

FEATURE EXTRACTION AND INTELLIGENT CLASSIFICATION OF POMEGRANATE LEAF DISEASES**¹Nirmal M.D., ²Dr.Pramod Jadhav and ³Dr. Santosh Pawar,**¹Dr. A. P. J. Abdul Kalam University, Indore²Associate Professor, Dr. A. P. J. Abdul Kalam University, Indore³Professor, Dr. A. P. J. Abdul Kalam University, Indore¹vedic1234@gmail.com**ABSTRACT**

The identification and classification of plant diseases through image analysis remain significant challenges in precision agriculture. This study presents an efficient framework for the detection and classification of diseases affecting pomegranate leaves using image processing and machine learning techniques. The proposed methodology involves a sequence of stages, including image acquisition, preprocessing, image enhancement, segmentation, and feature extraction. A publicly available pomegranate leaf image dataset obtained from Mendeley Data was utilized for experimentation. The dataset consists of 559 images, comprising 287 healthy leaf samples and 272 diseased leaf samples. For model development and evaluation, the dataset was partitioned into training and testing subsets using an 80:20 ratio.

During the feature extraction stage, the region of interest (ROI) was identified through image enhancement and segmentation procedures. The extracted features were subsequently employed for disease classification using a Support Vector Machine (SVM)-based supervised learning approach. The developed framework effectively distinguishes healthy and diseased pomegranate leaves and further categorizes disease-affected samples into their respective classes.

Experimental results demonstrate that the proposed framework achieved an accuracy of 95.54% in distinguishing healthy and diseased leaves. Furthermore, the disease classification stage attained an accuracy of 96.43%, indicating the effectiveness of the proposed methodology for automated pomegranate leaf disease analysis. The obtained results suggest that the framework can serve as a reliable decision-support tool for the early detection and classification of pomegranate leaf diseases in agricultural applications.

Keywords: Pomegranate Leaves Classification, Feature Extraction, Machine Learning Techniques

I. INTRODUCTION

Agriculture plays a pivotal role in the economic development of India and remains one of the primary sources of livelihood for a substantial portion of the population. The sector contributes significantly to the national economy and supports the socioeconomic well-being of rural communities. Over the years, advancements in agricultural practices and technologies have enhanced crop productivity and sustainability. A large percentage of the rural population continues to depend directly or indirectly on agriculture for income generation and employment opportunities [1].

Among the various horticultural crops cultivated in India, pomegranate (*Punica granatum L.*) occupies an important position due to its nutritional value, commercial significance, and adaptability to diverse climatic conditions. The crop exhibits excellent tolerance to arid and semi-arid environments, making it suitable for cultivation in regions with limited water availability. Pomegranate cultivation has expanded considerably in recent years owing to its high market demand and export potential. However, the productivity and quality of pomegranate crops are often affected by various leaf diseases and environmental stresses. These factors adversely influence plant growth, reduce fruit yield, and ultimately impact farmers' income.

The increasing demand for sustainable agricultural production has encouraged the adoption of advanced technologies for crop monitoring and disease management. In this context, image-processing-based techniques

have emerged as effective tools for the remote analysis of crop health using digital imaging sensors. Such approaches facilitate rapid, non-destructive, and cost-effective assessment of plant conditions while minimizing human intervention and diagnostic errors [1–3].

Plant diseases commonly manifest through visible symptoms such as discoloration, yellowing, wilting, browning, spotting, lesion formation, and abnormal growth patterns. These symptoms may arise due to fungal, bacterial, or viral infections, as well as environmental factors including temperature fluctuations and water stress [2–5]. Accurate and timely identification of disease symptoms is essential for implementing appropriate management strategies and reducing crop losses. Delayed or incorrect diagnosis can lead to ineffective treatment practices, resulting in increased expenditure of resources and reduced agricultural productivity [6–8].

Consequently, the development of automated disease detection systems has become an important research area in precision agriculture. Advances in digital image processing and machine learning provide opportunities to identify disease symptoms efficiently and objectively. Compared with traditional visual inspection, computer-assisted analysis offers improved consistency, faster diagnosis, and enhanced reliability, thereby supporting early disease detection and informed decision-making in agricultural applications.

The following characteristics are possessed by diseases, and Fig. 1 illustrates what these characteristics look like:

The features of the diseases [6] is described below and illustrated in Fig.1:

- When plants are infected with this fungus, the fruits and leaves produce spots that are initially circular and dark brown, but as the plant matures, the spots become more irregular. There is a faint halo that surrounds the lesions, which range in colour from a very dark rusty brown to almost completely black.



