# A Comprehensive Review of Machine Learning Techniques for Accurate Pomegranate Disease Classification

# **Abstract:**

The methods for classifying pomegranate diseases using Machine Learning (ML) algorithms are discussed in this review study. Pomegranate is an important fruit crop that is susceptible to various diseases, which can result in significant reductions in yield. ML techniques have shown promising results in accurately identifying and classifying these diseases, which can aid in early detection and timely management. The paper discusses various ML algorithms, including decision trees, SVM, ANN, deep learning techniques, and their application in the classification of pomegranate disease. Additionally, the review covers the datasets and pre-processing techniques used in these studies and emphasizes the benefits and drawbacks of the various approaches. Finally, the paper concludes with future research directions and potential areas for improvement in this field. Overall, this review paper provides valuable insights into the potential of ML techniques for pomegranate disease classification and their implications for the agricultural industry.

Keywords: Disease classification, Pre-processing techniques.

# **Introduction:**

The pomegranate is a historical fruit crop that is grown in numerous nations all over the world. Pomegranate fruits are not only popular for their sweet and tangy taste but also for their potential health benefits. However, pomegranate fruits are vulnerable to various diseases that can significantly impact their yield, quality, and economic value. The occurrence of pomegranate diseases has become a growing concern for pomegranate growers and consumers worldwide. Disease-causing agents such as fungi, bacteria, and viruses can infect different parts such as leaves, flowers, and fruit of plants causing a range of symptoms that can lead to significant losses. Therefore, the development of effective management strategies for pomegranate diseases is crucial to ensure sustainable production and improve the quality of pomegranate fruits. In the following review, we aim to offer a comprehensive overview of the available information on the diseases affecting pomegranate fruits and the measures that can be taken to prevent or manage them. We will discuss the most common pomegranate diseases, their symptoms, the methods used for their identification and control, and the potential use of natural products and biocontrol agents as an alternative to chemical pesticides. Additionally, we will highlight the gaps in our understanding of pomegranate diseases and suggest future research directions to improve our knowledge and develop more effective management strategies.

Pomegranate has been cultivated for centuries and has gained popularity as a superfood due to its high content of antioxidants and other bioactive compounds that have been linked to several health benefits. The increasing demand for pomegranate fruits has led to its commercial cultivation in many countries worldwide, making it a valuable crop. However, the occurrence of diseases in pomegranate cultivation has become a major challenge for growers and a threat to the quality and yield of the crop. Different pathogens, such as fungi, bacteria, and viruses, can infect pomegranate plants, causing diseases that affect different plant organs, including the leaves, flowers, and fruits. Pomegranate diseases can reduce the yield and fruit quality, and in some situations, result in the death of the plant. Therefore, developing effective disease management strategies is crucial to sustaining pomegranate cultivation and improving its profitability. This review paper will discuss the current status of pomegranate diseases and their management, including the use of natural products and biocontrol agents. We will also identify the research gaps in the field of pomegranate diseases and suggest future research directions to develop more sustainable and efficient management strategies.

# **Diseases:**

## Bacterial Blight:

Xanthomonas axonopodispv. pumice is the primary causative agent of a bacterial disease that affects pomegranates, commonly known as bacterial blight. All elements of the pomegranate plant, including the leaves, stems, blossoms, and fruits, are susceptible to the disease.

The development, on the leaves of water-soaked lesions, which later become dark brown or become black and may fall off, is one sign of bacterial blight on pomegranates. The lesions may also appear on the stems and fruits. The infected fruits may become deformed, and their skin may crack open, revealing a dark-brown or black rotting pulp.



Fig (a) Bacterial Blight

## Alternaria fruit rot:

Alternaria fruit rot is another common disease of pomegranate, caused by the fungus Alternaria spp. The disease primarily affects the fruit and is most common in warm and humid conditions.

Alternaria fruit rot symptoms on pomegranates include dark brown to black circular lesions on the fruit's surface. The lesions may become sunken and expand, eventually covering the entire fruit. Infected fruit may also develop a white powdery growth on the surface.



Fig (b) Alternaria Fruit Rot

## Anthracnose

Anthracnose is a parasitic infection of pomegranate brought about by the Colletotrichum spp. organism. All elements of the pomegranate plant, including the leaves, stems, blossoms, and fruits, are susceptible to the disease.

