

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Multimodal Classification of Mangoes

---

Son V.T. Dao

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.81356>

---

## Abstract

Grading, sorting, and classification of agricultural products are important steps to ensure a profitable and sustainable food industry. Human-intensive labors are replaced with better devices/machines that can be used in-line and generate sufficiently fast measurements for a high production volume. Most previous works focused on only one of the external quality parameters, such as color, size, mass, shape, and defects. In this work, we proposed an integrated machine vision system that can grade, sort, and classify mangoes using multiple features including weight, size, and external defects. We found that weight estimation using our proposed algorithm based on visual information was not statistically different from that of a conventional weight measurement using a static digital load cell; the estimation error is relatively small (4–5%). We also constructed an artificial neural network model to classify mango having multiple types of external defect; the classification error is less than 8% for the worst possible case. The results indicate that our system shows a great potential to be used in a real industrial setting. Future work will aim to investigate other features such as ripeness and bruises to increase the effectiveness and practicality of the system.

**Keywords:** mango, image processing, machine vision, weighing system, defect classification, real time, neural network

---

## 1. Introduction

Food standards are evolving both to ensure the sustainability of agriculture and to satisfy consumer needs. The reputation of producers and consequently their market share is based on the quality of the product, which makes quality controls very crucial. The market together with ever increasing social concerns about good agricultural practices, including environmental,

economic, and social sustainability and traceability, require guarantees of high quality from the earliest stages of the crop to postharvest storage and treatments.

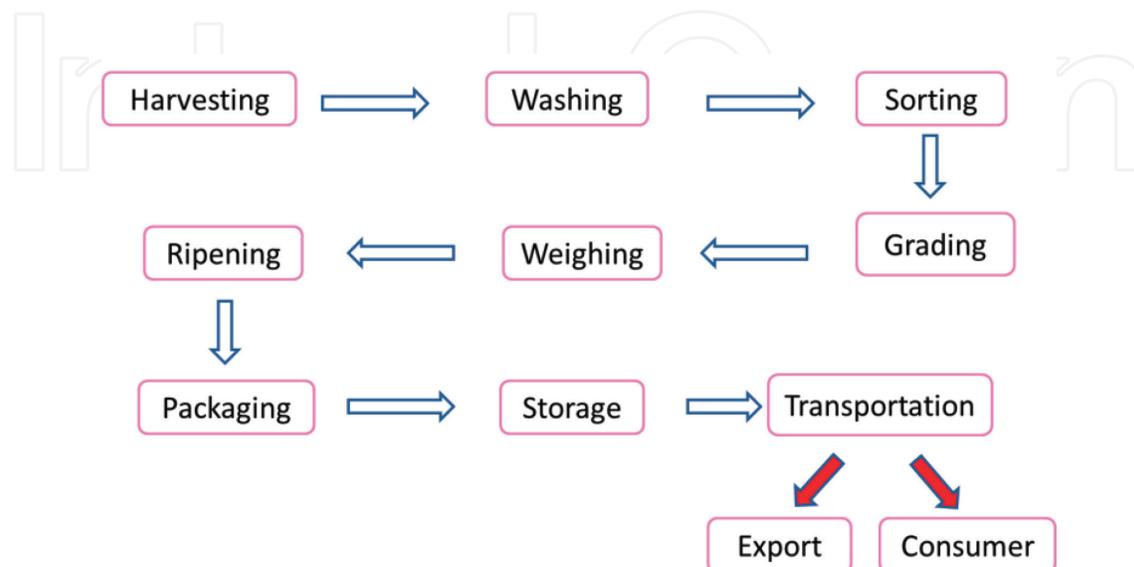
Optical sensors have been used extensively in the industry ranging from the automatic sorting of products into categories to the control of processes which are difficult to observe, for instance, because of their long duration [1]. At this point it is important to note that the quality of biological products is not easy to assess, as individuals of the same category may differ greatly from one to another in terms of color, shape, or size. Furthermore, because they are living products, their physiochemical properties evolve over time. Their inherent variability sometimes introduces a certain amount of subjectivity into quality control, thus increasing the difficulty involved in developing automated inspection systems. Addressing these challenges often requires research in advanced and multidisciplinary technologies and sometimes the use of expensive equipment.

In this work, our focus is on mango, *Mangifera indica*, especially the Cat-Chu cultivar due to the increasing export potential in our country – Vietnam.

Postharvest handling of mangoes is usually completed in several steps: washing, sorting, grading, packing, storage, and transportation as shown in the following **Figure 1**; among which, sorting and grading are considered the most important especially for fresh agricultural products.

Sorting of agricultural products is accomplished based on external quality parameters such as color, defects, shape, and sizes. Manual sorting is based on traditional visual quality inspection performed by trained human operators situated on one or both sides of a conveyor belt. They visually inspect the produce and remove those not satisfying the predetermined quality standards. Pieces are transported slowly enough to allow the workers to inspect all of them and even manipulate them to ensure the inspection of most of their surface. The process is normally tedious, time-consuming, subjective, slow, expensive, and nonconsistent. A cost-effective, consistent, faster, and accurate sorting can be achieved with a machine vision-assisted sorting.

In this work, we present an integrated machine vision-based inspection system including sorting, grading, and weighing of mangoes – particularly, the Cat-Chu cultivar.



