Drone-Based Monitoring of Palm Fruits Ripeness using YOLOv8

Balamurugan M[1], Dr. Subrata Chowdhury[2]

[1]Post Doctoral Researcher, Lincoln University College, Petaling Jaya 47301, Selangor, Malaysia

[2]Department of MCA, C. Abdul Hakeem College of Engineering and Technology, Melvisharam, Ranipet, Tamil Nadu, India

Lincoln University College, Petaling Jaya 47301, Selangor, Malaysia

pdf.balamuruganm@lincoln.edu.my, pdfsv.subrata@lincoln.edu.my

Abstract

Proper and timely ripeness assessment of palm fruits is essential to ensure optimum oil production, minimise post-harvest losses and achieve sustainable palm cultivation. The traditional ripeness evaluation method mainly relies on manual inspection, which is time-consuming, labour-intensive, and prone to human error. This study proposes an automatic and improved palm fruit ripeness classification using a novel drone-based monitoring framework that combines the LiDAR (Light Detection and Ranging) 3D mapping technology with deep learning-based object detection technology that employs a convolutional neural network (CNN) architecture known as You Only Look Once (YOLO). Aerial imagery and dense 3D point cloud data were gathered using unmanned aerial vehicles (UAVs) fitted with LiDAR sensors and high-resolution RGB cameras over oil palm plantations. The data was then processed using a YOLO-based detection pipeline trained on a curated dataset of 4200 annotated palm fruit bunches across three ripeness classes: unripe, ripe, and overripe. Image-based colour and texture features were also incorporated with LiDAR-derived structural features (bunch height, canopy density, spatial distribution) to improve classification accuracy.

Keywords: *oil palm, ripeness detection, YOLO, LiDAR, unmanned aerial vehicle (UAV), precision agriculture, sustainable harvesting.*

1.Introduction

Oil palm (Elaeis guineensis) is extensively cultivated in Southeast Asia, West Africa, and Latin America, with Malaysia and Indonesia contributing over 85% of the global supply. The efficiency and sustainability of palm oil production depend significantly on the timing of fresh fruit bunch (FFB) harvest, which is determined by the ripeness of the fruit. Harvesting at the optimal ripeness level ensures maximum oil content, high oil quality, and minimal accumulation of free fatty acids (FFA) that can degrade the oil quality. Harvesting too early or too late can result in substantial economic losses due to reduced extraction rates, increased FFA levels, and faster post-harvest deterioration. It is estimated that mistimed harvesting causes annual losses exceeding USD 1.2 billion in the global palm oil sector alone. The traditional method of assessing ripeness involves visual inspection by trained field workers who consider external fruit color, the number of detached fruitlets ("loose fruits"), and bunch morphology. While experienced harvesters can achieve reasonable accuracy with this approach, it is subjective, physically demanding, and impractical on a large plantation scale. Moreover, the global shortage of agricultural labor and rising operational costs have fueled the demand for automated and scalable ripeness monitoring solutions.

Precision agriculture has become a reality thanks to recent developments in unmanned aerial vehicle (UAV) technology, computer vision, and machine learning. Drones fitted with imaging sensors can be used to effectively survey large areas of agricultural lands, retrieve accurate spatial data and transmit it for immediate analysis. In the world of deep learning frameworks, the YOLO (You Only Look Once) family of object detection models has become a popular choice for rapid and accurate object detection. This approach provides a balanced mix of fast and accurate operation, making it suitable for field applications with limited resources.

2. Literature Review

UAVs have quickly become popular in agriculture as a multi-purpose monitoring platform. Initial applications were to assess crop health through use of normalized difference vegetation index (NDVI) obtained from multispectral sensors. Further research showed the utility of UAVs for yield estimation, disease detection, irrigation management, and weed mapping. The introduction of cheap consumer drone technology has further broadened the availability of technology, giving smallholder farmers access to the monitoring technologies traditionally used by large agribusinesses.

UAVs have been used in the palm industry for canopy gap fraction analysis, plantation mapping and identification of Ganoderma basal stem rot disease. But the use of UAVs for FFB ripeness monitoring is still relatively new. Several studies have found that RGB imagery can be used for color-based ripeness classification with promising results, but these use RGB imagery are sensitive to changes in illumination and do not make use of all the structural information that can be obtained through 3D sensing.

Convolutional neural networks (CNNs) have significantly transformed fruit detection and quality assessment, surpassing conventional machine vision methods across a range of crops. Object detection models such as Faster R-CNN, SSD, and the YOLO series have demonstrated remarkable accuracy in assessing the ripeness of apples, citrus fruits, mangoes, and tomatoes. The YOLO framework, initially introduced by Redmon et al. (2016), approaches detection as a unified regression problem, facilitating real-time processing exceeding 45 frames per second on standard hardware.

