30th Annual International IEEE EMBS Conference
Vancouver, British Columbia, Canada, August 20-24, 2008

# A supervised method to assist the diagnosis and classification of the status of Alzheimer's disease using data from an fMRI experiment

Evanthia E. Tripoliti, Dimitrios I. Fotiadis, *Senior Member, IEEE*, Maria Argyropoulou

*Abstract*—**The aim of this work is the development of a method to assist the diagnosis and classification of the status of Alzheimer's Disease (AD) using information that can be extracted from fMRI. The method consists of five stages: a) preprocessing of fMRI data to remove non-task related variability, b) modeling BOLD response depending on stimulus, c) feature extraction from fMRI data, d) feature selection and e) classification using the Random Forests (RF) algorithm. The proposed method is evaluated using data from 41 subjects (14 young adults, 14 non demented older adults and 13 demented older adults.**

## I. Introduction

THE Alzheimer disease (AD) is a progressive disease of the brain which is characterized by the impairment of memory and disturbance in at least one other thinking function (language or perception of reality). The main risk factor for AD is increased age [1].

Current consensus statements have emphasized the need for early recognition of AD [2]. However, there is no single comprehensive test to diagnose Alzheimer's disease. Based on medical history, examinations, laboratory and brain imaging results, a properly trained physician can diagnose the cause of patient's condition. The diagnosis is made using the National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association (NINCDS – ADRDA) criteria. Typically, it takes a few weeks to complete a diagnostic evaluation of Alzheimer's disease and related disorders. The only way to confirm the diagnosis of AD is through autopsy [2].

A variety of neuroimaging techniques, including PET, fMRI and structural MRI have made significant advances during the last years in an attempt to diagnose the disease earlier and monitor the progression of the disease pathologically, specifically to identify early markers. Functional MRI allows researchers to diagnose AD by pinpointing dysfunctional areas of the brain, structural MRI may aid in diagnosis by measuring overall brain shrinkage and PET may allow scientists to see the plaques and tangles of AD in living patients [2].

The Alzheimer's researchers are particularly interested in functional magnetic resonance imaging (fMRI) since it offers several distinct advantages over PET, including non-invasion, increased spatial and temporal resolution and repeatability as it does not require radiation exposure. fMRI enables the examination of brain activity during performance of cognitive, sensory and motor tasks and thus provides a powerful mean to identify disrupted neural circuits that underlie disorders such as AD. Relating this to Alzheimer's disease, a "cognitive stress test" can be used to bring out subtle brain abnormalities that would otherwise go undetected during a resting state [2].

A variety of differences between AD patients, normal young and elderly volunteers were revealed through the use of fMRI [2]. Differences concern the brain activation, the shape of BOLD response, metabolism and functional synchrony [3]-[12]. Shi-Jiang Li *et al.* [9] measured changes in functional synchrony in the hippocampus in patients with mild cognitive impairment and Alzheimer disease. Functional synchrony was quantified using the COSLOF index, which values were significantly lower in AD patients than in control subjects. Recent functional imaging studies have revealed co-activation in a distributed network of cortical regions that characterizes the resting state, or the default mode, of the human brain. Brain regions implicated in this network have shown decreased metabolism early in the course of AD. Greicius *et al.* [10] reasoned that the default-mode network activity might be abnormal in AD. To test this hypothesis, they have used fMRI and independent component analysis. J.P. Petrella *et al.* [12] focused their research on identifying brain regions in which task related changes in activation during a memory encoding task correlate with the degree of memory impairment across AD, mild cognitive impairment and elderly control subjects.

All of the studies reported previously provide increasing evidence that dementia of Alzheimer type represents a disease process that is different from the process of healthy aging. They present how activation, hemodynamic response function and other factors are differentiated between AD and healthy groups and provide useful information about cognitive functions which are affected by the presence of dementia. Only Shi-Jiang Li *et al.* [9] and Greicius *et al.* [10] quantified some of AD related changes in order to create an index that can serve as preclinical marker for the diagnosis of AD.