**Figure 1.** Typical postharvest steps of agricultural products.

## 2. Our research

### 2.1. Mass estimation

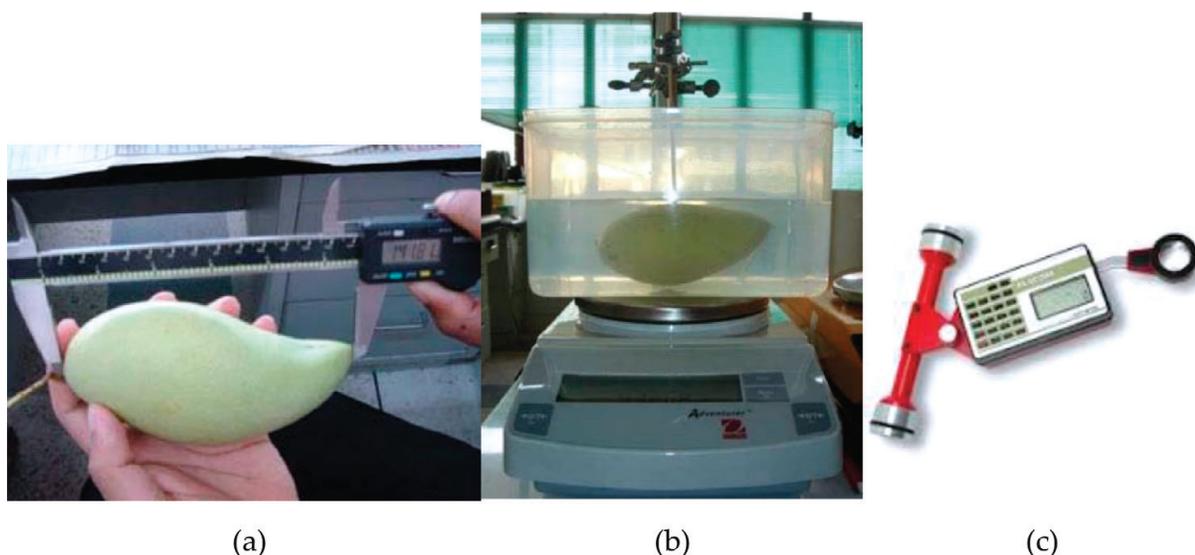
Consumers usually prefer fruits having almost uniform masses and shapes. This is also one of the requirements for export. However, one cannot easily model mango shapes which are not round or oval-shaped. Commonly accepted laboratory instruments are shown in **Figure 2** including a Vernier caliper for size/length measurements, a water replacement measurement setup to estimate volumes, and a planimeter to calculate areas. These methods are time-consuming and not suitable to be implemented into a real production line.

Several attempts have tried to formulate a relationship between mangoes' masses and their sizes [2–5]. Guzman-Estrada et al. [2] used a set of complicated geometrical parameters to estimate the mass of mangoes; most of the parameters can only be obtained using a mechanical measurement tool. Vasquez-Caicedo et al. [3] tried to use five parameters such as length, width, and thickness at maximum width and minimum width to estimate mango weight. Yimyam et al. [4] used four digital photographs to produce a three-dimensional model of Nam-Dokmai mangoes. Most of these methods did not provide easy-to-obtain parameters, except for Spreer et al. [5]; they provide an experimental weight-size correlation based on just three parameters—Length, Max Width (W), and Max Thickness (T) for a specific mango cultivar (Chok Anan) from Thailand.

The weight estimation method using Spreer's method is shown as follows:

$$\text{Estimated mass\_ChoknanMango (grams)} = 5.39 \times 10^{-4} \times L(\text{mm}) \times W(\text{mm}) \times T(\text{mm}) \quad (1)$$

In this work, we will try to use Spreer's approach to find a meaningful relationship between shape parameters and masses of Cat-Chu mangoes. We used over 200 mangoes as a training



**Figure 2.** Equipment setup for measurement by lab instruments: (a) using a Vernier caliper—a device to measure size, (b) using water replacement method—a device to measure volume, and (c) using a planimeter—a device to calculate area.

dataset to establish the necessary weight-size relationship. Fortunately, we also obtain a linear relationship as shown in the following **Figure 3**. The constant in our case is  $4.879 \times 10^{-4}$ . The obtained  $R^2$  is about 97.6%.

Our estimated mass is:

$$\text{Estimated mass\_CatChuMango (grams)} = 4.879.39 \times 10^{-4} \times L \times W \times T \quad (2)$$

