Potato Leaf Disease Detection

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

Trisha Pal, trishapal845@gmail.com

Under the Guidance of

Jay Rathod

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 work.

Firstly, I would like to thank my supervisor, **Jay Rathod**, for being a great mentor and the best advisor I could ever have. His advice, encouragement, and constructive criticism have been a source of innovative ideas and inspiration, leading to the successful completion of this project. The confidence he showed in me was the biggest source of motivation. It has been a privilege working with him over the past few months. His guidance not only helped me in the project but also shaped me into a better and more responsible professional.

I would also like to thank **AICTE, Microsoft**, and **SAP** for providing this wonderful opportunity through the **TechSaksham** program. The resources and support provided by them were instrumental in the completion of this project.

Lastly, I would like to thank my family and friends for their constant support and encouragement throughout this journey.

ABSTRACT

Potato leaf diseases are a significant threat to global agriculture, leading to substantial crop losses and economic damage. Early detection of these diseases is crucial for effective management and prevention. This project aims to develop a **Convolutional Neural Network (CNN)** model for the automatic detection of potato leaf diseases. The model is trained on a dataset of potato leaf images, categorized into healthy, Early Blight and Late Blight. The trained model is then hosted using **Streamlit**, providing a user-friendly interface for real-time disease detection.

The project's objectives include:

1. Building a robust CNN model for accurate disease detection.

2. Creating a web-based application for easy access and usability.

3. Evaluating the model's performance using metrics such as accuracy, precision, and recall.

The results demonstrate that the CNN model achieves high accuracy in detecting potato leaf diseases, making it a valuable tool for farmers and agricultural experts. The project's future scope includes expanding the dataset, improving model performance, and integrating additional features for broader agricultural applications.

Table of Contents

LIST OF FIGURES

CHAPTER 1: Introduction

1.1Problem Statement

Potato crops are susceptible to various diseases, such as Late Blight, Early Blight etc. which can significantly reduce yield and quality. Early detection of these diseases is challenging for farmers due to the lack of expertise and resources. Manual inspection is time-consuming and often inaccurate. Therefore, there is a need for an automated system that can accurately detect potato leaf diseases and assist farmers in taking timely preventive measures

1.2Motivation

The motivation behind this project is to leverage **Artificial Intelligence (AI)** and **Deep Learning** techniques to address a critical agricultural problem. By automating the detection process, we can help farmers reduce crop losses, improve yield, and minimize the use of pesticides through targeted interventions. This project also aligns with the global goal of achieving food security and sustainable agriculture.

1.3Objective

The primary objectives of this project are:

1. To develop a CNN-based model for detecting potato leaf diseases.

2. To create a user-friendly web application using Streamlit for real-time disease detection.

3. To evaluate the model's performance and ensure its accuracy and reliability.

1.4Scope of the Project

The scope of this project is limited to the detection of common potato leaf diseases using image data. The model is trained on a specific dataset and may require further refinement for broader applications. The web application is designed for ease of use, but it currently supports only potato leaf disease detection.

CHAPTER 2: Literature Survey

2.1Overview of Plant Disease Detection:

Plant disease detection is a critical aspect of agricultural practices, as diseases can significantly reduce crop yield and quality, leading to economic losses and food insecurity. Early detection of plant diseases is essential for effective disease management and prevention. Traditionally, farmers and agricultural experts relied on manual inspection to identify diseases, which was time-consuming, labor-intensive, and often inaccurate due to human error.

With the advancement of technology, automated methods for plant disease detection have emerged, leveraging **image processing**, **machine learning**, and **deep learning** techniques. These methods aim to provide accurate, efficient, and scalable solutions for detecting diseases in crops. The use of digital images and automated systems has revolutionized the field, enabling faster and more reliable disease detection, which is crucial for sustainable agriculture.

2.2Traditional Methods for Disease Detection:

Traditional methods for plant disease detection primarily relied on **image processing techniques** and **machine learning algorithms**. These methods can be broadly categorized into the following steps:

1. **Image Acquisition:**
    - Digital images of plant leaves are captured using cameras or smartphones.
    - The quality of the images plays a crucial role in the accuracy of disease detection.

