# A Deep Generative–Discriminative Learning for Multimodal Representation in Imaging Genetics

Wonjun Ko, Wonsik Jung, Eunjin Jeon, and Heung-Il Suk, *Senior Member, IEEE*

*Abstract*— **Imaging genetics, one of the foremost emerging topics in the medical imaging field, analyzes the inherent relations between neuroimaging and genetic data. As deep learning has gained widespread acceptance in many applications, pioneering studies employed deep learning frameworks for imaging genetics. However, existing approaches suffer from some limitations. First, they often adopt a simple strategy for joint learning of phenotypic and genotypic features. Second, their findings have not been extended to biomedical applications, e.g., degenerative brain disease diagnosis and cognitive score prediction. Finally, existing studies perform insufficient and inappropriate analyses from the perspective of data science and neuroscience. In this work, we propose a novel deep learning framework to simultaneously tackle the aforementioned issues. Our proposed framework learns to effectively represent the neuroimaging and the genetic data jointly, and achieves state-of-the-art performance when used for Alzheimer's disease and mild cognitive impairment identification. Furthermore, unlike the existing methods, the framework enables learning the relation between imaging phenotypes and genotypes in a nonlinear way without any prior neuroscientific knowledge. To demonstrate the validity of our proposed framework, we conducted experiments on a publicly available dataset and analyzed the results from diverse perspectives. Based on our experimental results, we believe that the proposed framework has immense potential to provide new insights and perspectives in deep learning-based imaging genetics studies.**

*Index Terms*— **Imaging genetics, deep learning, magnetic resonance imaging, single nucleotide polymorphism.**

## I. INTRODUCTION

IMAGING genetics, one of the foremost emerging topics in medical imaging, analyzes the inherent relations between neuroimaging, e.g., *magnetic resonance imaging* (MRI) and genetic data, e.g., *single nucleotide polymorphism* (SNP) in an integrated manner, and often combines other biomarker or clinical information for more robust analysis [1]. In this discipline, researchers have been devoting their efforts to obtain new insights into phenotypic and genotypic brain properties [1]. Specifically, studies are aimed at mining neuroimaging data-related information from big genetic data and further substantiating the *contribution* of the genetic data to the brain features [1], [2].

Existing imaging genetics studies can be categorized into two groups [1]: (i) *multivariate regression* methods and (ii) *bimultivariate correlation* methods. Multivariate regression methods [2]–[4] have explored complex gene associations. They have tried to regress neuroimaging data from given genetic data, mainly by exploiting regularized regression analysis. Here, the learned regression coefficients represent a direct relation between neuroimaging and genetic data, so interpretation is easy [5]. In addition, by using only a single model for the analysis, *multiple testing correction* is not necessary [1]. By contrast, the bimultivariate correlation methods [6]–[8] have revealed gene-neuroimaging associations from the high-dimensional data. Most bimultivariate models are based on regularized sparse canonical correlation analysis [6], [9]. One particular approach introduces parallel independent component analysis [10]. Furthermore, observed relations between neuroimaging and genetic data have often been employed for biomedical applications to discover their impact on *neurodegenerative brain disorders*, such as Alzheimer's disease (AD) [11], [12] or *cognitive measures*, such as mini-mental state examination (MMSE) [13].

Recent advances in deep learning have shown a methodological and practical impact on various research fields, having been used as de facto standards in many tasks, e.g., *computer vision*, *natural language processing*, and *healthcare*. Furthermore, although deep learning methods are relatively underexplored in imaging genetics, deep learning-based studies have achieved prominence in the field, having demonstrated their proficiency in joint analysis of neuroimaging and genes [14]–[17].

However, deep learning-based imaging genetics methods have some limitations. First, existing approaches [16], [17] have mostly adopted a relatively simple way to jointly analyze the phenotypic and the genotypic data, e.g., concatenating the features, instead of learning complex distributional relations of data. In other words, those methods have been

aimed to build multi-modal deep learning frameworks, instead of aiming specifically to gain new insights on their relations. Second, there are few biomedical applications of deep learning-based imaging genetics, and some existing application studies [14]–[18] have mainly focused on a single task, e.g., brain disease diagnosis. Finally, previous studies [1], [17] have involved very limited analyses on the neuroscientific plausibility of the learned feature representation because of the unfavorable *black-box* property of deep models.

In this work, we propose a novel framework to tackle the aforementioned issues simultaneously. Specifically, the proposed framework learns representations of SNP data for assisting MRI analysis by inferring the phenotype–genotype associations. Furthermore, the proposed framework not only represents SNP and MRI data, but also diagnoses brain disorders and predicts cognitive score using the SNP-attentive MRI features combined with demographic values. Finally, we investigate the learned representations from various perspectives with a generative mechanism in our framework as well as via a post-explainable method, thereby gaining neuroscientifically plausible knowledge purely in a data-driven manner. The main contributions of this work are as follows:

- First, we propose a novel deep generative and discriminative learning framework that jointly analyzes the phenotypic and genotypic data for brain disease diagnosis and cognitive score prediction.
- Second, we demonstrate the validity of our framework on a publicly available *Alzheimer's disease neuroimaging initiative* (ADNI) dataset.
- Finally, we perform a variety of analyses to interpret the nonlinear relations between phenotypic and genetic features learned by our framework and to discover potential imaging biomarkers guided by the genetic information.

This work is an extended version of our previous presentation in the 21st IEEE international conference on data mining [19]. We supplement our original work by conducting more classification experiments on earlier stages of AD, thereby validating the soundness of our proposed framework. We further perform an ablation study to support the ability of modules in our proposed framework. We also discover gene–neuroimaging associations and achieve neuroscientifically reasonable results.

This paper is organized as follows: Section II discusses the previous studies on imaging genetics methods and their biomedical applications. In Section III, we propose our novel deep learning-based imaging genetics framework that can classify brain disease and regress cognitive score. Section IV describes the dataset used in this work and detailed experimental settings. Further, we also report quantitative results of both classification and regression tasks. We analyze our framework in both quantitative and qualitative manner in Section V and finally conclude our work in Section VI.

## II. RELATED WORK

### A. Linear Model-Based Imaging Genetics

Over the past decade, statistical analyses and machine learning methods have been widely adopted in imaging genetics.

Among various methods, multivariate regression models aim to discover associations between neuroimaging and genetic data. Ge *et al.* [20] exploited a kernel machine method and trained it by a penalized squared-error loss function. Furthermore, they selected optimal kernels via the identity-by-state kernel estimation method. Alam *et al.* [21] also used a kernel machine method. They also trained the kernel machine using a penalized squared-error function and used the Gaussian kernel selection method. Hao *et al.* [2] proposed a tree-guided sparse learning method based on a group least absolute shrinkage and selection operator (LASSO), thereby regressing a set of SNPs to a single neuroimaging quantity. Vonou *et al.* [4] also regressed genome-wide SNP data to voxel-wise longitudinal neuroimaging quantities. They used penalized linear discriminant analysis and optimized it using a reduced rank loss function. Zhu *et al.* [5] introduced a robust version of graph regularized structured-sparse reduced rank regression for substantiating associations between multiple SNPs and multiple neuroimaging quantities.

By contrast, bimultivariate correlation methods focused on identifying the complex associations from high-dimensional data. Yan *et al.* [9] proposed structure-aware sparse canonical correlation analysis. Specifically, they conducted a group LASSO norm on the genetic data and graph Laplacian type norm on the neuroimaging data. Thus, they observed connected brain regions sharing similar patterns across the genetic features. Meda *et al.* [10] used a parallel independent component analysis to jointly maximize within-modality component independence and between-modality component correlation, thereby revealing multi-SNP-multi-neuroimaging quantity associations.

In this work, we propose deep learning framework-based imaging genetics. Our framework regresses neuroimaging quantities using genetic data, generates an attentive vector for assisting neuroimaging analysis, and performs biomedical applications of brain disease diagnosis and cognitive score prediction in a unified manner. Although the proposed framework performs multiple downstream tasks in a unified manner unlike existing linear model-based methods, our interpretation supports the notion that the framework learns neuroscientifically meaningful information similar to the existing linear model-based methods.

### B. Deep and Hierarchical Model-Based Imaging Genetics

Recent progress in deep learning methods has achieved remarkable results in imaging genetics. For instance, Wang *et al.* [14] devised an additive model via feedforward neural networks with random weights. They considered two advantages of modeling the nonlinear associations between neuroimaging data and genetic data and computational effectiveness over the network with backpropagation. By utilizing these characteristics, their method obtained more accurate associations between phenotypic and genotypic observations than conventional approaches.

Multi-modal data, i.e., neuroimaging and genes, have recently drawn more attention for complementary information
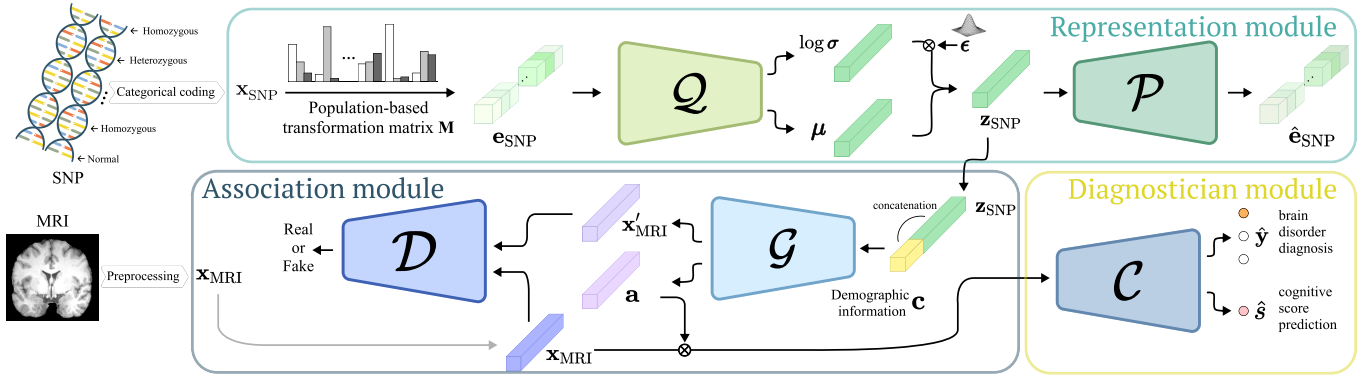
Fig. 1. Our proposed deep learning-based imaging genetics framework. In the proposed framework, a categorical genetic data vector $\mathbf{x}_{\text{SNP}}$ is transformed into a numerical embedding vector $\mathbf{e}_{\text{SNP}}$ using a population-based mapping matrix $\mathbf{M}$. Then, an encoder $\mathcal{Q}$ represents the embedding vector while a decoder $\mathcal{P}$ reconstructs the embedding. To sample the latent representation, the outputs of the encoder, $\mu$ and $\sigma$, are reparameterized with the white noise $\epsilon$. The latent representation vector $\mathbf{z}_{\text{SNP}}$ is concatenated with a demographic code $\mathbf{c}$ and fed into a generator $\mathcal{G}$. The generator outputs a fake neuroimaging data vector $\mathbf{x}'_{\text{MRI}}$ and an attentive mask vector $\mathbf{a}$. A discriminator $\mathcal{D}$ tries to distinguish the fake neuroimaging data from a real neuroimaging data $\mathbf{x}_{\text{MRI}}$. Finally, the attentive mask and the real neuroimaging are Hadamard multiplied ($\otimes$) by a diagnostician $\mathcal{C}$ and used for biomedical applications, brain disease diagnosis, $\hat{\mathbf{y}}$, and cognitive score prediction, $\hat{s}$.

