



## Article

# Offshore Hydrocarbon Exploitation Target Extraction Based on Time-Series Night Light Remote Sensing Images and Machine Learning Models: A Comparison of Six Machine Learning Algorithms and Their Multi-Feature Importance

Rui Ma <sup>1,†</sup>, Wenzhou Wu <sup>2,†</sup>, Qi Wang <sup>1,2,\*</sup>, Na Liu <sup>1</sup> and Yutong Chang <sup>1</sup><sup>1</sup> School of Surveying and Geo-Informatics, Shandong Jianzhu University, Jinan 250101, China<sup>2</sup> State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

\* Correspondence: wangqi19@sdjzu.edu.cn

† These authors contributed equally to this work.

**Abstract:** The continuous acquisition of spatial distribution information for offshore hydrocarbon exploitation (OHE) targets is crucial for the research of marine carbon emission activities. The methodological framework based on time-series night light remote sensing images with a feature increment strategy coupled with machine learning models has become one of the most novel techniques for OHE target extraction in recent years. Its performance is mainly influenced by machine learning models, target features, and regional differences. However, there is still a lack of internal comparative studies on the different influencing factors in this framework. Therefore, based on this framework, we selected four different typical experimental regions within the hydrocarbon basins in the South China Sea to validate the extraction performance of six machine learning models (the classification and regression tree (CART), random forest (RF), artificial neural networks (ANN), support vector machine (SVM), Mahalanobis distance (MaD), and maximum likelihood classification (MLC)) using time-series VIIRS night light remote sensing images. On this basis, the influence of the regional differences and the importance of the multi-features were evaluated and analyzed. The results showed that (1) the RF model performed the best, with an average accuracy of 90.74%, which was much higher than the ANN, CART, SVM, MLC, and MaD. (2) The OHE targets with a lower light radiant intensity as well as a closer spatial location were the main subjects of the omission extraction, while the incorrect extractions were mostly caused by the intensive ship activities. (3) The coefficient of variation was the most important feature that affected the accuracy of the OHE target extraction, with a contribution rate of 26%. This was different from the commonly believed frequency feature in the existing research. In the context of global warming, this study can provide a valuable information reference for studies on OHE target extraction, carbon emission activity monitoring, and carbon emission dynamic assessment.

**Keywords:** offshore hydrocarbon exploitation targets; night light remote sensing images; feature increment strategy; machine learning model; feature evaluation



**Citation:** Ma, R.; Wu, W.; Wang, Q.; Liu, N.; Chang, Y. Offshore Hydrocarbon Exploitation Target Extraction Based on Time-Series Night Light Remote Sensing Images and Machine Learning Models: A Comparison of Six Machine Learning Algorithms and Their Multi-Feature Importance. *Remote Sens.* **2023**, *15*, 1843. <https://doi.org/10.3390/rs15071843>

Academic Editor: Benoit Vozel

Received: 27 December 2022

Revised: 27 March 2023

Accepted: 28 March 2023

Published: 30 March 2023



**Copyright:** © 2023 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

Massive carbon emissions contribute to global warming and pose a great threat to the world's ecosystems and sustainable development [1]. Countries around the world are reducing carbon emissions in a global compact to improve their ability to combat climate change [2]. Most of the anthropogenic greenhouse gas emissions come from the combustion of fossil fuels [3–5]. According to the International Energy Agency (IEA), the total greenhouse gas emissions from energy reached a historic high in 2021 [6]. Of these, offshore hydrocarbon exploitation (OHE) provided nearly 30% of the global hydrocarbon

growth [7,8] and was an important source of marine carbon emissions since its exploitation process is accompanied by fossil fuel combustion, flare burning, venting and escape, and electricity consumption, all of which emit a large amount of methane, CO<sub>2</sub>, etc. [9,10]. Therefore, the continuous and dynamic acquisition of spatial distribution information for OHE targets (exploitation platforms, platform groups, FPSOs, etc.) provides an important basic database for monitoring and analyzing offshore carbon emission activities. It is also of great significance for carbon emission policy formulation and carbon trading.

With wide coverage and multiple sources of information, remote sensing technology can overcome the limitation of the difficulty of close proximity observation, thus becoming the most effective method for extracting OHE targets. As of now, many scholars have completed considerable work and gained outstanding headway in this research field.

Since OHE targets and ships have similar scattering properties in synthetic aperture radar (SAR) images, the offshore ship detection algorithm is also applied to the detection of OHE targets. For example, Chen et al. [11] implemented an extraction of OHE platforms in the western part of the South China Sea based on ENVISAT ASAR images using the two-parameter constant false alarm rate (CFAR) detection algorithm. Cheng et al. [12] achieved an automatic matching of OHE targets using the CFAR detection algorithm combined with a point clustering matching model based on invariant triangle rules. However, the CFAR algorithms are mostly pixel-based, which are highly susceptible to interference from sea clutter and the weak scattering of some HE targets. Since deep learning algorithms are very capable of feature learning, they provide new ideas for the recognition of OHE targets based on SAR images. Falqueto et al. [13] implemented the recognition of OHE platforms using the visual geometry group 16 (VGG-16) and visual geometry group 19 (VGG-19) convolutional neural network models based on Sentinel-1 SAR images. Liu et al. [14] proposed a target recognition algorithm coupling level set segmentation of the limited initial region with a convolutional neural network (CNN) and achieved an effective differentiation between the OHE platforms and the ships based on polarized synthetic aperture radar (PolSAR) images. Although SAR is capable of all-weather observation [15], its high data cost, low observation frequency, and low sea area coverage make it difficult to achieve large-scale and long time-series OHE target extraction [16].

Different from SAR images, optical images have a longer service life, wider coverage, and larger data inventory, which effectively compensate for the time and space span deficiency of SAR images [17–19]. Therefore, many studies on the extraction of OHE targets based on optical images have been conducted. For example, Xing et al. [18] proposed an iterative threshold segmentation algorithm to detect the distribution of ships and OHE platforms in the Bohai Sea based on Landsat images. Liu et al. [20] effectively extracted OHE platforms in the Thailand Gulf, the Persian Gulf, and the northern Mexico Gulf based on Landsat-8 OLI images using both time-series and multi-refinement strategies. In addition, this automated method was extended to different types of satellite images [17]. Zhu et al. [21] proposed a multi-temporal normalized difference water index (NDWI)-based OHE platform detection method and extracted OHE platforms in the Caspian Sea using multi-temporal Landsat-7 ETM+ images. However, since the thresholds used in the above methods are typically defined based on the actual situation of the study area and are largely chosen empirically, it is difficult to determine the optimal threshold. Furthermore, optical images are difficult to use to obtain high-quality long-term images of the same area in the sea due to the impact of cloudy and rainy weather, which leads to deficiencies in the continuous and dynamic acquisition of OHE targets at sea [22,23].

In comparison to the first two data sources, night light remote sensing image data represented by NPP/VIIRS images can cover a vast region and be dynamically updated in a long time series, which was used in the study of OHE gas flaring monitoring [24–27]. Although the spatial resolution of nocturnal light data is relatively low [20], its high temporal resolution and ability to cover long time periods and large areas not only provide a free and abundant source of data for extracting OHE targets, but also demonstrate great potential in meeting the more refined time scales required for monitoring carbon emission

activities. Currently, related products and datasets are published in this field, for example, the global gas flaring distribution data obtained by Elvidge et al. [25] based on the VIIRS Nightfire (VNF) algorithm [28], and the global industrial heat source products extracted from long time-series VIIRS images by Liu et al. [29]. It should be noted that this product does not include a non-gas central processing platform, and there may be omissions in the detection of OHE targets with a low radiation intensity. However, OHE targets without gas emissions or with a weak radiation intensity are also important sources of carbon emissions. To address this problem, Wang et al. [23] proposed a framework of a coupled feature increment strategy and machine learning model based on time-series VIIRS images. This framework successfully extracted the spatial distribution of OHE targets in the South China Sea from 2012 to 2019, including “weak targets” with a low light intensity and low occurrence frequency. This method avoids the difficulties of setting thresholds, such as frequency parameters, in the existing studies. However, the performance of this framework is affected by the type of machine learning model, target features, and regional differences. Currently there is insufficient understanding of which type of machine learning model performs best, which type of feature has the greatest impact, and what factors contribute to errors in this framework.

To address the aforementioned problems, this study intends to extract OHE targets based on time-series VIIRS night light remote sensing images combined with a machine learning model. The main study objectives are (1) to compare and analyze the performance of the different machine learning algorithms in the OHE targets extraction, (2) to evaluate the degree of importance for the multi-features, and (3) to analyze the error factors in the different regions. The research results can provide technical support for the extraction of offshore carbon emission targets and data support for the subsequent analysis of the spatial and temporal variations of offshore carbon emission activities.

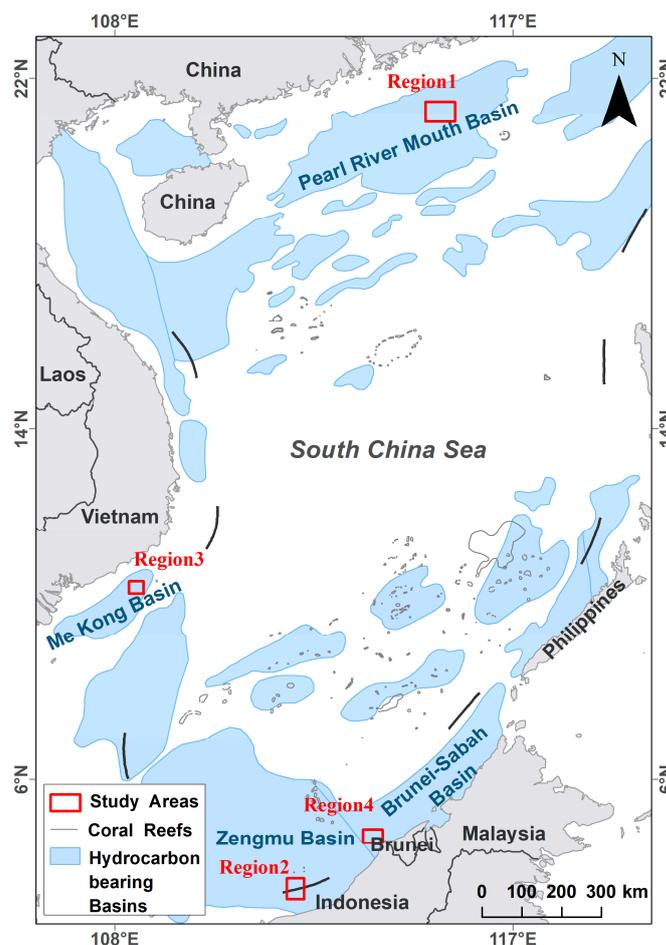
## 2. Study Area and Datasets

### 2.1. Study Area

In order to explore the impact of the regional differences on the extraction of the OHE targets, we selected four different typical regions in the South China Sea for the experiments. These four experimental areas were characterized by strong OHE activities, all of which were carbon emission intensive areas. Meanwhile, they displayed obvious regional differences.

