# Classification of Apple-leaf diseases for FGVC8 Plant Pathology

{i.shaheen,amroadel,mahashatta}@aucegypt.edu

Ismail Shaheen, Amr Abdallah, Maha Shatta

CSCE Department, AUC The American University in Cairo

## ABSTRACT

Plant pathology detection have been one of the most active areas of research recently due to the increasing reliability of convolutional neural networks and its economic value.

This paper presents a proposed method for classifying diseases in apple-leaf with a focus on decreasing computational resources and maintaining a competitive accuracy. Various computer vision techniques have been used to help the proposed deep learning model, and the results of our experimentation show an accuracy of 72% on FGVC8 Plant Pathology.

## INTRODUCTION

Plant pathology is one of the most important fields and that is because its main responsibility is the diagnosis of crops diseases and utilizing various methodologies to control this damage. Image processing has been widely used to detect plant diseases to guarantee sustainable agriculture. This is because of its advantages in early capturing the disease and improving the productivity of the crops.

Multiple computer vision techniques were conducted such as smoothing, segmentation or augmentation that are followed by a feature extraction model such as convolutional neural networks (CNN) and then classification layers. This has resulted in great accuracy that is above 90%. In addition, various challenges and competitions have been conducted such as FGVC8 plant pathology which concentrates more on apple leaf diseases. The main challenge in this competition is that its dataset has multiple disease symptoms variations for each leaf due to differences in leaf color, leaf morphology, age of infected tissues, various light brightness in addition to the non-homogeneous and analogous to leaf-color background .