and proved their validity at enhancing accuracies in brain disorder diagnosis. For example, Ning *et al.* [16] exploited volumetric measurements extracted from structural MRI as well as genotypic data including SNPs. Given those features, they trained a deep neural network and a linear regression model to predict each subject's risk of progression toward AD. Zhou *et al.* [15] proposed a three-stage deep feature learning framework to learn latent representations for each modality. They further tried to learn joint representations between multimodal data by training their deep neural network model using concatenated intermediate features. Venugopalan *et al.* [17] used MRI, SNP, and *electronic health record* data to predict AD by combining intermediate-level features. They executed feature extraction for each modality individually by employing deep neural networks to learn the intermediate features.

Compared to the existing deep network-based methods that usually learned modality-specific features first and then combined them for complementary information representation via feature concatenation [15]–[17], we devise an explicit modeling for discovering the relations between phenotypic and genetic features in a generative manner and obtain a task-predictive feature representation via genetic-induced attention learning.

## III. METHODS

In this section, we propose our novel framework to jointly represent both phenotype and genotype data. Based on the effective representation, the proposed framework diagnoses neurodegenerative brain disease and predicts cognitive score. The proposed framework has three network modules as schematized in Fig. 1. First, the proposed framework transforms a categorical-formed SNP vector $\mathbf{x}_{\text{SNP}}$ into a numerical embedding vector $\mathbf{e}_{\text{SNP}}$ and extracts a latent feature vector of the genetic data $\mathbf{z}_{\text{SNP}}$. Then, it exploits the learned latent features and demographic information $\mathbf{c}$ to generate an *attentive* mask vector $\mathbf{a}$, which is used for assisting biomedical application tasks. While generating the attentive vector, the framework also regresses fake neuroimaging $\mathbf{x}'_{\text{MRI}}$ from

the representations of $\mathbf{z}_{\text{SNP}}$ and the demographic features $\mathbf{c}$. Finally, the genotypic-attentive feature vector is used for the target tasks. Note that our proposed representation module and association module exploit generative networks to focus on probabilistic modeling between the input and the output whereas the diagnostician module tries to estimate a strict decision boundary by adopting a discriminative network.

### A. Nomenclature

In this work, we use a boldfaced capital letter $\mathbf{V}$ as a matrix. A boldfaced small letter $\mathbf{v}^{(i)}$ and a normal small letter $v^{(i,j)}$ denote respectively, the $i^{\text{th}}$ row vector and an element of the $i^{\text{th}}$ row and $j^{\text{th}}$ column of the matrix $\mathbf{V}$. Note that $v^{(i)}$ denotes the $i^{\text{th}}$ element of a column vector $\mathbf{v}^{\top}$. A normal capital letter $V$ denotes a scalar value or a set. A calligraphic letter $\mathcal{V}$ and $\Theta_{\mathcal{V}}$ are used for a neural network and its tunable parameters, respectively. $\mathcal{N}$ is used for the normal distribution.

### B. Representation Module

The proposed framework, before representing joint modalities to learn the neuroimaging-gene associations, represents the genetic data to reduce its high-dimensionality, which can be problematic [15]. To do so, the framework includes a *representation module*. Here, the input genotypic data is transformed for better representation. Then, the representation module learns latent representations for reducing the dimensionality and preparing for the attentive vector inference.

SNP values are categorical and composed of three values, normal (0), homogeneous mutant (1), and heterogeneous mutant (2), i.e., the SNP data matrix $\mathbf{X}_{\text{SNP}} \in \{0, 1, 2\}^{N_s \times F_{\text{SNP}}}$ where $N_s$ and $F_{\text{SNP}}$ respectively denote the number of training samples and the dimension of SNP features. As a deep neural network underperforms when dealing with categorical values inherently [22], we transform the categorical SNP data into numerical values. Specifically, we construct a mapping matrix $\mathbf{M} \in [0, 1]^{3 \times F_{\text{SNP}}}$ with distributional characteristics of SNPs

and define its $k^{\text{th}}$ row and $j^{\text{th}}$ column element $m^{(k,j)}$ as:

$$m^{(k,j)} \equiv \frac{1}{N_s} \sum_{i=1}^{N_s} \delta\left(x_{\text{SNP}}^{(i,j)}, k\right) \tag{1}$$

where $\delta(v, w)$ is the Kronecker-delta function that returns 1 if $v = w$ and 0 otherwise, and $x_{\text{SNP}}^{(k,j)}$ is the $j^{\text{th}}$ SNP value of the $k^{\text{th}}$ SNP sample $\mathbf{x}_{\text{SNP}}^{(k)}$. The mapping matrix $\mathbf{M}$ is population-based, as each element of this matrix shows normalized frequency of the category appearance in the training data, i.e., $\sum_k m^{(k,j)} = 1, \forall j$. By constructing the population-based mapping matrix, we suppress many zero values, which can cause challenges in SNP representation learning. Furthermore, we could better reflect importance of each SNP as, each SNP affects specific and distinct phenotypic expressions [23]. Let us assume that a particular SNP is highly related to brain disease while another one is not. Then, for brain disease diagnosis, information as to the disease-related SNP being mutant may be more important than a similar observation about the non-disease-related SNP. However, the categorical-formed SNP data has the same values for those two SNPs, and so it is not trivial to recognize different informativeness. Fortunately, population-based mapping can deal with this difference.

Using the population-based matrix, we transform the $i^{\text{th}}$ SNP vector $\mathbf{x}_{\text{SNP}}^{(i)} \in \{0, 1, 2\}^{F_{\text{SNP}}}$ into the $i^{\text{th}}$ embedded SNP vector $\mathbf{e}_{\text{SNP}}^{(i)} \in [0, 1]^{F_{\text{SNP}}}$:

$$\mathbf{e}_{\text{SNP}}^{(i)} = \left[ m^{\left(x_{\text{SNP}}^{(i,1)}, 1\right)}, m^{\left(x_{\text{SNP}}^{(i,2)}, 2\right)}, \cdots, m^{\left(x_{\text{SNP}}^{(i,F_{\text{SNP}})}, F_{\text{SNP}}\right)} \right]. \tag{2}$$

From the transformed data, we represent the genetic data to extract a latent feature vector. To represent input data, we use a *variational autoencoder* (VAE) [24]. The VAE is a branch of neural networks and learns data codings in an unsupervised manner. It has two neural networks; an encoder $\mathcal{Q}$ for the representation and a corresponding decoder $\mathcal{P}$ for the reconstruction.

In our representation module, the encoder network $\mathcal{Q}: \mathbb{R}^{F_{\text{SNP}}} \mapsto \mathbb{R}^{2 \cdot \dim(\mathbf{z}_{\text{SNP}})}$ estimates the distribution of the embedding vector for the given input where $\mathbf{z}_{\text{SNP}}$ is the represented vector and $\dim(\mathbf{v})$ denotes the dimension of $\mathbf{v}$. Here, the encoder network outputs:

$$\left(\boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{(i)}\right) = \mathcal{Q}\left(\mathbf{e}_{\text{SNP}}^{(i)}\right) \tag{3}$$

where $\boldsymbol{\mu} \in \mathbb{R}^{\dim(\mathbf{z}_{\text{SNP}})}$ and $\boldsymbol{\sigma} \in \mathbb{R}^{\dim(\mathbf{z}_{\text{SNP}})}$ denote respectively the mean and the variance of the latent distribution. Then, the latent feature vector $\mathbf{z}_{\text{SNP}}$ is sampled from the normal distribution:

$$\mathbf{z}_{\text{SNP}}^{(i)} \sim \mathcal{N}\left(\boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{(i)}\right). \tag{4}$$

Finally, the decoder network $\mathcal{P}: \mathbb{R}^{\dim(\mathbf{z}_{\text{SNP}})} \mapsto \mathbb{R}^{F_{\text{SNP}}}$ reconstructs the embedded SNP vector $\hat{\mathbf{e}}_{\text{SNP}}$ from the latent vector:

$$\hat{\mathbf{e}}_{\text{SNP}}^{(i)} = \mathcal{P}\left(\mathbf{z}_{\text{SNP}}^{(i)}\right). \tag{5}$$

## C. Association Module

In this *association module*, our framework generates the attentive vector from the genetic data representation and further learns the neuroimaging-gene associations inherently. To learn the joint representation of the neuroimaging and the genetic data, our framework also represents MRI data $\mathbf{x}_{\text{MRI}} \in \mathbb{R}^{F_{\text{MRI}}}$ where $F_{\text{MRI}}$ is the dimension of the MRI data.

In this regard, the association module, as a core module of the proposed framework, uses the latent representation $\mathbf{z}_{\text{SNP}}$ of the genetic data for generating an attentive mask vector $\mathbf{a}$ that assists neuroimaging analysis for the downstream tasks of brain disease diagnosis and cognitive score prediction. Let us denote that a generator network used for the attentive mask generation as $\mathcal{G}$.

