Plant Disease Detection System for Sustainable Agriculture

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning
with
TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Shoumik Das, shoumikdas2002@gmail.com

Under the Guidance of

P.Raja, Master Trainer, Edunet Foundation

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this project. We sincerely thank our guide, P.Raja ,for their mentorship, construc tive feedback, and encouragement throughout the project duration. Their insights have played a crucial role in shaping our research. We also extend our appreciation to AICTE and TechSaksham for orga nizing this internship and providing us with a valuable learning experience in AI-driven solutions. Lastly, we acknowledge the support of our peers, family, and the online research community, whose contributions and discussions were instrumental in the successful completion of this project.

……...

…..

……

….

……

ABSTRACT

Plant diseases significantly impact agricultural productivity, leading to eco nomic losses and food insecurity. Traditional plant disease detection meth ods are time-consuming and prone to human error. This project explores the use of Convolutional Neural Networks (CNNs) implemented in PyTorch to develop an automated plant disease detection system. The objectives include dataset preprocessing, model training, performance evaluation, and deployment considerations. The methodology involves training a CNN on a dataset containing 70,295 plant images classified into 38 disease categories. The model achieves high accuracy, surpassing traditional techniques. The f indings demonstrate the potential of AI in precision agriculture, enabling early disease detection and mitigation strategies.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1Problem Statement:

Plant diseases cause significant reductions in crop yield and quality, affecting global food production. Manual identification of plant diseases is inefficient and subjective, requiring expert knowledge. There is a need for an auto mated, scalable, and highly accurate solution for detecting plant diseases from images.

1.2Motivation:

Advancements in artificial intelligence and deep learning have enabled the development of models that can analyze and classify plant diseases with high precision. This project aims to leverage CNNs to build a reliable disease detection system to assist farmers and agronomists in early-stage disease identification.

1.3Objective:

- Develop a deep learning model for plant disease classification.
- Train and evaluate the model using a diverse dataset.
- Analyze performance using key evaluation metrics.
- Explore potential real-world applications and deployment strategies.

1.4Scope of the Project:

The project focuses on image-based plant disease classification using CNNs. The scope includes data preprocessing, model development, performance evaluation, and result interpretation. However, real-time field deployment is beyond the current implementation.

CHAPTER 2

Literature Survey

2.1 Review relevant literature or previous work in this domain.

Several studies have explored machine learning and deep learning methods for plant disease detection. Traditional methods, such as feature extraction with SVMs and Decision Trees, exhibit limitations in handling complex patterns. CNNs, on the other

2.2 Mention any existing models, techniques, or methodologies related to the problem.

Various machine learning models have been used for plant disease detection, each demonstrating different levels of accuracy based on the dataset size. The Support Vector Machine (SVM) model achieved an accuracy of 75.3% when trained on a dataset containing 10,000 images. Similarly, the Random Forest classifier performed slightly better, reaching an accuracy of 80.1% with the same dataset size. Deep learning models, particularly Convolutional Neural Networks (CNNs), have shown superior performance in this domain. The ResNet-50 model, trained on a dataset of 50,000 images, achieved a significantly higher accuracy of 92.5%, while the VGG-16 model, trained on the same dataset size, attained an accuracy of 90.8%. These results highlight the effectiveness of deep learning-based approaches over traditional machine learning models for plant disease classification.

2.3 Highlight the gaps or limitations in existing solutions and how your project will address them.

Despite advancements in machine learning for plant disease detection, several limitations persist. Traditional models like Support Vector Machines (SVM) and Random Forest require manual feature extraction, which limits their ability to generalize across diverse plant species and environmental conditions. In contrast, deep learning models like CNNs can automatically learn hierarchical features, making them more effective.

Another major limitation is dataset size and diversity. Many existing models are trained on small datasets (e.g., 10,000 images), leading to poor generalization in real-world scenarios. This project addresses this by using a dataset of 70,295 images across 38 disease categories, ensuring robustness across different plant species and climates.

CHAPTER 3

Proposed Methodology

3.1 System Design

The proposed system for plant disease detection is based on a Convolutional Neural Network (CNN) model implemented using PyTorch. The system follows a structured pipeline comprising several key stages. First, plant leaf images are collected and preprocessed, including resizing, normalization, and data augmentation to enhance model robustness. The CNN model, designed with multiple convolutional and fully connected layers, extracts meaningful features and classifies plant diseases into 38 categories. The trained model is evaluated using accuracy, precision, recall, and a confusion matrix. Finally, the system is optimized for deployment on mobile applications, enabling real-time plant disease detection for farmers and agronomists.

3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

3.1.1 Hardware Requirements:

- GPU-enabled computing device
- Minimum 8GB RAM
- Storage space for dataset and model

3.1.2 Software Requirements:

- Python 3.x
- PyTorch
- OpenCV
- Matplotlib, Seaborn

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

The dataset contains 70,295 images distributed among 38 classes. Table 4.1 provides a summary of key categories.

