50.039 Theory and Practice of Deep Learning
Project Proposal

Wang Tianduo 1002963, Wu Tianyu 1002961, Luo Yifan 1002975

April 11, 2020

1 Problem Statement

[1] Misdiagnosis of the many diseases impacting agricultural crops can lead to misuse of chemicals leading to the emergence of resistant pathogen strains, increased input costs, and more outbreaks with significant economic loss and environmental impacts. Current disease diagnosis based on human scouting is time-consuming and expensive, and although computer-vision based models have the promise to increase efficiency, the great variance in symptoms due to age of infected tissues, genetic variations, and light conditions within trees decreases the accuracy of detection.

2 Objectives

Objectives of 'Plant Pathology Challenge' are to train a model using images of training dataset to

1. Accurately classify a given image from testing dataset into different diseased category or a healthy leaf
2. Accurately distinguish between many diseases, sometimes more than one on a single leaf
3. Deal with rare classes and novel symptoms
4. Address depth perception—angle, light, shade, physiological age of the leaf
5. Incorporate expert knowledge in identification, annotation, quantification, and guiding computer vision to search for relevant features during learning

3 Expected Input and Output

Given a photo of an apple leaf, can you accurately assess its health? This competition will challenge you to distinguish between leaves which are healthy, those which are infected with apple rust, those that have apple scab, and those with more than one disease.

Files

1. train.csv
   - image_id: the foreign key for the parquet files
   - combinations: one of the target labels
   - healthy: one of the target labels
   - rust: one of the target labels
   - scab: one of the target labels
2. images/
   A folder containing the train and test images, in jpg format.

[1] This is a Kaggle Competition and you may find more details at https://www.kaggle.com/c/plant-pathology-2020-fgvc7/overview/description

3. test.csv
   - image_id: the foreign key for the parquet files

4. sample_submission.csv
   - image_id: the foreign key for the parquet files
   - combinations: one of the target labels
   - healthy: one of the target labels
   - rust: one of the target labels
   - scab: one of the target labels

4 Dataset

Provided by Cornell Initiative for Digital Agriculture (CIDA). This competition is part of the Fine-Grained Visual Categorization FGVC7 workshop at the CVPR 2020

5 Team members

Wang Tianduo 1002963
Wu Tianyu 1002961
Luo Yifan 1002975

6 Deliverables

1. Code for training
2. Code for deployment
3. An app with GUI to demonstrate the result
4. A detailed report covered everything we do
5. Team members and their contributions
6. 5-min video persentation