ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Paddy Leaf Disease Detection Classification and Remedy Finder Using K means Clustering and Image Mining

N.Neelaveni, S. Rajeswari

M.Phil Research Scholar, Dept. of C.S. Sree Saraswathi Thyagaraja College, Pollachi, Tamil Nadu, India

Assistant Professor, Dept. of MCA, Sree Saraswathi Thyagaraja College, Pollachi, Tamil Nadu, India

ABSTRACT: Data mining and knowledge discovery in databases have been attracting a significant amount of research, industry, and media attention of late. The paddy plant leaf disease, an impairment of the normal state of a plant that interrupts or modifies its better functions. Data mining and image mining techniques play an important role to identify the paddy leaf diseases. The proposed approach uses data mining and image mining based approach for identification, classification and remedy of paddy leaf diseases. The advisory may help farmers in effective decision making to secure their crop from diseases, thereby, increasing crop yield.

KEY WORDS: data mining, image mining, k-means clustering, neural network.

I. INTRODUCTION

Data mining, the extraction of hidden predictive information from large databases, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses.
Image Mining, Image Mining is the advanced field of Data Mining technique .The main objective of the Image Mining is to remove the data loss and extracting the meaningful information to the human expected needs. Image Mining is an extended branch of data mining that is concerned with the process of knowledge discovery concerning images**.**

Data Mining In Agriculture:

Data mining in agriculture is a very recent research topic. It consists in the application of data mining techniques to agriculture. Recent technologies are nowadays able to provide a lot of information on agricultural-related activities, which can then be analyzed in order to find important information.[1] A related, but not equivalent term is precision agriculture.

Objectives:

There are three objectives to achieve in this project:

- To get the processed paddy leaf image as an input.
- To segment the image using K-Means clustering algorithm.
- Finally provide the type of disease attacked in the leaf using NN classifier and severity level.

II. RELATED WORK

Takesh Saitoh examined about automatic method for recognizing a blooming flower based on a photograph. A method that extracts a boundary of blooming flower by selecting a route with minimizing a sum of the local cost divided by the route length[1].Nunik Noviana demonstrated on extracting paddy features through off line image.The

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

RGB image is converted into a binary image using variable, global and automatic threshold based on Otsu method. For removing noise a morphological algorithm is used by region filling technique. Then image characteristics like lesion type, lesion percentage, spot color, boundary color and broken paddy leaf color are extracted[2].An evaluation of framework for detection of plant leaf/stem diseases is proposed by Dheeb Al Bashish. Relying on pure naked-eye observation to detect such diseases can be expensive. The proposed framework is composed of the following steps1)the images at hand are segmented using the K-Means technique.2)the segmented images are passed through a pre-trained neural network. The proposed approach can significantly support accurate and automatic detection of leaf diseases[3].An application for detecting and classifying the quality of areca nuts image processing techniques and neural networks are used in Kuo-Yitluang Journal. Defects of areca nuts with diseases or insects were segmented by detection line (DL) method, for classifying 6 geometrical features, 3color features and defects area were used to sort the quality of areca nuts a back-propagation neural network classifier was used[4].

III. METHODOLOGY

EXISTING METHODOLOGY

Existing system uses hierarchical clustering algorithm. This algorithm is easy to implement and only requires that a kd-tree be built once for the given data points. Efficiency is achieved because the data points do not vary throughout the computation and hence, this data structure does not need to be recomputed at each stage. A natural question is whether the filtering algorithm can be improved. The most obvious source of inefficiency in the algorithm is that it passes no information from one stage to the next. But this algorithm is quite complex and does not provide significantly faster running time in practice. Firstly, the RGB images of leaves are acquired. Then RGB images are converted into Hue Saturation Value (HSV) color space representation. RGB is an ideal for color generation. Hue is a color attribute that describes pure color as perceived by an observer. Saturation refers to the relative purity or the amount of white light added to hue and Value means amplitude of light. After the transformation process, the Hue component is taken for further analysis. Saturation and Value are dropped since it does not give extra information. The proposed approach (K-Means clustering) is a valuable approach, which can significantly support an accurate detection of leaf diseases in a little computational effort.

