Detection of Alzheimer's Disease via Statistical Features from Brain Slices

Namita Aggarwal, Bharti Rana, R. K. Agrawal

School of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi-110067, India
namita_jnu@rediffmail.com, bhartirana.jnu@gmail.com, rkajnu@gmail.com

Abstract

In this study, we propose a model which may assist in diagnosis of Alzheimer's disease (AD) using T1 weighted MRI brain images. The proposed model involves construction of statistical features from multiple trans-axial slices from hippocampus and amygdala regions, which play a significant role in AD diagnosis. Features from multiple slices are then averaged, which resulted into a smaller set of relevant features. The reduced set of features enhances the performance of decision learning system, and takes less memory and computation time. Effectiveness of the proposed model is compared with recent voxel-based-morphometry work in terms of sensitivity, specificity and accuracy. Experimental results on a publicly available MRI dataset showed that the proposed method outperforms the recent voxel-based-morphometry model.

Introduction

Computer-aided image analysis is becoming increasingly important for early diagnosis of Alzheimer's disease (AD), a neurological disorder. Many research works have been proposed for automated classification of AD and controls. Few research works (Maitra and Chatterjee 2006; Chaplot, Patnaik, and Jagannathan 2006; Dahshan, Hosny, and Salem 2010) extracted relevant features from 2D trans-axial brain slices. However, they may have not considered the relevant slices of interest sufficiently. Kloppel et al. (2008) proposed approaches based on gray probability maps of 3D brain volumes. However, it generates a huge size feature vector and thus suffers from curse of dimensionality (Bellman 1961) as available number of samples was small. Savio et al. (2011) constructed a reduced set of features from the voxel clusters detected by automated voxel-based morphometry (VBM) (Ashburner and Friston 2000) on gray matter (GM) segmented volumes using SPM8 (Statistical Parametric Mapping-8) software (http://www.fil.ion.ucl.ac.uk/spm/). However, they pointed out that the mask employed to extract features may not be appropriate as it was created from modulated images. Moreover, images have to be manually reoriented into a right-handed coordinate system. Pre-processing involved in the research work (Savio, et al. 2011) is based on a template image which may not be representing the population appropriately, resulting in possible bias. Above all, they have only considered gray matter tissues while other brain tissues i.e. white matter tissues may also be important in AD diagnosis (Medina, et al. 2006).

In this paper, we propose a method which determines a smaller set of relevant features based on statistical characteristics from multiple slices of brain-extracted volume covering region of interest appropriately. Statistical features were extracted from each of the considered trans-axial 2D slice of a subject, and their averaged values are considered as features. Effectiveness of the proposed approach is investigated and its performance is compared with the research work of Savio et al. (2011).

Feature Extraction Methods

Feature extraction is designed to obtain a meaningful representation of observations and reduce the dimension of the feature vector by removing noisy, irrelevant and redundant features. A small set of relevant features may enhance the performance of decision learning system, and take less memory and computation time.

Voxel-based Morphometry (VBM)

VBM (Ashburner and Friston 2000) compares regional patterns of brain voxel by voxel between two groups of subjects. It spatially normalizes all the training images (3d volumes) into the same standard space. It is then followed

by segmentation of the images into grey matter, white matter, and cerebrospinal fluid. The segmented data can be modulated to correct for volume change that occurred during the spatial normalization. Also, smoothing is performed to correct noise and small variations. Finally a statistical parametric map is obtained by performing voxel-wise parametric statistical tests based on the general linear model.

Savio et al. (2011) obtained statistical parametric map using VBM on smoothed and modulated gray matter tissue probability maps. The research work proposed two feature extraction approaches based on the voxel clusters detected by VBM analysis using SPM8. The feature extraction approaches are as follows: 1) Mean and standard deviation of the GM voxel values of each voxel location cluster were used as features denoted by MSD. 2) A high-dimensional vector with all the GM segmentation values for the voxel locations included in each VBM detected cluster. These features were denoted by VV.

The Proposed Model

In this paper, we propose to construct a smaller set of relevant features based on statistical characteristics from multiple slices covering region of interest appropriately. Hippocampus and amygdala located in medial temporal lobe are considered as region of interest for feature construction as they play an important role in AD diagnosis (Basso et al. 2006). Unlike research works (Maitra and Chatterjee 2006; Chaplot, Patnaik, and Jagannathan 2006; Dahshan, Hosny, and Salem 2010) where slices are considered individually, we propose method that takes into account multiple slices at once. In addition, we considered all brain tissues. Irrelevant tissues external to the brain, such as skull, dura, and eyes were removed using brain extraction tool (BET) (Smith 2002) to enhance the performance of the decision system.

