PROJECT REPORT: PLANT DISEASE DETECTION FROM IMAGES

1. INTRODUCTION:

Plant diseases significantly impact crop yield and food security. Quick, automated detection can help farmers and gardeners take preventive action earlier. This project aims to build an end-to-end **deep learning system** to detect plant diseases from leaf images, and deploy it as a **Streamlit web application** for easy use.

2. DATASET:

- **Source:** New Plant Diseases Dataset (Augmented)
- **Content:** Images of healthy and diseased leaves across **38 classes**.
- **Augmentation:** Dataset already contains augmented samples; additional augmentation applied during training.

3. SYSTEM DESIGN & IMPLEMENTATION:

3.1 Preprocessing & Augmentation

- Images resized to **128×128×3**.
- Normalized pixel values (rescale=1. /255).
- On-the-fly augmentation:
    - Rotation (up to 25°)
    - Width/height shift (20%)
    - Shear (0.2)
    - Zoom (0.2)
    - Horizontal flip

Goal: Improve model generalization and reduce overfitting.

3.2 Model Architectures

Implemented and compared **four models**:

Model	Type	Input Size	Trainable Params (approx.)
CustomCNN	Custom-built	128×128×3	~5M
MobileNetV2	Transfer learning	128×128×3	Few M
VGG16	Transfer learning	128×128×3	~15M
ResNet50	Transfer learning	128×128×3	~23M

All transfer learning models used pretrained imagenet weights with frozen base layers, adding:

- GlobalAveragePooling2D
- Dense (256, relu)
- Dropout
- Dense (38, softmax)

3.3 Training Details

- Optimizer: **Adam** (lr=0.001)
- Loss: **Categorical cross entropy**
- Batch size: **32**
- Epochs: **20**
- Callbacks:
    - Early Stopping (patience=4, restore_best_weights=True)
    - ReduceLROnPlateau (factor=0.5, patience=2, min_lr=1e-6)

4. MODEL EVALUATION & RESULTS:

Model	Accuracy	Precision	Recall	F1-Score
CustomCNN	94.58%	95.37%	94.51%	94.57%
MobileNetV2	92.08%	92.50%	92.07%	92.11%
VGG16	87.46%	88.20%	87.53%	87.46%
ResNet50	21.88%	22.64%	21.67%	18.18%

4.1 Analysis

- **CustomCNN** performed best (F1 ≈ 94.57%), likely because it was lightweight, directly trained on the dataset, and tuned for this specific task.
- **MobileNetV2** also performed very well with ≈ 92% accuracy, showing that lightweight transfer learning is effective.
- **VGG16** had acceptable performance but slightly lower than MobileNetV2.
- **ResNet50** performed poorly (≈ 22% accuracy).
  - **Reason:** Due to system hardware limitations and functional constraints, the ResNet50 base layers were not fully trainable; model was underfitting the dataset.

5. OUTPUTS:

Generated and saved:

- Trained models (.h5): models/CustomCNN.h5, etc.
- Accuracy plots (.png): plots/CustomCNN_acc.png, etc.
- Metrics reports (.json): metrics_reports/CustomCNN_metrics.json, etc.

6. DEPLOYMENT:

Deployed as a **Stream lit web application**:

- **URL:** Streamlit App
- Features:
  - Upload leaf image (.jpg, .jpeg, .png)
  - Visual display of uploaded image
  - Get disease prediction + confidence

7. USER TESTING & FEEDBACK:

How to Interpret Predictions:

- The model predicts **one disease class** with a confidence score.
- Higher confidence (>80%) → prediction is usually more reliable.
- For borderline cases (50–70%), check the leaf visually and consult an expert if unsure.

Important: This tool is intended for educational & early-detection purposes. Always confirm with an expert or agronomist for critical decisions.

Troubleshooting & Known Issues

Issue	Explanation / Fix
Prediction seems wrong	Try uploading a clearer image (single leaf, good lighting).
App takes longer to load	Happens if server is idle for long; wait & retry.
Unusual predictions	Dataset may not cover all rare diseases or local plant varieties.
ResNet50 low accuracy	Model underperforms due to limited fine-tuning and hardware constraints.

- Tested by peers and students with various leaf images.
- Users appreciated:
  - Simplicity of interface
  - Quick prediction (<1 sec on CPU)
- **Suggested improvements:**
  - Add example images / dropdown for demo
  - Show top-3 predictions instead of one
  - Add explanation of disease names

Limitations:

- The system can only classify diseases present in the trained dataset (**38 classes**).
- Rare diseases or different crops may not be supported.
- Confidence depends on photo quality and dataset balance.
- Some deep models (e.g., ResNet50) are intentionally limited in this version due to system performance.

8. CONCLUSION:

- Successfully built and deployed a **real-time plant disease detection system**.
- Custom CNN outperformed popular pretrained models on this dataset.
- The system shows high accuracy and can help farmers/gardeners detect diseases early.