*ICIC Express Letters*
*Part B: Applications*
Volume **15**, Number **5**, May **2024**

ICIC International ©2024 ISSN 2185-2766
pp. 467–477

# QUALITY LEVEL IDENTIFICATION FOR *MANGIFERA INDICA* L. USING A DEEP LEARNING MODEL

Saowalak Arampongsanuwat*, Chanya Maneechai, Nuttachot Promrit and Sajjaporn Waijanya

Department of Computing
Faculty of Science
Silpakorn University
6, Rajamankha Nai Road, Muang District, Nakhon Pathom 73000, Thailand
maneechai_c@su.ac.th; { promrit_n; waijanya_s }@silpakorn.edu
*Corresponding author: Arampongsanuwat_s@su.ac.th

Received September 2023; accepted November 2023

Abstract. *Mangifera indica L. (Falan mango) is an economic plant prone to breakage and defects. Farmers thus need to grade their Falan mangoes before selling them. Traditionally, as mangoes are graded using the naked eye, especially with different eye values among individuals, errors in grading can happen. To avoid such errors, uplift grading accuracy and reduce the workforce cost, the construction of the Falan mangoes quality identification system was designed by integrating Mask R-convolutional neural network (CNN) to accurately segment Falan mango images. For the image segmentation and classification processes, the quality level identification model was created using a convolutional neural network (CNN) model. The input data were derived from collecting normal, defective and broken Falan mango images for training the model. After the tested images were graded, the results indicated the accuracy values of both quality levels at 85.64% and size classifications at 81.00%. The model was also applied to developing an application that receives image files from users' mobile devices for segmentation and classification processes. The result showed the summary of grades and sizes of Falan mangoes.*
**Keywords:** Deep learning, Mask R-CNN, Mango, Image processing, Falan mango

1. **Introduction.** Mangoes are perennials of the *Mangifera* genus, classified as tropical plants in the Anacardiaceae family and scientifically named *Mangijera indica* L. Mangoes originated from various regions across India, Bangladesh and Myanmar's northwest territories. They vary in species, shapes, sizes, weights, colours, textures and smells, among others. Mangoes are medium-sized shrubs, bearing fruits that are green when unripe and yellow when ripe, both of which are edible [1]. Mangoes are economically horticulture fruits registered in Thailand's farming areas at 599,369.61 rai (1 rai equals 1,600 $m^2$) across the country, with exports mainly to Malaysia, South Korea, Vietnam, Japan and Singapore [2].

The Falan mango, a popular variety consumed unripe, is long with a large head and pointed end. Its light green, crunchy flesh makes a cracking sound when peeled, earning its name 'Falan' (meaning thunder in Thai) [3]. Falan mangoes are more prone to defects than other types of mangoes. They are graded according to size and surface quality into normal and defective groups, with the latter subdivided into broken and blotched categories. Grading Falan mangoes based on size is a skilled task, often leading to variability in quality assessment and potential underpricing due to grading errors.

Agricultural items differ from one another in terms of colours, shapes and sizes [4]. Besides, their features vary concerning duration. This difference is a challenge for developing an accurate tool for detecting these items' features. In the agricultural domain, machine

DOI: 10.24507/icicelb.15.05.467

learning plays a crucial role in grading and evaluating the quality of fruits, which is an important measure in the fruit market, especially in detecting defects on the surface of fruits [5-7]. Deep learning techniques have been popularly deployed for analyzing agricultural images [8], increasing accurate prediction from cultivating to harvesting. These techniques can be used in various research fields, such as classification and clustering.

Among many deep learning techniques, the convolutional neural network (CNN) is one of the frequently used deep network architectures for image classification, segmentation and object detection, among others. Because CNN architectures can be determined flexibly, implementing the detection and classification of objects is significantly challenging, especially for fruits with diverse patterns, colours and textures.

Concerning studies on detection and classification based on deep learning, some were dedicated to applying deep learning to fruit detection. For instance, Arampongsanuwat and Chaowalit [9] used deep CNNs to classify the ripeness of mangosteen according to market segments, which were divided into four groups: export market, domestic market, local market and ungraded mangosteens. In this study, pretrained learning CNN models were deployed for classifying mangosteen images into different market segments, experimentally resulting in a classification accuracy of 85.00%. Meanwhile, Nithya et al. [10] developed a model for identifying defects in mangoes using a CNN deep learning model. In their study, 800 images of mangoes were evaluated and the classification accuracy was 98.5%. Albarrak et al. [11] trained a transfer learning model to classify eight different popular date fruits. The model provided 99% accuracy.

