Washbourne Lab Project for Machine Learning Identification of Synapses

Sarah Stednitz/Phil Washbourne

Background:
Neurons are the cells that perform computations in the brain, and they do this by communicating via specialized structures called synapses. The correct number, function, and location of synapses is critical for the appropriate function of the nervous system. Unsurprisingly, disruptions to many synaptic genes are associated with neurodevelopmental disorders, such as autism and schizophrenia. The Washbourne lab is broadly interested in understanding how changes in genes alter synapses and ultimately behavior.

What's the problem:
A critical part of this research is our ability to visually identify synapses in brain tissue. We achieve this using immunolabeling, which allows us to fluorescently tag synaptic proteins and image them using microscopy. A single image of a small section of brain tissue can contain many hundreds, if not thousands, of synapses that currently must be manually annotated. Ideally, an automated, unbiased approach would eliminate the need for manual annotation. However, there are distinct challenges presented by these types of images that prevent traditional computer vision techniques from being wholly effective:
1.) Synapses are very small and often difficult for a computer to distinguish from background noise.
2.) Variability in staining results from different antibodies, tissue quality, or even from experiment-to-experiment, causing changes in brightness in both the signal and the noise.
3.) Synapses are not uniformly bright even within the same image, so using a single threshold to binarize an image (required of most computer vision methods) can cause undersampling on one extreme and no ability to distinguish individual puncta on the other.

We believe that machine learning techniques could be applied to overcome these challenges. Because object recognition algorithms are capable of taking into account the textures of features (eg a bright spot surrounded by a dimmer halo) an appropriately trained network may be able to identify synapses among a diverse set of conditions as well, if not better, than manual annotation.

What you'd be doing:
1.) Develop a training set of manually annotated synapses in the correct format, using a variety of image qualities.
2.) Using Mask-RCNN or a similar implementation, train a neural network to automatically identify the hundreds of individual synapses.
3.) Validate the success of the system by comparing it to manual annotations by our intrepid undergraduates.

And in a perfect world,

4.) Develop a usable pipeline that supports multichannel images, identifying synapses in each channel, and further identifying areas where pre- and post-synaptic markers are colocalized.