IMPROVING SHORT-TERM PREDICTION FROM MCI TO AD BY APPLYING SEARCHLIGHT ANALYSIS

Juan Eloy Arco[1], *Javier Ramírez* [2], *Carlos G. Puntonet* [3], *Juan Manuel Górriz* [2], *María Ruz*[1]

[1] Department of Experimental Psychology
[2] Department of Signal Theory Networking and Communications
[3] Department of Architecture and Computer Technology
University of Granada, Granada 18071, Spain

ABSTRACT

Alzheimer's disease (AD) is the most common cause of dementia. Nowadays, 44 million people worldwide suffer from this neurodegenerative disease. Fortunately, the use of new technologies can help doctors in diagnosing this disease in an increasingly early stage, which is vital to prevent its advance. In this work we have developed a new automatic method to predict if patients suffering from mild cognitive impairment (MCI) will develop AD within one year or, conversely, its impairment will remain stable. This technique is based on the so-called *Searchlight*, a widely known approach in fMRI but which has not been previously used with structural images. Besides analyzing the intensity of the voxels in each of the subregions defined by the Searchlight, data from two neuropsychological tests were used during the classification process, achieving an accuracy of 84%.

Index Terms— Alzheimer's disease, mild cognitive impairment, searchlight, MRI, prediction.

1. INTRODUCTION

In the last few years, there has been a large increase in the number of studies attempting to develop systems for the diagnosis of AD ([1]). Most existing research focuses on only a single modality of biomarkers for diagnosis of AD and MCI, although recent studies have shown that different biomarkers may provide complementary information for the diagnosis of AD ([2]). Given the importance of carrying out a diagnosis as early as possible, in this study we propose a method to compare between MCI patients who had converted to AD within 12 months and MCI patients who had not converted to AD within 12 months, in order to predict whether the patient will develop the disease or not.

A critical decision in order to obtain the best possible performance is given by the choice of the appropriate brain region. Atlas-based parcellation using a predefined anatomical brain atlas is a simple feature extraction method with good interpretability and general versatility. However, the non concordance between different brain atlases makes that depending on which of them is used for parcellation different features will be provided. Another popular approach to look in the whole brain for discriminative pattern information is provided by the searchlight mapping ([3]). Thus, searchlight analysis can minimize the effects of the much feared 'curse of dimensionality' since relatively few voxels are typically included in each searchlight, in addition to the results potentially easy to interpret. These appealing aspects have led to a rapid increase in the number of studies using searchlight analyses. Nevertheless, none of these works is aimed at early prediction of a neurological disorder but they focus on fMRI experiments and the subsequent study of the areas with highest brain activity from some kind of stimulus ([4]). Given the good results that this method provides in fMRI, we believe that high performance can also be achieved in structural images (sMRI). Besides, combining data from different biomarkers or modalities can cause a considerable increase in the performance of the classifier ([5]). The main contributions of this paper are to show that searchlight-based methods can be used for an early prediction of AD combining both sMRI and two neuropsychological tests: MMSE (Mini-mental state examination) and ADAS-Cog (Alzheimer's disease assessment scale-cognitive subscale).

2. DATABASE

The data used in the preparation of this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessments can be combined to measure 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 develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. In this paper, only ADNI subjects with all corresponding MRI, MMSE and ADAS-Cog baseline data are included. This yields a total of 134 MCI subjects who had at least three longitudinal scans

(baseline image, and two subsequent images six months and twelve months after) including 73 MCI converters who had converted to AD within 12 months and 61 non-converters who had not converted to AD within 12 months.

3. METHODS

3.1. Geometric structure

The basis of the searchlight analysis is the geometric structure of the voxel-space in which the brain images are defined. The results that may be obtained from this method are closely related to the shape and especially the size of this structure. Although most studies set a spherical shape, in this work we used a cubic shape due to its simplicity. Regarding the size, it was necessary to strike a balance between performance and the size of the region to be analyzed. It is widely known that the regions which are associated with the appearance of AD have a considerable size. However, choosing too large structures can cause an exponential decay of the performance, as unrelated voxels to the central one would be used for classification. Besides, it should be noted that the longer the side of the cube, the higher the computational cost of the associated process with the analysis. Striking a balance between performance and execution time in the location of the region under study we found that a suitable size for the cubic side was 18mm long (9 voxels), meaning that a total of 729 voxels are used as input features of the classification process, in addition to the scores of the above-mentioned two neuropsychological tests.

