Detection of Plant Leaf Diseases Using Image Processing

K. VASUMATHI[1], S. SELVAKANI[2] , G. PRAKASH[3]

[1]PG Department of Computer Science,
Government Arts and Science College,
Arakkonam, Tamilnadu,
INDIA

[2]PG Department of Computer Science,
Government Arts and Science College,
Arakkonam, Tamilnadu,
INDIA

[3]PG Scholar, PG Department of Computer Science,
Government Arts and Science College,
Arakkonam, Tamilnadu,
INDIA

Abstract:- Agricultural productivity is strongly affected by plant diseases caused by microorganisms, pests, and bacteria that spread through parts of the plant like leaves, stems, and fruits. Early detection is crucial to prevent these diseases from reducing crop quality and yield, which harms the economy, especially in agricultural countries. Traditionally, farmers inspect plants manually for diseases, but this method is slow and often inaccurate, with detection rates as low as 60%. It also requires expert knowledge, making it impractical for large-scale farms. To solve this problem, modern technologies like image processing and computer vision are used to automate disease detection. Image processing involves analyzing plant images to identify symptoms using techniques such as image enhancement, segmentation, feature extraction, and classification. This process can convert images from RGB to other color spaces like YCbCr to make disease symptoms clearer. The system then extracts features like color, texture, and shape to accurately identify the disease. Machine learning methods, such as support vector machines (SVM), are used for classification. Automated systems can detect diseases with up to 99% accuracy, much higher than manual methods. Real-time crop image capturing and disease detection systems provide immediate feedback to farmers about plant health and specific diseases affecting their crops. These systems can also track environmental data, such as temperature and humidity, which helps in disease management and crop growth. Content-based image retrieval (CBIR) systems that compare features like color, shape, and texture improve disease detection accuracy. These systems help farmers make better decisions about irrigation, fertilization, and disease control. By automating disease detection, farmers can save time, reduce the need for expert help, and increase crop productivity, leading to more efficient farming and better food security.

Key-Words:- Agricultural productivity, Plant diseases, Image processing, Computer vision, Support vector machines (SVM), Environmental data, Content-based image retrieval (CBIR), Crop productivity, Disease detection, early detection.

Received: April 15, 2024. Revised: March 9, 2025. Accepted: April 11, 2025. Published: May 8, 2025.

1 Introduction

The traditional and classical methods for detecting and recognizing plant diseases largely rely on visual observation, a process that is not only slow but also prone to inaccuracies. This method often leads to delayed identification, which can hinder effective intervention and disease management. In many regions, the process of consulting experts for disease diagnosis is both costly and time-consuming, especially considering the scarcity of skilled professionals. Furthermore, infrequent monitoring of crops can allow plant diseases to spread unchecked, worsening the situation. Given the significance of agriculture in such regions, especially in countries like India, these limitations pose a significant challenge to effective plant health management.

India, as an agrarian nation, offers a diverse array of crops to its farmers, ranging from staple grains to various fruits and vegetables. However, despite the abundance of crops, the detection and management of plant diseases remain complex and resource-intensive. Traditional methods struggle to keep up with the scale of modern agriculture, especially with the diverse array of diseases that affect crops. In response to these challenges, research has led to the development of advanced computational systems capable of automatically identifying plant diseases. These systems analyze images of infected leaf spots, typically captured through digital cameras or mobile devices, to detect disease symptoms. The captured images undergo a process of image processing, where portions of the diseased leaves are isolated and analyzed for classification through advanced algorithms. This system merges image processing with modern computational techniques, allowing for more accurate, scalable, and timely disease detection compared to traditional visual methods.

Currently, manual observation by experts remains the predominant approach for disease detection and identification in many agricultural settings. While this method is effective in some cases, it requires constant monitoring of crops, making it labor-intensive and expensive. This challenge is particularly significant in large-scale farming operations, where expert consultation becomes even more costly and impractical. In several developing countries, farmers may need to travel long distances to consult with experts, further complicating the process and adding to the resource burden. Additionally, farmers may not be aware of emerging non-native diseases that could threaten their crops. These factors make the reliance on expert consultation increasingly less viable, especially as agricultural sectors expand. As a result, the need for automated systems capable of identifying plant diseases at scale has become more critical. These systems can aid significantly in monitoring vast agricultural fields, enabling early detection of diseases and reducing the burden on human experts. By identifying diseases based on the visible symptoms on leaves, these systems also allow for the integration of machine vision, which is useful for automatic inspection, process control, and even robotic guidance. When compared to traditional visual identification, these systems offer an efficient, scalable, and more accurate alternative.

