Brain Imaging and Behavior
https://doi.org/10.1007/s11682-019-00255-9

ORIGINAL RESEARCH

Diagnosis of early Alzheimer's disease based on dynamic high order networks

Baiying Lei[1] · Shuangzhi Yu[1] · Xin Zhao[1] · Alejandro F Frangi[3] · Ee-Leng Tan[4] · Ahmed Elazab[1] · Tianfu Wang[1] · Shuqiang Wang[2]

Abstract

Machine learning methods have been widely used for early diagnosis of Alzheimer's disease (AD) via functional connectivity networks (FCNs) analysis from neuroimaging data. The conventional low-order FCNs are obtained by time-series correlation of the whole brain based on resting-state functional magnetic resonance imaging (R-fMRI). However, FCNs overlook inter-region interactions, which limits application to brain disease diagnosis. To overcome this drawback, we develop a novel framework to exploit the high-level dynamic interactions among brain regions for early AD diagnosis. Specifically, a sliding window approach is employed to generate some R-fMRI sub-series. The correlations among these sub-series are then used to construct a series of dynamic FCNs. High-order FCNs based on the topographical similarity between each pair of the dynamic FCNs are then constructed. Afterward, a local weight clustering method is used to extract effective features of the network, and the least absolute shrinkage and selection operation method is chosen for feature selection. A support vector machine is employed for classification, and the dynamic high-order network approach is evaluated on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Our experimental results demonstrate that the proposed approach not only achieves promising results for AD classification, but also successfully recognizes disease-related biomarkers.

Keywords Alzheimer's disease · Resting-state functional magnetic resonance imaging · Functional connectivity networks · Dynamic high-order network

Introduction

Among the neurodegenerative diseases of the brain, Alzheimer's disease (AD) is one of the most common dementia, and it accounts for 60–80% of all dementia patients. AD is a chronic and progressive disease with a variety of higher-order cortical dysfunctions with cognitive function decline, reduced judgment and memory ability, and eventually leads to death. According to the World Alzheimer Report 2018, there are about 5.7 million people with AD in the United States. This number accounts for about 2% of the US population. It is estimated that the number of patients will reach 13.8 million (Alzheimer 2018) by 2050. China has the largest number of AD patients in the world, with over 9.5 million patients. Studies show that the global cost of dementia in 2015 exceeds 975.56 billion US dollars, which is much higher than the estimated value of 800 billion US dollars reported by the World Alzheimer Report in 2015. This cost will reach 2.54 trillion

✉ Tianfu Wang
tfwang@szu.edu.cn

✉ Shuqiang Wang
sq.wang@siat.ac.cn

1 National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Shenzhen University, Shenzhen 518060, China

2 Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong Province 518000, People's Republic of China

3 CISTIB Centre for Computational Imaging & Simulation technologies in Biomedicine, School of Computing and the School of Medicine, University of Leeds, Leeds, UK

4 School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

Published online: 12 August 2020

US dollars by 2030 (Jia et al. 2018). Most of AD patients rely on others' assistance to maintain a normal life, which imposes a huge burden to caregivers, including social life, psychological aspects, physical activities and economic pressure. AD has become a global problem, and it is of high urgency to find the pathogenesis of AD for early diagnosis and treatment. Mild cognitive impairment (MCI) is the early stage of AD. The annual conversion rate of MCI to AD is as high as 10–15%, while the annual conversion rate of healthy people to AD is only 1–2%. It is possible to delay AD progression in MCI patients after certain cognitive training and rehabilitation treatment, and some patients can even return to the normal state. However, once entering the AD stage, there are no known effective therapeutic drugs developed at present, and this process is irreversible. It was found that AD could start 20 years of preclinical diagnosis (Visser and Tijms 2017), which can be further classified as significant memory concern (SMC), early MCI (EMCI), and late MCI (LMCI). By early diagnosis and intervention treatment, it is possible to delay the onset of AD, reduce the suffering of patients, and reduce the burden on the family and society.

To diagnose AD (Zhou et al. 2019b), the neuroimaging technique is an effective tool. Magnetic resonance imaging (MRI), safe and non-invasive real-time imaging of the brain, has opened up new ways for early diagnosis of AD (Shi et al. 2017)by locating changes of imaging markers in the brain. As one of the MRI technologies, the functional MRI (fMRI) (Huettel et al. 2004) has been widely used in the research of brain diseases. The brain activities are associated with contraction and relaxation of blood vessels in the brain, which changes blood flow velocity and oxygen levels. The fMRI measures these hemodynamic changes, and real-time in vivo imaging of brain functions is performed using blood oxygenation level dependence (BOLD) (Ogawa et al. 1990). The fMRI uses a non-radioactive imaging modality with high spatial and temporal resolution and can provide effective functional information for the early diagnosis of AD.

According to the latest study, there are about 86 billion neurons in the normal adult brain (Herculano-Houzel 2012), and there are various neurons in the brain. Many neurons connect with each other through synapses, which forms a complex, multi-level and multi-functional brain network. Because of this characteristic structure, the brain forms a huge center of information transmission, exchange, storage, and calculation. If a neuron is treated as a node and the connections between neurons is treated as edges, the brain is essentially a complex neural network. At the micro scale, this is a huge complex network and can be solved by connectomics. To facilitate brain functionality analysis, brain atlases can greatly reduce the complexity of brain network connectivity at the macro scale and for the research and interpretation of brain diseases. Network analysis can reliably quantify brain networks with a few neurobiological measures easy to compute. Meanwhile, by clearly defining anatomical connectivity networks and functional connectivity networks (FCNs) on the same atlas of brain regions, network analysis can explore the relationship between structure and FCNs. By comparing the structure or functional network topology among subjects, we can further reveal the abnormal connections between neurological and mental diseases. Therefore, network construction and analysis methods are widely used in early diagnosis and prediction of brain diseases (Qi et al. 2015; Wang et al. 2010; Zhu et al. 2018b).

For brain disease analysis, Biswal et al. found that the time-series of low-frequency fluctuations in the resting brain have high temporal correlation between some brain regions, which can be a manifestation of brain functional connectivity (Biswal et al. 1995). Friston defined functional connectivity as the temporal correlation of the activation neurons in dissected brain regions to reflect the level of functional communication between brain regions (Friston 2011). Researchers are using FCNs method not only to discover the high-order between left and right hemisphere motor cortex in the function of the connection, but also to find new functional brain networks, such as visual network, auditory network, language, and cognitive network (van den Heuvel and Hulshoff Pol 2010). These studies have shown there are a large number of spontaneous activities even when the FCNs are not idle. In addition, these activities are highly correlated among different brain regions. Later, many researchers applied the FCNs method to the functional disconnection analysis of neurological and mental diseases, such as AD (Greicius et al. 2004), Parkinson's disease (Niethammer et al. 2012), schizophrenia (Jafri et al. 2008), Autism (Wang et al. 2017; Weng et al. 2010), and depression (Lui et al. 2011). Zhou et al. (2019a) exploited the underlying association among different modalities by projecting to a latent feature space, which achieves good performance in the classification of MCI. The above studies provided sufficient evidence to prove the reliability of the functional connectivity analysis methods.

