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-Salpetrière Hospital, Paris, France

junhao.wen@inria.fr

Several studies using **machine learning** have recently looked at the potential of **diffusion tensor imaging** (**DTI**) for Alzheimer's disease (**AD**) classification.[1,2,3] 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[4] 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

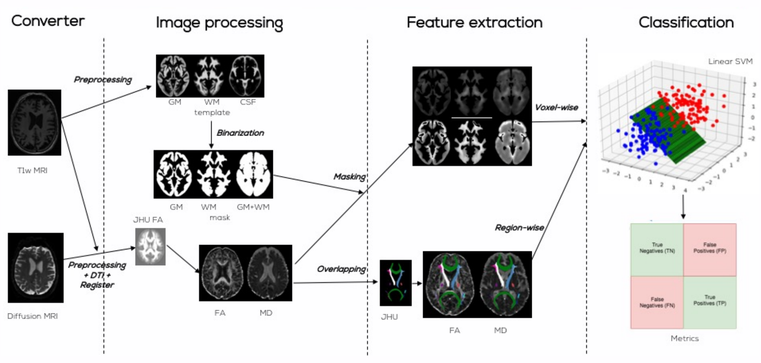

Figure 1: Classification framework for AD subjects recognition.

The classification framework is composed of several components (**Fig. 1**). Tools in Clinica software[5] allowed to automatically convert original ADNI data into the Brain Imaging Data Structure (BIDS) format. Tools were also implemented for subject selection based on different times of follow up and diagnoses. Image preprocessing and feature extraction pipelines were developed in Clinica, respectively for T1-weighted MRI and diffusion MRI.

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, To assess the impact of imbalanced data, a random under-sampling technique was used for each imbalanced task.

We demonstrate the use of the framework with 46 AD patients, 46 cognitively normal (**CN**) subjects, 97 mild cognitive impairment (**MCI**) subjects, including 54 stable MCI (**sMCI**) and 24 progressive MCI (**pMCI**) at 36 months. Additionally, the corrected resampled t-test was performed to statistically compare the performances across imaging modalities with a significant level $P < 0.05$.

We performed clinical diagnosis classification tasks, or "predictive" tasks of the evolution of MCI subjects. Four different classification tasks were considered in our analyses: CN vs AD, CN vs pMCI, sMCI vs pMCI and CN vs MCI (**Tab. 1**).

Tasks	Imaging Modality	Feature Type	Data
CN vs AD or CN vs pMCI or sMCI vs pMCI or CN vs MCI	Diffusion MRI	Voxel-based	Balanced
			Imbalanced
		Region-based	Balanced
			Imbalanced
	T1w MRI	Voxel-based	Balanced
			Imbalanced
		Region-based	Balanced
			Imbalanced

Table 1: Summary of all the classification experiments.

• Results

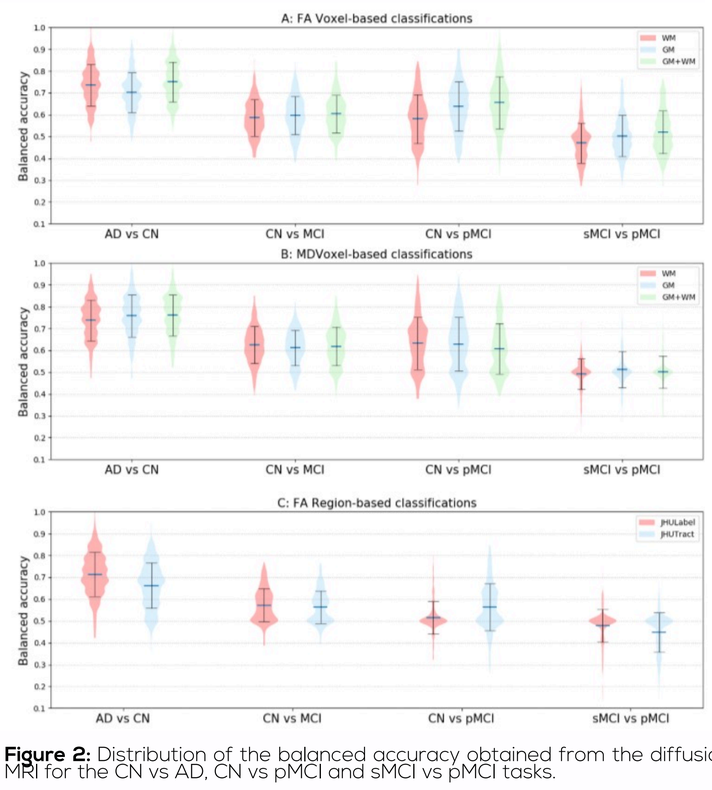

Figure 2: Distribution of the balanced accuracy obtained from the diffusion MRI for the CN vs AD, CN vs pMCI and sMCI vs pMCI tasks.

