





Letter

Ship Detection Using a Fully Convolutional Network with Compact Polarimetric SAR Images

Qiancong Fan ¹, Feng Chen ^{1,*} , Ming Cheng ¹, Shenlong Lou ¹, Rulin Xiao ², Biao Zhang ³ ,
Cheng Wang ¹  and Jonathan Li ^{1,4} 

¹ Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China; fanqc@stu.xmu.edu.cn (Q.F.); chm99@xmu.edu.cn (M.C.); loushenlong@stu.xmu.edu.cn (S.L.); cwang@xmu.edu.cn (C.W.); junli@uwaterloo.ca (J.L.)

² Center for Satellite Applications on Ecology and Environment, Ministry of Ecology and Environment, Beijing 100094, China; xiaorl@secmep.cn

³ School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China; zhangbiao@nuist.edu.cn

⁴ Department of Geography and Environmental Management and Department of Systems Design Engineering University of Waterloo, Waterloo, ON N2L 3G1, Canada

* Correspondence: chenfeng@xmu.edu.cn

Received: 30 May 2019; Accepted: 11 September 2019; Published: 18 September 2019



Abstract: Compact polarimetric synthetic aperture radar (CP SAR), as a new technique or observation system, has attracted much attention in recent years. Compared with quad-polarization SAR (QP SAR), CP SAR provides an observation with a wider swath, while, compared with linear dual-polarization SAR, retains more polarization information in observations. These characteristics make CP SAR a useful tool in marine environmental applications. Previous studies showed the potential of CP SAR images for ship detection. However, false alarms, caused by ocean clutter and the lack of detailed information about ships, largely hinder traditional methods from feature selection for ship discrimination. In this paper, a segmentation method designed specifically for ship detection from CP SAR images is proposed. The pixel-wise detection is based on a fully convolutional network (i.e., U-Net). In particular, three classes (ship, land, and sea) were considered in the classification scheme. To extract features, a series of down-samplings with several convolutions were employed. Then, to generate classifications, deep semantic and shallow high-resolution features were used in up-sampling. Experiments on several CP SAR images simulated from Gaofen-3 QP SAR images demonstrate the effectiveness of the proposed method. Compared with Faster RCNN (region-based convolutional neural network), which is considered a popular and effective deep learning network for object detection, the newly proposed method, with precision and recall greater than 90% and a F₁ score of 0.912, performs better at ship detection. Additionally, findings verify the advantages of the CP configuration compared with single polarization and linear dual-polarization.

Keywords: compact polarimetric SAR; ship detection; fully convolutional network; semantic segmentation; Gaofen-3

1. Introduction

Due to the all-day, all-weather capabilities of synthetic aperture radar (SAR) systems, SAR images play an important role in maritime monitoring. Among different polarimetric SAR modes, the quad-polarization SAR (QP SAR) mode with four channels (HH, HV, VH, and VV, where H and V denote horizontal and vertical polarization, respectively) captures the richest information of the observed area [1]. However, compared with the QP SAR mode, the linear dual-polarization SAR mode,

with lower system complexity, provides a wider swath width [1]. Likewise, the compact polarimetric SAR (CP SAR) mode, which has attracted much attention, provides a compromise between swath width and scattering information [2].

The first CP configuration, $\pi/4$ CP mode, was introduced in [3] which transmits a linear polarization microwave directed at 45° and receives signals in both H and V polarizations. Stacy and Preiss [4] proposed the dual circular polarization (DCP) configuration, which, compared with the $\pi/4$ CP configuration, has the advantage of rotational invariance. DCP transmits and receives circular polarizations (RR, RL or LR, LL, where R and L denote right and left circular polarization, respectively). Different from the DCP configuration, the circular transmit-linear receive (CTLR) CP configuration transmits circular polarization while receiving two linear polarizations (RH and RV or LH and LV) [5]. The CTLR CP configuration, with a unique self-calibrating property, has an architecture simpler than that of DCP. In addition, CTLR CP has been adopted in the following: Mini-RF aboard NASA's Lunar Reconnaissance Orbiter [6], Mini-SAR on India's lunar Chandrayaan-1 satellite [7], and Canadian RADARSAT Constellation Mission [8]. Therefore, further investigations in this study are based mainly on CTLR CP SAR images.

The potential for ship detection from CP SAR images has been explored, mainly through these three frameworks: Reconstruction, feature extraction, and distribution statistics.

Reconstruction framework. The first framework, initially proposed in [9], was based on QP SAR scattering covariance matrix reconstruction algorithms, in which an extrapolation between co-polarization and cross-polarization channels is designed for reconstruction. Based on an empirical model introduced for maritime applications [10], Souyris' method [9] was modified in [11] for application to complex (urban) areas. Additionally, a helical scattering mechanism was designed for reconstruction [12]. However, these methods, restricted by assumptions about the scattering mechanism and the relationships between different channels, result in limited reconstruction accuracy [13].

Feature extraction framework. The second framework is based on the features extracted from the CP SAR scattering matrices. The degree of polarization was used in [14] to discriminate a ship from the ocean. The dual-pol relative phase was used in [15] for ship detection. To classify the candidate targets, three m - χ decomposition parameters were employed in [8]. By combining the decomposition parameters with a transform, a new feature was formed in [13]. Furthermore, to reduce the effect of ocean clutter, post-processing was applied in [16] to the extracted features.

Distribution statistics framework. The third framework is to directly use the original CP SAR data without feature extraction. The likelihood ratio test algorithm under the assumption of Gaussian statistics was used for ships and ocean scattering components in CP SAR data [17]. Similarly, using this strategy, ship detection and sea ice discrimination were implemented [18,19]. To detect ships and oil slicks, an extended Bragg scattering model (X-Bragg) based method was proposed in [20]. Recently, to characterize the statistics of the notch distance of sea clutter, a notch filter was modified to be suitable for CP SAR data [21]. After that, the CFAR (constant false alarm rate) threshold of ship detection was mathematically derived [21].

At the same time, to detect ships using traditional SAR configuration (e.g., single polarization, dual-polarization, and quad-polarization), mainly two algorithms are being explored.

Object-wise method. It is developed in the pattern recognition community via sliding windows or region proposals [22]. Before deep learning was applied widely in object detection, ship detection had depended on predefined features (SIFT and HOG) and traditional classifiers (e.g., SVM and AdaBoost) [22]. Currently, deep learning-based methods dominate object-wise detection. In particular, the following show great ability in SAR ship detection: Region-based convolutional neural network (Faster RCNN) [23], you only look once (YOLO) [22], and single shot multiBox detector (SSD) [24]. Due to the strong ability of deep learning methods to extract and classify deep features, using these methods yields accurate results in ship detection [22–24].

Pixel-wise target discrimination. Another algorithm category is based on pixel-wise target discrimination, in which the most popular approach is CFAR. For CFAR, a threshold is set to keep the

false alarm rate constant [25,26]. To reduce the impact of SAR ambiguities and sea clutter in complex sea conditions, a bilateral CFAR algorithm was proposed in [27]. On the average, a detectability rate of about 71% (about 1% higher than for standard CFAR) was observed. A pixel-based CFAR detector, Search for Unidentified Maritime Objects (SUMO), automatically detects ships over a wide range of image types and environmental conditions [28]. More recently, the generalized-likelihood ratio test (GLRT) method was proposed in [26] to detect ships in real or near-real time. Mostly, sea–land segmentation is necessary for the pixel-wise algorithm, which also suffers from false alarms caused by ocean clutter [23]. Meanwhile, the lack of detailed information about ships in SAR images results in difficulties for object-wise detection methods [23]. In addition, SAR polarimetry provides significant information for ship detection. Compared against the co-polarized channels (HH or VV), a substantial improvement is observed when using the cross-polarized channel (HV) [29]. Generally, since it had been shown that QP SAR provides the best ship detection performance [30], polarimetric properties of ships have been employed for ship detection [31,32].

