Recognition of Diseases in Paddy Leaves Using kNN Classifier

Suresha M and Shreekanth K N
Department of P.G. Studies and Research in Computer Science, Kuvempu University, Shankaraghatta, Karnataka, India
srit_suresh@yahoo.com and shreekant.kn@gmail.com

Thirumalesh B V
Department of P.G. Studies and Research in Applied Botany, Kuvempu University, Shankaraghatta, Karnataka, India
thirubv@gmail.com

***Abstract*—Paddy is the most important crop in Asian country. Most of the people depend on rice for their food, so rice is considered as staple food in Asian country. Rice plant is affected by many diseases that affect the farmers in yield loss. In this paper, proposed a method for identification of Blast and Brown Spot diseases. Global threshold method has been applied and kNN classifier has been used to classify the data. The result 76.59% has been obtained for the proposed method.**

Keywords- Blast Disease; Brown Spot Disease; kNN Classifier; Paddy leaves; Segmentation;

I. INTRODUCTION

In agriculture, rice is a staple food for more than half of the human population. Rice has the major source of food in 144 countries and in all the continents [4]. Plant diseases are significant issues, as it causes significant reduction in the quantity and quality of the agriculture production. Detection and classification of diseases is an important task in agriculture field. In general, traditional approaches are made by experts to identify diseases by naked eye. Observing the changes in leaf color diseases are identified [6]. Usually diseases spread from one place to another place by carry agents, like water, air, soil, insects and infected seeds. The plant disease may be broadly classified into three types. They are bacterial, fungal and viral diseases [10]. In this paper Blast and Brown Spot diseases are identified by using image processing and pattern recognition approach. Diseased plant can greatly impair productivity and sometimes it destroys a crop. Losses include direct loss and indirect loss. Direct loss include reduction in plant stands, lodging, spotted kernels, fewer, smaller grains per plant and a general reduction in plant productivity. An indirect loss includes the cost of fungicides used in managing diseases. This application has got some drawbacks such as high cost and reduced yields, which is associated with special cultural practices that reduce diseases. This is not so helpful in producing maximum yields [3]. In this paper, brown spot and blast diseases have been identified using geometrical features like Area, Major axis, Minor axis and Perimeter of the infected area of the leaves.

II. LITERATURE REVIEW

Literature survey has been conducted pertaining to paddy leaves diseases and significant works have been done by the researchers. In [11] disease identification and classification has been done. Authors have proposed two methods for classification of diseased rice leaves and uninfected rice leaves. A radial features with eight direction have been extracted from the infected region of leaves. Histogram has been generated out of these features. In this first method based on peaks of a histogram the infected and uninfected diseases have been classified. In second method with same features comparison study has been conducted using Bayes classifier and Support Vector Machines (SVMs) classifiers and the authors have found 79.5% and 68.1% accuracy for these methods respectively. In [6], authors identified the disease spot using different color model like HSI, YCbCr and CIELAB. Using otsu's algorithm segmentation process is carried out. Before segmentation images are smoothened using Median filter to avoid noise. According to authors camera flash can be removed using CIELAB color model. For the experiments, authors used different family of leaves i.e. Monocot and Dicot plant leaves with both noise free and noisy background. In [5], classification and recognition of paddy diseases are made, RGB images were collected from Lousiana State University Auricular Center (www.lsuagcenter.com) and the RGB images were converted to HSV color model. Authors have applied novel Gaussian mean method for segmentation. Polar Statistics features have been used for classification. In [9], proposed a method for disease detection of an agricultural plant leaf disease by using image processing. At first, author collects the infected leaf images by digital camera and then RGB images were converted to HSV color Model. Segmentation of infected region has been done by considering hue component of HSV color space and segmented part is again further segmented into number of patches 32X32. The authors are chosen significant information such that no infected region is missing. Statistical features have been extracted out of GLCM. In [8], detected and diagnose the agricultural plant leaf disease based on morphological changes of feature. Preprocessing has been done by using analyzing histogram of an image and image contrast has been adjusted by applying threshold. Segmentation can be done using both classical and Fuzzy C-means (FCM) algorithms. Features like color, shape and texture are used for segmentation. Artificial Neural Network (ANN) is proposed to perform classification. Authors proposed to use statistical analysis technique called ANOVA for determining the optimal configuration of neural network.

