Analysis of structural brain MRI and multi-parameter classification for Alzheimer's disease

Article *in* Biomedizinische Technik/Biomedical Engineering · June 2017
DOI: 10.1515/bmt-2016-0239

2 authors:

Yingteng Zhang
South China University of Technology
2 PUBLICATIONS **1** CITATION
SEE PROFILE

Liu Shenquan
South China University of Technology
78 PUBLICATIONS **124** CITATIONS
SEE PROFILE

Yingteng Zhang and Shenquan Liu*

Analysis of structural brain MRI and multi-parameter classification for Alzheimer's disease

DOI 10.1515/bmt-2016-0239
Received August 25, 2016; accepted May 16, 2017

Abstract: Incorporating with machine learning technology, neuroimaging markers which extracted from structural Magnetic Resonance Images (sMRI), can help distinguish Alzheimer's Disease (AD) patients from Healthy Controls (HC). In the present study, we aim to investigate differences in atrophic regions between HC and AD and apply machine learning methods to classify these two groups. T1-weighted sMRI scans of 158 patients with AD and 145 age-matched HC were acquired from the ADNI database. Five kinds of parameters (i.e. cortical thickness, surface area, gray matter volume, curvature and sulcal depth) were obtained through the preprocessing steps. The recursive feature elimination (RFE) method for support vector machine (SVM) and leave-one-out cross validation (LOOCV) were applied to determine the optimal feature dimensions. Each kind of parameter was trained by SVM algorithm to acquire a classifier, which was used to classify HC and AD ultimately. Moreover, the ROC curves were depicted for testing the classifiers' performance and the SVM classifiers of two-dimensional spaces took the top two important features as classification features for separating HC and AD to the maximum extent. The results showed that the decreased cortical thickness and gray matter volume dramatically exhibited the trend of atrophy. The key differences between AD and HC existed in the cortical thickness and gray matter volume of the entorhinal cortex and medial orbitofrontal cortex. In terms of classification results, an optimal accuracy of 90.76% was obtained via multi-parameter combination (i.e. cortical thickness, gray matter volume and surface area). Meanwhile, the receiver operating characteristic (ROC) curves and area under the curve (AUC) were also verified multi-parameter combination could reach a better classification performance (AUC = 0.94) after the SVM-RFE method. The results could be well prove that multi-parameter combination could provide more useful classified features from multivariate anatomical structure than single parameter. In addition, as cortical thickness and multi-parameter combination contained more important classified information with fewer feature dimensions after feature selection, it could be optimum to separate HC from AD to take the top two important features of them to construct SVM classifiers in two-dimensional space. The proposed work is a promising approach suggesting an important role for machine-learning based diagnostic image analysis for clinical practice.

Keywords: Alzheimer's disease; classification; cortical feature; multi-parameter combination; structural MRI; support vector machine.

Introduction

Alzheimer's disease (AD) is a severe neurodegenerative disease which results in cognitive impairment and memory ability damage, even leading to unceasing deterioration of viability and death in the end [23]. Recent studies indicated that there were 24.3 million people with dementia in the world in 2001, and predicted that this would rise to 42.3 million in 2020 and 81.1 million by 2040 [21]. Traditionally neuropsychological tests are time-consuming with passable recognition rate for AD [45]. With the development of neuroimaging technology, sMRI has initiated a non-invasive and widely prevalent method, which can be used to detect more subtle morphological abnormalities in brain disorders [11, 18, 44]. Therefore, to discover the abnormalities of imaging characteristics of AD and prevent the progression of disease would be particularly urgent.

To our knowledge, gradual cerebral atrophy is one of the obvious changes of AD and the degree of atrophy can be observed via high-resolution MRI technology. Morphology-related cortical volume and cortical thickness measure have been used to better understand the underlying pathophysiology in AD diagnosis. Nowadays diverse tools such as statistical parametric mapping (SPM) [24], voxel-based morphometry (VBM) [4] and Freesurfer [22] could be applied in extracting morphologic whole-brain MRI features. However, the majority of these studies [9, 10, 26, 42] mainly focused on the differences of cortical

***Corresponding author: Shenquan Liu,** School of Mathematics, South China University of Technology, Guangzhou 510640, China, E-mail: mashqliu@scut.edu.cn
Yingteng Zhang: School of Mathematics, South China University of Technology, Guangzhou 510640, China

thickness and gray matter volume. It is still lacking with respect to the research on surface area, curvature and sulcal depth. In addition, previous studies [15, 47] had specifically investigated the use of hippocampal volumes and medial temporal lobe for diagnosis in AD. The combination of multiple cortical metrics (i.e. volumes and thicknesses) across multiple brain regions may provide better discrimination between AD and HC. Therefore, the advanced multivariate machine-learning based approaches, which can extract the fine-grained relationship among the multiple metrics and regions [8, 37, 51], have significant potential for assisting with diagnosis and prediction of AD.

Recently, a review [3] roundly summarized the single subject prediction of several brain disorders such as AD and involved several key aspects such as modality, machine learning algorithms, sample size and extracted features. This implies that the multi-modal neuroimaging techniques, advanced machine learning algorithms, big data, multi-parameter combination will become the mainstream of neuroimaging research. To our knowledge, the machine learning approaches can train a classifier to predict the label of an unseen subject by taking multi-regional brain features into account jointly. Specifically, machine learning can capture the size of weight coefficients among various anatomical regions, which have exhibited their importance to distinguish AD from HC [49]. There are several different aspects for classification research of machine learning. One is classification using the features extracted from T1-weighted images and diffusion tensor images (DTI). For example, Li and colleagues [33] reported an accuracy of 94.3% in discriminating between AD and HC among 36 individuals after combined the tract-based fractional anisotropy (FA) with gray matter volumes. The second one [2] compared four supervised learning methods [i.e. orthogonal projections to latent structures (OPLS), decision trees, artificial neural networks (ANN) and SVM], who indicated that SVM and OPLS were slightly superior to decision trees and ANN. The third one [7] researched the classification effect across the sample sizes and feature dimensions among AD, mild cognitive impairment (MCI) and HC. Given that multi-parameter combination and large sample has enormous potential for improving discriminative capability, it is likely that the above two factors can contribute to extraction of prominent features and stability of classification model.

In this study, we applied a machine learning approach to discriminate AD from HC. Structural MRI data was firstly preprocessed by FreeSurfer to receive five kinds of parameters and applied in query design estimate contrast (QDEC) for statistical analysis. Except common cortical thickness and gray matter volume, we also took surface area, curvature and sulcal depth as cortical parameters. Then the SVM algorithm was used to classify HC and AD. The ROC curve and SVM classifier of two-dimensional space were depicted to validate the classification performance of SVM classifier.