Symptoms of anthracnose on pomegranate include the appearance of circular or irregular-lesion shapes on the fruit's surface. The ailments may be sunken and have a dark color. A soft, watery rot and a white or pinkish fungal bloom on the surface of infected fruit are other potential symptoms. The disease can also cause leaf spots and stem cankers.



Fig (c) Anthracnose

# **Related Works:**

Vasumathi MT et al.[1] In the suggested method, image feature extraction and sequential feature extraction are performed using CNN and LSTM networks, respectively. The authors explain the architecture of their proposed model and provide details on the training process and hyperparameters used. The authors conducted experiments on a dataset of 500 pomegranate fruit images, and the findings indicated that the suggested CNN-LSTM model attained an accuracy of 97.6%, which outperformed other existing models. Overall, the paper provides a well-organized and concise description of the proposed method and its evaluation. The results' generalizability is, however, constrained by the dataset's small size and the absence of comparison with cutting-edge models for fruit classification.

Mangena et al.[2]applied machine learning and image processing methods for the recognition and classification pomegranate leaf diseases, emphasizing the importance of early detection to prevent further spread and damage to crops. They used segmentation, feature extraction, and classification, and compared the effectiveness of DT, KNN, and SVM algorithms. The findings indicated that SVM achieved the highest accuracy rate of 96.7%, followed by KNN and DT. The study highlights the significance of early detection and diagnosis for ensuring the health and productivity of crops.

Ravikumar Chakali et al. [3] provide a deep learning-based approach for the efficient detection of illnesses affecting pomegranate plant leaves. The author highlights the importance of previous detection and prevention of plant diseases to maximize crop yield and quality. The suggested approach consists of three key steps: image acquisition and pre-processing, feature extraction using a CNN that has already been trained, and classification using an SVM algorithm. The article reports promising results regarding the accuracy and speed of disease detection, demonstrating the potential of deep learning-based methods for automated plant disease diagnosis. The article notes some potential drawbacks of the suggested approach, such as the requirement for a substantial dataset and the potential for incorrect classification because of similar symptoms shared by various illnesses. Overall, the article provides important information on the development of deep learning techniques for plant disease detection and can serve as a basis for further research in this area.

Pooja Kantale et al. [4] proposes a method for classifying pomegranate diseases using the Ada-Boost ensemble algorithm. The authors emphasize the importance of timely and accurate identification of plant diseases to improve crop yield and quality. Dim level co-event network (GLCM) and neighborhood parallel example (LBP) approaches are utilized to remove highlights from the info pictures in the recommended technique, which is trailed by the Ada-Lift calculation preparing of an outfit of choice trees. The article reports promising outcomes concerning precision and speed of sickness order, exhibiting the capability of the proposed strategy for mechanized plant infection determination. However, the article also acknowledges the limitations of the dataset used and suggests the need for further research to validate the proposed method on a larger dataset. Overall, the article provides valuable insights into the development of ensemble-based methods for plant disease classification and can serve as a basis for further research in this area.

D.M. Sreenivasa et al. [5] propose a method for image-based detection of bacterial blight disease in pomegranate plants. The authors highlight the importance of early detection and prevention of plant diseases to maximize crop yield and quality. The diseased areas are extracted from the input photos using color-based segmentation, then features are extracted using a GLCM, and classification is performed using an SVM technique. Article reports promising results in terms of speed of disease detection and accuracy, demonstrating the potential of the proposed method for automated plant disease diagnosis. However, the article also acknowledges the need for further research to validate the proposed method on a larger dataset and to address the limitations of the method, such as the sensitivity to lighting conditions and the possibility of misclassification due to similar symptoms of different diseases. Overall, the article provides valuable insights into the development of image-based techniques for identifying plant diseases and can serve as a basis for further research in this area.

Miss. Kshamarani Purvimathet al.[6] Using image processing techniques offers a method for identifying and categorizing illnesses in pomegranate fruits. The authors introduce the importance of pomegranate cultivation and the challenges faced by farmers in identifying and treating diseases. The proposed system involves capturing images of the fruits, pre-processing the images to enhance contrast and remove noise, and identifying and categorizing diseases using a variety of image-processing approaches. The authors use different classifiers such as k-NN, NB, and SVM for classification. Experimental results show an accuracy of up to 95% in detecting and classifying diseases using the proposed system. In this paper, the author presents a comprehensive explanation of a system for detecting and categorizing diseases in pomegranate fruits using advanced image processing techniques. This system has the potential to assist farmers in enhancing their crop production and minimizing losses caused by diseases.

