Development of Machine Learning Algorithm for Classifying Banana Diseases from Leaf Image.

Davis David,Fredrick Apina,Ibrahimu Mtandu,Waziri Shebogholo,
Dr.Dina Machuve & Sophia Sanga

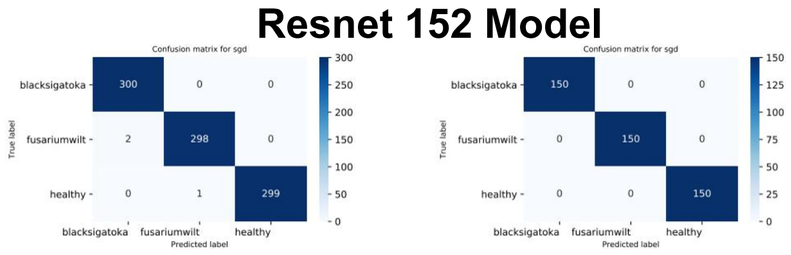
BACKGROUND & MOTIVATION

- Banana is an important crop for the livelihoods of people in the Great Lakes region of East and Central Africa (ECA) including Uganda, Rwanda, Burundi, Western Kenya. North western Tanzania and DR Congo.
- This crop contribute to national economies and play a key role in food security. In Tanzania.
- Studies indicated that banana is among the crops that are affected by several fungal diseases most commonly being black Sigatoka and Fusarium wilt.
- This project aim to develop a machine learning algorithm for early detection of banana disease to help smallholders' farmers and extension officers in developing better manage and control the incidence of the banana fungal diseases.

METHODS

- We evaluated the applicability of transfer learning from a deep Convolutional neural network (CNN) model for the banana image datasets.
- We decided to use three different resnet models (resnet18,resnet50 and resnet152) to find out which one can give out best results.
- We conducted hyper-parameter search with aim to find best hyper-parameter that would give better performance.
- These parameters are batch size, optimizer, momentum, learning rate, weight decay and model architecture.The objective was to increase true positives and decrease false negatives.

Resnet 152 Model



(a) Validation confusion matrix (b) Test confusion matrix

Performance Metrics

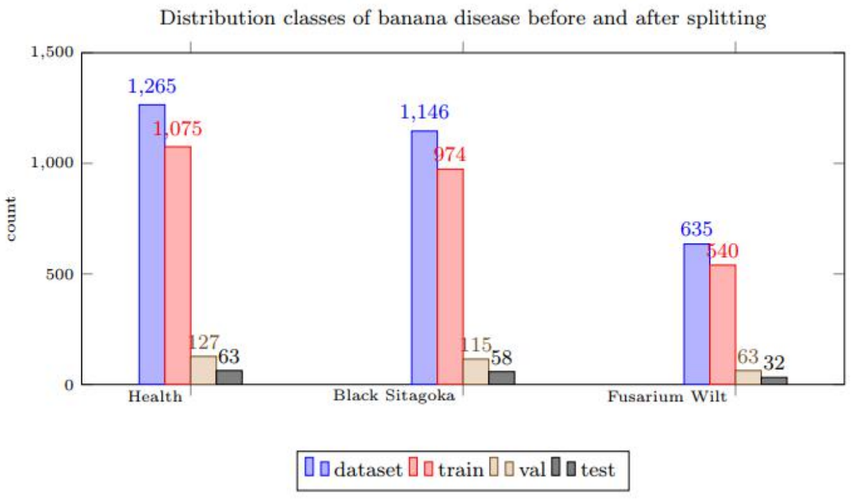
Model Architecture	Valid set (Accuracy)	Valid set (F1 score)	Test set (Accuracy)	Test Set (F1 Score)
Resent18	0.997	0.996	0.998	0.9977
Resnet50	0.997	0.996	0.998	0.9977
Resnet152	0.997	0.996	1	1

DATA & PREPROCESSING

The collected dataset contain images with two categories of banana disease namely Black Sigatoka and Fusarium and one category of healthy banana leaves.

(a) black Sigatoka (b) Fusarium Wilt (c) Health

- We used different augmentation techniques for training, validation and testing data.
- For the training, each image was resized to size 255, randomly cropped at 244 ratio, horizontally and vertically flipped with 0:5 probability,
- For testing and validation image was only cropped at the center at 244 ratio.
- Image dataset was then splitted in the following ratio: 85% training set, 10% validation set, 5% test set.

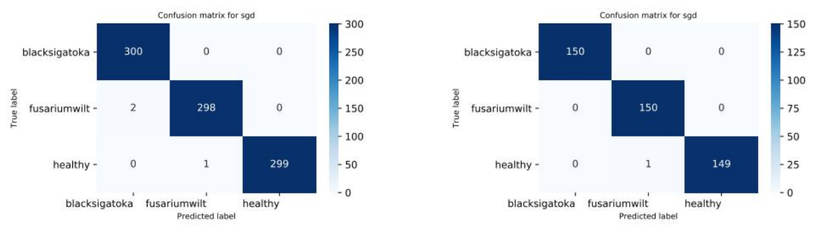


EXPERIMENTAL RESULTS

Hyper-parameters values that gave us best results

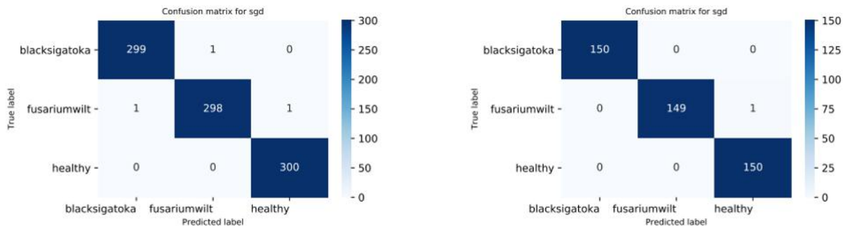
Parameter	Value
Optimizer	SGD
Batch-size	4
Learning rate	$1e^{-3}$
Weight decay	$1e^{-9}$
Model architecture	Resnet
Momentum	0.9

Resnet 18 Model



(a) Validation confusion matrix (b) Test confusion matrix

Resnet 50 Model



(a) Validation confusion matrix (b) Test confusion matrix

CONCLUSIONS

The best model is Resnet152 but Resnet18 can be further improved to increase the generalization performance by collecting more data from different regions in Tanzania (not only in Arusha and Mbeya).

REFERENCES

1. Godliver Owomugisha, John A. Quinn, Ernest Mwebaze, James Lwasa. Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease.
2. Vijai Singh, A.K. Misra.Detection of plant leaf diseases using image segmentation and soft computing techniques

CONTACT: d.david@parrotai.co.tz | +255 762 260 621 | www.parrotai.co.tz