



Article

Multi-Source Remote Sensing Image Fusion for Ship Target Detection and Recognition

Jinming Liu ¹, Hao Chen ^{2,*} and Yu Wang ²¹ Defense Engineering Institute Academy of Military Sciences, Beijing 100036, China; 17B905031@stu.hit.edu.cn² School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150006, China; 20b905065@stu.hit.edu.cn

* Correspondence: hit_hao@hit.edu.cn

Abstract: The active recognition of interesting targets has been a vital issue for remote sensing. In this paper, a novel multi-source fusion method for ship target detection and recognition is proposed. By introducing synthetic aperture radar (SAR) sensor images, the proposed method solves the problem of precision degradation in optical remote sensing image target detection and recognition caused by the limit of illumination and weather conditions. The proposed method obtains port slice images containing ship targets by fusing optical data with SAR data. On this basis, spectral residual saliency and region growth method are used to detect ship targets in optical image, while SAR data are introduced to improve the accuracy of ship detection based on joint shape analysis and multi-feature classification. Finally, feature point matching, contour extraction and brightness saliency are used to detect the ship parts, and the ship target types are identified according to the voting results of part information. The proposed ship detection method obtained 91.43% recognition accuracy. The results showed that this paper provides an effective and efficient ship target detection and recognition method based on multi-source remote sensing images fusion.

Keywords: multi-source remote sensing image; target detection and recognition; saliency; detection result voting



Citation: Liu, J.; Chen, H.; Wang, Y. Multi-Source Remote Sensing Image Fusion for Ship Target Detection and Recognition. *Remote Sens.* **2021**, *13*, 4852. <https://doi.org/10.3390/rs13234852>

Academic Editors: Chunhui Zhao, Xiuping Jia, Wei Li, Shou Feng, Nan Su and Yiming Yan

Received: 19 October 2021
Accepted: 24 November 2021
Published: 29 November 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Accurate target detection and recognition is of great scientific and practical significance in urban planning, air traffic control and traffic navigation, and has always been a hot research topic in the field of remote sensing. However, it is a challenge to accurately detect and identify targets from high-resolution remote sensing images in a timely manner. With the rapid development and innovation of sensor technology, wireless communication technology, aerospace technology and other related disciplines, a large number of optical remote sensing satellites and synthetic aperture radar (SAR) satellites have been successfully launched and operated worldwide [1,2]. The advancement of high temporal and spatial resolution data and multi-source data fusion technology has provided unprecedented opportunities for remote sensing information application fields.

Optical data and SAR data are the two most common data types in the field of satellite remote sensing; the two sensors have different advantages in Earth observation due to the different imaging mechanisms. Optical remote sensing images can intuitively reflect texture, color, shape and other information to users, but the ability of data acquisition is limited due to the limitation of light and weather [3]. The SAR sensor is capable of all-day and all-weather detection, which can penetrate clouds and fog and is not affected by shadow occlusion and light time. The SAR image can obtain completely different image information from the optical image by using an EM spectrum of different range, which can supplement the optical image information; however, it is difficult to interpret the obtained SAR data due to insufficient information for texture and ground target radiation [4–6]. SAR remote sensing images have advantages in scattering components

and parameters, while optical remote sensing images can extract rich spectral information in radiometric properties [7]. With the emergence of a large amount of remote sensing data, mining multi-source information from massive high-resolution remote sensing images for target interpretation has become an important part of the remote sensing information application field.

With the continuous development of remote sensing satellite technology, satellite imaging has developed from the earliest tens of meters ground image resolution to the sub-meter resolution now, and the remote sensing image target detection and recognition technology has also gradually emerged. Due to the difference of optical and SAR imaging mechanism, the methods used for target detection and recognition are different. In the optical remote sensing image target detection, the authors of [8] built a new high-resolution ship detection data set, where 2499 images and 9269 instances were collected from Google Earth with different resolutions. Han et al. [9] classified typical targets such as planes, ships and oil depots in large area remote sensing images according to the shape, and extracted saliency geometric features for target detection and recognition. Zhu et al. [10] proposed a method to extract and select new combined invariants for training classifiers when identifying aircraft types, which was conducive to controlling the stability of image invariants at all stages. Fang et al. [11] removed the interference area by the contour tracking method and left the target floating area in the image, normalized the moment invariants of the aircraft by extracting samples and then trained the BP neural network to identify suspicious areas in remote sensing images. In terms of target detection based on SAR data, the authors of [12] provided a huge ship detection data set, which, labeled by SAR experts, was created using 102 Chinese Gaofen-3 images and 108 Sentinel-1 images. Li et al. [13] proposed a superpixel constant false alarm rate detection method in high-resolution SAR images. The statistical properties of superpixels were described by weighted information entropy to distinguish between targets and clutter superpixels, and a two-stage constant false alarm rate detection scheme was proposed to detect the superpixels of target, including global detection and local detection. Han et al. [14] proposed an aircraft detection algorithm based on feature fusion, which took into account the dihedral characteristics of the aircraft tail and the strong scattering intensity of the fuselage; the above two polarization characteristics were combined to construct the aircraft target detection features, and then the constant false alarm rate was used for target detection. In view of the problems encountered in direct end-to-end feature learning for object detection and the close relationship between objects and auxiliary cues, the authors of [15] proposed a multitask learning-based object detector to distinguish ships in SAR images. Compared with traditional single-task-based object detectors, more discriminative object-specific features are learned by multitask learning without the extra cost of manual labeling.

Generally, image fusion technology is mainly divided into three fusion levels: pixel-level fusion, feature-level fusion and decision-level fusion [16,17]. Pixel-level image fusion generally processes the acquired raw pixel data directly, which can provide comprehensive and rich detailed information of the target scene. Li et al. [18] proposed an image fusion method based on wavelet transform, which reconstructed an information-rich image by decomposing the image into sub-images with different frequencies and then fusing them with different fusion rules to improve the image quality and meet the needs of visual applications. Feature-level image fusion performs the fusion process by preprocessing and extracting features from the edge contour of the image, which compresses the image information without losing important information and enhances the recognition ability of the image. You et al. [19] proposed a visual saliency detection model based on adaptive fusion of color and texture in 2017. Based on image preprocessing, the color saliency map and texture saliency map are extracted, respectively, in this method. Then, according to the texture complexity of each image, a simple and excellent fusion mechanism is used to fuse these images adaptively. Decision-level image fusion extracts and classifies information from multi-source image data and makes optimal decisions according to certain fusion rules and the confidence level of each decision. Zhao et al. [20] proposed a

robust recognition method based on decision-level fusion of infrared and visible images, which obtained a good recognition effect by combining the linear weighted sum and maximum matching score.

In recent years, there have been more and more research studies based on SAR and optical fusion. Han et al. [21] adopted Mallat algorithm and À trous algorithm to achieve SAR and multi-spectral image fusion based on wavelet transform, respectively, which preserves the spectral information of multi-spectral images and reduces distortion, and also highlights the texture information of SAR images, so that the fused images show more spatial detail. Byun et al. [22] proposed a region-based optical and SAR image fusion algorithm, which ultimately fused into a multi-spectral image by setting different fusion rules for each segmented region. Spröhnle et al. [23] proposed an object-based analysis and fusion of optical and SAR satellite data algorithm for dwelling detection in refugee camps. The authors of [24] proposed an algorithm to transfer knowledge from optical domains to SAR domains to eliminate the need for huge labeled data points in the SAR domains.

The present ship detection methods can be extensively utilized for optical, hyperspectral and SAR images [25]; however, the research on target detection and recognition based on multi-source remote sensing image fusion is still lacking. The existing ship detection and recognition based on multi-source fusion can locate the ship position, some methods can further identify the type of ship, but it is difficult to distinguish two types of ship that are relatively similar. For example, in this paper, destroyers and cruisers were similar in shape, size, color and texture, which made the recognition task more difficult.

The proposed method includes port detection, ship detection and ship recognition. In the paper, for the deficiency of single source information, saliency dock detection based on multi-source data level fusion is proposed to obtain the dock slice containing the ship target. Based on the fusion of SAR image features and optical image features, a ship target detection method based on joint shape analysis and multi-feature classification is proposed, and the ship target is detected successfully. Finally, a multi-part detection method is proposed and the target type identification is realized according to the voting result of position information. The overview of the proposed method is shown in Figure 1.

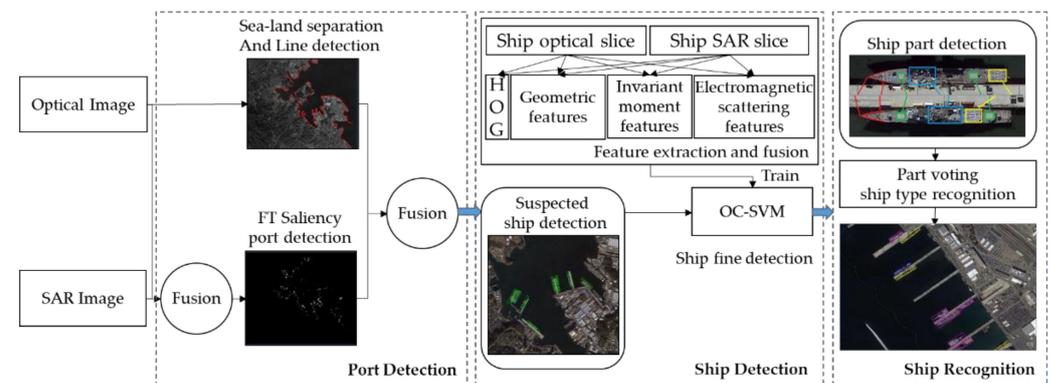


Figure 1. An overview of the proposed method.

The remainder of this paper is organized as follows. Section 2 introduces the methodology of the proposed method. Section 3 is a comparative analysis of experimental results. Discussions and conclusions are presented in Sections 4 and 5, respectively.

