JIAISE

Journal of Integrated Artificial Intelligence Science and Engineering

Vol. 1, No. 2, July 2025, pp. 29-34, DOI: 10.59190/jiaise.v1i2.328

PUBLISHED BY

LSPRAR

Detection of ripeness level of oil palm fresh fruit bunches using YOLOv4 model in automated harvesting system: A review

Nisa Arpyanti

Department of Physics, Universitas Riau, Pekanbaru 28293, Indonesia

ABSTRACT

This research discusses the development of an automated system for detecting ripe oil palm fruits (fresh fruit bunch, FFB) using computer vision and artificial intelligence (AI) technology. The main objective of this study is to improve efficiency and productivity in the oil palm harvesting process by adopting the latest technology. Several related studies have developed non-destructive methods for classifying the ripeness of FFB. This research utilizes the YOLOv4 deep learning model to detect ripe FFB in real-time. Visual data of FFB is obtained by capturing images using an Intel Realsense D435 camera on oil palm trees. The data is then labeled and divided into training and validation sets. Through evaluation, the YOLOv4 model with a network input size of 512 × 512 was found to be the best model for TBS detection task. The training process was conducted for 2000 iterations, achieving a mean average precision (mAP) of 87.9% in the final iteration. This model successfully detects ripe FFB with high accuracy. The results of this research indicate that the use of computer vision and artificial intelligence technology can help optimize the oil palm harvesting process. With this automated system, the palm oil industry can address labor shortages and improve production efficiency. This study provides an important contribution to the development of the oil palm industry with potential applications in broader fields.

ARTICLE INFO

Article history:

Received May 2, 2025
Revised Jun 5, 2025
Accepted Jul 28, 2025

Keywords:

Artificial Intelligent
Computer Vision
Fresh Fruit Bunch
Ripeness
YOLOv4

* **Corresponding Author**
E-mail address: nisa.arpyanti@gmail.com

1. INTRODUCTION

Ripe oil palm fruit is key to producing quality palm oil. The ripeness of oil palm fruit bunches (FFB) can generally be identified by the number of fruits detached from the bunch [1]. On oil palm plantations, FFB can only be harvested when the trees are mature, at three years old. Field workers harvest the FFB 10-14 days after harvest. Ripe FFB are typically identified by their bright red and yellow color, as opposed to the brown and black color of unripe FFB. The harvested FFB are collected and transported to the palm oil mill for oil extraction [2]. In practice, ideally, FFB should be delivered to the mill within 24 hours of harvest to maintain optimal fruit quality. However, this is not always guaranteed due to factors such as adverse weather during harvest, unforeseen logistical issues, and other factors. Furthermore, labor shortages have a significant impact on the economic growth of the traditional, labor-intensive palm oil industry [3].

In 2020, the palm oil industry contributed 15% to the country's foreign exchange earnings. Through palm oil exports, the industry contributed approximately USD 15 billion to the country's foreign exchange earnings that year. Meanwhile, in terms of imports, the palm oil industry also contributed approximately IDR 38 trillion in the same year [4]. Exported palm oil production continues to increase annually. However, the main obstacle faced is that the majority of palm fruit sorting processes are still carried out manually, due to the limited availability of reliable systems for classifying palm fruit [5]. Oil palm plantations have reported labor shortages of around 20-30%, resulting in potential yield losses of around 15-45% due to post-harvest issues [6]. To address these challenges, the horticultural industry needs to adopt the latest technology in the FFB harvesting process. Utilizing the latest technology, such as harvesting machines and automation systems, can help increase efficiency and productivity in the harvesting process. This way, the industry can mitigate the negative impacts of labor shortages and optimize harvest yields.

Several studies have been conducted on FFB maturity classification using non-destructive methods. A study used a computer vision algorithm implemented in a low-cost, portable processor to train a Convolutional Neural Network (CNN). Palm fruit ripeness detection was obtained from images of 100 palm fruits at varying levels of ripeness, which were then analyzed in Hue, Saturation, and Value (HSV) space [7]. Another study used multispectral imaging with a monochrome camera to determine the correlation between ripeness and fruit firmness [8]. Palm fruit was classified through image binarization, morphological processing, and color feature extraction. The average color intensity was calculated based on a previously developed RGB model [9].

