34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Identification of Mild Alzheimer's Disease through automated classification of structural MRI features*

Stefano Diciotti[1], *Member, IEEE*, Andrea Ginestroni[1], Valentina Bessi[2], Marco Giannelli[3], Carlo Tessa[4], Laura Bracco[2], Mario Mascalchi[1] and Nicola Toschi[5]

Abstract— **The significant potential for early and accurate detection of Alzheimer's disease (AD) through neuroimaging data is becoming increasingly attractive in view of the possible advent of drugs which are able to modify or delay disease progression. In this paper, we aimed at developing an effective machine learning scheme which leverages structural magnetic resonance imaging features in order to identify and discriminate individuals affected by mild AD on a single subject basis. Selected features included one- and two-way combinations of subcortical and cortical volumes as well as cortical thickness and curvature of numerous brain regions which are known to be vulnerable to AD. Additionally, several feature combinations were fed into support vector machines (SVMs) as well as Naïve Bayes classifiers in order to compare scheme accuracy. The most efficient combination of features and classification scheme, which employed both subcortical and cortical volumes feature vectors and a SVM classifier, was able to distinguish mild AD patients from healthy controls with 86% accuracy (82% sensitivity and 90% specificity). While this investigation is of preliminary nature, and further efforts are currently underway towards automated feature selection, best classifier determination and parameter optimization, our results appear very promising in terms of automated high-accuracy discrimination of disease stages which cannot easily be distinguished though routine clinical investigation.**

I. INTRODUCTION

Alzheimer's disease (AD) is the most common form of dementia, characterized by progressive brain atrophy as a consequence of the accumulation of abnormal proteins such as amyloid-beta and hyperphosphorilated tau [1]. Neurodegeneration begins with medial temporal lobe atrophy and progresses with the involvement of the neocortical structures [2].

An increased interest in the early detection of AD and the development of accurate markers of disease is prompted by the possible advent disease-modifying/delaying drugs [1]. Early, reliable identification and classification of AD patients is critical in planning therapeutic intervention and would allow a significant improvement of treatment efficacy and hence outcome in terms of both social cost and patient quality of life [3]. Further, amnestic mild cognitive impairment (MCI) is an intermediate state between healthy subjects and AD in which manifest memory disturbance is not accompanied by dementia [4]. However, clinical diagnostic criteria are not easily able to distinguish mild AD from MCI and normal aging [4].

Several structural and functional magnetic resonance (MR) techniques have been proposed as noninvasive imaging biomarkers for the evaluation of disease progression and early diagnosis of AD, and mainstream approaches for assessment of these biomarkers allow the study of differences between groups (e.g. diseased vs control group). However, such approaches are not applicable on a single subject level and hence do not improve diagnostic potential *per se* in a clinical setting. In order to overcome this issue, machine learning techniques have recently been identified as promising tools in neuroimaging data analysis, and such techniques are also able to concomitantly contribute to the construction and validation of novel and effective disease biomarkers. For example, structural (T1-weighted) MR imaging (MRI) methods have been employed in an automated diagnosis/discrimination framework of MCI and AD through regional [5] or voxel-based approaches [6], [7], and Diffusion Tensor Imaging (DTI) methods have been explored in the automatic classification of MCI patients [8]. Notwithstanding, most studies dealt with self-evident AD patients rather than mild AD, i.e. the early stage of the disease. The aim of this paper was to develop a dedicated and effective machine learning scheme for single-subject-level identification and classification of individuals affected by mild AD. Since the commonly and routinely available MRI modality is T1-weighted imaging, we chose to base out framework exclusively on T1-based MRI features.

