



Article

A New Machine Learning Approach in Detecting the Oil Palm Plantations Using Remote Sensing Data

Kaibin Xu ^{1,2}, Jing Qian ^{1,3,4,*} , Zengyun Hu ^{1,4,5} , Zheng Duan ⁶ , Chaoliang Chen ^{1,3,4,5}, Jun Liu ⁷, Jiayu Sun ¹, Shujie Wei ¹ and Xiuwei Xing ^{1,3,4}

- ¹ Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; kb.xu@siat.ac.cn (K.X.); huzengyun@ms.xjb.ac.cn (Z.H.); cl.chen@siat.ac.cn (C.C.); sun.jy@siat.ac.cn (J.S.); sj.wei@siat.ac.cn (S.W.); xw.xing@siat.ac.cn (X.X.)
- ² School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China
- ³ University of Chinese Academy of Sciences, Beijing 100049, China
- ⁴ State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
- ⁵ Research Center for Ecology and Environment of Central Asia, CAS, Urumqi 830011, China
- ⁶ Department of Physical Geography and Ecosystem Science, Lund University, S-223 62 Lund, Sweden; zheng.duan@nateko.lu.se
- ⁷ TripleSAI Technology, Shenzhen 518109, China; rsluijun@163.com
- * Correspondence: jing.qian@siat.ac.cn; Tel.: +86-189-264-071-98

Abstract: The rapid expansion of oil palm is a major driver of deforestation and other associated damage to the climate and ecosystem in tropical regions, especially Southeast Asia. It is therefore necessary to precisely detect and monitor oil palm plantations to safeguard the ecosystem services and biodiversity of tropical forests. Compared with optical data, which are vulnerable to cloud cover, the Sentinel-1 dual-polarization C-band synthetic aperture radar (SAR) acquires global observations under all weather conditions and times of day and shows good performance for oil palm detection in the humid tropics. However, because accurately distinguishing mature and young oil palm trees by using optical and SAR data is difficult and considering the strong dependence on the input parameter values when detecting oil palm plantations by employing existing classification algorithms, we propose an innovative method to improve the accuracy of classifying the oil palm type (mature or young) and detecting the oil palm planting area in Sumatra by fusing Landsat-8 and Sentinel-1 images. We extract multitemporal spectral characteristics, SAR backscattering values, vegetation indices, and texture features to establish different feature combinations. Then, we use the random forest algorithm based on improved grid search optimization (IGSO-RF) and select optimal feature subsets to establish a classification model and detect oil palm plantations. Based on the IGSO-RF classifier and optimal features, our method improved the oil palm detection accuracy and obtained the best model performance (OA = 96.08% and kappa = 0.9462). Moreover, the contributions of different features to oil palm detection are different; nevertheless, the optimal feature subset performed the best and demonstrated good potential for the detection of oil palm plantations.

Keywords: oil palm detection; Landsat; Sentinel; land cover classification; random forest



Citation: Xu, K.; Qian, J.; Hu, Z.; Duan, Z.; Chen, C.; Liu, J.; Sun, J.; Wei, S.; Xing, X. A New Machine Learning Approach in Detecting the Oil Palm Plantations Using Remote Sensing Data. *Remote Sens.* **2021**, *13*, 236. <https://doi.org/10.3390/rs13020236>

Received: 3 December 2020

Accepted: 4 January 2021

Published: 12 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Oil palm (*Elaeis guineensis*), whose planting areas are distributed mainly in humid tropical countries such as Indonesia, is one of the most rapidly expanding and productive equatorial crops in the world [1]. Because this crop has multiple uses, high yields, and low production costs, the global demand for palm oil has increased exponentially over the last few decades, generating considerable economic benefits in local areas [2]. However, the rapid expansion of oil palm plantations has also led to deforestation and a series of negative environmental impacts, such as forest estate losses, social costs, alternative revenue losses, reduced biodiversity, and diminished ecological connectivity [3]. In addition, oil palm

plantations are a substantial and frequent cause of fires in Indonesia, many palm oil producers take advantage of the conditions to clear vegetation for oil palm plantations using the slash-and-burn method, they often spin out of control and spread into protected forested areas, and these fires emit increasing quantities of greenhouse gases that threaten the global climate and ecosystem [4–6]. Therefore, to scientifically manage and supervise this activity and to safeguard forests beneficial for the global climate and ecosystem services, it is necessary to precisely detect and monitor oil palm plantations.

The detection of oil palm plantations using satellite remote sensing data has been carried out in many studies [7–10]. Optical detection methods rely on information extracted from the phenology or image characteristics of oil palm plantations. Phenology-based methods utilize temporal changes in the vegetation spectrum to detect the expansion of oil palm [7]. In addition, oil palm plantations can be detected from satellite images based on their unique textural features, such as the rectangular blocks and geometric shape of industrial plantations [8]. For example, employing an image-based method, Li et al. used the texture features of oil palm trees trained by neural networks to identify them from high-resolution remote sensing images [9,10]. However, some challenges continue to face the detection of oil palm plantations using optical methods; for instance, it remains difficult to separate oil palm plantations from other spectrally similar vegetation (e.g., forests and rubber trees) [11], and the frequent presence of clouds in the humid tropics hinders image-based method analysis [12]. In addition, most high-resolution images are not free.

To reduce the difficulty of detecting oil palm plantations in tropical regions, synthetic aperture radar (SAR), which provides global observations under all weather conditions and all times of day, has been the focus of some researchers [13–15], who employed radar satellite data, namely, L-band data from the Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) and C-band data from Sentinel-1, as the main source for the detection of oil palm plantations. With radar satellite data, oil palm plantations present a characteristic radar backscatter distribution, which can be easily separated from those of other tropical plantations [13]. However, due to the similar scattering values for palm trees of different ages, it is difficult to distinguish mature and young (<3-year-old) oil palm plantations using only SAR data [16].

To overcome the limitations of using SAR or optical data alone, several recent studies have detected oil palm plantations by using data fusion techniques [17–20]. These studies selected specific backscatter values and reflectance/emissivity characteristics from SAR and optical satellite combinations to identify oil palm plantations (including mature and young oil palm) and other land use types using specific machine learning algorithms. For example, Cheng et al. fused Landsat and PALSAR data to conduct the supervised classification of oil palm plantations in peninsular Malaysia [17], and Poortinga et al. combined Landsat-8, Sentinel-1, and Sentinel-2 to accurately map rubber and oil palm plantations [18]. The results show that the accuracies from data fusion are better than the accuracies from SAR or optical satellite data alone. In addition, several studies have shown that using the appropriate vegetation indices to analyze and select feature combinations can also yield improved results [21–23]. However, due to the low canopy coverage of young palm trees and the poor ability to differentiate young palm trees from bare soil, the detection of young plantations is still very challenging in most data fusion methods, as was reported in a previous study [24]. Moreover, most studies select only spectral bands and a small number of backscatter properties as the characteristic variables; as a result, the detection accuracy of oil palm plantations has remained at approximately 90% [25]. In addition, the selection of machine learning classifiers is fundamental to guarantee the oil palm detection accuracy. These classifiers include support vector machine (SVM) [26], naïve Bayes (NB) classifiers [27], classification and regression trees (CART) [28], and neural networks [29]. However, these classifiers depend on large amounts of sample information to improve the prediction accuracy, and the acquisition of these samples is considerably time- and labor-intensive. Compared with other classifiers, random forest (RF) has low preprocessing requirements for the training data because it is not sensitive to the differences of data

units and can cope with badly unbalanced data [30], and it can make predictions when an observation presents missing values [31]. Furthermore, several studies have shown that, in the selection of machine learning algorithms, using the RF algorithm to classify and select features can yield better results [32]. However, RF and other machine learning classifiers are strongly dependent on the values of the input hyper-parameters [33,34], and thus the detection accuracy of oil palm plantations might be easily affected by the selection of hyper-parameters.

