



## Identification of MCI individuals using structural and functional connectivity networks

Chong-Yaw Wee<sup>a</sup>, Pew-Thian Yap<sup>a</sup>, Daoqiang Zhang<sup>a</sup>, Kevin Denny<sup>b</sup>, Jeffrey N. Browndyke<sup>c,d,e</sup>, Guy G. Potter<sup>d</sup>, Kathleen A. Welsh-Bohmer<sup>c,d,f</sup>, Lihong Wang<sup>b</sup>, Dinggang Shen<sup>a,\*</sup>

<sup>a</sup> Image Display, Enhancement, and Analysis (IDEA) Laboratory, Biomedical Research Imaging Center (BRIC) and Department of Radiology, University of North Carolina at Chapel Hill, USA

<sup>b</sup> Brain Imaging and Analysis Center (BIAC), Duke University Medical Center, USA

<sup>c</sup> Joseph and Kathleen Bryan Alzheimer's Disease Research Center, Duke University Medical Center, USA

<sup>d</sup> Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, USA

<sup>e</sup> Duke Institute for Brain Sciences, Duke University, USA

<sup>f</sup> Department of Medicine, Division of Neurology, Duke University Medical Center, USA

### ARTICLE INFO

#### Article history:

Received 8 July 2011

Revised 15 September 2011

Accepted 5 October 2011

Available online 14 October 2011

#### Keywords:

Mild cognitive impairment (MCI)

Alzheimer's disease (AD)

Diffusion tensor imaging (DTI)

Resting-state functional magnetic resonance imaging (rs-fMRI)

Brain network analysis multiple-kernel Support Vector Machines (SVMs)

Multimodality representation

### ABSTRACT

Different imaging modalities provide essential complementary information that can be used to enhance our understanding of brain disorders. This study focuses on integrating multiple imaging modalities to identify individuals at risk for mild cognitive impairment (MCI). MCI, often an early stage of Alzheimer's disease (AD), is difficult to diagnose due to its very mild or insignificant symptoms of cognitive impairment. Recent emergence of brain network analysis has made characterization of neurological disorders at a whole-brain connectivity level possible, thus providing new avenues for brain diseases classification. Employing multiple-kernel Support Vector Machines (SVMs), we attempt to integrate information from diffusion tensor imaging (DTI) and resting-state functional magnetic resonance imaging (rs-fMRI) for improving classification performance. Our results indicate that the multimodality classification approach yields statistically significant improvement in accuracy over using each modality independently. The classification accuracy obtained by the proposed method is 96.3%, which is an increase of at least 7.4% from the single modality-based methods and the direct data fusion method. A cross-validation estimation of the generalization performance gives an area of 0.953 under the receiver operating characteristic (ROC) curve, indicating excellent diagnostic power. The multimodality classification approach hence allows more accurate early detection of brain abnormalities with greater sensitivity.

© 2011 Elsevier Inc. All rights reserved.

### Introduction

Alzheimer's disease (AD) is an irreversible neurodegenerative disease characterized by progressive cognitive and memory deficits, which are serious enough to interfere daily life. It has been reported that the prevalence of AD is expected to quadruple by year 2050, where 1 in every 85 persons will be affected by the disease (Brookmeyer et al., 2007). Thus, there is a pressing need for accurate diagnosis of AD at its early stage in order to reduce the burden of worldwide healthcare system.

Mild cognitive impairment (MCI), often a prodromal stage of AD, is a good target for early diagnosis and therapeutic interventions of AD. Recent studies show that individuals with MCI tend to progress to probable AD at a rate of approximately 10% to 15% per year (Grundman et al., 2004; Misra et al., 2009), compared with healthy controls who develop dementia at a rate of 1% to 2% per year (Bischkopf et al., 2002). Early

detection of MCI is of paramount importance for possible delay of the transition from MCI to AD. Nevertheless, MCI is difficult to diagnose due to its very mild symptoms of cognitive impairment. Most pattern classification algorithms proposed for MCI or early-AD identification achieve relatively inferior performance with low sensitivity rate.

Several imaging modalities have been applied for AD and MCI diagnosis. These modalities provide complementary information: diffusion tensor imaging (DTI) (Haller et al., 2010; Wee et al., 2011) provides local microstructural characteristics of water diffusion; structural magnetic resonance imaging (MRI) (Fan et al., 2008a; McEvoy et al., 2009) can be used to delineate brain atrophy; functional MRI (fMRI) (Machulda et al., 2009; Pihlajamäki and Sperling, 2008) characterizes hemodynamic response related to neural activity; positron emission tomography (PET) (Grady et al., 2003; Silveira et al., 2010) measures metabolic patterns. Most existing pattern classification methods, however, use only one of these modalities at a time for AD and MCI diagnosis. Greater effort should hence be focused on integration of different modalities since combining complementary information from the different imaging modalities can improve accuracy in disease diagnosis.

\* Corresponding author.

E-mail address: [dgshen@med.unc.edu](mailto:dgshen@med.unc.edu) (D. Shen).

To the best of our knowledge, to date, DTI and resting-state fMRI (rs-fMRI) have not been combined at a brain connectivity network level for identifying individuals with MCI, although they have been employed separately with reasonably good classification performance. In this study, we propose a high-dimensional multivariate classification framework to accurately identify individuals with MCI from those who undergo normal aging. The key idea of our approach involves employing multiple-kernel based support vector machine (SVM) algorithm to integrate anatomical and functional connectivity information extracted from DTI and rs-fMRI. The current study is the first attempt to integrate these two modalities to identify MCI individuals from cognitively normal controls. We seek to validate whether complementary structural and functional information can be combined to improve classification performance. We will also report brain regions that contribute most to classification. While confirming findings of previous studies, this paper sheds new light on the effectiveness of applying multimodality information for diagnosis of progressive neurodegenerative disorders.

The rest of the paper is organized as follows: **Method and materials** furnishes information on the image dataset, image acquisition protocols, and post-processing pipeline. This is followed by a comprehensive description of the multiple-kernel-SVM based multimodality classification framework. The proposed MCI classification framework is evaluated in **Experimental results**. Findings, methodological issues, and limitations of the proposed framework are discussed extensively in **Discussion**. **Conclusion** concludes this paper.

## Method and materials

### Subjects characterization and diagnosis

All subjects used in this study were recruited by the Duke-UNC Brain Imaging and Analysis Center (BIAC), Durham, North Carolina, USA. Informed consent was obtained from all participants, and the experimental protocols were approved by the institutional ethics board. This cohort involved of 27 participants, 10 individuals with MCI, and 17 socio-demographically matched normal controls. All the recruited subjects were diagnosed by expert consensus panels at the Joseph and Kathleen Bryan Alzheimer's Disease Research Center (Bryan ADRC) and the Department of Psychiatry at Duke University Medical Center. Diagnosis was based upon available data from a battery of general neurological examination, neuropsychological assessment evaluation, collateral and subject symptom and functional capacity reports. Conformation of diagnosis for MCI if subjects met the following inclusion criteria: 1) age >55 years and any race; 2) recent worsening of cognition, but still functioning independently; 3) Mini Mental State Examination (MMSE) score between 24 and 30; 4a) score  $\leq -1.5$  SD on at least two Bryan ADRC cognitive battery memory tests for single-domain amnesic MCI; or 4b) score  $\leq -1.5$  SD on at least one of the formal memory tests and score  $\leq -1.5$  SD on at least one other cognitive domain task (e.g., language, visuospatial-processing, or judgment/executive function) for multi-domain MCI; 5) 4 or lower for baseline Hachinski score; 6) does not meet the NINCDS-ADRDA or DSM-IV-TR criteria for dementia; 7) no psychological symptoms or history of depression; and 8) capacity to give informed consent and follow study procedures.

On the other hand, all healthy control individuals met the following criteria: 1) age >55 years and any race; 2) adequate visual and auditory acuity to properly complete neuropsychological testing; 3) no self-report of neurological or depressive illness; 4) shows no evidence of depression based on the Diagnostic Interview Schedule portion of the Duke Depression Evaluation Schedule; 5) normal score on a non-focal neurological examination; 6) a score  $> -1$  SD on any formal memory tests and a score  $> -1$  SD on any formal executive function or other cognitive test; and 7) demonstrates a capacity to give informed consent and follow study procedures. In order for safety purposes and minimizing biases, subjects were excluded from the study if they have: 1) any of the traditional MRI contraindications,

such as foreign metallic implants or pacemakers; 2) a past head injury or neurological disorder associated with MRI abnormalities, including dementia, brain tumors, epilepsy, Parkinson's disease, demyelinating diseases, etc.; 3) any physical or intellectual disability affecting completion of assessments; 4) documentation of other Axis I psychiatric disorders; and 5) any prescription medication (or nonprescription drugs) with known neurological effects. Noteworthy that the diagnosis of all cases were made on clinical grounds without any reference to MRI. Demographic information of the participants is shown in [Table 1](#).

