35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Identification of Brain White Matter Regions for Diagnosis of Alzheimer using Diffusion Tensor Imaging

Ravindra B Patil, Ranjan Piyush and S Ramakrishnan

Abstract— **Diffusion Tensor Imaging (DTI) technique is widely used to probe the white matter (WM) tracts, which is affected most by neurological disorders. The fractional anisotropy (FA) metric has been used predominantly to study changes in the WM tracts. Here an attempt is made to delineate specific regions of interest in the WM that may be probable indicators for the diagnosis of Alzheimer disease (AD). Genetic algorithm has been used as feature reduction method along with Adaptive Boosting (AdaBoost) machine learning technique to determine the most prominent regions in the WM that are indicators of AD. It is found in this study that Fornix region of WM is most affected by Alzheimer. Further, classification was done to differentiate between Alzheimer and Normal controls with accuracy of 84.5%. The results obtained were validated by comparing with the existing literature on Alzheimer.**

I. Introduction

Alzheimer's disease (AD) causes major anatomical degeneration to the neurons and synapses in specific regions of hippocampus in the brain. Statistical estimates reveal that 1 in every 85 persons will have developed this dementia by 2050 [1]. As AD progresses, substantial damage is caused to the white matter (WM) tracts which, if left untreated can result in subsequent loss of memory leading to death of the patient [4]. The clinical treatment available today is effective only in the early stages of disease thus detection of AD in its early stages is a pressing issue of the present times.

Diffusion tensor imaging (DTI) is one of the non-invasive in vivo techniques to probe the WM tracts in brain [3]. DTI works on the probabilistic determination of diffusion of water molecules over a given time period in a tissue structure. This imaging technique is a modified MRI method which uses a variable gradient of magnetic field in required co-ordinate directions to map the diffusion of water molecules. Images are acquired in a minimum of six directional planes and are used to generate the voxel map. Due to varying magnitude of diffusion in each direction, the diffusion tensor acquires an ellipsoidal shape in three dimensional spaces. This can be used very effectively to study the cause of delineation of WM tracts resulting in the development of AD.

The most widely studied index of DTI is Fractional Anisotropy (FA).This gives a measure of the molecular displacement anisotropy of water in space [5]. FA can be physically demonstrated by the eccentricity of ellipsoidal voxel. FA provides the degree of anisotropy of diffusion process which varies from 0 to 1 and decreases with the loss of myelin and axons.

The characteristic study on WM has been carried out since long. Analysis on WM tracts is done using atlas based approach followed by manual segmentation to identify the abnormalities present. Kenichi et.al [6] used Atlas based approach to map all the regions of the WM using large deformation diffeomorphic mapping and analyzed WM deformity AD subject. Study done by White et.al demonstrated Voxel Based Analysis (VBA) [10] to delineate the region of interest in WM. To date many machine learning approaches have been used for the diagnosis of AD using MRI images. Laurence et.al [7] used Support Vector Machine along with FA for classification of Mild Cognitive Impairment. M.Grana et.al [8] used Pearson's correlation for diagnosis of AD. Desikan et.al used volumetric analysis of hippocampus combined with logistic regression [9].

The aim of present study is to investigate the specific regions in the white matter which are most affected by AD with an automated approach. Feature reduction technique and machine learning algorithm are used to achieve this objective. In this study genetic algorithm is used to identify specific regions in white matter which describes the AD pathology most efficiently. Further, an adaptive method of classifier development AdaBoost is applied to ascertain the prominence of each affected region by classifying AD from Normal Control (NC). A tree classifier is initialized in this study which on repeated iteration with weighted voting renders a highly efficient classifier. The regions best describing the classification accuracy obtained from this method were compared with those mentioned in literature. Most of the regions were found to be consistent to regions responsible for AD.

The rest of the paper is organized as follows. Section II deals with the steps involved in acquiring data, feature extraction, genetic algorithm for feature reduction and classification using adaptive method. Section III describes the classification results for the proposed method and discusses its relevance. Section IV concludes the study and highlights important research prospects of the current approach.

II. Methodology

A. Data

The data used in this study are obtained from the Alzheimer's Disease Neuroscience Initiative (ADNI) database. This is an open repository containing datasets of AD patients recorded using various imaging modalities. The dataset comprising DTI images from the second phase of

Ravindra B. Patil is Research Scholar in Biomedical Engineering, Department of Applied Mechanics at Indian Institute of Technology Madras, Chennai, India. (e-mail: ravipatil3354@ gmail.com).

