International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:6, No:12, 2012

Automatic Classification of Initial Categories of Alzheimer's Disease from Structural MRI Phase Images: A Comparison of PSVM, KNN and ANN Methods

Ahsan Bin Tufail, Ali Abidi, Adil Masood Siddiqui, and Muhammad Shahzad Younis

***Abstract*—**An early and accurate detection of Alzheimer's disease (AD) is an important stage in the treatment of individuals suffering from AD. We present an approach based on the use of structural magnetic resonance imaging (sMRI) phase images to distinguish between normal controls (NC), mild cognitive impairment (MCI) and AD patients with clinical dementia rating (CDR) of 1. Independent component analysis (ICA) technique is used for extracting useful features which form the inputs to the support vector machines (SVM), K nearest neighbour (kNN) and multilayer artificial neural network (ANN) classifiers to discriminate between the three classes. The obtained results are encouraging in terms of classification accuracy and effectively ascertain the usefulness of phase images for the classification of different stages of Alzheimer's disease.

***Keywords*—**Biomedical image processing, classification algorithms, feature extraction, statistical learning.

I. Introduction

ALZHEIMER'S disease (AD) is a neurodegenerative brain disorder that is characterized by neurofibrillary tangles, amyloid plaques and histopathologic changes that are typically associated with neuronal loss and volume reductions [1]. Due to longevity and aging populations, the prevalence of AD among the elderly will increase in the coming years.

Due to the acute nature of AD, a number of attempts have been made in the literature to carefully diagnose its patients using a combination of biomarkers such as positron emission tomography (PET), magnetic resonance imaging (MRI) and cerebrospinal fluid (CSF) biomarkers with considerable better performance [2]. Region of interest based analysis methods such as those based on measuring the cortical thickness to differentiate between different dementias using MRI imaging modality proves to be a useful surrogate biomarker for differentiating among dementias [3]. Similarly, subtle and complex deformation patterns of hippocampus analyzed using machine learning approaches such as support vector machine (SVM) can differentiate among elderly healthy controls and AD patients with good accuracy rate [1] [4]–[6].

However, in recent years, focus has been shifted towards whole brain based multivariate methods that are geared towards the use of statistical machine learning methods such as support vector machines based on bootstrap resampling approach [7], integration of MRI and PET water scanning and their classification [8], wavelet decomposition and subsequent reduction of morphological features and their classification using nonlinear support vector machines to discriminate between the two genders [9] as well as inputs to self-organizing map (SOM) and SVM to discriminate the MRI images as normal or abnormal [10], the classification of single photon emission computed tomography (SPECT) images using 1-norm SVM classifier [11] and SVM and principal component analysis (PCA) classifiers [12], as well as using SPECT and PET images to diagnose AD using PCA and a posteriori Bayesian classifiers [13] or using only PET scan for discrimination among normal and AD subjects using neural network classifier [14].

A number of studies have been conducted recently to aid in the research effort geared towards the comparisons of different methodologies for effective aiding in the diagnosis of AD. For example, to differentiate among electroencephalograms (EEGs) of patients with varying degree of AD and their age-matched control subjects using a subset of methods like principal component linear discriminant analysis (PC LDA), partial least squares LDA (PLS LDA), principal component logistic regression (PC LR), partial least squares logistic regression (PLS LR), bagging, random forest, support vector machines (SVM) and feed-forward neural network [15] as well as sMRI to discriminate among MCI, AD and NC subjects [16].

SVM, K-Nearest Neighbour (KNN) and multilayer artificial neural network (ANN) are among the widely used classification methods in medical settings [17]. In the present work, comparison and analysis of the performance of these methods is performed to discriminate between normal controls (NC), mild cognitive impairment (MCI) and AD subjects using fixed point independent component analysis (ICA) features.

The contents of the paper are explained next. In section II,

Ahsan Bin Tufail, Adil Masood Siddiqui and Ali Abidi are with the Department of Electrical Engineering, Military College of Signals, National University of Sciences and Technology, Islamabad, Pakistan (e-mail:{ahsantufail, dradil, abidi}@mcs.edu.pk).

Muhammad Shahzad Younis is with the Department of Electrical Engi neering, School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan (e-mail: muhammad.shahzad@seecs.edu.pk).

the experimental subjects and the methods for analysis and discrimination for the feature extraction and classification purposes especially the concept of ICA, proximal support vector machines (PSVM), kNN and multilayer ANN classifiers are explained. In section III, the results of the classification process are presented followed by the discussion of the said results in section IV and finally section V concludes the present work.

