Peyton Wong

Plant Disease Detection using Transfer Learning in MobileNetV2 - A Report

The agricultural sector in Canada is raising concerns about declining crop production due to challenges with disease and infestation. Research has been done to identify and predict a crop's health by collecting image data on various plant leaves, gathered by farmers across the country. In *Plant Disease Detection and Classification by Deep Learning - A Review*, Li et al. states that, "early detection is the basis for effective prevention and control of plant diseases, and they play a vital role in the management and decision-making of agricultural production" (Li et al., 2021, p. 56683). By detecting early signs of disease and infestation through analyzing plant leaves, farmers can act accordingly in treating their crops. This research will focus on using transfer learning in MobileNetV2 to identify the state of a plant based on their leaves using the "Plant Disease Detection" dataset on Kaggle. I will walk through my steps in choosing transfer learning for multiclass classification, explain my findings, along with considerations that come with using transfer learning and this dataset.

Transfer learning holds prominents to multiclass classification like plant disease, according to Li et al., as data collection on plant leaves is rather limited due to budget (Li et al., 2021). Li et al. states that, "transfer learning enables the adaptation of pre-trained CNNs by retraining them with smaller datasets whose distribution is different from the larger datasets previously used to train the network from scratch" (Li et al., 2021, p. 56684). In context to the dataset, it records the plant leaves of five different crops which are apples, corn, bell peppers, potatoes and tomatoes. Each plant has a series of images that include their healthy and diseased states. The dataset holds approximately 35700 images, each folder can be classified as its own class (e.g: Apple__apple_scab, Potato__healthy, etc.). The dataset holds a total of 23 class labels, each crop group undergoing different diseases and infestations, which need to be identified in order for the pre-trained model to classify the image.

Layer (type)	Output Shape	Param #
input_layer_4 (InputLayer)	(None, 224, 224, 3)	0
mobilenetv2_1.00_224 (Functional)	(None, 1280)	2,257,984
dropout_1 (Dropout)	(None, 1280)	0
dense_1 (Dense)	(None, 23)	29,463

Total params: 2,287,447 (8.73 MB)
Trainable params: 29,463 (115.09 KB)
Non-trainable params: 2,257,984 (8.61 MB)

Fig 1. Screen capture of model layers including MobileNetV2 transfer learning

With consideration in using transfer learning, MobileNetV2 resulted in being the best model for this dataset. Using ImageNet as pre-trained weights, it reported 90-95% accuracy on the test data each run. In fig 1, the layers are shown as followed; input layer (pre-defined by keras, 224x224 images), mobile net v2 layer (functional transfer learning), dropout layer of 1280 to avoid classification bias, and dense output layer of 23 (identifying where the image lands among the 23 classes) using softmax activation. The model is fine tuned with early-layers frozen, learning rate of 1e-5, early stopping and best model checkpoint. By training at 10 epochs, the model was able to avoid overfitting on loss and accuracy.