II. SUBJECTS AND METHODS

A. Subjects

This study included 29 healthy controls (age 71.9±6.1, Mini Mental State Examination (MMSE) 29.4±0.9) and 51 consecutive patients - self-reporting a memory complaint - who were referred for baseline evaluation to the Centre for Alzheimer's Disease and Adult Cognitive Disorders of the Neurological Department of our Institution. Subjects were classified on the basis of a thorough clinical history as

*This work was not supported by any organization

[1]S. Diciotti, A. Ginestroni and M. Mascalchi are with the Computational Biomedical Imaging Laboratory, Department of Clinical Pathophysiology, University of Florence, Florence, Italy. stefano.diciotti at unifi.it, aginestroni at gmail.com, m.mascalchi at dfc.unifi.it

[2]V. Bessi and L. Bracco are with Azienda Ospedaliero-Universitaria Careggi, Florence, Italy. valentina.bessi at tin.it, bracco at neuro.unifi.it

[3]M. Giannelli is with the Unit of Medical Physics, Azienda Ospedaliero-Universitaria Pisana, Pisa, Italy. m.giannelli at ao-pisa.toscana.it

[4]C. Tessa is with the Versilia Hospital, Azienda USL 12 Viareggio, Lido di Camaiore (LU), Italy. c.tessa at sirm.org

[5]N. Toschi is with the Department of Biopathology and Imaging Diagnostics, University of Rome "Tor Vergata", Rome, Italy. toschi at med.uniroma2.it

well as neurological and neuropsychological examinations as affected by mild probable AD (21 subjects, age 74.4±6.9, MMSE 25.7±2.5) or MCI (30 subjects, age 68.8±7.8, MMSE 27.7±1.8). The diagnosis of AD was made according to a standardized protocol [9] that satisfies the NINCDS-ADRDA criteria [10]. MCI diagnosis was based on unanimously adopted criteria [11]. Each subject provided a written informed consent, and the study was approved by the local ethics committee.

B. Image acquisition

All subjects were examined on a clinical 1.5 T system (Intera, Philips Medical System, Best, The Netherlands) with 33 mT/m gradients strength and a 6-channel head coil technology. After the initial scout sequence, the examination protocol included a sagittal T1-weighted 3D turbo gradient echo sequence [repetition time (TR) = 8.1 ms, echo time (TE) = 3.7 ms, flip angle = 8°, inversion time (TI) = 764 ms, field of view (FOV) = 256x256 mm, matrix size = 256 × 256, 160 contiguous slices, slice thickness = 1 mm, acceleration factor (SENSE) = 2].

C. Automated cortical and subcortical segmentation and calculation of cortical thickness and curvature

Completely automated cortical and subcortical reconstruction and volumetric segmentation of each subject's structural T1-weighted MRI scan were performed with the FreeSurfer image analysis suite (http://surfer.nmr.mgh.harvard.edu/) [12]. Briefly, this includes removal of non-brain tissue using a hybrid watershed/surface deformation procedure, automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures, intensity normalization, tessellation of the gray matter white matter boundary, automated topology correction and surface deformation following intensity gradients to optimally place the gray/white and gray/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class. Volumetric region of interests (ROIs) delineating several subcortical structures including the left and right hippocampus and amygdala and basal ganglia in each subject's native T1 space were obtained by affine transformation. Also, cortical parcellation labels were converted into volumetric regions and successively registered into in each subject's native T1 space by affine transformation (Fig. 1). Successively, we performed surface inflation and registration to a spherical atlas which utilized individual cortical folding patterns to match cortical geometry across subjects, parcellation of the cerebral cortex into units based on gyral and sulcal structure, allowing creation of a variety of surface based data including maps of curvature (Fig. 2) and sulcal depth. This method uses both intensity and continuity information from the entire three dimensional MR volume in segmentation and deformation procedures to produce representations of cortical thickness, calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface.

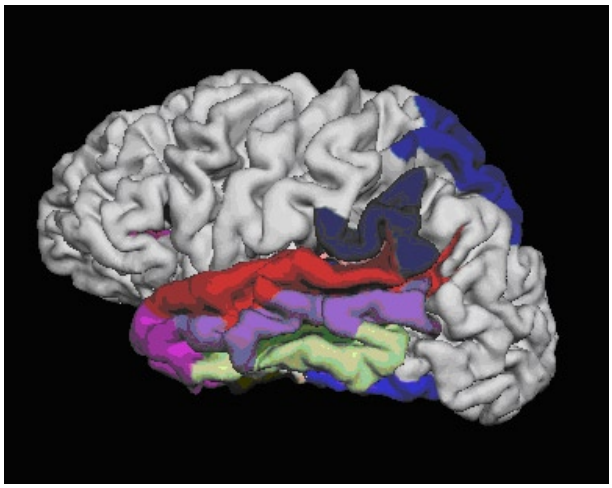

Fig. 1. Lateral view of the result of cortical parcellation highlighting the cortical regions preselected for classification.

