



Article

Polarimetric Synthetic Aperture Radar Ship Potential Area Extraction Based on Neighborhood Semantic Differences of the Latent Dirichlet Allocation Bag-of-Words Topic Model

Weixing Qiu ¹ and Zongxu Pan ^{2,3,4,*} ¹ School of Electronics and Information Engineering, Beihang University, Beijing 100191, China² Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China³ Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Chinese Academy of Sciences, Beijing 100190, China⁴ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101408, China

* Correspondence: zxpan@mail.ie.ac.cn

Abstract: Recently, deep learning methods have been widely studied in the field of polarimetric synthetic aperture radar (PolSAR) ship detection. However, extracting polarimetric and spatial features on the whole PolSAR image will result in high computational complexity. In addition, in the massive data ship detection task, the image to be detected contains a large number of invalid areas, such as land and seawater without ships. Therefore, using ship coarse detection methods to quickly locate the potential areas of ships, that is, ship potential area extraction, is an important prerequisite for PolSAR ship detection. Since existing unsupervised PolSAR ship detection methods based on pixel-level features often rely on fine sea-land segmentation pre-processing and have poor applicability to images with complex backgrounds, in order to solve the abovementioned issue, this paper proposes a PolSAR ship potential area extraction method based on the neighborhood semantic differences of an LDA bag-of-words topic model. Specifically, a polarimetric feature suitable for the scattering diversity condition is selected, and a polarimetric feature map is constructed; the superpixel segmentation method is used to generate the bag of words on the feature map, and latent high-level semantic features are extracted and classified with the improved LDA bag-of-words topic model method to obtain the PolSAR ship potential area extraction result, i.e., the PolSAR ship coarse detection result. The experimental results on the self-established PolSAR dataset validate the effectiveness and demonstrate the superiority of our method.



Citation: Qiu, W.; Pan, Z. Polarimetric Synthetic Aperture Radar Ship Potential Area Extraction Based on Neighborhood Semantic Differences of the Latent Dirichlet Allocation Bag-of-Words Topic Model. *Remote Sens.* **2023**, *15*, 5601. <https://doi.org/10.3390/rs15235601>

Academic Editors: Mohammad Awrangjeb, Danfeng Hong, Shou Feng, Nan Su, Chunhui Zhao, Yiming Yan and Qingsheng Xue

Received: 10 October 2023

Revised: 28 November 2023

Accepted: 30 November 2023

Published: 1 December 2023



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

Keywords: PolSAR ship detection; polarimetric features selection; superpixel; LDA topic model

1. Introduction

Synthetic aperture radar (SAR) is one of the main means in the field of remote sensing because of its all-day and all-weather imaging characteristics [1]. Polarization is an important attribute of electromagnetic waves. With the development of sensor technology, the SAR imaging mode has been extended from single polarization to full polarization. Applying all polarization information to the SAR system constitutes the polarimetric SAR (PolSAR) system. Compared with SAR, PolSAR can provide complete target electromagnetic scattering characteristics and polarization information [2]. Ship detection has been a research hotspot in the fields of SAR and PolSAR applications, which helps to strengthen maritime traffic management and has good application prospects in civilian and military fields, such as safeguarding maritime rights and improving maritime warning capabilities.

Since ship targets are generally active in the vast ocean, they have position uncertainty and target dispersion. In addition, the massive remote sensing data used for ship detection include a large number of land areas and seawater areas without ships. Therefore, how to

quickly and easily locate the area where ships exist in a complete remote sensing image through a ship coarse detection method, that is, extracting the potential areas of ships, is an important issue in ship detection tasks in massive remote sensing data. Another role of ship potential area extraction is to reduce the complexity of computation for polarimetric feature extraction in PolSAR ship detection tasks. In order to better extract the spatial and polarimetric features of PolSAR images and improve the detection effect, the latest PolSAR ship detection method based on deep learning [3] uses multiple polarimetric feature extraction methods to construct multi-channel data as input for deep learning networks. If polarimetric feature extraction is performed on the entire scene image, the computational complexity will be significant. Therefore, it is a feasible method to reduce the computational complexity by only calculating the polarimetric features of the potential areas through ship potential area extraction.

Currently, three types of methods for PolSAR ship potential area extraction, also known as ship coarse detection, are as follows: (1) Statistical distribution-based methods—since ship detection is looking for a specific target from the ocean background and ship targets have strong scattered echoes compared to sea clutter, ships can be detected by modeling sea clutter and searching for outliers through a statistical analysis. The Constant False-Alarm Rate (CFAR) method and its variants [4] belong to this category. The core of CFAR methods is to model sea clutter more accurately, e.g., Liu et al. [5] applied an adaptive truncation method to estimate the parameters of the statistical models in PolSAR images. (2) Polarimetric scattering-feature-based methods, including various polarization decomposition methods [6–10]—in a PolSAR target detection task, Bordbari et al. [11] categorized the scattering mechanism into target and non-target and used subspace projection to improve the detection performance. (3) Spatial-feature-based methods—spatial-feature-based methods use manually designed or automatically learned features extracted in the spatial domain to distinguish ships from the background. Grandi et al. [12] used wavelet features to detect targets in PolSAR images, which explains the dependence of texture measurements on the polarization state. All of the above ship potential area extraction methods have some limitations. On the one hand, CFAR-type methods usually need to perform an accurate sea–land segmentation first to ensure that the background is seawater. Although current methods based on GIS information can quickly and efficiently exclude large land areas, the fine segmentation at the sea–land boundaries still relies on specially designed sea–land segmentation methods. In addition, the non-homogeneous sea clutter under a complex sea state makes it difficult to model the clutter distribution uniformly on the whole PolSAR image. On the other hand, PolSAR ship potential area extraction methods based on polarimetric features as well as spatial features rely on the accurate description of the features. The backscattering from radar targets is sensitive to the relative geometric relationship between the target attitude and the radar line of sight, which leads to scattering diversity of the target [13,14], and the scattering diversity makes the polarimetric and spatial features of the ship variable, which makes it difficult to detect based on the pixel-level features. In addition, complex sea surface backgrounds, including islands, waves, ship wakes, defocusing, azimuth ambiguity, cross sidelobes, stripe noise, strong scattering artificial targets (e.g., lighthouses and buoys), etc., can interfere with the detection of ship targets in the case of imperfect feature descriptions, resulting in false alarms and missed alarms. The visualization of some false alarms is shown in Figure 1.

