

Identification of differential brain regions in MCI progression via clustering-evolutionary weighted SVM ensemble algorithm

Xia-an BI (✉)^{1,2}, Yiming XIE^{1,2}, Hao WU^{1,2}, Luyun XU (✉)³

1 Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing,
Hunan Normal University, Changsha 410081, China

2 College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China

3 Business School of Hunan Normal University, Changsha 410081, China

© Higher Education Press 2020

Abstract Mild cognitive impairment (MCI) as the potential sign of serious cognitive decline could be divided into two stages, i.e., late MCI (LMCI) and early MCI (EMCI). Although the different cognitive states in the MCI progression have been clinically defined, effective and accurate identification of differences in neuroimaging data between these stages still needs to be further studied. In this paper, a new method of clustering-evolutionary weighted support vector machine ensemble (CEWSVME) is presented to investigate the alterations from cognitively normal (CN) to EMCI to LMCI. The CEWSVME mainly includes two steps. The first step is to build multiple SVM classifiers by randomly selecting samples and features. The second step is to introduce the idea of clustering evolution to eliminate inefficient and highly similar SVMs, thereby improving the final classification performances. Additionally, we extracted the optimal features to detect the differential brain regions in MCI progression, and confirmed that these differential brain regions changed dynamically with the development of MCI. More exactly, this study found that some brain regions only have durative effects on MCI progression, such as parahippocampal gyrus, posterior cingulate gyrus and amygdala, while the superior temporal gyrus and the middle temporal gyrus have periodic effects on the progression. Our work contributes to understanding the pathogenesis of MCI and provide the guidance for its timely diagnosis.

Keywords machine learning, MCI progression, optimal feature extraction, differential brain regions, functional magnetic resonance imaging

1 Introduction

Mild cognitive impairment (MCI) posited as the clinical precursor of Alzheimer's disease (AD) represents a slight but detectable cognitive decline that is unusual for someone's age and education [1–3]. It has been reported that MCI patients have a relatively higher likelihood to develop AD dementia with an annual rate of 10%–15%, while cognitively normal (CN) progress

to AD with an estimated conversion rate of 1%–2% annually [4–6]. Owing to a high risk of conversion to AD, MCI has attracted increasing interest in the neuroimaging study.

In accordance with the Alzheimer's Disease Neuroimaging Initiative (ADNI) protocol, MCI is stratified between two levels, i.e., late MCI (LMCI) and early MCI (EMCI) [7]. The patients diagnosed with LMCI are similar to EMCI subjects but have severer objective memory impairment [8,9]. Multi-stage analysis in MCI is of great significance in distinguishing the different cognitive stages of MCI progression and exploring brain pathological changes.

Brain network is the visualization of brain functional or structural connections, which can characterize brain activities in different stages of neurodegenerative diseases. In the construction approaches of brain network, functional magnetic resonance imaging (fMRI) technology has become a mainstream method due to its high authenticity. As the changes of cognitive ability during the MCI progression are related to the interaction of many nodes in the brain network, it is particularly important to study the topological structure of the brain network which may provide objective intermediate markers for different stages of MCI [10]. Therefore, it is a promising approach to explore the MCI progression by analyzing the graph theory measures of brain network based on fMRI.

In order to identify the changes of brain network during the MCI progression, a large number of studies have been carried out. These studies have confirmed the altered functional connectivities or disruptive topological structures in the brain network of MCI patients [11,12]. Specifically, in the progression of MCI, the connection interruptions among brain network nodes are common, and the distribution of these interruptions is selective, which is easy to generate aggregation in default mode network and visual network [13,14]. Furthermore, researchers discovered that the abnormal brain network also changes dynamically with the development of MCI [15]. In addition, based on the analysis of brain network, some potential brain regions which play a key role in the MCI progression are also detected [16]. These studies provide a basis for the in-depth analysis of MCI and AD. However, it is worth noting that most studies

may underestimate the changes of brain networks in different stages of MCI and make undifferentiated researches on EMCI and LMCI owing to the similarity of brain networks in different MCI stages, which may miss important information in the development of MCI [17, 18].

To overcome such a potential drawback, an advanced clustering-evolutionary weighted support vector machine (SVM) ensemble (CEWSVME) is presented to detect the subtle differences in CN vs. EMCI and EMCI vs. LMCI respectively. The CEWSVME model integrates multiple SVMs and uses clustering evolution to enhance the differences between SVMs, which can obtain highly stable and robust classification results compared to traditional machine learning methods. With this method, we carried out independent experiments on the two stages of MCI development (e.g., CN to EMCI, EMCI to LMCI). The detection accuracy of EMCI and LMCI samples reached 83.47% and 84.30%, respectively. Moreover, some regions such as parahippocampal gyrus, posterior cingulate gyrus and amygdala are simultaneously detected in the two sets of experiments, indicating that these regions have continuous impact on the whole MCI progression. But there are some other regions that only occur in EMCI or LMCI stage such as superior temporal gyrus (STG.L) and middle temporal gyrus, which may serve as biomarkers for distinguishing EMCI and LMCI, and assist to understand the subtle changes of the brain in different cognitive states.

2 Materials and methods

Figure 1 displays the overall framework of our study for predicting class labels of subjects and detecting abnormal brain regions. More precisely, after fMRI data acquisition and preprocessing, the brain is divided into 90 regions of interest (ROIs) by anatomical automatic labeling (AAL) atlas, which are used to construct the brain connectivity matrix. The graph measures are calculated and used as the original inputs for CEWSVME classifier. We use the CEWSVME classifier to identify different stages of the MCI progression and extract optimal feature subset. The differential brain regions related to the detected optimal

features. In the following subsection, we mainly introduce the CEWSVME algorithm and its application in classification and differential brain region detection.

2.1 Dataset

The available experimental data is a subset collected by the publicly released ADNI database [19], which collects data on AD, EMCI, LMCI and elderly control, including MRI, blood biomarkers and others, to support early diagnosis, intervention, prevention, and treatment of AD. Briefly, there were 36 CN subjects (17 male, 74.46 ± 4.15 years old), 42 EMCI patients (18 male, 72.34 ± 6.87 years old) and 38 LMCI patients (23 male, 72.99 ± 7.79 years old) in our study. The CN group was free from other neurological diseases and drug abuse. In addition, two types of clinical scores for all participants were acquired, which were mini-mental state examination (MMSE) and clinical dementia rating (CDR). The two-sample T test and Chi-squared test were performed on the continuous variables (age, MMSE and CDR) and categorical variable (gender) respectively. The demographic information of the participants was exhibited in Table 1. This research was implemented according to the Internal Conference on Harmonization, etc. All subjects have signed informed consent.

