



Review

Deep Learning for SAR Ship Detection: Past, Present and Future

Jianwei Li ¹, Congan Xu ^{1,2,*} , Hang Su ¹, Long Gao ¹ and Taoyang Wang ³

¹ Information Fusion Institute, Naval Aviation University, Yantai 264000, China; lgm_jw@163.com (J.L.); shpersonal_email@163.com (H.S.); gaolong14@nudt.edu.cn (L.G.)

² Advanced Technology Research Institute, Beijing Institute of Technology, Jinan 250300, China

³ School of Remote Sensing Information Engineering, Wuhan University, Wuhan 430000, China; wangtaoyang@whu.edu.cn

* Correspondence: 7520220053@bit.edu.cn

Abstract: After the revival of deep learning in computer vision in 2012, SAR ship detection comes into the deep learning era too. The deep learning-based computer vision algorithms can work in an end-to-end pipeline, without the need of designing features manually, and they have amazing performance. As a result, it is also used to detect ships in SAR images. The beginning of this direction is the paper we published in 2017BIGSAR DATA, in which the first dataset SSDD was used and shared with peers. Since then, lots of researchers focus their attention on this field. In this paper, we analyze the past, present, and future of the deep learning-based ship detection algorithms in SAR images. In the past section, we analyze the difference between traditional CFAR (constant false alarm rate) based and deep learning-based detectors through theory and experiment. The traditional method is unsupervised while the deep learning is strongly supervised, and their performance varies several times. In the present part, we analyze the 177 published papers about SAR ship detection. We highlight the dataset, algorithm, performance, deep learning framework, country, timeline, etc. After that, we introduce the use of single-stage, two-stage, anchor-free, train from scratch, oriented bounding box, multi-scale, and real-time detectors in detail in the 177 papers. The advantages and disadvantages of speed and accuracy are also analyzed. In the future part, we list the problem and direction of this field. We can find that, in the past five years, the AP₅₀ has boosted from 78.8% in 2017 to 97.8 % in 2022 on SSDD. Additionally, we think that researchers should design algorithms according to the specific characteristics of SAR images. What we should do next is to bridge the gap between SAR ship detection and computer vision by merging the small datasets into a large one and formulating corresponding standards and benchmarks. We expect that this survey of 177 papers can make people better understand these algorithms and stimulate more research in this field.

Keywords: SAR ship detection; SAR dataset; single-stage detector; two-stage detector; anchor free; train from scratch; oriented bounding box; multi-scale detection; deep learning; computer vision



Citation: Li, J.; Xu, C.; Su, H.; Gao, L.; Wang, T. Deep Learning for SAR Ship Detection: Past, Present and Future. *Remote Sens.* **2022**, *14*, 2712. <https://doi.org/10.3390/rs14112712>

Academic Editor: Dusan Gleich

Received: 6 May 2022

Accepted: 2 June 2022

Published: 5 June 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Synthetic aperture radar (SAR) remote sensing has become one of the important methods for marine monitoring due to its all-day, all-weather advantage. Ship detection in SAR images has broad prospects in both military and civilian fields [1,2].

The traditional detection method includes three steps: sea-land segmentation, CFAR (constant false alarm rate) detection, and discrimination [3,4]. In the sea-land segmentation step, the land pixels are rejected to avoid interference with the CFAR step. The common method is based on GIS (geographic information system) or image features. The gray histogram is the classical feature used for segmentation. In the second step, CFAR is usually used for ship detection. The distribution function is assumed to fit the pixel distribution of the SAR image. K, Weibull, and Rayleigh distribution are usually used in this step. To keep the probability of a false alarm at a constant value, the CFAR algorithm compares

the testing pixel with an adaptive threshold that is generated by the local background surrounding the testing pixel. After the pre-screening by CFAR, a discriminator is needed to reject the background. Discriminator includes two procedures: feature designing and classifier designing. According to the feature difference between ship chips and non-ship chips, this step can reduce the number of false alarms. The traditional detection method dominated this field for a long time.

With the development of deep learning-based object detection algorithms in computer vision (CV) [5], SAR researchers also began to seek inspiration from computer vision. There are three reasons that can explain the revival of deep learning. They are the arising of computing power, big data, and corresponding algorithms. As SAR images are not easily accessed, the deep learning-based detection method cannot be used for SAR ship detection at the beginning.

This problem was solved in 2017, as the first dataset SSDD (SAR Ship Detection Dataset) was open to the public. SSDD provides the same data and evaluation criteria for researchers, and it solves the problem that the traditional algorithms lack data and are not comparable in this field. Since then, more and more researchers adopt a deep learning-based method in this area. The deep learning-based algorithms also show great results compared with the traditional CFAR-based method. The active and open characteristics of computer vision also further promote the development of this field. We think that the emergence of SSDD means that this field comes into the deep learning era.

As far as we know, there are 177 papers [6–182] that use deep learning-based algorithms to detect ships in SAR images. However, there are no papers that review them yet. In order to summarize the achievements of the 177 papers and show the way for the future, we specially wrote this paper, hoping to contribute to the development of this field.

The rest of this review is arranged as shown in Figure 1. Section 2 briefly analyzes some work related to our paper. Section 3 summarizes the past of the traditional detection algorithms in SAR images. It mainly includes CFAR, hand-crafted features, and limited shallow representation. Section 4 introduces the present in deep learning-based detectors. We review the 177 papers, divide them into 10 categories and analyze them, respectively. Section 5 shows the future direction of this field. Section 6 is the conclusion of the paper.

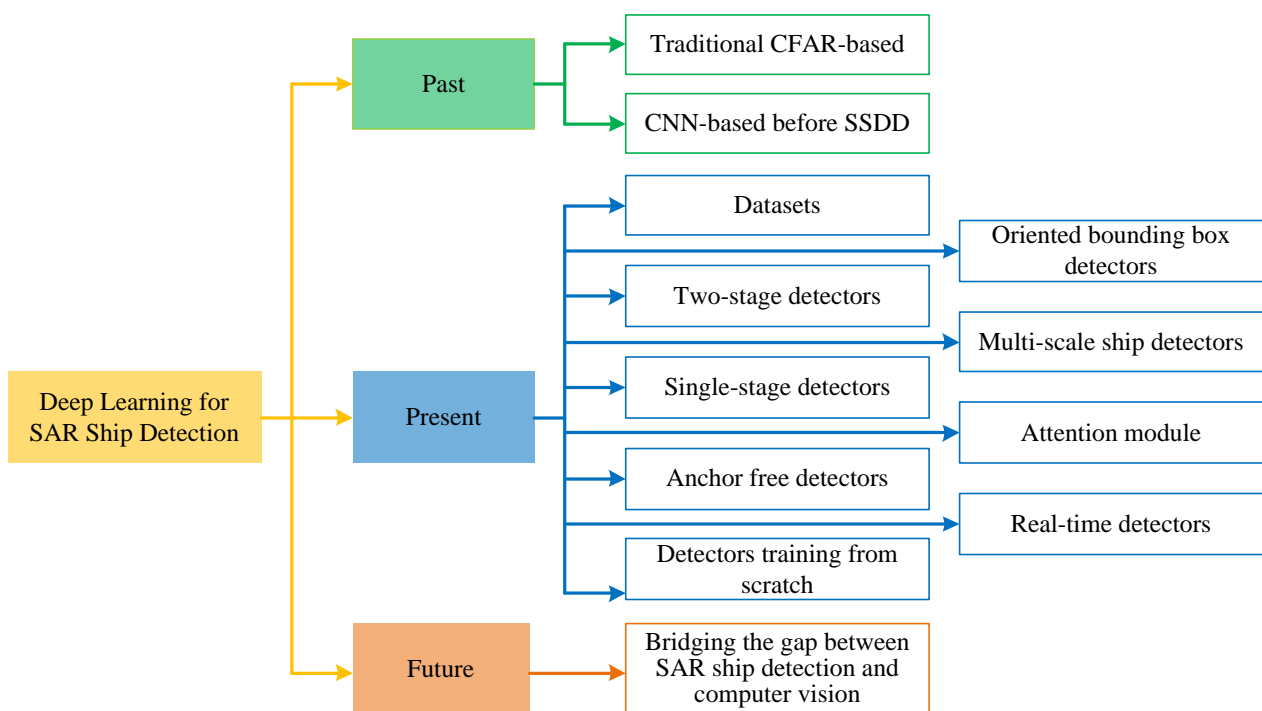


Figure 1. The overall architecture of the paper.

2. Related Work

As far as we know, few researchers have written review papers about this direction. This is partly due to the fact that this direction is new to some extent. At present, only three papers [105,127,170] have performed work related to our work.

Jerzy et al. [105] reviewed the papers from the last 5 years that discuss SAR ship detection. They mainly introduce the development of CFAR methods, CNN (convolutional neural network) based methods, GLRT (generalized likelihood ratio test) based methods, feature extraction-based methods, weighted information entropy-based methods, and variational Bayesian inference-based methods. Compared with paper [105], we mainly focus on the deep learning-based detection methods and do not focus on the traditional methods.

Mao et al. [127] solved the problem of the lack of performance benchmark for state-of-the-art methods on SSDD. Through this work, researchers can compare their work in the same experimental setup. They present 21 advanced detection models, including single-stage, two-stage, train from scratch detection algorithms, and so on. Compared with paper [127], we not only introduce the performance of different public datasets, but also classify all the papers, and summarize the principles and results of the algorithms.

Zhang et al. [170] solved the problem of the coarse annotations and ambiguous standards in SSDD. These improvements are beneficial for a fair comparison. It has played a great role in promoting the healthy development of this field. We suggest that researchers use the standards specified in this paper in the future. Compared with paper [170], our work is not limited to SSDD but introduces other datasets in this field. More importantly, our team has systematically analyzed, classified, and commented on the methods used, and pointed out the future research direction, which is beneficial to the development of this field.

In short, our work is different from the other papers. It is the first comprehensive review of SAR ship detection.

3. Past—The Traditional SAR Ship Detection Algorithms

Traditional detection algorithms in SAR images are based on hand-crafted features and limited shallow-learning representation. It can be divided into three steps: preprocessing, candidate region extraction, and discrimination.

CFAR is a common method for candidate region extraction. It can select potential ship regions. It first statistically models the clutter and then obtains the threshold value according to the false alarm rate. The pixels above the threshold are regarded as ship pixels, and those below the threshold are regarded as background. CFAR is essentially a segmentation-based algorithm, that is, the pixels are classified into two categories (ship or non-ship) according to the gray size, and then the ship pixel region is merged into the ship region. The performance of this method largely depends on the statistical modeling of sea clutter and the parameter estimation of the selected model. According to different SAR image products and practical application requirements, different statistical models such as Gaussian distribution, gamma distribution, log-normal distribution, Weibull distribution, and K distribution are proposed. Gaussian distribution and K distribution are the most commonly used. Generally speaking, when the scene is relatively simple, the CFAR method can achieve better results. However, for small ships and complex offshore scenes, due to the difficulty of modeling, it will have more false positives and poor detection performance.

Discrimination is generally realized by using artificially designed features and training classifiers. In addition to the simple features such as length, width, aspect ratio, and scattering point position, the features introduced from computer vision are also commonly used and have stronger robustness. Such as integral image features, HoG (histogram of oriented gradients), SURF (speeded up robust features), and LBP (local binary pattern). These features improve the performance of the detection algorithm. In classifier designing, decision trees, SVM, gradient boosting, and their improved versions also further improve the performance.

Feature and classifier designing have pushed this field forward in the past few years. However, since the rise of deep learning in 2012 [5], the above ideas are dwarfed in speed and accuracy. The object detection algorithm based on deep learning is an end-to-end processing method, as shown in Figure 2. It does not need to optimize multiple independent steps like the traditional method. It optimizes the whole detection system uniformly. It can adapt to various complex scenes (there is no need for sea-land segmentation in nearshore and port) and has very strong robustness. Therefore, in recent years, deep learning-based SAR ship detection algorithms have become a new research hotspot.

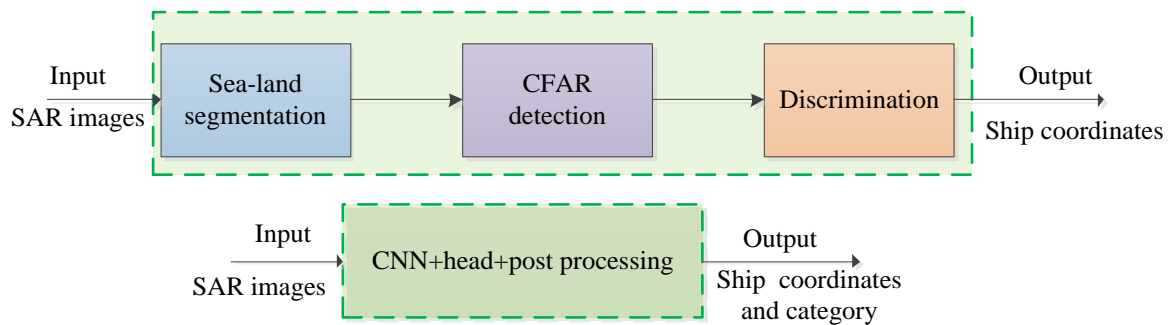


Figure 2. The differences between the CFAR-based detector and the deep learning-based detector.

The advantages and disadvantages of deep learning and traditional-based detection algorithms in SAR images can be proved by qualitative analysis or quantitative experiments.

The detection method based on CFAR has four main shortcomings through qualitative analysis. Firstly, CFAR needs to set the size of the protection window according to the size of the ship. It works well in the case of local uniform clutter in a single ship. If several ships with different sizes are close, the inconsistency between the change of ship size and the fixed protective window will lead to missed detection. Secondly, the CFAR algorithm needs to accurately model SAR images, which is difficult to implement. Thirdly, the essence of CFAR is an unsupervised algorithm, and its performance is essentially worse than the supervised algorithm (Faster R-CNN (region-based convolutional neural network), YOLO (you only look once), SSD (single shot detector), etc.). Fourthly, the CFAR algorithm and discrimination algorithm is a system pieced together after multiple links are debugged separately, and its performance cannot be compared with the end-to-end deep learning algorithm.

Sun Xian carried out a comparative experiment between the classical ship detection algorithm and the deep learning algorithm [65]. In the paper, the classical ship detection algorithms (optimal entropy automatic threshold method and CFAR method based on K distribution) are tested and analyzed on the AIR-SARShip-1.0 dataset. The experiment results are shown in Table 1. We can find that the performance of the deep learning algorithm is significantly better than that of the traditional algorithm.

Table 1. The performance difference between traditional-based detectors and deep learning-based detectors on the same condition [65].

Algorithms	AP (Average Precision)
CFAR method based on K distribution	19.2%
optimal entropy automatic threshold method	28.2%
Faster R-CNN	79.3%
SSD-512	74.3%

Before SSDD, there are six papers that use convolutional neural networks [1–6] to detect ships in SAR images. We think that these six papers are not based on deep learning. The reasons are as follows. Firstly, some algorithms are not end-to-end, they just use CNN as a component in the traditional detection process. Secondly, although some algorithms

are end-to-end, the dataset and evaluation criteria are not public, which is difficult for future researchers to reproduce, and the results are also not comparable.

Due to the important role of SSDD, we take the publication date of SSDD paper as the time separation point between the traditional and deep learning-based detection algorithms. Therefore, we believe that ship detection in SAR images entered the era of deep learning on 1 December 2017, as shown in Figure 3. A large number of researchers gradually began to abandon the traditional detection algorithm based on CFAR and adopt the advanced detection method based on deep learning [6–182]. The overview of these deep learning-based detectors in SAR images is the focus of this paper.

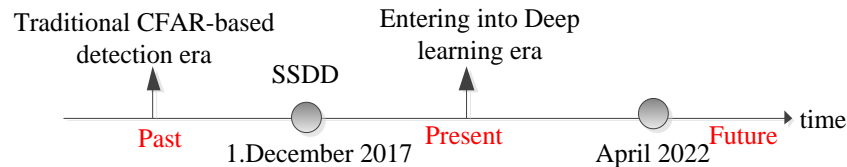


Figure 3. The time divisions of past, present and future.

4. Present—The Deep Learning-Based SAR Detection Algorithms

4.1. The General Overview of the 177 Papers

4.1.1. The Countries

In the country view, we can find that 90% of the papers' authors are Chinese, which is shown in Figure 4. There is no doubt that Chinese researchers have been the mainstream in this direction. Several public datasets are constructed by Chinese researchers, which further prove the above opinion.

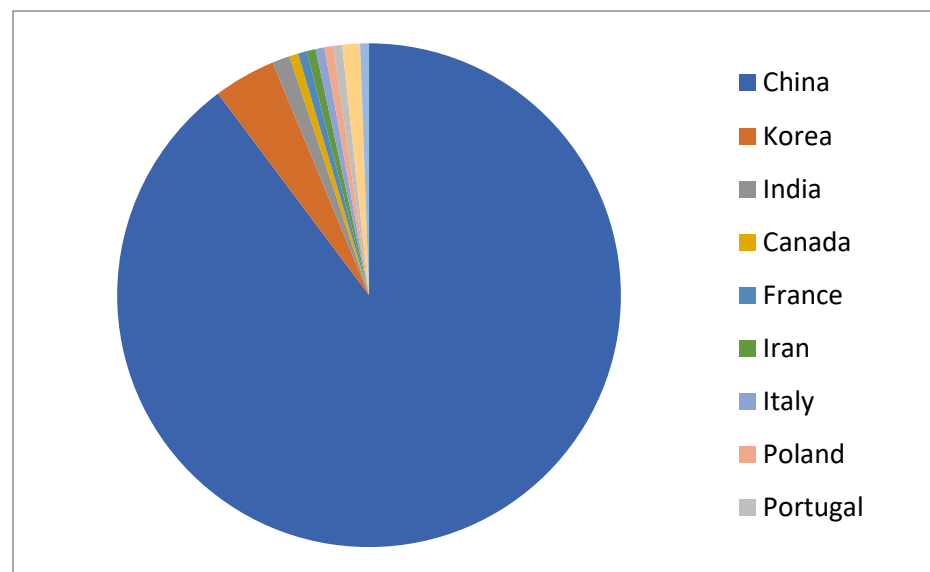


Figure 4. The percentages of the author countries.

4.1.2. Journal or Conference

A total of 63% of the 177 papers are published in journals, and 37% are in conferences. The most common journals and conferences are Remote Sensing and IEEE International Geoscience and Remote Sensing Symposium (IGARSS), respectively.

4.1.3. Timeline of the 177 Papers

The timeline of the deep learning based SAR ship detectors is shown in Figure 5.

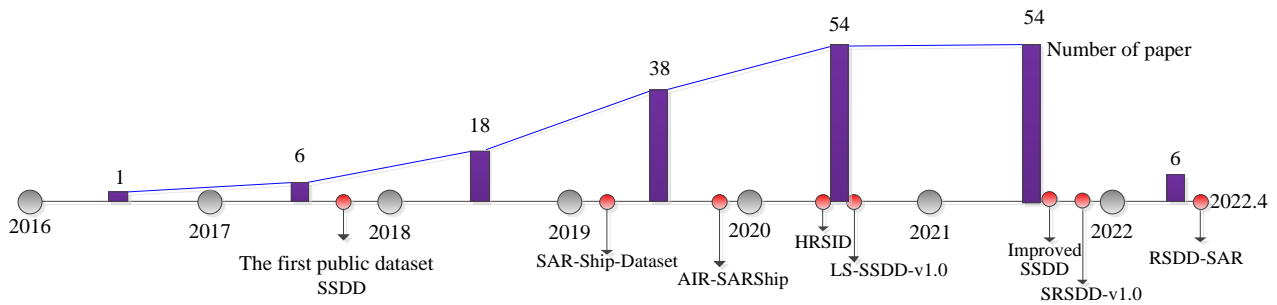


Figure 5. The timeline of the 177 papers.

Gray lines and gray circles in the figure represent the time, the purple bar represents the number of papers in the current year, and the red circles represent the public time of the dataset. From the timeline, we can find that in the passing five years, the number of papers about deep learning-based SAR ship detectors becoming more and more. The period 2016–2017 is in the transitional period between traditional and deep learning methods, and there are sporadic papers. Additionally, due to the lack of a unified dataset and the lack of in-depth understanding of deep learning and computer vision algorithms, the application of deep learning algorithms is not thorough. This situation did not change until the emergence of the first dataset (SSDD) paper at the end of 2017. SSDD discloses the data and evaluation criteria it used, which lays the foundation for the rapid development of this field. Since then, the fast lane of research was opened, and a large number of papers were published. The milestones of this field are the several open-access datasets, which are shown in the red circle above. With the increase in available datasets, more and more researchers are paying attention to this field.

4.1.4. The Datasets and Satellites That Are Used

The datasets that those papers used are shown in Table 2. We can see that SSDD is the most frequently used dataset for now. It was used 83 times, 62.4% percent of the total. Additionally, the usage of several public datasets shows a gradual upward trend.

