Evolution in Electrical and Electronic Engineering Vol. 3 No. 2 (2022) 286-293

EEEE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

Development of an Algorithm for Size and Weight Measurement of Intact Pineapples

Kelvin Chay[1], Chia Kim Seng[1]*

[1]Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussien Onn Malaysia, 86400, Batu Pahat, Johor MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2022.03.02.033
Received 26 June 2022; Accepted 14 September 2022; Available online 30 October 2022

Abstract: In the past, in order to classify pineapples according to their size and weight, farmers usually measured the size of each pineapple manually, it takes a lot of time and human resources to obtain pineapple information. The objective of this project is to develop an algorithm for size and weight measurement for intact pineapples by using OpenCV and Python as image processing algorithms. In this project, the webcam is used to capture the appearance of pineapples and the computer is used to execute the algorithm by providing reliable measure results. The developed system was able to capture a valid image of the pineapple and accurately estimate the size and predict the weight of the pineapple. The regression analysis will decide the reliability of the prediction. The size and weight measured by the proposed algorithm were strongly correlated with manually measured values. The final phase of the project is to provide a rapid, accurate, and non-destructive method for agriculture and improve work efficiency for farmers. The result shows the highest percentage error of the diameter is 2.77% and the length is 2.84%. For the weight prediction, there is a 94.56% of strong correlation in between the size and weight prediction. Next, hope that this project can reduce the physical strength and time required for farmers to work and thereby reduce the dependence on human resources for agriculture.

Keywords: Size, Weight, Pineapple, Vision System, Image Processing

1. Introduction

From 2016 to 2020, Malaysia's pineapple exports have been on an upward trend, especially local varieties such as MD2, N36, and Josapine, with increasing demand from China and the Middle East. At present, 95% of canned pineapple products are used in the export market and 5% in the domestic market. Fresh pineapples account for 30% of the export market and 70% of the domestic market. In 2020, the pineapple planting area was 17,228 hectares, accounting for 9.5% of the national fruit planting area [1]. The Minister of Agriculture and Food Industries Datuk Seri Ronald Kiandee said in April 2022 that Malaysia produces 350,000 metric tons of pineapples annually and has the potential to expand pineapple export productivity over the next five years. In order to meet the demand for pineapples, improving the technology in the agricultural field is an important goal at present [2]. One of them is to improve the efficiency of detecting the size and weight of pineapples.

**Corresponding author: kschia@uthm.edu.my*

Images represent a significant subset of all measurements made for example in DNA microarrays, astronomical observations, satellite maps, and robotic vision capture. Therefore, to explore these rich data sources, image processing technology is a popular practical technology in the computer field [3]. Image processing algorithms can solve various image tasks, which include digital image detection, image analysis, image reconstruction, image restoration, image enhancement, image data compression, spectral image estimation, and image estimation [4].

To implement the size and maturity index measurement by using an image processing algorithm, OpenCV and Python are the most suitable software tools. OpenCV is a huge open-source library for computer vision, machine learning, and image processing. It supports a wide variety of programming languages like Python, C++, and Java. Next, Python is a rising popular scientific programming language, together with the increasing availability of a large ecosystem of complementary tools, which makes it an ideal environment in which to produce an image processing toolkit [5]. Thus, this study aims to develop an effective system to determine the size and weight of pineapples to improve work efficiency for farmers while saving time and energy.

2. Materials and Methods

2.1 Vision System Setup

The setup of the vision system is as Figure 1. The vision system setup is built by a lighting device and a webcam. First, the lighting devices are set up on the three sides of the sensing area, and the fourth side or last side of the sensing area is to facilitate the placement of the pineapple. Lighting is an important factor that affects the vision system to obtain input, and it directly affects the quality of the input data and the application effect. In this setup, three 20W USB lithium battery bulbs are used and fixed on three sides inside the box. The advantage is that it provides sufficient brightness while also allowing easy setting of the angle of light exposure. The camera in the settings uses a webcam. The webcam is set and fixed at the top of the sensing area, and the shooting angle of the camera is set to be perpendicular to the pineapple.