Graph theory has been generally used in brain network analysis (Bullmore and Sporns 2009; Rubinov and Sporns 2010; Zhu et al. 2018a). The degree of a node, as the most basic network measurement parameter, is often defined as the number of other nodes connected to this node(Zhu et al. 2018b). Distribution of nodes in the network forms the degree distribution concept. It is found that the brain is a power-law degree distribution network form, and random network medium has an equal probability distribution, which illustrates the brain network has small world properties. Compared with the random network with higher overall efficiency, the brain network has higher local efficiency (Bassett and Bullmore 2006). The clustering coefficient is the ratio of the number of connected edges between the nearest neighbor nodes to the number of all possible connected edges, which is used to describe the efficiency of the network. The above network evaluation

are often compared between patients with encephalopathy and normal subjects for inter-group statistical analysis to find abnormal connections in brain areas(Shi et al. 2018). However, this method cannot analyze and diagnose a single subject. Different from the traditional graph theory methods, machine learning provides a new way of thinking through multivariate mode analysis, which can analyze individual subjects. Therefore, many researchers have carried out their research on brain network analysis using machine learning methods (Zhang et al. 2018; Zheng et al. 2018).

Machine learning analysis of brain network is mainly divided into four parts: the construction of brain network, feature extraction, feature learning, and classifier learning. This paper focuses on the research of brain network learning. The traditional low-order functional brain network is based on the correlation of whole brain time series of R-fMRI, which ignores the dynamic changes of connections among brain regions and limits its application in the diagnosis of brain diseases. To overcome this limitation, many researchers have explored the dynamic network methods, including sliding window methods (Allen et al. 2014; Damaraju et al. 2014), wavelet transform coherence methods (Chang and Glover 2010; Yaesoubi et al. 2015), and dynamic conditional correlation methods (Lindquist et al. 2014). The dynamic network provides a new research direction to study the brain network. In addition, some hybrid networks have been proposed for early diagnosis of AD. For example, Zhang et al (Zhang et al. 2017)propose a novel approach called "hybrid high-order FC networks" to characterize previously untouched inter-level interaction between the low- and the high-order FC networks. This method combined high-order with low-order brain networks and improves the accuracy of classification in some extent. However, such methods have certain limitations since it ignore the dynamic changes of connections among brain regions. Motivated by the aforementioned works, in this study, we use the high-order dynamic interaction features among brain regions for early classification of AD. Compared to hybrid networks, ours model has both high-order and dynamic characteristics and provides more characteristic information (e.g. dynamic information). Specifically, we adopt a sliding window method to generate subsequences of fluctuation signals in brain regions and construct a dynamic functional brain network by calculating the correlation among these subsequences. Then, we further construct a dynamic high order network based on the topological similarity between the paired dynamic networks. The local weight clustering coefficient (Rubinov and Sporns 2010) is used to extract network features, and feature selection is conducted by the least absolute shrinkage and selection operation (LASSO) method (Tibshirani 1996). Finally, the linear kernel support vector machine (SVM) is selected as the classifier (Chang and Lin 2011), and the leave-one-out cross-validation method is applied for model validation. This paper experiments based on the Alzheimer's disease neuroimaging initiative (ADNI) database (Mueller et al. 2005). The experiments verify the effectiveness of the high-order dynamic network method by comparing four groups of network construction methods. Specifically, static low-order functional network (s-LON), static high-order functional network (s-HON), dynamic low-order functional network (d-LON), and dynamic high-order functional network (d-HON). Overall, the contributions of this paper are summarized:

1) To our best knowledge, this is the first attempt using dynamic high-order model for early diagnosis of AD.
2) Local weight clustering method is used for feature extraction, which focuses on the local network density. Compared with local clustering method, the local weight clustering coefficient considers the influence of the weights during the process of network learning, which can effectively learn network feature for model classification.
3) A feature selection LASSO method is used. This method applies statistical criteria like filtering model to select several candidate feature subsets with a cardinality and the subset with the highest classification accuracy is selected with the wrapping model.

Materials and methods

Subjects and data acquisition

In this study, all the data are from the ADNI database (adni.loni.usc.edu), and the detailed statistical information is shown in Table 1. A total of 170 subjects are included in the experiment, including 38 LMCI patients, 44 EMCI patients, 44 SMC patients and 44 normal control (NC). All patients are scanned by a 3 T MRI scanner across various sites using these parameters: imaging matrix =64 × 64, 48 slices, 180 volumes, and voxel thickness = 3.3 mm.

Data preprocessing

The R-fMRI data preprocessing procedure is performed using GRETNA toolbox (Wang et al. 2015). To keep magnetization equilibrium, we discard the first 10 acquired R-fMRI volumes of each subject. The remaining 170 volumes are corrected for

Table 1 Summary of the subjects' information used in this study

Group	LMCI (38)	EMCI (44)	SMC(44)	NC(44)
Male/Female	19 M/19F	22 M/22F	22 M/22F	22 M/22F
Age (mean±SD)	76.0±7.6	76.5±6.1	76.3±5.4	76.5±4.5

the staggered order of slice acquisition during echo-planar scanning by matching all time points to intermediate time points. Afterward, the head-motion artifacts in the R-fMRI time-series are removed by the first volume of each subject as a reference to register all these volumes. After registration, the volumes are normalized to the MNI atlas space and resampled to the voxel size of $3 \times 3 \times 3$ mm. After spatial normalization, we spatially smooth the dataset with a 4 mm full width half maximum Gaussian kernel followed by band-pass filtering of frequency interval ($0.01\ \text{Hz} \le f \le 0.08\ \text{Hz}$) to minimize the effects of low-frequency drift and high-frequency noise. Each regional mean time series is regressed against the average cerebrospinal fluid, white-matter signals and the six parameters from motion correction.

A parcellation of the brain space into 90 ROIs is then performed by warping the automated anatomical labelling (AAL) (Tzourio-Mazoyer et al. 2002) atlas to the R-fMRI images. Finally, we obtain the time series of each individual ROI of each subject via the time series over all the voxels in that particular ROI. As Fig.1 shows, before feature selection, the ROI feature dimension is 90. After feature selection, the feature dimension is not fixed, we select the ROI features according to the weights based on LASSO method.

Network construction and analysis

Construction of the static low-order FCNs

In this paper, we construct the functional connectivity using pairwise Pearson correlation coefficients (PCC) between the ROI pairs. We consider the ROIs as nodes and PCC as edges connecting them.We can calculate PCC using:

$$\mathbf{PCC}_{ij} = corr\{\boldsymbol{x}(\boldsymbol{i}), \boldsymbol{x}(\boldsymbol{j})\}, \tag{1}$$

where *corr* is the pairwise correlation coefficient between $x(i)$ and $x(j)$.

Construction of the dynamic low-order FCNs

To obtain the dynamic FCNs, a sliding window approach is used to segment the entire R-fMRI time series into multiple sub-series. Given that an R-fMRI time series with M volumes, $D = [(M - L)\ \ s] + 1$ is the number of sub-series that can be generated, where L denotes the length of a sliding window and s denotes the step size, N denotes the number of subjects, the d-th sub-series is represented in the form of a matrix $\mathbf{x}^{(d)} = \left\{x_1^{(d)}, x_2^{(d)}, \ldots, x_N^{(d)}\right\} \in \mathbb{R}^{\mathrm{d} \times N}, d = 1, 2, \ldots, D$, and $R = 90$ is the total number of ROIs. A symmetric connectivity matrix $\mathbf{c}^{(d)} = \left[c_{ij}^{(d)}\right] \in \mathbb{R}^{R \times R}$ can be constructed using $\mathbf{x}^{(d)}$, where each element in the connectivity matrix defines the correlation strength between different ROIs. Specifically, the correlation strength is defined as:

$$c_{ij}^{(d)} = corr\left\{x^{(d)}(i), x^{(d)}(j)\right\}, \tag{2}$$

where $corr\{\cdot\}$ computes the correlation between region i and region j of the d-th sub-series. Then, the entry for the temporal FCNs matrix $\mathbf{C}^{(d)} = \left\{\mathrm{c}_1^{(\mathrm{d})}, \mathrm{c}_2^{(\mathrm{d})}, \ldots, \mathrm{c}_{\mathrm{N}}^{(\mathrm{d})}\right\} \in \mathbb{R}^{1 \times N}$. The traditional FCNs is an extreme case where window length is maximized to the entire time scale ($L = M$). Figure 2 illustrates the generation of ROIs sub-series using the sliding window based on the PCC of R-fMRI data.