978-1-5090-4307-1/17/$31.00 ©2017 IEEE

III. PROBLEM DEFINITION

There are many different types of diseases in paddy crop, disease identification is most difficult task and it will take more time to analyze. Based on morphological changes, leaf evaluations of diseases are made by experts in a naked eye and it is expensive. Traditional human inspection can be replaced by using machine vision technology in the field of paddy crop. In this paper identification of blast and brown spot disease of paddy leaves are addressed.

IV. PROPOSED METHODOLOGY

In the proposed methodology, segmentation has been carried out using otsu segmentation method. Geometrical feature like Area, Major Axis, Minor Axis and Perimeter of diseased part of paddy leaves have been extracted. kNN classifier has been used for classification of brown spot and blast disease of paddy leaves.

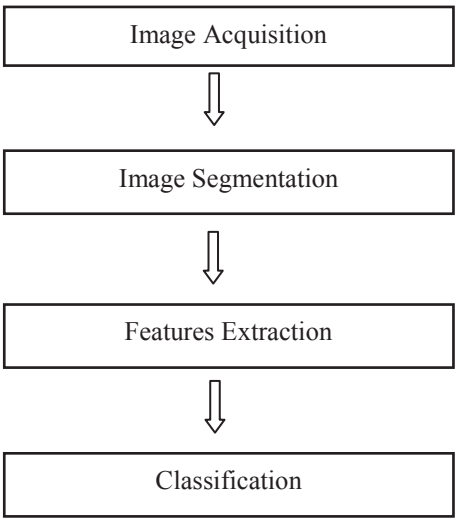


Fig 1. *Typical work flow for proposed methodology*

A. Image Acquisition

Images have been captured in paddy fields, Shivamogga district, Karnataka state. The images were captured with digital camera with 18.1 megapixels in day light with black background such that the leaf objects are filled the camera view.

Fig. 2. *Sample Blast disease paddy leaves*

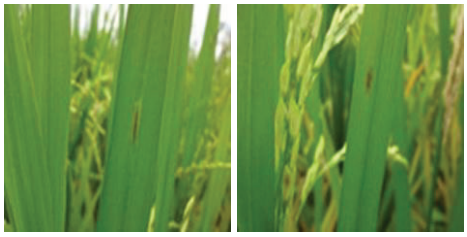


Fig. 3. *Sample Brown Spot diseased Paddy leaves*

B. Image Segmentation

Segmentation is a process of extracting a region of interest [7]. Segmentation is the most difficult task in paddy leaves, because there is so much contrast between diseased part of the leaf and healthy part of the leaf. RGB image of paddy leaves have been converted into HSV color model and the results of hue component are shown in figure 4 and figure 5. Hue component of HSV color model has been used for segmentation because Hue is dominant and discriminative intensity values between infected part and normal part of paddy leaves [12]. Otsu segmentation method with global threshold has been used for segmentation. Sample segmentation experimental results are shown in figure 6.