2. Methodology

Ship targets may generally appear in two types of areas in remote sensing images; one is on the sea surface and the other is in the port [26]. The proposed detection and recognition method is mainly designed for ship targets in port, but it is also applicable to the stationary ship target detection on the sea surface. First, a large number of suspected port slice images that may contain ships are obtained by port area detection, so that the range of the target detection area is reduced to improve the speed and accuracy of target

recognition. Then, the SAR images are used to screen suspected port slice images to further narrow down the range of the target detection area. Next, the joint shape analysis is used to detect the suspected ship target in the port slice image, and the feature vector is constructed by using the previously extracted multi-source and multi-feature to achieve the ship detection through the one-class support vector machine (OC-SVM) classification. Finally, feature point matching, contour extraction and brightness saliency are used to detect ship parts, and voting is performed according to the part detection results and the ship target type is identified.

2.1. Port Area Detection Based on Multi-Source Remote Sensing Image

Since the ships anchored near the shore are close to the straight coastline, it is considered that the area with linear features is the suspected area containing ships, and a large number of background areas can be excluded. Therefore, this paper conducts the detection of the port area first (suspected ship docking area). Firstly, the sea-land boundary zone is extracted by using the different reflectance of land and sea to determine the approximate position of the ship. Sea-land separation is made by randomly selecting sea surface seed points and then growing regionally. In the optical remote sensing image, land and sea have different reflectance, and the brightness of the sea area is low while that of the land area is high, so the sea-land separation is processed in the optical image.

Meanwhile, the docking direction of nearshore ships is the same as that of ports, so that the position and direction of ports need to be determined first. The linear characteristics of the port are obvious, so that the port position direction can be easily determined by the direction of the port line position. The line segment detector (LSD) is used for line detection, and the aforementioned sea-land separation is used to determine the effective area of the line segment detection, and the ship sensitive area is selected by Equation (1).

$$Region = \{A \oplus B - A \ominus B\} \quad (1)$$

where *Region* represents the sensitive region containing the ship, \oplus is the dilating operation, \ominus is the eroding operation and *A* is the sea-land separation region. *B* is the square structural element with side length of 200 pixels, which is determined by the length of the ship.

The sensitive area is determined by using the difference between ocean expansion and corrosion images, which can be considered as the transition area between ocean and land. The detected line segments mainly include port boundary, ship boundary and land boundary, and the ship or port neighborhood image can be obtained by extracting the neighborhood image of each retained straight line segment. Then, the ship docking direction can be determined by the angle of the straight line segment, and the ship can be detected in each ship/port neighborhood image successively. The line detection results are shown in Figure 2.

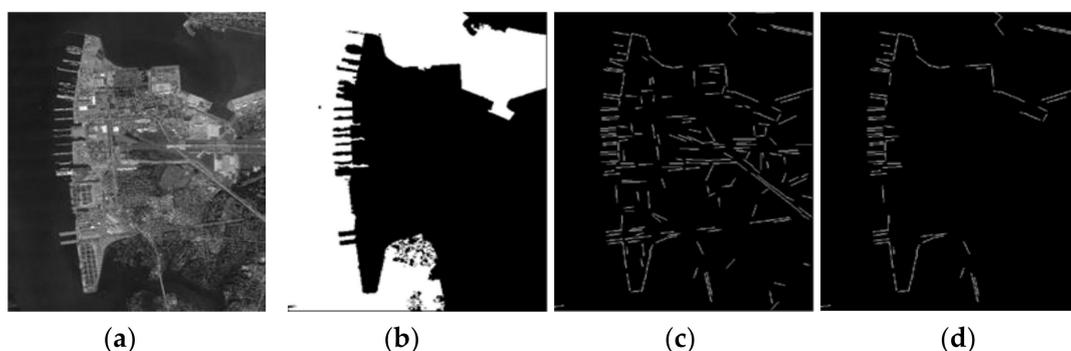


Figure 2. The results of line detection and retention. (a) Experimental image. (b) Sea-land separation. (c) Line detection. (d) Line retention.

Through this method, the optical image of Yokosuka Port is tested, and the results are shown in Figure 3.



Figure 3. The results of line detection and retention in Yokosuka Port.

A large number of slice images of suspected port areas are obtained by using line segment detection results (Figure 4a). However, due to the difficulty in defining the range of port areas, SAR data (Figure 4b) is introduced to constrain the detection results of optical image ports.

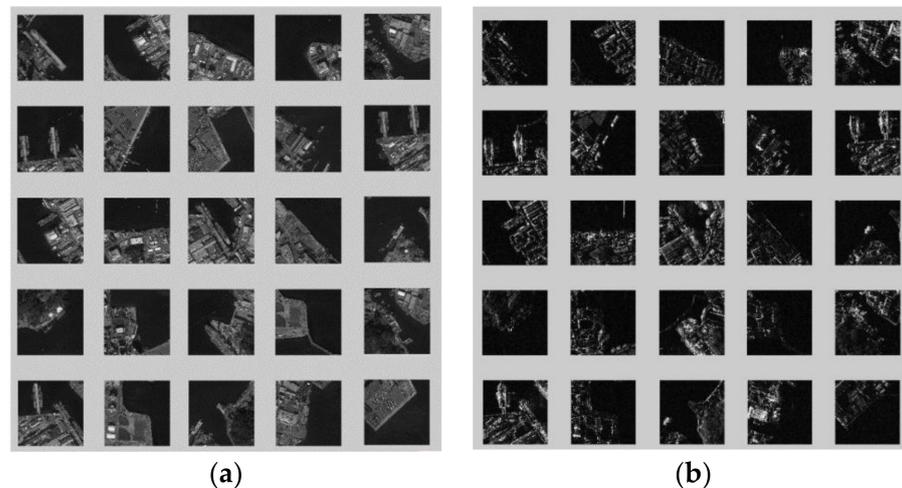


Figure 4. Slice images of suspected port area in Yokosuka. (a) Optical slice images of suspected port area in Yokosuka; (b) SAR slice images of suspected port area in Yokosuka.

Visual saliency estimation is one of the pre-attentive procedures for humans to focus their eyes on regions with attractive contents from scenes [27]; so, this technique is used to highlight valuable targets while suppressing backgrounds. In the paper, the pixel-level fusion of single-band SAR image and multi-band optical image based on HSV color space is carried out. The ship targets in the port area are highlighted by using the frequency-tuned (FT) saliency map [28]. FT saliency image is obtained by using the center-periphery operator of color features, as shown in Figure 5.

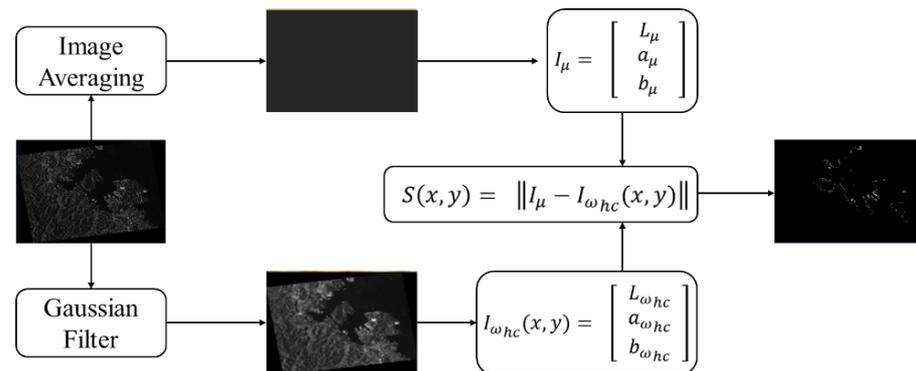


Figure 5. Saliency feature map calculation.

In the paper, firstly, the fusion remote sensing image is processed by Gaussian filtering, and the obtained Gaussian filtering image is converted from RGB color space to LAB color space. Then, the images of L, A and B channels in LAB color space are averaged. The saliency value $S(x, y)$ is obtained by calculating the Euclidean distance of the mean image and the filtered image of the three channels and summing, the maximum value is used to normalize the saliency image and the final saliency feature image is obtained. The detection results of ship-suspected area are shown in Figure 6.

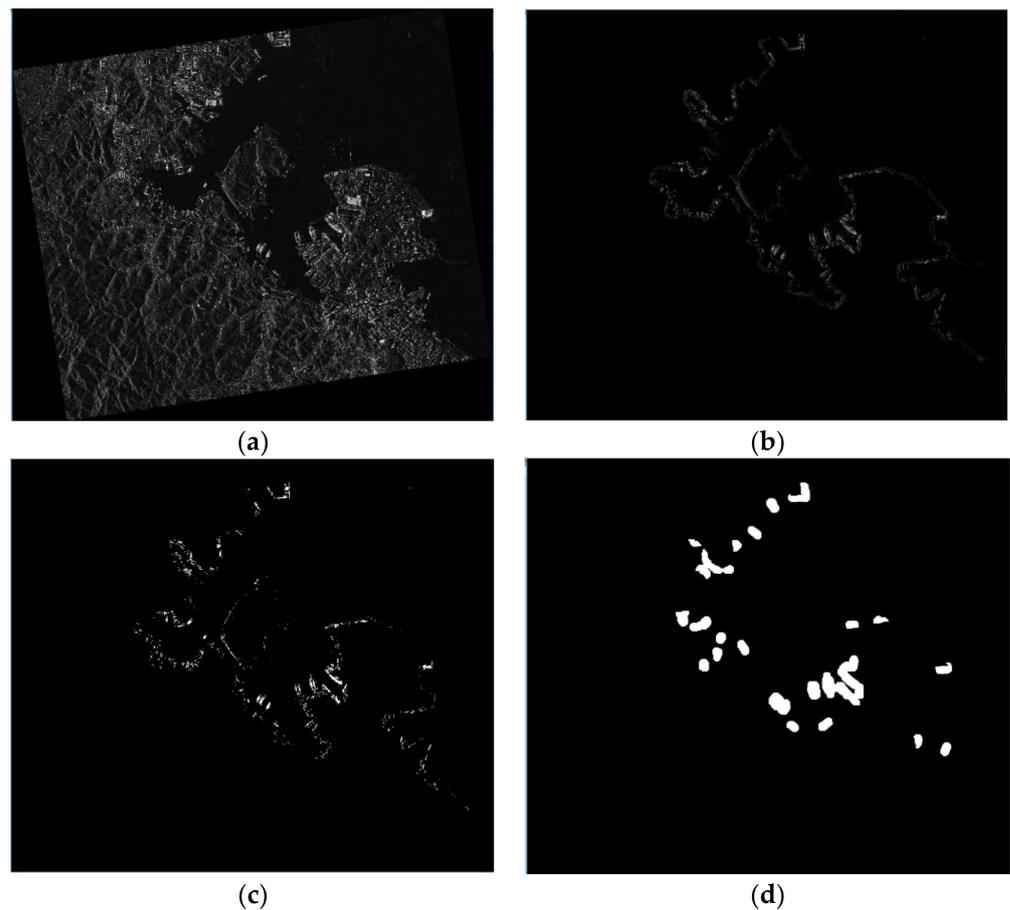


Figure 6. Saliency feature map calculation. (a) SAR image of the port. (b) FT saliency map of the land-sea boundary zone. (c) Binarization of saliency map. (d) Results of suspected ship area after morphological processing.

