

Dementia Diagnosis using MRI Cortical Thickness, Shape, Texture, and Volumetry

Lauge Sørensen¹, Akshay Pai¹, Cecilie Anker², Ioana Balas¹, Martin Lillholm³,
Christian Igel¹, Mads Nielsen¹

¹ Department of Computer Science, University of Copenhagen, Denmark

² DTU Compute, Technical University of Denmark, Denmark

³ Biomediq A/S, Denmark

Abstract. This paper proposes a structural magnetic resonance imaging (MRI) biomarker that combines a range of individual biomarkers (cortical thickness measurements, hippocampal shape, hippocampal texture, and volumetric measurements) for the purpose of multi-class classification of Alzheimer’s disease, mild cognitive impairment and, normal controls. The combination is achieved by entering the biomarkers as features in a linear discriminant analysis. The fully automated method is trained on a combination of two publicly available datasets and is evaluated on the training set from the CADDementia challenge. Test set scores using two different priors are submitted to the same challenge.

1 Introduction

Structural magnetic resonance imaging (MRI) is an integral part of the diagnostic work flow in many memory clinics. The modality allows for non-invasive in vivo inspection of the degree and the location of brain atrophy, a hallmark of several dementias including Alzheimer’s disease (AD), the most frequent type. The importance of structural MRI has been underlined by the inclusion of MRI volumetry as a surrogate biomarker of atrophy in international diagnostic guidelines for AD [13] and its prodromal stage, mild cognitive impairment (MCI) due to AD [2].

Volumetry, and in particular hippocampal volumetry, is in general the most widely studied and used MRI biomarker of AD, and there are already efforts towards standardization of this biomarker [8, 12]. However, it is evident that there are other sources of information, than what is captured by volumetry, to extract from a structural MRI scan. This include cortical thickness [14] as well as less established biomarkers such as the hippocampal shape [1] and the textural patterns within the hippocampal tissue [15, 17]. Both shape and texture have shown to provide volume-independent diagnostic or prognostic information, and to improve prediction of conversion from MCI to AD when combined with volume [1, 15, 17]. Cortical thickness may be more reliable than volume in detecting differences between MCI and AD [14] and combining cortical thickness measurements with volumetric measurements has shown good NC vs. AD discrimination [18].

In this study, we propose to combine a range of volumetric measurements with cortical thickness measurements, hippocampal texture, and shape, in order to obtain a combination biomarker that uses more of the information contained in a structural MRI scan. Such a biomarker has potential of improved diagnosis of MCI and AD compared to, e.g., a pure volumetry-based biomarker. To the best of our knowledge, this is a unique combination of basic MRI biomarkers not tried before. The combination is achieved by entering all biomarkers as features in a linear discriminant analysis (LDA). The proposed method is developed and trained on a combination of MRI scans from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) and the MRI imaging arm of the Australian Imaging, Biomarker & Lifestyle Flagship Study of Aging (AIBL). The trained method is then evaluated on the training set from the Computer-Aided Diagnosis of Dementia based on structural MRI data challenge⁴ (CADDementia), a challenge at the 17th International Conference on Medical Image Computing & Computer-Assisted Intervention. Two different test set scores using different priors are further submitted to the CADDementia challenge.

2 Data

The following five datasets are used: the “complete annual year 2 visits” 1.5-T dataset from the collection of standardized datasets recently released by ADNI [19]; a subset of manual hippocampal segmentations from the Harmonized Hippocampal Protocol (HHP) [8] and associated MRI scans; the MRI imaging arm of AIBL [5]; and the CADDementia training and test sets. Table 1 summarizes the characteristics of the datasets.

The ADNI dataset and the AIBL dataset are merged into one combined dataset (termed ADNI+AIBL) that is used for training, and HHP is used in a special purpose hippocampal segmentation method described in Section 3.3. The CADDementia training set is used for evaluation. Finally, the CADDementia test set is classified using the trained combination MRI biomarker, and the obtained scores are submitted to the CADDementia challenge.

