Comparison of DTI Features for the Classification of Alzheimer's Disease: A Reproducible Study

Junhao Wen[1,2], Jorge Samper-González[1,2], Simona Bottani[1,2], Alexandre Routier[1,3], Ninon Burgos[1,2], Thomas Jacquemont[1,2], Sabrina Fontanella[1,2], Stanley Durrleman[1,2], Anne Bertrand[1,4,5], Olivier Colliot[1,4,6]

[1]Inria Paris, Aramis project-team, Paris, France
[2]Sorbonne Université, Inserm, CNRS, Institut du Cerveau et la Moëlle (ICM), Paris, France
[3]Sorbonne Université, Inserm, CNRS, Institut du Cerveau et la Moëlle (ICM), FrontLab, Paris, France
[4]Sorbonne Université, Inserm, CNRS, Institut du Cerveau et la Moëlle (ICM), AP-HP, Paris, France
[5]AP-HP, Saint-Antoine Hospital, Department of Radiology, Paris, France
[6]AP-HP, Departments of Neuroradiology and Neurology, Pitié-Salpêtrière Hospital, Paris, France

Introduction

Several studies using machine learning have recently looked at the potential of diffusion tensor imaging (DTI) for Alzheimer's disease (AD) classification (O'Dwyer et al., 2012; Dyrba et al., 2013; Maggipinto et al., 2017). However, classification accuracies are not directly comparable across studies because of differences in subject selection, image processing, feature extraction and selection, and classification algorithms. Samper-Gonzalez et al. (2017) proposed a reproducible framework for automatic classification of AD from T1 MRI and PET data. Here, we extend this framework for DTI-based classification using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study. This work aims to facilitate replication of classification experiments based on ADNI and also to compare the classification performances with different DTI-based features.

Methods

The framework is composed of the following components. Tools were implemented to automatically convert original ADNI diffusion MRI into the Brain Imaging Data Structure (BIDS) format (Gorgolewski et al., 2016), a community standard for data organization, thus simplifying future data management. These tools also provide possibility for subject selection based on different times of follow up and diagnoses. An image preprocessing and feature extraction pipeline was developed using Nipype (Gorgolewski et al., 2011) and combined tools from FSL, MRtrix and ANTS. Artifact correction was performed as in (Jacquemont et al., 2017), including corrections for susceptibility-induced distortions, eddy current-induced distortions and head motion. The DTI model was then fitted generating FA and MD maps. FA and MD maps were nonlinearly registered onto the John Hopkins University (JHU) atlas template. We then extracted two types of features: regional features and voxel-based features. Regional features were the average FA and MD value in each region of the JHU atlas (two versions of the atlas were used: JHUTract and JHULabel). Voxel-based features were FA and MD maps in MNI space and masked using the WM, GM or GM+WM binarized maps. Classification was performed using a linear support vector machine (SVM) from scikit-learn. A repeated holdout cross validation (250 runs of stratified random splits with 20% of the data used for testing) with a 10-fold inner grid search for hyperparameter optimization was performed. Additionally, optimal margin hyperplane coefficient maps were reported to characterize potential anatomical patterns in AD (as in Cuingnet et al, 2013).
We demonstrate the use of the framework on the classification of 46 AD patients and 46 cognitively normal (CN) subjects.

Results

All classification results for AD vs CN are shown in table 1. Overall, voxel-wise features provided higher accuracies than regional features. MD region-based classification did not perform better than chance. The highest classification accuracy for FA was obtained with the WM+GM voxel features (77%), and for MD with the GM voxel features (76%).
From the hyperplane coefficient maps displayed in Fig 2, we observe that the hippocampus and medial temporal cortex are the main regions discriminating AD from CN subjects using

MD, and that the most influential WM regions for MD are absent from the JHULabel atlas, explaining the low accuracy obtained with region-based MD features.

Conclusions

We presented a framework for data organization, image processing, feature extraction and machine learning-based analysis that enables the comparability and reproducibility of DTI-based AD classification. The classification results obtained with this framework were in line with the state of the art and highlighted that the atlases used for region-based approaches should be chosen with care. The code will be made publicly available at the time of the conference at https://github.com/aramis-lab/AD-ML.

References

Dyrba, M., Ewers, M., Wegrzyn, M., Kilimann, I., Plant, C., Oswald, A., Meindl, T., Pievani, M., Bokde, A.L., Fellgiebel, A. and Filippi, M., 2013. Robust automated detection of microstructural white matter degeneration in Alzheimer's disease using machine learning classification of multicenter DTI data. PloS one, 8(5), p.e64925.

Maggipinto, T., Bellotti, R., Amoroso, N., Diacono, D., Donvito, G., Lella, E., Monaco, A., Scelsi, M.A., Tangaro, S. and Alzheimer's Disease Neuroimaging Initiative, 2017. DTI measurements for Alzheimer's classification. Physics in Medicine and Biology, 62(6), p.2361.

O'Dwyer, L., Lamberton, F., Bokde, A.L., Ewers, M., Faluyi, Y.O., Tanner, C., Mazoyer, B., O'Neill, D., Bartley, M., Collins, D.R. and Coughlan, T., 2012. Using support vector machines with multiple indices of diffusion for automated classification of mild cognitive impairment. PloS one, 7(2), p.e32441.

