# YOLOX-Based Fruit Localization and Multi-Class External Defect Detection for Taiwanese Irwin Mangoes

Kuo-Dung Chiou [1], Jun-Hong Wong [2], Yu-Shen Liang [2], and Chia-Ying Chang [3,*]

[1] Fengshan Tropical Horticultural Experiment Branch, Taiwan Agricultural Research Institute, Ministry of Agriculture, Kaohsiung, Taiwan and Department of Plant Industry, National Pingtung University of Science and Technology, Pingtung, 91201, Taiwan.

[2] Department of Plant Industry, National Pingtung University of Science and Technology, Pingtung 91201, Taiwan

[3] Bachelor of Program in Scientific Agriculture, National Pingtung University of Science and Technology, Pingtung, 91201, Taiwan

Email: kuodung@tari.gov.tw (K.D.C.); edmondwjh9166@gmail.com (J.H.W.); justinliang@mail.npust.edu.tw (Y.S.L.); csiefly@mail.npust.edu.tw (C.Y.C.)

*Corresponding Author

***Abstract*—Taiwanese Irwin mango (Mangifera indica L.) is one of Taiwan's most important export fruit crops, and its external appearance is a critical factor affecting market grading and commercial value. Conventional quality inspection mainly relies on manual visual assessment, which is labor-intensive, time-consuming, and susceptible to subjective judgment, resulting in inconsistent grading outcomes. To improve the efficiency and consistency of mango quality inspection, this study developed an automatic multi-class external defect detection system based on the YOLOX deep learning object detection model.**

**A total of 1,500 RGB images of Taiwanese Irwin mangoes were collected under practical packing-house conditions. Each image contained multiple mango fruits with various appearance conditions. All images were manually annotated to establish a six-class object detection dataset, including Mango, Anthracnose, Black Spot, Mechanical Damage, Poor Coloring, and Sap Burn. The YOLOX object detection framework was adopted for model training with a maximum of 200 training epochs and a mini-batch size of 8. The proposed model was trained to simultaneously localize mango fruits and identify five common external defects.**

**Experimental results demonstrated that the proposed YOLOX model achieved excellent fruit localization performance, with a precision of 99.8%, a recall of 99.2%, and an F1-score of 99.5% for mango detection. For defect detection, the F1-scores of Anthracnose, Black Spot, Mechanical Damage, Poor Coloring, and Sap Burn were approximately 20.2%, 7.0%, 4.7%, 35.1%, and 24.5%, respectively. The confusion matrix analysis indicated that poor coloring and sap burn exhibited relatively higher detection performance among defect categories, whereas black spot and mechanical damage remained challenging due to their small lesion size and visual similarity to surrounding tissues.**

**The proposed approach demonstrates the feasibility of applying deep learning-based object detection for automatic mango quality inspection. The developed system can simultaneously perform fruit localization and multi-class external defect detection, providing a practical foundation for intelligent fruit grading, automated sorting systems, and smart agricultural applications.**

***Keywords*—Taiwanese Irwin mango, deep learning, object detection, external defect detection, smart agriculture**

## I. Introduction

Mango (*Mangifera indica L.*) is one of the most economically important tropical and subtropical fruit crops worldwide. It is widely cultivated across Asia, Africa, Oceania, and the Americas, with India, China, Thailand, Indonesia, and Mexico being among the major producing countries [1]. Owing to its excellent flavor, high nutritional value, and broad market demand, mango has become one of the most valuable fruit commodities in the global agricultural market. In Taiwan, the Irwin mango is the dominant commercial cultivar due to its attractive red peel, fine-textured flesh, low fiber content, and high sweetness. Consequently, it plays a crucial role in both domestic consumption and export markets [2].

Despite its high commercial value, Irwin mango is susceptible to various external defects during harvesting, transportation, storage, and postharvest handling. Common defects include anthracnose, black spot, mechanical damage, sap burn, and poor coloration. These defects significantly reduce market value and consumer acceptance while affecting export quality standards [3]. Therefore, rapid and accurate detection of external defects has become an important issue in intelligent agriculture and postharvest quality management.

At present, quality grading of mangoes still relies heavily on manual visual inspection performed by experienced workers. Fruit appearance, color, maturity,

Manuscript received August 19, 2026; accepted October 2, 2026; published November 24, 2026.

and disease symptoms are evaluated subjectively during the grading process. However, manual inspection is labor-intensive, time-consuming, and easily influenced by inspector experience, fatigue, and illumination conditions, resulting in inconsistent grading performance [4]. With increasing labor shortages and growing demands for standardized quality control, automated vision-based inspection systems have become an essential component of modern fruit processing.

