

Research Article

Relation and Application Method of Deep Learning Sea Target Detection and Segmentation Algorithm

Guangfu Li,^{1,2} Zheng Wang,^{2,3} and Jia Ren ^{2,4}

¹Department of Network and Information, Shandong University of Art & Design, Jinan 250300, Shandong, China

²Institute of Data Science, City University of Macau, Taipa, Macau, China

³School of Information Engineering, Shandong Youth University of Political Science, Jinan 250100, Shandong, China

⁴School of Information and Communication Engineering, Hainan University, Haikou 570228, Hainan, China

Correspondence should be addressed to Jia Ren; renjiahnu@126.com

Received 24 April 2020; Revised 6 June 2020; Accepted 13 July 2020; Published 12 August 2020

Guest Editor: Chi-Hua Chen

Copyright © 2020 Guangfu Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Target detection and segmentation algorithms have long been one of the main research directions in the field of computer vision, especially in the study of sea surface image understanding, these two tasks often need to consider the collaborative work at the same time, which is very high for the computing processor performance requirements. This article aims to study the deep learning sea target detection and segmentation algorithm. This paper uses wavelet transform-based filtering method for speckle noise suppression, deep learning-based method for land masking, and the target detection part uses an improved CFAR cascade algorithm. Finally, the best separable features are selected to eliminate false alarms. In order to further illustrate the feasibility of the scheme, this paper uses measured data and simulation data to verify the scheme and discusses the effect of different signal-to-noise ratio, sea target type, and attitude on the algorithm performance. The research data show that the deep learning sea target detection and segmentation algorithm has good detection performance and is generally applicable to ship target detection of different types and different attitudes. The results show that the deep learning sea target detection and segmentation algorithm fully takes into account the irregular shape and texture of the interfering target detected in the optical remote sensing image so that the accuracy rate is 32.7% higher and the efficiency is increased by about 1.3 times. The deep learning sea target detection is compared with segmentation algorithm, and it has strong target characterization ability and can be applied to ship targets of different scales.

1. Introduction

Despite the rapid development of China's HNA industry, the current severe sea search and rescue conditions and environment have made the existing search and rescue system's emergency rescue capabilities facing severe tests. On the whole, China's existing emergency search technology equipment is still difficult to effectively meet the needs of the rapid development of maritime transportation, and its ability to deal with major disasters and accidents at sea is weak, and there is a large gap compared with the international advanced level. China's maritime search and rescue technology and equipment cannot meet the rapid search requirements for distress targets in the deep sea under harsh

sea conditions. Therefore, in order to ensure the safety of people's lives and property and escort the emerging marine industry, we have a set of reliable and efficient maritime. It is very necessary to search for equipment to carry out the corresponding maritime search work.

Sea surface target detection is based on specific purposes and needs, by enhancing the image contrast method to enhance specific information in the image, while reducing or removing unimportant or unnecessary information, thereby improving the overall image quality. For the images obtained by different methods, through appropriate grayscale enhancement processing, the original image that can reduce the contrast or even cannot see any contour information at all is processed into a clear image that is convenient for

visual observation, making it suitable for subsequent target detection and identification [1]. At present, sea surface target detection methods are diverse and can be divided into two major categories: the first category is image enhancement methods based on the spatial domain; the second category is image enhancement methods based on the transform domain. In actual applications, depending on the specific application objects, occasions, and purposes, the methods used for sea target detection processing will also be different, and sometimes a combination of several methods will achieve better enhancement effects [2].

Satellite platform is an indispensable platform for marine target monitoring. Multisource information fusion can improve detection accuracy, recognition reliability, and space-time coverage and reduce the ambiguity of inference. In order to improve the ability of wide area search, accurate identification, continuous surveillance, and operational response of marine targets, it is necessary to further study the space-based marine target surveillance system and multi-satellite information fusion technology. He and Yao have studied the characteristics and development trend of information perception and fusion of marine environment and marine targets. They reviewed the data characteristics and shortcomings of the space-based maritime surveillance system and discussed the key technologies and future development directions in the fields of satellite networking, space-time fusion, and information fusion [3], but their methods were not forward-looking. The governments of the United States, Canada, and the European Union set up the Atlantic Ocean mapping International Working Group (asmiwg) according to the "Galway statement on Atlantic Cooperation," and made plans to map the entire Atlantic Ocean. The first step in this effort is to establish a target area of 400×400 km for the pilot mapping project. Here, scholars such as Wölfl et al. introduced the selection algorithm used to define these experimental areas. The algorithm they studied is based on careful selection of various stakeholders and publicly available marine environment parameters. Wölfl et al.'s research method involves a geographic information system-based coverage technology, which takes the parameters of the marine environment as a separate layer and combines them to obtain information about whether a location is suitable for the target area. Their research results reveal the suitability of the whole North Atlantic region and highlight three potential experimental mapping points [4]. The implementation process of this method is relatively complex, and the accuracy is not high.

The main problems that our country currently has in the technology of sea surface infrared image target detection based on the analysis of time and space characteristics available for research projects are as follows:

- (1) Insufficient research on image preprocessing methods: at present, domestic scholars have not paid enough attention to image preprocessing when studying sea surface infrared image target detection algorithms. However, the special environment at sea determines that the infrared image cannot show high quality. For example, when detecting weak and small

targets under high wind and wave conditions, strong sea wave interference will suppress or obscure the spatial characteristics of the target; under back-lighting conditions, sea clutter affects the light strong reflection and refraction will seriously reduce the image quality, which will not only cause the image brightness distribution to be unbalanced but also destroy the original contour and contrast information of the target, which will bring great difficulties to the subsequent detection [5, 6]. Therefore, an excellent image preprocessing algorithm is extremely important for improving image quality, increasing the feature difference between the target and the background, improving the target detection rate, and improving the robustness of the algorithm.

- (2) The target single frame detection algorithm has poor environmental adaptability. The infrared single-frame detection algorithm for sea targets currently available in China often has a relatively single application scenario or application target. Although the limitation of application scenarios or application goals is beneficial to extract a priori information, and thus improve the detection rate of the algorithm, however, in the actual sea target infrared search process, the scene information or target information cannot be predicted or limited. Therefore, the algorithm of the application condition list is difficult to be applied in actual engineering; otherwise, it will cause high leakage alarm rate or false alarm rate [7, 8]. In addition, in some harsh marine environments (such as dense fog, backlighting, and heavy wind), the existing single-frame detection algorithm is also difficult to accurately detect the target, so it cannot guarantee the robustness of the algorithm in different marine environments.
- (3) The multiframe decision algorithm is seriously affected by the interframe jitter of the image sequence. In the actual sea target infrared search process, the imaging platform will be randomly shaken due to the influence of sea breezes or waves, which will cause the image sequence to appear "misaligned" between frames. This will destroy the continuity of the movement of the target between frames and the consistency of the trajectory, thereby adversely affecting the final target detection rate.
- (4) The algorithm lacks systematic experimental verification. In the currently available literature or published research results, the experimental verification stage often uses a small amount of sea video images captured in specific scenarios to verify the algorithm and lacks verification data under complex and diverse environmental conditions, so it is difficult to reflect or guarantee. The environmental adaptability and robustness of the algorithm is extremely unfavorable for practical engineering applications [9, 10].

