

Article

Infrared Ship Target Detection Based on Dual Channel Segmentation Combined with Multiple Features

Dongming Lu ^{1,2,*}, Jiangyun Tan ^{1,2}, Mengke Wang ^{1,2}, Longyin Teng ^{1,2}, Liping Wang ^{1,2} and Guohua Gu ^{1,2}

¹ School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; jiangyun@njust.edu.cn (J.T.)

² Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China

* Correspondence: e-mail@lu_dongming@aliyun.com

Abstract: In infrared images of the sea surface, apart from the complex background of the sea surface, there are often sky and island backgrounds. The disturbances caused by sea wind and the reflection of intense sunlight on the sea surface increase the complexity of the background, which seriously hinders the detection of targets. To achieve the detection of dark-polarity ship targets in such environments, a dual-channel threshold segmentation method based on local low-gray region detection and geometric features judgment is proposed in this paper. In one channel, adaptive threshold segmentation is performed on the low-gray regions of the acquired image and combined with geometric features to obtain a finer segmentation result. In the other channel, adaptive segmentation is performed on the preprocessed image, and potential backgrounds that may be finely segmented as targets are filtered out based on an area threshold. Finally, the results of the two channels are multiplied and fused to obtain an accurate segmentation result. Experimental results demonstrate that the proposed algorithm outperforms the comparison algorithm in subjective and objective evaluations. The proposed algorithm in this paper not only achieves a low false alarm rate but also exhibits a higher detection rate, and the average detection rate in the test sequence surpasses 95%.

Keywords: complex sea background; infrared ship detection; low grayscale region detection; dual-channel segment



Citation: Lu, D.; Tan, J.; Wang, M.; Teng, L.; Wang, L.; Gu, G. Infrared Ship Target Detection Based on Dual Channel Segmentation Combined with Multiple Features. *Appl. Sci.* **2023**, *13*, 12247. <https://doi.org/10.3390/app132212247>

Academic Editor: Yao Mao

Received: 18 October 2023

Revised: 6 November 2023

Accepted: 10 November 2023

Published: 12 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ships play a crucial role as the primary carriers of human maritime activities, and the safeguarding of sovereignty and maritime rights in territorial waters relies on the effective monitoring of ship targets on the sea surface. The accurate detection of ship targets at sea is essential in various contexts, including military operations for identifying potential threats, civilian applications such as ship monitoring, and search and rescue operations for shipwrecks. The methods for detecting ship targets in the marine environment can be broadly categorized into shore-based and space-based approaches. Shore-based detection methods encompass infrared imaging and visible light imaging. On the other hand, space-based detection methods include remote sensing techniques such as visible light imaging, infrared imaging, and synthetic aperture radar (SAR) imaging [1]. Visible light imaging offers high resolution and detailed information [2] but it is limited in its ability to operate effectively under all weather conditions.

Despite extensive research on the remote sensing-based detection of ship targets on the sea surface [3–6], several challenges still need to be addressed. Remote sensing images are susceptible to cloud interference, and broken cloud patterns often resemble actual targets resulting in a high rate of false alarms. Additionally, remote sensing images are more suitable for open sea scenes due to their long imaging distances, which often leads to missed detections of small ship targets. False detections can also occur due to the presence of islands, reefs, and shore-based structures. In the case of remote sensing

infrared images, the temperature difference between day and night often leads to grayscale inversions between the background and the target. These challenges significantly limit the applicability of existing algorithms in practice.

Deep learning-based object detection algorithms in the mainstream are primarily classified into two-stage algorithms and one-stage algorithms. Representative algorithms for the two-stage approach include R-CNN [7] and faster R-CNN [8]. These algorithms involve candidate region extraction, candidate region classification, and regression as the main processes. On the other hand, one-stage algorithms, exemplified by the SSD [9] and YOLO series [10], transform object detection into a regression problem for determining target position and category information. This approach is faster than the two-stage algorithms. Based on deep learning techniques, numerous researchers have made improvements to the respective algorithms in their specific research fields to detect ship targets at sea [11–14]. These advancements have yielded better detection results compared to comparative algorithms. However, deep learning necessitates a significant amount of training samples, while the available offshore ship datasets are limited in size. Furthermore, the complex sea surface background poses challenges in obtaining models with wide applicability.

To address the limitation of deep learning algorithms requiring a large number of samples for model training, deep transfer learning techniques can be employed to mitigate the scarcity of infrared target datasets for ships at sea. For instance, Wang et al. [15] utilized a significant amount of simulated infrared ship images for initial model training, followed by fine-tuning the trained model using a smaller quantity of actual infrared ship data. While this approach helps alleviate the demand for real infrared ship datasets, it still necessitates substantial computational resources. Additionally, the radiometric characteristics of simulated images may not entirely align with those of actual ship images, thereby increasing the complexity of the transfer learning process.

Infrared imaging technology is extensively utilized for target detection due to its compact size, simple structure, high concealment, and ability to operate in all weather conditions. Numerous scholars have made significant contributions to the field of infrared ship target detection, resulting in notable achievements in various applications. Traditional algorithms for infrared ship target detection at sea can be categorized into single-frame and multi-frame detection based on time series. In the context of multi-frame detection methods, researchers such as Lin et al. [16] and Wang et al. [17] have successfully suppressed interference clutter by leveraging the temporal characteristics of the image sequence. The target remains relatively stable, while the interference clutter exhibits fluctuation patterns, enabling effective target detection. However, this approach limits adaptability in scenarios with fixed interferences, such as islands in the sea surface background. Additionally, the computation of multi-frame images is time-consuming, and it cannot detect targets that lack continuous multi-frame information.

Regarding single-frame detection algorithms, there are relatively more approaches available. Different researchers have analyzed and described infrared ship targets from various perspectives, including the grayscale difference between targets and backgrounds, the combination of multiple target features, frequency domain characteristics of images, and saliency maps of images. Based on these analyses, researchers have designed different algorithms to achieve target detection.

A detection method based on multi-resolution difference filtering combined with a selective search algorithm was proposed by Xiu et al. [18]. This method utilizes the results of multi-resolution difference filtering and a selective search algorithm to obtain the detection outcome. However, this algorithm is typically suitable for scenarios with simple backgrounds, uniformity, and significant grayscale differences between targets and backgrounds. In the presence of complex backgrounds such as bright and dark patches on the sea surface, the effectiveness of multi-resolution difference filtering may deteriorate significantly.

Xing et al. [19] proposed a detection method based on multi-scale local edge gradients, leveraging the characteristic distribution of sea clutter and the sea horizon, which typically

exhibit significant vertical gradients and nearly zero horizontal gradients. In contrast, ships of specific sizes show substantial gradients in horizontal and vertical directions. To achieve accurate detection, this method relies on an effective preprocessing technique with exceptional clutter suppression capabilities. However, it is worth noting that currently, there is no algorithm available that can effectively suppress the abundance of bright and dark fluctuating sea clutter under sunlight conditions.

Xu et al. [20] employed the mean shift algorithm to filter an image and then calculated the mean gradient projection in the vertical direction to obtain the regions of interest for ship targets. Finally, they applied a row-column mean adaptive threshold segmentation algorithm to segment the infrared ship targets against a background with intense sunlight reflection. Although this method demonstrated good detection results in the tested scenarios, it relies on the accurate extraction of the locating rows. If the sea surface exhibits intense and widespread sunlight reflection, the estimation of locating rows based on gradients can be prone to errors.

Based on the analysis of the geometric features, grayscale statistical features, and texture features of local regions in the sky, waves, and targets, Hu et al. [21] proposed a ship target recognition technique, which utilized a combination of multiple features. Although the use of principal component analysis reduced the dimensionality of the feature vectors, the training and classification process still required considerable time. Additionally, in scenarios with fish-scale waves, this method exhibited a higher false alarm rate.

Liang et al. [22] proposed a segmentation method based on energy maps by leveraging the thermal infrared characteristics of infrared images at sea. This method utilizes the energy difference between the background and targets. It can effectively segment targets when there is minimal background interference. However, in scenarios with large waves and significant grayscale variations in clutter, the practical segmentation performance of this method may not be ideal.

Fractional Fourier transform, also known as the fractional order Fourier transform, is a mathematical transform that generalizes the classical Fourier transform. It provides a rich representation of spatial frequency information. Zhou et al. [23] utilized the differences in fractional Fourier frequencies between the target and the background to localize the vertical position of the target. This method showed good performance in locating bright targets but yielded poor results for dark targets. Furthermore, they discovered that modeling the fluctuating sea background in the frequency domain using Gaussian distributions was more stable, compared to modeling it in the spatial domain using pixel values. Therefore, they proposed a method [24] that utilizes Gaussian distribution to model the background in the Fourier domain. However, this method has numerous parameters, most of which are non-adaptive, making it challenging to apply in practical scenarios.

In terms of computing target saliency maps, both the image feature algorithm proposed by Hou et al. [25] and the visual attention model (VAM) proposed by Itti et al. [26] have achieved good results. Based on these methods, many researchers have proposed using them to assist in target detection under complex background conditions in maritime environments. Liu et al. [27] used the algorithm proposed by Hou et al. to compute the saliency regions of the image after background filtering. By combining multiple shape features and iteration threshold segments, they achieved target segmentation. Wang et al. [17] employed the visual attention model to obtain salient regions in consecutive frames. They combined the spatial saliency of the target with temporal consistency to distinguish real targets from salient interference regions, enabling ship detection under strong sunlight reflection on the sea surface. Dong et al. [28] improved the visual attention model for target detection in infrared backlight environments and successfully detected ship targets under backlight conditions. Although the aforementioned methods can achieve good results in ship detection under strong daylight conditions, they are prone to false alarms when there are large areas of waves on the sea surface or bright/dark bands traversing the image.

