Classification of Alzheimer's Disease Using Whole Brain Hierarchical Network

Jin Liu, Min Li, Wei Lan, Fang-Xiang Wu, Yi Pan, and Jianxin Wang

Abstract—Regions of interest (ROIs) based classification has been widely investigated for analysis of brain magnetic resonance imaging (MRI) images to assist the diagnosis of Alzheimer's disease (AD) including its early warning and developing stages, e.g., mild cognitive impairment (MCI) including MCI converted to AD (MCIc) and MCI not converted to AD (MCInc). Since an ROI representation of brain structures is obtained either by pre-definition or by adaptive parcellation, the corresponding ROI in different brains can be measured. However, due to noise and small sample size of MRI images, representations generated from single or multiple ROIs may not be sufficient to reveal the underlying anatomical differences between the groups of disease-affected patients and health controls (HC). In this paper, we employ a whole brain hierarchical network (WBHN) to represent each subject. The whole brain of each subject is divided into 90, 54, 14, and 1 regions based on Automated Anatomical Labeling (AAL) atlas. The connectivity between each pair of regions is computed in terms of Pearson's correlation coefficient and used as classification feature. Then, to reduce the dimensionality of features, we select the features with higher $F-$ scores. Finally, we use multiple kernel boosting (MKBoost) algorithm to perform the classification. Our proposed method is evaluated on MRI images of 710 subjects (200 AD, 120 MCIc, 160 MCInc, and 230 HC) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The experimental results show that our proposed method achieves an accuracy of 94.65 percent and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.954 for AD/HC classification, an accuracy of 89.63 percent and an AUC of 0.907 for AD/MCI classification, an accuracy of 85.79 percent and an AUC of 0.826 for MCI/HC classification, and an accuracy of 72.08 percent and an AUC of 0.716 for MCIc/MCInc classification, respectively. Our results demonstrate that our proposed method is efficient and promising for clinical applications for the diagnosis of AD via MRI images.

Index Terms—Alzheimer's disease, regions of interest, hierarchical network, classification

1 Introduction

Alzheimer's Disease (AD) is a degenerative brain disease and the most common cause of dementia. The symptoms of dementia include memory loss and difficulties with thinking, problem-solving or language, which seriously affect patient's daily life. Mild cognitive impairment (MCI) [1] is an early stage of AD characterized by significant cognitive impairment in the absence of dementia. In the United States, there are more than 5.2 million people with Alzheimer's disease in 2014, and it is estimated that 13.8 million Americans have AD [2] by 2050. Precise prediction and diagnosis of AD, especially at its early warning stage such as MCI, have become a crucial step to delay or even avoid dementia.

Neuroimaging techniques provide a way for clinicians to examine the structural and functional changes in the brain associated with the development of diseases *in vivo* [3], [4], [5], [6]. Commonly used modalities include magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), single-photon emission computed tomography (SPECT), and diffusion tensor imaging (DTI). Owning to its easy access in clinical settings, MRI receives the most attention of researchers compared with other modalities. The structural changes in the brain associated with AD can be non-invasively assessed using MRI. As shown in Fig. 1, AD patients typically have evidence of cortical atrophy, and enlarged ventricles in comparison with health controls (HC). Medial temporal lobe (MTL) atrophy as assessed using MRI has proven to be an effective clinical aid in the early diagnosis of AD [7], [8], [9], and histological studies show that the hippocampus, amygdala and entorhinal cortex are particularly vulnerable to AD pathology [3], [10], [11], [12].

Regions of interest (ROIs) based analysis has become one of the most popular approaches for automatic AD diagnosis. By directly accessing to the structures provided by MRI, brain ROIs can be utilized to identify the anatomical differences between populations of AD, MCI and HC, and subsequently determine the AD-related characteristics to assist diagnosis, prognosis, as well as evaluation of MCI progression and treatment effects. The determination of each ROI is the key for ROIs based analysis methods. Once ROIs have been determined either by predefinition or by adaptive parcellation, each ROI can be used as feature for classification. ROIs based methods can be divided into two groups according to the number of ROI: single ROI methods [8], [13] and multiple ROIs methods [14], [15], [16], [17]. Chupin et al. [13] proposed a segmentation method to obtain hippocampus and used the obtained hippocampal volumes to automatically discriminate among AD, MCI and HC

- *J. Liu, M. Li, W. Lan, and J. Wang are with the School of Information Science and Engineering, Central South University, Changsha 410083, China. E-mail: {liujin06, limin, weilan, jxwang}@mail.csu.edu.cn.*
- *F. X. Wu is with the Division of Biomedical Engineering and Department of Mechanical Engineering, University of Saskatchewan, Saskatoon SKS7N5A9, Canada. E-mail: faw341@mail.usask.ca.*
- *Y. Pan is with the Department of Computer Science, Georgia State University, Atlanta, GA 30302. E-mail: yipan@gsu.edu.*

Manuscript received 17 Apr. 2016; revised 26 July 2016; accepted 29 Nov. 2016. Date of publication 2 Dec. 2016; date of current version 30 Mar. 2018.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TCBB.2016.2635144

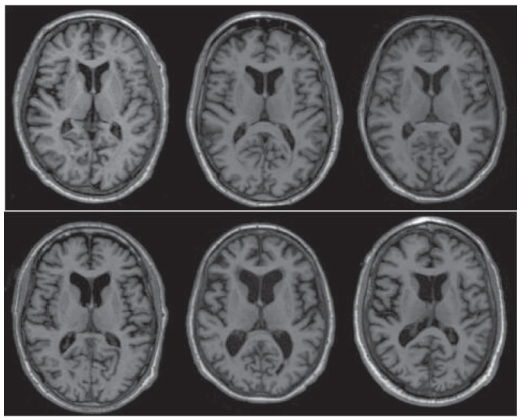


Fig. 1. Cross sections from MRI images of HC (the top row) and AD (the bottom row).

subjects. Ahmed et al. [15] combined hippocampus visual features and cerebrospinal fluid (CSF) volume to yield an automatic classification framework for AD diagnosis. To discriminate between AD and HC subjects, Magnin et al. [14] used 90 ROIs to represent whole brain features via a common template. Liu et al. [17] used 83 ROIs to represent whole brain features. In this paper, we focus on multiple ROIs methods of the whole brain because brain structural changes of neurodegeneration diseases such as AD may occur not in single ROI, but in multiple ROIs and even the whole brain. Furthermore, disease-induced brain structural changes may not only be ROIs themselves, but also the correlation among ROIs. Therefore, the correlation changes may be more efficient to discriminate AD, MCI and HC.