Ranjan Piyush is Research Scholar in Biomedical Engineering, Department of Applied Mechanics at Indian Institute of Technology Madras, Chennai, India. (e-mail: am12s024@smail.iitm.ac.in).

S. Ramakrishnan is Associate Professor with Biomedical Engineering Group, Department of Applied Mechanics at Indian Institute of Technology Madras, Chennai, India.(e-mail: sramki@iitm.ac.in).

ADNI2 and ADNIGO protocols are used for the present study. A total of 92 patient data is studied (58-Normal and 34-AD).

B. DTI Maps

DTI is a representation of time varying effects of diffusion of water molecules across the brain tissues. Each voxel in this image has components in all spatial directions which vary in magnitude. The magnitude of components is defined by the size of tissues present and their directionality. Maximum diffusion of water molecules occur along the direction of fiber tract and minimum perpendicular to it [2].

A diffusion tensor is modeled at each voxel [3] in the brain. Diffusion Tensor matrix obtained at each voxel is further decomposed into eigen values which are considered for calculation of anisotropy index. Scalar anisotropy maps are obtained from the resultant diffusion tensor eigen values (λ_1, λ_2 and λ_3).

Fractional anisotropy is the estimate of eccentricity of mapped voxels and is calculated as

$$\mathrm{FA} = \frac{3}{2} * \frac{\sqrt{(\lambda_1-\lambda_m)^2+(\lambda_2-\lambda_m)^2+(\lambda_3-\lambda_m)^2}}{\sqrt{\lambda_1{}^2+\lambda_2{}^2+\lambda_3{}^2}} \quad (1)$$

where, λ_m is the average of the trace of diffusion tensor matrix given by

$$\lambda_m = \frac{\lambda_1+\lambda_2+\lambda_3}{3}. \quad (2)$$

C. ROI Measures

Region of Interest (ROI) masks for various regions are applied to generate FA values. Since the images in database were acquired on a variety of machines with different plane of reference, alignment of all images to the same co-ordinate space is carried out before feature extraction. JHU DTI atlas [4] is used for initial registration of the images; subsequently the JHU "Eve" WM atlas labels are applied. A total of 115 ROIs are identified for JHU DTI for each set of patient data and reduced to 50 ROIs on application of JHU "Eve" WM atlas. The average value of FA in each of these 50 ROIs for all sets of the patient data is extracted. These 50 regions are sufficient to identify the white matter regions based on the atlas used.

D. Feature Extraction and Reduction

FA values obtained from 50 regions of White matter using DTI maps forms the feature vector for each subject. Genetic Algorithm (GA) a dimension reduction technique is used in order to remove the redundant features and extract the prominent features which best describe the AD pathology.

GA is influenced by natural selection and evolutionary growth [11] extensively explained by Charles Darwin. This optimization technique generates the best subset of available features. Binary encoding is used for the entire feature space which renders the preferred features as 1 and redundant features as 0. A random set of solutions is generated as the initial population.

Randomly selected solutions are tested for fitness. The preferred features obtained as a result of selection is then modified [12]. Further this process is iterated until an optimal solution is derived from the entire population. Each of the features is assigned a rank based on the predictive weight associated to each feature. Predictive weight for each feature vector is computed as follows[15].

Let i be the individual subject from the set of samples S of the form $f(x_1, x_2, x_3 \ldots x_n)$ consisting of n feature vectors .The classifier output corresponding to the subject i is given by

$$classifier_i(x_1, x_2, \ldots\ldots, x_n) = f(x) = \begin{cases} s_A, & f_i(x_1, x_2, \ldots\ldots, x_n) \le 0 \\ s_B, & fi\,(x_1, x_2, \ldots\ldots, x_n) > 0 \end{cases}$$

Where S_A , S_B are the two classes of the samples space S.The fitness of each subject is calculated based on the classification accuracy obtained by the below formula

$$fit_i = \frac{TN+TP}{FN+FP+TP+TN} \quad (3)$$

where TP is the number of subjects from s_A that are correctly classified as that class, TN is the number of subjects from s_B that are correctly classified as that class, FN is the number of subjects from s_A that are incorrectly classified as the class s_B and FP is the number of subjects from s_B that are incorrectly classified as class s_A.