One of dimensionality reduction techniques which provide a minimal set of salient features is based on first order (Papoulis 1991) and second order statistics (Haralick, Shanmugan, and Dinstein 1973). We employed first and second order statistics to extract 14 features from each of the considered trans-axial 2D slice. While 4 features were derived from first order statistics, 10 were constructed from second order statistics. Four first order features used were mean (m_1), variance (μ_2), skewness (μ_3), and kurtosis (μ_4). They are defined as follows:

$$m_1 = E[I] = \sum_{I=0}^{N_g-1} IP(I)$$

$$\mu_k = E\left[\left(I - E[I]\right)^k\right] = \sum_{I=0}^{N_g-1} (I - m_1)^k P(I), \text{ k} = 2, 3, 4$$

where random variable I represents the gray levels of image (trans-axial 2D slice) region, N_g is the number of possible gray levels and $P(I)$ denotes first-order histogram defined as:

$$P(I) = \frac{number\ of\ pixels\ with\ gray\ level\ I}{Total\ number\ of\ pixels\ in\ the\ region}$$

The variance measures deviation of gray levels from the mean. Skewness is a measure of degree of histogram asymmetry around the mean and kurtosis is a measure of the histogram sharpness.

Although first-order statistics based features are translation as well as rotation invariant and capture significant information about gray levels, it do not give any information about the relative positions of the various gray levels within the image. This information can be extracted from the gray-level co-occurrence matrix that measures second-order statistics. It determines how often gray values co-occur at two pixels which are separated by a fixed distance and an orientation. A co-occurrence matrix $P_{d,\theta}$, is a two-dimensional array of size $n \times n$, where n is the number of gray levels in an image. The $(i,j)^{th}$ element of $P_{d,\theta}$ is the probability of transition from a pixel with intensity i to a pixel with intensity j lying at distance d with a given orientation θ in the image.

Using co-occurrence matrix, features can be defined which quantify coarseness, smoothness and texture related information that have high discriminatory power. Among them, angular second moment (ASM), contrast, correlation, homogeneity and entropy are few such commonly used measures which are given by:

$$ASM = \sum_{i,j} P_{d,\theta}(i,j)^2$$

$$Contrast = \sum_{i,j} |i-j|^2 \log P_{d,\theta}(i,j)$$

$$Correlation = \sum_{i,j} \frac{(i-\mu_1)(j-\mu_2)P_{d,\theta}(i,j)}{\sigma_1\sigma_2}$$

$$Homogeneity = \sum_{i,j} \frac{P_{d,\theta}(i,j)}{1+|i-j|^2}$$

$$Entropy = -\sum_{i,j} P_{d,\theta}(i,j) \log P_{d,\theta}(i,j)$$

ASM measures the smoothness of the image. Less smooth the region is, more uniformly distributed is $P_{d,\theta}(i,j)$ and lower will be the value of ASM. Contrast is a measure of local level variations which takes high values for image of high contrast. Correlation is a measure of association between pixels in two different directions. Homogeneity is a measure that takes high values for low-contrast images. Entropy is a measure of randomness and takes low values for smooth images. Together all these features provide high discriminative power to distinguish two different kind of images. Second order statistics based features were built from co-occurrence matrix with d=1 and $\theta = \{0^0, 45^0, 90^0, 135^0\}$. For each of the five second order measures, mean

and range of the resulting values from the four directions were calculated resulting in 10 features.

Even though, only 14 features were extracted from each trans-axial slice, it became large in number when features from multiple slices considered all together. Also, atrophy may not be restricted to one particular slice. Thus, corresponding features from each slice were averaged out resulting into a reduced set of relevant features (14 in number). These averaged features constructed using first and second order statistics are denoted as FSOS.

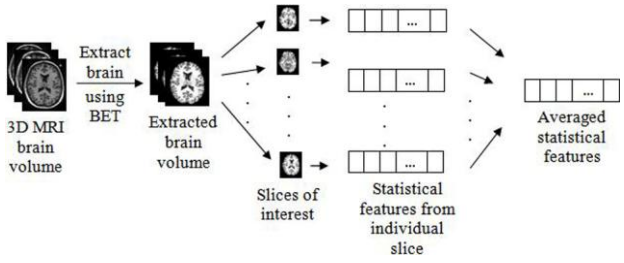


Figure 1 Proposed feature extraction technique

Outline of the proposed feature extraction method is shown in Figure 1 and described as follows. Brain was extracted from each 3D MRI brain volume using BET. A set of 14 features were extracted from each 2D slice that belong to hippocampus and amygdala region. Respective features from each slice were then averaged out resulting in small set of 14 features.

