Research Article

Enhancing sustainability in the production of palm oil: creative monitoring methods using YOLOv7 and YOLOv8 for effective plantation management

Istiyak Mudassir Shaikh [a], Mohammad Nishat Akhtar [a,*,1], Abdul Aabid [b,1], Omar Shabbir Ahmed [b]

[a] School of Aerospace Engineering, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia
[b] Department of Engineering Management, College of Engineering, Prince Sultan University, PO BOX 66833, Riyadh 11586, Saudi Arabia

A B S T R A C T

The You Only Look Once (YOLO) deep learning model iterations—YOLOv7–YOLOv8—were put through a rigorous evaluation process to see how well they could recognize oil palm plants. Precision, recall, F1-score, and detection time metrics are analyzed for a variety of configurations, including YOLOv7x, YOLOv7-W6, YOLOv7-D6, YOLOv8s, YOLOv8n, YOLOv8m, YOLOv8l, and YOLOv8x. YOLO label v1.2.1 was used to label a dataset of 80,486 images for training, and 482 drone-captured images, including 5,233 images of oil palms, were used for testing the models. The YOLOv8 series showed notable advancements; with 99.31 %, YOLOv8m obtained the greatest F1-score, signifying the highest detection accuracy. Furthermore, YOLOv8s showed a notable decrease in detection times, improving its suitability for comprehensive environmental surveys and in-the-moment monitoring. Precise identification of oil palm trees is beneficial for improved resource management and less environmental effect; this supports the use of these models in conjunction with drone and satellite imaging technologies for agricultural economic sustainability and optimal crop management.

1. Introduction

Oil Derived from stemless monocots, palm oil is essential for biodiversity conservation and plays a major role in tropical ecosystems. These monocots, which are widely distributed in South America, Africa, and Southeast Asia, play a significant role in the world's vegetable oil supply and have a significant impact on agricultural economies [1]. With Indonesia emerging as the world's top producer, followed by Malaysia, Thailand, Nigeria, and several Latin American countries, palm oil's exponential rise in significance on the global market has spurred not only economic advancement but also other developmental benefits like improvements in agricultural methodologies, significant progress in reducing poverty, and robust infrastructure development [2]. However, there are significant sustainability issues associated with the oil palm farms' rapid expansion. This sector's growth is being closely watched, as deforestation, environmental degradation, and socioeconomic tensions accompany it. This is raising questions about the sector's long-term economic viability and environmental sustainability [3], highlighting the need to maximize agricultural productivity while maintaining

environmental stewardship and fair economic benefits. Accurate palm tree counts, and monitoring has been one of the main areas of focus to address the issues with sustainable palm oil production. Accurate palm tree counting is necessary for evaluating the health of plantations, efficiently allocating resources, and ensuring compliance to environmental laws. Palm tree counting used to be a labor-intensive and error-prone procedure, but technological improvements have changed this.

Historically, the monitoring of palm oil fields has been dependent on labor-intensive methods like manual tree counting and identification, which work well for smaller plantations but not well enough for big commercial operations. The scope and frequency of monitoring operations are limited by these antiquated methods, which are not only biased and time-consuming but also error-prone [4]. As such, the use of contemporary technology represents a profound paradigm change in the way these issues are addressed. This study recommends that drones, machine learning, and remote sensing be used to monitor palm oil plants [5]. These cutting-edge technologies make it easier to collect and interpret data accurately and efficiently, giving real-time insights about plantation health, tree density, and changes in land usage. Plantation
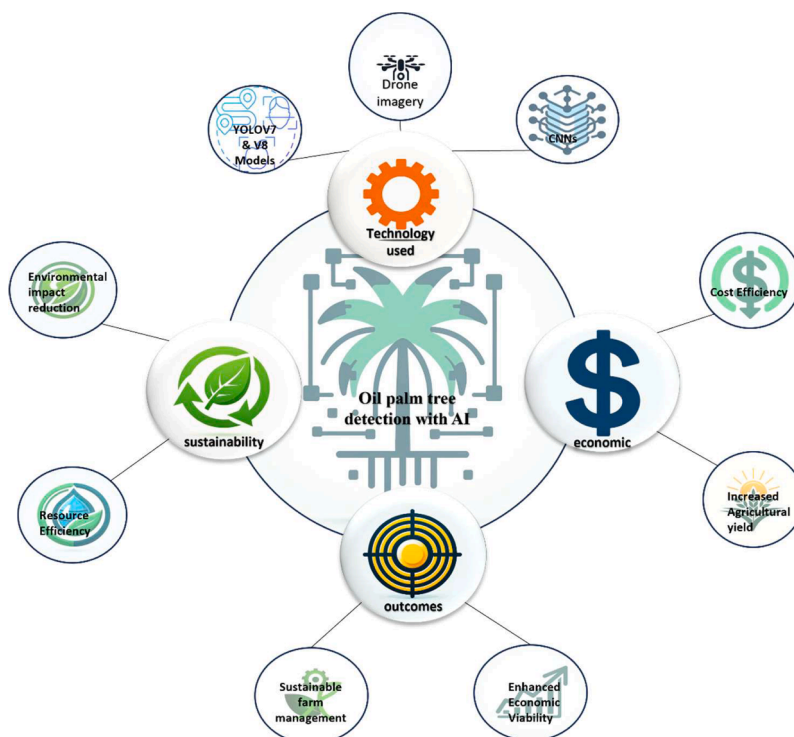

Fig. 1. Balancing Economic and Environmental Factors in Oil Palm Detection.

managers and environmental authorities can make well-informed decisions that improve sustainable practices and ensure environmental stewardship and economic efficiency with the use of such timely and accurate data [6].

Drones have emerged as a key instrument in the revolution of palm tree recognition within oil palm plantations thanks to artificial intelligence, computer vision, and machine learning [7]. Drones with sophisticated aerial photography equipment provide accurate measurements of tree density and canopy coverage. Their capacity to adapt to difficult terrains, such rocky or heavily forested areas, makes thorough crop monitoring possible. Drones' high-resolution photos are essential for precisely recognizing palm plants in the surrounding flora [8]. Drones are unique in that they gather data incredibly quickly, making real-time surveillance possible, which is crucial for the early identification of anomalies like disease outbreaks or unauthorized activity. Drones are essential for improving the detection of palm trees and play a major role in supporting sustainable agricultural practices by delivering accurate, timely, and efficient data. Because of their many advantages, they are vital resources for preserving the sustainability, long-term health, and environmental stewardship of palm oil fields [9]. Drone technology and remote sensing have become an effective pair for palm tree counts. Drones can produce high-resolution aerial imagery that makes it possible to map plantations in detail, which makes it possible to count and identify individual trees with amazing accuracy [10]. A growing number of machine learning algorithms are being utilized to automate the counting process by identifying palm trees from other vegetation by analyzing aerial images. These algorithms make use of deep learning models such as Convolutional Neural Networks (CNNs).

Research has demonstrated that precision agriculture—which is fueled by deep learning and advanced imaging techniques—can significantly boost output while reducing environmental damage [11,12]. Adopting these technologies has significant benefits from an economic standpoint as it can maximize resource utilization, reduce expenses, and boost profitability [13,14]. The YOLO algorithm and other AI technologies can be integrated into farming operations to increase both sustainability and financial returns. This technological progress in

agriculture supports sustainable practices and boosts economic efficiency [15]. The YOLO algorithm evaluates the entire image using Convolutional Neural Networks (CNNs), generating predictions for bounding box coordinates and class probabilities [16,17,18]. The YOLO family has evolved with variants such as YOLOv5 [19], YOLOX [20], PP-YOLO [21], PP-YOLOv2 [22], YOLOv6 (C. [23]), YOLOv7 [24], and YOLOv8 [25], each offering improvements in speed and accuracy. These modifications have broadened the YOLO algorithm's utility, embracing tasks such as tree counting, Fresh Fruit Bunches (FFB) counting, and even harvesting systems [26,27,28].

Deep learning research has improved oil palm tree identification and measurement methodologies. [29] used the sliding window approach, a CNN, and the LeNet architecture to achieve a stunning 96 % detection accuracy in distinguishing oil palm trees in high-resolution Quick Bird satellite pictures. Another study [30] provided a technique that included sliding window operations, the AlexNet architecture of the Deep Convolutional Neural Network (DCNN), and post-processing techniques, with an accuracy range of 92–97 %. [31] suggested a comprehensive system for data processing and storage that integrates the CNN method with the LeNet architecture and Geographic Information System (GIS) capabilities. Using high-resolution Worldview-3 satellite photos, they spotted juvenile and young oil palm plants with 95.11 % and 92.96 % accuracy, respectively. Similarly, [32] used Transfer Learning with the VGG-16 architecture and an SVM classifier to detect oil palm plants in UAV pictures, obtaining detection accuracy of 97–98 %. [33]developed the Faster R-CNN approach, which achieved outstanding accuracy rates of 97.06 %, 96.58 %, and 97.79 % for detecting and counting oil palm plants in three unique regions. The [34] demonstrated high accuracy in detecting oil palm trees, with F1-scores of 97.28 % for YOLOv3, 97.74 % for YOLOv4, and 94.94 % for YOLOv5m, demonstrating the utility of deep learning models in precision agriculture. There are still a few issues with oil palm tree detection even with improvements made in CNN- and previous YOLO-based techniques. High-quality input data and significant computational resources are needed for these techniques, which might not be easily accessible in remote locations. For best results, images with a high resolution are required, but they can be expensive and
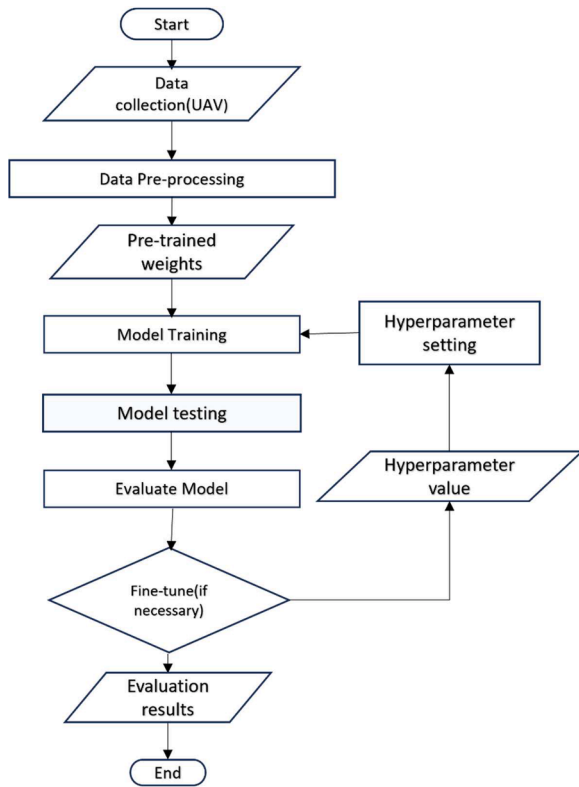
Fig. 2. Flow of the study.

