in the validation set. For the association analysis modules, an extra FC layer was added on $\boldsymbol{u}$, hyperparameters were chosen with AUC in validation set by utilizing Eqs. (5)–(7) and Eq. (9), and extra cross-entropy loss to train. After fixing the hyperparameters of the other modules, Eq. (12) and the AUC results were used to choose the hyperparameters of the disease diagnosis module. (b) After determining the architectures, the hyperparameters in the objective function were defined in the second step. To reduce the number of hyperparameters to be tuned, λ_{i1} and λ_{i2} were set to 1 and 1 by default in our study, respectively. Meanwhile, the predefined search range in the third step are detailed as follows: γ_1 ranged from 0.5 to 3.0 with the step of

Table 5
Hyperparameters used in DMAAN, and the hyperparameters not involved in the table can be determined by the hyperparameters in the table.

Hyperparameters	Network	Selected values
Layer number	$\mathbf{E}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$, $\mathbf{E}_i^{\text{Asso}}$, **Dis** / **C**	3/2
Node number of hidden layer	$\mathbf{E}_i^{\text{Img}}$ / $\mathbf{E}^{\text{Gene}}$ / **Dis** / $\mathbf{E}_i^{\text{Asso}}$	200/500/100/100
Node number of output layer	$\mathbf{E}_i^{\text{Img}}$, $\mathbf{E}^{\text{Gene}}$	50
λ_{i1}, λ_{i2}	–	1
γ_1 / γ_2 / γ_3 / γ_4	–	2.0/0.8/0.7/0.1
Initial learning rate	–	0.001
Batch size	–	32
Rate of dropout	–	0.5

0.5; γ_2, γ_3, and γ_4 were chosen from 0.1 to 1.0 in step of 0.1. (c) The hyperparameters of training settings were selected in the third step. In particular, an exponentially decreasing learning rate was adopted: the initial learning rate was selected from 10^{-1} to 10^{-5} with interval accumulation as multiples of 10^{-1}; batch size was selected from 16 to 128 with interval accumulation as multiples of 2; and the dropout rate ranged from 0.3 to 0.7 with step of 0.1.

In summary, two three-layer MLPs were adopted for two modalities as encoders of AAE ($\mathbf{E}_i^{\text{Img}}$) in the multimodality-disentangled module. All imaging data were 90-dimensional ROI-based features, and the number of nodes in the input layer was set to 90. Subsequently, the number of nodes in the hidden layers was set to 200, and the number of nodes in the output layer was set to 50. Subsequently, the output of the encoder was divided into two 25-dimensional features, namely, modality-shared and modality-specific features, with a FC layer ($\mathbf{FC}_i^{\text{Img}}$) of 50 nodes. In addition, the structures of the two corresponding decoders ($\mathbf{D}_i^{\text{Img}}$) were the same, each of which was composed of 50, 200, and 90 input, hidden, and output nodes, respectively. For the discriminator (**Dis**) of AAEs, a three-layer MLP, which had 50, 100, and 1 nodes at different layers, was applied. For the AAE of the genetic data in the association analysis module, the encoder ($\mathbf{E}^{\text{Gene}}$) and the decoder ($\mathbf{D}^{\text{Gene}}$) were both three-layer MLPs. The numbers of nodes of the encoder were set to 2960/3376 (AD and PD genetic data were 2960- and 3373-dimensional features, respectively), 500, and 50, respectively, whereas the setting of the decoder was reversed. For association networks ($\mathbf{E}_i^{\text{Asso}}$) in the association analysis module, the network consisted of two three-layer MLPs, each of which was composed of 54 input nodes (50-dimensional data were provided from the genetic representations and other 4-dimensional data were about information of age, one hot coding of sex, and education year), 100 hidden nodes, and 100 output nodes. Furthermore, a two-layer MLP was included in the disease diagnosis module (**C**), which had an input layer of 75 nodes and an output layer of 1 node. All hidden layers in the aforementioned MLPs were activated by the exponential linear unit. To alleviate overfitting, the dropout strategy (Krizhevsky et al., 2017) was utilized for all modules, and the rate of the dropout was set to 0.5. After determining the network structure, the best hyperparameters of the loss function were searched. Moreover, γ_1, γ_2, γ_3, and γ_4 were set to 2.0, 0.8, 0.7, and 0.1, respectively. The selected hyperparameters and corresponding search range are listed in Table 5.

At the training stage, all trainable parameters were initialized by Xavier initializer, and SGDM was used for optimizing, whose initial learning rate and batch size were set to 0.001 and 32, respectively. $\ell 1$- and $\ell 2$-regularizations were applied to avoid overfitting. Moreover, an exponentially decreasing learning rate was adopted, whose initial value and decreasing ratio per 1000 steps were set to 0.004 and 4%, respectively.

4.2.3. Ablation study

Four variants of DMAAN for ablation study were conducted to evaluate the effectiveness of each module, which were denoted as DMAAN-oMD, DMAAN-oAAE, DMAAN-oAsso, DMAAN-oImp, respectively.

For DMAAN-oMD, the latent imaging representations generated by encoders were not disentangled and were directly sent to the decoders to investigate the effectiveness of the multimodality-disentangled strategy. Correspondingly, the genetic representations were used to generate the modality mask vectors and fit the latent imaging representations directly in the association analysis module. Meanwhile, for DMAAN-oAAE, when discarding AAE, adversarial learning used in the multimodality-disentangled module was removed from the autoencoder. Herein, AAE enables the latent imaging representation to conform to the prior distribution. Thus, constraints from the discriminator were lost in the latent imaging representations without AAE. For DMAAN-oAsso, the validity of the associations between imaging and genetic data was considered. Through the genetic representations to fit common and specific imaging representations, modality-shared and modality-specific associations between imaging and genetic data inherently can be learned in the association analysis module. Thus, when removing these associations, the association analysis module was only used to generate modality mask vectors. For DMAAN-oImp, when the learning strategy for incomplete multimodal imaging data was removed, only the paired imaging data was used for training, which indicates the subjects with missing PET or DTI data were discarded. For DMAAN-mRec, the effectiveness of the reconstruction loss listed in Eq. (4) was explored. In Eq. (4), the reconstruction loss assumed that the common representations from both modalities contributed equally to the reconstruction task. To investigate whether such an assumption was reasonable, a parameter η was added in the reconstruction loss of DMAAN-mRec to control the contributions of common representations from other modalities. For the reconstruction of modality i, the common representation $\boldsymbol{v}_j^c$ in Eq. (4) was replaced by the combination of $(1-\eta)\,\boldsymbol{v}_j^c$ and $\eta\boldsymbol{v}_i^c$ (i.e., $\mathcal{L}_{\text{Recon}}^{\text{Img}} = \sum_{i=1}^{M}\sum_{j=1}^{M}\left\|\boldsymbol{x}_i - \mathbf{D}_i^{\text{Img}}\left((1-\eta)\,\boldsymbol{v}_j^c + \eta\boldsymbol{v}_i^c, \boldsymbol{v}_i^s\right)\right\|$). Herein, η was set as $0.97^{n_{\text{epoch}}}$, where n_{epoch} represents the number of epochs during training. As training progresses, the contributions from other common representations gradually increased.

The results of the ablation study are summarized in Table 6. As shown in Table 6, the best performance is achieved by using our complete DMAAN on most metrics in four classification tasks. Moreover, DMAAN achieves significant differences in AUC (p-value < 0.05) compared with DMAAN-oMD, DMAAN-oAAE, DMAAN-oAsso, and DMAAN-oImp. Compared with DMAAN-oMD, the classification performance is improved when incorporating the multimodality-disentangled strategy, which suggests that the redundancy among multimodal features may be reduced by using the multimodality-disentangled strategy (Yang et al., 2022). In addition, this strategy opens the possibility of detecting potential modality-shared and modality-specific imaging and genetic biomarkers. The performance degradation observed in DMAAN-oAAE demonstrates the importance of matching the aggregated posterior to the prior. When the latent representations were constrained in the prior distribution, the representations may be easily disentangled. The poor performance achieved by DMAAN-oAsso implies that the associations between imaging and genetic data facilitate classification. For DMAAN-oImp, the removal of the learning strategy for incomplete multimodal imaging data leads to a reduction in the number of training data and

Table 6
Ablation experiments of DMAAN. The results (%) are reported as mean ± standard deviation, where * indicates that DMAAN achieves a significant difference on AUC with a p-value < 0.05 compared with other methods.