Li et al. [29] proposed a detection method based on morphological reconstruction and multi-feature analysis. They first performed grayscale morphological reconstruction on the original image and calculated and fused intensity feature maps. This approach suppressed sea clutter and enhanced the targets, enabling candidate target segmentation. Multiple features were then used to restrict and identify the candidate targets. Although this algorithm can recognize bipolar targets, it cannot effectively identify targets when there is interference between bright or dark patches of similar size to the target region, leading to certain false alarms. Similarly, utilizing saliency detection, Mo [30] proposed a dual-channel image separation algorithm based on combining local grayscale dynamic range saliency mapping for nighttime maritime ship target detection. This algorithm can effectively detect small targets in nighttime ocean images.

In summary, although there is a considerable amount of research on infrared ship target detection in maritime environments, most of it focuses on relatively simple scenes. There is a lack of algorithms specifically designed for detecting ship targets in complex backgrounds such as intense sunlight reflection, and the robustness of existing algorithms is generally limited. Under strong sunlight, the phenomenon of glare on the sea surface is often severe. After infrared imaging, there may be areas of high brightness or bright bands in the sea clutter, making it difficult to distinguish targets from the background clutter. Additionally, the disturbance caused by sea wind and the fluctuation of sea waves results in irregular patches of various sizes and grayscale values in infrared images. If there are cloud formations in the sky, the image may contain dark areas or dark bands. Moreover, due to the diverse reflective properties of ships and the unknown polarity of imaged targets, targets often exhibit dark polarity in the presence of glare. Furthermore, factors like islands and waves in the background further increase the challenge of correctly detecting infrared ship targets in such complex maritime environments.

In this paper, the research focuses on a complex sea surface scene consisting of light and dark bands, as well as island and reef backgrounds on a bright sea. The target of interest is a dark polar patch of a specific size on an imaged ship. Traditional infrared ship target detection methods typically rely on grayscale and gradient analysis to identify differences between the target and the background, enabling target detection. However, in the scenario described in this paper, there are often interference regions that resemble the target, leading to a high false alarm rate with these conventional methods. To address this challenge, this paper proposes a dual-channel threshold segmentation method based on local low-gray region detection and geometric feature analysis. Figure 1 illustrates the overall algorithm flowchart. Experimental results demonstrate that the proposed algorithm effectively detects dark polar ship targets in complex sea surface environments.

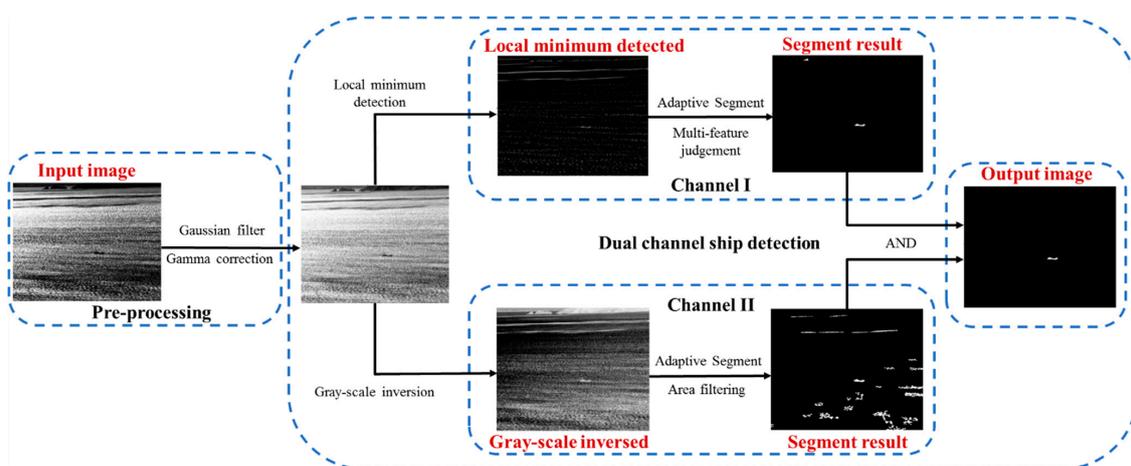


Figure 1. Algorithm flowchart.

2. Analysis of Image Characteristics

As depicted in Figure 2, images (a) to (d) illustrate a set of representative complex sea-surface infrared images. The uniformly dark regions enclosed within the red boxes represent the actual targets. In image (a), the background consists of a simple sea surface and the sky. The boundary between the sea surface and the sky is well-defined, and the strong sunlight results in a relatively uniform grayscale distribution across the sea surface, facilitating clear contrast with the targets. Consequently, separating the targets from the background is relatively straightforward, with the primary interference being two low-gray-level regions proximate to the target, exhibiting similar grayscale values. In contrast, the scenes depicted in images (b) to (d) are notably more complex. Due to the influence of sea winds and sunlight, the grayscale distribution of the sea surface becomes highly uneven. Various regions, both bright and dark, of different sizes emerge, and certain dark stripes exhibit grayscale values similar to the target. Moreover, significant areas of islands or landmasses with grayscale values resembling the target are present. In such scenes, conventional threshold segmentation or gradient-based edge detection methods become inadequate for ship target detection.

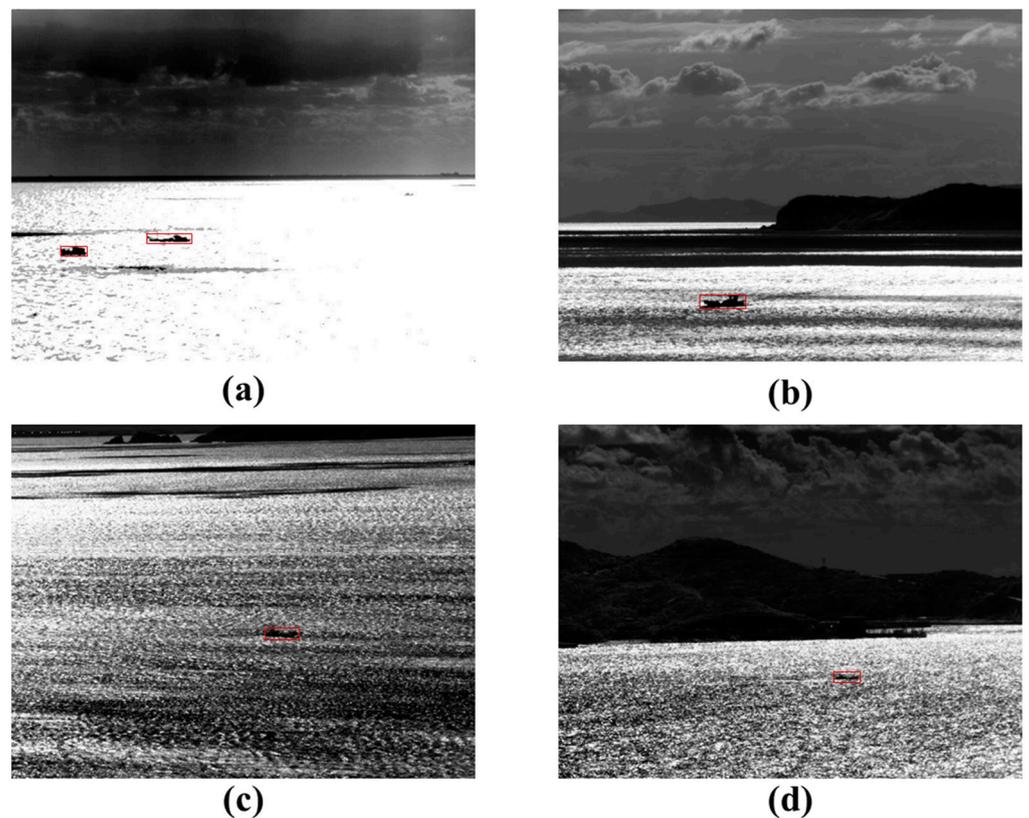


Figure 2. Typical complex sea-surface infrared images (a–d).

Figure 3 shows the histograms corresponding to the images in Figure 2. It can also be observed from the histograms that the low-gray regions where the targets are located do not exhibit distinct peak points. In other words, there is a gradual grayscale transition between the targets and certain background areas. Simple threshold segmentation can only accomplish the separation of brightness and darkness in the image. Further discrimination and separation are required to detect the targets within the obtained local dark regions that encompass both the targets and the background.

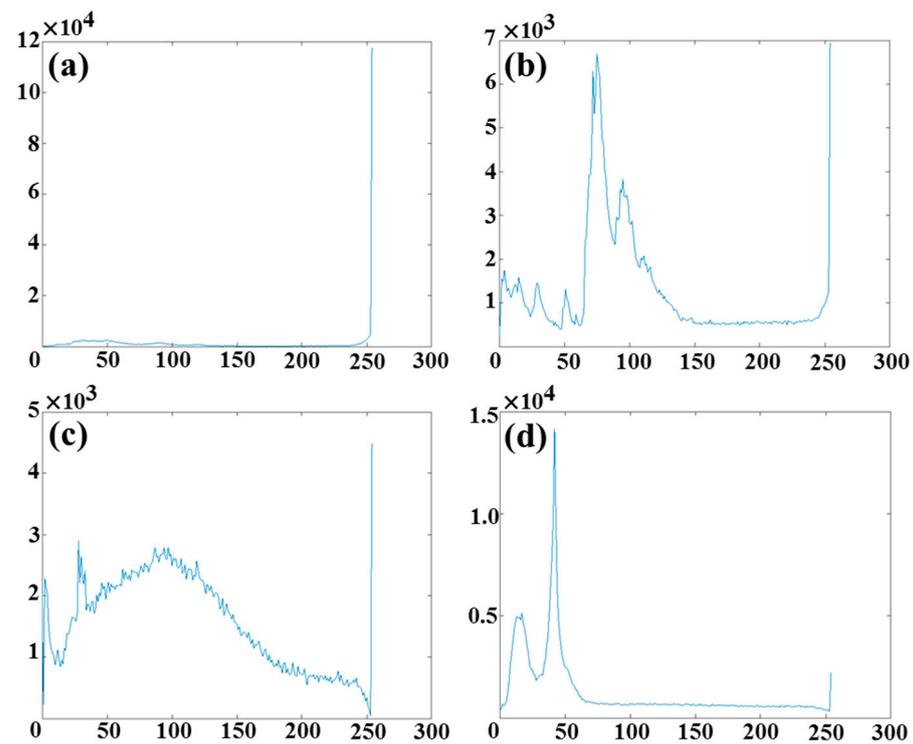


Figure 3. The histograms corresponding to images (a–d).