2. **Preprocessing:**
    - The images are preprocessed to enhance quality and remove noise.
    - Techniques such as resizing, normalization, and color correction are applied.

3. **Feature Extraction:**
    - Handcrafted features such as **color**, **texture**, and **shape** are extracted from the images.
    - For example, color histograms, texture descriptors (e.g., Gray-Level Co-occurrence Matrix), and shape-based features (e.g., leaf area, perimeter) are commonly used.

4. **Classification:**

- Machine learning algorithms such as **Support Vector Machines (SVM), Random Forests**, and **k-Nearest Neighbors (k-NN)** are used to classify the images into healthy or diseased categories based on the extracted features.

2.3 Gaps or limitations in existing solutions and how this project will address them.

2.3.1 Limitations in the Traditional Models:

1. Limited Dataset Diversity:

- Many existing models are trained on small or limited datasets that do not capture the full variability of plant diseases in real-world conditions.
- This limits the model's ability to generalize across different regions, lighting conditions, and growth stages.

2. High Computational Requirements:

- Deep learning models, especially those with complex architectures (e.g., VGGNet, ResNet), require significant computational resources for training and inference.
- This makes them unsuitable for resource-constrained environments, such as rural areas where farmers may not have access to high-end hardware.

3. Lack of Real-Time Detection Capabilities:

- While many models achieve high accuracy in controlled environments, their performance in real-time, field conditions is often suboptimal.
- There is a need for lightweight models that can perform real-time disease detection on mobile or edge devices.

4. Crop-Specific Models:

- Most existing models are designed for specific crops or diseases and may not generalize well to other crops or regions.
- Developing models that can detect diseases across multiple crops with high accuracy remains a challenge.

2.3.2 How This Project Addresses These Gaps:

1. Lightweight CNN Model:

- This project proposes a **Convolutional Neural Network (CNN)** model that is computationally efficient and suitable for real-time applications.
- The model is designed to be lightweight, making it accessible to farmers in resource-constrained environments.

2. Data Augmentation:

- To address the issue of limited dataset diversity, your project employs **data augmentation techniques** such as rotation, flipping, and scaling to increase the variability of the training dataset.
- This improves the model's ability to generalize across different conditions.

3. Real-Time Deployment:

- The trained model is hosted using **Streamlit**, a user-friendly web framework that allows for real-time disease detection.
- This makes the system accessible to farmers and agricultural experts without requiring high-end hardware.

4. **Focus on Potato Leaf Diseases:**

- While the model is currently focused on potato leaf diseases, the approach can be extended to other crops in the future.
- This provides a foundation for developing multi-crop disease detection systems.

5. **User-Friendly Interface:**

- The Streamlit-based web application provides a simple and intuitive interface for users to upload images and receive disease predictions.
- This makes the system accessible to non-technical users, such as farmers.

CHAPTER 3: Proposed Methodology

3.1System Design

The system design consists of the following components:

- **Data Collection:** A dataset of potato leaf images is collected, categorized into healthy, Early Blight and Late Blight classes.
- **Preprocessing:** The images are resized, normalized, and augmented to improve model performance.
- **Model Training:** A CNN model is trained on the preprocessed dataset.
- **Model Evaluation:** The model's performance is evaluated using metrics such as accuracy, precision, and recall.
- **Deployment:** The trained model is hosted using Streamlit, providing a web-based interface for real-time disease detection.

3.2Requirement Specification

3.1.1 Hardware Requirements:

- GPU-enabled system for training the CNN model.
- Minimum 8GB RAM.
- Storage for dataset and model files.

3.1.2 Software Requirements:

- Python 3.x
- TensorFlow/Keras for model development.
- Streamlit for web application deployment.
- Libraries such as NumPy, Pandas, and OpenCV for data processing.

