A Multi-Classification Accessment Framework for Reproducible Evaluation of Multimodal Learning in Alzheimer's Disease

Fengtao Nan, Shunbao Li, Jiayu Wang, Yahui Tang, Jun Qi, Menghui Zhou, Zhong Zhao, Yun Yang, and Po Yang

Abstract—Multimodal learning is widely used in automated early diagnosis of Alzheimer's disease. However, the current studies are based on an assumption that different modalities can provide more complementary information to help classify the samples from the public dataset Alzheimer's Disease Neuroimaging Initiative (ADNI). In addition, the combination of modalities and different tasks are external factors that affect the performance of multimodal learning. Above all, we summrise three main problems in the early diagnosis of Alzheimer's disease: (i) unimodal vs multimodal; (ii) different combinations of modalities; (iii) classification of different tasks. In this paper, to experimentally verify these three problems, a novel and reproducible multi-classification framework for Alzheimer's disease early automatic diagnosis is proposed to evaluate and verify the above issues. The multi-classification framework contains four layers, two types of feature representation methods, and two types of models to verify these three issues. At the same time, our framework is extensible, that is, it is compatible with new modalities generated by new technologies. Following that, a series of experiments based on the ADNI-1 dataset are conducted and some possible explanations for the early diagnosis of Alzheimer's disease are obtained through multimodal learning. Experimental results show that SNP has the highest accuracy rate of 57.09% in the early diagnosis of Alzheimer's disease. In the modality combination, the addition of Single Nucleotide Polymorphism modality improves the multi-modal machine learning performance by 3% to 7%. Furthermore, we analyse and discuss the most related Region of Interest and Single Nucleotide Polymorphism features of different modalities.

Index Terms—Multi-modal learning, multi-modality data, alzheimer's disease

1 Introduction

Alzheimers's disease (AD) is an irreversible neurodegenerative disease that slowly destroys memory and thinking abilities. According to the report from the World Health Organization (WHO) in December 2020, AD and other forms of dementia are one of the top ten causes of death in the world [1]. AD is the most common dementia (approximately 60% to 70%), and its cause is still unknown. Currently, there are no treatments for AD that can prevent or reverse the course of the disease, and only a few that can temporarily relieve or improve symptoms. The latest survey report of the Alzheimer's Association in 2021, by 2050, it is estimated that 12.7 million people aged 65 and over will have AD [2].

The increase in the number of AD patients and deaths is highly valued by governments all over the world. Due to the different early diagnosis methods and equipment for AD, different forms and manifestations of data have been generated, such as cerebrospinal fluid (CSF), Positron Emission Computed Tomography (PET), Magnetic Resonance Imaging (MRI), Single Nucleotide Polymorphism (SNP), Electrocardiography (ECG), Electroencephalography (EEG) and so on. This data, which comes from several different devices or contains many different manifestations can be defined as multimodal data [3], as shown in Fig. 1. From Fig. 1, it can be seen in the course of early diagnosis of AD, will produce multimodal medical data. However, we rely on human or clinic knowledge to identify biomarkers in these data for early diagnosis. In general, it is a kind of manual diagnosis using multiple modalities. Based on the doctor's diagnosis method and the five main multimodal biomarkers shown in Fig. 1, these two points have become the basis for researchers to build an automated system for the early diagnosis of AD. Currently, machine learning technology has been widely

- *Fengtao Nan and Yahui Tang are with the School of Information, Yunnan University, Kunming 650500, China. E-mail: fengtaonan@gmail.com, bytoo_hui@mail.ynu.edu.cn.*
- *Shunbao Li and Po Yang are with the Department of Computer Science, University of Sheffield, S1 4DP Sheffield, U.K. E-mail: {leeshunbao, Poyangcn}@gmail.com.*
- *Jiayu Wang is with the National Pilot School of Software, Yunnan University, Kunming 650500, China. E-mail: bisark@ra3.in.*
- *Jun Qi is with the Department of Computing, Xi'an JiaoTong-Liverpool University, Suzhou 215123, China. E-mail: Jun.Qi@xjtlu.edu.cn.*
- *Menghui Zhou is with the National Pilot School of Software, Yunnan University, Kunming 650500, China. E-mail: menghuizhoucn@gmail.com.*
- *Zhong Zhao is with the Department of Neurology, First People's Hospital of Yunnan Province, Kunming 650500, China. E-mail: wasx-1128new@163.com.*
- *Yun Yang is with the National Pilot School of Software, Yunnan University, Kunming 650500, China. E-mail: yunyang@ynu.edu.cn.*

Manuscript received 29 December 2021; revised 26 July 2022; accepted 28 August 2022. Date of publication 6 September 2022; date of current version 8 August 2024.
This work was supported in part by the Yunnan provincial major science and technology special plan projects: digitization research and application demonstration of Yunnan characteristic industry under Grant 202002AD080001 and in part by the Natural Science Foundation of China (NSFC) under Grants 61876166 and 61663046.
(Corresponding author: Po Yang.)
Digital Object Identifier no. 10.1109/TCBB.2022.3204619

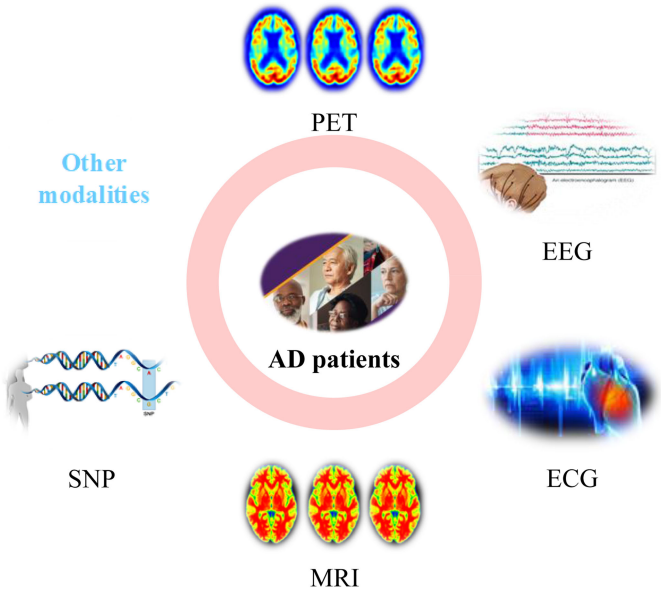


Fig. 1. AD patients produce different modal data in the hospital's diagnosis process. At present, MRI and PET are the main modalities.

used in the medical and health field[4], [5], [6], such as COVID-19 MRI reading, COVID-19 host prediction [7], Forecast of the development of COVID-19 [8] and early detection of AD and so on. In the early diagnosis of AD, many studies have been devoted to understanding the underlying biological or physiological mechanisms of AD.

Some research works [9], [10], [11], [12], [13], [14] have made unimodal analyses based on single-factor pathogenesis of AD [14], [15]. Others have conducted multimodal analyses based on multi-factor pathogenesis of AD [9], [12], [13], [16]. A large number of studies have used the public dataset Alzheimer's Disease Neuroimaging Initiative (ADNI) [17] to study the pathogenesis of AD based on Multi-modal Learning (MML), and explore the effects of mutations in brain lesions and gene fragments on AD [9], [12], [13], [18], [19].

However, a certain limitation of these methods is that they did not analyze in depth the feasibility of multimodal techniques in the early diagnosis of AD. According to previous studies [9], [12], [13], [18], [19], [20], we summarise that there are four problems in the application of MML in the early diagnosis and treatment of AD. 1) *Unimodal vs multimodal* (Explore the generalization ability of MML and traditional machine learning for early AD classification); 2) *Different combination of modalities* (Investigate the influence of the combination of modalities on the generalization ability of MML); 3) *Classification on different tasks* (Investigate the influence of the different tasks on the generalization ability of MML); 4) *Most related ROIs and SNP features* (Explore the features of the medical explanation).

To better explain and promote the reliability and effectiveness of MML in early AD diagnosis, these four issues need to be explained reasonably. The primary difficulty of this work is to adopt a unified framework to build an automated diagnosis system and to verify the four problems of MML in the early diagnosis of AD. It is difficult to obtain samples in the medical field. What is more difficult is to obtain samples with complete modalities. These factors severely limited the development of multimodal learning. In the early diagnosis of AD, the data quality was very low due to the use of different data processing techniques and standards. Incompleteness of modality is also a common problem. Our team spent a year manually processing and selecting the four modalities data on ADNI-1, and finally got 402 samples with complete modalities (MRI, PET, and SNP). Then, the choice of machine learning algorithms and the reproducibility of experimental results are other major difficulties. Different machine learning algorithms, parameters, and experimental setups greatly influence on the early diagnosis of AD and are related to the reproducibility of the automatic classification results. Finally, most researchers have only focused on the model's generalization ability but ignored these points.

In contrast to other previous work, our work is an important step in objectively evaluating the performance of MML in the task of AD diagnosis. We used the same benchmark dataset ADNI to perform detailed tests of the performance of MML in the early diagnosis of AD. At the same time, MML on AD, modal combination, number of tasks, modal contribution, and interpretable medical features were discussed. These can help researchers better design AD diagnosis experiments and understand the working mechanism of MML.

Data in the medical field is very scarce and precious. Existing studies of MML are based on an assumption that utilizing rich and completed modalities will improve the diagnosis performance of machine learning models. But the assumption of valid conditions is difficult to satisfy in practical situations, because although the improvement of the quality and quantity of multimodal data can improve the performance of machine learning models, in practical situations multimodal data always contain missing data or incomplete modal. Although Fig. 1 contains five main biomarkers, according to the previous research [9], [12], [13], [21], [22], researchers were more committed to fusing MRI, PET image data and SNP sequence data to build MML algorithms. To supplement and improve the application basis of MML in AD, we conducted detailed and precise experiments to fully and fairly verify the prior knowledge. Our team spent a year collecting valid data (MRI, PET and SNP) from 819 patients in ADNI-1, and using a unified data processing method for data preprocessing. After our manual selection, 402 patients with MRI, PET and SNP data (complete modalities) were selected for the experiment.

