International Journal of Engineering & Technology, 7 (4) (2018) 4880-4885

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.26553
Research paper

A scale-invariant lettuce leaf area calculation using machine vision and knowledge-based methods

Pocholo James M. Loresco [1] *, Elmer P. Dadios [2]

[1] Department of Electronics and Communications Engineering, De La Salle University, Manila Philippines
[2] Department of Manufacturing Engineering and Management, De La Salle University, Manila Philippines
**Corresponding author E-mail: pocholo_loresco@dlsu.edu.ph*

Abstract

Leaf area is one of the most significant reference tools to characterize plant growth and predict growth stages. Scale invariance in calculating leaf area needs to be understood in the lettuce growth monitoring context. Using machine vision and knowledge-based classifiers, this research produced a system for a scale invariant area calculation of lettuce leaf area by detecting a template marker with a known area for scaling area measurements. Results showed that knowledge-based algorithm can improve the performance of the machine vision classifiers for rejecting false positives even for a limited number of training datasets. Area measurements produced by the system performed well in terms of root-mean-square error (RMSE).

Keywords: *Knowledge-Based; Lettuce Leaf Area; Machine Vision; Scale Invariance.*

1. Introduction

Controlled-environment agriculture is an integrated science and engineering approach to horticultural technology with the aim of promoting crop production in controlled environments called as smart farms [1]. Growth analysis can account for environmental factors, such as radiation, temperature, and water in relation with the state of the crop, development stage and nutrient availability [2]. Leaf area is one of the most significant reference tools to characterize plant growth and predict growth stages [3 - 5]. Canopy area calculated from the segmented lettuce images can be compared in the learned thresholds that indicate the lettuce growth whether there are sowing, vegetative, or harvest stage [6]. The accurate determination of lettuce growth stage determines correctly the optimum supply or distribution of the related crop growth environmental factors such as radiation, water and temperature that can improve the number of harvestable products.

Although the vision-based lettuce growth monitoring has encountered several developments in recent years, some problems remain an open research area [7] [8]. Current studies have not satisfactorily resolved the challenges and requires improvement and extensive testing for functionality and robustness when applied in controlled greenhouse environment or smart farms [9]. Consideration of scale invariance also has remained open for research in the lettuce growth monitoring context [10] [11]. The accuracy of the calculation of area invariant to the scale at which the image is taken is very important in properly deriving other shape and morphological features.

The approach here used in this smart farm setup to address scale invariance is to embed a template or marker of known area into the image as shown in Figure 1. This known area of the marker is used for the system to scale area measurements. The number of pixels of a reference component shown in Figure (b) is determined and relate to the known area to produce the scaling factor in Equation 1. The area of the black square (1.8 cm by 1.8 cm) component of the marker used in this research bounded by red outline shown in Figure 1(b) is 3.24 cm^2. The scaling factor computed is then multiplied to the number of pixels of segmented lettuce leaf to calculate the lettuce leaf area in cm^2.

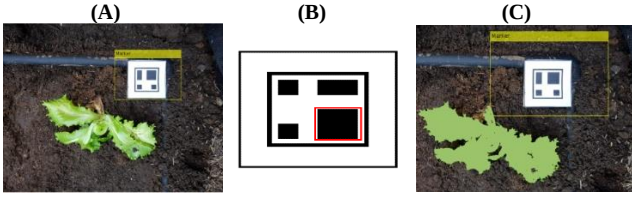


Fig. 1: (A) Marker to Detect in Yellow Bounding Box (B) Reference Component of Known Area Bounded by Red Box (C) Segmented Lettuce in Green where Pixels are Computed and Multiplied to the Scaling Factor to Give the Area in cm^2

$$L = P\frac{3.24}{C} \quad (1)$$

Where L is actual lettuce area in sq. cm, P is actual pixel area and C is reference component pixel area

This study proposes a scale invariant lettuce area calculation which uses combination machine vision and knowledge-based methods for object detection and spatial oriented image processing. Machine vision classifier is used for detecting the marker and knowledge-based methods are used to filter out the false positive detections. Spatial oriented image processing techniques are then used to extract the scaling factor and consequently the area. Computer vision metrics such as sensitivity, false positive rate, precision and accuracy are used to evaluate the performance of the proposed system. The paper is organized as follows: Section II covers the proposed methodology that discusses the combined machine vision and

Copyright © 2018 Pocholo James M. Loresco, Elmer P. Dadios. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

knowledge-based methods for object detection, lettuce leaf segmentation, area calculation and performance evaluation. Section III presents and analyzes the results. Finally, Section IV gives conclusions and future works.