Although the aforementioned approach can be a solution, we also exploit demographic information because it has shown its complementarities for the neuroimaging analysis and our target tasks [1], [15]. Demographic information statistically expresses *socioeconomic* information such as sex, age, and education years. In this framework, in order to have richer information, we exploit a demographic code vector $\mathbf{c}^{(i)}$:

$$\mathbf{z}^{(i)} = \text{Concat}\left(\mathbf{z}_{\text{SNP}}^{(i)}, \mathbf{c}^{(i)}\right) \tag{6}$$

where $\text{Concat}(\mathbf{v}, \mathbf{w}) = [\mathbf{v}, \mathbf{w}]$ is the concatenation operator between $\mathbf{v}$ and $\mathbf{w}$. With this enriched representation, the generator network is defined as $\mathcal{G}(\mathbf{z}) = \mathcal{G}(\text{Concat}(\mathbf{z}_{\text{SNP}}, \mathbf{c})) = \mathbf{a}$. Note that the demographic code is also used for controlling dynamics. The genetic information of a subject is never changed whereas the brain structure does during his/her life. Hence, this demographic code controls the static representation of the genetic information by providing the subject's dynamic information, e.g., age.

It should be noted that, as neuroimaging data is highly complicated, it could still be difficult to generate the neuroimaging-assisted attentive mask using only the genetic and the demographic information. To overcome this challenge, we devise a strategy for this generation module to learn distributional information of the neuroimaging data in an implicit manner. Specifically, we exploit another branch of neural networks, a *generative adversarial network* (GAN) [25]. The GAN is one of the most widely used generative models because of its ability to implicitly estimate data density. Therefore, to enjoy this benefit in our proposed framework, we finally define the generator network $\mathcal{G}: \mathbb{R}^{\dim(\mathbf{z})} \mapsto \mathbb{R}^{2 \cdot F_{\text{MRI}}}$:

$$\left(\mathbf{x}_{\text{MRI}}^{\prime(i)}, \mathbf{a}^{(i)}\right) = \mathcal{G}\left(\mathbf{z}^{(i)}\right), \tag{7}$$

where $\mathbf{x}_{\text{MRI}}^{\prime(i)}$ denotes a generated fake MRI sample. Furthermore, to help the GAN learn and generate more realistic MRI data, we also introduce a discriminator network $\mathcal{D}: \mathbb{R}^{F_{\text{MRI}}} \mapsto [0, 1]$. The discriminator network outputs a probability value as to whether an input vector $\mathbf{v}$ is real or fake, i.e., $\mathcal{D}(\mathbf{v}) \to 1$ if $\mathbf{v}$ is realistic and $\mathcal{D} \to 0$ if $\mathbf{v}$ is unrealistic. In the training steps, the discriminator network is learned to determine $\mathbf{x}_{\text{MRI}}^{\prime(i)}$ is fake while $\mathbf{x}_{\text{MRI}}^{(i)}$ is real. Note that by generating the fake MRI from the represented genetic data, the generator network inherently learns the phenotype–genotype associations. We therefore

hypothesize that, the attentive vector $\mathbf{a}$ synthesized from the generator is aware of such an association.

### D. Diagnostician Module

A *diagnostician module* is built for the downstream tasks of brain disease diagnosis and cognitive score prediction. The input neuroimaging $\mathbf{x}_{\text{MRI}}^{(i)}$ is modulated by the genetic-guided attentive mask $\mathbf{a}^{(i)}$. Then, a diagnostician network $\mathcal{C} : \mathbb{R}^{F_{\text{MRI}}} \mapsto \{0, 1\}^{|N_o|} \cup [0, 1]$ finally diagnoses brain disease $\hat{\mathbf{y}}$ and predicts a cognitive score $\hat{s}$ where $N_o$ is a set of brain statuses of interest and $|V|$ is the cardinality of a set $V$:

$$\left( \hat{\mathbf{y}}^{(i)}, \hat{s}^{(i)} \right) = \mathcal{C} \left( \mathbf{a}^{(i)} \otimes \mathbf{x}_{\text{MRI}}^{(i)} \right), \tag{9}$$

where $\mathbf{v} \otimes \mathbf{w}$ is the Hadamard multiplication between $\mathbf{v}$ and $\mathbf{w}$.

In this module, additionally, to learn multi-tasks, i.e., brain disease diagnosis and cognitive score prediction, the diagnostician network takes a hard parameter sharing strategy [26]. Let us consider the diagnostician network having multiple feedforward layers. Then, an intermediate feature vector $\mathbf{f}^{(i)}$ is extracted by the shared parameters $\mathbf{f}^{(i)} = \sigma \left( \left( \mathbf{a}^{(i)} \otimes \mathbf{x}_{\text{MRI}}^{(i)} \right) \mathbf{W} + \mathbf{b} \right)$ where $\sigma$, $\mathbf{W} \in \mathbb{R}^{F_{\text{MRI}} \times \dim(\mathbf{f})}$, and $\mathbf{b} \in \mathbb{R}^{\dim(\mathbf{f})}$ denote a nonlinear activation function, a weight matrix, and a bias vector, respectively; thus, $\{\mathbf{W}, \mathbf{b}\} \subset \Theta_{\mathcal{C}}$. It is also worth noting that the intermediate feature can be extracted by deep and hierarchical layers, but we ignore it for the sake of brevity. Finally, for the classification and the regression tasks, two different feedforward layers are defined:

$$\hat{\mathbf{y}}^{(i)} = \sigma_{\text{clf}} \left( \mathbf{f}^{(i)} \mathbf{W}_{\text{clf}} + \mathbf{b}_{\text{clf}} \right), \tag{10}$$

$$\hat{s}^{(i)} = \sigma_{\text{reg}} \left( \mathbf{f}^{(i)} \mathbf{W}_{\text{reg}} + \mathbf{b}_{\text{reg}} \right), \tag{11}$$

where $\sigma_{\text{clf}}$, $\sigma_{\text{reg}}$, $\mathbf{W}_{\text{clf}} \in \mathbb{R}^{\dim(\mathbf{f}) \times |N_o|}$, $\mathbf{W}_{\text{reg}} \in \mathbb{R}^{\dim(\mathbf{f}) \times 1}$, $\mathbf{b}_{\text{clf}} \in \mathbb{R}^{|N_o|}$, and $\mathbf{b}_{\text{reg}} \in \mathbb{R}$ are activation functions, weight matrices, and biases for the classification and the regression, respectively. Similarly, $\{\mathbf{W}_{\text{clf}}, \mathbf{W}_{\text{reg}}, \mathbf{b}_{\text{clf}}, \mathbf{b}_{\text{reg}}\} \subset \Theta_{\mathcal{C}}$. Deep and hierarchical layers can also be used for both tasks.

### E. Optimization and Learning Strategy

In our work, we train the proposed framework in an *end-to-end* manner. We define five objective functions. The first objective function optimizes the representation module. Here, we define a reconstruction loss used for the representation module by exploiting the VAE [24] to sample the latent feature vector $\mathbf{z}_{\text{SNP}}$. For the optimization, we use the *reparameterization trick* [24]; thus, we modify Eq. (4) to:

$$\boldsymbol{\epsilon} \sim \mathcal{N} \left( \mathbf{0}_{\dim(\mathbf{z}_{\text{SNP}})}, \mathbf{1}_{\dim(\mathbf{z}_{\text{SNP}})} \right)$$
$$\mathbf{z}_{\text{SNP}}^{(i)} = \boldsymbol{\mu}^{(i)} + \sqrt{\boldsymbol{\sigma}^{(i)}} \otimes \boldsymbol{\epsilon}, \tag{12}$$

where $\mathbf{0}_v$ and $\mathbf{1}_v$ respectively denote an all-zeros vector and an all-ones vector of $v$ size. Then, we estimate the *evidence*

*lower bound* (ELBO) as Eq (8), as shown at the bottom of the page, and minimize the negative ELBO for the reconstruction loss $L_{\text{rec}}$.

The second and the third functions are used for the association module. Here, we exploit a variant of GAN structures, the *least-square GAN* [27], which penalizes fake samples thereby forcing the generator to synthesize artificial samples closer toward the decision boundary. Hence, we describe a generator loss:

$$L_{\text{gen}}^{(i)} = \text{MSE} \left( \mathcal{D} \left( \mathbf{x}_{\text{MRI}}^{\prime(i)} \right), 1 \right), \tag{13}$$

where $\text{MSE}(v, w) = \mathbb{E}[(v - w)^2]$ is the mean squared error (MSE). In addition, a discriminator loss is defined as:

$$L_{\text{dis}}^{(i)} = \text{MSE} \left( \mathcal{D} \left( \mathbf{x}_{\text{MRI}}^{(i)} \right), 1 \right) + \text{MSE} \left( \mathcal{D} \left( \mathbf{x}_{\text{MRI}}^{\prime(i)} \right), 0 \right). \tag{14}$$

Finally, we estimate the fourth and the fifth functions for the classification and the regression respectively. These objective functions are used for training the diagnostician network. In this work, to control the class-imbalance issue, we use the focal cross entropy (FCE) [28] for the classification:

$$L_{\text{clf}}^{(i)} = \text{FCE} \left( \mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}, \zeta \right), \tag{15}$$

where $\text{FCE}(\mathbf{v}, \mathbf{w}, \zeta) = -\mathbb{E}[\mathbf{v}(1 - \mathbf{w})^{\zeta} \log \mathbf{w}]$ and $\mathbf{y}^{(i)}$ is the $i^{\text{th}}$ sample's disease status. Furthermore, we use the MSE function for the regression:

$$L_{\text{reg}}^{(i)} = \text{MSE} \left( s^{(i)}, \hat{s}^{(i)} \right), \tag{16}$$

where $s^{(i)}$ denotes the cognitive score of the $i^{\text{th}}$ sample.

For the overall optimization, we control the ratio of loss values by multiplying hyperparameters $\alpha_{\text{rec}}$, $\alpha_{\text{gen}}$, $\alpha_{\text{dis}}$, $\alpha_{\text{clf}}$, and $\alpha_{\text{reg}}$ to $L_{\text{rec}}$, $L_{\text{gen}}$, $L_{\text{dis}}$, $L_{\text{clf}}$, and $L_{\text{reg}}$, respectively. The complete pseudo-algorithm to train all the networks of the proposed framework is presented in Algorithm 1.

