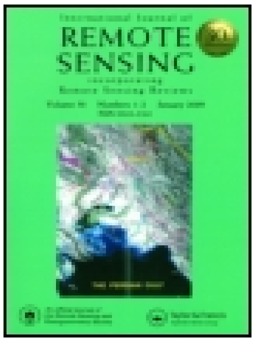

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

Young and mature oil palm tree detection and counting using convolutional neural network deep learning method

Nurulain Abd Mubin, Eiswary Nadarajoo, Helmi Zulhaidi Mohd Shafri & Alireza Hamedianfar

To cite this article: Nurulain Abd Mubin, Eiswary Nadarajoo, Helmi Zulhaidi Mohd Shafri & Alireza Hamedianfar (2019): Young and mature oil palm tree detection and counting using convolutional neural network deep learning method, International Journal of Remote Sensing

To link to this article: https://doi.org/10.1080/01431161.2019.1569282

Published online: 25 Jan 2019.

Submit your article to this journal

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tres20

INTERNATIONAL JOURNAL OF REMOTE SENSING
https://doi.org/10.1080/01431161.2019.1569282

Taylor & Francis
Taylor & Francis Group

Check for updates

Young and mature oil palm tree detection and counting using convolutional neural network deep learning method

Nurulain Abd Mubin[a], Eiswary Nadarajoo[a], Helmi Zulhaidi Mohd Shafri[a,b] and Alireza Hamedianfar[c]

[a]Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang, Malaysia; [b]Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang, Malaysia; [c]Department of Surveying Engineering, Estahban Branch, Islamic Azad University, Estahban, Iran

ABSTRACT

Detection and counting of oil palm are important in oil palm plantation management. In this article, we use a deep learning approach to predict and count oil palms in satellite imagery. Previous oil palm detections commonly focus on detecting oil palm trees that do not have overlapping crowns. Besides this, there is a lack of research that builds separate detection system for young and mature oil palm, utilizing deep learning approach for oil palm detection and combining geographic information system (GIS) with deep learning approach. This research attempts to fill this gap by utilizing two different convolution neural networks (CNNs) to detect young and mature oil palm separately and uses GIS during data processing and result storage process. The initial architecture developed is based on a CNN called LeNet. The training process reduces loss using adaptive gradient algorithm with a mini batch of size 20 for all the training sets used. Then, we exported prediction results to GIS software and created oil palm prediction map for mature and young oil palm. Based on the proposed method, the overall accuracies for young and mature oil palm are 95.11% and 92.96%, respectively. Overall, the classifier performs well on previously unseen datasets, and is able to accurately detect oil palm from background, including plant shadows and other plants.

ARTICLE HISTORY
Received 28 July 2018
Accepted 3 November 2018

1. Introduction

Gathering information of oil palm trees within a plantation is important in order to increase the revenue. Information gathered can be used to estimate yield, amount of fertilizer needed, number of workers needed, and periodic weeding costs (Ng 1972; Kiama, Raman, and Patrick 2014). Many plantation companies practice tree counting because it helps in yield estimation and monitoring. Tree counting can be done manually in the field, but this is labour-intensive, time-consuming, and expensive. Another approach that our plantation industry uses is multiplying total area with the number of palms per hectare (Chong et al. 2017). However, this approach is inaccurate due to land surface variation and features.

CONTACT Helmi Zulhaidi Mohd Shafri helmi@upm.edu.my Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang, Selangor 43400, Malaysia

© 2019 Informa UK Limited, trading as Taylor & Francis Group

The use of remote sensing as an alternative to traditional method has brought many researchers in finding various technique and ways to increase inventory for oil palm tree. Remote-sensing technology has shown great potential in evaluating age and location of oil palm trees in large coverage that contributes to effective productivity (Kanniah, Tan, and Cracknel 2012; Tan, Kanniah, and Cracknell 2012; Cracknell et al. 2013; Tan, Kanniah, and Cracknell 2013; Cracknell et al. 2015; Cheng et al. 2016; Rizeei et al. 2018). Oil palm has unique shape and plantation fashion, making it easy to be detected through aerial and satellite imagery (Shafri, Hamdan, and Saripan 2011). The approach of detecting oil palm also makes it possible to count the number of trees in a plantation (Srestasathiern and Rakwatin 2014). Additionally, counting trees in an image is more manageable than counting on the ground. This is because trees can be labelled easily on the image, and we can easily recount for verification. Individual tree counting has been recognized by means of visual interpretation or digitization. Accurate tree counting result can be achieved by using good quality, very high spatial resolution, and using images free from obstructionMansur, Mukhtar, and Al-Doksi 2014).

Various methods such as local maximum filtering, image binarization, scale analysis, and template matching have been applied for automatic individual tree crown detection (Ke and Quackenbush 2011). One method was developed in a study of tree counting based on unmanned aerial vehicle platform. The data used was an oil palm plantation in Thailand using normalized cross-correlation to detect and remove non-oil palm components (Wong-In et al. 2015). Srestasathiern and Rakwatin (2014) used non-maximal suppression algorithm to detect local peaks from the two-dimensional (2D) semi-variogram computation technique. The proposed method used local peak detection that applied on a vegetation index image, and then applied rank transformation derived from high-resolution multispectral satellite image. 90.00% accuracy has been achieved from this method.

The approach that combines multiple steps has been studied. Shafri, Hamdan, and Saripan (2011) proposed six major parts of scheme from high spatial resolution airborne imagery data. There were (1) discrimination of oil palms from non-oil palms using spectral analysis, (2) texture analysis, (3) edge enhancement, (4) segmentation process, (5) morphological analysis, and (6) blob analysis. One limitation of this approach is that segmentation parameters need to be changed to suit the maturity and size of oil palms. Wong-In et al. (2015) performed several steps such as removing non-tree components from an image, distinguishing oil palms from other components using a low-pass filter and normalized cross-correlation, identifying individual oil palm trees, and counting the number of oil palms.

A previous study by Cheang, Cheang, and Tay (2017) used a convolutional neural network (CNN) to count palm trees in satellite images. They performed the sliding window technique from a larger image, passed into classifier, and the outputs produced the probability that there was a tree crown centred on that window after the binary classification task of oil palm. While the approach uses similar architecture as proposed in this article, it did not compare the architecture performance with respect to oil palm age and did not utilize geographic information system (GIS) approach for detection and counting. Additionally, their Tensorflow frameworks have high learning curves and lead to slow implementation of sequential models such as LeNet. Another study by Li et al. (2017) introduced a deep learning-based method to oil palm tree detection for the first time. They have proposed CNN based on Tensorflow framework by collecting a number of manually interpreted samples to train and optimized the CNN to achieve the best CNN model. Then, the image dataset is used to predict labels for all the samples through the sliding window technique. Finally, the predicted palm coordinates were merged corresponding to the same palm tree into one coordinate and obtained final palm tree detection results.