2. Methodology

Figure 2 illustrates the proposed framework of the methodology. It includes input lettuce images with the said marker that undergo machine vision and knowledge-based object detection to determine the marker and extract out the scaling factor, determination of lettuce pixels from segmentation based on color, area calculation and output as lettuce leaf area in cm^2.

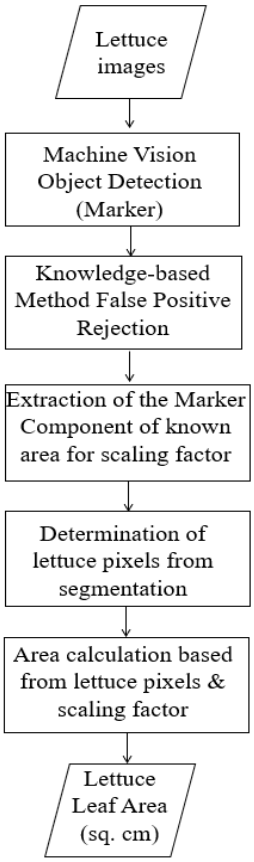


Fig. 2: Proposed Methodology for Area Calculation.

2.1. Proposed algorithm for object detection

The methodology of the proposed algorithm for object detection is illustrated in Figure 3. The objective here is to locate the object of interest which is the marker within an image.

The methodology consists of machine vision and knowledge-based methods. Machine vision methods composed of training phase and testing phase where lettuce images with or without markers were the inputs. Image pre-processing that included histogram equalization and median filtering enhanced the input lettuce image in preparation to feature extraction. Haar-like, Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) features were then extracted out of the images that were used to set up the machine vision classifier.

Training was followed by testing to validate the results. In the testing phase, since the system has enough training pairs, the system therefore was able to detect the marker when new inputs were introduced. However, the object detection performance was degraded by the occurrence of false positives. The knowledge-based methods addressed this by the eliminating the false positives.

2.1.1. Machine vision object detection

One of the most commonly used machine vision technique for object detection is Viola and Jones algorithm for its high accuracy, low false positive detection rate and speed of detection. This efficiency of the method is due to three key components in Viola and Jones's technique which are computing integral picture, adaptive boosting, and cascading [12].

The framework for the machine vision classifier was built from the representations from the extracted features Haar-like, HOG and LBP that created three XMLs. The XML files contained codes transformed from marker images that served as basis for the classifier to identify the marker thru the system.

In training the object detection system, positive and negative images were required. Positive samples contain image object (marker) to be detected while Negative samples do not contain the object of interest however should contain the backgrounds ordinarily associated with the object to be detected.

Similar set of training data were used for each feature. The same parameter values were also considered in the training as presented in Table 1. The object detector was trained using 1000 stages from a small training size 134 images using Haar-like, HOG and LBP features. There was limited training set of 134 images and this was compensated by adjusting the number of stages and setting a lower positive rate for each stage. The true positive rate of 0.999 and a false positive rate of 0.001 were used in this research. True positive rate is the rate of correctly identified markers while the false positive rate is the of incorrectly identified markers.

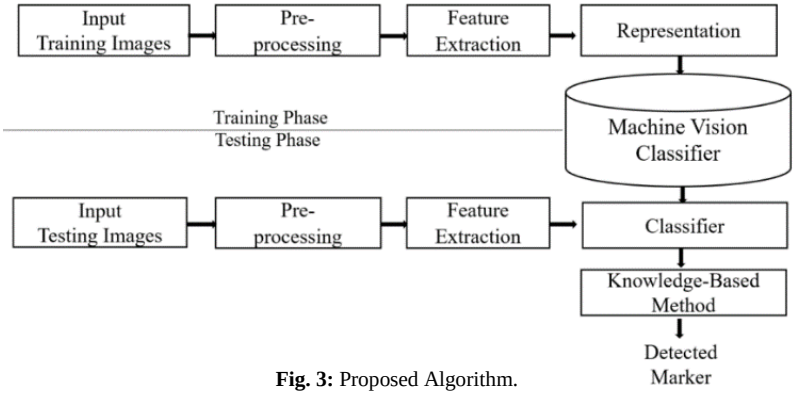


Fig. 3: Proposed Algorithm.