Generation of ship labels through pixel-wise detection methods are considered a semantic segmentation problem. In recent years, deep learning has dominated the field of semantic segmentation. For example, a fully convolutional network [33] was used for land cover mapping based on remote sensing images [34]. An architecture with an encoder–decoder in SegNet (a fully convolutional network for semantic segmentation) was introduced in [35]. Based on the architecture, U-Net was proposed in for the segmentation of biomedical images [36]. Currently, U-Net is used in crop mapping with SAR data [37] and road extraction with optical images [38]. It was proved in [37] that small numbers of training data could drive U-Net and obtain high accuracy in applications.

For ship detection, U-Net architecture is incorporated with CP SAR in the newly proposed method. Since, in SAR images (8 m in nominal resolution), in-shore ships can be confused with the infrastructure of harbors, the focus of this study was on off-shore ship detection. The following three classes were considered in end-to-end ship detection: Ship, land, and sea. The rest of the paper is organized as follows: The architecture and method proposed are detailed in Section 2. To compare our method with traditional object-based detection methods using deep learning, experiments and an analysis of Faster RCNN (which is considered a robust object detector for different image types and complex backgrounds [39]) are presented in Section 3. Discussions of different polarization SAR mode images and two measures for modifying the network, in which issues on validation are also highlighted, are discussed in Section 4. Section 5 concludes the paper.

2. Materials and Methods

2.1. Architecture of U-Net

Figure 1 shows the architecture of U-Net, which is shaped like the letter “U”. The encoder–decoder framework is adapted to extract features and predict labels, respectively.

Encoder. The encoder follows the original U-Net proposed in [36]. Especially, for CP SAR images, the input layer is adapted to have two channels (RH and RV). In U-Net, each “floor” (within a dotted rectangle) consists of two 3×3 convolutions, where each convolution is followed by a batch normalization (batch-norm) and a rectified linear unit (ReLU) operation. For down-sampling (designated by red arrows), each “floor” in the encoder is followed by a 2×2 max pooling operation with the stride of two. The maximum in the 2×2 area is retained during max pooling. After down-sampling, the number of feature channels is doubled by convolutions. As shown in Table 1, there are fourteen layers in the encoder, of which, ten layers are convolutional layers, and four layers are max pooling layers.

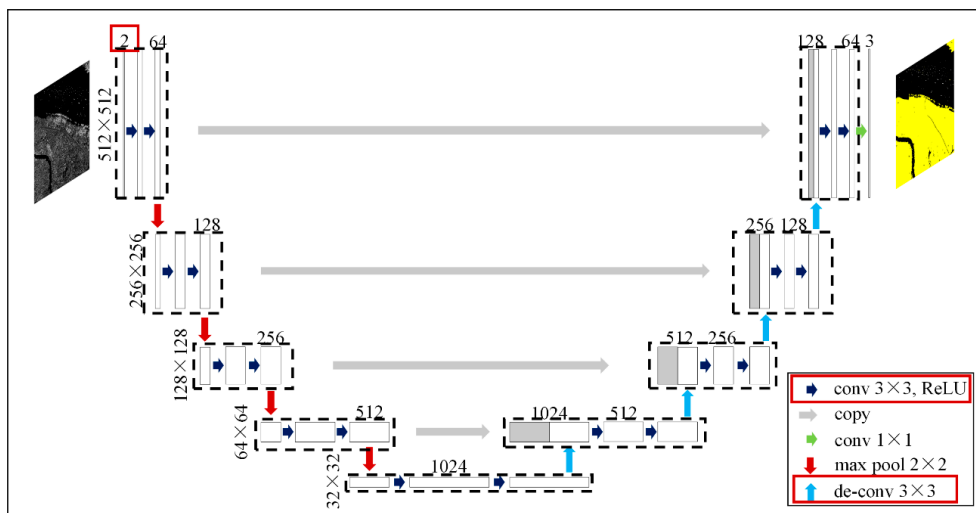


Figure 1. Architecture of the U-Net (example for 512×512 as the size of input data). Each white box corresponds to a multi-channel feature map. The number of channels is denoted at the top of the box. The size of feature maps for each layer is denoted on the left of the box. The gray boxes concatenated with white boxes are the copies of the feature maps from the encoder.

Table 1. Fully convolutional network (U-Net) architecture set for ship detection.

No	Type ¹	Input ²	Filters	Size/Stride	Output ²
1	conv	512 × 512 × 2	64	3 × 3/1	512 × 512 × 64
2	conv	512 × 512 × 64	64	3 × 3/1	512 × 512 × 64
3	max	512 × 512 × 64		2 × 2/2	256 × 256 × 64
4	conv	256 × 256 × 64	128	3 × 3/1	256 × 256 × 128
5	conv	256 × 256 × 128	128	3 × 3/1	256 × 256 × 128
6	max	256 × 256 × 128		2 × 2/2	128 × 128 × 128
7	conv	128 × 128 × 128	256	3 × 3/1	128 × 128 × 256
8	conv	128 × 128 × 256	256	3 × 3/1	128 × 128 × 256
9	max	128 × 128 × 256		2 × 2/2	64 × 64 × 256
10	conv	64 × 64 × 256	512	3 × 3/1	64 × 64 × 512
11	conv	64 × 64 × 512	512	3 × 3/1	64 × 64 × 512
12	max	64 × 64 × 512		2 × 2/2	32 × 32 × 512
13	conv	32 × 32 × 512	1024	3 × 3/1	32 × 32 × 1024
14	conv	32 × 32 × 1024	1024	3 × 3/1	32 × 32 × 1024
15	de_conv	32 × 32 × 1024	512	3 × 3/2	64 × 64 × 512
16	concat	64 × 64 × 512			64 × 64 × 1024
17	conv	64 × 64 × 1024	512	3 × 3/1	64 × 64 × 512
18	conv	64 × 64 × 512	512	3 × 3/1	64 × 64 × 512
19	de_conv	64 × 64 × 512	256	3 × 3/2	128 × 128 × 256
20	concat	128 × 128 × 256			128 × 128 × 512
21	conv	128 × 128 × 512	256	3 × 3/1	128 × 128 × 256
22	conv	128 × 128 × 256	256	3 × 3/1	128 × 128 × 256
23	de_conv	128 × 128 × 256	128	3 × 3/2	256 × 256 × 128
24	concat	256 × 256 × 128			256 × 256 × 256
25	conv	256 × 256 × 256	128	3 × 3/1	256 × 256 × 128
26	conv	256 × 256 × 128	128	3 × 3/1	256 × 256 × 128
27	de_conv	256 × 256 × 128	64	3 × 3/2	512 × 512 × 64
28	concat	512 × 512 × 64			512 × 512 × 128
29	conv	512 × 512 × 128	64	3 × 3/1	512 × 512 × 64
30	conv	512 × 512 × 64	64	3 × 3/1	512 × 512 × 64
31	conv	512 × 512 × 64	3	1 × 1/1	512 × 512 × 3

¹ Conv denotes convolutional layer, max denotes the max pooling layer, de_conv is the deconvolutional layer, and concat is concatenation. ² The size of the input and output is presented as height × width × channels.