Similarly, Figure 4 represents the sum of the horizontal row gradients and Figure 5 indicates the sum of the vertical column gradients for each image in Figure 2. Figures 6 and 7 show the sum of the column gradients for each row in the horizontal direction and the sum of the row gradients for each column in the vertical direction, respectively. The calculation of the horizontal and vertical gradients follows Formulas (1) and (2), while Formulas (3) and (4) illustrate the computation of the column gradients for each row in the horizontal direction and the row gradients for each column in the vertical direction, respectively. Formulas (5) and (6) calculate the overall horizontal gradient and vertical gradient, respectively.

$$gradx(i, j) = |g(i + 1, j) - g(i, j)| \tag{1}$$

$$grady(i, j) = |g(i, j + 1) - g(i, j)| \tag{2}$$

$$G_x(i) = \sum_{j=1}^{N-1} grady(i, j) \tag{3}$$

$$G_y(j) = \sum_{i=1}^{M-1} gradx(i, j) \tag{4}$$

$$G_{xx}(i) = \sum_{i=1}^{M-1} gradx(i, j) \tag{5}$$

$$G_{yy}(j) = \sum_{j=1}^{N-1} grady(i, j) \tag{6}$$

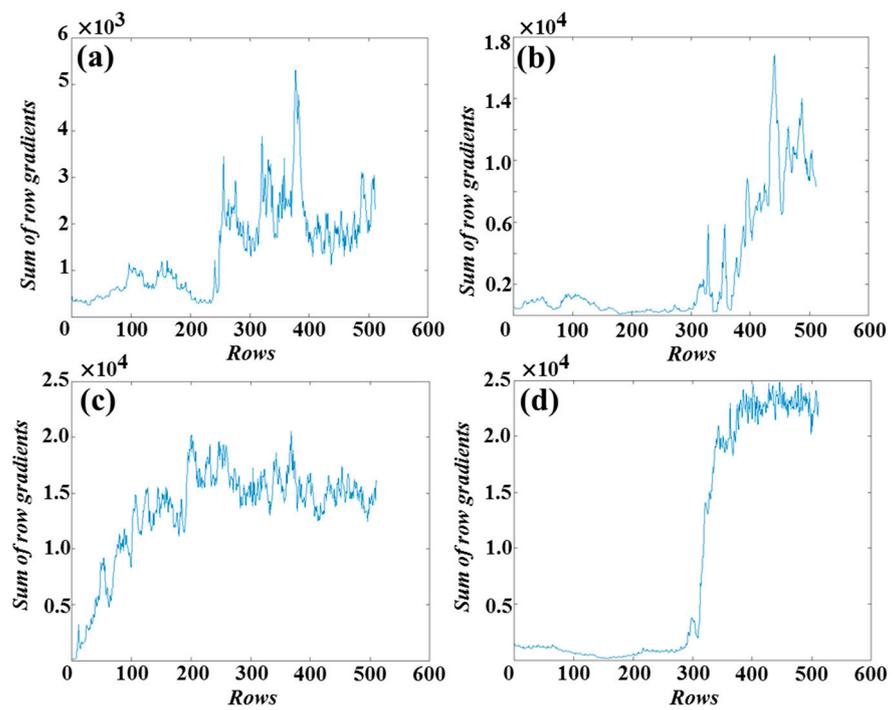


Figure 4. The sum of row gradients corresponding to images (a–d).

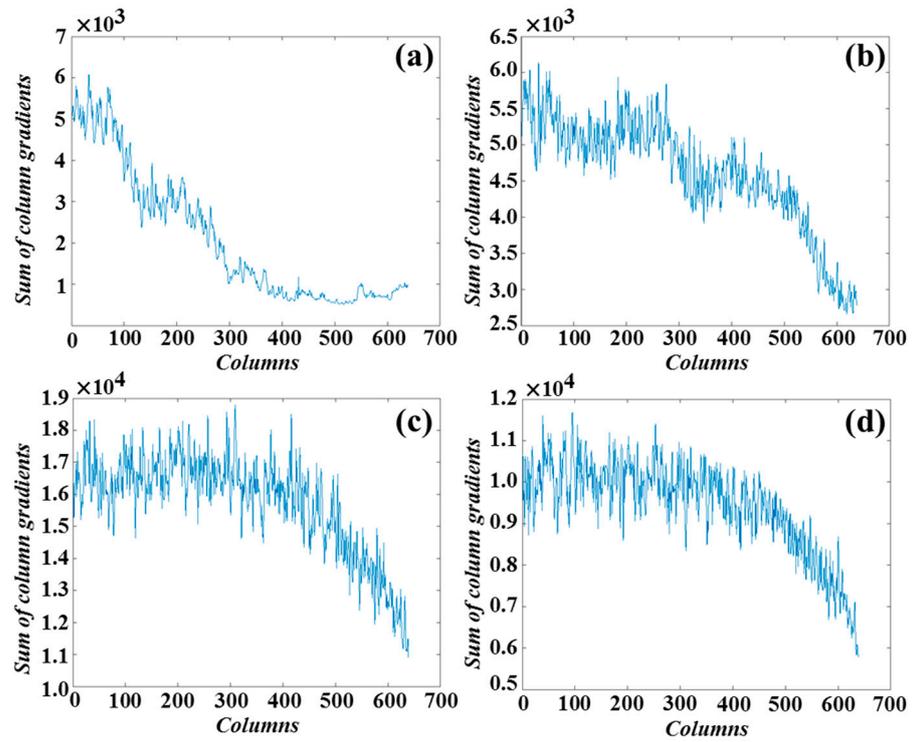


Figure 5. The sum of column gradients corresponding to images (a–d).

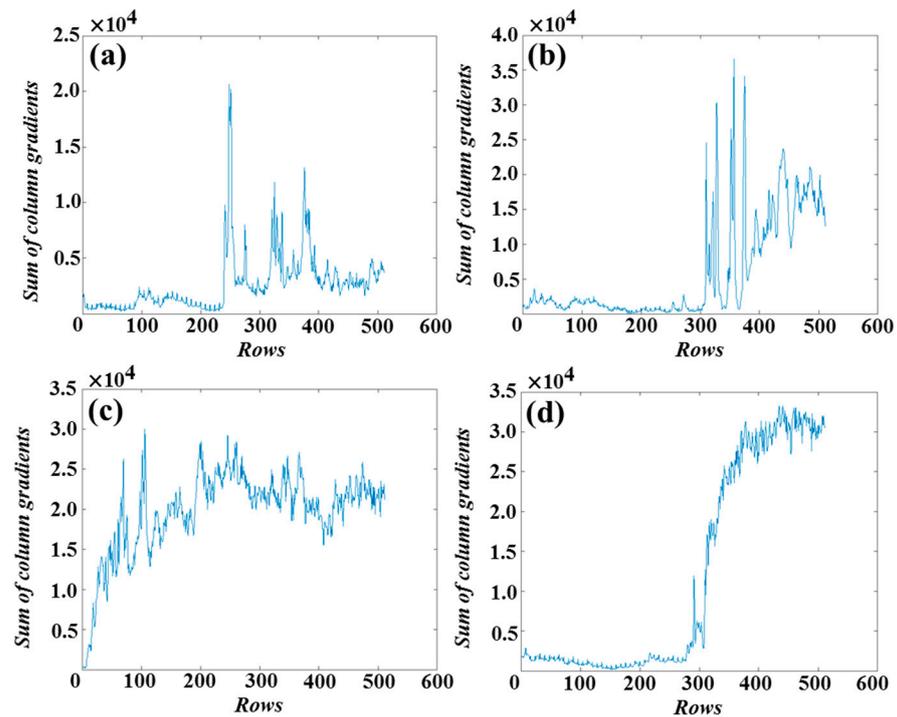


Figure 6. The sum of column gradients for each row in the horizontal direction corresponding to images (a–d).