don't always show environmental variation. Accurate detection may be impacted by weather and illumination sensitivity. Furthermore, smaller organizations may find it more difficult to use training models due to their complexity, which requires a significant investment of time and skill. The actual usefulness of these technologies is limited by the labor-intensive and costly nature of creating huge, datasets for training. Fig. 1 showcases how YOLOv7 and YOLOv8 can be utilized for AI-based detection and counting of palm oil trees, and how the results of their implementation influence such societal sectors as sustainability and economy. The main component is the utilization of these enhanced object detection algorithms to analyze drone-acquired images, for identifying and counting the palm oil trees in real-time. This is an addition to saving time, energy, and resources involved in completing the task manually should the technology not exist, the technology fosters sustainable working in that the time taken, and the efforts applied

towards the job are less than those required if done by hand. The impact goes economic where there is increased accuracy on the management and monitoring of the plantations hence added value of the increase in palm oil production through improved yields effectively making it more economically viable and sustainable.

With a focus on the various challenges posed by plantation conditions like sparse, densely populated, and overlapping oil palm trees as well as the presence of closely related vegetation, our research aims to improve the precision and accuracy of oil palm tree detection through high-resolution drone imagery. Our goal is to accurately identify oil palm trees in these challenging environments by utilizing the object recognition and training capabilities of sophisticated Convolutional Neural Networks (CNNs) on large, annotated datasets. To efficiently detect oil palm trees from other species, our technology combines multispectral and hyperspectral imaging techniques to exploit differential spectral fingerprints. With the goal of revolutionizing oil palm tree detection, this cutting-edge computational method combines the best aspects of machine learning, computer vision, and a variety of data sources. The goal is to overcome the difficulties associated with precisely counting oil palm trees in a variety of situations, moving away from labor-intensive, conventional techniques and toward a more effective, long-lasting, and economically viable monitoring system for the palm oil sector. This kind of research is essential for creating long-term, comprehensive strategies for managing palm oil resources and for meeting the urgent demand for precise, accurate, and sustainable monitoring systems.

The paper's structure is arranged as follows: The methodology is explained in Section 2; the results and discussions are presented in Sections 3 and 4, respectively; and the paper is concluded in Section 5.

2. Methodology

2.1. Flowchart of this study

Fig. 2. illustrates methods to detect oil palm trees using the YOLOv7 and YOLOv8 models. It starts with the model's initialization and then moves on to pre-process the input data. Following pre-processing, the data is loaded into YOLO models for inference. During the model testing phase, oil palm tree detection happens when the models use their trained algorithms to detect oil palm tree features in the input. If an oil palm tree is spotted, the model follows a specific branch of the procedure designed for positive detection; otherwise, it takes a different path for non-detection. The flow continues with post-processing stages and returns the output, which most likely comprises the location of the detected oil palm trees within the photos, as well as the detection's confidence score. This procedure illustrates a typical use of object
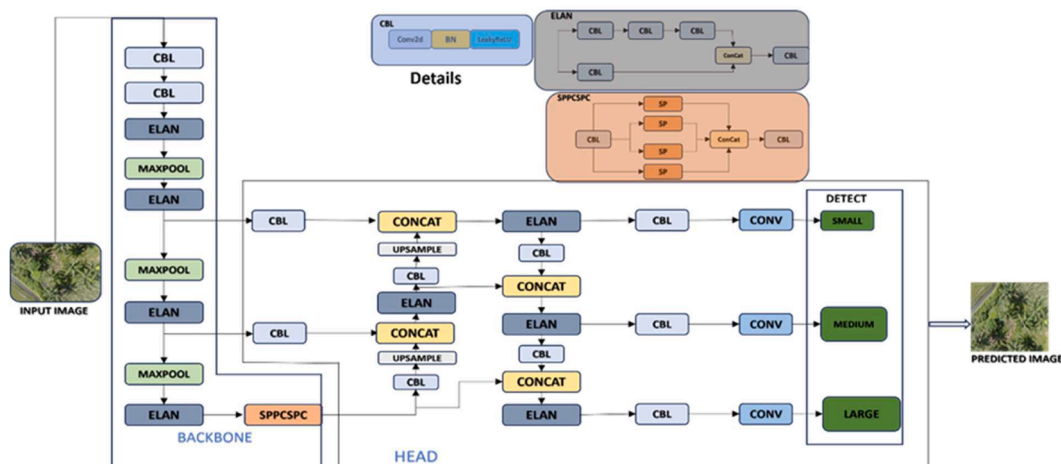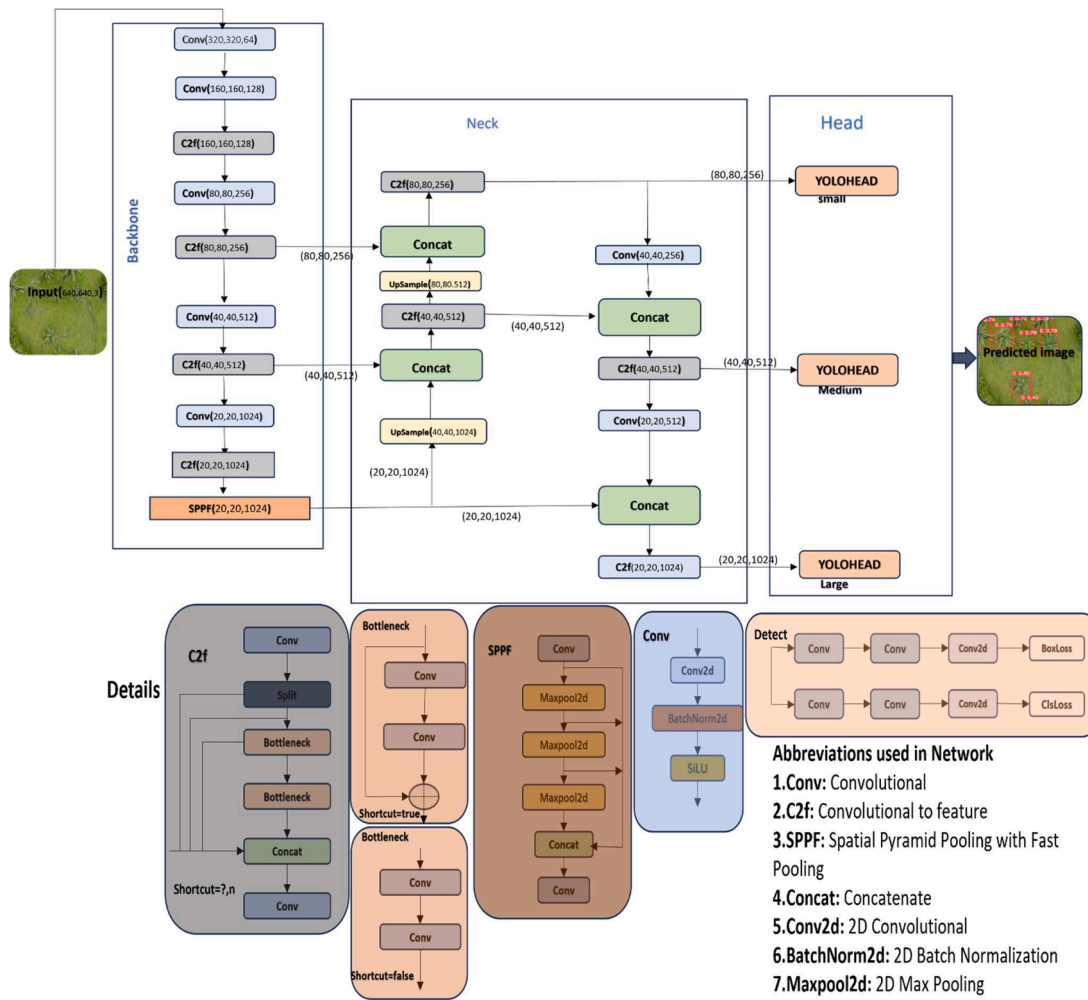

Fig. 3. YOLOv7 Architecture.

Fig. 4. YOLOv8 architecture.

identification models in machine learning, which are used to recognize specific items within visual data.

2.2. YOLOv7 architecture and improvements

In terms of object detection, YOLOv7 [24] is an evolution of the YOLO family, incorporating intricate structural innovations to improve detection precision and processing efficiency. The architecture, as shown in Fig. 3. begins with an image input, which is transformed within the neural network to detect and pinpoint objects.

At its core, the architecture is divided into three sections: the backbone, neck, and head. The backbone is built for feature extraction, with an arrangement of Convolutional Blocks (CBL), Efficient Local Attention Network(ELAN) modules for enriched feature extraction, and Max-Pooling layers to reduce spatial dimensionality. This backbone culminates in an improved variant of Spatial Pyramid Pooling with Channel and Spatial Pyramid Convolution(SPPCSPC), which incorporates multi-scale contextual information required for object recognition at various sizes.

The neck acts as an intermediate, improving features utilizing upsampling to recover spatial resolution and concatenation to combine semantically rich and geographically precise characteristics. This component of the network ensures reliable multi-scale detection, which is crucial for distinguishing objects across several dimensions.

The detection procedure takes place in the head, which is the final stage. It has three unique detectors for different scales: tiny, medium, and giant. Each is responsible for calculating bounding boxes, objectness

scores, and class probabilities, making it possible to identify numerous objects in a single inference cycle.