In the processing of images with multiple objects, the objects must be segmented before employing model processing. In this respect, Zhang et al. [12] provided a comprehensive survey of instance segmentation in computer vision, outlining its evolution from object detection and semantic segmentation covering methods like Mask R-CNN. Ganesh et al. [13] improved the mask segmentation performance. The developed framework was validated using images obtained from orange groves by introducing the instance segmentation framework for orange detection. The segmentation using a deep learning approach was presented by augmenting Mask R-CNN to include HSV input data, benefiting the harvest planning operation. Pan and Ahamed [14] used a fruit recognition method for systems developed to identify pears in a complex orchard environment using a three-dimensional (3D) stereo camera combined with Mask R-CNN deep learning technology to obtain targets by training for different shapes and sizes of fruit using deep learning algorithms. The Mask R-CNN algorithm had high accuracy in detecting the individual pears from the gathered pears in a complex orchard environment. Macías-Macías et al. [15] used Mask R-CNN to segment images of papayas to measure them according to their ripeness level: green, semiripe and ripe. Their proposed system successfully achieved an accuracy rate of 93.5%.

In addition, the study conducted by Liu et al. [16] focused on cucumber detection by applying Mask R-CNN in conjunction with the feature pyramid network to improve detection precision. The results of the comparison between Mask R-CNN and faster R-CNN indicated that the improved Mask R-CNN generated the highest scores and that its average elapsed time was lower than before the improvement. Meanwhile, the standard deviation of the location deviation in the improved Mask R-CNN was lower than those of other methods.

Dhiman et al. [17] compared different techniques used by researchers for fruit quality detection. They presented a review of various papers, emphasizing commonly used machine learning models for fruit quality classification. They reviewed, explored and discussed the progress in the field of fruit quality detection and several machine learning and deep learning approaches based on existing works. Although the preponderance of the approaches generated high accuracy, the quality improvement process can be enhanced in the future by extending to more fruits and their features for more dynamic research.

In connection with the aforementioned studies, many researchers not only developed quality detection and classification but also obtained satisfactory results using deep learning algorithms. This research proposes the application of deep learning techniques using the Mask R-CNN neural network for image segmentation. This provides a distinct mask for each mango in a batch, enabling the determination of the fruit's size before classifying the Falan mangoes' grade and quality. The overall concept is to start building a model through data collection, preparation, modelling and evaluation, followed by developing an application that presents classification results to users. This model categorizes Falan mangoes by size and surface condition into small, medium and large and normal, defective and broken, respectively. The study aims to streamline mango grading, boosting farmers' productivity with greater ease.

The implementation of this technology is intended to aid farmers by improving the precision and uniformity of mango grading, thereby reducing the reliance on subjective visual assessments, which can be affected by the evaluator's experience and fatigue. Additionally, this method takes advantage of the latest advancements in machine learning and image analysis to bolster the growth of Thailand's agricultural industry.

The following parts of this research are as follows: Section 2, Methodology; Section 3, Experimental Results; and Section 4, Discussion and Conclusions.

## 2. Methodology.

**2.1. Data preparation.** The data used for this research were from Falan mango images with different grades from a Falan mango farm in Samut Sakhon. In taking pictures of Falan mangoes, the mangoes were arranged in piles on a white cloth set to the specified dimensions ($90 \times 90$ cm), with sufficient brightness and from a height of 120 cm above the ground. There were 720 images at $600 \times 450$ pixels each and they were classified by an expert to be used as training data for classifying normal, defective and broken mangoes. The images used in this research were divided into training, testing and validation sets ($56 : 30 : 14$). Image data augmentation was implemented for the training images to increase the efficiency of the training set, enlarge the variety of the data and reduce any issues from overfitting of the model's recognition. The augmentation was set as follows: rotation range = 30°, width shift range = 20%, height shift range = 20%, shear range = 20%, zoom range = 30%, vertical flip = true and horizontal flip = true.

**2.2. System architecture.** The system architecture consisted of two key processes: 1) image segmentation process using the Mask R-CNN model to exclude objects in the image and separate mangoes from the background and 2) classification process to predict the grades of the mangoes.

The first process began with dataset labelling to identify mango images as JSON files before using them to train the Mask R-CNN model for image segmentation. The second process started with dataset labelling: out of 720 images, 70% (504 images) were used to train the model and 30% (216 images) to test it. Further, 20% of the training data (100 images) was set aside for validation to evaluate the model's performance during training and prevent overfitting.

For a greater variety of the dataset, the data underwent augmentation before being brought into the trained classification model (CNN model) to predict mango grades without deploying transfer learning as the model was not large but effective for classification, as shown in Figure 1.

The web application process began after receiving image files from users' devices into object segmentation. The segmented objects were classified by the classification model for grading mangoes based on surface features and sizes: normal, broken and defective and large, medium and small. The results were displayed via the application screen for users.