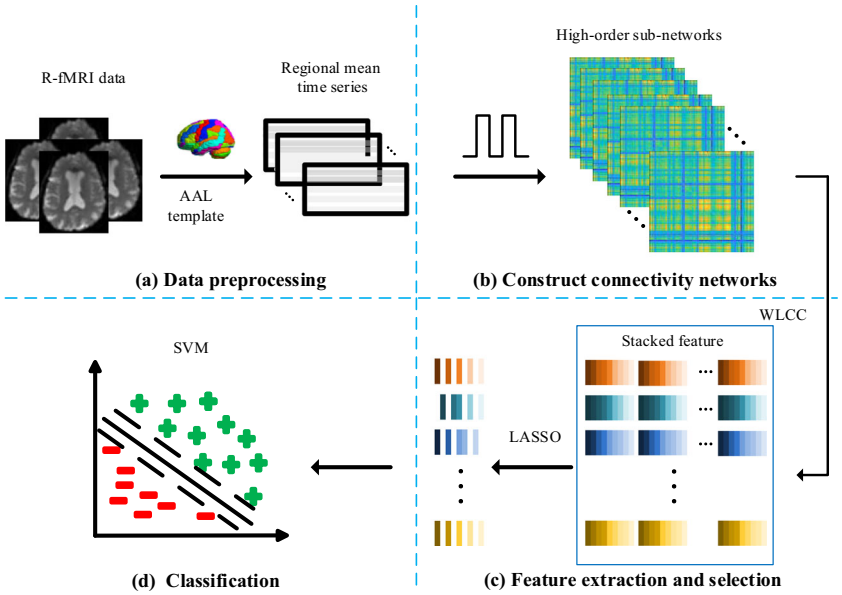

Fig. 1 The proposed framework of dynamic high order connected network for MCI identification

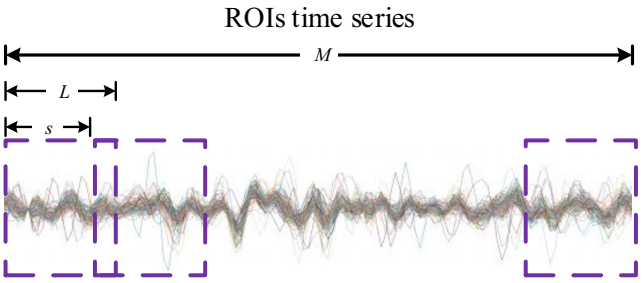

Fig. 2 Generation of the ROIs sub-series using the sliding window approach

Construction of the dynamic high-order FCNs

In the previous subsection, we construct FCNs using a sliding window technique. However, our main goal is to reveal high-level information and intrinsic relationships. Therefore, the FCNs is first built dynamically using the d-th R-fMRI time sub-series ($\boldsymbol{d=1,2,\ldots,D}$). We then construct high-order FCNs based on topographical similarities between each sub-networks $\boldsymbol{c^{(d)}}$ pair. The dynamic high-order FCNs can be equivalently written as:

$$\mathbf{h}_n{}^{(d)} = \left(\mathbf{c}_n{}^{(d)}\right)^T \mathbf{c}_n{}^{(d)}, \tag{3}$$

where $\mathbf{h}_n{}^{(d)}$ denotes dynamic high-order FCNs of the d-th sub-networks in the n-th subjects. Figure 3 illustrates the construction of the high-order FCNs.

Feature extraction and selection

After obtaining the dynamic high-order network, we utilize the method of local weight clustering coefficient to extract discriminant features of the functional brain network. This method quantifies the clustering of each node in the weighted network. Compared with the local clustering coefficient, it can represent the network more effectively, and the influence of the network weight is considered in the calculation process. We adopt local weight clustering coefficients to extract features from FCNs. The method quantifies the "cliqueness" of each node in a weighted network. A clique is originally a graph-theoretic concept, and refers to the the network's local topology for every node. This metric has been widely used for feature extraction in MCI diagnosis (Chen et al. 2016). Given

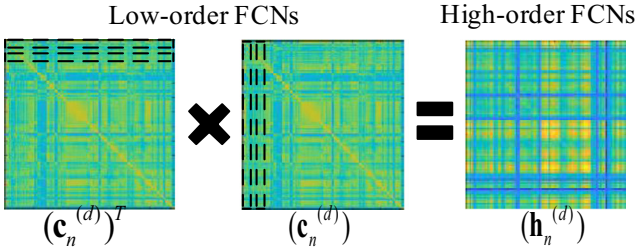

Fig. 3 Constructing high-order FCNs

a network with K vertices, the weight of the edge connecting vertex i and vertex j is denoted as $\boldsymbol{w_{ij}(1\le i,j\le K)}$, for each network $\boldsymbol{w_{ij}}$, the i-th node is

$$f_i = \frac{2\sum_{j:j\in\varepsilon_i}(w_{ij})^{\frac{1}{3}}}{|\varepsilon_i|(|\varepsilon_i|-1)}, \tag{4}$$

where $\boldsymbol{\varepsilon_i}$ denotes the set of vertices directly connected to vertex i and $|\boldsymbol{\varepsilon_i}|$ is the number of elements in $\boldsymbol{\varepsilon_i}$. After extracting the features from all the nodes, we concatenate them to form the feature vector. Then a supervised feature selection strategy based on LASSO is adopted to select the most discriminative features, while discarding redundant ones. LASSO can achieve the same precision as wrapping and the same efficiency as filtering. While eliminating redundant features, the most discriminative features are selected. Let $\mathbf{F}\in\mathbb{R}^{N\times K}$ be the sample feature matrix, where N is the number of samples and K is the number of features extracted by the method and $\mathbf{y}\in\mathbb{R}^N$ is the label vector of samples, the LASSO regression can be formulated using the following objective function:

$$\boldsymbol{min}_{(\mathbf{w},\mathbf{b})}\frac{\mathbf{1}}{\mathbf{2}}\|\mathbf{Fw}+\mathbf{b}-\mathbf{y}\|^2+\boldsymbol{\lambda}\|\mathbf{w}\|_1, \tag{5}$$

where $\boldsymbol{w_i}$ is the weight for the i-th feature, $\mathbf{b}$ is a bias term, $\boldsymbol{\lambda}$ is a balance parameter controlling the model sparsity based on the $\boldsymbol{l_1}$-norm regularization. The features corresponding to non-zero LASSO regression coefficients are retained as crucial features for classification. The SLEP toolbox (Liu et al. 2009) is used to implement feature selection.

