Decoding spontaneous brain activity from fMRI using Gaussian Processes: tracking brain reactivation

Jessica Schrouff*†, Caroline Kussé*, Louis Wehenkel†‡, Pierre Maquet*§ and Christophe Phillips*†
*Cyclotron Research Centre, University of Liège, Belgium
†Department of Electrical Engineering and Computer Science, University of Liège, Belgium
‡Giga-R, Systems Biology and Chemical Biology, University of Liège, Belgium
§Department of Neurology, Liège University Hospital, Belgium

***Abstract*—While Multi-Variate Pattern Analysis techniques based on machine learning have now been regularly applied to neuroimaging data, decoding brain activity is usually performed in highly controlled experimental paradigms. In more realistic conditions, the number, sequence and duration of mental states are unpredictably generated by the individual, resulting in complex and imbalanced fMRI data sets. Moreover, in the case of spontaneous brain activity, the mental states can not be linked to any external or internal stimulation, which makes it a highly difficult condition to decode. This study tests the classification of brain activity, acquired on 14 volunteers using fMRI, during mental imagery, a condition in which the number and duration of mental events were not externally imposed but self-generated. Application of the obtained model on rest sessions allowed classifying spontaneous brain activity linked to the task which, overall, correlated with their behavioural performance to the task.**

***Keywords*-brain decoding; fMRI; spontaneous activity; semi-constrained activity;**

I. Introduction

Multivariate methods, also known as brain decoding or mind reading, aim at associating a particular cognitive, behavioural or perceptual state to specific patterns of regional brain activity [1]. During the last years, methods such as Support Vector Machines (SVM, [2]) or Gaussian Processes (GP, [3]) classifiers were applied to functional Magnetic Resonance Imaging (fMRI) times series to predict, from individual brain activity, the patterns of perceived objects [4] mental states related to memory retrieval [5] or even hidden intentions [6]. GP classifiers, which provide a principled probabilistic approach to kernel machine learning, have been recently developed to allow for classifying more difficult data sets, such as predicting subjective pain intensity [7]. In most of these studies, the experimental design completely controlled the nature, timing and duration of experimental trials, and temporally isolated experimental conditions from one another. However, in some contexts, decoding more realistic data sets would be desirable. An example is the study of memory traces: patterns of brain activity present during a memory task should be reactivated during a post-task rest period [8], [9]. Due to the complexity of this phenomenon, few studies could investigate the reactivation of distributed memory traces in humans using non-invasive neuroimaging. To the best of our knowledge, Tambini et al [9] are the only group who could investigate off-line transfer of information using fMRI. They showed an enhanced marginal correlation between the hippocampus and neocortex (lateral occipital cortex) during post-task rest compared to baseline rest, which predicted individual differences in later associative memory. They however used seed regions in a univariate fashion whilst the reactivation of distributed patterns is an intrinsically multivariate problem. In this work, we acquired a memory task flanked by two rest sessions and followed by a recall or mental imagery session. The aim of this work was to (i) classify the mental imagery session comprising uneven numbers of short events with possibly overlapping patterns of brain activity, (ii) apply the computed model to the different rest sessions and (iii) establish the relationship between the outputs of such classifications and behavioural data.

II. Materials and Methods

A. Population

A group of 14 volunteers (7 females), aged between 19 and 29 years (mean 24.44), participated in the study. This study was approved by the Ethical Committee of the Faculty of Medicine of the University of Liège. All subjects were fully informed and gave their written informed consent. They were screened for anxiety, depression, sleep quality, chronotype and excessive daytime sleepiness.

B. Experimental design

All subjects underwent two main conditions: a memory condition and a control condition, their order being randomized (see fig.1). The memory task consisted in an *exploration* session, during which images of faces, buildings and animals were displayed one at a time for 3 seconds, each image being assigned a specific location on the screen. The order of presentation followed a predefined sequence of contiguous screen positions in such a way that volunteers had the impression of following a path throughout a bidimensional maze. The complete maze consisted of three blocks of 27 consecutive images within which the 3 categories of images

978-0-7695-4765-7/12 $26.00 © 2012 IEEE
DOI 10.1109/PRNI.2012.19

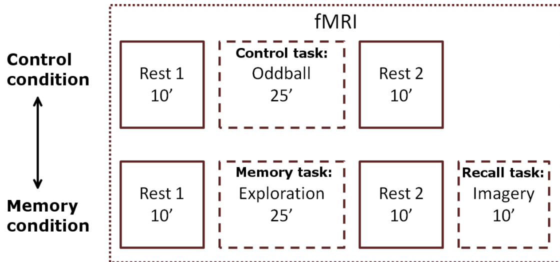

Figure 1. Experimental design: subjects underwent a control task flanked by two rest sessions and a memory task, flanked by two rest sessions and followed by a recall or mental imagery session.