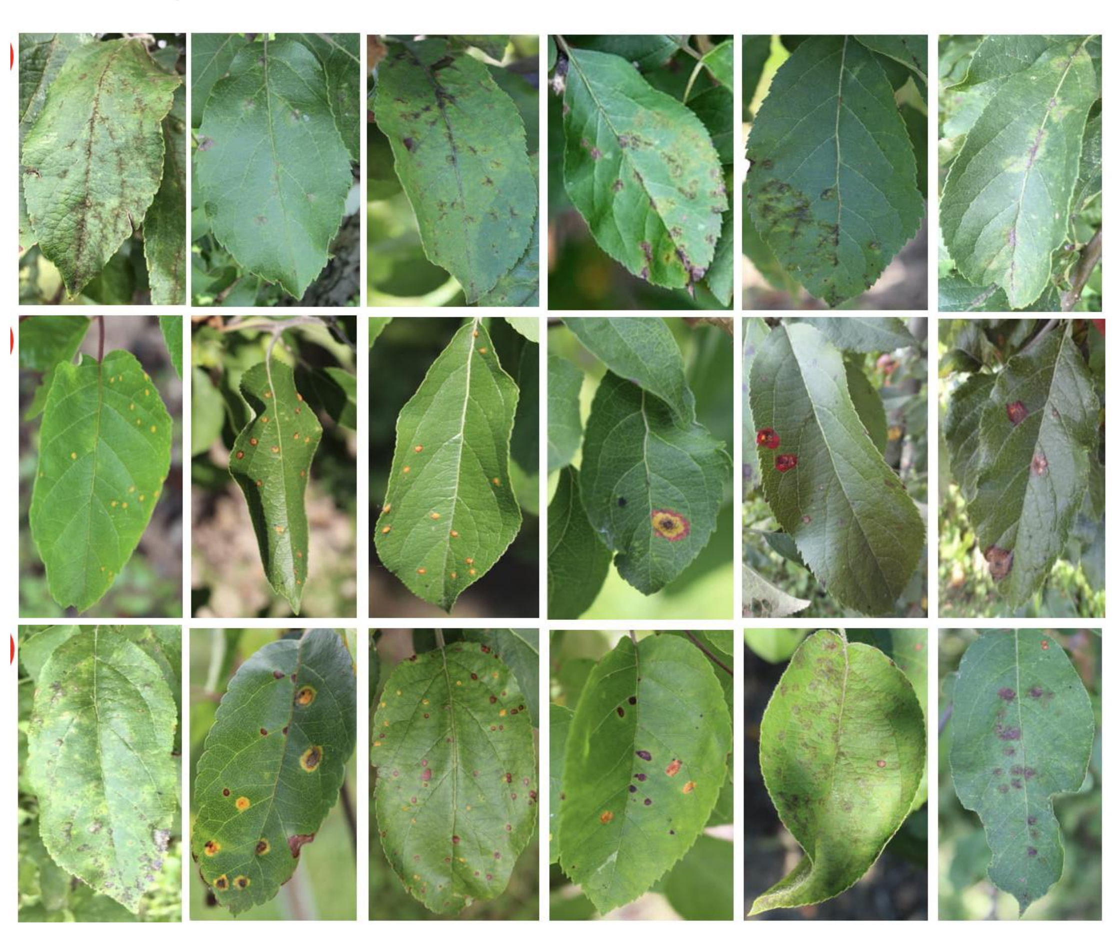


Figure 1 – Different Apple leaf pathologies

## PROPOSED SOLUTIONS

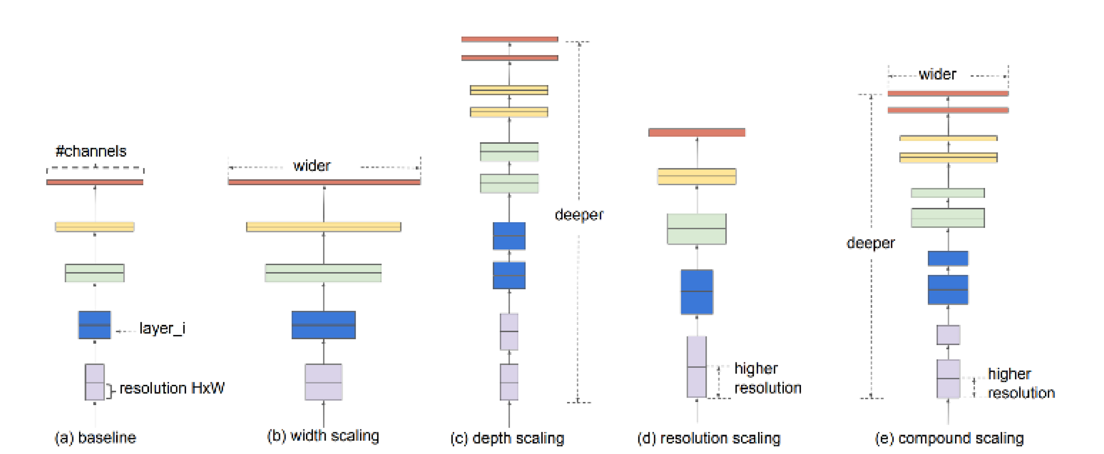
Our problem statement is to classify the leaf images to decide if they have a disease or not. Our proposed solution combines computer vision techniques with deep learning to better approach the problem. The computer vision part is mainly applying filters and realistic transforms to the images to better help the deep learning model generalize and avoid overfitting. The details of this part will be elaborated in more details in the Data Augmentation section.

Our proposed deep learning architecture is based on EfficientNet. EfficientNet is a convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth/width/resolution using a compound coefficient.

The base EfficientNet-B0 network is based on the inverted bottleneck residual blocks of MobileNetV2, in addition to squeeze-and-excitation blocks.

EfficientNets also transfer well and achieve state-of-the-art accuracy on CIFAR-100 (91.7%), Flowers (98.8%), and 3 other transfer learning datasets, with an order of magnitude fewer parameters.

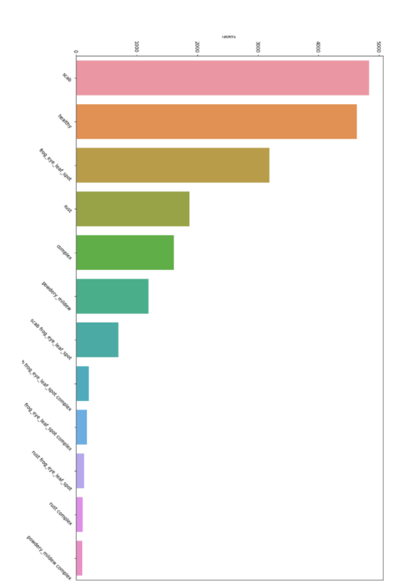
We used EfficientNet as our backbone by importing a pretrained model and removing its classification head. We then added dense layers to the end of the piplines that was trained independently at the begingin, then we used it to finetune the weights of the EfficientNet.