The Pearl River Mouth Basin is located between 18°30′N–23°30′N, 113°10′E–118°00′E, bordering the Qiongdongnan Basin in the west and the Taixinan Basin in the east. By 2021, 48 hydrocarbon fields and 36 hydrocarbon-bearing formations have been discovered in the Pearl River Mouth Basin, with the cumulative proven oil reserves exceeding  $10 \times 10^8$  t and the natural gas reserves reaching  $1800 \times 10^8$  m<sup>3</sup> [30]. Therefore, it has turned into the most intensively developed hydrocarbon basin of the South China Sea at present. We selected the first experimental area in the northeastern part of the Pearl River Mouth Basin (Figure 1, Region 1), which mainly covers the Huizhou 26-4, Huizhou 32-3, Xijiang 24-3, and Xijiang 30-2 hydrocarbon fields, and the distribution among OHE targets is relatively independent. In addition, the region is close to coastal cities such as Guangzhou and Hong Kong in southern China, and is busy with the ship traffic.

The Zengmu Basin is located between 2°33′N–7°08′N, 108°30′E–114°15′E. It borders the WanAn Basin to the northwest and is separated from the Brunei-Sabah Basin by the Tinja fault [31] to the east. It is bordered by Indonesia to the west and Malaysia to the east. Currently, the proven geological reserves of natural gas are about  $4.7 \times 10^{12}$  m<sup>3</sup>, the oil geological reserves are about  $8 \times 10^8$  t, and more than 500 wellheads have been drilled. We selected the second experimental area within the SK305 and SK309 hydrocarbon blocks in the southern part of the Zengmu Basin (Figure 1, Region 2), which mainly contains the gas fields E6, F23, F6, E11, and F13. The OHE targets in this experimental area are generally scattered, but some of them are distributed in groups, and the intensity of the light radiation varies greatly among the different targets in the VIIRS images.



**Figure 1.** Study area. The red rectangle indicates the selected experimental area.

The Mekong Basin is located on the continental shelf of Vietnam, bordered by the Vietnamese mainland to the north and the Wan An Basin to the south, covering an area of about 41,000 km<sup>2</sup> with a sediment thickness of more than 7 km. As of 2012, 46 hydrocarbon fields have been discovered in the Mekong Basin, with proven recoverable oil reserves of over 600 million m<sup>3</sup> and natural gas reserves of over 170 billion m<sup>3</sup>. To date, 232 wells have been drilled in the Mekong Basin by the Vietnamese and other foreign oil companies [32]. The number of OHE targets within blocks 15-1 and 01 in the northeastern Mekong Basin is large and very densely distributed, so we selected the third experimental area from within these two blocks (Figure 1, Region 3). It should be noted that this area is close to the Ho Chi Minh Port in Vietnam, and there is a relatively high volume of traffic of ships coming and going.

The Brunei-Sabah Basin is located offshore in northeastern Kalimantan, with an area of about 94,000 km<sup>2</sup> and Cenozoic deposits up to 12.5 km thick. Among them, the deep water area has not only become the main area of hydrocarbon discovery in the basin in recent years but also one of the important pillars of the future Malaysian oil industry [33]. The experimental area in this basin is located within the SK307 hydrocarbon block, which is adjacent to the coastline of Malaysia and Brunei. Although there are many OHE targets in this area, the radiation intensity of these targets is relatively weak, and there is a significant difference in the light radiation intensity of each target.

## 2.2. Datasets

The datasets used in this study include NPP, VIIRS, and NTL images, Sentinel-2 images, and offshore platform records.

- (1) VIIRS Day/Night Band Nighttime Lights Monthly Composite Images

The VIIRS sensor radiation detection range is wide, covering a total of 22 spectral bands. Among them, the day/night (DNB) band can detect both electric lighting sources and combustion sources [28], thus both the OHE targets with and without gas flaring can be detected using VIIRS images. In addition, the high temporal resolution of VIIRS night light images can meet our needs for the OHE targets extraction and carbon emission monitoring under a long time series. Therefore, we used VIIRS night light images for the extraction of the OHE targets. We downloaded 24 images in 2016 and 2017 from the Earth Observation Group, Payne Institute for Public Policy, Colorado School of Mines ([https://eogdata.mines.edu/nighttime\\_light/monthly/](https://eogdata.mines.edu/nighttime_light/monthly/) (accessed on 1 September 2022)). The spatial resolution of the data was approx. 500 m, the range was Tie 3\_75N060E, and the data format was VCMCFG.

Since this range was far beyond the study area, we clipped the monthly VIIRS images for 2016 and 2017 with the boundaries of the Pearl River Mouth Basin, the Zengmu Basin, the Mekong Basin, and the Brunei-Sabah Basin.

## (2) Offshore Platform Records and Sentinel-2 images

To evaluate the accuracy of the OHE targets extraction results based on the long time-series VIIRS images, we collected the offshore platform records and Sentinel-2 images in 2016 and 2017 and used them as the important evaluation data to verify the accuracy of the OHE targets extraction and analyze the influencing factors.

The data of the offshore platform records were obtained from thematic product and the field survey data provided by the State Key Laboratory of Resources and Environment Information System, Institute of Geographical Sciences and Natural Resources, Chinese Academy of Sciences (<http://www.lreis.ac.cn/> (accessed on 11 March 2023)). These data recorded the location and photos of offshore platforms in the South China Sea from 2014 to 2017. The Sentinel-2 images of the study area were acquired through the Google Earth Engine with a maximum spatial resolution of 10 m. We selected the images in 2017 using the time filter function on the Google Earth Engine platform and then used the quality control band (QA) to de-cloud, taking the image set as the validation data for the target extraction.

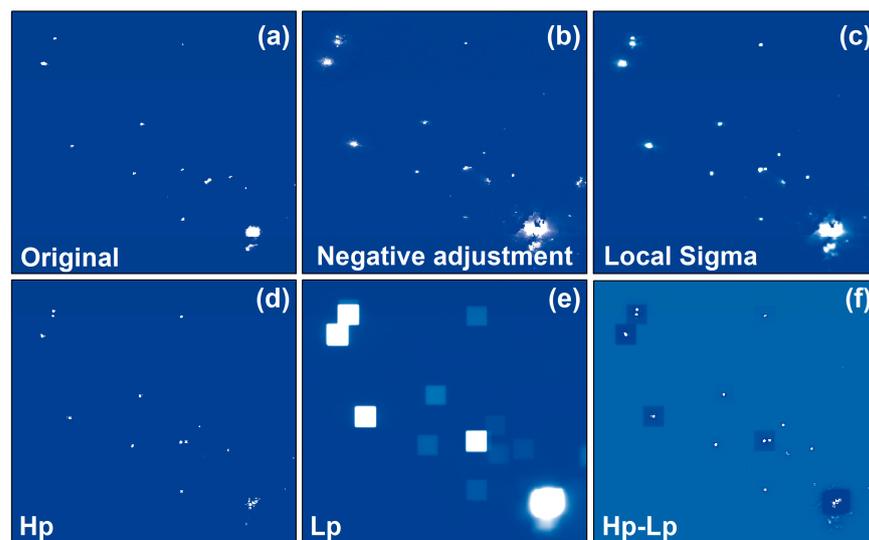
## 3. Methods

### 3.1. OHE Target Extraction Framework of the Coupling Feature Increment Strategy and Machine Learning Model

The framework mainly consisted of three parts: the data preprocessing, multi-features construction, and target extraction. The specific processes are as follows.

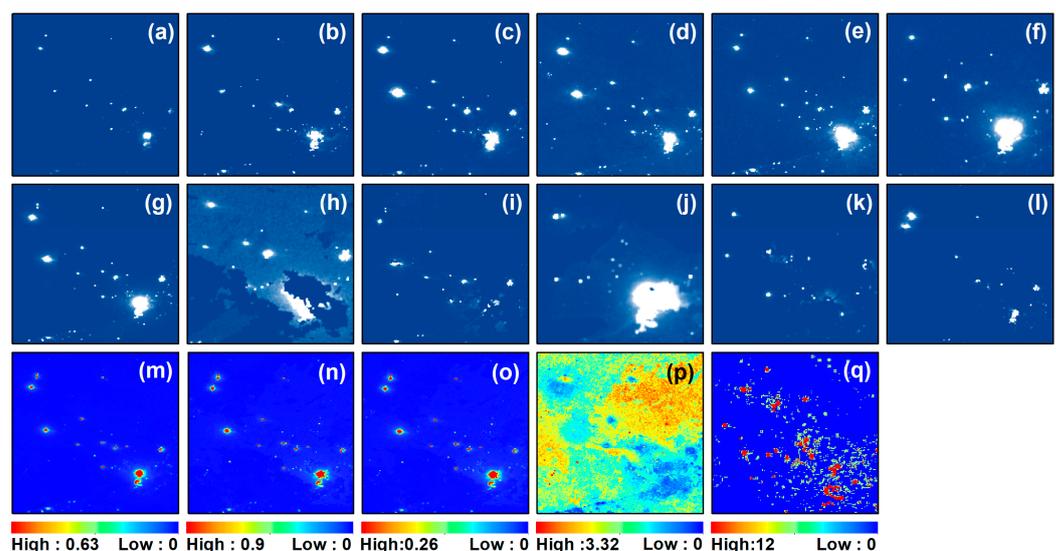
(1) First, the pixels with values less than 0 in the VIIRS images were adjusted. This was because the light radiation intensity values less than 0 were often not practically meaningful and tended to interfere with the subsequent calculations [23,34]. Then, the local sigma filtering [35] was used to further remove the large amount of noise contained in the monthly VIIRS images. This reduced the possibility of an incorrect extraction in the subsequent target extraction [23]. In this case, the window size of the filtering was set to  $3 \times 3$ , and the Sigma factor was set to four (Figure 2c).

(2) Convolution operations were used to enhance the contrast between the target and the background. First, a high-pass filter (Hp) was used to improve the pixel values of some targets with a low light intensity. However, the low resolution of the VIIRS images may have caused some targets with close positions to be blurred in the convolution operation. Therefore, the convolution kernel size of the high-pass filter was set to  $3 \times 3$  (Figure 2d). Then, a low-pass filter (Lp) was used to retain the low-frequency background information in the image, such as the halo generated by the exhaust gas combustion, seawater, etc. According to the halo diffusion range of the OHE targets, the low-pass filter convolution kernel size was set to  $27 \times 27$  (Figure 2e). Finally, the separation of the potential targets and the seawater background were enhanced by calculating the difference between the high-pass filtering and the low-pass filtering (Hp-Lp) (Figure 2f).



**Figure 2.** Data preprocessing: (a) is the original image; (b) is the negative adjustment image; (c) is the local sigma filtered image, the window size is  $3 \times 3$ , sigma factor is 4; (d) is the Hp image, convolution kernel is  $3 \times 3$ ; (e) is the Lp image, convolution kernel is  $27 \times 27$ ; and (f) is the difference between the Hp image and the Lp image.

(3) To avoid the difficulty of setting the optimal threshold in the existing threshold segmentation methods, this framework adopted a feature increment strategy, which added the brightness feature of the OHE targets on the basis of the occurrence frequency [23]. Firstly, the preprocessed monthly images were binarized using the threshold segmentation ( $t = 0.5$ ), and the binarization results of the twelve months were algebraically superimposed to obtain the occurrence frequency statistics of each target in a year. Subsequently, the negative-adjusted images were subjected to the max–min normalization on the basis of the local sigma filtering (Figure 3a–l). Meanwhile, four brightness indicators of the mean, standard deviation, maximum, and coefficient of variation were constructed (Figure 3m–p). Finally, the five constructed features were combined as a new data source for the subsequent target extraction based on the machine learning model.



**Figure 3.** Multi-features construction: (a–l) are the normalization results of the time-series images; (m) is the image of the mean statistics; (n) is the image of the maximum statistics; (o) is the image of the standard deviation statistics; (p) is the image of the coefficient of variation statistics; and (q) is the image of the occurrence frequency statistics.

