Potato Leaf classification disease

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

Tejavath Sriraj, srirajtejavath913@gmail.com

Under the Guidance of

Abdul Aziz MD

Master Trainer, Edunet Foundation

ACKNOWLEDGEMENT

We would like to take this opportunity to express our sincere gratitude to all individuals who contributed, directly or indirectly, to the successful completion of this thesis.

First and foremost, we extend our deepest appreciation to our mentors, **Abdul Aziz MD, Pavan, and Aditya Prashant Ardak**, for their invaluable guidance, unwavering support, and insightful advice throughout this project. Their encouragement, constructive feedback, and innovative ideas have been instrumental in shaping the direction of our research.

Their confidence in our abilities has been a constant source of motivation, and their mentorship has not only enriched our academic journey but has also contributed significantly to our professional and personal growth. The past year working under their guidance has been an enriching experience, providing us with knowledge and skills that extend beyond this project.

We are truly grateful for their mentorship and for the opportunity to learn from their expertise.

ABSTRACT

Potato crops are highly susceptible to various leaf diseases, which can significantly impact yield and quality. Early and accurate identification of these diseases is crucial for effective management and prevention of crop loss. This project presents a Convolutional Neural Network (CNN)-based approach for detecting and classifying potato leaf diseases using image processing techniques.

The proposed model is trained on a dataset consisting of images of healthy and diseased potato leaves, categorized into Early Blight, Late Blight, and Healthy classes. Preprocessing methods such as image resizing, normalization, and augmentation are applied to improve model performance. The CNN architecture is designed to extract meaningful features and classify the images with high accuracy. Model evaluation is conducted using metrics like accuracy, precision, recall, and F1-score to ensure robustness and reliability.

The results indicate that the model effectively distinguishes between healthy and diseased leaves, making it a valuable tool for farmers and agricultural professionals. Implementing this system can facilitate early disease detection, allowing timely intervention and improving overall crop productivity. Future developments may involve real-time detection via mobile applications and integration with smart farming technologies.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

1. Problem Statement:

Farmers every year face economic loss and crop waste due to various diseases in potato plants. Early blight and Late blight are major disease of potato leaf. It is estimated that the major loss occurred in potato yield is due to these diseases. Thus the images are classified into 3 classes :

1. Potato Late Blight,
2. Potato Early Blight,
3. Potato Healthy Leaf.

The process involves several steps, including importing libraries, loading and viewing data, splitting the data into training and validation sets, data pre-processing, building the CNN model, training the model, and analyzing the model's performance.

2. Motivation:

Potato is one of the most widely cultivated crops globally, playing a crucial role in food security and the agricultural economy. However, its productivity is often threatened by various leaf diseases, such as Early Blight and Late Blight, which can cause significant yield losses if not detected and treated in time. Traditional disease identification methods rely on manual inspection, which is time-consuming, prone to errors, and requires expert knowledge.

With advancements in deep learning and computer vision, automated plant disease detection has become a viable solution. Convolutional Neural Networks (CNNs) have shown remarkable performance in image classification tasks, making them a suitable choice for detecting potato leaf diseases with high accuracy. By leveraging CNN-based models, this project aims to provide a cost-effective, efficient, and scalable solution for early disease detection.

3. **Objective:**

The primary goal of this project is to develop an automated potato leaf disease detection system using Convolutional Neural Networks (CNNs). The specific objectives include:

1. To Design and Train a CNN Model – Develop a deep learning-based model capable of accurately detecting and classifying potato leaf diseases such as Early Blight, Late Blight, and Healthy leaves.

2. To Collect and Preprocess Image Data – Utilize a dataset of potato leaf images, applying preprocessing techniques like image resizing, normalization, and augmentation to enhance model performance.

3. To Achieve High Classification Accuracy – Optimize the CNN model by fine-tuning hyperparameters to maximize performance metrics such as accuracy, precision, recall, and F1-score.

4. To Enable Early Disease Detection – Provide farmers and agricultural experts with a tool that can detect diseases at an early stage, allowing for timely intervention and prevention of crop loss.

5. To Develop a Scalable and Efficient Solution – Ensure that the model is lightweight, efficient, and adaptable for real-world applications, including mobile-based and IoT-integrated systems for smart farming.

6. To Minimize Dependence on Manual Inspection – Reduce the reliance on traditional disease identification methods, which can be time-consuming and require expert knowledge, by providing an automated, user-friendly solution.

7. To Promote Sustainable Agriculture – Help in reducing excessive pesticide use by identifying diseases accurately, leading to targeted treatment and eco-friendly farming practices.

