Applied Ocean Research 106 (2021) 102448

Contents lists available at ScienceDirect

Applied Ocean Research

journal homepage: www.elsevier.com/locate/apor

Ship Detection in Multispectral Remote Sensing Images via Saliency Analysis

Wensheng Wang [*,a], Jianxin Ren [a], Chang Su [b,c], Min Huang [a]

[a] School of Mechanical and Electrical Engineering, Beijing Information Science and Technology University, Beijing 100192, China
[b] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
[c] College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:
visible remote sensing
multispectral
weighted least squares
saliency detection
ship detection

ABSTRACT

Despite the increasing visible optical remote sensing cameras equipped with panchromatic and four-band multispectral sensors, the application of multispectral data is still rarely used in the field of remote sensing ship target detection with the openly recognized challenge to improve detection precision. Towards this end, a ship target saliency detection method is proposed in the paper, which combines weighted least squares (WLS) with maximum symmetric surround (MSS), based on RGB-NIR four-band multispectral remote sensing images to locate the candidate area of ship targets quickly and accurately. The high frequency information of the NIR band image is extracted through the WLS filter and integrated into the RGB band image, and then the saliency of the image is analyzed. The detection results are combined with AIS to achieve complementary information for ship recognition. Some experiments show that the proposed method can effectively suppress the complex background information of clutter interference such as cloud waves and sea waves, highlight the ship target in low contrast scenes, and also increase recall and precision effectively.

1. Introduction

The monitoring of marine targets through optical remote sensing has become an important monitoring method, and has increasingly important value in the military and civilian fields. One of the biggest challenges for ship detection in optical remote sensing images is to locate candidate areas from complex backgrounds quickly and accurately.

At present, many visible remote sensing cameras in the world are equipped with visible panchromatic detectors and four-band multispectral (RGB-NIR) detectors, *e.g.*, Chinese "JL-1" wide format satellite camera (launched in January 15, 2020), China-Brazil Earth Resources Satellite 04A (launched in December 20, 2019), GF-12 Satellite (launched in November 28, 2019), and India ISRO Cartosat-3 (launched in November 27, 2019), etc. The same is true for well-known optical commercial remote sensing satellites such as QuickBird and SPOT series in some European and American countries. The panchromatic visible spectrum band is in the visible range (400-700 nm), and the NIR band (700-1000 nm) is beyond the visible range. It was shown in Chen et al. (2014) that the RGB-NIR data exhibits wider range of characteristics than the panchromatic visible spectrum. The combination of RGB and NIR data have provided good function and effect for image recognition and classification in the life scenes non-remote sensing, see Lezama et al. (2017); Pan et al. (2013) for examples. Therefore, it is both practically and theoretically important to study if we may use multi-spectral data for target detection in remote sensing images.

For the ships detection of optical remote sensing image, despite many research efforts (Han et al., 2014; Proia and Page, 2010; Qi et al., 2015; Song et al., 2018; Sun et al., 2012; Zhou et al., 2018), most of them are currently focused on the analysis of visible panchromatic optical remote sensing images, but relatively few results haven been obtained about the detection and analysis of ship target for multispectral images. The same is true of the saliency detection algorithm for remote sensing data. Later, many efforts are devoted to improve the existing saliency detection algorithms and apply them to the saliency detection of remote sensing images (Han et al., 2015; Nie et al., 2020; Wan et al., 2019; Zhang and Zhang, 2017). However, most of the research on such improved saliency detection algorithms for remote sensing images is focused on gray-scale panchromatic remote sensing images, and rarely involves multispectral and near-infrared images. Although wang *et al.* Wang et al. (2013a) proposed a saliency detection algorithm that uses the color and texture features of the near-infrared image to produce more accurate results, the method does not consider the influence of the

* Corresponding author.
E-mail address: ws_wang@bistu.edu.cn (W. Wang).

https://doi.org/10.1016/j.apor.2020.102448
Received 27 August 2020; Received in revised form 14 October 2020; Accepted 9 November 2020
Available online 20 November 2020

characteristics of the near-infrared image itself and the spectral information on the saliency, just add one-dimensional image data.

Scholars have studied and proved that near-infrared band images have a certain positive effect on image saliency detection. A multi-spectral image dataset of ordinary life scenes was produced as a new platform for saliency research(Wang et al., 2013b). Some regression model experiments have proved the effectiveness using the data set which includes RGB and NIR bands images in saliency detection. Due to the near-infrared image is not easy to obtain in life scenes, a method of directly generating near-infrared images from RGB images by a shallow neural network is proposed (DAI et al., 2019). This method effectively solves the problems in this direction that the difficulty of acquiring near-infrared images and hinders the development. However, the synthesized near-infrared image is not obtained by real imaging, and the imaging mechanism is essentially different from the real near-infrared image. The traditional SLR camera was modified into a four-band camera, which captured hundreds of color (RGB) and near-infrared (NIR) scene images (Brown and Susstrunk, 2011). The calculation of the maximum entropy and joint entropy of the data set shows that the NIR channel provides significantly different information from the R, G, and B channels, and the experiment shows that the addition of near-infrared information leads to a significant improvement of scene detection and recognition performance. The above researches prove that the near-infrared band is of great significance for the detection of image saliency. Unfortunately, these saliency detection algorithms focuse on daily life pictures rather than remote sensing data. So naturally, whether the introduction of near-infrared images in remote sensing data will bring positive effects is the question of this paper.

Optical remote sensing satellites can be used as independent systems to detect and track ships. However, it was unable to provide relevant information about the identity of the ship. Automatic Identification System (AIS) is a collaborative self-reporting system that can provide accurate static data (such as name, type, length and width), dynamic data (such as location, course over ground-COG and speed over ground-SOG), and some voyage-related data (Last et al., 2014). Currently, the main platforms of AIS are ground AIS and satellite AIS. Satellite AIS is not restricted by region, making it powerful, and it can provide the possibility of global maritime surveillance. On the one hand, AIS data can not only be used be used as ground-truth verification for ship detection in optical remote sensing satellite images, but also can be fused with optical satellite data to provide more information about the identity of noncooperative ships which are not broadcasting their positions. On the other hand, it is possible to pin-point those ship candidates that do not carry an AIS, and thereby take appropriate security or rescue actions. Therefore, the combination of optical satellite data ship detection and AIS can overcome the limitations of the two and effectively improve maritime surveillance capabilities.

It is openly recognized that locating the candidate area is one of the most important steps of target detection in order to be able to extract the target quickly and efficiently. By analyzing the characteristics of the four-band remote sensing image, we found that the longer the wavelength of the NIR image, the more prominent the high frequency information, including the high frequency information of the target and some background. After extracting the high frequency of the NIR image with WLS filter, Combined with the low frequency information of the RGB image, the reconstructed image high frequency can well surround the low frequency information, the saliency area extracted by the MSS method is more obvious, and it is easier to extract the saliency area. Motivated by this fact, a new method named Weighted least squares-maximum symmetric surround (WLS-MSS) is proposed in the paper to analyze the saliency of ship targets for RGB-NIR four-channel multi-spectral remote sensing images, and obtain candidate areas for ship target detection and recognition.

The remainder of this paper is structured as follows. Section 2 presents a comprehensive description of the WLS-MSS model and analyze the effectiveness of this model. In Section 3, the experiments and analysis of result is done. Finally, the conclusion is presented in Section 4.

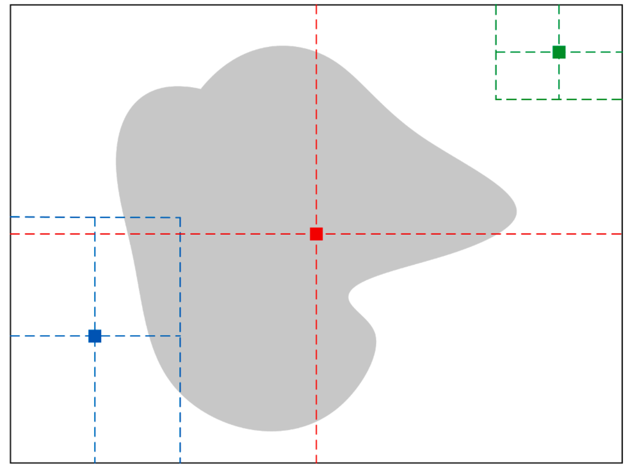


Fig. 1. Schematic diagram of maximum symmetrical surround

Highlights

- **C1** Instead of commonly used panchromatic single-band data, we utilize multi-spectrum data for ship targets detection. Using multi-spectral data can reflect the target spectral characteristics, which is more conducive to target detection.
- **C2** For ship targets saliency detection, we propose a novel model named WLS-MSS with superior precision and recall using four-band multi-spectral remote sensing data. The model is robust, and it can detect dark targets and small targets in complex backgrounds.

2. Relevant Theories

2.1. Maximum symmetric surround visual attention model

First, we introduce the maximum symmetric surround algorithm (Achanta and Susstrunk, 2010). Fig. 1 is a schematic diagram of the maximum symmetrical surrounding of each pixel, the area used is the sub-image of "maximum symmetrical surround". "Maximum symmetrical surround" is a maximum rectangular sub-image with the target pixel as the center of its own sub-image and the entire image boundary as the limit. As shown in the blue and green pixels in Fig. 1, the corresponding area is centered on blue or green pixels, and the area set with the minimum distance from the image in the x and y directions as the side length is the sub-image of "maximum symmetrical surround". As shown in the blue and green dotted boxes in Fig. 1, the largest sub-image in the whole picture is the image center point, that is, the sub-image corresponding to the red pixel point.

