International Journal of Technology 13(6) 1326-1335 (2022)
Received April 2022 / Revised August 2022 / Accepted September 2022

International Journal of Technology

http://ijtech.eng.ui.ac.id

Object Detection Algorithms for Ripeness Classification of Oil Palm Fresh Fruit Bunch

Mohamed Yasser Mohamed Ahmed Mansour[1], Katrina D. Dambul[1*], Kan Yeep Choo[1]

[1]*Faculty of Engineering, Multimedia University, Persiaran Multimedia, 63100, Cyberjaya, Selangor, Malaysia*

Abstract. Ripe oil palm fresh fruit bunch allows extraction of high-quality crude palm oil and kernel palm oil. As the fruit ripens, its surface color changes from black (unripe) or dark purple (unripe) to dark red (ripe). Thus, the surface color of the oil palm fresh fruit bunches may generally be used to indicate the maturity stage. Harvesting is commonly done by relying on human graders to harvest the bunches according to color and number of loose fruits on the ground. Non-destructive methods such as image processing and computer vision, including object detection algorithms have been proposed for the ripeness classification process. In this paper, several object detection algorithms were investigated to classify the ripeness of oil palm fresh fruit bunch. MobileNetV2 SSD, EfficientDet (Lite0, Lite1 and Lite2) and YOLOv5 (YOLOv5n, YOLOv5s and YOLOv5m) were simulated and compared in terms of their mean average precision, recall, precision and training time. The models were trained on a dataset with four main ripeness classes: ripe, unripe, half-ripe, and over-ripe. In conclusion, object detection algorithms can be used to classify different ripeness levels of oil palm fresh fruit bunch, and among the different models, YOLOv5m showed promising results with a mean average precision of 0.842 (0.5:0.95).

Keywords: Computer vision; Object detection; Oil palm fresh fruit bunch; Ripeness classification; YOLO

1. Introduction

Malaysia is one of the leading countries in the world, producing oil palm (Gan & Li, 2014). For the period of January to September 2022, Malaysia has produced more than 13 million tonnes of crude palm oil and exported over 17 million tonnes of oil palm products (MPOB), 2022). The government of Malaysia is encouraging the utilization of Industry Revolution 4.0 (IR 4.0) technologies to realize high crops yields, reduction of costs, and replacement of low-skilled and labor-intensive work with automated machinery, which can yield more sustainable development in agriculture industries of Malaysia (Ibrahim, 2021; Ghulam, 2021) despite the human workforce disruption caused by the COVID-19 pandemic (Ng, 2021). There is also a huge potential for adopting IR 4.0 in the oil palm industry in Malaysia, as pointed out by the research studies reported in (Parvand & Rasiah, 2022; Lazim et al., 2020) and demonstrated in the agriculture sectors of other countries (Heryani et al., 2022; Belousova & Danilina, 2021; Onibonoje et al., 2019).

*Corresponding author's email: katrina@mmu.edu.my, Tel.: +603-8312 5594; Fax: +603-8312 5300
doi: 10.14716/ijtech.v13i6.5932

Oil palm fresh fruit bunches (FFBs) in plantations are currently harvested by human graders based on the surface color of the fruit and the number of loose fruits on the ground as the indication of the ripeness level. This is done based on the standards and requirements specified by the Malaysian Palm Oil Board (MPOB) (Malaysia Department of Standards, 2007). However, relying on human graders may lead to misclassified bunches due to factors such as the height of the tree (on higher trees, the oil palm FFBs may not be clearly visible to the human graders), the position of the bunches (some FFBs may be hidden due to the branches of the tree), lighting conditions, unclear vision and miscount of the loose fruits on the ground. These misclassifications may lead to the harvesting of unripe FFBs, which will cause profit losses due to the production of lower-quality oil palm (Junkwon et al. 2009; Sunilkumar & Babu, 2013).