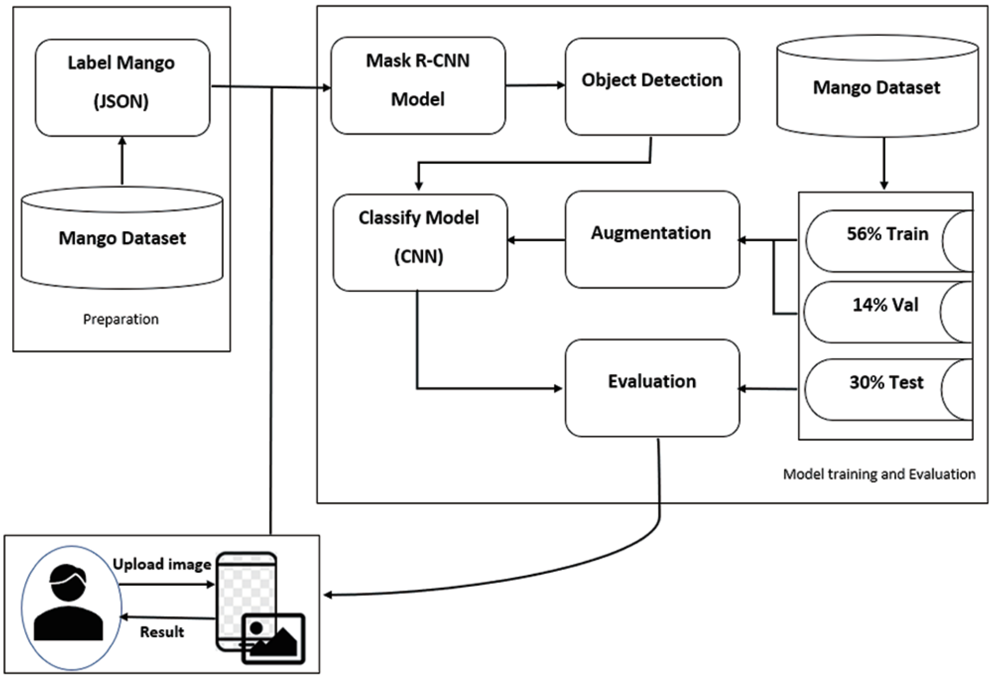


FIGURE 1. Overview of the methodology

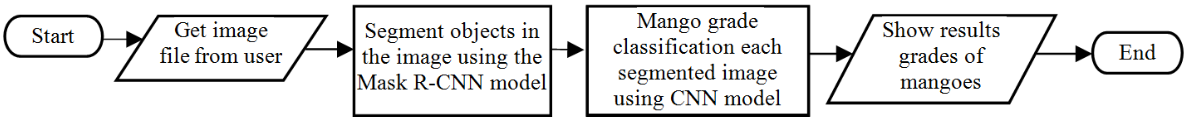


FIGURE 2. Prediction process

As illustrated in the prediction process in Figure 2, the system received image files from users' devices and then segmented the objects in the images using Mask R-CNN. The segmented objects (mangoes) were brought into the classification model to predict their grades in terms of normal, broken, and defective surface features and large, medium and small sizes before reporting the results to the users via the application.

2.3. **Training the image segmentation model using Mask R-CNN.** Falan mango images were labelled using the Make Sense programme by drawing different lines around the objects (mangoes). The drawn images were converted into JSON files, consisting of the data that define the shape of the objects represented as $x$ and $y$ coordinates and all the determined feature names. Next, the JSON files were trained using the Mask R-CNN model to define the regions of the mangoes in the images. Then, the JSON files were used to train Mask R-CNN for image segmentation.

For the image segmentation using the Mask R-CNN model, after importing mango images and loading the weights of the Mask-RCNN model for detecting mango images, the coordinates of the contours of the mangoes in the images were utilized in the following step. $x$, $y$, $w$ and $h$ were defined to frame and crop mango images available for contours and automate the cropping loops until all the mangoes were cropped, as illustrated in Figure 3, before continuing to the classification process.

We used the pretrained model with the Mask R-CNN pretrained on a coco dataset to reduce the time spent on creating a model for object detection.

2.4. **Deep learning network configuration.** The model was trained in 200 epochs with a batch size of 32, an Adam optimizer (an algorithm used to update the weights of the nodes) and a categorical cross-entropy loss function (an algorithm for suitably reducing

Figure 3. Masks of the objects detected (examples of mango images)

loss). The layers for the training included an input layer, which received 64 × 64 image sizes; a Conv2D layer, which extracted significant features from 3D into two-dimensional (2D) images; a Flatten layer, which converted the 2D images from the Conv2D layer into one-dimensional vectors; and a dense layer, a fully connected layer using ReLU activation function, which was popular and could increase accuracy.

