



Feature Selection and Combination of Information in the Functional Brain Connectome for Discrimination of Mild Cognitive Impairment and Analyses of Altered Brain Patterns

Xiaowen Xu¹, Weikai Li², Jian Mei³, Mengling Tao¹, Xiangbin Wang¹, Qianhua Zhao^{4,5}, Xiaoniu Liang^{4,5}, Wanqing Wu^{4,5}, Ding Ding^{4,5*} and Peijun Wang^{1*}

¹ Department of Medical Imaging, Tongji Hospital, Tongji University School of Medicine, Shanghai, China, ² College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China, ³ Physical Education College, Soochow University, Suzhou, China, ⁴ Institute of Neurology, Huashan Hospital, Fudan University, Shanghai, China, ⁵ National Clinical Research Center for Aging and Medicine, Huashan Hospital, Fudan University, Shanghai, China

OPEN ACCESS

Edited by:

Mohamad Habes,
University of Pennsylvania,
United States

Reviewed by:

Han Zhang,
The University of North Carolina
at Chapel Hill, United States
Sandhitsu Das,
University of Pennsylvania,
United States

*Correspondence:

Ding Ding
dingding@huashan.org.cn
Peijun Wang
tongjipjwang@vip.sina.com

Received: 31 May 2019

Accepted: 28 January 2020

Published: 19 February 2020

Citation:

Xu X, Li W, Mei J, Tao M, Wang X, Zhao Q, Liang X, Wu W, Ding D and Wang P (2020) Feature Selection and Combination of Information in the Functional Brain Connectome for Discrimination of Mild Cognitive Impairment and Analyses of Altered Brain Patterns.

Front. Aging Neurosci. 12:28.
doi: 10.3389/fnagi.2020.00028

Mild cognitive impairment (MCI) is often considered a critical time window for predicting early conversion to Alzheimer's disease (AD). Brain functional connectome data (i.e., functional connections, global and nodal graph metrics) based on resting-state functional magnetic resonance imaging (rs-fMRI) provides numerous information about brain networks and has been used to discriminate normal controls (NCs) from subjects with MCI. In this paper, Student's *t*-tests and group-least absolute shrinkage and selection operator (group-LASSO) were used to extract functional connections with significant differences and the most discriminative network nodes, respectively. Based on group-LASSO, the middle temporal, inferior temporal, lingual, posterior cingulate, and middle frontal gyri were the most predominant brain regions for nodal observation in MCI patients. Nodal graph metrics (within-module degree, participation coefficient, and degree centrality) showed the maximum discriminative ability. To effectively combine the multipattern information, we employed the multiple kernel learning support vector machine (MKL-SVM). Combined with functional connectome information, the MKL-SVM achieved a good classification performance (area under the receiving operating characteristic curve = 0.9728). Additionally, the altered brain connectome pattern revealed that functional connectivity was generally decreased in the whole-brain network, whereas graph theory topological attributes of some special nodes in the brain network were increased in MCI patients. Our findings demonstrate that optimal feature selection and combination of all connectome features (i.e., functional connections, global and nodal graph metrics) can achieve good performance in discriminating NCs from MCI subjects. Thus, the combination of functional connections and global and nodal graph metrics of brain networks can predict the occurrence of MCI and contribute to the early clinical diagnosis of AD.

Keywords: resting-state functional magnetic resonance imaging, functional connectivity, graph theory, multiple kernel learning, mild cognitive impairment

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by loss of memory and cognitive decline (Blennow et al., 2006). With the aging of the global population, there will be an estimated 115 million AD patients in the world by 2050, with an average of 1 new AD patient every 33 s (Ijaopo, 2017). Mild cognitive impairment (MCI) is an intermediate stage that precedes early AD. Evidence indicates that about 15% of MCI patients progress to AD per year (Petersen et al., 1999; Grundman et al., 2004). Therefore, MCI is regarded as the critical time window for early prediction of conversion to AD (Manly et al., 2008).

Components of the brain functional connectome, including functional connections and graph theory topological metrics, have become important imaging markers for exploring brain networks and predicting the classification of neurodegenerative diseases (Biswal et al., 2010; Wang et al., 2013; Filippi et al., 2018). The functional connectome systematically depicts global graph metrics (i.e., small world, modularity, global efficiency), nodal graph metrics (i.e., degree, participant coefficient, shortest path length), and functional connections of the network. It provides a novel approach for revealing altered brain network patterns (delEtoile and Adeli, 2017; Khazaei et al., 2017; Filippi et al., 2018). Given the large numbers of network features in the brain connectome, the Student's *t*-test (Qiao et al., 2016; Li W. et al., 2019) and sparse methods such as least absolute shrinkage and selection operator (LASSO) have been applied to select the critical features of brain networks (Wee et al., 2014; Li Y. et al., 2019). Nodal graph metrics naturally have a group topology (i.e., a node corresponds to a group of node-graph theoretical attributes). Group-LASSO is a regression-analysis method for group-feature selection and regularization that can be adopted to select nodal graph metrics (Liu et al., 2019) and maintain significant discrimination of nodal features.

In recent years, machine learning approaches with data-driven algorithms have been used to combine and classify brain features. Some classifiers such as support vector machines (SVMs) (Prasad et al., 2015; Khazaei et al., 2016), Naïve Bayes (Zhuo et al., 2018) and deep neural networks (Themistocleous et al., 2018) are applied to discriminate normal controls from subjects with MCI. However, most of these methods focus on a single modality of imaging, the functional connectome, or graph theory attributes separately, resulting in relatively poor classification performance (Suk et al., 2014). Therefore, the multimodal brain network (i.e., functional connections and graph theory topological metrics) should be used to provide a comprehensive and insightful understanding of the brain network in patients with MCI. Combined with information from different attributes, multiple kernel learning SVM (MKL-SVM) (Niu et al., 2017) can partially alleviate the high-dimensional curve of multiple features and measure the contributions of different features to the classification. These proposed methods could help select critical features and discriminate normal controls from subjects with diseases.

The main purposes of the present study were to select discriminative features of the brain connectome (i.e., functional

connections, global graph metrics, and nodal graph metrics) and develop a classification of MCI based on different attributes of the brain network. Altered patterns of discriminative features were further analyzed using the proposed methods. By combining the group-LASSO model and MKL-SVM, we (i) identified the most discriminative nodal features of the brain connectome and predominant brain regions in MCI patients, (ii) achieved accurate and automatic classification of MCI patients and normal controls (NCs), and (iii) analyzed the changed patterns in the brain network.

MATERIALS AND METHODS

Participants

Participants with MCI and NCs were recruited to establish a registry at Huashan Hospital. Each participant underwent a comprehensive evaluation, including clinical interview, neuropsychological assessment, laboratory tests, and multimodal magnetic resonance imaging (MRI) examinations of the brain. MCI was defined according to the following criteria (Petersen, 2004): (i) cognitive concern/complaint by the subject, nurse, or physician, with a Clinical Dementia Rating (CDR) = 0.5; (ii) objective impairment in ≥ 1 cognitive domain based on 1.5 standard deviations (SDs) below the mean using the norms obtained in the pilot study; (iii) basic normal functional activities (determined by CDR and daily living activity assessment); (iv) absence of dementia according to the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (Rabe-Jablonska and Bienkiewicz, 1994). The inclusion criteria of NCs were: (i) no neurology-related or cerebral vascular diseases (e.g., Parkinson's disease, intracranial aneurysms, or cerebral tumors); (ii) no severe mental retardation or schizophrenia; (iii) no severe problems in speaking, vision, or hearing; (iv) able to actively participate in the neuropsychological assessment. In the present study, 105 participants (41 MCI patients and 64 NCs) were selected. Two patients with MCI and four NCs were excluded due to incomplete data in resting state-functional MRI (rs-fMRI) and severe head motion at some time points. Finally, data from 99 individuals (39 MCI patients and 60 NCs) were included in the subsequent statistical analyses. The clinical and demographic data of these 99 participants were summarized. The study protocol was approved by the Ethics Committee of Huashan Hospital of Fudan University (Shanghai, China). Written informed consent was obtained from each participant (or his/her legal representative). In addition, we adopted the Alzheimer's Disease Neuroimaging Initiative (ADNI)¹ dataset as an independent test dataset to verify the performance of the pre-trained model.

Data Acquisition

Rs-fMRI and structural MR images were acquired on a 3T MR scanner (Magnetom® Verio; Siemens, Munich, Germany) using a 32-channel head coil. Before imaging, all participants were

¹<http://adni.loni.usc.edu/>

instructed to keep their eyes closed (but not to fall asleep), think of nothing, and move as little as possible during data acquisition. Three-dimensional (3D) T1-weighted sagittal images were acquired using magnetization-prepared rapid gradient echo with the following parameters: repetition time (TR) = 2530 ms, echo time (TE) = 2.34 ms, flip angle = 7°, inversion time (TI) = 1100 ms, matrix = 256 × 256, slice number = 192, thickness = 1.0 mm, and voxel size = 1 × 1 × 1 mm³. The scan lasted 6 min 03 s. The parameters of the rs-fMRI protocol were as follows: axial acquisition, TR = 2000 ms, TE = 30 ms, flip angle = 90°, slice thickness = 3.8 mm, slice number = 31, field of view = 220 × 220 mm², matrix size = 64 × 64, and voxel size = 3.4 × 3.4 × 3.8 mm³. Each scan collected 240 volumes with a scan time of 8 min 06 s. The ADNI dataset was acquired on the 3T Philips with the following scan parameters: TR = 3000 ms, TE = 30 mm, flip angle = 80°, slice thickness = 3.3 mm, slice number = 48, matrix size = 64 × 64, and measurements = 140.

