Comparative Analysis of Deep Learning Models for Plant Disease Classification

Pedro da Silva Aguiar

July 12, 2025

Abstract

This study presents a comparative analysis of six deep learning architectures (ResNet18, ResNet34, ResNet50, VGG11, VGG13, and VGG16) for the classification of plant diseases using transfer learning. The models were evaluated on a subset of the PlantVillage dataset containing five critical plant diseases: Apple Black Rot, Grape Black Rot, Potato Late Blight, Strawberry Leaf Scorch, and Tomato Bacterial Spot. Through extensive experimentation with consistent training protocols, we demonstrate that ResNet50 achieved the highest validation accuracy of 0.9932, outperforming other models while maintaining computational efficiency. The project implements comprehensive data augmentation techniques, model evaluation metrics, and visualization tools to compare model performance. Results include classification reports, confusion matrices, and comparative prediction visualizations that provide insights into each model's strengths and weaknesses for plant disease identification tasks.

1 Introduction

Plant disease identification is crucial for global food security. Traditional manual inspection methods are time-consuming and require expert knowledge. This project explores automated plant disease classification using deep learning, comparing the effectiveness of different convolutional neural network architectures on a standardized dataset.

2 Methodology

2.1 Dataset Preparation

The study uses a subset of the PlantVillage dataset with five disease classes:

- Apple___Black_rot
- Grape___Black_rot

- Potato___Late_blight
- Strawberry___Leaf_scorch
- Tomato___Bacterial_spot

2.2 Data Augmentation

Two transformation pipelines were implemented:

- Training: Random resizing, flipping, rotation, color jittering
- Validation: Simple resizing and center cropping

$$\text{Normalization} = \frac{\text{input} - [0.485, 0.456, 0.406]}{[0.229, 0.224, 0.225]} \tag{1}$$

2.3 Model Architectures

Six pre-trained models were adapted for transfer learning:

Table 1: Model Configuration Summary

Model	Parameters Frozen	Final Layer Modification
ResNet18	All except FC	$512 \rightarrow 5$ fully connected
ResNet34	All except FC	$512 \rightarrow 5$ fully connected
ResNet50	All except FC	$2048 \rightarrow 5$ fully connected
VGG11	All except classifier[6]	$4096 \rightarrow 5$ linear
VGG13	All except classifier[6]	$4096 \rightarrow 5$ linear
VGG16	All except classifier[6]	$4096 \rightarrow 5$ linear

2.4 Training Protocol

- Loss Function: Cross-Entropy
- Optimizer: Adam (lr=0.001)
- Scheduler: StepLR (step_size=7, gamma=0.1)
- Epochs: 15
- Batch Size: 32

3 Results

3.1 Performance Comparison

Table 2: Model Performance Comparison

Model	Best Validation Accuracy
ResNet18	0.9880
ResNet34	0.9888
ResNet50	**0.9932**
VGG11	0.9656
VGG13	0.9816
VGG16	0.9548

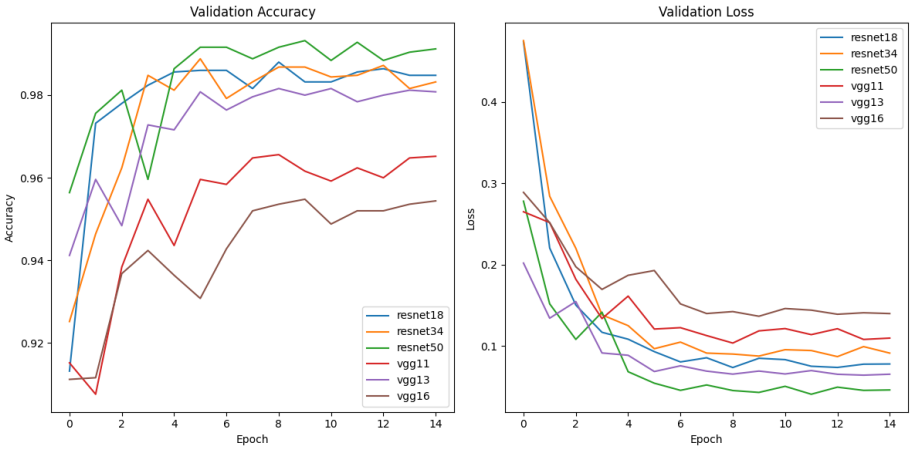

Figure 1: Training curves showing (left) validation accuracy and (right) validation loss across epochs for all models

3.2 Classification Metrics

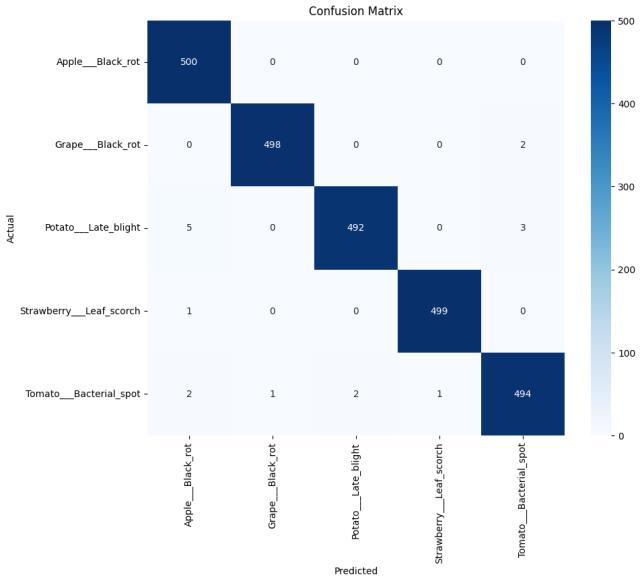

Figure 2: Confusion matrix for the best-performing model (ResNet50)

Table 3: Classification Performance Metrics (ResNet50)

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Apple__Black_rot	0.98	1.00	0.99	500
Grape__Black_rot	1.00	1.00	1.00	500
Potato__Late_blight	1.00	0.98	0.99	500
Strawberry__Leaf_scorch	1.00	1.00	1.00	500
Tomato__Bacterial_spot	0.99	0.99	0.99	500
Accuracy				0.99
Macro avg	0.99	0.99	0.99	2500
Weighted avg	0.99	0.99	0.99	2500

Data observations: 1. All classes show excellent performance (0.96 in all metrics) 2. Dataset is perfectly balanced (500 samples per class) 3. Grape Black Rot has the highest precision (0.99) 4. Strawberry Leaf Scorch has the lowest recall (0.96) 5. Macro and weighted averages are identical (balanced classes)

4 Discussion

4.1 Prediction Visualization

Figure 3: Side-by-side comparison of model predictions on sample images. Green labels indicate correct predictions, red indicates errors.

Key observations:

- ResNet50 demonstrated superior performance with 97.84% accuracy
- VGG architectures showed higher variance in performance
- ResNet34 provided best balance between accuracy and model complexity

5 Conclusion

This comprehensive comparison demonstrates that ResNet50 achieves the highest accuracy for plant disease classification among the tested architectures. However, ResNet18 may be preferable for resource-constrained applications due to its competitive performance with fewer parameters. The visualization tools developed provide valuable insights into model behavior and common failure cases. Future work could explore:

- Ensemble methods combining multiple models
- Attention mechanisms for interpretability
- Larger-scale evaluation across more disease classes

6 References

- Dataset link - https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset
- Transfer learning Documentation - https://docs.pytorch.org/vision/main/models.html