A Survey on Cluster Analysis Techniques for Plant Disease Diagnosis

T.Nagarathinam[1], Dr. K. Rameshkumar[2]
PhD Research Scholar[1], PhD Research Guide[2]
[1](Department of Computer Science, MASS College of Arts & Science, Kumbakonam, Tamil Nadu, India)
[2](Department of Computer Science, MASS College of Arts & Science, Kumbakonam, Tamil Nadu, India)

ABSTRACT: *Plants are one of the major key resources to crack the trouble of global warming in the world. But plant diseases like Blast, canker, Bacterial Leaf Blight; Rice tungro, black spot, Scab, Powdery, Downey, mildew, Speckle, early scorch ashen mold, tiny whiteness, cotton mold, late scorch and etc., stops the growth of the plants. If the diseases are not detected in the early hours then there are decreases in the production of plantation. In this paper we surveyed Cluster Analysis technique for plants diseases detection and diagnosis.*

KEYWORDS: *ANN, Brown sport, fuzzy C-means, k-mean, Leaf Blast.*

I. INTRODUCTION

Diagnosis of diseases in plant is the fortitude of the causes of a disease in a plant, and detection refers to the identification of microorganisms. Failures in disease diagnosis show the way openly to too little disease control and cutback in crop production and quality, and therefore trade. Whilst plants are affectionate to diseases they put on view a range of visual indication such as change in color, shape and size as the disease growth. Identification of types of diseases that appears in the plant is difficulty for most of the people. So the various technique and method are developed for identification of diseases in plant. From the techniques; clustering technique has been reviewed to easy references of the future researchers. The survey begins with the introduction and the paper is organized as in section-II exhibit some import diseases that affect the plant growth, in section-III deals survey of the paper with rice, and general plants. Finally the paper concluded with summary and Future work.

II. SOME IMPORTANT RICE DISEASES

2.1 RICE BLAST

Blast is often considered as the most serious rice disease because it spreads rapidly and is high destructive under favorable conditions. The blast may
infect rice plants at any stage of growth. It is widespread disease of rice.

This is also named as rice fever disease. The disease was first recorded in India in 1913. In 1919, devastating epidemic happened in Tanjore delta area of Tamil Nadu. Various studies in India found that 1 per cent panicle blast reduced yield by as little as 0.4 per cent or 1.4 per cent to as much as 17.4 percent. Besides percentage, time and degree of infection greatly affected the yield loss.

Symptoms: the fungus produces spots or lesions on leaves nodes, neck of the panicles and the grains. The fungus attaches both the seeding and the transplanted crop. In the nursery, the effected seeding will show numerous spots on the leaves and begin to wither and die. In severe cases of infection large number of seeding are destroyed. Types of rice blast diseases are

2.1.1. Leaf blast: Typical leaf spots are elliptical and somewhat pointed at both ends that are the spots are spindle shaped. The margins are usually brown to reddish brown and the centre is often gray or whitish. Both the color and shape of the spots vary depending upon the
Environmental conditions, Age of the spots and Varietal susceptibility.

Susceptible varieties

The spots begin as small, water-soaked, whitish, grayish or bluish dots on susceptible varieties. The make bigger rapidly under moist conditions. Fully developed spots reach a length of 1 to 1.5cm and a width 0.3 to 0.5cm with brown margins.

Highly resistant varieties

In this, no lesions or only pinhead sized brown specks.

Intermediate varieties

In this varieties reaction have small, round or short-elliptical lesions, a few millimeters long with a small necrotic centre and a brown border. The brown border indicates that the lesion has stopped enlarging because of varietal resistant or unfavorable condition. When numerous spots occur on a leaf it son dies. The pots coalesce together and cause quick death of leaves. Plants attacked between the seedlings and maximum tillering stages are often completely killed.

2.1.2. Nodal blast: when the node is infected, it turns black, the node becomes weak and breaks apart, remaining connected to the straw by the modal septum only. All parts above the infected node die. The following figure shows nodal blast infected rice node.

