Brief communication (Original)

Discrimination of Alzheimer's disease using hippocampus texture features from MRI

Jayapathy Rajeesh, Rama Swamy Moni, Suyambumuthu Palanikumar, Thankappan Gopalakrishnan
Noorul Islam College of Engineering, Kumaracoil 629180, India

Background: Alzheimer's disease is the commonest cause of dementia and is fatal. Early detection of Alzheimer's disease is important because treatment may be most useful if introduced early.
Objective: Analyse the utility of using the texture features of hippocampus MRI as the biomarker to identify Alzheimer's disease.
Materials and methods: We chose the MRI of 146 normal controls and 133 Alzheimer's disease patients from the Alzheimer's disease neuroimaging initiative website. Sixty-nine texture features were extracted from the hippocampus region of interest as obtained from an automatic segmentation procedure. The best features were selected based on principal component analysis to decide which ones resulted in the highest rate of good classification. The proposed method is validated using a Support Vector Machine classifier.
Results: The accuracy of the proposed method on discrimination of Alzheimer's disease patients is 93.6%.
Conclusion: The texture features taken only from hippocampus gives better discrimination between Alzheimer's disease and normal controls. Therefore, the textures of hippocampus are much affected by Alzheimer's disease.

Keyword: Alzheimer's disease, hippocampus, principal component analysis, support vector machine, texture features

Alzheimer's disease (AD) is a brain disorder. Alzheimer's destroys brain cells, causing problems with memory, thinking, and behaviour severe enough to affect work, lifelong hobbies, or social life. Alzheimer's gets worse over time, and is fatal. Today, it is the seventh-leading cause of death in the United States. There is an accelerating effort to find better ways to treat the disease, delay its onset, or prevent it from developing. It is estimated that there are currently about 18 million people worldwide with AD. This figure is projected to nearly double by 2025 to 34 million [1]. Recent research in India suggests that the risk of AD was possibly higher for urban as compared to rural areas. Accurate diagnosis of AD can be challenging, in particular at the earlier stage. Early diagnosis of AD patients is important because it allows early treatment with cholinesterase inhibitors, which have been shown to delay institutionalization and improve or stabilize cognition and behavioural

symptoms [2, 3]. As therapeutic interventions become available, there is a need for developing methodologies that will serve as an in vivo surrogate for these pathologic changes, and thus, accurately identify those cognitively impaired individuals who are in the earliest stages of AD.

In the literature, several approaches for the development of biomarker to identify AD are presented [4-9]. D.P. Devanand et al. proposed a method to predict the conversion of Mild Cognitive Impairment (MCI) to AD from MRI hippocampal and entorhinal cortex volume. Emilie Gerardin et al. developed an automated system that discriminate patients with AD or MCI from normal. It is based on hippocampal shape feature. To model the hippocampi shape, they use spherical harmonics coefficients. These coefficients are classified by Support Vector Machines (SVM). J.L. Whitwell et al. made a voxel based morphometry study to identify if the patterns of gray matter atrophy MRI correlates with neurofibrillary tangles (NFTs). Those are composed of hyperphosphorylated tau proteins, and are one of the hallmarks of AD. Rachael I. Scahill et al. used

Correspondence to: Jayapathy Rajeesh, Noorul Islam College of Engineering, Kumaracoil 629180, India. E-mail: rajeesh_j@ yahoo.co.in

fluid registration to localize longitudinal changes in individuals, and subsequently use Statistical Parametric Mapping (SPM) to determine consistent changes within the group. Rahul S. Desikan et al. studies if automated segmentation techniques are used as a biomarker to identify AD and the regions most affected by this disease. They concluded that automated MRI measures of entorhinal cortex thickness, hippocampal volume, and supramarginal gyrus thickness identify MCI and AD with excellent discrimination accuracy and specificity. In all cases, the volume of brain regions should be measured and it requires many MRI slices. It is a time consuming process. Peter A. Freeborough et al. applied MRI texture analysis to diagnose AD. They estimated the texture features derived by spatial gray level dependence method on the complete brain image and were able to obtain a classification rate of 91%. To improve the classification rate further, we choose only hippocampus area to determine the texture features.

