



Learning Brain Connectivity Sub-networks by Group- Constrained Sparse Inverse Covariance Estimation for Alzheimer's Disease Classification

Yang Li^{1,2,3,4}, Jingyu Liu¹, Jie Huang¹, Zuoyong Li⁴ and Peipeng Liang^{5,6*}

¹ School of Automation Sciences and Electrical Engineering, Beihang University, Beijing, China, ² Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China, ³ Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing, China, ⁴ Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fujian, China, ⁵ Department of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China, ⁶ School of Psychology, Capital Normal University, Beijing, China

OPEN ACCESS

Edited by:

Pedro Antonio Valdes-Sosa,
Clinical Hospital of Chengdu Brain
Science Institute, China

Reviewed by:

Mingli Zhang,
Montreal Neurological Institute,
McGill University, Canada

Yong Xu,
First Hospital of Shanxi Medical
University, China

*Correspondence:

Peipeng Liang
p.p.liang@163.com

Received: 09 March 2018

Accepted: 17 August 2018

Published: 07 September 2018

Citation:

Li Y, Liu J, Huang J, Li Z and Liang P
(2018) Learning Brain Connectivity
Sub-networks by Group- Constrained
Sparse Inverse Covariance Estimation
for Alzheimer's Disease Classification.
Front. Neuroinform. 12:58.
doi: 10.3389/fninf.2018.00058

Background/Aims: Brain functional connectivity networks constructed from resting-state functional magnetic resonance imaging (rs-fMRI) have been widely used for classifying Alzheimer's disease (AD) from normal controls (NC). However, conventional correlation analysis methods only capture the pairwise information, which may not be capable of revealing an adequate and accurate functional connectivity relationship among brain regions in the whole brain. Additionally, the non-sparse connectivity networks commonly contain a large number of spurious or insignificant connections, which are inconsistent with the sparse connectivity of actual brain networks in nature and may deteriorate the classification performance of Alzheimer's disease.

Methods: To address these problems, in this paper, a new classification framework is proposed by combining the Group-constrained topology structure detection with sparse inverse covariance estimation (SICE) method to build the functional brain sub-network for each brain region. Particularly, to tune the sensitive analysis of the regularized parameters in the SICE method, a nested leave-one-out cross-validation (LOOCV) method is adopted. Sparse functional connectivity networks are thus effectively constructed by using the optimal regularized parameters. Finally, a decision classification tree (DCT) classifier is trained for classifying AD from NC based on these optimal functional brain sub-networks. The convergence performance of our proposed method is furthermore evaluated by the trend of coefficient variation.

Results: Experiment results indicate that a LOOCV classification accuracy of 81.82% with a sensitivity of 80.00%, and a specificity of 83.33% can be obtained by using the proposed method for the classification AD from NC, and outperforms the most state-of-the-art methods in terms of the classification accuracy. Additionally, the experiment results of the convergence performance further suggest that our proposed scheme has a high rate of convergence. Particularly, the abnormal brain regions and functional connections identified by our proposed framework are highly associated with the

underpinning pathological mechanism of the AD, which are consistent with previous studies.

Conclusion: These results have demonstrated the effectiveness of the proposed Group-constrained SICE method, and are capable of clinical value to the diagnosis of Alzheimer's disease.

Keywords: Alzheimer's disease (AD), resting-state fMRI, Group-constrained topology structure detection, sparse inverse covariance estimation (SICE), functional connectivity network, classification

INTRODUCTION

Alzheimer's disease (AD) is one of the most common neurodegenerative diseases and generally causes impairments in multiple cognitive functions such as memory, attention, verbal, and visuospatial abilities (Alzheimer's Association, 2015). This disease adversely affects the daily life of AD patients, and eventually results in death. Currently, there is no clinical cure method for the AD so far (Alzheimer's Association, 2015). Considering the high incidence of the AD, it is highly required for the precise diagnosis of the AD individually so that some effective behavioral or pharmacological treatment can be adopted to alleviate the symptoms and delay the progression of the AD (Brookmeyer et al., 2007).

Recently, functional connectivity networks constructed from the functional magnetic resonance image (fMRI) hold great promise for distinguishing AD patients from NC (Rosa et al., 2015). Numerous functional connectivity modeling methods have been proposed to investigate brain functional connectivity activities, including correlation-based and sparse representation-based methods (Wee et al., 2014; Meszlényi et al., 2017). For example, Khazaei et al. (2016) constructed the functional connectivity networks for AD classification by using a Pearson correlation-based method and the weak connections were removed by maximizing the global cost efficiency (GCE) in the networks. Their findings demonstrated that the region of interests (ROI) of the Thalamus, Paracentral_Lobule, and Temporal_Pole_Mid are significantly different between the AD and NC. Furthermore, a high-order functional connectivity network was discussed based on the sliding window, the Pearson correlation coefficient and minimum spanning tree method for AD classification Guo et al. (2017). Their experiment results indicated that the functional connectivity between the Precentral and Supp_Motor_Area is the discriminative feature for identifying individuals with the AD from NC subjects. Among these methods, the correlation-based methods generally obtain relatively high sensitivity in detecting network connections due to the negative connections in the connectivity networks. However, correlation-based network analysis can only capture the pairwise information and does not effectively characterize the functional interactions among many brain regions working together (Huang et al., 2010; Li et al., 2014). In addition, correlation-based fully connected functional networks may contain a large number of spurious or insignificant connections, and may deteriorate the classification accuracy and generalization performance of the classifiers. Recent work has

shown that robust connections can be elucidated from a set of noisy connections when certain sparsity constraints are imposed on the connectivity networks (Supekar et al., 2008; Zanin et al., 2012). The sparsity constraint is validly correlated with the small-world property, where a single brain region usually only interacts with a small number of other brain regions (Stam et al., 2007; Supekar et al., 2008; Li et al., 2014).

Recently, the sparse inverse covariance estimation (SICE) method has been widely applied to constructing brain functional connectivity and detecting the connectivity difference between patients with neurological diseases and NC (Huang et al., 2010; Rosa et al., 2015; Zhang et al., 2015; Qureshi et al., 2017). For example, Huang et al. (2010) employed the SICE scheme in the positron emission tomography (PET) data to observe brain connectivity difference between AD and NC. Rosa et al. (2015) further combined the SICE with a sparse discriminative classifier (linear l_1 -norm support vector machine (SVM)) to discriminate major depressive disorder (MDD) patients from NC. However, the SICE is a sparse constraint method applied at the individual level, inevitably generating different network topologies for different subjects and causing the inter-subject variability (Wee et al., 2014). This deficient may degrade the generalization performance of the classifiers due to the difference of the different functional connectivity among subjects (Wee et al., 2014). Additionally, the SICE method commonly may encounter the overfitting problem due to the limitation of the small sample size, which is a main challenge in computer-aided diagnosis of the AD (Zhang et al., 2018).

To overcome these deficiencies, a new classification framework based the functional brain sub-network is proposed. The key of the proposed framework is that a Group-constrained topology structure detection algorithm is first used to select the most discriminative brain regions, indicating that the connection topology is identical among subjects. Then the SICE algorithm is further employed to preserve the individual information via different connectivity values. Specifically, the Group-constrained topology structure detection algorithm utilizes a $l_{2,1}$ -norm penalization term to encourage an identical connection topology among subjects and minimize the inter-subject variability, and thus a better classification performance was achieved (Wee et al., 2014). Furthermore, the sample size requirement of the Group-constrained topology structure detection is much weaker than that of traditional l_1 -norm sparse methods (Mittra and Zhang, 2016). Therefore, the proposed Group-constrained connectivity structure detection scheme can alleviate the small sample size problem. Finally, a nested LOOCV scheme is implemented

to tune the regularization parameter in the functional brain sub-network construction process and further to evaluate the effectiveness of the proposed framework. Experimental results demonstrate that our classification framework achieved a high cross-validation accuracy of 81.82%, outperforming the competing methods with a relatively large margin. Furthermore, the area under ROC curve (AUC), regarded as a metric of diagnostic power, of 0.9667 demonstrates the efficacy of our proposed framework in extracting discriminative information for diagnosing the AD. Particularly, the abnormal functional connections identified by our proposed framework are highly associated with the underpinning pathological mechanism of the AD, which are in line with previous studies.

The remaining sections are organized as follows. An explicit description is provided for fMRI dataset acquisition and the process of constructing functional brain sub-networks in Materials and Method sections. Then, the classification performances of different brain network construction methods as well as the most discriminative connections are presented in Results and Discussion sections. Finally, we conclude this study in Conclusion section.

