



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 6, June 2025

Automated Leaf Disease Identification via Image Processing and Multi-Class Support Vector Machine

Arshdeep Singh¹, Meenakshi Arora², Rohini Sharma³

P.G. Student, Department of CSE, Sat Kabir Institute of Technology and Management, Ladrawan, Haryana, India¹

Assistant Professor, Department of CSE, Sat Kabir Institute of Technology and Management, Ladrawan, Haryana, India²

Assistant Professor, Department of CS, GPGCW, Rohtak, Haryana, India³

ABSTRACT: Plant diseases affect crop output and food security, making them a serious danger to the global agricultural sector. For prompt intervention and efficient management, leaf diseases must be identified early and accurately. This study offers a thorough and automated method for detecting leaf illness that makes use of Multi-Class Support Vector Machine (SVM) classifier and image processing techniques. In order to increase visual quality, the suggested methodology uses a nine-stage pipeline that begins with image acquisition and ends with image augmentation. To isolate impacted areas, diseased regions are divided into segments, and the ROI is chosen. These regions are converted to grayscale to simplify further processing, and the area of infection is calculated for quantification. Key features such as color, shape, and texture are extracted from the diseased regions and used to train a multi-Class SVM, which classifies the leaf image into specific disease categories. The final output provides accurate classification results, aiding in efficient disease diagnosis. Experimental results demonstrate the effectiveness of the proposed system in identifying multiple leaf diseases with high precision, making it suitable for practical agricultural applications.

KEYWORDS: Plant Pathology, Automated Diagnosis, Precision Agriculture

I. INTRODUCTION

Many economies rely heavily on agriculture, particularly in developing nations where a sizable section of the populace makes their living from farming. The early and precise identification of plant diseases, which can significantly lower crop yield and quality if left untreated, is one of the biggest problems facing the agriculture industry. One of the most prevalent signs of plant health problems is leaf disease, and outward signs such as spots, yellowing, and wilting may indicate underlying bacterial, viral, or fungal infections [1]. Disease identification has historically depended on agricultural specialists performing manual inspections, which is labor-intensive, time-consuming, and frequently prone to human mistake. The need for automated, dependable, and scalable technologies that can help with early disease diagnosis and lower crop losses is therefore increasing. Machine learning and image processing methods have become more effective tools for automating the disease detection process in recent years [2].

This study introduces an automated approach for identifying leaf diseases that combines a Multi-Class Support Vector Machine (SVM) classifier with an organized image processing pipeline. The process of the suggested approach is as follows: segmentation and ROI extraction, grayscale conversion, area calculation of infected regions, feature extraction, and picture acquisition and enhancement come first. After these attributes are loaded into a multi-Class SVM, the input leaf picture is highly accurately classified into predetermined illness categories. The choice of SVM is motivated by its robustness in handling high-dimensional data and its effectiveness in multi-class classification problems in plant disease detection scenarios [3]. By integrating traditional image processing with machine learning, the proposed system aims to offer a practical and cost-effective solution that can be deployed in real agricultural environments for early disease detection and crop monitoring.