We designed a method for the PolSAR ship potential area extraction task, so that it can be applied to both the nearshore areas containing part of the land and the distant sea areas, and it can coarsely detect ships while excluding the interference of complex backgrounds without relying on the fine land–sea segmentation algorithms other than GIS information. Under the premise that extracting traditional pixel-level features is not effective, trying to use the latent high-level semantic information in PolSAR images is a good idea to solve the problem. The bag-of-words (BOW) models and the topic models were initially used for text data mining and natural language processing (NLP), which can extract the semantic information, especially latent topic information, in documents. They can also be applied in the field of remote sensing image processing if the images or image blocks are regarded

as the documents or bag of words [15]. Sivic et al. [16] introduced the bag-of-words model for the first time in the field of computer vision. The visual bag-of-words model treats an image as a set of local visual features within a bag and ignores the spatial layout information of the features. It borrows the idea of the traditional bag-of-words model, which treats the features extracted from an image as visual words and ignores the order of occurrence and grammatical structure of the words. By statistically modeling visual words, the features are reduced in dimensionality. For the case that multiple visual features correspond to a visual word, Yuan et al. [17] proposed a meaningful spatially co-occurrent pattern of visual words to eliminate the influence of polysemous visual words. For the topic model, Deerwester et al. [18] proposed the latent semantic analysis (LSA) model. Later, Hofmann [19] extended it to Probabilistic LSA (pLSA). Bosch et al. [20] regarded image classes as latent topics, using the pLSA method to automatically obtain these latent topics from bag-of-words features of images for classification. The Latent Dirichlet Allocation (LDA) model [21] is also a classic generative topic model, which introduces parameters that follow the Dirichlet distribution on the basis of pLSA to establish the probability distribution of the latent topic variable. Li et al. [22] used the LDA model for scene classification for the first time, while Zhong et al. [23] utilized an improved LDA topic model for natural image classification. There are the following problems to be solved when using the LDA topic model for PolSAR ship potential area extraction: Firstly, the bag-of-words generation method should be optimized, so that each bag of words contains homogeneous features as much as possible to facilitate the subsequent semantic information extraction. Secondly, the original PolSAR image has a large height and width, and when there are too many pixels, the LDA topic model has large computational complexity, so measures need to be taken to reduce the computational complexity. Thirdly, some of the targets with similar semantic features are not actually homogeneous targets, and further precise differentiation between them is needed to improve the precision rate as much as possible on the basis of ensuring a high recall rate for ship coarse detection.

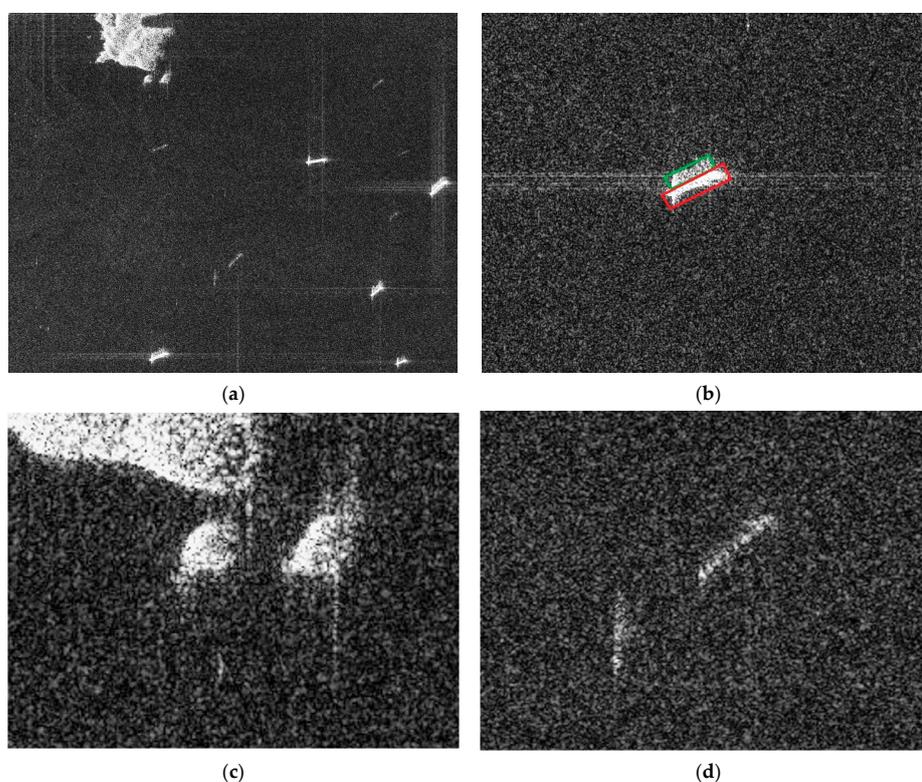


Figure 1. Complex backgrounds in the PolSAR ship potential area extraction (coarse detection) task: (a) overall view; (b) the green rectangle represents the false alarm of defocusing, and the red rectangle represents the real ship; (c) false alarm of islands; (d) false alarm of azimuth ambiguity.

In this article, we propose a PolSAR ship potential area extraction (coarse detection) method based on neighborhood semantic differences of the LDA bag-of-words topic model (NSD-LDA). Firstly, in order to reduce the effect of scattering diversity, the unified polarimetric rotation domain theory proposed by Chen et al. [24–26] is introduced. By selecting several typical polarimetric rotation domain feature parameters, a feature map suitable for extracting high-level semantic features is obtained, which not only maximizes the differences between the target and the background but also maximizes prior homogeneous regions to reduce the computational complexity of the subsequent semantic feature extraction. Secondly, we generate the bag of words via an improved superpixel segmentation method. The traditional superpixel segmentation method is not applicable to the selected feature maps, and a more suitable superpixel segmentation method can be obtained by improving the seed point selection, iteration strategy, and termination conditions. Then, on the bag of words obtained with the superpixel segmentation method, high-level semantic information is extracted using our proposed NSD-LDA method. Specifically, in order to enhance the correlation between polarimetric and spatial features, making the extracted high-level semantic information more accurate, on the basis of generating the bag of words using the superpixel method, the differences between the semantic vectors of the target bag of words and its neighboring bags of words are used to replace the original target semantic vectors as the extracted high-level semantic features. Finally, based on the extracted high-level semantic features, the PolSAR ship potential area extraction (coarse detection) is completed using an SVM classifier, prior knowledge, and morphological post-processing. The main contributions of this article are summarized as follows:

1. We propose an unsupervised PolSAR ship potential area extraction (coarse detection) method, which can effectively migrate images obtained from the same type of sensors and facilitate deployment on large-scale production lines.
2. By extracting high-level semantic features of the generated bag of words, our method has better applicability to complex backgrounds including parts of land.
3. Through polarimetric rotation domain feature selection, improved superpixel bag-of-words generation, and high-level semantic features extraction, our method further strengthens the correlation between polarimetric and spatial features, resulting in more robust ship detection results.

The innovations of our method are summarized as follows:

1. By selecting polarimetric rotation domain feature parameters under dual-constraint conditions, we improved the discrimination between the target and background while expanding prior homogeneous semantic regions, and obtained polarimetric feature maps suitable for subsequent bag-of-words generation and high-level semantic feature extraction.
2. By improving the superpixel segmentation method and using prior information guidance, the bag of words applicable to the selected polarimetric feature map is constructed, which combines polarimetric features with spatial features and significantly reduces the computational complexity of the subsequent semantic feature extraction.
3. With the proposed NSD-LDA method, polarimetric and spatial features are more correlated, and the extracted potential areas of ships are more accurate.

The remainder of this paper is organized as follows: The proposed method is detailed in Section 2, followed by experimental results in Section 3. Some discussions are presented in Section 4, and Section 5 concludes the paper.

2. Methods

In this paper, we propose a PolSAR ship potential area extraction method based on neighborhood semantic differences of an LDA bag-of-words topic model. A flowchart of this method is given in Figure 2. Firstly, several polarimetric rotation domain feature parameters are extracted from the original PolSAR image and compared for selection, and the feature parameters are selected to construct a polarimetric feature map containing

polarimetric information based on the maximization of the difference between the ship target and the background and the maximization of the number of pixels in the prior homogeneous region (seawater). Details will be presented in Section 2.1. Secondly, the selected polarimetric feature map is clustered to generate a bag of words by using an improved superpixel method. All details are discussed at length in Section 2.2. Thirdly, the high-level semantic information is extracted with the proposed NSD-LDA method for each bag of words. This part will be discussed thoroughly in Section 2.3. Finally, the extracted semantic information is classified using the SVM classifier and then post-processed using expert knowledge to obtain the results of ship potential area extraction. This part will be presented in Section 2.4.