This study used a deep learning model called YOLO (You Only Look Once) to detect ripe FFB [10]. The advantage of the YOLO model is its ability to detect ripe fruit bunches (FFB) with high accuracy and in real time thanks to its fast processing techniques. Previously, the YOLO model has been implemented to detect agricultural fruits such as apples [11], tomatoes [12], and pears [13]. In the palm oil sector, Junos et al. [14] developed an automated detection system incorporating the YOLO model for FFB detection.

In this project, the goal was to develop an automated system that can detect ripe and unharvested FFB in real-time using a combination of computer vision and artificial intelligence (AI). This system was designed for use in field applications as part of a robotic harvesting mechanism. In this system, images of oil palm trees were captured using an RGB camera (Intel Realsense D435) and sent to a tablet computer (Nvidia Jetson NX) equipped with an inference model based on YOLOv4. This model was trained to identify ripe FFB, record their coordinates, and send the location information to the robot's picking mechanism in the robot operating system (ROS). YOLOv4 was chosen because it is well integrated with ROS, which is a core component of the developed robotic harvesting system. Although not the latest version, YOLOv4 is still capable of detecting objects with high accuracy and speed [15].

2. THEORITICAL REVIEW

Computer vision plays a crucial role in image-based applications, evolving from a sensing modality to intelligent computing systems. Computer vision develops methods that mimic human visual ability (eyes) to infer real-world characteristics in three dimensions using light reflections received by sensors from objects. Based on the level of abstraction of the output information, three task categories can be distinguished: low-level, mid-level, and high-level computer vision.

Low-level computer vision focuses on image or video processing in various applications such as image matching, automatic identification, optical flow computing, and motion analysis. Mid-level computer vision can infer object geometry and capture and track visual motion. Meanwhile, high-level computer vision provides more complex information such as object recognition and relationship interpretation, enabling better recognition and understanding of objects within images [16].

The use of computer vision in image- and camera-based imaging has significant potential for handling identification, classification, and detection processes in various fields. In agriculture, this method has been widely applied in ripeness classification systems, disease detection, advanced harvesting, and crop care. Several studies related to computer vision applications in agriculture include leaf disease detection in apple orchards for intelligent spraying [17], automated tractor driving using computer vision [18], real-time peach detection, specifically in oil palm plantations, research related to fresh fruit bunch (FFB) analysis, oil palm ripeness classification, and crude palm oil (CPO) prediction [16, 19]. Research focused on the efficiency of optical imaging for oil palm FFB ripeness classification is a key focus in efforts to increase CPO productivity in palm oil mills.

YOLO (you only look once) is a development of an artificial neural network (ANN) computing system used for real-time object detection. This method localizes images at various areas and scales and makes detections based on the highest image value. By dividing the image into several regions, the ANN can predict bounding boxes and possible objects within each region. This approach is more efficient than classifier-based systems. The main advantage of YOLO is its speed in detecting objects without requiring extensive initial data, as is required by the region-convolutional neural network (R-CNN) method [17].

The YOLO algorithm has been widely applied in various studies due to its superior ability to detect objects with high accuracy and in real time. Some relevant studies include tomato detection using a modified version of YOLOv3, real-time pear detection and counting, and the development of a YOLO-based object detection model optimized for oil palm harvesting systems. Research related to oil palm plantations has promising prospects, especially in Indonesia, one of the world's largest palm oil producers. Real-time operations are invaluable in improving the efficiency of automated harvesting mechanisms in the field [11-15].

3. RESEARCH METHODS

To develop an artificial intelligence (AI)-based vision system for detecting FFB in oil palm trees, the algorithm must be trained using visual data or samples of ripe FFB from the trees. This section provides a detailed explanation of the work done for data acquisition, preparation, and training.

3.1. Data Acquisition

Visual data of FFB from oil palm trees was captured using an Intel RealSense D435 camera. The camera was mounted on a height-adjustable platform to capture images at the same height as the FFB on the tree. The captured images had a resolution of 1920 × 1080 pixels and were stored on a laptop computer connected via a USB-C cable. During data collection, an expert in the field assisted in identifying ripe and unripe FFB on the trees.

3.2. Data Preparation and Training

Video captured of oil palm trees was first extracted into images. These images can be categorized as positive and negative. Positive images are images of trees containing the desired object to be detected, namely ripe FFB. Conversely, negative images are images of trees without ripe FFB. There were a total of 240 positive images and 250 negative images. Ten images were then systematically selected for each tree from the large number of extracted images. The reason for selecting only 10 images was to avoid redundancy in the training images. These images show objects in various angles, backgrounds, positions, and lighting. This variation is crucial in the training process of deep learning algorithms, as they require this information to learn and detect patterns in images.Next, the positive images underwent a manual labeling process using labelImg software. The positive and negative images were further separated into two datasets: a training set and a validation set. The positive images were divided into 210 and 30 images for the training and validation sets, respectively. Meanwhile, the negative images were divided into 220 and 30 images for the training and validation sets, respectively. The images extracted from each tree were unique, meaning images from the same tree would not be included in the same training and validation sets.

