Decoding Brain States from fMRI Connectivity Graphs

Jonas Richiardi[a,b], Hamdi Eryilmaz[c], Sophie Schwartz[c], Patrik Vuilleumier[c], Dimitri Van De Ville[a,b]

[a] *Medical Image Processing Lab, Ecole Polytechnique Fédérale de Lausanne, Switzerland*
[b] *Medical Image Processing Lab, University of Geneva, Switzerland*
[c] *Laboratory of Neurology and Imaging of Cognition, University of Geneva, Switzerland*

Abstract

Functional connectivity analysis of fMRI data can reveal synchronised activity between anatomically distinct brain regions. Here, we extract the characteristic connectivity signatures of different brain states to perform classification, allowing us to decode the different states based on the functional connectivity patterns. Our approach is based on polythetic decision trees, which combine powerful discriminative ability with interpretability of results. We also propose to use ensemble of classifiers within specific frequency subbands, and show that they bring systematic improvement in classification accuracy. Exploiting multi-band classification of connectivity graphs is also proposed, and we explain theoretical reasons why the technique could bring further improvement in classification performance. The choice of decision trees as classifier is shown to provide a practical way to identify a subset of connections that distinguishes best between the conditions, permitting the extraction of very compact representations for differences between brain states, which we call discriminative graphs. Our experimental results based on strict train/test separation at all stages of processing show that the method is applicable to inter-subject brain decoding with relatively low error rates for the task considered.

Keywords: fMRI, brain decoding, functional connectivity, graphs, decision tree

1. Introduction

Traditional fMRI analysis consists of univariate statistical hypothesis testing to assess changes in the activity of each brain voxel induced by the

stimulation paradigm (Frackowiak et al. (1997)). More recently, approaches derived from supervised machine learning—commonly termed "brain decoding" in the field of neuroimaging—have shown that it is possible to exploit more subtle relationships in voxels' intensity patterns (Haxby et al. (2001); Haynes and Rees (2006); Norman et al. (2006)). These methods rely on a classifier to predict the subject's brain state from the BOLD responses in a set of selected voxels, such as visual (Cox and Savoy (2003); Haynes and Rees (2005); Kamitani and Tong (2005); Thirion et al. (2006); Kay et al. (2008); Miyawaki et al. (2008)) or auditory cortices (Ethofer et al. (2009)). Results from brain decoding are often remarkable since they clearly reach beyond the possibilities of univariate techniques, but also because they are able to uncover information from fine-grained cortical activity despite the relatively low spatial resolution of fMRI. Furthermore, instead of a fixed spatial window, one can also apply classification to a so-called "searchlight" that slides over the whole brain data, such that the classification success for each position of the spotlight can then be mapped to show brain regions that carry discriminative information between different conditions (Kriegeskorte et al. (2006)). Another interesting approach is to use spatiotemporal observations as an input to the classifier (Mitchell et al. (2004); Mourao-Miranda et al. (2007)).

The study of functional connectivity is concerned with the temporal coherence between neurophysiological events observed in spatially remote brain regions. In early work, correlation with a seed voxel was investigated and revealed bilateral coactivation between sensory cortices (Biswal et al. (1995); Lowe et al. (1998)). Further advances have been driven by unsupervised methods such as source separation—mainly principal components (Friston et al. (1993)) and independent components analysis (McKeown et al. (1998); Calhoun et al. (2002); Beckmann and Smith (2004)) which allow to identify large-scale cortical networks—but also by other methods such as dynamic causal modelling, which tries to establish effective connectivity and requires prior information about the neurological network to investigate (Friston et al. (2003)). A recent method related to our work proposed to use resting-state correlations between regions of interest as features for an SVM classifier (Craddock et al., 2009). Another attractive methodology to investigate functional networks is to rely on mathematical graph theory; i.e., constructing the (undirected) graph from temporal correlation matrices and computing related measures; e.g., node degree, hubbiness, and so on (Sporns et al. (2000); Salvador et al. (2005)). This methodology has brought new

insights in functional connectivity at resting state, such as the small-world organisation of cortical networks at low temporal frequencies (Achard et al. (2006)).

Here, we bring together brain decoding and graph representations based on functional connectivity measures. These measures, such as temporal correlation, are performed over a given period of time and reflect the timecourse resemblances between different regions. Moreover, we estimate connectivity at different temporal scales using the wavelet transform as a preprocessing step. Then, we build a classifier trained on functional connectivity graphs of a group a subjects to distinguish between different brain states of an unseen subject. The aim of our approach is to identify the connections that are most discriminative between brain states, and to obtain relevant visual representation of the data for neuroscience studies.

Previous brain decoding techniques primarily rely on linear support vector machines (SVMs), which use a soft-margin hyperplane to separate classes. In the present work, we propose polythetic decision trees that fit a hyperplane using the most discriminative features (i.e., connections) at each level, such that potentially complex and non-linear class boundaries can be obtained by multilevel trees. Efficient and effective learning of this type of decision trees relies on recent advances in pattern recognition (Friedman et al. (2000); Gama (2004); Landwehr et al. (2005)), and provides embedded feature selection which yields a compact discriminant function whose parameters are amenable to interpretation. Their variance properties make them good candidates for ensembling, steering the classification strategy towards slightly weaker but simpler (lower capacity) classifiers, which is a desirable behaviour in high-dimensional learning, a classical situation in fMRI where the number of dimensions is much higher than the number of training examples.

Our paper is organised as follows. In the Methods section, we describe our data processing pipeline together with the details of the proposed methodology. Next, we illustrate the feasibility of our approach by a proof of concept; i.e., an fMRI experiment with block-based stimulation paradigm (watching short movies) with long resting periods. We extract the set of connections that is the most discriminative between rest and stimulation at different temporal scales. The method is able to correctly classify the conditions in a leave-one-subject-out cross-validation setting. Interestingly, we find that the low-frequency correlations of the BOLD signal (below 0.11Hz) are the most informative. The discriminative network also confirms a differential modulation of sensory areas, in particular within the visual system, and midline

brain areas during movie and rest conditions.

2. Methods

2.1. Preprocessing and data representation

The preprocessing steps are illustrated schematically in Fig. 1, and explained in detail below.

After realignment of the functional volumes using SPM5[1], we use the IBASPM toolbox (Tzourio-Mazoyer et al. (2002); Alemán-Gómez et al. (2006)) to build an individual brain atlas based on the structural MRI, containing $M = 90$ anatomical regions. While this a relatively coarse atlas, it is an essential step to allow for inter-subject variability and enable inter-subject decoding with a good generalisation ability to unseen subjects - using group-level normalisation and atlasing is not an option in this setting. Furthermore, the structural atlas serves only as a basis for computing a much lower resolution functional atlas. Using a more fine-grained atlas might result in some regions disappearing completely in the functional atlas. Another benefit of using the AAL atlas is that it offer a way of comparing results with several other studies (Zalesky et al. (2010)).

We then obtain spatially-averaged timecourses from the voxels corresponding to these regions in the functional space. For N repetitions (which can be intra- or inter-subject, in our case N is the number of subjects) and C conditions, we obtain the matrix

$$\mathbf{X} : M \times T \times N \times C,$$

that contains $M \cdot N \cdot C$ timecourses of length T. We denote the submatrix $\mathbf{X}_{n,c} : M \times T$ for the M timecourses of subject n and condition c.

The timecourses are then decomposed using the (redundant) discrete wavelet transform (DWT) along the temporal dimension. This results into J matrices $\mathbf{X}^{(i)}$, $i = 1, \ldots, J$ that reflect the regional brain activities at different temporal scales. We use cubic Battle-Lemarié wavelets (Battle, 1987).

2.2. Functional connectivity graphs

We use pairwise Pearson correlation coefficients to form the correlation matrix $\mathbf{R}^{(i)}_{n,c} = \mathbb{E}[\mathbf{X}^{(i)}_{n,c}\mathbf{X}^{(i)\,T}_{n,c}] : M \times M$, where the temporal detrending (up

[1]Available at `http://www.fil.ion.ucl.ac.uk/spm/`.