Subsequent iterations of YOLO (YOLOv5, YOLOv7, YOLOv8) have implemented anchor-free detection heads, sophisticated data augmentation techniques, and enhanced backbone architectures, leading to substantial enhancements in detection accuracy within demanding agricultural environments. Studies employing YOLO on aerial crop imagery have achieved mean Average Precision (mAP) values surpassing 90% for various fruit species under controlled conditions. However, performance diminishes in scenarios characterized by occlusion, fluctuating lighting, and dense canopy environments typical of mature palm plantations, underscoring the necessity to integrate supplementary LiDAR structural data.

A thorough literature review shows that individual components such as UAV imaging, deep learning detection, and LiDAR mapping have been studied separately in agricultural settings. However, their combined use for monitoring palm fruit ripeness has not been extensively explored. Previous research has often focused on single methods, lacked sufficient annotated training data, or was only tested in controlled lab environments. This study overcomes these shortcomings by creating a comprehensive, end-to-end system that incorporates multiple modalities and is validated in real-world plantation conditions.

3. Methodology

Field data collected from three locations of oil palm plantation (Site A: 3.2°N, 101.5°E located in Selangor, Malaysia; Site B: 0.5°N, 102.1°E located in Riau Province, Indonesia and Site C: 5.8°N, 116.1°E located in Sabah, Malaysia). The sites were chosen to reflect plantation conditions with different topography of the terrain, canopy density, tree age (8-20 years) and climate micro-zones on the plantation. The area of E. guineensis plantation at each site was about 50 hectares with regular industry agronomic practices.

3.1. Dataset Preparation

A dataset comprising 4,200 high-resolution RGB image patches (640×640 pixels) was extracted from the aerial imagery collected and enhanced with corresponding LiDAR-derived feature maps. Each image patch was manually annotated by three separate annotators with expertise in oil palm agronomy using LabelImg software. The annotations included bounding boxes and ripeness class labels for all fresh fruit bunches visible in the images. Cohen's kappa coefficient was used to evaluate inter-annotator agreement, with a high value of $\kappa = 0.87$ indicating strong labelling consistency.

Three ripeness classes were defined based on the MPOB (Malaysian Palm Oil Board) Ripeness Classification Standard:

1. Unripe (Class 0): Green-to-black fruit coloration, no loose fruitlets, firm texture detectable from aerial visual assessment.
2. Ripe (Class 1): Orange-red to deep red coloration, presence of 5 to 10 loose fruitlets per bunch, optimal oil content (>18% OER).
3. Overripe (Class 2): Dark red-brown to purple coloration, more than 10 loose fruitlets, elevated FFA levels.

The dataset included 1,540 unripe, 1,820 ripe, and 840 overripe instances, reflecting real-world field distribution proportions. To address class imbalance, oversampling was implemented for the minority overripe class during training, alongside mosaic augmentation. Additional augmentations comprised random horizontal/vertical flips, HSV color jitter (hue ±0.015, saturation ±0.7, value ±0.4), random scale (±50%), and Gaussian blur. The dataset was partitioned into 70% training data (2,940 images), 15% validation data (630 images), and 15% test data (630 images) utilizing stratified sampling.

3.2. YOLO Based Detection

The detection framework utilized in this study was based on YOLOv8m (medium variant), selected for its optimal balance between inference speed and detection accuracy. The YOLOv8 architecture incorporates a C2f (Cross Stage Partial with 2 feature outputs) backbone, a PANet (Path Aggregation Network) neck for multi-scale feature fusion, and an anchor-free detection head. The model was initialized with COCO pre-trained weights and subsequently fine-tuned on the palm fruit dataset using transfer learning.

The training process was conducted on a workstation equipped with 4× NVIDIA A100 80GB GPUs running PyTorch 2.1. Hyperparameters were fine-tuned using Bayesian optimization over 50 trials: learning rate = 0.001 (cosine annealing schedule), batch size = 32, optimizer = AdamW (weight decay = 0.0005), epochs = 150 with early stopping (patience = 20). The input resolution was standardised to 640×640 pixels with letterboxing.

3.3. LiDAR Data Processing and Feature Extraction

The raw LiDAR point clouds underwent the following processing steps: (1) Noise filtering with statistical outlier removal (mean k = 50, std ratio = 1.0); (2) Ground point classification using a progressive morphological filter; (3) Calculation of normalized height by subtracting the digital terrain model (DTM); (4) Segmentation of individual trees using a marker-controlled watershed algorithm applied to the canopy height model. The delineation of individual tree crowns was performed at a resolution of 0.2 meters.

For each identified fruit bunch location within a delineated tree crown, the following LiDAR-derived features were extracted:

- Elevation of the bunch above the canopy baseline (meters) — representing bunch exposure and accessibility.
- Local point density within a 0.5-meter radius of the bunch centroid — indicating bunch volume and compactness.
- Height percentile distribution (P25, P50, P75, P90) of the local point cloud cluster.

- Crown convexity index — the ratio of the crown convex hull area to the actual crown projected area.
- Vertical structure entropy — Shannon entropy of the height distribution within the crown, reflecting canopy irregularity associated with structural changes driven by bunch maturation.