4. Scope of the Project:

The Potato Leaf Disease Detection using CNN project aims to help farmers detect diseases early using deep learning and image processing. The key areas covered by this project include:

1. Disease Detection & Classification

- Identifies and classifies common potato leaf diseases like Early Blight, Late Blight, and Healthy leaves.
- Uses CNN (Convolutional Neural Network) to accurately detect diseases from leaf images.

2. Image Processing & Model Training

- Processes images using techniques like resizing, normalization, and augmentation to improve accuracy.
- Trains and tests the CNN model on a dataset of potato leaf images.
- Evaluates model performance using accuracy, precision, and recall.

3. Real-World Applications

- Can be used by farmers, agricultural experts, and researchers for quick disease identification.
- Reduces the need for manual inspection, making disease detection faster and more reliable.
- Helps in early intervention, preventing crop loss and improving yield.

4. Future Possibilities

- Can be integrated into a mobile app where farmers can upload photos for instant disease detection.
- Potential for real-time monitoring using smart cameras or IoT devices.
- Can be expanded to detect diseases in other crops, making it more useful for agriculture.

5. Benefits to Farmers

- Saves time and effort by automating disease detection.
- Reduces costs by minimizing unnecessary pesticide use.
- Improves crop yield by allowing early disease treatment.

CHAPTER 2

Literature Survey

Potato Leaf Disease Classification Using CNN

Working:

The classification of potato leaf diseases using Convolutional Neural Networks (CNNs) has gained significant attention in recent years due to advancements in deep learning and computer vision. Several research studies have explored different approaches to detecting and classifying plant diseases, particularly in potatoes. This literature survey highlights key findings from previous works related to potato leaf disease classification using CNN models.

Limitations:

hile the **Potato Leaf Disease Detection Using CNN** project is useful, it has some challenges that need to be considered:

1. Requires High-Quality Data

- The model needs a large and diverse dataset to work accurately.
- Poor lighting, different angles, or unclear images may lead to wrong predictions.

2. May Not Work Well in All Conditions

- The model might not perform well in different weather, soil, or lighting conditions.
- Disease symptoms may look different based on location, affecting accuracy.

3. Needs Powerful Computers for Training

- Training deep learning models requires high-performance computers.
- Running the model on mobile devices or low-end systems may be slow.

4. Difficulty in Differentiating Similar Diseases

- Some diseases have similar symptoms, leading to misclassification.
- The model may not recognize cases where multiple diseases affect a leaf at the same time.

Improvments:

To enhance the **Potato Leaf Disease Detection Using CNN** project, several improvements can be made to increase accuracy, efficiency, and real-world usability. Here are some key areas for improvement:

1. Improve Dataset Quality & Diversity

2. Use Transfer Learning & Advanced CNN Architectures

3. Optimization for Mobile & Edge Devices

4. Enhance Real-Time Detection

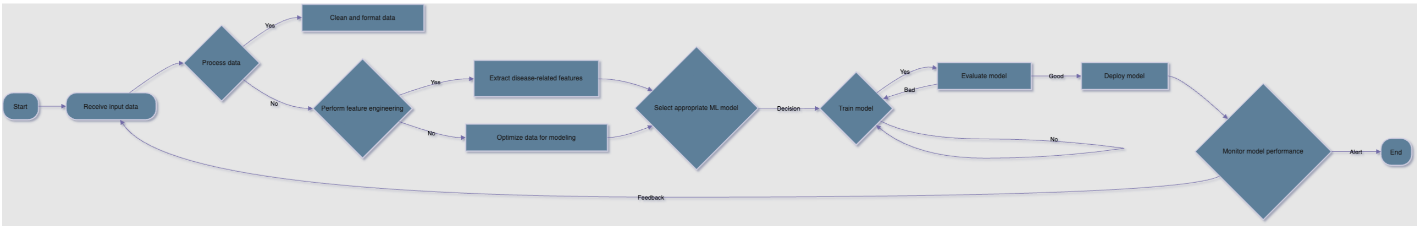
5. Improve Model Accuracy & Robustness

6. Reduce False Positives & Misclassifications

CHAPTER 3

Proposed Methodology

1. System Design



1. Start

- The process begins with receiving input data (images of potato leaves).
- These images may come from farmers, drones, or field cameras.

2. Receive Input Data

- The system takes raw images of potato leaves as input.
- These images might contain noise, inconsistent lighting, or other distortions.

3. Process Data

- This step involves cleaning and formatting the data if required.
- If the data needs preprocessing, it is sent to the "Clean and Format Data" step.
- If the data is already in a good format, it proceeds further.

4. Perform Feature Engineering

- The system checks whether feature engineering is needed.
- If required, disease-related features (such as color, texture, shape) are extracted.
- Otherwise, the data is optimized for better modeling.

5. Select Appropriate Machine Learning (ML) Model

- The system decides which CNN-based model (or another ML model) to use for classification.
- Common CNN architectures like ResNet, MobileNet, or EfficientNet can be used.

6. Train Model

- The selected model is trained using the preprocessed dataset.
- Training involves adjusting the model parameters to detect diseases effectively.

7. Evaluate Model

- After training, the model's performance is evaluated.
- If the model performs well, it is deployed.
- If it performs poorly, feedback is provided, and the model is re-trained with improvements.

8. Deploy Model

- If the model achieves high accuracy, it is deployed for real-world use.
- Deployment could be on a mobile app, web-based tool, or IoT devices in farms.

9. Monitor Model Performance

- The model is continuously monitored for accuracy and reliability.
- If issues arise, alerts are generated, and improvements are made.

10. End

- The process concludes after successful **model deployment and monitoring**.

2. Requirement Specification

2.1. Hardware Requirements:

- Processor (CPU): Intel Core i7/i9 or AMD Ryzen 7/9 (or better)
- Graphics Card (GPU): NVIDIA GPU (RTX 3060, RTX 4090, Tesla V100, or A100) for faster training
- RAM: Minimum 8 GB (Recommended: 16GB for large datasets)
- Storage: At least 512GB SSD (Recommended: 1TB SSD for handling image datasets)
- Power Supply: High-wattage power supply if using a dedicated GPU

2.2. Software Requirements:

A. Type of the system

- Windows 10/11, Linux (Ubuntu 20.04+), or macOS

B. Programming Languages & Frameworks

- Python 3.8+ (Primary language for deep learning models)
- TensorFlow / Keras / PyTorch (Deep learning frameworks)
- OpenCV (For image preprocessing and enhancement)
- NumPy & Pandas (For data handling)
- Matplotlib & Seaborn (For visualization)

C. Dataset & Storage

- Kaggle / Custom Dataset (Potato leaf disease images)
- Google Drive / AWS S3 / Firebase (For storing training data and models)
- SQLite / MySQL (If a database is required for storing results)

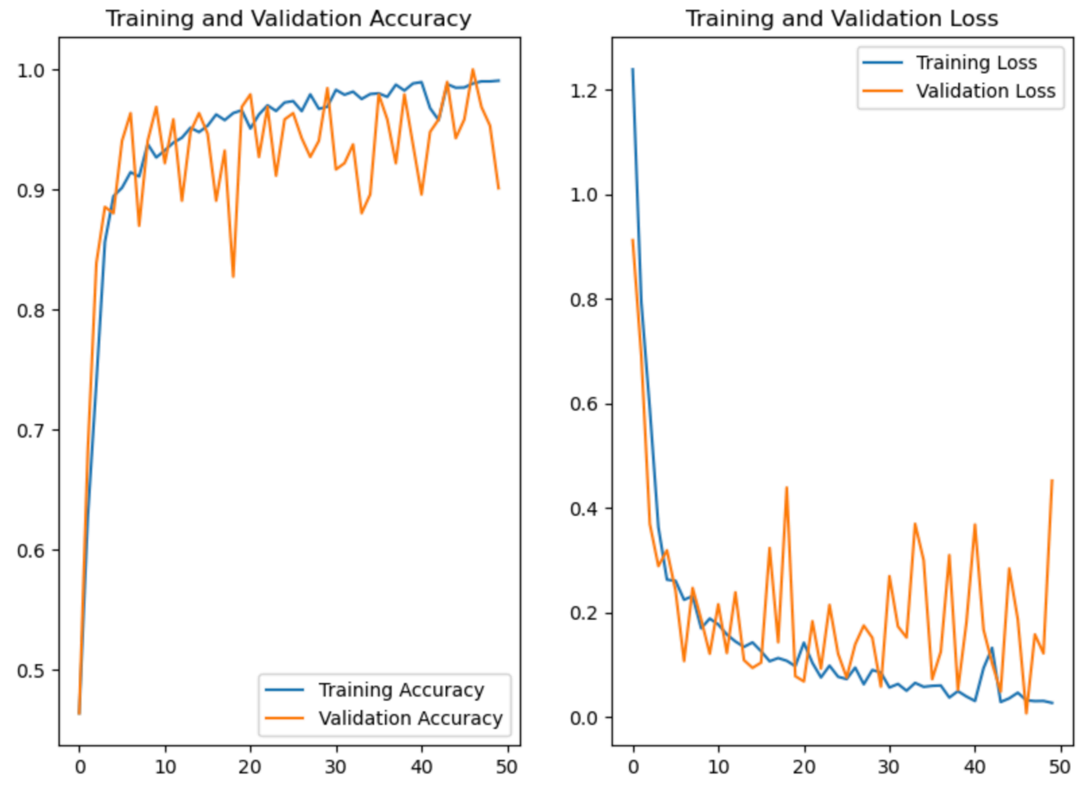
D. Model Training & Deployment

- Jupyter Notebook / Google Colab (For training models)