The fMRI scans for all participants were performed using the Siemens TRIO 3 Tesla machine based on the following sequence parameters: resolution = 64×64 matrix, pixel spacing [PS] Y/X = $3.313/3.313$ mm, repetition time [TR]/ flip angle [FA]/echo time [TE] = $3s/80^\circ/30ms$, slices number = 48, interslice gap = 0 mm, slice thickness = 3.313 mm and volume number = 140.

Table 1 The detailed demographic characteristic data of the subjects

Variable	CN	EMCI	LMCI	P value
Male/Female	17/19	18/24	23/15	0.70 ^a /0.11 ^b
Age (mean \pm sd)	74.46 \pm 4.15	72.34 \pm 6.87	72.99 \pm 7.79	0.11 ^a /0.69 ^b
MMSE (mean \pm sd)	28.90 \pm 1.66	28.10 \pm 1.57	27.11 \pm 2.44	0.02 ^a /0.03 ^b
CDR (mean \pm sd)	0.695 \pm 0.21	0.45 \pm 0.22	0.54 \pm 0.14	0.00 ^a /0.04 ^b

^a the P value represents the group comparison between CN and EMCI

^b the P value represents the EMCI and LMCI comparison

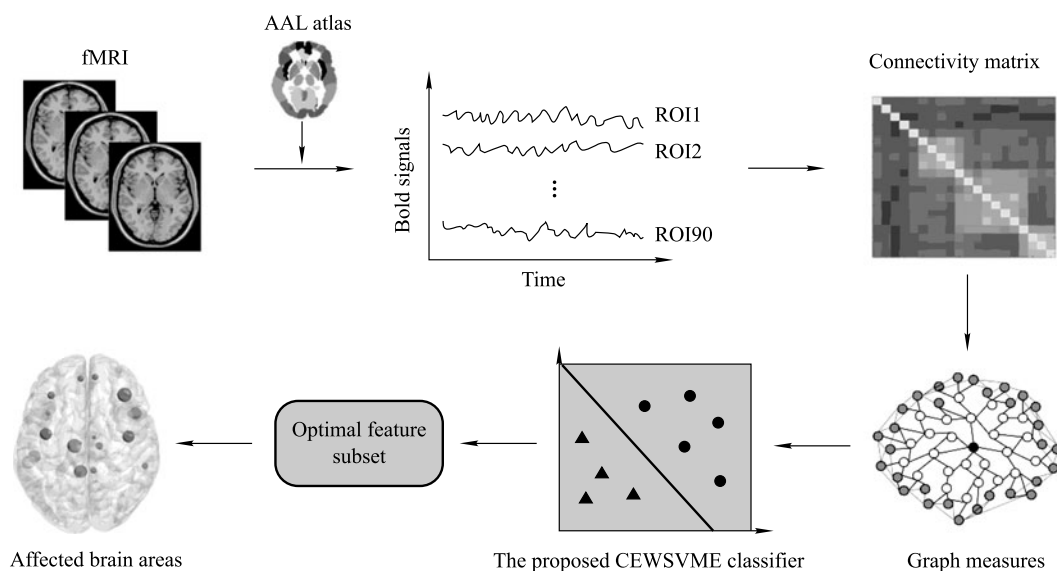


Fig. 1 The overall framework of our study

2.2 Image preprocessing and feature construction

Standard preprocessing procedures were conducted by data processing assistant for resting-state fMRI (DPARSF). Specifically, the preprocessing pipeline for every participant was made up of the following conventional steps: transforming the data into NIFTI format; discarding the top ten volumes of functional images; slice timing; head motion correction to meet the criteria of spatial movement $<2.5^\circ$ rotation or 2.5 mm translation; spatial transformation to the EPI space; Gaussian spatial smoothing using 6-mm full-width; elimination of the signal with linear time trend; temporal band-pass filtering with 0.01–0.08Hz frequency range and regression of several nuisance signals.

The human brain network is an intricate and dynamic network that supports highly effective information communication, which could be mathematically modeled as a graph composed of two fundamental elements (vertices and edges). To build the whole-brain functional network, we firstly use the AAL atlas to generate 90 automatically parcellated ROIs, which constitute the 90 vertices in the brain network. The edge is considered as the proxy of interregional resting-state functional connectivity and calculated by the Pearson correlation coefficients for any two vertices' time series. Therefore, the brain information of each sample is abstracted into a brain functional connection network, which is represented by a 90×90 matrix. The network is represented as $G = (D, E)$, where D denotes vertices and E denotes edges. Since Pearson correlation coefficient is directly used as the edge, the absolute value of smaller Pearson correlation coefficient means that the edge is not stable, which may decrease the robustness of the network. We further apply Eq. (1) to dealing with the edges.

$$E = \begin{cases} 0, & \text{if the absolute value of the edge} < w, \\ 1, & \text{if the absolute value of the edge} \geq w, \end{cases} \quad (1)$$

where w represents the threshold. If the absolute value of the edge is greater than w , the edge value is reassigned to 1. Otherwise it is reassigned to 0. Consequently, the undirected binary network of brain is acquired.

In the presented study, four typical local graph measures that quantify different network characteristics are considered in the above undirected binary network, which are respectively shortest path, local efficiency, degree and clustering coefficient. For each participant, a total of 4275 topological measures (including 4005 shortest paths, 90 clustering coefficients, 90 degrees and 90 local efficiencies) are acquired.

2.3 Clustering-evolutionary weighted SVM ensemble

It is commonly accepted that the small-subject size of high-dimensional medical imaging makes it difficult for the conventional single learner to accurately identify the subtle brain changes between groups. In this paper, a new ensemble learning method based on SVM is proposed, which is called clustering-evolutionary weighted SVM ensemble. Besides constructing a simple combination of SVMs, the model carries out dynamic selection of SVMs through hierarchical clustering, which maximizes the diversities among SVMs and improves the generalization performance of ensemble learner. The design idea of the method is shown in Fig. 2, and the training process is as follows.

