Magnetic Resonance Imaging 34 (2016) 252–263

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.mrijournal.com

Original contribution

Feature-ranking-based Alzheimer's disease classification from structural MRI

Iman Beheshti *, Hasan Demirel, for The Alzheimer's Disease Neuroimaging Initiative [1]

Biomedical Image Processing Lab, Department of Electrical & Electronic Engineering, Eastern Mediterranean University, Gazimagusa, Mersin 10, Turkey

ARTICLE INFO

Article history:
Received 11 May 2015
Revised 25 August 2015
Accepted 29 November 2015

Keywords:
Alzheimer's disease
Support vector machine
Feature ranking
Fisher Criterion
Data fusion
Voxel-based morphometry

ABSTRACT

High-dimensional classification approaches have been widely used to investigate magnetic resonance imaging (MRI) data for automatic classification of Alzheimer's disease (AD). This paper describes the use of t-test based feature-ranking approach as part of a novel feature selection procedure, where the number of top features is determined using the Fisher Criterion. The proposed classification system involves five systematic levels. First, voxel-based morphometry technique is used to compare the global and local differences of gray matter in patients with AD versus healthy controls (HCs). The significant local differences in gray matter volume are then selected as volumes of interests (VOIs). Second, the voxel clusters are employed as VOIs, where each voxel is considered to be a feature. Third, all the features are ranked using t-test scores. In this regard, the Fisher Criterion between the AD and HC groups is calculated for a changing number of ranked features, where the vector size maximizing the Fisher Criterion is selected as the optimal number of top discriminative features. Fourth, the classification is performed using support vector machine. Finally, data fusion methods among atrophy clusters are used to improve the classification performance. The experimental results indicate that the performance of the proposed system could compete well with the state-of-the-art techniques reported in the literature.

1. Introduction

Alzheimer's disease (AD), an irreversible neurodegenerative dementia that occurs most frequently in older adults, gradually destroys regions of the brain that are responsible for memory, learning, thinking, and behavior [1]. Current estimates indicate that 5.3 million Americans of all ages will suffer from AD in 2015. This number is expected to increase to 16 million people by 2050. AD is the only disease among the top ten causes of death in Americans that cannot be cured, prevented, or slowed. [1] Presently, no cure exists for AD, but early detection may aid in determining the root of AD mechanisms and improve the quality of life for patients who suffer from AD [1]. In recent years, the analysis of neuroimaging data has attracted much interest, given the recent improvements in early and accurate detection of AD [2,3]. Among the several available neuroimaging modalities, magnetic resonance imaging (MRI) is more widely used in AD related studies because of its excellent spatial resolution, high availability, good contrast, and the lack of a requirement for the radioactive pharmaceutical injection that is needed with positron emission tomography (PET) or single photon emission computed tomography (SPECT) [4–7]. Recently, several studies have used biomarkers to classify AD based on structural MRI [8–15], which can be utilized to specify brain atrophy; functional MRI [16–18], which can be employed to describe hemodynamic response relevant to neural activity; diffusion tensor imaging [19–21], which can be used for local microstructural characteristics of water diffusion; and functional/structural connectivity [22–24], which can be used to characterize neurological disorders in the whole brain at the connectivity level. In this paper, we focused only on AD classification using structural MRI. Atrophy measured by structural MRI is a powerful biomarker of the stage and intensity of the neurodegenerative aspect of AD pathology [25]. Several studies have used structural MRI feature extraction for AD classification. These studies are variously based on morphometric methods [26–28], region of interest (ROI)/volume of interest (VOI) [29–31], gray matter voxels in the automatic segmentation of images [32], and structural MRI measurement of the hippocampus and the medial temporal lobe [33–39]. Despite the recent improvements in early detection of AD, the prediction of disease progression using structural MRI alone remains challenging and requires more investigation. The present study describes the use of a statistical feature ranking approach using t-test as part of a novel feature selection process. The number of highest ranking features selected is determined by using the

* Corresponding author. Tel.: +90 392 630 1301x1093; fax: +90 392 63000201648.
E-mail addresses: iman.beheshti@cc.emu.edu.tr (I. Beheshti), hasan.demirel@emu.edu.tr (H. Demirel).

[1] Data used in this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). ADNI investigators other than those listed above contributed to the study design, implementation or data provision but did not participate in the analyses or writing of this report. The complete listing of ADNI investigators is available at http://www.loni.ucla.edu/ADNI/Data/ADNI_Authorship_List.pdf.

http://dx.doi.org/10.1016/j.mri.2015.11.009
0730-725X/

Fisher Criterion, which maximizes the class separation between AD and HC groups.

The Fisher Criterion aids in finding an optimal number of features with the most discriminative information for the classification process. The proposed feature selection method is applied to different atrophy clusters of voxels, which correspond to the volumes of interest (VOIs) in the gray matter of the MRI obtained through the voxel-based morphometry (VBM) analysis in the preprocessing. In this context, data fusion is introduced to increase the classification performance, which utilizes a majority-voting-based score fusion and a feature vector concatenation-based source fusion. In the proposed system, we use only MRI data, unlike several recent studies where MRI is combined with other different data such as PET, Cognitive Scores, and Mini Mental State Examination (MMSE) to increase the classifier performance [8,12,40,41]. The proposed system is accomplished by the systematic use of several ideas at five levels. At the first level, the VBM technique is employed to analyze group-wise comparisons between cross-sectional structural MRI scans, in order to find the MRI voxels that are best discriminated between the AD group and the HC group [14,42–44]. The inter-subject registration of the MRI images is promoted by employing the Diffeomorphic Anatomic Registration Through Exponentiated Lie algebra algorithm (DARTEL) [44]. This algorithm provides precise, accurate localization of structural damage of the MRI images [43,44]. Based on the VBM plus DARTEL approach, the overall and regional structural gray matter alterations are investigated to define regions with significant atrophy of gray matter in the patients who suffer from AD. The results obtained from 68 patients with AD, when compared to 68 HCs, show significant gray matter decline in right/left hippocampuses and in the inferior parietal and anterior cingulate regions in patients with AD. Instead of making a single global classifier, the multiple individual classifiers based on atrophy clusters obtained using VBM plus DARTEL analysis are proposed for use with data fusion techniques for more accurate classification. Based on these clusters, five different VOIs are defined as follows: 1) VOI_1 includes the right hippocampus region, 2) VOI_2 includes the left hippocampus region, 3) VOI_3 contains the right inferior parietal lobule region, 4) VOI_4 includes the right anterior cingulate region, and 5) VOI_{all} contains an accumulation of all atrophy cluster regions. At the second level, specified VOIs are used as 3D masks to extract voxel values from the VOIs to generate raw feature vectors. These raw feature vectors can be used in the data selection processes before use by the classifiers. At the third level, the extracted features are systematically ranked, based on the t-test values of the respective features obtained from the training set. The t-test can be considered as a statistical indicator showing the level of separation/discrimination between two groups (AD and HC) in the training set. For this reason, ranking according to the t-test, followed by the use of a subset of highest ranking features, would increase the classification performance. The t-test feature ranking has been used successfully in a number of pattern recognitions studies [45–47]. In addition, an automatic approach based on the Fisher Criterion is proposed to determine the number of top features. This approach adaptively determines the optimum number of top features and identifies a discriminative subset of high performance features based on training data in each fold, instead of using a fixed number of features. At the fourth level, the performance of the proposed feature selection technique is evaluated using support vector machine (SVM) classifiers. In the present work, the SVM classifier with both linear (linear SVM) and nonlinear (RBF SVM) kernels is trained to discriminate between the classes. In the final level, data fusion techniques among atrophy clusters (VOIs) are proposed to increase the overall performance. Data fusion improves the classification performance by integrating data (vectors, classifiers) from different atrophy clusters. To this purpose, source and score data fusion techniques were used to achieve higher performance. A direct comparison shows that the experimental results using the proposed t-test feature selection and data fusion-based approach indicate superior performance when compared to classifiers that use all raw features and a data reduction method involving principal component analysis (PCA). In summary, the aim of this study was to introduce a novel and automatic statistical feature selection method based on the combination of t-test feature ranking and the Fisher Criterion of the VOI, which can be considered a lower-dimensional feature vector representation of sMRI. The dimensionality of the feature vector can be adjusted by maximizing the Fisher Criterion in the training data-set. The proposed feature selection method not only selects the top discriminative features but also reduces the dimensionality of the input vectors to feature vectors. In addition, data fusion techniques are used to improve the AD classification performance among gray matter atrophy clusters. The performance of the proposed system is tested on 136 subjects (including 68 AD and 68 HC) from an ADNI dataset using 10-fold cross validation. The experimental results, when compared to those obtained with state-of-the-art techniques, show that the proposed system is highly competitive in terms of accuracy (96.32%), specificity (98.52%), and AUC (99.93%) for AD classification. The rest of this paper is arranged as follows: Section two gives the statistics for the data used in this work. Section three describes the proposed methodology for the design of an automatic, high performance AD classification system. The experimental results, discussion, and analysis of the proposed system in comparison to the state-of-the-art classification methods are given in section four. Section five presents some conclusions.

