



Journal home page: <http://ajarcde-safe-network.org> ISSN 2581-0405

Oil Palm Crown Detection and Tree Counting Using Roboflow Detection Transformer on UAV Imagery

Ani Dijah Rahajoe^{*1}, Denisa Septalian Alhamda¹, Angelo A Beltran Jr² and Muhammad Suriansyah³

¹ Department of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, Indonesia,

² Department of Electronics Engineering, Adamson University, Philippines

³ Postgraduate School, Universitas Diponegoro, Semarang, Indonesia

ARTICLE INFO

Article History:

Received: 05 June 2026

Final Revision: 17 June 2026

Accepted: 28 June 2026

Online Publication: 28 June 2026

KEYWORDS

oil palm crown detection, tree counting, UAV imagery, RF-DETR, CLAHE, deep learning

CORRESPONDING AUTHOR

*E-mail: anidijah.if@upnjatim.ac.id

ABSTRACT

Oil palm plantations require accurate and timely inventory data to support plantation management, productivity assessment, and sustainable agricultural practices. However, manual tree inventory in large plantation areas is labor-intensive, time-consuming, and prone to human error. This study proposes an automated approach for oil palm crown detection and tree counting using high-resolution Unmanned Aerial Vehicle (UAV) imagery. To improve image quality, Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied as a preprocessing step, and the Roboflow Detection Transformer (RF-DETR) was used as the object detection model. The proposed method was evaluated using 1,135 UAV images containing 56,547 annotated oil palm crowns collected from commercial plantations in West Kalimantan, Indonesia. Experimental results demonstrate that the proposed approach achieved mAP@50 of 97.5%, precision of 97.1%, recall of 96.0%, and F1-score of 96.6% for oil palm crown detection. In the tree-counting evaluation, the system successfully detected 1,105 of 1,120 ground-truth trees, achieving an overall accuracy of 98.5%. Furthermore, the proposed method achieved an average detection time of 16.1 ms, indicating high computational efficiency. These results demonstrate that the proposed framework provides an effective and practical solution for automated oil palm inventory and plantation monitoring using high-resolution UAV imagery.

Contribution to Sustainable Development Goals (SDGs):

SDG 2: Zero Hunger

SDG 9: Industry, Innovation and Infrastructure

SDG 12: Responsible Consumption and Production

SDG 15: Life on Land

1. INTRODUCTION

1.1. Research Background

Oil palm is a commodity that plays an important role in Indonesia's economy and contributes significantly to the development of the national agribusiness sector. Beyond the economic aspect, the palm oil industry also faces demands to

sustainably increase productivity and management efficiency to support more responsible development of the plantation sector [1]. Effective palm oil plantation management requires accurate information on the number of trees, plant distribution, and growth conditions in each plantation block. Information is very important for estimating production, planning maintenance, evaluating plant health, and making data-based harvesting decisions. However, the inventory process for oil palm plantations is still largely conducted manually through field surveys, which require



This work is licensed under a Creative Commons Attribution- NonCommercial - ShareAlike 4.0 International License

Published under license by SAFE Network

significant time, effort, and energy, especially in large plantation areas. In West Kalimantan, the Regency Kubu Raya.

Advances in sensing technology offer more alternatives for efficient plantation monitoring. One of the technologies used is a capable Unmanned Aerial Vehicle (UAV). They produce high-resolution aerial images with greater operational flexibility than conventional data acquisition methods [2], [3]. UAVs enable fast, detailed spatial data collection, so they are widely used in applications such as mapping, vegetation monitoring, and object-based image analysis [4], [5].

Although automatic detection of oil palm in UAV imagery remains challenging, it presents various obstacles. Variations in size, nonuniform lighting conditions, shadow vegetation, and high plant density often cause objects to overlap, complicating identification. Conditions that can lower system detection performance for objects, especially when the image has low contrast or lacks clear visual details.

To overcome the problem of low-quality images, the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique has been widely used to improve local contrast in digital images. CLAHE can increase object visibility without excessively amplifying noise, so it is widely used in various computer vision and image analysis applications. Previous research has shown that applying CLAHE to UAV imagery can improve visual quality and significantly enhance object detection performance under low-light conditions [6]. Additionally, CLAHE has been shown to effectively increase the representation of visual features required by deep learning models [7], [8].

On the other hand, the development of deep learning has produced various object detection models with increasing accuracy. Transformer-based models are among the most promising approaches because of their ability to capture global relations between image features. Unlike approaches based on more advanced convolutional neural networks focused on local information, transformers can model context more broadly, so they have the potential to achieve better performance on objects that are close together or overlapping. It is highly relevant for the detection of oil palm in plantation areas with tall, dense canopies.