Oil palm FFBs ripeness levels defined by MPOB include under-ripe, partially ripe, ripe, and over-ripe (MPOB, 2016; MJM (Palm Oil Mill) Sdn. Bhd, 2014). Only ripe FFBs, which produce high-quality palm oil with a higher yield, have the highest commodity values, followed by the over-ripe FFBs. On the other hand, under-ripe and partially ripe FFBs will be rejected by oil palm mill owners. Some ripeness levels are hard to be differentiable by human vision due to the very similar color appearance and the number of fruitlets that are ripe on an FFB. The situation is worsened by the shortage of plantation workers (Ng, 2021) which causes the oil palm estate owners to employ inexperienced workers to fill the vacancies. Classification of the ripeness level of FFBs by inexperienced workers imposes a potential risk of wrong classification that can lead to the whole lot of the FFBs being returned to the plantation owner by the oil palm millers, and the plantation owner can be given a fine by the authority. Hence, the productivity, cost efficiency, revenues, and reputation of the plantation owner will be affected. Thus, an experience-independent, visual-based FFB classification tool is proposed in this work to overcome the weaknesses faced by plantation workers.

Computer vision is a part of artificial intelligence technology which allows computers to extract features from images and other visual inputs and perform image classification, object detection and others (Szeliski, 2021). Object detection is a widely used application of computer vision. Object detection algorithms can predict the class and the location of the object in the image as well as image classification, which can identify different classes of images (Wu et al., 2020). Before the development of deep learning object detection algorithms (Lecun et al., 2015), traditional object detection algorithms such as Hough transform (Hough, 1960), sliding windows, and background extraction was divided into proposal generation, feature vector extraction, and region classification which were slow and computationally inefficient (Patkar et al., 2016;Tang et al., 2017;Wu et al., 2020).

Table 1 shows different ripeness classification methods for oil palm FFBs. Convolution neural network (CNN) (Ibrahim et al., 2018;Saleh & Liansitim, 2020; Arulnathan et al., 2022), achieved accuracies of 92-97%. However, the models consisted of a few CNN layers (one or two layers) and would not be able to capture the high-level features. Low-level features captured would only partially represent the FFBs. This is not robust or reliable for real-time applications. Pre-trained models such as AlexNet and DenseNet have also been developed, as shown in Table 1. (Herman et al., 2020; Herman et al., 2021) have tested the use of an attention mechanism module to improve the performance of DenseNet. Modifications using squeeze and excitation (SE) block (Hu et al., 2020) and ResATT (Herman et al., 2020) were also developed. AlexNet was able to achieve 100% (Ibrahim et al., 2018). However, with a bigger dataset, the performance dropped to 60-85% (Herman et al., 2020; Herman et al., 2021; Wong et al., 2020).

Table 1 Ripeness classification methods for oil palm fresh fruit bunch

Method (Reference)	Ripeness classes	Number of images	Accuracy (%)
CNN (Ibrahim et al., 2018)	4	120	92
CNN (Saleh & Liansitim, 2020)	2	628	97
CNN (Arulnathan et al., 2022)	3	126	96
AlexNet (Ibrahim et al., 2018)	4	120	100
AlexNet (Herman et al., 2020)	7	400	60
AlexNet (Wong et al., 2020)	2	200	85
AlexNet (Herman et al., 2021)	7	400	77
DenseNet (Herman et al., 2021)	7	400	89
DenseNet Sigmoid (Herman et al., 2020)	7	400	69
DenseNet and SE layer (Herman et al., 2020)	7	400	64
ResAtt DenseNet (Herman et al., 2020)	7	400	69
Faster R-CNN (Prasetyo et al., 2020)	-	100	86
YOLOv3 (Selvam et al., 2021)	3	4500	mAP = 0.91
YOLOv3 (Khamis, 2022)	3	229	mAP = 0.84

Deep learning-based object detection is divided into two types (Zhao et al., 2019). The first type is region proposal-based models (two-stage detectors) such as R-CNN (Girshick et al., 2014), fast R-CNN (Girshick, 2015), and faster R-CNN (Girshick et al., 2014), and the second type is bounding box regression-based models (single stage detectors) such as You Only Look Once (YOLO) (Redmon et al., 2016), single shot multi-box detector (Liu et al., 2016), YOLOv3 (Redmon & Farhadi, 2018) and EfficientDet (Tan et al., 2020).

Faster R-CNN (Prasetyo et al., 2020) developed to detect and count the bunches achieved 86% accuracy. However, faster R-CNN is slow for real-time applications and requires more computational power. YOLOv3 has also been developed for real-time models (Khamis, 2022; Selvam et al., 2021).