II. Subjects and Methods

A. Subjects and sMRI Protocols

The demographics of the subjects selected from Open Access Series of Imaging Studies (OASIS) database [18] are shown in Table I. The MRI acquisition protocols are shown in Table II. We made use of the tissue classification images in the dataset a sample of which is shown in Fig. 1.

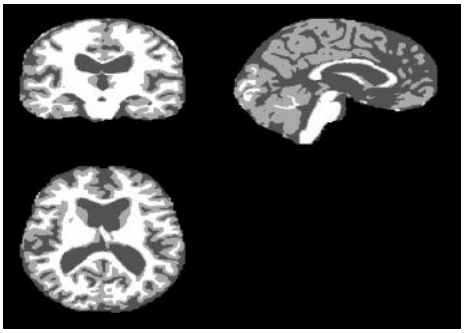

Fig. 1 A sample of the tissue classification image in the database

Fig. 2 Proposed solution

B. Methods

The framework of our proposed solution is shown in Fig. 2. We used the tissue classification images to form the phase images [19]. Then, we apply the trilinear interpolation method to reduce the size of the phase image from 176×208×176 to 176×208×4. All the reduced phase images are converted into one dimensional row vectors and the dimensionality of the corresponding matrix is reduced using principal component analysis (PCA) method to retain first hundred components. The idea of PCA is based on singular value decomposition (SVD) technique [20].

TABLE I
Subject Demographics and Dementia Status

	NC	MCI	AD
Number of Subjects	**83**	**62**	**25**
Male/Female	**20/63**	**28/35**	**8/17**
Age	**75.24±8.96**	**76.13±7.55**	**78.20±7.21**
Education	**3.34±1.30**	**2.92±1.27**	**2.60±1.32**
SES	**2.47±1.03**	**2.69±1.13**	**2.86±1.35**
CDR	**0**	**0.5**	**1**
MMSE	**29±1.20**	**25.5±3.66**	**21.64±3.97**

TABLE II
MRI Acquisition Details

Sequence	MP-RAGE
TR (ms)	9.7
TE (ms)	4
Flip Angle (°)	10
TI (ms)	20
TD (ms)	200
Orientation	Saggital
Thickness, gap (mm)	1.25, 0
Slice number	128
Resolution (pixels)	256×256 (1×1 mm)

1) Independent Component Analysis (ICA)

ICA [21] is a probabilistic and multivariate method for learning a linear transform of random vectors. The basic goal of ICA is to search for the components which are maximally as independent and non-Gaussian as possible. Its fundamental difference to classical multivariate statistical methods such as PCA and linear discriminant analysis (LDA) is in the assumption of non-gaussianity, which ensures the identification of original components, in comparison with these classical methods. ICA can be mathematically modelled as,

$$X = A \times S \quad (1)$$

In (1), X is the observed data vector, A is the mixing matrix and S is the source matrix. In our work, we made use of the FastICA matlab toolbox to compute both A and S from X. The mixing matrix A has been considered in the subsequent steps of feature selection and classification.

2) Proximal support vector machine (PSVM)

Support vector machine [22] is a versatile data classification method widely used in the machine learning domain. It can be used to classify both linearly and nonlinearly separable data. Kernel trick is used to separate examples that are non-linearly separable in the space of the inputs and might be separable in a higher dimensionality feature space given a suitable mapping. We made use of the inverse multiquadratic kernel [23] which is defined as follows:

$$1 / \sqrt{(\|x_i - x_j\|^2 + c)} \quad (2)$$

In (2), c is a constant greater than zero while x_i and x_j are variables dependent on the available data. In the present work, we made use of the value of c as 10^{-3}.

Proximal support vector machine (PSVM) [24] classifies datapoints depending on the proximity of the two parallel planes that are pushed as far apart as possible by solving a single system of linear equations. In the present scenario, we used the nonlinear version of PSVM that employed inverse multiquadratic kernel to classify the datapoints in the three classes.

3) K-Nearest Neighbour (KNN)

K-Nearest Neighbour (KNN) is a data mining algorithm with a wide range of applications in the image processing domain [22]. There are three key elements of this approach: a set of labeled training examples, a distance measure to compute the distance between the training set examples and the test example, and the value of k; i.e., the number of nearest neighbours to the testing example.

