



1-12-2017

Longitudinal Measurement and Hierarchical Classification Framework for the Prediction of Alzheimer's Disease

Meiyan Huang
Southern Medical University, China

Wei Yang
Southern Medical University, China

Qianjin Feng
Southern Medical University, China

Wufan Chen
Southern Medical University, China

Follow this and additional works at: https://uknowledge.uky.edu/neurology_facpub

Michael Weiner

University of California, San Francisco
Part of the [Diagnosis Commons](#), [Diseases Commons](#), [Longitudinal Data Analysis and Time Series Commons](#), and the [Neurology Commons](#)

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)
See next page for additional authors

Repository Citation

Huang, Meiyan; Yang, Wei; Feng, Qianjin; Chen, Wufan; Weiner, Michael; Aisen, Paul; Petersen, Ronald; Jack, Clifford R. Jr.; Jagust, William; Trojanowki, John; Toga, Arthur W.; Beckett, Laurel; Green, Robert C.; Saykin, Andrew; Morris, John; Shaw, Leslie M.; Kaye, Jeffrey; Quinn, Joseph; Silbert, Lisa; Lind, Betty; Carter, Raina; Dolen, Sara; Schneider, Lon S.; Pawluczyk, Sonia; Beccera, Mauricio; Teodoro, Liberty; Spann, Bryan; Brewer, James; Vanderswag, Helen; Fleisher, Adam; Smith, Charles D.; Jicha, Greg A.; Hardy, Peter A.; Sinha, Partha; Oates, Elizabeth; and Conrad, Gary, "Longitudinal Measurement and Hierarchical Classification Framework for the Prediction of Alzheimer's Disease" (2017). *Neurology Faculty Publications*. 18.

https://uknowledge.uky.edu/neurology_facpub/18

This Article is brought to you for free and open access by the Neurology at UKnowledge. It has been accepted for inclusion in Neurology Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Longitudinal Measurement and Hierarchical Classification Framework for the Prediction of Alzheimer's Disease

Digital Object Identifier (DOI)

<https://doi.org/10.1038/srep39880>

Notes/Citation Information

Published in *Scientific Reports*, v. 7, article no. 39880, p. 1-13.

© The Author(s) 2017

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.


Due to the large number of authors, only the first 30 and the authors affiliated with the University of Kentucky are listed in the author section above. For the complete list of authors, please download this article or visit: <https://doi.org/10.1038/srep39880>

This group of authors is collectively known as the Alzheimer's Disease Neuroimaging Initiative.

Authors

Meiyan Huang, Wei Yang, Qianjin Feng, Wufan Chen, Michael Weiner, Paul Aisen, Ronald Petersen, Clifford R. Jack Jr., William Jagust, John Trojanowki, Arthur W. Toga, Laurel Beckett, Robert C. Green, Andrew Saykin, John Morris, Leslie M. Shaw, Jeffrey Kaye, Joseph Quinn, Lisa Silbert, Betty Lind, Raina Carter, Sara Dolen, Lon S. Schneider, Sonia Pawluczyk, Mauricio Beccera, Liberty Teodoro, Bryan Spann, James Brewer, Helen Vanderswag, Adam Fleisher, Charles D. Smith, Greg A. Jicha, Peter A. Hardy, Partha Sinha, Elizabeth Oates, and Gary Conrad

SCIENTIFIC REPORTS



OPEN

Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease

Received: 03 August 2016
Accepted: 29 November 2016
Published: 12 January 2017

Meiyan Huang¹, Wei Yang¹, Qianjin Feng¹, Wufan Chen¹ & the Alzheimer's Disease Neuroimaging Initiative[#]

Accurate prediction of Alzheimer's disease (AD) is important for the early diagnosis and treatment of this condition. Mild cognitive impairment (MCI) is an early stage of AD. Therefore, patients with MCI who are at high risk of fully developing AD should be identified to accurately predict AD. However, the relationship between brain images and AD is difficult to construct because of the complex characteristics of neuroimaging data. To address this problem, we present a longitudinal measurement of MCI brain images and a hierarchical classification method for AD prediction. Longitudinal images obtained from individuals with MCI were investigated to acquire important information on the longitudinal changes, which can be used to classify MCI subjects as either MCI conversion (MCIC) or MCI non-conversion (MCINC) individuals. Moreover, a hierarchical framework was introduced to the classifier to manage high feature dimensionality issues and incorporate spatial information for improving the prediction accuracy. The proposed method was evaluated using 131 patients with MCI (70 MCIC and 61 MCINC) based on MRI scans taken at different time points. Results showed that the proposed method achieved 79.4% accuracy for the classification of MCIC versus MCINC, thereby demonstrating very promising performance for AD prediction.

Alzheimer's disease (AD) is characterized by the progressive impairment of cognitive and memory functions and is the most common form of dementia in elderly people. As the life expectancy increases, the number of AD patients increases accordingly, thereby causing a heavy socioeconomic burden^{1,2}. Mild cognitive impairment (MCI) is a prodromal stage of AD; existing studies have suggested that individuals with amnesic MCI tend to progress to probable AD at a rate of approximately 10% to 15% per year^{1,3}. Generally, patients with MCI who convert to AD after some time are called MCI converters (MCIC), whereas others who never convert to AD or even revert to a normal status are called MCI non-converters (MCINC). The classification of MCINC and MCIC is studied because of its importance in the early prediction of AD. Considering the limited period for which the symptomatic treatments are effective, patients with MCI who are at high risk of fully developing AD should be identified.

Recent studies showed that MRI can contribute significant progress to understand the neural changes related to AD and other diseases. Moreover, MRI data provide some brain structure information; this information can be used to identify the anatomical differences between populations of AD patients and normal controls (NC) and assist in the diagnosis and evaluation of MCI progression^{4,5}. Generally, most MRI-based classification methods consist of two major steps: (1) feature extraction and selection and (2) classifier learning. Basing on the type of features extracted from MRI, the MCINC/MCIC classification methods can be divided into three categories: the voxel-based approach⁶⁻⁹, the vertex-based approach^{1,10,11}, and the region of interest (ROI)-based approach¹²⁻¹⁵.

The vertex-based approach can be used to obtain information regarding the conversion from MCI to AD by using cortical thickness, sulcal depth, or cortical surface area as features. Although crucial disease progression information can be acquired through the vertex-based approach, this method depends on the accuracy of the surface registration¹. The ROI-based approach usually employs nonlinear registration to register a brain MRI image to a structurally or functionally predefined brain region template before extracting representative features from each

¹Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, Guangzhou, China. [#]A comprehensive list of consortium members appears at the end of the paper. Correspondence and requests for materials should be addressed to Q.F. (email: qianjinfeng08@gmail.com)

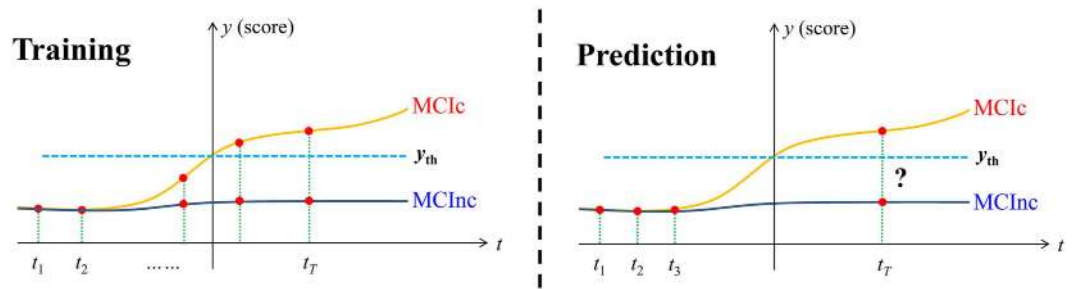


Figure 1. An example using longitudinal data to predict AD conversion.

region. Although the ROI-based approach can significantly reduce the feature dimensionality, the features extracted from ROIs are very coarse and cannot reflect small or subtle changes associated with the brain diseases¹⁶. In the voxel-based approach, the features are defined at the level of the MRI voxel, which is simple and intuitive in terms of the interpretation of the results. However, the main limitations of the voxel-based approach are the high dimensionality of feature vectors and the lack of spatial information¹⁶. On one hand, the high dimensionality of feature vectors often leads to low performance attributed to the “curse of dimensionality”²². To address this problem, feature selection is typically performed to reduce feature dimensionality and eliminate the redundant features. On the other hand, AD often affects spatially contiguous regions instead of isolated voxels. Thus, the local spatial contiguity of the selected discriminative features (voxels) should be carefully considered during feature selection or classification^{2,5,16}.

