

Team Members

Leavitt, Lee

- email: u0591788@gcloud.utah.edu
- UID: u0591788

Desabathina, Indumathi

- email: u1265976@utah.edu
- UID: u1265976

Cristiano, Deicy

- email: u1228538@utah.edu
- UID: u1228538

Introduction

Heterogeneous populations of cells provide massive potential for drug target discovery. Each cell has a multitude of receptors, ion channels and GPCRs, each of which has various probabilities of success for drug targeting. The system to be studied in this project is the Dorsal Root Ganglion. This is the sensory organ of the body and is responsible for the transmission of all sensations to the central nervous system. Within this system, and depending on the experimental design, *each cell is an experiment*. To understand which ion channels and GPCRs are on the surface of these cells we load them with a dye that binds to calcium. When these channels become activated calcium will flow into the cell. Measuring calcium from these cells (~3000) provides a functional approach to viewing and perturbing these potential drug targets. Once a target is identified this system can either be used as the drug screening platform or provide researchers with the target to further study and develop on higher throughput screens.

Project Objectives

We would like to automatically identify which molecules activate which kind of cells. In other words, we would like to get a classification over the cells attributed to their reaction under some different plant derived natural products. In addition, as the response is not uniformly across the cells, is our interest focuses on those cells which react under some component, which percentage is very low compared to all the cells that we are considering.

Methods

Cells from the dorsal ganglion are gently dissociated and plated in a minimal media. Figure 1A shows an intact dorsal root ganglion. Each green dot is an individual neuron. Figure 1B shows these cells after dissociation. The red and green colors on the surface of these cells identify 2 different types of pain sensing neurons. Cells recover for ~16-24 hours, the following day, cells are loaded with Fura2 (Figure 1C). This is a dye which binds to calcium. This is the signal for activation of receptors and ion channels on the surface of these cells. A region of interest is then identified using [CellProfiler](#). This is shown in figure 1D. These cells are then imaged every 2 seconds while different plant derived natural products are applied to these neurons. Follow this [link](#) to watch of video of how this data is collected.

Figure 1

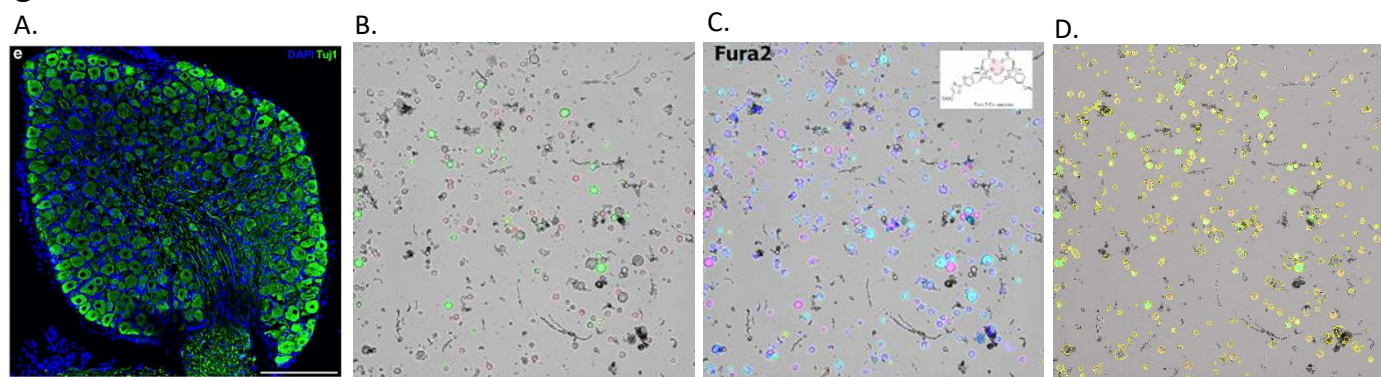
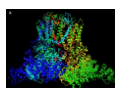


Figure 2

AITC: Wasabi, Horseradish
Sensation: noxious
Target: TRPA1



Menthol: Produced by Mint plants
Sensation: Cold
Target: TRPM8



Capsaicin: Produced by spicy chili
Sensation: Hot
Target: TRPV1



During the video three different plant derived natural products are added to the cells. Each molecule activates a different ion channels on the surface of the cells. As can be observed each cell has a different composition of ion channels present on the surface of the cell. Cell 1 senses cold since menthol activate the cold sensing ion channel TRPM8. Cell 3 senses heat due to capsaicin activating the heat sensing ion channel TRPV1. For a summary of these ion channels see figure 2.

