Multiview Feature Learning With Multiatlas-Based Functional Connectivity Networks for MCI Diagnosis

Yu Zhang, *Senior Member, IEEE*, Han Zhang, Ehsan Adeli, *Member, IEEE*, Xiaobo Chen, *Member, IEEE*, Mingxia Liu, *Senior Member, IEEE*, and Dinggang Shen, *Fellow, IEEE*

***Abstract*—Functional connectivity (FC) networks built from resting-state functional magnetic resonance imaging (rs-fMRI) has shown promising results for the diagnosis of Alzheimer's disease and its prodromal stage, that is, mild cognitive impairment (MCI). FC is usually estimated as a temporal correlation of regional mean rs-fMRI signals between any pair of brain regions, and these regions are traditionally parcellated with a particular brain atlas. Most existing studies have adopted a predefined brain atlas for all subjects. However, the constructed FC networks inevitably ignore the potentially important subject-specific information, particularly, the subject-specific brain parcellation. Similar to the drawback of the "single view" (versus the "multiview" learning) in medical image-based classification, FC networks constructed based on a single atlas may not be sufficient to reveal the underlying complicated differences between normal controls and disease-affected patients due to the potential bias from that particular atlas. In this study, we propose a multiview feature learning method with multiatlas-based FC networks to improve MCI diagnosis. Specifically, a three-step transformation is implemented to generate multiple individually specified atlases from the standard automated anatomical labeling template, from which a set of atlas exemplars is selected. Multiple FC networks are constructed based on these preselected atlas exemplars, providing multiple views of the FC network-based feature representations for each subject. We then devise a multitask learning algorithm for joint feature selection from the constructed multiple FC networks. The selected features are jointly fed into a support vector machine classifier for multiatlas-based MCI diagnosis. Extensive experimental comparisons are carried out between the proposed method and other competing approaches, including the traditional single-atlas-based method. The results indicate that our method significantly improves the MCI classification, demonstrating its promise in the brain connectome-based individualized diagnosis of brain diseases.**

***Index Terms*—Brain disease diagnosis, functional connectivity (FC), mild cognitive impairment (MCI), multinetwork classification, resting-state functional magnetic resonance imaging (rs-fMRI).**

Manuscript received 27 March 2020; revised 6 July 2020; accepted 5 August 2020. Date of publication 11 December 2020; date of current version 1 July 2022. This article was recommended by Associate Editor Q. Shen. *(Corresponding authors: Yu Zhang; Dinggang Shen.)*

Yu Zhang is with the Department of Bioengineering, Lehigh University, Bethlehem, PA 18015 USA (e-mail: yuzhang@lehigh.edu).

Han Zhang and Mingxia Liu are with the Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA.

Ehsan Adeli is with the Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA 94305 USA.

Xiaobo Chen is with the Automotive Engineering Research Institute, Jiangsu University, Zhenjiang 212013, China.

Dinggang Shen is with the Department of Research and Development, Shanghai United Imaging Intelligence Company, Ltd., Shanghai 200232, China, and also with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea (e-mail: dinggang.shen@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCYB.2020.3016953.

Digital Object Identifier 10.1109/TCYB.2020.3016953

I. Introduction

AS A TYPICAL neurological disorder, Alzheimer's disease (AD) is characterized by progressive perceptive deficits without effective treatment so far [1]. This disease is irreversible and often becomes more severe over time due to neurofibrillary tangles, neurotoxicity, neurodegeneration, and so on, which could ultimately cause death [2]. Hence, timely and accurate diagnosis of AD patients has unquestionable importance so that proper treatments can be applied to possibly slow down or reverse the degeneration process. As an intermediate stage between normal aging and AD, mild cognitive impairment (MCI) is known to be associated with the increased risk of developing AD [3]. Evidence shows that more than half of the MCI subjects with memory loss or other preclinical symptoms will progress to AD within five years [3], [4]. Such a high conversion rate could be reduced if timely interventions were applied to the MCI subjects after early detection. Therefore, identification of MCI subjects is of great clinical importance for effective treatment and possible prevention of AD progression. Accurate MCI diagnosis, however, is considerably challenging due to the subtle anatomical and functional changes during the early stages of AD [5].

In the past few years, various neuroimaging techniques, including structural magnetic resonance imaging (MRI) [6]; diffusion tensor imaging (DTI) [7]; functional MRI [8], [9]; and positron emission tomography (PET) [10], have been widely applied to investigate the pathological biomarkers related to the early detection of AD. Among them, resting-state fMRI (rs-fMRI), which measures the changes of blood oxygen level-dependent (BOLD) signals associated with spontaneous brain activity [11], has been increasingly adopted to explore the functional alterations in the large-scale brain functional connectome [i.e., functional connectivity network (FCN)] in MCI subjects for the early diagnosis of AD, because

of its increased sensitivity to the pathological changes in AD [12], [13]. To construct FCN based on rs-fMRI, the most widely used strategy is to calculate pairwise functional connectivities (FCs) as edges based on a set of predefined regions of interest (ROIs, or brain regions) as nodes by the brain parcellation with a specific atlas [7], [14], [15]. Most brain atlases are defined in a standard space [Montreal Neurological Institute (MNI) space]; therefore, the rs-fMRI data from each subject need to be first registered to the standard space so that the predefined standard atlas can be applied to each subject to obtain uniform brain parcellations for ROIs. ROI-averaged BOLD signals are then extracted to represent regional brain activities for the temporal synchronization-based FC calculations [16], [17] with various metrics, such as Pearson's correlation [18]; partial correlation [19], [20]; or high-order FC [5], [21], [22]. As indicated above, accurate construction of the brain FCN not only depends on the FC metrics chosen but also relies on accurate brain parcellation, or atlas, which can dramatically affect accuracy in the subsequent MCI classification [23].

So far, extensive efforts have been dedicated to developing various sophisticated FC measurements and the subsequent FCN-related features, such as the brain network properties extracted by graph-theoretical analysis [5], [15], [24]. However, the methodological development of appropriate brain parcellation has not been comprehensively explored and, to date, considerably ignored. Most existing studies adopt a common atlas, such as the automated anatomical labeling (AAL) template [25], to parcellate each subject's brain in a uniform manner [26], [27]. A potential issue of using a single atlas for all subjects is that it assumes a highly consistent functional border of every brain region across all subjects. However, such an assumption may not hold in practice, due to not only imperfect spatial registration [28], [29] but also mismatched anatomical-functional correspondence [30]. In fact, evidence has suggested that different subjects could have different functional parcellation, with the varied functional border of brain regions [30]. In other words, a common atlas that simply ignores individual differences in the brain functional areas could affect the FC and the subsequent brain functional connectome modeling, jeopardizing personalized brain disease diagnosis [31]. For example, a not well-defined group-shared AAL template could functionally blur the FC differences between the MCI and cognitive normal subjects, making it difficult to conduct individualized classification. Such a problem could become even more serious in the MCI study, as the FCN changes in the MCI stage could be subtle [32]. Generating individualized functional atlas for every subject could be an alternative way to accurately modeling normative or MCI-like FCNs [31], but can be error prone due to the heavy noise in rs-fMRI data, or methodologically less efficient when the number of ROIs is big, let alone the difficulty in searching for ROI correspondence across all subjects.

As an emerging direction in the machine learning and pattern recognition fields is multiview learning [33]–[35], which has been well adapted to take advantage of the features from multiple views to jointly and comprehensively represent an object [36], [37]. In multiview learning, each view captures a different set of features. For instance, color and texture features provide complementary information from different views for image recognition [38], [39]. Multiview learning aims at learning multiple representations of information jointly to improve the generalization performance of the model recognition [40], [41]. In the past decade, multiview learning has been applied to structural MRI studies by exploiting multiple anatomical templates, which achieved improved brain disease diagnosis. For instance, Koikkalainen *et al.* [42] developed a multiatlas-based strategy to reduce the MRI registration bias for AD classification and obtained increased accuracy in comparison with a single-atlas-based method. Min *et al.* [29] proposed to generate multiple feature representations of the anatomical structures based on multiple atlases. The derived features were then aggregated to form a comprehensive representation for enhancing the separability between AD and healthy subjects. To better explore multiatlas feature representations, Liu *et al.* [28], [43] developed a multitask learning strategy to jointly optimize anatomical features extracted from multiple MRI templates for improved AD/MCI classification. Inspired by these previous studies, for studying brain functional abnormality, we consider that the FCNs constructed based on multiple atlases can be an analog to the multiple views for better characterizing the functional connectome of the same brain. With each brain atlas as a specific view of functional parcellations, multiview feature representation can be achieved by generating multiple FCNs based on multiple individualized atlases, thus better representing a unique subject and improving the performance of subsequent disease diagnosis. However, to the best of our knowledge, previous studies have not investigated multiatlas-based FCN construction and feature learning for studying the abnormal functional architecture of the human brain, especially for MCI diagnosis.