Decoder. In the decoder, the first step of each “floor” is up-sampling (blue arrows), which doubles the height and width of the feature maps. In Table 1, for example, the size of the input feature maps in the 15th layer is 32×32 , and the output, after up-sampling (designated as de_conv) is 64×64 . Then, a concatenation with the copies (gray arrows) of the feature maps from the encoder is used to combine the deep semantic and shallow high-resolution features. As seen in Table 1, after concatenation (designated as concat), the channels of the output feature maps are twice the channels of the input. This operation is important to retain the original CP SAR information after repeated down-samplings and up-samplings. At the final layer, a 1×1 convolution is used to predict the probabilities of pixels. Padding is necessary in convolutions to ensure the uniform size of the input and output layers. In Figure 1, within the red rectangles, are the differences between the proposed method and the original network, including the number of input layers, padding, and up-sampling through deconvolution. Lastly, a pixel-wise softmax (a function used to normalize the input into a probability distribution) over the final feature map is computed. In the proposed method, the computational result, combined with the cross-entropy loss function, becomes the energy function of the U-Net.

In addition, deconvolution was used for up-sampling in the proposed method and also used in feature visualization and image generation [37,40]. By using a traditional up-sampling method, details are lost; therefore, a better choice is to learn rescaling during training. Included are the following two main steps: (1) Insert zeros between the consecutive inputs according to the resolution requirements and (2) produce a higher resolution output by convolution.

2.2. Data

Launched, as an ocean surveillance satellite, from the Taiyuan space center on 10 August 2016, the Chinese Gaofen-3 satellite, equipped with a multi-polarized C-band SAR at meter-level resolution, can operate in twelve different working modes [41,42]. With a design life of eight years [43], Gaofen-3 has been in operation officially since January 2017 [42]. The Gaofen-3 SAR image, which meets the accuracy for ship detection [44,45], is competent in numerous applications, such as in monitoring the global ocean and land resources [46]. Used in the experiments were eight Gaofen-3 QP SAR images, provided by the China Centre for Resource Satellite Data and Applications (CRESDA, <http://www.cresda.com>). Specific information about the images is given in Table 2. These acquired images (all single-look without multi-look processing) cover different coastal areas where some large ports are located (Figure 2). As in the previous study [16], after being calibrated through the PIE software (<http://www.piesat.com.cn/>), the CTRLR mode CP SAR images were generated by Equation (1) as follows [5]:

$$\vec{K}_{CTRLR} = \begin{bmatrix} E_{RH} \\ E_{RV} \end{bmatrix} = (1/\sqrt{2}) \begin{bmatrix} S_{HH} - S_{HVj} \\ S_{HV} - S_{VVj} \end{bmatrix} \quad (1)$$

where, S_{pq} are the elements of the scattering matrix with p transmitting and q receiving polarization; while H denotes the horizontal polarization; V denotes the vertical polarization, and j denotes the imaginary unit. E_{RH} and E_{RV} are the elements of CP SAR scattering vectors. Under the assumption of scattering reciprocity, the cross-polarization components in a QP SAR system are equal, that is, $S_{VH} = S_{HV}$ [9]. The intensity images of E_{RH} and E_{RV} was employed in experiments for ship detection.

Table 2. The information about the Gaofen-3 images employed in the experiment.

Image Number	Region	Nominal Resolution (m)	Acquisition Mode	Acquired Time (yyyy-mm-dd)	Incidence Angle (NearRange/FarRange) (degree)	Pixel Spacing (Rng/Az) # (m)
1	Taizhou Area	8	Ascending	2016-12-26	33.6833/35.6152	2.2484/5.1995
2	Taizhou Area	8	Ascending	2016-12-26	33.6830/35.6150	2.2484/5.1994
3	Taizhou Area	8	Ascending	2016-12-26	33.6828/35.6147	2.2484/5.2000
4	Taizhou Area	8	Ascending	2016-12-26	33.6827/35.6144	2.2484/5.1998
5	Taizhou Area	8	Ascending	2016-12-26	33.6836/35.6156	2.2484/5.1997
6	Shanghai Area	8	Ascending	2016-12-31	28.3065/30.6847	2.2484/4.7303
7	Qingdao Area	8	Ascending	2017-10-12	36.7626/38.1709	2.2484/5.2981
8	Haikou Area	8	Descending	2017-09-27	36.7564/38.1533	2.2484/4.7219

Rng and Az indicate range and azimuth, respectively.

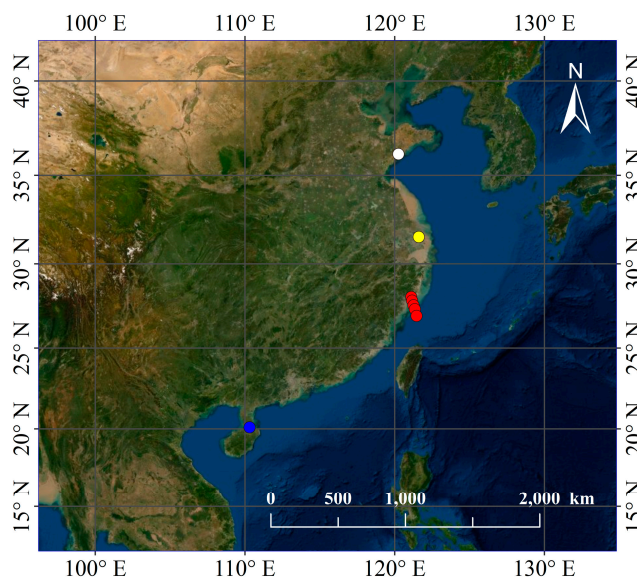


Figure 2. Spatial distribution of the data, the dots represent corresponding areas covered by the images we used. Respectively, the white dot represents the Qingdao area, the yellow dot denotes the Shanghai area, the red dots denote the Taizhou area, and the blue dot denotes the Haikou area.

2.3. Training

Images were manually labeled with image editing software. Figure 3 shows a general process for ship labeling, through the magnified image (top right) of a specific area (the rectangle area in the left CP SAR image) accompanied with automatic identification system (AIS) information (mid right) and labels (bottom right) correspondingly. Ideally, the length, course, and speed of ships from the AIS data (with latency) can be used to estimate the ships in corresponding SAR images. However, not all ships on the ocean carry AIS transponders [8]. AIS data, particularly the data for archived images, acquired synchronously with satellites is difficult to obtain. Therefore, as in previous studies [13,22–24,47–49], some images used in this study were labeled with experience and expert knowledge through visual interpretation. Since ship characteristics (including ship superstructure configuration, orientation of a ship with respect to radar beam, ship size, material from which the ship is made) affect the observation (e.g., intensity and geometry) in SAR images and the detectability of ships followed by [8]. Generally, ships appear as pixels with more brightness in intensity of the CP SAR image, compared against ocean background. Accordingly, a ship was determined firstly, mainly based on the characteristics (shape and texture) of the bright pixel cluster in the intensity image. The ship labeling was finally obtained by manually tracing the boundary between the target and background. However, spatial resolution of image (i.e., 8 m) might limit the credibility of labeled targets, especially for the small targets. Accordingly, in this investigation all ships were labeled through a collaborative way, with assumption that a target confirmed by more than two experts was credible for further investigations. As shown in Figures 3 and 4, land is yellow, ships are white, and the sea is black.