Materials and methods

Data

Our participants involved in the experimental analysis were obtained from the Alzheimer's disease neuroimaging initiative (ADNI) database (adni.loni.usc.edu), which is consisted of high-resolution T1-weighted sMRI of 303 participants. The screening process was described in detail in the database manuals. The T1-weighted structural image parameters of all participants were roughly described as follows: TR = 2400 ms, TE = 4 ms, slice thickness = 1.0 mm, voxel size 1.0 mm × 1.0 mm × 1.0 mm. The exact parameters were varied slightly across scanners. Before scanning, the participants experienced cognitive and behavioral assessments. There was no significant difference ($p > 0.05$) between HC group and AD group when comparing age and gender (See Table 1 for group characteristics). There were differences between groups for demographics including mini-mental state examination (MMSE), clinical dementia rating (CDR), such as AD had a lower score of MMSE but a higher score of CDR than HC. The statistical analysis of basic information was completed in SPSS 22.0.

Freesurfer analysis

The T1-weighted sMRI scans were processed with Freesurfer's recon-all preprocessing for cortical reconstruction and volumetric segmentation [22], freely available at http://surfer.nmr.mgh.harvard.edu. This method automatically generated reliable volume and thickness segmentations of white matter, gray matter, and subcortical volumes. The streamlined pipeline included the removal of non-brain tissue, Tailarach transformations, segmentation of subcortical gray

Table 1: Mean and standard deviation of sample demographics used in this study.

Characteristic	HC	AD	p-Value
Sample size	145	158	
Gender (male/female)	71/74	86/72	0.342
Age (years)	75.76 ± 4.42	75.21 ± 7.47	0.432
MMSE	29.14 ± 0.94	23.30 ± 2.05	<0.000001
CDR	0.04 ± 0.14	4.68 ± 1.74	<0.000001

The data are represented as mean ± standard deviation. Columns on the right display p-values for two sample t-tests for each sample characteristic except for gender, which displays p-value from a χ^2 test.

and white matter regions, intensity normalization and atlas registration. After these steps, a mesh model of the cortical surface was generated and the cortical surface was parcellated into 34 cortical regions based on gyral and sulcal landmarks for each hemisphere according to the Desikan-Killiany atlas [14]. For purpose of statistical analysis, smoothing was applied using recon-all with qcache option added. QDEC, a tool within Freesurfer, was used to identify differences in cortical thickness, surface area, gray matter volume and cortical folding (curvature and sulcal depth) between HC and AD. To control for multiple comparison, statistical significance levels were cluster corrected for both hemispheres using false discovery rate (FDR), $p < 0.05$ [48].

Machine learning methods and analysis

The machine-learning analysis was performed by using a scikit-learn open source package [38] (version 0.18.1, freely available http://scikit-learn.org/) in Python. Scikit-learn is a Python module integrating with a wide range of advanced machine learning algorithms for supervised and unsupervised problems. The basic function of scikit-learn mainly contains six parts: classification, regression, clustering, dimensionality reduction, model selection and preprocessing. For specific machine learning problem, it usually could be divided into three steps: data preparation and preprocessing; model selection and training; model validation and parameter optimization.

In terms of classification methods, support vector machine (SVM) [5] was by far the most popular method. SVM was already known as a tool that discovered informative patterns [27]. The present application demonstrated that SVM was also very effective for discovering informative features or attributes. Different forms of SVM such as linear, non-linear along with recursive feature elimination and regularization had been used for classification of various disorders. In this paper, linear support vector machine recursive feature elimination (SVM-RFE) [27] was applied in order to obtain a ranked list of features which could best distinguished HC from AD. The SVM-RFE method allowed one to minimize redundant and extraneous features that could potentially degrade classifier performance [19]. SVM-RFE worked backwards from the initial set of features and eliminated the least "useful" feature on each recursive pass and it had been applied successfully for feature selection across several functional neuroimaging studies [12, 13]. The performance of the classifier was evaluated by utilizing the leave-one-out cross validation (LOOCV) test. In each trial, the data from all but one (N – 1 of the N sample) to train the classifier, then the classifier tested the remaining one. This procedure was repeated N times and then each time leaving out a different sample [52]. The learning and classification process involved four steps: (i) dividing the subjects into a training set and a testing set, (ii) selecting discriminative regions, (iii) training the SVM classifier model using the training data, and (iv) evaluating the performance of the SVM model using the testing data. To determine the general performance of the SVM classifier, a LOOCV approach was taken. Every subject was selected once as the testing dataset, with the remaining 302 subjects forming the training dataset (step i). Step (ii) and step (iii) were performed to select the discriminative features and trained the SVM classifier model. The final step (step iv) was to evaluate the performance of the SVM model using the testing data. An overview of AD/HC process flow chart was shown in Figure 1.

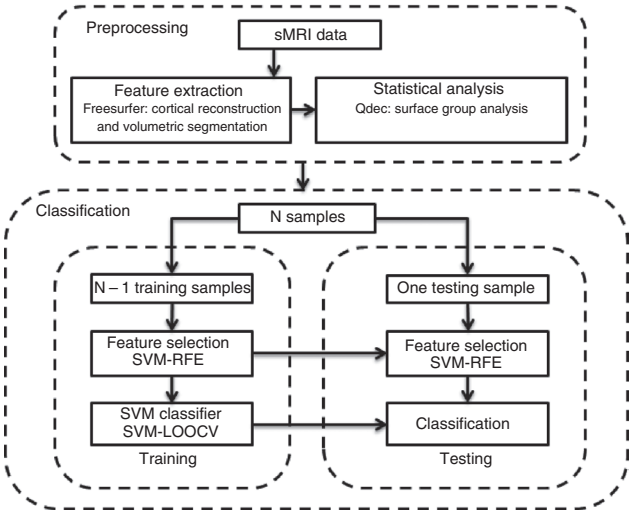

Figure 1: Process flow chart of AD/HC preprocessing and classification.

The results of classification were the mean accuracy, sensitivity and specificity. The accuracy was defined as accuracy = (TP + TN)/n where TP was the number of true positives, TN was the number of true negatives and n was the total number of subjects. Following a common convention, we defined correctly classified patients with AD as true positives. The sensitivity and the specificity measured the ability of a classifier to identify positive and negative instances, i.e. sensitivity = TP/(TP + FN), specificity = TN/(TN + FP), where FN and FP were the number of false negative and false positive instances, respectively. The classification performance computed from the confusion matrix was shown in Table 3. In order to understand the performance of a classifier, it was important to report the sensitivity or specificity along with the overall accuracy. The other common way of reporting results for a classifier was by plotting "receiver operating characteristic" (ROC) curve [53]. The ROC curve was the plot of sensitivity against "1-specificity" by changing the discrimination threshold and therefore provided a complete picture of classifier's performance. The ROC curve was usually summarized by the area under the curve (AUC), which was a number between 0 and 1 [20].

