

Identifying Mild Cognitive Impairment with Random Forest by Integrating Multiple MRI Morphological Metrics

Zhe Ma^{a,b,1}, Bin Jing^{a,b,1}, Yuxia Li^c, Huagang Yan^{a,b}, Zhaoxia Li^d, Xiangyu Ma^{a,b}, Zhizheng Zhuo^{a,b}, Lijiang Wei^{a,b} and Haiyun Li^{a,b,*} for the Alzheimer's Disease Neuroimaging Initiative²

^a*School of Biomedical Engineering, Capital Medical University, Beijing, China*

^b*Beijing Key Laboratory of Fundamental Research on Biomechanics in Clinical Application, Capital Medical University, Beijing, China*

^c*Department of Neurology, XuanWu Hospital of Capital Medical University, Beijing, China*

^d*School of Chinese Medicine, Capital Medical University, Beijing, China*

Handling Associate Editor: Tianyi Yan

Accepted 15 November 2019

Abstract. Mild cognitive impairment (MCI) exhibits a high risk of progression to Alzheimer's disease (AD), and it is commonly deemed as the precursor of AD. It is important to find effective and robust ways for the early diagnosis of MCI. In this paper, a random forest-based method combining multiple morphological metrics was proposed to identify MCI from normal controls (NC). Voxel-based morphometry, deformation-based morphometry, and surface-based morphometry were utilized to extract morphological metrics such as gray matter volume, Jacobian determinant value, cortical thickness, gyrification index, sulcus depth, and fractal dimension. An initial discovery dataset (56 MCI/55 NC) from the ADNI were used to construct classification models and the performances were testified with 10-fold cross validation. To test the generalization of the proposed method, two extra validation datasets including longitudinal ADNI data (30 MCI/16 NC) and collected data from Xuanwu Hospital (27 MCI/32 NC) were employed respectively to evaluate the performance. No matter whether testing was done on the discovery dataset or the extra validation datasets, the accuracies were about 80% with the combined morphological metrics, which were significantly superior to single metric (accuracy: 45%~76%) and also displayed good generalization across datasets. Additionally, gyrification index and cortical thickness derived from surface-based morphometry outperformed other features in MCI identification, suggesting they were some key morphological biomarkers for early MCI diagnosis. Combining the multiple morphological metrics together resulted in a significantly better and reliable identification model, which may be helpful to assist in the clinical diagnosis of MCI.

Keywords: Deformation-based morphometry, mild cognitive impairment, random forest, surface-based morphometry, voxel-based morphometry

¹These authors contributed equally to this work.

²Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate

in analysis or writing of this report. A complete list of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

*Correspondence to: Prof. Haiyun Li, School of Biomedical Engineering, Capital Medical University, Beijing 100069, China. Tel.: +86 13520973464; E-mail: haiyunli@ccmu.edu.cn.

INTRODUCTION

Mild cognitive impairment (MCI) is a syndrome with cognitive decline more serious than normal aging but not severe enough to cause notable impairments of daily function, which is generally regarded as the prodromal stage of Alzheimer's disease (AD). The early diagnosis and intervention for cognitive decline is crucial to postpone progression to AD, so MCI identification has been a persistent research focus for decades. Brain atrophy is a remarkable sign of neurodegeneration as measured by structural magnetic resonance images (sMRI) [1], and different morphological analysis methods have been used to investigate subtle structural alterations in MCI with sMRI [2, 3].

The common morphological analysis methods in sMRI include voxel-based morphometry (VBM) [4], deformation-based morphometry (DBM) [5], and surface-based morphometry (SBM) [6]. VBM is a popular morphometry analysis method which could provide voxel-wise volume/density estimations of segmented gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). DBM relies on the deformation descriptor generated from the spatial registration to reflect the whole brain structural changes. SBM measures types of cortical surface properties in a vertex-based manner. Different metrics could be extracted using these morphological analysis methods, representing distinctive structural properties of cerebral cortex. Scanlon et al. [7] used these morphological analysis methods to detect brain structural changes related with temporal lobe epilepsy, and found three methods could reveal different aspects of brain atrophy, which implied that integration of these three methods may further help in the detection of neurodegenerative diseases.

So far, several studies have adopted various morphological metrics to investigate MCI and AD patients. Schmitter et al. [8] used the volume of hippocampus extracted by VBM as the classification feature to discriminate MCI from normal controls (NC) and achieved a 71% detection rate. Koikkalainen et al. [9] utilized multi-template DBM to analyze sMRI images of MCI and the classification accuracy for stable versus progressive MCI subjects was 72.1%. Park et al. [10] computed two cortical features including cortical thickness (CTH) and sulcus depth (SD) with SBM and an accuracy of 86% was achieved in MCI identification. Madan et al. [11] demonstrated fractal dimension (FD) extracted

from SBM was a sensitive index for detecting age-related cortical folding changes due to its ability to characterize tiny morphometric deformations. However, single morphological feature may be insufficient for MCI identification because it may only reflect specific morphological abnormalities. Several studies used multiple morphological features to recognize MCI, which could provide comprehensive information about the complex structural changes in MCI. Bron et al. [12] found the best performance in an MCI prediction challenge was achieved using a combination of features including volume, CTH, shape, and intensity, and the best algorithm yielded an accuracy of 63%. Liu et al. [13] used thickness and volume of selected brain regions to differentiate MCI from NC, and obtained an 82% accuracy. Our prior studies [14, 15] demonstrated that integrating multiple features could improve the MCI classification accuracy. Above all, the classification accuracy varied largely in different studies, which emphasized the importance of the generalization in MCI classification. Ideally, a good prediction model should be able to perform well regardless of the input data; however, there are few studies to adopt more than one cohort for model validation. In addition, previous studies mainly focused on gray matter volume (GMV) and CTH of MCI patients, and there are only a few studies to use other structural metrics [16, 17]. It is still unknown whether the combination of morphological metrics obtained from VBM, DBM, and SBM could further improve the MCI classification and the prediction generalization.

Taken together, we speculate that metrics extracted from multiple morphological analysis methods would outperform those from single morphological analysis method in MCI identification, and the prediction generalization with multiple morphological metrics would also perform better than the one with single morphological metric. In this study, VBM, DBM, and SBM analyses were conducted to calculate the morphological metrics including GMV, Jacobian determinant value (JDV), CTH, gyrification index (GI), SD, and FD, and these features were selected by statistical analysis and served as classification features for random forest (RF) classifier to identify MCI with 10-fold cross validation. The classifier performances were finally validated by two extra datasets including a longitudinal Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset and a clinically collected dataset from Xuanwu Hospital.

