A Comprehensive Study on Drone-Based Monitoring of Palm Fruits Ripeness for Sustainable Harvesting

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

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

[2]Department of Computer Science Engineering, SVCET College, Chittoor, Andhra Pradesh, India

Lincoln University College, Petaling Jaya 47301, Selangor, Malaysia

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

ABSTRACT

Accurate and timely assessment of palm fruit ripeness is critical for maximizing oil yield, reducing post-harvest losses, and promoting sustainable harvesting practices. Traditional ripeness evaluation methods rely heavily on manual inspection, which is labor-intensive, time-consuming, and prone to human error. This study presents a novel drone-based monitoring framework that integrates LiDAR (Light Detection and Ranging) 3D mapping with deep learning-based object detection—specifically the You Only Look Once (YOLO) convolutional neural network architecture—to automate and enhance palm fruit ripeness classification. Unmanned aerial vehicles (UAVs) equipped with LiDAR sensors and high-resolution RGB cameras were deployed over oil palm plantations to acquire dense 3D point cloud data and aerial imagery. The collected data were processed through a YOLO-based detection pipeline trained on a curated dataset of 4,200 annotated palm fruit bunch images spanning three ripeness categories: unripe, ripe, and overripe. LiDAR-derived structural features including bunch height, canopy density, and spatial distribution were fused with image-based color and texture features to improve classification accuracy.

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

1. Introduction

Palm oil is one of the world's most widely consumed vegetable oils, accounting for approximately 35% of global vegetable oil production. Oil palm (Elaeis guineensis) is cultivated extensively across Southeast Asia, West Africa, and Latin America, with Malaysia and Indonesia together contributing over 85% of global supply. The efficiency and sustainability of palm oil production hinge critically on the timing of fresh fruit bunch (FFB) harvest—a stage governed entirely by the ripeness of the fruit.

Harvesting at optimal ripeness ensures maximum oil content, high oil quality, and minimal free fatty acid (FFA) accumulation, which degrades oil quality. Conversely, premature or delayed harvesting leads to significant economic losses through reduced extraction rates, increased FFA levels, and accelerated post-harvest deterioration. Industry estimates suggest that mistimed harvesting contributes to annual losses exceeding USD 1.2 billion globally in the palm oil sector alone.

Conventional ripeness assessment relies on visual inspection by trained field workers, who evaluate external fruit coloration, the number of detached fruitlets ("loose fruits"), and bunch morphology. While experienced harvesters achieve reasonable accuracy, this approach is inherently subjective, physically demanding, and impractical at plantation scale. Furthermore, the global shortage of agricultural labor and rising operational costs have intensified the demand for automated, scalable ripeness monitoring solutions.

Recent advances in unmanned aerial vehicle (UAV) technology, computer vision, and machine learning have created new opportunities for precision agriculture applications. Drones equipped with imaging sensors can rapidly cover large plantation areas, acquire high-resolution spatial data, and relay information for near-real-time analysis. Among deep learning frameworks, the YOLO (You Only Look Once) family of object detection models has emerged as a leading solution for fast, accurate object detection, offering a compelling combination of speed and precision suitable for deployment in resource-constrained field environments.

LiDAR (Light Detection and Ranging) technology adds a critical three-dimensional sensing dimension absent from camera-only systems. By generating dense point clouds of plantation canopies, LiDAR enables the extraction of structural features—such as bunch height relative to the canopy, bunch volume estimation, and spatial cluster analysis—that complement the spectral and textural information captured by cameras. The fusion of LiDAR-derived structural features with image-based deep learning features represents a particularly promising avenue for improving ripeness classification robustness.

2. Related Works

UAVs have rapidly gained traction as versatile platforms for agricultural monitoring. Early applications focused on crop health assessment using normalized difference vegetation index (NDVI) derived from multispectral sensors. Subsequent studies demonstrated UAV utility for yield estimation, disease detection, irrigation management, and weed mapping. The proliferation of affordable consumer drones has further democratized access, enabling smallholder farmers to benefit from aerial monitoring capabilities previously reserved for large agribusinesses.