Fig. 1. (1) H₂O control, (2) Black rot (internal rot), (3) Pathogenic black spot (4). Pathogenic black spot. *Courtesy -First report of Alternaria black spot of pomegranate caused by Alternaria alternata in Israel*

- To distinguish between bacterial and fungal leaf diseases, you should place the leaves in a humid chamber for two to three days and then look for fungal structures, which will appear as minute black dots in the lesions. Before they completely dry out, bacterial lesions will have a "water-soaked" or "glassy" appearance if the surrounding environment is damp.
- The spots on the leaves are rather tiny, brown, and surrounded by a bright yellow halo.
- Spots of either a circular or an irregular shape appear on ageing skin. These spots can be either brown or black.
- The underside is darker in colour because it contains more clumps of spore-bearing structures.
- The flower's sepals include numerous minute, spherical, dark spots all over their surface.

The proposed study utilizes image segmentation, texture feature extraction, and statistical image analysis techniques in combination with machine learning algorithms for the detection and classification of pomegranate

leaf diseases. The integration of these techniques aims to improve the efficiency and accuracy of the disease identification process by extracting meaningful features from leaf images and employing an effective classification mechanism.

The remainder of this paper is organized as follows. Section 2 reviews the existing literature related to plant leaf disease detection and classification. Section 3 presents the proposed methodology, including feature extraction and machine learning-based classification techniques used for identifying diseased leaf regions and categorizing disease classes. Section 4 discusses the experimental results and corresponding analysis. The final section concludes the study.

II. LITERATURE SURVEY

A. 2. Literature Review

The application of image processing and machine learning techniques for plant disease detection has attracted considerable attention from researchers in recent years. Advances in digital imaging technologies have enabled the development of automated systems capable of identifying and classifying plant diseases with improved accuracy and efficiency. Consequently, numerous studies have explored various image-processing and machine-learning approaches for disease diagnosis in agricultural crops.

Sharath D. M. et al. [1] proposed a computer-assisted system for the identification of plant diseases in leaves and fruits. Their methodology consisted of image acquisition, preprocessing, feature extraction, segmentation, classification, and estimation of disease severity. Grab Cut Segmentation (GCS) was employed to isolate the fruit region, while Canny Edge Detection (CED) was used to identify disease-affected areas. Disease severity was estimated by analyzing infected pixels within the segmented region.

Nikhil Shah et al. [2] utilized an Artificial Neural Network (ANN) for the classification of cotton leaf diseases. The proposed framework included image acquisition, preprocessing, feature extraction, and ANN-based classification. The extracted image features were used to determine disease severity and classify cotton leaf samples into different health categories.

Sumit Nema et al. [3] employed Support Vector Machine (SVM) classification for wheat leaf disease detection. K-means clustering was applied to segment disease-affected regions, while the Lab color space was used to distinguish healthy and infected areas. Statistical features such as variance, mean, mode, and standard deviation were extracted and utilized for disease classification.

Sa'ed Abed et al. [4] developed an SVM-based classification framework for the detection of diseases in bean plants. The system was trained and evaluated using images obtained from a publicly available database. Feature extraction techniques were applied to identify disease characteristics, which were subsequently classified using an SVM classifier.

Arya M. S. et al. [5] proposed a leaf-sorting approach based on Genetic Algorithms (GA). The methodology involved image acquisition, feature extraction, and automated sorting of leaves using an Arduino-controlled conveyor mechanism. Similarly, P. Krithika et al. [6] combined K-means clustering and Support Vector Machine classification for the detection and classification of cucumber leaf diseases.

Deep learning techniques have also been widely adopted for plant disease identification. Sladojevic et al. [9] employed Convolutional Neural Networks (CNNs) for the recognition and classification of multiple plant diseases. Their study reported classification accuracies ranging from 91% to 98%, with an average accuracy of 96.3%. Singh et al. [10] proposed a disease diagnosis framework that integrated image processing techniques with optimization methods for plant disease identification.

Khirade et al. [11] developed an ANN-based system for automated plant disease detection. Kirani et al. [12] further explored intelligent disease diagnosis techniques, while Singh et al. [13] combined Genetic Algorithms, K-

means clustering, and Support Vector Machines for plant disease classification. Their study reported classification accuracies of 93.6% using the minimum distance criterion and 95.7% using SVM classification.