Agriculture remains the backbone of India's economy, with approximately 70% of the population relying directly or indirectly on this sector for their livelihood. However, environmental changes, including erratic rainfall patterns and temperature fluctuations, often contribute to the proliferation of plant diseases. These diseases can manifest in various forms, such as leaf spots, discoloration, drying of leaves, and even defoliation. Many farmers struggle with accurately identifying these symptoms in time to intervene effectively. Given that expert diagnosis is typically required for accurate disease identification, the process remains costly, and it often takes too long to implement corrective measures. To address these challenges, this paper proposes a solution that leverages image processing techniques for the early detection of leaf infections. Advancements in digital image processing, supported by progress in microelectronics and computer technology, have enabled the widespread application of image analysis techniques in the biological sciences. This approach offers a promising solution for plant disease detection, enabling more efficient and accurate disease identification and management.

The automatic detection of plant diseases is vital for early intervention, as it allows farmers to take timely action before diseases spread and cause irreversible damage. This paper presents a MATLAB-based system designed to detect and identify plant diseases, with a focus on both leaf and fruit infections. The system incorporates advanced image processing techniques to enhance the accuracy of disease detection. To begin, high-resolution digital images of both healthy and diseased plants are captured and stored in a database for analysis. The images undergo preprocessing steps to enhance their quality, followed by segmentation using the K-means clustering method. This method divides the images into distinct clusters based on similarity, facilitating the isolation of diseased regions for further analysis. Feature extraction is then performed, identifying key characteristics of the leaf spots or other disease symptoms. After feature extraction, the system employs both the K-means clustering and Support Vector Machine (SVM) algorithms for training and classification, with the aim of accurately identifying the disease based on the processed data. Ultimately, the system classifies the disease based on the distinctive features derived from the images.

2 Related Work

[1]The paper titled "Deep Learning-Based Techniques for Plant Disease Recognition in Real-Field Scenarios," authored by A. Fuentes, S. Yoon, and D. S. Park, was published in 2020. This study investigates the use of deep learning methods, particularly convolutional neural networks (CNNs), for plant disease recognition in real-world agricultural settings. The authors highlight the challenges of implementing these techniques outside of controlled environments, where factors such as fluctuating lighting conditions, varying leaf orientations, and diverse backgrounds can significantly affect performance. To address these challenges, they propose several approaches aimed at enhancing the accuracy and robustness of plant disease recognition systems, with the ultimate goal of enabling their deployment in the field. This would support early disease detection and contribute to more effective crop management in dynamic, real-world conditions.

[2]The paper titled "Plant Leaf Disease Recognition Using Depth-Wise Separable Convolution-Based Models," authored by S. M. M. Hossain, K. Deb, P. K. Dhar, and T. Koshiba, was published in Symmetry in March 2021. This study introduces an efficient approach for recognizing plant leaf diseases using convolutional neural networks (CNNs) that incorporate depth-wise separable convolutions. The proposed method aims to address the limitations of traditional disease recognition techniques, which are often slow and challenging to scale. The authors likely trained the model using a dataset of leaf images representing various diseases and compared its performance with other existing methods, demonstrating superior accuracy. This research presents a promising solution for farmers by enabling the use of automated systems to detect plant diseases, ultimately enhancing crop health management and supporting more efficient agricultural practices.

[3]The paper titled "Survey on SVM and Their Application in Image Classification," authored by M. A. Chandra and S. S. Bedi, was published in the International Journal of Information Technology in October 2021. This paper provides a comprehensive overview of Support Vector Machines (SVMs) and their significant role in image classification tasks. SVMs are highly effective tools for handling classification problems, particularly when dealing with high-dimensional data. The paper explains the fundamental principles of SVMs, including the concepts of hyperplanes, kernels, and margin optimization. It also delves into various kernel types, such as linear, polynomial, and radial basis function (RBF), and their applications in different fields. The authors review the use of SVMs in areas like facial recognition, medical image analysis, and object detection, highlighting both the advantages and the challenges associated with their use in image classification.