were always presented in the same order (i.e. 9 faces, 9 buildings and 9 animals). This session was flanked by two rest sessions (referred to as '$R1_m$' and '$R2_m$' respectively, each lasting 10 minutes, [8]) and followed by a *mental imagery* session, in which volunteers were presented with 54 memory tests. During each test, two images, simultaneously displayed on the screen for 4 seconds, represented the starting and target positions of a path that the volunteers would have to follow mentally. The mental trajectories included 3 to 6 images (average 4.5) of a same category. Volunteers had to signal by a key press each image that they could conjure up during this mental travel. The control task consisted in an oddball experiment (discriminative auditory task), flanked by two rest sessions (10 minutes, referred to as '$R1_o$' and '$R2_o$' respectively).
A memory test was finally performed outside the scanner, in order to behaviourally assess the accuracy of both the spatial and content knowledge acquired by the volunteers.

C. Data acquisition and image preprocessing

fMRI time series were acquired on a 3T head-only scanner (Magnetom Allegra, Siemens Medical Solutions, Erlangen, Germany). The images were preprocessed using SPM8 [1] to correct for spatial deformations induced by the field inhomogeneities and differences in slice acquisition time. They were then simultaneously realigned and unwrapped to account for the subject movements in the scanner, and finally smoothed using a Gaussian function with a 4mm FWHM kernel.

D. Functional MRI data analysis

The selected methodology is based on [10], in which the authors compared different (combinations of) techniques to classify exploration and mental imagery sessions. The classification technique considered here was used in a within-subject and binary way.

1) Signal extraction: For the different sessions, the whole time series of all voxels were extracted. A GLM [11] was used to regress out movement effects and low frequency drifts. The signal corresponding to stimulus onsets was then extracted, considering a hemodynamic response function (HRF) delay of 6 seconds [12]. To avoid decoding the signal linked to motor activity in the mental imagery session, the scans selected for further classification were the ones preceding the key presses (after correction for HRF delay). Overlapping events were also handled with care, preventing the inclusion of two different stimuli in the same TR. For spontaneous brain activity, each scan of the rest session was considered as an 'event'.

2) Feature selection: From [10], the best results were obtained using a double-step feature selection comprising both a univariate and multivariate technique. The univariate technique consisted in a GLM which selected the subset of 1000 most 'active' voxels. The multivariate feature selection was based on an SVM classifier [2]. It consisted in a Recursive Feature Addition, which recursively adds the most discriminant voxels until a decrease or plateau in the global accuracy (i.e. the sum of the accuracies of each binary comparison) was observed. Care was taken to perform the feature selection on the training set only, to ensure unbiased estimations of the accuracy.

3) Classification method: Gaussian Process (GP) classification [3], was performed using the Expectation Propagation approximation of the posterior mode [3], which recursively updates local parameters of the distribution. The covariance matrix was modelled by a linear kernel matrix.

4) Mental imagery: For mental imagery, the outputs of the classification consisted in accuracy (balanced and class accuracies) measured on a Leave-One block-Out cross-validation. To obtain a multiclass accuracy measure for each subject (instead of one measure for each binary comparison), an Error-Correcting Output Code approach (ECOC, [13]) was considered, inspired from [14]. In this work, three classes were considered (i.e. faces, buildings and animals, $N = 3$), leading to codewords of length three. The codewords were defined in terms of probabilities obtained from each GP binary classification (table I), and the distance was computed as the sum of the absolute differences between the class-specific codewords and the probabilities obtained from each binary classifier. The significance of the results was then assessed using permutations of the training labels (100 permutations per block, i.e. 900 permutations in total).

5) Rest sessions: The model built on the mental imagery session was then applied to the rest sessions. In this case, the ECOC approach was not used to associate a label to each scan but to assess the 'confidence' of the prediction by taking the distance between the two most probable classes (in terms of distances to the table, [13]). A unique measure,

[1] www.fil.ion.ucl.ac.uk/spm

[2] LIBSVM, Chang C. C. and Lin, C. J., http://www.csie.ntu.edu.tw/~cjlin/libsvm/, interfaced in PROBID (A. Marquand and J. Mourao-Miranda, http://www.brainmap.co.uk/)

[3] GPML toolbox, C.E. Rasmussen and C.K.I. Williams, http://www.gaussianprocess.org/gpml/, interfaced in PROBID

Table I
ECOC PROBABILISTIC CODEWORDS FOR FACES (F), BUILDINGS (B) AND ANIMALS (A) CATEGORIES. THE 'TEST' LINE REPRESENTS THE OUTPUTS OF THE BINARY CLASSIFIERS FOR AN EXAMPLE TEST POINT, ITS DISTANCE TO EACH CLASS CODEWORD (D_i) BEING COMPUTED IN THE LAST COLUMN.