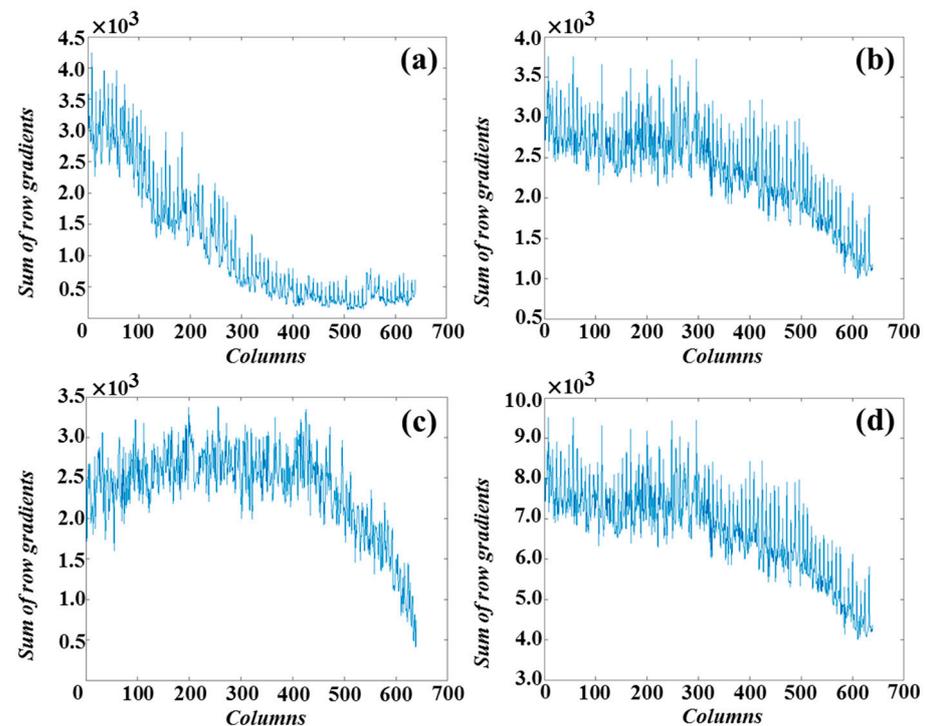


Figure 7. The sum of row gradients for each column in the vertical direction corresponding to images (a–d).

In the provided formulas, M and N represent the height and the width of the image, respectively. (i, j) denotes the coordinates of a pixel, $g(i, j)$ represents the grayscale value of the pixel, $grad_x$ and $grad_y$ are the gradient matrices computed in the horizontal and vertical directions, respectively. G_x represents the sum of vertical direction gradients for

each row, and G_y represents the sum of horizontal direction gradients for each column. G_{xx} represents the sum of horizontal gradients for each row, and G_{yy} represents the sum of vertical gradients for each column.

From Figure 4, it can be seen that in horizontally uniform regions with consistent grayscale values, the sum of row gradients is relatively small, such as the areas representing the sea-sky interface and horizontally distributed dark bands. However, the overall distribution of row gradient sums is not uniform. Except for the noticeable transition region in Figure 4d, the remaining images do not exhibit significant variations, making it difficult to segment the images based on row gradient sums. In the vertical direction, the grayscale distribution of the images is even more uneven. Whether we examine the column gradient sums in Figure 5 or the sum of row gradients in the vertical direction in Figure 6, the gradient sums are generally large and exhibit a chaotic distribution, lacking separability. Figure 6 reveals that there are rows with the maximum sum of vertical gradients in certain images, allowing for the detection of the sea-sky interface or the approximate row position of the targets. However, for the sea surface portions, the sum of the vertical gradients for each row is generally large, making it insufficient to differentiate between the targets and the noise caused by waves.

Most image saliency analyses start from the image's grayscale and gradients. But based on the analysis above, saliency maps calculated using grayscale or gradients will inevitably contain much interference. In fact, when there are regions in the image that are brighter or darker than the target, the computed saliency map may not include the target.

However, from the examples of typical infrared images of the sea surface, we can observe that within a certain imaging distance, the grayscale distribution of the target area is relatively uniform and mostly approximates a rectangular shape. Although there are some regions of interference with an even grayscale, the interference areas are irregular and constantly changing in size and shape due to the continuous fluctuation of the waves.

Based on the analysis above, considering the low grayscale characteristics of the targets, this paper proposes a method based on detecting local low grayscale regions. Specifically, a convolutional kernel with a negative value at its center is used to traverse the preprocessed image, aiming to identify regions with low gray levels. Subsequently, adaptive threshold segmentation is applied to these identified regions, leading to refined segmentation outcomes by incorporating shape, area, and other features. Furthermore, given the substantial grayscale difference typically observed between the sky background or land terrain background and the boundary of the sea surface, the proposed convolutional template may yield relatively high grayscale values in this region, introducing interference with the targets. Conversely, the grayscale values across the entire mountain or sky background region are similar. By performing adaptive segmentation on the preprocessed image, a connected region representing the mountain or sky background, which is significantly larger than the target, can be obtained. Therefore, by considering the area feature, the mountain or sky background can be eliminated to obtain a preliminary segmentation result. Consequently, by combining the refined segmentation result with the preliminary segmentation result, accurate detection outcomes can be achieved.

3. Detection Algorithm

3.1. Image Preprocessing

It is necessary to appropriately preprocess an image to improve the effectiveness of subsequent operations, before performing operations such as convolution and segmentation on it. From the typical infrared images shown in Figure 2, it can be seen that there are many isolated bright and dark spots on the complex sea surface under sunlight. In this paper, a Gaussian filter is used to smooth the image, since Gaussian filtering is a classic method for dealing with such salt-and-pepper noise. Additionally, gamma correction is employed to adjust the grayscale of the image and enhance its visibility. The formula for gamma correction is as follows:

$$f(i, j) = g(i, j)^\gamma \quad (7)$$

where $g(i, j)$ represents the grayscale value of the original image pixel, $f(i, j)$ represents the corrected grayscale value of the pixel, and γ is the correction coefficient. A commonly used value for γ is $1/2.2$, which is a classic value in gamma correction.

In Figure 8, image (a) and image (b) correspond to two distinct scenes. Images (a1) and (b1) represent the results obtained after applying a Gaussian filter to their respective original images, while (a2) and (b2) depict the outcomes following gamma transformation performed on (a1) and (b1). Through a careful examination and analysis of the two sets of images in Figure 8 and their preprocessing results, several observations can be made. After the application of the Gaussian filter, the images exhibit a smoother appearance, with a suppression of high-frequency bright and dark spots. Although this process may introduce a certain degree of blurring to the image edges, it does not significantly impact the subsequent detection process, as the focus of this paper lies on regions characterized by a relatively uniform grayscale and specific areas of interest. Furthermore, after gamma correction, the contrast of the images is enhanced. In the original images, background regions such as mountains and dark stripes that grayscale values similar to the target region exhibit improved differentiation following the correction.

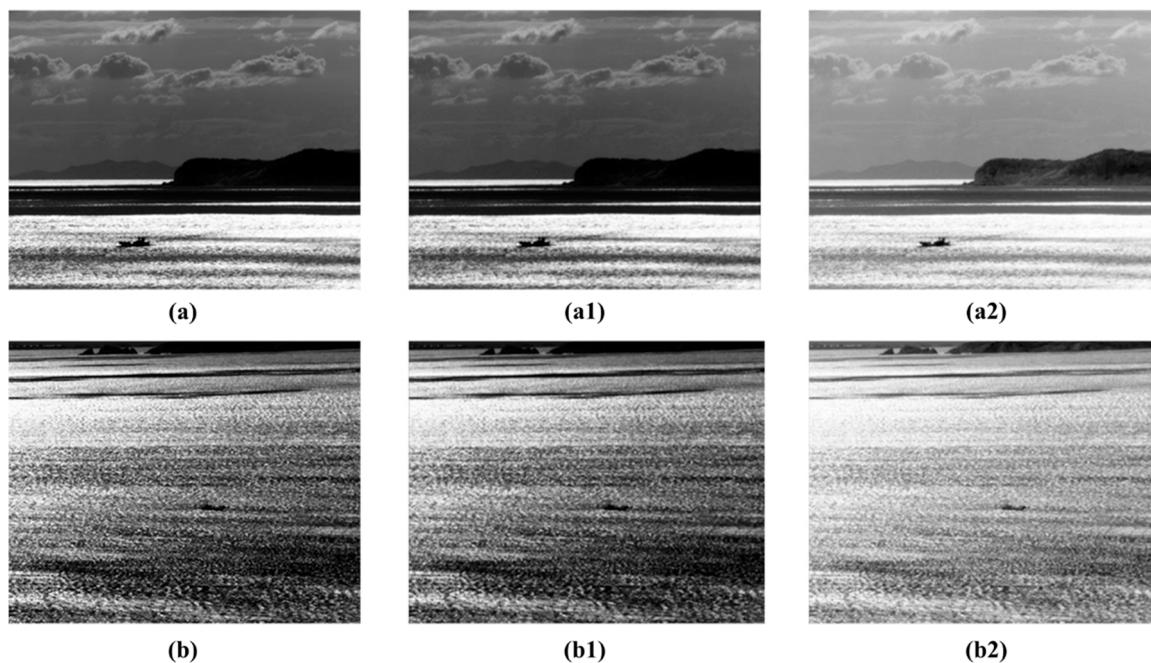


Figure 8. Image preprocessing diagram (a–b2).

3.2. Local Low Gray Region Extraction and Segmentation

3.2.1. Local Low Gray Region Extraction

Based on the preprocessing steps, the image is now ready for performing a pixel-wise convolution operation using the template defined by Equation (8), where $M \times N$ represents the size of the template.

$$template = \frac{1}{M \times N - 1} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & -(M \times N - 1) & \vdots \\ 1 & \dots & 1 \end{bmatrix} \quad (8)$$

By conducting an analysis of the template, it can be deduced that if the grayscale value of the central pixel surpasses the grayscale values of the surrounding pixels, the resulting convolution will yield a negative value. In situations where the pixel grayscale values within the template area are similar, indicating the presence of large bright or dark regions, the convolution result will be a small positive value. When the grayscale values

are identical, the convolution result becomes zero. Conversely, when the grayscale value of the central pixel is lower than the surrounding pixel values, the convolution result will be a relatively large positive value. Thus, we compute the convolution results by the following formula:

$$result(i, j) = \begin{cases} 0, & \text{if } res(i, j) < 0 \\ res(i, j), & \text{else.} \end{cases} \quad (9)$$

where res is the convolution result matrix of the original image, and $result$ is the judged result matrix.