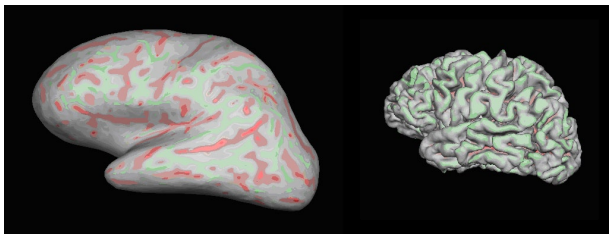

Fig. 2. Lateral view of the mean curvature (green-red color scale) superimposed on the pial surface (right) and inflated pial surface (left).

D. Features definition

Since structural biomarkers of early AD typically concern specific subcortical (hippocampus and amygdala) and cortical (temporo-parietal lobes) regions [1], we constrained our analysis *a priori* to these areas. Accordingly, 2 subcortical and 25 cortical ROIs (Fig. 2 and Fig. 3) were selected for each hemisphere from the output of the cortical parcellation and subcortical segmentation streams, resulting in a total of 4 subcortical and 50 cortical ROIs. Each subcortical and cortical ROI was described by its volume. Additionally, each cortical ROI was characterized by mean cortical thickness and mean cortical curvature. Both volume and thickness of several cortical regions are known to be reduced in AD patients due to brain atrophy, and a reduction of cortical curvature has also been observed, possibly as a result of increased sulcal widening [13]. In order to eliminate possible confounds in volume measurements due to differences in overall head size [5], the volume of each subcortical and cortical ROI was normalized to the total intracranial volume (TIV), defined as the total volume of whole-brain gray-matter, white matter and cerebrospinal fluid in ventricular and subarachnoid spaces. In order to be able to separate the efficacy of subcortical and cortical ROI descriptors in identifying and discriminating mild AD individuals, we considered four different feature vectors constituted of 1) normalized volumes of predefined subcortical structures, 2) normalized volumes of predefined cortical ROIs, 3) cortical

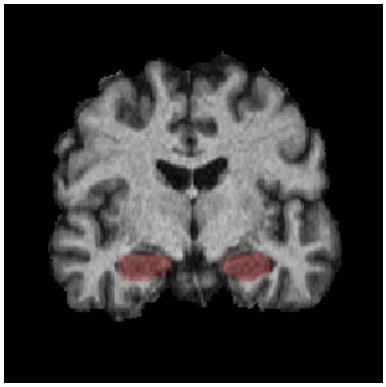

Fig. 3. Coronal view of bilateral hippocampal segmentation (red color) superimposed on a T1-weighted image.

thickness of predefined cortical ROIs and 4) mean curvature of predefined cortical ROIs.

E. Classification

In addition to studying the potential of the four structural MRI feature vectors in identifying mild AD individuals, we compared two popular machine learning classifiers: a Support Vector Machine (SVM) and a Naïve Bayes classifier [14]. The SVM was parameterised with a radial basis function kernel (width set to 0.01) and complexity parameter c equal to 1. Two different classification tasks were carried out: 1) mild AD vs healthy subjects and 2) mild AD vs MCI patients. For each classification task, we separately employed each feature vector as well as all possible two-way combinations of feature vectors. In the former case, each feature vector was fed into the classifiers separately; in the latter case, every pair of feature vectors were merged before being fed into the classifiers. In all experiments we adopted a 10 folds cross-validation scheme and performance evaluation through sensitivity, specificity and accuracy measurements. All machine learning tests were performed using the publicly available WEKA software package (http://www.cs.waikato.ac.nz/ml/weka, version 3.6.6) [14].