It is widely reported that brain is a complex network and recently, brain network based methods [18], [19] have been increasingly popular in medical image analysis by neuroimaging techniques. Brain network based methods have also been applied to accurately identify individuals with AD and MCI from HC [20], [21], [22], [23], [24]. For example, Yao et al. [21] calculated the network properties for AD, MCI, HC, respectively, to investigate the abnormality among the three groups. Dai et al. [23] established a network based on a kernel-based method for each subject using mean cortical thickness and the network edges were used as features to make prediction of AD/NC through support vector machines with radial basis function (RBF-SVM). Li et al. [22] constructed individual networks based on the correlation between the longitudinal thickness changes of different ROIs and used the network features to train an SVM to predict AD and MCI patients. Commonly used brain network methods can be categorized into two groups: group network and individual network. The former constructed a network for each group of subjects and investigated the abnormality of network measures in patients, while the latter established a network for each subject and these networks were used as features for disease diagnosis. Because of individual network differences, the latter is considered in this paper.

Texture analysis [25] is an important and useful area of study in machine vision. Most natural surfaces exhibit texture and a successful vision system must be able to deal with the textured world surrounding it. Texture analysis methods have been utilized in medical image analysis [26], [27], [28]. The gray level co-occurrence matrices (GLCM)[25] have become one of the most well-known and widely used texture features and have also been applied to AD studies [29], [30].

In this paper, in order to generate a rich representation of anatomical structures that will be more discriminative to separate different groups of subjects, we propose to construct a whole brain hierarchical network (WBHN) to measure the correlation of two ROIs. This hierarchical network includes four layers and each layer includes the different number of ROIs. In order to consider different information provided by ROIs of different sizes, unlike most existing ROIs based methods [13], [14], [15] which register their subjects to a single atlas via different registration strategy, we propose to register each subject to four different template based on Automated Anatomical Labeling (AAL) atlas [31]. In this hierarchical network, AAL [31] atlas is first applied to register the whole brain to 90 ROIs for each subject. Then, the original 90 ROIs are combined to form the other three templates via different rules. Next, we connect all the ROIs in the four templates to form this hierarchical network. Afterwards, texture features of each region are calculated and the connectivity between two regions is also calculated by Pearson correlation of texture features of them and used as classification features. Subsequently we select the top ranked features from the extracted features by an F-score method. Finally, we use the MKBoost-S2 [32] algorithm for the classification. The main contributions of this paper can be summarized as follows:

- A hierarchical network is constructed to provide the complementary information for classification.
- New ROIS combinations and feature selection methods are also proposed.
- The MKBoost based classification method is used for AD and MCI diagnosis based on MRI data.

By performing 10-fold cross validation with the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [33], our proposed method achieves a very promising performance for AD/HC, MCI/HC, AD/MCI, and MCIc/MCInc classification.

The rest of the article is organized as follows. we first describe the details of the proposed method in the section MATERIALS AND METHODS. Then, we illustrate the experiments and discuss comparison results in the section RESULTS AND DISCUSSION. Finally, we draw conclusions in the section CONCLUSION.

2 Materials and Methods

2.1 Data Description

Data used for evaluation of our proposed method are taken from ADNI database [33]. The primary goal of the ADNI has been to test whether MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. For more details with the ADNI database, please see http://adni.loni.usc.edu/.

TABLE 1
Demographic Information of 710 Subjects (from ADNI Database)

Diagnostic Type	Number	Age	Gender (M/F)	MMSE
HC	230	77.13 ± 6.24	116/84	29.16 ± 0.82
MCInc	160	76.26 ± 5.35	89/71	27.56 ± 1.18
MCIc	120	75.95 ± 6.27	53/67	26.38 ± 1.52
MCI	280	76.11 ± 5.98	142/138	26.97 ± 1.34
AD	200	76.63 ± 5.91	122/108	23.54 ± 2.07

The values are denoted as mean $\pm$ standard deviation.

In our experiments, we used the T_1-weighted MRI data obtained by the ADNI acquisition protocol in [34] from 710 ADNI participants. The cohort consisted of 200 patients with Alzheimer's Disease, 120 subjects that had mild cognitive impairment and converted to AD within 18 months (MCIc), 160 subjects with mild cognitive impairment that did not convert (MCInc) and 230 healthy controls (HC). Table 1 presents a summary of the demographic information of these 710 subjects from the ADNI database in this paper. Mini Mental State Examination is abbreviated to MMSE in Table 1.

2.2 Data Preprocessing

A four-step preprocessing procedure is applied to each MRI brain image. First, non-parametric non-uniform bias correction (N3) algorithm [35] is used to correct intensity inhomogeneity. Second, the skull and cerebellum are removed by using BET [36]. Third, each MRI brain image is further segmented into three tissue types, namely gray matter (GM), white matter (WM) and CSF by using FAST [37]. Finally, all brain images are affine aligned by FLIRT [38], [39].

2.3 Hierarchical Network Construction

2.3.1 Anatomical Parcellation

In order to parcellate the whole brain, AAL atlas previously proposed by Tzourio-Mazoyer et al. [31] is used, which is shown in Fig. 2. This atlas parcellates the brain into 90 cerebral regions (45 in each hemisphere) and 26 cerebellar regions (nine in each cerebellar hemisphere and eight in the vermis). Details of AAL regions can be found in a XML file (ROI_MNI_V4.xml) of AAL package (www.cyceron.fr/index.php/en/plateforme-en/freeware). Cerebellar regions are typically not considered to be related to AD and dementia [14], [40], and have therefore been removed from the analysis.