For a $w \times h$ image, the saliency values of image pixels defined by the MSS algorithm are

$$S(x,y) = \| I_\mu(x,y) - I_\omega(x,y) \|, \tag{1}$$

where $\| \cdot \|$ is $\mathcal{L}_2$ norm representing the Euclidean distance, $I_\omega(x,y)$ is an image smoothed by a Gaussian filter in Lab color space, $I_\mu(x,y)$ is the characteristic mean value of the sub-image of "maximum symmetrical surround" of (x,y) pixel in Lab color space. The calculation formula is as follow

$$I_\mu(x,y) = \frac{1}{S} \sum_{i=x-x_0}^{x+x_0} \sum_{j=y-y_0}^{y+y_0} I(i,j), \tag{2}$$

The sub-image in the above formula is the maximum possible symmetrical surrounding area of a given central pixel in Lab color space, the

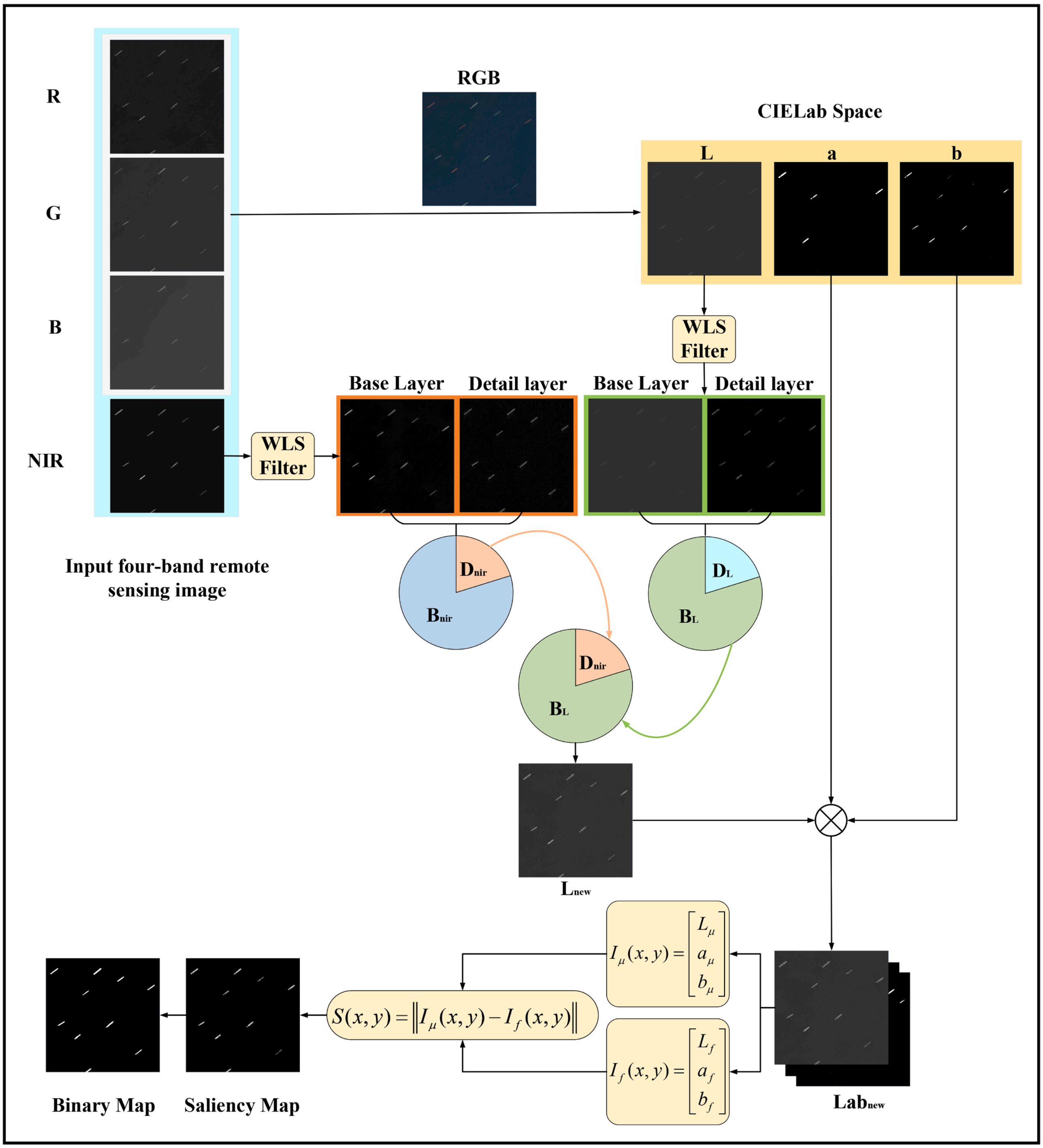
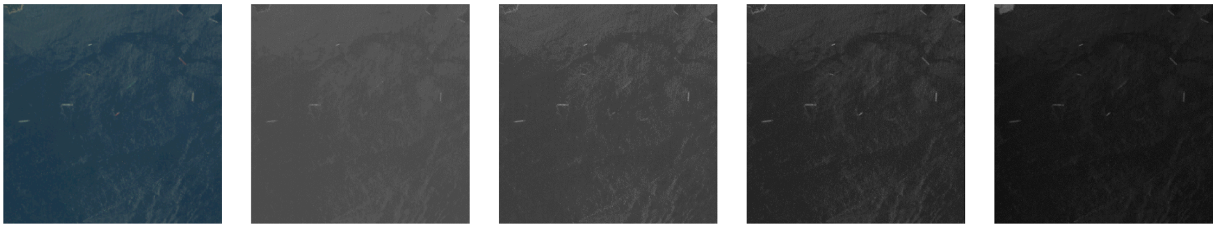


Fig. 2. Flow chart of ship saliency detection of four-band remote sensing image



(a) RGB (b) B-band (c) G-band (d) R-band (e) NIR-band

Fig. 3. Picture of ships under the interference of four-band waves. (a) RGB color image (b) Blue band image (c) Green band image (d) Red band image (e) NIR band image

offset x_0, y_0, and area S are calculated as

$$\begin{aligned} x_0 &= \min(x, w-x), \\ y_0 &= \min(y, h-y), \\ S &= (2x_0+1)(2y_0+1), \end{aligned} \tag{3}$$

Therefore, the closer the center pixel is to the edge of the object, the narrower it is, that is, the sub-image with the object partial area centered on the center pixel is smaller. The acquisition process of $I_\mu(x,y)$ can essentially be regarded as the mean filter of the adaptive window, which can average the energy of object boundary and interior with high energy and can average the low energy of the background area. The algorithm can suppress the low-brightness background of the image and highlight the foreground area of the high-brightness target. However, when there is a lot of high-frequency information in a complex background, the target will be insignificant.

2.2. Weighted least squares filtering

The weighted least squares optimization framework(Farbman et al., 2008) is a non-linear, edge-preserving smoothing method that can capture details at various scales through the multi-scale decomposition of preserved edges. An edge-preserving filter can be seen as a combination of two contradictory goals. We want the target image to be as close as possible. At the same time, except where there is a large change in the edge gradient, the smoother the better, which is expressed mathematically as:

$$\sum_p \left(\left(u_p - g_p\right)^2 + \lambda \left(a_{x,p}(g) \left(\frac{\partial u}{\partial x}\right)^2 + a_{x,p}(g) \left(\frac{\partial u}{\partial y}\right)^2 \right) \right), \tag{4}$$

where the subscript p represents the spatial position of the pixel. The goal of *data term* $(u_p - g_p)^2$ is to minimize the distance between u and g, while theregularization tries to achieve smoothing by minimizing the partial derivatives of u. The smoothness requirements vary spatially by the smoothness weights a_x and a_y depending on g. Finally, λ is responsible for the balance between the two terms; increasing the value of λ will cause the image u to gradually smooth.

The Eq. (4) can be rewritten into matrix form:

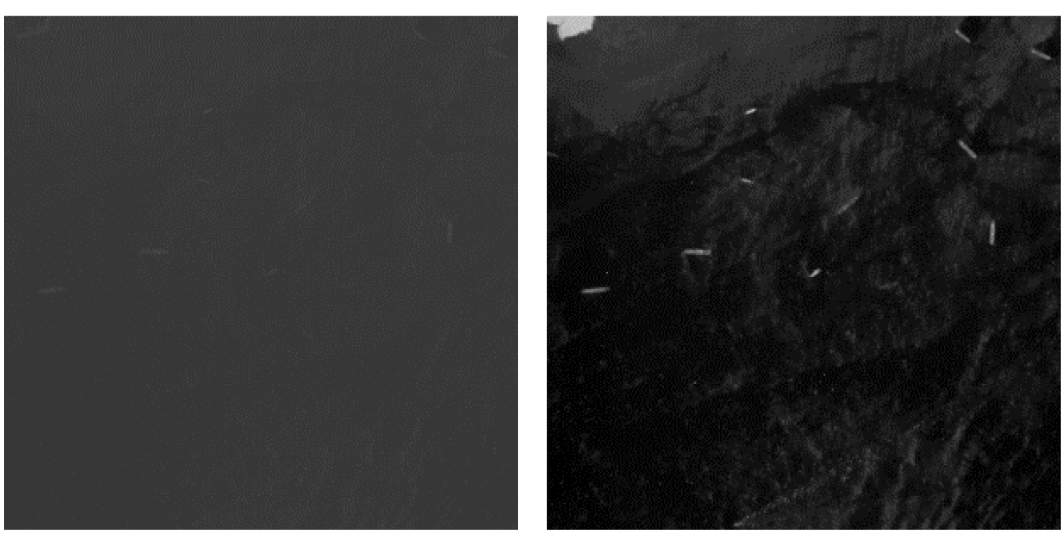
$$(u-g)^\top (u-g) + \lambda \left(u^\top D_x^\top A_x D_x u + u^\top D_y^\top A_y D_y u \right), \tag{5}$$

where A_x and A_y are diagonal matrices with $a_x, p(g)$ and $a_y, p(g)$ as diagonal elements, D_x and D_y are forward difference matrices, and are backward difference operators. To make the minimum value of Eq. (5), u should meet the following requirements:

$$(I + \lambda L_g)u = g, \tag{6}$$

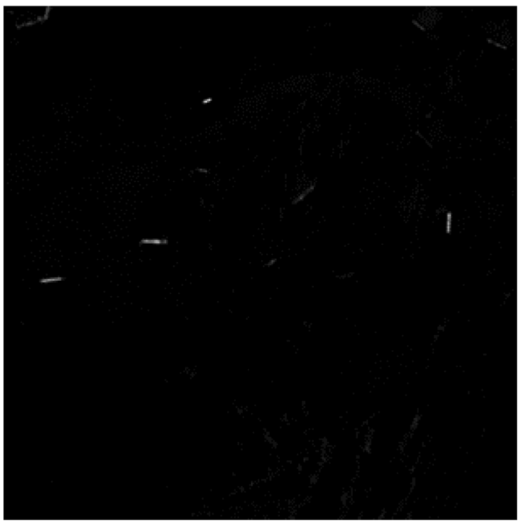
which $L_g = D_x^\top A_x D_x + u^\top D_y^\top A_y D_y$.