Table 2. The datasets that are used in the past five years.

Datasets	2016	2017	2018	2019	2020	2021	2022	Total
SSDD	0	1	2	19	28	29	4	83
SSDD+	0	0	1	2	1	2	0	6
SAR-Ship-Dataset	0	0	0	1	4	14	1	20
AIR-SARShip1.0/2.0	0	0	0	1	3	5	0	9
HRSID	0	0	0	0	2	6	1	9
LS-SSDD-v1.0	0	0	0	0	1	1	1	3
Official-SSDD	0	0	0	0	0	1	0	1
SRSD-v1.0	0	0	0	0	0	1	0	1
RSDD-SAR	0	0	0	0	0	0	1	1
Total	0	1	3	23	40	58	8	133

Before the first dataset paper was published in 2017, different researchers adopted different SAR images and indicators to test their detectors. Thus, the results of the papers are not comparable. This phenomenon is not beneficial for the development of this field. In order to overcome this, we constructed the first dataset SSDD and it is open to the public. Meanwhile, we provide another dataset called SSDD+, which shares the same images with SSDD but has an oriented bounding box. With the rapid development of deep learning-based computer vision algorithms after 2019, SSDD draws more attention to researchers. Zhang analyzed the usage situation of SSDD in paper [170]. From this paper, we can find that SSDD becomes the most popular dataset, though it has many drawbacks. In addition to this, SAR-Ship-Dataset and AIR-SARShip are showing great potential to be

the popular dataset. The other datasets were seldom used, as their public dates are a little late. As deep learning models need more data to prevent overfitting, the future of this field is to merge them into a large dataset and provide the benchmarks on the large dataset with the common detection algorithms in computer vision. We think that, if the dataset is big enough, the benchmark is whole enough, and the maintenance is regular enough, it will be accepted by most researchers. This is the focus of our future work.

Table 3 shows the SAR satellites that papers used besides the public dataset. We can find that SAR images from Sentinel-1 are the most frequently used all the time. This is because the data are easy to acquire, and can be downloaded for free.

Table 3. The satellites that are used in the paper beside the datasets.

Satellites	2016	2017	2018	2019	2020	2021	2022	Total
Sentinel-1	1	4	7	6	6	2	0	26
RadarSat-2	1	0	2	2	2	0	0	7
ALOS PALSAR	0	1	0	0	0	0	0	1
TerraSAR-X	0	1	0	0	3	0	0	4
Gaofen-3	0	1	5	6	5	2	1	20
COSMO_SKYMed	0	0	1	0	2	0	0	3
AISSAR	0	0	1	0	0	0	0	1

However, as China's first C-band multi-polarization SAR satellite Gaofen-3 was officially put into use on 23 January 2017, the policy of obtaining Gaofen-3 images has become easier and easier. More and more papers use Gaofen-3 as the source image.

4.1.5. Deep Learning Framework

A deep learning framework can reduce the workload of researchers [183–187]. So, since the emergence of CAFFE (Convolution Architecture For Feature Extraction) [188] in 2017, it gets more and more attention from researchers. Table 4 shows the deep learning framework those 177 papers used. We can find that in the beginning years (2017–2018), CAFFE is the most frequently used framework. It is because CAFFE is the first common deep learning framework that researchers use and most of the detection algorithms in computer vision are based on CAFFE, for example, Faster R-CNN and SSD. In order to improve the efficiency of the deep learning framework, Google provided TensorFlow [189] in 2017. Compared with CAFFE, Tensorflow is more powerful and easier to use. A lot of researchers adopt Tensorflow as their framework. After Tensorflow, PyTorch [190] was promoted by Facebook FAIR in 2017. PyTorch is more suitable for researchers, and the number of users surpasses the Tensorflow gradually. In addition to CAFFE, Tensorflow and Pytorch, Keras, DarkNet and PaddlePaddle are also used by some researchers. Due to the fact that most detection algorithms in computer vision are based on Tensorflow and Pytorch, we recommend researchers in this area use them as the deep learning framework.

Table 4. The deep learning framework those papers used.

Framework	2016	2017	2018	2019	2020	2021	2022	Total
Caffe	0	3	9	3	6	2	0	23
Tensorflow	0	2	3	12	5	7	0	29
Pytorch	0	0	0	3	19	18	6	44
Keras	0	0	0	1	3	3	0	7
Darknet	0	0	0	0	1	3	0	4
PaddlePaddle	0	0	0	0	0	1	0	1

4.1.6. Performance Evolution

Tables 5–11 show the performance of several public datasets. In the 'AP' column, the large number represents AP₅₀, and the small one represents AP. AP50 refers to the average

precision with IoU (intersection over union) = 50%. AP refers to the value of IoU from 50% to 95% in steps of 5% and then calculates the average value of AP under these IoUs. Normally, AP₅₀ is higher than AP. AP₅₀ and AP are usually used in PASCAL VOC and MS COCO, respectively. Since the dataset contains only one class of ships, the mAP value is the same as the AP value.

In the table, the italics represent the performance of two-stage detectors, and the non-italics represent the performance of single-stage detectors. The blue, red, green, purple, and golden colors represent anchor-free, train from scratch, oriented bounding box, multi-scale, and attention detectors. The underlines represent real-time detectors.

The number of papers in Tables 5–11 is less than in Table 2. This is because some of the papers in Table 2 did not use the AP or AP₅₀ as the evaluation indicator, so we do not show them in Tables 5–11.

From Table 5, we can see that there are 52 papers that are trained and tested on SSDD. Additionally, in the past five years, AP₅₀ of detectors on SSDD boosted from 78.8% in 2017 to 97.8% in 2022. The testing time is also getting faster and faster. What should be noticed is that as the train-test division is ambiguous in the original SSDD, so the AP in Table 5 is not comparable to some extent. That is also why we recommend the following researchers adopt the Improve SSDD [170] as the new standard.

Table 5. The performance evolution of detectors on SSDD (The data come from the 177 papers).

No.	Date	AP	Time	No.	Date	AP	Time
11	1 December 2017	78.8%	173 ms	<u>104</u>	<u>14 October 2020</u>	<u>92.6%</u> <u>56.5%</u>	<u>7.39 ms</u>
<u>15</u>	<u>9 March 2019</u>	<u>91.3%</u>	<u>96 ms</u>	111	16 November 2020	91.84%	
39	2 April 2019	89.76%	10.938 ms	<u>115</u>	<u>2 December 2020</u>	<u>89.79%</u>	
<u>40</u>	<u>21 May 2019</u>	<u>90.16%</u>	<u>21 ms</u>	116	3 December 2020	90.7%	13.6 ms
<u>43</u>	<u>24 July 2017</u>	<u>79.78%</u>	<u>28.4 ms</u>	117	4 December 2020	88.33%	74 FPS
<u>51</u>	<u>23 September 2019</u>	<u>80.12%</u>	<u>9.28 ms</u>	<u>118</u>	<u>7 December 2020</u>	<u>94.6%</u>	<u>3.9 ms</u> <u>258 FPS</u>
<u>54</u>	<u>24 October 2019</u>	<u>94.13%</u>	<u>9.03 ms</u> <u>111 FPS</u>	<u>121</u>	<u>28 December 2020</u>	<u>95.1%</u>	<u>33 ms</u>
55	24 October 2019	90.04%	87 ms	<u>131</u>	<u>12 February 2021</u>	<u>95.7%</u> <u>63.4%</u>	
56	14 November 2019	94.7%		<u>132</u>	<u>17 February 2021</u>	<u>93.78%</u>	<u>202 FPS</u>
62	14 November 2019	83.4%		134	27 February 2021	80.45%	
<u>63</u>	<u>14 November 2019</u>	<u>90.44%</u>	<u>96.04 ms</u>	149	17 March 2021	94.41%	31 FPS
68	December 2019	96.93%	8.72 ms	146	23 March 2021	95.52%	
69	2 January 2020	97.9%	103 ms	148	31 March 2021	92.09%	
74	19 March 2020	64.6%					
		96.4%	106.4 ms	151	13 May 2021	88.08%	12.25 ms
		67.4%					
78	30 March 2020	94.2%	0.93 M	154	9 June 2021	98.4%	
		59.5%					
<u>81</u>	<u>3 April 2020</u>	<u>94%</u>		<u>157</u>	<u>30 June 2021</u>	<u>61.4%</u>	<u>45 FPS</u>
<u>82</u>	<u>16 April 2020</u>	<u>90.08%</u> <u>68.1%</u>		<u>158</u>	<u>1 July 2021</u>	<u>96.8%</u> <u>62.7%</u>	<u>438 ms</u>
<u>84</u>	<u>22 April 2020</u>	<u>97.07%</u>	<u>233 FPS</u>	160	13 July 2021	97.2%	
90	25 May 2020	93.96%		<u>161</u>	<u>14 July 2021</u>	<u>95.29%</u>	<u>11 FPS</u>
93	24 June 2020	94.72%	63.2 ms	170	6 December 2021	97.8%	
						64.9%	
<u>96</u>	<u>21 July 2020</u>	<u>96.08%</u>	<u>4.51 ms</u> <u>222 FPS</u>	171	10 December 2021	82.2%	5.2 ms
<u>98</u>	<u>21 August 2020</u>	<u>81.17%</u>	<u>24 ms</u>	173	22 December 2021	97.4%	42.5 FPS
99	21 August 2020	83.4%		174	6 February 2022	95.6%	
<u>100</u>	<u>31 August 2020</u>	<u>90.57%</u>	<u>17.2 ms</u>	175	25 February 2022	61.1%	
<u>102</u>	<u>6 October 2020</u>	<u>86.3%</u>		176	25 February 2022	97.8%	17.5 FPS
		<u>95.6%</u>				95.03%	47 FPS
<u>103</u>	<u>14 October 2020</u>	<u>61.5%</u>		177	19 March 2022	97.0%	

From Table 6, we can see that there are only four papers that are trained and tested on SSDD+, and the AP₅₀ performance is increased from 84.2% in 2018 to 94.46% in 2021. The overall performance is a bit lower than that of SSDD. That is because the detectors on SSDD+ should predict an additional parameter (angle). We also find that the SSDD+ is seldom used compared with SSDD. That is, few researchers are interested in oriented bounding box detection in this area.

Table 6. The performance evolution of detectors on SSDD+ (the data come from the 177 papers).

No.	Date	AP	Time
20	29 August 2018	84.2%	40 FPS
41	26 June 2019	81.36%	
83	20 April 2020	90.11%	62.77 ms
124	8 January 2021	94.46%	

From Table 7, we can see that there are 14 papers that are trained and tested on SAR-Ship-Dataset, and the AP₅₀ performance is boosted from 89.07% in 2019 to 96.1% in 2021. The running speed is also accelerated to 60.4 FPS with 96.1% AP. The overall performance is a bit lower than that of SSDD. That is because this dataset is relatively larger than SSDD.

Table 7. The performance evolution of detectors on SAR-Ship-Dataset (the data come from the 177 papers).

No.	Date	AP	Time	No.	Date	AP	Time
38	29 March 2019	89.07%		138	17 February 2021	92.4%	
89	20 May 2020	94.7%	18 ms	157	19 May 2021	93.46%	339 FPS
113	30 November 2020	91.89%	12.05 FPS	158	8 June 2021	95.52%	
114	30 November 2020	91.07%		163	1 July 2021	95.8%	
123	5 January 2021	90.25%	22 ms	166	14 July 2021	94.39%	
133	17 February 2021	93.9%		178	22 December 2021	96.1	60.4 FPS
136	17 February 2021	95.1%		179	6 February 2022	95.1	

From Table 8, we can see that there are only four papers are trained and tested on AIR-SARShip, and the AP₅₀ performance is boosted from 88.01% in 2019 to 92.49% in 2021. In addition, the running speed becomes 7.98 times faster (from 41.6 ms to 5.22 ms). The overall performance is a bit lower than that of SSDD.

Table 8. The performance evolution of detectors on AIR-SARShip (the data come from the 177 papers).

No.	Date	AP	Time	Version
65	1 December 2019	88.01%	24 FPS	1.0
97	13 August 2020	86.99%		1.0
130	8 February 2021	80.9%		1.0
171	1 December 2021	92.49%	5.22 ms	2.0

From Table 9, we can see that there are only nine papers are trained and tested on HRSID, and the AP₅₀ performance is boosted from 89.3% in 2019 to 94.4% in 2021. The overall performance is a bit lower than that on SSDD. That is because this dataset is relatively larger than SSDD.

Table 9. The performance evolution of detectors on HRSID (the data come from the 177 papers).

No.	Date	AP	No.	Date	AP
94	29 June 2020	89.3% 69.4%	168	6 August 2021	89.2% 68.0%
110	10 November 2020	<i>not given</i> 84.4%	174	14 February 2022	91.4% 66.4%
120	23 December 2020	91.99% 68.5%	175	6 December 2021	94.4% 72%
131	12 February 2021	92.4% 69.5%	178	22 December 2021	88.3%
165	13 July 2021	90.7% 69.4%			

From Table 10, we can see that there are only three papers are trained and tested on LS-SSDD-v1.0, and the AP performance is boosted from 72.3% in 2019 to 75.5% in 2022. The overall performance is a bit lower than that on SSDD. LS-SSDD-v1.0 is specially used for large-scale SAR ship detection, which is fit for satellite-based SAR systems. It should be used more in the future.

Table 10. The performance evolution of detectors on LS-SSDD-v1.0 (the data come from the 177 papers).

No.	Date	AP
101	15 September 2020	75.3%
168	6 August 2021	71.7%
180	25 February 2022	75.5%

The above datasets are relatively smaller than the datasets used in computer vision. In order to improve the generalization ability of the detector, researchers should use a large dataset. Some researchers merge several datasets into a large one as shown in Table 11. From Table 11, we can see that there are three papers that are trained and tested on the composite dataset, and the AP performance is 81.13%, 71.4%, and 95.1%, respectively. As deep learning-based detectors are data-hungry, we should merge the public datasets into a large one to prevent over-fitting.

Table 11. The performance evolution of detectors on other datasets (the data come from the 177 papers).

No.	Date	AP ₅₀	Time	Datasets
108	30 October 2020	81.13%	35.5 ms	SSDD + SAR-Ship-Dataset
125	27 January 2021	71.4%	2920 ms	SAR-Ship-Dataset +AIRSAR-Ship-1.0
167	26 July 2021	95.1%		HRSID + SSDD + IEEE 2020 Gaofen Challenge

4.2. The Algorithm Taxonomy of the 177 Papers

We divide the 177 papers into 10 categories, they are papers about datasets, two-stage detectors, single-stage detectors, anchor-free detectors, train from scratch detectors, detectors with the oriented bounding box, multi-scale detectors, detectors with attention module, real-time detectors, and others. The percentages of each algorithm are shown in Table 12. What should be explained is that the summation of the percentages is larger than 1. This is because many algorithms in the papers have several attributes. For example, it not only belongs to the single-stage detector but is also trained from scratch.

Table 12. The percentage of each algorithm.

Algorithms	Datasets	Two-Stage	Single-Stage	Anchor Free	Scratch
Percentage	5%	26.7%	25.6%	5.1%	4.0%
Algorithms	Oriented	Multi-scale	Attention	Real-time	Others
Percentage	5.7%	14.2%	5.1%	13.1%	14.2%

From Table 12, we can find the following conclusions. Firstly, there are eight papers that introduce the datasets to the researchers. They make a great contribution to this field. Secondly, two-stage detectors used in this field are slightly more than single-stage detectors. This is partly because the two-stage detectors have higher accuracy than the single-stage detectors in most cases. In addition, accuracy is the first consideration at the moment. Thirdly, anchor-free detectors, detectors trained from scratch, oriented bounding box detectors, and detectors with attention modules almost have a percentage of 5–6% among the 177 papers. This is because, as the above four directions are rare, they are not yet noticed by many researchers. In fact, these directions can overcome the problems of the ship size distribution abnormal and the lack of SAR images. They should be paid more attention in the future. Fourthly, almost 14% of papers are about multi-scale SAR ship detection, which is a little higher than other directions. This is because, compared with objects in computer vision images, ships in SAR are rather small. In order to improve the performance, detectors should pay more attention to multi-scale ships. Fifthly, 14.20% of papers are classified as others, which represents that these papers do not belong to the nine categories. Sixthly, only three papers, which is 1.7% of the 177 papers are reviewed in this field. Considering the active research in this field, it is not enough for now. This is one of the motivations for our work.

4.3. The Public Datasets

4.3.1. Overview

As far as we know, there are 10 public datasets that are used for training and detecting ships in SAR images. They are SSDD(SSDD+) [11], SAR-Ship-Dataset [38], AIR-SARShip1.0 [65], HRSID [94], LS-SSDD-v1.0 [101], AIR-SARShip2.0 [191], Official-SSDD [170], SRSDDD-v1.0 [177] and RSDD-SAR [192]. Table 13 shows the detailed information of the 10 public datasets, in which the annotations of SSDD+, Official-SSDD, SRSDDD-v1.0, and RSDD-SAR are the oriented bounding box.

Table 13. Detail information of existing public datasets.

Dataset	Date	Source	Resolution	Image Size	Images/Ships	Annotation
SSDD (SSDD+)	1 December 2017	RadarSat-2 TerraSAR Sentinel-1	1 m–15 m	190–668	1160/2456	vertical oriented
SAR-Ship-Dataset	29 March 2019	Gaofen-3 Sentinel-1	3 m–25 m	256 × 256	43,918/59,535	vertical
AIR-SARShip-1.0	1 December 2019	Gaofen-3	1 m, 3 m	3000 × 3000	31	vertical
AIR-SARShip-2.0	25 August 2021			1000 × 1000	300	
HRSID	29 June 2020	Sentinel-1 TerraSAR	0.5 m, 1 m, 3 m	800 × 800	5604/16,951	polygon
LS-SSDD-v1.0	15 September 2020	Sentinel-1	5 m, 20 m	24,000 × 16,000	15/6015	vertical oriented
Official-SSDD	15 September 2021					
SRSDDD-v1.0	15 December 2021	Gaofen-3	1 m	1024 × 1024	666/2275	recognition
RSDD-SAR	April 2022	Gaofen-3 TerraSAR	2–20 m	512 × 512	7000/10,263	oriented

In addition to these, SMCDD [182] is a good dataset based on China's first commercial SAR satellite HISEA-1. It has 1851 bridges, 39,858 ships, 12,319 oil tanks, and 6368 aircraft. It shows a great advantage in multi-class ship detection.

In the future, it is very necessary to combine the above datasets into a large one to avoid the problem of overfitting.

In the following part, we will introduce the details of the datasets and evaluate their advantages and drawbacks.

4.3.2. SSDD, SSDD+ and Official-SSDD

We made our dataset SSDD publicly available at the conference of 2017BIGSAR DATA in Beijing [11]. SSDD is the first open dataset in this community. It can be a benchmark for researchers to train and evaluate their algorithms. In SSDD, there are a total of 1160 images and 2456 ships. The ships in SSDD have rich diversity, including small-size ships, complex backgrounds, and dense arrangements near the wharf. We also give the statistical results of the length, width, and aspect ratio of the ship bounding box in SSDD. The papers that used SSDD and their performance are shown in Table 5.

At the same time, based on 1160 SAR images of SSDD, we use the oriented bounding box to relabel the ship and obtain the dataset SSDD+. SSDD+ is the first dataset for SAR ship detection with an oriented bounding box. The papers that used SSDD+ and their performance are shown in Table 6.

At that time, there were some problems in SSDD due to the lack of understanding of computer vision and deep learning. The drawbacks of SSDD are the coarse annotations and ambiguous standards of use. It hinders fair comparisons and effective academic exchanges in this field.

In September 2021, Zhang [170] systematically analyzed and improved the problem of SSDD; they call it Official-SSDD. Zhang relabeled ships in SSDD and proposed three new datasets; they are bounding box SSDD, rotatable bounding box SSDD, and polygon segmentation SSDD. In addition, they also formulate some standards: the train-test division, the inshore-offshore protocol, the ship-size definition, the determination of the densely distributed small ship ships, and the determination of the densely parallel berthing at ports ship samples. We suggest that follow-up researchers use the Official-SSDD and standards proposed in paper [170] to carry out their relevant research.

4.3.3. SAR-Ship-Dataset

The training of the deep learning model depends on a large amount of data, and the amount of SSDD is relatively small. To solve this problem, Wang Chao [38] constructed a dataset called SAR-Ship-Dataset. SAR-Ship-Dataset contains 43,819 images and 59,535 ships, which are more than SSDD. The sources of SAR-Ship-Dataset are 102 Gaofen-3 images and 108 Sentinel-1 SAR images. These ships have distinct scales and backgrounds. The resolution, incident angle, polarization mode, and imaging mode are also diverse, which are helpful for the deep learning models to fit different conditions. The papers that used SAR-Ship-Dataset and their performance are shown in Table 7.