The purpose of the experiment is to verify the generalization ability of the model and reflect the general performance of most models in early AD diagnosis with guaranteed reproducibility, so there is no excessive adjustment of parameters. In this work, we present a framework for the reproducible assessment of MML in patients with AD and show its application in the classification task of PET, MRI, and SNP data. All the experimental results are tested under the data of 402 patients. Finally, we use feature selection algorithms and extraction methods to medically verify the selected features to prove that the features of the extracted band of machine learning are medically interpretable. Specifically, our contribution has four aspects:

- In 402 AD patients, MRI, PET, and SNP data of ADNI-1 were collected to validate that multimodal machine

TABLE 1
Participants' Demographic Information and Cognitive Scale (MMSE: Mini-Mental State Examination; CDR-SB: Clinical Dementia Rating-Sum of Boxes)

Group	Female/Male	Education	Age	MMSE	CDR-SB
NC	40/62	15.84 ± 3.13	75.84 ± 4.80	28.94 ± 1.11	0.039 ± 0.134
sMCI	27/68	15.60 ± 3.01	75.11 ± 7.33	27.35 ± 1.67	1.411 ± 0.778
pMCI	39/69	15.89 ± 2.71	74.72 ± 6.99	27.04 ± 1.69	1.634 ± 0.795
AD	39/58	14.65 ± 3.20	75.57 ± 7.28	23.47 ± 2.13	4.557 ± 1.650

learning outperforms traditional machine learning algorithms in early AD diagnosis. The experimental results show that multimodal machine learning can improve the generalization ability of the model by using complementary information.

- Through the analysis and testing of different modalities, it is found that each modality has different influences on the early diagnosis of AD. The experimental results show that different modalities contribute differently to the diagnosis of diseases. And perform medical validation on the features obtained by machine learning, analyze the consistency of these features and their importance in the medical field.
- As the number of modalities changes from unimodal to bimodal, and then to trimodal, the generalization ability of its multimodal algorithm gradually improves.

This work helps researchers to prove the effectiveness of MML in the field of smart medicine from an experimental perspective.

The rest of this paper is organized as follows. Section 2 described the materials used in this study and presented the data preprocessing steps. Section 3 explained the process of general multimodal learning in early disease diagnosis and the processing details. In Section 4, we further displayed the experimental setup and demonstrated the detailed experimental results to verify our motivation. The discussion and conclusion of this paper are in Section 5.

2 Materials and Data Preprocessing

ADNI continues to develop and standardize biomarker methods and provides qualified researchers with more depth and breadth of data. ADNI was established in 2005 and is a natural longitudinal study aimed at developing and verifying biomarkers for the selection of research subjects and as an alternative outcome measure for clinical trials of AD modification therapies. The initial 5-year study, known as ADNI-1, enrolled 819 participants from 56 study sites. The primary purpose of ADNI is to investigate the potential of fusing multimodal data, including neuroimaging, clinical, biological, and genetic biomarkers to diagnose AD at the early stage.We use data from the ADNI-1 database to verify four main issues in the multimodal medical field.

2.1 Studied Subjects

In the baseline ADNI-1 dataset, there are 819 patients' data, some of whom have incomplete modalities, such as PET data. In order to ensure that the data of each modality are complete, 402 patients with data on the three full modalities (MRI, PET and SNP) were selected from 809 patient data. Among 402 patients, there are 102 Normal Control (NC) patients, 203 Mild Cognitive Impairment (MCI) patients and 97 AD patients. MCI can be divided into stable Mild Cognitive Impairment (sMCI) and progressive Mild Cognitive Impairment (pMCI) according to whether the disease progresses to AD within a certain period of time (usually 36 months) [7]. Among them, sMCI means that MCI patients have not transformed into AD patients within a certain period, and pMCI means that MCI patients have transformed into AD patients within a certain period. In this study, there are 95 sMCI patients and 108 pMCI patients. Table 1 shows the demographic information and cognition scales of the subjects. ADNI uses different data processing methods, which leads to differences between the data.

2.2 Data Preprocessing

The 1.5T MRI baseline of ADNI-1 database are selected. According to the patient's ID, the corresponding PET and SNP are selected. In order to improve the quality of the data, MRI, PET and SNP data were processed separately. We use SPM12 to segment MRI into white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF), and extract the voxel features from 142 Region of Interests (ROIs) based on Neuromorphometrics-template. Neuromorphometrics -template is an anatomical atlas based on multiple disciplines. It was established by manual tracking of anatomical MRI of 30 healthy subjects. Hence, each issue was registered in the Montreal Neurological Institute (MNI) space to generate the maximum probability map.

Next, for each subject, we first aligned PET images to their corresponding T1 MRI using affine registration and then computed the average PET Standard Uptake Value (SUV) [17] of each brain area as a feature representation. SUV is a dimensionless ratio that has historically been used by nuclear medicine professionals to distinguish between "normal" and "abnormal" levels of uptake [17]. The SUV features we extracted from PET are of the same dimension of 142 as the MRI voxel feature. Specifically, the PET is processed under the following steps (as shown in Fig. 2):

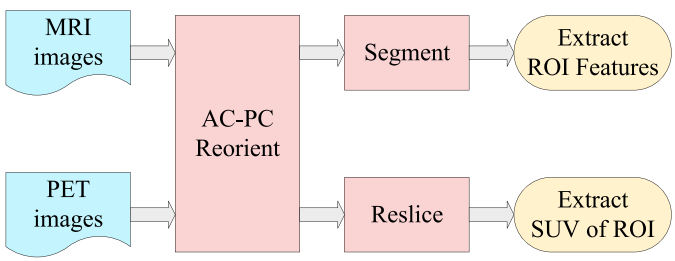


Fig. 2. MRI & PET feature extraction process.

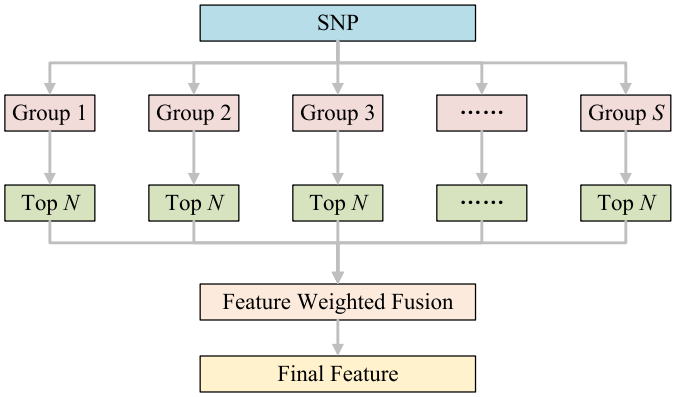


Fig. 3. Grouped dimensionality reduction and weighted fusion for high-dimensional SNP data. Where *S* represents how many groups are divided into, and *N* represents the top *N* features in each group.

1) Adjusting the PET position to anterior commissure-posterior commissure (AC-PC) correction;
2) Segmenting MRI T1 brain region to obtain the segmented deformation field;
3) Using the deformation field, Neuromorphometrics-template and MRI are used for standard space registration;
4) Matching the number of PET images per patient to the number of Neuromorphometrics-template layers;
5) Converting PET density information into SUV.

SNP is provided by ADNI and used as a genetic pathway in this study, which can provide us with microscopic information of AD. SNP samples were genotyped by using the Illumina Human610-Quad BeadChip and intensity data was processed with GenomeStudio v2.0 [18]. GenomeStudio software generates plots of all SNPs for B allele frequency (interpolated from known B allele frequencies of the three canonical clusters: 0, 0.5, and 1) and log *R* ratio ($log2_{(Robserved/Rexpected)}$), where *Rexpected* is interpolated from the observed allelic ratio concerning to the canonical genotype clusters 3, 4, 5 [19]. Finally, 402 patients with 620,901 SNPs were obtained. Due to the high dimensionality of the SNP data, which exceeds 620,901 dimensions, feature extraction methods or dimensionality reduction methods cannot be used directly, which would lead to a serious degradation in the performance and efficiency of the algorithm. To avoid this problem, we divided the 620,901 - dimensional SNP data into 10 groups, selecting the top 100 dimensions of significant features for each group, and then merging the extracted features to obtain 1000 - dimensional features. The flow chart of our grouped feature extraction is shown in Fig. 3. We used two feature extraction algorithms (random forest regression, logistic regression) for the SNP data, and finally combined the extracted features to obtain the 1000-dimensional features.

2.3 Multi-Modal Learning

In this work, we mainly investigated two mainstream multi-modal learning modeling methods [3], [23], [24]: Feature-based algorithms and Modal-based algorithms. Feature-based algorithms combine the features of heterogeneous data to realize the integration of different information. In the early feature-based fusion combined the different modal features were combined horizontally. The high-dimensional features after stitching adopt feature selection or dimensionality reductions [3], [25]. This approach reduces the efficiency of algorithms and less effective filtering of redundant information. Researchers were committed to exploring new representation methods. Latent space representation is currently the mainstream method of multi-modal feature fusion. This type of method seeks to find a latent representation to fuse different modalities [9], [18], [26].

Modal-based algorithms are another modeling strategy that explicitly solves the fusion problem in model construction — such as kernel-based methods [16], neural networks [3]. Liu et al. [16] used Multiple Kernel Learning (MKL) for multimodal fusion in AD classification. Neural networks have become very popular for multimodal fusion [16]. Using a neural network to build a multi-modal representation, each modality starts from several separate neural layers and then projected onto different modalities spatial representation of the potential hidden layer. In this section, we used latent space representation and neural networks to construct an automated diagnosis system for early AD diagnosis.

3 Framework and Application

The construction of MML automatic diagnosis system or model is currently a hot research topic in the medical field [7], [10], [12], [13], [15], [16]. In order to better improve and enrich the basic work of multimodal learning in the medical field, we verified the aforementioned four points of work by constructing a multi-modal diagnosis and treatment framework.

3.1 Multi-Classification Framework

The multi-classification framework aims to explore effective modality combinations and the performance of multi-modality methods in early AD diagnosis. The strategy is designed by using a feature representation, traditional machine learning, MML and experimental analysis approaches. The multi-classification framework is presented as a 4-layers structure as shown in Fig. 4.