## IV. EXPERIMENTS

In this section, we describe a dataset, experimental settings, and baseline settings used for all the experiments. In addition, we also present the experimental results of the proposed network as well as the competing methods considered in this work.

### A. Dataset and Preprocessing

In this work, we used the public ADNI database to demonstrate the validity of our work. Specifically, we used 734 ADNI-1 subjects, including 211 cognitively normal (CN), 350 mild cognitive impairment (MCI), a preclinical stage of dementia, and 173 AD subjects. We used a binary-code for sex ([1, 0] for male and [0, 1] for female), a scalar education years, and a scalar age as demographic information. For our experiments, we concatenated these demographic information, thus, $\dim(\mathbf{c}) = 4$. We also exploited the MMSE score for

$$L_{\text{rec}}^{(i)} = -\text{ELBO} \left( \mathbf{e}_{\text{SNP}}^{(i)}, \mathbf{z}_{\text{SNP}}^{(i)}, \hat{\mathbf{e}}_{\text{SNP}}^{(i)} \right) = -\mathbb{E}_{\mathcal{Q} \left( \mathbf{z}_{\text{SNP}}^{(i)} | \mathbf{e}_{\text{SNP}}^{(i)} \right)} \left[ \log \mathcal{P} \left( \mathbf{e}_{\text{SNP}}^{(i)} | \mathbf{z}_{\text{SNP}}^{(i)} \right) + \log \mathcal{P} \left( \mathbf{z}_{\text{SNP}}^{(i)} \right) - \log \mathcal{Q} \left( \mathbf{z}_{\text{SNP}}^{(i)} | \mathbf{e}_{\text{SNP}}^{(i)} \right) \right] \tag{8}$$

---

**Algorithm 1** Pseudo-Code for the Proposed Framework

---

**Input**: Training dataset $\mathbf{X}_{\text{MRI}}, \mathbf{X}_{\text{SNP}}, \mathbf{C}, \mathbf{Y}, \mathbf{s}^{\top}$; network architectures $\mathcal{Q}, \mathcal{P}, \mathcal{G}, \mathcal{D}, \mathcal{C}$ and their parameters $\Theta_{\mathcal{Q}}, \Theta_{\mathcal{P}}, \Theta_{\mathcal{G}}$, $\Theta_{\mathcal{D}}, \Theta_{\mathcal{C}}$; an optimizer SGD and hyperparameters $\zeta, \eta$; loss control hyperparameters $\alpha_{\text{rec}}, \alpha_{\text{gen}}, \alpha_{\text{dis}}, \alpha_{\text{clf}}, \alpha_{\text{reg}}$

1   Estimate the SNP mapping matrix $\mathbf{M}$ where $m^{(i,j)} = \sum_{k=1}^{N_s} \delta\left(x_{\text{SNP}}^{(k,j)}, i\right)/N_s$

2   **while** *Network parameters not converged* **do**

3     Draw data $\left\{\mathbf{x}_{\text{MRI}}^{(b)}, \mathbf{x}_{\text{SNP}}^{(b)}, \mathbf{c}^{(b)}, \mathbf{y}^{(b)}, s^{(b)}\right\}$   `# Mini-batch is ignored for the brevity`

4     Map the given SNP vector $\mathbf{e}_{\text{SNP}}^{(b)} = \left[ m^{\left(x_{\text{SNP}}^{(b,1)}, 1\right)}, m^{\left(x_{\text{SNP}}^{(b,2)}, 2\right)}, \cdots, m^{\left(x_{\text{SNP}}^{(b, F_{\text{SNP}})}, F_{\text{SNP}}\right)} \right]$

5     $\left(\boldsymbol{\mu}^{(b)}, \boldsymbol{\sigma}^{(b)}\right) \leftarrow \mathcal{Q}\left(\mathbf{e}_{\text{SNP}}^{(b)}\right)$   `# Encoding embedded SNP vector`

6     $\mathbf{z}_{\text{SNP}}^{(b)} \leftarrow \boldsymbol{\mu}^{(b)} + \sqrt{\boldsymbol{\sigma}^{(b)}} \otimes \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{1})$, $\hat{\mathbf{e}}_{\text{SNP}}^{(b)} \leftarrow \mathcal{P}\left(\mathbf{z}_{\text{SNP}}^{(b)}\right)$   `# Representation & reconstruction`

7     $L_{\text{rec}}^{(b)} \leftarrow -\alpha_{\text{rec}} \, \text{ELBO}\left(\mathbf{e}_{\text{SNP}}^{(b)}, \mathbf{z}_{\text{SNP}}^{(b)}, \hat{\mathbf{e}}_{\text{SNP}}^{(b)}\right)$   `# Reconstruction loss`

8     $\mathbf{z}^{(b)} \leftarrow \text{Concat}\left(\mathbf{z}_{\text{SNP}}^{(b)}, \mathbf{c}^{(b)}\right)$   `# Concatenating latent vector with demographic vector`

9     $\left(\mathbf{x}_{\text{MRI}}'^{(b)}, \mathbf{a}^{(b)}\right) \leftarrow \mathcal{G}\left(\mathbf{z}^{(b)}\right)$   `# Generating fake MRI and attentive mask`

10    $L_{\text{gen}}^{(b)} \leftarrow \alpha_{\text{gen}} \, \text{MSE}\left(\mathcal{D}\left(\mathbf{x}_{\text{MRI}}'^{(b)}\right), 1\right)$, $L_{\text{dis}}^{(b)} \leftarrow \alpha_{\text{dis}} \, \text{MSE}\left(\mathcal{D}\left(\mathbf{x}_{\text{MRI}}^{(b)}\right), 1\right) + \alpha_{\text{dis}} \, \text{MSE}\left(\mathcal{D}\left(\mathbf{x}_{\text{MRI}}'^{(b)}\right), 0\right)$   `# GAN loss`

11    $\hat{\mathbf{y}}^{(b)}, \hat{s}^{(b)} \leftarrow \mathcal{C}\left(\mathbf{a}^{(b)} \otimes \mathbf{x}_{\text{MRI}}^{(b)}\right)$   `# Analyzing neuroimaging data`

12    $L_{\text{clf}}^{(b)} \leftarrow \alpha_{\text{clf}} \, \text{FCE}\left(\mathbf{y}^{(b)}, \hat{\mathbf{y}}^{(b)}, \zeta\right)$, $L_{\text{reg}}^{(b)} \leftarrow \alpha_{\text{reg}} \, \text{MSE}\left(s^{(b)}, \hat{s}^{(b)}\right)$   `# Classification & regression loss`

13    $\Theta_{\mathcal{Q}} \leftarrow \Theta_{\mathcal{Q}} + \text{SGD}\left(L_{\text{rec}}^{(b)} + L_{\text{gen}}^{(b)}; \eta\right)$, $\Theta_{\mathcal{P}} \leftarrow \Theta_{\mathcal{P}} + \text{SGD}\left(L_{\text{rec}}^{(b)}; \eta\right)$   `# Update representation module`

14    $\Theta_{\mathcal{D}} \leftarrow \Theta_{\mathcal{D}} + \text{SGD}\left(L_{\text{dis}}^{(b)}; \eta\right)$, $\Theta_{\mathcal{G}} \leftarrow \Theta_{\mathcal{G}} + \text{SGD}\left(L_{\text{gen}}^{(b)} + L_{\text{clf}}^{(b)} + L_{\text{reg}}^{(b)}; \eta\right)$   `# Update association module`

15    $\Theta_{\mathcal{C}} \leftarrow \Theta_{\mathcal{C}} + \text{SGD}\left(L_{\text{clf}}^{(b)}; \eta\right)$ and/or $\text{SGD}\left(L_{\text{reg}}^{(b)}; \eta\right)$   `# Update diagnostician module respectively`

---

the cognitive score regression task. Note that we performed min–max normalization upon the education years, the age, and the MMSE score.

For a personalized cure, it is very important to identify a group that is more likely to progress toward AD among MCI subjects. Therefore, for the MCI subjects, we labeled those who progressed to AD after a predefined period as progressive MCI (pMCI) subjects and the remaining ones as stable MCI (sMCI). Similar to previous studies [15], [18], [29], we set the period as 30 months. Finally, we identified 194 sMCI and 156 pMCI subjects.

In this study, we exploited the preprocessed 1.5T MRI data from the ADNI website.[1] The MRI data were acquired from a variety of scanners with individualized protocols. The MRI data had been corrected for spatial distortion. Then, following previous studies [15], [30], we additionally preprocessed the neuroimaging data to extract region-of-interest (ROI) features. We conducted anterior commissure-posterior commissure correction by using the MIPAV program[2] and intensity inhomogeneity correction using the nonparametric nonuniform intensity normalization algorithm [31]. In addition, we extracted an image of the brain using the skull-stripping algorithm [32], removed the cerebellum, and segmented tissues using the FSL package.[3] Finally, we registered the obtained brain tissues to a template using the hierarchical attribute matching mechanism for elastic registration algorithm [33] and

computed the gray matter (GM) tissue volume for each ROI. All GM volumes are further normalized using the intracranial volume. After the preprocessing, we obtained 93 ROIs neuroimaging (MRI) features, i.e., $F_{\text{MRI}} = 93$.

In this work, we exploited genetic variations to obtain microscopic information about degenerative brain diseases. We genotyped the SNP data using the Human 610-Quad BeadChip. Then, we only selected SNPs and their linakge disequilibrium associations that are highly related to AD according to the AlzGene database.[4] We adopted the illumina annotation information to select a subset of SNPs [34]. Finally, we used 2,098 SNPs, i.e., $F_{\text{SNP}} = 2098$.

### B. Experimental Settings

In this work, we compared the proposed framework with linear and task-specific popular representation methods and state-of-the-art deep learning-based representation methods.