Golhani et al. [14] investigated the use of hyperspectral data and neural network models for plant disease detection. Various neural network architectures, including single-layer, multilayer, radial basis function, probabilistic neural network, and convolutional neural network models, were evaluated for disease identification. Mattihalli et al. [15–18] proposed an Internet of Things (IoT)-based architecture capable of monitoring plant health and supporting disease management during the early stages of infection. In addition, several studies [18–22] explored molecular and spectroscopy-based approaches, including fluorescence spectroscopy, visible and infrared spectroscopy, fluorescence imaging, hyperspectral imaging, electronic nose systems, and gas chromatography–mass spectrometry (GC-MS) for plant disease analysis.

Although significant progress has been achieved in plant disease detection, many existing approaches have been developed for generalized crop datasets. Since disease symptoms and characteristics vary across different plant species, crop-specific disease detection frameworks can provide improved classification performance. Motivated by this observation, the present work focuses specifically on pomegranate leaf disease classification using feature extraction and machine learning techniques to achieve reliable and accurate disease identification.

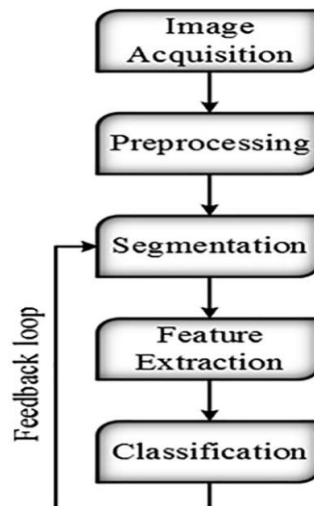


Fig. 2: Process flow steps.

The proposed framework for leaf disease detection is based on machine learning techniques, as illustrated in Figure 2. Machine learning provides an effective approach for the identification and classification of plant diseases by analyzing image-based characteristics and extracting meaningful features from leaf samples. The integration of image processing and machine learning techniques enables automated disease diagnosis and supports timely decision-making for disease management.

Machine-learning-based systems have demonstrated promising performance across a wide range of classification tasks and have been successfully applied in agricultural applications. However, existing approaches often exhibit limitations in terms of classification accuracy, feature representation, and adaptability to specific crop diseases. Therefore, there is a need to develop improved classification frameworks that can effectively identify disease symptoms and enhance diagnostic performance.

The application of machine learning for the recognition and classification of plant leaf diseases remains an active area of research. Continuous advancements in image processing, feature extraction, and classification algorithms provide opportunities for developing more accurate and reliable disease detection systems for agricultural crops.

III. PROPOSED METHODOLOGY

3.1 Input Image

As part of this experiment, we integrated data that came from a wide range of different sources to produce our very own original dataset. Images of pomegranate leaves make up this collection's subject matter. Figure 1 presents an illustration of the information that may be found in this data set.

- **Healthy Leaves:** These leaves do not have any infection and are green in color.
- **Alternaria Alternata Infected Leaves Images:** These are a fungus that causes dark spots on plant parts and rots them.
- **Anthracoze:** For this disease, dark and sunken type lesion is visible on leaves. It is an infection caused by fungus.
- **Bacterial Blight:** It is a bacteria-infected disease in which pale green spots appear on leaves and in later stage leaf appear water-soaked.
- **Cercospora Leaf Spot:** Black dots and grayish tanned lesions appear on the leaf.

Taking a picture of the leaf of the groundnut plant is the first and most important step in determining how well it is doing. A number of different tools, such as cameras, mobile phones, and satellites, have been utilized to capture and store images. One of the most important steps in capturing a picture that turns out well is getting an exact count of the amount of RGB colour pixels that are contained in the image. The sensitivity of photosensors to light, the resolution of their spatial arrays, and the precision of their digital focusing have all seen considerable advancements over time. The use of high-tech equipment like as digital cameras and cell phones is already ubiquitous in the hands of plant pathologists and farmers alike [13–17].

Pre – Processing: The data set images is then subjected to preprocessing once it has been captured. Both the creation of a database containing all of the photographs and the finishing of the pre-processing steps are complete.

Segmentation: The process of separating the diseased area from the rest of the image is referred to as segmentation (leaf). There are many various strategies to segment data, some of which include clustering, thresholding, edge detection, ANN-based algorithms, partial differential equation-based segmentation, and many others [18]. The k-mean clustering method for segmentation was given the highest weight in this paper. This method was chosen out of the strategies that were mentioned above. The k-means clustering algorithm is useful for situations in which there are very large datasets. After the formation of a cluster with the same level of intensity, the centre of mass was no longer able to shift. The finished product is a combination of leaves that are healthy and leaves that are unhealthy. After the segmentation process was complete, one infected cluster was chosen at random and used to evaluate the severity of the disease's impact on one or more clusters.