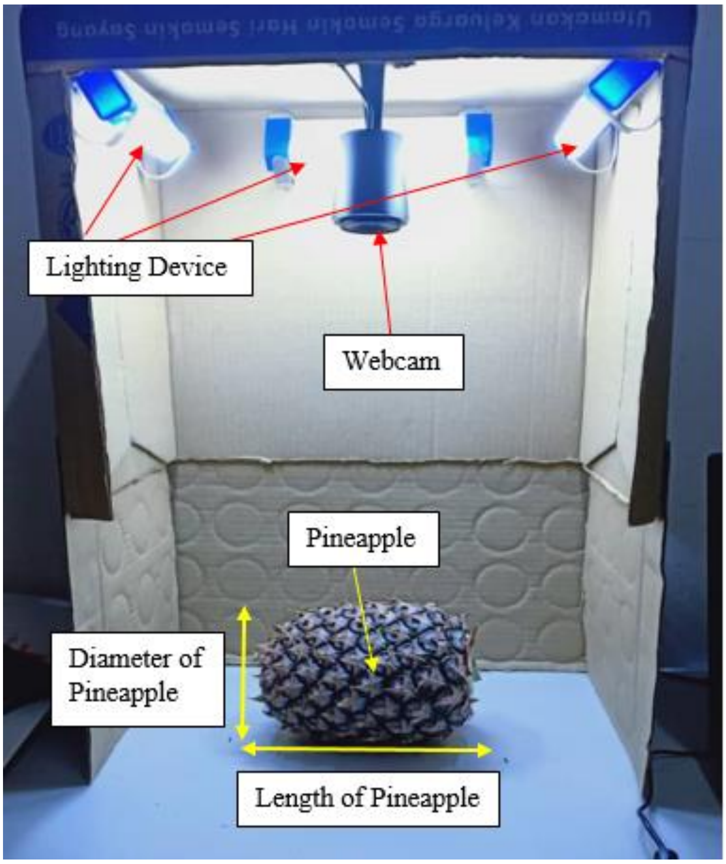

Figure 1: The setup of vision system

2.2 Data Acquisition

First, at the bottom of pineapples without crown measure and mark the angle at 120° and 240° as Figure 2. Next, place the pineapple inside the bottom part of the sensing area and capture it in five different positions. Then manually measure and record the diameter and length of the current pineapple side with a digital Vernier caliper. Next, the pineapple was rotated to 120° to capture the other side of the pineapple in 5 different positions. Then manually measured and recorded the diameter and length of the current pineapple side with a digital Vernier caliper. After that, Next, the pineapple was rotated again to 240° to capture the last side of the pineapple in 5 different positions. Then manually measured and recorded the diameter and length of the current pineapple side with a digital Vernier caliper. Next, manually measured and recorded the weight of the pineapple by using a digital weighing scale. Finally, repeat the above steps with other pineapples until reach 10 sets of the dataset.

Figure 2: Mark the angle at bottom of the pineapples

2.3 Image Processing

Images may contain different types of noise. The technique to reduce the noise in image smoothing. In OpenCV, image smoothing or blurring could be done in many ways. One of the ways is Gaussian Blur. Gaussian filters have the properties of having no overshoot to a step function input while minimizing the rise and fall time. In terms of image processing, any sharp edges in images are smoothed while minimizing too much blurring.

Morphological operations are a set of operations that process images based on shapes. They apply a structuring element to an input image and generate an output image. There are two basic Morphological operations which are Erosion and Dilation. The erosion is used to erode the boundaries of the foreground object and diminish the features of an image. The dilation is used to increase the object area, accentuate features, and is useful in joining the broken parts of an object.

Edges are characterized by sudden changes in pixel intensity. The detection edges are to find the changes in the neighboring pixels. Canny Edge Detection is one of the most popular edge-detection methods because it is robust and flexible. To avoid detecting the unnecessary edge, image blurring is a preprocessing step to reduce noise.

Color conversion from RGB to Grayscale is called gray scaling. It is the process of converting an image from other color spaces for example RGB, CMYK, or HSV to shades of gray. It varies between complete black and complete white. The gray scaling is important for dimension reduction and reduces model complexity. For example, in RGB images there are three color channels and three dimensions while grayscale images are single-dimensional.