Recently, several longitudinal neuroimaging studies have collected a rich set of longitudinal data to better understand the progress of neuropsychiatric and neurodegenerative diseases or normal brain development^{3,17–22}. The predictive value of early brain developmental trajectories should be studied for later brain and cognitive development and disease progression^{18,19}. Therefore, the longitudinal changes in MRI measures may be a crucial factor in the prediction of future conversion from MCI to AD^{3,7,20,23–25}. In the group-based approaches, longitudinal data have already been used for measuring longitudinal changes of the brain; until very recently, only a few researchers started to use longitudinal data for individual-based MCInc/MCIc classification^{3,7,20,26,27}. Li *et al.*²⁰ investigated the longitudinal cortical thickness changes of 75 MCI subjects to distinguish MCIc from MCInc. Moreover, Zhang *et al.*³ proposed an AD prediction method with ROI-based features from longitudinal data. The experiments were performed on 88 MCI subjects, and the results shown that the performance of their method with longitudinal data was better than that with baseline visit data³. Despite these efforts, extracting discriminative features from longitudinal data for the early diagnosis and prediction of AD progression is still challenging and requires more research.

In the present study, a longitudinal measurement-based hierarchical classification (LMHC) method for AD prediction is proposed. Specifically, longitudinal images obtained from individuals with MCI were investigated to acquire important information on the longitudinal changes that can be used to classify MCI subjects as MCIc or MCInc individuals. From a clinical perspective, an observed trend can show the tendency of an MCI subject to become an AD patient or to remain stable. If such trends are dynamically monitored with longitudinal data, AD-related changes can be determined; an AD prediction model can be constructed with the longitudinal data. In clinical settings, when a new MRI scan of an MCI subject is available, the future medical condition of the MCI subject (to develop AD or remain stable) can be predicted with his/her previous MRI scans and the constructed prediction model (Fig. 1). Thus, richer information can be extracted from the longitudinal data to help enhance the prediction accuracy. From a feature extraction perspective, several studies also suggested that voxel-based morphometry of longitudinal data can provide useful information regarding AD progression^{25,28,29}. Thus, voxel intensities in MRI images from longitudinal data are used as features for classification. From a classifier construction perspective, instead of building a single classifier with an optimal subset of features, an ensemble learning method was used in this study to improve the generalizability and robustness of individual classifiers. Recently, a hierarchical ensemble classification method that combined multilevel classifiers through gradual integration of numerous features from both local brain regions and interbrain regions was proposed in refs 16 and 30. Unlike the abovementioned studies, we propose a hierarchical classification method that builds multiple and multilevel classifiers with supervised learning and suitable thresholds to address the issues of high feature dimensionality and sensitivity to small changes for more accurate classification of MCI. Therefore, we can evaluate the classification abilities of the image features in various brain regions and at different levels. The performance of the proposed AD prediction method was tested on 131 patients with MCI with MRI scans taken at different time points. Overall, the findings show that the proposed method with longitudinal data and the hierarchical classification framework generate promising results for AD prediction.

Materials

Data were downloaded from the Alzheimer’s disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI, PI Michael M. Weiner). ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a \$60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial MRI, PET and other biological markers are useful in clinical trials of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their

Diagnosis	Number	Age	Gender (M/F)	MMSE
MCIc	70	74.26 ± 7.55	42/28	26.46 ± 1.76
MCInc	61	75.85 ± 6.49	32/29	27.05 ± 1.78

Table 1. Demographic information of the studied subjects from the ADNI database.

Time point	First 6 m	6–12 m	12–18 m	18–24 m	24–36 m	Total
number	11	28	17	7	7	70

Table 2. Number of MCI subjects who developed to AD during different time points (6 m, 12 m, 18 m, 24 m, and 36 m represent 6, 12, 24, and 36 months, respectively).

	Baseline	6 m	12 m	18 m	24 m	36 m	48 m	Total
MCIc	70	61	65	52	52	31	8	339
MCInc	61	55	49	30	27	12	1	235
Total	131	116	114	82	79	43	9	574

Table 3. Number of MCIc and MCInc subjects at different time points (6 m, 12 m, 18 m, 24 m, 36 m, and 48 m represent 6, 12, 24, 36, and 48 months, respectively).

effectiveness, as well as lessen the time and cost of clinical trials. ADNI subjects aged 55 to 90 from over 50 sites across the US and Canada participated in the research and more detailed information is available at www.adni-info.org.

T1-weighted MRI images were used in this study. The scanning parameters for the 1.5T MRI images can be found in ref. 31. A total of 131 subjects with MCI (70 MCIc and 61 MCInc) from ADNI1 were considered in this study. Demographic information of the studied subjects is presented in Table 1. The diagnosis of the 61 MCInc subjects was MCI at all available time points (0–48 months). Moreover, the diagnosis of the 70 MCIc subjects was MCI at baseline but conversion to AD was reported after baseline within 6, 12, 24, 36, or 48 months, and without reversion to MCI or NC at any available follow-up (0–48 months). From the 70 MCIc subjects, 11 subjects were converted to AD within the first 6 months, 28 subjects were converted to AD between the 6 and 12 months follow-up, 17 subjects were converted to AD between the 12 and 18 months follow-up, 7 subjects were converted to AD between the 18 and 24 months follow-up, and the remaining 7 subjects were converted to AD between the 24 and 36 months follow-up. The number of MCI subjects who converted from MCI to AD during different time points is displayed in Table 2. The MRI scans at the baseline visit, 6, 12, 24, 36, and 48 months were used in the AD prediction when available. A longitudinal study usually covers a relatively long period of time in the field of health sciences; some individuals almost always miss their scheduled visits or date of observation. Therefore, the sequence of observation times may vary across individuals¹⁹. The details of subject number at different time points are listed in Table 3.

Methods

Preprocessing. Given that the intensity values in MRI images do not indicate a fixed meaning and widely vary within or between subjects, the MRI images were preprocessed with the following steps. First, the N3 method³² was applied to remove bias field artifacts from the images. Second, a two-step method³³ was used to normalize the intensity values. Third, all images were spatially normalized to the publicly available ICBM152 average³⁴ via FLIRT (<http://www.fmrib.ox.ac.uk/fsl/>). The images were subsequently aligned to a standard template space with an image size of $193 \times 229 \times 193$ and a voxel size of $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$. Subsequently, the non-brain tissues were removed with the skull stripping method proposed in ref. 35. The intensity values were normalized again by the following steps. The intensity values for the brain region were calculated at 0.1% and 99.9% quantiles; both values were used to linearly scale the intensity values of voxels to the range of [0, 100]. Finally, for the simplicity of the proposed method, MRI data obtained before the missing time points were used as the missing data. For instance, if MRI scans are obtained at baseline visit, 6, 24, and 36 months, then we used the MRI scans from the 6 and 36 months as data for the 12 and 48 months, respectively. Thus, the number of time points remained the same among all the subjects considered for data collection.

Basic idea of LMHC. In this study, voxel intensities in MRI images were used as features for classification. The voxel intensity is high dimensional, which consists of much more voxels ($193 \times 229 \times 193 \approx 8.53 \times 10^6$) than the subjects (that is, hundreds at most). If all voxel intensities are used as classification features, the high dimensionality features will likely degrade the classification capability of the classifier. To solve this problem, a common strategy used is to select a set of useful voxels and apply a supervised classifier on these voxels to perform classification. However, an optimal subset of discriminative features is difficult to find by using only a single global classifier given that the discriminative features from the high dimensional neuroimaging data may lie in multiple low-dimensional feature subspaces². Moreover, disease-induced structural changes may occur at some relatively large regions of the brain^{36,37}. Therefore, the spatial information found in several voxel-grouped local regions should be considered to enhance the classification accuracy.

