Clinical Neurophysiology 131 (2020) 2429–2439

Contents lists available at ScienceDirect

Clinical Neurophysiology

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

# Predicting the progression of mild cognitive impairment to Alzheimer's disease by longitudinal magnetic resonance imaging-based dictionary learning

Yanyan Lin, Kexin Huang, Hanxiao Xu, Zhengzheng Qiao, Suping Cai, Yubo Wang, Liyu Huang *, for the Alzheimer's Disease Neuroimaging Initiative [1]

*School of Life Science and Technology, Xidian University, Xi'an, Shaanxi 710071, PR China*

ARTICLE INFO

*Article history:*
Accepted 2 July 2020
Available online 14 August 2020

*Keywords:*
Alzheimer's disease
Mild cognitive impairment
Longitudinal prediction
Dictionary learning

HIGHLIGHTS

- Developed an accurate prediction system for the progression of mild cognitive impairment to AD.
- Used longitudinal structural MRI data without the segmentation of regions of interest.
- Applied dictionary learning to exploit the nuances between mild cognitive impairment patients.

ABSTRACT

*Objective:* Efficient prediction of the progression of mild cognitive impairment (MCI) to Alzheimer's disease (AD) is important for the early intervention and management of AD. The aim of our study was to develop a longitudinal structural magnetic resonance imaging-based prediction system for MCI progression.
*Methods:* A total of 164 MCI patients with longitudinal data were collected from the Alzheimer's Disease Neuroimaging Initiative (ADNI). After preprocessing, a discriminative dictionary learning framework was applied to differentiate MCI patches, avoiding the segmentation of regions of interest. Then, the proportion of patches classified as more severe atrophy patches in a patient was calculated as his or her feature to be input into a simple support vector machine. Finally, a new subject was predicted with fourfold cross-validation (CV), and the area under the receiver operating characteristic curve (AUC) was determined.
*Results:* The average accuracy and AUC values after fourfold CV were 0.973 and 0.984, respectively. The effects of the data from one or two time points were also investigated.
*Conclusion:* The proposed prediction system achieves desirable and reliable performance in predicting progression for MCI patients. Additionally, the prediction of MCI progression with longitudinal data was more effective and accurate.
*Significance:* The developed scheme is expected to advance the clinical research and treatment of MCI patients.


## 1. Introduction

Alzheimer's disease (AD), an irreversible neurodegenerative disorder characterized by a chronic, progressive degeneration of cognitive functions, is the most common cause of dementia in older populations, affecting millions of people worldwide(Brookmeyer et al., 2007; Barnes and Yaffe, 2011). Mild cognitive impairment (MCI) is the prodromal, clinically detectable phase of the trajectory toward dementia and AD. It has been found that the rates of MCI progressing to AD range from 10 to 30% annually and 20 to 66%

* Corresponding author.
*E-mail address:* huangly@mail.xidian.edu.cn (L. Huang).
[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

https://doi.org/10.1016/j.clinph.2020.07.016
1388-2457/

over 3 to 4 years(Luis et al., 2003; Petersen, 2004). Thus, it is vital to accurately predict the possibility of progression to AD in MCI patients for the early intervention and management of AD.

Due to the importance of MCI in the early diagnosis of AD, numerous studies have been published that have examined the prediction of MCI progression using noninvasive neuroimaging techniques(Duchesne et al., 2009; Misra et al., 2009; Leung et al., 2010; Stonnington et al., 2010; Huang et al., 2020). To facilitate the prediction of the risk of MCI progressing to AD, MCI patients have been further classified into those who progress to AD (pMCI) and those who remain stable (sMCI) after a few years. Among the various neuroimaging techniques, structural magnetic resonance imaging (MRI) is a powerful technique used to visualize brain structures and detect anatomic abnormalities, lesions and damage and provides measures of the inevitable atrophy related to AD (Frisoni et al., 2010; Vemuri and Jack, 2010; Arbabshirani et al., 2017). Since accelerated tissue loss is an indicator of neurodegenerative diseases, a series of structural MRI scans are often analyzed to evaluate disease progression to avoid cross-sectional measurements that are insensitive to early pathological changes (Zhang and Shen, 2012; Yau et al., 2015; Chincarini et al., 2016; Mubeen et al., 2017).

Some longitudinal studies have focused on the atrophy of well-known regions of interest (ROIs). Devanand et al. used logistic regression to predict MCI patients at a 3-year follow-up using hippocampus volume, entorhinal cortex volume and the combination of both, and accuracies of 78.9%, 79.6% and 86.7%, respectively, were achieved(Devanand et al., 2007). Mubeen et al. used features based on the hippocampus and corpus callosum of MCI patients with 6-month longitudinal data and built a longitudinal random forest model with an accuracy of 80.2% (Mubeen et al., 2017). Moscoso et al. used hippocampus volume, entorhinal volume and a combination of both from 5-year follow-up MCI patients to predict pMCI and sMCI classes with AUCs of 80%, 80% and 84%, respectively (Moscoso et al., 2019). These findings indicate that using an ROI alone, such as the hippocampus, to predict the progression of pMCI to AD will result in prediction errors because the ROI method may not be sensitive enough to detect small and more diffuse changes in the MRI structure. The segmentation of ROIs is also observer- and algorithm-dependent and relies on a priori assumptions regarding the target structures between pMCI and sMCI(Chetelat and Baron, 2003), which may result in inconsistent results(Kaye et al., 1997; Jack et al., 2000; 2004). Determining the structural features of the entire image directly is an effective way to address the issue of preselected ROIs (Arbabshirani et al., 2017; Rathore et al., 2017).

Dictionary learning is a topic in signal processing aimed at discovering a framework for the sparse representation or approximation of signals (Tosic and Frossard, 2011; Wong et al., 2011). Dictionary learning-based techniques have been successfully applied to several tasks such as image restoration (Dong et al., 2011), image denoising(Knaus and Zwicker, 2013) and image classification (Nguyen et al., 2012; Vu and Monga, 2017). By using the label information available in the training set, supervised dictionary learning methods exploit the discriminating information for classification and deliver good classification performance in natural scene classification(Wang and Kong, 2014) and face recognition(Yang et al., 2014) tasks. Recently, dictionary learning schemes have also been utilized in medical applications and can be used to directly extract features from a sparse matrix without the segmentation of ROIs(Vu et al., 2016; Diamant et al., 2017; Li et al., 2017).

In this paper, we developed an objective and automated prediction system using patch-based dictionary learning on a longitudinal structural MRI to automatically obtain the subtle gray matter (GM) density differences between patients with pMCI and sMCI. The inevitable atrophy caused by neurodegeneration can be measured by structural MRI, which can also be utilized to compute the maps of GM density. Since GM density reduction is a proxy for atrophy in AD (Frisoni, 2002; Tondelli et al., 2012), and GM atrophy progression is different in patients with pMCI and sMCI (Chételat et al., 2005), GM density has the potential to be used for differentiating between patients with pMCI and sMCI(Rathore et al., 2017). Since the atrophy patterns between pMCI and sMCI overlap, and the whole brain atrophy rate of pMCI patients is higher than that of sMCI patients one or two years later(Jack et al., 2004, 2005), we assume that the regions with more severe brain atrophy in pMCI patients account for a larger proportion of the whole brain. As follow-up scanning is an effective method for the identification of AD progression, longitudinal structural MRI was employed in our prediction system to improve the accuracy and efficiency of diagnosis and reduce unnecessary variability due to the inherent noise associated with each individual measurement. In addition, the system performance depends on the accurate segmentation of ROIs in most conventional classification or prediction systems, and our prediction system circumvented the segmentation of ROIs by using the dictionary learning technique.