Results

Statistical analysis

The statistical analysis between HC and AD is performed on QDEC, a module of Freesurfer developed to design and execute surface analysis. Color areas showing significant distinction among groups are superimposed on the template (Figure 2).

The results of cortical thickness, surface area, gray matter volume, curvature and sulcal depth analysis are drawn in Figure 2. Based on the Desikan-Killiany atlas

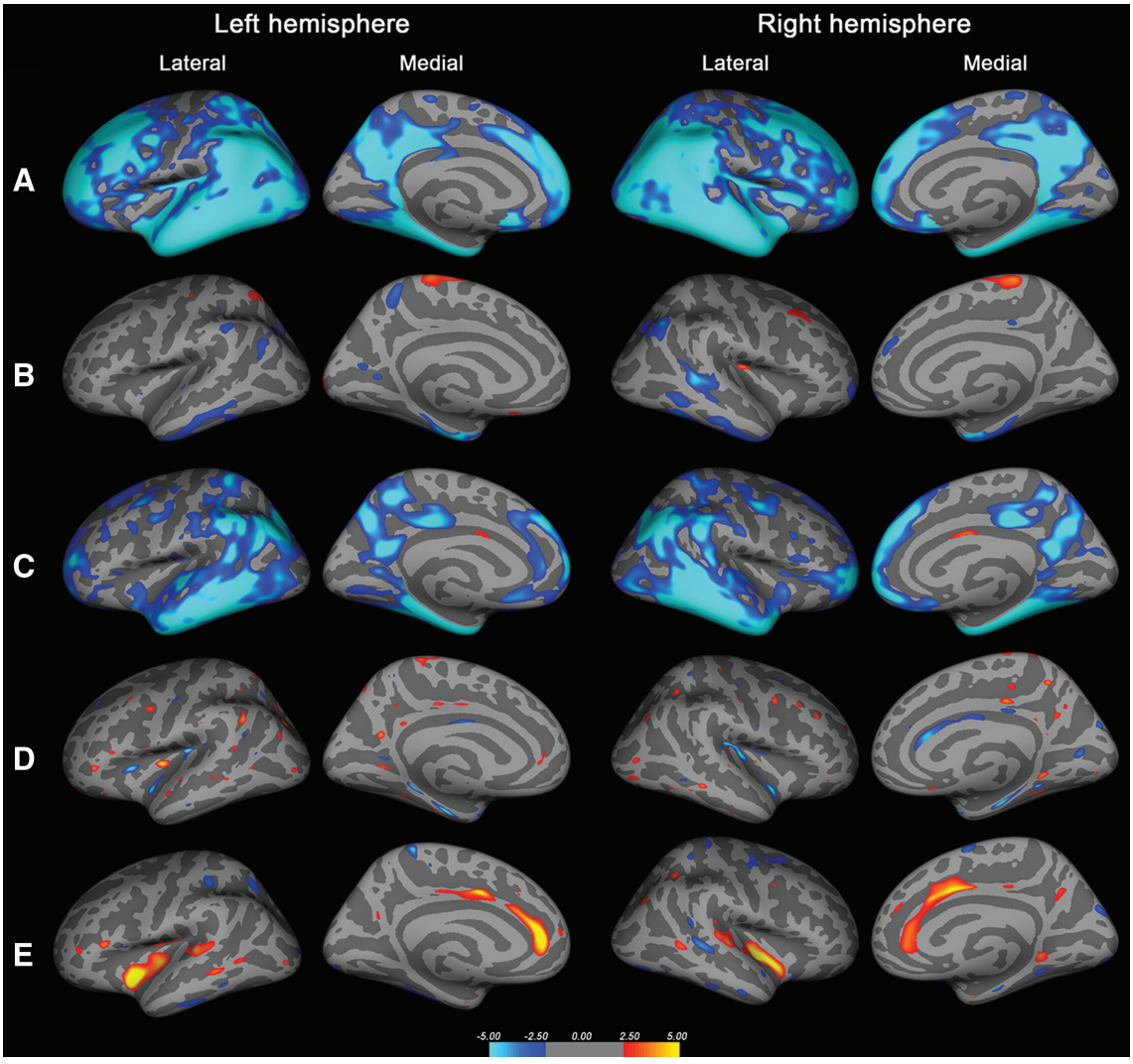

Figure 2: Analysis of five kinds of parameters (including cortical thickness, surface area, gray matter volume, curvature, sulcal depth) to differentiate between HC and AD.
For each row from top to bottom, statistical maps show parameter pattern in HC relative to AD for the (A) cortical thickness, (B) surface area, (C) gray matter volume, (D) curvature, (E) sulcal depth of the left and right hemisphere presented on the inflated cortical surface (dark grey = sulci; light gray = gyri). For each column from left to right, statistical maps show parameter pattern in lateral and medial of the left hemisphere and right hemisphere respectively. An increased parameter is marked by red and yellow, and a decreased parameter is marked by dark and pale blue. The significant thresholds are set at $p < 0.05$, FDR corrected. Cortical maps are smoothed with full-width at half-maximum (FWHM) Gaussian kernel set at 10 mm. The color bar scale represents maximum –log10 (p-value) in the cluster.

[14], the human cerebral cortex is divided into 34 cortical features in each hemisphere. As the number of discrepant features is much, only several top-ranked significant differences of features are listed for five kinds of parameters. Table 2 presents the position and involves a range of clusters of differences in cortical thickness, surface area, gray matter volume, curvature and sulcal depth at each vertex between HC and AD by QDEC controlling. In this table, only the top four features which have significant differences for each kind of parameter are provided. Combined with statistical maps Figure 2 and statistical Table 2, the cortical thickness of the bilateral entorhinal, bilateral paracentral, left medial orbitofrontal, left superior parietal, right cuneus, right postcentral was thinner in AD compared with HC. In the surface area, AD have smaller areas than HC in the left inferior temporal, left precuneus, right entorhinal, right bankssts, right inferior parietal and have greater areas than HC in the left precentral, left superior parietal, right insula. Compared with HC, the gray matter volume of the bilateral entorhinal, bilateral precentral, left superior frontal, left caudal middle frontal, right posterior cingulate and right rostral

Table 2: Clusters of differences in cortical thickness, surface area, gray matter volume, curvature, sulcal depth between HC and AD for each hemisphere.