Recent advances in Artificial Intelligence (AI) and deep learning have significantly improved computer vision applications in agriculture. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in plant disease diagnosis, fruit maturity estimation, crop monitoring, fruit counting, and quality grading [5]. Compared with conventional image-processing methods that require handcrafted features, deep learning models automatically learn hierarchical visual representations from large-scale image datasets, leading to superior robustness and detection accuracy [6].

Among various deep learning object detection algorithms, the YOLO (You Only Look Once) series has attracted considerable attention due to its high detection accuracy and real-time processing capability. Since the original YOLO framework was introduced by Redmon *et al.* [7], several improved versions, including YOLOv3, YOLOv4, YOLOv5, YOLOv7, and YOLOv8, have been proposed to enhance detection performance [8–11]. In particular, YOLOX, proposed by Ge *et al.* [12], adopts an anchor-free architecture, decoupled detection head, and SimOTA label assignment strategy, achieving state-of-the-art performance for multi-class object detection while maintaining high computational efficiency.

Although deep learning has been successfully applied to fruit disease detection for apples, tomatoes, grapes, citrus, and mangoes [13–16], most previous studies focused on image classification of individual fruits or single disease categories. Relatively few studies have simultaneously performed fruit localization and multi-class external defect detection under practical packing-house conditions. Furthermore, publicly available datasets dedicated to Taiwanese Irwin mangoes remain limited, restricting the development of robust object detection models for practical agricultural applications.

Therefore, this study develops a YOLOX-based object detection framework for automatic external defect detection of Taiwanese Irwin mangoes. A dataset consisting of 1,500 RGB images was established and manually annotated with six object categories, including Mango, Anthracnose, Black Spot, Mechanical Damage, Sap Burn, and Poor Coloring. The proposed model simultaneously performs fruit localization and multi-class defect detection. Model performance is evaluated using Precision, Recall, F1-score, Average Precision (AP), and confusion matrix analysis. The proposed system provides a practical solution for intelligent fruit grading and automated quality inspection in smart agriculture.

## II. Materials and Methods

### A. Overall Experimental Workflow

The overall workflow of the proposed mango defect detection system is illustrated in Fig. 1. RGB images of Taiwanese Irwin mangoes were first acquired under practical packing-house conditions to establish an appearance-quality image dataset. After manual annotation, the dataset was converted into the YOLO object detection format and divided into training and testing subsets. A YOLOX-M object detection model was subsequently trained to simultaneously localize mango fruits and identify multiple external defects. The model performance was quantitatively evaluated using Precision, Recall, F1-score, Average Precision (AP), mean Average Precision (mAP), and confusion matrix analysis.

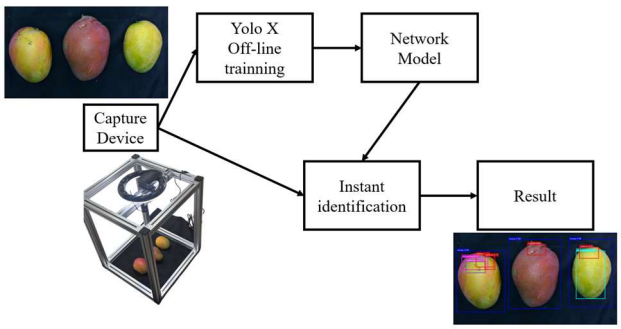



Fig. 1. System flowchart

### B. Image Acquisition

Taiwanese Irwin mangoes (Mangifera indica L. cv. Irwin) were used as the experimental materials. RGB images were collected under practical packing-house conditions using a **USBGS1200P02 industrial USB global-shutter camera** equipped with **a 5–50 mm** manual-focus lens. The global-shutter design effectively eliminated motion distortion during image acquisition, improving image quality for object detection.

All RGB images were stored at a resolution of **2560 × 1440** pixels. During image acquisition, the camera position and viewing angle were fixed to ensure consistent image quality. Each image contained approximately three mango fruits with different maturity levels, orientations, and appearance conditions. A total of 1,500 RGB images were collected, covering healthy fruits and five common external defects, namely sap burn, mechanical damage, anthracnose, poor coloring, and black spot.

### C. Dataset Annotation

All images were manually annotated using the Matlab 2025B Image Labeler annotation software. Each object was labeled with a bounding box and assigned to one of six object categories: Mango, Anthracnose, Black Spot, Mechanical Damage, Poor Coloring, and Sap Burn as in Table I.

