Image Processing-Based Detection of Pomegranate Leaf Diseases Using K-Means Clustering and SVM

Vanita S. Lokare[1], Dr. P.P. Belagali[2]

[1] P. G. Student / Department of Electronics and Telecommunication Engineering / Dr. J.J.M.C.O.E. Jaysingpur, Shivaji University, Kolhapur, India
[2] Associate Professor/ Department of Electronics and Telecommunication Engineering / Dr. J.J.M.C.O.E. Jaysingpur, Shivaji University, Kolhapur, India

Corresponding Author Email: lokarevanita91@gmail.com

How to Cite this Article:

Lokare, V. S. (2026). Image Processing-Based Detection of Pomegranate Leaf Diseases Using K-Means Clustering and SVM. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(05). https://doi.org/10.55041/ijcope.v2i5.746

https://doi.org/10.55041/ijcope.v2i5.746

Abstract—
Pomegranate is an economically and nutritionally significant fruit crop whose yield and quality are severely affected by diseases such as bacterial blight, anthracnose, fruit rot, and leaf spot. Early and accurate detection of these diseases is critical to minimizing crop losses and ensuring sustainable agricultural productivity. Conventional manual inspection methods are time-consuming and prone to error, highlighting the need for an automated solution. This study presents an automated pomegranate disease detection system based on image processing and machine learning. The proposed system integrates K-means clustering-based segmentation for isolating infected regions, followed by the extraction of color, texture, and shape features, and classification using a Support Vector Machine (SVM). The system was trained and validated on a dataset of pomegranate fruit and leaf images, achieving high detection accuracy in the early stages of disease development. The results demonstrate that the proposed approach provides an efficient, reliable, and farmer-friendly tool for disease identification, thereby reducing agricultural losses and advancing smart, precision farming practices. The proposed system has significant practical value, particularly for smallholder farmers who lack access to expert agricultural guidance, as it offers a cost-effective, easy-to-use alternative to manual scouting. In future work, the system can be further enhanced by incorporating deep learning architectures such as Convolutional Neural Networks (CNNs) to improve detection accuracy and robustness. Additionally, deploying the technology as a mobile application would enable real-time, field-level disease diagnosis, making it more accessible to farming communities. The framework can also be extended to detect diseases in other fruit and vegetable crops, broadening its applicability and contributing to large-scale precision agriculture.

Keywords— Pomegranate; Image processing; SVM; Bacterial blight; leaf spot.

I. INTRODUCTION

Pomegranate (*Punica granatum* L.) is one of the oldest and most economically important fruit crops cultivated across tropical and subtropical regions of the world, including India, Iran, Spain, Turkey, and China [1]. It is widely recognized for its exceptional nutritional value, being rich in antioxidants, vitamins, minerals, and polyphenolic compounds that offer significant health benefits [2]. In recent years, the global demand for pomegranate and its derived products, such as juice, powder, and extracts, has increased substantially, making it a crop of considerable commercial significance [3]. In India, pomegranate is predominantly cultivated in the states of Maharashtra, Karnataka, Gujarat, and Andhra Pradesh, contributing significantly to the horticultural economy of the country [4].

Despite its economic and nutritional importance, pomegranate cultivation is severely threatened by a range of fungal, bacterial, and environmental diseases that affect both the quality and quantity of the yield. Among the most prevalent and damaging diseases are bacterial blight (*Xanthomonas axonopodis* pv. *punicae*), anthracnose (*Colletotrichum gloeosporioides*), fruit rot (*Botrytis cinerea*), and leaf spot (*Cercospora punicae*) [5, 6]. These diseases manifest as visible symptoms on leaves, fruit skin, and internal plant tissue, and, if left undetected in their early stages, can result in significant crop losses ranging from 30% to 80%, depending on the severity of the infection and prevailing environmental conditions [7]. The economic losses incurred by farmers due to such diseases not only affect their livelihoods but also pose a challenge to food security and sustainable agricultural production [8].