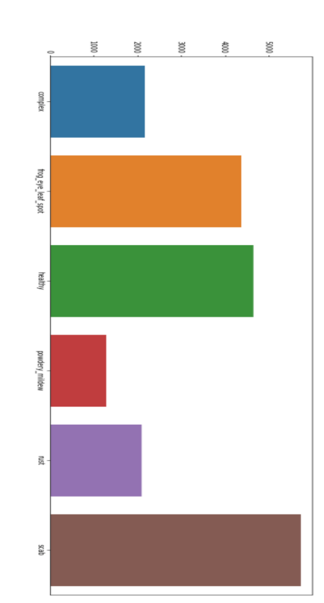



**Model Scaling.** (a) is a baseline network example; (b)-(d) are conventional scaling that only increases one dimension of network width, depth, or resolution. (e) is our proposed compound scaling method that uniformly scales all three dimensions with a fixed ratio.

## MODIFYING THE LABELS

In addition to data augmentation, we also tried to balance the dataset as shown in the graphs below.



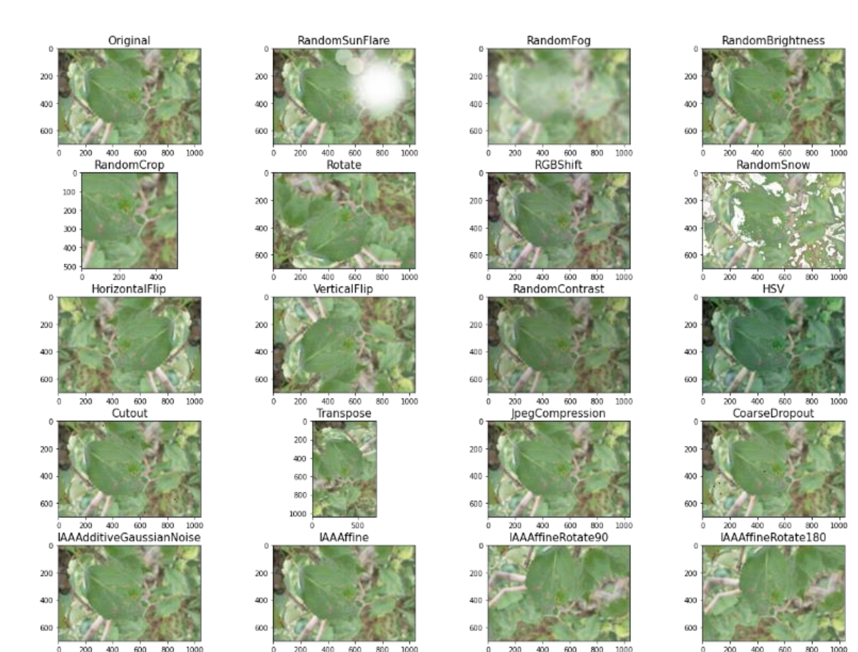


## DATASET

The Dataset used in this architecture is the dataset of FGVC8 Plant Pathology. The dataset contains about 20K high-quality RGB images of apple foliar diseases, including a large expert-annotated disease dataset. This dataset reflects real field scenarios by representing non-homogeneous backgrounds of leaf images taken at different maturity stages and at different times of day under different focal camera settings. The images are labels with 12 different labels, some of which are a mixture of several abstract labels.

Data augmentation is the process of applying random yet realistic transformation to the images in the dataset to increase its diversity. Data augmentation is a critical component of training a deep learning model. Artificially increasing the dataset helps reducing the probability of overfitting and improves model's generalization. Simple image transformation like rotation and flipping can already improve the performance of the model in most vision tasks. However, using task specific augmentation will provide better gains. In our case, task specific augmentation can be flares, random noise, hue transformation, …etc.