Table 1: Training Parameters

Training parameters	Value
No. of positive images	134
No. of negative images	134
No. of stages	1000
False positive rate	0.001
True positive rate	0.999

2.1.2. Knowledge-based object detection

Although the machine vision presented here works well in marker detection under most conditions, the misidentification of objects degrades its performances. Most of the markers were detected using any feature, however there was a high occurrence of false positives. An example is shown in Figure 4. This has a disastrous impact on the performance of the object detection as describe on the next section, Table 2. This poor performance may be attributed to the low number of training dataset used in the study [13].

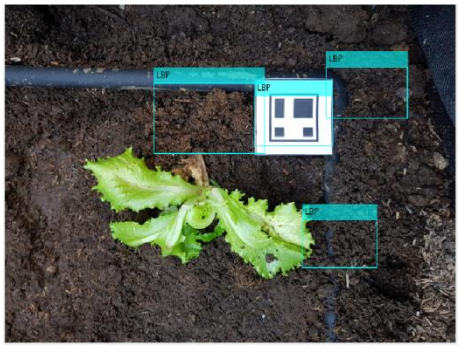


Fig. 4: A Detected Marker Using LBP Features with Several False Positives.

It is recommended therefore to address the filtering out of false positives. This paper presents an algorithm that eliminated false positives from the set of recognized objects. The models were refined by adding knowledge-based classifier corresponding to the unique image features of the markers such as color, length of the image and number of connected objects. HOG was selected as a feature for the machine vision classifier coupled to the knowledge- based features presented here. These are the features used: average RGB color, RGB color difference, marker image length, and number of connected objects. Rules were made according to these features and presented in Figure 5.

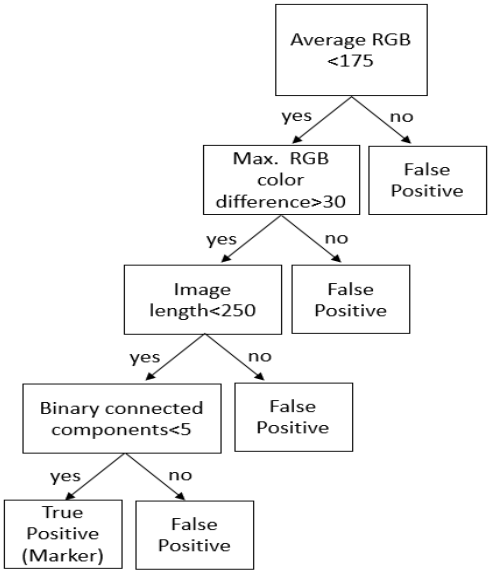


Fig. 5: Knowledge-Based Classification.

Average RGB color

Images can be represented in colors space in RGB components where colors are characterized as a combination of various intensities of red, green, and blue. Average RGB color is the average of value for R, G, B vectors extracted from the target images as shown in Equation 2. The value of average RGB color of the marker was said to be by experiment within a threshold value of less than 175.

$$\text{Average RGB color} = \frac{\frac{\sum_{i=i=1}^{n} R(i,j)}{n} + \frac{\sum_{i=i=1}^{n} G(i,j)}{n} + \frac{\sum_{i=i=1}^{n} B(i,j)}{n}}{3} \quad (2)$$

Where R(i,j), G(i,j), and B(i,j) are R,G,B pixels with spatial coordinates (i,j) and n is the total number of pixels in R,G,B vectors.

RGB color difference

The average RGB color information was not enough to discriminate true positives to false positives thus another color information was considered. The absolute maximum value of the difference between the R, G and B vectors were calculated as shown in Equation 3. Then the maximum value was determined among these three vectors were determined as in Equation 4. This value should be at least greater than the learned threshold value 30.

$$\begin{bmatrix} R_m \\ G_m \\ B_m \end{bmatrix} = \begin{bmatrix} \max(R-G) \\ \max(R-B) \\ \max(G-B) \end{bmatrix} \quad (3)$$

$$\text{RGB color difference} = \max\left(\begin{bmatrix} R_m \\ G_m \\ B_m \end{bmatrix}\right) \quad (4)$$

Where R_m, G_m, and B_m are maximum values from the difference of R, G, B vectors respectively.

