ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)
Vol. 2, Issue 4 (Oct. - Dec. 2014)
International Journal of Advanced Research in Computer Science & Technology (IJARCST 2014)

Evaluation of Various Statistical Classifiers Performance in Detecting the Sugarcane Diseased Leaves

[I]Umapathy Eaganathan, [II]Dr.Jothi Sophia, [III]Vinukumar Luckose
[I]Research Scholar in CS, Vels University-India
[II]Director & Principal of Nursing, CSI JACON-Madurai, India
[III]Faculty in EEE, Nilai University-Malaysia

Abstract
Sugarcane is one of the major staple foods in the world. Everyday human being need to have sugar to enrich their food items. In all other variety of food items the sugarcane plays vital role in the agriculture. Sugarcane gives different byproducts and each product may vary in its taste. In ancient days farmers easily protected their fields to produce good quality of sugarcane with the help of traditional methods. Nowadays there are different diseases which affecting the plants in diverse areas. The causes of diseases may be from bacterial, viral and environmental factors. In order to control these kinds of diseases the farmers need to monitor, produce remedy is a challenging issue. Farmers get difficulties in detecting the diseases and also need to spend high amount to control it. For their easy and minimal way of preventing these diseases different researches carried out in image processing. This paper discussed about different statistical classifiers performance identified in sugarcane diseased leaves detection. Here sugarcane leaves of Normal, Brown rust and Yellow rust are collected and investigated using texture analysis based on the Color Co-occurrence Method (CCM). The YCbCr features are derived from RGB color model. The features are categorized for the selected leaves using K-Nearest Neighbor (KNN) classifier, Linear Discriminate Analysis (LDA) classifier and Parzen classifier. The experimental results based on KNN classifier supports 97.6% quality.

Keywords
Segmentation, Texture Analysis, Normal Leaf, Brown Rust, Yellow Rust, Feature Extraction, Classifier

I. Introduction

Digital image processing is the most important technique used in agriculture field. It has helped in the access to technical data in digital, multi-wavelength, services of computers in terms of speed of processing the data and the possibilities of big storage. The area of digital image processing belongs to processing of digital images by using digital computer. Digital images are form of visual information captured or transmitted using camera or other digital imaging system. In the digital images various kinds of noise exist in an image and a variety of noise reduction techniques are available to remove noises. Considering about the sugarcane one of the major countries is India and 70% of the population depends on agriculture. The cultivation of the sugarcane plants for optimum yield and quality produce is highly technical. It can be improved by the aid of technological support and its innovations [1]. The management of sugarcane plant requires close monitoring especially for the management of diseases that can affect production significantly and subsequently the post-harvest life. In most of the cases pests or diseases are seen on the leaves or stems of the plant. Therefore identification of plants leaves and finding out the diseases plays a key role in successful cultivation of plants. Disease management is a challenging task. Mostly diseases are seen on the leaves or stems of the plant. Here image processing plays important role. In this paper, the most classifiers approaches are given for automatic classification of plant leaf diseases.

II. Materials and Methods

1. Image Acquisition

The sugarcane leaves for this proposed experiment were taken from Madurai, Tamilnadu, India during the month of Feb 2014. Three different sugarcane diseased leaves were taken for this study such as normal, brown rust and yellow rust shown in Fig.1. The image samples acquired by using Canon EOS 550D color camera with zoom length of 22m of 18 M-Pixel resolution in 32-bits color resolution. The taken images transferred to PC which is of Intel Pentium 4 Processor 2.4GHz and then stored in the form of JPEG. Image processing was performed using MATLAB 7.14 software. The segmented images consist of arrays which is of 24-bits for each of the red (R), green (G) and blue (B) colors. In order to reduce effects associated with illumination changes the normalized colors considered based on the following.

$r = R / I$
$g = G / I$
$b = B / I$ where $I = (R+G+B) / 3$ is the average intensity.

