Paddy Leaf Disease Detection and Quantification using Computational Techniques

L. Devichandana[1], Dr. Kunjam Nageswara Rao[2], Dr. G. Sita Ratnam[3]

Departments of Computer Science and Systems Engineering, Andhra University College of Engineering (A), Andhra University, Visakhapatnam, India.

[1](E-mail: chandana.lankalapalli@gmail.com), [2](E-mail:kunjamnag@gmail.com)

[3](E-mail:sitagokuruboyina@gmail.com)

***Abstract*:** *Plants are mainly multi-cellular, predominantly photosynthetic eukaryotes of the kingdom Plantae. Plants are significant resources to avoid global warming. Phyto-Pathology is the scientific study of plant diseases caused by pathogens. Plant diseases can cause significant damage to the agricultural products both qualitatively and quantitatively. Detection of plant leaf disease is an essential research topic as it may produce benefits in examining large field of crops and thus automatically detect the symptoms of pathogens as soon as they appear on plant leaves. Disease can damage paddy leaf at various stages of growth and all parts of the plants like at the leaf neck and at the node. In this paper, introduced modern technique to find out the disease related to paddy leaf for the fast and accurate result, since the naked eye observation takes some time to detect pathogens. In our proposed system, K-means segmentation and Gray level co-occurrence matrix (GLCM) to extract features to detect the plant leaf diagnosis are implemented.*

***Index terms*:** *Pathogens, K-means clustering, GLCM, classification, diagnosis.*

I. INTRODUCTION

Agriculture plays an important role for the Indian economy. Nearly 70% of Indian Economy depends on cultivation. Paddy is one of the essential food crops in all over world.Considering the growing population in the demand for items like rice is increasing day by day. So to detect the diseases related to paddy is a main theme of this project. To address this problem to develop an automated system to the farmers to identify the paddy leaf disease. The main objective is to offer detection and finding the intensity of the paddy leaf disease by computational techniques such as image processing. The traditional approach for detecting and recognition of plant leaf disease is based on naked eye observation, which is very slow method as well as gives less accuracy. In some countries, taking advice from experts to find out plant leaf disease is expensive and time consuming due to availability of expert. Uneven checkup of plants results in increasing of various pathogens on plant which requires more chemicals to cure it and also these chemicals are toxic to others. Automatic recognition of plant leaf diseases are required to detect the features of disease in early stages

when they appear on the growing leaf [4], [5], [6].

The image processing begins with capturing of digital high resolution images. Healthy and Unhealthy images are captured and stored in database for experiment. Then images are applied for pre-processing for image magnification. Captured leaf images are segmented using K-means clustering method to form clusters. Features are extracted,before applying K-means and classification. Finally diseases are recognized by the system. The techniques that are involved in the identification of paddy leaf diseases are acquisition of image, pre-processing of image, segmentation of image, feature extraction and classification [2].

Initially, to read the various format images like jpeg, jpg, gif, bmp can be used at early stage of the system in image acquisition. After capturing the image various pre-processing techniques are applied to enhancing and removing the unwanted noise of the image. In preprocessing, RGB images are converted into gray image using the color conversions. To increase the contrast, by using various contrast enhancement techniques like histogram equalization and contrast adjustment. In segmentation, partitions of an image at distinct regions that each pixel containing the similar attributes in the image by various techniques like K-means, Fuzzy c-mean (FCM), Principal Component Analysis (PCA) are processed. Classification is a method for identifying and detecting the plant leaf diseases. All the classification algorithms are based on assumption of the image that depicts one or more features. There are different types of classification approach like SVM, artificial neural networks (ANN), Fuzzy classification. The feature extraction uses different types of feature values like texture feature, structure feature and geometric feature [2].

II. RELATED WORK

Literature survey is the initial step in preparing new methodologies for the particular area of interest by moving research forward.

In this author aimed at the pathogens which causes the leaf pathogens and defined a variety of paddy leaf diseases such as Blast disease (BD), Brown spot disease (BPD), Narrow brown spot disease (NBSD), which stops the growth and protection of the paddy. Disease can infect paddy at various stages of growth and all parts of the plant as the leaf neck and node. A system for identifying the paddy leaf disease, it is based on image processing using K-means algorithm is applied to remove the noise and unnecessary spots of the image. The noise free images are filtered, the filtered image is converted in to enhanced image for segmentation. Extracting the features based on shape and color. By using ANN and Fuzzy classifier is used to detect the leaf disease [1].