Image Preprocessing

Preprocessing procedures were carried out using Data Processing Assistant for Resting-State fMRI (DPARSF)² and Statistical Parametric Mapping (SPM12)³. The first 10 time points were not used to ensure stabilization of the initial signal and adaptation of participants to the environment. Timing correction to the last slice was conducted. Realignment for compensation of head-movement effects was achieved using a six-parameter rigid-body spatial transformation. All spatial movement was <3 mm of displacement and <3° of rotation in any direction, and no participant was excluded. Next, rs-fMRI images were co-registered to the high-resolution 3D-T1 structural images. Normalization of 3D-T1 structural MRI images to Montreal Neurological Institute (MNI) space was undertaken by non-linear warping based on Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL). Then, rs-fMRI images were spatially normalized to the MNI space using the parameters derived from the normalization of structural images and simultaneously resampled into 3-mm isotropic voxels. All normalized fMRI images were smoothed with a 6-mm, full-width at half-maximum Gaussian kernel. Linear detrending and band-pass filtering at 0.01–0.1 Hz were carried out to control low-frequency drift and high-frequency physiological noise. Finally, nuisance covariates were regressed out, including the Friston 24-motion parameter model (six head-motion parameters, six head-motion parameters one time point before, and the 12 corresponding squared items), global mean, white matter, and cerebrospinal fluid signals.

Brain Network Construction

The average time series within each region based on the 264 putative functional area atlas were separately extracted to construct the connectivity brain network (Power et al., 2011). The Pearson’s correlation coefficients of all pairs of 264 regions of interest (ROIs) were applied separately to define the edges

²<http://restfmri.net/forum/index.php>

³<http://www.fil.ion.ucl.ac.uk/spm>

of functional connections. Thus, the functional connectivity matrix (adjacency matrix) was constructed (Li et al., 2017). The final functional connection networks produced N*(N-1)/2 edges, where N corresponded to the number of nodes in the networks. Considering the ambiguous interpretation of negative correlations, we restricted the analysis to positive correlations and set the negative correlation coefficients as zero. A thresholding method based on network sparsity was adopted to remove the less significant connections and to retain the topological properties of graph theory by setting an appropriate threshold for network sparsity (Dai et al., 2019). Sparsity thresholds (ranging from 0.02 to 0.5, with steps of 0.01) were set to acquire a binary undirected network (Chang et al., 2016). To avoid ambiguity, we used the area under the curve (AUC; i.e., the sum value of 49 values of the corresponding node attributes) as input for the node attribute to train the classifier.

Computation of Graph Metrics

Based on binary undirected matrices, we systematically analyzed the global and local properties of the functional brain network with the Graph Theoretical Network Analysis Toolbox (GRETN)⁴ based on Statistical Parametric Mapping (SPM8; see text footnote 3) with MATLAB R2013b. Global metrics [i.e., clustering coefficient (C_p), characteristic path length (L_p), normalized clustering coefficient (γ), normalized characteristic path length (λ), small-world σ, global efficiency (E_{global})], and nodal properties (i.e., degree centrality, nodal efficiency, betweenness centrality, shortest path length) were applied to characterize the different patterns of connections in the brain network (Table 1; Wang et al., 2015). The modularity (Q) of a brain network quantified the efficiency of segmenting a network into modules (Newman, 2006). A modified greedy optimization algorithm was used as follows:

$$Q = \sum_{i=1}^{N_m} [l_i/L - (d_i/2L)^2]$$

where N_m represents the number of modules, L is the total number of edges in the brain network, and l_i is the number of within-module edges in module i; d_i represents the sum of the linked edges at each node within module i. Modified greedy optimization was applied to detect the modular structure (Newman, 2004).

⁴www.nitrc.org/projects/gretna/

TABLE 1 | Global and local graph metrics of the brain connectome.

Global graph metrics	Local graph metrics
Clustering coefficient C _p	Betweenness centrality
Characteristic path length L _p	Degree centrality
Normalized clustering coefficient γ	Nodal clustering coefficient
Normalized characteristic path length λ	Local efficiency
Small-world σ	Shortest path length
Network efficiency	Participant coefficient
Modularity	Within-module degree

At the module level, the intra-module connectivity density (D_s) and intermodule connectivity density ($D_{s,t}$) were calculated as follows:

$$D_s = \frac{2 \sum_{i,j \in s} \varepsilon_{i,j}}{N_s(N_s - 1)}$$

where N_s is the number of nodes within module s , and $\varepsilon_{i,j}$ are the edges within module s .

$$D_{s,t} = \frac{\sum_{i \in s, j \in t} \varepsilon_{i,j}}{N_s * N_t}$$

where N_s is the number of nodes within module s , N_t represent the number of nodes within module t , and ε_{ij} is the number of edges between module s and module t .

Moreover, at the nodal level, within-module degree (WD) and the participation coefficient (PC) were measured as follows:

$$WD_i = \frac{e_i - \bar{e}_s}{\sigma_s}$$

where e_i is the nodal degree of node i within module s , \bar{e}_s is the average nodal degree of all nodes in module s , and σ_s is the standard deviation of the nodal degree within the module of all nodes in module s .

$$PC_i = 1 - \sum_{s=1}^{N_m} \left(\frac{k_{i,s}}{k_i} \right)^2$$

where N_m is the number of modules and $k_{i,s}$ is the number of connections between node i and module s . k_i represents the number of connections of node i to all other nodes within the N_m modules.

Nodes with a degree of 2 standard deviations higher than the mean of the degree of all nodes were identified as hub nodes (Rubinov and Sporns, 2010). Small-world attributes were applied to characterize an optimized balance between functional segregation and integration of the network.

Statistical Analyses

For demographics and clinical characteristics, two-sample Student's t -tests were carried out except for sex, which was tested by the chi-square test. $P < 0.05$ indicated a significant difference in the demographic data. First, functional connections and global and local metrics were regressed to remove potential effects of the covariates age, sex, and education duration. Then, differences pertaining to graph theory metrics between MCI patients and NCs were compared based on two-sample Student's t -tests. A procedure to ascertain the false discovery rate was performed to further correct for multiple comparisons. To localize the specific pairs of regions in which functional connections were altered in MCI patients, we used a network-based statistic (NBS) approach (Zalesky et al., 2010). A corrected P -value was calculated for each component using the null distribution of the maximal connected component size, which was empirically derived using a non-parametric permutation approach (10,000 permutations) (Zuo et al., 2012). $P < 0.01$ indicated a significant difference.

Feature Selection for Nodal Graph Metrics

As mentioned above, the brain was divided into 264 nodes based on the 264 putative functional area atlas (Power et al., 2011), and each node corresponded to seven local graph metrics (i.e., betweenness centrality, degree centrality, nodal clustering coefficient, local efficiency, shortest path length, participant coefficient, within-module degree). Thus, the nodal graph metrics naturally have a group topology, that is, a node corresponds to a group of node-graph theoretical attributes. Given the natural group attributes, we used group-LASSO as the feature-selection scheme for nodal graph metrics.

$$\min_w \sum_{i=1}^{nSub} \log \left(1 + \exp \left(-y_i \times \left(\sum_{j=1}^{nROI} \sum_{k=1}^7 w_{(j,k)} x_{(j,k)} + c \right) \right) \right) + \lambda \sum_{j=1}^{nROI} \|w_{jk}\|_q,$$

where y_i is the label of the i -th participant, and $w_{(j,k)}$ and $x_{(j,k)}$ are the weight and value of the j -th ROI and k -th Nodal Graphic Metric, respectively. Note that $x_{(j,k)}$ is normalized by Fisher Z-transformation to avoid scale imbalance. We used the SLEP toolbox⁵ to calculate $w_{(j,k)}$ with a default setting of $\lambda = 1$.

Classification

Combination of information provides an effective way to integrate multiple views of biomarkers (i.e., connections and graph metrics). The simplest way is to overlay the data directly, but this approach can be inappropriate due to the high-dimensional curve and small number of samples. Moreover, a modality with more dimensions can submerge a modality with fewer dimensions. To overcome this challenge, we used MKL-SVM for information combination because the kernel trick can partially alleviate the high-dimensional curve. MKL-SVM was conducted as shown below.

