ARACHIS HYPOGAEA LEAF DISEASE DETECTION MODEL USING IMAGE PROCESSING

[1] Chandan V, [2]Sheik Imran, [3]Aishwarya B. J., [4]Chaitali Shetty, [5]Sumanth R.S.

[1,2]Assistant Professor, [3, 4, 5] Students
Department of Information Science and Engineering,
Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India.

Abstract : Arachis Hypogae, which is also known as Peanut, is widely cultivated crop in tropic and sub-tropic regions, this crop is being important to small and large commercial and non-commercial farmers in India. Diseases in plants cause major production and economic losses, and also degrade the quality of agricultural products. So the identification and quantification of crop diseases are required for adequate crop production and determination of crop losses. The detection of infected leaves have been done by farmers using naked eyes observation, which is time consuming, not accurate and contribute to many errors. Our proposed work aims to provide an user friendly interface that helps farmers to continuously monitor the peanut crop and meantime identify the different diseases that infect the crop which in turn increase the crop yield . Detection of the disease is based on symptoms that can be observed on the parts of the plants such as leaf, stems, lesions and fruits.

***Index Terms* - A**rachis Hypogae, GLCM, SVM, K-means clustering, Fabaceae.

I. INTRODUCTION

India is the land of farmers and Agriculture is the main source of their income. Almost 70% people depend on this & shares major part for the Gross Domestic Product. Nearly two thirds of all groundnuts are crushed for oil [1] .India has the distinction of being the largest producer of groundnut in the world. But average yield is very low i.e., 7.5 qtr/ha. Groundnut crop is prone attack by numerous diseases to a much larger extent than any other crops. One of the most important factors contributing to low yield is disease attack. Some diseases are widely distributed and cause economic crop losses while others are restricted in distribution and are not considered to be economically important at present .The occurrence of the disease on the plant leaf may result in significant loss in both quality as well as the quantity of agricultural product [2].

Peanut which is also known as Groundnut is one of the top oil-seed crops grown in Asian countries. Peanut is widely grown in the tropics and subtropics, being important to both small and large commercial farmers. India is the second largest producer of Peanuts in Asia. Peanut belong to the family of “Fabaceae” and genus of “Arachis”. Peanuts are rich in essential nutrients. Usually the methods that are adopted for monitoring and management of plant leaf disease are manual. One such major approach is naked eye observation. But it needs continuous monitoring. Another approach is seeking advice from the expert which may add to the cost. Diagnosis of the disease on plant leaf can also be done in laboratory testing. But this method requires satisfactory conditions along with professional knowledge.

Different types of diseases are caused by fungi, bacteria and viruses on agricultural plants. Each leaf has its own features and carries significant information that can help people to recognize and classify the leaf disease by looking at it. So proper system has to be developed which should identify and classify the disease accurately from the leaf images and it should be user friendly and should give faster results in an accurate manner. Image processing techniques such as image acquisition i.e., the peanut leaf images are captured with the help of camera and data set of those images are created. Then image is pre-processed to make it as a standard set. Pre-processing involves image enhancement and segmentation. In feature extraction the diseased portion of the input image is extracted and is treated as Region of Interest. Classification mainly involves two phases namely training phase and testing phase. In training phase characteristic features of images are isolated. Based on these features a unique description of each classification category i.e. training class is created. In testing phase feature space partition are used to classify the image features. Support Vector Machine (SVM) algorithm is used as the classifier to detect and classify different diseases that are infected to the peanut leaf.

II. LITERATURE SURVREY

Various methods for identification and classification of plant leaf diseases like pattern recognition method, back propagation, neural network, support vector machine, back propagation neural network etc. and research gap for the existing works were discussed [2]. A digital image-based algorithm based on histogram for automatic plant disease identification that deals with several diseases and easily retrained when new diseases are included. It is robust to the condition under which the images were captured [3]. A system to detect the crop diseases using image processing methods for the following crops such as rice (drechslera oryzae), cotton (alternaria gossypina), groundnut (bud necrosis virus), sugarcane (collectotrichum falcatum), and wheat (puccinia recondite) diseases will be detected. The system is based on feature extraction of an image and various comparison techniques in image processing [5].

