Plant Disease Detection

Report

Submitted in partial fulfillment of the requirements of the degree of

Bachelor of Engineering

(Information Technology)

By

Nikita Chhabria (07)

Dimple Dalwani (08)

Nayaab Jindani (20)

Under the guidance of

Dr. Ravita Mishra

Department of Information Technology

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY,

Chembur, Mumbai 400074

(An Autonomous Institute, Affiliated to University of Mumbai) April 2024

Vivekanand Education Society's
Institute of Technology
(Autonomous Institute Affiliated to University of Mumbai, Approved by AICTE & Recognised by Govt. of Maharashtra)
NAAC accredited with 'A' grade

Certificate

This is to certify that project entitled

"Plant Disease Detection"

Group Members Names

Nikita Chhabria (Roll No. 07)
Dimple Dalwani (Roll No. 08)
Nayaab Jindani (Roll No. 20)

In fulfillment of degree of BE. (Sem. VI) in Information Technology for Project is approved.

Dr. Ravita Mishra
Project Mentor

External Examiner

Dr.(Mrs.)Shalu Chopra
H.O.D

Dr.(Mrs.) J.M.Nair
Principal

Date: / /2025
Place: VESIT, Chembur

College Seal

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Nikita Chhabria (Rollno. 7) **(Signature)** - - - - - - - - - - -

Dimple Dalwani (Rollno. 8) **(Signature)** - - - - - - - - - - -

Nayaab Jindani (Rollno. 20) **(Signature)** - - - - - - - - - - -

Abstract

In the face of growing agricultural challenges, the early detection and classification of plant diseases play a vital role in safeguarding crop health and ensuring sustainable food production. This project introduces a computer vision-based plant disease detection system capable of classifying plant leaf images into three categories: healthy, rusty, and powdery. The system is designed to assist farmers and agricultural experts by providing rapid and reliable disease identification through image uploads. Three state-of-the-art models are implemented and compared in this project: a traditional Convolutional Neural Network (CNN), EfficientNet, and YOLOv8. Each model leverages different architectures to detect and classify diseases with varying focuses on accuracy, speed, and generalization. The dataset undergoes comprehensive preprocessing steps such as image resizing, normalization, and augmentation to ensure robustness and improved model performance. Among the approaches, YOLOv8 demonstrates real-time detection capabilities with high precision, while EfficientNet balances accuracy and efficiency through its scalable architecture. Traditional CNN serves as a strong baseline for evaluating deep learning effectiveness in this context. Evaluation metrics including accuracy, precision, recall, and F1-score are used to assess model performance. The findings affirm that integrating deep learning with agricultural diagnostics significantly enhances disease detection and management. This project not only showcases the technical viability of automated plant disease classification but also emphasizes its potential impact in supporting precision agriculture and promoting global food security.

Contents

1 Introduction **7**
1.1 Introduction 7
1.2 Objectives 7
1.3 Motivation 7
1.4 Scope of the Work 7
1.5 Feasibility Study 7

2 Literature Survey **8**
2.1 Introduction 8
2.2 Problem Definition 8
2.3 Review of Literature Survey 8

3 Design and Implementation **9**
3.1 Introduction 9
3.2 Requirement Gathering 9
3.3 Proposed Design 9
3.4 Proposed Algorithm 9
3.5 Architectural Diagrams 10

4 Results and Discussion
4.1 Introduction.......... 10
4.2 Cost Estimation.......... 10
4.3 Feasibility Study.......... 10
4.4 Results of Implementation.......... 10
4.5 Result Analysis.......... 10
4.6 Observation/Remarks.......... 10

5 Conclusion **12**
5.1 Conclusion.......... 12
5.2 Future Scope.......... 12
5.3 Societal Impact.......... 12

ACKNOWLEDGEMENT

The project report on “Plant Disease Detection” is the outcome of the guidance, moral support and devotion bestowed on our group throughout our work. For this we acknowledge and express our profound sense of gratitude to everybody who has been the source of inspiration throughout project preparation. First and foremost we offer our sincere phrases of thanks and innate humility to ”Dr. Shalu Chopra, HOD”, ”Dr. Manoj Sabnis,Deputy HOD”, ”Dr. Ravita Mishra” for providing the valuable inputs and the consistent guidance and support provided by them. We can say in words that we must at outset tender our intimacy for receipt of affectionate care to Vivekanand Education Society’s Institute of Technology for providing such a stimulating atmosphere and conducive work environment.