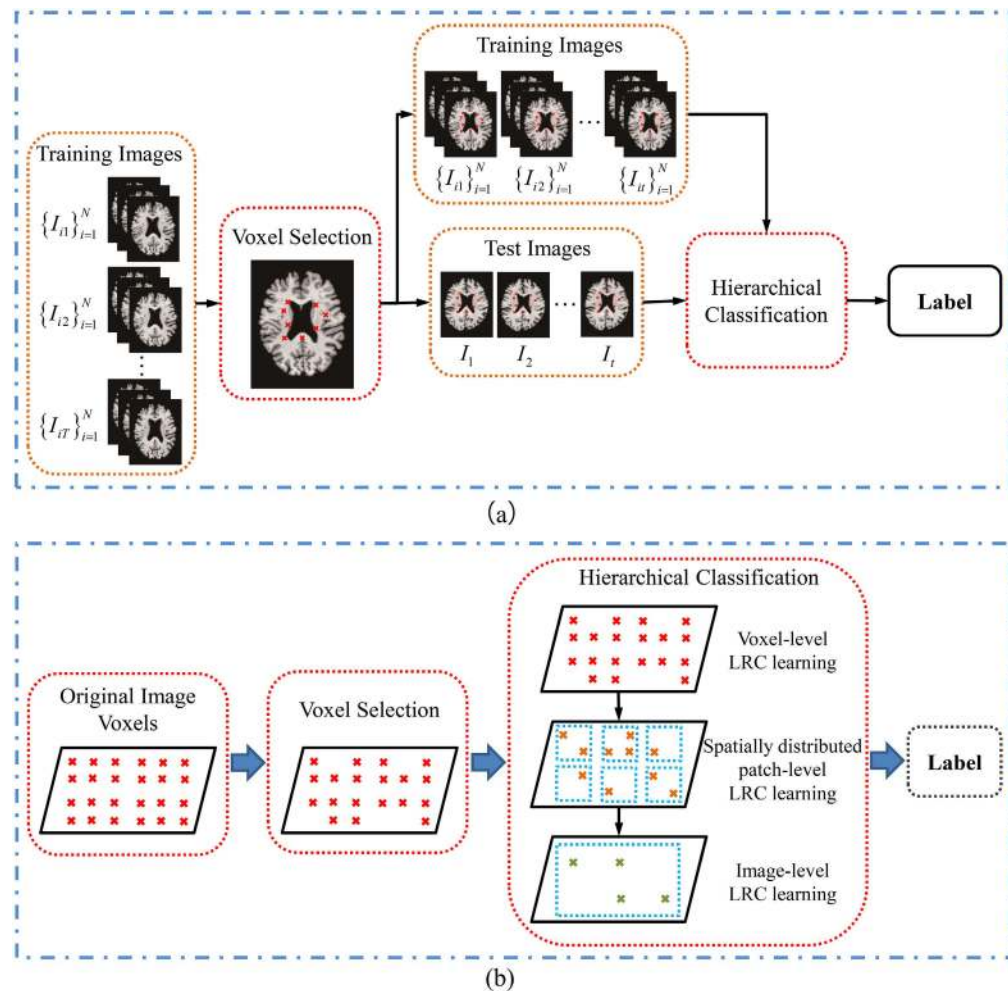


Figure 2. (a) Flowchart and (b) illustration of the proposed LMHC method.

To address the abovementioned problems, we proposed a longitudinal measurement-based hierarchical classification framework that hierarchically combined multiple individual classifiers for more accurate AD prediction. Specifically, the logical regression classifier (LRC), which is easily used and trained, was employed to design each individual classifier. To select a set of informative voxels, the longitudinal data from different time points (MRI data of MCI subjects after the conversion were also included) were used to train each individual classifier. With the selected voxel set, a hierarchical classification framework can be built on the selected voxel sites with the longitudinal data at time points that are at least 6 months ahead of the conversion. Given that each voxel or region feature defines a subspace of the whole brain feature space, each individual classifier can be trained more easily in a much smaller subspace, and thereby substantially improving in the dimensionality-to-subject ratio. The accuracy of the final classification can be further improved by replacing a single classifier with a hierarchical classification framework.

The overall schematic of the proposed classification method is shown in Fig. 2. The method consists of two fundamental steps: a voxel selection step that selects a good subset of MRI voxels for AD conversion prediction and a classification step that uses a hierarchical framework to make the final prediction. We provide details for each step in the following sections.

Selection of significant voxels. For accurate classification, the most useful voxels among all the MRI voxels were selected, whereas the noisy ones were excluded. For voxel selection, we used LRC on the longitudinal data. Given a training image set, $X = \{I_{ij}, L_i\}_{i=1, \dots, N, j=1, \dots, T}$, for the longitudinal data, the longitudinal feature of a voxel site v of the images is represented as $\{F_v\}_{v=1}^n = \{\{f_v^1, \dots, f_v^N\}\}_{v=1}^n$, where N is the subject number, T is the number of time points, $L_i \in \{0, 1\}$ is the label of the i th image, n is the voxel number, and f_v^i is a feature vector containing T elements. In this voxel selection step, the longitudinal data from different time points were included because the data from MCI to AD status can provide important conversion information on the classification of MCIC and MCInc. In addition, the longitudinal feature of each voxel site in the training images was used to learn an LRC, and the longitudinal features in the training images were fed to their corresponding classifiers to obtain a confident value for each voxel site. The confident values ranged from 0 to 1. Finally, the voxel sites with confident values higher than the threshold t_s were selected. Specifically, the cost function of LRC on the longitudinal features at voxel site v was calculated as

$$J(\mathbf{w}_v) = -\frac{1}{N} \left[\sum_{i=1}^N L_i \log \left(\frac{1}{1 + \exp(-\mathbf{w}_v^T \mathbf{f}_v^i)} \right) + (1 - L_i) \log \left(1 - \frac{1}{1 + \exp(-\mathbf{w}_v^T \mathbf{f}_v^i)} \right) \right] \quad (1)$$

Given that no closed-form technique can be used to solve for the minimum of $J(\mathbf{w}_v)$, we used the gradient descent method to iteratively optimize eq. (1). For the training image set, $X = \{I_{ij}, L_i\}_{i=1, \dots, N, j=1, \dots, T}$, and its corresponding longitudinal feature set, $\{\mathbf{F}_v\}_{v=1}^n = \{\mathbf{f}_v^1, \dots, \mathbf{f}_v^N\}_{v=1}^n$, the classification result of LRC for subject i at voxel site v was defined as

$$c_v^i = 1/(1 + \exp(-\mathbf{w}_v^T \mathbf{f}_v^i)) \quad (2)$$

Thus, the confident value at voxel site v was calculated as

$$y_v = \sum_{i=1}^N \mathbf{1}\{c_v^i > 0.5\} = L_i/N \quad (3)$$

Finally, the significant voxels were selected as $V_s = \{v | y_v > t_s\}$.

Hierarchical classification method. After selecting a set of significant voxels, a hierarchical classification framework was constructed on the selected voxel sites of the longitudinal data. In this step, only the longitudinal data at time points that are at least 6 months ahead of the conversion were used because of their importance in predicting the conversion of MCI in the clinic. In the hierarchical classification framework, a three-level classifier was built for forming decisions: voxel, patch, and image levels. For the first-level classifier, a classifier was built for each longitudinal feature at every significant voxel site. For the second- and third-level classifiers, the outputs from the lower-level classifiers were fed to the corresponding upper-level classifier (Fig. 2(b)). Specifically, for the voxel-level classification, an LRC was independently trained for each significant voxel site with the longitudinal feature as input. The output coming from the voxel-level LRC is a confident value, which was obtained as the output in the voxel selection step. We selected the confident values higher than the threshold t_h as inputs for the patch-level classifier. To incorporate the spatial information, an image with the original image size ($193 \times 229 \times 193$) was generated. In this image, the values at the selected voxel sites (that is, voxel sites with confident values higher than t_h) were set according to the corresponding confident values, whereas the rest of the values were set to 0. Subsequently, the confident values inside a patch with size w (image values equal to 0 were excluded) on the image were fed to an LRC for patch-level classification. Finally, the confident value from a patch-level classifier with a value higher than t_h was isolated and fed to the image-level classifier. The output obtained from the image-level classifier was the final decision for AD prediction. Notably, the threshold t_h at different levels shared the same value. Moreover, voxels that were not useful (i.e., $p < t_h$) for the classification were discarded, and each patch classifier covered various regions of the different brain areas.

Summary of LMHC. To elucidate the concept of LMHC, we provided a pseudo-code for LMHC, as illustrated in Algorithm 1.

Algorithm 1. LMHC

Input: Training image set $X = \{I_{ij}, L_i\}_{i=1, \dots, N, j=1, \dots, T}$ and testing image set $x = \{I_j\}_{j=1}^t$

Output: The classification label L of x .

