CS 229: Machine Learning

Classifying Modality of Pain from Calcium Imaging Florescence data

Amelia Christensen

Motivation. When you are bitten by a mosquito, you know that the bump itches. When you burn your hand on a match, you know that the match was hot. This ability to discriminate sensory modalities is present, despite the fact that these stimuli activate overlapping sets of peripheral receptors, and travel to your brain generally along the same paths, utilizing overlapping sets of neurons [1]. How, or where, this distinction emerges, remains unclear in the neuroscience literature.

To overcome this limitation, we designed a machine learning approach, whereby we built a classifier to determine, purely from neural data, what modality of stimulus we present to an animal. The logic is that if the brain region from whence we record neural data is actively involved in task of detecting the different stimuli modalities, are decoder will be able to easily distinguish between them. This approach is similar to one often taken in both visual and decision making literature [2].

Collection of neural data. To accomplish this goal, we undertook to record the activity of neurons in primary somatosensory cortex (of mice), a region that is heavily implicated through human FMRI work, to be important for distinguishing different tactile stimuli [3]. We utilized transgenic mice that were expressing GCaMP6f, a transgene that fluoresces whenever the given neuron it is expressed in fires an action potential [4]. Thus, by implanting a glass coverslip over the cortex of a mouse, we can use two-photon fluorescence microscopy to observe that activity in a large population of neurons at a time. In our experiments, we simultaneously recorded 177 neurons at a 1 hz frame rate, over around a mm^2 of cortex. See figure 1. Data were processed using a standard pipeline (pre-existent in our laboratory), whereby the images recorded from the microscope are first registered frame by frame to each other, and then individual cell bodies are segmented in a quasi-manual way, and finally fluorescent time series for each individual neuron are extracted.

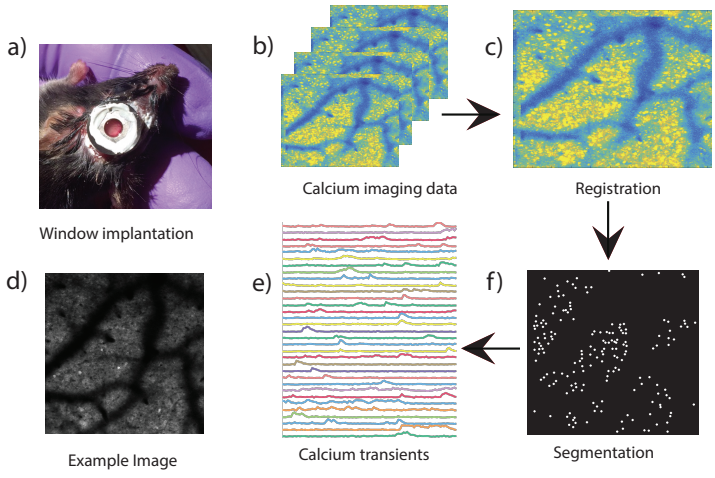

Figure 1: **Collection of neural data.** a). implanted window and head fixation device. d) example image of neural activity. b-e) Schematic of image processing pipeline.

Experimental Paradigm. We prepared seven distinct types of stimuli to present to the mouse. These stimuli types included cold (ice and acetone), hot (a heating pad), mechanical (a clip to put mechanical pressure on the mouses paw, and sticky tape), vibrational stimuli (a small vibrating motor), and finally, nothing. We alternated placing these different stimuli on the contralateral hind paw to the side of somatosensory cortex from which we were recording. Neural responses to each stimulus were recorded for one minute, and then the mouse was given a five minute break, after which we switched stimuli and recorded again. All told, we recorded five different trials for each of the seven stimuli, for a total of 35 minutes of neural recording. All recordings were performed under light isofluorane anesthesia, which, to the best of our ability, was kept constant throughout the recording. All data were recorded on one day, from a single mouse.

Determining features. Given a set of time series extracted from a given group of neurons, labelled by the trial from which they came, it is not immediately obvious what features should be defined as. Our overarching goal was to predict which stimulus was presented on a given trial, so our first thought was to use a slice of time from each trial as a training example, where the features are individual neurons at that slice of time. Unfortunately, our results with that approach were quite poor (see figure 2), and with all of the classifiers we tried, the classifier mislabelled almost everything as acetone, and in PCA acetone was the only cluster that emerged. While this is potentially interesting from a neuroscience perspective, it's not particularly helpful for the goals of this project. Next, we thought to try a paradigm where each neuron is a training example, and it's label is what trial type the observations of that neuron came from. In this case, the features time points. This approach was much more successful, and was the paradigm we used for the rest of the project.

