

Assessment of Linear and Non-linear Feature Projections for the Classification of 3-D MR Images on Cognitively Normal, Mild Cognitive Impairment and Alzheimer's Disease

Marcelo R. Moura Araújo, Katia M. Poloni, and Ricardo J. Ferrari(✉)

Department of Computing, Federal University of São Carlos, São Carlos, SP 13565-905, Brazil
rferrari@ufscar.br
https://www.bipgroup.dc.ufscar.br

Abstract. Alzheimer's disease (AD) is an age-related neurodegenerative disease and the most common form of dementia. It is a brain disorder that impacts the daily life of the patient due to memory loss and cognitive changes. Due to population aging and the fact that dementia incidence rising sharply at ages greater than 75, AD has become a major public health problem. Currently, hippocampal atrophy assessed on structural magnetic resonance (MR) images is the most used imaging biomarker of AD. Among many methods applied for automated classification of cognitively normal (CN), mild cognitive impairment (MCI), and AD, the linear PCA projection method, also known as eigenbrain, has shown effectiveness in AD subject prediction even with its restriction of using only linear projections. This study investigates both linear (PCA) and non-linear (kernel PCA) projection method performances to classify 3-D structural MR images on the CN, MCI, and AD classes. Support vector machines (SVMs) were trained using feature vectors with different space dimensions obtained by projecting MR study images into the previously created "eigenbrain spaces." We tested the method using different kernel functions for both cases, the eigenbrain projections and SVM classifiers. We also conducted our analyses separately using the whole-brain and the gray-matter (GM) regions. Comparison results between PCA and kernel PCA methods showed that non-linear projections improved the classification of MR images in both class groups, particularly when processing the GM region - CN × MCI (PCA: AUC = 0.63 versus KPCA: AUC = 0.76), MCI × DA (PCA: AUC = 0.67 versus KPCA: AUC = 0.74), and CN × AD (PCA: AUC = 0.86 versus KPCA: AUC = 0.89).

Keywords: Alzheimer's disease · Eigenbrain · kernel PCA · Support vector machines · Magnetic resonance imaging

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12950, pp. 18–33, 2021.
https://doi.org/10.1007/978-3-030-86960-1_2

1 Introduction

Alzheimer's disease (AD) is the most devastating brain disorder of elderly humans and the most common cause of dementia, accounting for 70 to 80% of all cases. There is currently no cure or even an effective treatment available that will alter the progressive course of AD, which can eventually lead to death. In 2019, nearly 50 million patients were diagnosed with AD globally, with a prediction of 115.5 million by 2050. As the world's population is rapidly aging, AD has become one of the most critical public health problems [35].

Histologically, AD is caused by neuronal death, caused mainly by two factors: the excessive accumulation of amyloid-β protein extracellular plaques that form senile plaques [11], which have a toxic effect on neurons and blocks the synaptic connections; and the deposition of hyperphosphorylated tau protein in conjunction with the neuronal microtubules disassociation, that creates intracellular neurofibrillary tangles [6], which causes a fail on transportation of nutrients and other essential elements between cells, causing neuronal death. While many neurotransmitters are affected, there is widespread loss of neurons containing acetylcholine, which correlates well with the degree of cognitive impairment observed clinically. Both neurofibrillary tangles and senile plaques occur in normal aging but are more numerous and widespread in AD [9]. This entire process is responsible for the inevitable downstream consequence of neurodegeneration that is brain atrophy.

Currently, the diagnosis of AD is based on clinical criteria in which doctors conduct tests to assess memory impairment and other thinking skills of a patient. They also perform a series of tests to rule out other possible causes of impairment. In addition to the clinical assessment, magnetic resonance imaging (MRI) has been playing an increasingly important role in helping doctors to identify structural brain changes and, consequently, to support their decision-making process [10,13], since this imaging technique allows to obtain high-quality images to observe subtle changes in anatomical brain structures [13]. Such information has also been used to build imaging biomarkers for AD.

Many computerized image analysis methods have been published in the last decade to help radiologists diagnose AD more accurately [2,4,5,8,25,36]. These methods use image information extracted either from magnetic resonance (MR), positron emission tomography (PET), or single-photon emission computed tomography (SPECT), with some methods concentrating the analysis on the whole brain [4,23] while others on specific brain regions - region of interest (ROI) - such as the hippocampus [5,19].

Among the methods, the eigenbrain [3,4,22,29,30,38–40], which is a relatively simple multivariate approach based on the Principal Component Analysis (PCA) technique [17], has shown a competitive performance compared to other more complex approaches, such as atlas-based [24] and deep learning [16,20]. Zhang et al. [39], for instance, selected 2D coronal slices with inter-class variance and created an eigenbrain set with PCA. They selected the first projection weights (PWs) for each 2D slice and used them as feature inputs to an SVM classifier to aid AD diagnosis. Moreover, Kheder et al. [22] exploited the eigen-

brain construction using partial least squares (PLS) since the PLS considers the samples' variance and the class labels and compared them with the PCA. They used only GM voxel intensities and inserted the most discriminant PWs as features into an SVM to distinguish patients with MCI compared to CN and AD. In the same year, Salvatore et al. [29] has also created eigenbrain sets. The authors have assessed the method with GM and WM voxel intensities and have ordered the PWs by the Fisher Discriminant Ratio to include the class discriminative power and used them to predict the conversion from MCI to AD.

Despite its performance, one limitation of the original eigenbrain method is the linear orthogonal projections that may fail to capture the underlying structure of the data. Aiming to overcome this limitation, Alam et al. [3] have created eigenbrains with a slightly modified version of this method that uses a non-linear generalization of PCA, the kernel PCA (KPCA). The authors have created the eigenbrain set from GM voxel intensities and used the LDA for feature selection and inclusion of the class as a criterion; they compared the PCA with the Gaussian kernel PCA. In this study, we conducted a comparative assessment of the original eigenbrain (standard PCA) with four different PCA kernels [31] using the GM and the whole-brain voxel intensities. We compared the sigmoid, radial basis function (RBF), cosine, and Polynomial PCA kernels with the (linear) PCA and evaluated their predictive power using only the PWs as features, without including any class information. The methods were evaluated using image information extracted from the whole-brain image and the gray matter (GM) tissue regions.

Using feature vectors obtained by projecting each ADNI training image from the CN, MCI, and AD groups into a previously defined PCA and KPCA "eigenbrain spaces," we designed SVM classifiers with three different kernels to classify unseen MR images in the CN $\times$ MCI, MCI $\times$ AD, and CN $\times$ AD cases. The classification results of our experiments were assessed using different quantitative metrics.

