Article

Ship Detection in SAR Images Based on Steady CFAR Detector and Knowledge-Oriented GBDT Classifier

Shuqi Sun and Junfeng Wang *

School of Electronic, Information, and Electrical Engineering, Shanghai Jiao Tong University,
Shanghai 200240, China; sunshuqi@sjtu.edu.cn
* Correspondence: junfengwang@sjtu.edu.cn

Abstract: Ship detection is a significant issue in remote sensing based on Synthetic Aperture Radar (SAR). This paper combines the advantages of a steady constant false alarm rate (CFAR) detector and a knowledge-oriented Gradient Boosting Decision Tree (GBDT) classifier to achieve the location and the classification of ship candidates. The steady CFAR detector smooths the image by a moving-average filter and models the probability distribution of the smoothed clutter as a Gaussian distribution. The mean and the standard deviation of the Gaussian distribution are estimated according to the left half of the histogram to remove the effect of land, ships, and other targets. From the Gaussian distribution and a preset constant false alarm rate, a threshold is obtained to segment land, ships, and other targets from the clutter. Then, a series of morphological operations are introduced to eliminate land and extract ships and other targets, and an active contour algorithm is utilized to refine ships and other targets. Finally, ships are recognized from other targets by a knowledge-oriented GBDT classifier. Based on the brain-like ship-recognition process, we change the way of the decision-tree generation and achieve a higher classification performance than the original GBDT. The results on the AIRSARShip-1.0 dataset demonstrate that this scheme has a competitive performance against deep learning, especially in the detection of offshore ships.

Keywords: SAR; ship detection; CFAR; knowledge-oriented GBDT

1. Introduction

Synthetic Aperture Radar (SAR) has the advantages of working in all types of weather, all day, having a long observation distance, and having a small actual aperture. Compared with optical sensors, the weather will affect the monitoring effect when there are clouds and fog, while SAR imaging is not interfered with by extreme weather. SAR is increasingly used in both civilian and military applications.

In the field of marine monitoring and marine rescue, ship target detection based on SAR images is undoubtedly a very meaningful topic, which is not only conducive to the maintenance of maritime rights but also provides information support for the management and supervision of marine ships. With the further development of SAR technology, the resolution of SAR imaging has gradually improved. Ship target detection in SAR images has developed from point target detection to distributed target detection. This requires a higher performance in ship detection. However, existing problems also make ship detection in SAR images challenging. The echo signal is the vector sum of the echoes of many ideal target signals, resulting in random fluctuations in the scattering echo intensity of the target based on the scattering coefficient, and finally, the speckle noise will be manifested in SAR images. Moreover, the azimuth defocus of the moving ship on the high-resolution SAR image will seriously affect the recognition of the ship target. Therefore, ship detection in SAR images is challenging and meaningful.

According to [1], as early as the 80s of the 20th century, Lincoln Laboratory proposed a hierarchical and modular three-level processing process for SAR Automatic Target Recognition (ATR), which mainly includes three stages: detection, identification, and classification.

This three-level process has a clear idea and reasonable structure. It has become a general process widely used in SAR ATR systems. The first stage is called object detection or prescreening. The purpose is to extract small suspected areas that may contain targets of interest from the large scene in SAR images and eliminate areas that do not contain targets, but this stage will produce a large number of clutter false alarms. The second level is called target identification, which is essentially a dichotomous problem (distinguishing between target class and clutter class); it is a post-processing stage after the detection stage, the purpose of which is to retain the real target while removing the natural clutter false alarm and some artificial clutter false alarms and obtain the target Region of Interest (ROI). The third stage is called target classification/recognition, and through more complex processing such as feature extraction, feature selection, and classifier calculation of the ROI region, the artificial target false alarm is further eliminated, and the target category, model, and other information are finally obtained.

The constant false alarm rate (CFAR) is the most widely used method in SAR target detection. The CFAR algorithm requires that the target has a strong contrast relative to the background. The threshold is calculated by fitting the hypothetical distribution and setting a constant false alarm rate. Since the distribution of targets and clutters must overlap, false alarms and missed detections are inevitable. By setting a small false alarm rate and using the threshold to separate the targets from the clutters, a fairly good detection result can be obtained.

Typical CFAR algorithms include the two-parameter CFAR, global CFAR, and two-stage CFAR. The two-parameter CFAR means that the CFAR detector fits a certain distribution by determining the shape and scale parameters of the K distribution [2] or the mean and standard deviation of the Gaussian distribution [3]. Once the hypothetical distribution has been determined, the threshold calculation equation can be deduced. The global CFAR is an algorithm for quickly extracting suspicious targets, which detects targets through a global threshold instead of setting sliding windows for local statistics and calculations [4]. Combining the advantages of the local CFAR and global CFAR, a two-stage CFAR is proposed. The first stage uses the global CFAR to coarsely filter the ROI from the SAR image. The fine local CFAR is used to achieve fine detection [5]. Typical CFAR detectors include Cell-Averaging CFAR (CA-CFAR) detectors [6], Greatest of CFAR (GO-CFAR) detectors [7], Smallest of CFAR (SO-CFAR) detectors [7], and order-statistic CFAR (OS CFAR) detectors [8].

Traditional machine learning differs from deep learning. Machine learning utilizes statistical methods, linear algebra, and optimization algorithms to classify or predict unknown data. Features, once manually designed and extracted, can be used to train models with machine learning. Classic machine learning classifiers include random forests [9], support vector machines (SVMs) [10], naive Bayes classifiers [11], logistic regression [12], etc.

Compared to machine learning, deep learning can automatically extract and learn features from data. Common deep learning models include Convolutional Neural Networks (CNNs) [13] and Recurrent Neural Networks (RNNs) [14]. In terms of object detection, deep learning networks include YOLO [15], Fast R-CNN [16], Cascade R-CNN [17], and others. Although optical image object detection has achieved significant success, there are still many challenges in SAR image object detection. For example, the speckle noise and defocusing specific to SAR images greatly interfere with object detection.

Following the procedure of SAR ATR proposed by Lincoln Laboratory, our work raises a steady CFAR algorithm to achieve the location of targets at the first stage, utilizes ACM to obtain a finer ROI to extract more accurate features, and modifies GBDT to achieve a better classification performance at the second stage. Since this paper aims at detection rather than the classification of target types, the third stage is not within our consideration. The procedure of our work is shown in Figure 1.

Our work makes contributions in the following areas:

- The proposal of a steady CFAR algorithm, which has a steady performance with the influence of the bright targets rejected;
- The design and extraction of four useful features and the use of ACM to refine the ROI;

- The proposal of a knowledge-oriented GBDT classifier, which generates the base learner based on certain prior criteria.
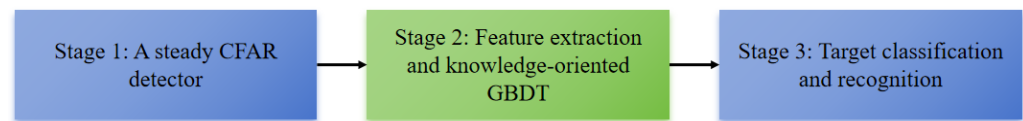
| Stage 1: A steady CFAR detector | → | Stage 2: Feature extraction and knowledge-oriented GBDT | → | Stage 3: Target classification and recognition |

Figure 1. Algorithm procedure diagram of our work.

