IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 11, Issue 1, January 2021, ||Series -I|| PP 44-51

Automatic Leaf Disease Detection Using Neural Network

Poonam Patil[1], Pallavi Patil[2]
[1](*Computer Engineering, GOVT. Residence women Polytechnic, Tasgaon, India)*
[2](*Computer Engineering, Latthe Education Society's Polytechnic, Sangli, India)*
Received 15 January 2021; Accepted 31 January 2021

Abstract:
In agriculture research, automatic leaf disease detection is essential research topic as it may prove benefits in monitoring large fields of crops, and thus automatically detect symptoms of disease as soon as they appear on plant leaves.To detect disease there are various steps- need to acquire image and preprocess image. Then image segmentation, feature extraction is done in order to go for statistical analysis. The proposed work includes image filtering and convert RGB image to CIE LAB color component. K-mediod technique is used for image segmentation. Lastly by using the neural networks disease is identified and classified.
Key Word: Leaf disease, Image processing, CIELAB color model, SGDM Matrix, Color Co-occurrence Method, k- medoids, Neural Network.

I. INTRODUCTION

India is a cultivated country. Farmers can select appropriate crops for his land from the huge diversified crops. Multiple research work has been done to identify diseases on various crops using advanced and automation techniques. For this various infected images have been taken from digital camera or drone camera or smart phones. These captured images are processed and classified undergoing trained and test sets. Various advanced computing techniques and image processing techniques is used for this to identify disease from the crops and leaves.

A. Image Analysis Can be useful For the Following Purposes:

1. To identify diseased leaf, stem, fruit.
2. To measure affected area by disease.
3. To find the boundaries of the affected area.
4. To find out the color of the affected area.
5. To determine size & shape of leaf.
6. To identify the Object correctly. Etc.

Disease management is a challenging task. Mostly diseases are visible on leaves and stems of plants. Plants contain various cells, different patterns and different textures which are very complex. As plants contain lots of complexity, the quantification of these diseases, pests and features have not yet studied. Hence there is huge demand for the study of specific patterns and image pattern analysis.

B. Different Types Of Leaf Spot Diseases:

- Bacterial
- Fungal
- Viral

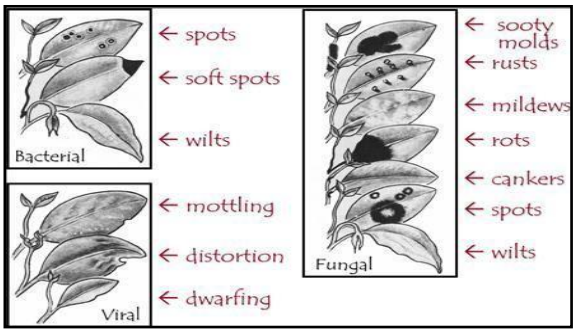

Fig. 1: Various types of diseases [11]

Fungi, bacteria and viruses are the major diseases that are spread on leaf of plants. Fungi are identified by observing their structures and forms. Major importance is given to their reproductive structures. Bacteria are measured more primitive than fungi and generally have simpler life cycles. It is said that bacteria has single cell and it increases by dividing itself into two cells by the process named binary fission. Viruses are minute particles. It consists of protein and genetic material [9]. Thousands of images are captured or being generated with single experiment in biological science. There image scan be required for further studies like classifying lesion, scoring quantitative traits, calculating area eaten by insects, etc. all the above tasks are done manually or some are done by using various software packages.