Figure 9 shows the detection results of low grayscale regions in the images. From the local low grayscale detection results, we can see that the proposed convolution template operation achieves grayscale inversion. The original low grayscale regions have transformed into high grayscale regions, while the previously large approximately grayscale regions have become low grayscale values approaching zero, making it easier to separate them from the target. For complex waves, although there are many regions with an uneven grayscale, the presence of irregular sea clutter results in discrete low-grayscale regions with an uneven grayscale distribution. Adaptive threshold segmentation can achieve the separation of the target from most interference.

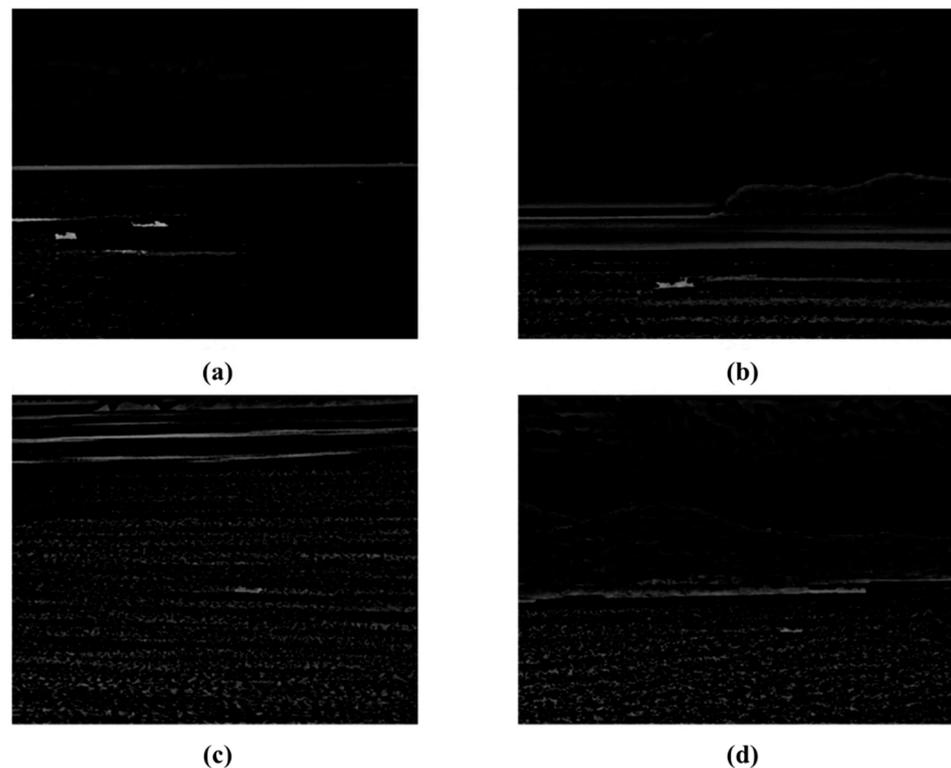


Figure 9. Local low gray regions detection results corresponding to images (a–d) in Figure 2.

3.2.2. Dual-Channel Segmentation and Multi-Feature Discrimination

From the detection results of local low grayscale regions, it can be seen that in areas with significant grayscale transitions, such as the sea-sky interface or the coastline, there are large grayscale differences. As a result, the convolution produces relatively large results, and some regions with a similar grayscale are not suppressed but are similar to the target, causing interference. Considering the overall nature of the background regions, this paper proposes a method to segment the image using two channels and fuse the segmentation results from both channels to achieve the final target detection. In this paper, the adaptive

threshold segmentation method is adopted to achieve image segmentation, and the specific segmentation formulas are as follows:

$$th = \mu + k * \sigma \quad (10)$$

$$bw(i, j) = \begin{cases} 1, & \text{if } result(i, j) > th, \\ 0, & \text{else.} \end{cases} \quad (11)$$

where th expresses the segmentation threshold, and μ and σ represent the mean and standard deviation of the input image $result$. The coefficient k is used for adjustment. In the first channel, the detection result of local low grayscale regions serves as the input image $result$. In the second channel, the input image img is the result of taking the complement of the pre-processed image $orin$. The complement calculation is performed according to Equation (12).

$$img(i, j) = 255 - orin(i, j) \quad (12)$$

Due to the existence of regions with a similar grayscale to the target, further analysis and evaluation are needed to obtain the potential target regions. In this paper, four feature quantities, namely aspect ratio (*Ratio*), compactness (*Compactness*), rectangularity (*Rectangularity*) and area ratio (*Area_{ratio}*), are used to describe the shape characteristics of the potential targets. The expressions of each feature quantity are as follows:

$$Ratio = \frac{width}{height} \quad (13)$$

$$Compactness = \frac{(Perimeter)^2}{Area} \quad (14)$$

$$Rectangularity = \frac{Area}{Rectangle} \quad (15)$$

$$Area_{ratio} = \frac{up}{down} \quad (16)$$

where $width$ and $height$ represent the width and height of the minimum enclosing rectangle of the candidate target region, $Perimeter$ and $Area$ express the perimeter and area of the candidate target, respectively. $Rectangle$ represents the area of the minimum enclosing rectangle, up and $down$ represent the areas of the upper and lower parts when the candidate target is divided into two parts evenly.

The aspect ratio criterion is employed to eliminate elongated strip-like interferences. Compactness and rectangularity measurements are utilized to filter out irregularly shaped interferences. Ship targets exhibit a characteristic of being “smaller on top and larger on the bottom.” By dividing the target region evenly into upper and lower parts, the upper part’s area should be smaller than the lower part, and this ratio is used to further remove interference. Additionally, this paper focuses on patch targets with a relatively uniform grayscale distribution (larger than 9×9 pixels), but the targets should not be too large. Larger targets tend to have clearer textures, leading to a non-uniform grayscale distribution and potential segmentation errors. Therefore, the number of pixels for the target is limited to a range of 81 to 1500 pixels in this study.

Figure 10 depicts the multi-feature discrimination process, where the connected area within the red box represents the actual ship target, while the remaining areas correspond to interference. In more detail, (a) in the figure illustrates the result of the adaptive segmentation of (d) in Figure 9. Subsequently, (b) shows the outcome of applying an area threshold to (a). Moving forward, (c) showcases the result of evaluating the Ratio, Rectangularity, and Compactness of (b). Finally, (d) displays the outcome of assessing the area ratio based on (c).

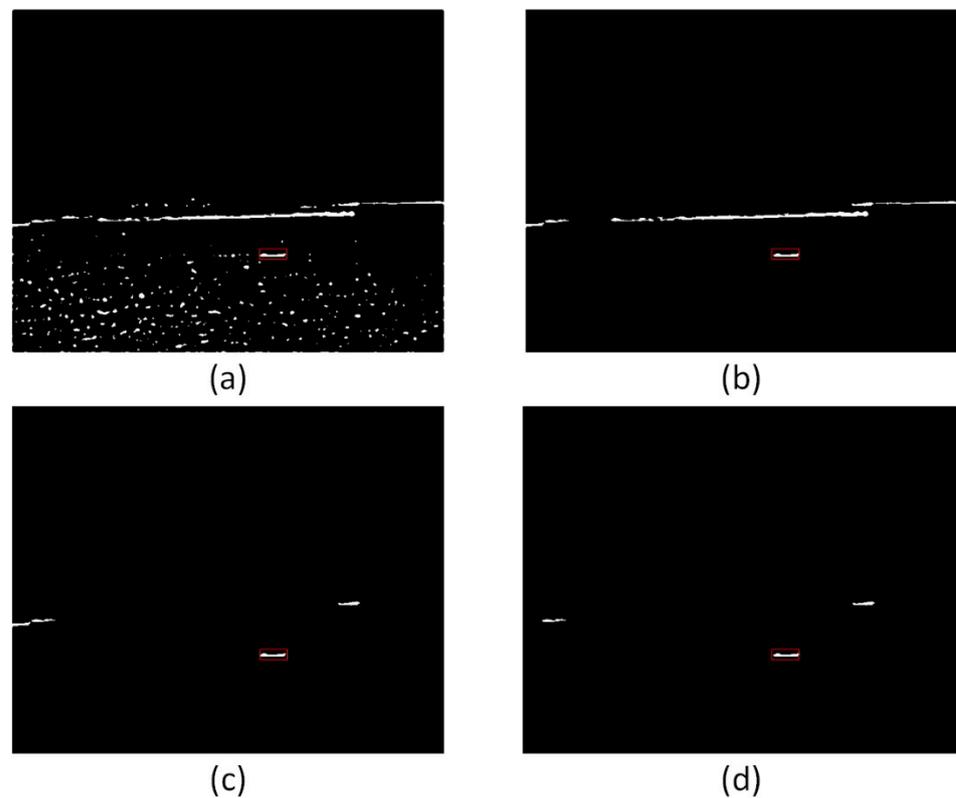


Figure 10. Schematic diagram of the multi-feature discrimination process. (a–d) corresponds to the result of different processing stages.

Firstly, regarding the adaptive segmentation of (d) in Figure 9, there are small patches of interference on the sea surface that have grayscale values similar to the target. Consequently, after segmentation, fragmented connected regions are generated. Since our focus is on patches of a certain size as targets, regions with fewer than 81 pixels or more than 1500 pixels are filtered out, resulting in the outcome shown in (b) of Figure 10. Secondly, due to the irregular nature of the sea–land boundary, the segmentation process produces multiple interfering regions. From (b), it is evident that some interference regions differ in length or width compared to the target, leading to dissimilar aspect ratios. By utilizing the aspect ratio, we can effectively filter out long coastlines or thin dark bands. The remaining interference regions in this case exhibit a rectangular shape and are relatively compact, so measures such as Rectangularity and Compactness do not significantly contribute to their filtering. However, the interference clutter on the sea surface displays a higher degree of randomness in terms of shape, making the application of these criteria useful for filtering such clutter. Lastly, typical ship targets often exhibit a distinctive “smaller on top and larger on the bottom” shape characteristic. Exploiting this characteristic, the candidate target region is divided vertically into two parts. If the area of the bottom part is greater than that of the top part, it is considered a candidate target; otherwise, it is classified as an interference region and subsequently filtered out. From (d), it can be observed that the $Area_{ratio}$ criterion successfully removes the interference region on the far left of (c). However, interference from the coastline region remains, necessitating the fusion and filtering of results from another channel.