To obtain an approximate image $g^{\mathtt{filter}}$, which is as close as possible to the input image g, while the obviously gradient along g is as smooth as possible at the same time, we seek solutions that minimize the following objective functions:

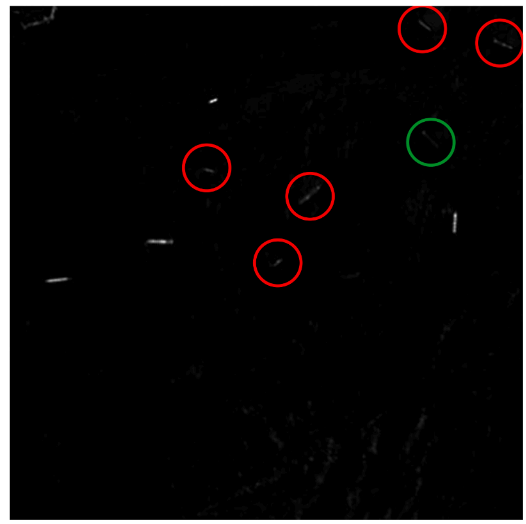


(a) L (b) L_{new}

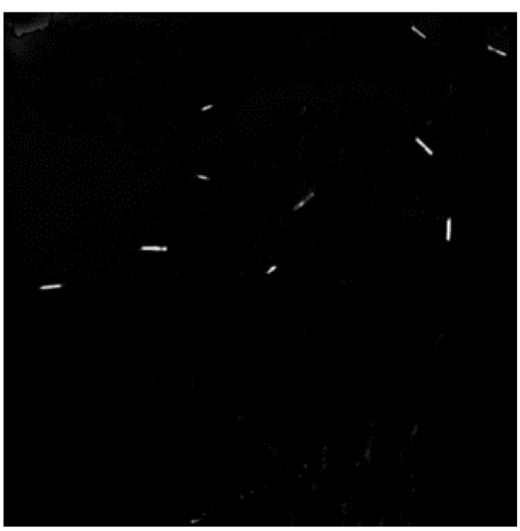
Fig. 4. The original L band image and the L_{new} image reconstructed by adding high-frequency information of NIR image



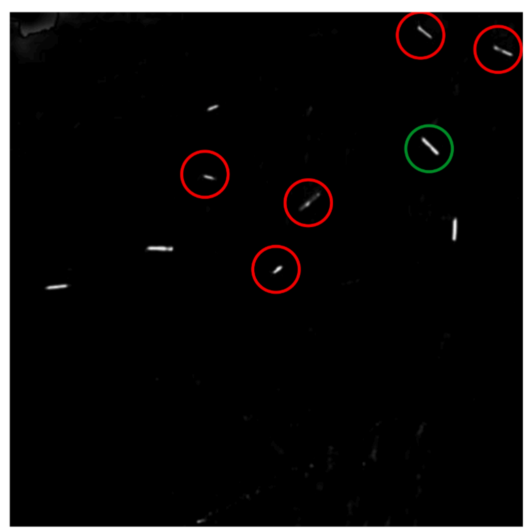
(a) the result of traditional MSS algorithm



(b) the emphasized result targets of traditional MSS



(c) the result of WLS-MSS method



(d) the emphasized result targets of WLS-MSS

Fig. 5. Comparison of saliency detection results between traditional MSS algorithm and WLS-MSS method

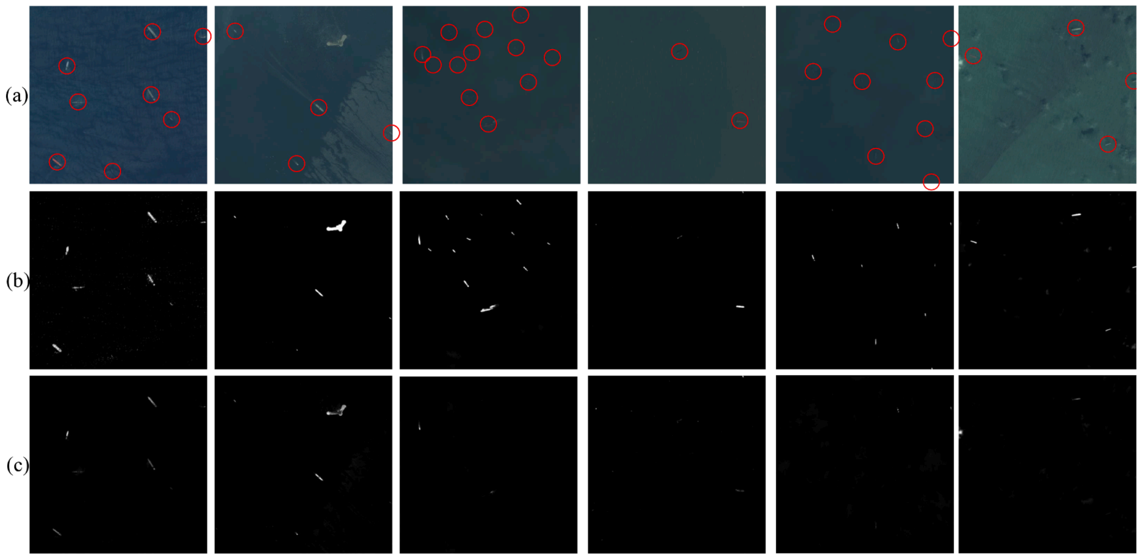


Fig. 6. Comparison of detection results in different scenes. Line (a) is the original RGB image, the ship target is circled, line (b) is the improved algorithm detection result, and line (c) is the original algorithm detection result.

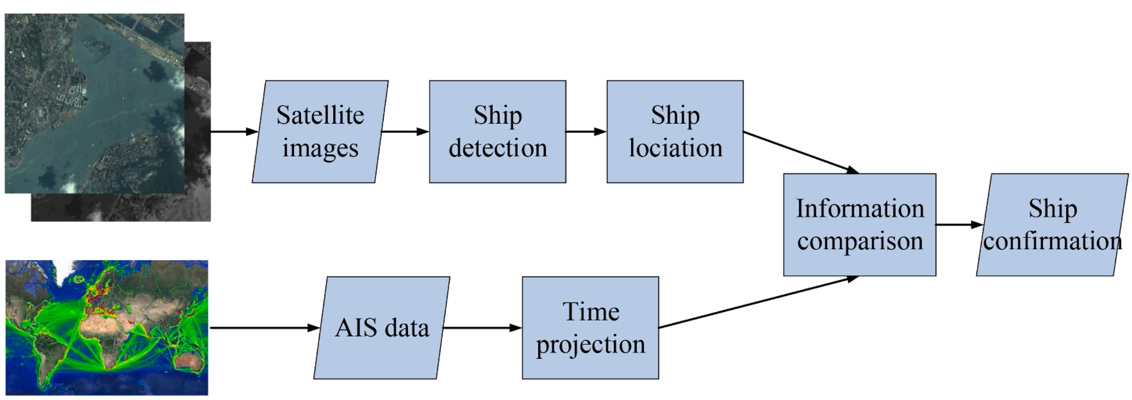


Fig. 7. Procedure of fusion between optical satellite remote sensing image and AIS data

$$g^{\text{filter}} = F_{\lambda}(g) = \left(I + L_g\right)^{-1} g, \quad (7)$$

with the subscript representing the pixel position in space. The first term of the objective function represents that the more similar the input image and the output image are, the better, and the second term is the regular term, which makes the output image as smooth as possible by minimizing the partial derivative. The weight of the smoothing term depends on the input image. When the edge gradient of the input image changes greatly, we want its constraint to be smaller to retain the structural information of the image, and when the edge gradient of the image changes very little, we think these details are not important. The constraint can naturally be larger, it is a regular term parameter. Balancing the proportion of the two, the larger the proportion of the two is, the smoother the image will be.

2.3. WLS-MSS saliency detection model

Because in the optical remote sensing images, the image contrast is often relatively low, especially in marine areas, sometimes targets such as ships are not prominent. Compared with the visible image, the near-infrared image can better reflect the image details, which is of great significance for image enhancement. In order to make the ship target more prominent, we combined the near-infrared image, used the WLS filtering method to enhance the high-frequency information of the image, and loaded the high-frequency area of the near-infrared image onto the visible image, and reconstructed the Lab image transformed from the visible remote sensing image. Finally, the reconstructed Lab image is used to calculate the saliency in the largest symmetric surrounding area.