2. Material

2.1. Image acquisition

MR images and data used in this work were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). The MRI scans were acquired using a 3 T, T1-weighted scanner (Siemens) with Acquisition Plane = SAGITTAL, Acquisition Type = 3D, Coil = Phased Arrays(PA), Flip Angle = 9.0°, Matrix X/Y/Z = 240.0 pixels/256 pixels/176 pixels, Mfg Model = Skyra, Pixel Spacing X/Y = 1.0 mm/1.0 mm, Pulse Sequence = Gradient Recalled(GR)/Inversion Recovery(IR), Slice Thickness = 1.2 mm, and Echo Time (TE)/Inversion Time (TI)/Repetition Time (TR) = 2.98 ms/900 ms/2300 ms.

2.2. Subjects

The diagnostic classification was conducted by selecting a total of 136 subjects from the ADNI database and grouping them as AD and HC. The AD group contained 68 subjects ranging in age from 61.4 to 89.2 (74.33 ± 6.41) years. The Mini Mental State Examination (MMSE) and Clinical Dementia Ratio (CDR) scores ranged from 15 to 25 (mean 22.83 ± 2.65) and 0.5 to2 (mean 0.75 ± 0.41), respectively. The HC group contained 68 healthy controls ranging in age from 60.8 to 84.4 (74.14 ± 4.95) years. The MMSE ranged from 28 to 30(mean 29.38 ± 0.71) and the CDR was zero. A direct comparison revealed that the AD patients' mean MMSE and CDR were significantly distinct when compared to the HC subjects. No significant group differences were noted in age or sex ratio. Details of the demographics and clinical characteristics of the sample used in this paper are presented in Table 1.

3. Proposed AD classification system

This section proposed a new AD classification system using a novel approach based on a combination of t-test feature ranking and the Fisher Criterion for the optimal selection of feature vectors for high performance MRI classification of AD. The system involves five levels of

Table 1
Demographic and clinical details of the patients with AD and HC subjects.

	AD (n = 68)	HC (n = 68)	t-value	M.D
Age	74.33 ± 6.41	74.14 ± 4.95	0.19	0.18[NS]
MMSE	22.83 ± 2.56	29.38 ± 0.71	14.76	−6.5*
CDR[0/0.5/1/2]	0.75 ± 0.41 [0/44/19/5]	0.0 ± 0 [68/0/0/0]	−20.26	0.75*

Note: All data presented in mean ± standard deviation mode. AD, Alzheimer's Disease patients; CDR, Clinical Dementia Rating; HC, Healthy Control patients; MMSE, Mini-Mental State Examination; MD, Mean Difference; NS, Non-Significant.
* $P < 0.0001$.

processing. The pipeline of the proposed system is illustrated in Fig. 1. First, the VBM plus DARTEL approach is employed to perform pre-processing on 3D MRI data. Second, a feature extraction method is used, based on VBM plus DARTEL analysis. Third, the extracted features are ranked based on the t-test values of the respective features, in the training set. In addition, an automatic approach based on the Fisher Criterion is adopted to determine the number of top ranking features. This approach adaptively determines the optimum number of top features and identifies a discriminative subset of high performance features based on training data in each fold. Hence, the feature vectors taken from VOIs of high dimensional s-MRI data are reduced into a low dimensional space, with improved discrimination capability. Fourth, the proposed technique is evaluated using state-of-the-art SVM classifiers. The performance analysis comprises an experimental setup based on 136 samples from the ADNI dataset. A 10-fold cross validation is employed throughout the performance analysis, which implies having 122 (90%) samples in the training and 14(10%) samples in the testing processes in each iteration. Finally, data fusion techniques among atrophy clusters are engaged to improve the classification performance.

3.1. MRI data preprocessing and statistical analysis

The MR images are pre-processed using the Statistical Parameter Mapping (SPM) software version 8 (Welcome Trust Centre for Neuroimaging, London, UK; available at: http://www.fil.ion.ucl.ac.uk/spm) and the VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm), implemented in MATLAB R2014a. VBM is an automated technique for assessment of the whole brain structure with voxel-by-voxel comparisons, developed to analyze tissue concentrations or volumes between subject groups for distinguishing degenerative diseases with dementia [42,43]. In more detail, VBM techniques investigate structural differences in areas with poorly defined structural landmarks (e.g., prefrontal areas) and provide explorative analysis of structural differences [48–50]. Recently, VBM has been applied to detect early atrophic changes in AD [44,51–53]. It can provide statistical results for comparisons of patients with AD and HCs [44]. The inter-subject alignment of the MRI images was increased by applying the DARTEL approach, which has been reported to optimize the sensitivity of this type of analysis over standard VBM by using the Levenberg–Marquardt strategy [49,54–57]. Moreover, the VBM8 toolbox benefits from the unified segmentation model with a maximum a posterior (MAP) technique [58] and partial volume estimation (PVE) to account for partial volume effects [59], which results in a more subtle segmentation of subcortical areas. In addition, the VBM toolbox uses a spatially adaptive nonlocal means (SANLM) filter for denoising and removal of MRI in homogeneities [60]. The signal-to-noise ratio is improved by employing a spatial constraint based on a classical Markov random field (MRF) model [61]. Registration to a standard MNI-space (http://www.mni.mcgill.ca/) consists of a linear affine transformation and a nonlinear deformation using high-dimensional DARTEL normalization [55].

In the current work, sample homogeneity prior to calculating 2nd level analyses is ensured by inspecting the quality of gray matter images using the VBM8 toolbox. All MR images are corrected for bias field in homogeneities and then they are normalized and segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). The normalized and segmented images are modulated using a nonlinear deformation. In this work, only GM images are used. Finally, the 8 mm full-width-half-maximum (FWHM) Gaussian kernel is used for spatial smoothing of the GM images. After spatial pre-processing, the normalized, smoothed, modulated, DARTEL-warped gray matter datasets are analyzed using a voxel-wise parametric mapping. The absolute threshold masking of around 0.1 is used to avoid possible edge effects around the border between gray matter and white matter or CSF.

The regional gray matter volume changes are generated by voxel-based analysis over the whole brain. The framework of the

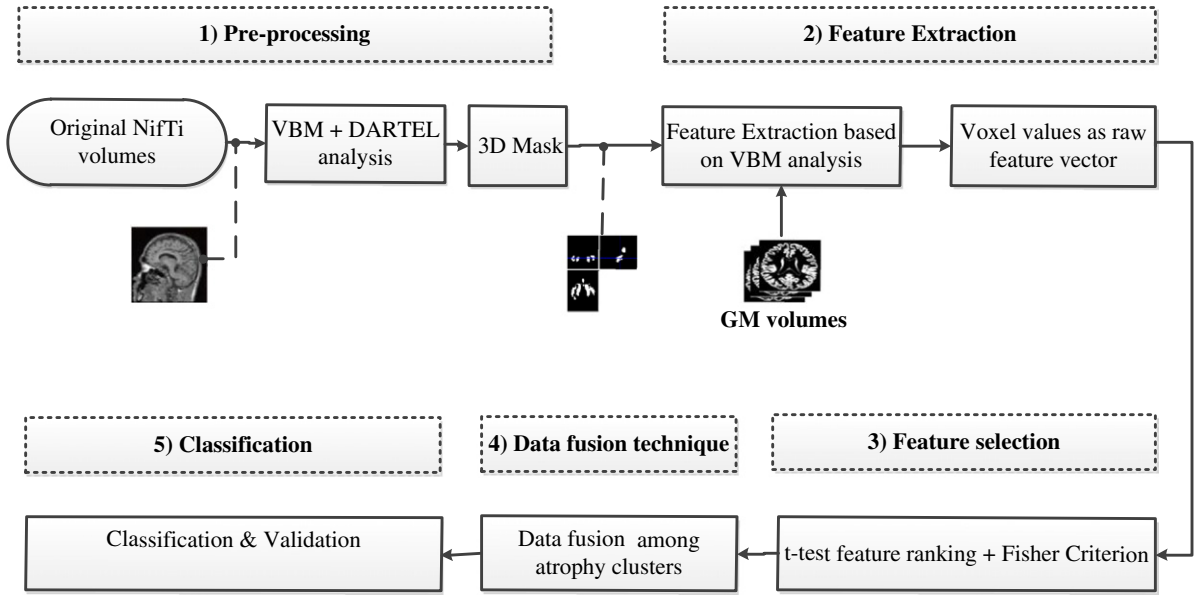

Fig. 1. The pipeline of proposed system for classifying AD.

