Plant Disease Detection System for Sustainable Agriculture

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

Pratyaoy Sarkar, sarkarpratyaoy@gmail.com

Under the Guidance of

P.Raja, Master Trainer, Edunet Foundation

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deep sense of gratitude to all individuals who helped me directly or indirectly during this thesis work.

Firstly, I am deeply grateful to Edunet Foundation for providing me with the opportunity to complete this internship successfully. I would like to thank my super trainer P.Raja sir, for being a great mentor and the best adviser I could ever have. His guidance and support were invaluable throughout this journey. His constant support and inspiration has helped me a lot to complete this project successfully . And I want to thank my mentor Pavan Sumohana sir, who constantly let me know all the schedules and the topics whatever would be covered by the session, on time. Your insights and encouragement have significantly contributed to my personal and professional growth. The confidence shown in me by him was the biggest source of inspiration for me. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

Moreover, I appreciate the entire team of Edunet Foundation for creating a conductive learning environment and for being welcoming and supportive. The experience has not only expanded my understanding but also inspired me to further my pursuits in this field.

Thank you once again for this wonderful opportunity.

Sincerely,

Pratyaoy Sarkar, sarkarpratyaoy@gmail.com

ABSTRACT

Agriculture is a critical sector for the global economy, but it faces significant challenges due to plant diseases, which can lead to substantial crop losses. Traditional methods of plant disease detection rely heavily on manual inspection, which is time-consuming, labor-intensive, and often inaccurate. With the advancement of technology, there is a need for an automated, efficient, and accurate system to detect plant diseases early, ensuring timely intervention and reducing crop losses.

To develop a deep learning-based system that can accurately detect and classify plant diseases from images of plant leaves. The system should be able to process images captured in various conditions (e.g., different lighting, angles, and stages of disease progression) and provide real-time feedback to farmers and agricultural experts.

A large amount of dataset was taken, which contain different images of defected area of plant with the name of the disease. Then this dataset was divided into training and testing or validation set. Then I have used CNN algorithm and Keras library, scikit-learn. Then we have built the model and added the image layers by applying different levels of filters into it. Then we have flattened all the layers and converted to single dimension. After that we have calculated the training accuracy and testing accuracy.

As a result, we got the visualization of the accuracy result, with training accuracy and validation accuracy. We also got the confusion matrix of the actual class and the predicted class of the images. Finally, we got 81% training accuracy and 92% validation accuracy.

The development and implementation of a deep learning-based system for plant disease detection have demonstrated significant potential in revolutionizing agricultural practices. By leveraging advanced techniques in image processing and deep learning, the project has successfully created a robust model capable of accurately identifying and classifying plant diseases from images of plant leaves.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1Problem Statement:

Plants play a vital role in economy and climate change. Many studies have proved that extinction of plants due to industry use have caused damage to ozone layer and thus resulting global warming. The rate of climate change forecast for the future is 10-100 times faster than the rate of DE glacial warming. Balance of global food production is also a major issue. Apart from this, in health care also plants play vital role. Overall plants are very essential for human survival therefore it's also a worldwide concern to take care of them.

Now a days due to excessive pollution of environment, not only the humans and animals, but plants are also suffering a lot. They are getting affected by various new type of disease also. But plants are concurrently connected with our daily life. So, to keep them safe from disease, we need to first detect the types of diseases from which they are suffering. After that, we have to treat them. But, when a bulk amount of crop production happen or in any forest, it's not possible to detect the disease of plants manually. So, we need to take help of some well learned learning algorithms to detect that and some IOT devices. That's why we are doing this project, to get a prototype of that broad scale.

1.2Motivation:

As I belong from a family, where farming is our main occupation. I have seen my father farming and growing crops in his fields, from my childhood. When I became adult, now I can understand, when plants get defected by diseases, how fast it spread and finishes all the production. That is a huge loss for environment and for the farmer himself also. This made me feel so upset. From that day, I decided, if I get any chance to contribute my knowledge to resolve this issue, that will be very proud moment for me.

Potential impact of this project is Precision Agriculture, Automated Monitoring, Research and Development, Supply Chain Management, Environment Monitoring.

Impact: Increased Crop Yields, Cost Reduction, Sustainability, Data-Driven Decisions, Global Food Security.