With the development of regression-based models (one-stage detectors), the use of region proposal-based models and convolution neural network (CNN) network for feature extraction and classification, and localization was eliminated (Tang et al., 2017; Zhao et al., 2019). Regression-based model is divided into the base model and the auxiliary model. The base model is an image classification model without the classifier layer and is responsible for the extraction of the features from the images, and the auxiliary model is responsible for the detection part. Single stage detector (SSD) such as YOLO is a single CNN where the front end works on the extraction of the features, and the last part is two fully connected layers for classification and regression (Redmon et al., 2016).

SSD combines the idea of YOLO by treating object detection as a regression problem and adds the concept of anchor boxes, such as in faster R-CNN (Liu et al., 2016). SSD utilizes multiscale feature maps, which allows the models to detect objects at multiple scales. However, YOLO uses only one feature map for detection (Liu et al., 2016). SSD is designed for speed with accuracy close to region-based object detectors, which is suitable for real-time applications. Most of the literature focused on image classification with less emphasis on object detection models for ripeness classification and localization.

In this paper, several object detection algorithms will be investigated to develop an object detection algorithm capable of the ripeness classification of oil palm FFBs. The algorithm can also be potentially implemented on mobile devices, which can help human graders to accurately harvest the ripe bunches only and to reduce wastage due to the harvesting of unripe bunches.

2. Object Detection Algorithms

The object detection algorithms investigated in this paper are MobileNetV2 SSD, EfficientDet-lite and YOLOv5. These algorithms were chosen as they are designed specifically for implementation on mobile devices such as mobile phones and have a high memory efficiency. All the algorithms will be tested using a dataset that contains oil palm FFBs with four different ripeness levels. The performance and effectiveness of each algorithm will be measured in terms of its accuracy in the ripeness classification of oil palm FFBs.

2.1. MobileNetV2 SSD

MobileNetV2 is a lightweight CNN network that is designed for implementation on mobile devices. MobilenetV2 as a backbone is combined with a SSD detector to develop MobileNetV2 SSD, which is to replace the original VGG16-SSD, which utilizes VGG16 as its backbone. MobileNetV2 consists of CNN layers and inverse residual modules. Inverse residual modules include depth-wise separable convolutional layers, batch normalization layer, and ReLU6 activation function, where ReLU stands for rectified linear unit. Together, these layers form a MBconv block which offers more efficient memory usage, especially for mobile applications (Chiu et al., 2020).

2.2. EfficientDet

EfficientDet is a single-shot object detector developed by Google. EfficientDet relies on EfficientNet (Tan & Le, 2019) as its backbone, which is a network in image classification. It employs new architecture, which allows it to extract complex features. For the neck, EfficientDet uses a bi-directional pyramid network (Bi-FPN) which is an improved path aggregation network (PANet), adding bottom-up and top-down paths, which help to develop a better feature fusion. EfficientDet is similar to EfficientNet, which utilizes the concept of model scaling, which allows it to change the width, resolution, and depth of the backbone to improve the performance of the algorithm. Feature levels extracted from the different layers are passed from the backbone to the neck and then sent to the head for prediction after fusion (Tan et al., 2020).

2.3. YOLOv5

YOLOv5 is developed by Ultralytics and is the latest improvement of the YOLO family. YOLOv3 is an incremental improvement to YOLOv2. Its improved architecture provides high real-time accuracy with a fast inference time. YOLOv5 network size is smaller than other object detection networks which makes it perfect for real-time applications and deployment on embedded devices (Yan et al., 2021). The original YOLOv5 architecture is divided into backbone, neck, and detect networks. The function of the backbone, which is inspired by cross stage partial network (CSPNet) (Wang et al., 2020) is to extract important features from the input images. The next part of the network is the neck which is based on PANet (Liu et al., 2018). PANet allows information to flow easily in bottom-up paths. It allows better use of the spatial information contained in the low-level features (Liu et al., 2018). The last part of the network is the head which is responsible for the detection part and is divided into parts to allow the model to detect objects on multiple scales (Xu et al., 2021).

