

Article

A Novel Cargo Ship Detection and Directional Discrimination Method for Remote Sensing Image Based on Lightweight Network

Pan Wang, Jianzhong Liu *, Yinbao Zhang, Zhiyang Zhi, Zhijian Cai and Nannan Song

School of Geoscience and Technology, Zhengzhou University, Zhengzhou 450001, China; panwang@gs.zzu.edu.cn (P.W.); zhangyinbao2005@sina.com (Y.Z.); hpuzhiyang@126.com (Z.Z.); zzucj0720@163.com (Z.C.); zzu_snn123@163.com (N.S.)

* Correspondence: liujianzhong@zzu.edu.cn



Citation: Wang, P.; Liu, J.; Zhang, Y.; Zhi, Z.; Cai, Z.; Song, N. A Novel Cargo Ship Detection and Directional Discrimination Method for Remote Sensing Image Based on Lightweight Network. *J. Mar. Sci. Eng.* **2021**, *9*, 932. <https://doi.org/10.3390/jmse9090932>

Academic Editors: Marco Cococcioni and Pierre Lermusiaux

Received: 10 August 2021

Accepted: 23 August 2021

Published: 28 August 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/>).

Abstract: Recently, cargo ship detection in remote sensing images based on deep learning is of great significance for cargo ship monitoring. However, the existing detection network is not only unable to realize autonomous operation on spaceborne platforms due to the limitation of computing and storage, but the detection result also lacks the directional information of the cargo ship. In order to address the above problems, we propose a novel cargo ship detection and directional discrimination method for remote sensing images based on a lightweight network. Specifically, we design an efficient and lightweight feature extraction network called the one-shot aggregation and depthwise separable network (OSADSNet), which is inspired by one-shot feature aggregation modules and depthwise separable convolutions. Additionally, we combine the RPN with the K-Mean++ algorithm to obtain the K-RPN, which can produce a more suitable region proposal for cargo ship detection. Furthermore, without introducing extra parameters, the directional discrimination of the cargo ship is transformed into a classification task, and the directional discrimination is completed when the detection task is completed. Experiments on a self-built remote sensing image cargo ship dataset indicate that our model can provide relatively accurate and fast detection for cargo ships (*mAP* of 91.96% and prediction time of 46 ms per image) and discriminate the directions (north, east, south, and west) of cargo ships, with fewer parameters (model size of 110 MB), which is more suitable for autonomous operation on spaceborne platforms. Therefore, the proposed method can meet the needs of cargo ship detection and directional discrimination in remote sensing images on spaceborne platforms.

Keywords: deep learning; remote sensing image; cargo ship; directional discrimination; faster-RCNN; lightweight

1. Introduction

With the progress of science and technology and the development of world trade, economic globalization has become the trend of world economic development. The trade links between countries are becoming closer and closer, and maritime transportation has become the main mode of transportation in foreign trade because of its advantages of large volume and low-cost transportation. Cargo ships are the main means of transportation for sea transportation. However, with the diversification of cargo ship types, the larger and higher speed of cargo ships, and the increase of ports, the navigation environment of cargo ships is becoming more and more complicated, causing the course of cargo ships to deviate from their planned course. This leads to frequent problems such as channel block, cargo ships collisions, and increased navigation costs. Therefore, if we can grasp the navigation information of the cargo ship in time, it is of great significance for improving the navigation environment of the cargo ship, ensuring the safe navigation of the cargo ship, improving the navigation efficiency, and shortening the navigation time [1–4].

In recent years, with the continuous development of satellite remote sensing technology and the emergence of a large number of remote sensing images, facing the vast ocean, remote sensing images have the advantages of wide coverage, high spatial resolution, fast update speed, and low cost, which makes the use of remote sensing images for real-time monitoring of ships in the ocean important [5,6]. However, in the face of massive remote sensing images, if we only rely on manual interpretation to obtain cargo ship information in remote sensing images, it can no longer meet the needs of modern society because of the huge workload and low efficiency. Therefore, obtaining cargo ship target information quickly and accurately from massive remote sensing images has become an urgent problem to be solved.

Traditional ship detection algorithms in remote sensing images rely heavily on manual design features, which require designers to understand relevant professional knowledge. It is difficult to realize the rapid processing of massive amounts of remote sensing data, and the accuracy of cargo ships under a complex background is low [7–9]. Compared with traditional object detection algorithms, the algorithms based on deep learning can independently extract object features, avoiding the complexity of manual design features, and the extracted features are more robust. Therefore, deep learning makes it possible to intellectualize cargo ship detection in remote sensing images.

Recently, with the substantial improvement of computer hardware performance and the emergence of large-scale training samples, the deep learning techniques represented by convolutional neural networks have shown strong performance in object detection applications [10–12]. At present, object detection algorithms based on the convolutional neural network can be divided into two-stage object detection algorithms and one-stage object detection algorithms according to whether region proposals are generated. Two-stage object detection algorithms, such as R-CNN [13], fast-RCNN [14], and faster-RCNN [15], firstly extract region proposals on the feature map where there may be objects, then further classify and locate these areas, and finally the detection results are obtained. This type of method achieves a high detection accuracy, but it cannot meet real-time requirements. One-stage object detection algorithms, such as YOLO [16–19] and SSD [20], use the regression method to establish an object detection framework, which removes the operation of generating region proposals. This type of method has fast detection speed, but the detection accuracy is low.

In view of the excellent performance of deep learning in computer vision, deep learning has been widely used in ship detection in remote sensing images [21–23]. However, there are some problems in remote sensing images, such as diversified ship sizes and complex backgrounds, which bring difficulties to ship detection. Many scholars have put forward their own solutions to such problems. To efficiently detect ships with various scales, Li et al. [24] proposed a hierarchical selective filtering layer based on the faster-RCNN algorithm to map features in different scales to the same scale space. Although this method can simultaneously detect inshore and offshore ships with dozens of pixels to thousands of pixels, the detection results lack the directional information of the cargo ships. Aiming at the problem of high false alarm ratio and the great influence of the sea surface in traditional ship detection methods, Zhang et al. [25] adopted the idea of deep networks and proposed a fast regional-based convolutional neural network (R-CNN) method to detect ships from high-resolution remote sensing imagery. However, the proposed method contains a complicated pre-processing stage, which cannot meet the requirements of real-time detection. Wang et al. [26] proposed a fast-RCNN method based on the adversary strategy, which adopted the adversarial spatial transformer network (ASTN) module to improve the classification ability for ships in remote sensing images under complex backgrounds. However, it is hard to realize autonomous operation on spaceborne platforms.

Although automatic identification system (AIS) data can provide some information about the cargo ship, the information will be lost if the AIS does not keep the normal opening statement or fails, and some cargo ships are not even loaded with AIS. However,

the remote sensing satellite will not have a similar situation, so the remote sensing satellite image has become an important data source to obtain cargo ship information.

In the face of massive satellite remote sensing images, the transmission bandwidth from the satellite to ground is limited. Therefore, the real-time intelligent processing of remote sensing images on spaceborne platforms must be the development trend in the future. However, due to the limitation of computing and storage on the spaceborne platform, complex and huge object detection models have not been widely used. In particular, although the two-stage target detection algorithm has high accuracy, it has high model complexity and low detection efficiency, resulting in poor availability of the algorithm on the spaceborne platform.

