Fuzzy Diagnostics of AD Subject Using Whole Brain Atrophy Patterns

Ali Farzan, M.B Moradi Gheshlagh, Mohammad Arjmand Khajeh and Farshad Arvin

Abstract - To assess capability of k-mean and fuzzy clustering method (FCM) in diagnosing Alzheimer's disease (AD) subjects using longitudinal whole brain atrophy percentage calculated via Magnetic Resonance Images (MRI). We included 60 subjects (30 AD and 30 controls) in this study and measured their whole brain atrophy percentage. Discriminative power of this measure is statistically analyzed and it is used as a feature in classifying subject using k-mean and FCM. It is revealed that FCM has higher specificity besides k-mean but the same sensitivity. Accuracies of both systems are remarkable. Pattern recognition algorithms even by unsupervised learning methods can help us to diagnose AD with considerable accuracy.

Keywords: **FCM; k-mean; Magnetic Resonance Image (MRI); Alzheimer's' disease; Diagnostic**

I. INTRODUCTION

Alzheimer's disease (AD) is known as the commonest form of dementia in subjects over 65 years old which has been influenced about 26 millions worldwide [1]. It starts with abnormal excessive agglomeration of amyloid β (Aβ) protein and then hyperphosphorylated tau in the brain. This leads to deterioration of the synopsis and axons in neurons. Consequently degeneration of brain structures happens and then memory lapses appear followed by functional and lingual decline. These changes always start in the same order but they may overlap each other in various clinical disease stages [2]. These orders and overlaps are revealed in Fig. 1.

Clinical measures for diagnosing AD are traditionaly based on two rightmost markers and some standard measures such as Mini Mental Score Exam (MMSE) or Clinical Dementia Rating (CDR) are used for clinically diagnosing people with AD. It is obvious that these measures are useful just in second and third stages of disease and cannot be used in first stage which there is not any manifest behavioral or memory impairment [3-4].

Manuscript received May 4, 2011; revised May 19, 2011. This work was supported by the Islamic Azad University, Shabestar Branch, Iran under Grant 51954880230008.

Ali Farzan is with the Islamic Azad University, Shabestar Branch and University of Putra Malaysia (Corresponding author to provide phone: +601-76737629; e-mail: alifarzanam@gmail.com).

M.B Moradi Gheshlagh is with the Islamic Azad University, Shabestar Branch (email: mb.moradi@gmail.com).

Mohammad Arjmand Khajeh is with the Islamic Azad University, Shabestar Branch (email: mhmd.arman@gmail.com).

Farshad Arvin is with the Islamic Azad University, Shabestar Branch (email: farshadarvin@yahoo.com).

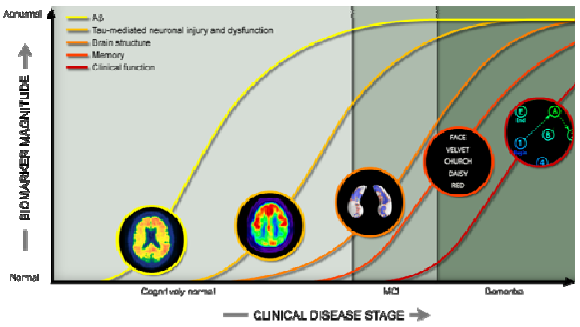

Fig.1. Various biomarkers of AD and the stage of disease they are affective. The first three biomarkers can be used to prognose AD prior to dementia diagnosis.

Furthermore, these scores alone are not accurate enough and some complementary biomarkers are needed for accurate diagnosis of AD [5]. The need for monitoring disease progression in designing new therapeutic trials encourages researchers to find noninvasive accurate biomarkers of AD [6-7]. MR images because of their high resolution and noninvasive nature, are good candidates for realising degeneration of brain structures and finding strong relationships between them and disease progression [8]. Various antomical structures of brain such as Entorhinal Cortex, Hippocampus, Cerebral Cortex and etc have been influenced of AD and their atrophic characteristics such as volume, shape and thickness can be used as biomarkers of AD [8-11]. Concentrating on atrophic characteristics of anatomical structures is prone to some imperfection. That is, disease related atrophies don't necessarily follow the anatomical boundaries of structures and any part of brain can be changed under influence of disease. The rate of whole brain volume change is almost constant in third satge of disease and this makes it usefull in monitoring the pharmacotherapeutic trials [10, 12-13]. Fig. 2 shows the profile of structural changes in AD.

