Plant Diseases Detection Based on Color Features and Kapur's method

EVa TUBA
University of Belgrade
Faculty of Mathematics
Studentski trg 16, Belgrade
SERBIA
etuba@acm.org

RAKA JOVANOVIC
Hamad bin Khalifa University
Qatar Env. and Energy Research Inst.
PO Box 5825, Doha
QATAR
rjovanovic@hbku.edu.qa

MILAN TUBA
John Naisbitt University
Graduate School of Computer Sci.
Bulevar umetnosti 29, Belgrade
SERBIA
tuba@ieee.org

Abstract: Plant diseases have become an important issue because they cause important reduction in each quality and amount of agricultural products. Automatic detection of plant diseases is an important analysis topic because it could significantly help in observation giant fields, and enable automatic detection the symptoms of diseases as soon as they appear on the plant leaves. In this paper an algorithm for plant disease detection using different color models is proposed and tested. Plant leaf images were first transformed into RGB, YCbCr, HSI or CIELAB color model. Noise in transformed image was reduced by applying median filter. At the end, disease spots were detected by using Kapur's thresholding method. Based on the experimental results, HSI color model is the most suitable for automatic plant disease detection, while RGB is practically unusable.

Key–Words: CIELAB, HSI, YCbCr, plant leaf disease detection, image thresholding, Kapur's method

1 Introduction

Digital images became part of everyday life. Digital image processing and analysis represents a rather popular scientific research topic. It was used in many different fields such as astronomy [1], biology [2], medicine [3], [4], etc. Increasing use of digital images resulted with numerous algorithms for image processing that are used in various applications such as multilevel thresholding [5], [6], JPEG compression [7], handwritten digit recognition [8], leaf recognition [9], lip detection [10], etc.

One of the areas where digital image analysis brought a great benefits is agriculture. Plant diseases have become an important issue because they cause important reduction in each quality and amount of agricultural products. Large amount of money is spent every year all around the world for efficient plant disease detection while the financial lost in case of late disease detection is 3, 4 times larger. One of the methods of plant disease detection and identification is an eye observation by the specialists is that. However, this needs continuous observation of specialists which could be time and financially expensive in giant farms. Further, in some developing countries, farmers could have to be compelled to go long distances to contact specialists which also is not practical and cheap.

Automatic detection of plant diseases is an important analysis topic because it could significantly help in observation giant fields, and enable automatic detection the symptoms of diseases as soon as they appear on the plant leaves. Therefore, searching for quick, automatic, more cost-effective and correct technique to find disease is of the great significance. Machine learning primarily detection and later recognition of plant diseases will offer clues to spot and treat the diseases in its early stages. Also, as it was explained earlier visually distinguishing plant diseases can be inefficient and troublesome since it needs the experience of trained plant scientist. Some researchers have used image processing techniques for quick and correct detection of plant diseases. The accuracy of result depends on technique used for disease spot detection. The biggest obstacle in disease spot detection is noise, that is introduced by camera flash, modification in illumination, droning background and presence of vein within the plant leaf. Thus a technique that wipes out the noise and provides higher disease spot segmentation is required.

Veins color is the same as plant leaf color solely intensity differs. On the other hand, disease spot color is completely different from plant leaf color. Thus if image is first transformed to some color model that

has intensity component threshold can be applied on it and this is one of the main techniques that are applied in plant diseases detection applications.

In this paper adjusted median filter is proposed for noise reduction. One of the well known threshold technique, Kapur's method is employed to convert filtered image into binary image Mistreatment on top of techniques disease spot is detected in totally different plants, during which largely veins are parallel and fewer visible. The proposed algorithm was tested by using different color models such as RGB, HSI and CIALAB.

The rest of the paper is organized in four sections. In Section 2 literature review of algorithms for plant diseases was presented. Our proposed algorithm is defined and described in Section 3 Experimental results are showed in Section 4. Conclusion of the paper is given in the last section.

2   Literature review

Numerous algorithms and methods were proposed for plant disease detection. Many of them tries to detect different diseases in various plant species. Some number of the papers deals detection and recognition of concrete diseases of one plant while the other papers classify numerous diseases. Extensive literature review can be found in [11] and [12].

