Research Paper

Application of Generalized Split Linearized Bregman Iteration algorithm for Alzheimer's disease prediction

Weimin Zheng[1,*], Bin Cui[1,*], Zeyu Sun[2,*], Xiuli Li[2], Xu Han[1], Yu Yang[3], Kuncheng Li[4], Lingjing Hu[5], Zhiqun Wang[1], Alzheimer's Disease Neuroimaging Initiative[#]

[1]Department of Radiology, Aerospace Center Hospital, Beijing 100049, China
[2]Deepwise AI lab, Beijing 100080, China
[3]Beijing Huading Jialiang Technology Co, Beijing 100000, China
[4]Department of Radiology, Xuanwu Hospital of Capital Medical University, Beijing 100053, China
[5]Yanjing Medical College, Capital Medical University, Beijing 101300, China
*Equal contribution

[#]In this article, when performing the cross-test, we used the Data which 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 listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Correspondence to: Lingjing Hu, Zhiqun Wang; **email:** hulj@ccmu.edu.cn, wangzhiqun@126.com
Keywords: Alzheimer's disease, machine learning, generalized split linearized Bregman iteration, voxel-based structural magnetic resonance imaging, feature selection
Received: November 5, 2019 **Accepted:** February 25, 2020 **Published:** April 5, 2020

ABSTRACT

In this paper, we applied a novel method for the detection of Alzheimer's disease (AD) based on a structural magnetic resonance imaging (sMRI) dataset. Specifically, the method involved a new classification algorithm of machine learning, named Generalized Split Linearized Bregman Iteration (GSplit LBI). It combines logistic regression and structural sparsity regularizations. In the study, 57 AD patients and 47 normal controls (NCs) were enrolled. We first extracted the entire brain gray matter volume values of all subjects and then used GSplit LBI to build a predictive classification model with a 10-fold full cross-validation method. The model accuracy achieved 90.44%. To further verify which voxels in the dataset have greater impact on the prediction results, we ranked the weight parameters and obtained the top 6% of the model parameters. To verify the generalization of model prediction and the stability of feature selection, we performed a cross-test on the Alzheimer's Disease Neuroimaging Initiative (ADNI) and a Chinese dataset and achieved good performances on different cohorts. Conclusively, based on the sMRI dataset, our algorithm not only had good performance in a local cohort with high accuracy but also had good generalization of model prediction and stability of feature selection in different cohorts.

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disease leading to dementia, typically manifesting as memory disturbance, attentional and executive deficits, and visuospatial and perceptual impairments. It is pathologically characterized by the deposition of amyloid-β plaques and tau-related neurofibrillary tangles, resulting in loss of neurons [1]. Currently, AD is still an irreversible condition, and there are no effective medications available today. Studies have found that early

diagnosis and administration of drugs can delay the progression of the disease. However, it is difficult to diagnose AD in the early stage in clinical practice. Thus, quantitative analysis based on imaging can possibly provide a potential method to make an early diagnosis of AD.

In the past several years, many neuroimaging studies have been performed to develop different biomarkers for the early diagnosis of AD at the individual level [2, 3]. Among various neuroimaging modalities, structural magnetic resonance imaging (sMRI) is most commonly used, possibly due to its wide operability and objectivity. AD pathological changes are mostly involved in the hippocampus, medial temporal gyrus, posterior cingulate gyrus, as well as precuneus [4–7]. Progressive brain atrophy due to neuropathology is often measured using sMRI and is taken as a valuable imaging biomarker for early individual diagnosis of AD [8, 9]. Many neuroimaging studies have used region-of-interest (ROI)-based analysis to explore the subtle local atrophy caused by AD, thus proposing imaging classifiers to distinguish AD from normal control (NC) individuals [10, 11]. Such studies relied solely on prior knowledge to guide the selection of ROI and features, thus ignoring the structural changes of the entire brain and the microstructural abnormalities in the anatomy, which made it difficult and challenging to establish reliable markers for diagnosing AD in the early stages. The potential considerations of classification in clinical practice have largely driven the development of machine learning, which can provide a systematic approach to developing complex, automated, and objective classification frameworks for analyzing high-dimensional data across the whole brain. Typically, the classification framework includes feature extraction and classification algorithms to build predictive models and develop imaging markers to perform classification with high sensitivity and specificity. Applying different classification algorithms on the extracted neuroimaging features for AD/mild cognitive impairment (MCI) showed great advantages for detecting AD at the prodromal stages, even before clinical manifestation [12, 13].

For AD classification, machine learning has attracted increasing attention by using the multimodal quantify patterns of atrophy together with different algorithms in recent years. For feature extraction, brain atrophy was most often quantified via tissue density maps, volume maps, cortical thickness measures, and geometric measures of the hippocampus from sMRI. However, several different classification algorithms have been proposed and applied for AD classification and have achieved promising results. Support vector machine (SVM) is the most popular algorithm for AD classification [14–18]. SVM can extract high-dimensional, informative features from MRI to build predictive classification models, resulting in the automation of clinical diagnosis. Multi-kernel learning, which is an extension of ordinary kernel-based classification algorithms, has also been increasingly used in AD classification [19–21]. Other less common classification algorithms used in AD research include linear discriminant analysis (LDA) [22, 23], orthogonal partial least square regression [24], random forest [25], regularization-based methods [26], voting-based ensemble methods [27], kernel SVM decision-tree [28], and spatially augmented linear programming boosting method (LPBM) [29]. Although several classification algorithms have been applied in AD, there were still some issues to be noted. For example, the algorithms with a general loss and L2 penalty, which could automatically select or extract classification-related features, were poorly interpretable or easy to overfit. Some other algorithms, such as L1 regularization-based methods, which had structural sparsity and interpretability, were inclined to ignore the effect of procedural bias on the classification. Moreover, in our recent study of AD, we found that the preprocessing steps could introduce one type of feature called procedural bias [30], referring to mistakenly enlarged gray matter volume during registration and segmentation in the preprocessing steps. Therefore, for neuroimaging-based AD classification, it has attracted increasing attention to derive a more precise classification algorithm that could take both lesion features and procedural bias into consideration and make better predictions on an individual basis.

Here, to achieve accurate classification of AD, we proposed a new classification algorithm in our preliminary study [30], named the Generalized Split Linearized Bregman Iteration (GSplit LBI) method, which combines logistic regression and structural sparsity regularizations to verify its capacity in a Chinese cohort. The model had a sparsity enforcement based on the idea of variable splitting; therefore, it can effectively leverage both procedural bias and lesion features into prediction, and it has a better regularization path and interpretability. By using this algorithm on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset in our previous study, the advantages of GSplit LBI were verified by the improved stability of selected lesion features and better classification results when compared with other algorithms, such as Maximum uncertainty Linear Discriminant Analysis (MLDA), SVM, Lasso, Graphnet, Elastic Net, Total Variation (TV + l_1) and Nonnegative Generalized Fused Lasso (n^2GFL). However, further analysis of the GSplit LBI in AD was not performed; therefore, the key brain regions that

determine the classification of AD and its association with clinical variables remain unknown. Most importantly, the generalization and stability of the GSplit LBI model are not very clear. Furthermore, other than the ADNI dataset, we do not know the prediction capacity of the model in other cohorts.

In the present study, we aim to use gray matter (GM) voxels and the GSplit LBI algorithm to distinguish AD from NC at the individual level in a Chinese cohort. We hypothesized that this classification was driven by a distributed pattern of GM voxel alterations that were involved in the temporal lobe, such as the entorhinal cortex, hippocampus, parahippocampal gyrus and other limbic system components. Furthermore, because structural changes might be affected in some specific cognitive-related regions early in the disease course, we expected that atrophy of specific regions could accurately describe and track disease progression, which could be applied as a valuable imaging biomarker for the early diagnosis of AD. In addition, we further extracted these specific regions and made correlations with cognitive performance as measured by mini-mental state examination (MMSE). Finally, we performed a cross-test on the ADNI dataset and our in-house dataset using the model parameters trained from our in-house dataset and the ADNI dataset, respectively. We speculated that the GSplit LBI algorithm not only had good performance in a local cohort but also had generalization of model prediction and stability of feature selection in different cohorts.

RESULTS

Demographic and neuropsychological tests

The demographic characteristics are shown in Table 1. There were no significant differences of gender, age and education between the AD and NC groups (both $P > 0.01$). However, the AD group exhibited significantly lower MMSE than the NC group ($P < 0.0001$).

Classification results for an in-house test dataset

After the 10-fold cross-validation was completed, we performed statistical analyses on the results of our model between 57 AD patients and 47 NCs and reported the results of each 10-fold cross-validation in Supplementary Table 1. The overall accuracy, which is the ratio of the correct classification number of AD/NC samples to the total number of AD/NC samples, was 90.44%. In ten cross-validations, the maximum accuracy was 91.36%, and the minimum accuracy was 89.55%. Moreover, the average of sensitivity was 91.17%, ranging from 89.67% to 91.67%, and the average of specificity was 89.50%, ranging from 87.00% to 92.00%. The area under the curve (AUC) value of our AD/NC classification model could reach 0.9090, and the receiver operating characteristic (ROC) curve is shown in Figure 1.

