NeuroImage 181 (2018) 521–538

Contents lists available at ScienceDirect

NeuroImage

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

Reconstructing subject-specific effect maps

Ender Konukoglu [a,*], Ben Glocker [b], for the Alzheimer's Disease Neuroimaging Initiative[1]

[a] Computer Vision Lab, ETH Zurich, Zurich, Switzerland
[b] Department of Computing, Imperial College London, London, United Kingdom

ABSTRACT

Predictive models allow subject-specific inference when analyzing disease related alterations in neuroimaging data. Given a subject's data, inference can be made at two levels: global, i.e. identifiying condition presence for the subject, and local, i.e. detecting condition effect on each individual measurement extracted from the subject's data. While global inference is widely used, local inference, which can be used to form subject-specific effect maps, is rarely used because existing models often yield noisy detections composed of dispersed isolated islands. In this article, we propose a reconstruction method, named RSM, to improve subject-specific detections of predictive modeling approaches and in particular, binary classifiers. RSM specifically aims to reduce noise due to sampling error associated with using a finite sample of examples to train classifiers. The proposed method is a wrapper-type algorithm that can be used with different binary classifiers in a diagnostic manner, i.e. without information on condition presence. Reconstruction is posed as a Maximum-A-Posteriori problem with a prior model whose parameters are estimated from training data in a classifier-specific fashion. Experimental evaluation is performed on synthetically generated data and data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Results on synthetic data demonstrate that using RSM yields higher detection accuracy compared to using models directly or with bootstrap averaging. Analyses on the ADNI dataset show that RSM can also improve correlation between subject-specific detections in cortical thickness data and non-imaging markers of Alzheimer's Disease (AD), such as the Mini Mental State Examination Score and Cerebrospinal Fluid amyloid-β levels. Further reliability studies on the longitudinal ADNI dataset show improvement on detection reliability when RSM is used.

1. Introduction

Statistical analysis methods for neuroimaging data are instrumental in detecting condition induced structural alterations. Available methods can process high number of measurements with complex spatial correlation and construct *effect maps*, for example, in the form of detailed volumetric (Ashburner and Friston, 2001) or surface-based (Greve, 2011; Fischl, 2012) maps that highlight changes statistically related to the condition. Effect maps are often used at the population level, where at each measurement, maps indicate statistical relationship between the condition and measurement across the entire population. Either a group analysis technique (Ashburner and Friston, 2001; Krishnanet al., 2011; Worsleyet al., 1997) or a machine learning based predictive model (Arbabshirani et al. Calhoun; Gaonkar and Davatzikos, 2013; Mwangi et al., 2014; Rahimet al., 2015; Ganzet al., 2015) is used to compare two cohorts of subjects, one showing the condition of interest and the other not, and estimate relationships. Population-wide effect maps constructed with existing methods have already provided valuable information on anatomical footprints of various diseases, e.g. (Thompsonet al., 2001a; Rosaset al., 2002; Burtonet al., 2004), lifestyle choices, e.g. (Garrido et al., 1993; Miller and OCallaghan, 2003; Kanai and Rees, 2011), as well as genetics and inherited traits, e.g. (Watkins et al., 2002; Peper et al., 2007; Thompson et al., 2001b).

Information provided in population-wide effect maps is useful, however, not *subject-specific*. When we consider a measurement extracted from a specific subject, population-wide effect maps do not tell us whether the measurement shows disease effect. Therefore, possible analyses on the extracted measurements are limited to population-wide questions. In order to perform subject-specific analyses, methods that can detect subject-specific effects and construct corresponding maps are needed. Furthermore, methods that can do this diagnostically, i.e. without having information on the presence of the condition for the subject, would be highly desirable. Constructing subject-specific effect maps in a diagnostic manner can have multiple applications. In clinical and neuroscience research, subject-specific detections can be used for stratification and identification of subpopulations (Iqbal, 2005). In engineering research, machine learning tools are often “black-box” components. Subject-specific maps can facilitate model improvement by

* Corresponding author.
E-mail addresses: ender.konukoglu@vision.ee.ethz.ch (E. Konukoglu), b.glocker@imperial.ac.uk (B. Glocker).

[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 completeblack listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

https://doi.org/10.1016/j.neuroimage.2018.07.032
Received 2 January 2018; Received in revised form 7 June 2018; Accepted 12 July 2018
Available online 23 July 2018
1053-8119/