The other sections of this paper are arranged as follows: Section 2 introduces the related work about our approach; Section 3 reveals the methodology and its principle; Section 4 shows the setting of the experiments and the results; and Section 5 concludes the whole work and looks into future work.

2. Related Work

Related work includes some typical CFAR algorithms, object detection in single-channel amplitude SAR images, and GBDT with its variants.

2.1. CFAR Algorithms

As mentioned above, the CFAR detects the targets of interest by giving an adaptive threshold. The CFAR calculates the distribution of the background clutters in the shifted window, and different CFAR detectors take different statistical processing strategies. The shifted window is usually hollow, which sets a protection window to remove the effect of target pixels on the parameter estimation of the clutter model. The other windows include the background window and the target window. The background window is used to calculate the clutter parameter. The detection window focuses on the pixel to be detected. The CFAR detector based on the hollow shifted window is shown in Figure 2.
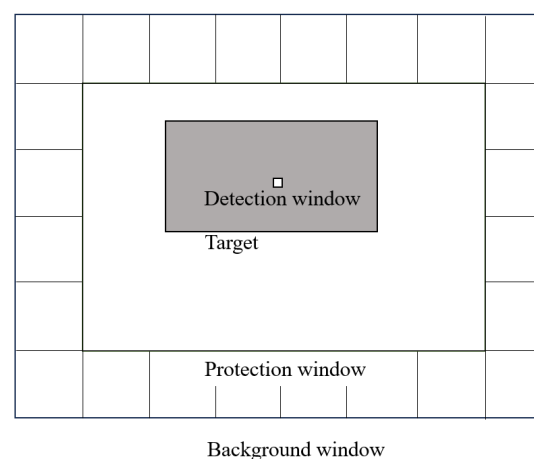
Figure 2. A diagram of the CFAR detector based on the hollow shifted window.

Researchers proposed a global CFAR to detect targets in the large scenes of SAR images, which improves the detection efficiency a lot compared to a typical two-parameter CFAR [4]. However, the detection precision cannot be guaranteed as the distribution model has severe mismatch when the scene varies in a complex way. Researchers designed a two-stage CFAR detector, which executes prescreening by a global threshold calculated by the Weibull distribution model at the first stage and executes fine detection by the K-distribution model [5]. This method represents an idea of combining a CFAR based on a global threshold and a CFAR based on a shifted window, which achieves a better performance in both precision and efficiency. Some researchers regarded the global CFAR as an initial clustering and classified the region of SAR images into targets and clutters by clustering [18–20].

2.2. Object Detection in Single-Channel Amplitude SAR Images

Besides the traditional hand-crafted feature extraction and classification, the great success of neural networks in general object detection promotes the use of deep learning in SAR image processing. Although the data-driven deep learning method has achieved some results in general image object detection, the characteristics of SAR images such as speckle noise and scattering interference and the lack of object-detection data samples in SAR images make the deep learning application challenging. Baseline models include YOLOv3 [21], YOLOv5 [22], SSD [23], Faster R-CNN [24], CenterNet [25], soft teacher [26], etc. Except for selecting a competitive network, some coping tricks and network components have a significant impact on detection performance. In the later comparison, we select YOLOv3, Faster R-CNN, and soft teacher as models to train and test. YOLOv3 is a real-time and effective object-detection algorithm. YOLOv3 utilizes a feature extraction network called Darknet-53 to construct a deeper network with residual units. Moreover, it uses Feature Pyramid Networks to achieve multi-scale detection. Considering the speed and precision of detection, YOLOv3 is a commonly used model. Faster R-CNN is an anchor-based detector with two stages. The Region Proposal Network is proposed to generate proposals for box classification and regression in the second stage, which accelerates the detection. Soft teacher is a semi-supervised object-detection method. It includes a teacher network and a student network. The teacher network is updated by conducting an Exponential Moving Average to the student network. Pseudo boxes are generated by the teacher network to be annotations of unlabeled data in the student network. A box-jittering mechanism is introduced to obtain pseudo boxes with higher quality. Soft teacher is a practical semi-supervised end-to-end network, and its performance is excellent.

The superpixel-based SAR ship-detection method is used to divide the image into regions with different grayscales, i.e., superpixels. Each superpixel is a representative region, which contains not only the color and orientation information of the bottom layer but also the structure information of the middle layer. By designing a certain algorithm, the superpixel is combined or otherwise processed to obtain the detection target. Pappas et al. combined superpixels with the sliding window constant false alarm method and replaced the background window and protection window with superpixels to make the extracted background clutter information more reasonable [27]. However, it is difficult for the superpixel-based SAR ship-detection method to detect the ships docked at the shore and the small ships in the image.

The visual saliency-based SAR ship-detection method uses image processing and computer vision technology to realize the automatic detection of ship targets by combining the saliency features of ships in Synthetic Aperture Radar (SAR) images. In this method, the SAR image is preprocessed, features are extracted at first, and then the visual saliency model is used to analyze the saliency of the image to highlight the saliency area of the ship target. Then, an appropriate algorithm is used to segment or binarize the saliency map to distinguish the ship target from the background. Finally, the detected ship targets are verified and filtered through post-processing steps to improve the accuracy and robustness of detection. The SAR ship-detection method based on visual saliency can effectively identify ship targets in SAR images and has a high detection performance and applicability. The detection methods of saliency SAR ships include airspace saliency detection methods represented by Itti [28] and frequency domain saliency detection methods represented by spectral residuals (SRs) [29].

2.3. GBDT and Its Variants

Gradient Boosting Decision Tree (GBDT) is an ensemble learning technique that sequentially builds decision trees to correct errors made by the existing ensemble. It is optimized by fitting the node values of new trees to the negative gradient of the loss function, using a learning rate parameter to control the contribution of each tree, and employing regularization techniques for model robustness [30].

GBDT has three powerful engineering realization libraries: CatBoost [31], Xgboost [32], and Lightgbm [33]. Some tried to enhance the classification performance of GBDT, while some made efforts to accelerate the serial calculation.

In addition, some scholars considered that the decision boundary of the decision tree is simply parallel to the attribute axis and modified the decision boundary to an oblique hyperplane, such as HHCART [34], OC1 [35], etc. Other machine learning structures can be added to the node of the decision tree to generate the hyperplane. Some impurity criterion is proposed to evaluate the distinguishing effect of the hyperplane. As the decision tree becomes an oblique decision tree, Gradient Boosting Oblique Decision Tree (GBODT) is a more accurate model in some cases.