The category and position information of cargo ships can be obtained by real-time detection of cargo ships in remote sensing images. However, due to the unique perspective of remote sensing images, the remote sensing images also contain rich directional information of cargo ship targets. If the directional information of cargo ship targets can be obtained at the same time, it not only can ensure the safe navigation of cargo ships but can also help cargo ships navigate accurately along the route and reduce navigation costs.

In this paper, cargo ships in remote sensing images are taken as the research objects, and the faster-RCNN algorithm is used as the basis to propose a novel cargo ship detection and directional discrimination method for remote sensing images based on a lightweight network. This model not only can accurately detect the category and position of cargo ship targets in remote sensing images in real time but can also discriminate the direction of cargo ships. Additionally, its lightweight advantage improves the availability of spaceborne platforms. The rest of the paper is summarized as follows: Section 2 introduces the proposed method, Section 3 describes the experimental dataset and the experimental results, and Section 4 summarizes the whole paper and puts forward some suggestions for future work.

2. Methodology

2.1. Faster-RCNN Network

Considering the application needs of cargo ship detection and direction recognition in remote sensing images, this paper adopts the faster-RCNN model with high comprehensive performance of detection accuracy and detection speed as the algorithm prototype. As shown in Figure 1, faster-RCNN is composed of three parts: feature extraction network, region proposal network (RPN), and region based convolutional neural networks (R-CNN). The feature extraction network uses convolution to extract the features of the images, and the extracted features are shared by the following RPN and R-CNN. The RPN is a fully convolutional neural network used to generate region proposals for potential cargo ships in the images, and the following ROI pooling layer is used to unify region proposals of different sizes. R-CNN is used for coordinating regression and classification of the extracted region of interest to realize the detection of the cargo ships.

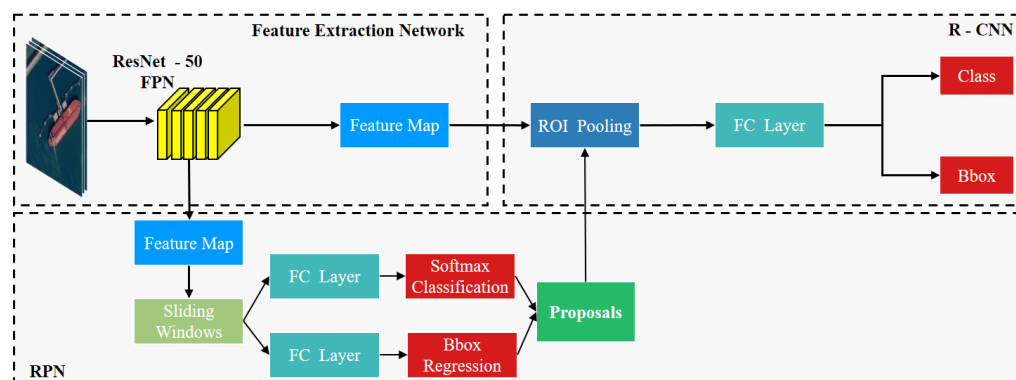


Figure 1. The architecture of faster-RCNN network.

2.2. Proposed Model

2.2.1. Model Overview

The model of cargo ship detection and directional discrimination in remote sensing images based on a lightweight network proposed in this paper is shown in Figure 2. In this model, one-shot aggregation module [27] and depthwise separable convolution [28] are used to build an efficient and lightweight feature extraction network, namely the one-shot aggregation and depthwise separable network (OSADSNet). Second, K-Means++ clustering algorithm is introduced into RPN to generate high-quality candidate boxes. Finally, without introducing extra parameters, the direction recognition problem of the cargo ship in remote sensing images is converted to a classification problem, and the directional discrimination is completed while the detection task is accomplished.

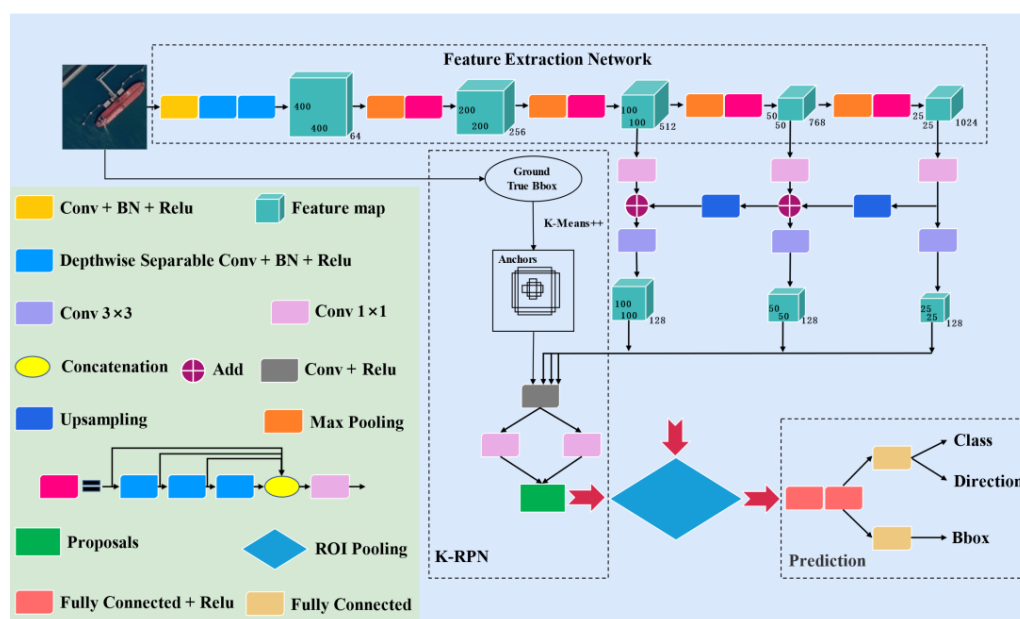


Figure 2. The overall model structure.

2.2.2. Feature Extraction Network

The feature extraction network is an important module in the object detection model, and its performance precisely influences the performance of the model. In this paper, we propose a novel and efficient feature extraction network, the one-shot aggregation and depthwise separable network (OSADSNet), to meet the specific needs of cargo ship detection in remote sensing images. The network draws on the idea of the one-shot aggregation module (Figure 3). Each layer of convolution has two connections: one is directly connected to the next layer to acquire a larger receptive field, and the other is connected to the last layer to aggregate features. Compared with the residual module [29], the network integrates more low-level features, and the connection is not the point-to-point addition of the feature map but the splicing between channels, which reduces the number of parameters between layers. Compared with the dense connection module [30], the connection is optimized, and all features are aggregated only before the last layer, which realizes the efficient aggregation of different convolutional layers, avoids feature redundancy, and performs at a faster speed. At the same time, the network uses depthwise separable convolution (Figure 4b) to replace the standard convolution (Figure 4a), which reduces the model parameters and effectively extracts features, finally realizing the light weight of the model. Table 1 shows the implementation details of OSADSNet.

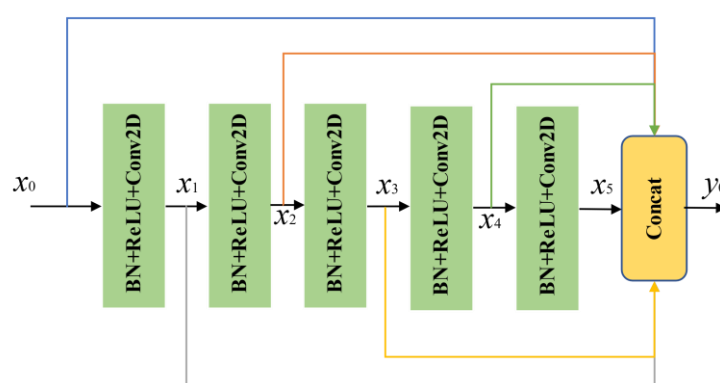


Figure 3. The one-shot aggregation module, where x_0 , x_1 , x_2 , x_3 , x_4 , and x_5 denote the input of each layer of the convolution layer; y_0 denotes the output; and concat denotes the feature concatenation operation. Rectified linear unit (ReLU) denotes an activation function. Batch normalization (BN) denotes centering and scaling activations within mini-batches. Conv2d denotes a 2D convolutional layer.

