Activity Report 2016

Project-Team PARIETAL

Modelling brain structure, function and variability based on high-field MRI data.

IN COLLABORATION WITH: CEA Neurospin

Table of contents

Project-Team PARIETAL

Creation of the Project-Team: 2009 July 01

> *The Parietal team is localized in two places: in the Alan Turing building of Inria Saclay, and the Neurospin building of CEA Saclay.*

Keywords:

Computer Science and Digital Science:

  3.3. - Data and knowledge analysis
  3.3.2. - Data mining
  3.3.3. - Big data analysis
  3.4. - Machine learning and statistics
  3.4.1. - Supervised learning
  3.4.2. - Unsupervised learning
  3.4.4. - Optimization and learning
  3.4.5. - Bayesian methods
  3.4.6. - Neural networks
  3.4.7. - Kernel methods
  3.4.8. - Deep learning
  5.3.2. - Sparse modeling and image representation
  5.3.3. - Pattern recognition
  5.9.1. - Sampling, acquisition
  5.9.2. - Estimation, modeling
  5.9.6. - Optimization tools
  6.2.4. - Statistical methods
  6.2.6. - Optimization
  8.2. - Machine learning
  8.3. - Signal analysis

Other Research Topics and Application Domains:

  1.3. - Neuroscience and cognitive science
  1.3.1. - Understanding and simulation of the brain and the nervous system
  1.3.2. - Cognitive science
  2.2.6. - Neurodegenerative diseases
  2.6.1. - Brain imaging

1. Members

Research Scientists
 Bertrand Thirion [Team leader, Inria, Research Scientist,Senior Researcher, HDR]
 Philippe Ciuciu [CEA, Research Scientist, Senior Researcher, HDR]
 Shan Liu [Inria, Research Scientist, Researcher]
 Gaël Varoquaux [Inria, Research Scientist, Researcher]

Faculty Member
 Matthieu Kowalski [Associate Professor,Univ. Paris XI, Faculty Member, until Aug 2016]

Engineers
    Kamalaker Reddy Dadi [CEA]
    Loïc Estève [Inria]
    Ana Luisa Grilo Pinho [Inria]

PhD Students
    Patricio Cerda Reyes [Inria, from Oct 2016]
    Jérome Dockès [Inria, from Apr 2016]
    Elvis Dohmatob [Inria, granted by Fondation Cooper. Scient. Campus Paris Saclay-DIGITEO]
    Loubna El Gueddari [Min. Ens. Sup. Recherche, from Oct 2016]
    Andres Hoyos Idrobo [Inria]
    Carole Lazarus [CEA]
    Arthur Mensch [Univ. Paris XI]

Post-Doctoral Fellows
    Darya Chyzhyk [Inria, from Aug 2016]
    Joke Durnez [Inria, from Nov 2016]
    Daria La Rocca [CEA]
    Andre Monteiro Manoel [CEA, from Dec 2016]
    Mehdi Rahim [Inria]

Administrative Assistants
    Régine Bricquet [Inria, until Jun 2016]
    Tiffany Caristan [Inria, from Jun 2016]

Others
    Alexandre Abraham [Inria, until Sep 2016]
    Moritz Boos [Inria, intern, from Jul 2016 until Oct 2016]
    Olivier Grisel [Inria, Engineers]
    Guillaume Lemaitre [Inria, Post-Doctoral Fellow, from Dec 2016]
    Joao Loula Guimaraes de Campos [Inria, intern, from Jun 2016]

2. Overall Objectives

2.1. Overall Objectives

The Parietal team focuses on mathematical methods for modeling and statistical inference based on neuroimaging data, with a particular interest in machine learning techniques and applications of human functional imaging. This general theme splits into four research axes:

- Modeling for neuroimaging population studies,
- Encoding and decoding models for cognitive imaging,
- Statistical and machine learning methods for large-scale data,
- Compressed-sensing for MRI.

Parietal is also strongly involved in open-source software development in scientific Python (machine learning) and for neuroimaging applications.

3. Research Program

3.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, brain population imaging is concerned with the *inverse problem* that consists in predicting individual information (behavior, phenotype) from neuroimaging data, while the corresponding *forward problem* boils down to explaining neuroimaging data with the behavioral variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data well enough ?), and a regularization scheme that represents a prior on the expected solution of the problem. These priors can be used to enforce some properties on the solutions, such as sparsity, smoothness or being piece-wise constant.

Let us detail the model used in typical inverse problem: Let $\mathbf{X}$ be a neuroimaging dataset as an $(n_{subjects}, n_{voxels})$ matrix, where $n_{subjects}$ and n_{voxels} are the number of subjects under study, and the image size respectively, $\mathbf{Y}$ a set of values that represent characteristics of interest in the observed population, written as $(n_{subjects}, n_{features})$ matrix, where $n_{features}$ is the number of characteristics that are tested, and β an array of shape $(n_{voxels}, n_{features})$ that represents a set of pattern-specific maps. In the first place, we may consider the columns $\mathbf{Y}_1, .., \mathbf{Y}_{n_{features}}$ of Y independently, yielding $n_{features}$ problems to be solved in parallel:

$$\mathbf{Y}_i = \mathbf{X}\beta_i + \epsilon_i, \forall i \in \{1, .., n_{features}\},$$

where the vector contains β_i is the ith row of β. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{\beta}_i = \mathrm{argmin}_{\beta_i} \|\mathbf{Y}_i - \mathbf{X}\beta_i\|^2 + \Psi(\beta_i), \tag{1}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\beta; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2 + \eta_1 \|\nabla\beta\|_{2,1} + \eta_2 \|\nabla\beta\|_{2,2} \tag{2}$$

with λ_1, λ_2, η_1, $\eta_2 \geq 0$ (this formulation particularly highlights the fact that convex regularizers are norms or quasi-norms). In general, only one or two of these constraints is considered (hence is enforced with a non-zero coefficient):

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution β is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.

- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piece-wise constant solution. It can be associated with Lasso to enforce both sparsity and sparse variations.

- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

Note that, while the qualitative aspect of the solutions are very different, the predictive power of these models is often very close.

The performance of the predictive model can simply be evaluated as the amount of variance in $\mathbf{Y}_i$ fitted by the model, for each $i \in \{1, .., n_{features}\}$. This can be computed through cross-validation, by *learning* $\widehat{\beta}_i$ on some part of the dataset, and then estimating $\|\mathbf{Y}_i - \mathbf{X}\widehat{\beta}_i\|^2$ using the remainder of the dataset.
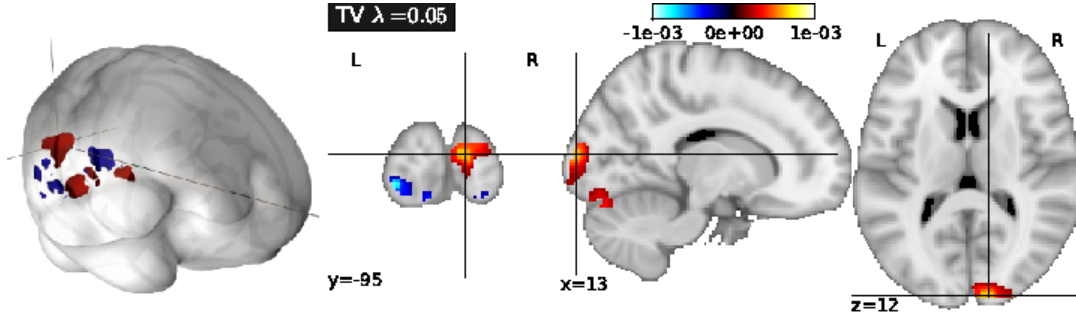
Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problem here is to predict the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the presentation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that also enjoys high predictive accuracy.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This amounts to enforcing structured priors on the problem solution.

- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.

- *Logistic and hinge regression*, where a non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.

- *Robustness to between-subject variability* to avoid the learned model overly reflecting a few outlying particular observations of the training set. Note that noise and deviating assumptions can be present in both $\mathbf{Y}$ and $\mathbf{X}$

- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector β to have a shared support across all these variables.
  For instance, when one of the variables $\mathbf{Y}_i$ is not well fitted by the model, the estimation of other variables $\mathbf{Y}_j, j \neq i$ may provide constraints on the support of β_i and thus, improve the prediction of $\mathbf{Y}_i$.

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \tag{3}$$

then

$$\widehat{\beta} = \text{argmin}_{\beta=(\beta_i), i=1..n_f} \sum_{i=1}^{n_f} \|\mathbf{Y_i} - \mathbf{X}\beta_\mathbf{i}\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \beta_{\mathbf{i,j}}^\mathbf{2}} \tag{4}$$

3.2. Multivariate decompositions

Multivariate decompositions provide a way to model complex data such as brain activation images: for instance, one might be interested in extracting an *atlas of brain regions* from a given dataset, such as regions exhibiting similar activity during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.

Let $\mathbf{X}$ be a neuroimaging dataset written as an $(n_{subjects}, n_{voxels})$ matrix, after proper centering; the model reads

$$\mathbf{X} = \mathbf{AD} + \epsilon, \tag{5}$$

where $\mathbf{D}$ represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and $\mathbf{A}$ the associated subject-wise loadings. While traditional PCA and independent components analysis are limited to reconstructing components $\mathbf{D}$ within the space spanned by the column of $\mathbf{X}$, it seems desirable to add some constraints on the rows of $\mathbf{D}$, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging. This yields the following estimation problem:

$$\min_{\mathbf{D},\mathbf{A}} \|\mathbf{X} - \mathbf{AD}\|^2 + \Psi(\mathbf{D}) \text{ s.t. } \|\mathbf{A}_i\| = 1 \ \forall i \in \{1..n_{features}\}, \tag{6}$$

where $(\mathbf{A}_i)$, $i \in \{1..n_{features}\}$ represents the columns of $\mathbf{A}$. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.

The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on $\mathbf{D}$ for $\mathbf{A}$ fixed, and conversely. This readily suggests an alternate optimization scheme, where $\mathbf{D}$ and $\mathbf{A}$ are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

Utlimately, the main limitations to these algorithms is the cost due to the memory requirements: holding datasets with large dimension and large number of samples (as in recent neuroimaging cohorts) leads to inefficient computation. To solve this issue, online method are particularly attractive.

3.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to study potential interactions between variables. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.

The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.

- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is still an open question.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-samples likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we consider different means of parametrizing covariance distributions and how these parametrizations impact the test of statistical differences across individuals.
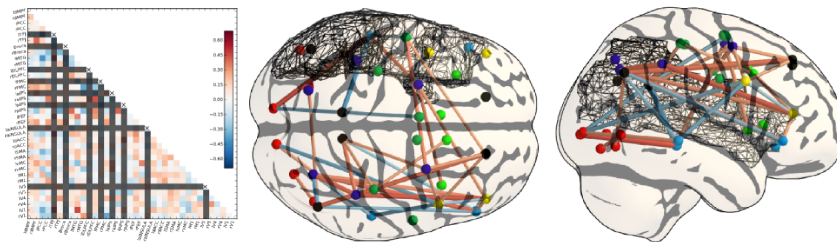
Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion volume outlined as a mesh) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

4. Application Domains

4.1. Cognitive neuroscience

4.1.1. *Macroscopic Functional cartography with functional Magnetic Resonance Imaging (fMRI)*

The brain as a highly structured organ, with both functional specialization and a complex newtork organization. While most of the knowledge historically comes from lesion studies and animal electophysiological recordings, the development of non-invasive imaging modalities, such as fMRI, has made it possible to study routinely high-level cognition in humans since the early 90's. This has opened major questions on the interplay between mind and brain , such as: How is the function of cortical territories constrained by anatomy (connectivity) ? How to assess the specificity of brain regions ? How can one characterize reliably inter-subject differences ?

4.1.2. *Analysis of brain Connectivity*

Functional connectivity is defined as the interaction structure that is underlies brain function. Since the beginning of fMRI, it has been observed that remote regions sustain high correlation in their spontaneous activity, i.e. in the absence of a driving task. This means that the signals observed during resting-state define a signature of the connectivity of brain regions. The main interest of retsing-state fMRI is that it provides easy-to-acquire functional markers that have recently been proved to be very powerful for population studies.

4.1.3. Modeling of brain processes (MEG)

While fMRI has been very useful in defining the function of regions at the mm scale, Magneto-encephalography (MEG) provides the other piece of the puzzle, namely temporal dynamics of brain activity, at the ms scale. MEG is also non-invasive. It makes it possible to keep track of precise schedule of mental operations and their interactions. It also opens the way toward a study of the rythmic activity of the brain. On the other hand, the localization of brain activity with MEG entails the solution of a hard inverse problem.

5. New Software and Platforms

5.1. Mayavi

FUNCTIONAL DESCRIPTION

Mayavi is the most used scientific 3D visualization Python software. Mayavi can be used as a visualization tool, through interactive command line or as a library. It is distributed under Linux through Ubuntu, Debian, Fedora and Mandriva, as well as in PythonXY and EPD Python scientific distributions. Mayavi is used by several software platforms, such as PDE solvers (fipy, sfepy), molecule visualization tools and brain connectivity analysis tools (connectomeViewer).

- Contact: Gaël Varoquaux
- URL: http://mayavi.sourceforge.net/

5.2. Nilearn

NeuroImaging with scikit learn
KEYWORDS: Health - Neuroimaging - Medical imaging
FUNCTIONAL DESCRIPTION

NiLearn is the neuroimaging library that adapts the concepts and tools of scikit-learn to neuroimaging problems. As a pure Python library, it depends on scikit-learn and nibabel, the main Python library for neuroimaging I/O. It is an open-source project, available under BSD license. The two key components of NiLearn are i) the analysis of functional connectivity (spatial decompositions and covariance learning) and ii) the most common tools for multivariate pattern analysis. A great deal of efforts has been put on the efficiency of the procedures both in terms of memory cost and computation time.

- Participants: Gaël Varoquaux, Bertrand Thirion, Loïc Estève, Alexandre Abraham, Michael Eickenberg, Alexandre Gramfort, Fabian Pedregosa Izquierdo, Elvis Dohmatob and Virgile Fritsch
- Contact: Bertrand Thirion
- URL: http://nilearn.github.io/

5.3. PyHRF

KEYWORDS: FMRI - Statistic analysis - Neurosciences - IRM - Brain - Health - Medical imaging
FUNCTIONAL DESCRIPTION

As part of fMRI data analysis, PyHRF provides a set of tools for addressing the two main issues involved in intra-subject fMRI data analysis : (i) the localization of cerebral regions that elicit evoked activity and (ii) the estimation of the activation dynamics also referenced to as the recovery of the Hemodynamic Response Function (HRF). To tackle these two problems, PyHRF implements the Joint Detection-Estimation framework (JDE) which recovers parcel-level HRFs and embeds an adaptive spatio-temporal regularization scheme of activation maps.

- Participants: Thomas Vincent, Solveig Badillo, Lotfi Chaari, Christine Bakhous, Florence Forbes, Philippe Ciuciu, Laurent Risser, Thomas Perret and Aina Frau Pascual
- Partners: CEA - NeuroSpin
- Contact: Florence Forbes
- URL: http://pyhrf.org

5.4. Scikit-learn

KEYWORDS: Classification - Learning - Clustering - Regession - Medical imaging

SCIENTIFIC DESCRIPTION

Scikit-learn is a Python module integrating classic machine learning algorithms in the tightly-knit scientific Python world. It aims to provide simple and efficient solutions to learning problems, accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.

FUNCTIONAL DESCRIPTION

Scikit-learn can be used as a middleware for prediction tasks. For example, many web startups adapt Scikitlearn to predict buying behavior of users, provide product recommendations, detect trends or abusive behavior (fraud, spam). Scikit-learn is used to extract the structure of complex data (text, images) and classify such data with techniques relevant to the state of the art.

Easy to use, efficient and accessible to non datascience experts, Scikit-learn is an increasingly popular machine learning library in Python. In a data exploration step, the user can enter a few lines on an interactive (but non-graphical) interface and immediately sees the results of his request. Scikitlearn is a prediction engine . Scikit-learn is developed in open source, and available under the BSD license.

- Participants: Olivier Grisel, Gaël Varoquaux, Bertrand Thirion, Michael Eickenberg, Loïc Estève, Alexandre Gramfort, Arthur Mensch
- Partners: CEA - Logilab - Nuxeo - Saint Gobain - Telecom Paris - Tinyclues
- Contact: Olivier Grisel
- URL: http://scikit-learn.org

6. New Results

6.1. Dictionary Learning for Massive Matrix Factorization

Sparse matrix factorization is a popular tool to obtain interpretable data decompositions, which are also effective to perform data completion or denoising. Its applicability to large datasets has been addressed with online and randomized methods, that reduce the complexity in one of the matrix dimension, but not in both of them. In this paper, we tackle very large matrices in both dimensions. We propose a new factoriza-tion method that scales gracefully to terabyte-scale datasets, that could not be processed by previous algorithms in a reasonable amount of time. We demonstrate the efficiency of our approach on massive functional Magnetic Resonance Imaging (fMRI) data, and on matrix completion problems for recommender systems, where we obtain significant speed-ups compared to state-of-the art coordinate descent methods.

See Fig. 3 for an illustration and [22] for more information.

6.2. Learning brain regions via large-scale online structured sparse dictionary-learning

We propose a multivariate online dictionary-learning method for obtaining de-compositions of brain images with structured and sparse components (aka atoms). Sparsity is to be understood in the usual sense: the dictionary atoms are constrained to contain mostly zeros. This is imposed via an 1-norm constraint. By "structured", we mean that the atoms are piece-wise smooth and compact, thus making up blobs, as opposed to scattered patterns of activation. We propose to use a Sobolev (Laplacian) penalty to impose this type of structure. Combining the two penalties, we obtain decompositions that properly delineate brain structures from functional images. This non-trivially extends the online dictionary-learning work of Mairal et al. (2010), at the price of only a factor of 2 or 3 on the overall running time. Just like the Mairal et al. (2010) reference method, the online nature of our proposed algorithm allows it to scale to arbitrarily sized datasets. Experiments on brain data show that our proposed method extracts structured and denoised dictionaries that are more intepretable and better capture inter-subject variability in small medium, and large-scale regimes alike, compared to state-of-the-art models.
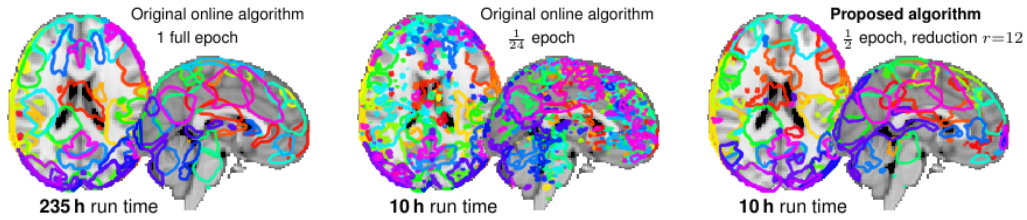
Figure 3. Brain atlases: outlines of each map obtained with dictionary learning. Left: the reference algorithm on the full dataset. Middle: the reference algorithm on a twentieth of the dataset. Right: the proposed algorithm with a similar run time: half the dataset and a compression factor of 9. Compared to a full run of the baseline algorithm, the figure explore two possible strategies to decrease computation time: processing less data (middle), or our approach (right). Our approach achieves a result closer to the gold standard in a given time budget. See [22] for more information.
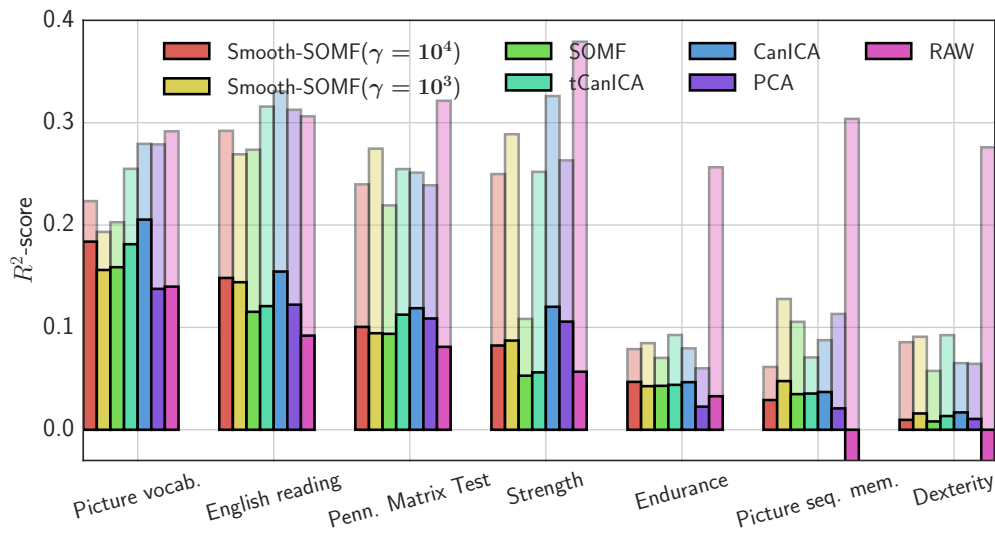
Figure 4. Predicting behavioral variables of the Human Connectome Project dataset using subject-level brain activity maps and various intermediate representations obtained with variants of dictionary learning. Bold bars represent performance on test set while faint bars in the background represent performance on train set. See [19] for more information.

See Fig. 4 for an illustration and [19] for more information.