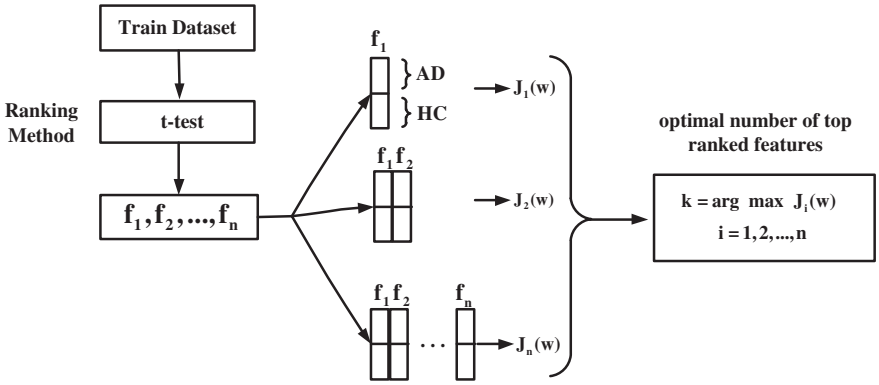

Fig. 2. Schematic representation of proposed feature selection approach.

general linear model is employed to detect gray matter volume changes in patients with AD using voxel-wise two sample t-test in SPM8. Age is engaged into the matrix design as a nuisance variable. The whole brain analysis is implemented using significance set at a P value of <0.01, with correction for family-wise error (FWE) and a minimum cluster size of 1400 voxels for two-sample comparisons. Between-group differences in demographics and clinical parameters among or between groups of this work are evaluated using an independent two-sample t-test with the SPSS 16.0 package. (http://www.spss.com/). $P < 0.05$ is set as the level of significance.

3.2. Feature extraction

The feature extraction procedure based on VBM plus DARTEL analysis is applied to isolate the VOIs. The brain regions that show significantly decreased gray matter volumes, obtained using VBM plus DARTEL analysis, in AD patients relative to HC are segmented using 3D masks. For the segmented regions, the MarsBaR region of interest toolbox is employed (http://marsbar.sourceforge.net/) to generate cluster-specific binary masks. The center coordinates of each mask are defined by the local maximum revealed by VBM plus DARTEL analysis on the whole brain. These masks are applied to all the smoothed gray matter density volumes resulting from the VBM plus DARTEL analysis, to extract voxel values as raw feature vectors.

3.3. Feature selection

The dimensionality of raw feature spaces in the VBM extracted s-MRI voxel features is very high in comparison to the number of samples. The feature vectors span a very small region in the high dimensional vector space; consequently, a feature selection mechanism is desired in the post-processing. Feature selection can be considered in the form of a standard dimensionality reduction via a standard method, such as PCA. Alternatively, feature selection can be considered in the form of choosing the most discriminative subset of the available features in the raw feature vector. In this context, the proposed method can be employed, as it is the combination of t-test feature ranking and the Fisher Criterion, which not only reduces the dimensionality, but also increases the discriminability.

3.3.1. PCA dimensionality reduction

Principal component analysis is a statistical dimensionality reduction method that extracts a set of orthogonal principal components (PCs) from an original dataset [62,63]. In this work, a 10-fold cross validation is used for measuring the performance of the

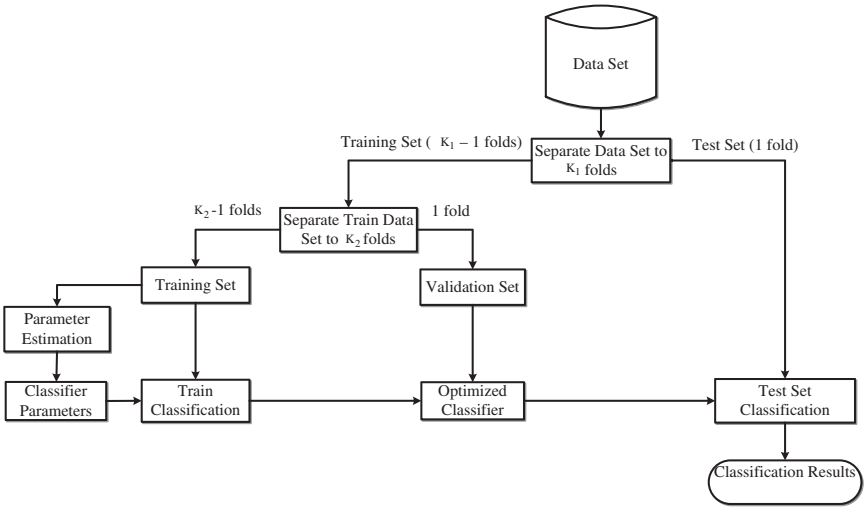

Fig. 3. 10-fold cross validation method used for parameter tuning and performance testing.

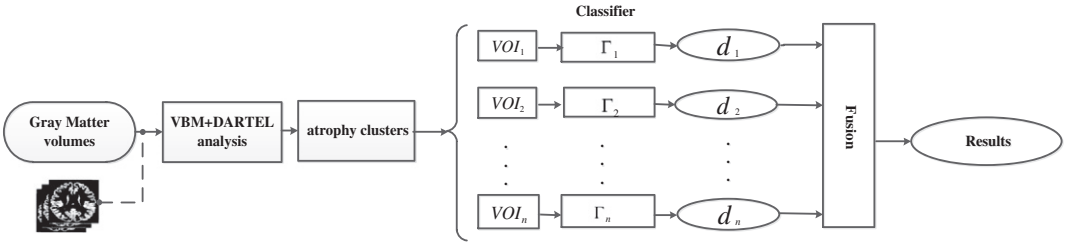

Fig. 4. Majority voting based score data fusion.

classifiers. With136 samples, a 10-fold cross validation implied having 122 PCs through the PCA process. The number of PCs, h, used to generate the projection vectors of the training and testing data was chosen as $h = 122$.

3.3.2. The general framework of feature ranking

The aim of feature ranking is to measure the relevance of features and class variables to aid in the selection of the most informative/discriminative features, thereby speeding up the learning process and promoting the performance of classifier models, especially when the dimensionality of the datasets is very large [64]. Let $D = [X_1, X_2, \ldots, X_N]^T$ be a dataset containing N samples, where $X_i = (x_{i1}, x_{i2}, \ldots, x_{iM})$ is a vector of M values and each value x_{ij} of this vector shows a feature of that sample. The vector $f_j = (x_{1j}, x_{2j}, \ldots, x_{Nj})^T$ is a vector of values of a feature f_j. On the other hand, D represents a $N \times M$ matrix, where row i is the subject X_i and each column j is the feature f_j. A feature-ranking algorithm applied to dataset D generates an ordered list of the features $\Psi = [f_*^1, f_*^2, \ldots, f_*^j]$, where the superscript denotes the position in the ranked list of a feature f_* and this list is ordered by reduction importance. Based on feature ranking, we can select the top k ranked features $[f_*^1, f_*^2, \ldots, f_*^k], k \geq M$ where k can be determined by the user or adjusted experimentally [65]. In this paper, we use t-test feature-ranking approach, as follows [66]:

$$T = \frac{\mu_{c1} - \mu_{c2}}{\sqrt{\frac{\sigma_{c1}^2}{n_{c1}} + \frac{\sigma_{c2}^2}{n_{c2}}}} \tag{1}$$

where T is the t-test value and μ_{c1}, σ_{c1}^2, n_{c1} and μ_{c2}, σ_{c2}^2, n_{c2} are the mean, variance values, and number of samples of two classes c_1 and c_2, respectively. The top informative features are selected by ranking all features according to their T values.

3.3.3. Optimal number of features based on Fisher Criterion

In addition to the feature-ranking algorithm based on the discriminative performance of the features, we propose the use of an automatic approach based on the Fisher Criterion, $J(w)$, given in Eq. (2), to determine the number of top discriminative features, thereby reducing the dimensionality of the prospective feature vectors [67,68].

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \tag{2}$$

Where S_B and S_W represent the determinants of between class and within class scatter matrices, respectively. For two classes c_1 and c_2, the between class scatter and within class scatter matrixes are defined as:

$$S_B = (\mu_{c1} - \mu_{c2})(\mu_{c1} - \mu_{c2})^T \tag{3}$$

$$S_W = \sum_{x_i \in c_1} (x_i - \mu_{c1})(x_i - \mu_{c1})^T + \sum_{x_i \in c_2} (x_i - \mu_{c2})(x_i - \mu_{c2})^T \tag{4}$$

Where $w = S_W^{-1}(\mu_{c1} - \mu_{c2})$ and μ_{ci} is the mean of data in each class. This approach helps in adaptively determining the k top discriminative features based on ranked t-test values using training data in each fold instead of using a fixed k. Once the features are ranked, the number of top ranked features iteratively increases from 1 to M (number of features) by calculating the respective Fisher Criterion. The number of top ranked features maximizing the Fisher Criterion is selected to be the optimal number of top ranked features k. The framework of the proposed feature selection method is illustrated in Fig. 2.

3.4. The SVM classifier

We classify AD patients apart from HCs by establishing the classification model using the SVM algorithm. The SVM is powerful classifier based on statistical learning principles. The SVM algorithm has been used successfully in a number of recent application machine learning studies [13,41,69–72]. During the training, SVM seeks the optimal class-separating hyper-plane in the maximal margin. Various kernels can be used during SVM training, such as linear, quadratic, polynomial, and radial basis function (RBF). In this work, SVM is performed using LIBSVM (http://www.csie.ntu.edu.tw/~cjlin/libsvm/)

Table 2
Clusters of gray matter atrophy (68 AD vs. 68 HC).

