

Comparison of Different Deep Learning Object Detection Algorithms on Fruit Drying Characterization

Umair Mohammad Yamin¹, Norazlianie Sazali^{1,2,*}, Maurice Kettner³, Mohd Azraai Mohd Razman¹, Robert Weiβ⁴

- 1 Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
- 2 Centre of Excellence for Advanced Research in Fluid Flow (CARIFF), Universiti Malaysia Pahang Al-Sultan Abdullah, Kampung Melayu Gambang, 26300 Kuantan, Pahang, Malaysia
- 3 Institut für Kälte-, Klima- und Umwelttechnik, Karlsruhe University of Applied Sciences, 76133 Karlsruhe, Germany

4 Fakultät für Maschinenbau und Mechatronik, Karlsruhe University of Applied Sciences, 76133 Karlsruhe, Germany

1. Introduction

In the areas of industrial machinery and automotive engineering, the significance of lubrication cannot be overstated. Lubricants are indispensable for minimizing friction and abrasion between surfaces in motion, thereby enhancing performance, extending the lifespan of equipment and lowering operational expenses [1,2]. Mineral oils are frequently selected as lubricants due to their cost-effectiveness and versatility. However, due to the growing demand for lubricants, there is a pressing need for sustainable alternatives that are both ecologically sound and efficient [3].

Bio-lubricants obtained from non-edible oil sources are seen as a promising solution to tackle this urgent issue [4]. Tamanu, scientifically known as Calophyllum inophyllum, is a promising source for

^{*} *Corresponding author.*

E-mail address: azlianie@umpsa.edu.my

producing bio-lubricants [5]. The widespread occurrence of Tamanu in tropical areas, combined with its high oil production and inedible properties, makes it a promising choice to produce environmentally-friendly lubricants [6]. The attempt to utilize Tamanu oil for bio-lubricants not only provides a more environmentally friendly option compared to mineral oils, but also takes advantage of local resources, promoting regional growth and independence [7].

However, obtaining and improving the quality of Tamanu oil present distinctive obstacles [8], such as the time-consuming nature of conventional approaches and the requirement for effective categorization methods to distinguish between raw and dry Tamanu fruits [9]. In order to overcome these obstacles, it is essential to employ innovative techniques that combine technology and research [8-10].

Furthermore, the latest progress in deep learning object detection algorithms presents possibilities for enhancing production processes [11-14]. In a study conducted by Groener and colleagues (2020), various object detection algorithms were compared to assess their precision and efficiency in detecting fracking well pads and small cars [15]. Their research emphasizes the efficacy of single-stage algorithms in detecting well pads, whereas two-stage and multi-stage models demonstrate exceptional performance in identifying small cars. In a similar manner, Kim *et al.,* [16] conducted a comparison of one-stage detector algorithms and found that YOLOv5s is the most efficient. It achieves accuracy like YOLOv5l and has a speed like MobileNetv2 FPN-lite.

This study aims to investigate the viability of using Tamanu oil as a feedstock for bio-lubricants. The focus is on improving the methods used for processing and classifying the fruit. The project aims to automate the classification of Tamanu fruit based on moisture content by utilizing advanced deep learning object detection algorithms, specifically YOLOv5 [17,18]. This will streamline production processes and improve efficiency. The research aims to determine the most appropriate algorithm for accurately distinguishing between raw and dry Tamanu fruits by evaluating the performance of different object detection models.

The purposes are to capture images of Tamanu fruits in controlled conditions, apply various deep learning algorithms for fruit classification and perform a thorough analysis to identify the algorithm with the highest classification accuracy. This study aims to contribute to the progress of sustainable lubricant production by utilizing local botanical resources for economic and environmental advantage. It is an interdisciplinary endeavour.

2. Methodology

This system was developed by capturing images of raw and dry Tamanu fruits in a controlled environment to ensure uniform lighting and a standardized background. High-resolution cameras with consistent settings were used and regular checks maintained the environment's integrity. The captured images were manually annotated to accurately label the fruits, followed by preprocessing steps like resizing, normalizing pixel values and augmenting the dataset with techniques such as rotation, flipping and scaling.