Therefore, this study aims to develop an innovative method to improve the classification accuracy of the oil palm type (mature or young) and to detect oil palm planting areas by fusing Landsat-8 and Sentinel-1 images in Sumatra, Indonesia. We first extract multi-temporal spectral characteristics, SAR backscattering values, vegetation indices and texture features to integrate feature combinations. Then, we use the random forest algorithm based on improved grid search optimization (IGSO-RF) and select the optimal feature subsets to establish a classification model and detect oil palm plantations.

2. Study Area and Materials

2.1. Study Area

The study area is located in the province of Riau (0.5333°N, 101.4500°E) on the island of Sumatra, Indonesia, with an area of 91,095 km² (Figure 1). Indonesia is one of the largest producers of oil palm in the world, and Riau is the largest producer of the oil palm-producing provinces in Indonesia, accounting for 24% of the national production [35]. In Riau Province, the area of oil palm plantations has reached 2,400,876 ha, of which more than half of all plantations belong to smallholders (57% or 1,354,503 ha), while the remainder (43% or 1,046,373 ha) belong to industry, the government, or private estates. Due to the expansion of smallholder plantations in Riau Province, the overall oil palm plantation area increased by 21% from 2004 to 2009 [35].

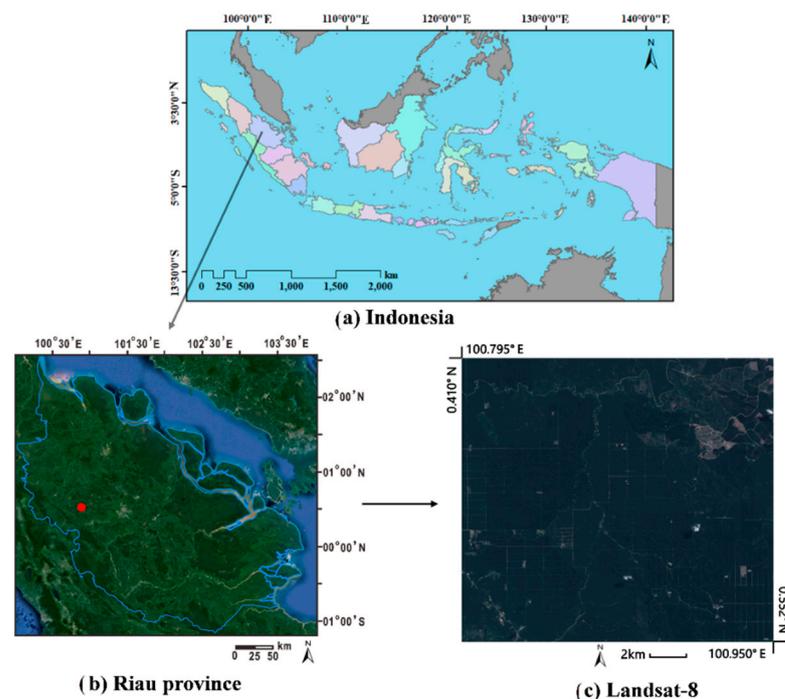


Figure 1. Oil palm plantations in the study area: Riau Province, Indonesia: (a) Overview of Indonesia; (b) Location of Riau province; (c) Landsat-8 composites in 2019 (R-G-B).

2.2. Datasets

This study used all Landsat-8 top-of-atmosphere reflectance (TOA) images and Sentinel-1A data from 1 January 2019 to 31 December 2019 from Google Earth Engine (GEE) platform and

selected the images with minimal cloud cover for the supervised classification of oil palm plantations. The details of the Landsat-8 images and Sentinel-1A data are presented in Table 1.

Table 1. Descriptions of the data used.

Sensor	Landsat-8	Sentinel-1A
Bands	Blue, Green, Red and Near Infrared(B2, B3, B4, B5)	Dual Polarization (VV, VH)
Sensor Type	Thermal Infrared Sensor (TIRS), Pushbroom	S1 Ground Range Detected Scenes
Spatial Resolution	30 m	10 m
Product Type	Top-of-Atmosphere Reflectance Images	Ground Range Detected Image

2.3. Training Data Collection

To accurately detect and map the oil palm plantations over Riau Province in 2019, the training dataset was based on 3000 data points collected from the visual interpretation of 2019 Landsat-8 images over Riau Province, and validated on high-resolution Google Earth images. We first identified the area of different land-cover types from images based on a previous study [24]. Then, for each land-cover type, we used the criteria of random sampling to obtain sample points, and the number of points were based on the area. The classes of points were visually distinguished as follows: 1. mature oil palm plantations (750 points), 2. young oil palm plantations (917 points), 3. bare land (388 points), and 4. other land uses that are not oil palm plantations (945 points).

Then, we plotted the points on the Sentinel-1 and Landsat-8 composite images to extract the features for the training datasets. The datasets were subdivided into 20% for training and 80% for validation by using fivefold cross-validation.

3. Methods

3.1. Overview

The detection of oil palm plantations starts with the compositing of Sentinel-1 images and compositing of Landsat-8 images in GEE platform, which were based on the median values of the daily Landsat-8 and Sentinel-1 images from 1 January 2019 to 31 December 2019, respectively. Then, the Landsat-8 composite images were resampled in GEE platform, the Sentinel-1 and Landsat-8 composites were combined into a single composite image for the classification process. In addition, multitemporal spectral bands, SAR backscattering values and feature extraction were employed to generate additional vegetation indices and texture features to improve the classification model. Then, the IGSO-RF model and feature selection were applied to establish a classification model for oil palm detection with four groups of feature combinations (Table 2): (1) Landsat-8 spectral bands, (2) Landsat-8 spectral bands and Sentinel-1 backscatter values, (3) spectral bands and backscatter values with vegetation indices and texture features, and (4) an optimal subset of all bands and features in group III. Figure 2 shows a schematic representation of the proposed method.

Table 2. Groups of feature combinations.

Group	Feature Combination
I	Spectral bands
II	Spectral bands and backscatter values
III	Spectral bands, backscatter values, vegetation indices and texture features
IV	Optimal subset of all bands and features

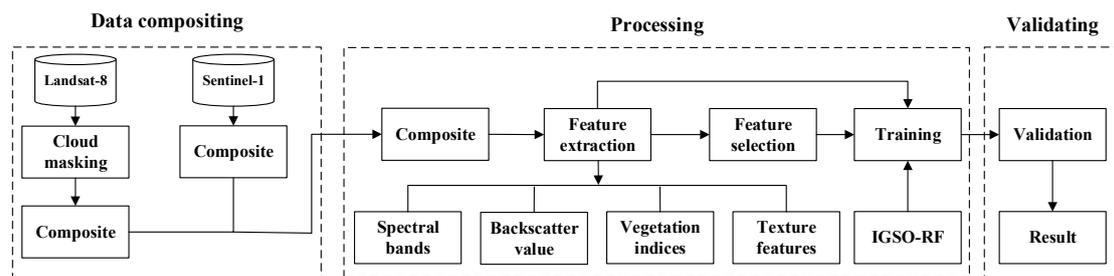


Figure 2. Diagram of the algorithm for oil palm detection.

This study was conducted in GEE platform, a cloud-based computing environment that includes access to the full archive of Landsat and Sentinel imagery [36]. This combination of a large data repository of satellite imagery with a computational platform enables scientists to conduct research on environmental issues at a variety of spatial and temporal scales.

3.2. Sentinel-1 and Landsat-8 Compositing

The Sentinel-1 composited images and Landsat-8 composited images were based on the median values of the daily Sentinel-1 and Landsat-8 images from 1 January 2019 to 31 December 2019, respectively. The images we used were automatically corrected by the GEE platform, and the process of compositing was also performed in GEE platform. For the Sentinel-1 composite, we used two backscatter bands: the single co-polarization VV (vertical transmit/vertical receive) band and the dual cross-polarization VH (vertical transmit/horizontal receive) band at 10 m. For the Landsat-8 composite, four spectral bands were used containing the reflectances in the blue, green, red, and near-infrared bands at 30 m. Additionally, in our case, in order to verify the possibility and ability of SAR images and our model in oil palm plantations detection under the high cloudy areas, we chose the area with high cloud covered from Landsat-8 composited image, and we roughly masked the clouds with a cloud-score algorithm provided in GEE platform [37]. The cloud-score algorithm uses the spectral and thermal properties of clouds to identify and remove pixels with cloud cover from the imagery [18]. The algorithm finds bright and cold pixels and uses Normalized Difference Snow Index (NDSI) to compare the spectral properties of snow and prevent snow from being masked, and the algorithm uses the visible, near-infrared, and shortwave infrared for a scaled cloud-score and then takes the minimum, which can remove pixels with cloud cover as much as possible.