### Data acquisition

Data acquisition was performed using a 3.0-Tesla GE Signa EXCITE scanner. Diffusion-weighted images of each participant were acquired axially parallel to the anterior and posterior commissures (AC–PC) line with twenty-five-direction diffusion-weighted whole-brain volumes using the following parameters:  $b = 0, 1000$  s/mm<sup>2</sup>, flip angle = 90°, TR/TE = 17000/78 ms, imaging matrix = 128 × 128, FOV = 256 × 256 mm<sup>2</sup>, resulting in a voxel dimension of 2 × 2 × 2 mm<sup>3</sup> reconstructed resolution. A total of 72 contiguous slices were acquired.

Resting-state functional images were acquired axially parallel to the horizontal plane connecting the AC–PC line with the following parameters: flip angle = 77°, TR/TE = 2000/32 ms, imaging matrix = 64 × 64, FOV = 256 × 256 mm<sup>2</sup>, resulting in a voxel resolution of 4 × 4 × 4 mm<sup>3</sup>. A total of 34 slices were acquired using a SENSE inverse-spiral pulse sequence in the same plane as the low resolution T1-weighted images. All the subjects were told to keep their eyes open and stare at a fixation cross in the middle of the screen during scanning, which lasted for 5 min. As we know, neurons get excited to changing stimuli across time. But when the stimuli such as the little cross sign in this study was presented steadily without changing across the five minute period, the neural excitation related to the stimuli can vanish quickly. Hence, this can ensure subjects not falling into sleep and avoid saccade-related activation which is unavoidable if eyes were closed.

### Method

Progressive degenerative neurological diseases such as Alzheimer's disease and similar dementias exhibit subtle, yet spatially and temporally diffuse pathology, where the brain is damaged in a large-scale, highly connected network, rather than in one single isolated region. In view of this, designing an informative description of interregional connections, which might be more sensitive in conveying the pathological information, is necessary for accurate diagnosis or prediction of neurological diseases.

It has been proven that information extracted from different modalities and different brain tissues might complement each other and provides more comprehensive characterization of brain abnormalities (Dukart et al., 2011; Fan et al., 2008b; Fjell et al., 2010; Jiang et al., 2008; Zhang et al., 2011). To integrate complementary information from different modalities and different brain tissue types, we proposed in this paper an approach that combines structural connectivity information from white matter (WM) regions and functional connectivity information from gray matter (GM) regions to distinguish individuals

**Table 1**  
Demographic information of the participants involved in this study.

Group	MCI	Normal
No. of subjects	10	17
Gender (M/F)	5 M/5 F	8 M/9 F
Age (mean ± SD)	74.2 ± 8.6	72.1 ± 8.2
Years of education (mean ± SD)	17.7 ± 4.2	16.3 ± 2.4
MMSE (mean ± SD)	28.4 ± 1.5	29.4 ± 0.9

with MCI from normal controls. The key of our multimodality classification framework includes:

- More informative description of the brain via multiple connectivity networks;
- Integration of DTI and rs-fMRI information at a whole-brain connectivity level;
- Multimodality data fusion using multiple-kernel SVM.

An overview of the proposed MCI classification framework is summarized graphically in Fig. 1.

#### Processing of DTI

The DTI images were first parcellated into 90 regions (45 for each hemisphere) by propagating the Automated Anatomical Labeling (AAL) ROIs (Tzourio-Mazoyer et al., 2002) to each image using a deformable DTI registration algorithm called F-TIMER (Yap et al., 2009, 2010) with tensor orientation corrected using the method described in (Xu et al., 2003). Whole-brain streamline fiber tractography was then performed on each image using ExploreDTI (Leemans et al., 2009). Whole-brain streamline fiber tractography was performed on each DTI image using ExploreDTI (Leemans et al., 2009), with minimal seed point fractional anisotropy (FA) of 0.45, stopping FA of 0.25, minimal fiber length of 20 mm, and maximal fiber length of 400 mm. The number of fibers passing through each pair of regions was counted. Two regions were considered as anatomically connected if fibers passing through their respective masks were present, giving us the connection topology of the network.

On top of the fiber count based connectivity network, averages of on-fiber FA, mean diffusivity (MD) and principal diffusivity values were computed to form another 5 connectivity networks with the same topology but conveying different biophysical properties (Wee et al., 2011). The fiber count network provides information on the number of fibers connecting a pair of regions. The FA network provides information on the degree of anisotropy along the fibers, the MD network provides information on the average diffusivity along fibers, the  $\lambda_1$  network provides information on the on-fiber average axial diffusivity (also called the longitudinal diffusivity); and the  $\lambda_2$  and  $\lambda_3$  networks provide information on the on-fiber average radial diffusivities (diffusivities in directions perpendicular to the axonal direction). By including these different biophysical properties, subtle information of WM atrophy caused by pathological disorder can be extracted and used to improve classification performance. It has been reported in a recent paper (Wee et al., 2011) that this informative description is capable of significantly improving classification accuracy compared with simple univariate description. Examples of the constructed connectivity maps are shown in Fig. 2.

#### Processing of resting-state fMRI (rs-fMRI)

Post-processing of the rs-fMRI images, such as slice timing correction and head-motion correction was performed using the Statistical Parametric Mapping software package (SPM8, <http://www.fil.ion.ucl.ac.uk/spm>). To ensure magnetization equilibrium, the first 10 acquired fMRI images of each subject were discarded. The remaining 140 images were first corrected for the acquisition time delay among different slices before they were realigned to the first volume of the remaining images for head-motion correction. We hypothesize that the variability of BOLD signal of GM regions is sensitive in delineating the alteration of connectivity patterns, caused by pathological attacks of MCI. Removing signal from the ventricles and WM is motivated by the fact that these regions contain a relatively high proportion of noise caused by the cardiac and respiratory cycles (Van Dijk et al., 2010). Accordingly, we first segmented the T1-weighted image of each subject into GM, WM and CSF. For each subject, the GM was then used to mask the fMRI images. This procedure eliminated the possible contribution from WM and CSF in the re-fMRI time series.

The first scan of fMRI time series was coregistered to the T1-weighted image of same subject. The estimated transformation was then applied

to other fMRI scans of the same subject. Deformation fields were estimated by warping the AAL (Tzourio-Mazoyer et al., 2002) template (T1-weighted image) to the subject T1-weighted images using a deformable registration method called HAMMER (Shen and Davatzikos, 2002). The brain space of each subject was then parcellated into 90 ROIs by warping the AAL region masks to the subject space using the estimated deformation fields. For each subject, the mean time series of each individual ROI was computed by averaging the GM-masked fMRI time series over all voxels in that particular ROI.

One crucial step in rs-fMRI analysis is temporal band-pass filtering. The frequency interval of band-pass filtering varies and depends on the application, but is normally within the interval of [0.01 – 0.10 Hz] since the fMRI dynamics of neuronal activities are most salient within this frequency interval. It provides a reasonable trade-off between avoiding the physiological noise associated with higher frequency oscillations (Cordes et al., 2001) and the measurement error associated with estimating very low frequency correlations from limited time series (Achard et al., 2008).

The analysis of rs-fMRI is normally performed on full spectrum of the filtered signals – a relatively global analysis which might not be sensitive enough to delineate complex yet subtle pathological patterns related to the neurological disease. Such global analysis on BOLD signal might cause local, subtle temporal changes to be averaged out, and thus deteriorate classification performance. A relatively local analysis, which is more sensitive to BOLD signal changes, is hence required.

In order to extract complex, yet subtle pathological influences of MCI, we employed a multi-spectrum characterization of the regional mean time series, which utilizes multiple frequency sub-bands, in contrast to the conventional full-spectrum description, to construct functional connectivity networks. The GM-masked mean time series of each region was band-pass filtered within frequency interval  $[0.025 \leq f \leq 0.100\text{Hz}]$  before it was decomposed into five distinct, equally divided frequency sub-bands using the Fast Fourier transform (FFT), enabling a relatively frequency specific analysis of the regional mean time series. By using this multi-spectral characterization, a relatively local analysis, which is more sensitive in delineating complex yet subtle pathological patterns related to the neurological disease, can be achieved.

