SpaceNet: Multivariate brain decoding and segmentation

Elvis Dohmatob, Michael Eickenberg, Bertrand Thirion, Gael Varoquaux

E. DOHMATOB, M. EICKENBERG, B. THIRION, G. VAROQUAUX, PRESENTED @ OHBM 2015, HONOLULU

N° 3980

SPACENET: MULTIVARIATE BRAIN DECODING AND SEGMENTATION

Abstract. We present *SpaceNet*, a multivariate method for brain decoding and segmentation. SpaceNet uses priors like TV (Total Variation) [Michel et al. 2011], TV-L1 [Baldassarre et al. 2012, Gramfort et al. 2013], and GraphNet / Smooth-Lasso [Hebiri et al. 2011, Grosenick et al. 2013] to regularize / penalize classification and regression problems in brain imaging. The result are brain maps which are both sparse (i.e regression coefficients are zero everywhere, except at predictive voxels) and structured (blobby). The superiority of such priors over methods without structured priors like the Lasso, SVM, ANOVA, Ridge, etc. for yielding more interpretable maps and improved classification / prediction scores is now well-established [Baldassarre et al. 2012, Gramfort et al. 2013, Grosenick et al. 2013]. In addition, such priors lead to state-of-the-art methods for extracting brain atlases [Abraham et al. 2013].

MODEL

$$\left.\begin{array}{l}\textbf{Minimize}\quad \overbrace{\mathcal{L}(X,y,w)}^{\text{loss / datafit term}} + \overbrace{\alpha\left(\rho\|w\|_1 + (1-\rho)\Omega(\nabla w)\right)}^{\text{penalty / regularization / prior term}}\\ \textbf{For } w \in \mathbb{R}^p,\\ \textbf{where } \Omega(z) = \begin{cases} \frac{1}{2}\|z\|_2^2 := \frac{1}{2}\sum_{j=1}^{p}\sum_{k=1}^{3} z_{j,k}^2, & \text{for GraphNet,}\\ \|z\|_{1,2} := \sum_{j=1}^{p}\sqrt{\sum_{k=1}^{3} z_{j,k}^2}, & \text{for TV-L1,}\\ \ldots \end{cases}\end{array}\right\} \quad (1)$$

where :

- $X \in \mathbb{R}^{n\times p}$ is the design matrix ; n = # samples (brain images, TRs, etc.) ; p = # voxels (the features). In practice $n \ll p$ (high-dimensions).
- $y \in \mathbb{R}^n$ is the response variable to be predicted.
- w is the brain map of regressor coefficients (one coefficient per voxel).
- $\mathcal{L}(X, y, w)$ is the loss / datafit term, which is $\frac{1}{2}\|Xw - y\|_2^2$ in regression settings, logistic loss in classification settings, etc.
- $\nabla = [\nabla_x, \nabla_y, \nabla_z]^T \in \mathbb{R}^{3p\times p}$ is the 3D discrete spatial gradient operator.
- $\alpha > 0$ is the total amount of regularization.
- ρ $(0 < \rho \leq 1)$ is a mixing constant between the sparsity-inducing ℓ_1 part and the structure-premoting $\Omega(\nabla w)$ part of the penalty term.

METHODS

- The SpaceNet model leads to difficult non-smooth mathematical optimization problems (1), making their implementation challenging.
- [Dohmatob et al. 2014 (PRNI)] benchmarked a rich variety of cutting-edge solvers for such problems, and gave recommendations on how to effectively implement these algorithms in practice. In these benchmarks, the FISTA algorithm [Beck & Teboulle 2009] emerged as the go-to algorithm for the TV-L1 problem, and has been implemented as the SpaceNet default solver (for problem (1)).

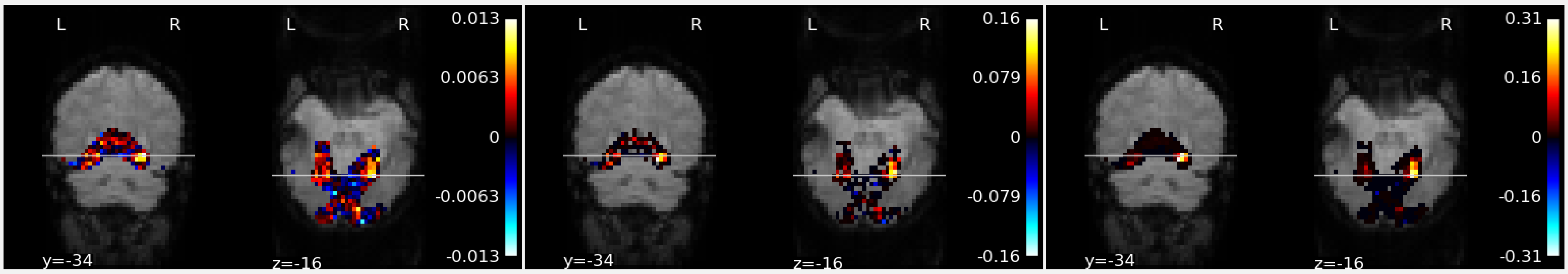
SPEEDUP HEURISTICS

Model-selection in SpaceNet is via automatic cross-validation on a 2D parameter grid of regularization parameters (α, ρ). Moreover, we have developed a number of heuristics that can result in up to 10-fold speedup in runtime without sacrificing the predictions / classifications :