After compositing the Sentinel-1 images and Landsat-8 images respectively, we combined them into a single image for the further classification process. Since GEE does all its computations at a given scale, regardless of the spatial resolution of the original image, we computed all the composite band images at 10 m.

3.3. Feature Extraction

The extraction and selection of features are processes that generate and select, respectively, a set of informative variables (features) from the original dataset to improve the accuracy of the classification model. This study employed four types of features: multi-temporal spectral characteristics, SAR backscatter values, vegetation indices, and texture features. Table 3 shows the different types of features, the input bands, and the computing formulas used in the extraction. Table 4 shows the features used in different groups of features combination.

Table 3. Feature variables and calculations.

Feature Group	Feature Variables	Input Bands or Calculation	Reference
Spectral Bands	Blue	B_2	
	Green	B_3	
	Red	B_4	
	Near Infrared	B_5	
SAR Backscatter	VV Polarization	VV	
	VH Polarization	VH	
	Difference Ratio	$VV - VH$ VV/VH	
	Normalized Difference Index (NDI)	$NDI = (VV - VH)/(VV + VH)$	[38]
Vegetation Indices	Difference Vegetation Index (DVI)	$DVI = B_5 - B_4$	[39]
	Ratio Vegetation Index (RVI)	$RVI = B_5/B_4$	[40]
	Greenness Index (GI)	$GI = B_3/B_2$	
	Normalized Difference Vegetation Index (NDVI)	$NDVI = (B_5 - B_4)/(B_5 + B_4)$	[41]
	Enhanced Vegetation Index (EVI)	$EVI = 2.5(B_5 - B_4)/(B_5 + 6.0B_4 - 7.5B_2 + 1)$	[42]
	Soil-Adjusted Vegetation Index (SAVI)	$OSAVI = (1 + 0.16)(B_5 - B_4)/(B_5 + B_4 + 0.16)$	[43]
Texture Features	Contrast (CON)	$CON = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - j)^2 \cdot p(i, j, d, \theta)$	[44]
	Angular Second Moment (ASM)	$ASM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j, d, \theta)^2$	[44]
	Entropy (ENT)	$ENT = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [p(i, j, d, \theta) \log p(i, j, d, \theta)]$	[44]
	Correlation (COR)	$COR = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [ijp(i, j, d, \theta) - \mu_1\mu_2]}{\sigma_1^2\sigma_2^2}$	[44]

Table 4. Feature variables in different groups of features combination.

Group	Feature Variables	Numbers of Features
I	Blue, Green, Red, Near Infrared	4
II	Blue, Green, Red, Near Infrared, VV, VH, Difference, Ratio, NDI	9
III	Blue, Green, Red, Near Infrared, VV, VH, Difference, Ratio, NDI, DVI, RVI, GI, NDVI, EVI, SAVI, CON, ASM, ENT, COR	19
IV	Most Relevant features	15

Multitemporal spectral features, which contain the most critical and direct information of images, are an important and direct basis for distinguishing and classifying various types of ground objects in remote sensing images. The reflectances in the blue, green, red, and near-infrared bands of nine temporal Landsat-8 data were selected as spectral features.

SAR backscattering values can provide observations under all weather conditions and all times of day, which is beneficial for studies in the tropics, and thus can compensate for the deficiency of spectral features caused by bad weather and clouds [45]. This study used the available polarization bands from Sentinel-1A data, namely, the dual-polarization VV and VH bands. Previous studies demonstrated that oil palm plantations are best separated by using VV backscatter values or the NDI [8,38]. In addition, Miettinen et al. found that the VV-VH backscatter difference for oil palm exhibits a unique histogram [38]. Therefore, this study also selected the NDI and VV-VH backscatter difference to complement the SAR backscatter value.

Vegetation indices, which are combinations of different bands in remote sensing image data, can reflect crop growth, crop structure, soil background and other related information [46]. Based on a relevant study [19] and the spectral features of oil palm in combination with the characteristics of the study area, this study selected six strongly applicable vegetation indices.

Texture features can fully reflect the features of vegetation in an image, and thus have great significance for the feature extraction and analysis of plants in images [47]. This study

derived several texture features from the Landsat-8 spectral bands using the median filter and texture analysis based on the gray-level co-occurrence matrix (GLCM). The specific process includes extracting a gray-level image, quantifying the gray levels, calculating the feature values, and generating a texture feature image. Ulaby et al. discovered that among the multiple texture features based on the GLCM, only the contrast (CON), angular second moment (ASM), correlation (COR), and entropy (ENT) are uncorrelated [48]. Fortunately, these four features are easy to calculate and can provide a high classification accuracy. Therefore, this paper selected these four features to compose the texture feature dataset and set the sliding window dimensions to 6×6 for the extraction of texture features based on the surface feature size and texture roughness in the study area.

3.4. Feature Selection

The importance of feature variables and the optimization of feature groups play important roles in remote sensing image classification. Feature selection can reduce the dimensionality of data, enhance the model generalizability, reduce overfitting, and enhance the relationship between features and values [49]. To rank the most relevant features in the classification model and filter out redundant and noninformative data, the extracted features from the optical and SAR images were analyzed with the Gini coefficient importance method. The Gini coefficient importance [50], which is calculated as the total decrease in node impurity averaged over all RF decision trees, is an implicit method in the RF classifier.

3.5. Random Forest Algorithm and Optimization of Parameters

The RF classification algorithm is an effective machine learning method based on decision tree [31]. In this study, the RF algorithm was selected for its fast computing time in model training and sample prediction, its low requirement for preprocessing the training data, and the ability to predict data when observations are missing.

RF is an ensemble machine learning algorithm that involves several decision trees $T = \{T_1(x), T_2(x), \dots, T_k(x)\}$. In the process of constructing the RF decision trees, the first step is to randomly select k samples from the original training dataset D of size k with replacement to generate a new self-service training dataset $D_k^{(B)}$ and construct k decision trees $T_k(x)$. In addition, the samples that are not selected each time constitute k out-of-bag (OOB) data. The second step is to randomly select a group of M features from a set of features in each node among the decision trees $T_k(x)$. Then, the RF tree is constructed by recursively repeating the above steps for each terminal node in the decision tree until the decision tree can accurately identify the training dataset $D_k^{(B)}$ with the minimum node size. During training, since the classification and regression tree (CART) can divide the datasets into two sub-datasets, we used it to split each node of the decision tree by randomly selecting m split features from among the M features, and the Gini coefficient importance method was used to select one of the m split features for the splitting process.

The accuracy of the RF algorithm depends on the hyper-parameters selected during the training process. It is difficult to select the optimal parameters by relying on experience alone. Fortunately, the grid search optimization (GSO) method [51], which searches the grid area of a variable to find the optimal grid point that satisfies the constraint function, has been widely used in the optimization of classification algorithm hyper-parameters. However, searching all the hyper-parameters on the grid requires a considerable amount of time. In this paper, we proposed the improved GSO (IGSO) algorithm to improve the training speed and construct a better model (IGSO-RF) for oil palm detection. To speed up the search time, we used a long-distance step size for a rough search over a large range and used small-distance steps to further refine the grid near the optimal point. Additionally, based on error rate and information entropy of OOB data, we proposed the estimate function f_{OOB} to estimate the generalization error of the objective function, which can evaluate the strength of a decision tree and the correlations between the decision trees [51]. Suppose that $OOB_N(x)$ is the OOB data of RF classification model, N is the number of OOB data, n is the number of correctly classified data in OOB data, e_n and $H(n)$ are error

rate and information entropy of OOB data, respectively. Then, the estimate function f_{OOB} can be defined as in Equation (1):

$$f_{OOB} = H(n) - \log \frac{1 - e_n}{e_n} \quad (1)$$

where $e_n = \frac{N-n}{N}$, $H(n) = - \sum_{n \in N} P(n) \times \text{lb}P(n)$, $P(n) = \frac{n}{N}$.

