IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Plant Leaf Disease Detection Using Image Processing

Onkar Heddurshetti
Electrical & Electronics Engineering
Hirasugar Institute of Technology, Nidasoshi, India

Hemalata Zinage
Electrical & Electronics Engineering
Hirasugar Institute of Technology, Nidasoshi, India

***Abstract*—Agricultural yield is something on which country's economy highly depends. This is one of the reason that plant disease detection is very important in agriculture field, as having plant diseases are quite common. If proper attention is not given in this area then it causes serious effects on plants and due to which respective product quality, quantity or productivity will reduce. Presently, the plant leaf disease can be detected physically by farmer by proper inspection. Farmers should go to the field and should identify leaves affected by the disease. Then we should find the solutions physically. This manual process of identifying plant leaf disease consumes more time and will give accuracy of about 60%.**

**In the proposed system, a real time crop image capturing facility is provided and there is 99% accuracy in detection of disease. The information about the leaf disease can be sent to the farmers and climatic parameters such as moisture, PH, humidity, temperature etc. can be measured and monitored for the better growth of the plant. This process consumes less time to detect and gives 99% accuracy. This helps farmer to grow quality product.**

***Keywords— Image processing, Segmentation, K-Mean clustering.***

## I.INTRODUCTION

India is rapidly developing country and agriculture plays a vital role for the country's development. Hence, agricultural yield is very important in the development of the country's economy. In India more than 65% population depends on agriculture. Plants are the basic source for supply of energy for human body. Productions based on agriculture get easily affected by various plant diseases. The crop losses due to diseases are approximately 10 to 30%.Farmer identifies the symptoms of diseases manually by calling the experts for detecting the diseases, but this is also time consuming way and the cost fetching process. The naked eye observation of experts is the main approach adopted in practice for detection and identification of plant diseases. But, this requires continuous monitoring of experts which might be more expensive in large farms. Further, in some developing countries, farmers may have to go long distances to contact experts, this makes consulting experts too expensive and time consuming. Early disease detection is not possible. If growth of disease exceeds its primary stage, then it will destroy complete crop and it's very effective to the farmer. They can loss their complete one year struggle and money too. Image processing techniques could be applied on various applications as follows:

- To detect plant leaf, stem, and fruit diseases.
- To quantify affected area by disease.
- To find the boundaries of the affected area.
- To determine the color of the affected area
- To determine size & shape of fruits.

The old and classical approach for detection and recognition of plant diseases is based on naked eye observation, which is very slow method also gives less accuracy. In some countries, consulting experts to find out plant disease is expensive and time consuming due to availability of expert. Irregular checkup of plant results in growing of various diseases on plant which requires more chemicals to cure it also these chemicals are toxic to other animals, insects and birds which are helpful for agriculture. Automatic detection of plant diseases is essential to detect the symptoms of diseases in early stages when they appear on the growing leaf and fruit of plant. Here introduces a MATLAB based system in which we focused on both leaf & fruit diseased area and used image processing technique for accurate detection and identification of plant diseases.

Following fig.1 shows that how diseases on cotton plant reduces the productivity. There is 20 to 25% of cotton loss due to diseases on plant.