The rest of this paper is structured as follows: Sect. 2 presents the image datasets used in this study. Section 3 describes the developed methods, providing details on how the feature extraction is performed. Section 4 present the results and discussions of our experiments, and, finally, Sect. 5 concludes the paper.

2 Image Datasets

This study uses two image datasets: the Alzheimer's Disease Neuroimaging Initiative[1] (ADNI) and the Neuroimage Analysis Center[2] (NAC).

The ADNI dataset includes T1-weighted (T1-w) images acquired using a wide variety of 1.5T and 3T scanners and protocols from the three leading manufacturers (Philips, General Electric, and Siemens), so the image specifications (size, resolution, etc.) are diverse and will be omitted here. For this reason, as we will discuss further, the images were standardized to a reference image previously to

[1] http://adni.loni.usc.edu/.
[2] http://projects.iw.harvard.edu/nac.

processing. A total of 774 T1-w images, 258 for each group (CN, MCI, and AD), were used in this study. According to the ADNI documentation, CN patients showed no signs of depression, mild cognitive impairment, or dementia. Also, mild-AD patients were evaluated and met established criteria for probable AD. Table 1 summarizes relevant patient information from this dataset, including sex, age, and clinical information.

Table 1. Clinical and demographic data for CN, MCI and AD patients.

Group	#Subjects	Age	Sex	MMSE score
		$(\mu \pm \sigma)$ [range]	(M/F)	$(\mu \pm \sigma)$ [range]
CN	258	74.0 ± 6.3 [58–90]	114/144	29.1 ± 1.1 [25–30]
MCI	258	74.7 ± 8.0 [54–97]	164/94	26.8 ± 1.9 [20–30]
AD	258	75.8 ± 8.1 [55–91]	140/1118	21.8 ± 4.4 [8–28]

MMSE stands for mini-mental state exam.

The NAC dataset provides a T1-w reference image with an isotropic spatial resolution of 1 mm and a matrix size of $256 \times 256 \times 256$ voxels used as a reference image for the standardization of all ADNI images.

3 Methods

In our proposed methodology, first, we preprocessed all ADNI images and extracted the GM tissue to delimited the two regions of analysis, the whole-brain and GM voxel intensities. Next, we applied the eigenbrain technique using two distinct projections, the linear PCA and the kernel PCA, with four different kernel functions. Finally, we trained SVM classifiers with three kernels using 10-fold stratified cross-validation to perform the brain MR image classification. All steps are described as follows.

3.1 Preprocessing

Several preprocessing steps were applied to the MR images to improve the image quality and allow proper analysis. First, the images were denoised using the Non-Local Means (NLM) method proposed by Buades, Coll and Morel [7]. Then, we used the N4-ITK method [34] to mitigate the bias field effect [18]. Next, using the T1-w template image from the NAC dataset (as indicated in Sect. 2) as a reference, we applied the image standardization method proposed by Nyul, Udupa, and Zhang [27] and aligned all ADNI images to the same spatial reference with affine transformations using the Nifty-Reg tool[3]. Then, we resampled all images to have an isotropic voxel size of 1mm and a matrix of dimension $256 \times 256 \times 256$ voxels. Finally, we applied a skull stripping technique [15] to remove unnecessary information from the analysis and reduce the computational time.

[3] http://cmictig.cs.ucl.ac.uk/wiki/index.php/NiftyReg.

3.2 Gray Matter Segmentation

Besides the whole-brain analysis, in this study, we also limited our analysis to the gray matter tissue region, which is the most affected tissue type in AD. For that, we segmented the three main tissues (GM, WM, and CSF) in all MR images using the fully automated brain tissue segmentation software FAST [37] (version 4.1), which is considered state of the art on tissue segmentation [28] and is available on the FSL package[4]. The FAST implementation is based on a Hidden Markov Random Field Model, optimized using the Expectation-Maximization (EM) algorithm. A detailed description of the algorithm is available in [37]. The input image required by the FAST is a skull stripped MR image, which we obtained as a result of the preprocessing step using the ROBEX [15] technique. The outputs are an intensity non-uniformity corrected version of the input image and the segmented GM, WM, and CSF brain tissues.

3.3 Eigenbrain Technique

Eigenbrain is a 3-D extension of the 2-D eigenface technique [33] applied to identify MR brain images of subjects with Alzheimer's disease [39]. The technique captures different characteristic variances of anatomical brain structures from a collection of 3-D brain images and uses them to encode and decode a brain, within a machine learning context, without considering the full image information. This process reduces computation and space complexity.

Mathematically, the original eigenbrain technique uses the PCA technique [17], which is a linear transform commonly used for dimensionality reduction, to calculate the linear brain space bases, represented by unidimensional vectors reformatted from a set of training 3D images. The base vectors correspond to the eigenvectors of the covariance matrix constructed from the training dataset and provide the directions of the greatest variances of the data, under the constraint that the base vectors are orthogonal to each other [1]. These eigenvectors, defined in the image space, look like ghost brain images after their reshaping to three-dimensional arrays and are usually referred to as eigenbrains. The first eigenbrain is the average brain, while the rest of the eigenbrains represent variations from the average brain. Figure 1 shows the formulation of the technique and illustrates the eigenbrains reshaped to the three-dimensional array using 25, 50, and 100 projection weights, respectively.

Eigenbrains
To compute the eigenbrains, we used a group of N_e MR brain images from the ADNI dataset. Then, to avoid unnecessary computation we used a brain mask from NAC to ignore background voxels containing no information. The resulting voxels were then transformed to unidimensional column vectors of size $M \times 1$, with $M = 1642084$ elements, corresponding to the number of useful brain voxels

[4] http://fsl.fmrib.ox.ac.uk/fsl/fslwiki.

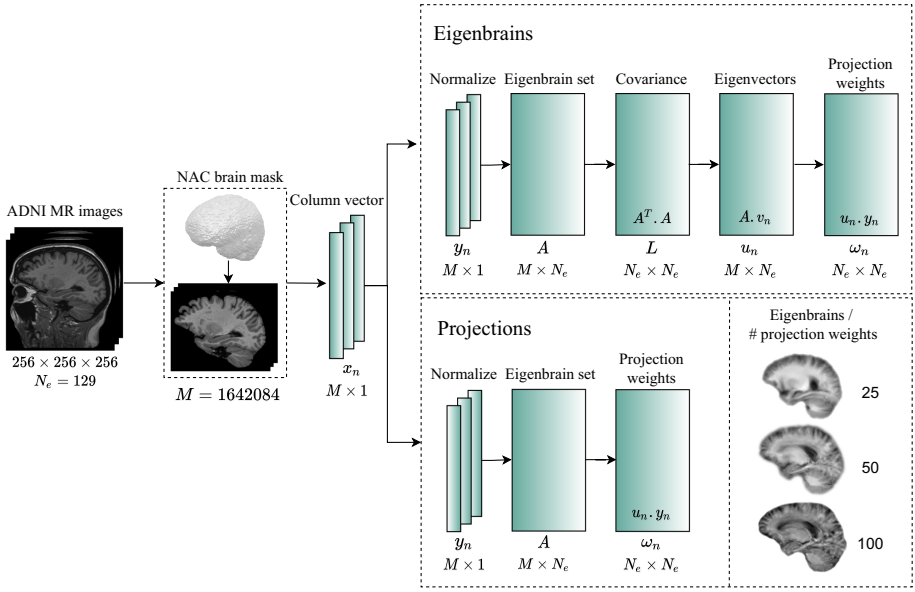

Fig. 1. Flowchart for the eigenbrain technique.