Functional connectivity, which represents interregional correlations in neuronal variability, was measured using pairwise Pearson correlation coefficients between the ROI pairs. Given a set of  $N$  random variables, the Pearson correlation matrix is a symmetric matrix in which each off-diagonal element is the correlation coefficient between a pair of variables. We considered the brain regions as a set of nodes and the correlation coefficients as signed weights represented by the edges connecting the nodes. Fisher's  $r$ -to- $z$  transformation was applied on the elements of the Pearson correlation matrix to improve the normality of the correlation coefficients as

$$z = \frac{1}{2} [\ln(1+r) - \ln(1-r)] \quad (1)$$

where  $r$  is the Pearson correlation coefficient and  $z$  is approximately a normal distribution with standard deviation  $\sigma_z = 1/\sqrt{N-3}$ . The functional connectivity networks are represented in the form of  $z$ -maps. Examples of the constructed functional connectivity maps for a normal control (NC) and an individual with MCI are shown in the top and bottom rows of Fig. 3, respectively.

#### Feature extraction and feature selection

The weighted local clustering coefficient, a measure that quantifies the cliquishness of the nodes, is extracted from all connectivity maps as

$$f(p) = 2 \times \frac{\sum_{q: q \neq p \in \mathcal{C}(p, q)} Z(p, q)}{k_p [k_p - 1]}, \quad (2)$$

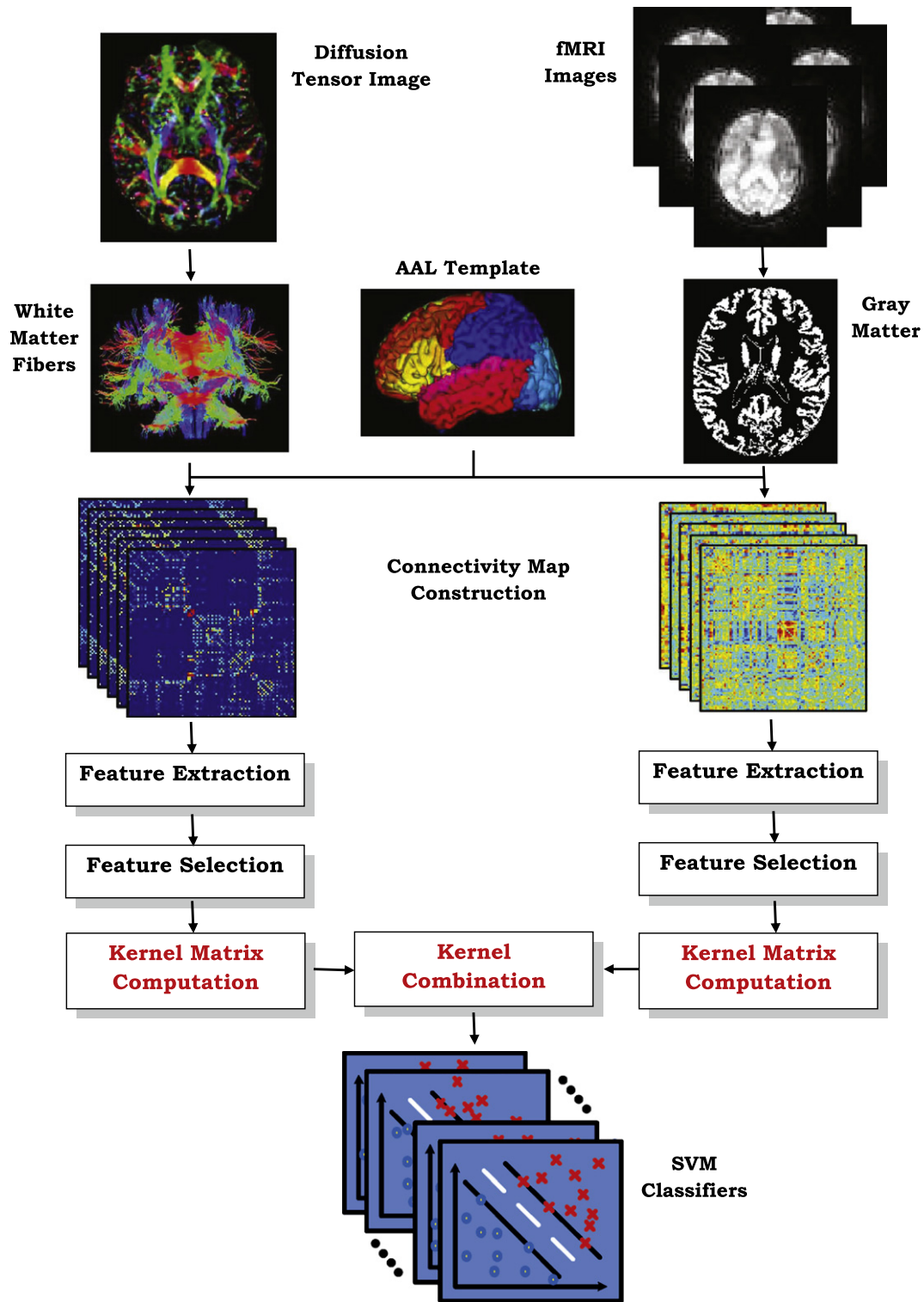


Fig. 1. Schematic diagram illustrating the proposed classification framework.

where  $\zeta$  is the subnetwork comprising of  $k_p$  nodes directly connected to the  $p$ -th node, and  $z(p, q)$  is the edge weight between the  $p$ -th ROI and  $q$ -th ROI. Hence, a total of 90 features can be obtained from each connectivity map, producing for each subject a pool of ( $6 \times 90 = 540$ ) and ( $5 \times 90 = 450$ ) features for DTI and fMRI modalities, respectively.

Statistical  $t$ -test was performed to select the most discriminative features for classification. Features with  $p$ -values smaller than a

predefined threshold will be selected to construct the individual kernel matrix for each imaging modality.

#### Multiple-kernel SVM

Given  $n$  training samples with  $\mathbf{X}_i = \{\mathbf{x}_{i,1}, \dots, \mathbf{x}_{i,D}\}$  denoting the feature vector of the  $i$ -th sample ( $D$  = number of connectivity maps,  $d = 1, \dots, D$  and  $\mathbf{x}_{i,d} = \{f_{i,d}(1), \dots, f_{i,d}(90)\}$ ),  $y_i \in \{-1, 1\}$  denoting the



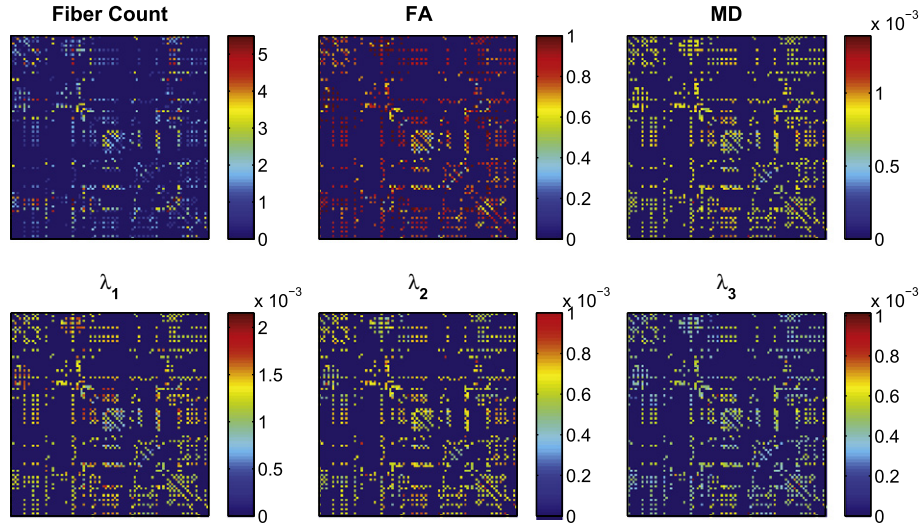


Fig. 2. Connectivity maps constructed from various DTI measures.

corresponding label, the primal optimization problem of a conventional single kernel SVM is given as

$$\begin{aligned} \min_{\mathbf{w}, b, \xi_i} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i, \\ \text{s.t.} \quad & y_i \times (\mathbf{w}^T \phi(\mathbf{X}_i) + b) \geq 1 - \xi_i \\ \text{with} \quad & \xi_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (3)$$

where  $\xi_i$ ,  $C$ ,  $b$ ,  $\mathbf{w}$  and  $\phi(\cdot)$  denote the distance of the  $i$ -th misclassified observation from its correct side of the margin, the model parameter that controls the amount of constraint violations introduced by  $\xi_i$ , the bias term, the normal vector of hyperplane, and the kernel-induced mapping function, respectively.

