



Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author: BAB 1
Assignment title: Deep learning-based palm tree detection in unmanned aerial ...
Submission title: Deep learning-based palm tree detection in unmanned aerial ...
File name: Deep_learning-based_palm_tree_detection_in_unmanned_aeri...
File size: 1.3M
Page count: 10
Word count: 4,907
Character count: 26,247
Submission date: 19-Mar-2025 08:00AM (UTC+0700)
Submission ID: 2618650747

TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 23, No. 1, February 2025, pp. 156-165
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v23i1.26244

156

Deep learning-based palm tree detection in unmanned aerial vehicle imagery with Mask R-CNN

Agung Syetawan¹, Danang Budi Susetyo^{1*}, Yustisi Lumban-Gaol², Susilo³, Mohammad Ardhah¹, Yunus Susilo⁴, Wahono⁵

¹National Research and Innovation Agency (BRIN), Cibinong, Indonesia
²Department of Geomatics Engineering, Faculty of Engineering, Dr. Soesono University, Surabaya, Indonesia
³Department of Agrotechnology, Faculty of Agriculture and Animal Science, University of Muhammadiyah Malang, Malang, Indonesia
⁴Department of Geomatics Engineering, Faculty of Civil Engineering, Yildiz Technical University, Istanbul, Turkey

Article Info

Article history:
Received Mar 29, 2024
Revised Oct 29, 2024
Accepted Nov 26, 2024

Keywords:
Deep learning
Mask region-based
convolutional neural network
Palm tree
Tree detection
Unmanned aerial vehicle

ABSTRACT

Oil palm is highly valuable in tropical regions like Southeast Asia, including Indonesia. Therefore, accurate monitoring of oil palm trees is necessary for operational efficiency and reducing its environmental impact. Geospatial data, such as orthomosaic imagery from the unmanned aerial vehicle (UAV), can facilitate this goal. This research aims to integrate UAV data with deep learning algorithms, specifically Mask region-based convolutional neural network (R-CNN), to detect oil palm trees in Indonesia. We utilized ResNet-50 as the backbone and trained the model using data sampled from the template matching tool in eCognition. Considering factors like cloud shadows and other features, such as other plants, buildings, and road segments, we divided the study area into three containing different feature combinations in each. The Mask R-CNN model achieved an accuracy exceeding 80%, which is sufficient and makes it suitable for large-scale oil palm tree detection using high-resolution images from UAV.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Danang Budi Susetyo
National Research and Innovation Agency (BRIN)
Jakarta-Bogor Km 46, Cibinong 16911, Indonesia
Email: danang.budi.susetyo@brin.go.id

1. INTRODUCTION

Oil palm holds high commercial value in tropical regions such as Southeast Asia [1], including Indonesia. It has become a national strategic issue that plays an important role in the Indonesian economy [2], which has led to significant growth in the oil palm industry development in this country. Unfortunately, it brings a heavy toll on forests, biodiversity, and carbon stocks [3], and causes deforestation and carbon emissions [4]. In addition, converting fresh fruit bunches into crude palm oil produces several types of waste [5]. Therefore, accurate palm tree monitoring is necessary to minimize their environmental impact.

Palm trees can be monitored effectively using geospatial data. Geospatial technology has undergone significant development, allowing the production of high-resolution and accurate data in recent times. Unmanned aerial vehicle (UAV) is one of the technologies that can be utilized for monitoring palm trees. The advantage of using UAV systems is that it can produce a more detailed view of the earth's surface with very good resolution [6]. Orthophotos generated from UAV can visually interpret individual palm tree structures and plant density [7], [8]. However, a reliable technique is required to optimize the geospatial data obtained through UAV monitoring to produce accurate analysis.

Deep learning, based on convolutional neural network (CNN), is a developed method in computer vision that can be applied for optimizing geospatial data. Since AlexNet [9] achieved satisfying results in the


Journal homepage: <http://telkomnika.iaind.ac.id>

BAB 1

Deep learning-based palm tree detection in unmanned aerial vehicle imagery with Mask R-CNN.pdf

 Deep learning-based palm tree detection in unmanned aerial vehicle imagery with Mask R-CNN

 Prodi Peternakan

 University of Muhammadiyah Malang

Document Details

Submission ID

trn:oid::1:3187466864

10 Pages

Submission Date

Mar 19, 2025, 8:00 AM GMT+7

4,907 Words

Download Date

Mar 19, 2025, 8:03 AM GMT+7

26,247 Characters

File Name

Deep_learning-based_palm_tree_detection_in_unmanned_aerial_vehicle_imagery_with_Mask_R-C....pdf

File Size

1.3 MB

13% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.





Filtered from the Report

- ▶ Bibliography
- ▶ Quoted Text




Exclusions

- ▶ 4 Excluded Matches

Match Groups

-  **24 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
-  **4 Missing Quotations 2%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 10%  Internet sources
- 7%  Publications
- 1%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **24 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
- **4 Missing Quotations 2%**
Matches that are still very similar to source material
- **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 10% Internet sources
- 7% Publications
- 1% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Publication	A Syetian, H Gularso, G I Kusnadi, G N Pramudita. "Precise topographic mappi...	2%
2	Internet	www.mdpi.com	2%
3	Internet	jurnal.big.go.id	1%
4	Internet	www2.mdpi.com	<1%
5	Publication	W Ambarwulan, A Syetian, A Rahadiati, G A Rahmawan. "Monitoring thirty yea...	<1%
6	Student papers	School of Business and Management ITB	<1%
7	Internet	telkomnika.uad.ac.id	<1%
8	Publication	Raja, Kalpana, Suresh Subramani, and Jeyakumar Natarajan. "A hybrid named en...	<1%
9	Internet	www.science.gov	<1%
10	Internet	jjbs.hu.edu.jo	<1%