PROPOSED METHODOLOGY

The proposed system focused on detecting the diseases of paddy leaves which help the farmers to detect disease and take proper prevention to enhance the production of paddy. We took the pictures of diseased paddy leaves and performed various preprocessing techniques on them for removing the boundary of the leaf. The main objectives of this research is to develop a data mining system for detect the paddy disease which are paddy Blast, Bacterial Leaf Blight, Rice tungro disease, Sheath Blight. This research concentrate on the image mining techniques used to enhance the quality of the image and neural network techniques to classify the paddy disease. The methodology involves image acquisition, preprocessing and segmentation, analysis and classification of the paddy disease. All the paddy sample will be passing through the RGP calculation before it proceed to the binary conversion. The proposed system involves K-Means clustering for segmentation takes less iteration and is efficient. It consumes less time to attempt disease identification. After segmentation the mostly green color pixels are masked based on specific threshold values. A neural network is trained for classification. K-Means clustering algorithm takes less number of iterations. It takes less running time and efficient for segmentation. A neural network is done for classification and is useful for identifying right type of disease. Image mining concept is used to take the input image and a neural network is to classify the disease, and clustering is used to image segmentation based on pixels. Detection of paddy leaf diseases can be done early and accurately using artificial neural network.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Steps Involved In Proposed System Are:

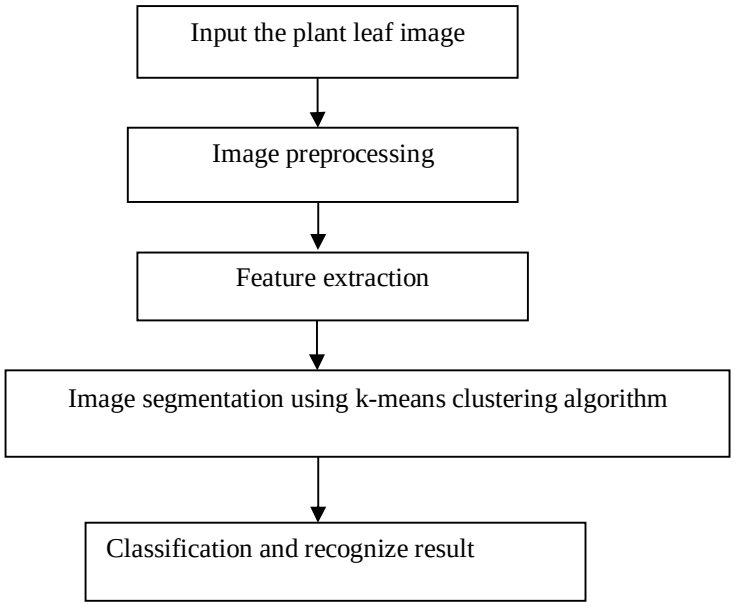


Fig: 1: Process of proposed system

Input the plant leaf:

Fig: 2: Affected leaf

Image preprocessing:

Image Pre-processing Noise gets added during acquisition of leaf images. So different types of filtering techniques are used to remove noise.

Image segmentation using K-means clustering:

In this research K-means clustering is used for segmentation. Convert Image from RGB Color Space to L*a*b* Color Space. The L*a*b* space consists of a luminosity layer 'L*', chromaticity-layer 'a*' and 'b*'.All of the color information is in the 'a*' and 'b*' layers. Classify the colors in a*b* color space using K-means clustering. Since the image has 3 colors create 3 clusters.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Extraction of Features

In feature extraction, we extract the features of diseased infected area and classify the disease according to the features. Feature extraction is used to extract texture feature of extracted diseases portion.

Classification and recognize result:

Classification task is done by using neural networks. Neural networks algorithm used to provide accurate detection. Result will show us the type of disease, severity level and also remedy for the particular disease.

K-Means Clustering Segmentation Technique

In this technique colors are segmented in an automated fashion using L*a*b*color space. K-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible.

The K-Means Algorithm Process

- ✓ The dataset is partitioned into K clusters and the data points are randomly assigned to the clusters resulting in clusters that have roughly the same number of data points.
- ✓ For each data point calculate the Euclidean distance from the data point to each cluster.
- ✓ If the data point is closest to its own cluster, leave it where it is. If the data point is not closest to its own cluster, move it into the closest cluster. Repeat the above step until a complete pass through all the data points' results in no data point moving from one cluster to another. At this point the clusters are stable and the clustering process ends.
- ✓ The choice of initial partition can greatly affect the final clusters that result, in terms of inter-cluster and intra-cluster distances and cohesion.

