fMRI Classification of Cognitive States Across Multiple Subjects

Lalla Mouatadid
Department of Computer Science
University of Toronto
Toronto, ON M5R 0A3
lalla@cs.toronto.edu

Abstract

With the evolvement of fMRI's, a great amount of attention has been given to classifying cognitive states of human beings. Several machine learning approaches have been used to train single-subject classifiers to do so. We present a different method using a neural network and a RBF SVM to train one classifier across all subjects. For the single-subject classifier case, we experiment with PCA as a feature selection step, as well as train a neural network and a RBF SVM on the preprocessed data to compare its performance to previous results.

1 Introduction

Studying the human brain activities has been a very popular topic recently. Functional Magnetic Resonance Imaging (fMRI) is a technique developed to determine which parts of the brain are active while a person is performing some sort of physical activities, or experiencing a mental stimulus. It is a 3D imaging tool used to measure the blood oxygen level dependent (BOLD) signal.

With this technology in hand, scientists have been able to predict brain patterns when a person carries out a certain task. The approach to such problems has been the gathering of 3D scans over a continuous interval of time, while subjects are doing some sort of activity and the goal is to determine which regions of the brain are stimulated to be able to predict the cognitive state of future subjects.

The experiments last a few seconds while the subject is not necessarily stimulated during the entire time interval, so the data is continuous and very noisy. The first problem that rises then is feature selection, which active voxels are really contributing to the experiment, and secondly, the classification method, we need to train a classifier that will distinguish between different cognitive states.

Various machine learning classification methods have been applied to different scenarios, such as the prediction of when a subject is looking at an image or reading a sentence, or if the subject is reading an ambiguous vs. unambiguous sentence. Various feature selections methods have also been proposed and implemented as a preprocessing to the classification.

The problem considered in this study is to differentiate between two cognitive states; whether the subject is reading a sentence or looking at an image, by training one classifier across multiple subjects. Human brains differ anatomically in shape and size, it is therefore complicated to generalize the outcome of fMRI scans, since the number of voxels differ from one subject's brain to another [4]. Our goal is therefore to explore if it is possible to train one classifier to use across multiple subjects. Succeeding to do this will allow us to associate brain activities to cognitive states independently from the anatomy of the brain. We will also briefly cover some results of using our methods on the subject-specific classifiers.

In section 2, we define the experiment in more details and provide on overview of previous work. In section 3, we provide a detailed description of our approach, in section 4 the experimental results and their interpretation, and in 5 our conclusion and some proposed research directions in this area.

2 Problem definition and background overview

In this study, the subjects are presented with a sentence to read, or a picture to look at (in no particular order) for 4 seconds, then they get 4 seconds of blank activity followed by the 2nd stimulus (a picture or a sentence - whichever they didnt get first), followed by 15 seconds of blank. A scan is taken every 500ms. Although only 8 seconds of the experiment are of interest to us, we consider 8 seconds per stimulus, keeping scans taken after the 4 seconds interval because, as explained by Mitchell et al [5], the BOLD signal lasts a few seconds after the neural activity is completed.

The problem becomes then, given an 8 second interval (16 scans), we need to train a classifier to distinguish between two cognitive states.

Mitchell et al [5] have trained different subject-specific classifiers for this specific experiment. The focus of our study will mainly be on training a multiple-subject classifier that can be applied on any fMRI scan. Previous studies [6] have used Gaussian Naive Bayes (GNB), Linear Support Vector Machine (L.SVM) and *k* Nearest Neighbor (*k*-NN). We train a Neural Network (NN) and a Radial Basis Function SVM (RBF SVM) using various feature selection methods, and compare their performance to what was previously accomplished.

Various feature selection methods have been proposed [5, 2], such as selecting the n-most discriminating voxels vs. the n-most active voxels where they have showed that selecting the n-most active voxels was the most successfully method. A third method is to use the n-most active voxels per region of interest (ROI), and [5, 2] have specified the seven ROIs of the brain that are the most relevant to this study. However, in order to avoid limiting a classifier to fixed spatial patterns, it is preferred to train on all regions [4], we therefore experiment training our classifiers on all 25 ROIs.

3 Our method and analysis

3.1 Data preprocessing

The focus of this study is to train a neural network and a RBF SVM across multiple subjects, and briefly cover subject specific classifiers. Preprocessing is a necessary step as our data is very high dimensional, it differs however for each classification approach (i.e single vs. multiple subject), since each one requires a different data representation.

3.1.1 Data representation

Every fMRI scan has on average 5000 voxels, and each experiment is represented by 16 images, thatś a total of 80000 voxels per experiment. Each subject gets 40 trials where one trial has a picture experiment and a sentence experiment, therefore in total, every subject gets 80 experiments, and we have 6 subjects in total.

3.1.2 Feature selection

Training a classifier to use across different subjects is a challenging task, mainly because peoples brains differ anatomically and therefore the fMRI scans are not similar. So the biggest challenge is to select the right voxels that will help the classification be more accurate.

