J. Trop. Agric. and Fd. Sc. 50(2022): 1 – 6

Detection and maturity index classification for tomato: Deep learning with computer vision-based

Nur Khalidah, Z*., Muhd Akhtar, M.T., Nuraini, A.A.S., Seri 'Aisyah, H., Siti Hawa, A.R., Mohamad Saiful Nizam, A., Mohd Daniel Hazeq, A.R., Mohd Shukry, H.B. and Muhamad Syahiran Afieff, A.

Engineering Research Centre, MARDI Headquarters, 43400 Serdang Selangor, Malaysia

Abstract

This paper is about the development of the tomato detection and maturity index classification model in a greenhouse using a deep learning technique. In total, two deep learning models were developed where the output of the first model will be the input for the second model. In this study, 2000 tomato image samples were captured for data acquisition. The annotated image samples were used to train the tomato detection model. On the other hand, the labeled image samples were used to train the tomato maturity index classification model. The confidence score of the first model was 0.958 whilst the accuracy of the second model was 92.33%. Lastly, both models were deployed and a dashboard was built where users can monitor the total distribution of tomatoes at one time.

Keywords: deep learning, maturity index classification, monitoring

Introduction

In line with the Industrial Revolution 4.0 (IR 4.0), the approach of using computer vision-based systems is seen to add efficiency compared to conventional methods. Traditionally, tomatoes are classified based on their physiological maturity by manual grading and picking. Manual grading depends on a person who has been specially trained in picking and sorting tomatoes. However, this technique has several disadvantages, such as low precision, labor-intensive and subjectivity.

Computer vision-based systems are focused on the theory of artificial intelligence (AI) that extract information from data and makes decisions based on patterns or trends from the data it learns. On top of that, this successful method has already been applied in vast field areas. In food and agricultural-based industry, this method has been applied to offer an automated solution such as for grading and sorting (Chaudhari et al. 2022; Fatima et al. 2022), crop monitoring (Bayazit et al. 2022; Lac et al. 2022), aerial surveying (Bouguettaya et al. 2022) and also pest and disease detection (Santhosh et al. 2022).

In a previous study, the Red-Green-Blue (RGB) images consisting of affected tomato leaves were classified and trained using Convolutional Neural Network (CNN) (Gonzales et al. 2021). This study explored four different CNN models namely MobileNetV2, NasNetMobile, Xception and MobileNetV3 for classification purposes. Among the four models, Xception is the best classifier with the best performance; however, its computational cost is higher as compared to the other models due to its number of parameters (Gonzales et al. 2021).

In a different study, the CNN model was successfully applied to extract the features for pest identification (Huang et al. 2022). This study used pre-trained models namely VGG16 and ResNet50 to extract the features and later combined the extracted features from the deep learning model with the three different machine learning models to perform the classification for tomato pests. Among the models, the ResNet50 with discriminant analysis model achieved the highest classification accuracy at 97.12% (Huang et al. 2022).

Through the IR 4.0 research and development (R&D) project initiative under the Ministry of Agriculture and Food Industries (MAFI), the tomato fruit maturity index classification model was developed to monitor tomato maturity in the greenhouse. Therefore, this study aimed to develop the tomato detection and classification model

Article history
Received: 19.10.22
Accepted: 30.11.22

Authors' full names: Nur Khalidah Zakaria, Muhd Akhtar Mohamad Tahir, Nuraini Ahmad Ariff Shah, Seri 'Aisyah Hassim, Siti Hawa Ahmad Ramli, Mohamad Saiful Nizam Azmi, Mohd Daniel Hazeq Abdul Rashid, Mohd Shukry Hassan Basri and Muhamad Syahiran Afieff Azman
Corresponding author: khalidah@mardi.gov.my

using a suitable method and algorithm that mainly focus on providing an automated solution for the end-users, which provides a simple and direct solution.

Materials and methods

This section elaborates in detail on the method used. To ease the explanation, the overall process methodology for this study can be divided into four main sections specifically image data acquisition, development of the tomato detection model, development of the tomato maturity classification model, and lastly deployment of the models. (*Figure 1*) shows the overall flowchart for this study.

Image data acquisition

The MT3 MAHA 2018 tomato variety was selected for this study. More than 100 plant samples were randomly selected based on its maturity at the Laman Sayur, MAEPS Serdang, Selangor. 2000 sample images were captured using a camera with a resolution of five megapixels. Image data acquisition was performed for two months, during the fruiting stage.

