



Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images



L. Khedher, J. Ramírez, J.M. Górriz*, A. Brahim, F. Segovia, the Alzheimer's Disease Neuroimaging Initiative¹

Department of Signal Theory, Networking and Communications, University of Granada, Spain

ARTICLE INFO

Article history:

Received 3 February 2014

Received in revised form

21 July 2014

Accepted 18 September 2014

Available online 27 October 2014

Keywords:

Computer aided diagnosis system

Alzheimer disease

PLS

PCA

Support vector machine

ADNI database

ABSTRACT

Computer aided diagnosis (CAD) systems using functional and structural imaging techniques enable physicians to detect early stages of the Alzheimer's disease (AD). For this purpose, magnetic resonance imaging (MRI) have been proved to be very useful in the assessment of pathological tissues in AD. This paper presents a new CAD system that allows the early AD diagnosis using tissue-segmented brain images. The proposed methodology aims to discriminate between AD, mild cognitive impairment (MCI) and elderly normal control (NC) subjects and is based on several multivariate approaches, such as partial least squares (PLS) and principal component analysis (PCA). In this study, 188 AD patients, 401 MCI patients and 229 control subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database were studied. Automated brain tissue segmentation was performed for each image obtaining gray matter (GM) and white matter (WM) tissue distributions. The validity of the analyzed methods was tested on the ADNI database by implementing support vector machine classifiers with linear or radial basis function (RBF) kernels to distinguish between normal subjects and AD patients. The performance of our methodology is validated using *k*-fold cross technique where the system based on PLS feature extraction and linear SVM classifier outperformed the PCA method. In addition, PLS feature extraction is found to be more effective for extracting discriminative information from the data. In this regard, the developed latter CAD system yielded maximum sensitivity, specificity and accuracy values of 85.11%, 91.27% and 88.49%, respectively.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Alzheimer's Disease (AD) is one of the most common forms of neurodegenerative brain disorders. This form of dementia is characterized by neurofibrillary tangles, amyloid plaques and histopathologic changes that are usually associated with neuronal loss and brain volume reductions [1]. This disease begins slowly, starting with memory loss and other cognitive functions getting worse until patients lose their ability to recognize very familiar things or persons. Furthermore, the occurrence and dominance of this disease will

increase, in the coming years [2], due to the growth of the older population in developed nations. In addition, it is considered to be one of the major causes of death around the globe. Deaths from heart disease have decreased by 16%, breast cancer by 2%, prostate cancer by 8% and stroke by 23% whereas deaths by AD increased by 68% since year 2000 [3]. Up to now, the pathogenesis and pathologies of AD is unknown. In addition, there is still no special treatment for AD, hence early diagnosis becomes an important way to improve AD survival rate. To clinically diagnose AD at an early stage [4,5], many biomedical imaging techniques have been used, such as Magnetic Resonance Imaging (MRI) [6,7]. The MRI [6,7] scan is a type of test that has been widely used for early detection and diagnosis of AD [4,5,8–11]. This imaging technique produces high quality images of the anatomical structures for the human body, specifically in the brain, and provides rich information for clinical diagnosis and biomedical research [12,13]. The early changes are reliable with the underlying pathology of AD. However, it is unclear which structures are most useful for early diagnosis because the pattern of AD pathology is complex and evolves as the disease progresses [4].

* Corresponding author.

E-mail address: gorriz@ugr.es (J.M. Górriz).

¹ Data used in the preparation of this paper were obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators include (complete listing available at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

Many studies have used manual delineation (segmentation) of the hippocampus in MR images [14–19]. These studies have demonstrated a high accuracy in distinguishing between AD patients and normal controls. Automatic methods of measuring hippocampal volumes have also been applied with similar results [20,21]. Entorhinal cortex measures have additionally been used to discriminate between subjects with AD and normal controls [15,16]. Hippocampal volumes and entorhinal cortex measures have been found to be equally accurate in distinguishing between AD and normal cognitive elderly subjects [22]. It is probable that a single structure such as the hippocampus is not sufficient for accurate diagnosis of the disease and the combination of different structures has proven to be more useful when distinguishing AD patients from cognitively normal elderly subject [23]. Therefore, multivariate approaches give the opportunity to analyze many variables at the same time and observe the intrinsic patterns of the complex data from different regions of the brain. Previous studies have utilized different techniques of multivariate approaches such as Principal Component Analysis (PCA), Partial Least Square (PLS) and support vector machine (SVM) to analyze MRI data. In this work, we have used the PLS and PCA methods for analyzing our database. The aim of our study was to compare different feature extraction methods for classification of patients with AD, MCI and NC based on MRI segmented data (GM, WM and (GM+WM)).

Furthermore, in the last years many researches have developed Computer Aided Diagnosis (CAD) systems that combine medical images of the brain and methods from Statistical Machine Learning [24–27] in order to automatically diagnose AD. In this way, several approaches for a CAD system based on functional and structural images have been proposed. A large number of these CAD systems for AD are composed of three steps: (i) preprocessing procedure, (ii) feature extraction algorithm, and (iii) classification. The first step ensures that different images from different subjects, with brains of different size and shape, are comparable. The feature extraction algorithm transforms the input data into small vectors in order to avoid the small sample size problem [28]. These vectors must include all the relevant information contained in the input data. Once it has been trained, the classifier determines if they are more similar to healthy subject vectors, to MCI patient vectors or to AD patient vectors, thus performing the diagnosis. In our case, this paper describes a CAD system for AD based on PLS and PCA feature extraction methods to detect early stages of the Alzheimer's disease (AD).

This paper is organized as follows. In Section 2 the database used to test the system, preprocessing technique, voxel selection, feature extraction methods, description of proposed methodology and classification methods used in this paper are presented. In Section 3 we propose some experiments, and different results obtained with each combination are detailed. Discussion is given in Section 4 in order to analyze the experimental results obtained in this work, and the conclusions are presented in the last section.

2. Materials and methods

2.1. ADNI database

Data used in preparation of this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a \$60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological

assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55–90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

The ADNI database contains 1.5 T and 3.0 T T1w MRI scans for AD, MCI, and cognitively normal controls (NC) at several time points. In addition, all structural MR scans used in this paper were acquired at 1.5 T MRI scanner and this database provides data for three groups of patients: healthy patients (Normal Controls, NC), Alzheimer disease patients (AD) and patients with mild cognitive symptoms (MCI). The database used in this work contains 1075 T1-weighted MRI images, comprising 229 NC, 401 MCI (312 stable MCI and 86 progressive MCI) and 188 AD. As only the first exam for each patient has been used in this work, 818 images were used for assessing the proposed approach. Demographic data of patients in the database is summarized in Table 1.

All subjects must be willing and able to undergo all test procedures including neuroimaging and agree to longitudinal follow-up. Specific psychoactive medications were excluded. General inclusion/exclusion criteria are as follows:

1. Normal control subjects: Mini-Mental State Examination (MMSE) (Folstein) scores between 24 and 30 (inclusive), a Clinical Dementia Rating (CDR) of 0, non-depressed, non-MCI, and non-demented. The age range of normal subjects was roughly matched to that of MCI and AD subjects. Therefore, there should be minimal enrollment of normals under the age of 70.
2. MCI subjects: MMSE scores between 24 and 30 (inclusive), a memory complaint, objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, a CDR of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia.
3. Mild AD subjects: MMSE scores between 20 and 26 (inclusive), CDR of 0.5 or 1.0, and meeting NINCDS/ADRDA [29] criteria for probable AD.

2.2. Image preprocessing

MRI images from the ADNI database were preprocessed and segmented using the Statistical Parametric Mapping (SPM)

Table 1
Demographic data of patients in the database (ADNI 1075-T1).

Diagnosis	Number	Age	Gender (M/F)	MMSE
NC	229	75.97 ± 5.0	119/110	29.00 ± 1.0
MCI	401	74.85 ± 7.4	258/143	27.01 ± 1.8
AD	188	75.36 ± 7.5	99/89	23.28 ± 2.0

software [30]. SPM was initially designed for functional images, but it also provides routines for realignment, smoothing and spatial normalization into a standard space of T1-weighted images. Moreover, the template from the VBF package [31] was used for this purpose. It is worth mentioning that normalization routines preserve the amount of tissues and not the intensities [30]. Thus, images from ADNI database were resized to $121 \times 145 \times 121$ voxels with voxel sizes of 1.5 mm (sagittal) \times 1.5 mm (coronal) \times 1.5 mm (axial). The segmentation performed by SPM provides probability maps for GM and WM, considering in values in the range (0,1) for each voxel, related to its membership probability.