## 2. Methods and materials

We applied a novel spatiotemporal discriminative dictionary learning method to obtain the difference in GM density between patients with pMCI and sMCI. As shown in Fig. 1, there were four phases in our study: image preprocessing, spatiotemporal dictionary learning, proportion calculation and prediction. The longitudinal images were registered, and GM segmentation was performed in the first step. In the spatiotemporal dictionary learning phase, dictionary bases were learned for pMCI and sMCI categories. Then, each patch in the training set was classified as a more severe atrophy patch (SAP) or a common atrophy patch (CAP). Next, the proportion of patches classified as SAP in each patient of the training set was calculated as a feature, which was input into a simple support vector machine (SVM) to train the prediction model(Vu et al., 2016). Finally, a new subject was predicted using a system trained with cross-validation (CV), and a corresponding area under the receiver operating characteristic curve (AUC) analysis was also performed.

### 2.1. Data acquisition

The data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership led by principal investigator Michael W. Weiner, MD. The primary goal of the ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Detailed MRI protocols are reported on the ADNI protocol website: (http://adni.loni.usc.edu/methods/documents/mri-protocols/).

### 2.2. Subjects

In this study, we obtained 3 T-MRI data from ADNI patients with MCI. After excluding patients with incomplete demographics or cognitive information, inability to match longitudinal images and poor image quality, 164 patients remained for which baseline and follow-up examinations were all available. According to their Clinical Dementia Rating (CDR) scores during the follow-up period,

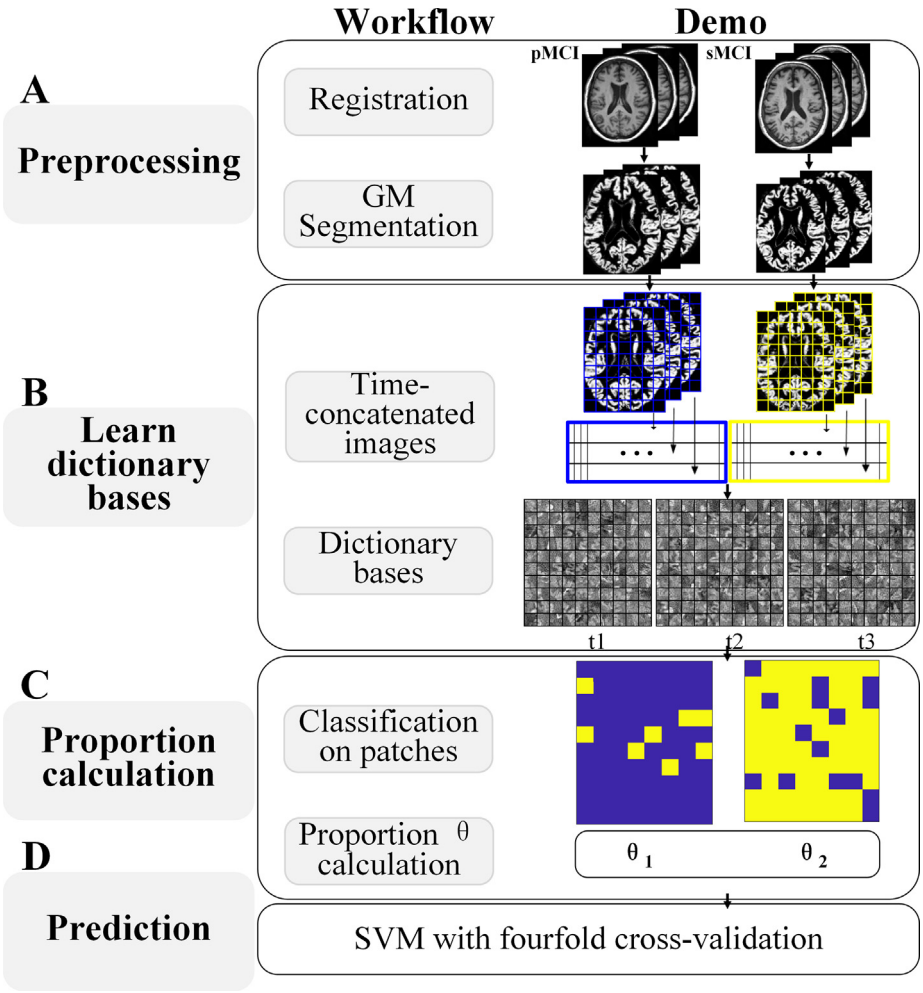


**Fig. 1.** Schematic for the prediction of patients who progress to Alzheimer's disease (pMCI) and those who remain stable (sMCI) after a few years based on spatiotemporal dictionary learning without ROI segmentation. GM: gray matter; θ: the proportion of patches classified as more severe atrophy patches in a patient; SVM: support vector machine.

the MCI patients were divided into pMCI (as having changes in CDR from 0.5 to 1) and sMCI (as CDR score maintaining at 0.5) cohorts (40 pMCI and 124 sMCI); the demographics and clinical characteristics of these patients are shown in Table 1. For each patient, we retrospectively selected serial structural MRI scans at three time points. In the case of pMCI, the last image when they confirmed AD diagnosis and two images proceed the AD diagnose were used. For sMCI, the last three images were selected. Hence, the diagnoses for each of the images were MCI-MCI-AD and MCI-MCI-MCI for pMCI patients and sMCI patients, respectively. Whole brain atrophy over 1 to 2 years is associated with the progression of MCI (Jack et al., 2005). Therefore, each image for a patient in this study was taken at an interval of 12 months. The ADNI image-IDs corresponding to time points in the pMCI and sMCI cohorts are shown in Supplementary Tables 1 and 2.

**Table 1**
Characteristics of MCI patients in the pMCI and sMCI Cohorts (baseline).

| Characteristics | pMCI | sMCI | P |
|---|---|---|---|
| **Demographic information** | | | |
| Mean (SD), years | | | |
| Age | 71.64 (7.56) | 70.81 (7.52) | 0.549 |
| Gender (M/F) | 21/19 | 69/55 | – |
| Education level | 16.10 (3.05) | 16.47 (2.69) | 0.468 |
| **Neuropsychological scales** | | | |
| Mean (SD) | | | |
| ADAS11 score | 12.76 (3.62) | 8.14 (3.39) | 0.000 |
| ADAS13 score | 20.40 (5.31) | 12.83 (5.30) | 0.000 |
| CDR score | 0.5 | 0.5 | – |
| FAQ score | 6.25 (5.11) | 1.76 (2.96) | 0.000 |
| GDS score | 1.83 (1.36) | 1.69 (1.43) | 0.610 |
| MMSE score | 26.75 (5.11) | 28.53 (1.41) | 0.000 |
| NPI-Q score | 2.775 (2.83) | 1.71 (2.63) | 0.031 |

pMCI: MCI patients who progress to AD; sMCI: MCI patients who remain stable after a few years; SD: standard deviation; ADAS: Alzheimer's Disease Assessment Scale (with 11 and 13 questionnaires, respectively); CDR: Clinical Dementia Rating; FAQ: Functional Activities Questionnaire; GDS: Geriatric Depression Scale; MMSE: Mini-Mental State Examination; NPI-Q: Neuropsychiatric Inventory Questionnaire.