2.3.2 Hierarchical Description

As some ROIs in the existing ROIs have a certain neighborhood and functional dependencies, for example, Cingulum_Ant, Cingulum_Mid and Cingulum_Post are a part of Cingulum; Cingulum_Ant, Cingulum_Mid, Cingulum_Post, Hippocampus, ParaHippocampal and Amygdala are a part of limbic system, these ROIs are integrated to generate new ROIs. In this article, we define the layers L_c, where $c = 1, 2, 3, 4$, to represent a brain of different number of ROIs as shown in Fig. 3. The details of each layer are as follows:

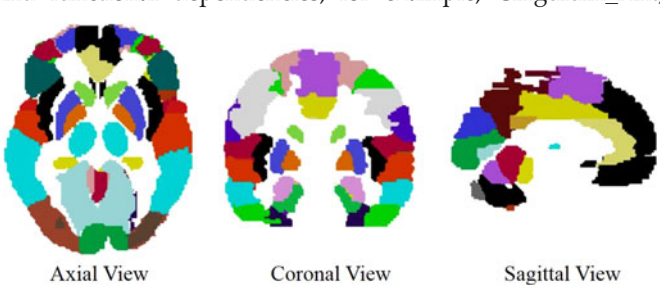


Fig. 2. AAL atlas. Different colors represent different regions.

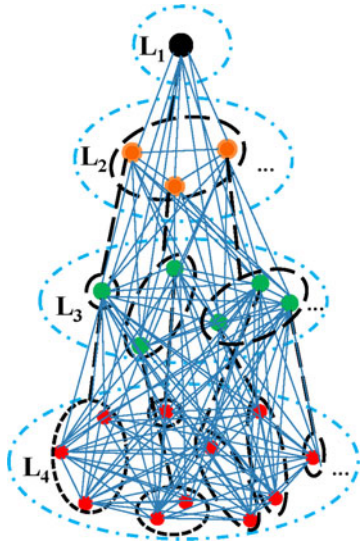


Fig. 3. Whole brain hierarchical network.

- The layer L_4 contains 90 original ROIs based on AAL atlas.
- The layer L_3 contains 54 ROIs. A group of ROIs in L_4 are merged into an ROI in L_3 if their first two digits and the last digit are the same, for example, Cingulum_Ant_L (4001), Cingulum_Mid_L (4011) and Cingulum_Post_L (4021) are considered as an ROI in the layer L_3.
- Similar to the process of the layer L_3. A group of ROIs in L_4 are merged into an ROI in L_2 if their first and the last digit are the same, for example, Cingulum_Ant_L (4001), Cingulum_Mid_L (4011), Cingulum_Post_L (4021), Hippocampus_L (4101), ParaHippocampal_L (4111) and Amygdala_L (4201) are considered as an ROI in the layer L_2. The layer L_2 contains 14 ROIs.
- The layer L_1 contains only one ROI, that is the whole brain.

2.3.3 The Representation and Connectivity of the ROIs

In this section, GLCM is used to extract 3D texture parameters[41] of each ROI in 0^0, 45^0, 90^0 and 135^0 directions. These parameters include energy, contrast, inverse difference moment, entropy, difference variance and difference entropy as shown in Table 2. In Table 2, $p_{x-y}(l) = \sum_{i=1} \sum_{j=1} p(i,j)$, u_{x-y} is the average value of $p_{x-y}(l)$, where $l = |i-j|$, $p(i,j)$ is the value of GLCM. For more details, please see [41]. We then calculate the average texture parameters of each ROI in four directions. A six-dimensional vector comprising the mean value of each texture parameter is used to represent each ROI. Afterwards, the connectivity between each pair of ROIs is evaluated by Pearson's correlation coefficients, resulting in $(90 \times 89 + 54 \times 53 + 14 \times 13)/2 + 0 + (90 \times 54 + 90 \times 14 + 90 \times 1) + (54 \times$

TABLE 2
Texture Parameter Formula

Parameter	Formula
Energy	$\sum_{i=1}\sum_{j=1} p(i,j)^2$
Contrast	$\sum_{i=1}\sum_{j=1} (i-j)^2 p(i,j)$
Inverse Difference Moment	$\sum_{i=1}\sum_{j=1} \frac{1}{1+(i-j)^2} p(i,j)$
Entropy	$-\sum_{i=1}\sum_{j=1} p(i,j)\log(p(i,j))$
Difference Variance	$\sum_{l=0} (l-u_{x-y})^2 p_{x-y}(l)$
Difference Entropy	$-\sum_{l=0} p_{x-y}(l)\log(p_{x-y}(l))$

$14 + 54 \times 1) + (14 \times 1) = 12,561$ dimensional connectivity feature vectors for each subject.

Finally, a whole brain hierarchical network can be constructed as shown in Fig. 3, the ROIs correspond to its nodes and the connectivity between two ROIs correspond to its edges. These connections are the features used in all subsequent analyses.

2.4 Feature Selection

Given that some features are uninformative, irrelevant or redundant for classification, reducing a number of features not only speed up computation, but also improve classification performance. Therefore, an initial feature selection step was adopted. The feature ranking approach [42], [43], [44], [45] has been widely used in feature selection. F-score [46] is a simple, generally quite effective technique which measures the discrimination of two sets of real numbers and used in previous studies [44], [45]. In this study, the F-score method is employed for feature ranking.

Given training vectors x_k, $k = 1, \ldots, m$, if the number of positive instances (i.e., HC) and negative instances (i.e., AD) are n_+ and n_-, respectively, then the F-score of the ith feature is defined as

$$F(i) = \frac{\left(\overline{x}_i^{(+)} - \overline{x}_i\right)^2 + \left(\overline{x}_i^{(-)} - \overline{x}_i\right)^2}{\frac{1}{n_+-1}\sum_{k=1}^{n_+}\left(x_{k,i}^{(+)} - \overline{x}_i^{(+)}\right)^2 + \frac{1}{n_--1}\sum_{k=1}^{n_-}\left(x_{k,i}^{(-)} - \overline{x}_i^{(-)}\right)^2}, \tag{1}$$

where $\overline{x}_i$, $\overline{x}_i^{(+)}$, $\overline{x}_i^{(-)}$ are the average of the ith feature of the whole, positive, and negative data sets, respectively; $x_{k,i}^{(+)}$ is the ith feature of the kth positive instance, and $x_{k,i}^{(-)}$ is the ith feature of the kth negative instance. The numerator indicates the discrimination between the positive and negative sets, and the denominator indicates the one within each of the two sets. The larger the F-score is, the more likely this feature is more discriminative. Therefore, we use this score as a feature selection criterion.