6.3. Social-sparsity brain decoders: faster spatial sparsity

Spatially-sparse predictors are good models for brain decoding: they give accurate predictions and their weight maps are interpretable as they focus on a small number of regions. However, the state of the art, based on total variation or graph-net, is computationally costly. Here we introduce sparsity in the local neighborhood of each voxel with social-sparsity, a structured shrinkage operator. We find that, on brain imaging classification problems, social-sparsity performs almost as well as total-variation models and better than graph-net, for a fraction of the computational cost. It also very clearly outlines predictive regions. We give details of the model and the algorithm.
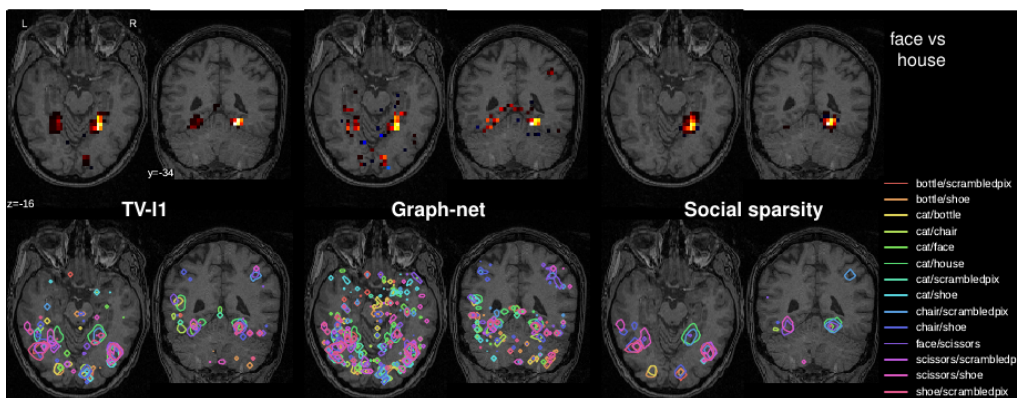
Figure 5. Decoder maps for the object-classification task – Top: weight maps for the face-versus-house task. Overall, the maps segment the right and left parahippocampal place area (PPA), a well-known place-specific regions, although the left PPA is weak in TV-l1, spotty in graph-net, and absent in social sparsity. Bottom: outlines at 0.01 of the other tasks. Beyond the PPA, several known functional regions stand out such as primary or secondary visual areas around the prestriate cortex as well as regions in the lateral occipital cortex, responding to structured objects. Note that the graphnet outlines display scattered small regions even thought the value of the contours is chosen at 0.01, well above numerical noise. See [32] for more information.

See Fig. 5 for an illustration and [32] for more information.

6.4. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example

Resting-state functional Magnetic Resonance Imaging (R-fMRI) holds the promise to reveal functional biomarkers of neuropsychiatric disorders. However, extracting such biomarkers is challenging for complex multi-faceted neuropatholo-gies, such as autism spectrum disorders. Large multi-site datasets increase sample sizes to compensate for this complexity, at the cost of uncontrolled heterogeneity. This heterogeneity raises new challenges, akin to those face in realistic diagnostic applications. Here, we demonstrate the feasibility of inter-site classification of neuropsychiatric status, with an application to the Autism Brain Imaging Data Exchange (ABIDE) database, a large (N=871) multi-site autism dataset. For this purpose, we investigate pipelines that extract the most predictive biomarkers from the data. These R-fMRI pipelines build participant-specific connectomes from functionally-defined brain areas. Connectomes are then compared across participants to learn patterns of connectivity that differentiate typical controls from individuals with autism. We predict this neuropsychiatric status for participants from the same acquisition sites or different, unseen, ones. Good choices

of methods for the various steps of the pipeline lead to 67% prediction accuracy on the full ABIDE data, which is significantly better than previously reported results. We perform extensive validation on multiple subsets of the data defined by different inclusion criteria. These enables detailed analysis of the factors contributing to successful connectome-based prediction. First, prediction accuracy improves as we include more subjects, up to the maximum amount of subjects available. Second, the definition of functional brain areas is of paramount importance for biomarker discovery: brain areas extracted from large R-fMRI datasets outperform reference atlases in the classification tasks.
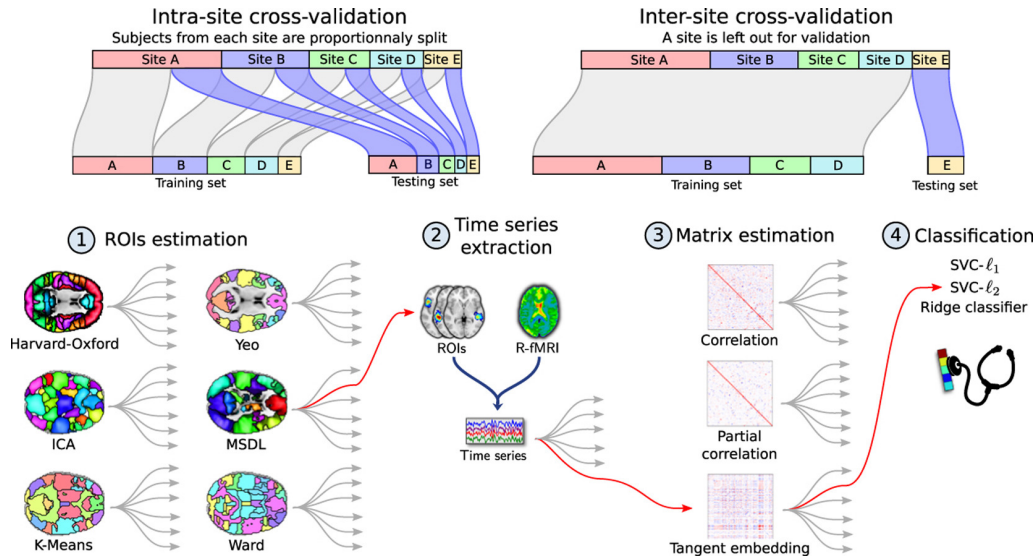
Figure 6. Validation of an fMRI-based pipeline for autism prediction. Several variants are considered for each pipeline step. See [1] for more information.

See Fig. 6 for an illustration and [1] for more information.

6.5. Seeing it all: Convolutional network layers map the function of the human visual system

Convolutional networks used for computer vision represent candidate models for the computations performed in mammalian visual systems. We use them as a detailed model of human brain activity during the viewing of natural images by constructing predictive models based on their different layers and BOLD fMRI activations. Analyzing the predictive performance across layers yields characteristic fingerprints for each visual brain region: early visual areas are better described by lower level convolutional net layers and later visual areas by higher level net layers, exhibiting a progression across ventral and dorsal streams. Our predictive model generalizes beyond brain responses to natural images. We illustrate this on two experiments, namely retinotopy and face-place oppositions, by synthesizing brain activity and performing classical brain mapping upon it. The synthesis recovers the activations observed in the corresponding fMRI studies, showing that this deep encoding model captures representations of brain function that are universal across experimental paradigms.

See Fig. 7 for an illustration and [10] for more information.

6.6. Formal Models of the Network Co-occurrence Underlying Mental Operations
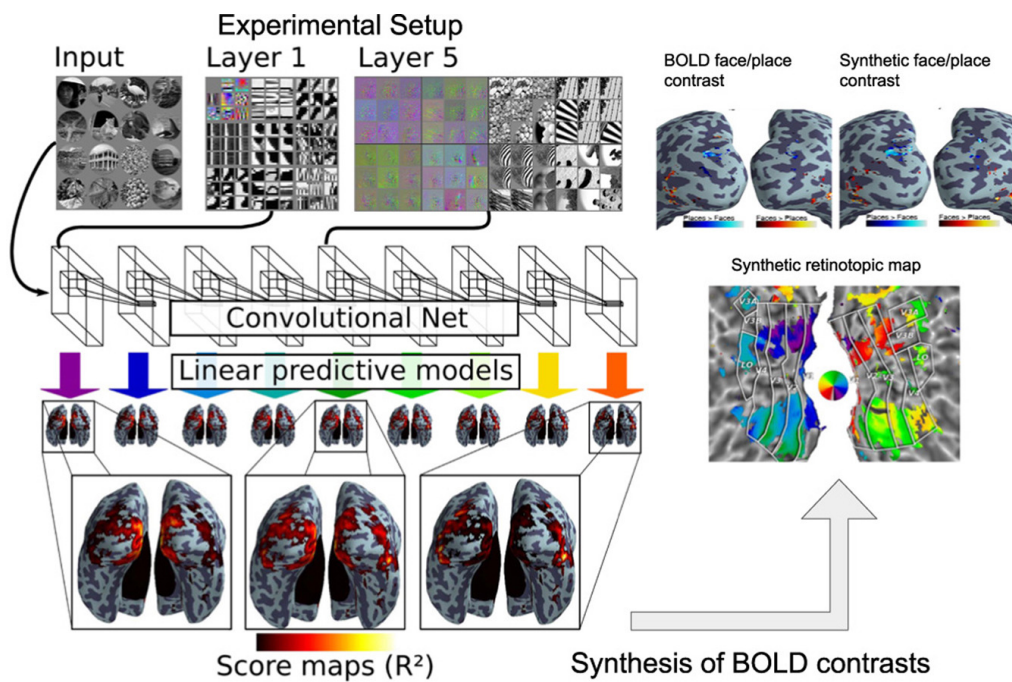
Figure 7. Overview of the vision mapping experiment: Convolutional network image representations of different layer depth explain brain activity throughout the full ventral visual stream. This mapping follows the known hierarchical organisation. Results from both static images and video stimuli. A model of brain activity for the full brain, based on the convolutional network, can synthesize brain maps for other visual experiments. Only deep models can reproduce observed BOLD activity. See [10] for more information.

Systems neuroscience has identified a set of canonical large-scale networks in humans. These have predominantly been characterized by resting-state analyses of the task-uncon-strained, mind-wandering brain. Their explicit relationship to defined task performance is largely unknown and remains challenging. The present work contributes a multivariate statistical learning approach that can extract the major brain networks and quantify their configuration during various psychological tasks. The method is validated in two extensive datasets (n = 500 and n = 81) by model-based generation of synthetic activity maps from recombination of shared network topographies. To study a use case, we formally revisited the poorly understood difference between neural activity underlying idling versus goal-directed behavior. We demonstrate that task-specific neural activity patterns can be explained by plausible combinations of resting-state networks. The possibility of decomposing a mental task into the relative contributions of major brain networks, the "network co-occurrence architecture" of a given task, opens an alternative access to the neural substrates of human cognition.

See Fig. 8 for an illustration and [6] for more information.