Marker length

The lengths of largest array dimension of the candidate objects were determined from equivalent binary images. Based on these values, the marker length should be less than the learned threshold 250, otherwise the candidate marker was rejected. This threshold was determined experimentally based on the camera range available.

Number of Connected Components

The candidate markers were binarized then the number of connected components were determined. A masked image was created to better isolate the connected objects. Logical AND was performed on the complemented binary and the masked image. The resulting image is shown in Figure 6. Morphological operations dilation, erosion and area opening [14] also helped in eliminating unwanted objects that were relatively smaller and bigger then experimental thresholds. As can be seen in the Figure, there are 4 connected components. The number of connected objects was identified using connectivity of two-dimensional eight-connected neighborhood. If the number of connected components was greater than 5, then the false positives were rejected.

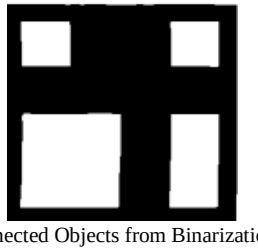


Fig. 6: Four Connected Objects from Binarization and Masking.

2.1.3. Testing the object detection algorithm

The proposed object detection methodology was evaluated at different altitudes and light illumination to test its effectiveness. A positioned camera taken top view of the canopy between 1 foot, 3 feet

and 5 feet of altitude away from the lettuce. Four variation of light illumination were considered with Photosynthetically Active Radiation (PAR) values as 92 µmol $m^{-2}s^{-1}$, 44 µmol $m^{-2}s^{-1}$, 7µmol $m^{-2}s^{-1}$, 3 µmol $m^{-2}s^{-1}$ measured from a Vernier PAR sensor.

Sensitivity, false positive rate, precision and accuracy in Equations (9) to (12) are the statistical evaluation criteria [15] used for the performance of the marker detection.

$$\text{sensitivity} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{precision} = \frac{TP}{TP + FP} \tag{6}$$

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$\text{false positive rate} = \frac{FP}{FP + TN} \tag{8}$$

Where TP true positive, correctly identified markers, FP false positive, incorrectly identified markers TN true negative, correctly identified non-markers and FN false negative, incorrectly identified non-markers.

2.2. Determination of the scaling factor from the image

After a true positive was extracted out from the false positives, bounding boxes were drawn to the marker image to indicate and verify detection as shown in Figure 1(a). This marker image was used to isolate the reference component of known area as shown in Figure 1(b)

The largest area among the found connected objects is the target reference component. The area in pixels of the region covered by this component was calculated. The actual area of the reference component is equal to 3.24 cm^2 was divided by the calculated area in pixels as shown in Equation 1, producing the scaling factor in cm^2 per pixel area.

2.3. Area calculation

Segmentation of the lettuce leaf pixels is required first before area calculation can be made. Segmentation differentiated the lettuce (foreground), from the background or environment the soil, irrigation pipes etc. K-Nearest neighbor algorithm using YCbCr color space was used in this research for this segmentation. In this color representation, the Y characterizes brightness, Cb characterizes the blue-yellow spectrum, and Cr component characterizes the red-green spectrum [16]. The transformation is given by Equation (9).

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{9}$$

Where Y, Cb, Cr represents the Luma, Blue Chroma and Red Chroma values respectively while R, G, B represents the Red, Green, Blue color values.

$$E(r,c) = \sqrt{\sum_{i=1}^{n} (Cb_r - Cb_c)^2 + (Cr_r - Cr_c)^2} \tag{10}$$

Where E (i, j): The Euclidean distance of a pixel with spatial coordinates (r, c) using color components Cb_i, Cr_i (YCbCr)

K-Nearest neighbor using Euclidian distance was used to classify objects within the image as lettuce or non-lettuce pixels [17]. The Euclidian distances between that pixels and each learned threshold color information were then computed using Equation 10. The smallest distance indicate that the pixel most closely matches that threshold. For instance, distances between a pixel and the background color threshold and lettuce color threshold were computed and compared. If the smallest distance was computed using lettuce color threshold, then the pixel would be labeled as a green pixel as shown in Figure 7. The pixel area of this green region was then determined and multiplied to the scaling factor. The resulting value was said to be the lettuce leaf area in cm^2.

A Graphical User Interface (GUI) was developed to aid the testing of marker detection and area calculation. Figure 7 illustrates the graphical user interface and its components. Using the GUI, the user was able to verify the following (1) original image of lettuce (2) detected marker with yellow bounding box (3) detected reference component (4) reference component pixel area (5) segmented lettuce leaf pixel area (6) scaling factor and (7) lettuce leaf area in cm^2.

