

Department of Computer Engineering

COMSATS University Islamabad – Lahore Campus

# PLANT DISEASE DETECTION USING CNN

CEP

Report By

|  |  |
| --- | --- |
| NAME | RegistrationNumber |
| Ayesha Butt | CUI/FA22-BCE-053/LHR |
| Maha Chaudhary | CUI/FA22-BCE-082/LHR |

For the course

# Semester6 Spring 2025

Supervised by:

ABU-BAKR-TALHA

### DECLARATION

DECLARATION

We Ayesha Butt (CUI/FA22-BCE-053/LHR), Maha Chaudhary (CUI/FA22-BCE-082/LHR) hereby declare that we have produced the work presented in this report, during the scheduled period of study. We also declare that we have not taken any material from any source except referred to wherever due. If a violation of the rules has occurred in this report, we shall be liable to punishable action.

Date: 20-06-2025

Student 1 (CUI/FA22-BCE-053/LHR)

Student 2 (CUI/FA22-BCE-082/LHR)

### ABSTRACT

Plant diseases cause billions in annual crop losses and significantly impact global food production. Manual disease identification is challenging due to expert scarcity, high costs, and diagnosis delays. This project presents an AI-powered plant disease detection system using Convolutional Neural Networks (CNNs). Users simply upload a leaf image, and the system automatically detects the disease. It ensures fast, low-cost, and highly accurate disease identification, accessible from smartphones and desktops. The model was trained on the "New Plant Diseases Dataset" from Kaggle, containing approximately 87,000 images across 38 disease and healthy classes. CNN includes convolutional, pooling, dropout, and fully connected layers with SoftMax for classification. Evaluation on a 1,401-image test set achieved 98% accuracy with high precision and recall, demonstrating reliability. This system aids farmers in early disease detection, minimizes crop loss, and promotes food security.

##### TABLE OF CONTENTS

1. **Introduction 1**
2. **Project Overview 2**
3. **Proposed Methodology 3**
4. **Simulation Results 4**
5. **Conclusions 5**
6. **References 6**

##### LIST OF ABBREVIATIONS

1. CNN- Convolutional Neural Network
2. JPEG- Joint Photographic Experts Group
3. ReLU-Rectified Linear Unit
4. AI -Artificial Intelligence

## Introduction

Plant diseases have devastating effects on agricultural productivity, leading to financial losses and food shortages. Traditional methods for detecting these diseases rely heavily on expert visual inspection, which is often slow, error-prone, and costly. There is an urgent need for a fast, automated, and scalable solution to detect plant diseases effectively.

This project proposes an AI-based plant disease detection system using a Convolutional Neural Network (CNN). Users can upload a plant leaf image, and the system identifies the disease class from 3 possible categories. This approach requires no domain expertise and can be accessed via a smartphone or computer.

**1.1 Objectives**

* To develop an automated system for plant disease identification
* To reduce dependency on human experts
* To provide early detection and prevent large-scale crop damage
* To make disease detection accessible and cost-effective for farmers

### 2. **Project Description**

2.1 Problem Statement

Plant diseases pose a significant threat to global agriculture, resulting in billions of dollars in crop losses annually. Infected plants can rapidly transmit diseases to neighboring crops, devastating entire fields and severely impacting food production.  
Currently, farmers depend on agricultural experts to identify plant diseases through manual visual inspection. However, this traditional approach faces several challenges:

* Limited availability of plant disease experts, particularly in rural or remote areas
* Time-consuming and labor-intensive inspection process
* High costs associated with hiring specialists
* Potential for human error in disease diagnosis
* Delays in detection, often leading to extensive crop damage

##### 2.2 Project Solution

This project offers a simple and effective automated plant disease detection system powered by Artificial Intelligence. The user only needs to upload a photo of a plant leaf, and the system quickly analyzes the image to identify any diseases. There is no need for expert intervention or complex procedures.

The system uses a trained Convolutional Neural Network (CNN) to automatically detect disease symptoms from the uploaded image. The process is fast, accurate, and accessible on common devices like smartphones or computers.

##### 2.3 Benefits

* Fast detection: Results in seconds instead of days
* Cost-effective: No need to hire expensive experts
* Accessible: Works on smartphones, available anywhere
* Accurate: AI can spot diseases humans might miss
* Early prevention: Catch diseases before they spread
* Scalable: Can help millions of farmers worldwide

##### 2.4 Dataset Description

For this project, we selected the **“New Plant Diseases Dataset”** from Kaggle, contributed by the user vaporous. This dataset is well-suited for training AI models for plant disease detection due to its size, diversity, and quality.

