Accuracy of Various Methods to Estimate Volume and Weight of Symmetrical and Non-Symmetrical Fruits using Computer Vision

Hurriyatul Fitriyah

Faculty of Computer Science, Universitas Brawijaya,
Jalan Veteran 8 Malang, Indonesia
E-mail: hfitriyah@ub.ac.id

Abstract. Many researchers have used images to measure the volume and weight of fruits so that the measurement can be done remotely and non-contact. There are various methods for fruit volume estimation based on images, i.e., Basic Shape, Solid of Revolution, Conical Frustum, and Regression. The weight estimation generally uses Regression. This study analyzed the accuracy of these methods. Tests were done by taking images of symmetrical fruits (represented by tangerines) and non-symmetrical fruits (represented by strawberries). The images were processed using segmentation in saturation color space to get binary images. The Regression method used Diameter, Projection Area, and Perimeter as features that were extracted from the binary images. For symmetrical fruits, the best accuracy was obtained with the Linear Regression based on Diameter (LDD), which gave the highest R^2 (0.96 for volume and 0.93 for weight) and the lowest RMSE (5.7 mm^3 for volume and 5.3 gram for volume). For non-symmetrical fruits, the highest accuracy for non-symmetric fruits was given by the Linear Regression based on Diameter (LRD) and Linear Regression based on Area (LRA) with an R^2 of 0.8 for volume and weight. The RMSE for LRD and LRA for strawberries was 3.3 mm^3 for volume and 1.4 grams for weight.

Keywords: *conical frustum; estimation; fruits; regression; solid of revolution.*

1 Introduction

Measuring the physical parameters of fruits, i.e., length, diameter, circumference, volume, and weight, is essential in fruit quality checking. Higher-quality fruits usually have specific measures. There are several standards for fruit quality such as those given by the Organization for Economic Co-operation and Development (OECD) [1] and SNI (Indonesian National Standard) [2]. Fruit ripeness is also detected based on size; unripe fruits are usually smaller compared to ripe fruits. For marketing, transporting and customer satisfaction purposes, a specific measure of fruits must also be met. Different types of fruit can be distinguished based on their physical measures. For example, mandarin oranges are small in size, blood oranges are medium in size, and navel oranges are big in size.

Received June 6th, 2022, Revised September 14th, 2022, Accepted for publication November 21st, 2022.
 ISSN: 2337-5787, DOI: 10.5614/itbj.ict.res.appl.2022.16.3.2

Measuring the parameters is usually performed manually. The weight is measured with a scale, the diameter with a tape measure, and the volume with the water displacement method. These measurements require contact with the fruits, hence, there is a possibility of degrading their quality by touching and applying pressure. The water displacement method also makes the fruit come into contact with water, which may cause it to rot. These classic types of measurement also require the fruits to be taken from their twig. The aforementioned drawbacks of this type of measurement have encouraged researchers to develop non-contact and non-destructive types of measurement. Some of these use electrical capacitance [3] and multi-directional 2D silhouettes to form a 3D wireframe [4], but most of them use 2D imaging as a tool to estimate various physical measures of the fruits. Image-based techniques are preferred since they require only an inexpensive tool, namely a camera, to acquire the data. Cameras are also portable, easy to install, and easy to connect to other devices. Image processing combined with machine learning has been widely applied to measure fruit size [5], classify fruit sizes [6], classify fruit types [7], and detect fruit ripeness [8].

Research on the volume measurement of fruits based on images has been done using many different methods. A very basic method for estimating fruit volume is based on the equation of solid objects [9]. This method has been applied to axisymmetric fruits like apples, sweet lime, lemon, and orange. First, the fruit's shape (circle, ellipse, or parabola) is determined using a cross-section image, after which the volume equation is applied. The estimation showed a high correlation of R^2 0.925. The study by Teerachaichayut, Yokswad, Terdwongworakul, Jannok & Fernandes [10] used the basic equation of sphere volume to estimate the volume of mangosteen fruits. The radius in the formula was an average value from six radii acquired from the XY plane and two radii from the XZ plane. Meanwhile, the study by [11] used the ellipsoid volume formula to estimate the volume of papaya fruits and achieved a high correlation of R^2 0.96 to the actual volume.

