pagepress

Parametric evaluation of segmentation techniques for paddy diseases analysis

Hemanthakumar R. Kappali,[1] Sadyojatha KM,[1] Prashanthi SK[2]

[1]Department of Electronics and Communication Engineering, Ballari Institute of Technology and Management, Ballari; Visvesvaraya Technological University, Belagavi, Karnataka; [2]Department of Plant Pathology, University of Agricultural Science, Dharwad, Karnataka, India

Abstract

In most paddy plant diseases, the leaf is the primary source of information for image-based disease identification and classification. Image segmentation is an important step in the plant disease analysis process. It is used to separate the normal part of the leaf from the disease-affected part of the leaf. In this paper diseases like Bacterial leaf blight (BLB), Brown spot (BS), and Leaf smut (LS) are segmented using existing, K-means clustering, the Otsu thresholding method. Color space-based segmentation is newly proposed for paddy disease analysis. The intelligence of segmentation techniques is evaluated using the Error Rate (ER) and Overlap Rate (OR) across the three paddy diseases namely, BLB, BS, and LS. The results were compared among the Otsu, K-means and color thresholding segmentation techniques. The results revealed that the color thresholding method using the Lab model emerged as a novel segmentation method for all three paddy diseases with an average ER and OR of [0.36, 0.95]. The proposed work is carried out in the Department of Electronics and Communication research center at Ballari Institute of Technology and Management, Ballari, Karnataka during the period from August 2022 to February 2023 with the expert suggestions of the plant pathologist, from the University of Agricultural Science, Dharwad, Karnataka.

Correspondence: Hemanthakumar R. Kappali, Ballari Institute of Technology and Management Ringgold standard institution, Department of Electronics and Communication Engineering Hosapete Road, Allipura, Ballari, Karnataka, India Ballari, Bellary, Karnataka 583104 India.
E-mail: hemanthakumar@bitm.edu.in

Key words: image segmentation; color space; paddy diseases.

Acknowledgments: the authors wish to express sincere gratitude to the Head, Dept of Electronics and Communication Engineering, Ballari Institute of Technology and Management, Bellary, Karnataka (India) for necessary arrangements and to the Plant Pathology Department, University of Agricultural University Dharwad, Karnataka (India) for continuous support.

Contributions: HRK, conducted this research, analyzed the results, and wrote a draft paper; SKM, designed the research and corrected the paper; PSK, assisted in domain-specific knowledge and field studies. All the authors approved the final version to be published.

Conflict of interest: the authors declare no potential conflict of interest.

Received: 17 April 2023.
Accepted: 15 July 2023.
Early view: 4 August 2023.

Journal of Agricultural Engineering 2023; LIV:1532
doi:10.4081/jae.2023.1532

Introduction

In the last three decades, paddy disease identification and classification research has been considerably in progress using traditional image processing and emerging machine learning algorithms. The main focus of this paper is to evaluate the image segmentation techniques used in the paddy disease analysis process. Segmentation is to select the region of interest that is, the disease-affected part from the normal part of the leaf. Segmenting a leaf from its complex background in the field is a challenging task even in a controlled environment because its performance is affected by variations in illumination, distance, and busy background.

The automation of the same using computer vision is still a challenging task. The complete process includes image acquisition, image preprocessing, image segmentation based on traditional image processing, feature extraction using pattern recognition, and a final step is disease identification and classification using a machine learning algorithm (Khirade Sachin & Patil, 2015).

Inconsistency in the segmentation task has an impact on the final result. This task is highly application-dependent and also a crucial part of the effective performance of the computer vision system (Barbedo 2016; Dhingra *et al*., 2018).

The segmentation techniques used in Paddy/Rice disease detection and classification are mainly thresholding based on the Local, Global, and Otsu methods (Kurniawati *et al.*, 2009; Orillo *et al*., 2014; Phadikar & Goswami, 2016; Guru *et al*., 2011; Zhang *et al*., 2018). The K means clustering segmentation (Gayathri & Neelamegam, 2019; Narmadha & Arulvadivu, 2017; Ramesh & Vydeki, 2020; Singh & Kaur, 2018). Active contour-based level set segmentation (Qiangqiang *et al*., 2015), Super pixel (Mai & Meng, 2016), Mean Shift (Devi & Muthukannan, 2014), and Edge based (Anthonys & Wickramarachchi, 2009) segmentation techniques are implemented. Color image segmentation based on different color spaces is a higher-level segmentation technique (Khattab *et al.*, 2014).