Integration Level: to collect the heterogeneous data from the different platform, we divided the currently collected data into three modalities: 1) structured modalities; 2) sequential modalities; 3) image modalities. Structured modalities mainly include MMSE, CDR-SB, Electronic Medical Record (EMR), etc. These data are of low dimensionality and often contain many missing data. Sequential data mainly includes EEG, ECG, SNP, etc. The dimensionality of these data is high. Image data mainly includes MRI, PET, CT, etc. These data are often expressed in the form of 2D or 3D images. In this paper, we fused sequence modalities (SNP) and image modalities (MRI and PET). For structured modalities, because of their low dimensionality, a common approach is to splice them directly. At present, most of the data in this experiment was obtained from ADNI. Since the ADNI database currently focuses on European human brain research, another current work of our team is to collect Asian human brain data. The collected Asian brain data includes MRI and PET, and a multi-modal data platform is being developed for sharing Asian brain data.

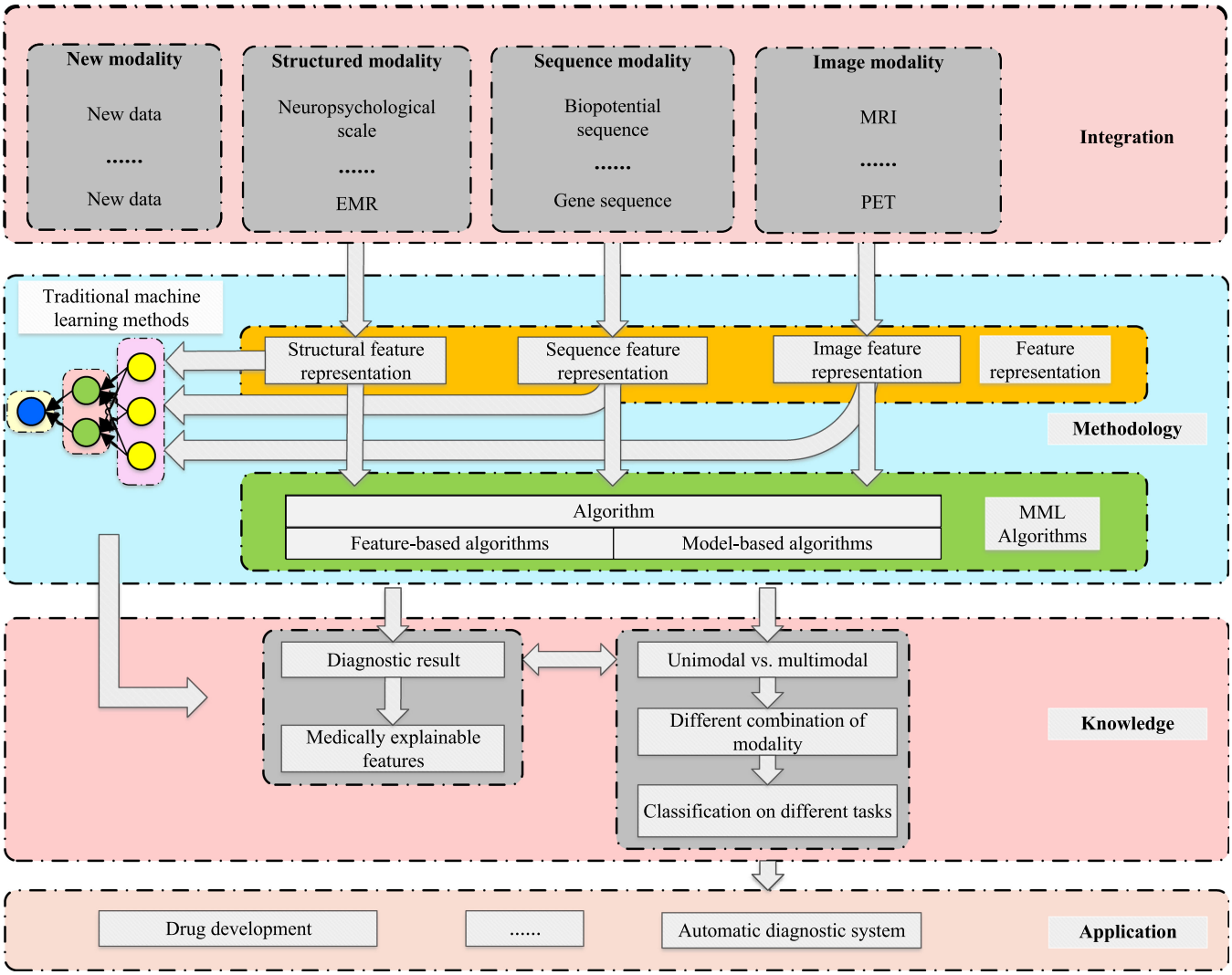


Fig. 4. Multi-classification framework for early diagnosis of AD.

Methodology Level: to represent the features of different modalities and build models. This layer mainly includes feature representation and algorithm application. At the feature representation level, different modalities are processed by different feature representations. The details of the feature representation method are described in Chapter 2. In the algorithm application layer, the traditional machine learning algorithm and MML algorithm are considered. In the choice of traditional machine learning algorithms, we introduced SVM, ensemble learning and neural networks as the main algorithms for the experiment. For the multi-modal learning methods, which are divided into two types: model-based methods and feature-based methods. Model-based methods include Multiple Kernel Learning (MKL), Multi-Task Learning (MTL), etc. Feature-based methods include Sparse learning (SL), Principal Component Analysis (PCA), Canonical Correlational Analysis (CCA), Kernel Canonical Correlation Analysis (KCCA), Deep Canonical Correlation Analysis (DCCA). Please see Section 3.2 for the details of the selected algorithm.

Knowledge Level: To summarize the output of the method level to obtain two parts of knowledge: 1) Alzheimer's related knowledge, including diagnostic results, interpretable medical features, etc. 2) Multimodal learning related knowledge, including modal combination forms, different classification tasks, etc.

Application Level: The rules and conditions obtained at the knowledge level can be applied to actual AD diagnosis and treatment, or new drug development. Specifically, to build an automated AD early diagnosis system, We can introduce some a priori knowledge, such as reasonable forms of modal combinations, valid special expressions, etc. To simplify and help the construction of the system.

Meanwhile, the purpose of the multi-classification framework is to discover some rules and factors and improve the current multi-modal automatic diagnosis system for AD by introducing prior knowledge and verifying the generalization performance of ML algorithms and MML algorithms while ensuring repeatability. We explained the knowledge level (Section 3.1) from two perspectives: AD diagnosis and Multi-Modal Learning respectively.

Unimodal Versus Multimodal: To evaluate the performance and effectiveness of MML in the early diagnosis of AD. Through a large number of experiments, we compared the performance of traditional machine learning algorithms

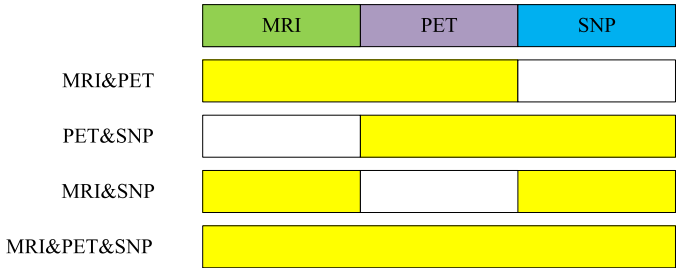


Fig. 5. The three modality combinations are formed by different combinations, and the two combinations have three forms, which are expressed as MRI & PET, MRI & PET and PET & SNP; the three modality combinations have one form, which is expressed as MRI & PET & SNP.

and multi-modal learning and verified the generalization ability of different multi-modal learning methods in early AD diagnosis.

Different Combination of Modality: To explore what kind of modal combination has the best performance of the model trained on early AD diagnosis, in the three modalities of MRI, PET and SNP. The combination of modalities is shown in Fig. 5. We also explored the relationship between the number of modalities and the performance of multi-modal learning. Last, we provided prior knowledge to avoid problems caused by modal selection.

Classification on Different Tasks: To explore the relationship between various classification tasks and multimodal learning performance, and find the best performing modal combination on the three classification tasks. This helps to improve the efficiency and accuracy of constructing a multimodal learning automated diagnosis of AD. As shown in Table 2, depicts the division of the three different tasks and the number of patients.

3.2 Classification Model

We describe the methods used by the knowledge layer, including 9 traditional machine learning algorithms and 9 multi-modal learning algorithms. Specifically, for traditional unimodal machine learning methods, we selected: k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Multilayer Perceptron (MLP), Decision Tree (DT), Random Forest (RF), Adaptive Boosting (AdaBoost), Extreme Gradient Boosting (XGBoost), Neural Networks (NNs), and Convolutional neural network (CNN). For multimodal learning methods, we selected: a representative Sparse Learning (SL) [27] algorithm: $L_{2,1}$-norm regularization [28], Principal Component Analysis (PCA), Canonical Correlation Analysis (CCA) [16], Kernel Canonical Correlation Analysis (KCCA) [21], Multi-View Multidimensional Scaling (MVMDS) [29], Multi-task learning (MTL), Co-regularized Multi-view Spectral Clustering (CMSC) [30], and Deep Semi-NMF for Multi-view Clustering (DMF-MVC) [31]. it is worth noting that MVMDS method expands the application scenarios of Multidimensional Scaling (MDS) (from unimodal to multimodal). This method can adaptively select discriminative views and enhance the contribution of information views. CMSC adopts the co-regularization method to constract a cluster representation for multimodal data. In the experiment, we used CMSC to represent multimodal data, and then trained a SVM model. DMF-MVC utilises a deep structure through semi-nonnegative matrix factorization to seek a common feature representation.