### 2.3. Image preprocessing

MRI scans from 3-T scanners using the ADNI-specified 3-dimensional T1-weighted magnetization-prepared rapid gradient-echo (MP-RAGE) sequence were used in this study. The structural images at the three time points for all patients had identical resolution and the same number of slices; then, all the patients' anterior commissures were coregistered to the central point of space, and the skull was extracted. After registering all image data to the standard Montreal Neurologic Institute (MNI) space, the same voxels across the longitudinal images were properly registered to one another. Brain GM density maps were segmented using the "New Segment + DARTEL" module of the Data

Processing & Analysis of Brain Imaging (DPABI) toolbox(Yan et al., 2016) running on MATLAB R2013b (MathWorks Inc., Natick, MA, USA). The resulting GM density maps were 121 × 145 × 121 voxels.

### 2.4. Construction of spatiotemporal patches

The longitudinal structural MRI for MCI patients provided the dynamic changes in GM atrophy. Since the amount of longitudinal data was large, it was difficult to learn a dictionary directly over the entire dataset. Thus, we adopted a patch-based dictionary learning approach from the spatiotemporal volumes (Qian et al., 2016). First, a slice with an axial coordinate of 10 in the standard Montreal Neurologic Institute (MNI) space (as shown in Fig. 2(a)) was selected from each longitudinal structural MRI. Let the size of an image patch over time be $l \times l \times t$, where $l$ is the spatial size of the patch, and $t$ represents the time point. Then, the patch was converted to a column vector denoted as $x \in R^m$, where $m = l^2 t$, as shown in Fig. 2(b). As a result, the labeled longitudinal images were divided into labeled 3D patches without overlapping, and a matrix $X = [x_1, \cdots, x_N] \in R^{m \times N}$ (from two categories) was obtained, where $N$ is the size of the training set.

### 2.5. Dictionary learning and proportion computation

Studies have shown that the GM density of pMCI and sMCI is different in some regions and similar in other areas (Chételat et al., 2005; Jack et al., 2005; Karas et al., 2008). To capture the essential distinctions between these two categories and improve prediction performance, the dictionary learning algorithm was adopted. The overall learning process proposed in this work has two phases. In phase one, 2-D images at three time points were provided to DL-COPAR, and a dictionary and its corresponding coefficient matrix were learned. With dictionary learning, we can assign each patch a value representing its closeness to either the pMCI or sMCI class. In phase two, SVM was employed for prediction.

The goal of dictionary learning is to find a dictionary $D \in R^{m \times K}$ and its corresponding coefficient matrix $A = [a_1, \cdots, a_N] \in R^{K \times N}$ from the training data so that each training data can be sparsely represented by the columns of the dictionary $D$. In our study, there were two classes, i.e., pMCI and sMCI. $X_1 \in R^{m \times N_1}$ and $X_2 \in R^{m \times N_2}$ represent the data from the pMCI and sMCI classes in the training set, respectively. Then, through dictionary learning, two class-specific dictionaries ($D_1$ and $D_2$) were learned that contained the information that can best discriminate between the pMCI and sMCI patients. Moreover, the common characteristics shared by both pMCI and sMCI were also learned and represented in the dictionary $D_3$. Then, the overall dictionary can be denoted as $D = [D_1, D_2, D_3] \in R^{m \times K}$, where $K = K_1 + K_2 + K_3$. To find the dictionary $D$ and its corresponding coefficient matrix $A$, the following objective function was solved:

$$f \equiv \sum_{c=1}^{2} \left\{ \| X_c - DA_c \|_F^2 + \sum_{\substack{j=1 \\ j \neq c}}^{2} \| A_j \|_F^2 + \| X_c - D_c A_c - D_3 A_3 \|_F^2 + \lambda \phi(A_c) \right\} + \eta \sum_{c=1}^{3} \sum_{\substack{j=1 \\ j \neq c}}^{3} \| D_c^T D_j \|_F^2$$

where $\phi(A_c) = \sum_{i=1}^{N_c} \| a_i^c \|_1$ for $A_c = [a_1^c, \cdots, a_i^c, \cdots, a_{N_c}^c]$, $\lambda$ and $\eta$ are constants. The subscript $F$ denotes the Frobenius norm.

The objective function was solved using the DL-COPAR method ((Kong and Wang, 2012, Zuo et al., 2015). Once the dictionary phase was completed, pMCI- and sMCI-specific dictionaries and a shared dictionary were obtained.

After the training phase of dictionary learning was completed, a dictionary $D$ and its corresponding coefficient matrix $A$ were obtained. Then, for each image patch in the training data, we can compute the reconstruction error $e_c$ as follows: $e_c = \min_{\tilde{a}_c} \| x - \tilde{D}_c \tilde{a}_c \|_2^2 + \lambda \| \tilde{a}_c \|_1 (c = 1, 2)$

where $c \in \{1, 2\}$ represents pMCI and sMCI, respectively. $\lambda$ is a constant and $\tilde{D}_c = [D_c, D_3] \in R^{m \times (K_c + K_3)}$, and $\tilde{a}_c \in R^{(K_c + K_3)}$ derives from $\tilde{A}_c = [A_c + A_3] \in R^{(K_c + K_3) \times N}$. The final identity of $x$ is $\hat{c} = \text{argmin}_c e_c$(Kong and Wang, 2012). Following the above procedure, all patches of a patient can be assigned to either SAP or CAP. For each patient, we computed the proportion value $\theta$ of patches classified as SAP,

$\theta = \frac{\text{the number of SAP}}{\text{the number of tatal patches in the patient}} \in (0, 1)$(Vu et al., 2016).

The proportion value $\theta$ is the feature of the patient to be put into the SVM to train a prediction model.

### 2.6. Prediction

For an unseen patient with scans of three time points, after preprocessing the images, we constructed a set of patches, converted them to column vectors, and then determined the identity of all these vectors of this patient by calculating the reconstruction error $e_c$ according to the learned dictionary $D$ and its corresponding coefficient matrix $A$. Next, the proportion $\theta$ of vectors classified to SAP was computed with the method described above as his or her feature to be put into the trained SVM predictor. Finally, the SVM predictor was used to predict the category of this patient.

## 3. Results

We chose a total of 164 MCI patients from the ADNI dataset, consisting of 40 pMCI patients and 124 sMCI patients. A longitudinal analysis with three time points was registered to the MNI space, and the GM density images were computed (Fig. 3). The normalized images (size: 145 × 121 pixels) were cropped to 116 × 90 pixels for computational efficiency. The parameters employed in our prediction system were empirical: the size of the specific dictionaries of pMCI and sMCI was 350, and the size of the shared dictionary was fixed at 5. The size of the patches was fixed at 12 × 12 pixels. Since the overall sample size is relatively small, a fourfold cross-validation scheme was employed in this work to ensure that sufficient data were retained in the testing dataset. Specifically, the patients in the pMCI group were randomly divided into four portions with 10 patients in each portion. Then, for each validation run, image data from the three portions were selected for training the model, and one portion was retained for testing. At each validation run, we randomly selected 30 sMCI patients to match the number of pMCI patients so that the number of samples per group was balanced in the training set. Importantly, the testing data of each fold had no overlap with the training data. The final accuracy was a fourfold average (and the SD was also calculated). Additionally, the corresponding AUC was calculated to validate the overall robustness of our prediction system.

