Classifying fMRI images using sparse coding: a project for CS229

Christopher Archibald
cja@stanford.ed

Evan Millar
emillar@stanford.edu

December 15, 2006

1 Introduction

Recent work has been done by Rajat Raina and other researchers at Stanford in applying sparse coding techniques to various classification problems. In this project we follow that tradition by applying sparse coding to the problem of classifying fMRI images. In particular, we try to classify fMRI images based on what the subject was doing when the fMRI image was obtained. The two classification possibilities we consider are whether the subject was looking at a sentence or looking at a picture.

Rajat Raina provided direction and advise on sparse coding techniques, as well as a code base for sparse coding and classification. We acknowledge him and thank him for his help.

1.1 FMRI images

FMRI (Functional Magnetic Resonance Imaging), is a method of using magnetic resonance imaging to measure the change in blood flow and oxygenation of blood in the brain. FMRI data is typically noisy and very high-dimensional. It is also difficult to obtain large amounts of data from single subjects. FMRI data must be pre-processed in a fairly significant way so that the resulting data is usable.

For our project, we started with the data obtained in the Star/Plus fMRI project at Carnegie Mellon. This data has already been preprocessed and arranged in a manner that makes dealing with it more practical. Each patient has 1280 "snapshots". Half of those snapshots the patient is looking at a sentence, and the other half the patient is looking at an image. Each

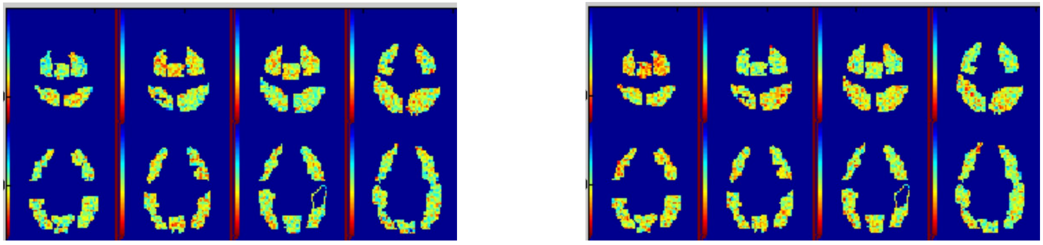

Figure 1: Sample fMRI images

snapshot contains eight two-dimensional images, each of which is a cross section or "slice" at a certain depth in the brain.

1.2 Sparse coding

Sparse coding is a technique for representing information that has been postulated to occur in the V1 portion of the visual cortex. Sparse coding algorithms learn bases that can be linearly combined to reconstruct the input data. Sparse coding algorithms specifically try to come up with a set of bases that will allow for accurate reconstruction of the input with as few images as possible.

Given input images $x^{(i)}$, the sparse coding algorithm finds activations $a^{(i)}$ and bases b to minimize equation **??**.

$$\min_{b,a} \sum_i ||x^{(i)} - \sum_j a_j^{(i)} b_j||_2^2 + \beta ||a||_1 \tag{1}$$

Our hope in applying sparse coding to fMRI data is to be able to identify regions in the brain that have correlated activity, much like ROIs (Regions of Interest) used in typical fMRI data analysis. ROIs must be hand-identified by someone with neuroscience knowledge, who can analyze the brain and designate the regions of interest as defined by specific anatomical landmarks. Through sparse coding, our hope was to achieve comparable performance without brain-specific knowledge.

2 Experiment

Originally we wanted to be able to train bases and a classifier that worked across multiple subjects. Our goal was to be able to train on data from a group of subjects, and then test on a new subject's data. However the shape and size of the brain varied substantially between subjects. We realized that in order to make this work we had to somehow normalize the fMRI data across subjects.

We decided to translate and scale each fMRI slice to a standard size of 32x32 pixels. This helped to line up the slices so that the brain size and shape across patients had more congruency. However upon calculating bases and training classifiers with this normalized data, we discovered that this data normalization was not sufficient to be able to compare data across subjects effectively. After doing more research into how this issue is dealt with in past studies, we found that this is actually a significant problem in machine learning of fMRI data. The most common solution is to use manually labeled regions of interest and then compare those regions across subjects. We decided that for the scope of this project we would stick with single subject bases and classifiers. This in itself is a hard problem, and one that has not been perfected as yet. We felt that this was still a significant topic to tackle, and one that could perhaps be improved upon by sparse coding techniques.

Once we decided to work to classify within a single subject, we tried several test runs of the whole process on one subject, with variations in the sparse coding and PCA parameters. We varied things like the PCA reduction dimension, the β value, or sparseness penalty, in the sparse coding algorithm, the number of bases learned, and the size of the test and training sets. From the results of these varied experiments on a single subject we decided to proceed with experiments involving the rest of the subjects, reducing our data to 400 dimension through PCA, using a β value of 1, and using 90% of the data as training data and 10% as test. We then proceeded to run the code with this setup on two other subjects. We present the results in the next section.