Accordingly, in this study, we propose a multiview feature learning method with multiatlas-based FCNs for improving MCI diagnosis. We first generate an individualized atlas by implementing a three-step transformation-based method on the standard AAL template. Specifically, a nonlinear registration is first implemented to transform each subject's rs-fMRI data into the standard MNI space. A deformation field is then derived to characterize the nonlinear mapping between the subject's native space and the standard space, with which we can warp the AAL template into the individual space by performing the inverse transformation. We then implement affinity propagation (AP) clustering [44] on the affine-transformed individual-space AALs (that are roughly in the same space but preserving a certain degree of anatomical variability) to generate a small set of representative individualized atlas exemplars. These exemplar atlases are selected for the extraction of regional mean time series and the construction of multiple FCNs for each subject. A multitask learning algorithm is designed for the optimization of multiview features that are finally integrated and used to train a support vector machine (SVM) for MCI versus normal control (NC) diagnosis. Extensive experimental comparisons are carried out between our proposed method and other competing approaches (using a single-atlas-based method or simply concatenating

multiview features). The experimental results indicate that our method achieves superior MCI classification performance.

The main contributions of this study can be summarized as follows.

1) An efficient method is proposed for representative multiview atlas generation.
2) A multinetwork classification framework is developed by exploiting supplementary discriminant features based on multiple individualized atlases for improved MCI diagnosis.
3) A novel relation-guided sparse group learning (SGL) algorithm is devised for joint feature selection from the multiple networks.

The remainder of this article is organized as follows. After describing data acquisition and preprocessing steps, we introduce the proposed MCI diagnosis framework in Section II. Then, we present the experiments and comparative results in Section III. In Section IV, we consider the influence of parameter choices, analyze the possible association of the most discriminative brain regions with the AD pathological attacks, and discuss the potential extensions of our method for future studies. Finally, we summarize this study in Section V.

II. Materials and Methods

A. Differences Between Our Method and Existing Approaches

So far, many multiatlas methods have been developed for improved brain disease diagnosis by exploiting multiple anatomical templates [28], [29], [42], [43]. However, all these existing methods were specifically designed for structural MRI analysis. Different from them, our proposed method was tailored for functional MRI. In particular, the newly developed three-step transformation strategy provided an effective way to generate individualized brain parcellations for multiview brain network analysis with functional MRI. The designed strategy is fundamentally different from the existing methods that are based on nonlinear registration with multiple atlases for structural MRI feature extraction. On the other hand, we further devised an SGL-based feature optimization algorithm under multiview feature learning framework. In recent years, SGL has been increasingly applied to neuroimaging data analysis for various applications. By modifying the standard sparse group model with a view-centralized regularization, a multiatlas classification method [28] has been developed for structural MRI-based AD diagnosis. The view-centralized multiatlas feature selection was able to identify the informative features from a main atlas with extra guidance from other atlases. By treating each time point as a task, a multitask learning method based on fused sparse group lasso [45] was proposed to identify pathological biomarkers for the AD progression measured by cognitive scores from multiple time points. The fused learning model incorporated the temporal smoothness for the effective selection of important biomarkers from different time points. In addition, a sparse group representation strategy [46] was recently developed for generic classification of electroencephalogram patterns in brain–computer interface applications. By extending sparse representation classification to a group-level analysis, the strategy was able to identify task-related neural patterns from group representation coefficients corresponding to a composite subject dictionary matrix. Importantly, different from these studies, our multiview feature learning model integrated the between-group sparsity, within-group sparsity, and intersubject relationship constraints. The designed model cannot only select informative atlases but also discriminative features within each atlas under the information guidance of intrinsic relationships among subjects in each of the atlas exemplars. None of the existing SGL-based methods is able to simultaneously achieve these critical functions. Our extensive experimental study demonstrated that integrating these critical functions achieved promising diagnosis performance of MCI.

B. Data Description and Preprocessing

The Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (http://adni.loni.usc.edu/) is used for validation of the proposed method in MCI diagnosis. The goal of ADNI was to identify biomarkers for early diagnosis of AD and their use in clinical trials. A total of 117 subjects (66 MCI patients and 51 NC subjects) are selected from ADNI-2 for our experimental study. All subjects were scanned using 3.0T Philips scanners and are age- and gender-matched between the two classes. The rs-fMRI data are preprocessed using the SPM8 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) according to the standard pipeline. In particular, we remove the first three volumes of each subject before preprocessing for magnetization equilibrium. The rigid-body registration is then implemented to correct head motion. We set a cutoff for the level of head motion and exclude subjects with large head motion (i.e., larger than 2 mm or 2°). The rs-fMRI data are normalized to the standard MNI space followed by spatial smoothing using a Gaussian kernel with full-width-at-half-maximum of $6\times6\times6$ mm^3. Subjects with framewise displacement for larger than 2.5-min data are further excluded in data screening.

C. Multiple Atlases Generation

In our study, we propose a multiview feature extraction method using multiatlas-based FCNs. One of the key steps in our proposed framework is to generate an appropriate set of atlas exemplars that are not only representative enough to cover the entire population for achieving better generalization capability but also able to accurately capture the discriminative biomarkers associated to MCI. To this end, we develop a study-specific atlas selection strategy to obtain multiple representative atlas exemplars from the data. A three-step transformation strategy was developed to generate an individualized atlas for each subject. An unsupervised clustering was then conducted to determine representative atlas exemplars for the subsequent FCN construction and analysis.

The flowchart of the three-step transformation strategy is illustrated in Fig. 1. A nonlinear registration with echo-planar imaging (EPI) template is implemented to transform the temporally averaged rs-fMRI data into the standard MNI space. A deformation field is derived from the transformation to characterize the nonlinear mapping relationship between the native

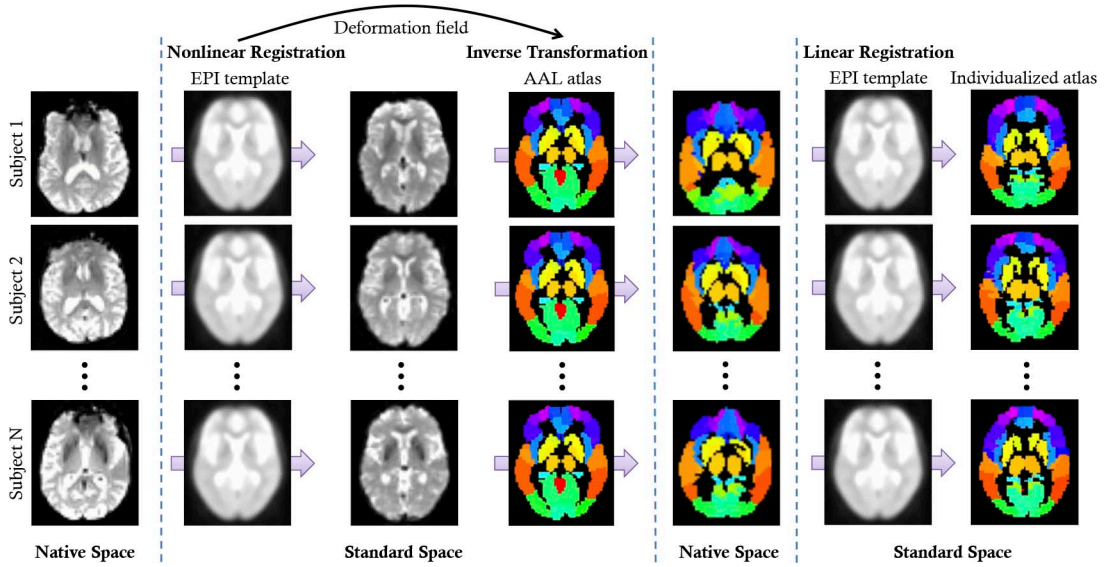


Fig. 1. Illustration of three-step transformation for the generation of individualized atlases. A nonlinear registration was first implemented with the EPI template to transform each subject's rs-fMRI data into the standard MNI space. A deformation field was generated to characterize the nonlinear mapping between the native space and the standard space. The standard AAL atlas was then warped into the native space for each subject by performing an inverse transformation with the deformation field. A linear registration was further implemented on these native-space atlases, which transformed them back to the standard space such that the derived individualized atlases become more comparable but still retained sufficient individual differences.

space and the standard space. Then, we warp the AAL atlas into the native space for each subject by performing the inverse transformation with the deformation field. A linear registration is further implemented on these native-space atlases to transform them back to the standard space such that the derived individualized atlases become more comparable but still retain sufficient individual differences, which can provide supplementary feature representations to each other. Note that these transformations are only used to generate individualized atlas without introducing additional noises to the rs-fMRI data.