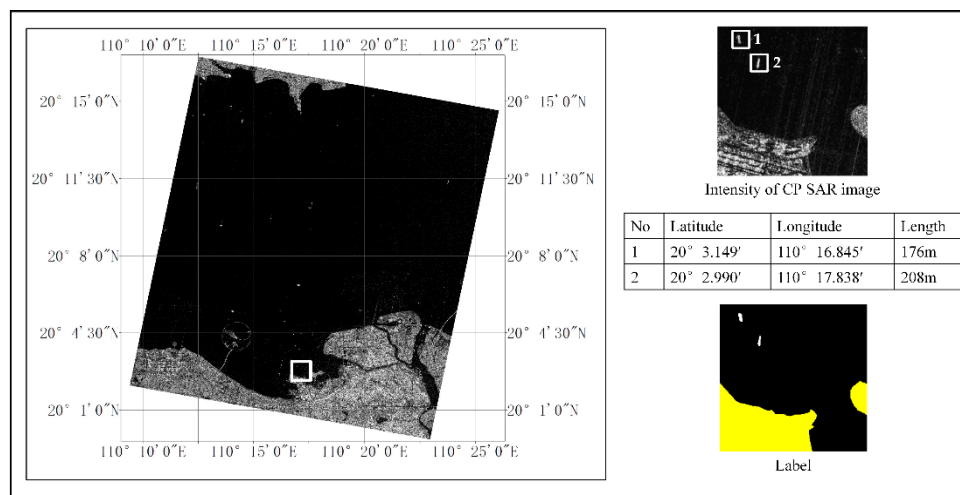


Figure 3. Illustrations of a compact polarimetric synthetic aperture radar (CP SAR) image (left), a sub-image (top right), corresponding automatic identification system (AIS) information (mid right), and label image (bottom right). The AIS information includes longitude, latitude, and ship length. Land is yellow, the ships are white, and the sea is black.

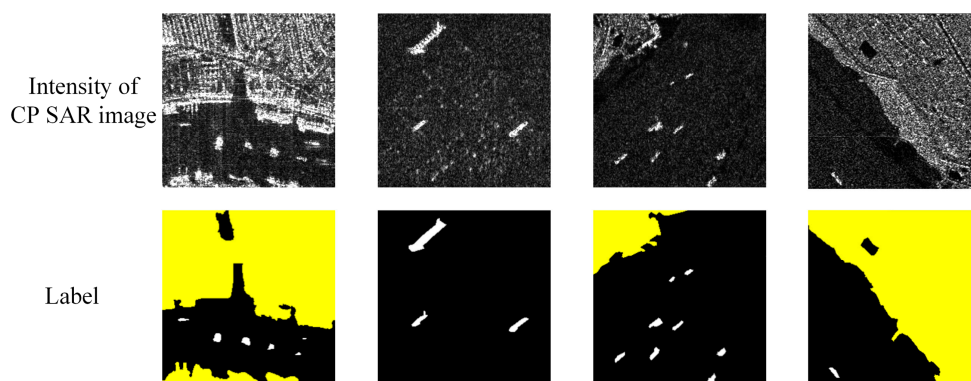


Figure 4. Illustrations of CP SAR images and corresponding label images. Land is yellow, ships are white, and the sea is black.

Each CP SAR image and the corresponding label image were divided into hundreds of 512×512 sub-images to generate training, validation, and test data randomly without overlapping in the proportion of 7:1:2. In general, the number of training samples, validation samples, and test samples was 1956, 280, and 560, respectively.

All deep learning experiments were performed based on the Tensorflow deep learning framework and were executed on the Ubuntu 16.04 operation system with a 12GB memory Titan X [50]. An Adam optimizer was used to optimize the network [51]. The learning rate was set at $1e-4$ with exponential decay. Batch size was set at 16. As shown in Figure 5, after about 70,000 iterations, the network converged. For Faster RCNN, against which a comparison was conducted to verify the correctness and effectiveness of the proposed method, an online open source project was used [52]. In particular, a VGG16 [53] model pre-trained on PASCAL VOC 2007 [54] was used to extract features. The same dataset labeled with LabelImg [55], a widely used annotation tool, was used to fine-tune the Faster RCNN network. The minimum enclosing rectangle was marked with the tool to generate labeled information. As shown in Figure 6, the labeled data used by the proposed method is semantic information. Accordingly, for the Faster RCNN, the labeled information is the bounding box information (provided in Extensible Markup Language (XML) data, which is an important input for Faster RCNN). The learning rate was set initially at $1e-6$. Batch size was set at 64. After about 30,000 iterations, the network converged.

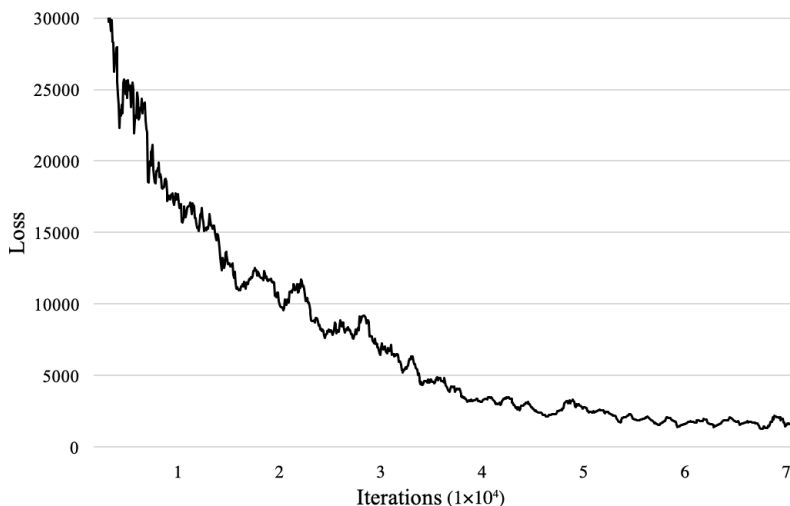


Figure 5. The train-loss curves of the U-Net training progress (with an average smooth).

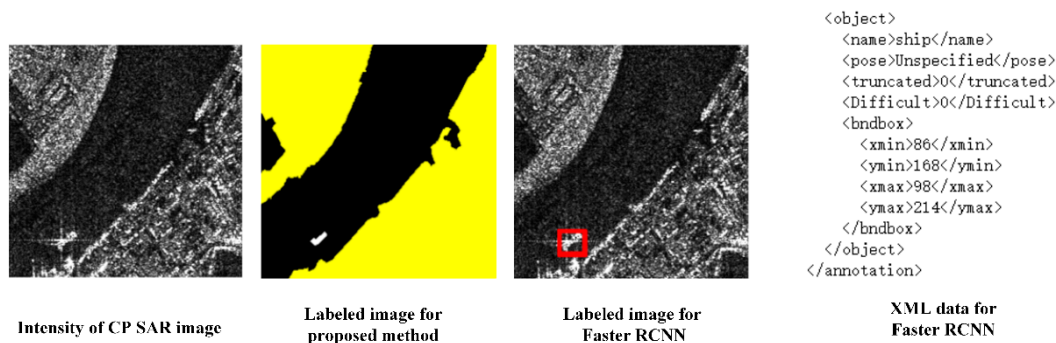


Figure 6. Comparison between the semantic segmentation label image and object detection label image with corresponding Extensible Markup Language (XML) data. The red rectangle denotes the bounding box of the ship. The bounding box information (xmin, ymin, xmax, and ymax represented the position of the bounding box) is provided in the XML file.