(4) Following the construction of the multi-features of the OHE targets, we performed sample tagging using the Sentinel-2 images and offshore platform records. For example, in 2016, we labeled 591 training samples, which included 65 target samples and 526 seawater samples. This was due to the fact that the OHE target samples were smaller compared to the seawater samples. Secondly, seawater as a background has a regional variability due to ship activities and other reasons. Thus, more seawater samples needed to be labeled to improve the model training. Subsequently, six machine learning algorithms (CART, RF, ANN, SVM, MaD, and MLC) were utilized to extract the OHE targets. We describe the parameters of each algorithm in Section 3.2.

(5) Based on the target extraction, the framework used morphological filtering techniques to eliminate the isolated noises at the target edges or away from the target. The filter kernel size was set to  $3 \times 3$  [23]. Finally, the method of building 10 km buffer zones seaward along the coastline of each country [20] was adopted to eliminate the interference from a large number of fixed lights, such as port facilities and docked ships along the coastline [36]. At the same time, masks were developed for the countries along the coastline to ensure that the information in the non-participating processing area was not affected while eliminating the false alarm targets in the mask-covered area [23].

### 3.2. Machine Learning Algorithms

Machine learning models are widely used in remote sensing image classification, and their feature learning ability can avoid the problem of difficult threshold optimization. Therefore, this article selected six machine learning models for comparison in order to obtain the optimal model for extracting the OHE targets. A brief description of each model and its training parameters are as follows.

#### (1) Classification and Regression Tree (CART)

The CART algorithm is an algorithm that selects the optimal classification feature and segmentation threshold by comparing the Gini coefficients of each feature and the recursively forms a binary tree structure [37]. In the candidate attribute set, the attribute with the smallest partition Gini coefficient is the final classification attribute. In the CART model, we set the minimum node size to five, the number of folds to 10, and the variable selection method to an unbiased estimation.

#### (2) Random Forest (RF)

The RF algorithm, combining the bootstrap resampling technique with the node random splitting technique, improves the classification performance by constructing a forest of multiple decision trees [38,39]. It firstly generates a training sample set by repeatedly drawing samples at random from the original dataset with put-backs. The remaining samples are collectively referred to as the out-of-bag (OOB) data and are used to estimate the classification error. Then, a certain number of features are randomly selected from the attribute set, and the division feature with the smallest Gini coefficient is used as the optimal feature for the node splitting. Finally, the constructed multiple decision trees are formed into a forest. The prediction results of all the decision trees are voted for, and the one with the most votes is the final winner. The parameters used in the RF modeling process and their values were as follows: the number of trees in the forest was 100, the minimum sample node was 1, and the number of features was selected using the square root method.

#### (3) Support Vector Machine (SVM)

The core of the support vector machine (SVM) is to find an optimal hyperplane to separate a linearly separable sample set based on the principle of maximizing the margin while minimizing the empirical error [40,41]. If the sample set is non-linearly separable, an appropriate non-linear algorithm is used to map the inseparable sample data to a high-dimensional space, and the sample is linearly separable in the high-dimensional feature space through inner product operations [42]. According to the principle that the

kernel function that satisfies the Mercer condition corresponds to the inner product in some transformation space [43], the SVM model selects an appropriate kernel function to achieve a linear classification after the non-linear operations, thereby simplifying the classification problem. Compared to the linear kernel functions, polynomial kernel functions, and sigmoid kernel functions, the radial basis kernel function had a good performance in both the large and small samples and had fewer parameters, which made it less prone to overfitting. Therefore, we chose the radial basis kernel function for the classification research.

#### (4) Artificial Neural Networks (ANN)

An ANN is a dynamic system built artificially with a directed graph topology, which process information by responding to continuous or intermittent input states [44]. First, the input layer is responsible for receiving and learning the features of the observed aspect of each part, and then processing the acquired information in the hidden layer. However, since the overall result is unknown, it is necessary to achieve the best fitting result by continuously training and adjusting the weights and feedback of the hidden layer. Finally, the computational results are passed from the output layer [45]. In the ANN algorithm, we set the activation function to the logarithm, the training contribution threshold to 0.9, the weight adjustment speed to 0.2, the number of hidden layers to 1, and the number of training iterations to 1000.

#### (5) Mahalanobis Distance (MaD)

The MaD is a classification model that measures the similarity of two sample sets using the covariance distance of the data as an indicator. Compared to the Euclidean distance, the Mahalanobis distance can better represent the overall characteristics of the sample set. It classifies remote sensing images by calculating the Mahalanobis distance of vector  $x$  to each class mean vector and assigning  $x$  to the class where the nearest mean belongs [46]. In the MaD algorithm, we set the maximum distance error to a single value.

#### (6) Maximum Likelihood Classification (MLC)

The MLC is a supervised classification. It assumes that each pixel to be classified is normally distributed in each category, and builds the probability distribution function for each class based on the Bayesian discrimination criteria with prior knowledge [47]. A sample  $x$  to be classified belongs to category  $i$  when the value of the discriminant function in category  $i$  is bigger than the value of the discriminant function in any of the remaining categories. In the MLC algorithm, we set the likelihood threshold to zero and the proportion coefficient to one by default.

### 3.3. Accuracy Evaluation Method

Limited by the low spatial resolution of the VIIRS images, the groups of the OHE platforms in close proximity are often difficult to distinguish as individual OHE targets. To address this problem, the framework performed a neighborhood analysis based on the existing knowledge and existing studies within a radius of one pixel [23]. The OHE platforms and facilities within a neighborhood are considered as one large OHE target.

Then, we used the F1-measure index to comprehensively evaluate the extraction accuracy of each model. The F1-measure index is widely used in natural language processing, machine learning, and other fields [48,49], and can effectively evaluate the accuracy of the research methods. A higher F1 value indicates a higher accuracy for the OHE target extraction. The F1-measure calculation formula is shown below.

$$\begin{cases} P = TP / (TP + FP) \\ R = TP / (TP + FN) \\ F1 = 2 * PR / (P + R) \end{cases} \quad (1)$$

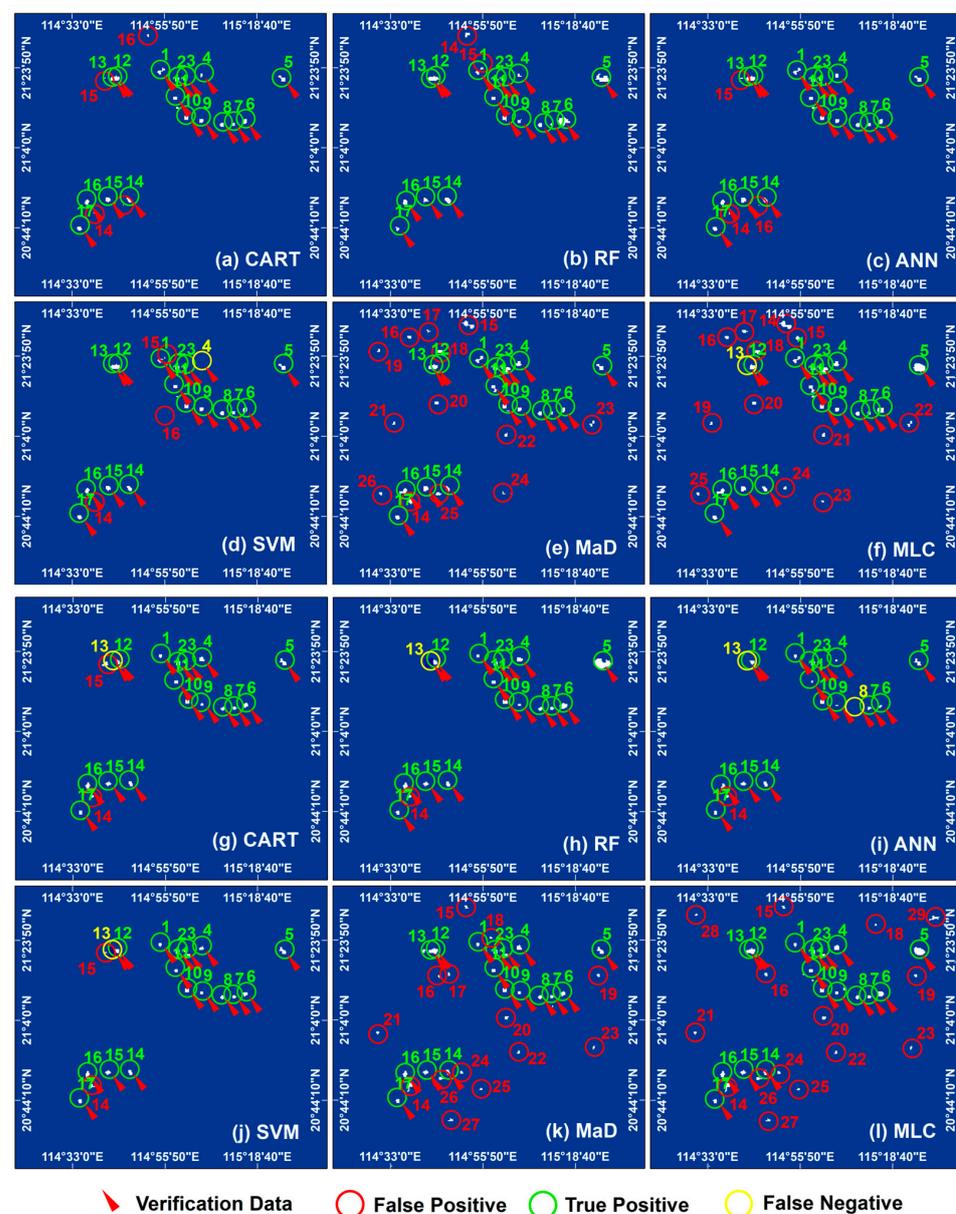
In Equation (1),  $TP$  (true positives) indicates the number of correct extractions;  $FP$  (false positives) indicates the number of incorrect extractions, i.e., the number of false alarms;  $FN$  (false negatives) indicates the number of omitted extractions;  $P$  (precision) indicates

the proportion of correct extractions to the actual number; and R (recall) indicates the proportion of correct extractions to the number of correct extractions that should be made.