In this work all the data that were received during the paradigm (fMR images, MR images, demographic details, behavioral data) were exploited in order to classify a person as AD or healthy and also to classify the status of the

E.E. Tripoliti and D.I. Fotiadis are with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Computer Science, University of Ioannina and Biomedical Research Institute – FORTH, GR 451 10 Ioannina, Greece (Tel: +30-2651-0-98803; fax: +30-26510-98889; e-mail: fotiadis@cs.uoi.gr).
M. Argyropoulou is with the Department of Radiology, Medical School, University of Ioannina, GR 451 10 Ioannina, Greece.

disease. A supervised method is proposed. The method consists of five stages (Section II-C – Section II-G). The points that differentiate this method from those reported in the literature are the following: a) it diagnoses the AD and classifies its status, b) it is independent from the type of the task and the design of the experiment, c) it is not based in one only feature but combines all features that express AD related changes and are extracted from fMRI data and d) it fuses features form different categories that depict as much as possible medical knowledge about AD. The proposed method is evaluated using a dataset of 41 subjects.

## II. MATERIALS AND METHODS

### A. Dataset

Forty one right-handed, English speaking individuals participated. Fourteen healthy young subjects (five male and nine female), thirteen subjects (six male and seven female) with very mild to mild AD and fourteen healthy elderly subjects (five male and nine female) were scanned during a simple sensory motor paradigm. The mean age of young participants was 21.1 years (range 18 – 24 years). The nondemented elderly participants ranged from 66-89 years old (mean age 74.9 years) while AD subjects ranged from 68-83years old (mean age 77.2 years).

Dementia status was established using the Washington University Alzheimer's Disease Research Center (ADRC) recruitment and assessment procedures. Non demented control subjects and those with mild dementia of Alzheimer type were assessed clinically with the Clinical Dementia Rating (CDR), where CDR 0 indicates no dementia and CDR 0.5 and CDR 1 indicate very mild and mild dementia of Alzheimer type, respectively [11].

### B. Imaging methods

Raw structural and functional data were received from the fMRI Data Center at Dartmouth College (Hanover, NH). Functional images were acquired on a Siemens 1.5-T Vision System with an asymmetric spin-echo sequence sensitive to BOLD-contrast. The following parameters were used: volume TR=2.68sec; 3.75x3.75mm in-plane resolution; T2* evolution time=50msec; alpha=90$^{o}$. Whole brain imaging was performed using 16 contiguous 8-mm thick axial oblique slices (acquired parallel to the plane of anterior-posterior commisures). Each functional run lasted approximately 5.5min and a 2min delay existed between runs, during which subjects were permitted to rest. High resolution structural images were also acquired in a series of three to four separate T1-weighted MP-RAGE anatomic images with the following parameters: 1x1x1.25mm resolution; TR=9.7msec; TE=4msec; flip angle=10$^{o}$; TI=20msec; TD=500msec [11].

### C. Data preprocessing

Data preprocessing is carried out in four steps. The first step is slice timing. It is done by a shift of the phase of the sines that make up the signal. The second step is motion correction. It is performed by registering the whole time series of images to the image acquired first. The third step is spatial normalization, where all image volumes are interpolated and scaled to conform the Talairach and Tournoux atlas (interpolated to 2-mm isotropic voxels). The fourth step in the preprocessing is spatial smoothing. It is performed through a discrete convolution with a 6-mm Gaussian kernel. The convolution is performed in all three directions of the fMRI data [13].

### D. Data modeling

The purpose of the preprocessing steps is to remove some confounding effects and make the data better meet the assumptions of the Gaussian Random Field and General Linear Model used to model fMRI data [13]. Modeling consists of two steps: a) decomposition of data into effects and error and b) formation of statistics, using estimates of effect and error. Through modeling we can make inference about the effects of interest, to detect, in a robust, sensitive, and valid way, those parts of the brain which show increased intensity at the points in time that stimulation was applied and to extract brain connectivity.

### E. Feature extraction

The output of modeling stage is essential to extract some features. The features are extracted from each functional run of each patient and can be grouped into the following categories: 1) Demographics, 2) Head motion, 3) Behavioral data, 4) Volumetric measures, 5) Activation patterns and 6) Hemodynamic modeling. The features that belong in each category are described in [14].

