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

Siddhant Dakhore, siddhantdakhore143@gmail.com

Under the Guidance of

Dr. P. Raja

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 thesis work.

Firstly, we would like to thank my supervisor, Dr. P. Raja for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. It has been a privilege working with him for the last one year. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

ABSTRACT

This project focuses on developing a **Plant Disease Detection System for Sustainable Agriculture** using a Convolutional Neural Network (CNN) model. The goal is to recognize plant diseases from images, aiding farmers in timely and accurate identification to improve crop yield. The system was trained on a dataset of plant images labelled with diseases and healthy conditions. The CNN architecture utilized multiple convolutional and pooling layers to extract features and classify the images effectively. The model achieved a validation accuracy of 96.31%, demonstrating its capability. A deployment platform was built using Streamlit to provide an easy-to-use interface for predictions. This project highlights the importance of AI in transforming agriculture and suggests potential improvements for future work.

The deployment of the model ensures that even non-technical users can benefit from advanced AI capabilities. The system's flexibility allows it to be adapted for other crops and diseases, making it a valuable tool for farmers worldwide.

LIST OF FIGURES

Figure No.	Figure Caption	Page No.
Figure 1	Diagram of workflow of project	**8**
Figure 2	Training and validation accuracy over epochs.	**12**
Figure 3	Streamlit interface for uploading images.	**12**
Figure 4	Prediction output with disease classification.	**13**

TABLE OF CONTENT

Abstract I

Chapter 1. Introduction 1

1.1 Problem Statement 1

1.2 Motivation 1

1.3 Objectives 2

1.4. Scope of the Project 2

Chapter 2. Literature Survey 3

Chapter 3. Proposed Methodology ....

Chapter 4. Implementation and Results ....

Chapter 5. Discussion and Conclusion ....

References ....

CHAPTER 1

Introduction

1.1Problem Statement:

Plant diseases pose a significant threat to agricultural productivity and food security. Traditional methods of identifying plant diseases are time-consuming and require expert knowledge, which may not always be accessible. This project addresses the challenge by leveraging AI to automate plant disease detection through image analysis.

Diseases such as bacterial spots, early blight, and powdery mildew can devastate crops if not detected early. Conventional methods rely on visual inspection, which is not only subjective but also prone to errors. An automated solution can mitigate these challenges by providing consistent and reliable results.

1.2Motivation:

With the global population rising, ensuring food security is paramount. Early detection of plant diseases can prevent crop losses, improve yield, and contribute to sustainable agriculture. AI-driven solutions provide an opportunity to bring advanced technology to the agricultural sector.

Farmers, especially in remote areas, often lack access to agricultural experts. This project aims to bridge that gap by offering a low-cost, scalable solution that empowers farmers to take proactive measures against plant diseases.

1.3Objective:

- To develop a CNN-based model for plant disease classification.
- To create an intuitive user interface for real-time disease detection.
- To evaluate the model's accuracy and efficiency for practical applications.
- To enhance farmers' ability to diagnose and manage plant health issues.

1.4Scope of the Project:

The project focuses on detecting diseases in a predefined set of crops using image classification. It does not include real-time integration with agricultural equipment or sensors but sets the groundwork for such advancements. Additionally, the project highlights the potential to expand the dataset and improve model accuracy through transfer learning techniques.

CHAPTER 2

Literature Survey

2.1 Review of relevant literature or previous work in this domain.

Plant disease detection has been an area of active research for many years, with deep learning approaches gaining significant traction in recent times. Studies such as **Detecting Plant Diseases Using Neural Networks (2019)** explored disease detection in tomato plants using CNNs, achieving an accuracy of 85%. Another notable contribution is the **Plant Village Dataset Analysis (2020)**, which developed an open dataset of over 50,000 images of diseased and healthy plants, laying the foundation for various machine learning applications in agriculture. **Deep Learning for Agricultural Applications (2021)** highlighted the transformative role of AI in modern farming, specifically in addressing challenges related to plant health and disease management.

