Crop Protection 103 (2018) 98–102

Contents lists available at ScienceDirect

Crop Protection

journal homepage: www.elsevier.com/locate/cropro

Short communication

AuToDiDAC: Automated Tool for Disease Detection and Assessment for Cacao Black Pod Rot

Daniel Stanley Tan[a,e,*], Robert Neil Leong[b,e,f], Ann Franchesca Laguna[c], Courtney Anne Ngo[a], Angelyn Lao[b,e,f], Divina M. Amalin[d,g], Dionisio G. Alvindia[h]

[a] Software Technology Department, De La Salle University, Manila, Philippines
[b] Mathematics Department, De La Salle University, Manila, Philippines
[c] Computer Technology Department, De La Salle University, Manila, Philippines
[d] Biology Department, De La Salle University, Manila, Philippines
[e] Center for Complexity and Emerging Technologies, AdRIC, De La Salle University, Manila, Philippines
[f] Mathematical and Statistical Modeling Unit, CENSER, De La Salle University, Manila, Philippines
[g] Biological Control Research Unit, CENSER, De La Salle University, Manila, Philippines
[h] Philippine Center for Postharvest Development and Mechanization, Nueva Ecija, Philippines

ARTICLE INFO

Keywords:
Cacao
Black pod rot
Defect segmentation
Infection level
Disease management tool
Decision support

ABSTRACT

Pest control strategies for crop diseases highly depend on visual inspection to assess the severity of the infection, which usually lead to inconsistencies: either over or under assessment. These inconsistencies could be attributed to the limitations of humans to perceive small differences. A more precise disease assessment is needed for better pest management decision, which will result to a more efficient utilization and allocation of resources for farm inputs. This translates to a better income for cacao farmers. This paper introduces a mobile application named AuToDiDAC or Automated Tool for Disease Detection and Assessment for Cacao Black Pod Rot (BPR). AuToDiDAC automatically detects, separates, and assesses the infection level of BPR in cacao through image processing and machine learning techniques. It gives the farmers the capacity to objectively monitor and report the infection level of the BPR compared to the common visual rating for plant disease level of infection. Pixel-level accuracy test of the tool showed an average of 84% accuracy on an independent test set of ten cacao pod images.

1. Introduction

Cacao, a tropical crop which grows best with humid climate and adequate rainfall, is a highly-valued crop among equatorial regions. (Miller et al., 2008). While the Philippines used to be a top producer of cacao, poor pests and disease management in the 1980s led to the fall of its cacao market (Hebbar, 2007; Ploetz, 2007). One main cause of this downfall is the black pod rot (BPR) (Adomako, 2007; Appiah et al., 2004; Guest, 2007).

BPR is caused by Phytophthora spp fungi. It mostly affects the pods, but can also be observed in any part of the tree. It is visually described as small, hard, and dark lesions. The disease spreads rapidly, as it can cover the entire surface and tissues in a span of 11 days since infection (Philip-Mora and Cerda, 2009). BPR can be devastating when left untreated. It accounts for up to 90% of loss worldwide; dealing an economic impact estimated to cost as much as 450 thousand metric tons amounting to 423 million dollars of losses every year (Bowers et al., 2001). To control the disease, many management procedures have been proposed, each having their own strengths and limitations. These include introduction of biological control agents, chemical pesticides application, phytosanitation, and genetic manipulation for resistant varieties. Newer management operations now consider an integrated approach with some combinations of the aforementioned procedures to take advantage of the strengths of one and mask the limitations of the other (Acebo-Guerrero et al., 2012). Regardless of the management approach, implementation of any control measure is reliant on the early and correct detection of the disease. So far, cacao farmers and agricultural technicians use visual inspection to assess the severity of the disease following the severity index scale of 0–7, which was based on the severity index rating for anthracnose in mangoes (Alvindia and Acda, 2015). A sample of the visual aid used is shown in supplementary materials figure 1. However, this visual assessment is more often not

* Corresponding author. Software Technology Department, De La Salle University, Manila, Philippines.
E-mail addresses: daniel.tan@dlsu.edu.ph (D.S. Tan), robert.leong@dlsu.edu.ph (R.N. Leong), ann.laguna@dlsu.edu.ph (A.F. Laguna), courtney.ngo@dlsu.edu.ph (C.A. Ngo), angelyn.lao@dlsu.edu.ph (A. Lao), divina.amalin@dlsu.edu.ph (D.M. Amalin), dgalvindia@yahoo.com (D.G. Alvindia).