2.2. Ship Target Detection Based on Multi-Source Remote Sensing Image

The angle of line on each port slice image has been obtained in the previous port detection; it can be considered that the angle of line is consistent with the direction angle of the ship, and the ship should be rotated to a horizontal direction according to the angle. Therefore, the problem of ship target detection of large and wide remote sensing images with different angles is transformed into detection of horizontally parked ships on port slices. The port slice images in the horizontal direction are shown in Figure 7.

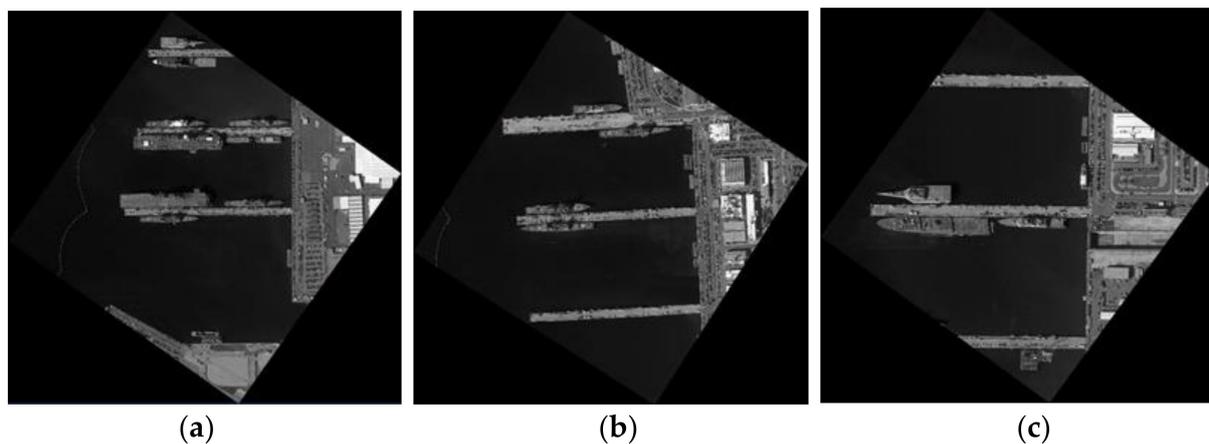


Figure 7. The port slice images in the horizontal direction. (a) port slice images of Yokosuka port in the horizontal direction. (b) port slice images of Santiago port in the horizontal direction. (c) port slice images of part of Santiago port in the horizontal direction.

The spectral residual saliency map and the region growth method are usually used to detect ship targets in typical optical remote sensing images. Since the spectral residual saliency map has a greater response to the region rich in high-frequency components such as ships, ships can be obtained by locating saliency points of ships and region growth. However, this method is mainly used to detect merchant ships rather than ships, and the ship target images will produce a number of shadows on the deck under lighting conditions due to the influence of the bridge and various weapons and equipment, which makes it a very poor method to obtain saliency points and regional growth by spectral residuals. Therefore, this paper introduces SAR data and proposes an optical and SAR image fusion ship detection method based on joint shape analysis and multi-feature classification, as shown in Figure 8.

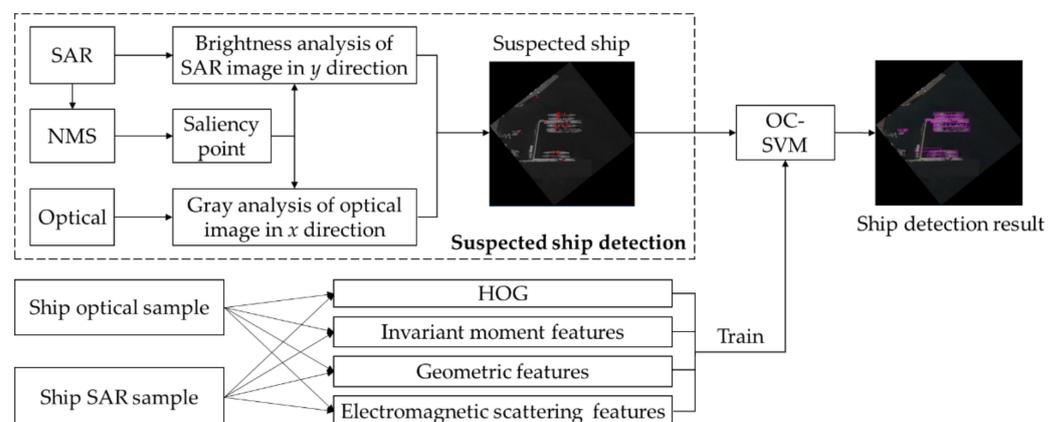


Figure 8. Ship detection model based on multi-feature fusion of SAR and optical image.

Firstly, saliency points are quickly determined by non-maximum suppression (NMS) of SAR images, which are used as seed points. The abscissa of the ship is obtained by X direction gray analysis of the optical image, and the ordinate of the ship is obtained by

Y direction brightness analysis of the SAR image, and the suspected ship coordinates are obtained. After that, the multi-source fusion features of the suspected ship target slice images are extracted, and the suspected targets are classified into ship targets and non-ship targets by a pre-trained support vector machine (SVM).

The specific contents are as follows:

- (1) Specifically, the bridge, turret and other parts of the ship target are made of metal and have rough surfaces, which have high backscattering characteristics and are shown as high brightness in SAR images (as shown in Figure 9b above). Compared with optical images (as shown in Figure 9a), SAR images can be used to locate ships better (as shown in Figure 9e). Therefore, non-maximum suppression is performed on SAR images directly, and the obtained local maximum is used to represent the metal objects with rough surfaces, such as the bridge and turret of the ship, so the significance points located inside the ship are obtained (as shown in Figure 9f).
- (2) The saliency points located inside the ships are used to determine the minimum bounding rectangle of each ship. Since the previous berthing area detection has rotated the regional image to the level of the ship target, it only needs to determine the length and width of the ship.

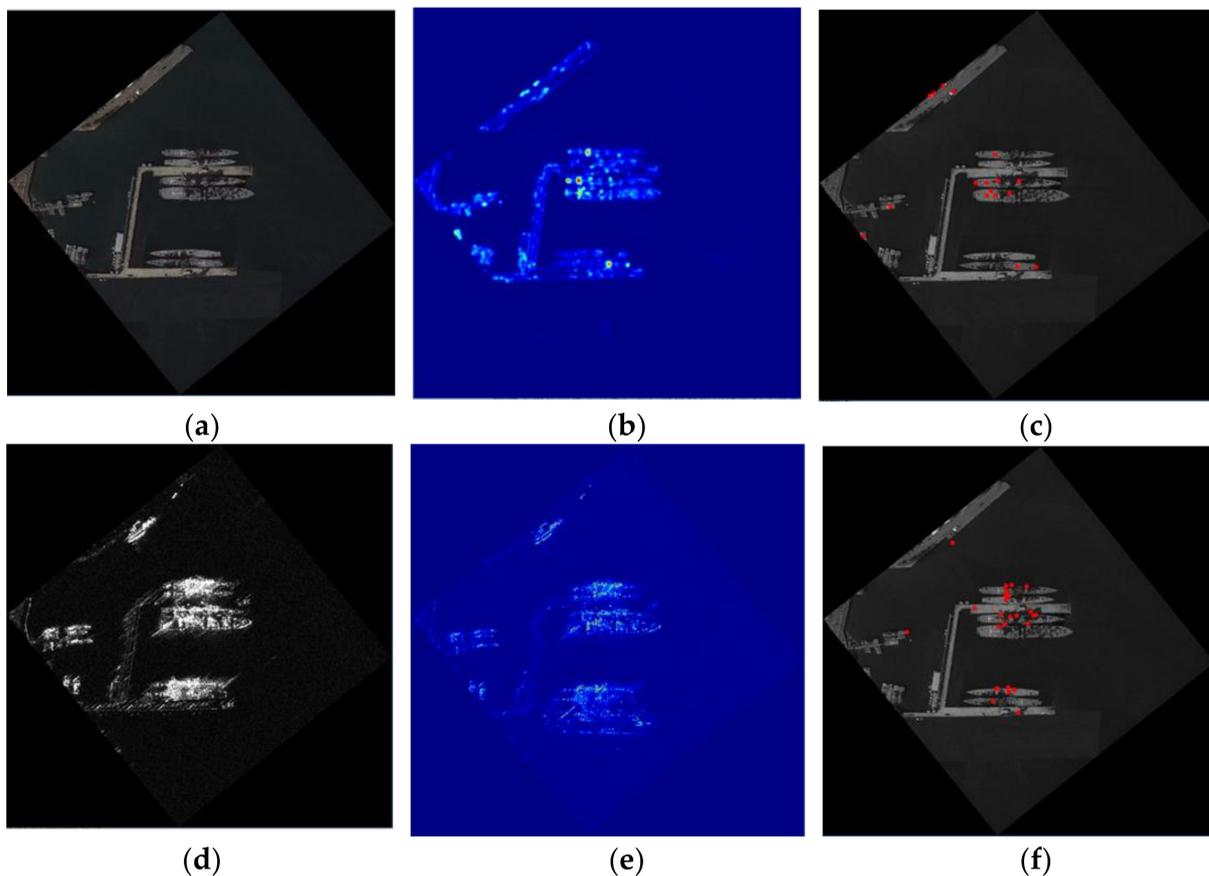


Figure 9. Obtaining results of saliency points of various ships. (a) Optical image of port area. (b) Optical image saliency map. (c) Optical image location result. (d) SAR image of port area. (e) SAR image saliency map. (f) SAR image location result.