11	Internet	mycoordinates.org	<1%
12	Internet	www.researchgate.net	<1%
13	Internet	archive.ismrm.org	<1%
14	Publication	Qian Pan, Maofang Gao, Pingbo Wu, Jingwen Yan, Shilei Li. "A Deep-Learning-Bas...	<1%
15	Internet	aj.tubitak.gov.tr	<1%

Deep learning-based palm tree detection in unmanned aerial vehicle imagery with Mask R-CNN

Agung Syetawan¹, Danang Budi Susetyo^{1,4}, Yustisi Lumban-Gaol¹, Susilo¹, Mohammad Ardha¹, Yunus Susilo², Wahono³

¹National Research and Innovation Agency (BRIN), Cibinong, Indonesia

²Department of Geomatics Engineering, Faculty of Engineering, Dr. Soetomo University, Surabaya, Indonesia

³Department of Agrotechnology, Faculty of Agriculture and Animal Science, University of Muhammadiyah Malang, Malang, Indonesia

⁴Department of Geomatic Engineering, Faculty of Civil Engineering, Yildiz Technical University, Istanbul, Türkiye

Article Info

Article history:

Received Mar 29, 2024

Revised Oct 29, 2024

Accepted Nov 26, 2024

Keywords:

Deep learning

Mask region-based

convolutional neural network

Palm tree

Tree detection

Unmanned aerial vehicle

ABSTRACT

Oil palm is highly valuable in tropical regions like Southeast Asia, including Indonesia. Therefore, accurate monitoring of oil palm trees is necessary for operational efficiency and reducing its environmental impact. Geospatial data, such as orthomosaic imagery from the unmanned aerial vehicle (UAV), can facilitate this goal. This research aims to integrate UAV data with deep learning algorithms, specifically Mask region-based convolutional neural network (R-CNN), to detect oil palm trees in Indonesia. We utilized Resnet-50 as the backbone and trained the model using data sampled from the template matching tool in eCognition. Considering factors like cloud shadows and other features, such as other plants, buildings, and road segments, we divided the study area into three containing different feature combinations in each. The Mask R-CNN model achieved an accuracy exceeding 80%, which is sufficient and makes it suitable for large-scale oil palm tree detection using high resolution images from UAV.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Danang Budi Susetyo

National Research and Innovation Agency (BRIN)

Jakarta-Bogor Km 46, Cibinong 16911, Indonesia

Email: danang.budi.susetyo@brin.go.id

1. INTRODUCTION

Oil palm holds high commercial value in tropical regions such as Southeast Asia [1], including Indonesia. It has become a national strategic issue that plays an important role in the Indonesian economy [2], which has led to significant growth in the oil palm industry development in this country. Unfortunately, it brings a heavy toll on forests, biodiversity, and carbon stocks [3], and causes deforestation and carbon emissions [4]. In addition, converting fresh fruit bunches into crude palm oil produces several types of waste [5]. Therefore, accurate palm tree monitoring is necessary to minimize their environmental impact.

Palm trees can be monitored effectively using geospatial data. Geospatial technology has undergone significant development, allowing the production of high-resolution and accurate data in recent times. Unmanned aerial vehicle (UAV) is one of the technologies that can be utilized for monitoring palm trees. The advantage of using UAV system is that it can produce a more detailed view of the earth's surface with very good resolution [6]. Orthophotos generated from UAV can visually interpret individual palm tree structures and plant density [7], [8]. However, a reliable technique is required to optimize the geospatial data obtained through UAV monitoring to produce accurate analysis.

Deep learning, based on convolutional neural network (CNN), is a developed method in computer vision that can be applied for optimizing geospatial data. Since AlexNet [9] achieved satisfying results in the

ImageNet large scale visual recognition challenge (ILSVRC) in 2012, CNN has gained popularity. Then, research on deep learning has significantly advanced, including in earth observation [10]. Deep learning has been widely utilized for various geospatial and remote sensing applications, including land cover classification, segmentation, and object detection. Among the models used for these applications, CNN is the most commonly applied [11]. Hence, deep learning should be considered for mapping, particularly in monitoring palm trees.

The application of deep learning for tree counting has become increasingly widespread. The detection and accurate counting of oil palm are crucial aspects of managing an oil palm plantation [12], including health and risk assessment, pest control, and sustainable management within the date palm industry [13]. Li *et al.* [14] successfully optimized artificial neural network (ANN) for detecting and counting oil palm trees, achieving a remarkable accuracy of over 78.35% in tree detection from high resolution satellite imagery. Ribera *et al.* [15] employed UAV in combination with CNN to accomplish a similar task and achieved a remarkably low mean absolute percentage error (MAPE) of only 6.7%.