Principal Component Analysis (PCA) -using Singular Value Decomposition (SVD) was considered, however because our data differs in dimensions, it would not be a very accurate approach to take for this particular case. It does however make sense to use it for the subject-specific classifiers. In fact, when dealing with one classifier per subject, there is no need to take into consideration the differences in the anatomy of the brain or how the scans differ from one subject to another.

The data is very high dimensional and noisy, we therefore experiment with various feature selection methods: ROI Supervoxels and n-active/ROI, proposed by [5], in addition to PCA for the subject-specific approach:

- ROI Supervoxel: for every ROI, we get the mean of all its active voxels.
- n-active/ROI: select only the n most active voxels per ROI, and calculate their mean.
- PCA where we experiment with different numbers of principal components (PCs) for every subject.

Both supervoxel methods are used for all regions of interest (25 in total) as well as the 7 most relevant ones to the experiment. Using the above methods reduce the dimensions drastically from 80000 voxels to 7*16 = 112 supervoxels for the 7ROI preprocessing, and 25*16 = 400 for all 25 regions.

3.2 Classifiers

The classification function for this problem is of the form:

$$f : fMRIscans(T_i) \rightarrow \{\textit{Sentence, Picture}\}$$

Where, given a time interval T_i of 8 seconds, 16 fMRI scans are taken, the function *f* will distinguish the cognitive state of the subject. Previous approaches to this classification problem were *k*-NN, GNB and linear SVMs, we have implemented a RBF SVM and a feed-forward back-propagation neural network with one layer of hidden units.

3.2.1 RBF SVM

When it comes to classification, SVM's have proven to work very well [3], in fact, as [6] have shown, linear SVM was successful for this particular classification. We experiment with SVMs using a Gaussian radial basis function as a kernel, of the form:

$$\kappa(\mathbf{x_i}, \mathbf{x_j}) = \mathbf{e}^{-\gamma\|\mathbf{x_i}-\mathbf{x_j}\|^2} \text{ where } \gamma > 0$$

Since the implementation of SVM's requires solving quadratic programming [1], we decided to use LIBSVM instead, for a detailed explanation on SVM as a classification function, we refer the reader to [2, 1]. For the multiple-subject classifiers, we get the best accuracy by setting γ to 0.16, and applying regularization with a cost equals to 8 for slack variables.

3.2.2 Neural Networks

We have trained a one layer neural network using cross-entropy as the objective function and sigmoid as the activation function for the output of the hidden units, which is then the new input to the outer layer, and we applied the softmax activation for the outer outputs. This softmax activation function estimates the classification probability (i.e P or S) as follow:

$$y_i = \frac{e^{x_i}}{\Sigma_j e^{x_j}}$$

Where x_j is the sigmoid of output of the hidden layer (and also the input to the outer layer). Given that we only deal with two classes (picture or sentence), out activation function becomes then:

$$y_i = \frac{e^{x_1}}{e^{x_1}+e^{x_0}} = \frac{1}{1+e^{-(x_1-x_0)}}$$

Although neural nets are a powerful tool when implemented properly, the major drawback is trying to find the right values of parameters which maximimze the accuracy rate. Using a grid search approach, we get the best accuracy with the following parameter:

- Number of Hidden Units: 20
- Learning Rate: 0.0001
- Momentum: 0.8

Table 1: Accuracies for the multiple subject classifiers

METHOD	NN	RBF SVM	GNB	L. SVM	1NN	3NN	5NN
7ROI Supervoxel	71.45%	74.16%	74.3%	75.3%	63.7%	67.3%	68.3%
7ROI n-Active Supervoxels(100)	66.54%	66.67%	72.8%	NA	66.0%	71.5%	72.0%
25ROI Supervoxels	73.75%	70.83%	NA	NA	NA	NA	NA
25ROI n-Active Supervoxels(100)	69.16%	68.75%	NA	NA	NA	NA	NA

Table 2: Accuracies for the multiple subject classifiers PS vs. SP vs. SP+PS

METHOD	CLASSIFIER	SP	PS	SP+PS
7ROI Supervoxels	NN	88.33%	80.03%	71.45%
7ROI Supervoxels	RBF SVM	91.66%	81.66%	74.16%
7ROI Supervoxels	GNB	88.8%	82.3%	74.3%
7ROI Supervoxels	L. SVM	86.5%	77.1%	75.3%
7ROI n-Active Supervoxels(100)	NN	82.91%	65.41%	66.54%
7ROI n-Active Supervoxels(100)	RBF SVM	83.75%	61.66%	66.67%
7ROI n-Active Supervoxels(20)	5NN	93.8%	87.5%	72.0%

When dealing with multiple-subject classifiers, leave-one-subject-out cross validation was the most suitable method to evaluate the error of our classifiers, since we only have 6 subjects in total. For the single-subject case, we split our data to keep a testing set. We present our results in the following section.