3. Methodology

In contrast to the algorithm we proposed in the previous work [36], we follow the basic module design and processing procedure but change some details in the novel algorithm. Moreover, a knowledge-oriented GBDT structure is proposed, aiming at solving the specific problem and complying with prior knowledge. The adjusted algorithm flow chart is shown in Figure 3.
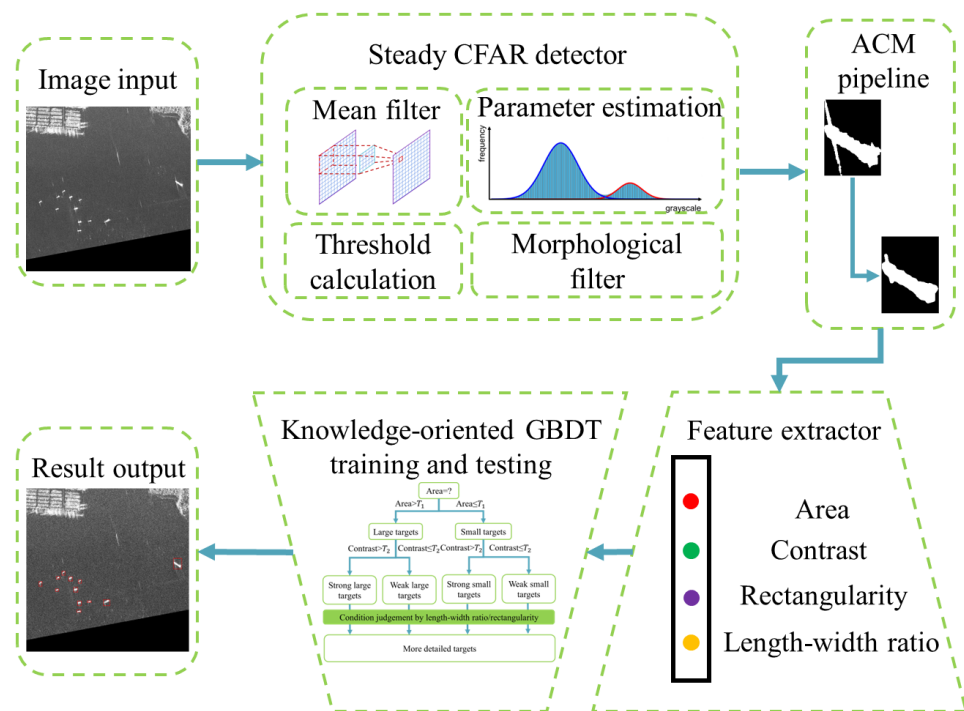

Figure 3. A flow chart of the proposed algorithm.

In short, after the input of the SAR images, they are fed into a steady CFAR detector, which eliminates the land at first and extracts the targets after obtaining the sea region in the image. Then we utilize the feature extraction module to extract four reliable features: rectangularity, length–width ratio, area, and contrast. Finally, the four features are trained and tested by the knowledge-oriented GBDT, and detection results are shown.

3.1. A Steady CFAR Detector

A CFAR detector is used to discover ships and other targets from the image. First, a mean filter is applied to the image. In the mean filtering, each pixel is assigned the mean of the pixels in its neighborhood. In most cases, each sea clutter pixel is assigned the mean of the sea clutter pixels in its neighborhood. According to the central limit theorem, the value

of this sea clutter pixel should be a Gaussian random variable whatever distribution it obeys originally. So, the probability density function of the sea clutter can be modeled as

$$f_c(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(x-\mu)^2}{2\sigma^2}), \tag{1}$$

where x is grayscale, μ is the mean of the Gaussian distribution, and σ is the standard deviation of the Gaussian distribution. The mean is estimated as the peak position of the histogram, and the standard deviation σ can be estimated by

$$\begin{aligned}\sigma &= \sqrt{\frac{\sum_{i=1}^{I} h_i(\mu-i)^2}{\sum_{i=1}^{I} h_i}} = \sqrt{\frac{\sum_{i=1}^{\mu-1} h_i(\mu-i)^2 + h_\mu(\mu-\mu)^2 + \sum_{i=\mu+1}^{I} h_i(\mu-i)^2}{\sum_{i=1}^{\mu-1} h_i + h_\mu + \sum_{i=\mu+1}^{I} h_i}} \\ &\approx \sqrt{\frac{2\sum_{i=1}^{\mu-1} h_i(\mu-i)^2}{2\sum_{i=1}^{\mu-1} h_i + h_\mu}} = \sqrt{\frac{\sum_{i=1}^{\mu-1} h_i(\mu-i)^2}{\sum_{i=1}^{\mu-1} h_i + 0.5h_\mu}},\end{aligned} \tag{2}$$

where i is grayscale, I is the maximum grayscale, and h_i is the histogram. Here, we use the peak and the left part of the histogram to estimate the parameters of the Gaussian distribution to remove the effect of land, ships, and other targets (Figure 4). The false alarm rate is written as

$$P_{fa} = \int_{x>T} f_c(x)dx. \tag{3}$$

where T is the detection threshold. The substitution of (1) into (3) yields

$$P_{fa} = \int_T^\infty \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(x-\mu)^2}{2\sigma^2})dx. \tag{4}$$

For a given P_{fa}, T is calculated according to (4). Typically, P_{fa} is chosen as 10^{-6}.
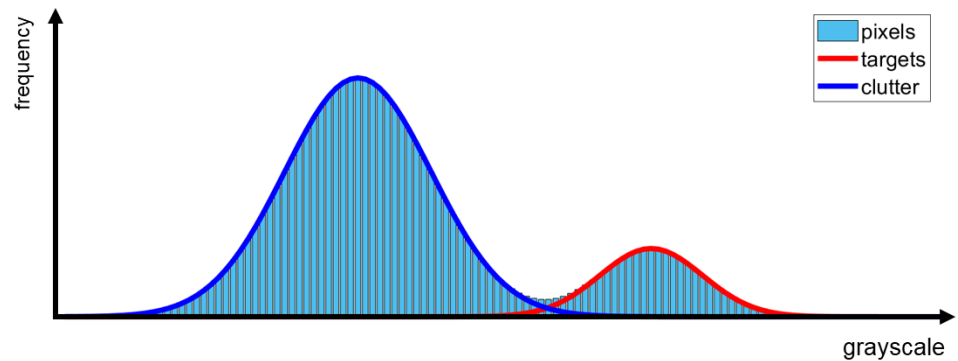

Figure 4. A sketch of the grayscale histogram of a sea image.

After the land extraction, we can generate a sea mask. We can eliminate the land by filtering the raw image with the sea mask. Then, we detect the targets in the target-extraction stage with the images without land. The morphological processing in the land-elimination stage follows the same procedure as mentioned in [36]. By the combination of erosion, dilation, hole filling, and logic reverse selection, the land connected to the edge of the images is eliminated. The similar operation helps to extract the targets.