In the context of palm cultivation, UAVs have been employed for canopy gap fraction analysis, plantation mapping, and detection of Ganoderma basal stem rot disease. However, the application of UAVs specifically for FFB ripeness monitoring remains relatively nascent. Several studies have reported promising results using RGB imagery for color-based ripeness classification, but these approaches are sensitive to illumination variability and do not exploit the full structural information available from 3D sensing.

Convolutional neural networks (CNNs) have transformed fruit detection and quality assessment, surpassing traditional machine vision methods across a wide range of crops. Object detection architectures such as Faster R-CNN, SSD, and the YOLO family have been applied to apple, citrus, mango, and tomato ripeness assessment with high accuracy. The YOLO architecture, first proposed by Redmon et al. (2016), frames detection as a single regression problem, enabling real-time inference at over 45 frames per second on standard hardware.

YOLOv5 and subsequent versions (YOLOv7, YOLOv8) introduced anchor-free detection heads, advanced data augmentation strategies, and improved backbone architectures, substantially boosting detection accuracy on challenging agricultural scenes. Studies applying YOLO to aerial crop imagery have demonstrated mAP values exceeding 90% for various fruit species under controlled conditions. However, performance typically degrades under occlusion, variable lighting, and dense canopy environments characteristic of mature palm plantations—conditions that motivate the inclusion of supplementary LiDAR structural information.

The integration of LiDAR with optical imagery, known as sensor fusion, has been shown to improve land cover classification, species identification, and biomass estimation beyond the capabilities of either modality alone. For fruit ripeness assessment specifically, LiDAR point cloud data can encode the physical protrusion and dimensional changes that accompany fruit bunch maturation, providing complementary cues to the color and texture features captured by cameras.

A comprehensive review of the literature reveals that while individual components—UAV imaging, deep learning detection, and LiDAR mapping—have each been studied in agricultural contexts, their integrated application for palm fruit ripeness monitoring remains underexplored. Prior work has typically relied on single modalities, lacked large annotated training datasets, or was validated only in controlled laboratory settings. This study addresses these limitations by developing a multimodal, end-to-end system validated in operational plantation conditions.

3. Methodology

Field data were collected from three oil palm plantation sites located in Selangor, Malaysia (Site A: 3.2°N, 101.5°E), Riau Province, Indonesia (Site B: 0.5°N, 102.1°E), and Sabah, Malaysia (Site C: 5.8°N, 116.1°E). Sites were selected to represent diverse plantation conditions including varying terrain topography, canopy density, tree age (8–20 years), and climate micro-zones. Each site covered approximately 50 hectares of mature E. guineensis cultivation, managed under standard industry agronomic practices.

A total of 12 UAV flight missions were conducted across the three sites during the fruiting season (June–September 2024), with each mission covering between 10 and 15 hectares at a flight altitude of 30 meters above ground level (AGL). All flights were conducted between 08:00 and 11:00 local time to minimize solar angle effects on image quality. Ground truth ripeness labels were collected concurrently by experienced plantation agronomists for a stratified random sample of 450 individual fruit bunches per site, resulting in 1,350 ground-labeled samples.

3.2 UAV Platform and Sensor Configuration

A custom-built hexacopter UAV platform was configured with the following sensor payload:

- RGB camera: Sony RX1R II (42.4 MP full-frame sensor, 35mm lens) for high-resolution optical imagery at 1.2 cm/pixel ground sampling distance (GSD) at 30 m AGL.
- LiDAR unit: Velodyne VLP-16 Puck LITE (16-channel, 360° horizontal FOV, 30° vertical FOV, 100 m range, ~300,000 points/second) for 3D point cloud acquisition.