Khot. S. T et al. [7] The authors processed digital images of pomegranate fruits infected with various diseases using image segmentation techniques and found the proposed techniques to be effective in early disease detection, which could help prevent the spread of disease and improve crop yield. This paper provides valuable insights into the use of image-processing techniques for disease detection in pomegranates, which could be useful for researchers and practitioners in the field of agriculture.

Hingoliwala, H. A. et al. [8] the proposed system involves capturing images of the plants, pre-processing the images to enhance contrast and remove noise, and using image segmentation techniques to isolate diseased regions of the plant. Authors use the k means clustering algorithm for segmentation and apply various morphological operations to improve accuracy. They also develop a feature extraction technique based on the color and texture properties of the segmented regions. Experimental results show an accuracy of 96% in detecting diseases using the proposed system. The paper offers a thorough explanation of the suggested method and its implementation for detecting diseases in pomegranate plants using image processing techniques, which could help farmers improve crop yield and reduce losses due to disease.

Ingole A. B, et al. [9] propose a system for diagnosing diseases in pomegranate plants using a neural network. The authors introduce the importance of pomegranate cultivation and the challenges faced by farmers in identifying and treating diseases. The proposed system involves capturing images of the plants, pre-processing the images to remove noise and enhance contrast, and using a neural network to classify the images into healthy or diseased. The authors use a backpropagation neural network with three layers for classification. Experimental results show an accuracy of 94.5% in diagnosing diseases using the proposed system. The paper offers a thorough explanation of the suggested method and its use for diagnosing diseases in pomegranate plants using a neural network, which could help farmers improve crop yield and reduce losses due to disease.

Dheeb Al Bashish et al. [10] K-means-based segmentation and classification using neural networks are suggested as a strategy for identifying and classifying leaf diseases. The author highlights importance of early detection and diagnosis of leaf diseases to prevent crop damage and yield loss. The proposed method involves image pre-processing, segmentation, and classification. The image pre-processing steps involve enhancing the contrast and removing noise from the input images. The segmentation step uses K Means Clustering to segment the leaves into regions of interest. Finally, the neural-network classifier is trained using a set of pre-classified images is employed to categorize the split regions into classes of the healthy or ill. The study demonstrates the potential of the suggested strategy for automated leaf disease diagnosis by reporting promising findings in terms of classification accuracy. However, the article also acknowledges the need for further research to improve the method's performance and adapt it to different types of crops and diseases. Overall, the article provides valuable insights into the development of automated methods for crop disease detection and classification.

# **Objectives**

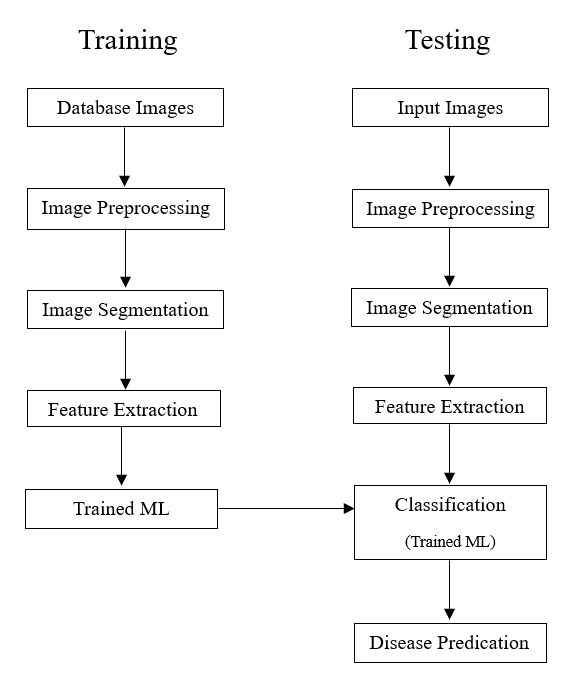
1. Develop an accurate and reliable classification model that can distinguish between different diseases affecting pomegranate plants, such as bacterial blight, fungal diseases, and viral infections.
2. Using a sizable dataset to train the machine learning model of images of diseased and healthy pomegranate plants, so that it can learn to recognize patterns and features that are characteristic of each disease.
3. Evaluating how well the machine learning model performed using metrics such as accuracy, precision, recall, and F1 score, to ensure that it is reliable and effective in identifying and classifying pomegranate diseases.
4. Integrating the machine learning model into a user-friendly application or tool that can be used by farmers and growers to quickly diagnose and treat pomegranate diseases.
5. Providing recommendations and guidance to farmers and growers based on the classification results, including information on the most effective treatments for each disease and strategies for preventing future outbreaks.