THE STEPS IN K-MEANS CLUSTERING IN THE COURSE OF DISEASE DETECTION

1. Read image.

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

3. Classify the Colors in a* b* Space using K-Means Clustering.

4. Label every pixel in the image using the results from k-means.

5. Create images that segment the image by color.

6. Separate the infected part and uninfected part.

Percentage Infection Calculation (P)

After calculation of the total leaf area (AT) as well as the diseased area (AD) of the leaf, the percentage infection (P) is calculated by using the following equation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

$P = AD/AT \times 100$

IV. RESULTS AND DISCUSSION

The large volume of data generated as well as the high complexity of its relations has generated. New proposals for models and algorithms those are able to handle this data efficiently and effectively. In this research problem of K-Means clustering and image mining is used for proposed system. Various experiments were conducted to judge the performance of the proposed system and in this chapter the result obtained are presented.

Pre-Processing Result:

Following results shows the enhanced image of the original image

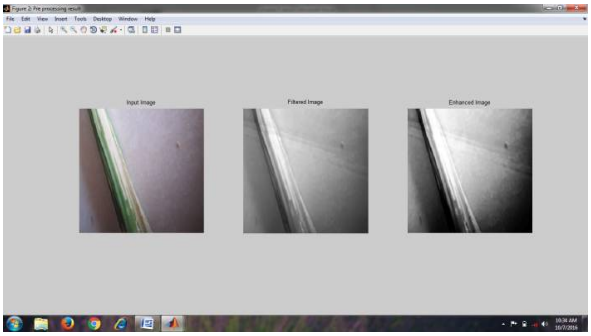


Fig: 3 : Pre-Processing Result

Histogram Result:

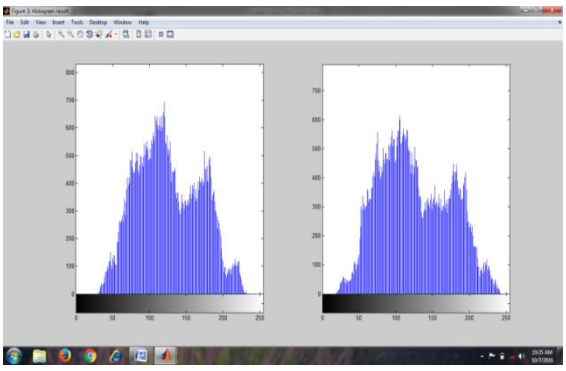


Fig: 4: Histogram Result

Above result shows the pixel difference between the original image and the gray image.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

K-means Result:

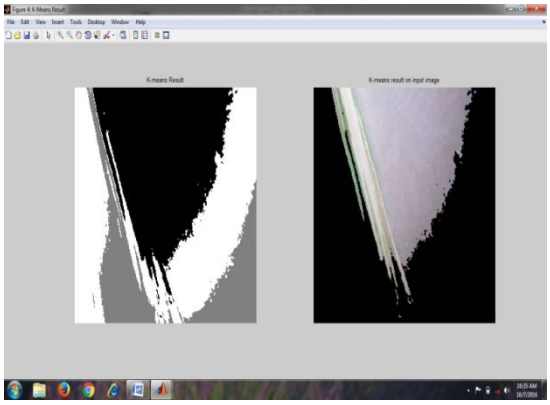
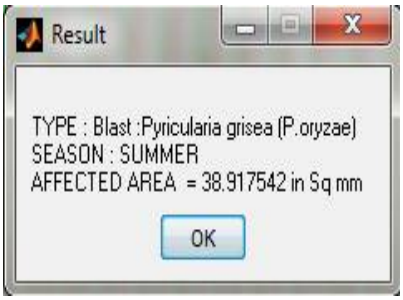


Fig: 5: K-means Result

Above result shows the infected area in the cluster 1 and green area is shown in the cluster. So K-means cluster is used to index the diseases affected area.