Experimental Setup and Results

Performance of the proposed approach was evaluated on a publicly available MRI data from Open Access Series of Imaging Studies database (Marcus et al. 2007). Here, average registered MRI volumes with corrected bias field were used. Details of data used in the experiment are summarized in Table 1. A global CDR of 0 indicates no dementia, and CDR of 0.5, 1, and 2 represent vey mild, mild and moderate dementia respectively. MMSE represents score of mini-mental state examination.

Performance was evaluated in terms of sensitivity = tp/(tp+fn), specificity = tn/(tn+fp) and accuracy = (tp+tn)/(tp+tn+fp+fn). Here tp, tn, fp and fn denote true positives, true negatives, false positives and false negatives respectively. Four widely used classifiers *i.e*. support vector machine with linear kernel (SVM), C4.5, linear discriminant classifier (LDC) and levenberg–marquardt neural classifier (LMNC) were used. Each experiment was executed 10 times on 10-fold cross-validation. Tools used in the experiment were Image Processing Toolbox from Matlab, SPM8 (http://www.fil.ion.ucl.ac.uk/spm/), BET (Smith 2002) and Prtools (Duin, et al. 2004).

Table 1 Demographic and clinical summaries of AD and controls

	AD	Control
No. of subjects	49	49
Age	78.08 (66-96)	77.77 (65-94)
Education	2.63 (1-5)	2.88 (1-5)
Socioeconomic Status	2.94 (1-5)	2.78 (1-5)
CDR(0.5/1/2)	(31/17/1)	0
MMSE	24.02 (15-30)	28.96 (26-30)

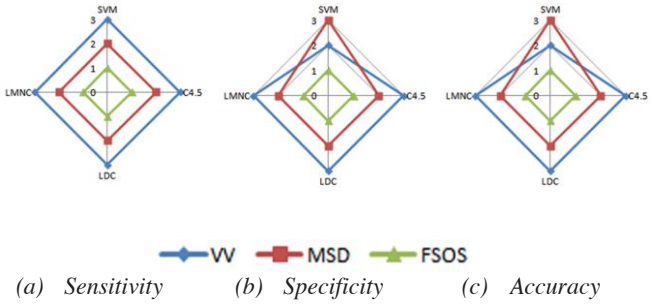
In this experiment the performance of FSOS is compared with VV and MSD techniques (Savio et al. 2011). Average performance measures along with its standard deviation are reported in Table 2. It also includes the performance of baseline approach (BFSOS) which considers all the features together from different slices. The best results achieved for each classifier corresponding to different performance measure is shown in bold.

Table 2 Comparison of performance measures

		Sensitivity		Specificity		Accuracy	
		Mean	Std	Mean	Std	Mean	Std
SVM	BFSOS	65.35	4.53	60.35	4.78	62.78	3.32
	FSOS	**74.40**	**2.78**	**71.80**	**2.51**	**73.08**	**1.73**
	VV	66.05	4.13	70.9	4.40	68.42	3.25
	MSD	67.05	3.63	67.00	2.69	66.97	1.92
C4.5	BFSOS	60.95	3.17	60.35	5.05	60.6	2.3
	FSOS	**63.60**	5.84	**68.10**	6.5	**65.80**	5.41
	VV	62.40	5.36	65.25	5.79	63.76	4.41
	MSD	63.50	6.07	66.7	5.56	64.89	4.63
LDC	BFSOS	-	-	-	-	-	-
	FSOS	**67.25**	**3.08**	**78.00**	**2.09**	**72.62**	**1.92**
	VV	-	-	-	-	-	-
	MSD	63.8	6.22	65.9	3.07	64.71	3.45
LMNC	BFSOS	66.25	4.85	66.15	5.28	66.08	3.53
	FSOS	**68.40**	**4.51**	**64.30**	8.49	**66.28**	**3.91**
	VV	-	-	-	-	-	-
	MSD	57.90	8.07	62.95	4.22	60.32	4.35

BFSOS resulted into lower performance in comparison to FSOS. It may be due to presence of irrelevant features. Moreover, decision model could not be built with LDC classifier. Thus BFSOS is not considered for further comparison. For each classifier, models were ranked based on individual performance measures where lowest rank 1 is given to the best model. Rankings of the feature extraction techniques for different classifiers based on different performance measures are shown in radar charts of Figure 2. We observed the following from Table 2 and Figure 2.

