Plant Disease Classification Using MobileNetV2

Ahmed Aboutaleb, Ismael Elsharkawi, and Marwah Hisham
The American University in Cairo

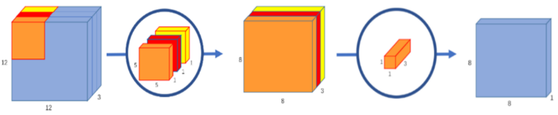
Abstract

Crop diseases is a major threat to the supply of food. As smartphone technology is becoming more widespread, there is a need for using lightweight Machine Learning algorithms. MobileNet V2, that uses depthwise separable convolution, is one of those. In this research, MobileNet V2 was used for plant disease classification on a 54,303 image dataset and it achieved and accuracy of 96.9%.

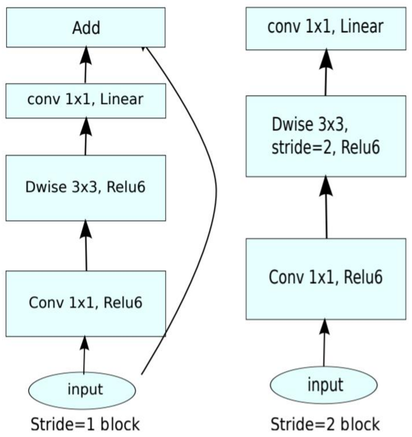
Literature Review

ResNet 50 achieved *99.44%* accuracy.
VGG 16 achieved an accuracy of *99.4%*.
Depthwise Separable Convolution is divided into 2 main parts.

- Depth wise convolution: made across each layer separately.
- Point wise convolution: made across all layers for one point only(1x1)



MobileNet V2, uses inverted residual blocks.



Methodology

Experiment A – Adding Only 4 Layers to the Base Model

After Adding the 4 layers, the total number of trainable parameters became 369,926. Layer details are summarized below:

Layer (type)	Output Shape	Param #
mobilenetv2_1.00_224 (Functi	(None, 7, 7, 1280)	2257984
conv2d_5 (Conv2D)	(None, 5, 5, 32)	368672
dropout_5 (Dropout)	(None, 5, 5, 32)	0
global_average_pooling2d_5 (	(None, 32)	0
dense_5 (Dense)	(None, 38)	1254

Total params: 2,627,910
Trainable params: 369,926
Non-trainable params: 2,257,984

Experiment B – Finetuned Experiment A model

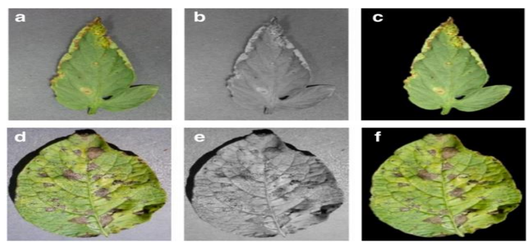
The last 50 layers of the base MobieNetV2 model were made trainable. The number of trainable parameters increased to 2,232,518.

Layer (type)	Output Shape	Param #
mobilenetv2_1.00_224 (Functi	(None, 7, 7, 1280)	2257984
conv2d (Conv2D)	(None, 5, 5, 32)	368672
dropout (Dropout)	(None, 5, 5, 32)	0
global_average_pooling2d (Gl	(None, 32)	0
dense (Dense)	(None, 38)	1254

Total params: 2,627,910
Trainable params: 2,232,518
Non-trainable params: 395,392

Dataset Used

Plant Village has 54,303 images with 38 classes. Raw, grayscale and segmented images are available for the whole dataset.



Results

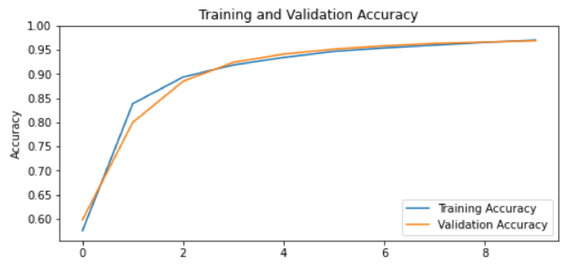
Experiment A

This experiment yields an accuracy of 92.819% and a cross entropy loss of 0.223 for segmented images. This indicates that MobileNet v2 is a good base model for architectures that are used to solve this problem.

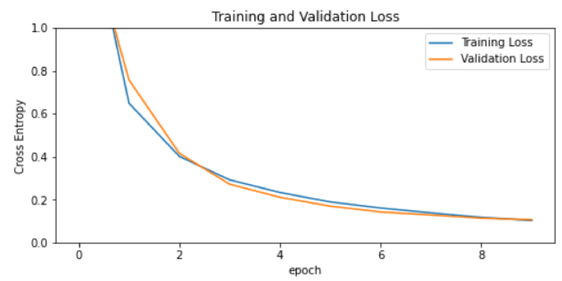
Experiment B

After finetuning the model from experiment 1, the results were significantly improved. The testing accuracy increased 96.91%, and the cross entropy loss dropped to 0.1007 for the non-segmented images. This is the best performing model since it improved the accuracy over our base experiment by 4.1% and its results are summarized below.

Accuracy for Experiment B with unsegmented images



Loss for Experiment B with unsegmented images



The Table Below Summarizes the results achieved in all experiments

Experiment	Dataset Used	Cross Entropy	Total Accuracy
A	Raw Images	0.231	92.34%
	Segmented Images	0.223	92.819%
B	Raw Images	***0.107***	***96.91%***
	Segmented Images	0.11	96.6%

The table above shows that training more layers from the MobileNet pre-trained layers improves the accuracy. However, using segmented images rather than raw images does not always improve the accuracy and if it improves it, the difference is not significant.

Conclusion

Concluding Remarks

The best model used was the one in experiment B with unsegmented images. Our model achieves an accuracy of 96.9%, which is lower than the state of the art, but it is more applicable to use in real life on low processing power mobile devices.

Recommendations

- Try with lower number of classes
- Fine tune even more
- Try MobileNet V3

References

Scan this code for the references