The dense layer with SoftMax activation function is commonly applied for data classification. The last output of this layer is divided into three categories, referring to the conditions of the Falan mangoes: normal mangoes (Normal), broken mangoes (Broken) and defective mangoes (Defective). The layers for training the recognition model had its architecture style as illustrated in Figure 4.

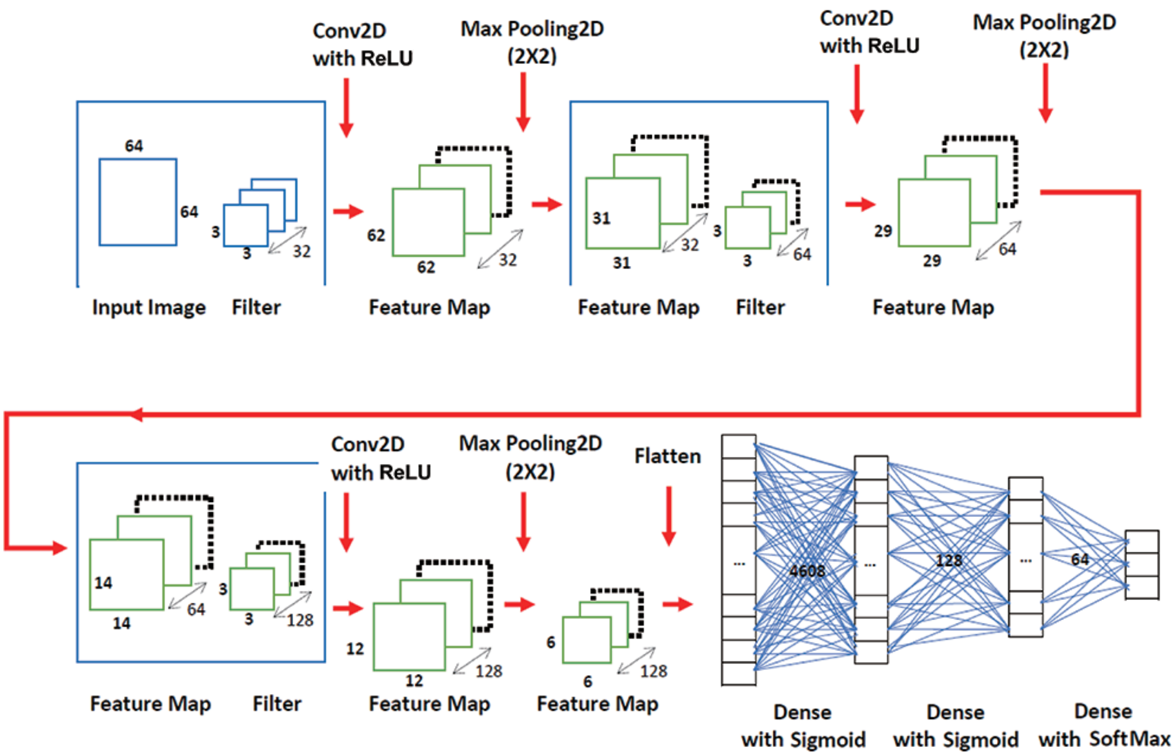


Figure 4. Layer architectures of recognition model for training

2.5. **Measuring the sizes of mangoes.** The segmented mango images were measured via pixel counting. The width (pixel) was multiplied by the height (pixel). The range of the pixels was divided into four groups from the average pixel values of each size of the mangoes: large (2,201-3,800), medium (1,643-2,200), small (800-1,642) and unknown.

2.6. **Classification measurements.** This research evaluated the efficiency of the deep learning network models by deploying a confusion matrix in conjunction with precision,

**Mango Size Measurement Algorithms**

---

**Input:** *mangoes image*
**Output:** *mango size (big, mid, small, unknown)*
***Initialization:***
1: Define *range of pixel* for *big; range of pixel* for *mid; range of pixel* for *small*
2: Red *mangoes image*
3: Measuring the *sizes* of *mangoes*

***Measuring the sizes of mangoes***
1: Red *mangoes image*
2: Segment *mangoes*
3: **For** each *mango* in *segmented mango images* **do**
  Get *width pixel* and *height pixel*
  *Each_size = width pixel* × *height pixel*
  **If** *Each_size in big* **then**
    count_big = +1
  **If** *Each_size in mid* **then**
    count_mid = +1
  **If** *Each_size in small* **then**
    count_small = +1
  **Else** count_unknown = +1
 ***End for***
4: Return *JSON data of mangoes sizes*

FIGURE 5. Mango size measurement algorithm

recall and F1 Score, calculated using the following formulas:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1\ \ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

$$Accuracy = \frac{TP + FN}{TP + FN + TN + FP} \tag{4}$$

where $TP$ = true positive, $TN$ = true negative, $FP$ = false positive and $FN$ = false negative.