2.3. Feature extraction methods

Some dimension reduction techniques were applied to reduce the information contained in the brain images. In this sense, this

section describes the methods for feature extraction used in this paper.

2.3.1. Feature extraction based on principal component analysis

Principal component analysis (PCA) [32,33] has been called one of the most important and valuable results from applied linear algebra (linear transformation). PCA is used frequently in all forms of analysis because it is an efficient tool for extracting the most significant features from a dataset. It is often used in neuroimaging in order to reduce the original high dimensional space of the brain images to a lower dimensional subspace [34,35]. Furthermore, it has been successfully applied in neuroimage classification problems [33,36].

Mathematically, PCA generates an orthonormal basis vector that maximizes the scatter of all the projected samples. After the preprocessing steps, the n remaining voxels for each subject are rearranged into a vector form. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ be the sample

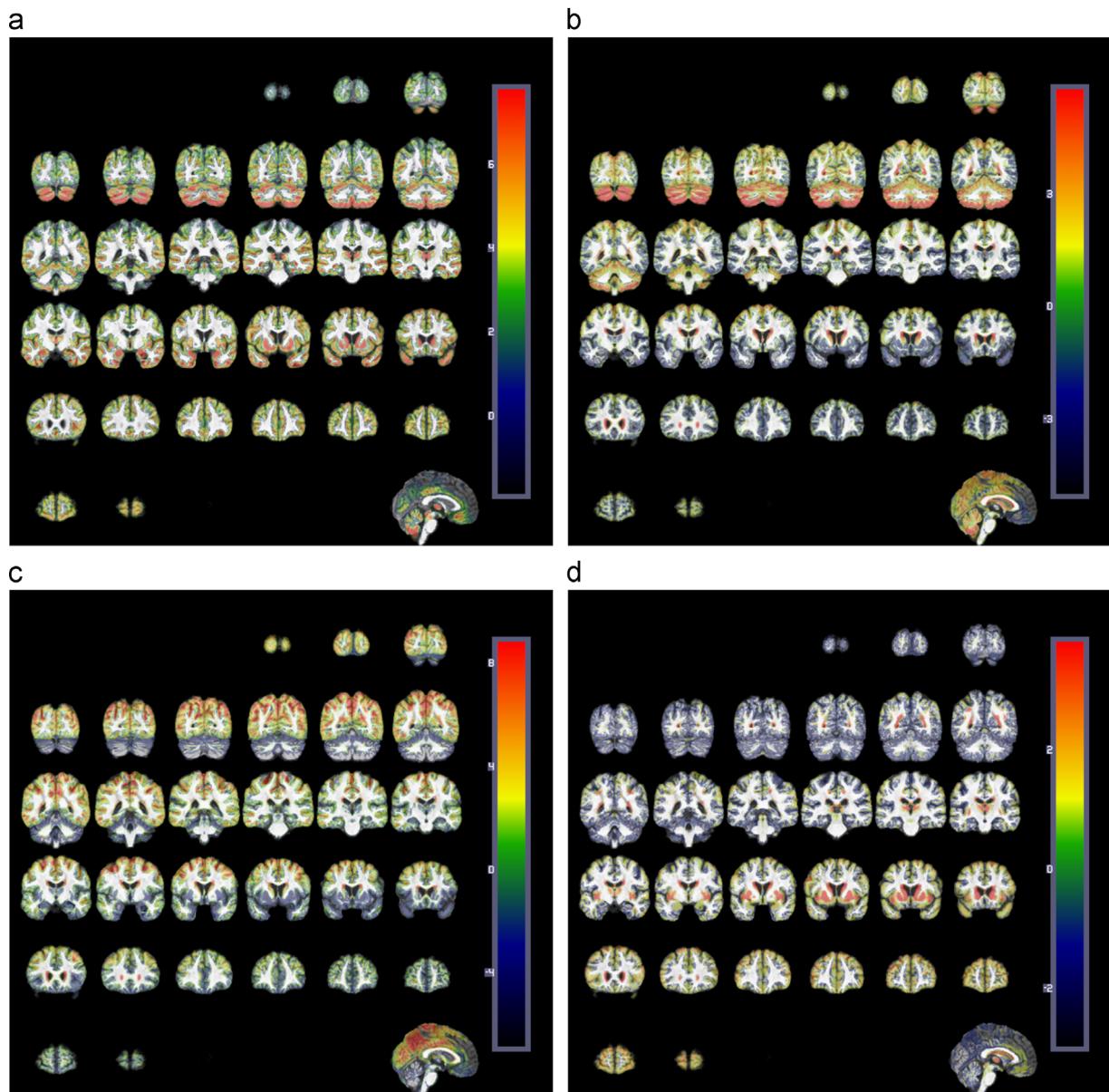


Fig. 1. The first four Eigenbrains extracted from an MRI reference image (gray level). They represent the principal components where original images will be projected onto to obtain a dimension reduction: (a) First eigenbrain; (b) second eigenbrain; (c) third eigenbrain; and (d) fourth eigenbrain.

set of these vectors, where N is the number of patients. After normalizing the vectors to unity norm and subtracting the large average, a new vector set $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ is obtained, where each \mathbf{y}_i represents an n -dimensional normalized vector, $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{in})^t$, $i = 1, 2, \dots, N$. The covariance matrix of the normalized vectors set is defined as:

$$\Sigma_Y = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_i \mathbf{y}_i^t = \frac{1}{N} \mathbf{Y} \mathbf{Y}^t \quad (1)$$

and the eigenvector and eigenvalue matrices Φ , Λ are computed as

$$\Sigma_Y \Phi = \Phi \Lambda \quad (2)$$

Note that $\mathbf{Y} \mathbf{Y}^t$ is an $n \times n$ matrix while $\mathbf{Y}^t \mathbf{Y}$ is an $N \times N$ matrix. If the sample size N is much smaller than the dimensionality n , then diagonalizing $\mathbf{Y}^t \mathbf{Y}$ instead of $\mathbf{Y} \mathbf{Y}^t$ reduces the computational complexity [15].

$$(\mathbf{Y}^t \mathbf{Y}) \Psi = \Psi \Lambda_1 \quad (3)$$

$$\mathbf{T} = \mathbf{Y} \Psi \quad (4)$$

where $\Lambda_1 = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ and $\mathbf{T} = [\Phi_1, \Phi_2, \dots, \Phi_N]$. Derived from the eigenface concept [37], and due to its still brain-like appearance, the eigenvectors or principal components (PCs) Φ_i , $i = 1, 2, \dots, N$ of the covariance matrix are called eigenbrains [38]. Fig. 1 shows the first four eigenbrains obtained with an MRI reference image (gray level) by the PCA algorithm.

2.3.2. Feature extraction based on partial least squares

PLS [39,40] is a statistical method that models sets of observed variables by means of latent variables. This method comprises regression and classification tasks as well as dimension reduction techniques and modeling tools. The objective general of PLS is to maximize the covariance between different sets of variables. Both the predictor (the observed variables) and predicted (response) variables are considered as a block of variables. It finds a linear regression model by projecting the predicted variables and the observable variables to a new space. PLS can be used as a regression tool or as a dimension reduction technique similar to PCA [41,42]. Mathematically, PLS is a linear algorithm that models the relation between two data sets: the observed variables $X \subset \mathbb{R}^N$ (representing the feature space of input) and $Y \subset \mathbb{R}^M$ (representing the labels). After observing n data samples from each block of variables, PLS decomposes the $n \times N$ matrix of zero mean variables \mathbf{X} and the $n \times M$ matrix of zero mean variables \mathbf{Y} into the regression models form [40,43]:

$$\mathbf{X} = \mathbf{T} \mathbf{P}^T + \mathbf{E} \quad (5)$$

$$\mathbf{Y} = \mathbf{U} \mathbf{Q}^T + \mathbf{F} \quad (6)$$

where \mathbf{T} and \mathbf{U} are $n \times p$ matrices of the p extracted score vectors (components, latent vectors), the $N \times p$ matrix \mathbf{P} and the $M \times p$ matrix \mathbf{Q} represent matrices of loadings with number of columns being the number of PLS components and $n \times N$ matrix \mathbf{E} and $n \times M$ matrix \mathbf{F} are the matrices of residuals (or error matrices). The x -score in \mathbf{T} are linear combinations of the x -variables, similarly, the y -score in \mathbf{U} are linear combinations of the y -variables. The model structure of PLS and PCA is the same in the sense that the data are first transformed into a set of a few intermediate linear latent variables (components) and these new variables are taken into account. The main difference between PLS and PCA is that the former creates orthogonal weight vectors by maximizing the covariance between the variables \mathbf{X} and \mathbf{Y} . Thus, PLS not only considers the variance of the samples but also considers the class labels [39].

