

Tomato Disease Classifier

Evie Klaassen
Ka Yam



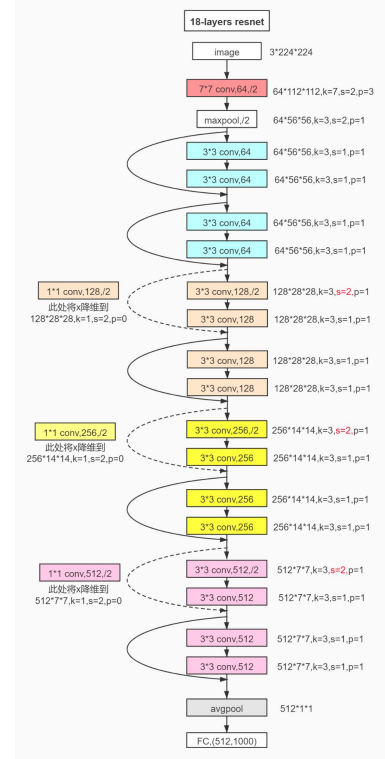
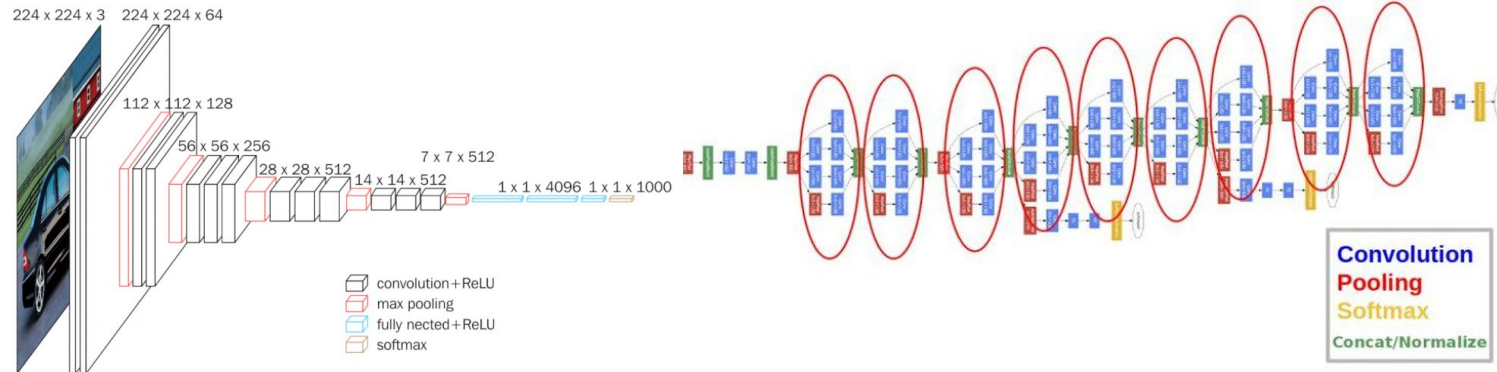
Data Sources:

- Kaggle – PlantVillage Dataset
 - Color, greyscale, and segmented images (~18k images in each category)
 - 10 labels (i.e. healthy, bacterial spot, early blight, mold, etc.)
 - Mendeley Data – Tomato Leaf Image Dataset
 - 306 images
 - 3 labels (healthy, mosaic virus, yellow leaf curl virus)
-



Preprocessing

- Resizing to fit model input needs
- Albumentations:
 - Spatial transformations
 - Horizontal flip, vertical flip, rotate
 - Pixel-level transformations
 - Random brightness contrast, blur, Gaussian noise
 - Normalization



Models

- Pretrained Models:
 - Fine-Tuned VGG16
 - GoogleNet
 - Partially Frozen ResNet18

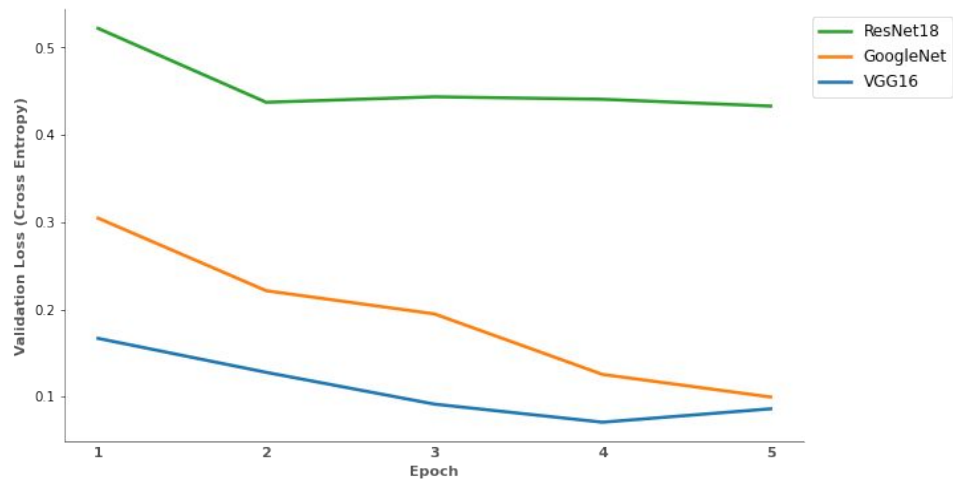


Results - Candidate Models*

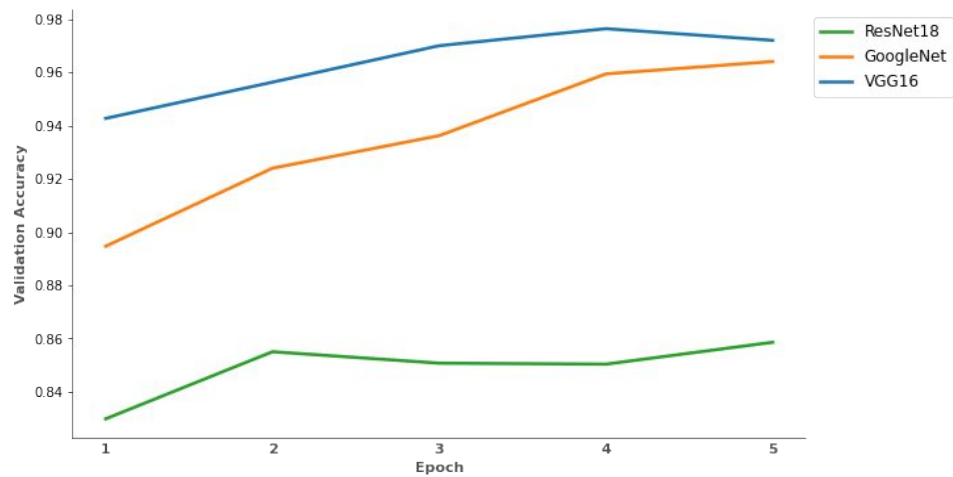
Model	Validation Loss	Validation Accuracy
Fine-tuned VGG16	0.086	0.97
GoogLeNet	0.0998	0.96
Partially Frozen ResNet18	0.433	0.86

*After training for 5 epochs

Validation Loss over 5 Epochs



Validation Accuracy over 5 Epochs





Final Model Performance

- Fine-Tuned VGG16
- Training dataset: 43,826 images (80%)
- Test dataset: 10,957 images (20%)
- Results:
 - **Test Loss: 0.057**
 - **Test Accuracy: 0.98**



Future Directions

- Figure out what classes are frequently confused with others due to high leaf pattern variation
- Integrate the model with built in cell phone cameras to provide real time tomato leaf disease detection and identification

Thank you!