Color features are very important for plant disease detection. In some papers image was first converted from RGB color model into HSI model and then threshold was applied on H element for segmenting infected areas in plant leaves. Cui et al. in [13] developed a quick manual threshold-setting technique supported HSI color model to section the disease spot. Kai et al. in [14] convert RGB image into YCbCr color aperture to discover the disease spot. In this paper the impact of YC_bC_r, HSI and CIELAB color area within the method of disease spot detection are compared. Experiments were carried upon various families plant leaves with each noise free (white) and clamorous background to induce the strategy that is individual of background signal and plant sort.

Kurniawati et al [15] introduced a technique for detection and classification of paddy disease. Otsu threshold method was employed for disease spot detection and unessential spots are removed by median filter. In [16] the tactics of image pre-processing for detection the diseases was studied.

In [17] an algorithm for plant diseases detection and classification based on K-mean segmentation and

neural network classification was proposed. The proposed algorithm had four phases where in the first one plant leaf image was transformed into device-independent color space, in the second K-mean algorithm was used to cluster the leaf image into four clusters. In the third phase texture features were calculated for segmented parts and at the end neural network was used for diseases classification. Accuracy of 93% of classification was reported.

Another four steps method for plant diseases detection and classification was proposed in [18]. The first step is creating transformation structure for RGB leaf image. The next step is applying mask to the green channel where some of them were removed by using specific threshold value.The third step is segmentation. At the end important segments were selected and the texture features were computed for them. Based on these features plant diseases detection and classification were done.

In [19] an algorithm for huanglongbing or citrus greening disease detection based on high resolution satellite images was proposed. Different classifiers were used for disease classification. Among the tested classification algorithms, support vector machine with RBF kernel obtained the best results. Linear support vector machine, linear discriminant analysis and quadratic discriminant analysis obtained also promising results.

A method for automatic disease diagnosis especially the identification and diagnosis of grape downy mildew and grape powdery was proposed in [20]. The method was based on K means clustering which was used to implement unsupervised segmentation of the disease images. Later fifty different shape, color and texture features were extracted from the segmented image. Support vector machine classifier was used for grape diseases detection. The testing recognition accuracy was 90% and 93.33% for grape downy mildew and grape powdery, respectively.

3   Our proposed algorithm

In this paper we propose a method for plant diseases detection based on the leaf images. The proposed method contains three main steps. The first one is image transformation into different color models, the second is noise reduction and the last one is image segmentation where disease spots are detected.

It was assumed that images of the plant leafs are available. Images can be obtained by putting cameras in the fields or satellite images can be used. Al-

l the images used in this paper are in JPEG format. These images are in RGB color model. The proposed method will be tested when the image is transformed into YC_bC_r, HIS and CIELAB color models. The color converted images are gone through median filter to get rid of spare spots and reduce the noise. In the last step Kapur's method is applied on different color components depending on the used color model. A element of CIELAB color area was used, H element of HSI color area and C_r element of YC_bC_r color area are employed to discover the disease spot. Diseases detection is obtained by all the three strategies and compared to induce the most effective methodology for disease spot detection.

In plants, leaf vein is totally different in intensity and disease spot is different in color compared to plant leaf. Therefore if Kapur's method is applied on grayscale image, vein will be detected in binary image with the disease spot. However the region of interest is simply disease spots, not vein. For minimize the effect of presence of vein, RGB color model is not suitable for segmentation. Thresholding method are often applied on color element to discover disease spot accurately.

The first color model that will be used is YC_bC_r. This color model is wide employed in digital video. In YC_bC_r color model, Y indicates brightness level element and C_b, C_r indicate color elements. C_b represents distinction between the blue elements while C_r is that the distinction between the red elements. This color model was often used in literature for different applications such as face detection [21], blood cell segmentation [22], etc. Using following formulas RGB image can be transformed into YC_bC_r color model.

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$
$$C_b = -0.168 * R - 0.331 * G + 0.500 * B$$
$$C_r = 0.500 * R - 0.418 * G - 0.081 * B \quad (1)$$

Color model HSI is device dependent color model and primarily based upon human color perception. In this color model H indicates hue, that describes a pure color and is mostly associated with the wavelength of the light. Component S indicates saturation, that measures the colorfulness in HSI color model while I indicates intensity, that shows the amplitude of the light. The first step in conversion is to represent the RGB components in the range $[0,1]$ by dividing each pixel value for each component by 255. Image further from