The network output is post-processed, including the use of confidence levels and Non-Maximum Suppression(NMS) to refine the predictions, resulting in a final image with correct bounding boxes and class labels. YOLOv7's design demonstrates the developments in object detection, presenting a system capable of speedy and precise object location and classification.

2.3. YOLOv8 architecture and improvements

YOLOv8 [25] is Ultralytics' latest YOLO model, known for object identification, image categorization, and instance segmentation. It builds on the success of YOLOv5 by making architectural changes and improving the developer experience. Ultralytics is actively developing and supporting YOLOv8, collaborating with the community to improve the model.

YOLOv8 is an enhanced object detecting system based on the original YOLO concept, which stands for "You Only Look Once." This novel methodology uses a grid style to identify and categorize things. It begins by scaling the input image to a standard size before splitting it into smaller portions, or grid cells as shown in Fig. 4. The YOLOv8 architecture combines several cutting-edge components to achieve great accuracy in real-time object recognition. The model begins with an input image of size (640,640,3) that is passed on the model's backbone. Multiple convolutional layers (Conv) and sophisticated C2f blocks make up the framework, which reduces the image's spatial dimensions and
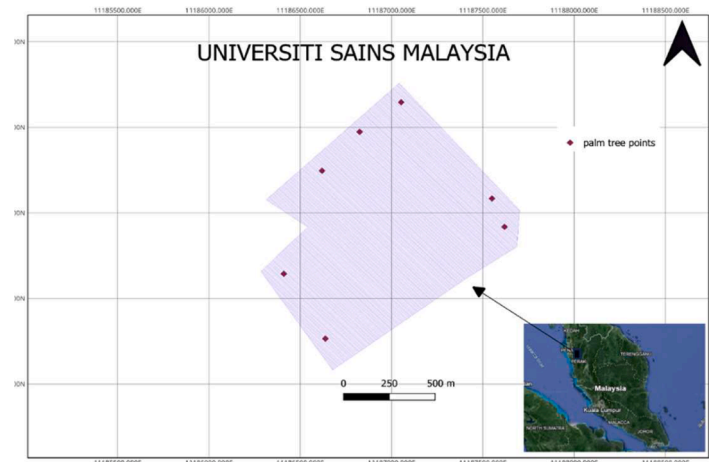
Fig. 5. Study area located at universiti sains Malaysia.

Table 1
Drone specifications [35] (https://dronespec.dronedesk.io/dji-mavic-air).

DJI Mavic Air

| Parameter	Value
Takeoff Weight	430 g
Maximum Takeoff Altitude	5000 m
Flight time	21 min
Camera sensor	1/2.3″ CMOS
Image size	4:3 4056×3040
Operating Frequency	16:9 4056×2280
Flight Battery capacity	2.400–2.4835 GHz
	5.725–5.850 GHz
	2375mAh

increases its depth while extracting information. An additional layer in the backbone, called Spatial Pyramid Pooling – Fast(SPPF), gathers global context data to improve the model's object detection performance at different scales.

Subsequently, the Neck of the model uses concatenation and upsampling techniques to merge information from multiple backbone layers, resulting in a multi-scale feature map that helps identify objects of varied sizes. The model's Head, comprising multiple YOLOHEAD layers that predict bounding boxes, class probabilities, and objectness scores at three distinct scales, receives these feature maps. This ensures that the model can detect tiny, medium, and big objects.

The C2f block, shown in the diagram, has shortcut connections and bottleneck layers that facilitate smoother gradient flow during training, hence aiding in effective feature extraction. The SPPF block enhances the overall performance of the model by focusing on different characteristics using max-pooling and concatenation. Bounding box coordinates and class predictions are the final outputs that the Detect layer generates after processing the features. The "Predicted Image," which highlights detected objects with bounding boxes and labels, is created using these outputs. All things considered, YOLOv8 incorporates updated architectural features to boost detection accuracy across a range of object scales while retaining the efficiency and speed of earlier YOLO models.

2.4. Study area

The research location in this study is an oil palm plantation located
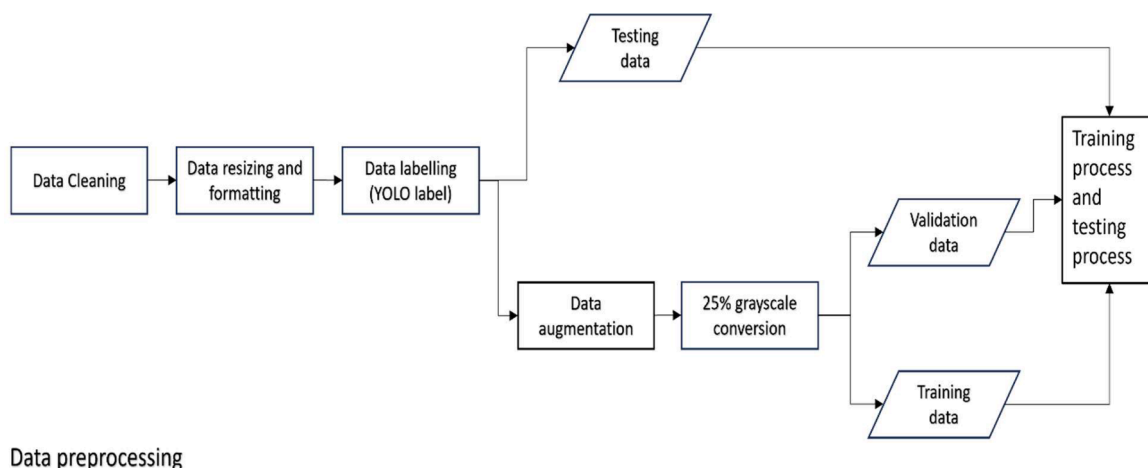
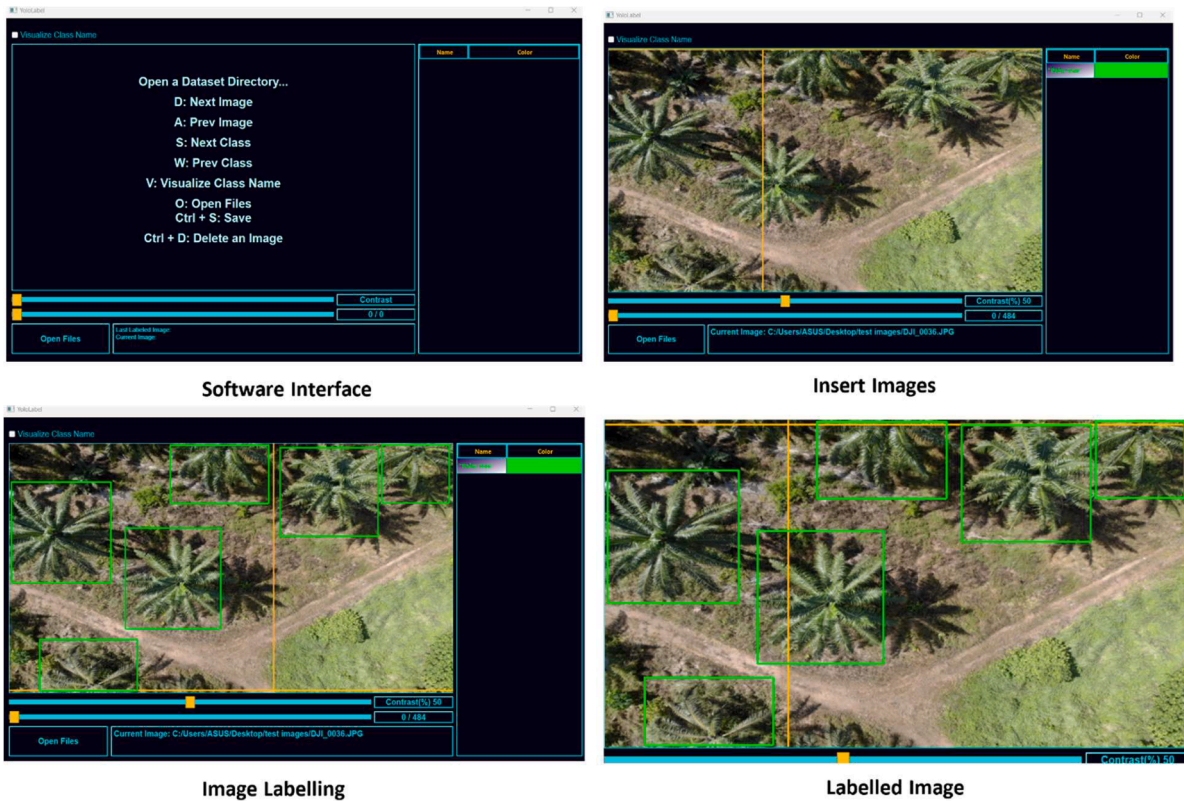**Fig. 6.** Flowchart for data preprocessing.

Fig. 7. Manual labelling by using YOLO label v1.2.1.

around university science Malaysia engineering campus nibong tebal, south seberang perai,penang Malaysia as shown in Fig. 5. The images were acquired in six different days in September and October 2023 using DJI Mavic Air at an altitude 36 m above sea level with a dimension of 4056×3040/image.

The drone specification can Table 1. Each of the sample dataset areas (training, validation, testing) includes young palm trees and mature palm trees in oil palm plantation with flat and hilly contours with sparse, dense, and overlapping canopy spacing conditions, as well as oil palm trees intersecting with other vegetations. The total images we have collected is about 1277.

2.5. Data preprocessing

In the YOLO, a preparatory stage of data preprocessing is required before beginning input data processing. This preliminary phase consists of a series of procedures as shown in Fig. 6. targeted at improving the quality and compatibility of the incoming data. The following are the primary preprocessing steps that have been implemented,

Data Cleaning: Cleaning data means that in the process of preparation of raw data for processing, various flaws, including inconsistencies, out of range values, and unrelated information, must be deleted. When it comes to oil palm tree detection resizing and formatting the images to 640×640 is very crucial especially to the models such as YOLO which are fixed in their input size. The 640×640 resolution was chosen because it is the optimal combination of the computational load of the algorithm, working memory, and detection performance. The composition of 416×416 is too low and may leave behind many features while 740×740 and 1080×1080 require much memory in addition to computation power hence the most suitable was 640×640 was effective enough in recognizing both small and large objects [36]. It confirms that the entire model has a fast capability to progress images and does not greatly burden GPU, which puts the model into practical use in real-time objective discovery. By means of this resolution, high accuracy can be preserved, but at the same time, the model remains relatively simple and applicable to different kinds of objects independently of their sizes.

