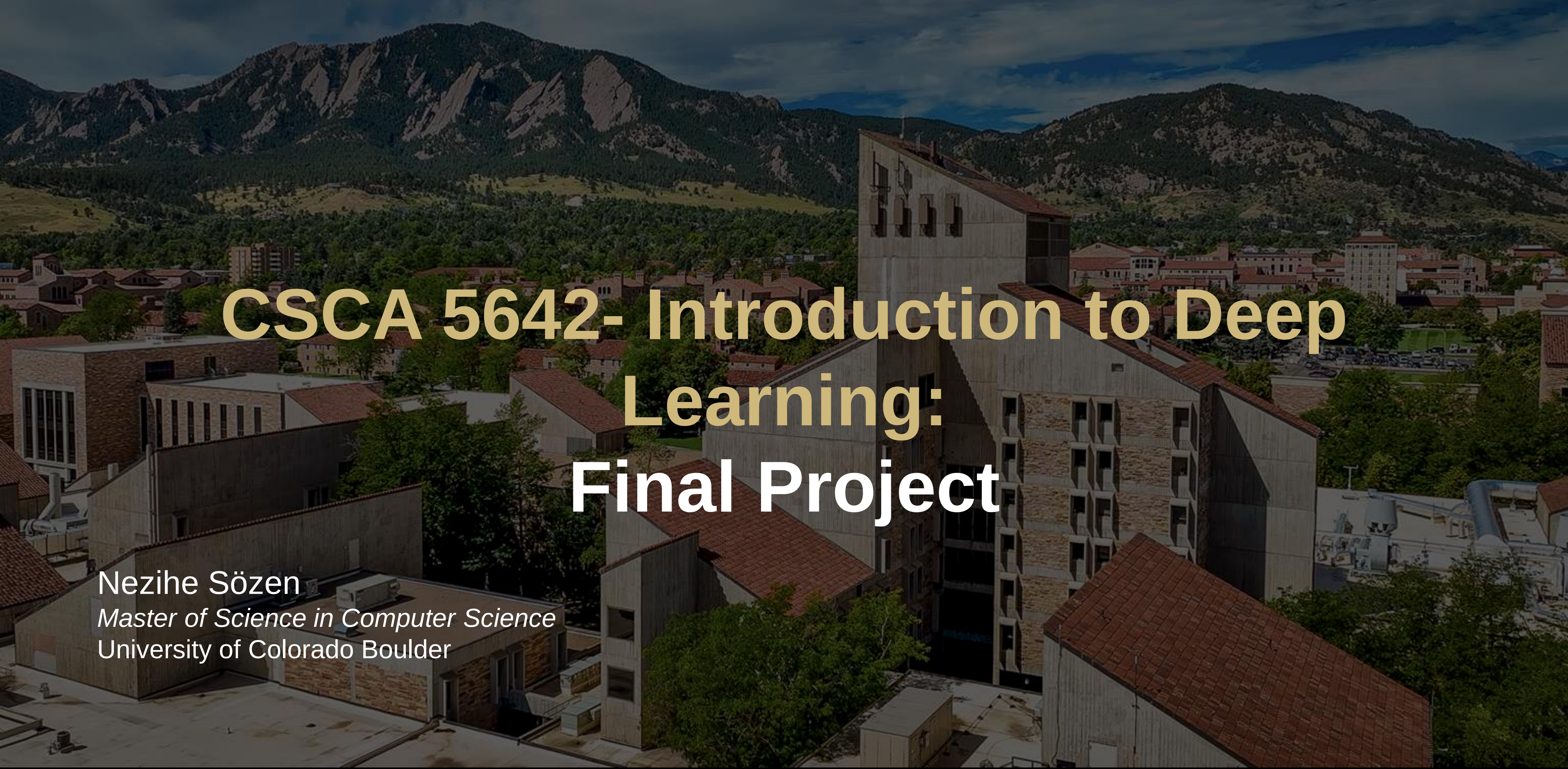


CSCA 5642- Introduction to Deep Learning:

Final Project

Nezihe Sözen
Master of Science in Computer Science
University of Colorado Boulder