2.5 Classification and Multiple Kernel Learning

Multiple Kernel Learning (MKL) is a promising family of machine learning algorithms using multiple kernel functions for various challenging data mining tasks [32], [47], [48], [49]. In this paper, we use MKBoost algorithm proposed by Hao et al. [32] for classification tasks. For each data set, we create a set of 13 base kernels, i.e.,

- Gaussian kernels with 10 different widths ($\{2^{-4}, 2^{-3}, \ldots, 2^4, 2^5\}$) on all features.
- Polynomial kernels of degree 1 to 3 on all features.

For the implementation of the final classification, by default, we set the total number of boosting trials T to 100, the boosting sampling ratio to 0.2, and the sampling decay factor to 2^{-5} for MKBoost-S2. For SVM, we adopt the popular LIBSVM toolbox [50] as the SVM solver, with a default value for the parameter C (= 1).

3 Results and Discussion

3.1 Evaluation Protocol

The evaluation of our proposed method is conducted on four different classifications in Alzheimer's disease diagnosis:

- D1: AD/HC classification;
- D2: AD/MCI classification;
- D3: MCI/HC classification;
- D4: MCIc/MCInc classification;

AD/HC classification is relatively simple, while MCIc/MCInc is the most difficult and has received relatively little attention in the previous studies. However, in order to possibly prevent the conversion from MCI to AD, from HC to MCI and from MCInc to MCIc via timely therapeutic interventions, the latter three diagnoses (D2∼D4) are the more important than the first diagnosis (D1).

In this article, a 10-fold cross validation strategy is employed to evaluate the performance of the classifier. In brief, the whole set of subject samples are equally partitioned into 10 subsets, and each time the subject samples within one subset are successively selected as the testing samples and all remaining subject samples in the other nine subsets are used for training the classifier. The classifier is built for each training set and tested with its corresponding testing subjects. This process is repeated for 100 times independently to avoid any bias introduced by randomly partitioning the dataset in the cross validation. Accuracy (ACC), sensitivity (SEN), and specificity (SPE) are used to quantify the performance of the classifier based on the results of 10-fold cross validation. The three metrics are calculated by the following formulas:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$SEN = \frac{TP}{TP + FN} \tag{3}$$

$$SPE = \frac{TN}{TN + FP}, \tag{4}$$

where TP, FP, TN, and FN are the numbers of true positives, false positives, true negatives, and false negatives, respectively. In addition, we also report the area under receiver operating characteristic (ROC) curve (AUC). The AUC value is an important index to measure the overall-performance of classification methods, and the higher the AUC value, the better the performance of the classification method.

3.2 Overall Classification Performance

An overview of the proposed classification framework is summarized in Fig. 4. As shown in Fig. 5, the accuracy of the classifier based on MKBoost algorithm could reach up

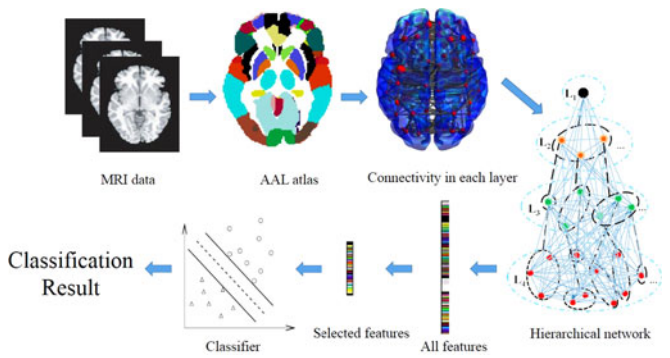


Fig. 4. Schematic diagram illustrating the proposed classification framework.

to 94.65 percent (95.03 percent for sensitivity, 91.76 percent for specificity) by using the 550 top ranked features for AD/HC classification, 89.63 percent (91.55 percent for sensitivity, 86.25 percent for specificity) by using the 600 top ranked features for AD/MCI classification, 85.79 percent (88.91 percent for sensitivity, 80.34 percent for specificity) by using the 450 top ranked features for MCI/HC classification and 72.08 percent (75.11 percent for sensitivity, 71.05 percent for specificity) by using the 300 top ranked features for MCIc/MCInc classification, respectively. Thus, by selecting the top ranked features with the best classification accuracy for the above four classifications, the AUC value of the classifier is 0.954, 0.907, 0.826, 0.716, respectively, as shown in Table 3.

3.3 Comparison of Different Features

All features in this article include the following:

- F_{L_4}: the features (4,005) from the correlation between each pair of ROIs of the layer L_4;
- F_{L_3}: the features (1,431) from the correlation between each pair of ROIs of the layer L_3;
- F_{L_2}: the features (91) from the correlation between each pair of ROIs of the layer L_2;
- F_{L_1}: the features (0) from the correlation between each pair of ROIs of the layer L_1;
- $F_{L_{43}}$: the features (4,860) from the correlation between each ROI of the layer L_4 and each ROI of the layer L_3;
- $F_{L_{42}}$: the features (1,260) from the correlation between each ROI of the layer L_4 and each ROI of the layer L_2;

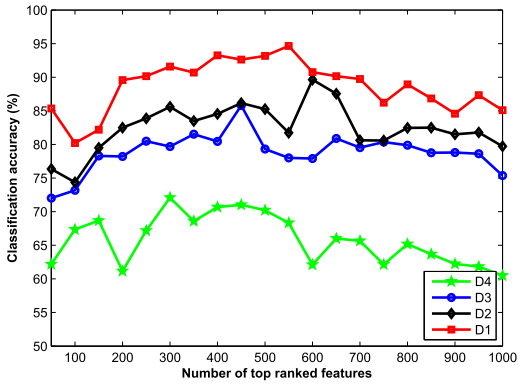


Fig. 5. Classification accuracy for four different classification process. The features are ranked according to F-scores in descending order.