Brain areas involved in the classification analysis

To verify which voxels in the data have greater impacts on the prediction results, we sorted the weight parameters in main parameter β from large to small and performed classification experiments with the largest top n (ranging from 1 to 2527; 2527 is the number of all input features) weight parameters in turn. We then computed the relationship between the prediction accuracy and the weight parameters used in experiments. As shown in Figure 2, at the beginning of the curve, with the increase of the weight number, the prediction accuracy increased gradually. However, when the weight number reached 163 (about the top 6%), the prediction accuracy attained its maximum and remained unchanged after that. This meant that adding more weight parameters would only produce information redundancy and could not improve performance. Therefore, we defined the voxels corresponding to these top 6% weight parameters as key voxels. The weight distribution map of these voxels is shown in Supplementary Figure 1. The brain regions with the largest weight ratio of our in-house dataset are shown in Figure 3A, and their information is reported in Table 2, including lateral temporal lobe, entorhinal cortex, the hippocampus, parahippocampal gyrus and the limbic system, motor cortex area (M1), cerebellum crus2 and thalamus.

Relationship between structural changes and cognitive behaviors

In the AD group, positive correlations were found between the MMSE scores and the structural changes of several regions obtained from GSplit LBI (i.e., bilateral middle temporal gyrus (MTG), bilateral angular gyrus (ANG), bilateral supramarginal gyrus (SMG), left inferior temporal gyrus (ITG), right superior temporal gyrus (STG), right precuneus (PCUN), right calcarine fissure and surrounding cortex (CAL), and left thalamus (THA)) ($P < 0.05$, with age, gender and education level as covariates). The details are shown in Figure 4.

Cross-test results on the ADNI dataset and in-house dataset

In cross-test, using the Chinese model trained from our in-house dataset, the accuracy, specificity and sensitivity of the prediction results on the ADNI dataset reached 86.36%, 80.30% and 90.00%, respectively. Simultaneously, using the ADNI model, the accuracy,

Table 1. Characteristics of AD patients and normal controls.

Characteristics	AD	Contrlos	*P* value
N (M/F)	57 (25/32)	47 (23/24)	0.61[a]
Age, years	65.21±9.14	63.94±8.06	0.46[b]
Education, years	10.56±6.47	11.36±3.59	0.43[b]
MMSE	14.02±5.90	26.49±4.02	<0.0001[b]
CDR	1.43±0.54(1-3)	0	-

MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating; plus–minus values are means±S.D.
a The P value for gender distribution in the two groups was obtained by Chi-square test.
b The P values were obtained by a two-sample two-tailed t-test.

specificity and sensitivity of the prediction results on our in-house dataset also reached 84.26%, 78.95% and 91.49%, respectively. The ROC curve of the cross-test is shown in Figure 5, and the AUC values of these two tests were 0.92 and 0.91, respectively. When using the model trained from our in-house dataset on the ADNI dataset, the brain regions with the largest weight ratios on the classification of AD patients and NCs are shown in Figure 3B.

DISCUSSION

Effective and accurate AD diagnosis is critical for early treatment. Therefore, many researchers have devoted their efforts to develop a computer-aided system that can diagnose AD in the early stages and on an individual basis [2, 3]. The present study demonstrated that patients with AD could be distinguished from NCs using a classified model based on GM voxels and the GSplit LBI algorithm, with good to excellent accuracy. This classification was driven by a distributed pattern of GM voxel alterations, involving the lateral temporal lobe, entorhinal cortex, hippocampus, parahippocampal gyrus, limbic system, M1, cerebellum crus2 and thalamus. In addition, we found that atrophy of several specific regions significantly correlated with cognitive performances as measured by MMSE. Finally, we performed a cross-test to verify the generalization of the

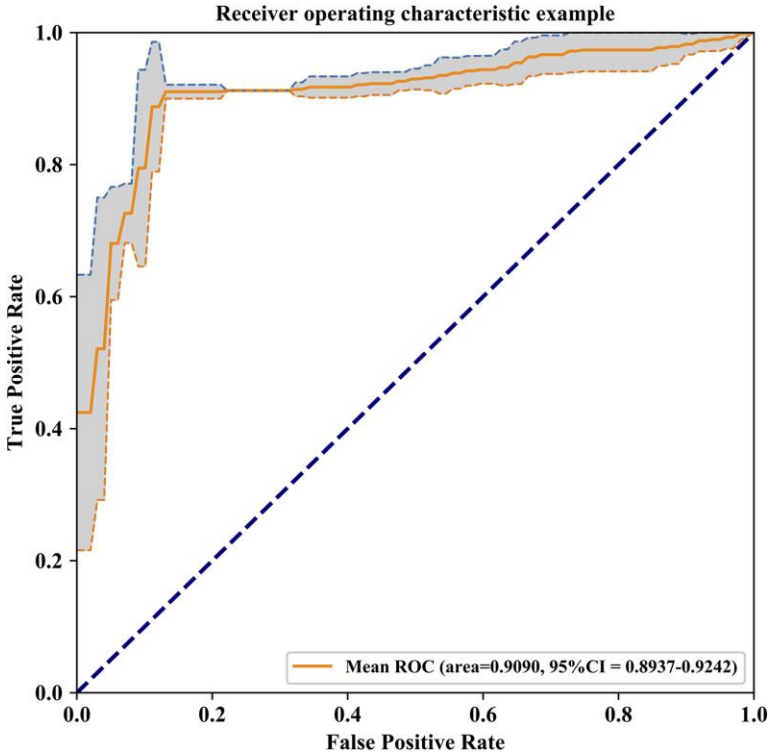


Figure 1. Receiver Operating Characteristic curve of the model prediction results, The AUC value of our AD/NC classification model is 0.9090 and 95% confidence interval is 0.8937-0.9242 (gray area in the graph).

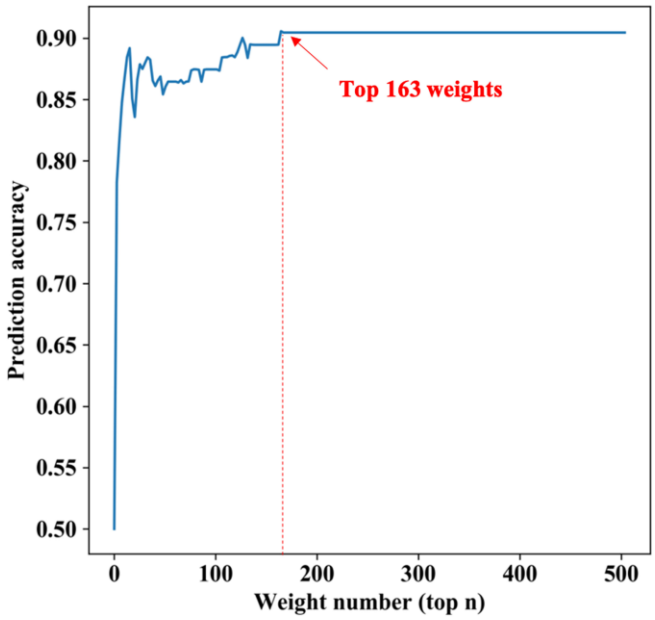


Figure 2. The relationship between the predicted results and the number of weights. The horizontal axis represents the largest top n weight parameters, the longitudinal axis represents the prediction accuracy under these parameters. At the beginning of the curve, with the increase of the weight number, the prediction accuracy increases gradually. However, when the largest top parameters number reaches 163 (about the top 6%), the prediction accuracy has the maximum and remains unchanged after that.