Three deep learning object detection algorithms—YOLOv5m, SSD MobileNet and EfficientDet were trained using these prepared datasets under identical conditions [19,20]. Performance evaluation utilized metrics such as mean average precision (mAP), precision, recall and F1 score, with models tested on a diverse validation set of images [21]. A comparative analysis considered detection speed, confidence scores and classification accuracy. The best-performing model was further optimized and fine-tuned to enhance its accuracy and efficiency. This methodology, outlined in the flow chart in Figure 1, aims to develop an efficient and reliable system for classifying Tamanu fruits, aiding the production of sustainable bio-lubricants.

3. Results

3.1 Collect Images

In this project, a custom dataset is used instead of publicly available datasets like MS COCO or PASCAL VOC 2008 [22,23]. Images of Tamanu fruits are taken in two conditions: raw and dry. Previous students from the TOPS group have already collected images of raw and dry Tamanu fruits, but these images were not captured in a suitable environment, negatively impacting detection accuracy. To achieve the first objective of capturing raw and dry Tamanu fruits in a suitable environment, images were taken directly inside the dryer device depicted in Figure 2.

This controlled environment ensures uniform lighting and background, minimizing external variations and enhancing the object detection models' performance [24]. High-resolution images were captured using a digital camera with consistent settings for exposure, focus and white balance. Each image was manually annotated to accurately label the raw and dry Tamanu fruits. The images were then pre-processed to enhance their quality and suitability for the object detection models, including resizing, normalizing pixel values and augmenting the dataset with techniques like rotation, flipping and scaling to increase training data diversity. By using this method, the project makes sure that the custom dataset is of high quality and can be used to train deep learning object detection

models [25]. This makes the models better at telling the difference between raw and dry Tamanu fruits, which increases their accuracy and dependability.

Fig. 2. Images taken from (a) previous TOPS group (b) dryer device

3.2 Collect Images

Once the images of raw and dry Tamanu fruits have been gathered, they are sorted into two distinct folders named "raw" and "dry." Subsequently, it is imperative to assign labels to these images, as this process is vital for the training of machine learning models. For the purpose of labelling, we utilize "labellmg," a commonly employed graphical image annotation tool, in this project. LabelImg facilitates efficient and precise image annotation, simplifying the process of generating datasets for image recognition tasks. LabelImg is the preferred choice among various software options for image labelling due to its user-friendly interface and strong functionality [26].

This step guarantees that our dataset is adequately prepared for subsequent analysis and model training.

Fig. 3. Folder of (a) raw images (b) dry images

The interface allows users to exhibit images and initiate the labelling procedure, which entails delineating bounding boxes around the objects to be detected and assigning the suitable class to each object. It is possible to assign labels to multiple objects in a single image. In order to optimize the labelling process and enhance the performance of the model, it is possible to implement various measures. First and foremost, make sure to draw the bounding boxes as tightly as feasible around the objects. The precision of the object detection model aids in enhancing its ability to accurately identify objects, thereby minimizing the occurrence of false positives and enhancing the overall performance of object detection. Furthermore, include photographs of the objects captured from multiple perspectives and in diverse lighting conditions. The presence of diverse data in the dataset enhances the model's ability to generalize, thereby increasing its robustness and accuracy in detecting objects across various real-world situations [27]. By adhering to these procedures, the calibre of the annotated data is improved, which in turn facilitates more efficient training of the object detection model and ultimately yields superior precision and dependability in the detection and classification of objects.

Fig. 4. Labelled raw and dry Tamanu fruits

3.3 Comparison Results 3.3.1 Results for YOLOv5m

The entire process of training YOLOv5m on a dataset consisting of 200 images of raw Tamanu fruits and 200 images of dry Tamanu fruits was completed in 1.409 hours. The model obtained a precision score of 0.994, a recall score of 0.996 and a mAP (mean Average Precision) of 0.974.