1.3Objective:

The objectives of the project is to develop an automated system that can accurately identify and diagnose diseases in plants. This involves training deep learning models on a large dataset of plant images to recognize patterns and features associated with various diseases.

Early Detection: Identify plant diseases at an early stage to allow for timely intervention and treatment, preventing the spread of the disease and minimizing crop loss.

Accuracy: Develop highly accurate models that can differentiate between healthy plants and those affected by specific diseases, reducing false positives and false negatives.

Scalability: Create a system that can be easily scaled to monitor large agricultural fields using drones, cameras, or other imaging devices.

Ease of Use: Ensure the system is user-friendly and accessible to farmers, researchers, and agricultural workers, even those with limited technical knowledge.

Cost-Effectiveness: Develop a solution that is affordable and reduces the overall cost of disease management by minimizing the need for manual inspections and chemical treatments.

Real-Time Monitoring: Enable real-time monitoring of plant health, providing instant feedback and recommendations to farmers for effective disease management.

1.4Scope of the Project:

Scope: Detection of Multiple Diseases, Deployment in Various Environment, Integration with IoT Devices, Real-Time Analysis, Support for Multiple Platforms.

Limitations:

Data Dependency: The accuracy of the model is highly dependent on the quality and quantity of the training data. Insufficient or poor-quality data can lead to inaccurate detection.

Environmental Factors: Variations in lighting, weather conditions, and camera quality can affect the accuracy of the disease detection.

Resource Intensive: It requires significant computational resources for training and deployment. Also other limitations are Initial Costs, Maintenance etc.

CHAPTER 2

Literature Survey

Related Literature: Several studies on ML and DL approaches used in agriculture have been evaluated, leaving a gap to comprehensively examine image-centered plant disease detection. Most recently, a review was made of the imperativeness of existing PDD methods and included segmentation, classification, localization, and disease techniques. The study concentrated on the performances of the CNN method in PDD, primarily on fruits, vegetables and various plants. Other studies like focused on barley, maize, rice, soy- beans, and wheat and compared the benefits and drawback of various IPs, segmentations, extractions, feature selections, and classifications. The main goal of this survey was to demonstrate the use of electric impedance spectroscopy, hyperspectral imaging, and fluorescence spectroscopy, among other things. They noticed that the segmentation strategy known as ‘‘k-means’’ was frequently used to identify plant diseases. A comprehensive review of existing approaches in the early 2000s was conducted, with implications for fully automated PDD and identification using digital approaches[1]. The review also discusses the various kinds of openly accessible datasets that scientists in the field use. In addition, this review examines the influence of current trends in DL, localization, TF, and attention processes, together with lightweight approaches in PDD.

Existing Techniques: Deep learning, particularly Convolution Neural Networks (CNNs), has become a powerful tool for plant disease detection. CNN can automatically learn features from images, making them highly effective for image-based disease classification. Some popular CNN architecture include AlexNet, VGGNet, and ResNet. Data augmentation techniques, such as rotation, flipping, and scaling are used to increase the diversity of training data, which helps improve model robustness and generalization. Remote sensing technologies, including satellite imagery and drones, can detect plant disease by capturing changes in plant reflectance patterns. This method is non-invasive and can cover large areas quickly. Traditional machine learning algorithms, such as Support Vector Machines (SVM) and Random Forests, are also used for plant disease detection. These algorithms require feature extraction and selection but can be effective when combined with other techniques.

