ABE

Agricultural and Biological Engineering

https://ph04.tci-thaijo.org/index.php/abe/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Measuring damaged skin of mangosteen using image processing

Thipat Seela[1, 2] and Jetsada Posom[1]*

[1]Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen, 40002, Thailand
[2]Agricultural Engineering Research Center of Chantaburi, Agricultural Engineering Research Institute, Chantaburi, Thailand

*Corresponding author.
Email address: jetspo@kku.ac.th

Received 15 October 2024
Accepted 6 November 2024

Abstract

Mangosteen is a major economic crop. Currently, commercial production still faces challenges in terms of quality sorting, particularly in adhering to the skin color standards which serve as quality criteria. Presently, quality sorting heavily relies on the expertise of individuals, especially for mangosteen with damaged skin, which cannot be exported. Advances in image processing technology allow for quality sorting, thus this research aims to examine mangosteen with damaged skin using image processing techniques. A sample of 60 mangosteen fruits at six maturity levels, with 20 fruits per level, images were taken from four sides using RGB cameras, totaling 480 images. These images were analyzed and models were built for distinguishing between good skin and damaged-skinned mangosteen using Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Decision Tree (Fine Tree) algorithms. Results showed that all three algorithms performed similarly performance. For levels 1 through 6, the average accuracy rates were approximately 100, 95.61, 93.03, 99.63, 99.40 and 100, respectively, with average recall rates of 100, 96.60, 94.45, 99.90, 99.73, and 100, respectively. Analysis revealed that evaluating damaged skin at levels 2 and 3 had the lowest effectiveness, as the good skin colors of mangosteen at levels 2 and 3 closely resembled the colors of the damaged skin. Therefore, the research demonstrates that image processing can effectively separate damaged-skinned mangosteen from good-skinned.

Keywords: Mangosteen, Color analysis, Image analysis

1. Introduction

Mangosteen, known as the queen of fruits, is a key agricultural product contributing to Thailand’s economy. According to a 2023 report, the export of fresh mangosteens reached 248,588.35 tons, valued at 17,191.23 million THB. Additionally, frozen mangosteens amounted to 231,574 kg, valued at 27.84 million THB [1]. The main importing countries are China, Hong Kong, Vietnam, the United States, South Korea, Japan, and Taiwan. Exporting mangosteens must meet certain agricultural product standards. These include quality: whole fruits with intact stems, no bruises, dents, cracks, and cleanliness. Mangosteens are categorized into three quality grades: Extra Class, Class I, and Class II. Extra Class, the highest quality, must have intact sepals and stems, with surface blemishes not exceeding 10% and translucent pulp or latex issues in no more than 5% of the fruits. Class I allows for minor imperfections, with surface blemishes up to 30% and pulp or latex issues in no more than 5%. Class II includes fruits that do not meet the standards for Extra Class or Class I, with surface blemishes not exceeding 50% and pulp or latex issues in no more than 20%. The ripening stage for export begins at stage 2, where the fruit is light green with red-purple specks, based on the six-stage ripeness index. Stage 1 is green fruit, stage 2 is light green with purple specks, stage 3 is pink throughout, stage 4 is reddish-purple, stage 5 is purplish-red, and stage 6 is dark purple or black [2].

Post-harvest handling of mangosteens for export involves labor-intensive and time-consuming sorting. Skilled workers are needed to grade the fruit, especially when using machines to sort by size. While this helps reduce time, no technology is currently available for sorting based on surface quality. Image processing techniques are now being studied for sorting agricultural products based on size and ripeness. For example, color analysis has been used to classify banana chips' ripeness with an accuracy of 86.2% ± 7% and precision of 84.06%. Bayesian image processing techniques have been applied to sort mangoes with 88% accuracy [3].