2.1.4. Grain infection: the fungus infects the grain also. On the glume, small, circular brown lesions are formed.

Management:

The important control measures include use of cultural practices, fungicides and resistant varieties. The Annexure – I shows the spray for Control the disease.

2.2. BACTERIAL LEAF BLIGHT

Bacterial leaf blight was first discovered in Japan in 1884. It is also occur in various countries. In India, it was first reported from Maharashtra in 1959 it produce the 50% losses.

Symptoms

Symptoms vary with the stage of infection and weather conditions. They are leaf blight, Kresek and yellow leaves.

2.2.1. Leaf blight

The disease can attack at any stage of plant. In nursery, the seedlings show circular, yellow spots on the margins which enlarge in size, eventually the leaves turn yellow, dry and wither completely. Grouping up in the main field are also attack. Small water soaked streaks appear initially at the margins of the lamina near the tip. As the disease advances, the streak enlarges in size on both dimensions and become yellow to straw colored then to grey necrotic stripes. The blight phase usually appears 4 to 6 weeks after transplanting.

1.2.2. Kresek or wilt phase

It results from easily systematic infection or from infected seed or when the bacterium is brought in Contact with the germinating seeds. This is notices in transplanted seedlings of 1 to 2 weeks old. Infected leaves become grayish green, begin to fold and roll completely along the midrib. They droop, turn yellow or grey ultimately the tillers in the hill will die. This type of disease symptoms will not come out easily.

1.2.3. Yellow leaf phase

In tropics production of pale yellow leaves was observed. Youngest leaves in a hill may turn to yellow or white. Such leaves later turn yellowish brown and wither away. The following sample figure shows appearance of bacterial leaf blight.

III. REVIEW OF LITERATURES

Di Cui, Oin Zhang & et al. in [1] proposed a method for fast & accurate detection & classification of plant diseases. They clustering the affected plant leaves using k-means clustering and otsu segmentation methods and the classification of the same was done by back propagation feed forward neural network. Here CCD camera and portable spectrometer used to collect the soybean leave images. They used three parameters; such as rust severity index, ratio of infected area and lesion color index; were extracted from the multispectral images and used to detect leaf and severity of infection.

IDheeb Al Bashish & et al. in[2] proposed work having the stages as In the first step the k-means techniques is applied in order to segment the images and the segmented images are inputting to a pre-trained neural network for classification of the segmented images. Through which it, five kinds of leave diseases were selected; Cottony mold, tiny whiteness, Early scorch, Ashen mold and late scorch. The investigational outcome shows that the neural network classifier detects the leaf diseases automatically with a accuracy of about 93%.

H. Al-Hiary, S. Bani-Ahmad et al is described in [3] Fast and Accurate Detection and Classification of Plant Diseases. In this paper, the K-means algorithm for clustering the diseases so as to affect on plant leave images and the Neural Networks (NNs) invent for classification of the same. The proposed algorithm tested on early and late scorch, ashen mold, cotton mold and tiny whiteness. Finally the authors bring to a close that it can support an accurate detection of leaf infection in a tiny computational attempt.

Toran Verma, Susanta Kumar Satpathy [4] applies digital image processing and fuzzy clustering techniques to estimating loss in the rice leaf due to Leaf Blast Diseases. Once the rice leaf image is captured which is further analyzed by extract the RGB features of the images. Finally Fuzzy C Mean Clustering algorithm is applied to classify diseases in the rice leaf from the extracted features and, Decision support system were developed using the clustered information.

R. L. Pugoy and V. Y. Mariano[5], proposed an automated system that can detect two different kinds of diseases present in a rice leaf using color image analysis. In the system, the outlier region is first obtained from a rice leaf image which is tested utilize histogram intersection between the healthy and tested rice leaf images. The k-means technique is applied to cluster the pixels into a number of similar and dissimilar groups. Lastly, these groups of pixels are compared with the system library to determine the suspected diseases of the rice leaf.