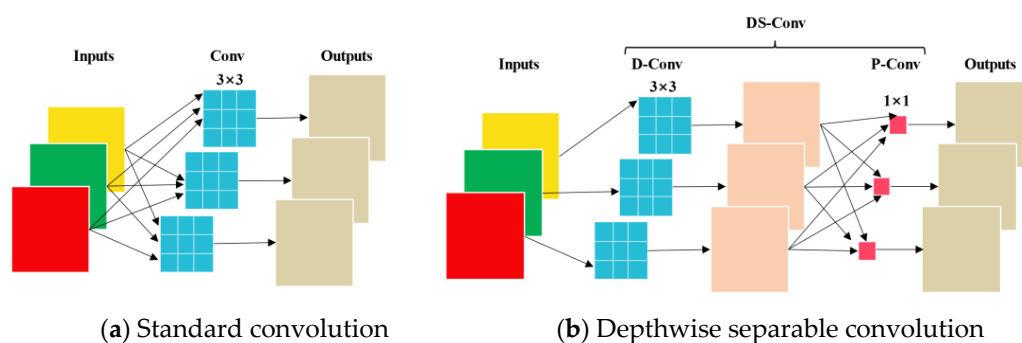


Figure 4. Comparison of different convolutions. Depthwise separable convolutions are special cases of standard convolution.

Table 1. An overview of the implementation details of OSADSNet.

Stage	Structure
Stage1	3×3 conv,64, stride = 2
	3×3 dsconv,64, stride = 1
	3×3 dsconv,64, stride = 1
	3×3 maxpool2d, stride = 2
Stage 2	$[3 \times 3 \text{ dsconv},128, \text{stride} = 1] \times 3$
	concat & 1×1 conv,256
Stage 3	$[3 \times 3 \text{ dsconv},160, \text{stride} = 1] \times 3$
	concat & 1×1 conv,512
Stage 4	$[3 \times 3 \text{ dsconv},192, \text{stride} = 1] \times 3$
	concat & 1×1 conv,768
Stage 5	$[3 \times 3 \text{ dsconv},224, \text{stride} = 1] \times 3$
	concat & 1×1 conv,1024

2.2.3. K-RPN

The faster-RCNN uses a region proposal network (RPN) to effectively generate high-quality region proposals, where each of them corresponds to the probability and position information of the cargo ship. Meanwhile, the network shares the feature maps with the feature extraction network to shorten the computing time.

In order to locate objects of different sizes effectively, the anchors of different scales are used in RPN, but the original anchors are designed manually without the prior information of cargo ship sizes in remote sensing images. Therefore, in this paper, the K-Mean++

algorithm [31] is used to cluster the bounding boxes of cargo ships in the dataset to obtain anchors suitable for cargo ships. The structure of K-RPN is shown in Figure 5.

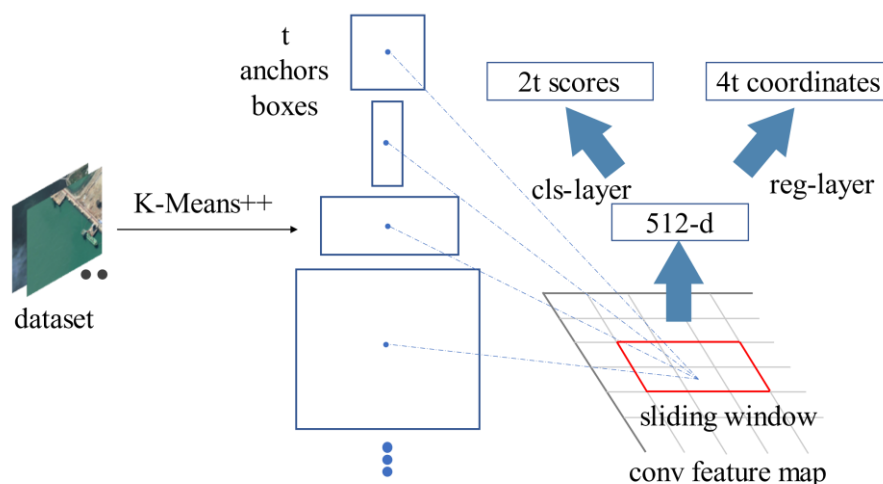


Figure 5. The structure of K-RPN in proposed model.

In the process of clustering, the Euclidean distance clustering method will lead to a greater error for the larger bounding boxes than for the smaller bounding boxes, resulting in a greater error in the intersection over union (IOU) between the anchors and the bounding boxes. In order to make the IOU of the anchors and the bounding boxes larger and obtain better anchors, the IOU of the bounding boxes and the cluster center bounding boxes is used as the distance function of the K-Mean++ algorithm to complete clustering:

$$d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid}) \quad (1)$$

where centroid represents the bounding boxes of the cluster centers; box represents the bounding boxes of the cargo ships; IOU represents the intersection over union of the cargo ship bounding boxes of cargo ships and the cluster center bounding boxes of the cluster centers; and d represents the distance between the bounding boxes of cluster centers and the bounding boxes of cargo ships.

2.2.4. Directional Discrimination

Since remote sensing images are taken above the cargo ships, the images contain not only the category and location information of the cargo ships but also the directional information of the cargo ships. Mastering the course information of cargo ships is important for cargo ships to sail along the expected route, which not only can ensure the safe navigation of the cargo ships but also can help to reduce the cost of navigation. In this paper, the problem of directional discrimination of cargo ships is transformed into a classification problem. The direction of the cargo ship is divided into four directions: east, south, west, and north (Figure 6). The object detection model is used to classify the directional information into category information to learn, and the model outputs the category of cargo ships and directional information at the same time. In this model, the directional information of cargo ships can be obtained without adding extra parameters.

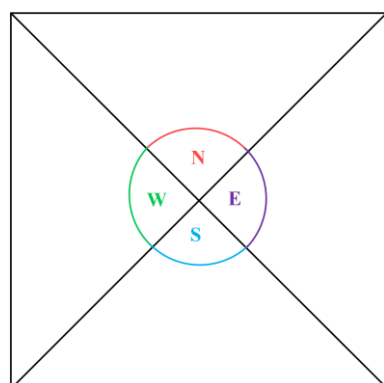


Figure 6. The division of direction. N, E, S, and W, respectively, denote the north, east, south, and west directions of the cargo ship.