Method	NC vs. AD					sMCI vs. pMCI				
	ACC	AUC	SPE	SEN	BCA	ACC	AUC	SPE	SEN	BCA
DMAAN-oMD *	90.04 ± 1.08	93.81 ± 0.77	92.07 ± 3.26	85.60 ± 6.42	88.84 ± 1.95	74.15 ± 2.47	75.18 ± 2.09	75.16 ± 6.96	71.76 ± 15.12	73.46 ± 5.40
DMAAN-oAAE *	87.28 ± 0.87	92.88 ± 1.19	87.93 ± 1.09	85.87 ± 2.02	86.90 ± 1.04	72.49 ± 3.13	72.81 ± 1.49	73.54 ± 3.82	70.00 ± 6.79	71.77 ± 3.66
DMAAN-oAsso *	89.54 ± 0.89	93.28 ± 1.36	91.83 ± 1.26	84.53 ± 3.72	88.18 ± 1.53	74.32 ± 2.72	74.19 ± 1.92	75.90 ± 2.42	70.59 ± 3.45	73.24 ± 2.92
DMAAN-oImp *	91.21 ± 1.70	94.74 ± 0.65	93.29 ± 2.76	86.67 ± 2.98	89.98 ± 1.52	74.67 ± 1.70	74.80 ± 0.65	76.52 ± 3.66	70.29 ± 3.35	73.41 ± 0.60
DMAAN-mRec	92.64 ± 1.41	96.40 ± 0.24	93.41 ± 1.09	90.93 ± 4.15	92.17 ± 2.07	76.19 ± 2.11	77.39 ± 1.42	**77.02 ± 2.48**	74.12 ± 8.74	75.57 ± 3.77
DMAAN	**93.81 ± 0.62**	**96.98 ± 0.52**	**95.00 ± 0.80**	**91.20 ± 1.52**	**93.10 ± 0.76**	**76.68 ± 1.05**	**78.08 ± 1.56**	76.40 ± 2.28	**77.35 ± 4.48**	**76.88 ± 1.56**
Method	NC vs. PD					PD vs. SWEDD				
	ACC	AUC	SPE	SEN	BCA	ACC	AUC	SPE	SEN	BCA
DMAAN-oMD *	73.93 ± 3.09	77.94 ± 4.23	74.86 ± 6.91	75.26 ± 5.60	75.06 ± 3.09	73.15 ± 4.12	71.85 ± 4.36	78.16 ± 6.37	72.52 ± 5.17	75.34 ± 3.04
DMAAN-oAAE *	76.43 ± 4.99	76.40 ± 4.39	74.97 ± 11.81	77.18 ± 8.68	76.07 ± 2.79	74.13 ± 4.67	73.63 ± 5.58	80.92 ± 6.08	73.17 ± 5.08	77.04 ± 4.47
DMAAN-oAsso *	74.81 ± 3.82	76.42 ± 1.76	74.72 ± 11.21	75.95 ± 7.75	75.34 ± 2.59	77.54 ± 4.50	78.58 ± 7.26	82.56 ± 7.68	76.80 ± 5.09	79.68 ± 4.49
DMAAN-oImp *	77.09 ± 4.55	78.51 ± 5.49	73.05 ± 14.96	79.71 ± 4.30	76.61 ± 2.82	79.20 ± 6.61	80.17 ± 5.63	81.22 ± 8.74	78.95 ± 6.97	80.09 ± 4.69
DMAAN-mRec	**80.18 ± 1.24**	80.79 ± 6.97	76.43 ± 9.06	80.21 ± 5.27	78.32 ± 6.69	79.63 ± 6.87	80.93 ± 5.65	82.68 ± 5.60	**79.20 ± 7.55**	80.94 ± 5.61
DMAAN	79.96 ± 5.03	**81.56 ± 6.83**	**77.27 ± 4.59**	**80.62 ± 6.10**	**78.95 ± 3.83**	**80.31 ± 7.03**	**82.51 ± 6.10**	**88.51 ± 5.56**	79.04 ± 8.39	**83.78 ± 4.16**

Table 7
Classification results when different types of data were used as inputs, where "Whole-brain" means the whole-brain images were used as inputs, and "fastGWAS-MLM" represents the SNPs selected by using fastGWAS-MLM were used as inputs. The results (%) are reported as mean ± standard deviation.

Task	Input	ACC	AUC	SPE	SEN	BCA
NC vs.AD	Whole-brain	93.05 ± 0.87	97.14 ± 1.97	94.27 ± 1.53	90.40 ± 3.82	92.33 ± 1.47
	fastGWAS-MLM	92.80 ± 1.30	96.13 ± 0.60	93.29 ± 1.78	91.73 ± 1.74	92.51 ± 1.20
	Ours	93.81 ± 0.62	96.98 ± 0.52	93.40 ± 1.76	92.16 ± 1.06	93.10 ± 0.76
sMCI vs.pMCI	Whole-brain	76.86 ± 1.27	77.10 ± 1.03	78.39 ± 3.91	73.24 ± 8.66	75.81 ± 2.75
	fastGWAS-MLM	75.90 ± 2.34	76.78 ± 0.77	76.27 ± 8.44	75.00 ± 13.56	75.64 ± 2.96
	Ours	76.68 ± 1.05	78.08 ± 1.56	76.40 ± 2.28	77.35 ± 4.48	76.88 ± 1.56
NC vs.PD	Whole-brain	80.14 ± 5.64	80.88 ± 5.48	75.67 ± 3.90	81.36 ± 6.81	78.52 ± 4.22
	fastGWAS-MLM	79.08 ± 6.45	78.88 ± 5.86	75.82 ± 7.97	79.99 ± 8.07	77.91 ± 5.37
	Ours	79.96 ± 5.03	81.56 ± 6.83	77.27 ± 4.59	80.62 ± 6.10	78.95 ± 3.83
PD vs.SWEDD	Whole-brain	79.53 ± 6.86	81.83 ± 6.36	83.81 ± 7.98	78.88 ± 7.00	81.34 ± 6.98
	fastGWAS-MLM	80.88 ± 5.68	79.80 ± 7.51	86.78 ± 5.92	80.01 ± 6.38	83.40 ± 4.70
	Ours	80.31 ± 7.03	82.51 ± 6.10	88.51 ± 5.56	79.04 ± 8.39	83.78 ± 4.16

a decline in the classification performance. Although higher SPE are observed by using DMAAN-oImp in the tasks of sMCI vs. pMCI, DMAAN is still considered superior in terms of classification performance with higher ACC (76.68% vs. 74.67%), AUC (78.08% vs. 74.80%), SEN (77.35% vs. 70.29%), and BCA (76.88% vs. 73.41%) in sMCI vs. pMCI classification. Moreover, DMAAN and DMAN-mRec perform similarly (no statistical difference on AUC), thereby illustrating the effectiveness of the reconstruction loss listed in Eq. (4). Moreover, previous studies proved that the reconstruction loss was effective to achieve self- and cross-reconstruction (Cheng et al., 2021; Hu et al., 2020). Although differences may be observed between the common representations from other modalities and the common representations from its own modality in the early stage of training process, this difference may not lead to the failure of the reconstruction task, which may benefit from the existence of the distance ratio loss listed in Eq. (3). In terms of SEN and SPE, a serious decline in SEN is observed in DMAAN in PD vs. SWEDD task, which may be caused by the severe imbalance in the number of subjects in various categories.

4.2.4. Effects of different inputs

Instead of using the extracted ROI data as the inputs of DMANN, the whole-brain images can also be used as the inputs to explore the different effects of using ROI data and whole-brain images on the proposed DMAAN. Specifically, the registered MRI and PET images in step (d) of MRI preprocessing pipeline from Section 4.2.1 were used as inputs. A two-step training strategy was used to train imaging feature extractor and DMAAN because the amount of network parameters increased greatly when using the whole-brain images as inputs. For the first step, a 3D-ResNet18 (Ruiz et al., 2020) was used to extract imaging features, and disease labels y were used to guide the feature extraction. Notably, two FC layers were connected after the convolution blocks. One layer contained 90 nodes as the feature extraction layer, and the other layer contains 2 nodes as the classification layer. Additionally, Adam optimizer with a learning rate of 0.01 was used for optimization. With the well-trained 3D-ResNet18, the outputs of the feature extraction layer can be used as imaging features. For the second step, the imaging features were used as the inputs of DMAAN. In this experiment, hyperparameters and data partitioning pattern remain the same as those in the experimental setup mentioned in Section 4.2.2. Moreover, Table 7 shows the classification results when using whole-brain images as inputs.

The comparison results of using different imaging inputs in all classification tasks show that the classification performance of using whole-brain images in the testing sets is similar to that of using ROI data as inputs (no statistical difference on AUC). However, the training cost greatly increased by using the whole-brain images as inputs, which is reflected in the training of the image feature extractor. Furthermore, the proposed biomarker detection method described in Section 3.5 can directly establish the relationship between imaging data and disease labels by using ROI data, whereas a more complex two-step operation is required when using whole-brain images. Therefore, the ROI data were still the main sources of imaging data for DMAAN, and subsequent biomarker detection was conducted on ROI data.

