

Learning Neural Representations of Human Cognition across Many fMRI Studies



Arthur Mensch⁽¹⁾

Julien Mairal⁽²⁾

Danilo Bzdok⁽³⁾

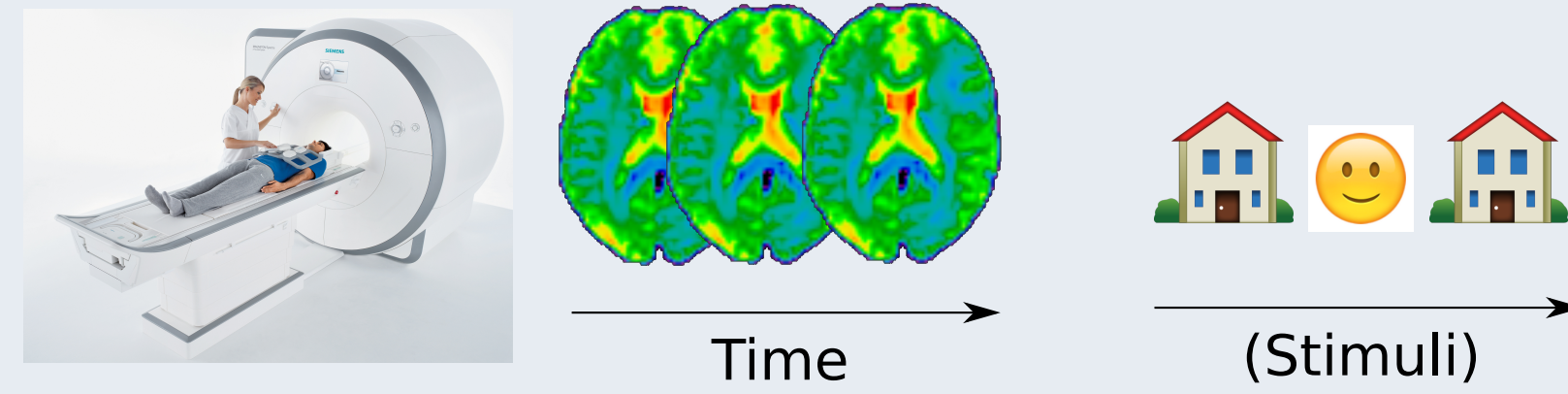
Bertrand Thirion⁽¹⁾

Gaël Varoquaux⁽¹⁾



⁽¹⁾Inria, CEA, Université Paris-Saclay, Gif-sur-Yvette, France ⁽²⁾Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, France ⁽³⁾Department of Psychiatry, RWTH, Aachen, Germany