MATERIALS AND METHODS

Participants

A total of 216 subjects were included in this study, including 111 subjects (56 MCI and 55 NC) in the discovery ADNI dataset, 46 subjects in the longitudinal ADNI validation dataset, and 59 subjects in Xuanwu validation dataset. The inclusion criteria for ADNI data were as follows: 1) normal subjects: the Mini-Mental State Examination (MMSE) scores between 26 and 30, Clinical Dementia Rating less than or equal to 0.5, non-MCI, and non-demented; 2) MCI subjects: objective memory loss measured by education adjusted score of Wechsler Memory Scale Logical Memory and a memory complaint, essentially preserved daily living function, and non-demented; 3) All subjects were scanned using magnetization prepared rapid gradient echo sequence by a 3 Tesla MRI system. The exclusion criteria were as follows: 1) MCI and NC without structural MRI scans and above-mentioned clinical information; 2) Poor image quality. In addition, the discovery dataset only contained the subjects scanned at baseline (first time), while the validation ADNI dataset were the longitudinal data of persons who had follow-up scanning in the discovery dataset. The other validation dataset were clinically collected patients at the clinic of the Department of Neurology, Xuanwu Hospital, and the health controls were recruited from the local community. Diagnoses of MCI due to AD were made by experienced neurologists using Petersen's criteria [18], and more inclusion/exclusion criteria were given in [19]. The study was approved by the Research Ethics Review Board of Xuanwu Hospital. All NC matched well with the MCI patients on age and gender. The detailed demographic information for all subjects is shown in Table 1.

Image acquisition

Structural images from ADNI were acquired with the following parameters: slice thickness = 1.2 mm, TE = 2.95 ms, TR = 2.3 s, TI = 900 ms, flip angle = 9°, slices = 176, and voxel size = 1 × 1 × 1.2 mm³. Data from Xuanwu hospital were obtained on a 3T Magnetom Trio Tim scanner (Siemens, Erlangen, Germany), and the parameters were set as following: thickness = 1 mm, TE = 2.2 ms, TR = 1.9 s, matrix = 448 × 512, slices = 176, and voxel size = 0.5 × 0.5 × 1 mm³.

Calculation of multiple morphological metrics

Data were preprocessed with the Computational Anatomy Toolbox (CAT12.3-r1317, <http://www.neuro.uni-jena.de/cat/>) ran under Statistical Parametric Mapping, Version 12 (SPM12, <http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). Notably, we took a two-step quality assurance: all raw images were visually inspected for artifacts and all segmented images were statistically controlled for inter-subject homogeneity. Previous literature showed that neurodegenerative diseases were mainly related to GM, therefore, three structural analysis methods were only conducted on the GM [20].

VBM could identify differences in local brain region in a voxel-wise manner. After segmentation, GM images were used and normalized to the MNI standard space with DARTEL algorithm. 'Modulation' was applied in the normalization step to preserve the volume of GM and all GM images were resampled to 1.5 × 1.5 × 1.5 mm³. Finally, the modulated GM images were smoothed with an isotropic 8 mm full width half maximum (FWHM) Gaussian kernel and the GMV was finally computed from the smoothed images.

Table 1
Demographic and neuropsychological information for all subjects

	Discovery dataset			Validation dataset 1 (longitudinal ADNI)			Validation dataset 2 (Xuanwu data)		
	MCI	NC	<i>p</i>	MCI	NC	<i>p</i>	MCI	NC	<i>p</i>
Sample size	56	55	—	30	16	—	27	32	—
Age (y)	74.6 ± 7.4	75.2 ± 6.7	0.669 *	74.4 ± 8.4	74.3 ± 5.8	0.070 *	67.4 ± 8.5	64.9 ± 7.5	0.223 *
Gender (M/F)	30 / 26	27 / 28	0.706 #	16 / 14	5/11	0.152 #	13 / 14	16 / 16	0.887 #
MMSE	25.1 ± 4.1	28.9 ± 1.2	<0.001~	24.7 ± 3.2	28.8 ± 1.1	<0.001~	23.5 ± 3.3	27.7 ± 1.7	<0.001~
CDR	0.5 ± 0.3	0.1 ± 0.2	<0.001~	0.5 ± 0.1	0	<0.001~	0.5 ± 0	0	<0.001~

M, male; F, female; MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating. * Two sample two-tailed *t* test, # Chi-square test, ~Mann-Whitney U test, *p* < 0.05.

DBM focuses on local deformations generated during the process of non-linear spatial registration, which was quantified by the Jacobian determinant. For each voxel registering from subject's image to reference image, the JDV greater than 1 represents an expansion of the voxel and below 1 indicates a compression of the voxel [21]. The resulting JDV maps were masked for GM and smoothed with an isotropic 8-mm FWHM Gaussian kernel.

SBM could measure cortical surface properties in a vertex-wise manner. CTH was calculated by adopting a fully automated projection-based thickness measurement method, which was different from other thickness measurement methods [22]. Other surface parameters including SD (based on the Euclidean distance between the central surface and its convex hull), GI (based on absolute mean curvature), and FD (based on spherical harmonics [23]) were also extracted. In addition, to enhance the statistical power of features, CTH was smoothed with a Gaussian kernel of FWHM 15 mm, and other three surface parameters were smoothed with an isotropic 25 mm FWHM Gaussian kernel [24].

The flowchart of the three morphological analysis methods was shown in Fig. 1.

Statistical analysis

Voxel/vertex-wise two-sample two-tailed *t*-test was performed for each structural metric, with age and gender as covariates, to determine the differences

between MCI patients and NC. Additionally, total intracranial volume was taken as a covariate to correct for individual head size differences in the VBM analysis [25]. For GMV, JDV, and CTH, the *p*-values were corrected using family wise error (FWE) with a threshold of 0.05. For GI, SD, and FD, we applied a statistical threshold of $p = 0.001$ (uncorrected). Of note, once used in the classifier training, statistical analysis was only conducted in the training data, which could decrease the possible overfitting in the subsequent classification.

