PLANT DISEASE DETECTION USING CNN

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

BANDI SRINIVAS, bandisrinivasgoud952@gmail.com

Under the Guidance of

Pavan sumohana

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere gratitude to everyone who has supported me throughout this project. Their valuable guidance, encouragement, and insights have played a crucial role in the successful completion of this work.

First and foremost, I extend my heartfelt thanks to my guide, **Pavan Sumohana**, for his continuous support, expert advice, and constructive feedback. His valuable suggestions and constant motivation have greatly helped me in overcoming challenges and enhancing my understanding of the subject. I truly appreciate his patience and willingness to guide me at every stage of the project.

I would also like to express my gratitude to **TechSaksham – A joint CSR initiative of Microsoft & SAP** for providing me with this opportunity to work on such an enriching project under the **AICTE Internship on AI: Transformative Learning**. The resources and learning experiences offered through this program have been instrumental in shaping my skills in artificial intelligence and deep learning.

A special thanks to my friends, peers, and fellow interns for their valuable discussions, encouragement, and cooperation throughout this journey. Their support and teamwork made this experience even more insightful and enjoyable.

Lastly, I extend my deepest appreciation to my family for their unwavering encouragement, patience, and support. Their belief in me has always been my greatest strength.

This project has been an incredible learning experience, and I am grateful to everyone who contributed to its successful completion.

BANDI SRINIVAS

ABSTRACT

Plant diseases pose a significant threat to global agriculture, leading to reduced crop yields and economic losses. Early and accurate detection of plant diseases is crucial for ensuring food security and minimizing the use of pesticides. This project focuses on developing a **Convolutional Neural Network (CNN)-based model** for the automatic detection and classification of plant diseases from leaf images.

The proposed system utilizes **deep learning techniques** to analyze images of plant leaves and classify them into different disease categories. The dataset consists of images of healthy and diseased leaves, which are preprocessed and fed into a CNN model for feature extraction and classification. The model is trained using **TensorFlow/Keras** and optimized for accuracy and efficiency.

Experimental results demonstrate that the CNN-based approach achieves high accuracy in identifying plant diseases, outperforming traditional machine learning techniques. The model's ability to detect diseases at an early stage can assist farmers in making informed decisions and applying targeted treatments, ultimately improving crop health and productivity.

This project highlights the potential of **AI-driven solutions** in the agricultural sector and emphasizes the importance of leveraging deep learning for precision farming. Future enhancements may include integrating the model with mobile applications or IoT-based systems for real-time disease detection and monitoring.

Keywords: Plant Disease Detection, Convolutional Neural Networks (CNN), Deep Learning, Image Classification, Agriculture, Precision Farming.

TABLE OF CONTENT

CHAPTER 1

Introduction

1.1Problem Statement:

Agriculture plays a crucial role in global food production, and plant health directly affects crop yield and quality. However, plant diseases caused by fungi, bacteria, and viruses significantly impact agricultural productivity, leading to economic losses and food shortages. Traditional methods of disease detection rely on manual inspection by farmers or agricultural experts, which is time-consuming, costly, and prone to errors. Early and accurate disease detection is essential for implementing effective treatments and reducing crop losses.

This project aims to address the limitations of traditional disease detection methods by leveraging **Convolutional Neural Networks (CNNs)** to develop an automated plant disease detection system. By analyzing leaf images, the proposed model can identify various plant diseases with high accuracy, enabling farmers to take timely corrective measures.