For diffusion MRI, our framework showed higher accuracies for task CN vs AD but lower accuracies for tasks involving MCI. FA & MD gave similar performances. Region-based MD does not perform better than chance (**Fig. 2**).

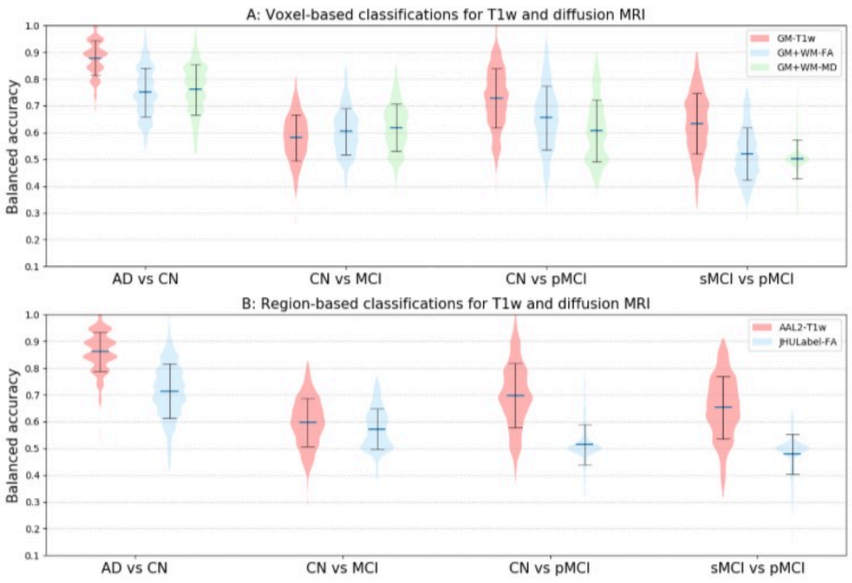

Figure 3: Distribution of the balanced accuracy obtained from both T1w and diffusion MRI.

Compared to diffusion MRI, T1w MRI lead to significantly higher balanced accuracies for tasks CN vs AD ($P = 0.011$) and sMCI vs pMCI ($P = 0.030$) between GM-T1w feature and GM+WM-MD feature (**Fig. 3**).

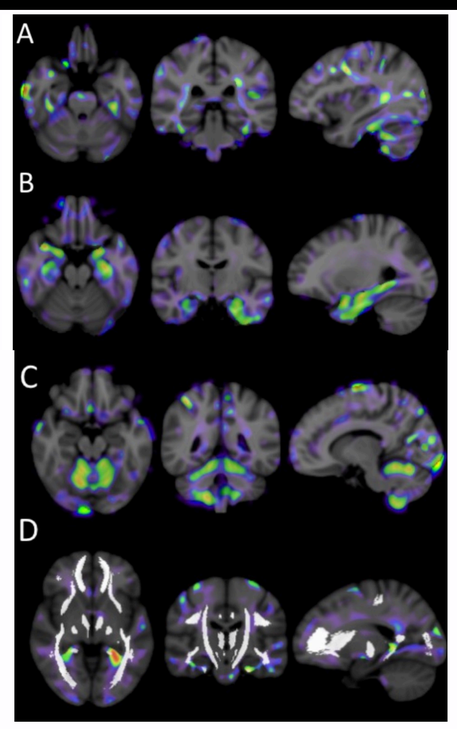

Figure 4: Normalized coefficient maps in MNI space with a background template. Task CN vs AD with A) GM+WM-FA feature; B) GM+WM-MD feature; C) T1-GM feature; D) WM-MD feature, which is on top of JHULabel atlas (in white). For each voxel in warm colors, it means higher likelihood of classification into AD.

The most discriminative voxels for MD are absent from JHU atlas (**Fig. 4 D**).

• Conclusion

We proposed an open-source framework for the reproducible evaluation of AD classifications for diffusion MR:

- Our baseline results from DTI-based features are in line with the state-of-the-art results.
- FA & MD voxel-based features give comparable results and the atlases used for regional approach should be taken with care.
- T1w MRI significantly outperformed diffusion MRI for task CN vs AD and sMCI vs pMCI.
- The poor performances obtained for tasks involving MCI could potentially be caused by the small data sample, rather than the imbalanced data.

We hope our work could be useful for researchers' future work in the AD field.

• References

1. Maggipinto, T. et al., 2017. Physics in medicine and biology, 62(6), pp.2361–2375.
2. Dyrba, M. et al., 2013. PloS one, 8(5), p.e64925
3. Demirhan, A. et al., 2015. IEEE International Symposium on Biomedical Imaging, 2015, pp.126–130.
4. Samper-Gonzalez, J., Burgos, N., Fontanella, In International Workshop on Machine Learning in Medical Imaging (pp. 53-60). Springer, Cham.
5. Routier, A., Guillon, J., Burgos, N., In Annual meeting of the Organization for Human Brain Mapping-OHBM 2018.