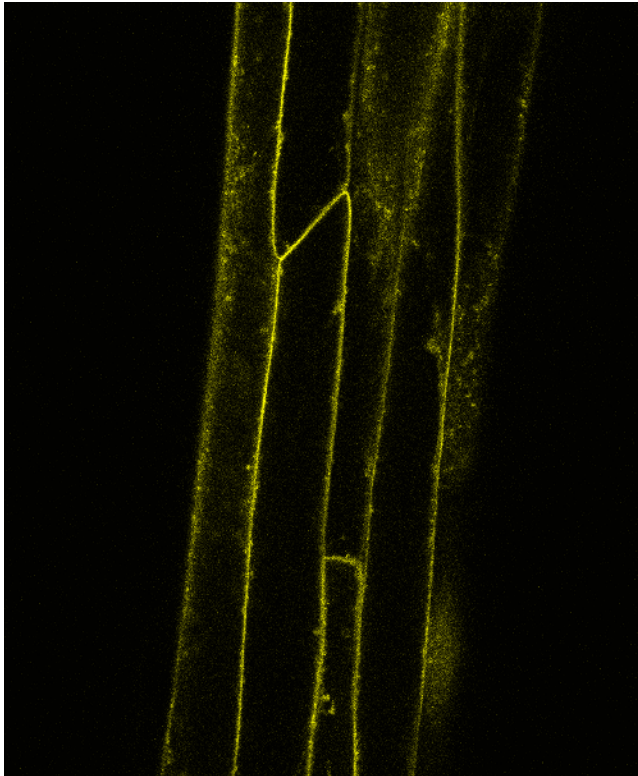


Applying Machine Learning to Analyze Protein Distribution in Plant Cells



- **By:** Jiayi Li
- **Major:** Biology
- **Faculty Advisor:** Alan Jones
- **Faculty Department:** Biology

Research Question:

How to automatically quantify the protein AtRGS1-YFP in the cytoplasm and on the membrane of *Arabidopsis thaliana* hypocotyl epidermal cells.

Why does it matter?

AtRGS1 is a membrane protein in *Arabidopsis thaliana*, which will be internalized inside the cell under some treatment like the presence of D-glucose, and it is related to signaling. So, analyzing the internalization of this protein quantitatively will help us understand this process better. People attached the AtRGS1 with the yellow fluorescent protein so that they can visualize the protein under a microscope, and then manually quantify the fluorescence intensity inside the cell as well as on the cell membrane by some software like ImageJ. However, it is time-consuming work. So, how to do the measurement automatically? We come up with the idea of applying Deep Learning (a type of machine learning) to distinguish different types of regions (e.g. cell membranes) in the image to speed up this image analysis.

Results

- Our Convolutional Neural Network (CNN) shows 74% accuracy of image segmentation (distinguishing region's type in the image).
- Current results suggest the quantification algorithms, based on the outputs of CNN, may perform well compared to the manual quantification results.
- The accuracy of our CNN need to be improved. And if software will be built based on our work in the future, it will be very helpful and easy for users to do the quantification job.