*1) Competing Baselines:* First, we built a *random forest classifier* (RFC) and a *random forest regressor* (RFR) for the classification and the regression tasks, respectively. These random forest-based methods exploited the Gini impurity for the information gain with the bootstrapping. We also exploited a *support vector classifier* (SVC) and a *support vector regressor* (SVR) for the classification and the regression tasks, respectively. We further used a *kernel ridge regressor* (KRR), which is identical to SVR but trained via the squared error loss function. For SVC, SVR, and KRR, we concatenated

---

[1] http://www.loni.usc.edu/ADNI

[2] http://mipav.cit.nih.gov/clickwrap.php

[3] http://fsl.fmrib.ox.ac.uk/fsl/fslwiki

[4] http://www.alzgene.org

the neuroimaging data, i.e., ROIs vector of MRI with the genetic data, i.e., SNP vector. In addition, we used linear feature representation methods for phenotype–genotype data analysis [15]. We implemented *principal component analysis* (PCA) to represent multi-modal data by reducing the dimension of the concatenated data. We also constructed a *locality preserving projection* (LPP) [35] to extract components. For PCA and LPP, we decided the optimal dimension of components based on their eigenvalues computed by the *generalized eigen-decomposition* method [35]. We classified and regressed the represented features extracted from PCA and LPP using SVC and SVR, respectively. Here, we concatenated demographic information to the raw data (SVC/R) or the represented features (PCA, LPP) before the classification and the regression. We also conducted the classification and the regression experiments on a *Gaussian process classifier* (GPC) and a *Gaussian process regressor* (GPR). Note that we used the Radial basis function for both GPC and GPR. Finally, we trained a scalable tree boosting system [36] on the disease diagnosis (XGBoostC) and the cognitive score prediction (XGBoostR) tasks. We took a grid search strategy for the hyperparameters selection and trained a set of classifiers via one-versus-rest for multi-class classification tasks for all these competing methods.

For deep learning-based methods, we exploited two diverse strategies [15], [17] to deal with neuroimaging and genetic data. First, we extracted the features of the neuroimaging and the genetic data using respective *multi-layer perceptrons* (MLPs). Then, we concatenated these features (MLP-concat) similar to [17]. Finally, we used another MLP for the biomedical application tasks. Specifically, we implemented an intermediate layer with 500 hidden units and the final output layer with 50 units for the genetic data representation. For the neuroimaging data, we directly projected inputs to the output layer with 50 units. Then, the concatenated 100-dimension feature vector is fed into the shared layer with 25 units and finally mapped into the classification and the regression output nodes. In addition, we built a high-level feature fusion method (MLP-fusion) based on [15]. Specifically, we represented the neuroimaging data and the genetic data feature using auxiliary classifiers in the first stage. We also concatenated the high-level features extracted in the first stage and fed them into an MLP for the tasks. For MLPs in this strategy, we used the same hyperparameters with the MLP-concat strategy. For these deep learning strategies, we additionally concatenated the demographic information vector to the concatenated multi-modal feature vector. Owing to its representation caliber for tabular data, we further exploited TabNet [37] which is composed of an encoder and a decoder. To make tabular samples, we concatenated the training genetic samples to the training phenotype features, i.e., $\text{Concat}(\mathbf{X}_{\text{SNP}}, \mathbf{X}_{\text{MRI}}) \in \mathbb{R}^{N_s \times (F_{\text{SNP}} + F_{\text{MRI}})}$. Finally, we randomly sampled 5 subjects to generate a tabular sample. Similar to [37], we only used the decoder network for self-supervised tubular learning. Finally, the TabNet encoder diagnosed brain disease and predicted cognitive score.

*2) Proposed Methods:* For the proposed framework, we used MLPs for deep neural networks in each module.

First, for the representation module, we used 2098 input nodes ($F_{\text{SNP}}$), 500 hidden nodes, and 100 output nodes ($\dim(\text{Concat}(\boldsymbol{\mu}, \boldsymbol{\sigma}))$) for the encoder while the decoder was composed of 50 input nodes ($\dim(\mathbf{z}_{\text{SNP}})$), 500 hidden nodes, and 2098 output nodes ($F_{\text{SNP}}$). For the association module, we implemented the generator network with 54 input nodes ($\dim(\text{Concat}(\mathbf{z}_{\text{SNP}}, \mathbf{c}))$), 100 hidden nodes, and 186 output nodes ($\dim \text{Concat}(\mathbf{x}'_{\text{MRI}}, \mathbf{a})$). Furthermore, 93 input nodes ($F_{\text{MRI}}$) and 1 output nodes were used for the discriminator network. For the diagnostician module, two biomedical applications shared 93 input nodes ($\dim(\mathbf{x}_{\text{MRI}} \otimes \mathbf{a}) = F_{\text{MRI}}$). Then, each task exploited respective output nodes. Finally, for the loss balance control, we set $\alpha_{\text{rec}}$, $\alpha_{\text{gen}}$, $\alpha_{\text{dis}}$, $\alpha_{\text{clf}}$, and $\alpha_{\text{reg}}$ as 0.7, 0.5, 1.0, 1.0, and 0.7, respectively.

In our experiments, for the competing baselines, we used the *exponential linear unit* (ELU) activation function for all intermediate layers, i.e., hidden nodes. Then, we used the *softmax* and the *sigmoid* function for the classification and the regression tasks, respectively. For the proposed framework, we also used the ELU for all intermediate activations. For the encoder and the decoder output, we used the linear activation and the sigmoid, respectively. We activated generator and discriminator outputs using the sigmoid. We similarly used the softmax and the sigmoid for the brain disease diagnosis and cognitive score prediction tasks. All deep learning methods were trained with a mini-batch size of 5, an exponentially decreasing learning rate (initial value: 0.004, decreasing ratio per 1,000 steps: 4%), and a total epoch of 200. All tunable parameters were initialized by the Xavier initializer, regularized by the elastic net regularizer ($\ell_1$: 0.001, $\ell_2$: 0.001), and optimized by the Adam optimizer. Note that we performed the grid search for tunable hyperparameters selection. Specifically, we predefined hyperparameter sets, $\{100, 200, \ldots, 900\}$ for the number of nodes in intermediate layers and $\{0, 1, 2\}$ for the number of intermediate layers to select network structures of $\mathcal{Q}$, $\mathcal{P}$, and $\mathcal{G}$. We also searched the optimal loss control hyperparameters, i.e., $\alpha_{\text{rec}}$, $\alpha_{\text{gen}}$, $\alpha_{\text{dis}}$, $\alpha_{\text{clf}}$, and $\alpha_{\text{reg}}$, the mini-batch size, and the initial learning rate from predefined sets, $\{0.5, 0.6, \ldots, 1.0\}$, $\{5, 10, 15\}$, and $\{0.001, 0.002, \ldots, 0.005\}$, respectively.

We implemented all the linear machine learning models used in our experiments using Python 3.7 and all the deep learning models considered in the experiments using TensorFlow 2. We trained the proposed framework on an Intel Xeon Silver 4216 CPU and on a single Nvidia Titan RTX GPU on Ubuntu 18.04. In our experimental setting, the proposed framework took approximately 0.1 sec per a sample. All codes used in the experiments are available at an online repository.[5]

### C. Experimental Results

To demonstrate the validity of the proposed framework, we evaluated five diagnosis tasks: (i) CN-versus-AD, (ii) CN-versus-MCI, (iii) sMCI-versus-pMCI, (iv) CN-versus-MCI-versus-AD, and (v) CN-versus-sMCI-versus-pMCI-versus-AD. Furthermore, we also conducted a cognitive score

[5]https://github.com/ku-milab/ENGINE

TABLE I

EXPERIMENTAL RESULTS. WE ESTIMATED THE (MULTI-CLASS) AREA UNDER A RECEIVER OPERATING CHARACTERISTIC CURVE ((M)AUC) AND ROOT MEAN SQUARED ERROR (RMSE) FOR THE CLASSIFICATION AND THE REGRESSION TASK, RESPECTIVELY. EACH CELL DENOTES THE MEAN AND THE STANDARD DEVIATION VALUES OF EACH METHOD (ROW) ON EACH TASK (COLUMN)