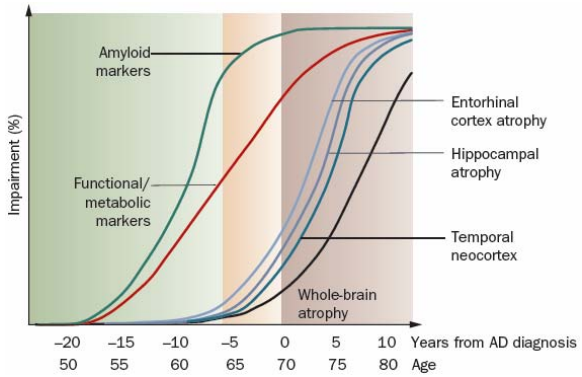

Fig. 2. Natural progression of cognitive and biological markers of Alzheimer disease: a theoretical model.

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

30 subjects with AD and 30 NCs from ADNI image database are used in this study. This paper uses the pipeline implemented by smith to measure the longitudinal volume changes between baseline and 2 years follow ups [13]. Fuzzy C-mean (FCM) is used as a clissifier to classify subjects into the AD and NC groups.

II. MATERIALS AND METHODS

A. Subjects

A total of 30 AD patients (46.7% female; 75 (7) years), and 30 age-matched healthy normal controls (50% female; 77 (5) years) were selected from the Alzheimer's disease NeuroImaging (ADNI) public database (http://www.loni.ucla.edu/ADNI/Data/). ADNI is a large five-year study launched in 2004 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and nonprofit organizations, as a $60 million public–private partnership. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessments acquired at multiple sites (as in a typical clinical trial), can replicate results from smaller single site studies measuring the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to monitor the effectiveness of new treatments, and lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W.Weiner, M.D., VA Medical Center and University of California, San Francisco.

All the AD and NC subjects in this study had successfully undergone magnetic resonance imaging (MRI), cognitive tests and clinical evaluation at baseline and 2 years follow up.

B. Statistical analysis

Some demographic parameters such as age, sex, years of education and etc have remarkable impact in brain atrophic measures and to avoid their influence in the study, subjects of two groups must be matched regarding them. Difference in gender among the two groups was tested with the Chi-square test. There were no significant differences in gender ($p = 0.796$). Independent two sample student t-test was used to test inter-group differences in age and years of education. As there were no significant differences in age ($p = 0.188$) and years of education ($p = 0.554$) among the two groups, they were ignored in diagnosing AD in this study. The MMSE score and also percentage of brain volume change show significant differences in between two groups (unpaired student t-test, $p < 0.00001$ and $p < 0.00001$ respectively). The former is one of the clinical biomarkers of AD and the aim of this study is to evaluate the discriminant power of the later in diagnosing AD and classify subjects based on this measure using fuzzy C-mean (FCM). Statistical analysis of demographic and clinical data is revealed in Table 1.

Table 1
Demographic and clinical variables by diagnostic group

	NC(n=30)	AD(n=30)	p
Gender(M/F)	15/15[a]	16/14	0.796
Age(M/SD)	77/5[a]	75/7	0.188
Years of Education(M/SD)	16.2/2.9[a]	15.7/2.7	0.554
Baseline MMSE(M/SD)	29.3/0.8[b]	23.5/2.2	<0.00001
PBVC(M/SD)	-1.65/1.05[b]	-4.13/1.85	<0.00001

Chi-square was used for gender comparison.
Unpaired student t-test was used for age, education-year, MMSE scores and percentage of whole brain volume change comparisons.
[a] Indicates insignificant compared to NC group.
[b] Indicates significant compared to NC group.