2.4 Size Prediction Model

The Size Prediction Model has used the method called pixels per metric ratio. The meaning is the number of pixels contained in each unit indicator. This concept is used as the role of reference. When the size of the reference pineapple on the picture is known, the reference object can be used to convert to the other pineapple picture and predict its size. The size prediction is predicted separately in diameter and length

2.5 Weight Prediction Model

Linear regression is the method used to predict the weight of pineapples when the length and diameter of the pineapple. Linear Regression is the supervised Machine Learning model in which the model finds the best fit linear line between the independent and dependent variable to know the relationship between the dependent and independent variable. Since there are two variables which are length and diameter, so the multiple linear regression will use and its formula is shown as equation 1 to get the fit line to find the relationship. y will be the predicted weight. b_0 is the intercept, b_1 and $b_2$is the coefficient or slope. $x_1$is the length of the pineapple and x_2 is the diameter of the pineapple.

$$y = b_0 + b_1x_1 + b_2x_2 \qquad \text{(Eq. 1)}$$

2.6 Model Optimization

During image processing, consideration must be given to image defects caused by illumination unevenness or artifacts in the imaging path. Use shading correction when a large portion of the image is darker or brighter than the rest of the image, for example, due to misaligned bulbs or poor optics used in the system. The relative difference between the feature of interest and the background is often the same, but a feature in one area of the image has a different grayscale range than a feature of the same type in another part of the image. The main method of shading correction uses a background reference image, whether real or artificial and a polynomial fit of the nearest neighbor pixels [6].

2.7 Performance Evaluation

To evaluate the performance, the result of pineapple was measured in manually compared with the results of the software measurement and got the difference between these two results. Finally, calculate the error of both results. Equation 2 is used to get the error between both results. For the weight prediction, the accuracy of the prediction is based on the correlation coefficient and it will calculate by the python code. If the accuracy or correlation coefficient is more toward 1 then the result will be more reliable.

$$Error\ (\%) = \frac{|\ Calculated - Measured\ |}{Measured} \times 100\% \qquad \text{(Eq. 2)}$$

3. Results and Discussion

3.1 The Size Detection Algorithm

The first part is to read the input image. Next make the image become a blur and to reduce the noise, the higher the value of the parameter the effect of blur will be higher. Then convert the BGR image into grayscale and it can minimize the complexity of the image. Next, detect the contour of pineapple by canny edge detection. The higher the value of canny, the less of the edge will be detected in the image. The using the bounding rectangle box to find the contour of the pineapple. Then get the pixel per metric ratio of the reference image from the length and width of the bounding rectangular box. The length of the bounding rectangular box will be the ratio length of the pineapple and the width will be the diameter of the pineapple. The second part is to read the dataset image. The image processing step is almost similar. After the step for getting the pixel-per-metric ratio of the dataset image, divide the ratio of the dataset image by the ratio of the reference image to get the size. If the dataset detection has not reached 150 sets, repeat the step by inputting other dataset images. All the steps are shown as Figure 3.