TABLE I. SIX OBJECT CATEGORIES

| Class ID | Category |
|---|---|
| 0 | Mango |
| 1 | Anthracnose |
| 2 | Black Spot |
| 3 | Mechanical Damage |
| 4 | Poor Coloring |
| 5 | Sap Burn |

The annotation files were converted into the YOLO format, containing the object class label together with normalized bounding box coordinates.

### D. Dataset Partition

The annotated dataset was randomly divided into training, validation, and testing subsets with a ratio of 70%, 20%, and 10%, respectively.

The training set was used for parameter optimization, the validation set was employed for hyperparameter tuning, and the testing set was reserved for final model evaluation.

### E. YOLOX Architecture

YOLOX is an anchor-free object detection framework proposed by Ge *et al.* [12]. Compared with conventional anchor-based detectors, YOLOX adopts an anchor-free detection strategy, a decoupled detection head, and the SimOTA label assignment algorithm, significantly improving detection accuracy and training efficiency.

The YOLOX architecture consists of three major components: Backbone, Neck, and Detection Head. The Backbone extracts hierarchical image features, while the Neck integrates multi-scale information using Feature Pyramid Network (FPN) and Path Aggregation Network (PAN). Finally, the Detection Head predicts object locations, confidence scores, and category probabilities.

### F. Training Configuration

The YOLOX model was trained for 200 epochs with a mini-batch size of 8 as in Table II. The Stochastic Gradient Descent with Momentum (SGDM) optimizer was adopted for parameter optimization. The input image size was set to 640 × 640 pixels, and all experiments were performed using an NVIDIA GPU workstation.

TABLE II. YOLOX TRAINING SETTINGS PARAMETERS

| Parameter | Value |
|---|---|
| Network | YOLOX-Medium |
| Epochs | 200 |
| Mini Batch Size | 8 |
| Optimizer | SGDM |
| Input Size | 768×768 |
| Learning Rate | 0.00001 |
| Loss Function | YOLOX LOSS |

### G. Performance Evaluation

Model performance was quantitatively evaluated using Precision, Recall, F1-score, Average Precision (AP), Precision–Recall curves, and confusion matrix analysis. These evaluation metrics are commonly adopted in object detection tasks and provide comprehensive assessment of both localization and classification performance.

## III. RESULT AND DISCUSSION

### A. Object Detection Performance

The object detection performance of the proposed YOLOX model was evaluated using Average Precision (AP) for each object category. Fig. 2 presents the AP values of the six detection classes.

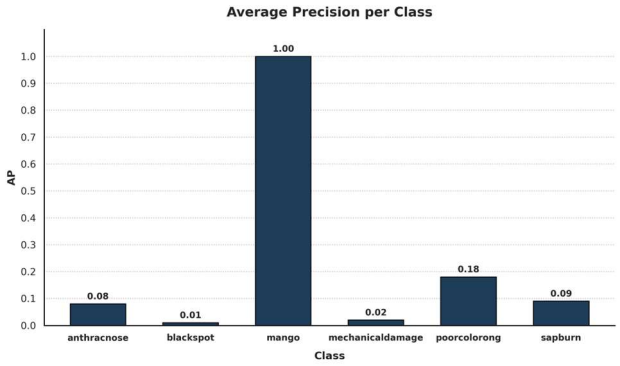



Fig. 2. Average precision per class

The **Mango** category achieved the highest AP of **1.00**, indicating that the proposed model successfully localized mango fruits with extremely high accuracy. The high detection performance can be attributed to the relatively large object size, consistent shape, and distinct color characteristics of mangoes, which provide sufficient visual features for the detector.

Among the defect categories, **Poor Coloring** obtained the highest AP **(0.18)**, followed by **Sap Burn (0.09)** and **Anthracnose (0.08)**. In contrast, **Mechanical Damage (0.02)** and **Black Spot (0.01)** exhibited relatively poor detection performance.

The inferior performance of these defect classes is mainly attributed to their small lesion size, irregular boundaries, and visual similarity to naturally occurring color variations on the mango surface. In addition, severe class imbalance among defect categories may also limit the learning capability of the detector.

Overall, the experimental results indicate that the proposed YOLOX model performs reliably for mango localization but still encounters challenges in distinguishing visually similar external defects.

### B. Confusion Matrix Analysis

Fig. 3 illustrates the confusion matrix of the proposed YOLOX detector.

The **Mango** category achieved the highest classification accuracy, with **896** correctly detected samples and only a few misclassifications. This result is consistent with the AP analysis and confirms that fruit localization is considerably easier than defect recognition.