Fruits are natural products, hence, their shape is slightly irregular. Rather than using the basic volume equation of solid objects, several researchers have proposed using the Solid of Revolution to accommodate the irregular shape of fruits. In the Solid of Revolution method, the total volume is a summation of its constituted smaller disk volume. This disk is assumed to be a cylinder with a height of 1 pixel. The diameter of each disk is determined by its length in pixels, either of each row or of each column of the fruit image. This method has been proposed by Ibrahim, Ahmad Saad, Zakaria & Md Shakaff [12] to estimate the volume of a regular shaped mango and achieved an R^2 of 0.9985 compared to its actual volume. It is assumed that the disk is more of an oval than a circle, thus the first and second diameter is obtained by taking the image twice, with a side view and a top view. A segmented binary image of the 2D image is used to produce a 3D image. The study by Rashidi, Seyfi & Gholami [13] used this method to estimate the volume of kiwi fruits and showed good results with the T-test and the Balt-Altman method compared to the actual volume obtained using the water displacement method. Although the Solid of Revolution method seems more advanced in volume estimation compared to the basic volume equation, the study by Koc [14] showed that for axisymmetric fruits the estimation error is similar, at 7.8% and 7.7%, respectively.

Solid of Revolution based on the disk method has a drawback because it assumes that the total volume is constituted by smaller cylinders. In reality, the fruit shape is irregular, thus each length has different top and bottom diameters, and left and right diameters. The conical frustum method was proposed to overcome this drawback. This method estimates the total volume as a summation of its constituted smaller conical frustums. The study by Omid, Khojastehnazhand & Tabatabaeefar [15] developed an algorithm to measure the volume of citrus fruits (lemon, lime, orange, tangerine) using this method, achieving an R^2 value of >0.95. The study by Sabliov, Boldor, Keene & Farkas [16] used the Conical Frustum method to estimate the volume of eggs, lime, peaches, and lemons with an error of less than 10%.

While volume estimation uses geometrical equations, mass or weight estimation mostly utilizes a learning method. This is a good approach to use for estimation, especially when there is a non-linear relation between the predicted value and the selected features. It is a black-box method based solely on input-output data. Most existing studies used Regression to estimate the mass or weight of fruits based on the volume that was determined beforehand. This method has been applied to estimate the volume of macaw palm fruits [17] with an R^2 of 0.837. Venkatesh, *et al.* [9] estimated the mass of apple, lemon, sweet lime, and orange fruits from their estimated volume, achieving an R^2 greater than 0.86. The study by Ponce Real, *et al.* [18] used Linear Regression based on the area of olive fruits shown in images to estimate their weight with an error of less than 1.3%. The study by Ashtiani, *et al.* [19] investigated various physical properties of lime fruits for weight estimation. They found that linear regression based on minor diameter, projected area, or estimated volume had an R^2 greater than 0.97.

Font, *et al.*[20] used linear regression not only for estimating individual fruits, but also to estimate the weight of clusters of grapes. The research used 2D images of the grape clusters and used an image segmentation method to obtain the region of the grape clusters as the region of interest (ROI). The weight was estimated using the area of the segmented region, achieving an R^2 of 0.9557. The research also estimated grape clusters based on volume estimation, achieving an R^2 of 0.9635. The volume of the grape clusters was determined using the Solid of Revolution method on the vertical axis. Although several rows that were revolved were not complete due to the irregular silhouette of the grape cluster, the diameter of the revolution was assumed to be from the most left to the most right edge.

The present study proposes an experiment using all the aforementioned methods for the volume and weight estimation of fruits. The methods were tested on axisymmetric fruits (represented by tangerines) and non-axisymmetric fruits (represented by strawberries). Tangerines are nearly round, hence, it is a suitable example of fruits that are axisymmetric from a lateral and a top view. Strawberry has nine highly different types of shapes, i.e., reniform, conical, cordate, ovoid, cylindrical, rhomboid, obloid, globose, and wedged. It also has uniform shapes that are affected by genetics and the environment [21], making this fruit a suitable example of non-axisymmetric fruit. The performance of the estimation was determined using the coefficient of determinant R^2 and RMSE.

2 Materials and Methods

Several different methods have been proposed by researchers to estimate the volume and weight of fruit. This paper explored various methods of volume estimation using the equations of the Basic Solid, the Solid of Revolution, and the Conical Frustum methods. This research also experimented with an alternative to estimate volume using regression straight from the 2D features. The features were taken from the segmented fruit region in 2D images using geometric measures such as diameter (horizontal and vertical), area, and perimeter, while the mass or weight of the fruits was estimated using regression from the determined volume. This research also experimented with an alternative method to estimate the weight of fruits using the Regression directly from geometric measures in order to analyze another possibility. This estimation was also based on 2D images of fruits.

2.1 Image Acquisition

Images of fruits were taken only once, with a frontal view. The fruits and the camera were placed at a fixed distance of 30 cm to maintain a pixel-to-centimeter ratio of 0.28. The image acquisition was performed in constant lighting conditions. The background of the fruit was white to simplify the segmentation process. The tangerines were placed in a position where the stalk was at the top. The strawberries were placed at rest, where the stalks or calyx were in a horizontal position. A common CCD camera was used to take the images.