Location of peak voxels	Hemisphere	Cluster size (no. of voxels)	Talairach coordinates (x,y,z)	MNI coordinates (x,y,z)	Z value (peak voxel)	T value (peak voxel)
Hippocampus–Amygdala	R	16069	26, −11, −9	27, −9, −15	Inf	10.94
Hippocampus–lateral globuspallidus	L	16974	−25, −15, −8	−26, −13, −14	Inf	10.36
Inferior Parietal Lobule	R	1454	55, −44, 25	56, −46, 25	7.22	8
Anterior Cingulate	R	2032	8, 42, 2	9, 47, 3	6.54	6.54

Note: Anatomical regions are derived from the Talairach Client program; L, heft hemisphere; R, right hemisphere; MNI, Montreal Neurological Institute; (FWE-corrected at $P < 0.01$).

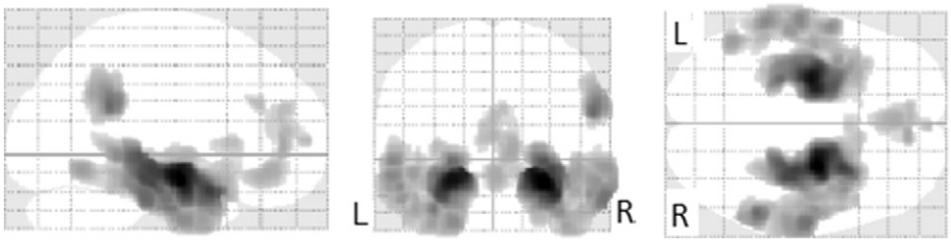

Fig. 5. Brain regions where there are significant gray matter reduction (atrophy) in 68 patients with AD and 68 age matched HC subjects (FWE corrected at P ˂ 0.01 and extend threshold K = 1400).

and the linear and nonlinear (RBF) kernels. A reliable measurement is achieved by obtaining all performance results using the 10-fold cross validation illustrated in Fig. 3. The RBF model has two parameters that need to be selected: C (regularization) and γ (controls the kernel width); the performance of the classifier depends on these parameters. The C and γ parameters are tuned using the training set, where two cross validation (CV) procedures with grid search are combined. This approach is performed to avoid unwarp bias in the estimation of accuracies produced by the CV procedure [73]. This procedure includes two nested loops. In the outer loop, the data are split into K_1 folds ($K_1 = 10$) at each step: one fold is used as a test and remaining $K_1 - 1$ folds for training and validation. In the inner loop, training data ($K_1 - 1$ folds) are further divided into K_2 folds ($K_2 = 10$). For each combination of *C* and γ, the classifier is trained using training data and its performance is assessed using the fold remaining for validation by estimating the classification accuracy. One fold is left for validation and the remaining $K_2 - 1$ folds are used for training, combined with grid search to determine the optimal parameters. In the grid search, the values of *C* and γ are varied among the candidate sets $\{2^{-5}, 2^{-4}, \ldots, 0, \ldots, 2^{19}, 2^{20}\}$ and $\{2^{-15}, 2^{-14}, \ldots, 0, \ldots, 2^{14}, 2^{15}\}$, respectively. The inner loop is repeated K_2 times, measuring the accuracy of the classifier across the K_2 folds for every combination of *C* and γ. The optimal parameters that produce maximum average accuracy across the K_2folds are selected, and then the class label of the test data is predicted, which is left out in the outer loop using the selected optimal parameters. The above procedure is repeated K_1 times by leaving a different fold as test data which are used to compute the classification accuracy. For SVM with a linear kernel, only the C parameter is tuned. Over-fitting is prevented by splitting the data into 10 parts, where the training set gets 9 parts and the test set gets 1 part. The data in the training set are used for parameter estimation, whereas the data in the test set are used to measure the performance. This process is repeated 10 times in the context of 10-fold cross validation, where no overlap of the testing sets occurs in this process [74].

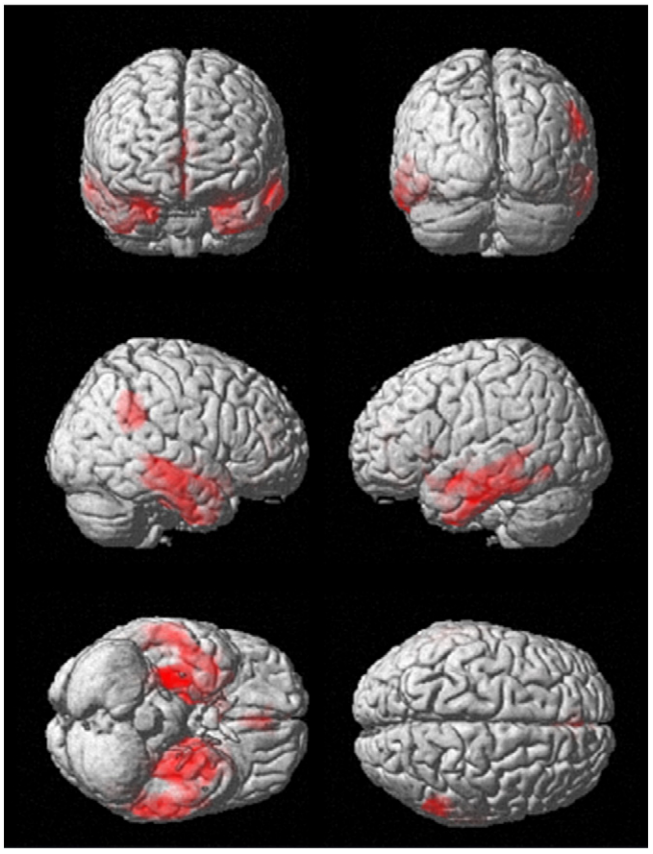

Fig. 6. Three-dimensional reconstruction of the brain showing gray matter atrophy using VBM technique plus DARTEL. The regions of gray matter loss are shown from anterior, posterior, right lateral, left lateral, inferior and superior view, respectively. The red region represents the region of gray matter loss.

The classification results are calculated by means of accuracy (ACC), sensitivity (SEN), specificity (SPE) and area under the curve (AUC), based on 10-fold cross validation. These parameters are defined as follows:

$$ACC = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{5}$$

$$SEN = \frac{TP}{TP + FN} \tag{6}$$

$$SPE = \frac{TN}{TN + FP} \tag{7}$$

where *TP* (the number of AD correctly identified as AD), *TN* (the number of HC patients correctly identified as HC), *FN* (the number of AD patients incorrectly identified as HC), and *FP* (the number of HC patients

Table 3
Raw feature vectors performance of atrophy clusters using 10 fold cross validation.

	Linear SVM				RBF SVM			
	ACC (%)	SEN (%)	SPE (%)	AUC (%)	ACC (%)	SEN (%)	SPE (%)	AUC (%)
VOI_1	80.14	79.41	80.88	85.37	82.35	80.88	83.82	88.71
VOI_2	77.20	77.94	76.47	84.93	79.41	76.47	82.35	87.69
VOI_3	71.32	70.58	72.05	75.65	75.00	72.05	77.94	80.75
VOI_4	69.85	69.11	70.58	77.82	70.58	73.52	67.64	77.99
VOI_{all}	77.20	79.41	75.00	84.49	83.82	83.82	83.82	86.00
Average	75.14	75.29	74.99	81.65	78.23	77.34	79.11	84.22

Note: ACC, Accuracy; SEN, Sensitivity; SPE, Specificity; AUC, Area Under Curve; SVM, Support Vector Machine; RBF, Radial Basis Function.

Table 4
PCA performance of atrophy clusters using 10 fold cross validation with 122 PCs.

	Linear SVM				RBF SVM			
	ACC (%)	SEN (%)	SPE (%)	AUC (%)	ACC (%)	SEN (%)	SPE (%)	AUC (%)
VOI_1	79.41	82.35	76.47	86.80	81.61	86.76	76.47	88.27
VOI_2	74.26	76.47	72.05	83.06	82.35	82.35	82.35	87.59
VOI_3	70.58	73.52	67.64	72.48	69.85	69.11	70.58	78.33
VOI_4	69.58	69.11	70.58	80.54	71.32	66.17	76.47	79.35
VOI_{all}	77.20	79.41	75.00	87.49	85.29	86.47	83.82	88.74
Average	74.20	76.17	72.34	82.07	78.08	78.17	77.93	84.45

Note: ACC, Accuracy; SEN, Sensitivity; SPE, Specificity; AUC, Area Under Curve; SVM, Support Vector Machine; RBF, Radial Basis Function.

incorrectly identified as AD) denote the number of true positive, true negative, false negative, and false positive cases, respectively.

3.5. Data fusion among atrophy clusters

This paper introduces data fusion technique among atrophy clusters (VOIs) to improve the performance of the proposed AD classification method. The aim of the data fusion technique is to integrate the data from two or more distinct multiple sources (vectors, classifiers) to improve performance. In the current work, two different fusion techniques are used: source fusion and score fusion.

3.5.1. Source data fusion

In the scheme of source data fusion, the top features selected based on our approach, described in Section 3.3, from different VOIs, are concatenated into a single feature vector. Assuming fv_1, fv_2, …, fv_n are feature vectors generated using proposed feature selection method for each atrophy cluster, the feature vector fusion (FVF) is then:

$$fvf = [fv_1, fv_2, \ldots, fv_n]_{1\times\sum_{i=1}^{n} m_i} \quad (8)$$

where m_i is the vector length for fv_i. This concatenated feature vector is then used for classification. The source data fusion relies on procedures for feature contraction.