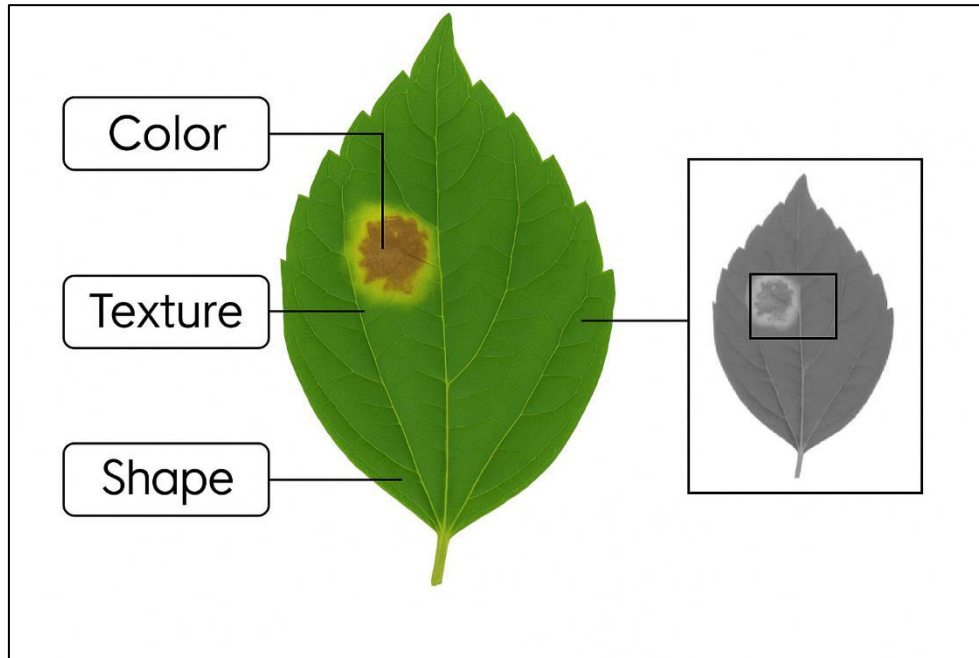


Figure 1: Color, Texture and Shape of a disease affected Leaf

Research Background: Thanks to developments in machine learning (ML) and image processing, automated plant disease diagnosis has become increasingly popular in recent years. In addition to being time-consuming, traditional disease diagnosis techniques—which mostly involve visual inspection by farmers or agronomists—lack impartiality and consistency, especially in large-scale farming. These restrictions have prompted studies into automated and digital alternatives that can improve diagnosis speed and accuracy. Primitive image processing methods including color-based thresholding and morphological operations for identifying afflicted areas in plant leaves were the main focus of early research in this field [4]. Although these techniques were straightforward and simple to use, they were frequently constrained by external elements including background noise, illumination, and differences in the texture and color of the leaves. Researchers started combining ML techniques with image processing to get over these obstacles. Color histograms, Gray Level Co-occurrence Matrix (GLCM), and texture analysis are examples of feature extraction approaches that have become commonplace for depicting the visual traits of sick areas [5]. To increase classification accuracy, these features were subsequently combined with classifiers such as Support Vector Machines (SVM), Decision Trees, and k-Nearest Neighbors (k-NN). SVM is one of them that has gained a lot of traction because of its excellent performance, particularly in high-dimensional feature spaces, and good generalization capabilities [6].

Deep learning methods, especially Convolutional Neural Networks (CNNs), have advanced recently and are capable of automatically learning characteristics from unprocessed picture data without the need for human involvement [7]. However, deep learning models sometimes call for substantial computational resources and extensive datasets, which might not be possible in settings with limited resources, like rural farms. The combination of lightweight machine learning models such as multi-Class SVM with conventional image processing makes hybrid systems extremely useful. The methodology proposed in this study builds upon this foundation by combining classical image processing steps with a multi-Class SVM classifier. This approach offers a balance between computational efficiency and diagnostic accuracy, making it suitable for real-time applications in agricultural settings. It addresses critical stages such as region segmentation, grayscale transformation, and precise area calculation, which are vital for reliable classification outcomes.

Researchers have investigated hybrid techniques, which integrate several steps of image analysis to increase accuracy and robustness, in addition to conventional machine learning approaches. Before feature extraction, for instance, sick regions are separated from healthy tissue using segmentation techniques including K-means clustering, Otsu thresholding, and watershed algorithms. The efficiency of later feature extraction and classification steps is greatly

impacted by accurate segmentation, especially in heterogeneous leaf images with complicated surroundings or overlapping symptoms [8]. Additionally crucial preprocessing methods that simplify the data and lessen computing load without sacrificing diagnostic detail are grayscale conversion and ROI selection. Research has demonstrated that by standardizing input image features and removing superfluous background information, this type of preprocessing enhances classification performance [9]. Despite being frequently disregarded, the area calculation of infected zones offers vital quantitative data regarding the severity of the illness, which can be helpful for crop management and decision-making in addition to classification. Numerous studies have looked into improving disease grading systems by using severity metrics [10].

When trained on robust feature sets, multi-class SVM classifiers have shown high classification accuracy for a variety of leaf diseases, including rust, bacterial spot, and early blight. Multi-Class SVM effectively solves multi-category classification issues by employing strategies like one-vs-one or one-vs-rest, in contrast to binary SVM, which only manages two classes at a time. Because of this, it is ideal for detecting several disease kinds in one crop or among various crops [11]. Although CNNs and other deep learning models have dominated recent research because of their automatic feature learning capabilities, these models frequently necessitate sophisticated hyperparameter tweaking, massive labeled datasets, and GPU acceleration. For users with limited resources, such smallholder farmers, this restricts their viability. On the other hand, conventional models such as multi-Class SVM provide a lightweight and interpretable substitute that can attain competitive performance with substantially lower overhead when paired with careful preprocessing and feature engineering [12].

II. PROPOSED METHODOLOGY

Figure 2 explain the complete methodology of detecting leaf disease and further classify it into different types of diseases.