http://dx.doi.org/10.1016/j.cropro.2017.09.017
Received 18 January 2017; Received in revised form 28 September 2017; Accepted 29 September 2017
Available online 07 October 2017

accurate because of human error and bias in visually rating the diseases. To address the need of accurately rating the severity of cacao pod rot, AuToDiDAC was developed to automatically inspect and detect cacao black pod rot. AuToDiDAC improves on human-intensive visual inspection procedures by minimizing subjectivity biases. An objective infection measure also assists in the development of a reliable decision framework for managing the incidence of cacao black pod rot.

The main objective of this study is to provide farmers the capacity to objectively monitor and report the infection level of the disease comparable to an expert's opinion by providing a mobile application (AuToDiDAC) that can isolate the cacao pod from its background and measure BPR infection level more accurately. The paper presents the framework detailing the flow in developing AuToDiDAC.

2. Brief overview of related works

Existing works on fruit defect or disease detection usually focus on distinct colored fruits such as oranges (Blasco et al., 2009; López-Garcia et al., 2010), mandarin (Kamalakannan and Rajamanickam, 2012; Wang et al., 2013), strawberries (Liming and Yanchao, 2010), and apples (Jhuria et al., 2013; Dubey et al., 2013). Han et al., 2016 also developed an automated detection and severity assessment system of crop diseases using image pattern recognition. Mahlein (2016) provides a summary of existing works on disease detection. These systems made used of different algorithms like marker-controlled watershed segmentation, multi- and hyperspectral sensors, thermography, and others. They mainly rely on color analysis and thresholding to separate out the defects. These approaches may not directly work for cacao pods due to the large color variations of cacao pods which can take on green, yellow, orange, red, purple or maroon. This makes it harder to impose a simple threshold on colors as the disease may have very similar colors to the natural purple or maroon colors of cacao pods. Our approach makes use of a learned support vector classifier which can adapt to the varying colors of cacao pods.

The objective of most of the previous systems are for fruit grading and sorting which is done after harvesting the fruits. This makes it easy to control the setting or environment at which they conduct the assessment. In contrast, we aim to provide a solution for disease detection and assessment of cacao pods in the wild and is applicable before, during and after harvesting. We also aim to make it accessible for farmers without the need for expensive machineries such as thermal sensors and hyperspectral sensors used in the works mentioned in Mahlein (2016).

3. Materials and methods

3.1. Cacao pod image collection and background removal

A total of 60 cacao pods with varying levels of infection were selected from the cacao plantation in Davao City managed by the Cocoa Foundation of the Philippines. AuToDiDAC includes a camera mobile application feature for image capturing to eliminate unnecessary complexities and variations. This mobile application ensured that the picture taken would have the cacao pod image centered with minimal variations on the size and the orientation of the cacao pod using an overlaid guide which can be manipulated (i.e. resizing and rotating) as necessary to accommodate the natural variations in the shape of the cacao pods. A sample screenshot of the mobile application can be found in supplementary materials figure 2.

Isolation of the cacao pod image was done by removing unwanted background elements in the image using graph cuts algorithm (Boykov and Jolly, 2001), which automatically extracts the foreground, the cacao pod, from its background, thereby ensuring that the program only reads the pod in measuring infection level. The guide on the mobile application provided a straightforward approach for marking parts of the image for the algorithm to differentiate the foreground and the background, thus eliminating human input.

A simple color balancing algorithm (Limare et al., 2011) was applied to adjust the colors and minimize the effect of varying lighting conditions. The graph cuts algorithm then identified the background pixels which was replaced with a plain black color, i.e. a value of 0 in all three RGB color channels.

3.2. Cacao pod infection identification

3.2.1. Clustering infected and uninfected parts of the cacao pod

Identifying the infected parts of the cacao pod from the captured image was done by grouping similar pixels together using Piotr's implementation of fast k-means clustering algorithm (Dollár, n.d.). Inspired by Dubey's work on apples (Dubey et al., 2013), images from the RGB color space were converted to the L*a*b* color space. Only the a* and b* channels were used for the k-means clustering algorithm. The outputs are pixel clusters which ideally should contain only “uninfected” pixels or “infected” pixels. This clustering step provided a fast separation of the “infected” pixels from the “uninfected” pixels, which was useful in estimating the spread of BPR.