The fact or performing as many classification procedures as voxels in the brain makes is a highly demanding computational process. Limiting the analysis to a particular anatomical region (such as the hippocampus) can lead to suboptimal results because of not taking into account the information that lies in other areas. A solution is to segment previously the MR images into white and grey matter (besides cerebrospinal fluid), what means that the searchlight structure is swept firstly across the voxels contained in the gray matter region of the sMRI scan, being then the procedure repeated for white matter. Thus, the process is speeded up without focusing on too small regions inasmuch as the whole grey/white matter is studied. It is said that there are specific networks within grey matter which are more vulnerable to age-related neurodegeneration, and that it degenerated sooner than other brain areas. For this reason, it is quite likely to get better results than by studying the regions belonging to the white matter. However, it is not possible to obviate its participation in the prediction of AD.

3.2. Classification and performance evaluation

The data within a searchlight and the scores from the two tests are then vectorized as the input information which the algorithm hast to learn from. Several authors have previously made comparisons between different types of classifiers in order to find which one gets better results inside a searchlight procedure ([6]). Their results show that linear classifiers perform better than the nonlinear, what could mean that the true distributions' Bayes-optimal decision boundaries were approximately linear. Besides, linear classification are easier to interpret because they let generate a map showing the contribution of each of the input features to the classification results. LDA (Linear Discriminant Analysis) was sucesfully used in [7], with a shrinkage estimator of the covariance matrix that sets the shrinkage parameter automatically.

However, when the dimensionality of the data is very large, the covariance-matrix estimation and inversion required for LDA become very computationally intensive. In the case of searchlight, hundreds or even thousands of voxels (depending on the size) are taken into account, so it is advisable to use a different algorithm. In this work, a linear support vector machine algorithm was used. As it is widely known, SVM chooses the hyperplane that has the maximum marging, i.e. the hyperplane that separates the classes with the maximum safety clearance to the closest training patterns on either side. The general case, where the training data points cannot be perfectly separated, is handle by allowing a few misclassifications among the training data points. A parameter $C > 0$ defines a penalty for misclassification, being quite important for good generalization performance as it controls regularization, which counteracts overfitting of the training data. We used the LIBSVM 3.20 package [8], with a parameter C equal to 1 for the SVM analyses.

To assess the classification performance of each searchlight subset we used a k-fold cross validation procedure. This scheme works in rounds: in each one of them, the dataset is randomly divided into groups of *k* observations, training the classifier with all groups but one, and testing it with the remaining group. This procedure assures that each classifier trained is tested with both positive and negative observations, being the influence of all observations in the performance estimate assures too, while allowing a potentially large number of training and test combinations. Many authors recommend the use of leave-one-out cross validation, a specific case of k-fold in which the dataset is divided into as many groups as samples are. However, it was not used for two reasons. The first one is that the database used in this work (with more than 130 subjects) is large enough to not need to carry out such a exhaustive procedure. The second one is related to the high computational cost that this involves, which together with the above mentioned high computing power required for the searchlight can make the execution time increase so that it is impracticable carry out this analysis in a low performance PC. For this reason, a 10-fold cross validation procedure was used, so that the data from the 134 patients available in the database were split into 10 folds, so that 9 of these 10 folds provided the training samples and the remaining provided the test samp-

Table 1. Results obtained using gray matter, white matter and neuropsychological tests from sessions 6/12 months before the conversion (MCI-converters). Besides, the average session in which the conversion happened was calculated, being used the data from the two sessions prior this one for MCI-no converters.

Prediction accuracy (%)

6 months before the conversion		
Approach	Grey Matter	White matter
Searchlight	84.3	82.05
Atlas	68.75	67.97
12 months before the conversion		
Approach	Grey Matter	White Matter
Searchlight	81.44	77.02
Atlas	65.55	65.55

les. The mean classification accuracy across the ten folds was used as the estimate of the classifier's performance.