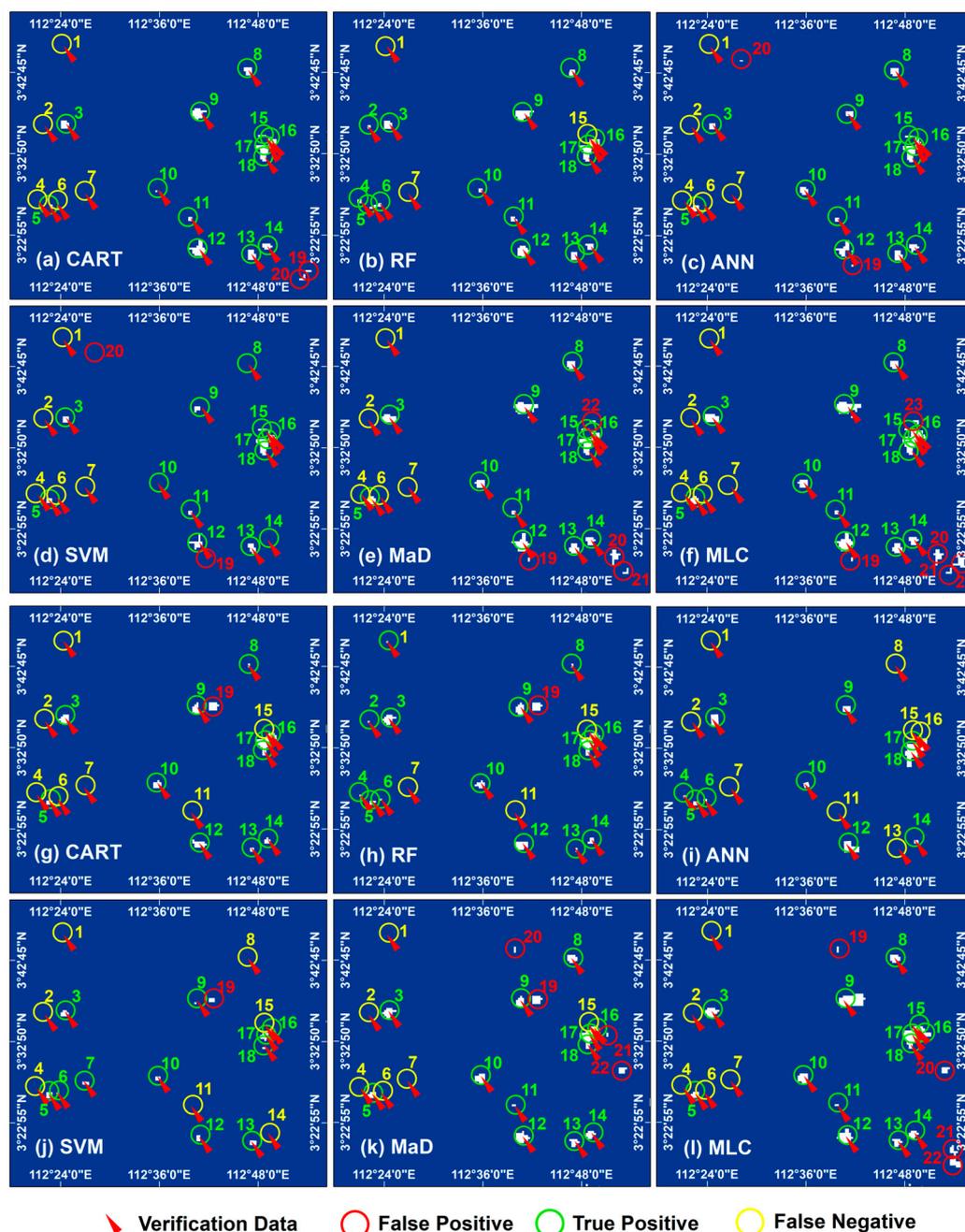
## 4. Results

### 4.1. OHE Target Extraction Results

We acquired the extraction results for each machine learning algorithm in the study areas based on the monthly VIIRS composite images from 2016 and 2017. Based on the offshore platform records and Sentinel-2 images, we classified the extraction results of the OHE targets into three categories (correct extraction, omission extraction, and incorrect extraction) and analyzed them. Figures 4–7 show the extraction results of the different machine learning algorithms in the four experimental areas. Since the numbers of OHE targets in 2016 and 2017 were similar, we chose the extraction results of 2017 for validation, as shown in Figure 8.

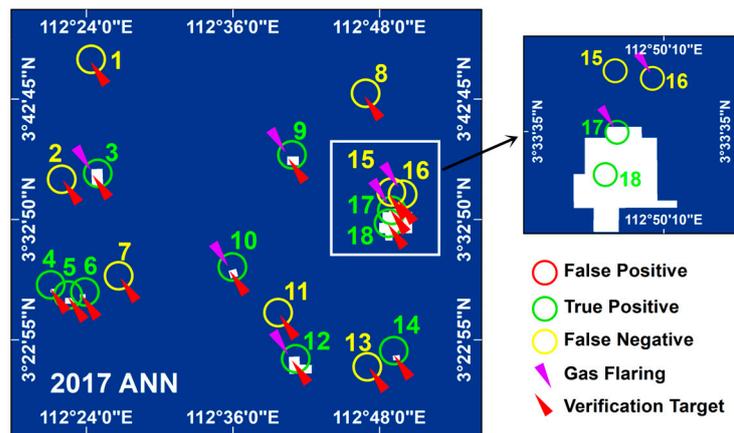


**Figure 4.** Extraction results of the targets in region 1, located in the Pearl River Mouth Basin: (a–f) are the extraction results in 2016; (g–l) are the extraction results in 2017. Red arrows are the validation targets; green circles are the correctly extracted targets; yellow circles are the omissive targets; red circles are the incorrectly extracted targets.

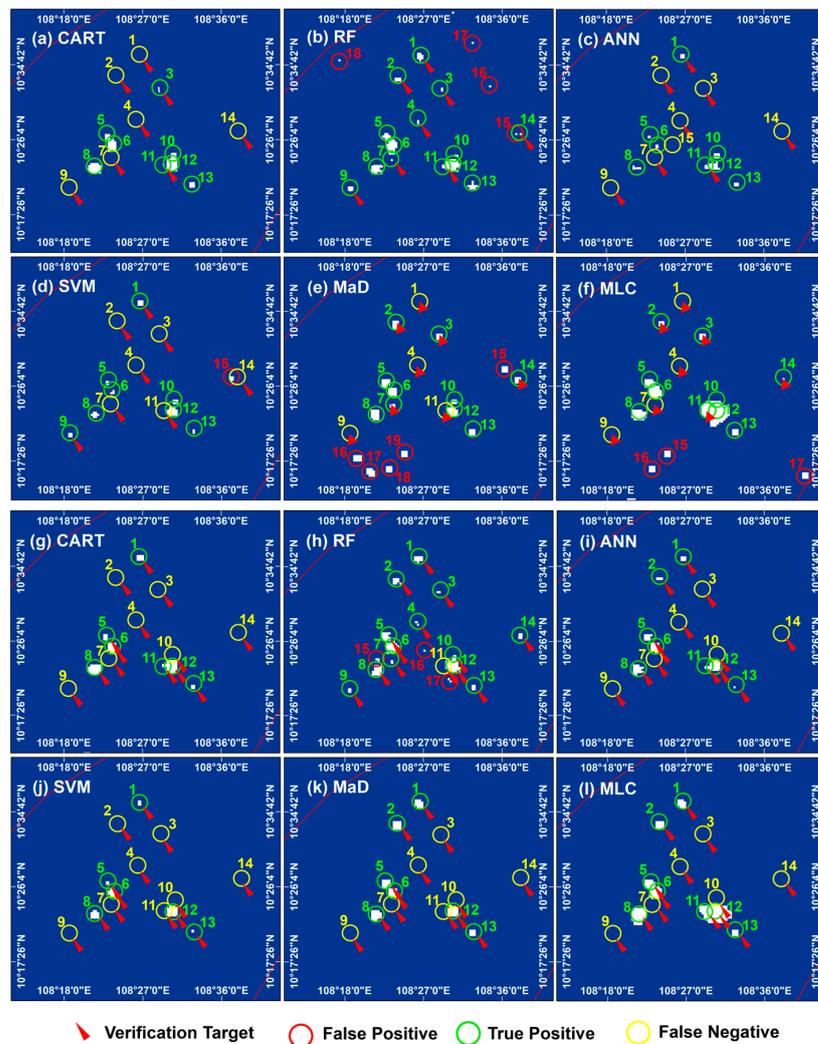


**Figure 5.** Extraction results of the targets in region 2, located in the Zengmu Basin: (a–f) are the extraction results in 2016; (g–l) are the extraction results in 2017. Red arrows are the validation targets; green circles are the correctly extracted targets; yellow circles are the omissive targets; and red circles are the incorrectly extracted targets.

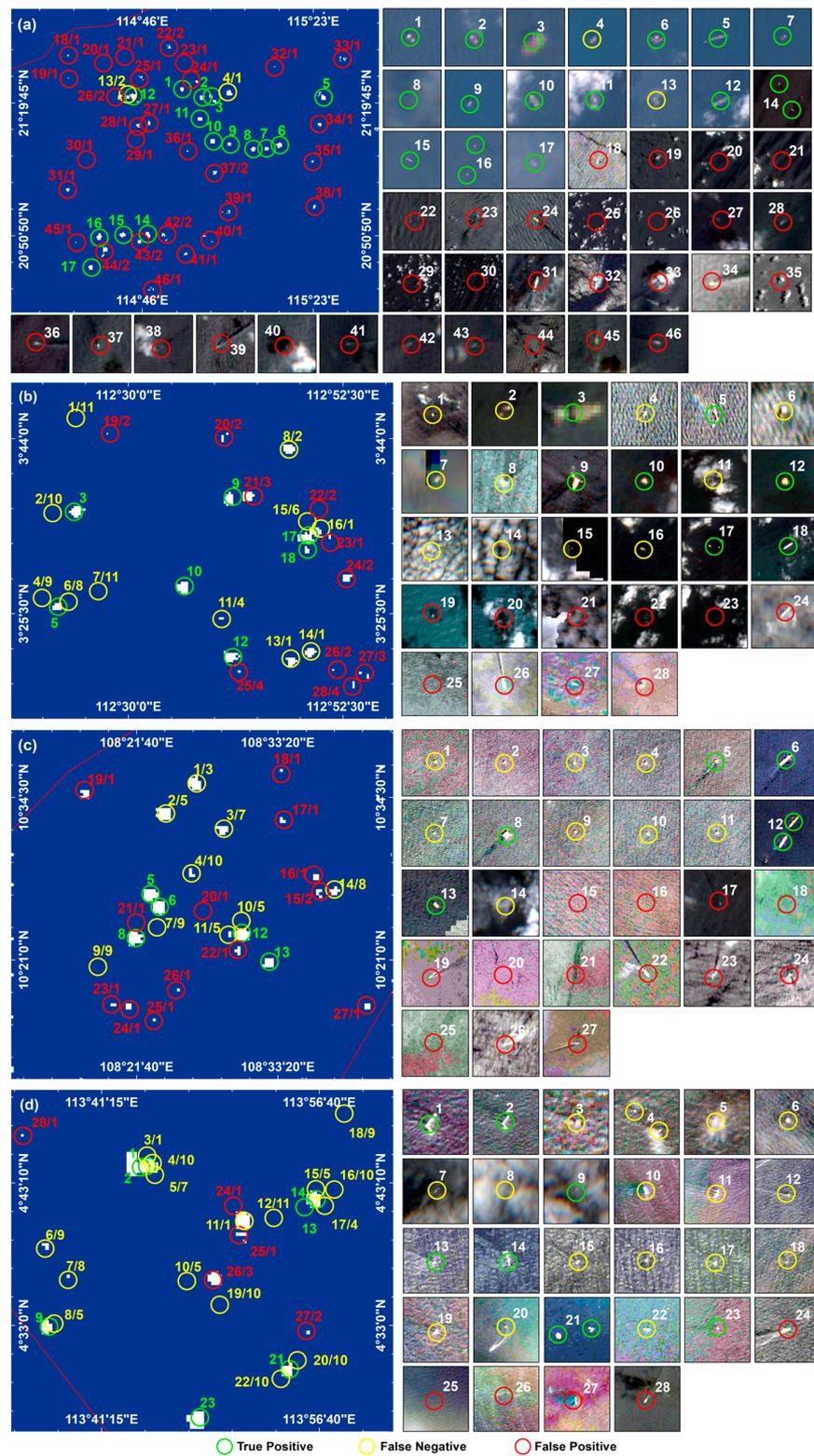
The results (Figure 4) show that a total of 17 OHE targets were contained in region 1. The majority of these targets were correctly extracted, except for target 13 which had a high frequency of omissions. This was because targets 12 and 13 were spatially closer to each other but target 13 had a lower light radiation intensity. Therefore, it was easy to overlook target 13 or combine both targets 12 and 13 into one during the extraction process. Furthermore, by combining the Sentinel-2 high-resolution image data (Figure 5), we found that the intensive operations of ships and fishing boats in this region were the primary reason of the incorrect extraction (Figure 8a).