In our study, we employed Mask region-based convolutional neural network (R-CNN), a powerful deep learning model, to detect and count palm trees using UAV imagery. The application of Mask R-CNN for palm tree detection and counting from UAV imagery has been explored in several studies. Yarak *et al.* [16] used Faster-RCNN to test the ability of automatic oil palm detection at different flying altitudes with an accuracy of 49.3% and 89.84% at flying heights of 160 m and 140 m, respectively. Ocer *et al.* [17] utilized Mask R-CNN and feature pyramid network (FPN) to extract trees from high-resolution RGB UAV data. Despite scale and content variations, their model retained its high level of accuracy. Yu *et al.* [18] compared several algorithms, including local maxima (LM), marker-controlled watershed segmentation (MCWS), and Mask R-CNN, and the results indicated that Mask R-CNN was the most effective in utilizing the available information, thus producing the most accurate detection results. Given its proven success in detecting and counting palm trees, Mask R-CNN appears suitable for oil palm tree detection and counting tasks. This research aims to integrate UAV data and deep learning algorithms, especially Mask R-CNN, to detect oil palm trees in Indonesia.

2. DATA AND METHOD

2.1. Data

Aerial photography was taken at an altitude of 420 meters above ground level. The total area mapped is approximately 500 hectares. Determination of sidelap and overlap by 70% and 80% produced 186 photos. The process of collecting aerial photo data uses a fixed-wing UAV. An illustration of the UAV flying height can be seen in Figure 1.

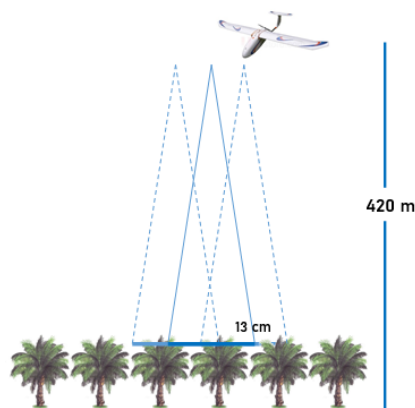


Figure 1. Flying height of fixed wings UAV

The vehicle used is a 1,880 mm wingspan skywalker with vehicle specifications as shown in Table 1. The unmanned vehicle uses a four-cell battery to fly for up to 60 minutes at optimal cruising altitude. In this oil palm mapping flight mission, the UAV used a Sony ILCE-Q1 CMOS type camera with a resolution of 20.1 mega pixels. The type of lens used is a mirrorless camera lens. The full specifications of the camera used can be seen in Table 1.

Table 1. Specification of UAV (left) and camera (right)

SkyWalker 1900 V2.0		Sony ILCE-QX1	
Airframe:	SW 2014 (Wingspan 1880 mm)	Lens	E-mount Sony 16 mm fixed, F2.8 ~
Flight controller:	3DR Pixhawk Cube 2.1	Pixel	20.1 MP
GPS+compass:	3DR ublox M8N	Sensor	Sensor CMOS Exmor
Radio controller:	2.4 Ghz 16 channels	Dimension	74×69.5×52.5 mm
Radio telemetry:	866-915 Mhz 100-1000 mw (adjustable)	Type	Mirrorless
Servo:	Digital servo 2.7 kg/0.13 sec	Sensor optical	APS-C type (23.2×15.4 mm)
Motor:	930 Kv brushless motor	Shutter speed	1/4000 to 30 sec
Sensor:	Digital airspeed sensor		
Battery:	Li-ion 4s5p-4s6p		

The aerial photo acquisition was performed for the entire research area to obtain orthomosaic photos. Apart from that, control point measurements were also carried out, including ground control point (GCP) to increase the geometric accuracy of aerial photos and independent check point (ICP) to test the accuracy of orthomosaic results. The number of control points is eight for GCP and seven for ICP. These control points were measured using geodetic global navigation satellite system (GNSS) equipment with rapid static mode. Next, processing aerial photos using the structure from motion (SfM) method includes the alignment process for each photo, GCP pricking, build dense cloud, build digital elevation model (DEM), and build orthomosaic. Finally, the geometric accuracy test results of orthomosaic photos use the root mean square error (RMSE) value of the ICP point at the 90% confidence level.

2.2. Methods

The deep learning algorithm used in this study is Mask R-CNN, which was developed by He *et al.* [19]. Mask R-CNN model can produce three distinct outputs: class, bounding box, and segment, making it an effective instance segmentation algorithm. Due to its ability to generate segments for each object, the Mask R-CNN model is well-suited for palm tree detection and counting.

The advancement of geographic information system (GIS) technology has paved the way for integrating deep learning models into mapping software. One example is Mask R-CNN, which has been seamlessly incorporated into ArcGIS Pro software. The process for utilizing this technology begins with generating appropriate training data that is compatible with the Mask R-CNN model. It entails using both images and Masks/labels in a shapefile format as input data for the “Export Training Data for Deep Learning” tool, resulting in training data in patch or tile format.

To enhance efficiency, we utilized the template matching tool in the eCognition software instead of manually creating labels. In order to minimize errors during automatic detection with this tool, we restricted the training area solely to the palm oil plantation. This approach reduces the possibility of misclassification due to other objects that may resemble palm trees, such as vegetation or trees around the settlement area. The training area, which is shown in Figure 2, encompasses an area of around 80 hectares.

The images and labels are used as inputs for the “Export Training Data for Deep Learning” tool in ArcGIS Pro software. This tool has several important parameters, including tile size, stride, and metadata format. To create the training dataset, we set the tile size value to 256, resulting in tiles of size 256×256 pixels. The stride parameter was set to 128, which creates an overlap of 50% for subsequent tiles. For this study, we used the Mask R-CNN model; therefore, the metadata format was set to RCNN Masks. Applying these parameters generated 3043 tile images, shown in Figure 3 as an example of the training data.