The precision-confidence curve demonstrates the relationship between the precision of the model and the confidence levels at which it operates [28]. It provides insight into the accuracy of the model's predictions as the confidence thresholds are modified. The recall-confidence curve illustrates the relationship between recall and confidence levels, indicating how well the model can identify true positives as the thresholds for confidence change [29].

The precision-recall curve illustrates the correlation between precision and recall, showcasing how these metrics are influenced by variations in the decision threshold [30]. The F1 score, a composite measure of precision and recall, varies depending on the confidence threshold. This is demonstrated by the F1 score vs. confidence curve. This curve illustrates the fluctuation of the model's trade-off between precision and recall at various confidence levels, offering valuable insights into its overall performance.

Fig. 5. (a) Precision-confidence curve (b) recall-confidence curve (c) precision-recall curve (d) F1 confidence curve

Fig. 6. YOLOv5m real-time detection

3.3.2 Results for SSD MobileNet

The process of training SSD MobileNet on a dataset consisting of 200 images of raw Tamanu fruits and 200 images of dry Tamanu fruits was completed in a mere 10 minutes. Nevertheless, the model's performance metrics demonstrate a diminished level of accuracy in comparison to YOLOv5m. The achieved precision value is 0.899, whereas the recall value is considerably lower at 0.089, resulting in a mAP (mean Average Precision) of 0.079. These metrics indicate that although the model can be trained quickly, its capacity to accurately identify and categorize Tamanu fruits is relatively restricted.

Fig. 7. (a) SSD MobileNet mAP graph (b) SSD MobileNet Recall graph (c) SSD MobileNet total loss graph

Fig. 8. SSD MobileNet real-time detection

3.3.3 Results for EfficientDet

Training EfficientDet on a dataset of 200 images of raw Tamanu fruits and 200 images of dry Tamanu fruits took 1.5 hours. The model achieved a precision value of 0.854 and a recall value of 0.084, resulting in a mAP (mean Average Precision) of 0.079. These performance metrics indicate that, despite the similar training duration to SSD MobileNet, EfficientDet also struggled with low

recall and mAP, reflecting limitations in accurately detecting and classifying Tamanu fruits in this specific setup.

Fig. 9. (a) EfficientDet mAP graph (b) EfficientDet recall graph (c) EfficientDet total loss graph

Table 1

Fig. 10. EfficientDet real-time detection

In order to make a comparison of which deep learning object detection algorithm is better, the performance of each method is evaluated. The test was performed with 200 images of raw Tamanu fruits and 200 images of dry Tamanu fruits. The important parameters which need to be look at to do comparison between these object detection models are the precision, recall, mean Average Precision (mAP), F1-score and the time taken for the training to be completed. The value of precision and recall at IoU=0.5:0.95 are chosen for all these algorithms to evaluate these performances.

The graph which has been generated for YOLOv5m is different with SSD MobileNet and EfficientDet due to different type of deep learning frameworks [18-20]. The results of all three deep learning object detection are summarized in Table 1 below.

YOLOv5m has outperformed the other two models with the highest value of accuracy where it has 0.974 compared to SSD MobileNet and EfficientDet with mAP value of 0.899 and 0.854 respectively. YOLOv5m has the highest value for recall with 0.996. However, both of object detection model by Tensorflow framework which are SSD MobileNet and EfficientDet achieve very low value of recall which are 0.089 and 0.084 respectively. The value of mAP and F1-score are calculated by using the precision and recall value, which means that YOLOv5m will always outperform both SSD MobileNet and EfficientDet in terms of mAP value and F1-score [18-20]. YOLOv5m is ranked as number one, followed by SSD MobileNet and EfficientDet.

The time taken for the images to be trained by YOLOv5m is 1.4 hour which is longer than SSD MobileNet, which only need 0.1 hour to train. The time taken for EfficientDet is 1.5 hour, the longest time needed to train among these 3 algorithms. However, when using real-time object detection, SSD MobileNet outperforms YOLOv5m and EfficientDet [18-20]. SSD MobileNet able to detect most of the Tamanu fruits correctly and has the high value of confidence interval for each of the detected Tamanu fruits if compared to YOLOv5m and EfficientDet for real-time detection.

