On the Use of Morphometry Based Features for Alzheimer's Disease Detection on MRI

Maite García-Sebastián[1], Alexandre Savio[1], Manuel Graña[1], Jorge Villanúa[2]

[1]Grupo de Inteligencia Computacional, www.ehu.es/ccwintco [2]Osatek, Hospital Donostia Paseo Dr. Beguiristain 109, 20014 San Sebastián, Spain.

Abstract. We have studied feature extraction processes for the detection of Alzheimer's disease on brain Magnetic Resonance Imaging (MRI) based on Voxel-based morphometry (VBM). The clusters of voxel locations detected by the VBM were applied to select the voxel intensity values upon which the classification features were computed. We have explored the use of the data from the original MRI volumes and the GM segmentation volumes. In this paper, we apply the Support Vector Machine (SVM) algorithm to perform classification of patients with mild Alzheimer's disease vs. control subjects. The study has been performed on MRI volumes of 98 females, after careful demographic selection from the Open Access Series of Imaging Studies (OASIS) database, which is a large number of subjects compared to current reported studies. [1]

Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder, which is one of the most common cause of dementia in old people. Currently, due to the socio-economic importance of the disease in occidental countries it is one of the most studied. The diagnosis of AD is done after the exclusion of other forms of dementia but definitive diagnosis can only be made after a post-mortem study of the brain tissue. This is one of the reasons why Magnetic Resonance Imaging (MRI) based early diagnosis is a current research goal in the neurosciences.

Morphometry analysis has become a common tool for computational brain anatomy studies. It allows a comprehensive measurement of structural differences within or across groups, not only in specific structures but throughout the entire brain. Voxel-based morphometry (VBM) is a computational approach to neuroanatomy that measures differences in local concentrations of brain tissue, through a voxel-wise comparison of multiple brain images [2]. For instance, VBM has been applied to study volumetric atrophy of the grey matter (GM) in areas of neocortex of AD patients vs. control subjects [4,16,9]. The procedure involves the spatial normalization of subject images into a standard space, segmentation of tissue classes using *a priori* probability maps, smoothing to correct noise and small variations, and voxel-wise statistical tests. Statistical analysis is based on

[1] Research partially supported by Saiotek research projects BRAINER and S-PR07UN02, and the MEC research project DPI2006-15346-C03-03.

the General Linear Model (GLM) to describe the data in terms of experimental and confounding effects, and residual variability. Classical statistical inference is used to test hypotheses that are expressed in terms of GLM estimated regression parameters. The computation of a given contrast provides a Statistical Parametric Map, which is thresholded according to the Random Field Theory.

Machine learning methods have become very popular to classify functional or structural brain images to discriminate them into two classes: normal or a specific neurodegenerative disorder. The Support Vector Machine (SVM) either with linear [10,14] or non-linear [7,11] kernels, have been extensively applied for this task. There are studies applying SVM to discriminate AD patients from controls based on Positron Emission Tomography (PET) or Single-Photon Emission Tomography (SPECT) functional volumes [14,15,1]. There are different ways to extract features from MRI for SVM classification: based on morphometric methods [6,7], based on ROIs (region of interest) [12,11] or GM voxels in automated segmentation images [10]. Work has also been reported on the selection of a small set of the most informative features for classification, such as the SVM-Recursive Feature Elimination [7], the selection based on statistical tests [12,14], the wavelet decomposition of the RAVENS maps [11], among others.

Many of the classification studies on the detection of AD were done over populations mixing men and women. However, it has been demonstrated that brains of women are different from men's to the extent that it is possible to discriminate the gender via MRI analysis [11]. Moreover, it has been shown that VBM is sensitive to the gender differences. For these reasons, we have been very cautious in this study. We have selected a set of 98 MRI women's brain volumes. It must be noted that this is a large number of subjects compared with the other studies referred above.

The approach taken in this paper is to use the clusters detected as result of VBM as a mask on the MRI and Grey Matter (GM) segmentation images to select the potentially most discriminating voxels. Features for classification are either the voxel values or some summary statistics of each cluster. We assume for classification the standard SVM, testing linear and non-linear (RBF) kernels. Section Materials and Methods gives a description of the subjects selected for the study, the image processing, feature extraction details and the classifier system. Section Results gives our classification performance results and section Conclusions gives some conclusions and further work suggestions.

Materials and Methods

Subjects

Ninety eight right-handed women (aged 65-96 yr) were selected from the Open Access Series of Imaging Studies (OASIS) database (http://www.oasis-brains.org) [13]. OASIS data set has a cross-sectional collection of 416 subjects covering the adult life span aged 18 to 96 including individuals with early-stage Alzheimer's Disease. We have ruled out a set of 200 subjects whose demographic, clinical

or derived anatomic volumes information was incomplete. For the present study there are 49 subjects who have been diagnosed with very mild to mild AD and 49 nondemented. A summary of subject demographics and dementia status is shown in table 1.