Chapter 1: Introduction

1.1. Introduction

In modern agriculture, ensuring plant health is vital for food security, sustainable farming, and economic stability. However, plant diseases often go undetected in early stages, leading to reduced crop yield and increased costs. This project aims to leverage machine learning techniques for early and accurate detection of plant diseases using images. By automating disease identification, farmers can take timely action, improving productivity and reducing losses.

1.2. Objective

- Develop a deep learning model that can classify plant diseases from leaf images with high accuracy.
- Train the model using the Kaggle Plant Disease Recognition Dataset and optimize it for better performance.
- Improve prediction accuracy through techniques such as data augmentation, transfer learning, and hyperparameter tuning.
- To evaluate the model using metrics such as accuracy, precision, recall, and F1-score.

1.3. Motivation

Plant diseases can spread rapidly and are often hard to identify without expert knowledge. Traditional manual inspection is time-consuming and subjective. With the rise in smartphone usage and IoT devices in agriculture, machine learning presents a scalable and objective solution for real-time disease detection, empowering farmers with timely insights.

1.4. Scope of the Work

- Focuses on supervised image classification for plant disease detection using deep learning techniques.
- Utilizes a labeled dataset of leaf images categorized into Healthy, Rust, and Powdery conditions.
- Implements and compares three models: Traditional CNN, EfficientNet, and YOLOv8 for performance benchmarking.
- Analyzes model effectiveness using classification reports, confusion matrices, and visualization of predictions.

1.5. Feasibility Study

- **Technical Feasibility**:
  Employs Python with frameworks such as TensorFlow, Keras, and OpenCV, along with YOLOv8 for object detection — all of which are well-supported and widely used in the machine learning ecosystem.
- **Operational Feasibility**:
  The system can be deployed as a web or mobile application allowing farmers to upload leaf images and get real-time predictions, making it accessible and practical in agricultural settings.
- **Economic Feasibility**:
  The solution is cost-effective as it relies on open-source tools and publicly available datasets, making it scalable even for low-resource environments.

Chapter 2: Literature Survey

2.1. Introduction

Plant disease detection has made significant strides with the advancements in machine learning and computer vision, addressing challenges in agriculture for early identification and efficient management of plant diseases. Traditional methods of detection, often dependent on visual inspections by experts, are now being enhanced or replaced by data-driven techniques. This literature survey presents a comparative study of three major approaches in plant disease detection: traditional Convolutional Neural Networks (CNN), EfficientNet, and YOLOv8, all of which utilize deep learning methods to classify plant diseases based on leaf images.

2.2. Problem Definition

The objective of this literature review is to explore the various methodologies employed in plant disease detection, focusing on their application to classify and diagnose plant diseases using deep learning techniques.

2.3. Review of Literature Survey

In the paper *"Plant Disease Detection Using Deep Convolutional Neural Network"* by J. Arun Pandian et al., published in *Applied Sciences* (2022), the authors present a 14-layer Deep Convolutional Neural Network (14-DCNN) for classifying plant leaf diseases. They used a large dataset of 147,500 images spanning 58 disease and healthy classes, including a 'no-leaf' category. To improve model performance and handle class imbalance, they applied data augmentation techniques such as Basic Image Manipulation (BIM), DCGAN, and Neural Style Transfer (NST). The model showed strong classification accuracy, evaluated through AUC-ROC curves and occlusion analysis, but its performance gains over existing methods were moderate. A noted limitation was its reliance on face-up leaf images, which could affect performance in varied real-world conditions.