Suppose that there are n training samples with connection values and graph metrics. For x_i^m , $m = 1, 2, 3$, which correspond to the connection value, the nodal graph metrics and global graph metrics respectively. y represent the corresponding class label of the i -th sample. MKL-SVM solves the following primal problem:

$$\begin{aligned} \min_w \quad & \frac{1}{2} \sum_{m=1}^2 \beta_m \|w^m\|^2 + C \sum_{i=1}^2 \xi_i \\ \text{s.t.} \quad & y_i \left(\sum_{m=1}^2 \beta_m (w^m)^T \phi^m(x_i^m) + b \right) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad i = 1, 2 \end{aligned}$$

where ϕ^m represents a mapping from the original space to the Represent Hilbert Kernel Space (RHKS), w^m represents the normal vector of the hyperplane in RHKS, and β_m denotes the

⁵www.yelab.net/software/SLEP

corresponding combining weight on the m -th modality. Then, the dual form of MKL-SVM can be represented as:

$$\begin{aligned} & \max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \sum_{m=1}^2 \beta_m k^m(x_i^m, y_i^m) \\ & \text{s.t.} \sum_{i=1}^n \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, 2 \end{aligned}$$

where $k^m(x_i^m, y_i^m) = \phi^m(x_i^m)^T \phi^m(y_i^m)$ and is the kernel matrix on the m -th modality. After we trained the model, we tested the new samples $x = \{x_1, x_2, \dots, x_M\}$. The kernel between the new test sample and the i -th training sample on the m -th modality is defined as $k^m(x_i^m, x^m) = \phi^m(x_i^m)^T \phi^m(x^m)$. In the end, the predictive level based on MKL-SVM can be formulated as follows:

$$f(x_1, x_2, \dots, x_M) = \text{sign} \left(\sum_{i=1}^n y_i \alpha_i \sum_{m=1}^M \beta_m k^m(x_i^m, x^m) + b \right)$$

The proposed formulation of MKL-SVM is similar to but different from existing multi-kernel learning methods because β_m is selected based on the cross-validation scheme on the grid-searching space with constraints $\sum_m \beta_m = 1$. The range of c was 2^{-5} to 2^5 . All data-processing and classification procedures used in our study are shown in **Figure 1**. Due to the small sample size, we used the leave-one-out cross-validation (LOOCV) strategy to verify the performance of the methods, in which only

one subject is left out for testing while the others are used to train the models and obtain the optimal parameters. For the choice of optimal parameters, an inner LOOCV was conducted on the training data using a grid-search strategy. Moreover, in order to verify the performance of the proposed model, we also tested the model on the independent ADNI dataset.

RESULTS

Demographics and Clinical Characteristics

The demographic data and clinical characteristics of all participants are summarized in **Table 2**. There were no significant differences in sex, age, or education level between the MCI and NC groups ($P > 0.05$ for all). However, the MCI group had significantly lower scores on the Mini Mental State Examination ($P < 0.001$) than the NC group. We also selected 50 samples (27 MCI and 23 NCs) from the independent ADNI dataset. The details of their demographic and clinical characteristics are listed in **Table 3**.

Significant Differences of Functional Connections in Brain-Network

The mean connection strengths of the whole brain network were compared between MCI and NC. A total of 3072 connections with significant differences were extracted between the MCI and NC groups within the range of fully sparse values from 0.02 to 0.5 ($P < 0.01$) using Student's t -tests. After permutation of NBS,

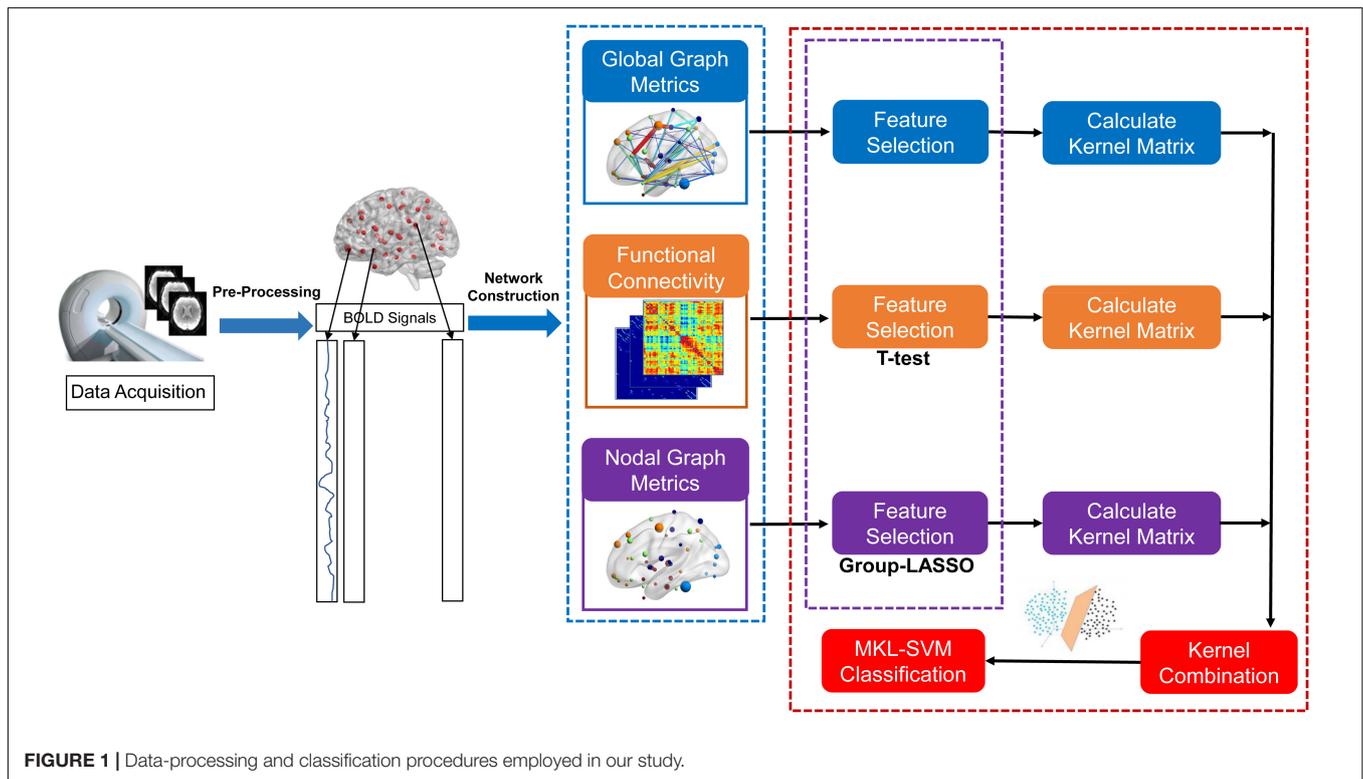


FIGURE 1 | Data-processing and classification procedures employed in our study.

TABLE 2 | Demographics and clinical characteristics of MCI patients and NCs in the current study.

Characteristic	MCI (n = 39)	NCs (n = 60)	P
Male/female, n	25/14	30/30	0.168 ^a
Age, years, mean ±SD	74.00 ± 7.67	71.25 ± 7.08	0.071 ^b
Education, years, mean ±SD	10.97 ± 4.29	12.42 ± 3.58	0.074 ^b
MMSE, mean ±D	26.77 ± 2.33	28.28 ± 1.35	<0.001 ^b
Hippocampal volume (×10 ³ mm ³)	6.80 ± 0.87	7.43 ± 0.69	0.002 ^b

MMSE, Mini mental state examination. ^aThe P-value was obtained by using the chi-square test. ^bThe P-value was obtained by using a two-sample t-test.

TABLE 3 | Demographics and clinical characteristics of the ADNI dataset.

Characteristic	MCI (n = 27)	NCs (n = 23)	P
Male/female, n	13/14	11/12	0.982 ^a
Age, years, mean ±SD	70.11 ± 8.17	75.22 ± 6.82	0.021 ^b
MMSE, mean ±SD	25.33 ± 1.07	27.17 ± 1.30	<0.001 ^b

MMSE, Mini mental state examination. ^aThe P-value was obtained by using the chi-square test. ^bThe P-value was obtained by using a two-sample t-test.

we retained the most significant 100 connections with the lowest P-values (Figure 2). We projected them into the corresponding subnetworks and found that the most discriminative network connections were mainly distributed in the default mode network (DMN), subcortical network, frontoparietal task control network,

dorsal attention network, and visual network. Compared with NCs, patients with MCI had significantly lower functional connection strength in brain-network connections ($P < 0.01$).

Global Graph Metrics of the Functional Brain Connectome

The global graph metrics of the MCI and NC groups showed the small-world topological attributes. That is, the functional brain networks had larger clustering coefficients and almost identical shortest path lengths compared with the matched random networks. With increasing connection density, C_p increased, whereas L_p , γ , λ , and small-world σ decreased in the MCI and NC groups. Statistical analysis revealed that the C_p of MCI patients was higher than that in the NC group, whereas λ and small-world σ were lower in the MCI group compared with the NC group ($P < 0.01$). However, these differences were only observed at a few network thresholds (Figure 3).