	Very mild to mild AD	Normal
No. of subjects	49	49
Age	78.08 (66-96)	77.77 (65-94)
Education	2.63 (1-5)	2.87 (1-5)
Socioeconomic status	2.94 (1-5)	2.88 (1-5)
CDR (0.5 / 1 / 2)	31 / 17 / 1	0
MMSE	24 (15-30)	28.96 (26-30)

Table 1. Summary of subject demographics and dementia status. Education codes correspond to the following levels of education: 1 less than high school grad., 2: high school grad., 3: some college, 4: college grad., 5: beyond college. Categories of socioeconomic status: from 1 (biggest status) to 5 (lowest status). MMSE score ranges from 0 (worst) to 30 (best).

Imaging Protocol

Multiple (three or four) high-resolution structural T1-weighted magnetization-prepared rapid gradient echo (MP-RAGE) images were acquired [8] on a 1.5-T Vision scanner (Siemens, Erlangen, Germany) in a single imaging session. Image parameters: TR= 9.7 msec., TE= 4.0 msec., Flip angle= 10, TI= 20 msec., TD= 200 msec., 128 sagittal 1.25 mm slices without gaps and pixels resolution of 256×256 (1×1mm).

Image Processing and VBM

We have used the average MRI volume for each subject, provided in the OASIS data set. These images are already registered and resampled into a 1-mm isotropic image in atlas space and the bias field has been already corrected [13]. The Statistical Parametric Mapping (SPM5) (http://www.fil.ion.ucl.ac.uk/spm/) was used to compute the VBM which gives us the spatial mask to obtain the classification features. Images were reoriented into a right-handed coordinate system to work with SPM5. The tissue segmentation step does not need to perform bias correction. We performed the modulation normalization for grey matter, because we are interested in this tissue for this study. We performed a spatial smoothing before computing the voxel-wise statistics, setting the Full-Width at Half-Maximum (FWHM) of the Gaussian kernel to 10mm isotropic. A GM mask was created from the average of the GM segmentation volumes of the subjects under study. Thresholding the average GM segmentation, we obtain a binary mask that includes all voxels with probability greater than 0.1 in the average

GM segmentation volume. This interpretation is not completely true, since the data are modulated, but it is close enough for the mask to be reasonable. We design the statistical analysis as a Two-sample t-test in which the first group corresponds with AD subjects. We also have done some experiments with nWBV (normalized whole brain volume) as the covariate. The general linear model contrast has been set as [-1 1], a right-tailed (groupN > groupAD), correction FWE and p-value=0.05. The VBM detected clusters are used for the MRI feature extraction for the SVM classification.

Support Vector Machine Classification

The Support Vector Machine (SVM) [17] algorithm used for this study is included in the libSVM (`http://www.csie.ntu.edu.tw/~cjlin/libsvm/`) software package. The implementation is described in detail in [5]. Given training vectors $x_i \in R_n, i = 1, \ldots, l$ of the subject features of the two classes, and a vector $y \in R^l$ such that $y_i \in \{-1, 1\}$ labels each subject with its class, in our case, for example, patients were labeled as -1 and control subject as 1. To construct a classifier, the SVM algorithm solves the following optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$

subject to $y_i(w^T \phi(x_i) + b) \geq (1 - \xi_i),\ \xi_i \geq 0, i = 1, 2, \ldots, n$. The dual optimization problem is

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha$$

subject to $y^T \alpha = 0,\ 0 \leq \alpha_i \leq C,\ i = 1, \ldots, l$. Where e is the vector of all ones, $C > 0$ is the upper bound on the error, Q is an l by l positive semidefinite matrix, $Q_{ij} \equiv y_i y_j K(x_i, x_j)$, and $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is the kernel function that describes the behaviour of the support vectors. Here, the training vectors x_i are mapped into a higher (maybe infinite) dimensional space by the function $\phi(x_i)$. The decision function is $sgn(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b)$.

The chosen kernel function results in different kinds of SVM with different performance levels, and the choice of the appropriate kernel for a specific application is a difficult task. In this study two different kernels were tested: the linear and the radial basis function (RBF) kernel. The linear kernel function is defined as $K(x_i, x_j) = 1 + x_i^T x_j$, this kernel shows good performance for linearly separable data. The RBF kernel is defined as $K(x_i, x_j) = exp(-\frac{||x_i - x_j||^2}{2\sigma^2})$. This kernel is basically suited best to deal with data that have a class-conditional probability distribution function approaching the Gaussian distribution [3]. One of the advantages of the RBF kernel is that given a kernel, the number of support vectors and the support vectors are all automatically obtained as part of the training procedure, i.e., they don't need to be specified by the training mechanism.