**Figure 6.** Comparison of the gas flaring data and the 2017 ANN extraction results. Red arrows are the validation targets; purple arrows are the 2017 natural gas flaring products; green circles are the correctly extracted targets; yellow circles are the omissive targets; and red circles are the incorrectly extracted targets.



**Figure 7.** Extraction results of the targets in region 3, located in the Mekong Basin: (a–f) are the extraction results in 2016; (g–l) are the extraction results in 2017. Red arrows are the validation targets; green circles are the correctly extracted targets; yellow circles are the omissive targets; and red circles are the incorrectly extracted targets.



**Figure 8.** Synthetic images of the target extraction results: (a) is the synthetic image of the extraction results of region 1 in 2017; (b) is the synthetic image of the extraction results of region 2 in 2017; (c) is the synthetic image of the extraction results of region 3 in 2017; (d) is the synthetic image of the extraction results of region 4 in 2017. The front and back of the “/” represent the label of each target and the frequency of the omission or incorrect extraction, respectively. The green circles are the correctly extracted targets, the yellow circles are the omissive targets, and the red circles are the incorrectly extracted targets.

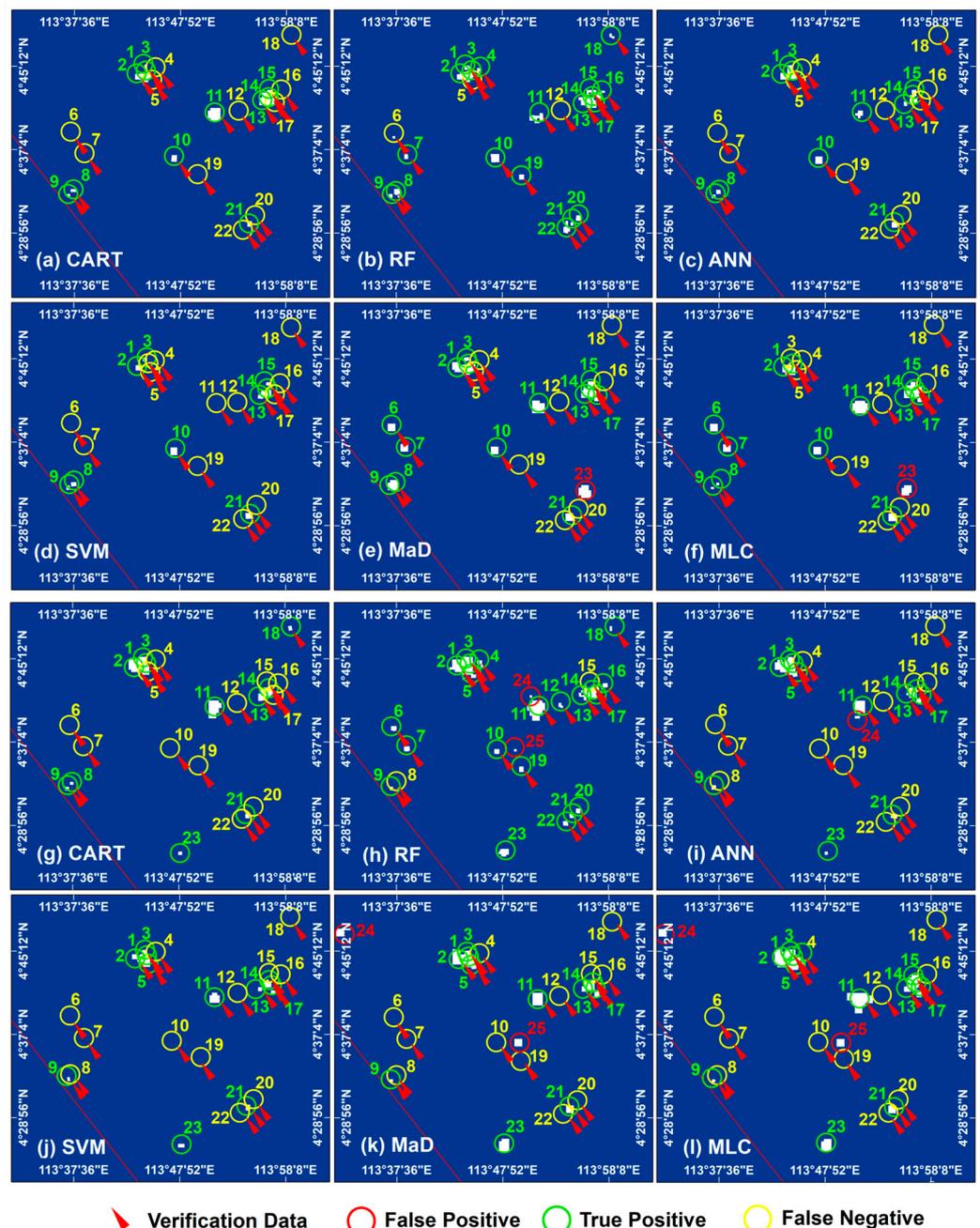
Region 2 was located within the offshore SK305 hydrocarbon block in the Zengmu Basin. Of the 18 OHE targets in this region, targets 1, 2, 4, 6, 7, and 15 were repeatedly missed in the extraction results of the different machine learning algorithms, as shown in Figure 5. By analyzing the VIIRS night light data, we found that these targets not only had small spots but also had extremely low light radiation intensity values. So, their distinction from the seawater background was not obvious, which in turn led to the omission of the extraction. Targets 4, 5, and 6, as well as 15, 16, and 17, were difficult to be distinguished in the VIIRS images due to their concentrated distribution. This was another crucial factor of the omission. In addition, in 2017, the ANN model omitted target 16, most likely since the VIIRS monthly data weakened the characteristic elements of the non-sustained flaring target during the synthesis process. In contrast, the global gas flaring product ([https://eogdata.mines.edu/products/vnf/global\\_gas\\_flare.html](https://eogdata.mines.edu/products/vnf/global_gas_flare.html) (accessed on 11 March 2023)) (Figure 6) released by the Payne Institute for Public Policy's Earth Observation Group based on the Nightfire data was calibrated for the targets with significant gas flares. Therefore, we combined the extracted results with this data (Figure 6). As can be seen in the figure, the product was able to effectively supplement the omission of target 16. The target 9 not only has a larger radiation range, but also a higher radiation intensity, which caused it to be mistakenly extracted as two targets multiple times in 2017. Furthermore, we discovered that targets 20, 24, 25, 26, 27, and 28 in Figure 9b were extracted incorrectly several times. This was due to the fact that this area was close to the Malaysian coast and was more disturbed by the lights of approaching ships.

Region 3 was located in hydrocarbon blocks 15-1 and 01 in the Mekong Basin. According to a combination of the target extraction results (Figure 7) and Sentinel-2 images (Figure 8c), we found that the 14 OHE targets in the area were more densely distributed, however there were significant differences in the light radiation intensity of the different targets. Targets 5, 6, 8, 12, and 13 had stronger light radiation values and consequently better extraction results, while targets 3, 7, 9, and 14 were more or less missed in the extraction results of the different machine learning algorithms due to their lower light radiation intensity. Targets 10 and 11 were distributed close to target 12, but their pixel values were both lower than target 12. Therefore, they were missed several times. Furthermore, since the Mekong Basin was close to the largest port in southern Vietnam (Ho Chi Minh Port), there were more intense ship activities in the area, which easily led to incorrect extractions for targets 17, 19, 22, 23, 24, 26, and 27, as shown in Figure 8.

Region 4 was located within the Tukai Timur field in the Brunei-Sabah Basin, which contained 22 OHE targets in 2016 and one new target in 2017. From the extraction results (Figure 9) and the comparison results (Figure 8d), it can be seen that the OHE targets with concentrated distribution were highly susceptible to omissions, such as targets 4, 5, 15, 16, and 17. Secondly, the targets with a weaker light radiation intensity were also not easily extracted, such as targets 6, 7, 8, 12, 18, 19, 20, and 22. The number of incorrectly extracted targets in this region was relatively less, which was mainly due to the disturbance of lights from navigable ships.

#### 4.2. Evaluation of Quantitative Accuracy

To quantitatively evaluate the reliability of the six machine learning algorithms, we first counted the number of correct, omission, and incorrect extractions of the OHE targets in the four regions. Then, the extraction results of the six machine learning models were verified using the F1-measure comprehensive evaluation index (Tables 1–4), and the extraction accuracy of each model was compared.



**Figure 9.** Extraction results of the targets in region 4, located in the Brunei-Sabah Basin: (a–f) are the extraction results in 2016; (g–l) are the extraction results in 2017. Red arrows are the validation targets; green circles are the correctly extracted targets; yellow circles are the omissive targets; and red circles are the incorrectly extracted targets.

The accuracy evaluation results showed that the extraction accuracy of the RF model was much higher than the other algorithms. In contrast, the CART, ANN, SVM, MLC, and MaD were influenced by regional differences, resulting in different rankings for their extraction accuracy in the different regions. Therefore, to better evaluate the comprehensive performance of these six models, we ranked them based on the mean extraction accuracy of each model in the four regions and two years.

According to Table 5, we found that the RF could better cope with the outliers and noise and performed outstandingly among the six algorithms, with an average F1-measure of 90.74%. While the ANN (77.06%), CART (75.16%), and SVM (72.18%) also had relatively good extraction effects, the omission rate was higher for the targets with a lower light

radiation intensity. The algorithms with a lower accuracy were the MLC and MaD, with a mean F1-measure of only 71.81% and 70.06%, respectively. This was because these two algorithms were susceptible to the interference from the false alarm targets such as ships, which led to considerable incorrect extraction and resulted in a lower accuracy.

#### 4.3. Feature Importance Evaluation

To further quantitatively analyze the differences in the feature values of the various target classes, we separately recorded the distribution ranges of the five feature values corresponding to the correct targets, incorrect targets, and omissive targets (Figures 10 and 11).