Methods

Two hundred seventy nine datasets were examined in this study. All datasets were selected from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). The ADNI is a large multi-site collaborative effort launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and non-profit organizations as a public–private partnership. One objective of the datasets is to test if serial MRI, PET, other biological markers and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The Principal Investigator of this initiative is Michael Weiner, MD. ADNI is the result of many co-investigations from a broad range of academic institutions and private corporations, with subjects recruited from over 50 sites across the United States and Canada. For more information, please see http://www.adni-info.org. The datasets are divided into two groups. One group is for training and it has 88 datasets belongs to AD and another 97 datasets from Normal. Another group is for validation and it has 45 datasets belongs to AD, and another 40 from Normal. The datasets belongs to AD has an average age of 74 years and male female ratio of 1:1.6. The dataset

belongs to Normal has an average age of 75 years and male female ratio of 1:1.7.

MRI Image acquisition

The MRI scans were acquired at multiple sites using either a GE or Siemens or Philips 3T system. High-resolution T1-weighted volumetric MP-RAGE scans were collected for each subject and the MINC format images were downloaded from the public ADNI site (http://www.loni.ucla.edu/ADNI/Data/index.shtml). Parameter values varied depending on scanning site. They can be found at http://www.loni.ucla.edu/ADNI/Research/Cores/.

Hippocampus Segmentation

To segment the hippocampus for the extraction of the texture features, an automatic segmentation scheme developed by J.Rajeesh et al. [10] were used. The first step of the algorithm automatically locates the LVs based on a previously well-defined procedure [11], and produces an initial hippocampus candidate set of voxels from these ventricle positions. To reduce the effect of variability of the tissues and to find the exact location of hippocampus, they initially located the approximate area where hippocampus may be present by an anatomical knowledge. Then they choose the initial bounding box based on the approximate size of hippocampus. Crop the area as ROI, which is occupied by the bounding box. Now the ROI histogram is calculated, which gives the multimodal histogram where the intensity peaks and ranges correspond to (White Matter) WM, (Gray Matter) GM, (Cerebra Spinal Fluids) CSF, and other classes. This histogram has substantial local variations that hamper determination of the global peaks and valleys. Therefore, the histogram should be smoothed [12] with the help of the Butterworth digital filter to get five peaks, which corresponds to CSF, CSF-GM interface, GM, GM-WM interface, and WM. While on smoothing, peak is ignored if its area is smaller than 1/30 of the cumulative area of the histogram, or if they are outside the intensity range [5%; 95%]. Then the smoothed histogram is modelled as a sum of five normal (Gaussian) distribution functions, one for each domain of CSF, CSF-GM, GM, GM-WM, and WM using (1).

$$F(x) = \sum_{i=1}^{5} h(i) \times \exp\left(\frac{-(x-\mu_i)^2}{2\sigma_i^2}\right) \qquad (1)$$

Where F(χ) Gaussian model of the histogram, χ is intensity value, μ mean, σ standard deviation, and is $h(i)$ the height of ith peak. Radiologically, on T1-weighted-SPGR MR images, the peak with the highest gray level corresponds to WM and the lowest to CSF. Measure the area of all peaks, if the third peak is larger, then find the seed point coordinate. If the third peak is not larger, move the bounding box to another area and follow the above procedure until the third peak becomes larger. Here third peak corresponds to the gray (hippocampus) region and the initial contour is taken from this point. The next step is to segment the hippocampus using level set algorithm. The level-set method was introduced [13] to track the evolution of interfaces by embedding a propagating interface as the zero level set of a higher dimensional level-set function and was found to overcome many limitations of the active contour methods (ACMs).