We utilized the “Train Deep Learning Model” tool to initiate the training process. During this stage, we set the number of epochs to 10. Considering the hardware specifications, we selected a batch size of one. To align with the tile size value, we opted for a chip size 256. Additionally, we selected Resnet-50 as the backbone. Lastly, we set aside 10% of the entire dataset for validation. Next, predictions are made on unseen data or images not used in training.

In this research, we divided the area into three criteria, assuming that the accuracy of the deep learning model can be influenced by cloud shadows and non-palm oil objects, like buildings and other plants. In Figure 4, we can observe a test image containing both an oil palm plantation and a settlement area, and it is also partially obscured by cloud shadows for area criteria 1 (red line). Then, area 2 is only an oil palm plantation area (light blue line), and area 3 is a subset of areas without cloud shadows (yellow line). Areas with extreme cloud shadows are not involved in oil palm detection (dark blue line). The total area used to detect palm oil is 114 hectares, including non-oil palm areas such as settlements, plantations, and roads, 74 hectares for only oil palm plantations, and 20 hectares for subset areas without cloud shadows. Finally, we used the “Detect Objects Using Deep Learning” tool and set the threshold value of 0.6. Additionally, we set the return_bboxes parameter to “False” to obtain segment results rather than bounding boxes.

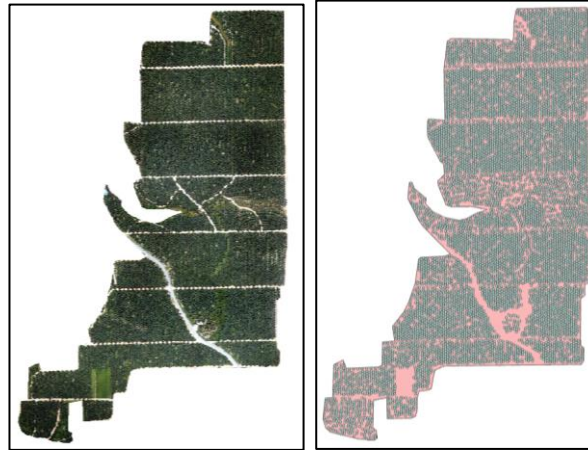


Figure 2. The training area consists of an image on the left and its corresponding label on the right, the training area is restricted solely to the oil palm plantation

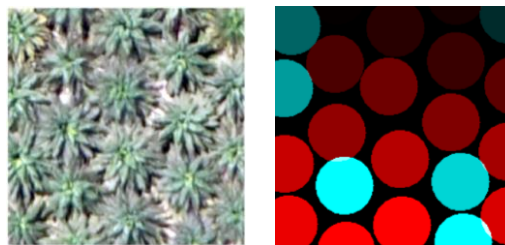


Figure 3. The training data, with the left panel showing the image and the right panel displaying its corresponding label

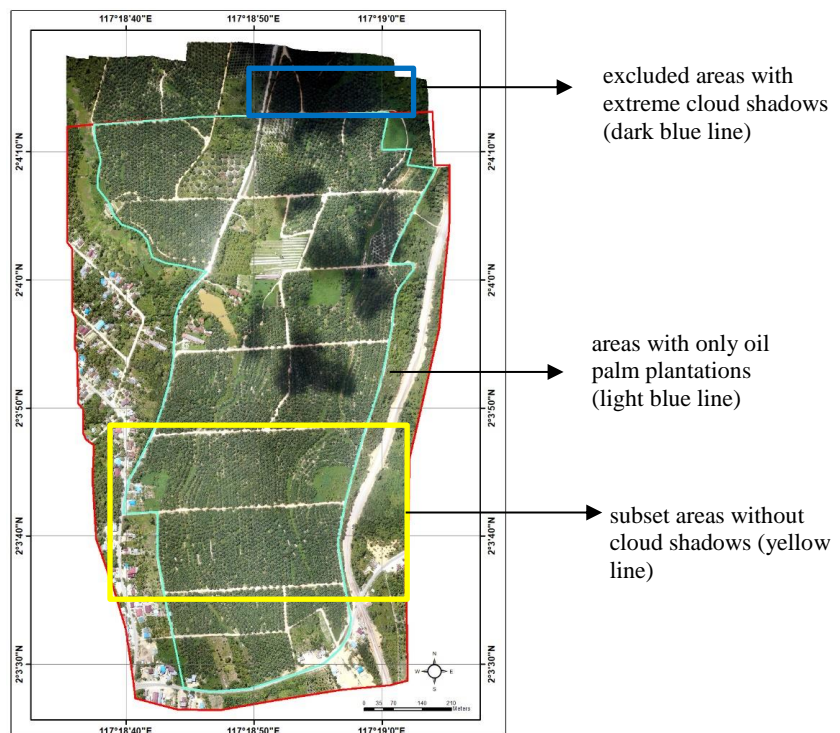


Figure 4. Test image used in the prediction process, this image was used as part of the testing process to evaluate the accuracy of the deep learning model

2.3. Accuracy assessment of oil palm detection

The accuracy of palm tree detection was evaluated by comparing the results of automatic extraction with ground truth values using the precision/recall method [20]. Precision/recall is an assessment method often used in deep learning to assess detection success. The process was conducted manually. The calculation formula for this method can be seen in (1) to (3):

$$Precision = TP / (TP + FP) \tag{1}$$

$$Recall = TP / (TP + FN) \tag{2}$$

$$F - measure = ((1 + \alpha) \times precision \times recall) / (\alpha \times precision + recall) \tag{3}$$

True positive (TP) is the total number of palm oil centroids successfully identified using deep learning, also known as the ground truth value. Meanwhile, false negative (FN) is the number of palm tree centroids that are not detected. False positive (FP) is the number of centroids recognized as oil palm trees, but after validation, they appear as other objects that are not oil palms. The α value is a non-negative scalar of the F-measure calculation, using the threshold of 0.5 [21]. Precision (P) can be interpreted as the amount of noise that can be tolerated during the object identification process, while recall (R) is the number of palm trees that are detected correctly (ground truth). F-measure defines the harmonic mean between precision and recall, where precision and recall are combined into a single performance measure [22].