Different SNP screening procedures may affect the number and loci of preselected SNPs. Therefore, a mixed linear model-based genome-wide association (fastGWAS-MLM) method, which is a highly resource-efficient tool for mixed linear model-based genome-wide association analysis of large-scale data to select candidate SNPs (Jiang et al., 2019), is applied to compare with our global sure independence screening procedure to evaluate the impact of different preselected SNPs on DMAAN

Table 8
Comparison of different methods on two different classification tasks about ND diagnosis. The results (%) are reported as mean ± standard deviation, where * indicates that DMAAN achieves a significant difference on AUC with a p-value < 0.05 compared with other methods.

Method	NC vs. AD					sMCI vs. pMCI				
	ACC	AUC	SPE	SEN	BCA	ACC	AUC	SPE	SEN	BCA
ScCNAA *	80.61 ± 3.16	85.85 ± 2.14	83.27 ± 6.19	74.79 ± 8.90	79.03 ± 2.72	68.12 ± 2.16	69.15 ± 2.47	75.28 ± 2.82	56.47 ± 4.42	68.64 ± 2.27
DCCSAE *	85.64 ± 1.29	89.30 ± 1.48	85.37 ± 4.46	86.13 ± 6.44	85.75 ± 1.35	70.39 ± 4.24	69.55 ± 1.48	73.91 ± 7.62	62.06 ± 14.72	67.99 ± 5.74
SWDNN *	88.62 ± 3.66	88.81 ± 2.80	88.29 ± 6.97	89.33 ± 6.99	88.81 ± 2.79	70.92 ± 3.35	72.09 ± 1.07	70.81 ± 4.17	71.18 ± 3.83	70.99 ± 3.10
DGDL *	90.10 ± 1.28	92.31 ± 1.22	92.93 ± 2.05	83.73 ± 3.82	88.33 ± 1.49	73.89 ± 2.38	73.13 ± 1.44	75.40 ± 2.58	70.29 ± 6.44	72.85 ± 3.24
DMAAN-oImp *	91.21 ± 1.70	94.74 ± 0.65	93.29 ± 2.76	86.67 ± 2.98	89.98 ± 1.52	74.67 ± 1.70	74.80 ± 0.65	**76.52 ± 3.66**	70.29 ± 3.35	73.41 ± 0.60
DMAAN	**93.81 ± 0.62**	**96.98 ± 0.52**	**95.00 ± 0.80**	**91.20 ± 1.52**	**93.10 ± 0.76**	**76.68 ± 1.05**	**78.08 ± 1.56**	76.40 ± 2.28	**77.35 ± 4.48**	**76.88 ± 1.56**
Method	NC vs. PD					PD vs. SWEDD				
	ACC	AUC	SPE	SEN	BCA	ACC	AUC	SPE	SEN	BCA
ScCNAA *	75.59 ± 6.34	69.94 ± 7.11	61.34 ± 13.85	79.66 ± 6.334	70.50 ± 8.05	72.21 ± 8.71	74.93 ± 8.82	73.85 ± 9.58	60.78 ± 10.88	67.32 ± 7.50
DCCSAE *	72.11 ± 5.75	72.21 ± 5.61	71.53 ± 12.34	73.63 ± 5.17	72.58 ± 8.04	70.02 ± 5.42	71.85 ± 4.36	73.19 ± 11.45	69.41 ± 5.76	71.30 ± 6.69
SWDNN *	76.01 ± 6.13	75.22 ± 5.77	72.49 ± 7.80	76.82 ± 6.38	74.65 ± 6.05	73.85 ± 5.35	73.63 ± 5.58	74.96 ± 3.68	73.35 ± 5.97	74.15 ± 3.82
DGDL *	76.54 ± 6.83	77.11 ± 5.22	71.15 ± 11.84	77.00 ± 5.44	74.08 ± 7.94	77.43 ± 4.38	78.58 ± 7.26	80.81 ± 10.89	77.02 ± 5.13	78.92 ± 5.55
DMAAN-oImp *	77.09 ± 4.55	78.51 ± 5.49	73.05 ± 14.96	79.71 ± 4.30	76.61 ± 2.82	79.20 ± 6.61	80.17 ± 5.63	81.22 ± 8.74	78.95 ± 6.97	80.09 ± 4.69
DMAAN	**79.96 ± 5.03**	**81.56 ± 6.83**	**77.27 ± 4.59**	**80.62 ± 6.10**	**78.95 ± 3.83**	**80.31 ± 7.03**	**82.51 ± 6.10**	**88.51 ± 5.56**	**79.04 ± 8.39**	**83.78 ± 4.16**

Table 9
Experimental results of handling missing data on real dataset, where "Rate" represents the data missing rate. The results (%) are reported as mean ± standard deviation, where * indicates that DMAAN achieves a significant difference on AUC with a p-value < 0.05 compared with other methods.

Method	Rate	ACC	AUC	SPE	SEN	BCA
ScCNAA *	10%	78.91 ± 8.20	86.74 ± 3.57	84.02 ± 15.59	67.73 ± 9.85	75.88 ± 4.38
ScCNAA *	20%	78.91 ± 9.56	86.00 ± 4.23	86.58 ± 19.72	62.13 ± 24.46	74.36 ± 8.45
ScCNAA *	30%	78.07 ± 9.04	85.73 ± 2.81	83.05 ± 19.55	67.20 ± 21.41	75.12 ± 6.61
SWDNN *	10%	87.03 ± 2.88	88.31 ± 3.12	86.71 ± 2.84	87.73 ± 4.56	87.22 ± 3.18
SWDNN *	20%	86.69 ± 3.48	87.83 ± 3.09	86.83 ± 3.48	86.40 ± 4.36	86.61 ± 3.11
SWDNN *	30%	86.19 ± 3.23	87.58 ± 2.87	86.59 ± 4.78	85.33 ± 2.49	85.96 ± 2.52
DMAAN	10%	90.45 ± 1.18	93.96 ± 0.39	90.36 ± 1.32	89.07 ± 1.12	90.00 ± 0.67
DMAAN	20%	89.21 ± 1.30	93.40 ± 0.62	90.00 ± 2.09	87.47 ± 1.19	88.77 ± 0.87
DMAAN	30%	88.62 ± 1.52	92.92 ± 0.85	89.27 ± 2.74	87.20 ± 2.76	88.23 ± 1.22

performance. In these two SNP screening procedures, similar numbers of SNPs were selected (2722 vs. 2960 in ADNI and 3159 vs. 3373 in PPMI datasets for fastGWAS-MLM and global sure independence screening procedure, respectively). Accordingly, the node number of the input layer in $\mathbf{E}^{\text{Gene}}$ was set to 2722 and 3159 when utilizing the SNP data from the new SNP screening procedure. The comparison results of the classification tasks on NC vs. AD, NC vs. PD, sMCI vs. pMCI, and PD vs. SWEDD are listed in Table 7.

According to the results listed in Table 7, our proposed DMAAN achieves similar performance with SNP data selected from two different SNP screening procedures (no statistical difference on AUC). The similar classification results imply that DMAAN is robust to SNP data obtained by different screening procedures. Our future research direction will further explore this aspect, such as developing a data-driven SNP screening procedure, or directly using unselected SNP data as inputs.

4.2.5. Comparison with other methods

In this section, four state-of-the-art methods were applied to further assess the performance of DMAAN. Moreover, the configurations of these compared methods are provided as follows:

- Structure-constrained combination-based nonlinear association analysis (ScCNAA) (Chen et al., 2022): a multivariate correlation method proposed for biomarker detection, which can handle the incomplete multimodal imaging data, and only SNP data from a single view was projected nonlinearly to the imaging data. Then, the selected biomarkers that were highly associated with the disease were fed to an SVM classifier to accomplish classification tasks.
- Deep canonically correlated sparse autoencoder (DCCSAE) (Li et al., 2020): a neural network designed for imaging genetics, where sparse autoencoder was used for projecting data from multiview, and deep canonical correlation analysis was used to explore the associations between imaging and genetic data. In addition, DCCSAE was able to achieve classifications using an additional SVM classifier.
- Stage-wise deep neural network (SWDNN) (Zhou et al., 2019): a three-stage deep neural network, which was trained stage-wise. The maximum number of available subjects was used at each stage for learning feature representations for a different combination of modalities. Then, the learned feature representations were fed to a softmax output layer to achieve classification. Moreover, a majority voting strategy was used for all output layers in three stages as the final classification result.
- Deep generative-discriminative learning (DGDL) (Ko et al., 2022): a novel deep learning framework that projects genetic data nonlinearly from a single view to multimodal imaging data to perform imaging genetic analysis, as well as predict disease status.