Cortical feature	Cluster region	Maximum –log10 (p-value) in the cluster	Surface area of cluster (mm^2)	Peak coordinates (TalX, TalY, TalZ)
Cortical thickness	Left entorhinal	−31.9509	36397.37	−20.3, −9.8, −30.4
	Left medial orbitofrontal	−10.4145	18753.17	−6.7, 21.3, −12.0
	Left paracentral	−3.5496	76.61	−6.0, −30.7, 60.7
	Left superior parietal	−2.8016	182.31	−10.4, −92.2, 20.1
	Right entorhinal	−28.2426	56846.33	23.4, −10.2, −32.2
	Right cuneus	−3.5091	138.44	6.0, −84.2, 18.1
	Right paracentral	−3.2123	132.92	4.1, −37.0, 63.7
	Right postcentral	−2.9332	147.51	49.5, −19.0, 56.4
Surface area	Left inferior temporal	−6.0970	2858.19	−41.3, −10.6, −32.4
	Left precentral	3.6319	500.20	−5.9, −28.8, 71.5
	Left precuneus	−3.3975	253.55	−9.4, −49.5, 59.9
	Left superior parietal	2.9183	177.05	−23.0, −61.0, 59.0
	Right entorhinal	−7.1872	3234.69	26.0, −0.4, −30.6
	Right bankssts	−4.8450	908.26	46.8, −43.0, 3.4
	Right inferior parietal	−4.3094	995.26	41.8, −64.5, 44.6
	Right insula	3.5801	93.18	34.6, −18.4, 19.6
Gray matter volume	Left entorhinal	−27.0247	28227.10	−21.3, −8.0, −31.5
	Left superior frontal	−6.1045	7019.18	−9.5, 59.9, 9.5
	Left caudal middle frontal	−4.5035	583.28	−38.6, 6.4, 54.1
	Left precentral	−3.9812	403.43	−50.4, −2.6, 37.9
	Right entorhinal	−23.8018	25123.34	23.9, −8.2, −33.1
	Right posterior cingulate	−9.2975	1270.68	5.3, −30.0, 40.6
	Right rostral middle frontal	−8.9783	8604.92	29.5, 49.8, −3.1
	Right precentral	−5.2569	1239.65	49.5, −5.3, 49.0
Curvature	Left insula	−6.0272	109.69	−30.9, −27.7, 15.8
	Left entorhinal	−5.9569	159.26	−21.6, −16.5, −29.2
	Left parahippocampal	−5.1017	51.80	−19.4, −32.3, −13.6
	Left superior temporal	−4.8077	75.11	−41.0, −10.3, −15.9
	Right supramarginal	−6.3087	219.14	36.2, −34.0, 16.9
	Right superior temporal	−5.3212	81.26	43.5, −8.2, −16.3
	Right parahippocampal	−5.0900	135.18	23.2, −19.6, −26.5
	Right temporal pole	−5.0422	49.17	30.3, 5.4, −32.5
Sulcal depth	Left insula	8.9027	995.78	−35.9, −2.6, −7.2
	Left rostral anterior cingulate	7.2678	504.93	−9.0, 38.9, 9.5
	Left posterior cingulate	5.1191	267.25	−5.8, 3.0, 37.5
	Left paracentral	−4.7826	135.83	−10.4, −41.1, 64.4
	Right superior temporal	6.4106	768.71	43.9, −17.7, −7.9
	Right posterior cingulate	5.2002	873.02	7.5, 5.9, 38.1
	Right bankssts	−3.8653	228.25	44.1, −38.6, 1.1
	Right parahippocampal	3.5100	125.27	19.1, −41.3, −6.3

middle frontal was smaller in AD. In the curvature, HC have smaller curvatures than AD in the bilateral parahippocampal, bilateral superior temporal, left insula, left entorhinal, right supramarginal and right temporal pole. Moreover, AD have greater depths than HC in the bilateral posterior cingulate, left insula, left rostral anterior cingulate, left posterior cingulate, right superior temporal, right parahippocampal and have smaller curvatures than HC in the left paracentral, right bankssts in the sulcal depth.

Feature selection and classification accuracy

The SVM-RFE method is firstly applied to 68 cortical features for five kinds of parameters to discriminate HC from AD (Figure 3A–E). The SVM-RFE method is then applied in multi-parameter combination to receive significant cortical features. (Fig. 3F). The multi-parameter combination was constructed by concatenating the cortical thickness (CT), gray matter volume (GMV) and surface area (SA) into a long feature vector.

