Neural Networks 70 (2015) 61–73

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Non-parametric temporal modeling of the hemodynamic response function via a liquid state machine

Paolo Avesani [a,b], Hananel Hazan [c,e], Ester Koilis [c], Larry M. Manevitz [c,*], Diego Sona [a,d]

[a] NeuroInformatics Laboratory (NILab), Fondazione Bruno Kessler, Trento, Italy
[b] Centro Interdipartimentale Mente e Cervello (CIMeC), Università di Trento, Italy
[c] Department of Computer Science, University of Haifa, Israel
[d] Pattern Analysis and Computer Vision, Istituto Italiano di Tecnologia, Genova, Italy
[e] Network Biology Research Laboratory, Technion, Haifa, Israel

HIGHLIGHTS

- A model for HRF learned directly from data not based on a priori assumptions.
- This allows the BOLD signal to be personalized and voxel specific.
- Voxels can be filtered as "relevant" based on their predictive ability.
- Temporal info stored in reservoir computing; trained feed forward NN produces HRF.
- Learning process robust both to noise and underlying shape of the "true" HRF signal.

ARTICLE INFO

Article history:
Received 7 October 2014
Received in revised form 26 April 2015
Accepted 27 April 2015
Available online 21 May 2015

Keywords:
HRF modeling
Temporal modeling
fMRI
Machine learning
Neural networks
Reservoir computing
Liquid state machines

ABSTRACT

Standard methods for the analysis of functional MRI data strongly rely on prior implicit and explicit hypotheses made to simplify the analysis. In this work the attention is focused on two such commonly accepted hypotheses: (i) the hemodynamic response function (HRF) to be searched in the BOLD signal can be described by a specific parametric model e.g., double-gamma; (ii) the effect of stimuli on the signal is taken to be linearly additive. While these assumptions have been empirically proven to generate high sensitivity for statistical methods, they also limit the identification of relevant voxels to what is already postulated in the signal, thus not allowing the discovery of unknown correlates in the data due to the presence of unexpected hemodynamics. This paper tries to overcome these limitations by proposing a method wherein the HRF is learned directly from data rather than induced from its basic form assumed in advance. This approach produces a set of voxel-wise models of HRF and, as a result, relevant voxels are filterable according to the accuracy of their prediction in a machine learning framework.

This approach is instantiated using a temporal architecture based on the paradigm of Reservoir Computing wherein a Liquid State Machine is combined with a decoding Feed-Forward Neural Network. This splits the modeling into two parts: first a representation of the complex temporal reactivity of the hemodynamic response is determined by a universal global "reservoir" which is essentially temporal; second an interpretation of the encoded representation is determined by a standard feed-forward neural network, which is trained by the data. Thus the reservoir models the temporal state of information during and following temporal stimuli in a feed-back system, while the neural network "translates" this data to fit the specific HRF response as given, e.g. by BOLD signal measurements in fMRI.

An empirical analysis on synthetic datasets shows that the learning process can be robust both to noise and to the varying shape of the underlying HRF. A similar investigation on real fMRI datasets provides evidence that BOLD predictability allows for discrimination between relevant and irrelevant voxels for a given set of stimuli.

1. Introduction

Modeling the HRF signal is a crucial prerequisite for the analysis of fMRI data. One needs to know what is expected a priori from a

* Corresponding author.
E-mail address: manevitz@cs.haifa.ac.il (L.M. Manevitz).

http://dx.doi.org/10.1016/j.neunet.2015.04.009

given stimulus that affects a voxel in order to, e.g., compare the actual readings and thereby be able to decide if the voxel is relevant to the stimulus or if it is pertinent to differentiate between contrasting stimuli. Most accepted techniques for accomplishing this task involve a mathematical model of the HRF signal. For example, this is the typical assumption in the basic GLM analysis wherein the effects of stimuli are thought to be linearly combined; or the somewhat more advanced methods wherein the time relationship is expanded by a temporal basis function, as in FIR (Goutte, Nielsen, & Hansen, 2000) or Fourier basis functions (Josephs, Turner, & Friston, 1997). In addition, the linear combination can be relaxed (Henson & Friston, 2007) by, e.g., using a Volterra expansion (Friston, Josephs, Rees, & Turner, 1998; Josephs & Henson, 1999) to allow higher order interactions.

In this study, building on previous work (Avesani, Hazan, Koilis, Manevitz, & Sona, 2011) a somewhat different direction is taken, using a temporal neural network approach. The basic idea is that a canonical (at least in the sense of brain modeling) recurrent model reflects the temporal interactions and the information retained over time; and then a second feed-forward network maps this into the readings of the BOLD signal. This method has several conceptual advantages: (i) by separating the feedback effects, the model acts as a Markovian system; (ii) the same recurrent framework can be used for all potential voxels and the difference in response of the voxel resides in a simple feed-forward network; (iii) the parameters of the model (parallel e.g. to "hyperparameters" (Henson & Friston, 2007)) are not related to a specific task but are chosen by general observations (Hazan & Manevitz, 2012). Thus the recurrence can in fact be considered a constant mechanism and is not learned; (iv) the parameters of the feed-forward model can be estimated directly from data.[1]

The application of this computational paradigm allows the use of the predictive success of the model to separate observed voxels into "relevant" and "non-relevant" which, in turn and, in principle, can be used to make studies of differentiation of voxels for specific stimulus related tasks.