Fig. 2 illustrates the entire process of the proposed method. Given the four-channel multispectral input image, we first convert the RGB three-channel image to the CIELab color space and obtain the brightness images L and two chroma images a and b. Then, the NIR channel image and the L channel image are decomposed into base and detail layers by WLS filter. The operation process only needs to subtract the low-frequency filtered image from the original image to obtain the high-frequency image with details. The base image contains low-frequency content, which is usually a smoother area, while the detail layer includes high-frequency content with edges and sharp transitions(such as noise). In this way, we get the base image B_{nir} and the detail image D_{nir} of the NIR image, as well as the base layers B_L and the detail layer D_L of the brightness channel L. The base layer of the RGB image B_L contains low-brightness information perceived by the human visual system, so the base layer of NIR is discarded. We combine the detail layer D_{nir} of the NIR image with the base layer B_L of the Lab color space image brightness map L obtained by using WLS to obtain a new brightness image L_{new}. This new brightness image L_{new} enhances the contrast and detail of the original image. Next, it is combined with the chrome images a and b of the Lab color space image to reconstruct the final image Lab_{new}. Finally,

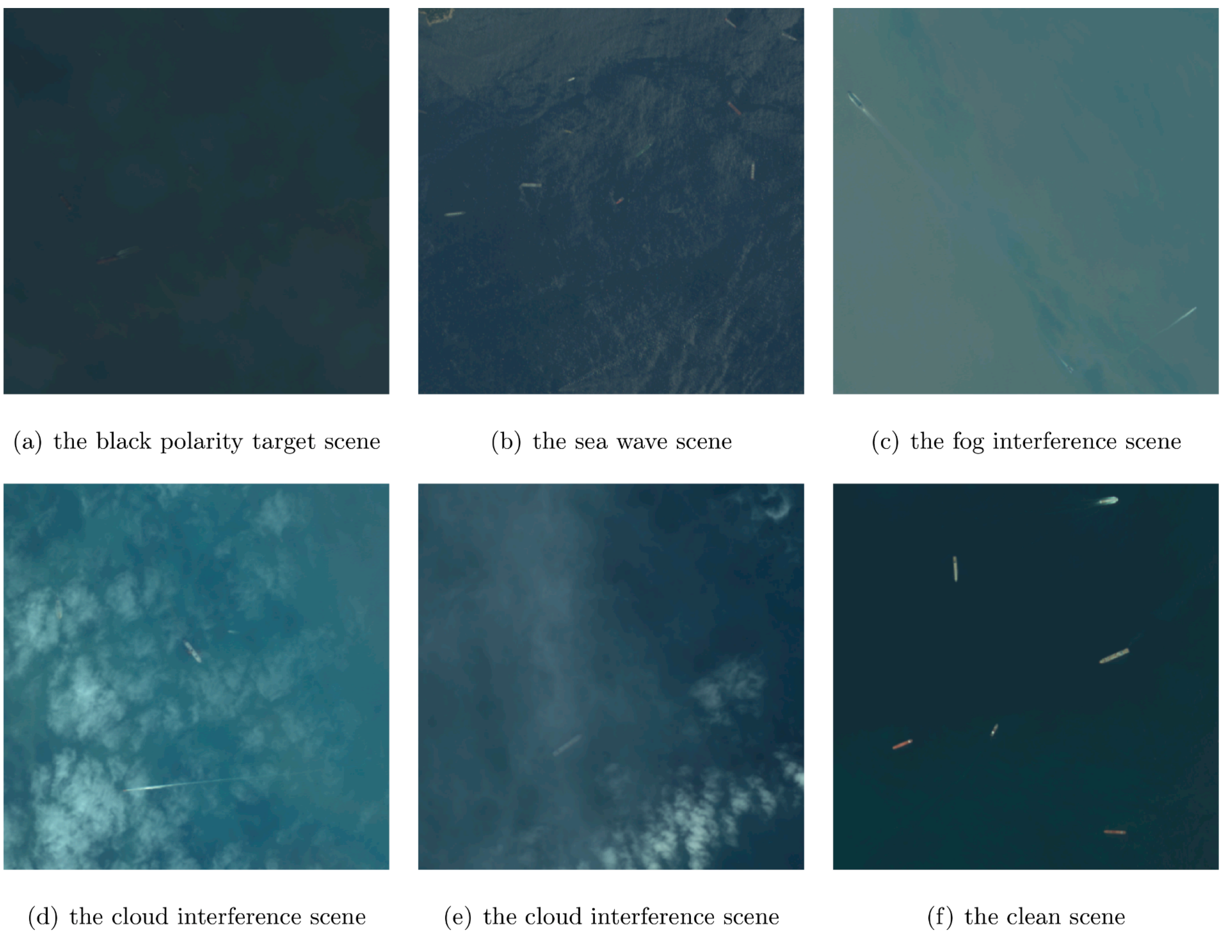
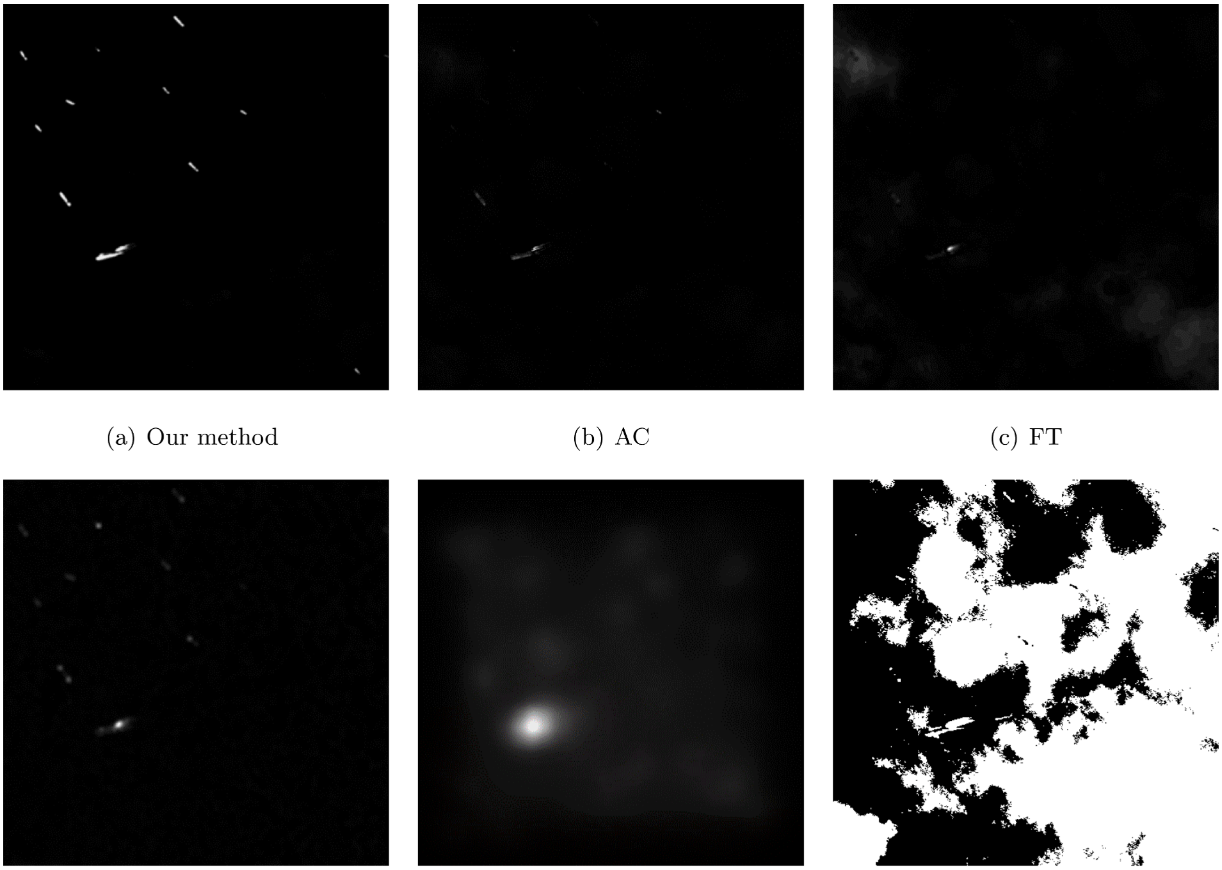


Fig. 8. The typical scenes of original color remote sensing image. (a) is the black polarity target scene, (b) is the sea wave scene, (c) is the fog interference scene, (d) and (e) are the cloud interference scenes, (f) is the clean scene.



(a) Our method (b) AC (c) FT

(d) SR (e) GBVS (f) OTSU

Fig. 9. The black polarity target scene result of image Fig. 8(a)

the new image is calculated by the maximum surround model to obtain the final saliency image.

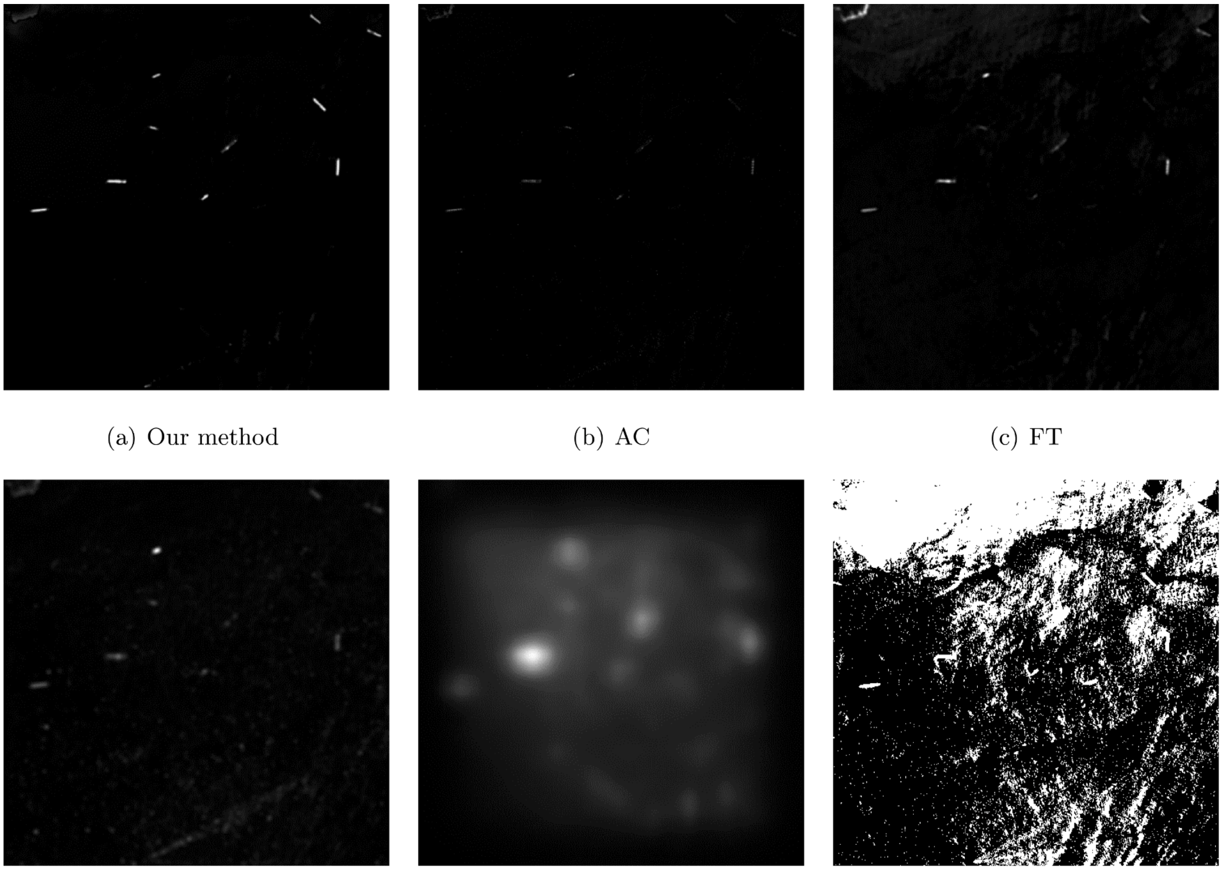
Fig. 3 shows the B, G, R, and NIR four-channel images of the input image, which contain ten ships. The background is filled with a lot of the clutter formed by the waves, and the clutter is not uniform. It can be seen from the Fig. 3 that the longer the wavelength of the NIR image, the cleaner the background and the more prominent the target. Fig. 4 shows the L-channel image under the original Lab transformation and the reconstructed L-channel image after adding the near-infrared image. In the Fig. 4, we can see that the reconstructed L-channel clearly highlights the target and suppresses the background after adding the near-infrared image Fig. 5 (a) is the saliency map obtained by the traditional MSS saliency detection algorithm, and (b) is the saliency map using WLS filtering. We can see that there are four ships in the traditional MSS algorithm that are not as obvious as the improved ones, as shown by the red circle; in addition, one ship is almost lost, as shown by the green circle. Fig. 6 shows the detection results of several other different scenes (including wave, low contrast and cloud and fog scenes), (a) (b) are the detection results under the interference of sea waves, (c) (d) is the detection results of dark polar ships, (e) is the detection result of low contrast under the influence of fog, and (f) is the detection result of cloud interference. It can be seen that the improved algorithm effect is significantly better than the traditional algorithm.