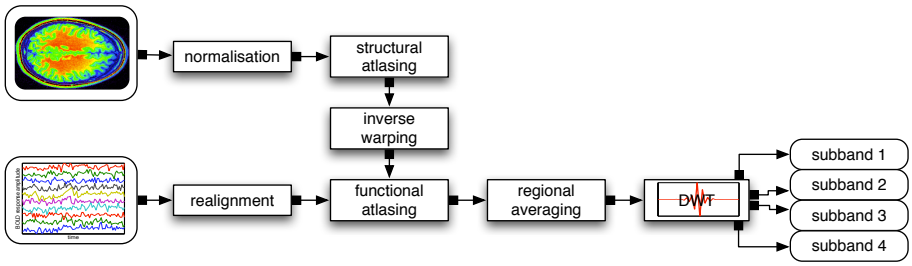


Figure 1: Flowchart of the preprocessing procedure. DWT stands for Discrete Wavelet Transform.

to third-degree polynomials) is provided by the vanishing moments of the wavelet decomposition.

We now consider the brain regions as a set of vertices V and the correlation coefficients as signed weights on the set of edges E, leading to an undirected complete weighted graph $\mathcal{G} = (V, E)$. The graph adjacency matrix $\mathbf{A}^{(i)}_{n,c}$ can be defined as $\mathbf{A}^{(i)}_{n,c} = \mathbf{R}^{(i)}_{n,c} - \mathbf{I}$.

2.3. Graph matching

Comparing graphs is an active field of research in computer vision and pattern recognition where numerous theoretical advances and practical algorithms have emerged recently (Conte et al. (2004)). *Full graph matching* algorithms (e.g., Umeyama (1988)) are commonly used to obtain distances between graphs, which is a basic operation of pattern recognition. Such algorithms operate by trying to find the best correspondence between the vertices of two graphs. They are typically robust (error-tolerant) with respect to changes in graph structure (such as a different number of vertices) (Bunke and Shearer, 1998), because in many fields graphs are computed from real-life data and their topology can fluctuate from instance to instance.

In our case, the graphs obtained have a fixed number of vertices (always the same number of atlas regions), as well as a fixed vertex ordering (regions are spatially defined), because of the atlasing procedure. This type of connectivity graphs form a particular restricted class of graphs, for which several general graph matching algorithms would not yield discriminative information (Richiardi et al. (2010)). For example full graph matching, by looking for the best possible permutation of vertices, would bring high computational cost and is unnecessary for our application where vertex correspondence between graphs fixed and one to one. Indeed, the main use of our graphs is

in computing discriminative information to distinguish between brain states, and this can be done by computing the distance between graphs as a function of the distance between the weights of the respective edges. For this reason, we opt for the strategy of representing the graphs in vector spaces instead of graph domains. This also allows us to use the large variety of existing pattern recognitions methods that operate on vector spaces.

2.4. Graph embedding and feature selection

We propose the following simple embedding procedure. The adjacency matrix $\mathbf{A}'^{(i)}_{n,c}$ is fully characterised by the upper triangular part above the main diagonal. Furthermore, the vertex ordering is constant across graph because of the atlasing procedure, as is the vertex set cardinality. For each graph, we thus generate a feature vector $\mathbf{F} : \binom{M}{2} \times 1$ from the edge weights of all the edges in the upper triangular part of $\mathbf{A}'^{(i)}_{n,c}$ by linearising this part of the matrix. This generates a high-dimensional feature space, and an adequate classifier is required to handle the ensuing high-dimensionality learning problem.

The variability of fMRI connectivity measures is usually high. Therefore, a statistical thresholding technique can be used to clean up the subsequent adjacency matrices. Achard et al. (2006) apply the false discovery rate (FDR) procedure to the adjacency matrices of multiple subjects; the significance of each connection within each experimental condition is assessed by building a t-value and then corrected for multiple comparisons by FDR (Benjamini and Hochberg, 1995). Such statistical edge pruning of the connectivity graph corresponds to feature selection of the embedding vector where pruned edges are mapped to zero. In the current context of comparing graphs for different conditions using a classification approach, applying the FDR procedure and considering the intersection is equivalent to a univariate feature selection procedure based on the presence of a connection. It also possible to make the feature selection discrimination-based; i.e., for two conditions the t-value can be constructed as in a two-sample t-test (an approach demonstrated in Section 4.4 in a post-hoc setting). Note that in this case this criterion (up to a constant) is equivalent to the Fisher ratio (Duda et al., 2001), which maximises inter-class distance while minimising intra-class scatter. It should be emphasised that the feature selection procedure is based on graphs from the training set only: in a leave-one-subject-out cross validation paradigms, only data from $N-1$ subjects is used to learn which connections are statistically significant, and the computed "significance mask" is then applied to the

test data of the left-out subject. Therefore, within a specific cross-validation folds, all graphs have the same number of edges.

Finally, we mention that univariate feature selection is limited because features might be individually irrelevant but provide discrimination when used together (Guyon et al. (2007)). Multivariate feature selection could be preferred, and has been applied to the context of brain decoding before (De Martino et al., 2008).

The multiple comparisons approach only considers feature subsets of cardinality 1, which corresponds to a simple ranking "search" algorithm. It is generally agreed in filter-style feature selection[2] that subsets of different cardinalities need to be considered and an optimum of the objective criterion found via search (annealing, genetic, or floating search are popular).

2.5. Classification

Decision trees are discriminative classifiers performing recursive partitioning of a feature space to yield a potentially non-linear decision boundary. At each decision node of the tree either a single feature f (monothetic trees) or a function of several features $f(\cdot)$ (polythetic trees) is considered so that the entropy of class labels in the partition is minimised. More precisely, if an entropy-based splitting criterion is used, the goal is to find cutpoints of f that minimise the conditional entropy on class labels $\mathcal{C} = \{1, \ldots, C\}$ attached to points in the corresponding subdomains of the discretised variable f'. Specifically, we can express the entropy of the dataset partitioned by the feature f' as

$$H(\mathcal{C}|f') \stackrel{\triangle}{=} -\sum_{j=1}^{2} P_j \sum_{c=1}^{C} P_{j,c} \log_2 P_{j,c}, \tag{1}$$

where P_j is the relative frequency of points in the subset that have value j for feature f', and $P_{j,c}$ is the relative frequency of points that belong to class c and have value j for feature f'.

The goal of decision tree growing is then to minimise (1), which is equivalent to maximising the mutual information between $I(\mathcal{C};f')$, and involves recursively selecting features (or discriminant functions) and computing the result of applying different cutpoints to them. Edge weights that are put to

[2]Where the merit of a feature subset is evaluated via an objective criterion. This is opposed to wrapper-style feature selection where features subset merit is evaluated directly via classification—better subsets provide lower error.

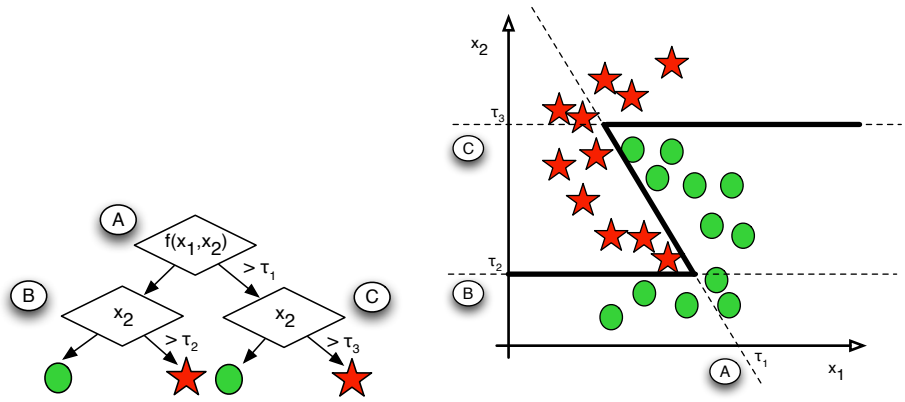
zero by the feature selection are never included as a feature or within a discriminant function since a random variable K that has constant value across all the dataset and classes will not decrease conditional entropy.