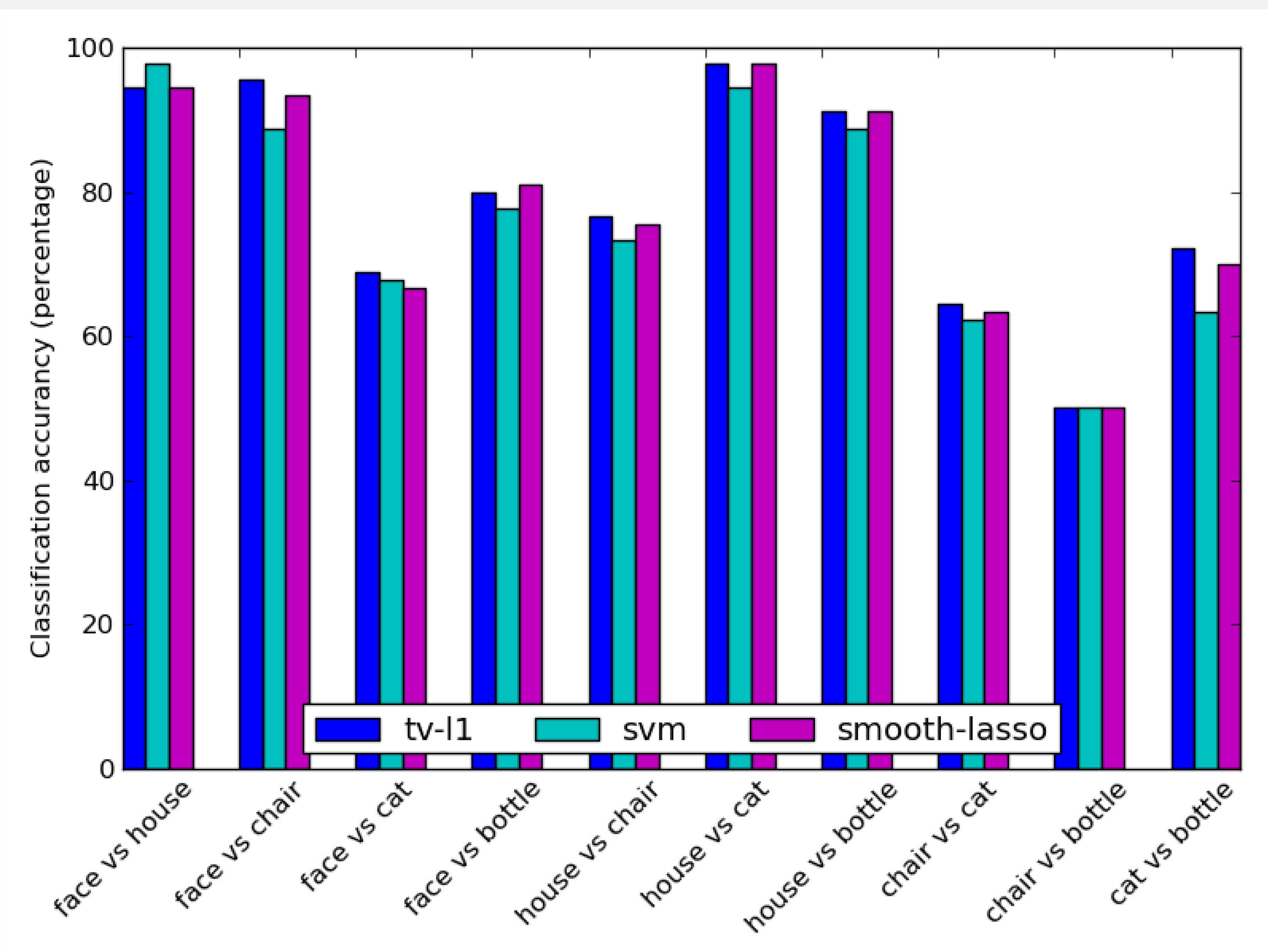
- Univariate feature-screening, whereby irrelavant voxels are detected and eliminated from the model even before solving problem (1).
- Early-stopping, whereby the optimization procedure is abrupted once the scores on cross-validation leftout data stop improving.

These speedup heuristics will appear in proceedings of PRNI 2015 in a slightly more technical discussion [Dohmatob et al. 2015 (PRNI)].

RESULTS & CONCLUDING REMARKS



(a) SVM regressor coefficients (b) Smooth-Lasso regressor coefficients (c) TV-L1 regressor coefficients



(d) Classification accuracy

FIGURE : The figure shows results of comparing SpaceNet TV-L1 and Smooth-Lasso priors against an SVM (Support Vector Machine) classifier on the visual-recognition dataset [Haxby et al. 2001]. As can be seen from the figure, SpaceNet priors yield stable and more interpretable maps, and also better regression / classification performance. The complete results will appear in [DOHMATOB et al. 2015 (OHBM)].

CONCLUSION

- We have presented, SpaceNet, a family of priors for brain decoding which enforce both sparsity and structure, leading to better prediction / classification scores and more interpretable brain maps.
- Indeed, they leverage a feature-selection function (since they limit the number of active voxels), and also a structuring function (since they penalize local differences in the values of the brain map).
- We believe that such priors will become commonplace in future.
- SpaceNet is now part of Nilearn [Abraham et al. 2014] : `http://nilearn.github.io`, an open-source Python library which features cutting-edge algorithms for machine learning in neuroimaging : data-cleaning, clustering, variance analysis, covariance estimation, decoding, and more.

PARIETAL informatics mathematics Inria NeuroSpin