Early and accurate detection of plant diseases is therefore a critical requirement in modern agriculture. Timely identification enables farmers to take appropriate preventive and corrective measures, such as targeted application of fungicides or bactericides, thereby minimizing damage and reducing the indiscriminate use of chemicals [9]. Conventionally, disease detection in pomegranate and other fruit crops has relied on visual inspection carried out by farmers or agricultural experts. However, this approach is inherently subjective, time-consuming, and dependent on the expertise and experience of the individual performing the inspection [10]. In many rural farming communities, access to trained agricultural experts is limited, making it difficult to detect and manage diseases at an early stage. These limitations highlight the pressing need for an automated, accurate, and cost-effective disease detection system that can be easily adopted at the farm level.

With the rapid advancement of digital imaging technology, image processing has emerged as a powerful tool for the non-destructive analysis of plant health. Image processing techniques enable the extraction of meaningful visual information from plant images, such as color changes, textural abnormalities, and morphological variations associated with disease symptoms [11]. Combined with machine learning algorithms, these techniques have demonstrated remarkable potential for the automated identification and classification of plant diseases with high accuracy [12]. Several studies have explored the application of image processing and machine learning for disease detection in crops such as tomato, grape, apple, rice, and wheat, yielding promising results [13, 14]. However, research specifically focused on pomegranate disease detection using such approaches remains relatively limited, indicating a significant research gap in this domain [15].

Among the various segmentation techniques available, K-means clustering has been widely adopted for plant disease detection due to its simplicity and effectiveness in separating diseased regions from healthy tissue based on color and intensity variations [16]. Similarly, the Support Vector Machine (SVM) classifier has been extensively used in plant disease classification owing to its superior performance in high-dimensional feature spaces, robustness against overfitting, and ability to handle both linear and non-linear classification problems [17]. The combination of K-means clustering for segmentation and SVM for classification has shown considerable promise in achieving accurate and reliable disease detection in agricultural applications [18].

In view of the above, this study proposes an automated system for the early detection of pomegranate diseases using image processing and machine learning techniques. The proposed system employs a systematic pipeline comprising image preprocessing, K-means clustering-based segmentation, multi-feature extraction encompassing color, texture, and shape descriptors, and SVM-based classification. The primary objectives of this work are to develop a reliable and accurate disease detection framework suitable for pomegranate fruit and leaf images, to evaluate its performance under different disease conditions, and to provide a practical and accessible tool that can assist farmers in identifying diseases at an early stage.

II. LITERATURE REVIEW

Kale and Shitole et al. [19] presented an automated system for early detection of crop diseases. The study uses image processing techniques for disease identification and classification. It helps in improving accuracy and supports timely diagnosis of plant diseases. To replace the expensive and time-consuming process of manually checking for diseases, the authors suggest a multi-stage process for the automated system. These stages include the acquisition of images, pre-processing to reduce noise, and clustering by using the K-means algorithm. In this research paper, the authors use a comparative analysis of three different models Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest to classify the health concerns related to apple crops such as black rot and cedar rust. The authors conclude that the Multivariate SVM algorithm offers the best results for accuracy to detect diseases.

The researchers in [20] were able to use image processing to transform images of infected leaves into digital form, specifically focusing on the region of interest with irrelevant background information removed. The process involves K-means clustering to accurately segment the images, with Gray Level Co-occurrence Matrix (GLCM) to extract vital features from the images. After this process, the images are then processed by a Convolutional Neural Network (CNN) to determine the type of disease affecting plants like pomegranates and tobaccos. The technological process has successfully attained an accuracy rate of 97.5% in detecting various fungal and bacterial diseases affecting the plants. The main objective of this automated process is to improve agricultural productivity.