3. Experiments

3.1. Experimental Dataset

- (1) **Data collection.** As far as we know, there is no publicly available remote sensing image cargo ship dataset with category and directional information. To detect cargo ships and discriminate directions of cargo ships, it is necessary to collect corresponding remote sensing images. Remote sensing images selected in this experiment are from Google Earth. The resolutions of images are 16, 17, and 18 levels. The bands of images are red, green, and blue. The data content covers different backgrounds and various positional relationships, which can meet the need for practical tasks. Due to the limitation of computer memory capacity, the collected images are cropped into 800×800 pixels, and the images containing three types of cargo ships are filtered out—bulk carrier, container, and tanker—ensuring that each image contains at least one cargo ship. The examples are shown in the Figure 7.
- (2) **Data annotation.** In order to establish a cargo ship dataset of remote sensing images, it is necessary to annotate the collected remote sensing images for the training and testing of the models. In this paper, labeling software (Labelimg) [32] is used on the cargo ships in remote sensing images. Labeled objects include bulk carrier, container, and tanker. According to the directions of cargo ships, they can be divided into four categories, east, south, west, and north, and the angle range of each category is 90 degrees. After the labeling is completed, a corresponding labeling file will be formed, which mainly records the location, category, and direction of the cargo ship. The obtained annotated dataset is uniformly processed into the format of VOC2007 [33] to provide a standard dataset for the training of models.
- (3) **Data augmentation and splitting.** In the process of model training, the larger and more comprehensive the dataset, the stronger the model recognition ability. Therefore, in this paper, the collected remote sensing images are rotated clockwise by 90 degrees, 180 degrees, and 270 degrees, as well as flipped horizontally and vertically, to expand the data. The number of the dataset is expanded by six times, and its directional label is adjusted accordingly. Finally, 15,654 remote sensing images are obtained. Table 2 shows the details of various cargo ships in the dataset. The dataset is randomly divided into training set, validation set, and test set according to the ratio of 8:1:1.

Table 2. The number of cargo ships used per category after applying augmentation techniques.

Category	East	South	West	North	Total
Bulk carrier	1505	1423	1505	1423	5856
Container	2207	2299	2207	2299	9012
Tanker	2013	1989	2013	1989	8004
Total	5725	5711	5725	5711	22,872



Figure 7. Examples of cargo ships in remote sensing images. (a) bulk carrier; (b) container; (c) tanker.

3.2. The Anchors Clustering

To obtain anchors suitable for cargo ships, this paper introduces the K-Mean++ algorithm into the RPN to propose the K-RPN, and the K-Mean++ algorithm is used to automatically generate the appropriate anchors instead of the manual method. If the anchors are not selected properly, the final detection results will be affected. Figure 8 shows the length and width distribution of cargo ships in the dataset and the clustering results of K-Mean++. Table 3 shows the comparison between original anchors and K-Mean++ anchors.

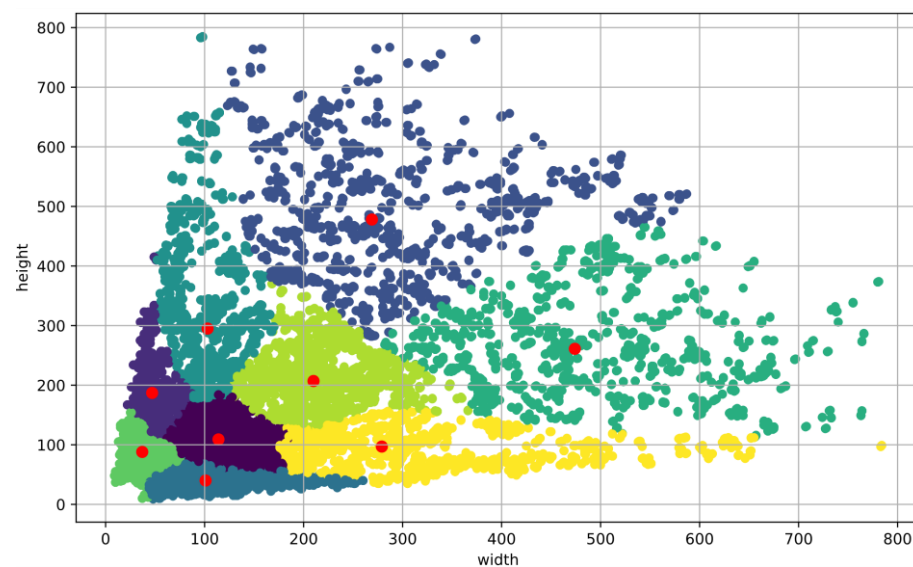


Figure 8. The distribution of width and height of annotated bounding boxes in the cargo ship dataset. The red dots represent the clustering results.

Table 3. Comparison between original anchors and K-Mean++ anchors.

Original Anchors	Width	91	128	181	181	256	362	362	512	724
	Height	181	128	91	362	256	181	724	512	362
K-Mean++ Anchors	Width	37	128	101	103	114	210	259	279	474
	Height	88	128	40	295	109	207	478	97	261

3.3. Implementation Details

The experiments are conducted on a Dell T3640 workstation. The operating system is Ubuntu 16.04 LTS. The model is written in Python and supported by Torch on the backend.

The model input size is set to 800×800 . Stochastic gradient descent is used in training to minimize the loss function. The initial learning rate is set to 0.025, and the learning rate is reduced by 0.1 times in 30,000 and 40,000 iterations. The batch size is two. The value of momentum is set to 0.9. The value of weight decay is set to 0.0005. The model is trained for 50,000 iterations. Figure 9 shows the change of loss during training.

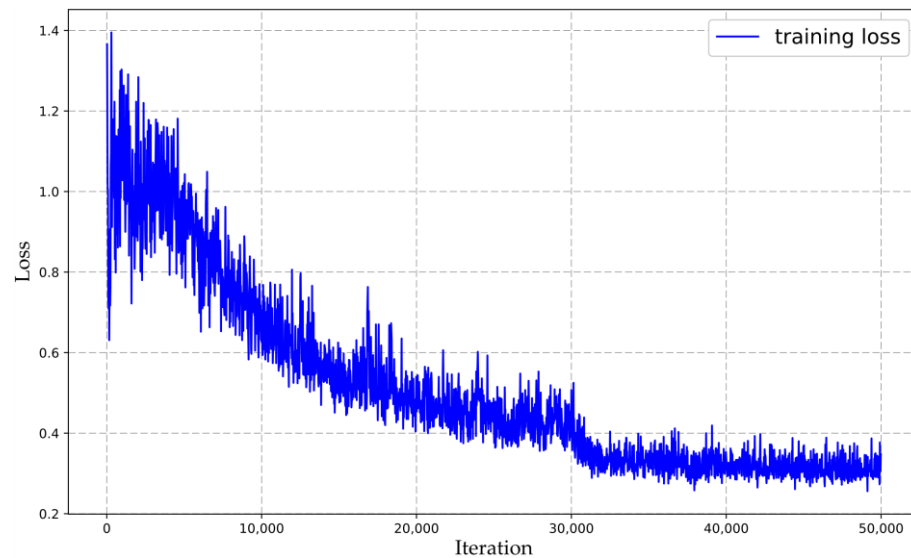


Figure 9. The curve of loss of training.

3.4. Evaluation Metrics

This paper used four evaluation metrics: precision (P), recall (R), average precision (AP), and mean average precision (mAP) to evaluate the performance of the model. Precision indicates the percentage of true positives of cargo ships in the sum of true positives of cargo ships and false positives of cargo ships. Recall indicates the percentage of true positives of cargo ships to the ground truths of cargo ships. AP is the area under the curve of precision-recall, which is the popular evaluation metric of object detection, and is often used to measure the advantages and disadvantages of object detection models. mAP indicates the mean of AP across all classes.

Precision:

$$P = \frac{Tp}{Tp + Fp} \quad (2)$$

Recall:

$$R = \frac{Tp}{Tp + Fn} \quad (3)$$

AP :

$$AP = \int_0^1 P(R) dR \quad (4)$$

mAP :

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (5)$$

where Tp represents the number of cargo ships correctly identified in the detection results, Fp represents the number of cargo ships falsely identified in the detection results, Fn represents the number of cargo ships that have been missed, and n represents the number of types of cargo ships.