RGB color model can be transformed into HSI model by the following equations:

$$H = \begin{cases} \theta & \text{if } B \leq G, \\ 360 - \theta & \text{if } B > G, \end{cases}$$

where θ is:

$$\theta = cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}} \right\}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R,G,B)] \quad (2)$$

$$I = \frac{1}{3}(R+G+B) \quad (3)$$

CIELAB system is device independent color model that is outlined by the CIE to classify color consistent with the human vision. Within the conversion method of an image from RGB color element to CIELAB color element, first RGB image is transform into CIEXYZ by the following equation:

$$X = 0.4124 * R + 0.3576 * G + 0.1805 * B$$
$$Y = -0.2126 * R + 0.7152 * G + 0.7220 * B$$
$$Z = 0.0193 * R + 0.1192 * G + 0.9505 * B \quad (4)$$

Brightness and color data of research lab color model is independent of every different. In CIELAB color model, L describes color brightness, A describes the color starting from green to red while B describes the color starting from blue to yellow. Conversion formula for research laboratory color model is defined by the following equation:

$$L = 116 * f\left(\frac{Y}{Y_n}\right) - 16$$
$$A = 500 * \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right)$$
$$B = 200 * \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \quad (5)$$

where function f is defined as:

$$f(t) = \begin{cases} \sqrt[3]{t} & \text{if } t > \beta^3, \\ \frac{t}{3\beta^2} + \frac{4}{29} & \text{otherwise} \end{cases}$$

where $\beta = \frac{6}{29}$. Constants X_n, Y_n and Z_n are equal to 95.047, 100.000 and 108.883, respectably.

3.1 Image Smoothing

The next step in our proposed algorithm is image smoothing. During image assortment, some noise is also introduced due to camera flash. This noise might have an effect on the detection of disease. To remove unneeded spots, image smoothing technique is required. In this paper adjusted median filter is employed for this purpose.

Median filter can be a higher order statistics filter. Median filter is nonlinear in nature and replaces the value of the central pixel by the median of the gray levels within the image area surrounded by the filter.

So in order to perform median filtering, first, window is rapt and the pixels enclosed by the window area are sorted. When median is computed, it is assigned to the center pixel. The number of the pixels in $K \times K$ window is odd. In this paper size of the window was empirically determined to be 5×5. Smaller window size was not able to remove some larger noise spots, while the larger window size cause damaging clear parts of the image.

3.2 Disease Detection by Kapur's Method

After image smoothing, a method to detect and isolate the disease spot is required. It is necessary to find a threshold value that will differentiate the disease spots from plant leaf. If the histogram has sharp and deep depression between two peaks, bottom of the depression may be chosen as threshold. However the problem is when depression does not exists of it is flat. In such case this method can not be accustomed to separate objects from background. One of the most used method for thresholding is Kapur's method that is based on the entropy. This method maximize the amount of information between the two parts of a intensity histogram that are separated by concrete threshold value or better to say maximize the entropy measure of the part of the histogram in order to each part has a more centralized distribution.

Kapur's method is defined as follows. Assume that image I has L gray levels. If the image has N pixels and with $h(i)$ is represented the number of pixels with i level intensity, then $P_i = \frac{h(i)}{N}$ represent the probability of intensity. Kapur's method determines optimal threshold value by maximizing the next objective function:

$$f(t) = H_0 + H_1 \qquad (6)$$

where

$$H_0 = -\sum_{i=0}^{t-1} \frac{P_i}{\omega_0} \ln \frac{P_i}{\omega_0}, \quad \omega_0 = \sum_{i=0}^{t-1} P_i$$

$$H_1 = -\sum_{i=t}^{L-1} \frac{P_i}{\omega_1} \ln \frac{P_i}{\omega_1}, \quad \omega_1 = \sum_{i=t}^{L-1} P_i \qquad (7)$$

In this paper, Kapur's method was used for diseases spot detection. Threshold value was searched for different components. In case of RGB image, threshold value was searched for gray level image obtained by averaging all pixels components. If HSI model was used, threshold value for H component was searched while in the case of the YC_bC_r model, Kapur's method was applied to C_r component. When CIELAB color model is used, Kapur's method was used to find optimal threshold value for component A.

4 Experimental results

In this paper, first was experimented with images in RGB model. Disease spots are detected by applying Kapur's threshold method on gray scale image that is result of averaging pixels components. In Fig. 1 are shown experimental results for one example. It is obvious that plant diseases were not recognized successfully. As it was assumed, RGB color model is not suitable for detection disease spots in plant leafs.