Polythetic trees bring two main advantages over monothetic ones; first, decision surfaces are not constrained to be (piecewise) perpendicular to the axis of the feature space, as they are in monothetic trees due to the fact that each node is a decision on a single feature. This is because polythetic nodes can be linear combinations of several features. Second, polythetic trees tend to be more shallow, because each node has more degrees of freedom in partitioning its subspace. Here, we propose to use functional trees (Gama (2004)) that can use multiple regression on a subset of features, both at decision nodes and at leaves. The learning procedure is divided in two phases: growing and pruning. In the growing phase, at each decision node, either a single feature in the original space, or a discriminant function based on linear combination of features is used, depending on which one optimises the splitting criterion. Thus, functional trees are hybrids between monothetic and polythetic decision trees, which results in decision boundaries that are piecewise hyperplanar oblique surfaces. In the pruning phase, functional leaves consisting of discriminant functions can be replaced by a simple function that predicts the class value. We also use logistic regression where the regression functions are learned iteratively by the LogitBoost algorithm (Friedman et al., 2000), which is a refinement suggested by (Landwehr et al. (2005)) and implemented in the Weka framework (Witten and Frank (2005)).

For illustration purposes, we show a functional tree with three nodes and constant leave functions in Fig. 2, along with the corresponding decision boundary in feature space. The first split is a polythetic functional discriminant function of two features x_1 and x_2, and the two splits in the induced subspaces are performed on features in the original space.

Compared to SVMs with either linear or radial basis function kernels, functional trees offer the convenience of adaptively adjusting the model according to feature space complexity. This means that an ad-hoc switch between linear and non-linear decision boundary is effected during training, and that very few parameters need to be optimised beforehand. The proper use of an SVM classifier requires the careful choice of several parameters, including at a minimum kernel type and cost parameter. This is typically performed using a multi-dimensional parameter sweep, for example a grid search for linear polynomial kernels.

Further on, the increased variance of decision trees compared to SVMs



(a) Functional tree classifier (b) Decision boundary

Figure 2: Functional decision tree classifier and corresponding decision boundary principle for a two-dimensional feature space. x_1 and x_2 correspond to 2 different adjacency matrix edge weights, forming a 2-dimensional feature space. In 2(b), green circles correspond to instances of edge weights in one class (condition), while red stars correspond to instances the other class. Hyperplanes corresponding to individual node decision boundaries are shown in dashed lines and identified with circled letters. The overall tree decision boundary is the piecewise linear thick line. Data is a synthetic scaled-down version of the classical *banana* machine learning dataset.

(in the sense of Kohavi and Wolpert (1996) and Geman et al. (1992)) can be turned to an advantage when ensembling several classifiers (see next Section). The decision trees' variance can also be reduced by bagging (Breiman, 1996), which creates multiple "bootstrapped" sets of data by repeatedly sampling with replacement from the training set within each frequency subband, and then averages the prediction of the diverse classifiers trained on the bootstrap samples.

To obtain a valid classification approach, we derive the training samples and grow the decision trees using a leave-one-subject-out cross-validation procedure, by which the data of $N-1$ subjects are used for training, and the data of 1 subject is used for testing. The training and testing partition is then rotated N times. This well-motivated evaluation procedure implies that inter-subject decoding takes place in our experimental setting.

2.6. Ensembling subband graph classifiers

Because of the differences in the functional connectivity graphs in different frequency subbands, it is likely that the subband-specific classifiers will

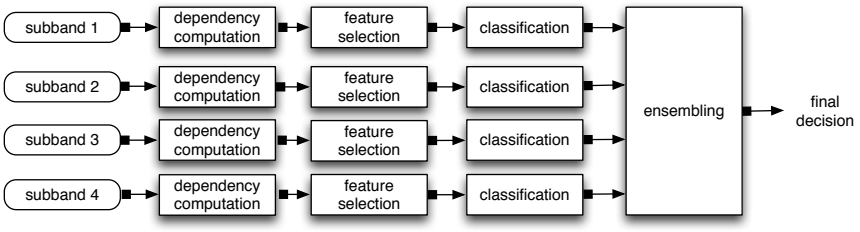


Figure 3: Flowchart of the classification and ensembling procedure, following the preprocessing procedure of Fig. 1.

learn substantially different parameters. Consequently, it is also likely that errors made by classifiers trained in different frequency subbands will be uncorrelated to a certain extent: by explicitly representing the fMRI signal at different scales, we create diversity in classifiers.

By combining the decisions of the classifiers in each subband, and assuming each such *base classifier* performs above chance, we can in theory obtain higher accuracy than the best single subband classifier. This interesting results rests on the *Condorcet jury theorem* (Boland (1989)), which states that the ensemble accuracy P_{MV} of L independent classifiers, each performing with accuracy $p_l > 0.5$ and combined through majority voting, increases monotonically as a function of the number of classifiers, and ultimately reaches perfect decoding for $L \to \infty$.

Given that the classifiers operate on subbands derived by an orthogonal DWT, it can be expected that their output will be quasi class-conditionally independent. In our particular case, we also note that L is at most 4. Lower and upper achievable accuracy bounds can be obtained on finite-sized ensembles (see Appendix).

The overall classification scheme is shown in schematic form in Fig. 3.

2.7. Discriminative graph construction

Based on the set of N trained decision tree classifiers, we propose to extract the subset of connections that is most discriminative for the cross-validation folds and give an easy-to-interpret feedback to the neuroscientists. For that purpose, we sum the regression weights of each connection appearing in a decision tree (the same connection can appear at various levels) and multiply them by the mean classification accuracy of the fold. This yields a composite weight for each connection indicating its discriminative power.

Several variations on using trees to judge feature relevance could be considered as well (Breiman et al., 1984).

This procedure allows us to build a discriminative graph that represents functional connectivity that is markedly different between the conditions of interest. The principle is also similar to the idea of extracting discriminative volumes from classifier parameters using voxels' intensity as features (Mourao-Miranda et al. (2005); Sato et al. (2009)).

3. Materials

3.1. Subjects and data acquisition

The $N = 15$ subjects (4 male, 11 female) were aged between 18 and 36 years old, without history of neurological disorders. They had given written informed consent to participate in the study, which was performed in accordance with the local Ethics Committee of the University of Geneva. Scanning was performed on a Siemens 3T Tim Trio. Functional imaging data were acquired in two sessions using gradient-echo echo-planar imaging (TR/TE/FA = 1.1s/27ms/90°, matrix = 64×64, voxel size = 3.75×3.75×4.2mm^3, 21 contiguous transverse slices, 1.05mm gap, 2598 volumes). Structural imaging data was acquired using a three-dimensional MPRAGE sequence (192 slices, TR/TE/FA = 1.9s/2.32ms/9°, matrix = 256 × 256, voxel size = 0.90 × 0.90 × 0.90mm^3).

3.2. Experimental design

The experimental design was a blocked design with alternating epochs of movie excerpts (50s) and resting periods (90s). All movies were excerpts taken from commercial series. During rest, subjects are instructed to close their eyes, relax, let their mind wander and avoid thinking of something in particular, as usually done in resting-state studies. The movies were projected on a screen through a mirror and the auditory stimulation was provided through MRI-compatible headphones. An empty grey display was projected before and after each movie. At the end of the rest period, a short beep sound was played to instruct subjects to open their eyes, followed by a display asking them to respond to a four-choices question about the content of their thoughts during the rest period. The data collection consists of 2 sessions with 9 movie-resting blocks each, for a total session duration of 23 minutes. The movie order was pseudo-randomised across the subjects.

One reason to use a movie task is that for reliable connectivity analysis, about 10 minutes of data are required. Since we acquire data in 9 task-resting blocks, using a motor task such as finger tapping would require (9x50s=) 7.5 minutes of tapping. This is likely to induce boredom and habituation in subjects, as would an n-back type task, and may lead to non-cooperation and frustration. Furthermore, in choosing an audio-visual task, we are also evoking a distributed pattern of activations in subjects (at a minimum involving the visual cortex in the occipital lobe and auditory areas in the temporal lobes), which is likely to be more difficult to distinguish from distributed resting-state activity than more localised tasks. This seems like a better proof-of-concept for the feasibility of using whole-brain functional connectivity as a basis for classification.

We build up the matrix $\mathbf{X}$ from the blocked experimental design; i.e., for each subject, 90 timecourses are constituted from blocks of the same condition that are concatenated after linear detrending. For the resting condition, this approach is known to only differ from continuous resting-state analysis in a few regions, where it tends to underestimate correlations (Fair et al., 2007).