All MRI scans were conformed to $1 \times 1 \times 1 \text{ mm}^3$ resolution followed by bias correction. Both operations were performed using FreeSurfer (version 5.1.0, default parameters) [7].

3 Individual MRI Biomarkers

A range of structural MRI biomarkers are used in the proposed combination biomarker, which are the following: volumetry of brain structures and of the ventricles, cortical thickness measurements, hippocampal shape, and hippocampal texture. These biomarkers are detailed in the following subsections, and an overview is provided in Table 2.

⁴ <http://caddementia.grand-challenge.org/>

Table 1: Characteristics of the datasets.

		n	Age, years (mean±std)	Male (%)	MRI field strength (1.5-T/3-T)
ADNI	Total	504	75.3±6.5	57.9	504/0
	NC	169	76.0±5.1	50.9	169/0
	MCI	234	74.8±7.0	66.5	234/0
	AD	101	75.3±7.4	50.5	101/0
HHP	Total	40	74.1±7.4	47.5	40/0
	NC	12	76.9±6.2	41.7	12/0
	MCI	11	70.9±6.8	54.6	11/0
	AD	17	74.2±8.6	47.1	17/0
AIBL	Total	145	75.4±7.4	44.6	1/144
	NC	88	75.2±7.2	47.7	1/87
	MCI	29	77.5±7.1	51.7	0/29
	AD	28	73.6±8.1	35.7	0/28
CADDementia train	Total	30	65.2±7.0	43.3	0/30
	NC	12	62.3±6.3	25.0	0/12
	MCI	9	68.0±8.5	44.4	0/9
	AD	9	66.1±5.2	66.7	0/9
CADDementia test	Total	354	65.1±7.8	60.2	0/354

3.1 FreeSurfer Volumetry

Sub-cortical and ventricular volumetric measurements were computed using cross-sectional FreeSurfer (version 5.1.0, default parameters) [7]. We used measurements from ROIs provided by FreeSurfer (i.e., in the Aseg atlas). Bilateral ROIs were joined. In addition to individual ROIs, we also computed total ventricular volume and whole brain volume, resulting in a total of 7 volumetric FreeSurfer measurements. All volumetric measurements were normalized for head size by dividing by the intra-cranial volume (ICV) also computed during the cross-sectional FreeSurfer pipeline.

3.2 FreeSurfer Cortical Thickness

Cortical thickness measurements were computed using cross-sectional FreeSurfer (version 5.1.0, default parameters) [6]. We used measurements from the ROIs in the Desikan-Killiany atlas that were joined into the four lobes and the cingulate cortex⁵. Left and right hemispheres were further joined, resulting in a total of 5 cortical thickness measurements. We did not normalize cortical thickness measurements for head size (i.e., ICV) [18].

⁵ <http://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation>

3.3 Hippocampal Volume

In addition to FreeSurfer's estimate of the hippocampal volume, we also computed the hippocampal volume using a special purpose algorithm. This was motivated by the fact that hippocampal volume is the most widely used MRI biomarker of AD [12] and FreeSurfer is not optimized for this structure specifically. The left and right hippocampus were segmented separately using a multi-atlas, affine registration, non-local patch-based segmentation (N-L Patch) technique [4, 3]. The atlas comprised 40 segmentations from HHP [8] (12 NC, 11 MCI, 17 AD). All 40 HHP segmentations were used as atlases during pre-selection, but only the 9 most similar contributed to the final segmentation. A subset of 15 HHP segmentations were used to cross-validate parameters (number of atlases used after pre-selection, cubic patch size, and search volume size) using Dice's coefficient. The bilateral volume was computed and divided by FreeSurfer's estimate of ICV. N-L Patch has previously demonstrated a better AD diagnostic performance than static FreeSurfer [4].

