International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Estimating Oil Palm Tree Yield and Soil Composition Using Multi-Scale CNN and Vegetation Indices

Edy Irwansyah [#1*], Alexander Agung Santoso Gunawan[#2], and Izzi Dzikri [#3]

Submitted: 28/10/2022 **Revised:** 20/12/2022 **Accepted**: 01/01/2023

Abstract: Indonesia is one of the countries in the world with the largest oil palm plantations besides Malaysia which is administratively a neighbour. The process of monitoring oil palm plantation is an important part of supporting the sustainability of the industry in a fluctuating market. Currently, monitoring of oil palm plantations is still using conventional methods through surveys at the plantation location which take a long time. The aim of the study was to count the number and assess the health level of oil palm trees and soil composition to assist plantation managers in making decisions to provide maximum economic benefits. By implementing the Multi-Scale CNN model, oil palm trees can be detected faster and accurately, where the combination of land cover and object classification architecture of ResNet-18 can produce accuracy with an average F1-score of 90.25%. Then by developing a prototype that integrates the Multi-Scale CNN model and the vegetation index, the yields and soil composition can be estimated, thus assisting the user in making a quick decision.

Keywords: *Convolutional Neural Network, yield estimation, vegetation index, remote sensing*

1. Introduction

Palm oil is one of the most widely used multipurpose oils. From products like soaps and lipsticks to biscuits and pizzas that contain palm oil. Reported from [1] the increase in demand for palm oil, production of palm oil has increased during the last two decades in Southeast Asia, especially in Malaysia and Indonesia. The two countries have become major producers and exporters of palm oil which have contributed more than 85% of world palm oil production, which is expected to reach 84 million by 2020. The massive expansion of oil palm has contributed to land use change with far-reaching environmental and socio-economic consequences.

The oil palm tree is characterized by its visible crown and single stemmed trunk. Fronds that emerge from the trunk apex and are extended outward spirally with eight fronds forming a rank in succession. Oil palm is a perennial crop, which resembles a forest tree more than any other agricultural crop [2]. As an industrial crop, oil palm is grown in monoculture. Most of the oil palm trees use in commercial plantations are tissue cultures clones with a small mix of hybrid oil palms (i.e., Dura X Pisifera), which gives the palms a uniform appearance except for the anomalous trees. This unique pattern makes the tree easy to distinguish from other forest in the satellite [3].

For economic and practical reasons, oil palm plantations can be divided into blocks / fields which include immature and mature planting to facilitate the implementation of planting operations. Immature oil palm is oil palm that is still young and no more than 5 years after planting in the field. Seen in figure 1, the trees are planted in a triangular pattern with a spacing of nine meters which is the industry standard to maximize yields with optimal sunlight [4]. In general, the planting density of oil palm trees that is practiced reach 130-140 trees per ha, but this number can also vary according to planting conditions and types of seeds [5].

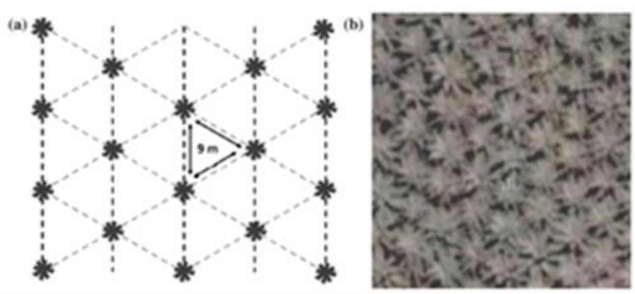


Fig 1. Oil palm tree planting patterns [6]

The development of the oil palm industry can cause environmental degradation, especially when plantation expansion is out of control. Remote sensing technology can provide timely, repeated information with a broad coverage and is more cost-effective than survey methods, thus makes

Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11480
[1]*ORCID ID: 0000-0002-3876-1943*
[2]*ORCID ID: 0000-0002-1097-5173*
[3]*ORCID ID: -*
**Corresponding Author Email: eirwansyah@binus.edu*

this technology an important part of supporting the sustainability of the palm oil industry [7]. Due to lack of collaboration between academics and industry, it hampers the development to utilize a more effective remote sensing technology [8]. Therefore, until today there are still many industries that rely on survey methods to monitor their plantations.