Training stage:

1. Significant voxel selection: A LRC was constructed for each voxel in the brain.
 - (1) The longitudinal features, $\{\mathbf{F}_v\}_{v=1}^n = \{\mathbf{f}_v^1, \dots, \mathbf{f}_v^N\}_{v=1}^n$, were extracted from X with all time point data for all voxels in the brain. An LRC was learned with eq. (1) for the longitudinal features at each voxel site.
 - (2) The confident values were computed with eqs (2) and (3); the significant voxels with confident values higher than the threshold t_s were selected: $V_s = \{v | y_v > t_s\}$.
2. Hierarchical classification: A hierarchical classification framework was constructed.
 - (1) Voxel-level classification: The longitudinal features, $\{\mathbf{F}_v\}_{v=1}^n = \{\mathbf{f}_v^1, \dots, \mathbf{f}_v^N\}_{v=1}^n$, were extracted from X with different time point data from MCI subjects (only the longitudinal data at time points that are at least 6 months ahead of the conversion were used) at selected significant voxel sites. An LRC was developed for each longitudinal feature.
 - (2) Patch-level classification: The confident values obtained from the voxel-level classifiers that were higher than the threshold t_h were selected, an image was generated with the selected confident values as image values, and an LRC was learned for the selected confident values inside a patch with a size of w on the generated image.
 - (3) Image-level classification: For the selection of input, the confident values obtained from the patch-level classifiers that were higher than the threshold t_h . An LRC was learned for all the selected confident values.

Test stage:

1. The longitudinal features, $\{\mathbf{F}_v\}_{v=1}^n = \{\mathbf{f}_v^1, \dots, \mathbf{f}_v^N\}_{v=1}^n$, were extracted from x with different time point data from MCI subjects (only the longitudinal data at time points that are at least 6 months ahead of the conversion were used) at selected significant voxel sites.
 2. Hierarchical classification was performed on the extracted features, and the label L of x was obtained.
-

Results

Experimental setting. The proposed method was evaluated with two nested cross-validation loops (10-fold for each loop)⁹. Specifically, for the external 10-fold cross-validation, all subject samples were divided into 10 subsets with the same proportion of each class label. For each run, all samples within one subset were successively chosen as the testing set, whereas the remaining samples in the other nine subsets were combined and used as the training set for voxel selection and classification. The final classification results were reported as the mean results from each run. Moreover, parameter tuning was evaluated with the inner 10-fold cross-validation on the training set. In particular, the training set can be further split into a training part and a validation part for each run of the external 10-fold cross-validation. By varying the values of the different parameters, the proposed classifier was developed using the samples in the training part. The classification results were obtained during validation. The parameters with the maximum average classification accuracy during validation were selected. Notably, all longitudinal data (MRI data of MCI subjects after the conversion were also included) on a subject were used in the training step to select a set of informative voxels. However, given that only NC and MCI data were available in practice for AD prediction, the longitudinal data at time points that were at least 6 months ahead of the conversion were used to train hierarchical classifiers. In addition, the longitudinal data at time points that were at least 6 months ahead of the conversion were used in the validation and testing steps.

In the experiments, four measurement criteria were applied to evaluate the classification performance: sensitivity (SEN), specificity (SPE), accuracy (ACC), and the area under the receiver operating characteristic (ROC) curve (AUC). Specifically, the accuracy is the proportion of subjects correctly predicted among all the studied subjects. The sensitivity is the proportion of correctly predicted MCIc, whereas the specificity is the proportion of correctly predicted MCInc.

Parameter optimization. The parameter settings of the proposed method were carefully considered to achieve optimum performance in our experiments. A summary of the parameter settings in the proposed method is presented in Table 4.

In the step for significant voxel selection, the threshold t_s was selected from the group {0.5, 0.55, 0.6, 0.65} to determine a set of significant voxels. Generally, the number of significant voxels is reduced when the threshold t_s is increased. However, few significant voxels were left when $t_s > 0.7$ was considered in our experiment. Therefore, the maximum value of t_s selected was 0.65 to reserve the useful information for the following classification step. Additionally, the threshold t_h and patch size w were set to 0.5 and 5, respectively. Table 5 shows the classification results with different t_s values. This table also shows that ACC is improved by the increase in t_s . In addition, all classification measurements reached their highest values when $t_s = 0.65$. Thus, the threshold t_s was fixed at 0.65 in the subsequent experiments to reduce the feature dimensionality and reserve useful information for the following classification step.

In the hierarchical classification step, the threshold t_h and patch size w are two crucial parameters that should be carefully determined. Thus, the two experiments were conducted to optimize the threshold t_h and patch size w , separately, during classification. In the first experiment, the proposed method was tested with different t_h values from 0 to 0.65. Moreover, the threshold t_s and patch size w were set to 0.65 and 5, respectively. Table 6 shows that the classification results with the threshold $t_h = 0.5$ were higher than the classification results without the threshold ($t_h = 0$). However, the classification accuracy was reduced when the threshold t_h increased at $t_h > 0.5$. A threshold of $t_h = 0.5$ was chosen for the subsequent experiments. In the second experiment, the patch size w was varied from 1 to 3, 5, 7, 9, and 11 to test the classification performance with these different values. In addition, the thresholds t_s and t_h were set to 0.65 and 0.5, respectively. As shown in Table 7, the ACC was improved by increasing w from 0 to 5, but further increasing the patch size to 7, 9, and 11 reduced ACC. Therefore, the patch size w was fixed at 5 for the subsequent experiments.

Effectiveness of the use of longitudinal data. To assess the effectiveness of the use of longitudinal data, the classification performance was evaluated with MRI scans of MCI subjects from the baseline visit data and the longitudinal data. In the experiment for baseline visit data, only the baseline visit data were used for comparison in the significant voxel selection and hierarchical classification steps. For fair comparison, we also used two nested cross-validation loops (10-fold for each loop) to carefully select the parameters with optimal performance for the baseline visit data. Moreover, the parameters t_s , t_h , and w were set to 0.65, 0.5, and 5, respectively, in accordance with the proposed method by using longitudinal data. The classification results of the proposed method with the baseline visit data and longitudinal data are shown in Table 8 and Fig. 3(a). The proposed method with longitudinal data consistently outperforms the proposed method that used baseline data in terms of ACC, SEN, SPE, and AUC. High SEN values indicated high confidence in AD prediction, which will significantly benefit the application of the method in real-life situations. The proposed method with longitudinal data significantly improved the sensitivity value (nearly 17% higher than the proposed method with baseline data). This high sensitivity may be advantageous for confident AD prediction and useful in practical applications.

Effectiveness of the hierarchical classification framework. To evaluate the effect of the hierarchical classification framework on the classification performance, we compared the obtained classification results by building a single global classifier and a hierarchical classification framework. For fair comparison, both classification methods were used on the same significant voxel set. We then used an LRC and the proposed hierarchical classification framework, respectively, to achieve the final classification result. Moreover, the parameters t_s , t_h , and w were set to 0.65, 0.5, and 5, respectively, in the proposed hierarchical classification method. Table 9 shows the classification results with respect to the single global classifier and the hierarchical classification framework. In addition, the ROC curves of different methods for classification of MCInc versus MCIc are illustrated in Fig. 3(b). These results demonstrate that the hierarchical classification framework performs better than the single classifier.

Parameter	Description	Setting
t_s	Threshold in significant voxel selection	0.65
t_h	Threshold in hierarchical classification framework	0.5
w	Patch size ($w \times w \times w$)	5

Table 4. Summary of the parameter settings in the proposed method for AD prediction.

t_s	0.5	0.55	0.6	0.65
ACC (%)	55.1	56.3	62.1	79.3
SEN (%)	62.0	67.6	64.2	87.7
SPE (%)	55.9	45.3	51.4	73.1

Table 5. Classification results of the proposed method with different t_s values.

t_h	0	0.5	0.55	0.6	0.65
ACC (%)	73.8	79.3	74.6	78.0	76.5
SEN (%)	83.8	87.7	85.0	84.1	83.8
SPE (%)	62.3	73.1	62.7	71.1	68.7

Table 6. Classification results of the proposed method with different t_h values.

w	1	3	5	7	9	11
ACC (%)	74.0	74.0	79.3	76.3	76.3	74.8
SEN (%)	86.8	79.7	87.7	86.2	90.2	92.8
SPE (%)	65.1	69.2	73.1	69.0	63.2	49.3

Table 7. Classification results of the proposed method with different w values.