3.4 Hippocampal Shape

Two hippocampal shape scores (for the left and right hippocampus, respectively) were computed as well. In a spirit similar to [1], a shape descriptor was computed by aligning each hippocampus surface to a template hippocampus using iterative closest point (ICP), followed by a mapping of 30 uniformly distributed landmarks from the template to the hippocampus. The set of hippocampi, each now represented by 30 landmarks, were all aligned using generalized Procrustes alignment [9]. Finally, principal component analysis was applied to the set of aligned hippocampus landmarks, and the components explaining 90 % of the variance were retained. This representation was used as features in a naive Bayes classifier. The feature extraction was performed on all data simultaneously, i.e., on the combination of ADNI+AIBL and the CADDementia data. Subsequently, only NC and AD observations from ADNI+AIBL were used for training of the naive Bayes classifier. The trained classifier was finally applied to score the CADDementia data. The FreeSurfer hippocampus segmentation was used to define the ROI in each MRI scan. The whole procedure was computed for the left and the right hippocampus separately, resulting in two hippocampal shape scores.

3.5 Hippocampal Texture

A hippocampal texture score was computed using a texture descriptor recently proposed for quantification of chronic obstructive pulmonary disease in computed tomography [16] in combination with a support vector machine (SVM) with a radial Gaussian kernel. This specific MRI biomarker has previously shown good results [15, 17]. The texture descriptor comprised marginal filter response histograms of a 3-dimensional, rotation-invariant, multi-scale, Gaussian derivative-based filter bank with the following scales: 0.6, 0.86, 1.2, and 1.7 mm. Compared to [16], the Gaussian filter was excluded in order to be invariant to the lack

of a standardized scale in MRI. The morphologically post-processed bilateral FreeSurfer hippocampus segmentation was used to define the ROI. The texture biomarker was trained on all NC and AD observations from ADNI+AIBL. The training involved estimation of adaptive histogram binning for the different filters, and training of the SVM to separate NC from AD. During SVM training, the width of the radial Gaussian kernel and the regularization parameter was estimated using grid search in a nested cross-validation loop. The SVM was subsequently trained on all training data using the optimal parameter combination.

4 Combination Biomarker

The individual MRI biomarkers $\{x^{(i)}\}_{i=1\dots N}$ were combined by entering them as features to a regularized linear discriminant analysis (LDA) with λ added to the diagonal of the covariance matrices [10]. Prior to entering the LDA, each individual biomarker x was z-score transformed dependent on the age of the subject according to $z = (x - \mu_{\text{age}})/\sigma_{\text{age}}$. The age-dependent weighted mean, μ_{age} , and the age-dependent weighted standard deviation, σ_{age} , of the biomarker used in the transformation were estimated from the training set using an adaptive width Gaussian interpolation kernel centered on the respective age. The age-dependent z-score transformation was applied within each group, resulting in a tripling of the features $\{z_{\text{NC}}^{(i)}, z_{\text{MCI}}^{(i)}, z_{\text{AD}}^{(i)}\}_{i=1\dots N}$. The LDA with $\lambda = 0.001$ was trained directly for the three-class problem of discriminating NC, MCI, and AD using ADNI+AIBL, and the Shark C++ library was used for this purpose [11].

Table 2: Overview of individual MRI biomarkers. Hippocampal shape and hippocampal texture uses the FreeSurfer hippocampal segmentation as ROI. FreeSurfer is not trained.

