PERFORMANCE ANALYSIS OF YOLO AND SSD-BASED DEEP LEARNING MODELS FOR DETECTION OF OIL PALM TREES IN DRONE IMAGES

ISTIYAK MUDASSIR SHAIKH

UNIVERSITI SAINS MALAYSIA

2025

PERFORMANCE ANALYSIS OF YOLO AND SSD-BASED DEEP LEARNING MODELS FOR DETECTION OF OIL PALM TREES IN DRONE IMAGES

by

ISTIYAK MUDASSIR SHAIKH

**Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science**

July 2025

ACKNOWLEDGEMENT

Alhamdulillah, all praises to Allah for His strength and blessings in completing this thesis. I extend my heartfelt gratitude to my supervisor, Dr. Mohammad Nishat Akhtar, for his unwavering support, guidance, and encouragement throughout my master's studies. This work would not have been possible without his invaluable support.

I am deeply grateful to the School of Aerospace Engineering, Universiti Sains Malaysia (USM), for providing excellent research facilities and to the Malaysian Government for supporting this work through grant number 304/PAERO/6315761.

I sincerely thank my beloved parents, Abdul Hussain and Gousiya Begum, for their unconditional love, prayers, and sacrifices, as well as my sister and brothers for their constant encouragement. Special appreciation goes to my senior researcher brothers, Dr. Abdul Aabid and Dr. Asrar Anjum, for their motivation and guidance, and to my sister-in-law, Nagma Parveen, for her support.

I also appreciate my colleagues, Dr. Junaid Mohammad Khan, Mohammad Uzair Gill, Zulfam Adnan, Mohd Hafiz Ab Satar, and Adeel Mohammad Khan, for their assistance and valuable discussions. A special thanks to Mr. Zihad Mahmood, the drone pilot, and Siva Aruth Sudhar, a final-year student, for their crucial help during data collection.

Lastly, I am grateful to all my friends who have supported me throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

GT	Ground Truth (Total number of instances in the dataset)
TP	True Positives (Correctly detected instances)
FP	False Positives (Incorrectly detected instances)
FN	False Negatives (Missed detections)
Precision (%)	Measure of correctly identified positive instances compared to total predicted positives
Recall (%)	Measure of correctly identified positive instances compared to total actual positives
F1-S	core (%) – Harmonic mean of precision and recall, representing the overall detection accuracy
D.T.	(sec) – Detection Time (Processing time per image or batch in seconds)
N_grayscale	Number of grayscale images in the dataset, (where N=total number of images)
IoU	Intersection over Union

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AUC	Area Under the Curve
BPNN	Backpropagation Neural Network
CNN	Convolutional Neural Networks
CNN+SVM	Convolutional Neural Networks combined with Support Vector Machine
D.T	Detection Time (in seconds)
F1	Score - F1-Score
FN	False Negatives
FP	False Positives
FPR	False Positive Rate
GIS	Geographic Information Systems
GPS	Global Positioning Systems
GT	Ground Truth
IoT	Internet of Things
IoU	Intersection over Union
LIDAR	Light Detection and Ranging
mAP	Mean Average Precision
mAP50	Mean Average Precision at Intersection over Union threshold 0.5
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
NDVI	Normalized Difference Vegetation Index
OA	Overall Accuracy
OBIA	Object-Based Image Analysis
P	Precision
PA	Precision Agriculture
PSNR	Peak Signal-to-Noise Ratio
R	Recall
R-CNN	Region-Based Convolutional Neural Network

ROC	Receiver Operating Characteristic
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SNR	Signal-to-Noise Ratio
SSD	Single Shot Multibox Detector
SVM	Support Vector Machine
SW26010	A processor model used in parallel processing
TN	True Negatives
TP	True Positives
TPR	True Positive Rate
UAV	Unmanned Aerial Vehicle
YOLO	You Only Look Once

LIST OF APPENDICES

ANALISIS PRESTASI MODEL DEEP LEARNING BERASASKAN YOLO DAN SSD UNTUK PENGESANAN POKOK KELAPA SAWIT DALAM IMEJ DRON

ABSTRAK

Kajian ini meneliti penggunaan model pembelajaran mendalam yang canggih bagi tujuan pengesanan dan pengiraan pokok kelapa sawit dalam bidang pertanian tepat, dengan menggunakan imej resolusi tinggi yang diperoleh melalui dron. Motivasi kajian ini berasal daripada kelemahan kaedah pemantauan manual yang lazimnya memakan masa, mudah terdedah kepada ralat, serta tidak efisien untuk ladang berskala besar. Memandangkan Malaysia merupakan antara pengeluar utama minyak sawit di peringkat global, sistem pengesanan automatik yang cekap amat diperlukan bagi menyokong pengurusan ladang yang mampan.