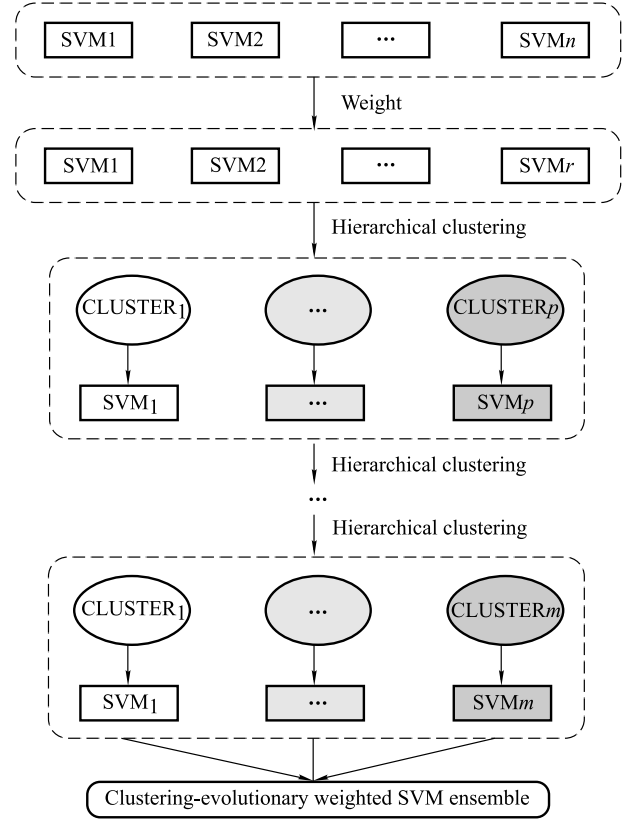


Fig. 2 A schematic diagram of the CEWSVME

Firstly, the 1:3 ratio is used to split all available data into the test examples and the “training & validation” examples. The validation examples and training examples are randomly extracted from the “training & validation” examples according to the ratio of 2:1 each time, and the $\sqrt{4275} = 65$ features are stochastically selected out from 4275 features to constitute input variables. These training examples and input variables are used to construct an SVM model. The cost parameter C and bandwidth γ of SVM are set to Inf and 3 respectively, and its performance is assessed by the remaining validation examples. Then, the step of building a single SVM is repeated for n times to get multiple SVMs that are randomly constructed. Then, we take the classification accuracies of the validation set as the weights to weight each SVM, and remove the SVM whose weight is lower than the specified threshold to get the weighted SVM ensemble which is called as the initial SVM ensemble. In order to improve the performance of the initial ensemble learner, the clustering evolution is carried out. The detailed process is as follows.

This paper introduces the disagreement measure as clustering index which could estimate the similarity of classification results produced by two classifiers. Specifically, the disagreement measure between SVM_h and SVM_k could be calculated according to Eq. (2).

$$D_{hk} = \frac{z_{01} + z_{10}}{z_{11} + z_{01} + z_{10} + z_{00}}. \quad (2)$$

We assume that z_{10} represents the quantity of examples correctly identified by SVM_h but misclassified by SVM_k . Similarly, z_{01} represents the quantity of examples correctly identi-

fied by SVM_k but misclassified by SVM_h . Lastly, z_{00} denotes the quantity of examples misclassified by SVM_h and SVM_k simultaneously, and z_{11} stands for the quantity of examples correctly identified by SVM_h and SVM_k simultaneously. The value of D_{hk} is between 0 and 1. The smaller the value is, the greater the similarity is.

The different SVMs in the original ensemble learner are grouped into different clusters according to the disagreement measure, and the SVM having the highest accuracy rate in each cluster is left. The process of clustering evolution will be repeated several times, and the number of classes will decrease at the step size of v in per evolution. When the number of clusters makes the classification performance of ensemble learners optimal, the clustering evolution is stopped and the final CEWSVME is obtained.

2.4 Feature selection

Clustering evolution is not only a crucial process to promote the performance of the ensemble learner, but also the preparation for feature selection. In this subsection, we mainly discuss how to search for features with strong identification abilities. Firstly, we extract the SVM base classifiers from the final CEWSVME and the features selected by these SVMs are retained. Then we calculate the frequencies of retained features and arrange the features in the decreasing order. The top u features are extracted as the important ones ($u < \text{the total number of features}$). Therefore, the feature extraction is a dimensionality reduction method which extracts the u important features from all the features.

Feature selection is an approach that further filters out some redundant features from the important ones, and the optimal feature subset most likely involved in disease pathology is obtained. Initially, we select out top q features from the important ones, where q is not more than the important feature number u . Subsequently, the 65 features are stochastically selected out from the top q ones and treated as input variables to construct the conventional SVM ensemble without clustering evolution. The quantity of SVMs in ensemble learner is set to the optimal amount of SVMs. The optimal q could be determined when the SVM ensemble achieved its highest accuracy. These top q dimensional features form the optimal feature subset.

2.5 The differential brain region identification

In the step of detecting differential brain regions, we extract the specific brain regions which are associated with the above-mentioned optimal subset of features during the feature selection. These brain regions are sorted in the decreasing order in accordance with their frequencies. The higher the frequency of a brain region is, the greater its abnormal degree is. In this paper, we pay close attention to the distinctive brain regions that are ranked higher in frequency in each group of experiment.

3 Results

3.1 Discovering differential brain regions between CN and EMCI

3.1.1 CEWSVME construction and performance comparison
The experiment 1 was designed for CN vs. EMCI. In the CN/EMCI classification, ensemble learners with different quantities of SVMs were likely to yield different classification ac-

curacies. Accordingly, the optimal SVM amount needed to be determined. The diagnostic accuracy rates of ensemble learners with various SVM amounts are shown in Fig. 3. It is learned that when the amount of the SVMs was equals to 400, the accuracy of CEWSVME peaked at 89.47%. As the hierarchical clustering continued, the accuracy of CEWSVME gradually decreased. Therefore, we set 400 as the optimal SVM number in our algorithm for the next analysis.

In order to validate the effectiveness of the CEWSVME algorithm in CN vs. EMCI, we performed the accuracy comparison of random forest (RF), random SVM cluster (RSVMC) and CEWSVME. The amounts of base classifiers for three models were set to 400. Figure 4 shows the 50 classification results for three algorithms. It can be observed that the CEWSVME algorithm has the higher average accuracy than the RF and RSVMC, and the interquartile range (IQR) of accuracy reported by the CEWSVME is smaller. The average accuracy, average sensitivity and average specificity of CEWSVME reach 83.47%, 90.90% and 87.50% respectively (please see Table 2). These results demonstrate that the proposed CEWSVME algorithm was

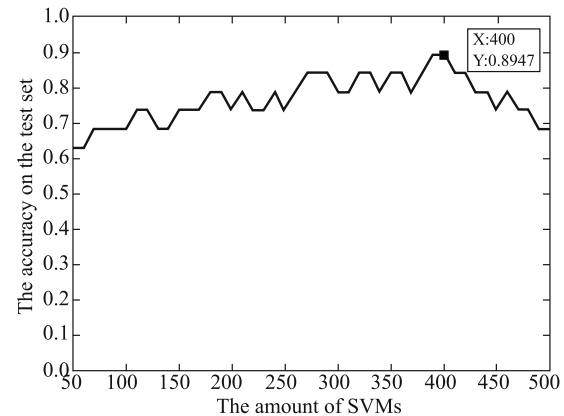


Fig. 3 Exploration of the optimal SVM amount of the CEWSVME in CN vs. EMCI

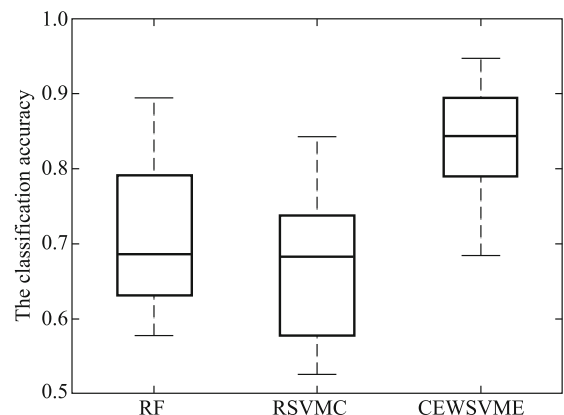


Fig. 4 The accuracy comparison of random forest, random SVM cluster and CEWSVME in CN vs. EMCI