We used Euclidean and Riemannian distance [25] measures in our work to classify the testing set examples from the three classes which can be mathematically expressed as:

$$\text{Euclidean distance} = \sqrt{(\sum_{i=1}^{4}(x_i - y_i)^2)} \quad (3)$$

$$\text{Riemannian distance} = \| \log x_i^{-1} y_i \| \quad (4)$$

4) Multilayer artificial neural network (ANN)

Artificial Neural Networks [26] [27] is the mainstay of modern data mining algorithms biologically inspired by connections inside brain used to carry information. In our work, we used matlab neural network toolbox to classify the test datapoints in the three classes comparing two classes at a time using tansigmoid transfer function with Levenberg-Marquardt (LM) optimization method to adjust the weights for all the neurons in all layers of the networks except the output layer in which the linear transfer function has been used.

Mathematically, the tan-sigmoid function can be expressed as:

$$\text{tansig}(n) = 2/(1+e^{-2n}) - 1 \quad (5)$$

while the linear transfer function can be represented as:

$$\text{linear}(n) = n \quad (6)$$

We used two different neural network configurations. For comparison between NC and MCI classes, we used a network with 8 neurons in the input, 8 neurons in the hidden layer 1, 8 neurons in the hidden layer 2 and one neuron with linear transfer function in the output layer. For comparison between MCI and AD and NC and AD, we used a network with 4 neurons in the input, 4 neurons in the hidden layer 1, 4 neurons in the hidden layer 2 and one neuron with linear transfer function in the output layer.

III. Results

TABLE III
KNN Classification Using Riemannian Distance between NC & MCI Classes

	True	False
Positive	**0.7286 (TP)**	**0.5679 (FP)**
Negative	0.2714 (FN)	0.4321 (TN)

TABLE IV
ANN Classification between NC & MCI Classes

	True	False
Positive	0.4286 (TP)	0.3571 (FP)
Negative	0.5714 (FN)	0.6429 (TN)

TABLE V
PSVM Classification between NC & MCI Classes

	True	False
Positive	0.2653 (TP)	0.3214 (FP)
Negative	0.7347 (FN)	0.6786 (TN)

TABLE VI
KNN Classification Using Euclidean Distance between NC & AD Classes

	True	False
Positive	0.8235 (TP)	0.5294 (FP)
Negative	0.1765 (FN)	0.4706 (TN)

TABLE VII
ANN Classification between NC & AD classes

	True	False
Positive	0.3333 (TP)	0.5294 (FP)
Negative	0.6667 (FN)	0.4706 (TN)

TABLE VIII
PSVM Classification between NC & AD Classes

	True	False
Positive	0.6615 (TP)	0.4881 (FP)
Negative	0.3385 (FN)	0.5119 (TN)

TABLE IX
KNN Classification Using Euclidean Distance between MCI & AD CLASSES

	True	False
Positive	0.9792 (TP)	0.6181 (FP)
Negative	0.0208 (FN)	0.3819 (TN)

TABLE X
ANN Classification between MCI & AD Classes

	True	False
Positive	0.3333 (TP)	0.5000 (FP)
Negative	0.6667 (FN)	0.5000 (TN)

TABLE XI
PSVM Classification between MCI & AD Classes

	True	False
Positive	0.6508 (TP)	0.4378 (FP)
Negative	0.3492 (FN)	0.5622 (TN)

We used the following convention in Tables III, IV and V: True positive (TP) is the number of MCI instances correctly diagnosed as MCI while false positive (FP) is the number of NC instances diagnosed as MCI, true negative (TN) is the number of NC subjects diagnosed as NC, while false negative (FN) is the number of MCI instances diagnosed as NC. In Tables VI, VII and VIII, we used the following convention: TP is the number of AD subjects correctly diagnosed as AD, FP is the number of NC subjects diagnosed as AD, TN is the number of NC subjects diagnosed as NC while FN is the number of AD subjects diagnosed as NC. In Tables IX, X and XI, we used the following convention: TP is the number of AD subjects correctly diagnosed as AD, FP is the number of MCI subjects diagnosed as AD, TN is the number of MCI subjects diagnosed as MCI while FN is the number of AD subjects incorrectly diagnosed as MCI.