3. RESULTS AND DISCUSSION

3.1. Orthomosaic result

The results of aerial photo processing produce an orthomosaic with a resolution of 13 cm, which is sufficient for detecting oil palm plants. According to Korom *et al.* [23], oil palm tree crowns are more straightforward to recognize using high-resolution aerial photos. Apart from spatial resolution, another thing that is no less important is the geometric accuracy of the resulting aerial photos. The results of the geometric evaluation show that the horizontal accuracy value using GCP obtained an accuracy of 0.250 meters (can be seen in Table 2).

Table 2. Horizontal accuracy of aerial orthomosaic

ID	Orthomosaic result (m)		GNSS result (m)		DX (Meters)	DY (Meters)	DX ²	DY ²	DX ² +DY ²
	X Ortho	Y Ortho	X_GNSS	Y_GNSS					
ICP1	534,004.513	227,363.575	534,004.713	227,363.597	0.200	0.022	0.04006	0.00050	0.04056
ICP2	534,816.148	226,443.593	534,816.249	226,443.425	0.101	-0.168	0.01013	0.02808	0.03821
ICP3	534,640.117	227,485.993	534,640.251	227,485.995	0.134	0.002	0.01804	0.00001	0.01804
ICP4	534,035.909	226,993.827	534,036.015	226,993.856	0.106	0.029	0.01119	0.00082	0.01201
ICP5	534,296.054	227,769.201	534,296.253	227,769.242	0.199	0.041	0.03946	0.00172	0.04119
ICP6	534,937.667	228,069.383	534,937.804	228,069.337	0.137	-0.046	0.01886	0.00214	0.02099
ICP7	533,938.515	228,126.433	533,938.597	228,126.545	0.082	0.112	0.00672	0.01247	0.01919
								Total (m)	0.190
								Variance	0.027
								STD (m)	0.165
								Accuracy (m)	0.250

Various indicators can be utilized to evaluate the training outcome. One such indicator is the loss graph that illustrates the progress of the training and validation over time, as demonstrated in Figure 5. It can be observed that the loss graph consistently decreases and eventually converges. Another measure of the model's performance is the prediction samples compared with the ground truth. Figure 6 depicts the prediction samples, demonstrating that the model provides accurate predictions.

3.2. Object detection

Figure 7 provides an example of successful segmentation, where the results are generally complete and correct, except for a few areas where several trees were not detected accurately. However, despite the successful results shown in Figure 7, there were still some errors in the model's predictions. For instance, as presented in Figure 8(a) and (b), some non-palm vegetation was incorrectly detected as palm trees, which could lead to inaccurate estimations in applications that rely on precise segmentation. Furthermore, cloud

cover also caused errors in the detection process, as objects covered by cloud shadows were not detected accurately. Clouds alter the energy radiation transmission between sun-surface sensors, making it difficult for object information under clouds to reach sensors accurately [24]. Additionally, its shadows' spectral characteristics are identical to those of wetlands, water, and other ground objects, reducing recognition accuracy beneath cloud shadows [25].

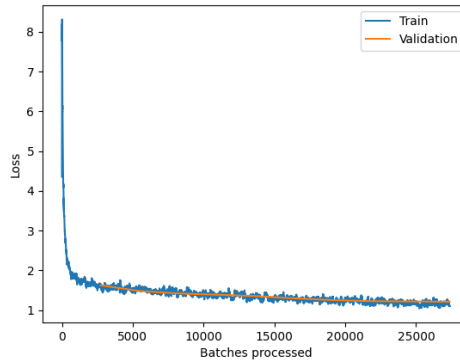


Figure 5. The loss graph for both training and validation data, the y-axis represents the loss values, while the x-axis denotes the number of batches processed

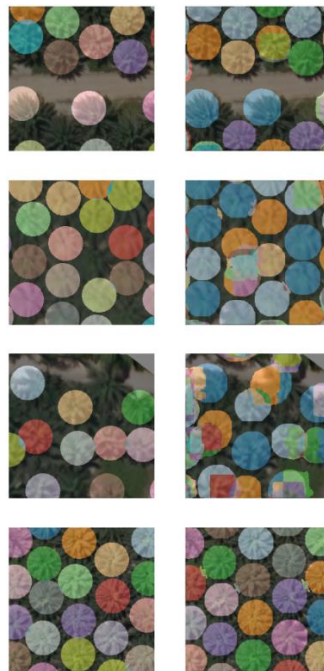


Figure 6. A side-by-side comparison of ground truth (left) and model predictions (right)



Figure 7. The successful prediction results in a palm plantation area, showing that the model can accurately detect and segment objects in complex environments

Deep learning-based palm tree detection in unmanned aerial vehicle imagery with ... (Agung Syetiawan)



Figure 8. The errors resulting from misclassification for: (a) non-palm vegetation was incorrectly identified as palm trees and (b) the errors caused by cloud shadow

3.3. Accuracy assessment

Figures 9 and 10 show the overall prediction results for palm oil detection using Mask R-CNN in areas 1 and 3, respectively. In area 1, we succeeded in detecting 5,780 centroids, while the FN and FP values were 1,968 and 916 centroids, respectively. Then, area 2 with 5,706, 1,820, and 319 centroids on TP, FN, and FP, respectively. Finally, area 3 with 1,729, 466, and 38 centroids on TP, FN, and FP, respectively. Table 3 shows the results of palm oil detection accuracy on various criteria. Area 1 produces a detection accuracy of 82.02%, area 2 is 87.43%, and area 3 is 90.52%. Table 3 indicates that oil palm trees are best detected in area 3 with an accuracy of 90.52%. The segments were accurately produced in areas that were not obstructed by clouds.