Random forest-based classification

RF is an ensemble of decision tree classifiers developed by Breiman, and many decision trees are built using randomized feature subset sampling and bagging [26]. The RF has important advantages in term of robustness to avoid overfitting, to handle highly non-linear data, and to conduct efficient parallel processing when applied on neuroimaging data [27, 28]. In order to construct the classification classifiers of single or combined morphological features, a RF procedure was carried out with random forests package (<https://cran.r-project.org/web/packages/randomForest>). The features were input separately into the RF classifiers with $n_{tree} = 100$ and $m_{try} = 2$ (n_{tree} indicates the number of trees and m_{try} represents the number of predictors sampled for splitting at each node). The classification performance was quantified by means of accuracy,

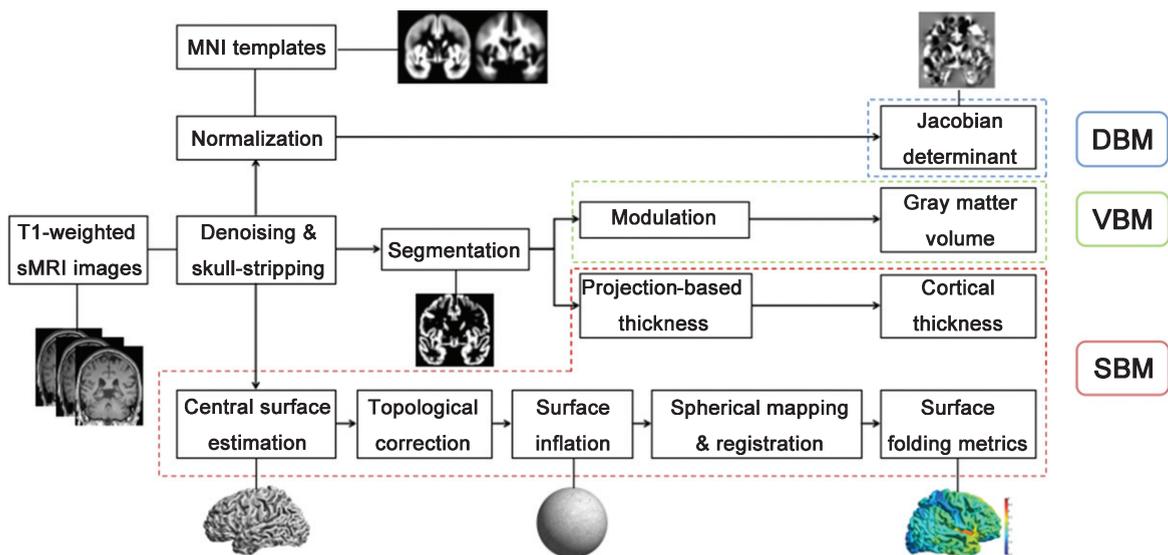


Fig. 1. The flowchart of the three morphological analysis methods.

which is defined as the percentage of correct identification cases with the optimal RF model. Considering the randomness of the RF performance, the ultimate result was assessed by the average accuracy of ten trials. Additionally, the mean decrease in accuracy and mean reduction in Gini index were respectively used to access the feature importance in classification. The mean decrease in accuracy is defined as the decrease of accuracy when a feature is changed into random numbers. The larger the value, the more importance the feature is. The Gini index is based on the principle of impurity reduction, and a greater decrease in Gini index means a higher importance of the feature. At last, the prediction model used for the extra validation datasets was slightly different in the training features compared with the model for the discovery dataset. The training features used in discovery dataset were generated from statistical differences in the training dataset (9 fold) while the training features adopted in the extra validation datasets were statistical differences in the whole discovery dataset.

The flowchart of our proposed method for MCI identification was shown in Fig. 2.

RESULTS

The between-group morphological metrics differences in the whole discovery dataset were summarized here. Using VBM, two large clusters with GMV loss were observed at bilateral temporal gyrus, hippocampus (HIP), parahippocampal gyrus (PHG), entorhinal cortex (ENT), and amygdala (AMYG). At similar locations, abnormal-

ities were also detected in JDV using DBM. MCI patients showed significantly thinner CTH in bilateral parietal, frontal, temporal, supramarginal, and left precentral gyrus than NC. If using FWE correction, there was no significant inter-group difference for three surface parameters (SD, GI, and FD), therefore, significant differences were set at $p < 0.001$ (uncorrected). The detected regions were illustrated in Fig. 3, and the detailed quantitative descriptions were shown in Table 2.

Table 3 listed the set of accuracy, sensitivity, and specificity for classification with different metrics. The classification performances of single or combined morphological features on extra validation sets were summarized in Fig. 4. In addition, the feature importance measured by mean decreased accuracy and mean reduction of Gini index were respectively shown in Fig. 5.

DISCUSSION

In this study, we proposed an MCI identification method which incorporated multiple morphological metrics. Multiple metrics like GMV, JDV, CTH, SD, GI, and FD were extracted by VBM, DBM, or SBM respectively, and RF was applied to quantify the classification performances of different combined features. No matter whether testing was done using the 10-fold cross validation on the discovery set or on two extra validation sets, nearly 80% accuracy was achieved, indicating that our method was robust and potentially valuable in distinguishing MCI from NC.

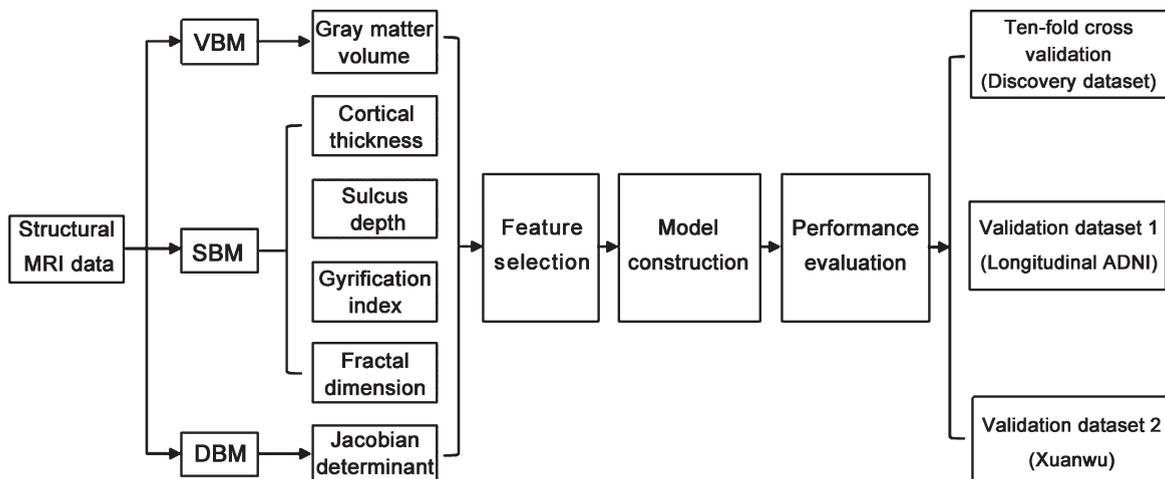


Fig. 2. The flowchart of our proposed method for MCI identification.