3.2 Feature Extraction

After constructing the Gray Level Co-occurrence Matrix (GLCM), texture features are extracted from the segmented image. The GLCM is a statistical method used to analyze the spatial relationship between pixel intensities and to characterize the texture properties of an image. It represents the frequency with which pairs of pixel intensity values occur at a specified spatial distance and orientation within the image.

Texture analysis based on GLCM provides valuable information regarding the structural arrangement and distribution of pixel intensities in the diseased region. Statistical measures derived from the GLCM capture the textural characteristics of the image and serve as discriminative features for subsequent classification. Feature extraction involves transforming the image data into a set of representative statistical descriptors that effectively characterize the underlying texture patterns.

The extracted texture features are utilized to distinguish healthy and diseased leaf regions by capturing variations in surface appearance and disease symptoms. Since the background is removed during the segmentation stage, the disease-affected region becomes more prominent, enabling the extraction of meaningful features and improving the effectiveness of the classification process.

The pretreatment steps are presented in graphical format in Figure 3, which may be seen here.

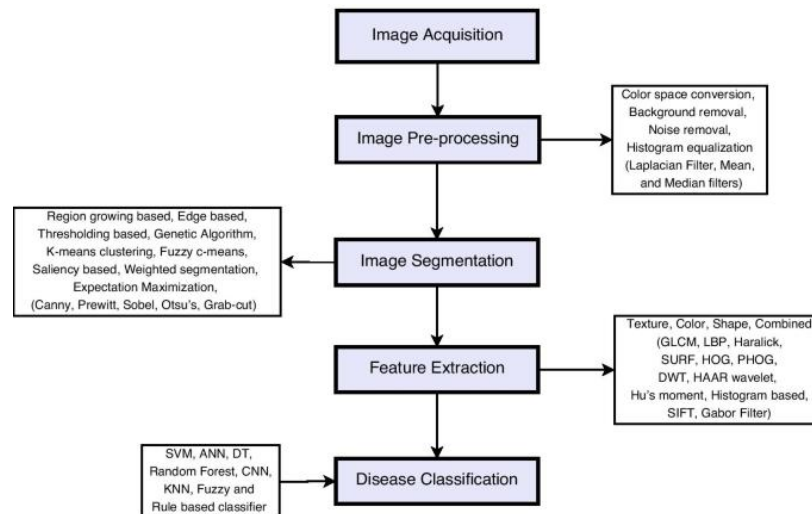


Fig 4: Feature Extraction steps for background elimination of diseased images.

3.3 LBP – Local Binary Dataset

Simple, yet successful: The Local Binary Pattern, often known as LBP, applies a threshold to each pixel that makes up an image and then uses the result of this thresholding to determine an integer value. This is done in order to identify each individual pixel that makes up the image. Despite the fact that it was not formally defined until 1994, this feature has been demonstrated to be useful in the classification of textured surfaces (in LBP).

3.4 Decision Tree

Decision Tree is one of the most widely used classification techniques in machine learning due to its simplicity and effectiveness in handling classification problems [20]. The algorithm constructs a tree-like structure by recursively partitioning the dataset based on the most informative attribute. At each stage, the data are divided into subsets in such a way that the resulting groups become increasingly homogeneous with respect to the target class. The partitioning process continues until a stopping criterion is satisfied or the resulting subsets contain instances belonging predominantly to a single class.

The ID3 algorithm follows a greedy approach for constructing the decision tree. During tree generation, the attribute that provides the highest information gain is selected for splitting the dataset at each node. The concepts of entropy and information gain, derived from information theory, play a crucial role in this process. Entropy measures the degree of uncertainty or impurity within a dataset. A lower entropy value indicates that the instances belong predominantly to the same class, whereas a higher entropy value reflects greater heterogeneity among the instances. Information gain quantifies the reduction in entropy achieved after partitioning the data and is used to determine the most suitable attribute for each split in the decision tree.

$$E = \sum_{i=1}^c -p_i \log_2 p_i$$

The sign c is used throughout the study of mathematics to denote the total number of classes. One is able to progress to the subsequent node in the tree if they have the ability to obtain new information. It is a good idea to

choose the attribute that has the biggest potential for producing valuable data because this will help you make better decisions.