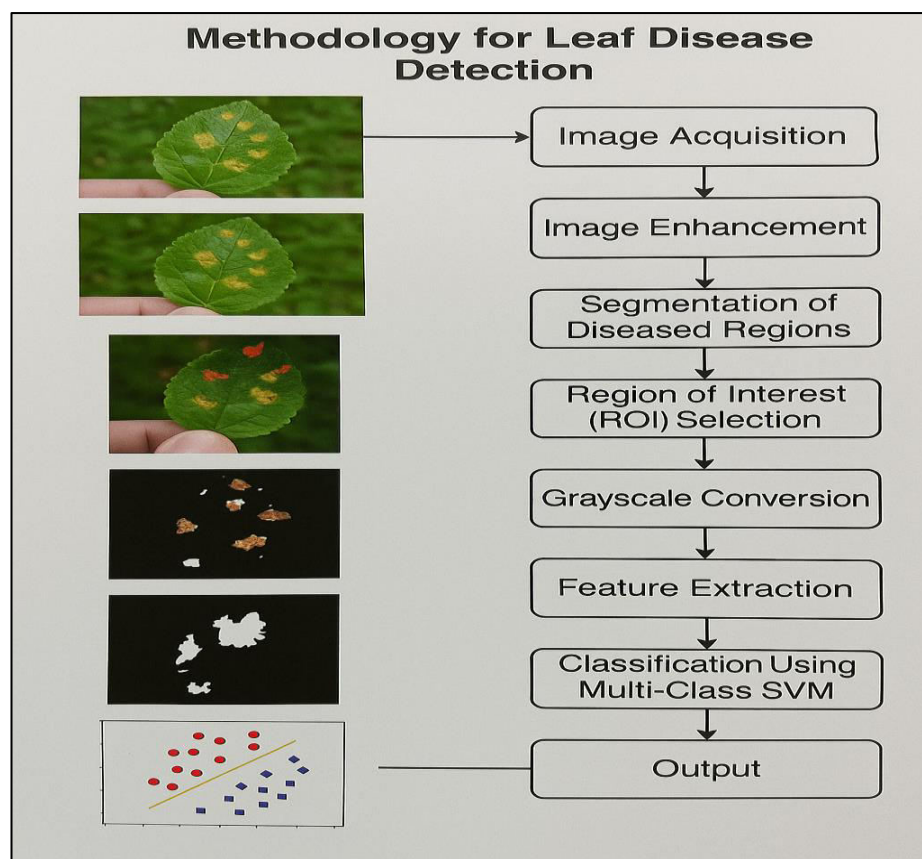


Figure 2: Methodology for Leaf Disease Detection

1. Image Acquisition

This is the first and foundational step where images of plant leaves are captured using digital cameras, smartphones, or specialized imaging devices under various lighting and environmental conditions. The quality, resolution, and clarity of these images directly affect the accuracy of downstream processing. The images should clearly show both healthy and potentially diseased regions.

2. Image Enhancement

Image enhancement techniques are applied to improve the visual quality and highlight critical features of the leaf image. This includes: Contrast adjustment, Histogram equalization, Noise removal using filters (e.g., Gaussian, Median) and color normalization.

3. Segmentation of Diseased Regions

Segmentation separates diseased regions from the rest of the image. Techniques such as: Color thresholding, K-means clustering, Otsu's method and edge detection are often used. This step isolates infected areas by identifying color, texture, or intensity variations, ensuring only disease-relevant areas are passed on for analysis.

4. Region of Interest (ROI) Selection

ROI selection focuses analysis only on the segmented diseased areas, ignoring irrelevant parts of the image like background, healthy leaf regions, or stem. This helps reduce computational overhead and enhances the specificity of feature extraction. The ROI typically contains the most prominent signs of disease (spots, lesions, mildew, etc.).

5. Grayscale Conversion

To simplify further processing and reduce computational complexity, the ROI is converted from RGB (color) to grayscale. Grayscale images represent only intensity values (0–255), removing color channels and focusing on texture and shape details which are essential for disease analysis.

6. Area Calculation

Area calculation quantifies the extent of the infected region within the ROI. This involves counting the number of pixels classified as diseased and may be used to estimate the **severity** of the disease. It can assist in disease grading (e.g., mild, moderate, severe) and inform treatment decisions.

7. Feature Extraction

From the ROI, various features are extracted to describe the disease characteristics. Common feature categories include: Color features, Texture features (contrast, homogeneity using GLCM), Shape features (area, perimeter, circularity). These features create a numerical representation of the image, forming the input to the classifier.

8. Classification Using Multi-Class SVM

The extracted features are fed into a Multi-Class Support Vector Machine (SVM) classifier. SVM constructs decision boundaries between different disease categories (e.g., Leaf Spot, Blight, Rust, etc.). Multi-class variants such as one-vs-one or one-vs-rest strategies allow SVM to handle more than two disease types. It assigns the leaf image to the most likely disease class based on learned patterns.

9. Output

The system generates the final result, typically in the form of: Disease name or label and Severity level (based on area)

III. SIMULATIONS AND OUTCOMES

We have implemented the proposed method in MATLAB 2024a. We have prepared a DATASET of numerous leaves. While some leaves are healthy, some are affected by some diseases. We have classified the diseases into following categories: Alternaria Alternata Leaf Disease, Anthracnose, Bacterial Blight and Cercospora leaf spot. Our method uses K means clustering to decide the region of interest and Multi SVM to classify a leaf image in to one of the four categories as mentioned above.

K-means clustering is a popular unsupervised machine learning algorithm used in image segmentation to group similar pixels into clusters based on their features (like color or intensity). For leaf disease detection, it's particularly useful in separating diseased regions from healthy areas without requiring labeled training data. The input image (usually in RGB format) is first reshaped into a 2D array where each pixel is treated as a data point. Each pixel is represented as a feature vector — typically based on color values (R, G, B) or other properties like intensity. We specified the number of clusters $K = 3$ for leaf disease segmentation. One cluster for healthy regions and two clusters for diseased areas. K-means randomly initializes K cluster centroids. Each pixel is assigned to the nearest centroid (based on Euclidean distance). Centroids are updated as the mean of all pixels assigned to each cluster. Steps repeat until convergence (no major change in cluster assignments). After convergence, each pixel belongs to a cluster. Clusters corresponding to diseased regions (typically brown, yellow, or black in color) are isolated based on color statistics or manual identification. The result is a segmented mask highlighting the infected portions of the leaf. A leaf image with green healthy areas and brown spots (disease). For $K = 3$ (green, brown, background), after clustering, we assign the cluster with brown pixels as "diseased region" and create a binary mask to isolate it.

