LARGE-SCALE OIL PALM TREE DETECTION FROM HIGH-RESOLUTION REMOTE SENSING IMAGES USING FASTER-RCNN

*Juepeng Zheng[1, 2, 3], Weijia Li[1, 2, *], Maocai Xia[1, 2], Runmin Dong[1, 2], Haohuan Fu[1, 2], Shuai Yuan[4]*

[1]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China

[2]Joint Center for Global Change Studies (JCGCS), Beijing 100084, China

[3]College of Surveying and Geo-informatics, Tongji University, Shanghai 100092, China

[4]Department of Eletronic Engineering, Tsinghua University, Beijing 100084, China

ABSTRACT

Oil palm is of great importance in agricultural productivity for many tropic developing countries and accordingly investigating as well as counting oil palms is a meaningful and valuable research. In this paper, we firstly apply Faster-RCNN, one of the most popular object detection algorithms, to detect tree crowns from satellite images. Although Faster-RCNN has an excellent performance in well-known datasets of general object detection, it does not have obvious advantages in oil palm tree detection in this study compared with other classical machine learning based methods. We argue two reasons accounting for the drawbacks of Faster-RCNN: (1) the size of each oil palm tree is too small (only 17 × 17 pixels on average) in 0.6m-resolution QuickBird satellite images; (2) there are lots of other similar trees around the oil palm trees that make it difficult to detect them correctly. In order to reach a satisfying accuracy, we tailored the Region Proposal Network (RPN) and proposed a simple but practical post-processing strategy based on empirical planting rules, filtering out the wrongly detected trees (False Positives) effectively. Eventually we achieved a higher average F1-score of 94.99% (using IOU based evaluation matrices) in our six study regions compared wtih state-of-the-art oil palm detection methods. In addition, we proposed a workflow of large-scale oil palm tree detection using high-resolution remotely sensed images based deep learning methods.

Index Terms— Oil palm trees, large-scale object detection, Faster-RCNN, high-resolution remote sensing images

1. INTRODUCTION

The oil palm tree, an agricultural economic crop, is of vital importance in many tropic countries such as Malaysia, Indonesia, Brazil and so on. The main usage of oil palm trees is to produce palm oil, which not only has been applied to producing edible vegetable oil, but also raw material of cosmetics, furniture, etc. When it comes to oil palm plantation, counting and detection of oil palm trees is one of the significant work, supplying helpful information for monitoring their growth, assessing their values and planning the planting layout. Due to the rapid advancement of high-resolution satellite [1-3] and unmanned aerial vehicle (UAV) [4] in recent years, automaticly detecting tree crowns from high-resolution remote sensed images is one of the most popular technique for oil palm detection and counting.

In previous studies, tree crown detection from very high-resolution remote sensing images are mostly based on traditional image processing methods, classical machine learning methods and deep learning methods. For traditional image processing methods, there are well-known algorithms such as image segmentation, template matching, and local maximum filter, etc. Although these methods can be implemented easily and do not need samples by manual work, as well as have good results in simple areas, they perform seriously bad especially in densely planted areas. Classical machine learning method includes maximum likelihood, support vector machine (SVM), random forest (RF), artificial neural network (ANN), etc. having a better effect than traditional image processing methods generally, but different kinds of feature extraction makes a considerable difference in the final results. In 2016, Li et al. [5] firstly applied deep learning based method to tree crown detection located in Malaysia, which proposed a convolutional neural network and sliding window method with a precision of 96%, exceeding three other classical crown detection methods. Recently, Li et al. [6] proposed the two-stage CNN (TS-CNN), including one CNN for land cover classification and one CNN for object classification, which achieves much higher average F1-scores of 92.80% (using the popular IOU based evaluation matrices) compared with single-stage CNN, SVM, RF and ANN.

* Correspondence: liwj14@mails.tsinghua.edu.cn

Authorized licensed use limited to: Tsinghua University. Downloaded on September 06,2020 at 02:40:39 UTC from IEEE Xplore. Restrictions apply.