The stalk and calyx were removed since the estimation was meant to be only of the main part of the fruits. The individual tangerines and strawberries had different sizes. Some images are shown in Figure 1. The strawberries had many different shapes and mostly were not axisymmetric. The fruits varied from unripe to very ripe to accommodate various sizes and colors of fruits.

Figure 1 Several images of tangerines and strawberries from the dataset.

2.2 Image Processing

The image processing step aims to obtain only the region that depicts the fruit. The background of the images was set to white since the saturation of the color white is very low compared to any other color. The saturation shows the impurity of a color. It has the highest value for a solid color and the lowest value for a greyish color. The images were originally in RGB color and were converted to the saturation color using Eq. (1). The values of R, G, and B are the intensities of each pixel in the red, green, and blue color space, with values of 0-255.

$$S = \begin{cases} 0, & C_{max} = 0 \\ \frac{C_{max} - C_{min}}{C_{max}}, & C_{max} \neq 0 \end{cases} \quad (1)$$

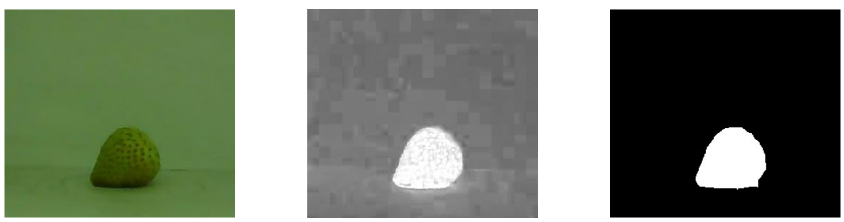
where $C_{max} = \max\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right)$ and $C_{min} = \min\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right)$

The saturation intensities have values between 0 and 1. Since the saturation of white is easily distinguishable, the segmentation process simply used Otsu's thresholding method. Otsu's thresholding method finds a certain threshold and puts the intensities into two classes: foreground and background. The threshold has the highest between-class variance and at the same time the lowest within-class variance between the two classes. Once the threshold has been found, the segmentation sets a value of 0 for pixels with saturation intensities below the threshold, and a value of 1 for pixels with saturation intensities above the threshold.

The segmentation produces a binary image, where fruits as the region of interest have a value of 1 and the background has a value of 0. The pseudocode for the image processing step is shown in Algorithm 1. An example of the image processing step applied to an RGB image to create a binary image is shown in Figure 2.

Algorithm 1 Image Processing Step
Input: RGB image
Output: Binary image

```
[m,n] size of RGB image
    for i equal to 1 until m
        for j equal to 1 until n
            convert RGB (i,j) into Saturation (i,j)
        end
    end
set threshold of segmentation
    for i equal to 1 until m
        for j equal to 1 until n
            if Saturation (i,j) < threshold
                Binary image (i,j) equal 0
            else
                Binary image (i,j) equal 1
             end
        end
    end
```



(a) (b) (c)

Figure 2 Image processing step: (a) RGB image, (b) Saturation image, (c) segmented image.

2.3 Geometric Measure Extraction

Simple geometric measures were extracted from the binary images. These measures will be used by the learning method and not by the solid of revolution method or the conical frustum method. The geometric measures consisted of (1) the horizontal length of the bounding box, (2) the vertical length of the bounding box, (3) the area, and (4) the perimeter. In this study, the horizontal length of the bounding box was determined by taking the maximum coordinate of pixel 1 subtracted by the lowest coordinate of pixel 1 in the same row for the horizontal length and the same column for the vertical length.

The area was determined by measuring the total number of white pixels in the fruit region; this is the area with the largest frontal cross-section of the fruit and not the surface area. The largest frontal cross-section is shown in the binary image as the fruit region. The perimeter is seldom used in similar research, but it was explored in this study as another possible feature, especially for the learning method. The perimeter was defined as the length of the outer shell of the largest frontal cross-section. The perimeter was determined by applying edge detection to the binary image. Once the edges were detected, they were given a value of 1 and 0 to the rest of the image. The perimeter was defined as the total number of 1. An illustration of the four geometric measures is shown in Figure 3. The algorithm to find the geometric measures is given in Algorithm 2.

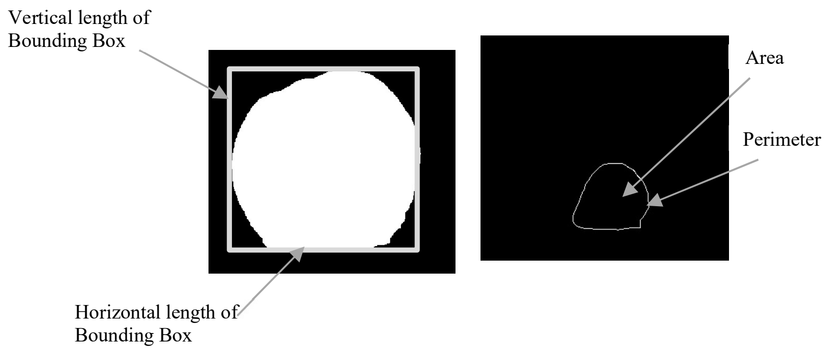


Figure 3 Geometric measures in the binary image of a strawberry.