2.4. Validation

As was done in several studies in the field of semantic segmentation, in this study, the mean intersection-over-union (mIoU) was used for pixel-wise evaluation [33–36], according to the following equations:

$$mIoU = \frac{\sum_i^N IoU_i}{N} \tag{2}$$

$$IoU = \frac{area(predicted) \cap area(label)}{area(predicted) \cup area(label)} \tag{3}$$

where $area()$ denotes the computation of the pixel number. In brackets, “*predicted*” denotes the pixels predicted as a ship; “*label*” denotes the pixels of a ship as ground truth; i denotes the indices of the class; N denotes the total number of classes. The result is better when the $mIoU$ is closer to one.

For object level analysis, four evaluation indices (probability of false detection (P_f) [23,24], precision [22], recall [22], and F_1 score [23]), formulated as the following four equations, were used in this study:

$$Probability \text{ of false detection } (P_f) = \frac{FP}{TP + FP} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} = 1 - P_f \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F_1 \text{ score} = \frac{2Precision \times Recall}{Precision + Recall} \quad (7)$$

where TP means the number of positive targets detected correctly. FP means the number of negative targets detected as positive targets. FN means the number of undetected positive targets. Generally, a connected pixel set (in white) was regarded as one target even if only one pixel was detected. This means that each target was counted as one during the evaluation. For example, we counted TP as one when a ship was detected correctly. F_1 score is the harmonic average of precision and recall, where an F_1 score attains its best value at one (perfect precision and recall) and worst at zero.

3. Results

The proposed method can detect ships accurately over nearshore and far away from the shore (Figure 7). However, caused by cross sidelobes, a false alarm is generated. As seen in Table 3, compared with standard CFAR and Faster RCNN, the proposed method is effective for ship detection. For the implementation of standard CFAR, the cell averaging CFAR [25] was used, and the size of guard window was set to 40. As in a previous study [27], K-distribution with a PFA of $1e-6$ was used in the application of CFAR, while land mask information was provided by the land segmentation results obtained by the proposed method. The comparison was based on object level evaluation indices (Equations (4)–(7)). Observed in the results for the standard CFAR (Table 3) are the highest P_f and the lowest F_1 scores, with a large number of false alarms. Compared with Faster RCNN, the proposed method, with increases of 6.54% and 8.28% in precision and recall, respectively, shows improved ability to detect ships with CP SAR images. As a result, a greater F_1 score (approximately 0.91) is obtained with the proposed method. On average, the proposed method, with fewer false alarms and missed detections, performs better than standard CFAR and Faster RCNN in ship detection.

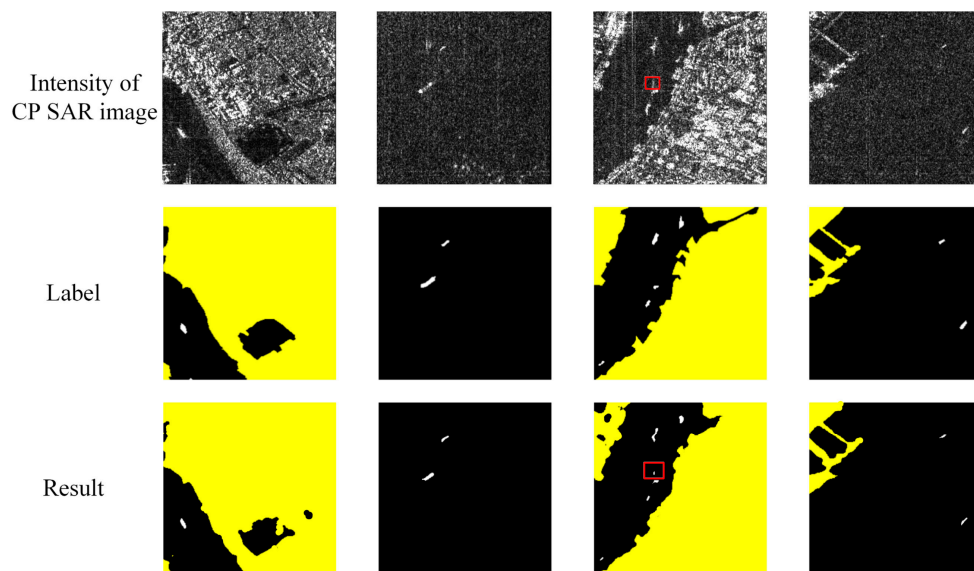


Figure 7. Illustrations of CP SAR images, corresponding label images, and detection results. The false alarm was marked by a red rectangle.

Table 3. Detection performance comparison between different methods (using CP SAR images).

Method	TP	FP	FN	P_f (%)	Precision (%)	Recall (%)	F_1 Score
Standard CFAR	609	307	74	33.52	66.48	89.17	0.762
Faster RCNN	565	95	124	14.39	85.61	82.00	0.838
U-Net	622	53	67	7.85	92.15	90.28	0.912

As shown in Figure 8, false alarms with the proposed method and Faster RCNN occur mainly in harbors. Moreover, false alarms caused by cross sidelobes hinder these two methods in detecting ships. False alarms are associated mainly with the characters on which each method concentrates. Accordingly, as shown in Figure 9, there is a difference in the false alarms of detected ships between the two methods. On one hand, the proposed method is sensitive to backscattering distributions; whereas, Faster RCNN focuses on shape and texture characteristics [23]. On the other hand, the missed targets of the proposed method all have weak intensity and small size. Likewise, some undetected targets, with a low confidence score assigned by Faster RCNN, are very near to brighter targets. Accordingly, it is difficult for Faster RCNN to detect ships in ship-intensive areas.

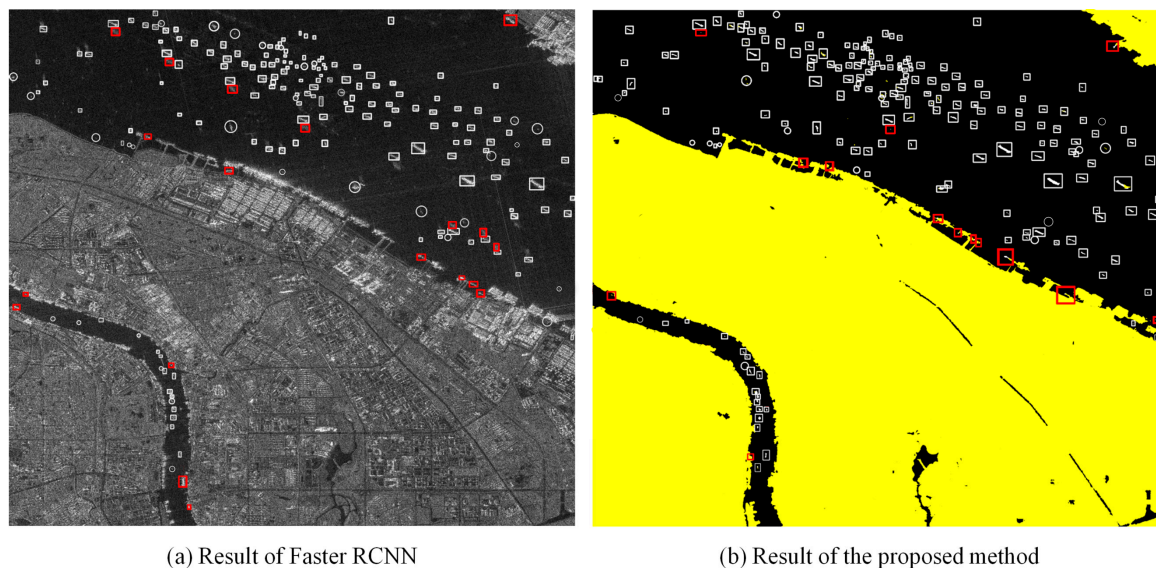


Figure 8. The comparison of the results between Faster region-based convolutional neural network (RCNN) and the proposed method. The white and red rectangles represent the detected targets and false alarms respectively, and the white circles represent the missed targets.