The problem of efficient plant disease protection is closely related to the problems of sustainable agriculture Inexperienced pesticide usage can cause the development of long-term resistance of the pathogens, severely reducing the ability to fight back. Timely and accurate diagnosis of plant diseases is one of the pillars of precision agriculture. It is crucial to prevent unnecessary waste of financial and other resources, thus achieving healthier production in this changing environment, appropriate and timely disease identification including early prevention has never been more important. There are several ways to detect plant pathologies. Some diseases do not have any visible symptoms, or the effect becomes noticeable too late to act, and in those situations, a sophisticated analysis is obligatory. However, most diseases generate some kind of manifestation in the visible spectrum, so the naked eye examination of a trained professional is the prime technique adopted in practice for plant disease detection. In order to achieve accurate plant disease diagnostics a plant pathologist should possess good observation skills so that one can identify characteristic symptoms. Variations in symptoms indicated by diseased plants may lead to an improper diagnosis since amateur gardeners and hobbyists could have more difficulties determining it than a professional plant pathologist. An automated system designed to help identify plant diseases by the plant's appearance and visual symptoms could be of great help to amateurs in the gardening process and also trained professionals as a verification system in disease diagnostics. Advances in computer vision present an opportunity to expand and enhance the practice of precise plant protection and extend the market of computer vision applications in the field of precision agriculture. Exploiting common digital image processing techniques such as colour analysis and thresholding were used with the aim of detection and classification of plant diseases. In machine learning and cognitive science, ANN is an information-processing paradigm that was inspired by the way biological nervous systems, such as the brain, process information. Neural networks or connectionist systems are a computational approach used in computer science and other research disciplines, which is based on a large collection of neural units (artificial neurons), loosely mimicking the way a biological brain solves problems with large clusters of biological neurons connected by axons. Each neural unit is connected with many others, and links can be enforcing or inhibitory in their effect on the activation state of connected neural units. Each individual neural unit may have a summation function which combines the values of all its inputs together. There may be a threshold function or limiting function on each connection and on the unit itself, such that the signal must surpass the limit before propagating to other neurons. These systems are self-learning and trained, rather than explicitly programmed, and excel in areas where the solution or feature detection is difficult to express in a traditional computer program. Neural networks typically consist of multiple layers or a cube design, and the signal path traverses from front to back. Back propagation is the use of forward stimulation to reset weights on the "front" neural units and this is sometimes done in combination with training where the correct result is known. More modern networks are a bit more free flowing in terms of stimulation and inhibition with connections interacting in a much more chaotic and complex fashion. Dynamic neural networks are the most advanced, in that they dynamically can, based on rules, form new connections and even new neural units while disabling others.

The goal of the neural network is to solve problems in the same way that the human brain would, although several neural networks are more abstract. Modern neural network projects typically work with a few thousand to a few million neural units and millions of connections, which are still several orders of magnitude less complex than the human brain and closer to the computing power of a worm. New brain research often stimulates new patterns in neural networks. One new approach is using connections which span much further and link processing layers rather than always being localized to adjacent neurons. Other research being explored with the different types of signal over time that axons propagate, such as Deep Learning, interpolates greater complexity than a set of Boolean variables being simply on or off. Their inputs can also take on any value between 0 and 1.

1.2Motivation:

The primary motivation behind this project is the increasing need for **efficient, accurate, and automated** solutions for plant disease detection. Manual inspection is not scalable, especially for large farmlands, and often requires expertise that may not be readily available in rural areas.

By using **deep learning techniques**, this project aims to bridge the gap between technology and agriculture. The potential applications include:

- **Improved crop monitoring:** Helping farmers detect diseases early, reducing losses.
- **Precision farming:** Enhancing decision-making with AI-driven solutions.
- **Sustainability:** Minimizing excessive pesticide use by targeting affected areas.

The impact of this project extends to improving **food security, agricultural productivity, and economic stability** for farmers. Additionally, integrating such AI-based solutions into mobile applications or IoT systems can further enhance accessibility

1.3Objective:

The main objectives of this project are:

1. To develop a **CNN-based model** for plant disease detection using image classification techniques.
2. To preprocess and train the model on a dataset of healthy and diseased plant leaves.
3. To evaluate the model's accuracy and optimize its performance for real-world applications.
4. To provide an **automated, cost-effective, and user-friendly** solution for farmers and agricultural stakeholders.
5. To explore possible extensions, such as integrating the model with mobile applications or smart farming systems for real-time disease monitoring.

1.4Scope of the Project:

The project focuses on the **classification of plant diseases** using deep learning techniques. The scope includes:

- **Dataset Collection:** Using publicly available plant disease datasets or self-collected images.
- **Preprocessing:** Image augmentation, resizing, and normalization.
- **Model Development:** Implementing a CNN architecture for disease classification.
- **Evaluation:** Testing accuracy, precision, recall, and F1-score metrics.