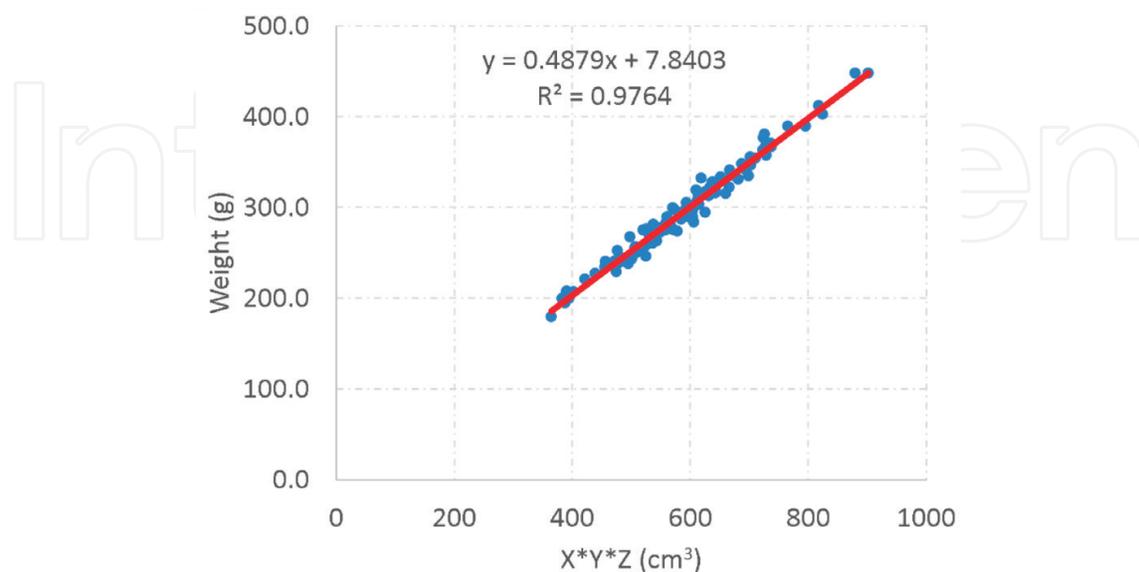
To validate our findings, we collected an additional 68 mangoes to be used as a validation dataset. The accuracy achieved is impressive, with an average error percentage of 3.23%. This further proves that the simple, linear correlation between mass and sizes can be used to estimate the corresponding mass effectively.

We also designed and constructed an image capturing platform to obtain the images from two different viewpoints (top and side views). The platform would also be used to test the algorithm's ability to estimate mangoes' masses solely based on their sizes. An algorithm was developed to capture and process the images while mangoes travel along a conveyor.

Top and side views of the mango were captured to estimate the mango mass using Eq. (2), and the result will also be compared with conventional mass measurement using a calibrated digital scale. We found that the difference between the masses estimated using this technique was not statistically different from the conventional method using a digital scale ( $p < 0.05$ ). Classification result showed an accuracy of 95–96% when grading mangoes solely based on masses.

## 2.2. Image segmentation

In this section we review a few methods for automatic selection of threshold values; the most important methods that we will discuss are Otsu's method and the valley-emphasis method.



**Figure 3.** Correlation between mango masses and their sizes: L, T, and W.

For a more general discussion regarding thresholding techniques, please read the reference “Machine Vision” by Davies [6].

### 2.2.1. Otsu’s method

This used to be one of the de facto algorithms in image segmentation [7]. An image is a two-dimensional matrix of  $N$  pixels, each with an intensity level between 0 and  $L-1$ , where  $L$  is the number of distinct gray levels. The number of pixels with a certain gray level  $i$  is denoted as  $f_i$ , and the probability of occurrence of gray level  $i$  is given by

$$P_i = f_i/N \tag{3}$$

The average of the intensity level of the whole image can be calculated as

$$\mu_T = \sum_{i=0}^{L-1} i p_i \tag{4}$$

By segmenting the image using a single threshold, we get two disjoint regions  $C_1$  and  $C_2$ , which are formed by the area of pixels with gray levels  $[1, \dots, t]$  and  $[t, \dots, L]$ , respectively, where  $t$  is the threshold level. Normally,  $C_1$  and  $C_2$  correspond to the object of interest and the background. The probability distributions of  $C_1$  and  $C_2$  are

$$C_1 = \frac{P_1}{\omega_1(t)}, \dots, \frac{P_t}{\omega_1(t)} \text{ and } C_2 = \frac{P_{t+1}}{\omega_2(t)}, \frac{P_{t+2}}{\omega_2(t)}, \dots, \frac{P_L}{\omega_2(t)} \tag{5}$$

$$\omega_1(t) = \sum_{i=1}^t P_i \text{ and } \omega_2(t) = \sum_{i=t+1}^L P_i \tag{6}$$

The mean gray-level values of the two classes can be computed as

$$\mu_1 = \sum_{i=1}^t \frac{i \cdot P_i}{\omega_1(t)} \text{ and } \mu_2 = \sum_{i=t+1}^L \frac{i \cdot P_i}{\omega_2(t)} \tag{7}$$