in the MR images. Next, as required by the PCA technique, we normalized the vector images by shifting them to have zero-mean, as

$$\boldsymbol{y}_n = \boldsymbol{x}_n - \bar{\boldsymbol{x}}, \tag{1}$$

where $\boldsymbol{x}_n$ is the nth vector image, and

$$\bar{\boldsymbol{x}} = \frac{1}{N_e} \sum_{n=1}^{N_e} \boldsymbol{x}_n, \tag{2}$$

is the average brain image. The normalized vector brain images, forming the eigenbrain set $\Delta = \{\boldsymbol{y}_n\}_{n=1}^{N_e}$, are organized as a matrix $\boldsymbol{A} = \left[\boldsymbol{y}_1\ \boldsymbol{y}_2\ \ldots\ \boldsymbol{y}_{N_e}\right]$ of size $M \times N_e$.

Now, if we consider finding the eigenvalues λ_n and eigenvector $\boldsymbol{u}_n$ for the covariance matrix $\boldsymbol{C}$

$$\boldsymbol{C} = \frac{1}{N_e} \sum_{n=1}^{N_e} \boldsymbol{y}_n \boldsymbol{y}_n^T = \boldsymbol{A}\boldsymbol{A}^T, \tag{3}$$

this would imply to solve

$$\boldsymbol{C}\boldsymbol{u}_n = \boldsymbol{A}\boldsymbol{A}^T\boldsymbol{u}_n = \lambda_n \boldsymbol{u}_n, \tag{4}$$

which would be a prohibitive task due to the large size ($M \times M$ elements) of matrix $\boldsymbol{C}$.

To overcome this issue, instead of solving Eq. 4 as it is, we replace the covariance matrix $\boldsymbol{C}$ by matrix

$$\boldsymbol{L} = \frac{1}{N_e} \sum_{n=1}^{N_e} \boldsymbol{y}_n^T \boldsymbol{y}_n = \boldsymbol{A}^T \boldsymbol{A}, \tag{5}$$

which has a size of $N_e \times N_e$ ($N_e \ll M$) and whose eigenvalues/vectors hold a close relationship with the ones of matrix $\boldsymbol{C}$. In fact, the eigenvectors $\boldsymbol{v}_n$ of matrix $\boldsymbol{L}$ can be determined by solving

$$\boldsymbol{L}\boldsymbol{v}_n = \boldsymbol{A}^T \boldsymbol{A}\boldsymbol{v}_n = \lambda_n \boldsymbol{v}_n, \tag{6}$$

which is equivalent to

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^T \boldsymbol{A}\boldsymbol{v}_n &= \lambda_n \boldsymbol{A}\boldsymbol{v}_n, \\ \boldsymbol{C}\boldsymbol{A}\boldsymbol{v}_n &= \lambda_n \boldsymbol{A}\boldsymbol{v}_n, \\ \boldsymbol{C}\boldsymbol{u}_n &= \lambda_n \boldsymbol{u}_n, \end{aligned} \tag{7}$$

where $\boldsymbol{u}_n = \boldsymbol{A}\boldsymbol{v}_n$. Thus, the matrices $\boldsymbol{C}$ and $\boldsymbol{L}$ have the same eigenvalues λ_n, and $\boldsymbol{u}_n = \boldsymbol{A}\boldsymbol{v}_n$ relates their eigenvectors. The eigenvalues/vectors of matrix $\boldsymbol{L} = \boldsymbol{A}^T \boldsymbol{A}$ represent the N_e largest eigenvalues/vectors of matrix $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{A}^T$, with the remaining eigenvectors being associated with eigenvalues of zero value [33].

Using the correspondence between matrices $\boldsymbol{C}$ and $\boldsymbol{L}$ as described above, all calculations are greatly reduced, from the order of the number of image voxels, $\mathcal{O}(M^3)$, to the number of images in the normalized set, $\mathcal{O}(N_e)$.

After normalizing the eigenvectors $\boldsymbol{u}_n$ to have a unit norm, i.e., $\|\boldsymbol{u}_n\| = 1$, we keep only the K eigenvectors ($K \leq N_e$) corresponding to the K largest eigenvalues to represent the images with minimal inaccuracy. By using all K bases, each brain image from our eigenbrain set Δ can be represented as a linear combination of the best K eigenvectors as

$$\hat{\boldsymbol{y}}_n = \sum_{j=1}^{K} \omega_j \boldsymbol{u}_j, \tag{8}$$

where $\omega_j = \boldsymbol{u}_j^T \boldsymbol{y}_n$ represents the PWs, and the $\boldsymbol{u}_j$ eigenvectors are referred to as the eigenbrains. Finally, with the aim of pattern classification, each normalized vector image $\boldsymbol{y}_n$ can be represented in this K bases by the vector $\boldsymbol{\Omega} = [\omega_1, \omega_2, \ldots, \omega_K]$

Kernel PCA

In eigenbrain analysis, both PCA and Kernel PCA [31,32] have a similar task, which is to extract relevant features from the brain images. However, while PCA extracts features via orthogonal projections, Kernel PCA uses a kernel function $\boldsymbol{\varphi}(\cdot)$ to implicitly map an image from the input space (image space $\boldsymbol{R}^M$) to a

higher dimensional feature space (kernel space $\boldsymbol{F}$), possibly in a non-linear way, as

$$\boldsymbol{\varphi} : \boldsymbol{R}^M \rightarrow \boldsymbol{F}, \tag{9}$$

and there it performs PCA on the images $\boldsymbol{y}_n$.