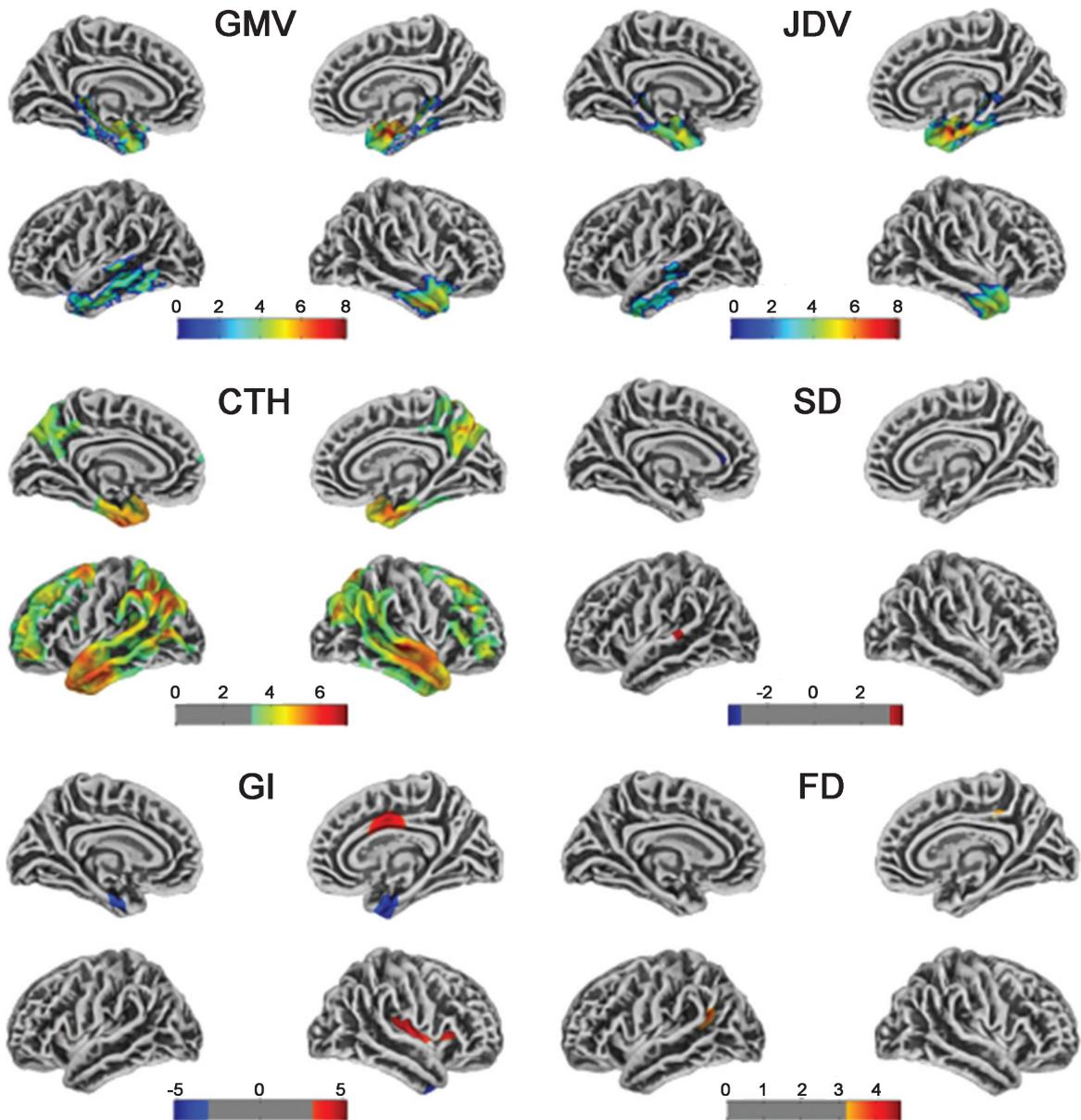


Fig. 3. T-value maps exhibiting significant differences with GMV, JDV, and CTH ($p < 0.05$, FWE correction) and SD, GI, and FD ($p < 0.001$, uncorrected). The green/red shades represent areas where the MCI group has lower measured values than NC group, whereas the blue regions represent areas where MCI group has higher value than the NC group.

In general, VBM, DBM, and SBM are widely used to detect the neuroanatomical changes in human brain. Among them, VBM and DBM are appropriate approaches for measuring alterations in subcortical regions. Our results found similar abnormal regions were detected in MCI patients by using both methods. VBM detected larger clusters with GMV loss while DBM detected smaller clusters with deformable changes (Table 2). One possible reason for the similar findings is that GMV is computed by multiply-

ing GM density with the Jacobian determinant at each voxel, which indicates an intrinsic association between each other [29]. However, JDV extracted from DBM performed better in classification compared with GMV from VBM (72% versus 61%). This may be due to some discrepancies between the two methods. First, VBM estimated the volume information from the segmented images (GM/WM/CSF), while DBM extracted the deformation information from the whole brain images. Second, DBM utilized

Table 2
The brain regions showing significant differences between MCI and NC

Method	Feature	No.	Brain regions	Cluster size	T value	
VBM	GMV	1	L, MT/IT/HIP/TMP/ENT/PHIP/AMYG	12876	6.52	
		2	R, TMP/HIP/PHIP/MT/IT/ENT/AMYG	12359	7.45	
DBM	JDV	3	L, HIP/MTG/TMP/PHIP/ENT/IT/AMYG	6033	5.91	
		4	R, TMP/HIP/ENT/IT/PHIP/MTG/AMYG	7966	7.78	
SBM	CTH	5	L, SP/IP/ST/SMG/MT/IT	39091	6.20	
		6	L, MF/SF	12488	5.40	
		7	L, PC	1523	4.00	
		8	R, IP/ST/SP/PC/MT/SMG/IT	39586	6.50	
		9	R, MF/SF	11649	5.30	
		SD	10	L, ST	222	3.40
			11	L, ACC	118	-3.30
		GI	12	L, ST	307	3.50
			13	L, INS	92	3.30
	14		L, ENT	555	-3.90	
	15		R, INS/ST	4677	4.60	
	16		R, PCC/ACC	1664	4.30	
	17		R, ENT	1219	-4.70	
	FD		18	L, IP/SMG/ST	1454	4.20
			19	R, PCC/PRE	279	3.50

L, left; R, right; GMV, gray matter volume; JDV, Jacobian determinant value; CTH, cortical thickness; SD, sulcus depth; GI, gyrification index; FD, fractal dimension; TMP, temporal pole; HIP, hippocampus; PHIP, parahippocampal gyrus; ENT, entorhinal cortex; AMYG, amygdala; SP, superior parietal gyrus; IP, inferior parietal gyrus; ST, superior temporal gyrus; SMG, supramarginal gyrus; MT, middle temporal gyrus; PRE, precuneus; IT, inferior temporal gyrus; MF, middle frontal gyrus; SF, superior frontal gyrus; PC, precentral gyrus; ACC, anterior cingulate; INS, insula; PCC, posterior cingulate.