### F. Feature selection

In this work feature selection is accomplished on the basis of correlation between features. A feature is good if it is highly correlated to the class but not highly correlated to any of the other features. The correlation measure is based on the information-theoretical concept of entropy which is a measure of the uncertainty of a random variable. The algorithm evaluates the worth of an attribute by measuring the symmetrical uncertainty with respect to the class [15].

### G. Classification

Selected features are fed into the classification algorithm. We use the Random Forests (RF) algorithm which is a classifier that consists of many decision trees. The predicted class is the one that most of the decision trees vote [16]-[17]. Also three improvements of RF algorithm, proposed by Marko Robnik-Sikonja [18], were applied to our data. These improvements are the following: a) ReliefF algorithm is used instead of Gini index, b) Five different attribute evaluation measures (Gini index, Gain ratio, Minimum Description Length, ReliefF, myopic Relief) for split selection is used instead of one [19]-[20], c) ordinary voting was replaced by weighted voting (wv). The reason for these improvements was to strengthen individual trees, without sacrificing variety between them or, alternatively increase variance without sacrificing strength. The classifier is

evaluated using 10-fold stratified cross validation.

## III. RESULTS

The proposed method was evaluated using three different datasets. The first consists of two groups: healthy old and demented old subjects. The second dataset is the same as the first one but demented old subjects are divided in subjects with very mild and mild AD. The third dataset is the same as the second one but also includes healthy young subjects. For each one of these datasets two repetitions were made. One using all extracted features and a run using only features extracted from fMRI data.

The input of the selection algorithm was a matrix which contains extracted features from all four functional runs for each subject. The last column of this matrix was completed with a number from one to four depending on the class that the subject belongs to. The number of instances is 108 (27 subjects x 4 functional runs) for the first and the second dataset and 164 (41 subjects x 4 functional runs) for the third one. Different feature selection algorithms were tested in order to conclude to the appropriate one (symmetrical uncertainty).

The performance of the proposed method is affected by two parameters. First the number of features (F) that will be retained from the selection algorithm and second the number of trees (T) that the RF algorithm will generate. In case of RF with weighted voting one more parameter influence the performance of our method. This parameter is the number of similar instances (K) and was derived from the weighting assignment procedure. We applied our method for different values of these parameters. The values of parameters that produce the results for each dataset, the achieved accuracy, sensitivity and specificity are reported in Table I and Table II. The improvement of the evaluation measures can be fully justified by the fact that age and loss of gray matter is two important markers for the diagnosis of the disease according to physicians and findings reported in the literature [3]-[12].

The features that were selected in both cases are the following: path length, behavioral data, total number of activated voxels, the value of primary peak, the size of cluster where the voxel with maximum z-score belongs to, the percentage of activated regions that belong to a region of interest, the total activations of these regions, the amplitude of BOLD response, the amplitude of undershoot, the amplitude of rCBF, of venous volume and deoxyHb. When features from different categories were fused demographic details and volumetric measures also participate in the classification. Furthermore transit time was used. On the other hand, when only fMRI data were used, transit time was replaced by the amplitude of the vascular signal. Different combinations of these features are selected depending on the classification problem and the version of the RF algorithm that was applied.

A comparison of different versions of RF algorithm leads us to the conclusion that weighted voting can be used to assist the diagnosis and to classify the status of AD. This conclusion is supported by the results of the comparison of our method with those reported in the literature (Table III). The comparison can be conducted only in the case of the two class problem since there are not methods reported in the literature that classify the status of AD.

According to these results the accuracy is 7% and 5% better than and specificity is 5.3% and 2.3% better than those reported by Greicius *et al.* [10] and Shi-Jiang Li *et al.* [9] respectively. The sensitivity of the proposed method is 1% lower than that of Greicius *et al.* and 4% higher than that of Shi-Jiang Li *et al.*.