The major diseases that affect the peanut plant are early leaf spot, late leaf spot, rust, web blotch etc. In recent years many researchers proposed many techniques, some of these are studied as discussed below:

Major Diseases affected to Peanut Leaf	Description
 Early leaf spot	Lesions are initially small, light brown in color. The lesions maybe surrounded by a yellow halo. Under favorable conditions, lesions enlarge up to one-half inch in diameter. Spots may expand or grow together.
 Late leaf spot	Lesions are dark brown to black. Dark brown to black spores are found on the lower leaf surface. Under extreme pressure, lesions can develop on stems.
 Rust	Rust is readily identified by the presence of numerous orange-to-red colored lesions (pustules) on the underside of the leaf.
 Web blotch	It has net like or webbing patterns, which are purplish, brown to tan on upper leaf surface. Circular, tan to brown lesions may grow together to cover the entire leaf.

Fig 2.1: Major Peanut Leaf diseases

The use of ANN methods for classification of disease in plants such as self-organizing feature map, back propagation algorithm, SVMs etc. can be efficiently used to accurately identify and classify various plant diseases using image processing techniques [6]. Feature Extraction is arguably the most important phase of all, and it is very important to choose the most appropriate set of feature for the images at hand. This proper selection sets the tone for better results by the classifier. The classifier based on neural network works better than others in most of the cases [7].

A new method was required for measuring the areas occupied by early and late leaf spots, which manifest in the leaves of peanut plants. The method is based on well-known mathematical morphology operations and on heuristics resulting from specialist knowledge. The characteristics of the leaves and lesions tackled in this work are quite similar to other combinations of plants and diseases [11]. Proposed method observes and identifies the diseases that attacked the plants, with the help of sensors that use image processing techniques to broadcast the captured image to the cloud. In turn, the image can then be viewed by us wherever and whenever we reside, in any part of the globe [8]. The segmentation refers to the process of classifying an image into plant and non-plant pixels. Good performance in this process is crucial for further analysis of the plant such as plant classification and effective action based on this analysis [10].

The Naked Eye Observation of experts is the main approach adopted in practice for detection and identification of plant diseases. This requires continuous monitoring of experts which might be expensive in large farms. Consulting experts is time consuming and moreover farmers are unaware of non-native diseases. Automatic detection of Plant diseases from the symptoms that appear on the plant leaves is an important research topic which may prove beneficial in monitoring large field of crops. Visual identification is labor intensive and less accurate. If disease symptoms are not obvious or more complex, it will be very difficult to distinguish the characteristics. Hence it delays the effective control of crop diseases [4] [5] [11].

III. METHODOLOGY

Initially the images of various leaves are captured using a digital camera or web cam and dataset of those images are created. Images are pre-processed to make it as a standard set which include image enhancement and segmentation. In Feature extraction the diseased portion from the input image is extracted and treated as a Region of Interest. Finally the disease infected to the peanut leaf is identified and classified using Support Vector Machine algorithm.

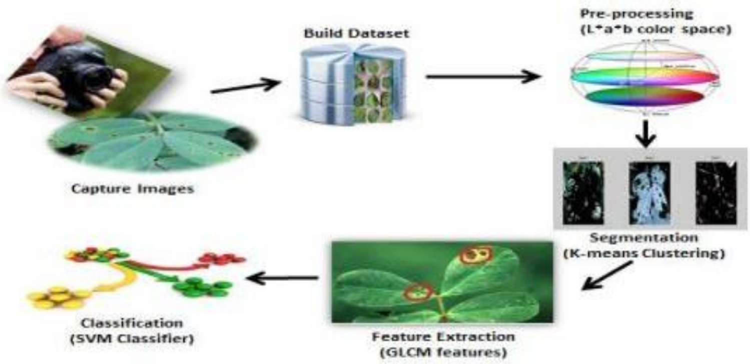


Fig 3.1: System Architecture

The proposed system architecture will give the farmers a handy approach to detect various diseases. It certainly brings an increase in yield or productivity. It also raises the income of farmers which replaces the traditional manual method.