TABLE 3
The Best Performance for Single and Multiple Kernel Classifiers

Classification	Classifier	ACC(%)	SEN(%)	SPE(%)	AUC	Features
AD/HC	C1	91.68	90.34	94.58	0.922	450
	C2	92.85	95.19	89.72	0.937	550
	C3	94.65	95.03	91.76	0.954	550
AD/MCI	C1	85.69	90.78	79.16	0.875	450
	C2	86.74	93.16	85.49	0.882	650
	C3	89.63	91.55	86.25	0.907	600
MCI/HC	C1	84.01	75.24	88.17	0.816	450
	C2	83.55	87.75	78.29	0.807	400
	C3	85.79	88.91	80.34	0.826	450
MCIc/MCInc	C1	70.56	62.15	73.46	0.689	300
	C2	70.97	74.80	68.33	0.702	300
	C3	72.08	75.11	71.05	0.716	300

- $F_{L_{41}}$: the features (90) from the correlation between each ROI of the layer L_4 and each ROI of the layer L_1;
- $F_{L_{32}}$: the features (756) from the correlation between each ROI of the layer L_3 and each ROI of the layer L_2;
- $F_{L_{31}}$: the features (54) from the correlation between each ROI of the layer L_3 and each ROI of the layer L_1;
- $F_{L_{21}}$: the features (14) from the correlation between each ROI of the layer L_2 and each ROI of the layer L_1;

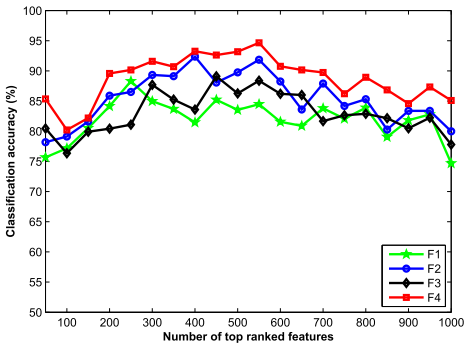
To investigate the classification performance of our proposed hierarchical network features, we put all the features into four different feature sets as follows:

- $F1 = F_{L_4}$;
- $F2 = F_{L_4} + F_{L_3} + F_{L_2} + F_{L_1}$;
- $F3 = F_{L_{43}} + F_{L_{42}} + F_{L_{41}} + F_{L_{32}} + F_{L_{31}} + F_{L_{21}}$;
- $F4 = F2 + F3$.

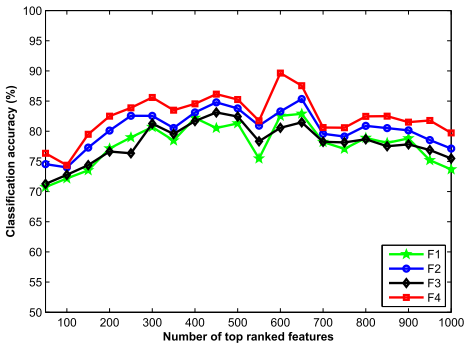
It is obvious that $F2$ contains $F1$ and the only difference between $F2$ and $F4$ is that the latter contains $F3$. Fig. 6 shows the classification accuracy by using different feature sets ($F1$, $F2$, $F3$ and $F4$) in the aforementioned four classification problems. As shown in Fig. 6, the best classification accuracy of $F2$ outperform $F1$ in AD/HC, AD/MCI, MCI/HC and MCIc/MCInc classification, and the best classification accuracy of $F4$ outperform $F2$ in AD/HC, AD/MCI, MCI/HC and MCIc/MCInc classification. Therefore, adding the hierarchical network features (including same-layer features ($F2 - F1$) and between-layer features ($F3$)) helps to improve classification performance. The results demonstrate that feature representation based on whole brain hierarchical network can provide more abundant and effective classification information.

3.4 Comparison between Single and Multiple Kernel Classifiers

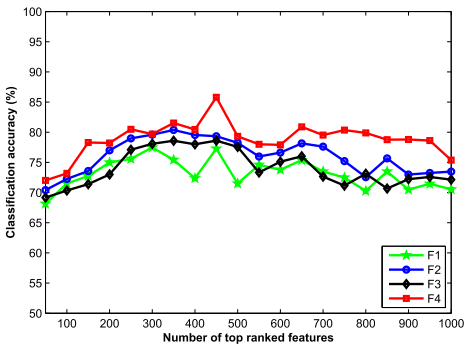
To further investigate the effectiveness of MKBoost classifiers, two single kernel classifiers, a linear kernel SVM classifier and a Gaussian kernel ($\sigma = 1$) SVM classifier, are trained and tested with the same procedure as in the proposed framework. For fair comparison, default parameters (i.e., C = 1) is used for the three classifiers. Fig. 7 shows that MKBoost classifier achieves the best accuracy when the top ranked features are used in the aforementioned four classification problems. Results in Table 3 show that MKBoost classifier performs better than single kernel SVM classifiers in terms of ACC, SEN, SPE, and AUC values. Linear SVM



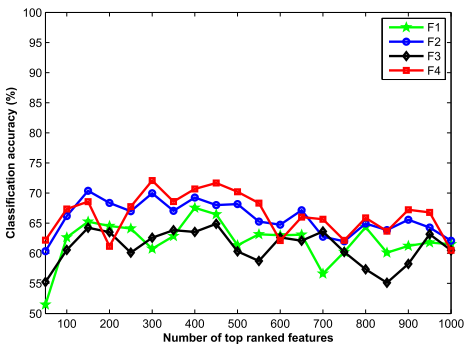
(a) AD/HC



(b) AD/MCI

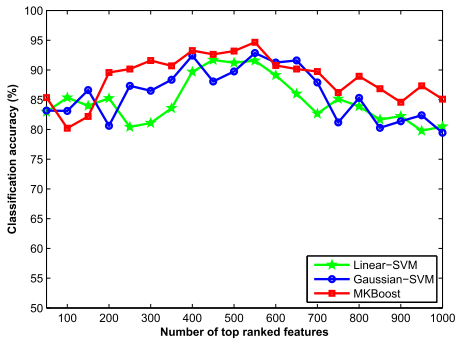


(c) MCI/HC

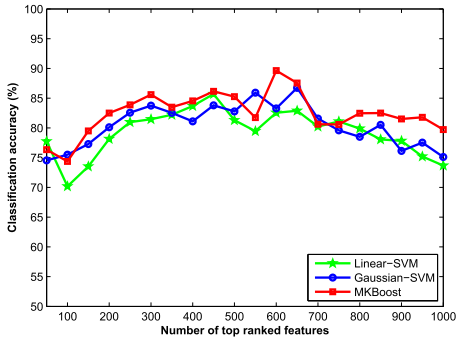


(d) MCIc/MCInc

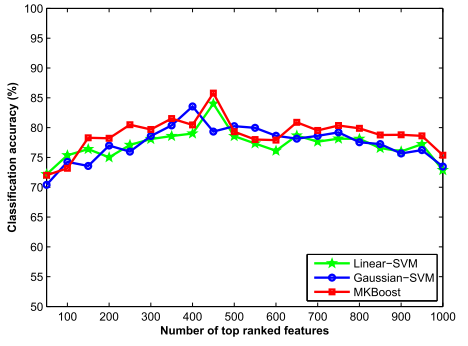
Fig. 6. The accuracy for classification by using different feature sets. The features of each set are ranked according to F-scores in descending order.