2

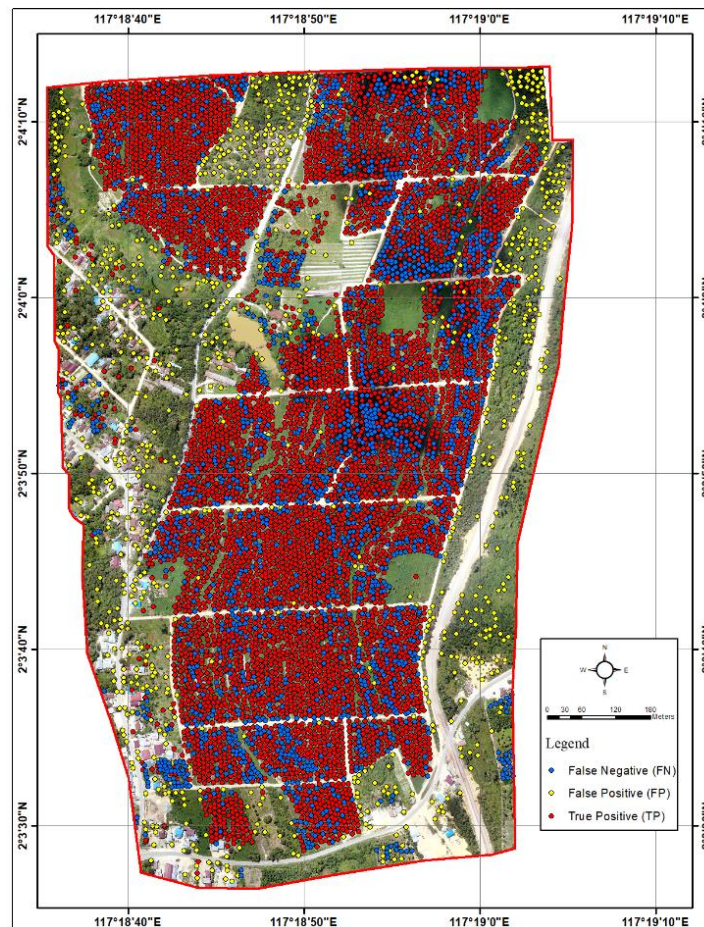


Figure 9. Detection results for oil palm plantation and a settlement area

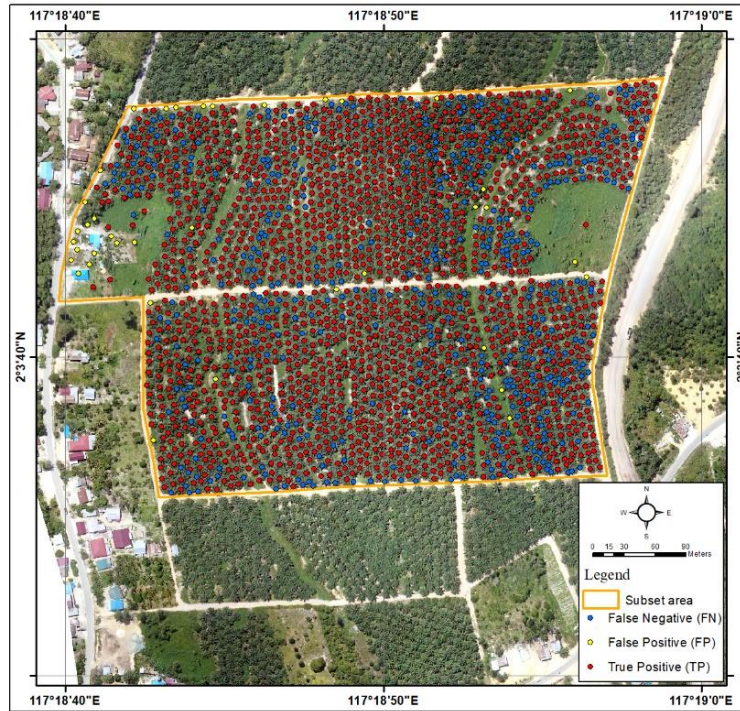


Figure 10. Detection results for subset areas without cloud shadows

Table 3. Automatic detection of oil palm trees at various criteria

No.	Criteria	Coverage (Ha)	Actual oil palm tree	Detected oil palm tree	Accuracy (%)
1.	Area 1	114	7,748	5,780	82.02
2.	Area 2	74	7,526	5,706	87.43
3.	Area 3	20	2,195	1,729	90.52

4. CONCLUSION

As the most significant economic plantation in Indonesia, oil palm monitoring techniques should be developed for efficient operation. In this study, we used a geospatial technique for data acquisition and a computer vision processing approach with deep learning Mask R-CNN to assess the ability of high-resolution images obtained from the UAV to detect oil palm trees. The results indicate that Mask R-CNN can distinguish oil palm trees from other plants, settlements, and road segments with more than 80% accuracy. Incorrect detection was found mainly in cloud shadow areas due to distinctive spectral values of oil palm trees compared to non-shadow areas. This observation was supported by the detection results excluding shadow areas, where an accuracy of 90.52% was achieved. An effort incorporating image preprocessing to enhance oil palm spectral in shadowed areas while preserving its spectral characteristics may be studied further for accuracy improvement. Overall, the obtained accuracy of oil palm tree detection using Mask R-CNN in this study is sufficient, thus becoming a promising technique to detect palm oil over a large area using high-resolution images from UAV.