Area calculation is the process of quantifying the extent of infection on a leaf by measuring the number of pixels (or region size) that represent the diseased portion. It plays a critical role in assessing disease severity, helping in grading infection levels (e.g., mild, moderate, severe) and guiding treatment decisions. After segmentation (via K-means), a binary mask is generated: Pixels in the diseased region = white (1) and all other pixels = black (0). Each white pixel in the binary mask is counted. The total number of white pixels = infected area in pixel units. To get the proportion of infected area relative to the entire leaf: Disease Ratio = Infected Pixels/ Total leaf Pixels as shown in figure 3.

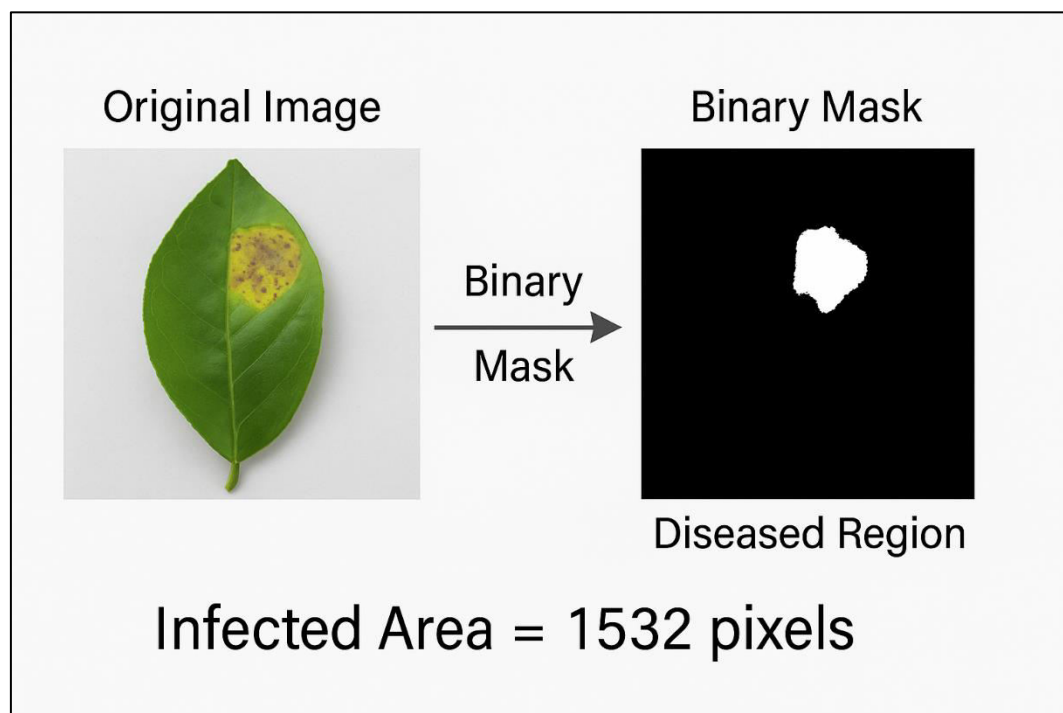


Figure 3: Area calculation

Support Vector Machine (SVM) is a powerful supervised machine learning algorithm widely used for classification problems. In leaf disease detection, Multi-Class SVM is used to categorize the detected leaf region into one of several disease types (e.g., leaf spot, blight, rust) based on extracted features. Standard SVM is a binary classifier, but for leaf disease detection, we often need to classify more than two diseases. To achieve this, we extend SVM using: One-vs-Rest (OvR) which trains N classifiers for N classes. Each one separates one class vs all others. The classifier with the highest confidence score determines the class. After preprocessing, segmentation, and feature extraction, a feature vector is generated (e.g., color, texture, shape features). The SVM model is trained using labeled examples of each

disease. The SVM learns optimal decision boundaries that separate the feature space into regions corresponding to each disease class. The feature vector from a new leaf image is input into the trained SVM. The SVM classifier predicts the disease category based on which side of the decision boundary the feature falls into.