Feature Extraction

We have tested three different feature extraction processes, based on the voxel location clusters obtained from the VBM analysis:

1. The first feature extraction process computes the ratio of GM voxels to the total number of voxels of each voxel location cluster
2. The second feature extraction process computes the mean and standard deviation of the GM voxel intensity values of each voxel location cluster
3. The third feature extraction process computes a very high dimensional vector with all the GM segmentation values for the voxel locations included in each VBM detected cluster. The GM segmentation voxel values were ordered in this feature vector according to the coordinate lexicographic order

Classifier Performance Indices

We evaluated the performance of the classifier using the 10-fold cross-validation test. To quantify the results we measured the accuracy, the ratio of the number of test volumes correctly classified to the total of tested volumes. We also quantified the specificity and sensitivity of each test defined as $Specificity = \frac{TP}{TP+FP}$ and $Sensitivity = \frac{TN}{TN+FN}$, where TP is the number of true positives: number of AD patient volumes correctly classified; TN is the number of true negatives: number of control volumes correctly classified; FP is the number of false positives: number of AD patient volumes classified as control; FN is the number of false negatives: number of control volumes classified as patient.

Results

In this section we present for each experiment the following data: the number of features, accuracy, specificity, which is related to AD patients and sensitivity, which is related to control subjects. We have performed the VBM twice, first without any covariate included in the GLM (Table 2) and second taking into account the normalized brain volume (nWBV) (Table 3). Each VBM process produces different sets of voxel location clusters, and, therefore, different sets of feature vectors. The covariate helps to focus the VBM, giving less and smaller clusters than the VBM without covariates. This implies that the feature vectors will be smaller. Each table entry contains the SVM results using the linear and RBF kernels upon the corresponding feature vector set. In both tables rows correspond to feature extraction processes as described in section 1.

The best accuracy result (Table 2) is 80.6% with the RBF kernel, but this result is not far away from the results of the linear SVM. The classification results of table 3, using the covariate nWBV in the design matrix of the GLM, confirm that the non-linear SVM is more accurate. However, as the size of the feature vectors is lower than in table 2, results in table 3 are systematically lower.

Overall the sensitivity results in tables 2 and 3 are much lower than the specificity. We hypothesize that the source of error is the confusion of mild

Feature extracted	Features	Accuracy (lk/nlk)	Sensitivity (lk/nlk)	Specificity (lk/nlk)
GM proportion	12	69.39% / 68.36%	0.88 / 0.90	0.63 / 0.61
Mean & StDev	24	78.57% / 80.61%	0.88 / 0.89	0.72 / 0.75
Voxel intensities	3611	73.47% / 76.53%	0.75 / 0.76	0.72 / 0.77

Table 2. Classification results with a linear kernel (lk) and a non-linear kernel (nlk). No covariates have been taken into account in the GLM used for the VBM. The values of $\gamma = \left(2\sigma^2\right)^{-1}$ for non linear kernel were 0.5, 0.031, 0.0078 for each feature extraction process, respectively.

Feature extracted	Features	Accuracy (lk/nlk)	Sensitivity (lk/nlk)	Specificity (lk/nlk)
GM proportion	2	51% / 51%	1 / 1	0.50 / 0.50
Mean & StDev	4	69.38% / 72.45%	0.79 / 0.79	0.65 / 0.68
Voxel intensities	265	66.32% / 75.51%	0.67 / 0.80	0.65 / 0.72

Table 3. Classification results with a linear kernel (lk) and a non-linear kernel (nlk). The normalized brain volume (nWBV) covariate has been taken into account in the GLM for the VBM. The values of γ for nlk were 0.5, 2.7, 0.004 for GM proportion, Mean & StDev and voxel intensities respectively.

demented AD patients with control subjects. Mild demented AD patients are subjects with CDR=0.5 (Clinical Dementia Ratio) and a high value for the MMSE (Minimental-State Examination), i.e. MMSE=30. Therefore we repeat the feature extraction and classification experiment taking out of the population 9 mild demented AD patients. The results for the RBF kernel SVM are given in table 4. The classification accuracy of the grows from 80.6% (in the best result of table 2) up to 87.5%. Also sensitivity and specificity improve if we compare table 2 and table 3 against table 4.

