Plant Disease Detection System for Sustainable Agriculture

Madhav K
AICTE Student Internship ID: STU66a1f414251071721889812

1. Problem Statement

In modern agriculture, early detection of plant diseases is critical to ensure crop health and maximize yield. Manual inspection is time-consuming and prone to human error. To address this, we propose a CNN (Convolutional Neural Network)-based model capable of automatically detecting and classifying diseases from leaf images of crops such as apple, cherry, grape, and corn.

Key Challenges

- Accurate classification of healthy vs. diseased leaves.
- Identification of specific disease types (e.g., apple scab, grape black rot).
- Handling variations in leaf appearance due to lighting, orientation, and background noise.

Solution Approach

- A deep learning-based system that processes leaf images and predicts disease conditions.
- Integration with precision agriculture tools for early intervention and sustainable farming.

2. Project Pipeline

The workflow follows a structured data-driven pipeline to ensure model accuracy and efficiency:

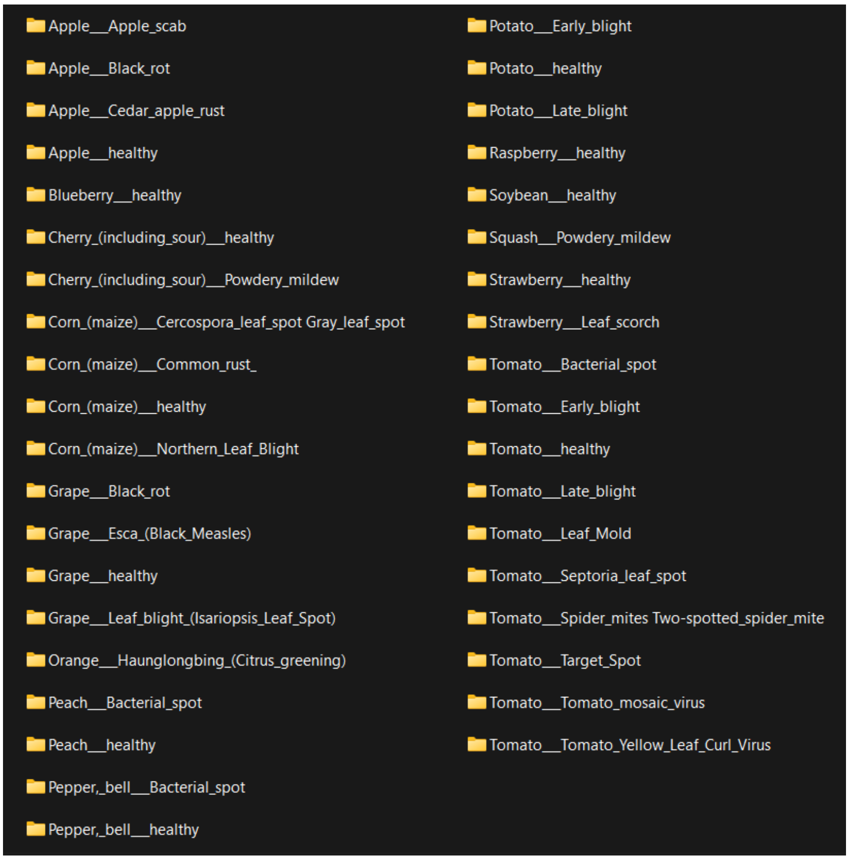
(1) Data Collection & Loading

- Source datasets from publicly available repositories such as PlantVillage and Kaggle.
- Dataset includes labeled images of healthy and diseased leaves across multiple crop types.

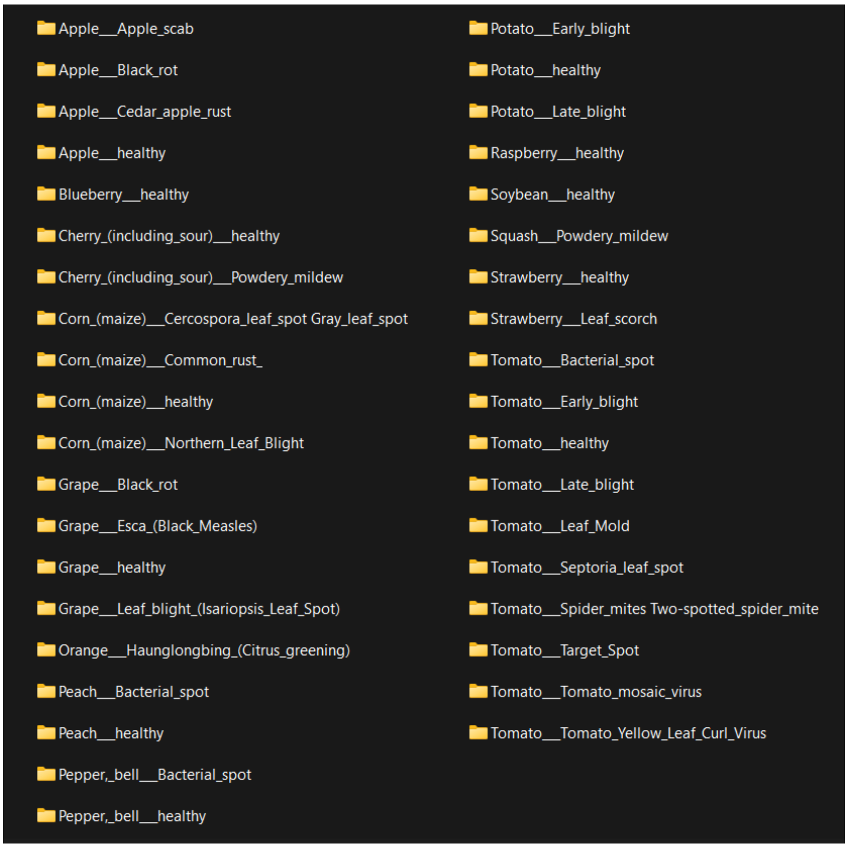
(2) Dataset Directory Structure

The dataset is organized into three subsets for training, validation, and testing:

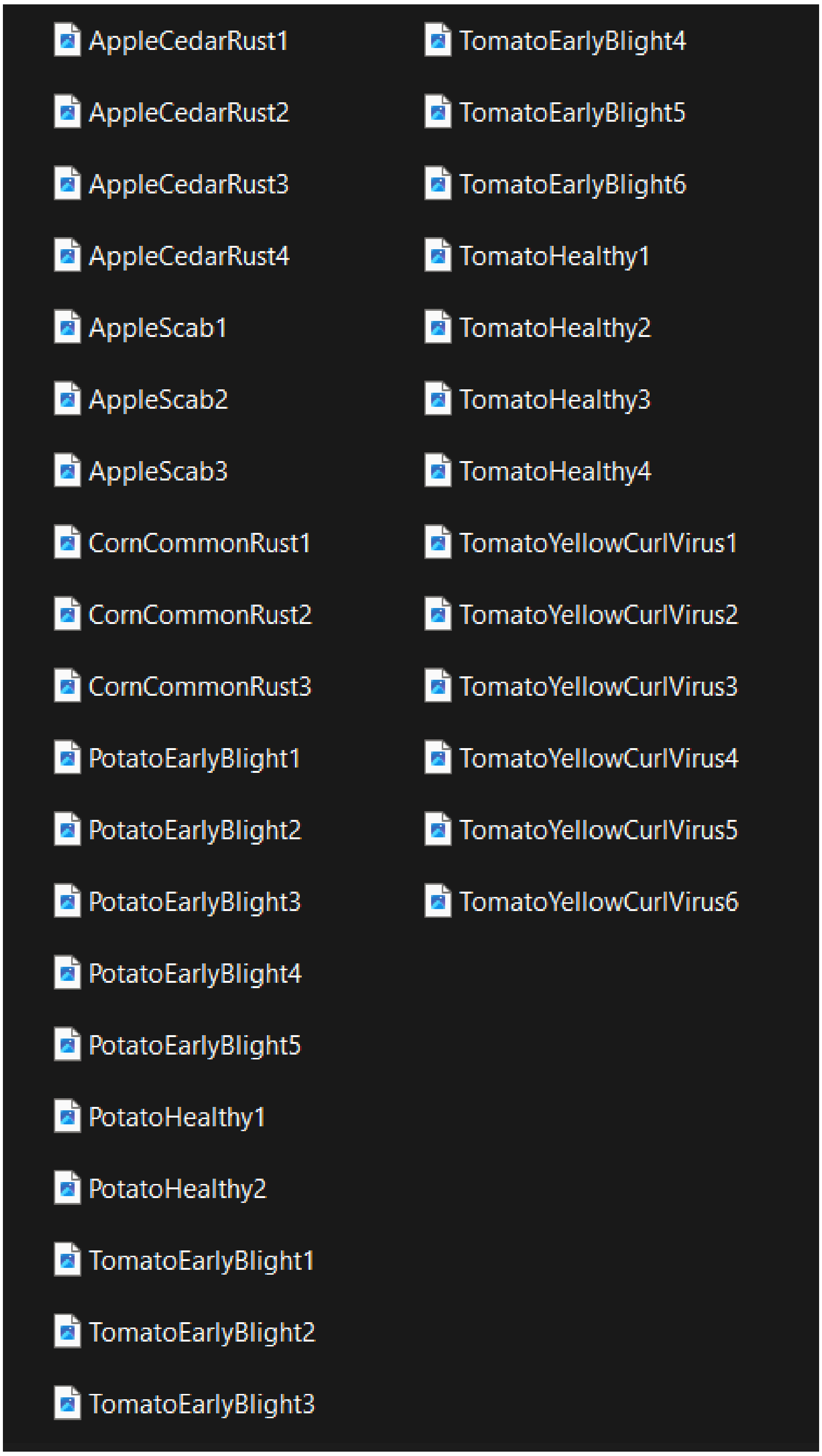
- **Train Directory**



- **Validation Directory**



- **Testing Directory**



(3) Data Upload & Extraction

- Compress the dataset into a .zip file.
- Upload the file to Google Drive and mount it in Google Colab.
- Use Python's `zipfile` library to extract and load images.

(4) Image Preprocessing & Augmentation

- Resize images to a uniform dimension (e.g., 128x128 pixels).
- Apply augmentation techniques such as rotation, flipping, and brightness adjustment to improve model robustness.

(5) CNN Model Architecture

- **Input Layer**: Accepts 128x128x3 RGB images.
- **Convolutional Layers**: Extract features using ReLU activation functions.
- **Pooling Layers**: Perform dimensionality reduction using MaxPooling.
- **Fully Connected Layers**: Classify features using Softmax activation.
- **Output Layer**: Predicts disease class (e.g., "healthy," "apple_scab").

(6) Model Training & Validation

- Train the model on the training set and validate it using the validation set.
- Optimize hyperparameters such as learning rate and batch size.

(7) Testing & Evaluation

- Evaluate the model's performance on an unseen test dataset.
- Use metrics such as Accuracy, Precision, Recall, and F1-Score.

3. Expected Outcomes

- A high-accuracy CNN model for classifying plant diseases.
- Potential integration with mobile applications or IoT-based farming systems.
- Contribution to sustainable agriculture by reducing crop loss and improving yield quality.