It can be seen that after the near-infrared image added, the target is obviously highlighted, the WLS algorithm is used to extract high-frequency information, and it is reconstructed with an L-channel image transformed from the RGB image and then combined with the original MSS algorithm, it has a good effectto extract the target's significance in a variety of complex scenes.

2.4. Ship identification and confirmation combined with AIS

Here, Fig. 7 describes the workflow of the proposed method of fusing optical satellite ship detection and AIS. First, the ship detection is performed on the optical remote sensing image through the WLS-MSS method. Then, the ship positioning is realized via coordinate conversion, and the AIS data is projected to the remote sensing image acquisition time. Finally, extracted in the ships detection step, features (mainly geometric features) are compared with the information provided by the AIS to confirm the targets.

1) Geographical location: To obtain the geographic location of the ships, the pixel coordinates need to be geo-referenced. For optical remote sensing data, the rational polynomial coefficients (RPCs) that directly describes the ground-to-image transformation by two-part cubic polynomial (Oh and Lee, 2015) is usually provided with the image. When the ship is at sea, the height can be set to approximately 0. The inverse transformation (image to ground) can be used to calculate the corresponding ground coordinates.
2) Time projection: Given remote sensing images, a spatiotemporal filtering is required to select only the AIS messages useful for data fusion. Based on the ships Maritime Mobile Service Identification (MMSI) number, AIS data is organized by location. A time projection step needs to be performed to obtain the ship positions of all AIS ship tracks at the time of remote sensing image acquisition, so as to obtain the time alignment between satellite image ship detection and AIS data. The approximate location of AIS data should be linearly interpolated or extrapolated to the acquisition time of satellite images because AIS transmission usually occurs continuously, but



(a) Our method (b) AC (c) FT (d) SR (e) GBVS (f) OTSU

Fig. 10. The sea wave scene result of image Fig. 8(b)

remote sensing image acquisition is only performed at fixed intervals (Yao et al., 2019).

3) Information association and target confirmation: In the process of simultaneously observing remote sensing image and AIS, the following situations may occur: target and AIS target represent the same ship, and AIS target has no counterpart in image target (e.g., due to cloud shadowing), or the target is not displayed in AIS (e.g., due to lack of AIS transponder), the two systems can complement each other. Through Yao et al. (2019) tracking correlation method, the correspondence between the same target in the two systems can be solved. Confirm the target by comparing the target features obtained by image detection with the information provided by AIS.

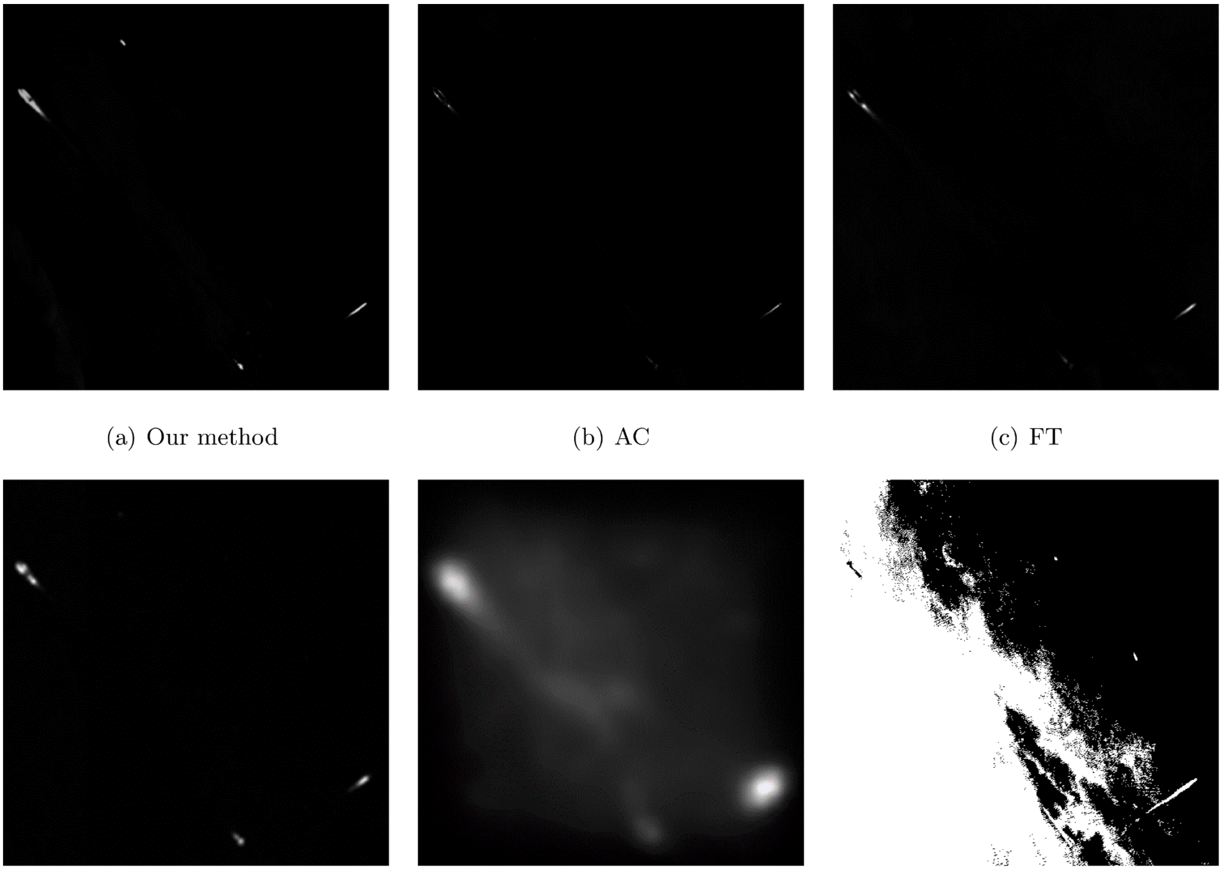
3. Experimental Analysis

3.1. Experimental data and environment

In the experiment, four-band multispectral images of GF-1 and ZY-3 optical remote sensing cameras were used as experimental materials. The spatial resolution of the multispectral detector is 8m, and the corresponding panchromatic image resolution is 2m. 500 groups of multispectral data of different scenes (including calm sea, low contrast, dark polarity, obvious sea clutter, and cloud interference, etc.) were selected for the experiment. The image contains ships of different types and sizes, and the image size was unified at 512*512 pixels.

3.2. Qualitative analysis of experimental results

This section gives the detection results and comparative experiments in different scenes and selects several representative extreme environments for display. The six groups of source images is presented in Fig. 8 (a)-(f) from the first to the sixth group. The current classic saliency detection algorithms AC(Cheng et al., 2015), FT(Yang et al., 2013), SR (Hou and Zhang, 2007), GBVS(Harel et al., 2006) algorithm and traditional adaptive algorithm based on pixel value segmentation (OTSU) (Otsu, 1979) are compared with the proposed method shown from Fig. 9 (a)-(f) to Fig. 14(a)-(f). The first group to the sixth group of experiments are shown from top to bottom. The ship in Fig. 9 is a dark polar target. The contrast between the target and the scene is very low and it is hard to distinguish by the naked eye. The algorithm in this paper detects all the targets very well, while other centralized algorithms lose a large number of targets. For large-area and non-uniform wave interference, as shown in Fig. 10, some ships are in the wave and some are outside the wave. The algorithm in this paper, FT and SR algorithm can detect the target, but the algorithm in this paper is significantly better than the other two algorithms, and the other two algorithms have some clutter not filtered; the AC algorithm misses some targets. For a scene covered by fog, as shown in Fig. 11, there are four targets in the scene and all of them are detected by the proposed method, and the saliency is the best. Several other algorithms will lose 1-2 targets, especially small targets. Several other algorithms will lose 1-2 targets, especially small targets. Fig. 12-13 is the interference of two typical cloud cover targets. The clouds in Fig. 12 are relatively broken and high-frequency, while the clouds in Fig. 13 are relatively slow. As can be seen from Fig. 12, in the background of clouds, the algorithm in this paper and SR algorithm have good anti-cloud interference ability and good resistance to cloud cover and surrounding, and can accurately detect targets, especially those whose upper left corner is obscured by shredded clouds are difficult to



(a) Our method (b) AC (c) FT

(d) SR (e) GBVS (f) OTSU

Fig. 11. The fog interference scene result of image Fig. 8(c)

see with naked eyes, both algorithms can detect well. However, in Fig. 13, the SR algorithm does not detect the target ship. From this we can see that the SR algorithm can be highlighted in the relatively high-frequency areas, but not in the relatively flat low-frequency areas. Although some clouds will be attached to the algorithm in this paper, the detection results do not miss the ship target, so as to achieve the purpose of no missed alarm in the candidate area screening, and then the clouds can be removed in the target screening process. The AC algorithm and the FT algorithm cannot detect ships covered by clouds or near the clouds. In addition, it can be seen from the figure that the overall effect of GBVS algorithm is not good, and it is not suitable for the detection of small targets such as marine targets in remote sensing image; OSTU algorithm is only effective for scenes with very clean calm sea surface, as shown in Fig. 14, while other algorithms of this scene also can basically complete the detection.