Other technologies include microcontrollers, skin color sensors for mangosteens, and specific gravity methods for sorting translucent pulp mangosteens. Image processing is ideal for assessing ripeness in agricultural products. Momin et al. [4] develop a machine vision system that can efficiently acquire images suitable for further processing and develop an algorithm for grading mango based on selected features using image analysis. A 2021 study used digital image analysis to grade mangosteens for export, finding that the technique was 99.54% effective in size sorting. The study recommended further research to improve accuracy in detecting cracked or latex-damaged fruits [5].

For durians, near-infrared (NIR) technology with short (450-1000 nm) and long wavelengths (860-1750 nm) was used in conjunction with supervised machine learning models like Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). The results showed that long-wave NIR had an accuracy of 83.15-88.04%, while short-wave NIR had an accuracy of 64.73-93.77%. The LDA model performed best at 97.28% and 100% accuracy for long- and short-wave analysis at the stem and skin positions, respectively [6].

Machine learning techniques such as one-class support vector machine and isolation forest models have been used to sort products based on training data. The models' performance was measured using four methods: Confusion Matrix, Accuracy, F1-score macro average, and ROC AUC. The results indicated similar predictive performance, though the isolation forest model showed higher prediction accuracy [7]. However, research is still lacking in sorting mangosteens with damaged or scarred skin caused by insect damage, making them unsuitable for export. The researcher believes it is feasible to use image processing techniques to identify defective or scarred mangosteens, leading to more efficient sorting machines for export.

2. Materials and methods

2.1 Equipments

The study used 120 mangosteens at six different ripening stages, a FUJI Film XA-5 digital camera with an XC15-45mm F3.5-5.6 OIS PZ lens, a 40x40x40 cm PULUZ portable photography box with 60 LED strip lights (5,500K color temperature), and a Digital lux meter model LX1010B.

2.2 Methods

2.2.1 Randomly select newly harvested mangosteen samples from a local collection center (called “Lhong”) using the standard harvesting index for mangosteens (TAS 2-2556) (Table 1). A total of 120 samples, with 20 fruits at each ripening stage were selected, including mangosteens with both normal skin (smooth) and defective skin (scarred) (Figure 1).

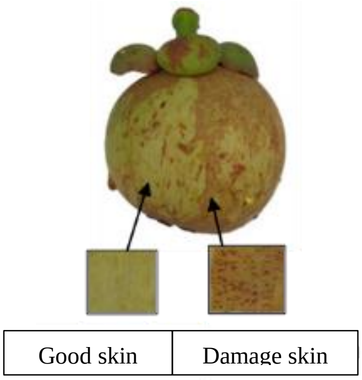


Figure 1 Good and damage skin of mangosteen

Table 1 Standard of mangosteen maurity

	Stages	Characteristics of Mangosteen
	1	The fruit is green, has a lot of latex in the rind, and the flesh and skin cannot be separated from each other.
	2	The fruit is light green with beginning purple-red spots, and the flesh and skin can be separated with difficulty to moderately.
	3	The fruit has a pink color distributed throughout, contains little latex in the rind, and the flesh and skin can be separated moderately.
	4	The fruit is reddish-purple and contains little latex in the rind.
	5	The fruit is purplish-red, contains no latex in the rind, and the flesh and skin cannot withstand pressure when separated. This is the stage where it begins to be edible.
	6	The fruit is dark purple or blackish-purple, contains no latex in the rind, and the flesh and skin cannot withstand pressure when separated.

Source: National Bureau of Agricultural Commodity and Food Standards [2].

2.2.2 Capturing images of the mangosteen samples was done using a digital camera inside the portable photography box (Figure 2). The process was carried out as follows:

(1) Select mangosteen samples with surface defects (scarred skin) at all six ripening stages. Place the mangosteen on a cylindrical stand to prevent movement, ensuring a distance of 25 cm between the camera and the mangosteen's surface.

(2) Set the camera to manual mode, with a shutter speed of 1/800, focus at 4, and ISO at 200.

(3) Turn on the LED lights to level 10 and check the light intensity using a lux meter. The lighting for the photos was set to 6,000-6,500 lux.