Classification

After the features with discriminant power are selected by LASSO method, the selected features are used to train SVM for identification. The SVM classifier optimization function can be expressed as:

$$min_{\mathbf{w},\beta,a}\frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{p=1}^{n}a_p \tag{6}$$

$$s.t.\mathrm{y}_p\left(\mathbf{w}^T\mathbf{x}_p+\beta\right)\ge 1-a_p, a_p\ge 0, p=1,\ldots,n.$$

Through the Lagrange duality transformation, the final objective function is

$$\mathcal{L}(\mathbf{w},\beta,a)=\sum_{p=1}^{n}a_p-\frac{1}{2}\sum_{p,q=1}^{n}y_p y_q a_p a_q \mathcal{K}\langle\mathbf{x}_p,\mathbf{x}_q\rangle \tag{7}$$

$$\mathrm{s.t.}\mathcal{K}\langle\mathbf{x}_p,\mathbf{x}_q\rangle=\mathbf{x}_p{}^T\mathbf{x}_q,$$

$$\mathbf{w}=\sum_{p=1}^{l}a_p y_p\mathbf{x}_p,\ \sum_{p=1}^{l}a_p y_p=0,$$

$$0\le a_p\le C, p=1,\ldots,n$$

where $\mathcal{K}\langle\mathbf{x}_p,\mathbf{x}_q\rangle$ represents a linear kernel function, $y_p\in\{-1,+1\}$ is the label, n is the training sample, a_p and

are the training parameter.

The steps to solve Eq.(7) are: First, according to the formula seven, we find the minimum values of **w** and β. Second, according to Eq.(7), we find the maximum values of α. Finally, the sequence minimum optimization algorithm (Fan et al. 2005) is used to solve the Lagrange multiplier in the dual problem. Linear SVM is the classifier in our experiments, because this study evaluates different functional network construction for MCI identification. Linear kernel SVM has the advantages of simple use and it is convenient for the comparison of experimental results.

Due to the limited number of samples in the experiments, we use the leave-one-out cross-validation method to evaluate the performance of the proposed method. Specifically, one subject is used for testing while the remaining subjects are used for training the model. We do not want the training set to participate in the testing process of the model as this leads to information leakage and artificially high performance of the model. In feature selection, the training set is further divided into a set for feature selection and an evaluation set.

Experiments and results

We set up 10 step sizes $s = 1, 2, \ldots, 10$ and 10 sliding window lengths $L = 10, 20, \ldots, 100$. LASSO λ feature selection parameter is set within (0, 1). Linear classifier parameter C of SVM is set to one. In this study, a nested leave-one-out cross-validation method is used to determine the model parameter values. This method takes any sample as the test sample and the rest as the training sample. Thus, we can use the training samples to find the optimal classification model, and uses the test sample to evaluate the model. In this way, the number of classifiers corresponding to the sample size and the test results are obtained. Finally, the average of the test results is calculated to measure the performance of the model. However, to determine the value of characteristic selection parameters and avoid data leakage, we further divide the training sample into one sample for evaluation test and the rest sample for model training. In this way, the repeated training model is used to select the optimal super-parameter, and finally, the original sample is utilized to test the optimal super-parameter model. Classification accuracy (ACC), sensitivity (SEN), specificity (SPEC), Youden index (YI), F-score, and balanced accuracy (BAC) are used as evaluation metrics to assess the obtained results. These evaluation metrics are calculated through true positive (TP), false positive (FP), true negative (TN), and false negative (FN) rates:

$$\mathrm{ACC} = (TP + TN)/(TP + TN + FP + FN),$$

$$\mathrm{SEN} = TP/(TP + FN);$$

$$\mathrm{SPEC} = TN/(TN + FP),$$

$$\mathrm{YI} = SEN + SPEC - 1,$$

$$precision = \frac{TP}{TP + FP},$$

$$recall = \frac{TP}{TP + FN},$$

$$\mathrm{F\text{-}score} = \frac{precision \times recall}{precision + recall},$$

$$\mathrm{BAC} = 0.5 \times (recall + SPEC).$$

Receiver operating characteristic curve (ROC) is another used metric to evaluate the performance of the proposed model. It reflects the comprehensive index of specificity and sensitivity. The larger the area under the curve (AUC) is, the better the model's classification performance will be.

Table 2 summarizes the comparative experimental results of different network construction methods in six classification tasks: LMCI vs. NC, EMCI vs. NC, SMC vs. NC, LMCI vs. EMCI, LMCI vs. SMC and EMCI vs. SMC. In this study, four network construction methods are compared, namely: s-LON, s-HON, d-LON, and d-HON. The classification performances with and without using LASSO feature selection are shown in Fig. 4. When the window length is set to 50 and the step size is

Table 2 Experimental results of different network construction methods in six classification tasks

Data	Method	ACC	SEN	SPEC	YI	F-score	BAC	AUC
SMC vs. NC	s-LON	0.56	0.583	0.603	0.175	0.648	0.476	0.554
	s-HON	0.625	0.602	0.689	0.297	0.718	0.592	0.698
	d-LON	0.732	0.698	0.742	0.460	0.744	0.694	0.754
	d-HON	0.789	0.746	0.804	0.543	0.79	0.796	0.834
EMCI vs. NC	s-LON	0.601	0.643	0.647	0.284	0.494	0.53	0.587
	s-HON	0.721	0.703	0.734	0.422	0.632	0.588	0.703
	d-LON	0.783	0.745	0.786	0.53	0.728	0.672	0.794
	d-HON	0.803	0.794	0.856	0.654	0.826	0.768	0.826
LMCI vs. NC	s-LON	0.623	0.616	0.675	0.29	0.628	0.63	0.618
	s-HON	0.742	0.69	0.763	0.456	0.728	0.684	0.732
	d-LON	0.796	0.762	0.842	0.612	0.822	0.752	0.784
	d-HON	0.852	0.805	0.874	0.673	0.864	0.826	0.870
LMCI vs. EMCI	s-LON	0.589	0.662	0.684	0.342	0.688	0.494	0.564
	s-HON	0.654	0.724	0.781	0.513	0.73	0.628	0.611
	d-LON	0.721	0.735	0.796	0.538	0.862	0.704	0.743
	d-HON	0.788	0.796	0.847	0.647	0.914	0.750	0.825
LMCI vs. SMC	s-LON	0.642	0.621	0.643	0.253	0.43	0.552	0.628
	s-HON	0.713	0.694	0.74	0.434	0.652	0.694	0.743
	d-LON	0.79	0.745	0.833	0.573	0.708	0.768	0.771
	d-HON	0.843	0.823	0.876	0.692	0.89	0.804	0.866
EMCI vs. SMC	s-LON	0.618	0.652	0.705	0.354	0.672	0.568	0.621
	s-HON	0.696	0.676	0.736	0.416	0.648	0.662	0.682
	d-LON	0.743	0.703	0.792	0.494	0.73	0.69	0.753
	d-HON	0.802	0.794	0.851	0.645	0.866	0.778	0.831