Due to the influence of bad weather such as foggy and rainy weather on the sea surface, long-distance transmission,

and atmospheric attenuation, the image after contrast stretching still has the problem of detail blur. In this paper, the wavelet transform-based filtering method is used for speckle noise suppression, and deep learning is used. The method is used for land masking, and the target detection part adopts the improved CFAR cascade algorithm and finally selects the best separable features for false alarm rejection. In order to further illustrate the feasibility of the scheme, this paper uses measured data and simulation data to verify the scheme and discusses the effect of different signal-to-noise ratio, sea target type, and attitude on the algorithm performance. The research data shows that the deep learning sea target detection and segmentation algorithm has good detection performance and is generally applicable to ship target detection of different types and different attitudes.

2. Proposed Method

2.1. Principle and Steps of Traditional Grabcut Image Target Segmentation Algorithm

2.1.1. Grabcut Image Segmentation. Grabcut image segmentation method is a graph-based segmentation method, which maps a graph into a network graph. By designing the energy function, the energy graph is segmented by the minimum cut, and the image is segmented by finding the minimum energy function. The color image is composed of pixels in the RGB color space [11, 12]. Because it is difficult to create a sufficient color space histogram, the GMM model (Gaussian mixture model) is used to build a color image data model. The basic steps are shown in Figure 1.

The current Grabcut algorithm uses an interactive method to manually mark the foreground target image and then complete the final segmentation operation according to the difference between the foreground and background. However, when processing a large number of images, manually marking the foreground target is a tedious and large amount of work with low efficiency [13, 14].

2.1.2. SSD Model Training. SSD convolutional neural network model training uses the open source TensorFlow deep learning framework provided by Google. There are 20 types of images in the PASCALVOC2007 database, including people, cars, airplanes, and various animals. During the SSD training process, you need to input the image and the real label box. After the convolution operation, the default boxes with different aspect ratios are evaluated at each position in the feature map with different scales. For each default box, it is predicted for all object categories, shape's offset, and confidence. During training, first match these default boxes to the real label boxes. As shown in Figure 2, the Jaccard similarity coefficient between the real target box and the default target box is calculated. The default target box with the largest similarity coefficient is paired with the real target box to ensure each real target frame has a default target frame corresponding to it [7, 15].

2.2. Characteristics of Sea Target Detection. In-depth study of sea surface infrared image target detection methods and the development of sea surface target detection mechanisms under complex sea conditions will help greatly improve China's maritime search and rescue capabilities, shorten the technical gap between China and developed countries, and thus provide life and property safety for the majority of marine personnel. Escort provides strong support for China's development to the world's major marine economy [16]. Target detection algorithm based on the analysis of space-time characteristics is the earliest and most classic infrared target detection algorithm [17, 18]. This kind of algorithm often uses a single frame detection and multiframe decision concatenation solution, first extract the suspect target area according to the spatial characteristics (such as gray contrast, texture, and gradient) in the single frame image, and then use the target to the continuity of motion and the consistency of the trajectory in the image sequence judge each suspected target, thereby eliminating the false target. At present, compared with the target detection algorithm based on the analysis of spatio-temporal characteristics, the target detection algorithm based on random signal processing can achieve more excellent target detection performance of sea surface infrared images. However, due to the extremely high-computational requirements of target detection algorithms based on narrow machine signal processing, even the use of program acceleration techniques (such as GPU and FPGA) cannot achieve real-time detection of targets [19, 20].

2.3. Features of Segmentation Algorithm. Image segmentation is an important research content of remote sensing image processing. The processing of visible light remote sensing images has a wide range of applications in military and civilian applications, and the segmentation of visible light remote sensing images is one of the important research contents [21]. Most of the existing remote sensing target extraction methods are based on a specific type of target on infrared images or SAR images, and the segmentation methods on visible light remote sensing images are mostly aimed at the classification of landforms on low and medium resolution images.

The feature classification is regarded as a kind of segmentation. Existing segmentation methods for such regional targets are usually inefficient and poor in real-time. For medium- and high-resolution object targets with clear boundaries, the current extraction method is basically in the semiautomatic processing stage of human experience interpretation or human-computer interaction and needs to solve the intelligence and automation of object extraction [21, 22].

In the infrared weak target detection algorithm designed in this paper, the feature extraction of the target adopts the method of extracting the minimum circumscribed rectangle of the target in the local neighborhood of the target, which is a very time-consuming method. In view of the fact that the system may be used in other occasions where the real-time requirement is higher or the target local neighborhood size is

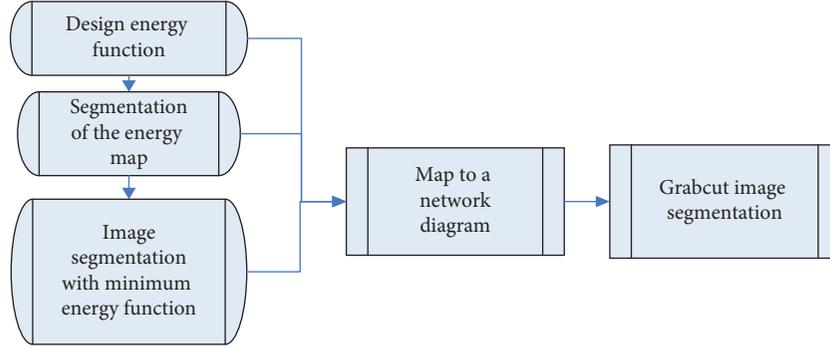


FIGURE 1: The basic steps.