C. Whole brain atrophy rate

First step in this pipeline is brain surface extraction which separates brain from other non-brain parts such as skull or scalp in both images of longitudinal study. To do this, a deformable tessellated mesh have been used which deforms under control of local parameters and finally matches the brain of head [14]. Afterward, base brain images must be registered to follow up counterparts. In this step, it is necessary to avoid rescaling artifacts which can change the atrophy size. With this in mind, it is supposed that the size of skull is constant and it is the normalization factor in scaling process. To escape unnecessary modifications of nonlinear registration which matches images as much as possible and so eliminates the atrophic differences between them, the linear registration is preferred in this study [15].

Now it is time to measure the differences between images. Thus, brain images are segmented into its three major tissues – Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) [16]. Boundary points of these tissues are used to measure differences between images. One 3 by 3 gradient operator is used to find the gradients in these points. In a peer to peer comparison of 3^{mm} intensity profile on these gradients the shift distance that maximizes the correlation between these profiles are considered as difference measure. Normalized sum of these measures over all boundary points indicates the overall differences between brain volumes and is called percentage of brain volume change (PBVC) [17]. This measure is used as a feature for discriminating AD from NCs. To evaluate the discriminative power of it, the unpaired student t-test is performed on it (Table 1). Results imply the discriminative power of this feature. That is, two different populations of AD and NCs have statistically significant differences in the average volume changes.

Until now, it is approved that our two groups are separable based on longitudinal volume changes. But there is not any way of classifying one individual subject into one of these groups. To this end, it is necessary to design a pattern classifier to classify patterns of subjects into two

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

separate groups, one for AD and the other for NC. Here in this study, K-mean and its fuzzy counterpart FCM were used to classify subjects.

D. K-mean algorithm

Suppose $X = \{x_i \in \mathbb{R}^p | i = 1 \dots n\}$ denotes the set of n observations of p–dimensional patterns and the goal is to classify those n observations into $K < n$ classes, $C_k \quad 1 < k < K$. A classifying rule (denoted by R is a many-to-one mapping, $R(x_i) = C_k \quad (1 < i < n, 1 < k < K)$. The K-means algorithm aims to minimize the overall objective function,

$$\sum_{k=1}^{K} \sum_{R(x_i)=C_k} \|x_i - \mu_k\|^2$$

with respect to the classification rule R, where μ_k is the means (or centroid) of patterns from cluster k. This paper, assumes $k = 2$ and therefore there is only two cluster means, μ_1 and μ_2. Beginning with an initial assignment of observations to clusters or an initial assignment of cluster means, the K-means algorithm iterates through the following two steps:

Step 1: Reassign each observation to the cluster whose mean is closest to that observation.

$$R(x_i) = C_k \quad \leftrightarrow \quad k = arg\ \min_l \|x_i - \mu_l\|^2$$

Step 2: Recalculate the new cluster means.

The convergence is reached if the cluster means do not change. An observation x is therefore classified to C_1 with cluster mean μ_1 if and only if

$$\|x_i - \mu_1\|^2 < \|x_i - \mu_2\|^2$$

E. FCM algorithm

FCM clustering technique imparts a degree of fuzziness to each data point corresponding to every cluster. The degree of fuzziness is represented by membership value $\mu_{ik} \in (0,1)$. μ_{ik} represents the membership value of $i'th$ pattern to $k'th$ class. Larger value of μ_{ik} implies greater proximity of the $i'th$ pattern to the center of $k'th$ class. The aim of the algorithm is to find well-defined membership value for every pattern. This is done by iteratively minimizing a modified objective function:

$$\sum_{k=1}^{K} \sum_{i=1}^{n} {\mu_{ik}}^m |x_i - w_k|$$

m is called fuzzification parameter. It controls the noise sensitivity and the extent of the effect of μ_{ik} in the computation of cluster centers. $m \in (1.5,2.5)$ has been found to be the optimal range for this parameter [19]. Correspondingly, $m = 2$ has been taken as it lies in the middle of the optimal range.