To identify/classify the different types of paddy disease support vector machine, k-nearest neighbor, and back-propagation, neural network-based machine learning algorithms are used with features like color, texture, shape, statistical, concurrence matrix, and lesion feature extracted from the segmented images as input vectors (Anthonys & Wickramarachchi, 2009; Guru *et al*., 2011; Kurniawati *et al*., 2009; Khirade Sachin & Patil, 2015; Mai & Meng, 2016; Narmadha & Arulvadivu, 2017; Phadikar & Goswami, 2016; Pinki *et al*., 2017).

The performance of the segmentation techniques used in vegetation disease analysis is evaluated by precision, recall, and dice parameters between reference and segmented image (Qiangqiang *et al*., 2015). Also, structural components, normalized cross-correlation, and peak signal noise ratio of the segmented image (Devi & Muthukannan, 2014).

The choice of the segmentation method is crucial for the effective performance of any computer vision system (Qiangqiang *et al*., 2015).

Visual assessment-based, properly segmented images gave better classification accuracy over poorly classified segmented images using a support vector machine classification algorithm (Kappali Hemanthakumar *et al*., 2023).

The literature review has revealed that image segmentation is a crucial step in the paddy disease analysis task. Though the segmentation task is application-dependent, features extracted from the segmented output solely depend on the intelligence of the segmentation task. Color models-based thresholding segmentation is rarely used in the agriculture domain and hence is implemented using RGB and lab color models. The performance of the segmentation has to be judged according to the context either by performance metrics or by visual intelligence which is not reflected in the literature review. In this paper performance of the segmentation techniques on various paddy diseases is evaluated using Overlap rate (OR) and Error rate (ER) to propose a novel segmentation method for paddy diseases identification and classification using machine learning.

Materials and Methods

The literature review has revealed that, in paddy disease identification and classification, Otsu thresholding and K-means segmentation methods are popularly used. In the case of the Otsu method, the intensities of the grey images are used as a local or global threshold value and the number of clusters or window size and seed point as an input in the case of K-means clustering. It uses Euclidean distance to identify and label the pixel to the nearest cluster in the image.

The color space-based segmentation is implemented for paddy diseases which includes the following steps: i) load image; ii) select the color space; iii) apply threshold; iv) output image; v) evaluate the performance metrics.

Load image

The proposed work considered various paddy disease-affected leaf images obtained from the freely available machine learning data repository over the internet (https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases). At the same time, the author is also collecting images from local farmers' fields. The 15 images of Bacterial leaf blight (BLB), Brown spot (BS), Leaf smut (LS), and paddy diseases are considered as input and each image is considered for further process in segmentation.

Select the color space

Most of the researchers converted color images into binary/grey images, though the color images give more information in various channels. The color images are represented using various color spaces like RGB, Lab, Hue Saturation Brightness, and YCbCr. The present paper considered only RGB and Lab color spaces. The RGB color image consists of three primary components red, green, and blue channels. The CIE Lab (CIE L*a*b*) most of the time termed as Lab color space consists of L-Lightness from black to white, a* from green to red, and b* from blue to yellow. The color wavelength can be defined by:

$$X = C\lambda \tag{1}$$

$$x = \frac{X}{X,Y,Z'} \tag{2}$$

$$y = \frac{Y}{X,Y,Z'} \tag{3}$$

where x is the color to get and C is the constant multiplied by length which is the speed of light. (X, Y) are considered color characteristics. X is luminance parameters, (X Y) coordinate chromatin and (X, Y, Z) triple emitting values which are the basis for all the colors. For example, the red and green colors found large values of (x, y) and sequentially and small values of (X, Y).

The RGB color space components are represented in terms of Equation 3 as follows:

$$\begin{aligned} R &= 3.24054X - 1.537138Y - 0.498531Z \\ G &= -0.96926X \; 1.876010Y \; 0.04155Z \\ B &= 0.055643X \; 0.204025Y \; 1.057225Z \end{aligned} \tag{4}$$

Similarly, the Lab color space components are represented as follows,

$$\begin{aligned} L^* &= 116 f\left(\frac{Y}{Yn}\right) - 16 \\ a^* &= 500\left\{f\left(\frac{X}{Xn}\right) - f\left(\frac{Y}{Yn}\right)\right\} \\ b^* &= 200\left\{f\left(\frac{Y}{Yn}\right) - f\left(\frac{Z}{Zn}\right)\right\} \end{aligned} \tag{5}$$

Apply threshold

Once the color space is selected the images are segmented using the color thresholding method applied for the different channels like R, G, B and L a* and b*. The terminating condition for segmentation is the visual interpretation of the author by comparing the various threshold values. The satisfactory interpretations are used as terminating criteria for segmentation and to extract the masked image and segmented image as an output.