Table 3
Classification performance of single or combined metrics using ten-fold cross validation
(Results were reported as mean with standard deviation in brackets)

Method	Feature	Accuracy (Mean \pm std.)	Sensitivity (Mean \pm std.)	Specificity (Mean \pm std.)
VBM	GMV	0.61 (0.14)	0.60 (0.21)	0.69 (0.12)
DBM	JDV	0.72 (0.10)	0.69 (0.17)	0.80 (0.14)
SBM	CTH	0.74 (0.13)	0.67 (0.18)	0.80 (0.17)
	SD	0.65 (0.10)	0.60 (0.19)	0.71 (0.15)
	GI	0.76 (0.15)	0.85 (0.13)	0.72 (0.21)
	FD	0.58 (0.10)	0.60 (0.13)	0.57 (0.27)
	Combined	All	0.80 (0.11)	0.75 (0.20)

GMV, gray matter volume; CTH, cortical thickness; JDV, Jacobian determinant value; SD, sulcus depth; GI, gyrification index; FD, fractal dimension.

DARTEL and Geodesic shooting normalization algorithms for registration, thus it is reasonably expected that DBM could detect subtle subcortical changes over VBM due to the superior registration [30, 31]. In contrast, SBM is suitable to detect cortical abnormalities through multiple metrics of cortical surface. An accuracy of 77% was achieved in identifying MCI from NC using four types of surface features (CTH, GI, SD, and FD), which was obviously superior to that from VBM and DBM, suggesting the combination of surface features could improve the classification performance. However, SBM has a potential drawback during the surface reconstruction process, which may artificially inflate surface areas that not perfectly match the underlying anatomy [24].

Regarding the classification accuracy, it varied across different morphological metrics (Table 3). We found GI provided the best accuracy of 76%, closely followed by CTH. GI is used as a measure of surface complexity and the cortical region with larger folding has a large GI, whereas the region with limited folding has a small GI. The cortical folding pattern is determined by genetic and early developmental factors and will keep stable throughout adulthood in healthy populations, implying that GI with deviation from normal populations has a high probability of brain abnormalities [32]. However, De Miras et al. [33] computed the GI according to the conventional method, which was defined as the ratio between the inner contour and the outer hull, and found MCI displayed no sig-

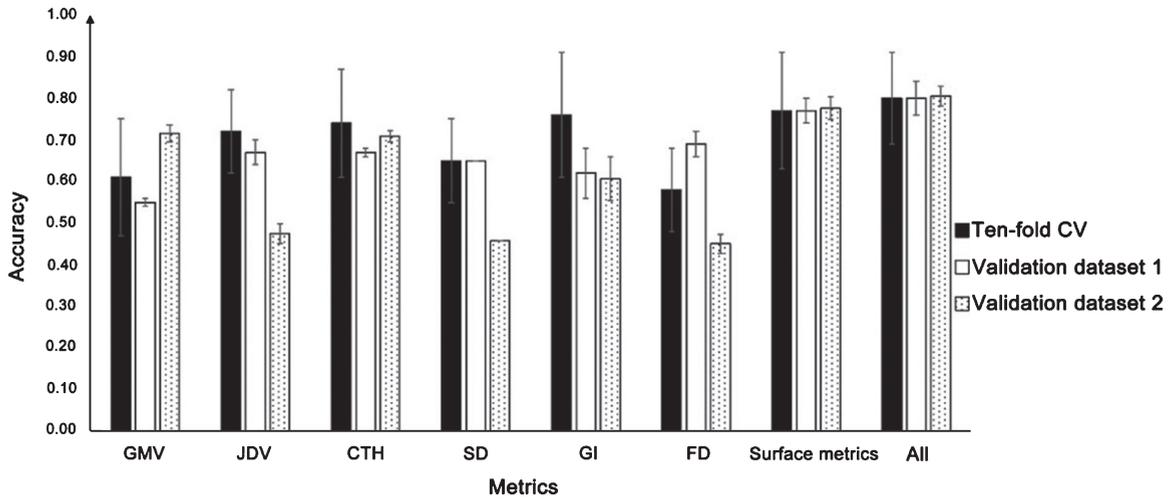


Fig. 4. Classification performance of the combinations of features extracted from different morphological analysis methods on different datasets.

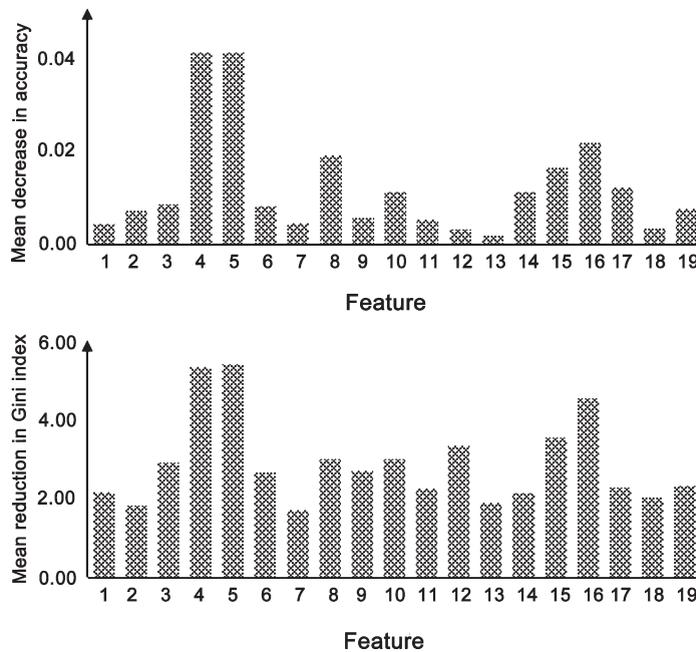


Fig. 5. The feature importance for classification using mean decrease in accuracy and mean reduction in Gini index as measures respectively (The brain regions corresponding to the feature ID could refer to Table 2).

nificant reductions of GI compared to NC. In our study, GI was defined via absolute mean curvature and was found with obvious differences between MCI and NC. Additionally, different normalization algorithms may also lead to differences in GI numerical calculation across studies. Taken together, the selection of the calculation method has a great influence on GI, which should be paid more attention in future studies.

The classification accuracies of single metrics varied across validation datasets largely, demonstrating the weak generalization ability (Fig. 4). However, using multiple metrics, the accuracies were almost consistent in three datasets, indicating the method we established was stable and reliable. Furthermore, combination of the features extracted from different morphological analyses could improve classification performances in discriminating MCI than single met-

ric, implying that different morphological metrics had complementary and specific contributions to the recognition accuracy. Besides using RF to classify MCI, we also established classification models with SVM. Though high accuracy (89.2%) was obtained for the discovery dataset, poor performance (accuracy of 65.2%) was found on the external validation datasets, demonstrating this model was likely to be overfitting. The results showed RF possessed a good between-cohort robustness compared to SVM, which was consistent with Lebedev's work [34].

With regard to the importance of each feature, similar results were discovered by using mean decrease in accuracy and the Gini index (Fig. 5). The GI of insula was found with weakest importance in MCI identification and lowest T values in the between group comparison. The insula, anterior cingulate, amygdala, and other subcortical structures are referred to as the salience network, which displays abnormalities in a wide range of neurological disorders, suggesting that insula has important roles in cognition. However, there is no consensus about the alterations of insula in MCI [35, 36], and future studies are needed to ascertain the role of salience network in MCI patients.

It was worth noting that we used different multiple comparison correction methods for the morphological metrics, which was FWE correction for GMV, JDV, and CTH, and uncorrected for SD, GI, and FD. So far, the selection of correction method in morphological analysis varies across studies [37, 38]. Westman et al. [39] identified MCI volume abnormalities in multiple brain regions, and the p -values were adjusted for multiple comparisons with Bonferroni correction with a significance threshold of $p < 0.05$. In Beheshti et al. study [40], significance level was set at $p < 0.01$ with FWE correction, and the cluster size threshold was set at 1400 voxels to report the VBM findings. Using DBM, Hua et al. [41] created Jacobian maps of MCI and NC respectively, and tested the overall significance of group differences using permutation tests for multiple comparisons. Maier et al. [42] investigated the cortical properties (including CTH, GI, and SD) of adults with autism spectrum disorder, and only uncorrected p -values exhibited significant group differences in region of interest (ROI)-based statistics. Taken together, for the volume and thickness metrics, rigorous correction method was used, while for other geometric metrics, loose correction method was used, which was probably due to complicated spatial patterns of geometric metrics. Moreover, the classification performance of uncorrected features was comparable to that of cor-

rected features in our study, thus, we inferred that it was not mandatory to use strict correction method for geometric metrics in a machine learning study.

Compared with NC, MCI patients showed significant lower values in GMV and JDV in the bilateral temporal gyrus, HIP and PHG, ENT and AMYG (Table 2), which also displayed high discriminative abilities for MCI identification according to the feature importance indices (Fig. 5). The GMV reduction in medial temporal gyrus is a typical trait of MCI [43, 44]. HIP atrophy is recognized as one of the most effective biomarkers for AD pathological progression [45]. ENT provides an interface between the HIP and neocortex, where early neurofibrillary tangles and tau protein are deposited [46–48]. The neurofibrillary tangles and amyloid plaques could result in the loss of neurons in AMYG, who has abundant neural connections with the HIP [49]. These brain regions are strongly implicated in memory functions and are widely used to study the pathological mechanism of MCI and AD [50].

Our results showed significant CTH thinning in some temporal, parietal, and frontal regions, which also showed high discriminative powers reflected by feature importance. These regions are parts of the widely studied default mode network (DMN), and a previous study reported that the activities of DMN diminished in MCI patients [51]. To date, some studies have examined cortical surface folding properties in MCI, which are linked with the progression of neuronal connections and cortical connectivity pattern. In our study, significant inter-group differences in surface properties (including SD, GI, and FD) were observed. Our results were similar to findings obtained by Li et al., who reported that MCI showed shallower SD and lower GI in the temporal gyrus. However, increased GI in MCI group was found in bilateral ENT in our study, while Li found increased GI lay in left ENT, right precuneus, and superior frontal gyrus [52]. Li et al. [53] found CTH and SD may be markers reflecting abnormal connectivities related with brain development and disease. In our experiment, MCI patients had lower FD values than NC in left inferior parietal, superior temporal gyrus, and right posterior cingulate gyrus and precuneus. To date, the differences in FD between MCI and NC have not been studied extensively, so FD should be given more attentions in the future which might provide new insights into the pathophysiological mechanism of MCI.

Several limitations should be considered in future work. First, in addition to structural biomarkers

extracted from different morphological analyses, the early identification of MCI can be improved by using more biomarkers, including proteins measured in the CSF and imaging biomarkers extracted from different modalities, such as positron emission tomography and functional MRI [54–57]. Second, further studies could utilize ROI-based analysis and explore specific ROI related to MCI pathology, which might benefit MCI early diagnosis or discrimination. Third, MCI patients are known to be a clinically heterogeneous group with different patterns of brain atrophy, of which just some cases are due to AD. However, we do not ascertain the classification between MCI due to AD and MCI due to other dementia in our paper. One interesting and significant problem to address in the future is the differential diagnosis of MCI and clinical subtypes of MCI, which would be very useful in clinic.

In conclusion, multiple metrics extracted based on different morphological analysis methods provide complementary information about the brain differences between MCI and NC; therefore, their combination could effectively and robustly identify the MCI patients and is potentially useful for the diagnosis of MCI.

ACKNOWLEDGMENTS

This work was supported by Beijing Natural Science Foundation (No. 7174282, No. 7182019). Bin Jing was supported by Beijing Natural Science Foundation (No. 4122018) and Scientific Research Common Program of Beijing Municipal Commission of Education. Yuxia Li was supported by China Postdoctoral Science Foundation (2018M641414) and Beijing Postdoctoral Research Foundation (ZZ2019-12).

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

Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; MesoScale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (<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 Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Authors' disclosures available online (<https://www.j-alz.com/manuscript-disclosures/19-0715r2>).