Previous studies have extensively investigated object detection models, including YOLO, Faster R-CNN, DETR, and other transformer-based architectures, for various agricultural monitoring applications. Most of these studies primarily focused on improving object detection accuracy by comparing different detection algorithms. However, relatively few studies have validated automated oil palm crown detection against manually verified tree inventory data collected from operational plantations [1] and [3]. Therefore, this study emphasizes the practical applicability of RF-DETR by evaluating automated oil palm crown detection and tree-counting results against expert-verified ground-truth data. This application-oriented evaluation provides practical evidence of the proposed method's effectiveness for plantation inventory and monitoring rather than solely assessing algorithmic detection performance.

Based on these challenges, this study proposes an automated oil palm crown detection and tree-counting framework that integrates Contrast Limited Adaptive Histogram Equalization (CLAHE) with the Roboflow Detection Transformer (RF-DETR) for high-resolution UAV imagery. Unlike previous studies that primarily emphasize comparisons among object detection algorithms, this research evaluates the practical applicability of RF-DETR by comparing its automated detection and tree-

counting results against manually verified ground-truth data. This application-oriented evaluation reflects operational plantation inventory practices, where the primary objective is to achieve reliable, efficient tree counting rather than to benchmark different detection architectures.

1.2. Literature Review

The development of sensing technology has enabled the use of high-resolution aerial imagery for various monitoring needs in natural and plantation environments. The use of Unmanned Aerial Vehicles (UAVs) is increasingly popular because they can produce high-resolution images at relatively low operational cost and offer greater flexibility for data acquisition than conventional sensing platforms. The characteristics of UAVs make them an effective data source for periodic mapping and monitoring of plantation areas.

Various studies have utilized UAV imagery to support the automatic detection of objects using computer vision and deep learning approaches. For example, research on the development of the YOLOv11 model improves object detection in complex environments while maintaining high computational efficiency [9]. Advances in deep learning enable automatic feature extraction, yielding better performance than methods based on manually extracted features. In the context of plantation management, this approach offers opportunities for object identification and counting in a faster, more consistent, and more accurate way than conventional field surveys.

Thus, image quality remains one of the main factors influencing the performance of object detection systems. Variations in lighting, shadows, vegetation, low contrast, and environmental conditions during image acquisition often make objects difficult to recognize. To overcome this problem, various techniques for improving image quality have been developed. One of the many methods used is Contrast Limited Adaptive Histogram Equalization (CLAHE), which can increase local contrast without amplifying excessive noise. Research conducted by [7] shows that CLAHE is effective in improving the visual quality of the image and supports computer vision analysis [7]. Similar results were also reported by [8], which showed that CLAHE implementation can enhance deep learning models' ability to extract features important for digital image classification, thereby improving performance [8].

In UAV environments, lighting quality often affects the success of object detection. Research [6] developed an improved CLAHE-based image enhancement method for UAV imagery under low-lighting conditions and demonstrated a significant improvement in object detection accuracy after the enhancement. Research results show that improving image quality before inference can help the model recognize objects more accurately, especially under challenging visual conditions.

Recent advances in object detection have shown a clear transition from convolutional neural network (CNN)-based detectors to transformer-based architectures. CNN-based models such as Faster R-CNN and the YOLO family have demonstrated strong performance in many agricultural object detection tasks due to their efficient feature extraction capabilities [9]. However, these approaches mainly rely on local receptive fields, making them less effective in scenes containing densely distributed or overlapping objects. Transformer-based detectors address this limitation by employing self-attention mechanisms that capture long-range spatial dependencies and global contextual

information, enabling more robust object localization in complex environments. These characteristics make transformer-based approaches particularly suitable for detecting densely distributed oil palm crowns in high-resolution UAV imagery [2], [5].

Although previous studies have reported promising results using CNN- and transformer-based object detectors, several limitations remain [6], [10], [11]. Most studies primarily emphasize detection accuracy while giving limited attention to practical plantation inventory, where the reliability of automated tree counting must be validated against manually verified field data. In addition, only a limited number of studies have investigated the integration of image enhancement techniques such as CLAHE with transformer-based detectors for oil palm crown detection in dense plantation environments [6], [10], [11], [12], [13]. These limitations indicate the need for an application-oriented framework that combines robust image enhancement, transformer-based object detection, and validation using expert-verified ground-truth data.

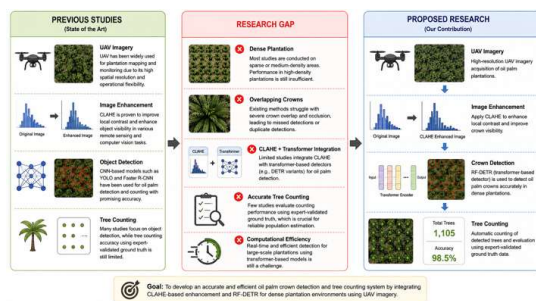


Fig. 1. Research Gap Analysis and Positioning of the Proposed Method

Based on Figure 1, previous research has shown successful utilization of UAV imagery for plantation mapping and monitoring because it is capable of generating spatial data with good flexibility and operations [4]. In terms of improving image quality, the CLAHE method has proven capable of increasing local contrast and clarifying object details, thereby supporting various computer vision and remote sensing applications [7] and [8]. In addition, CNN-based deep learning models such as YOLO and Faster R-CNN have been widely used for various object detection tasks and demonstrate strong performance on images.