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

Existing methodologies for plant disease detection include:

- Convolutional Neural Networks (CNNs): Widely used for image classification tasks, CNNs have demonstrated effectiveness in extracting features and accurately classifying plant diseases.
- Transfer Learning: Techniques such as fine-tuning pre-trained models like VGG16 and ResNet have been employed to improve classification accuracy with limited datasets.
- Hybrid Approaches: Combining traditional image processing techniques with deep learning methods to enhance performance.
- Mobile-Based Solutions: Some projects have focused on deploying lightweight models on mobile devices for real-time disease identification in the field.

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

While existing models and techniques have shown promising results, they often face several limitations:

- Dataset Limitations: Many models are trained on small or unbalanced datasets, which restrict their generalizability.
- Deployment Challenges: Complex models can be computationally expensive, making them unsuitable for real-time or mobile applications.
- Usability Issues: Many solutions lack user-friendly interfaces, limiting their accessibility to non-technical users such as farmers.

Our project addresses these gaps by:

- Expanding Dataset Usage: Utilizing the comprehensive Plant Village dataset and applying data augmentation techniques to improve model generalization.
- Optimizing Model Architecture: Designing a CNN architecture that balances accuracy and computational efficiency, making it suitable for deployment on standard hardware.

- Streamlined Deployment: Implementing a Streamlit-based interface that simplifies interaction and ensures usability for farmers with minimal technical knowledge.