The classification accuracy defined as (TP + TN) / (TP +

TN + FP + FN) is found to be 58.04%, 64.71% and 68.06% for the KNN classification between NC and MCI, NC and AD, and MCI and AD classes respectively. For ANN classification, the accuracy rate is 53.57%, 40.20%, 41.67% for classification between NC and MCI, NC and AD, and MCI and AD classes respectively. For PSVM classification, the accuracy rate is found to be 47.19% for MCI and NC, 58.67% for NC and AD classes, and 60.65% for testing between MCI and AD classes.

IV. Discussion

From the results, it can be concluded that the KNN classifier is a good option for the overall classification between either of the three classes due to its superior accuracy rate than its counterparts. However, PSVM classifier performs a lot better in identifying the true negatives (TN) than that of its counterparts.

ANN does not perform up to the mark in this application because of the addition of extra dimensions in the weight space which creates local minima in the lower dimensional subspaces. Also, the time taken by ANN training is several magnitudes higher than its counterparts.

V. Conclusion

In this paper, we compare the performance of the three widely used classifiers in data mining, that is, PSVM, KNN and ANN classifiers, for the classification of AD, MCI and NC subjects from 3D structural MRI data sets of OASIS database. Features in this study are extracted using independent component analysis technique. The results obtained were satisfactory in terms of both accuracy and computational speed for the KNN and PSVM classifiers. It would be interesting to see how the results vary by using other feature extraction and selection methods and large MRI data sets such as those provided by Alzheimer's Disease Neuroimaging Initiative (ADNI). There are also some other possible extensions to the current work such as increasing the number of classes, using neuropsychological information as features and adding more classification methods in comparative analysis.

References

[1] Li S, Shi F, Pu F, Li X, Jiang T, Xie S and Wang Y, "Hippocampal Shape Analysis of Alzheimer Disease Based on Machine Learning Methods," AJNR Am J Neuroradiol, vol. 28, no. 7, pp. 1339–1345, August 2007.

[2] Zhang D, Wang Y, Zhou L, Yuan H, Shen D and the Alzheimer's Disease Neuroimaging Initiative, "Multimodal classification of Alzheimer's disease and mild cognitive impairment," NeuroImage, vol. 55, no. 3, pp. 856–867, April 2011.

[3] An-Tao Du, Norbert Schuff, Joel H. Kramer, Howard J. Rosen, Maria Luisa Gorno-Tempini, Katherine Rankin, Bruce L. Miller and Michael W. Weiner, "Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia," Brain, vol. 130, no. 4, pp. 1159–1166, March 2007.

[4] Gerardin E, Chetelat G, Chupin M, Cuingnet R, Desgranges B, Kim HS, Niethammer M, Dubois B, Lehericy S, Garnero L, Eustache F, Colliot O and the Alzheimer's Disease Neuroimaging Initiative, "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging," NeuroImage, vol. 47, no. 4, pp. 1476–1486, October 2009.

[5] Jonathan H. Morra, Zhuowen Tu, Liana G. Apostolova, Amity E. Green, Christina Avedissian, Sarah K. Madsen, Neelroop Parikshak, Xue Hua, Arthur W. Toga, Clifford R. Jack, Michael W. Weiner, Paul M. Thompson and the Alzheimer's Disease Neuroimaging Initiative, "Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls," NeuroImage, vol. 43, no. 1, pp. 59–68, October 2008.

[6] Chupin M, Gerardin E, Cuingnet R, Boutet C, Lemieux L, Lehericy S, Benali H, Garnero L, Colliot O and the Alzheimer's Disease Neuroimaging Initiative, "Fully Automatic Hippocampus Segmentation and Classification in Alzheimer's Disease and Mild Cognitive Impairment Applied on Data From ADNI," Hippocampus, vol. 19, no. 6, pp. 579–587, June 2009.

[7] Magnin B, Mesrob L, Kinkingnehun S, Pelegrini-Issac M, Colliot O, Sarazin M, Dubois B, Lehericy S and Benali H, "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI," Neuroradiology, vol. 51, no. 2, pp. 73–83, February 2009.

[8] Fan Y, Resnick SM, Wu X and Davatzikos C, "Structural and functional biomarkers of prodromal Alzheimer's disease: A high-dimensional pattern classification study," NeuroImage, vol. 41, no. 2, pp. 277–285, June 2008.

[9] Lao Z, Shen D, Xue Z, Karacali B, Resnick SM and Davatzikos C, "Morphological classification of brains via high-dimensional shape transformations and machine learning methods," NeuroImage, vol. 21, no. 1, pp. 46–57, January 2004.