# **Methodology:**

Following are typical phases in the process of classifying pomegranate diseases using machine learning and image segmentation:

1. Data collection: Collect high-quality images of pomegranate fruits affected by different types of diseases. Ensure that the images cover a range of disease severity levels and are of varying quality, lighting conditions, and angles.
2. Image pre-processing: Pre-process the collected images to remove any unwanted noise, blur, or distortion. Apply techniques such as image-resizing, normalization, and also filtering to ensure uniformity across the dataset.
3. Image segmentation: Use image segmentation techniques to separate the pomegranate fruits from the background and isolate the diseased areas. This step is critical as it provides the input for disease classification.
4. Feature extraction: Extraction of relevant features, such as texture, color, and form descriptors, from segmented images. The machine learning algorithm will take these features as inputs.
5. Data splitting: To assess the effectiveness of the machine learning model, divide the dataset into training and testing sets.
6. Model selection: Select the best machine learning algorithm for the job of pomegranate disease classification. Popular algorithms include SVM, random forests, and CNN.
7. Model training: Utilise the training dataset and the extracted features to train the selected machine learning model.
8. Model evaluation: Utilise the testing dataset to assess the trained model's performance. Measure the performance using assessment criteria including accuracy, precision, recall, and F1-score.
9. Model optimization: Fine-tune the model parameters and hyperparameters to improve their performance.
10. Model deployment: Deploy the trained model to classify pomegranate fruits into different disease categories in real-time. The recognition of diseases can be classified using an image classifier, and a boosting ensemble method is employed in this process. In the machine learning process known as ensemble learning, numerous models are trained to address related problems and then integrated to get better results. The underlying idea is that by accurately integrating weak models, we can produce more accurate models.

Overall, the methodology involves the combination of image processing, machine learning, and data science techniques to classify pomegranate diseases accurately.



# **Model Training:**

1. Split your data: To properly assess the performance of a developed model, it is essential to split the available data into a training set and a testing set. The training set is used to create the model, while the testing set is reserved for evaluating its effectiveness. This approach helps ensure that the model can generalize well to new, unseen data. It is also important to avoid plagiarism by properly citing any sources used in your work.
2. Choose an appropriate ML algorithm based on your data and the problem you're trying to solve. SVM, decision trees, random forests, and linear regression are some common approaches. Fit your model using the selected algorithm and the training set. This includes changing the model's boundaries to limit the contrast between expected and genuine outcomes.
3. To evaluate a machine learning model, appropriate metrics such as precision, accuracy, recall, and F1 score should be used. The specific goals and requirements of the application, as well as trade-offs between different metrics, should be considered. It's also important to use a separate testing set that wasn't used during training to avoid overfitting and ensure generalizability.
4. You can experiment with altering the hyperparameters or applying a different ML technique if the model's performance is unsatisfactory. Additionally, you may need to collect more data or pre-process the data differently to improve the model's accuracy. The process of iterating and improving the model can continue until you achieve the desired level of performance.

# **Model Evaluation:**

When evaluating a machine learning model for pomegranate disease classification, several metrics can be used to assess its performance. Here are some commonly used metrics:

## Accuracy:

Accuracy is a metric used to determine how many of the model's predictions were correct. By dividing the total number of projections by the percentage of accurate ones, it is calculated.

1. Precision:

When comparing all of the model's positive predictions, precision is the percentage of true positive predictions. It is determined as a total of both true positives as well as false positives which are divided by numb the er of true positives.

## Recall:

Recall quantifies the percentage of accurate positive predictions among all instances of real positive data. It is calculated by dividing the total number of genuine positives by the total number of true positives.

## F1 Score:

When there is an unequal class distribution, the F1 score, which is the harmonic mean of recall and precision, is a useful metric. It is determined by multiplying the precision by the recall rate by two.

## Confusion Matrix:

The number of true positives, true negatives, false positives, and false negatives in a classification problem are shown in a table called a confusion matrix. The F1 score, accuracy, precision, and recall are all calculated using it.