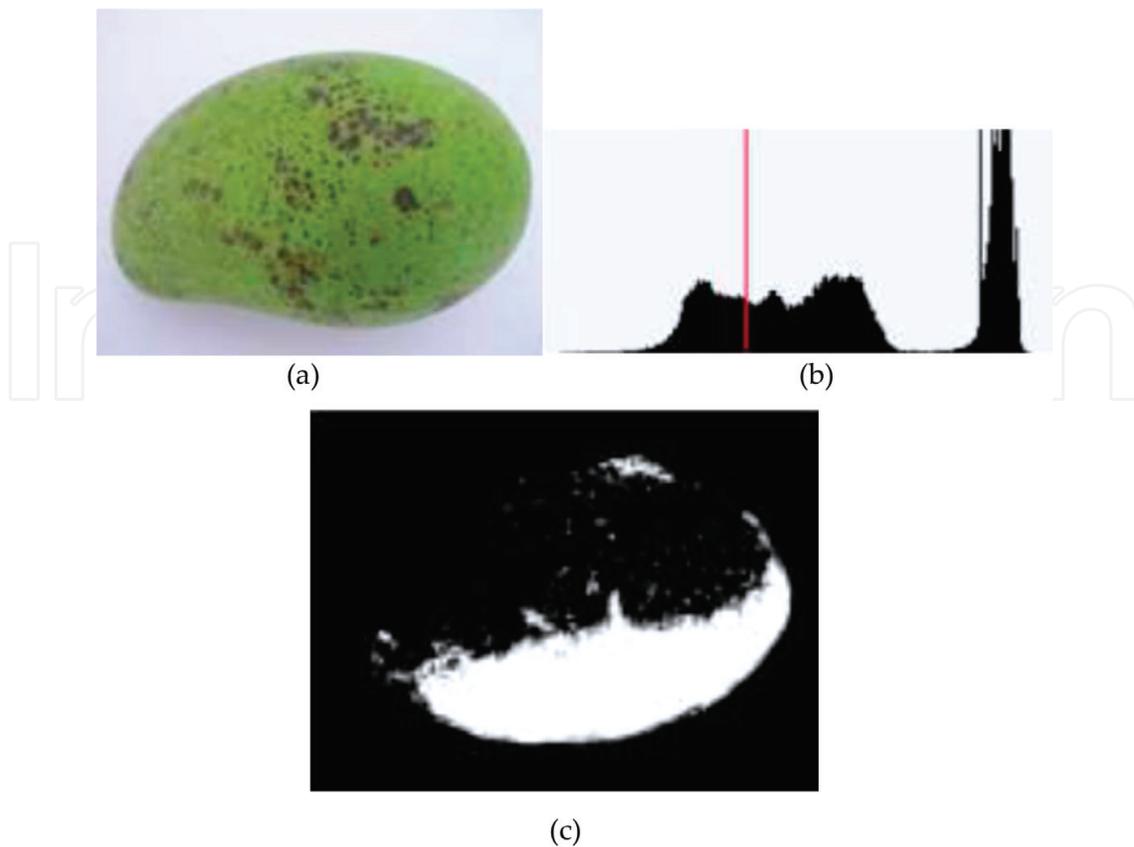
Using discriminant analysis, Otsu [7] showed that the optimal threshold  $t^*$  can be determined by maximizing the between-class variance, that is

$$t^* = \underset{0 \leq t < L}{\text{ArgMax}} \{ \sigma_B^2(t) \} \tag{8}$$

where the between-class variance  $\sigma_B$  is defined as

$$\sigma_B^2 = \omega_1 \cdot (\mu_1 - \mu_T)^2 + \omega_2 \cdot (\mu_2 - \mu_T)^2 \tag{9}$$

Otsu’s method works well when the images have clear peaks and valleys—in other words, it works for images whose histograms show clear bimodal or multimodal distributions. There are times when histograms of images contain several different types with widely varied number of pixels, such as external defects; Otsu’s method will not give the correct threshold level as shown in the following **Figure 4**.



**Figure 4.** Segmentation result using Otsu's algorithm: (a) original image, (b) histogram with the Otsu threshold level (red), and (c) resulting image.