The ship in the optical image has shadow interference in the center area of the ship, but the bow and stern areas are relatively flat and easy to identify. Therefore, the intersection point of the bow, stern and sea water can be found through gray analysis of the X direction (horizontal direction) of each saliency point. The bow and stern abscissa can be obtained by analyzing the X direction of the binary image. Due to the shadow in the center area of the ship in the optical image, the ship width obtained from the optical image is highly unstable, but the SAR image does not have this problem. The brightness curves were

obtained by the brightness values of saliency points in SAR images. It was found that there were obvious brightness changes in the gap between two ships docking closely side by side. After calculating the average brightness of sea surface, the boundary between the ship and the sea water in the Y direction (vertical direction) could also be obtained, so as to obtain the ordinate of the ship. A large number of suspected ship targets are obtained, as shown in Figure 10.

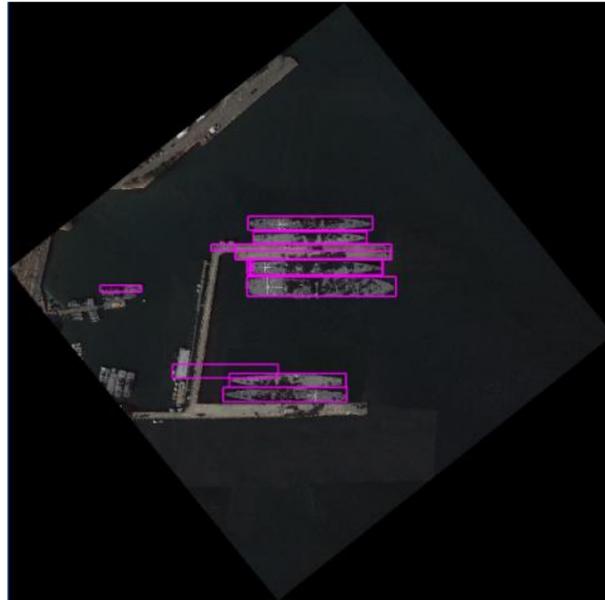


Figure 10. Suspected ship target.

Then, the corresponding optical and SAR target slice images are extracted according to the above positioning results. Since the center of some suspected ship targets is offset from the real ship to a certain extent, in this paper, a sliding window with a step length of 10 pixels is carried out around the center point of each connected domain of regional growth results to ensure the integrity of ship target slice images extraction.

For these optical and SAR image slice images, the geometric features, invariant moment features, histogram of oriented gradient (HOG) features of the optical and SAR slice images and scattering features of the SAR slice images are extracted, respectively, and a multi-feature fusion vector is constructed. Through the trained multi-feature fusion classification model, the false target can be eliminated, and the false alarm rate of ship detection can be reduced.

2.3. Ship Target Recognition Based on Multi-Source Remote Sensing Images

Ship models with different types and colors need to be further distinguished in ship target detection. In this paper, a variety of detection methods are used to detect different parts of the ship based on optical and SAR images slice images. Finally, the ship type is identified by combining the detection results of flight deck, prow contour, vertical launch system (VLS) and bridge. The part analysis diagram of different types of ships is shown in Figure 11.

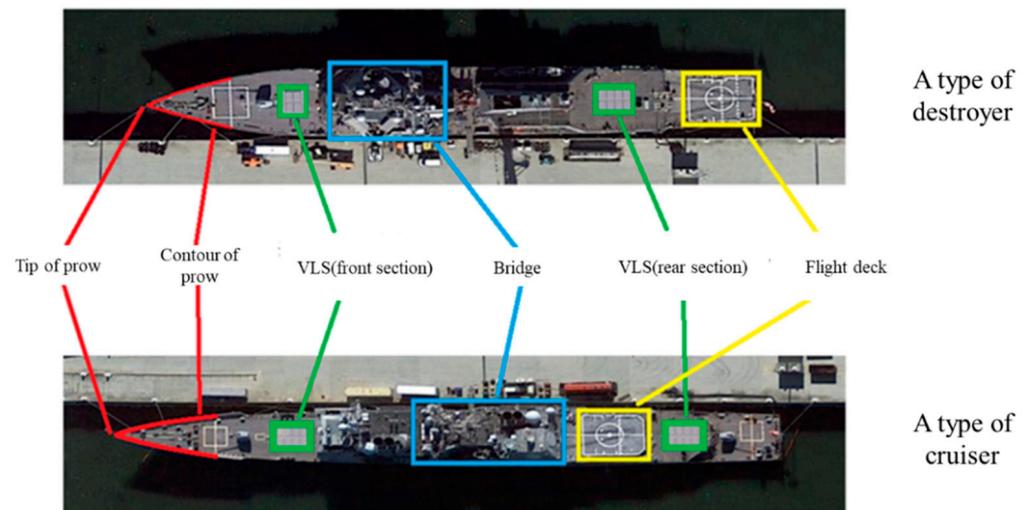


Figure 11. Part analysis of different types of ships.

- (1) Ship part (flight deck) detection based on feature point matching.

Considering the unique shape of the flight deck, scale-invariant feature transform (SIFT) [29] is used to extract feature points. SIFT feature points are extracted from the flight deck slice images, as shown in Figure 12.

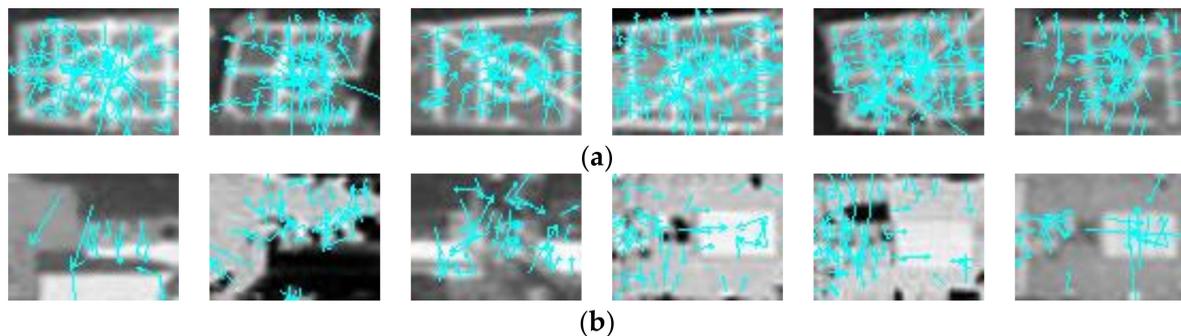


Figure 12. SIFT feature points extraction results of ship parts. (a) SIFT feature points extraction results of ship flight deck; (b) SIFT feature points extraction results of other parts.

Then, feature vectors of each positive and negative training sample are constructed based on bag of words (BOW) to describe the flight deck, which are then input into the SVM classifier for training.

- (2) Ship part (prow) detection based on contour extraction.

In order to further distinguish different types of ships in the same country, in this paper, a ship part detection method based on contour extraction is proposed to detect the prow contour radian and the position of ship VLS. In high-resolution optical satellite images, notable ship heads are usually important for ship detection [30]. It is found that the prow contour of different types of ships has obvious differences in shape and angle. Considering that the prow edge contour is not a regular triangle but a shuttle shape, which cannot be directly represented by numerical value, this paper adopts a convolution filter to identify the type of prow.

There are a large number of complex buildings on the ship target deck (as shown in Figure 13a). In this paper, the ship edge image is extracted by canny operator (as shown in Figure 13b), and then the binary image of the ship target is obtained by simple morphological processing (expansion, corrosion and hole filling) (as shown in Figure 13c). The outer contour of the ship target is further obtained (as shown in Figure 13d). The real

pre-trained prow contour (as shown in Figure 14a) is used as a convolution operator to conduct convolution filtering on the extracted ship overall contour and the prow contour response diagrams are formed (as shown in Figure 14b).

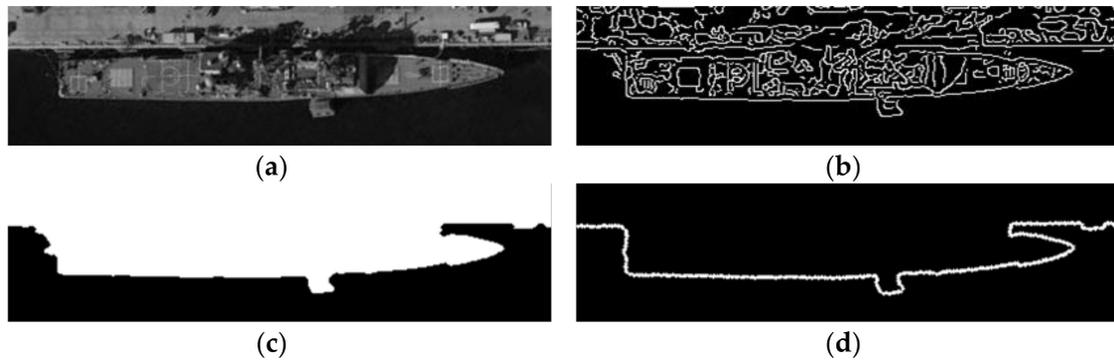


Figure 13. Extraction process of ship target edge contour. (a) Gray scale image of ship, (b) edge image of ship, (c) binary image of ship and (d) edge contour of ship.



Figure 14. Prow template and prow contour response diagram. (a) Prow model of destroyers and cruisers; (b) response diagram of cruiser prow contour.

(3) Ship part (bridge) detection based on brightness saliency.

In the image, pixels with less brightness distribution and higher brightness have higher saliency value. Therefore, this paper proposes a saliency algorithm based on brightness saliency image. The proposed saliency map has a good extraction effect for the targets with high brightness in SAR images.

I is the brightness map of the input remote sensing image with a size of $M \times N$. For the $\forall I_{ij} \in I$, the value of $BBSM_{ij}$ at any point in the brightness saliency map is shown in Equation (2).

$$BBSM_{ij} = \frac{I_{ij}}{M \times N} \sum_{m=1}^M \sum_{n=1}^N D(I_{ij}, I_{mn}) \quad (2)$$

$D(I_{ij}, I_{mn})$ is the absolute difference between I_{ij} and I_{mn} , as shown in Equation (3).

$$D(I_{ij}, I_{mn}) = |I_{ij} - I_{mn}| \quad (3)$$

The obtained $BBSM_{ij}$ constitutes the brightness saliency map and then the binary image of the saliency target is obtained by simple threshold segmentation. The center coordinates of the connected domains are obtained by screening the connected domains that meet the geometric conditions. Finally, the center coordinates of the brightness saliency targets are obtained, which is namely the center position of the bridge.

