

Cassava Image Classification


Melanie Butler - Final Springboard Capstone

[Google Slide Deck](#)

Agbeli (There is life)

- Ewe word for cassava
- Food security crop in vulnerable areas
- Robust to difficult climates
- Primary food staple for over 800 million people
- Potential biofuel source - extra income for farmers


CASSAVA





CASSAVA IS A **HIGH-STARCH CROP** (LIKE POTATOES) THAT THRIVES IN POOR SOIL


CASSAVA HAS MANY USES:

FOOD

FEED


FUEL

PRODUCTS



CASSAVA SUSTAINS 600 MILLION PEOPLE ACROSS THE **WORLD**

SOME PEOPLE IN DEVELOPING COUNTRIES DEPEND ON CASSAVA FOR **HALF** OF THEIR CALORIES



The Problem

- Disease and pestilence:
 - Cassava Mosaic Disease (CMD)
 - Cassava Brown Streak Disease (CBSD)
 - Cassava Bacterial Blight (CBB)
 - Cassava Green Mottle (CGM)
- Yield reduction of 10% to 70%, up to 100% with CBSD



Mitigation Strategies

- Cassava is propagated with plant cuttings - spreads disease
- Quarantine / burn diseased plants
- Critical to identify disease plants for separation



Project Goals and Metrics

1. Develop a model capable of classifying an image of a cassava plant into one of four disease categories or as healthy.
 - a. False disease classification → unwarranted destruction of valuable crops
 - b. False healthy classification → propagation of disease
 - c. Both are unacceptable - overall accuracy as metric
2. Deploy the model in an API that farmers could use to aid in their crop management decisions.

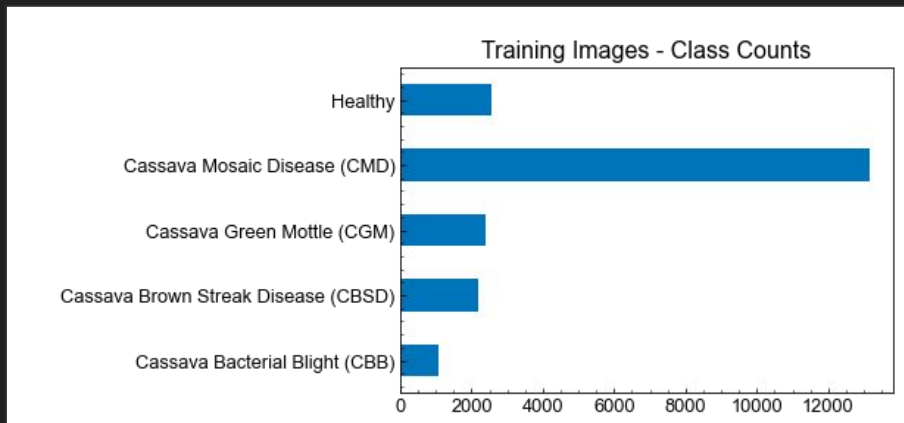
Data Source

- National Crops Resources Research Institute / AI lab at Makerere University, Kampala
- Crowd-sourced images from farmers in Uganda
- 21,367 Images
- Four disease categories, one healthy category



Preprocessing

- Highly class- imbalanced



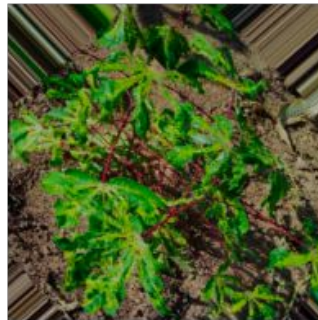
- Initial training - resampled 2000 images per class



Preprocessing

- Train / Val split (85:15)
- Hidden test set - Kaggle (~10,000 images)
- Random image augmentation:
 - Basic - variability expected from images in the field
 - Rotation, Brightness, Zoom
 - H / W / Channel Shift
 - Shearing
- Generated each batch - regularizing effect

Cassava Mosaic Disease (CMD)



Healthy



Cassava Bacterial Blight (CBB) Cassava Mosaic Disease (CMD)



Preprocessing

- Cutmix Image

Augmentation:

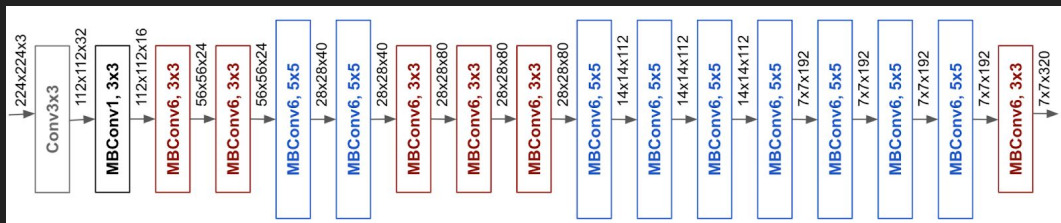
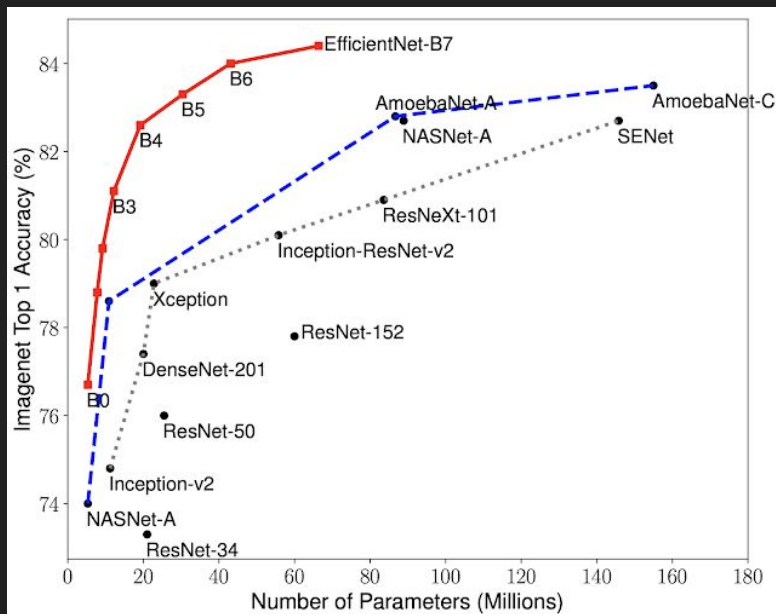
- Randomly crop two images together
- Label weighted according to area of each image
- Better model generalization, object localization
- Similar to dropout effect, more efficient pixel use

Cassava Brown Streak Disease (CBSD) / 72%
Cassava Mosaic Disease (CMD) / 28%



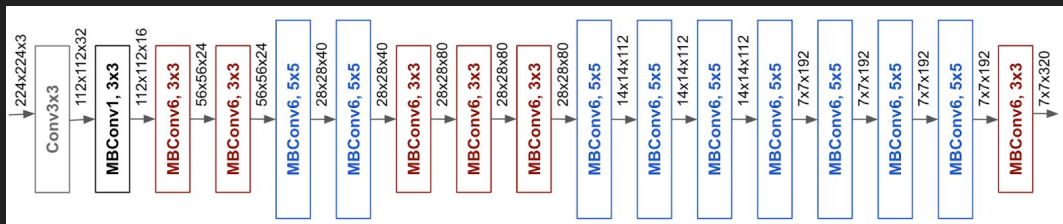
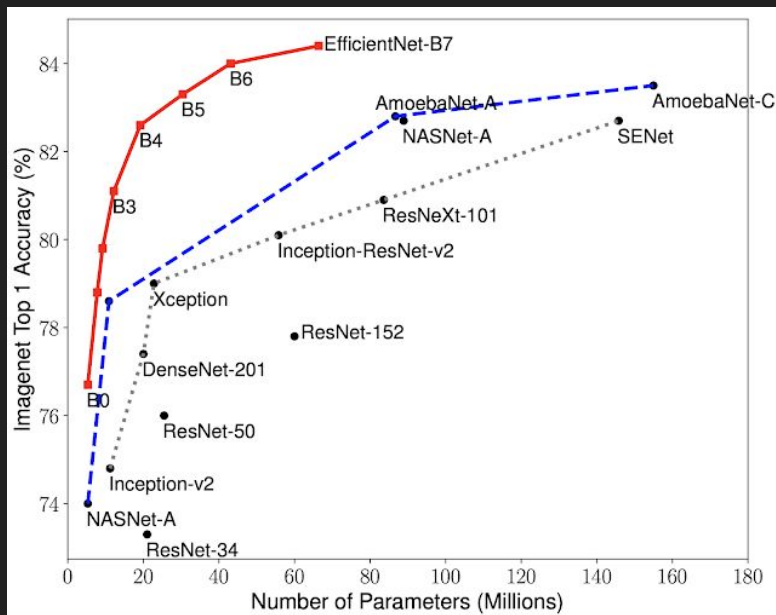
Model Architecture

- Keras Applications
- Preliminary training - balanced, basic augmentation:
 - ResNet50
 - ResNet152v2
 - Xception
 - **EfficientNet-B4**