Class	F vs B	F vs A	B vs A	D_i
F	1	1	0.5	1.4
B	0	0.5	1	0.8
A	0.5	0	0	1.2
test	0.5	0.4	0.8	Class B

referred to as 'L', is therefore attributed to each scan and is computed as:

$$L = min(|D_i - D_j|), i, j = 1..N, i \neq j \tag{1}$$

In table I, $L = |1.2 - 0.8| = 0.4$. The base-level of this L measure was then computed using 1000 permutations of the training labels. For each permutation, the maximum of L was considered, which allowed comparing the L value of each scan to the 1000 L values of the permutations and thereby associating a p-value to each scan. The proportion of scans linked to the memory task was then computed as the percentage of scans for which the associated p-value $<$0.05. For each subject, four values were obtained (one per rest session), referred to as *Pr*. The subjects were then sorted in descending order according to Pr($R2_m$)$-$ Pr($R1_m$) and correlations were computed between Pr($R2_m$)$-$ Pr($R1_m$) and the subjects' behavioural performance. To assess the significance of the correlation, non-parametric testing was performed by permuting 1000 times the behavioural performance and computing the correlations between the shuffled subjects' behavioural performance and both Pr($R2_m$)$-$Pr($R1_m$) and Pr($R2_o$)$-$Pr($R1_o$) (correlations referred to as 'C_m' and 'C_o', respectively). The difference between the two correlation values was also calculated at each permutation. P-values could then be associated to C_m, C_o and their difference.

III. RESULTS

A. Behavioural data

The screening revealed that subjects S4 and S8 were statistical outliers regarding anxiety and depression.

During mental imagery, the number of extracted events and their corresponding duration were variable depending on the volunteer's ability to retrieve the different images forming the requested mental path. `Dprime` measures [15] were computed from the memory test led outside the scanner to assess the behavioural performance of each subject. The results revealed that subject S6 was a statistical outlier regarding Dprime.

Table II
CORRELATION COEFFICIENT C BETWEEN PR($R2_{m/o}$)-PR($R1_{m/o}$) AND THE SUBJECTS' BEHAVIOURAL PERFORMANCE, AND THEIR P-VALUES. P(DIF) REPRESENTS P($C_m > C_o$).

Selection	C_m	p(C_m)	C_o	p(C_o)	p(dif)
All subjects	-0.0565	0.5880	0.0024	0.5060	0.3970
No outliers	0.4968	0.0580	-0.0137	0.5040	0.0400

B. Mental imagery

1) Feature selection: The selected voxels were mostly comprised in the ventral visual path (primary areas, Fusiform Face Area), parietal regions linked to spatial features and hippocampus related to navigation. Activation in these areas represented properly the different aspects of the task.

2) Classification: The optimal subsets of features defined by GLM and RFA were associated with mean balanced accuracies ranging from 32.98 to 69.78 % (figure 2). Non-significant results were found for subjects S10, S11 and S13 (p>0.05). Significant correlations were found between the number of events in the buildings and faces category and the corresponding class accuracy. Trends indicated a relationship between the behavioural performances of the subjects and the performance of the classifier.

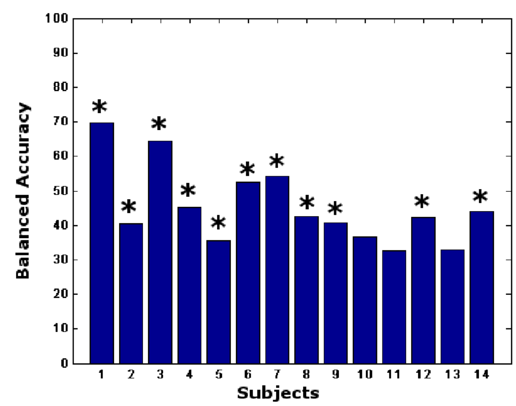

Figure 2. Balanced accuracy obtained from classifying the mental imagery session. Significant results are marked by a *. Please note that due to imbalances in the number of events, the chance level is different from 33% and varies across subjects.

C. Rest sessions

The values for C_m, C_o, p(C_m), p(C_o) and p($C_m > C_o$) are displayed in table II, when considering all subjects and when discarding the statistical outliers (i.e. S4, S6 and S8) in terms of behaviour ('No outliers').

Results show significant C_m value when discarding the statistical outliers in terms of behaviour. On the contrary, C_o values are close to zero with non-significant associated p-values. The difference between the correlations obtained from the 2 sessions were also significant for the second subset of subjects. These results suggest that Pr($R2_m$)−Pr($R1_m$) correlated significantly with the subjects' behavioural performance and that this correlation is significantly higher than when considering the control task.