From the related previous studies by Cheang, Cheang, and Tay (2017) and Li et al. (2017), the research gaps identified are as follows: (1) lack of investigation that explores the potential of deep learning technique to build a separate detection system for young and mature oil palm, (2) only detecting oil palm trees that do not have overlapping crowns, and (3) insufficient research that includes GIS technology during deep learning approach. It should be noted that oil palm plantations are planted in monoculture fashion, with fields divided for immature planting and mature planting. Immature oil palm has a unique plantation pattern. They are planted in a triangular pattern, while mature oil palm has no such pattern, but instead has crops with overlapping crowns (Chong et al. 2017). Due to these different patterns, developing two separate framework for mature and young oil palm would be a possible approach. This could prevent false positives and false negatives during oil palm tree detection.

In this article, we propose a deep learning framework for two different oil palm types, mature and young oil palms. The proposed method uses Keras framework due to faster implementation. Output for this model is prediction where '0' is for the image that does not contain oil palm and '1' for the image that contains oil palm. The predicted images were exported to GIS software for further counting as well as to produce oil palm detection map. Using this approach, we successfully combined deep learning model with GIS to allow existing GIS approach to be used together with deep learning technique. We also compared performance on mature and young oil palm field, providing potential reason for a difference in prediction accuracy and suggesting possible research direction.

2. Materials and methods

2.1. *Study area*

Malaysia contributes to 39% of world palm oil production and 44% of world palm oil exports as of 2017 (Malaysian Palm Oil Industry, 2017). Hence, we chose a study area within Malaysia. The study area is located at Universiti Putra Malaysia (UPM) campus in Serdang, Selangor, as shown in Figure 1. The location of study area is approximately 23 km south of Kuala Lumpur with a latitude of 2°58′32.90″N to 2°59′33.19″N and a longitude of 101°43′5.37″E to 101°43′7.07″E. The area consists of urbanized area and various natural features, with majority of oil palm and other vegetation types.

2.2. *Dataset*

The dataset used in this research is WorldView-3 (WV-3) image rendered in red, green, and blue (RGB) bands. To prepare this dataset, we add WV-3 image into GIS software. Next, we render the image in RGB bands. Following this, we exported the image to joint photographic experts group (JPEG) format for further processing and analysis. Image spatial resolution during export is 0.3 m per pixel. The exported RGB image was further split into 24 tiles, each measuring 991 × 1155 pixel. The tiles were numbered 1–24 according to their

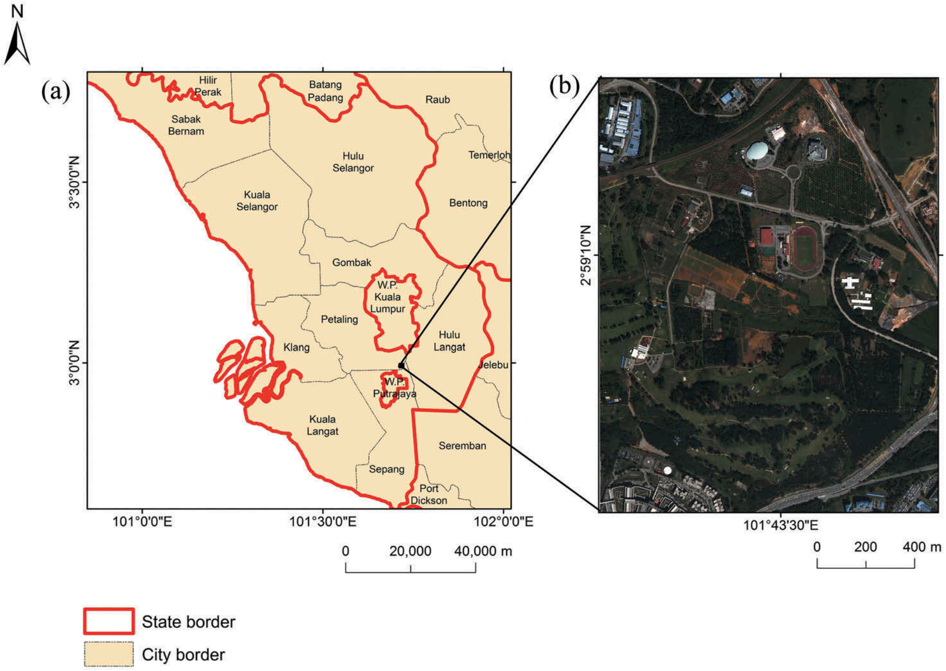

Figure 1. Location of study area: (a) administrative map of Selangor; (b) WorldView-3 image of the study area.

position in the study area, as in Figure 2. Tiles 3, 6, 7, and 8 consist of young oil palm plantation, while tiles 15, 16, 19, 20, 22, and 23 are mature oil palm plantation.

The 24 tiles were further split into smaller sub-tiles. Tiles 3, 6, 7, and 8 were split to 26 × 26 pixel because they contain young oil palm, while tiles 15, 16, 19, 20, 22, and 23 were split to 31 × 31 pixel because they contain mature oil palms. The size was selected such that the image has oil palm as well as background context, which includes bare soil, building, and other vegetation. We also made sure the size is small enough to ensure we have only one oil palm in one image.

We divided images further into training, testing, and validation datasets. Training datasets are used to train the CNN, testing datasets are used to check for overfitting, while validation datasets are used for CNN performance evaluation. We used sub-tiles from tiles 6 and 7 as training, sub-tiles from tile 3 for testing, and sub-tiles from tile 8 for validating young oil palm predictions. For mature oil palm, sub-tiles from tiles 20 and 23 were used as the training dataset, sub-tiles from tile 19 were used for testing, and sub-tiles from tile 16 are used for validating mature oil palm prediction. Table 1 shows the number of sub-tiles used for training, testing, and validation. Although there are more images for mature oil palm, we made sure the train-test split for mature oil palm and young oil palm are set constant at 80:20.

