

Capstone 3 Project Proposal:

Cassava Leaf Disease Classification

Ryan Slattery

Problem:

We are trying to classify images of cassava leaves to determine if they are healthy or which of four particular diseases they have.

Context:

“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)

Criteria of Success:

Kaggle measures the efficacy of a model for this competition by its accuracy. The best solutions so far have about 90% accuracy so beating this would be a clear success but getting near this value would also be acceptable.

Scope of Solution Space:

We'll be creating one or more models with PyTorch for the classification of cassava leaf images.

Constraints on Solution Space:

“Effective solutions for farmers must perform well under significant constraints, since African farmers may only have access to mobile-quality cameras with low-bandwidth.” (Cassava Leaf Disease Classification - Overview)

Key Data Sources:

As this is a Kaggle competition, all necessary data is provided at <https://www.kaggle.com/c/cassava-leaf-disease-classification/data>

Citations:

“Cassava Leaf Disease Classification - Overview” *Kaggle*, Kaggle,
<https://www.kaggle.com/c/cassava-leaf-disease-classification/overview>.