In the paper *"Enhanced YOLOv8 Algorithm for Leaf Disease Detection with Lightweight GOCR-ELAN Module and Loss Function: WSIoU"* by Guihao Wen et al., published in *Computers in Biology and Medicine* (2025), the authors propose an optimized version of the YOLOv8 algorithm for improved leaf disease detection. The enhanced model incorporates a lightweight GOCR-ELAN module to reduce the number of parameters while preserving essential feature representations. Additionally, the standard convolution layers are replaced with a new ADown module to improve feature extraction in complex and occluded environments. A novel loss function, WSIoU, is also introduced to boost localization accuracy and reduce false positives. Evaluated on a Roboflow dataset with 5,494 images across 12 disease categories, the model achieves a 28.7% reduction in parameters and a 43.2% drop in computational cost, while improving MAP50 to 87.7%. With a final model size of just 4.55 MB, the proposed architecture outperforms the baseline YOLOv8 in both efficiency and accuracy, making it well-suited for real-time deployment in resource-constrained agricultural settings.

Chapter 3: Design and Implementation

3.1. Introduction

This chapter outlines the design and implementation of the Plant Disease Detection System, developed using a structured deep learning workflow. The process involves dataset preparation, image preprocessing, model selection, training, and evaluation. Three key models—Convolutional Neural Networks (CNN), EfficientNet, and YOLOv8—are utilized to classify plant leaf diseases based on visual symptoms.

3.2. Requirement Gathering

Hardware Requirements:

- System with at least 4GB RAM
- Stable internet connectivity

Software Requirements:

- Python 3.x
- Python Libraries:
    - TensorFlow / Keras – for building and training CNN and EfficientNet models
    - Ultralytics – for implementing and running YOLOv8
    - OpenCV – for image preprocessing and augmentation
    - Matplotlib and Seaborn – for visualizing data and model performance
    - NumPy and Pandas – for data handling and manipulation

3.3. Proposed Design

The system design follows a structured and repeatable workflow for developing the plant disease detection system:

- Data Collection:
  A large dataset of plant leaf images was collected, containing both healthy and diseased leaves across various plant species(Kaggle dataset was used). The dataset includes labels indicating specific disease types.

- Data Cleaning and Preprocessing:
  Images were resized, normalized, and enhanced using data augmentation techniques such as rotation, flipping, and brightness adjustments to improve model generalization and address class imbalance.

- Feature Engineering:
  For the CNN and EfficientNet models, features were learned automatically through the convolution layers. For YOLOv8, labeled bounding boxes were used to assist the model in identifying and classifying diseases in real time.

- Model Development:
  Three different models were developed for comparison:

    - Convolutional Neural Network (CNN)
    - EfficientNet
    - YOLOv8

- **Model Evaluation:**
  Performance was evaluated using metrics such as Accuracy, Precision, Recall, F1-Score, and confusion matrices.

- **Visualization:**
  Model results were visualized using confusion matrices to assess the accuracy of the disease classification and highlight any misclassifications.