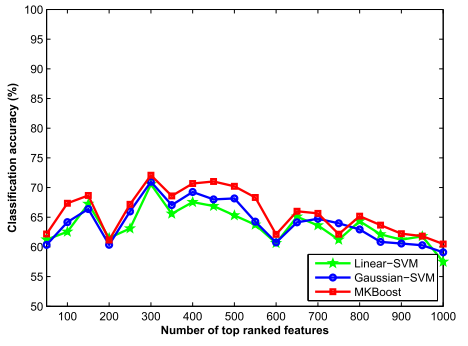
(a) AD/HC



(b) AD/MCI



(c) MCI/HC



(d) MCIc/MCInc

Fig. 7. The accuracy for classification by using linear SVM, Gaussian SVM and MKBoost classifiers. The features are ranked according to F-scores in descending order.

classifier, Gaussian SVM classifier and MKBoost classifier in Table 3 are denoted as C1, C2 and C3, respectively.

3.5 Comparison with Existing Classification Methods

In this section, we compare our results with several recently reported methods on AD classification, based on single ROI [13], multiple ROIs [15], and whole brain [51], [52], demonstrating the superiority of the proposed method. For fair comparison, the data sets of these experiments are the same. Although Suk et al. [51] proposed the multi-modality based approach, only the results based on MRI data are reported in this article. Tables 4, 5, 6, and 7 present the comparative results for AD/HC classification, AD/MCI

TABLE 4
Comparison to Existing Works Using MRI Data of ADNI for AD/HC Classification

Method	ACC(%)	SEN(%)	SPE(%)	AUC
Chupin et al. 2009 [13]	80.51	78.76	82.16	0.785
Ahmed et al. 2015 [15]	86.40	77.61	93.28	0.849
Suk et al. 2014 [51]	93.05	90.86	**94.57**	0.948
Khedher et al. 2015 [52]	88.96	92.35	86.24	0.926
Dai et al. 2013 [23]	90.81	92.59	90.33	0.943
WBHN	**94.65**	**95.03**	91.76	**0.954**

TABLE 5
Comparison to Existing Works Using MRI Data of ADNI for AD/MCI Classification

Method	ACC(%)	SEN(%)	SPE(%)	AUC
Chupin et al. 2009 [13]	73.48	70.35	76.98	0.733
Ahmed et al. 2015 [15]	74.51	77.94	71.23	0.756
Suk et al. 2014 [51]	**88.98**	82.11	**90.65**	0.901
Khedher et al. 2015 [52]	84.59	88.75	83.07	0.886
Dai et al. 2013 [23]	85.92	82.46	87.59	0.874
WBHN	**89.63**	**91.55**	86.25	**0.907**

TABLE 6
Comparison to Existing Works Using MRI Data of ADNI for MCI/HC Classification

Method	ACC(%)	SEN(%)	SPE(%)	AUC
Chupin et al. 2009 [13]	71.94	65.79	74.18	0.716
Ahmed et al. 2015 [15]	76.29	72.30	81.53	0.768
Suk et al. 2014 [51]	83.67	**96.79**	57.28	0.820
Khedher et al. 2015 [52]	82.41	84.12	80.48	0.813
Dai et al. 2013 [23]	81.92	78.51	**88.34**	0.812
WBHN	**85.79**	88.91	80.34	**0.826**

TABLE 7
Comparison to Existing Works Using MRI Data of ADNI for MCIc/MCInc Classification

Method	ACC(%)	SEN(%)	SPE(%)	AUC
Chupin et al. 2009 [13]	64.21	69.74	62.45	0.664
Ahmed et al. 2015 [15]	68.72	67.38	70.69	0.681
Suk et al. 2014 [51]	**72.86**	40.55	**88.49**	0.712
Khedher et al. 2015 [52]	70.11	68.61	74.16	0.708
Dai et al. 2013 [23]	71.04	65.98	75.56	0.709
WBHN	72.08	**75.11**	71.05	**0.716**

classification, MCI/HC classification and MCIc/MCInc classification, respectively. Details of each method are given in the tables.

For AD/HC classification, although the specificity (91.76 percent) of our method is not the best (lower than the other two methods proposed by Ahmed et al. [15] (93.28 percent) and Suk et al. [51] (94.57 percent)), its accuracy (94.65 percent) and sensitivity (95.03 percent) are the best. For AD/MCI classification, similar to AD/HC classification, although Suk et al. [51] obtains the best specificity, its accuracy (88.98 percent) and sensitivity (82.11 percent) is lower than our method (89.63, 91.55 percent). For MCI/HC classification, our method obtains the best accuracy and its specificity and sensitivity are relatively high (≥ 80 percent). Although Suk et al. [51] obtains the highest sensitivity, their specificity (57.28 percent) is very low and although Dai et al. [23] obtains the highest specificity, their specificity is relatively low (≤ 80 percent). The results of MCIc/MCInc classification are similar with AD/MCI classification, Suk et al. [51] also obtains the best sensitivity and its accuracy, but its sensitivity (40.55 percent) is nearly 35 percent lower than our method (75.11 percent). In clinical setting, the accuracy is more important than the sensitivity and the specificity.

As can be seen from Tables 4 to 7, the methods based on whole brain substantially outperform both single ROI method proposed by Chupin et al. [13] and multiple ROIs method proposed by Ahmed et al. [15], in terms of ACC, SEN, SPE and AUC. This result indicates that AD may be related to many regions of the brain, rather than the specific one or several regions, which is in agreement with many existing reports [23], [51], [52].

The AUC has widely been used to evaluate the performance of predictive and diagnostic tests in brain disease as well as other areas. In particular, the AUC can be thought as a measure of the overall performance of predictive and diagnostic tests. For the aforementioned four classifications, Tables 4, 5, 6, and 7 clearly show that our proposed method is better than any other competing methods in terms of AUC and demonstrate that our proposed method is more robust.