To select the representative atlases, we perform a clustering analysis using the AP algorithm [44] on all the individualized atlases. In the AP algorithm, similarity is measured by the normalized mutual information [47] and the appropriate preference value is determined by a bisection method [44]. The AP algorithm automatically identifies nine atlas exemplars that define the corresponding clusters. Finally, combining with the standard AAL atlas, a total of ten atlas exemplars (one AAL atlas + 9 atlas exemplars, as shown in Fig. 2) are used to construct multiple FCNs. In Fig. 2, exemplars 1–5 are from NC subjects, while exemplars 6–9 are from MCI subjects. To further investigate the representativeness and interdifferences of the generated atlas exemplars, we applied *t*-SNE [48] to reduce all individualized atlases into a 2-D distribution. Fig. 3 visualizes the distribution of these individualized atlases belonging to nine different clusters. The *t*-SNE distribution difference indicates the interdifferences of clusters and confirms the representativeness of the atlas exemplars.

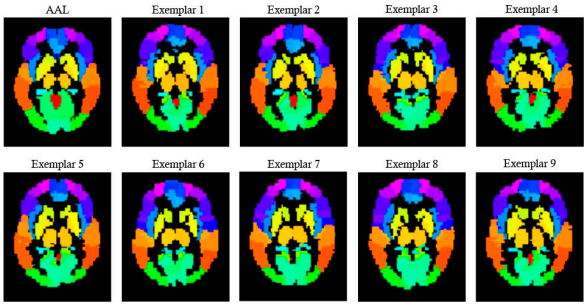


Fig. 2. AAL atlas and nine atlas exemplars selected using the AP clustering algorithm.

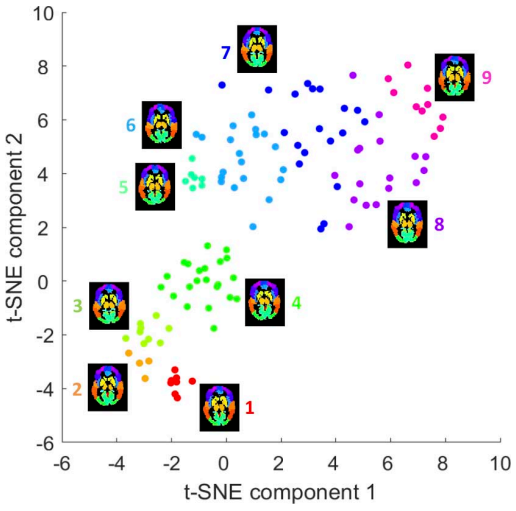


Fig. 3. Visualization of the *t*-SNE distribution of the individualized atlases and the nine atlas exemplars.

D. FCN Construction and Feature Extraction

According to each selected atlas, the preprocessed rs-fMRI data can be parcellated into 116 ROIs. Regional mean rs-fMRI time series of each ROI are computed and band-pass filtered between 0.015 and 0.15 Hz. An FCN is then constructed by computing Pearson's correlation between all pairs of ROIs for each atlas. As a result, we obtain ten FCNs for each subject. Fig. 4 depicts the FCNs constructed by ten atlases. It can be

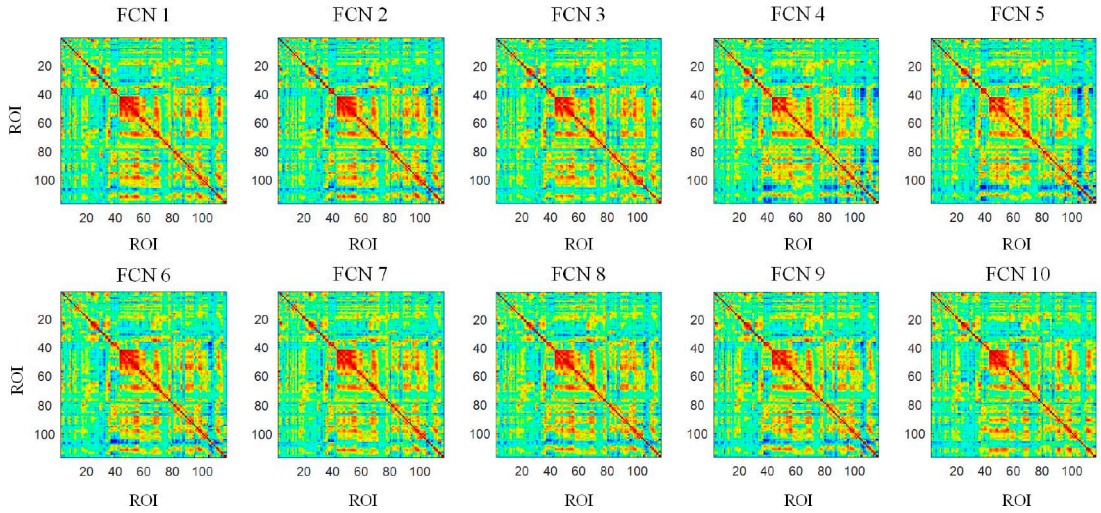


Fig. 4. FCNs constructed based on the selected atlas exemplars shown in Fig. 2.

seen that the constructed multiple FCNs present overall similar structures but locally different connectivity strength, which can provide supplementary discriminative information.

For each of the constructed FCNs, features are extracted using graph-theory-based method [14], [23]. In this study, the weighted local clustering coefficient (WLCC) [5] is computed for each of the constructed FCNs. The WLCC characterizes the local topology of each node in a weighted network and provides informative nodewise features for disease classification [15]. Assume C_{ij}^k denotes the FC strength between nodes i and j of the constructed FCN based on the kth atlas for a certain subject. The WLCC for node i is defined as

$$x_i^k = \frac{2\sum_{j:j\in\Theta}(C_{ij}^k)^{\frac{1}{3}}}{V_i(V_i-1)} \tag{1}$$

where Θ_i denotes a set of nodes directly connected to the ith node and V_i is the number of elements in Θ_i. As a result, a feature vector is formed by computing the WLCC for all brain regions $x^k = [x_1^k, x_2^k, \ldots, x_R^k]^T$ $(R = 116)$.

E. Feature Selection and Classification

Although we select the most representative atlases for FCN construction and perform the subsequent graphical feature extraction, some of the derived features may still be redundant or irrelevant for subsequent classification. To optimize the extracted features, we devise a multiview feature selection algorithm under the SGL framework [49] by treating each atlas as a specific view. As a result, the complementary discriminant features could be jointly exploited from multiple FCNs to provide improved MCI classification performance. Below, we provide detailed explanations of our proposed multiview feature selection algorithm.

Suppose that $\mathbf{X}^k = [\mathbf{x}_1^k, \ldots, \mathbf{x}_n^k, \ldots, \mathbf{x}_N^k]^T \in \mathbb{R}^{N\times R}$ contains the feature vectors, where x_n^k denotes the feature vector extracted from the FCN of the nth subject based on the kth atlas, and $\mathbf{y} = [y_1, \ldots, y_n, \ldots, y_N]^T \in \mathbb{R}^N$ denotes the class label vector, where $y_n \in \{-1, 1\}$ indicating NC or patient. An SGL-based feature learning algorithm [49] can be formulated as

$$\mathbf{w} = \arg\min_{\mathbf{w}} \ \frac{1}{2}\left\|\mathbf{y} - \sum_{k=1}^{K}\mathbf{X}^k\mathbf{w}^k\right\|_2^2 + \lambda_1\|\mathbf{w}\|_1 + \lambda_2\sum_{k=1}^{K}\left\|\mathbf{w}^k\right\|_2 \tag{2}$$

where $\mathbf{w}^k$ denotes the weight vector for the features corresponding to the kth view (i.e., kth atlas) and $\mathbf{w} = [\mathbf{w}^1, \mathbf{w}^2, \ldots, \mathbf{w}^K]$. The second term $\|\mathbf{w}\|_1$ is an l_1-norm regularization that is used to enforces some elements of $\mathbf{w}$ to be zero, and the third one $\sum_{k=1}^K \|\mathbf{w}^k\|_2$ denotes the group sparse regularization that is used to control the between-group sparsity of $\mathbf{w}$. The hyperparameters λ_1 and λ_2 are used to balance a tradeoff between the two regularization terms. Hence, SGL provides a good means to determine the most important atlases and the most informative features from these atlases. However, the model in (2) ignores the intrinsic relationship among subjects in each of the exemplar atlases, which has been suggested to be a valuable prior to guide multiview feature selection [43]. To further enhance the performance of the learned model, we propose to penalize the difference between linear transformations $(\mathbf{x}_i^k)^T\mathbf{w}^k$ and $(\mathbf{x}_j^k)^T\mathbf{w}^k$ if features of the two subjects $\mathbf{x}_i^k$ and $\mathbf{x}_j^k$ extracted in the kth view are very similar. To this end, we construct a graph Laplacian regularization term

$$\begin{aligned}\Omega &= \sum_{k=1}^{K}\sum_{i,j=1}^{N} s_{i,j}^k\left(\left(\mathbf{x}_i^k\right)^T\mathbf{w}^k - \left(\mathbf{x}_j^k\right)^T\mathbf{w}^k\right)^2 \\ &= \sum_{k=1}^{K}\left(\mathbf{X}^k\mathbf{w}^k\right)^T\mathbf{L}^k\left(\mathbf{X}^k\mathbf{w}^k\right)\end{aligned} \tag{3}$$

where $\mathbf{S}^k = [s_{i,j}^k] \in \mathbb{R}^{N\times N}$ denotes a similarity matrix with the element $s_{i,j}^k = \exp(-\|\mathbf{x}_i^k - \mathbf{x}_j^k\|_2^2)$ measuring the pairwise similarity of subjects in the kth view. Here, $\mathbf{L}^k = \mathbf{D}^k - \mathbf{S}^k$ is the Laplacian matrix for view k, and $\mathbf{D}^k \in \mathbb{R}^{N\times N}$ is a diagonal matrix with its diagonal elements defined as $d_{i,i}^k = \sum_j s_{i,j}^k$. By integrating the above-mentioned regularization term Ω into the SGL model, our newly proposed feature selection algorithm,

