Detecting Alzheimer's disease by morphological MRI using hippocampal grading and cortical thickness

Simon F. Eskildsen[a], Pierrick Coupé[b], Vladimir Fonov[c], D. Louis Collins[c]

[a] Center of Functionally Integrative Neuroscience and MINDLab, Aarhus University, Denmark. [b] Laboratoire Bordelais de Recherche en Informatique, Unité Mixte de Recherche CNRS (UMR 5800), PICTURA Group, Bordeaux, France. [c] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada.

Background

- Structural MRI is an important imaging biomarker in Alzheimer's disease as the cerebral atrophy has been shown to closely correlate with cognitive symptoms. Recognizing this, numerous methods have been developed for quantifying the disease related atrophy from MRI over the past decades. Effort has been dedicated to separate AD related modifications from normal aging for the purpose of early detection and prediction. Several groups have reported promising results using automatic methods; however, it is very difficult to compare these methods due to varying cohorts and different validation frameworks. To address this issue, the public challenge on Computer-Aided Diagnosis of Dementia (CADDementia) was proposed. The challenge calls for accurate classification of 354 MRI scans collected among AD patients, subjects with mild cognitive impairment and cognitively normal control. The true diagnosis is hidden from the participating groups, thus making the validation truly objective. This poster describes our proposed method to automatically classify the challenge data along with a validation on 30 scans with known diagnosis also provided for the challenge.

Results

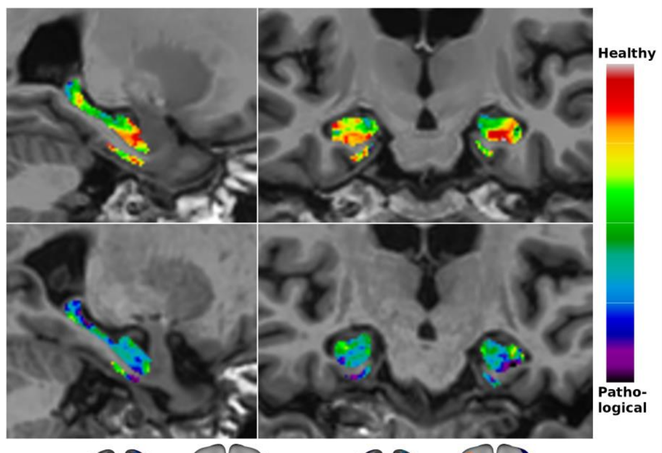

Figure 1. SNIPE grading maps generated from T1-weighted MRI. Sagittal and coronal slices visualizing hippocampus and entorhinal cortex. Top row: ADNI1 MCI non-progressor at baseline. Bottom row: ADNI1 MCI progressor at baseline.

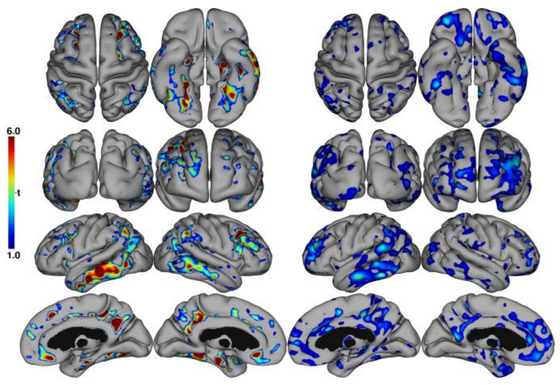

Figure 2. Thresholded t-maps for ADNI1 MCI/CN contrast (left) and ADNI2 MCI/CN contrast (right). Notice the difference in statistical strength due to sample size differences.

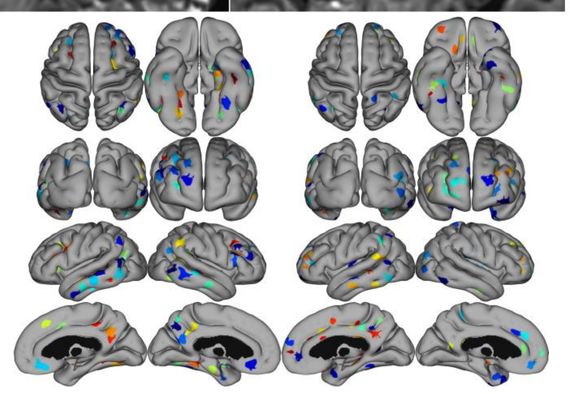

Figure 3. Cortical thickness ROIs generated by respectively ADNI1 cohorts (left) and ADNI2 cohorts (right).

Computational time

- The classification process is fully automatic.
- Test setup: single core (Intel Core i7 @3.40Ghz) per subject.
- Total computational time was approximately 55 minutes distributed on preprocessing (30 min), FACE (15 min), and SNIPE (10 min).
- Applying the classifier after it has been trained takes only a few seconds.

CADDementia training set	Classification accuracy			
Classifier	CN	MCI	AD	Overall
SNIPE/FACE ADNI1	75.0	66.7	66.7	70.0
SNIPE ADNI1	75.0	77.8	77.8	76.7
SNIPE/FACE ADNI2	58.3	66.7	77.8	66.7
SNIPE ADNI2	66.7	66.7	77.8	70.0
Combined	75.0	66.7	77.8	73.3

Table 2. Classification accuracies of the five classifiers when applied on CADDementia training data. All numbers are in percentage (%).

