INRIA Saclay, Équipe Parietal http://team.inria.fr/parietal NeuroSpin, bat 145, CEA Saclay

Deep models for sensory inputs encoding: Application to brain activity analysis

Research theme: machine learning, functional brain imaging
Keywords: regression, fMRI, deep learning, representationsx, cognition.
Duration & salary: 4 to 6 months, 550 € monthly
Research team: Parietal (INRIA Saclay and CEA)
Adviser: Bertrand Thirion,
Contact: bertrand.thirion@inria.fr
Application: Interested candidate should send CV and motivation letter

Context: Encoding models of the visual cortex

The understanding of brain functional architecture has long been driven by subtractive reasoning approaches, in which the activation patterns associated with different experimental conditions presented in event-related or block designs are contrasted in order to yield condition-specific maps [1]. A more ecological way of stimulating subjects consists in presenting complex continuous stimuli that are much more similar to everyday cognitive experiences. The analysis of the ensuing complex stimulation streams proceeds by extracting relevant features from the stimuli and correlating the occurrence of these features with the brain activity recorded simultaneously with the presentation of the stimuli. Such a procedure is called an *encoding model* [2]. Importantly, encoding models can be inverted, opening the possibility to decode brain activity (see Fig. 1 and [2]).
This framework has proved very successful for the study of the visual cortex, dues to the existing relationships between computer and biological vision [3, 4]; moreover, the large size of the human visual makes it easy to study it with functional neuroimaging at a resolution of 1 to 3mm (see e.g. [5]). Recent works have provided more advanced insights on the functional organization of the visual cortex [6, 7]. A particularly compelling application of this study is the capacity to reconstruct or predict an image seen by a participant, an approach called *decoding* [8, 9, 2].

Proposed work

Despite recent progress on the topic (see [6], [7] and Fig. 2), the decoding of visual experiments is currently limited by the capacity to interpret the units of the convolutional network. We will investigate whether state-of-the-art solutions in computer vision such as *deconv net* [10] can be used to clarify the functional significance of the units and yield a clear interpretation of the associated activity. The underlying question is the statistical reliability of such reconstruction, that we will study carefully.
Another important topic in the encoding is the presence of an hemodynamic filter between the assumed neural activation and the observed fMRI fluctuations. As this filter is not perfectly known, estimating it can have a large impact on encoding and decoding accuracy [11]. We will assess the impact of such modeling on the analysis of visual activation datasets in [7].

Implementation

The expected result of will be a publication describing the mapping of high-level features to brain territories. We will focus particularly on double dissociations, where different regions exhibit distinct behavior according to the analytic model.
We will provide analytic methods that can be used to identify the features that best explain the signal across brain regions. These will be implemented in Python and made available to the community.
Finally, we will identify carefully the computational bottlenecks entailed by this project and address them in dedicated libraries (Joblib, `https://pythonhosted.org/joblib`, Scikit-Learn, `http://nilearn.github.`

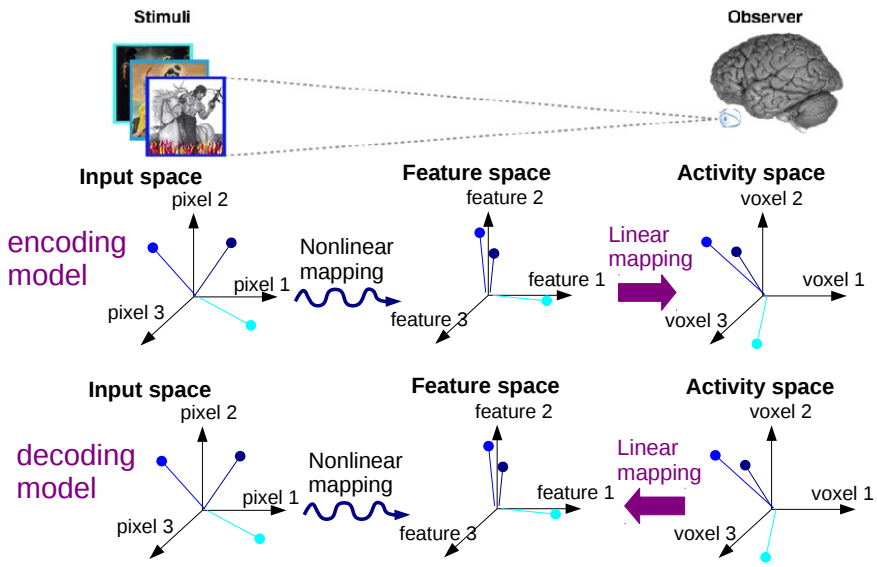

Figure 1: Encoding and decoding model of brain activation. (top) An encoding model can be conceptualized as a non-linear feature extraction from naturalistic stimuli followed by a linear mapping to brain activity. This linear mapping opens the possibility to invert the mapping, e.g. predict stimulus features from brain activity. The particular role of the assumed non-linearity is to make the mapping more efficient by extracting new representations of the stimuli that are more relevant to brain activity modeling.