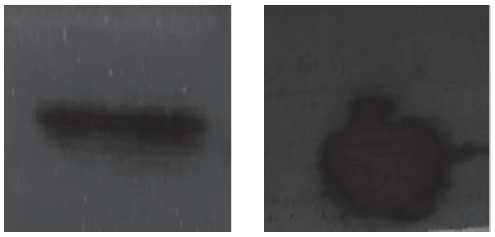


Fig. 4. *Hue Component of Blast*

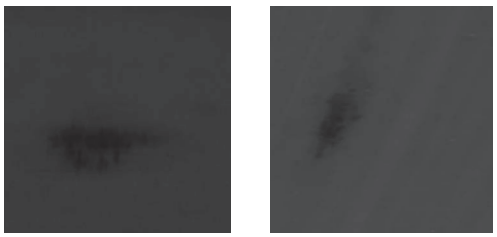
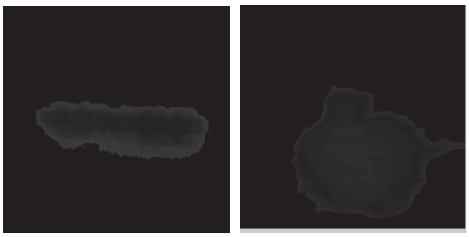


Fig. 5. *Hue Component of Brown Spot*

(a) Binary Segmentation Results of Blast Disease

(b) Binary Segmentation results of Brown Spot Disease



(c) Segmented grey scale image of Blast

(d) Segmented grey scale image of Brown Spot

Fig. 6: *Segmented results of Blast and Brown Spot Disease*

C. Features Extraction

The purpose of feature extraction using connected component reduced the image data by measuring certain properties of each segmented region. [1] Here Blast diseases are having elliptical or spindle shape or sometimes it resembles diamond shape (center part will be wide and end points towards both the end are pointed in nature). Brown Spot diseases are round in nature. So the following geometrical features have been extracted from paddy leaves for classification of Blast and Brown Spot disease.

1) Area: The number of pixels in the shape of the diseased area [2]. In proposed feature extracted attribute, area which calculates the area of infected region that area will differentiate from Blast and Brown spot diseased segmented images.

2) Minor Axis Length: To the minor axis is drawn perpendicular to the major axis where this line has the maximum length. Once the end points of the minor axis have been found, its length is given by the same equation as the major axis length. It is also called the object width [2]. In proposed attribute the length of the diseased segmented part will be considered, because the blast dieseased images have round or circle shape and blast diseased images are in diamond shape. When compared to brown spot diseased images are shorter in length compared to blast diseased segmented images length, so we considered Minor Axis as one of the attributes.

3) Major Axis Length: The major axis points are the two points in an object where the object is more elongated and where the straight line drawn between these two points is the longest. Major axis points are calculated by all possible combinations of perimeter pixels where the line is the longest. The length of the major axis is given by:

$$MajorAxisLength = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad (1)$$

Where (x_1,y_1) and (x_2,y_2) are the coordinates of the two end points of the major axis [2]. In proposed attribute the width with respect to the segmented image were difference in brown spot and blast segmented image. Centre part of the Brown spot images will more widen compare to blast images, so we are considered Major Axis as one of the attribute.

4) Perimeter: Perimeter is an important feature of an object. Contour based features which ignore the interior of a shape, depend on finding the perimeter or boundary points of the object [2]. The perimeter of an object is given by the integral as follows:

$$T = x(t) + y(t)dt \qquad (2)$$

D. Classification

Authors have used supervised classifier for classification of Blast disease and Brown spot disease. Training data set has a pair of order set of features and corresponding labels. Let's say h_x is a feature vector and y_i is the corresponding label vector. In training data set label is known where as in testing set label in unknown. kNN classifier at first determines k nearest neighbors and then determines labels for the sample based on neighbor weight. Let us say a training set D is has x_i training samples. The training and testing features are normalized to [0,1]. The training samples are labeled with the class label y_o. The objective is to classify testing sample q for each $x_i \in D$ and the distance between q and x_i is calculated using equation (3).

$$d(q, x_i) = \sum_{f \varepsilon F} w_f \delta(q_f, x_{if}) \quad (3)$$

The equation for finding k nearest neighbors is selected for continuous and discrete attributes is given below.

$$\delta(q_f, xi_f) = \begin{cases} 0 & f \ \ discrete \ \ and \ \ q_f = x_{if} \\ 1 & discrete \ \ and \ \ q_f \neq x_{if} \\ |q_f - x_{if}| & f \ \ continuous \end{cases} \quad (4)$$

These are variety of methods for identifying *k* nearest neighbors to determine the class. The most popular method is majority rule for assigning a label to unknown sample.