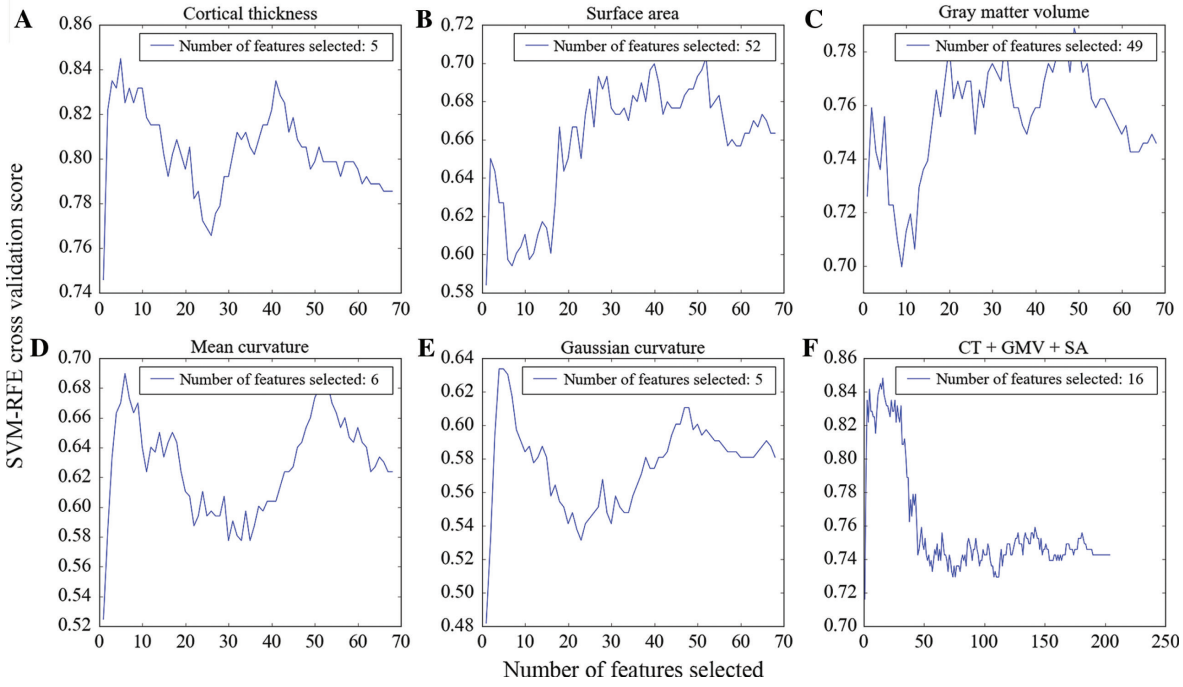

Figure 3: Selection of the optimal feature dimensions of cortical features via the SVM-RFE and LOOCV method.
(A)–(F) represent the SVM-RFE cross validation score corresponding to the number of features selected from cortical thickness, surface area, gray matter volume, mean curvature, Gaussian curvature and CT + GMV + SA. The optimal feature dimensions have been described in the rectangular frame for each kind of parameter. It is remarkable that except diagram (F) which is computed based on 204 features, the rest is based on 68 features including the whole cerebral cortex. CT, Cortical thickness; GMV, gray matter volume; SA, surface area.

Via the SVM-RFE method combined with LOOCV, the optimal feature dimensions for six kinds of parameters have been drawn in the corresponding rectangular frame (Figure 3). The peak value of the curve corresponding to the value of the X-axis is the best number of features. It seems that the range of feature selection for cortical thickness and Gaussian curvature is most remarkable (Table 3). The range of feature selection for mean curvature can take the second place, while the effect of feature selection for surface area and gray matter volume is not particularly significant. For the classified quality of single parameter, gray matter volume receives the highest accuracy, sensitivity and specificity which are 86.80%, 86.08% and 87.59% respectively. In addition, the classified quality of the rest four kinds of parameters also obtains improvement via the SVM-RFE method. By fusing cortical thickness, gray matter volume and surface area together, the classification accuracy gets a decent increase in comparison to single parameter classification. The accuracy, sensitivity and specificity increase from 74.25%, 76.58%, 71.72% to 90.76%, 89.87%, 91.72% respectively after SVM-RFE.

Evaluation of classifier performance

ROC curves typically depict true positive rate (TPR) on the Y-axis, and false positive rate (FPR) on the X-axis. This means that the top left corner of the plot is the “ideal”

Table 3: Classification performance for each kind of parameter.

Parameter	Optimal feature dimensions	Accuracy (%)		Sensitivity (%)		Specificity (%)	
		Before	After	Before	After	Before	After
Cortical thickness	5	78.55	85.48	79.11	84.81	77.93	86.21
Surface area	52	66.34	76.24	68.35	79.11	64.14	73.10
Gray matter volume	49	74.58	86.80	74.05	86.08	75.17	87.59
Mean curvature	6	62.38	69.97	58.86	69.62	66.21	70.34
Gaussian curvature	5	58.09	65.68	48.10	52.53	68.97	80.00
CT + GMV + SA	16	74.25	90.76	76.58	89.87	71.72	91.72

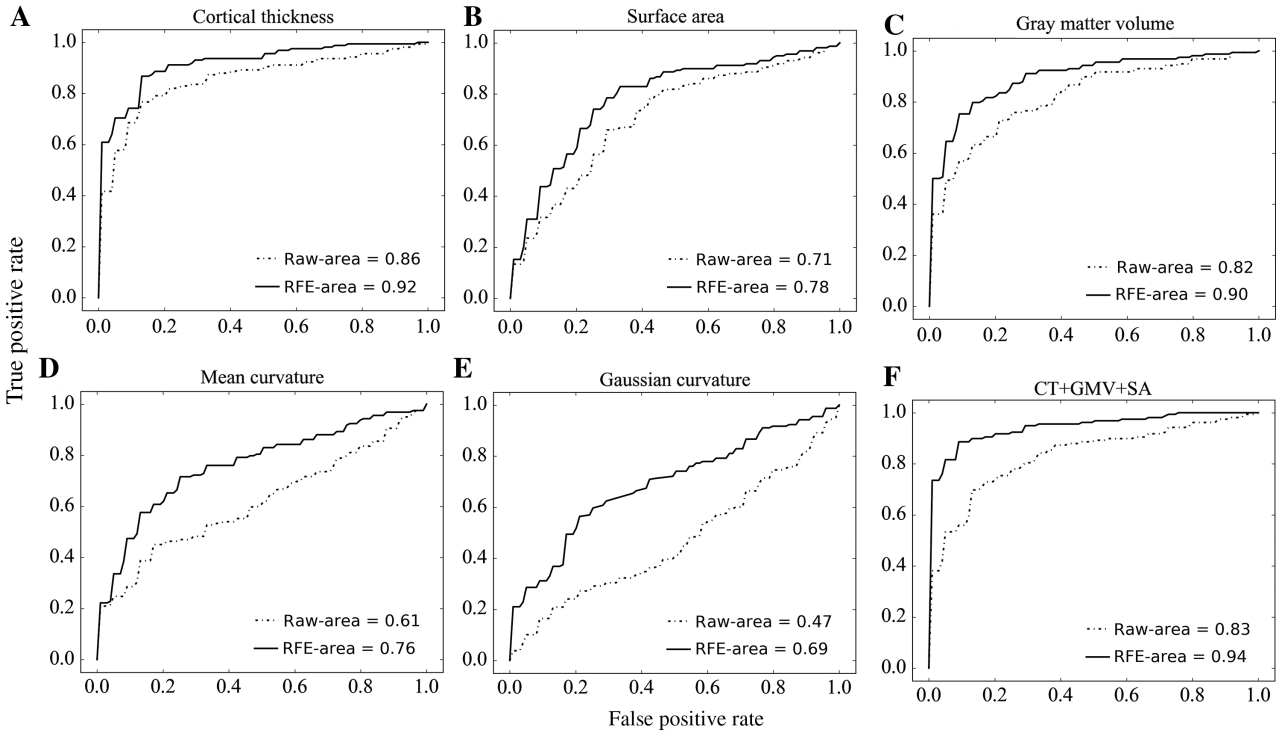

Figure 4: The ROC curves and AUCs for different kinds of parameters to evaluate classifier output quality using six-fold cross-validation. (A)–(F) represent the ROC curve and AUC for cortical thickness, surface area, gray matter volume, mean curvature, Gaussian curvature and CT + GMV + SA, respectively. The dot-dashed lines represent the raw ROC curve and the full lines represent the ROC with SVM-RFE method.

point – a FPR of zero, and a TPR of one. This is not very realistic, but it does mean that a larger AUC is usually better. Notably AUC = 0.5 stands for completely random predictions and AUC = 1.0 stands for perfect separation. The "steepness" of ROC curves is also important, since it is ideal to maximize the TPR while minimizing the FPR.