Functional MRI decoding



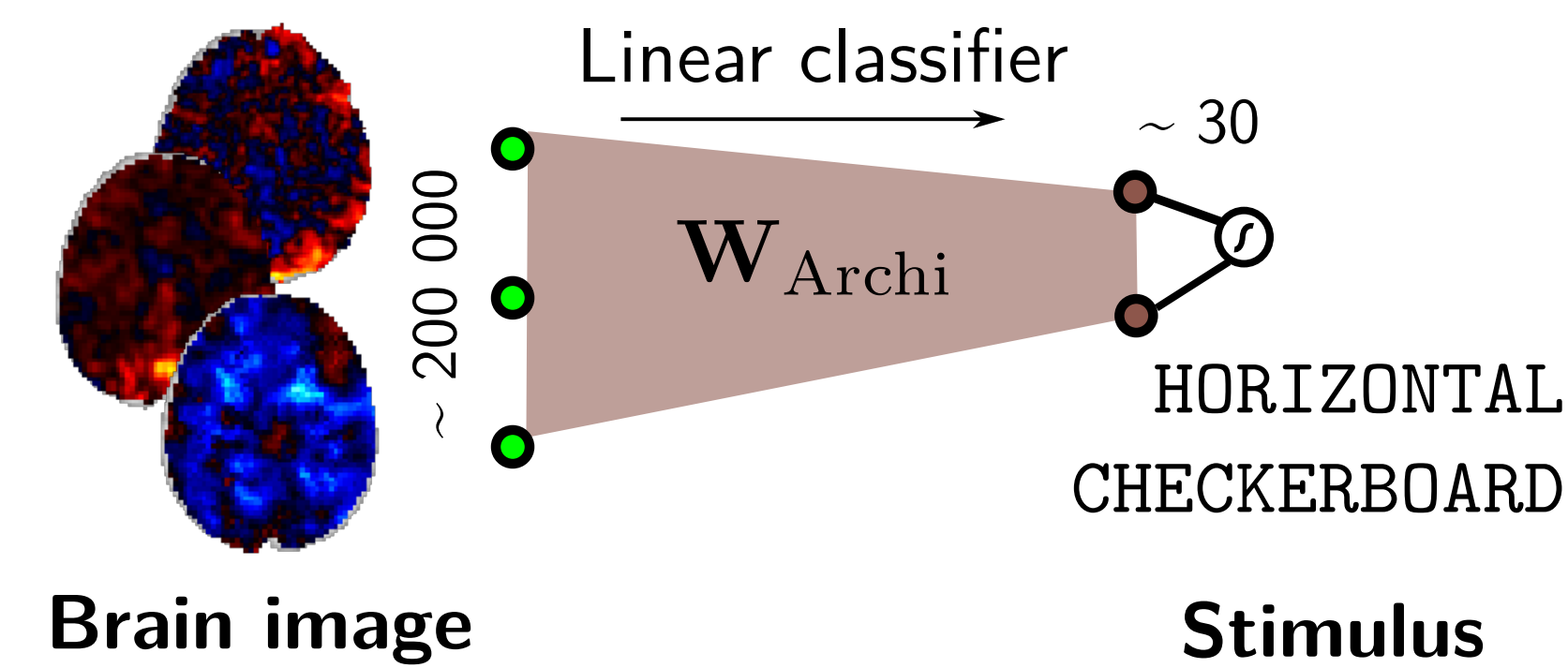
Relate brain activation to cognitive stimuli

- Many small studies with \neq psychological paradigms
- A few large scale studies (1000s subjects)

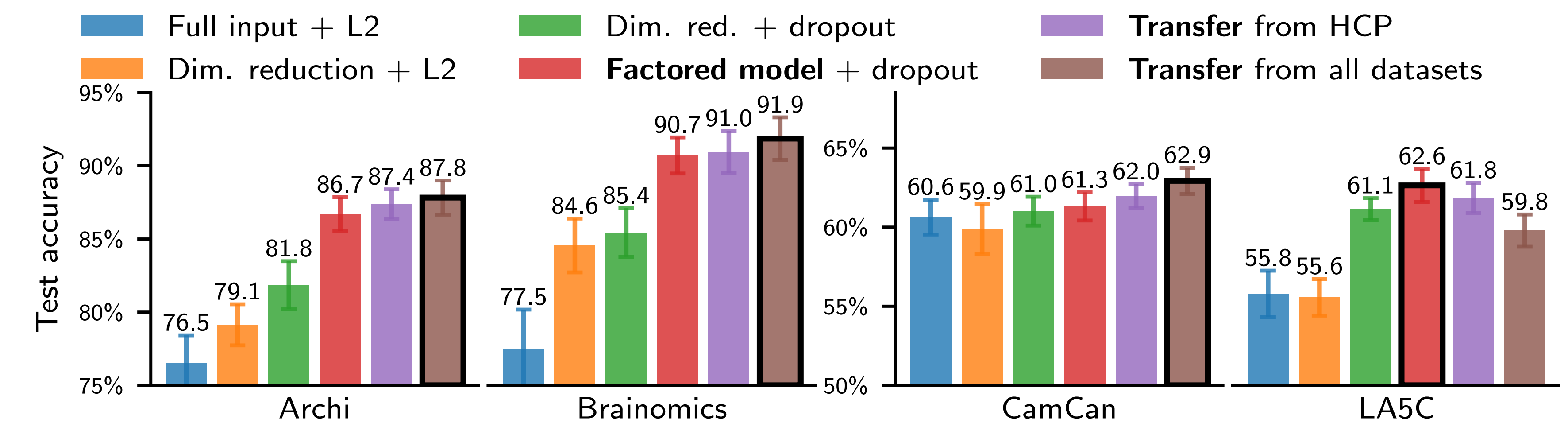
How to aggregate heterogeneous fMRI data into a common cognitive model ?