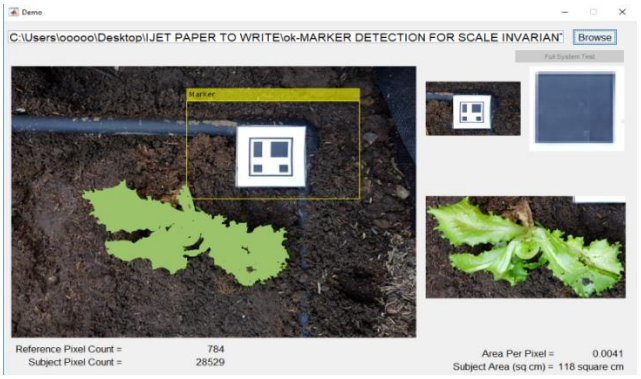


Fig. 7: GUI Developed to Test the Methodologies.

Finally, the root-mean-square error (RMSE) was used to calculate the performance of the system in terms of area calculation [18]. One hundred true positive images at varying illumination and altitude were used for testing. The reference component area in cm^2 was used for the computation. This formula described in Equation 11 was used to compare the observed data using the GUI with expected reference component area as 3.24cm^2.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} [x_{exp}(i) - x_{obs}(i)]^2}{n}} \tag{11}$$

Where x_{exp} is the expected value, x_{obs} is the observed data and n is the no. of samples

3. Discussion and analysis of results

This section presents the results of the study that includes object detection and area calculation. The object detection algorithms were evaluated on a total of 170 images of lettuce images obtained from Smart farm. Example of detected markers with bounding boxes are shown in Figures 8(a) to 8(d). The performances for each method which are mainly classified as machine vision object detection (Viola-Jones-Haar, Viola-Jones-HOG, and Viola-Jones-LBP) and machine vision-knowledge based methods are summarized at Table 2.

Table 2: Performance Summary

	Sensitivity	False positive rate	Precision	Accuracy
Viola-Jones Haar	0.238372	0.935406	0.173361	0.143044
Viola-Jones HOG	0.378787	0.772925	0.297619	0.297423
Viola-Jones LBP	0.389743	0.854077	0.276363	0.257009
Proposed method	0.970149	0	1	0.988304

As can be seen in Table 2, Viola-Jones-Haar, Viola-Jones-HOG, and Viola-Jones-LBP performed relatively poorer with the proposed method. These low computer vision metrics can be attributed to small training size [13]. With the addition of knowledge-based classification, sensitivities, precision, and accuracy of the detectors greatly increased, that means more correct markers were detected and more false positives rejected. The results of this test indicated that knowledge-based algorithm improved the performance of the classifiers even if there was limited number of training datasets.

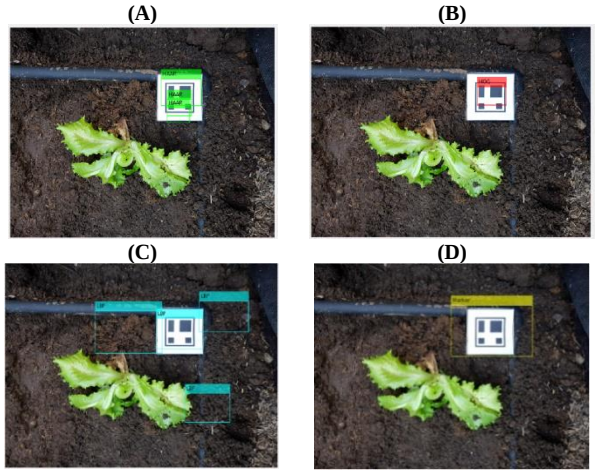


Fig. 8: Marker Detection Using (A) Viola-Jones-Haar, (B) Viola-Jones-HOG, (C) Viola-Jones-LBP and (D) Proposed Method.

