NeuroImage 124 (2016) 740–751

Contents lists available at ScienceDirect

NeuroImage

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

Analyzing neuroimaging data with subclasses: A shrinkage approach

Johannes Höhne [a,*], Daniel Bartz [b], Martin N. Hebart [c], Klaus-Robert Müller [b,d,e,f,*], Benjamin Blankertz [a,d,e,*]

[a] Neurotechnology Group, Department of Computer Science, Berlin Institute of Technology, Marchstr. 23, 10587 Berlin, Germany
[b] Machine Learning Group, Department of Computer Science, Berlin Institute of Technology, Marchstr. 23, 10587 Berlin, Germany
[c] Department of Systems Neuroscience, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany
[d] Bernstein Focus Neurotechnology, Berlin, Germany
[e] Bernstein Center for Computational Neuroscience, Berlin, Germany
[f] Department of Brain and Cognitive Engineering, Korea University, Anam-dong, Seongbuk-gu, Seoul 136-713, Republic of Korea

ARTICLE INFO

Article history:
Received 6 July 2015
Accepted 15 September 2015
Available online 25 September 2015

Keywords:
Pattern classification
Linear classifier
Shrinkage
Subclass
EEG
ERP
fMRI
Searchlight
Regularization profile
Data-driven clustering
BCI
Single-trial classification

ABSTRACT

Among the numerous methods used to analyze neuroimaging data, Linear Discriminant Analysis (LDA) is commonly applied for binary classification problems. LDAs popularity derives from its simplicity and its competitive classification performance, which has been reported for various types of neuroimaging data.

Yet the standard LDA approach proves less than optimal for binary classification problems when additional label information (i.e. subclass labels) is present. Subclass labels allow to model structure in the data, which can be used to facilitate the classification task. In this paper, we illustrate how neuroimaging data exhibit subclass labels that may contain valuable information. We also show that the standard LDA classifier is unable to exploit subclass labels.

We introduce a novel method that allows subclass labels to be incorporated efficiently into the classifier. The novel method, which we call Relevance Subclass LDA (RSLDA), computes an individual classification hyperplane for each subclass. It is based on regularized estimators of the subclass mean and uses other subclasses as regularization targets. We demonstrate the applicability and performance of our method on data drawn from two different neuroimaging modalities: (I) EEG data from brain–computer interfacing with event-related potentials, and (II) fMRI data in response to different levels of visual motion. We show that RSLDA outperforms the standard LDA approach for both types of datasets. These findings illustrate the benefits of exploiting subclass structure in neuroimaging data. Finally, we show that our classifier also outputs regularization profiles, enabling researchers to interpret the subclass structure in a meaningful way.

RSLDA therefore yields increased classification accuracy as well as a better interpretation of neuroimaging data. Since both results are highly favorable, we suggest to apply RSLDA for various classification problems within neuroimaging and beyond.

Introduction

Researchers commonly apply single trial analysis for the investigation of neuroimaging data. The main objective of such analysis is to study the temporal and spatial properties of neural processes that the experimental paradigm initiates. In a typical analysis scenario, a binary classifier is trained on neural responses to two types of stimuli, which can be measured with neuroimaging techniques such as EEG or fMRI. The two types of stimuli give rise to two conditions/classes,[1] and the analysis task is to discriminate the data from two such classes. Researchers have proposed various machine learning methods for this classification task (Garrett et al., 2003; Lotte et al., 2007; Pereira et al., 2009; Lemm et al., 2011). These methods differ in complexity (linear/non-linear) as well as in additional assumptions about the distribution of the data (Müller et al., 2003; Parra et al., 2005).

A problem with this approach, however, is that neuroimaging studies may employ complex experimental paradigms that do not allow for simple binary classification methods. Such complexity can arise from several subconditions/subclasses, as multiple peculiarities may appear within stimuli of the same type (i.e. the same class). Examples for an fMRI and an EEG study are therefore depicted in Fig. 1A–B and briefly described below.

Fig. 1A shows the experimental paradigm of a visual motion fMRI study with two classes. The paradigm investigates the neural correlates of upwards and downwards motion, with visual stimuli that have either low, medium or high motion coherence. With the motion direction

* Corresponding authors.
E-mail addresses: j.hoehne@tu-berlin.de (J. Höhne), klaus-robert.mueller@tu-berlin.de (K.-R. Müller), benjamin.blankertz@tu-berlin.de (B. Blankertz).

[1] The terms “class” and “condition” are considered to be equivalent. The same holds for “subclass” and “subcondition”. We use the terms “class” and “subclass” hereafter. In addition, $\{c,k\}$ denote the number of {classes, subclasses}, and $\{h,g\}$ denote the index of a {class, subclass}.