MATERIALS

Subjects

Sixty-two right-handed subjects (30 AD and 32 healthy controls) participated in this study after giving written informed consent. AD subjects were recruited from patients who had consulted a memory clinic for memory complaints at Xuanwu Hospital, Capital Medical University, Beijing, China. The healthy elderly controls were recruited from the local community through advertisements. This study was approved by the Medical Research Ethics Committee of Xuanwu Hospital, Capital Medical University.

All AD patients underwent a complete physical and neurological examination, standard laboratory tests, and an extensive battery of neuropsychological assessments. The diagnosis of AD fulfilled the Diagnostic and Statistical Manual of Mental Disorders 4th Edition criteria for dementia (Frances et al., 1994), and the National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer Disease and Related Disorders Association (NINCDS-ADRDA) criteria for possible or probable AD (McKhann et al., 1984). The subjects were assessed with the Clinical Dementia Rating (CDR) score (Morris, 1993) as (mainly) being in the early-stages of the AD (2 patients with CDR = 2, 14 patients with CDR = 1 and 14 patients with CDR = 0.5).

The criteria for healthy controls were as follows: (1) no neurological or psychiatric disorders such as stroke, depression, epilepsy; (2) no neurological deficiencies such as visual or hearing loss; (3) no abnormal findings such as infarction or focal lesion in conventional brain MR imaging; (4) no cognitive complaints; (5) MMSE score of 28 or higher; (6) CDR score of 0.

Data from seven subjects (five AD patients and two healthy controls) were excluded due to excessive motion. Thus, the remaining 55 participants were included in the following data analysis. There were no significant differences between the two

groups in gender, age, and years of education, but the MMSE scores were significantly different ($p < 0.001$) between two groups.

Data Acquisition and Pre-processing

MRI data acquisition was performed on a SIEMENS Trio 3T scanner (Siemens, Erlangen, Germany). Foam padding and headphones were used to limit head motion and reduce scanner noise. The subjects were instructed to hold still, keep their eyes closed and think nothing in particular. Functional images were collected axially by using an echo-planar imaging (EPI) sequence [repetition time (TR)/echo time (TE)/flip angle (FA)/field of view (FOV) = 2,000 ms/40 ms/90°/24 cm, resolution = 64 × 64 matrix, slices = 28, thickness = 4 mm, gap = 1 mm, bandwidth = 2232 Hz/pixel]. The scan lasted for 478 s. 3D T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) sagittal images were collected by using the following parameters: TR/TE/inversion time (TI)/FA = 1900 ms/2.2 ms/900 ms/9°, resolution = 256 × 256 matrix, slices = 176, thickness = 1 mm.

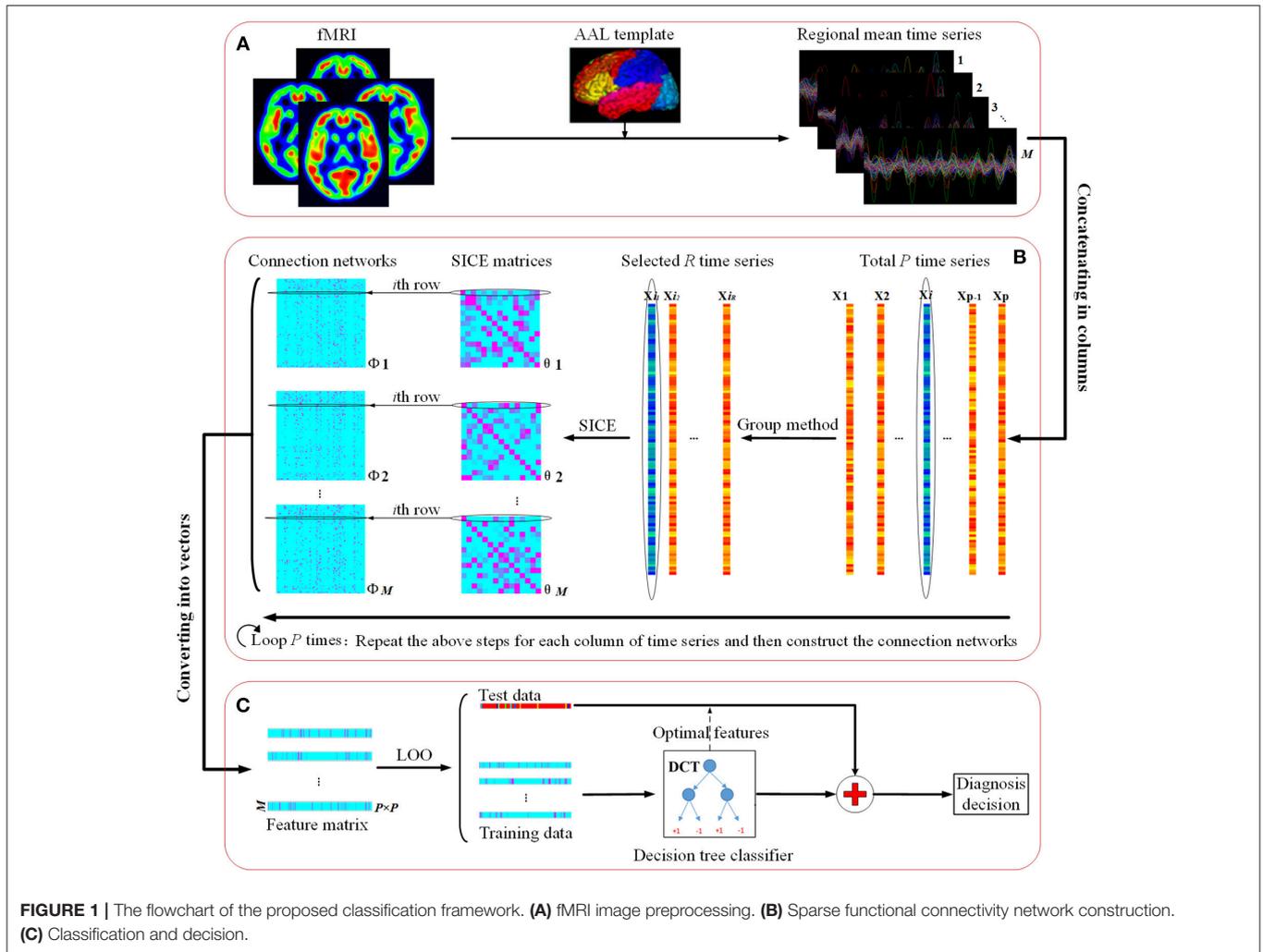
Unless otherwise stated, all preprocessing procedures were conducted using the toolbox of Data Processing Assistant for Resting-State fMRI (DPARSF) (Chao-Gan and Yu-Feng, 2010). Particularly, the first 10 volumes of the functional images were discarded for the signal equilibrium and participants' adaptation to the scanning noise. The remaining 229 fMRI images were first corrected for within-scan acquisition time differences between slices and then realigned to the first volume to correct for inter-scan head motions. No participant had a head motion of more than 1.5 mm maximum displacement in any of the x , y , or z directions and 1.5° of any angular motion throughout the course of the scan. The individual structural image was co-registered to the mean functional image after motion correction using a linear transformation. The transformed structural images were then segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) by using a unified segmentation algorithm (Ashburner and Friston, 2005). The motion corrected functional volumes were spatially normalized to the Montreal Neurological Institute (MNI) space and re-sampled to 3 mm isotropic voxels using the normalization parameters estimated during unified segmentation.

Subsequently, the functional images were spatially smoothed with a Gaussian kernel of 6 × 6 × 6 mm³ full width at half maximum (FWHM) to decrease spatial noise. Following this, temporal filtering (0.01–0.08 Hz) was applied to the time series of each voxel to reduce the effect of low-frequency drifts and high-frequency noise. To further reduce the effects of confounding factors, we also used a linear regression process to further remove the effects of head motion and other possible sources of artifacts (Fox et al., 2005): (1) six motion parameters; (2) whole-brain signal averaged over the entire brain; (3) linear drift.