3.5.2. Score data fusion

Score data fusion includes multiple classifiers and a combination method. The number of classifiers is determined based on the number of atrophy clusters obtained using the VBM plus DARTEL approach in the pre-processing. In this work, the majority voting method is employed as the score data fusion technique. Majority voting is one of the most versatile combination methods, because of its simplicity and performance on real data [75]. The adopted score data fusion framework is illustrated in Fig. 4.

4. Experimental results and discussion

This section considers the experimental results obtained through the pre-processing phase using VBM plus DARTEL analysis on 3D T1-weighted MR imaging, as an indicator disclosing the significance of decreased gray matter volumes in AD contributing to VOIs. The performance of the proposed feature selection method based on t-test ranking and the Fisher Criterion is also measured. Finally, the performance results obtained through data fusion are presented and analyzed. The performance of the classification using SVM classifiers with 10-fold cross validation is reported for the following cases: 1) performance of raw feature vectors directly extracted from VBM, 2) performance of the PCA data reduction method, 3) performance of proposed t-test feature-ranking technique using the optimal number of top features based on the Fisher Criterion, 4) performance of the proposed data fusion techniques among atrophy clusters of GM. The ACC (%), SEN (%), SPE (%) and AUC (%) performance metrics are used for the performance assessment.

4.1. Differences in gray matter volume between ADs and HCs

The gray matter volume atrophy differences between patients who suffer from AD and HC are summarized in Table 2. The group comparison by VBM plus DARTEL reveals a significant decline in GM volume in the right hippocampus (Talairach coordinates 26,−11,−9,x,y,z;z = Inf), left hippocampus (−25,−15,−8,x,y,z;z = Inf), right inferior parietal lobule (55,−44,25,x,y,z;z = 7.22), and right anterior cingulate (8,42,2,x,y,z;z = 6.54) (see Table 2 and Fig. 5 for more details) in patients with AD when compared to the HCs. Fig. 6 illustrates six three-dimensional views of group comparison representing relative gray matter atrophy in patients with AD compared to HCs. The voxel locations of the significant atrophy regions are used as 3D VOI masks. These 3D VOI masks are applied to the gray matter density volume results from the segmentation step in the VBM plus DARTEL analysis in order to extract voxel values into raw feature vectors for use in feature selection and classification. Based on these atrophy clusters, we define five different VOIs as follows:

i. VOI_1 includes the right hippocampus and amygdala regions. The center of this mask is at Talairach coordinates x = 26, y = −11, z = −9. VOI_1 contains 16,069 voxel values as a raw feature vector.

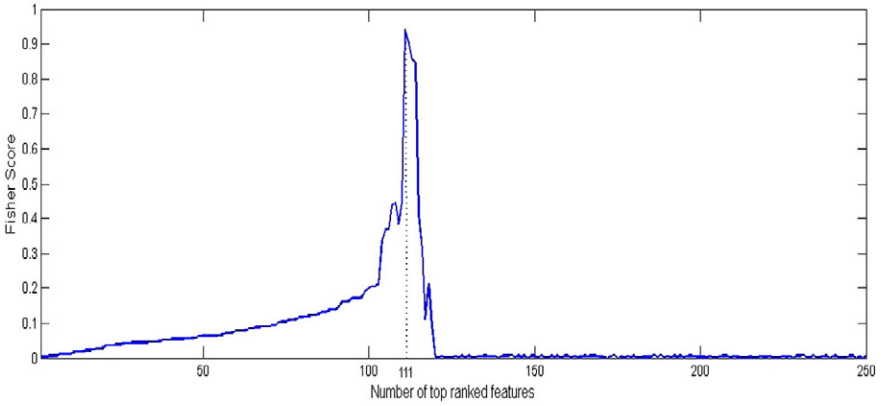

Fig. 7. Fisher scores for the respective ranked features in fold 1 training of VOI_{all}.

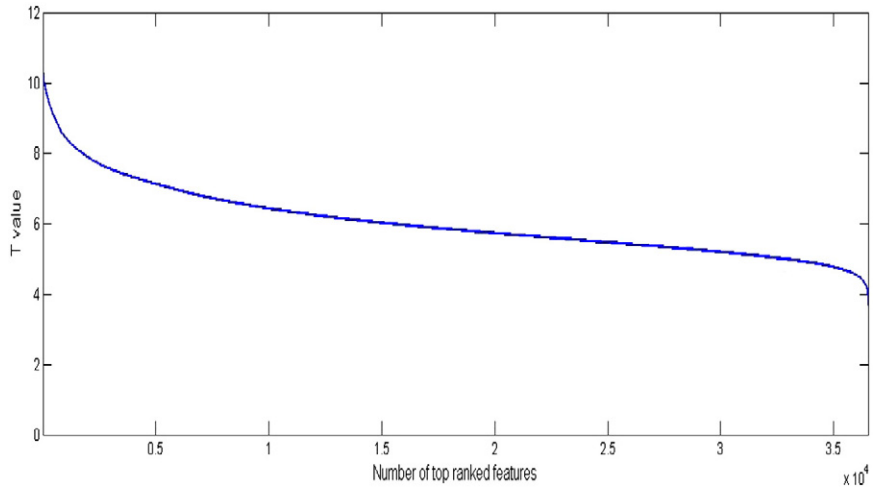

Fig. 8. t-test (T) values for the respective ranked features in fold 1 training of VOI_{all}.

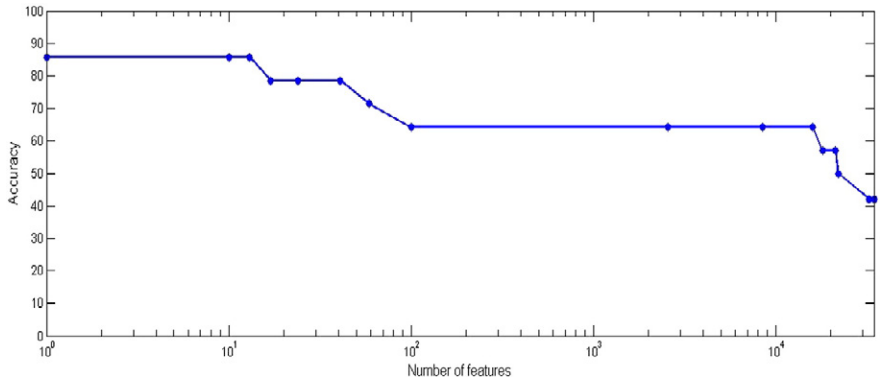

Fig. 9. Classification accuracies of linear SVM with respect to different numbers of features selected in fold 1 training of VOI_{all}.

ii. VOI_2 includes the left hippocampus–lateral globus pallidus regions. The center of this mask is at Talairach coordinates $x = -25$, $y = -15$, $z = -8$. VOI_2 contains 16,974 voxel values as a raw feature vector.
iii. VOI_3 includes the right inferior parietal lobule regions. The center of this mask is at Talairach coordinates $x = 55$, $y = -44$, $z = 25$. VOI_3 contains 1454 voxel values as a raw feature vector.
iv. VOI_4 includes the right anterior cingulate regions. The center of this mask is at Talairach coordinates $x = 8$, $y = 42$, $z = 2$. VOI_4 contains 2032 voxel values as a raw feature vector.
v. VOI_{all} includes all regions of gray matter loss (atrophy). VOI_{all} contains all four clusters above, with 36529 voxel values as a raw feature vector.

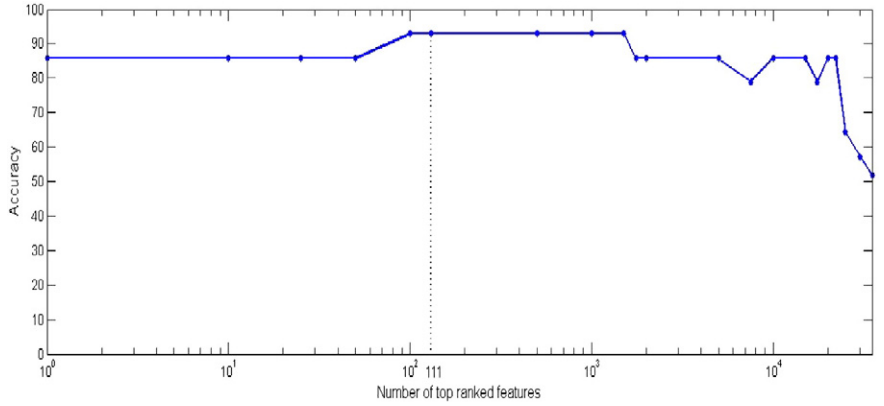

Fig. 10. Classification accuracies of linear SVM with respect to different numbers of top ranked features selected in fold 1 training of VOI_{all}.

Table 5
Performance results of the proposed feature selection method.

