

Optimized Cortical Subdivision for Classification of Alzheimer's Disease With Cortical Thickness

Mirco Richter, Dorit Merhof

Visual Computing, University of Konstanz
mirco.richter@uni-konstanz.de

Abstract. In several studies, brain atrophy measured by cortical thickness has shown to be a meaningful biomarker for Alzheimer's disease. In this research field, the level of granularity at which values are compared is an important aspect. Vertex- and voxel-based approaches can detect atrophy at a very fine scale, but are susceptible to noise from misregistrations and inter-subject differences in the population. Regional approaches are more robust to these kinds of noise, but cannot detect variances at a local scale. In this work, an optimized classifier is presented for a parcellation scheme that provides a trade-off between both paradigms by increasing the granularity of a regional approach. For this purpose, atlas regions are subdivided into gyral and sulcal parts at different height levels. Using two-stage feature selection, optimal gyral and sulcal subregions are determined for the final classification with sparse logistic regression. The robustness was assessed on clinical data by 10-fold cross-validation and by testing the prediction accuracy for unseen individuals. In every aspect, superior classification performance was observed as compared to the original parcellation scheme which can be explained by the increased locality of cortical thickness measures and the customized classification approach that reveals interacting regions.

1 Introduction

Various image processing and classification methods to diagnose Alzheimer's disease (AD) and mild cognitive impairment (MCI) have been developed and gradually improved over the past decades. With dedicated segmentation, registration and feature extraction methods, biomarkers based on structural and functional imaging can detect subtle neurodegenerative changes [1]. As a result, the amount of observed variance between health groups is increasing for the sake of improved discrimination. All these developments are important for an early classification and treatment of MCI and AD [2]. Amongst others, the structural biomarker cortical thickness (CoT) that is based on segmented data from magnetic resonance imaging (MRI) has shown to be a significant biomarker for dementia. Regarding MCI and AD, gray matter structures that are affected by atrophy include the hippocampus, the parahippocampal gyrus, the cingulate, parts of the temporal, parietal and frontal lobe, and the occipital pole [3].

For the estimation of CoT, there exist two competing paradigms, the voxel- and the surface-based approach [4, 5]. In both cases, measures can be compared

at different levels of granularity: fine-grained methods using voxel and vertex values or coarse-grained methods using mean values of anatomically distinct regions of interest (ROIs). Fine-grained methods provide high dimensional feature sets for the detection of local changes. However, they are susceptible to noise introduced e.g. by registration to normal space or inter-subject differences. Also, the relatively small set of observations compared to the large number of variables renders dimensionality reduction a crucial problem. In contrast, region-based approaches create low dimensional spaces that are easier to handle and averaging of measurements compensates for noise such as subtle misregistrations. However, as a consequence changes at a local scale cannot be revealed.

To close the gap between both extremes of the scale of granularity, an anatomical parcellation method was presented previously that subdivides the cortex into gyral and sulcal subregions at different levels of height for which mean CoT values are computed [6]. In this work, this parcellation scheme is employed by a customized feature selection and classification approach that determines the optimal set from the original ROIs and the gyral and sulcal subregions. Section 2 describes the subject population and the imaging data, presents our approach including a summary of preprocessing steps and CoT measurement, and describes the classification approach and the evaluation of robustness. In Sections 3 and 4, we present and discuss the improvements achieved on clinical data.

2 Materials and methods

2.1 Image data and preprocessing

For this study, MRI images of 84 subjects were collected from the database of the Alzheimer’s Disease Neuroimaging Initiative (ADNI) for the following four groups (diagnoses assigned according to the ADNI protocol available at <http://adni.loni.ucla.edu>): healthy elderly controls N (11 females / 10 males; age: 76.4 ± 6.6 ; education: 16.4 ± 2.8), patients with early MCI (eMCI) (10/11; age: 74.2 ± 8.0 ; edu.: 16.6 ± 2.8), late MCI (lMCI) (8/13; age: 72.3 ± 5.5 ; edu.: 16.7 ± 2.7), and AD (8/13; age: 75.4 ± 10.2 ; edu.: 15.0 ± 3.2). One-way ANOVA ensured that sex, age and education were not significantly different between the groups. For all subjects, MRI data was acquired on 3T GE Signa HDxt scanning devices in the same format, but from different sites. The distribution of sites was random across subjects. An IR-SPGR sequence was applied with TR = 6.98 ms, TE = 2.85 ms, and TI = 400 ms, resulting in 196 T1-weighted sagittal slices with 1.0×1.0 mm in plane spatial resolution of dimension 256×256 voxels and 1.2 mm slice thickness.