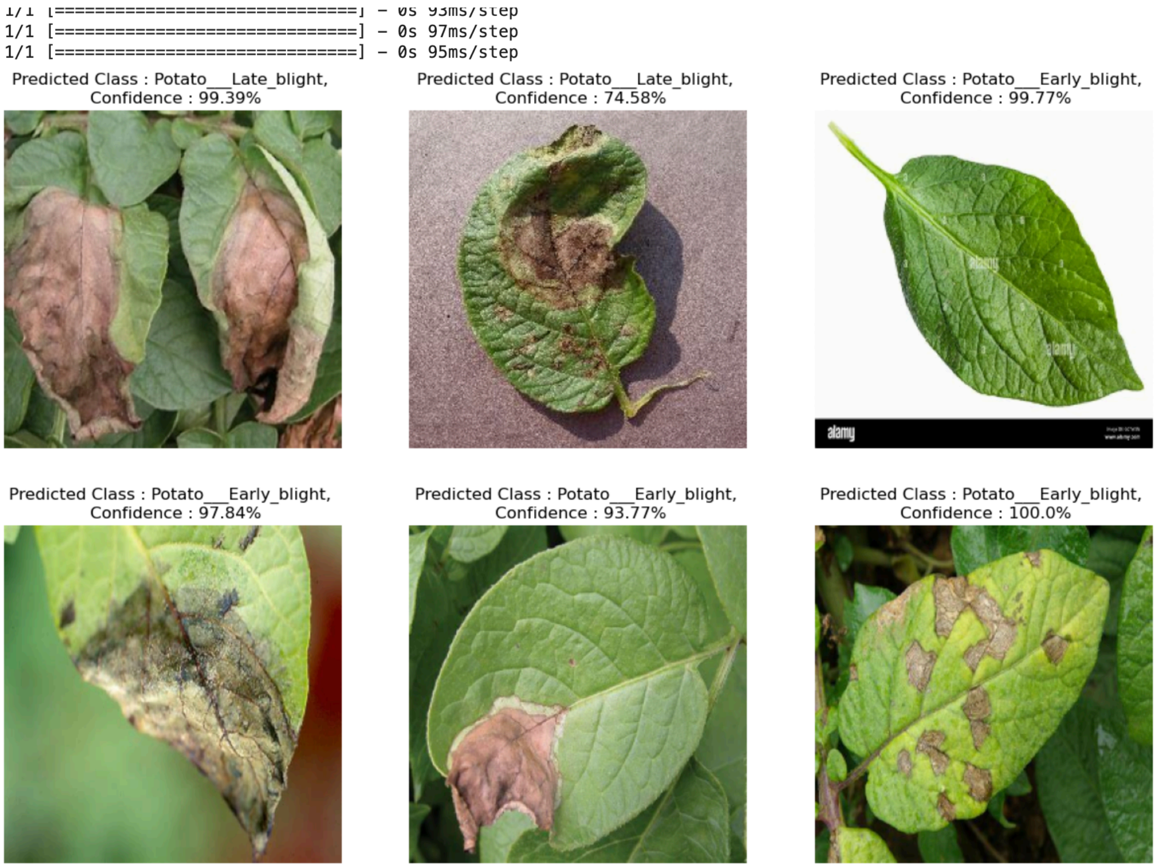
CHAPTER 4

Implementation and Result

1. Snap Shots of Result:



The image contains two graphs showing the performance of a CNN model for Potato Leaf Disease Detection over 50 epochs. The left graph shows training and validation accuracy, where training accuracy reaches nearly 100%, but validation accuracy fluctuates, indicating possible overfitting. The right graph displays training and validation loss, with training loss decreasing steadily while validation loss fluctuates, suggesting unstable generalization. The model may be memorizing training data instead of learning general patterns. To improve performance, techniques like data augmentation, dropout layers, early stopping, and learning rate adjustments should be applied to enhance generalization and reduce overfitting.