Development of the tomato detection model

In this study, the tomato detection model was developed to detect the presence of the tomato. After eliminating the damaged samples and low-quality images, the original image samples were annotated manually. Image annotation is the most common technique used to recognise an object (tomato) from the background for better understanding. During the annotation process, the tomatoes were annotated based on the shape of the fruits. During the annotation stage, several conditions were defined as follows:

a) All tomatoes must be annotated regardless of their index
b) All tomatoes must be annotated according to their full size
c) Blurred images of tomatoes must be excluded and ignored

Further, the images were randomly divided into two datasets with a ratio of 80:20 accordingly. The images in the training dataset were used to build the detection model, and the images in the testing dataset were used to verify the model using the confidence score. The confidence score shows the probability of the image being detected correctly by the algorithm. In other words, the score tells how efficiently the algorithm detects the presence of a tomato. The score is measured between 0 to 1 with a default threshold of 0.5, where the higher the score, the more confident the model is in predicting the correct result. In general, increasing the threshold will lower the sensitivity of the model to detect positive instances and improve the precision of the model by measuring the quality of a positive prediction made by the model.

The annotated training dataset was resized and fed to train the detection model. In this study, a pre-trained model named YOLOv3 was used. A pre-trained model is a deep learning model that is previously trained on large data to solve a problem. In this study, YOLOv3 was chosen because it is one of the best one-stage object detection models that works for dense object detection. Subsequently, the pre-trained model was applied to train the training dataset. During the training process, the selection of hyperparameters such as learning rate, LR (tuned to 0.001 and 0.005) and the number of epochs

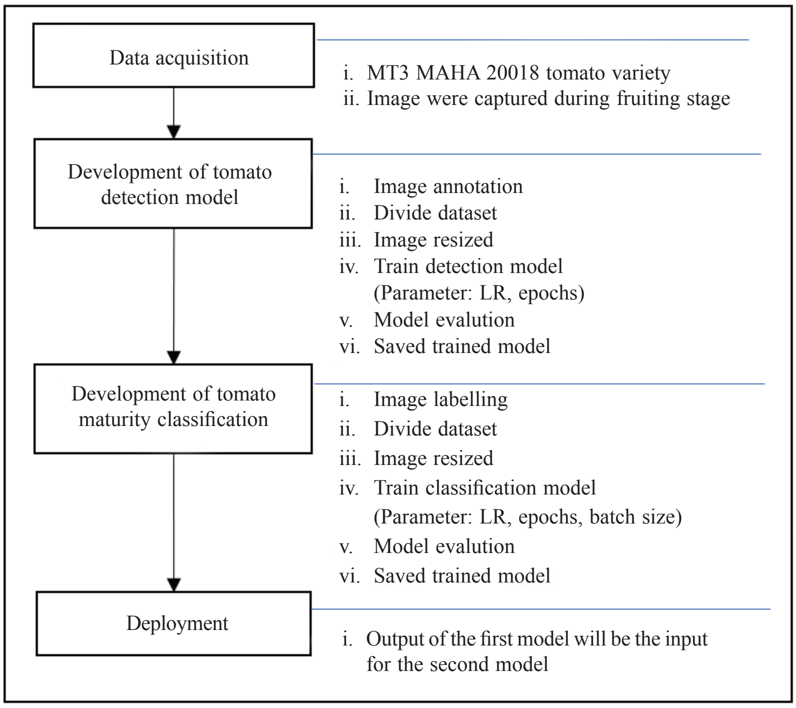

Figure 1. Overall flowchart

(tuned to 20, 30, 40, 50 and 60) was tested to increase the confidence score of the model. The evaluation of the model was performed using the annotated testing dataset and measured by the confidence score of detected objects.

Development of the tomato maturity classification model

The tomato maturity index classification model was developed to automatically classify the tomato maturity index. For the purpose of this study, 150 tomato image samples per maturity index were required. Original image samples were cropped and labeled manually by the agronomist at MARDI with indices namely Index 1, Index 2, Index 3, Index 4, Index 5 and Index 6. Indexes are categorised by color and differences in maturity samples as shown in *Figure 2*.

On top of that, the image augmentation technique was applied to balance the number of image samples for each group index. It is a technique of altering the existing data to create more data for the model training process. Next, the labeled datasets were randomly divided into two datasets with a ratio of 80:20 accordingly. The images in the training dataset were used to build the classification model and the images in the testing dataset were used to verify the model. For the classification model, a pre-trained model namely MobileNet was used as the base model for the classification.

Further, the selection of hyperparameters such as learning rate, LR (tuned to 0.001 and 0.005), number of epochs (tuned to 20, 30, 40, 50 and 60), batch size (tuned to 128, 266 and 512) were tested to increase the percentage of accuracy of the model during training process. The evaluation of the model was performed using the testing dataset and measured by the accuracy based on the confusion matrix of the classification. *Table 1* describes the confusion matrix that consists of actual and predicted information (Cho et al. 2022).