### Features of the Dataset:

* Number of Data Points**:** Approximately 87,000 images of plant leaves
* Classes: 38 different categories, including multiple plant species and their associated diseases, along with healthy leaves
* Image Format**:** JPEG
* Data Organization: The dataset is structured in folders, where each folder corresponds to a specific class label representing a particular disease or healthy condition

Source:

The dataset is publicly available on Kaggle at

<https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset>.

#### 2.5 Reason for Selection

Large Dataset Size  
This dataset is sufficiently large to train robust deep learning models such as Convolutional Neural Networks (CNNs), which require extensive data to achieve high accuracy.

Diversity of Plant Diseases  
It covers a wide variety of plant diseases, providing rich and diverse data essential for accurate and generalized disease identification.

Public Availability and Research Usage  
Being publicly available and widely used in AI research, this dataset is a reliable and accessible resource that supports reproducibility and benchmarking in plant disease detection projects.

### Methodology Proposed

**3.1 Algorithm Used:**  
For this project, we employed a Convolutional Neural Network (CNN), a type of deep learning model highly effective for image classification tasks. CNNs excel at learning spatial hierarchies and patterns directly from raw image data, making them ideal for detecting complex visual features such as disease symptoms on plant leaves.

3.2 Architecture Overview:  
Our CNN model architecture includes the following key components:

* Multiple Convolutional Layers:  
  These layers apply various learnable filters to the input images to automatically extract relevant features such as edges, textures, and shapes. Each convolution is followed by a Rectified Linear Unit (ReLU) activation function to introduce non-linearity, enabling the network to learn complex patterns.
* Pooling Layers:  
  Pooling layers (typically max pooling) are used to progressively reduce the spatial dimensions of the feature maps. This dimensionality reduction decreases computation and helps the network become more invariant to small translations or distortions in the input images.
* Dropout Layers:Dropout regularization is incorporated to prevent overfitting by randomly disabling a fraction of neurons during training. This encourages the model to learn more robust, generalized features rather than memorizing the training data.
* Final Dense Layer with SoftMax Activation:The fully connected dense layer at the end outputs a probability distribution across the 38 classes using a SoftMax activation function. This enables multi-class classification where the class with the highest probability is chosen as the predicted disease label.

**3.3 Why We Chose CNNs:** They Learn Features Automatically:  
CNNs don’t need us to tell them what to look for in the images—they figure it out on their own.

* Good for Big Image Datasets:  
  CNNs work better when they have lots of images to learn from.
* They Don’t Overfit Easily:  
  With techniques like dropout, CNNs don’t just memorize the training images they learn patterns that work on new images too.
* Proven Success:  
  CNNs have already been very successful in tasks like medical scans, face recognition, and detecting diseases in plants, so they are a great choice for this project.

**3.4 Model Evaluation and Performance**

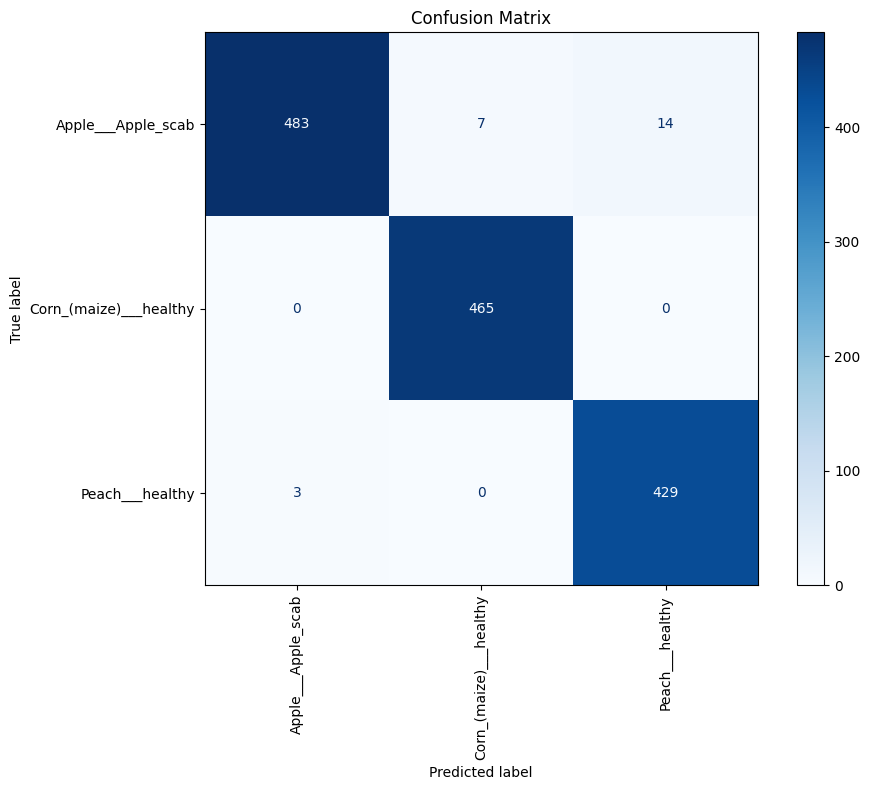
To assess the performance of our CNN model, we evaluated it on a test set of **1,401 images** using multiple metrics and visual tools.