(4) Ship target model recognition based on position detection results voting.

In this paper, optical image slice image is used to obtain the length information of the ship, the type of flight deck, the position of prow tip, the type of prow contour and the position of VLS, and SAR image slice image is used to obtain the width information of the ship and the position of the bridge. Based on this, a ship recognition method based on part detection result voting is proposed. According to the detection results of the above seven groups of parts, the detection results of each part are identified and voted on, and the ships

are divided into a certain type of destroyer, a certain type of cruiser and other ships. The class with the maximum cumulative value of voting results is taken as the ship recognition result to identify model recognition. The method is shown in Figure 15, and the voting criteria of the detection results of each part are shown in Table 1.

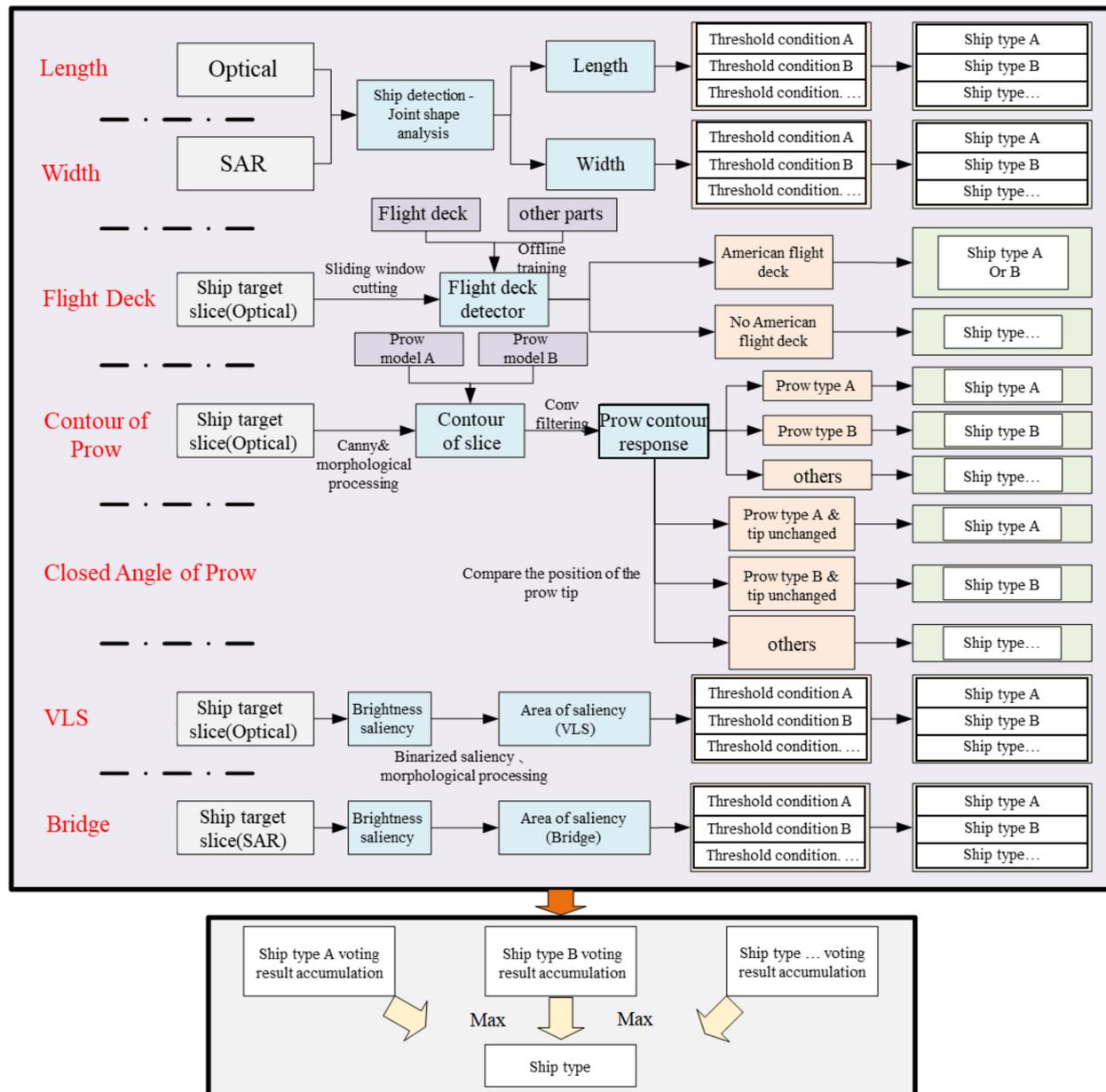


Figure 15. Schematic diagram of ship target type recognition method based on voting of part detection results, where ship types A and B represent cruiser and destroyer, respectively.

The voting rules are as follows:

For the length and width information, the actual length and width of the destroyer are 153.7 m and 20.4 m, respectively. The actual length and width of the cruiser are 172.8 m and 16.8 m, respectively, and the optical and SAR remote sensing image resolution is 0.5 m. Therefore, for the length of a ship, it is specified that a ship with a length of 285–325 pixels should be a destroyer, a ship with a length of 326–366 pixels should be a cruiser and other lengths should be other ships. For ship widths, it is specified that a ship with a width of 38–45 pixels should be a destroyer, a ship with a width of 30–37 pixels should be a cruiser and other widths should be other ships.

For the type of flight deck, ships that can detect the flight deck are considered to be American ships (destroyers or cruisers).

For the prow contour type, destroyers and cruisers are classified according to the detection results of the prow contour.

For the tip of the prow, the maximum point of the response of the contour, namely the center point of the prow, can be obtained by non-maximum suppression for the contour response point, so that the coordinates of the tip of the prow are obtained. The ship can be identified by the consistency between the coordinates of the tip of the prow and the coordinates of the front of the ship.

For the bridge, the distance between the bridge and the prow is taken as the classification standard. The ship with a distance between the bridge and the prow between 115 and 160 pixels is considered to be a destroyer, the ship with a distance between the bridge and the prow between 161 and 210 pixels is considered to be a cruiser and other ships are considered to be other ships.

For VLS, the distance from the prow can be used as a classification criterion; ships with front VLS within 60–80 pixels of the prow are considered destroyers, and ships with front VLS within 81–100 pixels are considered cruisers. Ships with rear VLS within 205–265 pixels of the prow are considered destroyers, and ships with rear VLS within 266–325 pixels are considered cruisers. Since some ships are shaded, and only one VLS can be detected, either the front end or the back end can be used as the judgment standard.

Table 1. Voting criteria for ship parts detection results.

	Destroyer	Cruiser	Other Ships
Length	285–325 (pixel)	326–366 (pixel)	Other pixel lengths
Width	38–45 (pixel)	30–37 (pixel)	Other pixel lengths
Bridge	115–160 (pixel)	161–210 (pixel)	Other pixel lengths
VLS	Front: 60–80 (pixel) Backend: 205–265 (pixel)	Front: 81–100 (pixel) Back end: 266–325 (pixel)	Other pixel lengths
Prow contour	Destroyer type	Cruiser type	Other types
Prow tip	Destroyer type with the same tip	Cruiser type with the same tip	Other types
Flight deck	American flight deck	American flight deck	Other types

3. Experimental Results and Analysis

The data set was collected in October 2009 over the Yokosuka and Santiago ports. The original data collection contains optical and SAR remote sensing images, which contain $7865 \times 11,729$ pixels and $12,980 \times 14,988$ pixels, respectively. There are three classes, including destroyers, cruisers and other ships. To fuse optical data with SAR data, SAR data are downsampled to 0.5 m. Specific data parameters are shown in Table 2.

Table 2. Voting criteria for ship parts detection results.

	Optical	SAR
Data acquisition time	27 October 2009	28 October 2009
Resolution ratio	0.5 m	1.1 m
Sensor	Google Earth	TerraSAR-X
SAR polarimetric	-	HH
Wavelength	-	3.2 m
Orbit type	-	Sun-synchronous orbit

3.1. Port Area Detection Results Based on Optical and SAR Image Fusion

FT saliency highlights the saliency target in the image and ignores high-frequency interference caused by texture, noise, etc. To improve the algorithm speed, the optical image slice images are fused with SAR images, and the port slice images without any suspected ship target are eliminated through FT saliency map detection, so that the port area is

obtained. Figure 16 shows the different port detection results on the 1000×1000 slices of Yokosuka Base.

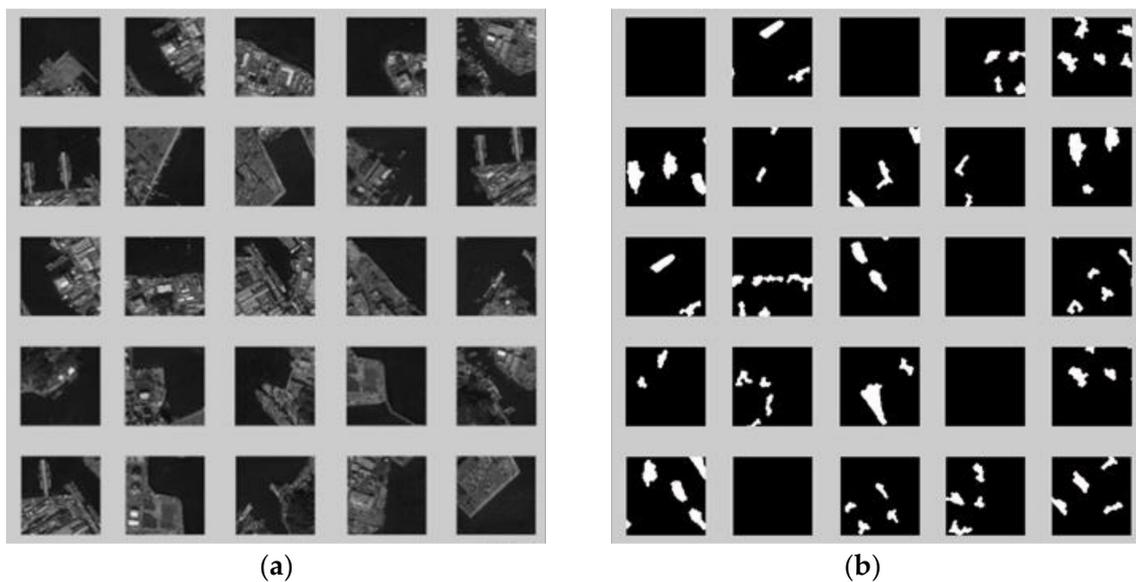


Figure 16. Detection results of port area fused with optical and SAR images. (a) Optical images of some ports. (b) Constraint results after SAR image fusion, where white area is the suspected area containing ships.