For all compared methods, the same data partitioning was used for training and testing, and the hyperparameters were turned carefully based on their corresponding papers to make a fair comparison. For DCCSAE and ScCNAA, the top or selected imaging and genetic biomarkers were concatenated and fed to the classifier to achieve classification results. Furthermore, given that DGDL was designed to use single-modal data, it was extended with another branch to handle PET/DTI data like the branch of MRI data to adapt DGDL to our data. Since DCCSAE and DGDL cannot handle missing data problem, the variant of the proposed method (DMAAN-oImp), which was trained on paired imaging data, was also included in this experiment to make a fair comparison.

Table 8 presents the classification performance of different methods. Table 8 illustrates that our proposed DMAAN has superior performance compared with other methods. For example, DMAAN achieves remarkable results on five metrics in the classification tasks of NC vs. AD, NC vs. PD, and PD vs. SWEDD. Moreover, DMAAN obtains higher ACC (76.68%), AUC (78.08%), SEN (77.35%), and BCA (76.88%) in the task

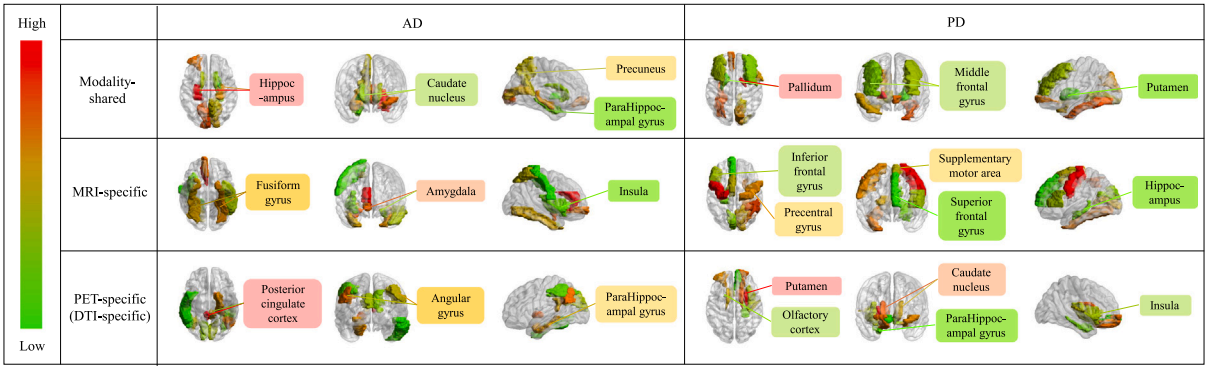

Fig. 2. The visualization results of top imaging biomarkers with the highest contribution scores in different groups of biomarkers, and different colors indicate the importance of different imaging biomarkers.

of sMCI vs. pMCI than other methods. In addition to the traditional classification tasks of NC vs. AD and NC vs. PD, the proposed DMAAN also exhibits superior performance on classification tasks of sMCI vs. pMCI and PD vs. SWEDD, thereby demonstrating that DMAAN has potential in disease classification or prediction. Meanwhile, superior classification performance can ensure that the sufficient and effective disease-related information can be captured by DMAAN. From the comparison of the results of DCCSAE and ScCNAA, the features obtained through multiview analysis are more discriminative, thereby also proving that multiview analysis has superior abilities to represent the nonlinear relationship between imaging and genetic data. In addition, the label information is discarded in DCCSAE and ScCNAA during biomarker detection, which may result in the loss of some key biomarkers, thereby reducing classification performance. SWDNN achieves the worse results than DGDL and DMAAN, implying that the association analysis between imaging and genetic data contains helpful information for classification. However, comparing the results of DGDL and DMAAN, the performance of DGDL is inferior. Apart from the effect of the established nonlinear associations between data from multiview, our learning strategy for incomplete multimodal imaging data also affects such results because DGDL cannot process the incomplete data.

To investigate the ability of DMAAN to handle missing data on real dataset, the paired imaging data were included in this section, and PET/DTI data were randomly discarded at the data missing rates of 10%, 20%, and 30%. Based on the consideration of data amount, the classification task of NC vs. AD was enrolled to evaluate the performance of different methods on missing data. Moreover, the compared methods that can handle missing data (i.e., ScCNAA and SWDNN) were selected for comparison with DMAAN. Herein, the ability of DMAAN to deal with missing data can be judged according to the classification performance of DMAAN, and the relevant results are illustrated in Table 9. According to the results in Table 9, our proposed DMAAN achieves the best performance in all missing rates, and a slight fluctuation is observed in all metrics with different missing rates, thereby showing that DMAAN has reasonable ability to handle missing data.

4.2.6. Detection of imaging biomarkers

Herein, the calculation method mentioned in Section 3.5 was used to calculate the contribution scores of different biomarkers to detect the potential biomarkers associated with NDs. Thus, the benefit of the multimodality-disentangled strategy, the modality-shared biomarkers that were prominent in both modalities, and modality-specific biomarkers that were particularly prominent in one modality can be obtained.

By averaging the calculation results of the training set, the top modality-shared, MRI-specific, and PET/DTI-specific imaging biomarkers are displayed in Fig. 2. For AD, the bilateral hippocampus was detected in the modality-shared biomarkers, and the hippocampus of AD patients shows volume atrophy on MRI and hypometabolism on PET (Chételat et al., 2008). Among the other detected modality-shared biomarkers for AD, the precuneus is correlated with a high level of cognitive function, whose atrophy and hypometabolism are characteristics of AD (Chételat et al., 2008); the caudate nucleus of AD patients is also different from that of normal subjects in MRI and PET (Koivunen et al., 2008; Madsen et al., 2010). For MRI-specific biomarkers of AD, the following imaging biomarkers are included: the amygdala atrophy is prominent in early AD, which can be observed through an MRI (Poulin et al., 2011); the fusiform gyrus contains a critical area for facial recognition, and patients with advanced AD frequently lose the ability to identify family members (Ma et al., 2020); and the volume atrophy is also present in the insula in AD patients (Sluimer et al., 2009). Meanwhile, for PET-specific AD biomarkers, a decrease of binding potential was detected in the parahippocampal gyrus from the AD brain by using PET images (Choo et al., 2014); the metabolic reduction in the posterior cingulate cortex in early AD can be observed on PET (Liang et al., 2008).

Some modality-shared biomarkers detected by using PD data have also been shown to be related to PD in some previous studies: grey matter loss and altered activation in frontal areas are related to PD (Bondi et al., 1993), where superior frontal, middle frontal, and inferior frontal gyrus were detected in this study; morphological alterations and functional connectivity changes exist in pallidum of PD patients (Garg et al., 2015; Rolinski et al., 2015), and pallidum is a modality-shared biomarker in this study. From the detected MRI-specific biomarkers, 30%–80% of PD patients eventually develop dementia in later disease stages, which is reflected in the atrophy of the hippocampus (Foo et al., 2017); precentral gyrus and supplementary motor area are associated with the bradykinesia in PD (Zhi et al., 2019; Casarotto et al., 2019). For DTI-specific biomarkers, impaired activation can be observed in PD patients' putamen (Péran et al., 2010), which is evident in functional imaging; olfactory dysfunction is often apparent in patients with PD, which is correlated with olfactory cortex and anomalies can be found in DTI (Ibarretxe-Bilbao et al., 2010); dopaminergic dysfunction in the caudate nucleus is implied in the pathophysiology of patients with PD (Péran et al., 2010); the insula is linked with cognitive decline, behavioral abnormalities, and somatosensory disturbances, and damage to the insula network occurs in patients with PD (Criaud et al., 2016). Furthermore, DMAAN provided some potential biomarkers, and their role in NDs still needs to be further investigated and proved.