Outline

- Introduction – Problem Statement & Importance
- Exploratory Data Analysis (EDA)
- Dataset Analysis and Key Insights
- GAN Process and Model Training
- Transfer Learning with EfficientNetB0
- Visualization & Interpretation of Results
- Conclusion & Future Work

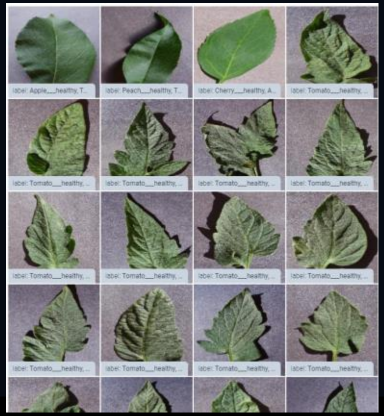
Introduction

Project Motivation & Objective

- Plant Disease Detection Using Deep Learning
- Accurate Disease Detection: Train a model to classify plant diseases effectively.
- Data Augmentation with GANs: Generate synthetic images for underrepresented disease classes.
- Transfer Learning with EfficientNetB0: Fine-tune a pre-trained model for better classification accuracy.

Challenges

- Limited data for rare diseases: Not enough real-world samples for all categories.
- GAN-generated image quality: Ensuring the synthetic images are realistic and useful.
- Transfer learning adaptation: EfficientNetB0 must generalize well for plant disease classification.



https://github.com/spMohanty/PlantVillage-Dataset

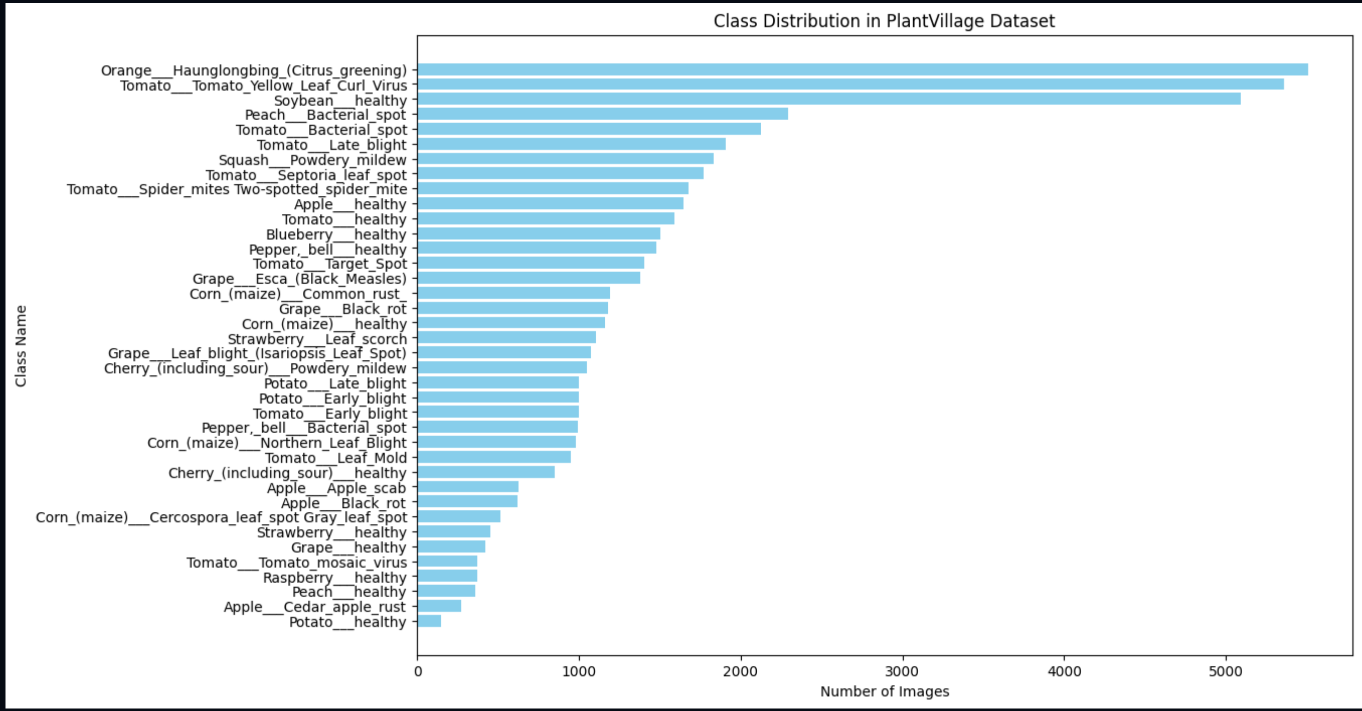
Exploratory Data Analysis (EDA)

- The dataset consists of multiple plant species and their diseases.