4. Results

4.1. Classification

Classifier training and testing are performed using the leave-one-subject-out procedure outlined in Section 2.5. Stratified classification results are given in Table 1 (column $L = 1, \alpha = 100\%$): in each of the 15 cross-validation folds, 2 tests are performed, one for the resting and one for the movies condition of a single subject. Thus, the granularity of results is about 3%. One striking result is that lower-frequency subbands (3 and 4) have much more discriminative power than higher-frequency subbands (1 and 2), which are at chance level. This can attributed to the fact that low-frequency activity correlations are substantially different between conditions, while the difference in higher-frequency correlations are much more difficult to pick up.

To provide insight into the experimental conditions, we also performed classification experiments in subband 3 by computing average regional *activations* per condition (thus generating a 90 dimensional vector space), after centering each timecourse with respect to that subject's and condition's regional temporal average vector. This yields about 76% leave-one-subject-out accuracy for the best classifier used. If the timecourses are centred using

Threshold		$\alpha = 100\%$	$\alpha = 5\%$	$\alpha = 2\%$	$\alpha = 1\%$	
L		1	21			
subband	1 (0.23-0.45 Hz)	47%	53%	47%	63%	53%
	2 (0.11-0.23 Hz)	53%	80%	80%	80%	87%
	3 (0.06-0.11 Hz)	87%	93%	90%	90%	90%
	4 (0.03-0.06 Hz)	80%	83%	83%	80%	83%
ensemble (bands 2-4)		90%	97%	90%	90%	93%

Table 1: Leave-one-subject-out overall accuracy with various significance thresholds α for feature selection. $\alpha = 100\%$ corresponds to no feature selection. L is the number of classifiers forming the in-band ensembles. The *ensemble* line refers to multi-band ensembling.

subject-specific, but inter-condition average vectors (a way to encode the difference with baseline rest activations), 100% accuracy is obtained, as would be expected from the conditions chosen and the GLM results in Fig. 7; in this case results cannot be directly compared with the connectivity-based classification results because these use only data from one subject and one condition individually for preprocessing. These results indicate that the tasks have very distinct activation patterns, but that differences in functional connectivity are slightly more subtle.

Note that when feature selection is applied, the significance masks (see Section 2.2) are computed in-fold using only the training data. They are applied to both the training and testing adjacency matrices. When performing feature selection, the number of connections passing the significance threshold was observed to be significantly higher for low frequency subbands.

4.2. Ensemble classification

4.2.1. In-band

Within each subband, we created ensembles of 21 functional trees on bootstrap replicas of the training data. The results are shown in Table 1. All frequency subbands benefit somewhat from the ensembling, with subband 2 showing the most improvement, and showing substantial accuracy. This indicates that the connectivity patterns in subband 2 are quite complex, and the decision boundary learned by the ensemble is certainly non-linear. Subband 1 sees very little improvement in accuracy, indicating that connectivity in this frequency subband is probably not discriminative between rest and movie conditions.

4.2.2. Multi-band

Given the close-to-chance accuracy of the classifier in subband 1 for most threshold settings, we excluded it from the majority voting ensemble to satisfy the Condorcet jury theorem (Section 2.6). Classification accuracy for multi-band ensembling, shown in Table 1, improves or at least is the same as the accuracy in the best individual subband ensemble. The results are consistent with theoretical lower bounds (see Appendix), which for the results in table 1 are 0.67 in the worst case ($L = 1, \alpha = 100\%$), and 0.8 in the best case ($L = 1, \alpha = 1\%$), while the theoretical upper bound is 1 in all cases.

An analysis of the classifier outputs suggests that the main cause for the improvement in the single-classifier case ($L = 1$) despite the very low accuracy of subband 2 is the negative correlation between the decisions of the classifier in subband 2 and that of the classifier in subband 4 (phi coefficient = -0.14).

4.3. Discriminative graphs

A discriminative graph $\mathcal{H}$ can be extracted for each subband from the distribution of classifier parameters as explained in Section 2.7. To ease visualisation and interpretation, we split the associated $\mathcal{H}$ in two discriminative subgraphs; i.e., from the sign of the contrast "movies > rest", obtained from the N adjacency matrices, we derive the subgraphs $\mathcal{H}_+$ and $\mathcal{H}_-$ for positive and negative values, respectively. It should be noted that both subgraphs carry discriminative power to distinguish conditions. However, $\mathcal{H}_+$ reflects those connections that are stronger in movies than rest condition, and vice versa. We thus have $\mathcal{H} = \mathcal{H}_+ \cup \mathcal{H}_-$. The discriminative graphs are shown in Fig. 4. Spheres represent the "functional connectedness" of a region with the rest of the brain - larger spheres means the region is more correlated with other regions.

For all frequency subbands, we observe that connections that are stronger in rest ($\mathcal{H}_-$) present more discriminative ability than those that are stronger in movies. The cuneus is a majorly connected region, present in all subbands, with particularly strong connections to occipital areas.

We also extract histograms depicting the importance (in terms of discriminative power) of each connection over the 15 cross-validation folds, shown in Fig. 5. The histograms are also colour-coded to reflect connections that are stronger in the resting condition (in blue) and those that are stronger in the movies condition (in red).

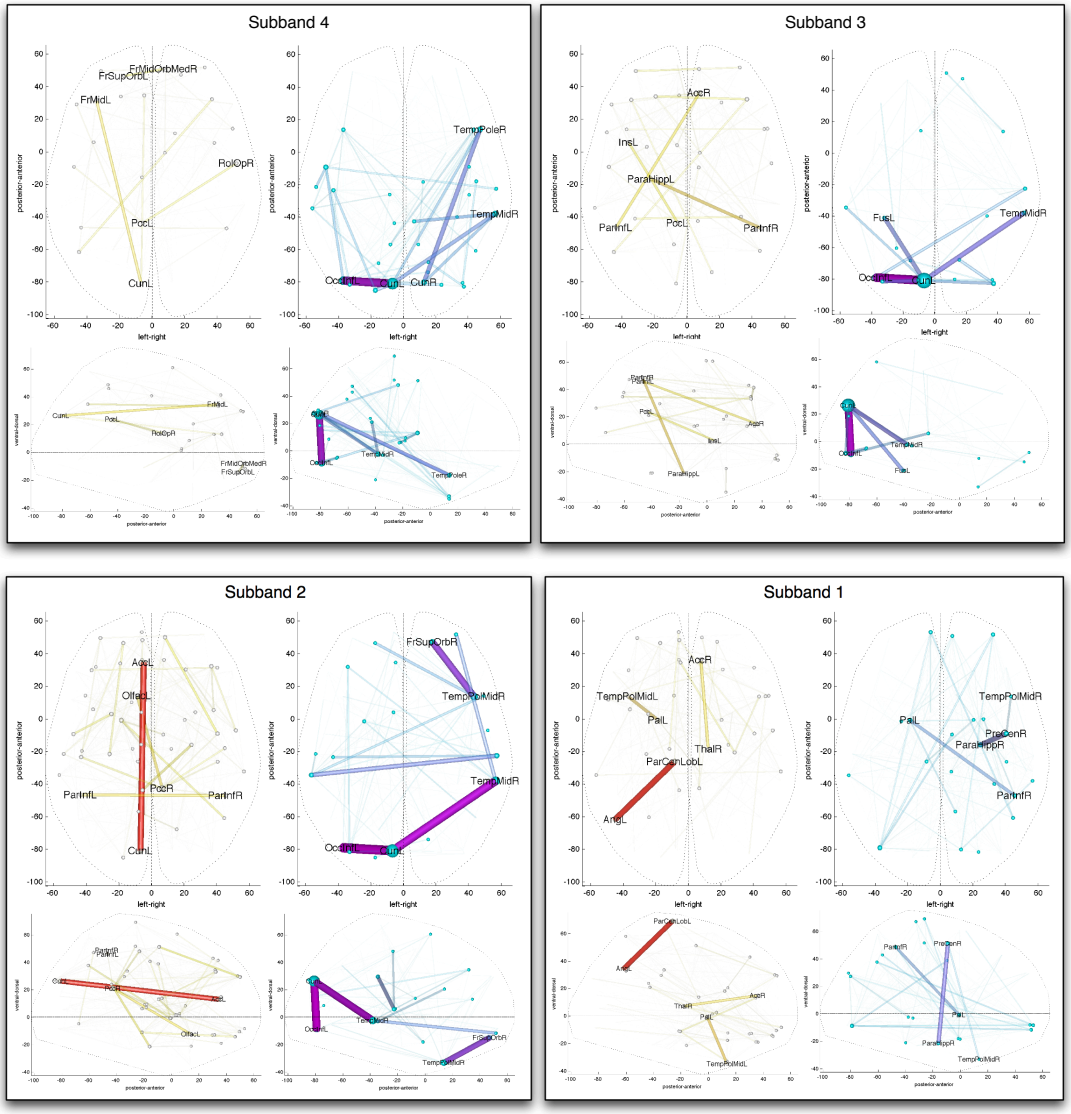


Figure 4: Axial (top rows) and sagittal (bottom rows) views of discriminative graphs $\mathcal{H}$ subdivided in subgraphs $\mathcal{H}_+$ (left columns, warm hues, connections stronger in movies than in resting) and $\mathcal{H}_-$ (right columns, cold hues, connections stronger in resting than in movies) for all subbands. Connections with darker colour and thicker lines correspond to more discriminative ability. Larger spheres correspond to more correlated regions. The same colormap scale is used for each subband.