4.2.7. Detection of genetic biomarkers

The relevant top SNPs were also identified using DMAAN, which are shown in Fig. 3. There was a high degree of overlap between modality-shared and modality-specific genetic biomarkers, which may be attributed to the effect of top genes that will have a great effect on all modalities. For AD, the following genes were detected in both modality-shared and modality-specific genetic biomarkers: CFAP57 and

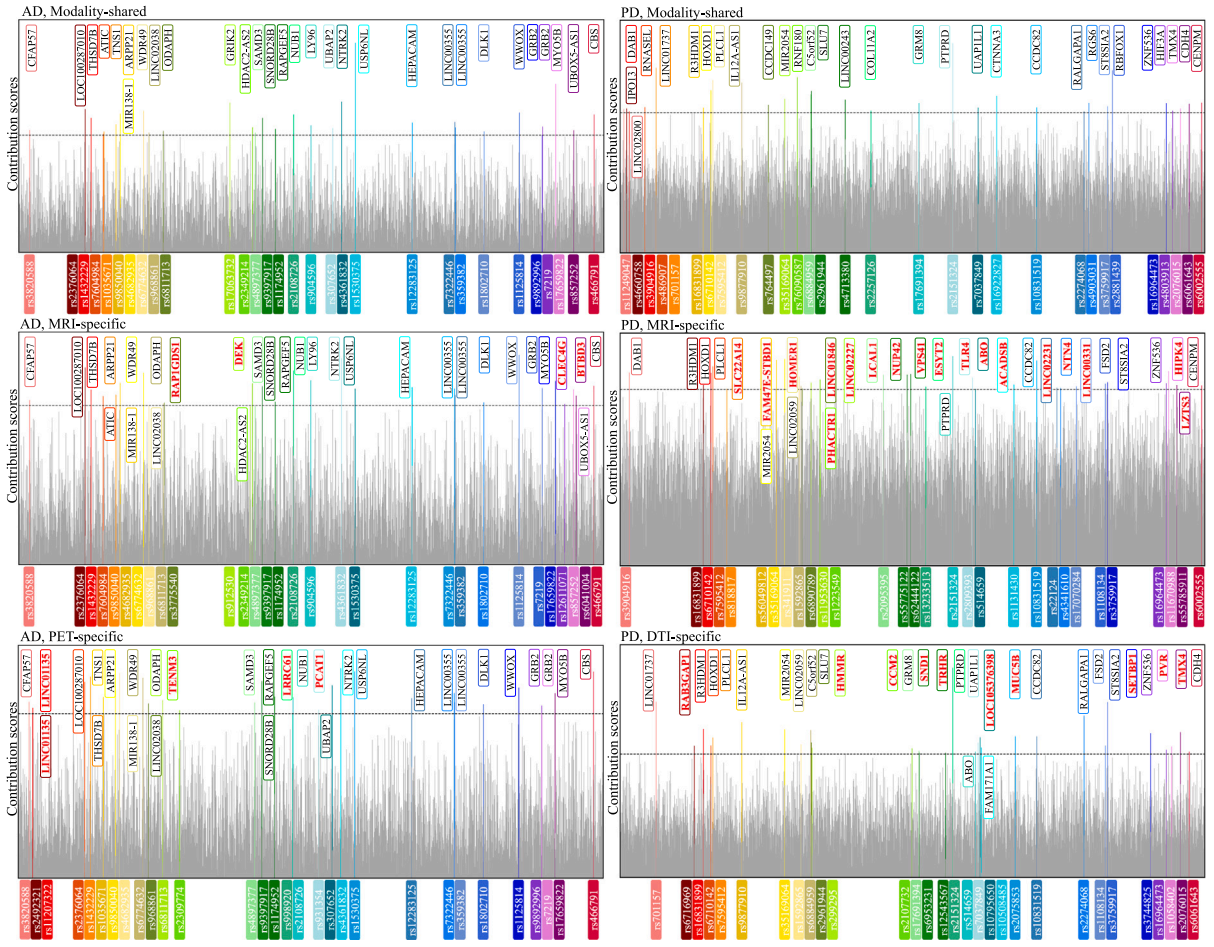

Fig. 3. The visualization results of the contribution scores of all SNPs for different NDs. Among them, the color-coded SNPs represent the top SNPs with the highest contribution score in different groups of biomarkers. Moreover, the genes marked in the red bold font are some unique modality-specific genetic biomarkers.

LINC02038 are related to cortical morphology, such as cortical thickness and cortical surface area measurement (Homann et al., 2022; Shadrin et al., 2021); NTRK2, THSD7B, and GRIK2 are linked to cognitive function and memory performance (Homann et al., 2022; Tan et al., 2021; Lee et al., 2018), and GRB2 and TNS1 affect the cognitive decline (Sherva et al., 2020; Kamboh et al., 2019); HDAC2-AS2, WWOX, USP6NL, and LINC00355 are associated with the increased AD risk (Bellenguez et al., 2022; Liu et al., 2018; Kunkle et al., 2019; Gouveia et al., 2022). However, there are still several genes that only exist in a single modality, such as DEK in MRI-specific biomarkers is highly expressed in the brain, particularly in regions relevant to learning and memory, and related to AD (Greene et al., 2020); RAP1GDS1 in MRI-specific biomarkers is associated with the volume of the hippocampus (Homann et al., 2022); and TENM3 in PET-specific biomarker is linked to schizophrenia (Goes et al., 2015).

Meanwhile, for PD, DAB1, PTPRD, CCDC82, and C5orf52 are detected by both modalities, which are reported to be highly related to PD (Tan et al., 2021; Weintraub et al., 2022; Davis et al., 2013; Ryu et al., 2020); R3HDM1, PLCL1, LINC02059, and CDH4, which also exist in both modalities, are related to dementia manifestations (Sherva et al., 2020; Bellenguez et al., 2022; Mez et al., 2017; Adewuyi et al., 2022); and GRM8 included in modality-shared and DTI-specific biomarkers is related to pallidum (Brouwer et al., 2022). Moreover, FAM47E-STBD1 and LZTS3 found in MRI-specific biomarkers are correlated with PD, where LZTS3 results in exacerbating neuronal dysfunction (Simon-Sanchez et al., 2009; Li et al., 2022); ACADSB detected in MRI-specific biomarkers is associated with Lewy body dementia (Beecham et al., 2014); and FAM171A1 and PVR belong to

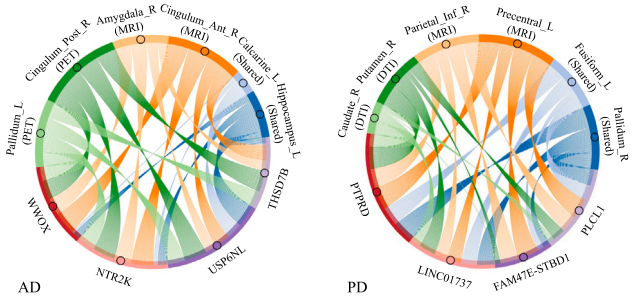

Fig. 4. Chord diagram of detected image-SNP paris, where "MRI", "PET", "DTI", and "shared" represent MRI-specific, PET-specific, DTI-specific, and modality-shared biomarkers, respectively. Each node represents an imaging biomarker or a genetic biomarker, and the thickness of the arc is proportional to the correlation coefficients. Pallidum_L/R, Cingulum_post_R, Amygdala_R, Cingulum_Ant_R, Calcarine_L, Hippocampus_L, Caudate_R, Putamen_R, Parietal_Inf_R, Precentral_L, and Fusiform_L are short for left/right pallidum, right posterior cingulate gyrus, right amygdala, right anterior cingulate gyrus, left calcarine fissure, left hippocampus, right caudate nucleus, right putamen, right inferior parietal gyrus, left precentral gyrus, and left fusiform gyrus, respectively.

DTI-specific biomarkers are linked to a polygenic risk score of cognitive decline (Gouveia et al., 2022).

4.2.8. *Biomarker detection of image-SNP pairs*

In addition to exploring modality-shared and modality-specific biomarkers that are highly associated with disease, the detection of

image-SNP pairs with high correlation is necessary. Thus, the calculation method described in Section 3.5 was modified to calculate the correlation coefficients between different imaging and genetic biomarkers. Specifically, the average in Eqs. (15)–(16) were removed, and the final contribution scores of each subject were obtained. Then, the final contribution scores of imaging biomarkers were transposed and multiplied with the final contribution scores of genetic biomarkers. Thus, the multiplied results of different contribution scores represented the correlation coefficients between different imaging and genetic biomarkers.

The top image-SNP pairs with the largest correlation coefficients can be achieved by using the above-mentioned strategy as shown in Fig. 4. Among the detected image-SNP pairs of AD, the hippocampus and amygdala are highly correlated with memory (Poulin et al., 2011); the metabolic reduction in the posterior cingulate cortex can be found in patients with AD, which is particularly prominent in PET (Liang et al., 2008); USP6NL is related to late-onset AD pathogenesis (Kunkle et al., 2019); WWOX and its binding proteins play a critical role in AD progression, and is a defined risk factor for AD (Liu et al., 2018); tropomyosin-related kinase mRNA levels related to NTRK2 is associated with the degree of cognitive impairment (Zhang et al., 2012); and THSD7B is correlated with cognitive performance (Lee et al., 2018). From detected image-SNP pairs of PD, the pallidum is a characteristic of MRI and DTI for PD (Menke et al., 2009); the volume atrophy of the hippocampus is significant for dementia patients in MRI (Foo et al., 2017); the activation of the putamen is impaired in patients with PD (Péran et al., 2010); PTPRD associated with PD achieves a high contribution score in each modality and may become a potential key gene for PD; LINC01737 contributes to the growth of basal ganglia, and basal ganglia are implicated in a wide range of movement disorders (Szekely et al., 2018); PCLC1 is related to cognitive decline and brain measurement (Mez et al., 2017; Van Der Meer et al., 2021); and FAM47E-STBD1 may be one of MRI related genes for PD. However, their associations need further exploration to confirm (Su et al., 2012).