Table 1. Image separation for training the YOLOv4 model.

Dataset	Trainingset	Validationset	Number of images
Positif	210	30	240
Negatif	220	30	250
Total	430	60	490

Several variations of the YOLOv4 model were evaluated to compare their performance on the TBS detection task. The evaluated model variations included YOLOv4-512, YOLOv4-608, YOLOv4-CSP-512, YOLOv4-CSP-608, YOLOv4-tiny-512, and YOLOv4-tiny-608. The evaluation was conducted by comparing the learning curves of each model, including the average loss and mean average precision. The evaluation results show that the YOLOv4 and YOLOv4-CSP models with an input network size of 512 × 512 have higher mAP than those with an input network size of 608 × 608. Furthermore, YOLOv4-CSP experiences a slower average loss rate compared to other YOLOv4 models. However, the YOLOv4-tiny model has a shorter training time but lower mAP compared to the YOLOv4 and YOLOv4-CSP models. Therefore, the YOLOv4 model with an input network size of 512 × 512 was selected as the best model for the TBS detection task in this study.

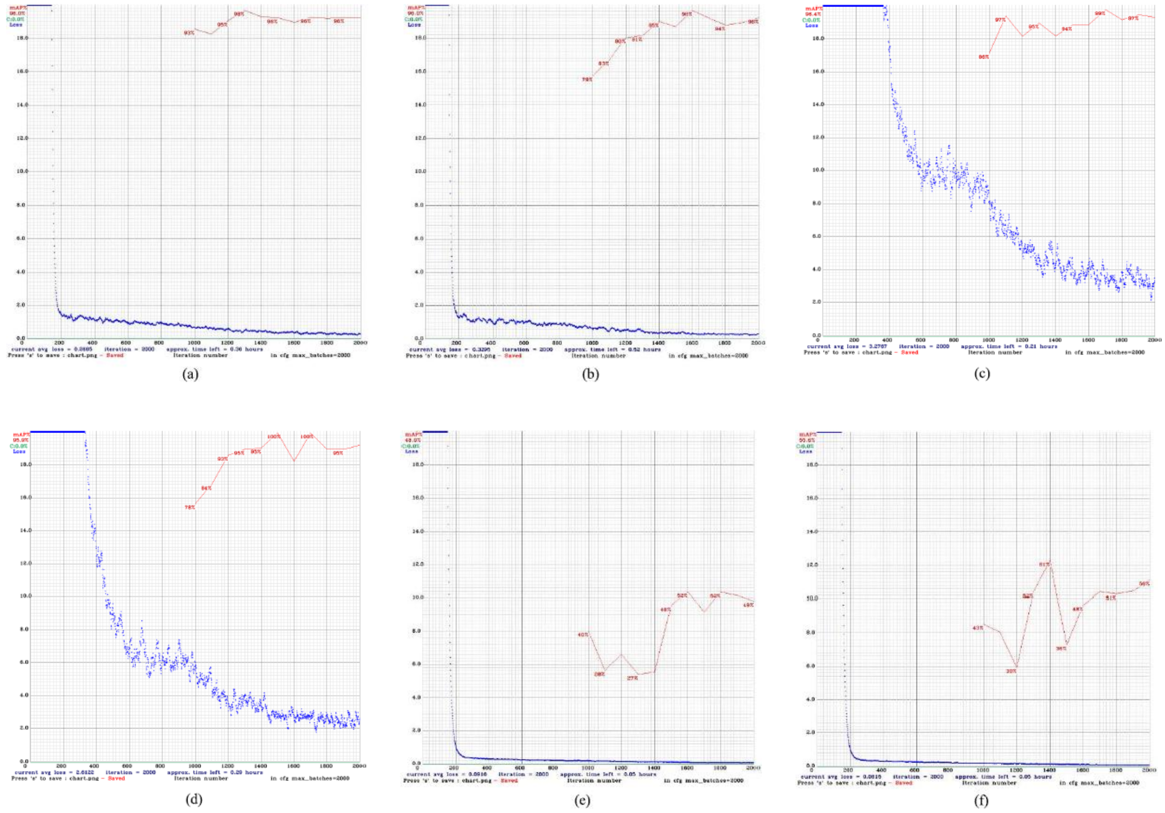

Figure 1. (a) YOLOv4-512, (b) YOLOv4-608, (c) YOLOv4-CSP-512, and (d) YOLOv4-CSP-608, (e) YOLOv4-tiny-512, and (f) YOLOv4-tiny-608.