6.7. Transmodal Learning of Functional Networks for Alzheimer's Disease Prediction

Functional connectivity describes neural activity from resting-state functional magnetic resonance imaging (rs-fMRI). This noninvasive modality is a promising imaging biomarker of neurodegenerative diseases, such as Alzheimer's disease (AD), where the connectome can be an indicator to assess and to understand the pathology. However, it only provides noisy measurements of brain activity. As a consequence, it has shown fairly limited discrimination power on clinical groups. So far, the reference functional marker of AD is the fluorodeoxyglucose positron emission tomography (FDG-PET). It gives a reliable quantification of metabolic activity, but it is costly and invasive. Here, our goal is to analyze AD populations solely based on rs-fMRI, as functional connectivity is correlated to metabolism. We introduce transmodal learning: leveraging a prior from one modality to improve results of another modality on different subjects. A metabolic prior is learned from an independent FDG-PET dataset to improve functional connectivity-based prediction of AD. The prior acts as a regularization of connectivity learning and improves the estimation of discriminative patterns from distinct rs-fMRI datasets. Our approach is a two-stage classification strategy that combines several seed-based connectivity maps to cover a large number of functional networks that identify AD physiopathology. Experimental results show that our transmodal approach increases classification accuracy compared to pure rs-fMRI approaches, without resorting to additional invasive acquisitions. The method successfully recovers brain regions known to be impacted by the disease.

6.8. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines

Decoding, ie prediction from brain images or signals, calls for empirical evaluation of its predictive power. Such evaluation is achieved via cross-validation, a method also used to tune decoders' hyper-parameters. This paper is a review on cross-validation procedures for decoding in neuroimaging. It includes a didactic overview of the relevant theoretical considerations. Practical aspects are highlighted with an extensive empirical study of the common decoders in within-and across-subject predictions, on multiple datasets –anatomical and functional MRI and MEG– and simulations. Theory and experiments outline that the popular " leave-one-out " strategy leads to unstable and biased estimates, and a repeated random splits method should be preferred. Experiments outline the large error bars of cross-validation in neuroimaging settings: typical confidence intervals of 10%. Nested cross-validation can tune decoders' parameters while avoiding circularity bias. However we find that it can be more favorable to use sane defaults, in particular for non-sparse decoders.

See Fig. 9 for an illustration and [16] for more information.

6.9. A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging
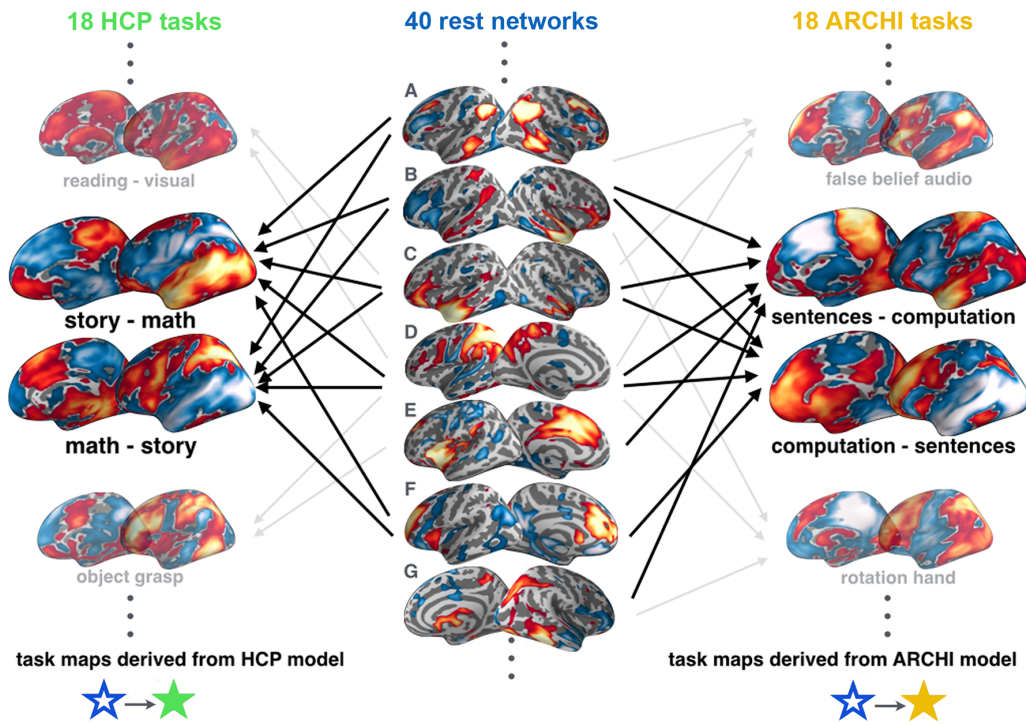
Figure 8. Task-rest correspondence: Reconstructing two similar tasks from two different datasets based on the same resting networks. 40 sparse PCA networks were discovered from the same rest data and used for feature engineering as a basis for classificationof 18 psychological tasks from HCP (left) and from ARCHI (right). Middle column: Examples of resting-state networks derived from decomposing rest data using sparse PCA. Networks B and C might be related to semantics processing in the anterior temporal lobe, network D covers extended parts of the parietal cortex, while networks E and F appear to be variants of the so-called "salience" network. Left/Right column: Examples of task-specific neural activity generated from network co-occurrence models of the HCP/ARCHI task batteries. Arrows: A diagnostic subanalysis indicated what rest networks were automatically ranked top-five in distinguishing a given task from the respective 17 other tasks. Although the experimental tasks in the HCP and ARCHI repositories, "story versus math" and "sentences versus computation" were the most similar cognitive contrasts in both datasets. For these four experimental conditions the model-derived task maps are highly similar. Consequently, two independent classification problems in two independent datasets with a six-fold difference in sample size resulted in two independent explicit models that, nevertheless, generated comparable task-specific maps. This indicated that network co-occurrence modeling indeed captures genuine aspects of neurobiology rather than arbitrary discriminatory aspects of the data. See [6] for more information.
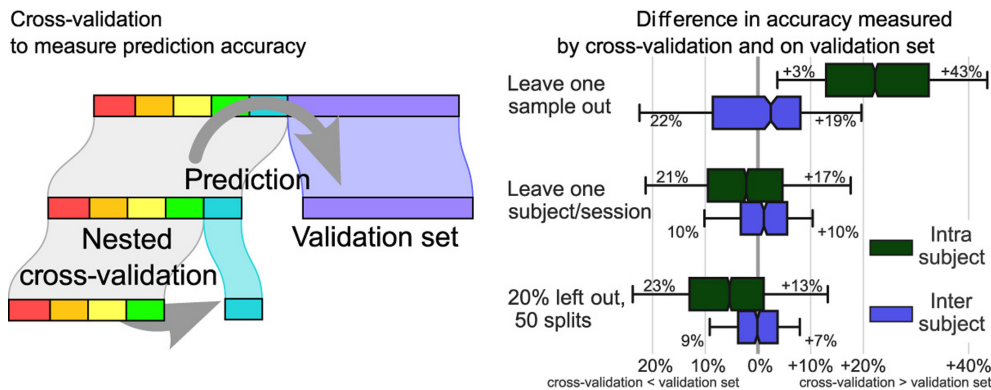
Figure 9. (Left) Illustration of the nested cross-validation principle. (Right) Typical cross-validated accuracy result: leave-one-out cross validation, when applied to imaging data, yields to optimistic bias (top) when used on dependent data, and in other cases leads to estimated with inflated variance. See [16] for more information.

Collecting the maximal amount of information in a given scanning time is a major concern in Magnetic Resonance Imaging (MRI) to speed up image acquisition. The hardware constraints (gradient magnitude, slew rate, ...), physical distortions (e.g., off-resonance effects) and sampling theorems (Shannon, compressed sensing) must be taken into account simultaneously, which makes this problem extremely challenging. To date, the main approach to design gradient waveform has consisted of selecting an initial shape (e.g. spiral, radial lines, ...) and then traversing it as fast as possible using optimal control. In this paper, we propose an alternative solution which first consists of defining a desired parameterization of the trajectory and then of optimizing for minimal deviation of the sampling points within gradient constraints. This method has various advantages. First, it better preserves the density of the input curve which is critical in sampling theory. Second, it allows to smooth high curvature areas making the acquisition time shorter in some cases. Third, it can be used both in the Shannon and CS sampling theories. Last, the optimized trajectory is computed as the solution of an efficient iterative algorithm based on convex programming. For piecewise linear trajectories, as compared to optimal control reparameterization, our approach generates a gain in scanning time of 10% in echo planar imaging while improving image quality in terms of signal-to-noise ratio (SNR) by more than 6 dB. We also investigate original trajectories relying on traveling salesman problem solutions. In this context, the sampling patterns obtained using the proposed projection algorithm are shown to provide significantly better reconstructions (more than 6 dB) while lasting the same scanning time.

See Fig. 10 for an illustration and [9] for more information.

6.10. Impact of perceptual learning on resting-state fMRI connectivity: A supervised classification study

Perceptual learning sculpts ongoing brain activity. This finding has been observed by statistically comparing the functional connectivity (FC) patterns computed from resting-state functional MRI (rs-fMRI) data recorded before and after intensive training to a visual attention task. Hence, functional connectivity serves a dynamic role in brain function, supporting the consolidation of previous experience. Following this line of research, we trained three groups of individuals to a visual discrimination task during a magneto-encephalography (MEG) experiment. The same individuals were then scanned in rs-fMRI. Here, in a supervised classification framework, we demonstrate that FC metrics computed on rs-fMRI data are able to predict the type of training the participants received. On top of that, we show that the prediction accuracies based on tangent embedding
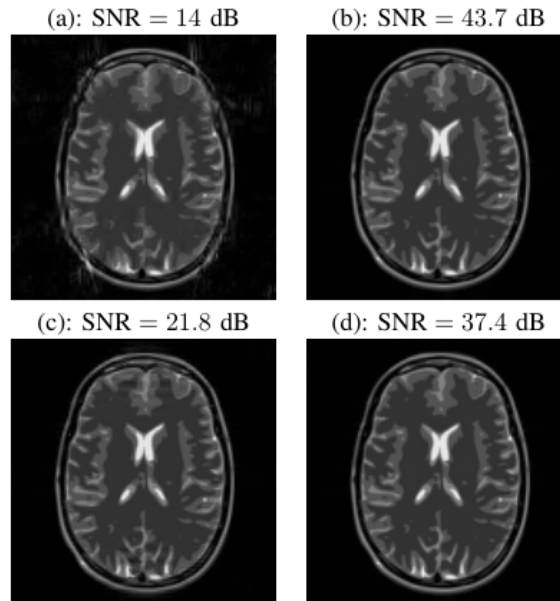
Figure 10. Reconstructed images from data collected along EPI-like trajectories. (a)-(b): Reconstruction results from the optimally reparameterized EPI readout. (c)-(d): Reconstructed results from data collected using the projected EPI trajectories. See [9] for more information.