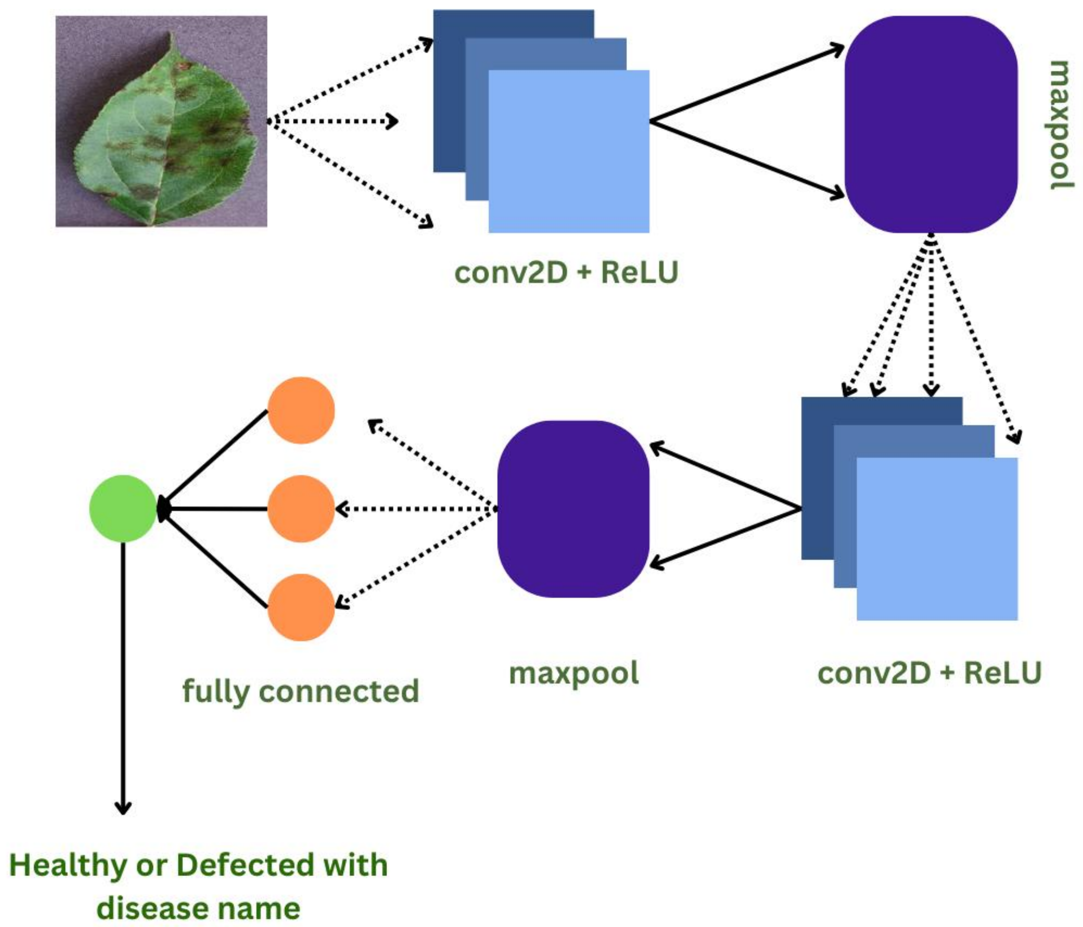
Limitations of existing solutions: High-quality, labeled datasets are crucial for training accurate models. However, obtaining such datasets can be challenging due to variations in lighting, angles, and backgrounds in real-world conditions. Detecting multiple diseases in a single leaf or plant remains a challenge due to the subtle differences between disease symptoms. While integrating deep learning with IoT devices (like drones) for field-based detection is promising, it still faces challenges in terms of data transmission, processing, and storage. Also cost and complexity is a matter of concern.

CHAPTER 3

Proposed Methodology

System Design:

Figure 1: Layout of the model



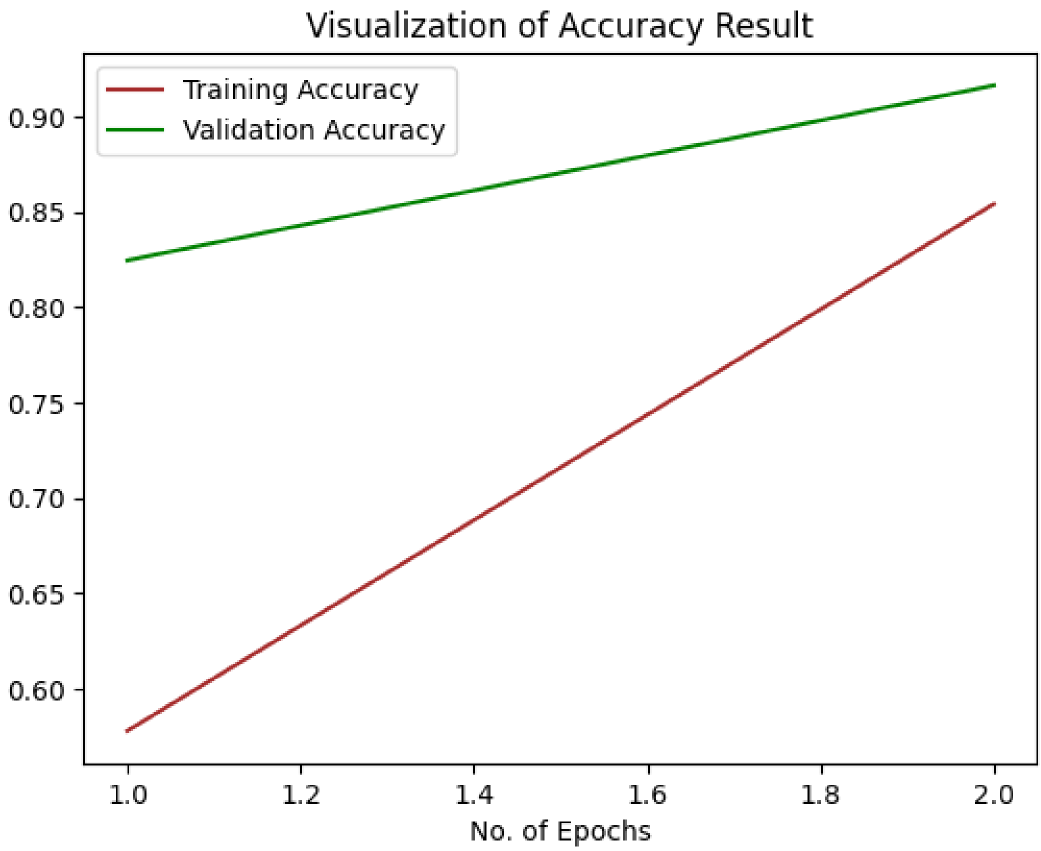
The first step of the process is image collection, which is based on Computer Vision. We collect healthy and diseases plant leaves from various sources through cameras, drones etc. Then the important step is **Annotations** or labelling the images to indicate the presence and type of disease. Next step is **Data Preprocessing** which include image enhancement and data augmentation. Then we choose the deep learning model, here we have used **Convolutional Neural Networks (CNN).** Here we have trained the model through various layers of CNN and the activation function we have used is **Rectified Linear Unit (ReLU).** After this we have flattened the all layers to convert in into a single array like structure. And we got this summary.

Figure 2: Summary after applying CNN layers.

Layer (type)	Output Shape	Param #
Conv2d(Conv2D)	(None, 128, 128, 32)	896
Conv2d_1(Conv2D)	(None, 126, 126, 32)	9248
max_pooling2d	(None, 63, 63, 32)	0
Conv2d_2(Conv2D)	(None, 63, 63, 64)	18,496
Conv2d_3(Conv2D)	(None, 61, 61, 64)	36,928
max_pooling2d_1	(None, 30, 30, 64)	0
Conv2d_4(Conv2D)	(None, 30, 30, 128)	73,856
Conv2d_5(Conv2D)	(None, 28, 28, 128)	147,584
max_pooling2d_2	(None, 14, 14, 128)	0

After that we have fit the training set and the validation set in the model with 2 epochs. Now the training of the model is done. Validate the model on a separate dataset to tune hyperparameters and avoid overfitting. After the model evaluation is done. Evaluated the model using metrics like accuracy, precision, recall, and F1-score. Using the confusion matrix we have analyzed the model's performance in identifying different diseases. After the deployment of the model. Integrated the trained model into an application or system for real-time detection. Developed a user-friendly interface for users to upload images and receive disease predictions. Here we have used the streamlit library to make the user interface. Now monitoring and maintenance. Updated the model with new data to improve accuracy over time. Regularly monitored the system's performance and make necessary adjustments.

Figure 2: Visualization of Accuracy Result



Requirement Specification

The tools and technologies required to implement the solution.

3.1.1 Hardware Requirements: A high-performance CPU (e.g., Intel i7/i9, AMD Ryzen 5/7/9) is recommended for handling complex computations. A dedicated GPU (e.g., AMD RADEON, NVIDIA RTX 3060) for significantly speeds up the training process by parallelizing computations. An IOT device and camera or drones to capture the image perfectly.

3.1.2 Software Requirements: We have used the CNN model to train the dataset. There various library was required as matplotlib to visualize the graph and plots. Tensorflow was used. Also keras, pandas and seaborn was used to create the heatmap.

