PRESIDENCY UNIVERSITY
Private University Estd. in Karnataka State by Act No. 41 of 2013
BANGALORE

A Project Report

On

"Detection of Plant Leaf Disease Using Image Processing and Recommendation of Pesticide"

Batch Details

Sl. No.	Roll Number	Student Name
1	20211CSE0231	SHREE BHUVAN S R
2	20211CSE0263	ANIKA NAYAK
3	20211CSE0381	SHREYAS D M
4	20211CSE0355	UDEEP S LOKESH

School of Computer Science,

Presidency University, Bengaluru.

Under the guidance of,

Ms. Tintu Vijayan

School of Computer Science,

Presidency University, Bengaluru

CONTENTS

1. INTRODUCTION

Agriculture is considered the strength of Indian budget. On the other hand, the farming of crops for optimal harvest and quality produce is very important. The project aims to leverage modern technology to address the critical challenge of plant diseases in agriculture. Timely and accurate identification of these diseases is essential for maintaining healthy crops and ensuring food security.

Traditional detection methods often rely on manual inspections, which can be inefficient and subjective. Project employs advanced image processing techniques to analyze leaf images and detect various diseases with high precision. By utilizing machine learning algorithms, the system can learn from a diverse dataset of healthy and diseased leaves, improving its accuracy over time.

Overall, this initiative not only enhances agricultural productivity and sustainability but also empowers farmers with the knowledge and tools needed to protect their crops effectively, ultimately contributing to global food security

- **INTRODUCTION TO THE DOMAIN OF THE PROBLEM**

The agricultural sector faces significant challenges from plant diseases, which can lead to reduced crop yields and financial losses for farmers. Traditional methods of disease detection often rely on visual inspections, which are time-consuming and prone to human error. As the global population grows and the demand for food increases, efficient and accurate methods for monitoring plant health become essential.

The advent of image processing and machine learning technologies offers a promising solution to this problem. By utilizing digital images of plant leaves, these technologies can analyze visual data to identify symptoms of diseases quickly and accurately. This automated approach not only enhances the detection process but also enables early intervention, which is crucial for minimizing damage.

In addition to disease detection, providing recommendations for pesticide application is vital for effective pest management. Many farmers lack access to expert advice on appropriate treatments, leading to either overuse or misuse of pesticides, which can harm the environment and human health. Integrating a recommendation system into disease detection can empower farmers with knowledge, helping them make informed decisions about pest control measures tailored to specific diseases.

This project, therefore, aims to bridge the gap between technology and agricultural practices by developing a comprehensive system for detecting plant leaf diseases and recommending suitable pesticides. By leveraging advancements in image processing and data analysis, the project seeks to promote sustainable agricultural practices, enhance crop productivity, and ultimately contribute to food security.

2. LITERATURE REVIEW

Here are some existing models and approaches related to the detection of plant leaf diseases using image processing, along with systems that recommend pesticides:

1. Convolutional Neural Networks (CNNs)

Overview:
CNNs are widely used for image classification tasks, including the detection of plant diseases. They excel at learning spatial hierarchies from images.
EX: AlexNet, VGGNet, and ResNet.

Advantages:
1. Automatic Feature Extraction: CNNs automatically learn relevant features from images, reducing the need for manual feature engineering.
2. Translation Invariance: They can recognize objects in images regardless of their position, orientation, or scale, making them robust to variations in the input data.
3. High Performance on Image Tasks: CNNs have achieved state-of-the-art performance in many computer vision tasks, such as image classification and object detection.

Limitations:
1. Data Requirement: CNNs typically require a large amount of labeled data for training to achieve optimal performance, which can be a challenge for some applications.
2. Computationally Intensive: Training CNNs can be resource-intensive, requiring significant computational power and time. especially for large models or datasets.

2. Image Processing Techniques

Traditional Methods: Techniques such as color segmentation, texture analysis, and morphological operations have been employed to preprocess images and highlight disease symptoms.
EX: Pipeline.

Advantages:
1. Enhanced Image Quality: Improves visual quality through techniques like filtering and noise reduction.
2. Feature Extraction: Enables extraction of relevant features for analysis and classification.
3. Automated Analysis: Facilitates quick analysis of large datasets, enhancing decision-making efficiency.