[10] Sandeep Chaplot, L.M. Patnaik and N.R. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network," Biomedical Signal Processing and Control, vol. 1, no. 1, pp. 86–92, September 2010.

[11] Stoeckel, J. and Fung, G., "SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information," Knowl Inf Syst, vol. 11, no. 2, pp. 243–258, September 2006.

[12] I. Alvarez, J.M. Gorriz, J. Ramirez, D. Salas-Gonzalez, M. Lopez, C.G. Puntonet and F. Segovia, "Alzheimer's Diagnosis Using Eigenbrains and Support Vector Machines," Electronics Letters, vol. 45, no. 7, pp. 342–343, March 2009.

[13] M. Lopez, J. Ramirez, J.M. Gorriz, D. Salas-Gonzalez, I. Alvarez, F. Segovia and C.G. Puntonet, "Automatic tool for Alzheimer's disease diagnosis using PCA and Bayesian classification rules," Electronics Letters, vol. 45, no. 8, pp. 389–391, April 2009.

[14] J. Shane Kippenhan, Warren W. Barker, Shlomo Pascal, Joachim Nagel and Ranjan Duara, "Evaluation of a Neural-Network Classifier for PET Scans of Normal and Alzheimer's Disease Subjects," The Journal of Nuclear Medicine, vol. 33, no. 8, pp. 1459–1467, August 1992.

[15] Lehmann C, Koenig T, Jelic V, Prichep L, John RE, Wahlund LO, Dodge Y and Dierks T, "Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)," Journal of Neuroscience Methods, vol. 161, no. 2, pp. 342–350, April 2007.

[16] Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehericy S, Habert MO, Chupin M, Benali H, Colliot O 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, vol. 56, no. 2, pp. 766–781, May 2011.

[17] S.R. Amendolia, G. Cossu, M.L. Ganadu, B. Golosio, G.L. Masala and G.M. Mura, "A comparative study of K-Nearest Neighbour, Support Vector Machine and Multi-Layer Perceptron for Thalassemia screening," Chemometrics and Intelligent Laboratory Systems, vol. 69, no. 1–2, pp. 13–20, November 2003.

[18] Daniel S. Marcus, Tracy H. Wang, Jamie Parker, John G. Csernansky, John C. Morris and Randy L. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults," Journal of Cognitive Neuroscience, vol. 19, no. 9, pp. 1498–1507, 2007.

[19] Lai Xu, Jingyu Liu, Tulay Adali and Vince D. Calhoun, "Source based morphometry using structural MRI phase images to identify sources of gray matter and white matter relative differences in schizophrenia versus controls," in ICASSP, 2008.

[20] Virginia C. Klema and Alan J. Laub, "The Singular Value Decomposition: Its Computation and some Applications," IEEE Transactions on Automatic Control, vol. 25, no. 2, pp. 164–176, April 1980.
[21] Aapo Hyvarinen, "A family of fixed-point algorithms for independent component analysis," in ICASSP, 1997.
[22] Xindong Wu, Vipin Kumar, J. Ross, Quinlan Joydeep, Ghosh Qiang Yang, Hiroshi Motoda, Geoffrey J. Mclachlan, Angus Ng, Bing Liu, Philip S. Yu, Dan Steinberg, X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang and H. Motoda , "Top 10 algorithms in data mining," Knowl Inf Syst, vol. 14, no. 1, pp. 1–37, December 2007.
[23] Paul Honeine and Cedric Richard, "Preimage Problem in Kernel-Based Machine Learning," IEEE Signal Processing Magazine, pp. 77–88, March 2011.
[24] Glenn Fung and Olvi L. Mangasarian, "Proximal Support Vector Machine Classifiers," Knowledge discovery and data mining (KDD), pp. 77–86, 2001.
[25] Xiaojing Long and Chris Wyatt, "An automatic unsupervised classification of MR images in Alzheimer's disease," in Computer Vision and Pattern Recognition (CVPR), June 2010, pp. 2910–2917.
[26] Martin T. Hagan and Mohammad B. Menhaj, "Training feedforward networks with the Marquardt algorithm," IEEE Transactions on Neural Networks, vol. 5, no. 6, pp. 989–993, November 1994.
[27] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning Representations by Back-Propagating Errors," Nature, vol. 323, pp. 533–536, October 1986.