Ryan Slattery

Cassava Leaf Disease Classification
Capstone Final Report

Problem Statement:

"As the second-largest provider of carbohydrates in Africa, cassava is a key food security crop grown by smallholder farmers because it can withstand harsh conditions. At least 80% of household farms in Sub-Saharan Africa grow this starchy root, but viral diseases are major sources of poor yields. With the help of data science, it may be possible to identify common diseases so they can be treated.

Existing methods of disease detection require farmers to solicit the help of government-funded agricultural experts to visually inspect and diagnose the plants. This suffers from being labor-intensive, low-supply and costly." (Cassava Leaf Disease Classification - Overview)

Given this, we'll be training a deep learning based image classifier to sort images into five categories: health, cassava bacterial blight, cassava brown streak disease, cassava green mottle, and cassava mosaic disease.

Data Cleaning and Preparation:

As a first step, I verified that all images were accounted for and all had labels. Next, I had to make my train/validation split. Then I created data loaders that apply transformations to the data when the data is loaded during training. These transformations include normalization as well as random mirroring, rotation, and cropping. Then I created an option to use weighted data loaders so that classes would appear with equal frequency during training, though I later decided to keep this feature disabled because it negatively impacted accuracy.

Models:

I fine-tuned several vision transformer variants, replacing the classifier head with either one or two linear layers. I used Weights and Biases to do a grid search over several hyperparameters.

Explored hyperparameters include the following:

ViT size: tiny, small, base, or large (all models took an input resolution of 384x384 and used image patches of size 16x16)
Learning rate: 3e-4 or 1e-4
Weight decay (L2 norm): 0 or 1e-5
Classifier head: 32 dimensional hidden layer or no hidden layer

More technical details about these can be found in model_metrics.txt

Findings:

The best ViT model had a validation accuracy of 85.8%.

More generally, the higher learning rate and including a hidden layer helped. Additionally, weight decay had little impact. Finally, larger transformers dramatically improved performance though the performance difference between 'base' and 'large' was small compared to previous size increases.

Recommendations:

1) Our model is highly sensitive to cassava mosaic disease. A negative result means we can say with near certainty that the plant is not infected.
2) Our model is also highly precise when detecting cassava green mottle. A positive result means we can say with near certainty that the plant is infected.
3) Our model otherwise struggles and shouldn't be used as a standalone method to determine disease status. Multiple images of one plant may help alert the user to incorrect results but how effective this is requires further research.

Further Research:

As mentioned, feeding the model multiple images of one plant and looking for a pattern in the results may help. Unfortunately, we don't have multiple images of individual plants in the current version of this dataset or if we do then we don't have labels indicating this.

Supplement:

See losses and accuracy over time for different models here:
https://wandb.ai/ryulord/cassava-leaf-disease-classifier/reports/Final-Report-Supplement--VmlldzoxMDU2NDg4