Limitations:
1. Sensitivity to Noise: Performance can degrade with noisy images or artifacts.
2. Computational Complexity: Some techniques require significant processing power and time.
3. Dependence on Quality of Input: Relies heavily on the quality of input images for accurate results

3. Hybrid Models

Overview:
Combining image processing techniques with machine learning models, such as Support Vector Machines (SVM) or Random Forests, can enhance disease detection accuracy.
EX: Framework

Advantages of Hybrid Models:
1. Improved Performance: Combining multiple techniques (e.g. machine learning and deep learning) often leads to enhanced accuracy and robustness in predictions.
2. Flexibility: Hybrid models can adapt to various types of data and tasks, allowing for tailored solutions based on exact requirements.
3. Leveraging Strengths: They can take advantage of the strengths of different approaches, such as using feature extraction from traditional methods with classification from advanced algorithms.

Limitations of Hybrid Models:
1. Increased Computational Cost: Combining multiple approaches often results in higher computational resource requirements and longer training times.
2. Difficult Implementation: Implementing hybrid models can be challenging due to the need for expertise in multiple domains and potential compatibility issues between different algorithms.

4. Deep Learning Frameworks

TensorFlow/Keras:
These frameworks are commonly used to build and train deep learning models for disease detection.
Ex: Adaptation

Advantages:
1. Ease of Use: They provide user-friendly APIs and high-level abstractions, simplifying the process of building and training deep learning models.
2. Flexibility: Many frameworks support various architectures and algorithms, allowing users to experiment with different models and customize them as needed.
3. Community and Support: Popular frameworks have large communities and extensive documentation, providing resources and support for troubleshooting and best practices.

Limitations:
1. Resource Intensive. Training deep learning models often requires significant computational resources, such as powerful GPUs, which may not be accessible to all users
2. Black Box Nature: The complexity of deep learning models can make it difficult to interpret their inner workings and understand how they make decisions
3. Overhead for Simple Tasks. For simpler tasks, the use of deep learning frameworks may introduce unnecessary overhead compared to traditional machine learning approaches.

5. Pest Management Systems

Integrated Pest Management (IPM):
Some systems incorporate disease detection with IPM strategies, recommending pesticides based on the type of disease detected.
Ex: Platform

Advantages:
1. Data-Driven Decision Making, Provides access to analytics for informed pest control strategies.
2. Targeted Treatment. Enables specific pest identification for precise and reduced pesticide usage.
3. Cost Efficiency: Optimizes pesticide application, les cring costs and improving profitability.

Limitations:
1.Dependence on Technology: Relies heavily on technology, which may be inaccessible in some areas.
2.Initial Setup Costs: Involves significant initial costs for implementations and training.
3.Data Quality and Accuracy: Effectiveness hinges on the quality of data collected, poor data can lead to ineffective strategies

6. Mobile Applications

Plantix:
A mobile app that allows users to take pictures of their plants and receive instant diagnoses and treatment recommendations, including pesticide suggestions.
Ex: App.

Advantages of Mobile Applications:
1. Accessibility: Mobile apps provide users with easy access to services and information anytime and anywhere, enhancing convenience and user engagement.
2. User Experience: They offer a more tailored and interactive user experience with optimized interfaces, features, and functionalities designed specifically for mobile devices.
3. Offline Functionality: Many mobile apps can work offline, allowing users to access certain features and data without an internet connection.

Limitations of Mobile Applications:
1. Device Compatibility: Apps may not be compatible with all devices or operating systems, limiting their reach and usability
2. Development Costs: Creating and maintaining mobile applications can be costly and time-consuming, especially when developing for multiple platforms
3. Internet Dependency: Despite some offline features, many apps still rely heavily on internet connectivity for full functionality, limiting usability in remote areas.

3. OBJECTIVES

1. Disease Detection: Develop a robust image processing algorithm to accurately detect and classify various plant leaf diseases from images.

2. Feature Extraction: Implement techniques for extracting relevant features (e.g., color, texture, shape) from leaf images to aid in disease identification.

3. Pesticide Recommendation: Create a system that links identified diseases to a database of pesticides. providing users with effective treatment options based on best practices.

4. User-Friendly Interface: Design an intuitive interface that allows users to upload images easily and receive instant feedback on disease diagnosis and treatment recommendations. I

5. Performance Evaluation: Evaluate the accuracy and efficiency of the detection algorithm through testing with a diverse dataset of plant leaf images.

6. Educational Resource: Provide users with information on plant diseases, symptoms, and best practices for pest management to promote sustainable agriculture.

7. Real-Time Analysis: Aim for the system to perform real-time analysis to assist farmers in timely decision-making regarding plant health management.

EXPERIMENTAL DETAILS/METHDOLOGY

Hardware:
1. Camera or Smartphone: To capture high-quality images of plant leaves for analysis.

2. Computer or Server: For running the image processing algorithms and deep learning models, which may include desktops, laptops, or dedicated servers,

3. GPU (Graphics Processing Unit): Useful for accelerating training and inference of deep learning models particularly when working with large datasets.