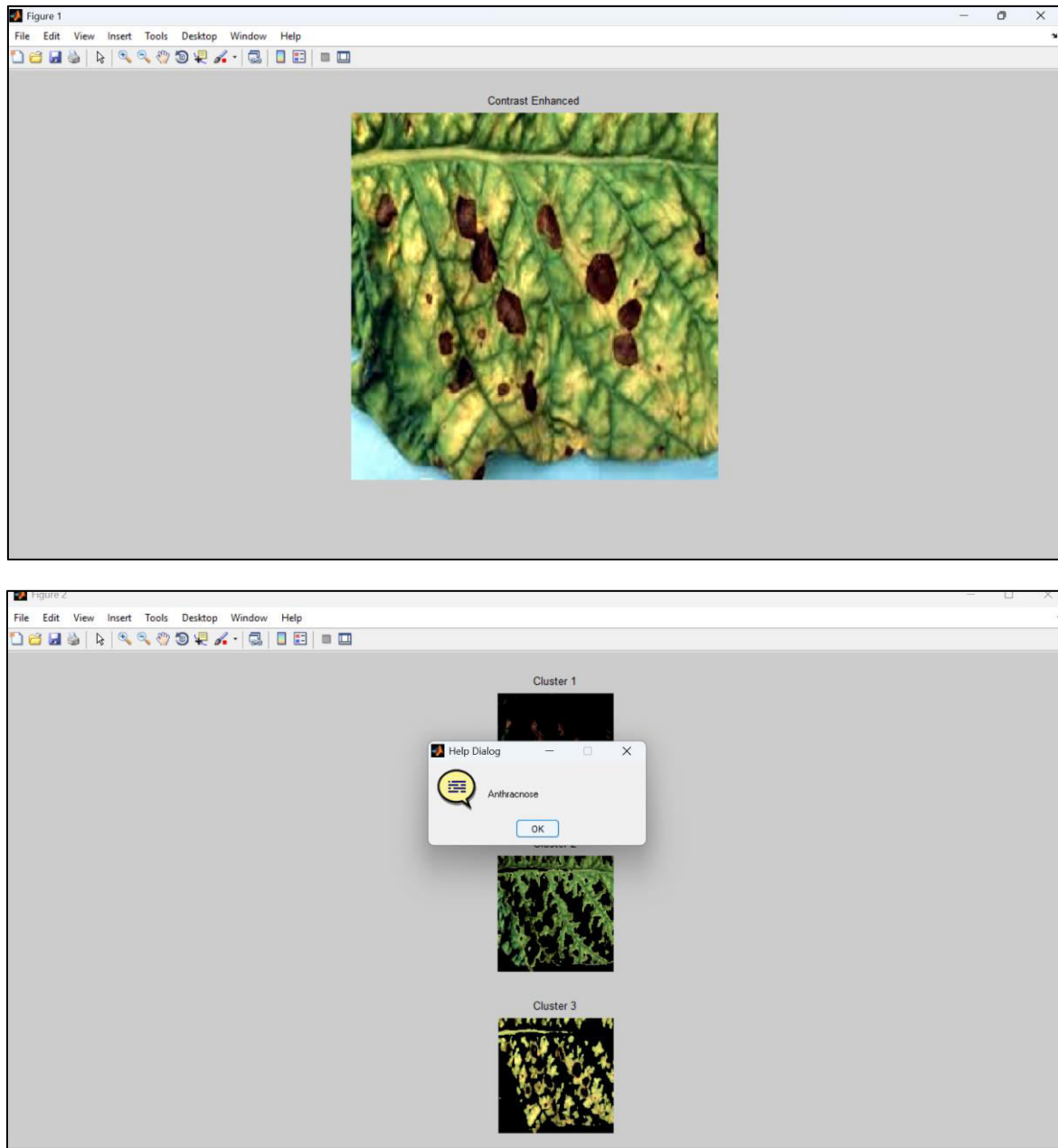


Figure 4a: Input Image of a Leaf, 4b: Detected Disease (Anthracnose)

Figure 4 identifies a disease named Anthracnose. The main features of this disease are as follows: Small, dark, water-soaked spots; often with yellow or brown margins. So, our simulation and method are able to correctly identify this disease. Affected Area is: 15.3035%. the leaf in figure 5 has 19.123% affected area.