These ROC curves (Figure 4A–F) show the classifier performance for different kinds of parameters, computed from six-fold cross-validation. It can be found that all AUCs for each kind of parameter receive enhancement after the SVM-RFE method. As well as the classification accuracy in Table 3, multi-parameter combination obtains higher classifier performance than single parameter.

SVM classifier performance of two-dimensional space

For depicting the classification results more visually, SVM classifiers in two-dimensional spaces are used in separating HC from AD. The feature dimensions have been reduced through SVM-RFE method at the beginning. Therefore, the top two of optimal features of each kind of parameter can be used for separating HC from AD.

Figure 5 shows that cortical thickness and multi-parameter combination receive a maximum of classification effect, while for surface area and gray matter volume the classification effect does not seem obvious. The reason for these phenomena is the optimal feature dimensions of cortical thickness and multi-parameter combinations which are small enough to represent the main features, while the optimal feature dimensions of surface area and gray matter volume still remain large. Therefore, in order to furthest separate AD from HC, all the optimal features should be utilized simultaneously.

Discussion

In this study, we applied both the statistical analysis and pattern classification to compare AD with HC. The participants were firstly preprocessed using Freesurfer tool. Then the statistical analysis was applied in QDEC and the classification was conducted by SVM, respectively.

For the difference comparison of the cortical metrics, five kinds of parameters (i.e. cortical thickness, surface area, gray matter volume, curvature, sulcal depth) were compared between HC and AD. Compared with HC, the

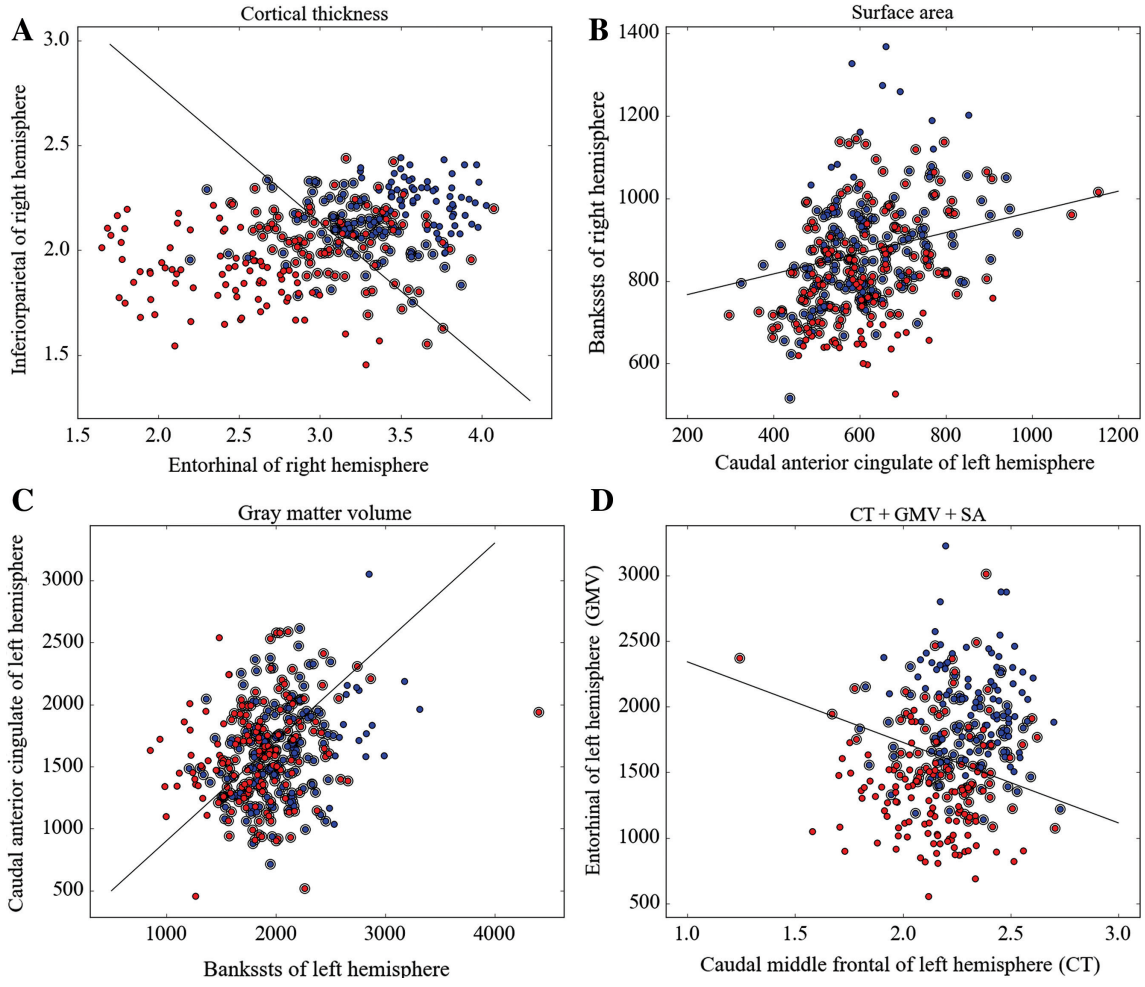

Figure 5: The dataset with the top two features for each kind of parameter in two-dimensional space using the SVM classifier with linear kernel.

The red color circles represent AD and the blue circles represent HC. (A) Right entorhinal versus right inferior parietal from cortical thickness. (B) Left caudal anterior cingulate versus right bankssts from surface area. (C) Left bankssts versus left caudal anterior cingulate from gray matter volume. (D) Left caudal middle frontal from cortical thickness versus left entorhinal from gray matter volume.