METHODOLOGY

Overview

Figure 1 illustrates the flowchart of our proposed classification framework, which includes the image preprocessing, sparse



network construction and the evaluation of the classification performance. Specifically, fMRI images are first parceled into 90 regions-of-interest (ROIs) based on the Automated Anatomical Labeling (AAL) template (Tzourio-Mazoyer et al., 2002), and thus regional mean time series are acquired for each ROI. Second, the most discriminative brain regions subset for each ROI is detected by using the proposed Group-constrained topology structure detection algorithm, where ROIs subset with the smallest Bayesian information criterion (BIC) score is finally selected as the most discriminative brain regions. Furthermore, a functional connectivity network for each subject is constructed by using the SICE method. Finally, a LOOCV framework, together with an optimal DCT classifier, is implemented to evaluate the generalization performance of classification tasks.

Construction of Sparse Functional Network and Feature Extraction

The Group-constrained topology structure detection algorithm is first used to detect the most discriminative ROIs subset for each brain region. Then, a detailed description of the SICE method is presented to construct the functional connectivity

network for each subject and a feature matrix is further built for the classification. Particularly, a non-sparse brain network construction method is also given as a comparison with the proposed method.

Group-Constrained Topology Structure Detection

Suppose that there are M subjects ($M = 55$), each subject includes the total of P ROIs ($P = 90$), and each ROI time series includes n_t time points ($n_t = 229$), the Group-constrained topology structure detection algorithm is first used to select the most discriminative ROIs by solving the following optimization rule (Wee et al., 2014):

$$\hat{\beta}_p = \operatorname{argmin} \left(\sum_{m=1}^M \left\| \mathbf{x}_p^m - \mathbf{X}_p^m \beta_p^m \right\|_2^2 + \varphi \left\| \beta_p \right\|_{2,1} \right), \quad (1)$$

where $\mathbf{x}_p^m = [x_p^m(1), x_p^m(2), \dots, x_p^m(n_t)]^T$ ($p = 1, 2, \dots, P$) denotes the fMRI time series for the p th ROI of the m th subject, $\mathbf{X}_p^m = [x_1^m, x_2^m, \dots, x_{p-1}^m, x_{p+1}^m, \dots, x_p^m]$ is a matrix including all ROIs time series of the m th subject except the p th

ROI, $\beta_p^m = [\beta_{p,1}^m, \beta_{p,2}^m, \dots, \beta_{p,p-1}^m, \beta_{p,p+1}^m, \dots, \beta_{p,p}^m]^T$ is the coefficient vector for the p th ROI of the m th subject, $\beta_p = [\beta_p^1, \beta_p^2, \dots, \beta_p^M]$ is the coefficient matrix for the p th ROI of all subject, $\|\bullet\|_2^2$ means the square of the l_2 -norm of a vector, i.e., the quadratic sum of all the elements in the vector, and $\|\bullet\|_{2,1}$ is the summation of the l_2 -norm of each row in the matrix. The nonzero elements in the coefficient matrix $\hat{\beta}_p$ are corresponding to the most discriminative ROIs for the p th ROI. The $l_{2,1}$ -norm in Equation (1) encourages an identical optimal ROIs subset for the p th ROI among subjects. The regularization parameter $\varphi \in (0, 1)$ controls the trade-off between fidelity term $\sum_{m=1}^M \|\mathbf{x}_p^m - \mathbf{X}_p^m \beta_p^m\|_2^2$ and regularization term $\|\beta_p\|_{2,1}$. Specifically, the Bayesian Information Criterion (BIC) (Schwarz, 1978) is used to optimize the regularization parameter φ :

$$BIC(\varphi) = -2L(\beta(\varphi)) + d(\varphi) \log n_t \quad (2)$$

where $L(\beta(\varphi))$ is the log-likelihood function and $d(\varphi)$ was the degree of freedom. The BIC score is defined by:

$$BIC = \sum_{m=1}^M \ln ESS(m) + \frac{d \ln n_t}{n_t} \quad (3)$$

where $ESS(m)$ is the residual sum of squares for the m th subject and d denotes the number of nonzero coefficients in the matrix $\hat{\beta}_p$. For each φ , a BIC score is calculated by using Equation (3) and the subset of ROIs with smallest BIC score are finally determined as the most discriminative brain regions for the p th ROI. Thus, an optimal ROI subset E_p for the p th ROI is obtained. Repeat the same steps for all ROIs, the most discriminative ROIs subset for each ROI can be obtained.

Sparse Network Construction

Many methods have been proposed to construct the human brain connectivity networks, among which the sparse representation-based method (Tibshirani et al., 2005; Wright et al., 2009; Wee et al., 2016) and the correlation-based method (Wee et al., 2012; Jie et al., 2014) are two popular approaches. Compared with the correlation-based method, the sparse representation-based method usually has much better discriminability due to the small-world properties and scale-free attributes of human brain connectivity networks. Therefore, a sparse representation-based method (SICE) in this study was adopted to construct the human brain networks. In addition, a non-sparse network was also used for a comparison to show the advantages of our proposed method.

Non-sparse network construction

Partial correlation matrix Π was used as the comparison in this study, which measures the relationship between two time series after excluding the effects of all other time series. Specifically, if the correlation matrix Σ is positively defined and therefore invertible, then partial correlation matrix Π can be acquired from

the full inverse covariance matrix Σ^{-1} by using the following formula:

$$\Pi_{ij} = \frac{-(\Sigma^{-1})_{ij}}{\sqrt{(\Sigma^{-1})_{ii}(\Sigma^{-1})_{jj}}} \quad (i \neq j), \quad (4)$$

Considering the fact that the length of the regional mean time series (229 points in this study) is larger than the number of brain regions (i.e., 90 ROIs), it is capable of calculating the full inverse covariance matrix Σ^{-1} from the time series data directly. Firstly, we computed the covariance matrix Σ for each pair of brain regions using the following formula:

$$\Sigma_{pq}^m = \frac{\sum_{i=1}^{n_t} (\mathbf{x}_p^m(i) - \bar{\mathbf{x}}_p^m)(\mathbf{x}_q^m(i) - \bar{\mathbf{x}}_q^m)}{n_t - 1}, \quad (5)$$

where Σ_{pq}^m is the covariance between the time series for the p th ROI \mathbf{x}_p^m and the q th ROI \mathbf{x}_q^m from the m th subject, $\bar{\mathbf{x}}_p^m$ and $\bar{\mathbf{x}}_q^m$ denote the mean of the \mathbf{x}_p^m and \mathbf{x}_q^m , respectively. Then, the full inverse covariance matrix Σ^{-1} is directly computed from the covariance matrix Σ .

Sparse network construction

The SICE method, which finds a sparse inverse covariance matrix by imposing a “sparsity” constraint on the maximum likelihood estimation of the inverse matrix, is used to construct the sparse functional brain sub-network.

Suppose the optimal ROIs time-series subset for the p th ROI selected in section Group-Constrained Topology Structure Detection is $E_p = [y_1^p, y_2^p, \dots, y_R^p]$ (n_t time-points $\times R$ regions), where y_r denotes the fMRI time series for the r th selected ROI. The expanded- E_p containing the target brain region (p th ROI) can be expressed as $\tilde{E}_p = [y_0^p, E_p] = [y_0^p, y_1^p, y_2^p, \dots, y_R^p]$, where y_0^p is the time series for the p th ROI.

The SICE method tries to find an estimate for the inverse covariance matrix $\hat{\Theta}$ for \tilde{E}_p by solving the following optimization:

$$\hat{\Theta} = \operatorname{argmin}_{\Theta > 0} (-\log(\det(\Theta)) + \operatorname{tr}(\mathbf{S}\Theta) + \lambda \|\Theta\|_1), \quad (6)$$

where $\det(\bullet)$ and $\operatorname{tr}(\bullet)$ denote the determinant and trace of a matrix, \mathbf{S} is the sample-based covariance matrix for \tilde{E}_p , $\|\bullet\|_1$ is the sum of absolute values of all the entries in a matrix, and λ is the regularization parameter, respectively. Considering a trade-off between the likelihood estimation and the regularization term $\|\Theta\|_1$, the Equation (6) can also be written as follows:

$$\hat{\Theta} = \operatorname{argmin}_{\Theta > 0} (-\log(\det(\Theta)) + \operatorname{tr}(\mathbf{S}\Theta)), \quad \text{subject to } \|\Theta\|_1 \leq \varepsilon, \quad (7)$$

where ε ($\varepsilon > 0$) is reversely related to λ . When ε is large enough (i.e., small λ), the constraint $\|\Theta\|_1 \leq \varepsilon$ has little effect and the SICE is just a usual Maximum likelihood estimation. Conversely, when ε is small enough (i.e., large λ), the SICE can produce a shrunken estimate for Θ , and effectively set certain coefficients

in $\hat{\Theta}$ exactly to zero. For example, if the ij th element in the connection matrix $\hat{\Theta}$ is zero, the region i and j are conditionally non-connection. Conversely, the nonzero elements in $\hat{\Theta}$ indicate the connection between two regions of ROI. Specifically, due to the time series for the target brain region is the first column of \tilde{E}_p . Thus, the first column/row of $\hat{\Theta}$ denotes the connection between the target brain region and its most discriminative ROIs subset. Therefore, the functional connectivity network can be constructed by using the sparse inverse covariance matrix $\hat{\Theta}$.