[4]The paper titled "An Aggregated Loss Function-Based Lightweight Few-Shot Model for Plant Leaf Disease Classification," authored by S. Garg and P. Singh, was published in Multimedia Tools and Applications in February 2023. This study presents a lightweight model for plant leaf disease classification that utilizes a few-shot learning approach. The model is designed to tackle the challenge of limited labeled data by incorporating an aggregated loss function, which enables the model to learn effectively with fewer samples. This approach improves the model's ability to classify diseases accurately, even when only a small number of images are available for training, making it both efficient and practical for real-world agricultural applications.

[5]The paper titled "Current and Emerging Molecular Technologies for the Diagnosis of Plant Diseases—An

Overview," authored by M. M. F. Azizi, N. H. Mardhiah, and H. Y. Lau, was published in the Journal of Experimental Biology and Agricultural Sciences in April 2022. This paper provides an overview of the latest molecular techniques used for diagnosing plant diseases. The authors examine traditional methods alongside emerging technologies, such as polymerase chain reaction (PCR), next-generation sequencing (NGS), and biosensors, which offer higher sensitivity, specificity, and rapid detection. The paper emphasizes the potential of these advanced methods to improve disease detection accuracy, enable early intervention, and support effective disease management strategies in agriculture.

[6]The paper titled "Weed Density Extraction Based on Few-Shot Learning Through UAV Remote Sensing RGB and Multispectral Images in Ecological Irrigation Areas," authored by S. Wang et al., was published in *Frontiers in Plant Science* in March 2022. This study explores the use of UAVs equipped with RGB and multispectral sensors for detecting weed density in irrigation areas. By employing a few-shot learning approach, the model can accurately identify weeds with limited labeled data, making it highly effective for real-world agricultural applications.

[7]The research paper, titled "Convolutional Neural Network Based Maize Plant Disease Identification," was authored by S. Jasrotia, J. Yadav, N. Rajpal, M. Arora, and J. Chaudhary. It was published in the Proceedings of Computer Science journal, which is a reputable publication in the field of computer science and technology. The paper appears in Volume *218* of the journal and was released in January 2023. It spans pages 1712 to 1721, providing a detailed exploration of the use of convolutional neural networks (CNNs) for the identification of diseases in maize plants. The citation follows the standard academic format used for technical and scientific papers, offering essential details such as the authors' names, the title of the study, the journal name, the volume and page numbers, and the publication date, all of which are necessary for proper reference and easy access to the original research.

[8]The research paper, titled "Performance of Deep Learning vs. Machine Learning in Plant Leaf Disease Detection," was authored by R. Sujatha, J. M. Chatterjee, N. Jhanjhi, and S. N. Brohi. It was published in the Microprocessors and Microsystems journal (Volume 80) in February 2021. This paper compares the effectiveness of deep learning and machine learning approaches in detecting plant leaf diseases, contributing to the ongoing research in the field of agricultural technology. The citation follows the standard academic format, providing essential details such as the authors' names, the paper's title, the journal name, volume, and the publication date, making it easier for researchers to locate and reference the work in scientific and technical literature.

[9]The research paper, titled "Identification of Plant Leaf Diseases by Deep Learning Based on Channel Attention and Channel Pruning," was authored by R. Chen, H. Qi, Y. Liang, and M. Yang. It was published in *Frontiers in Plant Science* (Volume 13) in November 2022. This paper explores the use of deep learning techniques, specifically channel attention and channel pruning, to identify plant leaf diseases, offering insights into advanced methods for improving disease detection. The citation provides key information such as the authors' names, the title of the paper, the journal name, the volume, and the publication date, following the standard academic citation format commonly used in scientific research for easy reference and access.

[10]The paper "Semi-supervised few-shot learning approach for plant diseases recognition" by Y. Li and X. Chao (published in Plant Methods, June 2021) presents a semi-supervised few-shot learning method for recognizing plant diseases. This approach addresses the challenge of limited labeled data, which is common in plant disease datasets, by leveraging both labeled sand unlabeled data for training. The model is designed to learn from a small number of labeled examples and a larger pool of unlabeled data, improving its ability to accurately identify plant diseases. This method enhances disease detection efficiency, making it practical for real-world agricultural applications with limited data.