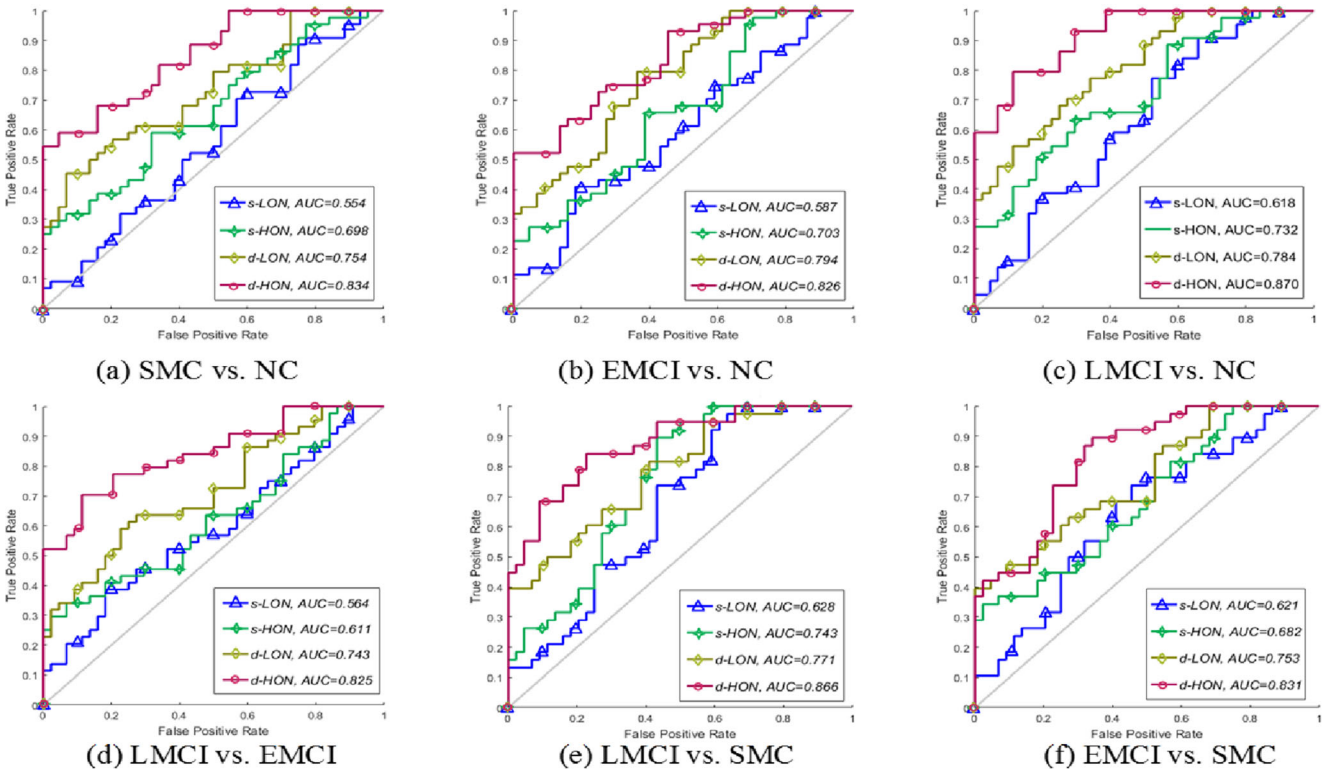

Fig. 4 ROC curve results of different network construction methods in six classification tasks

set to 4, and the best results are obtained with the LASSO feature selection method. We constructed the brain network with the d-HON method, and the experimental results of the six classification tasks without using the feature selection method and using the LASSO feature selection method are compared in Fig. 5. The effect of different window lengths on the experimental results is also investigated. In the six classification tasks, the AUC values of the different length and windows are shown in Fig. 6. In the d-HON method, we report the selected features of three classification tasks LMCI vs. NC, EMCI vs. NC and SMC vs. NC. The top 10 selected brain regions with the highest frequency of occurrence of each classification task are illustrated in Fig. 7.

As shown in Fig. 7, the most frequently identified 10 regions for LMCI vs. NC classification task are right superior gyrus, left and right medial frontal orbital gyrus, right inferior frontal orbital gyrus, left posterior cingulate gyrus, left and right hippocampus, right parahippocampal gyrus, left superior occipital gyrus and left angular gyrus. The most frequently identified 10 regions for EMCI vs. NC task are right superior frontal gyrus, left inferior frontal orbital gyrus, right gyrus rectus, left posterior cingulate gyrus, left and right hippocampus, right parahippocampal gyrus, left cuneus, right superior occipital gyrus and left middle temporal gyrus. The top 10 identified regions with the highest frequency for SMC vs. NC task are right superior frontal gyrus, left and right middle frontal gyrus, left and right middle frontal orbital gyrus, left and right hippocampus, right cuneus, left lingual gyrus and left middle temporal gyrus. Table 3 shows the comparison of this paper with other different methods. We can observe that our proposed method has achieved promising performance for the early identification of MCI.

Discussion

Table 2 shows the results of different classification tasks. We observe that the d-HON method achieves the best classification results in the six classification tasks. The best accuracy and AUC results for LMCI vs. NC classification are 85.2% and AUC is 0.87, respectively, which are higher than EMCI vs. NC and SMC vs. NC classification results. The reason is that LMCI brain function degradation is more severe than EMCI and SMC, so the classification feature is more discriminative using the proposed model. Results in Table 2 show that dynamic functional network and high-order functional network methods improve the classification accuracy by 13.8% and 6.9% on average compared with the static functional network and the low-order functional network methods, respectively. It can be seen that the dynamic high-order network model improves the disease classification due to considering dynamic and high-order characteristics of brain network. Similarly, it can be seen from the ROC curve in Fig. 4 that the d-HON method achieves the best classification effect

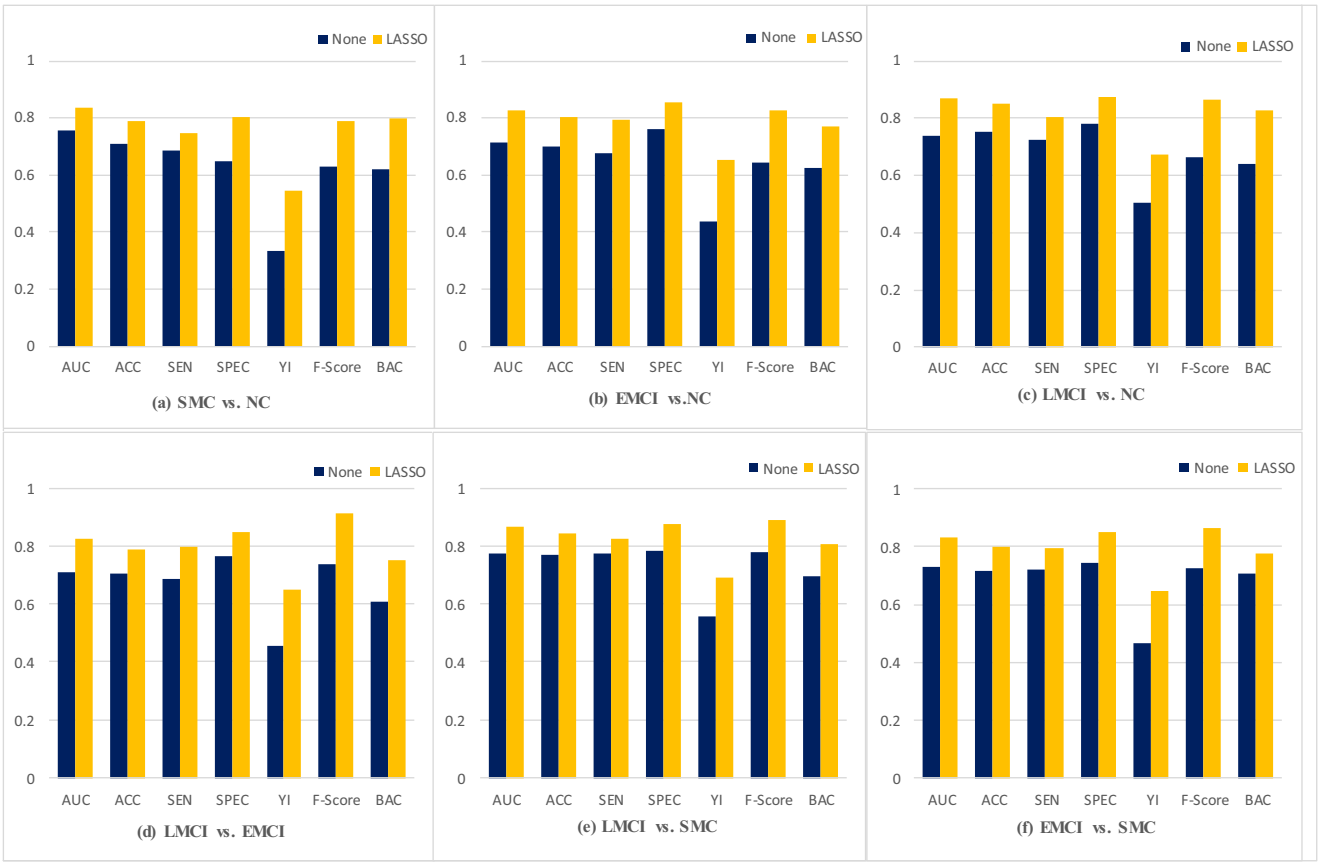

Fig. 5 The results of with and without feature selection methods in six classification tasks