IV. Discussion

In this paper, the authors applied machine learning based MVPA to two complex fMRI datasets: a mental imagery session characterized by imbalanced and self-paced trials and rest sessions, during which no external or internal stimulation was imposed. Based on [10], the combination of a univariate and a multivariate feature selection followed by Gaussian Processes Classification led to a significant classification of the mental imagery session for 11 subjects out of 14. The obtained models were then applied to the different rest sessions which resulted in the computation of the proportion of scans of spontaneous brain activity linked to the memory task (Pr).

We observed that the results from both classifications correlated significantly with behavioural data. For mental imagery, it is the number of events in each category which mostly affects the class accuracies. For spontaneous brain activity however, it was the performance of the subjects which correlated significantly with the results of most subjects (i.e. when discarding statistical outliers in terms of behaviour and performance). Moreover, this correlation could not be achieved when considering permutations of the behavioural measure and was significantly higher than the correlation value obtained from a control task (C_o). This result suggests that the larger the difference between the proportions of spontaneous brain activity linked to a task before and after this task (i.e. Pr($R2_m$)−Pr($R1_m$)), the better the memorization of task features by the subject. However, investigations on the subjects not verifying the hypothesis (i.e. the statistical outliers) should be performed. In particular, the effect of thresholding at $p < 0.05$ to determine the proportions Pr should be investigated.

V. Conclusion

The classification of rest sessions could be performed by applying previously built models on a mental imagery session. While the results should be more deeply investigated, a significant correlation between the proportions of spontaneous brain activity linked to the memory task and the subjects' performance to memorize the task features suggests that reactivations during post-experience rest are linked to the memorization of the conducted experiment.

Acknowledgment

This study was supported by the Belgian Fonds National de la Recherche Scientifique (F.R.S.-F.N.R.S., F.R.I.A.), the Fondation Reine Elisabeth, the Research Fund of ULg, the PAI/IAP Interuniversity Pole of Attraction and the BIAL Foundation.

References

[1] J.-D. Haynes and G. Rees, "Decoding mental states from brain activity in humans." *Nature Review Neuroscience*, vol. 7, no. 7, pp. 523–534, Jul 2006.

[2] C. Burges, "A tutorial on support vector machines for pattern recognition." *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, 1998.

[3] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, ser. Adaptive Computation and Machine Learning, T. Dietterich, Ed. the MIT Press, 2006.

[4] J. Haynes and G. Rees, "Predicting the orientation of invisible stimuli from activity in human primary visual cortex." *Nature Neuroscience*, vol. 8, pp. 686–691, 2005.

[5] M. Chadwick, D. Hassabis, N. Weiskopf, and E. Maguire, "Decoding individual episodic memory traces in the human hippocampus." *Current Biology*, vol. 20, pp. 1–4, 2010.

[6] J.-D. Haynes, K. Sakai, G. Rees, S. Gilbert, C. Frith, and R. E. Passingham, "Reading hidden intentions in the human brain." *Current Biology*, vol. 17, no. 4, pp. 323–328, Feb 2007.

[7] A. Marquand, M. Howard, M. Brammer, C. Chu, S. Coen, and J. Mourao-Miranda, "Quantitative prediction of subjective pain intensity from whole-brain fmri data using gaussian processes," *NeuroImage*, vol. 49, no. 3, pp. 2178 – 2189, 2010.

[8] K. Hoffman and B. McNaughton, "Coordinated reactivation of distributed memory traces in primate neocortex." *Science*, vol. 297, pp. 2070–2073, 2002.

[9] A. Tambini, N. Ketz, and L. Davachi, "Enhanced brain correlations during rest are related to memory for recent experiences," *Neuron*, vol. 65, pp. 280–290, 2010.

[10] J. Schrouff, C. Kussé, L. Wehenkel, P. Maquet, and C. Phillips, "Decoding semi-constrained brain activity from fMRI using Support Vector Machines and Gaussian Processes," *PLoS one*, vol. 7, p. e35860, 2012.

[11] K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny, *Statistical Parametric Mapping: the analysis of functional brain images*. London: Elsevier Academic Press, 2007.

[12] R. Frackowiak, K. Friston, C. Frith, C. Price, S. Zeki, J. Ashburner, and W. Penny, *Humain brain function*. New York: Elsevier Academic Press, 2004.

[13] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problem via error-correcting output codes," *Journal of Artificial Intelligence Research*, vol. 2, pp. 263–286, 1995.

[14] D. Hassabis, C. Chu, G. Rees, N. Weiskopf, P. Molyneux, and E. Maguire, "Decoding neural ensembles in the human hippocampus," *Current Biology*, vol. 19, pp. 546–554, 2009.

[15] D. Green and J. Swets, *Signal detection theory and psychophysics*. New York: Wiley, 1966.