Compared with classical machine learning based method, deep convolutional neural network based deep learning methods have distinct advantage in remote sensing. Since Faster-RCNN was proposed [7], it has been an excellent algorithm for general object detection in computer vision. However, when we try to apply Faster-RCNN to oil palm tree detection, it can not perform beyond the classical machine learning methods and will wrongly detect plenty of trees that are not actual oil palms. Because of its drawbacks and regular pattern for oil palm plantation, we proposed a simple but practical post-processing approach according to the general layout of oil palm, which achieves a much faster detection speed and a higher average F1-score of 94.99% than TS-CNN and other classical machine learning methods.

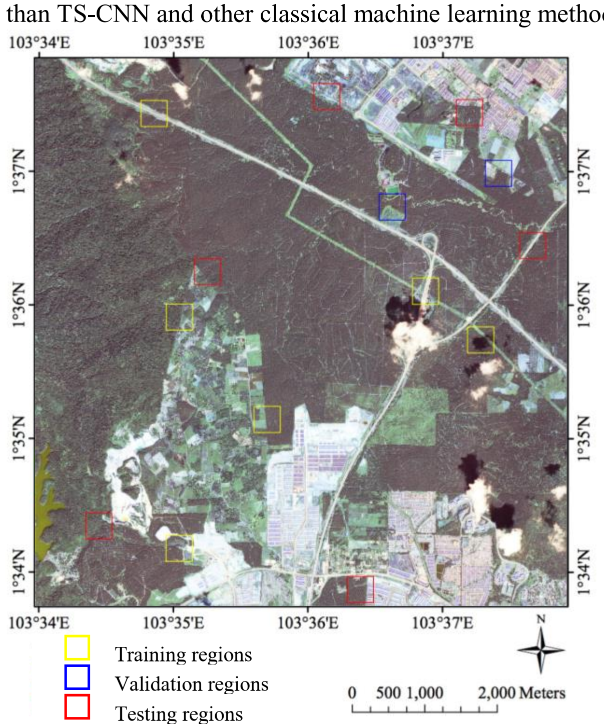

Figure 1. The study area of this research and the regions of training, validation and test are denoted by yellow, blue, red rectangles respectively.

2. DATASETS AND METHODS

2.1. Study area and datasets

As is shown in Figure 1, our study area is located in the South of Malaysia. The QuickBird image was photographed in 2006, with 0.6m spatial resolution in four spectral bands after preprocessing [5].

We selected 14 regions with 600 × 600 pixels and 6 of them are training areas, 2 are validation areas and the rest are test areas, which are represented by red, blue and yellow rectangles respectively in Figure 1. All oil palm trees of regions above are labelled manually.

In 6 training areas, we applied the bilinear interpolation so that the original images are resized to 2400 × 2400 pixels. After that, we randomly cropped the enlarged training images with 500 × 500 pixels and generated the training dataset of 5000 samples as the input of deep neural network. As for validation and testing datasets, we firstly resized them using the same scale (enlarge by 4 times both in width and height) as the training dataset and cropped them with 66 pixels overlapped. (Given that the size of an oil palm tree in QuickBird image with 0.6m spatial resolution is about 17 × 17 pixels and 65 × 65 pixels or so after image resizing.)

2.2. Methods

2.2.1. Data preprocessing and upsampling

The framework of our proposed oil palm detection method is represented in Figure 2. In the training phase, we divided large-scale oil palm images into blocks and employ the data augmentation of flipping. In the test phase, we divided the original QuickBird images with overlapping after enlarging it by 4 × 4 times on area. We adopted non-maximum suppression (NMS) on the detection oil palms based on the score to acquire the whole large-scale oil palm tree detection results.

2.2.2. Region proposal network for oil palm detection

In Faster-RCNN, RPN is the most impressing contribution of object detection that takes images of any size as input and provides us object proposals with categories.

We implemented and evaluated three types of RPN backbone structure in this study: (1) ZF model [8]; (2) VGG_CNN_M_1024 model [9]; (3) VGG16 model [10]. We display VGG16 as our backbone structure in this paper and we will show the effects of these 3 different models in Section 4.