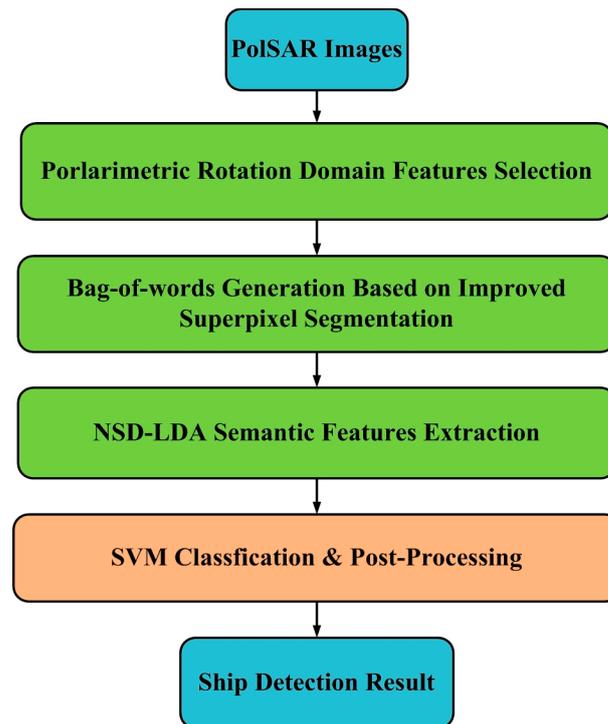


Figure 2. Flowchart of the proposed method.

2.1. Polarimetric Rotation Domain Features Selection

2.1.1. Characterization of Polarimetric Rotation Domain Feature Parameters

The scattering diversity of radar targets makes SAR/PolSAR information processing more difficult. In order to explore and utilize the information contained in this scattering diversity, Chen et al. extended the polarimetric information obtained under specific imaging geometric relations to the direction rotating around the radar line of sight, and they proposed a unified polarimetric rotation domain theory [24] and polarimetric correlation/coherence feature rotation domain interpretation tools [25,26].

Specifically, for PolSAR images, the polarimetric scattering matrix S can be represented as follows under horizontal (H) and vertical (V) polarization bases:

$$S = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix}, \quad (1)$$

where S_{HV} is the backscattered coefficient from vertical polarization transmission and horizontal polarization reception. Other terms are similarly defined.

By rotating the polarimetric scattering matrix S around the radar line of sight, the polarimetric scattering matrix in the rotation domain $S(\theta)$ can be obtained:

$$S(\theta) = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad (2)$$

where θ is the rotation angle, and $\theta \in [-\pi, \pi)$.

The correlation values and coherence values between different polarization channels in PolSAR images contain rich polarimetric information. For two arbitrary polarization channels s_X and s_Y , the polarimetric correlation pattern can be written as

$$|\hat{\gamma}_{X-Y}| = |\langle s_X s_Y^* \rangle|, \quad (3)$$

and the polarimetric coherence pattern can be written as

$$|\gamma_{X-Y}| = \frac{|\langle s_X s_Y^* \rangle|}{\sqrt{\langle |s_X|^2 \rangle \langle |s_Y|^2 \rangle}}, \quad (4)$$

where s_Y^* is the conjugate of s_Y . By extending the above two parameters to the polarimetric rotation domain, the polarimetric rotation domain correlation pattern can be written as

$$|\hat{\gamma}_{X-Y}(\theta)| = |\langle s_X(\theta) s_Y^*(\theta) \rangle|, \quad (5)$$

and the polarimetric rotation domain coherence pattern can be written as

$$|\gamma_{X-Y}(\theta)| = \frac{|\langle s_X(\theta) s_Y^*(\theta) \rangle|}{\sqrt{\langle |s_X(\theta)|^2 \rangle \langle |s_Y(\theta)|^2 \rangle}}, \quad (6)$$

where the value of $|\hat{\gamma}_{X-Y}(\theta)|$ is within $[0, +\infty)$, and the value of $|\gamma_{X-Y}(\theta)|$ is within $[0, 1)$.

Taking a polarimetric correlation pattern as an example, for four kinds of independent polarimetric correlation patterns, including $|\hat{\gamma}_{HH-HV}(\theta)|$, $|\hat{\gamma}_{HH-VV}(\theta)|$, $|\hat{\gamma}_{(HH+VV)-(HH-VV)}(\theta)|$, and $|\hat{\gamma}_{(HH-VV)-HV}(\theta)|$, the following seven amplitude class feature parameters are defined for feature characterization, including the original correlation $\hat{\gamma}_{-org} = |\hat{\gamma}_{X-Y}(0)|$, the mean value of correlation $\hat{\gamma}_{-mean} = \text{mean}\{|\hat{\gamma}_{X-Y}(\theta)|\}$, the maximum correlation $\hat{\gamma}_{-max} = \text{max}\{|\hat{\gamma}_{X-Y}(\theta)|\}$, the minimum correlation $\hat{\gamma}_{-min} = \text{min}\{|\hat{\gamma}_{X-Y}(\theta)|\}$, the standard deviation of correlation $\hat{\gamma}_{-std} = \text{std}\{|\hat{\gamma}_{X-Y}(\theta)|\}$, the correlation contrast $\hat{\gamma}_{-contrast} = \hat{\gamma}_{-max} - \hat{\gamma}_{-min}$, and the correlation anisotropy $\hat{\gamma}_{-Ani} = \hat{\gamma}_{-max} - \hat{\gamma}_{-min} / \hat{\gamma}_{-max} + \hat{\gamma}_{-min}$.

Similarly, for four kinds of independent polarimetric coherence patterns, seven polarimetric coherence feature parameters, which are consistent with polarimetric correlation feature parameters, are also defined for feature characterization, including $\gamma_{-org} = |\gamma_{X-Y}(0)|$, $\gamma_{-mean} = \text{mean}\{|\gamma_{X-Y}(\theta)|\}$, $\gamma_{-max} = \text{max}\{|\gamma_{X-Y}(\theta)|\}$, $\gamma_{-min} = \text{min}\{|\gamma_{X-Y}(\theta)|\}$, $\gamma_{-std} = \text{std}\{|\gamma_{X-Y}(\theta)|\}$, $\gamma_{-contrast} = \gamma_{-max} - \gamma_{-min}$, and $\gamma_{-Ani} = \gamma_{-max} - \gamma_{-min} / \gamma_{-max} + \gamma_{-min}$.

In summary, a total of 56 polarimetric rotation domain feature parameters are obtained, all of which contain rich polarimetric information and have clear physical meanings.