4. RESULTS AND DISCUSSIONS

4.1. YOLOv4 Model Training Performance

The training process was conducted on a computer with an Intel Core i7-10700K CPU, NVIDIA GeForce RTX 3080 GPU, and 32 GB of RAM. The model was trained for 2,000 iterations with a batch size of 64 and a learning rate of 0.001. During training, evaluations were conducted every 100 iterations to monitor model performance. The evaluation results showed that the YOLOv4 model achieved a mAP of 87.9% at the 2,000th iteration. Furthermore, the model also had a recall of 82% and an F1-score of 88%. These results indicate that the YOLOv4 model is capable of detecting mature FFB with high accuracy.

Table 2. YOLOv4 model analysis for 1000 iterations.

Iterations	1000	2000	3000
Precision	86%	100%	100%
Recall	80%	97%	100%
FI-score	83%	98%	100%
Average IOU	64.24%	77.93%	79.75%
mAP	87.88%	99.89%	100%

4.2. Comparison with Different YOLOv4 Models

Comparison of the performance of several YOLOv4 model variations with different input network sizes. The evaluation results show that the YOLOv4 model with an input network size of 512 × 512 has a higher mAP than the one with an input network size of 608 × 608. In addition, the YOLOv4-CSP model experiences a slower average loss decrease compared to the other YOLOv4 models. However, the YOLOv4-tiny model has a shorter training time but lower mAP compared to the YOLOv4 and YOLOv4-CSP models. Therefore, the YOLOv4 model with an input network size of 512 × 512 was selected as the best model for the TBS detection task in this study.

Table 3. Weight size in different models.

Model	Weight size (MB)
YOLOv4-512	256.0
YOLOv4-608	256.0
YOLOv4-CSP-512	210.2
YOLOv4-CSP-608	210.2
YOLOv4-tiny-512	23.5
YOLOv4-tiny-608	23.5

Table 4. Analysis of different YOLOv4 models.

Model name	Precision	Recall	F1-score	Average IoU	mAP
YOLOv4-512	97%	97%	97%	75.85%	96.00%
YOLOv4-608	97%	97%	97%	77.13%	96.22%
YOLOv4-CSP-512	90%	87%	88%	67.48%	95.89%
YOLOv4-CSP-608	84%	90%	87%	63.24%	96.43%
YOLOv4-tiny-512	57%	57%	57%	38.74%	55.60%
YOLOv4-tiny-608	48%	77%	59%	33.76%	48.89%

Table 5. Evaluation of the YOLOv4 model in on-site testing dataset.

Evaluation	Percentage
Precision	95%
Recall	82%
F1-score	88%
Average IoU	70.19%
mAP	87.9%

4.3. Real-Time on-Site Testing of the YOLOv4 Mode

The YOLOv4 model was tested in real time at various locations to evaluate its performance under different environmental conditions. Testing was conducted under seven different environmental conditions: bright sunlight, shaded lighting, near view, far view, leaf obstruction, motion blur, and unripe FFB.

The evaluation results showed that the YOLOv4 model was able to detect ripe FFB with high accuracy under all tested environmental conditions. In bright sunlight, the YOLOv4 model achieved a detection accuracy of 96.7%, while in far-field conditions, it achieved a detection accuracy of 93.3%. Although the YOLOv4 model's detection accuracy decreased slightly under conditions of leaf obstruction and motion blur, it still achieved quite high detection accuracies of 86.7% and 83.3%, respectively.

These evaluation results demonstrate that the YOLOv4 model can be used effectively in a real-time ripe FFB detection system in the field, even under varying environmental conditions. This demonstrates the YOLOv4 model's broad application potential in the oil palm plantation industry.