### 3.1. Dictionary learning

In the dictionary learning phase, class-specific dictionaries for pMCI and sMCI were obtained separately to describe the class-specific characteristics, and a shared dictionary was obtained to

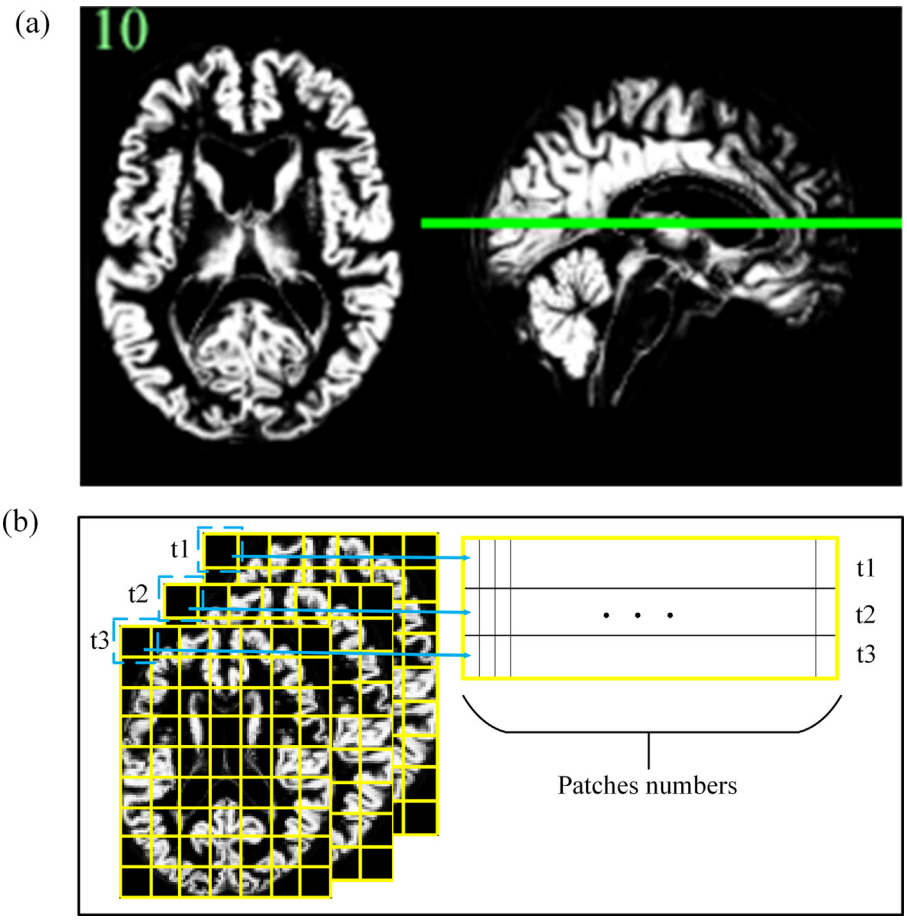


**Fig. 2.** Construction of spatiotemporal patches using selected slices. (a) A demonstration of the axial slice selected (z = 10) in the standard Montreal Neurological Institute (MNI) coordinates; (b) Construction of spatiotemporal patches for one patient. t: time point.

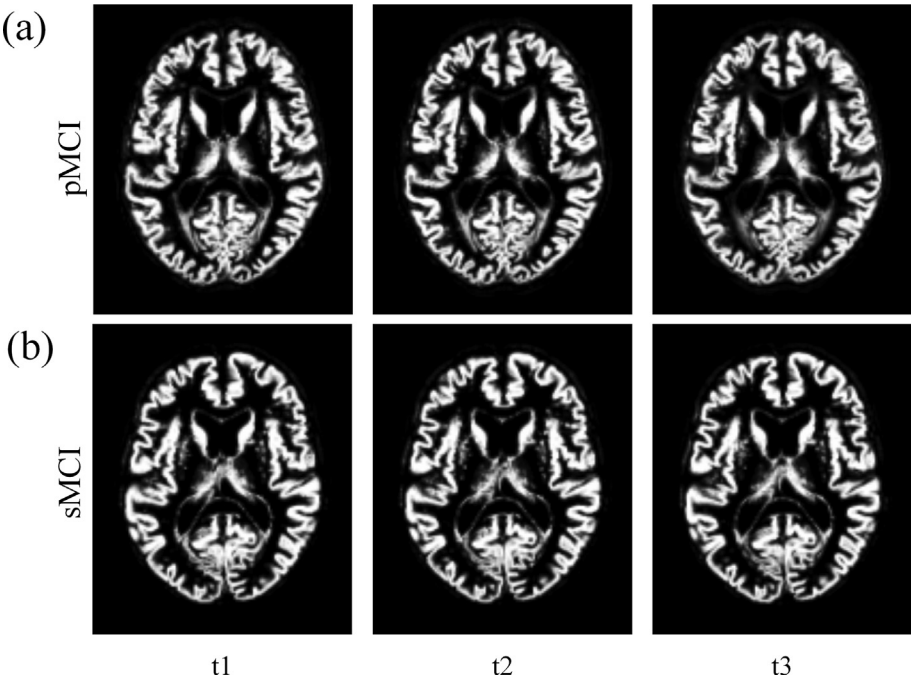


**Fig. 3.** Preprocessing performance. Preprocessing performance of structural MRI. Registered longitudinal GM images of patients (a) who progress to Alzheimer's disease (pMCI) and (b) those who remain stable (sMCI) after a few years at three time points.

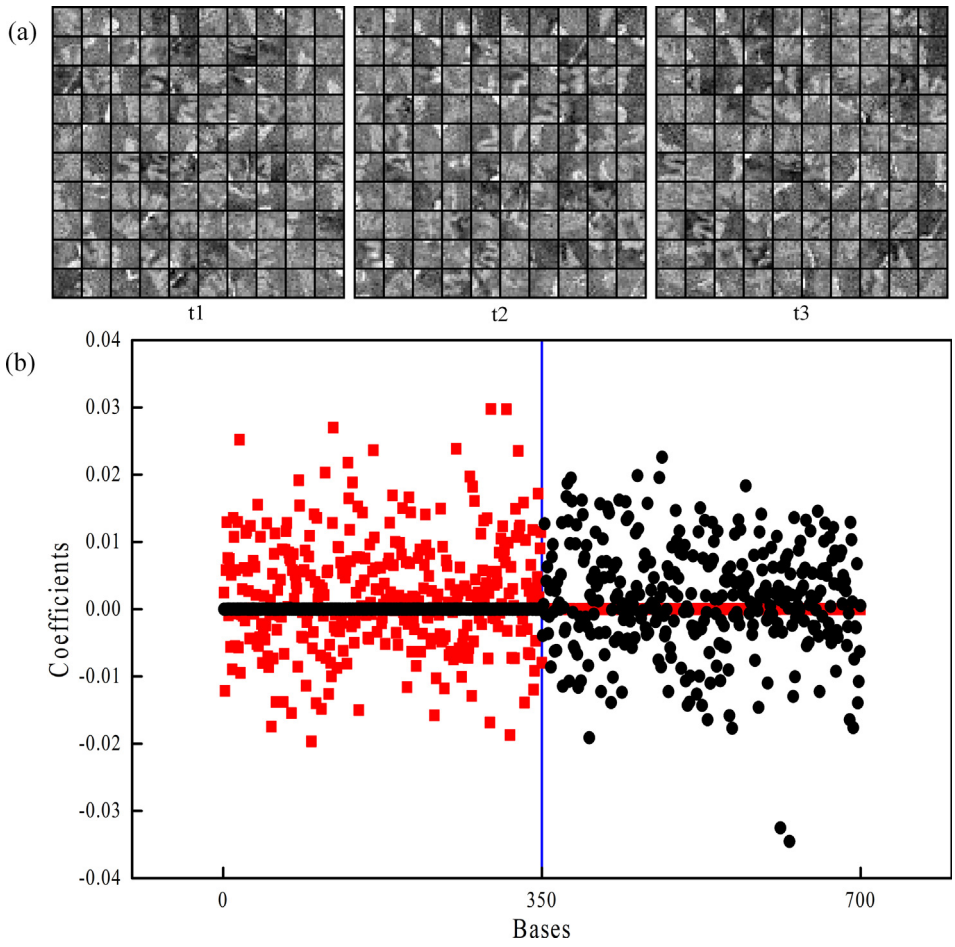


**Fig. 4.** The results during the dictionary learning phase. (a) The specific dictionaries of patients who progress to Alzheimer's disease (pMCI) at three time points (an example). (b) The coefficients corresponding to the specific dictionaries for pMCI and patients who remain stable (sMCI) after a few years, which indicated the dimensions of coefficients with coefficients on the left of blue line corresponding to the bases of dictionary related to pMCI and those on the right corresponding to the bases of dictionary related to sMCI (an example).