Methods

Training data (ADNI)

- 1.5T T1w MRI scans from the ADNI1 study
- 3.0T T1w MRI scans from the ADNI2 study

Testing data (CADDementia)

- 3.0T T1w MRI scans from three different sites in Europe
- No protocol harmonization
- 30 scans with known labels
- 354 scans with unknown labels

Dataset	N (females)	Age±sd
ADNI1 AD	181 (88)	75.3±7.5
ADNI1 MCI	381 (139)	74.8±7.4
ADNI1 CN	222 (105)	75.9±5.0
ADNI2 AD	48 (16)	75.6±8.8
ADNI2 MCI	183 (69)	71.7±7.6
ADNI2 CN	73 (36)	75.6±6.2
CAD AD	9 (6)	66.1±5.2
CAD MCI	9 (4)	68.0±8.5
CAD CN	12 (3)	62.3±6.3
CAD test set	354 (141)	65.1±7.8

Table 1. Demographics of the cohorts used in the analyses

Image preprocessing

- Fully automatic pipeline [1].
- Denoising [2]
- Bias field correction [3]
- Affine registration to MNI space [4].
- Population-specific template derived from the ADNI1 [5].
- Image intensity normalization [6]
- Skull stripping using BEaST [7].

Hippocampus and Entorhinal cortex

- Structural features of the hippocampal complex were estimated using SNIPE (Scoring by Nonlocal Image Patch Estimator) method [8, 9]. In this technique, the local structural information surrounding each voxel (i.e., 3D patch) of a test subject is compared to those in a training library of MRI datasets from ADNI1 and ADNI2 AD and CN subjects with segmentation of the considered structures (HC and ERC) (see Figure 1).

Neocortex

- Cortical thickness was calculated using FACE (Fast Accurate Cortex Extraction) [10,11] and mapped to the cortical surface of the population-specific average non-linear anatomical template [5] using an iterative, feature-based algorithm [12]. MCI and CN subjects were used to generate a statistical map of group differences in cortical thickness. From this t-map, cortical thickness features were derived with the procedure described in [13] using the proportion of the cortical surface with the 15% largest t-values corresponding to a threshold of t=4.3 and t=1.0 for ADNI1 and ADNI2 respectively (see Figure 2 and 3).

Classification

- Multinomial regression with lasso and L1/L2 elasticnet regularization.
- GLMNET matlab (http://www.stanford.edu/~hastie/glmnet_matlab/).
- During experiments ADNI1 or ADNI2 were used as training datasets.
- The classification framework was based on ensemble learning approach [14].
- In order to create the ensemble we used an iterative approach. For each iteration the following steps were performed:

1. Age correction based on the CN ADNI training population.
2. Over-sampling of the ADNI training dataset.
3. Grid search for optimal classifier parameters.
4. Ensemble learning classification – mean posterior probability

- This procedure was used in the four following scenarios:

1. ADNI1 as training dataset and SNIPE and FACE features (98 features in total)
2. ADNI1 as training dataset and SNIPE features (8 features)
3. ADNI2 as training dataset and SNIPE and FACE features (95 features in total)
4. Using ADNI2 as training dataset and SNIPE features (8 features)
5. The four scenarios above were combined by using the grand mean of all posterior probabilities from the scenarios. As in step four above, the maximum posterior probability was used to label the images.

Conclusion

The results on the CADDementia training data may be inflated due to the grid search for optimal parameters. Nevertheless the results indicate the ranking of the different classifiers. It seems that training on ADNI1 data provides better results than training on ADNI2 data even though ADNI2 data should better represent the CADDementia data using only 3T images. The difference is most likely due to more available training data in the ADNI1 cohorts leading to well-defined classes. Thus the morphological variation of the three populations (AD/MCI/CN) is better represented in the larger sample. Adding cortical thickness features to the SNIPE features does not seem to improve results. In fact, it seems to impair the classifiers, possibly due to adding noise. Perhaps fewer and more carefully selected ROIs in the neocortex would have better complemented the MTL features and thus added discriminative information.

References

[1] Aubert-Broche et al, NeuroImage 82C:393-402, 2013. «» **[2] Coupé** et al, IEEE TMI 27:425-441, 2008. «» **[3] Sled** et al, IEEE TMI 17:87-97, 1998. «» **[4] Collins** et al, J Comput Assist Tomogr 18:192-205, 1994. «» **[5] Fonov** et al, NeuroImage 54:313-327, 2011. «» **[6] Nyul & Udupa**, SPIE 2000«» **[7] Eskildsen** et al, NeuroImage 59:2362-2373 2012. «» **[8] Coupé** et al, NeuroImage 59(4):3736-47, 2012a. «» **[9] Coupé** et al, NeuroImage:Clin 1(1):141-152, 2012b. «» **[10] Eskildsen** et al, MICCAI 2006 «» **[11] Eskildsen** et al, MICCAI 2007. «» **[12] Eskildsen** et al, SIBGRAPI 2008. «» **[13] Eskildsen** et al, NeuroImage 65C, 511-521, 2013. «» **[14] Dietterich**, MCS 2000.

The Danish National Research Foundation's
Center of Functionally Integrative Neuroscience
University of Aarhus / Aarhus University Hospital
www.cfin.au.dk

Laboratoire Bordelais de Recherche en Informatique
Université de Bordeaux

Montreal Neurological Institute
McGill University

Contact: seskildsen@cfin.au.dk