Both, PCA and PLS decompose the data into two sets of variables named scores and loadings. Thus, we can perform an interesting analysis based on the concept of eigenbrains (similar to the concept of eigenfaces [37] used in face recognition systems). That way, loadings would be viewed as elementary brain images and scores would represent the quantity of loadings used for building a specific image. The PLS-based methodology maximizes the covariance taking into account the labels information, thus the PLS-brains contain the differences between the two classes. In addition, most of the variance is gathered by the first components and therefore, most of the differences are also gathered by the first components. Fig. 2 shows the representation of the first four PLS-brain obtained by the PLS algorithm. An implementation of the PLS and PCA algorithms has been developed using Matlab Software [44].

2.4. Support vector machines classifier

The support vector machine (SVM) is one of the most popular supervised learning algorithms that been applied to neuroimaging data in recent years [45–48]. SVM demonstrates good classification performance, and is computationally efficient for training with high dimensional data. Besides, SVM represents a set of related supervised learning methods widely used in pattern recognition, voice activity detection, classification and regression analysis [49–52]. It is introduced in order to separate a set of binary labeled training data with a hyperplane that is maximally distant from the two classes (called maximal margin hyperplane). The objective is to build a function $f: \mathbb{R}^N \rightarrow \pm 1$ using training data, that is, n -dimensional patterns \mathbf{x}_i and class labels y_i , so that f will correctly classify new examples (\mathbf{x}, y) :

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N) \in \mathbb{R}^N \times \pm 1 \quad (7)$$

Linear discriminant functions define decision hyperplanes in a multidimensional space, that is:

$$\mathbf{g}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 \quad (8)$$

where \mathbf{w} is the weight vector that is orthogonal to the decision hyperplane and w_0 is the threshold. The optimization task consists of finding the unknown parameters \mathbf{w}_i , $i = 1, 2, \dots, N$ and w_0 that define the decision hyperplane. Let \mathbf{x}_i , $i = 1, 2, \dots, n$ be the feature vectors of the training set, \mathbf{x} . These belong to either of the two classes, w_1 or w_2 . If the classes were linearly separable, the objective would be to design a hyperplane that classifies correctly all the training vectors. The hyperplane is not unique, and the selection process is focused on maximizing the generalization performance of the classifier, that is, the ability of the classifier, designed using the training set, to operate satisfactorily with new data. Among the different design criteria, the maximal margin hyperplane is usually selected since it leaves the maximum margin of separation between the two classes. Since the distance from a point \mathbf{x} to the hyperplane is given by $z = |\mathbf{g}(\mathbf{x})| / \|\mathbf{w}\|$, scaling \mathbf{w} and \mathbf{w}_0 so that the value of $\mathbf{g}(\mathbf{x})$ is $+1$ for the nearest point in w_1 and -1 for the nearest points in w_2 , reduces the optimization problem to maximizing the margin: $2 / \|\mathbf{w}\|$ with the constraints:

$$\mathbf{g}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \mathbf{w}_0 \geq 1, \quad \forall \mathbf{x} \in w_1 \quad (9)$$

$$\mathbf{g}(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \mathbf{w}_0 \leq -1, \quad \forall \mathbf{x} \in w_2 \quad (10)$$

When no linear separation of the training data is possible, SVM can work effectively in combination with kernel techniques such as quadratic, polynomial or radial basis function (RBF), so that the hyperplane defining the SVM corresponds to a non-linear decision boundary in the input space [53]. A kernel function is defined as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i) \varphi(\mathbf{x}_j) \quad (11)$$

The use of kernel functions avoids directly working in the high dimensional feature space, thus the training algorithm only depends on the data through dot products in Euclidean space, i.e., on terms of the form $\varphi(\mathbf{x}_i)\varphi(\mathbf{x}_j)$.

2.5. Automated classification of AD based on PLS and PCA methods with SVM classifier

In this work, we present a new classification approach for AD diagnosis which is based on segmented data (GM database and

WM database) and a voxel selection process. This process reduces the input space dimensionality in order to address the small sample size problem. Our developed CAD system is shown in Fig. 3. Firstly, structural MRI images are normalized and segmented (preprocessing). Then, a different binary mask for each tissue (GM and WM) is computed by averaging all the normal subject tissue images. Only the voxels that have an intensity above 10% of maximum intensity in the average image will be considered. This step reduces the number of voxels in the input space. For example, for data used in this work, the initial number of voxels per image

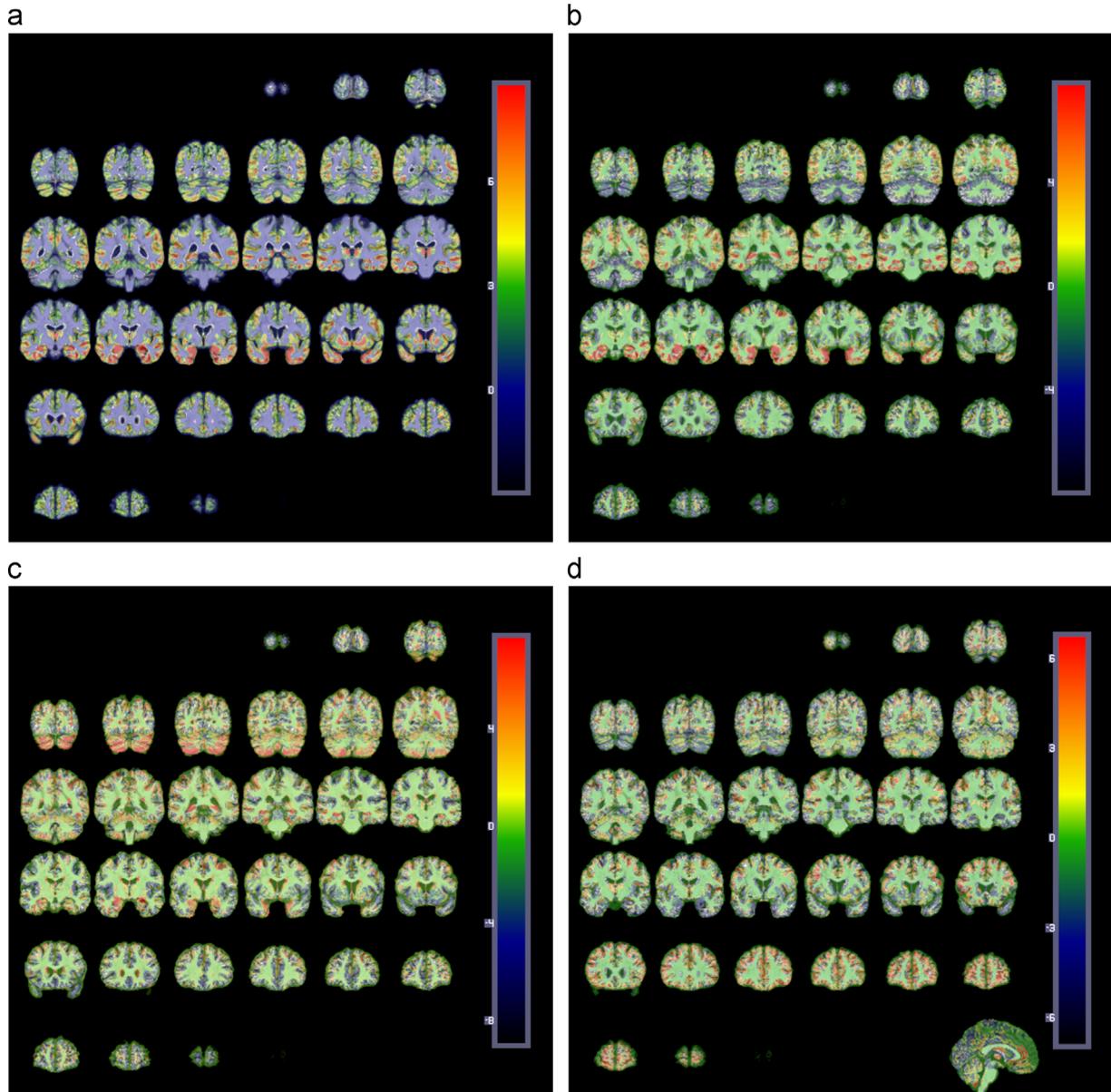


Fig. 2. Representation of the four PLS-brain of an MRI reference image (gray level) obtained by the PLS algorithm: (a) first PLSbrain; (b) second PLSbrain; (c) third PLSbrain; and (d) fourth PLSbrain.