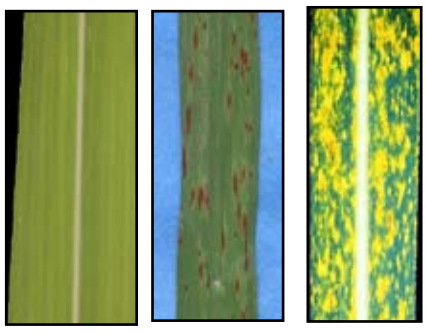


Fig. 1: Images of various sugarcane leaves

Sugarcane rust is caused by *Puccinia melanocephela*, an obligate parasitic fungus. Sugarcane rust is mainly a disease of the leaf. The earliest symptoms are small, elongated yellowish spots that are visible on both leaf surfaces. The spots increase in length, turn brown to orange-brown in color and develop a slight chlorotic halo. Lesions typically range from 2-10 mm in length but occasionally reach 30 mm. They are rarely more than 1-3 mm in width [2].

2. Methodology

Here the methodology of process based on the following step by step shown in Fig.2.

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

- Edge detection
- Filtering
- Extraction of YCbCr features
- Classification

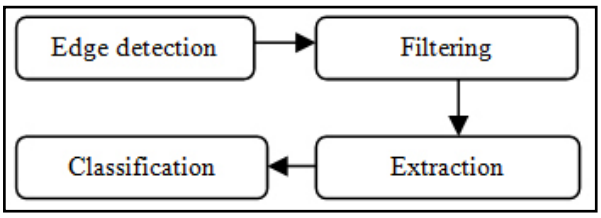


Fig. 2: Methodology

A. Edge Detection Using Sobel Operator

Segmentation is one of the steps in image processing and it refers to the process of partitioning a digital image into multiple regions. The main objective of image segmentation is to simplify the representation of an image into easier way to analyze shown in Fig.3. It is typically used to locate objects and boundaries in images. In this paper Edge detection type of segmentation is used. Edge detection is basic tool in image processing and in computer vision, mainly in the feature extraction.

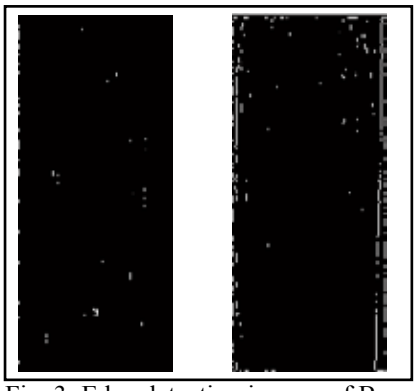


Fig. 3: Edge detection images of Brown rust and Yellow rust

An edge is the boundary between an object and the background, and indicates the boundary between overlapping objects. Edge detection refers to the process of identifying and locating sharp discontinuities in an image. Edges can be detected with the help of sobel operator is been used. The sobel operator is used in image processing, particularly within edge detection algorithms. The sobel operator is based on convolving the image with a small, separable and integer valued and it shown in Fig.4. If we denote A as the source image, and Gx and Gy are two images which at each point contain the horizontal and vertical derivative approximations, the computations are as follows [3].

Gx				Gy		
-1	0	1		1	2	1
-2	0	2		0	0	0
-1	0	1		-1	-2	-1

Fig. 4: Sobel Operators

In the second step, the sample images are pre-processed to improve the image data by enhancing for further processing. The processing includes color space conversion, image enhancement and image segmentation. The RGB leaves are converted into color space representation. The purpose of the color space is to facilitate the specification of colors in some standard accepted way that is converted into YCbCr color space representation. After the color transformations the Hue component is used for further analysis [4]. Median filter is a kind of higher order statistical filter and is nonlinear in nature and shown in Fig.5. The main idea behind the median filter is to find the median value by across the window, replacing each entry in the window with the median value of pixel and the pattern of neighbor's pixels is called the window [5].