Based on [6] as a commodity with a fluctuating market, oil palm yields must be estimated to produce maximum economic benefits, then the composition of oil palm soil is also an important factor that can affect oil production. Yields can be affected by the health of the oil palm tree, where the composition of the soil can be used as an indication the oil palm tree will lack nutrients. Soil composition such as nitrogen (N) and potassium (K) are important nutrients in the growth of oil palm trees, where the identification of nutrient deficiencies is one of the important variables in oil palm tree production. The amount of nitrogen affects the growth of oil palms while potassium affects size, number, and the prevention of disease disorders of fruit bunches. Oil palm tree detection is one of the techniques needed to produce accurate estimation results.

Tree counting is an important and necessary practice for yield estimation and monitoring, replanting and layout planning, etc. Conventionally this practice is expensive and prone to human error, as most plantations are forced to estimate the number of trees by multiplying the total area by the number of oil palms per hectare, which is clearly inaccurate due to the heterogeneity of land surface (hill or flat) and features (river, land, or Forest). Therefore, remote sensing is used to solve this problem because it provides an overview of the plantation and how to calculate trees can also be done automatically with a deep learning approach. Based on [9] the multi scale sliding windows method by combining 2 convolutional neural networks can detect oil palm trees well. In his research, the combination of the AlexNet architecture used for object classification and land cover obtained the best results with an average F1-score of 94.99%. The implementation of multi scale sliding windows is a continuation from the previous method that only using 1 sliding window to detect oil palm tree which still prone to misclassification in other vegetation and bare land areas [10].

Based on the above problems, we would like to develop an application that can provide information about the estimation of yield and soil composition more easily. To get good detection results, the authors implemented the Multi-Scale CNN method on the Pléiades satellite image. Because the datasets were made manually, the author need to test the combination of the CNN architecture between LeNet-5, AlexNet, and ResNet-18. This is done to test previous research methods with different datasets and evaluate their accuracy with other architectures. In this study we decided to develop web-based application to minimize computing cost on local hardware.

2. Methodology

This research methodology is split into 5 main parts based on the prototyping model, and these stages consist of:

- Communication
- Quick Plan
- Modeling Quick Design
- Construction
- Delivery and Feedback

2.1. Communication

This research begins by identifying the problem first in which, as mentioned in section 1, this research focuses on developing applications to estimate crop yields and soil composition of oil palm tree by implementing Multi-Scale CNN and vegetation indices. But before starting conducting the research surely the author needs to do a literature study first to enrich insights on how to develop a Multi-Scale CNN method and vegetation index which is then integrated into the application.

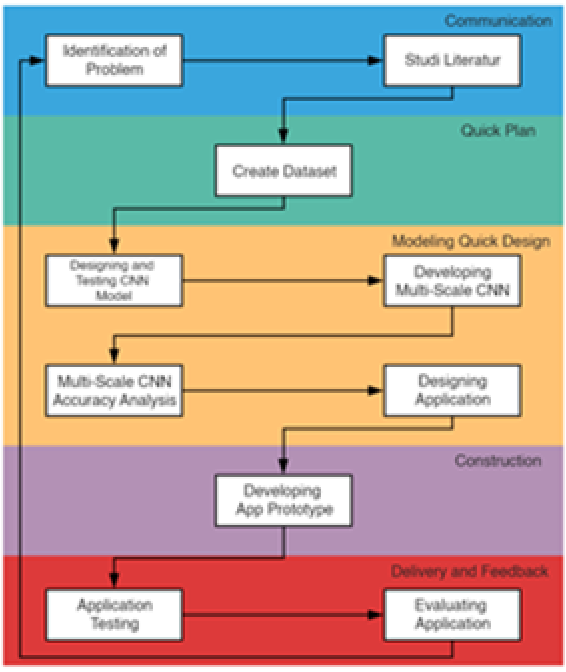


Fig 2. Research Methodology

2.2. Quick Plan

Then at the next stage the authors made a dataset that was acquired from one of the oil palm plantations in Prov. West Kalimantan and processed in advance by PT Citra Bhumi Indonesia. This image was acquired on December 9, 2018, using the Pléiades satellite, which uses a spatial resolution of 0.5 meters which based on [11] can be interpreted that each image pixel represents 0.5 × 0.5 meter and 4 spectral bands (RGB and NIR).