Features Extraction

Texture can be expressed in many ways, both in the spatial domain and in the frequency domain. Statistical texture analysis is based on statistical properties of the intensity histogram without considering spatial dependence. One possible approach to quantitatively describe a histogram is to look at the central moments. The general form of the central moments is

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \qquad (2)$$

where n is the moment order, L is the number of possible intensity values, z_i is a discrete random variable that denotes intensity levels in an image, $p(z_i)$ is a normalized histogram, and therefore, it represents an estimate of the probability of occurrence of intensity value, and m is the mean value defined as

$$m = \sum_{i=1}^{L-1} z_i\, p(z_i) \qquad (3)$$

In this study, we are interested in the mean value of the histogram, the square root of the second-order moment, the third-order moment, and entropy. From the preceding equations, it is clear that the second-order moment represents variance. The feature used is the standard deviation expressed by the square root of the variance. The third-order moment represents the skewness of the histogram. This is a measure

of symmetry of the histogram where a value of 0 represents a symmetric histogram relative to the mean, a positive value represents a histogram skewed to the right, and a negative value represents a histogram skewed to the left, relative to the mean. A fourth histogram property used in this paper is the randomness of the intensity distribution as expressed by the entropy [14]

$$e = -\sum_{i=1}^{L-1} p(z_i) \log_2 p(z_i) \qquad (4)$$

where e is the entropy, L the number of intensity levels, and is as defined earlier. Therefore, four features were derived in this category.

Statistical texture features based on the Gray Level Co-occurrence Matrix (GLCM) were also extracted from ROI. This matrix characterizes the spatial distribution of gray levels in an image or a region. An element at location (i, j) of the co-occurrence matrix expresses the joint probability density of the occurrence of gray levels i and j in a specific orientation θ and distance d from each other. The choice of different orientation and distance generates different matrices.

Here, four directions ($\theta=0^O$, $\theta=45^O$, $\theta=90^O$, $\theta=135^O$) and one distance ($d=1$pixel) were used to generate GLCMs. Texture measures can be extracted from each GLCM and the mean value was used as one feature. Hence, 56 texture features were computed for each closed regions according to Haralick's definition [15].

Wavelet-based methods for texture classification can be broadly classified into two categories, namely feature-based and model-based methods. Here, we used an approach [16] based on wavelet transformation and singular value decomposition (SVD). This approach consists of the following stages a) wavelet transformation of image textures, b) introduction of non-linearity on the wavelet transformation coefficients, c) SVD of the transformation coefficients after introduction of nonlinearity, d) truncation of lower singular values for effective classification of textures in the presence of noise, e) modeling the PDF of retained singular values using exponential function, and f) estimation of the model parameters using maximum likelihood estimation technique. Decomposition is made with db4 wavelets for three levels. Thus, nine features were derived.

Features Selection

Inclusion of unsuitable features may adversely affect the classification performance. The goal of the feature selection step is to choose the optimal feature vector consisting of only those features that minimize the classification error from 69 features. Many feature selection methods have been developed for specific contexts, each having their weaknesses. Principal Component Analysis (PCA) is a possible solution to find features that retain most of the information. Here four best features were selected for classification.

Classification

In this study, SVM is used for classification because of its powerful pattern classification capability [17]. Assume that we want to build a linear classifier that best separates two populations in high-dimensional space. This classifier is described by a hyperplane whose position and orientation must be determined with the help of a pre-classified training set.

The optimal parameters of the dividing hyperplane are determined via an iterative constrained quadratic optimization scheme, in which the training samples of one group are forced to be on one side of the hyperplane and the samples of the other group are forced to be on the opposite side. This problem is solved via a variety of non-linear programming techniques [18] that results in many "active" constraints that determine the solution. These constraints correspond to samples that are very close to or are on the interface between the two groups. These are called "support vectors". The rest of the training samples do not contribute to the expression of the dividing hyperplane. This reveals a very important aspect of SVM, which is one of the reasons for its effectiveness as a classifier. The hyperplane is determined only by a relatively small number of samples that are close to the opposite group. The samples that are far away have no influence on the results. The classifier inherently focuses on the subtleties of the morphological differences between the two groups and not on gross differences. Therefore, it is more effective. In practice, however, it is impossible to prevent the two groups from overlapping to some degree. Therefore, the constraints are relaxed to permit some training samples to be on the wrong side of the hyperplane.

Results and discussions

In order to test the proposed method, we selected 133 AD brain T1-weighted MRI dataset and 146 normal brain T1-weighted MRI dataset from ADNI database. From all the datasets, hippocampus is segmented from the coronal slice using the method listed above [10]. The slice of large area of hippocampus is selected for feature extraction. Sixty-nine texture features were extracted using the methods discussed above. Four best features were selected using PCA. Finally, SVM is used for classification. To train and test the SVM, out of 133 AD features 88 AD features were randomly assigned for training and remaining 45 AD features were assigned for testing and out of 146 normal features, 97 normal features were assigned for training and the remaining 49 normal features were assigned for testing.