Result and Remedy:



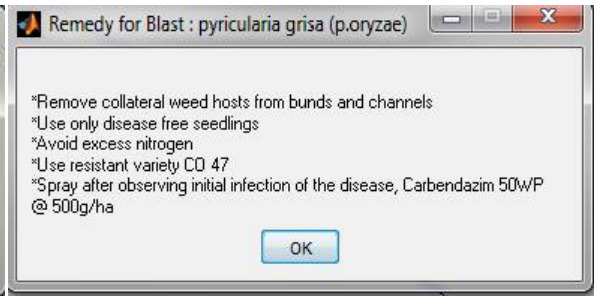


Fig:6: Result **Fig: 7: Remedy**

Above result shows the Result of the disease and Remedy for affected leaves

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Comparison of existing and proposed system:

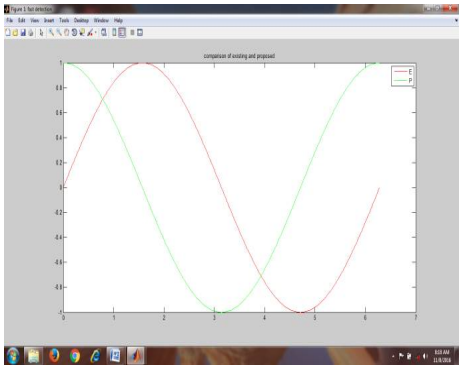


Fig: 8: Comparison of existing and proposed system (Fast detection)

Above showed result is the comparison of fast disease detection of the result in existing and proposed system.

V. CONCLUSION

The proposed system is accurate and fast detection of leaf disease. K means clustering and neural network is used for segmentation and classification of diseases which affected by paddy leaf diseases and provides treatment measures for classified disease type. This system use input as set of paddy leaf images and produce type of disease is attacked and severity level, and also it suggest the remedy for the particular attacked disease. It helps the farmers to yield in a better way and it reduces the cost and environmental pollution. The result obtained helps farmers/agronomists in making effective decisions early and efficiently to protect their crops from heavy loss.

REFERENCES

1. Takesh Saitoh, Kimiya Aoki and Toyohisa Kaneko, "Automatic Recognition of Blooming Flowers", Proc.of the 17th International conference on Pattern Recognition,Vol 1,pp.27-30,2004.
2. Nunik Noviana Kurniawati, Sili Norul Huda Sheikh Ad Abdullah, Salwani Ad Abdullah, Saad Abdullah "Texture Analysis for Diagnosing Paddy Disease", Electrical Engineering and Informatics"Vol.1,pp.23-27,2009.
3. Dheeb Al Bashish, Malih Braik and Suliman Bani Ahm "A Frame work for Detection and classification of plant leaf and stem Diseases" Signal and Image Processing,pp.113-118,2010.
4. Kuo-Yitluang "Detection and classification of areca nuts with machine vision" Computers and Mathematics with Application,Vol64,pp.739-746,2012.
5. Aakanksha Bhatnagar, Shweta P. Jadye, Madan Mohan Nagar," Data Mining Techniques & Distinct Applications: A Literature Review", International Journal of Engineering Research & Technology (IJERT) Vol. 1 Issue 9, November- 2012.
6. D Ramesh , B Vishnu Vardhan, "Data Mining Techniques and Applications to Agricultural Yield Data", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 9, September 2013.
7. Ruxandra-Ştefania PETRE, "Data mining in Cloud Computing", Database Systems Journal vol. III, no. 3/2012.
8. Nikita Jain, Vishal Srivastava ,"Data Mining Techniques: A Survey Paper", Ijret: International Journal of Research in Engineering and Technology, Volume: 02 Issue: 11 | Nov-2013.
9. Dr. M.H.Dunham, "Data Mining, Introductory and Advanced Topics", Prentice Hall, 2002.
10. Time Series Analysis and Forecasting with Weka, http://wiki.pentaho.com/display/DATAMINING.
11. Ghulam Mustafa Choudhary, Vikrant Gulati, ―Advance in Image Processing for Detection of Plant Diseases‖, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 7, July 2015
12. Sachin D. Khirade, A. B. Patil, ―Plant Disease Detection Using Image Processing‖, 2015 International Conference on Computing Communication Control and Automation, 978-1-4799-6892-3/15 $31.00 © 2015 IEEE.
13. Khushal Khairnar, Rahul Dagade, ―Disease Detection and Diagnosis on Plant using Image Processing – A Review‖, International Journal of Computer Applications (0975 – 8887) Volume 108 – No. 13, December 2014.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

BIOGRAPHY

Ms.N.NEELAVENI, is a Research scholar in the computer science Department, Sree Saraswathi Thyagaraja College,Pollachi. She received Master of Computer science (MSC CS) degree in 2015 from Bharathiar university, Coimbatore, Tamil nadu.

Ms.S.RAJESWARI, is a Assistant professor in the department of MCA, Sree Saraswathi Thyagaraja College,Pollachi, Coimbatore ,Tamil nadu.