123	125	126	130	140
122	**124**	**126**	**127**	135
118	**120**	**150**	**125**	134
119	**115**	**119**	**123**	133
111	116	110	120	130

Fig. 5: Median value calculation

It is just the center value after all the entries in the window are sorted numerically in ascending order.
115,119,120,123,124,125,126,127,150
Median value = 124

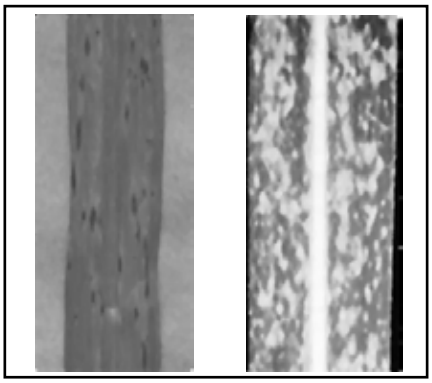


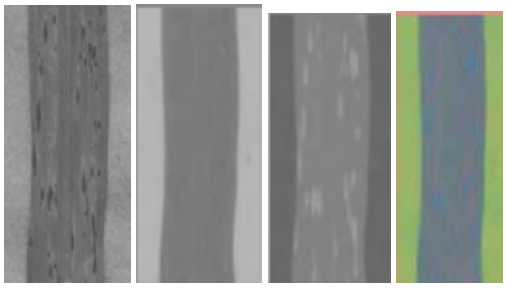
Fig. 6: After median filter the brown rust and yellow rust

B. YCBCR Color Space

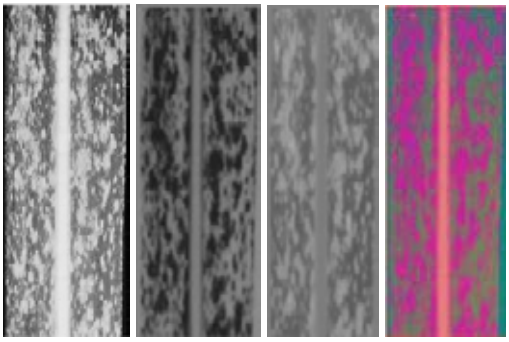
YCbCr is a family of color spaces used in video systems. Y is the luma component and Cb and Cr are the chroma components. Y is defined to have range from 16-235, Cb & Cr are defined to have a nominal range from 16-240. The basic equations to convert between RGB and YCbCr is given in the following.

$Y = 0.257R + 0.504G + 0.098B + 16$
$Cb = -0.148R - 0.291G + 0.439B + 128$
$Cr = 0.439R - 0.368G - 0.071B + 128$

In contrast to RGB, the YCbCr color space is luminance independent, where Cb and Cr components stand for difference of the blue and red with reference value respectively [6] shown in Fig.7.



Y Cb Cr YCbCr



Y Cb Cr YCbCr

Fig. 7: RGB to YCbCr color space of Brown rust and Yellow rust

3. Feature Extraction

After the color transformations the area of interest will be diseased part extracted. The significance features are extracted and those features can be used to determine the meaning of a given sample leaves. Normally, image features usually includes color, shape and texture features. Texture is a pattern of non-function spatial distribution of differing image intensities. Textured is termed as quantifying the spatial relationship between materials in an image. The extraction methods are used for extracting the relevant features from the input images [4].

The process of texture analysis usually produces kind of numeric descriptions of the texture, called texture features. The process of computing the texture feature is known as feature extraction. There are enormous number of texture analysis methods available and methods that used for the classification is also described. The color co-occurrence texture analysis method was developed through the use of Spatial Gray-level Dependence Matrices (SGDM's). The gray level co-occurrence methodology is a statistical way to described shape by statistically sampling the way certain grey-levels occur in relation to other grey-levels [7]. The GLCM function is MATLAB to create gray-level co-occurrence matrix; the number of gray levels is set to 8, and the symmetric value is set to 'true' and at the end offset is given a '0' value.