Computation cost. In this study, the experiments were implemented on a standard PC with an Intel Xeon E5-2620 v3 processor at 2.40 GHz. To classify a subject in the testing step, the processing time was approximately 5 min; of which, 4 min was allotted for intensity and spatial normalization, and 1 min was allotted for the actual classification. In the training step, 6 threads were used to perform voxel selection and hierarchical classifier learning, the parameters t_s , t_h , and w were set to 0.65, 0.5, and 5, respectively. The total processing time for 118 MCI subjects was approximately 10 h; of which, 8 h was allotted for voxel selection, and 2 h was allotted for hierarchical classifier learning.

Discussion

We proposed a novel classification method for AD prediction. Our study is twofold. First, we investigated the longitudinal images of 131 individuals with MCI to obtain important information on the longitudinal change. The data were subsequently used to classify MCI subjects into either MCIC or MCInc. The longitudinal change is a crucial factor in the prediction of possible conversion from MCI to AD. This factor is widely used in AD conversion analysis^{20,23,38}. In most methods, the selected time points of MRI scans must be the same among individuals to capture changes in the longitudinal data. However, given that a longitudinal study in the field of health sciences commonly covers a relatively long period of time, some individuals may miss their scheduled visits¹⁹. Therefore, the same time points of MRI scans are difficult to implement across individuals. In the present study, the voxel intensities in the MRI images from the longitudinal data were used as features. The MRI data gathered before the missing time point were used as the missing data. Thus, we could use the longitudinal data at different time points. Second, we developed a hierarchical classification framework to address the high feature dimensionality issue and incorporate spatial information, thereby improving the classification accuracy. Unlike other hierarchical classification methods^{16,30}, our proposed strategy established multiple and multilevel classifiers with supervised learning and suitable thresholds to address the issues of high feature dimensionality and sensitivity to small changes. These characteristics enhance the classification accuracy. In the experiment, the classification results with a threshold t_h are consistently higher than the classification results without this threshold (Table 6). Therefore, unusable information can be discarded through the proposed hierarchical classification method with a suitable threshold.

In the hierarchical classification framework, the patch size w was adjusted to optimize the classification performance. Small patches may lack the required information for good performance in patch-level classification, and

Method	ACC (%)	SEN (%)	SPE (%)	AUC
Baseline visit	71.7	69.9	77.7	0.754
Longitudinal data	79.4	86.5	78.2	0.812

Table 8. Classification results of the proposed method with baseline visit data and longitudinal data.

Method	ACC (%)	SEN (%)	SPE (%)	AUC
Single classifier	64.9	54.9	78.0	0.712
Hierarchical classification	79.4	86.5	78.2	0.812

Table 9. Comparison of single classifier and hierarchical classification for MCIc versus MCIc classification.

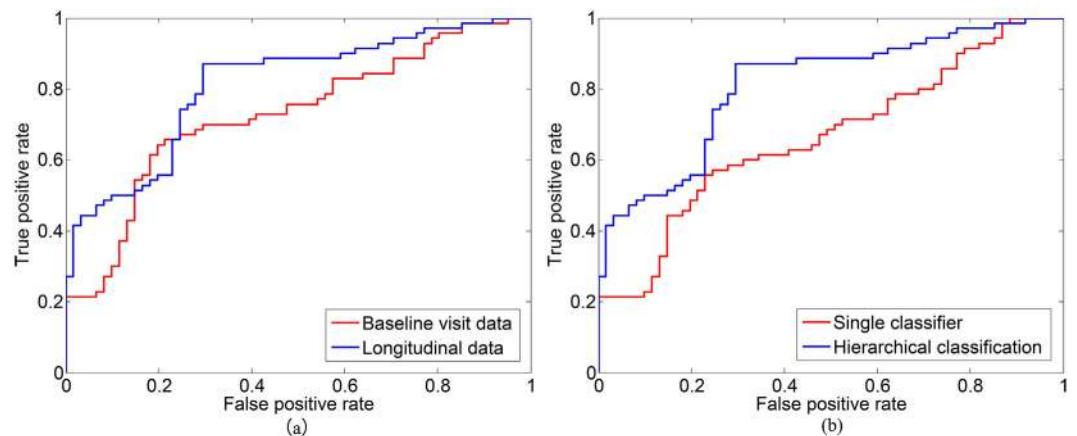


Figure 3. ROC curves for the classification of MCIc and MCInc obtained with (a) baseline visit data and longitudinal data and (b) a single classifier and hierarchical classification.

numerous patches or patch-level classifiers will significantly increase the computational cost for the classification. More redundant or even confounding information may be included in a large patch, thereby affecting the localization of informative brain regions and the ensemble classification results. In our experiment, the classification results with $w = 5$ were higher than those of other patch sizes (Table 7). Therefore, a moderate-sized patch was optimum in the proposed method as compared with other patch sizes.

Comparisons of baseline visit data versus longitudinal data and single classifier versus hierarchical classification were conducted to evaluate the performance of the proposed method in this study. Table 8 and Fig. 3(a) show that the classification results from longitudinal data are higher than those obtained with baseline visit data. These findings suggest that longitudinal change is a crucial factor for the prediction of future conversion of MCI to AD. Moreover, our experimental results show that the method with the hierarchical classification framework performs better than a single global classifier, probably because the hierarchical classification framework can better utilize local features and classifier decisions. In the hierarchical classification framework, the local spatial contiguity of image features is important during classification with a hierarchical spatial structure built from voxels to larger brain regions. The hierarchical spatial structure can utilize the local information better than the ROI-based methods³⁰. An ensemble method can also improve the generalizability and robustness of individual classifiers for better classification decisions as compared with individual classifiers². Therefore, the proposed hierarchical ensemble method can utilize the local features and make better classifier decisions than a single global classifier.

In the training stage, the proposed method requires that a classifier is trained for each voxel in the brain to select significant voxels and construct a hierarchical classification framework to train a hierarchical classifier. These two steps are off-line procedures; thus, each step is performed only once but used for all testing images. Moreover, both steps were run in multi-threads in this study, thereby significantly decreasing processing time.

Table 10 shows that the classification results of the proposed method (ACC = 79.4%, SEN = 86.5%) are comparable to the results of recently published papers. Tang *et al.*³⁹ and Wee *et al.*⁴⁰ extracted vertex-based features from MRI scans obtained from baseline visit data to classify MCI subjects into either MCIc or MCInc, and an accuracy of 75.0% and 75.1% were obtained, respectively. Liu *et al.*³⁰ proposed a hierarchical ensemble classification method to combine multilevel classifiers through gradual integration of a large number of features from local brain and interbrain regions. MRI scans from baseline visit data were used for AD prediction, and an accuracy of 64.8% was attained. Suk *et al.*¹⁶ first used the deep learning method to learn a high-level latent and shared feature representation. They then constructed a hierarchical classifier for the classification of MCIc versus MCInc. MRI and PET markers were included, and an accuracy of 75.9% was obtained. Zhang *et al.*³ used longitudinal data to

Method	Subjects (MCInc/MCIc)	Data source	Features	Classifier	ACC (%)	SEN (%)	SPE (%)
Korolev <i>et al.</i> ⁴¹	120/139 (baseline visit)	Risk factors, cognitive and functional assessments, MRI, plasma proteomic data	ROI-wise	Probabilistic multiple kernel learning	80.0	83.0	76.0
Tang <i>et al.</i> ³⁹	87/135 (baseline visit)	MRI	Vertex-based	LDA	75.0	77.0	71.0
Liu <i>et al.</i> ³⁰	128/ 76 (baseline visit)	MRI	Voxel-wise	Hierarchical ensemble	64.8	22.2	89.6
Suk <i>et al.</i> ¹⁶	128/76 (baseline visit)	MRI, PET	Voxel-wise	Hierarchical ensemble	75.9	48.0	95.2
Wee <i>et al.</i> ⁴⁰	111/89 (baseline visit)	MRI	Vertex-based	SVM	75.1	63.5	84.4
Zhang <i>et al.</i> ³	50/35 (longitudinal data)	MRI, PET, cognitive scores	ROI-wise	SVM	78.4	79.0	78.0
Proposed method	61/70 (longitudinal data)	MRI	Voxel-wise	Hierarchical ensemble	79.4	86.5	78.2

Table 10. Comparison of MCInc/MCIc classification accuracy in literature.