Nodal Graph Metrics of the Functional Brain Connectome

Two strategies were developed to investigate the discriminative features of nodal graph metrics and nodes based on local network parameters. On the one hand, we analyzed the most predominant brain regions with the greatest number of significant differences in nodal graph metrics. Before group-LASSO, 212 significantly different nodes were observed between MCI and NC groups ($P < 0.01$). However, after feature selection by group-LASSO,

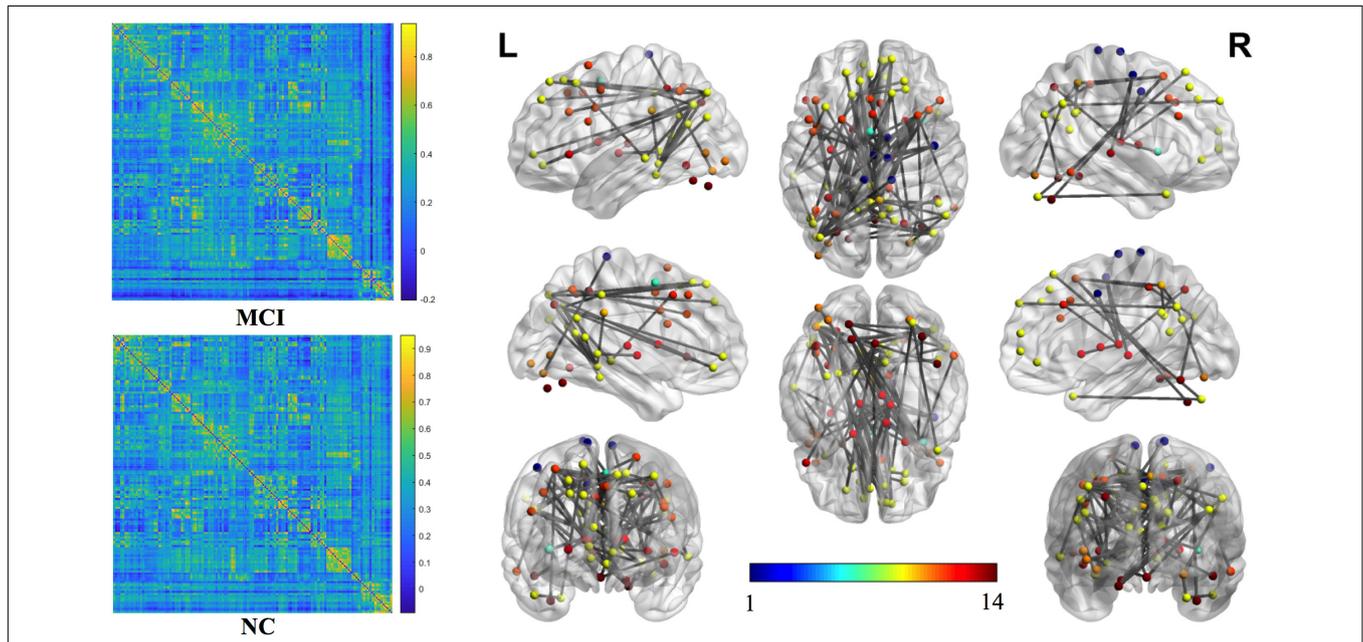


FIGURE 2 | The most significant 100 connections mapped on the ICBM 152 template using the BrainNet Viewer software package (<http://nitrc.org/projects/bnv/>). The connectivity matrices of the fully connected network of MCI patients and NCs are shown. The 100 most significant connections were retained, with gray indicating a reduction in connectivity strength. Plots in this figure were created by BrainNet Viewer (<http://nitrc.org/projects/bnv/>). The color-bar numbers represent the subnetworks with reference to the 264 putative functional area atlas proposed by Power et al. (2011). The details are: 1 sensory/somatomotor hand network; 2 sensory/somatomotor mouth network; 3 cingulo-opercular task control network; 4 auditory network; 5 default mode network; 6 memory retrieval network; 7 visual network; 8 frontoparietal task control network; 9 salience network; 10 subcortical network; 11 ventral attention network; 12 dorsal attention network; 13 cerebellar network; 14 unknown network.

we selected the nodal graph metrics from 76 ROIs as inputs. These 76 ROIs were considered as the extremely predominant nodes for discriminating MCI patients from NCs, and each ROI had ≥ 4 and ≤ 7 nodal topological metrics with significant differences. The locations of nodes in the 264 atlas were labeled according to the AAL_90 atlas (Figure 4 and Table 4). On the other hand, we identified the distinguishing features for each nodal graph theory attribute using the feature selection of group-LASSO (Table 5). The top-20 nodal graph topological features with maximum discriminative ability are listed in Table 6. Therefore, the most predominant brain regions with the greatest numbers of significant nodal graph measures and the most discriminative nodal graph features were distributed mainly in the temporal, cingulate, superior frontal, lingual, and parietal gyri, which corresponded to the DMN, dorsal attention network, and cingulo-opercular task network.

According to the definition of “hubs,” we identified hub nodes in MCI patients and NCs. Figure 5 shows the hub nodes in each group. In MCI patients and NCs, the common hub regions were mainly located in the left middle temporal gyrus,

right precuneus, left median cingulate gyrus, left cuneus, and paracingulate gyri. More importantly, some hub nodes were present only in MCI patients and absent in NCs: the left paracentral lobule, right paracentral lobule, left postcentral gyrus, and right cuneus. Simultaneously, there were also some hub nodes in NCs but not in MCI patients. These regions were located on the left Heschl, right superior temporal, left inferior occipital, and left middle occipital gyri. Hub nodes play critical roles in maintaining high-level cognitive functions by coordinating overall information flow and supporting the integrity of the brain connectome (Wang et al., 2013). The similar distributions suggested preservation of hubs in MCI.

Further comparisons of the predominant brain regions mentioned above revealed that MCI patients had significantly lower values of betweenness centrality and degree centrality and significantly higher values for the nodal shortest path in the frontal lobe (e.g., bilateral superior frontal gyrus), temporal lobe (e.g., bilateral inferior temporal gyrus), limbic lobe (e.g., left median cingulate and paracingulate gyri), and parietal lobe (e.g., left inferior parietal gyrus) compared with the NC group

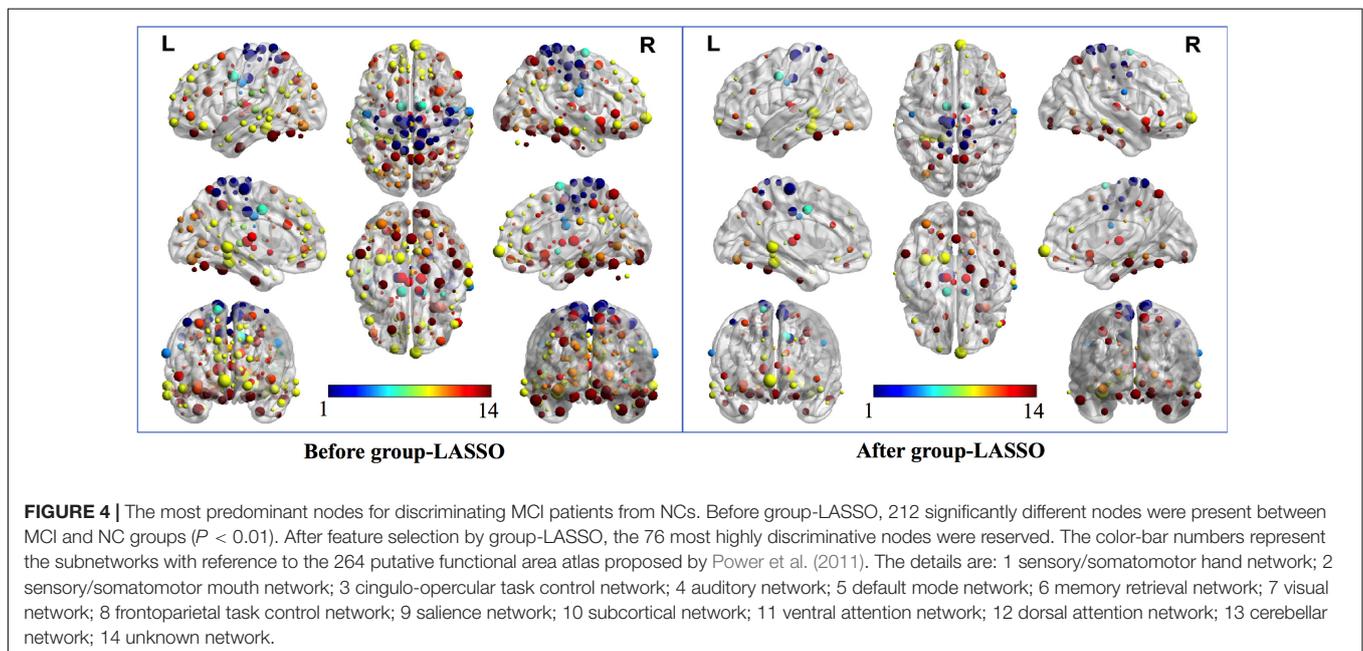
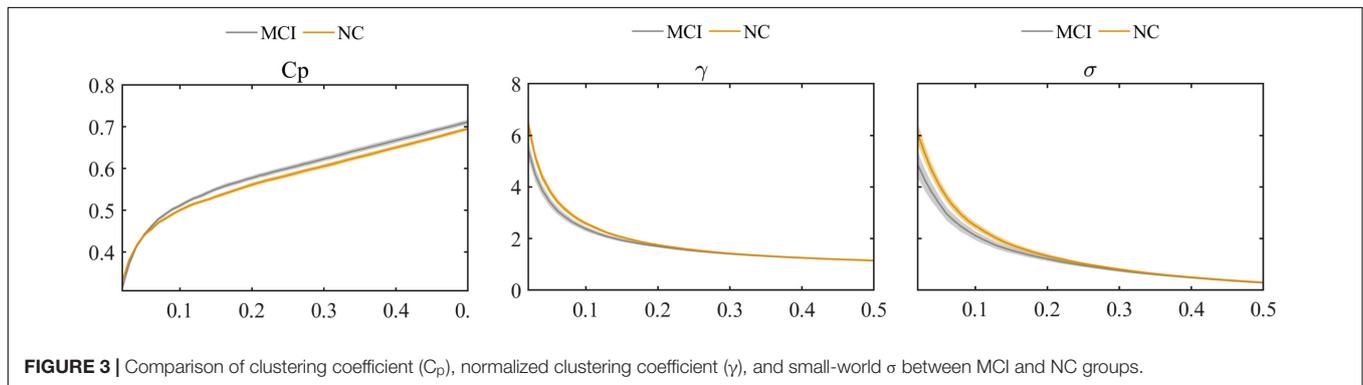


TABLE 4 | Top 20 most predominant nodes (brain regions) with the greatest number of significant differences in nodal graph metrics.