Biomarker	Segmentation method	Training
Cortical thickness		
Frontal lobe	FreeSurfer	-
Parietal lobe	FreeSurfer	-
Temporal lobe	FreeSurfer	-
Occipital lobe	FreeSurfer	-
Cingulate cortex	FreeSurfer	-
Volumetry		
Amygdala	FreeSurfer	-
Caudate nucleus	FreeSurfer	-
Hippocampus	FreeSurfer/N-L Patch	-/HHP
Pallidum	FreeSurfer	-
Putamen	FreeSurfer	-
Ventricular	FreeSurfer	-
Whole brain	FreeSurfer	-
Hippocampal shape	FreeSurfer	ADNI+AIBL
Hippocampal texture	FreeSurfer	ADNI+AIBL

5 Results

The combination biomarker was evaluated by training on ADNI+AIBL and scoring of the CADDementia training dataset. We also report average 10-fold cross-validation performance on ADNI+AIBL with splits stratified on group and dataset. The python script supplied by the CADDementia team is used for this purpose, and the computed performance measures are described on the CADDementia website⁶. The results are summarized in Table 3, and associated receiver operating characteristic (ROC) curves and areas under the ROC curves (AUCs) are shown in Figure 1 and confusion matrices in Table 4.

Table 3: Performance measures.

	classification accuracy	true positive fraction			AUC			
		NC	MCI	AD	Total	NC	MCI	AD
CADDementia train	73.3	91.7	44.4	77.8	83.2	86.6	68.3	95.8
ADNI+AIBL	62.2	79.8	53.2	45.7	78.4	85.5	68.1	82.7

The proposed method is fully automated. Approximate computation time in order to classify a new MRI scan is presented in Table 5 where we also provide the computation time of individual components of the method.

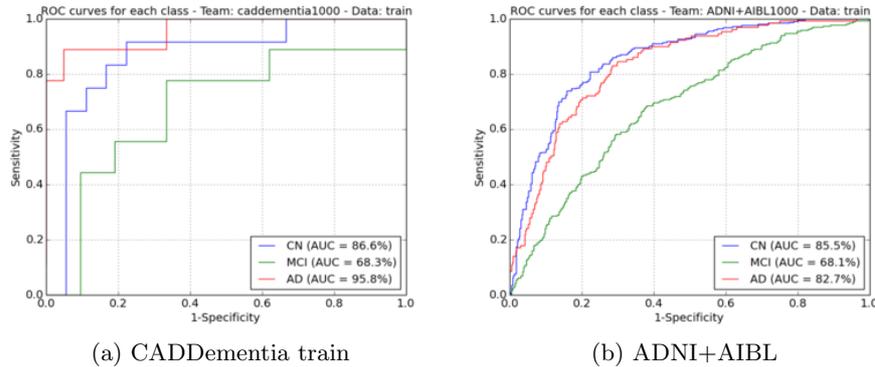


Fig. 1: Per-class ROC curves and AUCs for the proposed combination biomarker.

Table 4: Confusion matrices. Rows are predicted and columns are true class.

	(a) CADDementia train			(b) ADNI+AIBL		
	NC	MCI	AD	NC	MCI	AD
NC	11	3	0	205	74	13
MCI	1	5	2	48	140	57
AD	0	1	7	4	49	59

⁶ <http://caddementia.grand-challenge.org/evaluation/>

Table 5: Per subject computation time divided into components.

Total (+FreeSurfer)	FreeSurfer	N-L Patch	Shape	Texture	Combiner
~1 (19) h	~18 h ¹	~40 min	~0.5 min	~15 min	~1 sec

¹ Median processing time due to high variability with some extreme outliers.

As seen from the confusion matrices, there was a tendency of classifying subjects as too healthy (MCI as NC, AD as MCI). We therefore also produced a second score using the following priors optimized for the CADDementia training set: $P(\text{NC}) = 1/8$, $P(\text{MCI}) = 3/8$, $P(\text{AD}) = 1/2$. These priors resulted in a classification accuracy of 80 % on the CADDementia training set and in a more balanced confusion matrix. The results of using these priors were submitted to the challenge as *LDA-optimized-priors* whereas the previous results were submitted as *LDA-equal-priors*.