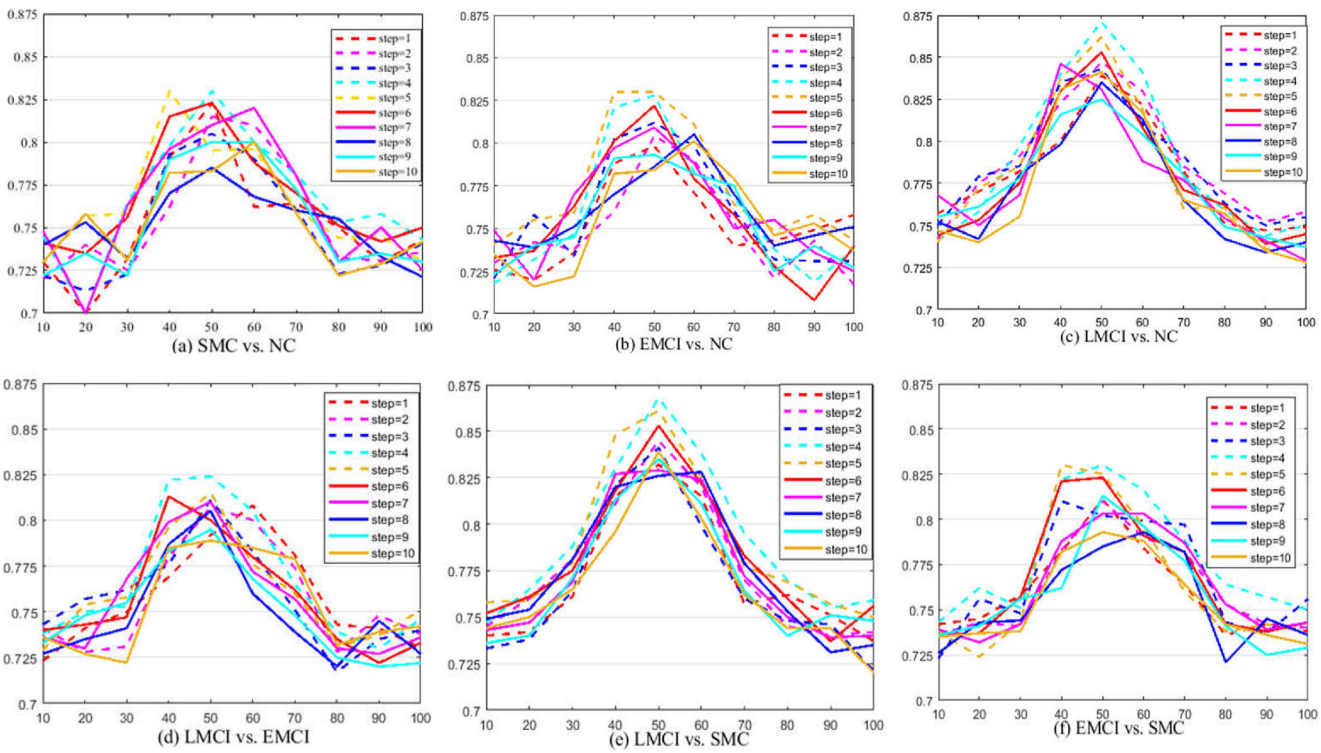

Fig. 6 The AUC values of asynchronous length and different windows in six classification tasks, (gray dotted line is the average of the results of 10 steps)

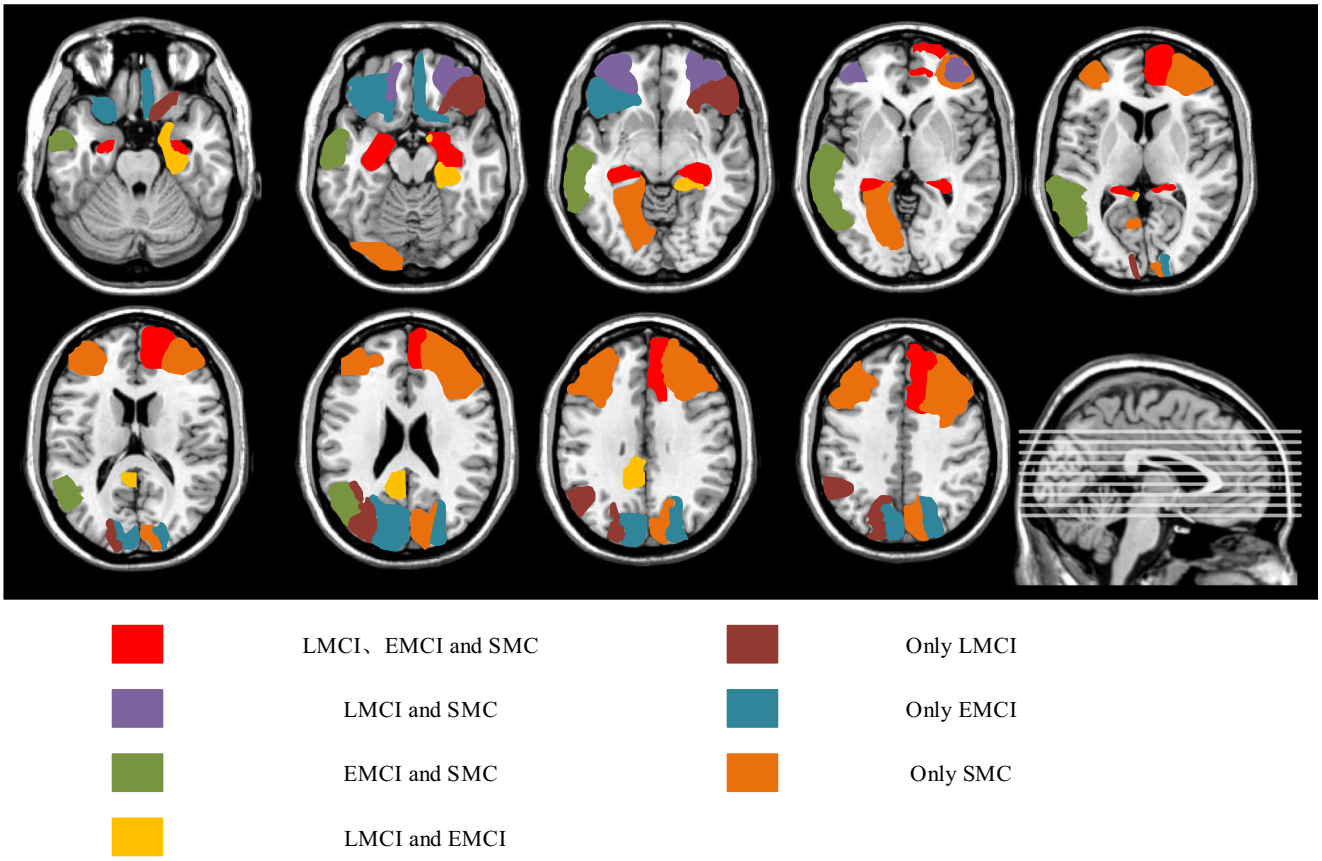

Fig. 7 For LMCI vs. NC, EMCI vs. NC and SMC vs. NC classification task, we select the top 10 most brain regions using statistical occurrences. The red area represents the three-group classification task appear in high frequency. The purple area in LMCI vs. NC and SMC vs. NC appear in high frequency. The green area in EMCI vs. NC and SMC vs. NC appear in high frequency. Yellow in LMCI vs. NC and EMCI vs. NC appear in a high frequency area, other colors said only in each classification task of statistics in high frequency

among the other four methods. We also analyze the effect of feature selection method on the performance of our model. Compared with the method without feature selection, it can be seen from Fig. 5 that adding LASSO feature selection method improves the classification accuracy by 9.6%. We also experiment on different window lengths and asynchronous lengths, as shown in Fig. 6. The results show that the window length between 40 and 60 have better classification