- For all classifiers, FSOS gave maximum average accuracy, sensitivity and specificity in comparison to both VV and MSD. Same can be observed from the radar chart where FSOS is focused more towards centre depicting its best performance.
- FSOS depicts comparatively less variation in the values of all three performance measures (i.e. standard deviation is low) with all classifiers except C4.5.
- Features obtained with VV were large and required huge memory. Hence, the learning model could not be built with LDC and LMNC classifiers.
- For all performance measures and three classifiers viz. LMNC, LDC and C4.5, rank 1, 2 and 3 were consistently achieved by FSOS, MSD and VV respectively.



(a) Sensitivity (b) Specificity (c) Accuracy

Figure 2 Radar charts of ranks obtained by various feature extraction techniques for different performance measures

Conclusion and Future Works

In this paper, we investigated the effectiveness of features based on first and second order statistics to distinguish AD from control. A smaller set of relevant features based on averaged statistical features from multiple slices of brain is extracted from hippocampus and amygdala region which are considered as good markers in AD diagnosis. Experiments were performed on a publicly available MRI brain dataset. Results were compared with VBM based approaches. Unlike considering only gray matter as in VBM, the proposed model considers all the relevant brain tissues and does not require any manual reorientation. For all classifiers, the proposed approach provides better sensitivity, specificity and accuracy in comparison to VBM based techniques. Although the proposed model outperforms the existing VBM based methods, it requires prior knowledge of region of interest (ROI). We plan to further enhance the model in future which will be independent of ROI.

References

Ashburner, J., and Friston, K. J. 2000. Voxel-based morphometry-the methods. *NeuroImage , 11* (6):805–821.

Basso, M., Yang, J., Warren, L., MacAvoy, M. G., Varma, P., Bronen, R. A., et al. 2006. Volumetry of amygdala and hippocampus and memory performance in Alzheimer's disease. *Psychiatry Research: Neuroimaging* , 146 (3): 251-261.

Bellman, R. 1961. *Adaptive control processes: A guided tour*. Princeton University Press.

Chaplot, S., Patnaik, L. M., and Jagannathan, N. R. 2006. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomedical Signal Processing And Control , 1* (1):86–92.

Dahshan, E.-S. A., Hosny, T., and Salem, A.-B. M. 2010. A hybrid technique for automatic MRI brain images classification. *Digital Signal Processing , 20*, 433-441.

Duin, R., Juszcak, P., Paclik, P., Pekalska, E., De Ridder, D., and Tax, D. 2004, January. *PrTools: The Matlab Toolbox for Pattern Recognition*. (Delft University of Technology) Retrieved from http://www.prtools.org

Haralick, R. M., Shanmugan, K., and Dinstein, I. 1973. Textural Features for Image Classification. *IEEE Transactions on Systems: Man, and Cybernetics SMC , 3* (6):610-621.

Kloppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. 2008. Automatic classification of MR scans in Alzheimer's disease. *Brain , 131* (3):681-689.

Maitra, M., and Chatterjee, A. 2006. A Slantlet transform based intelligent system for magnetic resonance brain image classification. *Biomedical Signal Processing and Control , 1* (4):299-306.

Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., and Buckner, R. L. 2007. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience , 19* (9):1498-14507.

Medina, D., DeToledo-Morrell, L., Urresta, F., Gabrieli, J. D., Moseley, M., Fleischman, D., et al. 2006. White matter changes in mild cognitive impairment and AD: A diffusion tensor imaging study. *Neurobiology of Aging , 27* (5):663-672.

Papoulis, A. 1991. *Probability, Random Variables and Stochastic Processes* (3rd ed.). New York: McGraw-Hill.

Savio, A., Garcia-Sebastian, M. T., Chyzyk, D., Hernandez, C., Grana, M., Sistiaga, A., et al. 2011. Neurocognitive disorder detection based on feature vectors extracted from VBM. *Computers in Biology and Medicine , 41* (8):600-610.

Smith, S. M. 2002. Fast robust automated brain extraction. *Human brain mapping , 17* (3):143-155.