By applying the kernel approach, Eq. (3) is normally solved using its dual form

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j \times k(\mathbf{X}_i, \mathbf{X}_j), \\ \text{s.t.} \quad & k(\mathbf{X}_i, \mathbf{X}_j) = \phi(\mathbf{X}_i)^T \phi(\mathbf{X}_j); \quad \sum_{i=1}^n \alpha_i y_i = 0; \\ \text{with} \quad & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \end{aligned} \quad (4)$$

where  $\alpha$  is the Lagrange multiplier and  $k(\mathbf{X}_i, \mathbf{X}_j)$  is the kernel function for training samples,  $\mathbf{X}_i$  and  $\mathbf{X}_j$ .

In order to integrate the biomarkers from  $M$  modalities, multiple-kernel SVM version of primal optimization problem can be written as

$$\begin{aligned} \min_{\mathbf{w}^{(m)}, b, \xi_i} \quad & \frac{1}{2} \sum_{m=1}^M \beta_m \|\mathbf{w}^{(m)}\|^2 + C \sum_{i=1}^n \xi_i, \\ \text{s.t.} \quad & y_i \left[ \sum_{m=1}^M \beta_m \left( (\mathbf{w}^{(m)})^T \phi^{(m)}(\mathbf{X}_i^{(m)}) + b \right) \right] \geq 1 - \xi_i \\ \text{with} \quad & \xi_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (5)$$

where  $\beta_m \geq 0$  denotes the weighting factor on the  $m$ -th modality. Accordingly, the corresponding dual form can be written as

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j \sum_{m=1}^M \beta_m k^{(m)}(\mathbf{X}_i^{(m)}, \mathbf{X}_j^{(m)}), \\ \text{s.t.} \quad & k^{(m)}(\mathbf{X}_i^{(m)}, \mathbf{X}_j^{(m)}) = \phi(\mathbf{X}_i^{(m)})^T \phi(\mathbf{X}_j^{(m)}); \quad \sum_{i=1}^n \alpha_i y_i = 0; \\ \text{with} \quad & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \end{aligned} \quad (6)$$

where  $k^{(m)}(\mathbf{X}_i^{(m)}, \mathbf{X}_j^{(m)})$  is the kernel function for the  $m$ -th modality.

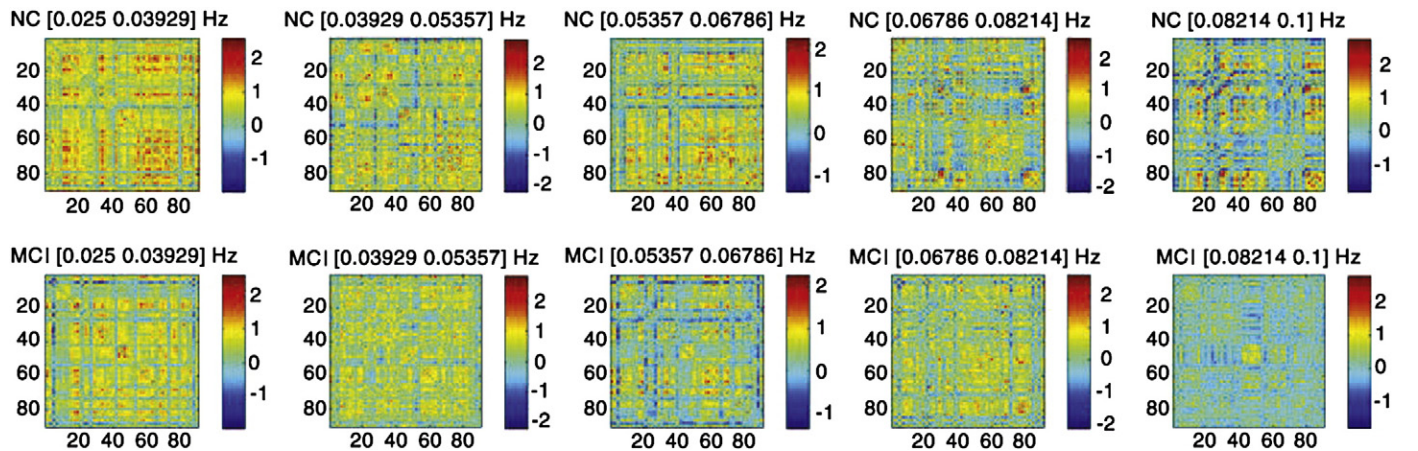


Fig. 3. Multi-spectral functional connectivity maps for NC (top) and MCI (bottom), respectively.

Given a new test sample  $\mathbf{X} = \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(M)}\}$ , the decision function for the predicted label can be determined as

$$F(\mathbf{X}) = \text{sign}\left(\sum_{i=1}^n y_i \alpha_i \sum_{m=1}^M \beta_m k^{(m)}(\mathbf{X}_i^{(m)}, \mathbf{X}^{(m)}) + b\right). \quad (7)$$

The multiple-kernel SVM can be naturally embedded into the conventional single kernel SVM framework by noting  $k(\mathbf{X}_i, \mathbf{X}_j) = \sum_m \beta_m k^{(m)}(\mathbf{X}_i^{(m)}, \mathbf{X}_j^{(m)})$  as a mixed-kernel between the multimodality training samples  $\mathbf{X}_i$  and  $\mathbf{X}_j$ , and  $k(\mathbf{X}_i, \mathbf{X}) = \sum_m \beta_m k^{(m)}(\mathbf{X}_i^{(m)}, \mathbf{X}^{(m)})$  as a mixed-kernel between the multimodality training sample between  $\mathbf{X}_i$  and the test sample  $\mathbf{X}$ .

#### SVM training and classification

A linear kernel based SVM classifier based on the LIBSVM library (Chang and Lin, 2001) was adopted in the current framework based on mixed-kernel. Before we computed the kernel between pairs of feature vectors, we first performed a normalization step on each feature vector to obtain unit norm vector (i.e.,  $\|\mathbf{x}\|_2 = 1$ ). In fact, linear kernel with this normalization step can be regarded as a “normalized linear kernel” and is defined as

$$k(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^T \cdot \mathbf{y}}{(\|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2)}; \text{ s.t. } \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1, \quad (8)$$

where  $T$  denotes matrix transpose,  $(\mathbf{x}, \mathbf{y})$  denotes the training set of instance-label pairs,  $\|\cdot\|_2$  denotes  $l_2$ -norm.

In order to obtain an unbiased estimation of the generalization performance of the complete framework and at the same time select the optimal SVM model, a nested cross-validation scheme was employed. Algorithmically, this can be described as two nested loops of cross-validation. Hyperparameter as part of the SVM model is determined in the inner loop of cross-validation. The selected models are then evaluated in terms of their generalization ability using an independent validation set. This procedure is repeated until all subjects have been left out. We note that in each run of the outer-loop cross-validation loop, the selected model and hence the hyperparameter can be different.

#### Summary of methodology

The proposed multiple-kernel SVM based multimodality classification framework is summarized as follows.

1. In the DTI modality, whole-brain streamline tractography is performed using the 90-region AAL atlas to construct structural connectivity network based on the number of fibers passing through each pair of regions. Based on the common fibers connecting each region pair, on-fiber average FA, MD, and principal diffusivities are derived to construct another five connectivity networks.
2. For each network, the clustering coefficients for each ROI are determined to compute a total of 90 features. Feature vectors from all networks of each subject were concatenated to form a long feature vector to provide an informative description of WM connections finally resulting in a total of 540 features.
3. In the fMRI processing, the first 10 fMRI images are discarded to cater for magnetization equilibrium of each subject. Slice timing correction and intra-subject registration of the acquired rs-fMRI time series are performed on the remaining images using the SPM software package to correct the slice acquisition time delay and head motion.
4. T1-weighted image of each subject is segmented and its GM regions are then used to mask the acquired fMRI time series. The brain space is further parcellated into 90 regions using the AAL template.