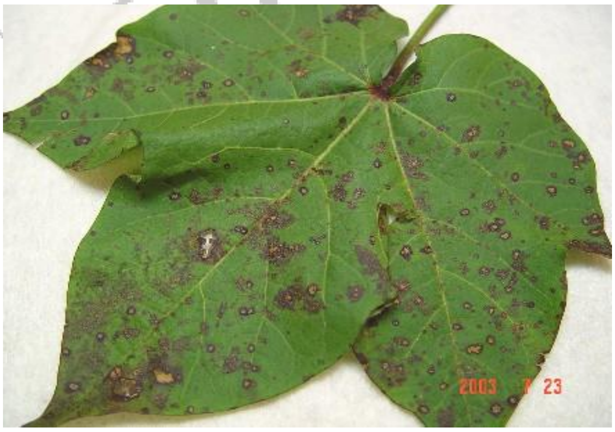


Fig.1. Disease of cotton plant

The MATLAB image processing starts with capturing of digital high resolution images. Healthy and unhealthy images are captured and stored for experiment. Then images are applied for pre-processing for image enhancement. Captured leaf & fruit images are segmented using k-means clustering method to form clusters. Features are extracted before applying K-means and SVM algorithm for training and classification. Finally diseases are recognized by the system. Most of the plant diseases reduce the food productivity and also more amount of losses to the farmers. The human eye cannot detect some of these diseases. Even if diseases are identify by human, it is difficult to detect and differentiate the plant diseases. So in modern days, Technologies are came to identifying and classifying a many plant diseases in short time. This technology are used to detect the leaf diseases at early stage, helps in loss reduction. The image processing techniques are necessary in this detection of leaf diseases automatically. Image processing is easiest technique to resolve the all the problems what we are facing at the time of production of food and after this technique we will get more product and healthy food to the people. Due to this image processing farmers will away from huge amount of loss production and farmers will get result before the time of production that is early stage leaf diseases detection.

## II.METHODOLOGY

The proposed model aims in the design and development of an automated system for early detection of diseases on bean crop and send diseased information to the farmer. This technique helps for classification and identification of various diseases in bean crop.

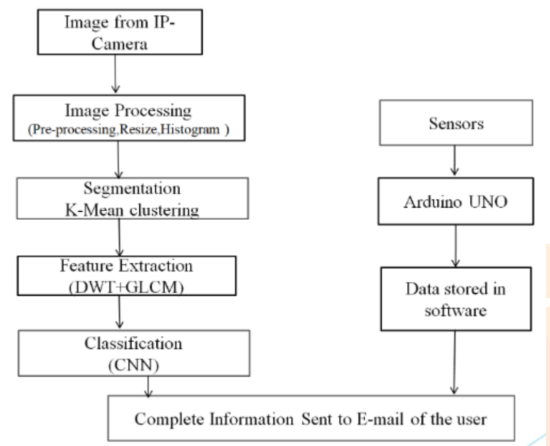



Fig 2. Block diagram of the model.

### *A. Test image*

First we need to select the plant which is affected by the disease and then collect the leaf of the plant and take a snapshot of leaf and load the leaf image into the system to detect the disease of that plant.

Fig 3. Test image

### *B. Pre-processing*

1) Image Scaling: Image scaling is applied because not all of the training and the testing images are of the same size. Some of these images are of big size that can cause a problem in the implementation including out of memory. Reducing the image size can hasten the processing time. Therefore, all the image sizes will be set to [11].
2) Min- Max Linear Contrast Stretch: The input images may have blurred or low contrast, which can affect the detecting process. That is why using min-max linear contrast stretch is necessary to improve the quality. This is because it reassigns the lowest and the highest values of the data into new set of values that utilize the full range of available brightness values [11]..

3) Colour transformation structure: The colour transformation structure is adopted instead of using greyscale images because of its ability to improve the image analysis process that would result in a better segmentation results. Although RGB is a colour model, HSI model is chosen as it presents colour information in different ways, which makes some calculations regarding feature extraction more convenient. Furthermore, the HSI model is based on human observation. Since different plant diseases have different colours, HSI model will be used. First, the RGB image representation (Red, Green and Blue) will be converted into (HSI) representation (Hue, Saturation and Intensity). Hue is the dominated colour as seen by the human eyes. Saturation on the other hand refers to the amount of white light that is added to the hue and Intensity refers to light amplitude. Then, the Hue component will be used and the other two components will be dropped; as they do not give any additional information. The Hue component is chosen between the other two components, as it is less sensitive to lighting variations [11].

### *C. Resize*

Image interpolation occurs when you resize or distort your image from one pixel grid to another. Image resizing is necessary when you need to increase or decrease the total number of pixels, whereas remapping can occur when you are correcting for lens distortion or rotating an image. Zooming refers to increase the quantity of pixels, so that when you zoom an image, you will see more detail. Interpolation works by using known data to estimate values at unknown points. Image interpolation works in two directions, and tries to achieve a best approximation of a pixel's intensity based on the values at surrounding pixels. Common interpolation algorithms can be grouped into two categories: adaptive and non-adaptive. Adaptive methods change depending on what they are interpolating, whereas non-adaptive methods treat all pixels equally. Non-adaptive algorithms include: nearest neighbor, bilinear, bicubic, spline, sinc, lanczos and others. Adaptive algorithms include many proprietary algorithms in licensed software such as: Qimage, Photo Zoom Pro and Genuine Fractals. Many compact digital cameras can perform both an optical and a digital zoom. A camera performs an optical zoom by moving the zoom lens so that it increases the magnification of light. However, a digital zoom degrades quality by simply interpolating the image. Even though the photo with digital zoom contains the same number of pixels, the detail is clearly far less than with optical zoom.