ROI number	Corresponding AAL area	Sub-network	Number of nodal metrics
77	Lingual_L	Default mode	7
126	Fusiform_L	Default mode	7
4	Temporal_Inf_L	Unknown	7
116	Temporal_Mid_R	Default mode	7
22	Precuneus_R	Sensory/somatomotor	6
17	Paracentral_Lobule_L	Sensory/somatomotor	6
251	Precuneus_R	Dorsal attention	6
259	Parietal_Inf_L	Dorsal attention	6
75	Frontal_Mid_Orb_R	Default mode	6
92	Cingulum_Post_R	Default mode	6
224	Thalamus_L	Subcortical	6
225	Thalamus_R	Subcortical	6
53	Supp_Motor_Area_R	Cingulo-opercular task	6
211	Insula_R	Salience	6
203	Cingulum_Mid_R	Salience	6
124	ParaHippocampal_L	Default mode	6
139	Frontal_Inf_Orb_R	Default mode	5
51	Cingulum_Mid_L	Cingulo-opercular task	5
172	Fusiform_L	Visual	5
263	Parietal_Sup_L	Dorsal attention	5

AAL, the automated anatomical labeling atlas.

($P < 0.01$ for all). Nevertheless, in the occipital lobe (e.g., left lingual and left fusiform gyri), the MCI group showed significantly higher values of betweenness centrality and degree centrality and significantly lower values of nodal shortest path, which was opposite to the pattern of nodal graph metrics in the brain lobes mentioned above (Figure 6).

Classification

After feature selection of functional connections with Student's *t*-tests and nodal graph metrics by group-LASSO, MKL-SVM was carried out to combine the brain connectome information. We evaluated the classification performance of different methods with a set of quantitative measures – accuracy, sensitivity, and specificity – which were defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN},$$

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

$$Specificity = \frac{TN}{TN + FP},$$

where TP, TN, FP, and FN denote the number of true-positive, true-negative, false-positive, and false-negative values, respectively. The area under the receiver operating characteristic curve (AUC) was calculated as a performance measure for binary classification of the MCI and NC groups. In particular, LOOCV was employed in this study due to the small sample size, which provided an optimistic estimate of the classification accuracy since all except one of the subjects are used to

TABLE 5 | Number of discriminative features for each nodal graph metrics from the feature-selection step of LASSO.

Nodal graph metric	Number of selected features
Betweenness centrality	33
Degree centrality	46
Nodal clustering coefficient	48
Nodal local efficiency	19
Nodal shortest path length	44
Participant coefficient	70
Within-module degree	81

TABLE 6 | Top 20 features corresponding to nodal graph metrics with maximum discriminative ability.

Nodal graph measure	ROI number	Corresponding AAL area	Subnetwork
Within-module degree	124	ParaHippocampal_L	Default mode
Within-module degree	89	Precuneus_R	Default mode
Within-module degree	191	Parietal_Sup_L	Frontoparietal task
Degree centrality	77	Lingual_L	Default mode
Participant coefficient	9	Temporal_Inf_R	Uncertain
Within-module degree	92	Cingulum_Post_R	Default mode
Degree centrality	225	Thalamus_R	Subcortical
Participant coefficient	118	Temporal_Mid_L	Default mode
Participant coefficient	75	Frontal_Mid_Orb_R	Default mode
Nodal clustering coefficient	75	Frontal_Mid_Orb_R	Default mode
Nodal shortest path length	75	Frontal_Mid_Orb_R	Default mode
Participant coefficient	17	Paracentral_Lobule_L	Sensory/somatomotor
Degree centrality	224	Thalamus_L	Subcortical
Nodal shortest path length	9	Temporal_Inf_R	Uncertain
Participant coefficient	83	Temporal_Inf_L	Default mode
Degree centrality	126	Fusiform_L	Default mode
Betweenness centrality	77	Lingual_L	Default mode
Nodal clustering coefficient	77	Lingual_L	Default mode
Betweenness centrality	51	Cingulum_Mid_L	Cingulo-opercular task
Nodal local efficiency	92	Cingulum_Post_R	Default mode

AAL, the automated anatomical labeling atlas.

train the classifier. For other approaches such as k-fold cross-validation, only N-k (N is the total number of participants in the dataset) participants are included during the training process, resulting in poorer performance due to the small dataset (Wee et al., 2012). For the functional connections (C), global metrics (G), and nodal metrics (N) of the brain network, we obtained AUCs of 0.9605, 0.7290, and 0.9576, respectively (Table 7). We also performed classification experiments by combining functional connections (C), global metrics (G), nodal metrics (N), global metrics (G), and nodal metrics (N). The results showed that despite the low classification performance of single global graph metrics, they still effectively increased the classification performance of nodal graph metrics and functional connections. For a direct combination of connections, global metrics, and nodal metrics, we obtained 87.88% accuracy and

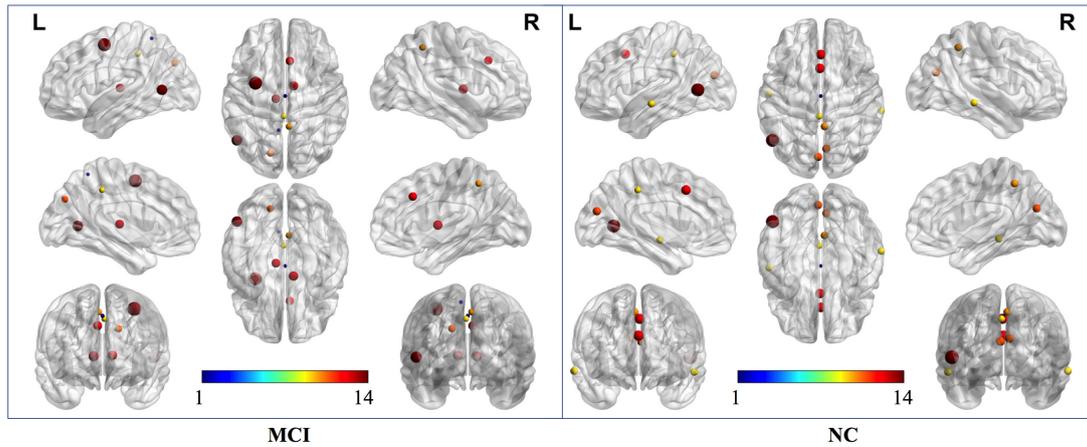


FIGURE 5 | Hub nodes of MCI and NC groups in the brain. The color-bar numbers represent the subnetworks with reference to the 264 putative functional area atlas proposed by Power et al. (2011). The details are: 1 sensory/somatomotor hand network; 2 sensory/somatomotor mouth network; 3 cingulo-opercular task control network; 4 auditory network; 5 default mode network; 6 memory retrieval network; 7 visual network; 8 frontoparietal task control network; 9 salience network; 10 subcortical network; 11 ventral attention network; 12 dorsal attention network; 13 cerebellar network; 14 unknown network.