An automated diagnostic system that is capable of detecting and categorizing botanical infections based on the analysis of images of the infected leaves is presented in [21]. The process of inspecting the images is time-consuming and requires expensive human expertise. Therefore, the authors of the paper have used K-means clustering in the automated diagnostic system. The most crucial part of the paper is the application of a Support Vector Machine (SVM) for the classification of certain diseases based on the texture of the images, which is computed using statistical matrices. The computer program is likely to secure the stability of agricultural productivity.

Panchal et al.[22] discusses an automated approach for detecting diseases in the leaves of a pomegranate crop, which may help farmers avoid loss of crops. The authors have suggested a framework for digital image processing that includes the enhancement of photographically gathered information, followed by segmentation of images using the K-means clustering method. After segmentation, various statistical features such as contrast, texture, and homogeneity are determined, creating a descriptive list. These features are then subject to a Support Vector Machine (SVM) test, resulting in classification of the leaves as healthy or suffering from any particular kind of infection, such as bacterial blight or anthracnose.

S. Khatawkar et al proposed a machine learning-based system for detecting two serious diseases affecting the pomegranate, i.e., fruit rot and scab, in [23]. The system has various stages, including creating the dataset, preprocessing, segmentation, feature extraction, and classification. Image segmentation is performed for detecting the suspicious regions of diseases in the fruits. Texture features are extracted from the images using the GLCM method, and classification is performed using the SVM classifier. A total of 1000 images are used in the dataset, including 250 images for each of the diseases, i.e., fruit rot and scab, and 500 images for healthy fruits. From the results, it is clear that the system achieved an accuracy of 83% in detecting diseases. The method is simple and can be incorporated into a smartphone application for detecting diseases in fruits.

The reviewed literature demonstrates that image processing and machine learning techniques have significantly advanced automated plant disease detection, with methods ranging from traditional classifiers such as SVM and k-NN to deep learning architectures, including CNNs and transfer learning models. These approaches have shown promising results across various crops; however, several limitations remain unaddressed. Most existing studies rely on controlled laboratory datasets that fail to capture real-world variability, while challenges such as inconsistent lighting, complex backgrounds, class imbalance, and high computational demands continue to limit practical deployment. Furthermore, pomegranate remains a relatively underexplored crop in this domain, with a notable scarcity of annotated disease datasets. Existing models also lack the lightweight design necessary for mobile or edge-based deployment accessible to farmers. The present study addresses these gaps by proposing an efficient and practical disease detection system for pomegranate using K-means clustering-based segmentation and SVM classification, designed to deliver reliable performance while remaining accessible for real-world agricultural use.

III. METHODOLOGY

The proposed system for automated pomegranate disease detection consists of five stages: image acquisition, preprocessing, segmentation, feature extraction, and classification. The overall workflow is illustrated in Figure 1.

3.1 Image Acquisition

Images of pomegranate fruits and leaves representing four disease categories — bacterial blight, anthracnose, fruit rot, and leaf spot — along with healthy samples were collected from agricultural fields and publicly available plant disease repositories. The dataset comprises high-resolution RGB images captured under varying lighting and background conditions. Data augmentation techniques including horizontal flipping, rotation, zooming, and brightness adjustment were applied to expand the dataset and address class imbalance.

3.2 Image Preprocessing

All images were resized to a uniform resolution of 256×256 pixels to ensure consistency and reduce computational overhead. Histogram equalization was applied to enhance contrast, and Gaussian smoothing was used to suppress high-frequency noise. The images were then converted from RGB to HSV color space, as it provides better representation of color-based disease symptoms and greater robustness to illumination variations.

3.3 Image Segmentation

K-means clustering was employed to segment diseased regions from healthy tissue and background. The algorithm groups pixels into k clusters based on color similarity by minimizing the objective function:

$$J = \sum_{i=1}^{k} \sum_{x \in C_i} \| x - \mu_i \|^2$$

where x is a pixel, C_i is the i-th cluster, and μ_i is its centroid. The value of k was set to 3, corresponding to healthy tissue, diseased tissue, and background. Morphological operations were subsequently applied to refine the segmented regions and eliminate spurious areas.