However, there are still a number of limitations that become research gaps in this study at the moment. Most studies have been conducted under low- to moderate-object density conditions, so their performance has not yet been extensively evaluated in high-density plantation areas. Research integrating the CLAHE technique with a transformer-based detection model remains relatively limited, especially in applications for detecting oil palm crowns. On the other hand, evaluation of the accuracy of tree counting using ground-truth data verified by experts is still seldom reported. Based on this gap, this research proposes integrating CLAHE and RF-DETR to detect oil palm titles and perform tree counting in a fully automatic way using high-resolution UAV imagery. Approach: This is expected to increase image input quality, improve detection performance on overlapping oil palm crowns, and produce more accurate and efficient calculations in high-density plantation areas.

Therefore, this research proposes integrating CLAHE and RF-DETR to detect oil palm crowns and to perform tree counting

using high-resolution UAV imagery. The proposed approach is expected to improve image input quality, enhance detection performance in the presence of overlapping oil palm crowns, and yield more accurate and efficient calculations than conventional approaches.

1.3. Research Objective

This study aim for develop system detection oil palm crown and tree counting in a way automatic use high-resolution UAV imagery tall through integration technique improvement quality Contrast Limited Adaptive Histogram Equalization (CLAHE) image and detection model object RF-DETR transformer -based approach This designed For overcome various challenges in plantation areas with level density height, condition overlapping oil palm crowns overlapping overlap, and variation quality frequent lighting found in the image UAV aerial. In particular, the research aims to evaluate the effectiveness of CLAHE in improving the quality of UAV imagery, measure RF-DETR's capabilities in detecting oil palm under dense plantation conditions, and assess the accuracy of tree-based calculations using ground-truth data verified by experts. In addition, this research aims to analyze the efficiency of the proposed computing method as a solution for supporting the monitoring and management of oil palm plantations in a faster, more accurate, and more sustainable way [4], [6].

2. MATERIALS AND METHODS

2.1. Study Area and Dataset

This study uses primary data obtained from an agro-estate in West Kalimantan, Indonesia. The data used consists of high-resolution aerial imagery acquired using a DJI Mavic 4 Enterprise and spatial data on plantation blocks in shapefile format (.shp). The aerial image obtained has a spatial resolution of 7.95 cm/pixel, which is sufficient to represent field conditions in detail, including planting patterns, vegetation distribution, and oil palm tree density.

Table 1. UAV Data Acquisition and Study Area Specifications

Parameter	Value
Study Area	West Kalimantan, Indonesia
UAV Platform	DJI Mavic 4 Enterprise
Camera	High-resolution RGB Camera
Flight Altitude	100 m
Image Format	TIFF
Total UAV Images	1,135
Ground Sampling Distance	7.95 cm/pixel
Oil Palm Age	3–10 years
Representative Sample	Approximately 5 years

In Table 1, the aerial survey was conducted using a DJI Mavic 4 Enterprise equipped with a high-resolution RGB camera. The UAV was operated at a flight altitude of 100 m to acquire high-resolution imagery over commercial oil palm plantations. A total of 1,135 UAV images in TIFF format were collected and processed to generate orthomosaic imagery with a ground sampling distance (GSD) of 7.95 cm/pixel. The surveyed plantations contained oil palm trees ranging from 3 to 10 years of age, while the representative examples presented in this paper were obtained from plantation blocks with an average tree age of approximately five years.

Utilization of UAV imagery in the study plantation has many uses because it can generate high-resolution, flexible spatial data, with better acquisition than conventional sensing platforms [4],[4], [14]. Various types of payloads, including high-resolution RGB cameras, enable UAVs to be used for mapping, vegetation monitoring, and image-based object analysis across a variety of application sectors. The image data used in the study comprises training, validation, and test sets from different plantations. To ensure the ability to generalize models to unspecified data not seen previously. Figure 2. This can represent the current condition of plantations, including planting patterns, vegetation distribution, and oil palm density, which is the object of study.



Fig. 2. Representative UAV Orthomosaic Tiles Used for Training and Validation.

Table 2. Dataset Distribution for RF-DETR Training

Dataset	Images	Bounding Boxes
Training	908	47,186
Validation	114	4,658
Testing	113	4,703
Total	1,135	56,547

Table 2 lists 1,135 UAV images with 56,547 manually annotated bounding boxes. The dataset was randomly divided into training, validation, and testing subsets using an 80:10:10 ratio. Specifically, 908 images containing 47,186 annotated objects were allocated for model training, while 114 validation images containing 4,658 annotated objects were used for model optimization. The remaining 113 test images, containing 4,703 annotated objects, were reserved exclusively for independent performance evaluation. This data partitioning minimizes overfitting and provides an unbiased assessment of the proposed RF-DETR model.