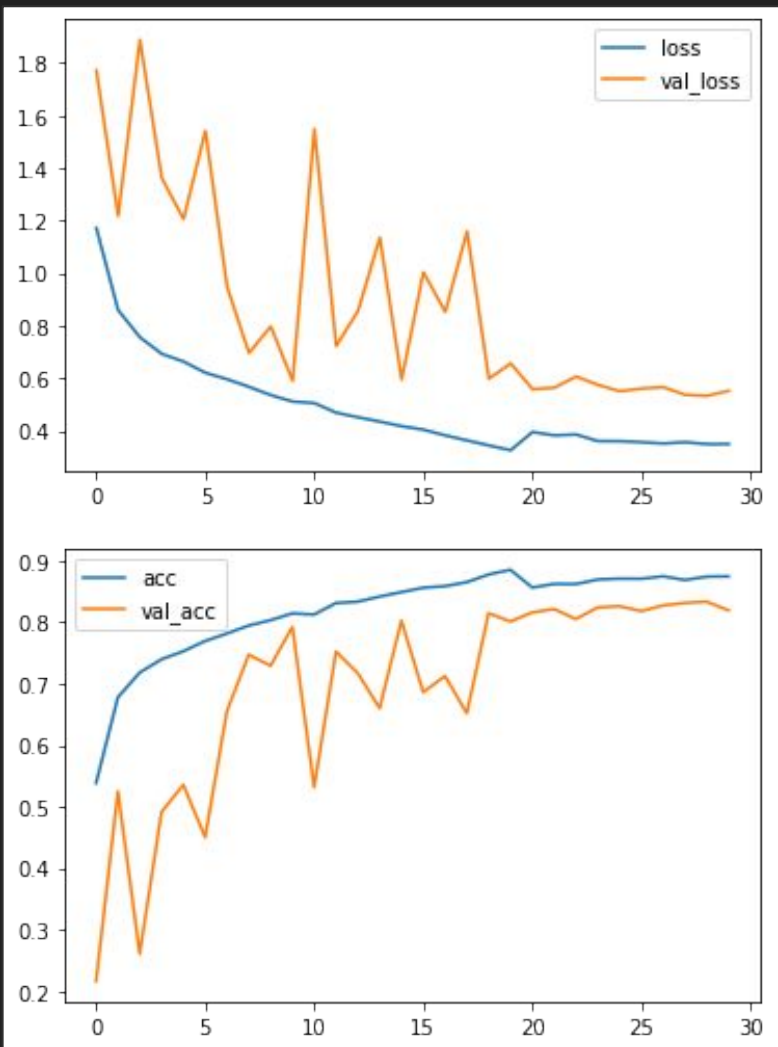
Model Architecture

- EfficientNet-B4 base
- Scaling factor
 - Optimize depth / width / resolution given resource constraints
- 2D Global Average Pooling
- Dropout (rate 0.2 \rightarrow 0.1)
- Dense output (softmax)