Data Labelling: For training and validation, we manually tagged 80,486 oil palm trees with YOLO Label v1.2.1. This careful labeling ensures a broad set of occurrences for strong model training. YOLO Label software labels items in pictures or videos to train YOLO object identification algorithms. After downloading and installing the software, users can input photos or videos, annotate objects with bounding boxes and class labels, store annotations, export annotated data in YOLO, Pascal VOC, or COCO format, and use this data to train the YOLO model. The YOLO Label software interface, shown in Fig. 7 is used to annotate photographs and create datasets for object detection. The interface includes features such as keyboard shortcuts for faster picture and class navigation, as well as image contrast tools for improved visibility. It demonstrates the use of bounding boxes for labeling, which is required for accurate oil palm tree detection. The interface also has tools for saving and loading photos, emphasizing its efficiency and usability for dataset creation in computer vision research.

Data Augmentation: By merging drone and internet platform data, we enhanced our collection from 1277 to 9667 photos. To improve model robustness and adaptability to a variety of environmental situations, we use grayscale transformation as a key data augmentation strategy. Specifically, we used this strategy on 25 % of the photos in our training dataset. This purposeful augmentation method is intended to reduce the model's reliance on color information, resulting in a greater emphasis on textural and structural qualities for detection tasks. After applying data augmentation, the data increased to approximately 106,543 oil palm trees in that 78 % (84,128) are for training, 15 % (17,182) are for validation, and 5 % (5, 233) are for testing.

The grayscale conversion process follows the conventional luminance model [37], which is represented by the following luminosity function:

$$Gray = 0.299 \times Red + 0.587 \times Green + 0.114 \times Blue \tag{1}$$

Table 2
YOLO models hyperparameter setting for training and validation.

Model	Input size	Batch size	Decay	Momentum	Learning rate
YOLOv7-W6	640×640	12	0.005	0.937	0.1
YOLOv7-D6	640×640	12	0.005	0.937	0.1
YOLOv7x	640×640	12	0.005	0.937	0.1
YOLOv8s	640×640	20	0.001	0.937	0.01
YOLOv8m	640×640	20	0.001	0.937	0.01
YOLOv8n	640×640	20	0.001	0.937	0.01
YOLOv8l	640×640	20	0.001	0.937	0.01
YOLOv8x	640×640	20	0.001	0.937	0.01

This equation converts each pixel's RGB values into a single intensity value, imitating the human eye's varying sensitivity to different color wavelengths—most notably, a stronger sensitivity to green light.

For a dataset *D* comprising *N* images, the subset *Ngrayscale* subjected to this transformation can be qualified as:

$$Ngrayscale = 0.25 \text{ x } N \tag{2}$$

Where *Ngrayscale* represents the number of photos converted to grayscale. This quantitative framework is useful for diversifying training data, thereby reinforcing the model against probable color fluctuations during real-world deployment, and lowering processing needs due to the reduced complexity of single-channel images.

Grayscale augmentation is expected to result in a more versatile and computationally efficient model capable of retaining high accuracy across a wide range of conditions where color fidelity is not guaranteed [38].

2.6. YOLO models development

To balance computational economy and model performance, the model development uses the PyTorch framework [39] for YOLOv7 and YOLOv8 with an input size of 640×640. Batch size, image size, momentum, decay, and learning rate were all carefully calibrated. To avoid GPU interruptions, a 640×640 input size was chosen, and batch size was adjusted correspondingly. The default parameters for momentum, decay, and learning rate were used. To improve accuracy and shorten training periods, pre-trained weights for YOLO convolutional layers were used. The training method was further enhanced through hyperparameter adjustment. Table 2. contains detailed network input sizes and hyperparameter setups for YOLOv7, and YOLOv8 models.

2.7. Evaluation metrics

We used key metrics such as Recall, Precision, and F1-score [40] to evaluate model performance, as described in the Eqns. (3)-(5). The model's effectiveness in detecting oil palm trees is measured by Recall,

the accuracy of the model's predictions is measured by Precision, and the F1-score is the harmonic mean of Recall and Precision. In addition, we used detection time as a statistic to compare model efficiency during detection. Average IoU (Intersection over Union) was used to evaluate the accuracy of bounding box locations, offering a full assessment of detection precision.

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{3}$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{4}$$

$$\text{F1} - -\text{score} = \frac{(2 \times Recall \times Precision)}{(Recall + Precision)} \tag{5}$$

The term True Positive (TP) refers to situations in which the model properly classified items as oil palm plants. False Positive (FP) refers to instances in which things other than oil palm trees were wrongly identified as such. The term False Negative (FN) refers to occasions in which actual oil palm trees were missed or were not spotted by the model [41] Fig. 8.

$$\text{IoU} = \frac{Area \text{ } of \text{ } Overlap}{Area \text{ } of \text{ } Union} \tag{6}$$

Overall, this study's methodology, which makes use of the YOLOv7 and YOLOv8 models, describes a thorough procedure for locating and counting oil palm plants. It starts with a flowchart that offers a clear outline of the study procedure. Next, the architectural enhancements in YOLOv7 and YOLOv8, which raise detection efficiency and accuracy, are thoroughly examined. In addition to thorough data preprocessing, which includes cleaning, labeling, and augmentation, the research region is carefully chosen to guarantee ideal testing conditions and high-quality inputs for model training. The performance of the carefully constructed and refined YOLO models is assessed using reliable metrics including Intersection over Union (IoU) (Fig.8), precision, recall, and F1 score. This well-organized approach highlights the study's emphasis on obtaining precise, effective, and practically useful outcomes in oil palm tree research.

3. Results

3.1. Training results

We investigated the capabilities of two latest well-known YOLO models— YOLOv7, and YOLOv8—to determine their ability to detect oil palm trees across a wide range of image sizes. Each model received intensive training over 500 epochs, including a preliminary warm-up phase to ensure steady optimization. Using pre-trained models accelerated the training process, allowing for faster adaptation to the individual
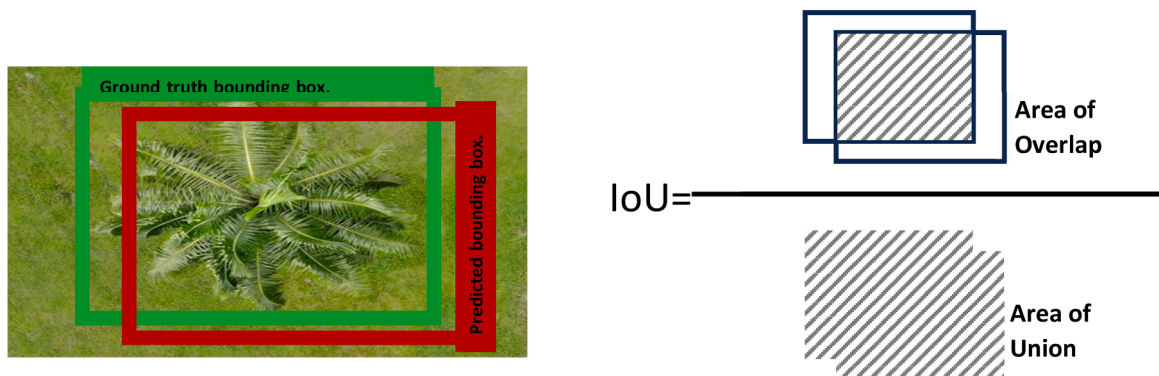
Fig. 8. Illustration of IoU(Intersection over Union).the green ground truth box from the manual labelling training data shows where our object is in the image, while the red predicted bounding box is from the trained model.

Table 3

The best YOLO model is based on the training and validation datasets.

Model	Peak Accuracy Epoch	Training Hours(hr)	Threshold	Precision(P)	Recall(R)	F1-Score	mAP50
YOLOv7x	654	65.12hr	0.2	0.95	0.94	0.94	0.97
YOLOv7-W6	675	60.65hr	0.2	0.93	0.938	0.92	0.97
YOLOv7-D6	600	93.29hr	0.2	0.95	0.94	0.94	0.98
YOLOv8s	137	5.799hr	0.7	0.94	0.95	0.94	0.98
YOLOv8n	407	11.71hr	0.7	0.946	0.953	0.94	0.986
YOLOv8m	247	20.127hr	0.7	0.949	0.952	0.95	0.986
YOLOv8l	186	21.345hr	0.7	0.945	0.956	0.95	0.986
YOLOv8x	220	44.239hr	0.7	0.942	0.957	0.94	0.984

Table 4

YOLO models' evaluation and comparison.