Feature Extraction

For the p th ROI, the Group-constrained topology structure detection algorithm is performed on all the other ROIs' time series to select the ROIs set E_p that included the most discriminative brain regions. Then, these ROIs time series together with the target ROI time series are used to construct the functional connectivity networks based on the SICE method. Specifically, the constructed functional connectivity networks Θ_m^p , where $m = 1, 2, \dots, M$, with the total number of subjects M , can be regarded as sub-networks of the whole brain from the p th ROI, since all brain regions used in ROIs set E_p are connected with the p th ROI. Additionally, we can take the first column of Θ_m^p as the weight vector, \mathbf{v}_m^p , that quantifies the degree of the influence of other ROIs to the p th ROI. Considering that the dimensions of the connection matrix Θ_m is $P \times P$, with the total number of ROIs P , we put the elements of \mathbf{v}_m^p in the corresponding positions of the p th column of Θ_m according to the ROI indexes.

Repeating the above processes M times, we can obtain a connection matrix Θ for each subject, with the p th column denoting the connections between the p th ROI and the other ROIs. Then, the connection matrix Θ is converted into a feature vector with 8,100 features (90×90 ROIs), where the element represented the connectivity strength between two ROIs for the given subject. Finally, a $N_{subjects} \times N_{edges}$ feature matrix was obtained, where $N_{subjects}$ is the number of subjects, and N_{edges} denotes the number of the features. A detailed procedure of the feature extraction is given in part B of **Figure 1**.

Classification and Validation

Considering the limited sample size, the following nested LOOCV scheme (**Figure 2**) is adopted to evaluate the classification performance of our proposed method. The LOOCV is reported as an effective method to obtain a reliable accuracy estimate for the small sample size classification (Wong, 2015). In the outer LOOCV loop, suppose the whole dataset consists of M subjects ($M = 55$ in this study), one subject is first left out for testing and the remaining $M - 1$ subjects are used for training the classifier. The above procedures will be repeated M times and each time a different subject will be left out for testing the performance of the classifier, which is trained based on the remaining $M - 1$ subjects. In this way, each subject is used as the test subject for one time and M classification results will be obtained. Finally, the average cross-validation classification accuracy is achieved.

In each repeat of the outer loop, we further performed an inner LOOCV loop on the $M - 1$ training data to optimize the hyper-parameter λ which is used for the network construction.

Specifically, the $M - 1$ training subjects will be further separated into $M - 2$ training subjects and one testing subject. Repeating the inner loop procedure $M - 1$ times and each time a different subject will be left out for testing. Finally, the parameter λ with the maximum classification accuracy in the inner loop is transmitted to the outer loop for classifying the AD from NC.

RESULTS

Comparison of Classification Performance

In this work, using the same dataset, the proposed Group-constrained topology structure detection + SICE (Group-constrained SICE) method is compared with other related works including the Partial method, the Group-constrained topology structure detection + Partial (Group-constrained Partial) method, and the SICE method. We further compare the performance of our proposed framework with the recent existing methods for AD classification, including the threshold correlation method (Khazaei et al., 2016) and the minimum spanning tree (MST) high-order method (Guo et al., 2017). In the threshold correlation approach (Khazaei et al., 2016), the brain network is constructed using the Pearson correlation coefficient to calculate the functional connectivity of all pairs of brain regions. Then, the weak connections are removed by maximizing the global cost efficiency (GCE) of the network. In the MST high-order approach (Guo et al., 2017), a high-order network is first constructed based on the sliding window and the Pearson correlation coefficient. Then, the high-order network is pruned by the MST method to construct the MST high-order network. The classification accuracy (ACC), sensitivity (SEN) (Wang et al., 2017), specificity (SPE), and AUC (Li et al., 2018) are used to measure the classification performance of different classification methods. The classification results are summarized in **Table 1**. **Figure 3** presents the ROC (Li et al., 2017) curves of different classification methods.

As shown in **Table 1**, our proposed classification method yielded a classification accuracy of 81.82% with a sensitivity of 80.0% and a specificity of 83.33%, while the best accuracy was only reached to 74.55% by the direct SICE method. The cross-validation estimation of the generalization performance showed 0.9667 of AUC with the proposed method, indicating the classification power of our proposed scheme.

The Optimal Lambda

In order to hunt for the optimal parameter λ , an inner LOOCV framework is exploited to determine λ in each leave-one-out fold. Therefore, different λ will be selected from the λ pool each time. The optimal λ in each fold is given in **Figure 4**. We can see that $\lambda = 0.55$ and $\lambda = 0.57$ are selected the 22 frequency times and 28 frequency times from the total number of 55 subjects, respectively. These results indicates that the functional connectivity networks based on $\lambda = 0.55$ and $\lambda = 0.57$ constructed have better discriminability between AD and NC.

Computational Complexity and Convergence

The computational complexity of the proposed method is evaluated by the computation time. All timings are carried out on

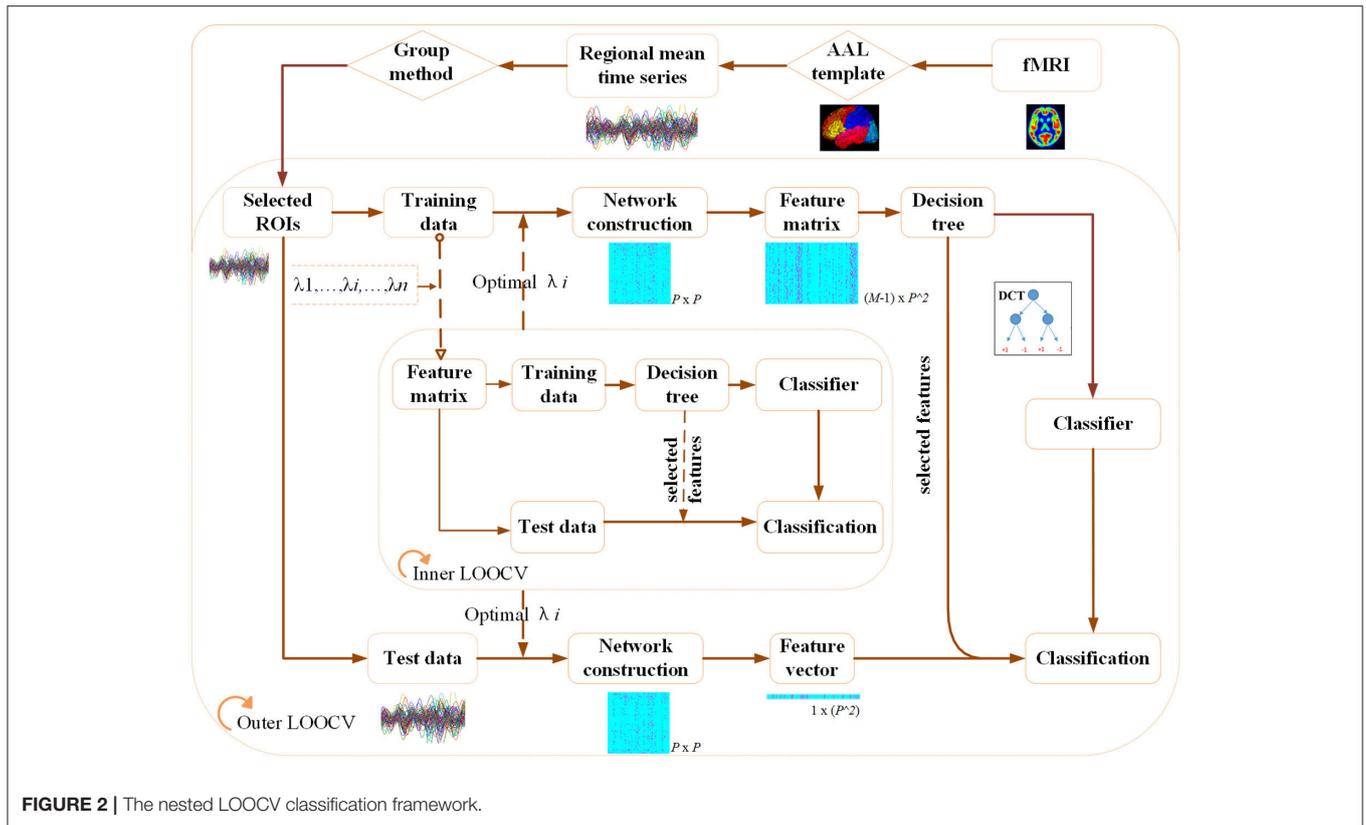


FIGURE 2 | The nested LOOCV classification framework.

