Plant Disease Detection from Leaf Images

---

1. Introduction

Plant diseases are one of the leading causes of reduced crop productivity and quality across the globe. In particular, apple leaf diseases such as scab, black rot, and cedar rust impact both small-scale and commercial farming. Traditionally, identification of these diseases relies on visual inspection by agricultural experts, which is not scalable for large farms and is prone to human error. With the advancement of artificial intelligence and computer vision, it is now possible to automate the detection process using image-based classification models. This project leverages deep learning techniques, specifically Convolutional Neural Networks (CNNs), to detect and classify apple leaf diseases from images.

---

2. Abstract

The primary objective of this project is to develop a deep learning model capable of identifying four apple leaf conditions: Apple Scab, Apple Black Rot, Apple Cedar Rust, and Healthy leaves. Using the PlantVillage dataset, a lightweight CNN model was trained on a filtered subset of the data. The dataset was preprocessed through resizing, normalization, and augmentation to improve generalization. After training, the model achieved a high level of accuracy while maintaining a small size (347 KB). A graphical interface was built using Streamlit to allow users to upload leaf images and receive real-time predictions. The model's compactness ensures efficient deployment, making it suitable for low-resource devices or integration into mobile or web-based diagnostic tools.

---

3. Tools Used

- **Programming Language:** Python
- **Libraries & Frameworks:** TensorFlow, Keras, NumPy, OpenCV, Pillow, Streamlit
- **Dataset Source:** PlantVillage (filtered for 4 apple-related classes)
- **Development Environment:** Visual Studio Code
- **Model Format:** HDF5 (.h5)
- **Deployment Interface:** Streamlit (Web App)

---

4. Steps Involved in Building the Project

1. **Dataset Acquisition and Cleaning**
    - Downloaded the PlantVillage dataset and extracted only four classes related to apple leaves.
    - Verified class balance and removed any corrupted or irrelevant images.
2. **Image Preprocessing**

- Standardized all input images to 128x128 pixels.
- Applied normalization (scaling pixel values between 0 and 1).
- Used ImageDataGenerator to apply data augmentation such as rotation, zoom, and flip.

3. **Model Design and Training**
    - Implemented a compact CNN with three convolutional layers, each followed by max pooling.
    - Used 'relu' activation in hidden layers and 'softmax' for output classification.
    - Compiled model with categorical crossentropy loss and Adam optimizer.
    - Trained the model over 10 epochs with batch size 32.
4. **Model Evaluation and Optimization**
    - Evaluated performance on a 20% validation split.
    - Achieved strong accuracy with low overfitting due to data augmentation.
    - Exported model to plant_model_small.h5 for deployment.
5. **GUI Development using Streamlit**
    - Created a simple GUI allowing users to upload a leaf image.
    - The uploaded image is preprocessed in real-time and passed to the model.
    - The app displays the predicted class and confidence percentage.
6. **Deployment and Testing**
    - Verified model behavior on unseen leaf images.
    - Ensured model and app file sizes are within GitHub limits for sharing.
    - Final system tested for usability, speed, and prediction accuracy.

---

5. Conclusion

This project demonstrates a practical and lightweight solution for plant disease detection using deep learning. The final model is compact, accurate, and easy to deploy, making it a viable tool for agricultural assistance. Through image preprocessing, efficient CNN design, and intuitive Streamlit integration, the system enables farmers or researchers to detect apple leaf diseases with minimal effort. The Streamlit app enhances accessibility, while the small model size ensures smooth performance even in resource-constrained environments. In future iterations, the project can be extended to support more plant species and deployed as a mobile app for real-time field diagnostics.