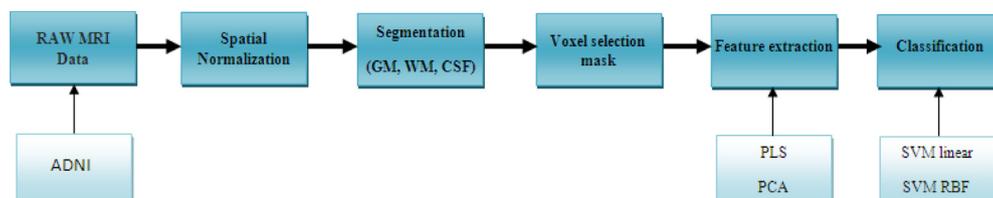


Fig. 3. Detailed schema of the proposed CAD system.

($121 \times 145 \times 121 = 2,122,945$) is reduced to 382,325. Secondly, we apply a PLS or PCA algorithm to compute score vectors for selected voxels. These methods are efficient and robust, and have been successfully used to model complex data [54] and, even, to develop CAD system for AD. Finally, these vectors are used as input for a statistical classifier. In order to test this approach, we developed a CAD system for AD using a Support Vector Machine (SVM) classifier. It was used to estimate the underlying class (NC, MCI or AD) of each subject. The performance of this system was estimated through a k -fold cross-validation strategy.

3. Experiments and results

In this paper, a CAD system has been developed using two feature extraction methods and different SVM classifiers. Several performance metrics are estimated by using cross-validation. Apart from the well-known accuracy rate of a classification procedure, which computes the proportion between correctly classified samples and total samples, sensitivity and specificity are the most widely used parameters to describe a diagnosis test. The accuracy (Acc), sensitivity (Sens) and

specificity (Spec) rates are defined as follows:

$$\text{Acc} = \frac{TP + TN}{TP + TN + FN + FP}, \quad \text{Sens} = \frac{TP}{TP + FN}, \quad \text{Spec} = \frac{TN}{TN + FP} \quad (12)$$

where TP is the number of true positives: number of AD patients correctly classified; TN is the number of true negatives: number of control subjects correctly classified; FP is the number of false positives: number of control subjects classified as AD patients; FN is the number of false negatives: number of AD patients classified as control subjects. Sensitivity and specificity are used to measure the proportion of actual positives or negatives which are identified correctly (e.g. the percentage of AD patients, or normal controls who are identified as such). These measures reveal the ability of a system to detect AD, MCI and NC patterns. Performance metrics are calculated using two cross-validation methods: k -fold cross-validation and Leave-One-Out validation [55,56]. k -fold has been used to assess the discriminative accuracy of different multivariate analysis methods applied to the discrimination of frontotemporal dementia from AD [25] and in classifying atrophy patterns based on MRI data [26]. It considers k randomly selected sets of AD patients, MCI and NC. It iteratively holds out a set for testing purposes, and train the classifier with the remaining sets, so that each set is

Table 2
Statistical measures of performances of the multivariate approach (PLS and PCA) for group 1 (NC vs AD), using 8 components for both features extraction techniques and two classifiers (SVM linear and RBF) with $k=10$.

	PLS			PCA			
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	
GM	87.53	88.65	86.17	85.61	89.08	81.38	SVM lin
WM	85.61	87.34	83.51	81.77	84.28	78.72	
(GM+WM)	88.49	91.27	85.11	87.77	89.96	85.11	
GM	87.29	87.77	86.70	83.93	86.26	81.38	SVM RBF
WM	84.41	85.59	82.98	81.29	83.41	78.72	
(GM+WM)	88.49	90.39	86.17	87.53	90.39	84.04	

Table 3
Statistical measures of performances of the multivariate approach (PLS and PCA) for group 2 (NC vs MCI) using 8 components for both features extraction techniques and two classifiers (SVM linear and RBF) with $k=10$.

	PLS			PCA			
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	
GM	77.57	76.76	71.90	75.41	78.38	72.43	SVM lin
WM	80.54	79.46	81.62	75.14	77.3	72.98	
(GM+WM)	81.89	82.16	81.62	78.92	80	77.84	
GM	76.22	81.62	70.81	72.97	75.68	70.27	SVM RBF
WM	81.35	76.22	80.54	72.70	64.86	80.54	
(GM+WM)	80.27	73.51	82.70	73.24	71.89	74.59	

Table 4
Statistical measures of performances of the multivariate approach (PLS and PCA) for group 3 (MCI vs AD), using 8 components for both feature extraction techniques and two classifiers (SVM linear and RBF) with $k=10$.

	PLS			PCA			
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	
GM	77.03	74.59	79.46	72.70	72.43	72.97	SVM lin
WM	87.03	88.65	85.41	79.19	82.16	76.22	
(GM+WM)	85.41	87.03	83.78	81.89	84.86	78.92	
GM	76.22	74.59	77.84	71.08	71.89	70.27	SVM RBF
WM	85.95	85.41	86.49	74.86	74.05	75.68	
(GM+WM)	85.41	85.95	84.86	76.77	74.59	78.92	