To validate and quantify the effectiveness of the proposed texture feature extraction, selection and classification algorithm, sensitivity, accuracy, specificity, and precision were measured [19]. Events that assign true positive (TP), false negative (FN), true negative (TN), and false positive (FP) used in the performance metrics equations are defined as follows;

$$\Pr ecision = \frac{TP}{TP+FP} \qquad (5)$$

$$Sensitivity = \frac{TP}{TP+FN} \qquad (6)$$

$$Specificity = \frac{TN}{FP+TN} \qquad (7)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (8)$$

Case I: Training and test features were taken from statistical properties of the intensity histogram of the hippocampus. The precision, sensitivity, specificity, and accuracy are 90.9%, 88.9%, 91.8%, and 90.4% respectively. The five misclassified datasets that belongs to ADs that are for above 80-years-old. Furthermore, one dataset belongs to a male and other two belongs to females. Four misclassified dataset that belongs to normal are above 85-years-old and both belongs to female. When the number of features was reduced, the validation parameters were also changed. However, it gave better values when all the features were taken into account.

Case II: Fifty-six texture features derived by GLCM is taken and the best four features were selected using PCA for classification. The validation parameters Precision, Sensitivity, Specificity, and Accuracy are 88.9%, 88.9%, 89.8%, and 89.4% respectively. Here the five misclassified datasets belongs to AD are above 85 years old and one dataset belongs to male and other two belongs to female. The five misclassified dataset belongs to normal are above 85-years-old and one belongs to female and other two belongs to male. Compared to the previous case here the Normal detection rate is high but its response is poor for AD detection.

Case III: Nine texture features taken using wavelet and SVD were reduced to four by PCA then classified by SVM. The validation parameters Precision, Sensitivity, Specificity, and Accuracy are 89.1%, 91.1%, 89.8%, and 90.4% respectively. Here the four misclassified datasets belongs to AD that are above

80-years-old and one dataset belongs to male and another belongs to female. The five misclassified dataset belongs to normal that are above 85-years-old and one belongs to female and other two belongs to male. The performance is better compared to the previous cases.

Case IV: All 69 texture features were reduced to best four by PCA and then classified using SVM. The precision, sensitivity, specificity, and accuracy are 95.3%, 91.1%, 95.9%, and 93.6% respectively. Here the four misclassified datasets belongs to AD that are above 85-years-old and one dataset belongs to male and another belongs to female. The two misclassified dataset belongs to normal that are above 85-years-old and it belongs to female. This show, it has highest classification rate with good accuracy and precision.