	Linear SVM				RBF SVM			
	ACC (%)	SEN (%)	SPE (%)	AUC (%)	ACC (%)	SEN (%)	SPE (%)	AUC (%)
VOI_1	91.17	92.64	89.70	96.9	90.44	89.70	91.17	95.07
VOI_2	92.64	91.17	94.11	97.93	94.11	92.64	95.58	98.74
VOI_3	76.47	73.52	79.41	84.93	76.47	75.00	77.94	84.66
VOI_4	79.41	75.00	83.82	86.67	80.14	73.52	86.76	89.29
VOI_{all}	94.11	95.58	92.64	98.33	92.64	94.11	91.17	98.13
Average	86.76	85.58	87.93	92.95	86.76	84.99	88.52	93.17

Note: ACC, Accuracy; SEN, Sensitivity; SPE, Specificity; AUC, Area Under Curve; SVM, Support Vector Machine; RBF, Radial Basis Function.

Note that the center of the mask in the Talairach coordinates corresponds to the center of the mass of the respective 3D VOI.

4.2. Performance of the raw feature vectors

The complete MRI dataset consists of 68 AD and 68 HC samples. The ACC, SEN, SPE, and AUC that were obtained by 10-fold cross validation using the SVM classifier (Linear and RBF kernels) on raw feature vectors from five different VOIs are shown in Table 3. The results indicate that the average performance, in terms of ACC, SEN, SPE, and AUC obtained from five atrophy clusters using RBF SVM, is marginally better than Linear SVM. The RBF kernel is generally more flexible than the linear kernel so it generally can model more functions with its function space.

4.3. Performance of the PCA method

The PCA based data reduction method is utilized to extract raw feature vectors. For each dataset, the features extracted are reduced to lower dimensional features using PCA, with 122 PCs. Table 4 presents the classifier performance obtained using 10-fold cross validation for SVM classifiers in terms of ACC, SEN, SPE, and AUC. The average accuracy of feature vectors with 122 PCs for linear and RBF SVM classifiers was74.20% and 78.08%, respectively, while the average accuracy using the raw feature vectors without dimensionality reduction was 75.14% and 78.23%, respectively. As observed, PCA introduces dimensionality reduction and generates comparable performance with the raw data.

4.4. Performance of the proposed feature selection using t-test ranking and the Fisher Criterion

As proposed in Section 3.3.3, the feature selection technique uses the t-test for ranking the features. The Fisher Criterion is used to determine the optimal number of top features. The Fisher scores for the samples in the training set from fold 1 of VOI_{all} are plotted in Fig. 7 for the top 250 ranked features. As Fig. 7 shows, the maximum Fisher score is located at 111, which means that 111 top-ranked features are to be used in the classification process. Typical Fisher scores are observed between 30 and 150 for all folds of 5 different VOIs. Fig. 8 shows all of the t-test values for the same data. The contribution of features on the accuracy is studied separately and plotted in Fig. 9 with linear SVM. As expected, the contribution of the features in relevance to their t-test values is highly correlated. A higher t-test rank implies higher performance of the respective feature. A logarithmic scale is used to cover the entire feature space. Additionally Fig. 10 is included to show the improvement in the accuracy obtained by using progressive inclusion of the ranked features in the feature vector with linear SVM. The performance increases with the increased number of ranked features used in the classification. However, after a certain maximum, which corresponds to 111 top ranked features in this fold, the performance does not increase further. The SVM-based classifiers are used to observe the classification performance of the selected feature vectors from five different VOIs. The results of classifiers are presented in Table 5. Examination of Tables 3 and 5 confirms that the proposed feature section method significantly improves the prediction capability of AD subjects when compared to prediction using raw features. The average accuracy for raw data for linear and RBF SVM classifiers is 75.14% and 78.23%, respectively, while the average accuracy for the proposed feature selection method is 86.76% and 86.76%, respectively. The improvement is around 10% for all performance indicators: ACC, SEN, SPE, and AUC.

4.5. Performance of data fusion among atrophy clusters

The performance improvement aided by data fusion of five clusters is shown in Table 6. The performance of both types of data fusion techniques is around 10% higher than the average performance obtained with individual clusters. The performance of the majority voting (score fusion) approach is always higher than or equal to the performance of the source concatenation (source fusion) approach. Table 6 shows that data fusion among atrophy clusters of GM volumes integrates information by improving the classification performance in all terms.

4.6. Performance comparison to the other methods

Several recent studies have reported classification results to distinguish AD and HC based on MRI. Zhang et al. [8] used multimodal classification of AD based on the combination of MRI, CSF, and PET. They reported an ACC of 86.2% in the classification of AD/HC by MRI image modality. They also achieved a high ACC performance of 93.2% by combining the MRI, CSF, and PET results. Westman et al. [12] reported an ACC of 87% from MRI data and increased it to 91.8% by combining MRI data with CSF measures. Zhou et al. [40] employed FreeSurfer software to calculate 55 volumetric variables from MRI. They reported an ACC of 78% for MRI data and 92.4% for combining MRI data with MMSE. In the present work, only the MRI modality with 136 samples from the ADNI dataset was used, with highly comparable results to those reported in other MRI-only studies. The performance of the proposed feature selection and data fusion techniques outperforms the alternative techniques given in Table 7. The detailed parameters of classification performance with different methods on MRI data are also provided

Table 6
Performance of proposed data fusion technique among atrophy clusters of GM.

	Linear SVM				RBF SVM			
	ACC (%)	SEN (%)	SPE (%)	AUC (%)	ACC (%)	SEN (%)	SPE (%)	AUC (%)
Source Concatenation	95.58	94.11	97.05	97.52	95.58	94.11	97.05	97.31
Majority Voting	96.32	94.11	98.52	99.93	95.59	94.11	97.05	99.82

Note: ACC, Accuracy; SEN, Sensitivity; SPE, Specificity; AUC, Area Under Curve; SVM, Support Vector Machine; RBF, Radial Basis Function.

Table 7
Supervised classification results of Alzheimer's disease and healthy control subjects on MRI data.

Author	Imaging Modality	Source of data	AD/HC	Validation method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
Zhang et al., 2011 [8]	MRI	ADNI	51/52	10 Fold	86.2	86.0	86.3	–
Zhang et al., 2011 [8]	MRI + CSF + PET	ADNI	51/52	10 Fold	93.2	93.0	93.3	–
Westman et al., 2012 [12]	MRI	ADNI	96/111	10 Fold	87	83.3	90.1	93.0
Westman et al., 2012 [12]	MRI + CSF	ADNI	96/111	10 Fold	91.8	88.5	94.6	95.8
Zhou et al., 2014 [40]	MRI	Private	127/59	2 Fold	78.2	68.5	81.7	–
Zhou et al., 2014 [40]	MRI + MMSE	Private	127/59	2 Fold	92.4	84.0	96.1	–
Kloppel et al., 2008 [32]	MRI (Group I)	Private	20/20	Leave-one-out	95.0	95.0	95.0	–
Kloppel et al., 2008 [32]	MRI (Group II)	Private	14/14	Leave-one-out	92.9	100	85.7	–
Kloppel et al., 2008 [32]	MRI (Group III)	Private	33/57	Leave-one-out	81.1	60.6	93.0	–
Hinrichs et al., 2011 [41]	MRI + PET	ADNI	48/66	10 Fold	87.6	78.9	93.8	–
Hinrichs et al., 2011 [41]	MRI + PET + CSF + APOE + Cognitive Scores	ADNI	48/66	10 Fold	92.4	86.7	96.6	–
Proposed method	**MRI**	**ADNI**	**68/68**	**10 Fold**	**96.32**	**94.11**	**98.52**	**99.93**

in Table 7.The results reported in Table 7 show that the performance of the proposed system is highly competitive for the performance terms including ACC, SPE, and AUC when compared to the other systems reported in the literature. The only exception is SPE, where the performance of the proposed system is lower than for results reported by Kloppel et al. (2008) [32] for groups I and II. Our results are highly competitive with the rest of the systems. The performance improvement over the previous work, shown in Table 7, can be attributed to the automatic statistical feature-selection method based on the combination of t-test feature ranking and the Fisher Criterion of the VOI. Due to t-test ranking, the proposed feature selection method is capable of sorting discriminative features in descending order. The optimal dimension of the feature vector is adjusted by maximizing the Fisher Criterion in the training dataset. Finally, data fusion techniques among gray matter atrophy clusters provide further improvement on the AD classification performance.

5. Conclusion

This paper proposes a feature selection method using t-test-based feature ranking, which is used for the classification of AD. The optimal size of the selected features is determined using the Fisher Criterion, which maximizes the class separation between AD and HC. The feature selection is applied to all voxels that pass through masks modeled by overall atrophy clusters, determined by using VBM analysis. Linear and RBF kernel-based SVM classifiers are used for the classification of the extracted feature vectors after the proposed feature selection method. A performance improvement is also proposed by applying data fusion among the individual atrophy clusters, as well as the overall atrophy clusters. Standard data fusion techniques, such as source and score fusion, are used to obtain improved performance in the classification of AD. The performance of the proposed system is measured on 136 subjects (68 AD and 68 HC) from the ADNI dataset using 10-fold cross validation. The experimental results show that the performance of the proposed approach for ACC, SPE, and AUC is highly competitive with the state-of-the-art techniques using MRI data reported in the literature.