3.2. Ship Target Detection Results Based on Optical and SAR Remote Sensing Image Fusion

The proposed method was tested in Yokosuka Port and Santiago Port. Experimental data are optical and SAR remote sensing images of Yokosuka and Santiago ports, including destroyers, cruisers and other ships, all with a resolution of 0.5 m. Part of the ship sample is shown in Figure 17.

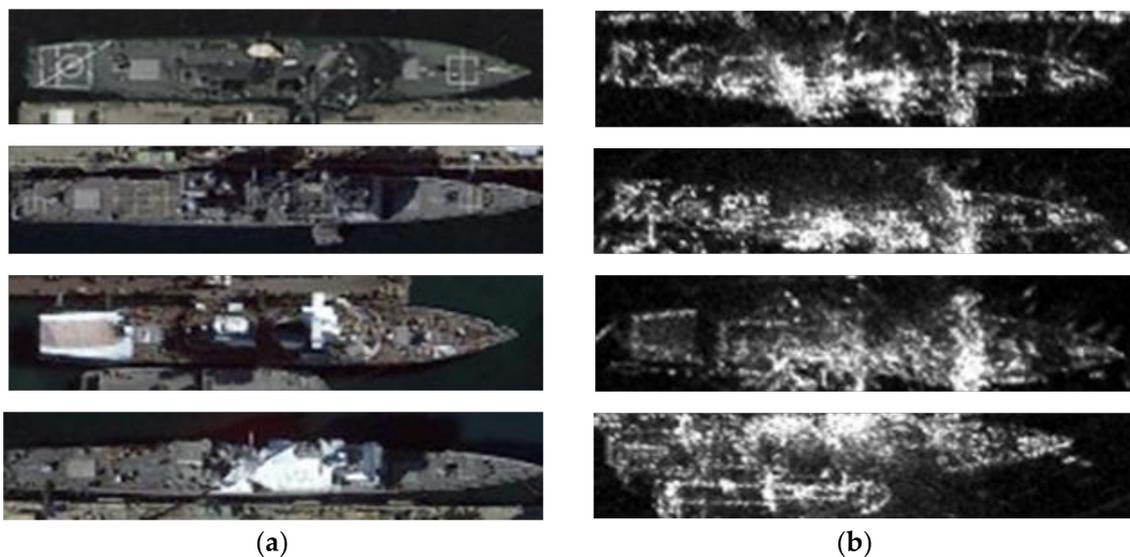


Figure 17. Optical and SAR images of ship samples. (a) Optical images of some ships. (b) SAR images of some ships.

The optical image in the experiment is obtained by the method introduced in the previous section, and the geometric features, invariant moment features and HOG features of the slice image are extracted to construct feature vectors, and the SVM classifier is used for classification. The experimental results are shown in Figure 18. The comparative results between the optical image ship detection and the proposed method are shown in Table 3.

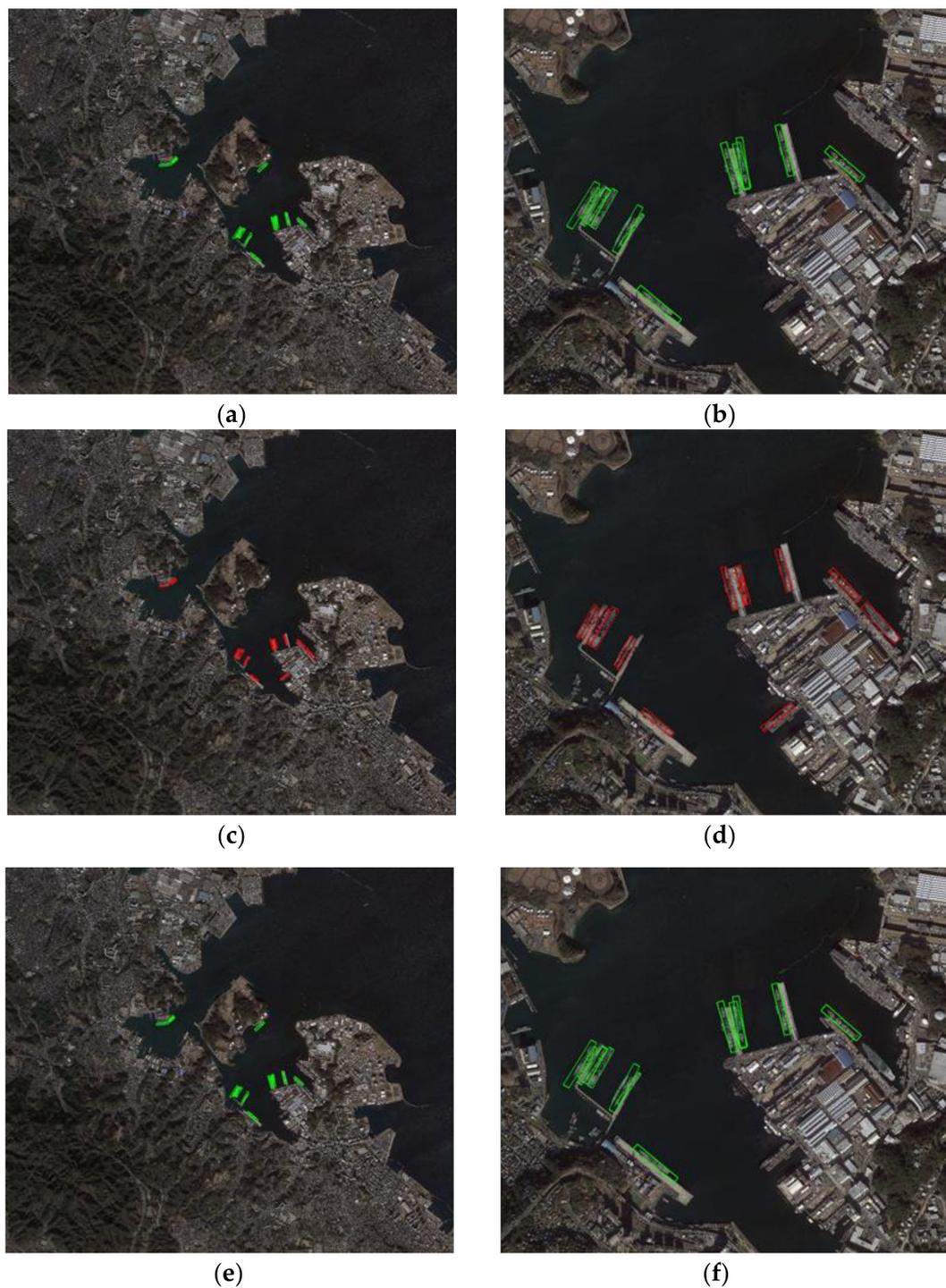


Figure 18. Ship target detection results of different methods. (a) Optical image detection results in Yokosuka Base, (b) partial enlarged drawing of (a), (c) detection results of fusion method (displayed on optical image), (d) partial enlarged drawing of (c), (e) detection results of fusion method (displayed on SAR image) and (f) partial enlarged drawing of (e).

Table 3. Comparison results of ship targets detection by different methods.

	Optical	Proposed Method	HSTM	AHSTM
Recall	85.29%	91.43%	88.23%	91.17%
Precision	87.88%	94.12%	91.45%	94.02%

It can be seen from the above test results, compared with the single sensor optical image target detection and heterogeneous support tensor machine (HSTM) and adaptive heterogeneous support tensor machine (AHSTM) [31], that each task performance of the proposed ship target detection method that is based on joint shape analysis and further characteristics of multi-source remote sensing image classification has been greatly improved, and the false detection of targets has been reduced.

3.3. Ship Target Recognition Results Based on Multi-Source Remote Sensing Images

(1) Ship part detection results based on feature point matching.

The ship slice image can be obtained through the obtained ship target detection result in which the prow is rotated to a uniform direction. Then, the small slice image is extracted through the sliding window on the ship section to construct the SIFT word bag feature [32], and the flight deck is detected through the trained SVM classifier. Figure 19 shows the detection results of flight decks of some ships. It can be seen that flight decks of American ships are accurately detected:

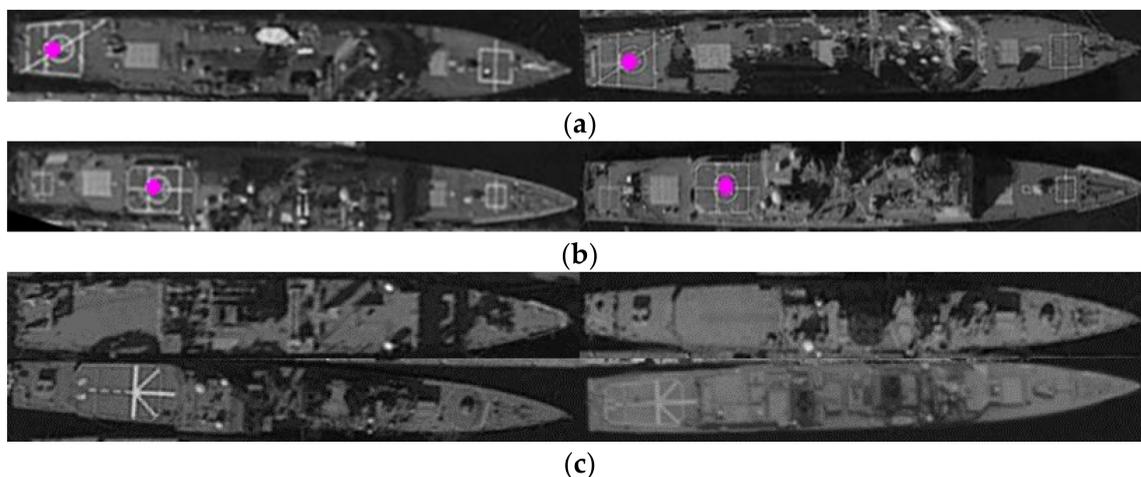


Figure 19. Flight deck recognition results of various types of ships. (a) Flight deck recognition results of destroyer. (b) Flight deck recognition results of cruiser. (c) Flight deck recognition results of other ships.