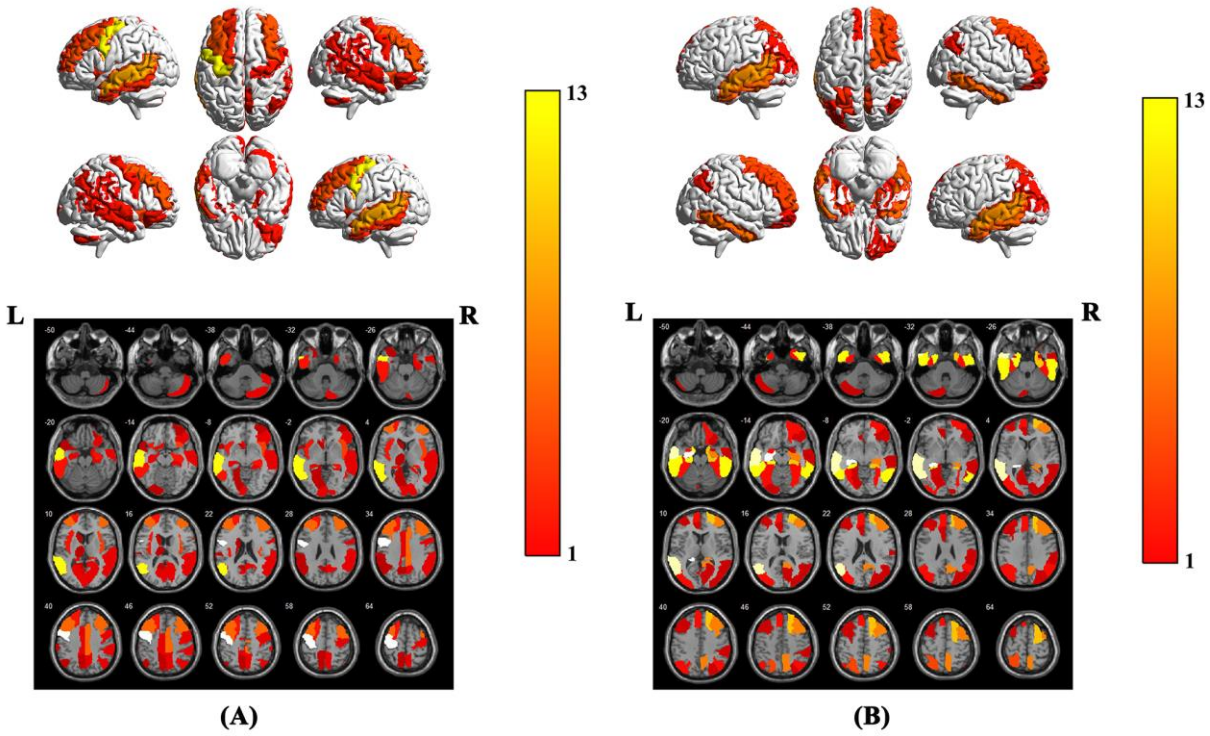


Figure 3. Brain region weight map based on AAL116 template, (**A**) represents the brain regions that have the greatest impact on the classification between AD patients and NCs when using the model parameters of this experiment on our in-house dataset, (**B**) represents the brain regions that have the greatest impact on the classification between AD patients and NCs. The weight value of each brain region is based on the average of the weight values in the brain region. The color bar represents the average weight value in each brain region, the larger the weight value of the model, the warmer the color in the graph.

Table 2. The brain regions that have the greatest impact on the classification between AD and NC.

Brain region	AAL index	Volume(mm^3)	Talairach			Weight (%)
			X(mm)	Y(mm)	Z(mm)	
Left precentral gyrus	1	28174	-38.65	-5.68	50.94	13.25
Left middle temporal gyrus	85	39353	-55.52	-33.80	-2.20	8.77
Right middle frontal gyrus	8	40374	37.59	33.06	34.04	4.80
Left middle frontal gyrus	7	38722	-33.43	32.73	35.46	4.80
Right median cingulate and paracingulate	34	17442	8.02	-8.83	39.79	4.70
Right insula	30	14128	39.02	6.25	2.08	4.34
Left inferior temporal gyrus	89	25647	-49.77	-28.05	-23.17	2.83
Left insula	29	15025	-35.13	6.65	3.44	2.72
Left superior frontal gyrus, dorsolateral	3	28915	-18.45	34.81	42.20	2.68
Right hippocampus	38	7606	29.23	-19.78	-10.33	2.63
Right supramarginal gyrus	64	15770	57.61	-31.50	34.48	2.50
Right superior temporal gyrus	82	25258	58.15	-21.78	-6.80	2.45
Right angular gyrus	66	14009	45.51	-59.98	38.63	2.40
Right middle temporal gyrus	86	35484	57.47	-37.23	-1.47	2.39
Right parahippocampal gyrus	40	9028	25.38	-15.15	-20.47	2.37
Right inferior frontal gyrus, orbital part	16	13747	41.22	32.23	-11.91	2.36
Right precentral gyrus	2	27058	41.37	-8.21	52.09	2.20
Left calcarine fissure and surrounding cortex	43	18157	-7.14	-78.67	6.44	2.03
Right precuneus	68	26083	9.98	-56.05	43.77	1.74
Left median cingulate and paracingulate gyri	33	15512	-5.48	-14.92	41.57	1.73
Right Calcarine fissure and surrounding	44	14885	15.99	-73.15	9.40	1.63
Right cerebellum crus2	94	17038	42.12	-69.97	-45.75	1.59
Left Hippocampus	37	7469	-25.03	-20.74	-10.13	1.53
Left olfactory cortex	21	2262	-8.06	15.05	-11.46	1.50
Left lingual gyrus	47	16932	-14.62	-67.56	-4.63	1.40
Left supramarginal gyrus	63	9907	-55.79	-33.64	30.45	1.31
Right middle frontal gyrus, orbital part	10	8057	33.18	52.59	-10.73	1.27
Left angular gyrus	65	9313	-44.14	-60.82	35.59	1.15
Left precuneus	67	28358	-7.24	-56.07	48.01	1.09
Left caudate nucleus	71	7682	-11.46	11.00	9.24	1.08
Left temporal pole: superior temporal gyrus	83	10228	-39.88	15.14	-20.18	0.96
Left thalamus	77	8700	-10.85	-17.56	7.98	0.88

These brain regions, indexes and volumes are based on AAL116 templates, the value of weight represents the weight ratio of this brain region.

model prediction and stability of feature selection in the classification.

Advantages of GSplit LBI for AD prediction

Multiple GM microstructural abnormalities in AD have been reported in previous studies, which have enabled the discrimination of AD from NC [4–7, 10, 31–33]. However, these studies only reported group-level differences of various brain structures and did not consider evaluations of single subjects [34]. Here, we used GSplit LBI to examine whether the GM microstructural abnormalities could be used to discriminate between AD patients and NCs at the individual level. In voxel-based neuroimage analysis, lesion features have been the main focus in disease prediction due to their interpretability with respect to the related diseases. However, the "Procedural Bias", which could be leveraged to improve classification accuracy, was introduced during the preprocessing

steps [35]. Among most existing models, the models with a general loss and L2 penalty are poorly interpretable and are easy to overfit. These kinds of models automatically select or extract features that are strongly related to classification according to the principle of minimizing classification errors. Other kinds of models, such as L1 regularization-based methods, have structural sparsity and interpretability, but it is easy to ignore the effect of procedural bias on AD/NC classification. In the preliminary study of our team members [30], empirical experiments were evaluated on the ADNI dataset. The advantage of GSplit LBI is verified by the improved stability of selected lesion features and better classification results when compared with other models, such as MLDA, SVM, Lasso, Graphnet, Elastic Net, TV+l1 and n2GFL. In the present study, we used GSplit LBI to distinguish AD from NC at the individual level in a Chinese cohort, and the average accuracy, sensitivity and specificity of this classifier reached 90.44%, 91.17% and 89.50%, respectively. The misclassified cases are mainly distributed in cases with low clinical dementia rating (CDR) scores. This is because the higher the CDR scores, the heavier the dementia, the easier it is to classify. Conversely, the lower the CDR scores, the lighter the dementia, the easier it is to classify errors. In addition, the AUC value of our AD/NC classification model reached 0.9, improved the stability of selected lesion features and led to better classification results. Moreover, the ROC curve shows that our model is highly sensitive and specific.

Brain areas involved in the classification analysis

In the present study, the discrimination was based not only on atrophy of the lateral temporal lobe, entorhinal cortex, the hippocampus, parahippocampal gyrus and limbic system but also on regions of M1, cerebellum crus2 and thalamus, which are not traditionally implicated in AD. This demonstrated the capacity of GSplit LBI to detect subtle and distributed GM alterations. Previous neuroimaging studies in AD have revealed alterations in the lateral temporal lobe, entorhinal cortex, hippocampus and limbic system [4–7, 10, 31–33], reflecting different disease stages and predicting the progression from MCI to AD [8]. In addition, AD patients accumulated abnormal proteins (Aβ and tau) in the form of amyloid plaques and neurofibrillary tangles, eventually resulting in loss of neurons in these areas [8, 36]. Alterations within these areas might explain memory problems, including difficulties in word finding and thinking processes,