3 Methodology

The block diagram of the proposed system is shown in Fig. 1. The proposed approach follows a step-by-step process, beginning with the collection of leaf and fruit image datasets. The images undergo pre-processing, which prepares them for further analysis.

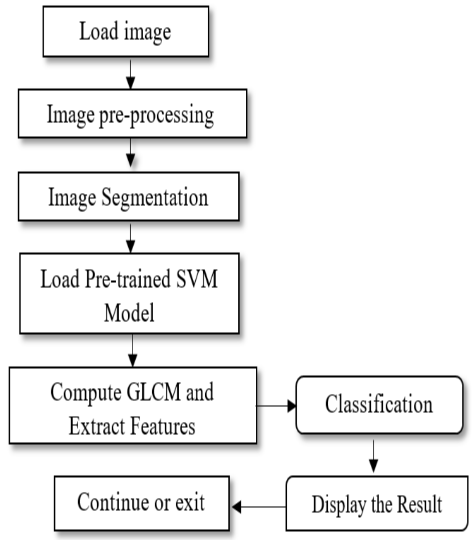

Fig.1 Framework of proposed system

3.1 Image selection

Image selection is a critical first step in digital image processing, involving the capture of images using a digital camera and storing them in a digital format such as JPEG or PNG. This process facilitates the transfer of the images from the camera hardware into software tools, such as MATLAB, for further analysis and processing. In our study, high-resolution images of both healthy and diseased leaves and fruits were taken with a digital camera to ensure that the images were clear and detailed, which is essential for accurate processing. These images then serve as the raw data input for the MATLAB image processing system, where advanced techniques like segmentation and classification are applied to assess plant health and identify any signs of disease. The success of these techniques is highly dependent on the quality of the captured images, making the image selection and acquisition process a vital starting point for effective analysis and interpretation in the study.

3.2 Image Pre-processing

The primary goal of image pre-processing is to improve the quality of the image by reducing unwanted distortions and enhancing specific features for more effective analysis. This step is crucial for preparing images that may contain noise or irrelevant details, ensuring that the subsequent processing stages are more accurate and efficient. In this study, several pre-processing techniques were applied using MATLAB code, starting with resizing the image to a consistent dimension, which is essential for uniformity in the processing pipeline. Contrast enhancement was also applied to improve the visibility of key details, making features like the disease-affected areas more prominent. Another important step involved converting the image from RGB (Red-Green-Blue) to grayscale, simplifying the image by focusing on intensity values rather than color, which facilitates tasks such as texture analysis or clustering. These pre-processing techniques, as help to create a cleaner and more consistent dataset, enabling more accurate segmentation and clustering for subsequent operations.

3.3 Image Segmentation

The image is first converted from the RGB color space to the Lab color space, which is a widely used model for color segmentation in image processing. The Lab color space is particularly advantageous because it separates the lightness information (L) from the chromatic components a and b, where L represents the brightness of the color, a and b indicate the color's position along the green-red and blue-yellow axes, respectively. This separation allows for more precise color-based segmentation, as it isolates color information from lightness variations. After the conversion, the a and b channels, which contain the chromatic data, are extracted and reshaped into a 2D array of pixel values. K-means clustering is then applied to this 2D data to divide the image into n color clusters, with n set to 3 in this case. The clustering process segments the image into distinct regions based on color characteristics, facilitating the identification of different parts of the image, such as the disease-affected areas.

3.4 Feature Extraction

Once the region containing the disease-affected portion of the leaf is selected, it is converted to a grayscale image if it is still in the RGB color format. This conversion simplifies the image by removing the color information and focusing solely on pixel intensity, which is essential for texture analysis. If the image is already in grayscale, no conversion is necessary, and it can be directly used for further analysis. The next step involves calculating the Gray Level Co-Occurrence Matrix (GLCM) for the grayscale image. The GLCM is a powerful tool used to capture the texture information in the image by examining the spatial relationships between pixel intensities. It works by measuring how frequently pairs of pixel with specific values occur in a defined spatial relationship, such as orientation and distance. From the GLCM, several texture features are extracted, such as **Contrast**, which highlights the difference in intensity between neighboring pixels; **Correlation**, indicating the degree of similarity between adjacent pixel values; **Energy**, which measures the uniformity of texture; and **Homogeneity**, reflecting the smoothness of the texture. These features are crucial for identifying and distinguishing the texture patterns characteristic of disease-affected regions in the leaf.