describe the common features of pMCI and sMCI via the DL-COPAR method. An example of the pMCI-specific dictionaries trained at three time points is shown in Fig. 4(a). With the DL-COPAR technique, the coefficients corresponding to the shared dictionary were excluded, and the average coefficients corresponding to the columns in specific dictionaries for pMCI and sMCI are shown in Fig. 4(b) to visually view the discriminating ability of learned columns. Note that the red squares in Fig. 4(b) correspond to pMCI, and most of the coefficients shown in red are to the left of the vertical line with values greater than those on the right, which meant that the specific dictionary for pMCI contributed more to reconstructing the pMCI patches than the specific dictionary for sMCI. Similarly, the black dots correspond to sMCI, and most of the active coefficients shown in black are located to the right of the blue line. These results agreed with our expectations and demonstrated that the dictionary we learned had a strong discriminative ability.

### 3.2. Performance of the prediction system

The performance of our prediction system in terms of the accuracy and AUC is shown in Fig. 5(a). The accuracy was 0.9736±0.0048 (mean±SD), indicating that the system has a promising prediction capability. The AUC for the performance was 0.9841±0.012. Our prediction system was developed based on clinical longitudinal scans at three time points. We also investigated whether the data from one or two time points could achieve a similar performance using the same parameters. The results are shown in Fig. 5(b) and (c). Our results indicated that the longitudinal data improved the prediction performance.

We also analyzed the cognitive functions of pMCI patients and sMCI patients at baseline. We included five cognitive scales in the revised manuscript, including the Alzheimer's Disease Assessment Scale with 11 questionnaires (ADAS11), the Alzheimer's Disease Assessment Scale with 13 questionnaires (ADAS13), the Functional Activities Questionnaire (FAQ), the Mini-Mental State Examination (MMSE) and the Neuropsychiatric Inventory Questionnaire (NPI-Q). The score of the five cognition scales was available for all selected patients at baseline. The statistical results are reported in Table 1. We observed that the mean value of these five scales differed significantly between the two groups. To verify whether the abovementioned cognitive scales can distinguish pMCI and sMCI patients, we conducted the following two experiments. In the case that the distribution of all patients did not change, the scores of the five scales were entered as features into SVM. First, we verified whether the scores of the five cognitive scales alone can predict the progression of MCI patients. Subsequently, we tested whether augmenting the image features with the baseline cognitive scores can improve the prediction performance. All results are shown as the mean accuracy and mean AUC.

The accuracy and AUC of 0.8101±0.0346 (mean±SD) and 0.8599±0.0533 were obtained when only employing baseline cognitive scores as features. When image features and baseline cogni-

(a)

| | Mean | SD |
|---|---|---|
| Accuracy | 0.973 | 0.0048 |
| AUC | 0.984 | 0.012 |

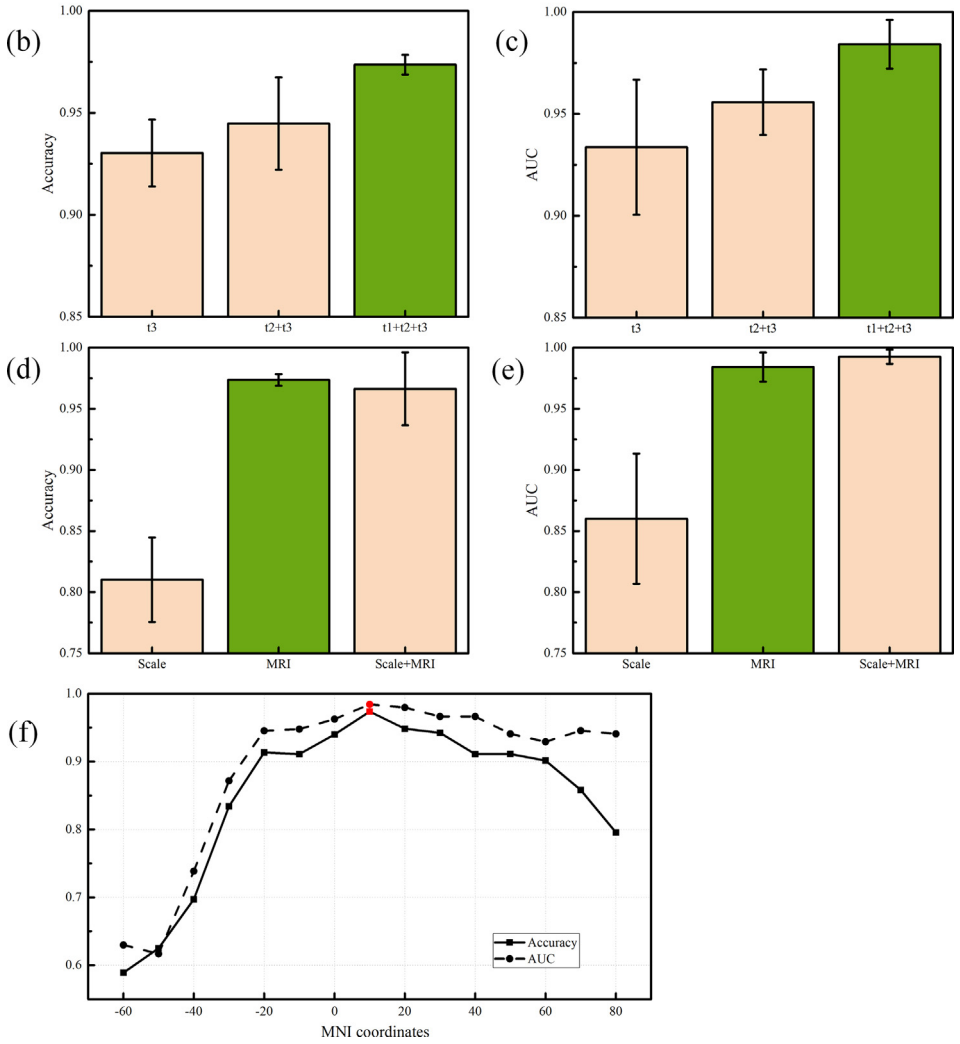


**Fig. 5.** Performance of the prediction system under different data. (a) Mean and standard deviation (SD) values for accuracy and the area under receiver operating characteristic curve (AUC) of the prediction system developed; (b) and (c) show the accuracy and AUC results using different datasets separately. t3: using data from t3 time point only. t2 + t3: using data from t2 and t3 time points. t1 + t2 + t3: using data from t1 to t3 time points totally; (d) and (e) show the accuracy and AUC results using cognitive scales or structure MRI alone and structure MRI with cognitive scales together; (f) Performance as a function of axial slices with the various standard Montreal Neurological Institute (MNI) coordinates. The slice corresponding to the red dot was selected in our prediction system.

tive scores were combined, the obtained accuracy and AUC were 0.9664±0.0299 and 0.9926±0.0058, respectively. When only image features were used to train the model, the accuracy and AUC were 0.9736±0.0048 and 0.9842±0.0119, respectively, as shown in Fig. 5(d) and (e). The results indicated that the prediction performance of using structural MRI alone was statistically higher than that of using the cognitive scales alone. Furthermore, we showed that combining baseline cognitive score and image features only marginally improved the prediction performance. Although screening neuropsychological testing is necessary to identify and monitor these high-risk subjects, to date, no fully accurate early AD cognitive markers have been identified (Chen et al., 2000). Moreover, cognitive performance depends not only on age and educational level but also on emotions and attention during the test, so it lacks extensive promotion(Chetelat and Baron, 2003).