It is not only incredible amount of work but also suffers from two major issues: excessive processing time and subjectiveness rising from different individuals. In order to carry out the high throughput experiments, biologist require computer software that automatically does process. This software extracts and analyze the important content that is available. In this concept, image processing is considered to be very important [1]. Hence this paper describes a survey for studying leaf diseases using different image processing techniques

II. LITERATURE REVIEW

Few papers are expressing the leaf disease detection using different types of methods suggesting the different ways for implementation as expressed here:

[2] In this paper consists of two phases to identify the affected part of the disease. At the beginning, we use Image segmentation by using Edge detection, and further we go through analysis on the image and finally disease classification has done. This work looks for Operating System to examine the provided input images using RGB (Red, Green, Blue) pixel count values technique and disease wise identification and also using homogenization techniques like Sobel and canny as an edge detection for identification of affected parts of the leaf and as a result it will recognize disease as a output. [3] For this paper detection of leaf diseases has been used methods are: 1) Use of K-Mean clustering for infected object identification; 2) Use of colour co-occurrence for texture analysis; 3) Use of NNs, for identification and classification of different types of diseases. Moreover, the presented scheme classifies the plant leaves into infected and not-infected classes. [4] colour models like HIS, YCbCr and CIELAB are differentiate by their effects and this effects are used in process of disease spot detection.[5] In this paper Support vector machines are a set of related supervised learning method used for classification and regression. SVM classifier are mainly used for improving the detection accuracy. [6] Otsu method is used to segmentation of leaf regions. Before that, the analysis of image segmentation has done. In the HSI colour system, H component was chosen to segment disease spot in order to reduce the disturbance of illumination changes and the vein. After that, disease spot regions were segmented with the help of Sobel operator just to examine disease spot edges. At the end, plant diseases are graded by calculating the quotient of disease spot and leaf areas. [7] This paper wills two techniques for feature extraction and comparison of two techniques. Otsu Threshold: thresholding creates binary image from grey level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one. K-Means clustering is all alone learning task where one seeks to identify a finite set of categories termed clusters to describe the data. [8] This paper express the segmentation composed in image conversion to HSV colour space and fuzzy c- means clustering in hue-saturation space to distributes several pixel classes. These classes are further integrate at the interactive stage into two final classes, where one of them determines the searched diseased areas. [13] This paper presents a colour image segmentation method, which divides colour space into clusters. Here we are also try to compare efficiency of available algorithm for segmentation of grey and noisy images. Different from K-Means algorithm, K-Medoids is not much sensitive for abnormal data. This particular algorithm is suitable for noisy data. It is seems that, the segmented images are highly dependent on the number of segments or centroids.

III. OBJECTIVES OF RESEARCH

The major objectives of this research are follows:

1. To collect image data sets of various common leaves diseases.
2. Identification of infected area using clustering algorithm.
3. Evaluation of the color co-occurrence method for disease detection in leaves.
4. Develop NNs strategies classification of leaves.

The objective behind this research work is to progress the performance of disease detection technique. The image data of the leaves selected for this study would be collected. Algorithms based on image processing techniques for image preprocessing, image segmentation, feature extraction, statistical analysis and classification would be designed. Comparison is also presented. Thus, the proposed algorithm was tested on five diseases are: Early scorch, cottony mold, and late scorch, brown spotand bacterial-fungal.

IV. THEORETICAL BACKGROUND

A. The Basic Approach Procedure:

The overall concept of this work for any vision related algorithm of image classification is almost the same. First, the digital images are acquired from the plants using a digital camera. Then image processing techniques are applied to the acquired images to take out useful features that are essential for further analysis. After that, several analytical selective techniques are used to classify the images according to the specific problem. Figure 3 show the basic procedure of the proposed detection algorithm in this research.

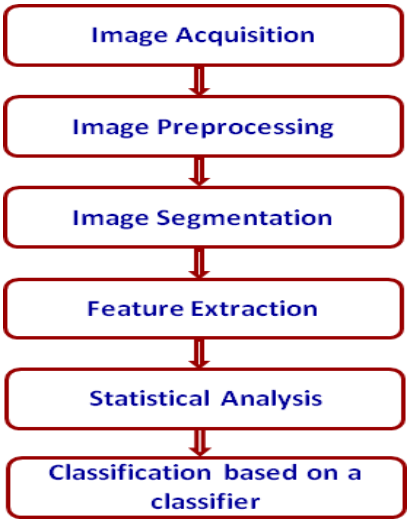

Fig. 2: The basic procedure based disease detection solution

B. Image Acquisition and Preprocessing:

Image Acquisition is nothing but the gathering of various images of plant leaves with the help of camera. These images are taken from various location. Images can also be collected from different agriculture centers or units. The images which have input are always not satisfactory regardless of what image acquisition devices are adopted.