6 Discussion and Conclusion

In this paper, we proposed to combine a range of structural MRI biomarkers for the purpose of multi-class classification of AD, MCI and NC. The individual biomarkers used in the combination were cortical thickness measurements, hippocampal shape and texture, and volumetric measurements. Combining such diverse biomarkers may potentially improve diagnostic performance from structural MRI. This is appealing because the modality is less invasive than state-of-the-art biomarkers based on lumbar puncture and positron emission tomography imaging that are directly measuring pathological hallmarks such as amyloid load. However, despite this potential improvement and despite the inclusion of biomarkers such as hippocampal texture that is sensitive to earlier stages of the disease process, there were problems discriminating MCI from AD and NC (see Table 4). Advancing performance further would probably need combination with other non-structural MRI biomarkers such as the aforementioned.

Irrespective of this, a dementia diagnosis is in practice based on several sources of information, such as neuropsychological assessment, physical examination, blood sampling, and visual inspection of some form of anatomical medical imaging (e.g., computed tomography or structural MRI), and structural MRI biomarkers should be used in conjunction with all this information. The proposed structural MRI combination biomarker is a promising direction for obtaining improved diagnostic information from MRI to be used in clinical assessment.

References

1. Achterberg, H.C., *et al.*: Hippocampal shape is predictive for the development of dementia in a normal, elderly population. *Hum Brain Mapp* 35(5), 2359–2371 (2014)

2. Albert, M.S., *et al.*: The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7(3), 270–279 (2011)
3. Anker, C.: Segmentation of Subcortical structures in T1 weighted MRI as a component of a Brain Atrophy Computation Pipeline. Master's thesis, Technical University of Denmark, Department of Applied Mathematics and Computer Science (2014)
4. Anker, C., *et al.*: Automated hippocampal segmentation using new standardized manual segmentations from the Harmonized Hippocampal Protocol. *Alzheimers Dement* , suppl. (2014), accepted.
5. Ellis, K.A., *et al.*: The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr* 21(4), 672–687 (2009)
6. Fischl, B., Dale, A.M.: Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci USA* 97(20), 11050–11055 (2000)
7. Fischl, B., *et al.*: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33(3), 341–355 (2002)
8. Frisoni, G.B., Jack, C.R.: Harmonization of magnetic resonance-based manual hippocampal segmentation: a mandatory step for wide clinical use. *Alzheimers Dement* 7(2), 171–174 (2011)
9. Gower, J.: Generalized procrustes analysis. *Psychometrika* 40(1), 33–51 (1975)
10. Guo, Y., Hastie, T., Tibshirani, R.: Regularized linear discriminant analysis and its application in microarrays. *Biostatistics* 8(1), 86–100 (2007)
11. Igel, C., Heidrich-Meisner, V., Glasmachers, T.: Shark. *Journal of Machine Learning Research* 9, 993–996 (2008)
12. Jack, Jr, C.R., *et al.*: Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. *Alzheimers Dement* 7(4), 474–485.e4 (2011)
13. McKhann, G.M., *et al.*: The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7(3), 263–269 (2011)
14. Singh, V., *et al.*: Spatial patterns of cortical thinning in mild cognitive impairment and Alzheimer's disease. *Brain* 129(Pt 11), 2885–2893 (2006)
15. Sørensen, L., *et al.*: Hippocampal texture provides volume-independent information for Alzheimer's disease diagnosis. *Alzheimers Dement* 8(4), suppl. (2012)
16. Sørensen, L., *et al.*: Texture-based analysis of COPD: A data-driven approach. *IEEE Trans Med Imaging* 31(1), 70–78 (2012)
17. Sørensen, L., *et al.*: Hippocampal texture predicts conversion from MCI to AD. *Alzheimers Dement* 9(4), suppl. (2013)
18. Westman, E., *et al.*: Regional magnetic resonance imaging measures for multivariate analysis in alzheimer's disease and mild cognitive impairment. *Brain Topogr* 26(1), 9–23 (2013)
19. Wyman, B.T., *et al.*: Standardization of analysis sets for reporting results from ADNI MRI data. *Alzheimers Dement* 9(3), 332–337 (2013)