Let $w_1, w_2, w_3 \ldots w_n$ be the weights of the feature vector

Initially all the features are considered equally important and the weight of the feature is assigned as

$$w_i(t=0) \leftarrow \frac{1}{n} \quad (4)$$

The weight of the i^{th} feature after t iterations is given by

$$w_i(t) \leftarrow w_i(t-1) + \sum_{j=1}^{N} \frac{count\ (j,i) fit_j}{\sum_{k=1}^{n} count\ (j,k)} \quad (5)$$

where fit_j is the fitness of the j^{th} program in the population, and $count(j,\ k)$ is the number of times the k^{th} feature appears in the j^{th} program in the population, and N is total number of subjects in sample space S.

The top ranked features representing the specific region of WM based on the weights are further fed to the classifier to ascertain the diagnostic relevance.

E. Classification Method

Adaptive Boosting (AdaBoost) is a classification algorithm based on the enhancement technique [13]. The performance of weak classifiers can be improved by repeated iteration using boosting methods. Each iteration aims at increasing the weights of misclassified data and decreasing that of classified data. This results in increased optimization of the classifier for a given dataset. The reduced feature vectors obtained after applying GA on the dataset are used as features for this classifier.

The algorithmic steps involved are as follows:

- Inputs given: $S = \{(x_1, y_1), \ldots\ldots, (x_N, y_N)\}$
- Initialize: Equal weights were assigned to all inputs:

$$d_n^{(1)} = \frac{1}{N} \text{ for all } n = 1, \ldots, N$$

- Loop for t =1,....., T iterations,

a) The classifier is trained with respect to the weighted sample set $\{S, d^{(t)}\}$ and hypothesis $h_t: X \to \{-1, +1\}$ is obtained. The weighted training error (ϵ_t) for h_t is calculated as,

$$\epsilon_t = \sum_{n=1}^{N} d_n^{(t)} I(y_n \neq h_t(x_n)) \quad (6)$$

A factor α is defined for updating the weights of each training data

b) Assign:

$$\alpha_t = \frac{1}{2} \log\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \quad (7)$$

c) Update weights:

$$d_n^{(t+1)} = \frac{d_n^{(t)} \exp(-\alpha_n y_n h_t(x_n))}{z_t} \quad (8)$$

The weight is normalized on each iteration by a factor Z_t such that $\sum_{n=1}^{N} d_n^{(t+1)} = 1$.

- Termination: The iteration ends if $\epsilon_t = 0$ or $\epsilon_t \geq \frac{1}{2}$ and set $T = t - 1$.
- Output :

$$f_T(x) = \sum_{t=1}^{T} \frac{\alpha_t}{\sum_{r=1}^{T} \alpha_r} h_t(x) \quad (9)$$

Hence, the final output $f_T(x)$ is the resultant of majority votes from the T weak assumption with α_t as the weight assigned to h_t.

The basic classifier used in this study is the tree. The learning rate for updating weights in each iteration is set at 0.01 after several trials, as this provided best convergence. On iterative update of the weights for different sets of training data the tree grew more robust.

III. Results

Analysis of prominent features representing the WM regions, which are obtained after applying GA on the data set are further performed. Predictive weights for each of the features contributing to the classification accuracy are calculated. Top 10 regions of WM based on the predictive weights are shown in Fig. 1. From Fig. 1 it can be inferred that Fornix which connects hippocampus to hypothalamus is the most affected region in AD subjects as this feature carries the maximum predictive weight.

Other prominent regions affected are Genu of Corpus Callosum Left, Cingulum, Body of Corpus Callosum, Cerebral Peduncle and Anterior Corona Radiate. These results are consistent with the atlas based study on white matter done by X.Fan et.al [14].