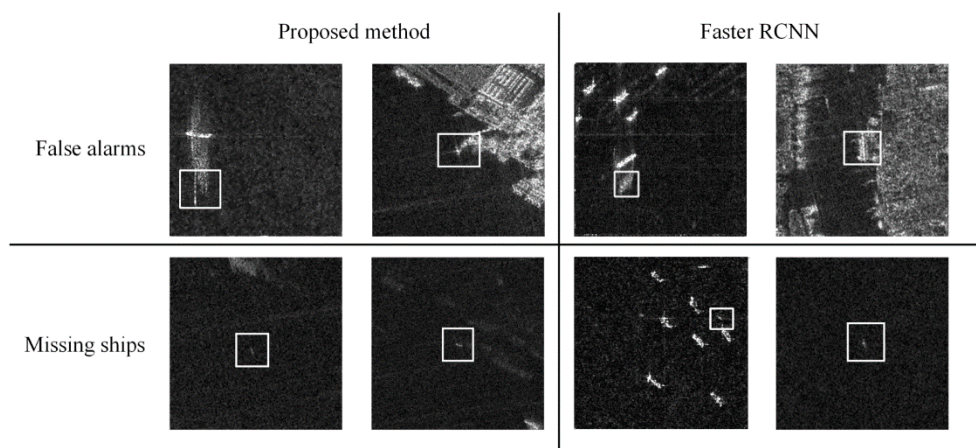


Figure 9. Demonstrations of typical false alarms and missing detection ships of the proposed method and Faster RCNN, which are marked by white boxes.

Figure 8 shows the results for a ROI sub-image. False alarms (marked by red rectangles) are mainly distributed over the harbor area and the area with heavy cross sidelobes. Generally, compared with the results from Faster RCNN, fewer false alarms, and missed targets were generated by our proposed method. As mentioned above, lack of detailed information in a SAR image is a challenge for ship detection by an object-wise detector. Meanwhile, similar to CFAR and GLRT, the proposed

method concentrates on neighbor information distribution, which, accordingly, shows a lower P_f for ship detection (see Table 3). However, in-shore ships are usually confused with the infrastructure of the harbors. In SAR images, even through manual interpretation, it is difficult to discriminate between in-shore ships and the infrastructure of the harbor.

4. Discussion

4.1. Comparison among Different Polarization Modes

As mentioned in Section 1, compared with the linear dual-polarization configuration, the CP SAR configuration potentially provides more information. In terms of ship detection ability, three combination modes of polarization channels (from one QP SAR data) were compared. The comparison of the detection results in ROI sub-images between different combination modes is shown in Figure 10. Due to the orientation of ship to radar beams, the ability of the single polarization mode (i.e., using HH) in ship detection and land segmentation is much lower compared with the other two modes. Furthermore, using the linear dual-polarization mode resulted in two false alarms and ten missed targets; whereas, using the CP mode resulted in two false alarms and six missed targets. As shown in Table 4, in terms of F_1 scores, the CP SAR configuration shows the best performance for ship detection. Due to a relatively small number of suspect targets detected from single polarization mode images, a lower P_f , and a much lower recall, are observed. According to the experimental results (Table 4 and Figure 10), compared with the single polarization and linear dual-polarization configurations, the CP SAR configuration is better at detecting ships.

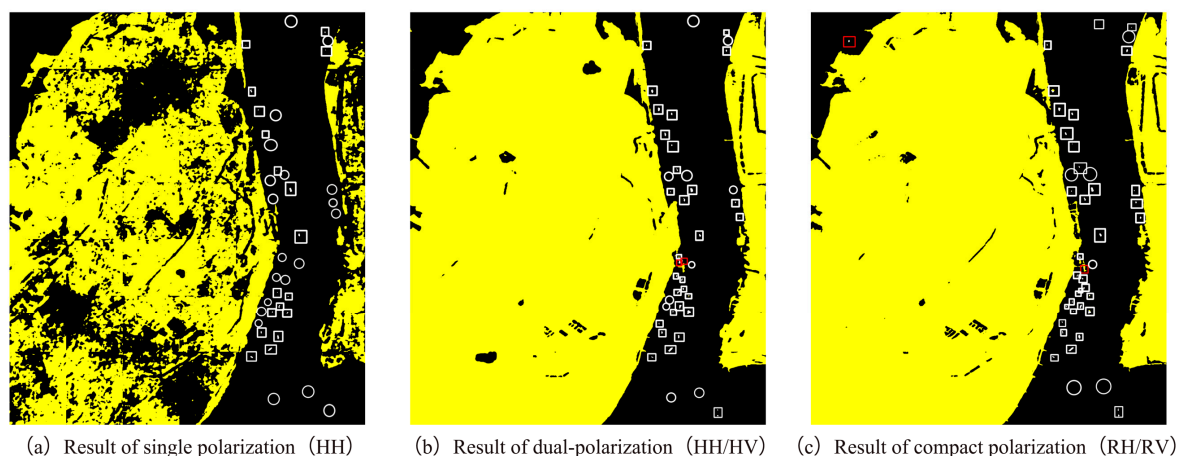


Figure 10. Illustration of results from different polarization modes. The white and red rectangles represent the detected targets and false alarms respectively, and the white circles represent the missed targets.

Table 4. Comparison between different combination modes of polarization channels in the proposed method.

Combination Modes	Channels	P_f (%)	Precision (%)	Recall (%)	F_1 score	mIoU
Single polarization	HH	6.86	93.14	49.87	0.650	0.601
Linear dual-polarization	HH+HV	8.22	91.88	81.37	0.863	0.734
Compact polarization	RH+RV	7.85	92.15	90.28	0.912	0.817

4.2. Comparison among Different Networks based on U-Net

Mainly two measures are used to test the possibility for improving the performance of the proposed method. One is to design deeper feature extraction using res-block (residual block of ResNet, which retains shallow information by a connection from the first to the last layers of a block layer). It has been proven that the res-block can take equal or higher feature expression ability [56]. Another is to expand

the receptive field by replacing the traditional convolutional kernel with the dilated convolutional kernel [57]. The main idea of dilated convolution is to insert “holes” (zeros) between pixels in convolutional kernels to increase the resolution of intermediate feature maps [57]. The detection results obtained by three methods are listed in Table 5. Findings show that the U-Net with dilated convolutional kernel is ineffective at detecting ships in CP SAR images. Minor improvement occurred in the detection ability of U-Net with res-block. However, it should be considered that, the deeper the network, the higher the computational complexity and the larger the resource consumption [56]. Accordingly, the proposed method, with ten layers, is more suitable for ship detection.

Table 5. Comparison between different methods based on the U-Net.