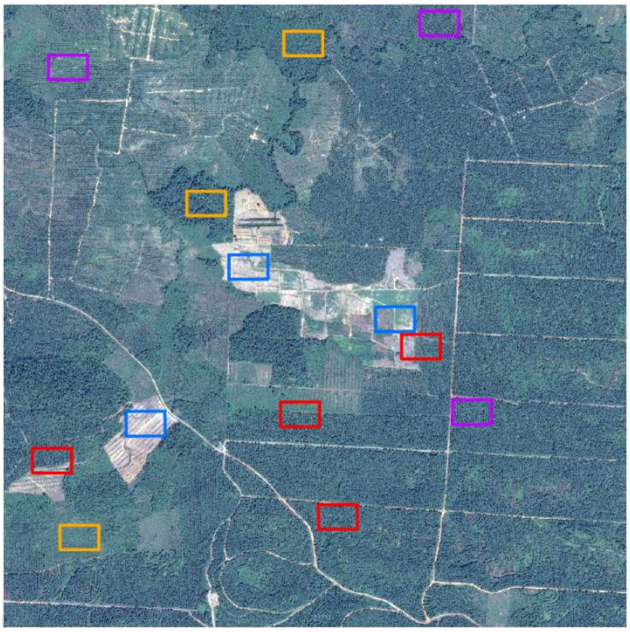
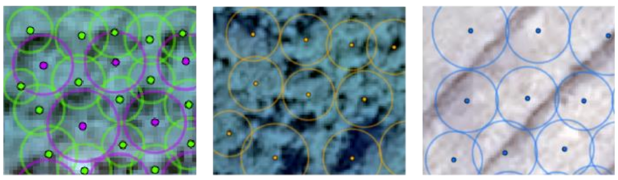


Fig. 3. Sampling Locations on Satellite Images

Shown in Figure 3 are some of the areas that are picked by the author as the sample dataset. The sample areas shown in the purple, yellow, and blue colored boxes are used to train the CNN model and the red boxes are used to evaluate the Multi-Scale CNN. Purple squares show areas containing oil palm trees, yellow boxes show other areas of vegetation and blue boxes show empty land. Basically, the labeling of the dataset can be divided into 2 types, namely land cover and object, where each serve to classify images of oil palm plantations with size 65 x 65 pixels and images of oil palm trees with sized 17 x 17 pixels. Specifically, for the labeling of objects or oil palm trees, the author determines an image labeled as an oil palm tree when the tree is in the center of the image, otherwise the image will be considered a background and here is an image that shows an example of labeling given.



(a) (b) (c)

Fig. 4. Examples of labeling object classification samples

Seen in Figure 4 (a) is an example of labeling oil palm trees and the background, respectively shown in purple and green, then Figures 4 (b) and (c) show examples of labeling other vegetation and bare land. From the manually labeled dataset, around 6,000 images were collected for land cover classification and about 8,000 images for object classification.

The author uses 10% of the dataset as testing data to evaluate the accuracy of classification model, and the rest is used as training data. In the training data, 10% of the data is also used as validation data that plays a role in tuning the hyperparameter. Then the authors use the red areas that already have the oil palm tree label to evaluate the accuracy of the Multi-Scale CNN, in detecting oil palm trees.

2.3. Modelling Quick Design

In this study, the authors developed and tested 3 CNN model architectures consisting of LeNet-5, AlexNet, and ResNet-18. The following figure shows a comparison of each architecture where LeNet-5 and ResNet-18 use the same parameters to classify land covers and objects, while AlexNet uses different parameters adjusted for research [9].