ACKNOWLEDGEMENTS

This study is financially funded by the Agriculture seed grant under the Research Organization of Agricultural and Food, National Research and Innovation Agency of Indonesia (BRIN) with grant number 29/III.11/HK/2022. The authors also thank the Geospatial Information Agency for the fixed-wing UAV used during the acquisition of this research.

REFERENCES

[1] X. J. Tan, W. L. Cheor, K. S. Yeo, and W. Z. Leow, "Expert systems in oil palm precision agriculture: A decade systematic review," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 4, pp. 1569–1594, 2022, doi: 10.1016/j.jksuci.2022.02.006.




[2] H. Varkkey, "The Growth and Prospects for the Oil Palm Plantation Industry in Indonesia," *Oil palm industry economic journal*,

Deep learning-based palm tree detection in unmanned aerial vehicle imagery with ... (Agung Syetiawan)




- vol. 12, no. No.2, pp. 1–13, 2012.
- [3] J. S. H. Lee, S. Abood, J. Ghazoul, B. Barus, K. Obidzinski, and L. P. Koh, "Environmental Impacts of Large-Scale Oil Palm Enterprises Exceed that of Smallholdings in Indonesia," *Conservation letters*, vol. 7, no. 1, pp. 25–33, 2014, doi: 10.1111/conl.12039.
 - [4] S. A. Abood, J. S. H. Lee, Z. Burivalova, J. Garcia-Ulloa, and L. P. Koh, "Relative Contributions of the Logging, Fiber, Oil Palm, and Mining Industries to Forest Loss in Indonesia," *Conservation letters*, vol. 8, no. 1, pp. 58–67, 2015, doi: 10.1111/conl.12103.
 - [5] E. Hambali and M. Rivai, "The Potential of Palm Oil Waste Biomass in Indonesia in 2020 and 2030," in *IOP Conference Series: Earth and Environmental Science*, 2017, p. 012050. doi: 10.1088/1755-1315/65/1/012050.
 - [6] A. Syetawan, H. Gularso, G. I. Kusnadi, and G. N. Pramudita, "Precise topographic mapping using direct georeferencing in UAV," in *IOP Conference Series: Earth and Environmental Science*, 2020, p. 012029. doi: 10.1088/1755-1315/500/1/012029.
 - [7] C. A. Rokhmana, "The potential of UAV-based remote sensing for supporting precision agriculture in Indonesia," *Procedia Environmental Sciences*, vol. 24, pp. 245 – 253, 2015, doi: 10.1016/j.proenv.2015.03.032.
 - [8] J. Zheng *et al.*, "Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images," *ISPRS J. Photogramm. Remote Sensing*, vol. 173, pp. 95–121, 2021, doi: 10.1016/j.isprsjprs.2021.01.008.
 - [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in neural information processing systems*, vol. 25, 2012, doi: 10.1201/9781420010749.
 - [10] T. Hoeser and C. Kuenzer, "Object detection and image segmentation with deep learning on Earth observation data: A review-part I: Evolution and recent trends," *Remote Sensing*, vol. 12, no. 10, 2020, doi: 10.3390/rs12101667.
 - [11] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications : A meta-analysis and review," *ISPRS journal of photogrammetry and remote sensing*, vol. 152, pp. 166–177, 2019, doi: 10.1016/j.isprsjprs.2019.04.015.
 - [12] N. A. Mubin, E. Nadarajoo, H. Z. M. Shaffri, and A. Hamedianfar, "Young and mature oil palm tree detection and counting using convolutional neural network deep learning method," *International Journal of Remote Sensing*, vol. 40, no. 19, pp. 7500–7515, 2019, doi: 10.1080/01431161.2019.1569282.
 - [13] M. B. A. Gibril *et al.*, "Deep convolutional neural networks and Swin transformer-based frameworks for individual date palm tree detection and mapping from large-scale UAV images," *Geocarto International*, pp. 1–31, 2022, doi: 10.1080/10106049.2022.2142966.
 - [14] W. Li, R. Dong, H. Fu, and L. Yu, "Large-Scale Oil Palm Tree Detection from High-Resolution Satellite Images Using Two-Stage Convolutional Neural Networks," *Remote Sensing*, vol. 11, no. 1, 2019, doi: 10.3390/rs11010011.
 - [15] J. Ribera, Y. Chen, C. Boomsma, and E. J. Delp, "Counting Plants Using Deep Learning," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Montreal, Canada, 2017, pp. 1344–1348, doi: 10.1109/GlobalSIP.2017.8309180.
 - [16] K. Yarak, A. Witayangkurn, K. Kritiyutanont, C. Arunplod, and R. Shibusaki, "Oil Palm Tree Detection and Health Classification on High-Resolution Imagery Using Deep Learning," *Agriculture*, vol. 11, no. 2, 2021, doi: 10.3390/agriculture11020183.
 - [17] N. E. Ocer, G. Kaplan, F. Erdem, D. K. Matci, and U. Avdan, "Tree extraction from multi-scale UAV images using Mask R-CNN with FPN," *Remote Sensing Letters*, vol. 11, no. 9, pp. 847–856, 2020, doi: 10.1080/2150704X.2020.1784491.
 - [18] K. Yu *et al.*, "Comparison of Classical Methods and Mask R-CNN for Automatic Tree Detection and Mapping Using UAV Imagery," *Remote Sensing*, vol. 14, 2022, doi: 10.3390/rs14020295.
 - [19] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 1 Feb. 2020, doi: 10.1109/TPAMI.2018.2844175.
 - [20] R. D. Nyland, L. S. Kenefic, K. K. Bohn, and S. L. Stout, *Silviculture: Concepts and Applications*, Third Edit. Illinois: Waveland Press, 2016.
 - [21] D. R. Martin, C. C. Fowlkes and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, May 2004, doi: 10.1109/TPAMI.2004.1273918.
 - [22] B. Kalantar, M. O. Idrees, S. Mansor, and A. A. Halin, "Smart Counting – Oil Palm tree inventory with UAV," *Coordinates*, vol. 13, no. 5, pp. 17–22, 2017.
 - [23] A. Korom, M. H. Phua, Y. Hirata, and T. Matsuura, "Extracting oil palm crown from WorldView 2 satellite image," in *IOP Conference Series: Earth and Environmental Science*, 2014, p. 012188. doi: 10.1088/1755-1315/18/1/012188.
 - [24] J. Zhang, Q. Zhou, X. Shen, and Y. Li, "Cloud Detection in High-Resolution Remote Sensing Images Using Multi-features of Ground Objects," *Journal of Geovisualization and Spatial Analysis*, vol. 3, no. 2, p. 14, 2019, doi: 10.1007/s41651-019-0037-y.
 - [25] Y. Zhang, C. Ye, R. Yang, and K. Li, "Reconstructing Snow Cover under Clouds and Cloud Shadows by Combining Sentinel-2 and Landsat 8 Images in a Mountainous Region," *Remote Sensing*, vol. 16, no. 1. 2024, doi: 10.3390/rs16010188.