Table 2 Classification performance comparison of three models

Model	Accuracy/%	Sensitivity/%	Specificity/%
RF	71.05±8.85	75.00±9.57	72.73±9.95
RSVMC	67.26±9.71	77.80±8.89	66.67±10.16
CEWSVME	83.47±8.03	90.90±8.84	87.50±9.56

The values are shown by mean ± standard deviation.

Table 3 The partial top-ranked important features in CN vs. EMCI

Frequency	Shortest path
16	SPG.L-TPOmid.L, ACG.L-STG.L
15	AMYG.R-PCL.R, PreCG.L-INS.R, PCG.L-ITG.L, OLF.L-PHG.R
14	INS.R-IFGoperc.L, IFGtriang.R-SMG.R, ITG.R-PCUN.R, FFG.L-PUT.L, ORBsupmed.L-PAL.R, IFGoperc.R-PCUN.L, FFG.L-PAL.R, SFGdor.L-THA.R, ANG.R-STG.L, SFGmed.R-ORBsupmed.L, ITG.R-MTG.R, CAU.R-TPOsup.R

efficient and stable.

3.1.2 Feature selection

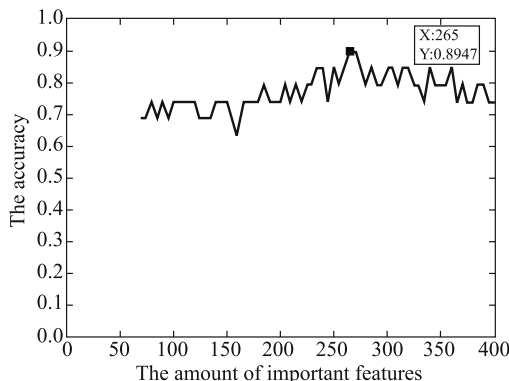
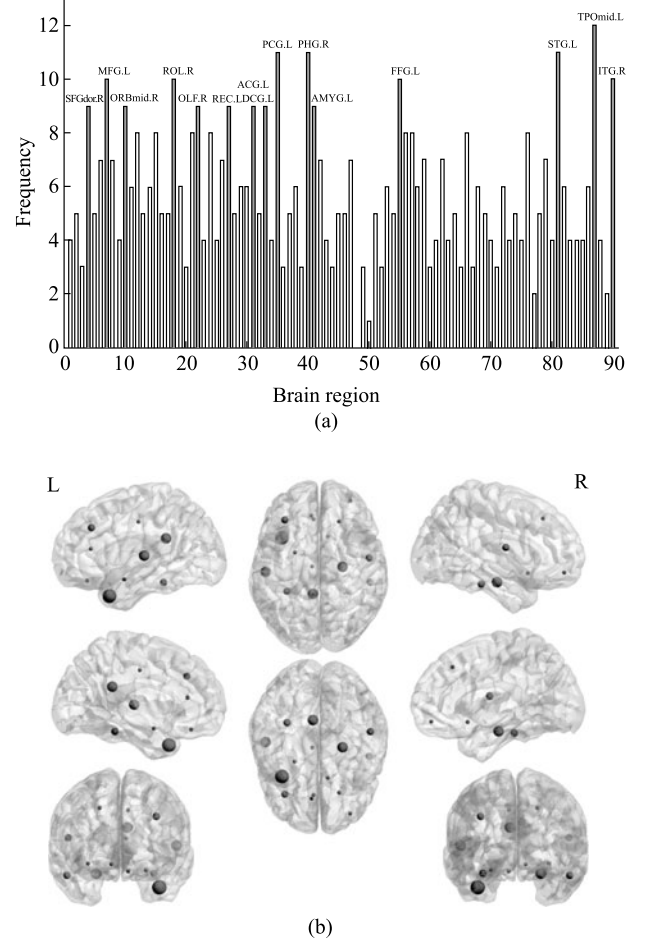
To find out the important features, we first extracted the 400 SVM base classifiers from the optimal ensemble learner (found in Fig. 3). Then the features selected by these SVMs were retained. The frequencies of the retained features were used to evaluate their discriminative power. The top 400 high-frequency features were extracted as the important ones. Table 3 shows the important features with frequencies of 16, 15 and 14. The graph features with the highest frequency of 16 were all the shortest paths between temporal pole: middle temporal gyrus (TPOmid.L) and superior parietal gyrus (SPG.L), anterior cingulate and paracingulate gyri (ACG.L) and STG.L.

In the feature selection, we randomly selected the top q features from 400 important ones, where q was within the range of $\{70, 75, \dots, 400\}$. Then the 65 features which were randomly sampled from the top q features were considered as input features of the SVM ensemble. As we could observe from Fig. 5, the CEWSVME algorithm reported the highest accuracy peaking at 89.47% when q was equals to 265. Therefore, the top 265 features were regarded as the optimal feature subset and employed to investigate corresponding differential regions in the “CN-EMCI” progression.

3.1.3 Differential brain regions detection

The locations of the differential brain regions with higher frequencies are described in Fig. 6(b). The vertices denoted the brain regions divided by the AAL template. The larger the vertex was, the higher the frequency of its corresponding brain region was and the more relevant this region was to the conversion from CN to EMCI.

Moreover, the differential brain regions which had higher frequencies in the optimal feature subset made great contribu-

**Fig. 5** Search for the amount of optimal features in CN vs. EMCI based on the CEWSVME algorithm**Fig. 6** The frequencies of brain regions obtained by utilizing the CEWSVME framework in CN vs. EMCI analysis. (a) The differential brain regions with their frequencies; (b) distribution for differential brain regions

tions to CN vs. EMCI. Figure 6(a) exhibits the frequencies of 90 brain regions. The specific regions which had higher frequencies were shown by the red bar. For example, the differential regions with frequencies of 12 and 11 were respectively TPOmid.L, STG.L,

posterior cingulate gyrus (PCG.L) and parahippocampal gyrus (PHG.R).