It appears clearly that the most discriminative connections are always those that are part of $\mathcal{H}_-$ (stronger in resting). It is also noteworthy that the connections belonging to $\mathcal{H}_+$ (stronger in movies) steadily gain discriminative power with rising frequency - the proportion of $\mathcal{H}_+$ connections rising into the top 50 most discriminative connections moves from 28% in subband 4 to 64% in subband 2. This is a clear indication that the "movies" network becomes more distinctive in higher frequency bands.

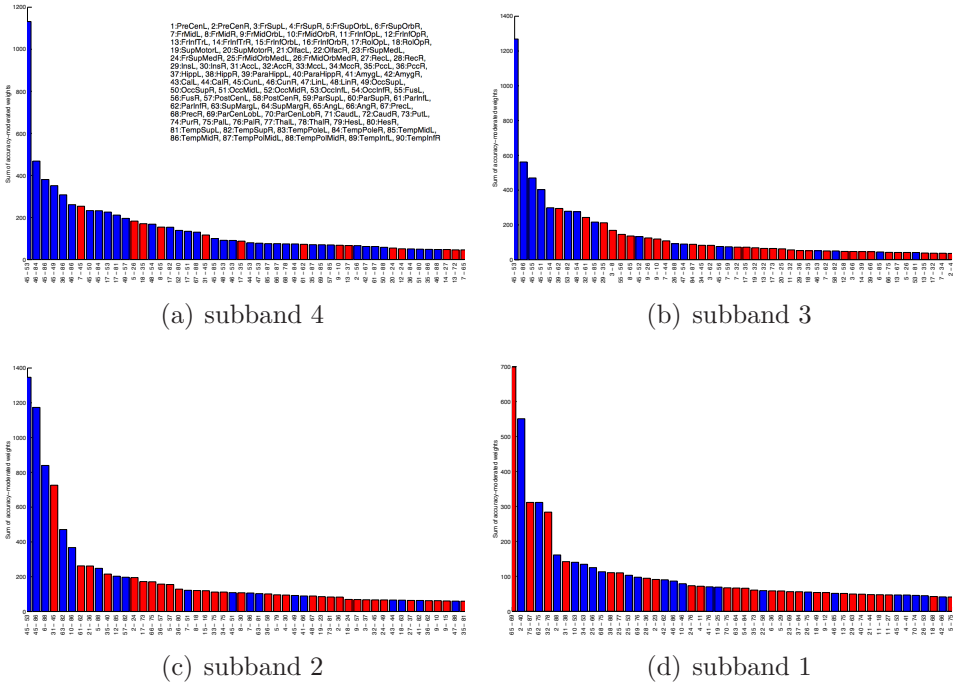
Another interesting aspect is that the number of unique connections picked up by the ensemble of functional trees classifier increases roughly in proportion classification accuracy: Over all cross-validation folds and all trees, 399 unique connections were selected by the classifier in subband 4, 306 in subband 3, and 568 in subband 2. This suggests that the discriminative networks have a larger spatial extent with increasing frequency, and that discriminative power is more distributed between connections at high frequencies.

4.4. Comparison with post-hoc whole-group contrasts

We now compare discriminative graphs obtained with a simple post-hoc group-level method in subband 3. In this method, no train/test separation is performed. All 15 correlation matrices for the two conditions were used to perform multiple t-tests for the difference in means of each correlation coefficient. FDR correction was applied. Subsequently, a significance threshold of $\alpha = 5\%$ was applied to prune the matrix. Only 23 coefficients passed the threshold.

Fig. 6 illustrates the mean condition-specific correlation matrices, the map of p-values, and the retained connections. Coherent with the classifier-based analysis (see Histogram of discriminative connections in Fig. 5(b)), the cuneus has the most connectivity to other regions, and most significant connections originate/terminate in the occipital lobe. A few connections concern the superior temporal lobe. The connections passing the significance threshold form a subset of the most discriminative connections found by the classifier-based methods.

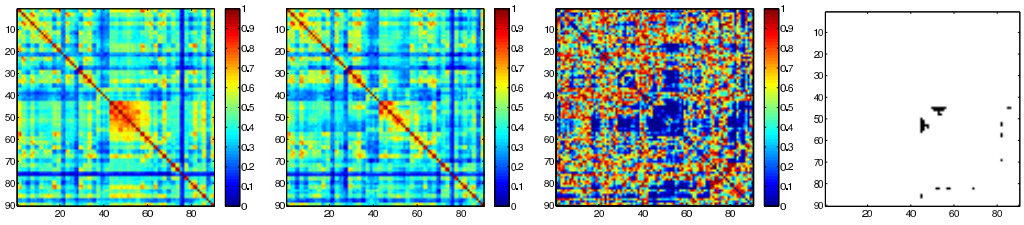
This provides a reassuring confirmation that the discriminative graphs extracted using the classifier-based method proposed make sense in light of a classical hypothesis testing analysis.



(a) subband 4 (b) subband 3

(c) subband 2 (d) subband 1

Figure 5: Histograms of connection discriminative ability for rest/movies classification across 15 cross-validation folds. Only 50 most discriminative connections are shown. Discriminative ability is computed as the sum of the regression weights of a connection across folds, multiplied by the accuracy of the classifier in each fold. Red bars correspond to connections which are stronger in movies than in resting, while blue bars correspond to the opposite situation. Note different scales between graphs.



(a) (b) (c) (d)

Figure 6: Whole-group mean correlation matrices for resting (6(a)) and movies (6(a)) conditions in subband 3, contrast t-test p-values (6(c)), and position of significant connections (6(d)). Region labels are as per AAL atlas and Fig. 5(a)

4.5. General linear model analysis

For comparison purposes, we also perform a standard confirmatory analysis based on the general linear model (GLM). The realigned data is analysed using conventional SPM methodology. Specifically, the design matrix is constructed for each subject to model movies versus rest, and the content question that is asked after each resting period is added as a covariate of no interest. The stimulation functions are then convolved with the canonical haemodynamic response function. We also add the realignment parameters and low-frequency components (cut- off frequency at 1/256Hz). The GLM model is fitted for each subject and the estimated contrast of interest is fed into a second level analysis after normalisation and regional averaging.

The positive (movies > rest) and negative (rest > movies) contrasts are shown in Fig. 7. During the movies condition, increase in activity is observed in visual, auditory, and multisensory regions (mostly occipital and large extents of the temporal cortex). During rest, brain regions associated with the default mode network (Raichle et al., 2001; Greicius et al., 2003) were significantly active, including posterior and anterior cingulate, bilateral insula, and bilateral inferior parietal lobules.

5. Discussion

5.1. Classification

The relatively high number of connections that are retained by statistical feature selection (Section 2.4) in the low-frequency subbands, with respect to the number retained in higher-frequency subbands, hints at the presence of resting-state networks that are consistent across subjects (Damoiseaux et al. (2006); Mantini et al. (2007)), which in turn yields relatively low inter-subject standard deviation on the weights of graph edges that these networks comprise. Not only are low-frequency connections more consistent across subjects, but those that are stronger in resting than in movies (edges of $\mathcal{H}_-$) are also much more discriminative. This strongly suggests that much of the inter-subject discriminative ability between the two cognitive states is due to the inter-subject topographical consistency of functional connectivity in resting state. Applying the methodology to other tasks, which may elicit responses that are generally coherent between normal subjects (such as face viewing), would probably yield different results, and it may be that discriminative characteristics would be more distributed between conditions.