5. Conclusion

In this study, we proposed an end-to-end deep learning framework to analyze the nonlinear associations between incomplete multimodal imaging and genetic data, and simultaneously realize disease diagnosis to detect disease-related biomarkers. Benefiting from the multimodality-disentangled strategy, we can explore the disease-related modality-shared and modality-specific biomarkers. In addition, the introduction of a multimodality-disentangled strategy improved the classification performance. Moreover, we projected the data in a nonlinear manner and established the association from multiview to accurately represent the complex relationships between imaging and genetic data, which can contribute to the subsequent biomarker detection. Furthermore, the modality mask vectors were synchronously synthesized when using genetic representing to fit the imaging representations at the association's analysis stage, and the modality mask vectors also implied the nonlinear associations between imaging and genetic data. One advantage of synthesizing the modality mask vectors was that we can use them as an attention vector and combine them with the different representations to achieve accurate classification. Another advantage was that we obtained disease-related information easily via the weights of modality mask vectors because the greater the relevance of the extracted representation dimension to the disease the higher the corresponding weights. Collectively, combined with all the above strategies, we obtained reasonable classification performance and verified the effectiveness of our proposed DMAAN on two types of NDs. Thus, DMAAN may be a new tool to understand the potential pathological mechanism of NDs.

Although competitive results were achieved compared with other methods, thus demonstrating the superiority of our proposed DMAAN, several limitations are still identified and should be addressed in future research. First, the small samples were used in the proposed method, which was still insufficient for the training of deep learning. Therefore, more samples should be collected to improve and evaluate the performance of DMAAN. Second, SNP data obtained through a global sure independence screening procedure were used in DMAAN, and severe overfitting caused by super-high-dimensional genetic data remains an open problem to be solved.

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

The authors do not have permission to share data.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 82272069, No. 81601562, No. 81974275, and No. 12126603), the Guangdong Basic and Applied Basic Research Foundation (No. 2021A1515012011), and the Science and Technology Planning Project of Guangzhou (No. 201904010417). Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu/) and the Parkinson's progression markers initiative (PPMI) database (https://www.ppmi-info.org). The investigators within the ADNI and PPMI contributed to the design and implementation of ADNI and PPMI, but did not participate in the analysis or writing of this paper.

References

Adewuyi, E.O., O'Brien, E.K., Nyholt, D.R., Porter, T., Laws, S.M., 2022. A large-scale genome-wide cross-trait analysis reveals shared genetic architecture between Alzheimer's disease and gastrointestinal tract disorders. Commun. Biol. 5 (1), 1–14.

Beecham, G.W., Hamilton, K., Naj, A.C., Martin, E.R., Huentelman, M., Myers, A.J., Corneveaux, J.J., Hardy, J., Vonsattel, J.-P., Younkin, S.G., et al., 2014. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. PLoS Genet. 10 (9), e1004606.

Bellenguez, C., Küçükali, F., Jansen, I.E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A.C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., et al., 2022. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nature Genet. 54 (4), 412–436.

Bi, X., Yang, L., Li, T., Wang, B., Zhu, H., Zhang, H., 2017. Genome-wide mediation analysis of psychiatric and cognitive traits through imaging phenotypes. Hum. Brain Mapp. 38 (8), 4088–4097.

Bonazzola, R., Ravikumar, N., Attar, R., Ferrante, E., Syeda-Mahmood, T., Frangi, A.F., 2021. Image-derived phenotype extraction for genetic discovery via unsupervised deep learning in CMR images. In: Medical Image Computing and Computer Assisted Intervention. MICCAI 2021, Springer, pp. 699–708.

Bondi, M.W., Kaszniak, A.W., Bayles, K.A., Vance, K.T., 1993. Contributions of frontal system dysfunction to memory and perceptual abilities in Parkinson's disease. Neuropsychology 7 (1), 89.

Brouwer, R.M., Klein, M., Grasby, K.L., Schnack, H.G., Jahanshad, N., Teeuw, J., Thomopoulos, S.I., Sprooten, E., Franz, C.E., Gogtay, N., et al., 2022. Genetic variants associated with longitudinal changes in brain structure across the lifespan. Nature Neurosci. 25 (4), 421–432.

Casarotto, S., Turco, F., Comanducci, A., Perretti, A., Marotta, G., Pezzoli, G., Rosanova, M., Isaias, I.U., 2019. Excitability of the supplementary motor area in Parkinson's disease depends on subcortical damage. Brain Stimul. 12 (1), 152–160.

Chandra, A., Valkimadi, P.-E., Pagano, G., Cousins, O., Dervenoulas, G., Politis, M., Alzheimer's Disease Neuroimaging Initiative, 2019. Applications of amyloid, tau, and neuroinflammation PET imaging to Alzheimer's disease and mild cognitive impairment. Hum. Brain Mapp. 40 (18), 5424–5442.

Chen, R.T., Li, X., Grosse, R.B., Duvenaud, D.K., 2018. Isolating sources of disentanglement in variational autoencoders. In: Advances in Neural Information Processing Systems. NIPS 2018, pp. 2615–2625.

Chen, X., Wang, T., Lai, H., Zhang, X., Feng, Q., Huang, M., 2022. Structure-constrained combination-based nonlinear association analysis between incomplete multimodal imaging and genetic data for biomarker detection of neurodegenerative diseases. Med. Image Anal. 78, 102419.

Cheng, J., Gao, M., Liu, J., Yue, H., Kuang, H., Liu, J., Wang, J., 2021. Multimodal disentangled variational autoencoder with game theoretic interpretability for glioma grading. IEEE J. Biomed. Health Inf. 26 (2), 673–684.

Chételat, G., Desgranges, B., Landeau, B., Mézenge, F., Poline, J., de La Sayette, V., Viader, F., Eustache, F., Baron, J.-C., 2008. Direct voxel-based comparison between grey matter hypometabolism and atrophy in Alzheimer's disease. Brain 131 (1), 60–71.

Choo, I., Carter, S.F., Schöll, M.L., Nordberg, A., 2014. Astrocytosis measured by 11C-deprenyl PET correlates with decrease in gray matter density in the parahippocampus of prodromal Alzheimer's patients. Eur. J. Nucl. Med. Mol. Imaging 41 (11), 2120–2126.

Criaud, M., Christopher, L., Boulinguez, P., Ballanger, B., Lang, A.E., Cho, S.S., Houle, S., Strafella, A.P., 2016. Contribution of insula in Parkinson's disease: A quantitative meta-analysis study. Hum. Brain Mapp. 37 (4), 1375–1392.

Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al., 2016. Next-generation genotype imputation service and methods. Nature Genet. 48 (10), 1284–1287.

Davis, M.F., Cummings, A., D'Aoust, L., Jiang, L., Velez Edwards, D., Laux, R., Reinhart-Mercer, L., Fuzzell, D., Scott, W., Pericak-Vance, M., et al., 2013. Parkinson disease loci in the mid-western Amish. Human Genetics 132 (11), 1213–1221.

Du, L., Liu, F., Liu, K., Yao, X., Risacher, S.L., Han, J., Saykin, A.J., Shen, L., 2020. Associating multi-modal brain imaging phenotypes and genetic risk factors via a dirty multi-task learning method. IEEE Trans. Med. Imaging 39 (11), 3416–3428.

Du, L., Zhang, J., Liu, F., Wang, H., Guo, L., Han, J., Alzheimer's Disease Neuroimaging Initiative, et al., 2021. Identifying associations among genomic, proteomic and imaging biomarkers via adaptive sparse multi-view canonical correlation analysis. Med. Image Anal. 70, 102003.

Foo, H., Mak, E., Chander, R.J., Ng, A., Au, W.L., Sitoh, Y.Y., Tan, L.C., Kandiah, N., 2017. Associations of hippocampal subfields in the progression of cognitive decline related to Parkinson's disease. NeuroImage Clin. 14, 37–42.

Garg, A., Appel-Cresswell, S., Popuri, K., McKeown, M.J., Beg, M.F., 2015. Morphological alterations in the caudate, putamen, pallidum, and thalamus in Parkinson's disease. Front. Neurosci. 9, 101.

Goes, F.S., McGrath, J., Avramopoulos, D., Wolyniec, P., Pirooznia, M., Ruczinski, I., Nestadt, G., Kenny, E.E., Vacic, V., Peters, I., et al., 2015. Genome-wide association study of schizophrenia in Ashkenazi jews. Am. J. Med. Genet. B Neuropsychiatr. Genet. 168 (8), 649–659.

Gouveia, C., Gibbons, E., Dehghani, N., Eapen, J., Guerreiro, R., Bras, J., 2022. Genome-wide association of polygenic risk extremes for Alzheimer's disease in the UK Biobank. Sci. Rep. 12 (1), 1–12.