3.2.2. Labeling clusters of the cacao pod image using Support Vector Machines

To infer whether the clusters formed in the previous step were “uninfected” or “infected”, a Support Vector Machine (SVM) was used. The job of the SVM classifier is to learn how to infer the labels of the clusters that it has never seen before based on examples that were initially provided. Training the SVM classifier was done by labeling the infected portions of the cacao pod in the training set using the brush tool in Adobe Photoshop. A sample output of this task is shown in Fig. 1. These labels determined whether the clusters given by the k-means algorithm were “uninfected” or “infected” and served as ground truth for evaluating the tool. The clusters were represented in terms of the mean, standard deviation, and skewness of each channel in the L*a*b* color space.

3.3. Measuring the infection level

Infection level was computed in terms of percentages by getting the area of the infection divided by the size of the crop. In formula form, it is expressed as $\text{Infection Level} = \frac{\text{Area of the infection}}{\text{Size of the crop}} * 100\%$, where the area of infection was estimated by counting the number of infected pixels, and the size of the crop was estimated by counting the number of pixels of the crop (excluding the background).

3.4. Calibration and evaluation of the tool

We randomly selected 50 images out of the 60 collected from the field were used to train and calibrate the model. The remaining 10 images were set aside as an independent test set to evaluate the performance of AuToDiDAC. This lets us estimate how the model performs to images that may be encountered on the field which the model has never seen before. AuToDiDAC was calibrated with respect to the accuracy of the output after performing k-means and SVM classification. Since the size of the training dataset is small, a five-fold cross-validation was performed to determine the best value for the parameter k for the k-means algorithm. In a five-fold cross-validation, the datasets are obtained from five equal partitions of the entire training set. A model will be trained using only four parts and then evaluated on the last one remaining part. The process is repeated for five rounds with each round having different combinations of parts for training and evaluation. This produces five different accuracy measures which are averaged to a single value that represents an estimate of the performance of the model.