(2) Ship part detection results based on brightness saliency.

The brightness saliency is used to detect ship parts, and the experimental results of bridge detection on ship SAR image slice images are shown in Figure 20.

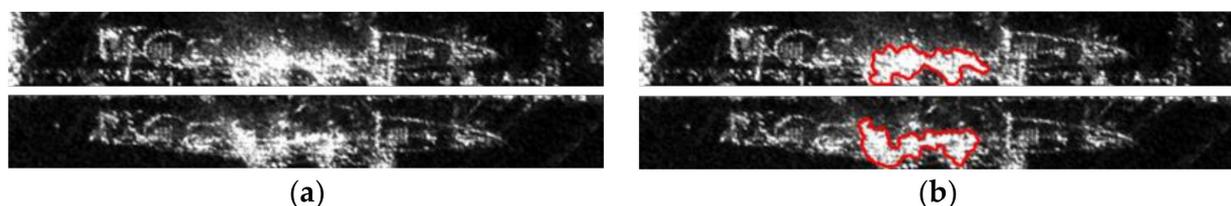


Figure 20. Detection results of ship bridge. (a) SAR images of ships. (b) Bridges detection results.

In the optical image, the brightness of the VLS is obviously different from that of the ship deck, and its shape is rectangular and the size is basically the same. Therefore, the VLS is identified by brightness saliency detection combined with geometric features. The experimental detection results of VLS on optical image slice images of ships using this method are shown in Figure 21.

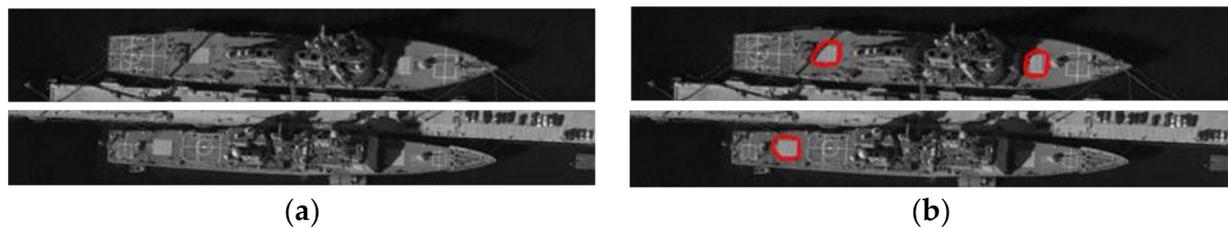


Figure 21. Detection results of ship VLS. (a) Gray scale of ships. (b) Detection results of ships VLS.

(3) Ship type recognition results based on voting results of part detection.

According to the criteria in Table 1, ships are classified through the detection results of each part. However, the detection error of single part will lead to the recognition error of ship type. Therefore, in the paper, the detection results of ship parts are divided into seven voting times, and the maximum voting result is taken as the result of ship type recognition. The recognition results are nested in the optical image and SAR image, respectively, which proves the feasibility of the proposed method. The recognition results of some ship types are shown in Figure 22.

Cruisers are marked in the yellow box, and destroyers are marked in the blue box and other types of ships are marked in the purple box. It can be seen from the above results that the proposed ship target model recognition method based on the voting results of part detection can make full use of the information provided by optical and SAR images for type recognition, and finally obtain excellent recognition results.

In Table 4, the proposed method is compared with the former algorithms in ship recognition, including HSTM [31], C-SVM [33], V-SVM [34] and C-STM [35]. These methods are mainly for different types of ships, and lack effective analysis for similar ships. It can be seen from the above results that the performance of the proposed method in the task is superior to other methods, as it uses multi-source fusion technology and voting algorithm to greatly improve the recognition precision.

Table 4. Comparison results of ship targets recognition by different methods.

Method	Proposed Method	HSTM	C-SVM	V-SVM	C-STM
Precision	91.43%	87.10%	74.19%	77.42%	74.19%

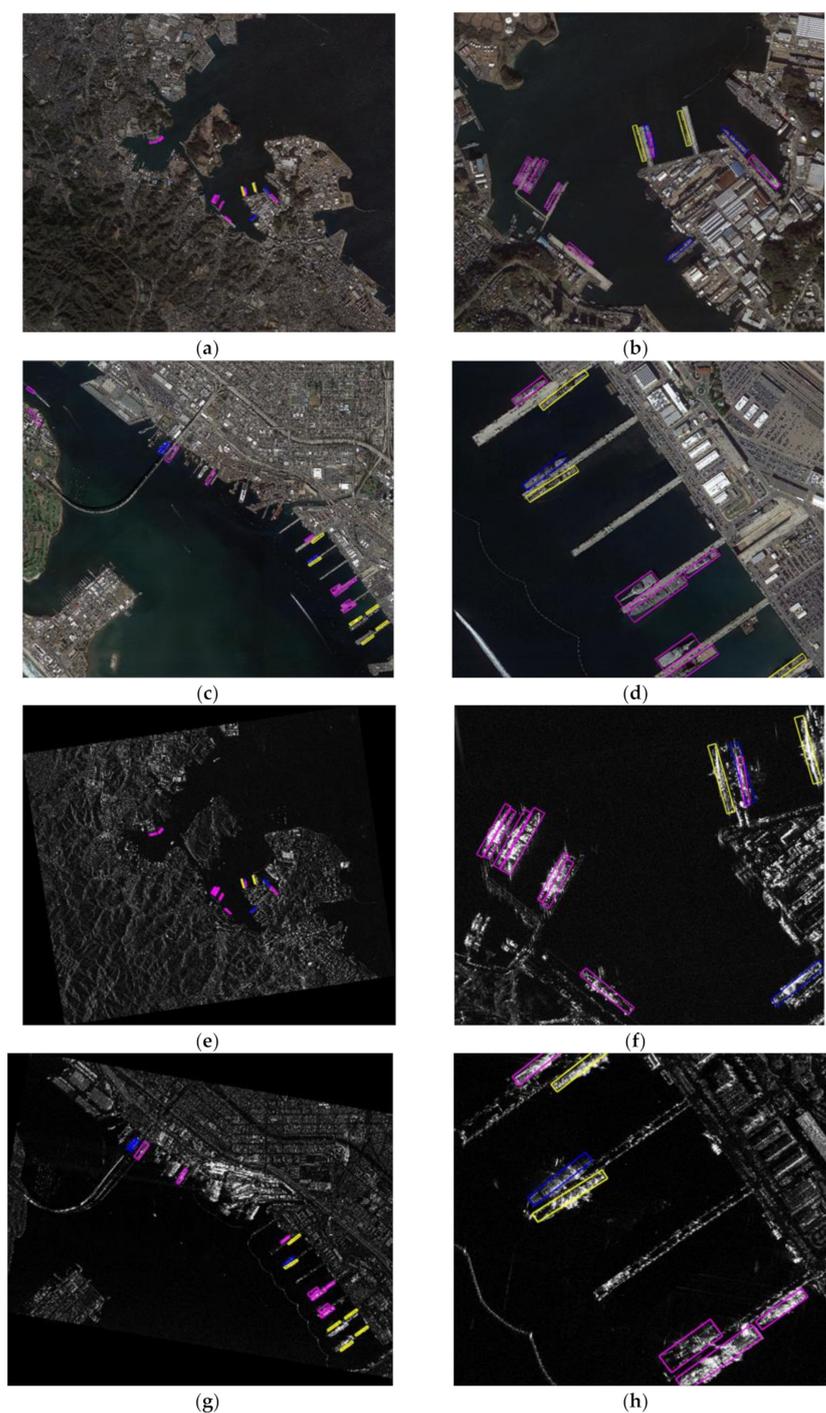


Figure 22. Ship type recognition results. (a) Ship type recognition results of Yokosuka Port on optical image, (b) partial enlarged drawing of (a,c) ship type recognition results in San Diego Port on optical image, (d) partial enlarged drawing of (c,e) ship type recognition results of Yokosuka Port on SARI image, (f) partial enlarged drawing of (e,g) ship type recognition results in San Diego Port on SARI image and (h) partial enlarged drawing of (g).

4. Discussion

The experimental results show that the proposed method obtained excellent target recognition effect by using multi-source remote sensing image fusion technology. A single optical image has the problems of fuzzy regional features and many candidate slice images in port, which is not conducive to ship target detection. Therefore, the port slice

classification method based on the saliency of fused SAR image is used to obtain the port slice containing the ship target, which can reduce the range of the target detection area and improve the speed and accuracy of target recognition. In addition, the complex port environment and the serious interference of the ship target shadow in the optical image lead to the problem of ship misdetection and missing detection. Therefore, by introducing SAR images to screen suspected dock slices, the range of the target area to be detected can be further reduced. It can be found that a single part detection cannot express the attributes of the target ship, so the accuracy of identification can be improved by using feature point matching, contour extraction and brightness saliency to identify the location detection of the ship, and voting occurs according to the part detection results. Only two types of typical targets are detected in this paper; to identify other types of ship targets and consider the multi-temporal impact of data will be the focus of further research.

The proposed method identified ship recognition through multi-task learning. Compared with other methods, in terms of fusion, this paper fully combines multi-task learning and multi-level fusion. In terms of fusion, SAR data are introduced for data level fusion to improve saliency detection results for large scale, while for small-scale ship detection, accurate description of ship features is obtained through feature level fusion. This fusion method for different learning tasks is effective and can make full use of the complementarity of multi-source information. In the aspect of ship recognition, most methods do not consider the problem of similar ships; the detection targets are often obviously different. The surface of the ship target is very complex, and there may be different interference in different parts. To ensure the accuracy of the recognition results and solve the possible detection error in a single part, in this paper, multiple parts are detected and the ship is identified by voting mechanism and better detection results are obtained.