Based on *Table 1*, TP is a positive value that has been predicted as true by the classifier and FP is defined as a positive value that has been predicted as false by the classifier. On the other hand, TN is defined as a negative value that has been predicted as true by the classifier whilst FP is defined as a negative value that has been defined as false by the classifier. Thus, the accuracy of the model can be defined in equation (*1*):

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \text{ x } 100\% \qquad (1)$$

Where, TP = True Positive;
TN = True Negative;
FP = False Positive;
FN = False Negative

Figure 2. Tomato maturity indexes

Table 1. Confusion table

		Predicted	
		Positive	Negative
Actual	Positive	True Positive, TP	False Negative, FN
	Negative	False Positive, FP	True Negative, TN

Deployment of models

In the last stage, the deployment of both models was executed. In this study, the models were stacked one by another, where the first model's output will be the second model's input. The first model is the tomato detection model and the second model is the tomato maturity index classification model.

Results and discussion

This section will discuss the results obtained from image data acquisition from both models. During image data acquisition, 2000 tomato image samples were collected. For the purpose of this study, the original image samples that have been annotated were used in the development of the tomato detection model whilst the original image samples that have been cropped were used in the development of the tomato maturity index classification model.

Upon eliminating the damaged samples and low-quality images, the original image samples were annotated manually. *Figure 3* shows the annotated images based on the conditions mentioned previously. Next, the annotated images were resized to 416 x 416 pixels and divided into two datasets. The training dataset consists of 1600 images whilst the testing dataset consists of 400 images respectively. Ten detection models were successfully developed based on the hyperparameters setup. To simplify the result, only the three best models were chosen. These models were chosen based on the highest confidence score and the results were tabulated in *Table 2*.

The leaning rate (LR) is an optimisation parameter that controls the weight of the model concerning the loss gradient. Generally, it defines how quickly the algorithm updates the weight it has learned. On the other hand, the number of epochs means that one complete training dataset had completely passed through the training algorithm. Thus, as the number of epochs increased, more weights were updated to achieve an optimal fitting curve. Referring to *Table 2*, the increase in the number of epochs significantly improved the confidence score of Model 1 to 0.958 and Model 2 to 0.950 given by the same LR. On the other hand, Model 2 and Model 3 gave significant output difference of 0.03 in confidence score for the same number of epochs, thus showing that smaller LR causes the model to quickly achieve its losses gradient. From *Table 2*, the highest confidence score was obtained by Model 1 with 0.958, followed by Model 2 with 0.950 and lastly Model 3 with 0.922. Therefore, Model 1 was chosen as the best tomato model detection.

The original image samples were cropped and labelled manually by the agronomist at MARDI as shown in *Figure 4*. Upon eliminating the noise image samples for example those with low resolution and blurred images, the original image samples were randomly chosen for augmentation. These augmented samples were added due to the small and unbalanced number of image samples for each index. It is necessary to balance the image samples for the distribution of the training and testing datasets. The imbalance distribution of samples for each group can cause the dataset to be biased towards a class or group during algorithm training. *Table 3* tabulates the number of images for the tomato image samples. Next, 1090 image samples were used in this study and divided in 80:20 ratios respectively. A total of 872 samples were used for the training dataset and 218 samples for the testing dataset. The distribution of the image samples of each index was tabulated in *Table 3*.

Thirty classification models were successfully developed based on the hyperparameters setup as mentioned in the previous section. To simplify the results, only the best models were chosen based on the highest accuracy given by the confusion matrix table in *Figure 5*.

(a) (b) (c)

Figure 3. The example of annotated images

Table 2. Hyperparameters and confidence level for the tomato detection model

Hyperparameters/model	Learning rate	Number of epochs	Confidence score
Model 1	0.001	50	0.958
Model 2	0.001	40	0.950
Model 3	0.005	40	0.922

Table 3. The total tomato image samples

	Index 1	Index 2	Index 3	Index 4	Index 5	Index 6	Total
Original	211	73	45	79	422	950	1780
Augmented	-	200	200	200	-	-	600
Selected	195	165	170	180	190	190	1090
Training data set	156	132	136	144	152	152	872
Testing data set	39	33	34	36	38	38	218