Greene, A.N., Parks, L.G., Solomon, M.B., Privette Vinnedge, L.M., 2020. Loss of DEK expression induces Alzheimer's disease phenotypes in differentiated SH-SY5Y cells. Front. Mol. Neurosci. 13, 594319.

Hastie, T., Mazumder, R., Lee, J.D., Zadeh, R., 2015. Matrix completion and low-rank SVD via fast alternating least squares. J. Mach. Learn. Res. 16 (1), 3367–3402.

Homann, J., Osburg, T., Ohlei, O., Dobricic, V., Deecke, L., Bos, I., Vandenberghe, R., Gabel, S., Scheltens, P., Teunissen, C.E., et al., 2022. Genome-wide association study of Alzheimer's disease brain imaging biomarkers and neuropsychological phenotypes in the European medical information framework for Alzheimer's disease multimodal biomarker discovery dataset. Front. Aging Neurosci. 14, 840651.

Hu, W., Meng, X., Bai, Y., Zhang, A., Qu, G., Cai, B., Zhang, G., Wilson, T.W., Stephen, J.M., Calhoun, V.D., et al., 2021. Interpretable multimodal fusion networks reveal mechanisms of brain cognition. IEEE Trans. Med. Imaging 40 (5), 1474–1483.

Hu, D., Zhang, H., Wu, Z., Wang, F., Wang, L., Smith, J.K., Lin, W., Li, G., Shen, D., 2020. Disentangled-multimodal adversarial autoencoder: Application to infant age prediction with incomplete multimodal neuroimages. IEEE Trans. Med. Imaging 39 (12), 4137–4149.

Huang, M., Chen, X., Yu, Y., Lai, H., Feng, Q., 2021a. Imaging genetics study based on a temporal group sparse regression and additive model for biomarker detection of Alzheimer's disease. IEEE Trans. Med. Imaging 40 (5), 1461–1473.

Huang, M., Lai, H., Yu, Y., Chen, X., Wang, T., Feng, Q., Alzheimer's Disease Neuroimaging Initiative, et al., 2021b. Deep-gated recurrent unit and diet network-based genome-wide association analysis for detecting the biomarkers of Alzheimer's disease. Med. Image Anal. 73, 102189.

Huang, M., Nichols, T., Huang, C., Yu, Y., Lu, Z., Knickmeyer, R.C., Feng, Q., Zhu, H., Alzheimer's Disease Neuroimaging Initiative, et al., 2015. FVGWAS: Fast voxelwise genome wide association analysis of large-scale imaging genetic data. Neuroimage 118, 613–627.

Huang, M., Yu, Y., Yang, W., Feng, Q., Alzheimer's Disease Neuroimaging Initiative, 2019. Incorporating spatial–anatomical similarity into the VGWAS framework for AD biomarker detection. Bioinformatics 35 (24), 5271–5280.

Ibarretxe-Bilbao, N., Junque, C., Marti, M.-J., Valldeoriola, F., Vendrell, P., Bargallo, N., Zarei, M., Tolosa, E., 2010. Olfactory impairment in Parkinson's disease and white matter abnormalities in central olfactory areas: a voxel-based diffusion tensor imaging study. Mov. Disorders 25 (12), 1888–1894.

Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., Wick, A., Schlemmer, H.-P., Heiland, S., Wick, W., et al., 2019. Automated brain extraction of multisequence MRI using artificial neural networks. Hum. Brain Mapp. 40 (17), 4952–4964.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. Fsl. Neuroimage 62 (2), 782–790.

Jiang, L., Zheng, Z., Qi, T., Kemper, K.E., Wray, N.R., Visscher, P.M., Yang, J., 2019. A resource-efficient tool for mixed model association analysis of large-scale data. Nature Genet. 51 (12), 1749–1755.

Kamboh, M.I., Fan, K.-H., Yan, Q., Beer, J.C., Snitz, B.E., Wang, X., Chang, C.-C.H., Demirci, F.Y., Feingold, E., Ganguli, M., 2019. Population-based genome-wide association study of cognitive decline in older adults free of dementia: identification of a novel locus for the attention domain. Neurobiol. Aging 84, 239–e15.

Kim, M., Min, E.J., Liu, K., Yan, J., Saykin, A.J., Moore, J.H., Long, Q., Shen, L., 2022. Multi-task learning based structured sparse canonical correlation analysis for brain imaging genetics. Med. Image Anal. 76, 102297.

Kim, M., Won, J.H., Youn, J., Park, H., 2019. Joint-connectivity-based sparse canonical correlation analysis of imaging genetics for detecting biomarkers of Parkinson's disease. IEEE Trans. Med. Imaging 39 (1), 23–34.

Ko, W., Jung, W., Jeon, E., Suk, H.-I., 2022. A deep generative–discriminative learning for multi-modal representation in imaging genetics. IEEE Trans. Med. Imaging 41 (9), 2348–2359.

Koivunen, J., Verkkoniemi, A., Aalto, S., Paetau, A., Ahonen, J.-P., Viitanen, M., Någren, K., Rokka, J., Haaparanta, M., Kalimo, H., et al., 2008. PET amyloid ligand [11C] PIB uptake shows predominantly striatal increase in variant Alzheimer's disease. Brain 131 (7), 1845–1853.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90.

Kunkle, B.W., Grenier-Boley, B., Sims, R., Bis, J.C., Damotte, V., Naj, A.C., Boland, A., Vronskaya, M., Van Der Lee, S.J., Amlie-Wolf, A., et al., 2019. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nature Genet. 51 (3), 414–430.

Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., et al., 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genet. 50 (8), 1112–1121.

Lei, B., Zhao, Y., Huang, Z., Hao, X., Zhou, F., Elazab, A., Qin, J., Lei, H., 2020. Adaptive sparse learning using multi-template for neurodegenerative disease diagnosis. Med. Image Anal. 61, 101632.

Li, J., Amoh, B.K., McCormick, E., Tarkunde, A., Zhu, K.F., Perez, A., Mair, M., Moore, J., Shulman, J.M., Al-Ramahi, I., et al., 2022. Integration of transcriptome-wide association study with neuronal dysfunction assays provides functional genomics evidence for Parkinson's disease genes. Hum. Mol. Gen. 32 (4), 685–695.

Li, G., Han, D., Wang, C., Hu, W., Calhoun, V.D., Wang, Y.-P., 2020. Application of deep canonically correlated sparse autoencoder for the classification of schizophrenia. Comput. Methods Programs Biomed. 183, 105073.

Liang, W.S., Reiman, E.M., Valla, J., Dunckley, T., Beach, T.G., Grover, A., Niedzielko, T.L., Schneider, L.E., Mastroeni, D., Caselli, R., et al., 2008. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc. Natl. Acad. Sci. 105 (11), 4441–4446.

Liu, C., Ho, P., Lee, I.-T., Chen, Y., Chu, C.-H., Teng, C.-C., Wu, S., Sze, C.-I., Chiang, M., Chang, N., 2018. WWOX phosphorylation, signaling, and role in neurodegeneration. Front. Neurosci. 12, 563.

Lu, D., Popuri, K., Ding, G.W., Balachandar, R., Beg, M.F., Alzheimer's Disease Neuroimaging Initiative, 2018. Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease. Med. Image Anal. 46, 26–34.

Ma, D., Fetahu, I.S., Wang, M., Fang, R., Li, J., Liu, H., Gramyk, T., Iwanicki, I., Gu, S., Xu, W., et al., 2020. The fusiform gyrus exhibits an epigenetic signature for Alzheimer's disease. Clin. Epigr. 12 (1), 1–16.

Madsen, S.K., Ho, A.J., Hua, X., Saharan, P.S., Toga, A.W., Jack, Jr., C.R., Weiner, M.W., Thompson, P.M., Alzheimer's Disease Neuroimaging Initiative, et al., 2010. 3D maps localize caudate nucleus atrophy in 400 Alzheimer's disease, mild cognitive impairment, and healthy elderly subjects. Neurobiol. Aging 31 (8), 1312–1325.

Mattsson-Carlgren, N., Palmqvist, S., Blennow, K., Hansson, O., 2020. Increasing the reproducibility of fluid biomarker studies in neurodegenerative studies. Nature Commun. 11 (1), 1–11.

Menke, R.A., Scholz, J., Miller, K.L., Deoni, S., Jbabdi, S., Matthews, P.M., Zarei, M., 2009. MRI characteristics of the substantia nigra in Parkinson's disease: a combined quantitative T1 and DTI study. Neuroimage 47 (2), 435–441.

Mez, J., Chung, J., Jun, G., Kriegel, J., Bourlas, A.P., Sherva, R., Logue, M.W., Barnes, L.L., Bennett, D.A., Buxbaum, J.D., et al., 2017. Two novel loci, COBL and SLC10A2, for Alzheimer's disease in African Americans. Alzheimer's Dementia 13 (2), 119–129.