Learning setting

- First level GLM \rightarrow **z-score maps**
- One map / record / base-condition
- Condition prediction on **new subjects**
- Baseline: **Multinomial regression**

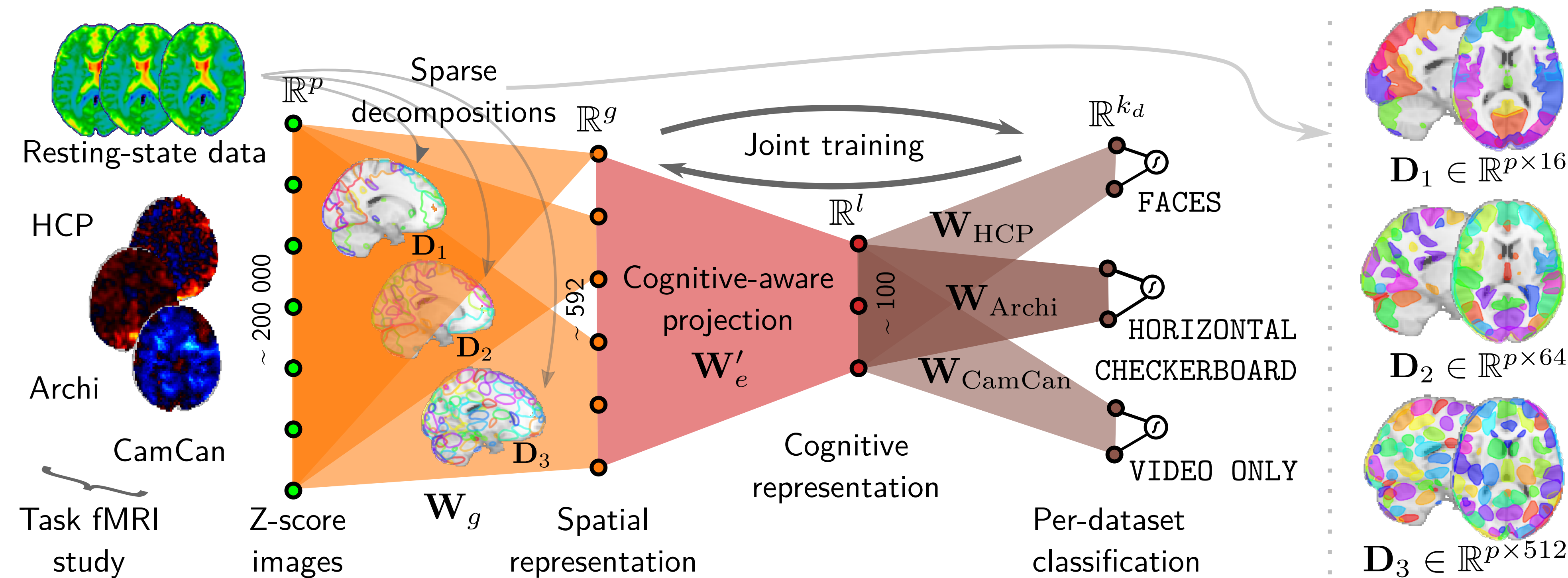


Performance



- Dimension reduction** using resting state data is efficient regularization (and \downarrow train cost)
- Extra efficient **latent layer + Dropout** (explains 2-layer \rightarrow 3-layer improvement)
- Transfer learning** occurs, and is stronger with more datasets (one \rightarrow two \rightarrow five datasets)

Model



Dimension reduction W_g

- Sparse dictionaries** from HCP resting-state [2]
 $\mathbf{X}_{rs} = \mathbf{D}\mathbf{A} \in \mathbb{R}^{p \times g} \times g \times n$, \mathbf{D} sparse
- Orthogonal projection \rightarrow 1 time-series / component
- Unsupervised setting \rightarrow number of components ?
- \rightarrow **Multi-scale** dictionaries