Equation (1) calculates the precision, which compares whether the model's prediction is true and the actual result (TP) or the prediction is true but the actual result is false (FP), to evaluate the accuracy of the model. Equation (2) calculates the recall, which represents the model's prediction as true compared with the times that the actual result was true, to evaluate the accuracy of the model. Equation (3) calculates the F1 Score, which is the harmonic mean between precision and recall, as a metric used to measure the efficiency of the model and the accuracy of the test, as shown in Equation (4).

## 3. Experimental Results.

3.1. **Testing the efficiency of the model.** This research used images of mangoes from a mango farm. A total of 720 fruits were collected to serve as data for both the training and validation sets. These sets were utilized for training the model and monitoring its performance during training, as illustrated in Figure 7. Additionally, 30% of the data was set aside to create a testing set, enabling the evaluation of the model's performance.

We trained the model to identify Falan mangoes into three grades (broken, defective and normal) by segmenting the mango images into greyscale, excluding each from the pile and adjusting the images to $64 \times 64$ pixel for training and validation. Then, 404 images were used for training the model (96 broken mango images, 140 defective mango images and 168 normal mango images) and 100 for validation (23 broken mango images, 35 defective mango images and 42 normal mango images), as portrayed in Figure 6.

Concerning the training, the model was trained in 32 batch sizes. The training data underwent image augmentation to train the model and there were 200 training epochs. Approximately 200 s were spent using Google Colab Pro. The loss and accuracy graphs demonstrated fairly good values, as depicted in Figure 7. Loss and validation loss showed constant reduction trends with relatively close gaps. Likewise, for the accuracy values,