2.3. *Research workflow*

Our research consists of three phases: preprocessing, processing, and postprocessing. We developed and tested our workflow using Jupyter notebook and utilized Keras

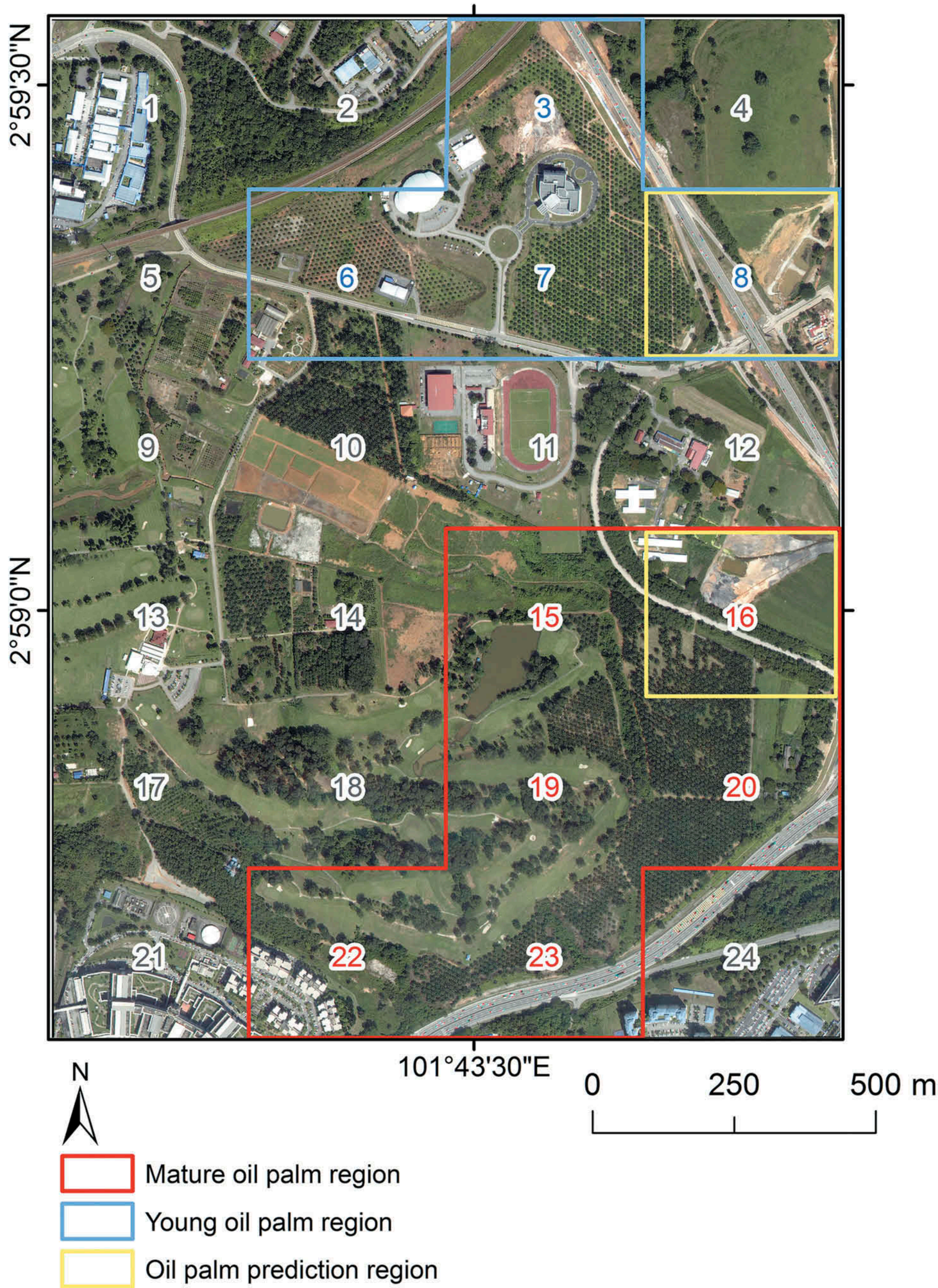

Figure 2. Subsets of study area. The red polygon shows regions where mature oil palm samples are collected while the blue polygon shows the region where young oil palm samples are collected. Yellow polygon shows the region where prediction samples are collected.

deep learning framework to build the deep learning architecture. GIS software was further used to build oil palm prediction map. The flow chart of our proposed method is shown in Figure 3.

Table 1. Dataset grouped in training and testing dataset.

Image	Young oil palm	Mature oil palm
Number of training tiles	452	114
Number of testing tiles	199	85
Number of validation tiles	1672	1185
Training tile: testing tiles	80:20	80:20

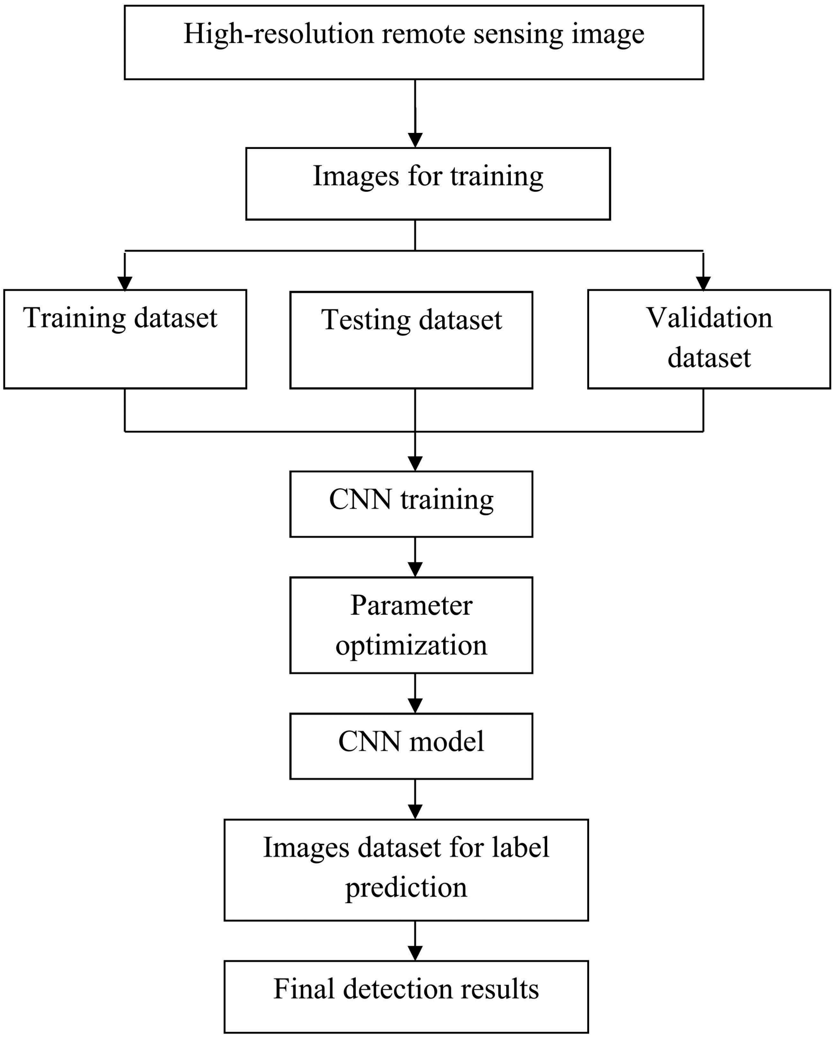

Figure 3. Methodology flow chart.