4. Conclusions

The raw and dry Tamanu fruits were gathered in a controlled setting using a drying apparatus, guaranteeing reliable and superior data for training purposes. These fruits were classified using three advanced deep learning object detection algorithms: YOLOv5m, SSD MobileNet and EfficientDet. Each algorithm exhibited efficacy in differentiating between unprocessed and dehydrated Tamanu fruits. After careful evaluation, YOLOv5m proved to be the superior model, demonstrating the highest precision and recall values. This indicates its exceptional performance in accurately classifying fruits. SSD MobileNet outperformed YOLOv5m and EfficientDet in real-time object detection, particularly in practical scenarios where fast detection is crucial. The discrepancy in performance underscores that distinct models, with their individual architectural foundations, provide diverse strengths. Although YOLOv5m offers superior mean average precision (mAP) and recall values, the real-time detection advantage of SSD MobileNet highlights the inherent trade-offs between accuracy and speed. In this context, EfficientDet did not achieve comparable precision, recall and mAP performance despite having similar training times to SSD MobileNet.

Acknowledgement

This research was funded by a grant from Universiti Malaysia Pahang for funding under grant RDU232707 and UIC231511.

References

- [1] Szpunar, Marcin, Tomasz Trzepieciński, Krzysztof Żaba, Robert Ostrowski, and Marek Zwolak. "Effect of lubricant type on the friction behaviours and surface topography in metal forming of Ti-6Al-4V titanium alloy sheets." *Materials* 14, no. 13 (2021): 3721. <https://doi.org/10.3390/ma14133721>
- [2] Garcia Tobar, Milton, Rafael Wilmer Contreras Urgiles, Bryan Jimenez Cordero, and Julio Guillen Matute. "Nanotechnology in Lubricants: A Systematic Review of the Use of Nanoparticles to Reduce the Friction Coefficient." *Lubricants* 12, no. 5 (2024): 166. <https://doi.org/10.3390/lubricants12050166>
- [3] Basiron, Juliana, Mohd Fadzli Bin Abdollah, Muhammad Ilman Chua Abdullah, and Hilmi Amiruddin. "Lubricant mechanisms of eco-friendly lubricant blended with mineral oil for steel-steel contact." *Tribology International* 186 (2023): 108653. <https://doi.org/10.1016/j.triboint.2023.108653>
- [4] Joshi, Jeimin R., Khodidas K. Bhanderi, and Jigar V. Patel. "A review on bio-lubricants from non-edible oils-recent advances, chemical modifications and applications." *Journal of the Indian Chemical Society* 100, no. 1 (2023): 100849. <https://doi.org/10.1016/j.jics.2022.100849>
- [5] Pribowo, Amadeus, Jyothsna Girish, Marsia Gustiananda, Rakrya Galih Nandhira, and Pietradewi Hartrianti. "Potential of Tamanu (Calophyllum inophyllum) Oil for Atopic Dermatitis Treatment." *Evidence-Based Complementary and Alternative Medicine* 2021, no. 1 (2021): 6332867. <https://doi.org/10.1155/2021/6332867>
- [6] Ulya, Nur Arifatul, Edwin Martin, Mamat Rahmat, Bambang Tejo Premono, Leo Rio Ependi Malau, Efendi Agus Waluyo, Andika Imanullah *et al.,* "Enabling Factors of NTFP Business Development for Ecosystem Restoration: The Case of Tamanu Oil in Indonesian Degraded Peatland." *Sustainability* 14, no. 17 (2022): 10681. <https://doi.org/10.3390/su141710681>
- [7] Malik, Muhammad Ali Ijaz, M. A. Kalam, M. A. Mujtaba, and Fares Almomani. "A review of recent advances in the synthesis of environmentally friendly, sustainable, and nontoxic bio-lubricants: Recommendations for the future implementations." *Environmental Technology & Innovation* (2023): 103366. <https://doi.org/10.1016/j.eti.2023.103366>