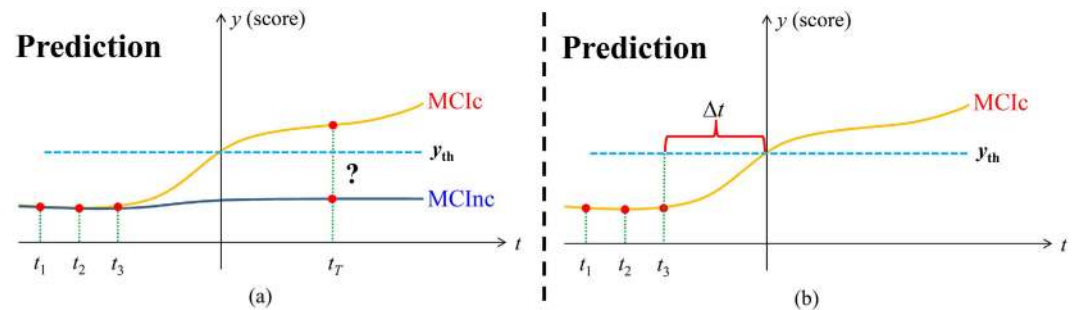


Figure 4. An example using longitudinal data to predict (a) AD conversion and clinical scores; (b) a rough time of AD conversion.

predict future conversion of patients with MCI, and an accuracy of 78.4% was obtained. Recently, Korolev *et al.*⁴¹ incorporated risk factors, cognitive and functional assessments, MRI, and plasma proteomic data for AD prediction. They obtained a high accuracy of 80.0%. Although the accuracy of our study is less than that of Korolev *et al.*'s study, our study is comparable to their model that incorporated only MRI data (ACC = 71.4%). Therefore, incorporating many data sources can potentially improve our prediction model in the future. However, only the results of different methods in literature are listed. Direct comparison of the performances of different methods is not reasonable because various datasets and methods for extracting features and building classifiers were used. Nonetheless, the proposed method showed the highest sensitivity and the second highest accuracy among the methods for MCInc/MCIc classification. These observations implied that the proposed method can potentially enhance confidence in AD prediction.

In this research, only MRI data were used. The data can be expanded to include other image modality data in future studies. Different image modalities can provide complementary information for disease diagnosis. Moreover, other data sources, such as clinical scores, genetic data, and demographic data, can be included to improve our prediction model in the future. Further advanced classifier ensemble methods, such as sparse multiple kernel learning⁴², can be investigated in future work to improve the classification performance. Finally, our method exclusively focused on the MCIc/MCInc classification. In the future, we aim to incorporate clinical scores, such as the Alzheimer's Disease Assessment Scale-Cognitive subscale and the Mini-Mental State Examination, to construct a joint regression and classification model. For instance, we can simultaneously perform AD and clinical score prediction with such a model. Furthermore, we can estimate the time when an MCIc subject develops AD by using the longitudinal data and the constructed model (Δt shown in Fig. 4).

Conclusion

This study presented a novel AD prediction method based on LMHC. Longitudinal images from individuals with MCI were investigated to obtain important information on the longitudinal changes for classifying MCI subjects into MCIc and MCInc. A hierarchical framework was introduced into the classifier to address the high feature dimensionality and incorporate spatial information for improved prediction accuracy. The performance of the proposed AD prediction method was tested on 131 patients with MCI with MRI scans at different time points. Our experimental results showed that longitudinal data and the hierarchical classification framework of our proposed method can improve the classification performance. To our knowledge, previous studies have not combined these two characteristics for AD prediction.

References

1. Cho, Y., Seong, J. K., Jeong, Y. & Shin, S. Y. Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *Neuroimage* **59**, 2217–2230 (2012).
2. Liu, M., Zhang, D. & Shen, D. Ensemble sparse classification of Alzheimer's disease. *Neuroimage* **60**, 1106–1116 (2012).

3. Zhang, D. Q., Shen, D. G. & Neuroimaging, A. S. D. Predicting Future Clinical Changes of MCI Patients Using Longitudinal and Multimodal Biomarkers. *PLoS one* **7**, doi: ARTN e3318210.1371/journal.pone.0033182 (2012).
4. Vos, F. D. *et al.* Combining Multiple Anatomical MRI Measures Improves Alzheimer's Disease Classification. *Human brain mapping* **37**, 1920–1929 (2016).
5. Liu, M., Zhang, D., Shen, D. & Alzheimer's Disease Neuroimaging, I. View-centralized multi-atlas classification for Alzheimer's disease diagnosis. *Human brain mapping* **36**, 1847–1865, doi: 10.1002/hbm.22741 (2015).
6. Magnin, B. *et al.* Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology* **51**, 73–83, doi: 10.1007/s00234-008-0463-x (2009).
7. 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. *Neuroimage* **44**, 1415–1422 (2009).
8. Lopez, M. *et al.* Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease. *Neurocomputing* **74**, 1260–1270 (2011).
9. Moradi, E. *et al.* Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* **104**, 398–412, doi: 10.1016/j.neuroimage.2014.10.002 (2015).
10. Park, H., Yang, J., Seo, J. & Lee, J. Dimensionality reduced cortical features and their use in the classification of Alzheimer's disease and mild cognitive impairment. *Neuroscience Letters* **529**, 123–127 (2012).
11. Westman, E., Muehlboeck, J. S. & Simmons, A. Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *Neuroimage* **62**, 229–238, doi: 10.1016/j.neuroimage.2012.04.056 (2012).
12. Coupe, P. *et al.* Simultaneous segmentation and grading of anatomical structures for patient's classification: application to Alzheimer's disease. *Neuroimage* **59**, 3736–3747, doi: 10.1016/j.neuroimage.2011.10.080 (2012).
13. Zhang, D., Shen, D. & Alzheimer's Disease Neuroimaging, I. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *Neuroimage* **59**, 895–907, doi: 10.1016/j.neuroimage.2011.09.069 (2012).
14. Zhu, X., Suk, H. I. & Shen, D. A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis. *Neuroimage* **100**, 91–105, doi: 10.1016/j.neuroimage.2014.05.078 (2014).
15. Cheng, B., Liu, M., Zhang, D., Munsell, B. C. & Shen, D. Domain Transfer Learning for MCI Conversion Prediction. *IEEE transactions on bio-medical engineering* **62**, 1805–1817, doi: 10.1109/TBME.2015.2404809 (2015).
16. Suk, H. I., Lee, S. W., Shen, D. & Alzheimer's Disease Neuroimaging, I. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *Neuroimage* **101**, 569–582, doi: 10.1016/j.neuroimage.2014.06.077 (2014).
17. Yendiki, A., Reuter, M., Wilkens, P., Rosas, H. D. & Fischl, B. Joint reconstruction of white-matter pathways from longitudinal diffusion MRI data with anatomical priors. *Neuroimage* **127**, 277–286, doi: 10.1016/j.neuroimage.2015.12.003 (2016).
18. Hyun, J. W. *et al.* STGP: Spatio-temporal Gaussian process models for longitudinal neuroimaging data. *Neuroimage* **134**, 550–562 (2016).
19. Li, Y. M. *et al.* Multiscale adaptive generalized estimating equations for longitudinal neuroimaging data. *Neuroimage* **72**, 91–105, doi: 10.1016/j.neuroimage.2013.01.034 (2013).
20. Li, Y. *et al.* Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. *Neurobiol Aging* **33**, doi: ARTN 427.e1510.1016/j.neurobiolaging.2010.11.008 (2012).
21. Petersen, R. C. *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI) Clinical characterization. *Neurology* **74**, 201–209 (2010).
22. Evans, A. C. & Grp, B. D. C. The NIH MRI study of normal brain development. *Neuroimage* **30**, 184–202, doi: 10.1016/j.neuroimage.2005.09.068 (2006).
23. Risacher, S. L. *et al.* Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort. *Neurobiol Aging* **31**, 1401–1418, doi: 10.1016/j.neurobiolaging.2010.04.029 (2010).
24. Davatzikos, C., Xu, F., An, Y., Fan, Y. & Resnick, S. M. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. *Brain* **132**, 2026–2035 (2009).
25. Chetelat, G. *et al.* Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: A longitudinal MRI study. *Neuroimage* **27**, 934–946, doi: 10.1016/j.neuroimage.2005.05.015 (2005).
26. McEvoy, L. K. *et al.* Mild Cognitive Impairment: Baseline and Longitudinal Structural MR Imaging Measures Improve Predictive Prognosis. *Radiology* **259**, 834–843, doi: 10.1148/radiol.11101975 (2011).
27. Hinrichs, C., Singh, V., Xu, G. F., Johnson, S. C. & Neuroimaging, A. D. Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population. *Neuroimage* **55**, 574–589, doi: 10.1016/j.neuroimage.2010.10.081 (2011).
28. Whitwell, J. L. *et al.* 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. *Brain* **130**, 1777–1786, doi: 10.1093/brain/awml12 (2007).
29. Hamalainen, A. *et al.* Voxel-based morphometry to detect brain atrophy in progressive mild cognitive impairment. *Neuroimage* **37**, 1122–1131, doi: 10.1016/j.neuroimage.2007.06.016 (2007).
30. Liu, M. H., Zhang, D. Q., Shen, D. G. & Initi, A. S. D. Hierarchical Fusion of Features and Classifier Decisions for Alzheimer's Disease Diagnosis. *Human brain mapping* **35**, 1305–1319, doi: 10.1002/hbm.22254 (2014).
31. Jack, C. R. *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging* **27**, 685–691, doi: 10.1002/jmri.21049 (2008).
32. Sled, J. G., Zijdenbos, A. P. & Evans, A. C. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging* **17**, 87–97, doi: 10.1109/42.668698 (1998).
33. Nyul, L. G., Udupa, J. K. & Zhang, X. New variants of a method of MRI scale standardization. *IEEE transactions on medical imaging* **19**, 143–150, doi: 10.1109/42.836373 (2000).
34. Mazziotta, J. *et al.* A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **356**, 1293–1322, doi: 10.1098/rstb.2001.0915 (2001).
35. Huang, M. Y. *et al.* Brain extraction based on locally linear representation-based classification. *Neuroimage* **92**, 322–339, doi: 10.1016/j.neuroimage.2014.01.059 (2014).
36. Hinrichs, C. *et al.* Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage* **48**, 138–149, doi: 10.1016/j.neuroimage.2009.05.056 (2009).
37. Huang, M. *et al.* FVGWAS: Fast voxelwise genome wide association analysis of large-scale imaging genetic data. *Neuroimage* **118**, 613–627, doi: 10.1016/j.neuroimage.2015.05.043 (2015).
38. Farzan, A., Mashohor, S., Ramli, A. R. & Mahmud, R. Boosting diagnosis accuracy of Alzheimer's disease using high dimensional recognition of longitudinal brain atrophy patterns. *Behav Brain Res* **290**, 124–130, doi: 10.1016/j.bbr.2015.04.010 (2015).
39. Tang, X., Holland, D., Dale, A. M., Younes, L. & Miller, M. I. Baseline Shape Diffeomorphometry Patterns of Subcortical and Ventricular Structures in Predicting Conversion of Mild Cognitive Impairment to Alzheimer's Disease. *Journal of Alzheimer's disease: JAD* **44**, 599–611, doi: 10.3233/JAD-141605 (2015).
40. Wee, C., Yap, P. & Shen, D. Prediction of Alzheimer's Disease and Mild Cognitive Impairment Using Baseline Cortical Morphological Abnormality Patterns. *Human brain mapping* **34**, 3411–3425, doi: 10.1002/hbm.22156 (2013).
41. Korolev, I. O., Symonds, L. L. & Bozoki, A. C. Predicting Progression from Mild Cognitive Impairment to Alzheimer's Dementia Using Clinical, MRI, and Plasma Biomarkers via Probabilistic Pattern Classification. *PLoS one* **11**, e0138866, doi: 10.1371/journal.pone.0138866 (2016).
42. Subrahmanya, N. & Shin, Y. C. Sparse Multiple Kernel Learning for Signal Processing Applications. *Ieee T Pattern Anal* **32**, 788–798, doi: 10.1109/Tpami.2009.98 (2010).

