Pattern Recognition Letters 37 (2014) 172–177

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

Semi-supervised ensemble update strategies for on-line classification of fMRI data

Catrin Oliver Plumpton *

School of Computer Science, Bangor University, Dean Street, Gwynedd LL57 1UT, United Kingdom

ARTICLE INFO

Article history:
Available online 6 April 2013

Keywords:
Semi-supervised learning
Online learning
Classifier ensemble
fMRI

ABSTRACT

Real time classification of fMRI data allows for neurofeedback experiments, whereby stimuli are updated in accordance with the response of the brain. In order to better facilitate real time fMRI classification, we propose a random subspace ensemble of online linear classifiers. In the absence of true class labels, classifiers are updated using the 'naive' label – the label predicted by the classifier. We propose three new ensemble update strategies, using the ensemble decision for updates. Our methods are tested on two emotion based fMRI data sets. We show that the best results are produced by an ensemble which updates using the ensemble decision, constrained by ensemble confidence.

1. Introduction

Recently there has been interest in the development and application of real time fMRI, in particular, interactive experiments where data analysis is required in real time (Hollmann et al., 2008, 2011; Eklund et al., 2009, 2010). This enables the psychologist or neuroscientist to update stimuli in real time, according to classification results. Such experiments are typically conducted via a brain computer interface (BCI). BCI technology is by no means unique to fMRI, and has also been used with EEG, PET and MEG (van Gerven et al., 2009; Weiskopf et al., 2007; DeCharms, 2008).

The efficiency and precision of real time fMRI for brain control has been demonstrated by participants carrying out tasks such as navigating through computer-generated mazes (Hollmann et al., 2008), balancing a virtual inverted pendulum (Eklund et al., 2009), predicting decisions in an economic game (Hollmann et al., 2011), and moving an arrow towards a target (LaConte et al., 2007). Real time fMRI classification also allows for self-regulation experiments. Self-regulation is the ability to regulate ones' own emotions or behaviour. Self-regulation is achieved by controlling brain subnetworks, for example those involved in pain perception (DeCharms et al., 2005) or sadness (Posse et al., 2003). Based on the measured brain state, feedback is given to the participant who then attempts to adjust his/her brain state to improve task performance. A possible application of self-regulation is treating alcohol or drug addiction.

fMRI data is complex in nature, and suffers from a very high feature-to-instance ratio, typically, with very few training examples available. Linear classifiers, in particular the support vector machine with linear kernel, are popular for fMRI analysis due to their speed and accuracy, and they are less prone to overfitting the data (Grosenick et al., 2008; Eklund et al., 2009; Mourao-Miranda et al., 2007; Wang et al., 2007; Yang et al., 2010). Classifier ensembles have also been used for fMRI classification. The Random Subspace (RS) ensemble framework (Ho, 1998) has been shown to be accurate for data sets with a high feature-to-instance ratio; specifically when there is high redundancy in the feature set (Skurichina and Duin, 2002). The RS ensemble therefore seems a natural choice to apply to fMRI data, especially as the algorithm is computationally inexpensive due to the reduced number of features per ensemble member. Traditionally, many fMRI analyses take a region of interest based approach. Whilst different brain regions associated with certain processes, voxels activated by a stimuli may be highly distributed throughout the brain. The random subspace approach allows us to capture these voxels, without specifying a priori where in the brain we expect activation to occur. Kuncheva and Rodríguez (2010), compare eighteen classifier methods for fMRI data. The experiments were conducted with a variety of voxel selection methods and parameters. The RS ensemble with SVM was shown to perform best across the trials conducted.

In a typical experiment, classifiers will be trained on data from one trial, and then used for real-time classification in subsequent trials (Eklund et al., 2009; Zotevand et al., 2011; Sitaram et al., 2011). Due to the expense and time taken to organise and conduct fMRI experiments, it is advantageous to be able to complete an experiment in a single trial. LaConte (2011), presents a series of future challenges when working with real time fMRI data. The author notes that future experiments may be designed in such a way as that the brain response is *expected* to change over time. Experiments which involve performance enhancement, rehabilitation or therapy expect the brain response to change over time, with trials being conducted weeks, months or even years apart. It is suggested

* Tel.: +44 01248 382857; fax: +44 01248 383982.
E-mail address: c.o.plumpton@bangor.ac.uk

0167-8655/$ - see front matter
http://dx.doi.org/10.1016/j.patrec.2013.03.029

that future classifiers should be able to adapt and learn with the data throughout the course of the run. In these cases, pre-trained supervised classifiers will become less relevant and there is a need for a classifier which adapts with the data as it trains over time. A weak classifier may be trained on the first few data points of a run, and then continue to learn and adapt through the course of the experiment.