2.1.2. Polarimetric Rotation Domain Feature Parameter Selection and Feature Map Construction

In order to construct a feature map containing polarimetric information for subsequent bag-of-words generation and high-level semantic feature extraction, the 56 polarimetric rotation domain feature parameters are selected to find which meets the following two conditions best. One is to maximize the difference between ship targets and various backgrounds. Enhancing the differentiation between targets and backgrounds can make the

ship potential area extraction results more accurate. The second is to maximize the number of pixels belonging to the prior homogeneous regions (seawater). Seawater is the most dominant background, and if it is excluded through prior information, the computational complexity of the subsequent semantic feature extraction step can be significantly reduced.

The Relief method [27] is a well-known filtered feature selection method that estimates the weight of each feature based on its ability to classify between different classes of samples. A total of 1500 ship pixels and 1500 background pixels, including 500 calm sea surface pixels, 500 land/island/reef pixels, and 100 pixels each of wave, defocus, azimuth ambiguity, cross sidelobe, and stripe noise, are randomly selected from the GF-3 dataset for the weight calculation. The results are shown in Tables 1 and 2. There are 56 polarimetric rotation domain feature parameters, and the larger the weight value is, the stronger the ability of this polarimetric rotation domain feature parameter to discriminate between the ships and the backgrounds.

Table 1. Classification weights of polarimetric rotation domain correlation feature parameters.

Feature Parameters	Polarimetric Correlation Patterns			
	$ \hat{\gamma}_{HH-HV}(\theta) $	$ \hat{\gamma}_{HH-VV}(\theta) $	$ \hat{\gamma}_{(HH+VV)-(HH-VV)}(\theta) $	$ \hat{\gamma}_{(HH-VV)-HV}(\theta) $
$\hat{\gamma}_{-org} = \hat{\gamma}_{X-Y}(0) $	0.85	0.19	0.26	0.91
$\hat{\gamma}_{-mean} = \text{mean}\{ \hat{\gamma}_{X-Y}(\theta) \}$	0.61	0.53	0.63	0.56
$\hat{\gamma}_{-max} = \text{max}\{ \hat{\gamma}_{X-Y}(\theta) \}$	0.51	0.29	0.46	0.52
$\hat{\gamma}_{-min} = \text{min}\{ \hat{\gamma}_{X-Y}(\theta) \}$	0.53	0.25	0.41	0.71
$\hat{\gamma}_{-std} = \text{std}\{ \hat{\gamma}_{X-Y}(\theta) \}$	0.51	0.53	0.37	0.35
$\hat{\gamma}_{-contrast} = \hat{\gamma}_{-max} - \hat{\gamma}_{-min}$	0.46	0.44	0.15	0.29
$\hat{\gamma}_{-Ani} = \hat{\gamma}_{-max} - \hat{\gamma}_{-min} / \hat{\gamma}_{-max} + \hat{\gamma}_{-min}$	0.41	0.38	0.39	0.31

Table 2. Classification weights of polarimetric rotation domain coherence feature parameters.

Feature Parameters	Polarimetric Coherence Patterns			
	$ \gamma_{HH-HV}(\theta) $	$ \gamma_{HH-VV}(\theta) $	$ \gamma_{(HH+VV)-(HH-VV)}(\theta) $	$ \gamma_{(HH-VV)-HV}(\theta) $
$\gamma_{-org} = \gamma_{X-Y}(0) $	0.31	0.36	0.39	0.33
$\gamma_{-mean} = \text{mean}\{ \gamma_{X-Y}(\theta) \}$	0.32	0.34	0.27	0.67
$\gamma_{-max} = \text{max}\{ \gamma_{X-Y}(\theta) \}$	0.23	0.78	0.32	0.75
$\gamma_{-min} = \text{min}\{ \gamma_{X-Y}(\theta) \}$	0.26	0.22	0.31	0.29
$\gamma_{-std} = \text{std}\{ \gamma_{X-Y}(\theta) \}$	0.32	0.39	0.43	0.56
$\gamma_{-contrast} = \gamma_{-max} - \gamma_{-min}$	0.44	0.52	0.33	0.59
$\gamma_{-Ani} = \gamma_{-max} - \gamma_{-min} / \gamma_{-max} + \gamma_{-min}$	0.33	0.37	0.35	0.48

After comparison, the classification weights of feature parameters $|\hat{\gamma}_{(HH-VV)-HV}(0)|$ and $|\hat{\gamma}_{HH-HV}(0)|$ are the highest, with values of 0.91 and 0.85, respectively. We choose these two feature parameters to construct polarimetric rotation domain feature maps separately and calculate the proportion of prior homogeneous region (seawater) pixels.

Five GF-3 PolSAR images each of a nearshore and distant ocean are selected as the dataset. For the polarimetric rotation domain correlation features, perform a truncation operation, set the pixel values of the original polarimetric correlation features exceeding 255 to 255, and round down other values to construct an 8-bit feature map. According to reference [28], if the target satisfies the reflection symmetry property, its cross-polarization scattering coefficients and co-polarization scattering coefficients are uncorrelated, which are represented as follows:

$$\begin{cases} \langle S_{HH}S_{HV}^* \rangle = 0, & \langle S_{HH}S_{VH}^* \rangle = 0 \\ \langle S_{VV}S_{HV}^* \rangle = 0, & \langle S_{VV}S_{VH}^* \rangle = 0 \end{cases} \quad (7)$$

In geophysical media, this symmetry can be observed on water surfaces in the upwind or downwind direction and isotropic and anisotropic scattering media, such as snow or sea ice. The feature parameters $|\hat{\gamma}_{(HH-VV)-HV}(0)|$ and $|\hat{\gamma}_{HH-HV}(0)|$ characterize the correlation between the cross-polarization scattering coefficient and co-polarization scattering coefficient. In the feature maps constructed with the feature parameters, for a calm sea surface, the pixel value is 0 due to the reflection symmetry. For sea clutter caused by waves, the pixel value is below 10. For artificial targets, the pixel value is large. This is determined by the physical properties of the targets and is general in PolSAR images. In order to choose the feature map that maximizes the number of pixels with semantic seawater, the proportion of pixels with a value of 0, a value not exceeding 10 and 20, is counted, and the results are shown in Table 3.

Table 3. The proportion of prior pixels in polarimetric rotation domain feature maps.

	Pixel Value of 0	Pixel Value Not Exceeding 10	Pixel Value Not Exceeding 20
Nearshore $ \hat{\gamma}_{HH-HV}(0) $	65.2	78.3	82.6
Distant ocean $ \hat{\gamma}_{HH-HV}(0) $	17.0	75.6	83.8
Nearshore $ \hat{\gamma}_{(HH-VV)-HV}(0) $	65.7	78.2	82.3
Distant ocean $ \hat{\gamma}_{(HH-VV)-HV}(0) $	17.3	76.0	83.4

The sea surface in nearshore PolSAR images is usually relatively calm, and based on reflection symmetry, pixels with a value of 0 can be identified as seawater. Due to the presence of a large amount of sea clutter caused by waves in the distant ocean PolSAR images, pixels with a value not exceeding 10 can be recognized as oceans. We compare the number of 0-value pixels in nearshore images and the number of pixels with a value not exceeding 10 in distant ocean images; feature $|\hat{\gamma}_{(HH-VV)-HV}(0)|$ is selected to construct the polarimetric rotation domain feature map, and subsequent processing is performed. The constructed polarimetric feature maps are shown in Figures 3 and 4.

Large areas of land can be excluded using GIS information. In order to expand the semantic prior areas and minimize the impact on subsequent semantic extraction of other targets, pixel values less than 2 in the nearshore feature maps are set to 0, and pixel values less than 10 in the distant ocean feature maps are set to 0.