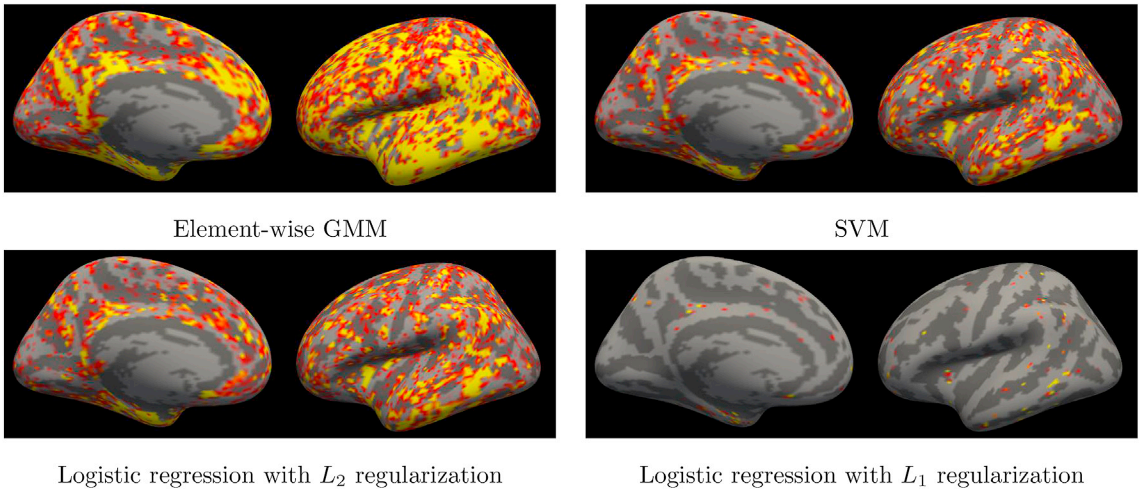

Fig. 1. Subject-specific effect maps of Alzheimer's Disease (AD) extracted from cortical thickness map of a patient with AD using different binary classifiers. We used the ADNI dataset to train the binary classifiers to distinguish between AD patients from healthy elderly using cortical thickness maps of the left-hemisphere extracted using Freesurfer from T1-weighted Magnetic Resonance Images (MRI) (further details on this dataset and the experiments are provided in Section 3). In the figure, we show subject-specific AD-effect maps for a case subject who was not in the training set. Regions highlighted with yellow are locations where algorithms suggest condition effects with yellow indicating highest degree and red lower. No thresholding is applied and for visualization the same colormap is used for all. Underlying is the inflated left-hemisphere surface with sulci and gyri indicated with different gray levels. Maps have numerous isolated islands highlighting areas that are not always associated with AD in the literature. However, they also highlight areas that are known to be associated with AD, such as the medial temporal lobe or entorhinal cortex. This is promising since it suggests that detections have both false and true positives. A reconstruction method that can suppress the former and highlight the latter would yield cleaner and potentially more useful subject-specific maps.

allowing to analyze cases where methods fail. Lastly, for clinical practice, subject-specific detections can help in diagnosis and grading.

There has been previous attempts to detect subject-specific effects by extending group analysis techniques, particularly using one-vs-all analysis (Maumet et al., 2013, 2016). This avenue is promising as the theory developed in group analysis can be applied. The main drawback, however, is the difficulty in applying this approach diagnostically, as condition information for the subject is needed in the analysis.

The main approach for detecting subject-specific effects in a diagnostic fashion is predictive modeling, in particular linear binary classifiers. When trained binary classifiers are applied to a new subject data, algorithms can readily output subject-specific effect maps without algorithmic modification. Despite their availability, these methods are rarely used for this purpose in practice. We believe one of the main reasons for this is that resulting subject-specific effect maps are often “noisy”. Detections form isolated small islands and can be dispersed to areas that may not be involved in the condition. We illustrate this with an example in Fig. 1 in the context of Alzheimer's disease. Alleviating the noise problem can facilitate subject-specific analyses of neuroimaging data.

In this article, we present a reconstruction method, named RSM (**R**econstruction **S**ubject-specific effect **M**aps), for improving subject-specific detections of binary classifiers. A main source of noise in subject-specific detections is sampling error associated with using finite training sets to train classifiers. The proposed method reduces this noise by using a Bayesian formulation with a prior probability model formulated as a Markov Random Field (MRF), whose parameters are estimated from training data, and solving a Maximum-A-Posteriori problem. RSM is a generic wrapper-type algorithm and can be used with various binary classifiers. We demonstrate RSM's use with four different models: element-wise Gaussian mixture models (ew-GMM), Support Vector Machines (SVM) (Cortes and Vapnik, 1995), Logistic Regression with L_2 and L_1 regularization (LR L_2 and LR L_1).

We focus on spatial maps of image-based measurements where local measurements are extracted densely at multiple points from the brain. Examples of such maps are voxel-wise gray matter density (Ashburner and Friston, 2001) and surface-based cortical thickness maps (Fischl, 2012). RSM can also be applied to other types of measurements, such as volumes of multiple anatomical structures, but it is especially designed for high-dimensional measurements with spatial context and makes use of the associated correlation structure. Although our interest is in neuroimaging, the method is not specific to the brain and can be used with other anatomical structures.

We first describe the proposed method in Section 2 and then evaluate it in Section 3. We performed evaluations both with synthetically generated data, where ground truth information is available, and data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), where the goal is to detect structural alterations due to Alzheimer's Disease (AD). We present quantitative results focusing on improvements in detection accuracy due to RSM. On the ADNI dataset, we present a correlation and a reliability study. The former analyzes correlation between subject-specific detections and auxiliary measures such as Mini Mental State Examination scores (MMSE) and Cerebrospinal Fluid amyloid-β (CSF a-β) measurements. The latter evaluates reliability of detections in the longitudinal setting. In both, we focus on the benefits of the proposed method by comparing detections obtained with and without using RSM for the different classifiers. In order to provide a bench-mark we also compare detections with an outlier detection method. We conclude with a discussion in Section 4.

2. Method

The proposed method is a statistical technique to analyze measurements across individuals. In the following we will assume that measurements extracted from different individuals are spatially normalized, which means they are aligned with a common template and corresponding measurements can be directly compared. Such normalization can be achieved, for instance, using publicly available tools, such as SPM and Freesurfer.[2]

2.1. Subject-specific effect maps

We denote with vector $\mathbf{f} = [f_1, \ldots, f_d] \in \mathbb{R}^d$ a set of measurements extracted from a subject's image. Some examples for $\mathbf{f}$ that are widely

[2] see http://www.fil.ion.ucl.ac.uk/spm/or https://freesurfer.net.