3.5 SVM Model Training and Disease Classification

To classify the disease type of the leaf in the test image, a pre-trained Support Vector Machine (SVM) model is first loaded from a file named Diseaseset.mat. This file contains a set of features that have been extracted from previously analyzed leaf images, along with their corresponding disease labels, which represent the known types of diseases affecting those leaves. The pre-trained SVM model has already been trained using this historical data, enabling it to learn the relationship between the extracted features and the associated disease labels. With this model in place, the classifier is ready to be used for predicting the disease type of the test leaf. The SVM classifier takes the features extracted from the current test leaf image as input and evaluates them based on the patterns learned during training. The result is a predicted disease label, which is the type of disease most likely affecting the test leaf based on its feature characteristics. This process allows for efficient and automated disease classification.

4 Experimental and Result

In this experiment, the primary objective was to accurately classify the type of disease affecting a leaf by leveraging a combination of advanced image processing techniques, feature extraction methods, and machine learning algorithms. The process began by acquiring digital images of the plant leaves, which were then processed to isolate and highlight the features associated with disease symptoms. This segmentation process helped in narrowing down the regions of interest that were relevant for further analysis.

Once the affected areas were segmented, texture features were extracted to quantify the visual patterns of the disease. By automating the process of disease classification, this method offers a reliable and efficient solution for plant health diagnosis, enabling early disease detection and helping to mitigate the impact of plant diseases in agricultural settings.