Different Types of Augmentations with Albumentations

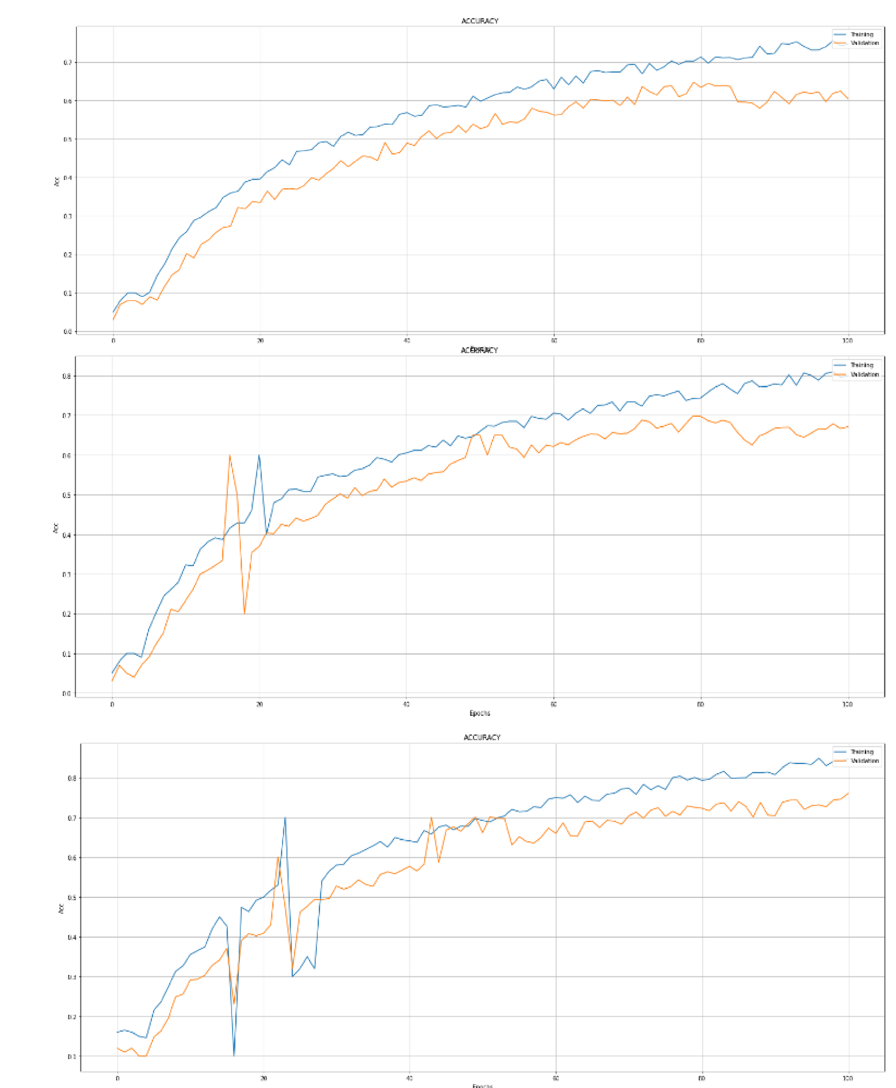


## EXPERIMENTAL RESULTS

We have conducted 3 experiments to choose the best performing model. Their training graphs can be shown below.

| Experiment | Accuracy |
|---|---|
| Without finetuning and augmentation | 63% |
| Without finetuning and with augmentation | 68% |
| With finetuning and augmentation | 72% |

Results of the three experiments



Graphs of the three experiments

## CONCLUSION

The results obtained by the experiments in this paper suggest that EfficientNet provides good results in vision problems; in addition, they show that data augmentation affect the results in a very desirable way as it increases the number of training examples. Our best achieving model yields accuracy of 72%, yet we assume that implementing the techniques recommended in the future work will improve the accuracy of the model.

## FUTURE WORK

There is a lot of potential in the data augmentation area, so one of our future strategies is to fine tune the augmentation to find the best combination the yields the best performance. An example of this is trying to include segmentation as one of our augmentations. Furthermore, special attention is a promising technique to increase the performance of vision models. Adding and attention layer to out model will make it easier for the CNN to focus on the important areas of the image. Finally, we will include ensemble to out pipeline as it will help our model avoid critical points while trying to converge.

## ACKNOWLEDGEMENTS

This work was supported by Dr. Mohamed Moustafa, AUC as well as AUC labs.