### *D. Histogram*

Histogram image processing is the act of modifying an image by controlling the parameters of the image intensity values. Intensity values for an image can be measured and mapped onto a histogram, representing either overall intensity or color intensity within a single color channel. A user can change the appearance of a picture through histogram image processing by interacting with and modifying that histogram displaying intensity values. A color's intensity throughout an image can be calculated, and the calculated values can be mapped onto a histogram. By changing the shape of that histogram or altering the midpoint of it, intensity values can be remapped, changing the appearance of an image. It's possible to modify a histogram for a single color channel, changing how a single color is treated throughout an image. Along with intensification or attenuation of selected colors, overall contrast within an image can also be increased or decreased.

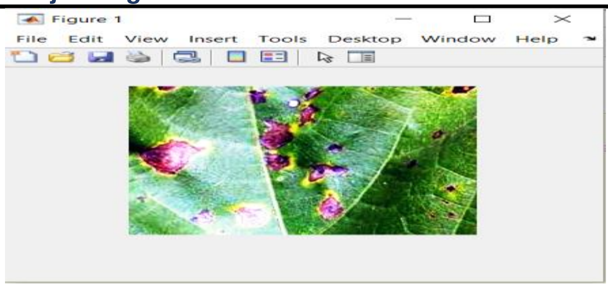



Fig.4. Pre-processed image

*E. Segmentation*

Image segmentation is the process of dividing an image into different parts [12].This is used to identify relevant information in digital images. In order to analyse the plant leaf diseases, there is a need to fetch the features of these images, which include segmentation techniques. There are different ways of performing image segmentation from the simple thresholding method to advanced colour image segmentation methods. Digital image processing plays a key role in agriculture field for the detection and recognition of diseases in plants with high accuracy. It is possible to perform the segmentation process based on various features found in the image like area, perimeter, and major axis length.In segmentation step to find out the infected region, we are using k-mean clustering. The image is segmented into various parts according to the region of interest. K-Mean clustering is used to get the data center of the image. In this we make the clusters of that image and find the center distance from the other cluster[12].

*F. K-Means Clustering*

K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoicells.The problem is computationally difficult (NP-hard); however, efficient heuristic algorithms converge quickly to a local optimum. These are usually similar to the maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both k-means and Gaussian mixture modeling. They both use cluster centers to model the data; however, k-means clustering tends to find clusters of comparable spatial extent, while the expectation-maximization mechanism allows clusters to have different shapes. The algorithm has a loose relationship to the k-nearest neighbor classifier, a popular machine learning technique for classification that is often confused with *k*-means due to the name. Applying the 1-nearest neighbor classifier to the cluster centers obtained by k-means classifies new data into the existing clusters. This is known as nearest centroid classifier or Rocchio algorithm.

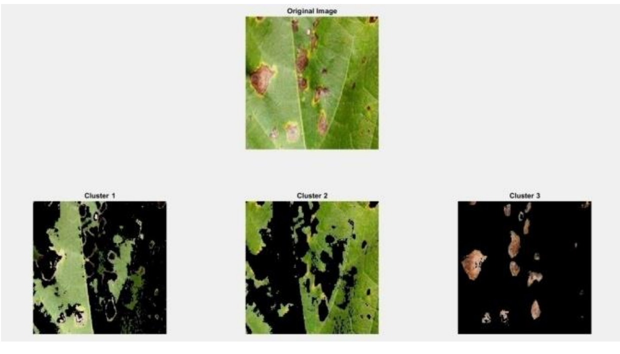



Fig.5. K-Mean clustering

*G. Contour tracing algorithm*

Contour tracing is a technique that is applied to digital images in order to extract their boundary. Contour tracing is one of many preprocessing techniques performed on digital images in order to extract information about their general shape. Once the contour of a given pattern is extracted, its different characteristics will be examined and used as features which will later on be used in pattern classification. Therefore, correct extraction of the contour will produce more accurate features which will increase the chances of correctly classifying a given pattern. The contour pixels are generally a small subset of the total number of pixels representing a pattern. Therefore, the amount of computation is greatly reduced when we run feature extracting algorithms on the contour instead of on the whole pattern. Since the contour shares a lot of features with the original pattern, the feature extraction process becomes much more efficient when performed on the contour rather on the original pattern. In conclusion, contour tracing is often a major contributor to the efficiency of the feature extraction process an essential process in the field.