- GNSS/IMU unit: Applanix APX-15 with real-time kinematic (RTK) correction for sub-5 cm positional accuracy.
- Onboard processor: NVIDIA Jetson Xavier NX for real-time YOLO inference during data collection missions.

Flight parameters were set to achieve 80% forward and 70% lateral image overlap, ensuring complete stereo coverage for structure-from-motion (SfM) photogrammetric reconstruction. LiDAR scan rate was synchronized with flight speed (5 m/s) to maintain point density exceeding 150 points/m² across the survey area.

3.3 Dataset Preparation and Annotation

A dataset of 4,200 high-resolution RGB image patches (640×640 pixels) was extracted from the acquired aerial imagery, augmented with corresponding LiDAR-derived feature maps. Each image patch was manually annotated using LabelImg software by three independent annotators with domain expertise in oil palm agronomy. Annotations consisted of bounding boxes and ripeness class labels for all visible fresh fruit bunches. Inter-annotator agreement was assessed using Cohen's kappa coefficient ($\kappa = 0.87$), indicating high labeling 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 comprised 1,540 unripe, 1,820 ripe, and 840 overripe instances, reflecting realistic field distribution proportions. To address class imbalance, oversampling with mosaic augmentation was applied to the minority overripe class during training. Additional augmentations included 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 (2,940 images), 15% validation (630 images), and 15% test (630 images) sets with stratified sampling.

3.4 YOLO-Based Detection Model

The detection backbone was built upon YOLOv8m (medium variant), selected for its balance of inference speed and detection accuracy. The YOLOv8 architecture employs 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 fine-tuned on the palm fruit dataset through transfer learning.

Training was conducted on a workstation equipped with 4× NVIDIA A100 80GB GPUs using PyTorch 2.1. Hyperparameters were optimized through 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 set to 640×640 pixels with letterboxing.

3.5 LiDAR Data Processing and Feature Extraction

Raw LiDAR point clouds were processed using the following pipeline: (1) noise filtering with statistical outlier removal (mean k = 50, std ratio = 1.0); (2) ground point classification using a progressive morphological filter; (3) normalized height computation by subtracting the digital terrain model (DTM); (4) individual tree segmentation using a marker-controlled watershed algorithm applied to the canopy height model. Individual tree crowns were delineated at 0.2 m resolution.

Within each delineated tree crown, the following LiDAR-derived features were extracted for each detected fruit bunch location:

- Bunch elevation above canopy baseline (meters) — proxy for bunch exposure and accessibility.
- Local point density within 0.5 m radius of bunch centroid — indicator of bunch volume and compactness.
- Height percentile distribution (P25, P50, P75, P90) of the local point cloud cluster.
- Crown convexity index — ratio of crown convex hull area to actual crown projected area.
- Vertical structure entropy — Shannon entropy of height distribution within the crown, indicating canopy irregularity associated with bunch maturation-driven structural change.

LiDAR features were co-registered with RGB image coordinates using the RTK-GNSS trajectory and rigid body transformation parameters. For each detected bounding box from the YOLO model, the corresponding LiDAR feature vector was retrieved via spatial join within a 0.3 m radius tolerance.

4. Conclusion

This study has presented and validated a drone-based monitoring framework for palm fruit ripeness assessment that integrates YOLO deep learning object detection with LiDAR 3D mapping in a multimodal feature fusion architecture. The methodology reported here are broadly applicable beyond oil palm to other tree crops where ripeness monitoring is critical, including date palm, coconut, and avocado, suggesting wide potential impact for UAV-based precision horticulture. This paper addresses that gap by presenting a comprehensive drone-based monitoring framework that combines LiDAR 3D mapping with YOLO-based deep learning for automated palm fruit ripeness classification.

References

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). 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. (2018). 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. (2016). 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). (2020). FFB Ripeness Grading Standard and Specification. MPOB Technical Document TT-001. Kuala Lumpur, Malaysia.

[9] Ampatzidis, Y., & Partel, V. (2019). 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.