The images were preprocessed in the same fashion as described in [6] with tools from the FMRIB Software Library (FSL, www.fmrib.ox.ac.uk/fsl). In short, manual neck cropping was followed by automatic skull stripping using the Brain Extraction Tool. The extracted brain was segmented into white matter (WM), gray matter (GM) and cerebrospinal fluid using FAST that provides probabilistic voxel-wise membership values for each of the tissue classes. ROIs were defined according to the labels of the Harvard-Oxford probabilistic atlas

(distributed with FSL) which comprises 48 cortical ROIs for each hemisphere. To transform the ROI labels from the atlas to each subject space, FLIRT and FNIRT performed spatial normalization of the original images to the MNI152 space by affine multi-resolution registration with normalized correlation as similarity function followed by non-linear free-form deformations. The labels of the Harvard-Oxford atlas were spatially transformed to the subject space by inverting these registrations. Using the GM segmentation, voxelwise CoT was estimated using minimum line integrals measured along an approximation of the medial GM layer [4, 6].

2.2 Feature selection and classification

The anatomical subdivision of the cortex into gyral and sulcal subregions was achieved by an adaption of a robust skeletonization approach for discrete volumetric objects with genus 0 [7]. Basically, the skeleton of the WM segmentation is computed that associates each skeleton point with at least two closest points on the WM surface. By measuring geodesic distances τ between those surface points, a pruning function is defined that can be thresholded to separate gyri and sulci at different height levels. An extension of this pruning function prevents the false detection of noisy features as gyral regions. Using thresholds $\tau \in [8, 20]$ mm in steps of 2 mm, seven gyral and seven sulcal subregions are created per ROI in addition to the ROI itself. These are denoted as \mathcal{G}_τ , \mathcal{S}_τ and \mathcal{W} , respectively. For the whole brain, this results in 15 regions per ROI times the 96 ROIs of the Harvard-Oxford atlas, each represented by a mean CoT value. For a collection of n subjects, T_{ij} denotes the feature vector containing the mean CoT values for all subjects for the cortex label $i \in [1, 96]$ and the subregion $j \in [1, 15]$. Each T_{ij} was corrected for the covariates sex, age, and education by a linear regression model estimated for the normal group and normalized to z-scores.

As rather subtle differences in mean CoT are expected between the groups, and as the subregions of each ROI may contain redundant information, direct application of sparse logistic regression or a support vector machine might not be sufficient to maximize the gain of this rich feature set. Therefore, we propose a two-stage feature selection before the final classification, which comprises the following steps: The first stage eliminates redundancy in each ROI by determining for the gyral and sulcal subregions \mathcal{G}_τ and \mathcal{S}_τ those that maximize the Pearson's correlation of their mean CoT values T_{ij} with the group membership variable Y . The resulting set T'_{ij} with $i \in [1, 96]$ and $j \in [1, 3]$ contains the mean CoT values of the whole ROI, the best gyral and the best sulcal subregion.

Within the second stage, interactions between pairs of variables of the set T'_{ij} were considered as the selection criterion since enforcing interactions might result in higher stability and better performance of the final classification. Similar to the linear discrimination for pairs of mean CoT values in [3], linear regression was applied to optimize the linear model $Y = [T'_{ij} \ T'_{lm}] \beta + \epsilon$ for each pair of variables T'_{ij} and T'_{lm} with $(i, j) \neq (l, m)$. Here, β and ϵ denote the regression weights and error variables. Ranking each variable pair according to the coefficients of determination R^2 , the best variables are filtered to form the set \mathcal{D}_{LR} . For

Table 1. AUC values of 10-fold cross-validation and prediction for the whole ROI set \mathcal{D}_W , the best gyral and best sulcal set $\mathcal{D}_{G\tau}$, $\mathcal{D}_{S\tau}$, and the sets $\mathcal{D}_{\text{Corr}}$ and \mathcal{D}_{LR} optimized by correlation and linear regression. First and second best values in bold.