4.3.4. AIR-SARShip

A dataset containing more diverse scenes and covering various types of ships will help to train a model with better performance, stronger robustness, and higher practicability. In order to achieve the above purpose, Sun Xian constructed a dataset based on the Gaofen-3 satellite, named AIR-SARShip-1.0 [65].

It contains a total of 31 large images. A total of 21 images are training data and the other 10 images are testing data. The image resolutions include 1 m and 3 m. The image size is about 3000×3000 pixels. The information of each image in the dataset includes image number, pixel size, resolution, sea state, scene, and the number of ships. The dataset has the characteristics of a large scene and a small ship.

On the basis of version 1.0, Sun Xian and other researchers added more Gaofen-3 data to build AIR-SARShip-2.0. The dataset contains 300 SAR images. The scene types include ports, islands, reefs, sea surfaces with different levels of sea conditions, etc. The annotation information includes the location of ships, which has been confirmed by professional interpreters.

The papers that used AIR-SARShip and their performance are shown in Table 8.

4.3.5. HRSID

The original SAR image used to construct HRSID [94] includes 99 Sentinel-1B images, 36 TerraSAR-X images, and 1 TanDEM-X image. HRSID has 5604 high-resolution SAR images with 800×800 pixels to meet the needs of actual training for GPU. It is designed for ship detection and segmentation based on CNN, and it only contains one category of ships. It is divided into 65% training set and 35% testing set. It uses polygons to label the ship. In order to reduce the deviation of the ship detection algorithm, the interference derived from the ship is marked as a part of the ship.

According to statistics, the total number of ships marked in HRSID is 16,951, and each SAR image contains an average of three ships. The number of small ships, medium ships, and large ships accounted for 54.5%, 43.5%, and 2% of all ships, respectively. The bounding box areas of small ships, medium ships, and large ships account for more than 0~0.16%, 0.16~1.5%, and 1.5% of SAR images, respectively. Therefore, ships are sparsely distributed in SAR images.

The papers that used HRSID and their performance are shown in Table 9.

4.3.6. LS-SSDD-v1.0

Zhang Xiaoling [101] constructed the SAR ship detection dataset LS-SSDD-v1.0 with a large scene and small ships. The dataset consists of 15 pieces with a size of $24,000 \times 16,000$ pixels Sentinel-1 SAR images. Each image is directly divided into 600 sub-images with 800×800 pixels. The dataset contains 6015 ships. LS-SSDD-v1 can support researchers to flexibly apply the dataset. The optical information provided in Google Earth software and ship information provided by AIS is used for the annotation of LS-SSDD-v1.0. The coastline of the imaging area in the dataset is relatively complex, the land area is smaller than the ocean area, and the ships in the inland river are densely distributed. The dataset has the following characteristics: contains large scenes, focus on the small ships, rich pure background, etc. It also provides a large number of performance benchmarks of detection algorithms on datasets.

The papers that used LS-SSDD-v1.0 and their performance are shown in Table 10.

4.3.7. SRSDD-v1.0

The original images of SRSDD-v1.0 are from Gaofen-3 [177]. It contains 30 panoramic SAR images of port areas. It is annotated with an oriented bounding box. Optical images (Google Earth or GF-2) are used to assist the annotation. The image size is set to 1024×1024 . The annotation format is the same as DOTA. The coordinates of the four corners of the box, the category, and whether it is difficult to identify are given in annotation files.

It contains 666 images. A total of 420 images with 2275 ships include the land cover. A total of 246 images with 609 ships only contain the sea in the background. It has six categories: ore-oil ships (166), bulk cargo ships (2053), fishing boats (288), law enforcement ships (25), dredger ships (263), and container ships (89). The dataset has a certain data imbalance problem.

4.3.8. RSDD-SAR

The RSDD-SAR dataset consists of 84 scenes of Gaofen-3 and 41 scenes of TerraSAR-X. RSDD-SAR has 7000 images, including 10,263 ships, of which 5000 are randomly selected as the training set and the other 2000 as the testing set. By analyzing the distribution of ship angle and aspect ratio in the dataset, it can be found that the angle of ships in the dataset is evenly distributed between 0° and 180° , and the aspect ratio is concentrated between two

and six. It indicates that the dataset has the characteristics of arbitrary rotation direction and a large aspect ratio. The dataset has the characteristics of a high proportion of small ships, which can be used to verify the performance of a small ship detection algorithm. The RSDD-SAR dataset contains vast sea areas, ports, docks, waterways, and other scenes with different resolutions, which are suitable for practical applications.

4.4. Two-Stage Detectors

The deep learning-based object detection algorithm can be divided into single-stage detectors and two-stage detectors. The single-stage detectors use a full convolution network to classify and regress these anchor boxes once to obtain the detection results. The two-stage detectors use a CNN to classify and regress these anchor boxes twice to obtain the detection results. The principles of single-stage and two-stage detection algorithms are shown in Figure 6.

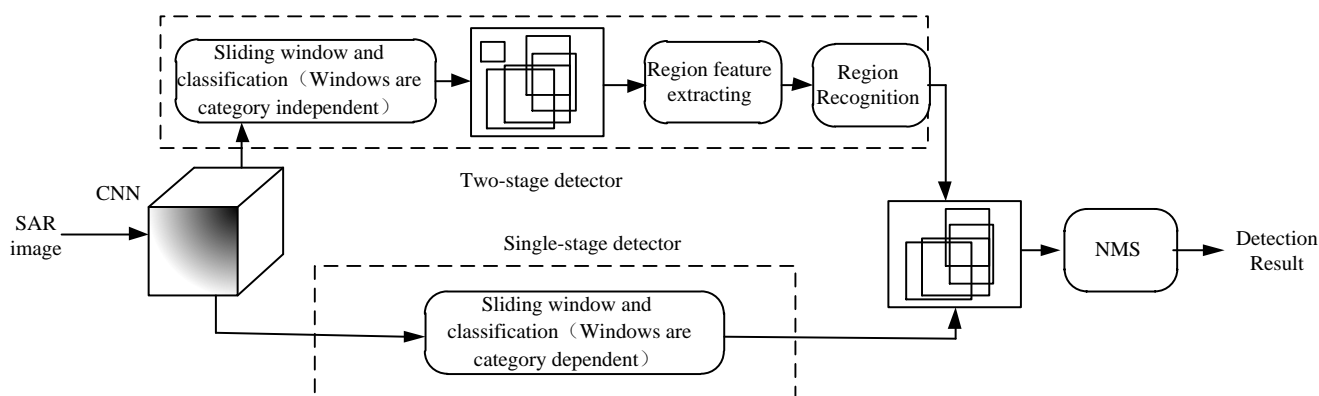


Figure 6. The principle of single-stage and two-stage detectors.

Classical two-stage detectors are Faster R-CNN, R-FCN (fully convolutional network) [193], feature pyramid networks (FPN) [194], Cascade R-CNN [195], Mask R-CNN [196], and so on [197]. Faster R-CNN is the foundation work, and most of the two-stage detectors are improved based on it.

Among the 177 papers, most of the papers are improved from the following aspects: backbone network, region proposal network (RPN), anchor box, loss function, and non-maximum suppressing (NMS). They are shown in Figure 7. Compared with computer vision, the research in this field lags behind, and other more advanced two-stage detection algorithms have not been used here.

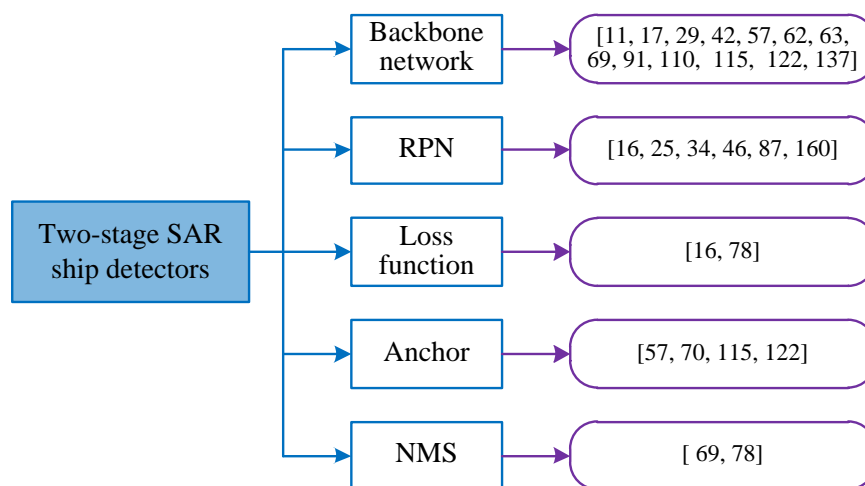


Figure 7. The two-stage SAR ship detectors.

4.4.1. Backbone Network

There are three main directions in the improvement of the backbone network, namely FPN, feature fusion, and attention.

FPN produces a feature pyramid structure that combines low-resolution, which has strong semantic features with high-resolution, which has weak semantic features. It includes a bottom-up channel, a top-down channel, and a skipping connection. It predicts independently at all levels, which only brings minimal additional calculation and storage consumption. It improves the detection result of small-size ships and thus is widely used. A lot of work has been completed to improve FPN in the computer vision field, such as ASFF [198], NAS-FPN [199], and BiFPN [200].

There are six papers [42,62,63,69,91,110] adopted and improved FPN in this field. Cui et al. [42] proposed a DAPN (dense attention pyramid network) structure. It densely connects the convolution block attention module from the top to the bottom of the pyramid network. By this, rich features including resolution and semantic information are extracted for multi-scale ship detection. Li et al. [62] used a convolution block attention module (CBAM) [201] to control the degree of upper- and lower-level feature fusion in FPN. Liu et al. [63] proposed a scale-transferrable pyramid network. It densely connects each feature map from top-to-down using scale-transfer layer. It can expand the resolution of feature maps, which is helpful for detection. Wei et al. [69] adopted a parallel high-resolution feature pyramid network to make full use of the feature mapping of high-resolution and low-resolution convolution for SAR ship detection. Zhao et al. [91] adopted receptive fields block and convolutional block attention module to build a top-down fine-grained feature pyramid. It can capture features of ships with large aspect ratios and enhance local features with their global dependences. Hu et al. [110] used a dense connection to a feature pyramid network, in which the shallow features and deep features are processed differently. It considered the differences between different levels.

There are three papers [11,57,115] that improved the backbone network through feature fusion. Li et al. [11] fused the feature maps from convolutional layer 3 to layer 5. The fusion includes the normalization and 1×1 convolution. Normalizing each ROI (region of interest) pooling tensor can reduce the scale differences between the following layers. It can prevent the 'larger' features from dominating the 'smaller' ones and make the algorithm more robust. This modification stabilizes the system and increases the accuracy. Yue et al. [57] fused the semantically strong features with the low-level high-resolution features, which is helpful for reducing false alarms. Li et al. [115] presented a jump connection structure to extract the features of each scale target in the SAR image. It can improve the ability of recognition and localization.

There are five papers [17,29,62,122,137] that improve the backbone network through the attention module (SENet). It squeezes the feature map along the space and the channel direction, which can explicitly model the interdependence between feature channels, and then automatically obtain the importance of each feature channel through learning. It can improve the useful features and suppress the features that are not useful for the current task according to the importance.

4.4.2. RPN

Another direction is improving the RPN module of Faster R-CNN. Paper [16,25,34,46,87] did not use a single feature map to generate proposals but generated proposals from each fused feature map. Liu et al. [36] designed a scale-independent proposal generation module, which extracts the features such as edge, super-pixel, and strong scattering component from SAR image to obtain ship proposals, and sorts whether the proposals contain ships from the integrity and tightness of the contour. In paper [160], candidate proposals are extracted from the original SAR image and the denoised SAR image, respectively, and then combined to reduce the impact of noise in the SAR image on ship detection. They can improve the performance of multi-size ships to some extent.

4.4.3. Loss Function

Faster R-CNN forces the ratio of positive and negative samples to 1:3 to solve the problem of unbalanced positive and negative proposals. Similar work in the field of computer vision includes focal loss, OHEM (online hard example mining) [202], GHM (gradient harmonizing mechanism) [203], and Libra R-CNN [204]. The paper [16,78], respectively, adopted focal loss to increase the weight of hard negative samples and reduce the weight of simple samples, so as to avoid the problem that a large number of simple samples cover a small number of hard negative samples in the training process.

4.4.4. Anchor and NMS

Faster R-CNN uses three scales and three aspect ratios, producing a total of $60 \times 60 \times 9$ anchor boxes. However, the ship size in the SAR image is extremely small and sparse. There will be a waste in using dense anchor boxes for ship detection in SAR images. Yue et al. [57] and Wang et al. [122] set the parameters of the anchor box based on the analysis of the actual size and distribution of the ship, mainly reducing the size of the anchor box and selecting the appropriate shape. Chen et al. [70] and Li et al. [115] used K-means to obtain the distribution of the ship size, so as to obtain the appropriate anchor box and reduce the difficulty of learning.

Wei et al. [69] and Wang et al. [78] used soft NMS [205] to replace NMS. Soft NMS improves the discrimination process of IoU and threshold in the cycle process and uses weights to attenuate scores to avoid accuracy loss.

4.4.5. Others

ISASDNet (instance segmentation assisted ship detection network) was proposed based on Mask R-CNN in the paper [163]. It has two branches: detection and segmentation. The two branches output interaction to improve the detection results. Gui et al. [34] proposed a lightweight detection head with a large separable convolution kernel and position-sensitive pooling, which improves the detection speed.

4.5. Single-Stage Detectors

The two-stage detectors generate a candidate box first and then identify and regress the candidate box, which is quite different from the principle of human eyes. The single-stage detectors only need to look at the picture once and can predict what the object is and where the object is. It is similar to the human eyes. In addition, they are quite faster than two-stage detectors.

Classical single-stage detectors are YOLO, SSD, RetinaNet [206], and CornerNet [207]. YOLO and SSD are the two most popular single-stage detection algorithms, and most of the subsequent single-stage works are based on them.

The single-stage ship detectors in SAR images are shown in Figure 8.

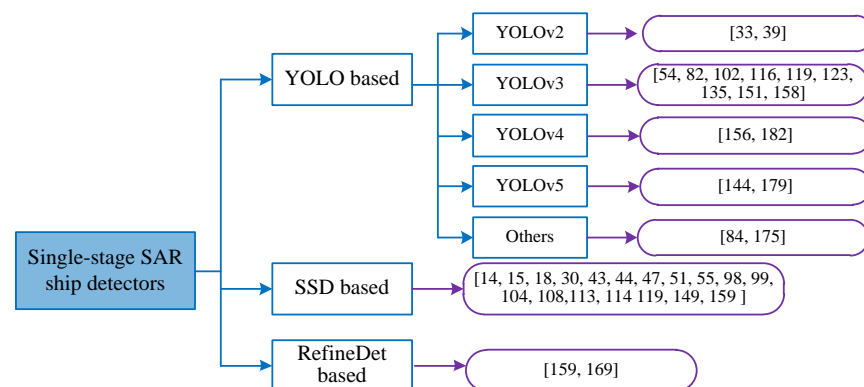


Figure 8. The single-stage SAR ship detectors.

4.5.1. YOLO and SSD Series in Computer Vision

YOLOv1 [186] regards object detection as a regression problem, and it outputs the spatially separated bounding box and related class probability simultaneously. A neural network can predict the bounding box and class probability from the image in one forward calculation. The speed is very fast, but it has an inaccurate location prediction, and the recall is low. YOLOv2 [208] uses the multi-scale training method. It predicts the offset rather than the parameter itself. The offset value is slightly smaller, which can increase the accuracy of prediction. It uses an anchor mechanism to obtain anchor box parameters by clustering the object size in the dataset. The backbone network adopts DarkNet-19. Although the detection head has changed from 7×7 to 13×13 , the detection result of a small object is still poor. The YOLOv3 [209] detection head includes three branches: 13×13 , 26×26 , and 52×52 , which can take into account large, medium, and small objects and make the location prediction more accurate. The anchor mechanism of YOLOv3 is the same as that of YOLOv2. YOLOv4 [210] uses two anchors for one ground truth, while YOLOv3 uses only one anchor for one ground truth. With this, the problem of imbalance between positive and negative samples is alleviated. CIoU loss is adopted to solve the problems of MSE (mean squared error) loss, IoU loss, GIoU, and DIoU [211–214]. YOLOv4 also uses several techniques to achieve state-of-the-art results. YOLOv5 adopts adaptive anchors and uses the network to learn anchor parameters. Its detection head is the same as YOLOv3 and YOLOv4. It is slightly weaker than YOLOv4 in performance, but much faster than YOLOv4, and has strong advantages in the rapid deployment of the model.

SSD detection algorithm combines the regression idea with the anchor box (default frame) mechanism. It eliminates the candidate region generation and subsequent pixel or feature resampling stage (RoI pooling) in the two-stage algorithm. It encapsulates all calculations in one network, making it easy to train and very fast. RFBNet [215] and M2Det [216] are two successors of SSD. They use receptive fields and multi-level feature pyramid networks to improve the classical SSD, respectively.

Single-stage SAR ship detection algorithms can be divided into three categories: SAR image ship detection based on the YOLO series, SAR ship detection based on the SSD series, and other algorithms.

4.5.2. SAR Ship Detection Based on YOLO Series

YOLO series are widely used in this field. The improvements mainly focus on lightweight backbone network designing, multi-layer feature fusion, anchor box generation, multi-feature map prediction, loss function, etc.

YOLOv2. Deng et al. [33] and Chang et al. [39] adopted YOLOv2 to detect ships in SAR images. Paper [39] proposed YOLOv2-reduced which reduces some layers of YOLOv2. YOLOv2-reduced has an AP of 89.76% with 10.937 ms and 44.72 BFLOPS compared with YOLOv2, which has an AP of 90.05% with 25.767 ms and 50.17 BFLOPS.

YOLOv3. Zhang et al. [82] accelerated the original YOLOv3 by using DarkNet-19 as the backbone network. Additionally, it reduces the repeated YOLOv3-scale1, YOLOv3-scale2, and YOLOv3-scale3. Zhu et al. [116], Chaudhary et al. [123], and Jiang et al. [135] used the classical YOLOv3 with some techniques to detect ships in SAR images. Wang et al. [119] proposed SSS-YOLO which redesigned the feature extractor network to enhance the spatial and semantic information of small ships. It adopts a PAFN (path argumentation fusion network) to fuse different features in a top-down and bottom-up manner. SSS-YOLO has a better performance for small ships in SAR images. Hong et al. [158] improved the performance of YOLOv3 with some techniques. The improved clustering algorithm K-means++ generates an anchor box, which improves the performance of YOLOv3 for multi-scale ships. The Gaussian parameter for ship detection is introduced to add an uncertainty estimator for the positioning of the bounding box. Four anchor boxes are assigned to each detection scale instead of three in YOLOv3. Zhang et al. [40] used the idea of the YOLO algorithm, the input image meshes, and the depth separable convolution is used to improve the detection speed. MobileNet is used as the feature extractor to detect

ships under three scales: 13×13 , 26×26 , and 52×52 . The size of the anchor box can be obtained by the K-means algorithm. D-CNN-13 has a big receptive field with anchor box widths and heights of (9, 11), (11, 22) (14, 26). D-CNN-26 has a medium receptive field with anchor box widths and heights of (16, 40), (17, 12) (27, 57). D-CNN-52 has a small receptive field with anchor box widths and heights of (28, 17), (57, 28) (69, 72). Zhang et al. [54] used depth convolution and point convolution to replace the traditional convolution neural network, and adopt a multi-scale detection mechanism, concatenation mechanism, and anchor box mechanism to improve the detection speed. The detection network is composed of three parts, which means that it can detect an input SAR image under three different scales (5×5 , 10×10 , and 20×20), and then obtain the final ship detection results. It has nine anchor boxes for three detection scales, so it can detect up to nine ships in the same grid cell. Zhou et al. [102] designed a CNN named LiraNet, which has low complexity, few parameters, and a strong feature representation ability. LiraNet combines the idea of dense connections, residual connections, and group convolution, and it includes stem blocks and extractor modules. The network is the feature extractor of Libra-YOLO. Lira-YOLO has only 2.980 Bflops, and the parameter quantity is only 4.3 MB. It has good accuracy with less memory and computational cost compared with tiny-YOLOv3. In [151], DarkNet-53 with the residual unit is used as the backbone to extract features, and a top-down pyramid structure is added for multi-scale feature fusion. Soft NMS, mix-up, mosaic data augmentation, multi-scale training, and hybrid optimization are used to boost the performance. The 13×13 , 26×26 , 52×52 feature maps with the large, medium, and small receptive fields are responsible for large, medium, and small ships, respectively. The model is trained from scratch to avoid the learning objective bias of pre-training. The detection speed is fast, about 72 frames per second.