(4) Capture images of the mangosteen sample by taking four photos per fruit, rotating the mangosteen 90 degrees each time to cover the entire surface. A total of 480 images were captured.

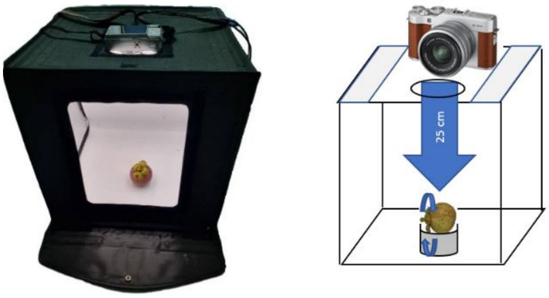
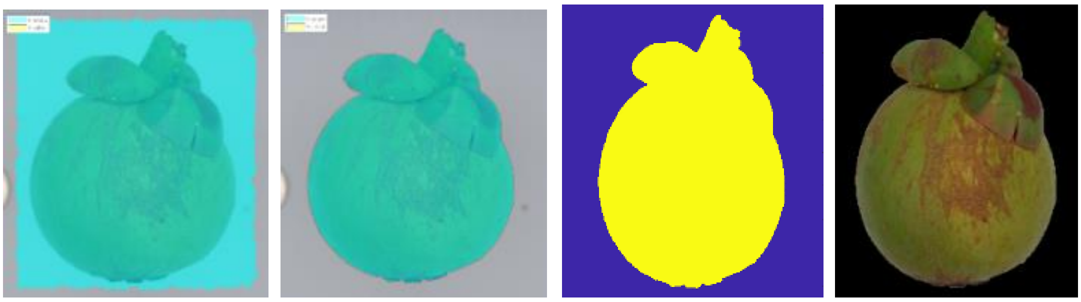


Figure 2 The process of mangosteen's photo collection

2.2.3 The captured images were pre-processed for analysis using Matlab 2018. The images were resized to 1000x1000 pixels, and the background was removed using image segmentation techniques, including Local Graph Cut, Active Contours, and Morphology Operations (Figure 3a). The sepals were removed, leaving only the mangosteen peel, using Graph Cut and Morphology Operations (Figure 3b).



(a)

(b)

Figure 3 Image pre-processing, a) Local graph cut, active contours and morphology operation, b) Graph cut and morphology operation

2.2.4 Class feature extraction was performed to distinguish between normal skin (smooth skin) and defective skin (scarred skin). Four sections of smooth skin and four sections of scarred skin were cropped, each with a resolution of 50x50 pixels. The RGB values of these sections were recorded as 3D data. This data was then converted into 2D format, and the RGB values were separated into two datasets: one for smooth skin and one for scarred skin. The data was also categorized according to the six different ripening stages (Figure 4).