Feature extracted	Features	γ	Accuracy	Sensitivity	Specificity
GM proportion	12	0.9	72.5%	0.84	0.66
Mean & StDev	24	0.6	87.5%	0.89	0.86
Voxel intensities	3611	1.5	86.25%	0.85	0.87

Table 4. Classification results of 40 AD patients vs. 49 control subjects with the SVM and a RBF kernel, 9 possible outliers were taken out from the AD patients subset.

Conclusions

In this work we have studied feature extraction processes based on VBM analysis, to classify MRI volumes of AD patients and normal subjects. We have analyzed different designs for the SPM of the VBM and we have found that the basic GLM design without covariates can detect subtle changes between AD patients and controls that lead to the construction of SVM classifiers with a discriminative

accuracy of 87.5%. In [6] they compare their results on a smaller population of controls and AD patients to the ones obtained with a standard VBM analysis using only one cluster and found a classification accuracy of 63.3% via cross-validation. Therefore, the results shown in this paper, along with the careful experimental methodology employed, can be of interest for the Neuroscience community researching on the AD. Further work may address the extraction of features based on other morphometric methods, such as Deformation-based Morphometry.

References

1. I. Alvarez, M. Lopez, J.M. Gorriz, J. Ramirez, D. Salas-Gonzalez, F. Segovia, and C.G. Puntonet. Automatic classification system for the diagnosis of alzheimer disease using Component-Based SVM aggregations. In *15th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Assembly (ICONIP 2008)*, 2008.
2. J. Ashburner and K. J. Friston. Voxel-based morphometry: The methods. *Neuroimage*, 11(6):805–821, 2000.
3. Christopher Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):167, 121, 1998.
4. G. F. Busatto, G. E. J. Garrido, O. P. Almeida, C. C. Castro, C. H. P. Camargo, C. G. Cid, C. A. Buchpiguel, S. Furuie, and C. M. Bottino. A voxel-based morphometry study of temporal lobe gray matter reductions in alzheimer's disease. *Neurobiology of Aging*, 24(2):221–231, 2003.
5. Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.
6. C. Davatzikos, Y. Fan, X. Wu, D. Shen, and S. M. Resnick. Detection of prodromal alzheimer's disease via pattern classification of magnetic resonance imaging. *Neurobiology of Aging*, 29(4):514–523, 2008.
7. Yong Fan, Dinggang Shen, and Christos Davatzikos. *Classification of Structural Images via High-Dimensional Image Warping, Robust Feature Extraction, and SVM*, pages 1–8. 2005.
8. A. F. Fotenos, A. Z. Snyder, L. E. Girton, J. C. Morris, and R. L. Buckner. Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology*, 64(6):1032–1039, 2005.
9. G. B. Frisoni, C. Testa, A. Zorzan, F. Sabattoli, A. Beltramello, H. Soininen, and M. P. Laakso. Detection of grey matter loss in mild alzheimer's disease with voxel based morphometry. *Journal of Neurology, Neurosurgery & Psychiatry*, 73(6):657–664, 2002.
10. S. Kloppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack Jr, J. Ashburner, and R. S. J. Frackowiak. Automatic classification of MR scans in alzheimer's disease. *Brain*, 131(3):681, 2008.
11. Z. Lao, D. Shen, Z. Xue, B. Karacali, S. M. Resnick, and C. Davatzikos. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage*, 21(1):46–57, 2004.
12. Y. Liu, L. Teverovskiy, O. Carmichael, R. Kikinis, M. Shenton, C. S. Carter, V. A. Stenger, S. Davis, H. Aizenstein, and J. T. Becker. Discriminative MR image feature analysis for automatic schizophrenia and alzheimer's disease classification. *Lecture Notes in Computer Science*, pages 393–401, 2004.

13. 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*, 19(9):1498–1507, September 2007. PMID: 17714011.
14. J. Ramirez, J.M. Gorriz, M. Lopez, D. Salas-Gonzalez, I. Alvarez, F. Segovia, and C.G. Puntonet. Early detection of the alzheimer disease combining feature selection and kernel machines. In *15th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Assembly (ICONIP 2008)*, 2008.
15. D. Salas-Gonzalez, J.M. Gorriz, J. Ramirez, M. Lopez, I. Alvarez, F. Segovia, and C.G. Puntonet. Computer aided diagnosis of alzheimer disease using support vector machines and classification trees. In *15th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Assembly (ICONIP 2008)*, 2008.
16. R. I. Scahill, J. M. Schott, J. M. Stevens, M. N. Rossor, and N. C. Fox. Mapping the evolution of regional atrophy in alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. *Proceedings of the National Academy of Sciences*, 99(7):4703, 2002.
17. Vladimir N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, September 1998.