Software:
1. Programming Languages:
Python: Commonly used for implementing machine learning, deep learning models, and image processing algorithms

2. Deep Learning Frameworks:
TensorFlow/Keras: For building and training convolutional neural networks (CNNs) used in image classification.

OpenCV: For image preprocessing and feature extraction tasks like segmentation, resizing, and filtering

3. Integrated Development Environments (IDEs):
Jupyter Notebook: For prototyping and visualizing data and models.

PsCharm or VS Code: For more extensive development work

4. METHODOLOGY

- DESIGN PROCEDURE

The methodology involved a systematic approach to building a robust plant disease detection system using machine learning. The project started with data collection, where images of various diseased and healthy plant leaves were gathered from publicly available agricultural datasets and field samples.

After acquiring the data, the images underwent a pre-processing phase where techniques like resizing, noise removal, and contrast enhancement were applied to improve clarity and ensure uniformity across all samples.

Next, image segmentation was performed to isolate regions of interest, particularly the diseased parts of the leaves. For this, techniques such as Otsu's thresholding and K-means clustering were explored to divide the images into meaningful segments. The segmented portions were then passed through a feature extraction pipeline, where key attributes such as color histograms, texture patterns using Gabor filters, and shape features were extracted.

The final step in the methodology was classification. The Random Forest algorithm was chosen due to its high accuracy and ability to handle high-dimensional datasets. By using a combination of decision trees, the random forest classifier built multiple models and aggregated their results to predict whether a leaf was healthy or affected by disease.

The classifier was trained on a labeled dataset of plant leaf images and validated using k-fold cross-validation to ensure consistent performance across different subsets of data. Performance metrics such as accuracy, precision, recall, and F1-score were used to evaluate the model’s effectiveness. Continuous optimization of hyperparameters, including the number of trees and the depth of each tree, was carried out to maximize the performance.

5. OUTCOMES

Once the model is fully developed and tested, it is projected to achieve a classification accuracy of over 90% on real-world plant leaf datasets. Through careful segmentation and feature extraction, the model is expected to successfully identify diseased regions with high precision. Key expected results include:

High Precision and Recall: The classifier should be able to accurately differentiate between healthy and diseased leaves, with precision and recall expected to exceed 90%. This will minimize both false positives (healthy leaves mistakenly classified as diseased) and false negatives (diseased leaves missed by the model), which is crucial for practical agricultural applications.

Adaptability to Field Conditions: Given the integration of image pre-processing techniques, the model is expected to perform robustly in varying field conditions, such as changes in lighting and background noise. This adaptability will be crucial for ensuring that the system can be deployed in real-world agricultural environments without the need for controlled setups.

Real-Time Detection Potential: Once optimized, the system is anticipated to be fast enough for near real-time disease detection, making it a valuable tool for continuous monitoring of large-scale crops. Farmers will be able to receive timely alerts about potential outbreaks, allowing for quick intervention to prevent disease spread and minimize crop damage.

6. TIMELINE OF THE PROJECT/ PROJECT EXECUTION PLAN

The project is divided into five major stages and each stage is preceded by a review and evaluation section. The timeline is detailed below:

Review 0 (Sep 12 - Sep 18):
Preliminary formulation of ideas regarding the project and the submission of the first design and problem formulation. In this phase, information related to the work domain will be obtained and the range and the techniques for the identification of plant diseases using machine learning methods will be determined.

Review 1 (Oct 15 - Oct 21):
This is the first formal mechanism for assessing the performance of the project. At this stage the general design and the general methodology should be clearer and first activities regarding data collection and the first conception of a model should be described. Important activities involve obtaining and cleaning the datasets for applying image processing methodologies.

Review 2 (Nov 19 - Nov 22):
The project following this stage will concentrate on developing as well as testing of models. The second review will evaluate the advancements in the usage of the machine learning algorithm (Random Forest) as well as image segmentations. What is presented here should be the preliminary outcomes from the testing of the model with special emphasis laid on how the various problems that are encountered are being solved.

Review 3 (Dec 17 - Dec 20):
The third review will focus considerable on the assessment of the final model's performance. By this phase, the system should have had some tests specifically to its reliability and efficiency in actual situations. Project results will be presented, some recommendations on its possible development and further application will be made, as well as proposals for possible further expansion will be provided.

Final Viva Voce (Jan 10 – Jan 17):
The last is the form submission and presentation of the case-archive. Consequently, all the results obtained, conclusions made together with the overview of future works and prospects will be demonstrated. The assessment of the system made at this phase relies on the evaluation criteria of the detection and classification of plant diseases.