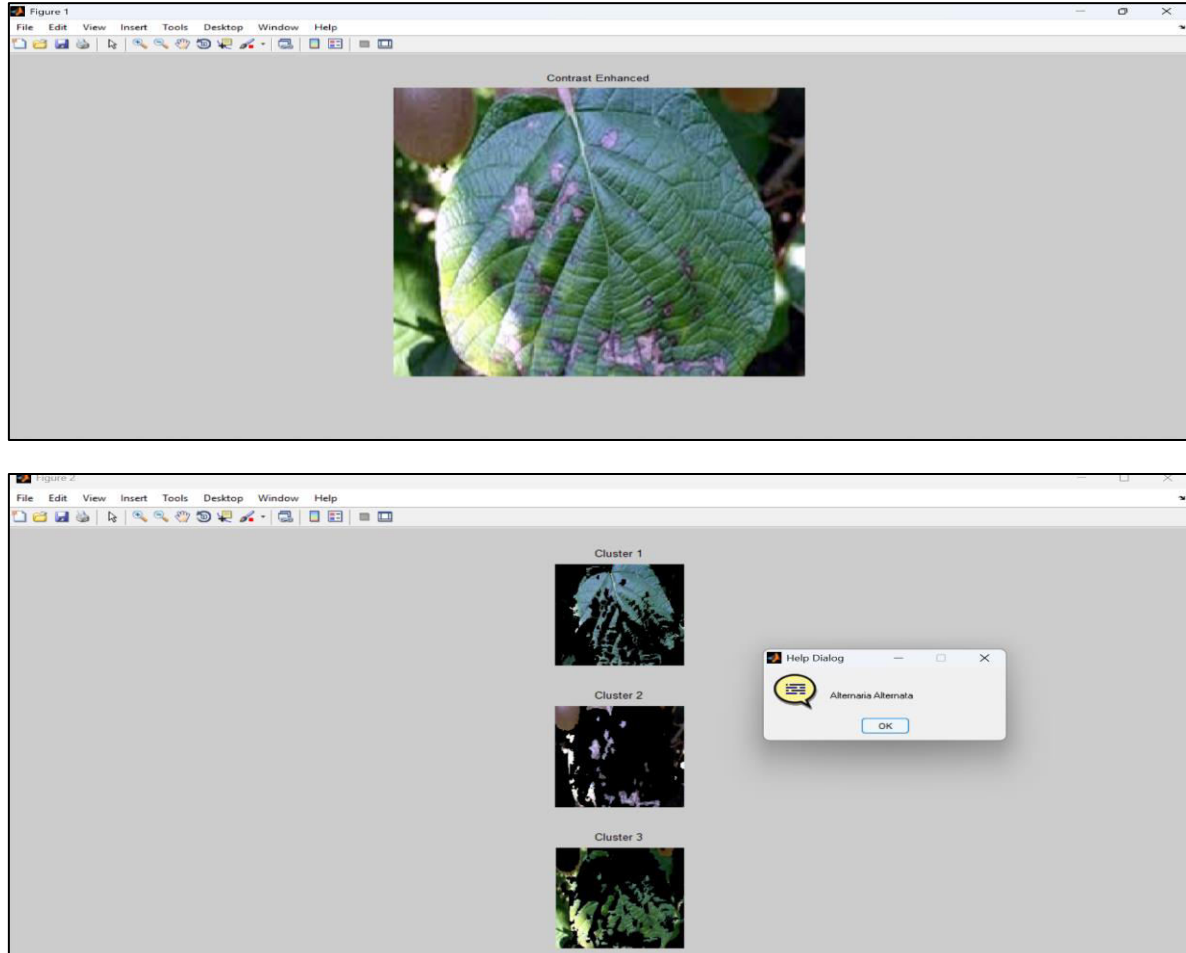
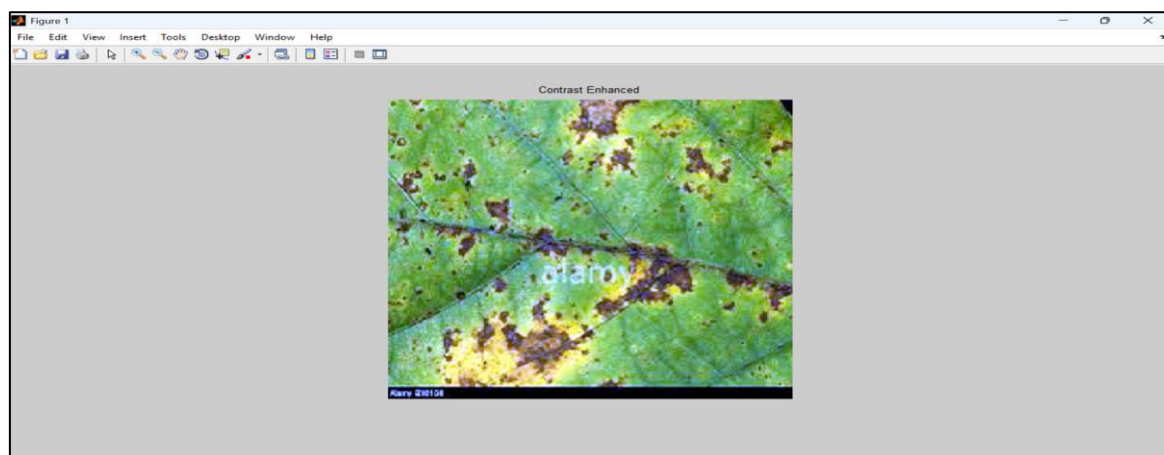