cortical thickness, surface area and gray matter volume of the entorhinal cortex are severely decreased in AD. The cortical thickness of the entorhinal cortex in the right hemisphere is thinner than that in the left hemisphere. These phenomena approximately agree with pre-existing literature involved in entorhinal cortex's atrophy [17, 28, 46]. About the difference research of cortical thickness, surface area and gray matter volume, the atrophy mainly appears in the temporal lobe, frontal lobe, occipital lobe, parietal lobe and cingulate gyrus. This phenomenon shows a strong agreement with previous findings related to the changes seen in HC and AD [18, 39]. These regions have been involved in complex cognitive behavior, motor execution, personality expression and decision making [50]. In addition, the attenuation of cortical thickness of left medial orbitofrontal reveals a cognitive inhibition related to the symptom of AD [29, 43]. Unlike pre-existing literature [16], there are no any atrophy of left entorhinal cortex in surface area. In the meantime, the surface areas of left precentral, left superior parietal and right insula show a small increase in AD. In the research of gray matter volume between HC and AD, apart from bilateral entorhinal cortex, the frontal lobe also has severe atrophy in AD, such as the left superior frontal and right rostral middle frontal, which agree with previous research [30, 41]. Previous research [6] had demonstrated that patients with AD showed an increasing degree of false recognition due to frontal lobe dysfunction. Our research on the atrophy of the frontal lobe is based on gray matter volume, while Salat's research [41] focused on white matter volume, they have some relevance to a certain extent. For the changes of sulcal shape (curvature in folded regions and

sulcal depth) between HC and AD, shallower sulcal depth mainly involved in left insula, bilateral cingulate gyrus for AD. While in other brain regions, the sulcal depths show different degrees of increase or decrease. These discoveries have some small deviations with previous research [31].

In the process of machine learning, the SVM-RFE technology is firstly used for the five kinds of parameters. Four kinds of single parameter and one kind of multi-parameter combination are used for constructing the classifiers. Among these parameters, multi-parameter combination yields the best classification result with an accuracy of 90.76%, a sensitivity of 89.87% sensitivity, a specificity of 91.72% specificity and an AUC of 0.94 after feature selection. The optimal feature dimensions for each kind of parameter rank from top to bottom are cortical thickness, Gaussian curvature, mean curvature, multi-parameter combination, gray matter volume and surface area in turn. The parameter with best improvement is multi-parameter combination whose accuracy increases from 74.25% to 90.76% via the SVM-RFE method. It shows that multi-parameter combination exhibit a better classification result than single parameter. In general, the classification performance receives favorable improvement through the SVM-RFE method. Combined with ROC curve and AUC, the performance of classifier receive improvement after feature selection. Moreover, the top two important features of different cortical parameters are used to build SVM classifier in a two-dimensional space to separate HC from AD. It is observed that cortical thickness or multi-parameter combination as a parameter could preferably separate HC from AD, which relates to the effect of feature selection. In other words, the fewer the significant features is, the more information the features carry.

Combining Freesurfer with QDEC tool, discrepant brain regions can be observed between groups. The discrepant brain regions can be contacted with previous research involving clinical manifestation. VBM just applies in single type of voxel-based comparison and the preprocessing steps are simple, while Freesurfer spends at least 10 h for each participant via numerous steps on reestablishing multi-type cerebral cortex. The existing literature with large sample analysis based on SPM [7, 32, 36] outdistance Freesurfer [25, 34, 40] so that it need to carry on more research based on Freesurfer. Beyond that, the majority of existing literature on classification research based on Freesurfer simply takes the cortical thickness and gray matter volume as classification parameters, this study even regards surface area, curvature and combined parameter as classification parameters. In addition, this study has proven the scikit-learn package could be successfully used in neuroimaging research. The articles of machine learning applied with scikit-learn for neuroimaging are still rare and just found in fMRI research [1, 35], not yet found in sMRI research. Taking SVM as a machine learning method, a high accuracy could be received via the SVM-RFE method. Combined with ROC curve, AUC and SVM classifier of two-dimensional space, classification performance could be shown vividly.

However, the achieved classification accuracy is still not optimal due to several factors. Firstly, the ADNI is a multicenter database (approximate 50 centers using different voxel sizes and acquisition parameters), and it does not take scanner or center effects into account. Next, potential brain vascular lesions in the participants may be a confounding factor. Finally, due to hundreds of images, the normalization quality of these MRI images cannot be manually validated. All of these aspects can influence the final classification accuracy.

Conclusion

This study applied cortical-based multiple parameters of MRI data to exhibit the differences of brain regions between groups. Then each kind of parameter for the data was trained by SVM algorithm to build the classification models for differentiating AD from HC. Via the SVM-RFE and LOOCV methods, the classification accuracy received a decent improvement. In addition, the ROC curve and AUC were used to verify the stability of constructed classifier model. The results indicated the key change in AD was the entorhinal cortex and medial orbitofrontal. The classification results revealed the multi-parameter combination could receive an accuracy of 90.76% and a better classification performance (AUC = 0.94). It proved that the classification performance of multi-parameter combination was superior to that of single parameter. Moreover, the classification performance of SVM classifiers in two-dimensional space indirectly related to the effect of SVM-RFE. This research reflects a promising advance for diagnostic image analysis based on machine learning. In the future, we will extend this work to a larger clinical data and include multiple modalities or more advanced machine learning techniques to further improve the classification accuracy.

Acknowledgements: This work was supported by the National Natural Science Foundation of China under Grant Nos. 11572127 and 11172103.

References

[1] Abraham A, Pedregosa F, Eickenberg M, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014; 8: 14.

[2] Aguilar C, Westman E, Muehlboeck JS, et al. Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. Psychiatry Res 2013; 212: 89–98.

[3] Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. NeuroImage 2016; 145(Pt B): 137.

[4] Ashburner J, Friston KJ. Voxel-based morphometry—the methods. NeuroImage 2000; 11: 805–821.

[5] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. Comput Learn Theory 1992: 144–152.

[6] Budson AE, Sullivan AL, Mayer E. Suppression of false recognition in Alzheimer's disease and in patients with frontal lobe lesions. Brain 2002; 125: 2750–2765.

[7] Casanova R, Hsu FC, Espeland MA; Alzheimer's disease Neuroimaging I. Classification of structural MRI images in Alzheimer's disease from the perspective of ill-posed problems. PLoS One 2012; 7: e44877.

[8] Challis E, Hurley P, Serra L, Bozzali M, Oliver S, Cercignani M. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. NeuroImage 2015; 112: 232–243.

[9] Cherbuin N, Shaw M, Sachdev PS, Anstey KJ. Validated dementia risk factor composite is associated with lower hippocampal volumes and cortical thickness. Alzheimers Dement 2015; 11: P813–P814.

[10] Cohn-Sheehy B, Ghosh P, Wirth M, Lehmann M, Madison C, Irwin W, et al. Temporoparietal cortical thickness outperforms hippocampal volume as a biomarker for atypical and early-onset Alzheimer's disease. Alzheimers Dement 2013; 9: P41–P42.

[11] Corbo V, Salat DH, Powell MA, Milberg WP, McGlinchey RE. Combat exposure is associated with cortical thickness in Veterans with a history of chronic pain. Psychiatry Res 2016; 249: 38–44.