Illumination and altitudes of capture were varied to test the methodologies. Distances such as 1 foot, 3 feet and 5 feet were used as point of capture away from the soil the lettuce crops were planted. Some of the results are presented in Figure 9(a) and 9(b). The performance of the proposed algorithm at increasing height and increasing illumination can be visualized by the graphs shown in Figures 10(a) and 10(b). Note that the numbers 1 to 4 corresponds to: 1-darkest up to 4-brightest, with PAR values as 1: 3 µmol $m^{-2}s^{-1}$, 2: 7µmol $m^{-2}s^{-1}$, 3: 44 µmol $m^{-2}s^{-1}$ and 4: 92 µmol $m^{-2}s^{-1}$.



Fig. 9: Marker Detection at (A) Various Illuminations, (B) Various Captures Height.

Figure 10(a) and 10(b) shows notably the stable and high performances in general of the proposed object detection, however a slight decrease in accuracy and sensitivity was observed when the images were taken at 5 feet and dark illumination.

Root-mean-square error were also computed to evaluate the performance of the system in terms of area calculation. RMSE gave the standard deviation of the differences between the actual values and the predicted values. Out of 100 images tested, the system gave a low RMSE value equal to 0.00024 in terms of predicting the area in terms of cm^2.

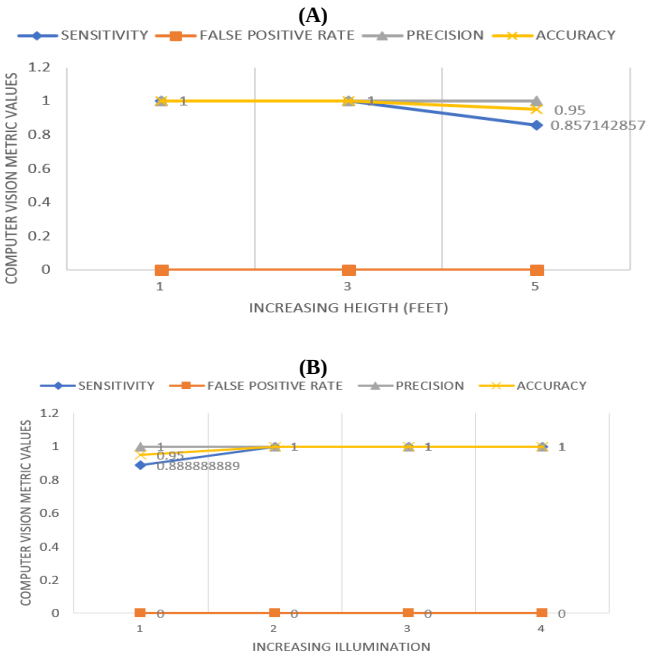


Fig. 10: Object Detection Performance Metrics (A) Increasing Height, (B) Increasing Illumination.

4. Conclusion

This research successfully presented a system for a scale invariant area calculation of lettuce leaf area. The developed algorithm combined machine vision and knowledge-based methods in detecting the marker with a known area for scaling the lettuce leaf area in pixels to area in terms of cm^2. The results of research showed that knowledge-based algorithm can improve the performance of the machine vision classifiers even for a limited number of training datasets. Area measurements performed well in terms of RMSE. This research can be a tool to assist in determining the lettuce canopy area, hence lettuce growth stage which is very useful in the determination of recommended environmental parameters for optimal harvestable products. It is recommended for future work the application of the system to other controlled farm environment such as hydroponics and the determination of relationship of the computed lettuce leaf area to growth stage as affected by other environmental parameters such as nutrients, water and lighting.

Acknowledgment

The authors are grateful for the support of USAID STRIDE, Department of Science and Technology – Philippine Council for Industry, Energy and Emerging Technology Research and Development (DOST-PCIEERD) and Commission on Higher Education (CHED). The authors also appreciate the Intelligent Systems Laboratory (ISL) for the use of their laboratory equipment and project site in the completion of this research.

References

[1] J. R. d. Cruz, R. G. Baldovino, A. A. Bandala and E. P. Dadios, "Water usage optimization of Smart Farm Automated Irrigation System using artificial neural network," *Proceedings of 2017 5th International Conference on Information and Communication Technology (ICoIC7)*, 2017.

[2] D. Escarabajal-Henarejos, J. Molina-Martinez, D. Fernandez-Pacheco and G. Garcia-Mateos, "Methodology for obtaining prediction models of the root depth of lettuce for its application in irrigation automation.," *Agriculture Water Management*, vol. 151, pp. 167-173, 2015. https://doi.org/10.1016/j.agwat.2014.10.012.