To avoid explicit calculations in the kernel space, a $N_e \times N_e$ kernel matrix $\boldsymbol{G}$, also known as Gram matrix [26], is defined, whose elements are

$$g_{i,j} = g(\boldsymbol{y}_i, \boldsymbol{y}_j) = \langle \boldsymbol{\varphi}\left(\boldsymbol{y}_i\right), \boldsymbol{\varphi}\left(\boldsymbol{y}_j\right) \rangle, \tag{10}$$

where $i, j = 1, 2, \ldots, N_e$ and $g(\cdot, \cdot)$ is a kernel function that allows computing inner products, denoted by $\langle \cdot, \cdot \rangle$, in $\boldsymbol{F}$ [31,32].

To obtain the eigenvectors in this new space, the covariance matrix is defined as

$$\boldsymbol{C}_{\boldsymbol{\varphi}} = \frac{1}{N_e} \sum_{n=1}^{N_e} \boldsymbol{\varphi}\left(\boldsymbol{y}_n\right) \boldsymbol{\varphi}\left(\boldsymbol{y}_n\right)^T = \frac{1}{N_e} \boldsymbol{A}_{\varphi} \boldsymbol{A}_{\varphi}^T, \tag{11}$$

where $\boldsymbol{A}_{\varphi} = \left[\varphi(\boldsymbol{y}_1)\ \varphi(\boldsymbol{y}_2)\ \ldots\ \varphi(\boldsymbol{y}_{N_e})\right]$.

Similarly to PCA, the eigenvalues λ_i and eigenvectors $\boldsymbol{w}_i$ of matrix $\boldsymbol{C}_{\varphi}$ are computed by solving

$$\boldsymbol{C}_{\boldsymbol{\varphi}} \boldsymbol{w}_i = \lambda_i \boldsymbol{w}_i. \tag{12}$$

By replacing $\boldsymbol{C}_{\boldsymbol{\varphi}}$ in the above equation with its definition in Eq. 11 and rearranging the terms we have

$$\begin{aligned} \boldsymbol{w}_i &= \sum_{n=1}^{N_e} \{ \frac{\boldsymbol{\varphi}\left(\boldsymbol{y}_n\right)^T \boldsymbol{w}_i}{\lambda_i N_e} \} \boldsymbol{\varphi}\left(\boldsymbol{y}_n\right) \\ &= \sum_{n=1}^{N_e} \alpha_{i,n} \boldsymbol{\varphi}\left(\boldsymbol{y}_n\right) = \boldsymbol{A}_{\varphi} \boldsymbol{\alpha}_i, \end{aligned} \tag{13}$$

where $\boldsymbol{\alpha}_i = [\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,M}]^T$. Therefore, since $\boldsymbol{\varphi}(\cdot)$ is a predetermined kernel function, to find the eigenvectors $\boldsymbol{w}_i$ is equivalent to find the coefficients $\alpha_{i,m}$.

In order to compute the $\alpha_{i,m}$, the eigenvalue problem from Eq. 12 is projected on each column of matrix $\boldsymbol{A}_{\varphi}$ as

$$\langle \boldsymbol{\varphi}_n\left(\boldsymbol{y}_n\right), \boldsymbol{C}_{\boldsymbol{\varphi}} \boldsymbol{w}_i \rangle = \lambda_i \langle \boldsymbol{\varphi}_n\left(\boldsymbol{y}_n\right), \boldsymbol{w}_i \rangle, \quad n = 1, \ldots, N_e. \tag{14}$$

Now, substituting Eq. 13 into Eq. 14 and using the definition of the Gram matrix $\boldsymbol{G}$, all $\boldsymbol{\alpha}_i$'s can be determined by solving

$$\frac{1}{N_e} \boldsymbol{G} \boldsymbol{\alpha} = \lambda \boldsymbol{\alpha}. \tag{15}$$

As the last step, the eigenvectors $\boldsymbol{w}_i$ need to be normalized in the feature space to ensure the unit norm, which is done by imposing the following condition $\langle \boldsymbol{w}_i, \boldsymbol{w}_i \rangle = 1$ and leads to

$$\begin{aligned} 1 &= \sum_{i,j=1}^{N_e} \alpha_i \alpha_j \langle \boldsymbol{\varphi}_i \left(\boldsymbol{y}_i\right), \boldsymbol{\varphi}_j \left(\boldsymbol{y}_j\right) \rangle, \\ &= \sum_{i,j=1}^{N_e} \alpha_i \alpha_j \boldsymbol{G}_{ij}, \\ &= \lambda \langle \boldsymbol{\alpha}, \boldsymbol{G}\boldsymbol{\alpha} \rangle, \end{aligned} \tag{16}$$

where the sum is taken over all values of i, j. The normalization implies that each α_i be divided by the square root of the corresponding eigenvalue ($\sqrt{\lambda_i}$).

Now, assuming that $\boldsymbol{y}$ and $\boldsymbol{\varphi}(\boldsymbol{y})$ are, respectively, a normalized vector brain image and its mapping in the feature space, then the projection of $\boldsymbol{\varphi}(\boldsymbol{y})$ on the eigenvectors $\boldsymbol{w}_i$ is computed as

$$\langle \boldsymbol{w}_i, \boldsymbol{\varphi}(\boldsymbol{y}) \rangle = \sum_{n=1}^{N_e} \alpha_n \langle \boldsymbol{\varphi}(\boldsymbol{y_n}), \boldsymbol{\varphi}(\boldsymbol{y}) \rangle = \sum_{n=1}^{N_e} \alpha_n g(\boldsymbol{y_n}, \boldsymbol{y}), \tag{17}$$

and provides non-linear principal components.

Therefore, we can extract the K non-linear principal components by working implicitly in the feature space.

3.4 Experiments Setup

In this study, we conducted six different classification experiments, i.e., CN × MCI, MCI × AD, and CN × AD, for features extracted from the whole-brain (whole-brain analysis - WBA) and for the GM tissue only (gray-matter analysis - GMA). We randomly split the entire ADNI dataset into two sets of equal sizes, with the number of class images for each experiment equal to 129 CN, 129 MCI, and 129 DA subjects for each split. Images from the first split were used to create the eigenspace. Images from the second split were projected into the eigenspace to produce the feature vectors for the training, validation, and test of the SVM classifiers. We applied this strategy to all experiments, i.e., whole-brain and GM analysis using both the PCA and KPCA techniques.

3.5 Pattern Classification Design

In this study, SVM classifiers were designed using the projection weights obtained from either the PCA or KPCA techniques using a training dataset as feature vectors. The SVM kernels used in the analyses were the Polynomial (degree 1), RBF, and the sigmoid kernels with $\gamma = \frac{1}{n_{PWs} * var(\omega_{ij})}$. The regularization parameter C, which controls the influence of the misclassification on the objective

function in SVM, was defined for each kernel using a 10-fold stratified cross-validation evaluation scheme [14], with each training set varying C penalty using the grid search algorithm from 0,01 to 100. The balanced accuracy metric [12] was used as the figure of merit to be maximized. We used this approach by increasing the number of PWs, i.e., we applied to the first PC, then with first and second, so forth.