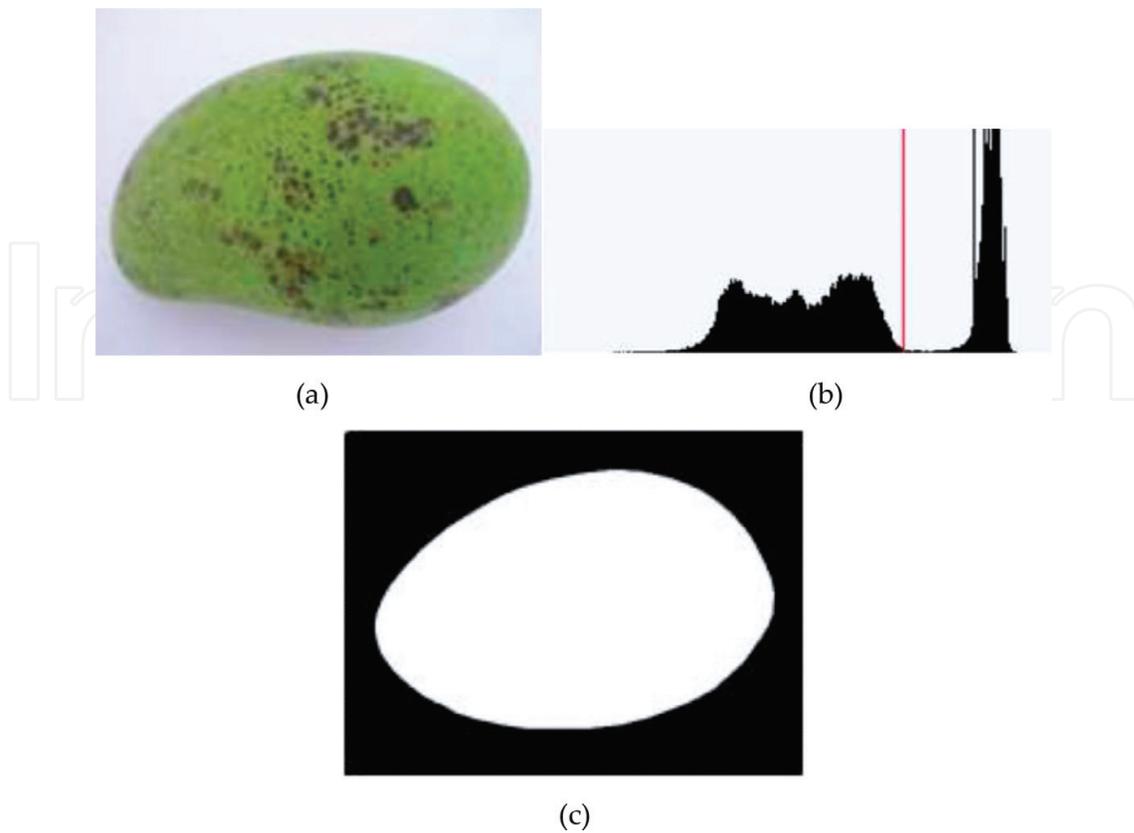
### 2.2.2. Valley-emphasis method

To improve drawbacks of Otsu's method, Ng et al. [8] proposed the valley-emphasis method. The idea of the valley-emphasis method is to select a threshold value that has a small probability of occurrence (valley in the gray-level histogram) and also maximize the between-group variance, as in Otsu's method. The formulation for the valley-emphasis method is

$$t^* = \underset{0 \leq t < L}{\text{ArgMax}} \left\{ (1 - P_t) (\omega_1(t) \mu_1^2(t) + \omega_2(t) \mu_2^2(t)) \right\} \quad (10)$$

The addition of an extra weight factor,  $(1 - P_t)$ , ensures the calculated threshold having a small probability of occurrence  $P_t$  will always be selected. Hence, the name valley-emphasis because the threshold level will always reside at the valley of the histogram. For images that have apparent bimodal distribution, the valley-emphasis method should give a threshold value that is very close to the value generated by Otsu's method because both methods attempt to maximize the between-group variance of the histogram.

The same segmentation experiment done previously using Otsu's method is repeated using the valley-emphasis method as shown in **Figure 5**. We can clearly observe that the segmentation result is much better. And, the result can be utilized for further analysis steps.

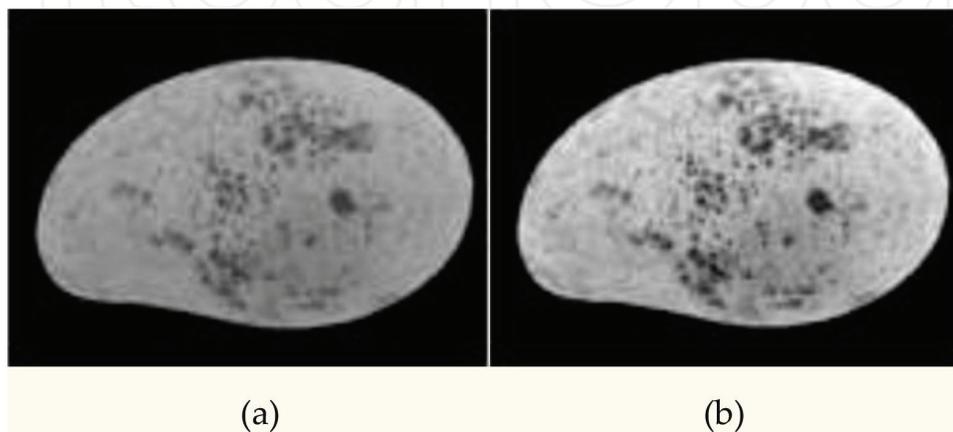


**Figure 5.** Segmentation result using valley-emphasis method: (a) original image, (b) histogram with the valley-emphasis threshold level (red), and (c) resulting image.

## 2.3. Defect isolation

### 2.3.1. Defect isolation

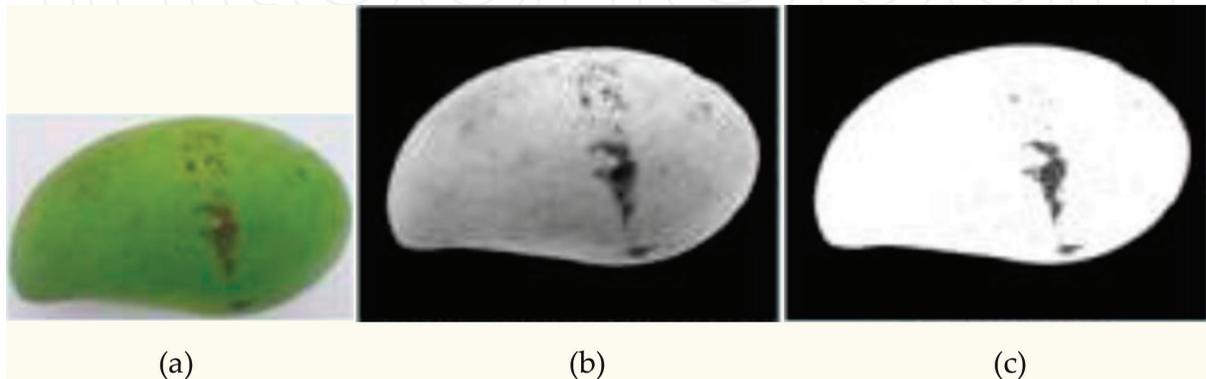
Due to their green appearances, we use G channel as the main channel, since it will be much easier to observe defects. To make the defects stand out, we use a simple linear contrast enhancement as shown in [3]. The results shown in **Figure 6** illustrate the effectiveness of the contrast enhancement.