3. Methods

3.1. Dataset Collection

In this paper, images of oil palm FFBs taken on the ground after the harvesting process was collected from an oil palm estate in Malaysia. The dataset consists of images of oil palm

FFBs taken on the ground due to the restrictions imposed on visiting oil palm plantations physically during the COVID-19 pandemic. Future datasets will include a combination of different scenarios (e.g., different lighting conditions, tree height, and bunch position) for the oil palm FFBs on the trees. The dataset consists of oil palm FFBs in four ripeness stages which are ripe, unripe, over-ripe, and half-ripe. Figure 1 illustrates the changes that happen during the oil palm FFBs ripeness process, where the fruit changes from unripe (in Figure 1(a)) to over-ripe (in Figure 1(d)). The images of the FFBs were classified based on the Malaysian Palm Oil Board (MPOB) standards (Malaysian Palm Oil Board (MPOB), 2016; Malaysia Department of Standards, 2007). The total number of images collected was 328 images. The dataset used to form the training and testing datasets were reduced to 304 images in order to have a balanced dataset of equal images per class. The dataset was divided into three sets which were training dataset, validation dataset, and testing dataset with a split ratio of 70%, 20%, and 10%. Image augmentation techniques will be implemented to induce variations to the dataset, such as horizontal and vertical flip, rotation, crop, zoom, and shear.

In this paper, three different object detection algorithms were tested and compared. The models investigated were MobileNetV2 SSD, EfficientDet, and YOLOv5. The images were first pre-processed and annotated for each algorithm. Image annotation was done by drawing a bounding box around all the objects of interest in each image, and then the images were resized to fit each model. Training and testing were done using Tesla K80 GPU on Google Colab. MobileNetV2 SSD, EfficientDet, and YOLOv5 algorithms were trained on images with a size of 640 x 640 pixels for 300 epochs.

Figure 1 Oil palm FFBs (a) unripe, (b) half-ripe, (c) ripe and (d) over-ripe bunch

4. Results and Discussion

YOLOv5 models were trained for 300 epochs with a batch size of 16 and an image size of 640 pixels. The hyperparameters and weights used for training were for the pre-trained model on COCO (Common Object in Context) dataset by Microsoft, which includes 80 classes of common objects and is used for object detection and benchmarking of algorithms using Pytorch (Lin et al., 2014). EfficientDet-lite0, EfficientDet-lite1, EfficientDet-lite2 and MobileNetV2 SSD were trained using TensorFlow backend. YOLOv5 and EfficientDet used model scaling (Tan & Le, 2019), which allows changes to the depth, width, and resolution of the model to produce other model sizes from the base model.

Table 1 shows the results of different object detection algorithms' performance and a comparison of their performance in terms of mean average precision (mAP), COCO mAP, parameters, and training time. The mAP is a COCO dataset benchmarking metric. It represents the mean average precision of intersection over the union between the prediction and ground truth of 0.5 to 0.95. For benchmarking purposes, the mean average precision results using COCO dataset is also shown, and the trend of the results is similar to those obtained in this paper. YOLOv5 is designed to provide a high-speed inference time for real-time applications. YOLOv5n is the smallest model of YOLOv5 in terms of parameters and size, which also gives the fastest training time. YOLOv5m shows a longer training time (235 minutes) compared with YOLOv5s (90 minutes). This is because the YOLOv5m learned better on the training dataset due to the depth and width parameters affecting the size of the model. But this creates a longer training time and a larger model size. Deeper models will have better feature extraction. However, the parameters, size, and training time will increase.

EfficientDet-lite0, EfficientDet-lite1 and EfficientDet-lite2 are derived from EfficientDet architecture for mobile applications where EfficientDet-lite0 is the base model and EfficientDet-lite1 and EfficientDet-lite2 are scaled versions based on compound scaling (Tan & Le, 2019). Performance wise, EfficientDet-lite models come second after YOLOv5 and then finally MobileNetV2 SSD in terms of mean average precision (mAP), training speed, and model size. EfficientDet-lite offers different model sizes similar to YOLOv5 ranging from EfficientDet-lite0 to EficientDet-lite2, with EficientDet-lite2 showing a higher mAP but longer training and inference time as well as more parameters and larger model size. YOLOv5 is designed for ease of implementation, and its different structure layers and parameters can be easily modified. From Table 2, it can be seen that the YOLOv5 models outperformed EfficientDet models in terms of mAP using both the COCO dataset and the dataset in this work.