2.3.1. *Preprocessing*

The steps for preprocessing include exporting image into JPEG format and further splitting images into formats that could be used in the deep learning architecture.

2.3.2. *Phase 1: preparation of data*

We used GIS software to render WV-3 in RGB bands. We then exported the image to JPEG format, making sure the resolution stays at 0.3 m per pixel. Following this, we use GIS tool to further split images into 26 × 26 pixel if we use the image for training young oil palm and 31 × 31 pixel if we used the image to train mature oil palm. All the datasets are manually labelled as 1 if they contain oil palm and 2 if they do not contain oil palm. During training and testing, we made sure background image does not contain any oil palm features and oil palm images have their image centered on one oil palm. Images that do not have these characteristics are removed from training and testing datasets.

We used 260 oil palm images and 192 background images to train CNN model for young oil palm prediction. Sixty-six oil palm images and 48 background images with a total of 114 were used as test dataset to evaluate whether the model is overfitting during training time. Figure 4 shows tiles and sub-tiles used to train a CNN that can detect young oil palm.

We used 119 oil palm images and 80 background images to train CNN model for mature oil palm detection. In validation dataset, 45 oil palms and 40 background images were selected to validate the model. Total of training dataset is 199, meanwhile total of validation dataset is 85. Figure 5 shows tiles and sub-tiles used to train a CNN that can detect mature oil palm.

2.3.3. *Phase 2: processing*

Processing includes building CNN architecture, training CNN architecture, and parameter optimization. The initial architecture developed is based on LeNet, introduced in the article 'Gradient-Based Learning Applied for Document Recognition', by LeCun et al. (1998). The proposed architecture used tan*h* function. In this research, rectified linear unit (RELU) activation function will be used since this function outperforms tan*h* and used to replace tan*h* in practice. The network architecture consists of four convolutional layers. Dropouts are used in the network to reduce overfitting and increase the accuracy of the model, with a dropout rate of 0.5.

The last two dense layers have filters of 64 and 2, respectively, where 2 corresponds to the number of classes to predict. The activation function used during convolution is RELU, while softmax is used as activation function during prediction. This choice of activation function is inline with previous practice where these combinations of activation function are suitable for LeNet-based architecture. All convolution filters have a stride of 3, and all maxpooling layer has a pool size of 2. The total trainable parameter for young oil palm network is 275,597 and the total trainable parameter is 440,597 for mature oil palm. Larger input and output image size attributes to the larger parameter for mature oil palm CNN. Tables 2 and 3 show the CNN architecture summary for young and mature oil palm, respectively.

The training process reduces loss using adaptive gradient algorithm called AdaGrad with a mini batch of size 20 for all the training sets used. Steps per epoch and validation steps were set such that each iteration sees all the images in the training and testing datasets. Table 4 shows the steps per epoch and validation steps of young and mature oil palm. Momentum was set as 0.2, with a learning rate of 0.05 and without weight decay. The training was done on Windows 10, with NVIDIA GPU with each epoch running for 1 s. The training was stopped at 50 epochs for young oil palm and 200 epoch for mature oil palm since no performance increase was observed after this epoch.

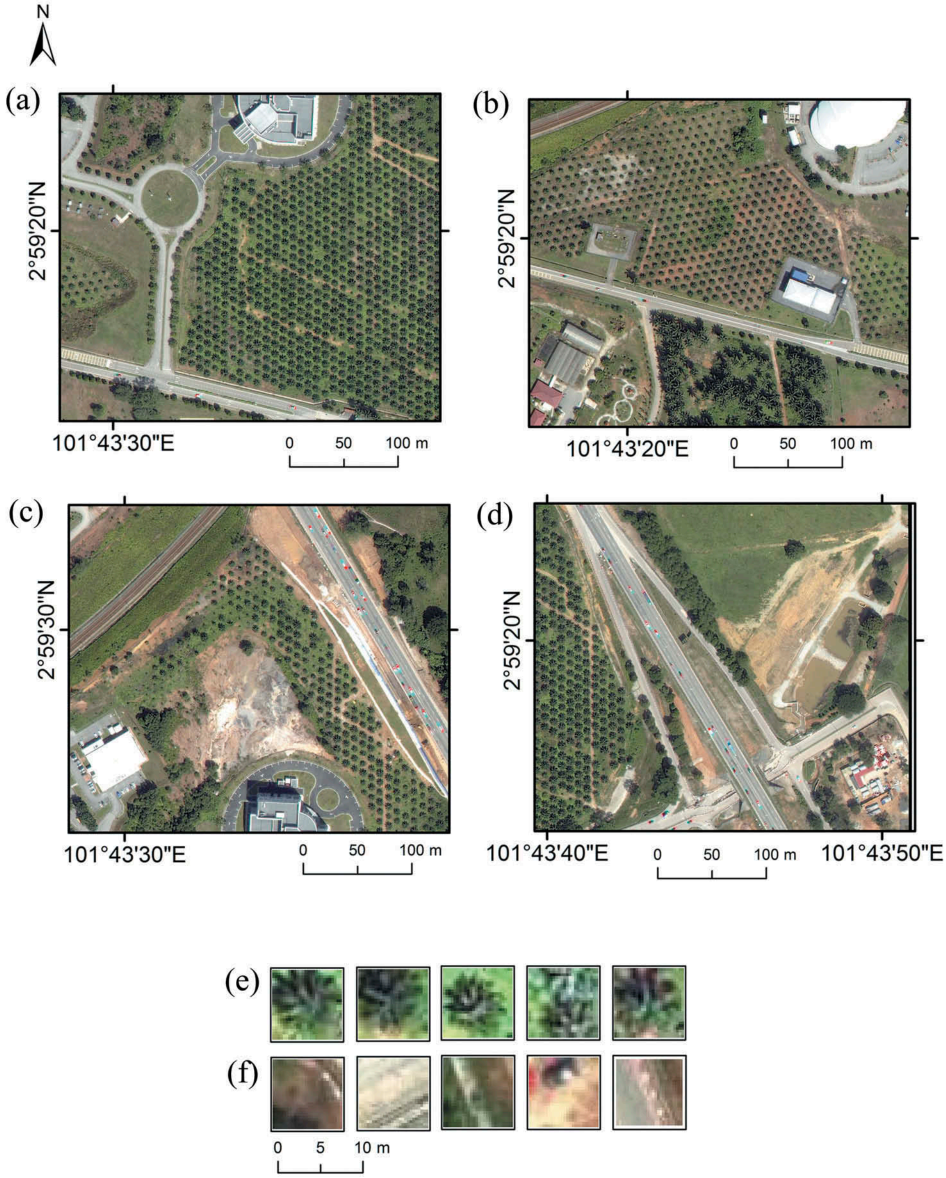

Figure 4. Tiles of the study area for young oil palm: (a) and (b) are used for training dataset; (c) and (d) are used for validation and testing dataset, respectively; (e) and (f) are sub-tiles labelled as young oil palm and background, respectively.