Figure 5: Identified Disease is Alternaria Alternata



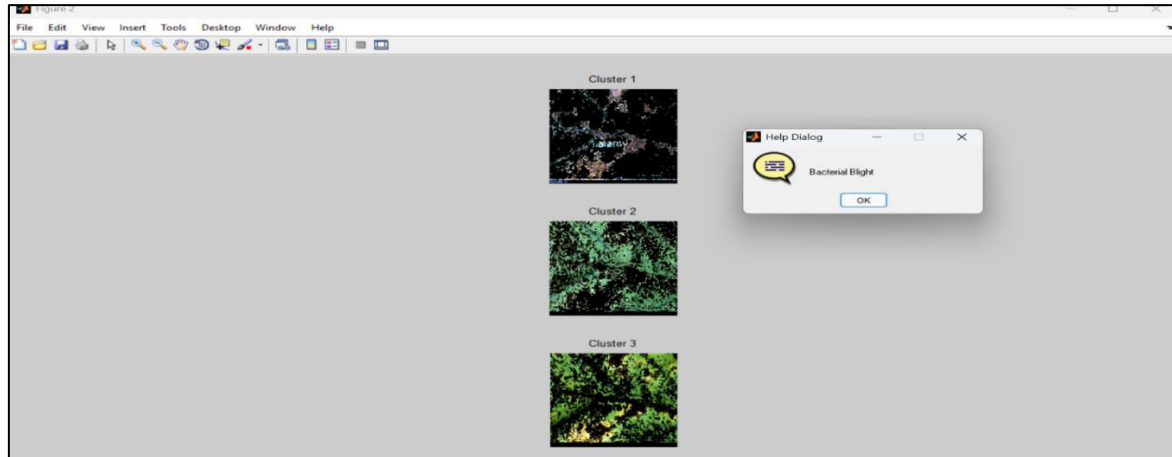


Figure 6: Identified Disease is Bacterial Blight, affected area = 15.0787%

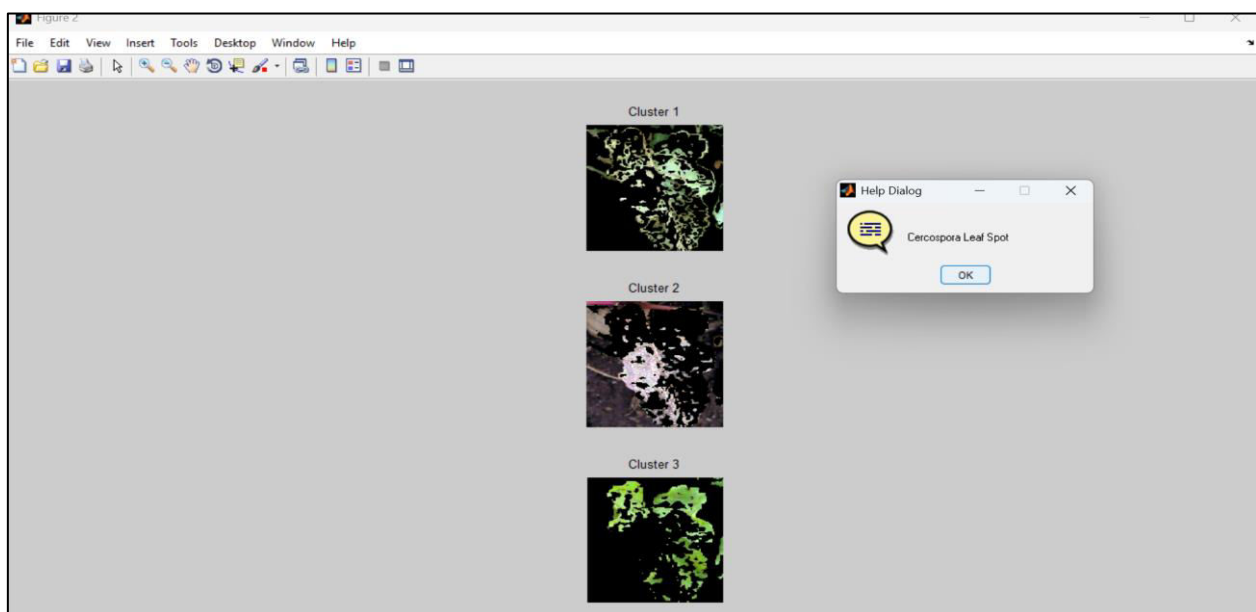
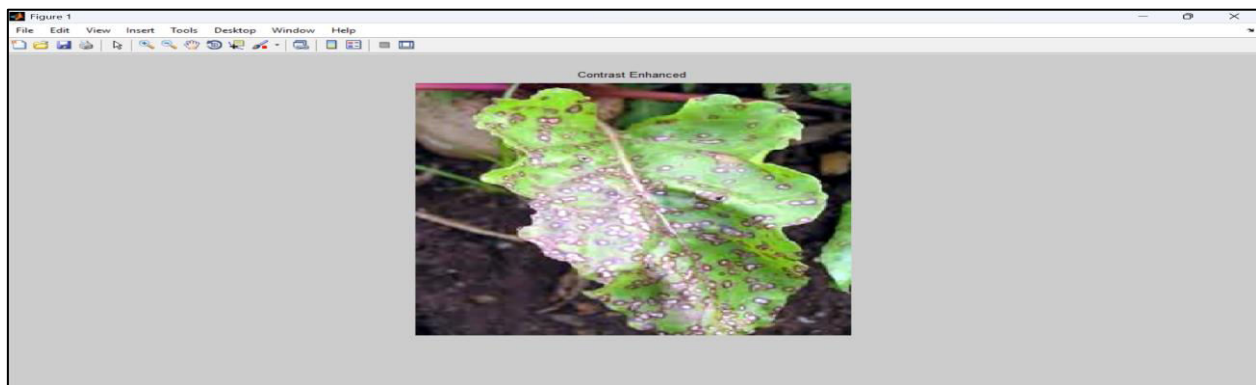


