ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 4, April 2018

Detection and Classification of soybean leaf diseases using K-means Clustering

[1] Sachin B. Jadhav, [2] Dr.Sanjay B. Patil
[1] Assistant professor, Department of E&TC. Bharati Vidyapeeth's College of Engineering, Kolhapur, India
Research scholar, V.T.U, Karnataka, India
[2] Professor & Principal, Department of E&TC Rajgad Dnyanpeeth Technical Campus Bhor, Pune, India

Abstract: **- Soybean Blight Brown Spot, Soybean powdery mildew and Downy Mildew are most common destructive foliar diseases of soybean and can cause significant yield loss. Timely application of fungicide, in the early stage of fungal infection, is important for effective control of the disease and is largely dependent upon the capability to quantitatively detection of the infection. The main purpose of this work is to identify and classify the soybean leaf disease based on the symptoms that are visible in leaf image. In this paper, Color-based segmentation method (K-means clustering) has been in corporate for separating the infected region from the leaf image. The infected stains are characterized by the features like color and textures. In the classification phase the color co-occurrence features, based on SGDM, are extracted and compared with the corresponding feature values stored in the feature library.**

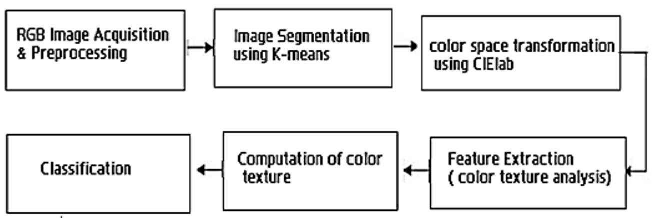
Keywords: **- K-means, Segmentation, Color feature, Color space, SGD..**

I. INTRODUCTION

Soybean rust is one of the most destructive foliar diseases of soybean. It produces copious amount of airborne spores that can infect large areas of soybeans and cause significant yield loss. Since 1994, the disease has been reported in countries such as Thailand, India, Southern China, and Japan. Yield losses were reported up to 80% in experimental trials in Asia (Hartman, Wang, &Tschanz, 1991). In the United States, this disease was first reported at the Louisiana State University AgCenter Research Farm in 2004 (Schneider, Hollier, &Hitam, 2005), but yield loss was not as high as those reported from other countries. An effective way to control soybean foliar diseases is by applying fungicides (Bravo et al.,2002; Heald, Thames, and Wiegand, 1972; Mueller et al., 2009;Foliar diseases can also cause changes in leaf color thus making it possible that the method of Schaberg et al. (2003) and Murakami et al.(2005) could be adapted to quantify different fungal foliar diseases. To test the method for disease assessment, black and white drawings from a manual of disease assessment keys showing foliar diseases with different disease severities (James, 1971) were digitized using a flatbed scanner, and then analyzed using Scion Image[4].Diagnosis the disease is a more complicated task to perform manually and consume much time. With the recent development in the field of image processing and pattern recognition techniques, it is possible to develop an automation system for disease detection and classification of crops based on the visual symptoms on leaf image.

II. RESEARCH METHOD

The Accuracy of the system based on the result of segmentation technique. Thresholding [1], [2], [11], [12], and 13] is a common Segmentation technique [3], [4], [5] that separate the region based on the threshold value. This method consumes much time and quality of the image is the main factor in the process. Clustering is a simple segmenting technique that forms the cluster with similar color pixel or similar texture. Easy deployment and plainness is one of the advantages of the clustering technique. In this work color based k-means clustering segmentation has been proposed to identify the infection based on the color of the leaf image. Based on the suggestions of field experts, the characteristic of the soybean foliar disease is dissimilar. The manual detection is based on the color and shape of the infected region. The color change in the infected region with respect to the background is measured as the one of the feature for classifying the disease. In this work the color features are computed using the mean, standard deviation of the infected regions, green pixels and background pixels with color change in the infected region compare with background in Red, Green, Blue color planes. Color, feature is considered as the main feature and some of the related features are computed for accurate classification. Based on the color feature the disease is predicted for verification. First, the images of various soybean foliar disease leaves are acquired using a digital camera. Then image-processing techniques are applied to the acquired images to extract useful features that are necessary for further analysis. To develop an automated system to detect and classify the soybean foliar diseases by performing the steps: (i) Image Acquisition and Preprocessing, (ii) Segment the infected regions, (iii) Color features Extraction (iv) Feature matching and Classification of Diseases. Figure 1 shows the basic procedure of the proposed vision-based detection algorithm in this paper.



