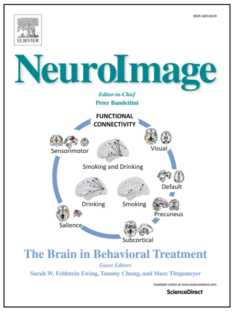
Accepted Manuscript



Decoding fMRI activity in the time domain improves classification performance

João Loula, Gaël Varoquaux, Bertrand Thirion

PII: S1053-8119(17)30665-1

DOI: 10.1016/j.neuroimage.2017.08.018

Reference: YNIMG 14249

To appear in: *NeuroImage*

Received Date: 15 March 2017

Revised Date: 1053-8119 1053-8119

Accepted Date: 4 August 2017

Please cite this article as: Loula, Joã., Varoquaux, Gaë., Thirion, B., Decoding fMRI activity in the time domain improves classification performance, *NeuroImage* (2017), doi: 10.1016/j.neuroimage.2017.08.018.

Decoding fMRI activity in the time domain improves classification performance

João Loula[a,1], Gaël Varoquaux[a], Bertrand Thirion[a]

[a]Parietal Team - Inria / CEA - Paris Saclay University, France
[b]Department of Computer Science - École Polytechnique, France

Abstract

Most current functional Magnetic Resonance Imaging (fMRI) decoding analyses rely on statistical summaries of the data resulting from a deconvolution approach: each stimulation event is associated with a brain response. This standard approach leads to simple learning procedures, yet it is ill-suited for decoding events with short inter-stimulus intervals. In order to overcome this issue, we propose a novel framework that separates the spatial and temporal components of the prediction by decoding the fMRI time-series continuously, i.e. scan-by-scan. The stimulation events can then be identified through a deconvolution of the reconstructed time series. We show that this model performs as well as or better than standard approaches across several datasets, most notably in regimes with small inter-stimuli intervals (3 to 5s), while also offering predictions that are highly interpretable in the time domain. This opens the way toward analyzing datasets not normally thought of as suitable for decoding and makes it possible to run decoding on studies with reduced scan time.

Keywords: Functional magnetic resonance imaging, Classification analysis, MVPA, Decoding, Rapid event-related design

1. Introduction

The application of multivariate analysis techniques to fMRI datasets, aka *decoding*, has become a popular approach to probe the relationships between stimuli and brain activity [20, 24, 14]. The very nature of fMRI data makes it a challenging problem: relatively few samples (events or blocks corresponding to stimulus presentation) are available, in comparison with the high dimensionality –number of voxels– of each observation. This mismatch leads to the so-called *curse of dimensionality*: learning distributed patterns from few samples is a hard problem. The power of high-dimensional regression methods is thus needed to achieve high accuracy and return an interpretable discriminative pattern (see e.g. [5]). However, the sluggishness of the Blood-Oxygen-Level-dependent (BOLD) response observed in fMRI implies that the occurrence of brain activity is not synchronous with the presentation of stimuli, but delayed by approximately 6s and smooth in time [9]. For the sake of statistical analysis, a preliminary regression step is thus typically performed, so that pairs of stimulus events and associated brain response can be considered. This prior regression is simply carried out by the traditional General Linear Model (GLM) used in standard statistical analyses of fMRI [8].

Although it is the standard solution used by nearly all practitioners, this two-step approach is not optimal; in particular, the intermediate event-related brain response estimates are very noisy, limiting decoding accuracy. The reason is that, unlike traditional brain mapping settings in which all events from one condition end up being one single regressor, for decoding purpose, events are split into different regressors, resulting in a loss in design efficiency and high-variance estimates. This approach is also bound to perform poorly on event-related tasks using small inter-stimuli inter-