Algorithm 2 Geometric Measures Extraction
Input: Binary Image
Output: Vertical length, horizontal length, area, perimeter

1: (x,y) is the indices of value 1 in the Binary Image
2: Vertical length of Bounding Box = maximum(x) – minimum(x)
3: Horizontal length of Bounding Box = maximum(y) – minimum(y)
4: Area = number of x
5: Find edges using Edge Detection, set edges to 1 and else to 0 in the Edge Image
6: Perimeter = sum of 1 in Edge Image

2.4 Basic Solid Method for Volume Estimation

The Basic Solid method uses the basic volume equation of a sphere or an ellipsoid. This method assumes fruits are axisymmetric and have a round or elliptical shape. Given a known diameter, d, the volume of a sphere and an ellipse is shown in Eqs. (2) and (3), respectively. In this paper, the diameter of the ellipse was taken from the horizontal and vertical length of the fruit's bounding box while the diameter of the sphere was taken from the longest diameter between the two. An ellipsoid usually has three different diameters on the XYZ-axis. Since the volume estimation in this study only had two diameters from the frontal view image, the equation shown in Eq. (3) was used, where d_1 is the shorter diameter and d_2 is the longer diameter. The algorithm to estimate the volume based on the Basic Solid method is given in Algorithm 3.

$$V_{sphe} = \frac{1}{6}\pi d^3 \tag{2}$$

$$V_{ellipsoid} = \frac{1}{6}\pi {d_1}^2 d_2 \tag{3}$$

Algorithm 3 Volume Estimation based on Basic Solid method
Input: Vertical length, Horizontal length
Output: Volume of Basic Solid Sphere (BSS), Volume of Basic Solid Ellipsoid (BSE)

If Vertical Length > Horizontal Length
    Volume of BSS = $\pi/6 \times$ (Vertical length)^3
    Volume of BSE = $\pi/6 \times$ (Horizontal Length)^2 $\times$ Vertical Lenght
else
    Volume of BSS = $\pi/6 \times$ (Horizontal length)^3
    Volume of BSE = $\pi/6 \times$ (Vertical Length)^2 $\times$ Horizontal Lenght
end

2.5 Solid of Revolution Method for Volume Estimation

The Solid of Revolution method estimates the volume of an object as the summation of the volumes of the total number n of smaller disks. The volume of each disk is defined as the area of a circle, A_i, times its length, Δx_i, where A_i is determined based on the diameter, d_i, at certain coordinates, i. In this study, the diameter was defined as the maximum coordinate of pixel 1 subtracted by the minimum coordinate of pixel 1 in one row for horizontal revolution and one column for vertical revolution. The Solid of Revolution formula is given in Eq. (4). The illustration of components in the Solid of Revolution method for the vertical revolution is shown in Figure 4. The algorithm to find the volume based on the Solid of the Revolution method is given in Algorithm 4.

$$V = \sum_{i=1}^{n} A_i \cdot \Delta x = \sum_{i=1}^{n} \frac{\pi}{4} {d_i}^2 \cdot \Delta x \tag{4}$$

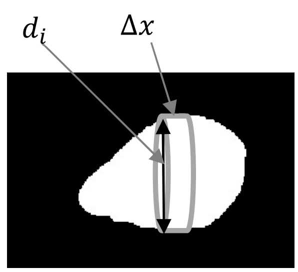


Figure 4 Illustration of the components in the solid of revolution method.

Algorithm 4 Volume based on the Solid of Revolution method
Input: Binary Image
Output: Volume of Solid Revolution Horizontal (SRH) and Vertical (SRV)

(x,y) is the indices of value 1 in the Binary Image
  for i equal to 1 until x
    Volume of SRH (i) = $\pi/4 \times$ (maximum y in row i – minimum y in row i)^2
  end
  Volume of SRH = sum of volume SRH of all i
  for j equal to 1 until y
    Volume of SRV (j) = $\pi/4 \times$ (maximum x in coloumn j – minimum x in coloumn j)^2
  end
  Volume of SRV = sum of Volume SRV of all j

2.6 Conical Frustum Method for Volume Estimation

The Conical Frustum method estimates the volume as the summation of the volumes of the total number n of smaller frustums. Instead of using the diameter of one side to calculate the volume of the disk as in the Solid of Revolution method, the Volume of Frustum uses the diameters of both sides, d_i and $d_{(i+1)}$. The diameter is defined as the maximum coordinate of pixel 1 subtracted by the minimum coordinate of pixel 1 in one row for the horizontal frustum and one column for the vertical frustum.

