Boosting Alzheimer Disease Diagnosis using PET images

Margarida Silveira*, Jorge Marques*
and the Alzheimers Disease Neuroimaging Initiative
*Instituto Superior Técnico
Instituto de Sistemas e Robótica
Lisbon, Portugal
msilveira@isr.ist.utl.pt jsm@isr.ist.utl.pt

***Abstract*—*Alzheimer's disease (AD) final is one of the most frequent type of dementia. Currently there is no cure for AD and early diagnosis is crucial to the development of treatments that can delay the disease progression.*

Brain imaging can be a biomarker for Alzheimer's disease. This has been shown in several works with MR Images, but in the case of functional imaging such as PET, further investigation is still needed to determine their ability to diagnose AD, especially at the early stage of Mild Cognitive Impairment (MCI).

In this paper we study the use of PET images of the ADNI database for the diagnosis of AD and MCI. We adopt a Boosting classification method, a technique based on a mixture of simple classifiers, which performs feature selection concurrently with the segmentation thus is well suited to high dimensional problems. The Boosting classifier achieved an accuracy of 90.97% in the detection of AD and 79.63% in the detection of MCI.

I. Introduction

Alzheimer's disease (AD) is one of the most frequent type of dementia. Most commonly it affects elderly people and therefore it is expected to increase due to the aging of the population. Currently there is no cure for AD and there is a great interest in the development of treatments that can delay its progression, especially if diagnosis is provided at an early stage where those treatments would have the most impact.

Mild Cognitive Impairment (MCI) is the diagnosis given to individuals who have cognitive impairments beyond what is expected for their age and education, but that do not interfere significantly with their daily activities. MCI is considered to be a transitional state between the normal cognitive changes of aging and the earliest clinical manifestations of dementia. It is presently a topic of great interest, particularly the amnestic subtype of MCI which is likely to be a clinical precursor of AD.

Currently, the diagnosis of AD and MCI is based primarily on clinical and neuropsychological assessments. Neuroimaging has also been recognized as a powerful tool to analyze brain changes. However, the analysis of brain images is a difficult task because the spatial pattern of brain degeneration in AD and MCI is highly variable and complex. Consequently, several attempts are being made to develop automated tools that will allow a more sensitive and consistent analysis.

Most of these attempts have focused on the diagnosis of AD from MRI. The diagnosis of MCI from MRI is a more challenging problem which received less attention. In the case of PET images, further investigation is still needed to determine their ability to diagnose AD, especially at the early stage of MCI.

A. State of the Art

The majority of the developed tools has focused on examining the volumetry or shape of a small number of brain structures such as the hippocampus [1], or the gray matter volume [2]. These techniques rely heavily on manual or semi-automatic extraction of the structures of interest, which is by no means a straightforward task. Furthermore, they are limited by the fact that the brain atrophy usually involves many brain regions and different regions are affected at different stages of the disease.

Therefore, current techniques are focusing on the use of the entire brain pattern. However, since a brain volume contains thousands of voxels which represent variables and the number of subjects is generally smaller, this task suffers from the so called 'curse of dimensionality', which results in highly non convex distributions. To handle non convex distributions globally nonlinear classifiers should be used.

Many of the traditional classifiers have been used such as discriminant analysis [3], neural networks [4], Nearest Mean Classifier (NMC) [5] and Fisher linear discriminant (FLD) [6]. Kernel methods, namely Support Vector Machines (SVM), have also been used in [6], [7], [8] with success.

To obviate the 'curse of dimensionality' problem, most methods use data reduction techniques such as Principal Components Analysis (PCA) [7], [9] or Partial Least-Squares (PLS) [3] which are not appropriate to nonlinear and heterogeneous data. In [8] an SVM was used for feature reduction prior to classification with another SVM, which is computationally very expensive. More recently, a variant of Linear Programming Boosting which imposes spatial continuity of the voxels selected by the classifier has been

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

proposed in [10], but MCI was not considered and no comparison was made with the original Boosting classifier.

In this paper, we propose the use of a Boosting classifier, a nonlinear method based on a mixture of classifiers (ensemble method), which is able to automatically select the most important features concurrently with the classification, thus is well suited to high dimensional problems. We apply this classifier to the intensity of FDG-PET brain images from the ADNI database and investigate its ability to differentiate between individuals with Alzheimer's disease (AD), mild cognitive impairment (MCI), and normal control subjects (NC).