The confusion matrix provides an easy and complete way to describe the knowledge about a classification result. It is given in **Figure 1.** First row
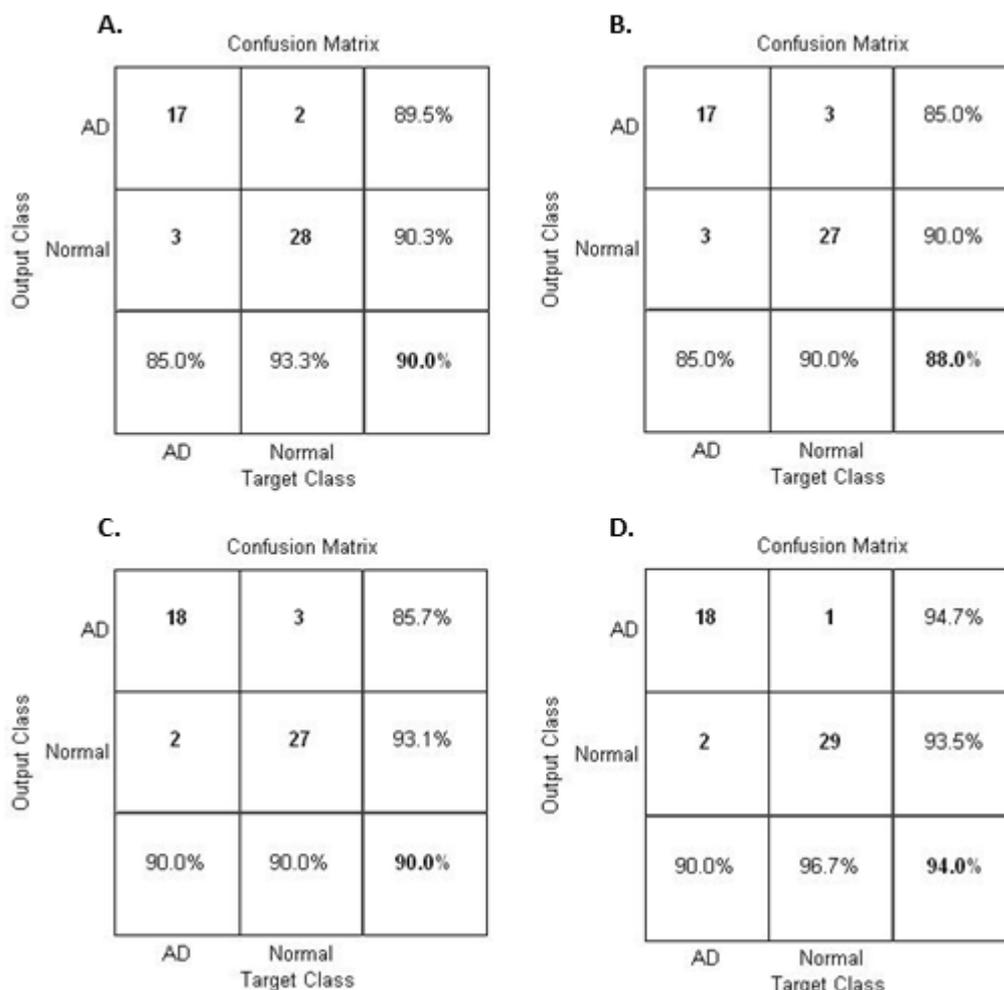
Figure 1. Confusion matrix of four cases; I (**A**), II (**B**), III (**C**), and IV (**D**).

of the third column shows the positive predictive value. The second row of third column shows the negative predictive value. The third row of first column shows the sensitivity, which gives the status, that how far the positive cases are correctly classified. The third row of second column gives the status, which is how far the negative cases are classified. Finally, the third row of third column shows the accuracy. The Receiver Operating Characteristic (ROC) curve is a plot with the false positive rate on the *X* axis and the true positive rate on the *Y* axis as shown in **Figure 2.**

Finally, the proposed method is applied to the entire brain image. The confusion matrix and ROC curve

for this analysis can be seen in **Figure 3**. This analysis shows texture features taken only from hippocampus gives better discrimination between AD and Normal. It proves that due to AD, the textures of hippocampus are much affected.