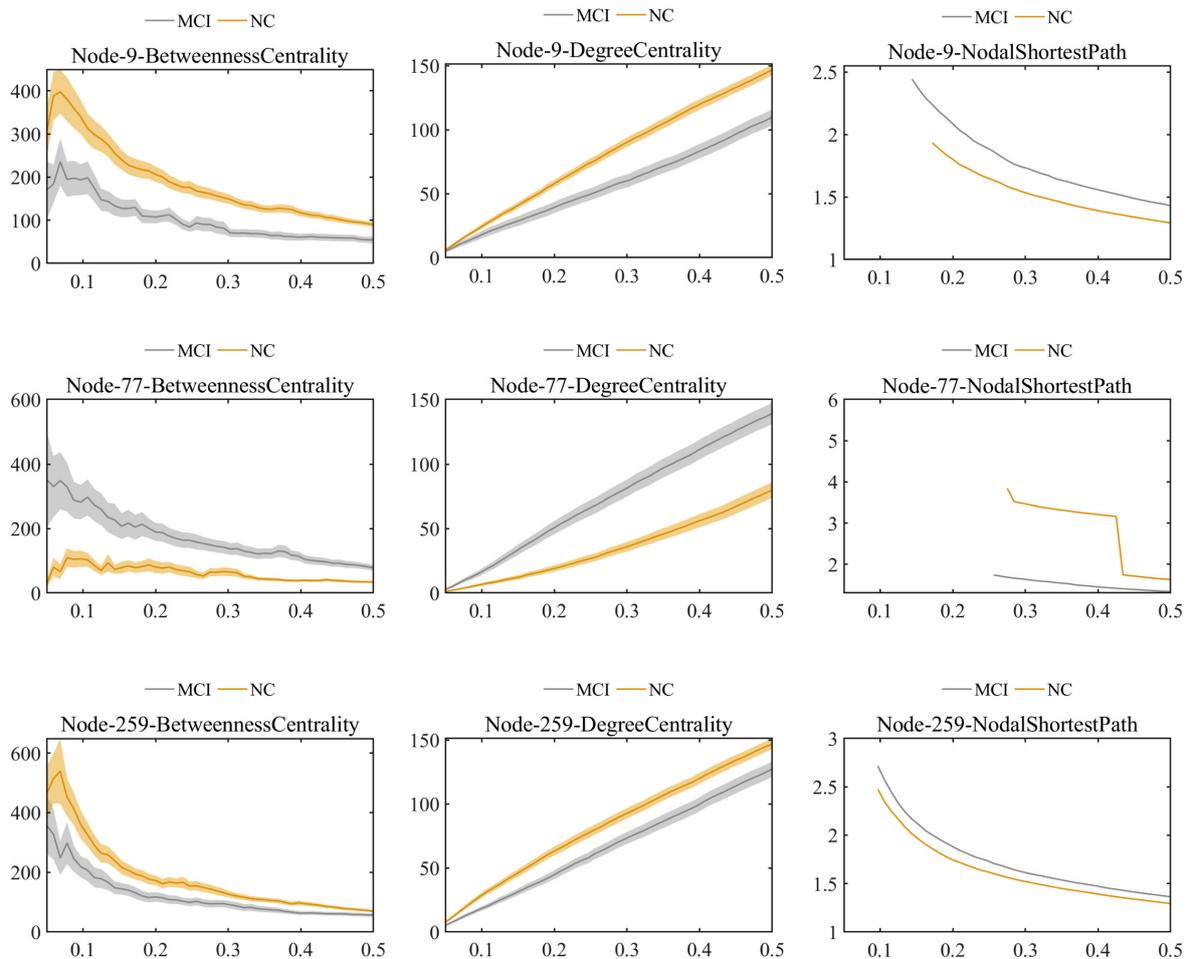


FIGURE 6 | Comparison of values of nodal graph metrics between MCI patients and NCs. Betweenness centrality, degree centrality, and nodal shortest path length of Node 9 (right inferior temporal gyrus). Betweenness centrality, degree centrality and nodal shortest path length of Node 259 (left inferior parietal). Betweenness centrality, degree centrality, and nodal shortest path length of Node 77 (left lingual gyrus).

TABLE 7 | The evaluation of classification performance corresponding to different functional connectome features.

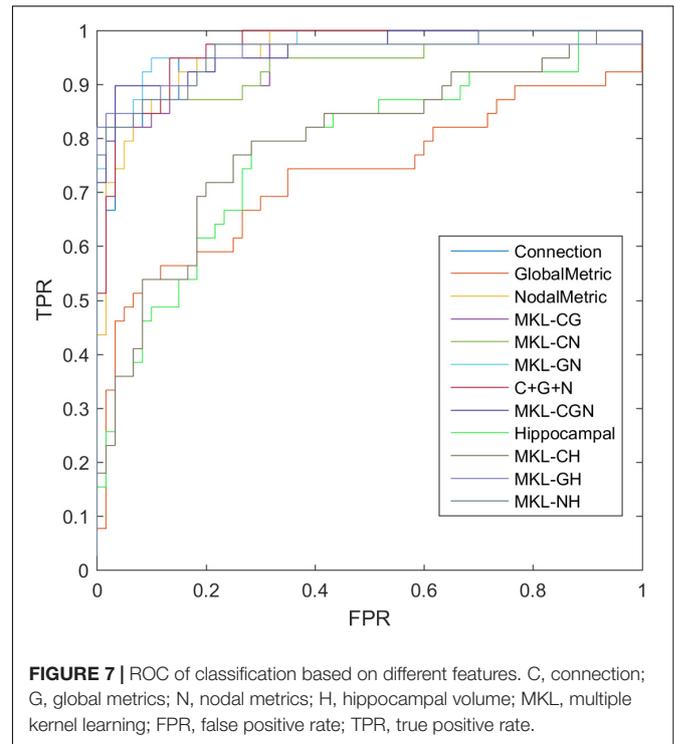
Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Connection (C)	85.86	82.05	88.33	0.9605
Global Metrics (G)	73.74	69.23	76.67	0.7290
Nodal Metrics (N)	87.88	82.05	91.67	0.9576
MKL_CG	86.87	82.05	90.00	0.9329
MKL_CN	90.91	84.62	95.00	0.9581
MKL_GN	89.90	84.62	93.33	0.9371
C + G + N	87.88	92.31	85.00	0.9666
MKL_CGN	92.93	89.74	95.00	0.9728
Hippocampal (H)	72.73	71.67	74.36	0.7005
MKL_CH	86.86	84.62	88.33	0.9509
MKL_GH	76.77	73.33	82.05	0.8117
MKL_NH	89.90	87.18	91.67	0.9647

MKL-SVM, multiple kernel learning support vector machine.

an AUC of 0.9666, which meant that simple combination did not effectively improve the classification performance. Finally, the combination of all connectome features based on MKL-SVM achieved the best classification performance, with 92.93% accuracy, 95.00% specificity, and an AUC of 0.9728. Moreover, the weight values (β) of functional connections, global metrics, and nodal metrics were 0.3, 0.01, and 0.6, respectively, indicating that the node attributes contributed most to the classification (Figure 7). It should be noted that MKL-SVM both combines the information of functional connectivity and graph theory attributes and provides a method to merge more useful information for MCI identification. Therefore, we also combined the traditional unimodal marker of hippocampal volume with the brain connectome; the results are listed in Table 7. Our results suggest that the AUC of the hippocampal volume was 0.7005, and the AUCs of the combination of hippocampal volume with functional connectivity, global graph theory attributes, or node graph theory attributes were 0.9509, 0.8117, and 0.9647, respectively. In addition, the independent ADNI dataset was then employed to verify the generalization of the pre-trained model. The all connectome features combination based on MKL-SVM achieved classification performance with 66.00% accuracy, 70.37% sensitivity, and 60.87% specificity.

DISCUSSION

In the present study, we selected discriminative features from different attributes of the brain connectome (i.e., functional connections, global graph metrics, and nodal graph metrics) and combined the information to train a classifier for distinguishing subjects with MCI from NCs. Based on the feature selection and combination of the proposed methods, we further described the altered patterns of the best distinguishing features of MCI through group comparison, aiming to further clarify disease pathogenesis. Our detailed results are listed as follows. First, the most predominant brain regions and most discriminative nodal graph metrics for discriminating NCs from MCI were selected



by the group-LASSO. Second, the information combination strategy (MKL-SVM) effectively improved the classification performance, and the nodal graph metrics of the connectome contributed most to the classification. Finally, the altered functional brain connectome pattern in MCI patients included a general decrease in functional connections in the whole brain network, whereas nodal topological attributes in some local brain regions were increased.

The Most Predominant Brain Regions and Discriminative Nodal Graph Metrics

The nodal graph metrics have a natural group topology; that is, a node corresponds to a group of node-graph theoretical attributes. Thus, we used group-LASSO as the feature-selection scheme for nodal graph metrics. It effectively extracted the group-structure information of nodal attributes. The most predominant brain regions (with seven significantly different nodal topological metrics) were mainly distributed in the left lingual, left fusiform, left inferior temporal, and right middle temporal gyri. These brain regions showed significant changes in nodal graph metrics and so could be regarded as the most sensitive observation areas for nodal topological attributes in MCI patients. Also, within-module degree, degree centrality, and participation coefficient showed the most significant discriminative ability among the selected nodal graph metrics. The corresponding brain regions with the three most discriminative nodal metrics considerably overlapped with the hub nodes found in MCI patients. Overall, our results emphasize the importance of analyzing the attributes of intra-modules and hub nodes for early discrimination of NCs from subjects with MCI.

By projecting brain regions with significant differences of functional connections and graph metrics in the brain network to subnetworks, we found that the differences between MCI patients and NCs were distributed mainly in the DMN, dorsal attention network, cingulo-opercular task network, and frontoparietal task network. Of these, the DMN had the most significant discriminative ability. Studies have verified the correlations between these subnetworks and cognitive functions in the human brain, corresponding to spatial attention (Rolle et al., 2017), visual attention (Wirth et al., 2017), and executive function (Talpos and Shoaib, 2015).

In this study, the DMN carried the most distinguishing information, which was verified by the proposed feature selection methods. Previous studies showed that the DMN is involved in episodic memory and is considered the major cognitive domain impaired in the early stage of AD (Meskaldji et al., 2016; Dillen et al., 2017). Besides validating the discriminative ability of the DMN for discriminating NCs from MCI, we accurately located the predominant brain regions (middle temporal, inferior temporal, lingual, posterior cingulate, and middle frontal gyri) in the DMN and the corresponding nodal graph metrics. These results may facilitate the early and accurate diagnosis of MCI. They also demonstrate the repeatability and verifiability of the proposed methods, which is an important contribution of our work.