2.3.4. *Phase 3: final detection and counting*

We evaluated the performance of our CNN on validation datasets. Validation datasets are images not used during model training and testing. The images are used to evaluate model performance on unseen dataset and hence can be used to represent model generalization capability and real-world performance.

Inputs for the CNN models are images with same size and resolution as the images used during training and testing. Hence, the input image for CNN model trained on

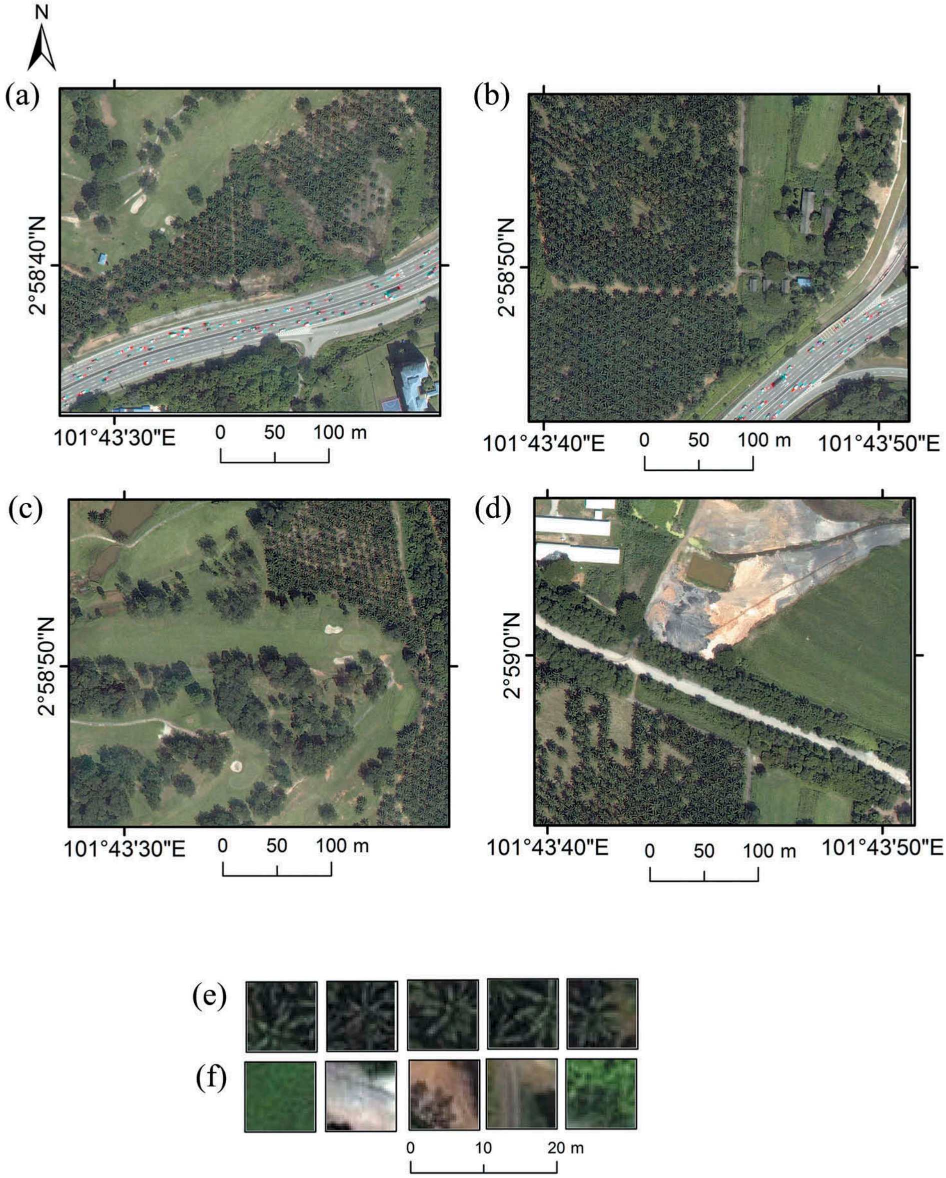

Figure 5. Tiles of study area for mature oil palm: (a) and (b) are used for training dataset; (c) and (d) are used for validation and testing dataset, respectively; (e) and (f) are sub-tiles labelled as mature oil palm and background, respectively.

young oil palm is 26 × 26 pixels and input image for CNN model trained on mature oil palm is 31 × 31 pixel. Outputs for both CNN models are comma-separated value (CSV) files with file name and prediction. The prediction output value is '0' for background and '1' for oil palm. We can directly count the number of oil palms by counting the number of '1' in the output CSV.

Table 2. LeNet architecture used in this research for young oil palm.

Layer	Output shape	Parameter
2D convolution	None, 22,22,30	2,280
Activation	None, 22,22,30	0
2D convolution	None, 18,18,30	22,530
Activation	None, 18,18,30	0
Max pooling	None, 9,9,30	0
Dropout	None, 9,9,30	0
2D convolution	None, 9,9,55	41,305
Activation	None, 9,9,55	0
2D convolution	None, 5,5,55	75,680
Activation	None, 5,5,55	0
Max pooling	None, 2,2,55	0
Dropout	None, 2,2,55	0
Flatten	None, 220	0
Dense	None, 600	132,600
Activation	None, 600	0
Dropout	None, 600	0
Dense	None, 2	1,202
Activation	None, 2	0

Table 3. LeNet architecture used in this research for mature oil palm.

Layer	Output shape	Parameters
2D convolution	None, 27,27,30	2,280
Activation	None, 27,27,30	0
2D convolution	None, 23,23,30	22,530
Activation	None, 23,23,30	0
Max pooling	None, 11,11,30	0
Dropout	None, 11,11,30	0
2D convolution	None, 11,11,55	41,305
Activation	None, 11,11,55	0
2D convolution	None, 7,7,55	75,680
Activation	None, 7,7,55	0
Max pooling	None, 3,3,55	0
Dropout	None, 3,3,55	0
Flatten	None, 495	0
Dense	None, 600	297,600
Activation	None, 600	0
Dropout	None, 600	0
Dense	None, 2	1,202
Activation	None, 2	0

Table 4. Steps per epoch and validation steps of young and mature oil palm.