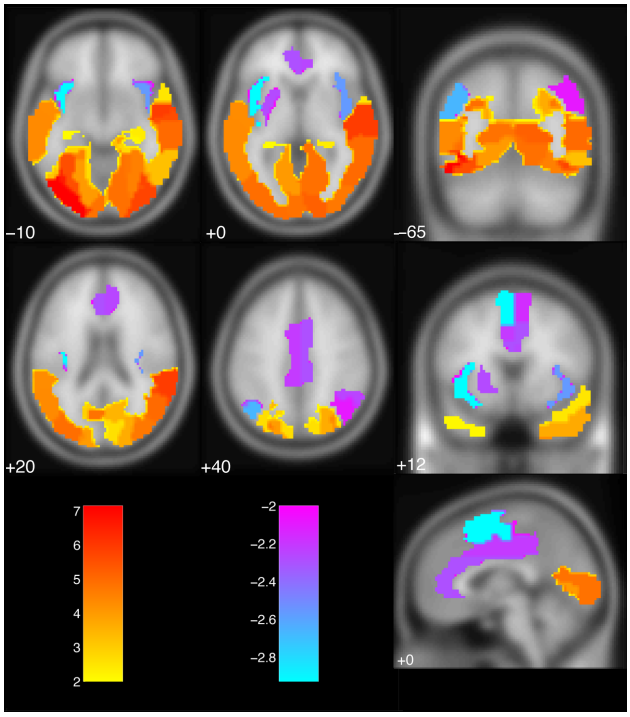


Figure 7: Axial, coronal, and sagittal slices (MNI normalised and region-averaged) showing the contrast movies versus rest at the group level. Colour correspond to t-values. The red-to-yellow colormap corresponds to movies, while the purple-to-blue colormap corresponds to resting.

In general, feature selection in low subbands seems to have no effect, while it might improve classification accuracy in higher subbands. The first explanation is that, in low subbands, even at $\alpha = 1\%$, most of the connections are retained, potentially due to the effect explained above. Thus, in low subbands, much lower significance values would be needed to have a noticeable effect. In high subbands, it is possible to find a setting of α that yields better accuracy, but it is by no means obvious to predict which threshold will produce the best result. This difficulty can be attributed to two factors. First, the search strategy, or lack thereof. In reproducing the common practice of FDR thresholding, we have shown that performing multiple-comparison correction for edge pruning is at its core a univariate ranking-based feature selection method. This means that feature that are irrelevant taken individually, but jointly predictive, will be missed. Second, our choice of classifier performs embedded feature selection, that is, classifier training includes a

specific objective function to select a subset of features that minimises prediction error. Preemptively throwing out features in this context, according to another objective criterion (t-test) can undermine good performance or happen to improve it.

However, it is likely that more weakly regularised classifiers or those that do not perform embedded feature selection might benefit from this approach to feature selection (Guyon et al. (2005)). This emphasises the fact that feature selection and classifier training should not be performed in isolation; i.e., optimising one objective criterion, such as the Fisher ratio, is not guaranteed to tighten the error bounds on the classifier.

In terms of comparing our classification results to other state-of-the-art machines, we note that our bagged subband classifiers perform the same as a linear polynomial SVM (parameters optimised via in-fold cross-validation) two subbands (subband 1 and 3) and slightly worse in two subbands (subbands 2 and 4). Even so, the small amount of test data causes these differences to be insignificant (McNemar test, p=0.05). The number of parameters used by decision trees to achieve our results is much smaller (typically at most 20 attributes are used in each tree, with significant overlap between trees trained over different bags), leading to more interpretable models.

5.2. Discriminative graphs

We start by examining the connection histograms in Fig. 5. For each subband, only a small fraction of the 4005 edges of the complete connectivity graph is retained by the classifiers (399 for subband 4, 306 for subband 3, 568 for subband 2). Moreover, less than about 10% of these connections are most discriminative, which is more than two orders of magnitude lower than the total number of edges. Further on, we observe a transition from resting to movies condition for the most discriminative connections as we go from low to high frequency subbands. The strong connections in subband 4 (0.03–0.06Hz) belong mainly to the resting condition; i.e., low-frequency coherent BOLD fluctuations are reminiscent of resting-state networks. Discriminative connections in the intermediate subband 3 (0.06–0.11Hz) are intermingled between resting and movies. For subband 2 (0.11–0.23Hz), the proportion of connections stronger in the movies condition sees an important increase. Finally, the discriminative importance of connections in subband 1 (0.23–0.45Hz) is very limited; this subband mainly contains noise due to the slow signal changes associated with the haemodynamic response. In general, the different connectivity patterns across subbands demonstrate the advantage

of using the wavelet transform. The fact that low-frequency subbands are mainly dominated by connections stronger in resting condition is in line with the fMRI resting-state literature.

Discriminative graphs in anatomical space for subbands 1–4 are shown in Fig. 4. During rest ($\mathcal{H}_-$), the cuneus appears as a hub in a number of highly discriminative connections toward occipito-temporal regions (subbands 2–4). The significant increase of functional connectivity within visual areas during rest or, equivalently, de-coherence during movie watching, has been reported before (Nir et al. (2006)). It is also known from visual mapping studies that spatial de-coherence breaks down into patterns with different functional specialisation (Grill-Spector and Malach (2004)). The fact that these connections are discriminative form a compelling argument that the underlying BOLD fluctuations have a neuronal basis and cannot attributed to non-neurophysiological sources such as cardiac and respiratory oscillations only. Furthermore, the connectivity patterns also nicely generalise over subjects (since we classify an unseen subject), in accordance with Hasson et al. (2004) who found correlating activities between different subjects watching the same movie.

The intermediate subbands 2 and 3 are particularly interesting since connections of both conditions contribute to successful classification. Intriguingly, BOLD activity relatively increases in dorsal areas during rest versus in ventral areas during movies (see Fig. 7). One possible explanation might be the distinction between intrinsic versus extrinsic activity in parietal versus sensory regions as described by Golland et al. (2007).

Finally, these results confirm that functional connectivity analysis can benefit from filtering timecourses in separate frequency bands as clear differences were observed between the subbands, allowing extraction of different brain networks and probably reflecting different functions (e.g., stimulation versus rest). This view is consistent with recent work in combined EEG-fMRI studies, where distinct functional networks are observed at different EEG frequency bands (Mantini et al., 2007), although at much higher frequencies for EEG than fMRI.

5.3. Activity versus connectivity

The results of the GLM analysis (see Fig. 7) reveal that, as expected, *activity* is increased during movies in visual, auditory, and multisensory regions, while typical default-mode regions are deactivated (posterior and anterior cingulate cortices, bilateral anterior insula, and bilateral inferior parietal

lobules). It is striking that functional connectivity between areas that successfully contributed to the classification task, was often stronger for regions during the condition for which they showed less activation.

For example, cuneus, occipital and temporal regions are clearly activated during movies, however, connectivity between cuneus and occipital and temporal regions increases significantly during rest and even constitutes a discriminative feature between both conditions at low frequencies (subbands 2–4). From the opposite side, midline brain areas are deactivated during movies while showing increased connectivity at high frequencies (subband 2).

It is possible that brain regions differentially engaged by active "task" processing appear not functionally connected according to pairwise linear correlations, but that (conditional) non-linear relationships such as provided by (conditional) mutual information measures would draw a different picture. In any event, these findings demonstrate that connectivity measures contribute to reveal functional organisation differently than sole activation. More generally, these data point to the fact that frequency information may provide very important information when investigation large-scale brain connectivity.

5.4. Potential for applications

One important feature of the proposed methodology is that the classification is based on the connectivity pattern of a single condition; i.e., we do not use connectivity differences between two conditions to create a relative baseline. This makes the method particularly interesting to be applied to the clinical setting; i.e., to distinguish between control and patients. In particular, there is an increasing amount of evidence favouring the presence of a specific anatomical connectivity subtending functional connectivity (Teipel et al. (2010)) that points to the important role that functional connectivity analysis could play in early diagnosis and differential diagnosis; e.g., neurodegenerative and inflammatory diseases.