2.2. Research Framework

The proposed research flow, as shown in Figure 3, consists of several main stages: UAV data acquisition, orthomosaic generation, image quality enhancement using CLAHE, dataset preparation and annotation, RF-DETR model training, oil palm canopy detection, and automatic tree counting. All stages are designed to produce a detection system that can work effectively in plantation conditions with high-density, interconnected canopies. overlapping overlap.

2.3. Orthomosaic Generation

The orthomosaic formation stage is carried out after the aerial image acquisition process using a UAV. Orthomosaics are the result of combining several overlapping images into a single, integrated spatial image. This process aims to produce a more accurate representation of the plantation area to serve as a basis for analysis and object detection. Orthomosaics in UAV-based

remote sensing applications are widely used because they provide more consistent and detailed spatial information than a single image [4].

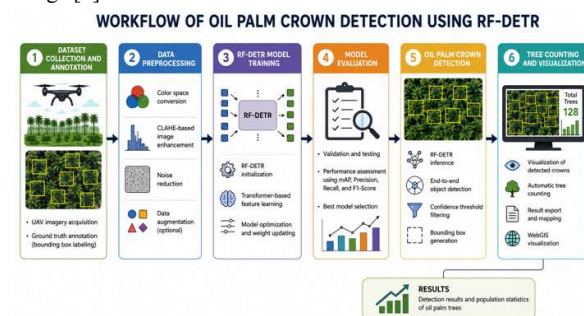


Fig. 3. Workflow of oil Palm tree Detection Using RF-DETR and CLAHE

In this study, UAV-acquired images were processed using photogrammetric software to produce high-resolution orthomosaics. The overlap between images is an important factor because it allows for feature matching. Matching and geometric correction are performed accurately. Next, image correction and alignment processes are performed to reduce geometric distortion, lighting differences, and positional mismatches between images.

Figure 4 shows the results of the successful orthomosaic process. Integrating a number of UAV images into one unified image with intact spatial integrity. This result retains the oil palm header details, which is good, so it can be used for object-based detection using deep learning. High spatial resolution allows the model to recognise variations in shape and size, including in areas with high tree density.



Figure 4. Orthomosaic Generation Results

The resulting orthomosaic has a spatial resolution of 7.95 cm/pixel and is used as the primary source for training, validating, and testing the RF-DETR model. In addition to producing a more accurate spatial representation, using orthomosaics helps avoid object duplication across datasets, thereby reducing the potential for overfitting and improving the model's generalizability.

2.4. CLAHE-Based Image Enhancement

Before the annotation and model training processes are carried out, all image Orthomosaics are also improved through stage-wise image quality enhancement using Contrast Limited Adaptive Histogram Equalization (CLAHE). The aim of this stage is to increase local image contrast so that the oil palm header limit can be seen more clearly, especially in areas with uneven lighting, shadowed vegetation, and low contrast.

CLAHE is derived from the Adaptive Histogram Equalization (AHE) method, which operates on the image in a

number of small blocks (tiles) and performs histogram equalization locally for each block. Different from AHE, CLAHE applies a clip limit mechanism to limit excessive contrast amplification, so that it can reduce increased noise in the image [12]. Various studies show that CLAHE is effective at increasing object visibility and preserving important details in low-contrast images, so it is often used as a preprocessing step in object detection systems and computer vision [7], [15], [16], [17], and [18].

In this research, CLAHE is applied to all Orthomosaic images before the dataset's cropping and annotation processes are carried out. The enhancement stage is performed to clarify the boundary of oil palm, increase object visibility, and strengthen the visual feature representation for the RF-DETR model to study. Improved image quality before training is known to help deep learning models extract features more effectively, especially under complex environmental conditions [8],[8], and [19]. To visually demonstrate the effectiveness of the proposed preprocessing technique, a comparison between the original Orthomosaic image and the CLAHE-enhanced image is presented in Figure 5.



Fig.5. Visual Comparison of Orthomosaic Images Before and After CLAHE Enhancement.

In addition to improving image visual quality, CLAHE has been shown to enhance object detection performance in UAV images under challenging lighting conditions [10]. Study [6] reports that applying CLAHE to UAV imagery significantly improves object quality, thereby enabling more accurate detection. Therefore, CLAHE was chosen as a preprocessing step in this research to enhance the RF-DETR model's ability to detect title oil palm in plantation areas with high density and overlapping titles.

2.5. Dataset Preparation and Annotation

After the quality image upgrade process using CLAHE is finished, the next stage is dataset preparation and annotation. The aim of this stage is to produce structured, appropriate datasets for training the RF-DETR model. All processed Orthomosaic images are then divided into smaller parts (tiles) using QGIS software. This process is done because the image size is very large, and an Orthomosaic can increase memory usage and computational time during model training.