Sekulska-Nalewajko and Goclawski [6] proposed a semi-automatic method to detect and quantify disease symptoms in two different plants leaves (pumpkin and cucumber). In this method, the leaves were removed from the plants and it was scanned by the flatbed scanner to test. Once the image scan is over, the image is transformed from RGB to HSV color space and the V is discarded. Pixels in the images grouped as diseased and healthy cluster using Fuzzy C-means algorithm to detect and quantify disease symptoms.

In [7] Jayamala K. Patil and Raj Kumar were tag a paper "Color Feature Extraction of Tomato Leaf Diseases". In this proposed method color of the image was used to extract the features of images by calculating first, second and third moment of color of the tomato leave samples. Then the clustering techniques were used for classifying the diseases that appears in the tomato leaves.

In [8] the authors title the paper as "A Hybrid Intelligent System for Automated Pomegranate Disease detection and Grading". The authors of the paper use various techniques to recognize the diseases on the Pomegranate. To do so, first the RGB images of leaves acquired then a color transformation structure is created for the RGB image. In the next step the RGB color values converted to the space specified in the color transformation structure. Then k-means technique would be applied to segment the images. In order to obtain HIS format, the green pixel in the segmented images were masked and pixels on the boundaries of the infected cluster and pixel with zeros red, green and blue values were completely removed then the infected cluster was converted from RGB format to HSI format. In this HIS format, the H and S values only used to generate SGDM matrix. Features are calculated by call the GLCM functions. After Texture Statistics Computation finally the recognition process was performed through a pre-trained Neural Networks.

In [9], Tushar H Jaware & et al. widen a Fast and accurate method for detection and classification of plant diseases. They have proposed algorithm which is tested on five diseases on the plants; the diseases were tiny whiteness, Early scorch, Ashen mold, Cottony mold, late scorch. Firstly, image is acquired using the camera and it is applied to color transformation structure. Secondly, k-means algorithm is utilized for segment the images and the green pixels are masked due to it does not contain diseased portion. Next the pixels on the boundaries of the infected object were completely removed. Then the infected cluster was obtained as HIS format from RGB format and generates the SGDM matrices. Finally the extracted feature was inputting to a pre-trained neural network for classifying the diseases. The authors conclude that the proposed system can successfully detect and classify the diseases with a precision between 83% and 94%.

In paper [10] N.Valliammal and Dr.S.N.Geethalakshmi designated a paper as Leaf Image Segmentation Based on the Combination of Wavelet Transform and K means clustering . They use Discrete Wavelet Transformation and K-means algorithm to segmenting the leaf images is called combined segmentation algorithm techniques. In this, features of the leaf images were extracted by Haar transformation tool. Whilst judge against to the on hand method, this combined technique determine better convergence result.

Dheeb Al Bashish &et al. in [11] authors uses k-means techniques to segment the plant image which is collected from the different sources. Once the segmentation is over, the segmented RGB images are converted into HIS color space. Then the features extracted from the images based on texture and color features. Finally, the features fed into the Minimum Distance Criterion and Support Vector Machine classifier for classifying the types of disease that it's appear in the plant.

Wang H, Li G, Ma Z, Li X. in [12] tagged a paper as "Application of neural networks to image recognition of plant diseases" proposed a method detect the diseases in wheat and grapevines. The images were segmented by k-means clustering algorithm and then fifty shape, texture and color features were extracted and inputted to four kinds of neural networks, such as MP, RBF, Generalized Regression, and Probabilistic, to classify the diseases.

In [13] Jagadeesh Devdas Pujari1, Rajesh et al, has proposed scheme consists of two phases. In the first phase is called segmentation techniques such as k-means clustering, thresholding, region growing, and watershed to found affected lesion areas from unaffected area. The affected areas were calculated graded by calculating the percentage of affected area. In the next phase, Run length Matrix was used to extract the texture features of the image. Finally the extracted features were fed into artificial neural network as a classifier for classify the diseases. The authors proposed method used to find the diseases in three types of fruit namely mango, grape and pomegranate the classification accuracies for normal fruit types(84.65%) and affected anthracnose fruit types(76.6%).