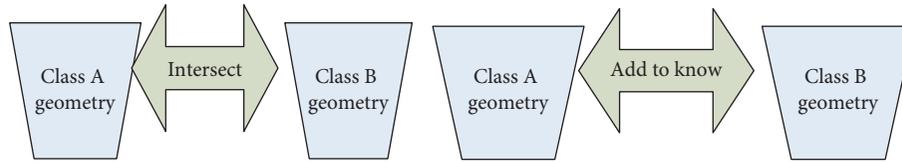


FIGURE 2: Jaccard similarity coefficient model.

larger, the method of circumscribed rectangles may not meet the real-time requirement, which requires the improvement of the target feature extraction method. Due to the problem of poor real-time target detection method based on convolutional neural network, in the next step, the CAFFE source code needs to be optimized and implemented, especially for the feedforward calculation of the CNN model, algorithm equivalent optimization and approximate optimization need to be considered, and performance is guaranteed under the premise of completing the hardware implementation of the prototype algorithm and through the laboratory injection simulation platform for testing and verification [23].

2.4. Grabcut Segmentation

2.4.1. Color Data Model. The Grabcut algorithm uses GM to establish a data model for color image target slices. It uses the full covariance GMM of K Gaussian components (generally $K = 5$) to model the foreground and background pixels of the image to obtain a Gaussian vector $k = \{k_1, \dots, k_n, \dots, k_N\}$, where k_n is the image. The n th pixel in the image belongs to the Gaussian component, k_n , $k_n \in [1, \dots, K]$. Therefore, each pixel in the image corresponds to a Gaussian component from the target GMM or the background GMM, so the Gibbs energy of the entire image is

$$E = V + U. \quad (1)$$

In the formula, it is opacity, $A \in [0, 1]$, 0 represents the image background and 1 represents the foreground target; it is the statistical gray histogram between the background and foreground pixels of the image, $B = \{h(C, D), E = 0, 1\}$, z is an array of pixel values, and $h(\cdot)$ calculates gradient statistics. The influence of the main GMM component k introduces the GMM model:

$$U = \sum N + V. \quad (2)$$

In the formula, U is the regional energy term, representing the negative logarithm of the probability that the pixel belongs to the background or the target and indicating that a pixel in the image is classified as a background or target penalty. The mixed Gaussian probability density model is defined as follows:

$$D = \sum 1 + \sum i, \quad (3)$$

$$g = \frac{1}{\sum \exp}. \quad (4)$$

The negative logarithm of equation (3) can be written as follows:

$$d = -\log w + \frac{1}{2 \log k}. \quad (5)$$

In the formula, the parameters of the GMM have R three items: the weight of each Gaussian component, $A = B$ ((the number of pixels belonging to the Gaussian component)/(total number of pixels)); the average value of each Gaussian component μ , for the three channels of rgb. There are three element vectors, the covariance matrix Σ , which is a 3×3 matrix, expressed as follows:

$$\theta = w + a + k\infty. \quad (6)$$

Once the three parameters w, μ , and Σ describing the background GMM and the target GMM of the image are determined, enter the RGB color values of the color image and bring it into the background GMM and the target GMM, and then the pixel can be obtained and classified into the background and target probability so that the U term of Gibbs energy can be determined.

The calculation formula of the boundary term energy V is as follows:

$$V = \sum \exp + \theta. \quad (7)$$

The boundary term reflects the penalty of discontinuity between neighboring pixels o and p . The greater the difference between adjacent pixels, the smaller the energy; the parameter β is determined by the pixel contrast; the Euclidean distance $|A|$ is used to measure the similarity between pixels. For images with low contrast, the difference between pixels o and $p|B|$ is relatively low, you need to multiply the larger β value to magnify the difference, otherwise multiply the smaller β value so that the energy of the V term can work normally; γ is a constant, and take $\gamma = 50$.

2.4.2. Iteratively Minimize Energy and Split

(1) Initialization:

- (a) Frame the target in the image so that the pixels outside the box area are background pixels TB , and the pixels inside the box area are potential target pixels TU .
- (b) For pixel $n \in TB$, let label $\alpha n = 0$, for pixel $n \in TU$, and let label $\alpha n = 1$.
- (c) The GMM of the background and the target can be estimated through the foregoing steps, and the background and target pixels are clustered into K categories using the k -means method, that is, K Gaussian models of GMM, each of which corresponds to some pixel samples.

(2) Iterative minimization

- (a) Assign Gaussian component K_n in GMM to each pixel:

$$K_n = \text{rig min}. \quad (8)$$

- (b) After the foregoing steps, each Gaussian model has some pixel samples. For the input image, learn the parameters w, μ of the optimization GMM optimization formula:

$$\theta = \text{rig min}. \quad (9)$$

- (c) Segmentation estimation: by minimizing Gibbs energy, the optimal formula of the initial segmentation algorithm is obtained:

$$\min = \min E. \quad (10)$$

- (d) Repeating formulae (1)–(3) iteratively make the energy gradually decrease. After each iteration, the GMM and segmentation results are optimized interactively until the energy converges.
- (e) Perform postprocessing such as smoothing on the boundary after image segmentation to complete the image segmentation process.

2.4.3. *Algorithm Flow Description.* The flow of this algorithm is described as follows. For an input image, the specific steps of the sea surface target detection and segmentation algorithm using this algorithm's deep learning are shown in Figures 3 and 4.

2.4.4. CFAR Algorithm Principle and Algorithm Model.

The formation methods of the detection threshold of automatic detection are roughly: (a) formed by fixed threshold; (b) formed by the average amplitude of external interference; (c) formed on the basis of the obtained prior information of the interference statistical distribution; (d) formed prior to the interference-free statistical distribution information by testing under a statistical assumption of free distribution. Here, (a) is fixed threshold detection, (b) and (c) are adaptive threshold CFAR detection, and (d) is nonparametric CFAR detection. The clutter background of CFAR detection, which belongs to the category of automatic detection in a broad sense, can also be summarized into three typical situations: uniform clutter background; clutter edge; multitarget background. The Rayleigh distribution is suitable for the description of general clutter. The mean CFAR detection technology can achieve a constant false alarm rate. The mean class CFAR is suitable for a statistically stable background. It has a sliding window covering several distance units before and after the detection unit. The average of the reference samples in the sliding window is used to form a local estimate of the front and back edges. Local estimation average, large selection, small selection, or weighted average are to determine the average power estimation of the background clutter of the detection unit. In view of the fact that the signal may cross into the front and back neighboring units, the detection unit and its neighboring front and back distance units are generally not included in the average window. If the signal amplitude in the detection unit is greater than K times the average value in the sliding window, it is considered a signal. The algorithm model is shown in Figure 5. The echo sequence is sent to the transmission delay line. The central unit of the delay line is the current detection unit. There is a protection unit on each of the two adjacent sides. The front and back windows are summed separately and sent to the detection logic. If the detection logic outputs the average value of the sum of the front and rear windows, it is CA-CFAR; the selected small value of the front and rear window sum is SO-CFAR; the front and rear windows are output. The largest value of the sum is GO-CFAR.