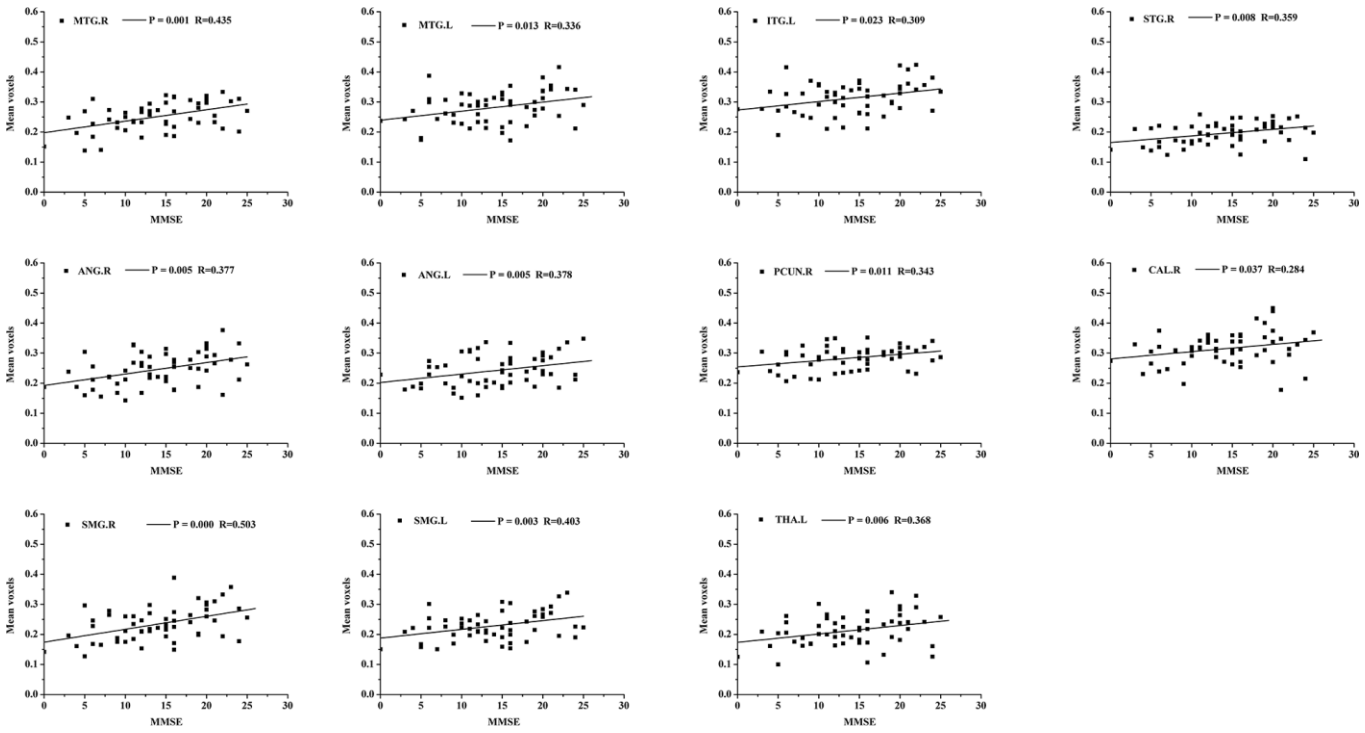


Figure 4. Scatterplot of mean voxels of the bilateral MTG, bilateral ANG, bilateral SMG, left ITG, right STG, right PCUN, right CAL and left THA plotted against MMSE scores (p<0.05, with age, gender and education level as covariates). MTG, middle temporal gyrus; ANG, angular gyrus; SMG, supramarginal gyrus; ITG, inferior temporal gyrus; STG, superior temporal gyrus; PCUN, precuneus; CAL, calcarine fissure and surrounding cortex; THA, thalamus.

abilities to reason, making judgments, communication and dealing with daily activities, which are the most common symptoms in AD [37, 38]. Consistent with these previous structural studies, our findings of the high discriminative values on the lateral temporal lobe, entorhinal cortex, hippocampus and limbic system provided new evidence that microstructural abnormalities within these areas were critically occurred in AD.

Except for the common GM atrophy areas, we also found that cerebellum crus2, M1 and thalamus contributed to the identification of AD patients. This result was consistent with other AD-related studies, which were also involved in cerebellar subregions. For instance, GM atrophy of the cerebellum has been detected in AD in several neuroimaging studies [39, 40]. Furthermore, AD pathological changes have now been revealed in the cerebellum, including deposits of amyloid-b plaques, neurofibrillary tangles, and increased microglia [41–43]. As the core region, the cerebellum has efferent and afferent fibers between the vermis, hypothalamus and limbic system [44]; thus, the alteration of the cerebellum might reflect abnormal multimodal functions. In our study, we also found that M1 contributed to the identification of AD from NC. In previous studies, some task-related functional magnetic resonance imaging (fMRI) studies have demonstrated reduced activation in the premotor cortex in AD patients when performing motor-related tasks [45, 46]. In addition, one resting state fMRI (rs-fMRI) study reported the significant functional abnormality of sensorimotor (SMN) cortex in AD patients [47]. In conjunction with the previous reports and our findings, we speculated that AD patients might present subtle motor impairment caused by the atrophy and dysfunction of the sensorimotor cortex. In the present study, we found that thalamus atrophy contributed to the identification of AD patients, which was consistent with several AD related studies. Reduction of thalamic volume was typically observed in some cases of amnestic MCI [48–51]. In 2010, de Oliveira performed a longitudinal analysis of thalamic tissue texture, finding a stepwise decline in thalamic status going from controls to amnestic MCI and AD cases [52]. Likewise, thalamic volume was found to be correlated with cognitive performance in MCI patients through MRI-based measurements [50, 51]. Furthermore, reduced thalamic volume was found in AD cases when compared to other non-dementing individuals who reported memory lapses [53], with thalamic volume again correlating with the decline of global cognitive performances.

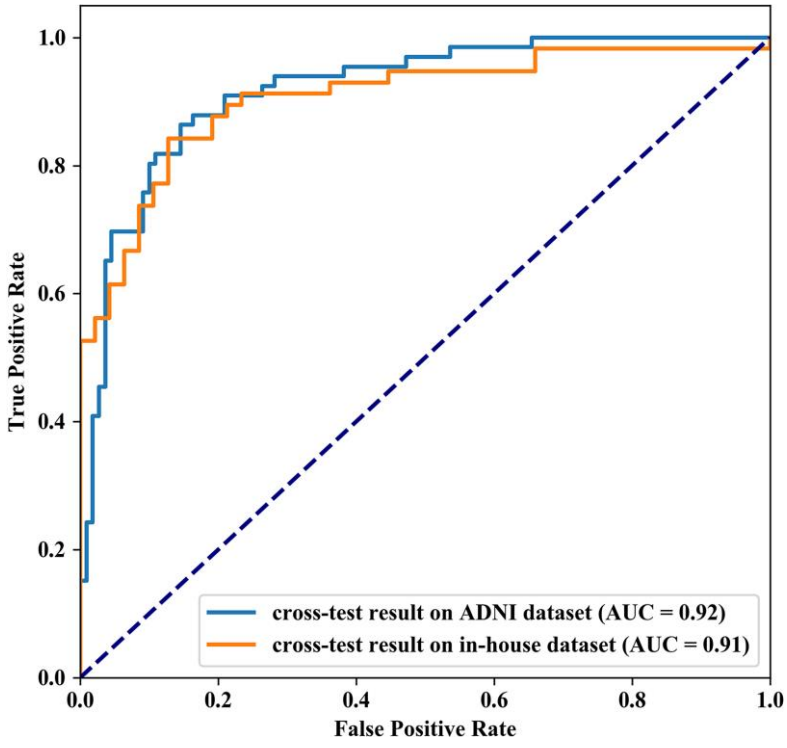


Figure 5. Receiver Operating Characteristic curve of the cross-test results. The AUC value of cross-test results on ADNI dataset and in-house dataset are 0.92 and 0.91 respectively.

Relationship between structural changes and cognitive behaviors

In addition, our study found a close relationship between cognitive impairment (MMSE scores) and mean voxels of several regions obtained from GSplit LBI in AD patients, including bilateral MTG, bilateral ANG, bilateral SMG, left ITG, right STG, right PCUN, right CAL and left THA, suggesting that structural changes in these regions could be used as imaging markers for tracking disease progression. Among these brain regions, the left MTG and left ITG account for a large weight on the classification between AD patients and NCs of GSplit LBI, further verifying the generalization of this model prediction and the stability of feature selection.

Reproducibility and Stability of GSplit LBI in different cohorts

In the cross-test on our in-house dataset and the public ADNI dataset, whether using the Chinese model or the ADNI model, we successfully achieved high accuracy, sensitivity and specificity. In addition, the AUC values of these two tests were 0.92 and 0.91 respectively, indicating that our GSplit LBI algorithm not only has good performance in a local cohort but also has generalization of model prediction in different cohorts. Moreover, during the cross-test on the ADNI dataset and our in-house dataset, when using the Chinese model, the brain regions that had the greatest impact on the classification between AD patients and NCs were mainly involved in the lateral temporal lobe, entorhinal cortex, hippocampus and limbic system and cerebellum crus2, which were highly consistent with the results when using the model parameters of this experiment on our in-house dataset, further suggesting the good reproducibility and stability of feature selection of GSplit LBI.

Future considerations

Several issues should be considered. First, these results were based on a relatively small sample size, although it provided preliminary support for the potential of GSplit LBI as a diagnostic aid for AD. In the future, large samples might be collected to confirm the results. Second, to explore whether the application of GSplit LBI could discriminate AD patients in different stages, future studies will add more samples of early stages of AD patients, such as MCI, as well as ApoE 4 carriers. In fact, in some previous studies [54–56], the author used different machine learning algorithms to classify stable MCI (sMCI) and progressive MCI (pMCI) in the ADNI dataset and achieved good results. We are currently collecting MCI patients and are performing follow-up. When the sample size is sufficient, we intend to test the GSplit LBI algorithm's classification effect on sMCI and pMCI in the in-house and ADNI datasets in future research. Third, considering the time complexity and space complexity of the algorithm, we currently only perform experiments on coarse scale voxels 8x8x8 mm^3 in size. In the future, some finer scale voxels will be extracted and other faster optimization methods will be introduced to our experiments. Finally, in this study, we only focus on the structural changes in early AD; future work will combine multimodal neuroimaging findings, such as structural, functional and perfusion MRI, to examine whether this combination analysis could lead to higher levels of diagnostic accuracy.

CONCLUSIONS

In summary, the present study revealed special patterns of GM abnormalities in patients with AD. By using the GSplit LBI algorithm, which has been proven to have good generalization, these GM abnormalities can be applied to accurately differentiate AD patients and NCs at the level of the individual.