The reason why the WLS-MSS model is better than the above-mentioned algorithms for detection is that it not only adds multiple band information of the image, but also organically combines the advantages of the two algorithms. First, the MSS algorithm can shield the background and extract the saliency area of the target. In the background of large waves, the ship will be submerged in the waves. Although the ship will be larger than the wave, the MSS algorithm will integrate the texture of the ship and the background area. Either the significant area cannot be extracted, or both the ship and the waves are extracted. After adding the WLS of the near-infrared image to extract the high-frequency information, the outline of the ship is more prominent. The model effectively increases the contrast between the ship and the sea wave. In this way, when the MSS algorithm is used to extract the saliency area, the ships can be extracted and waves are suppressed, effectively.

In general, in the above several scenarios, some algorithms can only be competent for the detection of one or two or three scenarios, and the algorithm in this paper can be well qualified for the detection task in the above several scenarios, and the saliency is the best. It can be seen from the figure that the algorithm in this paper is resistant to interference from complex backgrounds, and has good effects on targets with low contrast or dark polarity, and can also detect small targets very well.

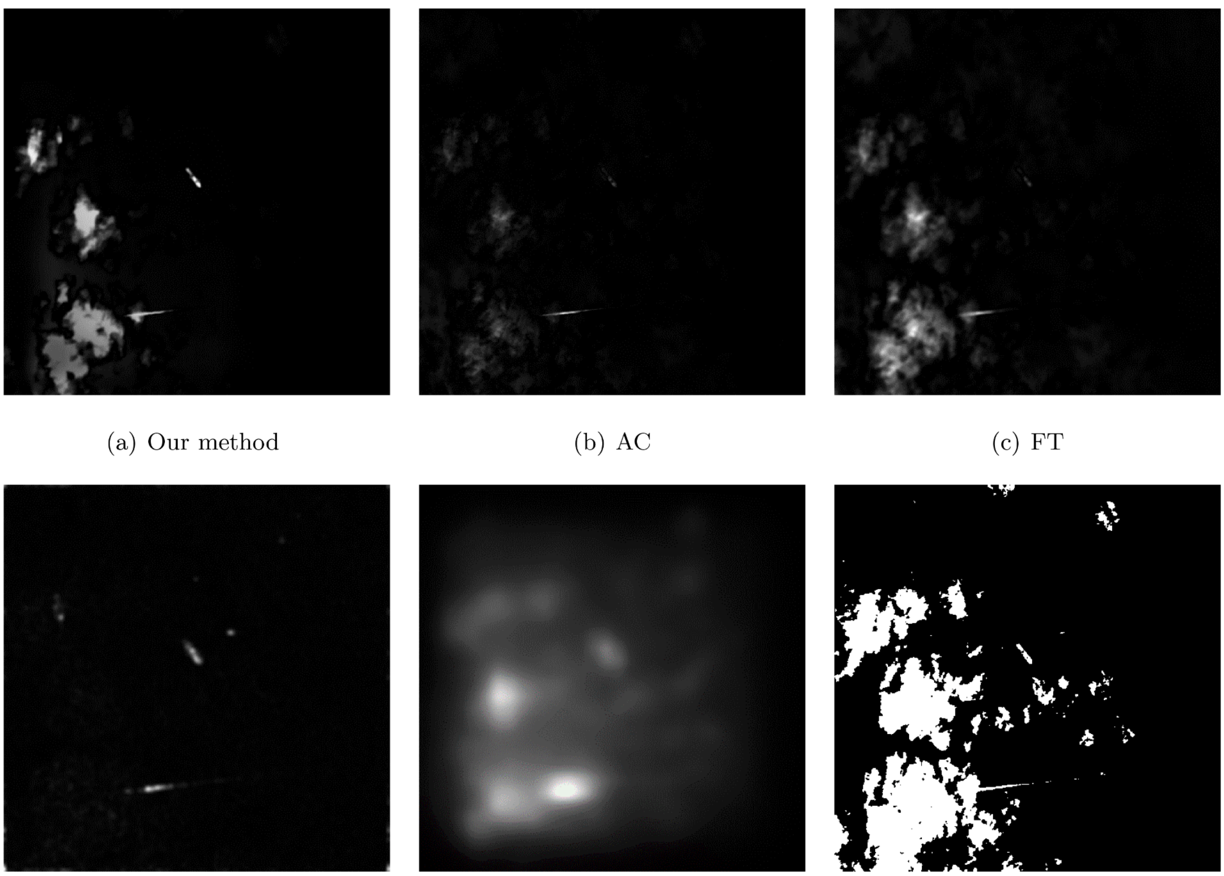
3.3. *Quantitative analysis of experimental results*

Compare the proposed algorithm with the AC algorithm, FT algorithm, and SR algorithm which have a better visual effect and are competent for more scenes, and carry out the quantitative calculation. In the test results, true positive (*TP*), false positive (*FP*), false negative(*FN*), and true negative (*TN*) are used to evaluate algorithm performance. The region in a bounding box is considered as *TP* or *FP* (also called as false alarm). On the other hand, the region of a target is considered as *FN* (also called as miss alarm) if its bounding box exists in the ground truth but no bounding boxes predict it. Otherwise, the region is considered as *TN* (also called as correct rejection). Hence, the precision, recall indicators are formulated as follows:

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{NP}, \tag{8}$$

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{NC}, \tag{9}$$

where *NT* is the total number of targets and *NC* is the total number of detected targets.



(a) Our method (b) AC (c) FT

(d) SR (e) GBVS (f) OTSU

Fig. 12. The cloud interference scenes result of image Fig. 8(d)

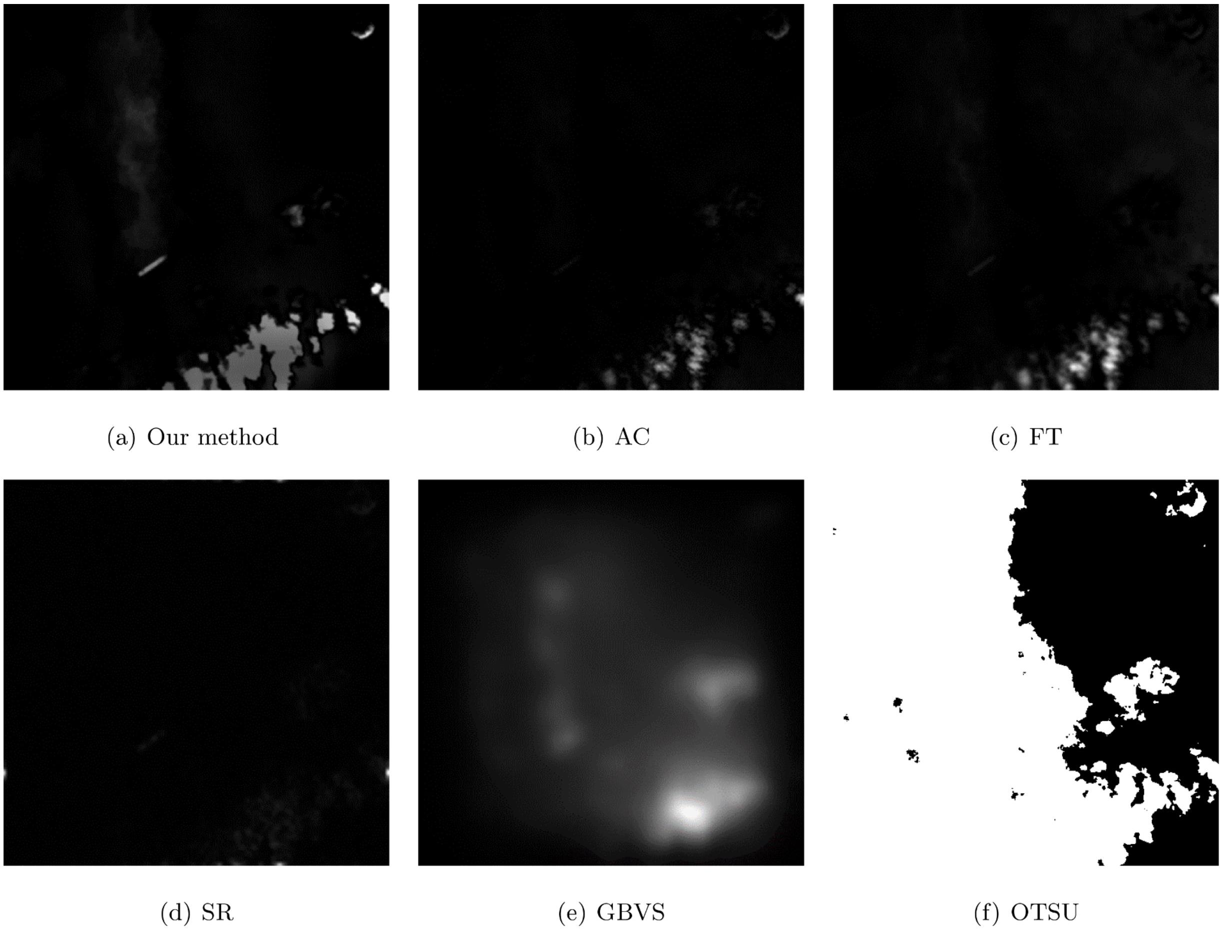


Fig. 13. The cloud interference scenes result of image Fig. 8(e)

The 450 groups of experimental materials including cloud cover scenes, strong sea clutter scenes, and stable sea surface scenes were calculated quantitatively. The results are shown in Tables 1-3.