REFERENCES

- [1] Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C (2017) A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *Neuroimage* **155**, 530-548.
- [2] Falahati F, Westman E, Simmons A (2014) Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J Alzheimers Dis* **41**, 685-708.
- [3] John JP, Lukose A, Bagepally BS, Halahalli HN, Moily NS, Vijayakumari AA, Jain S (2015) A systematic examination of brain volumetric abnormalities in recent-onset schizophrenia using voxel-based, surface-based and region-of-interest-based morphometric analyses. *J Negat Results Biomed* **14**, 11.
- [4] Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. *Neuroimage* **11**, 805-821.
- [5] Chung MK, Worsley KJ, Paus T, Cherif C, Collins DL, Giedd JN, Rapoport JL, Evans AC (2001) A unified statistical approach to deformation-based morphometry. *Neuroimage* **14**, 595-606.
- [6] Fischl B (2012) FreeSurfer. *Neuroimage* **62**, 774-781.
- [7] Scanlon C, Mueller SG, Tosun D, Cheong I, Garcia P, Barakos J, Weiner MW, Laxer KD (2011) Impact of methodologic choice for automatic detection of different aspects of brain atrophy by using temporal lobe epilepsy as a model. *AJNR Am J Neuroradiol* **32**, 1669-1676.
- [8] Schmitter D, Roche A, Marechal B, Ribes D, Abdulkadir A, Bach-Cuadra M, Daducci A, Granziera C, Kloppel S, Maeder P, Meuli R, Krueger G, Alzheimer's Disease Neu-

- roimaging Initiative (2015) An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *Neuroimage Clin* **7**, 7-17.
- [9] Koikkalainen J, Lotjonen J, Thurfjell L, Rueckert D, Waldemar G, Soininen H, Alzheimer's Disease Neuroimaging Initiative (2011) Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. *Neuroimage* **56**, 1134-1144.
- [10] Park H, Yang JJ, Seo J, Lee JM (2012) Dimensionality reduced cortical features and their use in the classification of Alzheimer's disease and mild cognitive impairment. *Neurosci Lett* **529**, 123-127.
- [11] Madan CR, Kensinger EA (2016) Cortical complexity as a measure of age-related brain atrophy. *Neuroimage* **134**, 617-629.
- [12] Bron EE, Smits M, van der Flier WM, Vrenken H, Barkhof F, Scheltens P, Papma JM, Steketee RME, Orellana CM, Meijboom R, Pinto M, Meireles JR, Garrett C, Bastos-Leite AJ, Abdulkadir A, Ronneberger O, Amoroso N, Bellotti R, Cardenas-Pena D, Alvarez-Meza AM, Dolph CV, Iftekharuddin KM, Eskildsen SF, Coupe P, Fonov VS, Franke K, Gaser C, Ledig C, Guerrero R, Tong T, Gray KR, Moradi E, Tohka J, Routier A, Durrleman S, Sarica A, Di Fatta G, Sensi F, Chincarini A, Smith GM, Stoyanov ZV, Sorensen L, Nielsen M, Tangaro S, Inglesse P, Wachinger C, Reuter M, van Swieten JC, Niessen WJ, Klein S, Alzheimer's Disease Neuroimaging Initiative (2015) Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CAD-Dementia challenge. *Neuroimage* **111**, 562-579.
- [13] Liu X, Tosun D, Weiner MW, Schuff N, Alzheimer's Disease Neuroimaging Initiative (2013) Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. *Neuroimage* **83**, 148-157.
- [14] Ma X, Li Z, Jing B, Liu H, Li D, Li H, Alzheimer's Disease Neuroimaging Initiative (2016) Identify the atrophy of Alzheimer's disease, mild cognitive impairment and normal aging using morphometric MRI analysis. *Front Aging Neurosci* **8**, 243.
- [15] Long Z, Jing B, Yan H, Dong J, Liu H, Mo X, Han Y, Li H (2016) A support vector machine-based method to identify mild cognitive impairment with multi-level characteristics of magnetic resonance imaging. *Neuroscience* **331**, 169-176.
- [16] Chincarini A, Bosco P, Calvini P, Gemme G, Esposito M, Olivieri C, Rei L, Squarcia S, Rodriguez G, Bellotti R, Cerello P, De Mitri I, Retico A, Nobili F, Alzheimer's Disease Neuroimaging Initiative (2011) Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. *Neuroimage* **58**, 469-480.
- [17] Singh V, Chertkow H, Lerch JP, Evans AC, Dorr AE, Kabani NJ (2006) Spatial patterns of cortical thinning in mild cognitive impairment and Alzheimer's disease. *Brain* **129**, 2885-2893.
- [18] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194.
- [19] Li Y, Jing B, Liu H, Li Y, Gao X, Li Y, Mu B, Yu H, Cheng J, Barker PB, Wang H, Han Y (2017) Frequency-dependent changes in the amplitude of low-frequency fluctuations in mild cognitive impairment with mild depression. *J Alzheimers Dis* **58**, 1175-1187.
- [20] Liu MX, Zhang DQ, Shen D, Alzheimer's Disease Neuroimaging Initiative (2015) View-centralized multi-atlas classification for Alzheimer's disease diagnosis. *Hum Brain Mapp* **36**, 1847-1865.
- [21] Savio A, Graña M (2013) Deformation based feature selection for computer aided diagnosis of Alzheimer's disease. *Expert Syst Appl* **40**, 1619-1628.
- [22] Dahnke R, Yotter RA, Gaser C (2013) Cortical thickness and central surface estimation. *Neuroimage* **65**, 336-348.
- [23] Yotter RA, Nenadic I, Ziegler G, Thompson PM, Gaser C (2011) Local cortical surface complexity maps from spherical harmonic reconstructions. *Neuroimage* **56**, 961-973.
- [24] Luders E, Thompson PM, Narr KL, Toga AW, Jancke L, Gaser C (2006) A curvature-based approach to estimate local gyrification on the cortical surface. *Neuroimage* **29**, 1224-1230.
- [25] Pell GS, Briellmann RS, Chan CH, Pardoe H, Abbott DF, Jackson GD (2008) Selection of the control group for VBM analysis: influence of covariates, matching and sample size. *Neuroimage* **41**, 1324-1335.
- [26] Breiman L (2001) Random forests. *Mach Learn* **45**, 5-32.
- [27] Sarica A, Cerasa A, Quattrone A (2017) Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Front Aging Neurosci* **9**, 329.
- [28] Ardekani BA, Bermudez E, Mubeen AM, Bachman AH, Alzheimer's Disease Neuroimaging Initiative (2017) Prediction of incipient Alzheimer's disease dementia in patients with mild cognitive impairment. *J Alzheimers Dis* **55**, 269-281.
- [29] Mietchen D, Gaser C (2009) Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution. *Front Neuroinform* **3**, 25.
- [30] Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95-113.
- [31] Ashburner J, Friston KJ (2011) Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation. *Neuroimage* **55**, 954-967.
- [32] Spalthoff R, Gaser C, Nenadic I (2018) Altered gyrification in schizophrenia and its relation to other morphometric markers. *Schizophr Res* **202**, 195-202.
- [33] de Miras JR, Costumero V, Belloch V, Escudero J, Avila C, Sepulcre J (2017) Complexity analysis of cortical surface detects changes in future Alzheimer's disease converters. *Hum Brain Mapp* **38**, 5905-5918.
- [34] Lebedev AV, Westman E, Van Westen GJ, Kramerberger MG, Lundervold A, Aarsland D, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Simmons A, Alzheimer's Disease Neuroimaging Initiative and the AddNeuroMed consortium (2014) Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *Neuroimage Clin* **6**, 115-125.
- [35] Uddin LQ, Nomi JS, Hebert-Seropian B, Ghaziri J, Boucher O (2017) Structure and function of the human insula. *J Clin Neurophysiol* **34**, 300-306.
- [36] Leech R, Sharp DJ (2014) The role of the posterior cingulate cortex in cognition and disease. *Brain* **137**, 12-32.
- [37] Righart R, Schmidt P, Dahnke R, Biberacher V, Beer A, Buck D, Hemmer B, Kirschke JS, Zimmer C, Gaser C, Muhlau M (2017) Volume versus surface-based cortical thickness measurements: A comparative study with healthy controls and multiple sclerosis patients. *PLoS One* **12**, e0179590.
- [38] Seiger R, Ganger S, Kranz GS, Hahn A, Lanzenberger R (2018) Cortical thickness estimations of FreeSurfer and the CAT12 Toolbox in patients with Alzheimer's disease and healthy controls. *J Neuroimaging* **28**, 515-523.