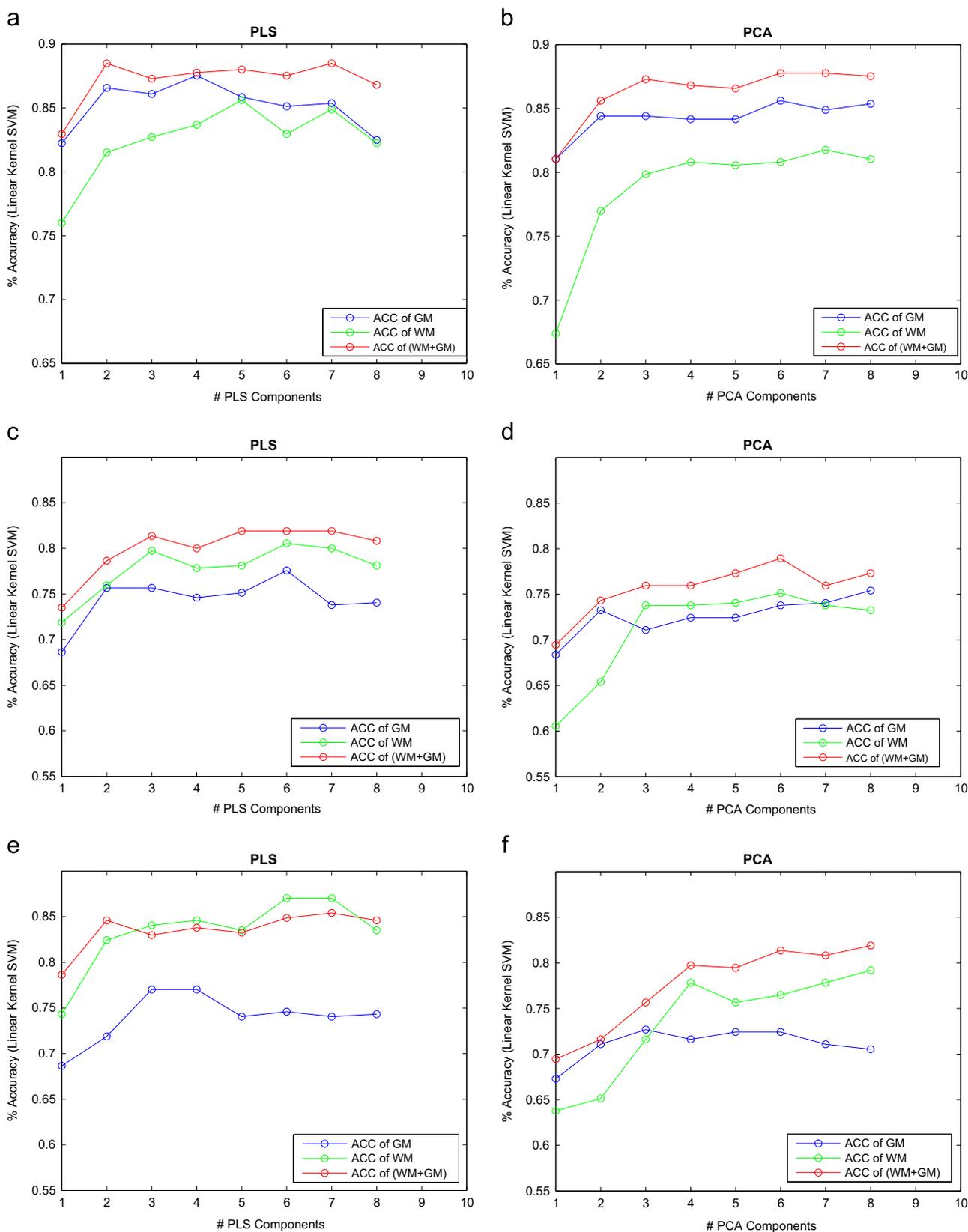


Fig. 4. SVM classification: Values of Accuracy (%) computed for ADNI database in function of number of component for features extraction techniques PLS (left) and PCA (right): (a) and (b) are the results of group 1, (c) and (d) are the results of group 2, and (e) and (f) are the results of group 3.

left out once. This is repeated for all the k sets, and an average value of the evaluation parameters is computed. Leave-One-Out is a particular case of k -fold, in which only one image is selected as test image.

In this work, two feature extraction methods (PCA and PLS) and two SVM classifiers (using linear or RBF kernel) were evaluated in order to develop more accurate CAD system for AD. Since our purpose is to distinguish between healthy subjects, MCI and AD

patients, firstly, we trained the CAD system with only normal controls and AD images (group 1). Secondly, we have tested our system using MCI and AD images (group 2) and finally we have trained the CAD system with MCI and AD images (group 3). Tables 2, 3 and 4 show the statistical measures obtained using PLS and PCA methods. Also, performance of these feature extraction methods was calculated by means of k -fold cross-validation with a number of folds equal to 10 ($k=10$).

3.1. Classification results for group 1 (NC vs. AD)

In this section we present the classification results obtained in the first experiment, which consisted on distinguishing between NC and AD subjects. In this group, we test our method with all the database (229 NC, 401 MCI and 188 AD). Table 2 shows the values of the accuracy, sensitivity and specificity for the different brain tissues and compares different feature extraction methods and SVM classifiers. Linear SVM yielded higher accuracy rates for both PLS and PCA methods than RBF. Furthermore, when we compare the results obtained from PLS and PCA, we find that PLS feature extraction and linear kernel SVM yielded the best accuracy rates. Combining features extracted from GM and WM segmentation reported a classification accuracy of 88.49% for PLS and linear SVM (sensitivity=91.27% and specificity=85.11%) compared to 87.53% for GM only (sensitivity=88.65% and specificity=86.17%) and 85.61% for WM only (sensitivity=87.34% and specificity=83.51%). As a conclusion, combining features extracted from both GM and WM tissue distributions increases the classification and accuracy of the classifier.

3.2. Classification results for group 2 (NC vs. MCI)

The most difficult classification task concerning the ADNI database is to distinguish between NC and MCI patients, due to the wide range spanned by the features extracted from MCI patients. In this group, we have used only 370 MRI images (185 NC and 185 MCI subjects) from the ADNI database described in Section 2.1. Using PLS and linear SVM, the combination of features extracted from GM and WM segmentation provided the highest accuracy, 81.89% (sensitivity=82.16% and specificity=81.62%), whereas the features extracted from GM or WM alone reported a classification accuracy of 77.57% and 80.54% respectively. Overall, we note that combining features extracted from both GM and WM

tissue yielded the highest accuracy value. All results related to this classification model are reported in Table 3.

3.3. Classification results for group 3 (MCI vs. AD)

Table 4 shows the classification results obtained in the last experiment, which consisted on distinguishing between MCI and AD subjects using different SVM classifiers. Also, in this group we have used the same number of images as group 2. Combining features extracted from GM and WM segmentation reported a classification accuracy of 85.41% for PLS and linear SVM (sensitivity=87.03% and specificity=83.78%) compared to 77.03% for GM only (sensitivity=74.59% and specificity=79.46%) and 87.03% for WM only (sensitivity=88.65% and specificity=85.41%). These results showed that the most important change in the brain occurs more in the white matter than in the gray matter tissue brain [59].

As shown in these tables, the both methods analyzed in this work highlight that the combination of features extracted from GM and WM tissue distributions gives better accuracy, sensitivity and specificity than using different brain tissues separately. As a result, combining the different features extracted from both brain tissues (GM+WM) of patients with classification methods produces a valid approach to perform a CAD system for AD.

The accuracy of the different approaches depends on the size of the feature vector. For PLS and PCA based algorithm, the maximum size of the feature vectors is the number of images which are in database minus two. However, this number may be reduced by selecting only the most important components. Fig. 4 shows the accuracy rates of the different groups achieved with both approaches in function of number of components selected (number of components=8).

4. Discussion

As shown in Section 3, the both methods analyzed in this paper are valid approaches to develop CAD system for AD. Besides, they achieve good values of accuracy, sensitivity and specificity. Fig. 4 compares the classification performance of the PLS and PCA feature extraction methods with linear SVM classifier. Note that, the performance of the CAD system improves with the number of PCA or PLS components used as input features for classification which up to a maximum stable value. PLS outperforms PCA as a feature extraction technique yielding peak values of

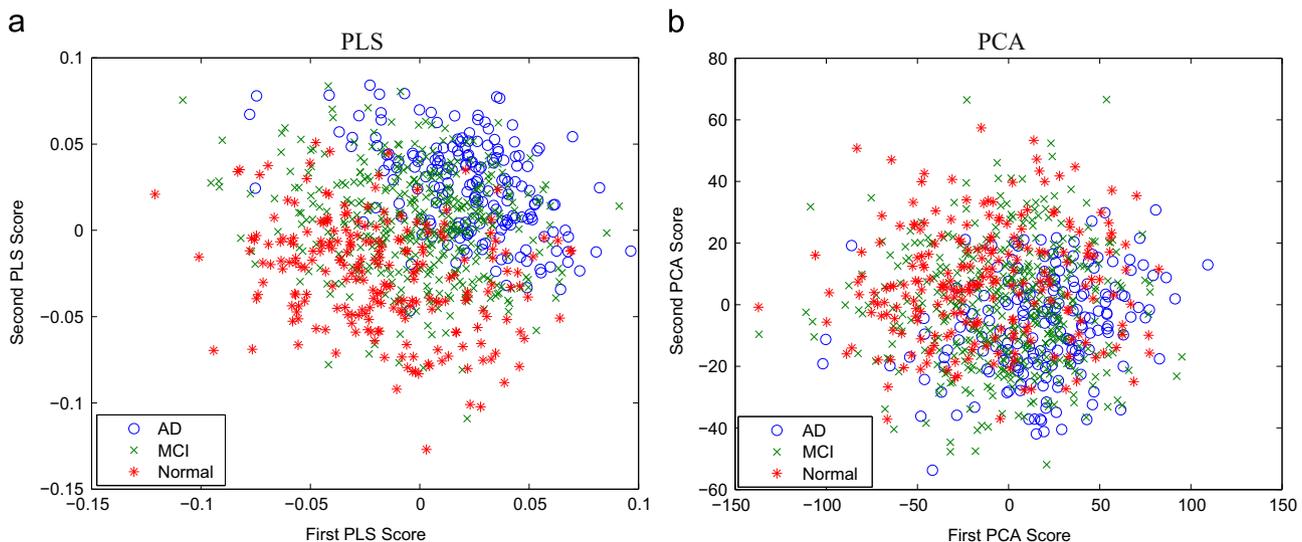


Fig. 5. Representation of all images of the database using only the two first Scores of PLS and PCA.

sensitivity=91.27%, specificity=85.11% and accuracy=88.49% when compared to PCA that just yields sensitivity=89.96%, specificity=85.11% and accuracy=87.77% (see Fig. 4(a) and (b)). The successful rate of PLS based method reached 88.49% for group 1. However, it is decreased for group 2 and 3 (78.92% ,85.4% respectively) when MCI images are included (see Fig. 4(c) and (d)). This is probably due to the high variability of the MCI pattern of each image. As a consequence, the classification task becomes more difficult (see Fig. 5).

