ISSN 2320-2602

International Journal of Advances in Computer Science and Technology (IJACST), Vol.3 , No.12, Pages :11-16
Special Issue of ICCEeT 2014 - Held on 22nd December 2014, Dubai

Identification of Sugarcane Leaf Scorch Diseases using K-means Clustering Segmentation and K-NN based Classification

Umapathy Eaganathan[1], Dr.Jothi Sophia[2], Vinukumar Luckose[3], Feroze Jacob Benjamin[4]

[1]Research Scholar in CS, Vels University-India, umapathy_mca2k5@yahoo.com
[2]Director & Principal of Nursing, CSI JACON-India, 3jroses@gmail.com
[3]Faculty in EEE, Nilai University-Malaysia, vinukumar@nilai.edu.my
[4]Student in Mechanical, Linton University, Malaysia

Abstract: Sugarcane is the most important plant in Indian agriculture as well as in economy also. In this world, India is a second largest production of sugarcane. Sugar is the by-product of sugarcane and it participates in our day today life in all the essential food items. Farmers are much interested to gain profit when they harvest from their land. Normally the plants may affect by the natural deficiency like water, sunlight and soil. Most of the plants leaves are affected by fungal diseases caused by viruses. Same way the sugarcane also affected by the bacterial diseases and it will leads the farmer to spend more time to identify the diseases and it increase cost to prevent from the diseases. The samples of sugarcane leaf scorch taken for the experiment. This leaf scorch caused by the bacterium which is popularly known as "Xylella Fastidiosa". This paper emphasizes to detect and classify the sugarcane leaf scorch diseases automatically by using image processing techniques they are acquisition, filtering, segmentation and feature extraction. The accuracy of this experiment is given 95%.

Keywords: Feature Extraction, K-Nearest Neighbor, Median Filtering, Otsu Method

INTRODUCTION

In India Sugarcane is one of the main crops in different states like Tamilnadu, Andhra, Telangana, Uttar Pradesh and Karnataka. In Tamilnadu the sugarcane called as "கரும்பு," "karumbu" and in the national language of India is Hindi, called the sugarcane is "गन्ना," "Ganna". Several diseases caused by the viruses to affect stem and leaves. The fungi caused diseases are the most predominant it appear as spots on the leaves. The various types of diseases on leaf determine the quality, cost and production of yield. Disease symptoms leaves are taken from a standard digital camera and smart phones etc. Fungi diseases are identified primarily from the plant morphology, with their reproductive structures [1]. Sugarcane is chewed in all of the producing countries because of its sweetness. Sugarcane juice is obtained by pressing the sugarcane and is mostly used to sweeten foodstuffs and it can also be consumed as fresh or fermented juice [2]. Plant diseases are important factors can cause significant reduction in both quality and quantity of crops in agriculture production. Therefore, identification and classification of diseases is an important task and it may prove benefits in monitoring large fields of crops. The disease leaves analyzed by computer image processing technology and extract the leaf features according to color, texture and other characteristics. Here median filter is used for smoothing. The following Fig. 1 shown the steps involved in this paper.

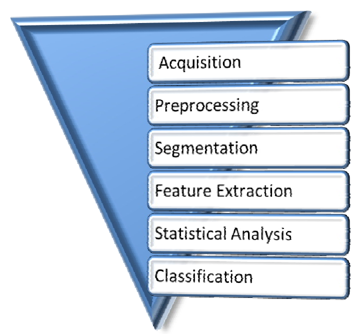


Fig.1 Procedure of disease detection

IMAGE ACQUISITION

In this paper the leaf scorch disease leaves taken for the consideration of disease detection from sugarcane field. First the images of sugarcane leaf scorch leaves are going to acquire by a latest digital camera. The images taken by using digital camera Canon EOS 600D 18MP Digital SLR Camera and images shown in Fig. 2. After this images can be transferred to PC (Personal Computer) and the image processing techniques applying by using MATLAB R2012a. If there are noises in the acquired image then different pre-processing methods should be followed. Different kinds of noises exist in an image and multiples of noise reduction techniques available to perform de-noising. For the detection and classification of certain diseases in plant leaves

ISSN 2320-2602

International Journal of Advances in Computer Science and Technology (IJACST), Vol.3 , No.12, Pages :11-16
Special Issue of ICCEeT 2014 - Held on 22nd December 2014, Dubai

different morphological features of the leaves are used.