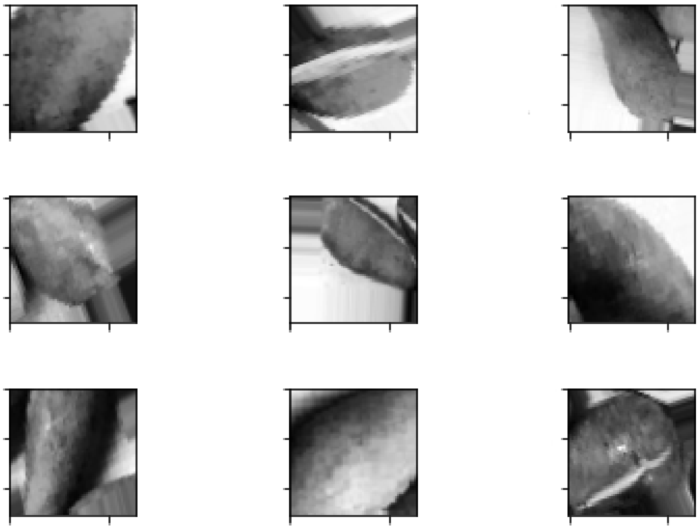

FIGURE 6. Displayed detected mangos

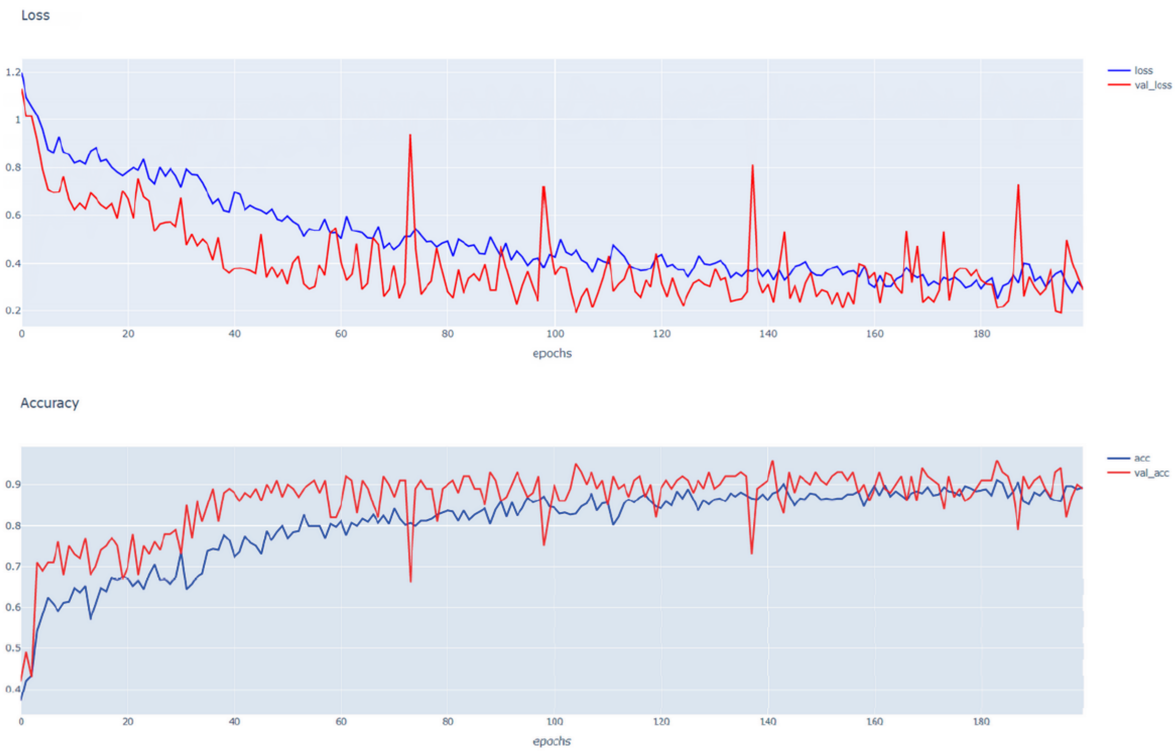


FIGURE 7. Training and validation accuracy of CNN architectures

both the accuracy and the validation accuracy values constantly increased with not too much of a gap.

The efficiency test of the prediction of mango grades using images of mangoes had various grades: 51 broken mangoes (Broken), 75 defective mangoes (Defective) and 90 normal mangoes (Normal). The model predicted the images as 44 broken (Broken), 59 defective (Defective) and 84 normal (Normal) mangoes, as exhibited in a confusion matrix in Table 1.

Table 1. Confusion matrix of the prediction of mango grades

| Actual \ Predicted | Broken | Defective | Normal |
|---|---|---|---|
| **Broken** | **44** | 1 | 6 |
| **Defective** | 9 | **59** | 7 |
| **Normal** | 2 | 4 | **84** |

The calculated results of the model classification into three grades and the results from the calculation of precision, recall and F1 Score for inspecting the model's accuracy using the divided data as the test set are indicated in Table 2.

Table 2. Precision, recall and F1 Score of the prediction of mango grades

| | Precision | Recall | F1 Score |
|---|---|---|---|
| **Broken** | 0.8000 | 0.8627 | 0.8302 |
| **Defective** | 0.9219 | 0.7867 | 0.8489 |
| **Normal** | 0.8698 | 0.9333 | 0.8984 |

The calculated accuracy value of the prediction of mango grades equals 85.64%.

The findings of the grading accuracy based on the mangoes' sizes via an image format (small, medium, and large), which were calculated using a confusion matrix, are presented in Table 3.

Table 3. Confusion matrix of mango grading based on size

| Actual \ Predicted | Big | Mid | Small |
|---|---|---|---|
| **Big** | **95** | 5 | 0 |
| **Mid** | 11 | **76** | 13 |
| **Small** | 0 | 28 | **72** |

The calculated results of the precision, recall, and F1 Score of the test are presented in Table 4.

Table 4. Precision, recall, and F1 Score of mango grading based on size

| | Precision | Recall | F1 Score |
|---|---|---|---|
| **Big** | 0.8962 | 0.9500 | 0.9223 |
| **Mid** | 0.6972 | 0.7600 | 0.7272 |
| **Small** | 0.8470 | 0.7200 | 0.7783 |

The calculated accuracy value of the grading based on sizes was 81.00%.

3.2. **Prototyping for Falan mango quality level identification.** The prototype for Falan mango quality level identification was developed via a web application and tested using mango images. On the home page of the web application, a 'Select File' button can be used to import mango image files for classification, as shown in Figure 8.

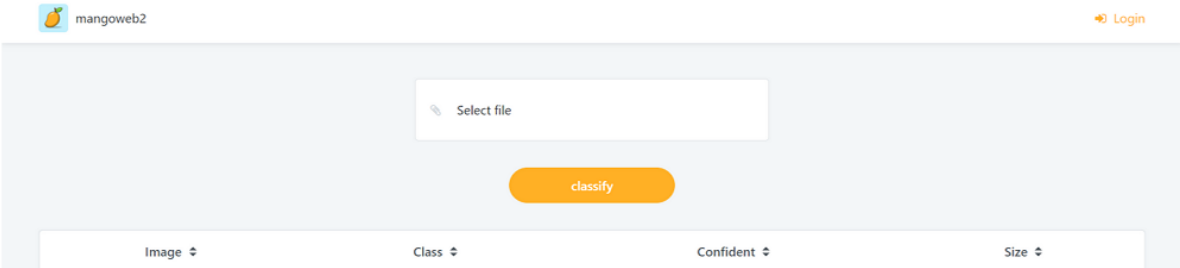


FIGURE 8. Web application's home page

The application screen displays the import file names, sizes, images, classes, confidence values, the total of each size and the total of each class, as illustrated in Figure 9.