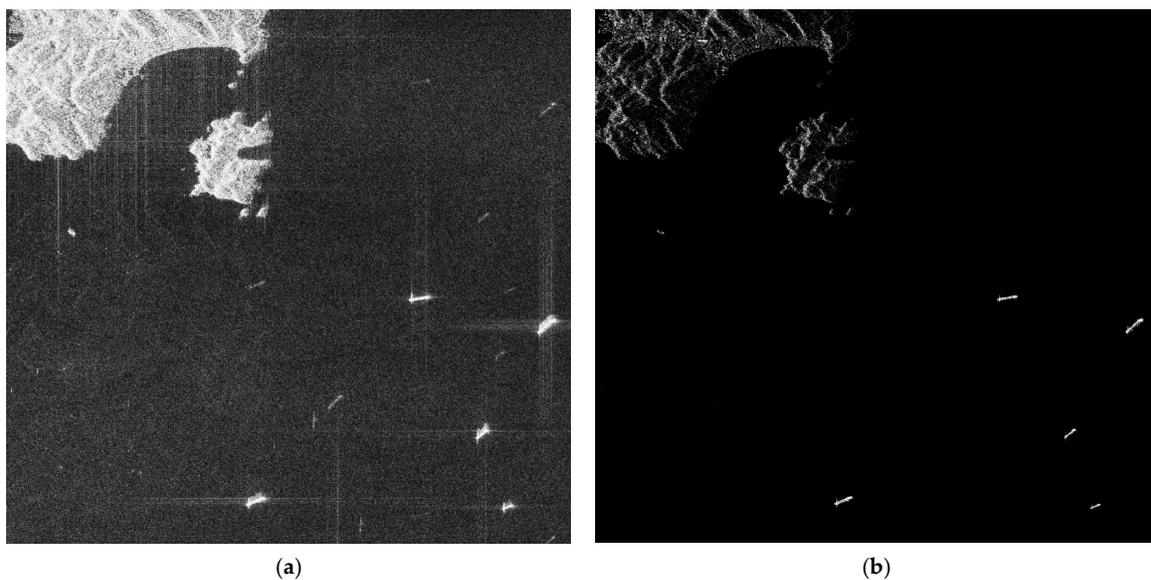


Figure 3. Nearshore polarimetric rotation domain feature map: (a) the HV channel of original PolSAR image; (b) constructed polarimetric feature map.

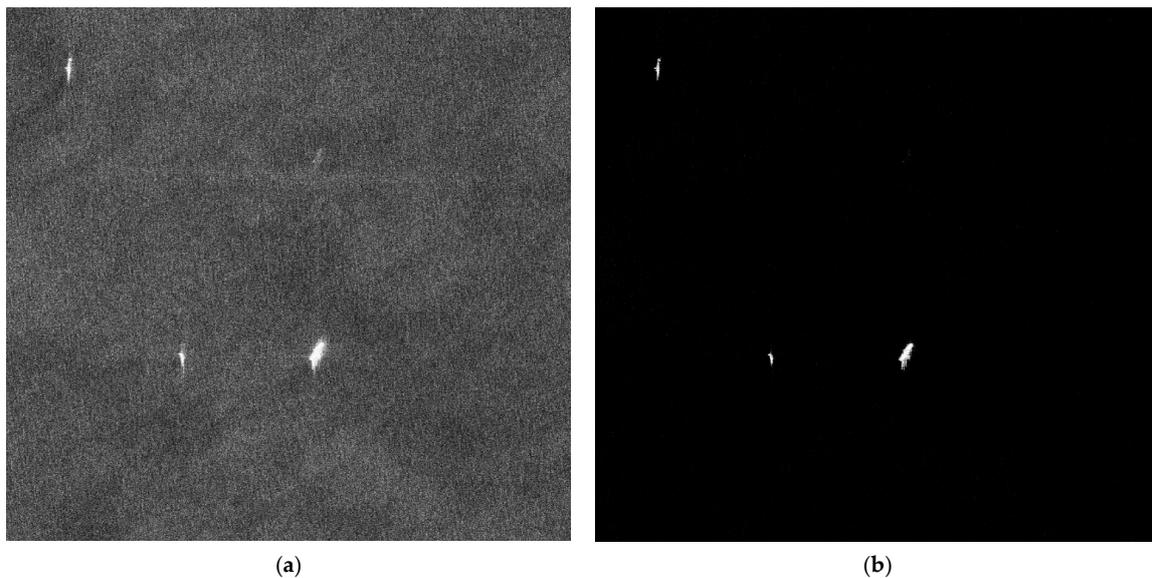


Figure 4. Distant ocean polarimetric rotation domain feature map: (a) the HV channel of original PolSAR image; (b) constructed polarimetric feature map.

2.2. Bag-of-Words Generation Based on Improved Superpixel Segmentation

The concept of a bag of words was first introduced in the field of natural language processing (NLP). The core of the concept is that if a text is treated as a bag, the word order of the words in it will not be considered, and it will only be treated as a set composed of words. This makes the bag of words also applicable in the field of computer vision (CV). If an image or a pixel block is treated as a bag of words, the pixels inside can be considered as words. A superpixel is a set of pixels with similar underlying features and similar spatial distances. By generating superpixels, an image can be dimensionally reduced for easy subsequent processing. In this article, we introduce the superpixel method to generate a bag of words. The most widely used superpixel segmentation methods include the following two types. One is based on changes in regional contours, which is represented as the watershed [29] method. The other is the clustering-based method, represented by Simple Linear Iterative Clustering (SLIC) [30].

After polarimetric rotation domain feature extraction, we obtain the feature map containing polarimetric information, and generating the bag of words on the feature map via the superpixel method will have the following problems: Firstly, like the original PolSAR image, the polarimetric rotation domain feature map also has speckle noise, which will have a negative effect on superpixel segmentation. Secondly, the seawater part of the polarimetric rotation domain feature map approaches a zero value, resulting in a large number of isolated points composed of one or several pixels scattered disorderly on the sea surface in the nearshore feature map. On the other hand, in the distant ocean feature map, there are some areas with irregular low-amplitude clutter pixel blocks. These can also have a negative effect on superpixel segmentation. Finally, the boundary of heterogeneous regions, such as ships, land, and sea clutter being blurred by speckle noise, degrades the accuracy of edge extraction as well as making the clustering results inaccurate.

To address the above problems, we use the following approach to optimize the superpixel segmentation process.

2.2.1. Edge Extraction of Polarimetric Feature Map

By extracting the edge information of the polarimetric rotation domain feature map as a constraint condition for generating superpixels, the generated superpixels can better fit the edge of the ground object. Considering the speckle noise, Gaussian Gamma-Shaped bi-windows (GGSBi) [31] are introduced to replace a conventional rectangular window.