In the other channel, to reduce computational complexity and consider that the segmentation performed on the preprocessed image is a coarse segmentation aimed at obtaining approximately grayscale regions, it is sufficient to filter out the segmented regions based on their area. The dual channel segmentation process is depicted in Figure 11.

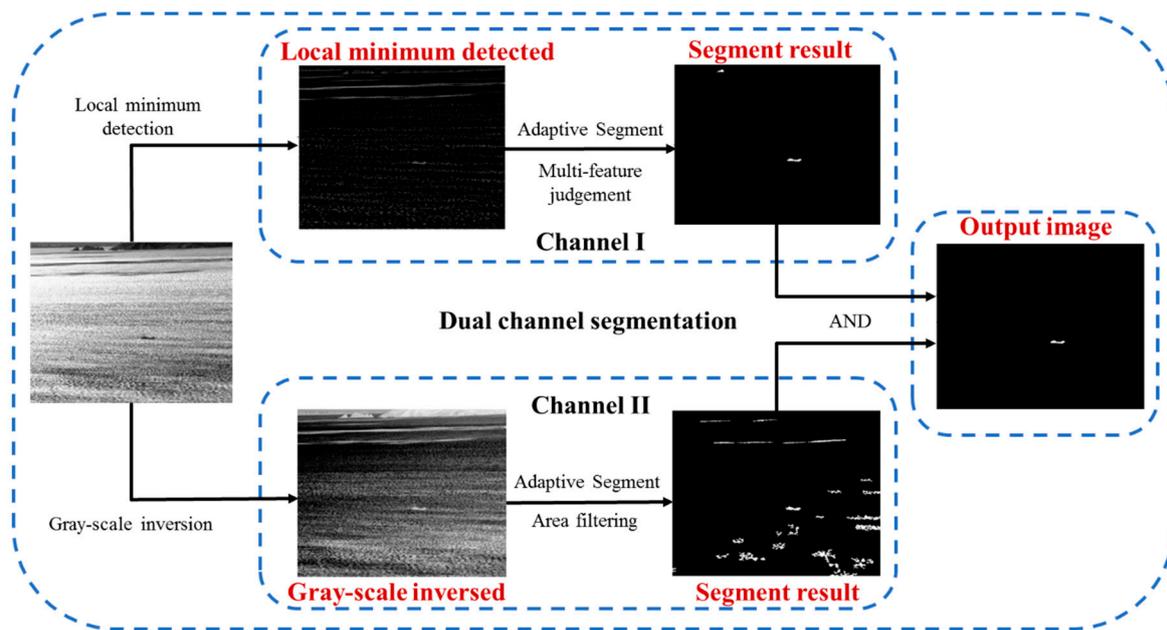


Figure 11. The dual channel segmentation process.

The dual-channel threshold segmentation technique is primarily employed to address the presence of small islands and reefs with a similar grayscale, shape, or irregular partial boundary areas between the sea and land, commonly encountered in complex sea surface backgrounds. Upon examining the dual-channel segmentation process depicted in Figure 9 of the paper, it is evident that even after segmenting and applying multi-feature discrimination to the low gray area detection results, there may still be interference from islands and reefs resembling the target. This issue arises because the convolutional template used for extracting the low gray area filters out regions with narrow shapes and small areas, thereby preventing the elimination of isolated island and reef interferences through multi-feature judgment alone. In the second channel, the inverted preprocessed image adopts a coarse segmentation strategy, causing the reefs to connect with the surrounding low-gray area, forming a larger area that can be eliminated by applying an area threshold after threshold segmentation. By subsequently multiplying and fusing the results from the two channels, accurate ship target detection results are obtained. This approach effectively addresses the challenge of interference from islands and reefs, leading to improved accuracy in ship target detection within complex sea surface backgrounds.

4. Experimental Results and Analysis

4.1. Introduction of Data Sets and Comparison Algorithms

The proposed algorithm in this paper was tested using six sets of video sequences. Figure 12 shows typical images from each set, and Table 1 provides specific information about each set.

In this paper, the algorithms proposed in references [20,28,29] were selected as comparative methods.

Ref. [20] proposes a self-adaptive segmentation algorithm for infrared ship targets against strong sunlight reflection background (SRISTAS). After applying mean-shift filtering to the image, it calculates the mean of the vertical gradient of each row, and compares the difference between the mean gradients of two rows separated by ten rows to locate rows with large gradients. Regions are then obtained by expanding from the identified rows, and a self-adaptive segmentation based on mean values in rows and columns is performed, respectively. Finally, the algorithm imposes restrictions on aspect ratio and area for the final results.

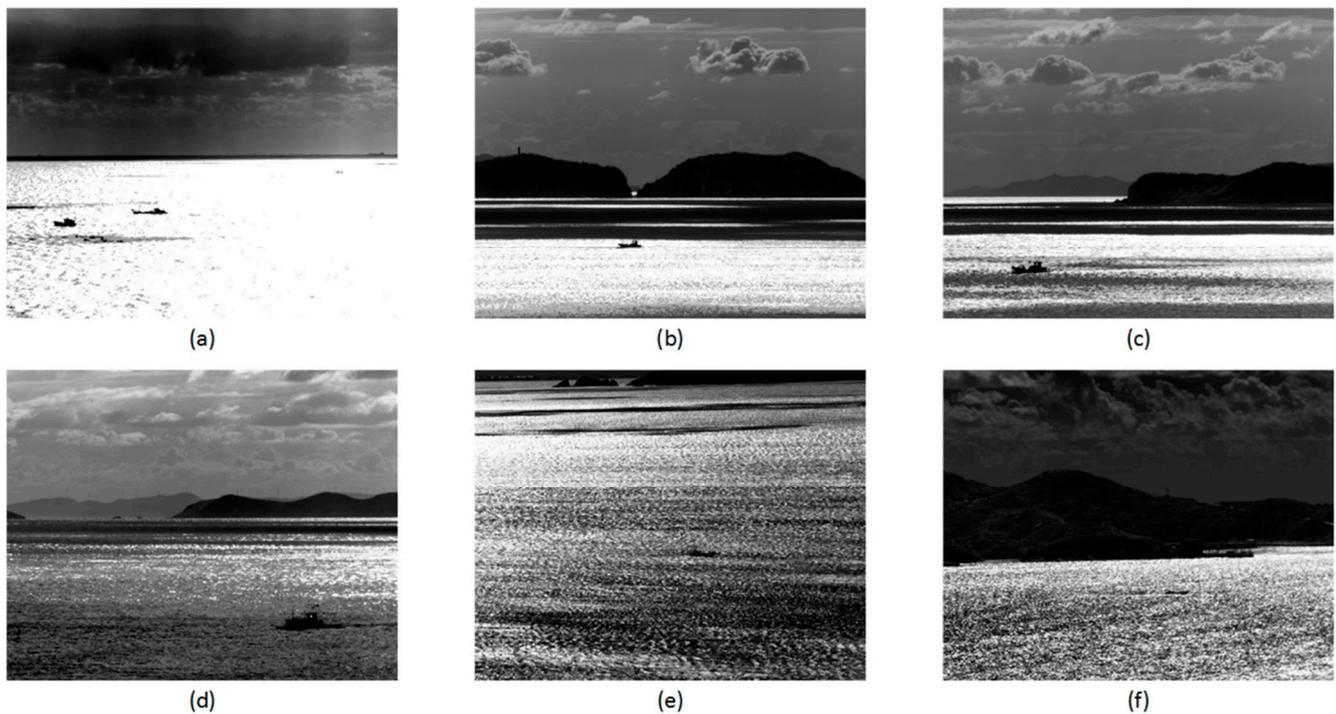


Figure 12. (a–f) corresponds to sequence 1–6.

Table 1. Information about each test sequence.

Sequence Number	Frame Number	Image Size	Numbers of Targets per Frame
1	650	640 × 512	2
2	300	640 × 512	1
3	400	640 × 512	1
4	400	640 × 512	1
5	500	640 × 512	1
6	320	640 × 512	1

Ref. [28] proposes an algorithm named infrared target detection in backlighting maritime environment based on the visual attention model (ITDBE), which enhances the attention model. After preprocessing the image with Gaussian difference, the algorithm selects the layer with the most prominent target saliency from the central-surround differences of each layer and performs segmentation using the maximum between-class variance method (OTSU). A regional growth method is used to preserve the size of the target as much as possible.

Ref. [29] introduces an algorithm called infrared small ship detection in sea clutter based on morphological reconstruction and multi-feature analysis (ISSD). It preprocesses the image using grayscale morphological reconstruction and obtains a saliency map by multiplying the brightness-contrast map and the intensity foreground saliency map. An adaptive segmentation method is then applied, followed by the incorporation of geometric feature quantities to eliminate false alarms.

4.2. Qualitative Contrast Analysis

In Figure 13, numbers (1)–(6) are the serial numbers for different scenes, (a), (b), (c) and (d) represent the detection results of the ISSD algorithm, ITDBE algorithm, SRISTAS algorithm, and the algorithm proposed in this paper for the leftmost column image, respectively.



Figure 13. (a–d) represent the detection results of each algorithm for typical images in each sequence.