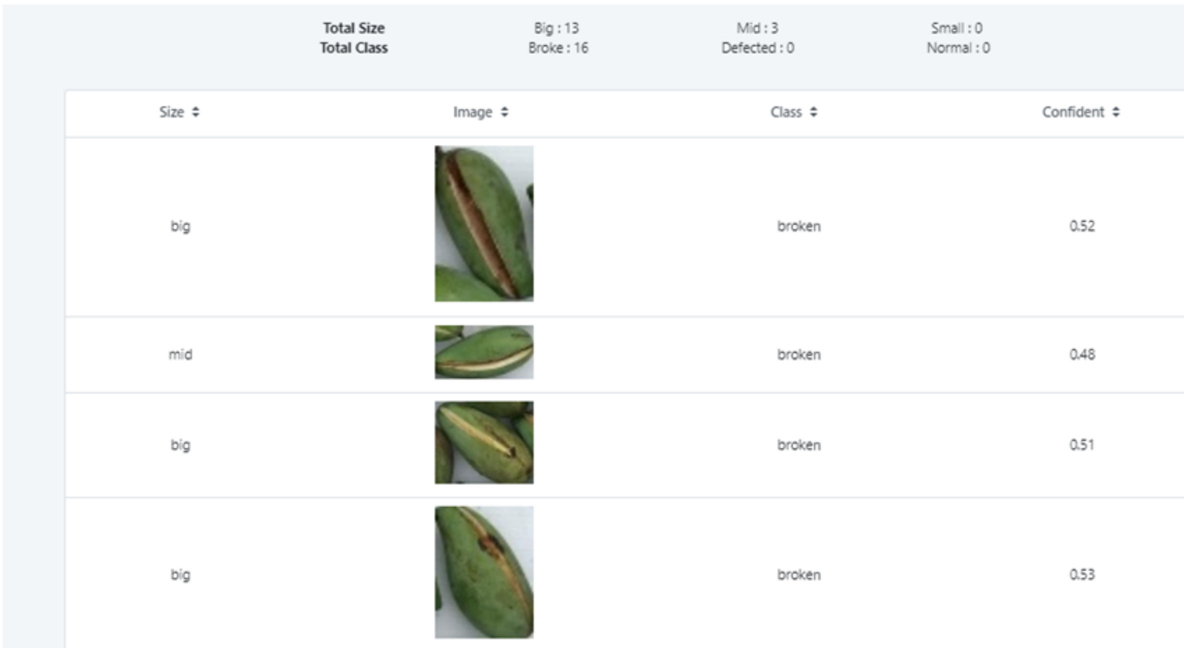


FIGURE 9. Screen displaying 'broken' mangoes

The results after the classification present mango images (for each mango), the class of each mango's grade (normal, broken and defective), a confidence value and the size of each mango (large, medium and small).

4. **Discussion and Conclusions.** This research aimed to indicate the quality levels and measure the sizes of each Falan mango from segmented mango images by receiving pile-of-mangoes images and segmenting them using Mask R-CNN before classifying them using CNN. The CNN model could classify the grades of mangoes into normal (Normal), broken (Broken) and defective (Defective). Its grading prediction accuracy was 85.64%, whereas its size measurement accuracy was 81.00%. Furthermore, we developed a web application for displaying the actual results of the mango classification. The users can import their image files and let the system grade and measure the sizes of the mangoes. This application can not only call API, grade and measure the mangoes but also display the total of each size and of each grade.

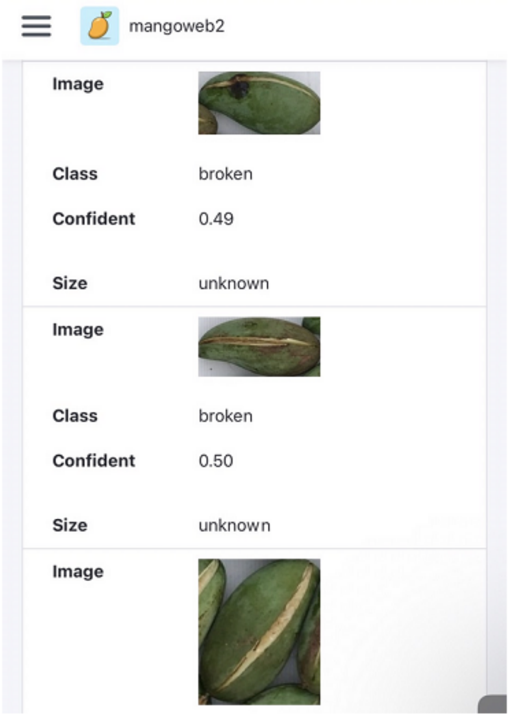


Figure 10. Results displayed via API on a smartphone

However, this research has limitations in terms of the data received, which were images taken from a height far from the ground at 120 cm. Therefore, the images may deviate from the designated height. When testing the images by counting their pixels, the results may be inaccurate. This can affect the accuracy of the testing in measuring the efficiency of mango size classification. Moreover, the mangoes should be placed on a white cloth according to the specified size (90 $\times$ 90 cm) with adequate light exposure as it can help leverage the accuracy of both segmentation and classification.