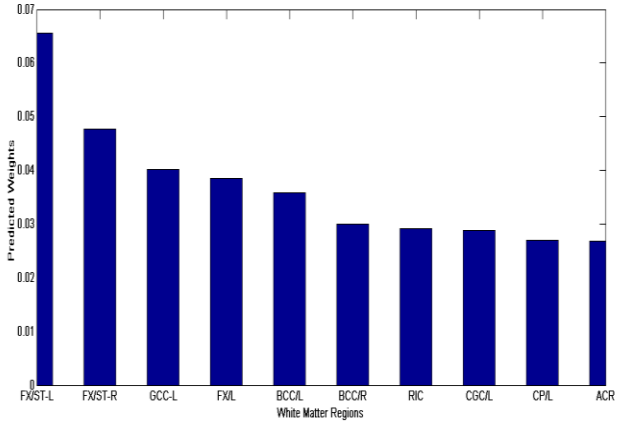

Figure 1. Bar plot of top 10 different white matter regions along with their predicted weights.

Abbreviations of white matter tracts are as below

FX/ST-L:Fornix/Stria Terminalis Left
FX/ST-R:Fornix/Stria Terminalis Right
GCC-L :Genu of Corpus Callosum Left
FX/L :Fornix Left
BCC/L :Body of Corpus Callosum Left
BCC/R :Body of Corpus Callosum Right
RIC :Retrolenticular part of Internal Capsule
CGC/L :Cingulum Left
CP/L :Cerebral Peduncle Left
ACR :Anterior Corona Radiata

The coloured FA map of the white matter depicting the prominent regions affected by AD based on the predictive weights are shown in Figure 2.

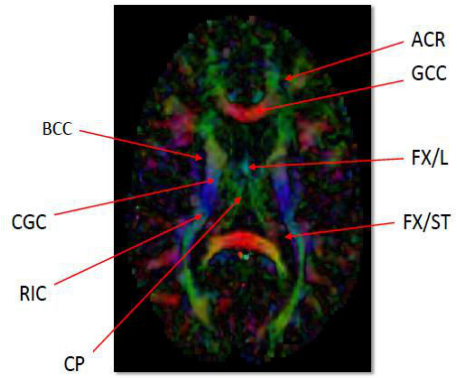

Figure 2. Axial view of FA map of human brain

Another part of this study is to classify AD from NC by analyzing the characteristic changes in the diffusion indices of white matter using AdaBoost classifier. A 10-fold cross validation technique is used to arrive at the classifier performance for selected features obtained after applying GA. Figure 3 shows the variation of the classifier accuracy along the number of selected features. The maximum classification accuracy of 84.5% is achieved with 10 selected features. Classification accuracy considering all the features

is 75.3%. Classification accuracy considering only tree classifier without boosting hovered around 62.34%, indicating that boosting mechanism aids in improving the classification accuracy by 20%. The sensitivity and specificity of different approaches used are tabulated in Table 1. The reduction in the classification accuracy when all the features are considered is due to over fitting of the data by the classifier. Thus feature reduction aids in improving the classification accuracy by eliminating the redundant or less useful features from the dataset. These results also demonstrate that only 8-12 prominent features which describe the AD pathology most are sufficient for the auto diagnosis of AD. Furthermore, considering all the features for the classification reduces the diagnosis of AD. Other DTI indices namely Mean Diffusivity (MD) was also considered for this study however there was no substantial change in the classification accuracy by inclusion of MD along with FA. Thus FA parameter is sufficient enough to provide the classification of AD from NC.

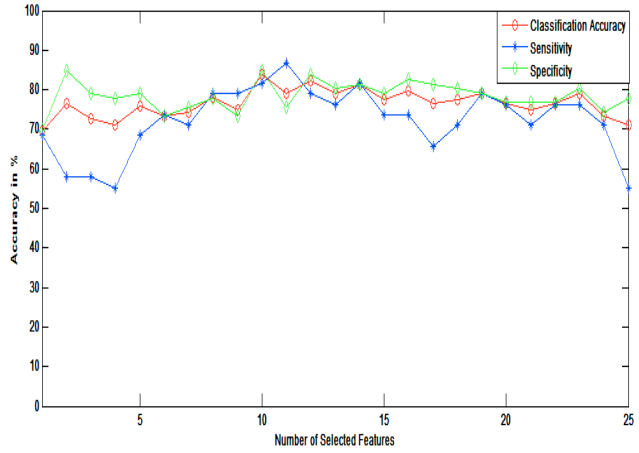

Figure 3 .Plot of accuracy with varied number of selected features

TABLE I
PERFORMANCE OF CLASSIFER WITH DIFFERENT APPROACHES USING BOOSTING

Approach	Accuracy	Sensitivity	Specificity
Top 10 Features	84.5%	80.2%	85.2%
All Features	75.3%	71%	76.7%