.Figure 1: Block diagram of proposed approach

A. Proposed Approach:

1. Image Acquisition and Preprocessing:

The image is acquired from the field directly by using a digital camera with white background. Just place leaves on the plain white paper and try to take the image without shadow and over lighting. Shadow and unwanted noise are removed in the preliminary process and preceded to the next stage.

2. Color based segmentation using K-means Clustering technique:

In this work, the color based segmentation using k-means for segmenting the leaf regions to find the infected spots. In First stage, the basic image is modernizing the basic color bands to the Spectral-band representation using CIELAB conversation. In the next stage the regions are grouped into the set of clusters based on the major colors in the input images performed by k-means clustering algorithm.

3. Color space Transformations:

CIELab can be a nonlinear conversion of XYZ into L*a*b* coordinates. RGB color planes are accurate values which are linearly associated with CIEXYZ [13]. The color transformation of the RGB to L*a*b* is achieved by using the following relations. Diagrammatic representation shown in Fig (2) RGB – XYZ conversation: (White points (W) are used as the reference W (n) in CIElab) X=CxR (Same white point while converting RGB-XYZ) X=B CxrR (new white point, Bradford correlation) XYZ - L*a*b*conversation: (reference white Xn)

$$X_1=\frac{X}{X_n};\quad Y_1=\frac{Y}{Y_n};\quad Z_1=\frac{Z}{Z_n}$$

$$X1=\begin{cases} X_1^{\frac{1}{3}} & \text{If } X_1 > 0.008856 \\ 7.787\,X_1+16/116 & \text{else} \end{cases}$$

$$Y1=\begin{cases} Y_1^{\frac{1}{3}} & \text{If } Y_1 > 0.008856 \\ 7.787\,Y_1+16/116 & \text{else} \end{cases}$$

$$Z1=\begin{cases} Z_1^{\frac{1}{3}} & \text{If } Z_1 > 0.008856 \\ 7.787\,Z_1+16/116 & \text{else} \end{cases}$$

$$L^* = 116\,Y_1-16$$
$$a^* = 500\,(X_1-Y_1)$$
$$b^* = 200(Y_1-Z_1) \qquad \text{Eq1}$$

Where
C-represents the R, G, B.
L-luminosity,
A-chromaticity red-green axis
B-chromaticity blue-yellow axis
Using the makecform and applycform function to process the procedure of RGB to lab transformation in Matlab.

4. Segmentation using k-means clustering

K-means is one of the easiest unsupervised data clustering technique. The clustering of image for identifying the grouping or clusters in NBANDS or color bands with using Euclidean distance is the main part in image sorting as shown in figure 2. The k-means technique is intended to lessen the sum of square distance between each pixel and the cluster head. It groups the each and every pixel into the predefined number of clusters.

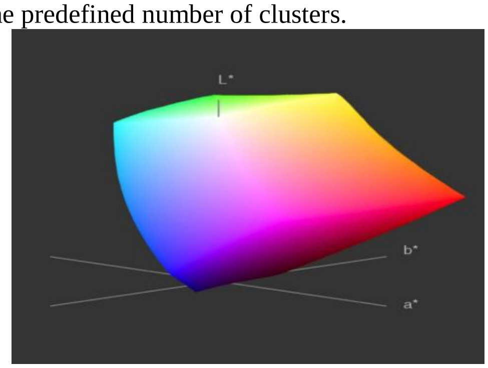


*Figure 2: L*a*b color representation*

Color information exists in the*, b* space, using the K-means to transform the pixel information into three clusters using the empty action function and create new cluster consists of a point far away from its centroid.

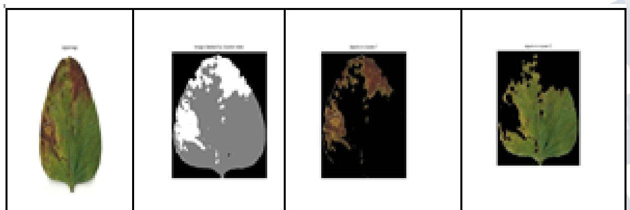
ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 4, April 2018

Cluster center is mentioned from the process and every pixel is labeled with its index number. Based on the label of pixel the image is segmented into three pixels. Image is segmented into the three clusters are shown in figure (3).