ACM is worth mentioning. We utilize the snake model to refine the masks of regions and set a high smooth factor to reject the sidelobe clutter. Moreover, the initial mask of ACM is of great importance, so we take the open operation to the initial ROI with a small disk template to weaken the sidelobe. The disk template size is 5 in our experiments. The template is chosen as a disk for smoothness. In addition, the initial image slice is filtered by a mean filter with the same size as conducted to the whole image. The trend to either contract or expand is set to be more likely to expand as the initial mask is smaller than the original one for the open operation. However, the masks of tenuous clutters and

lax bright targets are usually processed to be full zero matrices under the open operation for their intensity distribution or their shape. Then, we recover the masks to the original ones to keep the basic shape information. The ACM data-processing pipeline is shown in Figure 5.
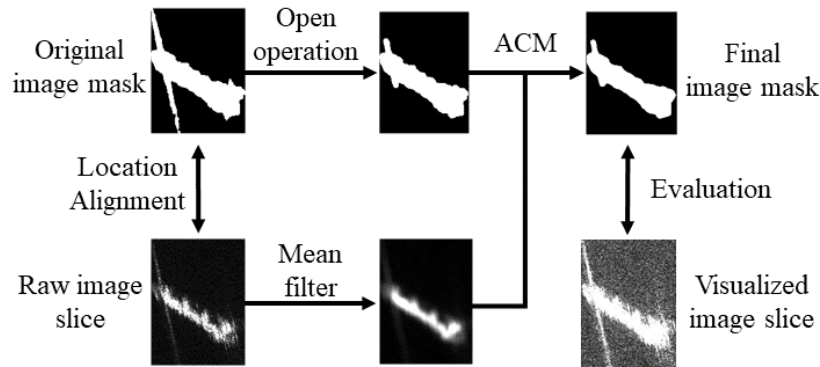

Figure 5. ACM data-processing pipeline.

3.2. Feature Extraction

After the fine processing of the ACM pipeline, four reliable hand-crafted features can be extracted from the refined targets. Three features are shape features, while the other one is grayscale features. The area feature of the targets includes information about the centralized range of the strong scattering points, representing the size of the targets. We can obtain the first feature as

$$Area = S_M, \tag{5}$$

where S_M is the target mask area. As for the grayscale feature, we keep the contrast feature in our previous work [36]. We calculate it by dividing the standard deviation of the target by its mean as

$$Contrast = \frac{\sigma_t}{\mu_t}, \tag{6}$$

where σ_t is the standard deviation of the target area and μ_t is the mean of the target area. A larger contrast reflects stronger fluctuation on the target, which indicates that the target is more likely to be a ship.

The other shape features are obtained by the rectangle fitting for the regions. In detail, we obtain the orientation of the ellipse with the same second-order moment as the region at first. Then, we rotate the major axis in the horizontal direction. In the next step, we calculate the standard deviation of the target and normalize it with its area. The grayscale-weighted second-order moment calculation has better robustness. The weight of the target refers to the sum of its grayscale. A typical result is shown in Figure 6.
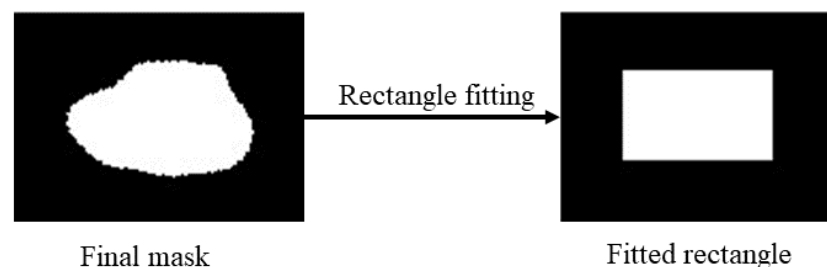

Figure 6. A typical rectangle fitting result.

The length–width ratio of the fitted rectangle can be obtained by

$$lwr = \frac{max(\sigma_x, \sigma_y)}{min(\sigma_x, \sigma_y)}, \tag{7}$$

where σ_x refers to the root of the second-order central moment along the x-axis and σ_y refers to the root of the second-order central moment along the y-axis. The length L and the width W can be described as

$$L = \sqrt{12 \times \sigma_x}, \tag{8}$$

$$W = \sqrt{12 \times \sigma_y}, \tag{9}$$

The centroid of the rectangle is determined by the centroid of the final mask. We define rectangularity as the intersection area of the target mask and the fitted rectangle divided by the geometric mean of the two region areas as

$$rect = \frac{S_i}{\sqrt{S_r \times S_t}}, \tag{10}$$

where S_i refers to the intersection area, S_r refers to the rectangle area, and S_t refers to the target area.

The appropriate *lwr* of ships ranges from 2.5 to 7.5, while the *lwr* of a ship can be close to one as sometimes the parts of ships are detected. Besides, the closer the *rect* is to 1, the more likely the target is a ship.

3.3. Knowledge-Oriented GBDT

As mentioned above, GBDT is a serial computation structure with each tree fitting the residual between the previous model prediction result and the true label. Different strategies are introduced to prevent overfitting, such as subsamples, dropout, and a small learning rate. We reconsider the generation criterion of the regression tree and set a customized attribute split order to make the regression tree structure closer to the ideal structure decided by prior knowledge.

According to the intensity of the targets, the targets are divided into strong and weak targets. According to the area of the targets, the targets are divided into large and small targets. Combining the two classifications, we can obtain four basic target types. As this GBDT is knowledge-oriented, the first two levels of the regression trees are meaningful. For example, after the area split at the first level and the contrast split at the second level, the weak and small targets are distributed to a node of the four nodes. The detection of weak and small targets is usually a difficult problem, and we can focus on solving problems in this field through this structure.

The structure of the novel regression tree is as follows. On the first level, we allocate the targets to the large target node and the small target node by splitting the parent node according to the area. Then, on the second level, we continue splitting the nodes into the strong target node and the weak target node. Furthermore, to prevent the contrast threshold from drifting, we set a fixed contrast threshold T_2 determined by knowledge. The contrast threshold is 0.79 in our experiments, which minimizes the contrast classification error on the training set. After the generation of the two levels, a combination of the *lwr* and *rect* is utilized to generate the deeper levels. The loss function is the Mean Square Error (MSE), which represents the purity of the split. In detail, it is the sum of the MSE of the left node and the MSE of the right node. The MSE loss function is shown as

$$\mathcal{L}(y, f(x)) = (y - f(x))^2. \tag{11}$$

The negative gradient error updated to the fitting truth of the next iteration can be expressed as

$$r = -\frac{\partial \mathcal{L}(y, f(x))}{\partial f(x)} = (y - f(x)), \tag{12}$$

where y is the fitting truth of the previous iteration and $f(x)$ is the current prediction value. By applying this loss function, we can obtain more precise prediction results by adding

the prediction value at each corresponding node in each tree, which reduces the residual after iteration.