Moreover, we estimated the prediction performance by varying the axial plane for extracting the 2-D image and repeated the same training and testing procedure as described. The results are shown in Fig. 5(f). The highest prediction performance between pMCI and sMCI was found at the axial MNI coordinate of 10.

## 4. Discussion

Because AD is a major cause of dementia and an important health and socioeconomic problem, how to intervene to delay the progression of AD and develop new remission treatments for

AD patients has become an urgent issue, which is especially important in patients with MCI. However, the effective differential technique of MCI patient progression remains a challenge for physicians due to the subtle differences between sMCI and pMCI. It has been suggested that pathological brain changes related to AD can be measured using in vivo neuroimaging techniques, and structural MRI scans are necessary for patients with suspected AD. Structural MRI measures the morphometry of the brain and thus can detect GM atrophy associated with AD at a microscopic level, and progressive cerebral atrophy associated with AD can be seen on high-resolution MRI scans, which makes MRI a powerful biomarker in differentiating the stages of AD in individuals.

Over the past few decades, prediction frameworks studying the progression of MCI patients have been used successfully in structural MRI data. In a prediction framework, accurate feature extraction is often more important than the selection of a prediction algorithm. The ROI-based method is a common approach used for feature extraction and can obtain discriminative AD-related information based on prior knowledge. However, the segmentation of ROIs is an essential and critical step in many traditional classification systems, and a priori knowledge of atrophy regions is not always available and comprehensive; therefore, some atrophy correlative details may be missed. In addition, the features finally obtained are not captured from the image directly but are obtained from the information included in the MRI image using statistical methods. Moreover, longitudinal changes in brain atrophy should also be considered; since AD is a progressive disease, cross-sectional studies have the limitations of interindividual variations in brain size and structure, and atrophy progression is inferred only from a single scan.

In this work, a dictionary learning technique that could capture the subtle difference between pMCI and sMCI directly from the sparse coefficient matrix to avoid ROI segmentation was utilized. The dictionary learning algorithm adopted in this paper has the ability to exclude features shared between pMCI and sMCI to exploit the nuances between these two phenotypes.

The effects of the specific dictionary size and the shared dictionary size on the performance of the dictionary learning algorithm were also investigated. As shown in Fig. 6(a), most dictionary sizes achieved good results, with accuracies and AUCs higher than 0.9. The best performance was observed with approximately 350 specific dictionary bases, which indicated that in our data, a larger dictionary size did not lead to a better prediction effect. Moreover, a larger dictionary size could lead to a longer runtime. This result indicates that the discriminative capability of coefficients would be weakened if the number of dictionary bases was too small, but if it was too large, too much redundant information would result in difficulties in matching identical categories. The effects of the shared dictionary sizes are shown in Fig. 6(b), which indicate that the exclusion of shared features between the two categories could significantly improve the performance of our system. Nevertheless, as the size of the shared dictionary increased, the results did not improve. The reason may be that when the size of the shared dictionary increased, DL-COPAR tended to absorb class-specific features into the shared dictionary.

From the DL-COPAR objective function, too large a training dataset can cause the learning process to converge slowly, resulting in longer calculation time. To achieve a balance between learning time and learning effect, only one 2-D slice per image was extracted to train the model. There is ample evidence showing that GM atrophy patterns are different between pMCI and sMCI patients. Additionally, GM atrophy beyond the temporal lobe is a characteristic of pMCI patients (Korf et al., 2004; Bell-McGinty et al., 2005; Chételat et al., 2005; Jack et al., 2005; Karas et al., 2008). Karas et al. found that the axial MNI coordinate of the peak value of the cluster with the largest volume of GM atrophy between pMCI and sMCI is 10 (Karas et al., 2008). We chose the same axial location to extract the 2-D image. Moreover, Fig. 5(f) shows that the highest prediction performance between pMCI and sMCI was found at the axial MNI coordinate of 10, which was consistent with previous findings. These results provided further confirmation that selecting a single 2-D image can serve to differentiate pMCI and sMCI patients.

To determine whether the GM density of the patients changed over time, we also used the "Statistical analysis" module of DPABI to perform a paired two-sample t-test at a threshold significance of $p < 0.0001$ (false discovery rate corrected for multiple comparisons) on the first and third time points of all pMCI and sMCI patients. The results of the selected slice are shown in Fig. 7(a). To analyze the results of dictionary learning in more detail, we obtained the average coefficients of columns in specific dictionaries, ranked the average coefficients corresponding to the two specific dictionaries from large to small, and found the columns in specific dictionaries that had the largest positive contribution to the average coefficients in the top twenty. The same operation was performed for the dictionary and coefficient matrix learned in each fold. The locations where the significant patches appear more frequently are shown in Fig. 7(b). The common atrophy pattern between pMCI and sMCI was eliminated, and the differential discriminative characteristics of the two categories could be effectively captured by the dictionary learning algorithm.

The performance of our prediction system was better than that of other longitudinal systems (Devanand et al., 2007; Mubeen

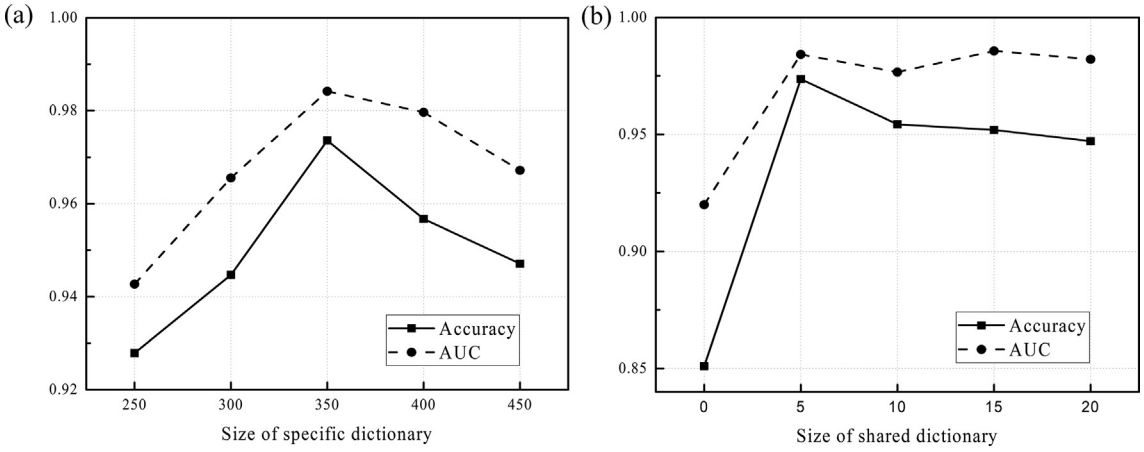


**Fig. 6.** Performance as a function of (a) the size of the specific dictionary and (b) the size of the shared dictionary.