CHAPTER 4: Implementation and Result

4.1Snap Shots of Result:

Below are some snapshots of the results obtained from the project:

1. **Snapshot 1:**

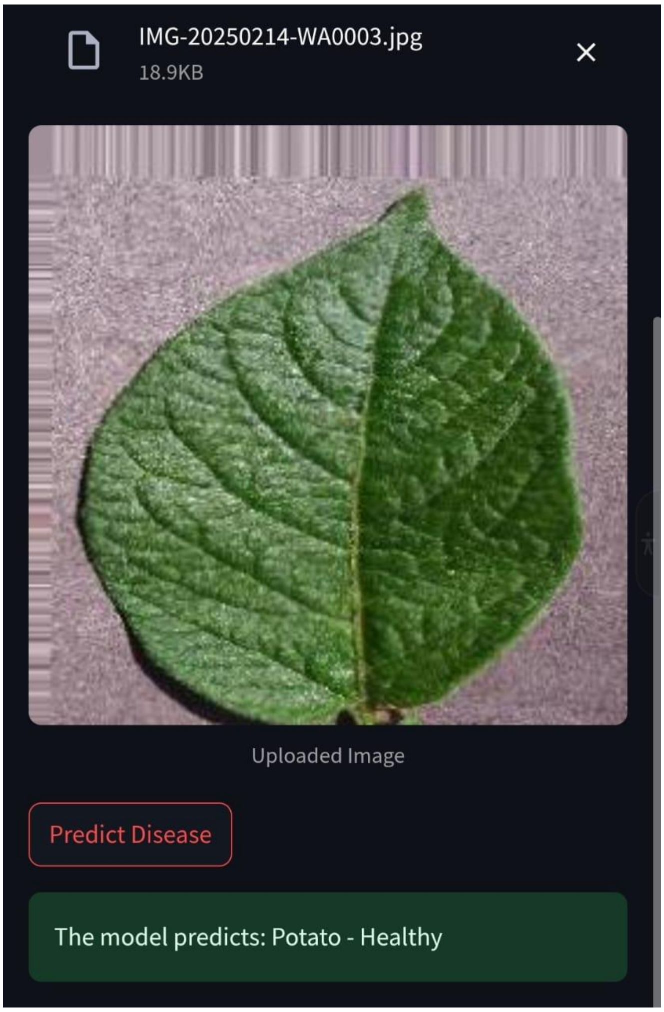

Figure 1: Snapshot of making a prediction on a healthy leaf

Description: This snapshot shows the output of the model when a healthy potato leaf image is uploaded. The model correctly classifies the leaf as healthy.

2. **Snapshot 2:**

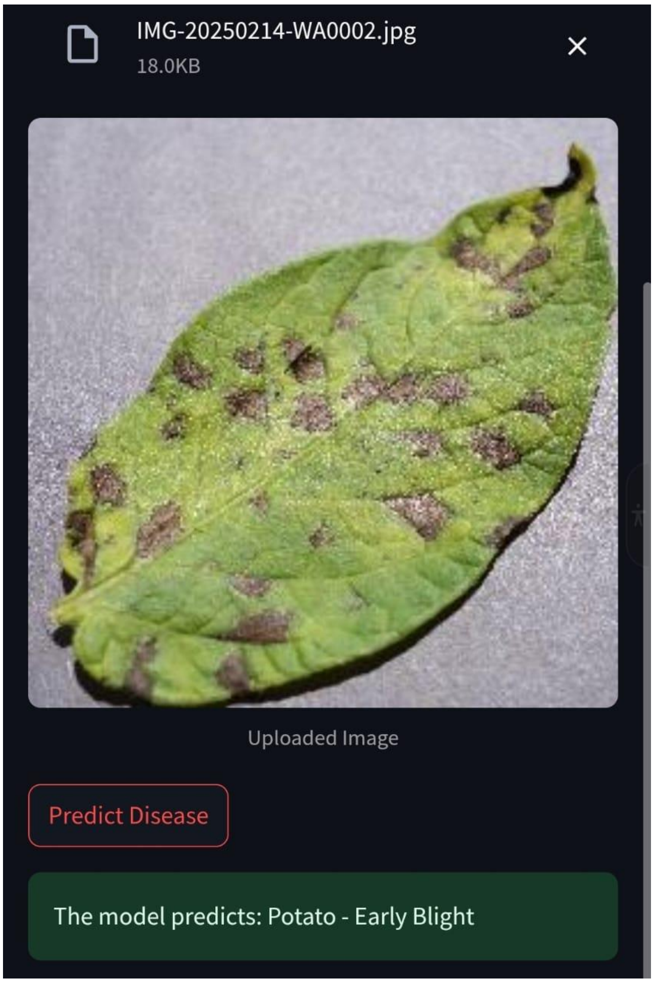

Figure 2: snapshot of the model prediction an Early Blight Disease

Description: This snapshot shows the output of the model when a diseased potato leaf image is uploaded. The model correctly identifies the disease Early Blight.