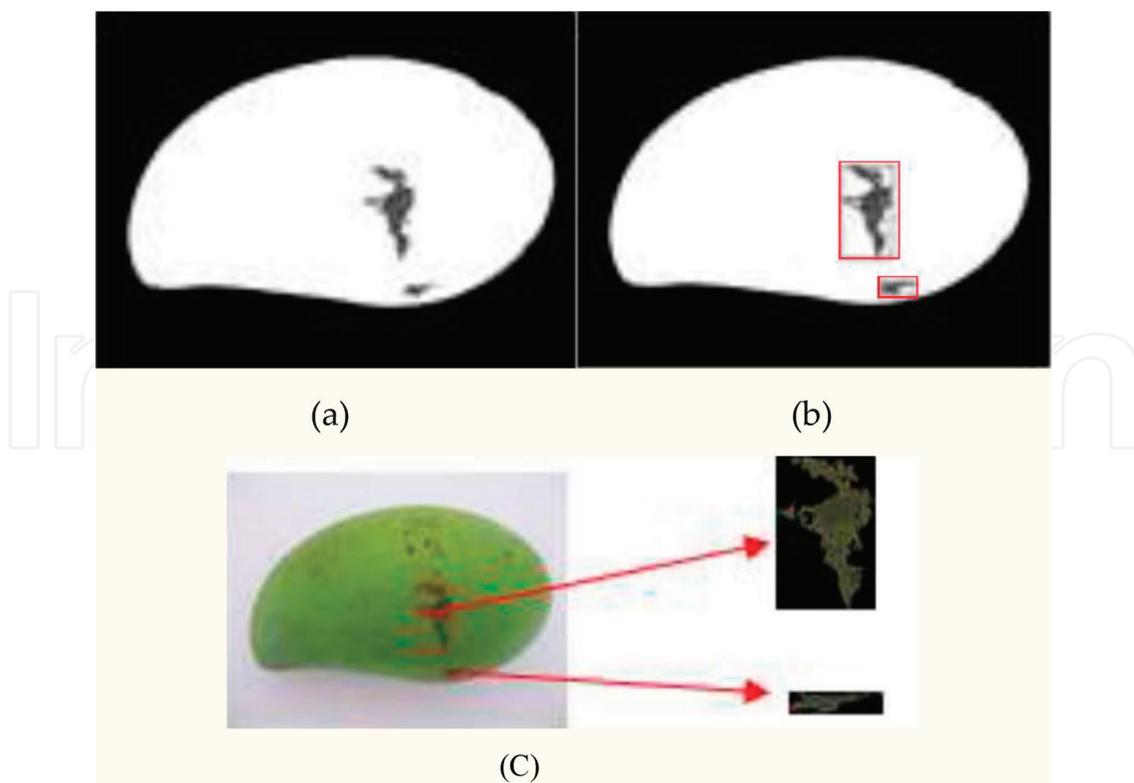
**Figure 6.** Contrast improvement after background removal: (a) before and (b) after.

After image enhancement, we apply another round of valley-emphasis segmentation on the area of the mango mask to isolate the defect zones. The result was illustrated as shown in **Figure 7**.

To simplify the calculation effort, we only concentrate on defects that are equal to or larger than 30 pixels. After segmenting the defect zones from the previous steps, we will use their sizes and locations on the original image to generate the new defect candidate for further classification steps as shown in **Figure 8**.



**Figure 7.** Defect zone isolation: (a) original image, (b) after background removal and contrast enhancement, and (c) defect isolation result.



**Figure 8.** Defect zones on the original image.

### 2.3.2. Defect classification

There are many kinds of defects that negatively degrade mangoes' quality [9]. Among them, four kinds that are most commonly seen are shown in **Figure 9** including stripe-type scars, dark patches, sap burns, and small spots. The defect classification steps will help us know how many kinds of defects are present on the fruit skin area as shown in **Figure 10**.