Methods	P_f (%)	Precision (%)	Recall (%)	F_1 Score	mIoU
Proposed (10 layers)	7.85	92.15	90.28	0.912	0.817
U-Net+Res-Block (32 layers)	7.63	92.37	91.36	0.919	0.826
U-Net+Dilated Convolutional Kernel	12.11	87.89	89.21	0.885	0.802

4.3. Validation

It has been suggested that a ground truth collection with plentiful samples is required for validation. In terms of AIS data, which are widely used to validate ship detection, the best choice is that the acquisition of SAR images and AIS data is synchronous. In most cases, AIS-SAR latency within several minutes is permissible [58–61]. The time difference allowed between the two observations depends on the ship density [61]. To project the targets from AIS data (with latency) to corresponding SAR images, mainly two methods are employed. One is to use dead reckoning to estimate the location of ships, based on the speed of a ship and course information [60,61]. Another is to use the correlation of the size of the ships between SAR images and AIS data [58]. Moreover, when it is difficult to obtain validation data, the only way to obtain labeled targets is based on experience and expert knowledge [13,22–24,47–49]. Although AIS information was used to verify the labeled dataset over several sub-areas, it is unfortunate that not all labeled ships in this study were supported by AIS data. It is a fact that not all ships on the ocean carry AIS transponders [8]. Likewise, especially for archived earth observations, AIS data simultaneously acquired during the overpass of a satellite is difficult to obtain. Therefore, as in previous studies [13,22–24,47–49], some images used in this study were labeled with experience and expert knowledge via collaboration. We assumed that a target confirmed by more than two experts was credible for further investigations. Nevertheless, to improve the ability of algorithms to discriminate confusing targets and detect small targets, a high-quality SAR ship dataset is required.

5. Conclusions

Compared with the quad-polarization SAR imaging mode, CP SAR provides a larger swath width. Likewise, compared with the linear dual-polarization SAR imaging mode, more information is available from an observed scene of CP SAR. We proposed a CP SAR ship detection method based on U-Net. Several CP SAR images, simulated from Gaofen-3 QP SAR images, were employed in the experiments. Experimental results verify the advantages of the CP configuration compared with single polarization and linear dual-polarization. Compared against the standard CFAR and Faster RCNN, our proposed method is more effective in detecting ships and especially more effective in reducing the impact of ocean clutter and SAR ambiguities. Additionally, in our proposed method, a deeper encoder contributes somewhat to the improvement in ship detection. Accordingly, worth considering is the trade-off between accuracy and resource consumption. Experimental results show that, while targets over harbors are confused with manmade infrastructures, ships with weak signals are likely not to be detected positively. It is necessary to have a standard dataset, in which accurate labels of small targets and artificial facilities (including harbor infrastructures and oil rigs) covering different ocean

backgrounds are contained. Since it will be helpful to train algorithms to learn differences among confusing targets, further investigation in this area will be considered.

Author Contributions: Q.F., F.C. and M.C. conceived and designed the experiments. Q.F. and S.L. performed the experiments and analyzed the data. R.X. and B.Z. contributed materials and data processing. All authors contributed to the writing—original draft preparation and editing.

Acknowledgments: This research was jointly supported by the National Key Research and Development Program of China under grants of 2016YFC1401001 and 2016YFC1401008, China Postdoctoral Science Foundation under grant of 2017M612124, and National Natural Science Foundation of China under grant of U1605254. The authors would also like to thank the anonymous reviewers for their very competent comments and helpful suggestions.

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

References

- Charbonneau, F.; Brisco, B.; Raney, R. Compact polarimetry overview and applications assessment. *Can. J. Remote Sens.* **2010**, *36*, 298–315. [[CrossRef](#)]
- Zhang, B.; Li, X.; Perrie, W.; Garcia-Pineda, O. Compact polarimetric synthetic aperture radar for marine oil platform and slick detection. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 1407–1423. [[CrossRef](#)]
- Souyris, J.C.; Mingot, S. Polarimetry based on one transmitting and two receiving polarizations: The $\pi/4$ mode. In Proceedings of the IEEE IGARSS, Toronto, ON, Canada, 24–28 June 2002; pp. 629–631.
- Stacy, N.; Preiss, M. Compact polarimetric analysis of X-band SAR data. In Proceedings of the EUSAR, Dresden, Germany, 16–18 May 2006.
- Raney, R. Comments on hybrid-polarity SAR architecture. In Proceedings of the IEEE IGARSS, Barcelona, Spain, 23–27 July 2007; pp. 2229–2231.
- Chin, G.; Brylow, S.; Foote, M.; Garvin, J.; Kasper, J.; Keller, J.; Litvak, M.; Mitrofanov, I.; Paige, D.; Raney, K.; et al. Lunar Reconnaissance Orbiter overview: The instrument suite and mission. *Space Sci. Rev.* **2007**, *129*, 391–419. [[CrossRef](#)]
- Goswami, J.; Annadurai, M. Chandrayaan-1: India's first planetary science mission to the moon. *Curr. Sci. India* **2009**, *96*, 486–490.
- Atteia Allah, G. On the Use of Hybrid Compact Polarimetric SAR for Ship Detection. Ph.D. Thesis, University of Calgary, Calgary, AB, Canada, 5 December 2014; p. 170. [[CrossRef](#)]
- Souyris, J.; Imbo, P.; Fjortoft, R. Compact polarimetry based on symmetry properties of geophysical media: The $\pi/4$ mode. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 634–646. [[CrossRef](#)]
- Collins, M.; Denbina, M.; Atteia, G. On the reconstruction of quad-pol SAR data from compact polarimetry data for ocean target detection. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 591–600. [[CrossRef](#)]
- Nord, M.; Ainsworth, T. Comparison of compact polarimetric synthetic aperture radar modes. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 174–188. [[CrossRef](#)]
- Yin, J.; Yang, J.; Zhang, X. On the ship detection performance with compact polarimetry. In Proceedings of the IEEE RADAR, Chengdu, China, 24–27 October 2011; pp. 675–680.
- Xu, L.; Zhang, H.; Wang, C. Compact polarimetric SAR ship detection with $m\text{-}\delta$ decomposition using visual attention model. *Remote Sens.* **2016**, *8*, 751. [[CrossRef](#)]
- Shirvany, R.; Chabert, M.; Tourneret, J. Ship and oils-pill detection using the degree of polarization in linear and hybrid/compact dual-pol SAR. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2012**, *5*, 885–892. [[CrossRef](#)]
- Li, H.; Perrie, W.; He, Y.; Lehner, S.; Brusch, S. Target detection on the ocean with the relative phase of compact polarimetry SAR. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3299–3305. [[CrossRef](#)]
- Fan, Q.; Chen, F.; Cheng, M.; Wang, C.; Li, J. A modified framework for ship detection from compact polarization SAR image. In Proceedings of the IEEE IGARSS, Valencia, Spain, 22–27 July 2018; pp. 3539–3542.
- Liu, C.; Vachon, P.W.; English, R.A.; Sandirasegaram, N. *Ship Detection Using RADARSAT-2 Fine Quad Mode and Simulated Compact Polarimetry Data*; Defence R&D Canada: Ottawa, ON, Canada, 2009; Tech. Rep. DRDC-O-TM-2009-285.
- Atteia, G.; Collins, M. On the use of compact polarimetric SAR for ship detection. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 1–9. [[CrossRef](#)]