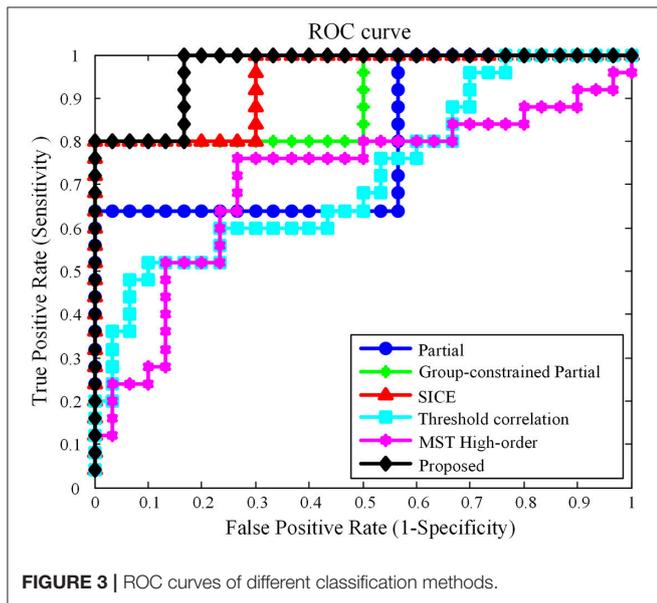


FIGURE 3 | ROC curves of different classification methods.

TABLE 1 | The classification performance of different classification methods.

Method	ACC (%)	SEN (%)	SPE (%)	AUC
Partial	58.18	64.00	53.33	0.7960
Group-constrained Partial	63.64	80.00	50.00	0.9000
SICE	74.55	80.00	70.00	0.9400
Threshold correlation	69.09	60.00	76.67	0.7200
MST High-order	74.55	76.00	73.33	0.7067
Group-constrained SICE	81.82	80.00	83.33	0.9667

Where bold values indicate the best results.

an Intel Core i7-4790 3.60GHz processor. Each subsection of the methods has been run 10 times and the mean of the computation time is shown in Table 2.

It is obvious that the computation time of Partial method is significantly shorter than other methods. However, with a classification accuracy of 58.18%, the Partial correlation is not

considered to be an effective method for AD classification. The computation times of the Group-constrained Partial method and Group-constrained SICE method (the proposed method) are similar, where the main computation time comes from the Group-constrained topology structure detection algorithm. Although the computation load of the proposed method is heavier than the SICE method, there is no significant of magnitude difference between them, and the computational issue becomes less critical when the performance of PC develops rapidly. Furthermore, applying the SICE at an individual level will inevitably cause the inter-subject variability, thus reducing the generalization performance. Therefore, it is crucial to adopt the Group-constrained method, which encourages an identical network topology across subjects, to overcome the limitation of SCIE. Interestingly, the computation time of the SICE in the

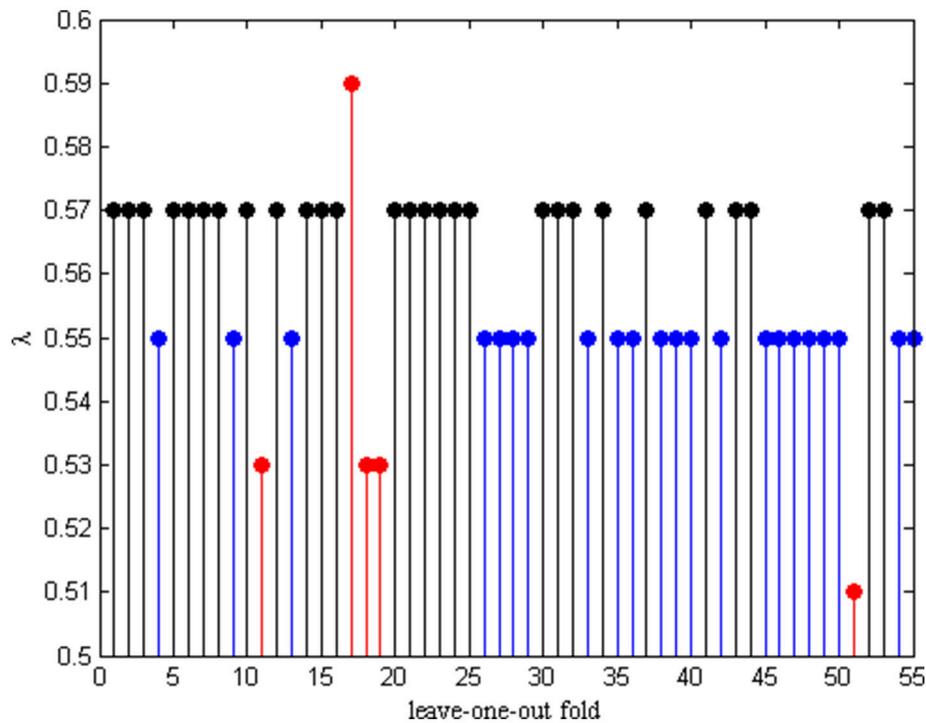


FIGURE 4 | The optimal λ in each LOO-fold.

TABLE 2 | The computation times (seconds) of all compared frameworks.

Method	Group	Partial	SICE	Other	Total
Partial	-	2.307	-	-	2.307
Group-constrained Partial	1,115.674	2.274	-	-	1,117.948
SICE	-	-	244.738	-	244.738
Threshold correlation	-	-	-	1,156.468	1,156.468
MST High-order	-	-	-	7,415.385	7,415.385
Group-constrained SICE	1,115.674	-	16.907	-	1,132.581

Where each column represents the computation time of the subsection of the method.

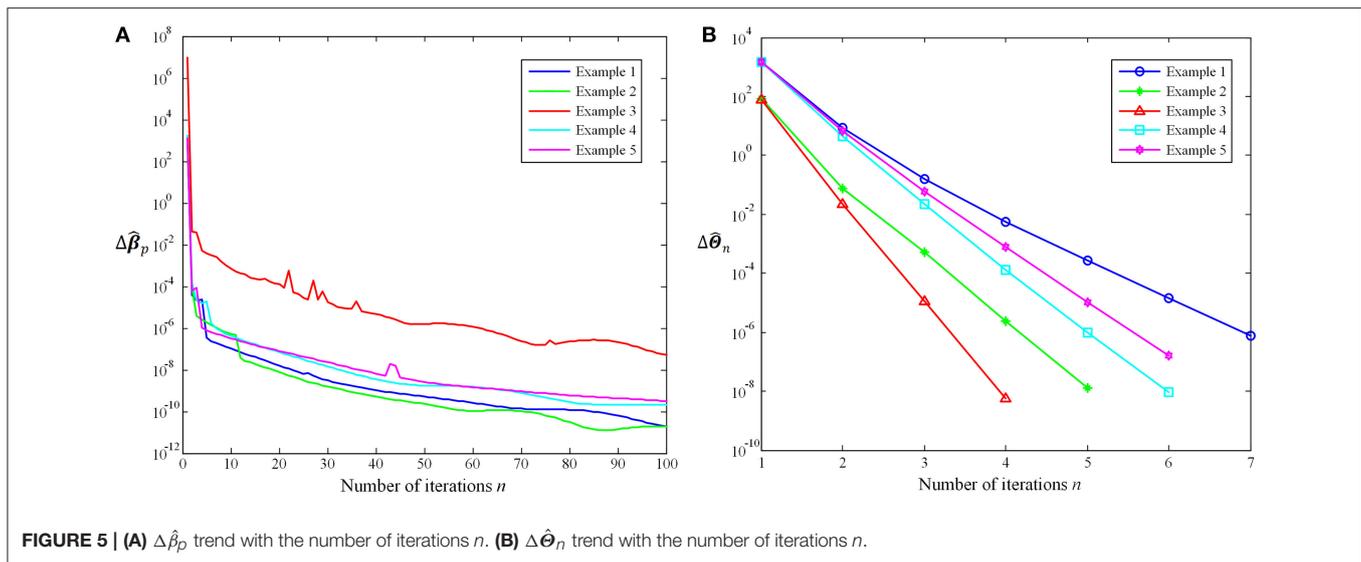
proposed method is shorter than that of using SICE alone. This phenomenon can be interpreted as that the Group-constrained topology structure detection algorithm is adopted to select the most relevant brain regions, reducing the number of brain regions calculated in the SICE. The computation time of the threshold correlation method (Khazaee et al., 2016) is slightly longer than that of our proposed method. Due to the large scale of the high-order network, the computation complexity of the MST high-order method (Guo et al., 2017) is significantly higher than other methods.