#### 2.3.2.1. Color features

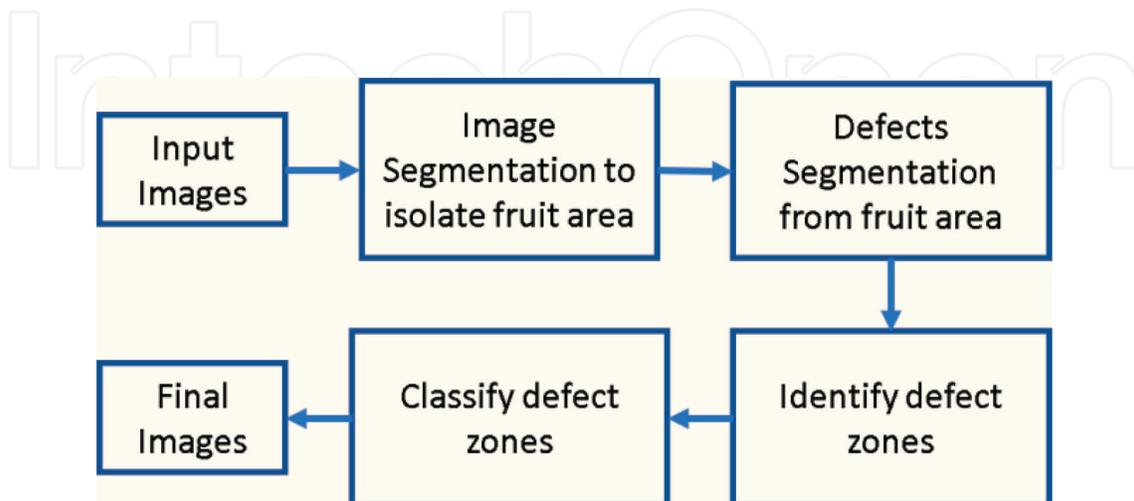
We use an artificial neural network with inputs as color features, shape features, and image statistical information. Li et al. [10] suggested that using HSV (HSI) instead of RGB color space improves segmentation results. In this research, there are 18H bins, 3S bins, and 3V bins. Therefore, we will have 162 features in HSV space.

#### 2.3.2.2. Shape features

To calculate shape features, we used the moment invariant proposed by Hu [11] with practical implementations by OpenCV as in [12].



**Figure 9.** Four typical types of defects (I) to (IV).



**Figure 10.** Proposed image processing procedures.

### 2.3.2.3. Histogram-based features

The histogram-based features used in this work are first-order statistics that include mean, variance, skewness, and kurtosis for all R, G, and B channels. Let  $z$  be a random variable denoting image gray levels and  $p(z_i)$ ,  $i = 0, 1, 2, 3, \dots, L-1$ , be the corresponding histogram, where  $L$  is the number of distinct gray levels. The five following features for each color channel are calculated using the abovementioned histogram:

$$\text{Mean} = \sum_{i=0}^{L-1} z_i p(z_i) \quad (11)$$

$$\text{Variance} = \sum_{i=0}^{L-1} (z_i - \text{Mean})^2 p(z_i) \quad (12)$$

$$\text{Skewness} = \sum_{i=0}^{L-1} (z_i - \text{Mean})^3 p(z_i) \quad (13)$$

$$\text{Kurtosis} = \sum_{i=0}^{L-1} (z_i - \text{Mean})^4 p(z_i) \quad (14)$$

$$\text{Range} = \frac{\text{max} - \text{min}}{255} \quad (15)$$

### 2.3.2.4. Manual labeling of training data

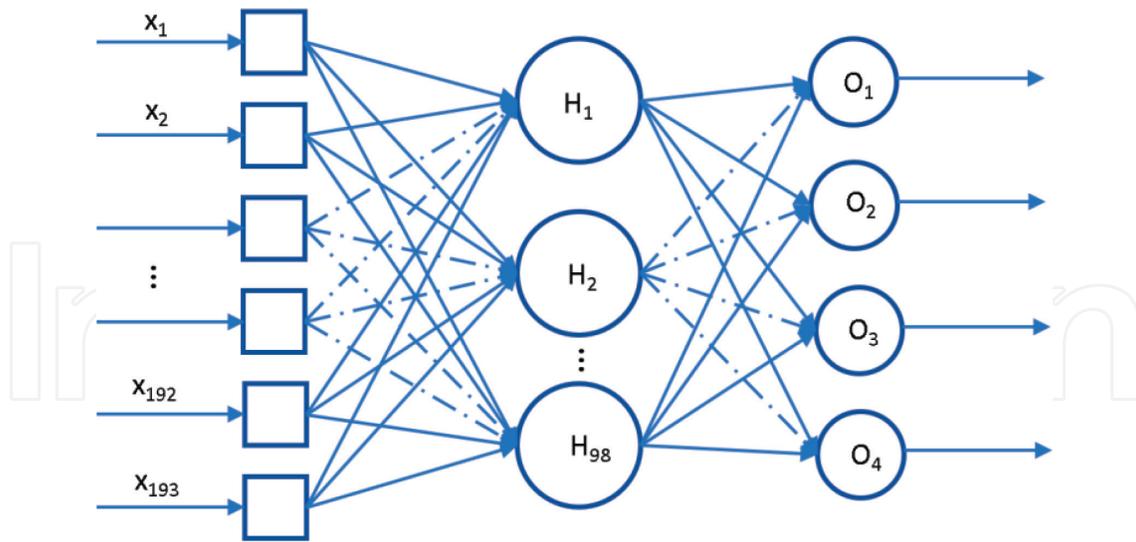
We prepare our dataset with standardized defect templates of  $20 \times 20$  using 193 abovementioned features: 162 color-based, 16 shape-based, and 15 histogram-based. We also manually label different images in the training dataset with different kinds of defects. For example, "Image 1" has nine defect zones, one is the first defect type and the rest are the fourth defect type. The procedure is applied similarly for the rest of the training images.

### 2.3.2.5. Building a neural network model

Our classification problem is a nonlinear one with 193 inputs corresponding to 193 chosen features and 4 outputs corresponding to four types of defects. Usually, a number of hidden layers are experimentally chosen to be a half of all the number of inputs and outputs [3]. Therefore, we chose 98 hidden neurons ( $= (193 + 4)/2$ ). The neural network model is illustrated in **Figure 11**.