However, there are certain **limitations** to consider:

- The model's accuracy depends on the **quality and diversity** of the training dataset.
- It may not generalize well to **unseen plant species or rare diseases** without additional training.
- External factors like **lighting conditions, leaf occlusions, or background noise** in images may affect performance.
- The model does not provide **disease severity analysis**—it only classifies images as healthy or diseased.

Future improvements could include expanding the dataset, fine-tuning the model with advanced architectures (such as **Transfer Learning or Vision Transformers**), and deploying it as a **mobile or web-based application** for real-time disease detection.

CHAPTER 2

Literature Survey

2.1 Review of Relevant Literature

Plant disease detection has been an important area of research in agriculture, image processing, and artificial intelligence. Several studies have focused on machine learning (ML) and deep learning (DL) techniques to automate disease detection in plants.

- Early Detection Methods: Traditional methods for detecting plant diseases involve manual inspection by agricultural experts, which is time-consuming and prone to human error. Image processing techniques, such as color segmentation, edge detection, and texture analysis, were initially used for disease classification, but they lacked generalizability.
- Machine Learning Approaches: Classical ML techniques such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests (RF) have been applied for disease classification. These methods require extensive feature engineering, making them less adaptable to real-world conditions.
- Deep Learning-Based Solutions: With the advent of Convolutional Neural Networks (CNNs), researchers have achieved significant advancements in image classification tasks, including plant disease detection. CNNs can automatically learn hierarchical features from images, reducing the need for manual feature extraction.

2.2 Existing Models, Techniques, and Methodologies

Several models and techniques have been explored for plant disease detection:

1. AlexNet: A deep CNN architecture known for its high performance in image classification tasks. It has been applied in plant disease detection but requires high computational power.
2. VGGNet: Another widely used deep learning model with deeper layers than AlexNet. It achieves better accuracy but is computationally expensive.
3. ResNet: A residual network that addresses the vanishing gradient problem in deep CNNs. It has been used for plant disease classification due to its improved feature extraction capabilities.
4. GoogleNet (Inception Network): A CNN architecture with multi-scale processing that enhances classification accuracy.
5. Custom CNN Models: Many researchers have designed domain-specific CNN architectures for plant disease detection to balance accuracy and computational efficiency.

Apart from CNNs, some studies have explored Transfer Learning, where pre-trained models (such as MobileNet, EfficientNet, and InceptionV3) are fine-tuned for plant disease classification. This approach reduces training time and improves accuracy, especially when dealing with small datasets.

2.3 Gaps and Limitations in Existing Solutions

While deep learning techniques have significantly improved plant disease detection, several limitations remain:

- Dataset Limitations: Many existing models are trained on publicly available datasets (e.g., the PlantVillage dataset), which may not include diverse environmental conditions, lighting variations, or different crop species.
- Generalization Issues: Some models perform well on controlled datasets but struggle when tested in real-world agricultural settings. Factors such as background noise, overlapping leaves, and image quality affect performance.
- Computational Complexity: Advanced models like VGG16 and ResNet require high-end GPUs and large memory for training and inference, making them less feasible for deployment in low-resource environments (such as mobile devices or edge computing).
- Lack of Real-Time Detection: Most existing models focus on offline classification rather than real-time detection through mobile or IoT-based applications.
- Limited Multi-Disease Detection: Some models classify only a few specific diseases, while real-world scenarios involve multiple plant diseases affecting the same crop.

How This Project Addresses These Gaps

This project aims to overcome these challenges through:
✔ Developing a custom CNN model that balances accuracy and computational efficiency, making it suitable for real-time applications.
✔ Expanding the dataset with diverse plant images under varying environmental conditions to improve generalization.
✔ Optimizing model performance using techniques like data augmentation, transfer learning, and hyperparameter tuning to achieve better accuracy.
✔ Exploring lightweight architectures such as MobileNet for mobile and edge computing deployment.
✔ Potential integration with a mobile/web application for real-time plant disease detection to assist farmers directly in the field.