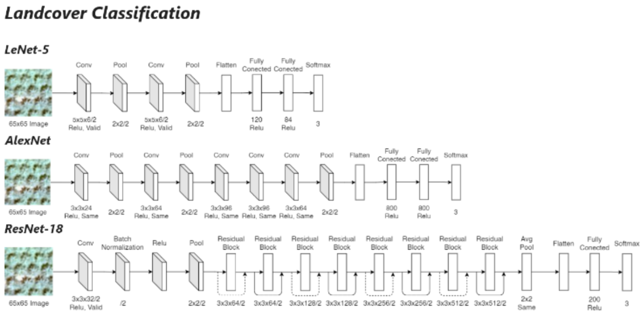


Fig. 5. CNN Architecture for the Land cover Classification

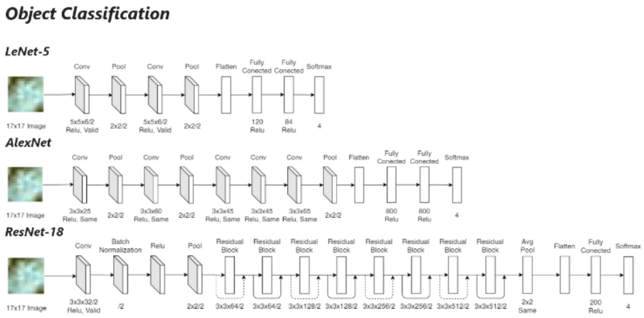


Fig. 6. CNN Architecture for Object Classification

Seen in figures 5 and 6 the ResNet-18 architecture implements a Residual Block that is different from other architectures. This is applied to overcome the vanishing gradient caused by adding a deeper layer of CNN [12]. In the following figure, you can see the detailed contents of the Residual Block applied to the ResNet-18 architecture, the image also depicts a shortcut with a dotted line in Figures 5 and 6, which is used to indicate which Residual Block that uses 1 × 1 convolution in its shortcut route to adjust the dimensions.

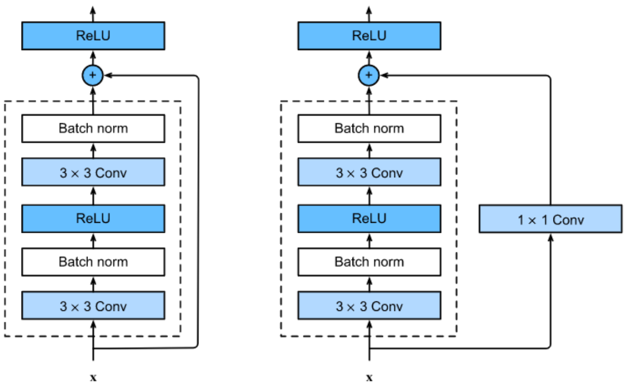


Fig. 7. Residual Block with and without 1 × 1 convolution [13]

After the CNN model has been developed and tested, the authors began to develop the Multi-Scale CNN to detect oil palm trees. At this stage, the author also implements agglomerative clustering with average linkage [14] as post-processing to collect all detection results that have been found by Multi-Scale CNN.

The next stage is continued by analyzing the Multi-Scale CNN using the 4 evaluation areas specified in section 2.2. The author compared the detection results and ground truth that were manually labeled in the 4 regions. The F1-score value is used to compare the accuracy of the model in detecting oil palm trees, where True Positive (TP) shows the number of oil palm trees that are detected correctly, False Positive (FP) shows the number of other objects detected as oil palm trees, and False Negative (FN) indicates the number of undetected trees. Oil palm trees will be considered correct if the Euclidean distance between the ground truth and the detection results is less than or equal to 5 pixels. The combination of CNN 1 (land cover classification) and CNN 2 (object classification) with the highest average F1-score will be used in the application.

After the combination of the CNN model with the F1-score average has been obtained, the writer can begin to design how to integrate the Multi-Scale CNN with the vegetation index to estimate yield and soil composition.

2.4. Construction

In this stage, the authors begin to develop a prototype application to integrate Multi-Scale CNN and vegetation indexes, using the flask framework as the backend of the website. At this stage, the author also adds various validations to avoid errors in the application.

2.5. Delivery and Feedback

In the next stage the writer will carry out functional testing of the prototype results that have been developed using the black box testing method. After that, the author evaluates the application in terms of User Interface based on 8 golden rules and checks the suitability of the application against user requirements with an acceptance test.