In [14] Jagadeesh D. Pujari et al, propose a paper to recognition and classification of visual symptoms affected by fungal disease. In this paper, color images of fungal diseases (leaf blight, leaf spot, powdery mildew, leaf rust, smut) affected in cereals used for the study. From the images affected regions were segmented by k-means cluster algorithm. Color Co-occurrence Matrix used for color texture analysis and Extracted features were inputted to the classifiers ANN and SVM to obtain the classification accuracy between 68.5% - 87% and 77.5 - 91.16 respectively.

Al. Tarawneh, Mokhled S [15] proposed a technique with combinations of auto-cropping segmentation and fuzzy c-means clustering algorithm for detect the olive oil diseases. In this proposal, crops the region of interest from diseased leaves for the segmentation using automatic polygon cropping, and classifies the diseases using fuzzy C-means algorithm. Finally, the result is compared with k-means using the parameter speed and accuracy and the author concludes that the fuzzy c-means produce better result than the k-means.

Vivek Chaudhari, C.Y. Patil [16] suggests "Disease Detection of Cotton Leaves Using Advanced Image Processing". In this proposal cotton diseases were first identified and classified. They make use of k-means for segment the cotton leaf images and apply the DWT for extract the features; PCA algorithm applied to reduce the extracted features. Lastly, the classification was done by NN techniques with 98% accuracy.

Niket Amoda et al.[17] proposed an automatic detection and classification of plant based diseases. In their software they use color transformation structure for the input leaf image. The major benefit of the software is no needs any preprocessing technique. So it significantly reduces time requirement for preprocessing. Here K-means clustering technique is used for segment the diseased portion from the image. Once the segmentation is over, the texture features were extracted from the segmented portion and it is input into neural network to classifying the leaves.

Rastogi, A et al.[18] are identify the leaf diseases using k-means clustering based segmentation algorithm. In this method k-number of clusters are generated using input image and the RGB space is converted into L*a*B space without consider of luminosity factor during processing it. Finally classification works done by using the classifier ANN.

In [19] Amit Kumar Singh, Rubiya .A, B.Senthil Raja proposed an approach to identify rice plant disease Leaf blast using the classifier Support Vector Machine classifier (SVM). In their approach, they used the k-means algorithm to segment the infected portion of the rice leaf image. Texture feature vectors were extracted from the segmented leaf image and it was given input to the SVM for classification. Finally authors conclude that the classifier classifies the disease with 82% accuracy.

In [20] Nikita Rishi1, Jagbir Singh Gill2 discussed various method and techniques; image cropping, de-noising, compression, Otsu method and k-means techniques; to detect the diseases in heterogeneous plant. They make use of Neural networks classifiers such as BPNN, RBF, GRNNs and PNNs to diagnose wheat and grape diseases. Sobel operator, canny filters and feature extraction applied to recognize the diseases on cotton and rice leaf.

Megha P Arakeri, Malavika Arun, Padmini R K[21] proposed a novel computer vision system for detect and analysis of late blight disease in Tomato to utilize k-means algorithm. The system classifies the leaves of tomato in diseased or healthy using thresholding algorithm. The given image is identified as diseased, the k-means algorithm is applied on the leaf image to attain three clusters such as background, healthy part of leaf and diseased part of the leaf. The developed method has been evaluated using the metrics

Sensitivity = (TP) / (TP+FN)

Specificity = (TN) / (TP+FN)

Accuracy = (TN +TP) / (TN+TP+FN +FP)

Where,

TP: True Positive, which is the proportion of positive cases that were correctly identified.

TN: True Negative, which is the proportion of negatives cases that were classified correctly.

FP: False Positive, is the proportion of negatives cases that were incorrectly classified as positive.

FN: False Negative, is the proportion of positives cases that were incorrectly classified as negative.

The developed method achieved sensitivity, specificity and an accuracy of 85%, 80% and 84% respectively.

In the paper [22] Bed Prakash1 Amit Yerpude2 proposed a system to identify the mango leaf diseases and control prediction of the same. Here the authors uses K-means algorithm to segment the diseases portion from the mango leaf image and find the optimal number of cluster to produce the best performance with the accuracy of about 94%. The classifier Back Propagation Neural Network (BPNN) technique used for the classification of the mango leaf diseases.