Based on the novel regression tree, we keep the Gradient Boosting (GB) method unchanged to combine the regression trees with an additive model. The results of the experiments and ablation study show the effectiveness of each part in the modified GBDT. The novel regression tree model is shown in Figure 7. Our novel tree simulates the judgment of the human brain in ship detection and has strong interpretability compared with neural networks. The human brain usually judges a target in a specific procedure. First, the area of the target will be considered. Second, the grayscale fluctuations relative to the background will be regarded as an important feature. Then, the basic shape features such as rectangularity are utilized to confirm whether the target is a ship or not.
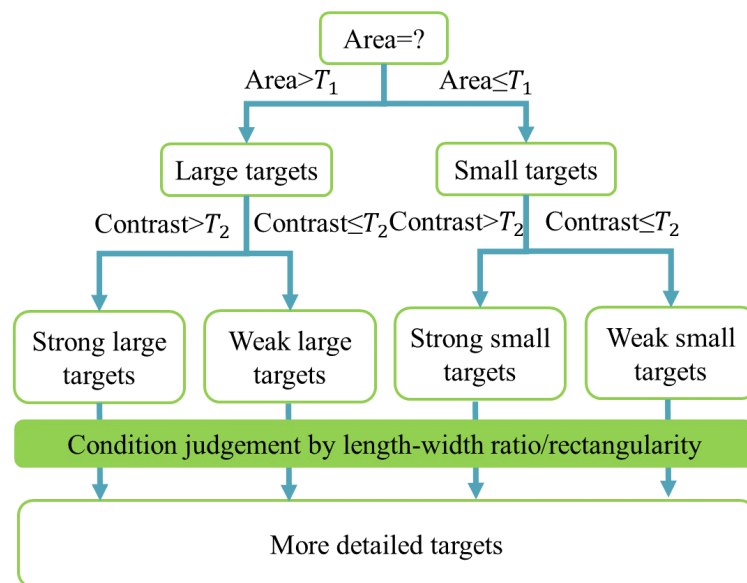

Figure 7. The novel regression tree structure.

4. Experiments and Results

4.1. Data Preparation

For the convenience of conducting the CFAR algorithm, we select the AIR-SARShip-1.0 dataset of [37] as our data source. Note that some images are rejected for their imaging quality and pollution level by noise. We split the indices into indices of the training set and indices of the testing set. The indices of the training and testing set are shown in Table 1. To construct a complete feature space, an appropriate split is required. Therefore, we finally chose the split in Table 1 to train our modified GBDT.

Table 1. Indices of split sets.

Indices of the Training Set	Indices of the Testing Set
6, 7, 8, 10, 12, 13, 16, 17, 20, 21, 22, 23, 24, 26, 27, 30	5, 14, 18, 19, 29, 31

As for the label preparation, we transform the bounding box information in the annotation files to the labels of the targets by checking if the target is in the bounding box. Our classification is binary. The label is 1 if the target is a ship, and the label is 0 if the target is not a ship. Moreover, the raw images with 16-bit depth are visualized by clipping them with three times their grayscale mean, and the detection results and some intermediate results will be shown on the visualized data.

4.2. Parameter Settings

Parameters for the steady CFAR detector are set as follows. As for the land-elimination stage of the steady CFAR detector, the mean filter size is 11×11, and the morphological filter size is 61×61. At the target-extraction stage, the mean filter size is 7×7 and the morphological filter size is 3×3. The hyperparameters for GBDT are discussed next.

4.3. Analysis of the Influence of Some Hyperparameters on the Model Performance

We choose the maximum depth, the number of base learners, the subsample rate, and the learning rate to be analyzed. All final accuracy results are obtained by calculating the mean of five experiments.

The influence of the maximum depth is shown in Table 2. As we can see, when the maximum depth is five, the model has the best performance. If the depth is too shallow, the classification ability of the model is lacking. If the depth is too deep, the model overfits. The number of base learners is 25, the subsample rate is 0.5, and the learning rate is 0.1.

Table 2. Analysis of the influence of maximum depth on the model.

Maximum Depth	Average Classification Accuracy
3	0.943
5	0.953
7	0.952
9	0.950

The influence of the number of base learners is shown in Table 3. As it indicates, an appropriate number of base learners helps to improve the accuracy of the model. The maximum depth is five, and other settings remain unchanged. We can infer from Table 3 that a sudden jump in accuracy from 0.793 to 0.905 occurs when the number of base learners changes from six to seven. It proves the necessity of a sufficient number of base learners, and the accuracy improves with a step effect at a certain number of base learners. With the increase in base learner number, the accuracy converges to an upper limit.

Table 3. Analysis of the influence of the number of base learners on the model.

The Number of Base Learners	Average Classification Accuracy
5	0.792
6	0.793
7	0.905
8	0.928
9	0.951
15	0.955
25	0.953
35	0.953

The influence of the subsample rate is shown in Table 4. As it indicates, the performance of the proposed GBDT is best when the subsample rate is 0.1. The subsample helps to prevent the model from overfitting, and the best subsample rate is relatively lower than that of the previous GBDT. It is inferred that the design of the first two levels and the contrast threshold determined by prior knowledge help the model to fit better on less data. Reasonably less data can relieve the overfitting. The superiority of the novel structure will be discussed in the next section. The number of base learners is 15, and other settings remain unchanged.

The influence of the learning rate is shown in Table 5. As we can see, the performance is best when the learning rate is 0.1. A lower learning rate needs more base learners and a higher learning rate leads to overfitting. The subsample rate is 0.1, and the other settings remain unchanged.

Table 4. Analysis of the influence of subsample rate on the model.

Subsample Rate	Average Classification Accuracy
0.05	0.928
0.1	0.967
0.3	0.958
0.5	0.957
0.7	0.938
0.9	0.932
1.0	0.925

Table 5. Analysis of the influence of learning rate on the model.