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum \frac{|sv|}{|S|}$$

Because A plays such an important role in this scenario, the portion of A that is denoted by *sv* is referred to as *sv*. With the help of a decision tree strategy that makes use of 10-fold cross validation, 94.9074 percent of the training data was successfully labelled despite the usage of only five criteria. The accuracy of the model was determined to be 97.9167 percent when it was tested using data taken from the real world.

sv is a subset of A in which A possesses the value v, where A is the attribute that is already known to be present in *sv* and is denoted by the letter v. With the help of a decision tree strategy that makes use of 10-fold cross validation, 91.9643% of the training data was successfully labelled despite the usage of only five criteria. The accuracy of the model was determined to be 90.1786 % when it was tested using data taken from the real world.

3.5 SVM - Support Vector Machine

1) Support Vector Machine (SVM) Classification

Support Vector Machine (SVM) is a widely used machine learning algorithm for classification and regression tasks. In the proposed framework, SVM is employed as the primary classifier due to its effectiveness in handling high-dimensional feature spaces and its ability to achieve reliable classification performance. The main objective of the SVM algorithm is to determine an optimal hyperplane that separates data samples belonging to different classes with maximum margin.

SVM can be applied to both linear and non-linear datasets through the use of kernel functions. For linearly separable data, a linear kernel can be utilized to construct the decision boundary. In the case of non-linear datasets, kernel functions such as polynomial and radial basis function (RBF) kernels are commonly employed to transform the data into a higher-dimensional feature space, enabling effective class separation.

The extracted texture and statistical features obtained from the leaf images are used as input to the SVM classifier for disease classification. Since plant disease identification involves recognizing complex patterns present in leaf images, supervised learning techniques are required to learn the relationship between feature values and disease categories. SVM has been widely adopted for such classification tasks because of its computational efficiency and capability to model complex decision boundaries.

In practical applications, disease classification may involve multiple categories rather than only two classes. Therefore, a multi-class SVM approach can be employed to classify leaf samples into their respective disease classes. The use of kernel functions further enhances the flexibility of the classifier by enabling effective handling of complex and non-linearly separable datasets. In the proposed framework, a linear kernel function is utilized throughout the classification process to generate the final disease prediction.

3.5.1 Mathematical Support of Multi-Class SVM

This method can be utilized to mathematically represent data that can be partitioned into two distinct groups and is capable of being represented graphically by means of a line or a plane. A hyperplane can be thought of as a line or plane that separates the data into two groups. In situations when there is only one dimension of data, the hyperplane can be represented by a single point. Due to the fact that a dataset only has two dimensions, the hyperplane is represented by a straightforward line. If there are less than three dimensions in the data set, then the hyperplane will include a plane within it. Following that, we will investigate the mathematical foundations of SVM's decision-making process in further detail. It is possible to improve data partitioning by making use of an SVM's hyperplane. This is a comprehensive list, which includes:

The SVM is a learning approach that makes advantage of assistance from outside sources. This method could be used to sort the data in the beginning if there are only two viable categories for the information. As technology progressed, SVM became capable of handling data from multiple classes simultaneously. The search that SVM is conducting for the hyperplane that most effectively partitions the data can be broken down into two parts.

- It is possible to locate every possible collection of hyperplanes.
- Out of all the hyperplanes that are possible, you need to figure out which one is the best.

These two phases are included in the depiction that is mathematical. When the two-dimensional data space is taken into consideration, the hyperplane will be represented by a line. The equation can be seen as a geometric representation of a line.

$$W_T X + B = 0$$

$$\text{where } W = \begin{pmatrix} a \\ -1 \end{pmatrix} \text{ and } X = \begin{pmatrix} x \\ y \end{pmatrix}$$

If you have an n-dimensional data space, you can represent the hyperplane using the equation $W_T X + B = 0$. Here, W is a column matrix with size $n \times 1$, and X is another column matrix with the same size. It is anticipated that the two-dimensional hyperplane will partition the data space into two distinct parts for the classifiers.