The metrics used to assess our classification performance were sensitivity, specificity, balanced accuracy, which correspond to the sensitivity and specificity average, and the area under the roc curve (AUC). A detailed description of these metrics can be found in [21].

4 Results and Discussion

As indicated in Sect. 3.4, our experiments were organized into two groups: (a) whole-brain analysis and (b) gray-matter analysis (GMA). Tables 2a to 3c are organized to show the obtained results using the combination of the feature extraction technique (PCA or KPCA) and the SVM classifier (kernels RBF or sigmoid or Polynomial). Also, for each combination of techniques, the number of projection weights retained for the analysis is indicated.

4.1 Results of the Whole-Brain Analysis

Table 2a shows the results for the CN × MCI experiments with an indication in bold for the best result obtained for each combination of feature extraction technique and SVM-kernel classifier. As can be noticed, the kernel PCA technique achieved slightly better results (AUC values of 0.74 ± 0.12 for the Polynomial-Sigmoid and Cosine-Polynomial combinations) compared to the traditional PCA (AUC value of 0.73 ± 0.11 for using the SVM-sigmoid classifier). The highest accuracy value (73.17 ± 9.12 %) for this experiment was obtained using the Cosine-RBF combination. However, it should be noted that the cross-validation standard deviation values for all metrics were relatively high. The best numbers of projection weights for most of the experiments were considerately small, especially for the two ones with the highest AUC values (13 out of 256).

The results for the MCI × AD experiments are shown in Table 2b. Similar to the CN × MCI experiments, the AUC values for the PCA and kernel PCA techniques were close, with a slightly higher average value for the kernel-Cosine/SVM-Polynomial case (AUC value of 0.72 ± 0.07). The numbers of projection weights were significantly higher than the CN × MCI experiments, which can be explained by the difficulty in differentiating the MCI and AD brains. The standard deviation values for all metrics were also considerably high.

Table 2c presents the results for the CN × AD experiments. Following the same trend seen in the two previous cases (CN × MCI and MCI × AD), it can be noticed that the results using the kernel PCA technique were slightly better than the PCA. Also, the numbers of projection weights were relatively higher when compared to the CN × MCI experiments.

As expected, the results have shown higher AUC values for the CN × AD experiments when compared to the CN × MCI and MCI × AD, which can be explained by the higher structural brain differences between these two extreme groups. That is corroborated by noticing that the CN × MCI case results are slightly better than the MCI × AD, for which the brains may show similar structural changes.

4.2 Results of the Gray-Matter Analysis

Considering the dynamically spreading wave of gray matter loss in the brains of patients with AD, which results in brain atrophy as a consequence, in this study, we investigate the PCA and kernel PCA techniques applied only to the gray-matter region of the brain. For that, as mentioned in Sect. 3.2, the three main brain tissues were segmented, and the GM was isolated for further analysis. Table 3a shows the results for the CN × MCI experiments. Different from the experiments using the whole-brain analysis, in this case, we can notice a significant average difference between the PCA and kernel PCA techniques, i.e., AUC value of 0.63 ± 0.08 for the PCA/SVM-polynomial classifier versus 0.76 ± 0.13 for the polynomial/SVM-polynomial classifier. The standard deviation values for these experiments have also shown to be high. Also, with a few exceptions, the numbers of projection weights were relatively small (35 out of 256, for the higher AUC value).

For the MCI × AD experiments (Table 3b), the AUC average values were marginally better than the CN × MCI case for the traditional PCA technique, with the highest AUC value of 0.67 ± 0.11, and slightly worse for the kernel PCA techniques, with the highest AUC value of 0.74 ± 0.11, achieved using the kernel PCA RBF and the SVM-polynomial classifier. The standard deviation values were similar or even smaller than all other experiments, while the number of projection weights for the best result was higher (145 out of 256).

Finally, Table 3c shows the results for the CN × AD experiments. In this case, we can see a significant improvement in the results, with the highest average AUC value of 0.89 achieved by using two different PCA kernels (RBF and Polynomial). The numbers of projection weights were relatively small (43 and 30) in these two cases. The best average AUC value obtained by the traditional PCA was 0.86 using the SVM-RBF classifier, which is higher than any other experiment.