Cabaran utama adalah untuk mengenal pasti pokok kelapa sawit secara tepat dalam keadaan yang kompleks seperti kanopi bertindih, vegetasi yang padat, pencahayaan tidak seragam, serta kewujudan tumbuhan lain yang serupa. Faktor-faktor ini mengehadkan keberkesanan kaedah pemprosesan imej secara tradisional, justeru mendorong kepada penerokaan rangka kerja pembelajaran mendalam yang lebih mantap dan berupaya mengendalikan keadaan lapangan sebenar.

Empat model pengesanan objek termaju telah dinilai dalam kajian ini, iaitu YOLOv5x, YOLOv7, YOLOv8, dan SSDv2FPN. Model-model ini dipilih berdasarkan keupayaan pengesanan masa nyata serta kebolehan dan ketepatannya dalam persekitaran pertanian. Dua set data telah digunakan: satu set kecil terdiri daripada 10 imej dron dengan 79 pokok kelapa sawit yang telah dilabel, dan satu set data berskala besar yang mengandungi 482 imej dengan sejumlah 5,233 pokok.

Penilaian dibuat berdasarkan metrik seperti Positif Benar, Positif Palsu, Negatif Palsu, Ketepatan , Kadar Kepekaan, Skor F1, dan Masa Pengesanan. Model SSDv2FPN mencatatkan ketepatan sempurna iaitu 100% dengan Skor F1 sebanyak 89.49%, namun memerlukan masa 83 saat untuk memproses setiap imej, menjadikannya kurang sesuai untuk aplikasi masa nyata. Sebaliknya, model YOLOv5x, YOLOv7x, dan YOLOv8x berjaya mengesan pokok dalam masa yang lebih pantas iaitu masing-masing 16, 12, dan 14 saat, dengan YOLOv5x mencatatkan Skor F1 tertinggi iaitu 97.36%. Keputusan ini menunjukkan dengan jelas kelebihan dari segi kelajuan yang dimiliki oleh model-model YOLO

Bagi set data yang lebih besar, model-model YOLOv8 menunjukkan prestasi terbaik berbanding model lain, dengan pencapaian Skor F1 antara 97.36% hingga 99.31%, nilai ketepatan antara 99.27% hingga 99.70%, dan kadar kepekaan antara 95.89% hingga 99.36%. Dalam kalangan varian YOLOv8, model YOLOv8s dan YOLOv8n mencatatkan masa pengesanan terpantas iaitu masing-masing 28 dan 33 saat, sekali gus menawarkan keseimbangan antara kelajuan dan prestasi pengesanan yang tinggi. Justeru, model-model ini dianggap paling sesuai untuk aplikasi pemantauan pertanian secara praktikal.

PERFORMANCE ANALYSIS OF YOLO AND SSD-BASED DEEP LEARNING MODELS FOR DETECTION OF OIL PALM TREES IN DRONE IMAGES

ABSTRACT

This study explores the use of advanced deep learning models for detecting and counting oil palm plants in precision agriculture using drone-based high-resolution images. The motivation stems from the limitations of manual monitoring methods, which are time-consuming, error-prone, and not feasible for large-scale plantations. Given Malaysia's significant role in global palm oil production, efficient and automated detection systems are essential to support sustainable plantation management. The primary challenge is to accurately identifying oil palm trees in complex conditions, such as overlapping canopies, dense vegetation, varying lighting, and similar surrounding plants. These factors limit traditional image processing techniques, prompting the use of robust deep learning frameworks. This study evaluates four state-of-the-art object detection models: YOLOv5x, YOLOv7, YOLOv8, and SSDv2FPN, selected for their real-time detection capabilities and accuracy in agricultural environments. Two datasets were used: a smaller set of 10 drone images containing 79 annotated palm trees, and a larger dataset of 482 images with 5,233 trees. Evaluation metrics included True Positives, False Positives, False Negatives, Precision, Recall, F1-Score, and Detection Time. SSDv2FPN achieved perfect precision at 100% with an F1-Score of 89.49%, but required 83 seconds per image, which limits its suitability for real-time applications. In contrast, YOLOv5x, YOLOv7x, and YOLOv8x detected palm trees in relatively lower execution time of 16, 12, and 14 seconds respectively, with YOLOv5x achieving an F1-Score of 97.36%.