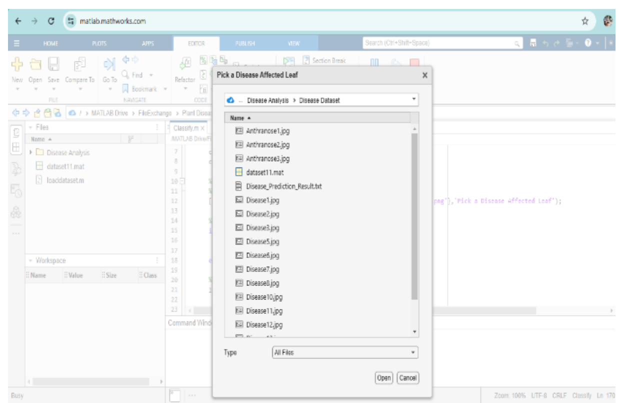

Fig 2: Data Upload

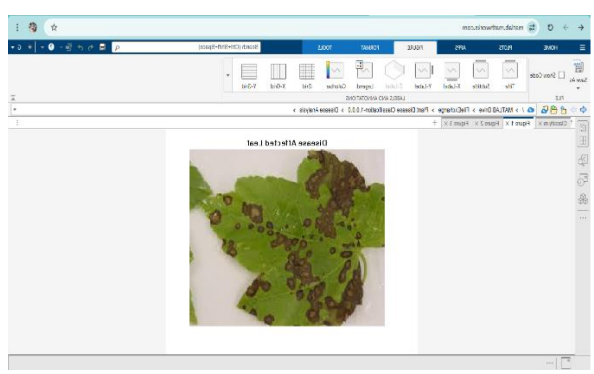

Fig 3: Sample image

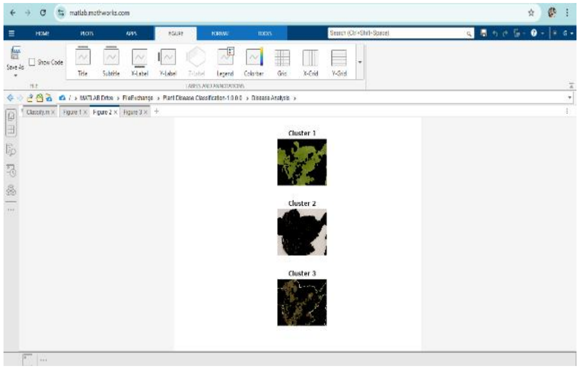

Fig 4: Three Classified Segmented Leaf

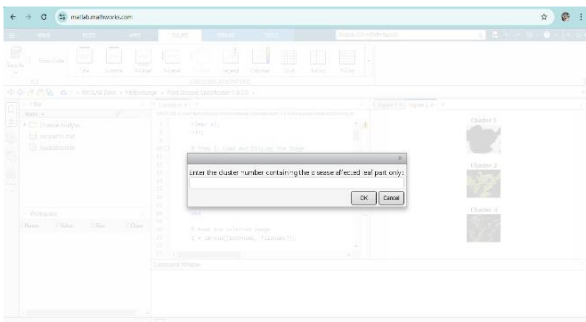

Fig 5: Select Affected Leaf Cluster

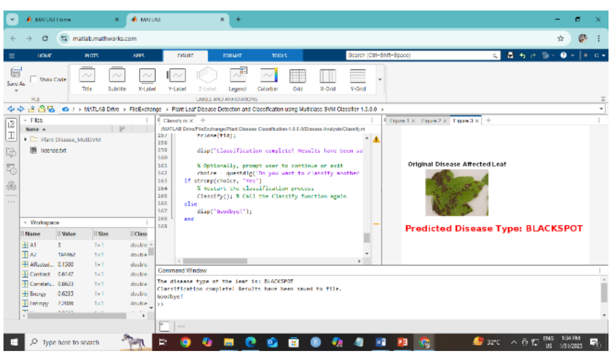

Fig 6: Final Output

5 Future Work

The disease classification method for identifying leaf diseases through image processing and machine learning provides a reliable solution, yet there are several areas where its effectiveness can be further enhanced to improve both accuracy and robustness in real-world applications. One of the key areas for improvement lies in the image pre-processing phase. While the current methods may handle basic noise and variations in image quality, more advanced techniques such as noise reduction, image sharpening, and enhanced segmentation could provide a significant performance boost. In particular, deep learning-based semantic segmentation methods, which segment images based on the meaning of pixels rather than just their colors or textures, could greatly improve the system's robustness. These methods would help to better handle challenging conditions, such as varying lighting, shadows, and complex backgrounds, which are frequently encountered in agricultural environments. The real-world agricultural settings often present images with fluctuating light conditions, motion blur, and non-uniform backgrounds, which can hinder the system’s performance. By incorporating these advanced pre-processing techniques, the system would be better able to enhance the quality of the input images, thus improving the extraction of relevant features for more accurate and reliable disease classification.

Another potential improvement area is expanding the dataset. The current dataset may be limited to a specific set of plant species and disease stages, which may result in reduced accuracy when the model encounters new, unseen leaf types or disease manifestations under different environmental conditions. To address this, the dataset could be expanded to include a broader variety of leaf images from different plant species, climates, and stages of infection. Incorporating these diverse datasets will enable the model to generalize more effectively and recognize a wider range of disease symptoms, improving its ability to handle the dynamic nature of plant diseases and their manifestations. Furthermore, advanced feature extraction methods, such as Local Binary Patterns (LBP) or Histogram of Oriented Gradients (HOG), could be employed to capture finer texture details on the leaf surface. These methods would enable the system to differentiate between subtle variations in diseases that might otherwise be overlooked.

Moreover, the integration of deep learning techniques, such as Convolutional Neural Networks (CNNs), could significantly enhance classification performance. CNNs have shown remarkable success in various image classification problems, and their application in plant disease detection could improve classification accuracy, especially when handling more complex or nuanced disease symptoms.

Incorporating real-time classification capabilities into mobile or web applications would also offer practical advantages. By enabling on-site disease detection, farmers and researchers could quickly identify and address plant health issues as they arise. Implementing real-time detection would allow for faster intervention and more proactive disease management. Additionally, optimizing the model with techniques such as Grid Search or Random Search to fine-tune hyper parameters could lead to a notable increase in classification accuracy, as these techniques help identify the optimal configuration for the machine learning model.

Finally, to ensure that the model's improvements are consistent and reliable, it would be essential to evaluate its performance using cross-validation and key metrics such as precision, recall, and F1 score. These metrics provide a thorough understanding of the model's robustness, enabling researchers to gauge its generalization to new, unseen leaf images and confirming its reliability in practical applications. This comprehensive evaluation approach would ensure that the model is not only highly accurate but also capable of making informed predictions in diverse, real-world scenarios.