The specific steps of IGSO algorithm are as follows:

1. The ranges of k and m , which represent the number of decision trees and the number of split features, respectively, are determined. Then, the step size is set, and a two-dimensional grid is established for the parameter search. The grid nodes are parameter pairs of k and m .
2. A RF decision tree is constructed for each set of hyper-parameters on the grid node, and estimate function f_{OOB} is utilized to estimate the classification error.
3. The parameters k and m with the minimum classification error are selected. If either the classification error or the step size meets the requirements, the optimal parameters and classification error are output; otherwise, the step size is reduced, the above steps are repeated, and the search continues.

Other classification models were tested to further justify the performance of IGSO-RF. We compared the performance of IGSO-RF, random forests (RF) [31], support vector machine (SVM) [26], classification and regression tree (CART) [27], naive Bayes (NB) [28], and minimum distance (MD) [52]. The comparison was done with four groups of feature combinations (Table 4), in which the kappa coefficient was evaluated for each model. For this analysis, we chose these classification models because these are implemented in GEE platform, and we used the default parameters set by GEE, thus the model comparison may serve for GEE users in future studies.

3.6. Validation

Each detection result with four groups was evaluated with the overall accuracy [53] and kappa coefficient [53]. Overall accuracy is the rate of correctly classified cells and the kappa coefficient is the most widely used measure for the performance of models generating presence–absence predictions [53]. The models were validated with fivefold cross-validation by using one folds (600 samples) for training and four fold (2400 samples) for validation. The cross-validation is widely adopted as the model selection criterion and validation. In K -fold cross-validation, a part of folds are used for model construction and the hand-out fold is allocated to model validation [54].

4. Results

To obtain the best classification results, we set up four groups of feature combinations in this paper, as shown in Tables 2 and 4. In these groups of feature combinations, the features of groups IV were the most relevant features derived from IGSO-RF model. These groups of feature combinations were input into the IGSO-RF classification model for oil palm detection. Combined with the 2400 validation samples, which included mature oil palm plantations (MOP), young oil palm plantations (YOP), bare land (BL), and other land uses that are not oil palm plantations (OLU), we calculated the confusion matrix for the classification results with different feature combinations and used the overall accuracy and kappa coefficient to compare and analyze the differences among the classification results, as depicted in Figure 3.

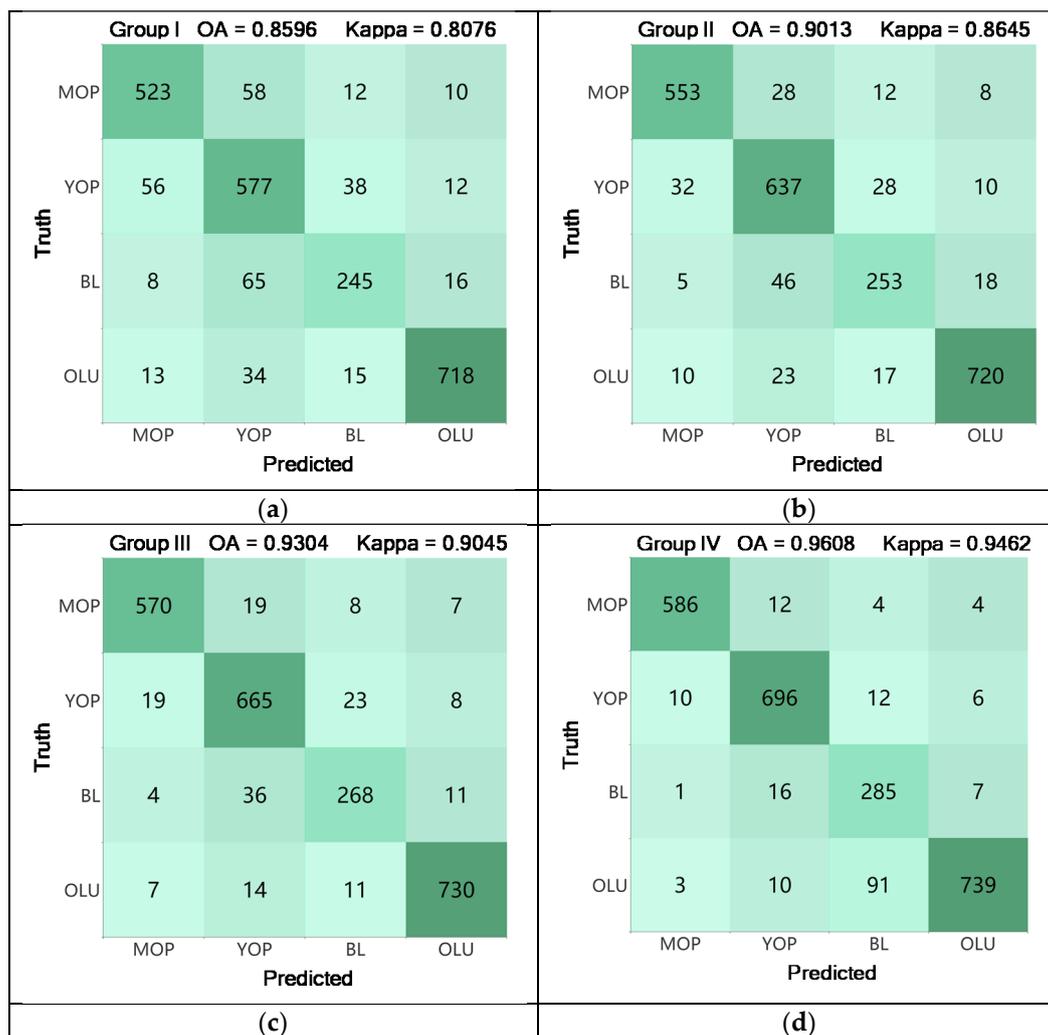


Figure 3. Confusion matrices of the models trained with random forest algorithm based on improved grid search optimization (IGSO-RF) and five land-cover classes, which included mature oil palm plantations (MOP), young oil palm plantations (YOP), bare land (BL), and other land uses that are not oil palm plantations (OLU). Each confusion matrix is the result of a different feature group: (a) Results of using feature group I; (b) Results of using feature group II; (c) Results of using feature group III; (d) Result of using group IV.

In Figure 3, the overall accuracy and kappa coefficient showed different growth trends among the four groups of feature combinations. For group I (Figure 3a), only the original spectral bands were selected as the features, the overall accuracy was 85.96%, and the kappa coefficient was 0.8076. On the basis of the spectral characteristics, by introducing the SAR backscatter value in group II (Figure 3b), the overall accuracy and kappa coefficient were improved to 90.13% and 0.8645, respectively, showing that the SAR backscatter value is helpful for oil palm detection and land classification. In group III (Figure 3c), after adding vegetation and texture features for oil palm detection, the overall accuracy and kappa coefficient further improved to 93.04% and 0.9045, respectively, reflecting the good detection performance of mature oil palm and young oil palm by the proposed algorithm. For group IV (Figure 3d), compared with those in group III, the classification accuracy was increased by 4.04%, and the kappa coefficient was increased by 4.61%. Hence, with the addition of feature variables, the classification accuracy was gradually improved. When the spectral bands, SAR backscatter values, vegetation indices and texture features were all integrated into the model, mature and young oil palm could be distinguished accurately, indicating that the synthesis of multisource features is conducive to distinguishing types of oil palm (mature and young). In addition, the classification accuracy has different

sensitivities to different feature types, and the optimal feature subset can improve the classification accuracy more effectively than the combination of features without a feature selection step; moreover, the best performance is achieved when distinguishing young oil palm plantations from bare ground.

Figure 4 shows the performance of the classification models with four groups of feature combinations. These models were used the default hyper-parameters set by GEE platform except our IGSO-RF model. For IGSO-RF model, the hyper-parameters k and m that derived by IGSO algorithm was set to 260 and 4, respectively. In Figure 4, the IGSO-RF model achieved the best accuracy among the four groups (kappa = 0.8076, 0.8645, 0.9045, and 0.9462 in groups I, II, III, and IV, respectively), followed by the SVM classifier in groups I and II (kappa = 0.7794 and 0.8471, respectively) and the RF algorithm in groups III and IV (kappa = 0.8741 and 0.9165, respectively). These results demonstrate that the performance of the RF algorithm is obviously better than that of the SVM and other classifiers with all feature combinations. In contrast, IGSO-RF uses the IGSO algorithm to optimize the RF hyper-parameters, which improves the classification accuracy and obtains better oil palm plantation detection results. Except for the NB and MD models in group I, the models display similar performance with a kappa coefficient above 75% in each of the four groups and exhibit a similar growth trend with the highest kappa in group IV.

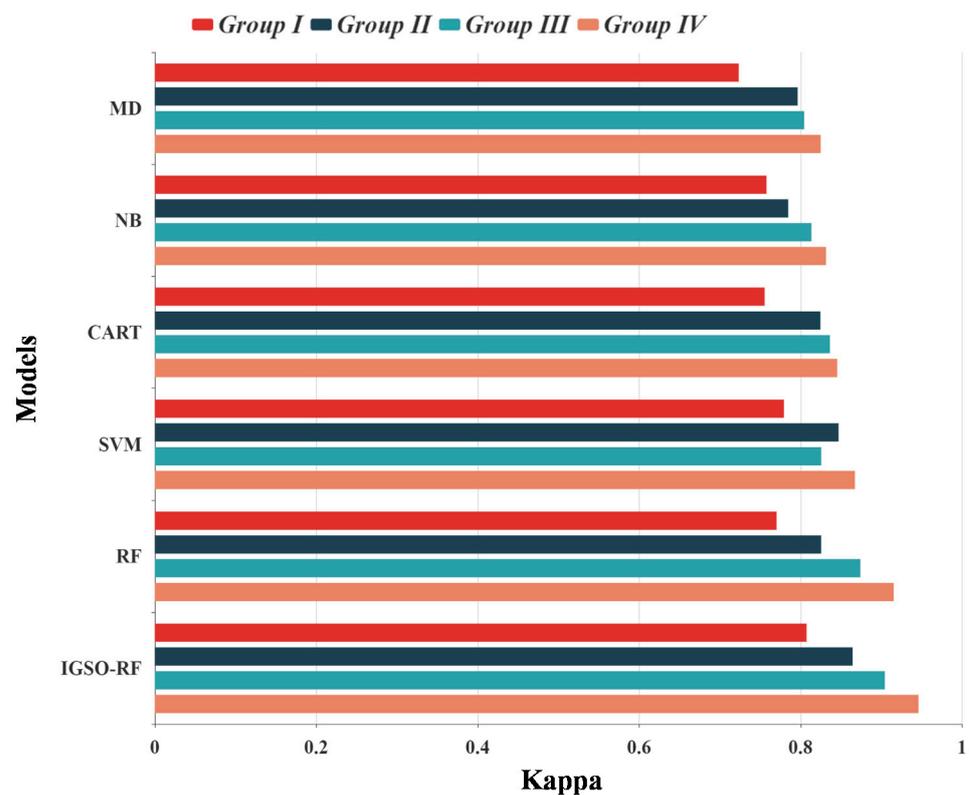


Figure 4. Results of six supervised classification models with different groups of selected features: random forest algorithm based on improved grid search optimization (IGSO-RF), Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART), Naive Bayes (NB), and Minimum Distance (MD).

The significant features selected by the Gini coefficient that driven from IGSO-RF are shown in Figure 5. The Gini coefficient method selected a total of 15 features derived from both Landsat-8 and Sentinel-1, with 9 features from Landsat-8 and 6 features from Sentinel-1 data. SAR data were sorted as the first feature, and the VH and VV bands and their difference showed high Gini coefficients in the classification model; hence, these features were important for the detection of oil palm plantations. Furthermore, texture information and vegetation indices were also useful for the classification of vegetation, and

texture features improved the oil palm detection accuracy more effectively than did the vegetation indices.

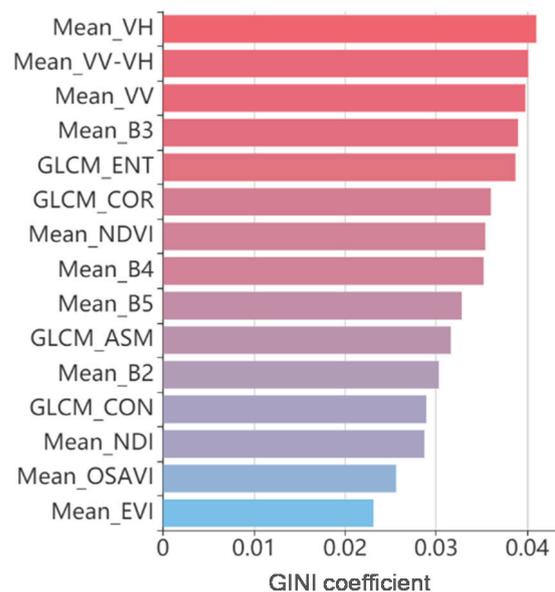


Figure 5. Ranking of the features according to the Gini coefficient method.

Figure 6 shows the oil palm plantation detection result using the RF algorithm in group III and the IGSO-RF model in group IV. For this analysis, we chose the area with high cloud covered to verify the possibility and ability of SAR images and models in oil palm plantations detection under the high cloudy areas. From the Landsat-8 composited images after cloud masked (Figure 6a), we can still see the high cloud occurrence in the study area, indicating the cloud-score algorithm of GEE platform still have limitations due to the frequent presence of clouds. Based on a visual comparison with the Landsat-8 composites (Figure 6a), Figure 6b shows the ability of SAR images in oil palm plantations detection under the cloud-covered areas, indicating that the fusion images of Landsat-8 and Sentinel-1 can avoid the impact of high cloud covered. Figure 6c exemplifies the improvement in the IGSO-RF model trained with the optimal features, indicating that our IGSO-RF classification model and optimal feature subset corrected the major issues associated with the detection of young oil palm. In addition, by using the IGSO algorithm to modify the RF hyper-parameters with a feature selection step, the overall detection accuracy of oil palm plantations improved remarkably by 8.2%.

Figure 7 shows the comparison results of RF algorithm using improved grid search optimization (IGSO) and traditional grid search optimization (GSO) [33]. For this analysis, firstly, we used all our sample points and divided into six groups, which means each group has 600 samples. In addition, four spectral bands (Table 3) were selected for features of these samples. Then, for the GSO algorithm, the search steps size of k and m will not be changed, which were set to 10 and 1, respectively. However, we used two different step sizes for parameters searching in IGSO algorithm, and the long-distance step size of k and m were set to 50 and 2, respectively, while the small-distance step size were set to the same value as the step sizes in GSO algorithm. Additionally, the range of parameters in two algorithm were set to the same value. Figure 7 shows the improvement in the IGSO algorithm, its average running time was only 964s while the average running time of GSO algorithm was 1750s, indicating that the IGSO algorithm can save a lot of time for the parameter optimization due to the long-distance step searching.