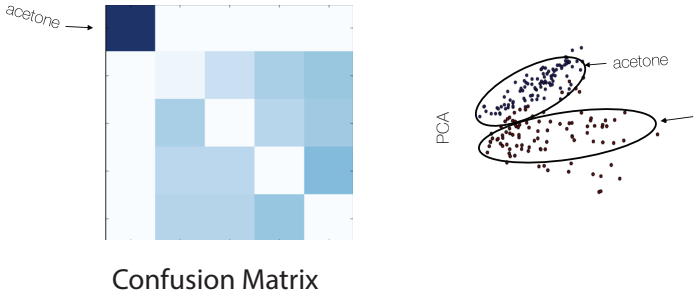

Figure 2: **Feature Extraction.** Left, confusion matrix between different stimuli modalities, using LDA as a classifier. Results were similar with multinomial regression, naive bayes, random forests, etc. Right, data projected on first two principal components.

Supervised learning of stimulus category. We trained four different classifiers on this data set, all of which achieved above chance performance. Specifically, we tested Random Forests, LDA, Multinomial Regression, and Naive Bayes. We used a 75 -25 train/test split. We found that on this data set, multinomial regression performed best, with LDA coming in second, and then naive bayes, and then random forests (see figure 3). These results make some sense, because LDA and Multinomial regression are very similar mathematically (in fact it can be shown that Multinomial regression is simply a restriction of LDA where the covariance matrix of the data is assumed to be diagonal). The Naive Bayes assumption is particularly bad for this type of dataset, so it's unsurprising that it's performance was lower than that of other algorithms. It is likely that random

forest, a very complex hypothesis class to utilize on so little data, was overfitting.

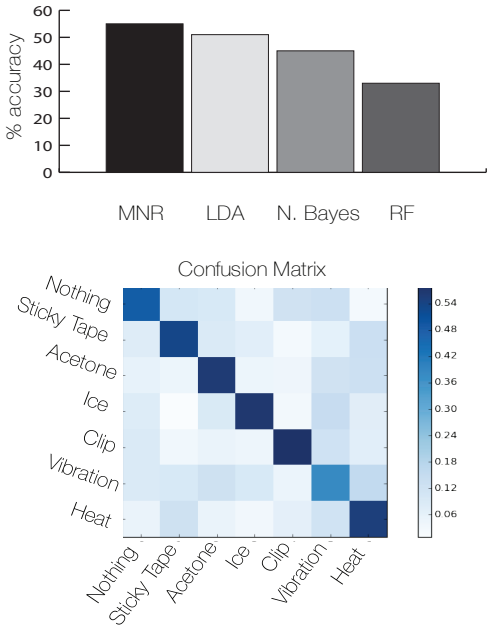

Figure 3: **Classifier Comparison.** Top: performance of different classifiers on validation set. Bottom: Confusion Matrix of results using Multinomial Regression as a classifier.

Although these results are much better than chance (14%), in the literature it is common for such classifiers trained on neural data to discriminate other types of stimuli, typically reach performance in the high 90% range, even using very vanilla machine learning algorithms [2]. It is likely that this difference stems from the fact that in our animals, the mice were not attempting to discriminate between the different stimuli, the animals were simply passively observing, so it's likely there would be less choice related information present in cortex. In future studies we would need to collect more data, or design an awake discrimination task, in order to increase performance.

Although it may be possible to engineer combinations of features in order to increase this performance a little bit, such manipulations are really uncommon in the neuroscience literature, and reduce interpretability of the results.

Clustering of stimulus category. Next we wanted to see whether, when we projected the data onto an orthogonal basis which preserved a as much variance as possible (using PCA), there were any discernible clusters of the different stimuli modality. PCA is very commonly used in neuroscience to segregate different categories of neural responses to stimuli [5]. When we used the data in the same way as above, we didn't see any clustering (at least in the 2D projection of the data), however when we average each neurons firing over all of the trials, leaving only one time series for each neuron in each trial type, suddenly we were able to see clusters. We have included LDA below for comparison.