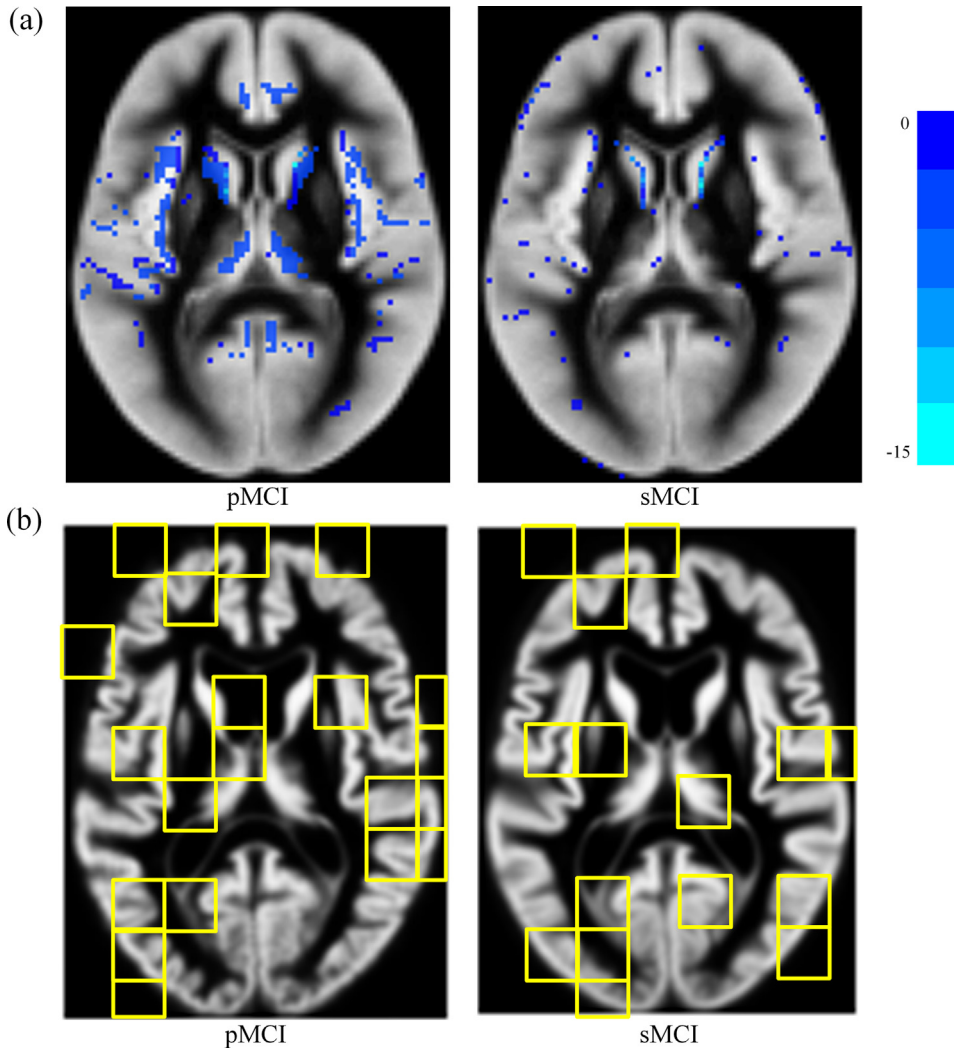


**Fig. 7.** Differences in longitudinal atrophy between pMCI (patients who progress to Alzheimer's disease) and sMCI (patients who remain stable) patients. (a) The results of paired two-sample t-test paired two-sample t-test at a threshold significance of $p < 0.0001$ (false discovery rate corrected for multiple comparisons) on the first and third time points of all pMCI and sMCI patients, respectively; (b) The locations where the patches that had the largest positive contribution to the average coefficients in the top twenty appear more frequently.

et al., 2017; Moscoso et al., 2019). The reasons may be that our system avoids ROI segmentation using a dictionary learning algorithm, and the dictionary learning algorithm used can eliminate the common atrophy pattern between pMCI and sMCI to better capture the differences between the two categories.

There are also limitations in our work. One limitation was that the dataset for pMCI was small because of our desire for completely longitudinal data, so to avoid overfitting, we chose one slice instead of the whole brain. To avoid the impact of data imbalance on the prediction performance of our system, we used the same number of images to train the specific dictionaries and the remaining images to test our system. The other limitation is that for an unseen MCI patient, structural MRI data needs to be provided at three time points with a time interval of approximately one year to obtain a more accurate prediction. In addition, we used only image data without considering the more essential effects of genes on the disease, which will be explored in a future study.

## 5. Conclusion

To conclude, we developed an objective prediction approach without segmentation to predict MCI-to-AD progression. As a promising imaging method for characterizing microstructural atrophy or differences in treatment, longitudinal structural MRI scans were applied in our system. A spatiotemporal and dictionary learning scheme for classification was utilized to capture nuances between pMCI and sMCI. The results indicated that our proposed prediction system can accurately predict pMCI and sMCI and that the prediction of MCI progression with longitudinal data was more effective and accurate. Therefore, this system will be a useful tool in the clinical diagnosis of patients with pMCI and sMCI.

## Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO

Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

## Funding

This work was supported by the National Natural Science Foundation of China [grant numbers 81671778, 81701787 and 81801789].

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinph.2020.07.016.

## References

Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls Available from. Neuroimage 2017;145:137–65. https://linkinghub.elsevier.com/retrieve/pii/S105381191600210X.

Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence Available from. Lancet Neurol 2011;10:819–28. https://linkinghub.elsevier.com/retrieve/pii/S1474442211700722.

Bell-McGinty S, Lopez OL, Meltzer CC, Scanlon JM, Whyte EM, DeKosky ST, et al. Differential cortical atrophy in subgroups of mild cognitive impairment. Arch Neurol 2005;62:1393–7.

Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease Available from. Alzheimer's Dement 2007;3:186–91. https://linkinghub.elsevier.com/retrieve/pii/S155252600700475X.

Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST, Ganguli M. Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. Neurology 2000;55:1847–53. Available from: https://doi.org/10.1212/wnl.55.12.1847

Chetelat G, Baron JC. Early diagnosis of Alzheimer's disease: Contribution of structural neuroimaging Available from. Neuroimage 2003;18:525–41. https://linkinghub.elsevier.com/retrieve/pii/S1053811902000265.

Chételat G, Landeau B, Eustache F, Mézenge F, Viader F, de la Sayette V, et al. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: A longitudinal MRI study Available from. Neuroimage 2005;27:934–46. https://linkinghub.elsevier.com/retrieve/pii/S1053811905003277.

Chincarini A, Sensi F, Rei L, Gemme G, Squarcia S, Longo R, et al. Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease Available from. Neuroimage 2016;125:834–47. https://linkinghub.elsevier.com/retrieve/pii/S1053811915009817.

Devanand DP, Pradhaban G, Liu X, Khandji A, De Santi S, Segal S, et al. Hippocampal and entorhinal atrophy in mild cognitive impairment: Prediction of Alzheimer disease. Neurology 2007;68:828–36.

Diamant I, Klang E, Amitai M, Konen E, Goldberger J, Greenspan H. Task-Driven Dictionary Learning Based on Mutual Information for Medical Image Classification. IEEE Trans Biomed Eng. 2017;64:1380–92.