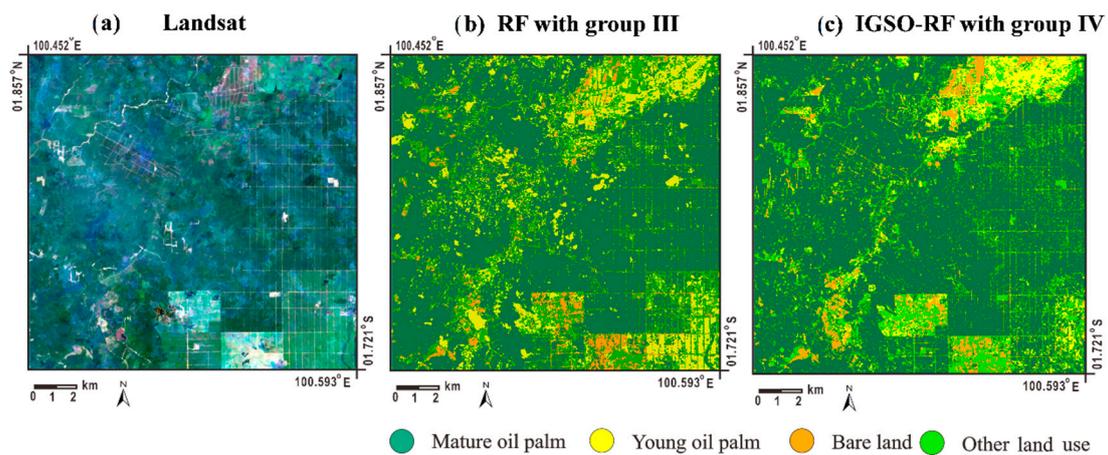


Figure 6. Examples of the improvements of using the IGSO-RF model and feature selection: (a) Landsat-8 images; (b) Results of using RF and multiple features; (c) Results of using the IGSO-RF model and feature selection.

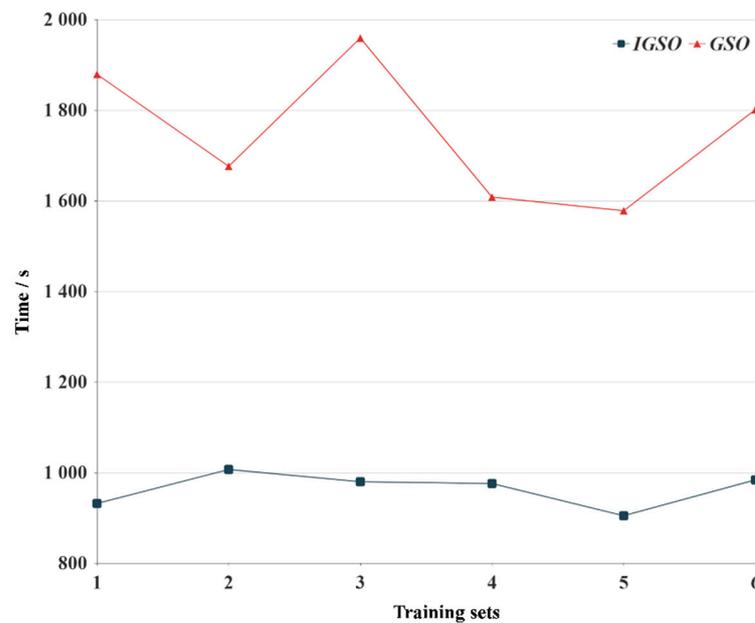


Figure 7. Comparison results of RF algorithm using improved grid search optimization (IGSO) or traditional grid search optimization (GSO).

Figure 8 shows the map of oil palm plantations in Riau Province generated from IGSO-RF model and feature combination group IV using the set of optimal features from Landsat-8 and Sentinel-1. The total area of oil palm plantations is 32,721 km², which represents 38.6% of the land surface of Riau on the Sumatran mainland. Of the total surface area of oil palm plantations, 70.8% is composed of mature oil palm plantations, and 29.2% comprises young oil palm plantations.

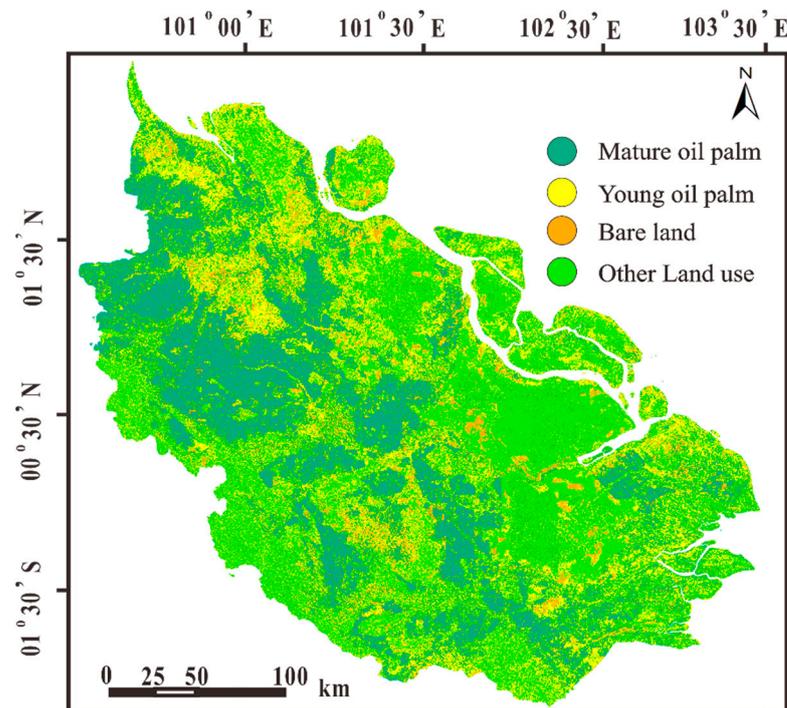


Figure 8. Result of oil palm plantation detection in Riau Province in 2019.

5. Discussion

This study showed the feasibility of detecting oil palm plantations by using Landsat-8 and Sentinel-1 datasets in the tropics. Moreover, this study developed an innovative method to improve the accuracy of detecting mature and young oil palm plantations by using the IGSO-RF classification model and optimal features with multisource remote sensing data. The bands of Landsat-8 and Sentinel-1 were selected, and further steps used in this study were unprecedented in the topic compared with previous studies that used Landsat-8 and Sentinel-1 images. Furthermore, the relative importance of optical and radar datasets and other features for the detection of oil palm plantations was analyzed, revealing that using the IGSO-RF model and optimal features improved the oil palm plantation detection accuracy. The improvement in the detection accuracy highlights the importance of data fusion and classifier optimization, which integrate multiple data features and optimize the parameters of classifiers to improve the detectability and accuracy of identified details. This is especially important in the tropics, which are characterized by the highest rate of oil palm plantation expansion and where optical-based detection is limited due to high extents of cloud cover.

The result of total area of oil palm plantation detection in Riau Province (32,721 km²), are comparable to the user's area of 31,020 km² obtained in a recent study [25] for oil palm plantation mapping. The accuracy obtained using the combination of Landsat-8 and Sentinel-1 (OA = 90.13% and kappa = 0.8645), for oil palm detection, confirmed the usefulness of SAR data for mature oil palm detection [55]. The results also show that feature extraction is necessary when detecting young oil palm trees. This is important for further studies that aim to detect precisely the type of oil palm plantations.

The characteristics of oil palm plantations are related to many different types of features. Without feature selection, the accuracy is increased by adding more types of feature variables involved in the classification to the feature combination. Combining all feature types, the overall accuracy of group III was improved to 93.04%, and the kappa coefficient was 0.9045, indicating that using multiple types of features can effectively improve the classification accuracy of mature and young oil palm plantations. In addition, after applying feature selection in group IV, using the optimal subset improved the detection accuracy more effectively than using a feature combination without feature selection

and exhibited the best performance for distinguishing young oil palm plantations from bare ground.

The model comparison emphasized the effectiveness of the IGSO-RF classifier for rapidly modeling and detecting oil palm plantations, even in cloudy areas such as Indonesia. Compared to other supervised classification models, the IGSO-RF classifier, which uses the IGSO algorithm to optimize the parameters selected for the traditional RF model, delivered higher accuracies. Although other classifiers such as SVM and neural network may perform well at recognizing individual oil palm trees, they require a large number of samples, and many parameters must be tuned in the training stage [26,29]. Instead, RF is a fast and easy-to-use classification model that requires less parameterization for training [31], and our IGSO-RF classifier optimized the parameter selection to improve the RF classification accuracy; thus, the proposed method can be used extensively for scientific vegetation detection.

The IGSO-RF model also shows the importance of parameter optimization. Since the (GSO) method can search the grid area of a variable to find the optimal grid point that satisfies the constraint function, it has been widely applied in many classification algorithm parameters, such as SVM [56] and RF [54]. In addition, our IGSO method can improve the classification performance of the random forest algorithm to a certain extent, and it also can save a lot of time compared with the traditional GSO algorithm. However, in rare cases, the continuing search near the best point generated by the long-distance steps searching may fall into the local optimum, which causes the optimal parameters cannot be found. This problem also occurs in many optimization algorithms such as artificial bee colony (ABC) optimization algorithm [57], and ant colony optimization algorithm [58]. Therefore, the step size of optimization algorithm needs to be set more reasonably to avoid this defect. This issue can be further studied in the future.

The Gini coefficient importance score was used to select the features of each feature combination, and the classification accuracy of each group was improved to different degrees. For the same combination of features, feature optimization can effectively improve the classification accuracy. The different feature types contributed to the detection of oil palm to different extents; SAR backscatter values contributed the most, followed by the spectral bands and texture features, while the vegetation indices contributed the least. The characteristic canopy of oil palm plantations might explain the high relevance of SAR backscatter values [25]. In particular, the shapes and structure of palm-like trees result in a characteristically high backscatter response in the VH dual band [38]. The high importance of the SAR backscatter value is further evidenced by the high relevance of the VH band and the VV-VH difference in the feature selection model, which is consistent with the results of a recent study on oil palm detection [38]. Despite the good results for Sentinel-1 in mature oil palm detection, the SAR backscatter value alone cannot distinguish young oil palm trees from other vegetation and requires additional features derived from Landsat-8 and Sentinel-1, which are effective at capturing the shape and density of other vegetation in oil palm plantations.

6. Conclusions

This study developed an innovative method to improve the detection accuracy of mature and young oil palm by using the IGSO-RF classification model and the selection of optimal features with the fusion of Landsat-8 and Sentinel-1 images. The proposed method performed better than other classifiers and improved the detection accuracy of oil palm plantations, thereby resolving the difficulties in distinguishing mature from young oil palm trees by using optical and SAR data; moreover, the proposed method optimized the input parameters in the classification model. This approach may be useful for our method to detect oil palm plantations in the humid tropics with insufficient oil palm samples and high-resolution images obscured by clouds. However, the feature combinations in this study need to be expanded. It is necessary to introduce other features, such as geometric and phenological features, and to thoroughly analyze the impacts of other feature types and

the number of features on the classification accuracy through feature selection. In addition, this study detected only the area of oil palm plantations; hence, it remains necessary to further expand the types of crops to be detected and to select research areas with more complex planting structures.

Author Contributions: Conceptualization, J.Q.; methodology, J.Q., K.X. and Z.D.; software, S.W. and J.S.; writing—original draft preparation, J.Q. and K.X.; writing—review and editing, Z.H. and J.L.; visualization, C.C. and X.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key R&D Program of China (No. 2017YFE0100700), Shenzhen International S&T Cooperation Project (GJHZ20190821155805960), the National Natural Science Foundation of China (Grant No. 41971386 and No.41801223), and the General Research Fund (HKBU 12301820).

Acknowledgments: We thank the academic editor and reviewers for their constructive comments which greatly helped us to improve the quality of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Rulli, M.C.; Casirati, S.; Dell'Angelo, J.; Davis, K.F.; Passera, C.; D'Odorico, P. Interdependencies and telecoupling of oil palm expansion at the expense of Indonesian rainforest. *Renew. Sustain. Energy Rev.* **2019**, *105*, 499–512. [[CrossRef](#)]
- Murphy, D.J. Oil palm: Future prospects for yield and quality improvements. *Lipid Technol.* **2009**, *21*, 257–260. [[CrossRef](#)]
- Fitzherbert, E.B.; Struebig, M.J.; Morel, A.; Danielsen, F.; Brühl, C.A.; Donald, P.F.; Phalan, B. How will oil palm expansion affect biodiversity? *Trends Ecol. Evol.* **2008**, *23*, 538–545. [[CrossRef](#)]
- Permpool, N.; Bonnet, S.; Gheewala, S.H. Greenhouse gas emissions from land use change due to oil palm expansion in Thailand for biodiesel production. *J. Clean. Prod.* **2016**, *134*, 532–538. [[CrossRef](#)]
- Carlson, K.M.; Curran, L.M.; Ratnasari, D.; Pittman, A.M.; Soares-Filho, B.S.; Asner, G.P.; Trigg, S.N.; Gaveau, D.A.; Lawrence, D.; Rodrigues, H.O. Committed carbon emissions, deforestation, and community land conversion from oil palm plantation expansion in West Kalimantan, Indonesia. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7559–7564. [[CrossRef](#)] [[PubMed](#)]
- Carlson, K.M.; Heilmayr, R.; Gibbs, H.K.; Noojipady, P.; Burns, D.N.; Morton, D.C.; Walker, N.F.; Paoli, G.D.; Kremen, C. Effect of oil palm sustainability certification on deforestation and fire in Indonesia. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 121–126. [[CrossRef](#)]
- Gutiérrez-Vélez, V.H.; DeFries, R.; Pinedo-Vásquez, M.; Uriarte, M.; Padoch, C.; Baethgen, W.; Fernandes, K.; Lim, Y. High-yield oil palm expansion spares land at the expense of forests in the Peruvian Amazon. *Environ. Res. Lett.* **2011**, *6*, 044029. [[CrossRef](#)]
- Srestasathien, P.; Rakwatin, P. Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sens.* **2014**, *6*, 9749–9774. [[CrossRef](#)]
- 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. [[CrossRef](#)]
- 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. [[CrossRef](#)]
- Morel, A.C.; Saatchi, S.S.; Malhi, Y.; Berry, N.J.; Banin, L.; Burslem, D.; Nilus, R.; Ong, R.C. Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data. *For. Ecol. Manag.* **2011**, *262*, 1786–1798. [[CrossRef](#)]
- Li, L.; Dong, J.; Njeudeng Tenku, S.; Xiao, X. Mapping oil palm plantations in Cameroon using PALSAR 50-m orthorectified mosaic images. *Remote Sens.* **2015**, *7*, 1206–1224. [[CrossRef](#)]
- Sum, A.F.W.; Shukor, S.A.A. Oil Palm Plantation Monitoring from Satellite Image. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *705*, 012043. [[CrossRef](#)]
- Oon, A.; Ngo, K.D.; Azhar, R.; Ashton-Butt, A.; Lechner, A.M.; Azhar, B. Assessment of ALOS-2 PALSAR-2L-band and Sentinel-1 C-band SAR backscatter for discriminating between large-scale oil palm plantations and smallholdings on tropical peatlands. *Remote Sens. Appl. Soc. Environ.* **2019**, *13*, 183–190. [[CrossRef](#)]
- Nomura, K.; Mitchard, E.T.; Patenaude, G. Oil palm concessions in southern Myanmar consist mostly of unconverted forest. *Sci. Rep.* **2019**, *9*, 1–9. [[CrossRef](#)]
- Lazecky, M.; Lhota, S.; Penaz, T.; Klushina, D. Application of Sentinel-1 satellite to identify oil palm plantations in Balikpapan Bay. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *169*, 012064. [[CrossRef](#)]
- Cheng, Y.; Yu, L.; Cracknell, A.P.; Gong, P. Oil palm mapping using Landsat and PALSAR: A case study in Malaysia. *Int. J. Remote Sens.* **2018**, *37*, 5431–5442. [[CrossRef](#)]