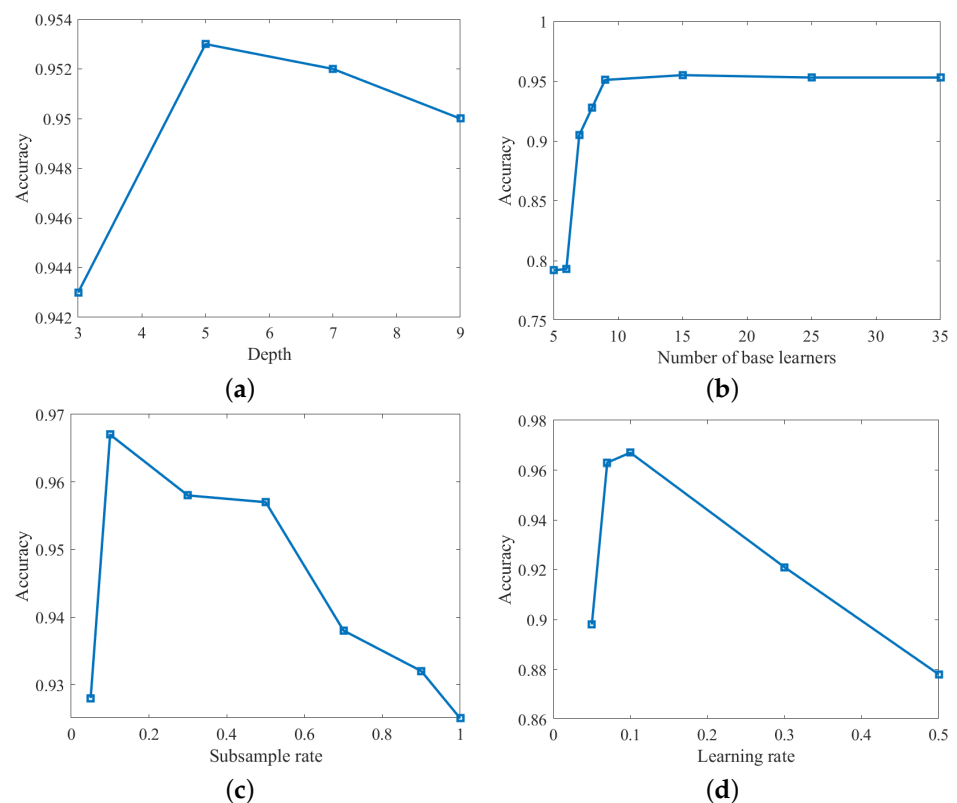| Learning Rate	Average Classification Accuracy
0.05	0.898
0.07	0.963
0.1	0.967
0.3	0.921
0.5	0.878

The data in Tables 2–5 are visualized in Figure 8.

Figure 8. The visualization of hyperparameter selection experiments. (**a**) The visualization of Table 2. (**b**) The visualization of Table 3. (**c**) The visualization of Table 4. (**d**) The visualization of Table 5.

Globally, there exists a trade-off in the selection of all the four factors. By conducting contrast experiments, we determine the best setting. We set the maximum depth to five, and the number of base learners is 15. Additionally, we set the subsample rate to 0.1, and the learning rate is 0.1. The following experiments are carried out based on this setting.

4.4. Contrast Experiments with the Original GBDT

We conducted contrast experiments with the original GBDT. The results show a stronger classification ability of our modified GBDT than the original one. An ablation study shows the necessity of our improvement. The accuracy is also obtained by calculating the mean of five experiments. The results are shown in Table 6.

Table 6. Ablation study results.

Gradient Boosting	Customized Split Order	Prior Contrast Threshold	Average Classification Accuracy
✗	✗	✗	0.933
✓	✗	✗	0.947
✓	✓	✗	0.957
✓	✓	✓	0.968

As we can see from Table 6, the classification accuracy gradually improves as the components are added. The baseline is a decision tree for binary classification. After introducing the gradient boosting method to implement ensemble learning, the residual gradually reduces and is fitted in each iteration. Considering the brain-like judgment process, we utilize a custom split order to split the nodes with obvious meaning at the first two levels, which achieves a better performance. To prevent the contrast threshold at the second level from drifting, we take a prior contrast threshold. The final result shows the improvement by adding this component. The significance of our improvements is verified and the knowledge-oriented GBDT achieves satisfactory results on the present issue. Aided by prior knowledge, we restrict the generation way of the novel GBDT classifier by useful standards, and the threshold calculation is needless at the second level of the regression, which accelerates the training process of our GBDT.

4.5. Contrast Experiments with the Advanced Techniques

Many scholars focus on the application of deep learning in the ship-detection field. Even if the data are limited and noises exist in SAR images, deep learning is powerful and full of possibilities. Our approach combines a classical machine learning method with a steady CFAR detector. The advantage of our work is high reliability and low training costs (compared to deep learning methods). Taking the training of the three deep learning models in the contrast experiments as an example, it takes several hours to achieve a good performance. However, the GBDT training of four low-dimensional features takes several seconds in this experiment. In addition, the training process of deep networks in large scenes is expensive, but our method is less sensitive to the size of the images. In offshore ship detection, the detection accuracy of our method is competitive with the advanced techniques.

The advanced techniques to be compared include YOLOv3 [21], Faster R-CNN [16], and soft teacher [26]. Among them, YOLOv3 and Faster R-CNN are supervised object-detection methods. Soft teacher is a practical semi-supervised object-detection method to overcome the problem that annotations are expensive and badly needed. In soft teacher, some images are annotated, while others are without labels. For the fairness of comparison, the proportion of data with the annotations is set to 0.1. The hyperparameters of all three methods are adjusted to obtain a better performance. SAR images are cropped into 512×512 slices according to the location of the annotations. Then, we conduct inference on the test images with big scenes. The optimizer is Stochastic Gradient Descent (SGD). As for the unsupervised method, the steady CFAR in our work can be regarded as a representative.

To compare the four methods quantitatively, we develop the following indices as

$$precision = \frac{TP}{TP + FP}, \tag{13}$$

$$recall = \frac{TP}{TP + TN}, \qquad (14)$$

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall}, \qquad (15)$$

where TP represents the true positive samples and FP represents the false positive samples. TN represents the true negative samples. Precision is related to false alarms and recall is related to missed detection. F_1 represents the F1 score, which is a harmonic mean of precision and recall. It is widely used in binary classification. The closer the three indices are to one, the better the detection performance is. Additionally, we calculate the three indices in all scenes and offshore scenes to verify the excellence of the proposed method.

The detection results of the four methods are presented in Figure 9. The result shows the effectiveness of our work, especially in the detection of offshore ships. In Figure 9, we achieved 100% precision and recall in this offshore scene. However, the inshore ships are ignored as it is considered a part of the land in the land-elimination stage of the steady CFAR detector. Therefore, our method works better in the offshore scenes. As we can judge, YOLOv3 and Faster R-CNN achieve better accuracy in inshore ship detection. There are some land false alarms in the detection result of Faster R-CNN. Soft teacher has a similar detection performance to our method in the offshore scene. As we can see in Figure 10, in the offshore scene, the sidelobe of the ship target is rejected by the ACM data-processing pipeline. Tenuous clutter is successfully classified into non-ship targets. The target that is similar to a ship on the bottom left of the image is identified as a non-ship target. It fully illustrates the powerful classification ability of our modified GBDT and the representativeness and accuracy of the data acquisition. Moreover, if the ship is on the edge of the image, it will also be eliminated by the morphological filter in the steady CFAR detector. A typical case is shown in Figure 11. As we can see, the ship on the bottom left of the image has very strong sidelobe interference, which leads to its disappearance in the sea mask. It is believed that with the movement of the imaging platform, the missed ship will be detected. The ships are regarded as the land in the calculation of inshore detection performance.
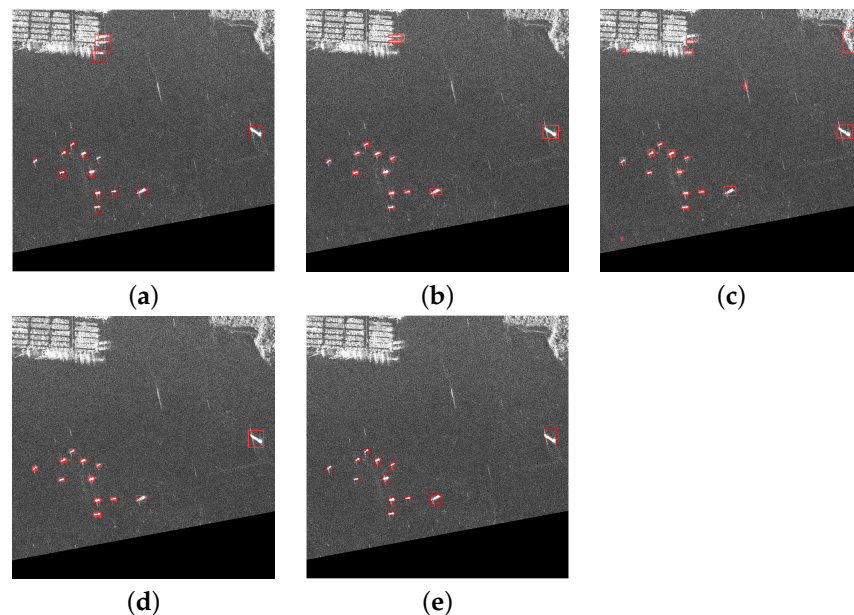

Figure 9. The comparison of detection results of some advanced techniques and ours. (**a**) The visualized result of ground truth. (**b**) The detection result of YOLOv3. (**c**) The detection result of Faster R-CNN. (**d**) The detection result of soft teacher. (**e**) The detection result of our work.
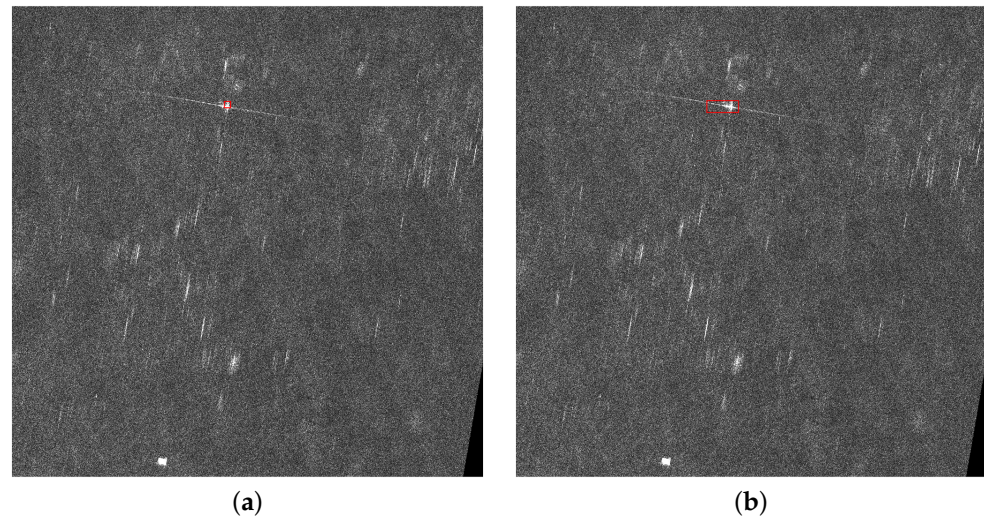
(**a**)   (**b**)

Figure 10. A detection result in the offshore scene. (**a**) The visualized result of ground truth. (**b**) The detection result of our work.
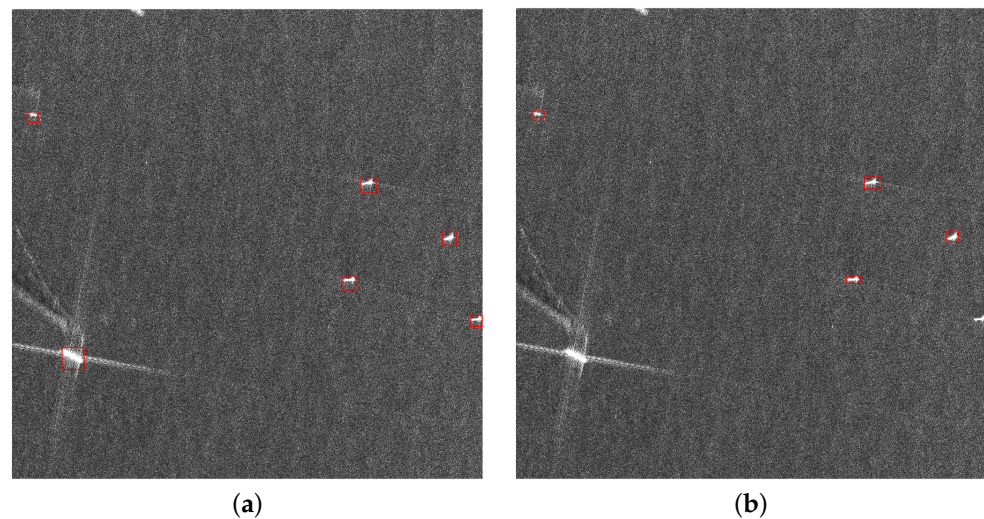

(**a**)   (**b**)

Figure 11. A detection result missing the ships on the edges. (**a**) The visualized result of ground truth. (**b**) The detection result of our work.

After obtaining the visualized data of the four methods, the three indices are calculated. The benchmark is shown in Table 7. M1 to M4 represent YOLOv3, Faster R-CNN, soft teacher, and our method, respectively.

Table 7. The benchmark of the four methods.

Model	Precision$_{All}$	Recall$_{All}$	F1 Score$_{All}$	Precision$_{In}$	Recall$_{In}$	F1 Score$_{In}$
M1	0.772	0.810	**0.791**	0.857	0.889	0.873
M2	0.565	**0.928**	0.702	0.711	**1.000**	0.831
M3	0.694	0.810	0.748	**1.000**	0.963	**0.981**
M4	**0.960**	0.619	0.753	0.960	0.889	0.923

As shown in Table 7, our method performs well, even in all scenes. For a few false alarms and a high detection rate, the F1 score of our method is high in inshore scenes. Although the method is free of neural networks and ignores the inshore ships, it is competitive with the deep learning methods. Among the deep learning methods, soft teacher has the highest F1 score in all scenes, and YOLOv3 has the highest F1 score in offshore scenes.

However, YOLOv3 and soft teacher generate some land false alarms in some scenes. Our method aims at rejecting the land false alarms by eliminating the land and rejecting the sea false alarms by training a powerful and robust classifier. The merit of our approach is its reliability and pertinence.