Table 2 Object detection models comparative analysis

Model	mAP_{val}	Training time (min)	COCO mAP	Parameters (million)
MobileNetv2 SSD	0.478	-	0.222	-
EfficientDet-lite0	0.743	65	0.264	3.2
EfficientDet-lite1	0.803	100	0.315	4.2
EfficientDet-lite2	0.812	140	0.351	5.3
Yolov5n	0.781	43	0.457	1.9
Yolov5s	0.832	90	0.568	7.2
Yolov5m	0.842	235	0.641	21.2

Although EfficientDet relies on EfficientNet, which is a strong image classifier, as a backbone, and utilizes Bi-FPN as neck, which is an improvement over PANet, YOLOv5 is still showing a higher mAP. YOLOv5 architecture, which is based on CSPNet, is working better on extracting features from the input images based on the results and the feature fusion between the head and neck in order to detect objects on different scales. YOLOv5X and EfficicentDet-D7 are both representing the strongest variation of the two models. Both models achieved 55 mAP, with EfficientDet having 77 M parameters and YOLOv5X having 86.7 M parameters. However, both models are not suitable for mobile application implementation, which requires the most efficient model with the highest accuracy and inference.

Another advantage of YOLOv5 is that the model can be continuously improved, and other techniques that can help the performance, such as ensemble, pruning, and test-time augmentation, can be added to the model. YOLOv5 has a faster detection speed (6-8 ms) and is more computationally efficient. It can also be implemented on mobile devices. YOLOv5m was able to achieve mAP of 0.84 (0.5: 0.95) for four different ripeness classes as compared to three ripeness class using YOLOv3 (Khamis, 2022). Further optimization of YOLOv5m, such as hyperparameter tuning, will help to improve its performance.

5. Conclusions

In this paper, the performance of three object detection algorithms which are MobileNetV2 SSD, EfficientDet, and YOLOv5, were simulated using different architectures to classify different ripeness levels of the oil palm FFBs. YOLOv5 is designed mainly for real-time application with the feasibility of improvement, modification, and ease of implementation. EfficientDet is a strong object detector but has not shown a similar performance to YOLOv5. MobileNetV2 SSD is based on MobileNet, which is designed for mobile applications but is not a strong backbone for object detection application compared to other models. In conclusion, YOLOv5m with a mean average precision of 0.842 (0.5:0.95) is proposed to be the object detection model for the application of ripeness classification of the oil palm FFBs with the possibilities for future improvements on the model. The use of object detection models to classify the ripeness of oil palm FFB supports the digitalization of the agriculture industry and its move towards the implementation of artificial intelligence (AI) in all applications. With the right object detection model, autonomous harvesters can outperform human workers with less cost and time. Future work will include an improvement to the dataset, such as adding images of oil palm FFBs on the trees and improving the algorithm's real-time testing accuracy. The algorithm will be further optimized using hyperparameter tuning to suit the dataset for the ripeness classification of oil palm FFBs. Advanced model ensemble techniques will be investigated to develop a more accurate algorithm by combining YOLOv5 and EfficientDet. Finally, a mobile phone application will be developed using the models developed, and real-time tests will be performed.

Acknowledgements

The authors acknowledged that this work was funded and supported by the Ministry of Higher Education, Malaysia (MOHE-FRGS) with grant number FRGS/1/2019/TK04/MMU/03/7, and the year of the grant received was 2019.

References

Arulnathan, D.N., Koay, B.C.W., Lai, W.K., Ong, T.K., Lim, L.L., 2022. Background Subtraction for Accurate Palm Oil Fruitlet Ripeness Detection. *In:* IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), pp. 48–53

Belousova, M., Danilina, O., 2021. Game-Theoretic Model of the Species and Varietal Composition of Fruit Plantations. *International Journal of Technology*. 12(7), pp. 1498–1507

Chiu, Y-C., Tsai, C-Y., Ruan, M-D., Shen, G-Y., Lee, T.T., 2020. Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems. *In:* International Conference on System Science and Engineering (ICSSE 2020)

Gan, P.Y.,Li, Z.D., 2014. Econometric Study on Malaysias Palm Oil Position in the World Market to 2035. *Renewable and Sustainable Energy Reviews,* Volume 39, pp. 740–747

Ghulam, K.A.P.,2021. Timely to invest in mechanisation. New Straits Times. Available online at https://www.nst.com.my/business/2021/10/738418/timely-investmechanisation, Accessed on September 7, 2022