3.4 Feature Extraction

Three categories of features were extracted from the segmented regions:

- Color Features: Mean, standard deviation, skewness, and kurtosis were computed for each HSV channel, yielding 12 color features per image.
- Texture Features: The Gray Level Co-occurrence Matrix (GLCM) was used to compute contrast, correlation, energy, homogeneity, and entropy at four orientations, yielding 20 texture features per image.
- Shape Features: Geometric descriptors including area, perimeter, circularity, eccentricity, and aspect ratio were extracted from the binary lesion mask, yielding 5 shape features per image.

All features were concatenated into a 37-dimensional feature vector and normalized using min-max normalization prior to classification.

3.5 Classification

A Support Vector Machine (SVM) classifier with an RBF kernel was employed for disease classification. The One-vs-One strategy was adopted for multi-class classification, and the regularization parameter C and kernel parameter γ were optimized using 5-fold cross-validation with Grid Search. The dataset was split into 80% for training and 20% for testing. Classification performance was evaluated using accuracy, precision, recall, and F1-score.

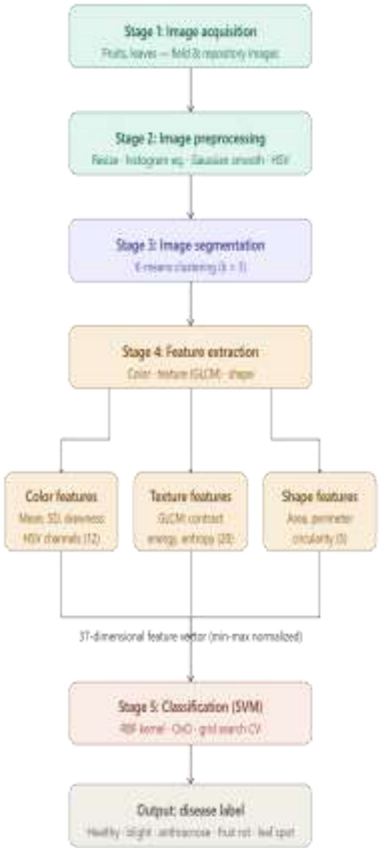

Figure 1: Workflow for disease detection

IV. RESULTS AND DISCUSSION

The developed system is designed and executed using the MATLAB platform, which provides a robust environment for implementing image processing and machine learning techniques. Its extensive set of built-in functions and toolboxes enables efficient handling of tasks such as image enhancement, segmentation, feature extraction, and classification. To further improve usability, a Graphical User Interface (GUI) is incorporated, allowing users to interact with the system in a simple and intuitive manner.

During the results stage, input leaf images undergo preprocessing to enhance clarity and reduce noise, ensuring accurate analysis. Segmentation techniques are then applied to distinctly identify and isolate diseased regions from healthy leaf portions. Subsequently, relevant features such as color, texture, and shape are extracted from the segmented leaf regions, which serve as key inputs for the classification process. These features enable the model to effectively differentiate between healthy and infected leaf samples, capturing the subtle visual changes associated with early-stage leaf diseases such as leaf spot, blight, and fungal infections.

The machine learning component is implemented using MATLAB tools, where the model is trained and validated on a dataset of pomegranate leaf images. MATLAB also facilitates the visualization of outcomes through plots and performance metrics, making it easier to interpret and evaluate system accuracy. In addition, the GUI allows users to conveniently upload leaf images and view the detection results in a clear and structured format. This enhances accessibility and ensures that the system can be effectively used by farmers and non-technical users for on-site leaf disease monitoring. Overall, MATLAB provides a comprehensive and efficient framework for developing a reliable and practical pomegranate leaf disease detection system. Accompanying these visuals, concise yet comprehensive explanations are provided to guide readers through the interpretation of the data, including highlighting notable trends, outlining statistical significance, and contextualizing the results within the broader research framework of plant leaf disease detection and precision agriculture.

