Project C1 – Leaf Wilting Detection in Soybean

Dr. Edgar Lobaton (ECE-NCSU), Dr. Anna Locke (USDA), Dr. Chris Reberg-Horton (CSSC-NCSU), Dr. Paula Ramos (CSSC-NCSU)

Motivation. Weather stress and climate change are major challenges for agriculture. Drought tolerant soybean varieties are needed but breeding for drought tolerance in soybean, is difficult. Soybean breeders evaluate thousands of field plots annually in order to select the best plants to advance for further testing or to release as a new variety. The most widely used indicator of drought tolerance/sensitivity in soybean is leaf wilting during periods of water stress[1], and breeders collect this data in the most low-tech way imaginable — walking the field with a notebook and writing down ratings to indicate how wilted each field plot looks [1,2].

Cameras capturing the wilting images at Sandhills Research Station, Jackson Springs, NC, 30 August 2019

There are some limitations to this strategy. Most obviously, it's time consuming. It's impossible for a breeder to rate all their plots in one day. On top of that, field sites are often located far away from the office, meaning that a breeder generally must drop everything and devote an entire day to collecting wilting ratings for just a subset of their plots. We also don't know when is best to collect wilting ratings: are the data more useful during periods of mild stress, moderate stress, or severe stress? What time of day or set of atmospheric conditions provides the most useful wilting information? How are leaf wilting ratings connected to plant physiology? These questions are hard to answer with manual annotations.

Our goal is to use temporally high-throughput image data, which represent two different soybean genotypes responding to a wide range of atmospheric and soil moisture conditions, as a training dataset to automate leaf wilting ratings. With automated ratings, we could generate larger datasets that would allow us to answer some of the questions above. We also eventually want to automate leaf wilting ratings for hundreds or thousands of plots across a field, which is a spatially high throughput application.

Data. Images from soybean crops at various times of the day using the configuration above are captured. The annotations corresponding to the level of wilting observed for each image are provided. See the "Illustration of Labels in Data" section for examples.

Objectives. This project will have two parts. In the first phase (Project C1), you are welcome to try any machine learning approach for prediction using either hand-crafted features from computer vision, or data-driven features (either using transfer learning or training networks from scratch) to predict the labels. See [3] for ideas of what can be done when you don't have much data for training a network

[1] In this context, water stress means some combination of *not enough soil moisture* + *hot, dry air*, which leads to more water being lost from the leaves than can be replaced from the soil

from scratch. Keep in mind that you will have limited time, so training of complex networks may not be possible. This will serve as your baseline for the second part of the project. We will provide you a test set for which you will turn in predictions in a format to be specified soon. The results from all the teams will be shared to the class in order to give you an idea of what to expect for performance. For the second phase, you will be expected to implement more complex neural networks for learning features from the data and improving on your results from the first part of this project. The groups with the best performance on this second phase will receive extra credit.

Deliverables:

1. Predictions for the provided test set
2. Technical report
3. name.csv file containing the unityID of group members

For your report, you should follow the following guidelines:

1. The length should not be more than 3 pages long. No introduction or abstract are needed. Your report should have three sections.
2. Section 1 - Methodology: Include a description of your approach including citations to papers describing the methodology, and references to toolboxes (and specific functions) used for the implementation.
3. Section 2 – Model Training and Hyperparameter Selection: Provide details about your procedure for hyper-parameter tuning.
4. Section 3 – Evaluation: Perform an evaluation of your method including error metrics. We will be giving you a test set for which you will give us your prediction for evaluation.
5. Include several plots showing results of your inference and the groundtruth.

Remember that you are welcome try some complex models for prediction, but this first round of the project is only meant to provide a baseline so classical machine learning techniques are acceptable.

References:

[1] King AC, Purcell LC, and Brye KR (2009) Differential wilting among soybean genotypes in response to water deficit. Crop Science 49:290-298. doi: 10.2135/cropsci2008.04.0219

[2] Pathan, SM et al. (2014) Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. Journal of Agronomy and Crop Science 200:231-236. doi: 10.1111/jac.12053

[3] B. Zhong, Q. Ge, B. Kanakiya, R. Mitra, T. Marchitto, E. Lobaton, “A Comparative Study of Image Classification Algorithms for Foraminifera Identification,” IEEE Symp. Series on Computational Intelligence (SSCI), 2017.

Illustration of Labels in Data:

0 – No wilting

1 - Leaflets folding inward at secondary pulvinus, no turgor loss in leaflets or petioles

2 - Slight leaflet or petiole turgor loss in upper canopy

3 - Moderate turgor loss in upper canopy

4 - Severe turgor loss throughout canopy