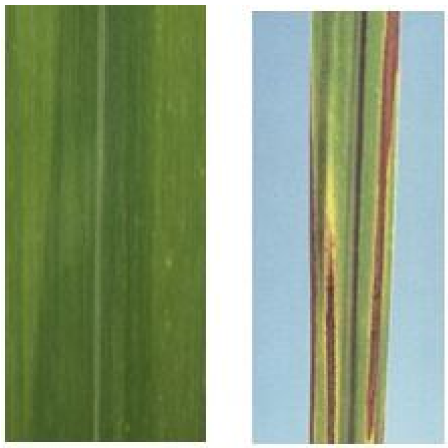


Fig.2 Normal Leaf Leaf Scorch

In leaf scorch disease first, the disease causes red to dark brown minute spots with a yellowish halo on the leaf and become as large as 5-17 x 0.3-1 cm, with yellow margin. The lesion centre becomes necrotic and pycnidia develop as minute blackish dots on the upper leaf surface.

IMAGE PRE-PROCESSING

The pre-processing task includes some procedures for the images need to be enhanced. The sugarcane leaf scorch image is in RGB color format. The RGB image is first converted to a grayscale image then gray-level threshold segmentation is adapted and the binary image is taken. The main obstacle in disease spot detection is noise which usually introduced by camera flash, change in illumination and noisy background. Hence median filter and morphological operators are applied and Fig. 3 shown histograms created for both normal leaf and leaf scorch to depict image statistics in an easily interpreted visual format. It can also be used to determine the type of processing has been applied to an image. Normally the histogram describes the frequency of the intensity values that occur in an image.

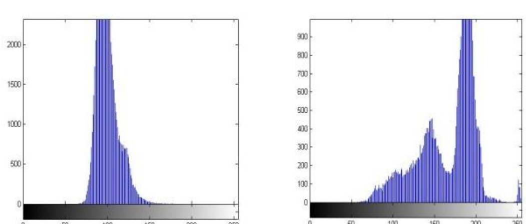


Fig.3 Histogram of normal and leaf scorch

After producing histogram for the leaf samples now the RGB image is converted into gray level images in order to proceed for the image process. The method followed to convert RGB to gray level in MATLAB is shown below

```
b=rgb2gray(a);
```

Here the argument 'a' reads the input image then RGB image to gray level is converted finally stored in the variable 'b'. Now the variable 'b' contains the gray level image. The conversion of RGB to grey level shown in the following Fig. 4.

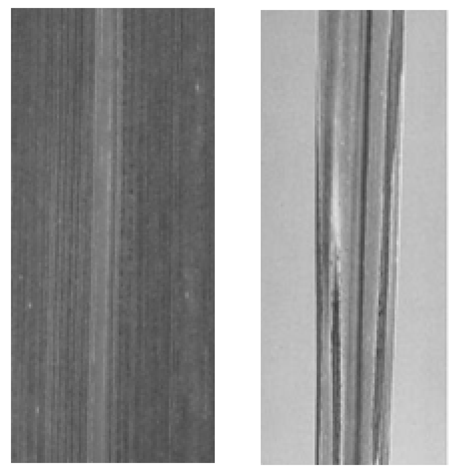


Fig.4 Gray level images of normal and leaf scorch

Noises which may be brought from the process of image collection and lot of information which may be easily leaded from the image. So the image with low quality must be smoothed by filter. Different kinds of noises exist in an image and different noise reduction techniques are available to perform de-noising. Selection of de-noising algorithm depends on the application. The chosen filter is a non-linear filter.

The Median filter performs better with salt and pepper noise. Median filter erases black dots called the pepper and fills in the white holes in the given image, called salt. It simply replaces each pixel values by the median of the intensity level in the neighborhood of the pixel [3]. The following Fig. 5 images shown after applying the median filtering techniques.

ISSN 2320-2602

International Journal of Advances in Computer Science and Technology (IJACST), Vol.3 , No.12, Pages :11-16
Special Issue of ICCEeT 2014 - Held on 22nd December 2014, Dubai

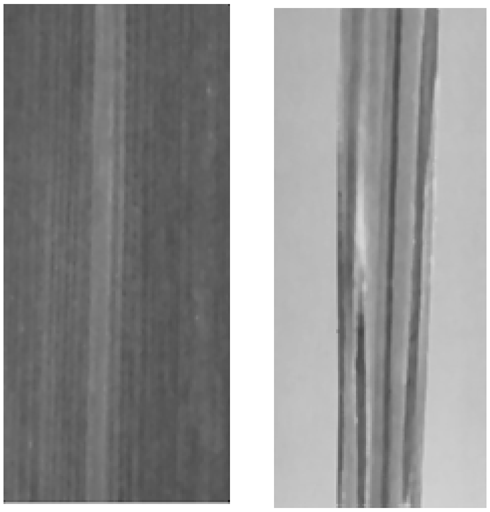


Fig.5 Filtered normal and leaf scorch leaves

Now the noises are removed from the images by the filters for the normal leaf and leaf scorch.