Table 2. Whole-brain analysis

	Feature extraction	kernel SVM	#PWs	Bal.Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
(a) CN × MCI							
kernel PCA	PCA	RBF	20	69.32 ± 9.57	71.41 ± 17.12	67.24 ± 17.49	0.71 ± 0.13
		Sigmoid	**8**	**70.42 ± 8.80**	**64.42 ± 18.47**	**76.41 ± 14.50**	**0.73 ± 0.11**
		Polynomial	10	70.38 ± 10.85	68.14 ± 18.17	72.63 ± 13.50	0.71 ± 0.13
	RBF	**RBF**	**12**	**71.54 ± 5.78**	**71.28 ± 17.94**	**71.79 ± 10.63**	**0.73 ± 0.07**
		Sigmoid	11	68.56 ± 6.79	65.19 ± 17.84	71.92 ± 13.38	0.71 ± 0.10
		Polynomial	18	71.22 ± 6.99	71.99 ± 18.16	70.45 ± 16.74	0.73 ± 0.11
	Sigmoid	RBF	20	65.29 ± 8.09	62.63 ± 15.48	67.95 ± 13.09	0.66 ± 0.11
		Sigmoid	42	68.11 ± 8.17	69.92 ± 15.23	66.60 ± 13.20	0.72 ± 0.11
		Polynomial	**50**	**71.19 ± 7.44**	**69.55 ± 15.92**	**72.82 ± 10.85**	**0.72 ± 0.10**
	Polynomial	RBF	20	68.56 ± 10.29	70.64 ± 18.69	66.47 ± 18.86	0.71 ± 0.13
		Sigmoid	**13**	**72.85 ± 7.71**	**72.12 ± 15.41**	**73.59 ± 15.74**	**0.74 ± 0.12**
		Polynomial	10	70.38 ± 10.85	68.14 ± 18.17	72.63 ± 13.50	0.73 ± 0.11
	Cosine	RBF	21	73.17 ± 9.12	75.26 ± 19.94	71.09 ± 13.78	0.74 ± 0.13
		Sigmoid	69	69.07 ± 12.22	70.58 ± 14.85	67.56 ± 17.69	0.71 ± 0.13
		Polynomial	**13**	**71.99 ± 8.53**	**70.45 ± 19.27**	**73.53 ± 16.48**	**0.74 ± 0.12**
(b) MCI × AD							
kernel PCA	PCA	RBF	98	64.17 ± 9.76	61.54 ± 18.76	66.79 ± 20.83	0.63 ± 0.09
		Sigmoid	77	65.61 ± 7.07	64.62 ± 14.65	66.60 ± 13.94	0.69 ± 0.08
		Polynomial	**31**	**65.42 ± 8.29**	**65.04 ± 15.00**	**65.19 ± 17.51**	**0.71 ± 0.07**
	RBF	RBF	51	63.53 ± 8.22	63.40 ± 14.38	63.65 ± 12.24	0.67 ± 0.11
		Sigmoid	**19**	**67.05 ± 9.59**	**61.15 ± 16.72**	**72.95 ± 14.19**	**0.71 ± 0.07**
		Polynomial	10	63.94 ± 10.51	60.38 ± 19.04	67.50 ± 11.68	0.64 ± 0.11
	Sigmoid	**RBF**	**53**	**64.10 ± 10.06**	**62.24 ± 18.29**	**65.96 ± 13.27**	**0.66 ± 0.07**
		Sigmoid	68	57.63 ± 8.56	57.88 ± 13.62	57.37 ± 15.04	0.57 ± 0.09
		Polynomial	68	62.40 ± 7.29	68.01 ± 14.27	56.79 ± 19.03	0.58 ± 0.09
	Polynomial	RBF	31	65.53 ± 4.25	69.55 ± 16.62	57.50 ± 16.73	0.65 ± 0.07
		Sigmoid	125	66.12 ± 6.56	67.82 ± 13.03	64.42 ± 11.32	0.67 ± 0.09
		Polynomial	**120**	**63.97 ± 7.07**	**68.14 ± 15.84**	**59.81 ± 20.43**	**0.68 ± 0.08**
	Cosine	RBF	34	67.34 ± 10.15	67.18 ± 14.97	67.50 ± 17.72	0.70 ± 0.13
		Sigmoid	28	64.58 ± 10.31	64.87 ± 15.00	64.29 ± 16.66	0.68 ± 0.12
		Poly	**85**	**64.13 ± 6.14**	**63.27 ± 7.74**	**65.00 ± 16.18**	**0.72 ± 0.07**
(c) CN ×AD							
kernel PCA	PCA	RBF	89	72.33 ± 8.04	76.47 ± 11.27	68.20 ± 13.95	0.734 ± 0.143
		Sigmoid	**55**	**75.06 ± 5.55**	**74.23 ± 13.60**	**75.89 ± 11.30**	**0.801 ± 0.092**
		Polynomial	92	73.87 ± 9.36	72.56 ± 12.57	75.19 ± 18.13	0.747 ± 0.117
	RBF	**RBF**	**62**	**72.76 ± 4.31**	**73.52 ± 7.33**	**71.41 ± 9.58**	**0.768 ± 0.054**
		Sigmoid	67	72.17 ± 11.44	71.85 ± 11.37	72.50 ± 17.23	0.759 ± 0.067
		Polynomial	28	72.91 ± 8.46	73.58 ± 16, 91	72.24 ± 6.96	0.758 ± 0.077
	Sigmoid	**RBF**	**43**	**71.25 ± 8.34**	**68.58 ± 14.80**	**73.91 ± 13.47**	**0.733 ± 0.111**
		Sigmoid	12	69.13 ± 6.05	66.08 ± 9.51	72.17 ± 11.64	0.652 ± 0.109
		Polynomial	38	71.31 ± 10.73	66.08 ± 10.22	76.53 ± 17.66	0.723 ± 0.104
	Polynomial	RBF	59	75.83 ± 6.47	76.02 ± 9.46	75.64 ± 12.60	0.798 ± 0.053
		Sigmoid	162	75.80 ± 7.79	74.29 ± 10.94	77.30 ± 13.76	0.758 ± 0.092
		Polynomial	**56**	**75.44 ± 6.12**	**73.65 ± 11.88**	**77.24 ± 10.98**	**0.804 ± 0.042**
	Cosine	**RBF**	**27**	**74.93 ± 6.43**	**74.42 ± 10.68**	**75.44 ± 12.61**	**0.793 ± 0.075**
		Sigmoid	28	75.44 ± 6.73	76.92 ± 7.88	73.97 ± 11.14	0.790 ± 0.093
		Polynomial	58	75.06 ± 8.2	72.05 ± 10.60	78.07 ± 11.09	0.763 ± 0.058