Acknowledgements

This work was supported by the Science and Technology Planning Project of Guangdong Province, China (grant number 2015B010131011); Major Program of National Natural Science Foundation of China (grant number U15012561016942); and National Natural Science Funds of China (NSFC, grant number 31371009 and 81601562). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, P fizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the US Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (<http://www.fnih.org>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Author Contributions

M.H. and Q.F. conceived and designed the experiments. M.H. and Y.W. performed the experiments. M.H. and W.C. analyzed the data. M.H. and Q.F. wrote the manuscript. The data used in this manuscript is obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI, PI Michael M. Weiner). 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.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Huang, M. *et al.* Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease. *Sci. Rep.* 7, 39880; doi: 10.1038/srep39880 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Consortia

The Alzheimer's Disease Neuroimaging Initiative Consortium

Michael W. Weiner², Paul Aisen³, Ronald Petersen^{4,5}, Clifford R. Jack Jr.⁵, William Jagust⁶, John O. Trojanowki⁷, Arthur W. Toga⁸, Laurel Beckett⁹, Robert C. Green¹⁰, Andrew J. Saykin¹¹, John Morris¹², Leslie M. Shaw¹², Jeffrey Kaye¹³, Joseph Quinn¹³, Lisa Silbert¹³, Betty Lind¹³, Raina Carter¹³, Sara Dolen¹³, Lon S. Schneider⁸, Sonia Pawluczyk⁸, Mauricio Beccera⁸, Liberty Teodoro⁸, Bryan M. Spann⁸, James Brewer¹⁴, Helen Vanderswag¹⁴, Adam Fleisher¹⁴, Judith L. Heidebrink¹⁵, Joanne L. Lord¹⁵, Sara S. Mason⁵, Colleen S. Albers⁵, David Knopman⁵, Kris Johnson⁵, Rachelle S. Doody¹⁶, Javier Villanueva-Meyer¹⁶, Munir Chowdhury¹⁶, Susan Rountree¹⁶, Mimi Dang¹⁶, Yaakov Stern¹⁷, Lawrence S. Honig¹⁷, Karen L. Bell¹⁷, Beau Ances¹², John C. Morris¹², Maria Carroll¹², Mary L. Creech¹², Erin Franklin¹², Mark A. Mintun¹², Stacy Schneider¹², Angela Oliver¹², Daniel Marson¹⁸, Randall Griffith¹⁸, David Clark¹⁸, David Geldmacher¹⁸, John Brockington¹⁸, Erik Roberson¹⁸, Marissa Natelson Love¹⁸, Hillel Grossman¹⁹, Effie Mitsis¹⁹, Raj C. Shah²⁰, Leyla deToledo-Morrell²⁰, Ranjan Duara²¹, Daniel Varon²¹, Maria T. Greig²¹, Peggy Roberts²¹, Marilyn Albert²², Chiadi Onyike²², Daniel D'Agostino²², Stephanie Kielb²², James E. Galvin²³, Brittany Cerbone²³, Christina A. Michel²³, Dana M. Pogorelec²³, Henry Rusinek²³, Mony J. de Leon²³, Lidia Glodzik²³, Susan De Santi²³, P. Murali Doraiswamy²⁴, Jeffrey R. Petrella²⁴, Salvador Borges-Neto²⁴, Terence Z. Wong²⁴, Edward Coleman²⁴, Charles D. Smith²⁵, Greg Jicha²⁵, Peter Hardy²⁵, Partha Sinha²⁵, Elizabeth Oates²⁵, Gary Conrad²⁵, Anton P. Porsteinsson²⁶, Bonnie S. Goldstein²⁶, Kim Martin²⁶, Kelly M. Makino²⁶, M. Saleem Ismail²⁶, Connie Brand²⁶, Ruth A. Mulnard²⁷, Gaby Thai²⁷, Catherine Mc-Adams-Ortiz²⁷, Kyle Womack²⁸, Dana Mathews²⁸, Mary Quiceno²⁸, Allan I. Levey²⁹, James J. Lah²⁹, Janet S. Cellar²⁹, Jeffrey M. Burns³⁰, Russell H. Swerdlow³⁰, William M. Brooks³⁰, Liana Apostolova³¹, Kathleen Tingus³¹, Ellen Woo³¹, Daniel H. S. Silverman³¹, Po H. Lu³¹, George Bartzokis³¹, Neill R. Graff-Radford³², Francine Parfitt³², Tracy Kendall³², Heather Johnson³², Martin R. Farlow¹¹, Ann Marie Hake¹¹, Brandy R. Matthews¹¹, Jared R. Brosch¹¹, Scott Herring¹¹, Cynthia Hunt¹¹, Christopher H. van Dyck³³, Richard E. Carson³³, Martha G. MacAvoy³³, Pradeep Varma³³, Howard Chertkow³⁴, Howard Bergman³⁴, Chris Hosein³⁴, Sandra Black³⁵, Bojana Stefanovic³⁵, Curtis Caldwell³⁵, Ging-Yuek Robin Hsiung³⁶, Howard Feldman³⁶, Benita Mudge³⁶, Michele Assaly³⁶, Elizabeth Finger³⁷, Stephen Pasternack³⁷, Irina Rachisky³⁷, Dick Trost³⁷, Andrew Kertesz³⁷, Charles Bernick³⁸, Donna Munic³⁸, Marek Marsel Mesulam³⁹, Kristine Lipowski³⁹, Sandra Weintraub³⁹, Borna Bonakdarpour³⁹, Diana Kerwin³⁹, Chuang-Kuo Wu³⁹, Nancy Johnson³⁹, Carl Sadowsky⁴⁰, Teresa Villena⁴⁰, Raymond Scott Turner⁴¹, Kathleen Johnson⁴¹, Brigid Reynolds⁴¹, Reisa A. Sperling⁴², Keith A. Johnson⁴², Gad Marshall⁴², Jerome Yesavage⁴³, Joy L. Taylor⁴³, Barton Lane⁴³, Allyson Rosen⁴³, Jared Tinklenberg⁴³, Marwan N. Sabbagh⁴⁴, Christine M. Belden⁴⁴, Sandra A. Jacobson⁴⁴, Sherye A. Sirrel⁴⁴, Neil Kowall⁴⁵, Ronald Killiany⁴⁵, Andrew E. Budson⁴⁵, Alexander Norbash⁴⁵, Patricia Lynn Johnson⁴⁵, Thomas O. Obisesan⁴⁶, Saba Wolday⁴⁶, Joanne Allard⁴⁶, Alan Lerner⁴⁷, Paula Ogrocki⁴⁷, Curtis Tatsuoka⁴⁷, Parianne Fatica⁴⁷, Evan