| Method | Metric | CN/AD | CN/MCI | sMCI/pMCI | CN/MCI/ AD | CN/sMCI/ pMCI/AD |
|---|---|---|---|---|---|---|
| RFC | (m)AUC | $.8196 \pm .0294$ | $.5832 \pm .0501$ | $.6704 \pm .0245$ | $.5255 \pm .0294$ | $.5136 \pm .0050$ |
| RFR | RMSE | $\mathbf{2.290 \pm .2577}$ | $\mathbf{1.768 \pm .1846}$ | $1.798 \pm .1248$ | $2.538 \pm .1452$ | $2.644 \pm .1671$ |
| SVC | (m)AUC | $.8544 \pm .0170$ | $.6309 \pm .0476$ | $.6500 \pm .0518$ | $.6539 \pm .0501$ | $.6566 \pm .0553$ |
| SVR | RMSE | $2.769 \pm .3770$ | $1.942 \pm .0949$ | $1.772 \pm .0804$ | $2.518 \pm .1209$ | $2.698 \pm .0481$ |
| KRR | RMSE | $2.557 \pm .3687$ | $2.073 \pm .1303$ | $2.087 \pm .1471$ | $2.619 \pm .1335$ | $2.633 \pm .1730$ |
| PCA-SVC | (m)AUC | $.8454 \pm .0350$ | $.6070 \pm .0651$ | $.6802 \pm .0557$ | $.6723 \pm .0332$ | $.6635 \pm .0493$ |
| PCA-SVR | RMSE | $2.684 \pm .2872$ | $1.926 \pm .1070$ | $1.777 \pm .0830$ | $2.716 \pm .1100$ | $2.829 \pm .1402$ |
| LPP [35]-SVC | (m)AUC | $.8811 \pm .0231$ | $.6309 \pm .0449$ | $.6571 \pm .0790$ | $.6761 \pm .0688$ | $.6274 \pm .0638$ |
| LPP [35]-SVR | RMSE | $2.861 \pm .3093$ | $1.990 \pm .1031$ | $1.900 \pm .0957$ | $2.524 \pm .1308$ | $2.545 \pm .1477$ |
| GPC | (m)AUC | $.8761 \pm .0492$ | $.5874 \pm .0279$ | $.6643 \pm .0661$ | $.6044 \pm .0236$ | $.6526 \pm .0238$ |
| GPR | RMSE | $2.774 \pm .2758$ | $1.808 \pm .1193$ | $1.765 \pm .1086$ | $2.564 \pm .8679$ | $2.676 \pm .1699$ |
| XGBoostC [36] | (m)AUC | $.8376 \pm .0348$ | $.6555 \pm .0310$ | $.6466 \pm .0352$ | $.6951 \pm .0257$ | $.6656 \pm .1886$ |
| XGBoostR [36] | RMSE | $2.722 \pm .2423$ | $1.975 \pm .1664$ | $1.921 \pm .0804$ | $2.541 \pm .1886$ | $2.609 \pm .2018$ |
| MLP- Concat [17] | (m)AUC | $.8488 \pm .0195$ | $.5940 \pm .0436$ | $.7134 \pm .0615$ | $.6122 \pm .0263$ | $.6074 \pm .0507$ |
| | RMSE | $3.035 \pm .2905$ | $1.801 \pm .1730$ | $1.812 \pm .1306$ | $2.580 \pm .1950$ | $2.594 \pm .2233$ |
| MLP- Fusion [15] | (m)AUC | $.9091 \pm .0211$ | $.6691 \pm .0311$ | $\mathbf{.7361 \pm .0544}$ | $.7014 \pm .0481$ | $.6700 \pm .0639$ |
| | RMSE | $3.121 \pm .3088$ | $1.882 \pm .3154$ | $1.933 \pm .1502$ | $2.588 \pm .2011$ | $2.760 \pm .3671$ |
| TabNet [37] | (m)AUC | $.8713 \pm .0390$ | $.6811 \pm .0109$ | $.7055 \pm .1036$ | $.6841 \pm .0710$ | $.6635 \pm .1049$ |
| | RMSE | $2.884 \pm .2458$ | $1.924 \pm .3710$ | $1.995 \pm .0936$ | $2.636 \pm .1782$ | $2.938 \pm .1835$ |
| Proposed | (m)AUC | $\mathbf{.9231 \pm .0248}$ | $\mathbf{.6973 \pm .0281}$ | $.7351 \pm .0355$ | $\mathbf{.7133 \pm .0513}$ | $\mathbf{.6931 \pm .0510}$ |
| | RMSE | $2.718 \pm .3025$ | $1.803 \pm .0811$ | $\mathbf{1.733 \pm .0811}$ | $\mathbf{2.490 \pm .1085}$ | $\mathbf{2.553 \pm .1322}$ |
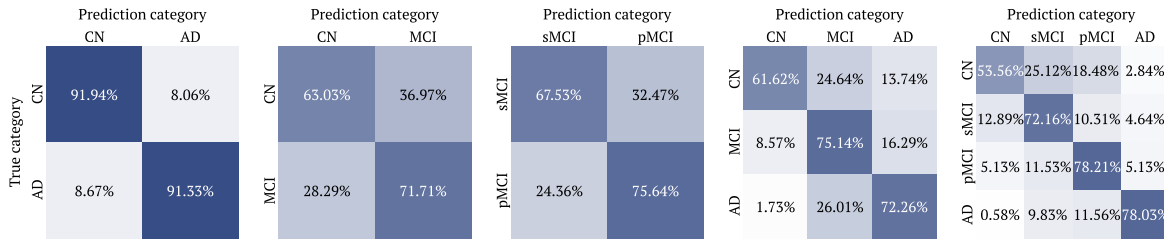


Fig. 2. Normalized averaged confusion matrices estimated by the proposed method on five diverse experimental cases.

prediction task for each diagnosis task. We used the *five-fold cross-validation* for performance estimations. Finally, we calculated the *area under the receiver operating characteristic curve* (AUC) or its multiple-class extension (mAUC) for classification tasks and estimated the *root mean squared error* (RMSE) for regression tasks. As reported in TABLE I, the proposed framework outperformed other baselines except in some cases. The proposed framework achieved the best performance on four classification tasks and three regression tasks. Note that our framework achieved slightly lower scores in cases that did not show the best performance. Additionally, even though linear model-based baselines, i.e., RFC/R, SVC/R, KRR, PCA-SVC/R, LPP-SVC/R, GPC/R, and XGBoostC/R, were trained in a single-task learning manner, the proposed framework scored the best or the second-best performance while it was optimized in a multi-task learning manner.

In addition, as shown in Fig. 2, we plotted the confusion matrices of the proposed framework for the conducted classification experiments. For each experiment, we summed confusion matrices estimated from each fold and divided each cell value by the total samples in the row-wise manner to normalize the performance. From these promising results, we concluded that the proposed framework can be useful

for biomedical applications and aid human diagnosticians in decision-making.

## V. ANALYSES

In this section, we analyze the proposed framework. We conducted an ablation study to show the importance of each component in our framework. Furthermore, we visualized the learned attentive mask vector to observe which ROIs gain attention for the neuroimaging data analysis. Finally, we discovered which SNPs were highly related in representing or generating MRI phenotypes, and the decision-making.

### A. Ablation Study

We conducted additional classification and regression experiments with six ablation cases. First, we removed the representation and the association modules and exploited only neuroimaging data for the biomedical applications (Case I). Second, we removed the representation module and used raw SNP data to generate the attentive mask vector (Case II). Third, we did not use the discriminator for the association module. Therefore, the generator network directly produces an attentive mask vector $\mathbf{a}$ without synthesizing $\mathbf{x}'_{\text{MRI}}$ (Case III). Fourth,

| Method | Metric | CN/AD | CN/MCI | sMCI/pMCI | CN/MCI/ AD | CN/sMCI/ pMCI/AD |
|---|---|---|---|---|---|---|
| Case I | (m)AUC | .8241 ± .0270 | .6417 ± .0512 | .6388 ± .0860 | .6798 ± .0851 | .6308 ± .0953 |
|  | RMSE | 3.540 ± .3819 | 2.133 ± .0855 | 2.003 ± .1105 | 2.688 ± .1860 | 2.881 ± .0571 |
| Case II | (m)AUC | .8320 ± .0421 | .6070 ± .0950 | .6125 ± .1047 | .6210 ± .1209 | .6032 ± .1370 |
|  | RMSE | 3.668 ± .3968 | 1.992 ± .1138 | 1.996 ± .0954 | 2.533 ± .1295 | 2.790 ± .1533 |
| Case III | (m)AUC | .8772 ± .0205 | .6610 ± .0443 | .6841 ± .2000 | .6624 ± .0406 | .6760 ± .0373 |
|  | RMSE | 3.304 ± .2935 | 1.833 ± .2069 | 1.941 ± .1319 | 2.581 ± .1776 | 2.633 ± .1090 |
| Case IV | (m)AUC | .8818 ± .0428 | .6646 ± .0420 | .6748 ± .1109 | .6833 ± .0318 | .6581 ± .0318 |
|  | RMSE | 3.349 ± .3492 | 2.016 ± .4709 | 1.906 ± .0971 | 2.570 ± .2025 | 2.631 ± .1129 |
| Case V | (m)AUC | .8631 ± .0320 | .6708 ± .0539 | .6935 ± .0562 | .6709 ± .0455 | .6609 ± .0688 |
|  | RMSE | 3.317 ± .3018 | 1.996 ± .3501 | 1.883 ± .1030 | 2.579 ± .1935 | 2.603 ± .1481 |
| Case VI | (m)AUC | .8850 ± .0194 | .6653 ± .0998 | .7004 ± .1049 | .6883 ± .0703 | .6814 ± .0410 |
|  | RMSE | 3.093 ± .0335 | 2.051 ± .1052 | 1.883 ± .2157 | 2.663 ± .0938 | 2.602 ± .1763 |
| Proposed | (m)AUC | **.9231 ± .0248** | **.6973 ± .0281** | **.7351 ± .0355** | **.7133 ± .0513** | **.6931 ± .0510** |
|  | RMSE | **2.718 ± .3025** | **1.803 ± .0811** | **1.733 ± .0811** | **2.490 ± .1085** | **2.553 ± .1322** |

we did not embed the genetic data and used the categorical data, i.e., exploited $\mathbf{x}_{SNP}$ instead of $\hat{\mathbf{e}}_{SNP}$ (Case IV). We further removed the decoder network as well as the reparameterization method, but only exploited the encoder network for the representation module (Case V). Finally, we ignored the demographic information $\mathbf{c}$ for estimating the attentive mask vector, i.e., $\mathbf{a} = \mathcal{G}(\mathbf{z}_{SNP})$ (Case VI).

As reported in TABLE II, the proposed framework achieved better performance than all ablation cases for two different biomedical application tasks. Specifically, without the genetic data (Case I) or its representation (Case II), the performance was dramatically lower than the proposed case. Furthermore, without the discriminator (Case III) or the SNP mapping (Case IV), there was performance degradation overall, which supports our choices. In other words, introducing the discriminator to learn the neuroimaging data distribution and transforming the categorical data into the numerical data are beneficial to the classification and the regression tasks. Finally, the representation module without the decoder network (Case V) and the association module without the demographic information (Case VI) reported performance degradation, hence we concluded that our proposed architectural designs and learning strategies effectively regularize our framework by appropriately representing the high-dimensional genetic data. Based on these results, we reached a conclusion that each module in the framework plays its own critical role in optimizing parameters with a limited number of samples.

### B. Visualizing Attentive Mask

To observe which ROIs are attended by the proposed framework, we visualized the attentive mask vector in Fig. 3. We gathered attentive mask vectors for correctly classified samples on the CN-versus-AD experiments of all folds. Then, we averaged the attentive vector in the AD class. Qualitatively, we observed that the CN attentive vector focused on brain regions evenly throughout whereas the AD attentive vector showed prominent differences between different regions. We further discovered the top-10 ROIs that showed the highest attentive value in the AD attentive vector. Remarkably, even

though the proposed framework was trained without any prior knowledge, we identified that those ROIs are highly related to existing neuroscience studies. Specifically, the proposed framework assigned a high attentive value to 'middle temporal gyrus,' 'left/right parahippocampal gyrus,' 'right hippocampal formation,' 'right entorhinal cortex,' 'left temporal pole,' 'right temporal lobe white matter,' 'right perirhinal cortex,' 'left superior frontal gyrus,' and 'left medial front-orbital gyrus' [38]–[42].

As described above, the association module in our framework requires no neuroscientific information; only genetic data is needed to generate the attentive mask vector. Nevertheless, our proposed framework showed plausible results for the attentive vector. Therefore, we believe that the proposed framework was able to learn meaningful features from the genetic data per se, and concluded that the framework takes us one step closer to data mining-based biomedical application commercialization.