This collected image may contain noise in it or there is possibility that the interested region of image is not clear. There is need to remove this noise for better result. This can be done by applying various preprocessing techniques.

There are three steps included in preprocessing phase: clipping, smoothing and enhancement [7]. But this process should be done with care because noises degrades the quality of image.So, the image with low quality must be smoothed by filter [7]. Different kinds of noises exist in an image and variety of noise reduction techniques is available to perform de- noising.

Application decides the selection procedure of de-noising algorithm. Median filter performs better with salt and pepper noise by choosing appropriate threshold. An image has salt and pepper noise will have dark pixels (black dots or pepper) in bright region and bright pixels (white dots or salt) in dark region. An effective method to remove this type of noise involves the use of median filter [5]. Median filter is non-linear filter. The median filter is to find the median value by across the window, replacing each entry in the window with the median value of the pixel [14]. It‘s best in removing salt and pepper noise and impulse noise. The function of median filter is to erase black dots. These dots are called as pepper. Similarly, filling white holes inside the image is called as salt. It simply replaces each pixel values by the median of the intensity level in the neighborhood of the pixel [6].

C. Image Segmentation (region of interest):

As per regions, the images will be segmented into various parts. Image segmentation is to divide the image into same meaningful regions. In a few words, image segmentation means to separate the object from background for further processing [7]. In this step the images are segmented using k- medoids clustering methods. K-medoids clustering is partitioning based clustering method. K-medoids or PAM (Partition around medoids): Every cluster is denoted by any one of the objects in the cluster. K-medoids more robust than k-means in the existence of noise and outlines; because a medoids is less influenced by outlines or other extreme values than a mean.

K-medoids Algorithm [12]:
Consider Input:
Number of cluster which is to be partitioned= k
Number of objects = n
Output: A set of k cluster that minimize the sum of the dissimilarities of all the objects to their nearest medoids.

Steps:
I. Choose k objects arbitrarily
II. Repeat,
a. Assign each remaining object to the cluster with the nearest medoids;
b. Randomly select a non- medoids object;
c. Compute the total cost of swapping old medoids object with newly selected non-medoids object.
d. If the total cost of swapping is than zero, than perform that swap operation to form the new set of k- medoids.
III. Until no change

Calculation of distance is done and each data object is connected the nearest mediod it has. Cost is calculated using Manhattan distance. Costs to the nearest medoid are shown bold in the table [17]. The Manhattan distance D is D-sum (abs(x-y))where cost between any two points is round using formula $costx$, $c= {}^{d}_{i=1}|x_i - c_i|$ where x is any data object, c is the medoid, and d is the dimension [12]. K-means and k-medoids – both the methods find out clusters from the image. K-means drawback is sensitivity to noisy data outlines. Compared to this, k-medoids is not sensitive to noisy data, outlines and effective for gray scale too.

An example of the output of K-medoids clustering for a leaf infected with *early scorch* disease is shown in Figure 3. It is observed from Figure 3 that cluster 3 contains infected object of early scorch disease. Furthermore, clusters 2 and 4 contain the intact parts of leaf, although they are distinct from each other. The black background from cluster 1 is removed firstly. Finally, the image in facilitates the segmentation procedure followedin K-medoids algorithm.