References

[1] Alzheimer's Association. Alzheimer's disease and dementia. http://www.alz.org/; 2015. [accessed April 5, 2015].
[2] Liu S, Cai W, Wen L, Feng DD, Pujol S, Kikinis R, et al. Multi-channel neurodegenerative pattern analysis and its application in Alzheimer's disease characterization. Comput Med Imaging Graph 2014;38:436–44. http://dx.doi.org/10.1016/j.compmedimag.2014.05.003.
[3] Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. 2014 Update of the Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimers Dement 2015;11:e1-120. http://dx.doi.org/10.1016/j.jalz.2014.11.001.
[4] Hanyu H, Sato T, Hirao K, Kanetaka H, Iwamoto T, Koizumi K. The progression of cognitive deterioration and regional cerebral blood flow patterns in Alzheimer's disease: a longitudinal SPECT study. J Neurol Sci 2010;290:96–101. http://dx.doi.org/10.1016/j.jns.2009.10.022.
[5] Gray KR, Wolz R, Heckemann RA, Aljabar P, Hammers A, Rueckert D. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. Neuroimage 2012;60:221–9. http://dx.doi.org/10.1016/j.neuroimage.2011.12.071.
[6] Chen YJ, Deutsch G, Satya R, Liu HG, Mountz JM. A semi-quantitative method for correlating brain disease groups with normal controls using SPECT: Alzheimer's disease versus vascular dementia. Comput Med Imaging Graph 2013;37:40–7. http://dx.doi.org/10.1016/j.compmedimag.2012.11.001.
[7] Górriz JM, Segovia F, Ramírez J, Lassl A, Salas-Gonzalez D. GMM based SPECT image classification for the diagnosis of Alzheimer's disease. Appl Soft Comput 2011;11:2313–25. http://dx.doi.org/10.1016/j.asoc.2010.08.012.
[8] Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 2011;55:856–67. http://dx.doi.org/10.1016/j.neuroimage.2011.01.008.
[9] Papakostas GA, Savio A, Graña M, Kaburlasos VG. A lattice computing approach to Alzheimer's disease computer assisted diagnosis based on MRI data. Neurocomputing 2015;150:37–42. http://dx.doi.org/10.1016/j.neucom.2014.02.076.
[10] Beheshti I, Demirel H. Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease. Comput Biol Med 2015;64:208–16. http://dx.doi.org/10.1016/j.compbiomed.2015.07.006.
[11] Aguilar C, Westman E, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, et al. Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. Psychiatry Res Neuroimaging 2013;212:89–98. http://dx.doi.org/10.1016/j.pscychresns.2012.11.005.
[12] Westman E, Muehlboeck JS, Simmons A. Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. Neuroimage 2012;62:229–38. http://dx.doi.org/10.1016/j.neuroimage.2012.04.056.
[13] Li M, Qin Y, Gao F, Zhu W, He X. Discriminative analysis of multivariate features from structural MRI and diffusion tensor images. Magn Reson Imaging 2014;32:1043–51. http://dx.doi.org/10.1016/j.mri.2014.05.008.
[14] Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage 2015;104:398–412. http://dx.doi.org/10.1016/j.neuroimage.2014.10.002.
[15] Bron EE, Smits M, van der Flier WM, Vrenken H, Barkhof F, Scheltens P, et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. Neuroimage 2015;111:562–79. http://dx.doi.org/10.1016/j.neuroimage.2015.01.048.
[16] Andersen AH, Rayens WS, Liu Y, Smith CD. Partial least squares for discrimination in fMRI data. Magn Reson Imaging 2012;30:446–52. http://dx.doi.org/10.1016/j.mri.2011.11.001.
[17] Fan Y, Resnick SM, Wu X, Davatzikos C. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. Neuroimage 2008;41:277–85. http://dx.doi.org/10.1016/j.neuroimage.2008.02.043.
[18] Dinesh E, Kumar MS, Vigneshwar M, Mohanraj T. Instinctive classification of Alzheimer's disease using FMRI, pet and SPECT images. 7th Int Conf Intell Syst Control; 2013. p. 405–9. http://dx.doi.org/10.1109/ISCO.2013.6481189.
[19] Mesrob L. DTI and structural MRI classification in Alzheimer's disease. Adv Mol Imaging 2012;02:12–20. http://dx.doi.org/10.4236/ami.2012.22003.
[20] Graña M, Termenon M, Savio A, Gonzalez-Pinto A, Echeveste J, Pérez JM, et al. Computer aided diagnosis system for Alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation. Neurosci Lett 2011;502:225–9. http://dx.doi.org/10.1016/j.neulet.2011.07.049.
[21] Lee W, Park B, Han K. Classification of diffusion tensor images for the early detection of Alzheimer's disease. Comput Biol Med 2013;43:1313–20. http://dx.doi.org/10.1016/j.compbiomed.2013.07.004.
[22] Shao J, Myers N, Yang Q, Feng J, Plant C, Böhm C, et al. Prediction of Alzheimer's disease using individual structural connectivity networks. Neurobiol Aging 2012;33:2756–65. http://dx.doi.org/10.1016/j.neurobiolaging.2012.01.017.
[23] Wee CY, Yap PT, Zhang D, Denny K, Browndyke JN, Potter GG, et al. Identification of MCI individuals using structural and functional connectivity networks.