**Table 1.** Accuracy evaluation in region 1.

Model	Year	TP	FN	FP	P (%)	R (%)	F1 (%)
CART	2016	17	0	4	80.95	100	89.47
	2017	16	1	2	84.21	94.12	88.89
	Mean				82.58	97.06	89.18
RF	2016	17	0	1	89.47	100.00	94.44
	2017	16	1	2	94.12	94.12	94.12
	Mean				91.80	97.06	94.28
ANN	2016	17	0	3	85.00	100	91.89
	2017	15	2	1	93.75	88.24	90.91
	Mean				89.38	94.12	91.40
SVM	2016	16	2	1	88.89	94.12	91.43
	2017	16	1	2	84.21	94.12	88.89
	Mean				86.55	94.12	90.16
MaD	2016	17	0	13	56.67	100.00	72.34
	2017	17	0	14	54.84	100.00	70.83
	Mean				55.58	100.00	71.59
MLC	2016	16	1	12	57.14	94.12	71.11
	2017	17	0	15	53.13	100.00	69.39
	Mean				55.14	97.06	70.25

**Table 2.** Accuracy evaluation in region 2.

Model	Year	TP	FN	FP	P (%)	R (%)	F1 (%)
CART	2016	13	5	2	86.67	72.22	78.79
	2017	11	7	1	91.67	61.11	73.33
	Mean				89.17	66.67	76.06
RF	2016	15	3	0	100.00	83.33	90.91
	2017	15	3	1	93.75	83.33	88.24
	Mean				96.88	83.33	89.58
ANN	2016	13	5	2	86.67	72.22	78.79
	2017	12	6	0	100.00	66.67	80.00
	Mean				93.34	69.45	79.34
SVM	2016	10	8	0	100.00	55.56	71.43
	2017	11	7	1	91.67	61.11	73.33
	Mean				95.84	58.34	72.38
MaD	2016	13	5	4	76.47	72.22	74.29
	2017	12	6	4	75.00	66.67	70.59
	Mean				75.74	69.45	72.44
MLC	2016	13	5	3	72.22	72.22	72.22
	2017	13	5	4	76.47	72.22	74.29
	Mean				74.35	72.22	73.26

**Table 3.** Accuracy evaluation in region 3.

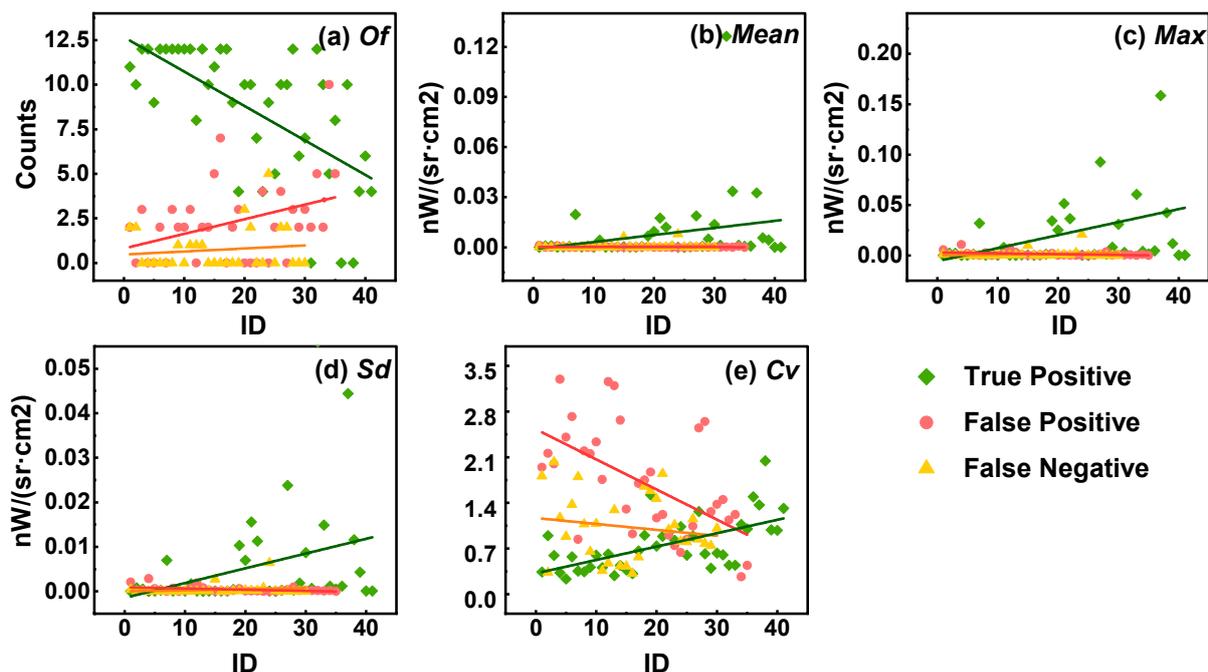
Model	Year	TP	FN	FP	P (%)	R (%)	F1 (%)
CART	2016	8	6	0	100.00	57.14	72.73
	2017	7	7	0	100.00	50.00	66.67
	Mean				100.00	53.57	69.70
RF	2016	14	0	4	77.78	100.00	87.50
	2017	13	1	3	81.25	92.86	86.67
	Mean				79.52	96.43	87.09
ANN	2016	8	6	0	100.00	57.14	72.73
	2017	8	6	0	100.00	57.14	72.73
	Mean				100.00	57.14	72.73
SVM	2016	8	6	1	88.89	57.14	69.57
	2017	6	8	0	100.00	42.86	60.00
	Mean				94.45	50.00	64.79
MaD	2016	10	4	5	66.67	71.43	68.97
	2017	7	7	0	100.00	50.00	66.67
	Mean				83.34	60.72	67.83
MLC	2016	10	4	3	76.92	71.43	74.07
	2017	8	6	0	100.00	57.14	72.73
	Mean				88.46	64.29	73.40

**Table 4.** Accuracy evaluation in region 4.

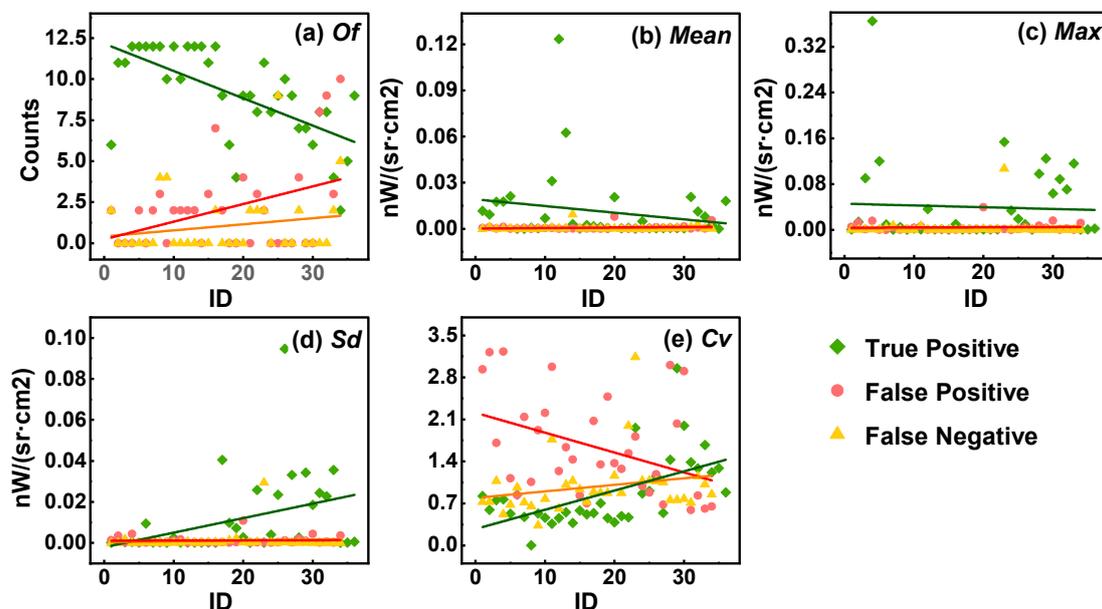
Model	Year	TP	FN	FP	P (%)	R (%)	F1 (%)
CART	2016	11	11	0	100.00	50.00	66.67
	2017	11	12	0	100.00	47.83	64.71
	Mean				100.00	48.92	65.69
RF	2016	19	3	0	100.00	86.36	92.68
	2017	21	2	2	91.3	91.30	91.30
	Mean				95.65	88.83	91.99
ANN	2016	11	11	0	100.00	50.00	66.67
	2017	11	12	1	91.67	47.83	62.86
	Mean				95.84	48.92	64.77
SVM	2016	9	13	0	100.00	40.91	58.06
	2017	11	12	0	100.00	47.83	64.71
	Mean				100.00	44.37	61.39
MaD	2016	14	8	1	93.33	63.64	75.68
	2017	11	12	2	84.62	47.83	61.12
	Mean				88.98	55.74	68.40
MLC	2016	13	9	1	92.86	59.09	72.22
	2017	13	10	2	86.67	56.52	68.42
	Mean				89.77	57.81	70.32

**Table 5.** Classification accuracy and mean value of each model in the study area.

Model	Region 1 (%)	Region 2 (%)	Region 3 (%)	Region 4 (%)	Mean (%)
CART	89.18	76.06	69.7	65.69	75.16
RF	94.28	89.58	87.09	91.99	90.74
ANN	91.4	79.34	72.73	64.77	77.06
SVM	90.16	72.38	64.79	61.39	72.18
MaD	71.59	72.44	67.83	68.4	70.07
MLC	70.25	73.26	73.4	70.32	71.81



**Figure 10.** Feature distribution of each target in 2016: (a) is the occurrence frequency distribution; (b) is the mean distribution; (c) is the maximum distribution; (d) is the standard deviation distribution; and (e) is the coefficient of variation distribution. The green diamond indicates the correctly extracted targets, the red circle is the incorrectly extracted targets, and the yellow triangle is the omissive targets. The green line is the fitted line of the feature distribution for the correct targets; the red line is the fitted line of the feature distribution for the incorrect targets, and the orange line is the fitted line of the feature distribution for the omissive targets.



**Figure 11.** Feature distribution of each target in 2017: (a) is the occurrence frequency distribution; (b) is the mean distribution; (c) is the maximum distribution; (d) is the standard deviation distribution; and (e) is the coefficient of variation distribution. The green diamond indicates the correctly extracted targets, the red circle is the incorrectly extracted targets, and the yellow triangle is the omissive targets. The green line is the fitted line of the feature distribution for the correct targets; the red line is the fitted line of the feature distribution for the incorrect targets, and the orange line is the fitted line of the feature distribution for the omissive targets.

It can be seen in Figures 10 and 11 that there were significant differences in the distribution ranges of the various features between the correctly extracted targets and the incorrectly or omitted extracted targets. Specifically, the frequency feature of the correctly extracted targets was much higher than the incorrectly and omitted extracted targets, with most of the correctly extracted targets appearing at a frequency of five or more in the temporal images. On the other hand, the mean, maximum, and standard deviation feature values for the incorrectly and omitted extracted targets are close to 0, which distinguished them from the correctly extracted targets. In addition, the coefficient of variation for the correctly extracted targets was often lower than the incorrectly extracted and omitted targets, mainly due to the relatively stable brightness characteristics of the OHE targets in the temporal images.

Combining the results analyses from each study area in Section 4.1, we found that the OHE targets that were spatially close and had low pixel values were more likely to be omitted. The targets with a high occurrence frequency and low standard deviation values were more likely to be falsely extracted, and they were mostly distributed around the targets with larger pixel radiation areas and located near the coastline.

To further determine which type of features contribute more to the results, the study selected the feature evaluation method built into the random forest algorithm with the highest extraction accuracy and ranked the importance of the five features. This effectively extracted the most distinctive features and provided ideas for the selection of the target features for the OHE. The magnitude of the calculated feature contribution varied slightly between the regions and years due to the regional differences. Therefore, we used the mean value to reflect the final importance of the features. First, we obtained the contribution values of each feature within the four study regions in the GEE platform in 2016 and 2017. Based on this, we calculated the average of the feature contribution values for the two periods. Finally, we calculated the average value of the feature values for the four regions as the final feature contribution evaluation results.

From the evaluation results (Figure 12), the three most important features that distinguished the OHE targets from the others were the coefficient of variation, mean, and standard deviation, with percentages of 26%, 25%, and 24%, respectively. The fourth ranked feature was the maximum value, with an importance share of 22%. The occurrence frequency contributed the least, with an importance share of only 4%.