6 Conclusion

This paper presents a straightforward yet highly effective approach for detecting and classifying plant diseases, making use of MATLAB's advanced image processing capabilities. The proposed method combines K-means clustering with Multi-SVM (Support Vector Machine) techniques, which work together to identify and classify plant diseases in both leaves and fruits. MATLAB, with its powerful image processing toolbox, offers an ideal platform for implementing such a system, providing the necessary tools for efficient disease detection and classification. The integration of K-means clustering, a popular unsupervised learning algorithm, and Multi-SVM, a robust supervised machine learning technique, ensures that the system delivers fast, accurate, and reliable results, making it particularly suitable for real-world agricultural applications.

One of the core advantages of this approach is its ability to detect plant diseases at their early stages, allowing farmers to take proactive measures before the disease spreads to critical parts of the plant. Early detection can significantly reduce crop losses, ultimately helping to safeguard the plant's health and yield. The system's efficiency and accuracy make it particularly useful in preventing the spread of diseases, ensuring that farmers can respond quickly with the right interventions. The automated nature of the system also alleviates the need for expert intervention, which is particularly valuable in remote areas where access to plant pathologists and other experts might be limited.

The method has been designed with user-friendliness in mind, making it accessible to farmers who may not have in-depth knowledge of plant diseases. By focusing on the key visual symptoms of diseases in plants, the system allows individuals with limited expertise to easily identify and diagnose common diseases affecting crops. The disease detection process in this system follows a well-defined procedure that includes steps like image acquisition, pre-processing (for enhancing image quality), segmentation (to isolate areas of interest), feature extraction (to capture relevant disease characteristics), and classification (for identifying the type of disease).

References

[1] A. Fuentes, S. Yoon, and D. S. Park, *Deep Learning-Based Techniques for Plant Diseases Recognition in Real-Field Scenarios,* 2020.

[2] S. M. M. Hossain, K. Deb, P. K. Dhar, and T. Koshiba, *Plant leaf disease recognition using depth-wise separable convolution-based models*, Symmetry, vol. 13, no. 3, pp. 511–529, Mar. 2021.

[3] M. A. Chandra and S. S. Bedi, *Survey on SVM and their application in image classification*, Int. J. Inf. Technol., vol. 13, no. 5, pp. 1–11, Oct. 2021.

[4] S. Garg and P. Singh, *An aggregated loss function based lightweight few shot model for plant leaf disease classification,* Multimedia Tools Appl., vol. 82, no. 15, pp. 23797–23815, Feb. 2023.

[5] M. M. F. Azizi, N. H. Mardhiah, and H. Y. Lau, *Current and emerging molecular technologies for the diagnosis of plant diseases an overview*, J. Experiment. Biol. Agriculture. Sci., vol. 10, no. 2, pp. 294–305, Apr. 2022.

[6] S. Wang, Y. Han, J. Chen, X. He, Z. Zhang, X. Liu, and K. Zhang, *Weed density extraction based on few-shot learning through UAV remote sensing RGB and multispectral images in ecological irrigation area*, Frontiers Plant Sci., vol. 12, Mar. 2022.

[7] S. Jasrotia, J. Yadav, N. Rajpal, M. Arora, and J. Chaudhary, *Convolutional neural network based maize plant disease identification*, Proc. Computer. Sci., vol. 218, pp. 1712–1721, Jan. 2023.

[8] R. Sujatha, J. M. Chatterjee, N.Jhanjhi, and S. N. Brohi, *Performance of deep learning vs machine learning in plant leaf disease detection*, Microprocessor. Microsystem, vol. 80, Feb. 2021.

[9] R. Chen, H. Qi, Y. Liang, and M. Yang, *Identification of plant leaf diseases by deep learning based on channel attention and channel pruning,* Frontiers Plant Sci., vol. 13, Nov. 2022.

[10] Y. Li and X. Chao, *Semi-supervised few-shot learning approach for plant diseases recognition*, Plant Methods, vol. 17, no. 1, pp. 1–10, Jun. 2021.