In this work, we further evaluate the proposed method convergence based on the trend of coefficients variation with the number of iterations. On the one hand, the iteration of the Group-constrained topology structure detection algorithm will be repeated until $n = 100$, where n is the number of iterations. $\Delta\hat{\beta}_p = \text{Frobenius norm of } (\hat{\beta}_p)_n - (\hat{\beta}_p)_{n-1}$ describes how

much the coefficient matrix changes in the n th iteration. In the 100th iterations, $\max(\Delta\hat{\beta}_p) = 1.17 \times 10^{-6}$ and $\text{mean}(\Delta\hat{\beta}_p) = 5.25 \times 10^{-9}$. Figure 5A shows the $\Delta\hat{\beta}_p$ trend with the number of iterations n using five examples. On the other hand, the iteration of the SICE will stop when $\Delta\hat{\Theta}_n < 10^{-6}$, where $\Delta\hat{\Theta}_n$ is the sum of absolute values of $\hat{\Theta}_n - \hat{\Theta}_{n-1}$. The average number of iterations is 5.57 and the maximum number is 13. Figure 5B shows the $\Delta\hat{\Theta}_n$ trend with the number of iterations n using five examples. It is obvious that as the number of iterations n increases, the variation of coefficients ($\Delta\hat{\beta}_p$ and $\Delta\hat{\Theta}_n$) decreases rapidly. The experimental results suggest that the proposed method is convergent and has a high rate of convergence.

Brain Regions Involved in Classification

Since the nested LOOCV is used to evaluate the performance of the proposed framework, different optimal feature (connection between two ROIs) subsets are selected for AD classification. Thus, those connections with the highest selected frequency times in the nested LOOCV are regarded as the most discriminative features for AD classification. Figures 6A,B show the selected connections among all LOO folds. A detailed feature index and its frequency times selected are summarized in Table 3. For example, the interconnection between Temporal_Pole_Sup_R and Temporal_Pole_Mid_R is selected 55 frequency times among all LOO folds, and showed very strong discriminative strength in AD diagnosis. In addition, the following four interconnections including



ParaHippocampal_L-Temporal_Pole_Sup_L, Cingulum_Post_L-Angular_L, Supp_Motor_Area_L-Frontal_Med_Orb_R, and Frontal_Sup_R-Cingulum_Ant_R are all selected with over 50 frequency times. These connections have much higher selected frequency times than others, thus can be regarded as effective biomarkers for identifying AD from healthy elderly.

It should be noted that the brain regions involved in the significant abnormal connectivity pathways (with over 50 selected frequency times) are located mainly within the default mode network (DMN), including Temporal_Pole_Sup_R, Temporal_Pole_Sup_L, Temporal_Pole_Mid_R, ParaHippocampal_L, Cingulum_Post_L, Angular_L, Supp_Motor_Area_L, Frontal_Med_Orb_R, Frontal_Sup_R, Cingulum_Ant_R. These brain regions, listed in **Table 4** and displayed in **Figure 6C**, are reported as highly associated with AD pathology (Rose et al., 2006; Matsuda, 2013; Salvatore et al., 2015; Scheff et al., 2015; Xu et al., 2016; Loewenstein et al., 2018).

DISCUSSION

Classification Performance

This paper proposed a new classification framework based on fMRI time series for diagnosing AD patients. The combination of the Group-constrained topology structure detection algorithm with the SICE, where a nested LOOCV method is employed to optimize the regularization parameter, are designed to construct the efficient functional brain sub-networks. Then, an optimal DCT classifier is trained for classifying AD from NC based on the optimal brain sub-networks. Experimental results demonstrate the effectiveness of the proposed method.

Specifically, in contrast to the other methods, experimental results show that the proposed classification method has at least 7.27% improvement of the diagnosis accuracy. The classification result indicates that the sparse-based method is more appropriate for brain network construction than the

traditional fully-connected correlation-based networks, which may contain a large number of spurious or insignificant connections among ROIs. It was also found that both the SICE and Partial method aided by the Group-constrained topology structure detection method can improve the classification performance. This may be due to the group-constrained topology structure detection algorithm encourages an identical network topology across subjects, minimizing the inter-subject variability problem which degrades generalization performance of trained classifiers.

The Most Discriminative Brain Regions and Connections

The top 5 brain connections listed in **Table 3** have much higher selected frequency times than others, which may serve as the more promising connectivity-based biomarker for AD diagnosis. These results are totally consistent with previous findings, and added new findings to the disconnection hypothesis of the AD (Delbeuck et al., 2003; Lacalle-Auriales et al., 2016). The brain regions identified in the top 5 connections are frequently reported as highly associated with the AD pathology. For example, Scheff et al. (2015) reported that the AD patients showed a significant decline in synaptic numbers in the Cingulum_Post (posterior cingulate gyrus) compared to healthy elderly. Furthermore, these brain regions mainly belong to DMN (Raichle et al., 2001), which is one of the earliest pathological sites of the AD (Greicius et al., 2004). The current findings are in line with the results reported in the related literature that both the impairment and compensation coexist in DMN of AD (Sorg et al., 2007; Liang et al., 2014). Given the disconnection within DMN can also be detected in the mild cognitive impairment (MCI) stage of the AD (Qi et al., 2010; Wang et al., 2012), it is expected that the proposed method may also be applied to the diagnosis of MCI patients. This will be discussed in our future studies.