The texture features are

The angular moment (E) is used to measure the homogeneity of the image and is defined as shown below.

$$E = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} [P(i,j)]^2 \quad (1)$$

The covariance of the intensity co-occurrence matrix is defined

$$Cov = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} (i - I^2)(j - I^2)\, P(i,j) \quad (2)$$

The entropy feature (e) is a measure of the amount of order in an image, and is computed as

$$e = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P(i,j)\, lnP(i,j) \quad (3)$$

The information measures of correlation *inf2h* is defined as

$$inf2h = [\, 1 - e^{-2(Hxy-e)}]^{½}$$

Where $Hxy = - \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P_x(i)P_x(j) \ln [P_x(i)P_x(j)]$ (4)

Contrast (id) of an image can be measured by the inverse difference moment as

$$id = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} \frac{P(i,j)}{1+(i-j)^2} \quad (5)$$

In all the above equations the value of P (i, j) as follows

$$P(i,j) = \frac{P(i,j,1,0)}{\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P(i,j,1,0)} \quad (6)$$

4. Classification

Classification of data is the challenging task in image processing based on the classifiers. The goal of classification is to correctly predict the value of a designated discrete class variable, given a vector of predictors or attributes. Comparison of the test image is being done with the trained dataset in order to achieve objective of classifying whether the image is infected by the diseases or not. The YCbCr features are derived from RGB color model. The features are categorized for the selected leaves using K-Nearest Neighbor (KNN) classifier, Linear Discriminate Analysis (LDA) classifier and Parzen classifier.

A. K-NEAREST NEIGHBOR (KNN) CLASSIFIER

K-NN classifier is a non-parametric classifier, used to calculate the minimum distance between the given point and other points to determine the given point belongs to which class. The target is to compute the distance from the query sample to every training and selects the neighbor that is having minimum distance. Table.1 shown and it classifies a point by assigning it the label most frequently occurring among the k nearest samples. If k = 1, then the target is simply appointed to the gathering of its nearest neighbor.

Table 1. Classification results of KNN

Test Dataset / True Labels	Trained Dataset / Estimated Labels			
	Normal	Brown rust	Yellow rust	Total
Normal	50	0	0	50
Brown rust	0	47	3	50
Yellow rust	0	2	48	50
Total	50	49	51	150

From the above KNN classification it is observed that, total 5 images are misclassified that is 3 brown rust and 2 yellow rust. The remaining leaves are correctly classified.

B. Linear Discriminate Analysis (LDA) Classifier

LDA classifier is under parametric discriminate analysis statistical technique to sort objects into manually through groups based on a set of measurable object's feature. In this multiclass problem, a pattern x is assigned to the class for which the discriminate function is the largest. Table.2 a linear discriminate function assumes that every class has equal priors. Each test sample is classified into

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

the class with largest discriminate function values.

Table 2. Classification results of LDA

Test Dataset / True Labels	Trained Dataset / Estimated Labels			
	Normal	Brown rust	Yellow rust	Total
Normal	50	0	0	50
Brown rust	0	45	5	50
Yellow rust	0	2	48	50
Total	50	47	53	150

From the above LDA classification it is observed that, total 7 images are misclassified that is 5 brown rust and 2 yellow rust. The remaining leaves are correctly classified.

C. Parzen Classifier

The Parzen windows classification is a technique for non-parametric density estimation used for classification. In this approach classification algorithm does not require any training phase and shown in Table.3, for estimating densities first need to fix the size and shape of the region R. Let the region R is a n-dimensional hypercube with side length of h_n and its volume V_n. The generalized equation for estimating the density is

$$P_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{V_n}\varphi\left(\frac{x-x_i}{h_n}\right) \quad (7)$$

Table 3. Classification results of Parzen

Test Dataset / True Labels	Trained Dataset / Estimated Labels			
	Normal	Brown rust	Yellow rust	Total
Normal	50	0	0	50
Brown rust	0	44	6	50
Yellow rust	0	4	46	50
Total	50	48	52	150

From the above Parzen classification it is observed that, total 10 images are misclassified that is 6 brown rust and 4 yellow rust. The remaining leaves are correctly classified.