TABLE 2
Description of the Dvision of Three Different Tasks and the Number of Patients

Tasks	Group	Number of subjects
2 classification tasks	sMCI	95
	pMCI	108
3 classification tasks	NC	102
	MCI	203
	AD	97
4 classification tasks	NC	102
	sMCI	95
	pMCI	108
	AD	97

4 Experiments

In this section, we first introduced nine traditional machine learning algorithms and nine MML methods. Next, the parameter settings of each algorithm are demonstrated. The parameters of all the algorithms in the experiments are determined by grid search, which is an exhaustive search of

TABLE 3
The Parameters of Traditional Machine Learning Algorithms and Their Parameter Value Ranges

Tasks	Group	Number of subjects
SVM	Kernel	[RBF, Linear]
	Penalty *C*	[1, 10 100,1000]
	Gamma	[1e-1,1e-2,1e-3, 1e-4]
NNs	Hidden layers	3
	Learning rate	0.001
	Loop	10000
Adaboost	*N* estimators	[10,30,50,100]
	Learning rate	[0.1,0.01]
XGboost	Learning rate	[0.1,0.01]
	N estimators	[10,100]
	Gamma	[1e-2, 1e-3]
KNN	Weights	[uniform, distance]
	N neighbors	[1, 2, 3, 4, 5, 6]
	Distance	[1, 2, 3, 4, 5, 6]
DT	Criterion	[gini, entropy]
	Max depth	[10, 20]
	Min samples split	[2, 4]
	Min samples leaf	[3, 4]
CNN	Dropout	0.5
	Conv1D	10*6
	Activation	Relu
MLP	Learning rate	0.1
	Max iteration	[5,10,15]
	Early stopping	True
RF	Max depth	[1,3,5]
	N estimators	[10,20,30]

TABLE 4
ACC of Classifiers on Different Unimodal of NC/MCI/AD Classification Tasks (MEAN±STD)

Algorithm	SVM	Adaboost	KNN	RF	DT	MLP	XGboost	NNs	CNN
MRI	48.01±6.70	50.24±4.70	44.30±7.27	50.48±5.86	49.76±4.26	49.43±6.69	52.74±5.64	**53.25 ±8.80**	49.44±2.52
PET	51.49±4.60	**51.73±4.57**	42.27±5.16	50.24±5.20	50.48±5.86	49.62±6.51	43.79±6.65	50.05±3.08	47.12±3.28
SNP	55.50±3.16	68.12±6.94	48.99±8.30	50.49±5.82	49.24±4.00	**68.39±4.37**	54.15±7.00	55.16±6.86	50.48±2.30

TABLE 5
AUC of Classifiers on Different Unimodal of NC/MCI/AD Classification Tasks (MEAN±STD)

Algorithm	SVM	Adaboost	KNN	RF	DT	MLP	XGboost	NNs	CNN
MRI	70.30±6.27	74.17±7.08	64.12±6.92	70.20±5.18	53.55±4.71	67.42±7.02	68.97±4.60	**75.71±9.31**	63.85±2.36
PET	**69.62±3.51**	69.27±3.85	53.63±3.76	62.25±7.15	61.19±2.83	68.28±5.19	61.61±5.73	54.36±6.88	60.63±3.08
SNP	**84.09±5.02**	82.06±5.96	62.98±7.99	67.59±6.73	58.55±-4.85	83.85±4.20	73.53±6.35	63.84±11.20	58.88±2.27

all the parameters in steps to obtain the best ones. Finally, we presented the Accuracy (ACC), Area Under Curve (AUC), average and standard deviation of each experiment. For the calculation of AUC, we chose the python *roc_auc_socre* function to calculate macro-AUC, the index of each label, and unweighted average.

4.1 Experimental Setup

We used a tenfold cross-validation strategy to evaluate all the approaches and the final results were obtained by averaging the ten repetition results. In order to verify our conjecture, we conducted a detailed and complete experimental verification on the ADNI-1 dataset. The parameter settings of the algorithm we selected are shown in Table 3. We chose ACC and AUC as our evaluation matrix. ACC is the evaluation standard used by most machine learning algorithms or deep learning algorithms, and is an important indicator to measure the quality of an algorithm. AUC compute area under the receiver operating characteristic curve from prediction scores.

4.2 Classification Results Using Unimodal and Multimodal Data

Tables 4 and 5 show the performance of 9 traditional algorithms on the three classification tasks (AD/MCI/NC) of MRI, PET and SNP, respectively. It can be clearly seen that on the MRI classification task, the NNs algorithm achieved the best performance, with ACC of 53.25% and AUC of 75.71%; on the PET classification task, the Adaboost algorithm achieved the best performance with ACC of 51.73%, AUC is 69.27%; on the SNP classification task, the MLP algorithm achieved the best performance, with ACC of 68.39% and AUC of 83.85%. As shown in Figs. 6 and 7, in order to compare the performance of traditional algorithm classification on unimodal and multi-modal more intuitively, we averaged the performance of 9 traditional machine learning algorithms and obtained their averages ACC and AUC on MRI, PET and SNP respectively. Then, the average performance of the 9 classifiers is used as the baseline to compare with the MML algorithm. Figs. 6 and 7 also reveal that each modality has a different degree of influence on the decision-

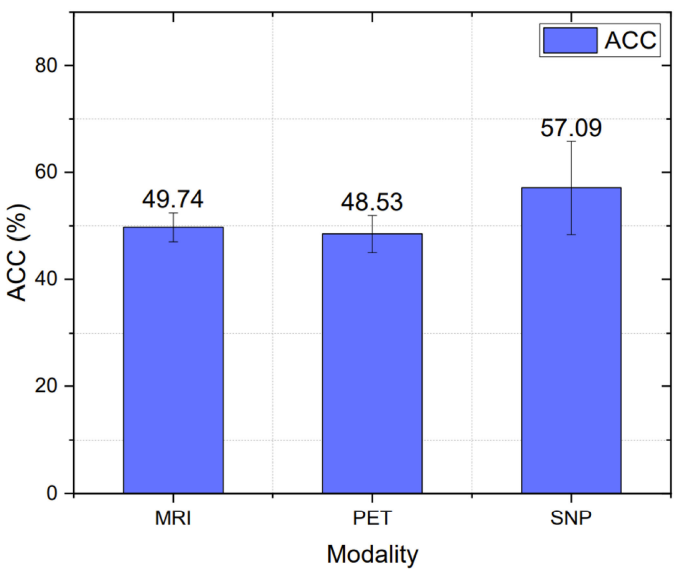


Fig. 6. Average ACC of nine traditional machine learning algorithms in unimodal.

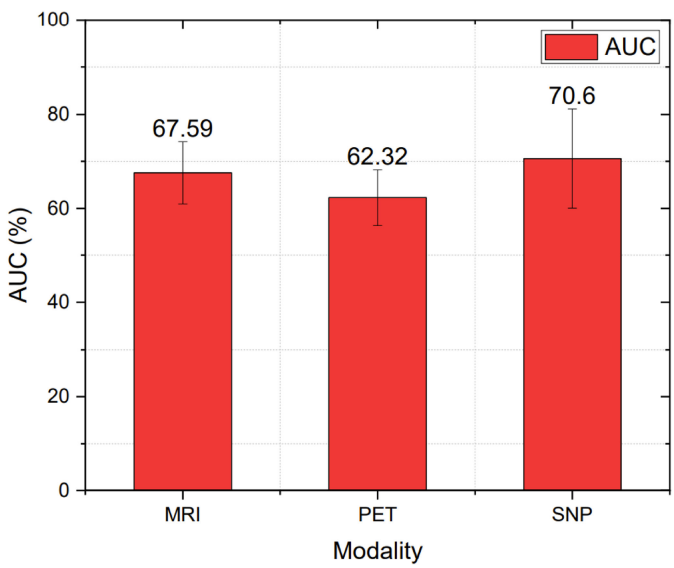
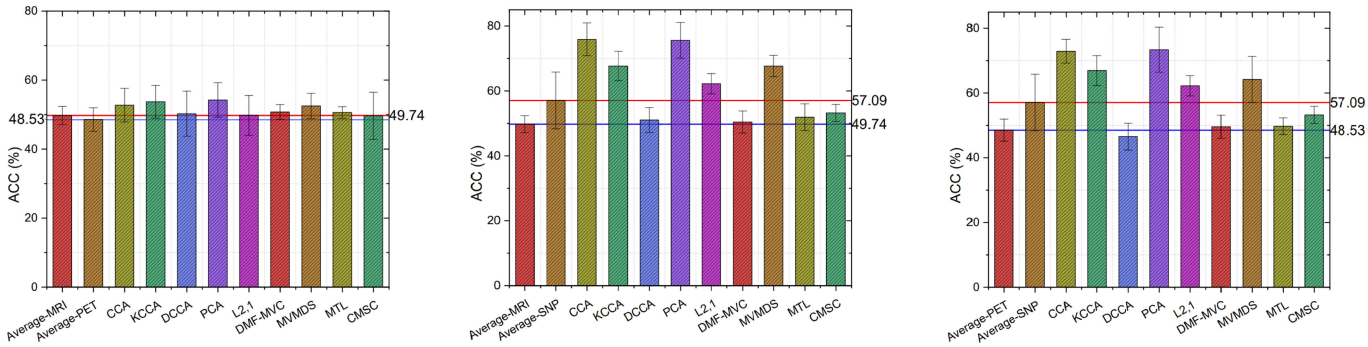
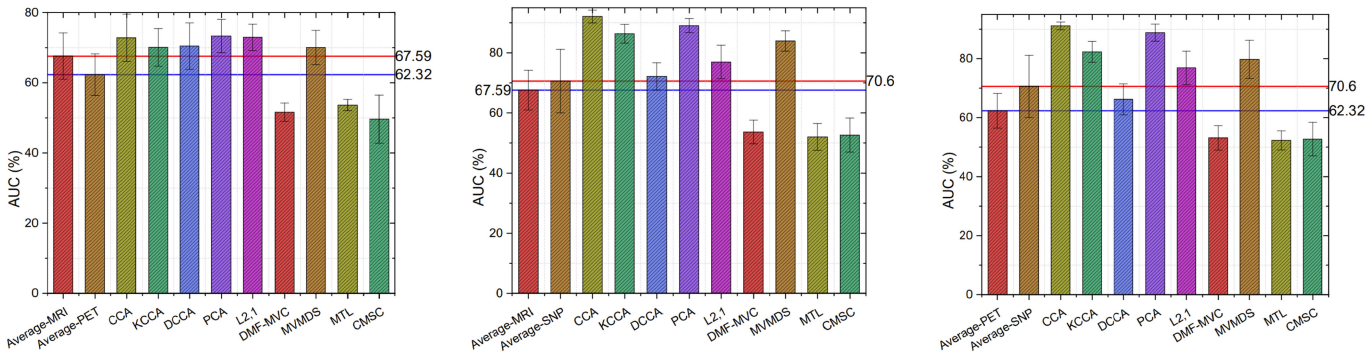


Fig. 7. Average AUC of nine traditional machine learning algorithms in unimodal.