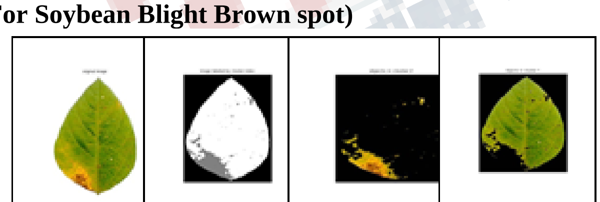
III. RESULT AND DISCUSSION

The proposed color based segmentation method and Otsu method are applied to the soybean diseased leaf image of samples in (3), (4), and (5) for Soybean Blight Brown spot. Based on the label of pixel the image is segmented into three pixels. Image is segmented into the three clusters are shown in figure (3), (4), and (5). The proposed color based segmentation method and Otsu method are applied to the soybean diseased leaf image of figure 5 for Downy Mildew, figure 6 for Soybean ceprospora and figure 7 for Soybean powdery Mildew, and segmented image are shown in figure 5, 6, and 7 respectively. Otsu method (fig 5a, fig6a, fig 7a) shows the infected pixels as non-infected pixels; the proposed method (fig 5b, 6b, 7b) shows the accurate detection.



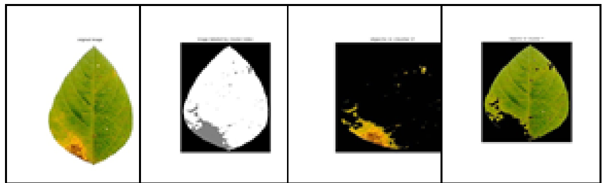
Sample image 1 (a)	Labeled diseased cluster (b)	Segmented Diseased region (c)	Segmented Healthy region (d)

Figure 3: (a) Sample Image & Segmentation of Three clusters with segmented region, (b) Otsu (c) Core Diseased regions and boundary of the diseased region. (d) Core Healthy Region, and boundary of the Healthy Region
(For Soybean Blight Brown spot)



Sample image 3	Labeled diseased cluster	Segmented Diseased region	Segmented Healthy region

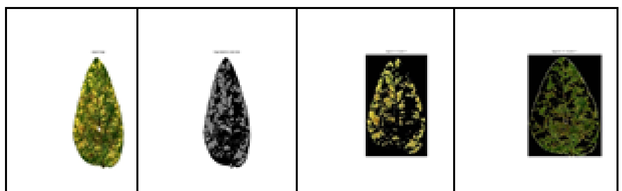
Figure 4: (a) Sample Image &Segmentation of Three clusters with segmented region, (b) Otsu (c) Core Diseased regions and boundary of the diseased region. (d) Core Healthy Region, and boundary of the Healthy Region.



Sample image 3	Labeled diseased cluster	Segmented Diseased region	Segmented Healthy region

Figure 5: (a) Sample Image &Segmentation of Three clusters with segmented region, (b) Otsu (c) Core Diseased regions and boundary of the diseased region. (d) Core Healthy Region, and boundary of the Healthy Region
(For Soybean Blight Brown spot)

The proposed color based segmentation method and Otsu method are applied to the soybean diseased leaf image of figure 6, for Downy Mildew, figure 7, for Soybean ceprospora and figure 8, for Soybean powdery Mildew, and segmented image are shown in figure 6, 7, and 8 respectively. Otsu method (fig 5a, fig6a, fig 7a) shows the infected pixels as non-infected pixels; the proposed method (fig 4b, 5b, 6b) shows the accurate detection.