3.5. Experimental Results

3.5.1. Performance of Different Models

To quantitatively analyze the effectiveness of the proposed model, the self-built dataset is used to train and test the model, and the evaluation indexes P , R , and AP are used to evaluate the test results of the model. Table 4 lists the test results of our model on the test set.

Table 4. The number of cargo ships for test set results and the evaluation of our model.

Class	Number				P (%)	R (%)	AP (%)
	G_T	T_p	F_p	F_n			
Bulk_Carrier_North	159	150	14	9	91.46	94.34	93.46
Bulk_Carrier_East	135	130	11	5	92.20	96.30	94.94
Bulk_Carrier_South	122	117	26	5	81.82	95.90	95.20
Bulk_Carrier_West	142	133	6	9	95.68	93.66	93.39
Container_North	192	181	27	11	87.02	94.27	93.17
Container_East	203	190	31	13	85.97	93.60	93.00
Container_South	226	207	34	19	85.89	91.59	90.10
Container_West	222	207	24	15	89.61	93.24	92.26
Tanker_North	207	185	41	22	81.86	89.37	88.10
Tanker_East	192	176	19	16	90.26	91.67	91.26
Tanker_South	205	184	26	21	87.62	89.76	88.77
Tanker_West	202	183	31	19	85.51	90.59	89.87

Table 5 gives statistics on model size, training duration, mAP , and single image prediction duration of different models. As can be seen from the experimental results, in terms of model size, our model has fewer parameters (model size of 110 MB) and a shorter training time (79.5 min), so it is more suitable to be deployed on the spaceborne platform. In terms of detection accuracy, our model can achieve a higher detection accuracy (mAP of 91.96%) similar to the two-stage object detection algorithm faster-RCNN (mAP of 94.41%). In terms of detection speed, our model maintains a fast detection speed (prediction time of 46ms per image) and meets the requirements of real-time detection.

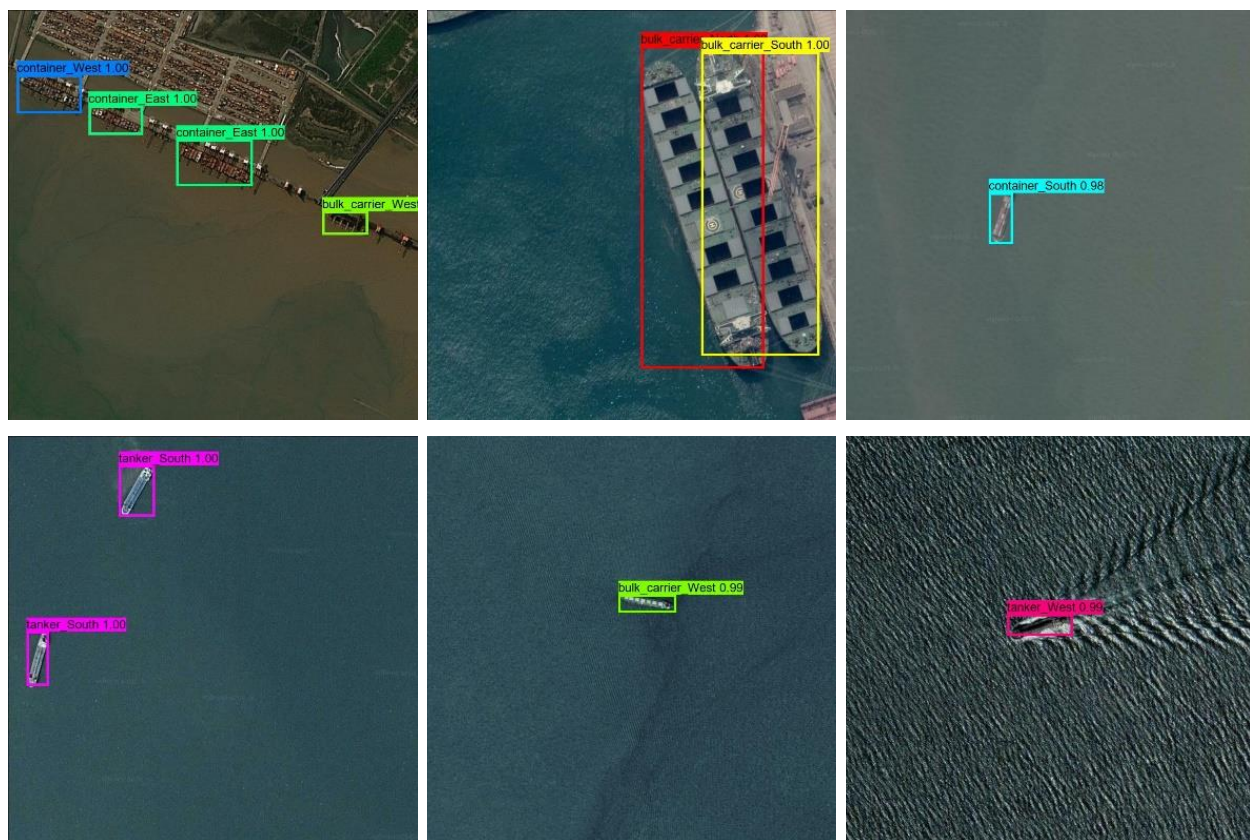
Table 5. Comparison of test performance between different models.

Model	Feature Extraction Network	Region Proposal Network	Weight Size/MB	Training Time/min	mAP	Input Size	Prediction Time/ms
Faster-RCNN	ResNet50	RPN	319.5	100.4	94.41%	800×800	56
YOLOv3	DarkNet53	None	235.0	758.0	83.98%	800×800	41
Our	OSASNet	K-RPN	110.0	79.5	91.96%	800×800	46