The remainder of this paper is organized as follows: section 2 describes the data used and the Boosting classifier, section 3 describes the experimental results and section 4 concludes the paper.

II. Materials and Methods

A. Data

Our data consists of 268 FDG-PET scans taken from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (http://www.loni.ucla.edu/ADNI/). ADNI provides a listing of all PET scans considered to have potential issues that may effect results of image analysis, none of these scans were used. Of the 268 scans, 81 were from Control Normals (CN), 113 from MCI subjects and 74 from AD subjects. Some clinical and Neuropsychological characteristics of each group are summarized in Table I.

Group	Normal	MCI	AD
Age (m±sd)	77.3±4.7	76.4±7.3	76.5±6.8
Sex (M/F)	53/28	75/38	45/29
MMSE (m±sd)	29.1±1.2	26.2±2.7	21.1±4.1

Table I
Clinical and Neuropsychological characteristics of each group (Mean ± Standard Deviation).

The PET images downloaded from the ADNI database had been processed in order to make PET data from different systems more similar. The processing included co-registration to their baseline PET scan and reorientation into a standard space, intensity normalization and conversion to a uniform isotropic resolution of 8 mm FWHM. The image matrices were 128x128x60 and intensity values ranged from 0 to 32700 in all scans. Extra-cranial voxels were excluded from the analysis.

B. The Boosting Algorithm

The Boosting algorithms belong to the class of ensemble methods which combine the output of several simple classifiers to form a complex one. In this work we use Adabost [11] where the simple classifiers are learned sequentially. They are called weak classifiers because they are not expected to classify data well. However, at each round of learning, the performance of the next weak classifier is boosted by a re-weighting of the examples in order to emphasize those that were incorrectly classified in the previous round. The final strong classifier is a weighted combination of the weak classifiers followed by a thresholding operation. Moreover each classifier is constrained to depend on a single feature, thus feature selection is performed.

Let $x_i \in \mathbb{R}^n$ denote the training patterns, $i = 1, ..., N$ and $y_i \in \{0, 1\}$ denote the corresponding classification. Each pattern has a weight $w_i \in [0, 1]$ which is initially given by $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives, respectively. Let h_j denote a weak classifier, $j = 1, ...,$. Each weak classifier h_j consists of a feature f_j (the j-th component of pattern x), a threshold θ_j and a parity $p_j \in \{-1, 1\}$ indicating the direction of the inequality sign:

$$h_j(x) = \begin{cases} 1 & p_j f_j(x) < \theta_j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The parameters of each weak classifier are learned from the data by minimizing:

$$\varepsilon_{t,j} = \sum_{i=1}^{N} w_{t,i} \left| h_j(x_i) - y_i \right| \quad (2)$$

and the weak classifier with the lowest error $\varepsilon_t = \min_j \varepsilon_{t,j}$ is chosen. After each round, t, the weights are updated by:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i} \quad (3)$$

where $\beta_t = \varepsilon_t / (1 - \varepsilon_t)$ and $e_i = 0$ if example x_i was correctly classified and $e_i = 1$ otherwise. The weights are subsequently normalized. This procedure is iterated T times, where T is the number of features to be selected and also the number of weak classifiers used.

The final classifier is given by:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $\alpha_t = \log(1/\beta_t)$.

III. Results

The Boosting method was applied to the detection of AD vs NC, MCI vs NC and AD vs MCI, using the voxels intensities as features. 150 features were used. In order to evaluate the generalization performance of the method, we used 10-fold cross validation, and averaged the testing set accuracy over the 10 folds. We compared these results with those obtained by the SVM classifier which is the most widely used in this context. For the SVM classifier, the RBF kernel was used and the model parameters (C, γ)

were estimated within each training fold by cross validation. Additionally we applied the SVM classifier to the features selected by the Boosting classifer (BSVM). In this case, those features that were selected by Boosting more than once were not replicated. Table II sumarizes the results.

Group	NC/AD	NC/MCI	MCI/AD
Boosting (%)	90.97	79.63	70.00
SVM (%)	86.11	74.07	68.75
BSVM (%)	86.80	71.52	59.72

Table II
ACCURACY OF THE DIFFERENT CLASSIFIERS.

