Crop Disease Detection & Risk Prediction System

Submitted by
Aaliyah Sheikh

Abstract

Agriculture plays a crucial role in food security and the global economy. However, plant diseases significantly reduce crop yield and quality. This project presents a system that detects crop diseases using Convolutional Neural Networks (CNN) and predicts the risk of disease occurrence based on weather conditions using a machine learning model. The system integrates image-based disease classification with weather-based risk assessment to assist farmers in taking early preventive actions.

Contents

Chapter 1

Introduction

Plant diseases are a major barrier to sustainable agricultural productivity. Traditional methods of disease detection require expert knowledge and are often time-consuming. With advancements in machine learning, automated disease detection has become possible using image analysis. Additionally, weather conditions such as temperature, humidity, and rainfall significantly influence the likelihood of disease outbreaks.

This project focuses on developing a two-stage intelligent system:

1. **Disease Classification Model (CNN)** — Detects the type of disease from a plant leaf image.

2. **Weather-Based Risk Prediction Model (Random Forest)** — Predicts whether the disease risk is "High" or "Low" using climatic parameters.

Chapter 2

System Architecture

The system follows the pipeline below:

1. Collect plant leaf images from dataset.

2. Preprocess and resize images.

3. Train CNN model for disease classification.

4. Create or collect weather dataset for disease conditions.

5. Train machine learning model (Random Forest) for risk prediction.

6. Integrate both models and provide final disease + risk output.

Chapter 3

Dataset Description

Image Dataset: Plant leaf images obtained from Kaggle. The dataset contains healthy and diseased leaves categorized by labels.

Weather Dataset: A structured dataset containing temperature, humidity, rainfall, pressure, and wind speed. The label indicates risk as high or low. A Random Forest classifier was trained on this dataset.

Chapter 4

Model Development

4.1 CNN Disease Classification

A Transfer Learning approach using MobileNetV2 was used due to its efficiency. Images were resized to 128×128 and normalized before training. The final output layer used Softmax activation for class probability prediction.

4.2 Weather Risk Prediction Model

A Random Forest model was trained on weather parameters. The model outputs whether current weather conditions indicate a high or low risk of disease spread.

Chapter 5

Implementation Tools

- Python

- TensorFlow / Keras

- Scikit-learn

- OpenCV

- Google Colab

Chapter 6

Results

The CNN model successfully classified leaf disease with high accuracy. The Random Forest model provided reliable risk prediction based on real or synthetic weather data.

Chapter 7

My Contribution (Integration & Final Inference System)

My primary role in the project was to **integrate both trained models** and build a final prediction system that produces:

- Disease Name from Leaf Image (using CNN)

- Disease Risk Level from Weather Data (using Random Forest)

I developed the final Python script / Jupyter Notebook which:

1. Loaded the trained CNN model and weather model.

2. Preprocessed user input leaf images for classification.

3. Accepted real-time or manual weather inputs.

4. Generated combined output showing:

   - Detected Disease Name
   - Probability / Confidence Level
   - Risk Level (High/Low)

This integration allows users to input a leaf image and weather conditions, and receive a complete diagnosis and prediction in one step.

Chapter 8

Conclusion

This project demonstrates that machine learning can greatly assist in early detection and prevention of crop diseases. By combining visual analysis and environmental risk prediction, the system enables proactive disease management, minimizing crop loss and improving agricultural efficiency. Future improvements may include deployment as a mobile app for real-time usage in farms.