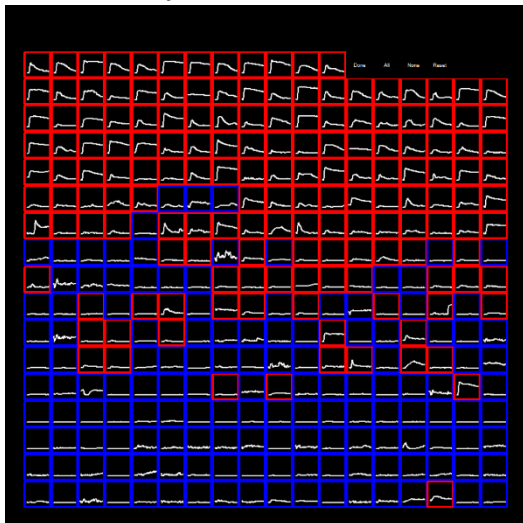
The video shows a small example of three cells imaged during these experiments. In practice the field of view is much larger since we use a 4x objective. This increased view of our cells allows us to collect data from ~3000 cells at a time. Each of these 3000 cells are scored by the researchers as responded or did not respond. Figure 3 shows the tool researchers use to correct the labels, each applications scoring is corrected as either 'yes the cell responded to this compound' (the trace with a red box surrounding it represents this). Every trace within a blue box represents 'no the cell did not respond to this compound'.

The researchers cell type definition and advanced analysis depend on the correct scoring of these responses and this process take ~2 hours. We have been collecting data for the past few years and have

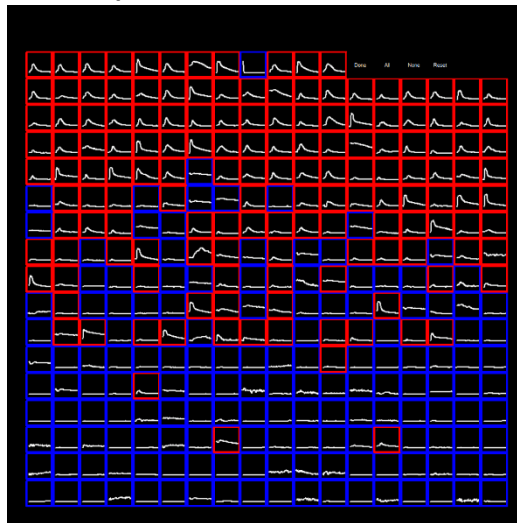
built up a large and clean dataset. Du to the hard work of these researchers our data set has a clean label space.