3 Results

Most of the subjects data contained about 5000 voxels. Using PCA, we reduced this data to 400 dimensional data. This reduction usually preserved between 60% and 70% of the total variance of the original data. Figure **??** shows an original fMRI image and then that same image reduced to 400 dimensions and then converted back to its original dimensionality. This is what we used to give us a feel for how the different PCA reduced dimension sizes were doing at preserving the original data.

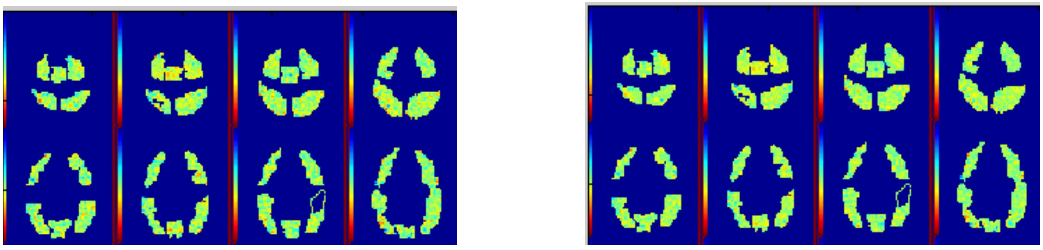

Figure 2: Comparison of original fMRI image (left) with PCA-reduced image

3.1 Sparse coding results

The sparse coding algorithm learned 50 bases for each of our subjects. The sparsity penalty β was 1, and this gave us average activation activities around 7%, meaning each fMRI image was reconstructed using only about 7% of the bases. With 50 bases learned, this means that on average we were using between 3 and 4 bases to reconstruct a given image. An example of a reconstruction that used only 2 bases is shown in figure **??**.

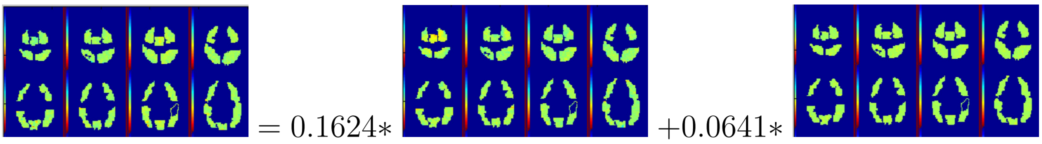

Figure 3: Reconstruction of fMRI image from bases 11 and 46

3.2 Classification results

We ran our code on the fMRI data of 3 subjects, and trained GDA classifiers on 90% of the data and tested on the other 10%. As a baseline for comparison we also trained classifiers on the raw voxel information after the PCA dimension reduction. This is shown in the table under raw features. The results of these experiments are shown in table **??**.

4 Conclusions and future work

Results of our various experiments were varied and inconsistent. Some indicated that sparse coding had a significant impact on the accuracy of classification of fMRI images, while others showed that sparse coding struggled to compete with use of raw voxel information to classify the images. Our conclusions are that more work needs to be done determining just how much

Table 1: Classification accuracies for GDA classifiers trained using raw and sparse features

	Training size	Raw features	Sparse features
Subject 4847			
	100	0.727	0.711
	200	0.742	0.727
	500	0.773	0.766
	1152	0.859	0.766
Subject 4820			
	100	0.570	0.570
	200	0.617	0.672
	500	0.672	0.703
	1152	0.789	0.672
Subject 5710			
	100	0.656	0.578
	200	0.703	0.578
	500	0.688	0.594
	1152	0.852	0.562

sparse coding can impact fMRI classification. The amount of time necessary to complete a cycle of bases learning and classification limited the number of experiment we were able to complete. As it was, we completed a fair number of experiments, but almost always with different parameters involved.

In the future, we would like to continue with a more comprehensive evaluation of some of the techniques we have used in this project. Specifically, it would be very helpful to use some form of cross-validation to evaluate the classifiers results on this data set. We plan also to vary the number of bases we learn, and see what effect this has on classification ability. We also would like to investigate the effectiveness of sparse coding as applied to time-sequences of fMRI images, as opposed to single time snapshots, as we attempted in this project. Bases learned on time-sequences might give us more robust classification possibilities and might contain new information about the relationships of different areas of the brain. We plan also to use sparse coding to help classify fMRI images or sequences into more categories than just two. This would be necessary in the Pittsburgh Brain Activity Interpretation Competition (www.ebc.pitt.edu), which would also provide a lot of data and opportunity to determine the effectiveness of a sparse coding approach. We are optimistic about the possibilities of sparse coding in the future of fMRI interpretation.