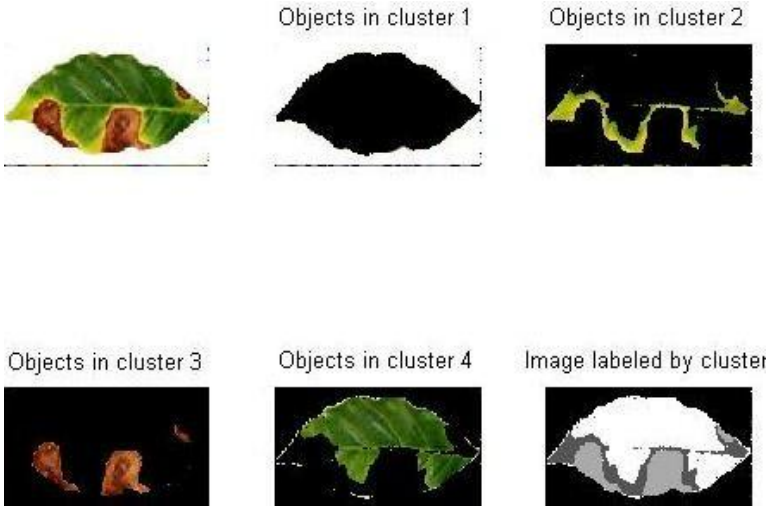

Fig. 3: Output of K-Medoids clustering for a leaf that is infected with early scorch disease

D. Extraction of Features And Statistics Analysis:

The input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a compact representation set of features. Feature extraction includes the extraction of features from input data. It means features are identified or taken out for further process. This features must be chosen carefully so that it will extract appropriate information to perform task [7].

The color co-occurrence texture analysis method is developed by the Spatial Gray- level Dependence Matrices (SGDM). The gray Level Co-occurrence Methodology (GLCM) is a statistical way to describe shape by statistically sampling the way certain gray level occurs in relative to other gray levels. For a position operator p, we can define a matrix Pij that counts the number of times a pixel with gray level I occurs at position p from a pixel with gray- level j. The SGDMs are represented by the function P (i, j, d, Θ) where I represent the gray level

of the location (x,y) at an orientation angle of Θ. The reference Pixel at image position (x, y) is shown as am matrix. All the neighbor from 1 to 8 are numbered in a clockwise direction Neighbors 1 and 5 are located on the same plant at a distance of l and an orientation of 0 degree. In this research, a one pixel offset distance and a zero degree orientation angle was used [3]. After the transformation, we calculated the feature set for H and S, we dropped (I) since it does not give extra information. However, we use GLCM function in Matlab to create gray-level co-occurrence matrix; the number of gray levels is set to 8, and symmetric value is set to-true‖,and finally, offsetisgivena-0‖value[3].Properties of Spatial Gray- level Dependence Matrices (SGDM) are compound for the content of the images are described below:

Table 1: Properties of SGDM [18]

Property	Description	Formula
Contrast	Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Range = [0 (size(GLCM,1)-1)^2] Contrast is 0 for a constant image.	$\sum_{i,j} i-j^2 P(i,j)$
Correlation	Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1] Correlation is 1 or -1 for a perfectly positively or negatively correlated image. Correlation is NaN for a Constant.	$\sum_{i,j} \frac{i-\mu ij-\mu jP(i,j}{\sigma\sigma ij}$
Energy	Returns the sum of squared elements in the GLCM. Range = [0 1] Energy is 1 for a constant image.	$\sum_{i,j} P(i,j)^2$
Homogeneity	Returns a value that measures the closeness of the distribution a\of element in the GLCM to the GLCM diagonal. Range = [0 1] Homogeneity is 1 for a diagonal GLCM.	$\sum_{i,j} \frac{P(i,j)}{1+\lvert i-j\rvert}$

E. Classification Based On Classifier:

In this paper, neural networks are used in the automatic detection of leaves disease. Neural network has well known technique of successful classifier. Hence neural network is chosen as classification tool for many applications. The dataset for training and validation processes consists of two parts; the training features set which are used to train the NN model; whilst a testing features sets are used to verify the accuracy of the trained using the feed-forward back propagation network. In training part, the connection weights are updated till it reaches stated iteration number or till any error occurs. The mean square error (MSE) method is used to measure correctness.