- [8] Hapsari, Safrina, Hakun Wirawasista Aparamarta, Nurul Jadid, and Setiyo Gunawan. "Extraction of Coumarin Mixture from Tamanu Oil using Food-Grade." *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 111, no. 2 (2023): 1-15. <https://doi.org/10.37934/arfmts.111.2.115>
- [9] Roads, Brett D., Buyun Xu, June K. Robinson, and James W. Tanaka. "The easy-to-hard training advantage with realworld medical images." *Cognitive Research: Principles and Implications* 3 (2018): 1-13. <https://doi.org/10.1186/s41235-018-0131-6>
- [10] Almaliki, Alaa Jabbar Qasim, Sajad Muhil Abd, Inam Abdullah Lafta, Roshidi Din, Osman Ghazali, Jabbar Qasim Almaliki, and Sunariya Utama. "Application of the Canny Filter in Digital Steganography." *Journal of Advanced Research in Computing and Applications* 35, no. 1 (2024): 21-30. <https://doi.org/10.37934/arca.35.1.2130>
- [11] Xiao, Youzi, Zhiqiang Tian, Jiachen Yu, Yinshu Zhang, Shuai Liu, Shaoyi Du, and Xuguang Lan. "A review of object detection based on deep learning." *Multimedia Tools and Applications* 79 (2020): 23729-23791. <https://doi.org/10.1007/s11042-020-08976-6>
- [12] Wu, Xiongwei, Doyen Sahoo, and Steven CH Hoi. "Recent advances in deep learning for object detection." *Neurocomputing* 396 (2020): 39-64. <https://doi.org/10.1016/j.neucom.2020.01.085>
- [13] Groener, Austen, Gary Chern, and Mark Pritt. "A comparison of deep learning object detection models for satellite imagery." In *2019 IEEE applied imagery pattern recognition workshop (AIPR)*, pp. 1-10. IEEE, 2019. <https://doi.org/10.1109/AIPR47015.2019.9174593>
- [14] Hashim, Mohd Ekram Alhafis, Nur Safinas Albakry, Wan Azani Mustafa, Banung Grahita, Miharaini Md Ghani, Hafizul Fahri Hanafi, Suraya Md Nasir, and Catherina ana Ugap. "Understanding the Impact of Animation Technology in Virtual Reality: A Systematic Literature Review." *International Journal of Computational Thinking and Data Science* 1, no. 1 (2024): 53-65. <https://doi.org/10.37934/CTDS.1.1.5365>
- [15] Kim, Jeong-ah, Ju-Yeong Sung, and Se-ho Park. "Comparison of Faster-RCNN, YOLO, and SSD for real-time vehicle type recognition." In *2020 IEEE international conference on consumer electronics-Asia (ICCE-Asia)*, pp. 1-4. IEEE, 2020. <https://doi.org/10.1109/ICCE-Asia49877.2020.9277040>
- [16] Arip, Afifuddin Arif Shihabuddin, Norazlianie Sazali, Kumaran Kadirgama, Ahmad Shahir Jamaludin, Faiz Mohd Turan, and Norhaida Ab Razak. "Object Detection for Safety Attire Using YOLO (You Only Look Once)." *Journal of Advanced Research in Applied Mechanics* 113, no. 1 (2024): 37-51. <https://doi.org/10.37934/aram.113.1.3751>
- [17] Shatnawi, Hashem, and Mohammad N. Alqahtani. "Delving into the Revolutionary Impact of Artificial Intelligence on Mechanical Systems: A Review." *Semarak International Journal of Machine Learning* 1, no. 1 (2024): 31-40. <https://doi.org/10.37934/sijml.1.1.3140>
- [18] Roslan, Shairatul Akma, Fitri Yakub, Shuib Rambat, Sharifah Munawwarah, Mokhtar Saidin, Farah Liana, Nurshafinaz Mohd Maruai, Mohamed Sukri Mat Ali, and Ahmad Faiz Mohammad. "The Novel Method in Validating the Spectral Wavelength Optimization to Determine Archaeological Proxies by the Integration of Aerial and Ground Platforms." Journal of Advanced Research in Applied Mechanics 108, no. 1 (2023): <https://doi.org/10.37934/sijese.2.1.115>