The ISSD algorithm essentially highlights the target by taking the difference between the image and the background, which means that it emphasizes areas where the target has significant contrast with the surrounding region. Consequently, when the grayscale values of the target and the neighboring background are relatively consistent but different, the extracted target features become more salient, making them easier to segment. This characteristic can be discovered in the segmentation results of the ISSD algorithm for the selected six sequences' typical images. However, in the image of sequence 6, there is a similar interfering region on the left side of the target area. Despite considering multiple geometric feature measurements, the ISSD algorithm fails to eliminate this interference.

The ITDBE algorithm uses Gaussian difference preprocessing to the image to highlight the target, so it performs better in regions with high contrast. This is demonstrated by the detection results of typical images in sequence 1, sequence 2, and sequence 3. However, this method selects only one saliency map that best highlights the target for segmentation. In some scenarios, the most salient map may not be fixed, making it challenging to accurately segment the target, as shown by the segmentation results of typical images in sequences 4 and 6. Furthermore, the ITDBE algorithm does not incorporate geometric feature quantities to discriminate the segmentation results. As a result, it tends to produce more false alarms in complex scenes.

The detection performance of the SRISTAS algorithm depends on the position of the located rows. If there are no suitable gradient values in the target region, it may result in missed detection. On the other hand, if multiple regions with significant gradient changes are present, it may lead to several regions for segmentation. Additionally, this method implements segmentation based on the mean grayscale values of rows and columns. Therefore, when there are regions of interest with uniform grayscale values but significant differences from other regions, such as cloud clusters, island backgrounds, or wave interferences with similar grayscale values to the target, the detection results may contain many false alarms. We can see it from the detection results of sequences 1, 2, 4, 5, and 6 easily. Furthermore, the SRISTAS algorithm only uses aspect ratio and target area to eliminate false alarms, which may not be effective in all cases.

It is evident that in complex backgrounds, there are often regions that resemble the target, and existing algorithms may struggle to remove these interferences perfectly. Therefore, it is crucial to combine multiple geometric feature measurements of the target to discriminate and improve the detection accuracy. The algorithm proposed in this paper achieves target detection by combining the coarse segmentation of the image and the fine segmentation of low-intensity regions detected and utilizing multiple geometric feature measurements of the target. In the provided test sequences, the proposed algorithm demonstrates accurate detection results. Considering the overall performance across the sequences, the proposed algorithm outperforms the comparative algorithms.

4.3. Quantitative Contrast Analysis

To objectively evaluate the performance of the proposed algorithm, this paper adopts widely used metrics, including true positive rate (TPR), false alarm rate (FAR), misclassification error (ME), and relative area error (RAE), as evaluation criteria. The calculation methods for these four metrics are as follows:

$$ME = 1 - \frac{|B_o \cap B_T| + |F_o \cap F_T|}{|B_o| + |F_o|} \quad (17)$$

$$RAE = \begin{cases} \frac{A_o - A_T}{A_o}, & A_o > A_T \\ \frac{A_T - A_o}{A_T}, & \text{others} \end{cases} \quad (18)$$

$$TPR = \frac{TP}{GT} \quad (19)$$

$$FAR = \frac{FP}{FP + TP} \quad (20)$$

where B_o and F_o represent the number of background and target pixels in the ground truth image, respectively. B_T and F_T represent the number of background and target pixels detected by the algorithm, respectively. A_o denotes the size of the ground truth target region and A_T represents the size of the detected target region. TP represents the number of correctly detected true targets and GT represents the number of actual targets in the reference image. FP represents the number of background pixels falsely detected as targets and TP represents the number of true targets correctly detected as targets.

The misclassification error (ME) reflects the ratio of background pixels misclassified as foreground and foreground pixels misclassified as background. The relative area error (RAE) indicates the degree of closeness between the size of the detected target and the actual target. The true positive rate (TPR) reflects the ability of an algorithm to detect targets from the given image, while the false alarm rate (FAR) indicates the proportion of non-true targets among the detected targets. Analysis of the above four formulas shows that the lower the ME, FAR, and RAE values, along with a higher TPR value, the better the detection and segmentation performance of the algorithm.

Tables 2 and 3 present the ME and RAE results of the four methods across different sequences. Tables 4 and 5 display the TPR and FAR values of the four algorithms in each test sequence. Table 6 shows the average values of the four evaluation metrics for the given sequences. The best results are indicated by bolded data. Figure 14 displays line graphs depicting the four metrics across different test sequences.

Table 2. The average misclassification error of different algorithms (ME).

Sequence	ISSD	ITDBE	SRISTAS	Proposed
1	0.0888	0.2074	0.0611	0.0969
2	0.0154	0.3597	0.1392	0.1679
3	0.5330	0.2280	0.3658	0.1177
4	0.0875	0.8401	0.0331	0.1464
5	0.1819	0.6930	0.3695	0.2359
6	0.3151	0.4421	0.3576	0.4604

Table 3. The average relative foreground area error of different algorithms (RAE).

Sequence	ISSD	ITDBE	SRISTAS	Proposed
1	0.0822	0.7655	0.8014	0.0793
2	0.0117	0.2469	0.6028	0.1616
3	0.1296	0.9296	0.7967	0.2754
4	0.1335	0.8597	0.4292	0.3370
5	0.3230	0.8968	0.8247	0.3593
6	0.3809	0.4883	0.3527	0.5081

Table 4. Detection rate of different algorithms (TPR).

Sequence	ISSD	ITDBE	SRISTAS	Proposed
1	1.0000	0.9416	1.0000	1.0000
2	0.9967	0.9336	0.9967	0.9900
3	0.9541	0.7545	0.9960	0.9980
4	0.9526	0.0873	0.9701	0.9302
5	0.9640	0.3700	0.7860	0.9980
6	0.8629	0.5950	0.8505	0.9065

Table 5. False alarm rate of different algorithms (FAR).

Sequence	ISSD	ITDBE	SRISTAS	Proposed
1	0.1179	0.5574	0.6824	0.0419
2	0.0033	0.3261	0.7710	0.0033
3	0.4213	0.9105	0.1199	0.0079
4	0.4050	0.9952	0.7471	0.1711
5	0.1378	0.9633	0.9301	0.1873
6	0.0889	0.7630	0.6746	0.0034

Table 6. The average value of each algorithm index in 6 sequences.

Algorithm	ME	RAE	TPR	FAR
ISSD	0.2036	0.1768	0.9551	0.1957
ITDBE	0.4617	0.6978	0.6137	0.7526
SRISTAS	0.2211	0.6346	0.9332	0.6542
Proposed	0.2042	0.2868	0.9705	0.0692

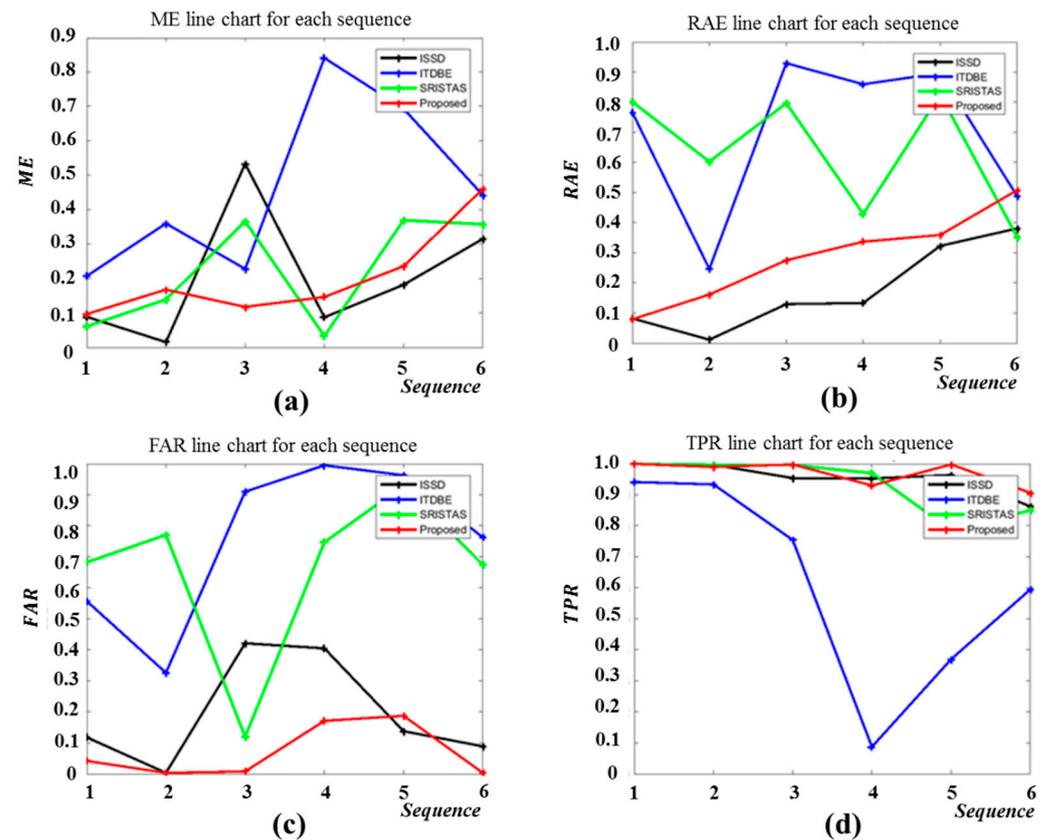


Figure 14. Line charts of the four metrics in each sequence.

The algorithm proposed in this paper often results in lower grayscale values for pixels on the target edge compared to the original higher grayscale values. Consequently, threshold segmentation may lead to the loss of the target edge. This observation is supported by the average misclassification error and the average relative foreground area error, which indicate that the proposed algorithm performs slightly worse than the ISSD algorithm in terms of these two evaluation metrics. However, the differences between the specific data of the proposed algorithm and the ISSD algorithm are not substantial. Moreover, the proposed algorithm outperforms the other two comparison algorithms in terms of both the average misclassification error and the average relative foreground area error in most of the test sequences.