Layer	Young	Mature
Batch size	20	20
Step per epoch	23	10
Validation step	6	5
Number of epoch	50	200

2.3.5. *Postprocessing*

The steps for postprocessing include evaluating model performance and preparing oil palm prediction map. We evaluate model performance by measuring precision (P), recall

(R), and overall accuracy (OA) for validation datasets. We then display the prediction result as a map using GIS software.

2.3.6. *Phase 4: detection results evaluation and accuracy assessment*

P, R, and OA were used to evaluate model performance as suggested by Zakharova (2017). P is the probability that a detected oil palm tree is valid, R is the probability that an oil palm tree in ground truth is detected, and OA is the average of precision and recall. Equations for this are as per below.

$$P = \frac{(\mathrm{TP})}{((\mathrm{TP}) + (\mathrm{FP}))} \tag{1}$$

$$R = \frac{(\mathrm{TP})}{((\mathrm{TP}) + (\mathrm{FN}))} \tag{2}$$

$$OA = \frac{P + R}{2} \tag{3}$$

where TP stands for true positive, i.e. the number of pixels correctly classified as oil palm. TN stands for true negative, i.e. the number of pixels correctly classified as background. FP stands for false positive, i.e. the number of pixels incorrectly classified as oil palm. FN stands for false negative, i.e. the number of pixels incorrectly classified as background.

2.3.7. *Phase 5: creation of map from deep learning result*

The first step in map creation is to export all 31 × 31 pixel sub-tiles used for mature oil palm validation and 26 × 26 pixel sub-tiles used in young oil palm validation in GIS software. Then, we used GIS image analysis tool to selectively change image hue based on model prediction. For this research, we further changed the hue to indicate TP and FP. We created two prediction maps, one for mature oil palm and another for young oil palm.

3. Results and discussion

3.1. *Accuracy assessment*

The result obtained is divided into two parts, one for young oil palm and one for mature oil palm. CNN model is trained with two classes of image: oil palm (the centre of oil palm tree crown) and background (images that do not have oil palm). Mature oil palm model took a longer time to train because it had more parameters. As there were no clear guidelines for selecting optimizer, we compared the performance of the model using stochastic gradient descent (SGD), root-mean-square propagation (RMSprop), AdaGrad algorithm, per-dimension learning rate method for gradient descent called Adadelta, adaptive moment (Adam) estimation, a variant of Adam based on the infinity norm called Adamax, and nesterov-accelerated adaptive moment (Nadam) estimation optimizers. Table 5 shows the performance of the CNN model for different activation functions. AdaGrad optimizer gives the best model accuracy and model loss for both young and mature oil palm. Changing learning rate, epsilon, and weight decay did not increase the

Table 5. The performance of different optimization methods evaluated in young oil palm.

Optimizer	Parameter	Model loss
SGD	L_r = 0.010, M = 0.0, D = 0.000, nesterov = False	0.0056
RMSprop	L_r = 0.001, ρ = 0.9, $\varepsilon = 10^{-7}$, D = 0.000	0.0075
AdaGrad	L_r = 0.010, $\varepsilon = 10^{-7}$, D = 0.000	0.0006
Adadelta	L_r = 1.000, ρ = 0.950, $\varepsilon = 10^{-7}$, D = 0.000	0.0067
Adam	L_r = 0.001, β_1 = 0.9, β_2 = 0.999, $\varepsilon = 10^{-7}$, D = 0.000, amsgrad = False	0.0030
Adamax	L_r = 0.002, β_1 = 0.900, β_2 = 0.999, $\varepsilon = 10^{-7}$, D = 0.000	0.0030
Nadam	L_r = 0.002, β_1 = 0.900, β_2 = 0.999, $\varepsilon = 10^{-7}$, S_d = 0.004	0.0019

accuracy or reduce loss during training and testing; hence the parameters were set at 0.01, 1×10^{-7}, and 0.0, respectively.

We monitored accuracy and loss during CNN training and testing. We stopped the training process when test accuracy and loss did not improve. Figures 6 and 7 show model accuracy and model loss for young oil palm detection and mature oil palm detection, respectively. Training was stopped at 50 epoch for young oil palm and 200 epoch for mature oil palm because performance did not increase after these epochs.

Table 6 shows overall accuracies for young and mature oil palm. The assumption made in this research is that each of the sub-tiles represents one oil palm. This assumption is made based on the size and resolution of the image used for validation. The tiles have the same size as one oil palm. From Table 6, for young oil palm classification, out of 273 oil palm detected only 22 has been misclassified as oil palm. Meanwhile for mature oil palm, out of 206 oil palm detected only 24 has been misclassified as oil palm. The number of oil palm that has been incorrectly classified as background is 4 for young oil palm and 4 for mature oil palm. Misclassification occurred when vegetation with similar shape and colour as oil is present in the training image. The overall accuracies are 95.11% and 92.96%, respectively, for young and mature oil palm. Accuracy for young oil palms is higher because the crowns did not overlap.

3.2. *Oil palm prediction map*

We evaluated model performance on new dataset. We counted the number of images predicted as oil palm and background. This is then compared with ground truth data.

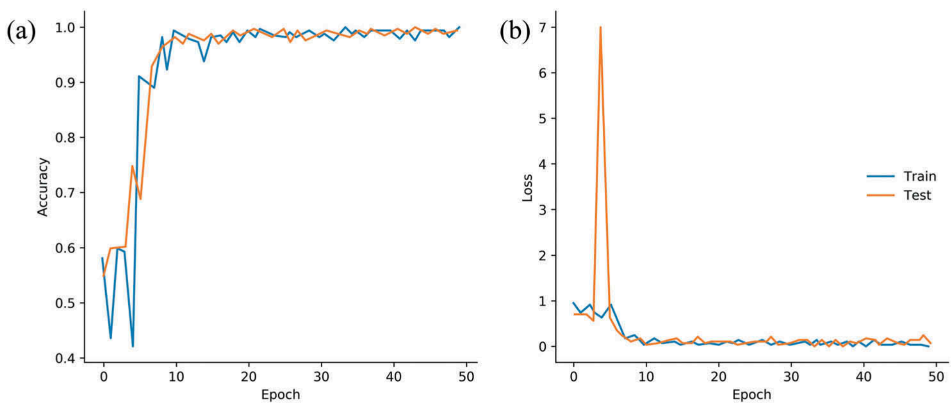

Figure 6. Model accuracy and model loss for young oil palm: (a) model accuracy for young oil palm; (b) model loss for young oil palm.