SEGMENTATION

The image will be segmented into different parts according to the region of interest. Image segmentation is to divide the image into different meaningful regions. Simply defined as image segmentation is to separate the object from background for the processing. Every cluster is represented by one of the objects in the cluster. Clustering is a way to divide groups of objects. K-means clustering treats each object as having a location in space. K-means method finds out clusters from the image [3]. The segmentation approaches are employed in the present work namely color based segmentation using K-means clustering. The K-means clustering followed the objects into three clusters.

OTSU METHOD

Among all the segmentation methods, Otsu method is one of the simplest methods for image thresholding because of its simple calculation. The Otsu method works on global thresholding and it requires computing a gray level histogram before running [5]. This method proposed based on discriminate analysis. In Otsu the pixels are separated into two classes C0 and C1 (e.g., objects and background) at grey-level t, i.e., C0 = {0, 1, 2,…, t} and C1 = {t + 1, t +2,….L-1}

$$\lambda = \frac{\sigma^2 B}{\sigma^2 W}, \ \eta = \frac{\sigma^2 B}{\sigma^2 T}, \ K = \frac{\sigma^2 T}{\sigma^2 W} \qquad (1)$$

Let $\sigma^2 W$, $\sigma^2 B$ and $\sigma^2 T$ be the within-class variance, between-class variance and the total variance respectively.

COLOR BASED K-MEANS SEGMENTATION

Pre-processed image that is filtered image is input to the K-means clustering algorithm with K value being two and Squared Euclidean Distance Metric is used to quantify how close two objects are to each other. Here, the value of K is chosen based on the heuristic that green pixels in the image belong to healthy portion of the leaf sample and should be separated out from that of diseased portion. Thus, K-means clustering approach divides objects into two clusters [4]. The procedure is as follows:

Step 1: Read in filtered image

Step 2: Convert image from RGB color space to L*a*b* color space

The L*a*b* space contains of a luminosity layer 'L*,' chromaticity layer 'a*' indicating where color falls along the red-green axis and chromaticity layer 'b*' indicating where the color falls along the blue-yellow axis.

Step 3: Classify the colors in 'a*b*' space using K-means clustering with Squared Euclidean Distance Metric.

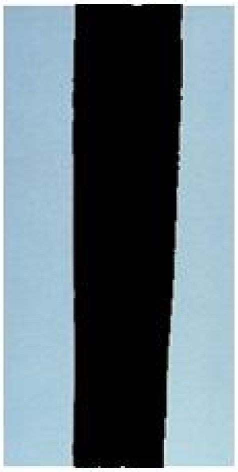
Step 4: Label every pixel in the image using the results from cluster indices returned by K-means.

Step 5: Create images that segment the input image by color based on pixel labeling.

The output of K-means clustering for a leaf infected with leaf scorch disease is shown in Fig. 6. It is clearly observed from Cluster 3 contains infected object of scorch disease. Furthermore, clusters 2 and image labeled by cluster contain the intact parts of leaf, although they are distinct from each other. However cluster 1 represents the black background of the leaf which can be discarded primarily. The result of this approach can be observed that objects in cluster.

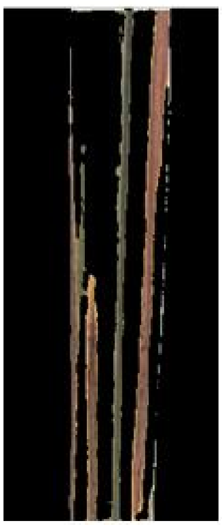
ISSN 2320-2602

International Journal of Advances in Computer Science and Technology (IJACST), Vol.3 , No.12, Pages :11-16
Special Issue of ICCEeT 2014 - Held on 22nd December 2014, Dubai



Cluster 1

Cluster 2



Cluster 3

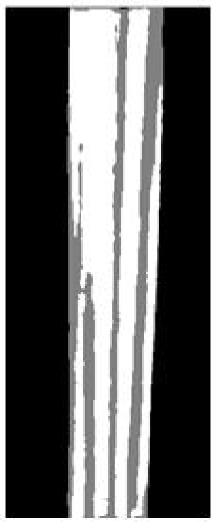


Image labeled by cluster

Fig.6 Output of K-means clustering for a leaf that is infected by scorch disease

FEATURE EXTRACTION

The input data to an algorithm will be transformed into a compact representation set of features. The input data transform into the set of feature is known as feature extraction. If the feature extraction is carefully chosen then the features will extract the relevant information from the input data.