2.2.3. Post-processing for filtering wrong trees

We proposed a post-processing strategy based on empirical planting rules, which can effectively filter out a large amount of wrongly detected trees (False Positives) in the results of the Faster-RCNN model.

As one of the most crucial economic sources in many tropic countries, most of the oil palm trees were planted intensively and uniform. We can rarely find the oil palm planted alone or gathered in two or three in remote sensing images. Depending on this helpful but simple prerequisites, we proposed the tree filtration strategy in which we drop out

the detected trees where there is fewer than or equal to 2 detected trees within 30 pixels, and iterate this strategy until the dropped-out oil palms in this region are fewer than the amount of 10. In Section 4.3 we will explore the significance of this post processing method.

Authorized licensed use limited to: Tsinghua University. Downloaded on September 06,2020 at 02:40:39 UTC from IEEE Xplore. Restrictions apply.

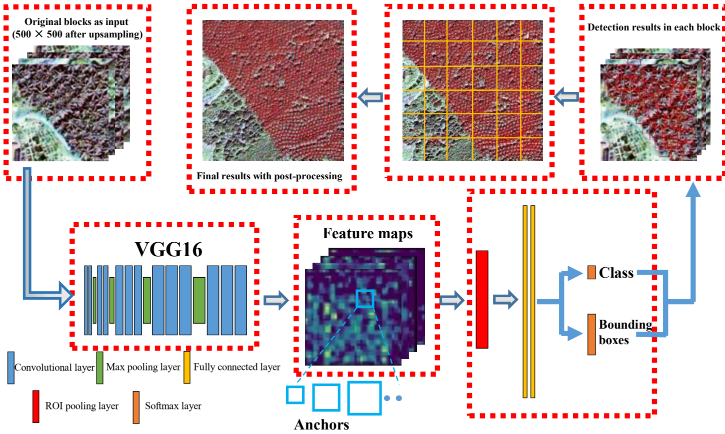

Figure 2. The flowchart of large-scale oil palm tree detection.

3. ABLATION EXPERIMENTS

3.1. Which backbone network performs best for oil palm detection?

Table 1. F1 score of each backbone network

Index	R 1	R 2	R 3	R 4	R 5	R 6
ZF	92.99	93.99	92.40	93.06	79.93	71.69
VGGM1024	92.86	92.95	92.49	92.11	77.18	65.63
VGG16	**95.11**	**96.62**	**94.81**	**96.00**	**91.02**	**83.58**
ZF + Post	94.56	96.22	94.38	94.64	90.52	79.88
VGGM1024 + Post	94.24	95.17	94.39	93.77	88.93	71.57
VGG16 + Post	**95.58**	**96.93**	**95.16**	**97.25**	**94.93**	**90.09**

In our experiment, an oil palm will be considered as correctly detected only if the Intersection of Union (IoU) between the detected oil palm and an oil palm in the ground truth dataset is more than or equal to 0.5. We adopted 3 backbone networks in RPN, including ZF model [8], VGG_CNN_M_1024 model [9] and VGG16 model [10]. Table 1 illustrates the effectiveness of this 3 backbone structures in oil palm tree detection.

Obviously, VGG16 has the highest average F1-score among this three models, followed by the ZF model, and the VGG_CNN_M_1024 model. VGG16 has the strong capacity of feature extraction because of deeper convolutional layers so that it performs successfully (average F1-score is over 90%) no matter whether adopting post-processing or not.

3.2. How important is the post-processing method?

As is shown in Table 1, we discover that the post-processing plays a vital role in the whole oil palm detection, increasing the average F1-score by about 2-4% generally. In our experiment, we remove the detected trees surrounded by fewer than or equal to 2 other detected trees within 30 pixels, iterating until the difference between the number of trees in the current step and the last step fewer than 10. On the contrary, our post-processing method may lead to side-effect, resulting in the decrease of the correct detected trees (true positives) and the recall. However, Table 2 and Table 3 reveal that fewer than 2 correct trees will drop through the post-processing while the wrongly detected trees will drop over 60 at most, which greatly improves the precision in our 6 study regions.