Based on previous work (Tan et al., 2016), a minimum number of

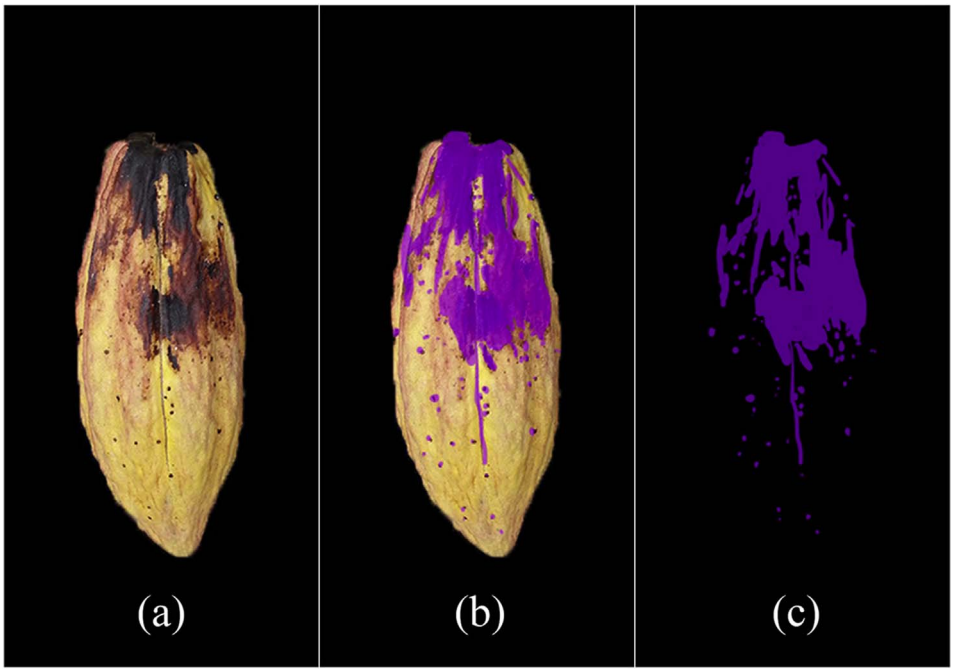

Fig. 1. (a) The original image of the cacao pod after background subtraction. (b) Labels (purple) of the infected parts overlaid on the image. (c) The label without the original image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$k = 4$ was required to satisfactorily separate "infected" from "uninfected" pixels so the parameter k for the k-means algorithm was varied from starting from 4 to 10. For each k, a five-fold cross-validation was performed. The optimal k was selected based on the value that generated the highest average accuracy. The pixel level accuracy, computed as the percentage of identified pixel labels by AuToDiDAC that matched the labels used in SVM, was used to compute evaluation metrics on the test set. The same set of test images was also visually rated by a plant pathologist specializing on cacao diseases using the rating in (Alvindia and Acda, 2015). Corresponding results of the visual rating were compared with AuToDiDAC's outputs. Wilcoxon's matched-pairs signed-rank test was used to assess the statistical significance of the difference between the two rating scales. Statistical analyses were performed in Stata 12.

4. Results and discussion

4.1. Calibrating the *k*-means algorithm

From the five-fold cross-validation, the best value of k was observed to be $k = 5$ achieving an average accuracy of 86.96% with a reasonable average processing time of 2.42 s per image. The accuracy further declined for $k = 6$, remained practically similar until $k = 8$, and then declined again for $k = 9$, showing signs that the model was over-fitting for the higher values of k (Fig. 2).

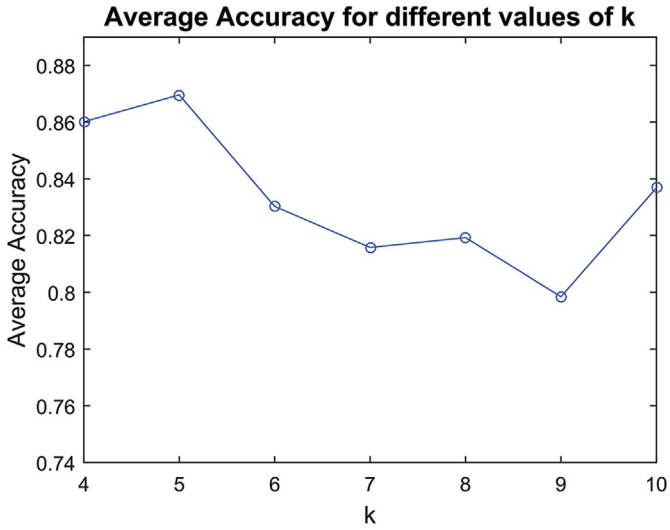

Fig. 2. The average accuracy of the five-fold cross-validation of the SVM classifier with respect to varying values of k.

4.2. Evaluation of AuToDiDAC

After setting k to five, the model was evaluated on the independent test set. Fig. 3 shows the resulting detections of AuToDiDAC on the test images. The actual accuracies, precision, and recall of every test image is reported in supplementary materials table 1. It was observed that on average, 84.59% of the pixels were accurately classified (SD = 0.0940). The labeled clusters formed from the previous steps were merged together to visualize the portions labeled as "uninfected" as well as the portions labeled as "infected", shown in Fig. 3. Visually evaluating the results, the tool adequately identified and separated the infected regions of the cacao pods. This implied that the computed infection levels were reliable, except for a few instances. For example, the infected regions of test image 9 was overestimated wherein the top part of the pod was classified as infected when in fact it was not. This might be due to the unfavorable lighting condition which oversaturated the top part of the pod. In such a case, color balancing does not work well since when forced to correct the top part, the bottom region will then be oversaturated resulting to poor classification performance. This can be avoided by ensuring proper lighting and positioning when taking pictures of the cacao pod. On the other hand, AuToDiDAC was sensitive enough to capture infected portions of the pod that humans tend to overlook. Examples of which are test images 3 and 10 shown in Fig. 3. Fig. 4a shows test image number 3, which noticeably had multiple sporadic infection sites. This made it difficult to quantify and annotate the infections, which is evident in the comparison between the detected infection and expert's annotation in Fig. 4b and c. This difficulty may have led to disparities between the results of AuToDiDAC and visual inspection. In the latter, the large portions of discoloration in the middle was easily recognizable as compared to the infection on the bottom part of the pod. These issues are the reason why visual inspection is not reliable, particularly when BPR are still on its early