4. RESULTS AND DISCUSSION

The aim of this work was to test the performance of the searchlight technique for AD prediction and has been proved. Although it is widely used in fMRI studies, searchlight had not been previously applied to structural images. The experiments carried out on the database formed by both structural MRI (segmented into gray and white matter) and neuropsychological tests (MMSE and ADAS-Cog). Several trials were made using different features, yielding an accuracy equal to 84%. These results are significantly higher than those obtained in our previous work [9], in which LDA was used as the classification algorithm and the selection of the anatomical regions to study were based on AAL atlas.

Table 1 shows the results obtained by the atlas approach besides searchlight approach used in this work. For MCI converters patients, the data from one and two sessions before their conversions can be used. Besides, the average conversion session was calculated for all these patients, resulting that this was the fourth session (month 18 of the longitudinal analysis). In the case of atlas approach, there is a lower performance 12 months before the conversion. Instead, using a searchlight approach let predict the development of AD with almost the same accuracy that could be obtained six months later, which means that this method is able to find information about the impairment caused by this disease where the other approach can not.

Figure 1 shows the regions both from gray and white matter with an accuracy higher than 70%. It can be seen that there are more informative areas in the case of the images 6 months before the conversion, which is logical given the progress of the disease. Similarly, the number of selected voxels from white matter images is higher than in gray matter, although the vast majority have a lower accuracy than in the other brain tissue. According to the AAL atlas, the brain regions from which the algorithm used in this paper get the best performance are the superior temporal gyrus and the supramarginal gyrus from gray matter whereas it focus on cerebellar regions from white matter.

It must be remembered that although there are other automated systems for predicting AD, this work represents a step further in predicting this disease because only patients with mild cognitive impairment are considered, as in the case of patients who have been diagnosed with AD only data prior to this conversion were used. Thus, searchlight approach finds significant differences in patients with identical clinical diagnosis. The fact that so it is not required to choose a prior anatomical region in which focus the analysis as this technique analyzes the whole brain by dividing it into small regions makes it an useful choice as an aid in the diagnosis of this disease. In the previous study we conducted there was a great improvement in performance when information from the two sessions previous to the conversion one were combined, that is, sMRI and neuropsychological tests 6 and 12 months earlier. We have not carried out this analysis because searchlight is computationally demanding, especially for large searchlight spheres and leave-one-out cross validation schemes. A suggestion for future research might be to develop a parallelized implementation based on graphics processing units (GPUs), which can be used to speedup such procedures and get a likely improvement in the performance.

5. CONCLUSION

In the current study, we have proved that applying searchlight approach to sMRI (which have not been previously done) can lead to a considerable improvement of the results compared with atlas-based approaches. Specifically, the developed system to predict Alzheimer's disease in mild cognitive impairment patients yields an accuracy of 84.3% 6 months before the possible conversion and 82.05% 12 months before, using the gray matter of the brain in both cases. We conclude that it would be quite interesting to develop a parallelized implementation based on graphics processing units (GPUs), besides that the combination of the information from the several images temporary prior to the conversion session could considerably improve the performance.

Acknowledgments

This work was partly supported by the MICINN under the TEC2012-34306 project and the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía, Spain) under the Excellence Project P11-TIC-7103.