Neuroimage 2012;59:2045–56. http://dx.doi.org/10.1016/j.neuroimage.2011.10.015.
[24] Challis E, Hurley P, Serra L, Bozzali M, Oliver S, Cercignani M. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. Neuroimage 2015;112:232–43. http://dx.doi.org/10.1016/j.neuroimage.2015.02.037.
[25] Vemuri P, Jack CR. Role of structural MRI in Alzheimer's disease. Alzheimers Res Ther 2010:2–23. http://dx.doi.org/10.1186/alzrt47.
[26] Huang CHC, BY Bin Yan, Jiang HJH, Wang DWD. Combining voxel-based morphometry with artifical neural network theory in the application research of diagnosing Alzheimer's disease. Int Conf Biomed Eng Inform 2008;1:250–4. http://dx.doi.org/10.1109/BMEI.2008.245.
[27] Zhang JZJ, BY Bin Yan, Huang XHX, Yang PYP, Huang CHC. The diagnosis of Alzheimer's disease based on voxel-based morphometry and support vector machine. Fourth Int Conf Nat Comput, 2. ; 2008. p. 197–201. http://dx.doi.org/10.1109/ICNC.2008.804.
[28] Savio A, García-Sebastián MT, Chyzyk D, Hernandez C, Graña M, Sistiaga A, et al. Neurocognitive disorder detection based on feature vectors extracted from VBM analysis of structural MRI. Comput Biol Med 2011;41:600–10. http://dx.doi.org/10.1016/j.compbiomed.2011.05.010.
[29] Yanxi Liu CCM. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification. Med Image Comput Comput Interv – MICCAI; 2004. p. 393–401.
[30] Lao Z, Shen D, Xue Z, Karacali B, Resnick SM, Davatzikos C. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. Neuroimage 2004;21:46–57. http://dx.doi.org/10.1016/j.neuroimage.2003.09.027.
[31] Fung G, Stoeckel J. SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. Knowl Inf Syst 2007;11:243–58. http://dx.doi.org/10.1007/s10115-006-0043-5.
[32] Klöppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, et al. Automatic classification of MR scans in Alzheimer's disease. Brain 2008;131:681–9. http://dx.doi.org/10.1093/brain/awm319.
[33] Chincarini A, Bosco P, Calvini P, Gemme G, Esposito M, Olivieri C, et al. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage 2011;58:469–80. http://dx.doi.org/10.1016/j.neuroimage.2011.05.083.
[34] Westman E, Cavallin L, Muehlboeck JS, Zhang Y, Mecocci P, Vellas B, et al. Sensitivity and specificity of medial temporal lobe visual ratings and multivariate regional MRI classification in Alzheimer's disease. PLoS One 2011;6. http://dx.doi.org/10.1371/journal.pone.0022506.
[35] Ben Ahmed O, Benois-Pineau J, Allard M, Ben Amar C, Catheline G. Classification of Alzheimer's disease subjects from MRI using hippocampal visual features. Multimed Tools Appl 2014;74:1249–66. http://dx.doi.org/10.1007/s11042-014-2123-y.
[36] Li S, Shi F, Pu F, Li X, Jiang T, Xie S, et al. Hippocampal shape analysis of Alzheimer disease based on machine learning methods. Am J Neuroradiol 2007;28:1339–45. http://dx.doi.org/10.3174/ajnr.A0620.
[37] Coupé P, Eskildsen SF, Manjón JV, Fonov VS, Collins DL. Simultaneous segmentation and grading of anatomical structures for patient's classification: application to Alzheimer's disease. Neuroimage 2012;59:3736–47. http://dx.doi.org/10.1016/j.neuroimage.2011.10.080.
[38] Gerardin E, Chételat G, Chupin M, Cuingnet R, Desgranges B, Kim HS, et al. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 2009;47:1476–86. http://dx.doi.org/10.1016/j.neuroimage.2009.05.036.
[39] Chupin M, Gérardin E, Cuingnet R, Boutet C, Lemieux L, Lehéricy S, et al. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 2009;19:579–87. http://dx.doi.org/10.1002/hipo.20626.
[40] Zhou Q, Goryawala M, Cabrerizo M, Wang J, Barker W, Loewenstein DA, et al. An optimal decisional space for the classification of Alzheimer's disease and mild cognitive impairment. Biomed Eng IEEE Trans 2014;61:2245–53. http://dx.doi.org/10.1109/TBME.2014.2310709.
[41] Hinrichs C, Singh V, Xu G, Johnson SC. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. Neuroimage 2011;55:574–89. http://dx.doi.org/10.1016/j.neuroimage.2010.10.081.
[42] Ashburner J, Friston KJ. Voxel-based morphometry—the methods. Neuroimage 2000;11:805–21. http://dx.doi.org/10.1006/nimg.2000.0582.
[43] Nakatsuka T, Imabayashi E, Matsuda H, Sakakibara R, Inaoka T, Terada H. Discrimination of dementia with Lewy bodies from Alzheimer's disease using voxel-based morphometry of white matter by statistical parametric mapping 8 plus diffeomorphic anatomic registration through exponentiated Lie algebra. Neuroradiology 2013;55:559–66. http://dx.doi.org/10.1007/s00234-013-1138-9.
[44] Matsuda H, Mizumura S, Nemoto K, Yamashita F, Imabayashi E, Sato N, et al. Automatic voxel-based morphometry of structural MRI by SPM8 plus diffeomorphic anatomic registration through exponentiated lie algebra improves the diagnosis of probable Alzheimer disease. Am J Neuroradiol 2012;33:1109–14. http://dx.doi.org/10.3174/ajnr.A2935.
[45] Liu M, Zhang D, Shen D. Ensemble sparse classification of Alzheimer's disease. Neuroimage 2012;60:1106–16. http://dx.doi.org/10.1016/j.neuroimage.2012.01.055.
[46] Chaves R, Ramírez J, Górriz JM, López M, Salas-Gonzalez D, Álvarez I, et al. SVM-based computer-aided diagnosis of the Alzheimer's disease using t-test NMSE feature selection with feature correlation weighting. Neurosci Lett 2009;461:293–7. http://dx.doi.org/10.1016/j.neulet.2009.06.052.
[47] Wang D, Zhang H, Liu R, Lv W, Wang D. T-test feature selection approach based on term frequency for text categorization. Pattern Recognit Lett 2014;45:1–10. http://dx.doi.org/10.1016/j.patrec.2014.02.013.
[48] Cousijn J, Wiers RW, Ridderinkhof KR, Van den Brink W, Veltman DJ, Goudriaan AE. Grey matter alterations associated with cannabis use: results of a VBM study in heavy cannabis users and healthy controls. Neuroimage 2012;59:3845–51. http://dx.doi.org/10.1016/j.neuroimage.2011.09.046.
[49] Ashburner J. Computational anatomy with the SPM software. Magn Reson Imaging 2009;27:1163–74. http://dx.doi.org/10.1016/j.mri.2009.01.006.
[50] Takao H, Hayashi N, Ohtomo K. Brain morphology is individual-specific information. Magn Reson Imaging 2015. http://dx.doi.org/10.1016/j.mri.2015.03.010.
[51] Hirata Y, Matsuda H, Nemoto K, Ohnishi T, Hirao K, Yamashita F, et al. Voxel-based morphometry to discriminate early Alzheimer's disease from controls. Neurosci Lett 2005;382:269–74. http://dx.doi.org/10.1016/j.neulet.2005.03.038.
[52] Karas GB, Burton EJ, Rombouts SARB, Van Schijndel RA, O'Brien JT, Scheltens P, et al. A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. Neuroimage 2003;18:895–907. http://dx.doi.org/10.1016/S1053-8119(03)00041-7.
[53] Chételat G, Landeau B, Eustache F, Mézenge F, Viader F, De La Sayette V, et al. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. Neuroimage 2005;27:934–46. http://dx.doi.org/10.1016/j.neuroimage.2005.05.015.
[54] Modi S, Bhattacharya M, Singh N, Tripathi RP, Khushu S. Effect of visual experience on structural organization of the human brain: a voxel based morphometric study using DARTEL. Eur J Radiol 2012;81:2811–9. http://dx.doi.org/10.1016/j.ejrad.2011.10.022.
[55] Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage 2007;38:95–113. http://dx.doi.org/10.1016/j.neuroimage.2007.07.007.
[56] Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 2009;46:786–802. http://dx.doi.org/10.1016/j.neuroimage.2008.12.037.
[57] Kasahara K, Hashimoto K, Abo M, Senoo A. Voxel- and atlas-based analysis of diffusion tensor imaging may reveal focal axonal injuries in mild traumatic brain injury — comparison with diffuse axonal injury. Magn Reson Imaging 2012;30:496–505. http://dx.doi.org/10.1016/j.mri.2011.12.018.
[58] Rajapakse JC, Giedd JN, Rapoport JL. Statistical approach to segmentation of single-channel cerebral MR images. IEEE Trans Med Imaging 1997;16:176–86. http://dx.doi.org/10.1109/42.563663.
[59] Tohka J, Zijdenbos A, Evans A. Fast and robust parameter estimation for statistical partial volume models in brain MRI. Neuroimage 2004;23:84–97. http://dx.doi.org/10.1016/j.neuroimage.2004.05.007.
[60] Manjón JV, Coupé P, Martí-Bonmatí L, Collins DL, Robles M. Adaptive non-local means denoising of MR images with spatially varying noise levels. J Magn Reson Imaging 2010;31:192–203. http://dx.doi.org/10.1002/jmri.22003.
[61] Cuadra MB, Cammoun L, Butz T, Cuisenaire O, Thiran JP. Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images. IEEE Trans Med Imaging 2005;24:1548–65. http://dx.doi.org/10.1109/TMI.2005.857652.
[62] Haq NF, Kozlowski P, Jones EC, Chang SD, Goldenberg SL, Moradi M. A data-driven approach to prostate cancer detection from dynamic contrast enhanced MRI. Comput Med Imaging Graph 2015;41:37–45. http://dx.doi.org/10.1016/j.compmedimag.2014.06.017.
[63] Tang L, Wen Y, Zhou Z, von Deneen KM, Huang D, Ma L. Reduced field-of-view DTI segmentation of cervical spine tissue. Magn Reson Imaging 2013;31:1507–14. http://dx.doi.org/10.1016/j.mri.2013.07.003.
[64] Zhou N, Wang L. A modified t-test feature selection method and its application on the HapMap genotype data. Genomics Proteomics Bioinformatics 2007;5:242–9. http://dx.doi.org/10.1016/S1672-0229(08)60011-X.
[65] Prati RC. Combining feature ranking algorithms through rank aggregation. Proc Int Jt Conf Neural Networks 2012;10–5. http://dx.doi.org/10.1109/IJCNN.2012.6252467.
[66] Kamkar I, Gupta SK, Phung D, Venkatesh S. Stable feature selection for clinical prediction: exploiting ICD tree structure using Tree-Lasso. J Biomed Inform 2014;53:277–90. http://dx.doi.org/10.1016/j.jbi.2014.11.013.
[67] Gao Q, Liu J, Zhang H, Hou J, Yang X. Enhanced Fisher discriminant criterion for image recognition. Pattern Recognit 2012;45:3717–24. http://dx.doi.org/10.1016/j.patcog.2012.03.024.
[68] Diaf A, Boufama B, Benlamri R. Non-parametric Fisher's discriminant analysis with kernels for data classification. Pattern Recognit Lett 2013;34:552–8. http://dx.doi.org/10.1016/j.patrec.2012.10.030.
[69] Dimitrovski I, Kocev D, Kitanovski I, Loskovska S, Džeroski S. Improved medical image modality classification using a combination of visual and textual features. Comput Med Imaging Graph 2015;39:14–26. http://dx.doi.org/10.1016/j.compmedimag.2014.06.005.
[70] Al-Kadi OS. A multiresolution clinical decision support system based on fractal model design for classification of histological brain tumours. Comput Med Imaging Graph 2014;41:67–79. http://dx.doi.org/10.1016/j.compmedimag.2014.05.013.
[71] Xue T, Bai L, Chen S, Zhong C, Feng Y, Wang H, et al. Neural specificity of acupuncture stimulation from support vector machine classification analysis. Magn Reson Imaging 2011;29:943–50. http://dx.doi.org/10.1016/j.mri.2011.03.003.

[72] Song X, Chen NK. A SVM-based quantitative fMRI method for resting-state functional network detection. Magn Reson Imaging 2014;32:819–31. http://dx.doi.org/10.1016/j.mri.2014.04.004.
[73] Casanova R, Maldjian JA, Espeland MA. Evaluating the impact of different factors on voxel-based classification methods of ADNI structural MRI brain images. Int J Biomed Data Min 2011;1:1–10. http://dx.doi.org/10.4303/ijbdm/B110102.
[74] Van Der Heijden VF, De Ridder D. Classification, parameter estimation, and state estimation: an engineering approach using MATLAB, vol. 32. John Wiley and Sons; 2004.
[75] Narasimhamurthy A. Theoretical bounds of majority voting performance for a binary classification problem. IEEE Trans Pattern Anal Mach Intell 2005;27:1988–95. http://dx.doi.org/10.1109/TPAMI.2005.249.