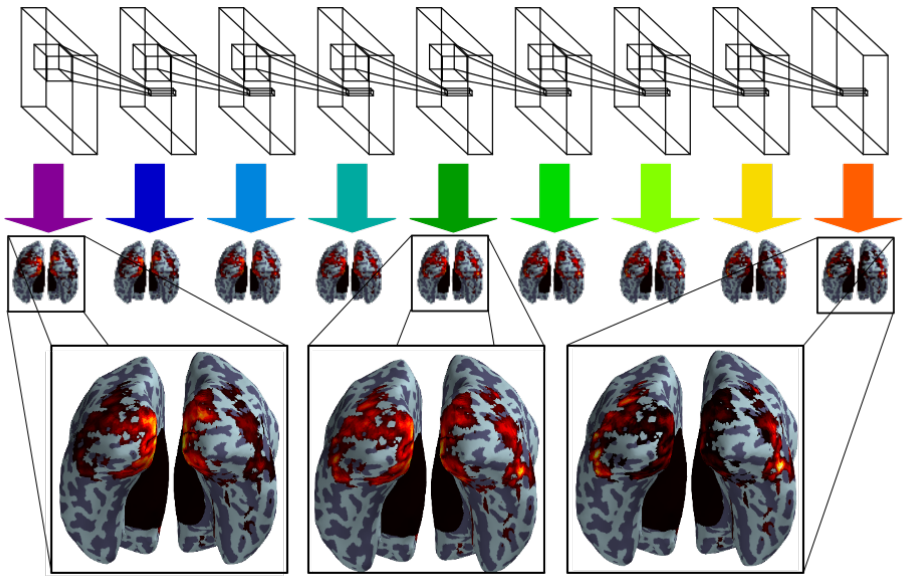

Figure 2: Deep representations of certain input stimuli may be used to map the brain's functional architecture by assessing specifically how well each layer of the deep model explains the activation in each brain area.

io, Nilearn, http://nilearn.github.io) and deep learning libraries (Pytorch, http://pytorch.org, Keras https://keras.io or the underlying libraries).

Required skills:

The successful candidate will be interested in applications of machine learning and in the understanding of human cognition.
Prior experience on deep model is a major asset, as it makes it easier for the candidate to understand the concepts and tools involved. Knowledge of scientific computing in Python (Numpy, Scipy) is required. All the work will be done in Python based on standard machine learning libraries and the Nilearn library http://nilearn.github.io for neuroimaging aspects. The candidate will benefit from the numerous development of the Tau Parietal teams for computational facilities and expertise in the various domaines involved (machine learning, optimization, statistics, neuroscience, psychology).

[1] J.-B. Poline and M. Brett, "The general linear model and fmri: does love last forever?," *Neuroimage*, vol. 62, pp. 871–880, Aug 2012.

[2] T. Naselaris, K. N. Kay, S. Nishimoto, and J. L. Gallant, "Encoding and decoding in fMRI," *Neuroimage*, vol. 56, p. 400, 2011.

[3] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms.," *IEEE Trans Pattern Anal Mach Intell*, vol. 29, pp. 411–426, Mar 2007.

[4] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Circuits and Systems (ISCAS)*, p. 253, 2010.

[5] S. O. Dumoulin and B. A. Wandell, "Population receptive field estimates in human visual cortex.," *Neuroimage*, vol. 39, pp. 647–660, Jan 2008.

[6] S.-M. Khaligh-Razavi and N. Kriegeskorte, "Deep supervised, but not unsupervised, models may explain it cortical representation.," *PLoS Comput Biol*, vol. 10, p. e1003915, Nov 2014.

[7] E. M., G. A., V. G., and T. B., "Seeing it all: Convolutional network layers map the function of the human visual system," *PLoS Comput Biol*, submitted 2015.

[8] B. Thirion, E. Duchesnay, E. Hubbard, J. Dubois, J.-B. Poline, D. Lebihan, and S. Dehaene, "Inverse retinotopy: inferring the visual content of images from brain activation patterns," *Neuroimage*, vol. 33, p. 1104, 2006.

[9] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant, "Identifying natural images from human brain activity," *Nature*, vol. 452, p. 352, 2008.

[10] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *CoRR*, vol. abs/1311.2901, 2013.

[11] F. Pedregosa, M. Eickenberg, P. Ciuciu, A. Gramfort, and B. Thirion, "Data-driven HRF estimation for encoding and decoding models," *submitted, arXiv preprint arXiv:1402.7015*, 2014.