5. The mean time series of each ROI, which is obtained by averaging all voxels within the GM-masked ROI, is band-pass filtered ( $0.025 \leq f \leq 0.100\text{Hz}$ ) before it is further decomposed into five equally-divided, distinct frequency sub-bands. Five functional connectivity networks are constructed using pairwise Pearson correlation coefficients, one for each frequency sub-band.
6. For each network, the local clustering coefficients for all ROIs are determined as features for classification. All clustering coefficients are concatenated to form a long feature vector consisting of 450 features.
7. For each modality, features with significant discriminative power are selected using  $t$ -tests.
8. Individual kernel matrices are then constructed from the selected features of each modality before they are integrated to form a single mixed-kernel matrix.
9. The constructed mixed-kernel is employed to train SVM classifiers in a leave-one-out fashion using the LIBSVM solver.

#### Experimental results

As with common practice, the classifiers are evaluated based on the classification accuracy and the area under ROC curve (AUC). Classification accuracy measures the effectiveness of predicting the true class label. The area under receiver operating characteristic (ROC) curve measures the probability that when one positive and one negative samples are drawn at random, the decision function assigns a higher value to the positive than to the negative sample.

In addition, we also employed other statistical measures to evaluate the diagnostic power of the various methods used for comparison. The Youden's index (Y), Balanced Accuracy (BAC) and F-score (F) are defined respectively as (Sokolova et al., 2006)

$$Y = \text{Sensitivity} + \text{Specificity} - 1 = \frac{TP}{FP} + \frac{TN}{FN} - 1, \quad (9)$$

$$BAC = \frac{1}{2} \times \left[ \frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right], \quad (10)$$

$$F = 2 \times \left[ \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \right], \quad (11)$$

where

$$\text{precision} = \frac{TP}{TP + FP}; \text{ recall} = \frac{TP}{TP + FN}$$

with TP, TN, FP and FN denoting true positive, true negative, false positive and false negative, respectively. F-score is a composite measure that favors algorithms with higher sensitivity and challenges those with higher specificity. Youden's index evaluates the ability of the algorithm to avoid failure by equally weighting its performance on positive and negative samples.

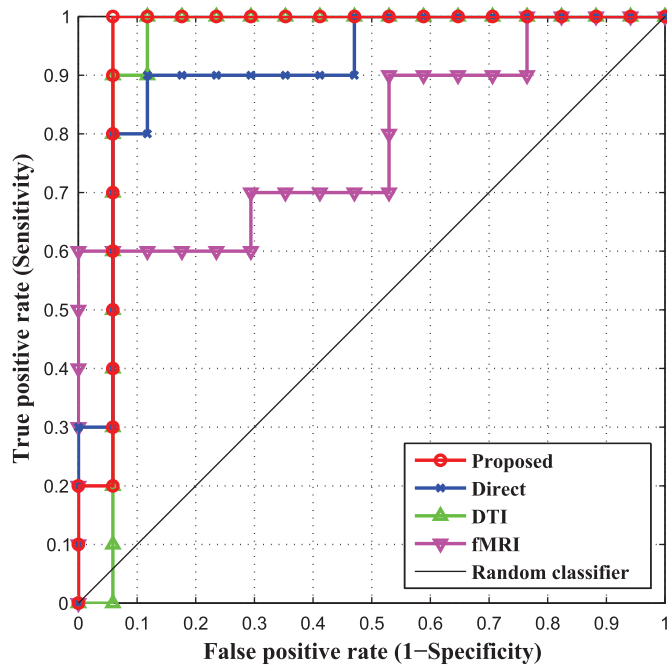
#### Classification based on multimodality connectivity maps

The proposed multiple-kernel SVM based multimodality classification approach was compared with the single modality approach and the direct data fusion method. In the single modality approach, only features selected from a single imaging modality (DTI or rs-fMRI) were applied for kernel matrix construction. In the direct data fusion method, all 990 features, including DTI and rs-fMRI features, were first concatenated into a long vector before feature selection and kernel matrix construction. In the multiple-kernel approach, the optimal weighting factor,  $\beta_m$ , was determined via grid search. Similarly, in all experiments parameters such as  $p$ -value for  $t$ -test feature selection method and SVM soft margin constraint parameter,  $C$ , were also determined via grid search.

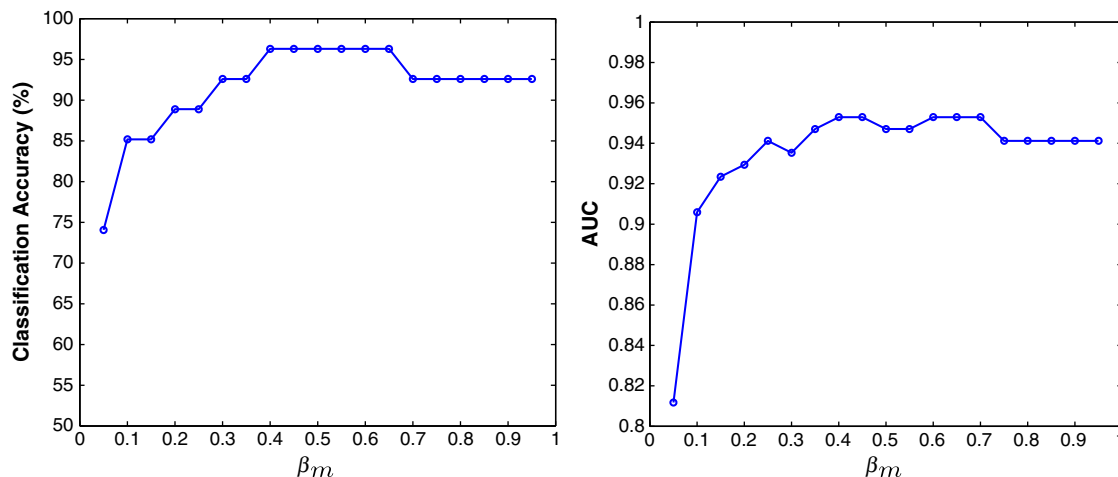
**Table 2**

Classification performance of single and multimodality connectivity networks. (ACC = Accuracy; SEN = SENSitivity; SPE = SPECificity).

Method	ACC (%)	SEN(%)	SPE(%)	AUC	Y	BAC	F
fMRI	70.37	70.00	70.59	0.7882	0.4059	0.7029	0.6364
DTI	88.89	80.00	94.12	0.9353	0.7412	0.8706	0.8421
Direct	88.89	90.00	88.24	0.9118	0.7824	0.8912	0.8571
Proposed	96.30	100.00	94.12	0.9529	0.9412	0.9524	0.9706

**Fig. 4.** ROC curves.

Comparison was performed via leave-one-out cross-validation due to the limited number of available samples. Classification performance for identification of individuals with MCI using single and multimodality connectivity networks is summarized in Table 2. The proposed method yields a classification accuracy of 96.30%, which is an increment of at least 7.41% from that of the single modality approach and the direct data fusion method. It also outperformed all other methods in the rest of the computed statistical indices.

**Fig. 5.** Performance of the multi-kernel SVM based multimodality classification with respect to the weighting factor,  $\beta_m$ . (Left: Classification accuracy; Right: AUC).

A cross-validation estimation of the generalization performance shows an area of 0.9529 under the ROC curve (AUC), indicating excellent diagnostic power. ROC curves for all compared methods are shown in Fig. 4.

#### Effect of weighting factor, $\beta_m$

The weighting factor,  $\beta_m$ , determines the contribution of each modality in multimodality classification as formulated in Eqs. (5), (6) and (7). A larger  $\beta_m$  value indicates larger contribution of DTI to the classification and vice versa. In this experiment, we seek to investigate how classification performance varies with respect to  $\beta_m$ . We fixed the constraint parameter,  $C$  to 3.0, and  $p$ -value to 0.01 for both modalities for feature selection. The classification accuracy and AUC value of the proposed method are shown graphically in Fig. 5.

It is observed that higher classification accuracy is achieved when larger  $\beta_m$  value is used. Accuracy of more than 90% can be achieved with  $\beta_m$  equal or larger than 0.3. The highest accuracy values are achieved within the range of  $0.40 \leq \beta_m \leq 0.65$ . Similar trends can be observed for AUC values. High and consistent AUC values are obtained over a relatively wide range of  $0.35 \leq \beta_m \leq 0.70$ . This indicates that the proposed method is relatively robust to the weighting factor. The skewness of the accuracy and AUC graphs toward larger  $\beta_m$  values indicates dominance of structural information carried by the DTI modality in the classification.