Model	GT	TP	FP	FN	Precision (%)	Recall (%)	F1-Score (%)	Detection time (in *sec*)
YOLOv7x	5233	4284	56	949	98.70	81.86	89.49	41.95
YOLOv7-W6	5233	3775	52	1458	98.64	72.13	83.32	41.80
YOLOv7-D6	5233	4259	54	974	98.74	81.38	89.22	48.66
YOLOv8s	5233	5197	38	36	99.27	99.31	99.28	28
YOLOv8n	5233	5018	35	215	99.30	95.89	97.56	33
YOLOv8m	5233	5190	28	43	99.46	99.17	99.31	40
YOLOv8l	5233	5200	33	33	99.36	99.36	97.36	43
YOLOv8x	5233	5093	15	140	99.70	97.32	98.49	51
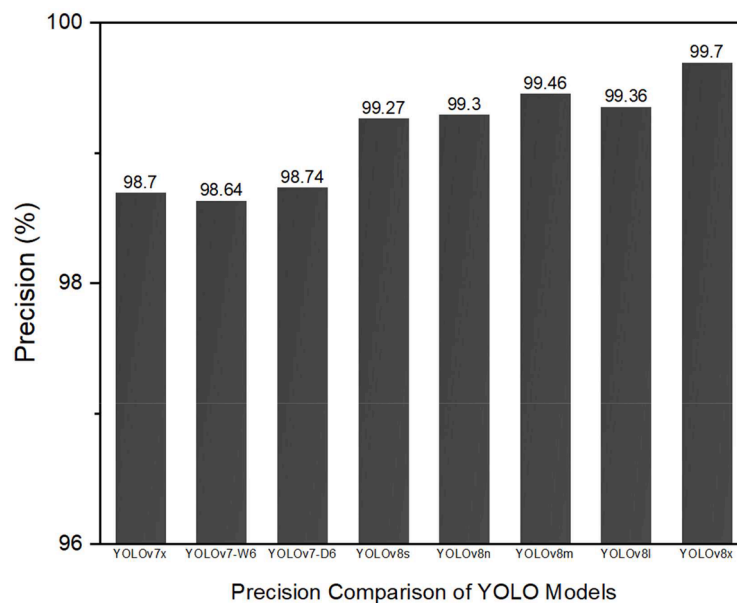
Fig. 9. Precision comparison of YOLOv7x, YOLOv7-W6, YOLOv7-D6, YOLOv8s, YOLOv8n, YOLOv8m, YOLOv8l, YOLOv8x.

task needs. Following each training iteration, rigorous validation was carried out, using uniform evaluation criteria across all models and image size scenarios. By examining accuracy, recall, F1-score, and mean Average accuracy (mAP) metrics [42] we gained useful insights into each model's comparative performance under various scenarios. This detailed analysis informed our understanding of the models' strengths and flaws. The results of the training evaluation for each epoch YOLOv7, and YOlOv8 models can be seen Microsoft excel sheet. We used Google Colab premium for fast training and evaluation of models, we used python3 programming language with hardware accelerator T4 GPU. The best models obtained from each epoch YOLOv7, and YOlOv8 are shown in Table 3. respectively, these models will be used for testing the test data.

3.2. Testing results

Testing data was acquired utilizing a UAV that flew 36 m above the ground and captured high-resolution photos of 4056×3040 pixels. Out of the 497 pictures gathered, 482 had oil palm trees, which were identified for proper tree counts. The ground truth count of oil palm trees in these pictures was 5233. Before testing, a confidence threshold of 0.7 was selected for each model, and the image size was standardized to 640×640 pixels. We tested YOLOv7x, YOLOv7-W6, YOLOv7-D6, YOLOv8s, YOLOv8n, YOLOv8m, YOLOv8l, and YOLOv8x for their ability to detect oil palm trees in the testing data. The Table 4. shows the performance metrics for these models, which include True Positives (TP), False Positives (FP), False Negatives (FN), Precision, Recall, F1-Score, and Detection Time. Models such as YOLOv8s and YOLOv8m demonstrated remarkable precision and recall rates, resulting in F1-
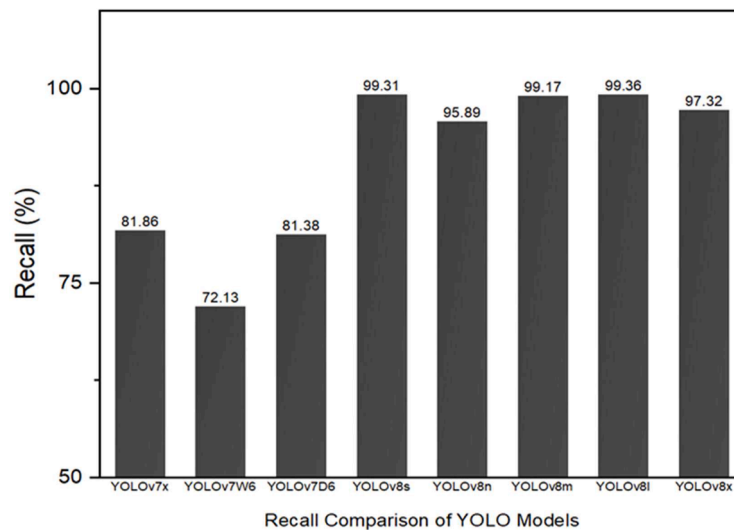
Fig. 10. Recall comparison of YOLOv7x, YOLOv7-W6, YOLOv7-D6, YOLOv8s, YOLOv8n, YOLOv8m, YOLOv8l, YOLOv8x.
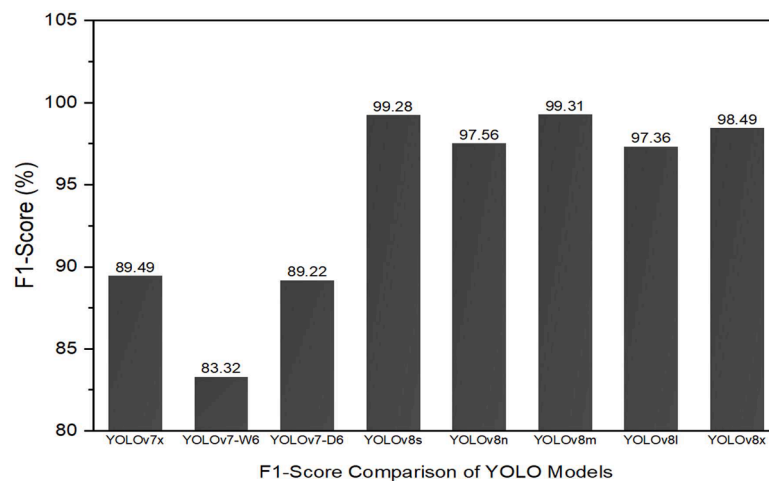

Fig. 11. F1-Score comparison of YOLOv7x, YOLOv7-W6, YOLOv7-D6, YOLOv8s, YOLOv8n, YOLOv8m, YOLOv8l, YOLOv8x.

scores of >99 %. YOLOv8s also had the fastest detection time of 28 s, demonstrating its effectiveness in oil palm tree recognition tasks.

The Fig. 9. that all models have great precision, with percentages starting at 98.7 % for YOLOv7x and trending upward with each subsequent iteration. Notably, YOLOv7-W6 has a significantly lower precision (98.64% vs. YOLOv7x). This could suggest a modest variation in the model's capacity to correctly detect oil palms tree without generating false positives.

As we advance to the YOLOv8 models, precision improves significantly. YOLOv8s is currently at 99.27 %, while YOLOv8n has increased to 99.3 %. The trend continues to improve, with YOLOv8m obtaining 99.46 % precision and YOLOv8l reaching 99.36 %. These enhancements propose changes to model architectures or training techniques that result in more accurate oil palm tree detection.

The maximum precision is seen with YOLOv8x at 99.7 %, indicating that the most recent iteration of the model series is the most accurate in predicting correct oil palm tree in the dataset used for this comparison. The YOLO series' persistent high performance across iterations indicates that it still has a strong competence in oil palm tree identification, with iterative enhancements contributing to incremental advances in precision.

Starting with YOLOv7x as shown in Fig. 10. The recall is 81.86 %, which is high but not flawless. This signifies that the model is missing

some oil palm trees. The next version, YOLOv7-W6, has a lower recall of 72.13 %, indicating that it misses more oil palm trees than the YOLOv7x, YOLOv7-D6 is showing higher recall of about 81.38 % better than YOLOv7-W6 model.

Moving on to the YOLOv8 series, there is a considerable improvement. YOLOv8s has a recall of 99.31 %, whereas YOLOv8n has lower recall to 95.89 %. The trend continues, with YOLOv8m at 99.17 % and YOLOv8l at 99.36 %, demonstrating that these models are extremely effective at detecting practically all oil palm trees.

The final one, YOLOv8x, had a tiny dip in recall to 97.32 %. It remains high, indicating that the model recognizes most oil palm trees, but it is lower than in prior YOLOv8 models.

The Fig. 11. compares F1-Scores from several YOLO (You Only Look Once) models. The F1-Score is a score that indicates how well a model recognizes oil palm trees in images while balancing accuracy and reliability. The first bar for YOLOv7x is at 89.49 %, which is quite good. However, YOLOv7-W6′s score lowers slightly to 83.32 %, indicating that it is less effective at detecting oil palm trees. Following this, all the other bars reflect various YOLOv8 models, and their scores are higher, which is phenomenal. They range from 97.56 % to 99.31 %, indicating that these models are highly good at identifying the correct oil palm trees. The highest score for YOLOv8m is 99.31 %, which is nearly perfect. The last bar, YOLOv8x, has a comparatively lower score of 98.49 %, but
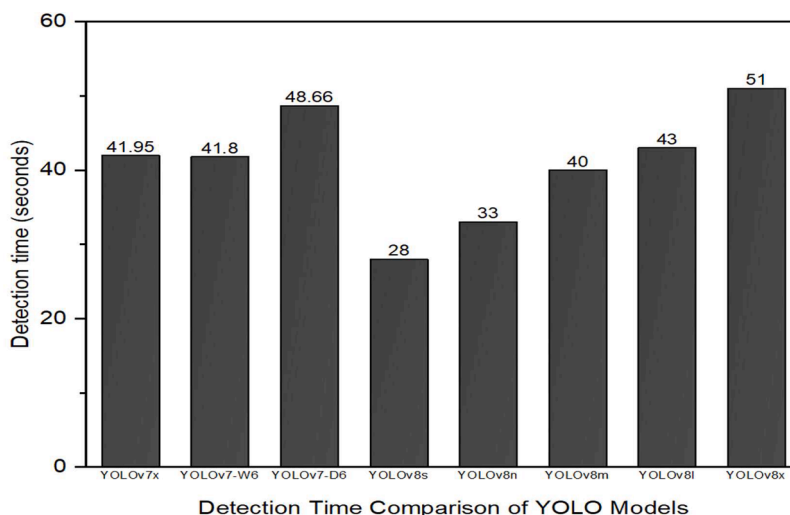
Fig. 12. Detection time comparison of YOLOv7x, YOLOv7-W6, YOLOv7-D6, YOLOv8s, YOLOv8n, YOLOv8m, YOLOv8l, YOLOv8x.

Table 5
Benefits and Drawbacks of each YOLO Model in Oil Palm Tree Detection.