Table 3. Gray-matter analysis

	Feature extraction	kernel SVM	#PWs	Bal.Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
(a) CN × MCI							
kernel PCA	PCA	RBF	25	63.69 ± 10.49	64.23 ± 10.33	63.14 ± 13.47	0.63 ± 0.13
		Sigmoid	77	62.82 ± 10.50	68.21 ± 10.58	57.44 ± 20.67	0.60 ± 0.11
		Polynomial	**30**	**63.33 ± 8.58**	**68.08 ± 11.67**	**58.59 ± 10.94**	**0.63 ± 0.08**
	RBF	RBF	33	67.66 ± 10.83	69.74 ± 19.92	65.58 ± 15.16	0.69 ± 0.13
		Sigmoid	11	68.30 ± 10.35	72.05 ± 8.69	64.55 ± 16.40	0.71 ± 0.15
		Polynomial	**36**	**71.60 ± 10.84**	**71.99 ± 14.94**	**71.22 ± 22.72**	**0.74 ± 0.11**
	Sigmoid	RBF	118	66.92 ± 11.47	73.59 ± 13.95	60.26 ± 22.16	0.66 ± 0.10
		Sigmoid	32	70.03 ± 9.97	70.58 ± 11.73	69.49 ± 13.23	0.65 ± 0.14
		Polynomial	**12**	**68.69 ± 11.15**	**71.28 ± 10.99**	**66.09 ± 13.19**	**0.72 ± 0.16**
	Polynomial	RBF	31	69.62 ± 8.69	69.62 ± 16.36	69.62 ± 15.01	0.71 ± 0.14
		Sigmoid	34	70.80 ± 9.03	71.28 ± 17.27	70.32 ± 20.66	0.73 ± 0.12
		Polynomial	**35**	**75.10 ± 8.82**	**77.44 ± 14.13**	**72.76 ± 16.23**	**0.76 ± 0.13**
	Cosine	RBF	34	68.43 ± 11.73	73.59 ± 13.95	63.27 ± 23.93	0.67 ± 0.16
		Sigmoid	4	69.13 ± 8.65	65.06 ± 9.44	73.21 ± 12.25	0.71 ± 0.13
		Polynomial	**37**	**70.71 ± 11.62**	**75.00 ± 16.96**	**66.41 ± 24.48**	**0.73 ± 0.13**
(b) MCI × AD							
kernel PCA	PCA	RBF	17	67.02 ± 10.94	64.23 ± 17.41	69.81 ± 13.87	0.66 ± 0.15
		Sigmoid	**18**	**66.67 ± 11.30**	**63.46 ± 17.96**	**69.87 ± 15.05**	**0.67 ± 0.11**
		Polynomial	15	67.34 ± 9.19	64.87 ± 20.11	69.81 ± 11.54	0.65 ± 0.10
	RBF	RBF	72	65.87 ± 9.11	61.79 ± 14.97	69.94 ± 18.59	0.68 ± 0.10
		Sigmoid	146	64.97 ± 4.74	60.96 ± 12.77	68.97 ± 9.76	0.59 ± 0.04
		Polynomial	**145**	**66.57 ± 5.42**	**65.64 ± 14.60**	**67.50 ± 13.11**	**0.74 ± 0.11**
	Sigmoid	**RBF**	**83**	**68.59 ± 8.11**	**65.71 ± 11.81**	**71.47 ± 15.70**	**0.72 ± 0.08**
		Sigmoid	88	65.42 ± 9.51	65.58 ± 9.38	65.26 ± 18.79	0.65 ± 0.13
		Polynomial	89	65.35 ± 8.89	66.28 ± 11.29	64.42 ± 17.48	0.67 ± 0.07
	Polynomial	RBF	74	63.43 ± 10.91	59.36 ± 16.66	67.50 ± 13.99	0.64 ± 0.11
		Sigmoid	64	64.97 ± 7.75	63.97 ± 11.84	65.96 ± 14.55	0.61 ± 0.11
		Polynomial	**55**	**63.85 ± 6.48**	**63.27 ± 10.91**	**64.42 ± 11.83**	**0.65 ± 0.06**
	Cosine	**RBF**	**141**	**64.90 ± 7.46**	**60.583 ± 17.43**	**68.97 ± 15.78**	**0.68 ± 0.08**
		Sigmoid	136	63.40 ± 6.74	59.36 ± 16.83	67.44 ± 18.13	0.64 ± 0.09
		Polynomial	138	65.29 ± 8.03	60.77 ± 20.46	69.81 ± 15.86	0.63 ± 0.07
(c) CN ×AD							
kernel PCA	PCA	**RBF**	**34**	**80.74 ± 9.25**	**81.22 ± 11.07**	**80.26 ± 19.80**	**0.86 ± 0.13**
		Sigmoid	78	77.24 ± 10.34	70.19 ± 16.07	84.29 ± 11.58	0.80 ± 0.12
		Polynomial	69	77.56 ± 11.17	69.42 ± 14.35	85.71 ± 17.20	0.81 ± 0.12
	RBF	**RBF**	**43**	**82.72 ± 8.33**	**84.42 ± 11.03**	**81.03 ± 18.88**	**0.89 ± 0.07**
		Sigmoid	139	83.46 ± 10.41	83.65 ± 10.15	83.27 ± 22.53	0.85 ± 0.14
		Polynomial	137	82.69 ± 9.85	82.12 ± 9.88	83.27 ± 22.53	0.86 ± 0.10
	Sigmoid	**RBF**	**146**	**82.31 ± 11.11**	**88.33 ± 7.17**	**76.28 ± 23.62**	**0.88 ± 0.13**
		Sigmoid	149	81.15 ± 12.64	83.72 ± 12.61	78.59 ± 24.67	0.87 ± 0.10
		Polynomial	148	81.89 ± 13.11	84.42 ± 12.05	79.36 ± 24.97	0.88 ± 0.10
	Polynomial	**RBF**	**30**	**83.11 ± 9.37**	**84.36 ± 7.84**	**81.86 ± 17.52**	**0.89 ± 0.11**
		Sigmoid	146	78.40 ± 10.95	73.53 ± 12.99	83.27 ± 22.79	0.74 ± 0.12
		Poly	144	78.43 ± 9.31	71.99 ± 11.50	84.87 ± 20.56	0.82 ± 0.13
	Cosine	**RBF**	**39**	**82.82 ± 6.09**	**85.26 ± 8.05**	**80.38 ± 15.04**	**0.88 ± 0.09**
		Sigmoid	134	78.46 ± 8.86	78.97 ± 10.89	77.95 ± 19.60	0.84 ± 0.07
		Polynomial	103	80.03 ± 9.99	84.42 ± 10.28	75.64 ± 17.51	0.87 ± 0.07

5 Conclusions

The eigenbrain method was initially proposed as a simple extension of the eigenfaces for the Alzheimer's disease classification problem, and, consequently, both techniques work by projecting the original data (3-D MR images) into PCA orthogonal bases.

In this study, we investigated the application of linear (PCA) and non-linear (kernel PCA) feature projections for the classification of 3-D MR images on CN, MCI, and AD, with the goal of contrast the results obtained from these two types of projections. In addition, our analyzes were performed using information extracted from the whole-brain MR images and from the gray-matter tissues, segmented from the T1-w MR images using the FAST segmentation technique.

With the use of the KPCA, compared with the standard PCA, we included information about the non-linearity of the features. These non-linear features were capable of outperforming the standard PCA in terms of classification accuracy when predicting CN $\times$ MCI, MCI $\times$ AD, and CN $\times$ AD, with a more advantage difference from the experiments performed for the GM analysis. In this case, the highest average AUC values (0.89 ± 0.07 and 0.89 ± 0.11) were obtained for the CN $\times$ AD experiment using two different combinations of the PCA kernels and SVM-RBF classifiers.

Acknowledgment. Funding for ADNI can be found at http://adni.loni.usc.edu/about/#fund-container.

Funding Statement. This study was financed by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant numbers 2018/08826-9 and 2018/06049-5) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001.