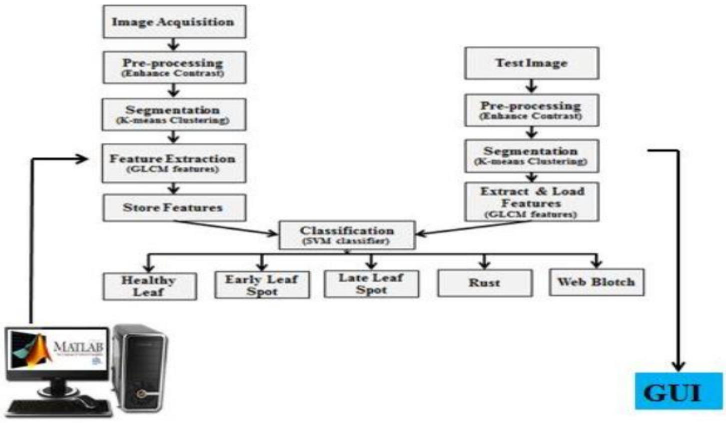


Fig 3.2: Methodology

The diagram gives the complete overview of the proposed system. The images from the dataset are given as input to the application. In training phase characteristic features of images are isolated. Based on the features a unique description of classification category i.e., a training class is created and is stored. In testing phase feature space partition are used to classify the different diseases such as early leaf Spot, Late leaf Spot, Rust and Web blotch that are affected to the Peanut leaf.

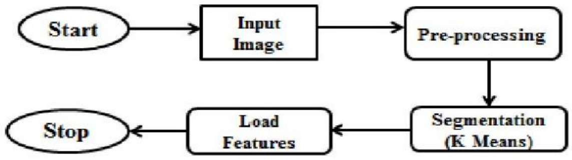


Fig 3.3: Training Phase

A dataflow diagram is a graphical representation of the flow of data through an information system, modeling its process aspects. It is often used as a preliminary step to create an overview of the system without going into great detail, which can later be elaborated. The above figure shows the data Flow diagram for the Training phase. The images are uploaded to the system and an image dataset is created. The uploaded images are Pre-Processed to enhance the contrast. Then the images are segmented using K-Means Clustering. Some of the features are extracted using Gray Level Co-occurrence Matrix (GLCM) and are stored.

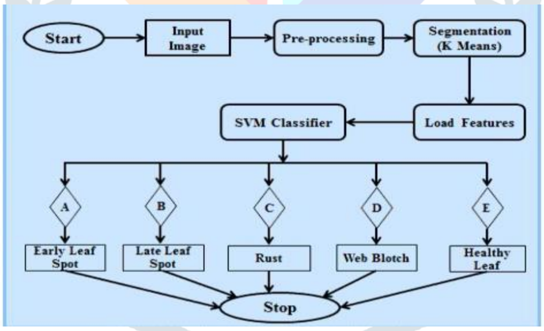


Fig 3.4: Testing Phase

The above figure shows the Data Flow diagram for testing phase. The user uploads the peanut leaf image as input. The uploaded image is Pre-Processed to enhance the contrast. Image is segmented using K Means Clustering. Features of the images are extracted using GLCM. Support Vector Machine (SVM) classifier is used to classify the disease affected to the peanut leaf.

IV. IMPLEMENTATION

Color based Segmentation using K-Means Clustering K-Means Clustering method is used for segmenting colors in an automated fashion using the L*a*b color space. The entire process can be summarized in the following steps below.

Step 1: Read the image.

Step 2: Convert image from RGB Color Space to L*a*b Color Space.

Step 3: Classify the colors in a*b* color space using K-Means Clustering.

Step 4: Label every Pixel in the image using results from K - Means.

Step 5: Create a blank cell array to store the results of clustering. Using pixel labels, we create the RGB labels. K-means clustering is a method of vector quantization 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.

Given an initial set of k-means m_1…...m_k, the algorithm precedes by alternating between two steps:

- **Assignment step:** Assign each observation to the cluster whose mean yields the latest within-cluster sum of squares (WCSS). Since the sum of squares is the squared Euclidean distance, this is intuitively the "nearest" mean.

$$S^{(t)} = \{ x_p : \| x_p - m_i^{(t)} \|^2 \leq \| x_p - m_j^{(t)} \|^2 \forall j, 1 \leq j \leq k \}$$

Where each x_p is assigned to exactly one S (t), even if it could be assigned to two or more of them.