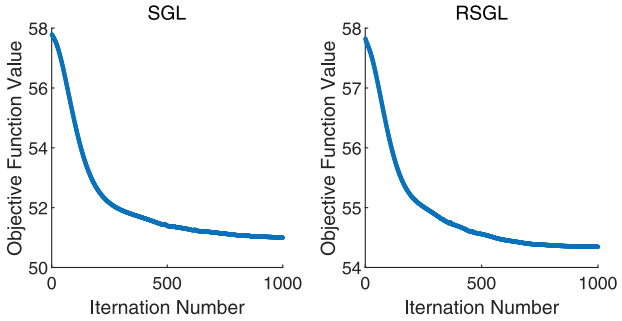


Fig. 5. Convergence curves of the proposed SGL and RSGL algorithms.

called relationship-guided SGL (RSGL), can be formulated as

$$\mathbf{w} = \arg\min_{\mathbf{w}} \frac{1}{2}\left\|\mathbf{y} - \sum_{k=1}^{K}\mathbf{X}^k\mathbf{w}^k\right\|_2^2 + \lambda_1\|\mathbf{w}\|_1 + \lambda_2\sum_{k=1}^{K}\left\|\mathbf{w}^k\right\|_2 + \lambda_3\sum_{k=1}^{K}\left(\mathbf{X}^k\mathbf{w}^k\right)^T\mathbf{L}^k\left(\mathbf{X}^k\mathbf{w}^k\right) \quad (4)$$

where λ_3 is the hyperparameter used to control the contributions of the intersubject relationship. Although the optimization problem in the above equation is convex, it is challenging to solve due to the second and third nonsmooth regularization terms. In this study, we adopt the accelerated proximal gradient (APG) method [50] to find the optimal solution. The optimization problem can be separated into a smooth part

$$f(\mathbf{w}) = \frac{1}{2}\left\|\mathbf{y} - \sum_{k=1}^{K}\mathbf{X}^k\mathbf{w}^k\right\|_2^2 + \lambda_3\sum_{k=1}^{K}\left(\mathbf{X}^k\mathbf{w}^k\right)^T\mathbf{L}^k\left(\mathbf{X}^k\mathbf{w}^k\right) \quad (5)$$

and a nonsmooth part

$$\mathcal{L}(\mathbf{w}) = \lambda_1\|\mathbf{w}\|_1 + \lambda_2\sum_{k=1}^{K}\left\|\mathbf{w}^k\right\|_2. \quad (6)$$

To approximate the composite function $\mathbf{f}(\mathbf{w}) + \mathcal{L}(\mathbf{w})$, we construct the following function:

$$\mathcal{G}(\mathbf{w}, \mathbf{w}(t)) = f(\mathbf{w}(t)) + < \mathbf{w} - \mathbf{w}(t), \bigtriangledown f(\mathbf{w}(t)) > + \frac{\mu}{2}\|\mathbf{w} - \mathbf{w}(t)\|_2^2 + \mathcal{L}(\mathbf{w}) \quad (7)$$

where $\bigtriangledown f(\mathbf{w}(t))$ denotes the gradient at the point $\mathbf{w}(t)$ and μ is the step size determined by a line search [51]. The iterative update rule of the proximal gradient method is given by

$$\mathbf{w}(t+1) = \arg\min_{\mathbf{w}} \frac{1}{2}\|\mathbf{w} - \mathbf{v}(t)\|_2^2 + \frac{1}{\mu}\mathcal{L}(\mathbf{w}) \quad (8)$$

where $\mathbf{v}(t) = \mathbf{w}(t) - \bigtriangledown f(\mathbf{w}(t))/\mu$ and $\mathbf{w}(t)$ is the search point obtained at the t-th iteration. For a fixed maximum iteration T, the APG algorithm is able to achieve a convergence rate of $\mathcal{O}(1/T^2)$ [52]. To confirm the convergence of our proposed method, we show the change of the objective function values in Fig. 5. For both the SGL and RSGL algorithms, the values of objective function decrease rapidly and achieve convergence within 600 iterations.

For each atlas exemplar, feature selection is conducted by excluding the noninformative features whose weights in the corresponding weight vector are zeros. The feature subsets selected based on all the atlas exemplars are then concatenated and used to train a linear SVM classifier for MCI diagnosis (i.e., MCI versus NC classification). Fig. 6 illustrates the MCI diagnosis framework based on our proposed multinetwork classification algorithm.

III. Experimental Study

A. MCI Classification and Performance Evaluation

To validate the effectiveness of our proposed multinetwork method for MCI diagnosis, we carried out extensive comparisons among the following six feature selection algorithms: 1) t-test_single and 2) Lasso_single: FCN is constructed using a single AAL atlas with t-test and Lasso for feature selection, respectively; 3) t-test_multi, 4) Lasso_multi, 5) SGL_multi, and 6) RSGL_multi: multiple FCNs are constructed using the selected atlases exemplars with t-test, Lasso, SGL, and RSGL for feature selection, respectively.

We evaluate results of MCI diagnosis based on various performance indices, including classification accuracy (ACC), area under the ROC curve (AUC), sensitivity (SEN), and specificity (SPE). ACC is defined as the ratio of the number of correctly predicted labels to the number of entire samples. AUC measures the probability. Also, SEN and SPE are calculated based on the true positive rate and the false-positive rate, respectively

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \quad (9)$$

where TP, TN, and FP are the true positive rate, true negative rate, and false-positive rate, respectively. The leave-one-out cross-validation (LOOCV) is implemented to assess the diagnosis performance of each compared method. It should be noted that LOOCV is performed on all procedures in our proposed diagnosis framework. That is, both the atlas generation and feature selection are implemented on the training data only and then applied to the test data for performance evaluation. In each fold of LOOCV, an additional inner LOOCV is also carried out on the training data to select the optimal hyperparameters.

B. Results

Fig. 7 depicts the classification results derived by different algorithms. Compared with the single-atlas-based methods, all the multiatlas-based methods improved the classification performance in varied degrees. Fig. 8 further shows the ROC curves derived by the comparison methods. To investigate the significance of performance difference between a pair of competing methods, we performed a nonparametric statistical analysis, namely, DeLong's test [53], for the comparison of each pair of ROC curves calculated on the dataset, with a confidence interval of 95%. It can be seen that all multinetwork-based methods significantly outperformed the single-network-based methods. More important, by incorporating a relation-guided SGL feature selection algorithm into the multinetwork diagnosis framework, our proposed method (i.e., RSGL_multi) achieved the best classification accuracy 85.5%,

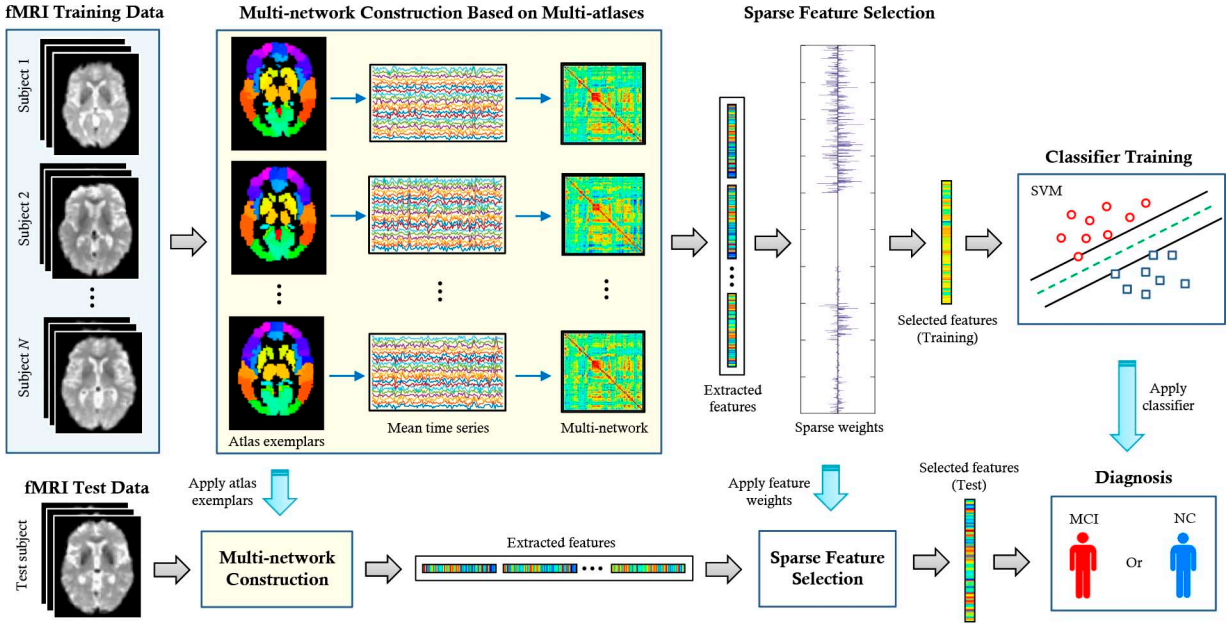


Fig. 6. Illustration of the multiview learning-based multinetwork classification framework for MCI diagnosis.