Fletcher⁴⁸, Pauline Maillard⁴⁸, John Olichney⁴⁸, Charles DeCarli⁴⁸, Owen Carmichael⁴⁸, Smita Kittur⁴⁹, Michael Borrie⁵⁰, T-Y Lee⁵⁰, Rob Bartha⁵⁰, Sterling Johnson⁵¹, Sanjay Asthana⁵¹, Cynthia M. Carlsson⁵¹, Steven G. Potkin⁵², Adrian Preda⁵², Dana Nguyen⁵², Pierre Tariot⁵³, Anna Burke⁵³, Nadira Trncic⁵³, Adam Fleisher⁵³, Stephanie Reeder⁵³, Vernice Bates⁵⁴, Horacio Capote⁵⁴, Michelle Rainka⁵⁴, Douglas W. Scharre⁵⁵, Maria Kataki⁵⁵, Anahita Adeli⁵⁵, Earl A. Zimmerman⁵⁶, Dzintra Celmins⁵⁶, Alice D. Brown⁵⁶, Godfrey D. Pearlson⁵⁷, Karen Blank⁵⁷, Karen Anderson⁵⁷, Laura A. Flashman⁵⁸, Marc Seltzer⁵⁸, Mary L. Hynes⁵⁸, Robert B. Santulli⁵⁸, Kaycee M. Sink⁵⁹, Leslie Gordineer⁵⁹, Jeff D. Williamson⁵⁹, Pradeep Garg⁵⁹, Franklin Watkins⁵⁹, Brian R. Ott⁶⁰, Henry Querfurth⁶⁰, Geoffrey Tremont⁶⁰, Stephen Salloway⁶¹, Paul Malloy⁶¹, Stephen Correia⁶¹, Howard J. Rosen⁶², Bruce L. Miller⁶², David Perry⁶², Jacobo Mintzer⁶³, Kenneth Spicer⁶³, David Bachman⁶³, Nunzio Pomara⁶⁴, Raymundo Hernando⁶⁴, Antero Sarrael⁶⁴, Norman Relkin⁶⁵, Gloria Chaing⁶⁵, Michael Lin⁶⁵, Lisa Ravdin⁶⁵, Amanda Smith⁶⁶, Balebail Ashok Raj⁶⁶ & Kristin Fargher⁶⁶

²Magnetic Resonance Unit at the VA Medical Center and Radiology, Medicine, Psychiatry and Neurology, University of California, San Francisco, USA. ³San Diego School of Medicine, University of California, California, USA. ⁴Mayo Clinic, Minnesota, USA. ⁵Mayo Clinic, Rochester, USA. ⁶University of California, Berkeley, USA. ⁷University of

Pennsylvania, Pennsylvania, USA. ⁸University of Southern California, California, USA. ⁹University of California, Davis, California, USA. ¹⁰MPH Brigham and Women's Hospital/Harvard Medical School, Massachusetts, USA. ¹¹Indiana University, Indiana, USA. ¹²Washington University St. Louis, Missouri, USA. ¹³Oregon Health and Science University, Oregon, USA. ¹⁴University of California–San Diego, California, USA. ¹⁵University of Michigan, Michigan, USA. ¹⁶Baylor College of Medicine, Houston, State of Texas, USA. ¹⁷Columbia University Medical Center, South Carolina, USA. ¹⁸University of Alabama–Birmingham, Alabama, USA. ¹⁹Mount Sinai School of Medicine, New York, USA. ²⁰Rush University Medical Center, Rush University, Illinois, USA. ²¹Wien Center, Florida, USA. ²²Johns Hopkins University, Maryland, USA. ²³New York University, NY, USA. ²⁴Duke University Medical Center, North Carolina, USA. ²⁵University of Kentucky, Kentucky, USA. ²⁶University of Rochester Medical Center, NY, USA. ²⁷University of California, Irvine, California, USA. ²⁸University of Texas Southwestern Medical School, Texas, USA. ²⁹Emory University, Georgia, USA. ³⁰University of Kansas, Medical Center, Kansas, USA. ³¹University of California, Los Angeles, California, USA. ³²Mayo Clinic, Jacksonville, USA. ³³Yale University School of Medicine, Connecticut, USA. ³⁴McGill University, Montreal-Jewish General Hospital, Montreal, Quebec, Canada. ³⁵Sunnybrook Health Sciences, Ontario, USA. ³⁶U.B.C. Clinic for AD & Related Disorders, Vancouver, Canada. ³⁷Cognitive Neurology - St. Joseph's, Ontario, USA. ³⁸Cleveland Clinic Lou Ruvo Center for Brain Health, Ohio, USA. ³⁹Northwestern University, Illinois, USA. ⁴⁰Premiere Research Inst (Palm Beach Neurology), Florida, USA. ⁴¹Georgetown University Medical Center, Washington D.C., USA. ⁴²Brigham and Women's Hospital, Massachusetts, USA. ⁴³Stanford University, California, USA. ⁴⁴Banner Sun Health Research Institute, Arizona, USA. ⁴⁵Boston University, Massachusetts, USA. ⁴⁶Howard University, Washington D.C., USA. ⁴⁷Case Western Reserve University, Ohio, USA. ⁴⁸University of California, Davis–Sacramento, California, USA. ⁴⁹Neurological Care of CNY, New York, USA. ⁵⁰Parkwood Hospital, Pennsylvania, USA. ⁵¹University of Wisconsin, Wisconsin, USA. ⁵²University of California, Irvine – BIC, California, USA. ⁵³Banner Alzheimer's Institute, Arizona, USA. ⁵⁴Dent Neurologic Institute, NY, USA. ⁵⁵Ohio State University, Ohio, USA. ⁵⁶Albany Medical College, NY, USA. ⁵⁷Hartford Hospital, Olin Neuropsychiatry Research Center, Connecticut, USA. ⁵⁸Dartmouth-Hitchcock Medical Center, New Hampshire, USA. ⁵⁹Wake Forest University Health Sciences, North Carolina, USA. ⁶⁰Rhode Island Hospital, state of Rhode Island, USA. ⁶¹Butler Hospital, Providence, Rhode Island, USA. ⁶²University of California, San Francisco, USA. ⁶³Medical University South Carolina, South Carolina, USA. ⁶⁴Nathan Kline Institute, Orangeburg, New York, USA. ⁶⁵Cornell University, Ithaca, New York, USA. ⁶⁶USF Health Byrd Alzheimer's Institute, University of South Florida, Florida, USA.