In this research, the Orthomosaic image was cropped to 416×416 pixels using the raster divider feature in QGIS. The cropping process is performed to improve computational efficiency while retaining the spatial details required for object detection in the title, oil palm. Besides that, the divided image becomes a larger size. Smaller data also simplifies the annotation process and dataset management. The resulting dataset is further split into training, validation, and testing data. Training data is used to learn the characteristics of oil palm trees, while validation data is used to monitor model performance during training.

Meanwhile, test data were used to evaluate the model's ability to generalize to previously unseen data. Stage annotation was performed using the LabelImg software with the bounding box method. Each oil palm crown is given a bounding box that surrounds the object, and the bounding box is labeled with the appropriate class. The annotation process is done manually to ensure accurate positioning and size of objects on each image.

Annotation was carried out under various conditions of objects, including small-sized titles, large-sized titles, interconnected objects close together, and objects that experience overlapping. Approach: This approach aims to increase the model's capability to recognize variation in the shape and size of oil palm crown under complex plantation conditions. Examples of annotation processes and results, including dataset labeling, are shown in Figure 6.



Fig. 6. Dataset Annotation Process Using LabelImg.

Figure 6 shows the annotation process. The object uses a bounding box method in the device software *LabelImg*. Each oil palm title is labeled by position and size, with objects visible in the image. Furthermore, Figure 7 shows the final annotations, ready for use in training the RF-DETR model. The resulting dataset annotation serves as the ground truth for the learning process and model performance evaluation.

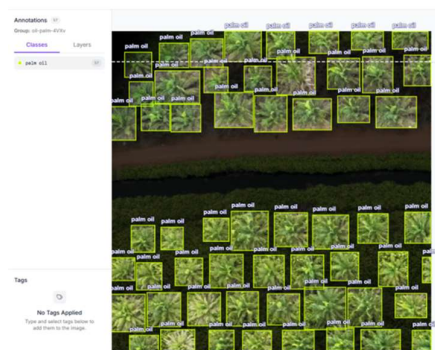


Fig. 7. Annotated Oil Palm Crown Dataset

2.6. RF-DETR Model Training

After the annotation process is complete, the dataset is used to train the RF-DETR model to automatically detect oil palm crowns in UAV imagery. RF-DETR is a transformer-based object detection model that utilizes self-attention mechanisms to learn global relations between features in an image. Unlike the Convolutional Neural Network (CNN)-based approach, the transformer can capture spatial context more comprehensively, making it effective at detecting mutually interacting objects that are close together and overlap. The use of transformer-based

models for high-resolution image analysis has improved performance across various computer vision and remote sensing applications by capturing global information from images. It is very appropriate for case detection in oil palm, which varies in size and shape and has high-density levels. Additionally, improved image quality from CLAHE at the previous stage is expected to help the model learn object features more effectively [6], [7], and [8].

The training process was performed using the cloud-based Google Colaboratory platform with NVIDIA Tesla T4 Graphics Processing Unit (GPU) support. GPU utilization enables deep learning training processes. It is more efficient than computing-central-processing-unit (CPU)-based methods, especially for high-resolution image datasets with many objects. The hardware and software configuration used during the training process is shown in Table 3.

Table 3. Hardware and Software Configuration Used for RF-DETR Training and CLAHE

Component	Specification
Platform	Google Collaboratory
Model	RF-DETR Nano
Python	3.10
GPU	Nvidia Tesla T4
RAM	16 GB
Storage	80 GB
CLAHE Clip Limit	2.0

During training, the model receives image data and annotations as input and learns the relationship between visual features and the object's location, represented by a bounding box. During this process, the weight network is iteratively updated via an optimization algorithm until the best-performing model is obtained. Validation data is used to monitor the model's ability to generalize and to prevent overfitting during training. Some studies have shown that improving image quality prior to training can enhance the performance of deep learning models on classification and object detection tasks [11], [13]. The model image was chosen based on the results evaluation on the validation dataset using object detection metrics, namely mean Average Precision (mAP), Precision, Recall, and F1-score. The model with the best performance was also used during testing to detect and count coconut palm trees in the dataset that had not been used during training. RF-DETR Nano was selected because it strikes a balance between detection accuracy and computational efficiency, making it suitable for analyzing high-resolution UAV imagery, which requires fast inference with consistent accuracy.

3. RESULTS AND DISCUSSION

3.1. Evaluation Model

The combination of Precision and Recall results in an F1-score value of 96.6%, which indicates a good balance between accuracy and completeness of detection. Overall, these results show that RF-DETR is highly effective at detecting oil palm crowns in high-resolution UAV imagery, even under high-density plantation conditions. Before testing on an independent dataset, the performance of the RF-DETR model is better than previously evaluated, as measured by several standard object detection

metrics: mean Average Precision (mAP@50), Precision, Recall, and F1-score. The metrics used to measure model capabilities in oil palm tree object localization and classification are accurate when applied to high-resolution UAV imagery. Evaluation results for model performance are shown in Table 4.