From the perspective of TPR and FAR, although the ISSD and SRISTAS methods have obtained higher detection rates, the false alarm rate is also relatively high. In contrast, the proposed algorithm in this paper ensures the highest TPR among almost all sequences while maintaining a lower FAR. This demonstrates the superiority of the proposed algorithm over the compared methods considered in this study.

From Figure 14, we can observe the performance of the four algorithms on different scene sequences. The line charts in Figure 14a–d depict the ME, RAE, FAR, and TPR for each test scenario, respectively. Figure 14c clearly shows that the proposed algorithm achieves a low false alarm rate. This can be attributed to the adoption of a dual-channel segmentation

method combined with multi-feature discrimination, which effectively eliminates interference regions resembling the target within the scene. Furthermore, Figure 14d indicates that the proposed algorithm outperforms the comparison algorithm in terms of detection rate across most scenarios. Although using the convolutional template to extract the low gray area may result in some loss of edge pixels for the target, the RAE and ME values of the proposed algorithm are comparable to those of the best performing ISSD algorithm. Notably, the RAE values of the proposed algorithm are significantly smaller than those of the other two algorithms. This suggests that the targets detected by the proposed algorithm closely resemble the real targets, as visualized in Figure 14a,b.

Considering the average results of the four metrics in the given sequences, the proposed algorithm outperforms the comparison methods in terms of TPR and FAR, with only a slight difference in ME and RAE compared to the best results. Therefore, overall, the detection performance of the proposed algorithm in this paper is superior to the ones used for comparison.

5. Conclusions

In infrared images of the sea surface, the presence of complex backgrounds such as the sea, sky, and islands poses a challenge for target detection. Factors like sea wind disturbance and intense sunlight reflection further contribute to the complexity of the sea background. Ship targets in such environments often exhibit low intensity, a specific size, and a uniform grayscale on imaging devices, and the sea surface also contains patches of varying brightness and darkness, and landmasses or islands exhibit low grayscale values, making the robust detection of ship targets highly challenging. This paper proposes a dual-channel segmentation method based on local low-intensity region detection and geometric feature analysis. The image is traversed using a convolutional kernel with a negative center point to identify the low-intensity regions. These regions are subjected to adaptive threshold segmentation, incorporating shape and area features, to achieve the fine segmentation of potential target regions. In the other channel, a rough low-intensity region is obtained by applying adaptive threshold segmentation to the preprocessed image, and large area regions with similar grayscale values, such as mountains, are filtered out based on area features. Finally, the two segmentation results are fused to generate the detection outcome. Through qualitative and quantitative comparisons with three other methods on test sequences, the proposed algorithm demonstrates superior performance and improved detection capability for infrared ship targets in complex sea scenes. It effectively enhances the detection capability of ship targets in challenging infrared sea surface backgrounds, particularly in flare environments. However, it is necessary to acknowledge that this paper focuses specifically on dark-polar targets. In some scenarios, both bright and dark polar targets may coexist. Therefore, future work will enhance the algorithm's applicability and robustness by addressing the detection of both bright and dark targets, thereby expanding its effectiveness in diverse scenarios.

Author Contributions: Conceptualization, D.L.; methodology, J.T.; software, J.T.; validation, J.T.; formal analysis, M.W. and L.T.; investigation, J.T.; resources, D.L.; data curation, D.L. and J.T.; writing—original draft preparation, J.T.; writing—review and editing, D.L.; supervision, L.W. and G.G.; project administration, L.W. and G.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

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

References

1. Xu, F.; Liu, J.H.; Sun, H.; Wang, T.L.; Wang, X. Research progress on vessel detection using optical remote sensing image. *Opt. Precis. Eng.* **2021**, *29*, 916–931. [[CrossRef](#)]
2. Liang, D.; Zhang, W.G.; Huang, Q.; Yang, F. Robust sea-sky-line detection for complex sea background. In Proceedings of the 2015 IEEE International Conference on Progress in Informatics and Computing (PIC), Nanjing, China, 18–20 December 2015; pp. 317–321. [[CrossRef](#)]
3. Zhang, F.; Wang, X.; Zhou, S.; Wang, Y.; Hou, Y. Arbitrary-Oriented Ship Detection Through Center-Head Point Extraction. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5612414. [[CrossRef](#)]
4. Song, W.T.; Hu, Y.; Kuang, D.B.; Gong, C.L.; Zhang, W.Q.; Huang, S. Detection of ship targets based on CFAR-DCRF in single infrared remote sensing images. *J. Infrared Millim. Waves* **2019**, *38*, 520–527.
5. Huang, Z.X.; Wu, F.L.; Fu, Y. Review of deep learning-based algorithms for ship target detection from remote sensing images. *Opt. Precis. Eng.* **2023**, *31*, 2295–2318. [[CrossRef](#)]
6. Wang, H.L.; Lin, C.B.; Chen, D.B.; Yang, H. Ship detection of complex sea background in optical remote sensing images. *Opt. Precis. Eng.* **2018**, *26*, 723–732. [[CrossRef](#)]
7. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
8. 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**, *39*, 1137–1149. [[CrossRef](#)]
9. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multi-box detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 21–37.
10. Redmon, J.; Farhadi, A. Yolo3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
11. Gu, J.J.; Li, B.Z.; Liu, K. Infrared Ship Target Detection Algorithm Based on Improved Faster R-CNN. *Infrared Technol.* **2021**, *43*, 170–178.
12. Liu, F.; Sun, J.; Zhang, S.; Sang, H.Q.; Sun, X.J. Infrared ship target detection algorithm based on YOLOv5. *Infrared Laser Eng.* **2023**, *52*, 20230006.
13. Wang, W.X.; Fu, Y.T.; Dong, F.; Li, F. Infrared Ship Target Detection Method Based on Deep Convolution Neural Network. *Acta Opt. Sin.* **2018**, *38*, 152–158.
14. Ding, R.L.; Han, C.Z.; Xie, B.R.; Wang, Y.; Zhang, Z. Ship Target Detection in Infrared Remote-Sensing Image. *Infrared Technol.* **2019**, *41*, 127–133.
15. Wang, Y.X.; Wu, Y.G.; Xu, C.G. Infrared Ship Target Detection Algorithm Based on Deep Transfer Learning. *Air Space Def.* **2021**, *4*, 61–66.
16. Li, L.; Liu, G.J.; Li, Z.Z.; Ding, Z.Q.; Qin, T.Q. Infrared ship detection based on time fluctuation feature and space structure feature in sun-glint scene. *Infrared Phys. Technol.* **2021**, *115*, 103693. [[CrossRef](#)]
17. Wang, B.; Mo, T.Y.; Dong, L.; Xu, W. Detecting Infrared Maritime Targets Overwhelmed in Sun Glitters by Antijitter Spatiotemporal Saliency. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5159–5173. [[CrossRef](#)]
18. Xiu, B.N.; Lv, J.W.; Lu, K.K. An infrared ship target detection method based on differential filtering. *J. Command. Control.* **2020**, *6*, 171–178.
19. Xing, S.; Ji, L.; Yong, Y.; Gong, T.; Yuan, L. Infrared ship target detection based on multi-scale local edge gradient. *Ordnance Ind. Autom.* **2014**, *33*, 4.
20. Xu, Q.; Chen, X.Z.; Bai, Z.G.; Cao, X.H.; Luo, Z.B.; Jin, D.Z. Infrared ship target adaptive segmentation under strong sunlight reflection background. *Infrared Technol.* **2015**, *37*, 6.
21. Hu, S.R.; Ma, F.M.; Qin, T.Q.; Ding, Z.Q. Infrared ship target recognition technology based on multi-feature combination. *Ship Electron. Eng.* **2022**, *42*, 185–189.
22. Liang, E.H.; Zhou, A.R.; Pei, J.H.; Xie, W.X. Target segmentation method for offshore infrared images based on energy maps. *Command. Inf. Syst. Technol.* **2018**, *9*, 79–84.
23. Zhou, A.R.; Xie, W.X.; Pei, J.H. Infrared maritime target detection using the high order statistic filtering in fractional Fourier domain. *Infrared Phys. Technol.* **2018**, 91123–91136. [[CrossRef](#)]
24. Zhou, A.R.; Xie, W.X.; Pei, J.H. Background Modeling in the Fourier Domain for Maritime Infrared Target Detection. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30*, 2634–2649. [[CrossRef](#)]
25. Hou, X.; Harel, J.; Koch, C. Image Signature: Highlighting Sparse Salient Regions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 194.
26. Itti, L.; Gold, C.; Koch, C. Visual attention and target detection in cluttered natural scenes. *Opt. Eng.* **2001**, *40*, 1784–1793.
27. Liu, Z.Y.; Zhou, F.G.; Chen, X.W.; Bai, X.Z.; Sun, C.M. Iterative infrared ship target segmentation based on multiple features. *Pattern Recognit.* **2014**, *47*, 2839–2852. [[CrossRef](#)]
28. Dong, L.L.; Ma, D.D.; Qin, G.; Zhang, T.; Xu, W.H. Infrared target detection in backlighting maritime environment based on visual attention model. *Infrared Phys. Technol.* **2019**, *99*, 193–200. [[CrossRef](#)]

29. Li, Y.; Li, Z.; Zhu, Y.; Li, B.; Xiong, W.; Huang, Y. Thermal Infrared Small Ship Detection in Sea Clutter Based on Morphological Reconstruction and Multi-Feature Analysis. *Appl. Sci.* **2019**, *9*, 3786. [[CrossRef](#)]
30. Mo, W.Y.; Pei, J.H. Nighttime infrared ship target detection based on Two-channel image separation combined with saliency mapping of local grayscale dynamic range. *Infrared Phys. Technol.* **2022**, *127*, 104416. [[CrossRef](#)]

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