Table 3 Comparison of other methods

Reference	Subject	Method	Task	ACC	SEN	SPEC
(Supekar et al. 2008)	21 AD,18NC	Wavelet correlation networks, 90 ROIs, clustering coefficient	–	–	0.72	0.78
(Chen et al. 2011)	20 AD,15MCI,20NC	PCC, 116 ROIs, large-scale network	AD vs. MCI + NC	0.82	0.85	0.80
			MCI vs.NC	0.91	0.93	0.90
(Challis et al. 2015)	27 AD,50MCI,39NC	Covariance, 82ROIs, Bayesian Gaussian process logistic regression	AD vs. MCI	0.8	0.7	0.9
			MCI vs.NC	0.75	1	0.5
(de Vos et al. 2018)	77 AD,173NC	Sparse partial correlation FCNs, 70 ROIs	–	0.75	0.79	0.71
		Sparse partial correlation dynamic FCNs, 70 ROIs	–	0.78	0.83	0.73
		Amplitude of low frequency fluctuation	–	0.76	0.71	0.82
		Combined	–	0.79	0.86	0.71
This work	38LMCI,44EMCI, 44SMC,44NC	d-HON, 90 ROIs	LMCI vs. NC	0.852	0.805	0.874
			EMCI vs.NC	0.803	0.794	0.856
			SMC vs. NC	0.789	0.746	0.804

performance and this also follows previous studies (Hutchison et al. 2013). For this work, we find that the best values of window length and step size are 50 and 4, respectively, which yields the best classification results.

According to the high-frequency brain regions extracted by the d-HON method, the brain regions of the three tasks are hippocampus and right superior frontal gyrus, as shown in red area in Fig.7. Relevant studies have found that left and right hippocampus are playing an important role in helping human short-term memory, long-term memory, and spatial localization, and are associated with the occurrence and development of AD. The right superior frontal gyrus is closely associated with depression, and studies have found that depression is also closely associated with AD. The left posterior cingulate gyrus and the right parahippocampal gyrus are shown in Fig. 7 in yellow area, the left middle temporal gyrus in green area, and the right middle frontal gyrus in purple area. The results of previous studies are further verified (Echavarri et al. 2011). These common high-frequency brain regions have been reported in AD-related studies.

Moreover, existing studies have shown disease-related biomarkers change in different stages of MCI. For example, Jagust et al. (Jagust et al. 2015) found florbetapir was associated with cognitive change in normal subjects only, whereas in the LMCI group, fluorodeoxyglucose (FDG) was more closely associated with cognitive change than was florbetapir. This follows a model in which amyloid changes precede neurodegeneration (measured by FDG), which is tied to subsequent cognitive decline. Also, Li et al.(Li et al. 2017) found there are cortical areas with significant association with Aβ deposition. Specifically, ROIs in C1(e.g., rh_inferior_temporal) exhibited steady negative significant correlation trends among healthy control (CN) and EMCI groups; C2(e.g., bi_fusiform, bi_precuneus, bi_superior_ temporal, bi_lateral_occipital, lh_medialorbitofrontal, lh_isthmuscingulate, rh_superamarginal, rh_middle_ temporal) showed reversed correlation in EMCI stage; C3(e.g., lh_inferior_temporal, lh_middle_temporal, lh_banks_of_the_ Superior_semporal_su-lcus(lh_bankssts), rh_inferior_parietal) showed such reversed relationship also in EMCI stage, which were maintained in the LMCI stage; C4(e.g., lh_inferior_parietal) showed negative correlation only in the LMCI stage. On the other hand, ROIs in C5 (e.g., rh_postcentral, rh_transversetemporal, rh_precentral)presented positive significant correlation in CN stage; C6(e.g., bi_entorhinal, bi_temporal_pole) showed such correlation in the AD stage.

In SMC vs. NC classification process, its particular brain regions for frontal medial gyrus, right cuneus and right lingual gyrus. This study found that the right frontal medial gyrus is related to working memory. The left frontal medial gyrus is mainly responsible for handling and processing in Chinese. In addition, it has function of memory and information coordination and find that its activity intensity gradually increased with the increase of complexity of text. The main manifestation of SMC is presented with cognitive impairment, in line with the performance of the SMC memory loss. Left lingual gyrus is mainly responsible for human visual perception, but studies have also found that lingual gyrus is involved in logical analysis. The right cuneus processes visual information, and studies have found that visual acuity is associated with spatial orientation. During the EMCI vs. NC classification, the unique brain regions are the left inferior frontal orbital gyrus, the right gyrus rectus, left cuneus and right superior occipital. Left inferior frontal orbital gyrus is mainly related to language processing. Some studies have found that the right gyrus rectus is related to AD, the left cuneus is mainly related to visual information processing, and some studies have found that the right superior occipital gyrus is related to MCI. In the classification process of LMCI vs. NC, its unique brain regions are right inferior frontal orbital gyrus, left superior occipital gyrus, and left angular gyrus. Some studies have found that the left superior occipital gyrus is related to MCI, while some studies have found that the left angular gyrus is related to MCI. Resection of the angular gyrus will lead to visual and auditory image disconnection, which will cause dyslexia and audio-visual aphasia. These results prove that the model is not only superior in classification results but also interpretable and can provide clinical diagnostic criteria.

Conclusions

In this paper, a new sliding window method is used to construct a dynamic functional network. Brain network dynamics provides richer information. This paper considers higher-order characteristics of the brain network and constructs a higher-order brain network through topology aware between brain networks. The model then uses local weight clustering to extract effective features of dynamic high-order networks and usies LASSO for network feature selection. In the six classification tasks, the proposed method obtains the best results. However, our study has some limitations. First, we do not consider biological mechanisms in our model and is purely data driven. Second, our model only uses single modality imaging (R-fMRI) and the results should be replicated on a larger database. Using multimodal imaging and more training samples can potentially improve the accuracy of classification. We will address these limitations in our future work.

Acknowledgments This work was supported partly by National Natural Science Foundation of China (Nos.61871274, 61801305 and 81571758), National Natural Science Foundation of Guangdong Province (No. 2017A030313377), Guangdong Pearl River Talents Plan (2016ZT06S220), Shenzhen Peacock Plan (Nos. KQTD2016053112051497 and KQTD2015033016 104926), and Shenzhen Key Basic Research Project (Nos.

JCYJ20170413152804728,JCYJ20180507184647636, JCYJ20170818142347251 and JCYJ20170818094109846).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. HoffmannLa Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; MesoScale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval This article contains no studies with human participants or animals performed by any of the authors.

References

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex, 24*(3), 663–676.

Alzheimer, A. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimer's & Dementia, 14*(3), 367–429.

Bassett, D. S., & Bullmore, E. (2006). Small-world brain networks. *The Neuroscientist, 12*, 512–523.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34*(4), 537–541.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience, 10*(3), 186–198.