BIOGRAPHIES OF AUTHORS






Agung Syetawan    is a researcher in the National Research and Innovation Agency of Indonesia. He received a Master's Degree in Geodesy and Geomatics Engineering from Institut Teknologi Bandung, Indonesia. He has published over 15 journal papers, 1 authored book, and 31 papers in conference proceedings. His research activities include oceanography, UAV, and GNSS processing. He also has experience in accurate UAV for disaster. Currently, his research uses geodetic data for coastal flooding and land subsidence modeling in north coastal Java, Indonesia. He can be contacted at email: agungsyetawan@gmail.com.






Danang Budi Susetyo    is a researcher at the National Research and Innovation Agency (BRIN) in Indonesia. He holds a Master of Engineering degree from the Geomatics Engineering Program, Universitas Gadjah Mada, Indonesia. His research interests are related to the processing and applications of photogrammetric and remote sensing data, particularly for mapping of urban area. Currently, he is pursuing a Doctoral study at Yildiz Technical University in Turkey, specifically in the Department of Geomatic Engineering. He can be contacted at email: danang.budi.susetyo@brin.go.id.






Yustisi Lumban-Gaol    received the M.Sc. degree (cum laude) in Geomatics from Delft University of Technology, Delft, The Netherlands, in 2021, where she is currently pursuing a Ph.D. degree in satellite radar interferometry. From 2022 to 2023, she actively worked with the National Research and Innovation Agency (BRIN) in Indonesia as a Researcher, where she was involved in land subsidence studies and slum mapping projects supporting Sustainable Development Goals. She can be contacted at email: yustisiardhitasari@gmail.com.






Susilo    is a Senior Researcher at Research Center for Geological Disaster, National Research and Innovation Agency of Indonesia (BRIN). He has experience using GPS observation to study tectonic deformation, Indonesian Geospatial Reference Frame, and much of his recent work has involved analysis and tectonic modeling of continuous data from the Indonesian CORS and the Sumatra GPS Array. He also has experience in using kinematic GPS data for studying earthquake magnitude estimation. His recent study is using continuous and campaign GPS data for land subsidence in Indonesia. He can be contacted at email: susilosarimun@gmail.com.






Mohammad Ardha    is a researcher at the National Research and innovation agency (BRIN) in Indonesia. He holds an undergraduate of geography from Gadjah Mada university, Indonesia. His research focuses on flood modeling in the coastal area particularly in north coastal Java. He can be contacted at email: moha054@brin.go.id.



Yunus Susilo    is a Lecture at Geomatics Engineering Dr. Soetomo University, which has experience using GPS observation to study Land Deformation and expert as border line village. He can be contacted at email: yunus.susilo@unitomo.ac.id



Wahono    is a researcher in the Agrotechnology Department at the Faculty of Agriculture and Animal Science, University of Muhammadiyah Malang (UMM), located in Malang, East Java, Indonesia. He possesses expertise in utilizing UAVs and remote sensing for the analysis of plant health and agronomic recommendations. His current study used unmanned aerial vehicles equipped with multispectral sensors to assess nitrogen requirements in tea plantations. He can be contacted at email: wahono@umm.ac.id.