In summary, the quality level identification model can be potentially developed to grade mangoes for commercial purposes. In future studies, to adjust the processes of the application to be more suitable for farmers' trading activities, we collect additional image datasets of mangoes to improve the efficiency of classification and refine the method of size measurement for better practical use. This will promote the distribution of mango products in the future.

## REFERENCES

[1] Ministry of Agriculture and Cooperative (MOAC), Sukhothai Provincial Agriculture and Cooperatives Office, *Large Mangoes Plots Data for Commercial Agriculture Development Planning: Large Plots Farming Extensions*, 2018.
[2] Ministry of Agriculture and Cooperative (MOAC), Department of Agricultural Extension, *The Production of High-Quality Mangoes for Exportation*, 2023.
[3] O. Salidnam, *Study of Genetic Diversity of Mango by SSR Marker Technique*, Research Report, Maejo University, 2015.
[4] J. Blasco, S. Munera, N. Aleixos, S. Cubero and E. Molto, Machine vision-based measurement systems for fruit and vegetable quality control in postharvest, *Advances in Biochemical Engineering/ Biotechnology*, vol.61, pp.71-99, 2017.
[5] K. G. Liakos, P. Busato, D. Moshou, S. Pearson and D. Bochtis, Machine learning in agriculture: A review, *Sensors*, vol.18, no.8, 2674, 2018.
[6] A. Koirala, K. B. Walsh, Z. Wang and C. McCarthy, Deep learning – Method overview and review of use for fruit detection and yield estimation, *Computers and Electronics in Agriculture*, vol.162, pp.219-234, 2019.

[7] H. Kang and C. Chen, Fast implementation of real-time fruit detection in apple orchards using deep learning, *Computers and Electronics in Agriculture*, vol.168, pp.105-108, 2020.
[8] C. C. Ukwuoma, Z. Qin, M. B. B. Heyat, L. Ali, Z. Almaspoor and H. N. Monday, Recent advancements in fruit detection and classification using deep learning techniques, *Mathematical Problems in Engineering*, pp.1-29, 2022.
[9] S. Arampongsanuwat and O. Chaowalit, Application of deep convolutional neural networks for mangosteen ripeness classification, *ICIC Express Letters*, vol.15, no.6, pp.649-657, 2021.
[10] R. Nithya, B. Santhi, R. Manikandan, M. Rahimi and A. H. Gandomi, Computer vision system for mango fruit defect detection using deep convolutional neural network, *Foods*, vol.11, no.21, 3483, 2022.
[11] K. Albarrak, Y. Gulzar, Y. Hamid, A. Mehmood and A. B. Soomro, A deep learning-based model for date fruit classification, *Sustainability*, vol.14, 6339, 2022.
[12] H. Zhang, H. Sun, W. Ao and G. Dimirovski, A survey on instance segmentation: Recent advances and challenges, *International Journal of Innovative Computing, Information and Control*, vol.17, no.3, pp.1041-1053, 2021.
[13] P. Ganesh, K. Volle, T. F. Burks and S. Mehta, Deep orange: Mask R-CNN based orange detection and segmentation, *IFAC-PapersOnLine*, vol.52, no.30, pp.70-75, 2019.
[14] S. Pan and T. Ahamed, Pear recognition in an orchard from 3D stereo camera datasets to develop a fruit picking mechanism using mask R-CNN, *Sensors*, vol.22, no.11, 4187, 2022.
[15] M. Macías-Macías, H. Sánchez-Santamaria, C. J. García Orellana, H. M. González-Velasco, R. Gallardo-Caballero and A. García-Manso, Mask R-CNN for quality control of table olives, *Multimedia Tools and Applications*, pp.1-15, 2023.
[16] X. Liu, D. Zhao, W. Jia, W. Ji, C. Ruan and Y. Sun, Cucumber fruits detection in greenhouses based on instance segmentation, *IEEE Access*, vol.7, pp.139635-139642, 2019.
[17] B. Dhiman, Y. Kumar and M. Kumar, Fruit quality evaluation using machine learning techniques: Review, motivation and future perspectives, *Multimedia Tools and Applications*, vol.81, pp.16255-16277, 2022.