As shown in Fig. 4(c) and (d), the classification result using only WM brain tissue is better than using only GM. This result confirms the previous study [57] in which the modification in the pattern of

brain atrophy in early study (MCI) of disease occurs in the white matter tissue. Besides, elder subjects are likely to have WM structural abnormalities caused by leucoaraiosis or other diseases [58]. This abnormality in the WM brain tissue for patients with AD or MCI can make the structure very different from normal controls. Thus, the classification result in the WM can be better than in GM of brain images. Furthermore, the classification results for group 3 (see 4(e) and 4(f)) confirm that neurodegeneration starts in the WM and spreads to GM with the progression of the disorder.

It is worth noting that CAD systems are reproducing current medical knowledge since they have been trained with samples labeled by physicians. For this reason, statistical measures reported

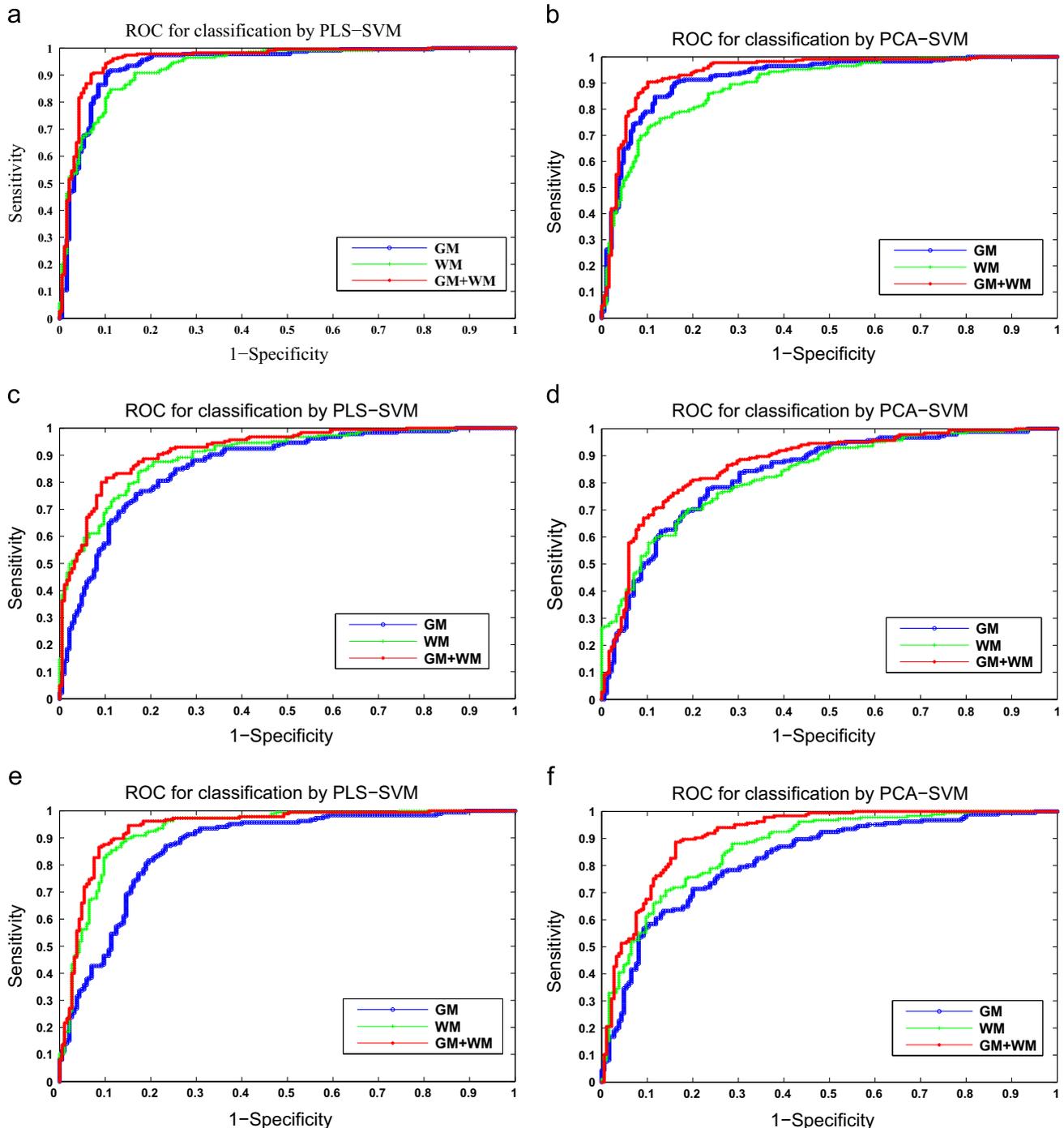


Fig. 6. ROC curves for statistical measures obtained with different feature extraction methods using a linear classifier (SVM) for all ADNI images: Group 1((a) and (b)), group 2 ((c) and (d)) and group 3 ((e) and (f)).

in this paper are an estimation about how a trained system is able to reproduce a medical diagnosis performed by experts [60]. Thus, some possible errors in the labeling process can modify the decision hyperplanes of the classifiers, specifically considering that the labels were assigned based on the scores obtained by patients in cognitive tests (as MMSE and CDR). In this work, we have previously shown that combining features extracted from GM and WM segmentation gives good classification accuracy using PLS and PCA methods. In addition, the PLS based method yields better classification results with smaller computational time than the PCA approach. The feature extraction approaches proposed in this paper achieve good classification performance when a linear SVM classifier is used. Furthermore, non-linear classifiers like RBF require more samples or smaller feature vectors than linear classifier in order to give good classification results. A more interesting approach consists of selecting only some components of feature extraction methods, such as FDR as it is described in [36] and the Out-Of-Bag (OOB) error in [61]. These previous methods achieve that using the first PCA/PLS components is optimal for classification purposes. In the development of our CAD system, we have selected only the 8 first PLS and PCA components. A higher number of components may worsen the classification results since it increases the input space. In order to evaluate our CAD system, we use the receiver operating characteristic (ROC) curves. Fig. 6 contains the sensitivity and 1-specificity values of the classification results in the ROC space considering the PLS and PCA techniques.

These plots show a trade-off between the specificity and sensitivity of the CAD system when varying any of the input parameters. In addition, the closer to the left upper corner values are the better [62] (the combination of brain tissues, i.e. GM+WM).

5. Conclusion

A computer aided diagnosis system (CAD) for assisting the early detection of the Alzheimer's disease was shown in this paper. The system was developed by combining the different brain tissues and exploiting the two features extraction methods (PLS and PCA) in order to improve the classification of MRI images and to diagnose Alzheimer disease. The both multivariate approaches used in our proposed methodology allow the dimensionality reduction of the feature vector in order to surmount the small sample size problem. The difficulty arises in classification problems when the dimension of the feature vector is very high compared to the number of available samples. The PLS based method uses score vectors as features and performs with the PCA method an initial reduction of the input space by applying a binary mask according an intensity threshold that discard low intensity voxels. Both methods are tested with two classifiers (linear and RBF kernel) based on SVM. The resulting CAD system was based on segmented MRI images classification from the ADNI database. It was estimated using a k -fold cross-validation methodology. The PLS method reached peak accuracy rate when we distinguish between controls and AD classes. In general, it outperforms several methods such as PCA approach.