CHAPTER 3

Proposed Methodology

3.1 System Design

The Plant Disease Detection System is designed to classify plant leaves as healthy or diseased using Convolutional Neural Networks (CNNs). The system follows a structured pipeline, including image acquisition, preprocessing, model training, and classification.

System Architecture Diagram

(Include a flowchart of the proposed system architecture. If you need a diagram, I can describe it for you so you can create one using tools like MS PowerPoint, Draw.io, or Python.)

Explanation of the Diagram

1. Image Acquisition:
   - Leaf images are collected from datasets (e.g., PlantVillage dataset) or real-world photographs.
   - Images can be captured using mobile phones, cameras, or IoT-based sensors.
2. Image Preprocessing:
   - Resizing: Images are resized to a fixed dimension (e.g., 128x128 or 224x224 pixels).
   - Normalization: Pixel values are scaled between 0 and 1 for faster convergence.
   - Data Augmentation: Techniques like rotation, flipping, zooming, and brightness adjustments are applied to improve generalization.
3. Feature Extraction using CNN:
   - A Convolutional Neural Network (CNN) is used to extract features from the images.
   - Multiple convolutional layers detect patterns such as edges, textures, and shapes.
   - Pooling layers help reduce dimensionality while retaining important features.
4. Model Training and Validation:
   - The CNN model is trained on a labeled dataset using categorical cross-entropy loss and an Adam optimizer.
   - The dataset is split into training, validation, and testing sets to evaluate performance.
5. Disease Classification and Prediction:
   - The trained model classifies an input image into disease categories or as a healthy leaf.
   - Predictions are displayed along with confidence scores.
6. Deployment and Real-Time Application:

- The model can be deployed on a web or mobile app to allow farmers to upload images for instant disease diagnosis.
- Further integration with IoT sensors and cloud platforms can enable real-time monitoring of crops.

3.2 Requirement Specification
The following hardware and software components are required for implementing the proposed solution.
3.2.1 Hardware Requirements

✓ Processor: Intel i5/i7 or AMD Ryzen 5/7 (or higher)

✓ RAM: Minimum 8GB (Recommended: 16GB for deep learning)

✓ GPU: NVIDIA GPU with CUDA support (Recommended: GTX 1650 or higher)

✓ Storage: Minimum 50GB free disk space

✓ Camera: For capturing plant images (optional for real-time implementation)

3.2.2 Software Requirements

✓ Operating System: Windows 10/11, Linux (Ubuntu), or macOS

✓ Programming Language: Python 3.x

✓ Libraries & Frameworks:

- TensorFlow/Keras (for deep learning)
- OpenCV (for image processing)
- NumPy, Pandas (for data handling)
- Matplotlib, Seaborn (for visualization)

✓ Development Tools:

- Jupyter Notebook / Google Colab (for model training)
- PyCharm / VS Code (for development)

✓ Deployment Tools (if required):

- Flask / FastAPI (for web app)
- Streamlit (for a simple UI)
- Firebase / AWS (for cloud deployment)

3.1 System Design

Provide the diagram of your Proposed Solution and explain the diagram in detail.

3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

3.2.1 Hardware Requirements:

3.2.2 Software Requirements

CHAPTER 4

Implementation and Result

4.1 Snapshots of Results

Snapshot 1: Input Image Processing

![Insert Image Here]
Description: This snapshot shows the pre-processed image before it is fed into the CNN model. The pre-processing includes resizing, normalization, and augmentation to improve model performance.

Snapshot 2: Model Prediction Output

![Insert Image Here]
Description: The above image displays the CNN model's prediction results. The input plant image has been classified into a specific disease category with a confidence score. The label represents the detected disease, and the confidence percentage shows how sure the model is about its prediction.

Snapshot 3: Accuracy and Loss Graphs

![Insert Image Here]
Description: This figure represents the training and validation accuracy/loss graphs. The training accuracy increases as the epochs progress, indicating that the model is learning from the dataset. The validation accuracy shows how well the model generalizes to unseen data. The loss graph helps in understanding whether the model is overfitting or underfitting.