YOLOv4. Ma et al. [156] proposed YOLOv4-light, which is tailored to reduce the model size, detection time, number of computational parameters, and memory consumption. The three-channel images are used for compensating for the loss of accuracy. Liu et al. [181] proposed a detection method based on YOLOv4-Lit [217], whose backbone is MobileNetv2. A receptive field block is used for multi-scale target detection. It has an AP of 95.03% with 47.16 FPS and 49.34 M model size.

YOLOv5. Tang et al. [144] proposed N-YOLO based on YOLOv5. N-YOLO adopts a noise level classifier to classify the noise level of SAR images. SAR ship potential area extraction module is used to extract the complete region of potential ships. Zhou et al. [179] proposed a multi-scale ship detection network based on YOLOv5. It has the cross-stage partial network to improve feature representation capability, and the feature pyramid network with fusion coefficients module to fuse feature maps adaptively. It has a good tradeoff between model size and inference time.

Others. Zhang et al. [84] proposed ShipDeNet-20. It has only 20 convolution layers, and the model size is smaller than 1 MB, which is lighter than the other state-of-the-art detectors. ShipDeNet-20 is based on YOLO and is trained from scratch. Feature fusion module, feature enhance module, and scale share feature pyramid module are proposed to make up the accuracy loss of the raw ShipDeNet-20. It has a good tradeoff between accuracy and speed. Zhu et al. [175] proposed DB-YOLO. It is composed of a feature extraction network, duplicate bilateral feature pyramid network, and detection network. The single-stage network can meet the requirements of real-time detection, and it uses cross-stage partial to reduce redundant parameters. A duplicate bilateral feature pyramid network can enhance the fusion of semantic and spatial information. It alleviates the problem of small ship detection. CIoU loss is used as the loss function, as it has a faster convergence speed and better performance.

4.5.3. SAR Ship Detection Based on SSD Series

Wang et al. [14,18] directly used SSD and do not improve it. Papers [51,98,108] are the detection algorithms trained from scratch based on SSD. Most of the other papers improve the backbone network of SSD to make the model have a stronger feature extraction ability.

Chen et al. [15] adopted a two-stage regression network based on SSD to improve the performance of small ships, namely R2RN (robust two-stage regression network). R2RN connected an anchor modified module and object detection module to inherit the essence of the feature pyramid. Ma et al. [30] proposed an SSD model with multi-resolution input, which can extract richer features. Papers [43,44] applied the attention mechanism to SSD and design a new loss function based on GIoU. Li et al. [47] analyzed the reasons for the low detection accuracy of small and medium-sized ships in SSD and puts forward improvement strategies. Firstly, the anchor box optimization method based on K-means clustering is adopted to improve the matching performance of the anchor box. Secondly, a feature fusion method based on deconvolution is proposed to improve the representation ability of the underlying feature map. Chen et al. [55] adopted the attention mechanism and multi-level features to improve the feature extraction ability of the backbone network. Han et al. [99] used deconvolution to enhance the representation of small ships in the pyramid and improved the detection accuracy of SSD. Zhang et al. [113] token the original SAR image and saliency map as the input and fused the fusion of their features to reduce the computational complexity and network parameters. Chen et al. [114] proposed SSDv2, which adds a deconvolution module and prediction module on the basis of SSD to improve the detection accuracy. Jin et al. [149] improved SSD by feature fusion and squeeze-excitation module.

Sun et al. [162] proposed SANet (semantic attention-based network). It combines semantic attention, focal loss, label, and anchor assigning to improve the performance without increasing computation. Papers [104,159] adopted M2Det to detect ships in SAR images.

4.5.4. Others

RefineDet adopts a two-step cascade regression strategy to predict the position and size of objects. It can make the single-stage detectors obtain the accuracy of the two-stage detector without increasing computation. It is widely used in computer vision. Zhu et al. [159] adopted RefineDet to detect ships in SAR Images, which achieve an AP of 98.4%. In [169], GHM was used as the loss function of RefineDet, so that the network can make full use of all examples, and adaptively increase the weight of difficult cases. A multi-scale feature attention module is added to the network to highlight important information and suppress the interference caused by clutter. It achieves 96.61% precision on AIR-SARShip-1.0.

4.6. Anchor Free Detectors

4.6.1. Development of Anchor Free Detection Algorithm in Computer Vision

The anchor box is the key to the success of Faster R-CNN and SSD. The backbone network extracts features from the input image to obtain the feature map, and each pixel on the map is the anchor point. Taking each anchor point as the central point and artificially setting different scales and aspect ratios, multiple anchor boxes can be obtained. Anchor box has the following two advantages: firstly, it can generate dense candidate boxes, which is convenient for the network to classify and regress the targets. Secondly, it can improve the recall ability and is suitable for small target detection.

However, the anchor box needs to be designed manually by experience, which has the following defects: firstly, hyper-parameters need to be set, such as the number, size, aspect ratio, IoU threshold, etc. Secondly, in order to match the ground box, a large number of anchor boxes need to be generated, which are computationally intensive. Thirdly, most of them are invalid, which will lead to an imbalance between positive and negative samples. Fourthly, it is necessary to adjust the anchor box according to the size and shape distribution of the dataset.

The anchor-free detector opens up another idea by eliminating the predefined anchor box. It can directly predict several key points of the target from the feature map. For example, CornerNet, ExtremeNet [218], CenterNet [219], Objects as Points [220], FCOS (fully convolutional one-stage) [221] and FoveaBox [222].

The anchor-free detectors can avoid various problems and has great application potential in SAR ship detection. For example, due to the small size and sparse distribution of ships, most of the candidate anchor boxes are invalid negative samples. The anchor-free detectors can neglect the invalid anchors and reduce the amount of the predicted boxes, thus improving the accuracy and speed simultaneously. The anchor-free ship detectors in SAR images are shown in Figure 9.

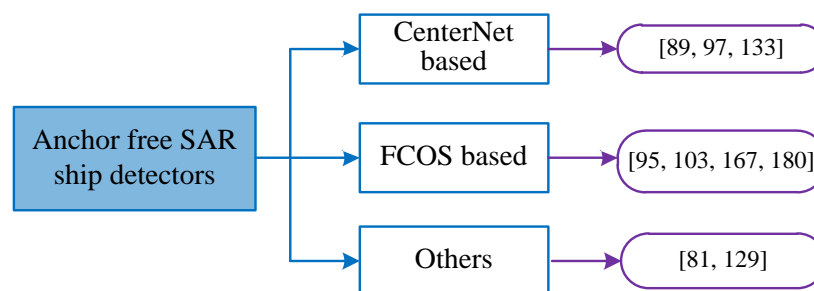


Figure 9. The anchor-free SAR ship detectors.

4.6.2. Development of Anchor-Free SAR Ship Detection Algorithm

Mao et al. [81] proposed a simplified U-Net [223] based anchor-free detector in SAR images. It includes ship bounding boxes regression network and score map regression network. The former is expected to be regressed based on each pixel in the input image. The latter is designed to predict a 2D probability distribution in which each score at each position indicates the likelihood of the current position in the center of any ship. Cui et al. [89] proposed a CenterNet (objects as points) based SAR ship detector. It predicts the center point of the target through key point estimation and uses the image information of the center point to obtain the size and position of the ship. There is no need to set anchors in advance and NMS is not needed, which greatly reduces the number of network parameters and calculations. Anchor mismatching of small ships is also reduced. Spatial shuffle-group enhance attention modules are used to extract features with more semantic information. Fu et al. [95] proposed an attention-guided balanced pyramid based on FCOS to improve the performance of small ships. Zhou et al. [97] proposed an anchor-free detector with dense attention feature aggregation. A lightweight feature extractor and dense attention feature aggregation are used to extract multi-scale features. A center-point-based ship predictor is used to regress the centers and sizes. There is no pre-set anchor and NMS, and thus the computational efficiency is high. Mao et al. [103] proposed a lightweight named ResSARNet with only 0.69 M parameters, and improved FCOS in four aspects: center-ness in bounding box regression branch, not in classification regression branch, center sampling, GloU loss, and adaptive training sample selection. The network only needs 1.17 M parameters and can achieve 61.5% AP and 70.9% AR. An et al. [129] designed an anchor-free rotatable detector. It designs center point-scale and angle prediction to convert the conventional rotatable prior box mechanism into the center point-scale and angle prediction. The training procedure includes positive sample selection, feature encoding, and loss function designing. Wang et al. [133] proposed a CenterNet-based detector in SAR images. The spatial group-wise enhanced attention module is used to extract more semantic features.

Sun et al. [167] proposed category-position FCOS. The category-position module is used to optimize the position regression branch in the FCOS network. The classification and regression branches are redesigned to alleviate the imbalance between positive and negative samples during training. Zhu et al. [180] adopted FCOS as the base model to reduce the effect of anchors. A new sample definition method is used to replace the IoU threshold according to the differences between SAR images and natural images. The same resolution feature convolution module, multi-resolution feature fusion module, and feature pyramid module are used to extract features. The focal loss and CIoU are used to improve the performance further.

In all, researchers in SAR ship detection realize the benefit of anchor-free detectors. Additionally, more and more papers are appearing in this field. However, there is still a problem: the innovation is relatively weak and some of the existing achievements of computer vision are not used in this field.

4.7. Detectors Trained from Scratch

At present, most SAR image detector backbone needs to pre-train on the classification dataset of natural images, and then fine-tune on the ship detection dataset of SAR images (for example SSDD). This transfer learning can make the detection algorithm initialize better and make up for the problem of insufficient samples. However, there will be the following problems: firstly, there is learning bias. The loss function and category distribution between classification and detection are contradictory in essence. The models trained on classification are not fit for detection. Secondly, most backbone networks will produce a high receptive field through multiple down sampling in the latter layers, which is good for the classification but is harmful for the location. Thirdly, the pre-trained backbone networks are redundant and cannot be modified, which hinders the researchers to design CNN flexibly according to their needs.

In order to solve the problems of transfer learning, algorithms trained from scratch are proposed in computer vision, for example, DSOD (deeply supervised object detectors), DetNet, ScratchDet, and so on [224–227].

The main idea of DSOD and GRP-DSOD to realize training from scratch is by designing the backbone and front-end network elaborately [224]. DetNet [225] retains a large scale in the last few layers, which can have more location information. ScratchDet [226] proposes to adopt the strategy of batch normalization in each layer and increase the learning rate, which can make the detection algorithm more robust and converge faster. Paper [227] replaced the original BN (batch normalization) with group normalization (GN) and asynchronous BN, so as to make the parameters of gradient normalization more accurate. This then made the descending direction of gradient more accurate, so as to accelerate convergence and improve the accuracy.

The model trained from scratch not only has high accuracy but also greatly reduces the size and amount of calculation of the model. Due to the above advantages, it is also used in SAR ship detection.

Most detectors that are trained from scratch in this field have well-designed networks. They are shown in Figure 10.

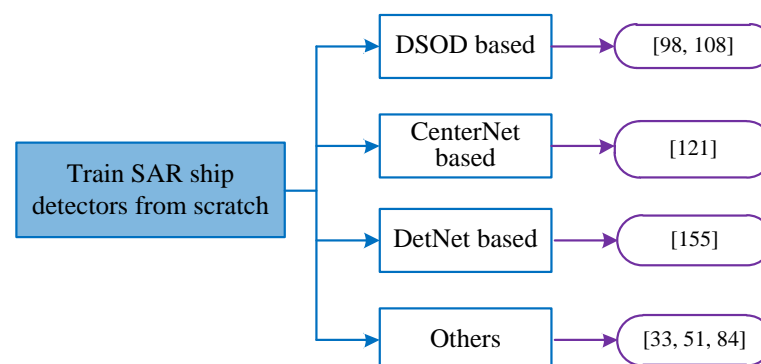


Figure 10. The detectors trained from scratch in SAR images.

Deng et al. [33] designed a dense backbone network composed of multiple dense blocks. The front layer can receive additional supervision from the objective function through dense connections, which makes the training easier, and adopts the feature reuse strategy to make the parameter highly efficient. Zhang et al. [51] designed a lightweight detection algorithm that can be trained from scratch, which can reduce the training and testing time without reducing the accuracy. It adopts the modules of semantic aggregation and feature reusing to improve the performance of multi-scale ships. Zhang et al. [84]

proposed a lightweight detection network ShipDeNet-20. It is designed with fewer layers and convolution kernels and depth separable convolution. It also adopts a feature fusion module, feature enhancement module, and proportionally shared feature pyramid module to improve detection accuracy. Han et al. [98] integrated the lightweight asymmetric square convolution block into SSD to realize training from scratch, and its accuracy and speed are better than the classical DSOD. Han et al. [100] proposed a parallel convolution block of multi-scale kernel and feature reusing convolution module to enhance feature representation and reduce information loss. Han et al. [108] designed two kinds of asymmetric convolution blocks: asymmetric and square convolution feature aggregation block, and asymmetric and square convolution feature fusion block. They replace all 3×3 convolution layers, which are embedded into the classic DSOD to achieve a better result of the training from scratch. Guo et al. [121] proposed an effective and stable single-stage algorithm that is trained from scratch, namely CenterNet++. The model mainly includes three modules: feature matching module, feature pyramid fusion module, and head enhancement module. Zhao et al. [155] used DetNet as the backbone network to realize training from scratch. It uses superposition convolution instead of down sampling to solve the problem of small ship detection and adopts a feature reusing strategy to improve parameter efficiency.

Compared with other directions, fewer researchers in SAR ship detection realize the benefit of training from scratch. Additionally, the papers using training from scratch techniques in this field are not advanced enough. We should adopt more advanced techniques in computer vision in this direction. In all, the detectors trained from scratch are not used to their full extent in this field. Some good conclusions in papers [226,227] should be considered and applied here.

4.8. Detectors with Oriented Bounding Box

The oriented bounding box was originally used in scene text detection. In addition, a large number of achievements have emerged, such as SegLink, RRPN (rotation region proposal network), TextBoxes, TextBoxes++, R2CNN (rotational region convolutional neural network), and so on [228–232]. The ships in remote sensing images also have multi-directional characteristics. The conventional vertical rectangular bounding box often cannot accurately surround the target. With the improvement of ship detection accuracy, the use of oriented bounding boxes to realize multi-directional ship detection has become a research hotspot [233–238]. DOTA (dataset for object detection in aerial images) is a commonly used aerial image target detection dataset in this field, which can be used to develop and evaluate the performance of detection algorithms. Similarly, there are many detection algorithms based on oriented bounding boxes in SAR images, which will be introduced here. At present, the dataset that can be used to train and test the oriented bounding box algorithm are SSDD+, RSDD-SAR, and SRSD-10, the details have been introduced earlier. The oriented bounding box detectors in SAR images are shown in Figure 11.

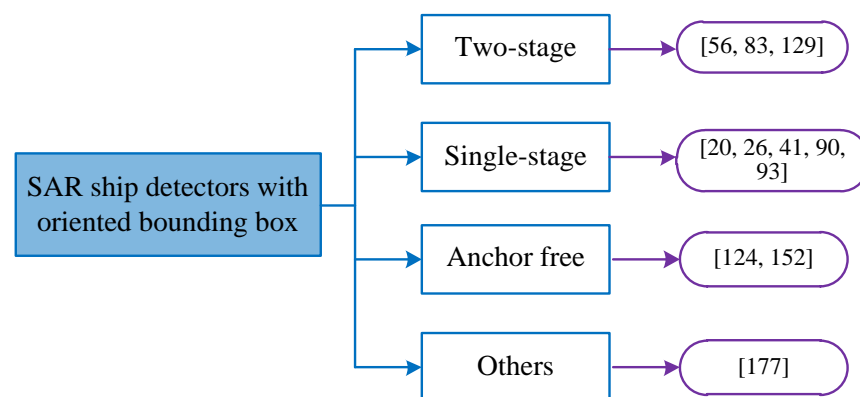


Figure 11. The SAR ship detectors with oriented bounding box.

Two-stage. Chen et al. [56] proposed a multi-scale adaptive recalibration network to detect multi-scale and arbitrarily oriented ships. It can learn the angle information of ships. The anchors, NMS, and loss function are also redesigned to fit the large aspect ratio and arbitrary directionality of ships in SAR images. Pan et al. [83] proposed a multi-stage rotational region-based network (MSR2N) to solve the problem of redundancy regions. MSR2N includes FPN, RRPN, and a multi-stage rotational detection network. It is more suitable and robust for SAR ship detection. An et al. [129] adopted an oriented detector as the based model to solve the problem that conventional CNN models have too many parameters, which increases the difficulty of transfer learning between different tasks.

Single-stage. Wang et al. [20] proposed a SAR ship detector with an oriented bounding box based on SSD. The detector can predict the class, location, and angle information of ships. The semantic aggregation module is used to capture abundant location and semantic information. The attention module is used to adaptively select meaningful features and neglect weak ones. Multi-orientation anchors, angular regression, and the loss function are used to fit the oriented bounding box. Liu et al. [26] adopted DR-Box [239] to detect ships in SAR images. DR-Box is specially designed to detect targets in any direction in remote sensing images. It can effectively reduce the interference of background pixels and locate the target more accurately. An et al. [41] proposed DR-Box-v2 to detect ships in SAR images. A multi-layer anchor box generation strategy for detecting small ships is proposed. A modified encoding scheme is proposed to estimate the position and orientation precisely. Focal loss and hard negative mining are also used to balance the positives and negatives. Yang et al. [90] regarded a rotatable bounding box detector as the base model to solve the problem of negative sample intra-class imbalance in the training stage. Chen et al. [93] proposed a rotated refined feature alignment detector to fit ships with large aspect ratios, arbitrary orientations, and dense distribution properties. A lightweight attention module, modified anchor mechanism, and feature-guided alignment module are proposed to boost the performance of the oriented detector.

Anchor free. Yang et al. [124] proposed R-RetinaNet to beat DRBox-v1, DRBox-v2, and MSR2N (multi-stage rotational region-based network) in this field. R-RetinaNet used a scale calibration method to align the scale distribution. Task-wise attention feature pyramid network is used to alleviate the contradiction of classification and localization. The adaptive IoU threshold training method is used to correct the unbalanced problem. He et al. [152] proposed a method to solve the problem of boundary discontinuity problem in oriented bounding box detectors by learning polar encodings. The encoding scheme uses a group of vectors pointing from the center of the ship to the boundary points to represent an oriented bounding box.

Others. Ding et al. [177] released the SRSDD-v1.0 dataset, which is used for oriented bounding box detectors. The details of the dataset have been described above. They present the performance of several advanced oriented bounding box detection algorithms on the dataset.

Summary. With the emergence of several datasets with oriented bounding boxes, ship detectors in SAR images based on oriented bounding boxes are becoming more and more advanced. However, it is not enough compared with DOTA. Some efforts should be taken in this direction.

4.9. Multi-Scale Ship Detectors

In MS COCO, the proportions of the small, medium, and large size objects are 41.43%, 34.32%, and 24.24%, respectively. However, in SAR images, the proportion of small-size ships is extremely high. For example, the proportion of small, medium, and large ships in the RSDD-SAR dataset are 81.175%, 18.776%, and 0.049%, respectively. In LS-SSDD-v1.0, the proportions are 99.80%, 0.20%, and 0.00%, respectively. Therefore, this field needs to focus on the problem of multi-scale ship detection, especially the small ships.

Although CNN has developed rapidly in computer vision, it has poor performance on small-size object detection. In order to improve the adaptability to multi-scale ships,

computer vision often fuses low-level and high-level features (such as FPN), increases the receptive field and improves the anchor box generation and matching strategy.

SAR ship detection also uses the above methods to improve the performance of multi-scale ship detection. The multi-scale ship detectors in SAR images are shown in Figure 12.

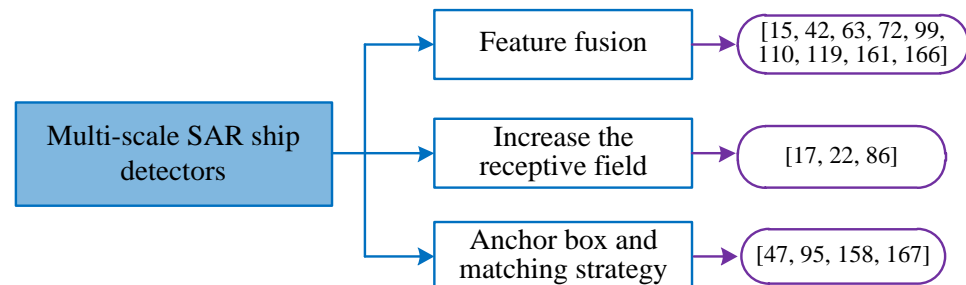


Figure 12. The multi-scale SAR ship detectors.