FC measure outperform those based on our recently developed multivariate wavelet-based Hurst exponent estimator, which captures low frequency fluctuations in ongoing brain activity too.
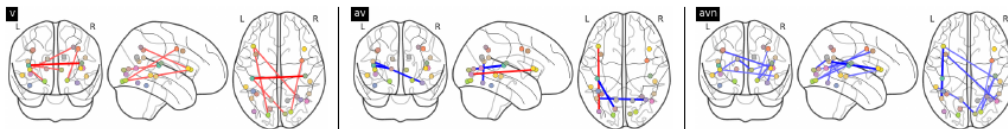
Figure 11. Statistical significant functional interactions (positive and negative values are color coded in red and blue, respectively) within each group of individuals (V: purely visual traing, AV: audio-visual training and AVn: unmatched audio-visual), Bonferroni-corrected for multiple comparisons at $\alpha = 0.05$. See [24] for more information.

See Fig. 11 for an illustration and [24] for more information.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

7.1.1. The Wendelin FUI project

The Wendelin project has been granted on December 3rd, 2014. It has been selected at the *Programme d'Investissements d'Avenir (PIA)* that supports "cloud computing et Big Data". It gives visibility and fosters the French technological big data sector, and in particular the scikit-learn library, the NoSQL "NEO" et the decentralized "SlapOS" cloud, three open-source software supported by the Systematic *pôle de compétitivité*.

Scikit-learn is a worldwide reference library for machine learning. Gaël Varoquaux, Olivier Grisel and Alexandre Gramfort have been major players in the design of the library and Scikit-learn has then been supported by the growing scientific Python community. It is currently used by major internet companies as well as dynamic start-ups, including Google, Airbnb, Spotify, Evernote, AWeber, TinyClues; it wins more than half of the data science "Kaggle" competitions. Scikit-learn makes it possible to predict future outcomes given a training data, and thus to optimize company decisions. Almost 1 million euros will be invested to improve the algorithmic core of scikit-learn through the Wendelin project thanks to the Inria, ENS and Institut Mines Télécom teams. In particular, scikit-learn will be extended in order to ease online prediction and to include recent stochastic gradient algorithms.

NEO is the native NoSQL base of the Python language. It was initially designed by Nexedi and is currently used and embedded in the main software of company information systems. More than one million euros will be invested into NEO, so that scikit-learn can process within 10 years (out-of-core) data of 1 exabyte size.

Paris13 university and the Mines Télécom institute will extend the SlapOS distributed mesh cloud to deploy Wendelin in *Big Data as a Service* (BDaaS) mode, to achieve the interoperability between the Grid5000 and Teralab infrastructures and to extend the cloud toward smart sensor systems.

The combination of scikit-learn, NEO and SlapOS will improve the predictive maintenance of industrial plants with two major use cases: connected windmills (GDF SUEZ, Woelfel) and customer satisfaction in car sale systems (MMC Rus). In both cases it is about non-personal, yet profitable big data. The Wendelin project actually demonstrates that Big data can improve infrastructure and everyday-life equipment without intrusive data collection. For more information, please see http://www.wendelin.io.

The project partners are:
- Nexedi (leader)
- GDF SUEZ
- Abilian
- 2ndQuadrant
- Institut Mines Télécom
- Inria
- Université Paris 13

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. *CoSmic project*

Participants: Philippe Ciuciu [Correspondant], Carole Lazarus, Loubna El Gueddari.

This is a collaborative project with Jean-Luc Stark, (CEA) funded by the CEA program drf-impulsion.

Compressed Sensing is a recent theory in maths that allows the perfect recovery of signals or images from compressive acquisition scenarios. This approach has been popularized in MRI over the last decade as well as in astrophysics (noticeably in radio-astronomy). So far, both of these fields have developed skills in CS separately. The aim of the COSMIC project is to foster collaborations between CEA experts in MRI (Inria-CEA Parietal team within NeuroSpin) and in astrophysics (CosmoStat lab within the Astrophysics Department). These interactions will allow us to share different expertise in order to improve image quality, either in MRI or in radio-astronomy (thanks to the interferometry principle). In this field, given the data delivered by radio-telescopet he goal will consist of extracting high temporal resolution information in order to study fast transient events.

8.1.2. *BrainAMP project*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Andre Monteiro Manoel.

This is a collaborative project with Lenka Zdeborová, Theoretical Physics Institute (CEA) funded by the CEA program drf-impulsion.

In many scientific fields, the data acquisition devices have benefited of hardware improvement to increase the resolution of the observed phenomena, leading to ever larger datasets. While the dimensionality has increased, the number of samples available is often limited, due to physical or financial limits. This is a problem when these data are processed with estimators that have a large sample complexity, such as multivariate statistical models. In that case it is very useful to rely on structured priors, so that the results reflect the state of knowledge on the phenomena of interest. The study of the human brain activity through high-field MRI belongs among these problems, with up to 10^6 features, yet a set of observations limited by cost and participant comfort.

We are missing fast estimators for multivariate models with structured priors, that furthermore provide statistical control on the solution. Approximate message passing methods are designed to work optimally with low-sample-complexity, they accommodate rather generic class of priors and come with an estimation of statistical significance. They are therefore well suited for our purposes.

We want to join forces to design a new generation of inverse problem solvers that can take into account the complex structure of brain images and provide guarantees in the low-sample-complexity regime. To this end, we will first adapt AMP to the brain mapping setting, using first standard sparsity priors (e.g. Gauss-Bernoulli) on the model. We will then consider more complex structured priors that control the variation of the learned image patterns in space. Crucial gains are expected from the use of the EM algorithm for parameter setting, that comes naturally with AMP. We will also examine the estimators provided by AMP for statistical significance. BrainAMP will design a reference inference toolbox released as a generic open source library. We expect a 3- to 10-fold improvement in CPU time, that will benefit to large-scale brain mapping investigations.

8.1.3. *iConnectom project*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Elvis Dohmatob.

This is a Digiteo project (2014-2017).

Mapping brain functional connectivity from functional Magnetic Resonance Imaging (MRI) data has become a very active field of research. However, analysis tools are limited and many important tasks, such as the empirical definition of brain networks, remain difficult due to the lack of a good framework for the statistical modeling of these networks. We propose to develop population models of anatomical and functional connectivity data to improve the alignment of subjects brain structures of interest while inferring an average template of these structures. Based on this essential contribution, we will design new statistical inference procedures to compare the functional connections between conditions or populations and improve the sensitivity of connectivity analysis performed on noisy data. Finally, we will test and validate the methods on multiple datasets and distribute them to the brain imaging community.

8.1.4. *MetaCog project*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Jérome Dockès.

This is a Digicosme project (2016-2019) and a collaboration with Fabian Suchanek (Telecom Paritech).

Understanding how cognition emerges from the billions of neurons that constitute the human brain is a major open problem in science that could bridge natural science –biology– to humanities –psychology. Psychology studies performed on humans with functional Magnetic Resonance Imaging (fMRI) can be used to probe the full repertoire of high-level cognitive functions. While analyzing the resulting image data for a given experiment is a relatively well-mastered process, the challenges in comparing data across multiple datasets poses serious limitation to the field. Indeed, such comparisons require to pool together brain images acquired under different settings and assess the effect of different *experimental conditions* that correspond to psychological effects studied by neuroscientists.

Such meta-analyses are now becoming possible thanks to the development of public data resources –OpenfMRI http://openfmri.org and NeuroVault http://neurovault.org. As many others, researchers of the Parietal team understand these data sources well and contribute to them. However, in such open-ended context, the description of experiments in terms of cognitive concepts is very difficult: there is no universal definition of cognitive terms that could be employed consistently by neuroscientists. Hence meta-analytic studies loose power and specificity. On the other hand, http://brainspell.org provide a set of curated annotation, albeit on much less data, that can serve as a seed or a ground truth to define a consensual ontology of cognitive concepts. Relating these terms to brain activity poses another challenge, of statistical nature, as brain patterns form high-dimensional data in perspective with the scarcity and the noise of the data.

The purpose of this project is to learn a semantic structure in cognitive terms from their occurrence in brain activations. This structure will simplify massive multi-label statistical-learning problems that arise in brain mapping by providing compact representations of cognitive concepts while capturing the imprecision on the definition these concepts.

8.1.5. CDS2

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Guillaume Lemaitre.

CDS2 is an "Strategic research initiatice" of the Paris Saclay University Idex http://datascience-paris-saclay.fr. Although it groups together many partners of the Paris Saclay ecosystem, Parietal has been deeply involved in the project. It currently funds a post-doc for Guillume Lemaitre.

8.2. National Initiatives

8.2.1. ANR

8.2.1.1. MultiFracs project
Participants: Philippe Ciuciu [Correspondant], Daria La Rocca.

The scale-free concept formalizes the intuition that, in many systems, the analysis of temporal dynamics cannot be grounded on specific and characteristic time scales. The scale-free paradigm has permitted the relevant analysis of numerous applications, very different in nature, ranging from natural phenomena (hydrodynamic turbulence, geophysics, body rhythms, brain activity,...) to human activities (Internet traffic, population, finance, art,...).

Yet, most successes of scale-free analysis were obtained in contexts where data are univariate, homogeneous along time (a single stationary time series), and well-characterized by simple-shape local singularities. For such situations, scale-free dynamics translate into global or local power laws, which significantly eases practical analyses. Numerous recent real-world applications (macroscopic spontaneous brain dynamics, the central application in this project, being one paradigm example), however, naturally entail large multivariate data (many signals), whose properties vary along time (non-stationarity) and across components (non-homogeneity), with potentially complex temporal dynamics, thus intricate local singular behaviors.

These three issues call into question the intuitive and founding identification of scale-free to power laws, and thus make uneasy multivariate scale-free and multifractal analyses, precluding the use of univariate methodologies. This explains why the concept of scale-free dynamics is barely used and with limited successes in such settings and highlights the overriding need for a systematic methodological study of multivariate scale-free and multifractal dynamics. The Core Theme of MULTIFRACS consists in laying the theoretical foundations of a practical robust statistical signal processing framework for multivariate non homogeneous scale-free and multifractal analyses, suited to varied types of rich singularities, as well as in performing accurate analyses of scale-free dynamics in spontaneous and task-related macroscopic brain activity, to assess their natures, functional roles and relevance, and their relations to behavioral performance in a timing estimation task using multimodal functional imaging techniques.

This overarching objective is organized into 4 Challenges:

1. Multivariate scale-free and multifractal analysis,
2. Second generation of local singularity indices,
3. Scale-free dynamics, non-stationarity and non-homogeneity,
4. Multivariate scale-free temporal dynamics analysis in macroscopic brain activity.

8.2.1.2. BrainPedia project

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux.

BrainPedia is an ANR JCJC (2011-2015) which addresses the following question: Neuroimaging produces huge amounts of complex data that are used to better understand the relations between brain structure and function. While the acquisition and analysis of this data is getting standardized in some aspects, the neuroimaging community is still largely missing appropriate tools to store and organize the knowledge related to the data. Taking advantage of common coordinate systems to represent the results of group studies, coordinate-based meta-analysis approaches associated with repositories of neuroimaging publications provide a crude solution to this problem, that does not yield reliable outputs and looses most of the data-related information. In this project, we propose to tackle the problem in a statistically rigorous framework, thus providing usable information to drive neuroscientific knowledge and questions.

