NeuroImage 122 (2015) 131–148

Contents lists available at ScienceDirect

NeuroImage

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

Relevant feature set estimation with a knock-out strategy and random forests

Melanie Ganz [a,b,c], Douglas N. Greve [b,c], Bruce Fischl [b,c,d], Ender Konukoglu [b,c,*], for the, Alzheimer's Disease Neuroimaging Initiative [1]

[a] Neurobiology Research Unit and Center for Integrated Molecular Brain Imaging, Rigshospitalet, Copenhagen, Denmark
[b] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Boston, MA, USA
[c] Harvard Medical School, Boston, MA, USA
[d] Computer Science and AI Lab/Division of Health Sciences and Technology, Massachusetts Institute for Technology, Boston, MA, USA

ARTICLE INFO

Article history:
Received 16 February 2015
Accepted 3 August 2015
Available online 10 August 2015

Keywords:
Multivariate pattern analysis
Interpretability
Relevant features
Random forests
Knock-out

ABSTRACT

Group analysis of neuroimaging data is a vital tool for identifying anatomical and functional variations related to diseases as well as normal biological processes. The analyses are often performed on a large number of highly correlated measurements using a relatively smaller number of samples. Despite the correlation structure, the most widely used approach is to analyze the data using univariate methods followed by post-hoc corrections that try to account for the data's multivariate nature. Although widely used, this approach may fail to recover from the adverse effects of the initial analysis when local effects are not strong. Multivariate pattern analysis (MVPA) is a powerful alternative to the univariate approach for identifying relevant variations. Jointly analyzing all the measures, MVPA techniques can detect global effects even when individual local effects are too weak to detect with univariate analysis. Current approaches are successful in identifying variations that yield highly predictive and compact models. However, they suffer from lessened sensitivity and instabilities in identification of relevant variations. Furthermore, current methods' user-defined parameters are often unintuitive and difficult to determine. In this article, we propose a novel MVPA method for group analysis of high-dimensional data that overcomes the drawbacks of the current techniques. Our approach explicitly aims to identify all relevant variations using a "knock-out" strategy and the Random Forest algorithm. In evaluations with synthetic datasets the proposed method achieved substantially higher sensitivity and accuracy than the state-of-the-art MVPA methods, and outperformed the univariate approach when the effect size is low. In experiments with real datasets the proposed method identified regions beyond the univariate approach, while other MVPA methods failed to replicate the univariate results. More importantly, in a reproducibility study with the well-known ADNI dataset the proposed method yielded higher stability and power than the univariate approach.

1. Introduction

In the study of psychiatric disorders and neurological diseases one of the fundamental questions is: which regions of the brain does the condition affect? This question is common in many neuroimaging studies, and the approach to answer it is to statistically analyze images acquired from a cohort of subjects to detect anatomical (and/or functional) variations related to the condition. The statistical analysis, referred to as *group analysis*, is often performed on densely extracted anatomical measurements, such as cortical thickness maps or gray matter densities. Such sets of measurements have a complex correlation structure none the least due to the spatial organization of the anatomical locations they are extracted from. Statistical methods that can leverage the correlation structure to improve the power of group analysis are of great interest. To this end, this article proposes a novel multivariate method that overcomes the major limitations of current algorithms.

The most commonly used statistical tool to identify group effects is still mass-univariate analysis (Friston et al., 1994; Ashburner and Friston, 2000). Univariate analysis tests each measurement independently for its statistical relationship with the condition of interest. Although useful and intuitive, univariate analysis ignores the correlation structure in the data. Hence, for problems with a very high number of measurements, which is typical for neuroimaging, the multiple comparison problem becomes critical. Univariate analysis results frequently do not survive family-wise error control such as

* Corresponding author at: Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Boston, MA, USA.
E-mail addresses: melanie.ganz@nru.dk (M. Ganz), enderk@nmr.mgh.harvard.edu, ender.konukoglu@gmail.com (E. Konukoglu).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

http://dx.doi.org/10.1016/j.neuroimage.2015.08.006
1053-8119/

Bonferroni correction (Bonferroni, 1935), and if the affected regions are small, they might not even survive false-discovery rate control (Benjamini and Hochberg, 1995; Storey and Tibshirani, 2003). To account for this, post-processing techniques such as Worsley et al. (1996), Andrade et al. (2001), Hagler et al. (2006), and Smith and Nichols (2009) attempt to integrate the multivariate information by assuming that "real" effects would be visible in larger spatial neighborhoods. However, these post-hoc corrections are applied to the results of the initial univariate analysis and hence propagate its adverse effects.

The natural way to analyze a set of measurements that has a correlation structure is to analyze the measurements jointly. Multivariate pattern analysis (MVPA) techniques provide the means to do this. Each measurement is considered as a *feature*, and detecting condition-related variation is formulated as identifying the subset of features that are useful in *predicting* the condition. In contrast to the theory of univariate analysis, multivariate analysis for neuroimaging is still an active field of research. Researchers have proposed a variety of techniques borrowing ideas from the machine learning literature (De Martino et al., 2008), (Mourao-Miranda et al., 2005; Kriegeskorte et al., 2006; Menze et al., 2009; Sabuncu and Van Leemput, 2012; Langs et al., 2011; Yamashita et al., 2008; Gaonkar and Davatzikos, 2012; Rondina et al., 2014) (We refer the reader to the recent review (Mwangi et al., 2014) for a more complete list.). However, direct applications of these methods have one major drawback. Most of the above methods aim to identify the set of features that yields the highest prediction accuracy, but not necessarily the complete set of relevant features with which the condition of interest can be predicted. As a consequence, the detection results of current MVPA methods are often not exhaustive nor reproducible (Rasmussen et al., 2012), and typically differ substantially from the regions detected by univariate analysis.