To improve detection further experiments were done. In the second technique RGB image is initial transformed into YC_bC_r color model by using color transform formula previously described. Then median filter is employed for image smoothing. Disease spots are detected by applying Kapur's threshold on C_r component of filtered YC_bC_r color image. Experimental results are shown in Fig. 2. Diseases were detected more successful in this case comparing to the detection in RGB model. When YC_bC_r model was used detections of the spots was larger then actual diseases. This means that some parts of the leafs without diseases were recognized as infected. False disease detection is not good characteristic of the application.

In next experiment RGB image was transformed into HSI color model and disease spots were detected like before by applying Kapur's threshold but on H component of filtered HSI color area. Detections of plant diseases by our proposed method by HSI color model are shown in Fig. 3. As it can be seen, diseases were mostly discovered correctly. Comparing to the

Figure 1: Plant diseases detection in RGB image

Figure 2: Plant diseases detection using YC_bC_r color model

previous examples this recognition was the most precise. Some not infected parts are still marked as a disease, which is visible in the third test example. The first two examples were almost completely correctly marked.

The last color model that need to be tested is CIALAB. Once more the proposed algorithm was used but with images in CIELAB color model. Dis-

ease spots are segmental by applying Kapur's threshold on component A of filtered laboratory color area and the results are presented in Fig. 4. It can be noticed that the diseases were not completely recognized. This is the most visible in the second test image. Even though three clear lines of disease are presented, our proposed method when CIALAB color

Figure 3: Plant diseases detection using HSI color model

Figure 4: Plant diseases detection using $CIELAB$ color model

model is used barely recognized only one. This leads to the conclusion that this color model is not appropriate for early disease detection which is one of the goals of the proposed algorithm.

Based on the experimental results some conclusions can be made. We tested the proposed algorithm and used different color models. Results can be summarized as the follows:

- Using threshold on RGB image disease spot is

almost unusable.

- Using threshold on A element of CIELAB color model and C_r element of YC_bC_r color model, disease spots are detected however not correctly. Therefore, results are not satisfying.

- Results show that exploitation threshold on H component of HSI color model disease spots are detected accurately and results better comparing to the other tested in this paper.
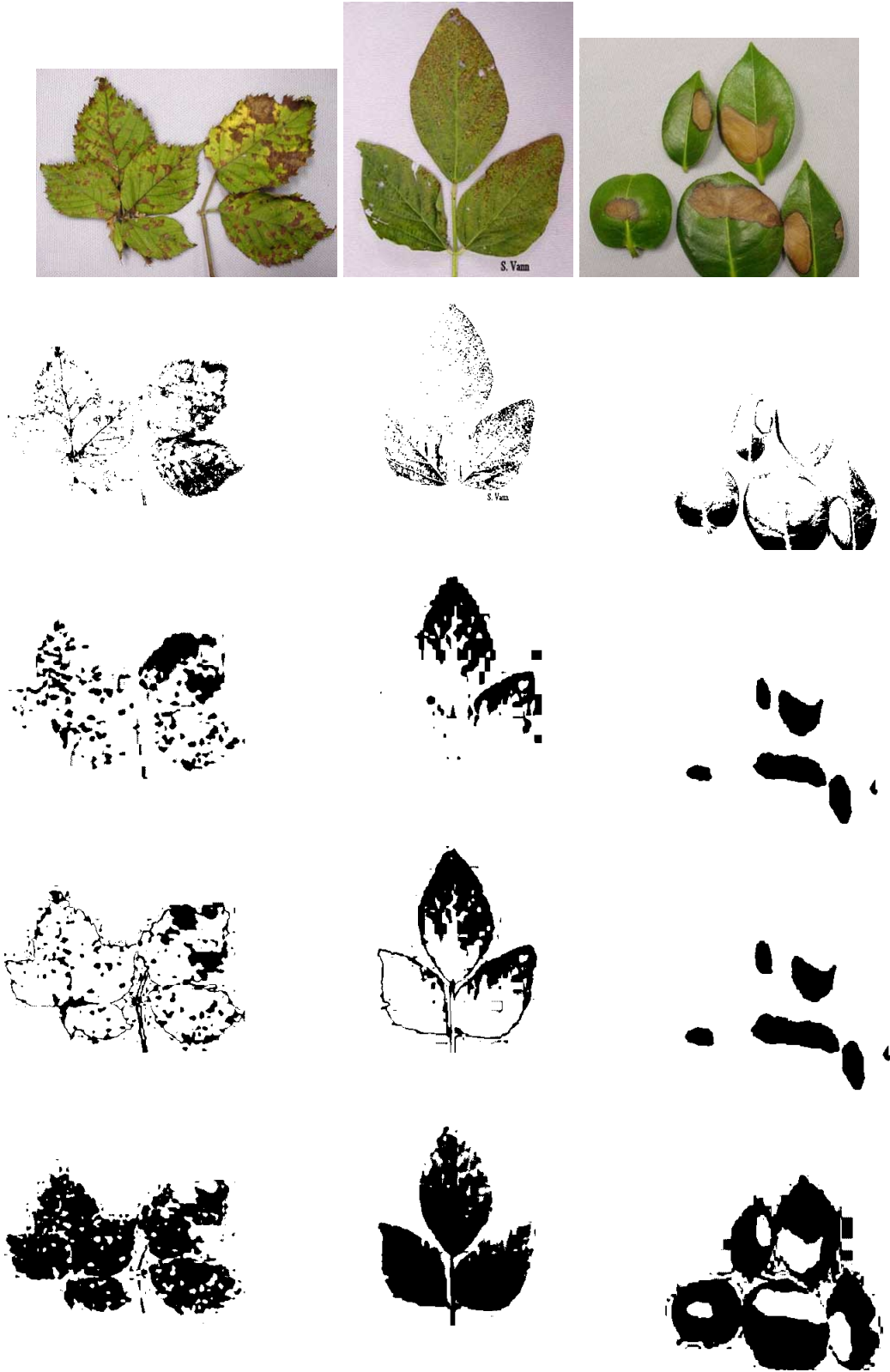
Figure 5: Diseases detection with different color models on images with the background