Model	Benefits	Drawbacks
YOLOv7-W6	High precision minimizes on false positives, moderate performance is useful under different conditions of detection.	Lower recall may serve to missed some of the oil palm trees which have influence on the degree of completion of the detection.
YOLOv7-D6	Having the capability to detect in less time is useful for real-time detections and or if many targets are to be detected.	Lower precision and recall can lead to better or worse of missed detections/ false positives, thereby probably decreasing total reliability.
YOLOv7x	Higher recall is beneficial in identifying more trees of oil palm that may have been unnoticed at initial instance.	Obvious drawbacks may be identified in the case of increased detection time when a fast operation is necessary.
YOLOv8s	An improvement in precision as well as in recall rates and the significant speed of the first detection gives the best throughput.	Not mentioned, which is appropriate, as this is the perfect place for fast and accurate work.
YOLOv8n	Hence in terms of the performance characteristic, a high precision and relatively good recall make it appropriate for accurate search with reasonable computational time.	A slight decrease in recall might possibly mean missing out on certain trees—otherwise adequate.
YOLOv8m	The highest F1-Score means very high accuracy in both Precision and Recall, which is so powerful in the identification of oil palm trees.	No specific feedforward, making it exceptionally trustworthy for precise and specific identification.
YOLOv8l	Great accuracy and relevancy work well in various setting and are not overburdened by complexity of the implemented search algorithm.	None have been mentioned but it responds with a rather favorable overall performance.
YOLOv8x	The highest level is absolute precision and in general case is best for minimizing false positive.	The longest detection time can be justified with only slightly lower recall and F1-Score, which could be acceptable if top priority is given on speed and slightly lower recall rate.

has a modest detection time, which is slightly improved in the YOLOv7-W6 version as shown in Fig. 12. Contrary to predictions, the YOLOv7-D6 model has a significant increase in detection time, requiring roughly 49 s. When we switch to the YOLOv8 series, we see a significant improvement in YOLOv8s, which exceeds its predecessors in a speedy 28 s. However, this speed improvement is not retained in later YOLOv8 versions, with each succeeding model showing a continuous increase in detection time, culminating with YOLOv8x, which takes the longest at 51 s. This increasing trend in oil palm detection durations from YOLOv8s to YOLOv8x could be the result of a design decision to improve accuracy or incorporate more extensive analysis at the expense of speed.

Table 5. helps in selecting the best YOLO model for oil palm tree detection depending on specific needs, including recall, speed, and precision.

A thorough examination of the YOLOv7 and YOLOv8 models for oil palm tree detection is provided in the results section. The models' capacity to effectively generalize with high precision and recall is demonstrated by the testing results, which also demonstrate considerable increases in accuracy and loss reduction over time. Table 5. contrasts the benefits and drawbacks of each YOLO model, emphasizing the benefits and drawbacks of each in the current instance. All things considered, the findings show that YOLO models are effective at precisely identifying oil palm trees, offering insightful information about both their usefulness and possible regions for development.

4. Discussion

In our complete evaluation of the YOLOv7 and YOLOv8 series for precise detection of oil palm trees, the chosen models—YOLOv7x, YOLOv7D6, YOLOv8s, YOLOv8l, and YOLOv8x—showed remarkable accuracy over a wide range of environments. This rigorous testing method, meant to represent real-world applications, included both sparsely and densely populated areas of oil palm trees. The models' robustness was especially noticeable in sparsely populated areas, where each algorithm correctly identified individual trees with high confidence. Fig. 13. and Fig. 14. exhibit the models' performance under sparse and dense conditions respectively. In the following, where there were 58 ground truth oil palm trees, different models performed to varying degrees. YOLOv7x, YOLOv8l, and YOLOv8x accurately recognized 56 oil palm trees, excluding two that were stressed or unhealthy during the labelling process and hence did not qualify as oil palms. Notably, lowering the confidence threshold to 0.2 may allow these oil palm trees to be considered, but it also increases the number of false positives. In contrast, YOLOv7D6 and YOLOv8s detected just 54 and 48
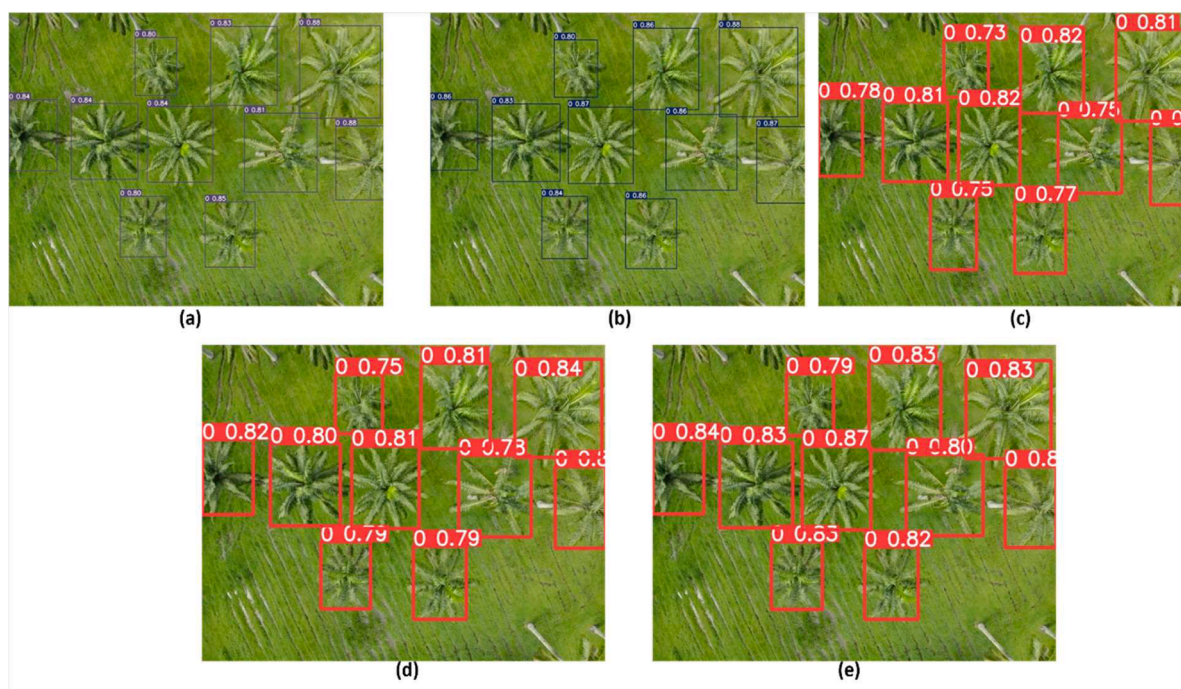
remains high. Overall, the graph shows that the YOLOv8 models are typically better at identifying oil palm trees in images than the prior YOLOv7 models.

After studying the detection timings over several iterations of the YOLO object detection models, there are speed variations. The YOLOv7x

Fig. 13. The sparse condition of oil palm trees: (a) YOLOv7x detection (b) YOLOv7D6 detection (c) YOLOv8s detection (d)YOLOv8l detection (e) YOLOv8x detection.
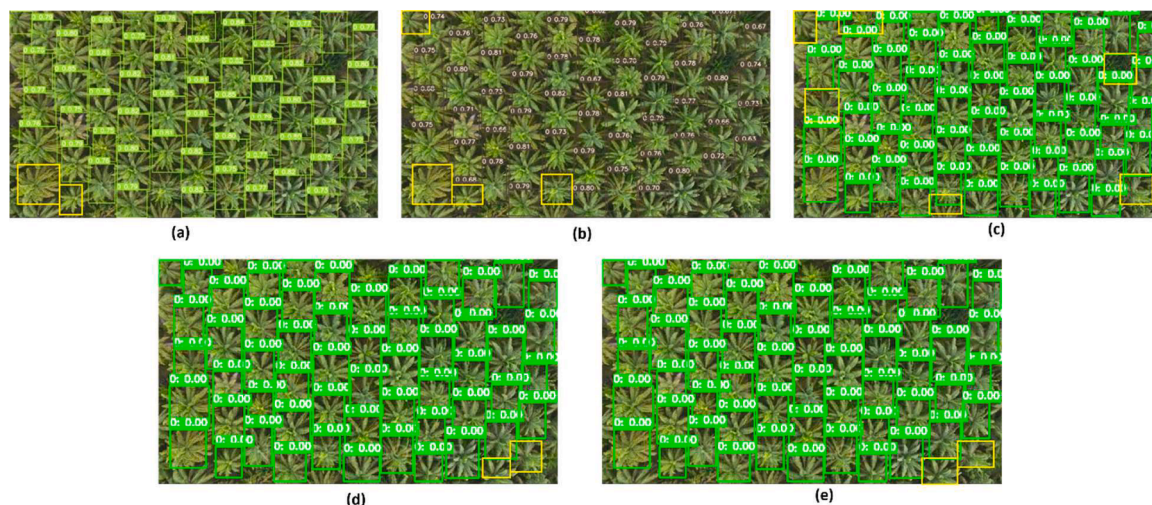

Fig. 14. The densely populated condition of oil palm trees: (a) YOLOv7x detection (b) YOLOv7D6 detection (c) YOLOv8s detection (d)YOLOv8l detection (e) YOLOv8x detection.

oil palm trees, respectively, showing a challenge for YOLOv8s in densely populated circumstances, although YOLOv7D6 performed marginally better.

In cases with tree overlap, as shown in Fig. 15. all models functioned brilliantly, except for YOLOv7x and YOLOv8s, which failed to detect a single tree due to excessive occlusion. This result highlights small issues that arise in highly overlapping circumstances, where even advanced models may struggle.

The Fig. 16. effectively shows the efficiency YOLO models in differentiating oil palm trees from coconuts. This is particularly difficult due to the morphological similarities between the two tree types. Each model's ability to correctly detect oil palms is significant, avoiding misdetection of coconut palms as oil palm tree. The confidence values that follow each detection demonstrate the models' high level of

precision. This accuracy is critical in precision agriculture, where small differences can have a significant influence on environmental monitoring and crop management.