- **Update step:** Calculate the new means to be the centroid of the observations in the new clusters. Since the arithmetic mean is a least-squares estimator, this also minimizes the within cluster sum of squares objective.

$$M_i^{(t+1)} = 1/|\, S_i^{(t)} \,|\ \sum_{x_j \in S_i^{(t)}} x_j$$

Pseudocode

Step 1: Read image.

```
I = imread([pathname,filename]);
I = imresize(I,[256,256]);
imshow(I2);title('Query Image');
```

Step 2: Convert image from RGB color space to L*a*b* color space using makecform and applycform

```
cform = makecform('srgb2lab');
lab_he = applycform(I,cform);
```

Step 3: Classify the colors in 'a*b*' space using K-means clustering. K-means clustering requires that you specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. Use k-means to cluster the objects into three clusters using the Euclidean distance metric.

Step 4: Label every pixel in the image using the results from k-means. For every object in your input, k-means returns an index corresponding to a cluster. Label every pixel in the image with its cluster index.

```
pixel_labels = reshape (cluster_idx,nrows,ncols);
```

Step 5: Create a blank cell array to store the results of clustering.

```
segmented_images = cell (1,3);
```

Step 6: Extract the features from the segmented images.

Different image features are extracted using Gray Level Co-occurrence Matrix (GLCM). Different statistics as given below are derived from GLCM.

```
glcms = graycomatrix(img);
stats = graycoprops(glcms,'Contrast Correlation Energy Homogeneity');
Contrast = stats.Contrast;
Correlation = stats.Correlation;
Energy = stats.Energy;
Homogeneity = stats.Homogeneity;
Mean = mean2(seg_img);
Standard_Deviation = std2(seg_img);
Entropy = entropy(seg_img);
RMS = mean2(rms(seg_img));
Variance = mean2(var(double(seg_img)));
 a = sum(double(seg_img(:)));
Smoothness = 1-(1/(1+a));
Kurtosis = kurtosis(double(seg_img(:)))
Skewness = skewness(double(seg_img(:)));
```

Step 7: Evaluate the disease affected area.

```
Affected_Area = 1-(A1/A2).
```

Where, A1=Area of the disease affected region. A2=Total leaf area.

Step 8: Load All The Features. Load all the features which are extracted in the earlier step.

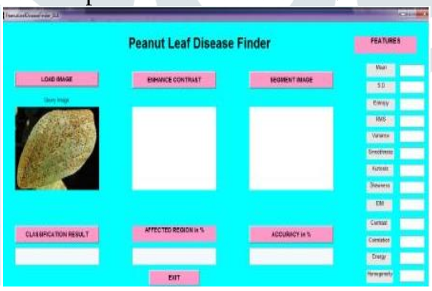
```
load('Training_Data.mat');
```

Step 9: Peanut leaf diseases are detected and classified using Support Vector Machine (SVM) classifier.

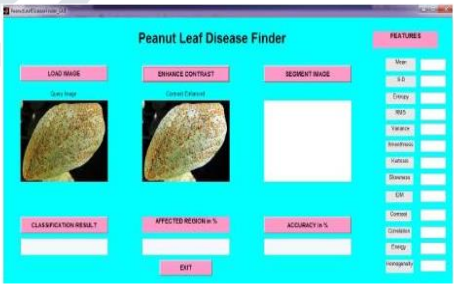
V. Results

The Snapshot 5.1 illustrates the Graphical User interface design of Peanut Disease Finder System. The Snapshot 5.2 and Snapshot 5.3 shows the image acquisition and enhancement process for rust disease.

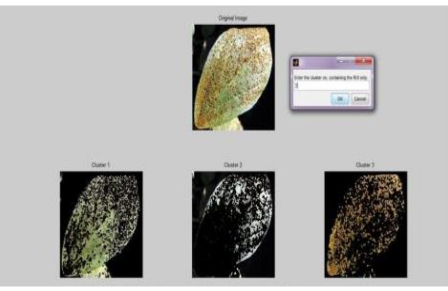
Snapshot 5.1: Shows GUI of proposed system