Next, experiments included three additional images of the plant but with the background. In these images noise was introduced by adding the background that can disturb diseases detection. Experimental results for disease spot detection of some leafs suffering from different diseases are shown in Fig. 5. In the first row are the original images. In the second row are shown results of our proposed method by using RGB color model, in the third YC_bC_r model, in the forth row are shown detection when HSI color model was used and at the end, in the fifth row are results when CIELAB color model.

Based on this results, it can be concluded that:

- Exploitation threshold on RGB image disease spots again are not detected and disturbance due to vein is present, thus this color model is not applicable in this application.

- Using threshold on C_r element of YCbCr color model, some disease spots area unit detected effectively, but again in the first two examples more parts of the leaf were recognized as disease than they really are.

- Using threshold on H element of HSI color model diseases were quite correctly recognized. Comparing to the previous color models, it is far the best one.

- Using threshold on A element of CIELAB color model disease spots may be detected accurately altogether. Obtained results are comparable to the one obtained with HSI color model. Again one recognition is not so good. The third example made a problem for this color model. It can be seen that the shadow made some recognition that are not accurate.

Result shows that disease spots might be detected accurately using HSI color model and CIELAB showed also promising results.

5 Conclusion

In this paper a method based on different color models and Kapur's thresholding for plant diseases detection was proposed. Four different color models were tested and compared: RGB, YCbCr, HSI and CIELAB color model. The best results were obtained when HSI color model was used. Component H was used for image segmentation where diseases were separated from the leaf. Median filter was applied to color transformed

image. At the end, disease spots area are determined by applying Kapur's threshold on different color components. Experimental result shows that noise that is introduced due to background, vein and camera flash makes the least problem for HSI color model. Following this technique totally different disease spots are detected accurately and results do not seem to be laid low with background, sort of leaf, type of disease spot and camera. In the further work, disease may be classified by calculative dimensions of disease spot.

References:

[1] C. Tsiotsios and M. Petrou, "On the choice of the parameters for anisotropic diffusion in image processing," *Pattern recognition*, vol. 46, no. 5, pp. 1369–1381, 2013.

[2] G. Stevenson, A. Welsh, L. Impey, J. A. Noble, and S. Collins, "An image processing technique for the visualization and quantification of blood flow entering the placenta using 3d power doppler ultrasound (pd-us)," *Placenta*, vol. 45, p. 85, 2016.

[3] M. Nikolic, E. Tuba, and M. Tuba, "Edge detection in medical ultrasound images using adjusted Canny edge detection algorithm," in *24th Telecommunications Forum (TELFOR)*. IEEE, 2016, pp. 691–694.