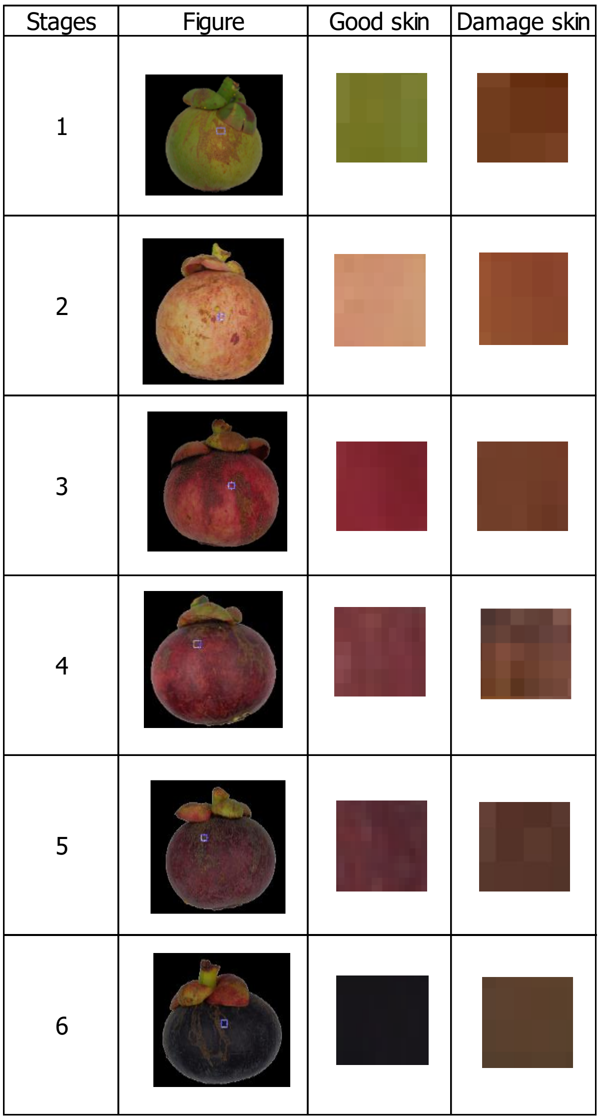


Figure 4 Class of feature extraction

2.2.5. Data analysis was conducted using Machine Learning techniques with the Classification Learner tool. The RGB data from step 4 was used for training, and three models were applied for prediction: Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Decision Tree. The performance of these models was evaluated based on the following metrics: Overall Accuracy, Precision, Recall, Specificity, and AUC-ROC, calculated from the confusion matrix.
The model's effectiveness was assessed using these metrics, where Overall Accuracy, Precision, Recall, and Specificity were calculated using equations (1), (2), (3), and (4), respectively. In these equations, the variables are defined as follows:

$$\text{Overall accuracy (\%)} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (1)$$

$$\text{Precision(\%)} = \frac{TP}{TP+FP} \times 100 \quad (2)$$

$$\text{Recall (\%)} = \frac{TP}{TP+FN} \times 100 \quad (3)$$

$$\text{Specificity (\%)} = \frac{TN}{TN+FP} \times 100 \quad (4)$$

where
TP (True Positive): the number of correctly predicted positive cases
TN (True Negative): the number of correctly predicted negative cases
FP (False Positive): the number of incorrectly predicted positive cases
FN (False Negative): the number of incorrectly predicted negative cases

3. Results and discussion

When considering the Overall Accuracy and Precision of each model at different ripening stages, it was found that at Ripeness Stage 1, both the LDA (Quadratic Discriminant) and SVM (Fine Gaussian Support Vector Machine) models provided an Overall

Accuracy of 100% and a Recall of 100%. At Ripeness Stage 2, the LDA (Quadratic Discriminant) model gave an Overall Accuracy of 95.61% and a Recall of 96.60%. At Ripeness Stage 3, the SVM (Fine Gaussian Support Vector Machine) model provided an Overall Accuracy of 93.30% and a Recall of 94.45%. At Ripeness Stage 4, the SVM (Fine Gaussian Support Vector Machine) model delivered an Overall Accuracy of 99.63% and a Recall of 99.90%. At Ripeness Stage 5, the SVM (Fine Gaussian Support Vector Machine) model showed an Overall Accuracy of 99.40% and a Recall of 99.08%. Finally, at Ripeness Stage 6, the SVM (Fine Gaussian Support Vector Machine) model achieved an Overall Accuracy of 100% and a Recall of 100% (Table 1).