3. Experiments

3.1. *Dataset.* The suppression of SAR speckle noise can be achieved by incoherent multiview processing, or it can be achieved by using spatial domain filtering. Incoherent multiview processing will reduce the ground resolution of the image. We can use spatial domain filtering methods, such as mean filtering, median filtering, Lee filtering, Kuan filtering, Frost filtering, Sigma filtering, and GammaMap filtering. However, this kind of algorithm has its own insurmountable contradictions: on the one hand, a larger filter

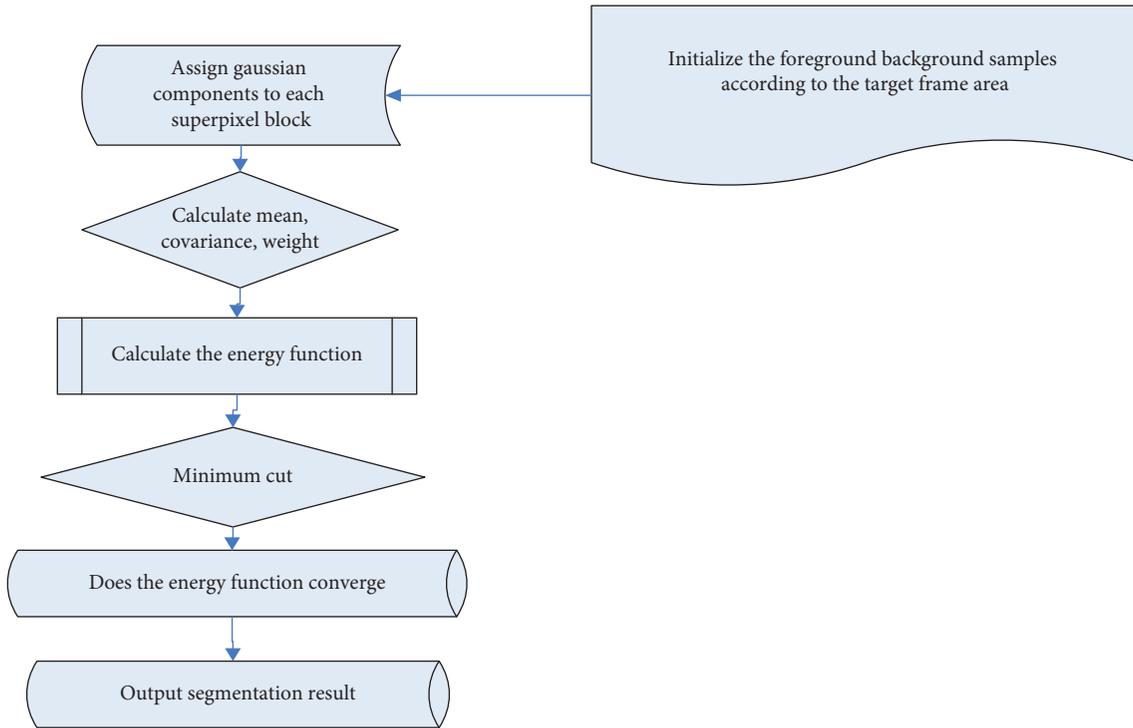


FIGURE 3: The specific steps of the detection and segmentation algorithm.

window needs to be selected to enhance the spot denoising effect, and on the other hand, the window selected must be small in order to maintain the actual resolution of the image. The wavelet transform filtering method can solve these problems, so in this paper, the wavelet transform filtering method of the sea surface image is used for speckle noise suppression, and the deep learning-based method is used for land masking. The target detection part uses an improved CFAR cascade algorithm, and finally selects the feature with the best separability is eliminated by false alarm. In order to further illustrate the feasibility of the scheme, this paper uses measured data and simulated data to verify the scheme and discusses the effects of different signal-to-noise ratios, sea target types, and attitude on the performance of the algorithm.

3.2. Basic Settings of the Experiment. In this paper, the wavelet transform filtering method of the sea surface image is used for speckle noise suppression, and the deep learning-based method is used for land masking. The target detection part uses an improved CFAR cascade algorithm and finally selects the best separable features for false alarm eliminate. In order to further illustrate the feasibility of the scheme, this paper uses measured data and simulation data to verify the scheme and discusses the effect of different signal-to-noise ratio, sea target type, and attitude on the algorithm performance.

3.3. Algorithm Description and Analysis. This algorithm has three main modules, and the specific content is shown below:

- (1) Target detection operation: using deep learning network model to automatically detect the target rectangular frame in the image.
- (2) Superpixel segmentation: calculate the objective function value through color information and position information. Update the cluster center according to the value of the objective function. By comparing the old and new cluster centers, if the cluster center changes, update the cluster information according to the new cluster center, then apply the visible light defogging algorithm to remove the infrared image blur, and finally get the super pixel segmentation result.
- (3) Automatically mark the target frame as the Grabcut algorithm to initialize the sample information of the foreground and background according to the target detection results and process the sea surface defogging image through superpixelization, reducing the number of nodes in the graph, building a simple graph structure and reducing Grabcut the iteration time of the image segmentation algorithm.

3.4. Experimental Procedure. Use the deep learning target detection model to intelligently extract the target frame in the image instead of manually marking the foreground target in the Grabcut algorithm, and select the clustering center from the specification, and the boundary update strategy to improve the SLIC superpixel segmentation algorithm in two aspects, treating similar pixels as a super Pixels, construct a streamlined network graph, reduce the