III. Experimental Results

In the classification both parametric and non-parametric classifiers are used and tested the features. The parametric classifier LDC and non-parametric classifier KNN and Parzen are used. Here the mean and co-variances for each classifier assumed as unknown. The data is randomly divided into 80% training and 20% test. The trained features which are obtained from sugarcane leaves using texture analysis are stored in a data set. Each and every class is tested differently which differentiates from normal and diseased leaves. The experiments for the proposed approach were conducted on a personal computer with an Intel i3 Processor and 2 GB RAM configured with Microsoft Windows 7 and MATLAB R2012a software with image processing toolbox. The Table.4 and Fig.8 shown the classification rate of KNN classifier is 96.7%, LDA is 95.3% and Parzen is 93.3%.

Table 4. Classification results

Classifier	Classification rate in %
KNN	96.7
LDA	95.3
Parzen	93.3

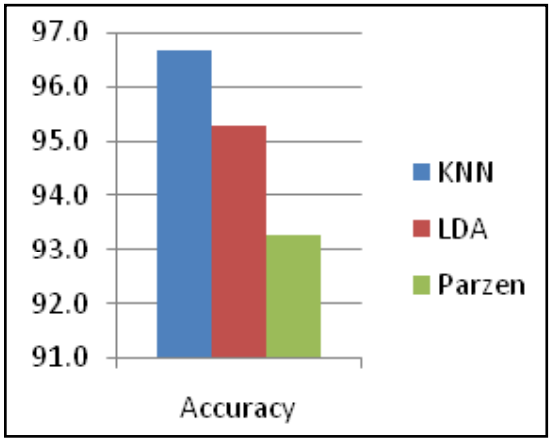


Fig 8. Comparison of KNN, LDA and Parzen classifier performance

IV. Conclusions

In this study, the classification based on different statistical classifiers involved and performed recognizing the leaves. In the image acquisition three sets of sugarcane leaves (Normal, Brown rust and Yellow rust) were collected and investigated from the sugarcane field, Madurai, Tamilnadu, India. Images considered in RGB then converted into YCbCr color space. The texture features have been extracted by using the GLCM algorithms. The features are tested on three statistical classifiers KNN, LDC and Parzen. Hence, it is conclude that KNN classifier provides better performance among the other classifiers. In future work, some other parametric and non-parametric classifiers may be used to classify the sugarcane leaves with more number of images.

V. AcknowledgEment

The authors say heartfelt thanks to the management and colleagues.

References

[1] *Jayamala K. Patil, Raj Kumar, "Advances in Image Processing for Detection of Plant Diseases", Journal of Advanced Bioinformatics Applications and Research, June 2011.*

[2] *http://www.pbcgov.com/coextension/agriculture/sugarcane/diseases/rust.htm.*

[3] *Manoj K.Vairalkar, S.U.Nimbhorkar, "Edge Detection of Images Using Sobel Operator", International Journal of Emerging Technology and Advanced Engineering, January 2012.*

[4] *Kiran R.Gavhale, Ujwalla Gawande, "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques", IOSR Journal of Computer Engineering, January 2014.*

[5] *Govindaraj.V, Sengottaiyan.G, "Survey of Image Denoising using Different Filters", International Journal*

of Science, Engineering and Technology Researh (IJSETR), Februrary 2013.

[6] E.Prathibha, A.Manjunath, R.Likitha, "RGB to YCbCr Color Conversion using VHDL approach", International Journal of Engineering Research and Development, June 2012.

[7] H.AI-Hiary, S.Bani-Ahmad, M.Reyalat, M.Barik and Z.ALRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, March 2011.

Authore Profiles

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 an 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.