

# Universal Score of Structural Abnormality in Alzheimer's Disease

V. Zagorodnov<sup>1</sup>, and O. Murphy<sup>1</sup>

<sup>1</sup>Nanyang Technological University, Singapore, NA, Singapore

**Introduction.** In-vivo structural brain measurements using magnetic resonance imaging (MRI) have already been established as effective diagnostic tool for Alzheimer's disease (AD) [1]. Of particular importance to diagnosis are the medial temporal regions, such hippocampus and entorhinal cortex, which start to undergo atrophy even before any symptoms can be detected, and other brain areas become affected as the disease progresses. While early techniques focused solely on hippocampal and entorhinal measurements, most recent approaches utilize the whole brain measurements [2], using a pattern classification framework that is common to most methods. In this framework, each subject's brain structure is first parameterized as a multi-dimensional feature vector  $X = (x_1, x_2, \dots, x_N)$  and then a training set is used to learn an optimal classifier or regressor. The elements of this feature vector vary between approaches and could mean gray matter densities of individual voxels, regional cortical thicknesses, subcortical volumes or ROI measurements. For most common linear classifiers, the learning implies estimating classifier weights  $w_0, w_1, w_2, \dots, w_N$ , which are then used to combine all feature vector elements into the classifier score:  $score = w_0 + w_1x_1 + w_2x_2 + \dots + w_Nx_N$ . The actual usage of this score varies between applications. For classification tasks, such as distinguishing between AD patients and normal controls (NC) or prediction of which of the subjects with mild cognitive impairment (MCI) will progress to AD (MCI-converters vs. MCI-non-converters), the classification score is compared to some predefined threshold (typically set to zero) and the sign of this comparison is used for classification. For prediction tasks, for example when the brain structure is used to predict a cognitive test score, the classifier score is used combined with the slope and offset, learned from the regression line between the scores. Typically, a new classifier score is learned for every new application, which makes it difficult to interpret it in terms of disease progression. We show that a single score learned on the basis of AD vs. NC classification can be reused for other tasks and thus be interpreted as a universal score of AD progression. Our score is based on regional cortical and subcortical measurements obtained by running FreeSurfer pipeline [3]. In terms of classification and prediction accuracy, our score performs equally well or better than ten methods that were compared in a recent study [2].

**Materials and methods.** All the data used in this article was obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (<http://www.loni.ucla.edu/ADNI>). The data set consisted of 135 AD, 161 NC and 209 MCI subjects. MCI participants were separated into 76 MCI Converters, those who were classified as having undergone conversion to AD based on changes in Global CDR of 0.5 to 1. The remaining 133 MCI participants were classified as non-converters. Segmentation and labeling of the structures were performed using FreeSurfer V4.2.0 segmentation pipeline. Mean thicknesses of 69 cortical regions, 36 subcortical region volumes as well as whole brain, CSF and WM volumes resulting in 113 aggregate measures were used as features for the derived score. Since the dimensionality of the data is similar to the number of the training samples for this data set, we used classifier based on Support Vector Machine (SVM) to avoid overfitting. The weights of the SVM classifier were trained on AD vs. NC classification task, using a subset of AD and NC subjects as a training set, and the rest of the subjects for validation. The training and validation sets were exactly the same as those used in ref. [2], to allow direct comparison between the results. The same score was then applied to MCI converters vs. non-converters classification task and prediction of MMSE and ADA-cog test scores. In each of these tasks, performance of the proposed universal score was compared with scores that were individually tuned to these tasks.

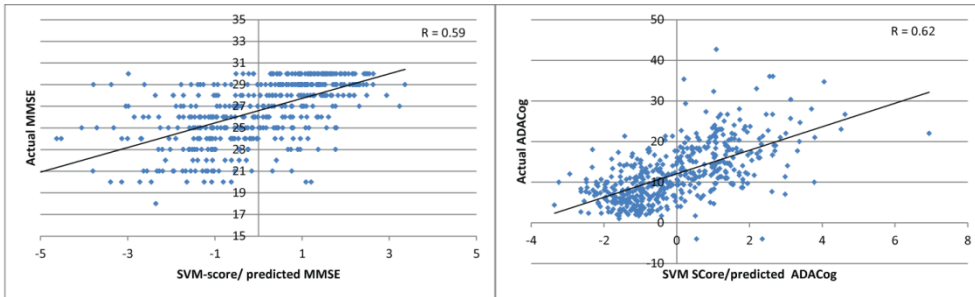


Fig. 1. 1 Relationship between structural SVM score and MMSE (left) and ADA-Cog (right)

**Results.** Our classifier achieved 87.3% classification accuracy (92% specificity and 85% sensitivity) on AD vs. NC classification task, which was similar to the best performance reported in the comparison study [2] (95% specificity and 0.81% sensitivity). When the same classifier was applied to the classification of MCI converters vs. non-converters, it achieved 72.1% accuracy (73% specificity and 70% sensitivity). This was better than the best result in the comparison study [2] (67% specificity and 62% sensitivity). Note that we had to adjust the

decision threshold to reflect the fact the structural abnormality in MCI subjects is less than that in AD patients. When the classifier was retrained using only MCI subject data, the accuracy became much lower (62.5%). This reduced accuracy is likely due to the fact that only 39 subjects from MCI converter class that were used for training. This number is much smaller than the dimensionality of the data (N=113), which can cause poor generalization performance, even for an SVM classifier. Finally, we evaluated the effectiveness of our classifier score in predicting the cognitive test scores. Our structural classifier score was highly correlated with both MMSE ( $r=0.59$ ) and ADA-cog ( $r=0.62$ ), see Fig. 1. These correlations were better than those reported in study [5],  $r=0.48$  for MMSE and  $r=0.57$  for ADA-cog. Retraining the score just for this task using a Support Vector Regression (SVR) did not change the correlations.

**Conclusions.** We proposed a single universal score of structural abnormality that can be interpreted as a measure of AD progression. The score was trained on AD vs. NC classification task using ADNI data set and performed equally well or better than the ten approaches that has recently been compared [2]. The score's performance did not improve (and actually became worse) when it was retrained for each specific task. This result agrees with growing evidence of the presence of significant AD pathology in many MCI patients [4], albeit at a smaller scale. It also suggests a direct relationship between the amount and regional distribution of gray matter atrophy and the decline in cognitive test scores.

**References.** [1] Vemuri, P. et al, Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. *Neuroimage*, 2008. 39(3): p. 1186-97. [2] Cuingnet, R. et al, Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. *Neuroimage*, 2010, *in press*. [3] Dale, A.M. et al, 1999. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage* 9, 179-194. [4] Kempainen, N.M., et al, PET amyloid ligand [11C] PIB uptake is increased in mild cognitive impairment. *Neurology*, 2007. 68(19): p. 1603-6. [5] Stonnington C. M., et al, Predicting clinical scores from magnetic resonance scans in Alzheimer's disease." *Neuroimage* 2010, 51(4): 1405-1413.