(a) MRI vs. PET vs. MRI & PET (b) MRI vs. SNP vs. MRI & SNP (c) PET vs. SNP vs. PET & SNP

Fig. 8. Classification ACC achieved by different methods using the different modality combinations. And the error bar denotes the standard deviation of the results



(a) MRI vs. PET vs. MRI & PET (b) MRI vs. SNP vs. MRI & SNP (c) PET vs. SNP vs. PET & SNP

Fig. 9. Classification AUC achieved by different methods using the different modality combinations. And the error bar denotes the standard deviation of the results

making results. In other words, there are differences in their contribution to the classification results.

Figs. 8a, 8b, and 8c shows the ACC of traditional machine learning algorithms on unimodal and the ACC of MML algorithms on multi-modal. Figs. 9a, 9b, and 9c provides the AUC of traditional machine learning algorithms on unimodal and the AUC of MML algorithms on multimodal. From Figs. 8 and 9, we have the following observations. First, from Fig. 8a, compared with MRI-average and PET-average, we can find that among the selected nine MML algorithms, the ACC of CCA KCCA, PCA and MVMDS algorithms are significantly higher than MRI-average and the ACC of CCA, KCCA, PCA, DMF-MVS, MVMDS, and MTL is significantly higher than PET-average. From Fig. 8b, compared with MRI-average and SNP-average, CCA, KCCA, PCA, L2,1, MVMDS and CMSC algorithms have higher ACC than MRI-average, and CCA, KCCA, PCA, L2,1 and MVMDS algorithms have higher ACC than SNP-average. From Fig. 8c, compared with PET-average and SNP-average, it is apparent that CCA, KCCA, PCA, L2,1, MVMDS and CMSC algorithms have higher accuracy than PET-average, and CCA, KCCA, PCA, L2,1 and MVMDS algorithms have higher accuracy than SNP-average. After the analysis in Figs. 8a, 8b, and 8c, it can be concluded that among the 9 compared MML algorithms, at least 5 MML algorithms are better than MRI-average, PET-average and SNP-average under different modal combinations. From the experimental results of AUC as shown in Figs. 9a-9c, We got similar results as in the ACC experiments. At the same time, Figs. 8 and 9 demonstrate the feasibility and effectiveness of MML in the early diagnosis of Alzheimer's disease.

Figs. 10 and 11 compare the classification ACC and AUC performance are achieved by different algorithms using different modal combinations. CCA is a linear combination of different modalities, while KCCA and DCCA are a nonlinear combination different modalities. An inspection of experimental results of the performance of CCA, KCCA

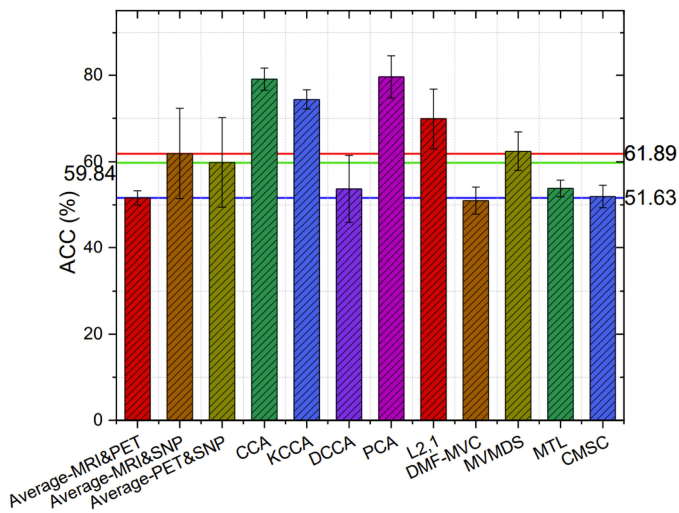


Fig. 10. Classification ACC achieved by different methods using the different modality combinations. Where the error bar denotes the standard deviation of the results.

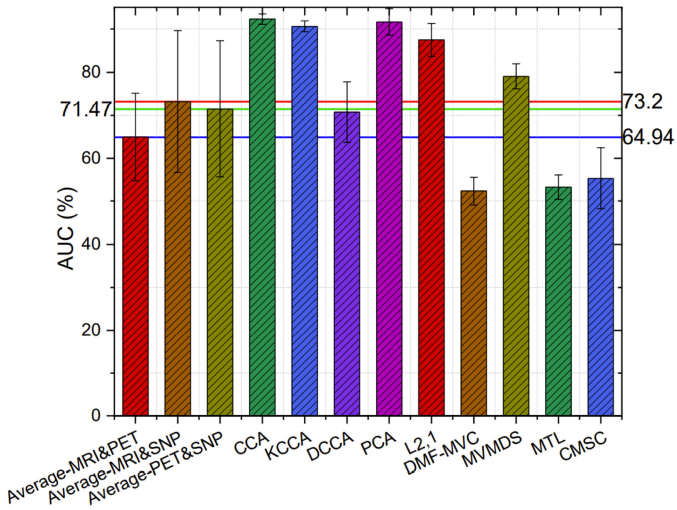


Fig. 11. Classification AUC achieved by different methods using the different modality combinations. Where the error bar denotes the standard deviation of the results.

and DCCA reveals that the generalization ability of the CCA algorithm is better than that of KCCA and DCCA. This shows to a certain extent that the linear combination is better than the non-linear combination for the early diagnosis AD. Furthermore, as a typical sparse learning algorithm, the $L_{2,1}$-norm regularization algorithm solves the problem of high-dimensional data through sparse learning of multimodal data of MRI, PET and SNP, and its classification performance ranks among the nine MML algorithms. MVMDS and CMSC are two representation learning methods, which have poor performance on the classification of AD. DMF-MVC is a matrix factorization method, and the model performance is not good. MTL trains a classifier on each modal and then uses shared parameters or shared features to diagnose early AD. The experimental results show that MTL is helpful in providing complementary information to help other tasks classify.

In summary, in the early AD diagnosis, the fusion of multimodal data helps to improve the generalization performance of the classifier. This is mainly because different modalities can provide complementary information about different modalities for AD disease diagnosis, which helps the model to diagnose AD more comprehensively. in addition, not all MML algorithms help in AD diagnosis, such as DCCA and DMF-MVC. these MML algorithms perform very poorly in early AD diagnosis, even worse than traditional machine learning algorithms.

4.3 Classification Results Using Different Combination of Modalities

In all the samples that we collected, we have four different combinations of the modalities (i.e., MRI & PET, MRI & SNP, PET & SNP, MRI & PET & SNP). These four combinations are divided into two groups, one containing two modalities of data (i.e., MRI & PET, MRI & SNP, PET & SNP), and the other containing three modalities of data (i.e., MRI & PET & SNP). In the experiment, we still used the three classification tasks (AD versus MCI versus NC) to carry out the experimental evaluation of the combinations of different modalities. Meanwhile, the ACC, AUC and standard deviation of the model were recorded.

In Section 4.2, we show the first combination pattern, including MRI & EPT, MRI & SNP, and SNP & PET. From Figs. 8 and 9, it can be observed that the ACC and AUC of the algorithms vary greatly depending on different combinations of modals. Specifically, the ACC and AUC of MRI & SNP and PET & SNP tend to be higher than that of MRI & PET. This indicates that the complementary information provided by SNP is more abundant than that provided by MRI and PET, which is more conducive to training a model with strong generalization ability. Further analysis shows that the combination of different modalities has different effects on the MML algorithms. For example, by comparing Figs. 8a and 8b, the classification ACC of MRI & SNP for the early AD diagnosis is higher than that of MRI & PET. One possible explanation is that MRI and PET features are extracted from the same ROI in the brain, and the extracted features are voxels and SUV. There is a relationship between voxel features and SUV features, resulting in less complementary information provided to each other than SNP.

Figs. 12 and 13 show the performance of average-MRI & PET, average-MRI & SNP and average-PET & SNP and the MML algorithms on three modalities (MEI, PET and SNP). By comparing the different combinations of the two modalities with the variety of the three modalities, we can observe that the algorithm performance of CCA, KCCA, PCA, and $L_{2,1}$ has been greatly improved, far exceeding other MML

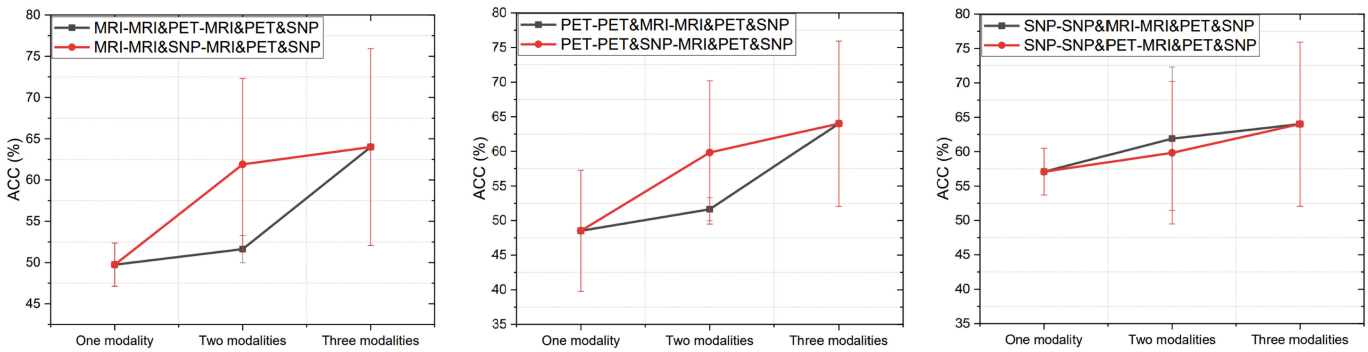


Fig. 12. The average classification ACC achieved by different algorithms using different modality combinations. The error bar denotes the standard deviation of the results