From Table II we conclude that the Boosting classifier performs better than SVM method in the three problems, leading to an accuracy improvement of 3-5%. The Boosting method achieves an accuracy of 90.97% in the detection of AD vs NC. The results obtained by this classifier in the detection of MCI are also very interesting, since accuracy is close to 80%. This clearly states that the proposed approach can still be useful for the automatic detection of MCI from PET images although the decision errors increased 10%. The most difficult problem is clearly the discrimination between AD and MCI. Even in this case, the Boosting classifier can provide a tentative classification with an accuracy of 70.00%. Boosting alone also performs better than Boosted SVM (BSVM).

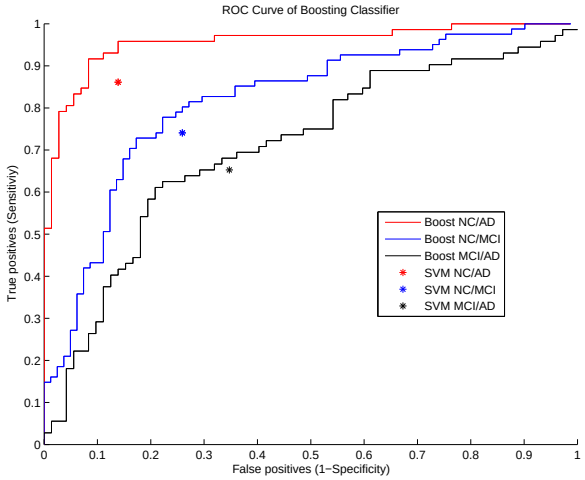

Figure 1. ROC curves for the three problems.

Fig. 1 shows the ROC curves for the boosting classifier. These curve were obtained by changing the threshold (right hand side in 4) in order to accept class 1 with larger (or lower) probability. This curve shows possible tradeoffs between sensitivity and specificity. The best solution depends on the application and can be chosen by the user. The performance of the SVM classifier is also displayed in the figure with an asterisk. We can easily conclude that the boosting classifier performs better in all cases (it is above).

In addition, the Boosting classifier was trained to discriminate between all AD and NC subjects and was then applied to the MCI cases. The output of the classifier for each PET scan is shown in Fig. 2 where NC subjects are shown in red, AD in blue and MCI in black. The training data patterns associated to AD and NC are well separated as expected but the same cannot be said about MCI test patterns. The classifier output is uniformly distributed in the range of interest in such cases, which suggests that there is a continuous transition between these two states.

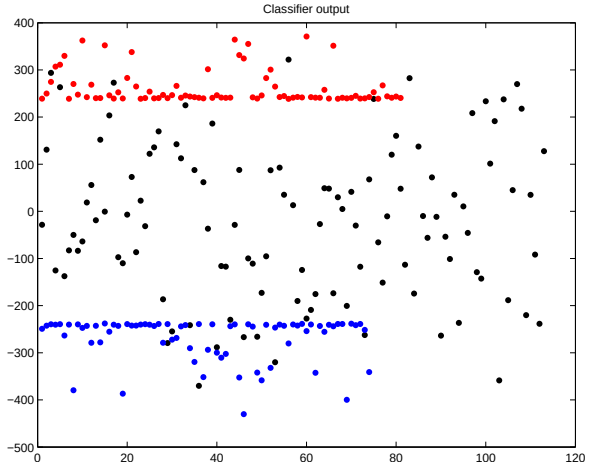

Figure 2. Output of Boosting classifier trained with NC and AD and tested on MCI data. (NC and AD training patterns are shown in red and blue, respectively, and MCI test patterns in black).

The features selected by the Boosting classifier were sparsely located in several regions throughout the brain, including the hippocampus, the posterior cyngulate and the inferior temporal lobe bilaterally, which suggests a widespread pattern of brain atrophy. This is illustrated in Fig.3 where some of these features are shown in red superimposed on one of the PET scans. Of the selected features, 51.5% were located in the left cerebral hemisphere while 48.5% were on the right hemisphere which indicates that the changes related to AD are bilateral.

IV. CONCLUSIONS

This paper describes the application of the Boosting classifier to detect Alzheimer's disease (AD), Mild Cognitive Impairment (MCI) and to discriminate between both conditions using PET scans. The Boosting classifier adopted in this paper is able to select a small number of voxels from the whole volume and provide a robust classification of the input using an ensemble of weak classifiers, each of them depending on the intensity of a single voxel.