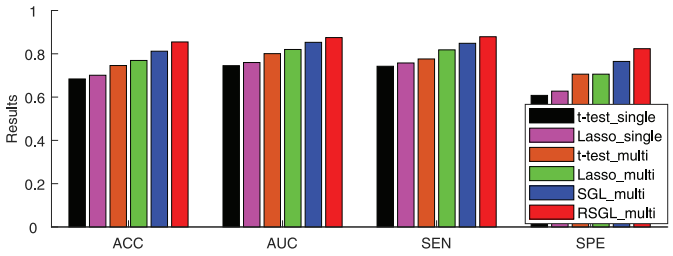


Fig. 7. Classification performance comparison among different methods.

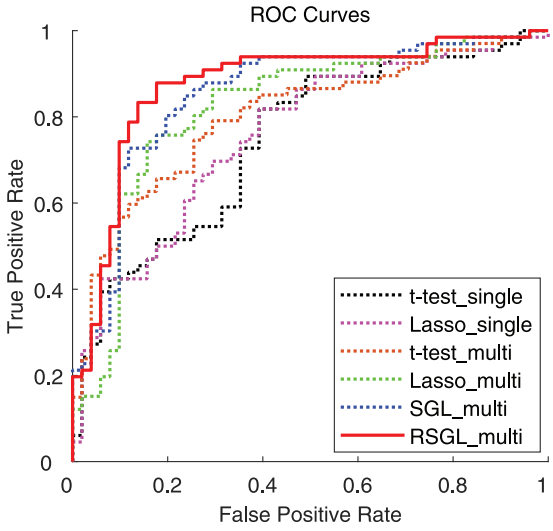


Fig. 8. ROC curves derived by different methods for MCI classification.

significantly higher than all the other methods, with p values of 0.0033, 0.0078, 0.0094, 0.013, and 0.025, respectively.

IV. Discussion

A. Parameter Sensitivity

Note that the performance of our proposed method is affected by the selection of hyperparameters, that is, λ_1 for the within-group sparsity, λ_2 for between-group sparsity, and λ_3 for intersubject relationship. Our experimental study used a grid search with an inner LOOCV to determine the optimal parameter values on the training data. To investigate the parameter sensitivity of our proposed feature selection method, we evaluated effects of varying values of these three hyperparameters on classification accuracy using LOOCV with all subjects. Fig. 9 shows the classification accuracies obtained using the multinetwork features selected by our proposed RSGL method with different settings for the aforementioned hyperparameters. The best accuracy of 88.9% is achieved by using $\lambda_1 = 0.05$, $\lambda_2 = 0.09$, and $\lambda_3 = 0.08$. This indicates that our method with parameters estimated from an inner LOOCV is able to achieve an accuracy of 85.5% that is comparable to the best accuracy obtained using parameters estimated with all subjects.

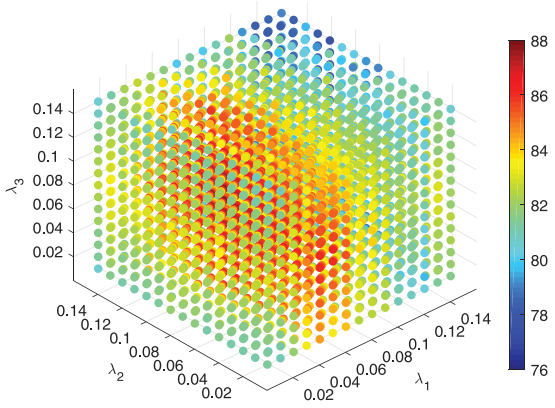


Fig. 9. Classification accuracy (%) derived based on multinetwork features selected by RSGL with different values of three hyperparameters (i.e., λ_1, λ_2, and λ_3). Each hyperparameter range is [0.01, 0.02, ..., 0.15]. The results are obtained using LOOCV on all subjects. The highest accuracy is 88.9% when $\lambda_1 = 0.05$, $\lambda_2 = 0.09$, and $\lambda_3 = 0.08$.

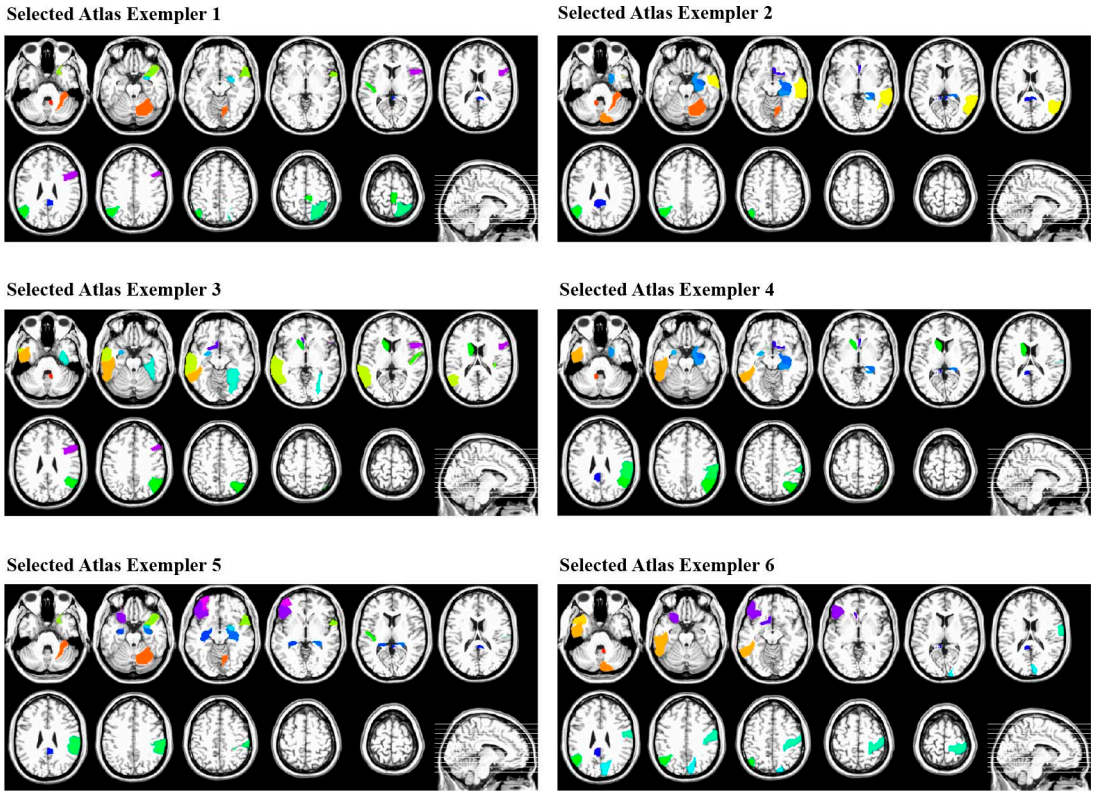


Fig. 10. The top ten discriminative ROIs for each of the six selected atlas exemplars by our proposed RSGL feature selection model. In each subplot, there are ten different colors that indicate the top ten discriminative ROIs for the corresponding atlas exemplar.

B. Most Discriminative Brain Regions

To interpret the identified biomarkers, we investigated the discriminative brain regions selected based on multinetwork features in the MCI diagnosis. In our experiment, the model weights of the trained SVM classifier reflected the importance of selected features for accurate MCI classification. Only six atlases were consistently selected by our RSGL algorithm across all the LOOCV runs. For each of these six selected atlases, we averaged SVM model weights for each feature over all the LOOCV folds during the training process. The top ten discriminative ROIs for each of the six selected atlases are shown in Fig. 10. Although some ROIs were consistently selected for multiple atlases, many nonoverlapped ROIs were selected from different atlases, which provided complementary discriminant information for achieving higher MCI diagnosis performance.