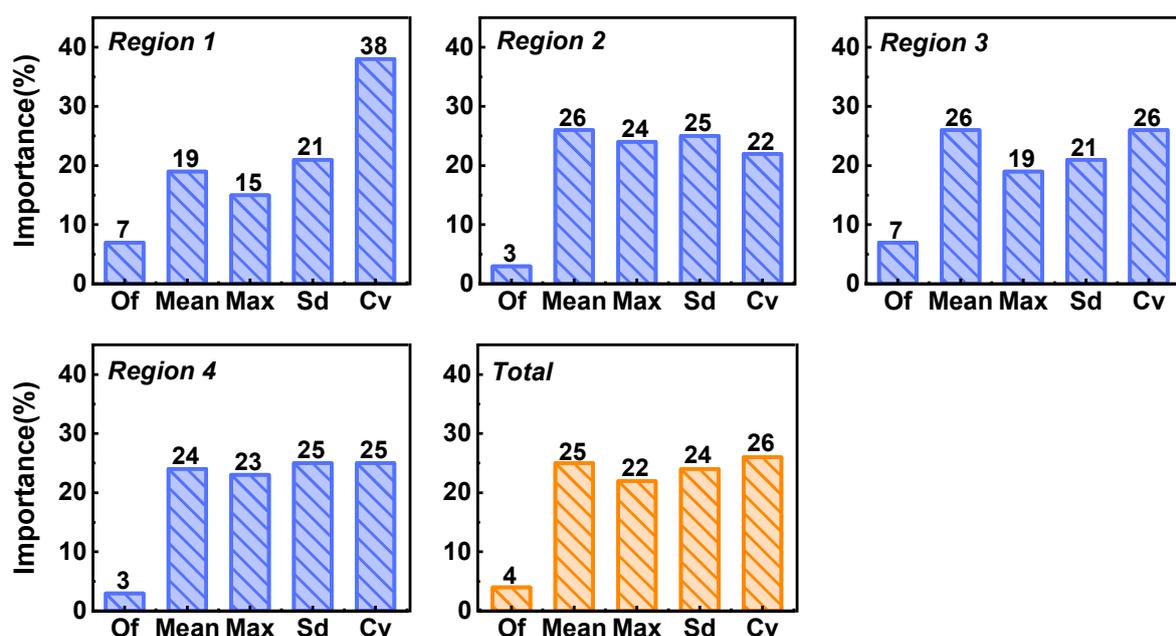


Figure 12. The feature importance of the four experimental areas and their average values.

In the time-series VIIRS images, the fluctuation degree of the light radiation intensity of the OHE and other targets showed significant differences. Among them, the moving targets, such as ships, were not fixed in position, and the light radiation intensity varied greatly. While the OHE targets were generally constant in position, the light radiation intensity fluctuated, but was relatively stable. Although, both the standard deviation and coefficient of variation could reflect the fluctuation, the coefficient of variation could eliminate the effect of the mean difference, and thus better measure the magnitude of the variation of the light radiation intensity for the moving and fixed targets. Therefore, the coefficient of variation feature became the most important feature to distinguish the OHE targets from the other targets. In addition, the average brightness values of the OHE targets were also different from those of the other targets. The pixel values of the background area were usually much lower than those of the target area, so the average value feature also played a more important role in the extraction of the OHE targets.

## 5. Discussion

In this study, OHE targets in four regions were extracted based on VIIRS time-series night light images using a feature increment strategy combined with six machine learning algorithms. By analyzing the error sources in the different regions and quantitatively evaluating the feature extraction performance of each model, we found that the targets with a weaker light radiation intensity were more likely to be missed, and the more frequently occurring ships were more mistakenly identified. Moreover, the comparative results of the models showed that the RF model outperformed the ANN, CART, SVM, MLC, and MaD models with an accuracy of 90.74%. Furthermore, the evaluation analysis of the feature importance indicated that the coefficient of variation feature was the most important feature in the machine learning-based extraction method to distinguish the OHE targets, followed by the mean, standard deviation, maximum, and frequency features.

The excellent classification performance and robustness of the RF algorithm were verified in landslide prediction [50,51], precipitation downscaling [52], soil downscaling [53], and land cover classification [54]. Wang et al. [23] also validated its feasibility and effectiveness in an OHE target extraction study. Based on this study, we further demonstrated the superior performance of the RF algorithm by comparing and evaluating the extraction performance of the different machine learning algorithms to the OHE targets. This was mainly because the random forest model had built-in randomly selected features and testing methods, which also randomly selects the features while selecting the samples. Thus, it was not over-fitted when the number of trees increased, and coped better with outliers and noise [55,56]. Therefore, the RF algorithm was widely used in long time-series remote sensing datasets [57,58].

Among the existing time-series-based methods for maritime target detection, the frequency feature constructed based on the “position invariance” were one of the most crucial parameters for extracting the fixed maritime targets, while the remaining features were mostly used as secondary indicators. [20,59]. However, after conducting the feature importance analysis on the constructed multi-dimensional features, we found that the coefficient of variation feature, rather than the frequency feature, had the greatest impact on the results extracted by the machine learning model. This was extremely different from existing studies. This was because, in sea areas with a high intensity of ship activities, moving targets such as ships are highly likely to appear multiple times at the same location in different periods, resulting in a large number of spurious high-frequency features in the time-series images. In addition, the frequency feature was highly susceptible to the influence of the target light radiation intensity, and it was difficult to effectively and statistically analyze the frequency features of the OHE targets when the pixel brightness values were extremely low. In contrast, the coefficient of variation distinguished the maritime fixed targets from the moving targets using the magnitude of brightness variation in the time series, which was more reliable and universal than the frequency feature.

Therefore, the coefficient of variation feature played a greater role in the machine learning-based OHE target extraction.

This study evaluated the performance of the different machine learning algorithms and the degree of importance of the multi-dimensional features of VIIRS night light remote sensing images based on a long time series. While providing a reference basis for the selection of the data sources for the maritime fixed target extraction, it also provided ideas for the improvement of the research methods. VIIRS images have high temporal resolution, which plays an important role in constructing the time-series features of the OHE targets and performing the dynamic change detection of the OHE targets. However, considering that the spatial resolution of the VIIRS data used in the study was about 500 m, the size of the OHE targets was generally within 120 m. This might have caused the extracted OHE targets to be less than the actual data and the existence of mixed targets, which was not conducive to the subsequent fine monitoring of carbon emission activities. Therefore, subsequent studies should consider decomposing the extraction results by overlaying the multi-source and multi-sensor remote sensing data to improve the fineness of the extraction. In addition, as the extraction results of the machine learning models are closely related to the construction of the feature variables, having too many feature variables or a low feature importance may affect the extraction performance. Therefore, in the subsequent study, we will consider adding other features or performing a feature dimensionality reduction to further filter the important features that distinguish the OHE targets. On the other hand, there are fewer feature elements of the OHE targets that can be acquired based on low spatial resolution VIIRS imagery. To address this issue, our subsequent research will take advantage of high-resolution images and combine the advantages of deep learning algorithms for feature learning to carry out research about the extraction of information, such as the OHE target attributes, types, and sources of carbon emissions. It is also important to note that the functional type of OHE targets serves as a crucial basis for evaluating and analyzing OHE carbon emissions. Therefore, we will perform the extraction of the OHE target functional type information based on the OHE target distribution data to build the foundation for the OHE carbon emission activity level assessment.

## 6. Conclusions

In order to explore the impact of machine learning algorithms, target features, and regional differences on the OHE target extraction, this study selected four typical experimental areas and evaluated the reliability of six machine learning algorithms for extracting OHE targets using time-series VIIRS night light remote sensing images. The error influencing factors were further analyzed and the degree of importance of the multi-dimensional features was revealed. The results showed that the random forest algorithm exhibited a better extraction performance than the other algorithms. The OHE targets with a closer spatial distribution and a lower light intensity were easily missed, while the ship targets with a higher frequency of occurrence were easily extracted incorrectly. In the machine learning-based OHE target extraction, the coefficient of variation was the most important feature to distinguish the OHE targets from the other targets.

This study provides an important reference for improving the OHE target extraction method and also provides important information support for subsequent OHE carbon emission activity analyses. The next step should focus on strengthening the collaborative role of the multi-source assessment of carbon emission activities and multi-sensor remote sensing imagery in order to improve the precision of the OHE target extraction. By combining the Nightfire data with the information related to OHE carbon emission monitoring, further evaluation of the carbon emission situation of the OHE targets can be carried out.

**Author Contributions:** Conceptualization, Q.W.; methodology, R.M. and W.W.; software, W.W.; formal analysis, R.M. and Y.C.; investigation, R.M. and Y.C.; resources, Q.W.; data curation, Q.W. and R.M.; writing—original draft preparation, R.M.; writing—review and editing, R.M. and Q.W.; visualization, R.M., N.L. and W.W.; supervision, N.L. and Q.W.; project administration, Q.W.; funding acquisition, Q.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the “Ph.D. Programs Foundation of Shandong Jianzhu University, grant number XNBS1984”, and the “State Key Laboratory of Resources and Environmental Information System”.

**Data Availability Statement:** The field survey data and offshore platform records presented in this study are available upon request from the corresponding authors. The data are not publicly available due to commercial secrets.