Sample image 2 (a)	Labeled diseased cluster (b)	Segmented Diseased region (c)	Segmented Healthy region (d)

Figure 6: (a) Sample Image &Segmentation of Three clusters with segmented region, (b) Otsu (c) Core Diseased regions and boundary of the diseased region. (d) Core Healthy Region, and boundary of the Healthy Region
(For Soybean Downy Mildew)

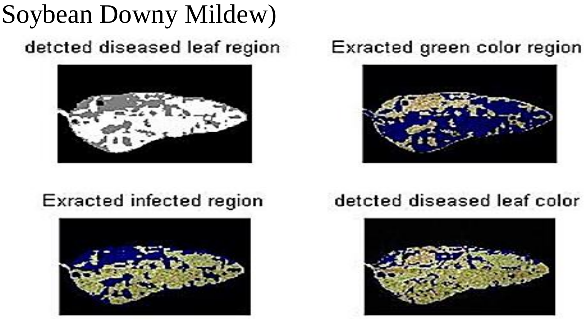


Figure 6 (a) Otsu, (b) K-means, (c) core infected region

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 4, April 2018

(d) Green pixel
(For Soybean Ceprospora)

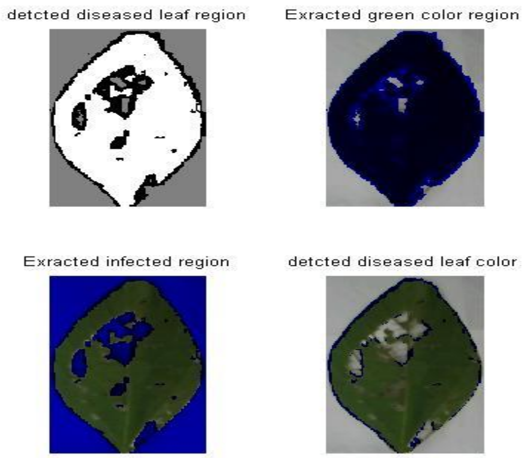


Figure 7 (a) Otsu, (b) K-means, (c) core infected region
(d) Green pixel (For Soybean Powdery mildew)

A. Color Feature extraction
Color changes in the plant leaf due to the fungal infection infected regions are considered as the major feature.

Color Feature
Color change in the infected region of the original color is a key feature for disease detection. The color of the infected region may vary for various diseases depend upon the significance of disease. Mean and standard deviation of the background region, infected regions, and green portions and change of color of the infected region with respect to background region in R, G, and B planes are measured as color features.

Mean
It provides average Color value in the image. It is calculated using following statics:

$$Mean(\mu) = \sum_{N}^{j=1} \frac{1}{N} P_{ij} \qquad (2)$$

Where, N is the total number of pixels in the image.

Deviation
It is called Standard Deviation. The standard deviation is the square root of the variance of the distribution. It is calculated using following statics:

$$Std\,Deviation(\sigma) = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - E_i)^2\right)} \qquad (3)$$

Classification of Disease:
The color of the infected region is the major feature for classifying the disease. Feature data set having the enough information about the color information about the disease. The color feature are taken as the constrains in IF then Rule. (IF condition color, then Color clusters) Based on the Rule disease are classified. Each disease has a specific value and based on the value the classification is done.

B.EXPERIMENTAL RESULTS AND DISCUSSION:
The proposed method applied on the sample data sets in Figure 4 (a), 5(a), 6(a), and segmented results are shown in Figure 4,5, 6 and 7. It shows the segmentation of blight brown, soybean Downey mildew ceprospora brown spot and Powdery Mildew. The given image is classified based on the color features. Table 1 shows color feature extraction using statics which shows variations within it.

TABLE I
Statistics of feature extraction for selected samples

Statistics	Blight Brown Spot	downy mildew	Ceprospora Brown Spot	Powdery mildew
contrast	2.2277e+04	2.7646e+04	1.1108e+04	3.3826e+04
correlation	-0.0012	0.0018	-0.0762	0.0012
Energy	1.3312e-05	9.2349e-06	5.0709e-05	1.0994e-05
Homogeity	0.0278	0.0241	0.0350	0.0232

IV. CONCLUSION

The experimental results indicate the proposed approach can recognize and classify the leaf diseases with a little computational effort. By this method, the plant diseases can be identified at the initial stage itself and the pest. The proposed method applied on the sample data sets Figure 4(a), 5(a), and 6(a) with segmented results are shown in Figure 4, 5, 6 and 7. It shows the segmentation of soybean blast and brown spot, soybean downy mildew, soybean powdery mildew and soybean ceprospora spot. With the minimal dataset of soybean blast and brown and downy Mildew, powdery mildew, and ceprospora spot disease. The given image is classified based on the color features. The reasons for misclassification are as follows: the Symptoms of the diseased plant leaves vary (at the beginning, tiny, dark brown to black spots, at later time, it has the phenomena of withered leaf, black or part leaf deletion), also the taken feature identification vectors need to further optimized. In order to improve disease identification rate at various stages, the training samples can be increased and shape feature and color feature along with the optimal features can be given as input condition of disease identification.