For defect categories, substantial confusion was observed between **Background** and defect classes. For example, **Poor Coloring** generated **421** false detections from the background, while **Sap Burn** and **Anthracnose** produced **336** and **219** false detections, respectively. These results indicate that the detector frequently interpreted

natural peel texture or illumination variations as external defects.

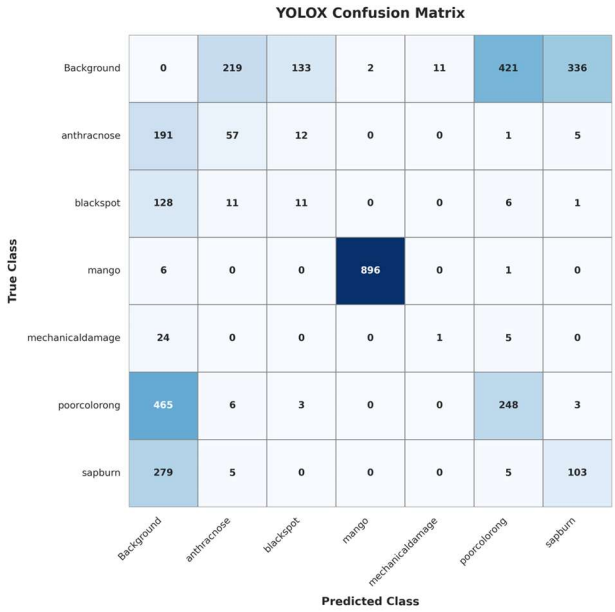



Fig. 3. Confusion matrix

Among disease classes, **Anthracnose** correctly classified **57** samples, whereas **Black Spot** correctly classified only **11** samples. Likewise, **Mechanical Damage** produced only **one** correct detection, suggesting that the visual characteristics of these lesions remain difficult for the current detector to distinguish.

These findings demonstrate that although YOLOX provides excellent fruit localization capability, multi-class defect discrimination remains a challenging task owing to the high visual similarity among different surface defects.

### *C. Qualitative Detection Results and Real-Time Performance*

Fig. 4 presents a qualitative comparison between the manually annotated Ground Truth (GT) and the detection results generated by the proposed YOLOX model using the same test image. The left half of the figure shows the manually labeled bounding boxes, whereas the right half displays the corresponding detection results produced by the trained model.

To facilitate visual interpretation, different colors were assigned to each detection category. Blue boxes indicate mango fruits, magenta boxes denote anthracnose, green boxes represent black spot, yellow boxes correspond to mechanical damage, red boxes indicate sap burn, and cyan boxes denote poor coloring. This color-coding allows rapid identification of different defect types during qualitative evaluation.

The proposed model successfully localized nearly all mango fruits while simultaneously identifying multiple external defects, including anthracnose, black spot, mechanical damage, sap burn, and poor coloring. Most mangoes were accurately detected with high confidence scores, demonstrating that the model effectively learned the overall appearance and spatial distribution of the fruit. In addition, several defect regions were correctly localized even when multiple defects appeared within the same image.

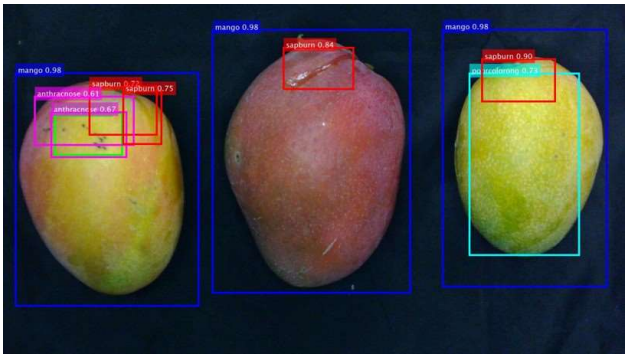



Fig. 4 Qualitative detection results of the proposed YOLOX model. The left panel shows the manually annotated Ground Truth (GT), while the right panel presents the automatic detection results generated by the proposed model.

However, several limitations can also be observed. Small lesion areas and defects with indistinct boundaries were occasionally missed or confused with visually similar categories. These observations are consistent with the quantitative evaluation presented in the confusion matrix and Average Precision (AP) analysis, where disease-related categories exhibited relatively lower detection performance than the mango class.