These results demonstrate the clear advantage of YOLO models with regard to high speed execution. On the larger dataset, YOLOv8 models outperformed other frameworks, thereby achieving F1-Scores between 97.36% and 99.31%, precision values ranging from 99.27% to 99.70%, and recall rates between 95.89% and 99.36%. Among the YOLOv8 variants, YOLOv8s and YOLOv8n demonstrated the fastest detection times of 28 and 33 seconds, respectively, effectively balancing rapid inference and detection performance. This makes them ideal for deployment in practical agricultural monitoring systems.

CHAPTER 1

INTRODUCTION

1.1 Significance of Palm Oil in Malaysia

The rising global demand for sustainable food sources, bio-based industrial products, and environmentally responsible agriculture, palm oil has emerged as a vital crop requiring efficient management and monitoring solutions. Its economic value, coupled with environmental implications, has made it a focal point for precision agriculture technologies and scientific research. The stemless monocot family, which includes palm oil plants, is an important component of tropical ecosystems and is well-known for its role in biodiversity conservation. These monocots are abundant in tropical places, particularly in Southeast Asia, Africa, and Latin America (Wilcove & Koh, 2010). Palm oil plants are highly appreciated for their function as primary producers of vegetable oils, having a significant impact on agricultural production around the world. The oil derived from these plants is used in a variety of food, cosmetics, and industrial items (I. Mukherjee & Sovacool, 2014). As a result, palm oil has risen to prominence in the global market.

Indonesia is the world's largest producer of palm oil, followed by Malaysia, Thailand, Nigeria, and numerous Latin American countries (Obidzinski et al., 2012). The expansion of palm oil plantations has had a tremendous impact not only on economic development but also on a variety of developmental domains. These contributions include better agricultural methods, poverty reduction, infrastructural development, and the rise of diverse enterprises (Gatto et al., 2015). While these advancements are significant, they also bring to light the challenges of sustaining such growth in the long run. However, the recent rapid increase of oil palm farming has prompted serious concerns regarding the long-term management of palm oil plants. Concerns over deforestation and habitat degradation, as well as labor and social issues, have heightened scrutiny of palm oil production (Union et al., 2018). In this regard, the European Union has introduced the EU Deforestation Regulation (EUDR) to prevent the import of commodities linked to deforestation, including palm oil. This is because deforestation, while it may offer short-term economic benefits through land clearing and increased agricultural output, contributes to biodiversity loss, greenhouse gas emissions, and long-term environmental degradation, ultimately undermining sustainable development and global climate goals**.** The EUDR, which is set to take effect in 2025, requires companies to ensure their supply chains are deforestation-free, presenting compliance challenges for Malaysian palm oil producers, particularly smallholders who contribute significantly to the nation's output. While initially viewed as discriminatory, the regulation has prompted Malaysia to enhance traceability and promote sustainable practices to align with these requirements (Reuters, 2024; SCMP, 2023).

As a result of these problems, it is critical to adopt accurate and fast monitoring systems to ease worries and assist informed decision-making. This is because earlier monitoring practices, though useful at a smaller scale, are no longer sufficient to manage large plantations efficiently, given the demand for real-time data, scalability, and precision. Traditional techniques of monitoring palm oil farms, such as tree counting and tree identification, have relied mainly on manual labor (Petri et al., 2022). While these traditional approaches are beneficial for smaller plantations, they are essentially insufficient for bigger, commercial-scale enterprises (Tang & Al Qahtani, 2020). They are biased, time-consuming, and frequently produce erroneous results. Furthermore, they necessitate large personnel, thereby limiting the frequency and extent of monitoring operations. Given these constraints, the use of modern technology such as remote sensing, drones, and machine learning has gained traction in the monitoring of palm oil farms (N. Khan et al., 2021). These tools not only improve operational efficiency but also offer strategic advantages to stakeholders by enabling timely decision-making, reducing manual workload, and supporting environmental compliance and sustainability certification. Recent studies published in the ISPRS Journal of Photogrammetry and Remote Sensing have begun exploring automated detection using UAVs and AI (J. Zheng et al., 2020), but few directly address scalable real-time object detection tailored to plantation environments. This research fills that gap by evaluating real-time deep learning models for palm detection under diverse field conditions, making it highly relevant to current agricultural monitoring needs.