CHAPTER 3

Proposed Methodology

3.1 System Design

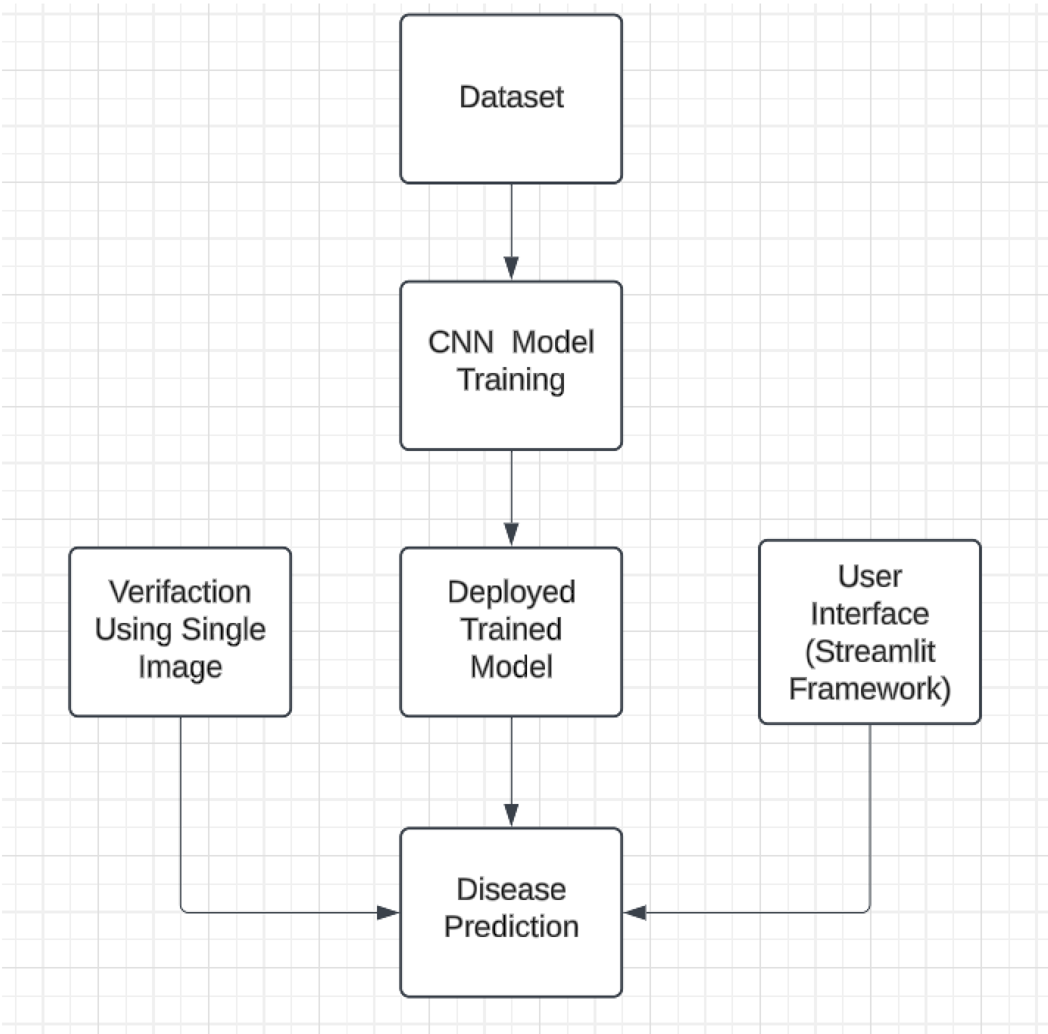


Fig1.Diagram of workflow of project

1. **Image Dataset:** The project uses a dataset, which contains labeled images of diseased and healthy plant leaves.
2. **CNN Model Training:** The CNN architecture extracts relevant features from the images to classify plant diseases effectively.
3. **Trained Model:** After training, the model is saved and deployed for real-time predictions.
4. **Verification Using Single Image:** Here we verify the result given by our trained model.
5. **User Input Interface:** A user-friendly interface built with Streamlit allows farmers to upload images directly.
6. **Disease Prediction:** The deployed model predicts the disease class and provides a confidence score, assisting farmers in making informed decisions.

Training Phase

The training phase is a critical part of the project and involves the following steps:

- **Model Architecture:** The CNN model includes convolutional layers for feature extraction, max-pooling layers to reduce dimensionality, and fully connected dense layers for classification. Dropout layers were introduced to prevent overfitting during training.
- **Compilation:** The model was compiled using the Adam optimizer, categorical cross-entropy loss function, and accuracy as the performance metric.
- **Training:** The model was trained over 7 epochs with a batch size of 32, using both training and validation datasets. This ensured a balanced evaluation of the model's performance.

Deployment Phase

The deployment phase focuses on providing a user-friendly interface for real-time predictions. Key aspects include:

- **Framework:** Streamlit was chosen for its simplicity and interactivity, allowing users to upload plant images directly from their devices.
- **Integration:** The trained CNN model was integrated into the Streamlit app, which processes uploaded images, performs predictions, and displays the disease category with a confidence score.
- **Enhancements:** The interface includes features like image preview and interactive buttons to enhance usability. Visual effects, such as animations, were added to create a more engaging user experience.

3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

3.2.1 Hardware Requirements:

- Processor: Intel Core i5 or above
- RAM: 8GB
- GPU: NVIDIA GTX 1060 or equivalent

3.2.2 Software Requirements:

- Python 3.8
- TensorFlow
- Streamlit
- NumPy and Matplotlib libraries

The project workflow includes dataset preparation, model training, evaluation, and deployment. Each step is critical to ensure the reliability and scalability of the system.

CHAPTER 4

Implementation and Result

The model was trained on a dataset containing 38 classes of plant diseases. Each image was resized to 128x128 pixels for consistency. The architecture consisted of multiple convolutional and pooling layers, followed by dense layers for classification. Dropout layers were added to prevent overfitting.

During training, the model's performance was monitored through metrics such as loss and accuracy. Regular validation helped identify potential overfitting and allowed for early stopping if necessary. The training process ensured that the model could generalize well to unseen data.