[3] A. Rajput, S. Rajput and G. Jha, ""Physiological Parameters Leaf Area Index, Crop Growth Rate, Relative Growth Rate and Net Assimilation Rate of Different Varieties of Rice Grown Under Different Planting Geometries and Depths in SRI,"," *International Journal of Pure and Applied Bioscience*, vol. 5, no. 1, pp. 362-267, 2017. https://doi.org/10.18782/2320-7051.2472.

[4] Q. Xie, W. Huang, D. Liang, P. Chen, C. Wu, G. Yang, J. Zhang, L. Huang and D. Zhang, "Leaf Area Index Estimation Using Vegetation Indices Derived from Airborne Hyperspectral Images in Winter Wheat," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 8, pp. 3586-3594, 2014. https://doi.org/10.1109/JSTARS.2014.2342291.

[5] D. Fernandez-Pacheco, D. Escarabajal-Henarejos, A. Ruiz-Canales, J. Conesa and J. Molina-Martinez, "A digital image-processing-based method for determining the crop coefficient of lettuce crops in the southeast of Spain," *Biosyst. Eng.*, vol. 117, pp. 23-34, 2014. https://doi.org/10.1016/j.biosystemseng.2013.07.014.

[6] P.J.M. Loresco, I.C. Valenzuela, E.P. Dadios, Color Space Analysis Using KNN for Lettuce Crop Stages Identification in Smart Farm Setup, *Proceedings of TENCON 2018 IEEE Region 10 Conference*, 2018.

[7] R. R. Fang Y., "Current and prospective methods for plant disease detection," *Biosensors*, vol. 5, no. 3, pp. 537-561, 2015. https://doi.org/10.3390/bios5030537.

[8] I. C. Valenzuela, J. C. V. Puno, A. A. Bandala, R. G. Baldovino, R. G. Luna, A. L. De ocampo, J. De Cuello and E. P. Dadios, "Quality assessment of lettuce using artificial neural network," *Proceedings of 2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, pp. 1-5, 2017. https://doi.org/10.1016/j.agwat.2014.10.012.

[9] D. Escarabajal-Henarejos, J. Molina-Martinez, D. Fernandez-Pacheco and G. Garcia-Mateos, "Methodology for obtaining prediction models of the root depth of lettuce for its application in irrigation automation.," *Agriculture Water Management*, vol. 151, pp. 167-173, 2015.

[10] P. Srivastava & A. Khare "Content-based image retrieval using scale invariant feature transform and moments" , *Proceedings of 2016 IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering (UPCON)*, pp. 162 – 166, 2016. https://doi.org/10.1109/UPCON.2016.7894645.

[11] C. Shi-Gang, L. Heng, W. Xing-Li, Z. Yong-Li and H. Lin, "Study on segmentation of lettuce image based on morphological reorganization and watershed algorithm," *Proceedings of 2018 Chinese Control and Decision Conference (CCDC)*, 2018. https://doi.org/10.1109/CCDC.2018.8408290.

[12] M. Nehru, S. Padmavathi, "Illumination invariant face detection using viola jones algorithm" *Proceedings of 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp.1-4, 2017. https://doi.org/10.1109/ICACCS.2017.8014571.

[13] Y. Ma, M. Kan, S. Shan, X. Chen, "Hierarchical training for large scale face recognition with few samples per subject" *Proceedings of 2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 2401 – 2405, 2018. https://doi.org/10.1109/ICIP.2018.8451561.

[14] P.J. M, Loresco, A. Africa, "ECG print-out features extraction using spatial-oriented image processing techniques", *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* Vol 10, No 1-5, pp. 15-20, 2018.

[15] S. Ervural and M. Ceylan, "Increasing lesion specificity with fusion of manually and automatically segmented liver MR images," *Proceedings of 2018 26th Signal Processing and Communications Applications Conference (SIU)*, 2018. https://doi.org/10.1109/SIU.2018.8404559.

[16] P.J.M. Loresco, R.Q. Neyra, A. A. Bandala, "Human gesture recognition using computer vision for robot navigation", *Proceedings of 5th International Conference on Communication and Computer Engineering (ICOCOE-2018)*, 2018.

[17] L. Greche, M. Jazouli, N. Es-Sbai, A. Majda, A. Zarghili, "Comparison between euclidean and manhattan distance measure for facial expressions classification", *Proceedings of 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)*, pp. 1-4, 2017. https://doi.org/10.1109/WITS.2017.7934618.