3.2 Discovering differential brain regions between EMCI and LMCI

To explore the pathological progression from EMCI to LMCI, the experiment 2 was designed. The diagnostic accuracy rates of ensembles with various SVM amounts are described in Fig. 7. The CEWSVME reported the highest accuracy rate of 90% when the SVM amount equals to 300 (After performing 20 clusterings). Therefore, we set 300 to be the optimal SVM amount for next analysis. Similar to experiment 1, we also performed the accuracy comparison of RF, RSVMC and CEWSVME. The enormous potential of the CEWSVME algorithm was further demonstrated in the comparison, the average accuracy, average sensitivity and average specificity of which reached 84.30%, 90.00% and 80.00% respectively (see the Table 4 and Fig. 8).

Using the same method as experiment 1, we calculated the frequencies of the features selected by SVMs in the optimal

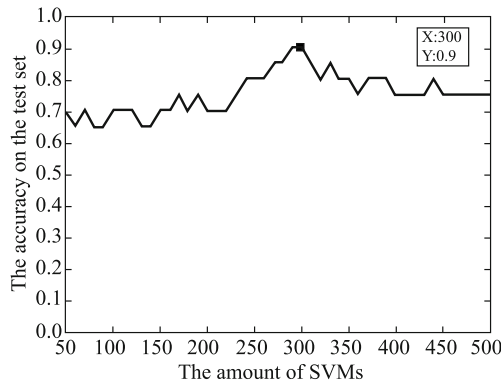


Fig. 7 Exploration of the optimal SVM amount of the CEWSVME in EMCI vs. LMCI

Table 4 Classification performance comparison of three models

Model	Accuracy/%	Sensitivity/%	Specificity/%
RF	74.10±8.31	70.00±7.15	80.00±6.93
RSVMC	77.21±7.96	72.73±7.69	77.78±10.35
CEWSVME	84.30±8.51	90.00±6.41	80.00±9.91

The values are shown by mean ± standard deviation.

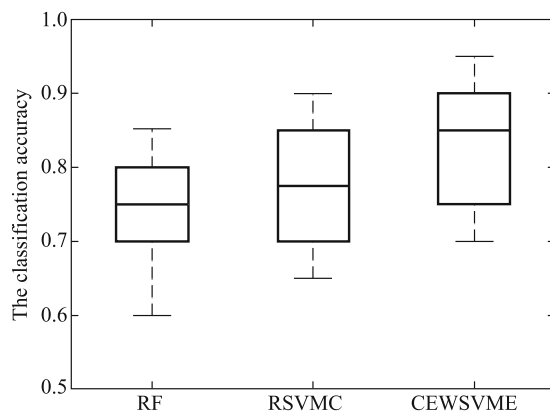


Fig. 8 The accuracy boxplots of random forest, random SVM cluster and CEWSVME in EMCI vs. LMCI

ensemble learner. The top 400 high-frequency features were extracted as important ones. The important features with frequencies of 13 and 12 are presented in Table 5. Then, the first q features of important features were used as inputs for the SVM ensemble to search out the optimal features in EMCI vs. LMCI, and the search process was showed in Fig. 9. The SVM ensemble reported the highest accuracy peaking at 90% when q equals to 280. Therefore, the top 280 ones of 400 important features are regarded as an optimal set of features, and employed for detection of the differential brain regions during the “EMCI-LMCI” progression.

Figure 10 (a) shows the differential brain regions with higher frequencies in EMCI vs. LMCI. For example, the brain regions with high frequencies of 13 and 11 were respectively parahip-

Table 5 The partial top-ranked important features in EMCI vs. LMCI

Frequency	Shortest path
13	HIP.L-DCG.L, TPOsup.R-DCG.R, PoCG.R-TPOsup.L
12	ROL.L-TPOmid.R, PHG.L-AMYG.L, REC.R-INS.L, MFG.L-PCG.L, ORBsup.L-PCUN.R, ORBmid.L-OLF.L, PCL.R-MTG.L, ITG.L-LING.R

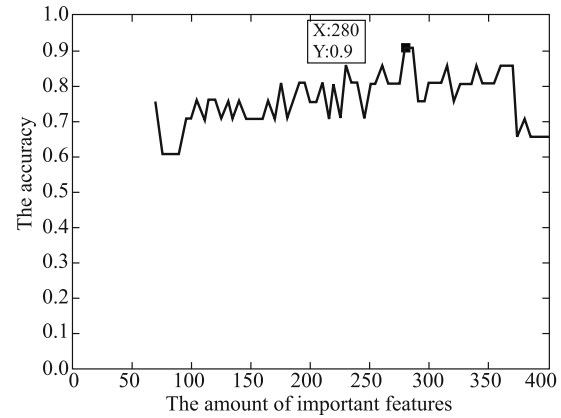


Fig. 9 Search for the amount of optimal features in EMCI vs. LMCI and EMCI vs. LMCI based on the CEWSVME algorithm

pocampal gyrus (PHG.L), amygdala (AMYG.L), postcentral gyrus (PoCG.L), middle temporal gyrus (MTG.R) and TPO-mid.L. The locations of these differential regions are described in Fig. 10(b).

3.3 Solitary-occurring and co-occurring differential brain regions

To clearly indicate the changes of abnormal brain regions during the MCI progression comprehensively, we analyzed brain

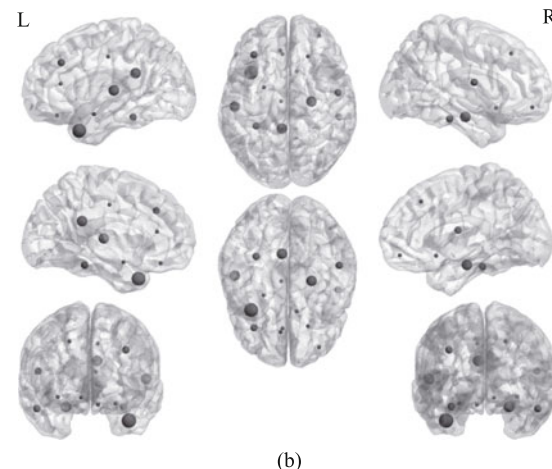
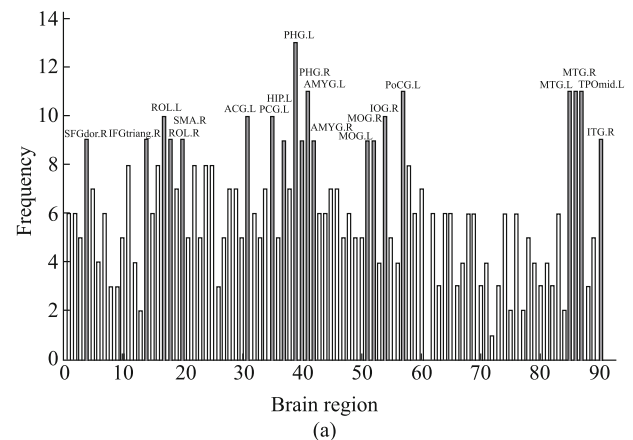