For increasing growth and productivity of crop field, farmers need automatic monitoring of leaf pathogens of paddy instead of traditional approach i.e., manual monitoring of disease do not give satisfactory result as naked eye observations in traditional method detection of paddy leaf pathogens require more time and also need expert hence it is non effective [7],[8]. So in this paper introduced a modern technique to find out disease related to leaf .To cope up with traditional eye observing technique, used digital image processing technique for quick and accurate paddy leaf disease detection. The proposed methodology is depends on K-means clustering algorithm and multi SVM technique to detect both leaf and fruit disease [3].

III. METHODOLOGY

To detect the diseases related to paddy leaf is the main theme of the project. To address this problem we are developing an automated system to the farmers to identify the paddy leaf diseases. The main objective is to offer a technological solution to address problem in effective way. In proposed system paddy leaf disease detection is done by using different techniques like Image Acquisition,Pre-Processing of Image, Segmentation of Image and classification algorithm so as to improve the accuracy of detecting the paddy leaf disease in a better way.

The following methodology is proposed for this project, it involves six stages of work:

A. Image Acquisition

Image Acquisition is the first step in digital image processing and it is capturing the image through digital camera and stores it in digital media for extreme MATLAB operations. It is also an action of retrieving an image from hardware, so it can be forwarded through further process. Here, we collect number of paddy leaf images with diseases.

B. Image Pre-Processing

Pre-Processing refers to the transformations applied to our data. Data Pre-processing is a technique that is used to convert raw data in to a clean dataset. In other words, whenever the data is gathered from different sources it is collected in raw format which is not realistic for the analysis.

The main cause of image pre-processing is to enhance the image data i.e., contained unwanted distortions or to enhance some image features for further processing. Pre-processing method uses various techniques like changing image size, shape, filtering of noise, image conversion, and transformation of images, image enhancement and morphological operations.

C. Image Segmentation

Image segmentation is the method for conversion of digital image in to several segments and rendering of an image in to something for easier analysis. Image segmentation is used for locating the objects and bounding line of that image. In segmentation, we used K-means clustering method for separation of images in to clusters. After segmentation, the intense cluster is the disease part of the leaf. The k-means clustering algorithm is applied to classify the segmented images in to k number of classes based on features. The codification is done by minimizing sum of square of distances between data objects and the corresponding cluster. From the results of K-mean clustering, labeling, of each pixel in the image is done and also segmented images are generated which contain diseases. This experiment uses segmentation technique so input image is partitioned in to three clusters for better segmentation result.

Algorithm: The procedure for K-Means Clustering Algorithm:

Step1: Classify the images into p number of groups where p should be known.

Step2: Mark p points at randomly in cluster centroid.

Step3: Mapping objects to their closest cluster centroid.

Step4: Calculate the mean, centroid or perimeter of all images in each cluster.

Step 5: Repeat steps 2, 3 and 4 until the equal points are mapped to each cluster.

D. Feature Extraction

In feature extraction, desired feature vectors such as color, texture, morphology and structures are extracted. Feature extraction is a method for resources required to describe a large set of data accurately. Statistical texture features are obtained by Gray level co-occurrence matrix (GLCM) formula for texture classification and texture features are calculated from statistical distribution of observed intensity combination at the specified position relative to others. Numbers of gray levels are important in GLCM also statistics are categorized in to order of first, second and higher for number of intensity points in each combination. Different statistical texture features of GLCM are energy, entropy, covariance, information measure of correlation, contrast etc.

No	Features	Formula
1	Contrast	$\sum_i \sum_j \lvert i-j \rvert^2 p(i,j,d,\theta)$
2	Correlation	$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}$
3	Energy	$\sum_i \sum_j p(i,j,d,\theta)^2$
4	Homogeneity	$\sum_i \sum_j p(i,j,d,\theta)/(1+\lvert i-j \rvert)$

Table 1. Shows list of formulae

E. Classification

Classification technique is applied on statistical features it depends on it by dividing the classes. For designing database, image is obtained and passed through pre-processing, segmentation, features extraction then disease name is selected for given leaf and lastly data is stored in database.

F. Finding Intensity

Determining the intensity of the diseased paddy leaf.

The step by step proposed methodology consists of leaf image acquisition, pre-processing the image, segmentation of the images using k-means clustering method, features withdrawal using GLCM method and finally classifies the system. The flow chart of the proposed system as shown in figure 1.