In the paper, the method was originally proposed to identify ships with similar features (destroyers and cruisers). To improve the accuracy of ship recognition and exclude ships irrelevant to the target, the relevant parameters of ships are regulated. This is why there are only three types of ships (destroyers, cruisers and others) in this paper. As shown in the figure, it can be seen that the length of the ship is regulated from 140 m to 200 m; in Figure 23a, the ships in the red box are identified as other ships due to their large scale; a similar situation occurs in the aircraft carrier of Figure 23b. This method has a certain extensibility, which can be extended to other types of ship detection and recognition under the condition of obtaining ship parameters, but the training samples of similar ship parts are cumbersome and need manual operation. In the future work, deep learning will be considered to solve the problem of ship part detection and the problem of the sliding window method of redundant calculations [36].



Figure 23. Recognition results of different types of ships. (a) The length of the vessel in the red box is 253.2 m. (b) The length of the vessel in the red box is 332.8 m.

5. Conclusions

In this work, a ship target detection and recognition method based on multi-source remote sensing image fusion was proposed. In order to improve the accuracy and speed of optical image ship detection and recognition, this paper introduced SAR data and research on port detection, ship detection and ship recognition, respectively. For the problem that the port area features of optical images are not obvious and there are many candidate slice images, the port slice images classification based on multi-feature fusion image was proposed to obtain the port slice images containing ship targets. For the ship detection errors caused by the complex optical image port environment and serious interference of the ship target shadow in the port slice image, the paper proposed a ship target detection method based on joint shape analysis and multi-feature classification and the ship target in scene was successfully detected. Finally, the ship position detection method based on feature point matching, contour extraction and brightness saliency was used to detect seven groups of position information, and the ship target model recognition was identified by the voting results of position information.

In conclusion, for one thing, compared with the single optical data target detection, the proposed ship detection and recognition method based on the optical and SAR images fusion can solve the problems that obscure optical image port area features and many candidate slice images. For another, the proposed method can solve the problems of ship false detection and missed detection caused by the complex port environment and serious interference of the ship target shadow in the optical image. In addition, more parts can be detected to improve the accuracy of recognition.

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

Funding: This work was supported in part by National Natural Science Foundation of China under Grant 61771170.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Xu, D. Reserch of High—Resolution SAR Images Feature Extraction and Pattern Retrieval. Master’s Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2016. (In Chinese)
2. Kayabol, K.; Zerubia, J. Unsupervised amplitude and texture classification of SAR images with multinomial latent model. *IEEE Trans. Image Process. (TIP)* **2013**, *22*, 561–572. [[CrossRef](#)] [[PubMed](#)]
3. Zhou, C.; Luo, J. *Geo—Computing of High Resolution Satellite Remote Sensing Images*; Science Press: Beijing, China, 2009; pp. 1–3. (In Chinese)
4. Long, T.; Zeng, T.; Hu, C. High resolution radar real—Time signal and information processing. *China Commun.* **2019**, *16*, 105–133.
5. Pavlov, V.A.; Belov, A.A.; Tuzova, A.A. Implementation of Synthetic Aperture Radar Processing Algorithms on the Jetson TX1 Platform. In Proceedings of the 2019 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech), St. Petersburg, Russia, 17–18 October 2019; pp. 90–93.
6. Tomiyasu, K.; Pacelli, J.L. Synthetic Aperture Radar Imaging from an Inclined Geosynchronous Orbit. *IEEE Trans. Geosci. Remote Sens.* **1983**, *21*, 324–329. [[CrossRef](#)]
7. Pour, A.B.; Hashim, M. The Application of ASTER Remote Sensing Data to Porphyry Copper and Epithermal Gold Deposits. *Ore Geol. Rev.* **2012**, *44*, 1–9. [[CrossRef](#)]
8. Liu, L.; Bai, Y.; Li, Y. Locality-Aware Rotated Ship Detection in High-Resolution Remote Sensing Imagery Based on Multiscale Convolutional Network. *IEEE Geosci. Remote Sens. Lett.* **2021**. [[CrossRef](#)]
9. Han, X. *Study on Key Technology of Typical Targets Recognition from Large-Field Optical Remote Sensing Images*; Harbin Institute of Technology: Harbin, China, 2013. (In Chinese)

10. Zhu, X.; Ma, C. The study of combined invariants optimization method on aircraft recognition. In Proceedings of the 2011 Symposium on Photonics and Optoelectronics (SOPO), Wuhan, China, 16–18 May 2011; pp. 1–4.
11. Fang, Z.; Yao, G.; Zhang, Y. Target recognition of aircraft based on moment invariants and BP neural network. In Proceedings of the World Automation Congress (WAC), Puerto Vallarta, Mexico, 24–28 June 2012; pp. 1–5.
12. Wang, Y.; Wang, C.; Zhang, H.; Dong, Y.; Wei, S. A SAR dataset of ship detection for deep learning under complex backgrounds. *Remote Sens.* **2019**, *11*, 765. [[CrossRef](#)]
13. Li, T.; Liu, Z.; Xie, R. An Improved Superpixel-Level CFAR Detection Method for Ship Targets in High-Resolution SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *99*, 1–11. [[CrossRef](#)]
14. Han, P.; Zhang, X.; Ge, P. Crashed airplane detection based on feature fusion in PolSAR image. In Proceedings of the 11th IEEE International Conference on Signal Processing, Beijing, China, 21–25 October 2012; pp. 2003–2006.
15. Zhang, X.; Huo, C.; Xu, N.; Jiang, H.; Cao, Y.; Ni, L.; Pan, C. Multitask Learning for Ship Detection from Synthetic Aperture Radar Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. (JSTAR)* **2021**, *14*, 8048–8062. [[CrossRef](#)]
16. Huang, L.; Chen, C. Study on image fusion algorithm of panoramic image stitching. *J. Electron. Inf. Technol.* **2014**, *36*, 1292–1298. (In Chinese)
17. Bhalerao, M.; Chandaliya, N.; Poojary, T.; Rathod, M. Review on Automatic Image Homogenization for Panoramic Images. In Proceedings of the 2nd International Conference on Advances in Science & Technology (ICAST), Maharashtra, India, 9 April 2019. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3375324 (accessed on 19 October 2021).
18. Li, M.; Dong, Y.; Wang, X. Pixel level image fusion based the wavelet transform. In Proceedings of the 2013 6th International Congress on Image and Signal Processing (CISP), Hangzhou, China, 16–18 December 2013; Volume 2, pp. 995–999.
19. You, T.; Tang, Y. Visual saliency detection based on adaptive fusion of color and texture features. In Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 13–16 December 2017; pp. 2034–2039.
20. Zhao, Y.; Yin, Y.; Fu, D. Decision-level fusion of infrared and visible images for face recognition. In Proceedings of the 2008 Chinese Control and Decision Conference, Yantai, China, 2–4 July 2008; pp. 2411–2414.
21. Han, N.; Hu, J.; Zhang, W. Multi-spectral and SAR images fusion via Mallat and À trous avelet transform. In Proceedings of the 2010 18th International Conference on Geoinformatics, Beijing, China, 18–20 June 2010; pp. 1–4.
22. Byun, Y.; Choi, J.; Han, Y. An Area-Based Image Fusion Scheme for the Integration of SAR and Optical Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2212–2220. [[CrossRef](#)]
23. Spröhnle, K.; Fuchs, E.-M.; Pelizari, P.A. Object-Based Analysis and Fusion of Optical and SAR Satellite Data for Dwelling Detection in Refugee Camps. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**. [[CrossRef](#)]
24. Rostami, M.; Kolouri, S.; Eaton, E.; Kim, K. SAR Image Classification Using Few-Shot Cross-Domain Transfer Learning. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 16–20 June 2019; pp. 907–915.
25. Park, K.A.; Park, J.J.; Jang, J.C.; Lee, J.H.; Oh, S.; Lee, M. Multi-spectral ship detection using optical, hyperspectral, and microwave SAR remote sensing data in coastal regions. *Sustainability* **2018**, *10*, 4064. [[CrossRef](#)]
26. Hou, X.; Ao, W.; Xu, F. End-to-end Automatic Ship Detection and Recognition in High-Resolution Gaofen-3 Spaceborne SAR Images. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 9486–9489.
27. Qi, S.; Ma, J.; Lin, J.; Li, Y.; Tian, J. Unsupervised Ship Detection Based on Saliency and S-HOG Descriptor From Optical Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1451–1455.
28. Achanta, R.; Hemami, S.; Estrada, F.; Susstrunk, S. Frequency-tuned salient region detection. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June 2009; pp. 1597–1604.
29. Lowe, D. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
30. Li, S.; Zhou, Z.; Wang, B.; Wu, F. A Novel Inshore Ship Detection via Ship Head Classification and Body Boundary Determination. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1920–1924. [[CrossRef](#)]
31. Gao, T.; Chen, H.; Chen, W. Adaptive Heterogeneous Support Tensor Machine: An Extended STM for Object Recognition Using an Arbitrary Combination of Multisource Heterogeneous Remote Sensing Data. *IEEE Trans. Geosci. Remote Sens.* **2021**. [[CrossRef](#)]
32. Koniusz, P.; Yan, F.; Gosselin, P.-H.; Mikolajczyk, K. Higher-Order Occurrence Pooling for Bags-of-Words: Visual Concept Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 313–326. [[CrossRef](#)] [[PubMed](#)]
33. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
34. Schölkopf, B.; Smola, A.J.; Williamson, R.C.; Bartlett, P.L. New support vector algorithms. *Neural Comput.* **2000**, *12*, 1207–1245. [[CrossRef](#)] [[PubMed](#)]
35. Tao, D.; Li, X.; Wu, X.; Hu, W.; Maybank, S.J. Supervised tensor learning. *Knowl. Inf. Syst.* **2007**, *13*, 1–42. [[CrossRef](#)]
36. Li, J.; Tian, J.; Gao, P.; Li, L. Ship Detection and Fine-Grained Recognition in Large-Format Remote Sensing Images Based on Convolutional Neural Network. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 2859–2862.