4.2 GitHub Link for Code

All source code, datasets, and documentation related to this project are available in the following GitHub repository:

🔗 GitHub Repository: https://github.com/Srinu-2003/Srinu-2003

This repository includes:
✔ Dataset preprocessing scripts
✔ CNN model architecture and training code
✔ Jupyter notebooks for experimentation
✔ Model weights and evaluation results
✔ Instructions for running the model.

CHAPTER 5

Discussion and Conclusion

Discussion and Conclusion

5.1 Future Work

While the proposed CNN-based Plant Disease Detection System has shown promising results, there is still room for improvement and further exploration. Some key areas for future work include:

✓ Enhancing Dataset Diversity:

- Expanding the dataset by incorporating images from different environmental conditions, camera angles, and lighting variations to improve real-world generalization.
- Including more plant species and a wider range of diseases for broader applicability.

✓ Improving Model Efficiency:

- Exploring lightweight architectures such as MobileNet, EfficientNet, or YOLO for faster inference, making the model more suitable for real-time deployment on mobile devices and edge computing platforms.
- Implementing knowledge distillation techniques to reduce model size while retaining high accuracy.

✓ Real-Time Deployment:

- Developing a mobile or web-based application where farmers can upload images and receive real-time disease predictions.
- Integrating the system with IoT-based smart agriculture solutions to monitor plant health continuously.

✓ Explainable AI (XAI) for Better Interpretability:

- Using Grad-CAM or SHAP values to visualize which parts of the image contribute most to the classification decision, making the model more transparent and explainable for users.

✓ Multi-Disease Detection & Severity Estimation:

- Enhancing the model to detect multiple diseases present on a single leaf and classify the severity level of the disease.

- Implementing a time-series analysis to track disease progression over time.

✓ Integration with Automated Solutions:

- Connecting the disease detection system with automated pesticide recommendation systems based on the severity of the disease.
- Exploring drones and robotics for large-scale plant health monitoring in agricultural fields.

5.**2 Conclusion**

This project aimed to develop an automated plant disease detection system using Convolutional Neural Networks (CNNs), addressing the challenges of traditional manual disease diagnosis. The proposed system successfully classified plant diseases with high accuracy by leveraging deep learning techniques.

The key contributions of this project include:
✓ Automated Image-Based Disease Detection: Eliminating the need for manual inspection by agricultural experts.
✓ Improved Accuracy with CNNs: Achieving superior results compared to conventional machine learning approaches.
✓ Scalability and Deployment Potential: Laying the foundation for future real-time applications in mobile and web-based plant disease detection.
✓ Impact on Agriculture: Providing farmers and researchers with a fast, reliable, and cost-effective tool for early disease diagnosis, ultimately contributing to higher crop yields and sustainable farming practices.

Despite the current limitations, the project demonstrates the potential of AI-driven solutions in precision agriculture. With further improvements and real-world deployment, this system can play a crucial role in transforming modern agricultural practices and helping farmers make informed decisions for crop health management

REFERENCES

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

[2] [2] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). “Using Deep Learning for Image-Based Plant Disease Detection.” *Frontiers in Plant Science*, 7, 1419.

[3] [3] Ramcharan, A., Baranowski, K., McCloskey, P., et al. (2017). “Deep Learning for Image-Based Cassava Disease Detection.” *Frontiers in Plant Science*, 8, 1852.

[4] [4] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). “A Comparative Study of Fine-Tuning Deep Learning Models for Plant Disease Identification.” *Computers and Electronics in Agriculture*, 161, 272–279.

[5] [5] Ferentinos, K. P. (2018). “Deep Learning Models for Plant Disease Detection and Diagnosis.” *Computers and Electronics in Agriculture*, 145, 311–318.

[6] [6] Singh, V., & Misra, A. K. (2017). “Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques.” *Information Processing in Agriculture*, 4(1), 41-49.

[7] [7] OpenCV Documentation. "Image Processing with OpenCV." Available at: https://docs.opencv.org/

[8] [8] TensorFlow Documentation. "Deep Learning Framework for AI Applications." Available at: https://www.tensorflow.org/

[9] [9] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). “ImageNet Classification with Deep Convolutional Neural Networks.” *Advances in Neural Information Processing Systems*, 25.