We evaluate YOLOv7 and YOLOv8 models for oil palm tree detection and show that YOLOv8s and YOLOv8m perform better in managing areas that are both highly and sparsely inhabited, proving their efficacy in a variety of scenarios. YOLOv7 models exhibit good precision but slightly worse performance in dense canopies. However, in scenarios where trees overlap, they perform quite well, demonstrating sophisticated detecting abilities. Effective crop management and sustainability are greatly aided by the exceptional accuracy of YOLOv8 models, particularly YOLOv8l and YOLOv8x, in differentiating oil palms from related plants. This accuracy also helps to minimize false positives. This improved detection capability supports sustainable farming practices
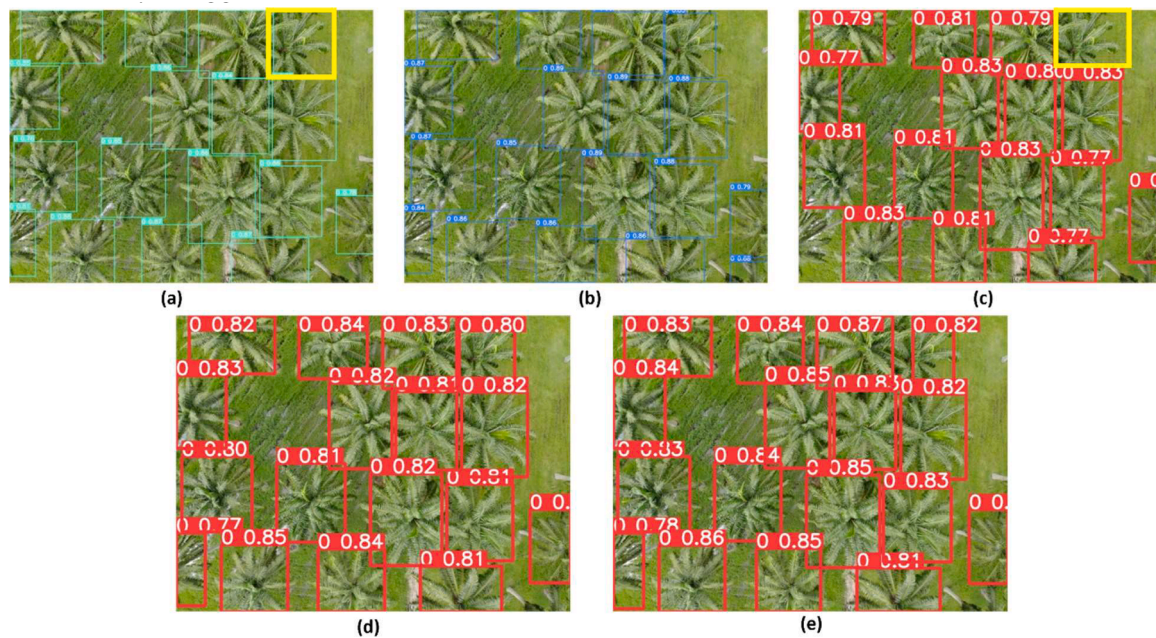
Fig. 15. The overlapping condition of oil palm trees: (a) YOLOv7x detection (b) YOLOv7D6 detection (c) YOLOv8s detection (d)YOLOv8l detection (e) YOLOv8x detection.
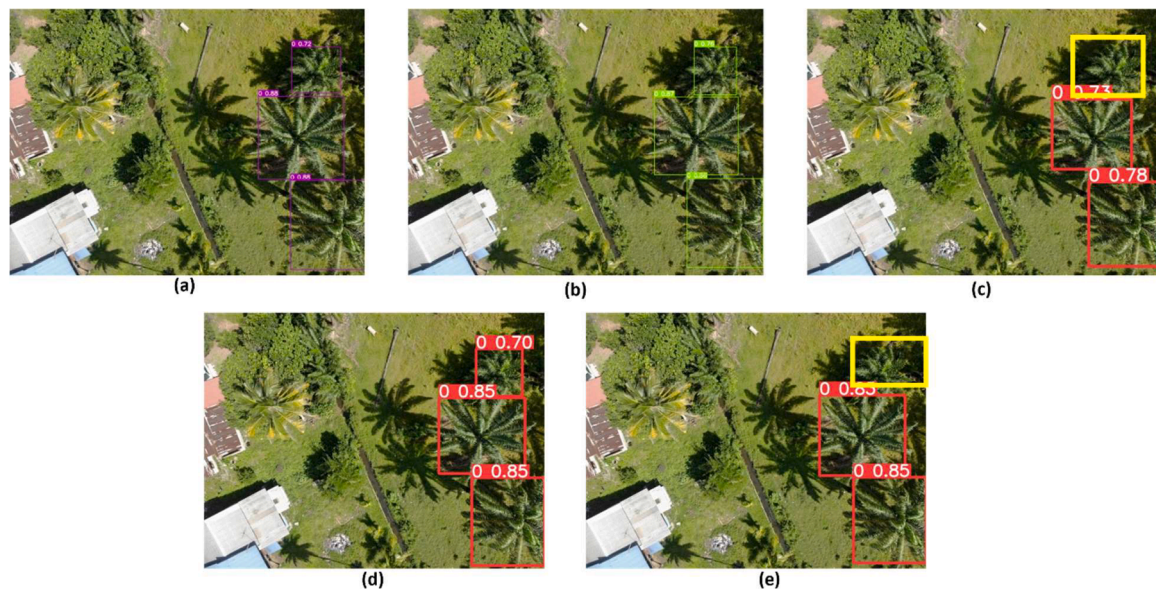

Fig. 16. Oil palm trees with other closely related vegetation (coconut tree): (a) YOLOv7x detection (b) YOLOv7D6 detection (c) YOLOv8s detection (d)YOLOv8l detection (e) YOLOv8x detection.

under a range of environmental conditions and not only increases agricultural productivity by guaranteeing precise interventions, but it also fosters economic sustainability by lowering operating costs and minimizing environmental impact.

Using both machine learning and deep learning techniques, our study provides a thorough comparison of our results with current benchmarks in the field of oil palm tree detection and counting. Our ability to identify oil palm trees with greater accuracy and precision now enables quicker processing times and more accurate counting. Our results have improved, and the number of false detections has decreased by establishing a confidence threshold of 0.7. Especially, our approach performs exceptionally well, with an average mean Average Precision (mAP) of 90 % or more at the Intersection over Union (IoU) threshold of 0.5 (mAP50). The capacity to reliably and effectively identify oil palm

trees on a broad scale is critical, highlighting its vital significance in promoting sustainable farming methods. This accuracy is essential for strategic resource allocation, yield optimization, and environmental effect assessment. Our study helps the agricultural industry make the shift to more productive and sustainable practices by increasing the accuracy and efficiency of detection, which improves large-scale monitoring and management of oil palm operations. Our developments ensure that the agricultural sector can effectively and responsibly fulfill future demands by raising the bar for current standards and making a substantial contribution to the larger objectives of environmental stewardship and sustainable agriculture.

5. Conclusions

We have used the most recent versions of the YOLO family—YOLOv7 and YOLOv8 in particular—to improve the recognition and counting of oil palm trees in our study. A variety of models, such as YOLOv7x, YOLOv7-W6, YOLOv7-D6, and all YOLOv8 variations, were painstakingly trained and evaluated using a dataset that included 482 images containing 5233 palm trees. Our models' precision varied impressively between 98.64 % and 99.7 %, demonstrating a consistently low false positive rate. The YOLOv7 variants exhibited strong recall rates, ranging from 72.13 % to 81.86 %. Meanwhile, the YOLOv8 models displayed even better performance, with recalls ranging from 95.89 % to 99.36 %, suggesting their efficacy in a variety of planting densities. By examining the F1-Score—a measure of precision and recall—YOLOv7 models scored between 83.32 % and 89.49 %, while YOLOv8 variations fared better, scoring between 97.36 % and 99.31 % and having detection times between 28 and 51 s. YOLOv7 and YOLOv8 versions, such YOLOv8s and YOLOv8n, are suggested for sparsely inhabited areas because of their accuracy and quick detection speed; YOLOv8s stands out for its quick detection time of only 28 s. On the other hand, YOLOv8l and YOLOv8x worked best in heavily populated areas, providing unmatched accuracy and precision.

These discoveries allow for more precise and effective monitoring and management of oil palm fields, which not only improves present approaches but also makes a substantial contribution to sustainable agricultural practices. This is essential for strategic resource planning, yield optimization, and environmental impact assessment. Our technologies also help the agriculture industry move toward more productive and sustainable operations, guaranteeing that oil palm tree detection is in line with the objectives of economic sustainability and environmental stewardship. Future work should concentrate on strengthening the models' resistance to occlusion and changing lighting, which are common problems in large-scale agricultural surveillance. Furthermore, early warnings for disease or insect outbreaks might be provided by integrating AI-driven anomaly detection systems, which would increase the productivity and sustainability of oil palm farming.

However, it's critical to recognize some of the YOLO models' shortcomings as applied to this research. Due to the sensitivity of the models to occlusions, palm trees that are completely or partially hidden by other trees, shadows, or vegetation may be missed or misclassified. Due to the possibility of detecting mistakes caused by nearby trees, this is especially difficult in places with dense plantings. Furthermore, inconsistent detection accuracy can result from the model's performance being affected by variations in sunlight, shadows, or reflections. For palm tree detection to be more reliable in a variety of dynamic agricultural situations, these constraints must be addressed.

CRediT authorship contribution statement

Istiyak Mudassir Shaikh: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Mohammad Nishat Akhtar:** Writing – review & editing, Supervision, Project administration, Formal analysis, Conceptualization. **Abdul Aabid:** Writing – review & editing, Project administration, Funding acquisition. **Omar Shabbir Ahmed:** Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Comte, Irina, François Colin, Joann K. Whalen, Olivier Grünberger, and Jean Pierre Caliman. 2012. "Agricultural practices in oil palm plantations and their impact on hydrological changes, nutrient fluxes and water quality in indonesia. a review." In *Advances in Agronomy*, 71–124. https://doi.org/10.1016/B978-0-12-394277-7.00003-8.