Fig. 10 The frequencies of brain regions obtained by utilizing the CEWSVME algorithm in EMCI vs. LMCI analysis. (a) The affected brain regions with their frequencies; (b) distribution for affected brain regions

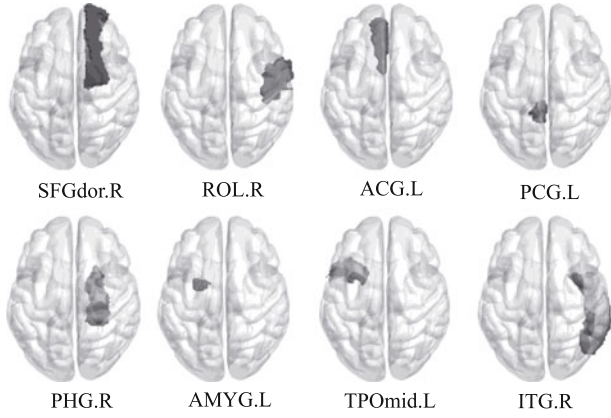


Fig. 11 The sagittal, coronal and axial positions of brain regions which have higher frequencies in two groups of experiments simultaneously

regions with frequencies up to 13, 12, 11, 10 and 9 in CN vs. EMCI and EMCI vs. LMCI respectively (see Fig. 6 and Fig. 10). Interestingly, some brain regions appeared simultaneously in two groups of experiments, which indicated that these brain regions were closely associated with the MCI progression (see Fig. 11). These differential regions were respectively TPOmid.L, PCG.L, inferior temporal gyrus (ITG.R), PHG.R, rolandic operculum (ROL.R), AMYG.L, ACG.L and superior frontal gyrus, dorsolateral (SFGdor.R). We will analyze these brain regions in the section of discussion.

4 Discussion

4.1 Classification effect

As the rapid development of neuroimaging technology, the classification and feature selection research on brain disorder have attracted the interest of many scientists [20,21]. The machine learning technology is very effective in disease identification, classification and feature extraction [22–25], and may be more intuitional and transparent compared with conventional statistical approaches [26–28]. In these studies on MCI progression, the SVM is the most widely-used [29] in machine learning methods, however, the accuracy seems to be unsatisfactory on account of the sensitivity to the selection of parameters and the effective discrimination between these stages was still difficult [30,31].

To circumvent this issue, we introduced the ensemble learning and the clustering-evolutionary idea to propose the CEWSVME framework for automatically investigating the alteration and transition from CN to EMCI to LMCI. The experiment results showed that the average classification accuracies achieved 83.47% and 84.30% respectively based on the optimal features in two sets of experiments for CN vs. EMCI and EMCI vs. LMCI, highlighting the fine performance. The core of the proposed framework was that an idea of clustering evolution was adopted into multiple SVMs to eliminate inefficient and highly similar SVMs. In the process of ensemble, the method performed many repeated experiments to acquire the optimal values of the cost parameter and bandwidth which were the Inf and 3, helping to optimize the performance of the base classifier. Furthermore, the optimal number of SVMs was found out to build the CEWSVME, which maximized the diversities of SVMs and ensured the generalization of ensemble

learners. On the other hand, a reasonable setting of the network threshold was of great importance, which could improve the final classification performance. The larger threshold value often preserved fewer and sparser connections. We applied multiple thresholds to revealing different levels of topological properties of brain connectivity network and found that 0.25 was still optimal. Consequently, the optimal features can be detected out in the case of maximizing the classification performance.

4.2 Analysis of solitary-occurring and co-occurring differential brain regions

In this paper, we extracted optimal features that greatly contributed to classifying patients and normal people, and further screened the differential brain regions for two brain diseases including EMCI and LMCI. The frequencies of differential brain regions were relatively higher than those of others, suggesting that they were related to EMCI (LMCI) pathology. Comparing two groups of experiments, we found that there were some co-occurring differential brain regions which had both relatively high frequencies in CN vs. EMCI and EMCI vs. LMCI, exhibiting that there were large changes in these brain regions at various stages of MCI. Furthermore, the solitary-occurring differential brain regions were observed. For example, some brain regions only showed the higher frequencies in the experiment with CN and EMCI, and the lower frequencies of occurrence in the experiment with EMCI and LMCI. It is indicated that these brain regions had little influence on the transformation from EMCI to LMCI and were only biomarkers of EMCI detected from normal people. On the contrary, some brain regions only had high frequencies in the EMCI and LMCI experiments and the changes of those were the most obvious, which were likely to be a sign of deterioration for LMCI.

In the co-occurring differential brain regions, multiple brain regions could be observed in two groups, which were highly implicated in the progression of MCI [32,33]. For instance, the parahippocampal gyrus showed a comparatively greater frequency compared to the others in two sets of experiments and was considered to be closely related to the MCI progression [34]. It is thought that the parahippocampal gyrus was involved in memory creation and recall of visual scenes [35,36]. Existing imaging researches on MCI have supported that it exhibited significantly consistent activation differences in the parahippocampal gyrus during episodic memory encoding [37]. The posterior cingulate gyrus also exhibited a comparatively greater frequency, which emphasized its importance in the MCI progression. Previous neuroscience researches on MCI have reported the posterior cingulate gyrus abnormalities [38–40], which were in line with our findings.

The solitary-occurring differential brain regions were found out in CN vs. EMCI (EMCI vs. LMCI) experiment. Since these brain regions individually appeared in the differential brain region analysis of one experiment, they only differentiate one stage rather than the entire progression of MCI, which might explain the differences between EMCI and LMCI. For example, the STG.L was one of solitary-occurring differential brain regions and only appears in CN vs. EMCI, indicating that the STG.L played a fundamental role in progression from CN to EMCI. Previous researches on EMCI have found out the STG.L

abnormalities [41,42]. Oppositely, the MTG.R showed a comparatively higher frequency in EMCI vs. LMCI, indicating that the MTG.R contributed greatly to the “EMCI-LMCI” stage. Several current researches on EMCI vs. LMCI have reported the MTG.R abnormalities [43–45]. Our study not only explored the brain regions that were in accordance with the development of the MCI progression, but also detected the brain regions that had independent effects on EMCI or LMCI, demonstrating the complexity of MCI progression.

5 Conclusion

As an effective and reliable imaging technique, the fMRI was utilized to investigate the intrinsic brain functional activities at the whole-brain level. The graph theory provided a novel way to explore the disturbed topological properties involved in neurodegenerative disorders. Based on the fMRI and graph-theoretic features, we proposed an advanced CEWSVME algorithm to differentiate the two stages in the MCI progression and explore early pathological changes. The solitary-occurring and co-occurring differential brain regions were detected in terms of their contributions to the algorithm, which were respectively implemented in two groups of CN vs. EMCI and EMCI vs. LMCI. Our work successfully opened new avenues for understanding brain pathological changes in the progression of MCI. Additionally, we will try our best to expand the data scale through various ways, and consider integrating more fMRI data from other databases into our research in the follow-up work.