Specifically, assuming the bi-windows are horizontal, the GGSBi function of pixel (x, y) is as follows:

$$\begin{aligned} W_U(x, y) &= \frac{|y|^{\alpha-1}}{\sqrt{2\pi}\sigma_x\Gamma(\alpha)\beta^\alpha} \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{|y|}{\beta}\right)\right), y \geq 0 \\ W_L(x, y) &= \frac{|y|^{\alpha-1}}{\sqrt{2\pi}\sigma_x\Gamma(\alpha)\beta^\alpha} \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{|y|}{\beta}\right)\right), y \leq 0 \end{aligned} \quad (8)$$

where $W_U(x, y)$ is the upper window and $W_L(x, y)$ is the lower window. Follow a Gaussian distribution along the x direction, with a parameter of σ_x controlling the window length, and the range of values is $\sigma_x > 1$. Follow a gamma distribution along the y direction, with parameters α and β . α controls the spacing of the two windows, $\Gamma(\alpha)$ represents the gamma function, β controls the window width, and the range of values is $\alpha > 1, \beta > 0$. In this article, the values are set to $\sigma_x = 6.5/\sqrt{\pi}, \alpha = 2, \beta = 1.6$.

Rotate the bi-windows counterclockwise along the centerline to obtain the bi-windows' function with orientation angle θ :

$$\begin{aligned} W_U^\theta(x, y) &= W_U(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) \\ W_L^\theta(x, y) &= W_L(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) \end{aligned} \quad (9)$$

At each orientation, two local mean functions are computed with the following convolutions:

$$\begin{aligned} m_U(x, y|\theta) &= \sum_{(x', y')} W_U^\theta(x', y') I(x - x', y - y') \\ m_L(x, y|\theta) &= \sum_{(x', y')} W_L^\theta(x', y') I(x - x', y - y') \end{aligned} \quad (10)$$

When the orientation angle θ is discretized into $\theta_p = 0, \pi/P, \dots, \pi(P-1)/P$, the ratio-based edge strength map $ESM(x, y)$ is

$$ESM(x, y) = 1 - \zeta_R(x, y), \quad (11)$$

where $\zeta_R(x, y)$ is calculated with

$$\zeta_R(x, y) = \min_{p=0,1,\dots,P-1} \left\{ \min \left\{ \frac{m_U(x, y|\theta_p)}{m_L(x, y|\theta_p)}, \frac{m_L(x, y|\theta_p)}{m_U(x, y|\theta_p)} \right\} \right\}. \quad (12)$$

The edge directional map $EDM(x, y)$ is

$$EDM(x, y) = \frac{\pi}{P} \operatorname{argmin}_p \left\{ \min \left\{ \frac{m_U(x, y|\theta_p)}{m_L(x, y|\theta_p)}, \frac{m_L(x, y|\theta_p)}{m_U(x, y|\theta_p)} \right\} \right\}. \quad (13)$$

The set of edge pixels can be obtained through the Non-Maximum Suppression (NMS) method.

The parameter setting of the GGS bi-windows is determined with the edge extraction effect. When the window is small and the distance between the two windows is large, it has better adaptability to edges with large curvature.

Due to the numerous evaluation indicators for edge extraction effectiveness and the difficulty in determining which one is most suitable, the accuracy of edge extraction is mainly obtained through visual interpretation. In addition, we use the following methods to assist in determining the accuracy of edge extraction: If the extracted edge pixel is within a specified tolerance of the ground truth pixel, then it is counted as a true edge pixel. Calculate the proportion of true edge pixels extracted from typical ground objects, such as ships, lands, islands, and defocusing, as well as the proportion of missed edge pixels caused by speckle noise to all edge pixels extracted. When the proportion of true edge pixels is high enough and the proportion of missed edge pixels is low, the edge extraction effect meets the requirements.

If the PolSAR image resolution changes, the GGS bi-windows' parameters need to be reset using the above method, provided that the task of extracting the ship potential area remains unchanged, i.e., the scene of edge extraction remains unchanged.

2.2.2. A Clustering Method Suitable for Speckle Noise and Low Amplitude, Low Discrimination Areas

This section proposes a method for clustering on the polarimetric rotation domain feature map to obtain initial superpixels. Since seawater contains a large number of low-amplitude pixel blocks and isolated points with values of 0 or approaching 0, when clustering, on the one hand, it is necessary to reduce the impact of speckle noise, and on the other hand, it should be suitable for a large number of low-amplitude, low-discrimination areas. Under edge constraint conditions, after cutting the polarimetric rotation domain feature map to obtain the initial blocks, seed points selection and pixels clustering are carried out for 0-value areas; low-amplitude, low-discrimination areas; and speckle noise areas in nearshore and distant sea scenes, respectively. The specific clustering steps are as follows:

1. Divide the original polarimetric rotation domain feature map into n blocks of size $S \times S$, where $n = MN/S^2$; M and N are the length and width of the original feature map.
2. Clustering of low-amplitude, low-discrimination areas and 0-value areas in the nearshore scene: For each initial block, if the original feature map is a nearshore feature map and the mean value of the pixels in the block is less than 10, find the point with the lowest pixel value and gradient from the center towards the edge as the initial seed point. When the value of the initial seed point is 0, if the value of its unlabeled neighbor pixel is also 0, it is merged into the superpixel to which the seed point belongs. When the value of the initial seed point is not 0, if the value of the unlabeled neighbor pixel has a difference with the seed point not greater than 3, or the difference with the pixel value of the superpixel's center point is not greater than 5, then it is merged into the superpixel to which the seed point belongs, and the seed point is updated to these neighbor pixels, and then the center-point position of the new superpixel is updated and the center-point amplitude is updated to the mean value of the new superpixel. Repeat this step until there are only isolated points left in the block.
3. Clustering of speckle noise areas in the nearshore scene: If the original feature map is a nearshore feature map and the mean value of the pixels in the block is not less than 10, find the point with the lowest gradient in the central 3×3 neighborhood as the initial seed point. For the unlabeled neighborhood pixels of the seed point, calculate its dissimilarity $\delta(i, j)$ with the seed point. Assuming the speckle noise follows a gamma distribution, the dissimilarity is defined as the likelihood ratio statistic of the 5×5 pixel block centered on two pixel points:

$$\delta(i, j) = 2M \ln \frac{\sum_{k=1}^M P_i(k) + \sum_{k=1}^M P_j(k)}{2 \sqrt{\sum_{k=1}^M P_i(k) \cdot \sum_{k=1}^M P_j(k)}}, \quad (14)$$

where M is the number of pixels in the pixel block around the pixel point, i.e., $5 \times 5 = 25$; $P_i(k)$ and $P_j(k)$ are the values of each pixel in the block. If the dissimilarity is less than 0.3, the neighboring pixel is merged into the superpixel to which the seed point belongs, and then the center-point position of the new superpixel is updated and the center-point amplitude is updated to the mean value of the new superpixel. Repeat this step until there are only isolated points left in the block.

4. Clustering of low-amplitude, low-discrimination areas and 0-value areas in the distant ocean scene: For each initial block, if the original feature map is a distant ocean feature

map and the mean value of the pixels in the block is less than 20, and when the value of the initial seed point is 0, the clustering method is consistent with the clustering method for the 0-value areas in step 2. When the value of the initial seed point is not 0, the thresholds for the difference between the unlabeled point and seed point, as well as between the unlabeled point and superpixel's center point, are set to 6 and 10, respectively. The clustering method is consistent with the clustering method for low-amplitude, low-discrimination areas in step 2.

5. Clustering of speckle noise areas in the distant ocean scene: If the original feature map is a distant ocean feature map and the mean value of the pixels in the block is not less than 20, the clustering method is consistent with the clustering method for speckle noise areas in step 3.
6. The edge information obtained from edge extraction constrains the clustering results mentioned above, so that the generated superpixel boundaries do not cross the edges.