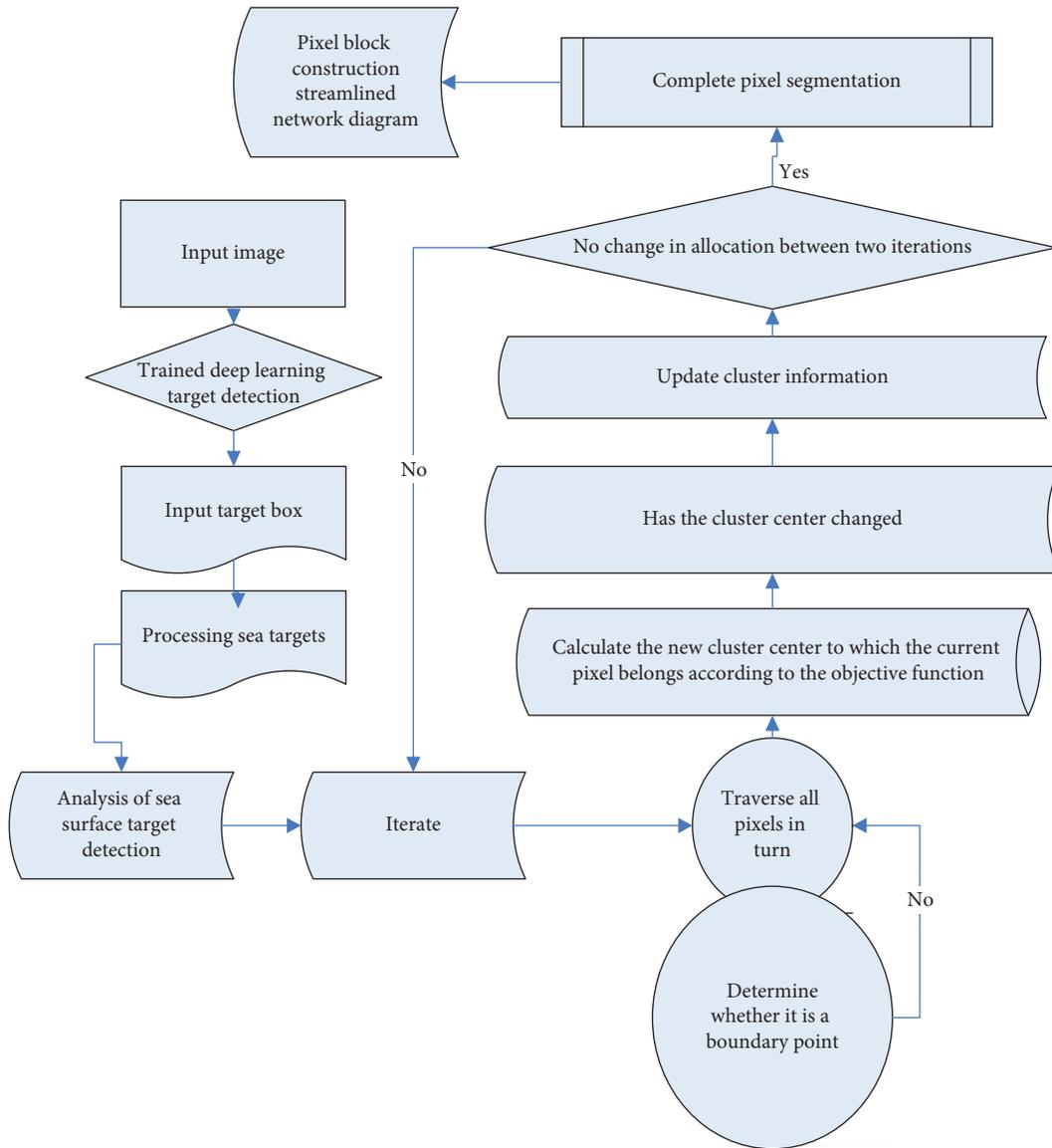


FIGURE 4: The specific steps of the detection and segmentation algorithm.

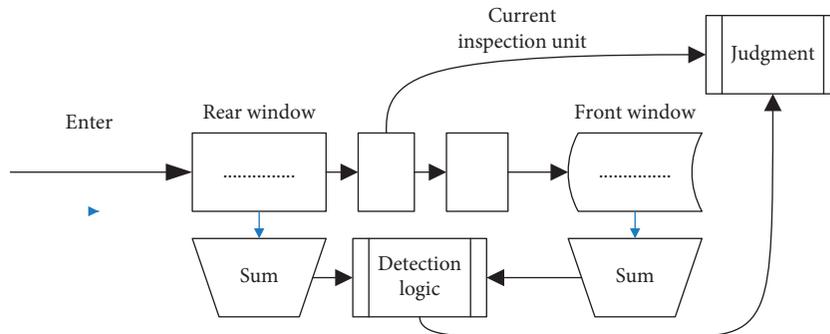


FIGURE 5: Algorithm model.

number of algorithm iterations, and improve the Grabcut segmentation algorithm from efficiency. Finally, to describe and analyze the steps of the improved Grabcut

segmentation algorithm, iteratively select the minimum cut and finally complete the automatic target segmentation.

4. Discussion

4.1. Sea Surface Target Detection and Analysis

- (1) As shown in Table 1 and Figure 6, the filtering method is used for speckle noise suppression, and the deep learning-based method is used for land masking. The target detection part uses an improved CFAR cascade algorithm, and finally the best separability is selected. False alarms are removed. The research found that the infrared imaging system has a long detection distance and the sea climate is complex and changeable. Due to the double impact of atmospheric scattering attenuation and bad weather on the sea surface, the attenuated transmitted light causes the reflected light received by the detector to reflect the essential information of the scene to change. After the contrast is enhanced, the infrared image of the sea target scene is still blurred; the target wheel, details, and other information are unclear. After the above measures have been taken after communication, the intensity contrast value of the original picture is more than 3 times higher than the screening value. The ratio will affect the subsequent target detection and recognition. The signal clutter must be deblurred after the contrast is enhanced. The blurred image after degraded seems to be covered with a layer of fog, which is similar to the effect of the fog compared to the visible light image. Inspired by this, we can use the visible light defogging algorithm to remove the infrared image blur.
- (2) As shown in Table 2 and Figure 7, it is the clarity of the sea defogging image compared with the original under several measures. It can be seen from the several algorithms of the sea image defogging of the collected filtered data, the basic changes of the data after multilevel filtering, the use of side suppression filtering to change the appearance of the image, and the appearance after CAS filtering, and finally apply the original image defogging technology to collect relevant data, solve the image degradation problem caused by fog through these methods, and use the deep learning target detection model to intelligently extract the target frame in the image instead of manually marking the foreground target in the Grabcut algorithm, and separately, from the specifications, select the clustering center and use similar pixels as a superpixel to construct a simplified network graph, reduce the number of algorithm iterations and improve the Grabcut segmentation algorithm from the efficiency. Judging from the number of image segmentation, feature coefficients, and the number of marks, the quality after the image is fitted. It can be seen that the quality distribution coefficients of the percentages of multilevel filtering and GAS filtering in dealing with blurred image segmentation are almost equal, and the percentage of sharpness of the filtered image is lower than the

original picture by about 50%. However, the results of feature visibility are not obvious. It shows that the sea surface image defogging process can increase the clarity and monitoring ability of the sea surface target and has a guiding significance for ocean detection and analysis.