Table 1 Results of machine learning model for mangosteen

Class	Model	N(pixel)	Precision	Recall	Specificity	Overall accuracy	AUC-ROC
1	LDA (Quadratic discriminant)	8000	100	100	100	100	1
	SVM (Fine GaussianSVM)	8000	100	100	100	100	1
	Tree (Fine Tree)	8000	99.85	99.92	99.85	99.89	1
2	LDA (Quadratic discriminant)	8000	96.60	94.73	96.53	95.61	0.98
	SVM (Fine GaussianSVM)	8000	94.38	93.03	94.29	93.65	0.98
	Tree (Fine Tree)	8000	95.80	93.67	95.70	94.66	0.98
3	LDA (Quadratic discriminant)	8000	85.73	84.02	85.43	84.71	0.94
	SVM (Fine GaussuanSVM)	8000	94.45	91.83	94.29	93.03	0.98
	Tree (Fine Tree)	8000	90.50	89.45	90.39	89.91	0.96
4	LDA (Quadratic discriminant)	8000	99.40	96.48	99.38	97.89	0.98
	SVM (Fine GaussianSVM)	8000	99.90	99.35	99.90	99.63	1
	Tree (Fine Tree)	8000	98.70	98.65	98.70	98.68	0.99
5	LDA (Quadratic discriminant)	8000	99.13	96.87	99.10	97.96	0.99
	SVM (Fine GaussianSVM)	8000	99.73	99.08	99.72	99.40	1
	Tree (Fine Tree)	8000	98.40	98.72	98.41	98.56	0.99
6	LDA (Quadratic discriminant)	8000	100	98.77	100	99.38	1
	SVM (Fine GaussianSVM)	8000	100	100	100	100	1
	Tree (Fine Tree)	8000	99.90	99.92	99.90	99.91	1

From the data displayed in the form of a scatter plot, it can be observed that mangosteens at ripeness stages 1 and 6 can be accurately and clearly separated into good skin (orange points) and scarred skin (blue points). Meanwhile, mangosteens at ripeness stages 2, 3, 4, and 5 can also be distinguished with accuracy levels exceeding 93% (Figure 5). Therefore, the model that is most suitable for developing a quality sorting system for mangosteen skin for export purposes is the SVM (Fine Gaussian Support Vector Machine). However, at stage 3, the accuracy in distinguishing between good and scarred skin is lower than at other stages due to the similar RGB values. Additionally, this study has limitations, as the image quality at the edges of the mangosteen fruit is lower than in other areas, which may lead to prediction errors (Figure 6).