3. Results and Discussion

Based on the results of the CNN training model, the following graph shows the performance of the training model in classifying 80% of training data and 10% of validation data. Then for the results of the model evaluation of 10% testing data can also be seen in the following table.

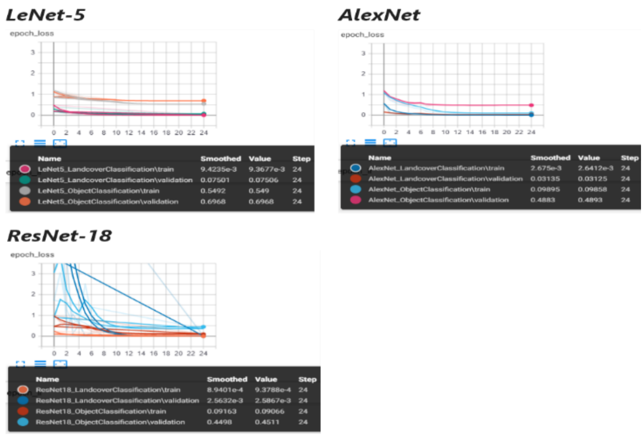


Fig. 8. Loss per epoch for each CNN model

Table 1: The accuracy value of the CNN model to the test dataset

CNN Models	Test Loss	Test Accuracy
LeNet-5 Land cover	0.05	98.33%
LeNet-5 Objek	0.7	69.83%
AlexNet Land cover	0.02	99.39%
AlexNet Objek	0.4	85.97%
ResNet-18 Land cover	**0.0005**	**100%**
***ResNet-18* Objek**	**0.43**	**87.73%**

It can be seen in Figure 8 that the loss value in training data and validation decreases during model training while having a low distance on each value, thus giving an indication that the model has reached generalization and low bias. Then based on table 1 it can also be concluded that the CNN ResNet-18 model is the best model in classifying oil palm plantations and oil palm trees.

It can be seen from the following table that the combination of the CNN ResNet-18 model in classifying land cover and objects has the highest average F1-score, which indicates high accuracy in detecting oil palm trees from the 4 evaluation areas.

Table 2: F1-score values for each CNN architecture from the 4 evaluation regions

CNN 1	*CNN* 2	Region 1	Region 2	Region 3	Region 4	*Avg*
LeNet-5	LeNet-5	78.23 %	82.87 %	54.97 %	89.37 %	76.36 %

LeNet-5	AlexNet	82.03 %	83.74 %	61.13 %	89.64 %	79.14 %
LeNet-5	ResNet-18	81.51 %	84.21 %	64.43 %	88.83 %	79.75 %
AlexNet	LeNet-5	83.28 %	94.62 %	73.93 %	94.12 %	86.49 %
AlexNet	AlexNet	85.63 %	**95.27 %**	75.52 %	**94.47 %**	87.72 %
AlexNet	ResNet-18	86.06 %	93.09 %	79.13 %	94.22 %	88.13 %
ResNet-18	LeNet-5	85.30 %	93.29 %	80.15 %	92.87 %	87.90 %
ResNet-18	AlexNet	87.21 %	93.95 %	86.61 %	92.63 %	90.10 %
ResNet-18	**ResNet-18**	**89.29 %**	92.15 %	**87.70 %**	91.84 %	**90.25 %**