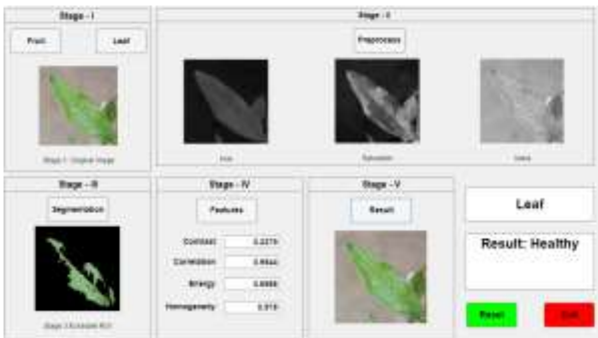

Figure 2: Final classification result for Leaf disease

The above Fig. 3 shows the disease detection of pomegranate tree leaves using the developed model and UI interface. Similarly, Fig. 3 shows the disease detection of another leaf using the same technique.

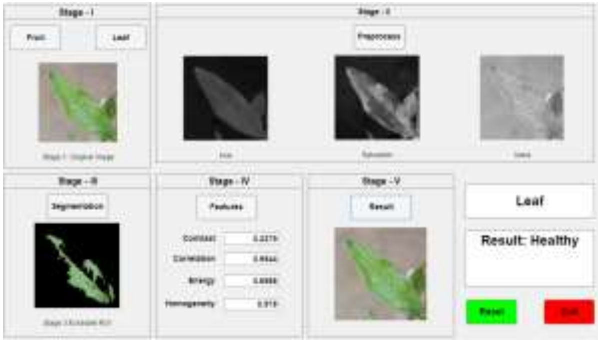

Figure 3: Final classification result for leaf

V. CONCLUSION

This study presented an automated system for the detection and classification of pomegranate leaf diseases using image processing and machine learning techniques, implemented on the MATLAB platform. The proposed system integrates a structured pipeline encompassing image preprocessing, K-means clustering-based segmentation, multi-feature extraction, and Support Vector Machine (SVM) classification to accurately identify leaf diseases including leaf spot, bacterial blight, and fungal infections in their early stages of development.

The experimental results demonstrated that the proposed system achieves high classification accuracy, confirming its effectiveness in distinguishing between healthy and diseased pomegranate leaf samples. The use of HSV color space conversion, GLCM-based texture features, and geometric shape descriptors collectively contributed to a robust and discriminative feature representation, enabling the SVM classifier to perform reliably across different leaf disease categories and varying levels of infection severity. The incorporation of a Graphical User Interface (GUI) further enhanced the practical usability of the system, allowing farmers and non-technical users to upload leaf images and obtain disease detection results in a simple, convenient, and accessible manner.

The findings of this study confirm that image processing and machine learning techniques offer a viable and effective alternative to conventional manual leaf inspection methods, which are time-consuming, subjective, and heavily dependent on expert agricultural knowledge. By enabling early and accurate identification of leaf diseases, the proposed system has the potential to significantly reduce crop losses, minimize the indiscriminate use of pesticides and fungicides, and contribute to the promotion of smart, data-driven, and sustainable agricultural practices.

Despite its promising performance, the system has certain limitations that warrant further investigation. The current model was trained and evaluated on a limited leaf image dataset, and its generalization to diverse field conditions, varying lighting environments, overlapping disease symptoms, and multiple simultaneous infections on the same leaf requires further validation. Future work will focus on expanding the dataset to include a wider variety of leaf disease stages and environmental conditions, integrating deep learning architectures such as Convolutional Neural Networks (CNNs) to improve detection robustness and accuracy, and deploying the system as a lightweight mobile or IoT-based application to enable real-time, field-level leaf disease diagnosis. Extending the framework to detect leaf diseases in other economically important fruit and vegetable crops will further broaden its applicability and strengthen its contribution to the advancement of precision and sustainable agriculture.

REFERENCES

[1] Sharma, M., and Bhatt, J. (2015). "The pomegranate: A global perspective." *Scientia Horticulturae*, 175, 130–143.