Table 4. Performance Evaluation Results of the RF-DETR Model

Metric	Value (%)
mAP@50	97.5
Precision	97.1
Recall	96.0
F1-Score	96.6

Based on the evaluation results shown in Table 4, the RF-DETR model achieved an mAP@50 of 97.5%, indicating its ability to detect and localize objects with high accuracy at the Intersection over Union (IoU) threshold of 0.5. The Precision value of 97.1% indicates that a large proportion of detected objects are correct detections, whereas the Recall value of 96.0% shows that the model is capable of identifying almost all objects in the testing dataset.

Besides that, the combination of Precision and Recall results in an F1-score value of 96.6%, which reflects a good balance between accuracy and completeness of detection. These results show that the RF-DETR model has stable and reliable performance in detecting tall oil palm in plantation areas with high density. Its high mark across all metric evaluations also indicates that the model is capable of maintaining high detection quality, although there is variation in title size, nonuniform lighting conditions, and mutually interacting objects in close proximity.

3.2. Oil Palm Crown Detection Results

After the RF-DETR model achieves good performance during the stage evaluation, the next step is to apply it to the test dataset to automatically detect oil palm crowns. Detection results are displayed as a bounding box around each oil palm tree recognized by the model. Visualization results of the detection are shown in Figure 8. Based on the detection results, the big oil palm crown is successfully identified with Good, although there is variation in object size, high plant density, as well as overlapping oil palm crowns close by. Deep model capabilities detect objects under conditions that demonstrate RF-DETR's ability to utilize spatial information effectively to differentiate objects with similar visual characteristics. The ability to model global relations between objects is one of the advantages of transformer-based approaches over purely feature-focused, local approaches [3].



Fig.8. Oil Palm Crown Detection Results Using RF-DETR

In addition to bounding box visualization, detection results are displayed as a spatial map, as shown in Figure 9. This mapping describes the distribution of oil palm in the research area and simplifies validation against ground-truth data. Spatial visualization also allows for identifying areas with high tree density and areas that have plants that are not distributed evenly.

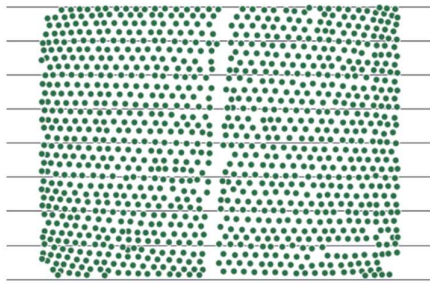


Fig.9. Spatial Distribution of Detected Oil Palm Trees

Visually, the detection results show high suitability with the actual situation in the field. Most of the detected oil palm trees are correctly identified, including in areas with dense planting and objects that are close together. This shows that the combination of CLAHE and RF-DETR can produce representations with sufficient features to support object detection in high-resolution UAV images. Thus, some objects still experience error detection, especially in the border areas of the Orthomosaic image. In that area, object quality tends to decrease due to geometric distortion and imperfections in the mosaicking process, so parts are not recorded intact. Condition: This potential cause is that objects are not detected or produce suboptimal predictions.

3.3. Tree Counting Performance

Besides being used to detect oil palm crown, results obtained from the RF-DETR model are also utilized for calculating the number of trees automatically. Evaluation was done by comparing the number of successful objects detected by the model against the ground truth data that has been verified by experts using the QGIS software. The test dataset covers a 1-hectare plantation area with tall, dense trees, which can represent actual field conditions.

Test results show that the model successfully detects up to 1105 oil palm trees out of a total of 1120 trees in the ground truth data. The difference between the number of detected trees and the current number of trees is relatively small, at only 15. This result shows that the proposed approach can produce an estimate very close to the actual tree condition in the field. To evaluate the performance of tree counting, Precision, Recall, and Overall Accuracy metrics were used. Summary results from the tree-counting evaluation are shown in Table 5.

Table 5. Tree Counting Performance Results

Parameters	Value
Ground Truth Trees	1120
Detected Trees	1105
Precision	0.99
Recall	0.98
Overall Accuracy	98.5%
Detection Time	16.1 ms

Based on Table 5, the obtained Precision value of 0.99 indicates that almost all detected objects are correct. The Recall value of 0.98 indicates that a large part of the tree contained in the ground truth data was successfully recognized by the model. The combination of the two metrics yields a detection-level accuracy of 98.5%, indicating a very high degree of suitability between the detection results and actual field conditions. In addition to high accuracy, the proposed method shows excellent computational efficiency. The entire process of detecting and calculating the tree can be completed in approximately 16.1 milliseconds. The relatively short processing time indicates that this method has potential for use in plantation monitoring systems in a fast and efficient monitoring, especially in large areas with many trees. However, some objects have not yet been detected optimally. Most of the detection errors occur in the border area of the orthomosaic image, which has undergone geometric distortion and a decline in visual quality due to the mosaicking process. In this area, the tree's shape is not always captured intact, leading to objects being missed during detection. Overall, the results show that integrating CLAHE and RF-DETR yields excellent performance on high-resolution UAV imagery.

