Feature extraction from structural MRI images based on VBM: data from OASIS database

Darya Chyzhyk, Alexandre Savio

September 30, 2010

Contents

1 Introduction

This research report gives a description of the feature extraction process that has been performed for the realization of the computational experiments reported in [8, 5, 7] on the classification of structural MRI data into Alzheimer's Disease (AD) patients and controls. The purpose of this publication and the accompanying data sets is to allow the independent validation of our results, and to allow new classification approaches to be tested on this data. The process consists in the realization of Voxel Based Morphometry analysis to detect clusters of voxels which can be discriminat about AD patients and controls. These clusters are used as an spatial mask to extract from the original volumes the values of the corresponding voxels. These raw values constitute the *voxel value* (VV) feature vectors. We also compute the mean and standard deviation of the voxel values inside each of the clusters, these constitute the MSD feature vectors.

In this report we will first recall some medical background on the AD in section 2. Then we specify the population extracted from the OASIS database that has been the subject of the referred publications in section 3. Section 4 recalls the definition of VBM and how it has been applied to the data.

2 Medical background. Alzheimer's Disease

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 socioeconomic importance of the disease in occidental countries it is one of the most studied. The diagnosis of AD can be done after the exclusion of other forms of dementia but a 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 hot topic in the neurosciences. The pharmaceutical companies have already recognized that imaging techniques especially MRI and Positron Emission Tomography (PET) provide "surrogate" information concerning the pattern and rate of neurodegeneration, which can be used to monitor the effects of treatments which slow the progression of neurodegeneration. Therefore, there is high interest in the development of automated detection procedures based on MRI and other medical imaging techniques.

Many of the classification studies on the detection of AD were done with both 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. 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.

3 OASIS subjects and imaging protocol

Ninety eight right-handed women (aged 65-96 yr) were selected from the Open Access Series of Imaging Studies (OASIS) database [6]. 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 non-demented. A summary of subject demographics and dementia status is shown in table 3.

3.1 OASIS Imaging Protocol

The OASIS database has been built following a strict imaging protocol, to avoid variations due to imaging protocol which would pose big image normalization problems. Multiple (three or four) high-resolution structural T1-weighted magnetization-prepared rapid gradient echo (MP-RAGE) images were acquired 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).

	Very mild to mild AD	CS
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).

4 Methods

In this section we will describe the Voxel Based Morphometry (VBM) process, and we give details of its independent application to the AD data. Then we describe the feature extraction processes.

4.1 Voxel-based Morphometry (VBM)

Morphometry analysis has become a common tool for computational brain anatomy studies. It allows a comprehensive measurement of structural differences within a group or across groups, not just 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 [3, 9, 4]. The processing pipeline of VBM is illustrated in figure 1. 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. Smoothing is done by convolution with a Gaussian kernel whose the Full-Width at Half-Maximum (FWHM) is tuned for the problem at hand. Statistical analysis is based on 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. This computation of given contrast provides a Statistical Parametric Map (SPM), which is thresholded according to the Random Field theory.

4.2 Image processing and VBM for OASIS

We have used the average MRI volume for each subject, provided in the OASIS data set. These images are already registered and re-sampled into a 1-mm
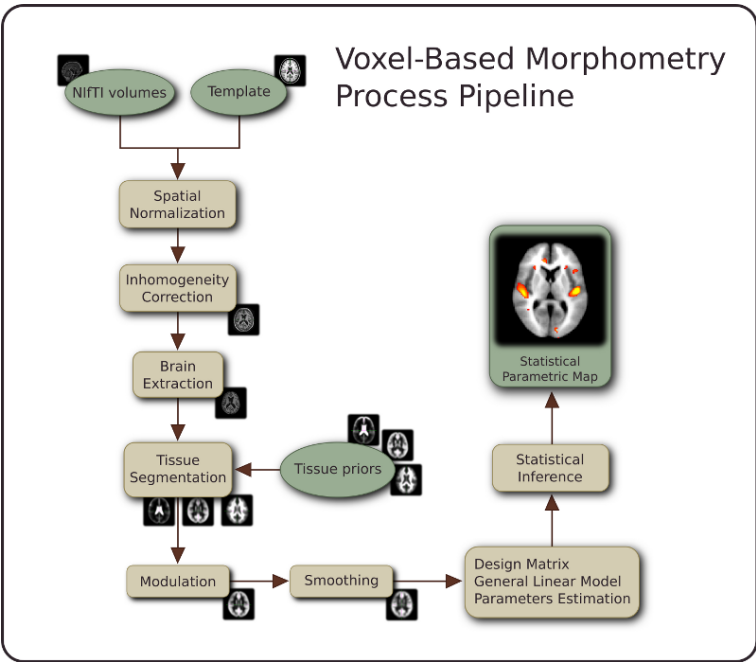
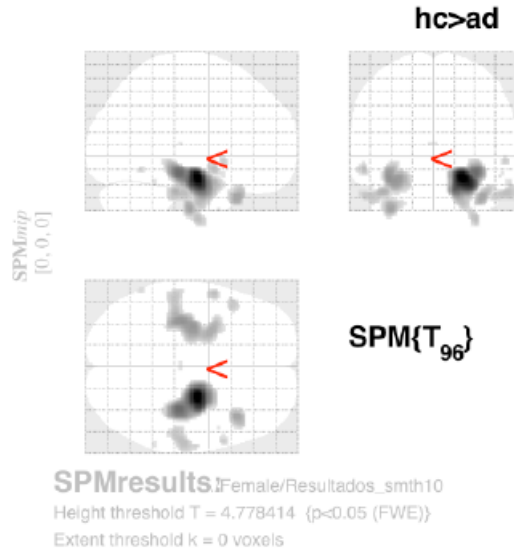Figure 1: The processing pipeline of the Voxel Based Morphometry (VBM) on structural MRI volumes.