MATERIALS AND METHODS

Participants

A total of 104 right-handed subjects participated in this study after providing written informed consent, including 57 patients with AD and 47 NCs. This study was carried out in accordance with the recommendations of the Medical Research Ethics Committee of Xuanwu Hospital. All subjects provided written informed consent in accordance with the Declaration of Helsinki. The AD subjects were recruited from patients who had consulted the memory clinic at Xuanwu Hospital for memory complaints. The NCs were recruited from the local community.

All participants underwent complete physical and neurological examinations, standard laboratory tests and neuropsychological assessments. The neuropsychological examinations included MMSE and CDR. The AD patients fulfilled the new research criteria for possible or probable AD [57, 58]. The new criteria emphasized the clinical history, neuropsychological assessment, sMRI, positron emission tomography (PET), and cerebrospinal fluid (CSF) examinations. Based on the new criteria, we carefully evaluated our study samples and confirmed each patient as AD. Among the 57 AD patients, 33 patients had a CDR [59] score of 1 and were thus assigned to a mild AD category. Additionally, 24 patients had a CDR of 2 or 3, suggesting moderate or severe AD categories.

The NCs met the following criteria: a) no neurological or psychiatric disorders, such as stroke, depression and epilepsy; b) no neurological deficiencies, such as visual or hearing loss; c) no abnormal findings, such as infarctions or focal lesions, in conventional brain MRI; d) no cognitive complaints; and e) CDR score of 0.

Participants with contraindications for MRI, such as pacemaker, cardiac defibrillator, implanted material with electric or magnetic systems, vascular clips or mechanical heart valve, cochlear implant or claustrophobia were excluded. In addition, patients with a history of stroke, psychiatric diseases, drug abuse, severe hypertension, systematic diseases and intellectual disability were excluded.

At last, we intend to briefly explain why we chose only right-handed subjects. First, the ratio of left-handed and right-handed people in the world is about 3:17. No matter in AD or normal people, samples of right-handed people are easier to collect. Indeed, most of the patients we collected were right-handed. In order to match the AD group, we selected the normal group with right-handed. Second, the size of the brain is different between the left-handed and right-handed people. Compared with the left-handed people, the right-handed people have smaller brain [60]. Therefore, it is necessary to consider the hand when quantifying the variability among individuals. Previous studies [61, 62] have consistently shown that people with smaller brain have a relatively higher percentage of gray matter, which can be used in most tissues to process local information. In this study, we focused on gray matter voxels, so only the right-handed subjects were selected.

MRI acquisition protocol

MRI data acquisition was performed on a SIEMENS verio 3-Tesla scanner (Siemens, Erlangen, Germany). The subjects were instructed to hold still, keep their eyes closed and think of nothing in particular. 3D T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) sagittal images were obtained with the following parameters: TR/TE/TI/FA = 1900 ms/2.2 ms/900 ms/9°, image matrix = 256×256, slice number = 176, thickness = 1 mm.

Preprocessing of structural MRI data

Preprocessing of the MRI data was carried out using SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) and the VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm). First, all 3D T1-weighted images were corrected for the bias field with regard to homogeneity in the VBM8 toolbox. Second, the corrected images were normalized and then segmented into GM, white matter (WM), and cerebrospinal fluid (CSF) components. Only GM components were considered in the current study. The normalization method we used was DARTEL [63], and the template we selected was the default DARTEL standard template. In the present preprocessing of structural MRI data, images were not smoothed. When no smoothing is employed, the sensitivity is increased. In addition, our model added the clustering constraints, which has its own denoising properties. After this processing, the total number of voxels is 121x145x121, and the voxel size is 1.5x1.5x1.5 mm^3. Third, a downsampling processing was used to change the voxel size to 8x8x8 mm^3 and the number of voxels is 24x28x24. Finally, a total number of 2,527 voxels with average values in the GM population template greater than 0.1 was extracted from the processed images and served as the input features.

GSplit LBI-based classifier

To achieve accurate classification of AD patients and NCs, we adopted the GSplit LBI method, combining logistic regression and sparse regularization of structure, and used the generalized image descent method to optimize the parameters of such binary classification problems. The model loss function was defined as follows:

$$\mathrm{L}(\beta,\gamma) = -\sum_{i=1}^{N}\Big[y_i * \log\big(\sigma(x_i\beta)\big) + (1-y_i) * \log\big(1-\sigma(x_i\beta)\big)\Big] + \frac{1}{2\nu}\|D\beta-\gamma\|_2^2$$

β is the main parameters of the model, γ is sparse constraint parameter for β, $\sigma(\cdot)$ is sigmoid function, and ν is a hyperparameter to balance loss term and regularization term. To obtain a better regularization path, we adopted the idea of variable splitting and introduced an additional sparse parameter γ in the regularization parameter. The structural sparsity of the model can be satisfied by controlling the square of 2-norms of sparse γ and $D\beta$. The goal of optimization is to minimize this loss function based on training set in each iteration. For more detail of Split LBI, please refer to the previous study [30].

Design of classification experiments

To make full use of the available data, a 10-fold full cross-validation method was used to train our model. In each cross-validation step, to ensure that all data were involved in the training and validation process, the dataset was randomly divided into ten subsets, nine

of which were used to train the model in turn; after that, the trained model was used to predict the results on the remaining subset. After all the subsets had been predicted, we combined the results of the ten subsets and got the prediction results for the entire dataset, shown in Figure 6. We then calculated the model's accuracy, sensitivity, specificity and other indicators on the results. We repeated this process 10 times and calculated the mean value and 95% CI range of these experimental results. In addition, the ROC analysis, which is a plot of the fraction of true-positive test results (sensitivity) versus the fraction of false-positive test results (1-specificity), was plotted after these experiments. Finally, the AUC was calculated to evaluate the model's classification performance of AD patients and NCs, and the mean and 95% CI of the ROC curve were used to evaluate the stability of the model.

Brain area analysis in the classification

We carried out a sensitivity analysis of our classification model by analyzing which parts of the gray matter voxels have greater impacts on the classification between AD patients and NCs. In this analysis, we take the parameters corresponding to each voxel in main parameter β as its contribution to classification. The weight parameters in parameter β was sorted from large to small and then performed the classification experiments with the largest top n (range from 1 to 2527) weight parameters in turn. The relationship between the prediction accuracy and the number of weight parameters used in the prediction was then recorded in a chart. When the accuracy reached the maximum and remained unchanged, we obtained all the key voxels that affected the prediction results.

In addition, we constructed a brain region weight map based on the Automated Anatomical Labeling (AAL) [64] atlas using the following steps. First, we mapped the weight parameters to their corresponding voxels on a DARTEL template and obtained a coarse weight distribution map with voxel size 8x8x8 mm^3. Second, the coarse weight distribution map was transformed to a finer weight distribution map with voxel size 1x1x1 mm^3 by the nearest-neighbor interpolation. This finer weight distribution map had the same dimensions as the AAL template. Third, we mapped the finer weight distribution map to the AAL template by multiplying the corresponding elements and calculated the sum of weight values in each brain region. These values were defined as the weight value of each brain region. Finally, according to the weight value of each brain region in AAL, we constructed a brain region weight map and calculated the relative weight ratio of each brain region.

Correlation analysis

To explore the relationships between cognitive impairment (MMSE scores) and mean voxels of several regions obtained from GSplit LBI in AD patients, a partial correlation analysis was performed, with age, gender and education being used as nuisance covariates (SPSS20, $P < 0.05$).

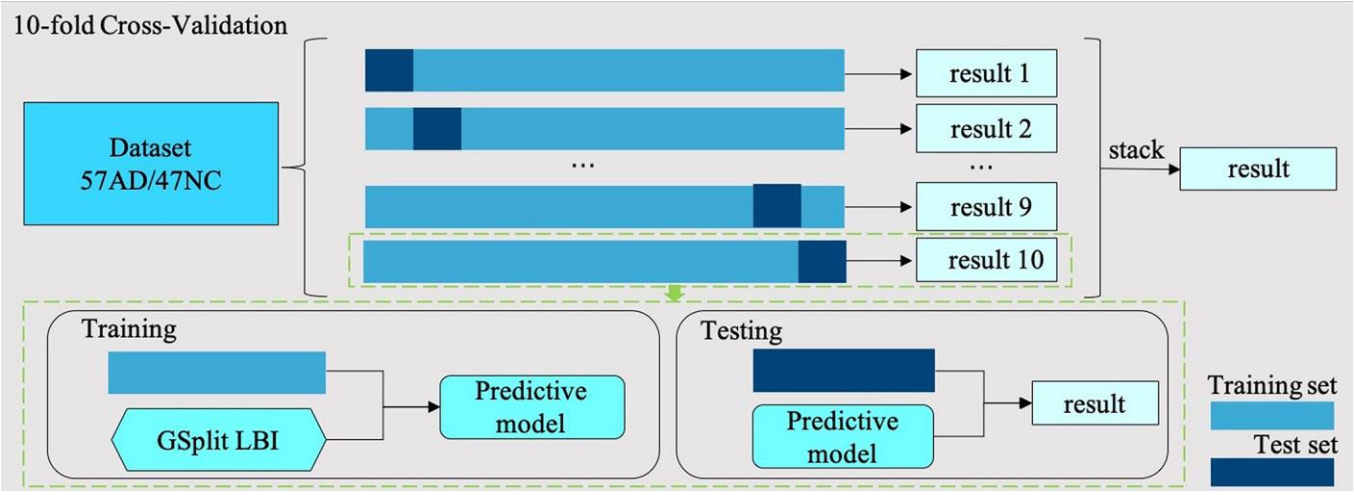


Figure 6. 10-fold cross-validation flow chart. In each cross-validation step, the dataset with 57 AD samples and 47 NC samples is divided into ten subsets, nine of which are used as training set and the rest as test set. In the training step, we use the training set training GSplit LBI model. In the test step, we use the trained model to predict the test set. Finally, the results of ten folds are stacked together as the results of this 10-fold cross-validation step.