[2] Truly Santika, Kerrie A. Wilson, Sugeng Budiharta, Elizabeth A. Law, Tun Min Poh, Marc Ancrenaz, Matthew J. Struebig, Erik Meijaard, Does oil palm agriculture help alleviate poverty? a multidimensional counterfactual assessment of oil palm development in indonesia, World Dev. 120 (2019) 105–117, https://doi.org/10.1016/j.worlddev.2019.04.012.

[3] Mohammad Nishat Akhtar, Emaad Ansari, Syed Sahal Nazli Alhady, Elmi Abu Bakar, Leveraging on advanced remote sensing- and artificial intelligence-based technologies to manage palm oil plantation for current global scenario: a review, Agriculture (Switzerland) 13 (2) (2023), https://doi.org/10.3390/agriculture13020504.

[4] Erik Meijaard, Thomas M. Brooks, Kimberly M. Carlson, Eleanor M. Slade, John Garcia-Ulloa, David L.A. Gaveau, Janice Ser Huay Lee, et al., The Environmental Impacts of Palm Oil in Context, Nat. Plants 6 (12) (2020) 1418–1426, https://doi.org/10.1038/s41477-020-00813-w.

[5] Karolayne Teixeira, Geovane Miguel, Hugerles S. Silva, Francisco Madeiro, A survey on applications of unmanned aerial vehicles using machine learning, IEEE Access 11 (2023) 117582–117621, https://doi.org/10.1109/ACCESS.2023.3326101.

[6] Søren Marcus, Pedersen Kim, Lind Martin, Progress (2017). http://www.springer.com/series/13782.

[7] Li, Jun, Yanqiu Pei, Shaohua Zhao, Rulin Xiao, Xiao Sang, and Chengye Zhang. 2020. "A review of remote sensing for environmental monitoring in china." *Remote Sens.* 12(7). https://doi.org/10.3390/rs12071130.

[8] Junichi Kurihara, Voon Chet Koo, Cheaw Wen Guey, Yang Ping Lee, Haryati Abidin, Early detection of basal stem rot disease in oil palm tree using unmanned aerial vehicle-based hyperspectral imaging, Remote Sens. 14 (3) (2022), https://doi.org/10.3390/rs14030799.

[9] Siti Nurul Afiah Mohd Johari, Siti Khairunniza-Bejo, Abdul Rashid Mohamed Shariff, Nur Azuan Husin, Mohamed Mazmira Mohd Masri, Noorhazwani Kamarudin, Detection of bagworm infestation area in oil palm plantation based on UAV remote sensing using machine learning approach, Agriculture (Switzerland) 13 (10) (2023), https://doi.org/10.3390/agriculture13101886.

[10] Adel Ammar, Anis Koubaa, Bilel Benjdira, Deep-learning-based automated palm tree counting and geolocation in large farms from aerial geotagged images, Agronomy 11 (8) (2021), https://doi.org/10.3390/agronomy11081458.

[11] Helmi Z.M. Shafri, Nasrulhapiza Hamdan, M. Iqbal Saripan, Semi-automatic detection and counting of oil palm trees from high spatial resolution Airborne Imagery, Int. J. Remote Sens. 32 (8) (2011) 2095–2115, https://doi.org/10.1080/01431161003662928.

[12] S.N.H. Syed Hanapi, S.A.A. Shukor, J. Johari, A review on remote sensing-based method for tree detection and delineation, in: IOP Conference Series: Materials Science and Engineering, IOP Publishing Ltd, 2019, https://doi.org/10.1088/1757-899X/705/1/012024.

[13] Weijia Li, Haohuan Fu, Le Yu, Peng Gong, Duole Feng, Congcong Li, Nicholas Clinton, Stacked autoencoder-based deep learning for remote-sensing image classification: a case study of african land-cover mapping, Int. J. Remote Sens. 37 (23) (2016) 5632–5646, https://doi.org/10.1080/01431161.2016.1246775.

[14] Lionel Pibre, Marc Chaumont, Gérard Subsol, Dino Ienco, Derras Mustapha, How to deal with multi-source data for tree detection based on deep learning, 5th IEEE global conference, 2017. https://doi.org/10.1109/Global.

[15] Bipul Neupane, Teerayut Horanont, Nguyen Duy Hung, Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV), PLoS ONE 14 (10) (2019), https://doi.org/10.1371/journal.pone.0223906.

[16] Hanae Moussaoui, Nabil El Akkad, Mohamed Benslimane, Walid El-Shafai, Abdullah Baihan, Chaminda Hewage, Rajkumar Singh Rathore, Enhancing automated vehicle identification by integrating YOLO v8 and OCR techniques for high-precision license plate detection and recognition, Sci. Rep. 14 (1) (2024) 1–17, https://doi.org/10.1038/s41598-024-65272-1.

[17] Joseph Redmon, Divvala Santosh, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection, 2016. http://pjreddie.com/yolo/.

[18] Noor Ul Ain Tahir, Zhe Long, Zuping Zhang, Muhammad Asim, Mohammed ELAffendi, PVswin-YOLOv8s: UAV-based pedestrian and vehicle detection for traffic management in smart cities using improved YOLOv8, Drones 8 (3) (2024) 1–20, https://doi.org/10.3390/drones8030084.

[19] G. Jocher, A. Stoken, J. Borovec, A Chaurasia… - …, 2021. 2020. "ultralytics/yolov5: v5. 0-YOLOv5-P6 1280 Models, AWS, Supervise. Ly and YouTube Integrations".

[20] Ge, Zheng, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun. 2021. "YOLOX: exceeding YOLO Series in 2021." http://arxiv.org/abs/2107.08430.

[21] Long, Xiang, Kaipeng Deng, Guanzhong Wang, Yang Zhang, Qingqing Dang, Yuan Gao, Hui Shen, et al. 2020. "PP-YOLO: an effective and implementation of object detector." http://arxiv.org/abs/2007.12099.

[22] Huang, Xin, Xinxin Wang, Wenyu Lv, Xiaying Bai, Xiang Long, Kaipeng Deng, Qingqing Dang, et al. 2021. "PP-YOLOv2: a practical object detector." http://arxiv.org/abs/2104.10419.

[23] Li, Chuyi, Lulu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, Liang Li, Zaidan Ke, et al. 2022. "YOLOv6: a single-stage object detection framework for industrial applications." http://arxiv.org/abs/2209.02976.

[24] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. 2022. "YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." http://arxiv.org/abs/2207.02696.

[25] Reis, Dillon, Jordan Kupec, Jacqueline Hong, and Ahmad Daoudi. 2023. "Real-time flying object detection with YOLOv8." http://arxiv.org/abs/2305.09972.

[26] Aripriharta Aripriharta, Adim Firmansah, Nandang Mufti, Gwo-Jiun Horng, Norzanah Rosmin, Smartphone for palm oil fruit counting to reduce embezzlement in harvesting season, Bullet. Social Informat. Theory Appli. 4 (2) (2020) 76–82, https://doi.org/10.31763/businta.v4i2.283.

[27] Pinaki Nath Chowdhury, Palaiahnakote Shivakumara, Lokesh Nandanwar, Faizal Samiron, Umapada Pal, Tong Lu, Oil palm tree counting in drone images, Pattern Recognit. Lett. 153 (2022) 1–9, https://doi.org/10.1016/j.patrec.2021.11.016.

[28] Mohamad Haniff Junos, Anis Salwa Mohd Khairuddin, Subbiah Thannirmalai, Mahidzal Dahari, An optimized YOLO-based object detection model for crop harvesting system, IET Image Process. 15 (9) (2021) 2112–2125, https://doi.org/10.1049/ipr2.12181.

[29] Li, Weijia, Haohuan Fu, Le Yu, and Arthur Cracknell. 2017. "Deep learning based oil palm tree detection and counting for high-resolution remote sensing images." *Remote Sens.* 9(1). https://doi.org/10.3390/rs9010022.

[30] Weijia Li, Haohuan Fu, Le Yu, Deep Convolutional Neural Network Based Large-Scale Oil Palm Tree Detection for High-Resolution Remote Sensing Images, 2017. 10.1109/IGARSS.2017.8127085.

[31] Nurulain Abd Mubin, Eiswary Nadarajoo, Helmi Zulhaidi Mohd Shafri, Alireza Hamedianfar, Young and mature oil palm tree detection and counting using convolutional neural network deep learning method, Int. J. Remote Sens. 40 (19) (2019) 7500–7515, https://doi.org/10.1080/01431161.2019.1569282.

[32] Isis Bonet, Fabio Caraffini, Alejandro Peña, Puerta Alejandro, Gongora Mario. Oil palm detection via deep transfer learning, 2020. 10.1109/CEC48606.2020.9185838.

[33] Xinni Liu, Kamarul Hawari Ghazali, Fengrong Han, Izzeldin Ibrahim Mohamed, Automatic detection of oil palm tree from UAV images based on the deep learning method, Appl. Artificial Intell. 35 (1) (2021) 13–24, https://doi.org/10.1080/08839514.2020.1831226.

[34] Hery Wibowo, Imas Sukaesih Sitanggang, Mushthofa Mushthofa, Hari Agung Adrianto, Large-scale oil palm trees detection from high-resolution remote sensing images using deep learning, Big Data Cognit. Comput. 6 (3) (2022), https://doi.org/10.3390/bdcc6030089.

[35] DJI Mavic air drone specifications datasheet. https://dronespec.dronedesk.io/dji-mavic-air.

[36] Redmon, Joseph, and Ali Farhadi. 2018. "YOLOv3: an incremental improvement." http://arxiv.org/abs/1804.02767.

[37] Szeliski, Richard. 2010. *Computer Vision: Algorithms and Applications*.

[38] Connor Shorten, Taghi M. Khoshgoftaar, A survey on image data augmentation for deep learning, J. Big Data 6 (1) (2019), https://doi.org/10.1186/s40537-019-0197-0.

[39] Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury Google, Gregory Chanan, Trevor Killeen, et al. 2019. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*.

[40] Reda Yacouby, Dustin Axman, Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models Amazon Alexa (2020).

[41] Subrata Goswami, "False Detection *(Positives and Negatives) in Object Detection*", *arXiv preprint, (2020), doi:https://doi.org/10.48550/arXiv.2008.06986.*

[42] Cyril Goutte, Gaussier Eric. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation Cyril, 2005.