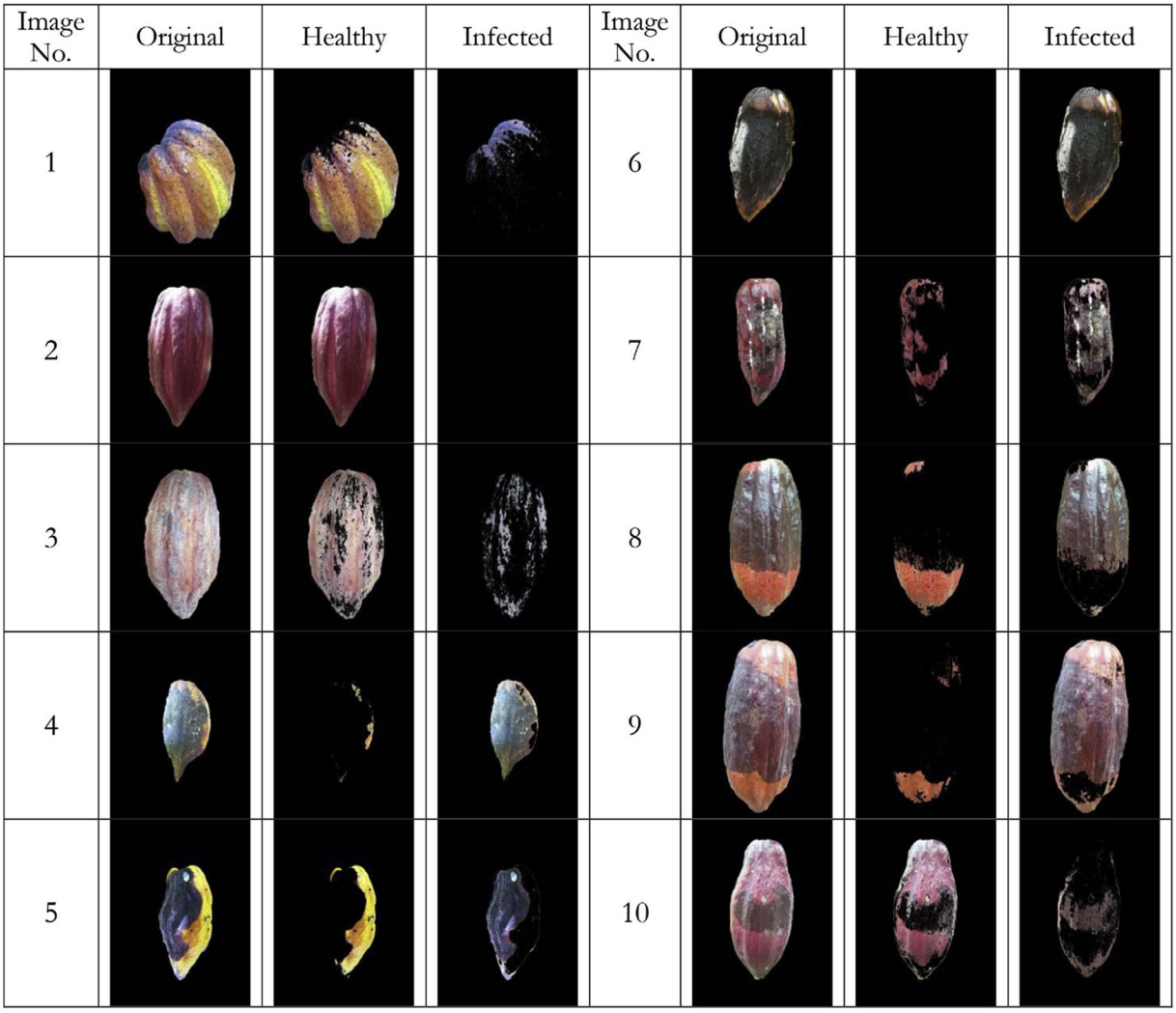

Fig. 3. Results of AuToDiDAC's separation and classification on the test images.

stages where it is easily overlooked.

For each test image, the output of AuToDiDAC were converted into their corresponding levels using the same scale as proposed by (Alvindia and Acda, 2015). Scale rating by the plant pathologist on the same set of images were compared to the output of AuToDiDAC (Table 1). Only 4 matches were found out of 10 test images. For the other 6 that did not match, AuToDiDAC revealed higher estimates than the plant pathologist's ratings. Such difference was found to be statistically significant (*Wilcoxon Matched-Pairs Signed-Rank Test p-value = 0.0165*). It is speculated that this was because the tool captured the small and subtle discolorations that human eyes tend to overlook as noted earlier. On the other hand, another reason may be because of the limitations in the visual perspective of the tool, which only saw one side projected in a 2D space as compared to humans which considered the entire 3D fruit.