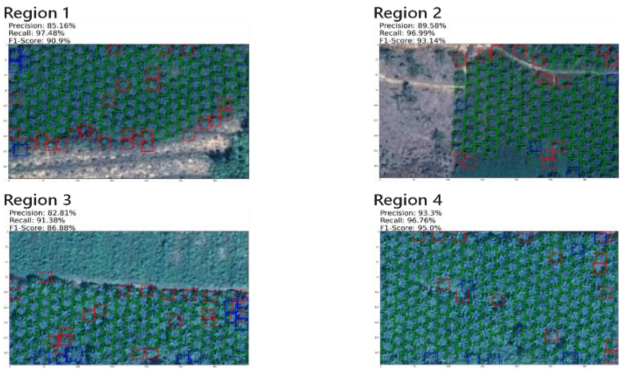


Fig. 9. Detection Results of Oil Palm Trees in the Evaluation Area

It can be seen in Figure 9 that the precision, recall, and F1-Score values are from the CNN Multi-Scale detection results for the combination of the ResNet-18 architecture on land cover and objects. Precision is the ratio of the model results to detect oil palm trees to the total of all objects detected by the model as oil palm trees. Meanwhile, recall measures the ratio of the model results to detect oil palm trees with the total labels detected and not. Precision and recall values can be averaged into the harmonic mean, where the average is also known as the F1-Score. Based on table 2 and figure 9, it can be concluded that the combination of ResNet-18 architecture on land cover and objects can detect oil palm trees very well when compared to other models. The minimum value of 82.81% precision and 91.38% recall given by the model indicates that the prediction model of the model has low errors in classification and undetected objects.

To estimate the yield and soil composition, the bounding box that has been obtained from the object detection results is converted into a circle with a diameter of 17 pixels first, then the results of the average vegetation index value for each tree can be calculated, after that by implementing the following equations which obtained from [15] and [16] research, yields and soil composition especially nitrogen and potassium substances can be estimated with.

$$yield\_estimation = \big(0.6731 + (0.0061 \times rvi\_transform)\big) \times 0.1$$

$$nitrogen = \frac{ndvi + 0.2321}{383.27} \times 100$$

$$pottasium = \frac{ndvi - 0.236}{0.4971}$$

Some of the equations above implement the results of the research that has been carried out but there are some equations that are converted to adjust the information provided to the user such as the conversion to kg/m^2 of the yield estimation and the percentage of the nitrogen. After the results of the estimation of soil composition are obtained, the value is classified based on [17] research, and the percentage of the entire classification is represented in a pie chart where detailed information can be seen in the following table.

Table 3: Classification Of Estimated Soil Composition [17]

Nutrient	Very Low	Low	Moderate	High	Very High
Total N (%)	< 0.08	0.08-0.12	0.12-0.15	0.15-0.25	> 0.25
K (cmol/kg)	< 0.08	0.08-0.20	0.20-0.25	0.25-0.30	> 0.30

Table 4 : Percentage description of the pie chart

Class	Description
Healthy	When nitrogen and potassium value fall into high – very high classification
Nitrogen deficiencies	When nitrogen value falls into moderate - very low classification
Potassium deficiencies	When potassium value falls into moderate - very low classification
Nitrogen and Potassium deficiencies	When nitrogen and potassium values fall into moderate - very low classification

After the yield estimates and soil composition have been obtained, the authors conducted the following experiments to ensure the consistency of the data which can be seen in the following table.

Table 5: Estimated Value of Yield and Soil Composition per Experiment

Image Name	Experiment	number of trees	Average Estimated Yield in blocks	Percentage classification of soil composition
Evaluasi_1.tif	1st	182	12.3 ton/ha/thn	100% healthy
K (cmol/kg)	2nd	182	12.3 ton/ha/thn	100% healthy
	3rd	182	12.3 ton/ha/thn	100% healthy
	4th	182	12.3 ton/ha/thn	100% healthy
	5th	182	12.3 ton/ha/thn	100% healthy

Based on table 5 it can be concluded that with the same input the model gives the same results, thus indicating that the data that are generated by the application is consistent.

4. Conclusion

Based on the results of the research and analysis conducted by the author, it can be concluded that: Oil palm trees can be detected accurately by using the Multi-Scale CNN method. From the experimental results by combining land cover and object classification architectures, the ResNet-18 architecture used in both classifiers resulted in the most accurate detection of oil palm trees with an average F1-score of 90.25%. Yield and soil composition for each oil palm tree can be estimated using a vegetation index. Estimates based on implementation of linear regression models from the previous research. Based on the results of the interviews conducted, it can be concluded that the application prototype can help the performance of the task compared to conventional methods.

Author contributions

Edy Irwansyah: Methodology, Verifying and Correcting Original Draft, Validate Deep Learning Results. **Alexander Agung Santoso Gunawan:** Methodology, Verifying and Correcting Original draft, Validate Deep Learning Results, **Izzi Dzkiri:** Writing Original Draft Preparation, Implementation of Deep Learning Models.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] M. Euler, V. Krishna, S. Schwarze, H. Siregar, and M. Qaim, "Oil Palm Adoption, Household Welfare, and Nutrition Among Smallholder Farmers in Indonesia," *World Dev*, vol. 93, pp. 219–235, May 2017, doi: 10.1016/j.worlddev.2016.12.019.