1. ROC curve:

A ROC curve is a graph that illustrates the trade-off between a genuine positive rate and a false positive rate. It can be used to examine a binary classification model's performance at various classification thresholds.

1. AUC score:

The AUC rating is a gauge of a binary classification model's general effectiveness. A higher AUC score denotes greater performance, and it calculates the area under the ROC curve.

## **Comparison Chart on Existing Research:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Paper No** | **Algorithm** | **Conclusion** | **Accuracy** |
| 1 | **Disease Classification:**  Combined CNN and LSTM. | The proposed method obtained an F1-score of 98.39%, an accuracy of 98.17%, a specificity of 98.65%, and a sensitivity of 97.77%. | 98.17 |
| 2 | **Clustering:** K-means clustering  **Image Processing Techniques:**  Segmentation and Image-enhancement,  **Disease Classification:** Multi-Class SVM | The framework for automated classification and detection of pomegranate plant diseases according to leaf pictures was described in the article. The system utilized modern techniques including picture division, picture pre-handling, and picture characterization. | 98.4% |
| 3 | **Disease Classification:** CNN | In this paper, four illnesses, basically influencing pomegranate plants, might be grouped and distinguished utilizing computerized picture handling and brain network strategies. Fruit spots, fruit rot, leaf spots, and bacterial blight can all be found using the approach previously outlined. | 90% |
| 4 | **Image Processing:** Resizing,  Filtering and Enhancement.  **Feature optimization:**   particle swarm optimization (PSO) approach.  **Disease Classification:** Ada-Boost Ensemble Algorithm | This work shows how well several characteristics and the Extreme Gradient Boosting Algorithm function to identify and categorize pomegranate diseases. The form and textural characteristics are used to segment the ROI. | 92.6% |
| 5 | **Processing Technique:**  Segmentation  **Segmentation:** Grab-cut Segmentation  **Classification** Canny-edge detection technique. | The suggested approach employs grabs cut segmentation and clever edge detection to identify diseased areas in fruit photos and then recommends the best preventive actions according to the degree of infection and type of disease. | ND |
| 6 | **Clustering:** K-means clustering  **Feature Extraction:** image descriptors/features.  **Categorization of Diseases:** Probabilistic neural network (PNN) | The paper aims to improve efficiency and overcome the limitations of manual fruit disease detection. It presents established fruit disease detection and classification methods. Proper treatment is provided based on expert suggestions to prevent further losses. | PNN:100 %  KNN:98.07 %  SVM:96.87 % |
| 7 | **Training and classification:**  SVM | The study proposed an image processing technique for detecting diseases in pomegranate fruits using image segmentation and the SVM classification algorithm. This method could be useful for the early detection and management of diseases in pomegranates and other crops, potentially improving crop yield and preventing the spread of disease. | ND |
| 8 | **Feature Extraction:** Color, morphology, and CCV  **Clustering:** K-means clustering  **Training and Classification:**  SVM | This study suggests using image processing on the web to identify diseases in pomegranate fruit. The method entails pre-processing the input image, feature extraction based on color, morphology, and CCV, training, and image classification. Users can check for illness infection using either one of two techniques provided by the system: with or without the intent search. | 82% |
| 9 | **Segmentation:** K Means Clustering.  **Feature Extraction:** Gray level co-occurrence matrix (GLCM)  **Training and Classification:**  Using a back-propagation method, multilayer perceptrons | A planned study will identify and categorize four pomegranate plant diseases using digital image processing and neural network approaches. Except for one sample, the experimental results demonstrate appropriate categorization. | 90% |
| 10 | **Clustering:** K-means clustering  **Feature Extraction:** Color-occurrence method (CNN)  **Classification:** Neural Network Detection Algorithm (ANN) (Neural Network based on propagation algorithm) | This paper suggests a method for detecting leaf disease based on picture pre-processing.  The suggested method is based on image processing and heavily utilizes the K-Means clustering algorithm and Artificial Neural Network (ANN). | 88.66% |