Table 2. Detection results without post processing

Index	R 1	R 2	R 3	R 4	R 5	R 6
TP	1137	1044	1024	971	451	402
FP	62	40	58	61	69	104
FN	55	33	54	20	20	54
Precision (%)	94.83	96.31	94.64	94.09	86.73	79.45
Recall (%)	95.39	96.94	94.99	97.98	95.75	88.16
F1-score (%)	95.11	96.62	94.81	96.00	91.02	83.58

Table 3. Detection results post processing

Index	R 1	R 2	R 3	R 4	R 5	R 6
TP	1136	1043	1023	971	449	400
FP	49	32	49	35	26	32
FN	56	34	55	20	22	56
Precision (%)	95.86	97.02	95.43	96.52	94.53	92.59
Recall (%)	95.30	96.84	94.90	97.98	95.33	87.72
F1-score (%)	95.58	96.93	95.16	97.25	94.93	90.09

3.3. Comparisons with state-of-the-art oil palm tree detection methods

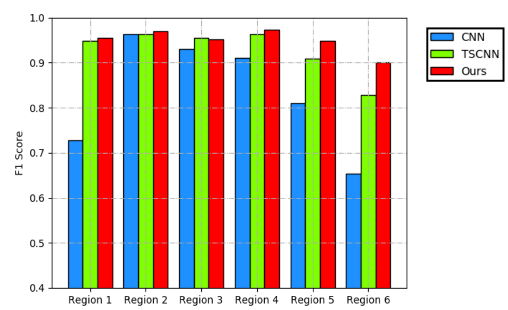

Figure 3. Average F1-score of each method

Figure 3 illustrates the results of three of oil palm detection methods including CNN, TS-CNN [6] and ours, and indicates that our method outperforms the other two state-of-the-art oil palm tree detection methods in all our study regions except region 3. In terms of the average F1-score in 6 study regions, these three methods achieve 86.48%, 92.80% and 94.99% respectively, and we can clearly discover the differences among the results from these methods in Figure 4.

4. CONCLUSIONS

In this paper, we argued the limitations of oil palm detection using Faster-RCNN, a popular object detection algorithm,

Authorized licensed use limited to: Tsinghua University. Downloaded on September 06,2020 at 02:40:39 UTC from IEEE Xplore. Restrictions apply.