8.2.1.3. Niconnect project

Participants: Bertrand Thirion, Gaël Varoquaux [Correspondant], Alexandre Abraham, Kamalaker Reddy Dadi, Darya Chyzhyk, Mehdi Rahim.

- **Context:** The NiConnect project (2012-2016) arises from an increasing need of medical imaging tools to diagnose efficiently brain pathologies, such as neuro-degenerative and psychiatric diseases or lesions related to stroke. Brain imaging provides a non-invasive and widespread probe of various features of brain organization, that are then used to make an accurate diagnosis, assess brain rehabilitation, or make a prognostic on the chance of recovery of a patient. Among different measures extracted from brain imaging, functional connectivity is particularly attractive, as it readily probes the integrity of brain networks, considered as providing the most complete view on brain functional organization.

- **Challenges:** To turn methods research into popular tool widely usable by non specialists, the NiConnect project puts specific emphasis on producing high-quality open-source software. NiConnect addresses the many data analysis tasks that extract relevant information from resting-state fMRI datasets. Specifically, the scientific difficulties are *i)* conducting proper validation of the models and tools, and *ii)* providing statistically controlled information to neuroscientists or medical doctors. More importantly, these procedures should be robust enough to perform analysis on limited quality data, as acquiring data on diseased populations is challenging and artifacts can hardly be controlled in clinical settings.

- **Outcome of the project:** In the scope of computer science and statistics, NiConnect pushes forward algorithms and statistical models for brain functional connectivity. In particular, we are investigating structured and multi-task graphical models to learn high-dimensional multi-subject brain connectivity models, as well as spatially-informed sparse decompositions for segmenting structures from brain imaging. With regards to neuroimaging methods development, NiConnect provides systematic comparisons and evaluations of connectivity biomarkers and a software library embedding best-performing state-of-the-art approaches. Finally, with regards to medical applications, the NiConnect project also plays a support role in on going medical studies and clinical trials on neurodegenerative diseases.

- **Consortium**
  - Parietal Inria research team: applied mathematics and computer science to model the brain from MRI

- LIF INSERM research team: medical image data analysis and modeling for clinical applications
- CATI center: medical image processing center for large scale brain imaging studies
- Henri-Mondor hospital neurosurgery and neuroradiology: clinical teams conducting research on treatments for neurodegenerative diseases, in particular Huntington and Parkinson diseases
- Logilab: consulting in scientific computing

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

8.3.1.1. HBP

Title: The Human Brain Project

Programm: FP7

Duration: October 2013 - September 2016

Coordinator: EPFL

Partners: 100 across Europe

Inria contact: Olivier Faugeras

Understanding the human brain is one of the greatest challenges facing 21st century science. If we can rise to the challenge, we can gain profound insights into what makes us human, develop new treatments for brain diseases and build revolutionary new computing technologies. Today, for the first time, modern ICT has brought these goals within sight. The goal of the Human Brain Project, part of the FET Flagship Programme, is to translate this vision into reality, using ICT as a catalyst for a global collaborative effort to understand the human brain and its diseases and ultimately to emulate its computational capabilities. The Human Brain Project will last ten years and will consist of a ramp-up phase (from month 1 to month 36) and subsequent operational phases. This Grant Agreement covers the ramp-up phase. During this phase the strategic goals of the project will be to design, develop and deploy the first versions of six ICT platforms dedicated to Neuroinformatics, Brain Simulation, High Performance Computing, Medical Informatics, Neuromorphic Computing and Neurorobotics, and create a user community of research groups from within and outside the HBP, set up a European Institute for Theoretical Neuroscience, complete a set of pilot projects providing a first demonstration of the scientific value of the platforms and the Institute, develop the scientific and technological capabilities required by future versions of the platforms, implement a policy of Responsible Innovation, and a programme of transdisciplinary education, and develop a framework for collaboration that links the partners under strong scientific leadership and professional project management, providing a coherent European approach and ensuring effective alignment of regional, national and European research and programmes. The project work plan is organized in the form of thirteen subprojects, each dedicated to a specific area of activity. A significant part of the budget will be used for competitive calls to complement the collective skills of the Consortium with additional expertise.

8.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: Marie Curie

Project acronym: Neuroimaging Power

Project title: Effect size and power for neuroimaging.

Duration: mois année début - mois année fin

Coordinator: Inria

Other partners: Univ. Stanford, USA

Abstract: There is an increasing concern about statistical power in neuroscience research. Critically, an underpowered study has poor predictive power. Findings from a low-power study are unlikely to be reproducible, and thus a power analysis is a critical component of any paper. This project aims to promote and facilitate the use of power analyses.A key component of a power analysis is the specification of an effect size. However, in neuroimaging, there is no standardised way to communicate effect sizes, which makes the choice of an appropriate effect size a formidable task. The best way today to perform a power analysis is by collecting a pilot data set, a very expensive practice. To eliminate the need for pilot data, we will develop a standardised measure of effect size taking into account the spatial variance and the uncertainty of the measurements. Communicating effect sizes in new publications will facilitate the use of power analyses.To further alleviate the need for pilot data, we will provide a library of effect sizes for different tasks and contrasts, using open data projects in neuroimaging. We will integrate our effect size estimator in open repositories NeuroVault and OpenfMRI. Consequently, these effect sizes can then serve as a proxy for a pilot study, and as such, a huge cost in the design of an experiment is eliminated.A new experiment will not be identical to the open data and as such the hypothesised parameters might not be fully accurate. To address this issue, we present a flexible framework to analyse data mid-way without harming the control of the type I error rate. Such a procedure will allow re-evaluating halfway an experiment whether it is useful to continue a study, and how many more subjects are needed for statistically sound inferences.To make our methods maximally available, we will write a software suite including all these methods in different programming platforms and we will provide a GUI to further increase the use of power analyses.

8.4. International Initiatives

8.4.1. MetaMRI

Title: Machine learning for meta-analysis of functional neuroimaging data

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Department of Psychology - Russ Poldrack

Start year: 2015

See also: https://team.inria.fr/metamri

Neuroimaging produces huge amounts of complex data that are used to better understand the relations between brain structure and function. Observing that the neuroimaging community is still largely missing appropriate tools to store and organize the knowledge related to the data, Parietal team and Poldrack's lab, have decided to join forces to set up a framework for functional brain image meta-analysis, i.e. a framework in which several datasets can be jointly analyzed in order to accumulate information on the functional specialization of brain regions. MetaMRI will build upon Poldrack's lab expertise in handling, sharing and analyzing multi-protocol data and Parietal's recent developments of machine learning libraries to develop a new generation of meta-analytic tools.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- **Bertrand Thirion**: Organization for Human Brain Mapping.

9.1.1.2. Reviewer

- **Philippe Ciuciu**: IEEE ISBI (15 papers), IEEE ICASSP (10 papers), IEEE ICIP (5 papers), NIPS (4 papers), EUSIPCO (5 papers).
- **Bertrand Thirion**: IPMI, MICCAI, NIPS, ISBI, PRNI, AISTATS
- **Gaël Varoquaux**: IEEE ICASSP, MICCAI, NIPS, IPMI, ICML

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- **Bertrand Thirion**: Medical Image Analysis, Frontiers in brain imaging
- **Gaël Varoquaux**: Frontiers in NeuroInformatics, Frontiers in brain imaging methods, NeuroImage

9.1.2.2. Reviewer - Reviewing Activities

- **Philippe Ciuciu**: Reviewer for Neuroimage, IEEE Signal Processing Letters, Signal Processing, IEEE Trans. Medical Imaging, Plos One, Plos Comput. Biology, Frontiers in Neuroscience.
- **Bertrand Thirion**: Human Brain Mapping, IEEE TMI, MedIA, NeuroImage, PNAS
- **Gaël Varoquaux**: NeuroImage, JSTSP, PNAS, HBM, PLOS Comp Bio, Gigascience
- **Olivier Grisel**: Journal of Machine Learning Research (software track).

9.1.3. Invited Talks

9.1.3.1. Bertrand Thirion

- February: invited talk at the *Imagerie du Vivant* National congress, entitled *Large-scale analyses in functional brain Imaging*.
- February: presentation at the Pasadena working group of the Digicosme Labex.
- April: invited presentation at European Neuroscience institute, Paris, entitled *Seeing it all: Convolutional network layers map the function of the human visual system*.
- April: presentation Functional connectomicts, at DTU Copenhagen, entitled *from large-scale estimators to empirical validation*.
- May: Talk at Atlas workshop, Grenoble, entitled *Learning representations from functional brain imaging*.
- June: organizer of a table ronde at the *Futur en Seine* event entitled *Computational methods for neurosciences & medical imaging*.
- October: talk at MPI Psychiatry, Munich, entitled *Machine learning for neuroimaging: current challenges and solutions*.
- June: Talk at Neurostic workshop, Grenoble, entitled *Learning representations from functional brain imaging*.
- October: Invited talk by the ITMO Neuroscience, Bordeaux, entitled *Working with large data samples: the case of human brain imaging*.

9.1.3.2. Philippe Ciuciu

- 12/16: IEEE Lecture at University of British Columbia (Vancouver, Canada): *Sparkling: Novel non-Cartesian sampling schemes for accelerated 2D anatomical imaging at 7 Tesla*.
- 12/16: Pacific Parkinson's research center (Vancouver, Canada): *Impact of perceptual learning on resting-state brain dynamics in fMRI: A supervised classification study*.
- 09/16: GdR d'Analyse Multifractale (Avignon, France): *Convergence of neural activity to multifractal attractors in MEG predicts learning*.
- 08/16: invitation to the Special session entitled "Unraveling brain networks from functional neuroimaging data" at EUSIPCO'16 (Budapest, Hungary): *Impact of perceptual learning on resting-state fMRI connectivity: A supervised classification study*.
- 06/16: Journées scientifiques d'Inria (Rennes, France): *Compressive Sampling in MRI*.

- 06/16: Inria Sophia-Antipolis, équipe Athena. *New physically plausible compressive sampling schemes for MRI: First results at 7 Tesla*

- 05/16: University of Geneva (Campus BioTech, Geneva, Switzlerand): *Convergence to asymptotic Multifractal dynamics in the brain predicts learning*.

- 02/16: Grenoble Institut of Neurosciences (Grenoble, France): *Physically plausible trajectories for Compressed Sensing in MRI*.

- 02/16: Workshop on 7 Tesla scanner at NeuroSpin (Gif-sur-Yvette, France) *Compressed sensing for high resolution MRI at 7 Tesla*.

- 01/16: Cosmostat lab, IRFU/CEA. *On the generation of compressed sampling schemes in MRI*.