[12] Craddock RC, Holtzheimer PE, 3rd, Hu XP, Mayberg HS. Disease state prediction from resting state functional connectivity. Magn Reson Med 2009; 62: 1619–1628.

[13] De Martino F, Valente G, Staeren N, Ashburner J, Goebel R, Formisano E. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 2008; 43: 44–58.

[14] Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 2006; 31: 968–980.

[15] Dhikav V, Duraisamy S, Anand KS, Garga UC. Hippocampal volumes among older Indian adults: Comparison with Alzheimer's disease and mild cognitive impairment. Ann Indian Acad Neurol 2016; 19: 195–200.

[16] Dickerson BC, Feczko E, Augustinack JC, et al. Differential effects of aging and Alzheimer's disease on medial temporal lobe cortical thickness and surface area. Neurobiol Aging 2009; 30: 432–440.

[17] Du AT, Schuff N, Amend D. Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. J Neurol Neurosurg Psychiatry 2001; 71: 441–447.

[18] Du AT, Schuff N, Kramer JH, et al. Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain 2007; 130: 1159–1166.

[19] Farahat AK, Ghodsi A, Kamel MS. An efficient greedy method for unsupervised feature selection. IEEE International Conference on Data Mining, ICDM 2011, Vancouver, BC, Canada, December DBLP. 2011: 160–170.

[20] Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006; 27: 861–874.

[21] Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. The Lancet 2005; 366: 2112–2117.

[22] Fischl B. FreeSurfer. NeuroImage 2012; 62: 774–781.

[23] Fjell AM, Walhovd KB, Fennema-Notestine C, et al. CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. J Neurosci 2010; 30: 2088–2101.

[24] Friston KJ, Holmes AP, Worsley KJ. Statistical parametric maps in functional imaging: a general linear approach. Human Brain Mapp 1994; 2: 189–210.

[25] Goryawala M, Zhou Q, Barker W, Loewenstein DA, Duara R, Adjouadi M. Inclusion of neuropsychological scores in atrophy models improves diagnostic classification of Alzheimer's disease and mild cognitive impairment. Comput Intell Neurosci 2015; 2015: 56.

[26] Guo X, Li Z, Chen K, Yao L, Wang Z, Li K. Mapping gray matter volume and cortical thickness in Alzheimer's disease. Proc SPIE Int Soc Opt Eng 2010; 7626: 76260B1–76260B9.

[27] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Mach Learn 2002; 46: 34.

[28] Hoesen GWV, Hyman BT, Damasio AR. Entorhinal cortex pathology in Alzheimer's disease. Hippocampus 1991; 1: 1–8.

[29] Hoesen GWV, Parvizi J, Chu CC. Orbitofrontal cortex pathology in Alzheimer's disease. Cereb Cortex 2000; 10: 243–251.

[30] Ikonomovic MD, Abrahamson EE, Isanski BA. Superior frontal cortex cholinergic axon density in mild cognitive impairment and early Alzheimer disease. Arch Neurol 2007; 64: 1312–1317.

[31] Im K, Lee JM, Seo SW, et al. Sulcal morphology changes and their relationship with cortical thickness and gyral white matter volume in mild cognitive impairment and Alzheimer's disease. NeuroImage 2008; 43: 103–113.

[32] Kloppel S, Peter J, Ludl A, et al. Applying automated MR-based diagnostic methods to the memory clinic: a prospective study. J Alzheimers Dis 2015; 47: 939–954.

[33] Li M, Qin Y, Gao F, Zhu W, He X. Discriminative analysis of multivariate features from structural MRI and diffusion tensor images. Magn Reson Imaging 2014; 32: 1043–1051.

[34] Lillemark L, Sørensen L, Pai A, Dam EB, Nielsen M. Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. BMC Med Imaging 2014; 14: 1–12.

[35] Michel V, Gramfort A, Varoquaux G, Eger E, Keribin C, Thirion B. A supervised clustering approach for fMRI-based inference of brain states. Pattern Recognit 2012; 45: 2041–2049.

[36] Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J. Alzheimer's disease neuroimaging I. Machine learning framework for early

MRI-based Alzheimer's conversion prediction in MCI subjects. NeuroImage 2015; 104: 398–412.
[37] Nir TM, Villalon-Reina JE, Prasad G, et al. Diffusion weighted imaging-based maximum density path analysis and classification of Alzheimer's disease. Neurobiol Aging 2015; 36 Suppl 1: S132–S140.
[38] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2011; 12: 6.
[39] Richards BA, Chertkow H, Singh V, et al. Patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Neurobiol Aging 2009; 30: 1626–1636.
[40] Sabuncu MR, Leemput KV. The relevance voxel machine (RVoxM): a self-tuning Bayesian model for informative image-based prediction. IEEE Trans Med Imaging 2012; 31: 2290–2306.
[41] Salat DH, Greve DN, Pacheco JL, et al. Regional white matter volume differences in nondemented aging and Alzheimer's disease. NeuroImage 2009; 44: 1247–1258.
[42] Simmons A, Westman E, Zhang Y, et al. Multivariate data analysis of regional MRI volumes and cortical thickness measures to distinguish between Alzheimer's disease, mild cognitive impairment and healthy controls. Alzheimers Dement 2009; 5: e16.
[43] Szatkowska I, Szymanska O, Bojarski P, Grabowska A. Cognitive inhibition in patients with medial orbitofrontal damage. Exp Brain Res 2007; 181: 109–115.
[44] Tessitore A, Santangelo G, De Micco R, et al. Cortical thickness changes in patients with Parkinson's disease and impulse control disorders. Parkinsonism Relat Disord 2016; 24: 119–125.
[45] Tierney MC, Yao C, Kiss A, McDowell I. Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. Neurology 2005; 64: 1853–1859.
[46] Velayudhan L, Proitsi P, Westman E, et al. Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. J Alzheimers Dis 2013; 33: 755–766.
[47] Visser PJ, Verhey FR, Hofman PA, Scheltens P, Jolles J. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. J Neurol Neurosurg Psychiatry 2002; 72: 491–497.
[48] Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process, and purpose. Am Stat 2016; 70: 129–133.
[49] Wee CY, Yap PT, Shen D, Alzheimer's disease neuroimaging I. Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. Hum Brain Mapp 2013; 34: 3411–3425.
[50] Yang Y, Raine A. Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. Psychiatry Res 2009; 174: 81–88.
[51] Zhang Y, Dong Z, Phillips P, et al. Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Front Comput Neurosci 2015; 9: 66.
[52] Zhao M, Zhao C, Zheng C. Identifying concealed information using wavelet feature extraction and support vector machine. Procedia Environ Sci 2011; 8: 337–343.
[53] Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993; 39: 17.