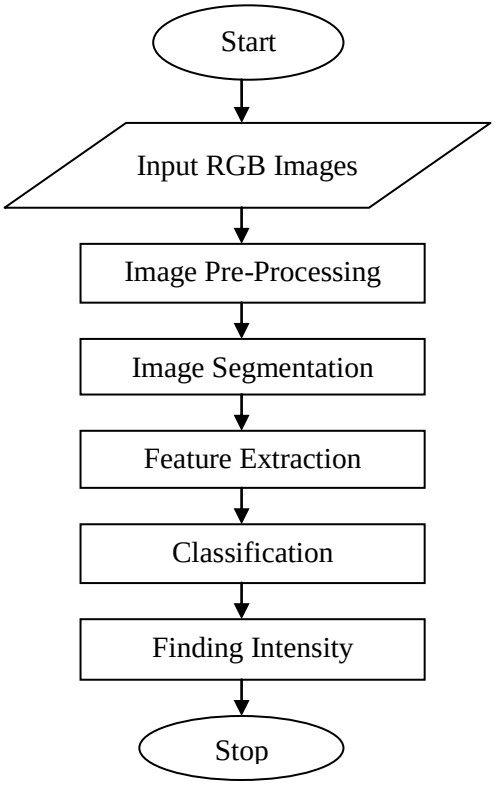


Figure 1. Dataflow diagram of the proposed system

Algorithm

Step 1: Start

Step 2: Initially taking RGB image i.e., paddy leaf image as input.

Step 3: Applying Pre-processing techniques to that leaf image.

Step 4: Applying color based segmentation using k-means clustering

Step 5: Identifying the affected regions by disease in paddy leaf.

Step 6: By extracting the features of the diseased part by using GLCM.

Step 7: Applying Classification techniques to enhance the paddy leaf disease.

Step 8: Finding the percentage of affected region.

Step 9: Stop

IV RESULTS

- Captured diseased images of paddy leaf for MATLAB image processing system.In Image Acquisition, by acquiring the desired paddy leaf which is diseased as shown in figure 2.

Figure 2. Original Image

- In Image pre-processing, to enhance the capturing image by changing image size and shape, filtering of noise, image conversion and contrast. Here, various MATLAB code to resize image, to enhance variance and RGB to gray scale conversion for image enhancement as shown in figure 3.

Figure 3 . Image Enhancement (Pre-Processed Image)

The following output screen shows the cropped Image where the diseased part located as shown in figure 4.

Figure 4. Cropped Image

After cropping that image, to remove noise.Noise is the result of errors in the image acquisition process that result in pixel values that do not reflect the true intensities of the real scene. There are several ways that noise can be introduced into an image, depending on how the image is created. Here to remove salt and pepper noise and to apply median filter as shown in figure 5.

Figure 5. Filtered Image

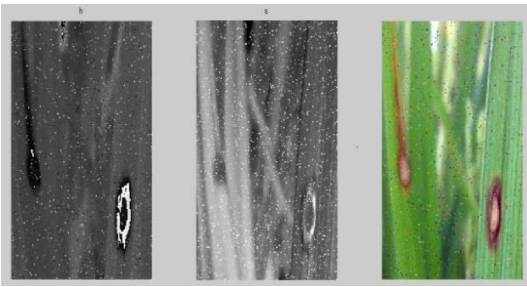


Figure 6. Converting RGB to HSV (Hue Saturation Value)

Here, HSV color space describes colors in terms of Hue, Saturation and value.HSV model describes color similarity to how the human eye tends to perceive color. The comparisons rely on color, vibrancy and brightness as shown in figure 6.

- The image labeled by cluster index as shown in figure 7.

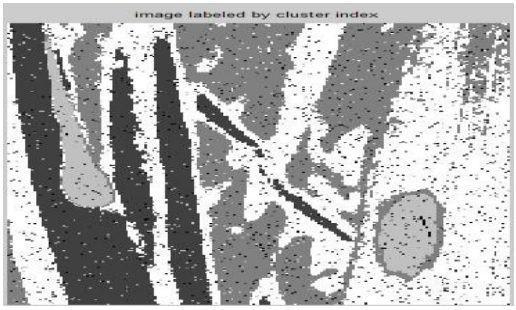


Figure 7. Clustered indexed image

- In segmentation, K-means clustering algorithm is used for separation of images into clusters. After segmentation, the intense cluster is the disease part of the leaf. This algorithm applied to classify the paddy leaf image which results in segmentation of images into clusters. The segmented leaf image is divided into 5 clusters, in which one or more cluster contains the disease as shown in figure 8.

Figure 8. Object in fourth Clustered Image

- After segmentation, the intense cluster is the disease part of the paddy.
- Finally, identified the defected region of the given leaf image as shown in figure 9.