V. Experimental results

The paddy leaves images are resized to size 255 X 255 pixel resolution to improve accuracy of the algorithms. Image data has been created with 330 images out of these 60% of images are used for training and 40% of images are used for testing. For geometrical features kNN classifier has given 76.59% of accuracy. S. Phadikar et al,. 2012 have used SVM Classifier for the radial features and obtained accuracy of 68.1% and Bayes classifier has given accuracy of 79.5%. The limitation of Bayes classifier is it works for only offline data. It shows that our proposed method has given better results.

VI. Conclusion

In this paper classification of paddy leaves diseases like Brown Spot and Blast is carried out. RGB images are converted into HSV color images for segmentation. Classification has been done using kNN classifier with geometrical features like Area, Major Axis, Minor Axis and perimeter. The plant disease may be broadly classified into three types. They are bacterial, fungal and viral diseases [10]. In the proposed method only fungal diseases have been considered. In the future work will be considering mentioned three types of diseases.

Acknowledgment

Authors like to thank faculty members of the Department for providing facilities to carry out the research work and the support during the work and also thank farmers for giving me an opportunity for image data collection and providing knowledge about the paddy leaves diseases.

References

[1] Adam Sparks, N P Castilla, and CM Vera Cruz, "Brown spot and Blast", Rice Knowledge Bank, International Rice research institute, IRRI.

[2] Costa, L. F., Cesar, R. M. "Shape Analysis and Classification," Boca Raton, Florida. CRC Press. 2000

[3] Don Groth, Clayton Hollier and Chuck Rush, "Louisiana Rice Production Handbook - Disease Management," chapter 6 Pub, 2321, (3M) Rev. 5/14, 2009, pp. 82-105.

[4] Jae Ok Kim, Seung Heon Lee and Kyu sang, "Efforts to improve biodiversity in paddy field ecosystem of South Korea," Invited Article 1, 2011, pp. 25-30.

[5] Md. Zahangir Alom and Hyo Jong Lee, "Gaussian Mean Based Paddy Disease Segmentation," 3rd International Conference on Machine Vision (ICMV 2010), IEEE 2010, pp. 522-525.

[6] Piyush Chaudhary, Anand K. Chaudhari, Dr. A. N. Cheeran and Sharda Godara, "Color Transform Based Approach for Disease Spot Detection on Plant Leaf," International Journal of Computer Science and Telecommunications, Vol. 3, Issue 6, June 2012 pp 65-70.

[7] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins "Digital Image Processing Using MATLAB," Second Edition. Mc Graw Hill Education, 2010, pp.532.

[8] Sachin B. Jagtap, Shailesh M. Hambarde "Agricultural Plant Leaf Disease Detection and Diagnosis Using Image Processing Based on Morphological Feature Extraction," IOSR Juunal of VLSI and Signal Processing, Vol. 4, Ver. I, Sep – Oct. 2014, ISSN No. : 2319–4197, pp 24-30

[9] Sanjay B. Dhaygude and Nitin P.Kumbhar "Agricultural plant Leaf Disease Detection Using Image Processing," International Journal Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, 1, January 2013, ISSN : 2278 – 8875, pp. 599-602.

[10] S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," Agric Eng Int : CIGR Journal Vol. 15, No.1 211, March 2013.

[11] S. Phadikar, J.Sil, and A. K. Das "Classification of Rice Leaf Diseases Based on Morphological Changes," International Journal of Information and Electronics Engineering, Vol. 2, No.3, May 2012 pp. 460-463.

[12] Suresha M and Ajit Danti "Construction of Co- occurrence Matrix using Gabor Wavelets for Classification of Arecanuts by Decision Trees," International Journal of Applied Information Systems, Vol. 4 - No. 6, December 2012.ISSN : 2249 -0868, pp. 33-39.