V. Basic steps of the proposed algorithm

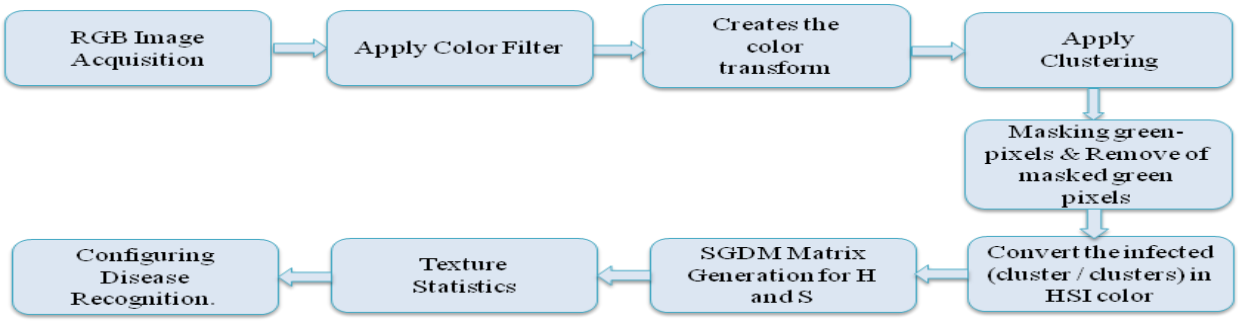

Fig. 4: Disease detection System Architecture

Algorithm:

1. RGB image acquisition.
2. Apply Color Median Filtering byMedian filter.
3. Create the color transformation structure RGB to CIELAB Color structure.
4. Apply K-medoids clustering.

5. Masking green-pixels & Remove of masked green pixels.
6. Convert the infected (cluster / clusters) in HIS Color.
7. SGDM Matrix Generation for H and S.
8. The GLCM function to calculate the Texture features Statistics.
9. Configuring Neural Networks for Recognition.

VI. EXPERIMENTAL RESULTS AND OBSERVATIONS

A. Input Data Preparation and Experimental Settings:

In this experiment, two main dataset were generated, namely: (i) Training texture feature data, and (ii) Testing texture feature data. Each row had a unique number (1 to 6) which represented the class (i.e., the disease) of the particular row of data. 1st, represented early scorch disease infected leaf. 2nd, represented Cottony mold disease infected leaf. 3rd, represented Late scorch disease infected leaf. 4th, represented Brown spot disease infected leaf. 5th, represented Bacterial-Fungal disease infected leaf, and 6th, represented normal leaf. Then, a software program was written in MATLAB that would take in .mat files representing the training and testing data, train the classifier using the ―train dataset‖, and then use the ―test dataset‖ to perform the classification task on the test data. Consequently, a Matlab routine would load all the data files (training and testing data files) and make modifications to the data according to the proposed model chosen [3].The architecture of the network used in this study was as follows. A set of 10 hidden layers in the neural network was used with the number of inputs to the neural network (i.e. the number of neurons) is equal to the number of texture features listed above. The total output is 6. This output consists of number of classes. Each class represents a disease for normal leaf. Those diseases are early scorch, cottony mold, late scorch, brown spot, Bacterial-Fungal. Here the neural network uses feed forward back propagation and also Mean Square error. Total 1000 iterations were considered.

B. Experimental Results:

Table 2 shows the results for NN classification strategy for testing. The results were drawn with NN classifier. In Color feature HS get the highest overall classification accuracy, in which it achieved an overall accuracy of 96% compared to the 94% accuracy achieved in [3]. Also, Figure 5 shows a graph that representing the percentage classification of various disease of all the color features models shown in Table 2.