5. Conclusions

To conclude, we achieve a ship-detection approach combining the advantages of a steady CFAR detector and the knowledge-oriented GBDT. The CFAR detector is utilized to locate the latent targets and eliminate most land false alarms. The novelty of the steady CFAR is its method of modeling the clutter and the well-designed processing pipeline.

The knowledge-oriented GBDT is modified by the brain-like process of judging prior knowledge. The targets are divided into large and small targets according to their area. Then, they are divided into strong and weak targets according to their intensity. The contrast threshold is fixed to prevent the fitting of the data from losing prior knowledge. It is proved by the experiments that our improvement achieved remarkable results, and the whole method has a competitive performance with the deep learning methods.

Future work should include adding the inshore ship-detection module, the classification of the ships, and improving the stability and reliability of the whole procedure. The inshore ship-detection module is beneficial to the detection of all the ships on the scene. In addition, the further classification of the detected ships is beneficial to complete the whole SAR ATR process. A stabler classifier and more accurate feature extraction help to improve the robustness of the detector.

Author Contributions: Conceptualization, S.S. and J.W.; data curation, J.W.; formal analysis, S.S. and J.W.; funding acquisition, J.W.; investigation, S.S. and J.W.; project administration, S.S. and J.W.; resources, J.W.; software, S.S.; supervision, S.S. and J.W.; validation, S.S.; visualization, S.S. and J.W.; writing—original draft, S.S.; writing—review and editing, S.S. and J.W. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are contained within the article.

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

References

1. Du, L.; Wang, Z.; Wang, Y.; Di, W.; Lu, L. Survey of research progress on target detection and discrimination of single-channel SAR images for complex scenes. *J. Radars* **2020**, *9*, 34–54.
2. Kuttikkad, S.; Chellappa, R. Non-Gaussian CFAR techniques for target detection in high resolution SAR images. In Proceedings of the 1st International Conference on Image Processing, Austin, TX, USA, 13–16 November 1994; Volume 1, pp. 910–914.
3. Ai, J.; Luo, Q.; Yang, X.; Yin, Z.; Xu, H. Outliers-robust CFAR detector of Gaussian clutter based on the truncated-maximum-likelihood-estimator in SAR imagery. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 2039–2049. [CrossRef]
4. Gao, G.; Jiang, Y.M.; Zhang, Q.; Kuang, G.Y.; Li, D.R. Fast acquirement of vehicle targets from high resolution SAR images based on combining multi-feature. *Acta Electron. Sin.* **2006**, *34*, 1663–1667.
5. Xing, X.W.; Chen, Z.L.; Zou, H.X.; Zhou, S.L. A fast algorithm based on two-stage CFAR for detecting ships in SAR images. In Proceedings of the Asian-Pacific Conference on Synthetic Aperture Radar, Xi'an, China, 26–30 October 2009; pp. 506–509.
6. El-Darymli, K.; McGuire, P.; Power, D.; Moloney, C. Target detection in synthetic aperture radar imagery: A state-of-the-art survey. *J. Appl. Remote Sens.* **2013**, *7*, 071598. [CrossRef]
7. Zhao, M.; He, J.; Fu, Q. Survey on fast CFAR detection algorithms for SAR image targets. *Acta Autom. Sin.* **2012**, *38*, 1885–1895. [CrossRef]
8. Rohling, H. Radar CFAR thresholding in clutter and multiple target situations. *IEEE Trans. Aerosp. Electron. Syst.* **1983**, 608–621. [CrossRef]
9. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
10. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]
11. Rish, I. An empirical study of the naive Bayes classifier. In Proceedings of the IJCAI Workshop on Empirical Methods in Artificial Intelligence, Seattle, WA, USA, 4–10 August 2001; Volume 3, pp. 41–46.

12. Kleinbaum, D.G.; Dietz, K.; Gail, M.; Klein, M.; Klein, M. *Logistic Regression*; Springer: New York, NY, USA, 2002.

13. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

14. Zaremba, W.; Sutskever, I.; Vinyals, O. Recurrent neural network regularization. *arXiv* **2014**, arXiv:1409.2329.

15. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, realtime object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

16. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Indore, India, 10–12 September 2015; pp. 1440–1448.

17. 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, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6154–6162.

18. Pan, Z.; Gao, X.; Wang, Y.-F.; Fan, L.-J. Clustering-based target fast detection for SAR imagery. *Appl. Res. Comput.* **2008**, *25*, 2416–2419.

19. Wang, Y.H.; Han, C.Z. PolSAR image segmentation by mean shift clustering in the tensor space. *Acta Autom. Sin.* **2010**, *36*, 798–806. [CrossRef]

20. Liu, H.Y. Fast target detection for SAR images based on weighted Parzen-window clustering algorithm. In Proceedings of the 2010 International Conference on Communications and Intelligence Information Security (ICCIIS), Xi'an, China, 13–14 October 2010; pp. 164–167.

21. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.

22. Ting, L.; Baijun, Z.; Yongsheng, Z.; Shun, Y. Ship detection algorithm based on improved YOLO V5. In Proceedings of the 2021 6th International Conference on Automation, Control and Robotics Engineering (CACRE), Dalian, China, 15–17 July 2021; pp. 483–487.

23. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Part I 14; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 21–37.

24. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 1137–1149. [CrossRef] [PubMed]

25. 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, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6569–6578.

26. Xu, M.; Zhang, Z.; Hu, H.; Wang, J.; Wang, L.; Wei, F.; Bai, X.; Liu, Z. End-to-end semi-supervised object detection with soft teacher. In Proceedings of the IEEE/CVF iNternational Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 3060–3069.

27. Pappas, O.; Achim, A.; Bull, D. Superpixel-level CFAR detectors for ship detection in SAR imagery. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1397–1401. [CrossRef]

28. Itti, L.; Koch, C.; Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 1254–1259. [CrossRef]

29. Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–2 June 2007; pp. 1–8.

30. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]

31. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 6639–6649.

32. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

33. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 7–11.

34. Wickramarachchi, D.C.; Robertson, B.L.; Reale, M.; Price, C.J.; Brown, J. HHCART: An oblique decision tree. *Comput. Stat. Data Anal.* **2016**, *96*, 12–23. [CrossRef]

35. Murthy, S.K.; Kasif, S.; Salzberg, S.; Beigel, R. OC1: A randomized algorithm for building oblique decision trees. *Proc. AAAI* **1993**, *93*, 322–327.

36. Sun, S.; Wang, J.; Tan, D. Ship Detection in SAR Images Based on CFAR and Machine Learning. In Proceedings of the 2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Taizhou, China, 28–30 October 2023; pp. 1–6.

37. Sun, X.; Wang, Z.; Sun, Y.; Diao, W.; Zhang, Y.; Fu, K. AIR-SARShip-1.0: Highresolution SAR ship detection dataset. *J. Radars* **2019**, *8*, 852–886.