Cassava Leaf Disease Classification

Group Name: Never Mind
Group Members: Millicent Omondi and Leonard Sanya

May 26, 2024

Introduction

Crop diseases are a threat to all types of cultivated plants, including cassava and other crops grown for human and animal consumption. These illnesses reduce gross yields, hamper the vigor and growth of affected plants, and deform plant outputs. [2].

Cassava is a perennial tuberous root crop. It thrives well in tropical regions, tolerates poor soil and drought, requires minimal input, but delivers higher outputs. Small-scale farmers mostly cultivate it, making it one of the most significant staple foods in developing countries. More than half of the world's production currently comes from Africa. [2].

Problem Statement

The major goal of this study was to assist farmers in recognizing healthy crops and distinguishing between four of the most prevalent cassava diseases, namely cmd, cgm, cbsd, and cbb, as the majority of farmers are unable to do so. We accomplished this by using a few computer vision algorithms to aid in image classification of the diseases. We made use of a Kaggle dataset that included images captured by Ugandan farmers of both good and diseased cassava leaves. [1].

Experiments and Results

Using various pretrained model architectures, including ResNet34, ResNet50, ResNet101, DenseNet161, MobileNet V2, and many more, we carried out multiple cassava leaf disease classifications. The results of the experiment indicate that the accuracy of the MobileNet V2 model architecture on the training and validation sets surpasses that of all other model architectures. Squeeze-and-excitation (SE) blocks, depthwise separable convolutions, inverted residuals, bottleneck design, and linear bottlenecks are the components of MobileNet V2 [3]. The stochastic gradient descent (SGD) optimizer, with a learning rate of 0.001, a momentum of 0.9, and a cross-entropy loss function, was used to train the model across 100 training epochs.

A 256 by 256 pixel resizing was applied to the input image. Our accuracy on the training and validation sets was 94.19% and 85.07%, respectively. Table 1 presents further findings from the experiments conducted with various model architectures.

Table 1: Experimentation result summary table

Model	**No. of Parameter**	**No. of Epochs**	**Training Loss**	**Training Accuracy**	**Validation Loss**	**Validation Accuracy**
ResNet34	21.8M	20	0.24478	0.9091	0.4589	0.8411
ResNet50	25.56M	20	0.4930	-	-	0.8400
ResNet101	44.5M	20	0.2618	0.9025	0.5285	0.8394
DenseNet161	28.7M	20	0.4518	-	-	0.8320
MobileNet V2	13.6M	100	0.1549	0.9419	0.4279	0.8507

Conclusion

In conclusion, based on our explored techniques, the MobileNet V2 architecture trained with SGD optimizer outperformed all the other model architectures. Conversely, we encountered the following issues:

- Difficulty in fine tuning the pretrained models,
- Too much training time for large models,
- Poor model generalization due to imbalanced data,

We recommend further studies to consider various transfer learning model fine-tuning strategies and the use of additional images to enhance the training data and improve the model generalization, while fine-tuning the outermost layers and retraining all the layers.

NOTE: We obtained a better model but not able to submit on kaggle since our slots had been field already.

References

[1] `https://arxiv.org/pdf/1908.02900.pdf`

[2] `https://www.mdpi.com/1999-4915/12/12/1388`

[3] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520).