IV. SUMMARY OF THE SURVEY

The following graph shows the summary of the literature survey based on the year of publication with accuracy (figure-1) and classification tool used with accuracy(figure-2).

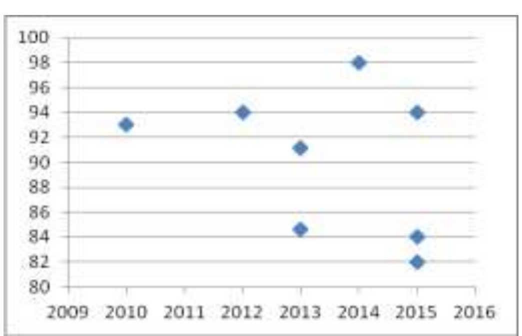

Figure – 1. Comparison of year of publication with Accuracy Names of the Classifier

Figure – 2. Comparison of classifier name with Accuracy

V. CONCLUSION AND FUTURE WORK

Since this review, we can conclude that there are number of ways by which we can detect diseases in plant. Each technique has some pros as well as limitation. This paper evaluates the techniques in data mining and image processing in use by researchers designed for detection, diagnosis and recognition of plant diseases. Still lot of research is going on for using various techniques to produce automated plant diseases detection via mobile phone captured images. Our upcoming research article will discover a cluster analysis algorithm for image classification for mobile devices with limited memory, higher processing speed and more accuracy.

REFERENCES

[1] Di Cui, Oin Zhang, Minzan Li, Youfu Zhao and Glen L. Hartman[2009] Detection of soybean rust using a multispectral image sensor, *SpringerLink –Journal article.*

[2] Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad ,"A Framework for Detection and Classification of Plant Leaf and Stem Diseases", *2010 International Conference on Signal and Image Processing*.

[3] H. Al-Hiary, S. Bani-Ahmad et al, Fast and Accurate Detection and Classification of Plant Diseases, *International Journal of Computer Applications (0975 – 8887) Volume 17– No.1, March 2011*

[4] Toran Verma, Susanta Kumar Satpathy-"A Step towards Precision Farming of Rice Crop by Estimating Loss Caused by Leaf Blast Disease Using Digital Image Processing and Fuzzy Clustering*", International Journal of Computer Trends and Technology- May to June Issue 2011.*

[5] R. L. Pugoy and V. Y. Mariano "Automated rice leaf disease detection using color image analysis*", Proc. SPIE 8009, Third International Conference on Digital Image Processing (ICDIP 2011), (July 08, 2011)*

[6] Sekulska-Nalewajko J, Goclawski J." A semi-automatic method for the discrimination of diseased regions in detached leaf images using fuzzy c-means clustering", *7th international conference on perspective technologies and methods in MEMS design. Polyana-Svalyava: IEEE; 2011. pp. 172–175*.

[7] Jayamala K. Patil, Raj Kumar- "Color Feature Extraction of Tomato Leaf Diseases", *International Journal of Engineering Trends and Technology- Volume2Issue2-2011. ISSN: 2231-5381 http://www.internationaljournalssrg.org Page 72.*

[8] Sannakki S.S., Rajpurohit V.S., Nargund V.B., Arun Kumar R.and Yallur P.S.," A Hybrid Intelligent System for Automated Pomegranate Disease detection and Grading", *International Journal of Machine Intelligence ISSN: 0975–2927. 2011*.