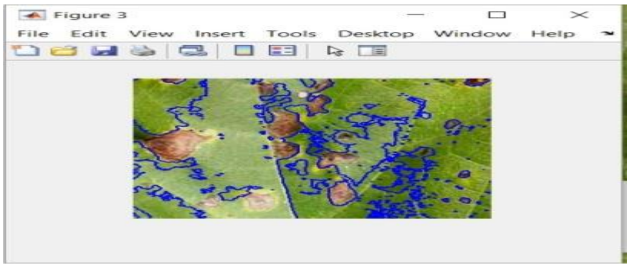



Fig.6. Contour tracing

*H. Feature extraction*

This process is done after the segmentation. According to the segmented information and predefined dataset some features of the image should be extracted. In this process we use the Grey Level Co-occurrence Matrices (GLCM).By using GLCM we measure the some parameters of the segmented image. For a example standard deviation, size, shape and intensity. After segmenting the region of interest using K-means clustering, the logical image will be converted into greyscale image in order to use the Grey Level Colour Co-occurrence Matrix (GLCM) for feature extraction. Thirteen different features will be calculated from the greyscale image. These features include Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, Smoothness, Variance, Kurtosis, Skewness, Inverse Difference Moment (IDM) and Root Mean Square (RMS) [11].

*I. The Discrete Wavelet Transform (DWT)*

The DWT represents the signal in dynamic sub-band decomposition. Generation of the DWT in a wavelet packet allows sub-band analysis without the constraint of dynamic decomposition. The discrete wavelet packet transform (DWPT) performs an adaptive decomposition of frequency axis. The specific decomposition will be selected according to an optimization criterion[11].

*Algorithm*

Step.1: capture the image by using digital camera.

Step.2: load the input images.

Step.3: resize of image and applying histogram for specific image, a viewer will able to judge the entire tonal distribution at a glance.
Step.4: segmentation is done by using K-Mean clustering method.
Step.5: contour tracing algorithm is used for tracing and extracting contour pixels, because these are simple and useful for detecting objects.
Step.6: features are extracted by using method as DWT, PCA and GLCM.
Step.7: CNN classifier is used for classification.

Step.8: finally disease will be detected and accuracy will be measure in terms of parameters.

## III. RESULTS AND DISCUSSIONS

Recognizing the disease is main purpose of the proposed system. The result shows the valuable approach which support accurate detection of the diseased leaf. Image processing technique is applied to detect the affected part of leaf from the input image. K-means algorithm is used for clustering of images. Disease detection is main motive of this system. In near future work can be extended for developing of hybrid algorithms using NNs to improve the recognition rate. Thus this technique would be useful for saving the farmers from a huge loss. The main characteristics of disease detection are speed and accuracy. Hence there is working on development of fast, automatic, efficient and accurate system, which is use for detection disease on unhealthy leaf.

```
training started...Wait for ~200 seconds...
training started...
Elapsed time is 2.033151 seconds.
Elapsed time is 2.239313 seconds.
...training finished.
testing started....
test error is
Elapsed time is 1.085832 seconds.
CNN Accuracy   =99.0909
CNN Precision    =0.9913
CNN Sensitivity    =0.99091
CNN Specificity    =0.99773
CNN Confutionmatrix    =

confmatrix =

    22     1     0     0     0
     0    21     0     0     0
     0     0    22     0     0
     0     0     0    22     0
     0     0     0     0    22
```