http://dx.doi.org/10.1016/j.neuroimage.2015.09.031
1053-8119/

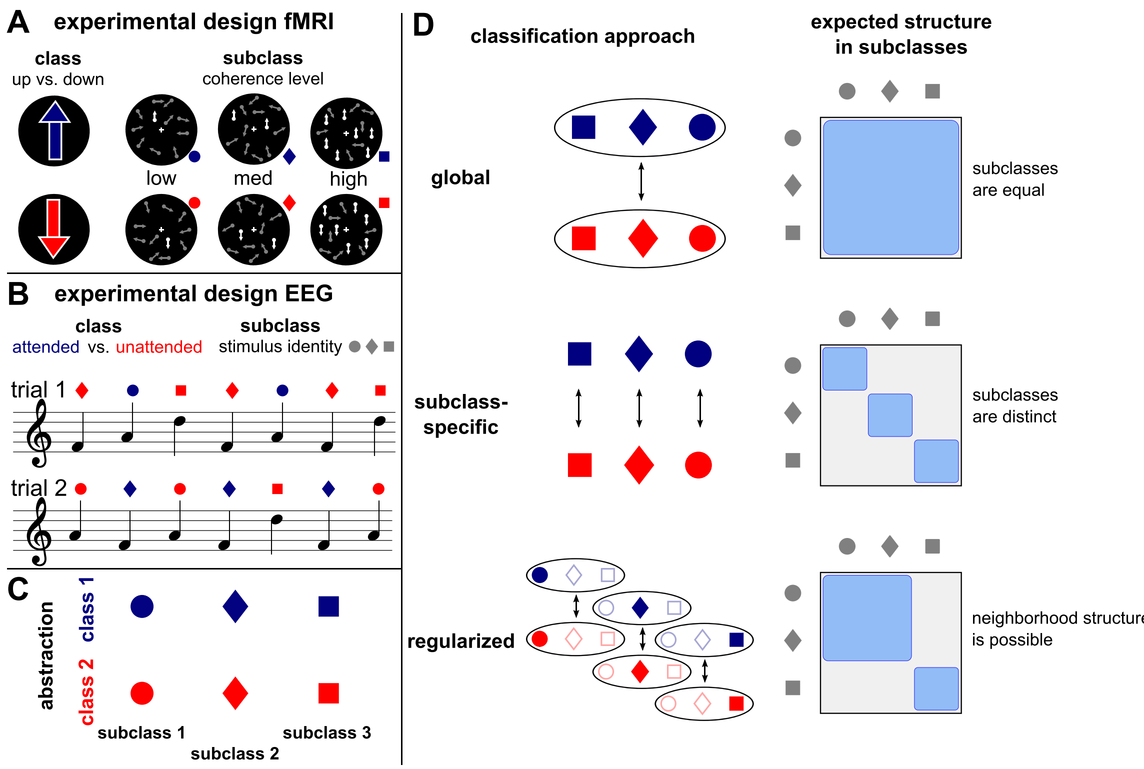

Fig. 1. Illustration of subclass structure in neuroimaging studies. Plot A shows the experimental paradigm of an fMRI study investigating upwards and downwards motion with several coherence levels. The coherence level can be considered to be a subclass. Plot B depicts the design of an EEG study in which the subject's task was to attend to specific stimuli. The data from both studies can be analyzed as a binary problem with subclass structure, as shown in plot C. Plot D visualizes three classification approaches for such data. The right column shows the underlying assumptions about the subclasses for each approach.

being the two classes (i.e. upwards vs. downwards), one can treat the coherence level as a subclass.

Fig. 1B shows the experimental paradigm of an auditory EEG study with two classes: attended and unattended stimuli. While random sequences of three types of stimuli are presented, the subjects have the task to attend to only one of them and ignore the other two stimuli. When training a classifier on the single-trial event-related potentials (ERPs) for attended vs. unattended stimuli, one can treat the stimulus identity as subclass information.

Both above mentioned studies seek for neural correlates of a binary classification problem, as illustrated in Fig. 1C. However, subclass information is available for both studies. Subclass labels are marked with different symbols and considering such information might be favorable for the classification task. Fig. 1D, therefore, depicts three classification approaches that researchers can apply for this data.

The *global* approach disregards any subclass information and thereby assumes all subclasses of each class to be equal. This approach pools data across all subclasses and computes only one classifier for the entire dataset.

The *subclass-specific* classification approach is based on one classifier for each subclass and thereby assumes each subclass to be distinct. This approach reduces the amount of available data on which to train each classifier.

The *regularized* approach presents a trade-off between the global and the subclass-specific approaches. It computes a classifier for each subclass separately, and uses the remaining subclasses for regularization. The regularized approach thereby enables the researcher to exploit any dependency or neighborhood structure that is present in the data. The drawback to this approach, however, is that it requires the researcher to estimate the additional regularization parameters on which it is based.

The aim of this article is to discuss the binary classification problem with subclass information in the context of neuroimaging data. We compare the three above mentioned approaches based on a reanalysis of existing EEG and fMRI data. We also derive a novel regularized approach – called Relevance Subclass LDA – that enables to exploit subclass information in a highly efficient way. We show that the proposed method outperforms the global and subclass-specific approach. We further show that Relevance Subclass LDA also delivers a distribution of regularization parameters. One can visualize these parameters as regularization profiles that can serve as a valuable tool for interpreting the underlying subclass structure in the data.

The remainder of this article is organized into the following sections. In the section 'Methods', we introduce the methodological details of state-of-the-art classification methods and explain why they are less than optimal in the presence of subclass structure. We then review the concept of analytic shrinkage, demonstrating how it allows the researcher efficiently to determine optimal regularization parameters. Next, we introduce the novel classification method "Relevance Subclass LDA," which is based on shrinkage. We also describe two evaluation data sets. Finally, we present results in the 'Results' section and conclude with a discussion in the 'Discussion' section.

Methods

Linear classification for neuroimaging data

Linear methods such as linear support vector machines (SVMs) or linear discriminant analysis (LDA) are commonly applied to analyze neuroimaging data. There are three main reasons why researchers prefer linear methods to more elaborate nonlinear methods (Misaki et al., 2010).

- (Performance) After applying suitable steps for feature extraction and processing, the classification performance of linear methods is on the same level as non-linear methods—or even better (LaConte et al., 2005; Krusienski et al., 2006; Misaki et al., 2010).