Latent space embedding

- Finding a common representation of brain images
- That is easy to classify for **multiple datasets**
- Factorized linear model (with 2 layers, $l < g$)
 $\forall d \in D, \mathbf{W}_d = \mathbf{W}_e \mathbf{W}'_d \in \mathbb{R}^{g \times l} \times \mathbb{R}^{l \times k_d}$
- \mathbf{W}_e is **shared**: multi-task/**transfer learning**

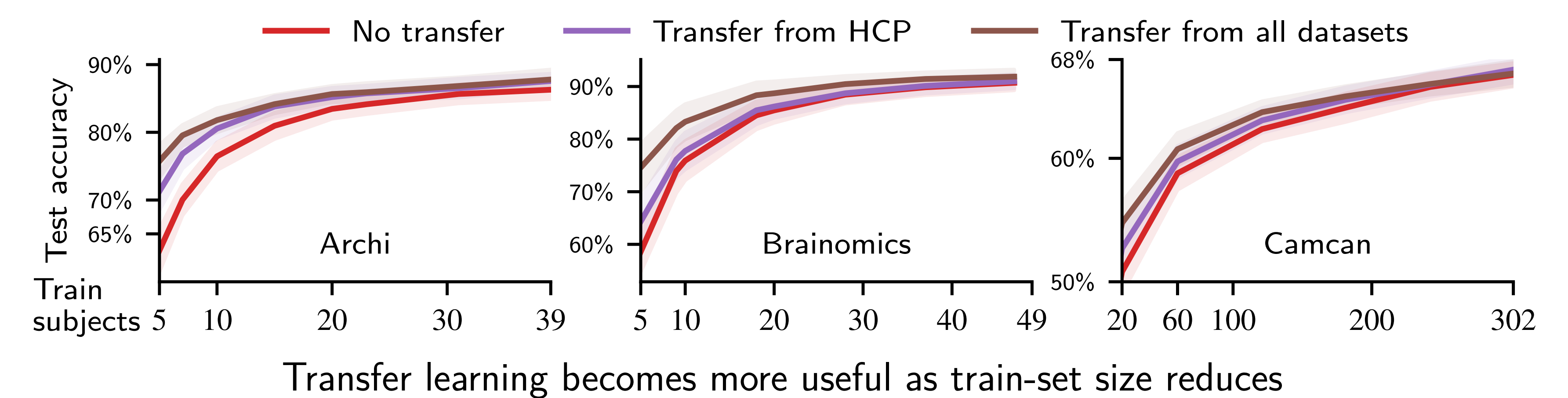
References

- [1] A. Mensch, J. Mairal, D. Bzdok, B. Thirion, and G. Varoquaux. Learning neural representations of human cognition across many fmri studies. In *Advances in Neural Information Processing Systems*, 2017.
- [2] A. Mensch, J. Mairal, B. Thirion, and G. Varoquaux. Stochastic Subsampling for Factorizing Huge Matrices. *IEEE Transactions on Signal Processing*, 66(1):113–128, 2018.