The hypothesis $H(x_i)$ will be defined as in equation:

$$H(x_i) = \begin{cases} +1, & \text{if } wx_i + b \geq 0 \\ -1, & \text{if } wx_i + b \leq 0 \end{cases}$$

In this scenario, there will be no restrictions on the number of hyperplanes that can exist, which means that there will be room for more of them. It is absolutely necessary to zero in on the ideal hyperplane from among the several possibilities. The two components of the hypothesis that we have gone over up to this point can be combined into an equation if we want to express them together.

$$y_i(wx_i + b) - 1 \geq 0 \text{ for } y_i = +1, -1$$

In the equation $y_i = +1$ that was just presented, there are two hyperplanes denoted by the letters H_1 and H_2 . We refer to the first of these as the H_1 hyperplane. This indicates that the distance between H_1 and the starting point is currently calculated as $((1-b)/|W|)$ and $((-1-b)/|W|)$. As a direct consequence of this, the margin may be found. In order to create an ideal hyperplane, it is necessary to maximize the distance that the margins have from the origin. If you say this, it means that you want to maximize $2/|W|$ while minimising W or $|W|/2$, depending on which you choose. The most effective hyperplane can be discovered in this manner.

3.6 Random Forest

1) Random Forest Classification

Random Forest is a supervised machine learning algorithm that is widely used for both classification and regression tasks. It is an ensemble learning technique that combines the predictions of multiple decision trees to improve classification performance and enhance model robustness. By aggregating the outputs of several individual classifiers, Random Forest is capable of handling complex classification problems effectively.

The Random Forest classifier constructs multiple decision trees using different subsets of the original dataset. Each tree is trained independently, and the final prediction is obtained by combining the predictions of all trees through a voting mechanism. For classification tasks, the class receiving the highest number of votes is selected as the final output. This ensemble approach reduces the risk of overfitting and improves the generalization capability of the model.

The performance of the Random Forest classifier is influenced by the number of trees used in the ensemble. Increasing the number of decision trees generally improves classification accuracy and reduces prediction variance, resulting in a more stable and reliable model.

In the proposed framework, the Random Forest classifier is employed to classify pomegranate leaf samples based on the extracted feature set. The implementation is carried out using the Random Forest Classifier available in the sklearn.ensemble library, which provides an efficient and scalable approach for disease classification.

IV. RESULT ANALYSIS

The dataset used in this study consists of 559 pomegranate leaf images, including 287 healthy samples and 272 diseased samples. For model development and performance evaluation, the dataset was divided into training and testing subsets. Among the diseased samples, images corresponding to *Alternaria Alternata*, Anthracnose, *Cercospora*, and Bacterial Blight were considered for disease classification. Figure 5 presents representative examples of the raw leaf images used in the study.



Figure 5: Shows the Raw Image of Data Collected



Figure 6: a. Shows the ROI Image of Data Collected

Figure 6. (a). depicts the ROI - region of interest image of Pomegranate Leaf and (b). Resized Pomegranate image.



Figure 6: b. Shows the resized image Data Collected

Figure 7. (a). Shows the Grayscale of pomegranate leaf and it is a range of gray shades from white to black, as used in a monochrome display or printout. Grayscale images are most commonly used in image processing because smaller data enables developers to do more complex operations in a shorter time.

Gray image

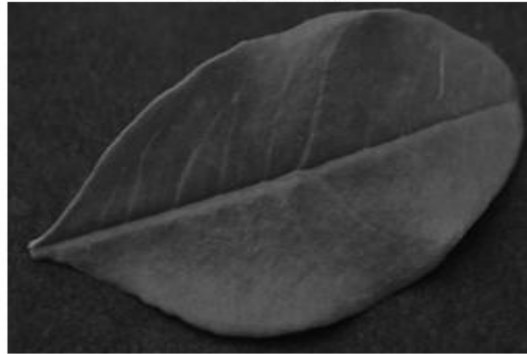


Figure 7: a. Shows the Gray Scale Image Data Collected

Filtered Image

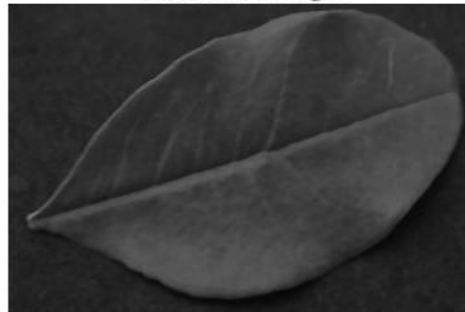


Figure 7: b. Shows the Filtered Image Data Collected after noise removal

Figure 7.b. shows the median filter of a pomegranate leaf after noise removal. Median filter is a nonlinear filter in which each output sample is computed as the median value of the input samples under the window – that is, the result is the middle value after the input values have been sorted. Ordinarily, an odd number of taps is used.