Group Pair	10-fold Cross Validation					Prediction				
	\mathcal{D}_W	\mathcal{D}_G	\mathcal{D}_S	$\mathcal{D}_{\text{Corr}}$	\mathcal{D}_{LR}	\mathcal{D}_W	\mathcal{D}_G	\mathcal{D}_S	$\mathcal{D}_{\text{Corr}}$	\mathcal{D}_{LR}
N-eMCI	.74	.58	.59	.80	.87	.64	.58	.50	.70	.81
N-IMCI	.69	.52	.44	.65	.83	.58	.47	.54	.67	.75
N-AD	.93	.93	.86	.97	.99	.87	.89	.88	.94	.96
eMCI-IMCI	.53	.48	.70	.62	.93	.50	.51	.67	.65	.82
eMCI-AD	.84	.93	.89	.81	.99	.70	.78	.79	.75	.90
IMCI-AD	.60	.81	.71	.72	.86	.64	.76	.72	.72	.82

comparison, $\mathcal{D}_{\text{Corr}}$ was created by selecting the best variables according to the correlation between T'_{ij} and Y , similar to the approach in [8]. For both sets, their size was fixed to 96, equal to the size of the original partitions.

For the six group pairs created from the four groups, the following feature sets were compared: $\mathcal{D}_j = \{T_{ij}\}$ with $j \in [1, 15]$ which denotes the 15 sets (i.e. whole ROI, seven gyral and seven sulcal sets), $\mathcal{D}_{\text{Corr}}$ and \mathcal{D}_{LR} . For this purpose, sparse logistic regression with the elastic net (ENLR) was applied to each set [9]. The regression factor α that combines features of lasso and ridge regression was set to 0.5 for a trade-off between low and high number of features. Two tests of robustness were applied: At first, stratified 10-fold cross-validation with 20 repetitions for bias-free estimates, and second, evaluation of predictive accuracy by splitting the population into two equally sized training and test tests by random stratification and assessing mean values from 20 repetitions.

3 Results

For the first comparison using 10-fold cross-validation, curves of the receiver operator characteristic (ROC) are presented in Figure 1 (top) for the group pairs N-eMCI, N-IMCI and N-AD. The corresponding discriminative regions are highlighted on the surface of one healthy subject in Figure 1 (bottom). The mean values of the area under the ROC (AUC) are listed for both validations and for all groups in Table 1. In all cases, \mathcal{D}_{LR} shows superior performance, while $\mathcal{D}_{\text{Corr}}$ is better than all sets \mathcal{D}_j in only two cases for 10-fold cross-validation, and in three cases for prediction. The three top regions identified using \mathcal{D}_{LR} for N-eMCI are \mathcal{G}_8 in the right frontal pole, \mathcal{S}_{14} in the left central opercular cortex, and \mathcal{W} of the right planum temporale. For N-IMCI, the regions are \mathcal{S}_{20} in the right middle temporal gyrus, \mathcal{G}_8 in the left parietal operculum cortex, and \mathcal{S}_{20} in the right superior temporal gyrus. For N-AD, the regions are \mathcal{G}_{18} in the right inferior temporal gyrus, \mathcal{S}_{20} in the right inferior temporal gyrus, and \mathcal{W} of the right cingulate gyrus.

4 Discussion

In this work, a classification method for Alzheimer's disease and mild cognitive impairment is presented that optimizes the selection of mean cortical thickness variables from a set provided by a cortical parcellation scheme. For this purpose, features are selected from this set by assessing the power to predict the group membership variable Y from each pair of mean CoT variables. In this way, the contribution of interacting variables could be enforced which was observed by a strong improvement in classification power validated for robustness by two different resampling techniques.

Traditional feature selection methods consider the correlation between a single predictor and Y [8]. This is beneficial for large feature spaces that require efficient methods. However, classification performance might get impaired as interactions between multiple variables are completely ignored. In this work, this problem was solved properly with linear regression using pairs of predictors. Due to the quadratic complexity, this approach is not directly applicable to high dimensional data, but a divide-and-conquer approach could achieve at least a good approximation of the complete ranking for the sake of quality.