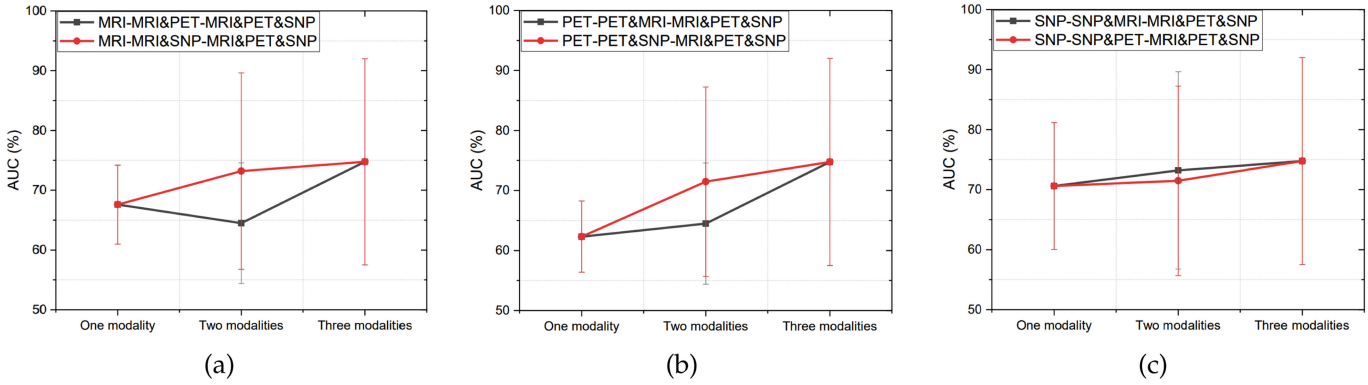


Fig. 13. The average classification AUC achieved by different algorithms using different modality combinations. The error bar denotes the standard deviation of the results

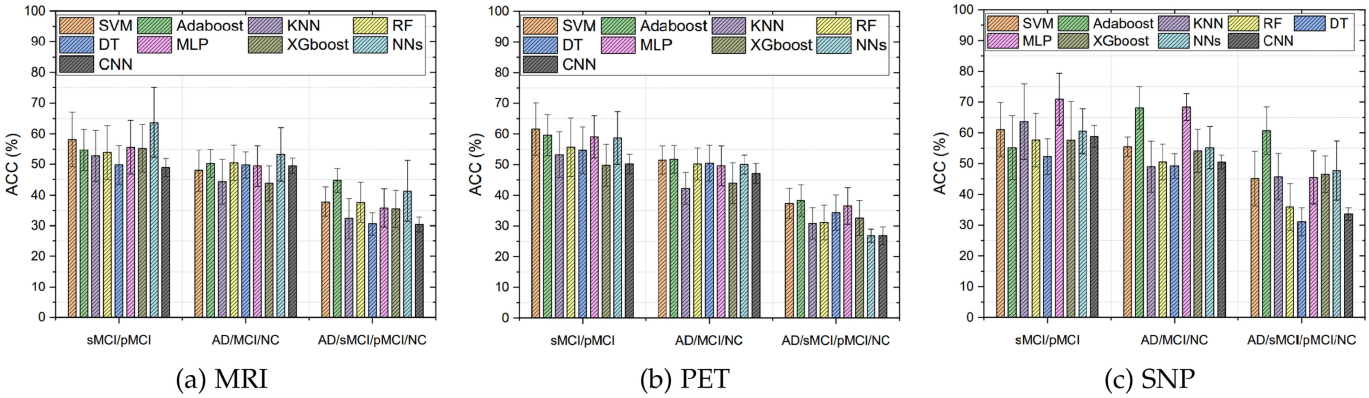


Fig. 14. The average classification ACC achieved by different algorithms on different task,when only one modality (MRI/PET/SNP) is included. The error bar denotes the standard deviation of the results.

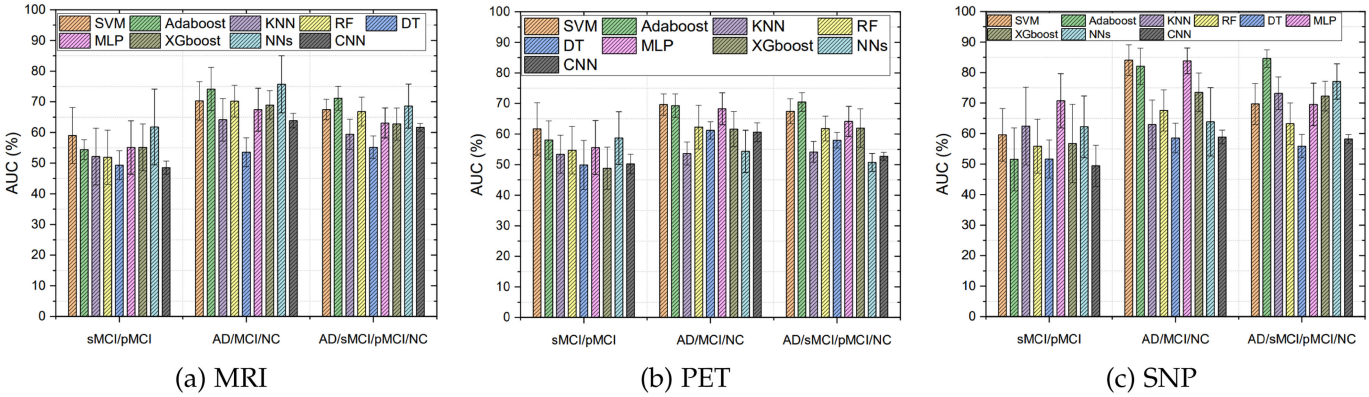
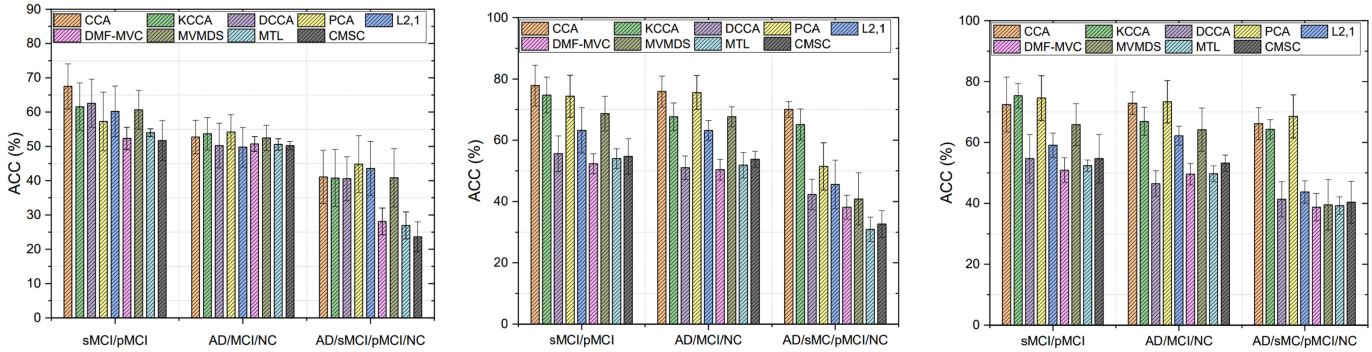


Fig. 15. The average classification AUC achieved by different algorithms on different task,when only one modality (MRI/PET/SNP) is included. The error bar denotes the standard deviation of the results.

algorithms in the early AD diagnosis. At the same time, it can also be explained that adding a modality helps to improve the generalization ability of some MML algorithms.

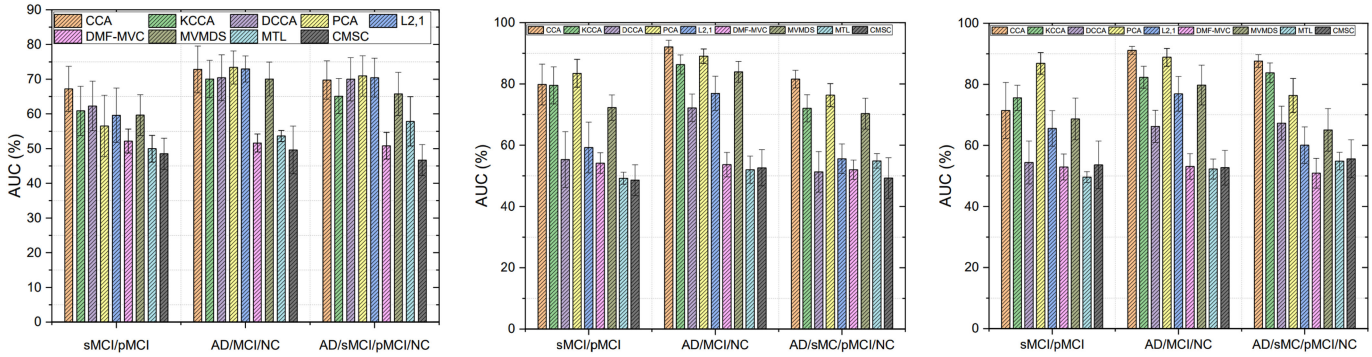
Figs. 12 and 13 show the average ACC and AUC for different modality combinations. As can be seen from Figs. 12a , 12b, and 12c, different modality combinations have a great impact on the performance of the model with great differences. As shown in Figs. 12a, 12b, 13a, and 13b, compared with MRI and PET, the model performance was greatly improved after the introduction of SNP. However, compared with SNP, the performance of the model was not significantly improved after introducing of MRI and PET. This also suggests that, in the case of limited computing resources, reducing some modalities will not greatly affect the classifier's performance, but will improve the efficiency.

In the early diagnosis of AD, MRI, PET and SNP were extracted using the feature extraction method, and it was found that SNP contributed the most to the diagnosis of AD. In addition, it provides the most complementary information for other modalities and has significantly improved performance compared with the traditional classifier trained in single-modality.



(a) MRI & PET (b) MRI & SNP (c) PET & SNP

Fig. 16. The average classification ACC achieved by different algorithms on different task,when only two modalities (MRI & PET, MRI & SNP and PET & SNP) are included. The error bar denotes the standard deviation of the results.



(a) MRI & PET (b) MRI & SNP (c) MRI & SNP

Fig. 17. The average classification AUC achieved by different algorithms on different task,when only two modalities (MRI & PET, MRI & SNP and PET & SNP) are included. The error bar denotes the standard deviation of the results.