9.1.3.3. Loïc Estève

- EuroScipy 2016: scikit-learn tutorial

- Budapest BI 2016 : scikit-learn tutorial and talk "Recent developments in scikit-learn and joblib"

9.1.3.4. Olivier Grisel

- PyData Berlin and PyData Paris 2016: *"Predictive modeling with Python, trends and tools*

- invited talk on *Some recent developments in Deep Learning researc* at Strata London 2016.

9.1.3.5. Gaël Varoquaux

- Paris Open Source summit 2016: scikit-learn, the vision and the community

- EuroScipy 2016 (Erlangen): keynote: "On writing code the science"

- Open Data Science Conference 2017 (London): keynote: "The code of data science"

- EuroPython 2016 (Bilbao): keynote "Scientists meet web dev: how Python became the language of data"

- PiterPy 2016 (St Petersbourg): keynote: "Python for data"

- Facebook AI Research: some statistical learning problems in brain imaging

- GDR ISIS Imagerie medicale: prediction de pathologies psychiatriques à partir d'imagerie fonctionnelle de repos

- Brain network analysis workshop, MICCAI 2016 (Athenes): keynote

- Journée Graphes et neuroscience à Marseilles: Machine learning on brain graphes

- Séminaire débat sur le Big data en Neuroscience, Lyon

- Seminar Max Planck Institute Leipzig: data mining for neuroimaging

- Seminar Telecom ParisTech: randomized methods for high-dimensional statistical learning

- Séminaire d'équipe Asclepios: Quelques problèmes d'apprentissage sur des images cérébrales

9.1.4. Leadership within the Scientific Community

- Gaël Varoquaux: Chair of the steering committee, IEEE PRNI

- Bertrand Thirion: member of the *Committee on Best Practices in Data Analysis and Sharing* for the OHBM community.

9.1.5. Scientific Expertise

- Philippe Ciuciu: ANR JC, NSERC au Canada, FWO

- Bertrand Thirion: ANR, NWO, NSF

- Gaël Varoquaux: Membre de la Commission Expertises Scientifiques, (CE23) ANR

- Olivier Grisel did 3 days of consulting with the CTO of the Therapixel startup to share expertise on the use of Deep Learning for the predictive analysis of 3D imaging data.

9.1.6. Research Administration

9.1.6.1. Philippe Ciuciu

- 03/16: Involvement in the CEA visiting committee on High Performance Computing.
- 05/16: Member of a Comité de sélection for hiring an Assistant Professor in Paris-Saclay University (Section 61 of CNU).
- 06/16: Member of the Inria scientific commission in charge of ranking PhD and post-doctoral applicants as well as delegations of Assistant Professors to Inria.

9.1.6.2. Bertrand Thirion

- Leader of the Datasense axis of the Digicosme Labex
- Member of the STIC department committee Paris-Saclay University and of the bureau thereof.
- DSA Saclay.

9.1.6.3. Gaël Varoquaux

- Member of "Comité de suivi doctoral", Inria Saclay
- Member of "Comité cluster", Inria Saclay
- Member of "Commission de Développement Technologique", Inria Saclay
- Member of the directorate of the Paris-Saclay CDS (Center for Data Science)

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

9.2.1.1. Bertrand Thirion

Master : Brain Computer interface and Functional Neuroimaging, 12 heures équivalent TD, niveau M2, ENS Cachan

9.2.1.2. Philippe Ciuciu

Master 2 : "Functional MRI: From data acquisition to analysis", 3h, Univ. Paris V René Descartes & Télécom-Paristech, Master of Biomedical Engineering

Master 2 : "FMRI data analysis", 3h, Univ. Paris-Saclay, Master of medical Physics

9.2.1.3. Gaël Varoquaux

Master 2 : "Brain functional connectivity analysis", 7h, Univ. Paris V René Descartes & Télécom-Paristech, Master of Biomedical Engineering

Master 2 : "Machine learning with scikit-learn", 2h, ENSAE

Master 2 : "Advanced Machine learning with scikit-learn", 3h, Centrale Paris, MSc in data sciences & business analytics

Ecole d'été multidisciplinaire analyse de données, Rennes, 1h

OHBM 2016: course on machine learning for cognitive neuroimaging 30mn

PRNI 2016: nilearn for machine learning on brain images, 8h

Max Planck Institute Leipzig: nilearn for machine learning on brain images, 8h

9.2.2. Supervision

9.2.2.1. Bertrand Thirion

PhD in progress: Elvis Dohmatob,

PhD in progress: Arthur Mensch,

PhD in progress: Andrés Hoyos Idrobo

9.2.2.2. Philippe Ciuciu

PhD defended: Aina Frau-Pascual, "Statistical models for the analysis of BOLD and ASL Magnetic Resonance modalities to study brain function and disease", University of Grenoble-Alpes (doctoral school: Mathématiques, Sciences et Technologies de l'Information, Informatique), defense: 19/12/2016, Advisors: Florence Forbes (Dir), Philippe Ciuciu (Co-Dir)

PhD in progress: Carole Lazarus, "Physically plausible compressed sensing for high resolution MRI at 7 Tesla in Humans" starting date: October 2015 (Univ. Paris-Saclay, doctoral school: EOBE). Advisors: Philippe Ciuciu (Dir), Alexandre Vignaud (Co-Dir)

PhD in progress: Loubna El Gueddari, "Parallel proximal algorithms for compressed sensing MRI reconstruction. Applications in ultra-high magnetic field imaging", starting date: October 2016 (Univ. Paris-Saclay, doctoral school: EOBE). Advisors: Philippe Ciuciu (Dir) and Jean-Christophe Pesquet (Co-Dir, Prof. at Centrale-Supélec)

9.2.2.3. Gael Varoquaux

PhD defended: Alexandre Abraham

PhD in progress: Elvis Dohmatob,

PhD in progress: Arthur Mensch,

PhD in progress: Andrés Hoyos Idrobo

9.2.3. Juries

9.2.3.1. Bertrand Thirion

- 04/29: Reviewer of Niklas Kasenburg PhD Thesis , Univ. Copenhagen, Denmark.
- 01/12: Examiner of Simona Schiavi PhD Thesis, Univ. Paris Saclay.
- 14/12: Reviewer of Olivier Marre habilitation, Paris.
- 15/12: Reviewer of Maite Termenon PhD thesis, Univ. Grenoble.

9.2.3.2. Philippe Ciuciu

- 04/16: Reviewer of Aiping Liu's PhD thesis (ECCS Dpt, Univ. British Columbia, Vancouvern Canada) entitled *"Brain Connectivity Network Modeling using fMRI signals"*
- 05/16: Reviewer of Andrea Laruelo-Fernandez's PhD thesis (INP Toulouse-IRIT- ENSEEIHT) entitled *"Integration of magnetic resonance spectroscopic imaging into the radiotherapy treatment planning"*
- 09/16: Examinor of Mohanad Albughdadi's PhD thesis (INP Toulouse-IRIT- ENSEEIHT) entitled *" Bayesian joint detection-estimation in functional MRI with automatic parcellation and functional constraints"*
- 10/16: Reviewer of Sébastien Combrexelle's PhD thesis (INP Toulouse-IRIT- ENSEEIHT) entitled *"Multifractal analysis for multivariate data with application to remote sensing"*.
- 10/16: Co-director of Aina Frau-Pascual's PhD thesis (see above).

9.2.3.3. Gaël Varoquaux

- 06/16: Examiner of Alberto García Durán, PhD Thesis, UTC Compiegne.

9.3. Popularization

9.3.1. Gaël Varoquaux

- Unithé ou Café, Inria Saclay Ile de France
- Atelier IHEST Les mots du numérique - 17 novembre

9.3.2. Loïc Estève

Software Carpentry workshops:

- git course at UNIC in Gif-sur-Yvette March 29-30
- helper at "Scientific Programming with Python and Software Engineering Best Practices" workshop, April 28-29 at Télécom Paris
- git course at Proto 204, May 24-25

Mentor at Startup Weekend Artificial Intelligence, November 4-6.

9.3.3. Olivier Grisel

"La tête au carré" radio show on France Inter in January 2016 to share his expertise and opinion on the use and impacts of Big Data and predictive algorithms [1].

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journals

[1] A. ABRAHAM, M. MILHAM, A. DI MARTINO, R. C. CRADDOCK, D. SAMARAS, B. THIRION, G. VAROQUAUX. *Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example*, in "NeuroImage", November 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.10.045], https://hal.inria.fr/hal-01398867

[2] M. ALBUGHDADI, L. CHAARI, J.-Y. TOURNERET, F. FORBES, P. CIUCIU. *A Bayesian Non-Parametric Hidden Markov Random Model for Hemodynamic Brain Parcellation*, in "Signal Processing", 2017 [*DOI :* 10.1016/J.SIGPRO.2017.01.005], https://hal.archives-ouvertes.fr/hal-01426385

[3] V. BORGHESANI, F. PEDREGOSA, M. BUIATTI, A. AMADON, E. EGER, M. PIAZZA. *Word meaning in the ventral visual path: a perceptual to conceptual gradient of semantic coding*, in "NeuroImage", September 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.08.068], http://hal.upmc.fr/hal-01372551

[4] C. BOYER, N. CHAUFFERT, P. CIUCIU, J. KAHN, P. WEISS. *On the generation of sampling schemes for Magnetic Resonance Imaging*, in "SIAM Journal on Imaging Sciences", December 2016, vol. 9, n⁰ 4, pp. 2039-20972, https://hal.archives-ouvertes.fr/hal-01373758

[5] D. BZDOK, G. HARTWIGSEN, A. REID, A. R. LAIRD, P. T. FOX, S. EICKHOFF. *Left inferior parietal lobe engagement in social cognition and language*, in "Neuroscience and Biobehavioral Reviews", April 2016 [*DOI :* 10.1016/J.NEUBIOREV.2016.02.024], https://hal.archives-ouvertes.fr/hal-01350512

[6] D. BZDOK, G. VAROQUAUX, O. GRISEL, M. EICKENBERG, C. POUPON, B. THIRION. *Formal Models of the Network Co-occurrence Underlying Mental Operations*, in "PLoS Computational Biology", June 2016 [*DOI :* 10.1371/JOURNAL.PCBI.1004994], https://hal.archives-ouvertes.fr/hal-01338307

[7] D. BZDOK, G. VAROQUAUX, B. THIRION. *Neuroimaging Research: From Null-Hypothesis Falsification to Out-of-sample Generalization*, in "Educational and Psychological Measurement", August 2016, https://hal.archives-ouvertes.fr/hal-01338313

[8] N. CHAUFFERT, P. CIUCIU, J. KAHN, P. WEISS. *A Projection Method on Measures Sets*, in "Constructive Approximation", February 2017, vol. 45, n⁰ 1, pp. 83 - 111 [*DOI :* 10.1007/S00365-016-9346-2], https://hal.inria.fr/hal-01432720

