

Predictive models based on functional connectivity: a benchmark across multiple datasets

Kamalaker DADI, Alexandre ABRAHAM, Mehdi RAHIM, Darya CHYZHYK,
Bertrand THIRION, Gael VAROQUAUX

Parietal Team, INRIA, Paris-Saclay University, France.

CEA, DSV, I2BM, Neurospin Center, Paris-Saclay University, 91191, Gif-sur-Yvette, France.

Introduction

Recently, a fully-automated pipeline has been proposed to extract autism biomarkers [1] from large and multi-site resting state fMRI datasets based on functional connectomes. It quantifies the impact of parameter variation on the analysis of a large multi-centric datasets. It allows pipeline tuning even in the situation of multi-site studies. However, the question of finding an optimal pipeline setting that generalizes to multiple diagnosis problems remains.

We study whether parameters can be reliably selected to answer diverse clinical questions based on population analysis of connectomes from three different datasets. We outline effects of pipeline steps consistent across studies, and compare the impact of each pipeline step on the prediction accuracy on these datasets that target different disease characterizations.

Methods

We use a standard functional connectivity analysis pipeline. It consists of four steps: Delineation of brain regions of interest (ROIs), ROI-level rs-fMRI time series signal extraction, Function Connectivity (FC) estimation and classification using the FC features.

For each step, we explore various methodological choices: ROI set selection, where we choose three predefined atlases - AAL, Harvard Oxford, Bootstrap Analysis of Stable Clusters [2] (BASC) and four data-driven methods - K-Means, hierarchical agglomerative clustering using Ward's algorithm, multi-subject Independent Component Analysis [3] (ICA), Online Dictionary Learning [4] (DictLearn). FC metrics such as correlation, partial correlation and tangent space embedding [5]. Linear classifiers to compare and evaluate the dominant strategies for the sake of prediction accuracy.

All computations and models are implemented based on scikit-learn [6] and nilearn [7].

We experiment our classification pipeline on three rs-fMRI datasets. The first dataset is from the Center for Biomedical Research Excellence (COBRE) [8]. The pipeline predicts the schizophrenia diagnosis of the subjects. The second dataset is the Alzheimer's Disease Neuroimaging Initiative (ADNI) [9]. We discriminate Alzheimer's Disease (AD) from Mild Cognitive Impairment (MCI) group. The third dataset is the Addiction Connectome Preprocessed Initiative (ACPI) [10], where we discriminate Marijuana consumers versus control subjects.

Results

We measure the impact of the different options on the prediction scores relative to the mean prediction using a full- factorial analysis of variance (ANOVA) on prediction scores. In the linear model, the pipeline step alternatives are considered as a categorical variable and their contribution is given by their coefficient. Fig. 1 gives the results on each dataset (on the right) and a summary analysis on the pooled data (on the left). Error bars give the 95% confidence interval.

Across all datasets, l2-penalized classifiers clearly outperform l1 penalization. In addition, using the tangent embedding as a connectivity measure gives a consistent improvement over correlations and partial correlation.

With regards to functional regions, the most effective atlases are those learned on functional data (BASC, ICA, Dictionary Learning).

Fig. 2, gives the accuracy and best strategy for selection of brain ROIs per dataset. Conclusions vary. Schizophrenic subjects (COBRE) are most easily discriminated with functional data-driven ROIs, whereas no such effect is observed in the other datasets (ADNI or ACPI). Note however that prediction accuracy is lower in these datasets.