Figure 7: Identified Disease is Cercospora leaf spot, affected area = 17.087%

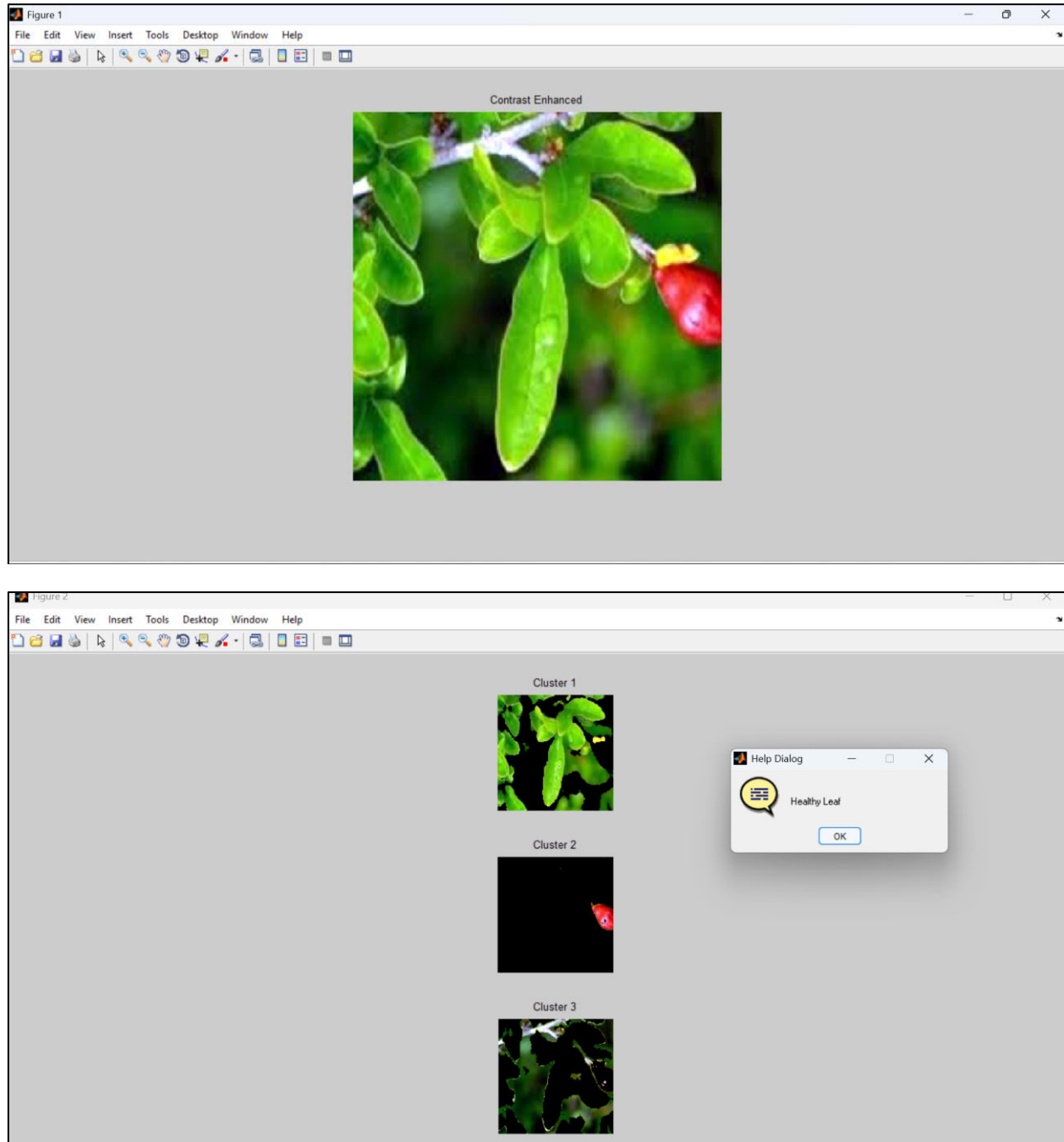


Figure 8: Healthy leaf

IV. CONCLUSION

In this study, we have successfully developed and implemented an automated leaf disease detection system using a structured image processing pipeline integrated with a Multi-Class Support Vector Machine (SVM) classifier. The proposed method efficiently performs multiple stages—including image enhancement, segmentation, ROI extraction, grayscale conversion, area calculation, and feature extraction—to accurately identify both the presence and extent of leaf disease. The classification model effectively categorizes leaf samples into four distinct disease types: Anthracnose,

Alternaria Alternata, Bacterial Blight, and Cercospora Leaf Spot, while also accurately identifying healthy leaves. Furthermore, the system quantifies the infected region, providing valuable information about the severity of infection, which can support disease monitoring and timely intervention. The results demonstrate the robustness and reliability of the approach, making it suitable for real-time deployment in precision agriculture. By offering accurate diagnosis and infection analysis, the system can serve as a practical tool for farmers, agronomists, and agricultural researchers aiming to enhance crop health management and productivity.

REFERENCES

1. Mahlein, A.-K. (2016). Plant disease detection by imaging sensors – Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, 100(2), 241–251.
2. Barbedo, J.G.A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus*, 2(1), 660.
3. Phadikar, S., Sil, J., & Paul, M. (2013). Rice diseases classification using feature selection and rule generation techniques. *Computers and Electronics in Agriculture*, 90, 76–85.
4. Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., & ALRahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications*, 17(1), 31–38.
5. Arivazhagan, S., Newlin Shebiah, R., Ananthi, S., & Vishnu Varthini, S. (2013). Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agricultural Engineering International: CIGR Journal*, 15(1), 211–217.
6. Camargo, A., & Smith, J. S. (2009). An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosystems Engineering*, 102(1), 9–21.
7. Mohanty, S.P., Hughes, D.P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419.
8. Pujari, J. D., Yakkundimath, R., & Byadgi, A. S. (2016). Image processing-based detection of fungal diseases in plants. *Procedia Computer Science*, 46, 1802–1808.
9. Patil, S. B., & Kumar, R. (2011). Feature extraction of diseased leaf images. *International Journal of Advanced Computer Science and Applications*, 2(6), 130–134.
10. Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, 72(1), 1–13.
11. Zhang, S., Wu, X., & You, Z. (2017). Leaf image based cucumber disease recognition using sparse representation classification. *Computers and Electronics in Agriculture*, 134, 135–141.
12. Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*, 74(1), 91–99.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details