#### Effect of constraint parameter, $C$

We investigate the influence of different values of the constraint parameter,  $C$ , on the classification performance. We fixed the weighting factor,  $\beta_m$  to 0.6 and varied the constraint parameter,  $C$ , from 1 to 10, in steps of 1. The  $p$ -value used for feature selection was maintained as 0.01 for both modalities. The classification accuracies and AUC values with respect to  $C$  are shown in Fig. 6.

Classification performance is relatively robust to  $C$ . Variation of constraint parameter,  $C$ , does not significantly affect the performance of the proposed method. In addition, we also tested our method using several much larger  $C$  values, i.e., 50 and 100. The outcome again showed that the proposed method consistently maintains high classification accuracy, despite the high  $C$  values.

#### Effect of predefined threshold for feature selection

In the proposed framework, the statistical  $t$ -test is applied to select a subset of features with most discriminative power. Features with their  $p$ -value smaller than a predefined threshold are retained for

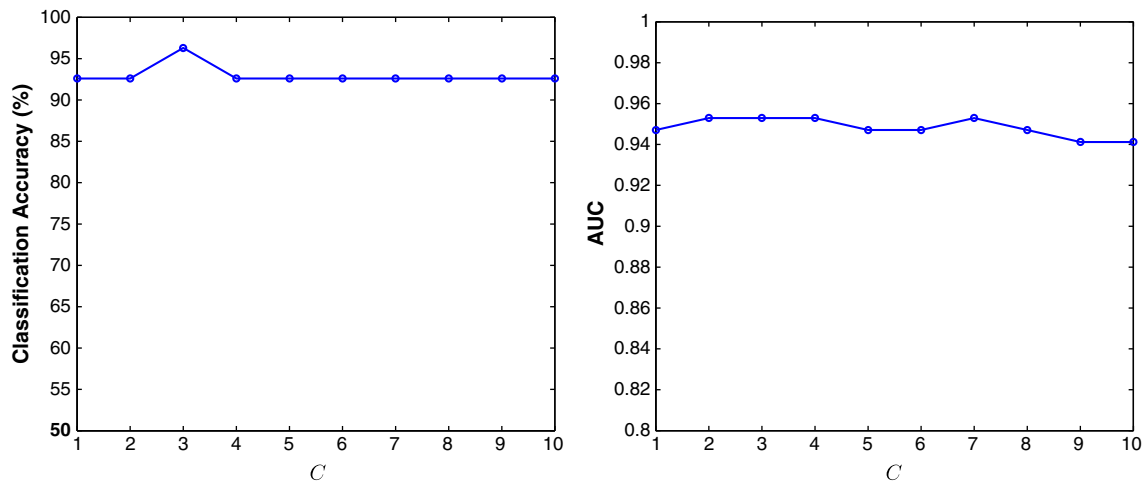


Fig. 6. Performance of the multi-kernel SVM based multimodality classification with respect to the constraint parameter,  $C$ . (Left: Classification accuracy; Right: AUC).

construction of SVM classifiers. Thus, the number of features that will be selected is determined by the predefined threshold. In this subsection, an experiment is performed to explore the robustness of the proposed framework over the number of features selected. The performance is summarized in Fig. 7.

A high classification performance is consistently observed with the variation of  $p$ -value (variation in the number of selected features). Classification accuracy is more than 88.0% for most of  $p$ -values. For large  $p$ -values, classification accuracy deteriorates because more features, including some redundant and confounding features, are selected for constructing the classifier. The proposed framework shows a very consistent AUC value, always larger than 0.92, over the  $p$ -values, indicating excellent diagnostic power. The numbers of features selected for different  $p$ -values are provided in Table 3.

#### The most discriminant regions

The most discriminant regions that are selected during classifier construction include the orbitofrontal cortex, temporal pole, anterior and posterior cingulate gyrus, frontal gyrus, amygdala, precuneus, thalamus, parahippocampal gyrus and insula, in line with results reported in previous studies. Regions which are selected by  $t$ -test feature selection method for mixed kernel construction are shown in Table 4.

The top twelve selected regions used for multimodality classification are graphically shown in Fig. 8.

#### Classification performance using linear and nonlinear kernel functions

In this subsection, we compared the classification performance of our proposed framework using the linear and nonlinear kernel-based SVMs. The best classification performance of the linear kernel as well as two other nonlinear kernels, i.e., radial basis function (RBF) and polynomial, is provided after choosing the optimal parameters via grid search. The polynomial kernel used in this study is a homogeneous kernel defined as

$$k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \cdot \mathbf{y})^d, \quad (12)$$

where  $T$  denotes matrix transpose,  $(\mathbf{x}, \mathbf{y})$  denotes the training set of instance-label pairs and  $d$  denotes the polynomial degree. The best classification performance achieved for all compared methods is shown in Table 5.

It can be observed that the performance of polynomial kernel is the same as the linear kernel. In fact, we found that the best performance is found to be when  $d = 1$ . RBF kernel shows significant improvement in classification performance if compared to the linear

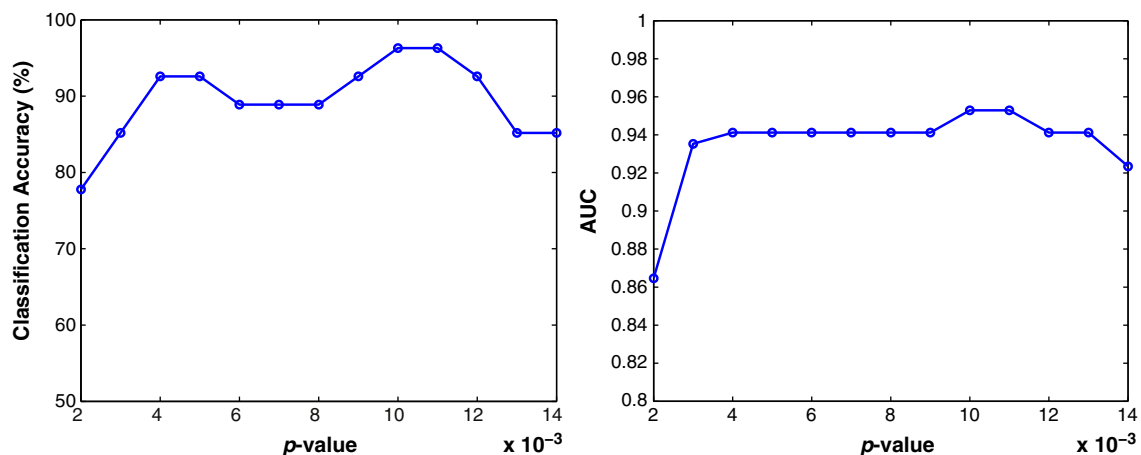


Fig. 7. Performance of the multi-kernel SVM based multimodality classification with respect to the  $p$ -value. (Left: Classification accuracy; Right: AUC).



**Table 3**Numbers of features selected for different  $p$ -values. (LOO = leave-one-out).

LOO	$p$ -value											
	0.004		0.006		0.008		0.010		0.012		0.014	
	DTI	fMRI	DTI	fMRI	DTI	fMRI	DTI	fMRI	DTI	fMRI	DTI	fMRI
1	7	1	7	7	7	16	7	24	7	26	7	28
2	7	0	7	3	7	10	7	15	8	22	8	24
3	7	5	7	10	7	13	7	20	7	24	7	31
4	6	8	7	17	7	21	7	25	7	30	7	36
5	5	4	7	10	7	14	7	15	8	20	8	22
6	7	2	7	6	7	10	7	12	8	12	8	15
7	7	0	7	3	7	7	7	8	7	12	7	15
8	7	0	7	3	7	7	7	11	8	14	8	20
9	7	0	7	4	7	8	7	10	8	13	8	16
10	7	0	7	1	7	3	8	10	8	14	8	20
11	7	0	7	2	7	6	7	10	8	10	8	12
12	7	5	7	10	7	14	7	17	7	20	7	25
13	6	1	7	3	8	4	9	6	10	11	11	14
14	7	2	7	5	7	10	7	17	7	18	8	20
15	7	2	7	4	7	7	7	11	7	16	7	21
16	7	0	7	1	7	3	7	6	7	8	8	10
17	8	0	8	0	8	1	8	3	8	5	9	9
18	7	0	8	0	8	2	9	3	10	5	12	9
19	7	0	7	1	7	3	7	7	7	9	8	12
20	7	5	7	9	7	12	7	14	7	18	7	19
21	7	1	7	5	7	8	7	10	7	11	7	13
22	7	0	7	1	7	4	7	7	7	9	8	10
23	7	5	7	8	7	18	7	20	7	29	8	32
24	7	2	7	4	7	8	8	17	8	22	8	27
25	7	0	8	0	8	0	8	2	8	4	8	5
26	6	0	7	0	7	3	7	4	8	5	8	13
27	7	0	7	0	7	0	2	7	4	7	5	5

and polynomial kernels. However, the proposed normalized linear kernel performs better than the RBF kernel. We adopted the normalized linear kernel in this study since choosing the appropriate value of  $\sigma$  for RBF kernel is very challenging in practice.