- [39] Westman E, Simmons A, Zhang Y, Muehlboeck JS, Tunard C, Liu Y, Collins L, Evans A, Mecocci P, Vellas BJN (2011) Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* **54**, 1178-1187.
- [40] Beheshti I, Demirel H, Matsuda H, Alzheimer's Disease Neuroimaging Initiative (2017) Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Comput Biol Med* **83**, 109-119.
- [41] Hua X, Leow AD, Parikshak N, Lee S, Chiang MC, Toga AW, Jack CR, Weiner MW, Thompson PM, Alzheimer's Disease Neuroimaging Initiative (2008) Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects. *Neuroimage* **43**, 458-469.
- [42] Maier S, Tebartz van Elst L, Perlov E, Duppers AL, Nickel K, Fangmeier T, Endres D, Riedel A (2018) Cortical properties of adults with autism spectrum disorder and an IQ>100. *Psychiatry Res Neuroimaging* **279**, 8-13.
- [43] Philippi N, Noblet V, Duron E, Cretin B, Bouilly C, Wisniewski I, Seux ML, Martin-Hunyadi C, Chaussade E, Demuynck C, Kremer S, Lehericy S, Gounot D, Armspach JP, Hanon O, Blanc F (2016) Exploring anterograde memory: a volumetric MRI study in patients with mild cognitive impairment. *Alzheimers Res Ther* **8**, 26.
- [44] Karas GB, Scheltens P, Rombouts SARB, Visser PJ, van Schijndel RA, Fox NC, Barkhof F (2004) Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. *Neuroimage* **23**, 708-716.
- [45] Platero C, Lopez ME, Carmen Tobar MD, Yus M, Maestu F (2019) Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: The local surface roughness. *Hum Brain Mapp* **40**, 1666-1676.
- [46] Braak H, Del Tredici K (2015) The preclinical phase of the pathological process underlying sporadic Alzheimer's disease. *Brain* **138**, 2814-2833.
- [47] Pennanen C, Kivipelto M, Tuomainen S, Hartikainen P, Hanninen T, Laakso MP, Hallikainen M, Vanhanen M, Nissinen A, Helkala EL, Vainio P, Vanninen R, Partanen K, Soininen H (2004) Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. *Neurobiol Aging* **25**, 303-310.
- [48] Gomez-Sancho M, Tohka J, Gomez-Verdejo V, Alzheimer's Disease Neuroimaging Initiative (2018) Comparison of feature representations in MRI-based MCI-to-AD conversion prediction. *Magn Reson Imaging* **50**, 84-95.
- [49] Poulin SP, Dautoff R, Morris JC, Barrett LF, Dickerson BC, Alzheimer's Disease Neuroimaging Initiative (2011) Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatry Res* **194**, 7-13.
- [50] Bubb EJ, Kinnavane L, Aggleton JP (2017) Hippocampal - diencephalic - cingulate networks for memory and emotion: An anatomical guide. *Brain Neurosci Adv* **1**, 2398212817723443.
- [51] Das SR, Pluta J, Mancuso L, Kliot D, Yushkevich PA, Wolk DA (2015) Anterior and posterior MTL networks in aging and MCI. *Neurobiol Aging* **36** Suppl 1, S141-150, S150.e1.
- [52] Li S, Yuan X, Pu F, Li D, Fan Y, Wu L, Chao W, Chen N, He Y, Han Y (2014) Abnormal changes of multidimensional surface features using multivariate pattern classification in amnesic mild cognitive impairment patients. *J Neurosci* **34**, 10541-10553.
- [53] Li Q, Li X, Wang X, Li Y, Li K, Yu Y, Yin C, Li S, Han Y (2016) Topological properties of large-scale cortical networks based on multiple morphological features in amnesic mild cognitive impairment. *Neural Plast* **2016**, 3462309.
- [54] Melah KE, Lu SY, Hoscheidt SM, Alexander AL, Adluru N, Destiche DJ, Carlsson CM, Zetterberg H, Blennow K, Okonkwo OC, Gleason CE, Dowling NM, Bratzke LC, Rowley HA, Sager MA, Asthana S, Johnson SC, Bendlin BB (2016) Cerebrospinal fluid markers of Alzheimer's disease pathology and microglial activation are associated with altered white matter microstructure in asymptomatic adults at risk for Alzheimer's disease. *J Alzheimers Dis* **50**, 873-886.
- [55] Zhou T, Thung KH, Zhu X, Shen D (2019) Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Hum Brain Mapp* **40**, 1001-1016.
- [56] Yan T, Wang Y, Weng Z, Du W, Liu T, Chen D, Li X, Wu J, Han Y (2019) Early-stage identification and pathological development of Alzheimer's disease using multimodal MRI. *J Alzheimers Dis* **68**, 1013-1027.
- [57] Zhou T, Liu M, Thung KH, Shen D (2019) Latent representation learning for Alzheimer's disease diagnosis with incomplete multi-modality neuroimaging and genetic data. *IEEE Trans Med Imaging* **38**, 2411-2422.