- Class imbalance is present, with some diseases having significantly fewer images.

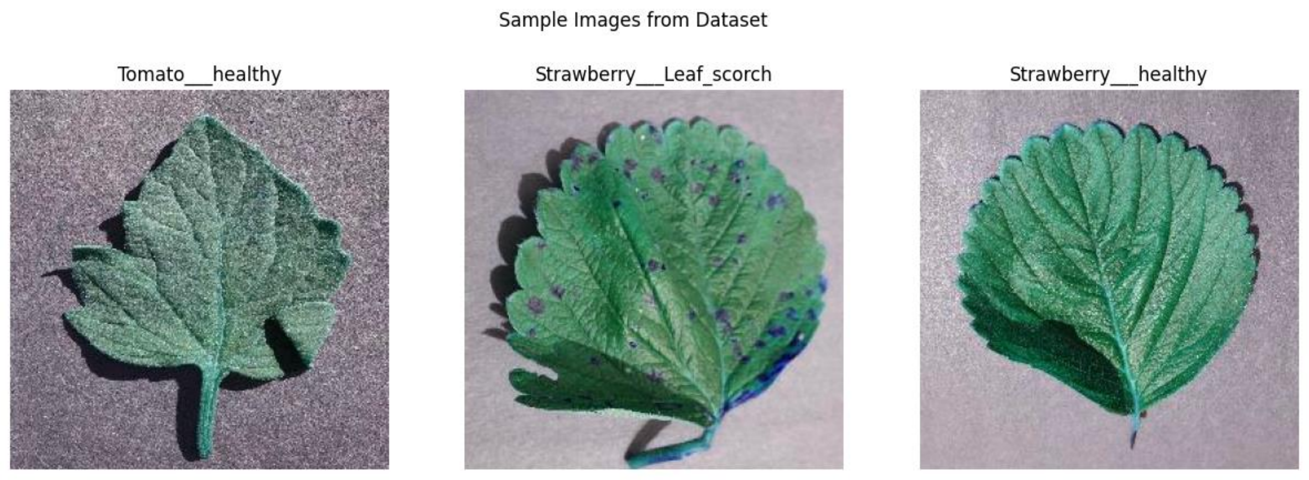
- Some underrepresented diseases have less than 500 images, requiring augmentation.

- GAN-based data augmentation will be applied to balance rare disease classes.



Key Insights

- The dataset contains 38 classes with significant class imbalance.
- Images are 256x256 pixels, standardized to .JPG, with no missing or corrupt files.
- Green is the dominant color, while diseased leaves show variations in red and blue channels.
- GAN-based augmentation will focus on underrepresented classes like Potato - Healthy and Apple - Cedar Apple Rust.
- Balancing the dataset and maintaining realistic color distributions are essential for model performance.