4.2. Analysis of Sea Targets under the Algorithm

- (1) The wavelet transform of the sea image is developed on the basis of two-dimensional discrete wavelet transform. Each time, through two-dimensional wavelet transform, we can get a low-frequency subimage and three high-frequency subimages in different directions. The low-frequency subgraph gives the main body contour information of the original image, while the high-frequency subgraph gives the original image edge details in all directions. After a wavelet transform, the obtained first-level low-frequency subgraph can be wavelet-transformed again to obtain a second-level subgraph, and the second-level low-frequency subgraph can still continue to undergo wavelet transform to obtain a third-level subgraph, and so on, as shown in Figure 8. In this schematic diagram, BB represents the low-frequency subgraph and BH, HB, and HH distributions represent the high-frequency subgraph in three different directions. The wavelet transform of the image is a pyramid structure, the bottom layer is the original image, and then each layer is decomposed upwards to obtain a layer of subimages, and the resolution of the image is gradually reduced from bottom to top. The original image is finally expressed as a combination of the top-level low-pass image and a series of prediction residual images of each layer image.

The basic idea of the speckle noise suppression method based on wavelet transform is to perform low-pass filtering on the low-frequency subgraph obtained after wavelet transform to filter out speckle noise, and then add the edge detail information contained in the three high-frequency subgraphs. It can achieve the dual purpose of not only removing speckle noise but also retaining the edge details of the image.
- (2) The image sequence is an uninterrupted input when searching for sea targets. At this time, the target multiframe decision algorithm of the deep learning sea target detection and segmentation algorithm proposed in this paper can well achieve the removal of pseudotargets and background interference. As shown in Table 3 and Figure 9, select images are collected at different time periods in a certain sea area for analysis. First, the dark primary color priori theory is used to estimate the transmission image, and the sharpness of the image is tested by a professional measuring instrument. The transmission diagram reflects the degree of the image affected

TABLE 1: Percentage of clarity of pictures in various situations.

	Filtering	Multistage filtering	Side suppression filtering	CAS filtering	Original image
Contrast	30	48	69	85	113
Signal to clutter	15	17	9	11	8
Scroll	65	68	94	58	74

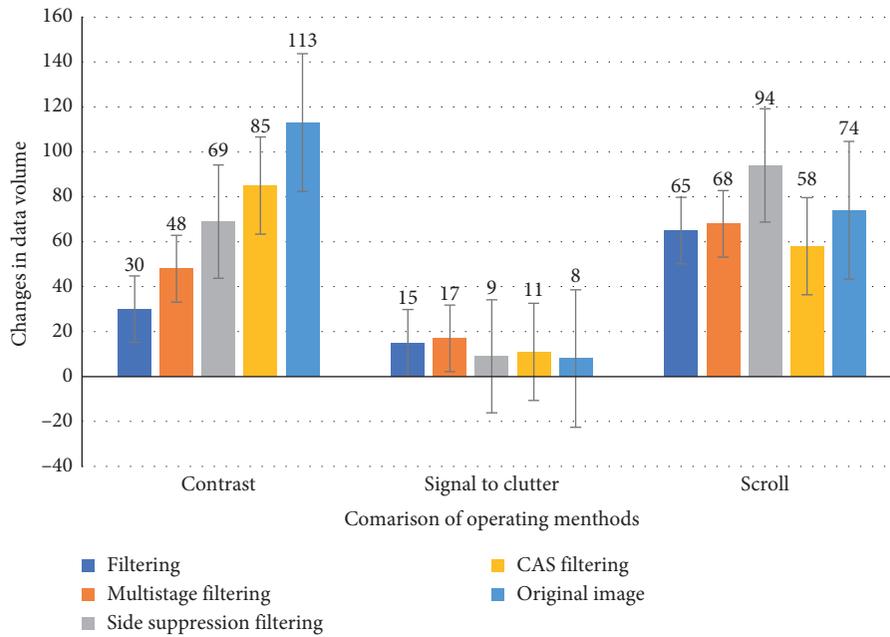


FIGURE 6: Percentage of clarity of pictures in various situations.

TABLE 2: Percentage of sea fog image clarity under the algorithm.

	Filtering	Multistage filtering	Side suppression filtering	CAS filtering	Original image
Image segmentation	27.3	29.8	26.4	29.7	20.8
Characteristic coefficient	45	34	35	46	46
Mark	245	643	543	564	453

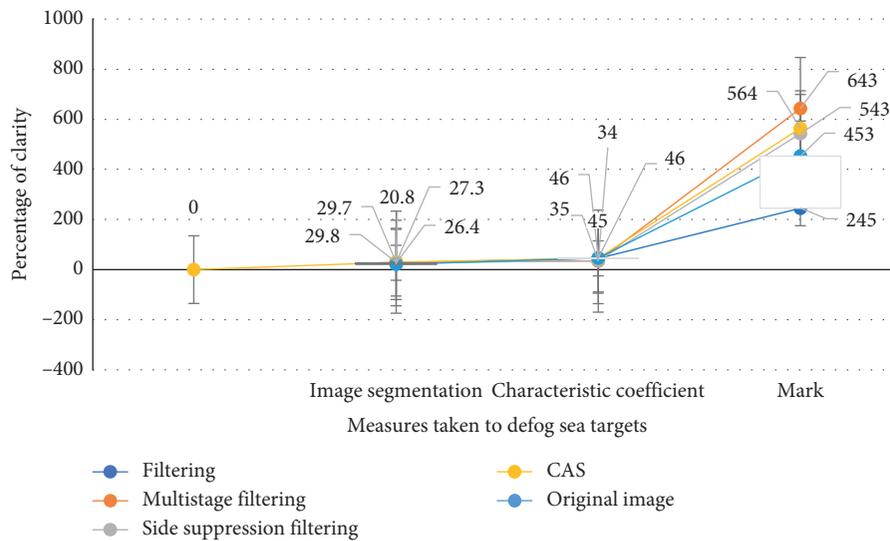


FIGURE 7: Percentage of sea fog image clarity under the algorithm.

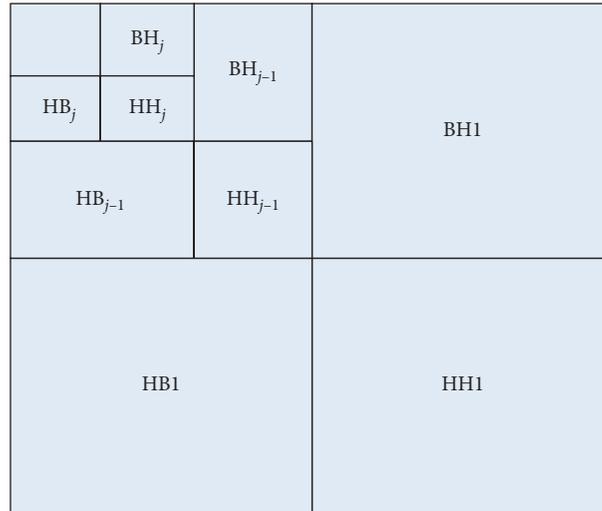


FIGURE 8: Speckle reduction of wavelet transform to suppress high-frequency subgraphs in different directions.