Fig.7. Overall Output

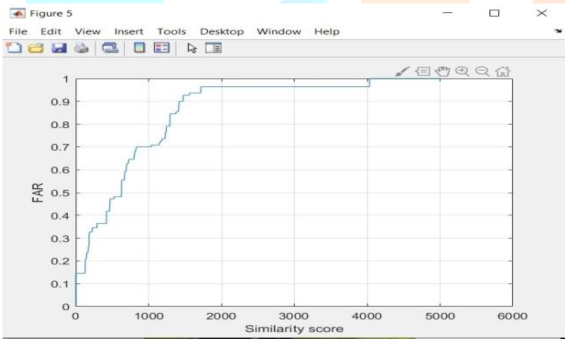



Fig.8. FAR Vs Similarity Score

FAR means False Acceptance Rate. Here we are using FAR Vs Similarity Score of that particular database. Similarity score shows that the percentage of the similar values or the how much percentage values are matching each other. In the similarity score if distance is more then it will give high percentage of non- similar value and if distance is less then it will shows the low percentage of non-similar values. Similarity score is depends on the distance or length of the feature database.

FAR gives the percentage of the false acceptance rate of the database. Far values are depends on the database and neural networks. In neural network if we use more number of iterations then FAR will be less for that particular database. Now here we are using FAR Vs Similarity Score graphs to show the results of database or proposed method. Here 'X' axis showing the FAR and 'Y' axis Showing Similarity Score. Above graphs we can see that at 4000 we are getting FAR = 1. This shows that at that point FAR is high means is allows more false no. data. If we decrease the distance value of the Similarity Score then it allows less no. of false data. Finally at 0 of similarity score we are getting 0 percentage of FAR Value.

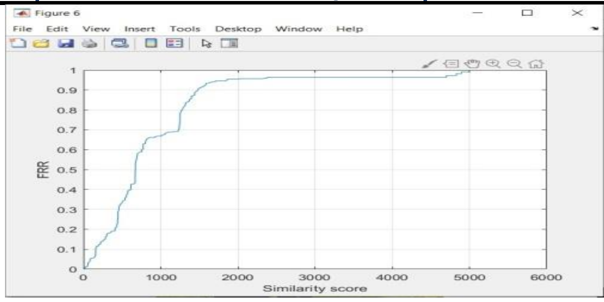



Fig.9. FRR Vs Similarity Score.

FRR means False Rejection Rate. Here we are using FRR Vs Similarity Score of that particular database. Similarity score shows that the percentage of the similar values or the how much percentage values are matching each other. In the similarity score if distance is more then it will give high percentage of non- similar value and if distance is less then it will shows the low percentage of non-similar values. Similarity score is depends on the distance or length of the feature database.

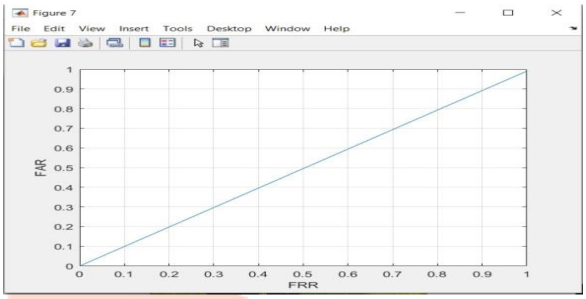



Fig.10. FARVs FRR

FRR gives the percentage of the false Rejection rate of the database. FRR value are depends on the database and neural networks. In neural network if we use more number of iterations then FAR will be less for that particular database. Now here we are using FRR Vs Similarity Score graphs to show the results of database or proposed method. Here 'X' axis showing the FAR and 'Y' axis Showing Similarity Score. Above graphs we can see that at 5000 we are getting FAR = 1. This shows that at that point FRR is high means it reject more true no. data. If we decrease the distance value of the Similarity Score then it Reject less no. of True data. Finally at 0 of similarity score we are getting '0' percentage of FRR Value.

Now in this graphs we are using FAR values and FRR values. In above graphs we can see that both FRR and FAR values are almost same at all point. Because of that in this graph we are getting linear line as a graph of both FRR and FAR values. It means in this proposed method FAR and FRR values are same (50% of FRR value and 50% of FAR value.)