IV. CONCLUSION

In this study an attempt is made to identify the prominent regions in WM which gets affected most by AD, using GA feature reduction technique and AdaBoost classifier. With this approach it is possible to concentrate on the specific regions for feature extraction rather than considering entire WM. GA reduction techniques shows that Fornix region of WM is most affected due to AD. AdaBoost classifier was used for the auto diagnosis of the AD. Combination of GA and AdaBoost classifier along with diffusion indices feature, FA extracted from DTI images is able to classify AD and NC with the accuracy of 84.5%.

This approach when used for mass screening of AD reduces the memory and time complexity of the algorithm due to feature extraction from the specified regions. The findings in this study are consistent with the literature in identification of most affected regions in WM due to AD. Furthermore study can be carried out to find better feature reduction techniques and high end classifiers to improve the classification accuracy for diagnosis of Alzheimer.

REFERENCES

[1] Ron,Brookmeyer and Elizabeth "Forecasting the global burden of Alzheimer's disease," The Journal of the Alzheimer's Association Volume 3, Issue 3 , pp 186-191, July 2007.

[2] Beaulieu C,"The Biological Basis of Diffusion Anisotropy". Diffusion MRI.San Diego: Academic Press. pp 105–126, May 2011.

[3] Andrew L. Alexander, Jee Eun Lee, Mariana Lazar and Aaron S. Field , Diffusion Tensor Imaging of the Brain, The Journal of the American Society for Experimental NeuroTherapeutics, volume 4, issue 3, pp 316-329, 2007.

[4] Braak H and Braak E. "Evolution of neuronal changes in the course of Alzheimer's disease,". J Neural Transm Suppl, pp. 127–140, 1998.

[5] Susumu Mori and Jiangyang Zhang , "Principles of Diffusion Tensor Imaging and Its Applications to Basic Neuroscience Research", Neuron 51, pp 527–539, 2006.

[6] Kenichi Oishi, Andreia Faria, "Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: Application to normal elderly and Alzheimer's disease participants", Neuroimage, volume 46, issue 2, pp 486–499, 2009.

[7] O'Dwyer L, Lamberton F, Bokde ALW, Ewers M, Faluyi YO, et al, "Using Support Vector Machines with Multiple Indices of Diffusion for Automated Classification of Mild Cognitive Impairment". PLoS ONE, volume 7, issue 2, e32441. May 2012.

[8] M. Granaa, M. Termenona, A. Savioa, A et.al ," Computer Aided Diagnosis system for Alzheimer Disease using brain Diffusion Tensor Imaging features selected by Pearson's correlation" Neuroscience Letters 502 ,pp 225– 229, 2011.

[9] Desikan RS, Cabral HJ, Hess CP, Dillon WP, Glastonbury CM, et al. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. Brain 132, pp. 2048–2057, 2009.

[10] White T, Kendi AT, Lehéricy S, Kendi M, Karatekin C, Guimaraes A, Davenport N, Schulz SC, Lim KO. "Disruption of hippocampal connectivity in children and adolescents with schizophrenia--a voxel-based diffusion tensor imaging study", Schizophr. Res. 90(1-3), pp 302-307, 2007.

[11] A. S. Fraser, "Simulation of genetic systems by automatic digital computers—I: Introduction," Australian J. Biol. Sci., vol. 10, pp. 484–491, 1957.

[12] D. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning". Reading, MA: Addison-Wesley, 1989.

[13] Yoav Freund Robert E. Schapire ,"A Short Introduction to Boosting Journal of Japanese Society for Artificial Intelligence", 14(5):771-780, Sept 1999.

[14] X. Fan, G. Xiao, K. Martin-Cook, R. Rosenberg, M. Weiner, and H. Huang, " Atlas-based approach to study white matter disruption in Alzheimer's disease", ISMRM Joint Annual Meeting, Oct 2010.

[15] Friedlander A, Neshatian K and Zhang M, "Meta-Learning and feature Ranking Using Genetic Programming for Classification: Variable Terminal Weighting", Evolutionary Computation (CEC), IEEE Congress : 941-948, June 2011.