Currently we are working towards the prediction of AD from MCI using the proposed method. For this analysis, seven MCI patients are involved. On every visit, MRI, t-tau/AB and MMSE scores are recorded. At the end of this analysis, it is possible to compare the proposed method with other single and multi-test models.
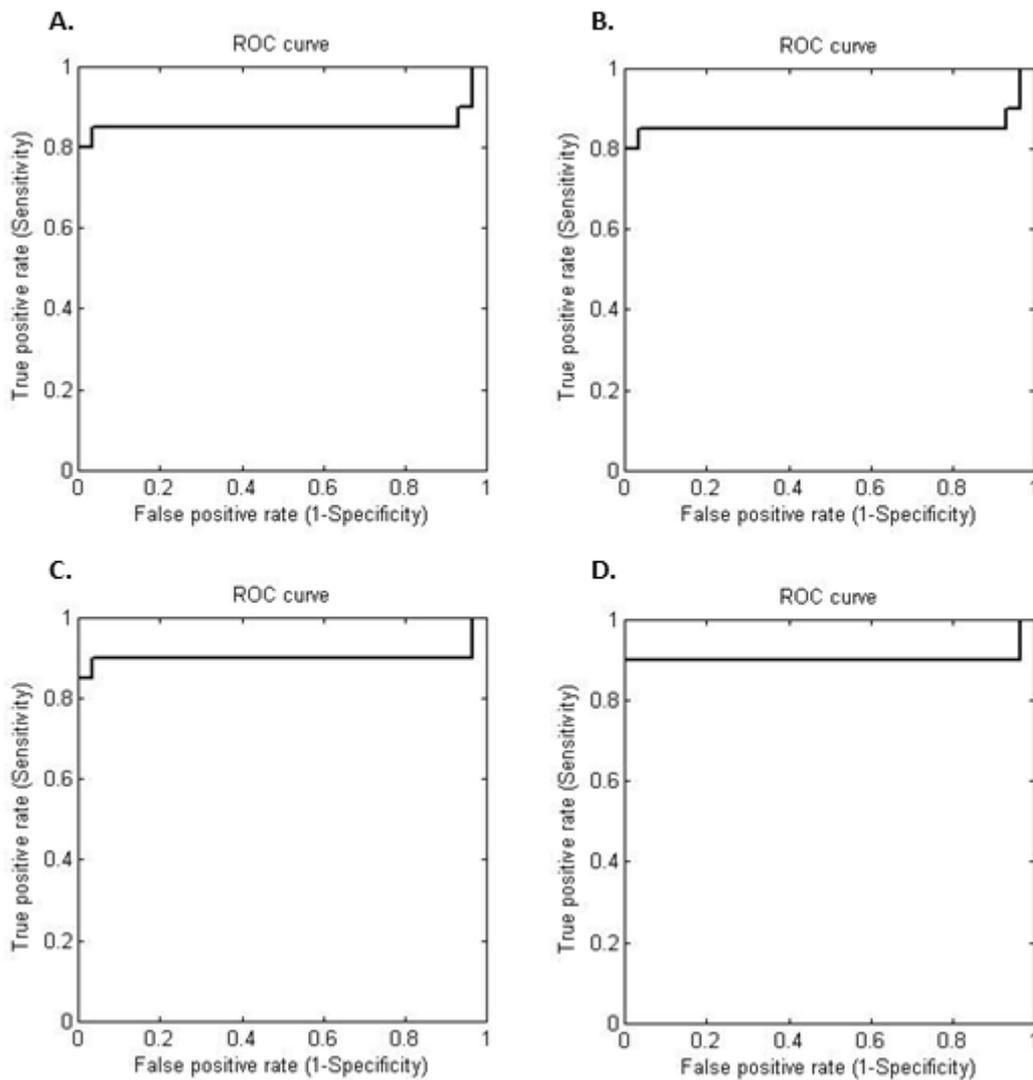
Figure 2. ROC curve of four cases; I (A), II (B), III (C), and IV (D).
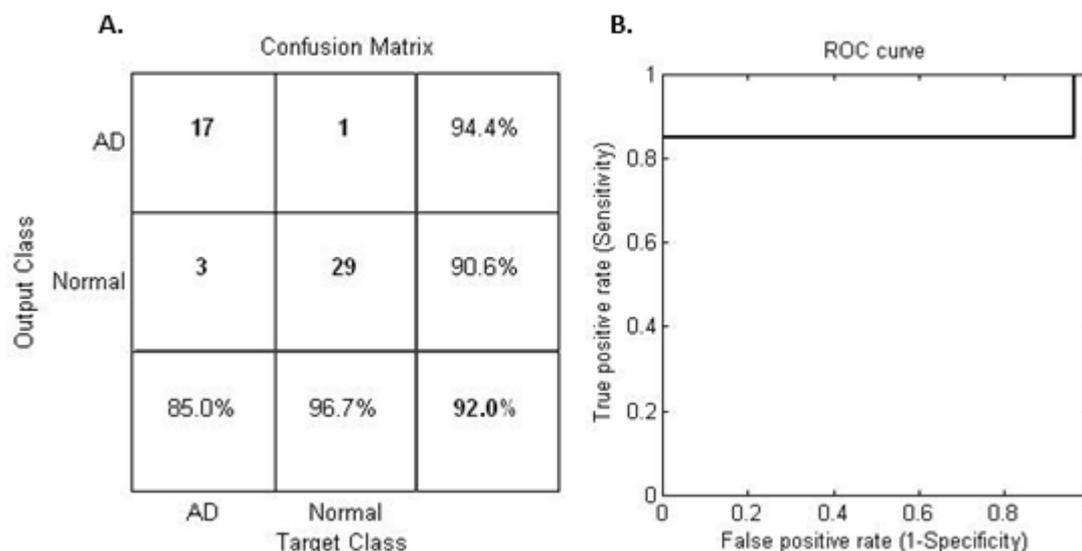
Figure 3. The analysis made on whole brain image; confusion matrix (**A**) and ROC curve (**B**).

