Generation of Classified Image Libraries to Train Machine Learning Algorithms to Identify Different Marine Phytoplankton

Banik, Sreya (School: Lincoln Park Academy)

The purpose of this in silico experiment was to classify images using a machine learning algorithm. Specifically a neural network (NN). The image library was obtained through FlowCam, which group the pictures all into one big image. These large images were classified based on date and location in the Indian River Lagoon. Therefore ImageJ was used to crop the pictures from the larger images and sorted into 5 categories: Diatom, Dirt, Ciliate, Dinoflagellate, and Miscellaneous. After classifying and cropping the particles, a code was used to run a neural network to train a machine to be able to categorize the FlowCam images. Next, the neural network code was made in reference to the book, "Neural Networks and Deep Learning". After writing the code, the images needed to be rescaled using ImageJ because the NN will only work if all the images are the same size. Certain parameters are set to analyze particle size and shape accurately, this will be the independent variable. The results from the code, which were graphs, show the accuracy over the ranges of different input parameters (one graph for each parameter) and this will show us what parameter values result in the most accurate classifications. Among the parameters that were changed within the code, we have deduced that when the learning rate is 0.1, hidden layer 1 is 190, and hidden layer 2 is 95, the maximum accuracy of the code is 88.060%. This helps us conclude that it is possible to create a neural network that will classify these images.

Awards Won: