Automatic Detection of Individual Oil Palm Trees from UAV Images Using HOG Features and an SVM Classifier

Yiran Wang, Xiaolin Zhu, Bo Wu

Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Abstract

Oil palm trees are important economic crops in tropical areas. Accurate knowledge of the number of oil palm trees in a plantation area is important to predict the yield of palm oil, manage the growing situation of the palm trees and maximise their productivity. In this study, we propose a novel automatic method for detection and enumeration of individual oil palm trees using images from unmanned aerial vehicles. This method required three major steps. First, images from unmanned aerial vehicles were classified as vegetation or non-vegetation by the support vector machine classifier. Second, a feature descriptor based on the histogram of oriented gradient was designed for palm trees and used to extract features for machine learning. Finally, a support vector machine classifier was trained and optimised using the histogram of oriented gradient features from positive (i.e., oil palm trees) and negative samples (i.e., objects other than oil palm trees). The trained classifier was then applied to detect individual oil palm trees using adaptive moving windows that allowed it to also return the crown size of each oil palm tree. The method was trained at one site and validated independently at four other sites with different situations. The overall accuracy of palm tree detection was 99.21% at the training site and 99.39%, 99.06%, 99.90% and 94.63% at the four validation sites; the last one was for the most challenging site, in which palm trees were mixed with other trees. These tests confirm the effectiveness of the proposed method. The simplicity and great efficiency of the proposed method allow it to support oil palm tree counting for large areas using imagery from unmanned aerial vehicles.

Keywords: Oil palm trees; UAV photos; HOG; SVM; Object detection; Classification

1. Introduction

Oil palm trees are important economic crops in tropical areas. Information regarding the locations, numbers and diameters of oil palm trees in a plantation area is important in many aspects. For example, it is vital for monitoring the growing situation of palm trees after planting, such as their age and survival rate. It is also essential to predict the yield of palm oil, which is one of the world's most widely used vegetable oils. Information about oil palm trees is also important to plan future cultivation and for biodiversity conversation. According to Malaysian Palm Oil Council, the 4.49 million hectares of land in Malaysia under oil palm cultivation produce 17.73 million tonnes of palm oil each year.

The large cultivation areas of palm trees have motivated the use of remote sensing to produce such data automatically. Previous studies of palm tree detection were usually based on commercial high-resolution satellite images (Jusoff and Pathan, 2009; Korom et al., 2014; Li et al., 2016; Srestasathiern and Rakwatin, 2014). However, satellite imagery has a relatively high cost, and data availability is often limited by weather and by the satellite revisit cycle. In recent years, unmanned aerial vehicles (UAVs) have been increasingly adopted as cost-effective remote-sensing data acquisition systems that allow the land surface to be mapped and monitored at an extremely high resolution and can reach the desired point of observation in just a few minutes, which permits interactive measurements according to a customer's specific needs. Their flexibility and competitive prices have made UAVs a practical solution for many agricultural applications such as vegetation monitoring (Berni et al., 2009; Uto et al., 2013) and precision agriculture (Xiang and Tian, 2011; Zarco-Tejada et al., 2013). However, in contrast, the improved spatial resolutions of the UAV images can collect more detailed information about objects

on the ground, which renders many methods developed for traditional satellite images less applicable.

In recent years, researchers have designed several methods of palm tree detection from UAV images. Malek et al. (2014) used a supervised extreme learning machine classifier to detect palm trees from UAV images. They used the extreme learning machine classifier to analyse the scale-invariant feature transform (SIFT) key points. The shape of each palm tree was detected by merging key points with an active contour method, followed by textural analysis of palm trees with local binary patterns (LBP) to distinguish palm trees from other vegetation. The accuracy of this method ranged from 89.4% to 96.4% for various experimental areas. However, their experimental areas presented relatively easy cases for palm tree detection because the palm trees were sparsely distributed without canopy overlapping and differed significantly in appearance from the bare soil background. Manandhar et al. (2016) noted that in dense cultivation regions, the method proposed by Malek et al. (2014) would have difficulty finding the arbitrary boundaries between the overlapped palm trees. Therefore, they proposed a new method to detect and count palm trees. This method first uses circular autocorrelation of the polar shape matrix and the linear support vector machine (SVM) to extract image features and then uses a local maximum detection algorithm to detect palm trees. This method can achieve an overall accuracy ranging from 60.9% to 95.1% for eight images. The performance of this method is not desirable for tough scenarios, and the palm trees are recorded as points without including the diameter information for individual trees.

In this study, we present a novel automatic method for palm tree detection in UAV images based on histogram of oriented gradient（HOG) features and the SVM classifier (HOG-SVM). The HOG-SVM method has proved effective in object detection. It was first introduced for pedestrian detection (Dalal and Triggs, 2005) and has since been

applied to detect many different objects, such as vehicle logos (Llorca et al., 2013) and traffic signs (Yao et al., 2014). In this study, the HOG-SVM method was extended to detect oil palm trees in various scenarios, including the most challenging cases in which palm trees overlap and are mixed with other tree species.

The remainder of this paper is organised as follows. Section 2 describes the study area and datasets used in this study. Section 3 presents the details of the proposed HOG-SVM method for palm tree detection in UAV images. Section 4 shows the detection results in five areas and presents the accuracy assessment. Section 5 concludes this study and discusses directions for future study.

2. Study Area and Data

The study area was in Jeram, a city on the west coast of Malaysia. The images were acquired on sunny days during 2017 with a commercial camera mounted on a UAV system. The camera had red, green, and blue bands with a spatial resolution of 0.04 m. As Figure 1 shows, five pilot areas with various scenarios were selected for training and validation. Figure 1 (a) is the site from which the training samples were selected. Figures 1 (b-e) show various situations for validation purposes. Figure 1 (b) shows both large mature trees and small young trees. Figure 1 (c) shows palm trees planted in diverse spatial patterns. Figure 1 (d) shows very crowded palm trees whose crowns overlap. Figure 1 (e) shows palm trees mixed with other tree species. Each of these situations presents different challenges for automatic detection of individual palm trees.

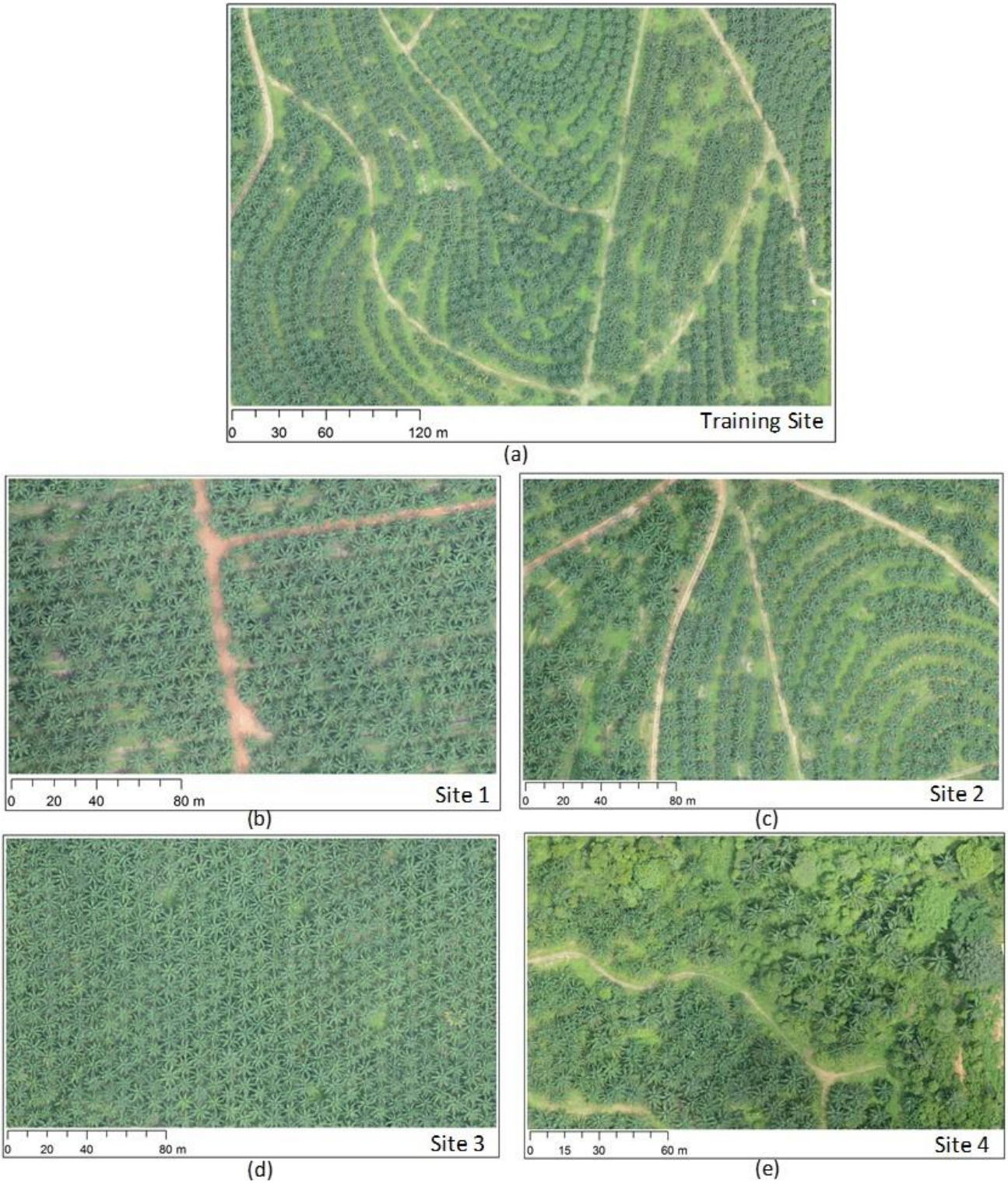


Figure 1. UAV images acquired in Malaysia. (a) Training area from which the manually interpreted training samples were collected. (b), (c), (d) and (e) Four areas used for independent validation of the performance of the proposed method.

3. HOG-SVM Method for Palm Tree Detection from UAV Images

3.1. Overview of the approach

Figure 2 presents a flowchart of the proposed approach. It is a hierarchical method. In the first hierarchy, images are classified into vegetation and non-vegetation classes at the pixel level using a supervised classifier, SVM. In the second hierarchy, individual palm trees are detected in the vegetation areas at the image-patch level. Both hierarchies use a supervised classifier, which requires reference samples for training. Palm tree detection begins with manual selection of positive and negative samples (i.e., palm tree and background image patches, respectively). The HOG method is then applied to extract the image features for all samples. The extracted features are used to train the SVM to classify palm tree and background patches. In the detection process, the SVM classifier moves across all vegetation areas in the image. To detect palm trees of various sizes, the image patch size changes in the detection process. In the final detection results, each detected palm tree is recorded with a circle that describes its centre and canopy diameter. Accuracy assessment is conducted with manually labelled results.

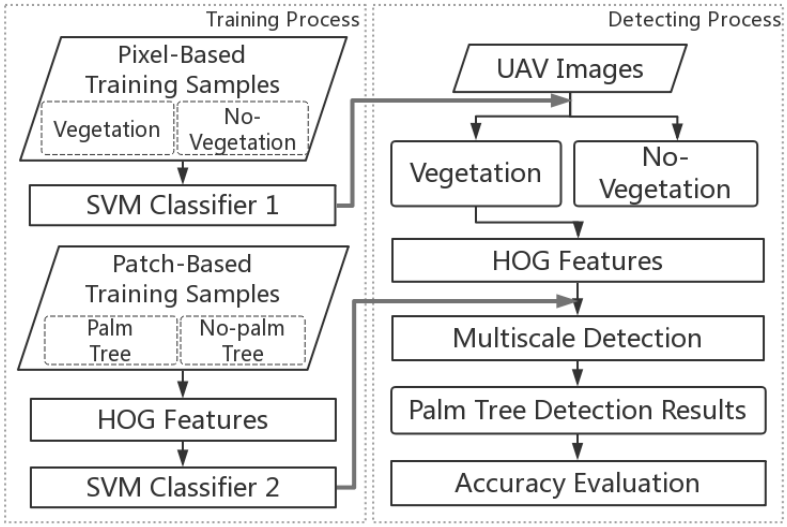


Figure 2. Flowchart of the proposed method for detection of oil palm trees.

3.2. Vegetation and non-vegetation classification

To increase efficiency and avoid the disturbance of non-vegetation pixels in palm tree detection, only the vegetation areas should be processed. Therefore, a classification of vegetation and non-vegetation is needed. An SVM classifier is used for this classification task in this study. Initially conceived by Cortes and Vapnik (1995), SVM is used to solve binary classification problems. Their goal is to determine an optimal hyperplane $h(x)$; they can not only separate two class labels of training samples, but also determine this hyperplane in a way that would make it as far as possible from the closest members of both classes. SVMs are particularly appealing in the remote sensing field due to their ability to generalise well, even with limited training samples, which is a common limitation of remote sensing applications (Adelson et al., 1984). They have been widely used in land cover classification studies (Pouteau and Collin, 2013; Zhu and Liu, 2014). C-support vector classification with a radial basis function kernel is used in this step. The parameters C, P and gamma are set to 100, 0 and 1, respectively, after optimisation.

The vegetation and non-vegetation training samples for SVM were selected by visual interpretation of the UAV image. The trained SVM classifier identifies non-vegetation pixels well (see example in Figure 3). The classification maps of the five pilot areas are shown in Figures 6 through 10. The classification result works as a binary mask in the subsequent detection process. The rule to exclude non-vegetation pixels is defined as follows. When the image patch moves through the entire scene, a small circle is drawn in the centre with a diameter of one third the length of the image patch. If the circle has a non-vegetation area of more than 20%, the image patch is skipped during the detection process.

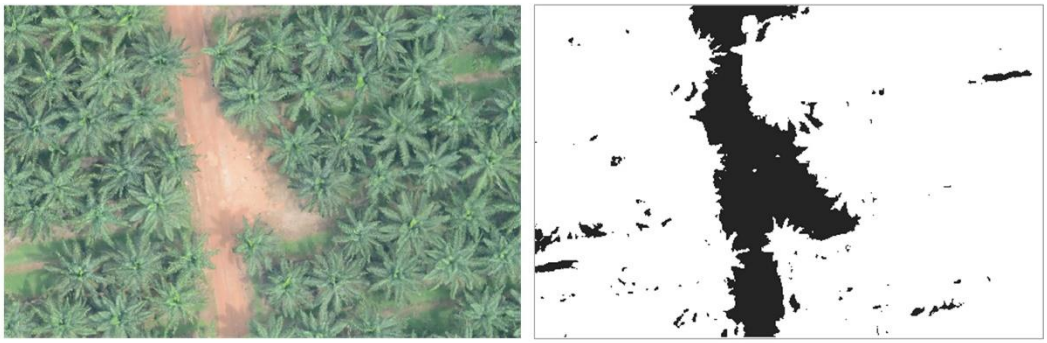
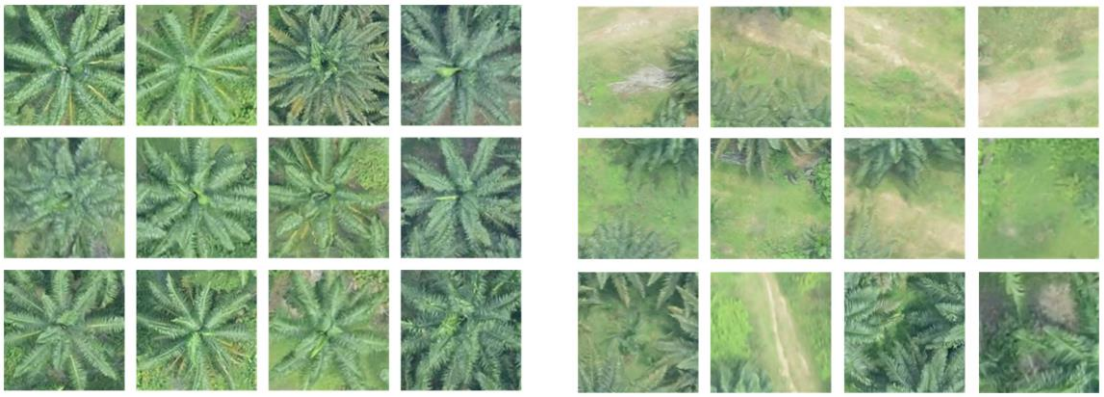


Figure 3. (a) Original UAV image with RGB bands. (b) Classification result. Regions in white indicate vegetation areas, and those in black denote non-vegetation areas.

3.3. Training sample selection for palm tree detection

Positive samples (i.e., image patches that contain a complete palm tree) and negative samples (i.e., image patches that do not contain an entire palm tree or other vegetation) are manually selected. Figure 4 shows some examples of positive and negative samples. The palm tree samples manually selected from the UAV images are cropped and adjusted to the same size (i.e., 64×64 pixels, an appropriate size to reveal the texture of palm trees, as shown in Figure 4 (a), and a favourable size for efficient extraction of features in the next step). In our approach, positive samples are carefully selected as a square which just contains one complete palm tree. The strict cropping of palm trees in positive samples can bring two benefits: 1) the extracted features in the image patch completely represent individual palm trees, and 2) the size of the image patch represents the crown size of the palm trees. For the negative samples, the size of the image patch is the same with positive samples. 500 positive samples and 1000 negative samples were manually selected from the site in Figure 1 (a). To increase the number of samples, both positive and negative samples were rotated and mirrored to generate new samples. Finally, a total of 2,500 positive and 5,000 negative samples were generated for this study. The HOG algorithm was then applied to extract features for the training samples.



(a) (b)

Figure 4. (a) Positive samples (i.e., image patches that contain a complete palm tree) and (b) negative samples (i.e., image patches that do not contain an entire palm tree or other vegetation). Each image patch is 64×64 pixels.

3.4. HOG feature descriptor for palm trees

Dalal and Triggs (2005) proposed the HOG algorithm as an image feature descriptor based on gradient direction. The basic idea of HOG is to calculate the histogram of the oriented gradient in local image patches. The steps of the HOG algorithm are explained as follows. First, the colour images with RGB bands are converted into one-band grey images. The gradient components of the images can then be calculated using the one-dimensional centred method in horizontal and vertical directions. The formula for the calculation is:

$$G_x(x,y) = H(x+1,y) - H(x-1,y) \quad (1)$$

$$G_y(x,y) = H(x,y+1) - H(x,y-1) \quad (2)$$

where $H(x,y)$ is the pixel value and $G_x(x,y)$ and $G_y(x,y)$ denote gradients in the vertical and horizontal directions of pixel (x,y), respectively. The gradient magnitude $\mathrm{G(x,y)}$ and the gradient direction $\alpha(\mathrm{x,y})$ of pixel (x, y) can then be calculated as follows:

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \quad (3)$$

$$\alpha(x,y) = tan^{-1}\left(\frac{G_y(x,y)}{G_x(x,y)}\right) \quad (4)$$

We first divide the image patch of the samples (64×64 pixels) into small cells of equal size (8×8 pixels) and then calculate the gradient histogram of the pixels in each cell by dividing the orientations (0 to 180 degrees) into nine bins. Instead of using the gradient of each pixel as a feature, calculation of a nine-bin histogram for each cell makes the representation more compact and more robust to noise. A 9×1 vector can be used to describe the histogram for each cell. Figure 5 shows the extracted HOG features in small cells for positive and negative samples. The gradient histogram for each cell is shown by white lines whose length and angle indicate the gradient magnitude and direction, respectively. In the positive sample, it shows that the direction of the largest gradient is generally perpendicular to the edge of the leaves and thus forms an obvious pattern of concentric circles. The negative sample shows an irregular distribution of gradient histograms.

Meanwhile, to make the descriptor independent of illuminative variation, the HOG algorithm takes a plurality of cells to compose a block of 16×16 pixels. A 16×16 block has four histograms that can be concatenated to form a 36×1 element vector. The block window is then moved by steps (8 pixels per step) to normalise the histograms, which generates a normalised 36×1 vector for each movement. The normalised 36×1 vectors of all movements are then concatenated into one giant vector as the final HOG features for each image patch.

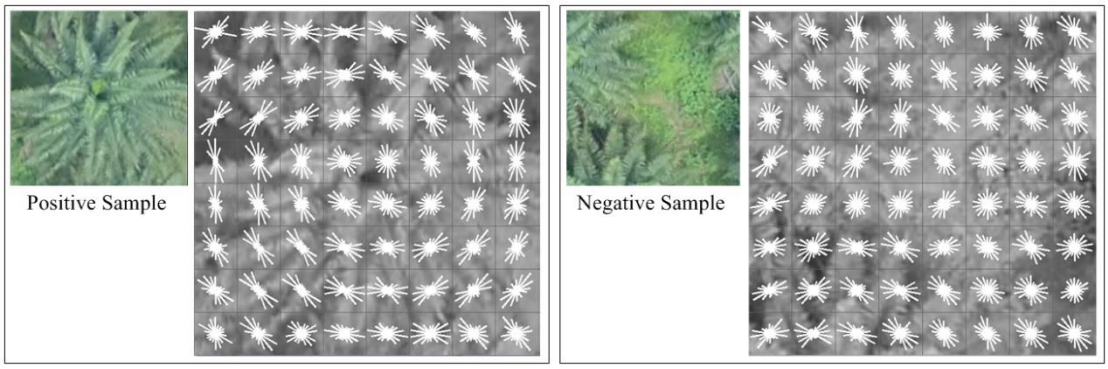


Figure 5. Extracted HOG features for positive samples and negative samples

3.5. Multiscale detection of palm trees by SVM

An SVM classifier is then trained using the HOG features of the positive and negative samples. In this study, C-support vector classification with linear kernel, as suggested in a previous study for object detection (Llorca et al., 2013), was used to detect palm trees. The parameters C, P and gamma were set to 2.5, 0 and 1.0 after optimisation.

To detect the palm trees in the unlabelled UAV images, the search window (i.e., image patch) was moved across the image, and each search window was checked by the trained SVM classifier. To avoid processing the non-vegetation areas, the vegetation classification map (described in Section 3.2) was used as a binary mask. Areas marked as non-vegetation were skipped during the moving window classification. To detect palm trees of various sizes, the image pyramid method (Adelson et al., 1984) was used for multiple-scale detection. The original UAV images were up-scaled to a sequence of layers with various spatial resolutions. The scale difference between two neighbouring layers in this sequence was 1.1. Each layer of the image pyramid was then traversed by a square searching window of 64×64 pixels, the same size as the training samples. One palm tree might be detected by SVM in multiple layers, and the average scale of these layers was used to decide the size of the palm tree. For example, if a palm tree was detected in two

layers with scales of 1.13 and 1.14 to the original image, the diameter of this palm tree was 64×0.04×(1.13+1.14)/2 m, where 0.04 refers to the pixel resolution in metres per pixel. The output of the detection was a set of circles that described the location and size of individual palm trees. The inscribed circle of the squares described the crown of the detected palm trees. Considering the reasonable range of palm tree diameters in the study area, detection results that included diameters larger than 14 m or smaller than 4.5 m were filtered out.

3.6. Accuracy assessment

The actual palm trees in the UAV images were manually labelled by visual inspection. The producer's accuracy (P_{acc}), user's accuracy (U_{acc}) and overall accuracy are used to evaluate the performance of the proposed method:

$$P_{acc} = \frac{TP}{TP+FN} = \frac{TP}{N_{palm}} \tag{8}$$

$$U_{acc} = \frac{TP}{TP+FP} \tag{9}$$

$$Overall\ accuracy = \frac{P_{acc}+U_{acc}}{2} \tag{10}$$

where N_{palm} is the total number of palm trees in the image; TP is the total number of palm trees correctly detected; FN is the false-negative detections, indicating the number of palm trees not detected (i.e., omission errors); and FP is the false-positive detections, indicating the number of non-palm tree objects that were detected as palm trees (i.e., commission errors).

4. Experimental Results

For the site where the proposed method was trained (Figure 1 a), the palm trees were successfully detected (Figure 6). The producer's, user's and overall accuracy rates

reached 99.25%, 99.16% and 99.21%, respectively. To test the effectiveness and adaptability of the proposed method for palm tree detection, the classifier trained was directly applied to four validation sites (Figures 1 (b-e)). Figures 7, 8, 9 and 10 show the detection results of these four validating sites. Yellow circles indicate correct detection (i.e., the detected palm trees are actual palm trees), blue circles indicate missed detection (i.e., undetected palm trees) and red circles indicate false detection (i.e., detection of other objects as palm trees). In these figures, the enlarged regions highlight some missed and false detection. For the training site and four testing sites, the vegetation classification successfully distinguished vegetation and bare soil. Manual checking of the detection results shows that most of the palm trees were successfully detected and that their diameters are also accurate. Few false-positive detections occurred on the grassland or in the gaps between palm trees (see the enlarged regions in Figures 6 through 9). This might have occurred when a portion of the palm tree's leaves formed a star-shaped object in the search window, which had a gradient histogram similar to that of individual palm trees. Most false-negative detections occurred when the palm trees had a different shape than typical palm trees (see the enlarged region in Figure 6) or were partly covered by other trees (see the enlarged region in Figure 10). In some extreme cases, the size of a palm tree was overestimated, as shown in the enlarged region of Figure 8.

From these detection results, we also obtained the distribution of tree size (see the histograms in Figures 6 through 10), which helped us to determine the trees' growing stages. For example, site 3 had more small palm trees than the other sites, suggesting that most of the palm trees at this site were young trees that required more attention from the grower in fertilisation and weed removal (Corley et al., 2008).

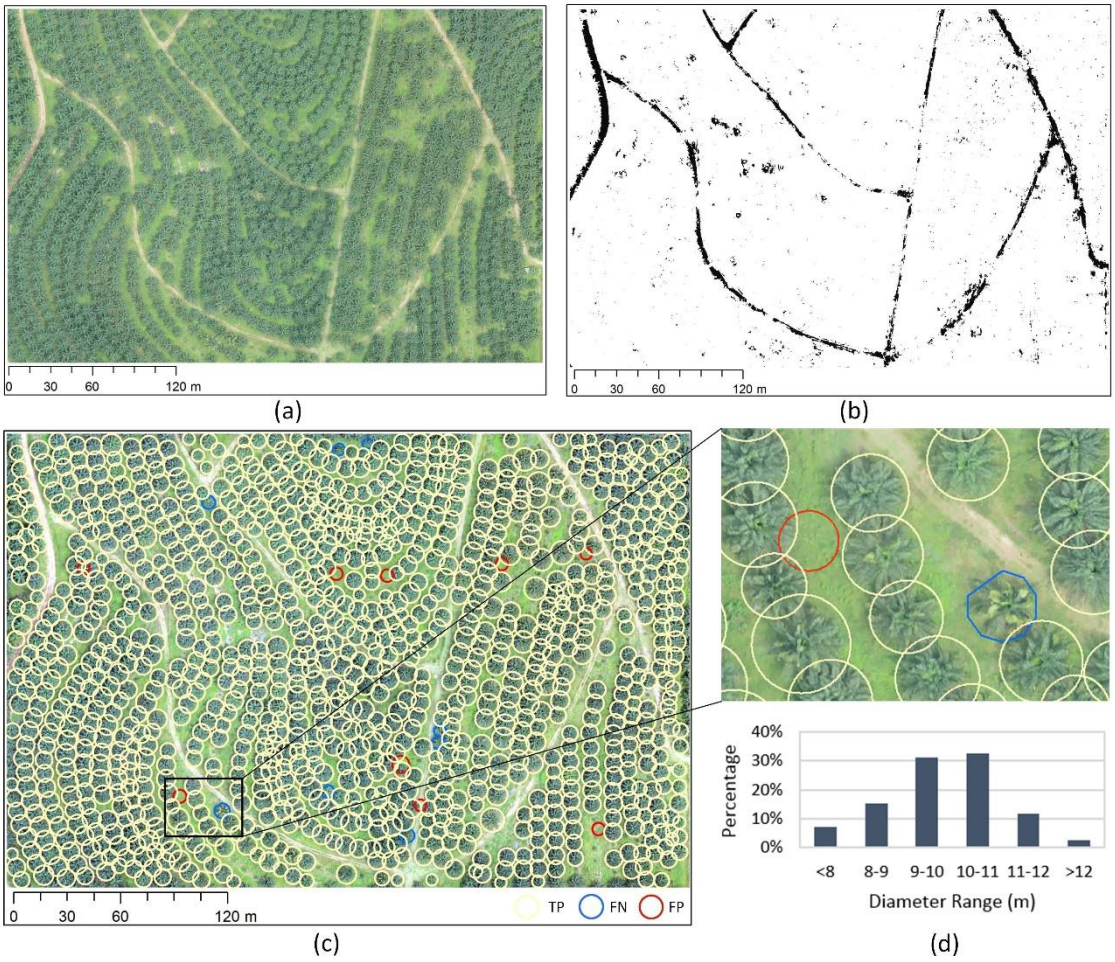


Figure 6. Palm tree detection at the training site. (a) Original UAV image. (b) Vegetation and non-vegetation classification results. (c) Palm tree detection results. (d) Histogram of diameter distribution.

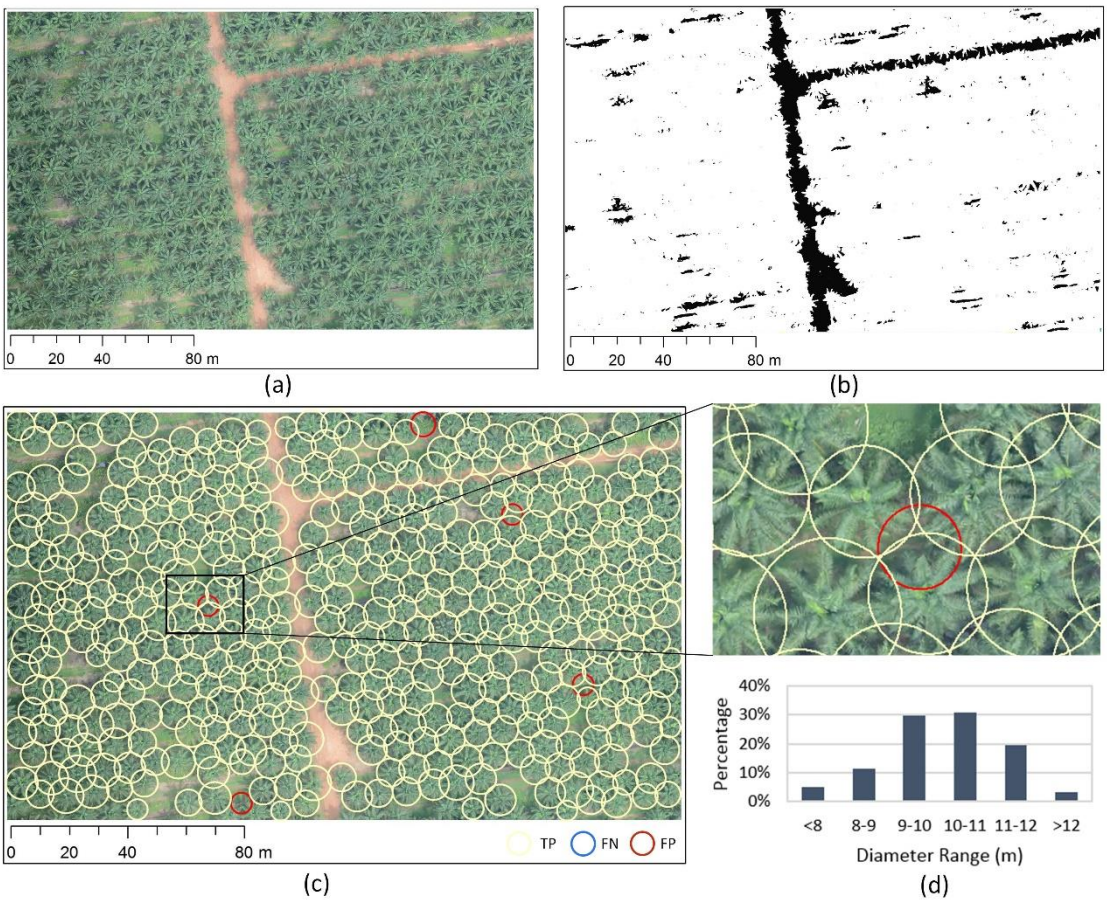


Figure 7. Palm tree detection at validation site 1. (a) Original UAV image. (b) Vegetation and non-vegetation classification results. (c) Palm tree detection results. (d) Histogram of diameter distribution.

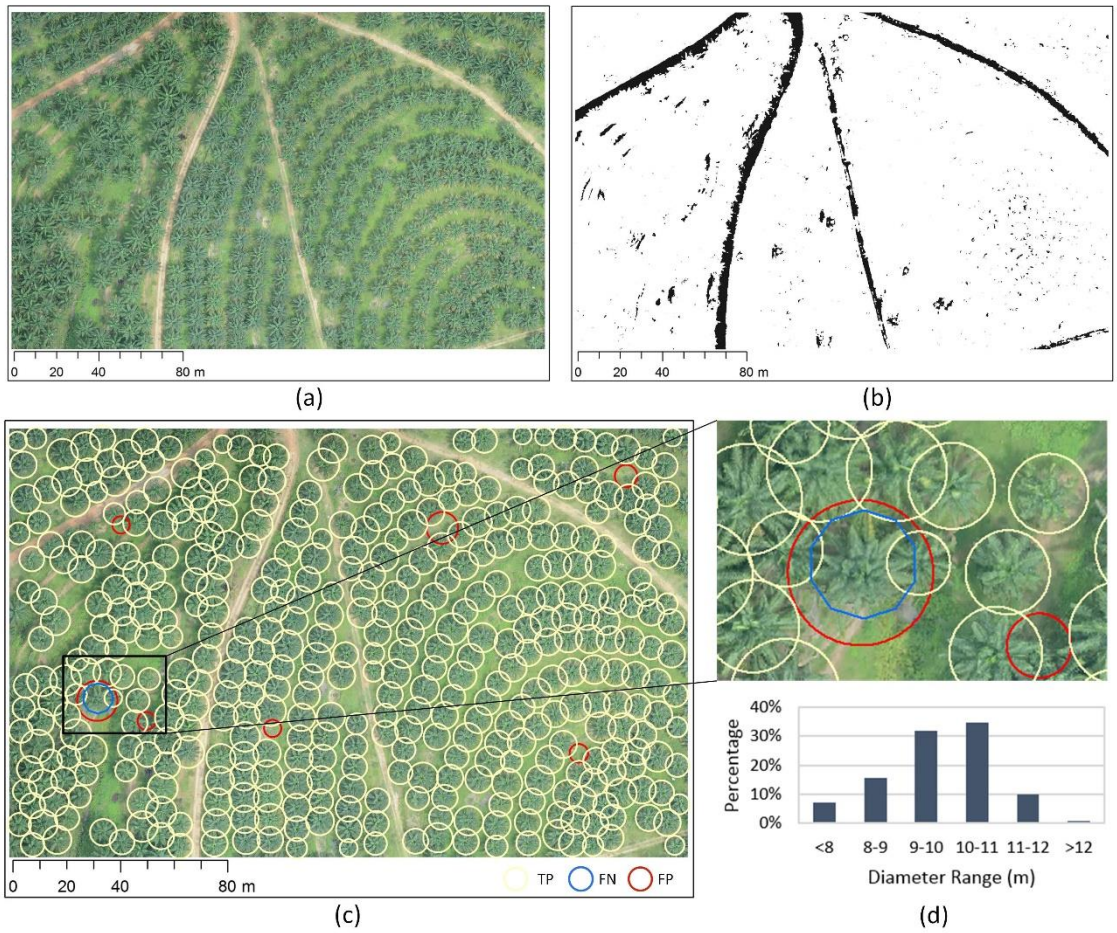


Figure 8. Palm tree detection at validation site 2. (a) Original UAV image. (b) Vegetation and non-vegetation classification results. (c) Palm tree detection results. (d) Histogram of diameter distribution.

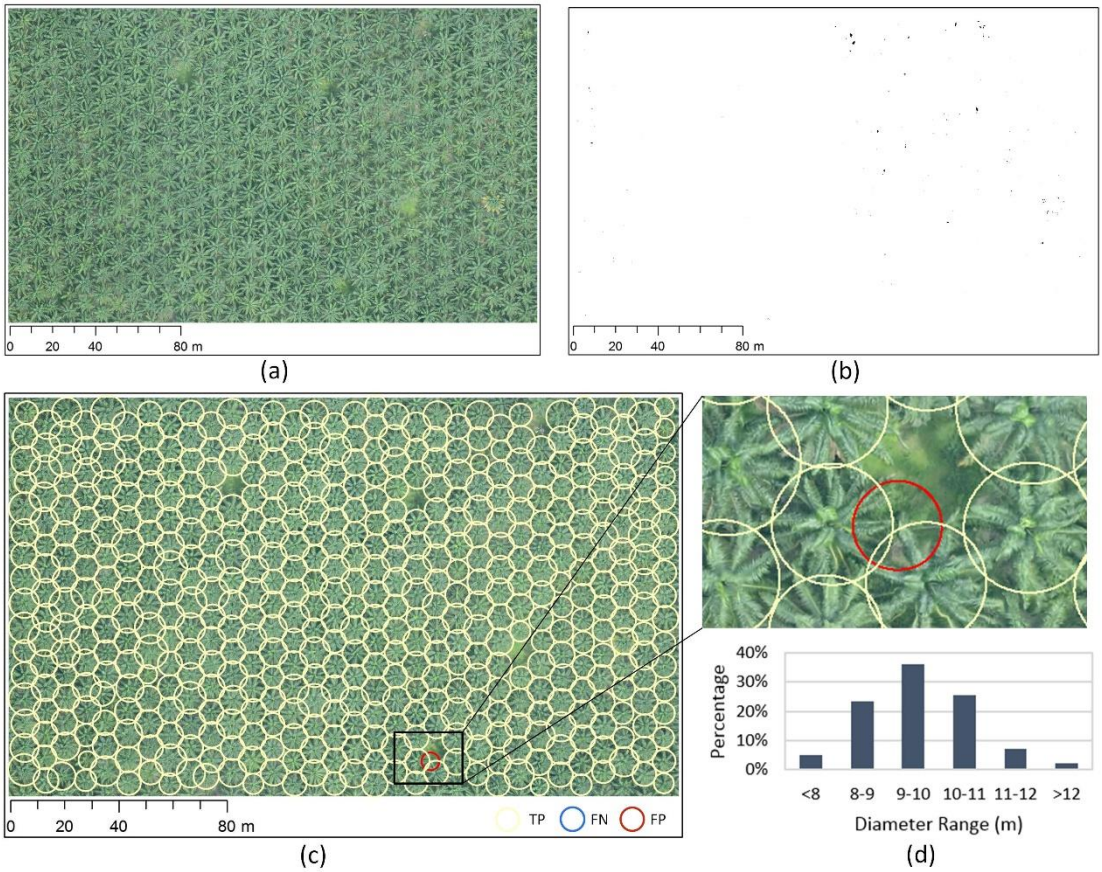


Figure 9. Palm tree detection at validation site 3. (a) Original UAV image. (b) Vegetation and non-vegetation classification results. (c) Palm tree detection results. (d) Histogram of diameter distribution.

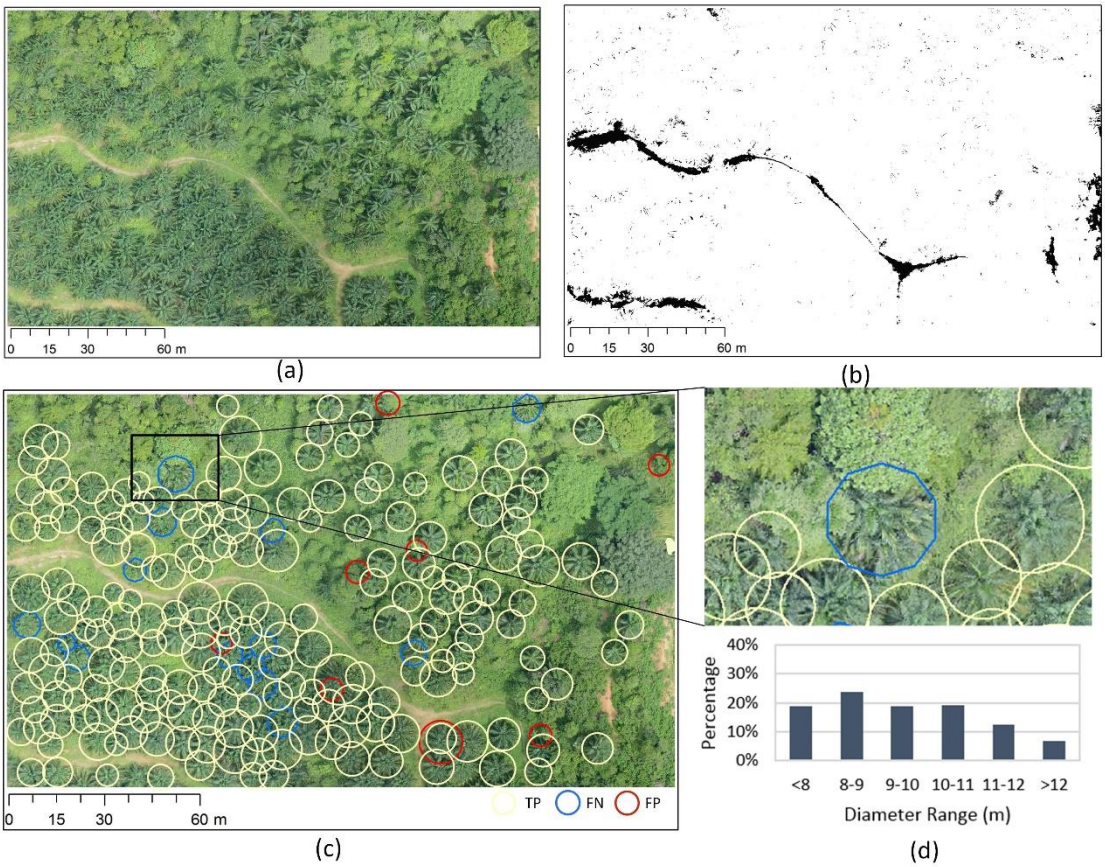


Figure 10. Palm tree detection at validation site 4. (a) Original UAV image. (b) Vegetation and non-vegetation classification results. (c) Palm tree detection results. (d) Histogram of diameter distribution.

Table 1 shows the accuracy assessment of the training site and four validation sites. Overall, 2,590 palm trees were detected by the proposed method. The overall accuracy rate was 99.21% for the training site and 99.39%, 99.06%, 99.90% and 94.63% for the four testing sites, respectively. This proposed method performed well in various scenarios from relatively easy to tough cases. Validation site 4 was particularly challenging for palm tree detection because the palm trees were mixed with or even partly covered by other tree species. Although the accuracy rate of 94.63% was not as good as those at the other three sites, it was still satisfactory and higher than the accuracy rates reported in previous studies (Malek et al., 2014; Manandhar et al., 2016). This

complicated situation is not very common and exists only on the boundaries between palm tree farmland and natural forestland.

Table 1. Accuracy assessment of training site and four testing sites for palm tree detection.

	TP	***FN***	***FP***	P_{acc}	U_{acc}	***Overall accuracy***
Training	1,063	8	9	99.25%	99.16%	99.21%
Site1	404	0	5	100.00%	98.78%	99.39%
Site2	421	1	7	99.76%	98.36%	99.06%
Site3	509	0	1	100.00%	99.80%	99.90%
Site4	193	14	8	93.24%	96.02%	94.63%

5. Conclusions and Discussion

In this study, we proposed a novel automatic method for detection of individual oil palm trees from UAV imagery. Considering the unique texture of palm trees, this method used the HOG algorithm to extract features that were capable of describing the texture of palm trees and then used an SVM classifier to implement the binary classification task. The proposed method was trained with images from one site and was then applied to four other representative sites for independent accuracy assessment. The results show the effectiveness and robustness of our proposed method for detection and counting of palm trees. The overall accuracy rates were 99.21% for the training site and 99.39%, 99.06%, 99.90% and 94.63% for the four testing sites, respectively. Even for the most challenging site, the accuracy of palm tree detection was still higher than that reported for the two methods proposed by Malek et al. (2014) and Manandhar et al. (2016), respectively. The principle of this method is relatively simple, so it has great potential for palm tree detection over large areas.

For future development, detection accuracy could be further improved by enlarging the training sample datasets to obtain more representative images, which is a critical factor in machine-learning-based classification. In addition, we believe this approach can be applied to the detection of other types of trees (e.g., pine trees or some broadleaf trees) because the HOG-SVM algorithm is able to automatically select meaningful features for classification problems.

Acknowledgments

The study described in this paper was funded by grants from the National Natural Science Foundation of China (Project Nos. 41671426 and 41701378) and from the Hong Kong Polytechnic University (Project Nos. G-YBN8 and 4-BCE5).

References

Adelson, E.H., Anderson, C.H., Bergen, J.R., Burt, P.J., Ogden, J.M., 1984. Pyramid methods in image processing. RCA Engineer 29, 33-41.

Berni, J.A., Zarco-Tejada, P.J., Suárez, L., Fereres, E., 2009. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Transactions on Geoscience and Remote Sensing 47, 722-738.

Corley, R., Hereward V., Tinker, P., 2008. The oil palm. John Wiley & Sons, pp. 89-127.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning 20, 273-297.

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, pp. 886-893.

Jusoff, K., Pathan, M., 2009. Mapping of individual oil palm trees using airborne hyperspectral sensing: An overview. Applied Physics Research 1, 15.

Korom, A., Phua, M., Hirata, Y., Matsuura, T., 2014. Extracting oil palm crown from WorldView-2 satellite image, IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 012188.

Li, W., Fu, H., Yu, L., Cracknell, A., 2016. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. Remote Sensing 9, 22.

Llorca, D.F., Arroyo, R., Sotelo, M., 2013. Vehicle logo recognition in traffic images using HOG features and SVM, Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on. IEEE, pp. 2229-2234.

Malek, S., Bazi, Y., Alajlan, N., Al Hichri, H., Melgani, F., 2014. Efficient framework for palm tree detection in UAV images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 4692-4703.

Manandhar, A., Hoegner, L., Stilla, U., 2016. Palm tree detection using circular autocorrelation of polar shape matrix. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 465-472. 3, p.465.

Pouteau, R., Collin, A., 2013. Spatial location and ecological content of support vectors in an SVM classification of tropical vegetation. Remote Sensing Letters 4, 686-695.

Srestasathiern, P., Rakwatin, P., 2014. Oil palm tree detection with high resolution multi-spectral satellite imagery. Remote Sensing 6, 9749-9774.

Uto, K., Seki, H., Saito, G., Kosugi, Y., 2013. Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6, 851-860.

Xiang, H., Tian, L., 2011. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). Biosystems Engineering 108, 174-190.

Yao, C., Wu, F., Chen, H.-J., Hao, X.-L., Shen, Y., 2014. Traffic sign recognition using HOG-SVM and grid search, Signal Processing (ICSP), 2014 12th International Conference on. IEEE, pp. 962-965.

Zarco-Tejada, P.J., Guillén-Climent, M., Hernández-Clemente, R., Catalina, A., González, M., Martín, P., 2013. Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). Agricultural and Forest Meteorology 171, 281-294.

Zhu, X., Liu, D., 2014. Accurate mapping of forest types using dense seasonal Landsat time-series. ISPRS Journal of Photogrammetry and Remote Sensing 96, 1-11.