In this article we propose a new method within the MVPA framework that aims to explicitly identify all the relevant features, i.e. all the measurements that display condition-related variation and hence have the ability to predict the condition or effect. The main principle of our method relies on the observation that for any learning algorithm, if the identified set of relevant features is not exhaustive, the learning algorithm would still have been able to predict the label if these features were absent in the first place (Konukoglu et al., 2013a). Based on this observation we build an iterative algorithm, where the main idea is to iteratively detect and *knock-out* sets of relevant features using a learning method. The main advantage of this approach compared to existing literature is that it is designed to construct an exhaustive set of relevant features and not directly maximize the prediction accuracy. This is contrary to most previous feature selection techniques in the machine learning literature. The proposed method is a wrapper algorithm that encapsulates a predictive method, which is chosen to be the Random Forest algorithm (Amit and Geman, 1997; Breiman, 2001). The design of the wrapper is such that it iterates around the predictive model, iteratively taking out the detected relevant features, until the predictive model is statistically not better than random guessing. We also take advantage of the recent developments in the theory of feature selection of Random Forest (Konukoglu and Ganz, 2014). These advancements allow us to make the tuning parameters of our algorithm *intuitive*, a characteristic missing in other MVPA methods.

We tested the proposed method on both synthetic and real datasets, and compared the results with other MVPA algorithms as well as univariate analysis. In the experiments with synthetic datasets we evaluated the performance of all algorithms by comparing the relevant feature sets identified by each algorithm with the ground truth sets using DICE score and sensitivity. Additionally, we also evaluated the quality of the identified sets in a condition-prediction experiment. In the real data experiments, we qualitatively compared identified features for different algorithms on four different datasets (one included in the main article and three in the supplementary materials). Furthermore, we studied the reproducibility of feature identification using the proposed method and the univariate analysis on the ADNI dataset. Lastly, the proposed knock-out strategy is a generic wrapper algorithm with which any MVPA method can be used. To illustrate the advantages of using Random Forests, we experimented with using LASSO (Tibshirani, 1996) within the knock-out strategy.

The rest of the article is structured as follows. We first present an overview of the related work on multivariate methods in Section 2. Next, we detail the proposed algorithm in Section 3. Additionally, we introduce a multiple comparison correction technique for it in Section 5. Then, we describe our experimental methodology in Section 6, present the results in Section 7 and discuss them. We conclude the article in Section 8.

2. Related work on multivariate methods

Earlier multivariate methods in neuroimaging focused mainly on applications in functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). The most popular amongst these are partial least square correlations (McIntosh et al., 1996; McIntosh and Lobaugh, 2004), (Krishnan et al., 2011), canonical variant analysis (Friston, 1997), (Friston et al., 1995) and multivariate linear modeling (Worsley et al., 1997). The common idea is to use linear dimensionality reduction to find the directions in the feature space that show the highest correlation with the condition. Each measurement gets assigned a weight indicating its contribution to the strength of the correlation with the condition, relative to the other measurements. Although higher weights suggest stronger condition-related effects, it is not obvious how to set a threshold to separate affected from non-affected regions. Bootstrapping (Krishnan et al., 2011) can provide a partial solution to this by quantifying the stability of weights, but it does not mitigate the relative assignment problem.

More recent work on multivariate analysis focused on the MVPA framework. These methods identify concrete sets of measurements, referred to as relevant features, instead of assigning relative weights. MVPA techniques, can be coarsely divided into local (Kriegeskorte et al., 2006; Zhang and Davatzikos, 2011) and global approaches(De Martino et al., 2008), (Mourao-Miranda et al., 2005; Yamashita et al., 2008; Sabuncu and Van Leemput, 2012; Langs et al., 2011; Rondina et al., 2014; Menze et al., 2009; Gaonkar and Davatzikos, 2012; Rasmussen et al., 2012; Rondina et al., 2013; Haufe et al., 2014; Konukoglu et al., 2013a). Local techniques extend univariate analysis by taking into account the neighborhood of a feature when detecting its relevance to the condition. While the statistical analyses are multivariate, based on Mahalanobis distance in (Kriegeskorte et al., 2006) and optimal filtering through nonnegative discriminative projection in (Zhang and Davatzikos, 2011), they are confined to small neighborhoods around each feature location. Although both approaches are interesting, they only explore local relationships and do not account for long distance spatially distributed patterns.

Global MVPA approaches take into account the entire set of measurements at once and are able to capture spatially distributed patterns. However, the main problem with current predictive modeling approaches is that they aim to identify the subset of features that yields the highest prediction accuracy. The remaining features get discarded, even though they might include features that are also informative. This strategy might be ideal to derive accurate and compact predictive models, however it does not guarantee feature exhaustivity nor reproducibility. But both of these aspects are important for detecting condition-related anatomical variations. This point has also been made by Rasmussen et al. in Rasmussen et al. (2012) and Rodina et al. in Rondina et al. (2013). In fact, in Rasmussen et al. (2012) the authors even demonstrated a trade-off between reproducibility and prediction accuracy, though without providing a solution. We believe this problem is the main reason why current predictive models do not produce results that are comparable to univariate analysis.