Fusion Classification of MKL-SVM and Identification of Maximum Contribution

Group-LASSO is valid for nodal feature selection because it can retain significant features with the most discriminative ability while avoiding data redundancy. We carried out reduction of nodal features according to group-LASSO and selected optimal features to achieve the best performance for discriminating NCs from MCI. This is an effective way to integrate multiple views of biomarkers for AD classification. The simplest way is to directly splice the data. Studies using multivariate pattern analysis [e.g., linear discriminant analysis (Alam et al., 2017), artificial neural networks (Quintana et al., 2012), and random forest (Sarica et al., 2017)] have been undertaken to identify MCI using complex network characteristics. However, those approaches could be inappropriate due to the high-dimensional curves and small samples. Information with higher dimensions can submerge the low-dimension information. To overcome these challenges, we employed MKL-SVM for information combination. MKL (Niu et al., 2017) is a sparse machine-learning method that allows identification of the most relevant classification sources. The results suggested that the performance of classification by combining multiple brain connectome features was better than that of individual connectome features. The weight value (β) of functional connections, global metrics, and nodal metrics emphasized that nodal graph attributes had the greatest contribution to classification. It also indicated that MCI patients had significant changes in nodal properties. More surprisingly, although global metrics showed the worst classification performance, they can still provide important information about functional connections and nodal metrics.

After combining functional connections and global metrics (C + G), functional connections and nodal metrics (C + N), and global metrics and nodal metrics (G + N), the results indicated that classification performance was effectively improved by combining the information of global metrics.

To verify this significant improvement, the DeLong test was applied (DeLong et al., 1988). We found that the proposed method significantly outperformed the global graph attributes, functional connection, and nodal graph attributes under the 95% confidence interval with *P*-values of 0.0002, 0.0227, and 0.0419, respectively. Although MKL-SVM did not yield significant improvements compared to the feature concatenation method (*P* = 0.1627), it still had two advantages. First, MKL-SVM could address the imbalanced dimension issue across modalities to some extent and better embody the contribution of different information sources to distinguish MCI patients from NCs. Second, experimental results demonstrated that the proposed method outperformed the single modality of the functional connectome in the brain network. It should also be noted that both methods are simple attempts to verify information effectiveness.

The classification results based on the traditional marker of hippocampal volume suggested that the combination of hippocampal volume and connectome features could also improve classification accuracy. The MKL-SVM can be used to combine multiple features of the brain connectome and effectively integrate multimodal information to discriminate NCs from patients with MCI.

During validation of the proposed model, the classification performance of the independent ADNI dataset was not as good as the pre-trained sample. This may be due to heterogeneity in scanning machines, parameters, and physiological structures between western and eastern samples, which obviously violates the independently identically distribution assumption of SVM.

Altered Pattern of the Brain Network Connectome in MCI

At the global brain level, we found that MCI patients had weaker functional connections in the brain network, which was consistent with previous functional network studies of AD (Li et al., 2016) and MCI (Wang et al., 2013; Lee et al., 2016). Some results demonstrated that these abnormal functional connections were directly related to the global topological attributes of brain networks (Wang et al., 2013). In our study, we first found that patients with MCI and NCs fit the features of a small-world network in a global network topology. That is, the brain network supported rapid, real-time integration of information across separate sensory brain regions to confer resilience against pathology and maximize efficiency with minimal cost for effective information processing between brain regions (Sporns and Zwi, 2004; Achard and Bullmore, 2007; Sporns, 2011). Further comparison suggested that the value of small-world σ in MCI patients was lower than that in NCs, indicating “economic small-world” disruption (Liao et al., 2017) (i.e., reduction of the segregation and integration functions of effective information in the brain network). Moreover, we found changes in the

functional segregation of brain networks in MCI patients (increased C_p). C_p is a measure of local network connectivity (Bullmore and Sporns, 2009) that reflects the efficiency of local information transfer and the ability to defend against random attacks against a network. A higher value of C_p represents a more concentrated clustering of local connections and a stronger capacity for processing local information. It is notable that previous studies reported decreased C_p in AD patients (Zhao et al., 2012). The reason for this difference might be related to the compensatory change of segregation function in the transition stage of MCI. Therefore, our results suggested that functional connections in the whole-brain network were generally decreased, whereas the network segregation of local information processing was increased.

At the local brain level, further analyses of the hub nodes and nodes with the most discriminative ability for MCI showed that MCI patients had significantly lower values of betweenness centrality and degree centrality and higher values of nodal shortest path in some brain regions (the frontal, temporal, limbic, and parietal lobes) compared with NCs. These data suggested that the network integration and local transmission capability of these lobes were decreased in MCI patients. However, in critical nodes in the occipital lobe, the increased betweenness/degree centrality and decreased shortest path indicated enhanced integration function and greater local transmission efficiency. We speculated that enhanced variation of these nodal graph metrics in some occipital nodes suggests compensation to maintain high-level cognitive performance despite the pathological process of amyloid accumulation during the earliest phases of AD. This functional variation in the occipital lobe was also mentioned in previous studies. For example, Dai et al. found that the left fusiform gyrus exhibited higher functional connections in the AD group (Dai et al., 2015). Bokde et al. (2010) found significantly greater activation in the right middle occipital gyrus during the location-matching task.

Therefore, the altered brain connectome patterns in our study revealed that functional connections generally decreased in the whole brain network but increased for nodal graph topological attributes of local brain regions. This might suggest functional compensation in some brain regions to maintain normal cognitive function in the early stage of AD.

Limitations and Future Directions

There are still several limitations that need to be considered further. First, the class imbalance issue. Although there are several approaches (e.g., resampling or reweighting) to overcome imbalance, taking them makes it difficult to estimate whether the improvement of performance is based on these adjustments or on the proposed methods. In the future, we plan to investigate high-quality data with more balanced samples for feature selection and classification or develop a more robust algorithm that improves classification accuracy and generalization.

Second, we assessed a small sample size. The optimization of parameters and hyperparameters inevitably leads to overfitting for small samples. To avoid this issue, we empirically chose parameters with a default setting of $\lambda = 1$ and $C = 1$ instead of optimized parameters and hyperparameters. In the

future, we will conduct parameter optimization based on a larger sample size.

Third, we must consider the generalization of the model. For the independent ADNI dataset, classification performance was not as good as observed for the pre-trained sample, which suggests a limitation in modal generalization for different centers. We intend to improve the classification performance of multicenter data sources by combining domain adaptation.

Finally, our cross-validation approach may have been insufficient. Evaluation of classification by k -fold cross-validation might be more precise when sufficient data are available. Therefore, in the future, it is necessary to compare the results obtained by different cross-validation methods (i.e., LOOCV and k -fold cross-validation).

CONCLUSION

In the present study, the discriminative features of functional connections and nodal graph metrics were selected by Student's t -tests and group-LASSO, respectively. The combination of all connectome information using MKL-SVM achieved the best classification performance (AUC = 0.9728). In addition, the altered brain connectome pattern revealed that functional connectivity was generally decreased in the whole-brain network, whereas graph theory topological attributes of some special nodes were increased in MCI patients. Our findings demonstrate that optimal feature selection and the combination of all connectome features could achieve good performance for discriminating NCs from MCI. The combination of functional connections and global and nodal graph metrics of brain networks can predict the occurrence of MCI and contribute to the early clinical diagnosis of AD.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Huashan Hospital within Fudan University (Shanghai, China). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XX and PW designed the study. QZ diagnosed patients. XL and WW administered the neuropsychological tests. XW acquired the MRI data. XX, WL, and JM analyzed and interpreted the results of the data. XX and MT drafted the manuscript. DD and PW revised the manuscript. All authors contributed equally to this work and approved the final manuscript.

FUNDING

This work was partially supported by the National Natural Science Foundation of China (Grant Nos. 81830059, 81571655, and 81771889), the Shanghai Municipal Commission of Health and Family Planning Smart Medical Special Research Project