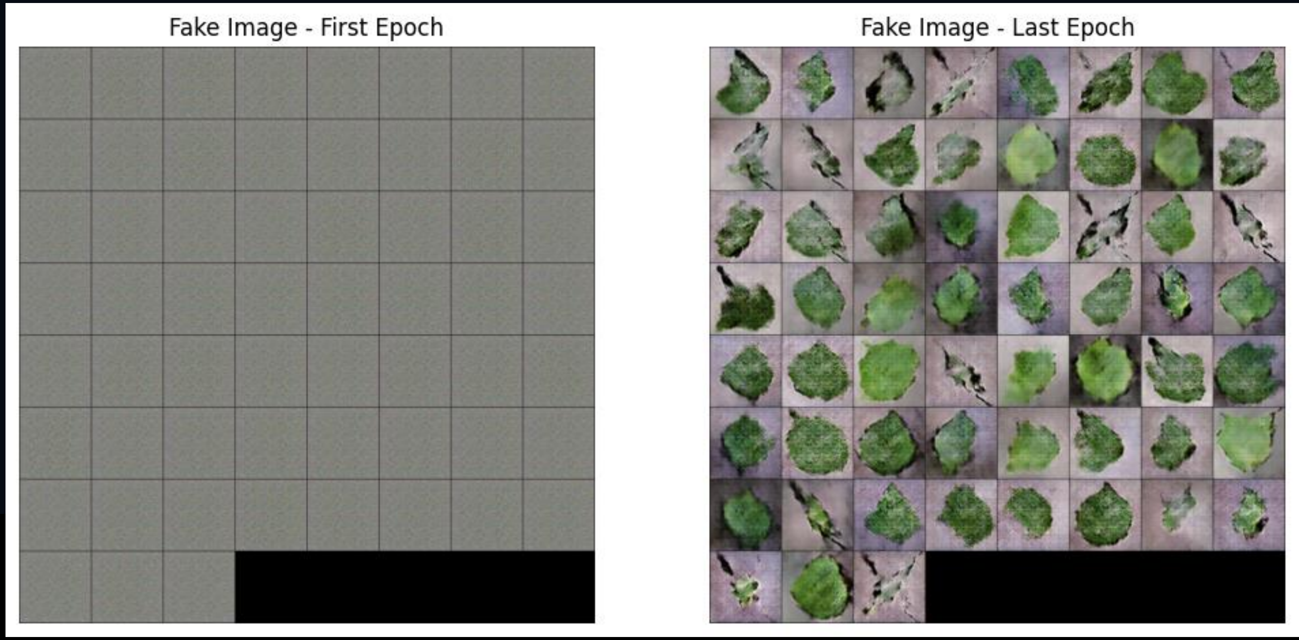


GAN Process

- The quality of the images produced was not high enough.
- More powerful hardware or alternative data augmentation methods should be evaluated to obtain better results.

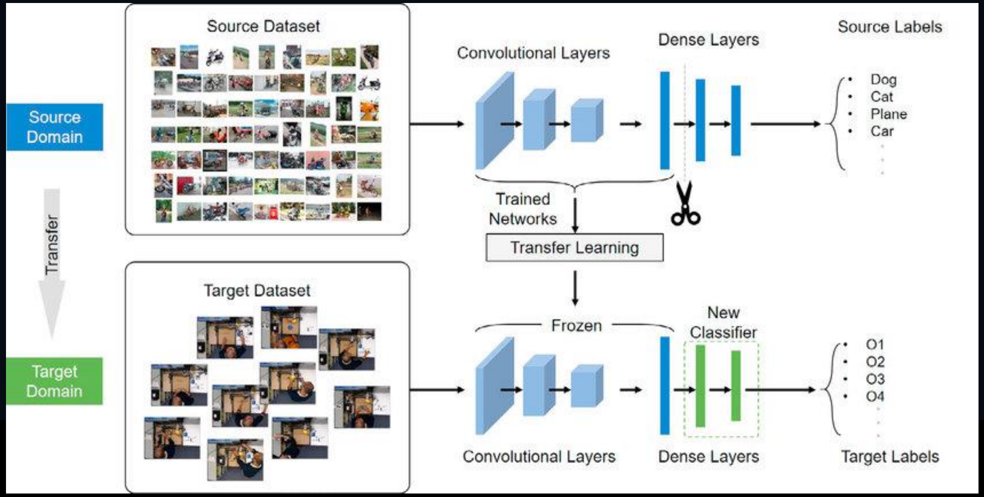
Architecture:

- Generator (G): Creates fake images from random noise.
- Discriminator (D): Distinguishes real images from fake ones.
- Deep Convolutional GAN (DCGAN)



Transfer Learning as a Deep Learning Approach

- Pretrained Model: Used EfficientNetB0 pretrained on ImageNet, with only the final layers trained for plant disease classification.
- Data Preprocessing: Applied ImageDataGenerator with EfficientNet's preprocessing function and an 80-20 train-validation split.
- Model Architecture: Added GlobalAveragePooling2D, Dense (128 units, ReLU), Dropout (0.3), and Softmax output for 38 classes.
- Training Details: Used Adam optimizer (lr=0.0001) and categorical cross-entropy loss, training for 20 epochs with batch size 32.