Output image

There are two images taken as an output, masked images are used to focus area of interest which is easier to analyze by neglecting the irrelevant information and segmented images are obtained by input image and masked images provide the disease symptoms as the output. These two images are stored in the computer memory for further performance evaluation steps.

Evaluate the performance metrics

The performance of the segmentation has to be evaluated either by performance metrics or by visual impressions (Kappali Hemanthakumar *et al.*, 2023) because the classification accuracy significantly depends on the segmentation performance. In this paper performance of the segmentation techniques is evaluated by ER and OR between the input image and the masked image. ER measures the proportion of pixels in the segmented image that are incorrectly classified compared to the ground truth input image given by:

$$ER = (FP + FN)/(TP + TN + FP + FN) \tag{6}$$

The OR measures the degree of agreement between the segmented image and the ground truth image is given by:

$$OR = (TP)/(TP + FP + FN) \tag{7}$$

where, true positives (TP) - number of pixels classified as foreground in the segmented image and ground truth input image.

True negatives (TN) - number of pixels classified as background in the segmented image and ground truth input image.

False positives (FP) - number of pixels classified as foreground in the segmented image but are background in the ground truth input image.

False negatives (FN) - number of pixels classified as background in the segmented image but are foreground in the ground truth input image.

The obtained results used for paddy disease analysis are discussed in the following section.

Results and Discussion

The results of BLB, LF, and BS paddy diseases using the existing segmentation methods like simple Otsu thresholding and K-means segmentation are shown in Figures 1 and 2.

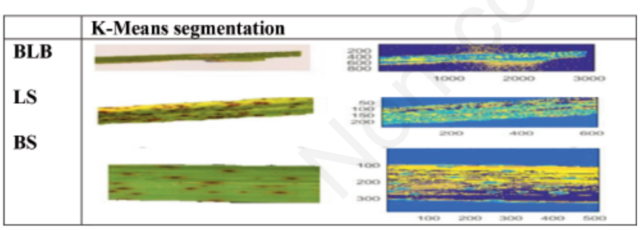

Figure 1. Segmentation results of Otsu thresholding. BLB, Bacterial leaf blight; LS, Leaf smut; BS, Brown spot.

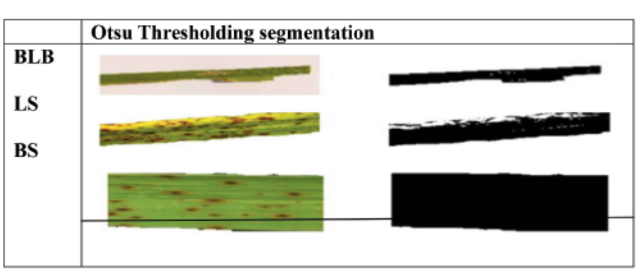

Figure 2. Segmentation results of K-means clustering method. BLB, Bacterial leaf blight; LS, Leaf smut; BS, Brown spot.

The complexity of the Otsu method is when the pixel distribution in an image is not bimodal and there are multiple pixel values acting as threshold values to segment the region of interest, whereas the uncertainty of the K-means method is to automate the identification of initial cluster centers, and also the number of clusters, the resulting in both the methods are failed to qualify the paddy disease segmentation phase unless they are equipped with higher level domain-specific key factors.

The segmentation using RGB and Lab color space results is newly implemented on paddy diseases with evaluating criteria to verify the contribution of the pixel details embedded in various color channels in segmenting the disease part from a healthy part of the leaf the results are as shown in Figure 2. The paddy disease images are segmented better using Lab color spaces over the RGB color models only the limitation is human intervention in choosing the threshold values for all the color channels with visual interpretation capability of the author.

The key factor in image-based disease analysis is the intelligence of the segmentation method. This can be evaluated by either visual impressions which we have discussed according to the results shown in Figure 3, or by appropriate segmentation performance metrics.

In this paper, the performances of the segmentation techniques are evaluated using OR and ER defined using equations 6 and 7. The higher value of OR and lower value of ER give better segmentation results.