[9] Tushar H Jaware, Ravindra D Badgujar and Prashant G Patil," Crop disease detection using image segmentation", *Proceedings of Conference on Advances in Communication and Computing (NCACC'12)", April 21, 2012.*

[10] N.Valliammal and Dr.S.N.Geethalakshmi.,"Leaf Image Segmentation Based On the Combination of Wavelet Transform and K Means Clustering", *(IJARAI)*

International Journal of Advanced Research in Artificial Intelligence, Vol. 1, No. 3, 2012

[11] Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad, "A Framework for Detection and Classification of Plant leaf and Stem Diseases*", 2012 International conference on signal and Image Processing, pg.no.113-118, Chennai, India*

[12] Wang H, Li G, Ma Z, Li X. "Application of neural networks to image recognition of plant diseases*", Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI) Yantai: IEEE; 2012.; pp. 2159–2164.*

[13] Jagadeesh Devdas Pujari1, Rajesh et al, Grading and Classification of Anthracnose Fungal Disease of Fruits based on Statistical Texture Features, *International Journal of Advanced Science and Technology Vol. 52, March, 2013*

[14] Jagadeesh D. Pujari, Rajesh Yakkundimath and Abdulmunaf S. Byadgi. "Classification of Fungal Disease Symptoms affected on Cereals using Color Texture Features*", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.6, No.6 (2013), pp.321-330 http://dx.doi.org/10.14257/ijsip.2013.6.29 . ISSN: 2005-4254 IJSIP.*

[15] Al-Tarawneh, Mokhled S. "An Empirical Investigation of Olive Leave Spot Disease Using Auto-Cropping Segmentation and Fuzzy C-Means Classification." *World Applied Sciences Journal 23.9 (2013*).

[16] Vivek Chaudhari,C.Y. Patil, "Disease detection of cotton leaves using advanced image processing", *International journal of advanced computer research, June-2014 (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-4 Number-2 Issue-15,pg 653-659,June-2014.*

[17] Niket Amoda, Bharat Jadhav, Smeeta Naikwadi. "Detection and Classification of Plant Diseases using Image Processing", *International Journal of Innovative Science, Engineering and Technology 2014;1- 7.*

[18] Rastogi, A.; Arora, R.; Sharma, S., "Leaf disease detection and grading using computer vision technology & fuzzy logic," *in Signal Processing and Integrated Networks (SPIN), 2015 2nd International Conference, vol., no., pp.500-505, 19-20 Feb. 2015.*

[19] Amit Kumar Singh, Rubiya .A, B.Senthil Raja- " Classification of Rice disease using digital image processing and SVM Classifier", *International Journal of Electrical and Electronics Engineers ISSN- 2321-2055 (E) http://www.arresearchpublication.com IJEEE, Volume 07, Issue 01, Jan- June 2015.*

[20] Nikita Rishi, Jagbir Singh Gill- "An Overview on Detection and Classification of Plant Diseases in Image Processing*", International Journal of Scientific Engineering and Research (IJSER) www.ijser.in ISSN (Online): 2347-3878, Impact Factor (2014): 3.05 Volume 3 Issue 5, May 2015.*

[21] Megha P Arakeri, Malavika Arun, Padmini R K " Analysis of Late Blight Disease in Tomato Leaf Using Image Processing Techniques*", I.J. Engineering and Manufacturing, 2015, 4, 12-22 Published Online November 2015 DOI: 10.5815/ijem.2015.04.02 Available online at http://www.mecs-press.net/ijem.*

[22] Bed Prakash , Amit Yerpude – " Identification of Mango Leaf Disease and Control Prediction using Image Processing and Neural Network*", IJSRD - International Journal for Scientific Research & Development/ Vol. 3, Issue 05, 2015 | ISSN (online): 2321-0613*.

Annexure – I

S.No	Spray in nursery	Spray in main field	During transplanting
1	1. Carbendaizm- 25g or edifenphos- 25ml	2. Carbendaizm- 200g or edifenphos- 250ml or probenphos- 500ml or Tricyclazole-400g or Thiophanate methyl-500g or phroquilon – 500g/ha	P fluorescens

S. No	Varieties	Sample varieties names
1	Highly resistant	Co.4, Co.25, Co.26, TKM.1, ADT.23, Tetep, Tadukan, Peta.
2	Resistant	ADT 40, Co.25, Ponmani, ADT36, ADT49, Co37, Co43.
3	Moderately	IR20, ADT37, ASD18, Co45, IR62, IR64, JAYA, VIKAS.
4	Highly Susceptible	IR50.
5	Susceptible	IR36, PONNI, WHITE PONNI.