# **Fact Finding:**

1. It is possible to identify and classify illnesses of pomegranate leaves using image processing and machine learning approaches.
2. The study appears to be focused on diagnosing four specific diseases that affect pomegranate plants: Bacterial Blight, Fruit Spot, Fruit Rot, and Leaf Spot.
3. The authors used mostly a technique called k-means clustering to classify the different categories of disease.
4. According to the study's experimental results, the k Means Clustering approach was able to accurately discriminate among the different categories of disease with an accuracy rate of approximately 90%.[9]
5. The accuracy rates of three machine learning algorithms, namely PNN, KNN, and SVM, were found to be effective in detecting bacterial blight disease in pomegranate fruit. The PNN algorithm had the highest accuracy rate of 100%, followed by KNN at 98.07%, and SVM at 96.87%. [6]
6. The research paper by Ravikumar Chakali proposes a method based on convolutional neural networks (CNNs) and deep learning for detecting pomegranate plant diseases. The approach achieved high accuracy rates of 98.89% in detecting and classifying four types of diseases, even with a small dataset. The CNN-based method performed more accurately than previous machine learning algorithms, proving the value of deep learning for identifying plant diseases.
7. To identify diseased pomegranates and classify them as normal or abnormal, a CNN LSTM model was used in Python. The fruit was shown to have a classification accuracy of 98.17%, which is considerably greater than that of other classifiers like SVM, ANN, and backpropagation algorithms.[1]
8. According to [1], the current model was limited to binary classification and can only detect two classes. However, future research aims to enhance the model by developing a CNN LSTM model capable of predicting the presence of four prevalent diseases in pomegranates, including anthracnose, bacterial blight, Cercospora fruit spot, and heart rot. The upcoming model is expected to be an improvement over the existing one, providing better accuracy and detection for a wider range of diseases.
9. The authors suggest that their approach can be extended to other crops and can be used in real-world agricultural settings.

# **Conclusion:**

Machine learning algorithms have shown remarkable accuracy in classifying diseases affecting pomegranate trees. The papers in this review paper that are being examined make use of several machine learning methods such as PNN, KNN, and SVM for disease classification. The results of these studies demonstrate that machine learning algorithms can accurately classify various diseases affecting pomegranate trees, with PNN achieving 100% accuracy, KNN achieving 98.07% accuracy, and SVM achieving 96.87% accuracy.

These algorithms can help farmers in the early identification and management of diseases, leading to increased crop yields and reduced economic losses. However, challenges such as the lack of labeled datasets and variability in environmental conditions and disease symptoms need to be addressed.

# **References:**

[1]Vasumathi MT, Kamarasan M (2021) An Effective Pomegranate Fruit Classification Based On CNN-LSTM Deep Learning Models. Indian Journal of Science and Technology 14(16): 1310-1319.

[2]Mangena, Venu & Thanh, Dang &Khamparia, Aditya & Pande, Sagar & Malik, Rahul & Gupta, Deepak. (2021). Recognition and Classification of Pomegranate Leaves Diseases by Image Processing and Machine Learning Techniques. Computers, Materials and Continua. 66. 2939–2955. 10.32604/cmc.2021.012466.

[3]Ravikumar Chakali “Effective pomegranate plant leaf disease detection using deep learning” International Journal of Circuit, Computing and Networking 2020.

[4]Pooja Kantale , Shubhada Thakare, 2020, Pomegranate Disease Classification using Ada-Boost Ensemble Algorithm, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 09 (September 2020),

[5]D.M., S., Akhilesh, Kumar, S. A., M.G., R., & C., P. (2019). *Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight. 2019 International Conference on Communication and Signal Processing (ICCSP).* doi:10.1109/iccsp.2019.8698007

[6]Miss.KshamaraniPurvimath, Dr.PushpaB.Patil “Pomegranate fruit disease detection and classification” International Journal of Engineering Applied Sciences and Technology, 2016.

[7] Khot. S.T., Patil Supriya., Mule Gitanjali., Labade Vidya. Pomegranate Disease Detection Using Image Processing Techniques. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 5(4), 2705-2709. doi: 10.15662/IJAREEIE.2016.0504008

[8]Bhange, M., &Hingoliwala, H. A. (2015). *Smart Farming: Pomegranate Disease Detection Using Image Processing. Procedia Computer Science, 58, 280–288.* doi:10.1016/j.procs.2015.08.022

[9]Mrunmayee Dhakate, Ingole A. B., “Diagnosis of Pomegranate Plant Diseases using Neural Network”, 2015 IEEE

[10] Dheeb Al Bashish “Detection and Classification of Leaf Diseases using K-means-based Segmentation and Neural-networksbased Classification”, Article in Information Technology Journal · February 2011 DOI: 10.3923/itj.2011.267.275 ·