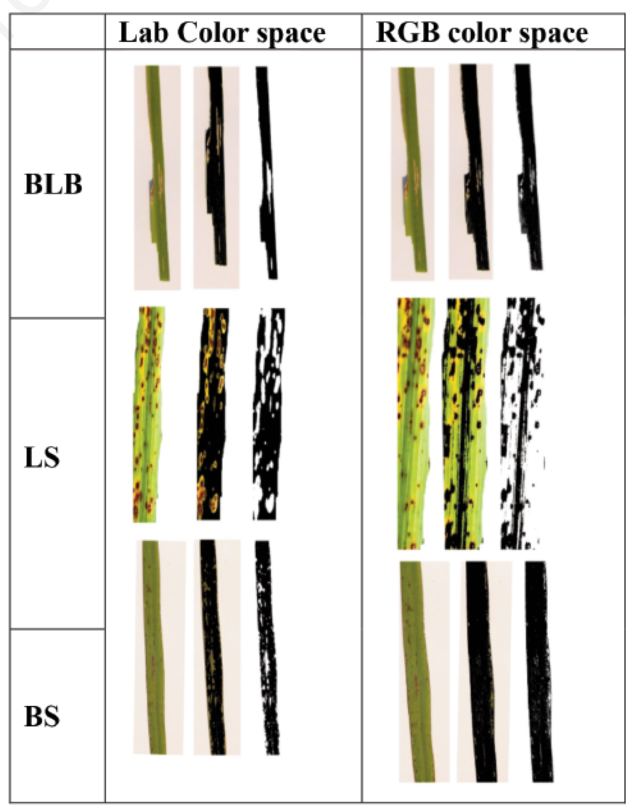

Figure 3. Segmentation results of RGB and Lab model color thresholding. BLB, Bacterial leaf blight; LS, Leaf smut; BS, Brown spot.

Table 1. Average performance metrics (error rate, overlap rate).

	Color threshold segmentation		Segmentation using	
	Lab color model	RGB color model	Otsu	K-Means
BLB	(0.56, 0.94)	(0.54, 0.94)	(0.65, 1)	(1, 0.65)
LS	(0.21, 0.94)	(0.30, 0.92)	(0.52, 1)	(1, 0.52)
BS	(0.29, 0.96)	(0.37, 0.93)	(0.46, 1)	(1, 0.62)

BLB, Bacterial leaf blight; LS, Leaf smut; BS, Brown spot.

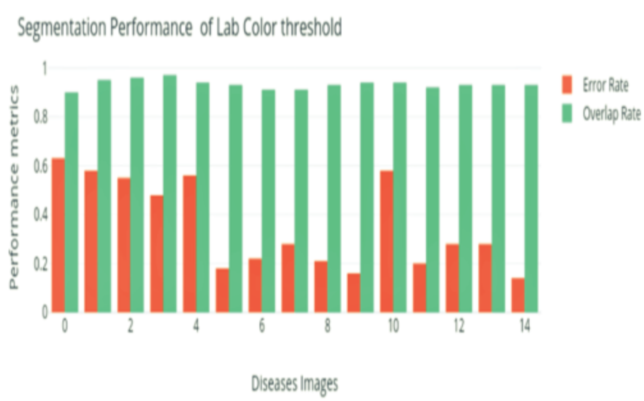

Figure 4. Segmentation performance of color threshold using Lab color model.

Figure 5. Overlap rate and error rate of segmentation methods.

These metrics ER and OR are calculated for all the segmented images of three paddy diseases and the average values for each disease are tabulated in Table 1.

From the table performance of RGB and Lab model-based color thresholding methods are performing better in comparison with the other K-means and Otsu segmentation methods. The Lab color model-based segmentation technique in all three diseases has an average error rate and overlap rate of [0.36, 0.95] and emerged as a novel segmentation method for paddy vegetation as shown in Table 1.

The ER and OR of the qualified color thresholding segmentation technique using Lab are shown in Figure 4.

The correlation between ER and OR is highly acceptable for all three diseases in the case of the novel color space-based threshold segmentation method, whereas the K-means and Otsu methods failed to qualify the performance evaluation as well as there is no correlation with the performance metrics as shown in Figure 5 showing the 15 BLB disease images along the x-axis and performance metrics values along y axis.

Conclusions

In the last three decades, research has been on vegetation disease analysis for the identification and classification of plant diseases. The diseases are characterized by their symptoms exhibited in visible parts of the plants. In this context image segmentation is the crucial step, which separates the diseased part and the healthy part of the leaf. In this paper, the novelty of the existing segmentation methods like Otsu, K-Means and color thresholding using RGB and Lab color space are evaluated using visual inspection and performance metrics using OR and ER on BLB, BS, and LS diseases of paddy vegetation.