COLOR FEATURE EXTRACTION

The color conversion is performed before extracting the color features. The color images are recognized by quantifying the distribution of color throughout the image and change in the color. The quantification obtained by computing mean and standard deviation for a given image. The color features represent the global characterization of an image. The mean and standard deviation are the features extracted as color features. These are calculated using the following formulae [5].

Standard deviation : $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2}$ (2)

Mean: $\mu = \frac{1}{N} \sum_{i=0}^{N-1}(x_i)$ (3)

The total two color features are extracted from two color channels from color model. In L*a*b color model, 'a' and 'b' are the color channels. The Haralick assumed that the texture feature information is contained in this matrix and texture features are then calculated from it. Haralick extracted 14 parameters from the co-occurrence matrix, but only three features used to give good results in a classification task [6].

Energy: i,j=0,1,…N-1 is the gray level

$$\sum_{i,j=0}^{N-1} P(i,j)^2 \quad (4)$$

Energy reflects the image gray scale distribution uniformity degree.

Entropy: $-\sum_{i,j=0}^{N-1} P(i,j)\, log\, (i,j)$ (5)

Entropy is a measure of the image with the amount of information it may be image texture information. It says the texture in the heterogeneous degree or complexity.

Contrast: $\sum_{i,j=0}^{N-1} (i,j)^2 P\,(i,j)$ (6)

The contrast reflects the degree of image sharpness and texture of deep grooves.

CLASSIFIER SELECTION

In pattern recognition, the K-nearest neighbor algorithms (K-NN) is a method for classifying objects based on closest training examples in the feature space [7]. K-NN is a type of instance-based learning where the function and computation is deferred until classification. So the K-nearest neighbor algorithm is the powerful of all machine learning algorithms. The object is classified of its neighbors and being assigned to the class most common to its K nearest neighbors. Here K is an integer and if K = 1 then the object is assigned to the class of its nearest neighbor.

International Journal of Advances in Computer Science and Technology (IJACST), Vol.3 , No.12, Pages :11-16
Special Issue of ICCEeT 2014 - Held on 22nd December 2014, Dubai

EXPERIMENT RESULTS

The trained features are obtained from the sugarcane leaves. Each and every class is tested differently to differentiate from normal and diseased leaves and shown in Table 1. The experiments for the proposed approach were conducted on a personal computer with an Intel Core 2 Duo Processor and 2 GB RAM configured with Microsoft Windows 7 and MATLAB R2012a software with image processing toolbox.

Test Dataset	Trained Dataset	
	Normal	Scorch
Normal	20	0
Scorch	1	19

Table 1. Classification Results for K-NN Classifier

The results shown in Table 1 were obtained by using the K-NN Classifier. The number in a particular cell indicates that correctly classified leaves against all the conditions of the leaves. The overall accuracy using K-NN Classifier is 95% shown in Fig.7. The CR called as classification rate.

$$CR = \frac{No.of\ correctly\ classified\ test\ data}{Total\ No.of\ test\ data}\ X\ 100$$

From the results it is clear that texture analysis can be used for feature extraction and also it is used to observe that normal and scorch spot diseased leaves can be classified easily.

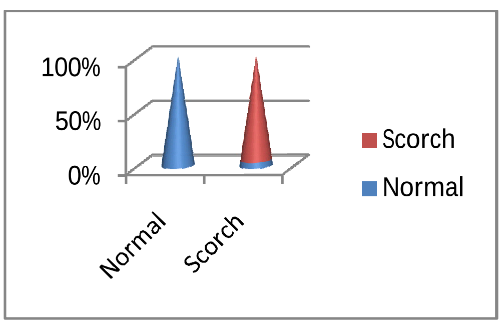


Fig.7 Classification of normal and scorch leaf in percentage

CONCLUSION

Initially two sets of sugarcane leaves (Normal and Leaf Scorch) were collected and investigated from the sugarcane field in Madurai, Tamilnadu, India. For filtering Median filter performs better with salt and pepper noise. Image processing algorithms were used for feature extraction and classification. Images of the leaf surface were extracted from the original RGB images and then converted into L*a*b color model with maximum quantization level. In color model L*a*b is accurately detected disease and results are not affected by background, type of leaf, type of disease spot and camera flash. The K-means algorithm is working for gray scale images and for better performance. The features obtained are used for classification using K-NN classifier. For future research, work can be extended for development of hybrid algorithms such as other clustering methods and Neural Networks in order to improve the recognition rate of the final classification process. Further needed to compute amount of disease present on leaf.