III. RESULTS

A. Mild AD vs healthy controls

The most accurate classification (74% accuracy) amongst the one-way feature vector analyses was obtained when using the cortical volume feature vector in conjunction with the SVM classifier (67% sensitivity and 80% specificity). Detailed results for all tests are reported in Table I. The two-way combinations of feature vectors yielded a substantial improvement of overall performance, reaching 86% accuracy (82% sensitivity and 90% specificity) for the pair composed of subcortical and cortical volumes feature vectors in conjunction with the SVM classifier (Table II).

B. Mild AD vs MCI

Amongst all one-way feature vector analyses, subcortical volumes ranked highest in the ability to discriminate between mild AD and MCI (74% accuracy, 67% sensitivity and 80% specificity in conjunction with the Naïve Bayes classifier). Detailed results are listed in Table III. The two-way combinations only afforded a slight improvement; specifically, the subcortical and cortical volumes feature vector pair was able to identify mild AD patients with 75% accuracy, 73% sensitivity and 77% specificity (Table IV).

TABLE I
ONE-WAY FEATURE CLASSIFICATION BETWEEN HEALTHY CONTROLS AND AD PATIENTS

Feature vector	SVM			Naïve Bayes		
	Sens	Spec	Accu	Sens	Spec	Accu
Subcort volumes	62%	80%	72%	62%	77%	70%
Cort volumes	67%	80%	74%	67%	72%	70%
Cort thickness	62%	80%	72%	55%	75%	68%
Cort mean curvature	45%	77%	64%	32%	77%	58%

TABLE II
TWO-WAY FEATURE CLASSIFICATION BETWEEN HEALTHY CONTROLS AND AD PATIENTS

Feature vector pair	SVM			Naïve Bayes		
	Sens	Spec	Accu	Sens	Spec	Accu
Subcort & cort volumes	82%	90%	86%	72%	83%	78%
Subcort volumes & cort thickness	72%	80%	77%	65%	83%	76%
Subcort volumes & cort mean curvature	62%	80%	72%	52%	80%	68%
Cort volumes & cort thickness	60%	77%	70%	65%	78%	74%
Cort volumes & mean curvature	62%	73%	68%	47%	83%	68%
Cort thickness & mean curvature	62%	77%	70%	40%	80%	64%

TABLE III
ONE-WAY FEATURE CLASSIFICATION BETWEEN MCI AND AD PATIENTS

Feature vector	SVM			Naïve Bayes		
	Sens	Spec	Accu	Sens	Spec	Accu
Subcort volumes	58%	83%	73%	67%	80%	74%
Cort volumes	72%	73%	72%	68%	77%	73%
Cort thickness	53%	63%	59%	52%	67%	61%
Cort mean curvature	53%	80%	69%	33%	87%	64%

IV. DISCUSSION

In this report, we described a machine learning approach for identification of mild AD individuals based on structural features extracted exclusively from T1-weighted MRI.

We used a hypothesis-driven selection of anatomical ROIs based on established knowledge of early localized neurodegeneration in AD, and evaluated the discriminating ability of subcortical and cortical volumes along with cortical thickness and curvature through both SVM and Naïve Bayes classifiers.

TABLE IV
TWO-WAY FEATURE CLASSIFICATION BETWEEN MCI AND AD PATIENTS

Feature vector pair	SVM			Naïve Bayes		
	Sens	Spec	Accu	Sens	Spec	Accu
Subcort & cort volumes	72%	77%	74%	73%	77%	75%
Subcort volumes & cort thickness	67%	73%	70%	67%	63%	65%
Subcort volumes & cort mean curvature	67%	80%	74%	43%	83%	67%
Cort volumes & cort thickness	63%	73%	69%	67%	73%	70%
Cort volumes & mean curvature	63%	77%	71%	53%	80%	69%
Cort thickness & mean curvature	58%	73%	67%	38%	80%	63%

The feature vectors which proved most useful were cortical volumes in mild AD vs healthy subjects and subcortical volumes in mild AD vs MCI patients. Nevertheless, both volume vectors showed similarly high performance, and the feature pair resulting from their combination proved to be the most accurate discriminant amongst all two-way feature combinations in both classification tasks.