Future work:

In this work the basic technique is used for classification and feature matching as well as minimum data set is used to classify the disease. In future, large numbers of images are acquired directly from the field to train the system. Improvise the segmentation model and use k-means classifier system for modeling and classification of leaf disease.

REFERENCES

[1] DaeGwan Kim, Thomas F. Burks, Jianwei Qin, Duke M. Bulanon, "Classification of grapefruit peel diseases using color texture feature analysis", International Journal on Agriculture and Biological Engineering, Vol:2, No:3,September 2009.

[2] H.Al-Hiary, S. Bani-Ahmad, M.Reyalat, M.Braik and Z.AlRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications (0975-8887), Volume 17-No.1.March 2011.

[3] Mohammed El.Helly, Ahmed Rafea and Salwa-El-Gammal, "An Integrated image Processing System for Leaf Disease Detection and Diagnosis".

[4] Al-Bashish, D., M. Braik and S. Bani-Ahmad, 2011. "Detection and classification of leaf diseases using K-means-based segmentation and neural networks based classification". Inform. Technol. J., 10: 267-275. DOI:10.3923/itj.2011.267.275, January, 2011.

[5] Qin J, Burks T F, Kim M S, Chao K, Ritenour M A. "CitrusPCA-based image classification method". Sens. Instrum. FoodQual. Saf., 2008.

[6] F.Argenti, L.Alparone and G.Benelli , Fast Algorithms for Texture Analysis using Co-occurrence Matrices, IEEE proceedings,Vol.137, No.6, pp. 443-448, 1990.

[7] Sanjeev S Sannakki, Vijay S Rajpurohit, V B Nargund, et.al(2011), " Leaf Disease Grading by Machine Vision and FuzzyLogic", Int. J. Comp. Tech. Appl., Vol 2 (5), 1709-1716

[8] Prasad Babu, M. S. and SrinivasaRao, B. "Leaves recognition using back-propagation neural network - advice for pest and disease control on crops'. Technical report, Department of Computer Science &system Engineering, Andhra University, India.

[9] J. MacQueen. "Some methods for classification and analysis of multivariate observations". In L. M. LeCam and J. Neyman, editors,Proceedings of the Fifth Berkeley Symposium on MathematicaStatistics and Probability, volume 1, pages 281--297, Berkeley, CA.

[10] A. Camargo, J.S. Smith, An image-processing based algorithm toautomatically identify plant disease visual symptoms, Biosystems Engineering, Volume 102, Issue 1, January 2009, Pages 9-21, ISSN1537-5110,DOI:10.1016/j.biosystemseng.2008.09.030.

[11] Dr.Sanjay B. Patil, Dr. S.K.Bhodhe, "Betel Leaf Area Measurement Using Image Processing."International Journal on Computer Science and Engineering (IJCSE)ISSN: 0975-3397Vol. 3 No. 7 July 2011.pp 2656-2660.

[12] Dr.Sanjay B. Patil, Dr. S.K.Bhodhe, "leaf disease severity Measurement using image Processing."International Journal of Engineering and Technology Vol.3 (5), 2011,pp 297-301.

BIBLIOGRAPHY OF AUTHORS

	Mr.Sachin.B.Jadhav has completed his B.E in Electronics & Telecommunication from Shivaji university Kolhapur and M.E in Electronics from shivaji university Kolhapur. Currently he is working as an Assistant Professor in Department of Electronics &Telecommunication Engineering at Bharatividyapeeth College of Engineering, Kolhapur. He has keen interest in the field of Image processing.
	Prof. Dr. Sanjay Bapuso Patil has Awarded Ph.D in subject of Electronics in 2013 from NMIMS, Mumbai. His fields of study are Image Processing; Digital signal processing. He had worked as Professor at SCSCOE, Pune and lecturer in T.K.I.E.T Warananagar. Currently working as Principal at RDTC, SCSCOE, Bhor, Pune. He has written more than 27 technical papers in reputed international Journal and Conferences. He has filed two patents to Indian Patent Autorney. He is life member of ISTE and many International Journals.