Acknowledgements This work was supported by the Hunan Provincial Science and Technology Project Foundation (2018TP1018), the National Science Foundation of China (61502167).

References

- Sherman D S, Mauser J, Nuno M, Sherzai D. The efficacy of cognitive intervention in mild cognitive impairment (MCI): a meta-analysis of outcomes on neuropsychological measures. *Neuropsychology Review*, 2017, 27(4): 440–484
- Li J Q, Tan L, Wang H F, Tan M S, Tan L, Xu W, Zhao Q F, Wang J, Jiang T, Yu J T. Risk factors for predicting progression from mild cognitive impairment to alzheimer's disease: a systematic review and meta-analysis of cohort studies. *Journal of Neurology, Neurosurgery & Psychiatry*, 2016, 87(5): 476–484
- Yi H A, Möller C, Dieleman N, Bouwman F H, Barkhof F, Scheltens P, van der Flier W M, Vrenken H. Relation between subcortical grey matter atrophy and conversion from mild cognitive impairment to alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 2016, 87(4): 425–432
- Ramírez J, Górriz J M, Ortiz A, Martínez-Murcia F J, Segovia F, Salas-Gonzalez D, Castillo-Barnes D, Illán I A, Puntónet C G. Ensemble of random forests one vs. rest classifiers for MCI and ad prediction using anova cortical and subcortical feature selection and partial least squares. *Journal of Neuroscience Methods*, 2018, 302: 47–57
- ten Brinke L F, Bolandzadeh N, Nagamatsu L S, Hsu C L, Davis J C, Miran-Khan K, Liu-Ambrose T. Aerobic exercise increases hippocampal volume in older women with probable mild cognitive impairment: a 6-month randomised controlled trial. *British Journal of Sports Medicine*, 2015, 49(4): 248–254
- Spulber G, Simmons A, Muehlboeck J S, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Soininen H, Spenger C, Lovestone S, Wahlund L O, Westman E, et al. An MRI-based index to measure the severity of alzheimer's disease-like structural pattern in subjects with mild cognitive impairment. *Journal of Internal Medicine*, 2013, 273(4): 396–409
- Mecca A P, Michalak H R, McDonald J W, Kemp E C, Pugh E A, Becker M L, Mecca M C, van Dyck C H. Sleep disturbance and the risk of cognitive decline or clinical conversion in the adni cohort. *Dementia and Geriatric Cognitive Disorders*, 2018, 45(3–4): 232–242
- Jagust W J, Landau S M, Koeppe R A, Reiman E M, Chen K, Mathis C A, Price J C, Foster N L, Wang A Y. The alzheimer's disease neuroimaging initiative 2 pet core: 2015. *Alzheimer's & Dementia*, 2015, 11(7): 757–771
- Lee E S, Yoo K, Lee Y B, Chung J, Lim J E, Yoon B, Jeong Y. Default mode network functional connectivity in early and late mild cognitive impairment. *Alzheimer Disease & Associated Disorders*, 2016, 30(4): 289–296
- Cai S, Chong T, Peng Y, Shen W, Li J, von Deneen K M, Huang L. Altered functional brain networks in amnesic mild cognitive impairment: a resting-state fMRI study. *Brain Imaging and Behavior*, 2017, 11(3): 619–631
- Fei F, Jie B, Zhang D. Frequent and discriminative subnetwork mining for mild cognitive impairment classification. *Brain Connectivity*, 2014, 4(5): 347–360
- Bi X A, Xu Q, Luo X, Sun Q, Wang Z. Weighted random support vector machine clusters analysis of resting-state fMRI in mild cognitive impairment. *Frontiers in Psychiatry*, 2018, 9: 340
- McKenna F, Koo B B, Killiany R, et al. Comparison of apoe-related brain connectivity differences in early MCI and normal aging populations: an fMRI study. *Brain Imaging and Behavior*, 2016, 10(4): 970–983
- Wee C Y, Yang S, Yap P T, Shen D, et al. Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging and Behavior*, 2016, 10(2): 342–356
- Jie B, Liu M, Shen D. Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease. *Medical Image Analysis*, 2018, 47: 81–94
- Grajski K A, Bressler S L. Differential medial temporal lobe and default-mode network functional connectivity and morphometric changes in alzheimer's disease. *NeuroImage: Clinical*, 2019, 23: 101860
- Daianu M, Jahanshad N, Nir T M, Jack Jr C R, Weiner M W, Bernstein M A, Thompson P M, et al. Rich club analysis in the alzheimer's disease connectome reveals a relatively undisturbed structural core network. *Human Brain Mapping*, 2015, 36(8): 3087–3103
- Jie B, Liu M, Zhang D, Shen D. Sub-network kernels for measuring similarity of brain connectivity networks in disease diagnosis. *IEEE Transactions on Image Processing*, 2018, 27(5): 2340–2353
- Ding X, Charnigo R J, Schmitt F A, Kryscio R J, Abner E L, et al. Evaluating trajectories of episodic memory in normal cognition and mild cognitive impairment: results from adni. *PLoS ONE*, 2019, 14(2): e0212435
- Schetinin V, Jakaite L, Nyah N, Novakovic D, Krzanowski W. Feature extraction with gmdh-type neural networks for eeg-based person identification. *International Journal of Neural Systems*, 2017, 28(6): 1750064
- Du L, Liu K, Zhu L, Yao X, Risacher S L, Guo L, Saykin A J, Shen L, et al. Identifying progressive imaging genetic patterns via multi-task sparse canonical correlation analysis: a longitudinal study of the adni cohort. *Bioinformatics (Oxford, England)*, 2019, 35(14): i474–i483
- Yan K, Xu Y, Fang X, Zheng C, Liu B. Protein fold recognition based on sparse representation based classification. *Artificial Intelligence in Medicine*, 2017, 79: 1–8
- Wu D, Zheng S J, Zhang X P, Yuan C A, Cheng F, Zhao Y, Lin Y J, Zhao Z Q, Jiang Y L, Huang D S. Deep learning-based methods for person re-identification: a comprehensive review. *Neurocomputing*, 2019, 337: 354–371
- Jin Q, Meng Z, Pham T D, Chen Q, Wei L, Su R. Dunet: a deformable network for retinal vessel segmentation. *Knowledge-Based Systems*, 2019, 178: 149–162
- Su R, Liu X, Wei L, Zou Q. Deep-resp-forest: a deep forest model to