To determine objective function, the matrix $U = [\mu_{ik}]$ and the vector $W = [w_k]$ are determined, which are given as

$$\mu_{ik} = \frac{\left(\frac{1}{{d_{ik}}^2}\right)^{1/(m-1)}}{\sum_{k=1}^{K}\left(\frac{1}{{d_{ik}}^2}\right)^{1/(m-1)}}$$

$$w_k = \frac{\sum_{i=1}^{n} \mu_{ik}^m x_i}{\sum_{i=1}^{n} \mu_{ik}^m}$$

where, d_{ik} is the distance between $i'th$ and $j'th$ cluster center.

The matrix U and vector W corresponding to minimized objective function represents the final classification of pixels and cluster centers.

III. Results and discussion

Two different classifiers with unsupervised learning methods are used in this study. Classification results impart their capacity in diagnosing AD.

A. Classification using k-mean

Pbvc as a feature is used to classify subjects into two classes, AD and NC. Discriminative power of this measure is tested using unpaired student t-test and its significance level is presented in Table 1. Classification results based on sensitivity and specificity are summarized in Table 2.

B. Classification using FCM

To compare the power of FCM with classic k-mean, patterns of subjects are classified using FCM. Results imply the higher sensitivity and specificity of FCM with respect to c-mean. Results are revealed in table 2.

Table 2
Classification Results of k-mean and FCM

	Sensitivity	Specificity	*Accuracy*
K-mean	73.3%	93.3%	83.3%
FCM	80%	93.3%	86.67%

Both classifiers have the same specificity, but FCM shows higher sensitivity and hence higher accuracy besides k-mean.

Herein, in order to more robust evaluation of the relative prediction performance of two classifiers, Receiver Operating Characteristics (ROC) curve is also adopted beside the traditional comparison of relative error. This analysis is held to provide more robust comparative evaluation of expected performance on target data than simple comparison of error, which assumes the observed class distribution and does not reflect any differences in the cost of different types of error. ROC curve for k-mean and FCM classifiers are illustrated in Fig. 3.

IV. Conclusion

According to the previous findings, pbvc measure has the power of discriminating populations of AD from controls. This study revealed the usefulness of it in diagnosing individual subjects using various classifiers with remarkable accuracy even in unsupervised mode.