These technologies make data collection and processing more accurate and efficient. They can give plantation managers and environmental authorities a real-time data into plantation health, tree density, and land use changes, thereby allowing them to make informed decisions about sustainable practices (Khuzaimah et al., 2022). Palm oil plantations are critical to both global agriculture production and tropical local economies (Ayompe et al., 2021). While they have provided enormous benefits, they have also generated questions about their long-term viability and environmental impact. To address these challenges and promote sustainable management practices in the palm oil business, accurate and timely monitoring systems are required (Ahmad et al., 2023). The shift from manual labor-intensive approaches to technology-driven solutions has the potential to increase the accuracy and efficiency of palm oil plantation monitoring dramatically.

1.2 Significance of Palm Tree Detection by UAV

The accurate detection of palm trees is fundamental for effective yield estimation, plantation planning, and early intervention strategies. As plantations scale up in size and complexity, traditional ground-based methods become inefficient, making aerial-based solutions essential for timely and comprehensive monitoring. The use of unmanned aerial vehicles, or drones, has emerged as a transformational instrument among these technical breakthroughs (Chowdhury et al., 2022). Furthermore, the unrivaled speed with which drones collect data enables constant and real-time monitoring operations, allowing for the early detection of deviations from the norm, such as disease outbreaks, insect infestations, and illicit activities (L. Wang et al., 2022). Drones equipped with advanced sensors, such as multispectral, hyperspectral, and thermal cameras, can capture high-resolution images and provide detailed insights into the health and spatial distribution of palm trees (Adão et al., 2017). These capabilities allow plantation managers to assess tree vitality, monitor stress levels, and identify potential threats, ensuring timely intervention. Additionally, UAVs can access hard-to-reach areas within plantations, overcoming physical barriers that would otherwise impede ground-based inspections (Ghazali et al., 2022).

The integration of UAVs with machine learning and computer vision algorithms has further enhanced their utility. Automated systems can process drone-acquired imagery to detect, classify, and count palm trees with high precision, reducing human error and labor costs (X. Liu et al., 2021). Furthermore, drones facilitate the creation of precise geospatial maps, enabling plantation managers to implement precision agriculture practices such as targeted irrigation, fertilization, and pest control (Puri et al., 2017). As a sustainable and cost-effective approach, UAV-based monitoring significantly contributes to improving plantation management efficiency. By optimizing resource utilization and minimizing environmental impact, this technology supports the broader goals of sustainable agriculture and food security (Reddy Maddikunta et al., 2021).

1.3 Computer Vision Techniques of Palm Tree Detection

Artificial intelligence (AI)-driven object detection frameworks, especially those built on deep learning, have become effective methods for tackling these issues in recent years. Traditional computer vision techniques, such as sliding window approaches and handcrafted feature extraction using algorithms like HOG (Histogram of Oriented Gradients) and SIFT (Scale-Invariant Feature Transform) (Dalal et al., 2005; Lowe, 2004; Ortiz Laguna et al., 2011) , were the mainstay of object detection prior to the development of sophisticated deep learning models. These methods were computationally demanding, involved a lot of human labor, and frequently had issues with scalability and real-time processing. When machine learning was introduced, handcrafted features were combined with techniques like Random Forests and Support Vector Machines (SVM) to improve the system (Jamie Shotton et al., 2008; Jin et al., 2020; Vapnik, 1999; Zhang Hao, Berg A. , Maire M., 2006). Although these methods increased accuracy, they were still constrained by their reliance on pre-established feature sets, which made it difficult for them to generalize effectively across a variety of datasets or intricate situations like occlusions and changing environmental conditions.

With Convolutional Neural Networks (CNNs) allowing models to automatically build hierarchical feature representations directly from data, the move towards deep learning marked a revolutionary step in object detection (Chauhan et al., 2018). By introducing region proposal networks and simplifying the detection procedure, frameworks such as R-CNN (Regions with CNN features) (Girshick et al., 2014), Fast R-CNN