The classifier was trained and evaluated by cross-validation, using a subset of the ADNI database with 268

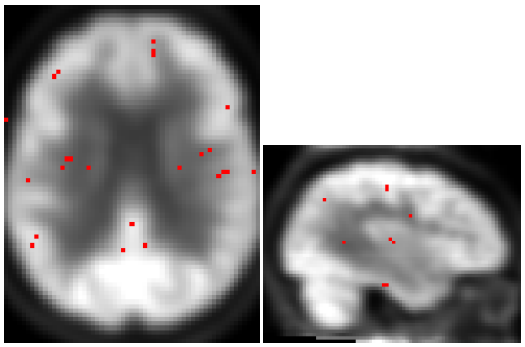

Figure 3. Some of the features selected by the Boosting Classifier.

scans. Despite the fact that intensities depend on registration and normalization, we obtained accuracies of 90.97% in the detection of AD, 79.63% for MCI and an accuracy of 70.00% in the discrimination between both. The method outperformed the widely used SVM classifier, and has lower computational complexity. The method also outperformed other state-of-the-art techniques which used PET images of the same database [10].

We have also applied the Boosting classifier, trained to discriminate AD from NC, to scans of people with MCI. In the case of MCI scans, the output of the classifier is uniformly distributed in the range of interest, suggesting that there is a continuous transition between both states.

We conclude that the Boosting classifier can be used for AD and MCI detection with PET images. Future work will include the use of multi-class classifiers in order to jointly distinguish between the three classes (AD, MCI, NC).

Acknowledgments

This work was supported by Fundação para a Ciência e a Tecnologia (ISR/IST and ICIST/IST plurianual funding) through the POS Conhecimento Program that includes FEDER funds. Data was provided by the Alzheimers Disease Neuroimaging Initiative (ADNI).

References

[1] E. Gerardin, G. Chételat, M. Chupin, R. Cuingnet, B. Desgranges, H. Kim, M. Niethammer, B. Dubois, S. Lehéricy, L. Garnero, F. Eustache, and O. Colliot, "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging," *NeuroImage*, vol. 47, no. 4, pp. 1476–1486, October 2009.

[2] Y. Fan, N. Batmanghelicha, C. Clark, and C. Davatzikos, "Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline," *NeuroImage*, vol. 39, no. 4, pp. 1731–1743, February 2008.

[3] R. Higdon, N. L. Foster, R. A. Koeppe, C. S. Decarli, W. J. Jagust, C. M. Clark, N. R. Barbas, S. E. Arnold, R. S. Turner, J. L. Heidebrink, and S. Minoshima, "A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG-PET imaging," *Statistics in Medicine*, vol. 23, no. 2, pp. 315–326, 2004.

[4] Y. Kosugi, K. Uto, R. Hagiwara, N. Kosaka, A. Abe, M. Kameyama, and T. Momose, "Neural networks for cerebral diagnosis using PET and SPECT," *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*, pp. 24–27, October 2003.

[5] J. Stoeckel, G. Malandain, O. Migneco, P. Koulibaly, P.and Robert, N. Ayache, and J. Darcourt, "Classification of SPECT images of normal subjects versus images of Alzheimer's disease patients," *4th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, vol. 2208, 2001.

[6] G. Fung and J. Stoeckel, "SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information," *Knowledge and Information Systems*, vol. 11, no. 2, pp. 243–258, 2007.

[7] C. Davatzikos, S. M. Resnick, X. Wu, P. Parmpi, and C. M. Clark, "Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI," *NeuroImage*, vol. 41, no. 4, pp. 1220–1227, July 2008.

[8] P. Vemuri, J. L. Gunter, M. L. Senjem, J. L. Whitwell, K. Kantarci, D. S. Knopman, B. F. Boeve, R. C. Petersen, and C. R. Jack Jr., "Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies," *NeuroImage*, vol. 39, no. 3, pp. 1186–1197, February 2008.

[9] S. Duchesne, A. Caroli, C. Geroldi, C. Barillot, G.B. Frisoni, and D.L. Collins, "MRI-based automated computer classification of probable AD versus normal controls," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 509–520, April 2008.

[10] C Hinrichs, V Singh, L Mukherjee, G Xu, MK Chung, SC Johnson, and Disease Neuroimaging Initiative Alzheimer's, "Spatially augmented LPBoosting for AD classification with evaluations on the ADNI dataset.," *Neuroimage*, vol. 48, no. 1, pp. 138–49, 2009.

[11] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.