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

Chang, C., & Glover, G. H. (2010). Time–frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage, 50*, 81–98.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST), 2*(3), 27–65.

Chen, G., Ward, B. D., Xie, C., Li, W., Wu, Z., et al. (2011). Classification of Alzheimer disease, mild cognitive impairment, and Normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Radiology, 259*(1), 213–221.

Chen, X., Zhang, H., Gao, Y., Wee, C. Y., Li, G., & Shen, D. (2016). High-order resting-state functional connectivity network for MCI classification. *Human Brain Mapping, 37*(9), 3282–3296.

Damaraju, E., Allen, E., Belger, A., Ford, J., McEwen, S., et al. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical, 5*, 298–308.

De Vos, F., Koini, M., Schouten, T. M., Seiler, S., Van der Grond, J., et al. (2018). A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. *NeuroImage, 167*, 62–72.

Echavarri, C., Aalten, P., Uylings, H. B., Jacobs, H., Visser, P. J., et al. (2011). Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease. *Brain Structure and Function, 215*, 265–271.

Fan, R.-E., Chen, P.-H., & Lin, C.-J. (2005). Working set selection using second order information for training support vector machines. *Journal of Machine Learning Research, 6*, 1889–1918.

Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity, 1*(1), 13–36.

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences, 101*(13), 4637–4642.

Herculano-Houzel, S. (2012). The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proceedings of the National Academy of Sciences, 109*, 10661–10668.

Huettel SA, Song AW, McCarthy G. (2004). Functional magnetic resonance imaging. Sinauer Associates Sunderland.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., Handwerker, D. A., Keilholz, S., Kiviniemi, V., Leopold, D. A., de Pasquale, F., Sporns, O., Walter, M., & Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage, 80*, 360–378.

Jafri, M. J., Pearlson, G. D., Stevens, M., & Calhoun, V. D. (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage, 39*(4), 1666–1681.

Jagust, W. J., Landau, S. M., Koeppe, R. A., Reiman, E. M., Chen, K., et al. (2015). The Alzheimer's disease neuroimaging initiative 2 PET core: 2015. *Alzheimer's & Dementia, 11*(7), 757–771.

Jia, J., Wei, C., Chen, S., Li, F., Tang, Y., et al. (2018). The cost of Alzheimer's disease in China and re-estimation of costs worldwide. *Alzheimer's & Dementia, 14*(4), 483–491.

Li C, Fang C, Cabrerizo M, Barreto A, Andrian J, et al. (2017) Pattern analysis of the interaction of regional amyloid load, cortical thickness and APOE genotype in the progression of Alzheimer's disease. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)2017*: 2171–2176. IEEE.

Lindquist, M. A., Xu, Y., Nebel, M. B., & Caffo, B. S. (2014). Evaluating dynamic bivariate correlations in resting-state fMRI: A comparison study and a new approach. *NeuroImage, 101*, 531–546.

Liu, J., Ji, S., & Ye, J. (2009). SLEP: sparse learning with efficient projections. *Arizona State University, 6* (191), 7.

Lui, S., Wu, Q., Qiu, L., Yang, X., Kuang, W., Chan, R. C., Huang, X., Kemp, G. J., Mechelli, A., & Gong, Q. (2011). Resting-state functional connectivity in treatment-resistant depression. *American Journal of Psychiatry, 168*(6), 642–648.

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C., et al. (2005). The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics, 15*(4), 869–877.

Niethammer, M., Feigin, A., & Eidelberg, D. (2012). Functional neuroimaging in Parkinson's disease. *Cold Spring Harbor Perspectives in Medicine*. https://doi.org/10.1101/cshperspect.a009274.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences, 87*(24), 9868–9872.

Qi, S., Meesters, S., Nicolay, K., ter Haar Romeny, B. M., & Ossenblok, P. (2015). The influence of construction methodology on structural brain network measures: A review. *Journal of Neuroscience Methods, 253*, 170–182.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage, 52*(3), 1059–1069.

Shi, J., Zheng, X., Li, Y., Zhang, Q., & Ying, S. (2017). Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics, 22*(1), 173–183.

Shi, J., Xue, Z., Dai, Y., Peng, B., Dong, Y., et al. (2018). Cascaded multi-column RVFL+ classifier for single-modal neuroimaging-based diagnosis of Parkinson's disease. *IEEE Transactions on Biomedical Engineering, 66*(8), 2362–2371.

Supekar, K., Menon, V., Rubin, D., Musen, M., & Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Computational Biology, 4*(6), e1000100.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological), 58*(1), 267–288.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical Parcellation of the MNI MRI single-subject brain. *NeuroImage, 15*(1), 273–289.

van den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology, 20*(8), 519–534.

Visser, P., & Tijms, B. (2017). Brain amyloid pathology and cognitive function: Alzheimer disease without dementia? *JAMA, 317*(32), 2285–2287.

Wang, J., Zuo, X., & He, Y. (2010). Graph-based network analysis of resting-state functional MRI. *Frontiers in Systems Neuroscience, 4*, 16–31.

Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., & He, Y. (2015). GRETNA: A graph theoretical network analysis toolbox for imaging connectomics. *Frontiers in Human Neuroscience, 9*, 386.

Wang, J., Wang, Q., Peng, J., Nie, D., Zhao, F., et al. (2017). Multi-task diagnosis for autism spectrum disorders using multi-modality features: A multi-center study. *Human Brain Mapping, 38*(6), 3081–3097.

Weng, S.-J., Wiggins, J. L., Peltier, S. J., Carrasco, M., Risi, S., Lord, C., & Monk, C. S. (2010). Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders. *Brain Research, 1313*, 202–214.

Yaesoubi, M., Allen, E. A., Miller, R. L., & Calhoun, V. D. (2015). Dynamic coherence analysis of resting fMRI data to jointly capture state-based phase, frequency, and time-domain information. *Neuroimage, 120*, 133–142.

Zhang, Y., Zhang, H., Chen, X., Lee, S.-W., & Shen, D. (2017). Hybrid high-order functional connectivity networks using resting-state functional MRI for mild cognitive impairment diagnosis. *Scientific Reports, 7*(1), 6530–6544.

Zhang, D., Zun, Q., Hao, X., Shao, W., Wang, M., et al. (2018). Intelligent analysis of brain images. *SCIENTIA SINICA Informationis, 48*, 589–602.

Zheng W, Zhu X, Wen G, Zhu Y, Yu H, Gan J. 2018. Unsupervised feature selection by self-paced learning regularization. *Pattern Recognition Letters*: https://doi.org/10.1016/j.patrec.2018.06.029.

Zhou, T., Liu, M., Thung, K. H., & Shen, D. (2019a). Latent representation learning for Alzheimer's disease diagnosis with incomplete multi-modality neuroimaging and genetic data. *IEEE Transactions on Medical Imaging DOI*. https://doi.org/10.1109/TMI.2019.2913158.

Zhou, T., Thung, K. H., Zhu, X., & Shen, D. (2019b). Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Human Brain Mapping, 40*(3), 1001–1016.

Zhu, X., Zhang, S., Hu, R., He, W., Lei, C., & Zhu, P. (2018a). One-step multi-view spectral clustering. *IEEE Transactions on Knowledge Data Engineering Human Brain Mapping, 40*(3), 1001–1016.

Zhu, X., Zhang, S., Li, Y., Zhang, J., Yang, L., & Fang, Y. (2018b). Low-rank sparse subspace for spectral clustering. *IEEE Transactions on Knowledge Data Engineering, 31*(8), 1532–1543.