[9] N. CHAUFFERT, P. WEISS, J. KAHN, P. CIUCIU. *A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging*, in "IEEE Transactions on Medical Imaging", 2016 [*DOI :* 10.1109/TMI.2016.2544251], https://hal.inria.fr/hal-01317939

---

[1] https://www.franceinter.fr/emissions/la-tete-au-carre/la-tete-au-carre-05-janvier-2016

[10] M. Eickenberg, A. Gramfort, G. Varoquaux, B. Thirion. *Seeing it all: Convolutional network layers map the function of the human visual system*, in "NeuroImage", 2016 [*DOI : 10.1016/j.neuroimage.2016.10.001*], https://hal.inria.fr/hal-01389809

[11] S. Eickhoff, B. Thirion, G. Varoquaux, D. Bzdok. *Connectivity-Based Parcellation: Critique and Implications*, in "Human Brain Mapping", January 2016, 22 p. [*DOI : 10.1002/hbm.22933*], https://hal.archives-ouvertes.fr/hal-01184563

[12] K. J. Gorgolewski, T. Auer, V. D. Calhoun, C. R. Craddock, S. Das, E. P. Duff, G. Flandin, S. S. Ghosh, T. Glatard, Y. O. Halchenko, D. A. Handwerker, M. Hanke, D. Keator, X. Li, Z. Michael, C. Maumet, N. B. Nichols, T. E. Nichols, J. Pellman, J.-B. Poline, A. Rokem, G. Schaefer, V. Sochat, W. Triplett, J. A. Turner, G. Varoquaux, R. A. Poldrack. *The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments*, in "Scientific Data ", June 2016, vol. 3, 160044 [*DOI : 10.1038/sdata.2016.44*], http://www.hal.inserm.fr/inserm-01345616

[13] A. Hoyos-Idrobo, G. Varoquaux, J. Kahn, B. Thirion. *Recursive nearest agglomeration (ReNA): fast clustering for approximation of structured signals*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", August 2016, https://hal.inria.fr/hal-01366651

[14] F. Liem, G. Varoquaux, J. Kynast, F. Beyer, S. K. Masouleh, J. M. Huntenburg, L. Lampe, M. Rahim, A. Abraham, R. C. Craddock, S. Riedel-Heller, T. Luck, M. Loeffler, M. L. Schroeter, A. V. Witte, D. S. Margulies. *Predicting brain-age from multimodal imaging data captures cognitive impairment*, in "NeuroImage", November 2016 [*DOI : 10.1016/j.neuroimage.2016.11.005*], https://hal.inria.fr/hal-01403005

[15] M. Rahim, B. Thirion, C. Comtat, G. Varoquaux. *Transmodal Learning of Functional Networks for Alzheimer's Disease Prediction*, in "IEEE Journal of Selected Topics in Signal Processing", October 2016, vol. 10, n° 7, pp. 1204 - 1213 [*DOI : 10.1109/JSTSP.2016.2600400*], https://hal.inria.fr/hal-01353728

[16] G. Varoquaux, P. A. Reddy Raamana, D. A. Engemann, A. A. Hoyos-Idrobo, Y. A. Schwartz, B. A. Thirion. *Assessing and tuning brain decoders: cross-validation, caveats, and guidelines*, in "NeuroImage", 2016 [*DOI : 10.1016/j.neuroimage.2016.10.038*], https://hal.archives-ouvertes.fr/hal-01332785

Invited Conferences

[17] E. Dohmatob. *A simple algorithm for computing Nash-equilibria in incomplete information games*, in "OPT2016 – NIPS workshop on optimization for machine learning", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01384460

[18] E. Dohmatob, M. Eickenberg, B. Thirion, G. Varoquaux. *Local Q-Linear Convergence and Finite-time Active Set Identification of ADMM on a Class of Penalized Regression Problems*, in "ICASSP, International Conference on Acoustics, Speach, and Signal Processing", China, March 2016, http://icassp2016.org/default.asp , https://hal.archives-ouvertes.fr/hal-01265372

[19] E. Dohmatob, A. Mensch, G. Varoquaux, B. Thirion. *Learning brain regions via large-scale online structured sparse dictionary-learning*, in "Neural Information Processing Systems (NIPS)", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01369134

International Conferences with Proceedings

[20] M. ALBUGHDADI, L. CHAARI, F. FORBES, J.-Y. TOURNERET, P. CIUCIU. *Multi-subject joint parcellation detection estimation in functional MRI*, in "13th IEEE International Symposium on Biomedical Imaging", Prague, Czech Republic, April 2016, https://hal.inria.fr/hal-01261982

[21] E. BELILOVSKY, G. VAROQUAUX, M. B. BLASCHKO. *Testing for Differences in Gaussian Graphical Models: Applications to Brain Connectivity*, in "Neural Information Processing Systems (NIPS) 2016", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01248844

[22] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX. *Dictionary Learning for Massive Matrix Factorization*, in "International Conference on Machine Learning", New York, United States, Proceedings of the 33rd Internation Conference on Machine Learning, June 2016, vol. 48, pp. 1737–1746, https://hal.archives-ouvertes.fr/hal-01308934

[23] H. PELLÉ, P. CIUCIU, M. RAHIM, E. DOHMATOB, P. ABRY, V. VAN WASSENHOVE. *Multivariate Hurst Exponent Estimation in FMRI. Application to Brain Decoding of Perceptual Learning*, in "13th IEEE International Symposium on Biomedical Imaging", Prague, Czech Republic, April 2016, https://hal.inria.fr/hal-01261976

[24] M. RAHIM, P. CIUCIU, S. BOUGACHA. *Impact of perceptual learning on resting-state fMRI connectivity: A supervised classification study*, in "Eusipco 2016", Budapest, Hungary, August 2016, https://hal.inria.fr/hal-01355478

Conferences without Proceedings

[25] Y. BEKHTI, D. STROHMEIER, M. JAS, R. BADEAU, A. GRAMFORT. *M/EEG source localization with multi-scale time-frequency dictionaries*, in "6th International Workshop on Pattern Recognition in Neuroimaging (PRNI)", Trento, Italy, June 2016 [*DOI : 10.1109/PRNI.2016.7552337*], https://hal.archives-ouvertes.fr/hal-01313567

[26] K. DADI, A. ABRAHAM, M. RAHIM, B. THIRION, G. VAROQUAUX. *Comparing functional connectivity based predictive models across datasets*, in "PRNI 2016: 6th International Workshop on Pattern Recognition in Neuroimaging", Trento, Italy, June 2016, https://hal.inria.fr/hal-01319131

[27] A. HOYOS-IDROBO, G. VAROQUAUX, B. THIRION. *Fast brain decoding with random sampling and random projections*, in "PRNI 2016: the 6th International Workshop on Pattern Recognition in Neuroimaging", Trento, Italy, June 2016, https://hal.inria.fr/hal-01313814

[28] C. LAROCHE, H. PAPADOPOULOS, M. KOWALSKI, G. RICHARD. *Genre specific dictionaries for harmonic/percussive source separation*, in "ISMIR The 17th International Society for Music Information Retrieval Conference", New York, United States, August 2016, https://hal.archives-ouvertes.fr/hal-01353252

[29] C. LAROCHE, H. PAPADOPOULOS, M. KOWALSKI, G. RICHARD. *Drum extraction in single channel audio signals using multi-layer non negative matrix factor deconvolution*, in "ICASSP", Nouvelle Orleans, United States, March 2017, https://hal.archives-ouvertes.fr/hal-01438851

[30] A. MENSCH, J. MAIRAL, G. VAROQUAUX, B. THIRION. *Subsampled online matrix factorization with convergence guarantees*, in "NIPS Workshop on Optimization for Machine Learning", Barcelone, Spain, December 2016, https://hal.archives-ouvertes.fr/hal-01405058

[31] A. MENSCH, G. VAROQUAUX, B. THIRION. *Compressed Online Dictionary Learning for Fast Resting-State fMRI Decomposition*, in "International Symposium on Biomedical Imaging (ISBI 2016) "From Nano to Macro"", Prague, Czech Republic, 13th International Symposium on Biomedical Imaging (ISBI), IEEE, April 2016, pp. 1282-1285 [*DOI :* 10.1109/ISBI.2016.7493501], https://hal.archives-ouvertes.fr/hal-01271033

[32] G. VAROQUAUX, M. KOWALSKI, B. THIRION. *Social-sparsity brain decoders: faster spatial sparsity*, in "Pattern Recognition in NeuroImaging", Trento, Italy, June 2016, https://hal.archives-ouvertes.fr/hal-01334551

Scientific Books (or Scientific Book chapters)

[33] B. THIRION. *Functional Neuroimaging Group Studies*, in "Handbook of Neuroimaging Data Analysis", H. OMBAO, M. LINDQUIST, W. THOMPSON, J. ASTON (editors), Handbook of Modern Statistical methods, Chapman & Hall / CRC, November 2016, https://hal.inria.fr/hal-01419347

Other Publications

[34] M. ALBUGHDADI, L. CHAARI, J.-Y. TOURNERET, F. FORBES, P. CIUCIU. *Hemodynamic Brain Parcellation Using A Non-Parametric Bayesian Approach*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01275622

[35] E. BELILOVSKY, K. KASTNER, G. VAROQUAUX, M. BLASCHKO. *Learning to Discover Graphical Model Structures*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01306491

[36] D. BZDOK, B. T. YEO. *The Future of Data Analysis in the Neurosciences*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01356923

[37] L. CHAARI, S. BADILLO, T. VINCENT, G. DEHAENE-LAMBERTZ, F. FORBES, P. CIUCIU. *Subject-level Joint Parcellation-Detection-Estimation in fMRI*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01255465

[38] F. FENG, M. KOWALSKI. *A Unified Approach for Over and Under-Determined Blind Source Separation Based on Both Sparsity and Decorrelation*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01297471

[39] F. HADJ-SELEM, T. LÖFSTEDT, E. DOHMATOB, V. FROUIN, M. DUBOIS, V. GUILLEMOT, E. DUCHESNAY. *Supplementary materials: Iterative Smoothing Proximal Gradient for Regression with Structured Sparsity*, May 2016, working paper or preprint, https://hal-cea.archives-ouvertes.fr/cea-01324021

[40] M. KOWALSKI, A. MEYNARD, H.-T. WU. *Convex Optimization approach to signals with fast varying instantaneous frequency*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01199615

[41] C. LAZARUS, N. CHAUFFERT, J. KAHN, P. WEISS, A. VIGNAUD, P. CIUCIU. *Physically plausible K-space trajectories for Compressed Sensing in MRI: From simulations to real acquisitions*, March 2016, CEA Visiting committee on High Performance Computing, Poster, https://hal.inria.fr/hal-01296496

[42] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX. *Stochastic Subsampling for Factorizing Huge Matrices*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01431618

[43] M. RAHIM, P. CIUCIU, S. BOUGACHA. *Functional connectivity outperforms scale-free brain dynamics as fMRI predictive feature of perceptual learning underwent in MEG*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01297845