Reproducibility and Stability analysis

To verify the generalization of model prediction and the stability of feature selection, we performed a cross-test on our in-house dataset and the public ADNI dataset. We used the model parameters trained from this experiment to test the ADNI dataset (to facilitate the reference later, here we name it the Chinese model); we also used the model parameters trained from the ADNI dataset of our previous work [30] to test the current dataset (to facilitate the reference later, here we name it the ADNI model).

In this cross-test experiment, we used the same ADNI subjects as our previous work [30] In the ADNI dataset, all subjects are divided into 1.5 Tesla and 3.0 Tesla MRI scan datasets. For comparison, we chose the 3.0T ADNI dataset in our experiment, which contains a total of 176 subjects (66 AD patients and 110 NCs) with ages ranging from 55-90 years old. The average educational levels of AD and NC are 15.59 and 16.65 respectively; meanwhile, the average of MMSE scores for AD and NC are 23.11 and 28.93, respectively. Subject IDs are shown in the Supplemental material in our previous paper [30].

Abbreviations

AD: Alzheimer disease**;** sMRI: structural magnetic resonance imaging; GSplit LBI: Generalized Split Linearized Bregman Iteration; NC: normal control; ADNI: Alzheimer's disease Neuroimaging Initiative; ROI: region of interest; MCI: mild cognitive impairment; SVM: support vector machine; LDA: linear discriminant analysis; LPBM: spatially augmented linear programming boosting method; MLDA: Maximum uncertainty Linear Discriminant Analysis; TV $+l_1$: Total Variation with l1 penalty; n^2GFL: Nonnegative Generalized Fused Lasso; MMSE: mini-mental state examination; CDR: clinical dementia rating; GM: gray matter; sMCI: stable MCI; pMCI: progressive MCI; PET: positron emission tomography; CSF: cerebrospinal fluid; DARTEL: Diffeomorphic Anatomical Registration Exponentiated Lie Algebra; WM: white matter; MNI: Montreal Neurological Institute; ROC: receiver operating characteristic curve; AUC: area under the curve; M1: motor cortex area; MTG: middle temporal gyrus; ANG: angular gyrus; SMG: supramarginal gyrus; IT: inferior temporal gyrus; STG: superior temporal gyrus; PCUN: precuneus; CAL: calcarine fissure and surrounding cortex; THA: thalamus; fMRI: functional magnetic resonance imaging; rs-fMRI: resting state functional magnetic resonance imaging; SMN: sensorimotor**.**

AUTHOR CONTRIBUTIONS

ZQW, LJH, KCL and WMZ were responsible for the study concept and design. XH contributed to the acquisition of MRI data. WMZ, ZYS, XLL and YY assisted with data analysis and interpretation of data. WMZ, BC, ZYS and ZQW drafted the manuscript. All authors provided critical revision of the manuscript for important intellectual content. All authors critically reviewed content and approved final version for publication.

ACKNOWLEDGMENTS

The authors thank the patients and their families for the time and effort they dedicated to the research. In addition, we are grateful to our colleagues from Deepwise AI Lab, Beijing, China. Also we would like to thank Stephanie A Yu from Hong Kong West Island School for editing the English text of this manuscript.

During the cross-test, we used the ADNI dataset. ADNI Data collection and sharing was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co.,Medpace, Inc.,Merck and Co., Inc., Novartis AG, Pfizer Inc., F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (https://fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles.

CONFLICTS OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING

This work was supported by the National natural scientific foundation of China (Grant No. 81571648, 81873892), the Key project of the National Natural Science Foundation of China (No. 81830057), the

scientific research fund of Aerospace Center Hospital (No. YN201901), the scientific research fund of China Aerospace Science and Industry Corp (No. 2019-LCYL-010), Beijing natural scientific foundation of China (No.7182105).

REFERENCES

1. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 1991; 82:239–59. https://doi.org/10.1007/BF00308809 PMID:1759558

2. Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. Neuroimage. 2017; 155:530–48. https://doi.org/10.1016/j.neuroimage.2017.03.057 PMID:28414186

3. Vieira S, Pinaya WH, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. Neurosci Biobehav Rev. 2017; 74:58–75. https://doi.org/10.1016/j.neubiorev.2017.01.002 PMID:28087243

4. Baron JC, Chételat G, Desgranges B, Perchey G, Landeau B, de la Sayette V, Eustache F. In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. Neuroimage. 2001; 14:298–309. https://doi.org/10.1006/nimg.2001.0848 PMID:11467904

5. Busatto GF, Garrido GE, Almeida OP, Castro CC, Camargo CH, Cid CG, Buchpiguel CA, Furuie S, Bottino CM. A voxel-based morphometry study of temporal lobe gray matter reductions in Alzheimer's disease. Neurobiol Aging. 2003; 24:221–31. https://doi.org/10.1016/S0197-4580(02)00084-2 PMID:12498956

6. Frisoni GB, Testa C, Zorzan A, Sabattoli F, Beltramello A, Soininen H, Laakso MP. Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. J Neurol Neurosurg Psychiatry. 2002; 73:657–64. https://doi.org/10.1136/jnnp.73.6.657 PMID:12438466

7. Ishii K, Kawachi T, Sasaki H, Kono AK, Fukuda T, Kojima Y, Mori E. Voxel-based morphometric comparison between early- and late-onset mild Alzheimer's disease and assessment of diagnostic performance of z score images. AJNR Am J Neuroradiol. 2005; 26:333–40. PMID:15709131

8. Frisoni GB, Fox NC, Jack CR Jr, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010; 6:67–77. https://doi.org/10.1038/nrneurol.2009.215 PMID:20139996

9. Vemuri P, Jack CR Jr. Role of structural MRI in Alzheimer's disease. Alzheimers Res Ther. 2010; 2:23. https://doi.org/10.1186/alzrt47 PMID:20807454

10. Chetelat G, Baron JC. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. Neuroimage. 2003; 18:525–41. https://doi.org/10.1016/S1053-8119(02)00026-5 PMID:12595205

11. Lerch JP, Pruessner J, Zijdenbos AP, Collins DL, Teipel SJ, Hampel H, Evans AC. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. Neurobiol Aging. 2008; 29:23–30. https://doi.org/10.1016/j.neurobiolaging.2006.09.013 PMID:17097767

12. Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. Neuroimage. 2009; 44:1415–22. https://doi.org/10.1016/j.neuroimage.2008.10.031 PMID:19027862

13. Park H, Yang JJ, Seo J, Lee JM, and ADNI. Dimensionality reduced cortical features and their use in predicting longitudinal changes in Alzheimer's disease. Neurosci Lett. 2013; 550:17–22. https://doi.org/10.1016/j.neulet.2013.06.042 PMID:23827219

14. Apostolova LG, Hwang KS, Kohannim O, Avila D, Elashoff D, Jack CR Jr, Shaw L, Trojanowski JQ, Weiner MW, Thompson PM, and Alzheimer's Disease Neuroimaging Initiative. ApoE4 effects on automated diagnostic classifiers for mild cognitive impairment and Alzheimer's disease. Neuroimage Clin. 2014; 4:461–72. https://doi.org/10.1016/j.nicl.2013.12.012 PMID:24634832

15. Cui Y, Wen W, Lipnicki DM, Beg MF, Jin JS, Luo S, Zhu W, Kochan NA, Reppermund S, Zhuang L, Raamana PR, Liu T, Trollor JN, et al. Automated detection of amnestic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. Neuroimage. 2012; 59:1209–17. https://doi.org/10.1016/j.neuroimage.2011.08.013 PMID:21864688

16. Kohannim O, Hua X, Hibar DP, Lee S, Chou YY, Toga AW, Jack CR Jr, Weiner MW, Thompson PM, and

Alzheimer's Disease Neuroimaging Initiative. Boosting power for clinical trials using classifiers based on multiple biomarkers. Neurobiol Aging. 2010; 31:1429–42.
https://doi.org/10.1016/j.neurobiolaging.2010.04.022
PMID:20541286

17. Padilla P, López M, Górriz JM, Ramírez J, Salas-González D, Álvarez I, and Alzheimer's Disease Neuroimaging Initiative. NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease. IEEE Trans Med Imaging. 2012; 31:207–16.
https://doi.org/10.1109/TMI.2011.2167628
PMID:21914569