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

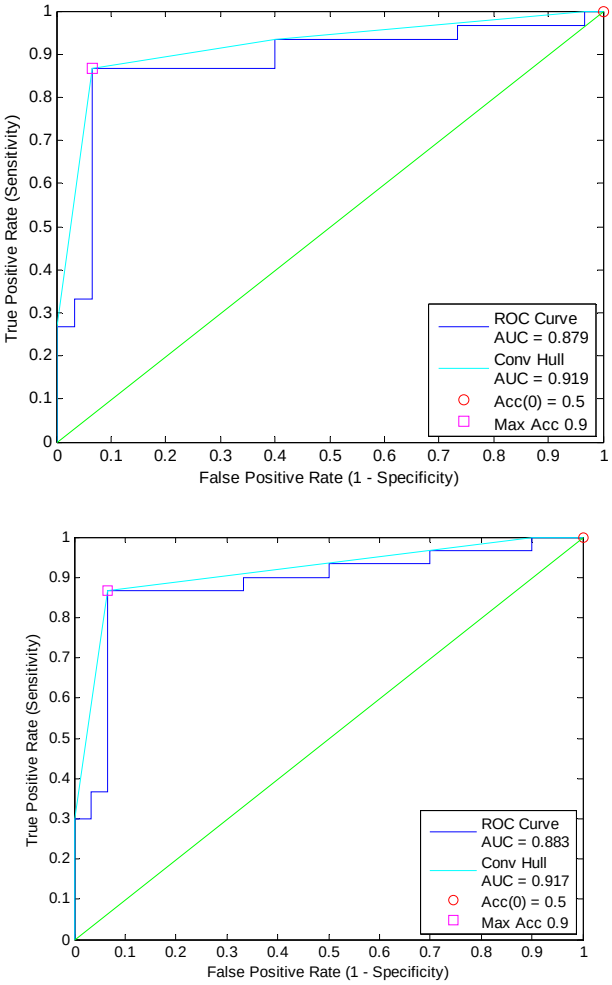

Fig. 3. ROC curve of k-mean (left) and FCM (right). Maximum accuracy of both classifiers are the same but the total area under the curve (AUC) of FOC is higher than k-mean.

REFERENCES

[1] Hua X, Leow AD, Lee S, Klunder AD, Toga AW, Lepore N, et al. 3D characterization of brain atrophy in Alzheimer's disease and mild cognitive impairment using tensor-based morphometry. Neuroimage. 2008;41(1):19-34.

[2] Frisoni G, Fox N, Jack C, Scheltens P, Thompson P. The clinical use of structural MRI in Alzheimer disease. Nature Reviews Neurology. 2010;6(2):67-77.

[3] Ridha B, Anderson V, Barnes J, Boyes R, Price S, Rossor M, et al. Volumetric MRI and cognitive measures in Alzheimer disease. Journal of neurology. 2008;255(4):567-74.

[4] Fox N, Crum W, Scahill R, Stevens J, Janssen J, Rossor M. Imaging of onset and progression of Alzheimer's disease with voxel-compression mapping of serial magnetic resonance images. The Lancet. 2001;358(9277):201-5.

[5] Hua X, Lee S, Yanovsky I, Leow AD, Chou Y-Y, Ho AJ, et al. Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: An ADNI study of 515 subjects. Neuroimage. 2009;48(4):668-81.

[6] Hampel H, Broich K. Enrichment of MCI and early Alzheimer's disease treatment trials using neurochemical & imaging candidate biomarkers. The Journal of Nutrition, Health and Aging. 2009;13(4):373-5.

[7] Visser P, Scheltens P, Verhey F. Do MCI criteria in drug trials accurately identify subjects with predementia Alzheimer's disease? Journal of Neurology, Neurosurgery & Psychiatry. 2005;76(10):1348.

[8] Wang L, Miller JP, Gado MH, McKeel DW, Rothermich M, Miller MI, et al. Abnormalities of hippocampal surface structure in very mild dementia of the Alzheimer type. Neuroimage. 2006;30(1):52-60.

[9] Teipel SJ, Born C, Ewers M, Bokde ALW, Reiser MF, Möller H-J, et al. Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment. Neuroimage. 2007;38(1):13-24.

[10] Plant C, Teipel SJ, Oswald A, Böhm C, Meindl T, Mourao-Miranda J, et al. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. Neuroimage. 2010;50(1):162-74.

[11] Teipel SJ, Ewers M, Wolf S, Jessen F, Kölsch H, Arlt S, et al. Multicentre variability of MRI-based medial temporal lobe volumetry in Alzheimer's disease. Psychiatry Research: Neuroimaging. 2010;182(3):244-50.

[12] Sluimer JD, Bouwman FH, Vrenken H, Blankenstein MA, Barkhof F, van der Flier WM, et al. Whole-brain atrophy rate and CSF biomarker levels in MCI and AD: A longitudinal study. Neurobiology of Aging. 2010;31(5):758-64.

[13] Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. Neuroimage. 2002;17(1):479-89.

[14] Smith S. Fast robust automated brain extraction. Human Brain Mapping. 2002;17(3):143-55.

[15] Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Medical Image Analysis. 2001;5(2):143-56.

[16] Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. Medical Imaging, IEEE Transactions on. 2002;20(1):45-57.

[17] Smith S, De Stefano N, Jenkinson M, Matthews P. SIENA -- Normalised accurate measurement of longitudinal brain change. Neuroimage. 2000;11(5, Supplement 1):S659-S.

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)