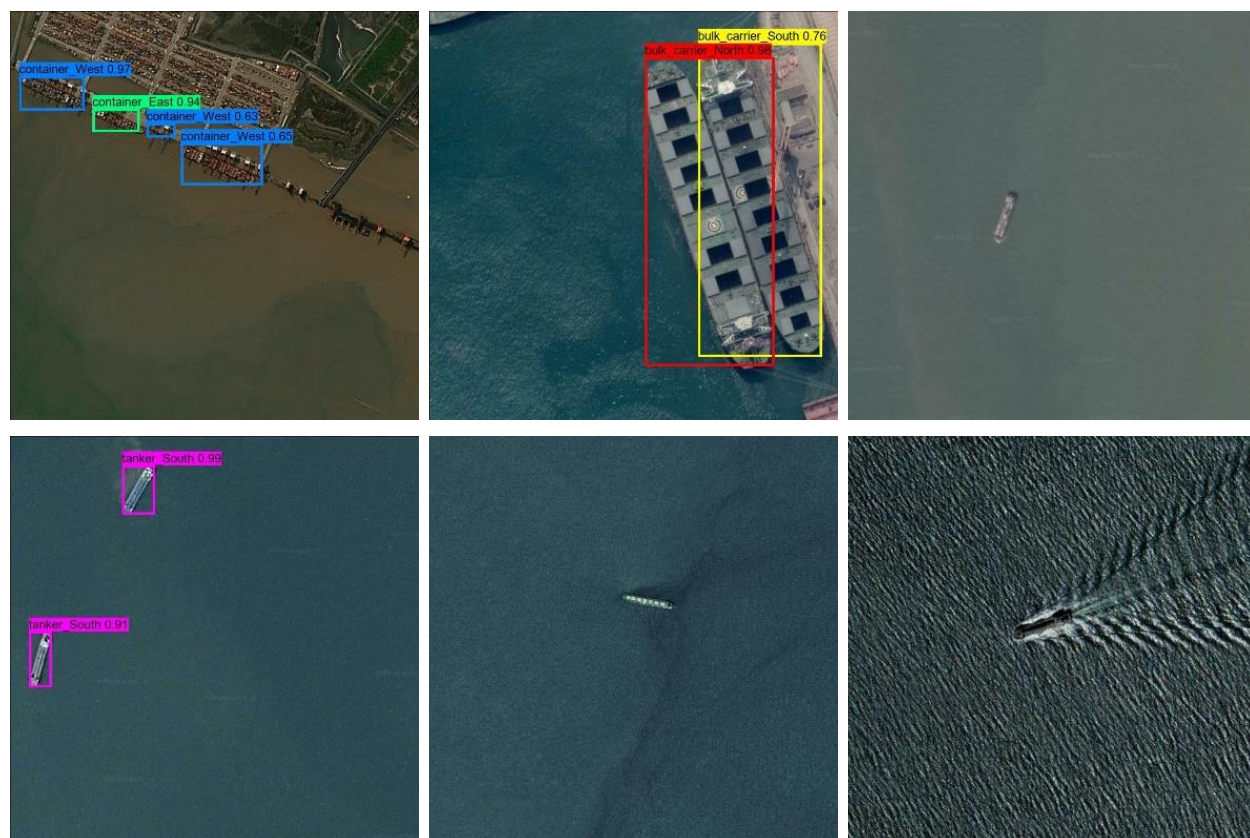
Figure 10 shows the detection results of different models on test sets. It can be seen from the figure that our model and faster-RCNN can complete the classification, positioning, and directional discrimination of cargo ships in remote sensing images well, while the YOLOv3 model has different degrees of missed detection.

3.5.2. Performance of Other Remote Sensing Images

Given that Google Earth images are the integration of multiple satellite images and aerial images, this paper conducts tests on high-resolution remote sensing images of multiple satellites, including Skysat 1.0 m/pixel resolution images (Figure 11), Deimos-2 0.5~2 m/pixel resolution images (Figure 12), and QuickBird 0.5~2 m/pixel resolution images (Figure 13), to verify the performance of our model on remote sensing images of a single source.

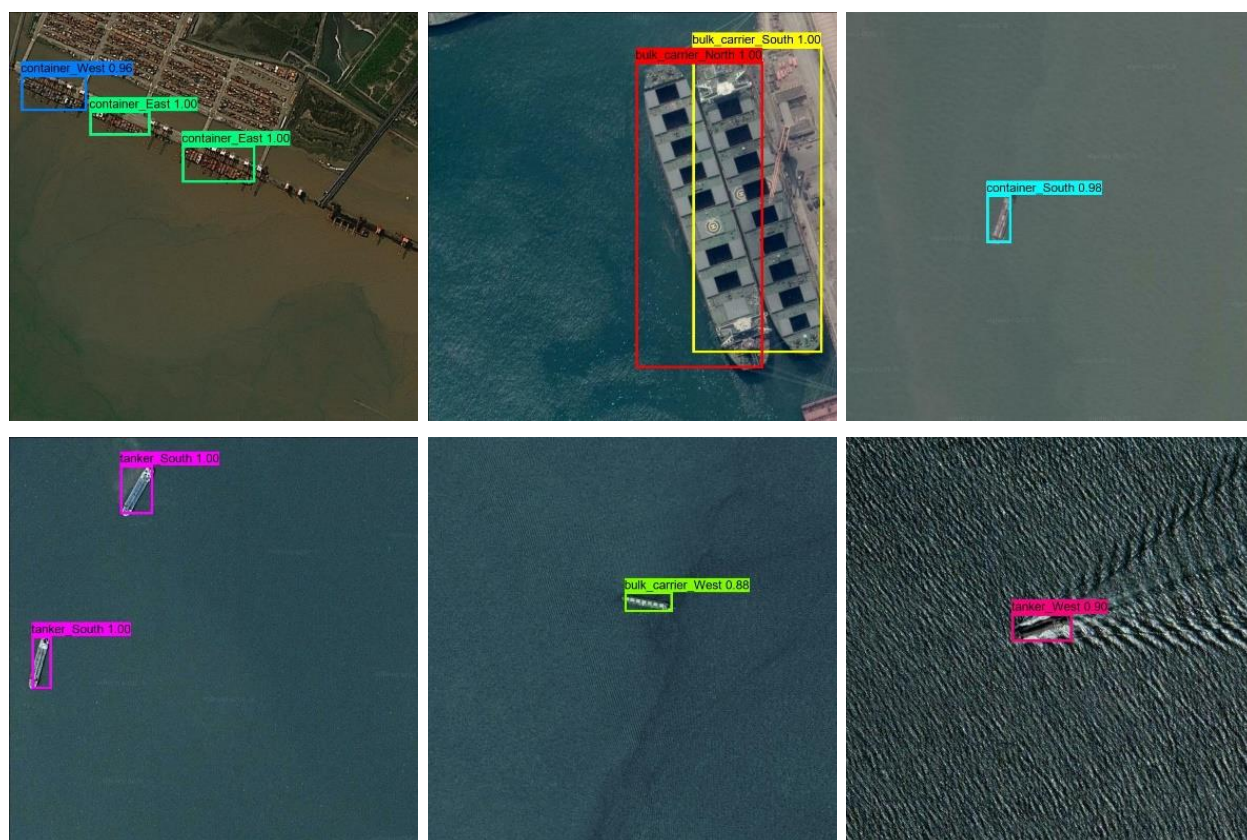


(a) Faster-RCNN



(b) YOLOv3

Figure 10. Cont.



(c) Our Model

Figure 10. The experiment result for cargo ships. (a–c) are the detection results of cargo ships by faster-RCNN, YOLOv3, and our model, respectively.