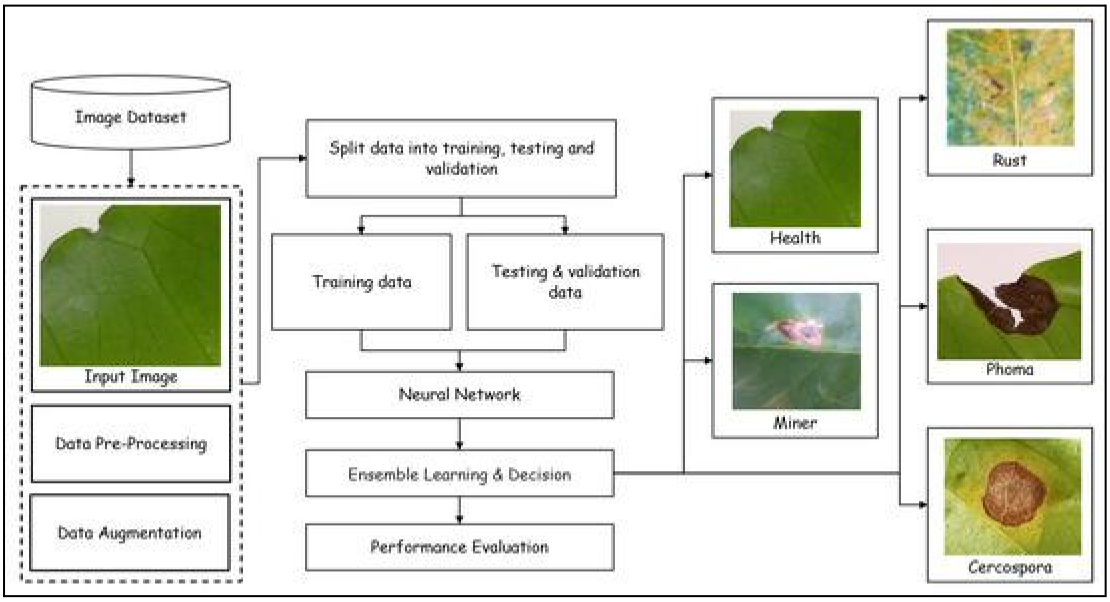
3.4. Proposed Algorithm

Convolutional Neural Network (CNN)
The CNN model plays a critical role in the plant disease detection system by learning hierarchical feature representations from leaf images. The network is designed to automatically detect patterns such as color, texture, and shape, which are indicative of various diseases. The architecture consists of multiple convolutional layers followed by pooling layers to reduce dimensionality and capture important features. During the training process, the model learns to classify images of both healthy and diseased leaves by minimizing the error in prediction. This approach allows for the detection of subtle disease symptoms in complex image data, making it highly effective for plant disease classification.

EfficientNet
EfficientNet is a state-of-the-art deep learning model known for its scalability and efficiency. By using a compound scaling method, EfficientNet improves accuracy while reducing computational requirements compared to traditional CNN architectures. In this project, EfficientNet was used to classify leaf diseases, taking advantage of its ability to achieve higher accuracy with fewer parameters. The model learns to capture intricate features in the leaf images, and it is trained on a variety of augmented data to handle different lighting conditions, rotations, and orientations of the leaves. EfficientNet demonstrated excellent performance in distinguishing between healthy and diseased plant leaves.

YOLOv8 (You Only Look Once)
YOLOv8 is an advanced object detection model that identifies and classifies diseases in real-time by drawing bounding boxes around the affected areas on the leaf images. It operates by detecting the disease symptoms directly on the plant's leaves, offering fast and accurate predictions, which is essential for real-time applications. YOLOv8 is particularly effective in localizing the disease regions, even in complex and cluttered images. The model was fine-tuned for the specific task of plant disease detection, enabling high-speed inference with minimal computational resources, making it suitable for on-site agricultural use.

3.5. Architectural Diagrams



Chapter 4: Results and Discussion

4.1. Introduction

To assess the effectiveness of the developed plant disease detection system, several deep learning models were trained and evaluated using key metrics such as Accuracy, Precision, Recall, F1-Score.These evaluation criteria offered a comprehensive view of each model's ability to classify healthy and diseased leaves accurately while reducing misclassifications.

4.2. Cost Estimation

The entire project was implemented and executed locally using the Jupyter Notebook format within Visual Studio Code (VS Code). This approach required no paid cloud services, helping to keep the development cost minimal. By utilizing open-source Python libraries and leveraging the system's existing hardware, the solution remained both budget-friendly and accessible for academic or small-scale agricultural applications.

4.3. Feasibility Study

The system demonstrated promising results for practical use, particularly in small to medium-scale farming environments. Models were able to function efficiently without requiring high-end infrastructure. The use of preprocessed leaf images allowed for accurate predictions using relatively lightweight models, proving the approach to be viable even in resource-constrained scenarios.

4.4. Results of Implementation

- **Traditional CNN**:
  Served as the baseline model but achieved a low overall accuracy of 40%. It showed poor precision, recall, and F1-scores across all three classes (Healthy, Powdery, Rust). The confusion matrix revealed significant misclassifications, indicating the model failed to learn distinguishing features effectively. It requires architectural improvements or additional preprocessing.

- **EfficientNet**:
  Delivered significantly better results with an overall accuracy of 97%. It demonstrated balanced precision and recall for all classes, with especially strong performance in classifying Rust and Powdery leaves. EfficientNet achieved high F1-scores, making it a reliable model for plant disease classification while maintaining computational efficiency.