4 Conclusion

In summary, we have developed a whole brain hierarchial network to extract features for MRI-based classification of AD in this paper. In the hierarchial network, in addition to the texture features of the ROIs, the spatial-correlations are also used for the better classification accuracy. To deal with the challenge of high-dimension features obtained from raw feature spaces during AD classification, the features are selected according to F-scores. In order to improve performance, we have used a multiple kernel classifier (MKBoost) to do the final classification. Experimental results on the MRI data from ADNI database show that our proposed method is efficient and is promising for clinical applications for the diagnosis of AD.

Acknowledgments

The authors would like to express their gratitude for the support from the National Natural Science Foundation of China under Grant Nos. 61232001, 61420106009 and 61622213, and Doctoral Student Independent Exploration Innovative Project No. 2015zzts051. Jianxin Wang is the corresponding author.

References

[1] R. C. Petersen, "Mild cognitive impairment as a diagnostic entity," *J. Internal Med.*, vol. 256, no. 3, pp. 183–194, 2004.
[2] A. Association, et al., "2014 Alzheimer's disease facts and figures," *Alzheimer's Dementia*, vol. 10, no. 2, pp. e47–e92, 2014.
[3] N. Schuff, et al., "MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers," *Brain*, vol. 132, no. 4, pp. 1067–1077, 2009.

[4] X. Wu, et al., "The receiver operational characteristic for binary classification with multiple indices and its application to the neuroimaging study of Alzheimer's disease," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 10, no. 1, pp. 173–180, Jan./Feb. 2013.

[5] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Sci. Technol.*, vol. 19, no. 6, pp. 578–595, 2014.

[6] Y. Xiao, et al., "Patch-based label fusion segmentation of brainstem structures with dual-contrast MRI for Parkinson's disease," *Int. J. Comput. Assisted Radiol. Surg.*, vol. 10, no. 7, pp. 1029–1041, 2015.

[7] C. R. Jack, et al., "Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease," *Neurology*, vol. 49, no. 3, pp. 786–794, 1997.

[8] P. Visser, F. Verhey, P. Hofman, P. Scheltens, and J. Jolles, "Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment," *J. Neurology Neurosurgery Psychiatry*, vol. 72, no. 4, pp. 491–497, 2002.

[9] C. DeCarli, et al., "Qualitative estimates of medial temporal atrophy as a predictor of progression from mild cognitive impairment to dementia," *Archives Neurology*, vol. 64, no. 1, pp. 108–115, 2007.

[10] H. Braak and E. Braak, *Evolution of Neuronal Changes in the Course of Alzheimers Disease*. Berlin, Germany: Springer, 1998.

[11] C. Jack, et al., "Atrophy rates accelerate in amnestic mild cognitive impairment," *Neurology*, vol. 70, no. 19 Part 2, pp. 1740–1752, 2008.

[12] L. Clerx, et al., "Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment," *Neurobiology Aging*, vol. 34, no. 8, pp. 2003–2013, 2013.

[13] M. Chupin, et al., "Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI," *Hippocampus*, vol. 19, no. 6, pp. 579–587, 2009.

[14] 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.

[15] O. B. Ahmed, J. Benois-Pineau, M. Allard, C. B. Amar, and G. Catheline, and A. D. N. Initiative, "Classification of Alzheimer's disease subjects from MRI using hippocampal visual features," *Multimedia Tools Appl.*, vol. 74, no. 4, pp. 1249–1266, 2015.

[16] Q. Zhou, et al., "An optimal decisional space for the classification of Alzheimer's disease and mild cognitive impairment," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 8, pp. 2245–2253, Aug. 2014.

[17] S. Liu, et al., "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1132–1140, Apr. 2015.

[18] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[19] A. Griffa, P. S. Baumann, J.-P. Thiran, and P. Hagmann, "Structural connectomics in brain diseases," *Neuroimage*, vol. 80, pp. 515–526, 2013.

[20] Y. He, Z. Chen, and A. Evans, "Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease," *J. Neuroscience*, vol. 28, no. 18, pp. 4756–4766, 2008.

[21] 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.

[22] Y. Li, et al., "Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features," *Neurobiology Aging*, vol. 33, no. 2, pp. 427.e15–427.e30, 2012.

[23] D. Dai, H. He, J. T. Vogelstein, and Z. Hou, "Accurate prediction of AD patients using cortical thickness networks," *Mach. Vis. Appl.*, vol. 24, no. 7, pp. 1445–1457, 2013.

[24] W. Zheng, Z. Yao, B. Hu, X. Gao, H. Cai, and P. Moore, "Novel cortical thickness pattern for accurate detection of Alzheimer's disease," *J. Alzheimer's Disease*, vol. 48, no. 4, pp. 995–1008, 2015.

[25] M. Tuceryan and A. K. Jain, "Texture analysis," in *Handbook of Pattern Recognition and Computer Vision*, vol. 2. River Edge, NJ, USA: World Sci., 1993, pp. 207–248.

[26] M. F. Insana, R. F. Wagner, B. S. Garra, D. G. Brown, and T. H. Shawker, "Analysis of ultrasound image texture via generalized rician statistics," *Opt. Eng.*, vol. 25, no. 6, pp. 256 743–256 743, 1986.

[27] G. Castellano, L. Bonilha, L. Li, and F. Cendes, "Texture analysis of medical images," *Clinical Radiol.*, vol. 59, no. 12, pp. 1061–1069, 2004.

[28] S. Bauer, L.-P. Nolte, and M. Reyes, "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*. Berlin, Germany: Springer, 2011, pp. 354–361.

[29] M. Torabi, R. D. Ardekani, and E. Fatemizadeh, "Discrimination between Alzheimer's disease and control group in MR-images based on texture analysis using artificial neural network," in *Proc. IEEE Int. Conf. Biomed. Pharmaceutical Eng.*, 2006, pp. 79–83.

[30] J. Zhang, C. Yu, G. Jiang, W. Liu, and L. Tong, "3D texture analysis on MRI images of Alzheimer's disease," *Brain Imag. Behavior*, vol. 6, no. 1, pp. 61–69, 2012.

[31] 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.

[32] H. Xia and S. C. Hoi, "MKBoost: A framework of multiple kernel boosting," *IEEE Trans. Knowl. Data Eng*, vol. 25, no. 7, pp. 1574–1586, Jul. 2013.