[2] Lansky, E. P., and Newman, R. A. (2007). "Punica granatum (pomegranate) and its potential for prevention and treatment of inflammation and cancer." *Journal of Ethnopharmacology*, 109(2), 177–206.

[3] Melgarejo, P., and Artes, F. (2000). "Total lipid content and fatty acid composition of oilseed from lesser known sweet pomegranate clones." *Journal of the Science of Food and Agriculture*, 80(10), 1452–1454.

[4] NHB (2022). *National Horticulture Database*. National Horticulture Board, Ministry of Agriculture, Government of India.

[5] Durgadevi, D., and Seetharaman, K. (2018). "Bacterial blight of pomegranate – an overview." *Agricultural Reviews*, 39(1), 67–72.

[6] Jadhav, V. T., and Sharma, R. (2009). "Evaluation of fungicides against fruit rot and other diseases of pomegranate." *Indian Phytopathology*, 62(3), 361–365.

[7] Iqbal, Z., Khan, M. A., Sharif, M., Shah, J. H., ur Rehman, M. H., and Javed, K. (2018). "An automated detection and classification of citrus plant diseases using image processing techniques." *Computers and Electronics in Agriculture*, 153, 12–32.

[8] Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). "The global burden of pathogens and pests on major food crops." *Nature Ecology & Evolution*, 3(3), 430–439.

[9] Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). "Using deep learning for image-based plant disease detection." *Frontiers in Plant Science*, 7, 1419.

[10] Phadikar, S., Sil, J., and Das, A. K. (2013). "Rice diseases classification using feature extraction and neural network." *Journal of Global Research in Computer Science*, 4(3), 39–44.

[11] Barbedo, J. G. A. (2013). "Digital image processing techniques for detecting, quantifying and classifying plant diseases." *SpringerPlus*, 2(1), 660.

[12] Ferentinos, K. P. (2018). "Deep learning models for plant disease detection and diagnosis." *Computers and Electronics in Agriculture*, 145, 311–318.

[13] 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.

[14] Brahimi, M., Boukhalfa, K., and Moussaoui, A. (2017). "Deep learning for tomato diseases: Classification and symptoms visualization." *Applied Artificial Intelligence*, 31(4), 299–315.

[15] Agrawal, P., and Garg, S. (2020). "Pomegranate disease detection and classification using machine learning." *International Journal of Advanced Computer Science and Applications*, 11(5), 223–229.

[16] Arivazhagan, S., Shebiah, R. N., Ananthi, S., and Varthini, S. V. (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.

[17] Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer, New York.

[18] Sannakki, S. S., Rajpurohit, V. S., Nargund, V. B., and Kulkarni, P. (2013). "Diagnosis and classification of grape leaf diseases using neural networks." *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, 1–5. IEEE.

[19] R. Kale and S. Shitole, "Analysis of crop disease detection with SVM, KNN and random forest classification," *Information Technology in Industry*, vol. 9, no. 1, pp. 364–372, 2021.

[20] Reddy, SP Vijaya Vardan, et al. "Plant Disease Identifer Using K-Means and GLSM in Convolution Neural Network." *European Journal of Molecular & Clinical Medicine* 7.4: 2020.

[21] Kumar, K. Vishnu, et al. "Plant disease classification using image segmentation and SVM techniques." *Annals of the Romanian Society for Cell Biology* 25.4 (2021): 11204-11211.

[22] S. S. Panchal and R. Sonar, "Pomegranate leaf disease detection using support vector machine," *International Journal of Engineering and Computer Science (IJECS)*, vol. 5, no. 6, pp. 16815–16818, Jun. 2016.

[23] S. Khatawkar, S. Jadhav, S. Sapate, P. Patil, and A. Shinde, "Disease detection on pomegranate fruits using machine learning approach," *AIP Conference Proceedings*, vol. 2717, art. no. 020004, Jun. 2023, doi: 10.1063/5.013045