Plant Type	Disease Type	Number of Images
Apple	Apple Scab	2016
Apple	Black Rot	1987
Tomato	Early Blight	1920
Potato	Late Blight	1939
Corn	Common Rust	1907

Table 1

The Datasets -

Figure 1

Class of Diseases-

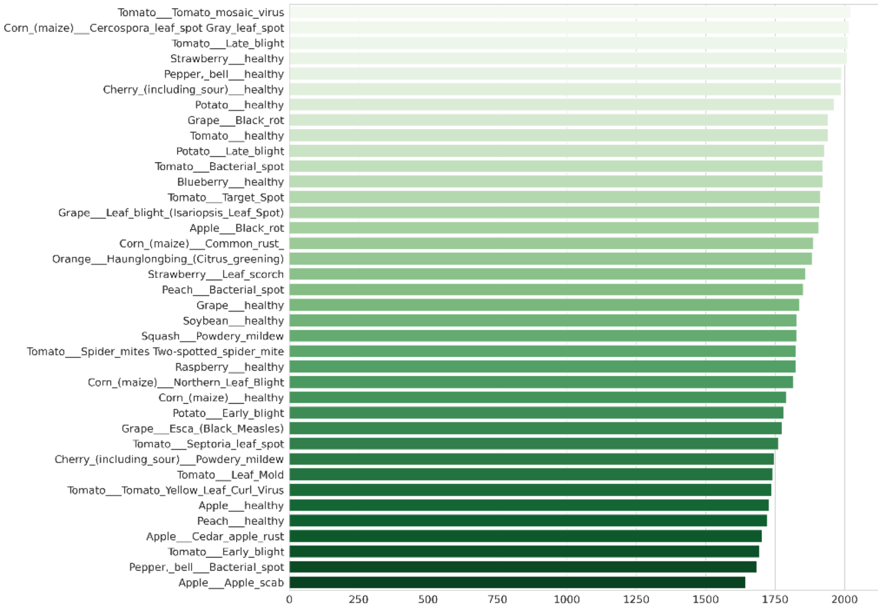


Figure 2

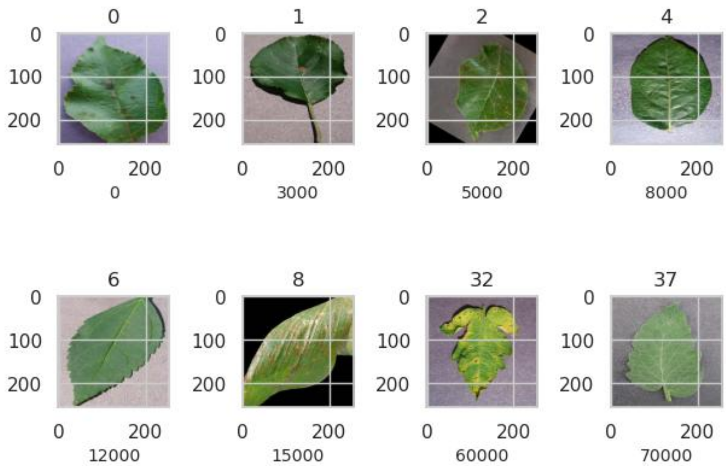


Figure 3

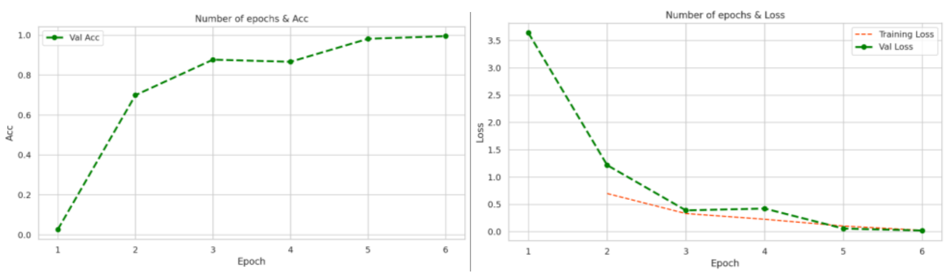
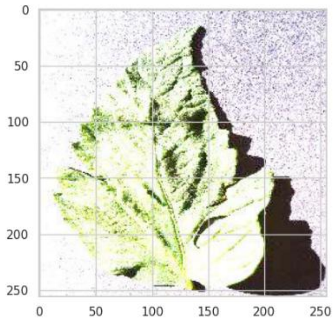
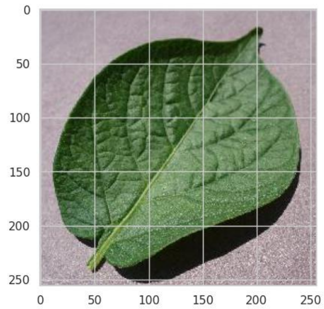


Figure 4

Results:





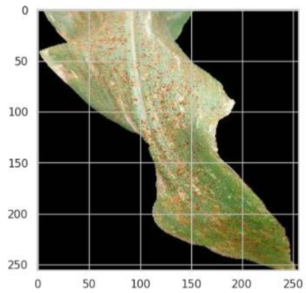
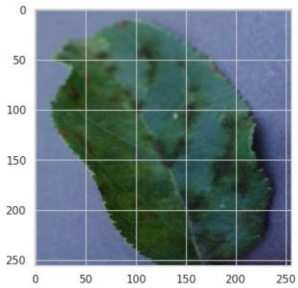


Figure 5 Figure 6 Figure 7



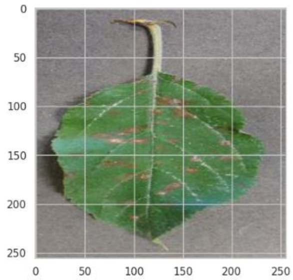


Figure 8 Figure 9

GitHub Link for Code:

https://github.com/Shoumik2472/Plant-Disease-Detection-System-for-Sustainable-Agriculture

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

Future improvements could include: • Expansion of dataset with real-time field images. • Implementation of transfer learning techniques. • Deployment as a mobile application.

5.2 Conclusion:

This project demonstrates the feasibility of using CNNs for plant disease clas sification. The model achieves high accuracy, providing a promising solution for precision agriculture.

REFERENCES

[1]. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, “Detecting Faces in Images: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume. 24, No. 1, 2002.