[33] M. C. Carrillo, L. J. Bain, G. B. Frisoni, and M. W. Weiner, "Worldwide Alzheimer's disease neuroimaging initiative," *Alzheimer's Dementia*, vol. 8, no. 4, pp. 337–342, 2012.

[34] C. R. Jack, et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Resonance Imag.*, vol. 27, no. 4, pp. 685–691, 2008.

[35] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1998.

[36] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping*, vol. 17, no. 3, pp. 143–155, 2002.

[37] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.

[38] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.

[39] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[40] J. Zhang, L. Zhou, L. Wang, and W. Li, "Functional brain network classification with compact representation of sice matrices," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1623–1634, Jun. 2015.

[41] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[42] C. Chu, A.-L. Hsu, K.-H. Chou, P. Bandettini, C. Lin, and A. D. N. Initiative, "Does feature selection improve classification accuracy? impact of sample size and feature selection on classification using anatomical magnetic resonance images," *Neuroimage*, vol. 60, no. 1, pp. 59–70, 2012.

[43] F. Liu, H.-I. Suk, C.-Y. Wee, H. Chen, and D. Shen, "High-order graph matching based feature selection for Alzheimer's disease identification," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Berlin, Germany: Springer, 2013, pp. 311–318.

[44] J. Gao, et al., "A novel approach for lie detection based on F-score and extreme learning machine," *PloS One*, vol. 8, no. 6, 2013, Art. no. e64704.

[45] F. Liu, et al., "Multivariate classification of social anxiety disorder using whole brain functional connectivity," *Brain Struct. Function*, vol. 220, no. 1, pp. 101–115, 2015.

[46] Y.-W. Chen and C.-J. Lin, "Combining SVMs with various feature selection strategies," in *Feature Extraction*. Berlin, Germany: Springer, 2006, pp. 315–324.

[47] S. Sonnenburg, G. Rätsch, C. Schäfer, and B. Schölkopf, "Large scale multiple kernel learning," *J. Mach. Learn. Res.*, vol. 7, pp. 1531–1565, 2006.

[48] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," *J. Mach. Learn. Res.*, vol. 9, pp. 2491–2521, 2008.

[49] Z. Xu, R. Jin, I. King, and M. Lyu, "An extended level method for efficient multiple kernel learning," in *Proc. Advances Neural Inf. Process. Syst.*, 2009, pp. 1825–1832.

[50] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. no. 27.

[51] H.-I. Suk, S.-W. Lee, D. Shen, and A. D. N. Initiative, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," *NeuroImage*, vol. 101, pp. 569–582, 2014.
[52] L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim, F. Segovia, and A. D. N. Initiative, "Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images," *Neurocomputing*, vol. 151, pp. 139–150, 2015.

Jin Liu is working toward the PhD degree in the School of Information Science and Engineering, Central South University, Changsha, Hunan, P.R. China. His current research interests include medical image analysis, machine learning, and pattern recognition.

Min Li received the PhD degree in computer science from Central South University, China, in 2008. She is currently a professor in the School of Information Science and Engineering, Central South University, Changsha, Hunan, P.R. China. Her main research interests include bioinformatics and systems biology.

Wei Lan received the BSc degree from Henan Polytechnical University, in 2009, and the MSc degree from Guangxi University, China, in 2012. He is currently working toward the PhD degree in bioinformatics at Central South University. His current research interests include bioinformatics and data mining.

Fang-Xiang Wu (M'06-SM'11) received the BSc and MSc degrees in applied mathematics from the Dalian University of Technology, Dalian, China, in 1990 and 1993, respectively. He received the first PhD degree in control theory and its applications from Northwestern Polytechnical University, Xi'an, China, in 1998, and the second PhD degree in biomedical engineering from the University of Saskatchewan (U of S), Saskatoon, Canada, in 2004. During 2004-2005, he worked as a postdoctoral fellow with the Laval University Medical Research Center, Quebec City, Canada. He is currently a professor in the Division of Biomedical Engineering and the Department of Mechanical Engineering, U of S. His current research interests include computational and systems biology, genomic and proteomic data analysis, biological system identification and parameter estimation, and applications of control theory to biological systems. He has published more than 260 technical papers in refereed journals and conference proceedings. He is serving as an editorial board member of three international journals, the guest editor of several international journals, and as the program committee chair or member of several international conferences. He has also reviewed papers for many international journals. He is a senior member of the IEEE.

Yi Pan received the BEng and MEng degrees in computer engineering from Tsinghua University, China, in 1982 and 1984, respectively, and the PhD degree in computer science from the University of Pittsburgh, Pittsburgh, Pennsylvania, in 1991. He is a regents' professor of computer science and an interim associate dean and chair of biology with Georgia State University, Atlanta, Georgia. He joined Georgia State University in 2000 and was promoted to full professor in 2004, named a distinguished university professor in 2013, and designated a regents' professor (the highest recognition given to a faculty member by the University System of Georgia) in 2015. He served as the chair of the Computer Science Department from 2005-2013. He is also a visiting Changjiang chair professor with Central South University, China. His profile has been featured as a distinguished alumnus in both the *Tsinghua Alumni Newsletter* and the *University of Pittsburgh CS Alumni Newsletter*. His research interests include parallel and cloud computing, wire-less networks, and bioinformatics. He has published more than 330 papers including more than 180 SCI journal papers and 60 IEEE/ACM Transactions papers. In addition, he has edited/authored 40 books. His work has been cited more than 6,500 times. He has served as an editor-in-chief or editorial board member of 15 journals including seven IEEE Transactions. He received many awards including IEEE Transactions Best Paper Award, four other international conference or journal Best Paper Awards, four IBM Faculty Awards, two JSPS Senior Invitation Fellowships, IEEE BIBE Outstanding Achievement Award, NSF Research Opportunity Award, and AFOSR Summer Faculty Research Fellowship. He has organized many international conferences and delivered keynote speeches at more than 50 international conferences around the world.

Jianxin Wang received the BEng and MEng degrees in computer engineering from Central South University, China, in 1992 and 1996, respectively, and the PhD degree in computer science from Central South University, China, in 2001. He is the chair of and a professor in the Department of Computer Science, Central South University, Changsha, Hunan, P.R. China. His current research interests include algorithm analysis and optimization, parameterized algorithm, bioinformatics, and computer network. He is a senior member of the IEEE.