Figure 9. Defected regions identification

- The statistical texture features of GLCM are energy, correlation, mean, variance, entropy, contrast, standard-deviation etc. among those diseases, bacterial blight, false smut, leaf blast, leaf spot are found.

Feature Extraction	Healthy Image Range
Entropy	<5
Mean	<5
Standard Deviation	<20

Table 2. Healthy Leaf Image Range

The table 2 shows the healthy leaf ranges. Those ranges are obtained based on entropy, Mean and standard deviation of healthy leaf images other than this range than that leaf is considered as diseased.

Diseased leaf feature values as shown in table 3:

A Leaf is taken here are values for that. That is a diseased leaf so the entropy value obtained is 6, mean obtained is 5 and the standard deviation value is 20**.**

Entropy	6
Mean	5
Standard deviation	20

Table 3. Diseased leaf values for Entropy, Mean and Standard Deviation.

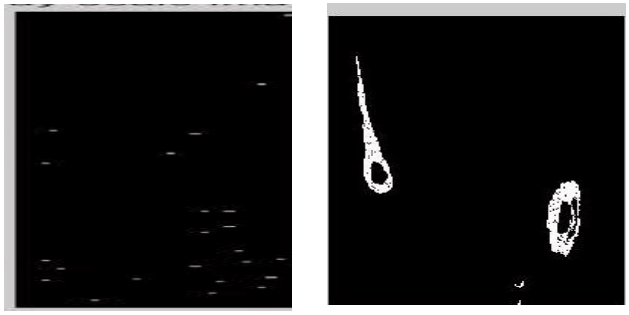
Healthy leaf feature values as shown in the table 4.

A Leaf is taken here are values for that. That is a healthy leaf so the entropy value obtained is 1, mean obtained is 1 and the standard deviation value is 10**.**

Entropy	1
Mean	1
Standard Deviation	10

Table 4. Healthy Leaf values for entropy, Mean and Standard Deviation.

Comparison of Healthy leaf and diseased leaf images as shown in figure 10.



a) **Healthy Leaf** **b) Diseased Leaf**

Figure 10. Healthy and Diseased leaf images

V. CONCLUSION

It provides efficient and transparent plant leaf disease detection and classification by using MATLAB image processing. The proposed methodology in this paper depends on k-means clustering and Classification techniques which are configured for paddy leaf disease detection. MATLAB software is ultimate for digital image processing. This system detects Leaf spot and stem rot disease. It provides fast and efficient time for entire processing.

Further enhancement of this work involves more experimentation's with large training sets to recognize various damaged leaves due to insects or diseases and to develop an expert system. The diseased images are to be collected in large number then sent to experts in agricultural research stations to estimate the effect of the diseases on the crops and also to propose the remedial measures for it. This type of methods can be used for disease identification and classification in planets like fruits, leaf, vegetables etc.

REFERNCES

[1] R.P.Narmadha&GArulvadivu,"Detection And Measurement of Paddy Leaf DiseaseSymptoms using Image Processing, International Conference on Computer Communication and Informatics ,2017".

[2] PoojaKulinavar "Classification of leaf disease based on Multiclass SVM classifier,International Journal of Innovative Research in Science, Engineering and Technology,Vol.3,Issue 4, pp.321-327, 2017".

[3] SandeshRaut "Plant Disease Detection in Image Processing Using MATLAB, International Journal of Innovative Research in Science, Engineering and Technology, Vol.6,Issue 6, pp.10373-10381, June 2017".

[4] T.N.Shaikh,Dr.S.M.Mukane"Segmentation of Pomegranate Leaf for Detection of Disease Using Image Processing, International Journal of Innovative Research in Computer and Communication Engineering, Vol.5, Special Issue 4,pp.231-235, June 2017".

[5] Manoj Mukherjee, Titan Paland DebabrataSamanta" Damaged Paddy Leaf Detection Using Image Processing,Journal of Global Research in Computer Science Vol.3, No. 10,pp.7-10, October 2012".

[6] Jitesh P. Shah ,Harshadkumar B. Prajapati and Vipul K. Dabhi "A Survey on Detection and Classification of Rice Plant Diseases, 2016 IEEE"

[7] Prithviraj D, Sumangala G S "Leaf Infection Detection and Diagnosis Using Image Processing Technique".

[8] Lipsa Barik "A Survey on Region Identification of Rice Diseases Using Image Processing, International Journal of Research and Scientific Innovation, Vol.V, Issue I,pp.21-23, January 2018".