18. Poortinga, A.; Tenneson, K.; Shapiro, A.; Nquyen, Q.; San Aung, K.; Chishtie, F.; Saah, D. Mapping plantations in Myanmar by fusing landsat-8, sentinel-2 and sentinel-1 data along with systematic error quantification. *Remote Sens.* **2019**, *11*, 831. [[CrossRef](#)]
19. Oon, A.; Mohd Shafri, H.Z.; Lechner, A.M.; Azhar, B. Discriminating between large-scale oil palm plantations and smallholdings on tropical peatlands using vegetation indices and supervised classification of LANDSAT-8. *Int. J. Remote Sens.* **2019**, *40*, 7312–7328. [[CrossRef](#)]
20. Shaharum, N.S.N.; Shafri, H.Z.M.; Ghani, W.A.W.A.K.; Samsatli, S.; Prince, H.M.; Yusuf, B.; Hamud, A.M. Mapping the spatial distribution and changes of oil palm land cover using an open source cloud-based mapping platform. *Int. J. Remote Sens.* **2019**, *40*, 7459–7476. [[CrossRef](#)]
21. Daliman, S.; Rahman, S.A.; Bakar, S.A.; Busu, I. Segmentation of oil palm area based on GLCM-SVM and NDVI. In Proceedings of the 2014 IEEE Region 10 Symposium, Kuala Lumpur, Malaysia, 14–16 April 2014; pp. 645–650. [[CrossRef](#)]
22. Morel, A.C.; Fisher, J.B.; Malhi, Y. Evaluating the potential to monitor aboveground biomass in forest and oil palm in Sabah, Malaysia, for 2000–2008 with Landsat ETM+ and ALOS-PALSAR. *Int. J. Remote Sens.* **2012**, *33*, 3614–3639. [[CrossRef](#)]
23. Carolita, I.; Darmawan, S.; Permana, R.; Dirgahayu, D.; Wiratmoko, D.; Kartika, T.; Arifin, S. Comparison of Optic Landsat-8 and SAR Sentinel-1 in Oil Palm Monitoring, Case Study: Asahan, North Sumatera, Indonesia. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2019; Volume 280, p. 012015. [[CrossRef](#)]
24. Miettinen, J.; Gaveau, D.L.; Liew, S.C. Comparison of visual and automated oil palm mapping in Borneo. *Int. J. Remote Sens.* **2019**, *40*, 8174–8185. [[CrossRef](#)]
25. Descals, A.; Szantoi, Z.; Meijaard, E.; Sutikno, H.; Rindanata, G.; Wich, S. Oil Palm (*Elaeis guineensis*) Mapping with Details: Smallholder versus Industrial Plantations and their Extent in Riau, Sumatra. *Remote Sens.* **2019**, *11*, 2590. [[CrossRef](#)]
26. Nooni, I.K.; Duker, A.A.; Van Duren, I.; Addae-Wireko, L.; Osei Jnr, E.M. Support vector machine to map oil palm in a heterogeneous environment. *Int. J. Remote Sens.* **2014**, *35*, 4778–4794. [[CrossRef](#)]
27. Sitthi, A.; Nagai, M.; Dailey, M.; Ninsawat, S. Exploring land use and land cover of geotagged social-sensing images using naive bayes classifier. *Sustainability* **2016**, *8*, 921. [[CrossRef](#)]
28. Shaharum, N.S.N.; Shafri, H.Z.M.; Ghani, W.A.W.A.K.; Samsatli, S.; Al-Habshi, M.M.A.; Yusuf, B. Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sens. Appl. Soc. Environ.* **2020**, *17*, 100287. [[CrossRef](#)]
29. Mubin, N.A.; Nadarajoo, E.; Shafri, H.Z.M.; Hamedianfar, A. Young and mature oil palm tree detection and counting using convolutional neural network deep learning method. *Int. J. Remote Sens.* **2019**, *40*, 7500–7515. [[CrossRef](#)]
30. Liu, M.; Wang, M.; Wang, J.; Li, D. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sens. Actuator B-Chem.* **2013**, *177*, 970–980. [[CrossRef](#)]
31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
32. Tan, K.P.; Kanniah, K.D.; Cracknell, A.P. Use of UK-DMC 2 and ALOS PALSAR for studying the age of oil palm trees in southern peninsular Malaysia. *Int. J. Remote Sens.* **2013**, *34*, 7424–7446. [[CrossRef](#)]
33. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305. [[CrossRef](#)]
34. Chappelle, O.; Vapnik, V.; Bousquet, O.; Mukherjee, S. Choosing multiple parameters for support vector machines. *Mach. Learn.* **2002**, *46*, 131–159. [[CrossRef](#)]
35. Susanti, A. *Oil Palm Expansion in Indonesia: Serving People, Planet and Profit?* Eburon Academic Publishers: Utrecht, The Netherlands, 2016.
36. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [[CrossRef](#)]
37. Huang, H.; Chen, Y.; Clinton, N.; Wang, J.; Wang, X.; Liu, C.; Gong, P.; Yang, J.; Bai, Y.; Zheng, Y.; et al. Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. *Remote Sens. Environ.* **2017**, *202*, 166–176. [[CrossRef](#)]
38. Miettinen, J.; Liew, S.C.; Kwok, L.K. Usability of Sentinel-1 dual polarization C-band data for plantation detection in insular Southeast Asia. In Proceedings of the 36th Asian Conference Remote Sensing, Manila, Philippines, 19–23 October 2015; pp. 19–23.
39. Richardson, A.J.; Everitt, J.H. Using spectral vegetation indices to estimate rangeland productivity. *Geocarto Int.* **1992**, *7*, 63–69. [[CrossRef](#)]
40. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
41. Tucker, C.J.; Pinzon, J.E.; Brown, M.E.; Slayback, D.A.; Pak, E.W.; Mahoney, R.; Vermote, E.F.; El Saleous, N. An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *Int. J. Remote Sens.* **2005**, *26*, 4485–4498. [[CrossRef](#)]
42. De Petris, S.; Boccardo, P.; Borgogno-Mondino, E. Detection and characterization of oil palm plantations through MODIS EVI time series. *Int. J. Remote Sens.* **2019**, *40*, 7297–7311. [[CrossRef](#)]
43. Huete, A. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
44. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [[CrossRef](#)]
45. Kee, Y.W.; Shariff, A.R.M.; Sood, A.M.; Nordin, L. Application of SAR data for oil palm tree discrimination. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2018; Volume 169, p. 012065. [[CrossRef](#)]
46. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. *Remote Sens. Rev.* **1995**, *13*, 95–120. [[CrossRef](#)]

47. Mohanaiah, P.; Sathyanarayana, P.; GuruKumar, L. Image texture feature extraction using GLCM approach. *Int. J. Sci. Res. publications*. **2013**, *3*, 1.
48. Ulaby, F.T.; Kouyate, F.; Brisco, B.; Williams, T.L. Textural information in SAR images. *IEEE Trans. Geosci. Remote Sens.* **1986**, *2*, 235–245. [[CrossRef](#)]
49. Chandrashekar, G.; Sahin, F. A survey on feature selection methods. *Comput. Electr. Eng.* **2014**, *40*, 16–28. [[CrossRef](#)]
50. Breiman, L. Some properties of splitting criteria. *Mach. Learn.* **1996**, *24*, 41–47. [[CrossRef](#)]
51. Wang, X.; Gong, G.; Li, N.; Qiu, S. Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization. *Front. Hum. Neurosci.* **2019**, *13*, 52. [[CrossRef](#)]
52. Lee, J.S.H.; Wich, S.; Widayati, A.; Koh, L.P. Detecting industrial oil palm plantations on Landsat images with Google Earth Engine. *Remote Sens. Appl. Soc. Environ.* **2016**, *4*, 219–224. [[CrossRef](#)]
53. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232. [[CrossRef](#)]
54. Jung, Y. Multiple predicting K-fold cross-validation for model selection. *J. Nonparametr. Stat.* **2018**, *30*, 197–215. [[CrossRef](#)]
55. Miettinen, J.; Liew, S.C. Separability of insular Southeast Asian woody plantation species in the 50 m resolution ALOS PALSAR mosaic product. *Remote Sens. Lett.* **2011**, *2*, 299–307. [[CrossRef](#)]
56. Wenwen, L.; Xiaoxue, X.; Fu, L.; Yu, Z. Application of improved grid search algorithm on SVM for classification of tumor gene. *Int. J. Multimed. Ubiquitous Eng.* **2014**, *9*, 181–188. [[CrossRef](#)]
57. Karaboga, D.; Basturk, B. Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems. In *International Fuzzy Systems Association World Congress*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 789–798.
58. Dorigo, M.; Birattari, M.; Stutzle, T. Ant colony optimization. *IEEE Comput. Intell. Mag.* **2006**, *1*, 28–39. [[CrossRef](#)]