A total of 32 ROIs were indicated as the most discriminative brain regions. Their indices and names are summarized in Table I. Most of these discriminative brain regions have been reported in previous AD studies. In particular, many selected regions, including PCC, angular gyrus, hippocampus, and supramarginal gyrus, are involved in the default mode network (DMN), which is believed to be closely related to high-level cognitive functioning, such as episodic memory [54]. Dysfunction of the DMN was also observed across a range of neurological disorders, including AD and MCI [55]. It has been reported that abnormal structural, functional, and metabolic changes occur in the posterior cingulate gyrus of individuals with MCI [56], [57], which may be closely associated with the deficits in memory functions, object recognition, or evaluation of information [58]. The middle temporal gyrus has been reported as a significant biomarker for identifying AD from NC subjects [59]. Prominent atrophy has been observed on the left lateral temporal lobe, especially on the middle temporal gyrus, for the MCI individuals who converted to AD in contrast to those stable MCI ones [60]. Extensive research found that the hippocampus was sensitive to the pathological attack in the early stage of AD [61], [62]. Compared with NC subjects, MCI individuals have been also found to show decreased centrality in the left angular gyrus that is an important part of DMN and is responsible for complex language functions, especially the language comprehension [63].

In addition, the above-mentioned DMN regions, most of the other selected brain regions have also proven to be important for early AD diagnosis. The olfactory cortex of AD patients presented decreased network wiring efficiency [64], which is a sensitive and early behavioral marker associated with olfactory dysfunction in neurodegenerative diseases [65]. The left orbitofrontal cortex was found to be associated with value assignment and deteriorated motivation that are well-known clinical AD impairment [66]. The inferior temporal gyrus is closely related to high-level brain functions [67]. As a multimodal association region, the inferior temporal gyrus was affected in the early stage of AD [68]. Several other selected regions were from the cerebellar cortex, including the right crus II, left lobule IX, right lobule IX, and right lobule VI,

TABLE I
NAMES OF THE MOST DISCRIMINATIVE ROIS SELECTED BY OUR METHOD

Index	ROI Name	Index	ROI Name
9	Left orbitofrontal cortex (middle)	12	Right inferior frontal gyrus
15	Left inferior frontal gyrus	21	Left olfactory cortex
22	Right olfactory cortex	35	Left posterior cingulate gyrus
36	Right posterior cingulate gyrus	37	Left hippocampus
38	Right hippocampus	40	Right parahippocampal gyrus
41	Left amygdala	42	Right amygdala
46	Right cuneus	56	Right fusiform gyrus
58	Right postcentral gyrus	60	Right superior parietal lobule
64	Right supramarginal gyrus	65	Left angular gyrus
66	Right angular gyrus	70	Right paracentral lobule
71	Left caudate nucleus	79	Left transverse temporal gyrus
80	Right transverse temporal gyrus	84	Right superior temporal pole
85	Left middle temporal gyrus	86	Right middle temporal gyrus
87	Left middle temporal pole	89	Left inferior temporal gyrus
94	Right crus II of cerebellar hemisphere	100	Right lobule VI of cerebellar hemisphere
105	Left lobule IX of cerebellar hemisphere	106	Right lobule IX of cerebellar hemisphere

and may also be influenced the early functional abnormality of AD [69].

C. Effect of the Number of Atlases

We also investigated the effect of varying the number of atlases on the classification performance of our proposed multinetwork-based diagnosis framework. Fig. 11 depicts the classification accuracies obtained by different feature selection algorithms with various numbers of atlas exemplars, respectively. By incorporating the manifold regularization (i.e., intersubject relationship guidance) into the SGL model, our proposed method exploited more supplementary discriminant features from multiple FCNs and thus further improved classification accuracy. For all of the four feature selection algorithms, the classification accuracy rises gradually with the increase of atlas number, and becomes relatively stable when using more than ten atlases. This indicates that the choice of ten atlas exemplars in our experimental study is an appropriate compromise between classification accuracy and computational efficiency.

D. Computational Cost

To assess the computational efficiency of the developed algorithm, we further implemented an experimental comparison of computational time among these compared methods (see Fig. 12). The computational time was evaluated for conducting one run of the cross-validation with MATLAB R2018a on a computer (3.41 GHz 8-core CPU, i7-6700, 64-GB RAM). The multiatlas analysis took longer computational time than single-atlas-based methods mainly due to the generation of individualized atlases. Compared with the simple t-test or Lasso-based feature selection, our multiview learning algorithms cost additional time for hyperparameter determination with inner-loop cross-validation. However, since both the generation of individualized atlases and the determination of optimal hyperparameters were parallelly implemented in our experiment, the increased computational cost was acceptable

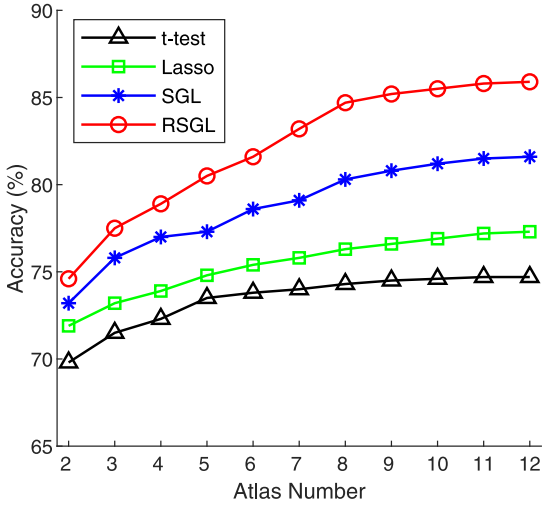


Fig. 11. MCI classification accuracies derived by multinetwork-based methods using different feature selection algorithms with various numbers of atlas exemplars.

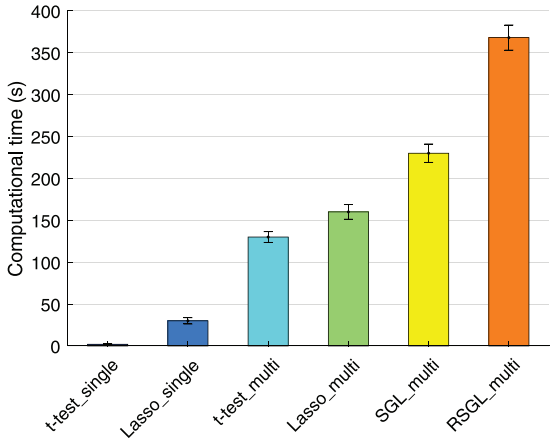


Fig. 12. Comparison of computational time cost by different methods.

given the significantly improved classification performance (see Figs. 7 and 8). The computational cost could be further reduced with a more powerful computer (e.g., a high-performance GPU). It should be noted that both the atlas generation and hyperparameter determination are only needed for the training phase. The diagnosis of new patients can be efficiently implemented with our developed algorithm.

E. Extensions

In this study, we constructed FCNs based on Pearson's correlation. In spite of its biological intuitiveness, Pearson's correlation only simply models the pairwise linear relationship between brain regions (also called low-order correlation [21]), ignoring the complex interaction among multiple brain regions. On the contrary, several other partial correlation-based methods, such as graphical Lasso [70] and sparse representation [71], are designed to model FCN by regressing out the potential confounding variables and considering the effects of multiple regions. On the other hand, high-order FC modeling methods [5], [21] have recently arisen, which are able to characterize the high-level interregional

interactions by quantifying how the low-order correlations between different pairs of brain regions interact with each other. The high-order FCNs have been successfully applied to MCI diagnosis [22], [26] with improved performance than the low-order FCNs. Thus, combining these more complex brain network modeling methods with our proposed multinetwork diagnosis framework could further improve the classification performance, which will be investigated in our future work.

Our study adopted the standard AAL atlas for individualized atlas generation. Many other standard atlases, including the Harvard–Oxford (HO) cortical atlas [72], Hammersmith atlas [73], and Yeo 2011 parcellation [74], could also be used in our multinetwork diagnosis framework. Furthermore, multiple FCNs constructed using various atlases may provide richer discriminative information than using just a single type of atlas. Thus, investigating performance of our method based on different parcellations is worthy of a future study.

It is worth further investigating how our proposed analysis framework would work with other types of clustering algorithms, such as k-means and hierarchical clustering. One potential limitation of AP clustering is that no feature selection has been taken into account for the clustering. Such a clustering procedure without feature optimization may not provide the optimal solution since the true underlying clusters presented in the data may differ only with respect to a subset of the features. A sparse learning-based clustering algorithm is considered as a good alternative to overcome this issue. By exploiting a sparse-induced constraint, sparse clustering [75] has been developed to exclude those noninformative features for achieving a more accurate clustering solution. Thus, incorporating sparse clustering into our analysis framework is also worth our further study.

V. Conclusion

In this study, we have proposed a multiview feature learning method with the multiatlas-based FCNs for accurate identification of MCI patients. A three-step transformation was implemented to generate individualized atlases from the standard AAL template. We constructed multiple FCNs based on multiple preselected atlas exemplars for each subject. By treating each of the FCNs as a view for characterizing the functional connectome of the same brain, we further developed a novel subject-relationship-guided SGL algorithm for multiview feature selection among multiple FCNs. Then, an SVM was trained with these selected features, to provide supplementary discriminant information for improved brain disease diagnosis. Extensive experimental comparisons were carried out between our proposed method and other competing approaches. The experimental results showed superior classification performance of our method, suggesting its potential in the connectome-based individualized diagnosis of brain disease.