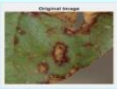
| | | | |
|---|---|---|---|
|  | Date- 02/03/2021<br>Temperature- 28 °C<br>Humidity- 82<br>Soil Moisture- -45% | Bacterial blight disease | Three sprays of fungicides at the interval of 12-15 days. |
|  | Date- 15/02/2021<br>Temperature- 32 °C<br>Humidity- 53<br>Soil Moisture- -20% | Fungus | Spray the copper content fungicides. |
|  | Date- 27/1/2021<br>Temperature- 19 °C<br>Humidity- 96<br>Soil Moisture- 10% | Cercospora leaf spot | Fertilize to keep your plant healthy. |

Fig. 11. Readings taken at different instants

| | | |
|---|---|---|
|  | Soil moisture -45 % Temperature 30°C Humidity 84 | DATE- 24/12/2020 |
|  | Soil moisture 50 % Temperature 24°C Humidity 76 | DATE- 6/1/2021 |
|  | Soil moisture 10 % Temperature 29°C Humidity 45 | DATE- 19/1/2021 |

Fig. 12. Hardware results at different instants

## IV. CONCLUSION

This paper gives the comparison on different techniques of plant diseases that can be used for plant leaf disease detection and an algorithm for image segmentation technique used for automatic detection as well as classification of plant leaf diseases has been described. Hence, related diseases for these sample plants have been taken for identification. Using very less computational techniques the optimum results have been obtained which also shows the efficiency of proposed algorithm in recognition and classification of the leaf diseases. Another advantage of using this method is that the plant leaf diseases can be identified at early stage to improve recognition rate in classification process.

## V. FUTURE SCOPE

In order to obtain better accuracy value, more leaf image data need to be used. I this paper output is displayed on system, in future it can be displayed on mobile.Solution of that particular disease will be sent to mobile and E-mail.

## REFERENCES

[1].RajneetKaur, Miss. ManjeetKaur, “A Brief Review on Plant Disease Detection using in Image Processing.”, International Journal of Computer Science and Mobile Computing, Vol.6 Issue.2, pp. 101-106, February- 2017.
[2].Sujatha R, Y Sravan Kumar and Garine Uma Akhil, “Leaf disease detection using image processing.”, Journal of Chemical and Pharmaceutical Sciences, Vol.10 Issue.1,January - March 2017.
[3]. Surender Kumar, RupinderKaur, “Plant Disease Detection using Image Processing- A Review.”, International Journal of Computer Applications (0975 – 8887) Vol.124 – no.16, August 2015.
[4].KhushalKhairnar, Rahul Dagade, “Disease Detection and Diagnosis on Plant using Image Processing – A Review.”, International Journal of Computer Applications (0975 – 8887) Vol.108 – no.13, December 2016.
[5]. Vaijinath B. Batule, Gaurav U. Chavan, Vishal P. Sanap and Kiran D. Wadkar, “Leaf Disease Detection using Image Processing and Support Vector Machine(SVM).”, Journal for Research, Vol.02, no.02,April 2016.
[6]. Sushil R. Kamlapurkar, “Detection of Plant Leaf Disease Using Image Processing Approach”, International Journal of Scientific and Research Publications, Vol. 06, no. 02, 2016.
[7]. RakeshChaware, RohitKarpe, PrithviPakhale and Prof. Smita Desai “Detection and Recognition of Leaf Disease Using Image Processing”, Research Article, Vol. 07, no. 05, 2014.
[8]. Arti N. Rathod, Bhavesh A. Tanawala and Vatsal H. Shah “Leaf Disease Detection Using Image Processing and Neural Network ”,International Journal of Advance Engineering and Research Development(IJAERD), Vol. 01, no. 06.
[9]. SujeetVarshney and TarunDalal “Plant Disease Prediction using Image Processing Techniques –A Review”, International Journal of Computer Science and Mobile Computing, Vol. 05, no. 05, pp.394-398, 2016.
[10]. Dr. S. Baskaran, P.Sampath, P. Sarathkumar, S. Sivashankar and K. VasanthKumar “Advances in Image Processing for Detection of Plant Disease”, The SIJ Transactions on Computer Science Engineering and its Applications (CSEA),Vol.03, no. 03, 2017.
[11]. Sa'ed Abed, Anwar Ali Esmaeel. "A novel approach to classify and detect bean diseases based on image processing”, 2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE), 2018.
[12]. M S Arya, K Anjali, Divya Unni. "Detection of unhealthy plant leaves using image processing and genetic algorithm with Arduino" , 2018 International Conference on Power, Signals, Control and Computation (EPSCICON), 2018