Table I shows the accuracy of three classification algorithms after performing 10- fold cross validation on dataset. Where best four attributes were selected.

Table I: Accuracy on Training and Test Dataset

MODEL	ACCURACY	SENSITIVITY	SPECIFICITY	RECALL	PRECISION
RF	93.75	90.385	96.6667	90.3846	95.9184
DT	90.1786	86.538	93.3333	86.5385	91.8367
SVM	95.5357	96.078	95.082	96.0784	94.2308

The comparison between the accuracy of the three classification algorithms are represented in Figures8(a) and (b). Besides accuracy, other performance measures like Precision, Accuracy, Recall, Sensitivity, Specificity are also evaluated to compare among the three algorithms.

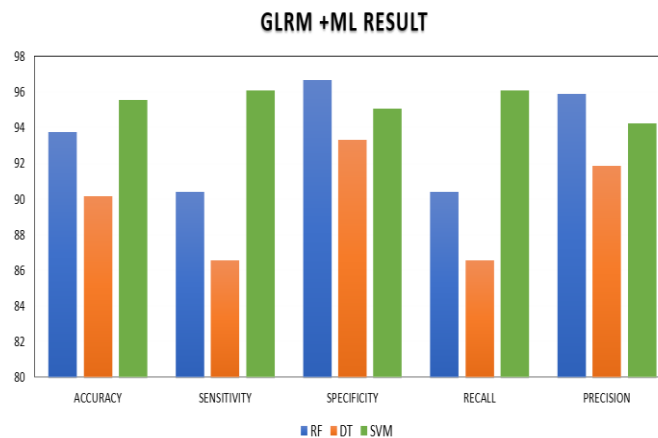


Figure 8: a. Shows the Filtered Image Data Collected after noise removal

Table II: Accuracy on Training and Test Dataset

MODEL	ACCURACY	SENSITIVITY	SPECIFICITY	RECALL	PRECISION
RF	94.6429	94.118	95.082	94.1176	94.1176
DT	91.9643	92	91.9355	92	90.1961
SVM	96.4286	94.118	98.3607	94.1176	97.9592

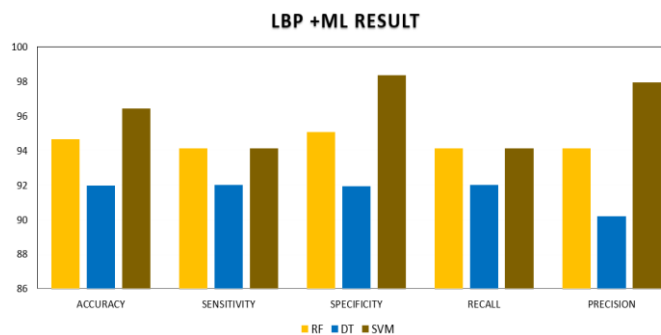


Figure 8: b. Shows the Filtered Image Data Collected after noise removal

Furthermore, it reveals from Table I and Table II that in each case, decision tree algorithm outperforms all other algorithms in detecting and classifying the diseases.

V. CONCLUSION

B. Conclusion

This study presented a machine-learning-based framework for the detection and classification of pomegranate leaf diseases using image processing techniques. The proposed methodology incorporated image preprocessing, segmentation, feature extraction, and classification to identify disease symptoms from leaf images. The framework was evaluated using a dataset containing healthy and diseased pomegranate leaf samples.

Four major pomegranate leaf diseases, namely Alternaria Alternata, Anthracnose, Bacterial Blight, and Cercospora Leaf Spot, were considered in the classification process. Experimental results demonstrated that the proposed framework achieved an accuracy of 95.5357% in distinguishing healthy and diseased leaf samples. Furthermore, the disease classification stage achieved an overall accuracy of 96.4286%, indicating the effectiveness of the proposed approach for automated disease diagnosis.

The obtained results demonstrate the potential of machine learning techniques for supporting disease identification in pomegranate cultivation. Future work may focus on the incorporation of deep learning techniques to further enhance classification performance. The use of Convolutional Neural Network (CNN)-based architectures with deeper feature learning capabilities may contribute to improved disease recognition accuracy and more robust classification results.