ACKNOWLEDGMENT

The authors thank the corresponding universities and colleges for their support and encouragement of this research.

References

[1] Sanjay B.Dhaygude, Nitin P.Kumbhar, "Agricultural Plant Leaf Disease Detection Using Image Processing," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 2, pp. 592-602, Jan. 2013.

[2] Franz Augstburger, Jorn Berger, Udo Censkowsky, Petra Heid, Joachim Milz, Christine Streit, *Organic Farming in the Tropics and Subtropics Exemplary Description of 20 Crops*, 1st ed. Naturland e.V, German, 2000, ch. 2, pp. 6-12.

[3] Arti N.Rathod, Bhavesh A. Tanawala, Vatsal H.Shah, "Leaf Disease Detection Using Image Processing and Neural Network," *International Journal of Advance Engineering and Research Development*, vol. 1, pp. 28-31, June 2014.

[4] Sunita I.Naik, Vivekanandreddy, S.S.Sannaki, "Plant Disease Diagnosis System for Improve Crop Yield," *International Journal of Innovations in Engineering and Technology*, vol. 4, pp. 198-204, June 2014.

[5] Ch.Hima Bindu, K.Satya Prasad, "An Efficient Medical Image Segmentation Using Conventional OTSU Method," *International Journal of Advanced Science and Technology*, vol. 38, pp. 67-74, Jan. 2012.

[6] Miss Hetal J.Vala, Prof.Astha Baxi, "A Review on Otsu Image Segmentation Algorithm," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 2, pp. 387-389, Feb. 2013.

[7] Neelamma K.Patil, Virendra S.Malemath, Ravi M. Yadahalli, "Color and Texture Based Identification and Classification of food Grains using different Color Models and Haralick features," *International Journal on Computer Science and Engineering*, vol. 3, pp. 3669-3680, Dec. 2011.

ISSN 2320-2602

International Journal of Advances in Computer Science and Technology (IJACST), Vol.3 , No.12, Pages :11-16
Special Issue of ICCEeT 2014 - Held on 22nd December 2014, Dubai

Mr. Umapathy Eaganathan was born in Walajapet, India in 1982. He received his Bachelor Degree in 2002 from APCOS, Arcot, India. He has a Masters Degree in Computer Applications in 2005 from Vels College of Science, Chennai, India. He has completed his Master of Philosophy (Computer Science) in 2008 from Vinayakka Mission University, Salem. He has 8+ years of teaching experience. He is now a PhD research scholar at the department of Computer Science, Vels University, India, since 2010. Currently his specialization of research is under image processing and he has fine interest in web and mobile technologies.

Dr.Jothi Sophia Christopher is a Principal and Director of Nursing in CSI Jeyaraj Annapackiam College, Madurai, Tamilnadu, India. She has obtained Bachelor degree in Nursing, MSc Degree in Nursing for Child Health and PhD in Nursing for Child Health in the year 1992, 1995 and 2011 respectively. She has published more than 20 research papers under Nursing in peer reviewed International Journals and Conferences. Since 2010, she is one of Research Guide in Tamilnadu Dr.MGR Medical University, Chennai, India. She is a member of Indian Nursing Council (INC), All India Council of Christian Higher Education (AICCHE) and Tamilnadu Nursers Midwives (TNM). Her research areas of interest are computer in Nursing, Awareness of IT in Nursing students, Computer Image Applications.

Mr.Vinukumar Luckose has a Master Degree in Power Electronics and Drives Engineering from Anna University, India and has a Bachelor Degree in Electrical and Electronics Engineering from Manonmaniam Sundaranar University, Tirunelveli, India. He has 9 years of teaching experience in India and abroad. He also has worked in Johonson Lift Ltd. Company in India as a Senior Engineer. He is now currently working as a Senior Faculty of Engineering, Science and Technology (FEST) at Nilai University, Malaysia. His specialization is power electronics, mobile technologies, image processing and communication engineering.