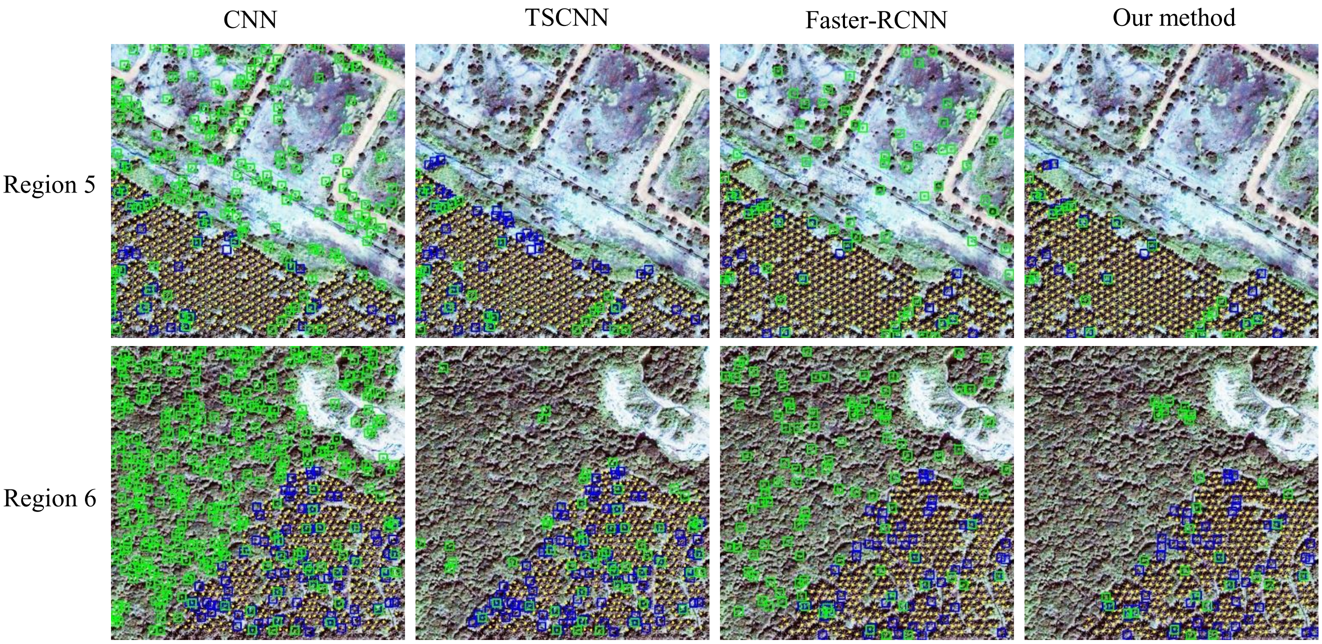

Figure 4. Detection image of our study regions. Yellow points denote detected palm trees. Blue squares denote palm trees not detected correctly. Green squares denote other objects detected as palm trees by mistake.

and proposed a new oil palm detection method based on the sample preprocessing, tailored RPN of Faster-RCNN, and a simple but effective post-processing method, achieving a higher average F1-score of 94.99% in our six study regions compared with existing deep learning methods. In addition, we analyzed the performance of different deep learning backbones for oil palm detection. In our future work, we will further improve the oil palm detection method to achieve higher detection accuracy, and apply our proposed method to larger-scale study regions.

5. ACKNOWLEDGMENTS

This research was supported in part by the National Key R&D Program of China (Grant No. 2017YFA0604500 and 2017YFA0604401), by the National Natural Science Foundation of China (Grant No. 51761135015), and by the Center for High Performance Computing and System Simulation, Pilot National Laboratory for Marine Science and Technology (Qingdao).

6. REFERENCES

[1] Li, W., He, C., Fang, J., Zheng, J., Fu, H., & Yu, L. (2019). Semantic Segmentation-Based Building Footprint Extraction Using Very High-Resolution Satellite Images and Multi-Source GIS Data. Remote Sensing, 11(4), 403.

[2] Nie Y, Shen Y, Chen Q. Combination Analysis of Future Polar-Type Gravity Mission and GRACE Follow-On[J]. Remote Sensing, 2019, 11(2): 200.

[3] Wu, H.; Xu, Z.; Wu, G. A Novel Method of Missing Road Generation in City Blocks Based on Big Mobile Navigation Trajectory Data. ISPRS Int. J. Geo-Inf. 2019, 8, 142.

[4] Zortea, M., Nery, M., Ruga, B., Carvalho, L. B., & Bastos, A. C. (2018, July). Oil-Palm Tree Detection in Aerial Images Combining Deep Learning Classifiers. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium (pp. 657-660). IEEE.

[5] Li, W.; Fu, H.; Yu, L.; Cracknell, A. Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. Remote Sens. 2017, 9, 22.

[6] Li, W.; Dong, R.; Fu, H.; Yu, L. Large-Scale Oil Palm Tree Detection from High-Resolution Satellite Images Using Two-Stage Convolutional Neural Networks. Remote Sens. 2019, 11, 11.

[7] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems, Montréal, QC, Canada, 7–12 December 2015; pp. 91–99.

[8] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[C]//European conference on computer vision. Springer, Cham, 2014: 818-833.

[9] Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. arXiv preprint arXiv:1405.3531.

[10] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

Authorized licensed use limited to: Tsinghua University. Downloaded on September 06,2020 at 02:40:39 UTC from IEEE Xplore. Restrictions apply.