Figure 2: SPM results: clusters of significant voxels with increased gray matter density in the controls relative to the patient subjects, detected by the VBM process.

isotropic image in atlas space and the bias field has been already corrected [6]. The Statistical Parametric Mapping software (SPM8) [1] 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 SPM8. The tissue segmentation step does not need to perform bias correction. We performed the modulation normalization for GM, because we are interested in this tissue for this study. We performed a spatial smoothing before performing the voxel-wise statistics, setting the 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 is modulated, but it is close enough for the mask to be reasonable. We designed the statistical analysis as a Two-sample t-test in which the first group corresponds with AD subjects. In SPM software jargon: the contrast has been set to [-1 1], a right-tailed (groupN > groupAD), correction FWE, p-value=0.05. The VBM detected clusters are used for the feature extraction for the classification procedures. Statistical significance was determined using an extent threshold of 0 adjacent voxels for two sample comparisons. The clusters of significant voxels detected by the VBM analysis are displayed in figure 2.
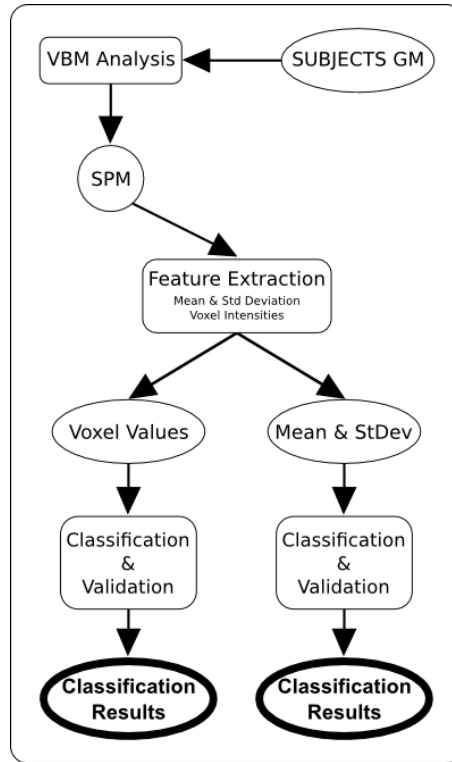
5

Figure 3: Flow diagram of the feature extraction process from the subjects' GM segmentation volumes.

4.3 Feature extraction

We have tested two different feature vector extraction processes, based on the voxel location clusters detected as a result of the VBM analysis. The process is illustrated in figure 3. The VBM detected clusters are used as masks to determine the voxel positions where the features are extracted. These masks are applied to the GM density volumes result of the segmentation step in the VBM analysis.

1. The first feature extraction process computes the mean and standard deviation of the GM voxel values of each voxel location cluster, we denote these features as MSD in the result tables given below.

2. The second 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 voxel values were ordered in this feature vector according to the coordinate lexicographical ordering. We denote these features as VV in the result tables below.

References

[1] http://www.fil.ion.ucl.ac.uk/spm/.

[2] John Ashburner and Karl J. Friston. Voxel-Based Morphometry–The methods. *NeuroImage*, 11(6):805–821, June 2000.

[3] 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.

[4] 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.

[5] M. García-Sebastián, A. Savio, M. Graña, and J. Villanúa. On the use of morphometry based features for alzheimer's disease detection on mri. *Bio-Inspired Systems: Computational and Ambient Intelligence. / IWANN 2009 (Part I) Joan Cabestany, Francisco Sandoval, Alberto Prieto, Juan M. Corchado (Editors), LNCS 5517*, pages 957–964, 2009.

[6] 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.

[7] A. Savio, M. García-Sebastián, M. Graña, and J. Villanúa. Results of an adaboost approach on alzheimer's disease detection on mri. *Bioinspired applications in Artificial and Natural Computation. J. Mira, J. M. FerrÃ¡ndez, J.R. Alvarez, F. dela Paz, F.J. Tolede (Eds.) LNCS 5602*, pages 114–123, 2009.

[8] A. Savio, M. García-Sebastián, C. Hernández, M. Graña, and J. Villanúa. Classification results of artificial neural networks for alzheimer's disease detection. *Intelligent Data Engineering and Automated Learning- IDEAL 2009, Emilio Corchado, Hujun Yin (eds) LNCS 5788*, pages 641–648, 2009.

[9] 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.