References

1. Abdi, H., Williams, L.J.: Principal Component Analysis. Wiley Interdisc. Rev. Comput. Stat. **2**(4), 433–459 (2010)
2. Ahmed, M.R., Zhang, Y., Feng, Z., Lo, B., Inan, O.T., Liao, H.: Neuroimaging and machine learning for dementia diagnosis: recent advancements and future prospects. IEEE Rev. Biomed. Eng. **12**, 19–33 (2018)
3. Alam, S., Kwon, G., Initiative, A.D.N.: Alzheimer disease classification using kpca, lda, and multi-kernel learning svm. Int. J. Imaging Syst. Technol. **27**(2), 133–143 (2017)
4. Álvarez, I., et al.: Alzheimer's diagnosis using eigenbrains and support vector machines. Electron. Lett. **45**(7), 342–343 (2009)
5. Amoroso, N., et al.: Alzheimer's disease diagnosis based on the hippocampal unified multi-atlas network (human) algorithm. Biomed. Eng. Online **17**(1), 1–16 (2018)
6. Bassiony, H.S., Zickri, M.B., Metwally, H.G., Elsherif, H.A., Alghandour, S.M., Sakr, W.: Comparative histological study on the therapeutic effect of green tea and stem cells in Alzheimer's disease complicating experimentally induced diabetes. Int. J. Stem Cells **8**(2), 181–190 (2015)
7. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. Multiscale Model. Simul. **4**(2), 490–530 (2005)
8. Cao, P., et al.: Nonlinearity-aware based dimensionality reduction and over-sampling for AD/MCI classification from MRI measures. Comput. Biol. Med. **91**, 21–37 (2017)

9. Deture, M., Dickson, D.: The neuropathological diagnosis of Alzheimer's disease. Mol. Neurodegeneration **14**(32), 1–18 (2019)
10. Devanand, D.P., Bansal, R., Liu, J., Hao, X., Pradhaban, G., Peterson, B.S.: MRI hippocampal and entorhinal cortex mapping in predicting conversion to Alzheimer's disease. Neuroimage **60**(3), 1622–1629 (2012)
11. Elahi, F.M., Miller, B.L.: A clinicopathological approach to the diagnosis of dementia. Nat. Rev. Neurol. **13**(8), 457–476 (2017)
12. Fawcett, T.: An introduction of ROC analysis. Pattern Recogn. Lett. **27**(8), 861–874 (2006)
13. Frisoni, G.B., Fox, N.C., Jack-Jr, C.R., Scheltens, P., Thompson, P.M.: The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. **6**(2), 67–77 (2010)
14. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification. Technical Report, Department of Computer Science, National Taiwan University, Taiwan (May 2016). https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
15. Iglesias, J.E., Liu, C.Y., Thompson, P.M., Tu, Z.: Robust brain extraction across datasets and comparison with publicly available methods. IEEE Trans. Med. Imaging **30**(9), 1617–1634 (2011)
16. Jo, T., Nho, K., Saykin, A.: Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. Front. Aging Neurosci. **11**(220), 1–14 (2019)
17. Jolliffe, I.T., Cadima, J.: Principal component analysis: a review and recent developments. Phil. Trans. Royal Soc. Math. Phys. Eng. Sci. **374**(2065), 1–16 (2016)
18. Juntu, J., Sijbers, J., Van Dyck, D., Gielen, J.: Bias Field Correction for MRI Images. In: Kurzy?ski, M., Puchała, E., Woźniak, M., żołnierek, A. (eds) Computer Recognition Systems. Advances in Soft Computing, vol. 30, pp. 543-551. Springer, Berlin (2005) https://doi.org/10.1007/3-540-32390-2_64
19. Hett, K., Ta, V.-T., Manjón, J.V., Coupé, P.: Graph of hippocampal subfields grading for Alzheimer's disease prediction. In: Shi, Y., Suk, H.-I., Liu, M. (eds.) MLMI 2018. LNCS, vol. 11046, pp. 259–266. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00919-9_30
20. Kanghan, O., Young-Chul, C., Ko, W., Woo-Sung, K., Il-Seok, O.: Classifcation and visualization of Alzheimer's disease using volumetric convolutional neural network and transfer learning. Sci. Reports **9**(18150), 1–16 (2019)
21. Kelleher, J.D., Mac Namee, B., D'arcy, A.: Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. 1 edn. MIT Press, Cambridge (2015)
22. Khedher, L., Ramirez, J., Gorriz, J.M., Brahim, A., Segovia, F., Initiative, A.D.N., et al.: Early diagnosis of alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented mri images. Neurocomputing **151**, 139–150 (2015)
23. Liu, J., Li, M., Lan, W., Wu, F., Pan, Y., Wang, J.: Classification of alzheimer's disease using whole brain hierarchical network. IEEE Trans. Comput. Biol. Bioinf. **15**(2), 624–632 (2018)
24. Liu, M., Zhang, D., Shen, D.: View-centralized multi-atlas classification for Alzheimer's disease diagnosis. Hum. Brain Mapp. **36**(5), 1847–1865 (2015)
25. Liu, M., Zhang, J., Nie, D., Yap, P.T., Shen, D.: Anatomical landmark based deep feature representation for MR images in brain disease diagnosis. IEEE J. Biomed. Health Inf. **22**(5), 1476–1485 (2018)
26. Nell, C., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge (2000)

27. Nyúl, L.G., Udupa, J.K., Zhang, X.: New variants of a method of MRI scale standardization. IEEE Trans. Med. Imaging **19**(2), 143–150 (2000)
28. Rehman, H.Z.U., Hwang, H., Lee, S.: Conventional and deep learning methods for skull stripping in brain MRI. Appl. Sci. **10**(5), 1773 (2020)
29. Salvatore, C., Cerasa, A., Battista, P., Gilardi, M.C., Quattrone, A., Castiglioni, I.: Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. Front. Neurosci. **9**, 307 (2015)
30. Sarwinda, D., Arymurthy, A.M.: Feature selection using kernel PCA for alzheimer's disease detection with 3D MR images of brain. In: 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp. 329–333. IEEE, Bali, IDN (September 2013)
31. Schölkopf, B.: Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput. **10**(5), 1299–1319 (1998)
32. Schölkopf, B., Smola, A., Müller, K.-R.: Kernel principal component analysis. In: Gerstner, W., Germond, A., Hasler, M., Nicoud, J.-D. (eds.) ICANN 1997. LNCS, vol. 1327, pp. 583–588. Springer, Heidelberg (1997). https://doi.org/10.1007/BFb0020217
33. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
34. Tustison, N.J., et al.: N4itk: improved N3 bias correction. IEEE Trans. Med. Imaging **29**(6), 1310–1320 (2010)
35. World Health Organization: Dementia: A public health priority. Technical Report, World Health Organization, Geneva, Switzerland (2019)
36. Zhang, J., Liu, M., An, L., Gao, Y., Shen, D.: Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images. IEEE J. Biomed. Health Inf. **21**(5), 1607–1616 (2017)
37. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging **20**(1), 45–57 (2001)
38. Zhang, Y.D., Wang, S., Dong, Z.: Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. Prog. Electromagnet. Res. **144**, 171–184 (2014)
39. Zhang, Y., et al.: Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Front. Comput. Neurosci. **9**, 66 (2015)
40. Zhang, Y., Wang, S., Phillips, P., Yang, J., Yuan, T.F.: Three-dimensional eigenbrain for the detection of subjects and brain regions related with Alzheimer's disease. J. Alzheimer's Dis. **50**(4), 1163–1179 (2016)