7. CONCLUSION

The ongoing development of this project promises to deliver an innovative solution for early detection of plant diseases using machine learning and image processing techniques. By utilizing advanced segmentation methods and the random forest classification algorithm, the model is expected to achieve high accuracy in identifying and classifying plant diseases from leaf images. Once fully implemented, the system will be adaptable to real-world agricultural environments and capable of near real-time monitoring, making it a valuable asset in modern precision farming.

The anticipated results will not only help in early disease detection but also offer farmers a practical tool to minimize crop losses, thus contributing to higher yields and more efficient agricultural practices. Future work will focus on optimizing the model further and exploring integration with IoT-based systems for automated and large-scale deployment. As agriculture faces increasing challenges due to climate change and the growing demand for food, the successful implementation of this project has the potential to make a significant impact on sustainable farming practices.

8. REFERENCES

1.
Samajpati BJ, Degadwala SD (2016) Hybrid approach for apple fruit diseases detec-tion and classification using random forest classifier. In: 2016 International conference on communication and signal processing (ICCSP). IEEE, pp 1015–1019
2.
Sandika B, Avil S, Sanat S, Srinivasu P (2016) Random forest based classification of diseases in grapes from images captured in uncontrolled environments. In: 2016 IEEE 13th international conference on signal processing (ICSP). IEEE, pp 1775–1780
3.
Zawbaa, HM, Hazman M, Abbass M, Hassanien AE (2014) Automatic fruit classification using random forest algorithm. In: 2014 14th International conference on hybrid intelligent systems. IEEE, pp 164–168
4.
Iqbal MA, Talukder KH (2020) Detection of potato disease using image segmentation and machine learning. In: 2020 International conference on wireless communications signal processing and networking (WiSPNET). IEEE, pp 43–47
5.
Chauhan MD (2021) Detection of maize disease using random forest classification algorithm. Turkish J Comput Math Educ (TURCOMAT) 12(9):715–720
6.
Deepak AH, Gupta A, Choudhary M, Meghana S (2019) Disease detection in tomato plants and remote Monitoring of agricultural parameters. In: 2019 11th International conference on advanced computing (ICoAC). IEEE, pp 28–33
7.
Hidayatuloh A, Nursalman M, Nugraha E (2018) Identification of tomato plant diseases by leaf image using squeeze net model. In: 2018 International conference on information technology systems and innovation (ICITSI). IEEE, pp 199–204
8.
Sarangdhar AA, Pawar VR (2017) Machine learning regression technique for cotton leaf disease detection and controlling using IoT. In: 2017 International conference of electronics, communi-cation and aerospace technology (ICECA), pp 449–454. https://doi.org/10.1109/ICECA.2017. 8212855
9.
Hatuwal BK, Shakya A, Joshi B (2020) Plant leaf disease recognition using random forest, KNN, SVM and CNN. Polibits 62:13–19

10.
Dhingra G, Kumar V, Joshi HD (2018) Study of digital image processing techniques for leaf disease detection and classification. Multimedia Tools Appl 77(15):19951–20000
11.
Sreenivasa BR, Nirmala CR, Manoj Kumar MV (2022) Session-based personalized recom-mender system for online shopping. In: Shetty NR, Patnaik LM, Nagaraj HC, Hamsavath PN, Nalini N (eds) Emerging research in computing, information, communication, and applica-tions. Lecture notes in electrical engineering, vol 790. Springer, Singapore. https://doi.org/10. 1007/978-981-16-1342-5_49
12.
Sreenivasa BR, Nirmala CR (2022) Hybrid time-centric recommendation model for e-commerce applications using behavioral traits of user. Inf Technol Manag. https://doi.org/ 10.1007/s10799-022-00358-8
13.
Das D, Singh M, Mohanty SS, Chakravarty S (2020) Leaf disease detection using support vector machine. In: 2020 International conference on communication and signal processing (ICCSP). IEEE, pp 1036–1040
14.
Albattah W, Nawaz M, Javed A, Masood M, Albahli S (2022) A novel deep learning method for detection and classification of plant diseases. Complex Intell Syst 8(1):507–524
15.
Hungilo GG, Emmanuel G, Emanuel AW (2019) Image processing techniques for detecting and classification of plant disease: a review. In: Proceedings of the 2019 international conference on intelligent medicine and image processing, pp 48–52
16.
Khirade SD, Patil AB (2015) Plant disease detection using image processing. In: 2015 International conference on computing communication control and automation. IEEE, pp 768–771
17.
Ramesh S, Vydeki D (2020) Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm. Information processing in agriculture 7(2):249–260