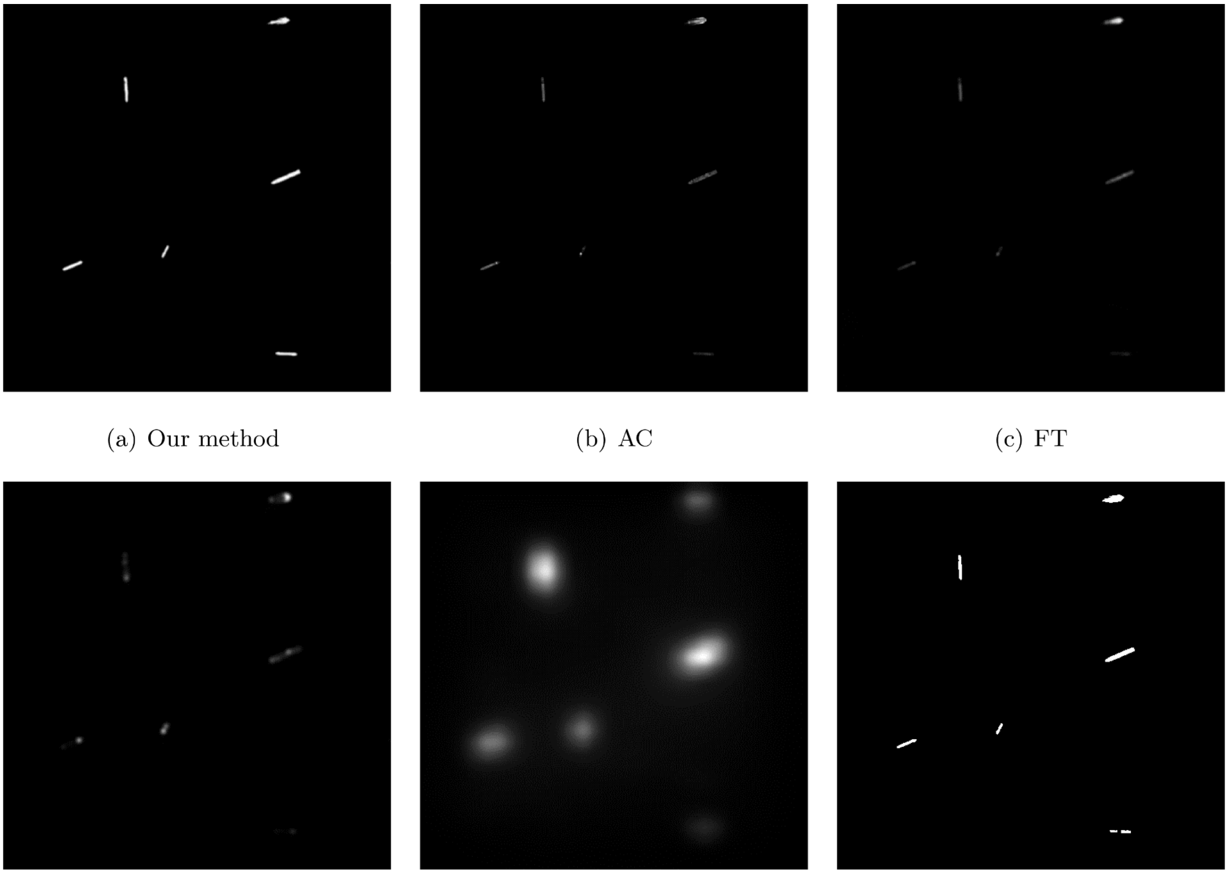
Tables 1 -3 give the quantitative evaluation index results of the proposed method and other algorithms. It can be seen from the table that for the targets of different application scenes, several algorithms have higher detection rates and lower false alarm rates in simple scenes, and the detection effect is good. However, in the case of strong clutter interference, the detection ability of the AC saliency algorithm is not stable, and the detection error rate is high (83.3%). The detection ability of the FT algorithm is slightly better than the AC saliency algorithm, but it is also sensitive to clutter and has a high detection error rate (75%). The SR algorithm has a high detection ability for all three targets, but it is accompanied by the highest false alarm rate for all three targets. The Precision-Recall Curves of three scenes are shown in Fig. 15.

3.4. *The result matching AIS*

After the targets are located by saliency detection, the geometric features are trained and extracted through the Support Vector Machine (SVM) algorithm in the location small box. The final detection result is obtained. Fig. 16 is the result of ship detection and matching with the AIS at the location of E114.2,N22.1 taken by the GF-1 satellite. The red frame is the target segmentation by SVM after the saliency detection locates the ship target. Through the use of inverse transformation (image to the ground) to calculate the corresponding ground coordinates, and time projection, and ship information in AIS for comparison. The green dot indicates the position of the ship in the AIS. Judging whether the matching result is correct through information such as location information and ship area. At the same time, more specific information such as ship model can be obtained according to the AIS. The specific information of the ship circled in yellow in the Fig. 16 is shown in Table 4 In addition, through the figure we can see that some ships have been detected without AIS. In this way, the two systems complement each other and play a more active role in ship inspection and ocean monitoring.

4. Conclusion

The saliency detection algorithm combined with RGB-NIR four-band multispectral remote sensing image marine target proposed in this paper effectively solves the problem of accurate detection of the weak signal saliency detection of ship target image under complex background to accurately locate the candidate area. The algorithm makes full use of the four-band detectors commonly used in existing visible remote sensing cameras, incorporates the characteristics of high-frequency highlights of near-infrared images in imaging into the algorithm, and uses the WLS-MSS method for image saliency analysis and obtain the saliency image of the ship targets. Compared with the existing classical saliency detection algorithms in different scenes, the method proposed in this



(a) Our method (b) AC (c) FT

(d) SR (e) GBVS (f) OTSU

Fig. 14. The clean scene result of image Fig. 8(f)

Table 1
Performance comparison in the simple scene between the proposed method and other algorithms(200 groups of 512*512 multispectral data)

	NT	NC	TP	FP	Recall	Precision
Ours	589	592	580	12	98.5%	98.0%
AC		606	547	59	92.9%	90.3%
FT		615	550	65	93.4%	89.4%
SR		708	560	148	95.1%	79.1%

Table 2
Performance comparison in the scene of strong clutter interference (such as cloud, fog, sea waves) between the proposed method and other algorithms(200 groups of 512*512 multispectral data)

	NT	NC	TP	FP	Recall	Precision
Ours	368	354	332	22	90.2%	93.8%
AC		330	289	41	78.5%	87.6%
FT		348	296	52	80.4%	85.1%
SR		475	315	160	85.6%	66.3%

paper is more excellent in detection ability and false alarm suppression. The experimental results show that the algorithm can fully suppress the complex background information of clutter interference such as cloud wave wake, highlight the ship target, and can effectively reduce the false alarm rate while maintaining a high detection rate, and achieve the detection of single-frame ship target candidate areas. In addition, the combination of optical satellite remote sensing images and AIS can provide more detailed ship information for detection results, which is more conducive to target recognition.

Table 3
Performance comparison in the low contrast and black polarity target scene between the proposed method and other algorithms(200 groups of 512*512 multispectral data)

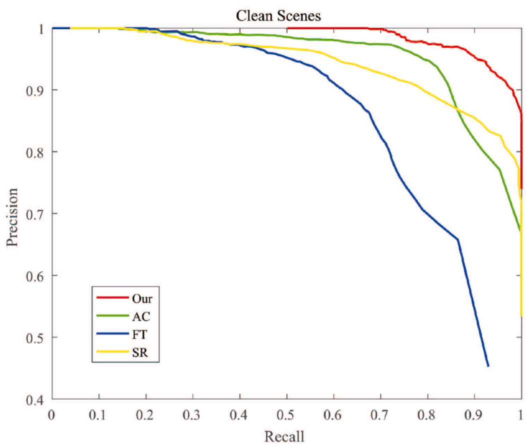
	NT	NC	TP	FP	Recall	Precision
Ours	165	161	158	3	95.8%	98.1%
AC		108	101	7	61.2%	93.5%
FT		126	116	10	70.3%	92.1%
SR		136	124	12	75.2%	91.2%

CRediT authorship contribution statement

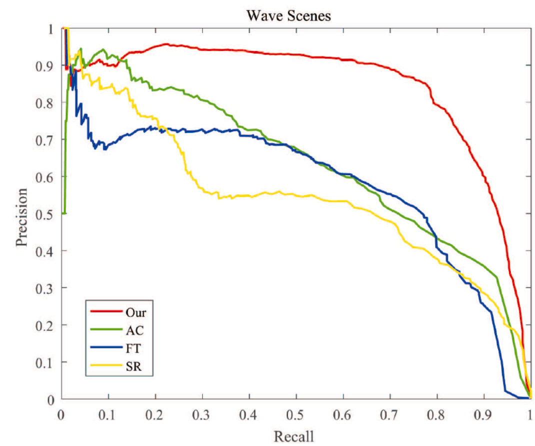
Wensheng Wang: Conceptualization, Methodology, Software, Data curation, Writing - original draft. **Jianxin Ren:** Investigation, Writing - review & editing. **Chang Su:** Validation. **Min Huang:** Resources, Supervision, Project administration.

Declaration of Competing Interest

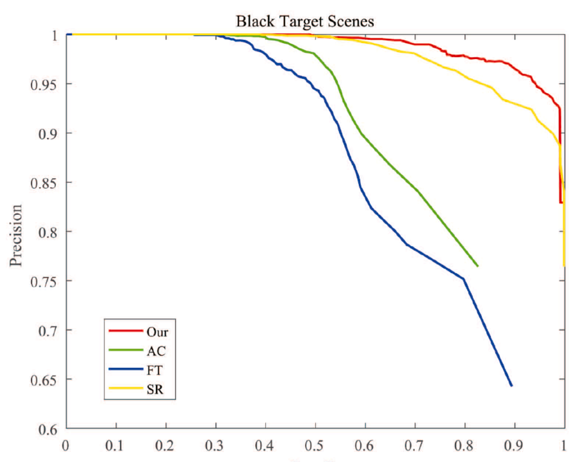
We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, th ere is no professional or other personal interest of any nature or kind in any product, servi ce and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.