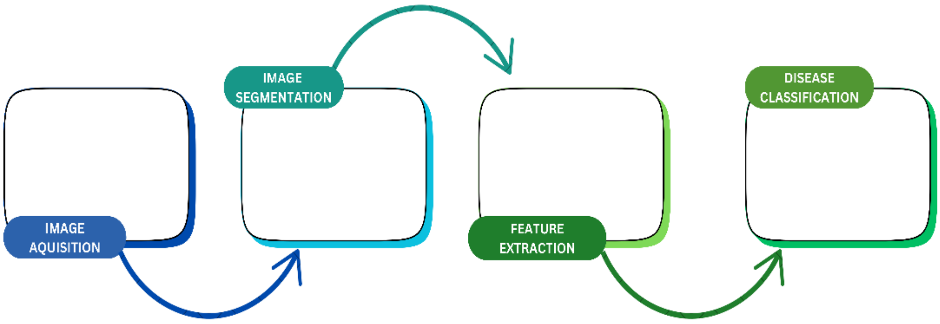


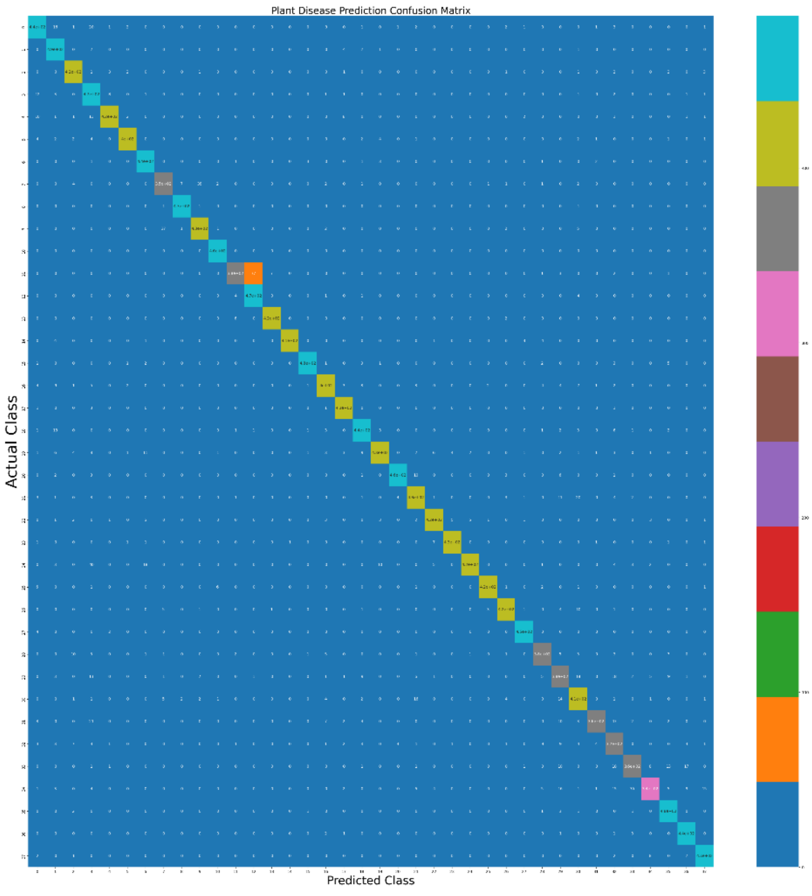
Figure 2: Process of the model building

CHAPTER 4

Implementation and Result

Snap Shots of Result:

Figure 3: Plant Disease Prediction Confusion Matrix

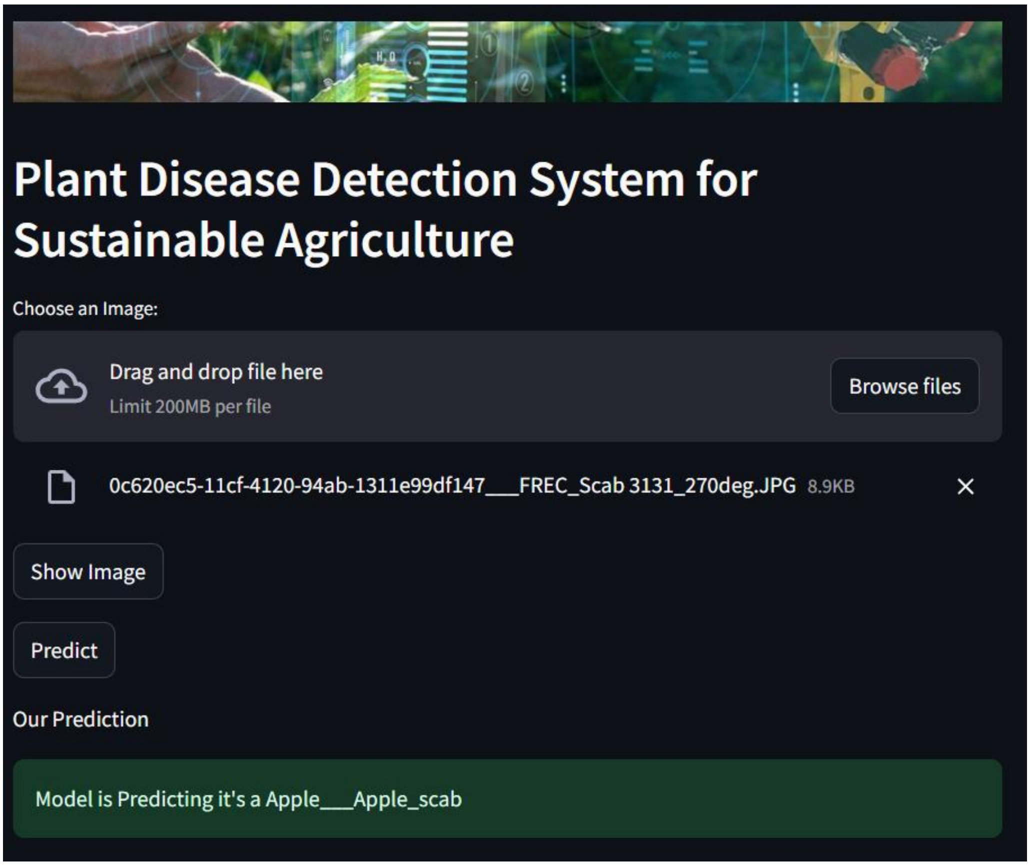


This above image represent the confusion matrix generated from the model as a result, where the X-axis represent the predicted class and the Y-axis represent the actual class.

Figure 4: The UI of the model

This is the first user interface of the model and it is done by using the streamlit library.

Figure 5: Prediction of the image in the UI



GitHub Link for Code: [PratyaoySarkar/Plant_Disease_Detection: This is a deep learning project based on Plant Disease Detection]

CHAPTER 5

Discussion and Conclusion

Future Work:

This is the very beginning of this technology in modern world. In future there are various opportunity to work on it and improve it. As, developing models capable of real-time detection is crucial for timely intervention and management of plant diseases. This requires advancements in hardware and software to process images quickly and accurately in the field. Transfer learning and Few-shot learning can help models generalize better to new diseases and conditions with limited training data. This is particularly useful for detecting rare or emerging plant diseases. Combining different types of data (e.g., images, spectral data, environmental conditions) can provide a more comprehensive understanding of plant health and improve detection accuracy. Ensuring that deep learning models are interpretable and explainable is important for gaining the trust of farmers and agronomists. This involves developing methods to visualize and explain model predictions. Creating models that are robust to variations in lighting, weather conditions, and other environmental factors is essential for reliable plant disease detection in different settings.

Conclusion:

The overall impact and contribution of a project focused on plant disease detection using deep learning can be quite significant. Enabling early and accurate detection of plant diseases, the project helps farmers take timely action to mitigate damage, leading to higher crop yields and improved food security. The project empowers farmers with accessible and reliable tools to monitor plant health, making advanced agricultural technology available to a broader audience. By improving the health and productivity of crops, the project contributes to the global effort to ensure a stable and secure food supply for the growing population. Overall, this project holds the potential to transform agriculture by making it more efficient, sustainable, and resilient to plant diseases, ultimately benefiting farmers, consumers, and the environment.

REFERENCES

[1] WASSWA SHAFIK 1, (Member, IEEE), ALI TUFAIL 1, (Senior Member, IEEE), ABDALLAH NAMOUN 2, (Member, IEEE), LIYANAGE CHANDRATILAK DE SILVA 1, (Senior Member, IEEE),
AND ROSYZIE ANNA AWG HAJI MOHD APONG1