The formula for The Conical Frustum method is given in Eq. (5). An illustration of the components in the Conical Frustum method on the horizontal axis is shown in Figure 5. The algorithm to find the volume based on the Conical Frustum method is given in Algorithm 5.

$$V = \sum_{i=1}^{n} \frac{\pi}{12}\left({d_i}^2 + {d_{i+1}}^2 + (d_i \cdot d_{i+1})^2\right) \tag{5}$$

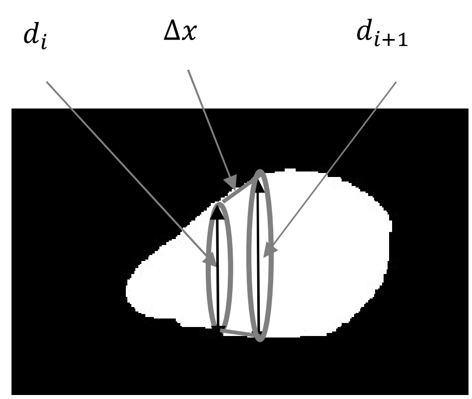


Figure 5 Illustration of the components in the conical frustum method.

Algorithm 5 Volume based on the Conical Frustum method
Input: Binary Image
Output: Volume of Conical Frustum Horizontal (CFH) and Vertical (CFV)

(x,y) is the indices of value 1 in the Binary Image
  for i equal to 1 until x-1
    diameter_i (i) = maximum y in row i – minimum y in row i
    diameter_i+1 (i) = maximum y in row (i+1) – minimum y in row (i+1)
    volume of CFH (i) = $\pi/12$ × ((diameter_i)^2 + (diameter_i+1)^2 + (diameter_i × diameter_1+1)^2)
  end
  Volume of CFH = sum of volume CFH of all i
  for j equal to 1 until y-1
    diameter_i (i) = maximum x in coloumn i – minimum x in coloumn i
    diameter_i+1 (i) = maximum x in coloumn (i+1) – minimum x in coloumn (i+1)
    volume of CFV (i) = $\pi/12$ × ((diameter_i)^2 + (diameter_i+1)^2 + (diameter_i × diameter_1+1)^2)
  end
  Volume of CFV = sum of volume CFV of all i

2.7 Learning Method for Volume and Weight Estimation: Regression

Regression is a method that is used to predict outcomes based on data learning. It learns the relation between previous independent-dependent data pairs that can be linear, logarithmic, exponential, or otherwise. In Linear Regression, the relation is shown as a line with a specific coefficient. Linear Regression can be applied to univariate data and multivariate data.

The formula for Multiple Linear Regression is shown in Eq. (6), where Y denotes a response (dependent) variable, X denotes a predictor (independent) variable, β denotes the vector of the coefficient to be estimated, and ε denotes an independent random variable. Coefficient β is estimated using the method of least square, as shown in Eq. (7). Given a predictor in data testing, $\hat{X}$, the response $\hat{Y}$ is determined using Eq. (8).

$$Y = \beta X + \varepsilon \tag{6}$$

$$\beta = (X^T X)^{-1} X^T Y \tag{7}$$

$$\hat{Y} = \hat{X}\beta \tag{8}$$

The linear Regression method is used to estimate the volume and weight of the fruits as the dependent variables. The dependent variables are the geometric measures, i.e., length, area, and perimeter. The volume is only estimated by the Linear Regression method since it is a non-spatial measure. The algorithm to find the volume and weight using the Linear regression methos is given in Algorithm 6.

Algorithm 6 Volume and weight estimation using Linear Regression method
Input: Vertical length, Horizontal Length, Area, Perimeter
Output: Volume and weight based on Diameter, Area, Perimeter, and all their sub-sets (LRD, LRA, LRP, LRDA, LRDP, LRAP, LRDAP)

1: Determine diameter as the longest between Vertical Length and Horizontal Length
2: Calculate vector of Coefficient for Volume based on Diameter, Area, Perimeter, Diameter-Area, Diameter-Perimeter, Area-Perimeter, Diameter-Area-Perimeter
3: Calculate Volume based on the coefficients of each Diameter, Area, Perimeter, Diameter-Area, Diameter-Perimeter, Area-Perimeter, Diameter-Area-Perimeter
4: Calculate vector of Coefficient for Weight based on Diameter, Area, Perimeter, Diameter-Area, Diameter-Perimeter, Area-Perimeter, Diameter-Area-Perimeter
5: Calculate Weight based on the coefficients of each Diameter, Area, Perimeter, Diameter-Area, Diameter-Perimeter, Area-Perimeter, Diameter-Area-Perimeter

3 Result and Discussion

The aforementioned methods to estimate fruit volume and weight were tested on tangerines and strawberries. The tangerines were chosen as an example of axisymmetric fruits, while the strawberries were chosen as an example of non-axisymmetric fruits. A total of 62 images of tangerines and 62 images of strawberries were acquired. The fruits had different ripeness to accommodate different sizes of tangerines and strawberries. The properties of the fruits' measures, i.e., range, mean and standard deviation, are shown in Table 1 for the tangerines and in Table 2 for the strawberries.