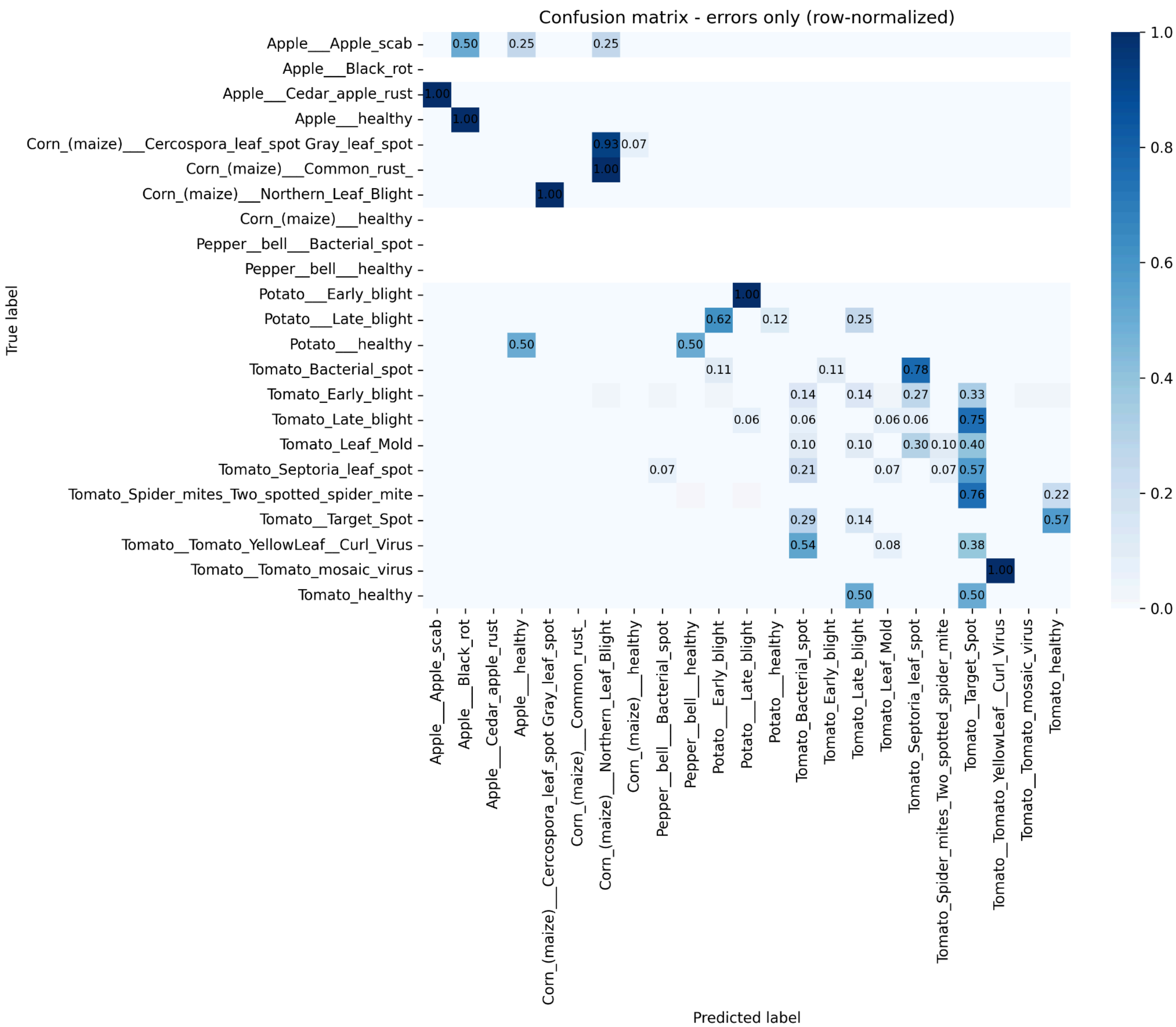

Fig 2. Confusion matrix based on errors only (row-normalized)