Figure 3

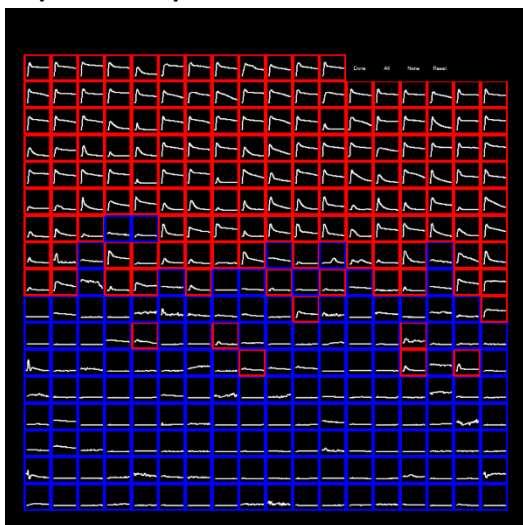
Menthol Responses



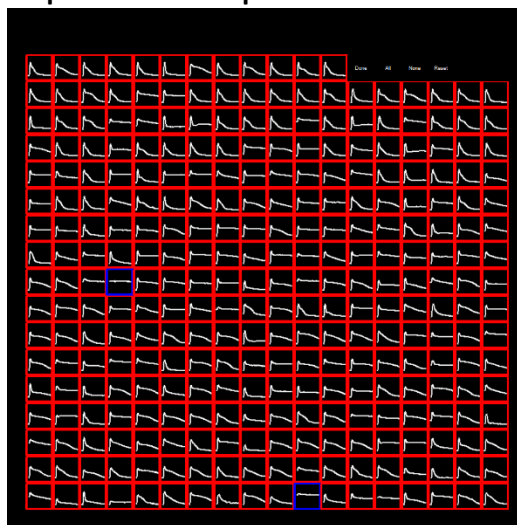
AITC Responses



Capsaicin Responses



Depolarization Responses



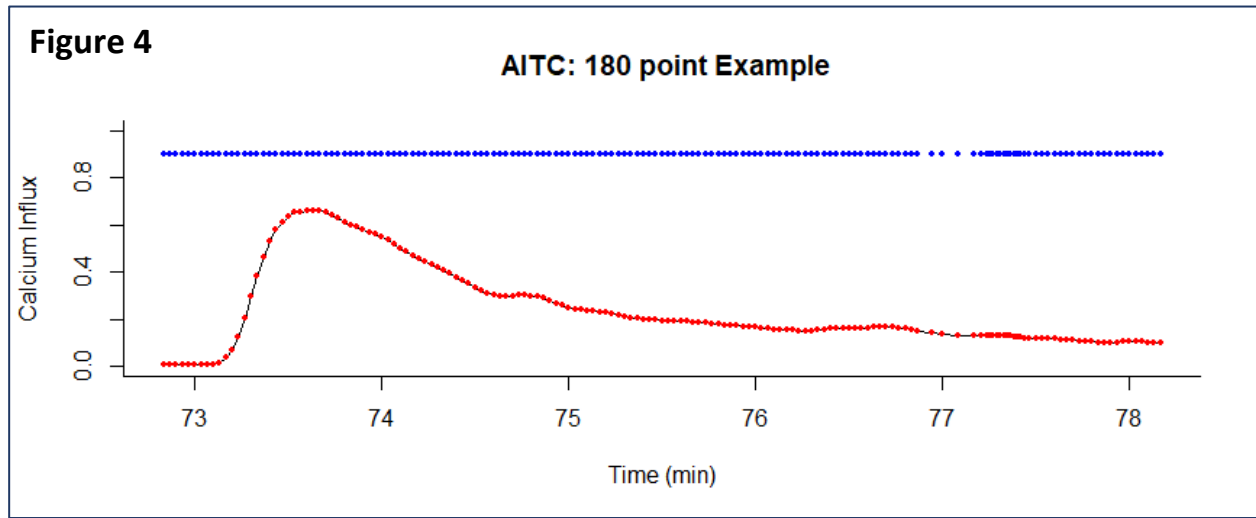
Data

We are interested in checking the effect of some molecules on some cells. We will use the data from the laboratory of Dr. Baldomero Olivera. These networks have ~3000 cells, and it is recorded by minutes. We have a data with ~15000 cells, but the cells that respond only make up 13% of all cells.

Figure 4 shows a representation of a single trace from a single cell. This will be the data to feed into the neural networks and will be used to create a neural network which predicts whether the cells response or not to some molecule. The red dots represent each time point of the response. The x axis is time in

minutes. The y axis represents the influx on calcium into a single cell. The data is scaled between a value of 0 and 1. A value of one is determined from a maximal depolarizing pulse at the end of the experiment. This should tell us what is the maximum at which this cell can respond.

The blue circles are to represent the timing of the image capture. One thing to note is the timing of the image capture separating at ~77 minutes, shortly have the timing of image collection has extended the timing then rapidly increases. In addition to this there is a pattern which can be observed in the timing.



Approach

The pattern analysis provides the diagnosis of the response of cells. For that reason, the signals will be used as input to a deep learning network architecture for classification since in the literature some neural network architectures has been proposed showing better results than traditional machine learning applications. In fact, a previous attempt has been made in machine learning. Using SVM's, Lasso Ridge and Elastic Net regression the success rate of these method proved to be unsatisfactory. Although the total success rate is 93% the success rate for 1's was 65.2%. This is due to the fact that cells that respond only make up 13 % of all cells. So a more sensitive approach is needed to succeed at this problem. For comparison purposes, we will use our previous findings using traditional machine learning applications.

We want to find alternative tools to analyze and identify the pattern directly from the time image. Train a deep learning network in the time domain in order to classify or diagnose patterns. So, we will use different techniques that each researcher will implement.

1. We establish a baseline using Fully connected networks in order to compare accuracy to other procedures.
2. Using recurrent neural networks we can use this type of data to input into the recurrent networks. Transforming this data into ~ 18 bins where the mean and standard deviation for 10 data points at a

time, each bin would be forced into a recurrent neural network. LSTM seem very popular in this space. We will read and follow [this paper](#).

3. We will explore different methods to analyze the data by splitting our data into training, testing, and validation sets and training some possible models discussed in class. The [second paper](#) we will follow points to the use of CNN's used in time series analysis which we will also use to guide the third part of the project. The goal is to predict the response of the cell with enough accuracy for that we can rely on the classification for those which show response.