Nazmuddin, M., Van Dalen, J.-W., Borra, R.J., Stormezand, G.N., van der Horn, H.J., van der Zee, S., Boertien, J., van Laar, T., 2021. Postural and gait symptoms in de novo Parkinson's disease patients correlate with cholinergic white matter pathology. Parkinsonism Rel. Disord. 93, 43–49.

Ning, K., Chen, B., Sun, F., Hobel, Z., Zhao, L., Matloff, W., Toga, A.W., Alzheimer's Disease Neuroimaging Initiative, et al., 2018. Classifying Alzheimer's disease with brain imaging and genetic data using a neural network framework. Neurobiol. Aging 68, 151–158.

Pan, Y., Liu, M., Xia, Y., Shen, D., 2021. Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data. IEEE Trans. Pattern Anal. Mach. Intell. 44 (10), 6839–6853.

Péran, P., Cherubini, A., Assogna, F., Piras, F., Quattrocchi, C., Peppe, A., Celsis, P., Rascol, O., Demonet, J.-F., Stefani, A., et al., 2010. Magnetic resonance imaging markers of Parkinson's disease nigrostriatal signature. Brain 133 (11), 3423–3433.

Pirinen, M., Donnelly, P., Spencer, C.C., 2012. Including known covariates can reduce power to detect genetic effects in case-control studies. Nature Genet. 44 (8), 848–851.

Poulin, S.P., Dautoff, R., Morris, J.C., Barrett, L.F., Dickerson, B.C., Alzheimer's Disease Neuroimaging Initiative, et al., 2011. Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. Psychiatry Res. Neuroimaging 194 (1), 7–13.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Rolinski, M., Griffanti, L., Szewczyk-Krolikowski, K., Menke, R.A., Wilcock, G.K., Filippini, N., Zamboni, G., Hu, M.T., Mackay, C.E., 2015. Aberrant functional connectivity within the basal ganglia of patients with Parkinson's disease. NeuroImage Clin. 8, 126–132.

Ruiz, J., Mahmud, M., Modasshir, M., Shamim Kaiser, M., Alzheimer's Disease Neuroimaging Initiative, 2020. 3D DenseNet ensemble in 4-way classification of Alzheimer's disease. In: International Conference on Brain Informatics. BI 2020, Springer, pp. 85–96.

Ryu, H.-S., Park, K.W., Choi, N., Kim, J., Park, Y.-M., Jo, S., Kim, M.-J., Kim, Y.J., Kim, J., Kim, K., et al., 2020. Genomic analysis identifies new loci associated with motor complications in Parkinson's disease. Front. Neurol. 11, 570.

Shadrin, A.A., Kaufmann, T., van der Meer, D., Palmer, C.E., Makowski, C., Loughnan, R., Jernigan, T.L., Seibert, T.M., Hagler, D.J., Smeland, O.B., et al., 2021. Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. NeuroImage 244, 118603.

Shen, L., Thompson, P.M., 2019. Brain imaging genomics: integrated analysis and machine learning. Proc. IEEE 108 (1), 125–162.

Sherva, R., Gross, A., Mukherjee, S., Koesterer, R., Amouyel, P., Bellenguez, C., Dufouil, C., Bennett, D.A., Chibnik, L., Cruchaga, C., et al., 2020. Genome-wide association study of rate of cognitive decline in Alzheimer's disease patients identifies novel genes and pathways. Alzheimer's Dementia 16 (8), 1134–1145.

Simon-Sanchez, J., Schulte, C., Bras, J.M., Sharma, M., Gibbs, J.R., Berg, D., Paisan-Ruiz, C., Lichtner, P., Scholz, S.W., Hernandez, D.G., et al., 2009. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nature Genet. 41 (12), 1308–1312.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imaging 17 (1), 87–97.

Sluimer, J.D., van der Flier, W.M., Karas, G.B., van Schijndel, R., Barnes, J., Boyes, R.G., Cover, K.S., Olabarriaga, S.D., Fox, N.C., Scheltens, P., et al., 2009. Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. Eur. Radiol. 19 (12), 2826–2833.

Stephens, M.C., Brandt, V., Botas, J., 2022. The developmental roots of neurodegeneration. Neuron 110 (1), 1–3.

Su, F.-C., Chen, C.-M., Chen, Y.-C., Wu, Y.-R., 2012. LINGO-2 polymorphism and the risk of Parkinson's disease in Taiwan. Parkinsonism Rel. Disord. 18 (5), 609–611.

Szekely, E., Schwantes-An, T.-H.L., Justice, C.M., Sabourin, J.A., Jansen, P.R., Muetzel, R.L., Sharp, W., Tiemeier, H., Sung, H., White, T.J., et al., 2018. Genetic associations with childhood brain growth, defined in two longitudinal cohorts. Genet. Epidemiol. 42 (4), 405–414.

Tan, M.M., Lawton, M.A., Jabbari, E., Reynolds, R.H., Iwaki, H., Blauwendraat, C., Kanavou, S., Pollard, M.I., Hubbard, L., Malek, N., et al., 2021. Genome-wide association studies of cognitive and motor progression in parkinson's disease. Mov. Disorders 36 (2), 424–433.

The 1000 Genomes Project Consortium, et al., 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491 (7422), 56.

Tustison, N.J., Cook, P.A., Klein, A., Song, G., Das, S.R., Duda, J.T., Kandel, B.M., van Strien, N., Stone, J.R., Gee, J.C., et al., 2014. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage 99, 166–179.

Van Der Meer, D., Kaufmann, T., Shadrin, A.A., Makowski, C., Frei, O., Roelfs, D., Monereo-Sánchez, J., Linden, D.E., Rokicki, J., Alnæs, D., et al., 2021. The genetic architecture of human cortical folding. Sci. Adv. 7 (51), eabj9446.

Wang, J.X., Li, Y., Li, X., Lu, Z.-H., 2022. Alzheimer's disease classification through imaging genetic data with ignet. Front. Neurosci. 16, 846638.

Weintraub, D., Posavi, M., Fontanillas, P., Tropea, T.F., Mamikonyan, E., Suh, E., Trojanowski, J.Q., Cannon, P., Van Deerlin, V.M., 23andMe Research Team, et al., 2022. Genetic prediction of impulse control disorders in Parkinson's disease. Ann. Clin. Transl. Neurol. 9 (7), 936–949.

Wen, C., Ba, H., Pan, W., Huang, M., Alzheimer's Disease Neuroimaging Initiative, 2020. Co-sparse reduced-rank regression for association analysis between imaging phenotypes and genetic variants. Bioinformatics 36 (21), 5214–5222.

Xu, Z., Wu, C., Pan, W., Alzheimer's Disease Neuroimaging Initiative, et al., 2017. Imaging-wide association study: integrating imaging endophenotypes in GWAS. Neuroimage 159, 159–169.

Yang, D., Huang, S., Kuang, H., Du, Y., Zhang, L., 2022. Disentangled representation learning for multimodal emotion recognition. In: Proceedings of the 30th ACM International Conference on Multimedia. ACM MM 2022, pp. 1642–1651.

Yoshida, K., Yoshimoto, J., Doya, K., 2017. Sparse kernel canonical correlation analysis for discovery of nonlinear interactions in high-dimensional data. BMC Bioinformatics 18 (1), 1–11.

Zhang, F., Kang, Z., Li, W., Xiao, Z., Zhou, X., 2012. Roles of brain-derived neurotrophic factor/tropomyosin-related kinase B (BDNF/TrkB) signalling in Alzheimer's disease. J. Clin. Neurosci. 19 (7), 946–949.

Zhi, Y., Wang, M., Yuan, Y.-S., Shen, Y.-T., Ma, K.-W., Gan, C.-T., Si, Q.-Q., Wang, L.-N., Cao, S.-W., Zhang, K.-Z., 2019. The increased gray matter volumes of precentral gyri in Parkinson's disease patients with diphasic dyskinesia. Aging (Albany NY) 11 (21), 9661.

Zhou, T., Thung, K.-H., Liu, M., Shen, D., 2018. Brain-wide genome-wide association study for Alzheimer's disease via joint projection learning and sparse regression model. IEEE Trans. Biomed. Eng. 66 (1), 165–175.

Zhou, T., Thung, K.-H., Liu, M., Shi, F., Zhang, C., Shen, D., 2020. Multi-modal latent space inducing ensemble SVM classifier for early dementia diagnosis with neuroimaging data. Med. Image Anal. 60, 101630.

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

Zhu, X., Zhang, W., Fan, Y., 2018. A robust reduced rank graph regression method for neuroimaging genetic analysis. Neuroinformatics 16 (3), 351–361.