While this investigation is of preliminary nature, and further efforts are currently underway towards automated feature selection, best classifier determination and parameter optimization, our results on identification of mild AD patients appear very promising. Interestingly, discrimination of AD from healthy controls generally accomplished accuracies which range from 58% to 100% [15]. One of the main factors underlying this large variability is the heterogeneity of AD populations under study. To this end, our patient base included exclusively mild AD patients, i.e. a very early disease stage which long precedes self-evident AD in which brain atrophy is macroscopically visible. Kloppel et al. dealt with mild AD versus healthy controls separation reaching 81.1% accuracy (60.6% sensitivity, 93.0% specificity) by using a voxel-based approach. An improvement to 85.6% accuracy (75.8% sensitivity, 91.2% specificity) was also observed when the analysis was restricted to the medial temporal lobe region.

V. CONCLUSIONS

This report demonstrates the effectiveness of machine learning techniques, which allow patient classification on a single subject basis, for the identification of mild AD patients using structural MRI measurements. Future developments will target the application of these techniques to multimodal MR imaging (diffusion imaging, functional MRI, etc.) features with the prospect of a further overall improvement of early detection of AD.

REFERENCES

[1] G. B. Frisoni, N. C. Fox, P. Scheltens, and P. M. Thompson. The clinical use of structural MRI in Alzheimer disease. *Nat Rev Neurol*, 6(2):67–77, 2010.

[2] H. Braak and E. Braak. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol*, 82(4):239–259, 1991.

[3] R. Williams. Biomarkers: warning signs. *Nature*, 475(7355):S5–S7, 2011.

[4] C. R. Jack Jr, M. S. Albert, D. S. Knopman, G. M. McKhann, R. A. Sperling, M. C. Carrillo, B. Thies, and C. H. Phelps. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroup on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*, 7(3):257–262, 2003.

[5] R. S. Desikan, H. J. Cabral, C. P. Hess, W. P. Dillon, C. M. Glastonbury, M. W. Weiner, N. J. Schmansky, D. N. Greve, D. H. Salat, R. L. Buckner, and Bruce Fischl; Alzheimer's Disease Neuroimaging Initiative. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. *Brain*, 132:2048–2057, 2009.

[6] 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*, 131:681–689, 2008.

[7] B. Magnin, L. Mesrob, S. Kinkingnhun, M. Plgrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehricy, and H. Benali. Support vector-machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology*, 51:73–83, 2009.

[8] L. O'Dwyer L, F. Lamberton, A. L. Bokde, M. Ewers M, Y. O. Faluyi, C. Tanner, B. Mazoyer B, D. O'Neill D, M. Bartley, D. R. Collins, T. Coughlan, D. Prvulovic, and H. Hampel. Using support vector machines with multiple indices of diffusion for automated classification of mild cognitive impairment. *PLoS One*, 7(2):e32441, 2012.

[9] L. Bracco and L. Amaducci. Italian Multicenter Study on Dementia: a protocol for data collection and clinical diagnosis of Alzheimer's disease. the SMID group. *Neuroepidemiology*, 11(1):39–45, 1992.

[10] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Proce, and E. M. Stadlan. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34(7):939–944, 1984.

[11] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen. Mild cognitive impairment: a clinical characterization and outcome. *Arch Neurol*, 56(3):303–308, 1999.

[12] B. Fischl. Freesurfer (accepted for publication). *Neuroimage*, to be published (available online).

[13] K. Im, J-M. Lee, S. W. Seo, S. H. Kim, S. I. Kim, and D. L. Na. Sulcal morphology changes and their relationships with cortical thickness and gyral white matter volume in mild cognitive impairment and Alzheimer's disease. *Neuroimage*, 43:103–113, 2008.

[14] E. Frank I. H. Witten and M. A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Third Edition. Morgan Kaufmann, Burlington, MA, USA, 2011.

[15] G. Chetelat and J. C. Baron. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. *Neuroimage*, 18(2):525–541, 2003.