References

[1] A. Association *et al.*, "Alzheimer's disease facts and figures," *Alzheimer's Dementia*, vol. 9, no. 2, pp. 208–245, 2013.

[2] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "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*, vol. 34, no. 7, p. 939, 1984.

[3] S. Gauthier *et al.*, "Mild cognitive impairment," *Lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.

[4] C. DeCarli, "Mild cognitive impairment: Prevalence, prognosis, aetiology, and treatment," *Lancet Neurol.*, vol. 2, no. 1, pp. 15–21, 2003.

[5] X. Chen *et al.*, "High-order resting-state functional connectivity network for MCI classification," *Human Brain Map.*, vol. 37, no. 9, pp. 3282–3296, 2016.

[6] G. Wu, F. Qi, and D. Shen, "Learning-based deformable registration of MR brain images," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1145–1157, Sep. 2006.

[7] L. Zhang, H. Zhang, X. Chen, Q. Wang, P.-T. Yap, and D. Shen, "Learning-based structurally-guided construction of resting-state functional correlation tensors," *Magn. Resonance Imag.*, vol. 43, pp. 110–121, 2017.

[8] W. Koch *et al.*, "Diagnostic power of default mode network resting state FMRI in the detection of alzheimer's disease," *Neurobiol. Aging*, vol. 33, no. 3, pp. 466–478, 2012.

[9] Y. Fan *et al.*, "Multivariate examination of brain abnormality using both structural and functional MRI," *NeuroImage*, vol. 36, no. 4, pp. 1189–1199, 2007.

[10] W. J. Jagust *et al.*, "The alzheimer's disease neuroimaging initiative positron emission tomography core," *Alzheimer's Dementia*, vol. 6, no. 3, pp. 221–229, 2010.

[11] M. H. Lee, C. D. Smyser, and J. S. Shimony, "Resting-state fMRI: A review of methods and clinical applications," *Amer. J. Neuroradiol.*, vol. 34, no. 10, pp. 1866–1872, 2012.

[12] F. Peters, F. Collette, C. Degueldre, V. Sterpenich, S. Majerus, and E. Salmon, "The neural correlates of verbal short-term memory in alzheimer's disease: An fMRI study," *Brain*, vol. 132, no. 7, pp. 1833–1846, 2009.

[13] M. P. Van Den Heuvel and H. E. H. Pol, "Exploring the brain network: A review on resting-state fMRI functional connectivity," *Eur. Neuropsychopharmacol.*, vol. 20, no. 8, pp. 519–534, 2010.

[14] S. M. Smith *et al.*, "Resting-state fMRI in the human connectome project," *NeuroImage*, vol. 80, pp. 144–168, Oct. 2013.

[15] O. Sporns, "The human connectome: A complex network," *Ann. NewYork Acad. Sci.*, vol. 1224, no. 1, pp. 109–125, 2011.

[16] M. R. Brier *et al.*, "Functional connectivity and graph theory in preclinical alzheimer's disease," *Neurobiol. Aging*, vol. 35, no. 4, pp. 757–768, 2014.

[17] J. S. Damoiseaux, K. E. Prater, B. L. Miller, and M. D. Greicius, "Functional connectivity tracks clinical deterioration in alzheimer's disease," *Neurobiol. Aging*, vol. 33, no. 4, p. 828, 2012.

[18] Z. Yao *et al.*, "Abnormal cortical networks in mild cognitive impairment and alzheimer's disease," *PLoS Comput. Biol.*, vol. 6, no. 11, 2010, Art. no. e1001006.

[19] S. Huang *et al.*, "Learning brain connectivity of alzheimer's disease by sparse inverse covariance estimation," *NeuroImage*, vol. 50, no. 3, pp. 935–949, 2010.

[20] H. Lee, D. S. Lee, H. Kang, B.-N. Kim, and M. K. Chung, "Sparse brain network recovery under compressed sensing," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1154–1165, May 2011.

[21] H. Zhang *et al.*, "Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment," *J. Alzheimers Disease*, vol. 54, no. 3, pp. 1095–1112, 2016.

[22] H. Zhang, X. Chen, Y. Zhang, and D. Shen, "Test-retest reliability of 'high-order' functional connectivity in young healthy adults," *Front. Neurosci.*, vol. 11, p. 439, Aug. 2017.

[23] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nat. Rev. Neurosci.*, vol. 10, no. 3, p. 186, 2009.

[24] S. M. Smith *et al.*, "Network modelling methods for fMRI," *NeuroImage*, vol. 54, no. 2, pp. 875–891, 2011.

[25] N. Tzourio-Mazoyer *et al.*, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, 2002.

[26] X. Chen *et al.*, "Hierarchical high-order functional connectivity networks and selective feature fusion for mci classification," *Neuroinformatics*, vol. 15, no. 3, pp. 271–284, 2017.

[27] J. D. Power *et al.*, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[28] M. Liu, D. Zhang, D. Shen, and A. D. N. Initiative, "View-centralized multi-atlas classification for alzheimer's disease diagnosis," *Human Brain Map.*, vol. 36, no. 5, pp. 1847–1865, 2015.

[29] R. Min, G. Wu, J. Cheng, Q. Wang, D. Shen, and A. D. N. Initiative, "Multi-atlas based representations for alzheimer's disease diagnosis," *Human Brain Map.*, vol. 35, no. 10, pp. 5052–5070, 2014.

[30] Y. Zhou *et al.*, "Functional MRI registration with tissue-specific patch-based functional correlation tensors," *Human Brain Map.*, vol. 39, no. 6, pp. 2303–2316, 2018.

[31] D. Wang *et al.*, "Parcellating cortical functional networks in individuals," *Nat. Neurosci.*, vol. 18, no. 12, p. 1853, 2015.

[32] D. J. Hawellek, J. F. Hipp, C. M. Lewis, M. Corbetta, and A. K. Engel, "Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 47, pp. 19066–19071, 2011.

[33] X. Zhu, X. Li, and S. Zhang, "Block-row sparse multiview multilabel learning for image classification," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 450–461, Feb. 2016.

[34] J. Wang, Q. Wang, H. Zhang, J. Chen, S. Wang, and D. Shen, "Sparse multiview task-centralized ensemble learning for asd diagnosis based on age-and sex-related functional connectivity patterns," *IEEE Trans. Cybern.*, vol. 49, no. 8, pp. 3141–3154, Aug. 2019.

[35] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Temporally constrained sparse group spatial patterns for motor imagery BCI," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3322–3332, Sep. 2019.

[36] X. Xue, F. Nie, Z. Li, S. Wang, X. Li, and M. Yao, "A multiview learning framework with a linear computational cost," *IEEE Trans. Cybern.*, vol. 48, no. 8, pp. 2416–2425, Aug. 2018.

[37] J. Wang, F. Tian, H. Yu, C. H. Liu, K. Zhan, and X. Wang, "Diverse non-negative matrix factorization for multiview data representation," *IEEE Trans. Cybern.*, vol. 48, no. 9, pp. 2620–2632, Sep. 2018.

[38] C. Xu, D. Tao, and C. Xu. (2013). *A Survey on Multi-View Learning*. [Online]. Available: https://arxiv.org/abs/1304.5634

[39] S. Sun, X. Xie, and C. Dong, "Multiview learning with generalized eigenvalue proximal support vector machines," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 688–697, Feb. 2019.

[40] D. Nie, L. Wang, E. Adeli, C. Lao, W. Lin, and D. Shen, "3-D fully convolutional networks for multimodal isointense infant brain image segmentation," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1123–1136, Mar. 2019.

[41] B. Lei, P. Yang, T. Wang, S. Chen, and D. Ni, "Relational-regularized discriminative sparse learning for alzheimer's disease diagnosis," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 1102–1113, Apr. 2017.

[42] J. Koikkalainen *et al.*, "Multi-template tensor-based morphometry: Application to analysis of alzheimer's disease," *NeuroImage*, vol. 56, no. 3, pp. 1134–1144, 2011.

[43] M. Liu, D. Zhang, and D. Shen, "Relationship induced multi-template learning for diagnosis of alzheimers disease and mild cognitive impairment," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1463–1474, Jun. 2016.

[44] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007.

[45] J. Zhou, J. Liu, V. A. Narayan, and J. Ye, "Modeling disease progression via fused sparse group lasso," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2012, pp. 1095–1103.

[46] Y. Jiao *et al.*, "Sparse group representation model for motor imagery EEG classification," *IEEE J. Biomed. Health. Inf.*, vol. 23, no. 2, pp. 631–641, Mar. 2019.

[47] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.

[48] L. v. d. Maaten and G. Hinton, "Visualizing data using T-SNE," *J. Mach. Learn. Res.*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[49] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "A sparse-group lasso," *J. Comput. Graph. Stat.*, vol. 22, no. 2, pp. 231–245, 2013.

[50] Y. Nesterov, "Smooth minimization of non-smooth functions," *Math. Program.*, vol. 103, no. 1, pp. 127–152, 2005.