[2] G. M. Foody *et al.*, "Mapping the biomass of Bornean tropical rain forest from remotely sensed data," *Global Ecology and Biogeography*, vol. 10, no. 4, pp. 379–387, 2001, doi: 10.1046/j.1466-822X.2001.00248. x.

[3] H. Z. M. Shafri, N. Hamdan, and M. I. Saripan, "Semi-automatic detection and counting of oil palm trees from high spatial resolution airborne imagery," *Int J Remote Sens*, vol. 32, no. 8, pp. 2095–2115, 2011, doi: 10.1080/01431161003662928.

[4] Y. Basiron, "Palm oil production through sustainable plantations," *European Journal of Lipid Science and Technology*, vol. 109, no. 4, pp. 289–295, 2007, doi: 10.1002/ejlt.200600223.

[5] Corley; R.H.V. and Tinker; P.B., "The Oil Palm: Fifth Edition," *Blackwell Science*, vol. 5th, pp. 1–680, 2015, doi: 10.1002/9781118953297.

[6] K. L. Chong, K. D. Kanniah, C. Pohl, and K. P. Tan, "A review of remote sensing applications for oil palm studies," *Geo-Spatial Information Science*, vol. 20, no. 2, pp. 184–200, Apr. 2017, doi: 10.1080/10095020.2017.1337317.

[7] United Nations Environment Programme, "Oil Palm Plantations: Threats and Opportunities for Tropical Ecosystems," *UNEP Global Environmental Alert Service (GEAS)*, 2011. https://wedocs.unep.org/20.500.11822/32263 (accessed Dec. 28, 2022).

[8] S. B. Hansen, R. Padfield, K. Syayuti, S. Evers, Z. Zakariah, and S. Mastura, "Trends in global palm oil sustainability research," *J Clean Prod*, vol. 100, pp. 140–149, Aug. 2015, doi: 10.1016/j.jclepro.2015.03.051.

[9] 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 Sens (Basel)*, vol. 11, no. 1, Jan. 2019, doi: 10.3390/rs11010011.

[10] W. Li, H. Fu, L. Yu, and A. Cracknell, "Deep learning-based oil palm tree detection and counting for high-resolution remote sensing images," *Remote Sens (Basel)*, vol. 9, no. 1, 2017, doi: 10.3390/rs9010022.

[11] Government of Canada, "Spatial Resolution, Pixel Size, and Scale," *Natural Resources Canada*, 2015. https://www.nrcan.gc.ca/maps-tools-and-publications/satellite-imagery-and-air-photos/tutorial-fundamentals-remote-sensing/satellites-and-sensor

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. [Online]. Available: http://image-net.org/challenges/LSVRC/2015/

[13] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*, vol. v3. 2022. doi: 10.48550/ARXIV.2106.11342.

[14] J. Han, M. Kamber, and J. Pei, *Data mining: concepts and techniques*, vol. 4th. 2022.

[15] S. K. Balasundram, H. Memarian, R. Khosla, and R. Khosla, “Estimating Oil Palm Yields using Vegetation Indices Derived from Quickbird,” *Life Sci J*, vol. 10, no. 4, pp. 1097–8135, 2013, doi: 10.4: 851-860.

[16] F. Marzukhi, A. L. Elahami, and S. N. Bohari, “Detecting nutrients deficiencies of oil palm trees using remotely sensed data,” in *IOP Conference Series: Earth and Environmental Science*, Jul. 2016, vol. 37, no. 1, pp. 1–11. doi: 10.1088/1755-1315/37/1/012040.

[17] K. J. Goh, “Fertilizer recommendation systems for oil palm: estimating the fertilizer rates,” in *Proceedings of MOSTA Best practices workshops-agronomy and crop management. Malaysian Oil Scientists’ and Technologists’ Association*, 2005, pp. 1–37.