Acknowledgement. The data used was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.ucla.edu>), NIH grants U01 AG024904, P30 AG010129, K01 AG030514. As such, the investiga-

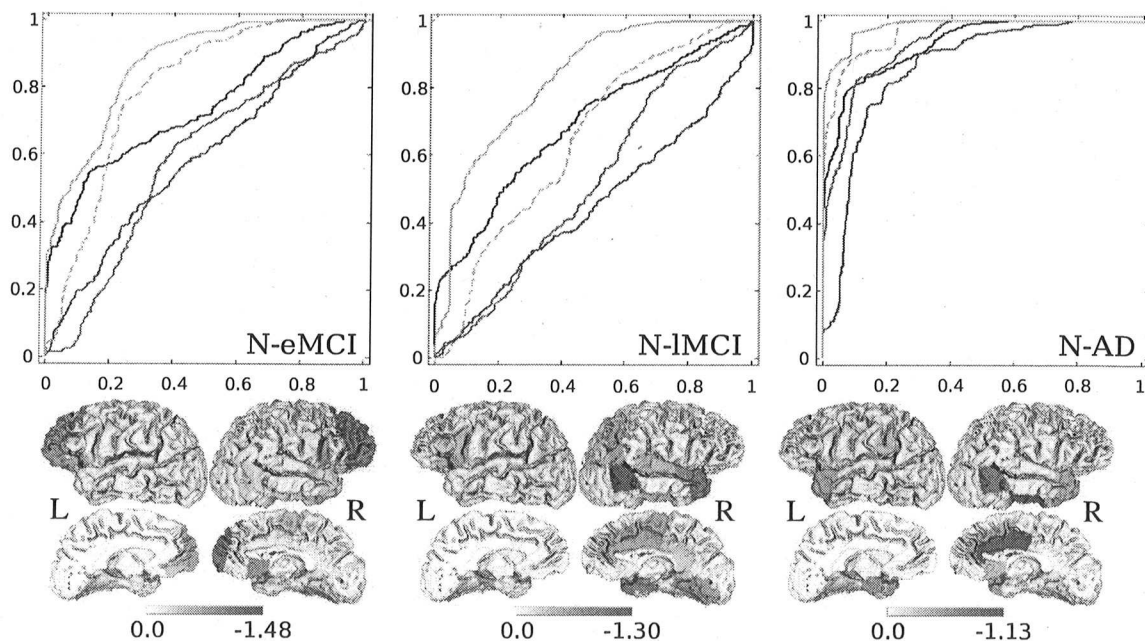


Fig. 1. *Top:* ROC curves assessed by 20-times repeated stratified 10-fold cross-validation of ENLR. D_W (black), best $D_{G\tau}$ (red), best $D_{S\tau}$ (blue), D_{Corr} (green dashed line), and D_{LR} (green solid line). False positive rate on abscissa, true positive rate on ordinate. *Bottom:* Discriminative regions projected on the surface of one healthy subject for the group pairs N-eMCI (left), N-IMCI (middle) and N-AD (right). Colors from white to red are proportional to the negative regression weight (decreasing CoT).

tors 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. A complete listing of ADNI investigators is given in the Internet¹. Furthermore, the work was supported by the DFG Research Training Group GK-1042 "Explorative Analysis and Visualization of Large Information Spaces".

References

1. Hampel H, Bürger K, Teipel SJ, et al. Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. *Alzheimer's & Dementia*. 2008;4(1):38–48.
2. Small G, Bullock R. Defining optimal treatment with cholinesterase inhibitors in Alzheimer's disease. *Alzheimer's & Dementia*. 2011;7(2):177–84.
3. Lerch JP, Pruessner J, Zijdenbos AP, et al. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. *Neurobiology of Aging*. 2008;29(1):23–30.
4. Aganj I, Sapiro G, Parikshak N, et al. Measurement of cortical thickness from MRI by minimum mine integrals on soft-classified tissue. *Hum Brain Mapp*. 2009;30(11):3188–99.
5. Fischl B, Dale A. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci USA*. 2000;97(20):11050–5.
6. Richter M, Bishop CA, Dukart J, et al. Skeleton-based gyri sulci separation for improved assessment of cortical thickness. In: *IEEE 9th International Symposium on Biomedical Imaging*; 2012. p. FR-PO.PA.37.
7. Reniers D, Jalba A, Telea A. Robust classification and analysis of anatomical surfaces using 3D skeletons. *Eurographics Workshop on Visual Computing for Biomedicine*. 2008; p. 61–8.
8. Zhou L, Wang Y, Li Y, et al. Hierarchical anatomical brain networks for MCI prediction: revisiting volumetric measures. *PLoS One*. 2011;6(7):e21935.
9. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*. 2005;67(2):301–20.