2.2.3. Post-Processing of Homogeneous Region Merging

This section proposes a method for merging homogeneous superpixels. After clustering, the initial superpixels are obtained, but a large number of superpixels have boundaries falling on the initial block boundaries. In addition, small-area superpixels and isolated points make the generated superpixels discontinuous, requiring post-processing steps to merge homogeneous regions. Under edge constraint conditions, merge the cross-edge homogeneous regions of the initial superpixels obtained by clustering, and merge isolated points and small-area superpixels into the neighboring superpixels with the smallest dissimilarity. The specific steps are as follows:

1. For the superpixels on both sides of the initial block boundary, if they are homogeneous regions, merge them. Homogeneous regions include 0-value regions; low-amplitude, low-discrimination regions; and regions with speckle noise. The merging conditions are consistent with the clustering conditions of each region. Among them, low-amplitude, low-discrimination regions are calculated as thresholds based on the mean of superpixels, while regions with speckle noise have a threshold of the dissimilarity of superpixels $\delta(SP_i, SP_j)$ less than 0.3. The dissimilarity is defined as follows:

$$\delta(SP_i, SP_j) = 2\min(M(SP_i), M(SP_j)) \ln \frac{\sum_{k=1}^{M(SP_i)} P_{SP_i}(k) + \sum_{k=1}^{M(SP_j)} P_{SP_j}(k)}{2\sqrt{\sum_{k=1}^{M(SP_i)} P_{SP_i}(k) \cdot \sum_{k=1}^{M(SP_j)} P_{SP_j}(k)}}, \quad (15)$$

where $M(SP_i)$ and $M(SP_j)$ are the number of pixels in the superpixel, and $P_{SP_i}(k)$ and $P_{SP_j}(k)$ are the values of each pixel in the superpixel.

2. For small-area superpixels, calculate the dissimilarity $\delta(SP_i, SP_j)$ with their neighboring superpixels to merge them into the superpixel with the smallest dissimilarity. When the number of pixels in a superpixel is less than $0.3 S^2$, the superpixel is considered to be a small-area superpixel, and S is the initial block edge length.
3. For isolated points, calculate the dissimilarity $\delta(i, SP_j)$ with their neighboring superpixels to merge them into the superpixel with the smallest dissimilarity.
4. The edge information obtained from edge extraction constrains the post-processing results mentioned above, so that the generated superpixel boundaries do not cross the edges.

After post-processing, semantic labels are directly assigned to some of the superpixels based on prior knowledge. Among them, the 0-value superpixels and low-amplitude, low-discrimination superpixels are seawater, and the merged superpixels have the same semantics as the superpixels, which merge other superpixels, rather than the superpixels that are merged. A bag of words is generated for the remaining unassigned semantic

labeled superpixels for subsequent semantic feature extraction. By pre-assigning labels with prior knowledge, a large number of seawater regions can be identified, reducing the computational complexity of subsequent semantic feature extraction. The comparison results of superpixel segmentation are shown in Figures 5 and 6. Compared with the classic watershed and SLIC methods, our method has better applicability in low-amplitude, low-discrimination areas on the basis of reducing the effect of speckle noise, and the generated superpixel edges are more closely matched to the actual target.

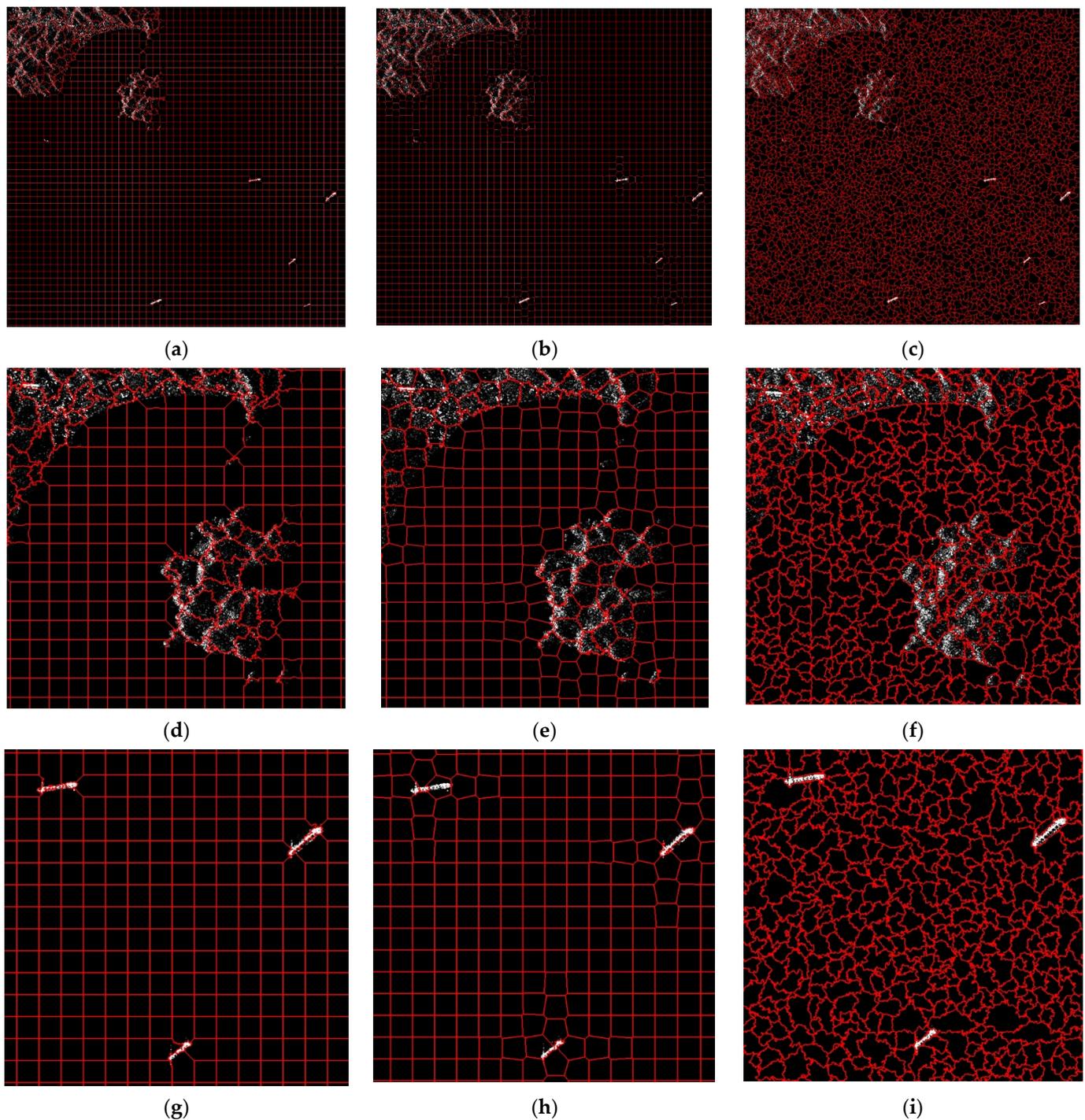


Figure 5. Comparison plot of superpixel segmentation results for nearshore feature map: (a–c) overall; (d–f) land area; (g–i) ship area; (a,d,g) watershed method; (b,e,h) SLIC method; (c,f,i) our method.