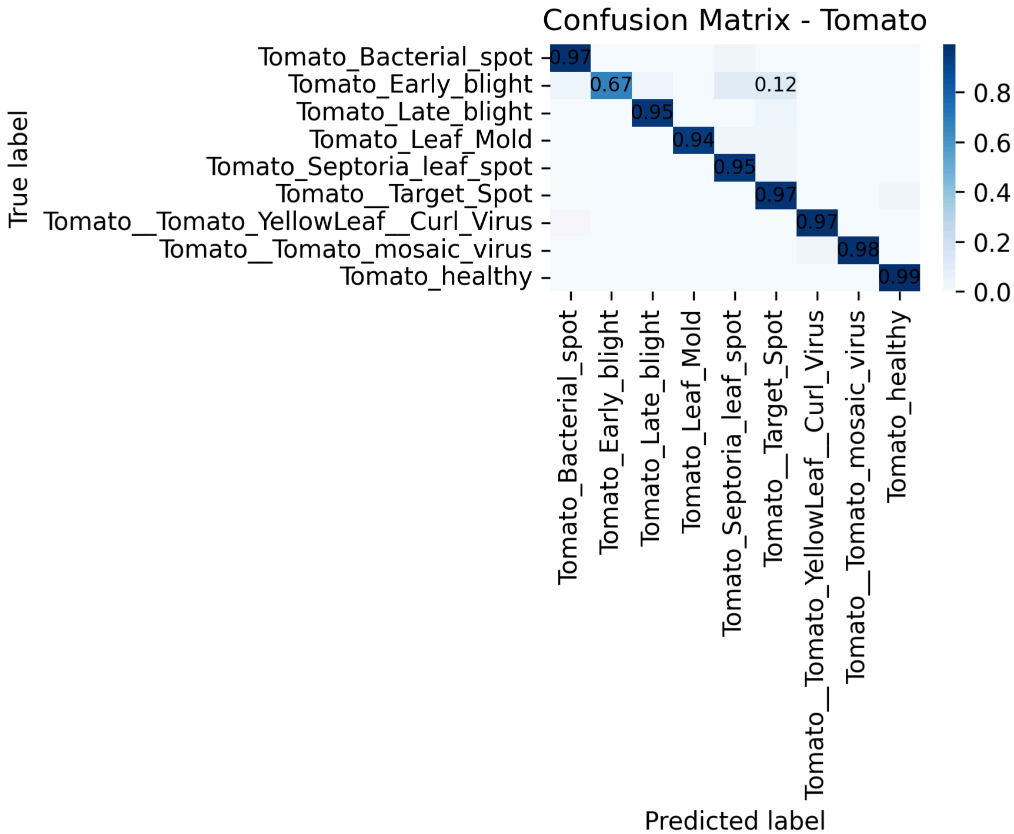

Fig 3. Closer look on confusion matrix based on images with tomato leaves

Evaluation on the test data are shown respectively by confusion matrices in fig 2 & 3. Fig 2 demonstrates errors made by the model in classifying the images, which pinpoints the model's miscalculation on plant leaves with four or more classes, such as tomatoes. Fig 3 is a closer look on the predicted and true labels of the tomato classes. According to the classification report done after evaluation, the 'Tomato___Early_blight' class had the lowest recall of roughly 0.45-0.65 each run with a f1 score of 0.59-0.7. Classification accuracy based on Tomato___Early_blight and Tomato___Spider_mites_Two_spotted_spider_mite scored the lowest by 60-68%. Therefore, fig 3 is a result of identifying where the model miscalculated the true tomato classes. The model created false positives on Tomato___Target_Spot when it's Tomato___Early_blight. Likely considerations to this miscalculation is image quality and collection (images from different classes looked too similar, not enough class balance, and possible data augmentation mix-up).

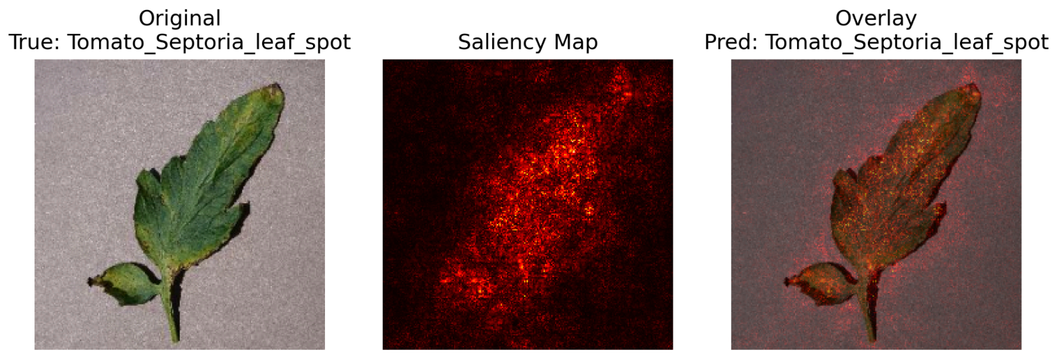

Fig 4. Saliency Map based on extracted sample from test dataset - Tomato Septoria Leaf Spots