Neural trajectories in State Space. Next, we wished to use an analysis called GPFA [6], sometimes used in neuroscience, to determine whether the trajectories of each trial through neural

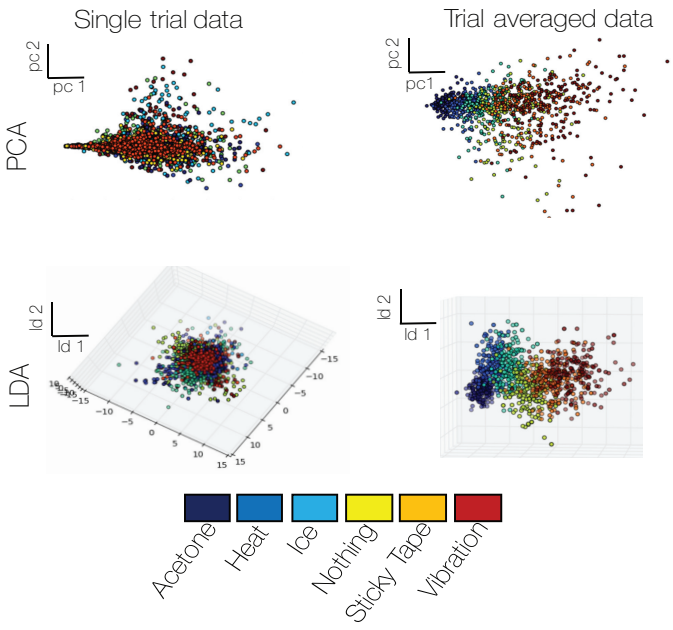

Figure 4: **unsupervised clustering of data.** Left column: single trial data. Right column: trial averaged data. Here each point is a single neuron.

state space was illuminating. Consider that neural state space is space where each axis is a firing rate of a given neuron. Then, each time point in a trial can be thought of as a point in state space. But this space is very high dimensional, and activity of neurons is sparse in this high dimensional space. GPFA plots neural trajectories through a dimensionality reduced (using Factor Analysis) state space, with some constraints of smoothness imposed. When we performed this analysis on our data, we noticed that trials in which the stimuli were the most salient, the firing rate of that neuron seemed to cover the most ground in state space. However, there were no other immediately obvious aspects of the trajectories where were particularly interesting, as they didn't display particularly stereotyped behavior. Similarly to the classification section, I think if we collected more data, in a more careful manor, possibly in an active perception task, this would possibly be a more fruitful analysis.

Figure 5: **Neural trajectories in state space.** Leftmost: Clip and Ice, two salient stimuli, similarly show a high amount of variation. In the middle, Clip is compared to Nothing, where the nothing trajectory stays in a compact ball.

Conclusion. We demonstrated that we were able to train a classifier from the neural firing from somatosensory cortex of a mouse, to determine what stimuli we presented to his paw. These results were far better than chance for all of the classifiers we tried, however even for the best classifier, the results were still less than we would expect if the brain region was actively involved in perceiving the difference between the stimuli. We also experimented with different methods for visualizing the data, including PCA, which was only helpful when we trial averaged the data, and GPFA, which was perhaps more useful for generating abstract art in this context, than insight into our dataset.

Acknowledgements. First and foremost I should acknowledge my labmate Saurabh Vyas, who wrote the initial image processing necessary for this project. I should also thank my advisor Scott Delp, and the Stanford Microscopy Core, for the usage of their microscope.

Citations.

1. J. Braz. et. al. Transmitting pain and itch messages: A contemporary view of the spinal cord circuits that generate Gate Control. *Neuron*. 2014.

2. R.M. Haefner et. al. Inferring decoding strategies from choice probabilities in the presence of correlated variability. *nature neuroscience*. 2013.

3. M.C. Lee and I Tracey. Imaging pain: a potent means for investigating pain mechanisms in patients. *BJA*. (2013)

4. T.W. Chen et. al. Ultra-sensitive fluorescent proteins for imaging neuronal activity. *Nature*. (2013)

5. J.P. Cunningham and B. Yu. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*. (2014)

6. B. Yu. et. al. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *NIPS*. (2009)