3.4. Discussion

One of the contributing factors to improved detection performance is the application of Contrast Limited Adaptive Histogram Equalization (CLAHE) at the preprocessing stage. CLAHE can increase local contrast and clarify object boundaries without significantly increasing noise, making the visual features the model receives more representative. Various studies previously report that CLAHE is effectively used to increase the quality of digital images, clarify object details, and increase the performance of deep learning models on various classification and object detection tasks [6], [7], [8], [20]

Besides that, improved image quality prior to training is known to help the model extract more features more effectively, especially in UAV images, which exhibit varying lighting, shadows, vegetation, and nonuniform contrast [4], [6], [8]. Using high-resolution UAV imagery also provides a significant contribution to object detection success because it provides more detailed spatial information compared to medium-resolution images. Some studies show that UAVs are effective platforms for vegetation monitoring, object mapping, and deep learning-based analysis due to their operational flexibility and the quality of the data they generate [4].

Overall, the results of this study show that combining high-quality modern deep learning images and models can increase object detection accuracy in complex environments. Findings: This is consistent with various studies previously reported that image quality and feature representation are important factors that influence computer vision system performance [3], [6], [7], and [20]. Unlike benchmark studies, the proposed framework was designed to evaluate agreement with manually verified plantation inventory data, reflecting real operational requirements.

4. CONCLUSION

This study proposes an approach for detecting oil palm and estimating tree density by automatically integrating Contrast Limited Adaptive Histogram Equalization (CLAHE) with RF-DETR on high-resolution UAV imagery. CLAHE is used to increase image quality through local contrast improvement, while

RF-DETR is utilized to detect the target oil palm in plantation areas with high density and overlapping oil palm crowns close together. Research results show that the proposed model achieves an mAP@50 of 97.5%, precision of 97.1%, recall of 96.0%, and F1-score of 96.6%. At this stage, the tree-counting system successfully detected 1105 of 1120 trees in the ground-truth data, achieving an overall accuracy of 98.5%. In addition to producing high accuracy at the pixel level, the proposed method also shows excellent computational efficiency, with an average detection time of about 16.1 milliseconds. The results show that the effective integration of CLAHE and RF-DETR is used to detect tree titles in oil palm and perform automatic tree counting on high-resolution UAV imagery. This approach potentially supports the inventory, monitoring, and management of oil palm plantations more quickly, accurately, and efficiently than a manual approach.

However, some objects have not yet been successfully detected, especially in the border areas of Orthomosaics, where geometric distortion and a decline in visual quality are evident. Therefore, research can further focus on using more diverse datasets, optimizing the UAV imagery acquisition process, and evaluating models on larger plantation areas to improve generalization.