19. Denbina, M.; Collins, M.J. Iceberg detection using compact polarimetric synthetic aperture radar. *Atmos. Ocean.* **2012**, *50*, 437–446. [[CrossRef](#)]
20. Yin, J.; Yang, J.; Zhou, Z.; Song, J. The extended Bragg scattering model-based method for ship and oil-spill observation using compact polarimetric SAR. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *8*, 3760–3772. [[CrossRef](#)]
21. Gao, G.; Gao, S.; He, J.; Li, G. Ship detection using compact polarimetric SAR based on the notch filter. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5380–5393. [[CrossRef](#)]
22. Chang, Y.L.; Anagaw, A.; Chang, L.; Wang, Y.C.; Hsiao, C.; Lee, W. Ship detection based on YOLOv2 for SAR imagery. *Remote Sens.* **2019**, *11*, 786. [[CrossRef](#)]
23. Kang, M.; Ji, K.; Leng, X.; Lin, Z. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection. *Remote Sens.* **2017**, *9*, 860. [[CrossRef](#)]
24. Wang, Y.; Wang, C.; Zhang, H. Combining a single shot multibox detector with transfer learning for ship detection using sentinel-1 SAR images. *Remote Sens. Lett.* **2018**, *9*, 780–788. [[CrossRef](#)]
25. Crisp, D.J. *The State-of-the-Art in Ship Detection in Synthetic Aperture Radar Imagery*; Defence Science and Technology Organisation Salisbury (Australia) Info Sciences Lab: Edinburgh, Australia, 2004.
26. Iervolino, P.; Guida, R. A novel ship detector based on the generalized-likelihood ratio test for SAR imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 3616–3630. [[CrossRef](#)]
27. Leng, X.; Ji, K.; Yang, K.; Zou, H. A bilateral CFAR algorithm for ship detection in SAR images. *IEEE Geosci. Sens. Lett.* **2015**, *12*, 1536–1540. [[CrossRef](#)]
28. Greidanus, H.; Alvarez, M.; Santamaria, C.; Thoorens, F.; Kourti, N.; Argentieri, P. The SUMO ship detector algorithm for satellite radar images. *Remote Sens.* **2017**, *9*, 246. [[CrossRef](#)]
29. Touzi, R. On the use of polarimetric SAR data for ship detection. In Proceedings of the IEEE IGARSS, Hamburg, Germany, 28 June–2 July 1999; pp. 812–814.
30. Liu, C.; Vachon, P.W.; Geling, G.W. Improved ship detection using polarimetric SAR data. In Proceedings of the IEEE IGARSS, Anchorage, AK, USA, 20–24 September 2004; pp. 1800–1803.
31. Nunziata, F.; Migliaccio, M.; Brown, C.E. Reflection symmetry for polarimetric observation of manmade metallic targets at sea. *IEEE J. Oceanic Eng.* **2012**, *37*, 384–394. [[CrossRef](#)]
32. Marino, A. A notch filter for ship detection with polarimetric SAR data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2013**, *6*, 1219–1232. [[CrossRef](#)]
33. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the CVPR, Boston, MA, USA, 7–13 June 2015; pp. 8–10.
34. Kampffmeyer, M.; Salberg, A.B.; Jenssen, R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In Proceedings of the CVPRW, Las Vegas, NV, USA, 27–30 June 2016; pp. 680–688.
35. Vijay, B.; Kendall, A.; Cipolla, R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv* **2015**, arXiv:1511.00561.
36. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the MICCAI, Munich, Germany, 5–9 October 2015; pp. 234–241.
37. Wei, S.; Zhang, H.; Wang, C.; Wang, Y.; Xu, L. Multi-temporal SAR data large-scale crop mapping based on U-Net model. *Remote Sens.* **2019**, *11*, 68. [[CrossRef](#)]
38. Zhang, Z.; Liu, Q.; Wang, Y. Road extraction by deep residual u-net. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 749–753. [[CrossRef](#)]
39. Huang, J.; Rathod, V.; Sun, C.; Zhu, M.; Korattikara, A.; Fathi, A.; Fischer, I.; Wojna, Z.; Song, Y.; Guadarrama, S.; et al. Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the CVPR, Hawaii, HI, USA, 21–26 July 2017; pp. 7310–7311.
40. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [[CrossRef](#)] [[PubMed](#)]
41. Sun, J.; Yu, W.; Deng, Y. The SAR payload design and performance for the GF-3 mission. *Sensors* **2017**, *17*, 2419. [[CrossRef](#)]
42. Kang, W.; Xiang, Y.; Wang, F.; Wan, L.; You, H. Flood detection in Gaofen-3 SAR images via fully convolutional networks. *Sensors* **2018**, *18*, 2915. [[CrossRef](#)]
43. An, Q.; Pan, Z.; You, H. Ship detection in Gaofen-3 SAR images based on sea clutter distribution analysis and deep convolutional neural network. *Sensors* **2018**, *18*, 334. [[CrossRef](#)]

44. Chang, Y.; Li, P.; Yang, J.; Zhao, J.; Zhao, L.; Shi, L. Polarimetric calibration and quality assessment of the GF-3 satellite images. *Sensors* **2018**, *18*, 403. [[CrossRef](#)]
45. Liu, J.; Qiu, X.; Hong, W. Automated ortho-rectified SAR image of GF-3 satellite using Reverse-Range-Doppler method. In Proceedings of the IEEE IGARSS, Beijing, China, 10–15 July 2016; pp. 4445–4448.
46. Zhang, Q. System Design and Key Technologies of the GF-3 Satellite. *Acta Geod. Cartogr. Sin.* **2017**, *46*. [[CrossRef](#)]
47. Wang, Y.; Wang, C.; Zhang, H.; Dong, Y.; Wei, S. A SAR dataset of ship detection for deep learning under complex backgrounds. *Remote Sens.* **2019**, *11*, 765. [[CrossRef](#)]
48. Li, J.; Qu, C.; Shao, J. Ship detection in SAR images based on an improved faster R-CNN. In Proceedings of the 2017 BIGSAR DATA, Beijing, China, 13–14 November 2017; pp. 1–6.
49. Huang, X.; Yang, W.; Zhang, H.; Xia, G.S. Automatic ship detection in SAR images using multi-scale heterogeneities and an Contrario decision. *Remote Sens.* **2015**, *7*, 7695–7711. [[CrossRef](#)]
50. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* **2016**, arXiv:1603.04467.
51. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
52. Faster RCNN. Available online: https://github.com/smallcorgi/Faster-RCNN_TF (accessed on 18 April 2018).
53. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the ICLR, San Diego, CA, USA, 7–9 May 2015.
54. Everingham, M.; Gool, L.V.; Williams, C.K.I.; Winn, J.; Zisserman, A. The pascal visual object classes (VOC) challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [[CrossRef](#)]
55. LabelImg. Available online: <https://github.com/tzutalin/labelImg> (accessed on 18 April 2018).
56. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the CVPR, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
57. Wang, P.; Chen, P.; Yuan, Y.; Liu, D.; Huang, Z.; Hou, X.; Cottrell, G. Understanding convolution for semantic segmentation. In Proceedings of the 2018 IEEE WACV, Lake Tahoe, NV, USA, 12–15 March 2018; pp. 1451–1460.
58. Vachon, P.W.; Campbell, J.; Bjerkelund, C.; Dobson, F.; Rey, M. Ship detection by the RADARSAT SAR: Validation of detection model predictions. *Can. J. Remote Sens.* **1997**, *23*, 48–59. [[CrossRef](#)]
59. Vachon, P.W.; Thomas, S.J.; Cranton, J.; Edel, H.R.; Henschel, M.D. Validation of ship detection by the RADARSAT Synthetic Aperture Radar and the Ocean Monitoring Workstation. *Can. J. Remote Sens.* **2000**, *26*, 200–212. [[CrossRef](#)]
60. Vachon, P.W.; English, R.A.; Wolfe, J. Validation of RADARSAT-1 vessel signatures with AISLive data. *Can. J. Remote Sens.* **2007**, *33*, 20–26. [[CrossRef](#)]
61. Vachon, P.W.; Kabatoff, C.; Quinn, R. Operational ship detection in Canada using RADARSAT. In Proceedings of the IEEE IGARSS, Quebec City, QC, Canada, 13–18 July 2014; pp. 998–1001.