VI. REFERENCES

- [1] Sharath D M, Akhilesh, S Arun Kumar, Rohan M G and Prathap C, ‘Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight’, International Conference on Communication and Signal Processing, April 4-6, 2019, India.
- [2] Nikhil Shah, Sarika Jain, “Detection of Disease in Cotton Leaf using Artificial Neural Network”, 2019 IEEE.
- [3] Sumit Nema, Aarju Dixit, “Wheat Leaf Detection and Prevention Using Support Vector Machine”, 2018 IEEE International Conference on Circuits and Systems in Digital Enterprise Technology.
- [4] Sa’ed Abed and Anwar Ali Esmael, “A Novel Approach to Classify and Detect Bean Diseases based on Image Processing”, 2018 IEEE.
- [5] P. Krithika and S. Veni, “Leaf Disease Detection on Cucumber Leaves Using Multiclass Support Vector Machine”, IEEE WiSPNET 2017 conference.
- [6] P. Krithika and S. Veni, “Leaf Disease Detection on Cucumber Leaves Using Multiclass Support Vector Machine”, IEEE WiSPNET 2017 conference.
- [7] Monzurul Islam, Anh Dinh, Khan Wahid, Pankaj Bhowmik, “Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine”, 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE).
- [8] Agarwal, M., Singh, A., Arjaria, S., Sinha, A., and Gupta, S. (2020). ToLeD: tomato leaf disease detection using convolution neural network. *Proc. Comput. Sci.* 167, 293–301. doi: 10.1016/j.procs.2020.03.225
- Albert, B. A. (2020). Deep learning from limited training data: novel segmentation and ensemble algorithms applied to automatic melanoma diagnosis. *IEEE Access* 8, 31254–31269. doi: 10.1109/ACCESS.2020.2973188
- [9] Pugoy R, & Mariano V.Y., (2011). Automated rice leaf disease detection using color image analysis. 3rd international conference on digital image processing, 8009. Chengdu: SPIE; F1-F7.
- [10] M. Dhakate and Ingole A. B., "Diagnosis of pomegranate plant diseases using neural network," 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2015, pp. 1-4, doi: 10.1109/NCVPRIPG.2015.7490056.
- [11] N. Goel, D. Jain and A. Sinha, "Prediction Model for Automated Leaf Disease Detection & Analysis," 2018 IEEE 8th International Advance Computing Conference (IACC), 2018, pp. 360-365, doi: 10.1109/IADCC.2018.8692116.
- [12] P. B. Wakhare, S. Neduncheliyan and K. R. Thakur, "Study of Disease Identification in Pomegranate Using Leaf Detection Technique," 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), 2022, pp. 1-6, doi: 10.1109/ESCI53509.2022.9758262.
- [13] Anthracnose is a disease which affects not only pomegranate but also other fruits such as mango, papaya and Guava.
- [14] MK Hingorani and PP Mehta, "Bacterial leaf spot of pomegranate", *Indian Phytopathol*, vol. 5, pp. 55-56, 1952.

- [15] J Raju, Benagi Vitthal, Krishnappa Jayalakshmi, Nargund Vijay, Nirmalanath Jones, Sonavane Priti, et al., Bacterial Blight of Pomegranate: A Menace in Quality Fruit Production, 2015
- [16] [online] Available: <https://plantix.net/en/library/plant-diseases/100202/cercospora-fruit-and-leaf-spot>.
- [17] M. Dhakate and A. B. Ingole, "Diagnosis of pomegranate plant diseases using neural network", 2015 Fifth National Conference on Computer Vision Pattern Recognition Image Processing and Graphics (NCVPRIPG), pp. 1-4, 2015.
- [18] K. Nirmalakumari, H. Rajaguru and P. Rajkumar, "Plant Leaf Disease Detection Using Modified Segmentation Process and Classification" in ICCVBIC 2019. Advances in Intelligent Systems and Computing, Springer, Cham, vol. 1108, 2020.
- [19] S. D.M. Akhilesh, R. M. G, S. A. Kumar and P. C, "Disease Detection in Pomegranate using Image Processing", 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), pp. 994-999, 2020.
- [20] N. Goel, D. Jain and A. Sinha, "Prediction Model for Automated Leaf Disease Detection & Analysis", 2018 IEEE 8th International Advance Computing Conference (IACC), pp. 360-365, 2018.
- [21] G. Dhingra, V. Kumar, H.D. Joshi A novel computer vision based neutrosophic approach for leaf disease identification and classification Measurement, 135 (2019), pp. 782-794
- [22] A.D. Nidhis, C.N.V. Pardhu, K.C. Reddy, K. Deepa Cluster based paddy leaf disease detection, classification and diagnosis in crop health monitoring unit Comput Aided Interven Diagnost Clinic Med Images, 31 (2019), pp. 281-291.