[4] A. Stojak, E. Tuba, and M. Tuba, "Framework for abnormality detection in magnetic resonance brain images," in *24th Telecommunications Forum (TELFOR)*. IEEE, 2016, pp. 687–690.

[5] A. Alihodzic and M. Tuba, "Bat algorithm (BA) for image thresholding," *Recent Researches in Telecommunications, Informatics, Electronics and Signal Processing*, pp. 364–369, 2013.

[6] I. Brajevic and M. Tuba, "Cuckoo search and firefly algorithm applied to multilevel image thresholding," in *LNCS, Cuckoo Search and Firefly Algorithm: Theory and Applications*, X.-S. Yang, Ed. Springer International Publishing, 2014, vol. 516, pp. 115–139.

[7] M. Tuba and N. Bacanin, "JPEG quantization tables selection by the firefly algorithm," in *International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE, 2014, pp. 153–158.

[8] E. Tuba and N. Bacanin, "An algorithm for handwritten digit recognition using projection histograms and SVM classifier," in *23rd Telecommunications Forum Telfor (TELFOR)*. IEEE, 2015, pp. 464–467.

[9] M. Lukic, E. Tuba, and M. Tuba, "Leaf recognition algorithm using support vector machine with Hu moments and local binary patterns," in *15th International Symposium onApplied Machine Intelligence and Informatics (SAMI)*. IEEE, 2017, pp. 000 485–000 490.

[10] A. Arsic, M. Jordanski, and M. Tuba, "Improved lip detection algorithm based on region segmentation and edge detection," in *23rd Telecommunications Forum (TELFOR)*. IEEE, 2015, pp. 472–475.

[11] J. G. A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," *SpringerPlus*, vol. 2, no. 1, p. 660, 2013.

[12] C. Bock, G. Poole, P. Parker, and T. Gottwald, "Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging," *Critical Reviews in Plant Sciences*, vol. 29, no. 2, pp. 59–107, 2010.

[13] D. Cui, Q. Zhang, M. Li, G. L. Hartman, and Y. Zhao, "Image processing methods for quantitatively detecting soybean rust from multispectral images," *Biosystems engineering*, vol. 107, no. 3, pp. 186–193, 2010.

[14] S. Kai, L. Zhikun, S. Hang, and G. Chunhong, "A research of maize disease image recognition of corn based on bp networks," in *Third International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, vol. 1. IEEE, 2011, pp. 246–249.

[15] N. N. Kurniawati, S. N. Abdullah, S. Abdullah, and S. Abdullah, "Investigation on image processing techniques for diagnosing paddy diseases," pp. 272–277, 2009.

[16] G. Ying, L. Miao, Y. Yuan, and H. Zelin, "A study on the method of image pre-processing for recognition of crop diseases," in *International Conference on Advanced Computer Control*. IEEE, 2009, pp. 202–206.

[17] D. Al Bashish, M. Braik, and S. Bani-Ahmad, "Detection and classification of leaf diseases using k-means-based segmentation and," *Information Technology Journal*, vol. 10, no. 2, pp. 267–275, 2011.

[18] S. Ananthi and S. V. Varthini, "Detection and classification of plant leaf diseases," *International Journal of research in engineering & applied Sciences*, vol. 2, no. 2, pp. 763–773, 2012.

[19] F. Garcia-Ruiz, S. Sankaran, J. M. Maja, W. S. Lee, J. Rasmussen, and R. Ehsani, "Comparison of two aerial imaging platforms for identification of huanglongbing-infected citrus trees," *Computers and Electronics in Agriculture*, vol. 91, pp. 106–115, 2013.

[20] G. Li, Z. Ma, and H. Wang, "Image recognition of grape downy mildew and grape powdery mildew based on support vector machine," *Computer and Computing Technologies in Agriculture V*, pp. 151–162, 2012.

[21] R. C. Hrosik, M. Tuba, and M. Vukovic, "Face detection algorithm based on skin detection and invariant moments," in *Recent Advances in Knowledge Engineering and Systems Science*, 2013, vol. 10, pp. 110–115.

[22] N. Abbas, D. Mohamad *et al.*, "Microscopic RGB color images enhancement for blood cells segmentation in YCbCr color space for k-means clustering," *Journal of Theoretical and Applied Information Technology*, vol. 55, no. 1, pp. 117–125, 2013.