Model	Color Features	Early scorch	Cottony mold	Late scorch	Brown spot	Bacterial-Fungal	Normal	Overall average
M1	H	92	92	89	90	93	98	92.33
M2	S	93	90	85	89	85	98	90
M3	I	92	89	87	88	86	99	90.16
M4	HS	95	96	89	96	95	100	95.17
M5	HIS	87	85	80	79	84	99	85.67

Table2: Percentage classification of various diseases

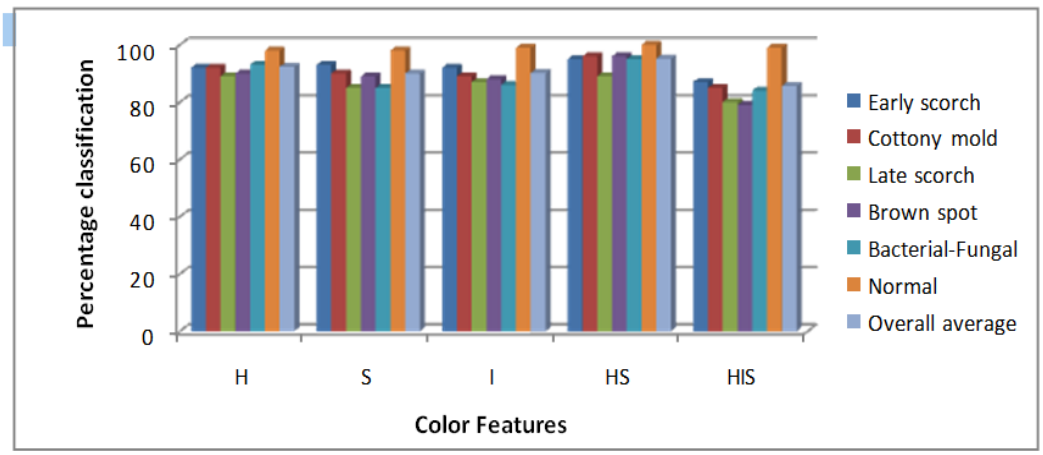

Fig. 5: Percentage classification of various diseases

It can be implied from Table 2 that Model M4 which has used only the H and S components in computing the texture features, has emerged as the best model among the various models. The numbers of leaf samples that were classified into each of the five tested categories using HS model with specific threshold value are shown in Table 3 and Figure 6. It is observed from Table-3 that only two samples from brown spot leaves were misclassified. Similarly, in the case of bacterial- fungal images, only three test images from the class were misclassified.

From species	Early scorch	Cottony mold	Late scorch	Brown spot	Bacterial-Fungal	Normal	Accuracy
Early scorch	25	0	0	0	0	1	100
Cottony mold	0	24	0	1	0	1	96
Late scorch	0	0	25	0	0	0	100
Brown spot	0	0	0	24	1	0	96
Bacterial-Fungal	0	1	1	0	23	0	92
Normal	0	0	0	1	2	23	92

Table 3: Classification results per class for neural network.

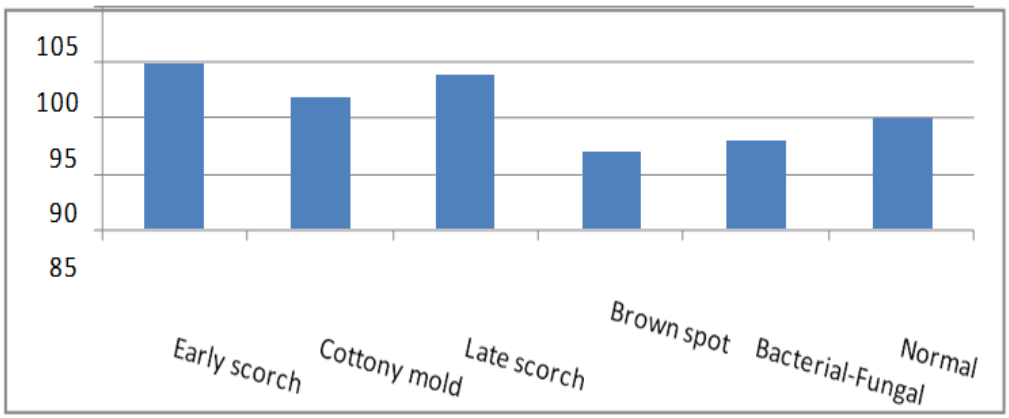

Fig. 6: Classification results per class for neural network with back propagation.