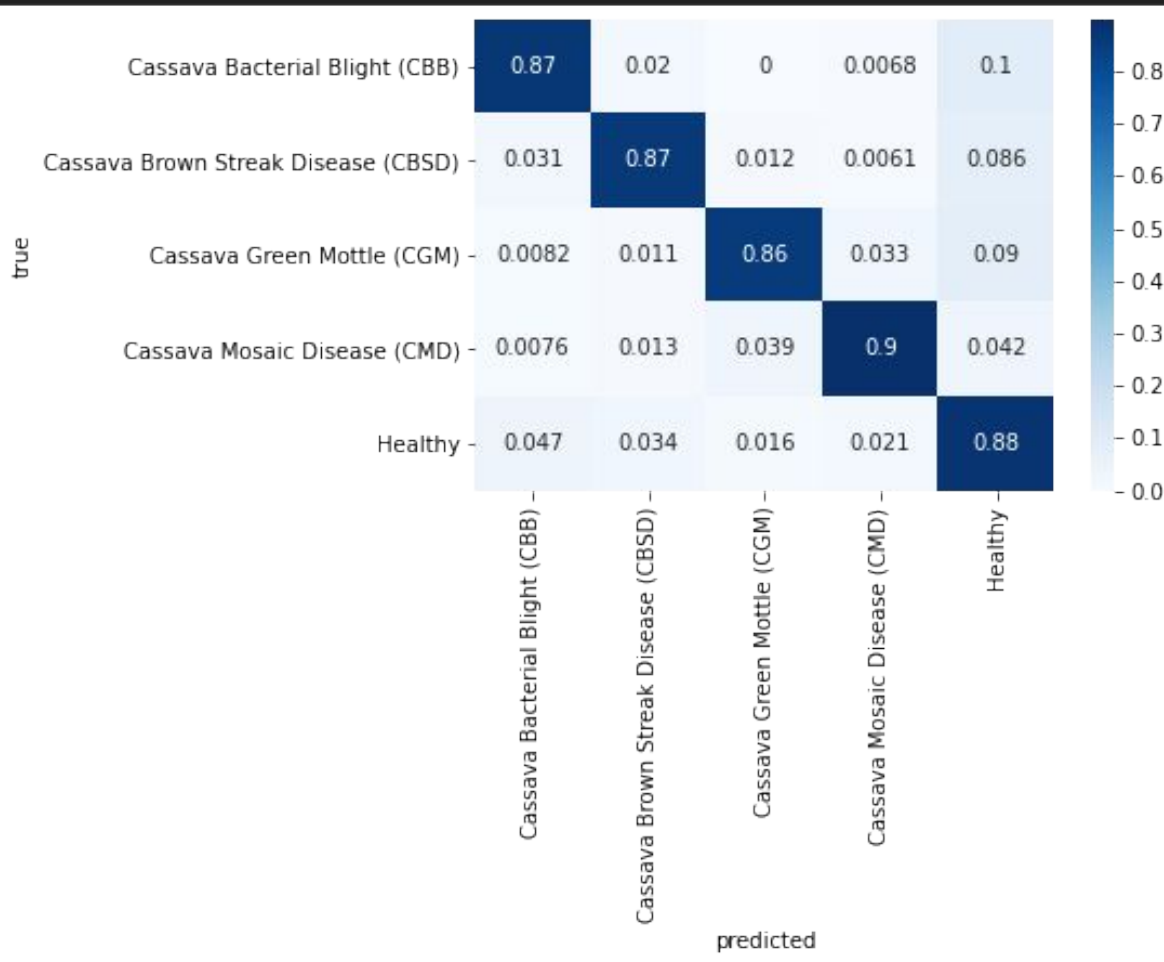


Model Training

- Adamax optimizer
- Initial LR of 0.001, reduce on plateau callback
- First 30 epochs -
 - balanced training
 - basic augmentation
 - first two blocks frozen
- Unfroze layers
- Cutmix augmentation
- Decreased dropout to 0.1



Model Evaluation



Model Evaluation

Weakness

	Precision	Recall	F1-Score	Support
Cassava Bacterial Blight	0.74	0.87	0.8	148
Cassava Brown Streak Disease	0.86	0.87	0.86	326
Cassava Green Mottle	0.78	0.86	0.82	365
Cassava Mosaic Disease	0.99	0.9	0.94	1984
Healthy	0.68	0.88	0.77	386
Accuracy			0.89	3209
Macro Average	0.81	0.88	0.84	3209
Weighted Average	0.9	0.89	0.89	3209

Test Accuracy on Kaggle - 87.8%



Can you tell the difference?

The Problem

Cassava, a woody shrub native to South America, is one of the most important crops of the tropics, behind rice and corn. It is grown in all tropical and subtropical ecological zones where other crops such as sugarcane, banana, and mango are grown. More than 800 million people in the world consume cassava products.

Cassava is a major source of food and income for millions of people in the tropics. However, it is highly susceptible to Cassava Brown Streak Disease (CBSD), a viral disease that causes significant losses in yield and quality. CBSD is caused by the Cassava Brown Streak Virus (CBSDV) and is transmitted by the cassava mealybug. The disease is characterized by yellowing and mottling of the leaves, and by the appearance of brown streaks in the roots. CBSD is a major constraint to cassava production in many countries, particularly in Africa. It is estimated that CBSD causes losses of up to 100% in susceptible varieties.

One Possible Solution

An EfficientNet-B4 image classifier has been developed that can achieve 87% accuracy in classifying images of cassava plants as being either healthy or having one of the four diseases described above. The model was trained using a Kaggle dataset of 21,367 images crowdsourced from farmers in Uganda and annotated by experts at the National Crops Resources Research Institute (NaCRRI) in collaboration with the AI lab at Makerere University, Kampala. For more information, please refer to the project Github repository, linked at the top of the page.

Try uploading your own image of a cassava plant here.

Select a file to upload

Label: Cassava Brown Streak Disease (CBSD) Model Probability: 87.0

No file chosen

References

1. <https://www.apsnet.org/edcenter/apsnetfeatures/Pages/cassava.aspx>
2. <https://www.bangkokpost.com/business/2042767/e20-set-to-be-leading-fuel-by-july-2021>
3. https://plantvillage.psu.edu/topics/cassava-manioc/infos/diseases_and_pests_description_uses_propagation
4. <https://www.kaggle.com/c/cassava-leaf-disease-classification/overview>

Model Deployment

- Website Developed on Flask
- Upload Image
- Receive Classification / Model Probability

Model Output

Future Work

- Improve model performance w.r.t Healthy class
- Binary classification appropriate for application - disease response is generally the same regardless of disease:
 - a. Healthy - no action needed
 - b. Diseased (CBB, CBSD, CMD, CGM) - quarantine / burning
- Train binary classifier, able to tune for desired precision / recall

Thank You

Thank you for your attention, feedback, questions!



**Mentor -
Giovanni Bruner**