TABLE 3: Analysis of images collected at different time periods in a sea area (%).

	17-19 time period	19-21 time period	21-23 time period	23-1 time period	1-3 time period
Dark primary colors measure the clarity of the image	56	78	65	76	78
Percentage of atmospheric light value	35	47	58	56	39
Percentage of image clarity of atmospheric scattering model	47	46	78	56	89

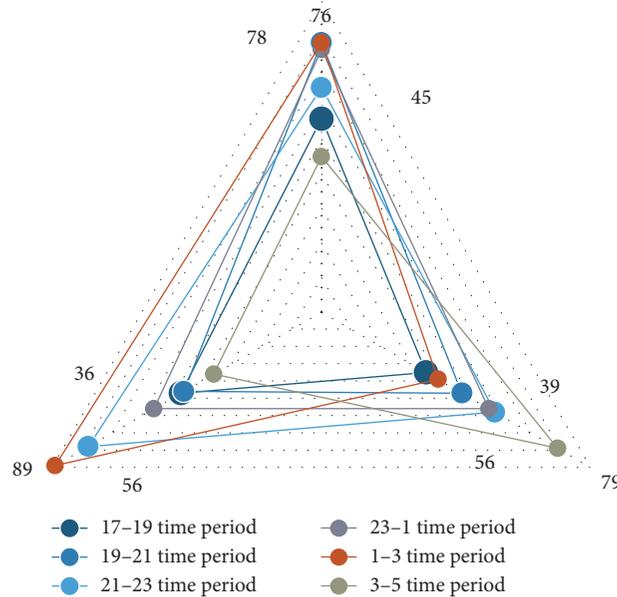


FIGURE 9: Analysis of images collected at different time periods in a sea area (%).

by the fog. According to the humidity and temperature of the air in the weather of the day, the percentage of the atmospheric light value is estimated, and the two are brought back to the atmospheric scattering model to obtain the image clarity after defogging percentage. Since the target single-frame detection method can effectively suppress

background clutter interference in the conventional sea environment, while only a small amount of background clutter remains, it can achieve accurate detection of different sea targets with a local contrast of not less than 0.07. It can be seen from the figure that the overall clarity between 23 o'clock and 1 o'clock in the morning is high, and the clarity is

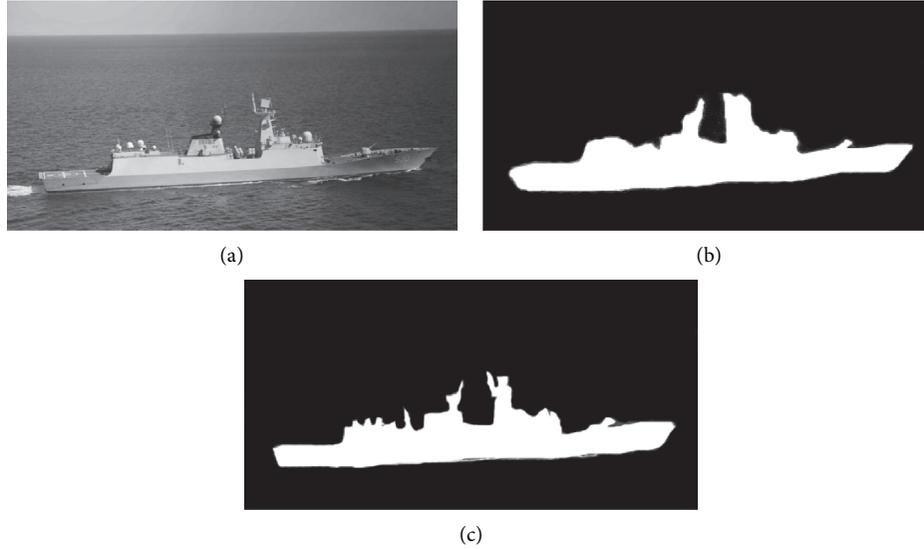


FIGURE 10: The step effect diagram combined with visual significance segmentation. (a) Original image. (b) Adaptive segmentation. (c) Grabcut image segmentation.

about twice the percentage of other time periods. The degree of clarity is higher than other times, so when using the atmospheric scattering model, a clear comparison is made.

4.3. Effect Comparison of Segmentation. In order to solve this problem, the prior information in the calculation of saliency map is fused into the energy function to improve the accuracy of segmentation. The improved calculation of energy function is shown as follows:

$$E_2(x) = \sum_{i \in v} \Phi(x_i) \varphi(x_i) + \beta \sum_{i, j \in N} \phi(x_i, x_j). \quad (11)$$

Among them, $\Phi(x_i)$ is the significant value after normalization, $\varphi(x_i)$ is the energy equation of regional term, and $\phi(x_i, x_j)$ is the energy of boundary constraint term. In order to solve the problem of the integrity of segmentation, after each iteration, the objects from the previous segmentation are expanded and corroded to get a new ternary image for the next iteration. Figure 10 is the step effect diagram combined with visual significance segmentation, in which Figure 10(b) is the result of adaptive segmentation and Figure 10(c) is the Grabcut segmentation.

5. Conclusions

At present, the Grabcut algorithm requires manual intervention and uses the method of manually labeling the rectangular frame to obtain the approximate area of the target. When processing complex images, the user cannot effectively label the rectangular frame. Based on the existing network model, this paper proposes a deep learning sea surface target detection and segmentation algorithm, and this method mainly uses the deep learning sea target detection and target detection model provided by Google and

combines the superpixel segmentation algorithm to optimize the Grabcut algorithm, which realizes the deep learning sea target detection and segmentation algorithm, which can be accurate target contour and semantic information.

In the search and rescue of distressed targets in the deep sea, the ability to quickly and accurately find the distressed target is the determinant of the success of the search and rescue. Therefore, the detection and tracking technology of infrared targets has always been a hot spot for domestic and foreign scholars. Research on deep learning sea target detection and segmentation algorithms found that location prediction and category prediction in target detection tasks have different requirements for feature map features. Location prediction requires feature maps to contain rich details and positioning information, while category prediction requires feature maps to have abstraction semantic features and translation invariance. However, the existing algorithms use the same feature map for prediction in both position prediction and category prediction. To solve this problem, this paper redesigned the network structure based on the idea of separating the corresponding feature maps of location prediction and category prediction and the deep learning sea target detection and segmentation algorithm model. Two channels are added to the deep learning sea target detection and segmentation algorithm model to predict the location and category, respectively, to achieve a certain degree of separation of the two tasks. The feature map used for position prediction is merged with the low-level feature map through downsampling, which adds rich details and positioning information.