4 Results

4.1 Multiple Subject Classifiers

For every method, we present the average accuracy across all LOO-subjects in table 1. For comparison, we also include Wang et al's results for their GNB, linear SVm and *k*-NN classifiers [2].

Looking at table 1, we see that our classifiers did not surpass the accuracy of L. SVM, or even GNB. Although the RBF SVM has a lower error rate than the neural network, it is still being outperformed by the Gaussian Naive Bayes. Using all region of interests did not achieve a higher accuracy either. It would be interesting to see if, given a bigger data set, we will achieve a better classification. In fact, with such a small data set, and a reduced number of dimensions, it is hard to tell how fast we would start over-fitting.

For the n-Active supervoxels method, we use 100 active voxels, where as [6] uses only 20. We noticed in the experiment that using 100 or more voxels, the error rate does not vary much, we were not able to achieve a better accuracy rate using less than 100 voxels however.

4.2 Different data sets

In [6]'s experiment, they have focused on 3 different sets; the set of trials that start with a picture followed by a sentence, denoted *SP data set*, the trials where sentences were presented first followed by a picture *PS data set*, and finally, the union of *SP* and *PS* (which is what we have done so far).

Table 3: Accuracies per subject using RBF SVM and NN

METHOD	Subj. 05680	04847	04799	05710	04820	05675
RBF SVM	70.0%	75.0%	60.0%	65.0%	65.0%	70.0%
NN	60.0%	60.00%	75.0%	60.0%	60.0%	55.0%

Table 4: Average of accuracies over all single-subject classifiers

METHOD	AVERAGE ACCURACY
RBF SVM	67.5%
NN	61.66%
GNB	82.0%

We train our classifiers on *SP* and *PS* as well to compare their performances. Table 2 shows our results for 7ROI's vs. Wang et al's best results (we only display the best accuracy out of all 3 classifiers they trained).

When looking at the 7 regions only, our RBF SVM performs better than all other classifiers on the *SP* set, and is only *0.64%* away from the best accuracy (GNB) for *PS*. We have achieved both results by setting the cost to 4, and γ to 0.1. It is surprising to see that the linear SVM is outperformed by all classifiers when looking at the *PS* and *SP* sets separately. The order of the stimulus does indeed increase the accuracy, we explain this by the fact that the brain is in a similar environment when the trials are being conducted, so the data is more correlated.

4.3 Single Subject Classifiers

Although the focus of this study is not the single-subject classification, we explore the usage of PCA as a feature selection method, and apply our SVM and NN as classifiers. For each subject, the optimal number of principal components was chosen to minimize the classification error. When training our SVM (resp NN), we chose the cost and γ (resp. the number of hidden units, the learning rate and the momentum) which give us the best accuracy. We split our data randomly into a training, validation and test sets, and take the average accuracy over all random splits. We present in table 3 our results for each subject. For comparison purposes, table 4 has the average accuracy for our classifiers (using PCA) as well as Wang et al's best GNB results (using the 240 most active voxels). We compare the average only since we do not know if we are training on the same subjects and we can see that using PCA to reduce dimensions was not as successful as we expected.

5 Conclusions

The goal of this study is to present a new approach to classifying cognitive states of human subjects, and the focus was on training one classifier across multiple subjects. Because FMRI's data is large and sparse, we compressed most of it during the preprocessing phase, one concern about this technique is whether it introduces noise to our input, and if so, then how much does this noise affect the data set and the classification. And as mentioned before, we need to look at over-fitting closely. The data set we have used is small, and our feature selection approaches have reduced the dimensions drastically, what would the classifiers performance be if we are given larger data sets?

As we saw on table 2, separating the experiments resulted in better accuracy, this suggests that voxels activation differs significantly when the context changes even slightly, so there are two things to consider here; first the voxels being selected for the training; ROI's differ in size, and therefore in the number of voxels, this suggests that some ROI's will have a higher impact on the classification than others. The question that rises here is how to select the voxels within the regions. Can we select *k* voxels per ROI's, instead of having *n* or all of them. Secondly, we need to consider the actual cognitive state we are trying to classify, since the *PS* vs. *SP* classifications produced better results, we need to consider using a mixture of experts approach, where we train an expert for every outcome *picture* or *sentence*, would that outperform GNB and linear SVM? Or consider training an expert for every active voxel or for every cluster of voxels that have identical behavior.

References

[1] C.M. Bishop et al. *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[2] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May 2011.

[3] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[4] J.A. Etzel, V. Gazzola, and C. Keysers. An introduction to anatomical roi-based fmri classification analysis. *Brain research*, 1282:114–125, 2009.

[5] Tom M. Mitchell, Rebecca Hutchinson, Radu S. Niculescu, Francisco Pereira, Xuerui Wang, Marcel Just, and Sharlene Newman. Learning to decode cognitive states from brain images. *Mach. Learn.*, 57(1-2):145–175, October 2004.

[6] X. Wang, R. Hutchinson, and T.M. Mitchell. Training fmri classifiers to detect cognitive states across multiple human subjects. *NIPS03*, 2003.