3. **Snapshot 3:**

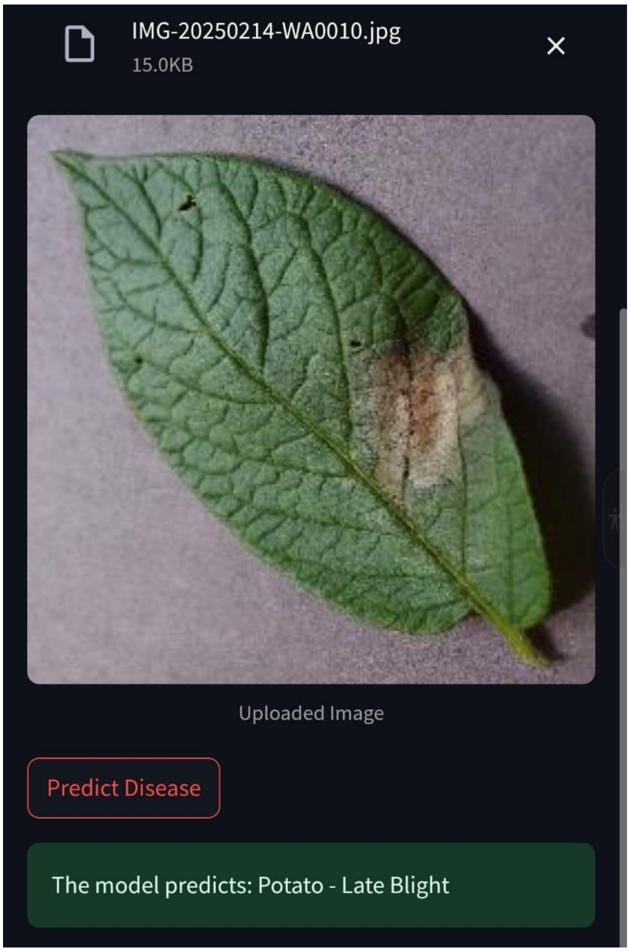

Figure 3: snapshot of the model predicting the disease Late Blight

Description: This snapshot shows the output of the model when a diseased potato leaf image is uploaded. The model correctly identifies the disease Late Blight.

4.2GitHub Link for Code:

https://github.com/trisha194/Potato-Leaf-Disease-Detection

Discussion and Conclusion

Future Work

1. **Dataset Expansion:** The model can be further improved by training it on a larger and more diverse dataset.
2. **Multi-Crop Support:** The system can be extended to detect diseases in other crops, such as tomatoes, rice, and wheat.
3**. Mobile Application:** Developing a mobile app for easier access by farmers in remote areas.

Conclusion:

This project successfully demonstrates the use of a CNN model for detecting potato leaf diseases. The model achieves high accuracy and is deployed using Streamlit, making it accessible to farmers and agricultural experts. The project has the potential to significantly impact agriculture by enabling early disease detection and reducing crop losses.

REFERENCES

[1] S. P. Mohanty, D. P. Hughes, M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection", Frontiers in Plant Science, Volume 7, 2016.

[2] K. P. Ferentinos, "Deep Learning Models for Plant Disease Detection and Diagnosis", Computers and Electronics in Agriculture, Volume 145, 2018.

[3] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification", Computational Intelligence and Neuroscience, Volume 2016, 2016.

[4] J. G. A. Barbedo, "A Review on the Main Challenges in Automatic Plant Disease Identification Based on Visible Range Images", Biosystems Engineering, Volume 144, 2016.

[5] S. Zhang, S. Zhang, C. Zhang, X. Wang, Y. Shi, "Cucumber Leaf Disease Identification with Global Pooling Dilated Convolutional Neural Network", Computers and Electronics in Agriculture, Volume 162, 2019.

[6] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems, Volume 25, 2012.

[7] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.