TABLE I
RESULTS OF THE PROPOSED METHOD-ONLY FMRI FEATURES

| Method | Dataset | Sensitivity | Specificity | Accuracy |
|---|---|---|---|---|
| Classical RF | 2 classes F=14, T=35 | 82% | 87% | 84% |
| | 3 classes F=17, T=82 | - | - | 78% |
| | 4 classes F=15, T=75 | - | - | 85% |
| RF with ReliefF | 2 classes F=14, T=42 | 86% | 86% | 86% |
| | 3 classes F=17, T=53 | - | - | 80% |
| | 4 classes F=16, T=89 | - | - | 81% |
| RF with multiple estimators | 2 classes F=14,T=100 | 84% | 90,3% | 87% |
| | 3 classes F=16, T=18 | - | - | 80,5% |
| | 4 classes F=18, T=84 | - | - | 83% |
| RF with weighted voting | 2 classes F=14,T=55,K =15 | 84% | 92,3% | 88% |
| | 3 classes F=17,t=82,K=15 | - | - | 80% |
| | 4 classes F=16,t=81,K=15 | - | - | 87% |

TABLE II
RESULTS OF THE PROPOSED METHOD-ALL EXTRACTED FEATURES

| Method | Dataset | Sensitivity | Specificity | Accuracy |
|---|---|---|---|---|
| Classical RF | 2 classes F=19, T=19 | 98% | 98% | 98% |
| | 3 classes F=18, T=36 | - | - | 97% |
| | 4 classes F=16, T=52 | - | - | 98% |
| RF with ReliefF | 2 classes F=23, T=27 | 98% | 98% | 98% |
| | 3 classes F=18, T=67 | - | - | 99% |
| | 4 classes F=16, T=47 | - | - | 99% |
| RF with multiple estimators | 2 classes F=17,T=48 | 98% | 98% | 98% |
| | 3 classes F=18, T=41 | - | - | 98% |
| | 4 classes F=16, T=51 | - | - | 99% |
| RF with weighted voting | 2 classes F=17,T=42,K =5 | 99% | 98% | 99% |
| | 3 classes F=15,t=22,K=7 | - | - | 99% |
| | 4 classes F=16,t=91,K=7 | - | - | 99% |

TABLE III
COMPARISON OF THE PROPOSED METHOD WITH OTHER METHODS EXISTING IN THE LITERATURE

| Method | | Sensitivity | Specificity | Accuracy |
|---|---|---|---|---|
| Our work | Classical RF 1.17sec* | 82% | 87% | 84% |
| | RF with ReliefF 15.94sec* | 86% | 86% | 86% |
| | RF with multiple estimators 11.67sec* | 84% | 90,3% | 87% |
| | RF with wv 1.78sec* | 84% | 92,3% | 88% |
| Greicius *et al.* | | 85% | 77% | 81% |
| Shi-Jiang *Li et al.* | | 80% | 90% | 83% |

* running time of the method

## IV. CONCLUSIONS

We proposed a supervised method to assist the diagnosis and classification of the status of AD using data from an fMRI experiment. The method employs five components: preprocessing, modeling, features extraction, feature selection and classification. Different versions of RF classification algorithm were applied in order to decrease the correlation between the trees and retain their strength and consequently increase the performance of the method. The method is differentiated from those reported in the literature since it is independent from the type of the stimulus and the design of the experiment; it is not limited in detecting how cognitive functions are differentiated between different groups but quantifies these differences. Also the classification can be conducted by using only fMRI features that express AD related changes or by fusing features form different categories, such as demographic details, head motion, activation patterns, hemodynamic features and behavioral data, that depict as much as possible medical knowledge about AD. The method was evaluated using three different datasets described in Section III. The method achieved to classify a patient as healthy or demented with 88% accuracy while the methods reported in the literature conducted classification from 81% - 83%. Furthermore, the proposed method moved one step forward and succeed in classifying the status of the disease with 80,5% accuracy in the case of three class problem and 87% accuracy in the case of four class problem. These results were higher in the case that features from different categories were fused. Although the achieved sensitivity, specificity and accuracy are quite high the improvement or the implementation of a different weighted voting procedure, the selection of several attribute evaluation measures, the utilized estimators and their combination, and finally the possibility to combine the different versions of RF algorithm will be addressed in the future.

## ACKNOWLEDGMENTS

The authors thank Randy Buckner and his colleagues for making their data available via the fMRI Data Center. Information dissemination of this work was supported by the European Union in the framework of the project "Support of Computer Science Studies in the University of Ioannina" of the Operational Program of Education and Initial Vocational Training" of the 3rd Community Support Framework of the Hellenic Ministry of Education, funded by national sources and by the European Social Fund (ESF).