REFERENCES

- [1] Astari , AJ, Lovett, JC, & Wasesa , M. Sustainable pathways in Indonesia's palm oil industry through historical institutionalism. *World Development Sustainability*. <https://doi.org/10.1016/j.wds.2024.100200>. 2025.
- [2] Laghari, AA, Jumani , AK, Laghari, RA, & Nawaz, H. Unmanned aerial vehicles: A review. In *Cognitive Robotics* (Vol. 3). <https://doi.org/10.1016/j.cogr.2022.12.004>. 2023.
- [3] Chinaramanamma , D., & Anuradha, B. Image Enhancement of Satellite Images Using Contrast Limited Adaptive Histogram Equalization and NLM. *International Journal of Scientific Research in Science and Technology International Journal of Scientific Research in Science and Technology (Wwww.Ijsrst.Com)* | , 11 (2). 2024.
- [4] Lyu, M., Zhao, Y., Huang, C., & Huang, H. Unmanned Aerial Vehicles for Search and Rescue: A Survey. In *Remote Sensing* (Vol. 15, Number 13). <https://doi.org/10.3390/rs15133266>. 2023
- [5] Arafat, MY, Alam, MM, & Moh, S. (2023). Vision-Based Navigation Techniques for Unmanned Aerial Vehicles: Review and Challenges. In *Drones* (Vol. 7, Number 2). <https://doi.org/10.3390/drones7020089>. 2023.
- [6] Yuan, Z., Zeng, J., Wei, Z., Jin, L., Zhao, S., Liu, X., Zhang, Y., & Zhou, G. (2023). CLAHE-Based Low-Light Image Enhancement for Robust Object Detection in Overhead Power Transmission Systems. *IEEE Transactions on Power Delivery* , 38 (3). <https://doi.org/10.1109/TPWRD.2023.3269206>. 2023.
- [7] Kryjak , T., Blachut , K., Szolc , H., & Wasala , M. (2022). Real-Time CLAHE Algorithm Implementation in SoC FPGA Device for 4K UHD Video Stream. *Electronics (Switzerland)* , 11 (14). <https://doi.org/10.3390/electronics11142248>. 2022.
- [8] Wilny Wilson, P., & Jayaseeli , J. D. D. (2025). Enhanced pneumonia detection: leveraging CLAHE in a mobile application. *Biomedical Engineering Communications*,4(4). <https://doi.org/10.53388/BMEC2025023>. 2025.
- [9] Zhou, Y., Zhang, T., Li, Z., & Qiu, J. (2025). Improved Space Object Detection Based on YOLO11. *Aerospace* , 12 (7). <https://doi.org/10.3390/aerospace12070568>. 2025.
- [10] Luo, J., Tian, Y., & Wang, Z. (2024). Research on Unmanned Aerial Vehicle Path Planning. In *Drones* (Vol. 8, Number 2). <https://doi.org/10.3390/drones8020051>. 2024.
- [11] Ganesh Kumar, S.S., & Gudipalli , A. (2024). A comprehensive review on payloads of unmanned aerial vehicles. In *Egyptian Journal of Remote Sensing and Space Science* (Vol. 27, Number 4). <https://doi.org/10.1016/j.ejrs.2024.08.001>. 2024.
- [12] Rentapalli , S., & Khairuzzaman , A. K. M. (2026). Enhancement of Satellite Images Using Contrast-Limited Adaptive Histogram Equalization with Hybrid Gray Wolf and Particle Swarm Optimization. *IEEE Access*.<https://doi.org/10.1109/ACCESS.2026.3651193> . 2026.
- [13] Said, H., Mohamed, S., Shalash , O., Khatab , E., Aman, O., Shaaban, R., & Hesham, M. (2024). Forearm Intravenous Detection and Localization for Autonomous Vein Injection Using Contrast-Limited Adaptive Histogram Equalization Algorithm. *Applied Sciences (Switzerland)* , 14 (16). <https://doi.org/10.3390/app14167115>. 2024.
- [14] Hj Wan Yussof , WNJ, Man, M., Umar, R., Zulkeflee , AN, Awalludin , EA, & Ahmad, N. (2022). Enhancing Moon Crescent Visibility Using Contrast-Limited Adaptive Histogram Equalization and Bilateral Filtering Techniques. *Journal of Telecommunications and Information Technology* , 2022 (1). <https://doi.org/10.26636/jtit.2022.155721>. 2022.
- [15] Hu, C., Li, H., Ma, T., Zeng, C., & Ji, X. (2024). An improved image enhancement algorithm: Radial contrast-limited adaptive histogram equalization. *Multimedia Tools and Applications* , 83 (36). <https://doi.org/10.1007/s11042-024-18922-5>. 2024.
- [16] Azhar, AN bin, Sani, NS, & Wei, LLX (2025). Enhancing COVID-19 Detection in X-Ray Images Through Deep Learning Models with Different Image Preprocessing Techniques. *International Journal of Advanced Computer Science and Applications* , 16 (1). <https://doi.org/10.14569/IJACSA.2025.0160161>. 2025.
- [17] Mehdizadeh , M., Tavakoli Tafti , K., & Soltani , P. (2023). Evaluation of histogram equalization and contrast limited adaptive histogram equalization effect on image quality and fractal dimensions of digital periapical radiographs. *Oral Radiology* , 39 (2). <https://doi.org/10.1007/s11282-022-00654-7>. 2023.
- [18] Chen, RC, Dewi, C., Zhuang, YC, & Chen, JK (2023). Contrast Limited Adaptive Histogram Equalization for Recognizing Road Marking at Night Based on Yolo

- Models. *IEEE Access*, 11. <https://doi.org/10.1109/ACCESS.2023.3309410>. 2023.
- [19] Priyah, R., & Kamalakkannan, S. (2025). Hybrid contrast-limited adaptive histogram equalization and Deep Learning techniques for improving liver tumor detection. *Future Technology*, 4 (3). <https://doi.org/10.55670/fpll.futech.4.3.7>. 2025.
- [20] Dewi, C., Chernovita, HP, Philemon, SA, Ananta, CA, Dai, G., & Chen, APS (2025). Integration of YOLOv9 and Contrast Limited Adaptive Histogram Equalization for Nighttime Traffic Sign Detection. *Mathematical Modeling of Engineering Problems*, 12 (1). <https://doi.org/10.18280/mmep.120105>. 2025.
- [21] Nia, SN, & Shih, FY (2024). Medical X-Ray Image Enhancement Using Global Contrast-Limited Adaptive Histogram Equalization. *International Journal of Pattern Recognition and Artificial Intelligence*, 38 (12). <https://doi.org/10.1142/S0218001424570106>. 2024.
- [22] Manjunath, A., Neelappa, Prakash, Veeramma Yatnalli, & Saroja, S. B. (2023). Performance Analysis of Graph theory-based Contrast Limited Adaptive Histogram Equalization for Image Enhancement. *WSEAS Transactions on Systems*, 22. <https://doi.org/10.37394/23202.2023.22.23>. 2023.