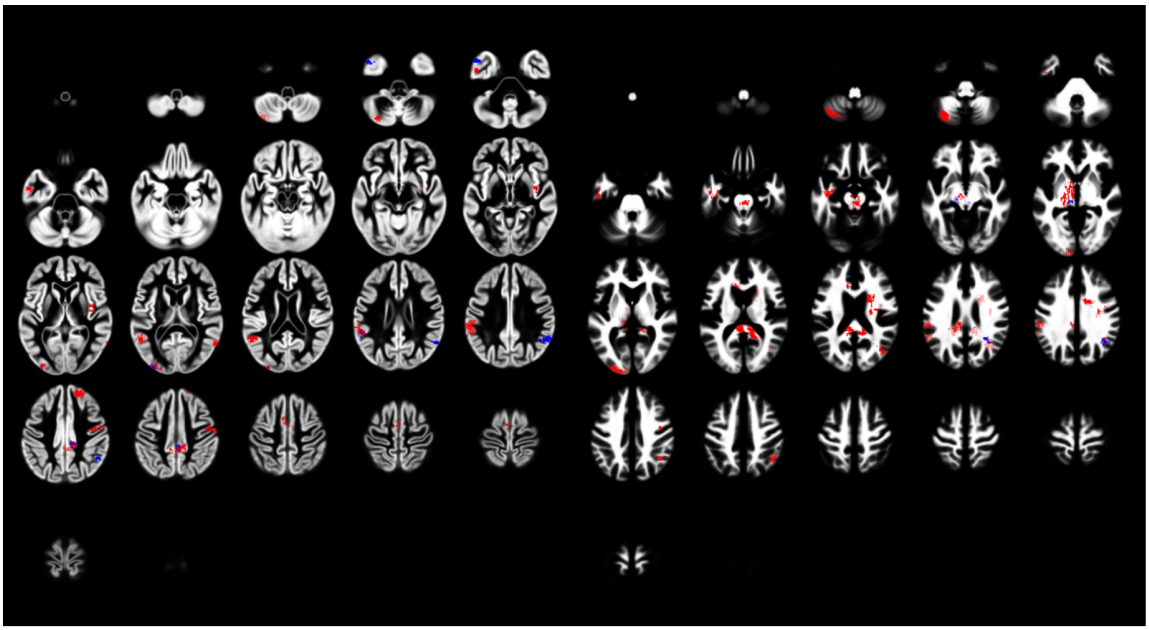

Fig. 1. Representation of the brain regions that achieve an accuracy in the classificatioon higher than 70% when the gray (left) and white (right) matter are used. The selected areas 6/12 months prior to the conversion are indicated with red/blue colour.

6. REFERENCES

[1] A. Nazeri, H. Ganjgahi, T. Roostaei, T. Nichols, M. Zarei, and Alzheimer's Disease Neuroimaging Initiative, "Imaging proteomics for diagnosis, monitoring and prediction of alzheimer's disease," *NeuroImage*, vol. 102 Pt 2, pp. 657665, November 2014.

[2] K. Ota, N. Oishi, K. Ito, H. Fukuyama, SEAD-J Study Group;Collab, and Alzheimer's Disease Neuroimaging Initiative;Collab, "Effects of imaging modalities, brain atlases and feature selection on prediction of alzheimer's disease," *Journal of neuroscience methods*, August 2015.

[3] N. Kriegeskorte, R. Goebel, and P. Bandettini, "Information-based functional brain mapping," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 10, pp. 3863–3868, 2006.

[4] D. Rivolta, A. Woolgar, R. Palermo, M. Butko, L. Schmalzl, and M.A. Williams, "Multi-voxel pattern analysis (MVPA) reveals abnormal fMRI activity in both the "core" and "extended" face network in congenital prosopagnosia.," *Frontiers in Human Neuroscience*, vol. 8, no. 925, 2014.

[5] Z. Dai, C. Yan, Z. Wang, J. Wang, M. Xia, K. Li, and Y. He, "Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3).," *NeuroImage*, vol. 59, no. 3, pp. 2187 – 2195, 2012.

[6] M. Misaki, Y. Kim, P.A. Bandettini, and N. Kriegeskorte, "Comparison of multivariate classifiers and response normalizations for pattern-information fMRI.," *NeuroImage*, vol. 53, no. 1, pp. 103 – 118, 2010.

[7] F. Pereira and M. Botvinick, "Classification of Functional Magnetic Resonance Imaging Data Using Informative Pattern Features.," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2011, KDD '11, pp. 940–946, ACM.

[8] Chang, C.C. and Lin, C.J., "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.

[9] J.E. Arco, J. Ramírez, C.G. Puntonet, J. M. Górriz, and M. Ruz, "Short-term Prediction of MCI to AD conversion based on Longitudinal MRI analysis and neuropsychological tests.," in *Innovation in Medicine Healthcare 2015*, 2015, pp. 385–394.