Table 1 Measures of the tangerine data set.

Properties	Volume (mm^3)	Weight (gr)
Range	131.00	111.42
Mean	88.65	81.99
Standard deviation	26.83	21.91

Table 2 Measures of the strawberry data set

Properties	Volume (mm^3)	Weight (gr)
Range	18.00	14.67
Mean	9.81	5.70
Standard deviation	5.06	2.92

The estimated (observed) volume of the fruits was compared to the actual (expected) volume that was measured using the Water Displacement Method (WDM). The estimated volume was determined using: (1) The Basic Solid Sphere method (BSS), (2) The Basic Solid Ellipsoid method (BSE), (3) The Solid of Revolution method (SR), (4) The Conical Frustum method (CF). This research also tested the Linear Regression method to estimate the volume. The regression used all combinations of the three parameters that could be taken from the fruits' 2D images as features, i.e., diameter, area, and perimeter. The methods used to estimate the fruits' volume are (1) The Linear Regression method using the

Diameter (LRD), (2) The Linear Regression method using the Area (LRA), (3) The Linear Regression method using the Perimeter (LRP), (4) The Linear Regression method using The Diameter and Area (LRDA), (5) The Linear Regression method using The Diameter and Perimeter (LRDP), (6) The Linear Regression method using The Area and Perimeter (LRAP), and (7) The Linear Regression method using The Diameter, Area and Perimeter (LRDAP). The total number of methods that were evaluated in this research was eleven.

The estimated weight was compared to the actual weight that was measured on an electronic milligram scale. The estimated weight was determined using Linear Regression. It used all combinations of the three parameters diameter, area, and perimeter as features. In total there were seven methods that were evaluated in this research are 7 which are: (1) The Linear Regression method using the Diameter (LRD), (2) The Linear Regression method using the Area (LRA), (3) The Linear Regression method using the Perimeter (LRP), (4) The Linear Regression method using the Diameter and Area (LRDA), (5) The Linear Regression method using the Diameter and Perimeter (LRDP), (6) The Linear Regression method using the Area and Perimeter (LRAP), (7) Linear Regression method using the Diameter, Area, and Perimeter (LRDAP).

The first performance analysis metric was the same as the one used in other studies, i.e., the Coefficient of Determinant, R^2. The result is shown in Table 3 for the volume estimation and Table 4 for the weight estimation. In the Linear Regression method, a total of 62 data was randomly divided into 40 training data and 22 testing data (ratio ± 70:30). The R^2 in the Linear Regression was calculated only on the testing data.

Table 3 R^2 for fruit volume estimation.

Fruits	BSS	BSE	SRH	SRV	CFH	CFV	
Tangerines	0.93	0.92	0.96	0.94	0.96	0.94	
	LRD	**LRA**	**LRP**	**LRDA**	**LRDP**	**LRAP**	**LRDAP**
	0.96	0.93	0.92	0.95	0.95	0.93	0.95
Strawberries	**BSS**	**BSE**	**SRH**	**SRV**	**CFH**	**CFV**	
	0.56	0.59	0.00	0.00	0.00	0.00	
	LRD	**LRA**	**LRP**	**LRDA**	**LRDP**	**LRAP**	**LRDAP**
	0.80	0.79	0.77	0.80	0.80	0.80	0.80

Table 4 R^2 for fruit weight estimation.

Fruits	LRD	LRA	LRP	LRDA	LRDP	LRAP	LRDAP
Tangerines	0.93	0.92	0.90	0.93	0.93	0.91	0.93
Strawberries	0.81	0.76	0.71	0.81	0.81	0.75	0.81

As shown in Table 3 for the volume estimation, the tangerines as an example of an axisymmetric fruit had a high R^2 (>0.91) for every method. The strawberries as an example of a non-axisymmetric fruit had a lower R^2. The BSS and BSE method still had considerably good results with an R^2 of 0.56 and 0.59, respectively. The SRH/SRV and the CFH/CFV had very poor volume estimations with R^2 near 0. A better R^2 on the volume estimation of strawberries was obtained by the Regression method. The LRD, LRA, LRP, LRDA, LRDP, LRAP, and LRDAP had a good R^2 for the volume estimation, in the range of 0.77 to 0.80. For the weight estimation, shown in Table 4, the Linear Regression method obtained a similar result. The weight estimation of tangerines had a high R^2 (>0.90), while strawberries had a lower but still good R^2 (> 0.71).