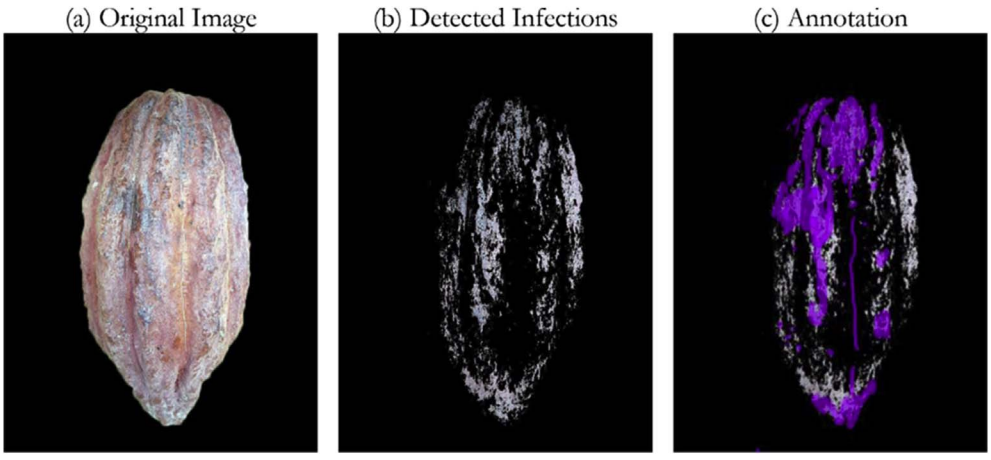

Fig. 4. (a) The original photo of Image no. 3; (b) Detected infections by AuToDiDAC; (c) Infected labels annotated by the plant pathologist.

Table 1
Computed ratings of AuToDiDAC compared with visual rating.

Image No.	Infection Level	Scale Rating[a]	
		AuToDiDAC	Visual
1	11.25%	2	2
2	0.00%	0	0
3	19.33%	2	1
4	93.81%	7	7
5	61.93%	5	3
6	100.00%	7	4
7	66.89%	5	2
8	78.52%	6	6
9	90.07%	7	6
10	54.76%	5	2

[a] Wilcoxon Matched-Pairs Signed-Rank Test p-value = 0.0165.

5. Conclusion

In summary, this paper detailed how AuToDiDAC automatically measures the infection level of BPR in cacao pods. While there have been many similar works that looked into defect detection for various fruits (Blasco et al., 2009; Dubey et al., 2013; Jhuria et al., 2013; Liming and Yanchao, 2010; Omid et al., 2010), this is so far the first documented work for cacao. Almost all the previous works were focused on fruit grading and sorting that required the fruit to be plucked from the tree and placed in a controlled setting with optimal lighting and camera positioning. In contrast, the advantage of AuToDiDAC is that it does not require destructive sampling of the pods. In early detection of BPR the fruit is required to be attached to the tree. Additionally, AuToDiDAC is a simple and user-friendly tool that does not require special machineries and can run on low cost android phones.

Our study showed that AuToDiDAC provided a more precise BPR infection rating for better disease assessment and control if necessary. Even though this work focused on cacao pods, the tool can easily be extended and calibrated to accommodate other crops showing related diseases. It can be used by farmers and agricultural technicians in assessing crop disease infection with minimal effort. Researchers and plant pathologists can also leverage on AuToDiDAC to measure the spread of the infection over time and assess the effectiveness of various control strategies. Future research endeavors to further improve this tool include accommodating multiple views to provide more accurate measures through reconstruction of the crop in 3D space. Additionally, repeated trials in different seasons and lighting conditions will be conducted to ascertain robustness of the detection and severity assessment.

Acknowledgements

The authors would like to acknowledge De La Salle University – University Research Council Office (URCO) for funding this research with project number #12 IR 4TAY14-1TAY16. The authors would also like to thank Cocoa Foundation of the Philippines for providing access to their farms for data collection.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cropro.2017.09.017.