It is well known that the default-mode network is affected by dementia and other diseases (Buckner et al. (2008); Wang et al. (2007)), and functional connectivity changes are known to be present in a variety of cognitive deficits, even when structural damage is not apparent. For example, differences in inter-hemispheric intra-parietal sulcus correlations can indicate post-stroke spatial neglect even when no structural damage is present (He et al. (2007)). Other diseases or trauma characterised by diffuse white

matter lesions, such as multiple sclerosis or axonal damage sustained from head injuries, are likely candidates for causing significant alterations in functional connectivity. The use of the technique in psychiatric diseases where a dysfunction of distributed networks exist, such as schizophrenia (Bluhm et al., 2007; Garrity et al., 2007), autism (Kennedy and Courchesne, 2008), or obsessive-compulsive disorder (Harrison et al., 2009), could also be particularly valuable where traditional univariate or single ROI methods are likely to be difficult to apply. Our method is data-driven and therefore does not rely on strong hypotheses about existing resting-state networks. Specifically, it is well possible that connections belonging to a subnetwork turn out to be most discriminative.

6. Conclusion

In summary, we have proposed a classification approach to infer brain states from functional connectivity graphs, instead of the commonly used brain voxel activation values. We have shown that the approach is applicable to inter-subject brain decoding with good results, and that interpretable output can be generated. We have demonstrated the feasibility using a cognitive task and compared the discriminative connectivity graph with SPM-style activation patterns. The potential of the proposed methodology lies in situations where connectivity measures are the most readily available, such as for comparing resting-state fMRI between patients and control groups, but also provides complementary information to task-based acquisition paradigms. The method fits well with current trends in clinical neuroscience where multivariate pattern recognition techniques are increasingly used to find increasingly subtle effects in data, inaccessible to mass-univariate methods (Bray et al., 2009).

The current atlasing procedure uses rather coarse-grained brain regions. While this might reduce inter-subject variability, future improvements could be obtained by using more sophisticated segmentation and atlasing methods; e.g., surface-based segmentation (Dale et al., 1999; Fischl et al., 1999) and function-based inter-subject mapping (Sabuncu et al., 2010) or "alignment-free" methods (Anderson et al., 2010).

7. Acknowledgements

The authors wish to thank the anonymous reviewers for their valuable input. This work was supported in part by the Swiss National Science Foun-

dation (grant PP00P2-123438), in part by the Société Académique de Genève and the FOREMANE foundation, and in part by the Center for Biomedical Imaging (CIBM) of the Geneva and Lausanne Universities, EPFL, and the Leenaards and Louis-Jeantet foundations.

Appendix A. Accuracy bounds for majority voting combination

Matan (1996) has shown that for an ensemble of L classifiers, the upper and lower bounds on achievable majority voting accuracy are given by:

$$P_{\max} = \min(1, f(\tau), f(\tau-1), \ldots, f(1)), \quad \text{(A.1)}$$

$$P_{\min} = \max(0, g(\tau), g(\tau-1), \ldots, g(1)), \quad \text{(A.2)}$$

where the functions $f(\tau)$ and $g(\tau)$ are defined in terms of a specific majority decision threshold τ' (an integer) and base classifier accuracies p_l:

$$f(\tau') = \frac{1}{\tau'} \sum_{l=1}^{L-\tau+\tau'} p_l. \quad \text{(A.3)}$$

$$g(\tau') = \frac{1}{\tau'} \sum_{\tau-\tau'+1}^{L} p_l - \frac{L-\tau}{\tau'}. \quad \text{(A.4)}$$

Thus, it can be seen that majority voting ensemble accuracy is a linear function of component classifier accuracies. Note however that this bound does not take into account classifier diversity, a very important and somewhat ill-defined parameter of ensembling. The diversity-accuracy tradeoff is explored in more detail in Meynet and Thiran (2007).

As an illustration, depending on the diversity of the ensemble, majority voting with 3 classifiers barely above chance (51% accuracy on a two-class problem) can yield accuracies between 27% and 77%, while a stronger ensemble (say, three classifiers at 70% accuracy) can yield accuracies between 55% and 100%.

Bibliography

Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., Jan. 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. The Journal of Neuroscience 26 (1), 63–72.

Alemán-Gómez, Y., Melie-García, L., Valdés-Hernandez, P., June 2006. IBASPM: Toolbox for automatic parcellation of brain structures. In: 12th Annual Meeting of the Organization for Human Brain Mapping. Vol. 27 of NeuroImage. Florence, Italy.

Anderson, A., Dinov, I. D., Sherin, J. E., Quintana, J., Yuille, A. L., Cohen, M. S., Feb 2010. Classification of spatially unaligned fmri scans. Neuroimage 49 (3), 2509–2519.

Battle, G., 1987. A block spin construction of ondelettes. Part I: Lemarié functions. Commun. Math. Phys. 110, 601–615.

Beckmann, C., Smith, S., Feb. 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Transactions on Medical Imaging 23 (2), 137–152.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57 (1), 289–300.

Biswal, B., Yetkin, F., Haughton, V., Hyde, J., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magnetic Resonance in Medicine 34 (4), 537–541.

Bluhm, R. L., Miller, J., Lanius, R. A., Osuch, E. A., Boksman, K., Neufeld, R. W. J., Théberge, J., Schaefer, B., Williamson, P., Jul 2007. Spontaneous low-frequency fluctuations in the BOLD signal in schizophrenic patients: anomalies in the default network. Schizophr Bull 33 (4), 1004–1012.

Boland, P. J., 1989. Majority system and the Condorcet jury theorem. Statistician 38 (3), 181–189.

Bray, S., Chang, C., Hoeft, F., 2009. Applications of multivariate pattern classification analyses in developmental neuroimaging of healthy and clinical populations. Front Hum Neurosci 3, 32.

Breiman, L., Aug. 1996. Bagging predictors. Machine Learning 24 (2), 123–140.

Breiman, L., Friedman, J., Stone, C. J., Olshen, R., 1984. Classification and Regression Trees. CRC Press.

Buckner, R., Andrews-Hanna, J., Schacter, D., 2008. The brain's default network: Anatomy, function, and relevance to disease. Annals of the New York Academy of Sciences 1124, 1–38.

Bunke, H., Shearer, K., Mar. 1998. A graph distance metric based on the maximal common subgraph. Pattern Recognition Letters 19 (3-4), 255–259.

Calhoun, V. D., Adali, T., Pearlson, G. D., van Zijl, P. C. M., Pekar, J. J., Jul. 2002. Independent component analysis of fMRI data in the complex domain. Magnetic Resonance in Medicine 48 (1), 180–192.

Conte, D., Foggia, P., Sansone, C., Vento, M., 2004. Thirty years of graph matching in pattern recognition. International Journal of Pattern Recognition and Artificial Intelligence 18 (3), 265–298.

Cox, D., Savoy, R., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19, 261–270.

Craddock, R. C., Holtzheimer, P. E., Hu, X. P., Mayberg, H. S., Dec 2009. Disease state prediction from resting state functional connectivity. Magn Reson Med 62 (6), 1619–1628.

Dale, A. M., Fischl, B., Sereno, M. I., Feb 1999. Cortical surface-based analysis. I. segmentation and surface reconstruction. Neuroimage 9 (2), 179–194.

Damoiseaux, J. S., Rombouts, S. A., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., Beckmann, C. F., 2006. Consistent resting-state networks across healthy subjects. Proceedings of the National Academy of Sciences 103, 13848–13853.

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 43, 44–58.

Duda, R. O., Hart, P. E., Stork, D. G., 2001. Pattern Classification, 2nd Edition. Wiley.

Ethofer, T., Van De Ville, D., Scherer, K., Vuilleumier, P., June 23, 2009. Decoding of emotional information in voice-sensitive cortices. Current Biology 19 (12), 1028–1033.

Fair, D. A., Schlaggar, B. L., Cohen, A. L., Miezin, F. M., Dosenbach, N. U., Wenger, K. K., Fox, M. D., Snyder, A. Z., Raichle, M. E., Petersen, S. E., Mar. 2007. A method for using blocked and event-related fmri data to study "resting state" functional connectivity. NeuroImage 35 (1), 396–405.

Fischl, B., Sereno, M. I., Dale, A. M., Feb 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9 (2), 195–207.

Frackowiak, R., Friston, K., Frith, C., Dolan, R., Mazziotta, J., 1997. Human Brain Function. Academic Press.