(a) The precision-recall curves of clean scenes



(b) The precision-recall curves of wave scenes



(c) The precision-recall curves of black polarity target scenes

Fig. 15. The Precision-Recall Curves of three typical scenes

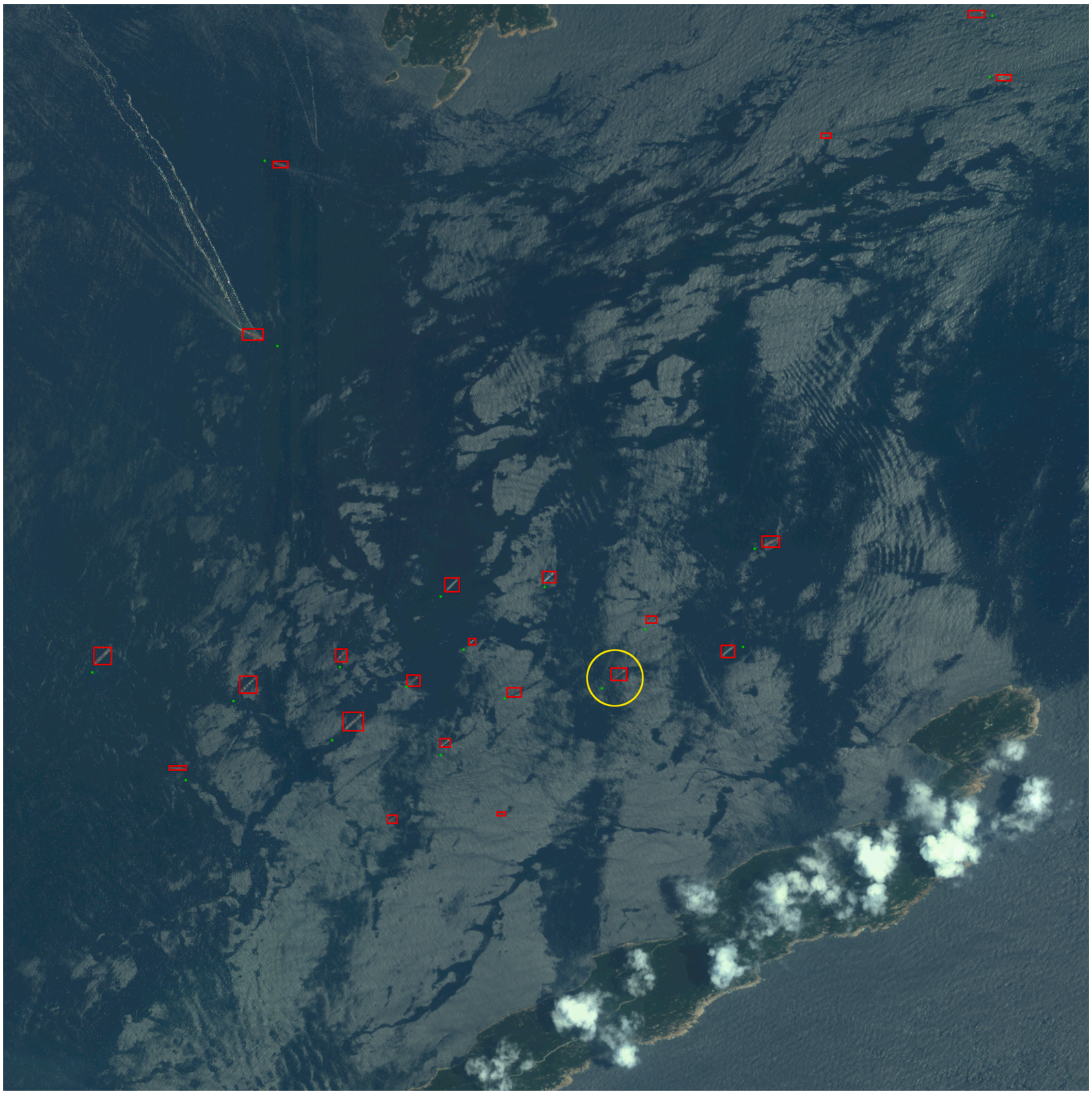


Fig. 16. The detection result at E114.2,N22.1 taken by the GF-1 satellite

Table 4
The information of the ship marked with yellow circle in AIS

MMSI: 564559000	Pre-arrival: HONGKONG
IMO: 9167435	Scheduled time: 10-06 09:00
Callsign: S6ST	Longitude: E 114 degrees 16.2186 minutes
Chinese name:	Latitude: N 22 degrees 4.4232 minutes
Nationality: Singapore	Heading: 66 degrees
Ship length: 144 meters	Course direction: 137.0 degrees
Ship width: 23 meters	Status/speed: undefined/0.5 knots
Deadweight: 13064	Draft: 5.5 meters
GT: 9422	Updated: 2020-10-10 11:20:38
Net Tonnage: 13064	Ship Type: Cargo Ship

Acknowledgements

This work was supported by the Beijing Municipal Science and Technology Project No.Z191100001419009.

References

Achanta, R., Susstrunk, S., 2010. Saliency detection using maximum symmetric surround, 2653–2656.
Brown, M., Susstrunk, S., 2011. Multi-spectral sift for scene category recognition, 177–184.
Chen, Z., Wang, X., Liang, R., 2014. Rgb-nir multispectral camera. Optics Express 22 (5), 4985–4994.
Cheng, M., Mitra, N., Huang, X., Torr, P., Hu, S., 2015. Global contrast based salient region detection. IEEE Transactions on Pattern Analysis & Machine Intelligence 37 (3), 569–582.
DAI, Y., ZHANG, J., HE, M., PORIKLI, F., LIU, B., 2019. Salient object detection from multi-spectral remote sensing images with deep residual network. Acta Geodaetica et Cartographica Sinica 002 (002), 101–110.
Farbman, Z., Fattal, R., Lischinski, D., Szeliski, R., 2008. Edge-preserving decompositions for multi-scale tone and detail manipulation, 27 (3), 67.
Han, J., Zhang, D., Cheng, G., Guo, L., Ren, J., 2015. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. IEEE Transactions on Geoscience and Remote Sensing 53 (6), 3325–3337.
Han, J., Zhou, P., Zhang, D., Cheng, G., Guo, L., Liu, Z., Bu, S., Wu, J., 2014. Efficient, simultaneous detection of multi-class geospatial targets based on visual saliency modeling and discriminative learning of sparse coding. Isprs Journal of Photogrammetry and Remote Sensing 89, 37–48.
Harel, J., Koch, C., Perona, P., 2006. Graph-based visual saliency, pp. 545–552.
Hou, X., Zhang, L., 2007. Saliency detection: A spectral residual approach, pp. 1–8.
Last, P., Bahlke, C., Hering-Bertram, M., Linsen, L., 2014. Comprehensive analysis of automatic identification system (ais) data in regard to vessel movement prediction. Journal of Navigation 67 (5), 791–809.
Lezama, J., Qiu, Q., Sapiro, G., 2017. Not afraid of the dark: Nir-vis face recognition via cross-spectral hallucination and low-rank embedding, pp. 6807–6816.
Nie, T., Han, X., He, B., Li, X., Liu, H., Bi, G., 2020. Ship detection in panchromatic optical remote sensing images based on visual saliency and multi-dimensional feature description. Remote Sensing 12 (1), 152.

Oh, J., Lee, C., 2015. Automated bias-compensation of rational polynomial coefficients of high resolution satellite imagery based on topographic maps. Isprs Journal of Photogrammetry & Remote Sensing 100, 14–22.(feb.)
Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9 (1), 62–66.
Pan, S., Chen, A., Zhang, P., 2013. Securitas: user identification through rgb-nir camera pair on mobile devices, pp. 99–104.
Proia, N., Page, V., 2010. Characterization of a bayesian ship detection method in optical satellite images. IEEE Geoscience and Remote Sensing Letters 7 (2), 226–230.
Qi, S., Ma, J., Lin, J., Li, Y., Tian, J., 2015. Unsupervised ship detection based on saliency and s-hog descriptor from optical satellite images. IEEE Geoscience and Remote Sensing Letters 12 (7), 1451–1455.
Song, M., Hongsong, Q.U., Zhang, G., Jin, G., 2018. Detection of small ship targets from an optical remote sensing image. Frontiers of Optoelectronics 11 (3), 275–284.
Sun, H., Sun, X., Wang, H., Li, Y., Li, X., 2012. Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model. IEEE Geoscience and Remote Sensing Letters 9 (1), 109–113.
Wan, X., Liu, J., Li, S., Dawson, J., Yan, H., 2019. An illumination-invariant change detection method based on disparity saliency map for multitemporal optical remotely sensed images. IEEE Transactions on Geoence and Remote Sensing 57 (3), 1311–1324.
Wang, Q., Yan, P., Yuan, Y., Li, X., 2013a. Multi-spectral saliency detection. Pattern Recognition Letters 34 (1), 34–41.
Wang, Q., Zhu, G., Yuan, Y., 2013b. Multi-spectral dataset and its application in saliency detection. Computer Vision and Image Understanding 117 (12), 1748–1754.
Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.H., 2013. Saliency detection via graph-based manifold ranking. 2013 IEEE Conference on Computer Vision and Pattern Recognition.
Yao, Libo, Xiong, Wei, Zhou, Zhimin, Liu, Yong, 2019. Gf-4 satellite and automatic identification system data fusion for ship tracking. IEEE Geoscience and Remote Sensing Letters.
Zhang, L., Zhang, Y., 2017. Airport detection and aircraft recognition based on two-layer saliency model in high spatial resolution remote-sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (4), 1511–1524.
Zhou, M., Jing, M., Liu, D., Xia, Z., Zou, Z., Shi, Z., 2018. Multi-resolution networks for ship detection in infrared remote sensing images. Infrared Physics & Technology 92, 183–189.