References

Acebo-Guerrero, Y., Hernández-Rodriguez, A., Heydrich-Pérez, M., El Jaziri, M., Hernández-Lauzardo, A.N., 2012. Management of black pod rot in cacao (Theobroma cacao L.): a review. Fruits 67 (1), 41–48.
Adomako, B., 2007. Causes and extent of yield losses in cocoa progenies. Trop. Sci. 47 (1), 22–25.
Alvindia, D.G., Acda, M.A., 2015. Revisiting the efficacy of hot water treatment in managing anthracnose and stem-end rot diseases of mango cv.“Carabao.”. Crop Prot. 67, 96–101.
Appiah, A.A., Opoku, I.Y., Akrofi, A.Y., 2004. Natural occurrence and distribution of stem cankers caused by Phytophthora megakarya and Phytophthora palmivoraon cocoa. Eur. J. Plant Pathology 110 (10), 983–990.
Blasco, J., Aleixos, N., Cubero, S., Gómez-Sanchis, J., Moltó, E., 2009. Automatic sorting of satsuma (Citrus unshiu) segments using computer vision and morphological features. Comput. Electron. Agric. 66 (1), 1–8.
Bowers, J.H., Bailey, B.A., Hebbar, P.K., Sanogo, S., Lumsden, R.D., others, 2001. The impact of plant diseases on world chocolate production. Plant Health Prog. 10.
Boykov, Y.Y., Jolly, M.-P., 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. 1. pp. 105–112.
Dollár, P. (n.d.). {P}iotr's {C}omputer {V}ision {M}atlab {T}oolbox ({PMT}).
Dubey, S.R., Dixit, P., Singh, N., Gupta, J.P., 2013. Infected fruit part detection using K-means clustering segmentation technique. IJIMAI 2 (2), 65–72.
Guest, D., 2007. Black pod: diverse pathogens with a global impact on cocoa yield. Phytopathology 97 (12), 1650–1653.
Han, L., Haleem, M.S., Taylor, M., 2016. Automatic detection and severity assessment of crop diseases using image pattern recognition. Stud. Comput. Intell. 647. https://doi.org/10.1007/978-3-319-33353-3_15.
Hebbar, P.K., 2007. Cacao diseases: a global perspective from an industry point of view. Phytopathology 97 (12), 1658–1663.
Jhuria, M., Kumar, A., Borse, R., 2013. Image processing for smart farming: detection of disease and fruit grading. In: Image Information Processing (ICIIP), 2013 IEEE Second International Conference on, pp. 521–526.
Kamalakannan, A., Rajamanickam, G., 2012. Surface defect detection and classification in mandarin fruits using fuzzy image thresholding, binary wavelet transform and linear classifier model. In: 4th International Conference on Advanced Computing, ICoAC 2012, . https://doi.org/10.1109/ICoAC.2012.6416829.
Limare, N., Lisani, J.-L., Morel, J.-M., Petro, A.B., Sbert, C., 2011. Simplest color balance. Image Process. Line 1.
Liming, X., Yanchao, Z., 2010. Automated strawberry grading system based on image processing. Comput. Electron. Agric. 71, S32–S39.
López-Garcia, F., Andreu-Garcia, G., Blasco, J., Aleixos, N., Valiente, J.-M., 2010. Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach. Comput. Electron. Agric. 71 (2), 189–197.
Mahlein, A.-K., 2016. Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping. Plant Dis. 100 (2), 241–251. https://doi.org/10.1094/PDIS-03-15-0340-FE.
Miller, K.B., Hurst, W.J., Payne, M.J., Stuart, D.A., Apgar, J., Sweigart, D.S., Ou, B., 2008. Impact of alkalization on the antioxidant and flavanol content of commercial cocoa powders. J. Agric. Food Chem. 56 (18), 8527–8533.
Omid, M., Abbasgolipour, M., Keyhani, A., Mohtasebi, S.S., 2010. Implementation of an efficient image processing algorithm for grading raisins. Int. J. Signal Image Process. 1 (1), 31–34.
Philip-Mora, W., Cerda, R., 2009. Catalog: Cacao Diseases in Central America.
Ploetz, R.C., 2007. Cacao diseases: important threats to chocolate production worldwide. Phytopathology 97 (12), 1634–1639.
Tan, D.S., Leong, R.N., Laguna, A.F., Ngo, C.A., Lao, A., Amalin, D., Alvindia, D., 2016. A framework for measuring infection level on cacao pods. In: 2016 IEEE Region 10 Symposium (TENSYMP), pp. 384–389.
Wang, L., Li, A., Tian, X., 2013. Detection of fruit skin defects using machine vision system. In: Business Intelligence and Financial Engineering (BIFE), 2013 Sixth International Conference on, pp. 44–48.