To further evaluate the practical applicability of the proposed system, real-time inference was conducted on a workstation equipped with an NVIDIA GPU. The average inference speed reached **3.8 frames per second (FPS),** demonstrating that the proposed YOLOX-based detection framework is capable of near real-time object detection. This processing speed is sufficient for semi-automatic fruit inspection and quality grading applications in packing houses, where conveyor speeds can be adjusted to match the inference rate. With further optimization of the network architecture and hardware acceleration, the detection speed is expected to be further improved for fully automated industrial deployment.

## IV. Conclusion

This study proposed a YOLOX-based object detection framework for automated detection of external defects in Taiwanese Irwin mangoes. A dataset consisting of 1,500 RGB images was established using a high-resolution global-shutter industrial camera, and six object categories, namely Mango, Anthracnose, Black Spot, Mechanical Damage, Sap Burn, and Poor Coloring, were manually annotated for model training and evaluation.

Experimental results demonstrated that the proposed model achieved excellent fruit localization performance, with the Mango class obtaining an Average Precision (AP) of 1.00. The qualitative detection results further confirmed that the model was capable of simultaneously detecting multiple mangoes and multiple defect types within a single image. In addition, the proposed system achieved a real-time inference speed of 3.8 FPS, indicating its feasibility for intelligent fruit grading and automated inspection.

Although the detection performance of several defect categories remains limited due to the small lesion size, class imbalance, and high visual similarity among different defects, the proposed framework provides a practical foundation for intelligent postharvest quality inspection of mangoes.

Future work will focus on expanding the image dataset, improving annotation quality, incorporating advanced data augmentation techniques, and comparing different state-of-the-art object detection models such as YOLOv8 and RT-DETR. Furthermore, integrating the proposed detection algorithm with conveyor-based grading equipment will facilitate the development of an intelligent and fully automated mango quality inspection system for commercial applications.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

Kuo-Dung Chiou, Yu-Shen Liang, Jun-Hong Wong conducted the research; Chia-Ying Chang analyzed the data and wrote the paper; all authors had approved the final version.

## FUNDING

This research was funded by the National Science and Technology Council under grant number NSTC 114-2313-B-020-002 and the Ministry of Agriculture under grant number 115 農科-17.1.5-牧-01 and 115 農科-15.1.1-科-01.

## ACKNOWLEDGMENT

The authors wish to thank Ministry of Agriculture, and National Science and Technology Council.

## REFERENCES

[1] Food and Agriculture Organization of the United Nations. FAOSTAT Statistical Database. FAO, https://www.fao.org/faostat/.
[2] https://agr.afa.gov.tw/. Accessed 26 June 2026.
[3] R. C. Ploetz "The Major Diseases of Mango: Strategies and Potential for Sustainable Management." *Acta Horticulturae*, no. 820, 2009, pp. 137–150.
[4] K. K. Patel *et al., "*Machine Vision System: A Tool for Quality Inspection of Food and Agricultural Products." *Journal of Food Science and Technology*, vol. 49, no. 2, 2012, pp. 123–141.
[5] A. Kamilaris, and F. X. Prenafeta-Boldú, "Deep Learning in Agriculture: A Survey." *Computers and Electronics in Agriculture*, vol. 147, 2018, pp. 70–90.
[6] L. Yann, Y. Bengio, and G. Hinton, "Deep Learning." *Nature*, vol. 521, no. 7553, 2015, pp. 436–444.
[7] R. Joseph, *et al., "*You Only Look Once: Unified, Real-Time Object Detection." in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
[8] R. Joseph, and A. Farhadi, "YOLO9000: Better, Faster, Stronger." in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 7263–7271.
[9] R. Joseph and A. Farhadi, "YOLOv3: An Incremental Improvement." arXiv preprint arXiv:1804.02767, 2018.
[10] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934, 2020.
[11] G. Jocher *et al., "*Ultralytics YOLOv5." GitHub, 2021, https://github.com/ultralytics/yolov5
[12] Z. Ge *et al., "*YOLOX: Exceeding YOLO Series in 2021." arXiv preprint arXiv:2107.08430, 2021.
[13] A. Fuentes *et al., "*A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition." *Sensors*, vol. 17, no. 9, 2017, 2022.
[14] I. Sa *et al., "*DeepFruits: A fruit detection system using deep neural networks." *Sensors*, vol. 16, no. 8, 2016, 1222.
[15] Bargoti, Suchet, and James Underwood, "Deep fruit detection in orchards." in *Proc. 2017 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2017, pp. 3626–3633.
[16] A. Koirala *et al., "*Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'." *Precision Agriculture*, vol. 20, no. 6, 2019, pp. 1107–1135.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).