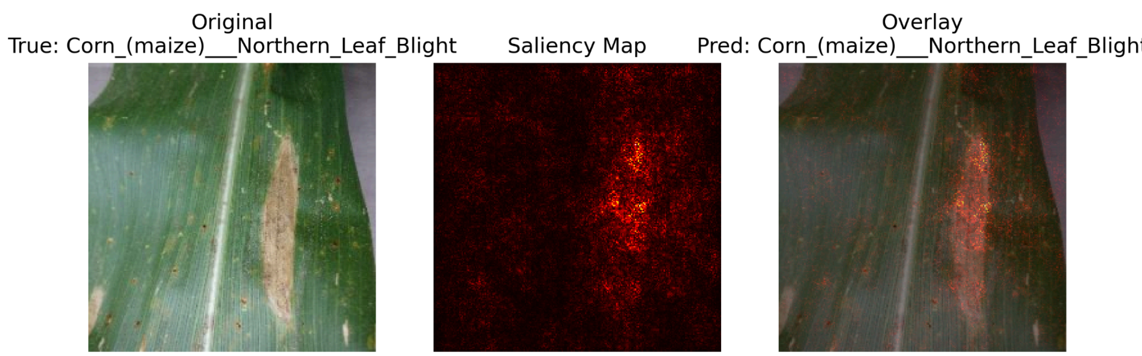

Fig 5. Saliency Map based on extracted sample from test dataset - Corn (Maize) Northern Leaf Blight

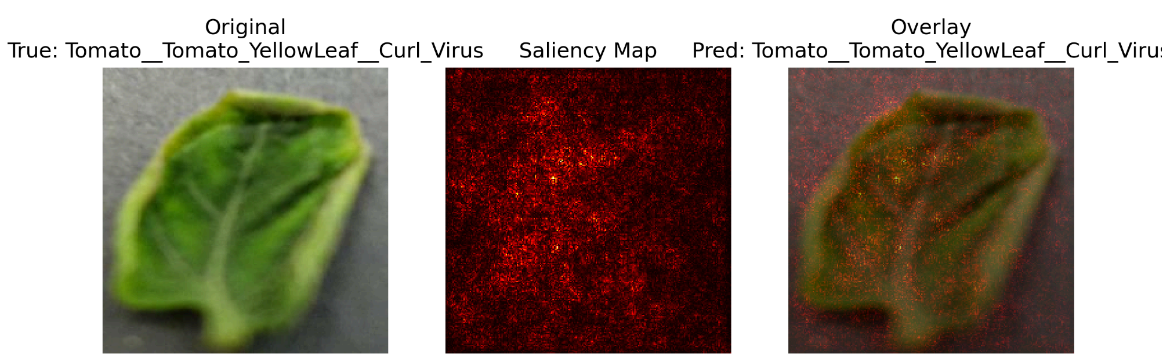

Fig 6. Saliency Map based on extracted sample from test dataset - Tomato YellowLeaf Curl Virus

By providing a heatmap to overlay the images, we can identify how MobileNetV2 transfer learning was looking at an image for classification. Li et al. explains that, “disease-infected plants usually show obvious marks or lesions on leaves, stems, flowers, or fruit” (Li et al., 2021, p. 56683). Basically, we need to see that the model is actively looking at lesions on the plant’s leaves in identifying the disease, not the background. Based on these assumptions, a series of saliency tests were done on the test dataset. Fig 4 & 5 show effective classification by the model as we look at the middle images, saliency map, and see how the model analyzed the lesions. Fig 5 is a zoomed-in image of a corn leaf with an obvious blight mark, seeing the map showed that the model clearly identifies abnormalities in the leaf. Fig 6 is an example of data augmentation that found its way in our test dataset. However, this is an example of the model’s prediction to be true even with slight data augmentation.

Transfer learning in MobileNetV2 effectively multi-classifies the health of a plant's leaves by having enough generalization to avoid overfitting. However, some key considerations of analyzing this model are with the dataset’s class balance and possible lack of collection in some classes. The dataset holds 35700 images, but some classes have 1000 or less images collected, which likely caused classification mix-up and bias (according to our confusion matrices). For future troubleshooting of this model, class weights need to be considered. In consideration to class weights, plant groups like tomato held nine classes total. These plant groups could have benefitted more in its own predictive model, as it seemed to complicate a model training on a more general dataset. Overall, recommendation for more data collection in classes with the fewest images, predictive modelling on plant leaves with multiple classes for better detection is key beneficial for farmers to readily predict early signs of declining health rates, and therefore, treat accordingly.

References

Lili Li et al. (2021). Plant Disease Detection and Classification by Deep Learning—A Review. *IEEE Access*. 56683-56698. https://ieeexplore.ieee.org/document/9399342/