[51] J. Liu, S. Ji, and J. Ye, "Multi-task feature learning via efficient l 2, 1-norm minimization," in *Proc. 25th Conf. Uncertainty Artif. Intell.*, 2009, pp. 339–348.

[52] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, vol. 87. New York, NY, USA: Springer, 2013.

[53] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.

[54] K. Mevel, G. Chételat, F. Eustache, and B. Desgranges, "The default mode network in healthy aging and alzheimer's disease," *Int. J. Alzheimer's Disease*, vol. 2011, Jun. 2011, Art. no. 535816.

[55] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network," *Ann. New York Acad. Sci.*, vol. 1124, no. 1, pp. 1–38, 2008.

[56] R. L. Buckner *et al.*, "Molecular, structural, and functional characterization of alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory," *J. Neurosci.*, vol. 25, no. 34, pp. 7709–7717, 2005.

[57] H. Jacobs *et al.*, "Functional integration of parietal lobe activity in early alzheimer disease," *Neurology*, vol. 78, no. 5, pp. 352–360, 2012.

[58] M. L. Ries, T. W. Schmitz, T. N. Kawahara, B. M. Torgerson, M. A. Trivedi, and S. C. Johnson, "Task-dependent posterior cingulate activation in mild cognitive impairment," *NeuroImage*, vol. 29, no. 2, pp. 485–492, 2006.

[59] B. Magnin *et al.*, "Support vector machine-based classification of alzheimer's disease from whole-brain anatomical MRI," *Neuroradiology*, vol. 51, no. 2, pp. 73–83, 2009.

[60] G. Karas *et al.*, "Amnestic mild cognitive impairment: Structural MR imaging findings predictive of conversion to alzheimer disease," *Amer. J. Neuroradiol.*, vol. 29, no. 5, pp. 944–949, 2008.

[61] S. A. Rombouts, F. Barkhof, R. Goekoop, C. J. Stam, and P. Scheltens, "Altered resting state networks in mild cognitive impairment and mild alzheimer's disease: An fMRI study," *Human Brain Map.*, vol. 26, no. 4, pp. 231–239, 2005.

[62] L. Wang *et al.*, "Changes in hippocampal connectivity in the early stages of alzheimer's disease: evidence from resting state fMRI," *NeuroImage*, vol. 31, no. 2, pp. 496–504, 2006.

[63] H. Ni, L. Zhou, X. Ning, and L. Wang, "Exploring multifractal-based features for mild alzheimer's disease classification," *Magn. Reson. Med.*, vol. 76, no. 1, pp. 259–269, 2016.

[64] Y. Li *et al.*, "Discriminant analysis of longitudinal cortical thickness changes in alzheimer's disease using dynamic and network features," *Neurobiol. Aging*, vol. 33, no. 2, p. 427, 2012.

[65] G. B. Frisoni, A. Prestia, P. E. Rasser, M. Bonetti, and P. M. Thompson, "In vivo mapping of incremental cortical atrophy from incipient to overt alzheimer's disease," *J. Neurol.*, vol. 256, no. 6, p. 916, 2009.

[66] N. Filippini *et al.*, "Anatomically-distinct genetic associations of apoe E4 allele load with regional cortical atrophy in alzheimer's disease," *NeuroImage*, vol. 44, no. 3, pp. 724–728, 2009.

[67] S. W. Scheff, D. A. Price, F. A. Schmitt, M. A. Scheff, and E. J. Mufson, "Synaptic loss in the inferior temporal gyrus in mild cognitive impairment and alzheimer's disease," *J. Alzheimers Disease*, vol. 24, no. 3, pp. 547–557, 2011.

[68] H. Braak, I. Alafuzoff, T. Arzberger, H. Kretzschmar, and K. Del Tredici, "Staging of alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry," *Acta Neuropathol.*, vol. 112, no. 4, pp. 389–404, 2006.

[69] C.-Y. Wee, S. Yang, P.-T. Yap, and D. Shen, "Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification," *Brain Imag. Behav.*, vol. 10, no. 2, pp. 342–356, 2016.

[70] M. J. Rosa *et al.*, "Sparse network-based models for patient classification using FMRI," *NeuroImage*, vol. 105, pp. 493–506, Jan. 2015.

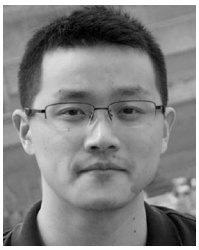
[71] J. Lv *et al.*, "Sparse representation of whole-brain fMRI signals for identification of functional networks," *Med. Image Anal.*, vol. 20, no. 1, pp. 112–134, 2015.

[72] J. A. Frazier *et al.*, "Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder," *Amer. J. Psychiatry*, vol. 162, no. 7, pp. 1256–1265, 2005.

[73] A. Hammers *et al.*, "Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe," *Human Brain Map.*, vol. 19, no. 4, pp. 224–247, 2003.

[74] B. T. Yeo *et al.*, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *J. Neurophysiol.*, vol. 106, no. 3, pp. 1125–1165, 2011.

[75] D. M. Witten and R. Tibshirani, "A framework for feature selection in clustering," *J. Amer. Stat. Assoc.*, vol. 105, no. 490, pp. 713–726, 2010.



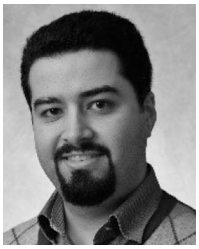
Yu Zhang (Senior Member, IEEE) received the Ph.D. degree from the East China University of Science and Technology, Shanghai, China.

He is an Assistant Professor of bioengineering with Lehigh University, Bethlehem, PA, USA. He received postdoctoral training with the Department of Psychiatry and Behavior Sciences, Stanford University, Stanford, CA, USA, and the Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. He has authored over 100 peer-reviewed papers that have been published in the prestigious journals, such as *Nature Biotechnology*, *Nature Human Behavior*, the PROCEEDINGS OF THE IEEE, the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, and the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING. His research interests include computational neuroscience, brain network, pattern recognition, machine learning, signal processing, artificial intelligence, brain–computer interface, and medical imaging computing.

Han Zhang received the B.S. degree in electronic science and technology from Zhejiang University, Hangzhou, China, in 2005, and the Ph.D. degree in cognitive neuroscience from the National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China, in 2011.

He is currently a Research Assistant Professor with the Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. He has more than 14 years of research experience in neuroimage computing, with more than 100 publications in journals and conferences. His research interests lie in the methodological development of brain network modeling for individualized brain disease diagnosis based on multimodal neuroimaging and machine learning.

Dr. Zhang is actively serving as a Reviewer for the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, the IEEE TRANSACTIONS ON MEDICAL IMAGING, the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, and the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, along with other prestigious journals, such as the *Journal of Neuroscience*, *Cerebral Cortex*, *Neural Networks*, *NeuroImage*, *Human Brain Mapping*, and *Medical Image Analysis*. He has been a member of the IEEE Engineering in Medicine and Biology Society since 2018.

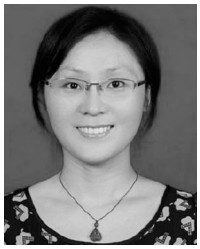


Ehsan Adeli (Member, IEEE) received the Ph.D. degree from the Iran University of Science and Technology, Tehran, Iran.

He is a Faculty Member with the Computational Neuroscience Laboratory, Department of Psychiatry and Behavioral Sciences, and is affiliated with the Stanford Vision and Learning Laboratory, Stanford AI Laboratory, Department of Computer Science, Stanford University, Stanford, CA, USA. His research falls at the intersection of machine learning, computer vision, and computational neuroscience, as it focuses on the investigation of various computational and statistical learning-based methods in processing both natural and neuroimages.

Xiaobo Chen (Member, IEEE) received the Ph.D. degree in pattern recognition and intelligent systems from the Nanjing University of Science and Technology, Nanjing, China, in 2013.

In 2011, he served as a Research Assistant with Hong Kong Polytechnic University, Hong Kong. From 2015 to 2017, he served as a Postdoctoral Research Associate with the University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. He is currently a Professor with the Automotive Engineering Research Institute, Jiangsu University, Zhenjiang, China. His major research interests include pattern recognition and its applications.



Mingxia Liu (Senior Member, IEEE) received the B.S. and M.S. degrees from Shandong Normal University, Shandong, China, in 2003 and 2006, respectively, and the Ph.D. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2015.

Her current research interests include machine learning, pattern recognition, and medical image analysis.

Dinggang Shen (Fellow, IEEE) received the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China.

He has published more than 1100 peer-reviewed papers in international journals and conference proceedings, with an H-index of 101. His research interests include medical image analysis, computer vision, and pattern recognition.

Dr. Shen serves as an editorial board member for eight international journals. He has served on the Board of Directors with the Medical Image Computing and Computer Assisted Intervention Society from 2012 to 2015. He was the General Chair for MICCAI 2019. He is a Fellow of the American Institute for Medical and Biological Engineering and the International Association for Pattern Recognition.