VII. CONCLUSION AND FUTURE WORK

This paper describes the system that identifies the disease on plants leaf using technique named image processing. For filtering Median filter performs better with salt and pepper noise. In the Color model, CIELAB color model accurately detects disease. Any background, type of leaf, type of disease spot and camera flash do not affect the results. The k-medoids algorithm is working for gray scale images and better performs for large databases not sensitive to noisy data and outliers, too. K-medoids performs reasonably better than the K-Means algorithm. The applications of Median filter, CIELAB color model, and clustering and texture analysis have been formulated for clustering and classification of diseases that affect plant leaves. Neural networks increases the recognition rate.The experimental results show that the proposed approach is important in accurately detecting leaf disease. Disease recognition is main aim of proposed approach.

For future research, Work can be extended for development of hybrid algorithms such as other clustering method and NNs in order to improve the recognition rate of the final classification process. Further to needed to compute amount of disease present on leaf.

REFERENCES

[1] Jayamala K. Patil, Raj Kumar, ―Advances In Image Processing For Detection of Plant Diseases ‖, JABAR, 2011, 2(2), 135-141.

[2] P.Revathi, M.Hemalatha, ― Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques‖, ISBN, 2012, 169-173, IEEE.

[3] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, ― Fast and Accurate Detection and Classification of Plant Diseases‖, IJCA, 2011, 17(1), 31-38, IEEE-2010.

[4] Piyush Chaudhary, Anand K. Chaudhari, Dr. A. N. Cheeran and Sharda Godara, ― Color Transform Based Approach for Disease Spot Detection on Plant Leaf‖, IJCST, 2012, 3(6), 65 -70.

[5] S. Arivazhagan, R. NewlinShebiah, S. Ananthi, S. Vishnu Varthini, ―Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features‖, CIGR, 2013, 15(1), 211-217.
[6] Chanchal Srivastava, Saurabh Kumar Mishra, Pallavi Asthana, G.R. Mishra, O.P. Singh, ―Performance Comparison Of Various Filters And Wavelet Transform For Image De-Noising‖, IOSR-JCE, 2013, 10(1), 55-63.
[7] MrunaliniR. Badnakhe, Prashant R. Deshmukh, ― Infected Leaf Analysis and Comparison by Otsu Threshold andk-Means Clustering‖, IJARCSSE, 2012, 2(3), 449-452.
[8] Mr. Salem Saleh Al-amri, Dr. N.V. Kalyankar and Dr. Khamitkar S.D, ―A Comparative Study of Removal Noise from Remote Sensing Image‖, IJCSI, 2010, 7(1).
[9] Prof. Sanjay B. Dhaygude, Mr.NitinP.Kumbhar, ―Agricultural plant Leaf Disease Detection Using ImageProcessing‖ IJAREEIE, 2013, 2(1), 599-602.
[10] Dheed AlHiary, S. Bani-Ahmad, M. Braik, ―A Framework for Detection and Classification of Plant Leaf and Stem Diseases‖, IEEE, 2010, 113-118.
[11] S. Ananthi, S. Vishnu Varthini, ―Detection and classification of plant leaf diseases‖, IJREAS, 2012, 2(2), 763-773.

Poonam Patil. "Automatic Leaf Disease Detection Using Neural Network." *IOSR Journal of Engineering (IOSRJEN)*, 11(01), 2021, pp. 44-51.