1.1. Characterizing the information in the BOLD signal

Functional Magnetic Resonance Imaging (fMRI) is a widely used modality in the studies of brain perception and cognition, with applications to a broad variety of neuroscientific questions including brain mapping, which corresponds to the set of data analysis methods designed to detect and map brain areas relevant to specific cognitive or perceptual tasks. The fMRI experiments are usually designed by contrasting categories of stimuli (e.g., visual representation of faces versus houses) and analyzing the recorded data to find brain areas related to the classes of stimuli. The underlying assumption is that the brain areas allowing for discrimination between the contrasting categories are related to the corresponding cognitive or perceptual task.

The existing brain mapping methods can be divided into two categories: the hypothesis-driven methods and the data-driven methods. While the former methods use different prior assumptions on the basic characteristics of the BOLD (Blood Oxygenation Level Dependent) signal, the later class of methods, usually having an exploratory nature, tries to infer these characteristics directly from the data.

The hypothesis-driven methods for the fMRI data analysis are used in the majority of brain mapping studies. A survey (Grinbald, Wager, Lindquist, & Hirsch, 2008) of 170 fMRI studies shows that 96% of experiments were based on the hypothesis-driven analysis methods. These methods have a strong statistical framework for assessing the relevant regional activation areas. However, they rely on prior assumptions on the BOLD signal underlying the real brain activity. Given a stimulation protocol, prior knowledge enables the definition of the expected BOLD response as a parametric Hemodynamics Response Function (HRF), which is used in a General Linear Model (GLM) framework (Friston et al., 1994; Monti, 2011). The outcome is based on the univariate analysis of the correlation between the real signal and the estimated HRF. These methods require, therefore, an accurate definition of the expected HRF shape, although it is allowed to vary significantly in different populations, between subjects, and between different brain areas (Aguirre, Zarahn, & D'Esposito, 1998).

There are recent and sophisticated parametric HRF models that attempt to better capture the complex structure of the BOLD response (Zheng et al., 2002). Nonetheless, these models still encode some ideal BOLD shape without considering possible uncommon or irregular hemodynamics due both to the brain structure (e.g., proximity to a large vascular vessel or changes caused by different brain injuries) and to the kind of cognitive tasks under investigation.

To avoid these issues, one can try to obtain a correct HRF voxel by voxel with a data-driven approach, thereby finding the unknown functional dependencies between the BOLD signal and the known set of stimuli, without any prior assumption on the expected HRF. Some data-driven methods are reported in the literature, such as selective averaging with a long inter-stimulus interval assuming non-overlapping responses (Bandettini & Cox, 2000; Buckner et al., 1996) or the FIR methodology (Goutte et al., 2000). Other methods, like Bayesian approaches (Woolrich, Jenkinson, Brady, & Smith, 2004) or wavelet deconvolution (Wink, Hoogduin, & Roerdink, 2008) are computationally expensive or require some additional prior assumptions (e.g., separability of signal and noise in the frequency domain for the wavelet methods).

A promising approach in this direction is the FIR approach (Goutte et al., 2000; Henson & Friston, 2007). In FIR the HRF is modeled as a linear combination of "impulse basis functions", a set of adjacent boxcar functions, specified over the period of time fitting the expected duration of BOLD response. Then the corresponding coefficients are calculated from the data. The combination of FIR basis functions can capture any shape of response up to a given timescale (Henson & Friston, 2007). However, the precision of FIR model depends on a correct specification of the expected duration of hemodynamic response. Moreover, the FIR model is linear and it can suffer low sensitivity with experimental protocols generating non-linear effects in the HRF.

In another approach, Wang (2009) proposed a method for brain mapping based on machine learning techniques. This method is a combination of the HRF data-driven analysis and the hypothesis-driven GLM inference. In this method, the HRF profile extracted from the Support Vector Machine (SVM) classifier (Burges, 1998; Vapnik, 1995) is used as the data-driven regressor for the consequent GLM analysis. Note that SVM is a machine learning technique that has been demonstrated to be successful for the analysis of neuroimaging data in many applications (Atir-Sharon, Gilboa, Hazan, Koilis, & Manevitz, 2015; Boehm, Hardoon, & Manevitz, 2011; Cox & Savoy, 2003; Hardoon & Manevitz, 2005; Mitchell et al., 2004; Mourão-Miranda, Bokde, Born, Hampel, & Stetter, 2005) including, in particular, brain-decoding using a framework referred to as multivariate pattern analysis (MVPA). In this approach one tries to *classify the contrasted stimuli* from the existing BOLD signal. However, the HRF derivation under discussion here is a reverse task, compatible with a standard brain-mapping analysis, in which one wishes to *reconstruct the expected BOLD signal*based on the stimuli sequence.

[1] The architecture of the feed-forward model can be considered a separate parameter as in most neural network research. However, we neglect this issue because it is consistent with the model to replace the feed-forward network with any other machine learning scheme. See below.