- [19] Chiu, Yu-Chen, Chi-Yi Tsai, Mind-Da Ruan, Guan-Yu Shen, and Tsu-Tian Lee. "Mobilenet-SSDv2: An improved object detection model for embedded systems." In *2020 International conference on system science and engineering (ICSSE)*, pp. 1-5. IEEE, 2020. <https://doi.org/10.1109/ICSSE50014.2020.9219319>
- [20] Afif, Mouna, Riadh Ayachi, Yahia Said, and Mohamed Atri. "An evaluation of EfficientDet for object detection used for indoor robots assistance navigation." *Journal of Real-Time Image Processing* 19, no. 3 (2022): 651-661. <https://doi.org/10.1007/s11554-022-01212-4>
- [21] Reddy, B. Haranadh, and P. R. Karthikeyan. "Classification of fire and smoke images using decision tree algorithm in comparison with logistic regression to measure accuracy, precision, recall, F-score." In *2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, pp. 1-5. IEEE, 2022. <https://doi.org/10.1109/MACS56771.2022.10022449>
- [22] Tong, Kang, and Yiquan Wu. "Rethinking PASCAL-VOC and MS-COCO dataset for small object detection." *Journal of Visual Communication and Image Representation* 93 (2023): 103830.<https://doi.org/10.1016/j.jvcir.2023.103830>
- [23] Perronnin, Florent, Jorge Sánchez, and Thomas Mensink. "Improving the fisher kernel for large-scale image classification." In *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11*, pp. 143-156. Springer Berlin Heidelberg, 2010. https://doi.org/10.1007/978-3-642-15561-1_11
- [24] Sankaranarayanan, Aswin C., Ashok Veeraraghavan, and Rama Chellappa. "Object detection, tracking and recognition for multiple smart cameras." *Proceedings of the IEEE* 96, no. 10 (2008): 1606-1624. <https://doi.org/10.1109/JPROC.2008.928758>
- [25] Xiao, Bo, and Shih-Chung Kang. "Development of an image data set of construction machines for deep learning object detection." *Journal of Computing in Civil Engineering* 35, no. 2 (2021): 05020005. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000945](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000945)
- [26] Xiong, Yijin, Xin Gao, and Guoying Zhang. "Interactive object annotation based on one-click guidance." *The Journal of Supercomputing* 79, no. 14 (2023): 16098-16117.<https://doi.org/10.1007/s11227-023-05279-z>
- [27] Nuri, Nur Rashid Mat, Mohd Idain Fahmy Rosley, Mohd Erdi Ayob, Shafiee Mohamad, and Nur Adlin Abu Bakar. "Real-time Water Quality Monitoring System using IoT: Application at Melaka River, Malaysia." *Journal of Advanced Research in Applied Mechanics* 121, no. 1 (2024): 66-78[. https://doi.org/10.37934/aram.121.1.6678](https://doi.org/10.37934/aram.121.1.6678)
- [28] Yang, Qifan, Minyi Su, Yan Li, and Renxiao Wang. "Revisiting the relationship between correlation coefficient, confidence level, and sample size." *Journal of chemical information and modeling* 59, no. 11 (2019): 4602-4612. <https://doi.org/10.1021/acs.jcim.9b00214>
- [29] Zawadzka, Katarzyna, and Philip A. Higham. "Judgments of learning index relative confidence, not subjective probability." *Memory & cognition* 43 (2015): 1168-1179.<https://doi.org/10.3758/s13421-015-0532-4>
- [30] Sofaer, Helen R., Jennifer A. Hoeting, and Catherine S. Jarnevich. "The area under the precision-recall curve as a performance metric for rare binary events." *Methods in Ecology and Evolution* 10, no. 4 (2019): 565-577. <https://doi.org/10.1111/2041-210X.13140>