**Acknowledgments:** The authors thank the Earth Observation Group of the Payne Institute for Public Policy at the Colorado School of Mines for providing the VIIRS/DNB monthly synthetic images; the State Key Laboratory of Resource and Environmental Information Systems (SKLREIS), Institute of Geographical Sciences and Resources, Chinese Academy of Sciences for providing the offshore platform record data; and Google Earth Engine for providing the Sentinel-2 images.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Meinshausen, M.; Meinshausen, N.; Hare, W.; Raper, S.C.B.; Frieler, K.; Knutti, R.; Frame, D.J.; Allen, M.R. Greenhouse-gas emission targets for limiting global warming to 2 °C. *Nature* **2009**, *458*, 1158–1162. [CrossRef] [PubMed]
2. Salman, M.; Long, X.; Wang, G.; Zha, D. Paris climate agreement and global environmental efficiency: New evidence from fuzzy regression discontinuity design. *Energy Policy* **2022**, *168*, 113128. [CrossRef]
3. Xue, H. Research on distributed streaming parallel computing of large scale wind DFigs from the perspective of Ecological Marxism. *Energy Rep.* **2022**, *8*, 304–312. [CrossRef]
4. Balat, M. Influence of coal as an energy source on environmental pollution. *Energy Sources Part A* **2007**, *29*, 581–589. [CrossRef]
5. EIA. Greenhouse Gases Effect on Climate. Available online: <https://www.eia.gov/energyexplained/energy-and-the-environment/greenhouse-gases-and-the-climate.php> (accessed on 5 November 2022).
6. IEA. International Energy Agency. Available online: <https://www.iea.org> (accessed on 6 November 2022).
7. Irakulis-Loitxate, I.; Gorrono, J.; Zavala-Araiza, D.; Guanter, L. Satellites Detect a Methane Ultra-emission Event from an Offshore Platform in the Gulf of Mexico. *Environ. Sci. Technol. Lett.* **2022**, *9*, 520–525. [CrossRef]
8. EIA. Offshore Production Nearly 30% of Global Crude Oil Output in 2015. Available online: <https://www.eia.gov/todayinenergy/detail.php?id=28492> (accessed on 8 November 2022).
9. Watson, S.M. Greenhouse gas emissions from offshore oil and gas activities—Relevance of the Paris Agreement, Law of the Sea, and Regional Seas Programmes. *Ocean Coast. Manag.* **2020**, *185*, 104942. [CrossRef]
10. Nguyen, T.-V.; Barbosa, Y.M.; da Silva, J.A.M.; de Oliveira Junior, S. A novel methodology for the design and optimisation of oil and gas offshore platforms. *Energy* **2019**, *185*, 158–175. [CrossRef]
11. Chen, P.; Wang, J.; Li, D. Oil Platform Investigation by Multi-Temporal SAR Remote Sensing Image. In *SPIE Remote Sensing International Society for Optics and Photonics*; International Society for Optics and Photonics: Prague, Czech Republic, 2011; Volume 28.
12. Cheng, L.; Yang, K.; Tong, L.; Liu, Y.; Li, M. Invariant triangle-based stationary oil platform detection from multitemporal synthetic aperture radar data. *J. Appl. Remote Sens.* **2013**, *7*, 73537. [CrossRef]
13. Falqueto, L.E.; Sa, J.A.S.; Paes, R.L.; Passaro, A. Oil Rig Recognition Using Convolutional Neural Network on Sentinel-1 SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1329–1333. [CrossRef]
14. Liu, C.; Yang, J.; Ou, J.; Fan, D. Offshore Oil Platform Detection in Polarimetric SAR Images Using Level Set Segmentation of Limited Initial Region and Convolutional Neural Network. *Remote Sens.* **2022**, *14*, 1729. [CrossRef]
15. Zhang, J.; Wang, Q.; Su, F. Automatic Extraction of Offshore Platforms in Single SAR Images Based on a Dual-Step-Modified Model. *Sensors* **2019**, *19*, 231. [CrossRef]
16. Liu, Y.; Hu, C.; Dong, Y.; Xu, B.; Zhan, W.; Sun, C. Geometric accuracy of remote sensing images over oceans: The use of global offshore platforms. *Remote Sens. Environ.* **2019**, *222*, 244–266. [CrossRef]
17. Liu, Y.; Sun, C.; Sun, J.; Li, H.; Zhan, W.; Yang, Y.; Zhang, S. Satellite data lift the veil on offshore platforms in the South China Sea. *Sci. Rep.* **2016**, *6*, 33623. [CrossRef]
18. Xing, Q.; Meng, R.; Lou, M.; Bing, L.; Liu, X. Remote Sensing of Ships and Offshore Oil Platforms and Mapping the Marine Oil Spill Risk Source in the Bohai Sea. *Aquat. Procedia* **2015**, *3*, 127–132. [CrossRef]
19. Anejionu, O.; Blackburn, G.; Whyatt, D. Satellite survey of gas flares: Development and application of a Landsat-based technique in the Niger Delta. *Int. J. Remote Sens.* **2014**, *35*, 1900–1925. [CrossRef]
20. Liu, Y.; Sun, C.; Yang, Y.; Zhou, M.; Zhan, W.; Cheng, W. Automatic extraction of offshore platforms using time-series Landsat-8 Operational Land Imager data. *Remote Sens. Environ.* **2016**, *175*, 73–91. [CrossRef]
21. Zhu, H.; Jia, G.; Zhang, Q.; Zhang, S.; Lin, X.; Shuai, Y. Detecting Offshore Drilling Rigs with Multitemporal NDWI: A Case Study in the Caspian Sea. *Remote Sens.* **2021**, *13*, 1576. [CrossRef]
22. Anejionu, O.C.D.; Blackburn, G.A.; Whyatt, J.D. Detecting gas flares and estimating flaring volumes at individual flow stations using MODIS data. *Remote Sens. Environ.* **2015**, *158*, 81–94. [CrossRef]

23. Wang, Q.; Wu, W.; Su, F.; Xiao, H.; Wu, Y.; Yao, G. Offshore Hydrocarbon Exploitation Observations from VIIRS NTL Images: Analyzing the Intensity Changes and Development Trends in the South China Sea from 2012 to 2019. *Remote Sens.* **2021**, *13*, 946. [[CrossRef](#)]
24. Oliva, P.; Schroeder, W. Assessment of VIIRS 375m active fire detection product for direct burned area mapping. *Remote Sens. Environ.* **2015**, *160*, 144–155. [[CrossRef](#)]
25. Elvidge, C.; Zhizhin, M.; Baugh, K.; Hsu, F.-C.; Ghosh, T. Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data. *Energies* **2015**, *9*, 14. [[CrossRef](#)]
26. Elvidge, C.D.; Ziskin, D.; Baugh, K.E.; Tuttle, B.T.; Ghosh, T.; Pack, D.W.; Erwin, E.H.; Zhizhin, M. A Fifteen Year Record of Global Natural Gas Flaring Derived from Satellite Data. *Energies* **2009**, *2*, 595–622. [[CrossRef](#)]
27. Lu, W.; Liu, Y.; Wang, J.; Xu, W.; Wu, W.; Liu, Y.; Zhao, B.; Li, H.; Li, P. Global proliferation of offshore gas flaring areas. *J. Maps* **2020**, *16*, 396–404. [[CrossRef](#)]
28. Elvidge, C.; Zhizhin, M.; Hsu, F.-C.; Baugh, K. VIIRS Nightfire: Satellite Pyrometry at Night. *Remote Sens.* **2013**, *5*, 4423–4449. [[CrossRef](#)]
29. Liu, Y.; Hu, C.; Zhan, W.; Sun, C.; Murch, B.; Ma, L. Identifying industrial heat sources using time-series of the VIIRS Nightfire product with an object-oriented approach. *Remote Sens. Environ.* **2018**, *204*, 347–365. [[CrossRef](#)]
30. Wenzhao, Z.; Houhe, Z.; Chunrong, L.I.; Han, Y.; Fanyi, L.I.; Jing, H. Petroleum Exploration History and Enlightenment in Pearl River Mouth Basin. *Xinjiang Pet. Geol.* **2021**, *42*, 346. [[CrossRef](#)]
31. Cullen, A. Reprint of: Nature and significance of the West Baram and Tinjar Lines, NW Borneo. *Mar. Pet. Geol.* **2014**, *58*, 674–686. [[CrossRef](#)]
32. Guo, J.; Yang, S.C.; Hu, W.B.; Song, S.; Wang, Y.B.; Wang, L. Difference Analysis of Hydrocarbon Generation in the Southern Part of the Western Continental Margin of the South China Sea. *Mar. Geo. Front.* **2021**, *37*, 1–7. (In Chinese) [[CrossRef](#)]
33. Morley, C.K.; King, R.; Hillis, R.; Tingay, M.; Backe, G. Deepwater fold and thrust belt classification, tectonics, structure and hydrocarbon prospectivity: A review. *Earth-Sci. Rev.* **2011**, *104*, 41–91. [[CrossRef](#)]
34. Bennett, M.M.; Smith, L.C. Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics. *Remote Sens. Environ.* **2017**, *192*, 176–197. [[CrossRef](#)]
35. Lee, J.-S.; Wen, J.-H.; Ainsworth, T.L.; Chen, K.-S.; Chen, A.J. Improved Sigma Filter for Speckle Filtering of SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 202–213. [[CrossRef](#)]
36. Wackerman, C.C.; Friedman, K.S.; Pichel, W.G.; Clemente-Colón, P.; Li, X. Automatic Detection of Ships in RADARSAT-1 SAR Imagery. *Can. J. Remote Sens.* **2001**, *27*, 371–378. [[CrossRef](#)]
37. Hansen, M.; Dubayah, R.; Defries, R. Classification trees: An alternative to traditional land cover classifiers. *Int. J. Remote Sens.* **1996**, *17*, 1075–1081. [[CrossRef](#)]
38. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *R News* **2002**, *2*, 5.
39. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
40. Kaper, M.; Meinicke, P.; Grossekoetter, U.; Lingner, T.; Ritter, H. BCI competition 2003-data set IIb: Support vector machines for the P300 speller paradigm. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1073–1076. [[CrossRef](#)]
41. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995; ISBN 978-147-572-442-4.
42. Tong, S.; Chang, E. Support vector machine active learning for image retrieval. In Proceedings of the MULTIMEDIA '01: Proceedings of the Ninth ACM International Conference on Multimedia, Ottawa, ON, Canada, 30 September–5 October 2001. [[CrossRef](#)]
43. Qi, H. Support Vector Machines and Application Research Overview. *Comput. Eng.* **2004**, *30*, 6–9. [[CrossRef](#)]
44. Hsu, K.; Gupta, H.V.; Sorooshian, S. Artificial Neural Network Modeling of the Rainfall-Runoff Process. *Water Resour. Res.* **1995**, *31*, 2517–2530. [[CrossRef](#)]
45. Chojaczyk, A.A.; Teixeira, A.P.; Neves, L.C.; Cardoso, J.B.; Guedes Soares, C. Review and application of Artificial Neural Networks models in reliability analysis of steel structures. *Struct. Saf.* **2015**, *52*, 78–89. [[CrossRef](#)]
46. De Maesschalck, R.; Jouan-Rimbaud, D.; Massart, D.L. The Mahalanobis distance. *Chemom. Intell. Lab. Syst.* **2000**, *50*, 1–18. [[CrossRef](#)]
47. Otukei, J.R.; Blaschke, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, S27–S31. [[CrossRef](#)]
48. Wang, B.; Li, C.; Pavlu, V.; Aslam, J. A Pipeline for Optimizing F1-Measure in Multi-Label Text Classification. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 913–918.
49. Sammut, C.; Webb, G.I. F1-Measure. In *Encyclopedia of Machine Learning and Data Mining*; Sammut, C., Webb, G.I., Eds.; Springer: Boston, MA, USA, 2017; p. 497.
50. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Bui, D.T.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* **2017**, *151*, 147–160. [[CrossRef](#)]
51. Youssef, A.M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Al-Katheeri, M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides* **2016**, *13*, 839–856. [[CrossRef](#)]

52. Jing, W.; Yang, Y.; Yue, X.; Zhao, X. A Comparison of Different Regression Algorithms for Downscaling Monthly Satellite-Based Precipitation over North China. *Remote Sens.* **2016**, *8*, 835. [[CrossRef](#)]
53. Liu, Y.; Yang, Y.; Jing, W.; Yue, X. Comparison of Different Machine Learning Approaches for Monthly Satellite-Based Soil Moisture Downscaling over Northeast China. *Remote Sens.* **2018**, *10*, 31. [[CrossRef](#)]
54. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [[CrossRef](#)]
55. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [[CrossRef](#)]
56. Yu, X.; Hyyppä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37. [[CrossRef](#)]
57. Tulbure, M.G.; Broich, M.; Stehman, S.V.; Kommareddy, A. Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sens. Environ.* **2016**, *178*, 142–157. [[CrossRef](#)]
58. Deng, Y.; Jiang, W.; Tang, Z.; Li, J.; Lv, J.; Chen, Z.; Jia, K. Spatio-Temporal Change of Lake Water Extent in Wuhan Urban Agglomeration Based on Landsat Images from 1987 to 2015. *Remote Sens.* **2017**, *9*, 270. [[CrossRef](#)]
59. Sun, C.; Liu, Y.; Zhao, S.; Jin, S. Estimating offshore oil production using DMSP-OLS annual composites. *ISPRS J. Photogramm. Remote Sens.* **2020**, *165*, 152–171. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.