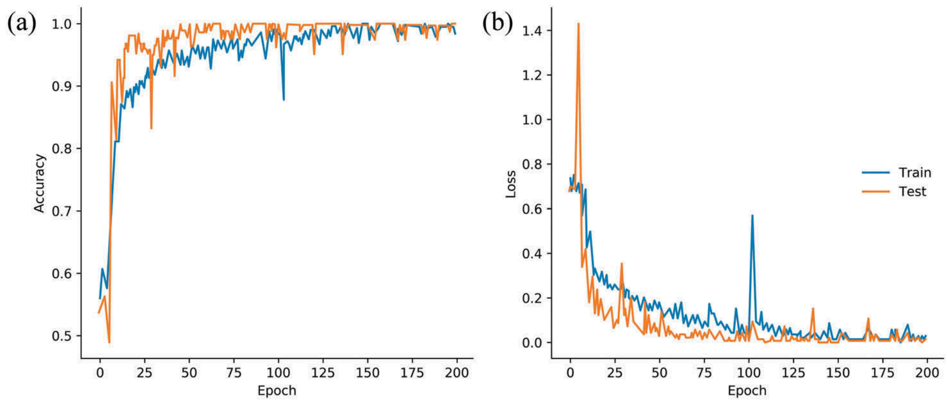

Figure 7. Model accuracy and model loss for mature oil palm: (a) model accuracy for mature oil palm; (b) model loss for mature oil palm.

Table 6. Overall accuracies of young and mature oil palm classification prediction of unseen image dataset.

Evaluation index	Young oil palm	Mature oil palm
TP	247	178
TN	1399	979
FP	22	24
FN	4	4
P (%)	91.82	88.12
R (%)	98.41	97.80
OA (%)	95.11	92.96

We then prepared a prediction map that shows the spatial position of prediction as well as display location of TP and FP, as in Figure 8. TP indicates oil palms that have been correctly predicted, while FP indicates oil palm that has been incorrectly labelled as oil palm. False positives occurred mostly because we used RGB bands. Notable success is for mature oil palm classification, shadows are often correctly classified as background.

4. Discussion

This study investigated the use of deep learning technique to identify two categories of young and mature oil palm trees using WV-3 data. We believe the current work added new findings to the existing literature on oil palm tree detection. A previous study by Li et al. (2017) utilized QuickBird image with 2.4 m spatial resolution and performed deep learning technique to only detect one feature class of oil palm trees without considering different ages. One of the main contributions of the current work is consideration of pansharpened WV-3 image with 0.3 m pixel size. The higher spatial resolution of WV-3 provides heterogenous and complex information from tree crowns, requiring more efficient processing to separate young and mature oil palm trees. The other contributions compared with Li et al. (2017) are as follows: (1) GIS technique was used instead of the sliding window technique, for oil palm tree counting and mapping. GIS serves as a better approach for spatial data management and including GIS system in deep learning approach leads to seamless integration with existing GIS-based research for oil palm management; (2) hyperparameter tuning with different optimizers

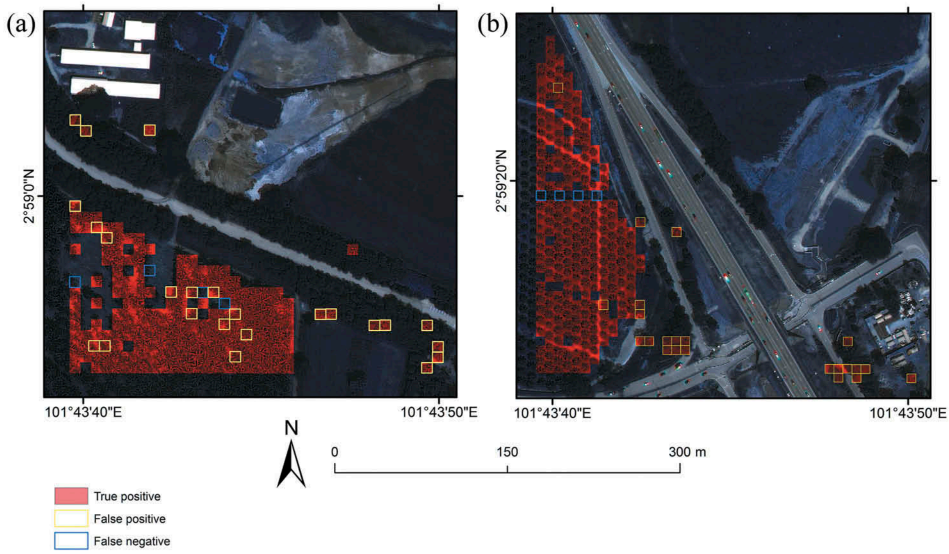

Figure 8. Result for oil palm and background prediction: (a) prediction for mature oil palm region; (b) prediction for young oil palm region.

such as SGD, RMSprop, AdaGrad, Adadelta, Adam, Adamax, and Nadam was explored and finally Adagrad was selected as the best optimizer for this research; and (3) Keras deep learning framework offered an efficient implementation and was achieved smaller learning curve.

5. Conclusion

Oil palm counting plays an important role in plantation management. Manual observation is inaccurate, inefficient, and not practical in a large plantation, while other approaches used may result in false positives and false negative since one framework is used to detect both mature and young oil palm. Remote-sensing approach proves to be a better approach, but current limitations include lack of research that builds separate detection system for young and mature oil palm, lack of approach that utilized deep learning, as well as insufficient research that includes GIS technology in detection and data storage process.

This research serves to close these research gaps by using deep learning method for oil palm detection and counting. We changed model architecture and input image size to best fit two different oil palm plantation types, mature and young oil palm types. We then assessed the performance of the CNN models separately for each oil palm type using validation sets. AdaGrad optimizer has been chosen because it has the lowest loss value and high accuracy. Preparing suitable input dataset for training and testing is vital in getting good model accuracy and model loss. To ensure that deep learning approach works seamlessly with existing GIS approach, we took measures to process images using GIS software, hence ensuring the spatial information for each image is maintained. We also stored the images in GIS database for easy retrieval and further processing. Our result shows that accuracy was at 95.11% for young oil palm plantation and 92.96% for mature oil palm plantation. This indicates there is a potential to improve mature and young oil palm detection using deep learning architecture.