### C. Subgroup Analysis in Mild Cognitive Impairment

Since MCI is considered as a preclinical stage of dementia, it can be progressed into varying disease such as AD, Parkinson's disease, etc. [43]. Therefore, it is importantly regarded to perform subgroup analysis in the literature. In this work, we also conducted clustering analysis to attentive vectors of correctly classified MCI samples in CN-versus-MCI experiments. Specifically, we clustered the synthesized attentive vectors by $k$-Means++ algorithm [44]. We set the number of clusters as $\{2, 3, \ldots, 6\}$ and estimated silhouette scores [45] for results. We obtained the best score of 0.6718 from the 5-cluster case. Finally, we extracted five centroids from the clusters.

We analyzed top-5 attended ROIs from each centroid. As shown in Fig. 4, we discovered that 'left hippocampal formation,' 'fusiform gyrus,' 'right insula,' 'right angular gyrus,' and 'superior frontal gyrus' for the first subgroup, 'middle temporal/frontal gryus,' 'paracentral lobule,' 'middle occipital gyrus,' and 'inferior frontal gyrus' for the second subgroup,
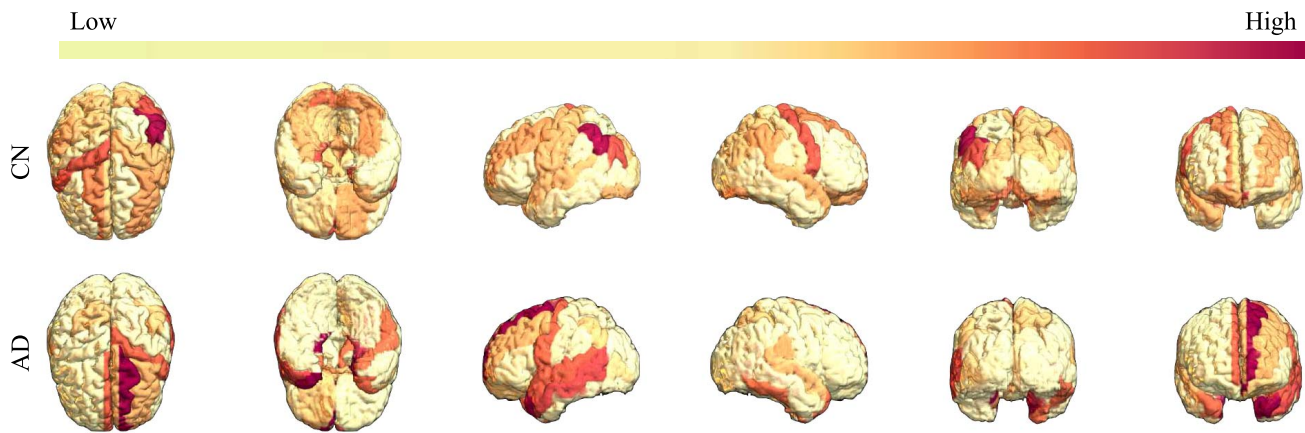
Fig. 3. Visualized attentive mask vectors of correctly classified control normal (top) and Alzheimer's disease (bottom) samples.
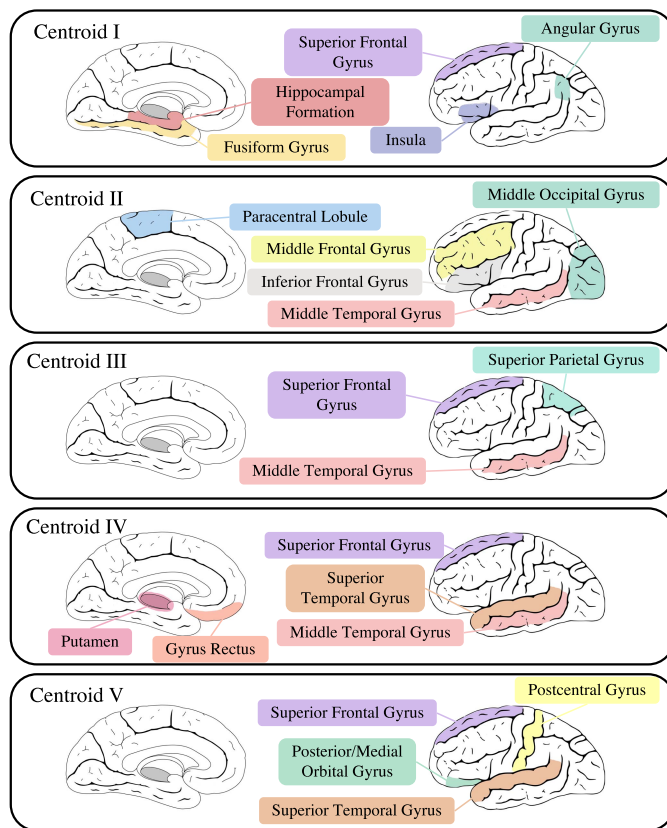


Fig. 4. Top-5 attended region of interests of centroids. All centroids were estimated by *k*-means++ clustering algorithm from correctly classified mild cognitive impairment (MCI) samples of cognitively normal-versus-MCI experiments.

'left/right middle temporal gyrus,' 'left/right superior parietal gyrus,' and 'right superior frontal gyrus' for the third subgroup, 'left superior frontal/temporal gyrus,' 'right middle temporal gyrus,' 'right putamen,' and 'left gyrus rectus' for the fourth subgroup, and 'right superior frontal/temporal gyrus,' right posterior orbital gyrus,' 'left medial orbital gyrus,' and 'right postcentral gyrus' for the final subgroup. We expect that our proposed framework can be also beneficial for subgroup analysis in the biological viewpoint, because each centroid shows diverse attended ROIs.

### D. Discovering Neuroimaging and Gene Associations

We also interpreted our trained model by means of a *gradient-weighted class activation mapping* (Grad-CAM) [46]. Here, we applied the Grad-CAM to correctly classified AD samples from all CN-versus-AD classification experiments of five folds. Specifically, we computed scores of three different flows: (i) from the diagnosis decision to the embedded SNP vector, (ii) from the generated MRI to the embedding vector, and (iii) from the high attentive ROIs to the embedding vector. We averaged the estimated scores. Then, we selected SNPs that achieve the higher relevancy scores than the predefined threshold ($5.0 \times 10^{-3}$) for each case. Finally, we discovered the genes that have the selected SNPs and visualized them in Fig. 5.

In this analysis, for the flow of $\hat{y} \rightarrow e_{SNP}$, we observed well-known top candidate genes that are associated with AD progression in previous imaging genetics studies. Our proposed framework assigned higher attention to KCNMA1, CTNNA3, DAPK1, OLR1, SOAT1, and IGF1R [15], [29], [47].

We also explored genes that are related to neuroimaging quantities. As the fake MRI is generated from the represented SNP in our generator network, this process can be considered as a traditional regression task in imaging genetics. Thus, for the case of $x'^{(i)}_{MRI} \rightarrow e_{SNP}$, we took into account the relations between specific ROIs and SNPs. We found genes that showed high relevancy with left/right hippocampal formation as follows: MTHFR, KCNMA1, FAS, and ABCG1 [48]–[50].

Finally, we computed relevancy scores between high attentive ROIs, i.e., the selected top-10 ROIs, and genes. We observed that SOAT1, CFH, APBB2, ABCA1, CTNNA3, KCNMA1, SORCS1, PICALM, and ABCG1 have strong association with the ROIs.

Based on our interpretation, we noticed that our framework could mine SNPs that were highly related to the CN-versus-AD identification and the hippocampal formation regression. Hence, we expected that the proposed framework could be used as a *starting point* for the future deep learning-based
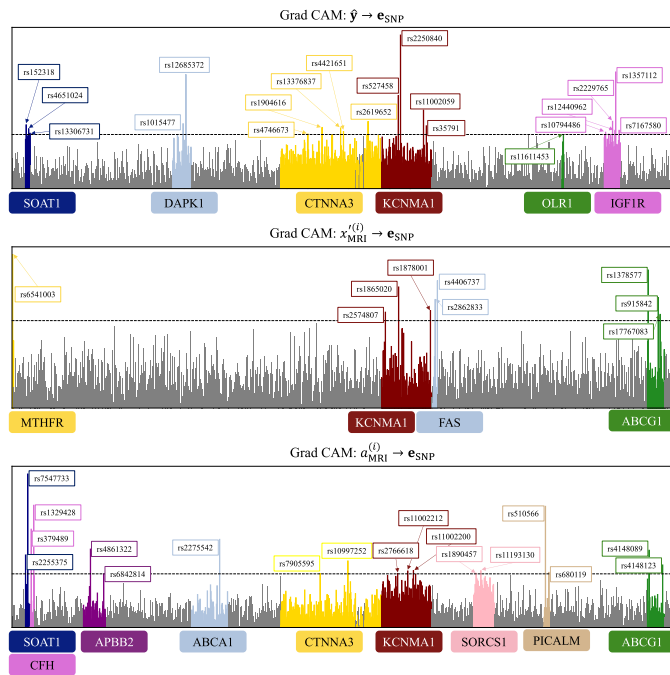
**Fig. 5.** Visualized single nucleotide polymorphism and corresponding genes showing high relevancy scores with the decision-making node (top), the left/right hippocampal formation (middle), and the high attentive regions (bottom).

imaging genetics studies as we finally explored relevancy scores between the attentive regions and genes.

## VI. CONCLUSION

In this work, we propose a novel deep learning framework for imaging genetics. The proposed framework learns the distributional properties of the neuroimaging and the genetic data, and then uses learned information for biomedical applications. Based on a variety of analyses, we observed that each module in the framework plays its own vital role, and our integrated framework represents neuroscientifically plausible features that are consistent with the results of earlier studies.

From a practical standpoint, there are still some limitations in our proposed framework. In terms of feature representation, we have used handcrafted imaging features, i.e., regional volumes, extracted by a series of preprocessing steps. Thus, it may be necessary to design a new architecture for different types of features, e.g., cortical thickness. Ultimately, it is desirable to learn task-related features from raw brain images. From the perspective of biomedical applications, diagnosing an earlier stage of MCI [51] or forecasting the progression of brain disease [52] have not been addressed. Additionally, estimating the uncertainty of decisions made by the deep learning framework is also an important concern. Therefore, it would be an interesting research direction to study Monte Carlo sampling of the attentive vector with appropriate regularization and objective functions. Based on recent trends in multi-modal learning, combining with other neuroimaging modalities, e.g., *positron emission tomography* (PET), can be a meaningful approach to improve our current work. Finally, to reduce the need for all human-centered approaches, mining a larger SNP database instead of using AlzGene-based selected SNPs will be better suited for the future deep learning-based imaging genetics study.