- **YOLOv8**:
  Achieved the highest accuracy of 98%, with near-perfect precision, recall, and F1-scores across all disease classes. It accurately detected and classified leaf conditions with minimal false positives and false negatives. YOLOv8 also offers real-time performance benefits, making it ideal for deployment in mobile or field applications.

4.5. Result Analysis

Confusion Matrices:

- YOLOv8 and EfficientNet had very few misclassifications, showcasing strong learning of class-specific patterns.
- CNN showed widespread misclassification, with a near-random spread across classes, especially failing to distinguish between Rust and Powdery.

Classification Metrics:

- YOLOv8: F1-Scores > 0.97 across all classes (Healthy, Powdery, Rust).
- EfficientNet: F1-Scores also high (above 0.95), slightly lower for Healthy.
- CNN: F1-Scores remained low (0.37–0.44 range), indicating weak model learning.

Model Efficiency:

- YOLOv8 offered the best trade-off between speed and accuracy.
- EfficientNet was computationally lighter than YOLOv8 but still highly effective.
- CNN was fast to train but failed in learning useful representations.

4.6. Observation/Remarks

- The Traditional CNN model struggled to learn class-specific features, likely due to limited depth, lack of data augmentation, or poor regularization.
- EfficientNet provided an excellent balance of performance and computational cost, ideal for deployment on moderately powered devices.
- YOLOv8 demonstrated the most robust results, excelling in both accuracy and speed, making it highly suitable for real-time detection tasks.
- Advanced architectures like EfficientNet and YOLOv8 leveraged pre-trained weights and deeper representations, resulting in significantly better generalization.

Chapter 5: Conclusion

5.1. Conclusion

The project titled "Plant Disease Detection using Deep Learning" presented a practical and effective approach for identifying diseases in plant leaves through image-based classification. By utilizing advanced deep learning architectures—specifically Traditional CNN, EfficientNet, and YOLOv8—we developed a system capable of learning visual patterns from leaf images and accurately detecting signs of disease such as Rust and Powdery Mildew, as well as healthy conditions.

Throughout the development process, strong emphasis was placed on image preprocessing, model tuning, and comprehensive evaluation metrics including precision, recall, and F1-score. Each model contributed distinct strengths: the CNN model provided a foundational benchmark, EfficientNet demonstrated powerful feature extraction with optimized resource usage, and YOLOv8 delivered outstanding accuracy with real-time detection capabilities.

5.2. Future Scope

While the current implementation lays a strong foundation for automated plant disease detection, there are several directions for future improvement:

- **Real-time Deployment**: Integrating the models into a mobile or web-based application could enable farmers or agronomists to detect diseases in real time using smartphone cameras.
- **Expanded Dataset**: Including a broader range of crops and disease types will help generalize the model, making it suitable for diverse agricultural conditions.
- **Field Condition Images**: Training the models with real-world, in-field leaf images (with varied backgrounds, lighting, and orientations) would improve robustness and performance in practical environments.
- **Multi-modal Learning**: Combining image data with sensor data (like temperature, humidity, or soil health) could enhance prediction reliability and offer deeper insights into plant health.

5.3. Societal Impact

The deployment of AI-based plant disease detection systems carries several benefits for society, particularly in the context of sustainable agriculture and food security:

- Early Detection for Better Yield: Timely identification of plant diseases helps prevent widespread crop damage, supporting better yield and reducing financial losses for farmers.
- Accessibility for Small Farmers: Cost-effective and easy-to-use tools can empower small-scale farmers with limited access to agricultural experts.
- Reduced Pesticide Use: Targeted diagnosis allows for more precise application of pesticides, reducing environmental impact and promoting healthier farming practices.
- Boosting Agri-Tech Adoption: Projects like this contribute to the growing ecosystem of agricultural technology, encouraging innovation and digital transformation in farming communities.
- Educational Value: The methodologies applied serve as learning material for students and researchers interested in agricultural AI, promoting interdisciplinary research and awareness.