Girshick, R., 2015. Fast R-CNN. ArXiv1504.08083v2-1. Available online at https://arxiv.org/abs/1504.08083, Accessed on April 15, 2022

Girshick, R., Donahue, J., Darrell, T., Malik. J., 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *In:* IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587

Herman, H., Cenggoro, T.W., Susanto, A., Pardamean, B., 2021. Deep Learning for Oil Palm Fruit Ripeness Classification with Densenet. *In:* 2021 International Conference on Information Management and Technology, pp. 116-119

Herman, H., Susanto, A., Cenggoro, T.W., Suharjito, 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,* Volume 12(2), pp. 21–27

Heryani, H., Legowo, A.C., Yanti, N.R., Marimin, Raharja, S., Machfud, Djatna, T., Martini, S., Baidawi, T., Afrianto, I., 2022. Institutional Development in the Supply Chain System of Oil Palm Agroindustry in South Kalimantan. *International Journal of Technology*, Volume 13(3), pp. 643–654

Hough, P.V.C., 1960. Method And Means For Recognizing Complex Patterns. Google Patents. Available online at https://patents.google.com/patent/US3069654A/en, Accessed on April 25, 2022

Hu, Jie, Shen, L., Albanie, S., Sun, G., Wu, E., 2020. Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 42(8), pp. 2011–2023

Ibrahim, A., 2021. Invest in IR4.0 Tech to Reduce Dependence on Migrant Labour. New Straits Times. Available online at https://www.nst.com.my/opinion/columnists/2021/07/705128/invest-ir40-tech-reduce-dependence-migrant-labour. Accessed on September 7, 2022

Ibrahim, Z., Sabri, N., Isa, D., 2018. Palm Oil Fresh Fruit Bunch Ripeness Grading Recognition Using Convolutional Neural Network. *Journal of Telecommunication, Electronic and Computer Engineering,* Volume 10(3–2), pp. 109–113

Junkwon, P., Takigawa, T., Okamoto, H., Hasegawa, H., Koike, M., Sakai, K., Siruntawineti, J., Chaeychomsri, W., Vanavichit, A., Tittinuchanon, P., Bahalayodhin, B., 2009. Hyperspectral Imaging for Nondestructive Determination of Internal Qualities for Oil Palm (Elaeis Guineensis Jacq. Var. Tenera). *Agricultural Information Research,* Volume 18(3), pp. 130–141

Khamis, N., 2022. Comparison of Palm Oil Fresh Fruit Bunches (FFB) Ripeness Classification Technique Using Deep Learning Method. *In:* 2022 13th Asian Control Conference (ASCC), pp. 64–68

Lazim, R.M., Nawi, N.M., Masroon M.H., Abdullah, N., Iskandar, M.C.M., 2020. Adoption of IR4.0 in agriculturalsector in Malaysia: Potential and challenges.*Advances in Agricultural and Food Research Journal*, Volume 1(2), p. A0000140

Lecun, Y, Bengio, Y., Hinton, G., 2015. Deep Learning. *Nature,* Volume 521(7553), pp. 436–444

Lin, T-Y, Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., Dollár, P., 2014. Microsoft COCO: Common Objects in Context. *Computer Vision and Pattern Recognition*, Volume 8693, pp. 740–755

Liu, S, Qi, L., Qin, H., Shi, J., Jia, J., 2018. Path Aggregation Network for Instance Segmentation. *In:* Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,pp. 8759–8768

Liu, W, Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. SSD: Single Shot Multibox Detector. *Computer Vision and Pattern Recognition,* Volume 9905, pp. 21–37

Malaysia Department of Standards, 2007. *MS 814:2007 Palm Oil - Specification (Second Revision)*, Volume 814, Malaysia

Malaysian Palm Oil Board (MPOB), 2016. Buku Manual Penggredan Buah Kelapa Sawit (Oil

Palm Fruits Grading Manual), 3rd Edition. Bangi, Malaysia

Malaysia Palm Oil Board (MPOB), 2022. Production of Crude Palm Oil 2022. https://bepi.mpob.gov.my/index.php/en/production/production-2022/production-of-crude-oil-palm-2022, Accessed on October 26, 2022

Malaysia Palm Oil Board (MPOB), 2022. Monthly Export of Oil Palm Products 2022. https://bepi.mpob.gov.my/index.php/en/export/export-2022/monthly-export-of-oil-palm-products-2022, Accessed on October 26, 2022