https://www.researchgate.net/figure/The-architecture-of-our-transfer-learning-model_fig4_342400905

	precision	recall	f1-score	support
Apple___Apple_scab	0.98	0.94	0.96	126
Apple___Black_rot	0.99	1.00	1.00	124
Apple___Cedar_apple_rust	1.00	1.00	1.00	55
Apple___healthy	0.96	0.99	0.98	329
Blueberry___healthy	1.00	1.00	1.00	300
Cherry_(including_sour)___Powdery_mildew	1.00	1.00	1.00	210
Cherry_(including_sour)___healthy	1.00	0.99	0.99	170
Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot	0.90	0.81	0.86	102
Corn_(maize)___Common_rust_	1.00	1.00	1.00	238
Corn_(maize)___Northern_Leaf_Blight	0.90	0.95	0.93	197
Corn_(maize)___healthy	1.00	1.00	1.00	232
Grape___Black_rot	0.98	0.97	0.97	236
Grape___Esca_(Black_Measles)	0.97	0.99	0.98	276
Grape___Leaf_blight_(Isariopsis_Leaf_Spot)	1.00	1.00	1.00	215
Grape___healthy	1.00	0.99	0.99	84
Orange___Haunglongbing_(Citrus_greening)	1.00	1.00	1.00	1101
Peach___Bacterial_spot	0.99	0.99	0.99	459
Peach___healthy	0.97	1.00	0.99	72
Pepper,_bell___Bacterial_spot	0.99	0.97	0.98	199
Pepper,_bell___healthy	0.98	0.99	0.98	295
Potato___Early_blight	1.00	0.99	1.00	200
Potato___Late_blight	0.99	0.97	0.98	200
Potato___healthy	0.96	0.87	0.91	30
Raspberry___healthy	1.00	1.00	1.00	74
Soybean___healthy	0.99	1.00	1.00	1018
Squash___Powdery_mildew	1.00	1.00	1.00	367
Strawberry___Leaf_scorch	1.00	1.00	1.00	221
Strawberry___healthy	1.00	0.99	0.99	91
Tomato___Bacterial_spot	0.96	0.96	0.96	425
Tomato___Early_blight	0.91	0.78	0.84	200
Tomato___Late_blight	0.93	0.95	0.94	381
Tomato___Leaf_Mold	0.94	0.95	0.94	190
Tomato___Septoria_leaf_spot	0.94	0.96	0.95	354
Tomato___Spider_mites Two-spotted_spider_mite	0.88	0.97	0.92	335
Tomato___Target_Spot	0.92	0.83	0.87	280
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.99	0.99	0.99	1071
Tomato___Tomato_mosaic_virus	1.00	0.96	0.98	74
Tomato___healthy	0.99	0.98	0.99	318
accuracy			0.98	10849
macro avg	0.97	0.97	0.97	10849
weighted avg	0.98	0.98	0.98	10849

Visualization & Interpretation of Results

- **High Overall Accuracy: The model achieved 98% accuracy, demonstrating strong performance.**
- **Balanced Performance: Macro F1-score = 97%, indicating that even less frequent classes were classified well.**
- **Strong Precision & Recall: Most classes have precision and recall values above 0.95, ensuring both high detection rates and low false positives.**
- **Challenging Classes: Some classes, like Tomato___Spider_mites Two-spotted_spider_mite (F1-score: 0.92) and Tomato___Target_Spot (F1-score: 0.87), had relatively lower performance.**
- **Well-Handled Class Imbalance: Even underrepresented classes like Potato___healthy (F1-score: 0.97) were classified with high accuracy.**

Conclusion & Future Work

Conclusion

- Transfer learning with EfficientNetB0 achieved 98% accuracy for plant disease classification.
- The model showed minimal overfitting, with validation accuracy closely following training accuracy.

Future Works

- Improve class balance using weighted loss functions or oversampling.
- Explore alternative augmentation methods, such as Mixup or CutMix, instead of GANs.
- Test real-world performance on unseen plant images outside the dataset.