Conclusion

In this paper, we have made an analysis, whether the texture features of hippocampus on MRI can be used as a biomarker to identify AD. The procedure involved are segmenting the hippocampus, extracting the features, selecting the best features using PCA, and discriminating the AD from normal using SVM. Analysis shows that proper selection of texture features can discriminate the AD from Normal from hippocampus texture features. The work can be further extended to include the texture features of MCI datasets. In addition, the datasets from various scanners can also be included for the analysis. Furthermore, the MCI patients can be monitored for two to three years and the possibility of conversion from MCI to AD can be predicted by this method.

Acknowledgment

References

1. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. Lancet. 2005; 366:2112-7.

2. Ritchie CW, Ames D, Clayton T, et al. Metaanalysis of randomized trials of the efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer disease. Am. J. Geriatr. Psychiatry. 2004; 12: 358-69.

3. Whitehead A, Perdomo C, Pratt RD, et al. Donepezil for the symptomatic treatment of patients with mild to moderate Alzheimer's disease: a meta-analysis of individual patient data from randomized controlled trials. Int. J. Geriatr. Psychiatry. 2004; 19:624-33.

4. Devanand DP, Pradhaban G, Liu X, et al. Hippocampal and entorhinal atrophy in mild cognitive impairment. Neurology. 2007; 68:828-36.

5. Gerardinn E, Ch telat G, Chupin M, et al. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. NeuroImage. 2009; 47:1476-86.

6. Whitwell JL, Josephs KA, Murray ME, et al. MRI correlates of neurofibrillary tangle pathology at autopsy. Neurology. 2008; 71:743-9.

7. Scahill RI, Schott JM, Stevens JM, et al. Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. PNAS. 2002; 99:4703-7.

8.   Desikan RS, Cabral HJ, Hess CP, et al. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. Brain. 2009; 132: 1-10.

9.   Freeborough PA, Fox NC. MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease. IEEE transactions on medical imaging. 1998; 17:475-9.

10. Rajeesh J, Moni RS, Palanikumar S, et al. A versatile algorithm for the automatic segmentation of hippocampus based on level set. International Journal of Biomedical Engineering and Technology. 2011; 7: 213-24.

11. Xia Y, Hu Q, Aziz A, et al. A knowledge-driven algorithm for a rapid and automatic extraction of the human cerebral ventricular system from MR neuroimages, Neuro Image. 2004; 21:269-82.

12. Rajeesh J, Moni RS, Palanikumar S. A comparative analysis of filtes for histogram smoothing in automatic segmentation of CN from brain MR image. National Journal of Technology. 2009; 5:84-7.

13. Sethian JA. Level set methods and fast marching methods. In: Cambridge, UK: Cambridge University Press.1996.

14. Gonzalez RC, Woods RE, Eddins SL. Digital image processing using matlab. Englewood Cliffs, NJ: Prentice-Hall, 2004.

15. Haralick RM, Shanmugam K, Dinstein I. Textural features for.image classification. IEEE Trans. Syst., Man, Cybern. 1973; 6:610-21.

16. Selvan S, Ramakrishnan S. SVD-based modeling for image texture classification using wavelet transformation. IEEE Transactions on Image Processing. 2007; 11:2688-96.

17. Lao Z, Shen D, Xue Z, et al. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. Neuroimage. 2004; 21:46-57.

18. Burges CJC. A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. 1998; 2:121-67.

19. Sonka M, Fitzpatrick JM. Handbook of Medical Imaging: Medical Image Processing and Analysis. Bellingham, WA: SPIE; 2004.