The list of R^2 in Tables 3 and 4 shows that all methods had high accuracy for both the volume and weight estimation of axisymmetric fruits. In the case of non-axisymmetric fruits such as strawberries, the volume and weight estimations were better during when using Linear Regression. BSS, BSE, SRH/SRV, and CFH/CFV cannot be used to estimate the volume and weight of non-axisymmetric fruits since they obtained a very low R^2.

The second analysis was error calculation. The most common method to calculate the error between two or more measurements is Root Mean Squared Error (RMSE). The RMSE for the volume estimation is shown in Table 5, while the weight estimation is shown in Table 6. The RMSE for volume estimation is expressed in mm^3.

Table 5 RMSE for volume estimation in mm^3.

RMSE of Volume Estimation	**BSS**	**BSE**	**SRH**	**SRV**	**CFH**	**CFV**
Tangerine	19.00	7.84	12.88	7.52	12.87	7.52
Strawberry	5.89	4.30	4.07	4.99	4.08	4.99

RMSE of Volume Estimation	**LRD**	**LRA**	**LRP**	**LRDA**	**LRDP**	**LRAP**	**LRDAP**
Tangerine	5.71	8.29	7.51	61.87	27.05	180.96	20.03
Strawberry	3.36	3.26	3.65	13.55	3.56	47.38	13.22

Table 6 RMSE for weight estimation in grams.

RMSE of Weight Estimation	**LRD**	**LRA**	**LRP**	**LRDA**	**LRDP**	**LRAP**	**LRDAP**
Tangerine	5.30	6.05	6.60	14.79	18.44	116.55	18.81
Strawberry	1.33	1.47	1.98	7.13	4.01	23.54	6.27

The RMSEs for the volume estimation of the tangerines (axisymmetric fruit) showed a similar trend to the R^2 analysis. The learning method, i.e., Linear Regression, performed better than BSS, BSE, SRH/SRV, and CFH/CFV. From Table 5 it can be seen that the lowest RMSE for the volume estimation of the tangerines was achieved by Linear Regression based on the Diameter (LRD). It also obtained the lowest RMSE in the weight estimation, as shown in Table 6. Thus, this research suggests that the diameter is the best geometric feature to use for volume and weight estimation of axisymmetric fruit when using only one side of the fruit in image acquisition.

The RMSEs for the volume estimation of the strawberries (non-axisymmetric fruit) also showed a similar trend to the R^2 analysis. Linear Regression based on the Diameter (LRD) and the Linear Regression based on the Area (LRA) obtained the best results, both for the volume and the weight estimation. Thus, this research suggests that diameter or area are the best geometric features to use for volume and weight estimation of non-axisymmetric fruits when using only one side of the fruit in image acquisition.

Overall, the volume estimation based on learning had higher accuracy compared to the other methods. Usually adding more parameters in the Linear Regression gives higher accuracy, but this was not confirmed by the result of this study. Combining two to three parameters gave lower accuracies in both the volume and the weight estimation.

4 Conclusion

This research performed experiments and analyses on the accuracy of the volume and weight estimation of axisymmetric and non-axisymmetric fruit based on frontal image acquisition using various methods. The learning method based on Linear Regression showed better results compared to the geometric method (the Solid of Revolution and the Conical Frustum method). This research suggests that for the axisymmetric fruits, the Linear Regression based on the Diameter (LRD) gave the highest accuracy, both in volume and weight estimation. Meanwhile, for the non-axisymmetric fruits, the Linear Regression based on the Diameter (LRD) or the Linear Regression based on the Area (LRA) gave the highest accuracy in the volume and the weight estimation. The accuracy for non-axisymmetric fruits was moderate and needs further study, probably by acquiring multiple images from various sides. Further research should also study more types of fruit and come up with regression equations for each fruit to estimate their volume and weight.

References

[1] OECD, *International Standards for Fruit and Vegetables*, Paris, OECD Publishing, 2011.

[2] National Standardization Agency of Indonesia, *SNI Tangerine,* 3165:2009, 2009. (Text in Indonesian)

[3] Jarimopas, B., Nunak, T. & Nunak, N., *Electronic Device for Measuring Volume of Selected Fruit and Vegetables*, Postharvest Biology and Technology, **35**(1), pp. 25-31, Jan. 2005.

[4] Concha-Meyer, A., Eifert, J., Wang, H. & Sanglay, G., *Volume Estimation of Strawberries, Mushrooms and Tomatoes with a Machine Vision System*, International Journal of Food Properties, **21**(1), pp. 1867-1874, Aug. 2018.