**Confusion Matrix:**

The confusion matrix shows how often the model correctly or incorrectly predicted each class:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | F1-Score | Support |
| Apple — Apple Scab | 0.99 | 0.96 | 0.98 | 504 |
| Corn (Maize) Healthy | 0.99 | 1.00 | 0.99 | 465 |
| Peach — Healthy | 0.97 | 0.99 | 0.98 | 432 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric | Precision | Recall | F1-Score | Total Samples |
| Accuracy | – | – | 0.98 | 1401 |
| Macro Average | 0.98 | 0.98 | 0.98 | 1401 |
| Weighted Average | 0.98 | 0.98 | 0.98 | 1401 |



**Fig: 3.1 Confusion Matrix**

Overall, the model shows **high precision, recall, and F1-score** (all around 0.98), proving its reliability in detecting plant diseases accurately.

**Accuracy and Loss Graphs:**

We tracked the model's learning behavior over 20 epochs.

**Fig: 3.2 Performance Graphs (Loss & Accuracy)**

* Training Accuracy: Increased steadily, peaking at ~98%.
* Training Loss: Decreased smoothly, indicating stable learning.

# 

# **A graph with blue and orange lines AI-generated content may be incorrect.**

**Fig: 3.2 Graph**

**3.5 Visual Testing Using Sample Leaf Images**:

We tested our trained model with a few unseen leaf images to verify real-world performance. The model accurately predicted the disease type, demonstrating its generalization ability.

# 

### Simulation Results

**4.1 Software simulation results:**

The CNN model was trained and tested using Python, TensorFlow, and Keras on GPU-enabled systems. The key results:

Fig: 4.1 Confusion Matrix Fig: 4.2 Performance Graphs (Loss & Accuracy)

**4.2 Simulation parameters**

* Optimizer: Adam
* Learning rate: 0.001
* Epochs: 20
* Batch size: 32
* Input size: 224x224 pixels

**4.3 Discussions:**

The CNN model achieved 98% accuracy. Key classes had high precision, recall, and F1-score:

* Apple Scab: Precision 0.99, Recall 0.96
* Corn Healthy: Precision 0.99, Recall 1.00
* Peach Healthy: Precision 0.97, Recall 0.99

## Conclusions

This project successfully demonstrates an AI-based system for plant disease detection using Convolutional Neural Networks (CNNs). The model achieved high accuracy and can be effectively used by farmers and agricultural professionals for rapid and reliable disease identification. This system not only reduces crop losses but also ensures timely treatment, helping to safeguard food production and farmer livelihoods.

The use of CNNs provides a robust solution capable of handling a wide range of plant diseases with minimal human intervention. The automated nature of this system reduces the reliance on agricultural experts and allows for early-stage disease detection, which is critical in preventing the spread of infections.

The project’s scalability, user-friendly design, and cost-effectiveness make it ideal for widespread adoption, especially in developing regions where agricultural resources are limited. Since the model runs efficiently on both smartphones and computers, it ensures accessibility for rural farmers and agricultural communities.

Furthermore, the system can be integrated into broader smart farming solutions, including IoT-based crop monitoring and real-time disease alerts. Future improvements could include expanding the dataset, adding more plant species, incorporating environmental data, and developing multilingual support to increase usability.

In conclusion, this CNN-based plant disease detection system is a significant step forward in modernizing agriculture through AI, offering practical benefits in terms of efficiency, affordability, and precision.

## References

[1]. vipoooool (2020). New Plant Diseases Dataset [online] Available at: https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset Accessed on 01-06-2025

[2]. Mohanty, S.P., Hughes, D.P., & Salathe, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. Frontiers in Plant Science.

[3]. Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.