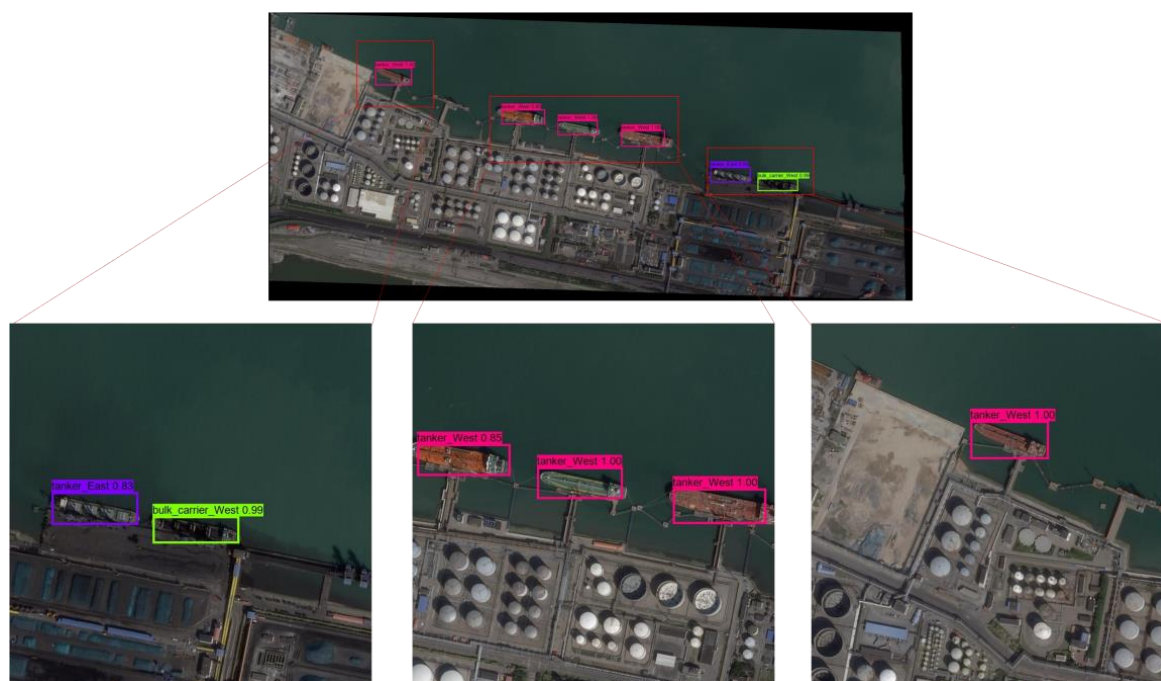


Figure 11. The Skysat image of Tianjin Port.



Figure 12. The Deimos-2 image of Dalian Port.

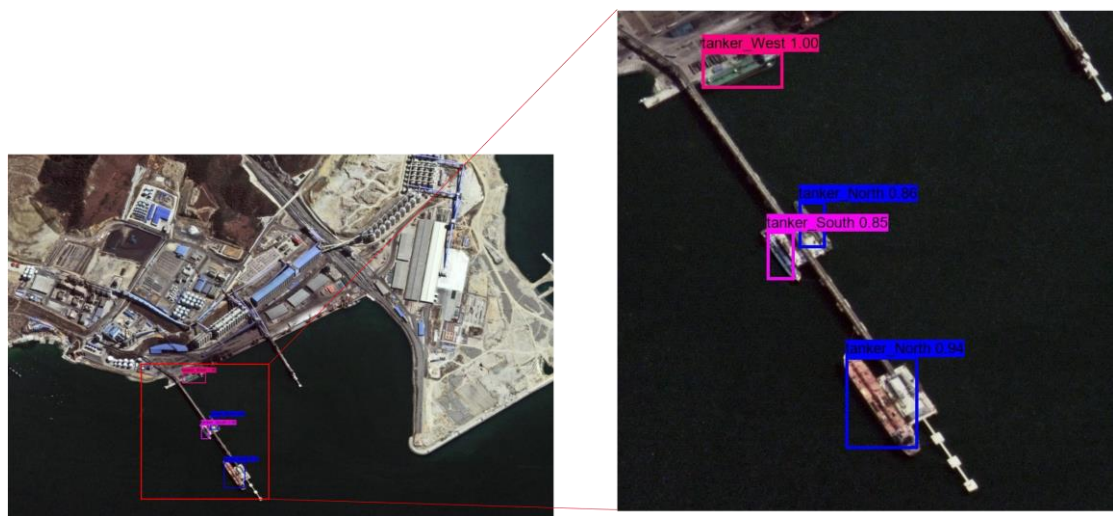


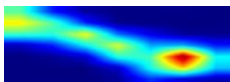
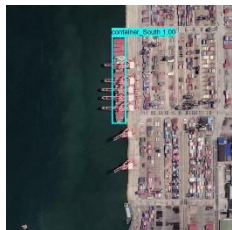
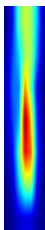
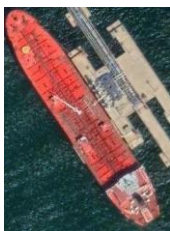
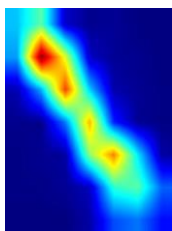
Figure 13. The QuickBird image of Pescara Port.

4. Discussion

Using remote sensing technology to obtain the remote sensing images of the cargo ships, combined with the deep learning object detection algorithms to obtain the classification, position, and directional information of the cargo ship in real time, is of great significance to cargo ship monitoring. In this paper, to further excavate the directional information of cargo ships in remote sensing images, by using the unique perspective of remote sensing images, the directional recognition problem of cargo ships is transformed into a classification problem, so that the model can complete the directional discrimination while completing the detection task. By extracting the heatmap of the feature map corresponding to the detection box, it is found that the model pays more attention to the head of the cargo ship, which is consistent with the actual judgment of the direction of the cargo ship (Table 6). At the same time, real-time processing of remote sensing images on the space-borne platform is the future trend. However, due to the limited computing and storage resources of the spaceborne platform, the complex and huge target detection

model has not been widely used. Therefore, we propose a novel cargo ship detection and directional discrimination method for remote sensing images based on a lightweight network, which has important practical significance.

Table 6. Detection results and corresponding heatmaps.

Detection Result	Region Proposal	Heatmap	Class Prediction	Direction Prediction
			Bulk carrier	East
			Container	South
			Tanker	North

5. Conclusions

This paper proposes a novel cargo ship detection and directional discrimination method for remote sensing images based on a lightweight network, which can efficiently and accurately classify, locate, and discriminate the direction of cargo ships in remote sensing images. Aiming at the problem in which the complex and large object detection network is not conducive to real-time autonomous operation on the spaceborne platform, we use one-shot feature aggregation modules and depthwise separable convolutions to design an efficient and lightweight feature extraction network, namely the one-shot aggregation and depthwise separable network (OSADSNet). At the same time, the K-Mean++ algorithm is introduced into the RPN to form the K-RPN to generate a more suitable region proposal for cargo ship detection. The current detection results of the cargo ship in the remote sensing image are output as the category and location of the cargo ship, but the directional information of the cargo ship is lacking. To solve this problem, we transform the directional recognition problem of the cargo ship in the remote sensing image into a classification problem without introducing additional parameters, which completes the directional estimation while completing the detection task. Finally, a comparative experiment is conducted based on the self-built dataset. Experimental results show that our model can meet the requirements of spaceborne platform cargo ship detection and directional discrimination in terms of model size (110 MB), detection accuracy (mAP of 91.96%), and detection speed (prediction time of 46 ms per image). Considering that Google Earth images are the integration of multiple satellite images and aerial images, to test the effectiveness of the model for remote sensing images that are from a single spaceborne

platform, remote sensing images from different satellites are selected for testing, and the experimental results show the effectiveness of the model.

This study has some limitations. The proposed method can complete the classification, positioning, and directional discrimination of cargo ships in remote sensing images well. The direction of the cargo ship is divided into four directions: east, south, west, and north. Remote sensing images of cargo ship contain rich information, and this paper only mined part of this information. Therefore, the direction of cargo ship can be further refined to obtain more detailed directional information. In the future, we will consider refining the direction of the cargo ship into north, northeast, east, southeast, south, southwest, west, northwest, or even a degree.

Author Contributions: Conceptualization, P.W., J.L., and Y.Z.; methodology, P.W. and J.L.; software, P.W.; validation, P.W., J.L., and Y.Z.; data curation, P.W., J.L., Y.Z., Z.Z., Z.C., and N.S.; writing—original draft preparation, P.W. and J.L.; writing—review and editing, P.W., J.L., and Y.Z.; visualization, P.W.; supervision, P.W. and J.L.; project administration, P.W., J.L., Y.Z., Z.Z., Z.C., and N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by High-Level Talent Research of Zhengzhou University, grant number 32310216.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data could be available on request from the corresponding author (liujianzhong@zzu.edu.cn).