18. Möller C, Pijnenburg YA, van der Flier WM, Versteeg A, Tijms B, de Munck JC, Hafkemeijer A, Rombouts SA, van der Grond J, van Swieten J, Dopper E, Scheltens P, Barkhof F, et al. Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. Radiology. 2016; 279:838–48.
https://doi.org/10.1148/radiol.2015150220
PMID:26653846

19. Dyrba M, Grothe M, Kirste T, Teipel SJ. Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. Hum Brain Mapp. 2015; 36:2118–31.
https://doi.org/10.1002/hbm.22759
PMID:25664619

20. Liu F, Zhou L, Shen C, Yin J. Multiple kernel learning in the primal for multimodal Alzheimer's disease classification. IEEE J Biomed Health Inform. 2014; 18:984–90.
https://doi.org/10.1109/JBHI.2013.2285378
PMID:24132030

21. Zu C, Jie B, Liu M, Chen S, Shen D, Zhang D, and Alzheimer's Disease Neuroimaging Initiative. Label-aligned multi-task feature learning for multimodal classification of Alzheimer's disease and mild cognitive impairment. Brain Imaging Behav. 2016; 10:1148–59.
https://doi.org/10.1007/s11682-015-9480-7
PMID:26572145

22. Lillemark L, Sørensen L, Pai A, Dam EB, Nielsen M, and Alzheimer's Disease Neuroimaging Initiative. Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. BMC Med Imaging. 2014; 14:21.
https://doi.org/10.1186/1471-2342-14-21
PMID:24889999

23. Tang X, Holland D, Dale AM, Younes L, Miller MI. Baseline shape diffeomorphometry patterns of subcortical and ventricular structures in predicting conversion of mild cognitive impairment to Alzheimer's disease. J Alzheimers Dis. 2015; 44:599–611.
https://doi.org/10.3233/JAD-141605
PMID:25318546

24. Westman E, Simmons A, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Soininen H, Weiner MW, Lovestone S, Spenger C, Wahlund LO, and AddNeuroMed consortium, and Alzheimer's Disease Neuroimaging Initiative. AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. Neuroimage. 2011; 58:818–28.
https://doi.org/10.1016/j.neuroimage.2011.06.065
PMID:21763442

25. Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J, and Alzheimer's Disease Neuroimaging Initiative. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage. 2015; 104:398–412.
https://doi.org/10.1016/j.neuroimage.2014.10.002
PMID:25312773

26. Casanova R, Whitlow CT, Wagner B, Williamson J, Shumaker SA, Maldjian JA, Espeland MA. High dimensional classification of structural MRI Alzheimer's disease data based on large scale regularization. Front Neuroinform. 2011; 5:22.
https://doi.org/10.3389/fninf.2011.00022
PMID:22016732

27. Liu M, Zhang D, Shen D, and Alzheimer's Disease Neuroimaging Initiative. View-centralized multi-atlas classification for Alzheimer's disease diagnosis. Hum Brain Mapp. 2015; 36:1847–65.
https://doi.org/10.1002/hbm.22741
PMID:25624081

28. Zhang Y, Wang S, Dong Z. Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. Prog Electromagn Res. 2014; 144:171–84.
https://doi.org/10.2528/PIER13121310

29. Hinrichs C, Singh V, Mukherjee L, Xu G, Chung MK, Johnson SC, and Alzheimer's Disease Neuroimaging Initiative. Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. Neuroimage. 2009; 48:138–49.
https://doi.org/10.1016/j.neuroimage.2009.05.056
PMID:19481161

30. Sun X, Hu L, Yuan Y, Wang Y. GSplit LBI: Taming the Procedural Bias in Neuroimaging for Disease Prediction. Medical Image Computing and Computer Assisted Intervention. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins D, Duchesne S (eds). Medical Image Computing and Computer Assisted Intervention – MICCAI 2017. MICCAI; 2017. Lecture Notes in Computer Science, vol 10435, pp. 107–15. Springer,

Cham
https://doi.org/10.1007/978-3-319-66179-7_13

31. Hirata Y, Matsuda H, Nemoto K, Ohnishi T, Hirao K, Yamashita F, Asada T, Iwabuchi S, Samejima H. Voxel-based morphometry to discriminate early Alzheimer's disease from controls. Neurosci Lett. 2005; 382:269–74.
https://doi.org/10.1016/j.neulet.2005.03.038
PMID:15925102

32. Johnson KA, Fox NC, Sperling RA, Klunk WE. Brain imaging in Alzheimer disease. Cold Spring Harb Perspect Med. 2012; 2:a006213.
https://doi.org/10.1101/cshperspect.a006213
PMID:22474610

33. Karas GB, Burton EJ, Rombouts SA, van Schijndel RA, O'Brien JT, Scheltens P, McKeith IG, Williams D, Ballard C, Barkhof F. A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. Neuroimage. 2003; 18:895–907.
https://doi.org/10.1016/S1053-8119(03)00041-7
PMID:12725765

34. Whitwell JL, Jack CR Jr. Comparisons between Alzheimer disease, frontotemporal lobar degeneration, and normal aging with brain mapping. Top Magn Reson Imaging. 2005; 16:409–25.
https://doi.org/10.1097/01.rmr.0000245457.98029.e1
PMID:17088691

35. Ashburner J, Friston KJ. Why voxel-based morphometry should be used. Neuroimage. 2001; 14:1238–43.
https://doi.org/10.1006/nimg.2001.0961
PMID:11707080

36. Jagust W. Vulnerable neural systems and the borderland of brain aging and neurodegeneration. Neuron. 2013; 77:219–34.
https://doi.org/10.1016/j.neuron.2013.01.002
PMID:23352159

37. Schmitter D, Roche A, Maréchal B, Ribes D, Abdulkadir A, Bach-Cuadra M, Daducci A, Granziera C, Klöppel S, Maeder P, Meuli R, Krueger G, and Alzheimer's Disease Neuroimaging Initiative. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. Neuroimage Clin. 2014; 7:7–17.
https://doi.org/10.1016/j.nicl.2014.11.001
PMID:25429357

38. Jeong J. EEG dynamics in patients with Alzheimer's disease. Clin Neurophysiol. 2004; 115:1490–505.
https://doi.org/10.1016/j.clinph.2004.01.001
PMID:15203050

39. Thomann PA, Schläfer C, Seidl U, Santos VD, Essig M, Schröder J. The cerebellum in mild cognitive impairment and Alzheimer's disease - a structural MRI study. J Psychiatr Res. 2008; 42:1198–202.
https://doi.org/10.1016/j.jpsychires.2007.12.002
PMID:18215400

40. Bas O, Acer N, Mas N, Karabekir HS, Kusbeci OY, Sahin B. Stereological evaluation of the volume and volume fraction of intracranial structures in magnetic resonance images of patients with Alzheimer's disease. Ann Anat. 2009; 191:186–95.
https://doi.org/10.1016/j.aanat.2008.12.003
PMID:19269145

41. Larner AJ. The cerebellum in Alzheimer's disease. Dement Geriatr Cogn Disord. 1997; 8:203–09.
https://doi.org/10.1159/000106632
PMID:9213064

42. Ciavardelli D, Silvestri E, Del Viscovo A, Bomba M, De Gregorio D, Moreno M, Di Ilio C, Goglia F, Canzoniero LM, Sensi SL. Alterations of brain and cerebellar proteomes linked to Aβ and tau pathology in a female triple-transgenic murine model of Alzheimer's disease. Cell Death Dis. 2010; 1:e90.
https://doi.org/10.1038/cddis.2010.68
PMID:21368863

43. Sepulveda-Falla D, Matschke J, Bernreuther C, Hagel C, Puig B, Villegas A, Garcia G, Zea J, Gomez-Mancilla B, Ferrer I, Lopera F, Glatzel M. Deposition of hyperphosphorylated tau in cerebellum of PS1 E280A Alzheimer's disease. Brain Pathol. 2011; 21:452–63.
https://doi.org/10.1111/j.1750-3639.2010.00469.x
PMID:21159009

44. Strata P. The emotional cerebellum. Cerebellum. 2015; 14:570–77.
https://doi.org/10.1007/s12311-015-0649-9
PMID:25626523

45. Agosta F, Rocca MA, Pagani E, Absinta M, Magnani G, Marcone A, Falautano M, Comi G, Gorno-Tempini ML, Filippi M. Sensorimotor network rewiring in mild cognitive impairment and Alzheimer's disease. Hum Brain Mapp. 2010; 31:515–25.
https://doi.org/10.1002/hbm.20883
PMID:19777557

46. Vidoni ED, Thomas GP, Honea RA, Loskutova N, Burns JM. Evidence of altered corticomotor system connectivity in early-stage Alzheimer's disease. J Neurol Phys Ther. 2012; 36:8–16.
https://doi.org/10.1097/NPT.0b013e3182462ea6
PMID:22333920

47. Brier MR, Thomas JB, Snyder AZ, Benzinger TL, Zhang D, Raichle ME, Holtzman DM, Morris JC, Ances BM. Loss of intranetwork and internetwork resting state functional connections with Alzheimer's disease

progression. J Neurosci. 2012; 32:8890–99. https://doi.org/10.1523/JNEUROSCI.5698-11.2012 PMID:22745490