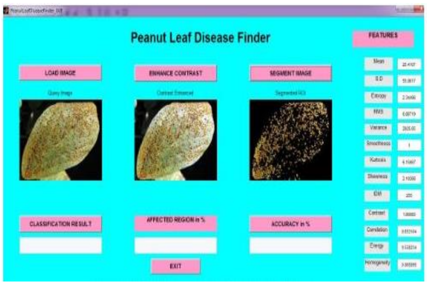
Snapshot 5.2: Illustrates Image Acquisition



Snapshot 5.3: Illustrates Image Enhancement



Snapshot 5.4: Illustrates Image Segmentation



Snapshot 5.5: Illustrates Feature Extraction

Snapshot 5.6: Classification and Fungicide suggestion

Feature Extraction is done using Gray Level Co-occurrence Matrix (GLCM). Features such as Mean, Standard Deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, Skewness, IDM, Contrast, Correlation, Entropy and Homogeneity are extracted. Classification is done using Support Vector Machine (SVM) algorithm and also suggests the appropriate fungicide to be applied for Rust disease.

Snapshot 5.7, 5.6, 5.9: Illustrates the Accuracies for Rust Disease, Early Leaf Spot Disease and Late Leaf Spot Disease

Snapshot 5.10: Illustrates the Accuracy for an Healthy Leaf

VI. CONCLUSION

The accurate detection and classification of the plant leaf diseases is very important for the successful cultivation of crop and this can be done using image processing. A system for diagnosis the leaf disease has been developed. The image processing techniques are applied to improve and enhance the image to a better quality. The methodology involves image acquisition, pre-processing, segmentation, analysis and classification of the peanut leaf. In the present work only four crop diseases were found. The future work concerns with large database of crop, fruit, and flowers by using advance feature of color extraction to find more diseases. The work carried out in this system includes the selection of clusters which is best suited from three different clusters of the loaded peanut leaf image during the application of k-means clustering technique. This selection can be done automatically by applying improved techniques for segmentation. Different classifiers can be used to get better accuracy rate.

REFERENCES

[1] Ishu Kumar Khute*, S.S. Rao, " Screening of Groundnut Genotypes for Early and Late Leaf Spot Disease" International Journal of Current Microbiology and Applied Sciences, volume 7 issue 2, August 2018.

[2] Prabhjeet kaur., Dr.Sanjay Singla., "A Review on the Plant Leaf Disease Detection Techniques", International Journal of Innovations and Technology(IJIET),volume 7 issue 2, August 2016.

[3] Jayme Garcia Arnal Barbedo, Luciano Vieira Koenigkan, Thiago Teixeira Santos, _"Identifying multiple plant diseases using digital Image Processing" ,Science_Direct, 2016.

[4] Shreya B. Bhavsar., Prof. Dipti B. Shah.," Plant Disease Identification Techniques using Image Processing: A Study", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 8, August 2015.

[5] Daisy., Deepinder kaur., _"Detection of Crop Diseases On Basis of Various Image Processing Techniques", International Journal Of Engineering Sciences & Research Technology, July 2015.

[6] Sachin D. Khirade, A. B. Patil, "Plant Disease Detection Using Image Processing" International Conference on Computing Communication Control and Automation, 2015.

[7] Diptesh Majumdar ., Dipak Kumar Kole ., Aruna Chakraborty ., Dwijesh Dutta Majumder., "Detection & Diagnosis Of Plant Leaf Disease Using Integrated Image Processing Approach", International Journal of Computer Engineering and Applications,Volume VI, Issue-III June 2014.

[8] K. R. Gokulakrishnan., Kapilya., _"Detecting the Plant Diseases and Issues by Image Processing Technique and Broadcasting", International Journal of Science and Research (IJSR), Volume 3 Issue 5, May 2014.

[9] Kevizhalhou Kuotsu, Anup Das, R. Lal, G.C. Munda, P.K. Ghosh, S.V. Ngachan,"Land forming and tillage effects on soil properties and productivity of rain fed groundnut cropping system in northeastern India", Science Direct, 2014.

[10] Esmael Hamuda , Martin Glavin, Edward Jones, "A survey of image processing techniques for plant extraction and segmentation in the field", Science Direct, 2014.

[11] Jayme Garcia Arnal Barbedo.,"Automatically Measuring Early and Late Leaf Spot Lesions in Peanut Plants Using Digital Image Processing", Conference Paper(October 2013).