Feature fusion. Chen et al. [15] proposed a densely connected multi-scale neural network to solve the problem of multi-scale SAR ship detection. It closely connects each feature map with other feature maps from top-to-bottom and generates proposals from each fused feature map. Cui et al. [42] proposed a dense attention pyramid network, which closely connects the convolutional attention module from the top to the bottom of the pyramid network to each feature map, so as to extract rich features containing location information and semantic information for adapting to multi-scale ships. Liu et al. [63] proposed a scale transferable pyramid network to adapt to the detection of multi-scale ships. It constructs a feature pyramid network through horizontal connection and uses a scale transfer layer to closely connect each feature graph from top to bottom. A horizontal connection introduces more semantic information, and a dense scale transfer connection can expand the resolution of the feature map. Jin et al. [72] combined all feature maps from top to bottom to make use of contextual semantic information at all scales and uses extended convolution to increase the receptive field exponentially. Han et al. [99] used deconvolution to enhance the feature representation of small and medium-sized ships in FPN, so as to improve the detection accuracy of SSD. Hu et al. [110] proposed a dense feature pyramid network, which processes shallow features and deep features differently. Compared with traditional FPN, it has stronger adaptability to multi-scale ships. Wang et al. [119] proposed a path argumentation fusion network to fuse different feature maps. It uses bottom-up and top-down methods to fuse more location information and semantic information. Hu et al. [161] proposed a two-way revolution network based on a bidirectional convolution structure, which can effectively process shallow and deep feature information and avoid the loss of small ship information. Zhang et al. [166] proposed a quad feature pyramid network to detect multi-scale ships. It includes deformable convolutional FPN, a content-aware feature reassembly FPN, a path aggregation space attention FPN, and a Balance Scale Global Attention FPN.

Increase the receptive field. Deng et al. [17] designed a feature extractor with multiple receptive fields through ReLU and inception modules. It generates candidate regions in multiple middle layers to match ships of different scales and fuses multiple feature maps so that small-scale ships have a stronger response. Zhao et al. [22] proposed a coupled CNN to detect small-scale ships. It includes a network that generates candidate areas from multiple receptive fields and improves the recognition accuracy by using the context information of each candidate box. Dai et al. [86] did not use a single feature map but fused the feature map in a bottom-up and top-down manner, and generated candidate boxes from each fused feature map.

Anchor box generation and matching strategy. Li et al. [47] first analyzed the reasons for the low detection accuracy of small and medium-sized ships in SSD and made some improvements. The anchor box optimization method based on K-means clustering solves the problem of less positive samples and more negative samples. The feature fusion

method based on deconvolution improves the representation ability of the low-level feature map and solves the weak recognition ability of the low-level feature map to small ships. Fu et al. [95] proposed a feature balance and matching network, which uses the anchor-free strategy to eliminate the influence of anchors and uses the attention-guided balance pyramid to balance multiple features at different levels semantically. It has a good performance in the detection of small-scale ships. Hong et al. [158] improved the anchor generation based on an improved K-means++ in YOLOv3. It alleviates the difficulty of multi-scale ship detection in YOLOv3 and changes the number of anchor boxes in the YOLO layer from three to four. Sun et al. [167] show that anchor-free detectors have good adaptability to small ships and have a fast speed.

Summary. Small ship detection is extremely hard but is also extremely important for some applications. That is because people hope to find targets within a long distance, and at this point, the targets must be small in size. SAR ship detection also proves this point. Although the above detection methods for small-size ships have certain effects, they are still far from enough. Innovative work needs to be continued.

4.10. Attention Module

The basic idea of the attention mechanism in computer vision is to make the model ignore irrelevant information and focus on key information. It can be divided into hard attention, soft attention, gaussian attention, spatial transformation, and so on. Attention can be calculated from the spatial domain, channel domain, layer domain, and mixed domain. Representative algorithms include SENet (squeeze and excitation network), SKNet (selective kernel network), CBAM (convolutional block attention module), CCNet (criss-cross attention), OCNNet (object context network), DANet (dual attention network), etc. [240–244]. Transformer [245] adopted encoder–decoder architecture, which is the extreme of the attention. It abandons CNN and RNN (recurrent neural network) used in previous deep learning tasks and shows great advantages in the field of NLP (natural language processing) and CV. Swin Transformer [246] makes it compatible with image classification and object detection. It demonstrates the potential of transformer-based models as vision backbones.

Chen et al. [43,44] proposed an attention-based detector. The attention model is mainly composed of the convolution branch and mask branch. Elements in mask maps are similar to the weight of feature maps, which enhance regions of interest and suppress non-target regions. Cui et al. [89] introduced the space shuffle group enhanced attention module to CenterNet. It can extract stronger semantic features and suppress some noise at the same time, so as to reduce false positives caused by inshore and inland interference. Zhao et al. [91] combined the receptive field module and convolution block attention module to construct a top-down fine-grained feature pyramid. Wang et al. [122] designed a feature enhancement module based on a self-attention mechanism. Its spatial attention and channel attention work at the same time to highlight the target and suppress the spot to a certain extent. Wang et al. [131] embedded a soft attention module in the network to suppress the influence of noise and complex background. Zhu et al. [136] proposed a SAR ship detection method based on a hierarchical attention mechanism. The method includes a global attention module and a local attention module. Hierarchical attention strategies are proposed from the image layer and target layer, respectively. Sun et al. [162] introduced a semantic attention mechanism, which highlights the regional characteristics of ships and enhanced the classification ability of the detector. Du et al. [169] embedded the multi-scale feature attention module in the network. By applying the channel and spatial attention mechanism to the multi-scale feature map, it can highlight important information and suppress the interference caused by clutter.

CRTransSar [182] is the first to use a transformer for SAR image ship detection. It is based on Swin Transformer and shows great advantages. CRTransSar combines the global contextual information perception of transformers and the local feature representation capabilities of convolutional neural networks. It innovatively proposes a visual transformer

framework based on contextual joint-representation learning. Experiments on SSDD and SMCDD show the effectiveness of the method.

4.11. Real-Time Detectors

At present, deep learning-based detectors need large computation and storage resources, which hinders the application in real-time prediction. In order to solve this problem, there are a lot of acceleration ideas in the evolution of object detection algorithms. Firstly, researchers usually speed up the detection process. This idea is reflected in the evolution process of R-CNN, Fast R-CNN, Faster R-CNN, R-FCN, and Light-Head R-CNN. The above detectors share the features gradually, and the network structures become thinner and faster. Secondly, researchers usually design lightweight detection networks. The backbone network and the detection head can both be accelerated. Thirdly, researchers usually compress and accelerate CNN models. It includes lightweight neural network designing, model pruning, model quantization, and knowledge distillation [247–253].

The exploration of real-time detection algorithms in SAR ship detection can be divided into three directions, which are shown in Figure 13.

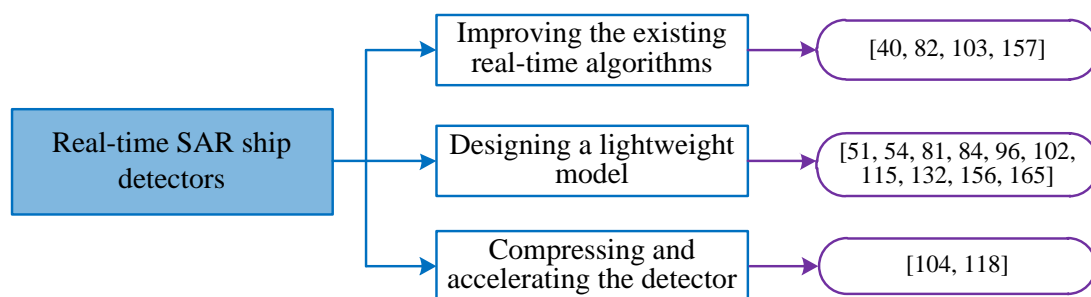


Figure 13. The real-time SAR ship detectors.

4.11.1. Improving the Existing Real-Time Algorithms

Many improvements in this field are based on the YOLO and SSD series, because they have great advantages in running time, especially the YOLO series. Zhang et al. [40] used the idea of the YOLO algorithm and adopted depth separable convolution to accelerate the speed. MobileNet is used as the backbone network to improve the detection speed under the condition of ensuring detection accuracy. Zhang et al. [82] proposed an improved YOLOv3 (using DarkNet-19 and deleting repeated layers). It achieved 90.08% AP₅₀ and 68.1% AP on the SSDD dataset. Mao et al. [103] adopted the FCOS detection algorithm with ResSARNet as the backbone network, and center-ness on bounding box regression branch, center sampling, GIoU loss, and adaptive training sample selection were used. It can achieve 61.5% AP with only 1.17 M parameters. Zhong et al. [157] adopted CFAR and YOLOv4 to realize real-time ship detection on China HISEA-1 SAR images.

4.11.2. Designing a Lightweight Model

Zhang et al. [51] designed a lightweight feature optimization network LFO-Net based on SSD. It can be trained from scratch and reduce the training and testing time without reducing the accuracy. The detection performance is further improved by the bidirectional feature fusion module and attention mechanism. It achieved 80.12% AP₅₀ with 9.28 ms testing time on SSDD. Zhang et al. [54] used multi-scale detection, cascade, and anchor box mechanism to design a lightweight network for real-time SAR ship detection. It uses depthwise and pointwise to replace the traditional convolution. It achieved 94.13% AP₅₀ with 9.03 ms testing time on SSDD. Mao et al. [81] used the simplified U-Net as the feature extraction network, which has only 0.47 million learnable weights, it improves the operation speed and solves the problem caused by the anchor box through the anchor-free method. It has a total of 0.93 million learnable weights, and the AP on the SSDD dataset is 68.1%. Zhang et al. [84] proposed ShipDeNet-20, which has 20 convolution layers and a

0.82 MB model size. It uses fewer layers and kernels, and depthwise separable convolution is also used. It improves the accuracy through the feature fusion module, feature enhancement module, and scale share feature pyramid module. It achieved 97.07% AP₅₀ with 233 FPS on SSDD. Zhang et al. [96] proposed HyperLiNet. It realizes high precision through five modules, namely multi receptive field module, divided revolution module, channel and spatial attention module, feature fusion module, and feature pyramid module. It realizes high speed through five modules, namely region-free model, small kernel, narrow channel, separate revolution, and batch normalization fusion. Zhou et al. [102] proposed a lightweight detector Lira YOLO. It combines the idea of dense connections, residual connections, and group convolution, including stem blocks and extractor modules. It achieved 85.46% AP₅₀ with a 4.3 MB model size. Li et al. [115] designed a lightweight network of feature relay amplification and multi-scale feature jump connection structure based on Faster R-CNN and improves the selection of anchor boxes and RoI pooling. It achieved 89.8% AP₅₀ and the speed increased a lot. Zhang et al. [132] proposed a lightweight detection algorithm ShipDeNet-18, which has fewer layers and fewer convolution kernels. The deep and shallow feature fusion module and a feature pyramid module are adopted to improve the detection accuracy. It achieved 93.78% AP₅₀ with 202 FPS. Ma et al. [156] proposed YOLOv4-tiny. It reduces the number of convolutional layers in CSPDarkNet53. It achieves 88.08% AP₅₀ with 12.25 ms compared with YOLOv4 with 96.32% AP and 44.21 ms. Sun et al. [165] proposed a lightweight densely connected sparsely activated detector. It can construct a lightweight backbone network, so as to achieve a balance between performance and computational complexity. It achieved 97.2% AP₅₀ and 61.5% AP on SSDD.

4.11.3. Compressing and Accelerating the Detector

Mao et al. [104] proposed a knowledge distillation-based network slimming method. YOLOv3 and Darknet-53 are pruned on filter-level to obtain lightweight models. Kullback Leibler Divergence (KLD) knowledge distillation is used to train student and teacher networks (YOLOv3@EfficientNet-B7). The model has only 15.4 M parameters, and the AP decreases by only 1%. Chen et al. [118] proposed the algorithm of Tiny-YOLO-Lite. It designs and prunes the backbone structure, strengthens the channel level sparsity, and uses knowledge partition to make up for the performance degradation caused by pruning. Tiny-YOLO-Lite reduces the size of the model, reduces the number of floating-point operations, and obtains faster accuracy.

4.11.4. Summary

From the above discussion, we can find that real-time ship detection is also a hot topic in SAR images. However, the above works are not enough. It is obvious that the transferred deep learning models from computer vision are abundant in this field. Researchers should do the following work to realize real-time detection. Firstly, the anchor-free and the training from scratch method should be used to design lightweight detection algorithms. Secondly, some model compressing, and accelerating techniques should be used to improve the speed further. Thirdly, the lightweight models should be transplanted to high-performance AI chips (NVIDIA Jetson TX2) to achieve the purpose of running at the edge (satellite, airplane).

4.12. Other Detectors

In this part, we mainly introduce weakly supervised, GAN (generative adversarial network) and data augmentation, which are shown in Figure 14.

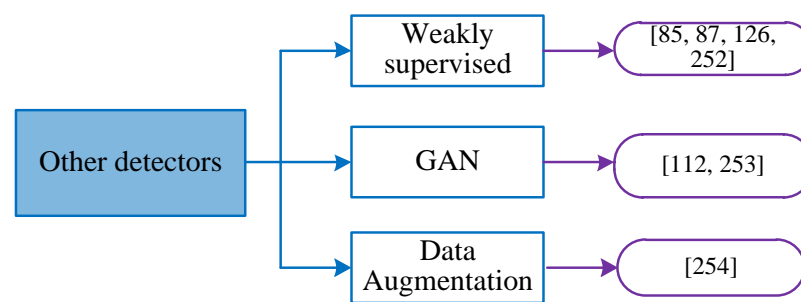


Figure 14. The other SAR ship detectors.

4.12.1. Weakly Supervised

The supervised methods, such as deep learning approaches, need substantial time and manpower to make training samples [254]. Papers [85,87,126] adopted weakly supervised to train ship detection algorithms. The model is trained by two global labels, namely, “ship” and “non-ship,” and produces a ship location heatmap, ship bounding box, and pixel-level segmentation product. They can alleviate the problem of annotation partly. However, the accuracy is lower than the supervised method.

4.12.2. GAN

The insufficient SAR samples restrict the performance of the algorithm. Zou et al. [112] used a multi-scale Wasserstein auxiliary classifier generative adversarial network [255] to generate high-resolution SAR ship images. Then, the original dataset and the generated data are combined into a composite dataset to train the YOLOv3 network, so as to solve the problem of low detection accuracy under a small dataset. Based on the idea of generative adversarial networks, an image enhancement module driven by target features is designed. The quality of the ships in the image is improved. The experimental results verify the effectiveness of this method.

4.12.3. Data Augmentation

Data augmentation can expand the size of the dataset several times, so as to improve the detection accuracy [256]. The training method based on a feature mapping mask eliminates the gradient noise introduced by random clipping, so as to improve the detection performance. The SAR images with ships are generated by electromagnetic numerical analysis technology, and the sea clutter model is used to simulate the real SAR image patch containing various SAR slices, so as to improve the performance of SSD.

4.13. Problems

From the 177 papers, we can see that most of the detection algorithms in this field are borrowed from computer vision. Additionally, its development is also behind the detectors in computer vision. Due to the large difference between natural image and SAR image (for example, SAR image is single-channel, ship size is small, and distribution is very sparse), some detection algorithms are not suitable for SAR ship detection. So, we should design detectors according to the real characteristics of ships in SAR images.

The 177 papers mainly use the image essences of SAR images, and the research and application of the scattering mechanism are not enough. This is one problem we should solve in the future.

At present, there are several public small datasets, but we lack a large dataset. The models trained on a small dataset face the problem of over-fitting. What we should do next is merge the small datasets into a large one, and make sure the train-test division standards, evaluation indicators, and benchmarks are clear. These works can promote the development of this field.

5. Future—The Direction of the Deep Learning-Based SAR Ship Detectors

5.1. Anchor Free Detector Deserves Special Attention

The anchor-free detection algorithm has many advantages, which have been introduced in Section 4.6. It should be emphasized that the detection algorithm without an anchor box is especially suitable for SAR images. As SAR images have sparse and small size ships, it can greatly improve the detection speed and avoid various problems in anchor box designing and matching. Therefore, the anchor-free detection algorithm needs to be paid more attention. Fortunately, researchers in this field have realized this and many research results have emerged.

5.2. Train Detector from Scratch Deserves More Attention

At present, there are the following generally accepted conclusions about training from scratch: firstly, pre-training accelerates the convergence speed, especially in the early stage of training. However, the training time of scratch is roughly equivalent to the total time of pre-training and fine-tuning. Secondly, if there are enough target images and computing resources, pre-training is not necessary. Thirdly, if the cost of image collection and image cleaning is considered, a general large-scale classification dataset is not an ideal choice, and collecting images on detection tasks will be a more effective approach. Fourthly, when the target task is to predict spatial positioning (such as ship detection), pre-training does not show any benefits.

Collecting images for detection and training is a solution worth considering, especially when there is a significant gap between the pre-training task and the detection task (such as ImageNet image and SAR image). Therefore, in the field of SAR ship detection, it is very necessary to combine the existing public datasets into a large dataset, so as to ensure training models from scratch.

Due to the difference between natural images and SAR images, it is very necessary to adopt training from scratch detection algorithms in this field, as it can obtain a detection algorithm with stronger adaptability to SAR images and smaller model size. However, the work at this stage is far from enough, so we need to pay more attention to the detection algorithms of training from scratch.

5.3. Many Other Works Need to Be Used for Oriented Bounding Box Detector

The ship in the SAR image has very changeable directionality, and the vertical bounding box cannot adapt to this scene. It is necessary to use an oriented bounding box. In an inshore scenario, a vertical bounding box is susceptible to interference from onshore buildings and other ships, affecting detection performance, while an oriented bounding box can accurately represent the ship target and reduce redundant interference. In addition, for ship targets in an offshore scenario, an oriented bounding box can obtain information such as heading and aspect ratio, which is of great significance for subsequent trajectory prediction and situation estimation tasks. The scene text detection and the aerial remote sensing image dataset DOTA have conducted in-depth research on the oriented bounding box and achieved many results. We should learn from them.

5.4. Small Ship Detection Is an Eternal Topic

The main reasons for the poor detection result of small-size ships are as follows: firstly, the features extracted from small-scale ships are few, and the size and receptive field of the anchor are too large for small ships. Secondly, the size of the anchor is discrete (for example, 16, 32, 64, etc.), while the size of the ship is continuous, which makes the recall rate of small-size ships low. Thirdly, the anchor of a small ship matches less with the ground truth bounding box, resulting in fewer positive samples and too many negative samples.

As the proportion of small ships in SAR images is very high, small object detection is difficult, especially in this field. So, it is an eternal topic to study how to improve the detection effect of small ships.

5.5. Real-Time Detection Is the Key to Application

Real-time SAR ship detection needs to start from many aspects. For example, we can design a lightweight detection network, compress and accelerate the model to improve the speed, and transplant the detection algorithm to high-performance AI chips at the edge (NVIDIA Jetson TX2). At present, most of the work in this field is focused on the first two aspects, and there is less research on the third aspect, which needs to be focused on in the future. Only by realizing this technology can we realize the real-time detection and recognition of ships on satellite or aircraft platforms.

5.6. Transformer Is the Future Trend

In the past two years, transformer shows great advantage in object detection compared with CNN, for example, DETR (detection transformer) [257] and Swin Transformer. DINO [258] can achieve 63.6% AP on the COCO test-dev, which surpasses CNN-based detector by a large margin. Nowadays, the hot topic of computer vision is the transformer. CRTransSar [182] is the first to use a transformer for SAR image ship detection. It shows a great advantage in accuracy (97% AP on SSDD). Although there are still some problems when a transformer is used for SAR ship detection, there is no doubt that the transformer will be the research trend in the future due to its great advantages.

5.7. Bridging the Gap between SAR Ship Detection and Computer Vision

Compared with the field of computer vision, the field of ship detection in SAR images is relatively small and not active enough. Therefore, it is necessary to bring this field to computer vision, and systematically learn from the rich achievements in computer vision.

We should also learn about its openness, standardized evaluation, and easily accessed codes. This work can promote this field to develop rapidly. What we should do is as follows: firstly, the existing public datasets (SSDD, SAR-Ship-Dataset, AIR-SARShip, HRSID, LS-SSDD-v1.0, SRSDD-v1.0, and RSDD-SAR) need to be combined into a large dataset, which can be called LargeSARDataset here. The modes trained on it can avoid over-fitting. Secondly, determine the training samples and testing samples. Thirdly, determine the evaluation indicators. Fourthly, release the benchmark. Fifthly, bring it into the field of computer vision. As shown in Figure 15. Through this work, we can bridge the gap between SAR ship detection and computer vision.