## Discussion

### Significance of results

Single modality of biomarkers has been widely employed for AD/MCI diagnosis and prediction, often achieving relatively high performance. Single-modality-based methods rely on simpler scanning protocols and hence require less image acquisition effort, making them

more clinically feasible. However, information from different imaging biomarkers complements to each other and potentially improves prediction accuracy. Integration of biomarkers of different modalities with different data fusion methods for diseases diagnosis and prediction is still an open area of research. Our study demonstrates that multimodality connectivity network analysis can be employed to distinguish individuals with MCI from normal controls with high accuracy.

It is well known that misclassifying a healthy person to be a patient may be troublesome, however, misclassifying a patient to be a healthy person may cause severe consequences. If a patient is mis-diagnosed as a healthy person, necessary treatments to delay or cure the disease may not be provided on time during the critical treatment period. This will accelerate the progression of disease from mild to severe, a point where no effective treatment is available, eventually causing death of patient. Hence, it is tremendously crucial for a classifier to alleviate such circumstance by providing high sensitivity to the disease. Our results show an improvement in sensitivity when both imaging modalities are combined through the multiple-kernel SVM algorithm. Statistically, our method performed better than any single modality method due to the improvement of sensitivity while maintaining high specificity. Furthermore, our method shows a relatively stable and robust performance with respect to the modality weighting factor and SVM constraint parameter.

The regions selected in the course of classification by our method are in agreement with previous studies and include the prefrontal cortex and orbitofrontal cortex (Grady et al., 2003), temporal pole (Nobili et al., 2008), anterior and posterior cingulate gyrus, precuneus and insula (Davatzikos et al., 2011), amygdala (Dai et al., 2009), thalamus (Wang et al., in press), parahippocampal gyrus (Celone et al., 2006; Machulda et al., 2009; Pihlajamäki and Sperling, 2008), inferior temporal left, superior temporal gyrus right and precentral gyrus right (Lenzi et al., 2011), inferior temporal left, inferior frontal gyrus left, insula and superior temporal gyrus (Han et al., 2010).

### Methodological issues/limitations

The performance of our method may be affected by the unbalanced data between individuals with MCI and normal controls that were used for SVM model construction. This is typically the case in real world disease diagnosis, where patient samples are normally more difficult to collect, depending on the prevalence of the disease. A classifier will normally try to adapt itself for better prediction of the majority class; this will increase the overall classification accuracy. Although the sensitivity has been improved by using multimodality connectivity networks, the proposed framework at its current stage is not designed to handle this issue. Full investigation focusing on handling unbalanced data will be our future work.

Whole-brain connectivity network analysis, a relative new neuroimaging approach for disease diagnosis and prediction, is still facing some fundamental problems. One of the problems is the lack of standardized approaches for brain parcellation, resulting in network nodes that are often defined either anatomically or randomly (Zalesky et al., 2010). The scale of parcellation can affect network properties such as clustering coefficient, global and local efficiencies, path length, small-worldness, etc.

There are many other imaging modalities that have been used for AD/MCI classification and prediction, as mentioned in Introduction. Integration of imaging modalities other than DTI and fMRI will be explored in the future to further improve the classification performance based on the proposed framework. We also plan to apply our method to distinguish between MCI subjects who progressed to AD, and those who remained stable for a certain number of years.

Our current study is limited by the small sample and statistical power is hence a potential concern. Leave-one-out cross-validation, as employed in this study, provides an optimistic estimate of the classification accuracy since all except one of the subjects are used to

**Table 4**

Regions selected during classification.

Regions
Rectus gyrus right
Insula right
Orbitofrontal cortex (medial) left
Temporal pole (superior) right
Orbitofrontal cortex (medial) right
Amygdala right
Posterior cingulate gyrus left
Posterior cingulate gyrus right
Anterior cingulate gyrus left
Temporal pole (superior) left
Orbitofrontal cortex (superior) right
Anterior cingulate gyrus right
Middle frontal gyrus right
Inferior frontal gyrus (opercular) right
Precentral gyrus right
Precuneus left
Thalamus right
Temporal pole (middle) left
ParaHippocampal gyrus left
Superior occipital gyrus left

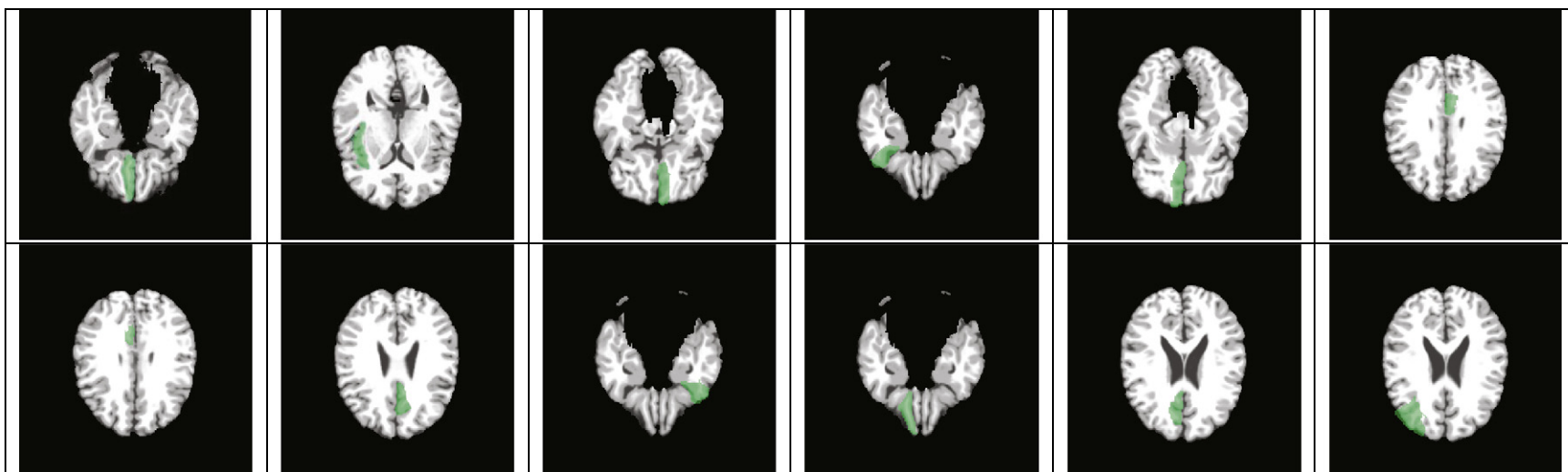


Fig. 8. Top twelve selected regions used for classification.

**Table 5**  
Classification performance of linear and nonlinear kernels based SVM.

Kernel	ACC (%)	AUC
linear	85.16	0.8941
polynomial	85.16	0.8941
RBF	92.59	0.9353
Proposed	96.30	0.9529

train the classifier. Hence, almost all information available in the dataset is used in classifier model construction. For other classification approaches such as  $k$ -fold cross-validation, only  $N-k$  ( $N$  is the total number of subjects in the dataset) subjects are included during training process, and hence the performance obtained may be less optimistic to the small dataset used in this study. However, estimation of the classification accuracy using  $k$ -fold cross-validation might be more precise when compared to leave-one-out cross-validation provided that there are sufficient data to accurately train the classifier. Our dataset is quite diverse, and it includes both sexes and all ages between 55 to 84 for individuals with MCI and 55 to 88 for normal controls. However, the results obtained have to be verified in the future with larger datasets to reduce individual effects and to validate the effectiveness of the proposed technique.