5. CONCLUSION

The conclusion of this review article is that the use of the YOLOv4 model in real-time detection of ripe FFB in the field shows effective results, even under varying environmental conditions. This article also illustrates the potential for broad applications of the YOLOv4 model in the oil palm plantation industry. Although this article has several shortcomings, such as the lack of detailed information about the dataset used in training the YOLOv4 model, the absence of comparisons with other ripe FFB detection methods, and the absence of discussion on the large-scale application of the YOLOv4 model in the oil palm plantation industry, it still provides useful information about the use of the YOLOv4 model in real-time detection of ripe FFB in the field and its potential applications in the oil palm plantation industry.

REFERENCES

[1] Chew, C. L., Tan, B. A., Low, J. Y. S., Mohd Hakimi, N. I. N., Kua, S. F., & Lim, C. M. (2021). Exogenous ethylene application on postharvest oil palm fruit bunches improves crude palm oil quality. *Food Science & Nutrition*, **9**(10), 5335–5343.

[2] Chew, C. L., Ng, C. Y., Hong, W. O., Wu, T. Y., Lee, Y. Y., Low, L. E., Kong, P. S., & Chan, E. S. (2021). Improving sustainability of palm oil production by increasing oil extraction rate: A review. *Food and Bioprocess Technology*, **14**(4), 573–586.

[3] Ramli, A., Azman, I., & Ayatollah, K. (2011). Labour requirements in the Malaysian oil palm industry 2010. *Malaysian Palm Oil Journal*, **11**(2), 1–12.

[4] Limaho, H., Sugiarto, Pramono, R., & Christiawan, R. (2022). The need for global green marketing for the palm oil industry in Indonesia. *Sustainability*, **14**(14), 8621.

[5] Herman, H., Susanto, A., Cenggoro, T. W., Suharjito, S., & Pardamean, B. (2020). Oil palm fruit image ripeness classification with computer vision using deep learning and visual attention. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, **12**(2), 21–27.

[6] Ibragimov, A., Sidique, S. F., & Tey, Y. S. (2019). Productivity for sustainable growth in Malaysian oil palm production: A system dynamics modeling approach. *Journal of Cleaner Production*, **213**, 1051–1062.

[7] Wong, Z. Y., Chew, W. J., & Phang, S. K. (2020). Computer vision algorithm development for classification of palm fruit ripeness. *AIP Conference Proceedings*, **2233**(1), 030012.

[8] Ningsih, S. A., Shiddiq, M., Arief, D. S., & Husein, I. R. (2020). Pengunaan pencitraan multispektral pada panjang gelombang 520 nm dan 800 nm untuk mengevaluasi tingkat kematangan TBS kelapa sawit. *Komunikasi Fisika Indonesia*, **17**(3), 144.

[9] Sulaiman, N. F., Pauline, O., Kiow, L. W., Huong, L. K., & Fong, G. S. (2020). Automatic grading system for oil palm fruit ripeness. *Communications in Computational and Applied Mathematics*, **2**(1).

[10] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.

[11] Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., & Liang, Z. (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and Electronics in Agriculture*, **157**, 417–426.

[12] Lawal, M. O. (2021). Tomato detection based on modified YOLOv3 framework. *Sci. Rep.*, **11**.

[13] Parico, A. I. B. & Ahamed, T. (2021). Real time pear fruit detection and counting using YOLOv4 models and deep SORT. *Sensors*, **21**(14), 4803.

[14] Junos, M. H., Mohd Khairuddin, A. S., Thannirmalai, S., & Dahari, M. (2021). An optimized YOLO-based object detection model for crop harvesting system. *IET Image Processing*, **15**(9).

[15] Lai, J. W., Ramli, H. R., Ismail, L. I., & Hasan, W. Z. W. (2022). Real-time detection of ripe oil palm fresh fruit bunch based on YOLOv4. *IEEE Access*, **10**, 95763–95770.

[16] Khan, S., Rahmani, H., Shah, S. A. A., Bennamoun, M., Medioni, G., & Dickinson, S. (2018). A guide to convolutional neural networks for computer vision. *Synth. Lect. Comput. Vis.*, **8**(1).

[17] Storey, G., Meng, Q., & Li, B. (2022). Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture. *Sustainability*, **14**(3), 1458.

[18] Vrochidou, E., Oustadakis, D., Kefalas, A., & Papakostas, G. A. (2022). Computer vision in self-steering tractors. *Machines*, **10**(2), 129.

[19] Jupiyandi. S., Pratama, F. R., Dharmawan, Y., Dharmawan, M. R., & Cholissodin, I., (2019). Pengembangan deteksi citra mobil untuk mengetahui jumlah tempat parkir menggunakan cuda dan modified YOLO. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, **6**(4). 413–419.