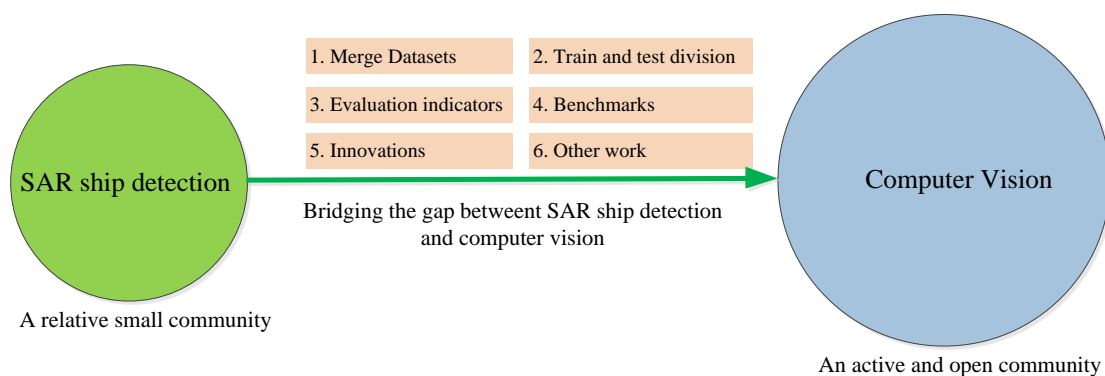


Figure 15. Bridging the gap between SAR ship detection and computer vision.

In addition to detection, classification and segmentation of SAR images also enter into the deep learning era [259–265]. Classification and segmentation algorithms borrowed from computer vision are extensively used in SAR images. We will review them in the future. In the process of detection, only ship and non-ship targets are considered, and the specific content of non-ship targets is not analyzed [266–268]. Some icebergs have great similarities in shape and size with ships, and the algorithms are difficult to distinguish them. So, we will study how to solve this problem in the future.

6. Conclusions

This paper introduces the past, present, and future of deep learning-based ship detection algorithms in SAR images.

Firstly, the history of SAR ship detection is reviewed (before SSD was public on 1 December 2017). This part mainly introduces the detection algorithm based on CFAR and analyzes the great advantages of deep learning-based algorithms. In addition, they are compared in theory and experiment.

After that, there is a comprehensive overview of the current (from 1 December 2017 to now) ship detection algorithms based on deep learning. This part first analyzes the datasets, country, timeline, deep learning framework, and the performance evolution of the 177 papers. The basic situation of 10 datasets in this field is introduced especially. The 177 papers were classified, and they are two-stage, single-stage, anchor free, train from scratch, oriented bounding box, multi-scale, attention model, real-time detection, and so on. The specific algorithms in those papers are analyzed, including the principle, innovation, performance, and the summary.

Finally, the problems existing in this field and the future development direction are described. The main ideas are to design the detection algorithm according to the specific characteristics of SAR image, focus on the detection algorithm without an anchor box, pay enough attention to the detection algorithm of training from scratch, and learn from the existing achievements of natural scene text detection and DOTA, improve the performance of small ships continuously, pay attention to realize the real-time detection of ships through model acceleration and AI chip. It is emphasized that the future important work is to bridge the gap between SAR ship detection and computer vision by merging the existing small datasets into a larger one and making relevant standards.

This review can provide a reference for researchers in this field or researchers interested in this field so that they can quickly understand the current situation and future development direction of this field.

Author Contributions: Conceptualization, J.L. and C.X.; methodology, H.S., L.G. and T.W.; investigation, J.L.; writing—original draft preparation, J.L.; writing—review and editing, J.L. and C.X.; supervision, C.X.; funding acquisition, C.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, No. 61790550, No. 61790554, No. 61971432, No. 62022092.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

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

References

1. Moreira, A.; Prats-Iraola, P.; Younis, M.; Krieger, G.; Hajnsek, I.; Papathanassiou, K.P. A tutorial on synthetic aperture radar. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–43. [[CrossRef](#)]
2. Reigber, A.; Scheiber, R.; Jager, M.; Prats-Iraola, P.; Hajnsek, I.; Jagdhuber, T.; Papathanassiou, K.P.; Nannini, M.; Aguilera, E.; Baumgartner, S.; et al. Very-high-resolution airborne synthetic aperture radar imaging: Signal processing and applications. *Proc. IEEE* **2013**, *101*, 759–783. [[CrossRef](#)]
3. Li, H.; Hong, W.; Wu, Y.; Fan, P. An efficient and flexible statistical model based on generalized Gamma distribution for amplitude SAR images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2711–2722. [[CrossRef](#)]
4. Achim, A.; Kuruoglu, E.E.; Zerubia, J. SAR image filtering based on the heavy-tailed Rayleigh model. *IEEE Trans. Image Process.* **2006**, *15*, 2686–2693. [[CrossRef](#)]
5. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
6. Schwegmann, C.P.; Kleynhans, W.; Salmon, B.P.; Mdakane, L.W.; Meyer, R.G.V. Very deep learning for ship discrimination in Synthetic Aperture Radar imagery. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 104–107. [[CrossRef](#)]
7. Miao, K.; Leng, X.; Zhao, L.; Ji, K. A modified faster R-CNN based on CFAR algorithm for SAR ship detection. In Proceedings of the 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), Shanghai, China, 18–21 May 2017. [[CrossRef](#)]

8. Liu, Y.; Zhang, M.H.; Xu, P.; Guo, Z. SAR ship detection using sea-land segmentation-based convolutional neural network. In Proceedings of the 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), Shanghai, China, 18–21 May 2017. [[CrossRef](#)]
9. 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](#)]
10. Wang, Y.; Chao, W.; Hong, Z. Combining single shot multibox detector with transfer learning for ship detection using Sentinel-1 images. In Proceedings of the 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA), Singapore, 19–22 November 2017.
11. Li, J.; Qu, C.; Shao, J. Ship detection in SAR images based on an improved faster R-CNN. In Proceedings of the Sar in Big Data Era: Models, Methods & Applications, Beijing, China, 13–14 November 2017. [[CrossRef](#)]
12. Cozzolino, D.; Martino, G.D.; Poggi, G.; Verdoliva, L. A fully convolutional neural network for low-complexity single-stage ship detection in Sentinel-1 SAR images. In Proceedings of the Geoscience & Remote Sensing Symposium, Fort Worth, TX, USA, 23–28 July 2017. [[CrossRef](#)]
13. 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](#)]
14. Wang, Y.; Wang, C.; Zhang, H.; Zhang, C.; Fu, Q. Combining Single Shot Multibox Detector with transfer learning for ship detection using Chinese Gaofen-3 images. In Proceedings of the 2017 Progress in Electromagnetics Research Symposium-Fall (PIERS-FALL), Singapore, 19–22 November 2017. [[CrossRef](#)]
15. Chen, S.Q.; Zhan, R.H.; Zhang, J. Robust single stage detector based on two-stage regression for SAR ship detection. In Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence, Shanghai China, 9–12 March 2018; pp. 169–174. [[CrossRef](#)]
16. Jiao, J.; Zhang, Y.; Sun, H.; Yang, X.; Gao, X.; Hong, W.; Fu, K.; Sun, X.; Wen, H. A Densely Connected End-to-End Neural Network for Multiscale and Multiscene SAR Ship Detection. *IEEE Access* **2018**, *6*, 20881–20892. [[CrossRef](#)]
17. Deng, Z.; Sun, H.; Zhou, S.; Zhao, J.; Lei, L.; Zou, H. Multi-scale object detection in remote sensing imagery with convolutional neural networks. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 3–22. [[CrossRef](#)]
18. 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.* **2019**, *9*, 780–788. [[CrossRef](#)]
19. Wang, R.; Li, J.; Duan, Y.; Cao, H.; Zhao, Y. Study on the Combined Application of CFAR and Deep Learning in Ship Detection. *J. Indian Soc. Remote Sens.* **2018**, *46*, 1413–1421. [[CrossRef](#)]
20. Wang, J.; Lu, C.; Jiang, W. Simultaneous Ship Detection and Orientation Estimation in SAR Images Based on Attention Module and Angle Regression. *Sensors* **2018**, *18*, 2851. [[CrossRef](#)]
21. Zhao, J.; Zhang, Z.; Yu, W.; Truong, T. A Cascade Coupled Convolutional Neural Network Guided Visual Attention Method for Ship Detection from SAR Images. *IEEE Access* **2018**, *6*, 50693–50708. [[CrossRef](#)]
22. Zhao, J.; Guo, W.; Zhang, Z.; Yu, W. A coupled convolutional neural network for small and densely clustered ship detection in SAR images. *Sci. China Inf. Sci.* **2019**, *62*, 42301. [[CrossRef](#)]
23. Khan, H.M.; Cai, Y. Ship detection in SAR Image using YOLOv2. In Proceedings of the 2018 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018. [[CrossRef](#)]
24. Sharifzadeh, F.; Akbarizadeh, G.; Kaviani, Y.S. Ship Classification in SAR Images Using a New Hybrid CNN–MLP Classifier. *J. Indian Soc. Remote Sens.* **2018**, *47*, 551–562. [[CrossRef](#)]
25. Zhou, F.; Fan, W.; Sheng, Q.; Tao, M. Ship Detection Based on Deep Convolutional Neural Networks for PolSAR Images. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018. [[CrossRef](#)]
26. Lei, L.; Chen, G.; Pan, Z.; Lei, B.; An, Q. Inshore Ship Detection in Sar Images Based on Deep Neural Networks. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018. [[CrossRef](#)]
27. Wang, Y.; Chao, W.; Hong, Z. Ship Discrimination with Deep Convolutional Neural Networks in Sar Images. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018. [[CrossRef](#)]
28. Schwegma, N.C.P.; Kleynhans, W.; Salmon, B.P.; Mdakane, L.W.; Meyer, R.G.V. Synthetic Aperture Radar Ship Detection Using Capsule Networks. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018. [[CrossRef](#)]
29. Lin, Z.; Ji, K.; Leng, X.; Kuang, G. Squeeze and Excitation Rank Faster R-CNN for Ship Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 751–755. [[CrossRef](#)]
30. Ma, M.; Chen, J.; Liu, W.; Yang, W. Ship Classification and Detection Based on CNN Using GF-3 SAR Images. *Remote Sens.* **2018**, *10*, 2043. [[CrossRef](#)]
31. Chen, S.W.; Tao, C.S.; Wang, X.S.; Xiao, S.P. Polarimetric SAR Targets Detection and Classification with Deep Convolutional Neural Network. In Proceedings of the 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama), Toyama, Japan, 1–4 August 2018. [[CrossRef](#)]
32. Wang, Z.; Yang, T.; Zhang, H. Land contained sea area ship detection using spaceborne image. *Pattern Recognit. Lett.* **2019**, *130*, 125–131. [[CrossRef](#)]
33. Deng, Z.; Sun, H.; Zhou, S.; Zhao, J. Learning Deep Ship Detector in SAR Images from Scratch. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4021–4039. [[CrossRef](#)]

34. Gui, Y.; Li, X.; Xue, L. A Multilayer Fusion Light-Head Detector for SAR Ship Detection. *Sensors* **2019**, *19*, 1124. [[CrossRef](#)] [[PubMed](#)]
35. Wang, Y.; Wang, C.; Zhang, H.; Dong, Y.; Wei, S. Automatic Ship Detection Based on RetinaNet Using Multi-Resolution Gaofen-3 Imagery. *Remote Sens.* **2019**, *11*, 531. [[CrossRef](#)]
36. Liu, N.; Cao, Z.; Cui, Z.; Pi, Y.; Dang, S. Multi-Scale Proposal Generation for Ship Detection in SAR Images. *Remote Sens.* **2019**, *11*, 526. [[CrossRef](#)]
37. Wang, J.; Zheng, T.; Lei, P.; Bai, X. A Hierarchical Convolution Neural Network (CNN)-Based Ship Target Detection Method in Spaceborne SAR Imagery. *Remote Sens.* **2019**, *11*, 620. [[CrossRef](#)]
38. Wang, Y.; Wang, C.; Zhang, H.; Dong, Y.; Wei, S.L. A SAR Dataset of Ship Detection for Deep Learning under Complex Backgrounds. *Remote Sens.* **2019**, *11*, 765. [[CrossRef](#)]
39. Chang, Y.L.; Anagaw, A.; Chang, L.; Wang, Y.C.; Hsiao, C.Y.; Lee, W.H. Ship detection based on YOLOv2 for SAR imagery. *Remote Sens.* **2019**, *11*, 786. [[CrossRef](#)]
40. Zhang, T.; Zhang, X. High-Speed Ship Detection in SAR Images Based on a Grid Convolutional Neural Network. *Remote Sens.* **2019**, *11*, 1206. [[CrossRef](#)]
41. An, Q.; Pan, Z.; Liu, L.; You, H. DRBox-v2: An Improved Detector with Rotatable Boxes for Target Detection in SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8333–8349. [[CrossRef](#)]
42. Cui, Z.; Li, Q.; Cao, Z.; Liu, N. Dense Attention Pyramid Networks for Multi-Scale Ship Detection in SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8983–8997. [[CrossRef](#)]
43. Chen, C.; Hu, C.; He, C.; Pei, H.; Pang, H.; Zhao, T. SAR ship detection under complex background based on attention mechanism. In *Chinese Conference on Image and Graphics Technologies*; Springer: Singapore, 2019; pp. 565–578.
44. Chen, C.; He, C.; Hu, C.; Pei, H.; Jiao, L. A Deep Neural Network Based on an Attention Mechanism for SAR Ship Detection in Multiscale and Complex Scenarios. *IEEE Access* **2019**, *7*, 104848–104863. [[CrossRef](#)]
45. Guo, Q.; Wang, H.; Kang, L.; Li, Z.; Xu, F. Aircraft Target Detection from Spaceborne SAR Image. In *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019*. [[CrossRef](#)]
46. Gui, Y.; Li, X.; Xue, L.; Lv, J. A scale transfer convolution network for small ship detection in SAR images. In *Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 24–26 May 2019*. [[CrossRef](#)]
47. Li, Y.; Chen, J.; Ke, M.; Li, L.; Ding, Z.; Wang, Y. Small targets recognition in SAR ship image based on improved SSD. In *Proceedings of the 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 11–13 December 2019*. [[CrossRef](#)]
48. Ai, J.; Tian, R.; Luo, Q.; Jin, J.; Tang, B. Multi-Scale Rotation-Invariant Haar-Like Feature Integrated CNN-Based Ship Detection Algorithm of Multiple-Target Environment in SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 10070–10087. [[CrossRef](#)]
49. Fan, Q.; Chen, F.; Cheng, M.; Lou, S.; Xiao, R.; Zhang, B.; Wang, C.; Li, J. Ship Detection Using a Fully Convolutional Network with Compact Polarimetric SAR Images. *Remote Sens.* **2019**, *11*, 2171. [[CrossRef](#)]
50. Ayhan, N.; Sen, N. Ship detection in synthetic aperture radar (SAR) images by deep learning. In *Proceedings of the Artificial Intelligence and Machine Learning in Defense Applications, Strasbourg, France, 19 September 2019*.
51. Zhang, X.; Wang, H.; Xu, C.; Lv, Y.; Fu, C.; Xiao, H.; He, Y. Lightweight Feature Optimizing Network for Ship Detection in SAR Image. *IEEE Access* **2019**, *7*, 141662–141678. [[CrossRef](#)]
52. Yang, T.; Zhu, J.; Liu, J. SAR Image Target Detection and Recognition based on Deep Network. In *Proceedings of the 2019 SAR in Big Data Era (BIGSARDATA), Beijing, China, 5–6 August 2019*. [[CrossRef](#)]
53. Hong, S.J.; Baek, W.K.; Jung, H.S. Ship Detection from X-Band SAR Images Using M2Det Deep Learning Model. *Appl. Sci.* **2020**, *10*, 7751. [[CrossRef](#)]
54. Zhang, T.; Zhang, X.; Shi, J.; Wei, S. Depthwise Separable Convolution Neural Network for High-Speed SAR Ship Detection. *Remote Sens.* **2019**, *11*, 2483. [[CrossRef](#)]
55. Chen, S.; Zhan, R.; Zhang, J. Regional attention-based single shot detector for SAR ship detection. *J. Eng.* **2019**, *2019*, 7381–7384. [[CrossRef](#)]
56. Chen, C.; He, C.; Hu, C.; Pei, H.; Jiao, L. MSARN: A Deep Neural Network Based on an Adaptive Recalibration Mechanism for Multiscale and Arbitrary-oriented SAR Ship Detection. *IEEE Access* **2019**, *7*, 159262–159283. [[CrossRef](#)]
57. Yue, B.; Zhao, W.; Han, S. SAR Ship Detection Method Based on Convolutional Neural Network and Multi-layer Feature Fusion. In *The International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*; Springer: Cham, Switzerland, 2019.
58. Wang, Z.; Yang, W.; Chen, J.; Li, C. A Level Set Based Method for Land Masking in Ship Detection Using SAR Images. In *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019*. [[CrossRef](#)]
59. Hou, X.; Ao, W.; Xu, F. End-to-end Automatic Ship Detection and Recognition in High-Resolution Gaofen-3 Spaceborne SAR Images. In *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019*. [[CrossRef](#)]
60. Wang, R.; Xu, F.; Pei, J.; Wang, C.; Huang, Y.; Yang, J.; Wu, J. An Improved Faster R-CNN Based on MSER Decision Criterion for SAR Image Ship Detection in Harbor. In *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019*. [[CrossRef](#)]

61. Li, Y.; Ding, Z.; Zhang, C.; Wang, Y.; Chen, J. SAR Ship Detection Based on Resnet and Transfer Learning. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019. [[CrossRef](#)]
62. Li, Q.; Min, R.; Cui, Z.; Pi, Y.; Xu, Z. Multiscale Ship Detection Based on Dense Attention Pyramid Network in Sar Images. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019. [[CrossRef](#)]
63. Liu, N.; Cui, Z.; Cao, Z.; Pi, Y.; Lan, H. Scale-Transferrable Pyramid Network for Multi-Scale Ship Detection in Sar Images. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019. [[CrossRef](#)]
64. Gao, F.; Shi, W.; Wang, J.; Yang, E.; Zhou, H. Enhanced Feature Extraction for Ship Detection from Multi-Resolution and Multi-Scene Synthetic Aperture Radar (SAR) Images. *Remote Sens.* **2019**, *11*, 2694. [[CrossRef](#)]
65. Sun, X.; Wang, Z.; Sun, Y.; Diao, W.; Zhang, Y.; Fu, K. AIR-SARShip-1.0: High-resolution SAR ship detection dataset. *J. Radars* **2019**, *8*, 852–862.
66. Fan, W.; Zhou, F.; Bai, X.; Tao, M.; Tian, T. Ship Detection Using Deep Convolutional Neural Networks for PolSAR Images. *Remote Sens.* **2019**, *11*, 2862. [[CrossRef](#)]
67. Dechesne, C.; Lefèvre, S.; Vadaine, R.; Hajduch, G.; Fablet, R. Ship Identification and Characterization in Sentinel-1 SAR Images with Multi-Task Deep Learning. *Remote Sens.* **2019**, *11*, 2997. [[CrossRef](#)]
68. Zhang, X.; Zhang, T.; Shi, J.; Wei, S. High-speed and High-accurate SAR ship detection based on a depthwise separable convolution neural network. *J. Radars* **2019**, *8*, 841–851.
69. Wei, S.; Su, H.; Ming, J.; Wang, C.; Yan, M.; Kumar, D.; Shi, J.; Zhang, X. Precise and Robust Ship Detection for High-Resolution SAR Imagery Based on HR-SDNet. *Remote Sens.* **2020**, *12*, 167. [[CrossRef](#)]
70. Chen, P.; Li, Y.; Zhou, H.; Liu, B.; Liu, P. Detection of Small Ship Objects Using Anchor Boxes Cluster and Feature Pyramid Network Model for SAR Imagery. *J. Mar. Sci. Eng.* **2020**, *8*, 112. [[CrossRef](#)]
71. Milios, A.; Bereta, K.; Chatzikokolakis, K.; Zissis, D.; Matwin, S. Automatic fusion of satellite imagery and AIS data for vessel detection. In Proceedings of the 2019 22th International Conference on Information Fusion (FUSION), Ottawa, ON, Canada, 2–5 July 2019; pp. 1–5.
72. Jin, K.; Chen, Y.; Xu, B.; Yin, J.; Yang, J. A Patch-to-Pixel Convolutional Neural Network for Small Ship Detection with PolSAR Images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6623–6638. [[CrossRef](#)]
73. Hou, X.; Ao, W.; Song, Q.; Lai, J.; Wang, H.; Xu, F. FUSAR-Ship: Building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition. *Sci. China Inf. Sci.* **2020**, *63*, 140303. [[CrossRef](#)]
74. Su, H.; Wei, S.; Liu, S.; Liang, J.; Wang, C.; Shi, J.; Zhang, X. HQ-ISNet: High-Quality Instance Segmentation for Remote Sensing Imagery. *Remote Sens.* **2020**, *12*, 989. [[CrossRef](#)]
75. Tanveer, H.; Balz, T.; Mohamdi, B. Using convolutional neural network (CNN) approach for ship detection in Sentinel-1 SAR imagery. In Proceedings of the 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Xiamen, China, 26–29 November 2019. [[CrossRef](#)]
76. Wang, J.; Chen, J.; Wang, P.; Zhao, C.; Pan, X.; Gao, A. An Algorithm for Azimuth Ambiguities Detection in SAR Images Using Faster-RCNN. In Proceedings of the 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Xiamen, China, 26–29 November 2019. [[CrossRef](#)]
77. Zheng, T.; Wang, J.; Lei, P. Deep learning based target detection method with multi-features in SAR imagery. In Proceedings of the 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Xiamen, China, 26–29 November 2019. [[CrossRef](#)]
78. Su, H.; Wei, S.; Wang, M.; Zhou, L.; Shi, J.; Zhang, X. Ship Detection Based on RetinaNet-Plus for High-Resolution SAR Imagery. In Proceedings of the 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Xiamen, China, 26–29 November 2019. [[CrossRef](#)]
79. Wang, C.; Pei, J.; Wang, R.; Huang, Y.; Yang, J. A new ship detection and classification method of spaceborne SAR images under complex scene. In Proceedings of the 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Xiamen, China, 26–29 November 2019. [[CrossRef](#)]
80. Xiao, Q.; Cheng, Y.; Xiao, M.; Zhang, J.; Shi, H.; Niu, L.; Ge, C.; Lang, H. Improved region convolutional neural network for ship detection in multiresolution synthetic aperture radar images. *Concurr. Comput. Pract. Exp.* **2020**, *32*, 5820. [[CrossRef](#)]
81. Mao, Y.; Yang, Y.; Ma, Z.; Li, M.; Su, H.; Zhang, J. Efficient Low-Cost Ship Detection for SAR Imagery Based on Simplified U-Net. *IEEE Access* **2020**, *8*, 69742–69753. [[CrossRef](#)]
82. Zhang, T.; Zhang, X.; Shi, J.; Wei, S. High-Speed Ship Detection in SAR Images by Improved Yolov3. In Proceedings of the 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing, Chengdu, China, 14–15 December 2019; pp. 149–152. [[CrossRef](#)]
83. Pan, Z.; Yang, R.; Zhang, A.Z. MSR2N: Multi-Stage Rotational Region Based Network for Arbitrary-Oriented Ship Detection in SAR Images. *Sensors* **2020**, *20*, 2340. [[CrossRef](#)]
84. Zhang, T.; Zhang, X. ShipDeNet-20: An Only 20 Convolution Layers and <1-MB Lightweight SAR Ship Detector. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1234–1238. [[CrossRef](#)]
85. Gu, F.; Zhang, H.; Wang, C.; Zhang, B. Weakly supervised ship detection from SAR images based on a three-component CNN-CAM-CRF model. *J. Appl. Remote Sens.* **2020**, *14*, 026506. [[CrossRef](#)]

86. Dai, W.; Mao, Y.; Yuan, R.; Liu, Y.; Pu, X.; Li, C. A Novel Detector Based on Convolution Neural Networks for Multiscale SAR Ship Detection in Complex Background. *Sensors* **2020**, *20*, 2547. [[CrossRef](#)]
87. Zhou, Y.; Cai, Z.; Zhu, Y.; Yan, J. Automatic ship detection in SAR Image based on Multi-scale Faster R-CNN. *J. Phys. Conf. Ser.* **2020**, *1550*, 042006. [[CrossRef](#)]
88. Kang, K.M. Automated Procurement of Training Data for Machine Learning Algorithm on Ship Detection Using AIS Information. *Remote Sens.* **2020**, *12*, 1443. [[CrossRef](#)]
89. Cui, Z.; Wang, X.; Liu, N.; Cao, Z.; Yang, J. Ship Detection in Large-Scale SAR Images Via Spatial Shuffle-Group Enhance Attention. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 379–391. [[CrossRef](#)]
90. Yang, R.; Wang, G.; Pan, Z.; Lu, H.; Zhang, H.; Jia, X. A Novel False Alarm Suppression Method for CNN-Based SAR Ship Detector. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1401–1405. [[CrossRef](#)]
91. Zhao, Y.; Zhao, L.; Xiong, B.; Kuang, G. Attention Receptive Pyramid Network for Ship Detection in SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 2738–2756. [[CrossRef](#)]
92. Han, L.; Zheng, T.; Ye, W.; Ran, D. Analysis of Detection Preference to CNN Based SAR Ship Detectors. In Proceedings of the 2020 Information Communication Technologies Conference (ICTC), Nanjing, China, 29–31 May 2020. [[CrossRef](#)]
93. Chen, S.; Zhang, J.; Zhan, R. R2FA-Det: Delving into High-Quality Rotatable Boxes for Ship Detection in SAR Images. *Remote Sens.* **2020**, *12*, 2031. [[CrossRef](#)]
94. Wei, S.; Zeng, X.; Qu, Q.; Wang, M.; Su, H.; Shi, J. HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation. *IEEE Access* **2020**, *8*, 2031. [[CrossRef](#)]
95. Fu, J.; Sun, X.; Wang, Z.; Fu, K. An Anchor-Free Method Based on Feature Balancing and Refinement Network for Multiscale Ship Detection in SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1331–1344. [[CrossRef](#)]
96. Zhang, T.; Zhang, X.; Shi, J.; Wei, S. HyperLi-Net: A hyper-light deep learning network for high-accurate and high-speed ship detection from synthetic aperture radar imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 123–153. [[CrossRef](#)]
97. Zhou, H. Anchor-free Convolutional Network with Dense Attention Feature Aggregation for Ship Detection in SAR Images. *Remote Sens.* **2020**, *12*, 2649.
98. Han, L.; Zhao, X.; Ye, W.; Ran, D. Asymmetric and square convolutional neural network for SAR ship detection from scratch. In Proceedings of the 2020 5th International Conference on Biomedical Signal and Image Processing, Suzhou, China, 21–23 August 2020; pp. 80–85.
99. Han, L.; Ye, W.; Li, J.; Ran, D. Small ship detection in SAR images based on modified SSD. In Proceedings of the 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 11–13 December 2019. [[CrossRef](#)]
100. Han, L.; Ran, D.; Ye, W.; Yang, W.; Wu, X. Multi-size Convolution and Learning Deep Network for SAR Ship Detection from Scratch. *IEEE Access* **2020**, *8*, 158996–159016. [[CrossRef](#)]
101. Zhang, T.; Zhang, X.; Ke, X.; Zhan, X.; Shi, J.; Wei, S.; Pan, D.; Li, J.; Su, H.; Zhou, Y. LS-SSDD-v1.0: A Deep Learning Dataset Dedicated to Small Ship Detection from Large-Scale Sentinel-1 SAR Images. *Remote Sens.* **2020**, *12*, 2997. [[CrossRef](#)]
102. Zhou, L.; Wei, S.; Cui, Z.; Fang, J.; Yang, X.; Ding, W. Lira-YOLO: A lightweight model for ship detection in radar images. *J. Syst. Eng. Electron.* **2020**, *31*, 950–956. [[CrossRef](#)]
103. Mao, Y.; Li, X.; Li, Z.; Li, M.; Chen, S. An Anchor-free SAR Ship Detector with Only 1.17M Parameters. In Proceedings of the ICASIT 2020: 2020 International Conference on Aviation Safety and Information Technology, Weihai, China, 14–16 October 2020.
104. Mao, Y.; Li, X.; Li, Z.; Li, M.; Chen, S. Network slimming method for SAR ship detection based on knowledge distillation. In Proceedings of the 2020 International Conference on Aviation Safety and Information Technology, Weihai, China, 14–16 October 2020; pp. 177–181.
105. Stefanowicz, J.; Ali, I.; Andersson, S. Current trends in ship detection in single polarization synthetic aperture radar imagery. In Proceedings of the Proc. SPIE 11581, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, Wilga, Poland, 14 October 2020; p. 1158109.
106. Li, K.; Luan, S.; Zhou, D. An Optical-to-SAR Transformation Method for SAR Ship Image Augmentation. In Proceedings of the 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP), Shanghai, China, 12–15 September 2020. [[CrossRef](#)]
107. Xu, C.; Yin, C.; Wang, D.; Han, W. Fast ship detection combining visual saliency and a cascade CNN in SAR images. *IET Radar Sonar Navig.* **2020**, *14*, 1879–1887. [[CrossRef](#)]
108. Han, L.; Ran, D.; Ye, W.; Wu, X. Asymmetric convolution-based neural network for SAR ship detection from scratch. In Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition, Xiamen China, 30 October–1 November 2020; pp. 90–95.
109. Idicula, S.M.; Paul, B. Real time SAR Ship Detection using novel SarNeDe method. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 2198–2201. [[CrossRef](#)]
110. Hu, W.; Tian, Z.; Chen, S.; Zhan, R.; Zhang, J. Dense feature pyramid network for ship detection in SAR images. In Proceedings of the Third International Conference on Image, Video Processing and Artificial Intelligence, Shanghai, China, 23–24 October 2020. [[CrossRef](#)]
111. Zhang, T.; Zhang, X.; Shi, J.; Wei, S.; Wang, J.; Li, J.; Su, H.; Zhou, Y. Balance Scene Learning Mechanism for Offshore and Inshore Ship Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 4004905. [[CrossRef](#)]

112. Zou, L.; Zhang, H.; Wang, C.; Wu, F. MW-ACGAN: Generating Multiscale High-Resolution SAR Images for Ship Detection. *Sensors* **2020**, *20*, 6673. [[CrossRef](#)]
113. Zhang, G.; Li, Z.; Li, X.; Yin, C.; Shi, Z. A Novel Salient Feature Fusion Method for Ship Detection in Synthetic Aperture Radar Images. *IEEE Access* **2020**, *8*, 215904–215914. [[CrossRef](#)]
114. Chen, Y.; Yu, J.; Xu, Y. SAR Ship Target Detection for SSDv2 under Complex Backgrounds. In Proceedings of the 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), Chongqing, China, 10–12 July 2020. [[CrossRef](#)]
115. Li, Y.; Zhang, S.; Wang, W.Q. A Lightweight Faster R-CNN for Ship Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 4006105. [[CrossRef](#)]
116. Zhu, M.; Hu, G.; Zhou, H.; Lu, C. Rapid Ship Detection in SAR Images Based on YOLOv3. In Proceedings of the 2020 5th International Conference on Communication, Image and Signal Processing (CCISP), Chengdu, China, 13–15 November 2020. [[CrossRef](#)]
117. Zhang, T.; Zhang, X.; Shi, J.; Wei, S.; Wang, J.; Li, J. Balanced Feature Pyramid Network for Ship Detection in Synthetic Aperture Radar Images. In Proceedings of the 2020 IEEE Radar Conference (RadarConf20), Florence, Italy, 21–25 September 2020. [[CrossRef](#)]
118. Chen, S.; Zhan, R.; Wang, W.; Zhang, J. Learning Slimming SAR Ship Object Detector Through Network Pruning and Knowledge Distillation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *14*, 1267–1282. [[CrossRef](#)]
119. Wang, J.; Lin, Y.; Guo, J.; Zhuang, L. SSS-YOLO: Towards more accurate detection for small ships in SAR image. *Remote Sens. Lett.* **2021**, *12*, 93–102. [[CrossRef](#)]
120. Yang, R.; Wang, R.; Deng, Y.; Jia, X.; Zhang, H. Rethinking the Random Cropping Data Augmentation Method Used in the Training of CNN-based SAR Image Ship Detector. *Remote Sens.* **2020**, *13*, 34. [[CrossRef](#)]
121. Guo, H.; Yang, X.; Wang, N.; Gao, X. A CenterNet++ model for ship detection in SAR images. *Pattern Recognit.* **2021**, *112*, 107787. [[CrossRef](#)]
122. Wang, C.; Su, W.; Gu, H. Two-stage ship detection in synthetic aperture radar images based on attention mechanism and extended pooling. *J. Appl. Remote Sens.* **2020**, *14*, 044522. [[CrossRef](#)]
123. Chaudhary, Y.; Mehta, M.; Goel, N.; Bhardwaj, P.; Gupta, D.; Khanna, A. YOLOv3 Remote Sensing SAR Ship Image Detection. In *Data Analytics and Management*; Springer: Singapore, 2021; pp. 519–531.
124. Yang, R.; Pan, Z.; Jia, X.; Zhang, L.; Deng, Y. A Novel CNN-Based Detector for Ship Detection Based on Rotatable Bounding Box in SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1938–1958. [[CrossRef](#)]
125. Liu, C.; Zhu, W. An improved algorithm for ship detection in SAR images based on CNN. In Proceedings of the Twelfth International Conference on Graphics and Image Processing, Xi'an, China, 13–15 November 2021.
126. Wang, J.; Wen, Z.; Lu, Y.; Wang, X.; Pan, Q. Weakly Supervised SAR Ship Segmentation Based on Variational Gaussian G (A) (0) Mixture Model A Learning. In Proceedings of the 2020 Chinese Automation Congress (CAC), Shanghai, China, 6–8 November 2020.
127. Mao, Y.; Li, X.; Su, H.; Zhou, Y.; Li, J. Ship Detection for SAR Imagery Based on Deep Learning: A Benchmark. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020. [[CrossRef](#)]
128. Zhao, K.; Zhou, Y.; Chen, X. A Dense Connection Based SAR Ship Detection network. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020. [[CrossRef](#)]
129. An, Q.; Pan, Z.; You, H.; Hu, Y. Transitive Transfer Learning Based Anchor Free Rotatable Detector for SAR Target Detection With Few Samples. *IEEE Access* **2021**, *9*, 24011–24025. [[CrossRef](#)]
130. Zhang, P.; Luo, H.; Ju, M.; He, M.; Chang, Z.; Hui, B. Brain-Inspired Fast Saliency-Based Filtering Algorithm for Ship Detection in High-Resolution SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5201709. [[CrossRef](#)]
131. Wang, R.; Shao, S.; An, M.; Li, J.; Wang, S.; Xu, X. Soft Thresholding Attention Network for Adaptive Feature Denoising in SAR Ship Detection. *IEEE Access* **2021**, *9*, 29090–29105. [[CrossRef](#)]
132. Zhang, T.; Zhang, X.; Shi, J.; Wei, S. Shipdenet-18: An Only 1 Mb With Only 18 Convolution Layers Light-Weight Deep Learning Network for Sar Ship Detection. In Proceedings of the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [[CrossRef](#)]
133. Wang, X.; Cui, Z.; Cao, Z.; Dang, S. Dense Docked Ship Detection via Spatial Group-Wise Enhance Attention in SAR Images. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [[CrossRef](#)]
134. Hou, S.; Ma, X.; Wang, X.; Fu, Z.; Wang, J.; Wang, H. SAR Image Ship Detection Based on Scene Interpretation. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [[CrossRef](#)]
135. Jiang, S.; Zhu, M.; He, Y.; Zheng, Z.; Zhou, F.; Zhou, G. Ship Detection with Sar Based on Yolo. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [[CrossRef](#)]
136. Zhu, C.; Zhao, D.; Liu, Z.; Mao, Y. Hierarchical Attention for Ship Detection in SAR Images. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [[CrossRef](#)]

137. Hou, Z.; Cui, Z.; Cao, Z.; Liu, N. An Integrated Method of Ship Detection and Recognition in Sar Images based on Deep Learning. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [\[CrossRef\]](#)
138. Wang, X.; Cui, Z.; Cao, Z.; Tian, Y. Ship Detection in Large Scale Sar Images Based on Bias Classification. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [\[CrossRef\]](#)
139. Shao, P.; Lu, X.; Huang, P.; Xu, W.; Dong, Y. Impact Analysis of Radio Frequency Interference on SAR Image Ship Detection Based on Deep Learning. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [\[CrossRef\]](#)
140. Li, J.; Guo, C.; Gou, S.; Wang, M.; Chen, J. Ship Segmentation on High-Resolution Sar Image by a 3D Dilated Multiscale U-Net. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [\[CrossRef\]](#)
141. Ferreira, N.; Silveira, M. Ship Detection in SAR Images Using Convolutional Variational Autoencoders. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020.
142. Lee, S.J.; Chang, J.Y.; Lee, K.J.; Oh, K.Y. Data Augmentation for Ship Detection using Kompsat-5 Images and Deep Learning Model. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [\[CrossRef\]](#)
143. Song, J.; Kim, D.J. Fine Acquisition of Vessel Training Data for Machine Learning from Sentinel-1 SAR Images Accompanied by AIS Information. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [\[CrossRef\]](#)
144. Tang, G.; Zhuge, Y.; Claramunt, C.; Men, S. N-YOLO: A SAR Ship Detection Using Noise-Classifying and Complete-Target Extraction. *Remote Sens.* **2021**, *13*, 871. [\[CrossRef\]](#)
145. Raj, J.A.; Idicula, S.M.; Paul, B. A novel Ship detection method from SAR image with reduced false alarm. *J. Phys. Conf. Ser.* **2021**, *1817*, 012010. [\[CrossRef\]](#)
146. Jiang, K.; Cao, Y. SAR Image Ship Detection Based on Deep Learning. In Proceedings of the 2020 International Conference on Computer Engineering and Intelligent Control (ICCEIC), Chongqing, China, 6–8 November 2020. [\[CrossRef\]](#)
147. Li, D.; Liang, Q.; Liu, H.; Liu, Q.; Liu, H.; Liao, G. A Novel Multidimensional Domain Deep Learning Network for SAR Ship Detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5203213. [\[CrossRef\]](#)
148. Geng, X.; Shi, L.; Yang, J.; Li, P.; Zhao, L.; Sun, W.; Zhao, J. Ship Detection and Feature Visualization Analysis Based on Lightweight CNN in VH and VV Polarization Images. *Remote Sens.* **2021**, *13*, 1184. [\[CrossRef\]](#)
149. Jin, L.; Liu, G. An Approach on Image Processing of Deep Learning Based on Improved SSD. *Symmetry* **2021**, *13*, 495. [\[CrossRef\]](#)
150. Ren, Y.; Li, X.; Xu, H. A Deep Learning Model to Extract Ship Size from Sentinel-1 SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5203414. [\[CrossRef\]](#)
151. Chen, Y.; Duan, T.; Wang, C.; Zhang, Y.; Huang, M. End-to-End Ship Detection in SAR Images for Complex Scenes Based on Deep CNNs. *J. Sens.* **2021**, *2021*, 8893182. [\[CrossRef\]](#)
152. He, Y.; Gao, F.; Wang, J.; Hussain, A.; Yang, E.; Zhou, H. Learning Polar Encodings for Arbitrary-Oriented Ship Detection in SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3846–3859. [\[CrossRef\]](#)
153. Tian, L.; Cao, Y.; He, B.; Zhang, Y.; He, C.; Li, D. Image Enhancement Driven by Object Characteristics and Dense Feature Reuse Network for Ship Target Detection in Remote Sensing Imagery. *Remote Sens.* **2021**, *13*, 1327. [\[CrossRef\]](#)
154. Li, Y.; Zhu, W.; Zhu, B. SAR image nearshore ship target detection in complex environment. In Proceedings of the 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12–14 March 2021; Volume 5, pp. 1964–1968. [\[CrossRef\]](#)
155. Zhao, K.; Zhou, Y.; Chen, X.; Wang, B.; Zhang, Y. Ship detection from scratch in Synthetic Aperture Radar (SAR) images. *Int. J. Remote Sens.* **2021**, *42*, 5010–5024. [\[CrossRef\]](#)
156. Ma, Z. High-Speed Lightweight Ship Detection Algorithm Based on YOLO-V4 for Three-Channels RGB SAR Image. *Remote Sens.* **2021**, *13*, 1909. [\[CrossRef\]](#)
157. Zhong, R. On-Board Real-Time Ship Detection in HISEA-1 SAR Images Based on CFAR and Lightweight Deep Learning. *Remote Sens.* **2021**, *13*, 1995. [\[CrossRef\]](#)
158. Hong, Z.; Yang, T.; Tong, X.; Zhang, Y.; Jiang, S.; Zhou, R.; Han, Y.; Wang, J.; Yang, S.; Liu, S. Multi-Scale Ship Detection from SAR and Optical Imagery via A More Accurate YOLOv3. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6083–6101. [\[CrossRef\]](#)
159. Zhu, M.; Hu, G.; Li, S.; Liu, S.; Wang, S. An Effective Ship Detection Method Based on RefineDet in SAR Images. In Proceedings of the 2021 International Conference on Communications, Information System and Computer Engineering (CISCE), Beijing, China, 14–16 May 2021. [\[CrossRef\]](#)
160. Shin, S.; Kim, Y.; Hwang, I.; Kim, J.; Kim, S. Coupling Denoising to Detection for SAR Imagery. *Appl. Sci.* **2021**, *11*, 5569. [\[CrossRef\]](#)
161. Hu, H. TWC-Net: A SAR Ship Detection Using Two-Way Convolution and Multiscale Feature Mapping. *Remote Sens.* **2021**, *13*, 2558. [\[CrossRef\]](#)
162. Sun, W.; Huang, X. Semantic attention-based network for inshore SAR ship detection. In Proceedings of the SPIE 11878, Thirteenth International Conference on Digital Image Processing (ICDIP 2021), Singapore, 30 June 2021; Volume 118782A.

163. Wu, Z.; Hou, B.; Ren, B.; Ren, Z.; Wang, S.; Jiao, L. A Deep Detection Network Based on Interaction of Instance Segmentation and Object Detection for SAR Images. *Remote Sens.* **2021**, *13*, 2582. [CrossRef]
164. Dong, Y.; Zhang, H.; Wang, C.; Zhang, B.; Li, L. Ship Detection based on M2Det for SAR images under Heavy Sea State. In Proceedings of the EUSAR 2021 13th European Conference on Synthetic Aperture Radar. VDE, Online, 29 March–1 April 2021; pp. 1–4.
165. Sun, K.; Liang, Y.; Ma, X.; Huai, Y.; Xing, M. DSDet: A Lightweight Densely Connected Sparsely Activated Detector for Ship Target Detection in High-Resolution SAR Images. *Remote Sens.* **2021**, *13*, 2743. [CrossRef]
166. Zhang, T.; Zhang, X.; Ke, X. Quad-FPN: A Novel Quad Feature Pyramid Network for SAR Ship Detection. *Remote Sens.* **2021**, *13*, 2771. [CrossRef]
167. Sun, Z.; Dai, M.; Leng, X.; Lei, Y.; Xiong, B.; Ji, K.; Kuang, G. An Anchor-free Detection Method for Ship Targets in High-Resolution SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7799–7816. [CrossRef]
168. Zhang, X.; Huo, C.; Xu, N.; Jiang, H.; Cao, Y.; Ni, L.; Pan, C. Multitask Learning for Ship Detection from Synthetic Aperture Radar Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8048–8062. [CrossRef]
169. Du, Y.; Du, L.; Li, L. An SAR Target Detector Based on Gradient Harmonized Mechanism and Attention Mechanism. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 4017005. [CrossRef]
170. Zhang, T.; Zhang, X.; Li, J.; Xu, X.; Wang, B.; Zhan, X.; Xu, Y.; Ke, X.; Zeng, T.; Su, H. SAR Ship Detection Dataset (SSDD): Official Release and Comprehensive Data Analysis. *Remote Sens.* **2021**, *13*, 3690. [CrossRef]
171. Liu, F.; Li, Y. SAR remote sensing image ship detection method NanoDet based on visual saliency. *J. Radars* **2021**, *10*, 885–894.
172. Zhao, Y.; Zhao, L.; Liu, Z.; Hu, D.; Kuang, G.; Liu, L. Attentional Feature Refinement and Alignment Network for Aircraft Detection in SAR Imagery. *arXiv* **2022**, arXiv:2201.07124. [CrossRef]
173. Li, S.; Xiao, Y.; Zhang, Y.; Chu, L.; Qiu, R.C. Learning Efficient Representations for Enhanced Object Detection on Large-scene SAR Images. *arXiv* **2022**, arXiv:2201.08958.
174. Song, T.; Kim, S.; Kim, S.T.; Lee, J.; Sohn, K. Context-Preserving Instance-Level Augmentation and Deformable Convolution Networks for SAR Ship Detection. *arXiv* **2022**, arXiv:2202.06513.
175. Zhu, H.; Xie, Y.; Huang, H.; Jing, C.; Rong, Y.; Wang, C. DB-YOLO: A Duplicate Bilateral YOLO Network for Multi-Scale Ship Detection in SAR Images. *Sensors* **2021**, *21*, 8146. [CrossRef] [PubMed]
176. Lin, S. A Lightweight Detection Model for SAR Aircraft in a Complex Environment. *Remote Sens.* **2021**, *13*, 5020. [CrossRef]
177. Ding, C. SRSDD-v1.0: A High-Resolution SAR Rotation Ship Detection Dataset. *Remote Sens.* **2021**, *13*, 5104. [CrossRef]
178. Qin, M. A Fast and Lightweight Detection Network for Multi-Scale SAR Ship Detection under Complex Backgrounds. *Remote Sens.* **2021**, *14*, 31. [CrossRef]
179. Zhou, K.; Zhang, M.; Wang, H.; Tan, J. Ship Detection in SAR Images Based on Multi-Scale Feature Extraction and Adaptive Feature Fusion. *Remote Sens.* **2022**, *14*, 755. [CrossRef]
180. Zhu, M.; Hu, G.; Zhou, H.; Wang, S.; Feng, Z.; Yue, S. A Ship Detection Method via Redesigned FCOS in Large-Scale SAR Images. *Remote Sens.* **2022**, *14*, 1153. [CrossRef]
181. Liu, S.; Kong, W.; Chen, X.; Xu, M.; Yasir, M.; Zhao, L.; Li, J. Multi-Scale Ship Detection Algorithm Based on a Lightweight Neural Network for Spaceborne SAR Images. *Remote Sens.* **2022**, *14*, 1149. [CrossRef]
182. Xia, R.; Chen, J.; Huang, Z.; Wan, H.; Wu, B.; Sun, L.; Yao, B.; Xiang, H.; Xing, M. CRTransSar: A Visual Transformer Based on Contextual Joint Representation Learning for SAR Ship Detection. *Remote Sens.* **2022**, *14*, 1488. [CrossRef]
183. Everingham, M.; Eslami, S.; Gool, L.V.; Williams, C.; Winn, J.; Zisserman, A. The pascal visual object classes challenge: A retrospective. *Int. J. Comput. Vis.* **2015**, *111*, 98–136. [CrossRef]
184. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, P.; Dollár, P.; Zitnick, C.L. Microsoft Coco: Common Objects in Context. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2014; pp. 740–755. [CrossRef]
185. 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]
186. 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, Las Vegas, NV, USA, 26 June–1 July 2016; IEEE Press: Piscataway, NJ, USA, 2016; pp. 779–788. [CrossRef]
187. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox detector. *Computer Vision—ECCV 2016*; Springer International Publishing: Cham, Switzerland, 2016; pp. 21–37. [CrossRef]
188. Jia, Y.Q.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM International Conference on Multimedia; ACM Press: New York, NY, USA, 2014; pp. 675–678.
189. 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. [CrossRef]
190. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.M.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 8024–8035.
191. Available online: <http://radars.ie.ac.cn/web/data/getData?newsColumnId=74fe223a-0b01-4830-8d99-1ba276e67ad8&pageType=en> (accessed on 23 April 2021).
192. Xu, C.; Su, H.; Li, J.; Li, Y.; Yao, L.; Gao, L.; Yan, W. RSDD-SAR: Rotated Ship Detection Dataset in SAR Images. *J. Radars*. in press.

193. Dai, J.; Li, Y.; He, K.; Sun, J. R-fcn: Object detection via region-based fully convolutional networks. *Adv. Neural Inf. Process. Syst.* **2016**, *29*. [[CrossRef](#)]
194. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; IEEE Press: Piscataway, NJ, USA, 2017; pp. 936–944.
195. Cai, Z.; Vasconcelos, N. Cascade r-cnn: Delving into high quality object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 18–23 June 2018; pp. 6154–6162. [[CrossRef](#)]
196. He, K.M.; Gkioxari, G.; Dollár, P.; Girshick, R.B. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 386–397. [[CrossRef](#)]
197. Li, Z.; Peng, C.; Yu, G.; Zhang, X.Y.; Deng, Y.D.; Sun, J. Light-head r-cnn: In defense of two-stage object detector. *arXiv* **2017**, arXiv:1711.07264. [[CrossRef](#)]
198. Liu, S.; Huang, D.; Wang, Y. Learning spatial fusion for single-shot object detection. *arXiv* **2019**, arXiv:1911.09516. [[CrossRef](#)]
199. Golnaz, G.; Lin, T.; Le, Q.V. Nas-fpn: Learning scalable feature pyramid architecture for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–17 June 2019.
200. Tan, M.; Pang, R.; Le, Q.V. EfficientDet: Scalable and Efficient Object Detection. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020. [[CrossRef](#)]
201. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV), Glasgow, UK, 23–28 August 2018; pp. 3–19. [[CrossRef](#)]
202. Shrivastava, A.; Gupta, A.; Girshick, R. Training Region-based Object Detectors with Online Hard Example Mining. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, 27–30 June 2016; pp. 761–769. [[CrossRef](#)]
203. Li, B.; Liu, Y.; Wang, X. Gradient Harmonized Single-Stage Detector. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2019; Volume 33, pp. 8577–8584. [[CrossRef](#)]
204. Pang, J.; Chen, K.; Shi, J.; Feng, H.J.; Ouyang, W.L.; Lin, D.H. Libra R-CNN: Towards Balanced Learning for Object Detection. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2020. [[CrossRef](#)]
205. Bodla, N.; Singh, B.; Chellappa, R.; Davis, L.S. Soft-NMS—improving object detection with one line of code. In Proceedings of the IEEE International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2017; pp. 5561–5569. [[CrossRef](#)]
206. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *99*, 2999–3007. [[CrossRef](#)]
207. Law, H.; Deng, J. CornerNet: Detecting Objects as Paired Keypoints. *Int. J. Comput. Vis.* **2018**, 734–750. [[CrossRef](#)]
208. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; IEEE Press: Piscataway, NJ, USA, 2017; pp. 6517–6525. [[CrossRef](#)]
209. Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767. [[CrossRef](#)]
210. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934. [[CrossRef](#)]
211. Zheng, Z.; Wang, P.; Ren, D.; Liu, W.; Ye, R.; Hu, Q.; Zuo, W. Enhancing Geometric Factors in Model Learning and Inference for Object Detection and Instance Segmentation. *IEEE Trans. Cybern.* **2020**, *7*, 1–13. [[CrossRef](#)] [[PubMed](#)]
212. Yu, J.; Jiang, Y.; Wang, Z.; Cao, Z.; Huang, T. UnitBox: An Advanced Object Detection Network. In Proceedings of the 24th ACM International Conference on Multimedia, Amsterdam, The Netherlands, 15–19 October 2016. [[CrossRef](#)]
213. Rezatofghi, H.; Tsoi, N.; Gwak, J.Y.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2019. [[CrossRef](#)]
214. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU loss: Faster and better learning for bounding box regression. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12993–13000. [[CrossRef](#)]
215. Liu, S.; Huang, D.; Wang, Y. 2018. Receptive field block net for accurate and fast object detection. *arXiv* **2017**, arXiv:1711.07767. [[CrossRef](#)]
216. Zhao, Q.; Sheng, T.; Wang, Y.; Tang, Z.; Chen, Y.; Cai, L.; Ling, H. M2Det: A Single-Shot Object Detector Based on Multi-Level Feature Pyramid Network. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2019; Volume 33, pp. 9259–9266. [[CrossRef](#)]
217. Huang, R.; Pedoem, J.; Chen, C. YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018. [[CrossRef](#)]
218. Zhou, X.; Zhuo, J.; Krhenbühl, P. Bottom-Up Object Detection by Grouping Extreme and Center Points. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2020. [[CrossRef](#)]
219. Duan, K.; Bai, S.; Xie, L.; Qi, H.; Huang, Q.; Tian, Q. Centernet: Keypoint triplets for object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2019; pp. 6569–6578. [[CrossRef](#)]
220. Zhou, X.; Wang, D.; Krhenbühl, P. Objects as Points. *arXiv* **2019**, arXiv:1904.07850. [[CrossRef](#)]
221. Tian, Z.; Shen, C.; Chen, H.; He, T. FCOS: Fully Convolutional One-Stage Object Detection. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2020. [[CrossRef](#)]

222. Kong, T.; Sun, F.; Liu, H.; Jiang, Y.; Shi, J. FoveaBox: Beyond anchor-based object detector. *IEEE Trans. Image Process.* **2020**, *29*, 7389–7398. [[CrossRef](#)]
223. Ronneberger, O.; Fischer, P.; Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*; Springer International Publishing: Berlin/Heidelberg, Germany, 2015. [[CrossRef](#)]
224. Shen, Z.; Liu, Z.; Li, J.; Jiang, Y.; Chen, Y.; Xue, X. Object detection from scratch with deep supervision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 398–412. [[CrossRef](#)] [[PubMed](#)]
225. Li, Z.; Peng, C.; Yu, G.; Zhang, X.; Deng, Y.; Sun, J. Detnet: A backbone network for object detection. *arXiv* **2018**, arXiv:1804.06215. [[CrossRef](#)]
226. Zhu, R.; Zhang, S.; Wang, X.; Wen, L.; Shi, H.; Bo, L.; Mei, T. ScratchDet: Training Single-Shot Object Detectors from Scratch. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2019. [[CrossRef](#)]
227. He, K.; Girshick, R.; Dollár, P. Rethinking imagenet pre-training. In Proceedings of the IEEE International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2019; pp. 4918–4927. [[CrossRef](#)]
228. Shi, B.; Bai, X.; Belongie, S. Detecting Oriented Text in Natural Images by Linking Segments. *IEEE Comput. Soc.* **2017**, 2550–2558. [[CrossRef](#)]
229. Ma, J.; Shao, W.; Ye, H.; Wang, L.; Wang, H.; Zheng, Y.; Xue, X. Arbitrary-Oriented Scene Text Detection via Rotation Proposals. *IEEE Trans. Multimed.* **2018**, *20*, 3111–3122. [[CrossRef](#)]
230. Liao, M.; Shi, B.; Bai, X.; Wang, X.; Liu, W. Textboxes: A fast text detector with a single deep neural network. In Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2017.
231. Liao, M.; Shi, B.; Bai, X. TextBoxes++: A Single-Shot Oriented Scene Text Detector. *IEEE Trans. Image Process.* **2018**, *27*, 3676–3690. [[CrossRef](#)] [[PubMed](#)]
232. Jiang, Y.; Zhu, X.; Wang, X.; Yang, S.; Li, W.; Wang, H.; Fu, P.; Luo, Z. R2CNN: Rotational region CNN for orientation robust scene text detection. *arXiv* **2017**, arXiv:1706.09579. [[CrossRef](#)]
233. Liu, W.; Ma, L.; Chen, H. Arbitrary-oriented ship detection framework in optical remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *15*, 937–941. [[CrossRef](#)]
234. Xue, Y.; Sun, H.; Fu, K.; Yang, J.; Sun, X.; Yan, M.; 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*, 32.
235. Yang, X.; Sun, H.; Sun, X.; Yan, M.; Guo, Z.; Fu, K. Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network. *IEEE Access* **2018**, *6*, 50839–50849. [[CrossRef](#)]
236. Li, K.; Cheng, G.; Bu, S.; You, X. Rotation-insensitive and context-augmented object detection in remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 2337–2348. [[CrossRef](#)]
237. Ding, J.; Xue, N.; Long, Y.; Xia, G.; Lu, Q. Learning RoI Transformer for Oriented Object Detection in Aerial Images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2019; pp. 2849–2858. [[CrossRef](#)]
238. Gong, W.; Shi, Z.; Wu, Z.; Luo, J. Arbitrary-oriented ship detection via feature fusion and visual attention for high-resolution optical remote sensing imagery. *Int. J. Remote Sens.* **2021**, *42*, 2622–2640. [[CrossRef](#)]
239. Liu, L.; Pan, Z.; Lei, B. Learning a rotation invariant detector with rotatable bounding box. *arXiv* **2017**, arXiv:1711.09405. [[CrossRef](#)]
240. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective kernel networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2019; pp. 510–519. [[CrossRef](#)]
241. Huang, Z.; Wang, X.; Wei, Y.; Huang, C.; Wei, Y.; Liu, W. CCNet: Criss-Cross Attention for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**. [[CrossRef](#)]
242. Yuan, Y.; Wang, J. OCNet: Object Context Network for Scene Parsing. *arXiv* **2018**, arXiv:1809.00916. [[CrossRef](#)]
243. Lin, X.; Guo, Y.; Wang, J. Global Correlation Network: End-to-End Joint Multi-Object Detection and Tracking. *arXiv* **2021**, arXiv:2103.12511. [[CrossRef](#)]
244. Fu, J.; Liu, J.; Tian, H.; Li, Y.; Bao, Y.; Fang, Z.; Lu, H. Dual Attention Network for Scene Segmentation. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2020. [[CrossRef](#)]
245. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszko-reit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
246. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 10012–10022. [[CrossRef](#)]
247. Wu, J.; Leng, C.; Wang, Y.; Hu, Q.; Cheng, J. Quantized convolutional neural networks for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 July 2016. [[CrossRef](#)]
248. He, Y.; Zhang, X.; Sun, J. Channel pruning for accelerating very deep neural networks. In Proceedings of the IEEE International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2017; pp. 1389–1397. [[CrossRef](#)]
249. Gong, Y.; Liu, L.; Yang, M.; Bourdev, L.D. Compressing deep convolutional networks using vector quantization. *arXiv* **2014**, arXiv:1412.6115.
250. Molchanov, P.; Tyree, S.; Karras, T.; Aila, T.; Kautz, J. Pruning convolutional neural networks for resource efficient inference. *arXiv* **2016**, arXiv:1611.06440. [[CrossRef](#)]

251. Han, S.; Mao, H.; Dally, W.J. Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding. *arXiv* **2016**, arXiv:1510.00149.
252. Han, S.; Pool, J.; Tran, J.; Dally, W. Learning both weights and connections for efficient neural network. *Adv. Neural Inf. Process. Syst.* **2015**, *28*. [[CrossRef](#)]
253. Molchanov, P.; Tyree, S.; Karras, T.; Aila, T.; Kautz, J. Pruning convolutional neural networks for resource efficient transfer learning. *arXiv* **2017**, arXiv:1611.06440.
254. Shao, F.; Chen, L.; Shao, J.; Ji, W.; Xiao, S.; Ye, L.; Zhuang, Y.; Xiao, J. Deep Learning for Weakly-Supervised Object Detection and Object Localization: A Survey. *Neurocomputing* **2022**, 192–207. [[CrossRef](#)]
255. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [[CrossRef](#)]
256. Zoph, B.; Cubuk, E.D.; Ghiasi, G.; Lin, T.; Shlens, J.; Le, Q.V. Learning data augmentation strategies for object detection. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2020; pp. 566–583. [[CrossRef](#)]
257. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-end object detection with transformers. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2020; pp. 213–229. [[CrossRef](#)]
258. Zhang, H.; Li, F.; Liu, S.; Zhang, L.; Su, H.; Zhu, J.; Ni, L.M.; Shum, H. DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection. *arXiv* **2022**. e-prints.
259. Zhang, T.; Zhang, X. Injection of Traditional Hand-Crafted Features into Modern CNN-Based Models for SAR Ship Classification: What, Why, Where, and How. *Remote Sens.* **2021**, *13*, 2091. [[CrossRef](#)]
260. Zhang, T.; Zhang, X. HTC+ for SAR Ship Instance Segmentation. *Remote Sens.* **2022**, *14*, 2395. [[CrossRef](#)]
261. Zhang, T.; Zhang, X.; Ke, X.; Liu, C.; Xu, X.; Zhan, X.; Wang, C.; Ahmad, I.; Zhou, Y.; Pan, D.; et al. HOG-ShipCLSNet: A novel deep learning network with hog feature fusion for SAR ship classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–22. [[CrossRef](#)]
262. Zhang, T.; Zhang, X. Squeeze-and-excitation Laplacian pyramid network with dual-polarization feature fusion for ship classification in sar images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [[CrossRef](#)]
263. Zhang, T.; Zhang, X. A Full-Level Context Squeeze-and-Excitation ROI Extractor for SAR Ship Instance Segmentation. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
264. Zhang, T.; Zhang, X.; Liu, C.; Shi, J.; Wei, S.; Ahmad, I.; Zhan, X.; Zhou, Y.; Pan, D.; Li, J.; et al. Balance learning for ship detection from synthetic aperture radar remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *182*, 190–207. [[CrossRef](#)]
265. Zhang, T.; Zhang, X. A polarization fusion network with geometric feature embedding for SAR ship classification. *Pattern Recognit.* **2022**, *123*, 108365. [[CrossRef](#)]
266. Heiselberg, H. Ship-iceberg classification in SAR and multispectral satellite images with neural networks. *Remote Sens.* **2020**, *12*, 2353. [[CrossRef](#)]
267. Heiselberg, P.; Sørensen, K.A.; Heiselberg, H.; Andersen, O.B. SAR Ship–Iceberg Discrimination in Arctic Conditions Using Deep Learning. *Remote Sens.* **2022**, *14*, 2236. [[CrossRef](#)]
268. Heiselberg, H.; Stateczny, A. Remote sensing in vessel detection and navigation. *Sensors* **2020**, *20*, 5841. [[CrossRef](#)] [[PubMed](#)]