Online classification techniques have already been applied for EEG classification in BCI settings (Buttfield et al., 2006; Lehtonen et al., 2008). Xi et al. note that despite the advantages of online learning for fMRI data, little work has been done in the field. The authors propose an adaptation of the SVM, designed to update online, whilst addressing the dimensionality challenges associated with fMRI data (Xi et al., 2011). A similar approach was used by Hollmann et al., whereby an online classifier was used to predict a participants' decision. Following the revelation of the decision, the classifier was updated online (Hollmann et al., 2011). In one application, real time fMRI allowed subjects to form words using a character map (Eklund et al., 2010). The cursor was moved by different motor control tasks. In such a scenario, beyond the initial training phase, it is impossible to know the true class label. Such experiments provide the motivation behind exploring semi-supervised learning techniques. Previous work has shown that an RS ensemble of online linear classifiers, updating using naive labelling, has outperformed a fixed pre-trained classifier (Kuncheva et al., 2008; Plumpton et al., 2010, 2011). The drawback is the possibility of a 'runaway' classifier, where the uncorrected classifier progressively learns 'the wrong thing' (Cozman et al., 2003). The ensemble environment appears to constrain this 'runaway' behaviour. Recently, we showed that using the ensemble decision rather than individual classifier decision for updates also lead to improvement (Plumpton, 2012).

In an approach similar to co-training, in conjunction with the RS ensemble, we use a notion of confidence (Blum and Mitchell, 1998). Data whose labels are most confidently predicted are used for updates. Rather than measure the confidence across a pool of unlabelled data, we take confidence as a measure of agreement within the ensemble output. We then use this as a criterion for determining when to update the ensemble members. As an addition to this, we consider disagreement within the ensemble, updating only those classifiers whose individual predictions do not match the ensemble decision.

2. Methods

Following on from previous work (Plumpton et al., 2010, 2011, 2012), we use the online linear discriminant classifier (O-LDC) in conjunction with a random subspace (RS) ensemble (Kuncheva and Plumpton, 2008; Ho, 1998).

The RS ensemble requires two parameters. These correspond to the number of classifiers in the ensemble, L, and the cardinality of the feature subsets, M. Define $\mathbf{X} = \{x_1, \ldots, x_n\}$ to be the total set of n features. To create an RS ensemble, L feature subsets of size $M < n$ are generated by drawing at random without replacement from a uniform distribution over $\mathbf{X}$. Each of these L subsets makes up the feature set for one of the L classifiers. Features for each of the L classifiers are drawn independently. The L member classifiers are trained and tested using the respective M features. The reduction in the dimensionality of the feature set while retaining the number of training data points makes RS ensembles particularly suitable for data sets with a large feature-to-instance ratio.

Updates in the online phase are carried out using naive labelling. Training a classifier with naive labelling does not come without risk. There is a danger that the classifier may become progressively less accurate if it is trained on incorrect labels (Cozman and Cohen, 2002; Cozman et al., 2003). The likelihood of such runaway classifiers is related to the amount of offline training data and on how well the underlying data distribution model is guessed when designing the classifier (Kuncheva et al., 2008). It is expected that the lower the amount of training data is, the higher are the chances of a runaway classifier appearing in the ensemble.

In an earlier work we introduced the concept of a 'guided' update (Plumpton, 2012). Intuitively, the decision from a classifier ensemble is likely to be more accurate than the predicted label from an individual classifier within the ensemble. By using the ensemble decision to update the individual ensemble members, the likelihood of runaway classifiers is reduced. The ensemble with 'guided' updates was shown to perform better than an ensemble where its members are updated using their individual predicted labels.

Acknowledging that there is a potential for concept drift in fMRI data, and that there is uncertainty around the transition period between brain states (and thus class labels at these time points), we introduce two update criteria. These criteria correspond to the confidence of the ensemble decision, and the agreement between individual classifier predictions and ensemble decision. By using what we term *confidence* and *error* to regulate the updates, three further guided update strategies are generated:

Confidence driven. The confidence with which the ensemble has made its prediction is calculated. Denote by y the outputted (predicted) label of the ensemble, and by y_i the predicted label of the component classifiers. Confidence is calculated as

$$\text{confidence} = \frac{\sum_{i=1}^{L}\{y_i = y\}}{L}$$

Classifiers within the ensemble are updated using the ensemble decision, when the confidence in the predicted label is above a threshold. For data sets of two classes this threshold may be between 50% and 100%, here 75% is chosen, as the mid-point of the available range. This prevents the ensemble from updating during uncertain phases, which in turn should reduce the frequency of updating with an incorrect label.

Error driven. Classifiers within the ensemble whose individual predicted label does not agree with the ensemble decision are updated using the ensemble decision. We acknowledge that updating all classifiers with the same label may compromise the diversity of the ensemble. By only updating some classifiers within the ensemble, we hope to maintain diversity.

Error and confidence driven. Classifiers within the ensemble whose predicted label does not agree with the ensemble decision are updated using the ensemble decision, when the confidence in the predicted label is above a threshold.

3. Data sets

Experiments were carried out on two emotion detection data sets, supplied by the School of Psychology, Bangor University (Johnston et al., 2010). The data sets are summarised in Table 1. EN1 and EN2 are two runs of the same experiment, corresponding to single runs with two different participants. Participants were instructed to up-regulate their target region activity, evoking emotion, for periods of 20 s using negative emotional imagery. Periods of emotion are alternated with baseline periods of rest, of 14 s. There were 12 blocks of up-regulation and rest. The classification task is to distinguish between periods of emotion, and periods of rest.

Data was collected on a 3 Tesla Philips Achieva MR scanner (TR = 2 s, TE = 30 ms, 30 slices, in-plane resolution 2×2 mm^2, 3 mm slice thickness). Slices were positioned such that the bottom slice was 30 mm ventral to the anterior commissure and angled to encompass all of the ventral prefrontal cortex. The data were