[5] Hu, M.-H., Dong, Q.-L., Malakar, P.K., Liu, B.L. & Jaganathan, G.K., *Determining Banana Size Based on Computer Vision*, International Journal of Food Properties, **18**(3), pp. 508-520, Dec. 2015.

[6] de Luna, R., Dadios, E.P., Bandala, A.A. & Vicerra, R.R.P., *Size Classification of Tomato Fruit Using Thresholding, Machine Learning, and Deep Learning Techniques*, Agrivita, Journal of Agricultural ScienceAgrivita, **41**(3), Oct. 2019.

[7] Ercisli, S., Sayinci, B., Kara, M., Yildiz, C. & Ozturk, I., *Determination of Size and Shape Features of Walnut (Juglans Regia L.) Cultivars Using Image Processing*, Scientia Horticulturae, **133**, pp. 47-55, Jan. 2012.

[8] Bargoti, S. & Underwood, J.P., *Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards*, Journal of Field Robotic, **34**(6), pp. 1039-1060, Oct. 2017.

[9] Venkatesh, G.V., Md. Iqbal, S., Gopal, A. & Ganesan, D., *Estimation of Volume and Mass of Axi-symmetric Fruits using Image Processing Technique*, International Journal of Food Properties, **18**(3), pp. 608-626, Dec. 2014.

[10] Teerachaichayut, S., Yokswad, S., Terdwongworakul, A., Jannok, P. & Fernandes, S.V., *Application of Image Analysis for Determination of Mangosteen Density*, Journal of Advanced Agricultural Technologies, **2**(2), pp. 92-97, Dec. 2015.

[11] Cortes, D.F.M., Catarina, R.S., Barros, G.B. d. A., Arêdes, F.A.S., da Silveira S.F., Ferreguetti, G.A., Ramos, H.C.C., Viana, A.P. & Pereiral, M.G., *Model-Assisted Phenotyping by Digital Images in Papaya Breeding Program*, Scientia Agricola, **74**(4), pp. 294-302, Aug. 2017.

[12] Ibrahim, M.F., Ahmad Saad, F., Zakaria, A. & md Shakaff, A.Y., *In-Line Sorting of Harumanis Mango based on External Quality using Visible Imaging*, Sensors, **16**(11), Oct. 2016.

[13] Rashidi, M., Seyfi, K. & Gholami, M., *Determination of Kiwi Fruit Volume Using Image Processing*, ARPN Journal of Agricultural and Biological Science, **2**(6), pp. 17-22, Jan. 2007.

[14] Koc, A.B., *Determination of Watermelon Volume Using Ellipsoid Approximation and Image Processing*, Postharvest Biology, and Technology, **45**(3), pp. 366-371, Jan. 2007.
[15] Omid, M., Khojastehnazhand, M. & Tabatabaeefar, A., *Estimating Volume and Mass of Citrus Fruits by Image Processing Technique*, Journal of Food Engineering, **100**(1), pp. 315-321, Sept. 2010.
[16] Sabliov, C.M., Boldor, D., Keener, K.M. & Farkas, B.E., *Image Processing Method to Determine Surface Area and Volume of Axi-symmetric Aggricultural Products*, International Journal of Food Properties, **5**(3), pp. 641-653, Feb. 2002.
[17] Costa, A.G., Elisângela, R., Braga, R.A. & Pinto, F.A., *Measurement of Volume of Macaw Palm Fruit Using Traditional and the Digital Moiré Techniques*, Revista Brasileira de Engenharia Agrícola e Ambiental, **20**(2), pp. 152-157, Feb. 2016.
[18] Ponce Real, J., Aquino, A., Millan, B. & Andujar Marquez, J., *Automatic Counting and Individual Size and Mass Estimation of Olive-Fruits through Computer Vision Techniques*, IEEE Access, **7**(1), pp. 59451-59465, May. 2019.
[19] Ashtiani, S.-H.M., Motie, J.B., Emadi, B. & Aghkhani, M.-H. *Models for Predicting the Mass of Lime Fruits by Some Engineering Properties*, Journal of Food Science Technology, **51**(11), pp. 3411-3417, Nov. 2014.
[20] Font, D., Tresanchez, M., Martínez, D., Moreno, J., Clotet, E. & Palacin, J., *Vineyard Yield Estimation Based on the Analysis of High-Resolution Images Obtained with Artificial Illumination at Night*, Sensors, **15**(4), pp. 8284-8301, Apr. 2015.
[21] Li, B., Cockerton, H.M., Johnson, A.W., Karlström, A., Stavridou, E., Deakin, G. & Harrison, R. J. *Defining Strawberry Shape Uniformity Using 3D Imaging and Genetic Mapping*. Horticulture Research, **7**(115). Aug. 2020.