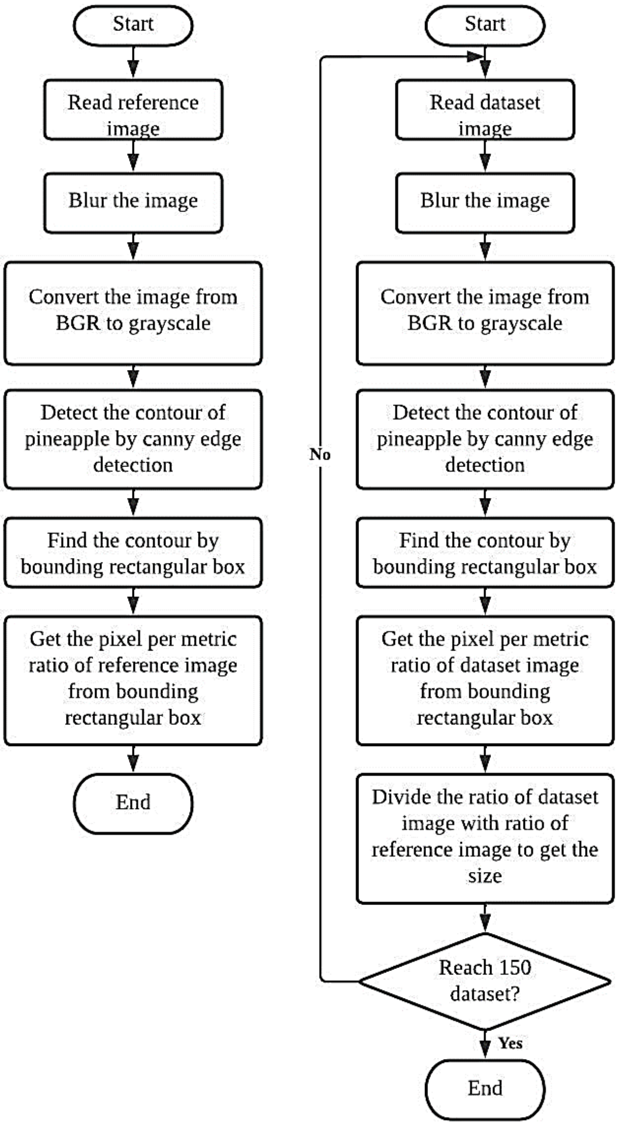

Figure 3: The flowchart of the size detection algorithm

3.2 Reference Image Detection

The reference image detection is using the reference image as Figure 4. The ratio of diameter is 11.9862 and the ratio of length is 13.4462. Both results are detected by the image processing algorithm.

Figure 4: Reference image

3.3 Pineapple data detection for all datasets

Table 1 to Table 3 shows the pineapple image data. The manual data is measured by using Vernier caliper and software data proceeds by using Python. From this table, the highest percentage error of the diameter is image 26 which is 2.39% and the highest percentage error of the length is image 44 which is 2.0%. Based on the data, since just a few of the data have more than a 2% of percentage error, it can assume the result of the software is reliable. These data sets will be the input data for the multiple linear regression.

Table 1: The manual data and software data with percentage error for 1618 g pineapple

Image	Manual Data (cm)		Software data (cm)		Error (%)	
	Diameter	Length	Diameter	Length	Diameter	Length
1	13.00	15.20	13.18	15.10	1.38	0.66
2	13.00	15.20	13.12	15.29	0.92	0.59
3	13.00	15.20	13.03	15.15	0.23	0.33
4	13.00	15.20	13.22	15.31	1.69	0.72
5	13.00	15.20	13.22	15.36	1.69	1.05
6	13.20	15.20	13.18	15.10	0.15	0.66
7	13.20	15.20	13.40	15.16	1.52	0.26
8	13.20	15.20	13.33	14.97	0.98	1.51
9	13.20	15.20	13.51	15.08	2.35	0.79
10	13.20	15.20	13.39	15.04	1.44	1.05
11	13.30	15.10	13.18	15.10	0.90	0.00
12	13.30	15.10	13.40	15.16	0.75	0.40
13	13.30	15.10	13.33	14.97	0.23	0.86
14	13.30	15.10	13.51	15.08	1.58	0.13
15	13.30	15.10	13.39	15.04	0.68	0.40

Table 2: The manual data and software data with percentage error for 1214 g pineapple

Image	Manual Data (cm)		Software data (cm)		Error (%)	
	Diameter	Length	Diameter	Length	Diameter	Length
16	11.40	15.80	11.34	15.76	0.53	0.25
17	11.40	15.80	11.43	15.82	0.26	0.13
18	11.40	15.80	11.26	15.73	1.23	0.44
19	11.40	15.80	11.26	15.83	1.23	0.19
20	11.40	15.80	11.23	15.70	1.49	0.63
21	11.50	15.60	11.57	15.74	0.61	0.90
22	11.50	15.60	11.53	15.69	0.26	0.58

23	11.50	15.60	11.45	15.58	0.43	0.13
24	11.50	15.60	11.56	15.72	0.52	0.77
25	11.50	15.60	11.57	15.70	0.61	0.64
26	11.30	15.60	11.57	15.74	2.39	0.90
27	11.30	15.60	11.27	15.62	0.27	0.13
28	11.30	15.60	11.03	15.50	2.39	0.64
29	11.30	15.60	11.47	15.44	1.50	1.03
30	11.30	15.60	11.45	15.39	1.33	1.35

Table 3: The manual data and software data with percentage error for 644 g pineapple

Image	Manual Data (cm)		Software data (cm)		Error (%)	
	Diameter	Length	Diameter	Length	Diameter	Length
31	10.00	10.60	9.93	10.63	0.70	0.28
32	10.00	10.60	10.05	10.76	0.50	1.51
33	10.00	10.60	10.03	10.62	0.30	0.19
34	10.00	10.60	10.15	10.58	1.50	0.19
35	10.00	10.60	10.18	10.63	1.80	0.28
36	10.00	10.60	9.93	10.49	0.70	1.04
37	10.00	10.60	10.0	10.67	0.00	0.66
38	10.00	10.60	10.06	10.61	0.60	0.09
39	10.00	10.60	10.08	10.55	0.80	0.47
40	10.00	10.60	10.15	10.64	1.50	0.38
41	9.90	10.50	9.68	10.49	2.22	0.10
42	9.90	10.50	9.95	10.46	0.51	0.38
43	9.90	10.50	9.96	10.39	0.61	1.05
44	9.90	10.50	9.71	10.29	1.92	2.00
45	9.90	10.50	9.85	10.38	0.51	1.14