Dong W, Zhang L, Shi G. Centralized sparse representation for image restoration. In: 2011 International Conference on Computer Vision. IEEE; 2011. p. 1259–66. Available from: http://ieeexplore.ieee.org/document/6126377/.

Duchesne S, Caroli A, Geroldi C, Collins DL, Frisoni GB. Relating one-year cognitive change in mild cognitive impairment to baseline MRI features Available from. Neuroimage 2009;47:1363–70. https://linkinghub.elsevier.com/retrieve/pii/S1053811909003693.

Frisoni GB. Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry Available from. J Neurol Neurosurg Psychiatry 2002;73:657–64. http://jnnp.bmj.com/cgi/doi/10.1136/jnnp.73.6.657.

Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 2010;6:67–77.

Huang K, Lin Y, Yang L, Wang Y, Cai S, Pang L, et al. A multipredictor model to predict the conversion of mild cognitive impairment to Alzheimer's disease by using a predictive nomogram Available from. Neuropsychopharmacology 2020;45:358–66. http://www.nature.com/articles/s41386-019-0551-0.

Jack CR, Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology 2000;55:484–9.

Jack CR, Shiung MM, Gunter JL, O'Brien PC, Weigand SD, Knopman DS, et al. Comparison of different MRI brain athrophy rate measures with clinical disease progression in AD. Neurology 2004;62:591–600.

Jack CR, Shiung MM, Weigand SD, O'Brien PC, Gunter JL, Boeve BF, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology 2005;65:1227–31.

Karas G, Sluimer J, Goekoop R, van der Flier W, Rombouts SARB, Vrenken H, et al. Amnestic Mild Cognitive Impairment: Structural MR Imaging Findings Predictive of Conversion to Alzheimer Disease Available from. Am J Neuroradiol 2008;29:944–9. http://www.ajnr.org/lookup/doi/10.3174/ajnr.A0949.

Kaye JA, Swihart T, Howieson D, Dame A, Moore MM, Karnos T, et al. Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. Neurology 1997;48:1297–304.

Knaus C, Zwicker M. Dual-domain image denoising. In: 2013 IEEE International Conference on Image Processing. IEEE; 2013. p. 440–4.

Kong S, Wang D. A Dictionary Learning Approach for Classification: Separating the Particularity and the Commonality. In: European Conference on Computer Vision. Berlin, Heidelberg: Springer; 2012. p. 186–99.

Korf ESC, Wahlund LO, Visser PJ, Scheltens P. Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment. Neurology 2004;63:94–100.

Leung KK, Shen K-K, Barnes J, Ridgway GR, Clarkson MJ, Fripp J, et al. In: Increasing Power to Predict Mild Cognitive Impairment Conversion to Alzheimer's Disease Using Hippocampal Atrophy Rate and Statistical Shape Models. Berlin, Heidelberg: Springer; 2010. p. 125–32.

Li Q, Wu X, Xu L, Chen K, Yao L, Li R. Multi-modal discriminative dictionary learning for Alzheimer's disease and mild cognitive impairment Available from. Comput Methods Programs Biomed. 2017;150:1–8. https://linkinghub.elsevier.com/retrieve/pii/S0169260716313244.

Luis CA, Loewenstein DA, Acevedo A, Barker WW, Duara R. Mild cognitive impairment: Directions for future research. Neurology 2003;61:438–44.

Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI Available from. Neuroimage 2009;44:1415–22. https://linkinghub.elsevier.com/retrieve/pii/S1053811908011257.

Moscoso A, Silva-Rodríguez J, Aldrey JM, Cortés J, Fernández-Ferreiro A, Gómez-Lado N, et al. Prediction of Alzheimer's disease dementia with MRI beyond the short-term: Implications for the design of predictive models. NeuroImage Clin 2019;23:101837. https://doi.org/10.1016/j.nicl.2019.101837.

Mubeen AM, Asaei A, Bachman AH, Sidtis JJ, Ardekani BA. A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment. J Neuroradiol 2017;44:381–7.

Nguyen H Van, Patel VM, Nasrabadi NM, Chellappa R. Kernel dictionary learning. ICASSP, IEEE Int Conf Acoust Speech Signal Process - Proc. 2012;2021–4.

Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004;256:183–94.

Qian X, Tan H, Zhang J, Zhao W, Chan MD, Zhou X. Stratification of pseudoprogression and true progression of glioblastoma multiform based on longitudinal diffusion tensor imaging without segmentation. Med Phys 2016;43:5889–902. https://doi.org/10.1118/1.4963812.

Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages Available from. Neuroimage 2017;155:530–48. https://linkinghub.elsevier.com/retrieve/pii/S1053811917302823.

Stonnington CM, Chu C, Klöppel S, Jack CR, Ashburner J, Frackowiak RSJ. Predicting clinical scores from magnetic resonance scans in Alzheimer's disease Available from. Neuroimage 2010 Jul;51:1405–13. https://linkinghub.elsevier.com/retrieve/pii/S1053811910003381.

Tondelli M, Wilcock GK, Nichelli P, de Jager CA, Jenkinson M, Zamboni G. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. Neurobiol Aging 2012;33:825.e25–36.

Tosic I, Frossard P. Dictionary Learning Available from. IEEE Signal Process Mag 2011;28:27–38. http://ieeexplore.ieee.org/document/5714407/.

Vemuri P, Jack CR. Role of structural MRI in Alzheimer's disease Available from. Alzheimers Res Ther 2010;2:23. http://alzres.biomedcentral.com/articles/10.1186/alzrt47.

Vu TH, Monga V. Fast low-rank shared dictionary learning for image classification. IEEE Trans Image Process 2017;26:5160–75.

Vu TH, Mousavi HS, Monga V, Rao G, Rao UKA. Histopathological Image Classification Using Discriminative Feature-Oriented Dictionary Learning Available from. IEEE Trans Med Imaging 2016;35:738–51. http://ieeexplore.ieee.org/document/7303944/.

Wang D, Kong S. A classification-oriented dictionary learning model: Explicitly learning the particularity and commonality across categories Available from. Pattern Recognit 2014;47:885–98. https://linkinghub.elsevier.com/retrieve/pii/S0031320313003245.

Wong K-W, Lin Q, Chen J. Error detection in arithmetic coding with artificial markers Available from. Comput Math with Appl 2011;62:359–66. https://linkinghub.elsevier.com/retrieve/pii/S0898122111004056.

Yan C-G, Wang X-D, Zuo X-N, Zang Y-F. DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging Available from. Neuroinformatics 2016;14:339–51. http://link.springer.com/10.1007/s12021-016-9299-4.

Yang M, Zhang L, Feng X, Zhang D. Sparse Representation Based Fisher Discrimination Dictionary Learning for Image Classification Available from. Int J Comput Vis 2014;109:209–32. http://link.springer.com/10.1007/s11263-014-0722-8.

Yau WYW, Tudorascu DL, McDade EM, Ikonomovic S, James JA, Minhas D, et al. Longitudinal assessment of neuroimaging and clinical markers in autosomal dominant Alzheimer's disease: A prospective cohort study. Lancet Neurol 2015;14:804–13.

Zhang D, Shen D. Predicting Future Clinical Changes of MCI Patients Using Longitudinal and Multimodal Biomarkers. Chen K, editor. PLoS One 2012;7:e33182. Available from: http://dx.plos.org/10.1371/journal.pone.0033182.

Zuo Z, Wang G, Shuai B, Zhao L, Yang Q. Exemplar based Deep Discriminative and Shareable Feature Learning for scene image classification. Pattern Recognit 2015;48:3004–15. https://doi.org/10.1016/j.patcog.2015.02.003.