In future works, the palm tree detection results could be further improved by using high-resolution hyperspectral remote-sensing images or multispectral images with more spectral bands. Current implementation of deep learning framework is time-consuming and has a high learning curve. Future research could also develop tools that work with current GIS systems, or as a stand-alone tool that can be easily utilized in oil palm plantation managers. Current technique also relies on an image with fixed image size and resolution. In this research, the image has a spatial resolution of 0.3 m per pixel and requires different input sizes for mature and young oil palm detection. Future research may work on developing an architecture which takes an image with different spatial resolutions and image sizes.

Acknowledgments

The authors would like to acknowledge the funding and facilities provided by UPM as well as NVIDIA Corporation for the donation of Quadro M5000 GPU used for this research. Comments from the anonymous reviewers are also highly appreciated.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Nvidia; Universiti Putra Malaysia under grant UPM/700-1/2/GPB/2017/9543100.

References

Britz, D. 2017. Understanding Convolutional Neural Networks for NLP. Accessed June 8 2017. http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-fornlp/

Cheang, E. K., T. K. Cheang, and Y. H. Tay. 2017. "Using Convolutional Neural Network to Count Palm Trees in Satellite Images." Accessed May 21 2018. https://arxiv.org/abs/1701.06462.https://arxiv.org/abs/1701.06462

Cheng, Y., L. Yu, A. P. Cracknell, and P. Gong. 2016. "Oil Palm Mapping Using Landsat and PALSAR: A Case Study in Malaysia." *International Journal of Remote Sensing* 37 (22): 5431–5442. doi:10.1080/01431161.2016.1241448.

Chong, K. L., K. D. Kanniah, C. Pohl, and K. P. Tan. 2017. "A Review of Remote Sensing Applications for Oil Palm Studies." *Geo-Spatial Information Science* 20 (2): 184–200. doi:10.1080/10095020.2017.1337317.

Cracknell, A. P., K. D. Kanniah, K. P. Tan, and L. Wang. 2013. "Evaluation of MODIS Gross Primary Productivity and Land Cover Products for the Humid Tropics Using Oil Palm Trees in Peninsular Malaysia and Google Earth Imagery." *International Journal of Remote Sensing* 34 (20): 7400–7423. doi:10.1080/01431161.2013.820367.

Cracknell, A. P., K. D. Kanniah, K. P. Tan, and L. Wang. 2015. "Towards the Development of a Regional Version of MOD17 for the Determination of Gross and Net Primary Productivity of

Oil Palm Trees." *International Journal of Remote Sensing* 36 (1): 262–289. doi:10.1080/01431161.2014.995278.

Global Palm Oil Production. 2016. Accessed June 8 2017. http://greenpalm.org/about-palm-oil/where-is-palm-oil-grown-2

Kanniah, K. D., K. P. Tan, and A. P. Cracknel 2012. "UK-DMC 2 Satellite Data for Deriving Biophysical Parameters of Oil Palm Trees in Malaysia." Paper presented at the Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 2012. IEEE International.

Ke, Y., and L. J. Quackenbush. 2011. "A Review of Methods for Automatic Individual Tree-Crown Detection and Delineation from Passive Remote Sensing." *International Journal of Remote Sensing* 32 (17): 4725–4747. doi:10.1080/01431161.2010.494184.

Kiama, J., V. Raman, and T. Patrick 2014. Low-Cost RFID-based Palm Oil Monitoring System (PMS): First Prototype. Paper presented at the IOP Conference Series: Earth and Environmental Science.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86 (11): 2278–2324. doi:10.1109/5.726791.

Li, W., H. Fu, L. Yu, and A. Cracknell. 2017. "Deep Learning Based Oil Palm Tree Detection and Counting for High-resolution Remote Sensing Images." *Remote Sensing* 9 (1): 22. doi: 10.3390/rs9010022.

Malaysian Palm Oil Industry. Accessed June 6 2017 http://www.mpoc.org.my/Malaysian_Palm_Oil_Industry.aspx

Mansur, M. A., R. B. Mukhtar, and J. Al-Doksi. 2014. "The Usefulness of Unmanned Airborne Vehicle (UAV) Imagery for Automated Palm Oil Tree Counting." *Reseacrhjournali's Journal of Forestry* 1 (1): 1–12.

Ng, S. K. 1972. The Oil Palm, Its Culture, Manuring and Utilisation. International Potash Institute. http://www.ipipotash.org/udocs/the_oil_palm-its_culturemanuring_and_utilisation.pdf

Palm Oil Facts and Figures. Accessed June 6 2017 http://www.simedarby.com/upload/Palm_Oil_Facts_and_Figures.pdf

Rizeei, H. M., H. Z. Shafri, M. A. Mohamoud, B. Pradhan, and B. Kalantar. 2018. "Oil Palm Counting and Age Estimation from WorldView-3 Imagery and LiDAR Data Using an Integrated OBIA Height Model and Regression Analysis." *Journal of Sensors* 2018: 1–13. doi:10.1155/2018/2536327.

Shafri, H. Z., N. Hamdan, and M. I. Saripan. 2011. "Semi-Automatic Detection and Counting of Oil Palm Trees from High Spatial Resolution Airborne Imagery." *International Journal of Remote Sensing* 32 (8): 2095–2115. doi:10.1080/01431161003662928.

Srestasathiern, P., and P. Rakwatin. 2014. "Oil Palm Tree Detection with High Resolution Multi-Spectral Satellite Imagery." *Remote Sensing* 6 (10): 9749–9774. doi:10.3390/rs6109749.

Tan, K. P., K. D. Kanniah, and A. P. Cracknell. 2012. "A Review of Remote Sensing Based Productivity Models and Their Suitability for Studying Oil Palm Productivity in Tropical Regions." *Progress in Physical Geography* 36 (5): 655–679. doi:10.1177/0309133312452187.

Tan, K. P., K. D. Kanniah, and A. P. Cracknell. 2013. "Use of UK-DMC 2 and ALOS PALSAR for Studying the Age of Oil Palm Trees in Southern Peninsular Malaysia." *International Journal of Remote Sensing* 34 (20): 7424–7446. doi:10.1080/01431161.2013.822601.

Wong-In, T., T. Kaewkongka, N. Cooharojananone, and R. Lipikorn. 2015. "Automatic Oil Palm Detection and Identification from Multi-Scale Clustering and Normalized Cross Correlation." In *Industrial Engineering, Management Science and Applications 2015*,, edited by M. Gen, K. J.Kim, X. Huang, and Y. Hiroshi, 403–410. Berlin: Springer-Verlag.

Zakharova, M. 2017. Automated Coconut Tree Detection in Aerial Imagery Using Deep Learning. Master Thesis. Technology Campus De Nayer.