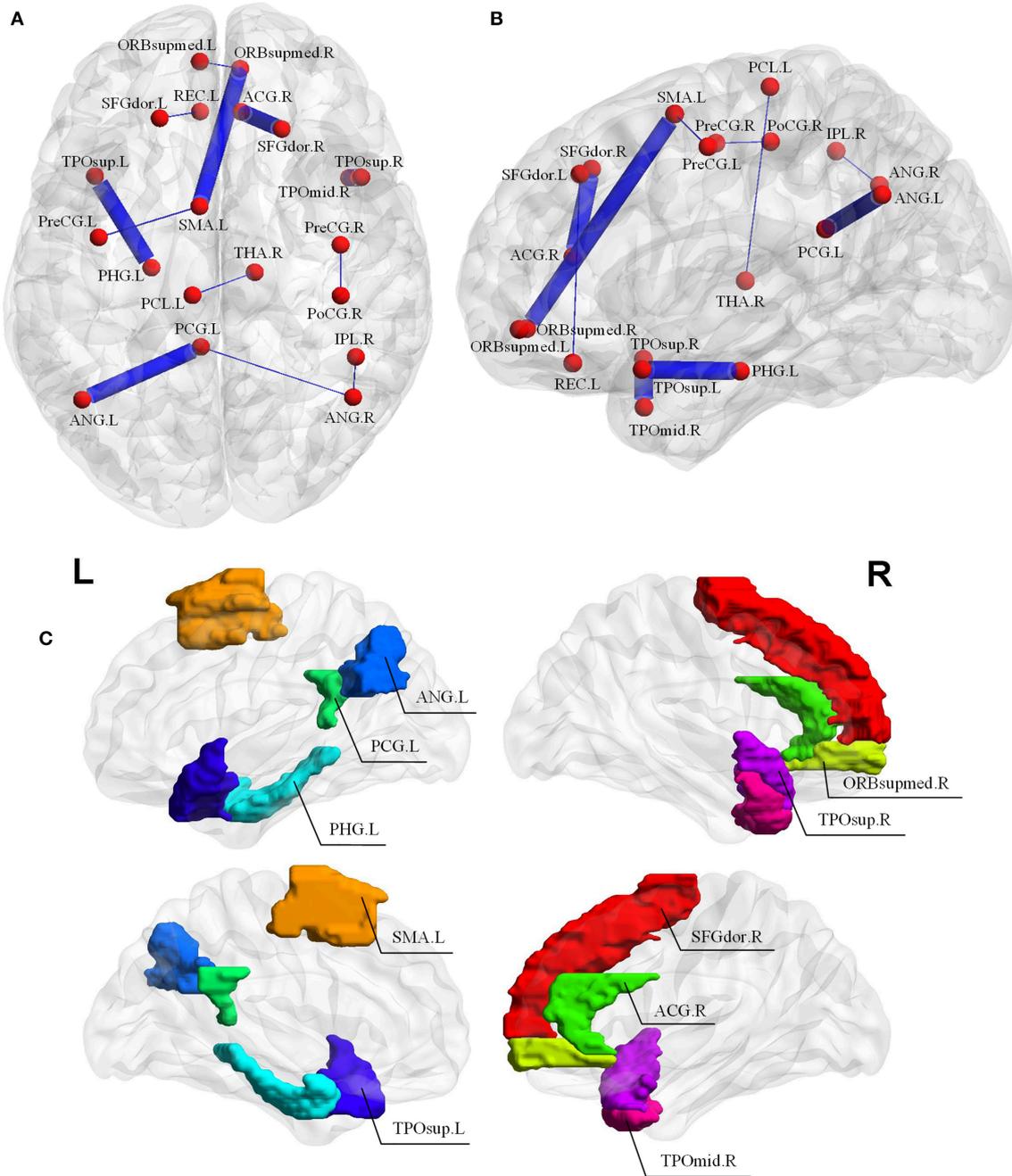


FIGURE 6 | (A,B) Selected connections in the LOOCV folds. The width of edges connecting two ROIs corresponds to the degree of discrimination. **(C)** The discriminative brain regions selected by our proposed method for AD classification. The corresponding ROI names of the abbreviations are as follows: TPOsup.R, Temporal_Pole_Sup_R; TPOmid.R, Temporal_Pole_Mid_R; PHG.L, ParaHippocampal_L; TPOsup.L, Temporal_Pole_Sup_L; CG.L, Cingulum_Post_L; ANG.L, Angular_L; SMA.L, Supp_Motor_Area_L; ORBsupmed.R, Frontal_Med_Orb_R; SFGdor.R, Frontal_Sup_R; ACG.R, Cingulum_Ant_R; PreCG.L, Precentral_L; PreCG.R, Precentral_R; PoCG.R, Postcentral_R; REC.L, Rectus_L; ORBsupmed.L, Frontal_Mid_Orb_L; IPL.R, Parietal_Inf_R; ANG.R, Angular_R; THA.R, Thalamus_R; SFGdor.L, Frontal_Sup_L; PCL.L, Paracentral_Lobule_L.

Methodological Limitations

There are still two limitations in this study. One limitation is that we set the tuning parameter λ of the SICE for identifying for different subjects, which may affect the classification

performance, since the optimal parameter may vary across subjects due to individual differences. To overcome this issue, one possible solution is to optimize the parameter λ for each subject using the BIC method. In this way, we can construct

TABLE 3 | Selected connections by the proposed classification framework.

No.	Functional connectivity	No. of frequency
1	Temporal_Pole_Sup_R---Temporal_Pole_Mid_R	55
2	ParaHippocampal_L----Temporal_Pole_Sup_L	53
3	Cingulum_Post_L---Angular_L	51
4	Supp_Motor_Area_L---Frontal_Med_Orb_R	51
5	Frontal_Sup_R---Cingulum_Ant_R	51
6	Precentral_L---Supp_Motor_Area_L	2
7	Precentral_R---Postcentral_R	1
8	Frontal_Sup_L---Rectus_L	1
9	Frontal_Med_Orb_L---Frontal_Med_Orb_R	1
10	Parietal_Inf_R---Angular_R	1
11	Angular_R---Cingulum_Post_L	1
12	Thalamus_R ----Paracentral_Lobule_L	1

TABLE 4 | The discriminative brain regions.

No.	ROI index	ROI abbr.	ROI name	References
1	84	TPOsup.R	Temporal_Pole_Sup_R	Salvatore et al., 2015
2	83	TPOsup.L	Temporal_Pole_Sup_L	Salvatore et al., 2015
3	88	TPOmid.R	Temporal_Pole_Mid_R	Salvatore et al., 2015
4	39	PHG.L	ParaHippocampal_L	Matsuda, 2013
5	35	PCG.L	Cingulum_Post_L	Scheff et al., 2015
6	65	ANG.L	Angular_L	Xu et al., 2016
7	19	SMA.L	Supp_Motor_Area_L	Rose et al., 2006
8	26	ORBsupmed.R	Frontal_Med_Orb_R	Loewenstein et al., 2018
9	4	SFGdor.R	Frontal_Sup_R	Salvatore et al., 2015
10	32	ACG.R	Cingulum_Ant_R	Salvatore et al., 2015

optimal connectivity networks for each subject and this will be investigated in our future works. Another limitation lies in the brain atlas used in the MRI data analysis. Given the participants in this study are from Chinese populations as well as the significant morphological difference between Chinese and Caucasian population (Tang et al., 2010), the statistical Chinese brain atlas (Liang et al., 2015) rather than the Caucasian brain atlas (as implemented in SPM8) should be used during the image segmentation and registration. This may be helpful for extracting the more exact MRI features of the two groups of participants, which thus may improve the diagnostic accuracy of AD patients.

REFERENCES

- Alzheimer's Association (2015). 2015 Alzheimer's disease facts and figures. *Alzheimers Dement.* 11, 332–384. doi: 10.1016/j.jalz.2015.02.003
- Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *Neuroimage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* 3, 186–191. doi: 10.1016/j.jalz.2007.04.381

CONCLUSION

In this paper, we have proposed a novel sub-network based classification framework to construct brain functional sub-connectivity and explore its diagnostic power in distinguishing AD patients from NC. Different from the method based on the whole-brain level, we constructed brain sub-networks with the most discriminative brain regions. Experimental results have verified the validity of the proposed classification framework.

ETHICAL APPROVAL

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

AUTHOR CONTRIBUTIONS

YL designed the framework, explained the results and wrote the paper. JL supplemented the experiment and revised the paper. JH made the program and wrote the paper. ZL made the program and wrote the paper. PL designed the study, explained the results, and wrote the paper.

FUNDING

This study was supported by the grant from the National Natural Science Foundation of China (61671042, 61403016 and 61473196), as well as Beijing Talents foundation (2016000021223TD07), Open Fund Project of Fujian Provincial Key Laboratory of Information Processing and Intelligent Control in Minjiang University (No. MJUKF201702), the Specialized Research Fund for the Doctoral Program of Higher Education (20131102120008), Project Sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, and the Fundamental Research Funds for the Central Universities.

- Chao-Gan, Y., Yu-Feng, Z. (2010). DPARSF: a MATLAB toolbox for "Pipeline" data analysis of resting-state fMRI. *Front. Syst. Neurosci.* 4:13. doi: 10.3389/fnsys.2010.00013
- Delbeuck, X., Van der Linden, M., and Collette, F. (2003). Alzheimer's disease as a disconnection syndrome? *Neuropsychol. Rev.* 13, 79–92. doi: 10.1023/A:1023832305702
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc.*

- Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102
- Frances, A., Pincus, H. A., and First, M. (1994). *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV*. Washington, DC: American Psychiatric Association.
- Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4637–4642. doi: 10.1073/pnas.0308627101
- Guo, H., Liu, L., Chen, J., Xu, Y., and Jie, X. (2017). Alzheimer classification using a minimum spanning tree of high-order functional network on fMRI dataset. *Front. Neurosci.* 11:639. doi: 10.3389/fnins.2017.00639
- Huang, S., Li, J., Sun, L., Ye, J. P., Fleisher, A., Wu, T., et al. (2010). Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *Neuroimage* 50, 935–949. doi: 10.1016/j.neuroimage.2009.12.120
- Jie, B., Zhang, D., Gao, W., Wang, Q., Wee, C. Y., and Shen, D. G. (2014). Integration of network topological and connectivity properties for neuroimaging classification. *IEEE Trans. Biomed. Eng.* 61, 576–589. doi: 10.1109/TBME.2013.2284195
- Khazaei, A., Ebrahimzadeh, A., and Babajani-Feremi, A. (2016). Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease. *Brain Imaging Behav.* 10, 799–817. doi: 10.1007/s11682-015-9448-7
- Lacalle-Aurioles, M., Navas-Sánchez, F. J., Alemán-Gómez, Y., Olazáran, J., Guzman-De-Villoria, J. A., Cruz-Orduna, I., et al. (2016). The disconnection hypothesis in Alzheimer's disease studied through multimodal magnetic resonance imaging: structural, perfusion, and diffusion tensor imaging. *J. Alzheimers Dis.* 50, 1051–1064. doi: 10.3233/JAD-150288
- Li, Y., Cui, W., Luo, M., Li, K., and Wang, L. N. (2018). Epileptic seizure detection based on time-frequency images of EEG signals using gaussian mixture model and gray level co-occurrence matrix features. *Int. J. Neural Syst.* 28:1850003. doi: 10.1142/S012906571850003X
- Li, Y., Wang, X. D., Luo, M. L., Li, K., Yang, X., and Guo, Q. (2017). Epileptic seizure classification of eegs using time-frequency analysis based multiscale radial basis functions. *IEEE J. Biomed. Health Inform.* 22, 386–397. doi: 10.1109/JBHI.2017.2654479
- Li, Y., Wee, C. Y., Jie, B., Peng, Z., and Shen, D. G. (2014). Sparse multivariate autoregressive modeling for mild cognitive impairment classification. *Neuroinformatics* 12, 455–469. doi: 10.1007/s12021-014-9221-x
- Liang, P., Shi, L., Chen, N., Luo, Y. S., Wang, X., Liu, K., et al. (2015). Construction of brain atlases based on a multi-center MRI dataset of 2020 Chinese adults. *Sci. Rep.* 5:18216. doi: 10.1038/srep18216
- Liang, P. P., Wang, Z. Q., Qian, T. Y., and Li, K. C. (2014). Acupuncture stimulation of taichong (Liv3) and hegu (LI4) modulates the default mode network activity in Alzheimer's disease. *Am. J. Alzheimers Dis. Dement.* 29, 739–748. doi: 10.1177/1533317514536600
- Loewenstein, D. A., Curiel, R. E., Rosselli, M., Penate, A., Greig-Custo, M. T., Bauer, R. M., et al. (2018). Semantic intrusions and failure to recover from semantic interference in mild cognitive impairment: relationship to amyloid and cortical thickness. *Curr. Alzheimer Res.* 15, 848–855. doi: 10.2174/156720510566180427122746
- Matsuda, H. (2013). Voxel-based morphometry of brain MRI in normal aging and Alzheimer's disease. *Aging Dis.* 4, 29–37.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 34, 939–944. doi: 10.1212/WNL.34.7.939
- Meszlényi, R. J., Buza, K., and Vidnyánszky, Z. (2017). Resting state fMRI functional connectivity-based classification using a convolutional neural network architecture. *Front. Neuroinform.* 11:61. doi: 10.3389/fninf.2017.00061
- Mitra, R., and Zhang, C. H. (2016). The benefit of group sparsity in group inference with de-biased scaled group Lasso. *Electron. J. Stat.* 10, 1829–1873. doi: 10.1214/16-EJS1120
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414. doi: 10.1212/WNL.43.11.2412-a
- Qi, Z., Wu, X., Wang, Z. Q., Zhang, N., Dong, H. Q., Yao, L., et al. (2010). Impairment and compensation coexist in amnesic MCI default mode network. *Neuroimage* 50, 48–55. doi: 10.1016/j.neuroimage.2009.12.025
- Qureshi, M. N. I., Oh, J., Cho, D., Jo, H. J., and Lee, B. (2017). Multimodal discrimination of schizophrenia using hybrid weighted feature concatenation of brain functional connectivity and anatomical features with an extreme learning machine. *Front. Neuroinform.* 11:59. doi: 10.3389/fninf.2017.00059
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676
- Rosa, M. J., Portugal, L., Hahn, T., Fallgatter, A. J., Garrido, M. I., Shawe-Taylor, J., et al. (2015). Sparse network-based models for patient classification using fMRI. *Neuroimage* 105, 493–506. doi: 10.1016/j.neuroimage.2014.11.021
- Rose, S. E., McMahon, K. L., Janke, A. L., O'Dowd, B., De, Z. G., Strudwick, M. W., et al. (2006). Diffusion indices on magnetic resonance imaging and neuropsychological performance in amnesic mild cognitive impairment. *J. Neurol. Neurosurg. Psychiatry* 77, 1122–1128. doi: 10.1136/jnnp.2005.074336
- Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., and Castiglioni, I. (2015). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307
- Scheff, S. W., Price, D. A., Ansari, M. A., Roberts, K. N., Schmitt, F. A., Ikonovic, M. D., et al. (2015). Synaptic change in the posterior cingulate gyrus in the progression of Alzheimer's disease. *J. Alzheimers Dis.* 43, 1073–1090. doi: 10.3233/JAD-141518
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 15–18. doi: 10.1214/aos/1176344136
- Sorg, C., Riedl, V., Mühlau, M., Calhoun, V. D., Eichele, T., Läer, L., et al. (2007). Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18760–18765. doi: 10.1073/pnas.0708803104
- Stam, C. J., Jones, B. F., Nolte, G., Breakspear, M., and Scheltens, P. (2007). Small-world networks and functional connectivity in Alzheimer's disease. *Cerebral Cortex* 17, 92–99. doi: 10.1093/cercor/bhj127
- Supekar, K., Menon, V., Rubin, D., Musen, M., and Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Comput. Biol.* 4:e1000100. doi: 10.1371/journal.pcbi.1000100
- Tang, Y., Hojatkashani, C., Dinov, I. D., Sun, B., Fan, L. Z., Lin, X. T., et al. (2010). The construction of a Chinese MRI brain atlas: a morphometric comparison study between Chinese and caucasian cohorts. *Neuroimage* 51, 33–41. doi: 10.1016/j.neuroimage.2010.01.111
- Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., and Knight, K. (2005). Sparsity and smoothness via the fused lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67, 91–108. doi: 10.1111/j.1467-9868.2005.00490.x
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Wang, L., Xue, W., Li, Y., Luo, M., Huang, J., Cui, W., et al. (2017). Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis. *Entropy* 19:222. doi: 10.3390/e19060222
- Wang, Z., Liang, P. P., Jia, X. Q., Jin, G. W., Song, H. Q., Han, Y., et al. (2012). The baseline and longitudinal changes of PCC connectivity in mild cognitive impairment: a combined structure and resting-state fMRI study. *PLoS ONE* 7:e36838. doi: 10.1371/journal.pone.0036838
- Wee, C. Y., Yang, S., Yap, P. T., Shen, D. G., and Alzheimer's Disease Neuroimaging Initiative (2016). Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging Behav.* 10, 342–356. doi: 10.1007/s11682-015-9408-2
- Wee, C. Y., Yap, P. T., Denny, K., Brownadyke, J. N., Potter, G. G., Welsh-Bohmer, K. A., et al. (2012). Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. *PLoS ONE* 7:e37828. doi: 10.1371/journal.pone.0037828
- Wee, C. Y., Yap, P. T., Zhang, D. Q., Wang, L. H., and Shen, D. G. (2014). Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. *Brain Struct. Funct.* 219, 641–656. doi: 10.1007/s00429-013-0524-8

- Wong, T. T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognit.* 48, 2839–2846. doi: 10.1016/j.patcog.2015.03.009
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., and Ma, Y. (2009). Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 210–227. doi: 10.1109/TPAMI.2008.79
- Xu, L., Wu, X., Li, R., Chen, K. W., Long, Z. Y., Zhang, J. C., et al. (2016). Prediction of progressive mild cognitive impairment by multi-modal neuroimaging biomarkers. *J. Alzheimers Dis.* 51, 1045–1056. doi: 10.3233/JAD-151010
- Zanin, M., Sousa, P., Papo, D., Bajo, R., Garcia-Prieto, J., del Pozo, F., et al. (2012). Optimizing functional network representation of multivariate time series. *Sci. Rep.* 2:630. doi: 10.1038/srep00630
- Zhang, J., Tu, Y., Li, Q., Caselli, R. J., Thompson, P. M., Ye, J., et al. (2018). Multi-task sparse screening for predicting future clinical scores using longitudinal cortical thickness measures. *Proc. IEEE Int. Symp. Biomed. Imaging* 2018, 1406–1410. doi: 10.1109/ISBI.2018.8363835
- Zhang, J., Zhou, L. P., Wang, L., and Li, W. Q. (2015). Functional brain network classification with compact representation of SICE matrices. *IEEE Trans. Biomed. Eng.* 62, 1623–1634. doi: 10.1109/TBME.2015.2399495

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

Copyright © 2018 Li, Liu, Huang, Li and Liang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.