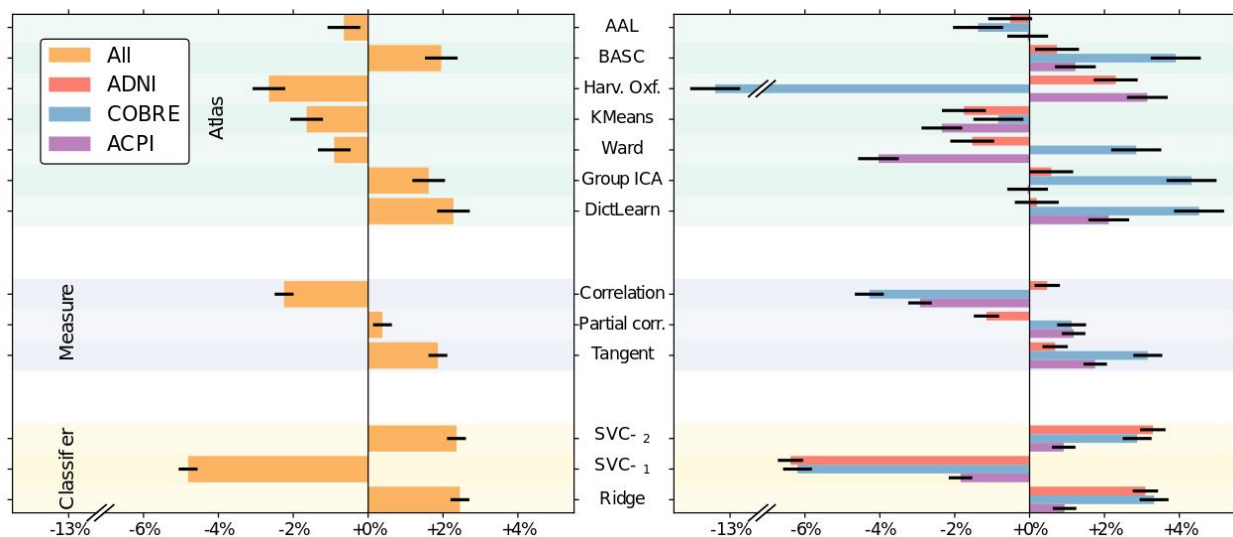


Figure 1: **Post-hoc comparisons between pipeline choices on the rs-fMRI datasets.** Single factor analysis (right) with each dataset (ADNI, COBRE, ACPI) and on pooled data (left). We observe that: (i) Tangent embedding performs better than correlation or partial correlation in all datasets. (ii) l2 regularized classifiers SVC and Ridge are more accurate than SVC- l1 classifier. (iii) With regards to brain atlases, decomposition methods (ICA, DictLearn) are generally the best choices, though with striking cross-datasets differences.

| | COBRE | ADNI | ACPI |
|-----------|---------------------------------|---------------------------------|---------------------------------|
| AAL | .79 \pm .09 | .66 \pm .09 | .58 \pm .08 |
| BASC | .83 \pm .08 | .67 \pm .09 | .59 \pm .08 |
| DictLearn | .83 \pm .08 | .66 \pm .07 | .62 \pm .07 |
| HO | .62 \pm .11 | .68 \pm .09 | .62 \pm .09 |
| ICA | .83 \pm .08 | .68 \pm .09 | .56 \pm .08 |
| K-Means | .79 \pm .08 | .64 \pm .09 | .56 \pm .08 |
| Ward | .83 \pm .08 | .64 \pm .08 | .53 \pm .09 |

Figure 2: **Comparison of the atlas impact on prediction accuracy.** Reported values are the mean and standard deviation of AUC over 100 iterations. Best predictions are in bold.

Conclusions

We present a functional connectivity-based analysis pipeline to predict diverse behavioral targets from rs- fMRI data.

We show guidelines for classifier selection and covariance estimation: Rely on l2 classifiers; Use the tangent space embedding of the covariance matrix. The brain atlas selection in the pipeline does not give a clear trend. Overall, functionally-driven regions with Dictionary Learning, group ICA or BASC atlas perform well across datasets.

References

- [1] A. Abraham, (2016), ‘Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example’, *Neuroimage*, Elsevier, 2016.
- [7] A. Abraham, (2014), ‘Machine learning for neuroimaging with scikit-learn’, *Frontiers in neuroinformatics*, 8: 14.
- [6] F. Pedregosa, (2011), ‘Scikit-learn: Machine learning in Python’, *Journal of Machine Learning Research*, **vol-12**, pages. 2825-2830.
- [3] G. Varoquaux (2010), ‘A group model for stable multi-subject ICA on fMRI datasets’, *NeuroImage*, **vol. 51**, p. 288.

- [5] G. Varoquaux (2010), 'Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling', *MICCAI*.
- [4] J. Mairal (2009), 'Online dictionary learning for sparse coding', *ICML*. pages. 689-696
- [2] P. Bellec (2010), 'Multi-level bootstrap analysis of stable clusters in resting-state fMRI', *NeuroImage*, **vol. 51**, p. 1126.
- [9] S. G. Mueller, (2005), 'The alzheimer's disease neuroimaging initiative', *Neuroimaging Clinics of North America*, **vol. 15**, no. 4, pp. 869-877.
- [8] V. D. Calhoun, (2011), 'Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder', *Frontiers in psychiatry*. 2:75.
- [10] <http://fcon1000.projects.nitrc.org/indi/ACPI/html/>