## Conclusion

This study investigated the diagnostic power of multimodality connectivity networks derived from the DTI and rs-fMRI for identifying individuals with MCI from normal controls. In this framework, multiple modalities integration was achieved via a multiple-kernel SVM algorithm. SVM classifiers were trained using a mixed kernel that was constructed from the individual kernel of each modality. This framework improves classification performance and hence justifies the hypothesis that the DTI and rs-fMRI contain complementary information, and each of them is indispensable particularly for achieving better diagnostic power. High sensitivity provided by our framework ensures that necessary treatments can be provided to individuals with MCI during critical treatment period and hence possibly delays the progression of MCI to AD or other cognitive impairment related dementias. The promising results shed new light on the effectiveness of applying multimodality information for diagnosis of progressive neurodegenerative disorders such as AD.

## Financial disclosures

There are no conflicts of interest including any financial, personal, or other relationships with people or organizations for any of the co-authors related to the work described in the article.

## Acknowledgments

This work was supported in part by NIH grants EB006733, EB008374, EB009634, MH088520, NIA L30-AG029001, P30 AG028377-02, K23-AG028982, as well as a National Alliance for Research in Schizophrenia and Depression Young Investigator Award (LW).

## References

- Achard, S., Bassett, D.S., Meyer-Lindenberg, A., Bullmore, E., 2008. Fractal connectivity of long-memory networks. *Phys. Rev. E* 77, 036104.
- Bischkopf, J., Busse, A., Angermeyer, M.C., 2002. Mild cognitive impairment – a review of prevalence, incidence and outcome according to current approaches. *Acta Psychiatr. Scand.* 106, 403–414.
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K., Arrighi, H.M., 2007. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* 3, 186–191.
- Celone, K.A., Calhoun, V.D., Dickerson, B.C., Atri, A., Chua, E.F., Miller, S.L., DePeau, K., Rentz, D.M., Selkoe, D.J., Blacker, D., Albert, M.S., Sperling, R.A., 2006. Alterations

- in memory networks in mild cognitive impairment and Alzheimer's disease: an independent component analysis. *J. Neurosci.* 40, 10222–10231.
- Chang, C.C., Lin, C.J., 2001. LIBSVM: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E., 2001. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *Am. J. Neuroradiol.* 22, 1326–1333.
- Dai, W., Lopez, O.L., Carmichael, O.T., Becker, J.T., Kuller, L.H., Gach, H.M., 2009. Mild cognitive impairment and Alzheimer disease: patterns of altered cerebral blood flow at MR imaging. *Radiology* 250, 856–866.
- Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.N., Trojanowski, J.Q., 2011. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol. Aging* 32, 2322.e19–2322.e27.
- Dukart, J., Mueller, K., Horstmann, A., Barthel, H., Möller, H.E., Villringer, A., Sabri, O., Schroeter, M.L., 2011. Combined evaluation of FDG-PET and MRI improves detection and differentiation of dementia. *PLoS One* 6, e18111.
- Fan, Y., Batmanghelich, N., Clark, C.M., Davatzikos, C., the Alzheimer's Disease Neuroimaging Initiative, 2008a. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *NeuroImage* 39, 1731–1743.
- Fan, Y., Resnick, S.M., Wu, X., Davatzikos, C., 2008b. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. *NeuroImage* 41, 277–285.
- Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M., for the Alzheimer's Disease Neuroimaging Initiative, 2010. CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *J. Neurosci.* 30, 2088–2101.
- Grady, C.L., McIntosh, A.R., Beig, S., Keightley, M.L., Burian, H., Black, S.E., 2003. Evidence from functional neuroimaging of a compensatory prefrontal network in Alzheimer's disease. *J. Neurosci.* 23, 986–993.
- Grundman, M., Petersen, R.C., Ferris, S.H., Thomas, R.G., Aisen, P.S., Bennett, D.A., et al., 2004. Mild cognitive impairment can be distinguished from Alzheimer's disease and normal aging for clinical trials. *Arch. Neurol.* 61, 59–66.
- Haller, S., Nguyen, D., Rodriguez, C., Emch, J., Gold, G., Bartsch, A., Lovblad, K., Giannakopoulos, P., 2010. Individual prediction of cognitive decline in mild cognitive impairment using support vector machine-based analysis of diffusion tensor imaging data. *J. Alzheimers Dis.* 22, 315–327.
- Han, Y., Wang, J., Zhao, Z., Min, B., Lu, J., Li, K., He, Y., Jia, J., 2010. Frequency-dependent changes in the amplitude of low-frequency fluctuations in amnesic mild cognitive impairment: a resting-state fMRI study. *NeuroImage* 55, 287–295.
- Jiang, T., Liu, Y., Shi, F., Shu, N., Liu, B., Jiang, J., Zhou, Y., 2008. Multimodal magnetic resonance imaging for brain disorders: advances and perspectives. *Brain Imaging Behav.* 2, 249–257.
- Leemans, A., Jeurissen, B., Sijbers, J., Jones, D.K., 2009. ExploreDTI: A graphical toolbox for processing, analyzing, and visualizing diffusion MR data. 17th Annual Meeting of Intl Soc Mag Reson Med, p. 3537.
- Lenzi, D., Serra, L., Perri, R., Pantano, P., Lenzi, G., Paulesu, E., Caltagirone, C., Bozzali, M., Macaluso, E., 2011. Single domain amnesic MCI: a multiple cognitive domains fMRI investigation. *Neurobiol. Aging* 32, 1542–1557.
- Machulda, M.M., Senjem, M.L., Weigand, S.D., Smith, G.E., Ivnik, R.J., Boeve, B.F., Knopman, D.S., Petersen, R.C., Jack Jr., C.R., 2009. Functional MRI changes in amnesic and non-amnesic MCI during encoding and recognition tasks. *J. Int. Neuropsychol. Soc.* 15, 372–382.
- McEvoy, L.K., Fennema-Notestine, C., Roddey, J.C., Hagler Jr., D.J., Holland, D., Karow, D.S., Pung, C.J., Brewer, J.B., Dale, A.M., 2009. Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology* 251, 195–205.
- Misra, C., Fan, Y., Davatzikos, C., 2009. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. *NeuroImage* 44, 1414–1422.
- Nobili, F., Salmaso, D., Morbelli, S., Girtler, N., Piccardo, A., Brugnolo, A., Dessi, B., Larsson, S.A., Rodriguez, G., Pagani, M., 2008. Principal component analysis of FDG PET in amnesic MCI. *Eur. J. Nucl. Med. Mol. Imaging* 35, 2191–2202.
- Pihlajamäki, M., Sperling, R.A., 2008. fMRI: use in early Alzheimer's disease and in clinical trials. *Futur. Neurol.* 3, 409–421.
- Shen, D., Davatzikos, C., 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE Trans. Med. Imaging* 21 (1421–1), 439.
- Silveira, M., Marques, J., The Alzheimer's Disease Neuroimaging Initiative, 2010. Boosting Alzheimer disease diagnosis using PET images. 2010 International Conference on Pattern Recognition, International Association for Pattern Recognition (IAPR), Istanbul, Turkey, pp. 2556–2559.
- Sokolova, M., Japkowicz, N., Szpakowicz, S., 2006. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. *AI 2006: Advances in Artificial Intelligence*, pp. 1015–1021.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289.
- Van Dijk, K.R.A., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties and optimization. *J. Neurophysiol.* 103, 297–321.
- Wang, Z., Jia, X., Liang, P., Qi, Z., Yang, Y., Zhou, W., Li, K., in press. Changes in thalamus connectivity in mild cognitive impairment: evidence from resting state fMRI. *Eur. J. Radiol.*
- Wee, C.Y., Yap, P.T., Li, W., Denny, K., Brownhyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D., 2011. Enriched white matter connectivity networks for accurate identification of mci patients. *NeuroImage* 54, 1812–1822.

- Xu, D., Mori, S., Shen, D., van Zijl, P.C.M., Davatzikos, C., 2003. Spatial normalization of diffusion tensor fields. *Magn. Reson. Med.* 50, 175–182.
- Yap, P.T., Wu, G., Zhu, H., Lin, W., Shen, D., 2009. Fast tensor image morphing for elastic registration. *MICCAI* 2009, pp. 721–729.
- Yap, P.T., Wu, G., Zhu, H., Lin, W., Shen, D., 2010. F-TIMER: fast tensor image morphing for elastic registration. *IEEE Trans. Med. Imaging* 29, 1192–1203.
- Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore, E.T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage* 50, 970–983.
- Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., The Alzheimers Disease Neuroimaging Initiative, 2011. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage* 55, 856–867.