Acknowledgments: The authors would like to thank the anonymous reviewers and editors for their useful comments and suggestions, which were of great help in improving this paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

AP	Average precision
AIS	Automatic identification system
BN	Batch normalization
CNN	Convolutional neural network
F_n	False negatives
F_p	False positives
G_T	Ground truth
IOU	Intersection over union
mAP	Mean average precision
P	Precision
R	Recall
R-CNN	Region-based convolutional neural networks
ReLU	Rectified linear unit
ROI	Regions of interest
RPN	Region proposal network
SSD	Single shot multibox detector
T_p	True positives
YOLO	You only look once

References

1. Tang, J.X.; Deng, C.W.; Huang, G.B.; Zhao, B.J. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine. *IEEE Trans. Geosci. Remote Sensing* **2015**, *53*, 1174–1185. [\[CrossRef\]](#)
2. Cheng, G.; Han, J.W. A survey on object detection in optical remote sensing images. *ISPRS-J. Photogramm. Remote Sens.* **2016**, *117*, 11–28. [\[CrossRef\]](#)
3. Yang, X.; Sun, H.; Fu, K.; Yang, J.R.; Sun, X.; Yan, M.L.; Guo, Z. Automatic Ship Detection in Remote Sensing Images from Google Earth of Complex Scenes Based on Multiscale Rotation Dense Feature Pyramid Networks. *Remote Sens.* **2018**, *10*, 132. [\[CrossRef\]](#)

4. Liu, T.; Pang, B.; Zhang, L.; Yang, W.; Sun, X.Q. Sea Surface Object Detection Algorithm Based on YOLO v4 Fused with Reverse Depthwise Separable Convolution (RDSC) for USV. *J. Mar. Sci. Eng.* **2021**, *9*, 753. [[CrossRef](#)]
5. Li, X.G.; Li, Z.X.; Lv, S.S.; Cao, J.; Pan, M.; Ma, Q.; Yu, H.B. Ship detection of optical remote sensing image in multiple scenes. *Int. J. Remote Sens.* **2021**, *29*. [[CrossRef](#)]
6. Wang, Q.; Shen, F.Y.; Cheng, L.F.; Jiang, J.F.; He, G.H.; Sheng, W.G.; Jing, N.F.; Mao, Z.G. Ship detection based on fused features and rebuilt YOLOv3 networks in optical remote-sensing images. *Int. J. Remote Sens.* **2021**, *42*, 520–536. [[CrossRef](#)]
7. Yokoya, N.; Iwasaki, A. Object localization based on sparse representation for remote sensing imagery. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2293–2296.
8. Li, Z.M.; Yang, D.Q.; Chen, Z.Z. Multi-Layer Sparse Coding Based Ship Detection for Remote Sensing Images. In Proceedings of the 2015 IEEE International Conference on Information Reuse and Integration, San Francisco, CA, USA, 13–15 August 2015; pp. 122–125.
9. Zhou, H.T.; Zhuang, Y.; Chen, L.; Shi, H. Ship Detection in Optical Satellite Images Based on Sparse Representation. In *Signal and Information Processing, Networking and Computers*; Sun, S., Chen, N., Tian, T., Eds.; Springer: Singapore, 2018; Volume 473, pp. 164–171.
10. Bhagya, C.; Shyna, A. An Overview of Deep Learning Based Object Detection Techniques. In Proceedings of the 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Chennai, India, 25–26 April 2019.
11. Wu, X.W.; Sahoo, D.; Hoi, S.C.H. Recent advances in deep learning for object detection. *Neurocomputing* **2020**, *396*, 39–64. [[CrossRef](#)]
12. Kwak, N.-J.; Kim, D. Object detection technology trend and development direction using deep learning. *Int. J. Adv. Cult. Technol.* **2020**, *8*, 119–128.
13. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
14. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision, Las Condes, Chile, 11–18 December 2015; IEEE: New York, NY, USA, 2015; pp. 1440–1448.
15. Ren, S.Q.; He, K.M.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
16. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: New York, NY, USA, 2016; pp. 779–788.
17. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2016; IEEE: New York, NY, USA, 2017; pp. 6517–6525.
18. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
19. Bochkovskiy, A.; Chien-Yao, W.; Liao, H.Y.M. YOLOv4: Optimal speed and accuracy of object detection. *arXiv* **2020**, *17*, 17.
20. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In *Computer Vision—ECCV 2016, Pt I*; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer International Publishing Ag: Cham, Switzerland, 2016; Volume 9905, pp. 21–37.
21. Lin, H.N.; Shi, Z.W.; Zou, Z.X. Fully Convolutional Network With Task Partitioning for Inshore Ship Detection in Optical Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1665–1669. [[CrossRef](#)]
22. Sun, Y.J.; Lei, W.H.; Ren, X.D. Remote sensing image ship target detection method based on visual attention model. In *Lidar Imaging Detection and Target Recognition 2017*; Lv, D., Lv, Y., Bao, W., Eds.; Spie-Int Soc Optical Engineering: Bellingham, WA, USA, 2017; Volume 10605.
23. Wei, S.H.; Chen, H.M.; Zhu, X.J.; Zhang, H.S. Ship Detection in Remote Sensing Image based on Faster R-CNN with Dilated Convolution. In Proceedings of the 39th Chinese Control Conference, Shenyang, China, 27–29 July 2020; pp. 7148–7153.
24. Li, Q.P.; Mou, L.C.; Liu, Q.J.; Wang, Y.H.; Zhu, X.X. HSF-Net: Multiscale Deep Feature Embedding for Ship Detection in Optical Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sensing* **2018**, *56*, 7147–7161. [[CrossRef](#)]
25. Zhang, S.M.; Wu, R.Z.; Xu, K.Y.; Wang, J.M.; Sun, W.W. R-CNN-Based Ship Detection from High Resolution Remote Sensing Imagery. *Remote Sens.* **2019**, *11*, 631. [[CrossRef](#)]
26. Yun, W. Ship Detection Method for Remote Sensing Images via Adversary Strategy. In Proceedings of the 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 11–13 December 2019; p. 4. [[CrossRef](#)]
27. Lee, Y.W.; Hwang, J.W.; Lee, S.; Bae, Y.; Park, J. An Energy and GPU-Computation Efficient Backbone Network for Real-Time Object Detection. In Proceedings of the 2019 IEEE/Cvf Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019; pp. 752–760.
28. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2016; IEEE: New York, NY, USA, 2017; pp. 1800–1807.
29. Kaiming, H.; Xiangyu, Z.; Shaoqing, R.; Jian, S. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]

-
30. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2016; IEEE: New York, NY, USA, 2017; pp. 2261–2269.
 31. Olukanmi, P.O.; Nelwamondo, F.; Marwala, T. k-Means-Lite plus plus: The Combined Advantage of Sampling and Seeding. In Proceedings of the 2019 6th International Conference on Soft Computing & Machine Intelligence, Johannesburg, South Africa, 30 June 2019; pp. 223–227.
 32. Tzutalin, D. Labelimg. 2019. Available online: <https://github.com/tzutalin/labelImg> (accessed on 14 April 2020).
 33. Everingham, M.; Van Gool, L.; Williams, C.K.I.; Winn, J.; Zisserman, A. The Pascal Visual Object Classes (VOC) Challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [[CrossRef](#)]