Acknowledgments

This work was partly supported by the MICINN under the TEC2012–34306 project and the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía, Spain) under the Excellence Projects P09–TIC–4530 and P11–TIC–7103.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI

(Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

- [1] S. Li, F. Shi, F. Pu, X. Li, T. Jiang, S. Xie, Y. Wang, Hippocampal shape analysis of Alzheimer disease based on machine learning methods, *AJNR Am. J. Neuroradiol.* 28 (7) (2007) 1339–1345.
- [2] R. Brookmeyer, E. Johnson, K. Ziegler Graham, H. Arrighi, Forecasting the global burden of Alzheimer's disease, *Alzheimer's Dement.* 3 (3) (2007) 186–191.
- [3] Alzheimer's Association, Alzheimer's disease facts and figures, *Alzheimer's Dement.* 9 (2) (2013) 208–245.
- [4] J.T. O'Brien, Role of imaging techniques in the diagnosis of dementia, *Br. J. Radiol.* 80 (2007) 71–77.
- [5] M.L. Ries, C.M. Carlsson, H.A. Rowley, M.A. Sager, C.E. Gleason, S. Asthana, S.C. Johnson, Magnetic resonance imaging characterization of brain structure and function in mild cognitive impairment, a review, *J. Am. Geriatr. Soc.* 56 (2008) 920–934.
- [6] C. Davatzikos, Y. Fan, X. Wu, D. Shen, S.M. Resnick, Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging, *Neurobiol. Aging* 29 (4) (2008) 514–523.
- [7] S. Kloppel, C.M. Stonnington, C. Chu, B. Draganski, R.I. Scahill, J.D. Rohrer, N.C. Fox, C.R. Jack Jr, J. Ashburner, R.S.J. Frackowiak, Automatic classification of MR scans in Alzheimer's disease, *Brain* 131 (3) (2008) 681–689.
- [8] N.C. Fox, S. Cousins, R. Scahill, R.J. Harvey, M.N. Rossor, Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects, *Arch. Neurol.* 57 (2000) 339–344.
- [9] C.R. Jack Jr, M. Slomkowski, S. Gracon, T.M. Hoover, J.P. Felmlee, K. Stewart, Y. Xu, M. Shiung, P.C. O'Brien, R. Cha, D. Knopman, R.C. Petersen, MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD, *Neurology* 60 (2003) 253–260.
- [10] P. Scheltens, N. Fox, F. Barkhof, C. De Carli, Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion, *Lancet. Neurol.* 1 (2002) 13–21.
- [11] P.M. Thompson, K.M. Hayashi, G. deZubicar, A.L. Janke, S.E. Rose, J. Semple, D. Herman, M.S. Hong, S.S. Dittmer, D.M. Doddrell, A.W. Toga, Dynamics of gray matter loss in Alzheimer's disease, *J. Neurosci.* 23 (2003) 994–1005.
- [12] Y. Zhang, L. Wu, S. Wang, Magnetic resonance brain image classification by an improved artificial bee colony algorithm, *Progr. Electromagnet. Res.* 116 (2011) 65–79.
- [13] A. Oikonomou, I.S. Karanasiou, N.K. Uzunoglu, Phased-array near field radio-metry for brain intracranial applications, *Progr. Electromagnet. Res.* 109 (2010) 345–360.
- [14] N.C. Fox, E.K. Warrington, P.A. Freeborough, P. Hartikainen, A.M. Kennedy, J.M. Stevens, M.N. Rossor, Presymptomatic hippocampal atrophy in Alzheimer's disease: a longitudinal MRI study, *Brain* 119 (1996) 2001–2007.
- [15] C.R. Jack Jr, R.C. Petersen, Y.C. Xu, S.C. Waring, P.C. O'Brien, E.G. Tangalos, G.E. Smith, R.J. Ivnik, E. Kokmen, Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease, *Neurology* 49 (1997) 786–794.
- [16] K. Juottonen, M.P. Laakso, K. Partanen, H. Soininen, Comparative MR analysis of the entorhinal cortex and hippocampus in diagnosing Alzheimer disease, *AJNR Am. J. Neuroradiol.* 20 (1999) 139–144.
- [17] R.J. Killiany, M.B. Moss, M.S. Albert, T. Sandor, J. Tieman, F. Jolesz, Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease, *Arch. Neurol.* 50 (1993) 949–954.