The same method can be applied for real-time disease images collected from paddy fields and also for all the paddy diseases available. These original and segmented images can be applied for machine learning-based classification algorithms to evaluate the contribution of segmentation over the performance of the disease classification/identification algorithm in the future. The hybrid method of image processing and machine learning-based disease identification/classification techniques is the only means for vegetation health analysis to improve the yield and quality of the crop.

References

Anthonys G., Wickramarachchi N. 2009. An Image Recognition System for Crop Disease Identification of Paddy Fields in Sri Lanka. 2009 Int. Conf. Ind. Info. Syst. (ICIIS). 403-7.

Barbedo J.G.A. 2016. A Review on the Main Challenges in Automatic Plant Disease Identification Based on Visible Range Images. Biosyst. Eng. 144:52-60.

Devi D.A., Muthukannan K. 2014. Analysis of Segmentation Scheme for Diseased Rice Leaves. 2014 IEEE Int. Conf. Adv. Commun. Control Comput. Technol. 1374-78.

Dhingra G., Kumar V., Joshi H.D. 2018. Study of Digital Image Processing Techniques for Leaf Disease Detection and Classification. Multimed. Tools Appl. 77:19951-0.

Gayathri Devi T., Neelamegam P. 2019. Image Processing Based Rice Plant Leaves Diseases in Thanjavur, Tamilnadu. Cluster Comput. 22:13415-28.

Guru D.S., Mallikarjuna P.B., Manjunath S. 2011. Segmentation and Classification of Tobacco Seedling Diseases. COMPUTE '11: Proceedings of the Fourth Annual ACM Bangalore Conference, 1-5. New York, NY, USA: Association for Computing Machinery.

Kappali Hemanthakumar R., Sadyojatha K.M., Prashanthi S.K. 2023. Computer Vision and Machine Learning in Paddy

Diseases Identification and Classification: A Review. Indian J. Agric. Res.
Khattab D., Ebied H.M., Hussein A.S., Tolba M.F. 2014. Color Image Segmentation Based on Different Color Space Models Using Automatic GrabCut. Sci. World J. 2014:e126025.
Khirade Sachin D., Patil A.B. 2015. Plant Disease Detection Using Image Processing. 2015 Int. Conf. Comput. Commun. Control Autom. 768-71.
Kurniawati N.N., Sheikh Abdullah S.N.H., Abdullah S., Abdullah S. 2009. Investigation on Image Processing Techniques for Diagnosing Paddy Diseases. Int. Conf. Soft Comput. Pattern Recognit. 272-77.
Mai X., Meng M.Q.-H. 2016. Automatic Lesion Segmentation from Rice Leaf Blast Field Images Based on Random Forest. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR), 255-59.
Narmadha R.P., Arulvadivu G. 2017. Detection and Measurement of Paddy Leaf Disease Symptoms Using Image Processing. Int. Conf. Comput. Commun. Inform. (ICCCI), 1-4.
Orillo J.W., Dela Cruz J., Agapito L., Satimbre P.J., Valenzuela I. 2014. Identification of Diseases in Rice Plant (Oryza Sativa) Using Back Propagation Artificial Neural Network. Int. Conf. Humanoid Nanotechnol. Inform. Technol. Commun. Control Environ. Manage. (HNICEM), 1-6.
Phadikar S., Goswami J. 2016. Vegetation Indices Based Segmentation for Automatic Classification of Brown Spot and Blast Diseases of Rice. 3rd Int. Conf. Recent Adv. Inform. Technol. (RAIT) 284-89.
Pinki F.T., Khatun N., Mohidul Islam S.M. 2017. Content Based Paddy Leaf Disease Recognition and Remedy Prediction Using Support Vector Machine. 20th Int. Conf. Comput. Inform. Technol. (ICCIT) 1-5.
Qiangqiang Z., Zhicheng W., Weidong Z., Yufei C. 2015. Contour-Based Plant Leaf Image Segmentation Using Visual Saliency. Image and Graphics, edited by Yu-Jin Zhang, 48-59. Lecture Notes in Computer Science. Cham: Springer International Publishing.
Ramesh S., Vydeki D. 2020. Rice Disease Detection and Classification Using Deep Neural Network Algorithm. Microelectron. Telecommun. Eng. 555-66. Lecture Notes in Networks and Systems. Singapore: Springer.
Singh J., Kaur H. 2018. A Review on: Various Techniques of Plant Leaf Disease Detection. 2nd Int. Conf. Inven. Syst. Control (ICISC), 232-38.
Zhang J., Yan L., Hou J. 2018. Recognition of Rice Leaf Diseases Based on Salient Characteristics. 13th World Congress on Intelligent Control and Automation (WCICA), 801-6.