4.4 Classification Results on Different Tasks

We conducted experiments to explore the impact of different tasks on the early diagnosis of AD and to find the best combination of modalities. First, we tested the performance of traditional machine learning algorithms for classifying unimodal data. Second, we evaluated the performance of the multimodal learning algorithm on different learning tasks.

Figs. 14 and 15 displays the ACC and AUC for different traditional machine algorithms for classification of different unimodal data, respectively. In terms of unimodal, we found that traditional machine learning algorithms perform better in SNP than MRI and PET modalities under different classification tasks. Also, we found that for early AD diagnosis of unimodal data, the SVM, Adaboost and NNs algorithms performed better than the other algorithms. Comparing Figs. 14 and 15, we can see that the AUCs for the same tasks are roughly the same as their ACCs when using different unimodal data.

Figs. 16 and 17 show the performance of different classification tasks and different multi-modal algorithms on pairwise combination of multi-modal data (MRI & PET, MRI & SNP and MRI & PET). Among the different classification tasks, two classifications (sMCI/pMCI) gives the best results, followed by three classification tasks (AD/MCI/NC), while four classification tasks give the worst results (AD/pMCI/sMCI/NC). A possible explanation for this might be that reducing the number of classification tasks reduces the complexity and thus improves the algorithm's performance. Furthermore, by comparing the different modal combinations in Fig. 18, we can see that MRI & SNP perform better than the other two combinations. One possible explanation is that the SUV features of PET are classified according to MRI brain regions. There are similarities

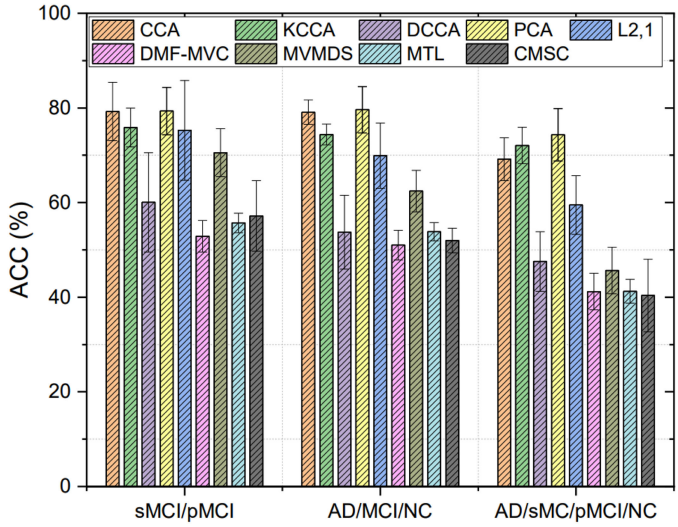


Fig. 18. The average classification ACC achieved by different algorithms on different tasks when three modalities (MRI & PET & SNP) are included. The error bar denotes the standard deviation of the results.

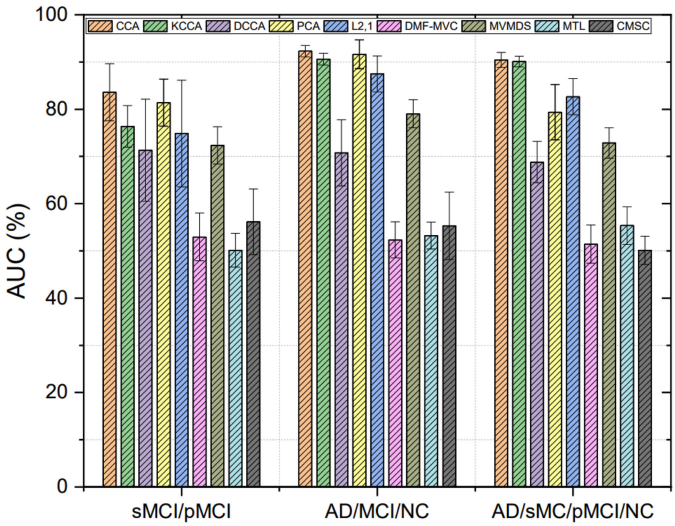


Fig. 19. The average classification AUC achieved by different algorithms on different task when three modalities (MRI & PET & SNP) are included. The error bar denotes the standard deviationof the results.

between the two features, and the complementary information provided is less than the SNP features. Last but not least, by combining the performance of all the algorithms in different classification tasks, we can see that CCA, KCCA and PCA outperformed the others.

From another point of view, in the early diagnosis of AD, there may be a linear relationship between the modalities. Figs. 18 and 19 show the performance of different multimodal learning algorithms on different classification tasks under the combination of the three modalities (MRI & PET & SNP). We found an interesting phenomenon in the combination of three modalities. Considering the performance of different classification tasks on different multi-modal learning, we found that the results of multi-modal learning algorithms are very similar.

In summary, from Figs. 14, 15, 16, 17, 18, and 19, we have verified and explained the performance of different classification tasks through experiments. Specifically, considering different classification tasks, SNP performs best in unimodal multi-classification tasks; MRI & SNP modality performs best in the multimodal multi-classfication tasks. This reveals that SNP provides more useful information than MRI and PET.

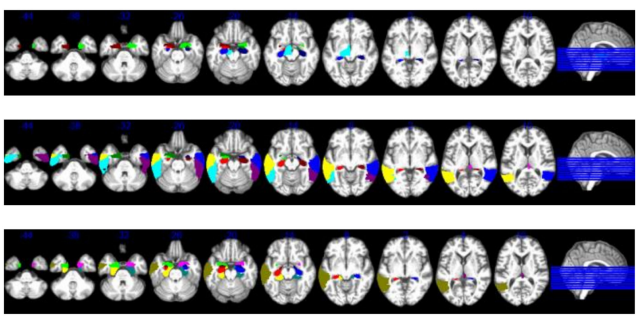


Fig. 20. Top ten selected ROIs for MRI modality in different classification tasks. From top to bottom: pMCI/sMCI, AD/MCI/NC, and AD/sMCI/pMCI/MC. Here, different colors denote different ROIs.

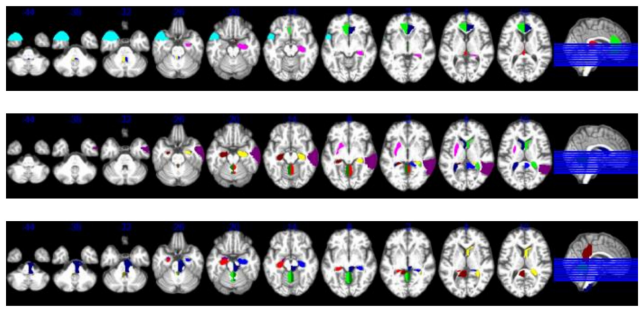


Fig. 21. Top ten selected ROIs for PET modality in different classification tasks. From top to bottom: pMCI/sMCI, AD/MCI/NC, and AD/sMCI/pMCI/MC. Here, different colors denote different ROIs.

5 Discussion and Conclusion

In this section, we first demonstrated the top ten features of MRI, PET and SNP. Hence, according to the previous research [9], [32], [33], [34], [35], [36], [37], we compared the selected features with the current ROIs and SNPs that may cause AD. Then we described the current problems in the early diagnosis of Alzheimer's disease. Finally, we introduced in detail some of our future research work and briefly introduces.

Our experiments show that the top ten ROIs are associated with MRI and PET on different classification tasks. As shown in Fig. 20, in the MRI modality, ROIs such as *Left Hippocampus*, *Right Hippocampus*, *Left Inferior Temporal Gyrus* and *Right Inferior Temporal Gyrus* has a strong correlation with the diagnosis of early AD. This finding is consistent with that of many previous studies and it shows that these ROIs have a strong connection to the diagnosis of AD. For PET modalities, as shown in Fig. 21, ROIs such as *Left Inferior Frontal Orbital Gyrus*, *Right Inferior Frontal Orbital Gyrus*, *Left Posterior Cingulate Gyrus* and *Right Posterior Cingulate Gyrus* has a strong correlation to the early diagnosis of Alzheimer's disease. Again, these identified ROIs are consistent with those reported in previous AD-related studies. For SNP modality, we choose the top 1000 most relevant features. *rs164975*, *rs1120643* are the two most relevant features in our feature extraction method and the early diagnosis of AD. This also shows that it is difficult to extract effective features for high-dimensional gene sequences, and this will also be a key research work for us in the future.

In this paper, we proposed a multi-classification framework for evaluating the early diagnosis of Alzheimer's disease. It contains a four-layer structure, from data input to knowledge application. Under this framework, we conducted research on multimodal learning and the early diagnosis of AD and made three contributions: First of all, our experiments not only reflected the usefulness of the multi-classification framework but also effectively verified the three problems. Second, our experiments prove the effectiveness of multimodal learning in the medical field that multimodal data have more complementary information or prior knowledge than single-modal data. Therefore, this complementary information help train a machine learning model with strong generalization ability. Third, with the addition of different modal data, machine learning models are becoming more and more capable of diagnosing diseases. The three points are used as prior knowledge for multi-modal learning in the medical field.

REFERENCES

[1] "WHO reveals leading causes of death and disability worldwide," 2000–2019. [Online]. Available: https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000–2019

[2] A. Alzheimer's, "2021 Alzheimer;s disease facts and figures," *Alzheimers Dement*, vol. 17, pp. 327–406, 2021.

[3] T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 423–443, Feb. 2018.

[4] X. Zhou, X. Xu, W. Liang, Z. Zeng, and Z. Yan, "Deep-learning-enhanced multitarget detection for end–edge–cloud surveillance in smart IoT," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12588–12596, Aug. 2021.

[5] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 912–921, May/Jun. 2020.

[6] X. Zhou, W. Liang, I. Kevin, K. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous big data recommendations," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 1, pp. 171–178, Feb. 2020.

[7] Y. Y. et al., "Reservoir hosts prediction for COVID-19 by hybrid transfer learning model," *J. Biomed. Inform.*, vol. 117, May 2021, Art. no. 103736.

[8] P. Yang, J. Qi, Y. Yang, S. Zhang, and G. Bi, "Rolling interventions for controlling COVID-19 outbreaks in the UK to reduce healthcare demand," in *Proc. Int. Conf. Knowl. Discov. Data Mining*, 2020, pp. 1839–1844.