- [18] M.P. Laakso, H. Soininen, K. Partanen, M. Lehtovirta, M. Hallikainen, T. Hanninen, E.L. Helkala, P. Vainio, P.J. Riekkinen Sr, MRI of the hippocampus in Alzheimer's disease: sensitivity, specificity, and analysis of the incorrectly classified subjects, *Neurobiol. Aging* 19 (1) (1998) 23–31.
- [19] S. Lehericy, M. Baulac, J. Chiras, L. Pierot, N. Martin, B. Pillon, B. Deweer, B. Dubois, C. Marsault, Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease, *AJNR Am. J. Neuroradiol.* 15 (1994) 929–937.
- [20] O. Colliot, G. Chetelat, M. Chupin, B. Desgranges, B. Magnin, H. Benali, B. Dubois, L. Garnero, F. Eustache, S. Lehericy, Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus, *Radiology* 248 (1) (2008) 194–201.
- [21] J.H. Morra, Z. Tu, L.G. Apostolova, A.E. Green, C. Avedissian, S.K. Madsen, N. Parikshak, A.W. Toga, C.R. Jack Jr, N. Schuff, M.W. Weiner, P.M. Thompson, Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment and elderly controls, *Neuroimage* 45 (1) (2009) S3–S15 <http://dx.doi.org/10.1016/j.neuroimage.2008.10.043>.
- [22] K. Kantarci, Magnetic resonance markers for early diagnosis and progression of Alzheimer's disease, *Expert Rev. Neurother.* 5 (5) (2005) 663–670.
- [23] E. Westman, A. Simmons, Y. Zhang, J.S. Muehlboeck, C. Tunndard, Y. Liu, L. Collins, A. Evans, P. Mecocci, B. Vellas, M. Tsolaki, I. Kloszewska, H. Soininen, S. Lovestone, C. Spenger, L.O. Wahlund, A. Consortium, Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls, *Neuroimage* 54 (2) (2011) 1178–1187.
- [24] P. Padilla, J.M. Górriz, J. Ramírez, D. Salas-González, I. Illan, NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease, *IEEE Trans. Med. Imaging* 99 (1) (2011).
- [25] J.M. Górriz, F. Segovia, J. Ramírez, A. Lassl, D. Salas-González, GMM based spect image classification for the diagnosis of Alzheimer's disease, *Appl. Soft Comput.* 11 (2) (2011) 2313–2325.
- [26] J. Ramírez, J.M. Górriz, F. Segovia, R. Chaves, D. Salas-González, M. López, Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification, *Neurosci. Lett.* 472 (2) (2010) 99–103.
- [27] F. Segovia, J.M. Górriz, J. Ramírez, D. Salas-González, I. Álvarez, M. López, Classification of functional brain images using a GMM-based multivariate approach, *Neurosci. Lett.* 474 (1) (2010) 58–62.
- [28] R.P.W. Duin, Classifiers in almost empty spaces, in: *Proceedings 15th International Conference on Pattern Recognition*, IEEE vol. 2, 2000, pp. 1–7.
- [29] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, E.M. Stadlan, Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease, *Neurology* 34 (7) (1984) 939–944.
- [30] J. Ashburner, Group T SPM8, Functional Imaging Laboratory, Institute of Neurology 12, Queen Square, London WC1N 3BG, UK, 2011.
- [31] Psychiatriy SBM8, Vbm toolboxes, University of Jena. URL (<http://dbm.neuro.uni-jena.de/vbm8/VBM8-Manual.pdf>), 2013.
- [32] I. Jolliffe, *Principal Component Analysis*, Springer Verlag, New York, 1986.
- [33] M. López, J. Ramírez, J.M. Górriz, D. Salas-Gonzalez, I. Álvarez, F. Segovia, Automatic tool for the Alzheimers disease diagnosis using PCA and Bayesian classification rules, *IET Electron. Lett.* 45 (8) (2009) 389–391.
- [34] A. Andersen, D.M. Gash, M.J. Avison, Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework, *J. Magn. Reson. Imaging* 17 (1999) 795–815.
- [35] U. Yoon, J.M. Lee, K. Im, Y.W. Shin, B.H. Cho, I.Y. Kim, J.S. Kwon, S.I. Kim, Pattern classification using principal components of cortical thickness and its discriminative pattern in Schizophrenia, *NeuroImage* 34 (2007) 1405–1415.
- [36] M. López, J. Ramírez, J.M. Górriz, I. Álvarez, D. Salas-Gonzalez, F. Segovia, R. Chaves, SVM-based cad system for early detection of the Alzheimer's disease using kernel PCA and LDA, *Neurosci. Lett.* 464 (4) (2009) 233–238.
- [37] M. Turk, A. Pentland, Eigenfaces for recognition, *J. Cognit. Neurosci.* 3 (1) (1991) 71–86.
- [38] I. Álvarez, J.M. Górriz, J. Ramírez, D. Salas-Gonzalez, M. López, C.G. Puntonet, F. Segovia, Alzheimer's diagnosis using eigenbrains and support vector machines, *IET Electron. Lett.* 45 (7) (2009) 342–343.
- [39] S. Wold, H. Ruhe, H. Wold, W.J. Dunn III, The collinearity problem in linear regression: the partial least squares approach to generalized inverse, *J. Sci. Stat. Comput.* 5 (1984) 735–743.
- [40] J. Ramírez, J.M. Górriz, F. Segovia, R. Chaves, D. Salas-Gonzalez, M. López, Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification, *Neurosci. Lett.* 472 (2010) 99–103.
- [41] D.V. Nguyen, D.M. Rocke, Tumor classification by partial least squares using microarray gene expression data, *Bioinformatics* 18 (1) (2002) 39–50.
- [42] R. Rosipal, L.J. Trejo, Kernel PLS-SVC for linear and nonlinear classification, in: *Proceedings of the Twentieth International Conference on Machine Learning*, 2003, pp. 640–647.
- [43] P. Bastien, V.E. Vinzi, M. Tenenhaus, PLS generalised linear regression, *Comput. Stat. Data Anal.* 48 (2005) 17–46.
- [44] F. Segovia, Ch. Bastin, E. Salmon, J.M. Górriz, J. Ramírez, Ch. Phillips, Combining PET Images and neuropsychological test data for automatic diagnosis of Alzheimer's disease, *PLoS One* 9 (2) (2014).
- [45] R.C. Craddock, P.E. Holtzheimer III, X.P. Hu, H.S. Mayberg, Disease state prediction from resting state functional connectivity, *Magn. Reson. Med.* 62 (2009) 1619–1628.
- [46] Y. Fan, S.M. Resnick, X. Wu, C. Davatzikos, Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study, *NeuroImage* 41 (2) (2008) 277–285.
- [47] S. Kloppel, C.M. Stonnington, C. Chu, B. Draganski, R.I. Scahill, J.D. Rohrer, N.C. Fox, C.R. Jack Jr, J. Ashburner, R.S. Frackowiak, Automatic classification of MR scans in Alzheimer's disease, *Brain* 131 (2008) 681–689.
- [48] J. Mourao Miranda, A.L. Bokde, C. Born, H. Hampel, M. Stetter, Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data, *NeuroImage* 28 (2005) 980–995.
- [49] V.N. Vapnik, *Estimation of Dependences Based on Empirical Data*, Springer Verlag, New York, 1982.
- [50] C.J.C. Burges, A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discov.* 2 (2) (1998) 121–167.
- [51] J. Shawe-Taylor, N. Cristianini, *Support Vector Machines and Other Kernel-Based, Learning Methods*, Cambridge University Press, UK, 2000.
- [52] B. Scholkopf, A.J. Smola, *Learning with Kernels*, The MIT Press, USA, 2001.
- [53] R. Chaves, J. Ramírez, J.M. Górriz, M. López, D. SalasGonzalez, I. Alvarez, F. Segovia, SVM-based computer-aided diagnosis of the Alzheimer's disease using t-test NMSE feature selection with feature correlation weighting, *Neurosci. Lett.* 461 (2009) 293–297.
- [54] S. Wiklund, E. Johansson, L. Sjöström, E.J. Mellerowicz, U. Edlund, J.P. Shockcor, J. Gottfries, T. Moritz, J. Trygg, Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models, *Anal. Chem.* 80 (1) (2008) 115–122.
- [55] Y. Fan, D. Shen, C. Davatzikos, Classification of structural images via high-dimensional image warping, robust feature extraction, and SVM, *Med. Image Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv.* 8 (2005) 1–8.
- [56] Z. Lao, D. Shen, Z. Xue, B. Karacali, S.M. Resnick, C. Davatzikos, Morphological classification of brains via high-dimensional shape transformations and machine learning methods, *NeuroImage* 21 (1) (2004) 46–57.
- [57] T. Gold Brian, David K. Powell, Anders H. Andersen, Charles D. Smith, Alterations in multiple measures of white matter integrity in normal women at high risk for Alzheimer's disease, *NeuroImage* 52 (2010) 1487–1494.
- [58] Remi Cuingnet, Emilie Gerardin, Jérôme Tessieras, Guillaume Auzias, Stéphane Lehricy, Marie-Odile Habert, Marie Chupin, Habib Benali, Olivier Colliot, and the Alzheimer's disease neuroimaging initiative, Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database, *NeuroImage* 56 (2) (2011) 766–781.
- [59] Xekardaki Aikaterini, Giannakopoulos Panteleimon, Haller Sven, White matter changes in bipolar disorder, Alzheimer disease, and mild cognitive impairment: new insights from DTI, *J. Aging Res.* 2011 (2011) 1–10.
- [60] F. Segovia, J.M. Górriz, J. Ramírez, D. Salas-Gonzalez, I. Álvarez, M. López, R. Chaves, The Alzheimer's Disease Neuroimaging Initiative1, a comparative study of feature extraction methods for the diagnosis of Alzheimer's disease using the ADNI database, *Neurocomputing* 75 (2012) 64–71.
- [61] F. Segovia, J.M. Górriz, J. Ramírez, D. Salas-Gonzalez, I. Álvarez, Early diagnosis of Alzheimer's disease based on partial least squares and support vector machine, *Expert Syst. Appl.* 40 (2013) 677–683.
- [62] C. Metz, Basic principles of ROC analysis, *Semin. Nucl. Med.* 4 (8) (1978) 283–298.



L. Khedher received the B.Sc. degree in electronic Engineering and the Master degree in microelectronic and instrumentation from the University of Monastir, Tunisia in 2009 and 2011 respectively. Then, she is actually a Phd student in the SIPBA research group, in the Dept. Signal Theory, Networking and Communications at the University of Granada (2013). Her research interests at present lie in the field of Signal Processing and Biomedical Applications.



J. Ramírez received the M.A.Sc. degree in Electronic Engineering in 1998, and the Ph.D degree in Electronic Engineering in 2001, all from the University of Granada. Since 2001, he is a Full professor at the Dept. of Signal Theory Networking and Communications of the University of Granada (Spain). His research interest includes signal processing and biomedical applications including brain image processing, robust speech recognition, speech enhancement, voice activity detection, seismic signal processing and implementation of high performance digital signal processing systems. He has coauthored more than 150 technical journal and conference papers in these areas. He has served as reviewer for several international journals and conferences.



J.M. Górriz received a BSc degree in physics (2000) and a BSc in electronic engineering (2001) from the University of Granada, Spain. He then earned a PhD from the University of Cádiz, Spain (2003) and a PhD from the University of Granada, Spain (2006). He is currently a Full Professor with the Department of Signal Theory, Networking and Communications at the University of Granada. He has coauthored more than 200 technical journals and conference papers in these areas and has served as editor for several journals and books. His present interests lie in the field of statistical signal processing and its application to speech and medical image processing.



Fermin Segovia received the B.Sc. degree in Computer Science in 2006, and the Master degree in Computer Engineering and Networks in 2009, both from the University of Granada, Spain. He is currently a Ph.D. student of the Department of Signal Theory, Networking and Communications at the University of Granada. His main research interests include Image Processing and Biomedical Applications.



A. Brahim received the B.Sc. degree in electronic Engineering and the Master degree in microelectronic and instrumentation from the University of Monastir, Tunisia in 2009 and 2011 respectively. Then, he earned a Ph.D. Grant from Erasmus Mundus Al.idrisi with the SIPBA research group, in the Dept. Signal Theory, Networking and Communications at the University of Granada (2012). His research interests at present lie in the field of Signal Processing and Biomedical Applications.