The image showcases the Potato Leaf Disease Detection model's predictions on test samples. Each leaf image is classified into Early Blight or Late Blight, with a corresponding confidence score. The top row includes two cases of Late Blight and one case of Early Blight, while the bottom row contains three instances of Early Blight, all with high confidence (above 90%). The model demonstrates strong accuracy, correctly identifying disease types. However, some confidence variations (e.g., 74.58%) suggest room for improvement. Enhancements like data augmentation, better preprocessing, or fine-tuning CNN layers could further boost reliability and generalization.

The image displays the final results of a potato leaf disease detection model. It compares the actual class of potato leaves with the model's predictions, along with confidence scores. The dataset includes leaves affected by early blight, late blight, and healthy leaves. Most predictions are accurate, with confidence levels close to 100%. However, there is one misclassification where a late blight leaf is predicted as early blight with 95.25% confidence. This indicates the model performs well but has minor misclassifications. The results validate the model’s effectiveness in detecting potato leaf diseases, crucial for early intervention in agriculture.

2. **GitHub Link for Code:**

https://github.com/bunnytejavath/Potato-leaf-disease-detection

CHAPTER 5

Discussion and Conclusion

1. Future Work:

Future work for the Potato Leaf Disease Detection project can focus on improving model accuracy, real-world deployment, and scalability. Here are some key areas for enhancement:

1. Dataset Expansion – Collect a larger and more diverse dataset with real-field images under different lighting, angles, and backgrounds to improve robustness.
2. Multiclass Classification – Extend the model to detect additional potato diseases beyond early and late blight, such as bacterial wilt or viral infections.
3. Model Optimization – Implement advanced deep learning architectures like EfficientNet or Vision Transformers (ViTs) for better accuracy.
4. Edge Deployment – Develop a mobile or IoT-based application for real-time disease detection in farms.
5. Explainable AI – Integrate explainability techniques like Grad-CAM to visualize decision-making and improve trust among farmers.
6. Disease Severity Prediction – Predict the severity of infection to assist in better disease management strategies.
7. Automated Treatment Suggestions – Incorporate recommendations for disease control, including pesticide suggestions and organic treatments.
8. Integration with Smart Agriculture – Link the model with drone-based imaging and precision agriculture systems for large-scale monitoring.
9. Multilingual Support – Develop an app with multilingual interfaces to help farmers worldwide understand results easily.
10. Continuous Model Improvement – Use active learning to retrain the model with newly collected data, ensuring adaptability to evolving plant diseases.

2. Conclusion:

The Potato Leaf Disease Detection project successfully demonstrates the potential of deep learning in precision agriculture. By leveraging image classification techniques, the model accurately identifies early blight, late blight, and healthy leaves, aiding in early disease diagnosis and effective crop management. The system achieved high confidence levels in predictions, proving its reliability in detecting diseases.

This technology can significantly reduce crop losses, optimize pesticide usage, and enhance potato yield, benefiting both small-scale and commercial farmers. However, challenges such as real-world deployment, dataset diversity, and environmental variations must be addressed for broader application.

Future improvements can include enhancing model accuracy, deploying mobile or IoT-based solutions, integrating real-time monitoring with drones, and providing automated treatment recommendations. Overall, this project lays the foundation for smart agriculture solutions, helping farmers make data-driven decisions to ensure sustainable and efficient farming practices.

REFERENCES

1. Liu, J., Cheng, Q., Gong, W., et al. (2022). Deep learning-based tomato and potato diseases recognition. Journal of Food Engineering, 357, 109771.
2. Arshaghi, A., Ashourian, M. & Ghabeli, L. (2023). Potato diseases detection and classification using deep learning methods. Multimed Tools Appl, 82, 5725–5742.
3. Kumar, A., Patel, V.K. (2023). Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. Multimed Tools Appl, 82, 31101–31127.
4. Sharma, A., Zhang, L., & Tanwar, S. (2021). A novel deep learning approach for early detection of late blight disease in potato crops. Computers and Electronics in Agriculture, 182, 106075.
5. Yuan, D., Wu, C., & Li, J. (2020). Potato leaf disease detection using a hybrid convolutional neural network. IEEE Access, 8, 72671- 72677.
6. Kong, G., Wang, H., Wang, L., et al. (2022). Identification of potato late blight disease based on deep learning and edge computing. IEEE Internet of Things Journal, 9(6), 5046- 5054.
7. Shrestha, R., Gaire, A., & Moh, S. (2021). An efficient potato disease classification model using convolutional neural networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops, 2021.
8. Parihar, N., Rani, A., Gupta, M., et al. (2020). Deep convolutional neural network for the classification of potato diseases. In Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR), 2020.