## REFERENCES

[1] L. Shen and P. M. Thompson, "Brain imaging genomics: Integrated analysis and machine learning," *Proc. IEEE*, vol. 108, no. 1, pp. 125–162, Jan. 2020.

[2] X. Hao *et al.*, "Identifying candidate genetic associations with MRI-derived AD-related ROI via tree-guided sparse learning," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 16, no. 6, pp. 1986–1996, Nov. 2019.

[3] H. Wang *et al.*, "Identifying quantitative trait loci via group-sparse multitask regression and feature selection: An imaging genetics study of the ADNI cohort," *Bioinformatics*, vol. 28, no. 2, pp. 229–237, 2012.

[4] M. Vounou *et al.*, "Sparse reduced-rank regression detects genetic associations with voxel-wise longitudinal phenotypes in Alzheimer's disease," *NeuroImage*, vol. 60, no. 1, pp. 700–716, 2012.

[5] X. Zhu, W. Zhang, and Y. Fan, "A robust reduced rank graph regression method for neuroimaging genetic analysis," *Neuroinformatics*, vol. 16, nos. 3–4, pp. 351–361, 2018.

[6] L. Du *et al.*, "A novel structure-aware sparse learning algorithm for brain imaging genetics," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*. Boston, MA, USA: Springer, 2014, pp. 329–336.

[7] A. Gossmann, P. Zille, V. Calhoun, and Y.-P. Wang, "FDR-corrected sparse canonical correlation analysis with applications to imaging genomics," *IEEE Trans. Med. Imag.*, vol. 37, no. 8, pp. 1761–1774, Aug. 2018.

[8] W. Hu, A. Zhang, B. Cai, V. Calhoun, and Y.-P. Wang, "Distance canonical correlation analysis with application to an imaging-genetic study," *J. Med. Imag.*, vol. 6, no. 2, 2019, Art. no. 026501.

[9] J. Yan *et al.*, "Transcriptome-guided amyloid imaging genetic analysis via a novel structured sparse learning algorithm," *Bioinformatics*, vol. 30, no. 17, pp. i564–i571, Sep. 2014.

[10] S. A. Meda *et al.*, "A large scale multivariate parallel ICA method reveals novel imaging–genetic relationships for Alzheimer's disease in the ADNI cohort," *NeuroImage*, vol. 60, no. 3, pp. 1608–1621, 2012.

[11] J. Dukart, F. Sambataro, and A. Bertolino, "Accurate prediction of conversion to Alzheimer's disease using imaging, genetic, and neuropsychological biomarkers," *J. Alzheimer's Disease*, vol. 49, no. 4, pp. 1143–1159, 2015.

[12] R. Filipovych, B. Gaonkar, and C. Davatzikos, "A composite multivariate polygenic and neuroimaging score for prediction of conversion to Alzheimer's disease," in *Proc. 2nd Int. Workshop Pattern Recognit. NeuroImaging*, Jul. 2012, pp. 105–108.

[13] K. Kauppi *et al.*, "Combining polygenic hazard score with volumetric MRI and cognitive measures improves prediction of progression from mild cognitive impairment to Alzheimer's disease," *Frontiers Neurosci.*, vol. 12, p. 260, Apr. 2018.

[14] X. Wang *et al.*, "Quantitative trait loci identification for brain endophenotypes via new additive model with random networks," *Bioinformatics*, vol. 34, no. 17, pp. i866–i874, Sep. 2018.

[15] T. Zhou, K.-H. Thung, X. Zhu, and D. Shen, "Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis," *Hum. Brain Mapping*, vol. 40, no. 3, pp. 1001–1016, 2019.

[16] K. Ning *et al.*, "Classifying Alzheimer's disease with brain imaging and genetic data using a neural network framework," *Neurobiol. Aging*, vol. 68, pp. 151–158, Aug. 2018.

[17] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," *Sci. Rep.*, vol. 11, no. 1, p. 3254, 2021.

[18] T. Zhou, M. Liu, K.-H. Thung, and D. Shen, "Latent representation learning for Alzheimer's disease diagnosis with incomplete multi-modality neuroimaging and genetic data," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2411–2422, Oct. 2019.

[19] W. Ko, W. Jung, E. Jeon, A. W. Mulyadi, and H.-I. Suk, "ENGINE: Enhancing neuroimaging and genetic information by neural embedding," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Dec. 2021, pp. 1162–1167.

[20] T. Ge *et al.*, "A kernel machine method for detecting effects of interaction between multidimensional variable sets: An imaging genetics application," *NeuroImage*, vol. 109, pp. 505–514, Apr. 2015.

[21] M. A. Alam, H.-Y. Lin, H.-W. Deng, V. D. Calhoun, and Y.-P. Wang, "A kernel machine method for detecting higher order interactions in multimodal datasets: Application to schizophrenia," *J. Neurosci. Methods*, vol. 309, pp. 161–174, Nov. 2018.

[22] J. Fürnkranz, E. Hüllermeier, E. Loza Mencía, and M. Rapp, "Learning structured declarative rule sets—A challenge for deep discrete learning," 2020, *arXiv:2012.04377*.

[23] L. B. Jorde, "Linkage disequilibrium and the search for complex disease genes," *Genome Res.*, vol. 10, no. 10, pp. 1435–1444, Oct. 2000.

[24] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2014. [Online]. Available: https://iclr.cc/archive/2014/conference-proceedings/

[25] I. J. Goodfellow *et al.*, "Generative adversarial networks," 2014, *arXiv:1406.2661*.

[26] S. Ruder, J. Bingel, I. Augenstein, and A. Søgaard, "Latent multi-task architecture learning," in *Proc. AAAI Conf. Artif. Intell.*, 2019, vol. 33, no. 1, pp. 4822–4829.

[27] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2794–2802.

[28] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.

[29] L. An, E. Adeli, M. Liu, J. Zhang, S.-W. Lee, and D. Shen, "A hierarchical feature and sample selection framework and its application for Alzheimer's disease diagnosis," *Sci. Rep.*, vol. 7, no. 1, p. 45269, 2017.

[30] H. I. Suk *et al.*, "Latent feature representation with stacked auto-encoder for AD/MCI diagnosis," *Brain Struct. Function*, vol. 220, no. 2, pp. 841–859, 2015.

[31] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1998.

[32] Y. Wang *et al.*, "Knowledge-guided robust MRI brain extraction for diverse large-scale neuroimaging studies on humans and non-human primates," *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e77810.

[33] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 11, pp. 1421–1439, Nov. 2002.

[34] L. Bertram, M. B. McQueen, K. Mullin, D. Blacker, and R. E. Tanzi, "Systematic meta-analyses of Alzheimer disease genetic association studies: The AlzGene database," *Nature Genet.*, vol. 39, no. 1, pp. 17–23, Jan. 2007.

[35] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 18, 2005, pp. 507–514.

[36] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794.

[37] S. O. Arík and T. Pfister, "TabNet: Attentive interpretable tabular learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, 2021, pp. 6679–6687.

[38] C. R. Jack *et al.*, "Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease," *Neurology*, vol. 51, no. 4, pp. 993–999, Oct. 1986.

[39] A. Biffi *et al.*, "Genetic variation and neuroimaging measures in Alzheimer disease," *Arch. Neurol.*, vol. 67, no. 6, pp. 677–685, 2010.

[40] M. Tondelli, G. K. Wilcock, P. Nichelli, C. A. De Jager, M. Jenkinson, and G. Zamboni, "Structural MRI changes detectable up to ten years before clinical Alzheimer's disease," *Neurobiol. Aging*, vol. 33, no. 4, p. 825, 2012.

[41] K. Juottonen *et al.*, "Volumes of the entorhinal and perirhinal cortices in Alzheimer's disease," *Neurobiol. Aging*, vol. 19, no. 1, pp. 15–22, 1998.

[42] K. Leng *et al.*, "Molecular characterization of selectively vulnerable neurons in Alzheimer's disease," *Nature Neurosci.*, vol. 24, no. 2, pp. 276–287, 2021.

[43] L. M. Besser *et al.*, "Mild cognitive impairment in Parkinson's disease versus Alzheimer's disease," *Parkinsonism Related Disorders*, vol. 27, pp. 54–60, Jun. 2016.

[44] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. 18th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*, Jan. 2007, pp. 1027–1035.

[45] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, 1987.

[46] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.

[47] M. Sohrabi, A. M. Floden, G. D. Manocha, M. G. Klug, and C. K. Combs, "IGF-1R inhibitor ameliorates neuroinflammation in an Alzheimer's disease transgenic mouse model," *Frontiers Cell. Neurosci.*, vol. 14, p. 200, Mar. 2020.

[48] C. S. Bailey, H. J. Moldenhauer, S. M. Park, S. Keros, and A. L. Meredith, "KCNMA1-linked channelopathy," *J. Gen. Physiol.*, vol. 151, no. 10, pp. 1173–1189, Oct. 2019.

[49] S. C. Tauber *et al.*, "Modulation of hippocampal neuroplasticity by Fas/CD95 regulatory protein 2 (Faim2) in the course of bacterial meningitis," *J. Neuropathol. Exp. Neurol.*, vol. 73, no. 1, pp. 2–13, Jan. 2014.

[50] S. Bélanger Jasmin, V. Pearson, D. Lalonde, D. Domenger, L. Théroux, and J. Poirier, "Differential regulation of ABCA1 and ABCG1 gene expressions in the remodeling mouse hippocampus after entorhinal cortex lesion and liver-X receptor agonist treatment," *Brain Res.*, vol. 1562, pp. 39–51, May 2014.

[51] J. Lee, W. Ko, E. Kang, and H.-I. Suk, "A unified framework for personalized regions selection and functional relation modeling for early MCI identification," *NeuroImage*, vol. 236, Aug. 2021, Art. no. 118048.

[52] W. Jung, E. Jun, and H.-I. Suk, "Deep recurrent model for individualized prediction of Alzheimer's disease progression," *NeuroImage*, vol. 237, Aug. 2021, Art. no. 118143.