REFERENCES

- Achard, S., and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* 3:e17. doi: 10.1371/journal.pcbi.0030017
- Alam, S., Kwon, G. R., Kim, J. I., and Park, C. S. (2017). Twin SVM-based classification of Alzheimer's disease using complex dual-tree wavelet principal coefficients and LDA. *J. Healthc. Eng.* 2017:8750506. doi: 10.1155/2017/8750506
- Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739. doi: 10.1073/pnas.0911855107
- Blennow, K., de Leon, M. J., and Zetterberg, H. (2006). Alzheimer's disease. *Lancet* 368, 387–403. doi: 10.1016/S0140-6736(06)69113-7
- Bokde, A. L., Lopez-Bayo, P., Born, C., Ewers, M., Meindl, T., Teipel, S. J., et al. (2010). Alzheimer disease: functional abnormalities in the dorsal visual pathway. *Radiology* 254, 219–226. doi: 10.1148/radiol.2541090558
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Chang, T. Y., Huang, K. L., Ho, M. Y., Ho, P. S., Chang, C. H., Liu, C. H., et al. (2016). Graph theoretical analysis of functional networks and its relationship to cognitive decline in patients with carotid stenosis. *J. Cereb. Blood Flow Metab.* 36, 808–818. doi: 10.1177/0271678X15608390
- Dai, Z., Lin, Q., Li, T., Wang, X., Yuan, H., Yu, X., et al. (2019). Disrupted structural and functional brain networks in Alzheimer's disease. *Neurobiol. Aging* 75, 71–82. doi: 10.1016/j.neurobiolaging.2018.11.005
- Dai, Z., Yan, C., Li, K., Wang, Z., Wang, J., Cao, M., et al. (2015). Identifying and mapping connectivity patterns of brain network hubs in Alzheimer's disease. *Cereb. Cortex* 25, 3723–3742. doi: 10.1093/cercor/bhu246
- delEtoile, J., and Adeli, H. (2017). Graph theory and brain connectivity in Alzheimer's disease. *Neuroscientist* 23, 616–626. doi: 10.1177/1073858417702621
- DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845.
- Dillen, K. N. H., Jacobs, H. I. L., Kukulja, J., Richter, N., von Reutern, B., Onur, O. A., et al. (2017). Functional disintegration of the default mode network in prodromal Alzheimer's disease. *J. Alzheimers Dis.* 59, 169–187. doi: 10.3233/JAD-161120
- Filippi, M., Basaia, S., Canu, E., Imperiale, F., Magnani, G., Falautano, M., et al. (2018). Changes in functional and structural brain connectome along the Alzheimer's disease continuum. *Mol. Psychiatry* 25, 230–239. doi: 10.1038/s41380-018-0067-8
- Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., et al. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch. Neurol.* 61, 59–66. doi: 10.1001/archneur.61.1.59
- Ijaopo, E. O. (2017). Dementia-related agitation: a review of non-pharmacological interventions and analysis of risks and benefits of pharmacotherapy. *Transl. Psychiatry* 7:e1250. doi: 10.1038/tp.2017.199
- 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
- Khazaei, A., Ebrahimzadeh, A., Babajani-Feremi, A., and Alzheimer's Disease Neuroimaging Initiative, (2017). Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI. *Behav. Brain Res.* 322(Pt B), 339–350. doi: 10.1016/j.bbr.2016.06.043
- (Grant No. 2018ZHYL0105), the Fundamental Research Funds for the Central Universities (Grant No. 22120190219), Natural Science Foundation Project of CQCSTC (cstc2018jcyjA2756), Shanghai Municipal Planning Commission of Science and Research Fund (201740010), Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01), and ZJLab.
- Lee, E. S., Yoo, K., Lee, Y. B., Chung, J., Lim, J. E., Yoon, B., et al. (2016). Default mode network functional connectivity in early and late mild cognitive impairment: results from the Alzheimer's disease neuroimaging initiative. *Alzheimer. Dis. Assoc. Disord.* 30, 289–296. doi: 10.1097/WAD.0000000000000143
- Li, W., Qiao, L., Zhang, L., Wang, Z., and Shen, D. (2019). Functional brain network estimation with time series self-scrubbing. *IEEE J. Biomed. Health Inform.* 23, 2494–2504. doi: 10.1109/JBHI.2019.2893880
- Li, W., Wang, Z., Zhang, L., Qiao, L., and Shen, D. (2017). Remodeling Pearson's correlation for functional brain network estimation and autism spectrum disorder identification. *Front. Neuroinform.* 11:55. doi: 10.3389/fninf.2017.00055
- Li, Y., Liu, J., Gao, X., Jie, B., Kim, M., Yap, P. T., et al. (2019). Multimodal hyper-connectivity of functional networks using functionally-weighted LASSO for MCI classification. *Med. Image Anal.* 52, 80–96. doi: 10.1016/j.media.2018.11.006
- Li, Y., Wang, X., Li, Y., Sun, Y., Sheng, C., Li, H., et al. (2016). Abnormal resting-state functional connectivity strength in mild cognitive impairment and its conversion to Alzheimer's disease. *Neural Plast.* 2016:4680972. doi: 10.1155/2016/4680972
- Liao, X., Vasilakos, A. V., and He, Y. (2017). Small-world human brain networks: perspectives and challenges. *Neurosci. Biobehav. Rev.* 77, 286–300. doi: 10.1016/j.neubiorev.2017.03.018
- Liu, X., Cao, P., Wang, J., Kong, J., and Zhao, D. (2019). Fused group lasso regularized multi-task feature learning and its application to the cognitive performance prediction of Alzheimer's disease. *Neuroinformatics* 17, 271–294. doi: 10.1007/s12021-018-9398-5
- Manly, J. J., Tang, M. X., Schupf, N., Stern, Y., Vonsattel, J. P., and Mayeux, R. (2008). Frequency and course of mild cognitive impairment in a multiethnic community. *Ann. Neurol.* 63, 494–506. doi: 10.1002/ana.21326
- Meskaldji, D. E., Preti, M. G., Bolton, T. A., Montandon, M. L., Rodriguez, C., Morgenthaler, S., et al. (2016). Prediction of long-term memory scores in MCI based on resting-state fMRI. *Neuroimage Clin.* 12, 785–795. doi: 10.1016/j.nicl.2016.10.004
- Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. *Phys. Rev. E* 69(Pt 2):066133. doi: 10.1103/PhysRevE.69.066133
- Newman, M. E. (2006). Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* 74(Pt 2):036104. doi: 10.1103/PhysRevE.74.036104
- Niu, W., Xia, K., Zu, B., and Bai, J. (2017). Efficient multiple kernel learning algorithms using low-rank representation. *Comput. Intell. Neurosci.* 2017:3678487. doi: 10.1155/2017/3678487
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* 256, 183–194. doi: 10.1111/j.1365-2796.2004.01388.x
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* 56, 303–308.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678. doi: 10.1016/j.neuron.2011.09.006
- Prasad, G., Joshi, S. H., Nir, T. M., Toga, A. W., Thompson, P. M., and Alzheimer's Disease Neuroimaging Initiative, (2015). Brain connectivity and novel network measures for Alzheimer's disease classification. *Neurobiol. Aging* 36(Suppl. 1), S121–S131. doi: 10.1016/j.neurobiolaging.2014.04.037
- Qiao, L., Zhang, H., Kim, M., Teng, S., Zhang, L., and Shen, D. (2016). Estimating functional brain networks by incorporating a modularity prior. *Neuroimage* 141, 399–407. doi: 10.1016/j.neuroimage.2016.07.058

- Quintana, M., Guardia, J., Sanchez-Benavides, G., Aguilar, M., Molinuevo, J. L., Robles, A., et al. (2012). Using artificial neural networks in clinical neuropsychology: high performance in mild cognitive impairment and Alzheimer's disease. *J. Clin. Exp. Neuropsychol.* 34, 195–208. doi: 10.1080/13803395.2011.630651
- Rabe-Jablonska, J., and Bienkiewicz, W. (1994). Anxiety disorders in the fourth edition of the classification of mental disorders prepared by the American Psychiatric Association: diagnostic and statistical manual of mental disorders (DMS-IV – options book). *Psychiatr. Pol.* 28, 255–268.
- Rolle, C. E., Anguera, J. A., Skinner, S. N., Voytek, B., and Gazzaley, A. (2017). Enhancing spatial attention and working memory in younger and older adults. *J. Cogn. Neurosci.* 29, 1483–1497. doi: 10.1162/jocn_a_01159
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Sarica, A., Cerasa, A., and Quattrone, A. (2017). Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Front. Aging Neurosci.* 9:329. doi: 10.3389/fnagi.2017.00329
- Sporns, O. (2011). The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125. doi: 10.1111/j.1749-6632.2010.05888.x
- Sporns, O., and Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics* 2, 145–162. doi: 10.1385/ni:2:2:145
- Suk, H. I., Lee, S. W., Shen, D., and Alzheimers Disease Neuroimaging, I. (2014). Subclass-based multi-task learning for Alzheimer's disease diagnosis. *Front. Aging Neurosci.* 6:168. doi: 10.3389/fnagi.2014.00168
- Talpos, J., and Shoaib, M. (2015). Executive function. *Handb. Exp. Pharmacol.* 228, 191–213. doi: 10.1007/978-3-319-16522-6_6
- Themistocleous, C., Eckerstrom, M., and Kokkinakis, D. (2018). Identification of mild cognitive impairment from speech in Swedish using deep sequential neural networks. *Front. Neurol.* 9:975. doi: 10.3389/fneur.2018.00975
- Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., and He, Y. (2015). GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. *Front. Hum. Neurosci.* 9:386. doi: 10.3389/fnhum.2015.00386
- Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., et al. (2013). Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biol. Psychiatry* 73, 472–481. doi: 10.1016/j.biopsych.2012.03.026
- Wee, C. Y., Yap, P. T., Zhang, D., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2012). Identification of MCI individuals using structural and functional connectivity networks. *Neuroimage* 59, 2045–2056. doi: 10.1016/j.neuroimage.2011.10.015
- Wee, C. Y., Yap, P. T., Zhang, D., Wang, L., and Shen, D. (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
- Wirth, M., Isaacowitz, D. M., and Kunzmann, U. (2017). Visual attention and emotional reactions to negative stimuli: the role of age and cognitive reappraisal. *Psychol. Aging* 32, 543–556. doi: 10.1037/pag0000188
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207. doi: 10.1016/j.neuroimage.2010.06.041
- Zhao, X., Liu, Y., Wang, X., Liu, B., Xi, Q., Guo, Q., et al. (2012). Disrupted small-world brain networks in moderate Alzheimer's disease: a resting-state fMRI study. *PLoS One* 7:e33540. doi: 10.1371/journal.pone.0033540
- Zhuo, Z., Mo, X., Ma, X., Han, Y., and Li, H. (2018). Identifying aMCI with functional connectivity network characteristics based on subtle AAL atlas. *Brain Res.* 1696, 81–90. doi: 10.1016/j.brainres.2018.04.042
- Zuo, X. N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F. X., Sporns, O., et al. (2012). Network centrality in the human functional connectome. *Cereb. Cortex* 22, 1862–1875. doi: 10.1093/cercor/bhr269

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

Copyright © 2020 Xu, Li, Mei, Tao, Wang, Zhao, Liang, Wu, Ding and Wang. 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.