48. Chételat G, Landeau B, Eustache F, Mézenge F, Viader F, de la Sayette V, Desgranges B, Baron JC. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. Neuroimage. 2005; 27:934–46. https://doi.org/10.1016/j.neuroimage.2005.05.015 PMID:15979341

49. Sorg C, Riedl V, Mühlau M, Calhoun VD, Eichele T, Läer L, Drzezga A, Förstl H, Kurz A, Zimmer C, Wohlschläger AM. Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. Proc Natl Acad Sci USA. 2007; 104:18760–65. https://doi.org/10.1073/pnas.0708803104 PMID:18003904

50. Pedro T, Weiler M, Yasuda CL, D'Abreu A, Damasceno BP, Cendes F, Balthazar ML. Volumetric brain changes in thalamus, corpus callosum and medial temporal structures: mild Alzheimer's disease compared with amnestic mild cognitive impairment. Dement Geriatr Cogn Disord. 2012; 34:149–55. https://doi.org/10.1159/000342118 PMID:22986782

51. Yi HA, Möller C, Dieleman N, Bouwman FH, Barkhof F, Scheltens P, van der Flier WM, Vrenken H. Relation between subcortical grey matter atrophy and conversion from mild cognitive impairment to Alzheimer's disease. J Neurol Neurosurg Psychiatry. 2016; 87:425–32. https://doi.org/10.1136/jnnp-2014-309105 PMID:25904810

52. de Oliveira MS, Balthazar ML, D'Abreu A, Yasuda CL, Damasceno BP, Cendes F, Castellano G. MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease. AJNR Am J Neuroradiol. 2011; 32:60–66. https://doi.org/10.3174/ajnr.A2232 PMID:20966061

53. de Jong LW, van der Hiele K, Veer IM, Houwing JJ, Westendorp RG, Bollen EL, de Bruin PW, Middelkoop HA, van Buchem MA, van der Grond J. Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. Brain. 2008; 131:3277–85. https://doi.org/10.1093/brain/awn278 PMID:19022861

54. Beheshti I, Demirel H, Matsuda H, and Alzheimer's Disease Neuroimaging Initiative. 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. 2017; 83:109–19. https://doi.org/10.1016/j.compbiomed.2017.02.011 PMID:28260614

55. Spasov S, Passamonti L, Duggento A, Liò P, Toschi N, and Alzheimer's Disease Neuroimaging Initiative. A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. Neuroimage. 2019; 189:276–87. https://doi.org/10.1016/j.neuroimage.2019.01.031 PMID:30654174

56. Cao P, Liu X, Yang J, Zhao D, Huang M, Zhang J, Zaiane O. Nonlinearity-aware based dimensionality reduction and over-sampling for AD/MCI classification from MRI measures. Comput Biol Med. 2017; 91:21–37. https://doi.org/10.1016/j.compbiomed.2017.10.002 PMID:29031664

57. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, Delacourte A, Galasko D, Gauthier S, Jicha G, Meguro K, O'Brien J, Pasquier F, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007; 6:734–46. https://doi.org/10.1016/S1474-4422(07)70178-3 PMID:17616482

58. Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, et al. Revising the definition of Alzheimer's disease: a new lexicon. Lancet Neurol. 2010; 9:1118–27. https://doi.org/10.1016/S1474-4422(10)70223-4 PMID:20934914

59. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993; 43:2412–14. https://doi.org/10.1212/WNL.43.11.2412-a PMID:8232972

60. Peters M, Jäncke L, Staiger JF, Schlaug G, Huang Y, Steinmetz H. Unsolved problems in comparing brain sizes in Homo sapiens. Brain Cogn. 1998; 37:254–85. https://doi.org/10.1006/brcg.1998.0983 PMID:9665746

61. Leonard CM, Towler S, Welcome S, Halderman LK, Otto R, Eckert MA, Chiarello C. Size matters: cerebral volume influences sex differences in neuroanatomy. Cereb Cortex. 2008; 18:2920–31. https://doi.org/10.1093/cercor/bhn052 PMID:18440950

62. Im K, Lee JM, Lyttelton O, Kim SH, Evans AC, Kim SI. Brain size and cortical structure in the adult human brain. Cereb Cortex. 2008; 18:2181–91. https://doi.org/10.1093/cercor/bhm244 PMID:18234686

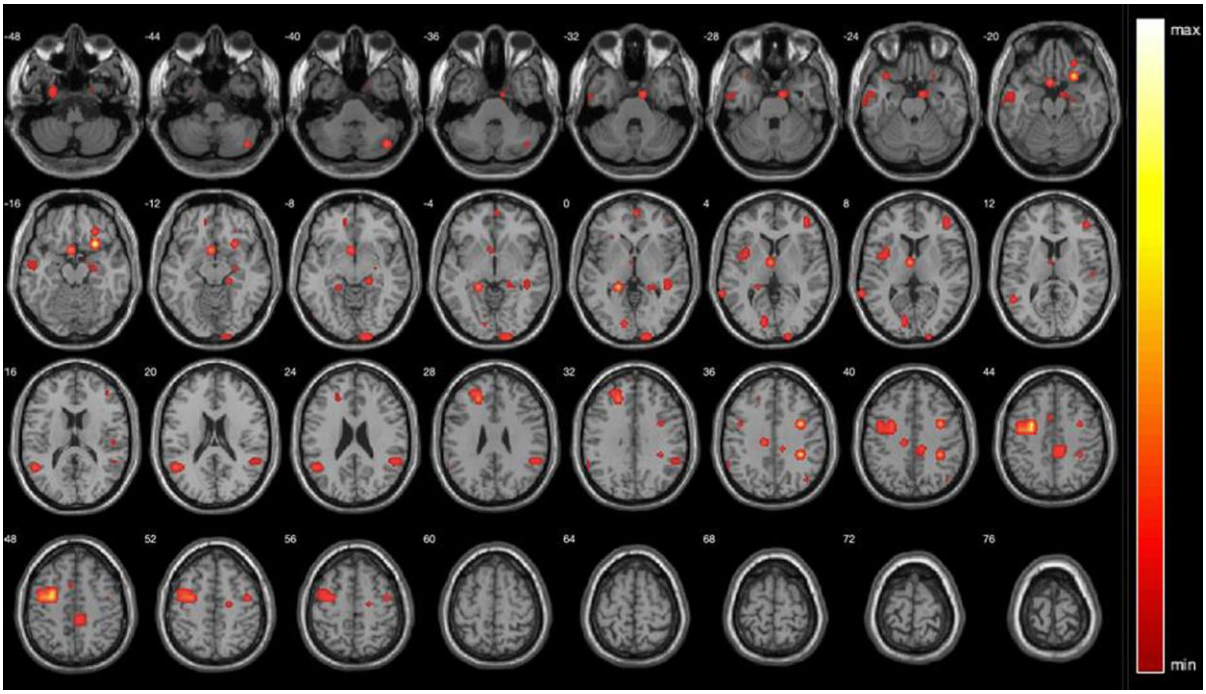
63. Ashburner J. A fast diffeomorphic image registration

algorithm. Neuroimage. 2007; 38:95–113. https://doi.org/10.1016/j.neuroimage.2007.07.007 PMID:17761438

64. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002; 15:273–89. https://doi.org/10.1006/nimg.2001.0978 PMID:11771995

SUPPLEMENTARY MATERIALS

Supplementary Figure



Supplementary Figure 1. Weight distribution map to classification between AD patients and NCs which is reconstructed by combining voxel information that is entered into the model and top 6% of the model parameters. The color bar represents the weight value form GSplit LBI model, the larger the weight value of the model, the warmer the color in the graph.

Supplementary Tables

Supplementary Table 1. Performance (accuracy, sensitivity, specificity) of ten experiments, each experiment is a ten-fold cross-validation.

	accuracy	sensitivity	specificity
1st	90.45%	91.67%	88.50%
2nd	89.55%	91.00%	87.50%
3rd	91.36%	91.67%	92.00%
4th	90.45%	91.67%	90.00%
5th	90.45%	91.00%	89.50%
6th	90.55%	91.00%	89.50%
7th	89.55%	91.33%	87.00%
8th	90.45%	89.67%	91.50%
9th	91.36%	91.33%	91.00%
10th	90.27%	91.33%	88.50%
average	90.44%	91.17%	89.50%

Supplementary Table 2. In previous study, the GSplit LBI-based classifier had been compared with other classifiers including MLDA, SVM, Lasso, Graphnet, Elastic Net, TV+l1 and n^2GFL and obtained better results than other models on ADNI dataset.

	SVM	Lasso	Elastic Net	MLDA	Graphnet	TV+l1	n^2GFL	GSplit LBI
ADNI	87.50%	87.50%	89.2%	86.93%	88.64%	87.50%	87.50%	90.91%
In-house	87.20%	87.20%	89.50%	-	-	-	-	90.44%

In this study, some comparative experiments including SVM, Lasso and Elastic Net also have been done to prove the performance of the GSplit LBI-based model on our in-house dataset. The result of comparative experiments between GSplit LBI-based classifier and other classifiers on ADNI dataset and our in-house dataset are shown in this table.