Samper-Gonzalez, J., Burgos, N., Fontanella, S., Bertin, H., Habert, M.O., Durrleman, S., Evgeniou, T., Colliot, O. and Alzheimer's Disease Neuroimaging Initiative, 2017, September. Yet Another ADNI Machine Learning Paper? Paving The Way Towards Fully-reproducible Research on Classification of Alzheimer's Disease. In International Workshop on Machine Learning in Medical Imaging (pp. 53-60). Springer, Cham.

K. Gorgolewski et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data, 3:160044, 2016.

Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L. and Ghosh, S.S., 2011. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Frontiers in neuroinformatics, 5.

Jacquemont, T., Fallani, F.D.V., Bertrand, A., Epelbaum, S., Routier, A., Dubois, B., Hampel, H., Durrleman, S., Colliot, O. and Alzheimer's Disease Neuroimaging Initiative, 2017. Amyloidosis and neurodegeneration result in distinct structural connectivity patterns in mild cognitive impairment. Neurobiology of Aging, 55, pp.177-189.

Cuingnet, R., Glaunès, J.A., Chupin, M., Benali, H. and Colliot, O., 2013. Spatial and anatomical regularization of SVM: a general framework for neuroimaging data. IEEE transactions on pattern analysis and machine intelligence, 35(3), pp.682-696.

Approach	Feature	AUC	Balanced Accuracy	Sensitivity	Specificity
ROI-wise	JHULabel-FA	0.89	0.71	0.71	0.70
	JHULabel-MD	0.47	0.5	0.52	0.48
	JHUTract-FA	0.60	0.69	0.74	0.63
	JHUTract-MD	51.4	0.5	0.49	0.51
Voxel-wise	WM-FA	0.82	0.76	0.77	0.74
	GM-FA	0.77	0.71	0.70	0.74
	WM+GM-FA	0.82	0.77	0.76	0.78
	WM-MD	0.82	0.74	0.63	0.86
	GM-MD	0.83	0.76	0.68	0.83
	WM+GM-MD	0.82	0.75	0.67	0.82

Table 1 Classification results obtained with DTI-based features using a linear SVM classifier when differentiating AD (n=46) from CN (n=46) subjects. The balanced accuracy, sensitivity and specificity were averaged over 250 runs. JHULabel corresponds to the ICBM-DTI-81 white-matter label atlas and and JHUTract to the JHU white-matter tractography atlas. AUC: area under the receiver operating characteristic (ROC) curve.

Table 1 Classification results obtained with DTI-based features using a linear SVM classifier when differentiating AD (n=46) from CN (n=46) subjects. The balanced accuracy, sensitivity and specificity were averaged over 250 runs. JHULabel corresponds to the ICBM-DTI-81

white-matter label atlas and and JHUTract to the JHU white-matter tractography atlas. AUC: area under the receiver operating characteristic (ROC) curve.

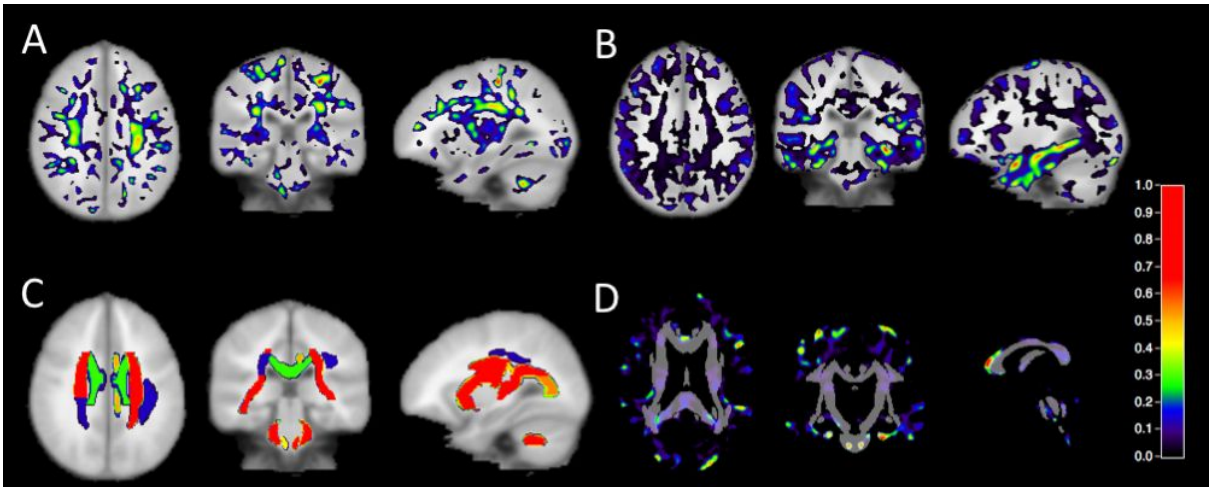

Fig 1. Normalized coefficient map based on different features. In each voxel or ROI, higher coefficient means higher likelihood of classification into AD. (A) WM-FA map (B) WM+GM-MD map (C) JHULabel-FA map (D) WM-MD map superimposed on JHULabel atlas