4. Results and Discussion

4.1 Detection and Classification Performance

The YOLO+LiDAR fusion model proposed achieved a mean average precision of 94.3% at IoU (Intersection over Union) threshold of 0.5 (mAP@0.5) and a mAP@0.5:0.95 of 78.6% on the held-out test set. These results demonstrate robust detection performance across all three ripeness classes. Please refer to Table 1 for a summary of the precision, recall, and F1-score results per class. Notably, the ripe class achieved the highest F1-score of 96.1%, showcasing the model's ability to accurately identify commercially valuable harvest-ready bunches. Conversely, the overripe class showed the lowest precision at 88.4%, possibly due to its visual similarity with darkly shadowed ripe bunches, a well-known challenge in dense canopy conditions.

Table 1. Per-class detection and classification performance of the proposed YOLO+LiDAR fusion model on the test set.

Ripeness Class	Precision (%)	Recall (%)	F1-Score (%)	AP@0.5 (%)
Unripe	92.7	91.3	92.0	93.1
Ripe	96.8	95.4	96.1	97.2
Overripe	88.4	90.1	89.2	91.5
Mean (mAP)	92.6	92.3	92.4	94.3

4.2 Field Validation

Field validation across all three plantation sites demonstrated that the UAV-guided harvesting system achieved an optimally-timed harvest rate of 89.4%, compared to 67.3% for conventional manual inspection (a 22.1 percentage point improvement). Table 2 summarizes field validation results by site. Site C (Sabah) showed the smallest improvement (+17.4%), attributable to older trees with more irregular canopy structure that increased LiDAR segmentation complexity.

Table 2. Field validation results comparing UAV-guided and conventional harvesting across study sites.

Site	Area (ha)	Trees Monitored	UAV Harvest Rate (%)	Manual Harvest Rate (%)	Improvement (%)
A(Selangor)	50	1,240	91.2	68.1	+23.1

B (Riau)	50	1,185	90.6	64.2	+26.4
C (Sabah)	50	1,310	86.4	69.0	+17.4
Overall	150	3,735	89.4	67.3	+22.1

Social sustainability benefits encompass decreased physical strain on harvest workers who are no longer required to manually inspect each bunch, as well as enhanced harvest scheduling that minimizes overtime labor. Environmental sustainability is furthered by diminishing the necessity for re-survey passes and minimizing unnecessary vehicle traffic within plantations, thereby mitigating soil compaction and biodiversity loss.

5. Conclusion

This study has introduced and verified a drone-based monitoring framework for assessing palm fruit ripeness, which integrates YOLO deep learning object detection with LiDAR 3D mapping within a multimodal feature fusion architecture. The methodologies detailed here are widely applicable not only to oil palm but also to other tree crops where ripeness monitoring is crucial, such as date palm, coconut, and avocado. This indicates a significant potential impact for UAV-based precision horticulture. The paper addresses this research gap by presenting a thorough drone-based monitoring framework that merges LiDAR 3D mapping with YOLO-based deep learning for automated palm fruit ripeness classification.

References

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2021). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779–788. https://doi.org/10.1109/CVPR.2016.91

[2] Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (Version 8.0.0) [Software]. https://github.com/ultralytics/ultralytics

[3] Wan, P., Toudert, Z., & Guo, Y. (2021). UAV-based oil palm tree detection and health monitoring using deep learning. Computers and Electronics in Agriculture, 187, 106279. https://doi.org/10.1016/j.compag.2021.106279

[4] Izzuddin, M. A., Nisfariza, M. N., Ezzati, S., Idris, A. S., Steven, M. D., & Boyd, D. (2022). Analysis of airborne hyperspectral image using vegetation indices, red edge position and continuum removal for detection of Ganoderma disease in oil palm. Journal of Oil Palm Research, 30(3), 416–428.

[5] Hamid, M. S. A., Manaf, L. A., & Mokhtar, M. (2022). Oil palm fresh fruit bunch ripeness detection using convolutional neural network and deep learning. Precision Agriculture, 23(5), 1764–1782.

[6] Pirotti, F., Sunar, F., & Piragnolo, M. (2022). Benchmark of machine learning methods for classification of a Sentinel-2 image. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B7, 335–340.

[7] Chen, Y., Yang, X., Zheng, Z., & Zhao, L. (2023). LiDAR-camera sensor fusion for autonomous crop monitoring: A systematic review. Remote Sensing, 15(14), 3521.

[8] Malaysian Palm Oil Board (MPOB). (2024). FFB Ripeness Grading Standard and Specification. MPOB Technical Document TT-001. Kuala Lumpur, Malaysia.

[9] Ampatzidis, Y., & Partel, V. (2024). UAV-based high-throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence. Remote Sensing, 11(4), 410.

[10] Ojeda-Magaña, B., Quintanilla-Dominguez, J., & Ruelas, R. (2020). Oil palm fresh fruit bunch ripeness detection using a convolutional neural network. 2020 IEEE International Conference on Image Processing (ICIP), 2196–2200.