3.4 Weight prediction with multiple linear regression

All the detected diameter and length, and measured weight will be the data of the multiple linear regression. The diameter and length are the independent variables and weight is the dependent variable. By using these two types of variables to predict the weight of pineapple. The b_0 of multiple linear regression generated by python is -2033.7320, b_0 is 179.8374, and b_1 is 79.4787. All the bias can be sub into equation 1 to get the regression line equation as equation 3. Next, the regression analysis has a 94.56 % strong correlation between size and weight prediction. It means that the weight prediction model is reliable. Figure 5 is the 3D graph for multiple linear regression. It shows the independent variable (length and diameter) is on the x and y-axis, and the dependent variable is on the z-axis.

$$y = -2033.7320 + 179.8374x_1 + 79.4787x_2 \qquad \text{(Eq. 3)}$$

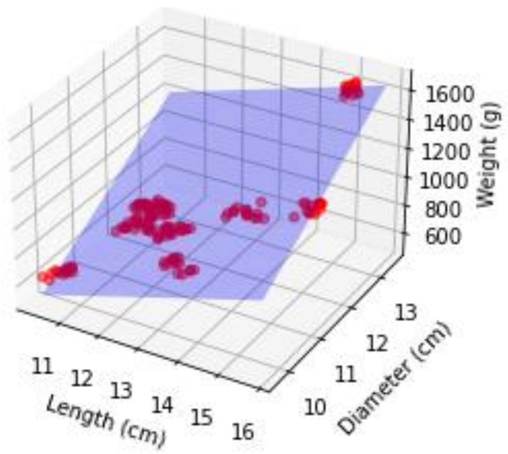

Figure 5: The 3D graph for multiple linear regression.

4. Conclusion

A vision system is developed by using a webcam and three USB lithium battery bulbs to capture the image of pineapples in five different positions. The image processing algorithm is developed by using blur for noise reduction, morphological operations for accentuating image features, colour conversion (from RGB to grayscale) is reduce the image complexity, and edge detection for finding the edge of the pineapple in the image. The results of the software were recorded and compared to the manual results. The size estimation results show that the highest percentage error of the diameter is 2.77% and the highest percentage error of the length is 2.84%. The weight prediction result shows that it has a 94.56% strong correlation between the predicted size and weight.

Acknowledgement

This research was fully supported by the Faculty of Electrical and Electronic Engineering (FKEE) and Universiti Tun Hussein Onn Malaysia (UTHM).

References

[1] Aili Hamzah, A. F., Hamzah, M. H., Che Man, H., Jamali, N. S., Siajam, S. I., & Ismail, M. H. (2021). Recent updates on the conversion of Pineapple Waste (ananas comosus) to value-added products, future perspectives and challenges. Agronomy, 11(11), 2221. https://doi.org/10.3390/agronomy11112221

[2] Siti Rashima, R., Maizura, M., Wan Nur Hafzan, W.M., & Hazzeman, H. (n.d.). Home. Food Research. Retrieved June 3, 2022, from https://www.myfoodresearch.com/

[3] He, D. and Xiong, S., 2021. Image Processing Design and Algorithm Research Based on Cloud Computing. from https://www.hindawi.com/journals/js/2021/9198884/.

[4] Pre-scient.com. 2021. Image Processing Algorithms | Prescient Technologies. from https://www.pre-scient.com/knowledge-center/image|processing/image-processing-algorithms.html

[5] Van der Walt, S., Schönberger, J., Nunez-Iglesias, J., Boulogne, F., Warner, J., Yager, N., Gouillart, E. and Yu, T., 2014. scikit-image: image processing in Python. PeerJ, 2.

[6] Grande, J.C. Principles of Image Analysis. Metallogr. Microstruct. Anal. 1, 227–243 (2012). https://doi.org/10.1007/s13632-012-0037-5