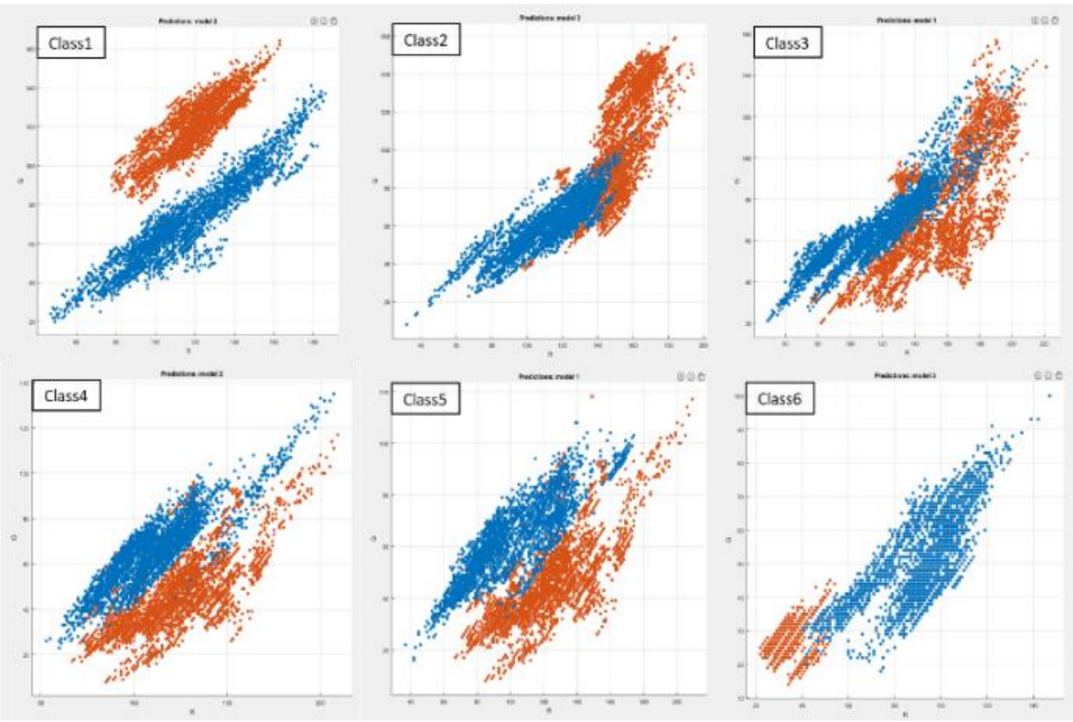


Figure 5 Scatter plot type of mangosteen skin analysis by machine learning

Figure 6 Result from binary pixel to original pixel

4. Conclusions

From the study on the quality classification of mangosteen skin using image analysis through Machine Learning, the quality of mangosteen at six ripeness stages was analyzed. A total of 480 images of mangosteens were used to predict scarred skin and good skin employing three models: Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Decision Tree. The evaluation metrics included Overall Accuracy, Precision, Recall, Specificity, and AUC-ROC. The results indicated that mangosteens at ripeness stages 1 and 6 could be distinctly and accurately separated into good skin (orange points) and scarred skin (blue points), with the SVM (Fine Gaussian Support Vector Machine) model achieving an Overall Accuracy of 100% and a Recall of 100%. Additionally, for mangosteen at ripeness stages 2, 3, 4, and 5, the SVM model demonstrated accuracy levels ranging from 93.03% to 99.63% and Recall values from 91.83% to 99.35% (Figure 5). Therefore, the SVM (Fine Gaussian Support Vector Machine) model is deemed suitable for developing a quality sorting system for mangosteen skin for export purposes. Although, at ripeness stage 3, the accuracy in distinguishing between good skin and scarred skin was lower than at other stages due to similar RGB values. This study also faced limitations, as the image quality at the edges of the mangosteen fruit was lower than in other areas, leading to potential prediction errors.

5. Acknowledgements

This research report has been successfully completed due to the support from the owner of the mangosteen farm in Chanthaburi and the staff of the Chanthaburi Agricultural Engineering Research Center for their assistance in data collection and provision of research equipment. The researcher would like to express sincere gratitude to them on this occasion.

6. References

[1] Office of Agricultural Economics. Export and import statistics [Internet]. 2024 [cited 2024 June 20]. Available from: https://impexpth.oae.go.th/export.

[2] National Bureau of Agricultural Commodity and Food Standards. Agricultural product standards, mangosteen. Bangkok: National Bureau of Agricultural Commodity and Food Standards. 2014.

[3] Rajamangala University of Technology Krungthep. Report on the performance of the project under the strategic plan for sustainable development [Internet]. 2022 [cited 2024 June 20]. Available from: http://sdg.rmutk.ac.th/files/projects/2022.

[4] Momin MA, Rahman MT, Sultana MS, Igathinathane C, Ziauddin ATM, Grift TE. Geometry-based mass grading of mango fruits using image processing. Information Processing in Agriculture. 2017;4(2):150-160.

[5] Thammastitkul A, Klayjumlang T. Mangosteen quality grading for export markets using digital image processing techniques. International Journal of Advanced Science Engineering Information Technology. 2021;11(6):2452-2458.

[6] Ditcharoen S, Sirisomboon P, Saengprachatanarug K, Phuphaphud A, Rittiron R, Terdwongworakul A, Posom J. Improving the non-destructive maturity classification model for durian fruit using near-infrared spectroscopy. Artificial Intelligence in Agriculture. 2023;7:35-43.

[7] Chotichatmala S. A single-type fruit sorting model using machine learning for apple detection. Nakhon Nayok: Srinakharinwirot University. 2020.