## REFERENCES

[1] DB. Carr, A. Goate, D. Phil, J.C. Morris, "Current concepts in the pathogenesis of Alzheimer's disease," *Am. J. Med.*, vol. 103 (suppl 3A), pp. 3S-10S, 1997.

[2] J.R. Petrella, R.E. Coleman, P.M. Doraiswamy, "Neuroimaging and early diagnosis of Alzheimer Disease: A Look to the Future," *Radiology,* vol. 226, pp. 315-336, 2003.

[3] K.R. Thulborn, C. Martin, J.T. Voyvodic, "Functional MR imaging using a visually guided saccade paradigm for comparing activation patterns in patients with probable Alzheimer's disease and in cognitively able elderly volunteers," *AJNR Am. J. Neuroradiol.*, vol. 21, pp. 524-531, 2000.

[4] S.S. Bassett *et al.*, "Familial risk for Alzheimer's disease alters fMRI activation patterns," *Brain,* vol. 129, pp. 1229-1239, 2006.

[5] G. Gronn, D. Bittner, B. Schmitz, A.P. Wunderlich, M.W. Riepe, "Subjective memory complaints: objective neural markers in patients with Alzheimer's disease and major depressive disorder," *Ann. Neuro.,* vol. 51, pp. 491-498, 2002.

[6] M. Grossman *et al.*, "Neural basis for verb processing in Alzheimer's disease: an fMRI study," *Neuropsychology,* vol. 17, pp. 658-674, 2003.

[7] C. Lustig *et al.*, "Functional deactivations: change with age and dementia of the Alzheimer type," *Proc. Natl. Acad. Sci. USA,* vol. 100, pp. 14504-14509, 2003.

[8] R.A. Sperling *et al.*, "FMRI studies of associative encoding in young and elderly controls and mild Alzheimer's disease," *J. Neurol. Neurosurg, Psychiatry,* vol. 74, pp. 44-50, 2003.

[9] S.J. Li *et al.*, "Alzheimer Disease: Evaluation of a functional MR imaging index as a marker," *Radiology,* vol. 255, pp. 253-259, 2002.

[10] M.D. Greicius, G. Srivastava, A.L. Reiss, V. Menon, "Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI," *PNAS,* vol. 101, pp. 4637-4642, 2004.

[11] R.J. Buckner, A.Z. Snyder, A.L. Sanders, M.E. Raichle, J.C. Morris, "Functional brain imaging of young nondemented, and demented older adults," *J. Cogn. Neurosc.,* vol. 12, pp. 24-34, 2000.

[12] J. P. Petrella *et al.*, "Cortical Deactivation in Mild Cognitive Impairment: High-Field Strength Functional MR Imaging", *Radiology,* vol. 245(1), pp.224-235, 2007.

[13] P. Jezzard, P. M. Matthews, S. M. Smith, "Functional MRI: An Introduction to Methods", Oxford University Press, USA, 2001.

[14] E. E. Tripoliti *et al.*, "A supervised method to assist the diagnosis of Alzheimer's Disease based on functional magnetic resonance imaging", *Proceeding of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society in conjunction with the Biennial Conference of the French Society of Biological and Medical Engineering (SFGBM)* pp. 3426-3429, 2007

[15] L. Yu, H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," *Proceedings of the Twentieth International Conference on Machine Learning,* pp. 856-863, 2003.

[16] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, pp. 5-32, 2001.

[17] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining", Addison Wesley Higher Education, 2006.

[18] Marko Robnik-Sikonja, "Improving Random Forests", *Proceedings of the European Conference on Machine Learning,* pp. 359-369, 2004.

[19] Igor Kononeko, "Estimating Attributes: Analysis and Extensions of RELIEF", *Proceedings of the European Conference on Machine Learning,* pp. 171-182, 1994.

[20] Igor Kononeko, "On Biases in Estimating Multi-Valued Attributes", *Proceedings of IJCAI-95*, pp. 1034-1040, 1995.