Figure 4. The example of the cropped and labelled image sample

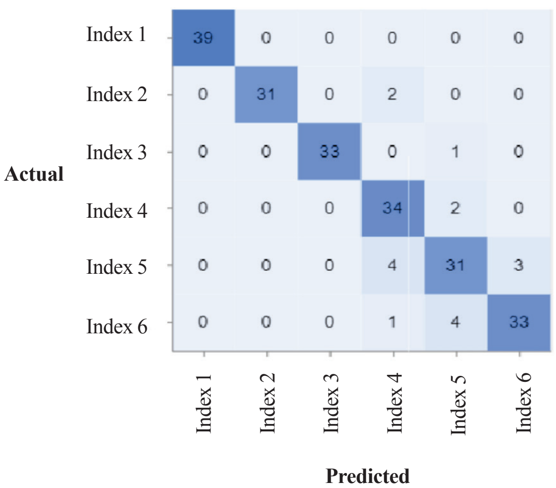

Figure 5. The confusion matrix of the tomato maturity index classification model

From the confusion matrix, 39 image samples from Index 1 are correctly classified in Index 1. On the other hand, two image samples from Index 2 are mistakenly classified into Index 4. Meanwhile, 33 out 34 image samples from Index 3 were correctly classified to their group whilst one sample was falsely classified to Index 5. Next, two image samples from Index 4 were wrongly classified into Index 5. Seven images from Index 5 were misclassified into Index 4 and Index 6 respectively. On top of that, 33 out of 38 image samples from Index 6 are correctly classified to their group whilst one sample is misclassified into Index 4 and four samples in Index 5. *Table 4* tabulates the accuracy per group attained by the confusion matrix in *Figure 5*. From the table, Index 1 attained 100.00% of accuracy as the group did not have misclassification. On the other hand, Index 2, Index 3 and Index 4 achieved accuracies above 90%. These groups have less than three image misclassification and thus attained accuracies with 94.00%, 97.00% and 94.00% respectively. Subsequently, Index 5 and Index 6 achieved accuracies with 82.00% and 87.0% accordingly. Overall, the average accuracy for this model is 92.33% with the number of epochs equal to 60, the batch size equal to 256 and LR equal to 0.005.

Table 4. The accuracy per group for the tomato maturity index classification model

Index	Number of image samples	Accuracy
Index 1	39	1.00
Index 2	33	0.94
Index 3	34	0.97
Index 4	36	0.94
Index 5	38	0.82
Index 6	38	0.87
Total	218	0.92

In the last stage, the deployment of both models were executed. The models developed are not stand alone, as the output from the tomato detection model will be fed as input for the second model; the tomato maturity index classification model. A dashboard for monitoring the maturity of the tomato fruit maturity index was also developed. From the dashboard, users can identify the total distribution of tomatoes at one time.

Conclusion

In conclusion, it was proven that deep learning with computer vision-based can improve and automate the process of tomato detection and tomato maturity index classification. In addition, the transfer learning method significantly enhanced and speed up the training process. This method also demonstrated its suitability for the given image samples. Future work can include implementations with the development of growth monitoring and pest and disease control models. These work can also be added to the dashboard as an additional function that significantly enhances the dashboard interface.

Acknowledgment

The authors would like to thank the Ministry of Agriculture and Food Industries (MAFI) and Malaysian Agricultural Research and Development Institute (MARDI) for the financial support and laboratory facilities for this study.

References

Bayazit, U., Altilar, D.T. and Bayazit, N.G. (2022). Classification and phenological staging of crops from in situ image sequences bydeep learning. *Turkish Journal of Electrical Engineering and Computer Sciences* 30(4): 1299 – 1316

Bouguettaya, A., Zarzour, H., Kechida, A. and Taberkit, A.M. (2022). Deep learning techniques to classify agricultural crops through UAV imagery: a review. *Neural Computing and Applications*, 1 – 26

Chaudhari, D. and Waghmare, S. (2022). Machine Vision Based Fruit Classification and Grading–A Review. *ICCCE 2021*: 775 – 781

Cho, B.H., Kim, Y.H., Lee, K.B., Hong, Y.K. and Kim, K.C. (2022). Potential of Snapshot-Type Hyperspectral Imagery Using Support Vector Classifier for the Classification of Tomatoes Maturity. *Sensors* 22(12): 4378

Fatima, S. and Seshashayee, M. (2022). Feature fusion of fruit image categorization using machine learning. *International Journal of Nonlinear Analysis and Applications, 13* (Special Issue for selected papers of ICDACT-2021): 71 – 76

Gonzalez-Huitron, V., León-Borges, J.A., Rodriguez-Mata, A., Amabilis-Sosa, L.E., Ramírez-Pereda, B. and Rodriguez, H. (2021). Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4. *Computers and Electronics in Agriculture* 181: 105951

Huang, M.-L., Chuang, T.-C. and Liao, Y.-C. (2022). Application of transfer learning and image augmentation technology for tomato pest identification. *Sustainable Computing: Informatics and Systems* 33: 100646

Lac, L., Keresztes, B., Louargant, M., Donias, M. and Da Costa, J.-P. (2022). An annotated image dataset of vegetable crops at an early stage of growth for proximal sensing applications. *Data in Brief* 42: 108035

Santhosh Kumar, P., Balakrishna, R. and Vinod Kumar, K. (2022). *Review on disease detection of plants using image processing and machine learning techniques.* Paper presented at the AIP Conference Proceedings