We split our dataset into five smaller ones with different characteristics:

- Set 1: All images show no defects.
- Set 2: Images show only one type of defects.
- Set 3: Images show only two types of defects.
- Set 4: Images show three types of defects.
- Set 5: Images have all four types of defects.



**Figure 11.** Our proposed feed-forward neural network with 193 inputs, 98 neurons in a hidden layer, and 4 outputs.

	Set 1	Set 2	Set 3	Set 4	Set 5	Total
Number of photos	70	50	40	30	30	220
Number of defects	70	106	272	425	722	1595
Number of defects correctly identified	70	103	258	392	657	1480
Number of wrong identifications	0	3	14	33	65	115
Accuracy (%)	100	97.17	94.85	92.24	91.0	92.79

**Table 1.** Summary of classification results.

The classification results are summarized in **Table 1**. From the statistics, we can see that the classification accuracy reduces with an increasing number of defect zones and it also takes more computation time. The result is quite promising to be applicable to an automated sorting and grading system. In the current version, no acceleration techniques have been applied; in the near future, advanced parallel programming technique using graphics processing units (GPU) can be utilized to speed up the process, hopefully, to achieve a real-time performance level.

### 3. Conclusion

In this work, we have established an integrated framework for an automated grading, sorting, and weighing system of Cat-Chu mangoes using features including weight, size, and external defects. We found a simple, easy-to-calculate formulation between simple parameters and mango mass. The estimation error is very small, less than 3% if we use a mechanical measurement tool and less than 5% if we use an optical measurement using top- and side-view image captures. We also proposed an innovative procedure to classify external defects based on an

artificial neural network. The classification error is less than 8% for the worst possible case. The results indicate that our system has a great potential to be used in a real industrial setting. Future work will aim to investigate other features such as ripeness and bruises to increase the effectiveness and practicality of the system and possible speedup to real-time performance using advanced graphics processing unit (GPU) and further code parallelism.

## Acknowledgements

This work is being carried out at the International University, Vietnam National University Ho Chi Minh City, and being funded by Ho Chi Minh City Department of Science and Technology under the Contract Number 236/2017/HD-SKHCN. The project title is “Research, design, fabricate a prototype of a mango sorter for export.”

## Author details

Son V.T. Dao

Address all correspondence to: [dvtson@hcmiu.edu.vn](mailto:dvtson@hcmiu.edu.vn)

Department of Industrial and Systems Engineering, International University—Vietnam National University, Ho Chi Minh City, Vietnam

## References

- [1] Cubero S, Aleixos N, Moltó E, Gómez-Sanchis J, Blasco J. Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology*. 2011;4:487-504. DOI: 10.1007/s11947-010-0411-8
- [2] Guzman-Estrada C, Mosqueda Vazques R, Alcalde-Blanco S, Marinez-Garza A. Equation to estimate volume and growth dynamics of mango ‘Manila’ fruit. *Acta Horticulturae*. 1997;455:449-454. DOI: 10.17660/ActaHortic.1997.455.57
- [3] Vasquez-Caicedo AL, Neidhart S, Pathomrungsinyounggul P, Wiriyacharee P, Chattrakul A, Sruamsiri P, et al. Physical, chemical and sensory properties of 9 Thai mango cultivars and evaluation of their technological and nutritional potential. In: *Conference for Sustaining Food Security and Managing Natural Resources in Southeast Asia*; 8-11 January; Chiang Mai, Thailand. 2002
- [4] Chalidabhongse T, Yimyam P, Sirisomboon P. 2D/3D vision-based mango’s feature extraction and sorting. In: *9th International Conference on Control, Automation, Robotics and Vision, ICARV’06*; 5-8 December; Singapore. 2006. pp. 1-6
- [5] Spreer W, Müller J. Estimating the mass of mango fruit (*Mangifera indica*, cv. Chok Anan) from its geometric dimensions by optical measurement. *Computers and Electronics in Agriculture*. 2011;75:125-131. DOI: 10.1016/j.compag.2010.10.007

- [6] Davies ER. Machine Vision. 3rd ed. Morgan Kaufman; 2004. p. 938. DOI: 10.1016/B978-0-12-206093-9.X5000-X
- [7] Otsu N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics. 1979;9(1):62-66. DOI: 10.1109/TSMC.1979.4310076
- [8] Ng HF. Automatic thresholding for defect detection. Pattern Recognition Letters. 2006; 27(14):1644-1649. DOI: 10.1016/j.patrec.2006.03.009
- [9] Mango Quality Standard. Available from: <https://www.industry.mangoes.net.au/resource-collection/2016/9/29/2016-mango-industry-quality-standards> [Accessed: 31-08-2018]
- [10] Qingzhong L, Maohua W, Weikang G. Computer vision based system for apple surface defect detection. Computers and Electronics in Agriculture. 2002;36(2):215-223. DOI: 10.1016/S0168-1699(02)00093-5
- [11] Hu MK. Visual pattern recognition by moment invariants. IRE Transactions on Information Theory. 1962;8(2):179-187. DOI: 10.1109/TIT.1962.1057692
- [12] OpenCV Tutorial. Available from: <http://docs.opencv.org/doc/tutorials/tutorials.html> [Accessed: 30-07-2018]

IntechOpen