- predict anti-cancer drug response. *Methods*, 2019, 166: 91–102
26. Zeng X, Yuan S, Huang X, Zou Q. Identification of cytokine via an improved genetic algorithm. *Frontiers of Computer Science*, 2015, 9(4): 643–651
 27. Chen X, Zhu C C, Yin J. Ensemble of decision tree reveals potential mirna-disease associations. *PLoS Computational Biology*, 2019, 15(7): e1007209
 28. Peng J, Hui W, Li Q, Chen B, Hao J, Jiang Q, Shang X, Wei Z. A learning-based framework for mirna-disease association identification using neural networks. *Bioinformatics (Oxford, England)*, 2019, 35(21): 4364–4371
 29. Cui H, Zhang X. Alignment-free supervised classification of metagenomes by recursive SVM. *BMC Genomics*, 2013, 14: 641
 30. Prasad G, Joshi S H, Nir T M, Toga A W, Thompson P M. Brain connectivity and novel network measures for alzheimer's disease classification. *Neurobiology of Aging*, 2015, 36: S121–S131
 31. Khazaei A, Ebrahimzadeh A, Babajani-Feremi A. Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and alzheimer's disease. *Brain Imaging and Behavior*, 2016, 10(3): 799–817
 32. Echávarri C, Aalten P, Uylings H B M, Jacobs H I L, Visser P J, Gronenschild E H B M, Verhey F R J, Burgmans S. Atrophy in the parahippocampal gyrus as an early biomarker of alzheimer's disease. *Brain Structure and Function*, 2011, 215(3): 265–271
 33. Chao L L, Mueller S G, Buckley S T, Peek K, Raptentsetseng S, Elman J, Yaffe K, Miller B L, Kramer J H, Madison C, Mungas D, Schuff N, Weiner M W. Evidence of neurodegeneration in brains of older adults who do not yet fulfill MCI criteria. *Neurobiology of Aging*, 2010, 31(3): 368–377
 34. Kim S M, Kim M J, Rhee H Y, Ryu C W, Kim E J, Petersen E T, Jahng G H. Regional cerebral perfusion in patients with alzheimer's disease and mild cognitive impairment: effect of apoe epsilon4 allele. *Neuroradiology*, 2013, 55(1): 25–34
 35. Ward A M, Schultz A P, Huijbers W, Van Dijk K R A, Hedden T, Sperling R A. The parahippocampal gyrus links the default-mode cortical network with the medial temporal lobe memory system. *Human Brain Mapping*, 2014, 35(3): 1061–1073
 36. Luck D, Danion J M, Marrer C, Pham B T, Gounot D, Foucher J. The right parahippocampal gyrus contributes to the formation and maintenance of bound information in working memory. *Brain and Cognition*, 2010, 72(2): 255–263
 37. Browndyke J N, Giovanello K, Petrella J, Hayden K, Chiba-Falek O, Tucker K A, Burke J R, Welsh-Bohmer K A. Phenotypic regional functional imaging patterns during memory encoding in mild cognitive impairment and alzheimer's disease. *Alzheimer's & Dementia*, 2013, 9(3): 284–294
 38. Kantarci K, Jack C R, Xu Y C, Campeau N G, O'Brien P C, Smith G E, Ivnik R J, Boeve B F, Kokmen E, Tangalos E G, Petersen R C. Regional metabolic patterns in mild cognitive impairment and alzheimer's disease: a 1h mrs study. *Neurology*, 2000, 55(2): 210–217
 39. Camus V, Payoux P, Barré L, Desgranges B, Voisin T, Tauber C, La Joie R, Tafani M, Hommet C, Chételat G, Mondon K, de La Sayette V, Cottier J P, Beauvils E, Ribeiro M J, Gissot V, Vierron E, Vercouillie J, Vellas B, Eustache F, Guilloteau D. Using pet with 18f-av-45 (florbetapir) to quantify brain amyloid load in a clinical environment. *European Journal of Nuclear Medicine and Molecular Imaging*, 2012, 39(4): 621–631
 40. Bailly M, Destrieux C, Hommet C, Mondon K, Cottier J P, Beauvils E, Vierron E, Vercouillie J, Ibazizene M, Voisin T, Payoux P, Barré L, Camus V, Guilloteau D, Ribeiro M J. Precuneus and cingulate cortex atrophy and hypometabolism in patients with alzheimer's disease and mild cognitive impairment: MRI and 18f-fdg pet quantitative analysis using freesurfer. *BioMed Research International*, 2015, 2015: 583931
 41. Cai S, Huang L, Zou J, Jing L, Zhai B, Ji G, von Deneen K M, Ren J, Ren A, et al. Changes in thalamic connectivity in the early and late stages of amnesic mild cognitive impairment: a resting-state functional magnetic resonance study from ADNI. *PLoS ONE*, 2015, 10(2): e0115573
 42. Li H, Fang S, Contreras J A, West J D, Risacher S L, Wang Y, Sporns O, Saykin A J, Goñi J, Shen L, et al. Brain explorer for connectomic analysis. *Brain Informatics*, 2017, 4(4): 253–269
 43. Xiang J, Guo H, Cao R, Liang H, Chen J. An abnormal resting-state functional brain network indicates progression towards alzheimer's disease. *Neural Regeneration Research*, 2013, 8(30): 2789–2799
 44. Wei H, Kong M, Zhang C, Guan L, Ba M, et al. The structural MRI markers and cognitive decline in prodromal alzheimer's disease: a 2-year longitudinal study. *Quantitative Imaging in Medicine and Surgery*, 2018, 8(10): 1004–1019
 45. Ribeiro A S, Lacerda L M, Silva N A d, Ferreira H A. Multimodal imaging of brain connectivity using the mibca toolbox: preliminary application to alzheimer's disease. *IEEE Transactions on Nuclear Science*, 2015, 62(3): 604–611



Xia-an Bi is currently an associate professor in the College of Information Science and Engineering in Hunan Normal University, China. He received the PhD degree in computer science and technology from the College of Information Science and Engineering in Hunan University, China in 2012. His current research interests include machine learning, brain science and artificial intelligence.



Yiming Xie, received the BE degree in Network Engineering from Chengdu Technological University, China in 2019. He is currently pursuing the MS degree in College of Information Science and Engineering, Hunan Normal University, China. The major of him is computer science and technology. His main research fields include data mining, brain science and artificial intelligence.



Hao Wu, received the BE degree in Computer Science and Technology from Jinggangshan University, China in 2019. He is currently pursuing the MS degree in College of Information Science and Engineering, Hunan Normal University, China. The major of him is computer technology. His main research fields include data mining, brain science and artificial intelligence.



Luyun Xu, received the PhD degree in business administration from Hunan University, China in 2018. She is currently an assistant professor in Business School in Hunan Normal University, China. Her research interests focus on knowledge management, data mining and machine learning. She has published in the *Journal of Technology Transfer*, *Technological Analysis & Strategic Management*, *Computational and Mathematical Organization Theory*.