MJM (Palm Oil Mill) Sdn. Bhd, 2014. FFB Grading Guideline. Available online at http://www.mjmpom.com/ffb-grading-guideline/, Accessed on September 7, 2022

Ng. J., 2021. Malaysia's Palm Oil Yield to Continue Declining on Labour Shortage. The Edge Markets, Malaysia. Available online at https://www.theedgemarkets.com /article/malaysias-palm-oil-yield-continue-declining-labour-shortage, Accessed on September 7, 2022

Onibonoje, M.O., Ojo, A.O., Ejidokun, T.O., 2019. A Mathematical Modeling Approach for Optimal Trade-Offs in a Wireless Sensor Network for a Granary Monitoring System. *International Journal of Technology*, Volume 10(2), pp. 332–338

Parvand, S., Rasiah, R., 2022. Adoption of Advanced Technologies in Palm Oil Milling Firms in Malaysia: The Role of Technology Attributes, and Environmental and Organizational Factors. *Sustainability*, Volume 14(1), pp. 260–286

Patkar, G., Anjaneyulu, G.S.G.N.,Mouli, P.V.S.S.R.C., 2016. Palm Fruit Harvester Algorithm for Elaeis Guineensis Oil Palm Fruit Grading Using UML. *In:* 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC 2015)

Prasetyo, N.A., Pranowo., Santoso, A.J., 2020. Automatic Detection and Calculation of Palm Oil Fresh Fruit Bunches Using Faster R-CNN. *International Journal of Applied Science and Engineering*, 17(2), pp. 121–134

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection. *In:* Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 779–88

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. Computer Vision and Pattern Recognition, pp. 1–6

Saleh, Yahya, A., Liansitim, E., 2020. Palm Oil Classification Using Deep Learning. *Science in Information Technology Letters,* Volume 1(1), pp. 1–8

Selvam, N.A.M.B., Ahmad, Z., Mohtar, I.T., 2021. Real Time Ripe Palm Oil Bunch Detection Using YOLO V3 Algorithm. *In:* 19th IEEE Student Conference on Research and Development: Sustainable Engineering and Technology towards Industry Revolution, pp. 323–328

Sunilkumar, K., Babu, D.S.S., 2013. Surface Color Based Prediction of Oil Content in Oil Palm (Elaeis Guineensis Jacq.) Fresh Fruit Bunch. *African Journal of Agricultural Research,* Volume 8(6), pp. 564–569

Szeliski, R., 2021. *Computer Vision : Algorithms and Applications 2nd Edition*. Springer

Tan, M., Le, Q.V., 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *In:* 36th International Conference on Machine Learning (ICML 2019), pp. 10691–10700

Tan, M., Pang, R., Le, Q.V., 2020. EfficientDet: Scalable and Efficient Object Detection. *In:* Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 10778–10787

Tang, Cong, Feng, Y., Yang, X., Zheng, C., Zhou, Y., 2017. The Object Detection Based on Deep Learning. *In:* 4th International Conference on Information Science and Control Engineering (ICISCE 2017), pp. 723–728

Wang, Yao, C., Liao, H.Y.M., Wu, Y.H., Chen, P.Y., Hsieh, J.Y., Yeh, I.Y., 2020. CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. *In:* IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1571–1580

Wong, Z.Y., Chew, W.J., Phang, S.K., 2020. Computer Vision Algorithm Development for Classification of Palm Fruit Ripeness. *In*: Proceedings of AIP Conference Proceedings, Volume 2233(1), p. 030012

Wu, X., Sahoo, D., Hoi, S.C.H., 2020. Recent Advances in Deep Learning for Object Detection. *Neurocomputing,* Volume 396, pp. 39–64

Xu, R, Lin, H., Lu, K., Cao, L., Liu, Y., 2021. A Forest Fire Detection System Based on Ensemble Learning. *Forests,* Volume 12(2), pp. 1–17

Yan, B., Fan, P., Lei, X., Liu, Z., Yang, F., 2021. A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5. *Remote Sensing,* Volume 13(9), pp. 1–23

Zhao, Z-Q., Zheng, P., Xu, S-T., Wu, X., 2019. Object Detection with Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, Volume 30(11), pp. 3212–3232