Through the above research work, the contrast of the infrared image of the sea target scene is significantly improved, and the details are clearer. In this paper, the wavelet transform-based filtering method is used for speckle noise suppression, and the deep learning-based method is used for land masking. The target detection part uses an improved

CFAR cascade algorithm and finally selects the best separable features for false alarm removal. In order to further illustrate the feasibility of the scheme, this paper uses measured data and simulation data to verify the scheme and discusses the effect of different signal-to-noise ratio, sea target type, and attitude on the algorithm performance. The research data shows that the deep learning sea target detection and segmentation algorithm has good detection performance and is generally applicable to ship target detection of different types and different attitudes. The results show that the deep learning sea target detection and segmentation algorithm fully takes into account the irregular shape and texture of the interfering target detected in the optical remote sensing image so that the accuracy rate is 32.7% higher and the efficiency is increased by about 1.3 times. The deep learning sea target detection compared with segmentation algorithm has strong target characterization ability and can be applied to ship targets of different scales.

Data Availability

All the data in the article are real and available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Hainan Provincial Natural Science Foundation of China (519QN180), Hainan Provincial Key R&D Plan (ZDYF2019014), National Natural Science Foundation of China and Macau Science and Technology Development Joint Fund (0066/2019/AFJ), and the program and the Scientific Research Foundation of Hainan University (KYQD(ZR)1859).

References

- [1] G. Gao, K. Ouyang, Y. Luo, S. Liang, and S. Zhou, "Scheme of parameter estimation for generalized gamma distribution and its application to ship detection in SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1812–1832, 2017.
- [2] D. Wang, F. X. Cui, and Z. Z. ShiJiang, "Multi-feature sea-land segmentation based on pixel-wise learning for optical remote-sensing imagery," *International Journal of Remote Sensing*, vol. 38, no. 15, pp. 4327–4347, 2017.
- [3] Y. He and L. Yao, "Space-based sea target information awareness and fusion," *Geomatics & Information Science of Wuhan University*, vol. 42, no. 11, pp. 1530–1536, 2017.
- [4] A.C. Wölfel, J. Jencks, and G. Johnston, "Identifying deep-sea target areas for a pilot Atlantic seabed mapping project using GIS techniques," *Journal of Ocean Technology*, vol. 12, no. 4, pp. 28–42, 2017.
- [5] X. Qiao, L. J. Bao, and D. J. LiZou, "An automatic active contour method for sea cucumber segmentation in natural underwater environments," *Computers and Electronics in Agriculture*, vol. 135, pp. 134–142, 2017.
- [6] R. Fu, B. Li, Y. Gao, and P. Wang, "Fully automatic figure-ground segmentation algorithm based on deep convolutional neural network and grabcut," *Iet Image Processing*, vol. 10, no. 12, pp. 937–942, 2016.
- [7] R. Liu, D. Wang, and P. Jia, "Overview on small target detection technology in infrared image," *Laser & Optoelectronics Progress*, vol. 53, no. 5, Article ID 050004, 2016.
- [8] Y. Ren, "An image segmentation algorithm based on K-means algorithm and otsu," *Boletín Técnico/technical Bulletin*, vol. 55, no. 9, pp. 78–83, 2017.
- [9] O. Paula López, F. Laura Docío, and M. Carmen García, "SEA_AP: a segmentation and labelling tool for prosodic analysis," *Dialectologia*, no. 6, pp. 223–244, 2016.
- [10] K. Yan, Y. Bai, and H. C. Wu, "Robust target detection within sea clutter based on graphs," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 99, pp. 1–11, 2019.
- [11] L. Sun, Y. Li, and Y. Zou, "Pig image segmentation method based on improved graph cut algorithm," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 33, no. 16, pp. 196–202, 2017.
- [12] Y. Yanfeng, Y. Wang, P. Liu, H. Luo, B. Cheng, and H. Sun, "Infrared LSS-target detection via adaptive TCAIE-LGM smoothing and pixel-based background subtraction," *Photon Sensor: English Version*, vol. 9, no. 2, pp. 179–188, 2019.
- [13] L. Lu and J. Wang, "Image processing and recognition algorithm for target tracking," *International Journal on Smart Sensing and Intelligent Systems*, vol. 9, no. 1, pp. 353–376, 2016.
- [14] W. Yu, Y. Wang, and H. Liu, "Superpixel-based CFAR target detection for high-resolution SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 1–5, 2016.
- [15] J. Enyuan and W. Xuejun, "Vehicle abnormal behavior detection based on trajectory tracking," *Journal of Jilin University (Information Science Edition)*, vol. 34, no. 1, pp. 98–103, 2016.
- [16] A. N. Katulev and A. A. Khramichev, "Automatic wavelet-based segmentation of a background-and-target frame from an optoelectronic device for detection of dynamic objects in 2D images," *Journal of Optical Technology*, vol. 83, no. 2, p. 98, 2016.
- [17] L. P. Dorado-Munoz and D. W. Messinger, "Initial study of Schroedinger eigenmaps for spectral target detection," *Optical Engineering*, vol. 55, no. 8, Article ID 083101, 2016.
- [18] Z. Yuan, Z. Jianming, and L. Xiao, "Infrared small target detection using weighting LoG operator," *Journal of Applied Optics*, vol. 38, no. 1, pp. 114–119, 2017.
- [19] P. Lv, C. S. Sun, and G. Liu, "A method for weak target detection based on human visual contrast mechanism," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 2, pp. 261–265, 2019.
- [20] K. Liu, "Research on improved fingerprint image compression and texture region segmentation algorithm," *International Journal Bioautomation*, vol. 21, no. 3, pp. 231–240, 2017.
- [21] Q. H. Wu, J. Wu, and L. Zhu, "Image segmentation algorithm based on graph theory and FCM," *Chinese Journal of Liquid Crystals & Displays*, vol. 31, no. 1, pp. 112–116, 2016.
- [22] Y. Liu and S. Hu, "Blind restoration based on detection and segmentation of motion blurred image," *Xi Tong Gong Cheng Yu Dian Zi Ji Shu/Systems Engineering and Electronics*, vol. 39, no. 3, pp. 662–667, 2017.
- [23] B. Gao, X. Li, W. L. Woo, and G. Y. Tian, "Physics-based image segmentation using first order statistical properties and genetic algorithm for inductive thermography imaging," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2160–2175, 2018.