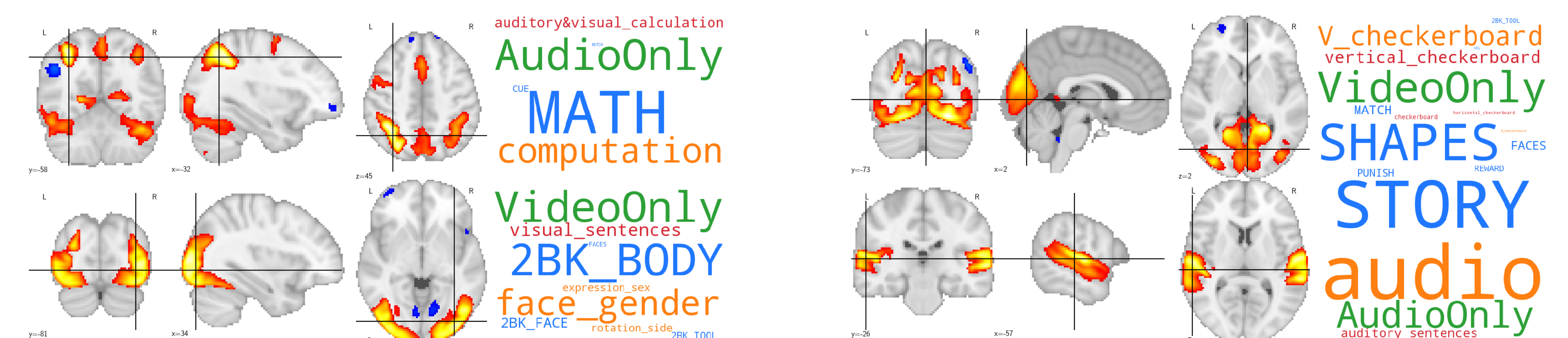
Regularization

- We allow $l > k$: not a reduced rank regression
- Trivial with ℓ_2 regularization: no transfer
- Dropout** allows transfer despite \mathbf{W}_d full rank:
(training) $\hat{\mathbf{y}} = \mathbf{x} \mathbf{W}_e \mathbf{M} \mathbf{W}'_d$, \mathbf{M} masking matrix

Transfer improves learning curves

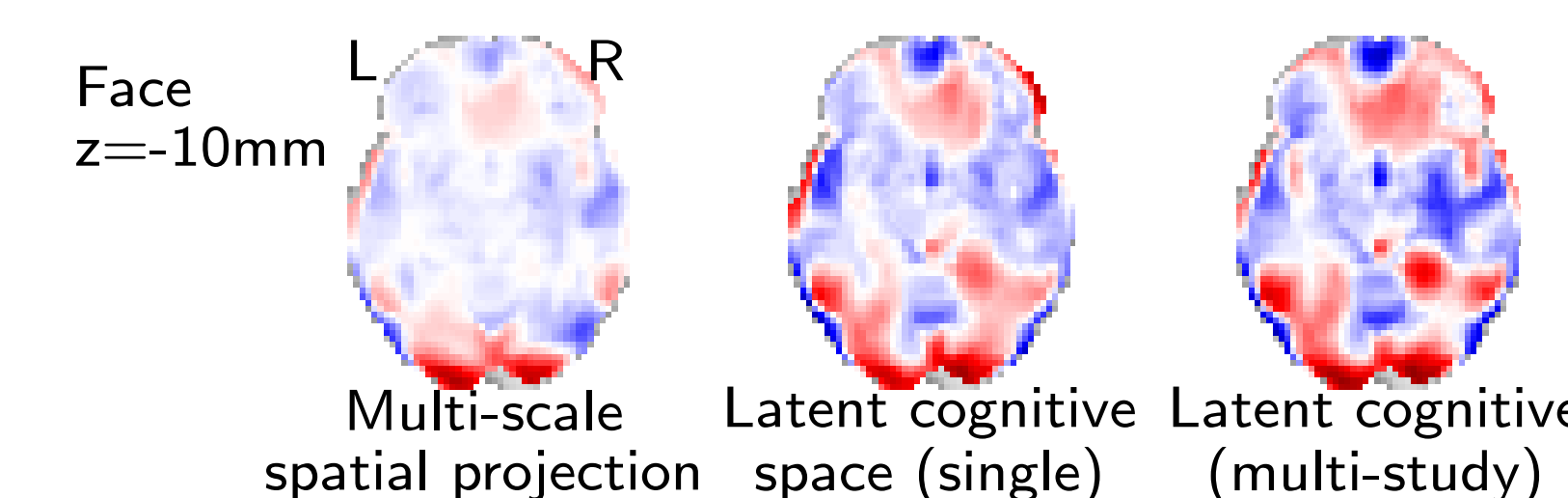


Cognitive space visualization



Meaningful template images associated to multi-dataset predictions (one color per dataset)

Interpretability



Model \rightarrow Higher-level regions (e.g. FFA)

Conclusion and future work

- Scalable** and **paradigm-agnostic** model
- With evidence of **transfer learning**
- Dictionaries + python code available
github.com/arthurmensch/cogspaces
- Model to be tested on the whole **openFMRI** repository