Friedman, J., Hastie, T., Tibshirani, R., 2000. Additive logistic regression: A statistical view of boosting. Annals of Statistics 28 (2), 337–407.

Friston, K., Frith, C., Liddle, P., Frackowiak, R., 1993. Functional connectivity: The principal-component analysis of large (pet) data sets. Journal of Cerebral Blood Flow and Metabolism 13 (1), 5–14.

Friston, K. J., Harrison, L., Penny, W. D., 2003. Dynamic causal modeling. NeuroImage 19, 1273–1302.

Gama, J., Jun. 2004. Functional trees. Machine Learning 55 (3), 219–250.

Garrity, A. G., Pearlson, G. D., McKiernan, K., Lloyd, D., Kiehl, K. A., Calhoun, V. D., Mar 2007. Aberrant "default mode" functional connectivity in schizophrenia. The American Journal of Psychiatry 164 (3), 450–457.

Geman, S., Bienenstock, E., Doursat, R., Jan. 1992. Neural networks and the bias/variance dilemma. Neural Computation 4 (1), 1–58.

Golland, Y., Bentin, S., Gelbard, H., Benjamini, Y., Heller, R., Nir, Y., Hasson, U., Malach, R., 2007. Extrinsic and Intrinsic Systems in the Posterior Cortex of the Human Brain Revealed during Natural Sensory Stimulation. Cereb. Cortex 17 (4), 766–777.

Greicius, M. D., Krasnow, B., Reiss, A. L., Menon, V., Jan. 2003. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proceedings of the National Academy of Sciences of the United States of America 100 (1), 253–258.

Grill-Spector, K., Malach, R., 2004. The human visual cortex. Annual Review of Neuroscience 27, 649–677.

Guyon, I., Aliferis, C., Elisseeff, A., 2007. Computational Methods of Feature Selection. Chapman and Hall, Ch. Causal Feature Selection.

Guyon, I., Gunn, S., Hur, A. B., Dror, G., 2005. Result analysis of the NIPS 2003 feature selection challenge. In: Advances in Neural Information Processing Systems 17. pp. 545–552.

Harrison, B. J., Soriano-Mas, C., Pujol, J., Ortiz, H., López-Solà, M., Hernández-Ribas, R., Deus, J., Alonso, P., Yücel, M., Pantelis, C., Menchon, J. M., Cardoner, N., Nov 2009. Altered corticostriatal functional connectivity in obsessive-compulsive disorder. Arch Gen Psychiatry 66 (11), 1189–1200.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., Mar 2004. Intersubject synchronization of cortical activity during natural vision. Science 303 (5664), 1634–1640.

Haxby, J. V., Gobbini, I. M., Furey, M. L., Ishai, A., Schouten, J. L., Pietrini, P., September 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293 (5539), 2425–2430.

Haynes, J., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. 8, 686–691.

Haynes, J.-D., Rees, G., Jul 2006. Decoding mental states from brain activity in humans. Nature Reviews Neuroscience 7 (7), 523–534.

He, B. J., Snyder, A. Z., Vincent, J. L., Epstein, A., Shulman, G. L., Corbetta, M., Mar. 2007. Breakdown of functional connectivity in frontoparietal networks underlies behavioral deficits in spatial neglect. Neuron 53 (6), 905–918.

Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8, 679–685.

Kay, K., Naselaris, T., Prenger, R., Gallant, J., 2008. Identifying natural images from human brain activity. Nature 452, 352–355. 452, 352–355.

Kennedy, D. P., Courchesne, E., Feb 2008. The intrinsic functional organization of the brain is altered in autism. Neuroimage 39 (4), 1877–1885.

Kohavi, R., Wolpert, D. H., 1996. Bias plus variance for zero-one loss functions. Proc. 13th Int. Machine Learning Conf.

Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. PNAS 103, 3863–3868.

Landwehr, N., Hall, M., Frank, E., May 2005. Logistic model trees. Machine Learning 59 (1), 161–205.

Lowe, M. J., Mock, B. J., Sorenson, J. A., Feb. 1998. Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations,. NeuroImage 7 (2), 119–132.

Mantini, D., Perrucci, M. G., Gratta, C. D., Romani, G. L., Corbetta, M., Aug 2007. Electrophysiological signatures of resting state networks in the human brain. Proc National Academy of Sciences of the USA 104 (32), 13170–13175.

Matan, O., 1996. On voting ensembles of classifiers (extended abstract). In: Working Notes of the Workshop on Integrating Multiple Learned Models for Improving and Scaling Machine Learning Algorithms. Portland, USA, held in conjunction with the 13th Nat. Conf. on Artificial Intelligence (AAAI-96).

McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., Sejnowski, T., 1998. Analysis of fMRI data by blind separation into independent spatial components. Human Brain Mapping 6, 160–188.

Meynet, J., Thiran, J.-P., 2007. Information theoretic combination of classifiers with application to adaboost. In: Proc. 7th Int. Workshop on Multiple Classifier Systems. pp. 171–179.

Mitchell, T., Hutchinson, R., Niculescu, R., Pereira, F., Wang, X., Just, M., Newman, S., 2004. Learning to decode cognitive states from brain images. Machine Learning 57, 145–175.

Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M., Morito, Y., Tanabe, H. C., Sadato, N., Kamitan, Y., 2008. Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron 60, 915–929.

Mourao-Miranda, J., Bokde, A. L., Born, C., Hampel, H., Stetter, M., Dec. 2005. Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. NeuroImage 28 (4), 980–995.

Mourao-Miranda, J., Friston, K. J., Brammer, M., May 2007. Dynamic discrimination analysis: A spatial-temporal SVM. NeuroImage 36 (1), 88–99.

Nir, Y., Hasson, U., Levy, I., Yeshurun, Y., Malach, R., 2006. Widespread functional connectivity and fmri fluctuations in human visual cortex in the absence of visual stimulation. NeuroImage 30 (4), 1313 – 1324.

Norman, K., Polyn, S., Detre, G., Haxby, J., 2006. Beyond mind-reading: multivoxel pattern analysis of fMRI data. Trends Cogn. Sci. 10, 424–430.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., Shulman, G. L., Jan 2001. A default mode of brain function. Proceedings of the National Academy of Sciences of the United States of America 98 (2), 676–682.

Richiardi, J., Van De Ville, D., Riesen, K., Bunke, H., 2010. Vector space embedding of undirected graphs with fixed-cardinality vertex sequences for classification. In: Proc. 20th Int. Conf. on Pattern Recognition (ICPR).

Sabuncu, M. R., Singer, B. D., Conroy, B., Bryan, R. E., Ramadge, P. J., Haxby, J. V., Jan 2010. Function-based intersubject alignment of human cortical anatomy. Cerebral Cortex 20 (1), 130–140.

Salvador, R., Suckling, J., Schwarzbauer, C., Bullmore, E., 2005. Undirected graphs of frequency-dependent functional connectivity in whole brain networks. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 937–946.

Sato, J. R., Fujita, A., Thomaz, C. E., Martin, M. d. G. M., Mourao-Miranda, J., Brammer, M. J., Junior, E. A., May 2009. Evaluating SVM and MLDA in the extraction of discriminant regions for mental state prediction. NeuroImage 46 (1), 105–114.

Sporns, O., Tononi, G., Edelman, G., 2000. Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. Neural Netw. 13, 909–922.

Teipel, S. J., Bokde, A. L., Meindl, T., Amaro Jr., E., Soldner, J., Reiser, M. F., Herpertz, S. C., Müller, H.-J., Hampel, H., 2010. White matter microstructure underlying default mode network connectivity in the human brain. NeuroImage 49 (3), 2021–2032.

Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J., Lebihan, D., Dehaene, S., 2006. Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage 33, 1104–1116.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289.

Umeyama, S., 1988. An eigendecomposition approach to weighted graph matching problems. IEEE Trans. on Pattern Analysis and Machine Intelligence 10 (5), 695–703.

Wang, K., Liang, M., Wang, L., Tian, L., Zhang, X., Li, K., Jiang, T., 2007. Altered functional connectivity in early alzheimer's disease: A resting-state fMRI study. Human Brain Mapping 28 (10), 967–978.

Witten, I. H., Frank, E., 2005. Data Mining: Practical machine learning tools and techniques, 2nd Edition. Morgan Kaufman.

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore, E. T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? NeuroImage 50, 970–983.