[9] 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.

[10] M. Liu, D. Zhang, and D. Shen, "Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1463–1474, Jun. 2016.

[11] M. W. Weiner et al., "Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials," *Alzheimer's Dement.*, vol. 13, no. 4, pp. e1–e85, 2017.

[12] Q. Wang, M. Sun, L. Zhan, P. Thompson, S. Ji, and J. Zhou, "Multi-modality disease modeling via collective deep matrix factorization," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2017, pp. 1155–1164.

[13] Q. Wang, L. Zhan, P. Thompson, and J. Zhou, "Multimodal Learning with Incomplete Modalities by Knowledge Distillation," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2020, pp. 1828–1838.

[14] S. Liu, C. Yadav, C. Fernandez-Granda, and N. Razavian, "On the design of convolutional neural networks for automatic detection of Alzheimer;s disease," in *Mach. Learn. Health Workshop*, 2020, pp. 184–201.

[15] A. Retico, P. Bosco, P. Cerello, E. Fiorina, A. Chincarini, and M. E. Fantacci, "Predictive models based on support vector machines: Whole-brain versus regional analysis of structural MRI in the Alzheimer's disease," *J. Neuroimaging*, vol. 25, no. 4, pp. 552–563, 2015.

[16] F. Liu, L. Zhou, C. Shen, and J. Yin, "Multiple kernel learning in the primal for multimodal Alzheimer;s disease classification," *IEEE J. Biomed. Health Inform*, vol. 18, no. 3, pp. 984–990, May 2013.

[17] C. R. Jack Jr et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Reson. Imag.: An Official J. Int. Soc. Magn. Reson. Med.*, vol. 27, no. 4, pp. 685–691, 2008.

[18] K. Kwak et al., "Multi-modality sparse representation for Alzheimer;s disease classification," *J. Alzheimer's Dis.*, vol. 65, no. 3, pp. 807–817, 2018.

[19] P. S. Pillai and T.-Y. Leong, "Knowledge-driven generative subspaces for modeling multi-view dependencies in medical data," 2018, *arXiv:1812.00509*.

[20] B. Pu, K. Li, S. Li, and N. Zhu, "Automatic fetal ultrasound standard plane recognition based on deep learning and iiot," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7771–7780, Nov. 2021.

[21] C. Zhang, Y. Cui, Z. Han, J. T. Zhou, H. Fu, and Q. Hu, "Deep partial multi-view learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 5, pp. 2402–2415, May 2020.

[22] J. Peng, X. Zhu, Y. Wang, L. An, and D. Shen, "Structured sparsity regularized multiple kernel learning for Alzheimer;s disease diagnosis," *Pattern Recognit.*, vol. 88, pp. 370–382, 2019.

[23] X. Zhou, W. Liang, K. I. K. Wang, and S. Shimizu, "Multi-modality behavioral influence analysis for personalized recommendations in health social media environment," *IEEE Trans. Computat. Social Syst.*, vol. 6, no. 5, pp. 888–897, Oct. 2019.

[24] B. Pu, N. Zhu, K. Li, and S. Li, "Fetal cardiac cycle detection in multi-resource echocardiograms using hybrid classification framework," *Future Gener. Comput. Syst.*, vol. 115, pp. 825–836, 2021.

[25] X. Zhou, W. Liang, I. Kevin, K. Wang, R. Huang, and Q. Jin, "Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly big data," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 1, pp. 246–257, Jan. 2018.

[26] V. M. Patel, H. Van Nguyen, and R. Vidal, "Latent space sparse and low-rank subspace clustering," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 691–701, Jun. 2015.

[27] K. Huang, D. Zheng, J. Sun, Y. Hotta, K. Fujimoto, and S. Naoi, "Sparse learning for support vector classification," *Pattern Recognit. Lett.*, vol. 31, no. 13, pp. 1944–1951, 2010.

[28] J. Liu, S. Ji, and J. Ye, "Multi-task feature learning via efficient l2, 1-norm minimization," 2012, *arXiv:1205.2631*.

[29] S. Bai, X. Bai, L. J. Latecki, and Q. Tian, "Multidimensional scaling on multiple input distance matrices," in *Proc. AAAI Conf. Artif. Intell.*, 2017, pp. 1281–1287.

[30] A. Kumar, P. Rai, and H. Daume, "Co-regularized multi-view spectral clustering," in *Proc. Adv. Neural Inf. Process. Syst.*, 2011, pp. 1413–1421.

[31] H. Zhao, Z. Ding, and Y. Fu, "Multi-view clustering via deep matrix factorization," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2921–2927.

[32] D. Salas-Gonzalez et al., "Feature selection using factor analysis for Alzheimer's diagnosis using PET images," *Med. Phys.*, vol. 37, no. 11, pp. 6084–6095, 2010.

[33] X. Zhu, H.-I. Suk, S.-W. Lee, and D. Shen, "Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 607–618, Mar. 2015.

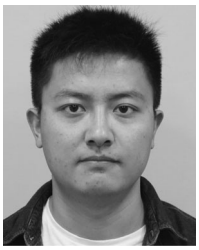
[34] M. Liu, D. Zhang, S. Chen, and H. Xue, "Joint binary classifier learning for ECOC-based multi-class classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2335–2341, Nov. 2015.

[35] T. M. Nir et al., "Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging," *Neuroimage: Clin.*, vol. 3, pp. 180–195, 2013.

[36] D. R. Westfall et al., "Increased brain activation during working memory processing after pediatric mild traumatic brain injury (mTBI)," *J. Pediatr. Rehabil. Med.*, vol. 8, no. 4, pp. 297–308, 2015.

[37] T. Zhou, K.-H. Thung, M. Liu, and D. Shen, "Brain-wide genome-wide association study for Alzheimer's disease via joint projection learning and sparse regression model," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 165–175, Jan. 2019.

Fengtao Nan received the BSc degree in computer science from Kunming University Science and Technology, China, in 2016, and the MSc degree in system analysis and integration from the University of Yunnan University, in 2019. He is currently working toward the computer science and computing degree with the School of Information, Yunnan University. His research interests include semi-supervised learning and Multimodal machine learning.

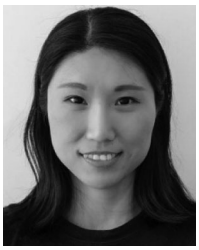


Shunbao Li received the BSc degree in automation from Yanshan University, China, in 2017, and the MSc degree in software systems and internet technology from the University of Sheffield, in 2019. Currently, he is working toward the PhD degree in computer science with the University of Sheffield. His research interests include Multimodality data fusion and Deep learning.

Jiayu Wang received the BSc degree in public affairs management major with electronic information engineering minor from Harbin Engineering University, China, in 2018. He is currently working toward the postgraduate degree majoring in artificial intelligence and machine learning with Yunnan University. His research interests include Multi-task learning and Multimodal machine learning.

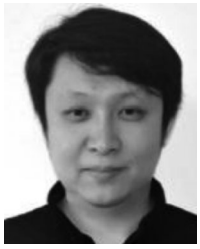
Yahui Tang received the BS degree in computer science and technology from the Kunming University of Science and Technology, Kunming, China, in 2016. She is currently working toward the PhD degree in computer science with Yunnan University. Her main research interests include process mining, business process, and software process.

Menghui Zhou received the bachelor's degree from Sun Yat-sen University, in 2017, and is currently working toward the graduate degree with Yunnan University. His research interests include machine learning, data mining and their applications to biomedical informatics.



Jun Qi received the BSc and MSc degrees in computer science and technology from Changzhou University, Changzhou, China, in 2010 and 2013, respectively, and the PhD degree in computer science with the Department of Computing Sciences from Liverpool John Moores University, U.K., in 2019. She is currently a lecturer with the Department of Computer Science and Software Engineering, Xi'an JiaoTong-Liverpool University, Suzhou, China. Her research focused on sensor-based data analysis for patients and healthy adults, using machine learning and expert systems. Special emphases of her work have been "lifelogging" physical activity patterns, and the analysis of time-series data from wearable sensors. Her research interests include predictive models for clinical applications, and investigating the effects of activity on Alzheimer's disease, Parkinson's disease, Stroke and other neurodegenerative conditions. She has published more than 30 papers in the past five years including 4 Q1 journal papers and 1 ESI highly cited paper.

Zhong Zhao is working with the Department of Neurology, the First People's Hospital of Yunnan Province, as an attending physician. He is a member of the Standing Committee of the Cognitive Disorders Branch of the Chinese Geriatric Society. His research interests include dementia and cognitive impairment, neuropsychology, neuroimage.



Yun Yang received the BSc degree (Hons.) in information technology and telecommunication from Lancaster University, Lancaster, U.K., in 2004, the MSc degree in advanced computing from Bristol University, Bristol, U.K., in 2005, and the MPhil degree in informatics, and the PhD degree in computer science from The University of Manchester, Manchester, U.K., in 2006 and 2011, respectively. He was a research fellow with the University of Surrey, Surrey, U.K., from 2012 to 2013. He is currently a full professor of machine learning with the National Pilot School of Software, Yunnan University, Kunming, China, the director of the Key Laboratory of Data Science and Intelligent Computing, Yunnan Education Department, and the director of the Kunming Key Laboratory of Data Science and Intelligent Computing. His current research interests include machine learning, data mining, pattern recognition, and temporal data process and analysis. He serves as an associate editor for the Journal of Yunnan University (Natural Sciences Edition).

Po Yang (Senior Member, IEEE) received the BSc degree (Hons.) in computer science from Wuhan University, China, in 2004, the MSc degree in computer science from the University of Bristol, in 2006, and the PhD degree in electronic engineering from the University of Staffordshire, in 2010. He is currently a senior lecturer in computer science with the University of Sheffield. Since 2006, he has been generating more than 150 international journal and conference papers in the fields of pervasive healthcare, image processing, parallel computing, and RFID-related Internet of Things (IoT) applications. His research interests include the Internet of Things, pervasive healthcare, image processing, and parallel computing. He serves as an associate editor for *IEEE Journal of Translational Engineering in Health and Medicine* and *IEEE Access*.