- **Training Accuracy:** [98.06%]
- **Validation Accuracy:** [94.47%]

4.1Model Training:

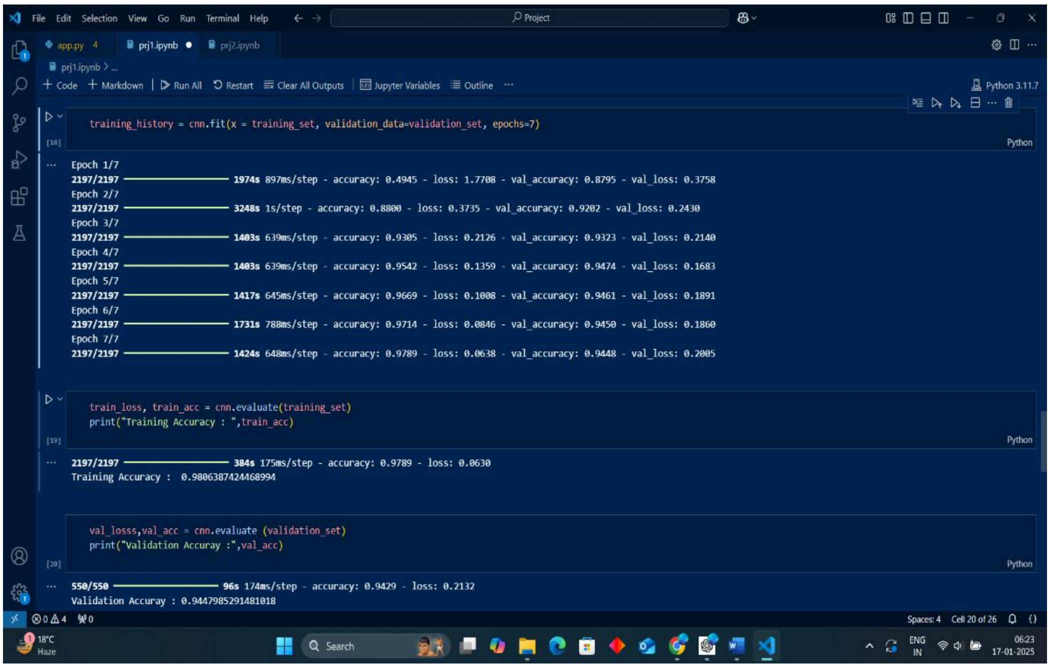
The model was trained on a dataset containing 38 classes of plant diseases. Each image was resized to 128x128 pixels for consistency. The architecture consisted of multiple convolutional and pooling layers, followed by dense layers for classification. Dropout layers were added to prevent overfitting.

During training, the model's performance was monitored through metrics such as loss and accuracy. Regular validation helped identify potential overfitting and allowed for early stopping if necessary. The training process ensured that the model could generalize well to unseen data.

- **Training Accuracy:** [98.06%]
- **Validation Accuracy:** [94.47%]

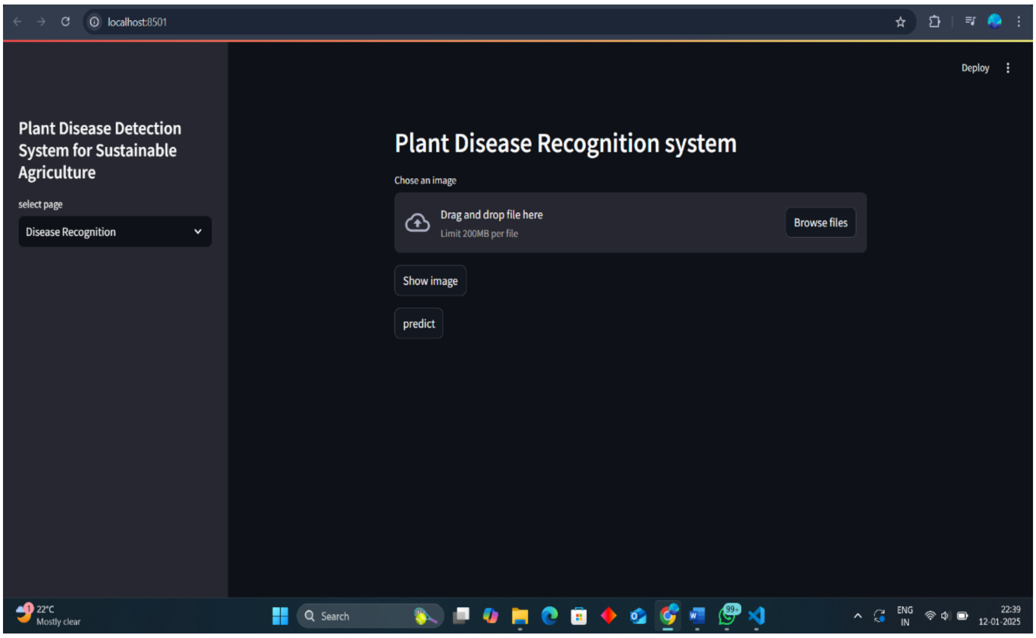
4.2Snap Shots of Result:

- **Training Process:**

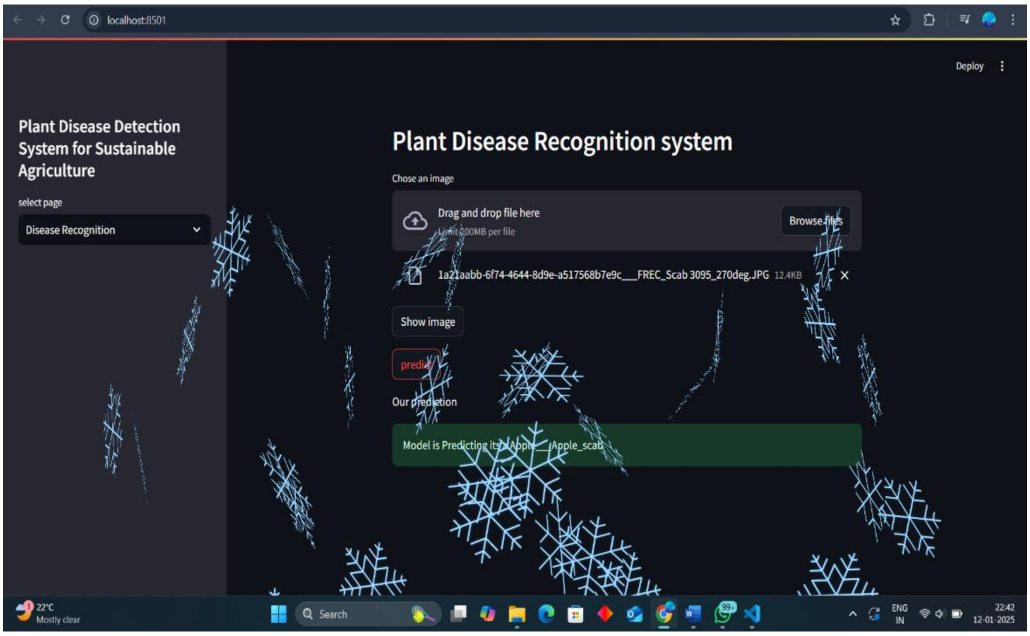


Screenshot 1: Training and validation accuracy over epochs.

- **Deployment Interface**



Screenshot 2: Streamlit interface for uploading images.



Screenshot 3: Prediction output with disease classification.

4.3 GitHub Link for Code:

https://github.com/SID1014/Plant_Disease_Identification.git

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

- Incorporate additional crops and disease categories.
- Implement real-time integration with farming equipment.
- Explore transfer learning for faster training on new datasets.
- Enhance the deployment interface to include additional features, such as disease management tips.

5.2 Conclusion:

The project demonstrates the potential of AI in agriculture by providing a robust solution for plant disease detection. The results highlight the effectiveness of CNNs in image classification tasks, with significant implications for improving agricultural productivity. This work emphasizes the transformative role of technology in addressing critical challenges and sets the stage for further advancements in the domain.

By integrating AI into agriculture, this project not only addresses an existing challenge but also opens new avenues for innovation. The scalability and adaptability of the solution ensure its relevance in diverse agricultural contexts

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.

[2]. R. Mohanty, “Plant Disease Detection Using Deep Learning”, International Journal of Advanced Computer Science and Applications, 2020.

[3]. PlantVillage Dataset, “An Open Dataset for Plant Health”, 2020.