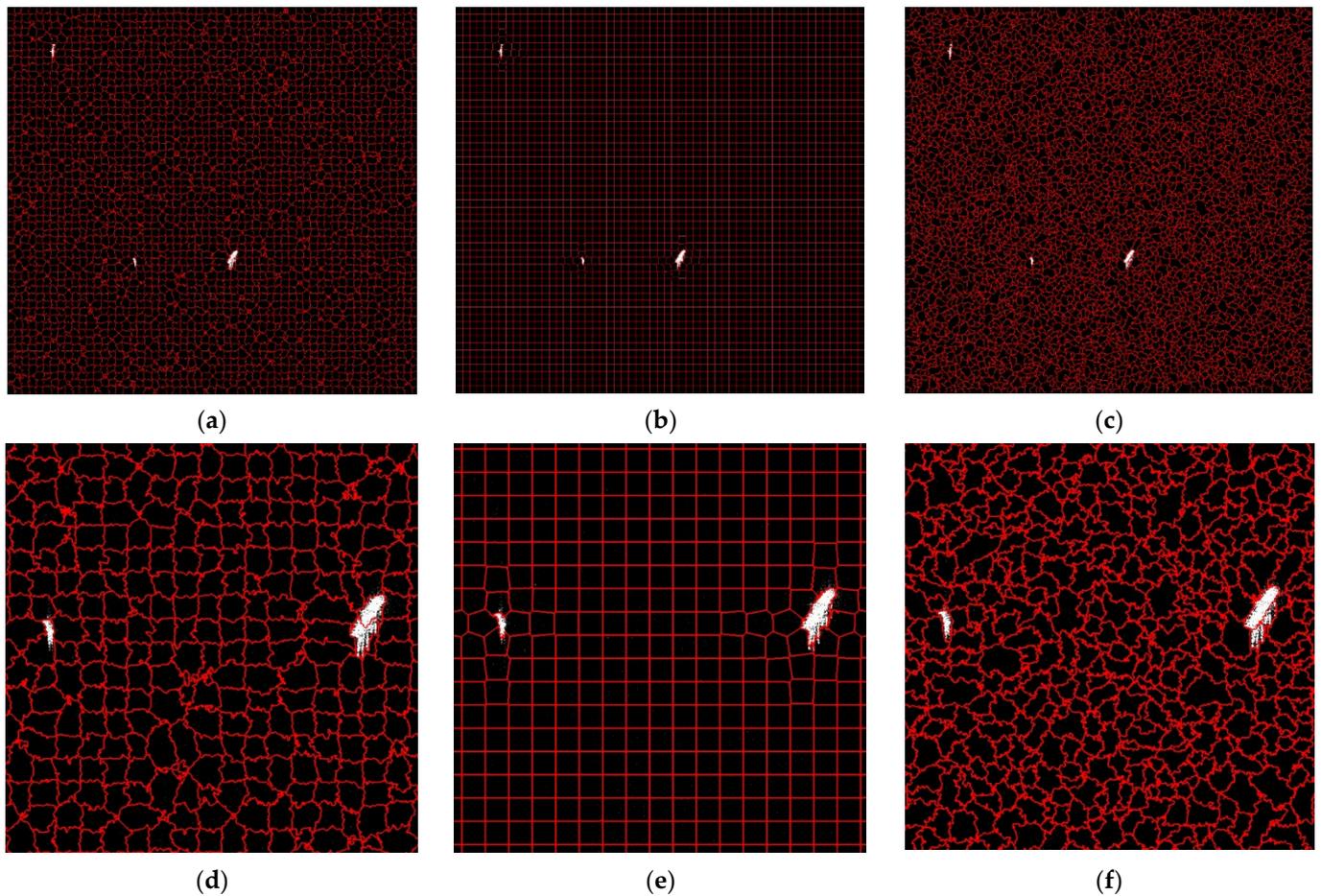


Figure 6. Comparison plot of superpixel segmentation results for distant ocean feature map: (a–c) overall; (d–f) ship area; (a,d) watershed method; (b,e) SLIC method; (c,f) our method.

2.3. Neighborhood Semantic Differences Extraction Based on LDA Bag-of-Words Topic Model

The topic model was originally applied in the field of text mining to extract semantic information implicit in the text. Latent Dirichlet Allocation (LDA) [21] is a bag-of-words-based topic model, so the order of words in a document can be disregarded. If an image is considered as a collection of pixels, then the image is a bag of words and the pixels are the words in it, so the LDA topic model can also be introduced into the field of computer vision (CV) [23]. LDA is based on a generative probabilistic model, with the core idea of learning a set of latent topics, and each document or image can be represented as a mixture of topics from that set. Therefore, after generating the bag of words via the superpixel method above, all superpixel blocks can be regarded as a set of documents, and pixels can be regarded as words in each document. By extracting the high-level semantic information implied by each superpixel block, i.e., the distribution of topics of that superpixel block, feature vectors are generated and classified to obtain superpixel blocks with the semantics of ships. This process is also a process of dimensionality reduction for features.

A sketch map of the LDA topic model is shown in Figure 7, where K is the number of topics; M is the number of documents, which is the number of superpixels in a polarimetric feature map; N_m is the number of words contained in the m th document, which is the number of pixels in the m th superpixel of the feature map; w_{mn} represents the value of the n th pixel in the m th superpixel; z_{mn} represents the topic of the n th pixel in the m th superpixel; θ_m represents the topic distribution of the m th superpixel; and β_k represents the pixel value distribution of the k th topic. Then, for this article, the pixel values in each superpixel are generated with the following process:

1. A certain topic is selected with a certain probability based on the topic distribution of the superpixel.
2. A certain pixel value is selected with a certain probability based on the word distribution of this topic, which is also the pixel value distribution.

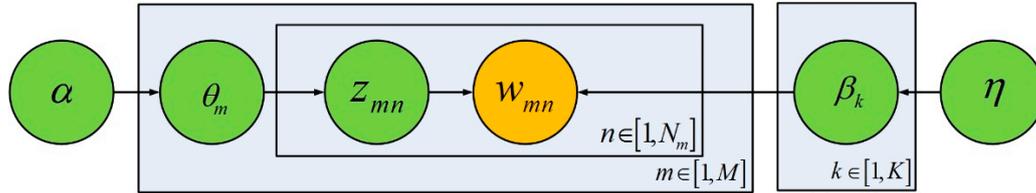


Figure 7. Sketch map of the LDA topic model.

The joint probability distribution function of this process is as follows:

$$p(\theta, z, w, \beta) = \left(\prod_{k=1}^K p(\beta_k | \eta) \right) \left(\prod_{m=1}^M p(\theta_m | \alpha) \prod_{n=1}^{N_m} p(z_{mn} | \theta_m) \prod_{m=1}^M p(w_{mn} | z_{mn}, \beta_k) \right), \quad (16)$$

where β_k follows the Dirichlet distribution with parameter η , θ_m follows the Dirichlet distribution with parameter α , z_{mn} follows the polynomial distribution with parameter θ_m , and w_{mn} follows the polynomial distribution with parameter $\beta_{z_{mn}}$. Repeat the above process to generate all superpixels and the whole feature map.

We solve the LDA parameters via the Gibbs sampling method [32], which is a special case of the Markov-Chain Monte Carlo algorithm. The core of this method is to randomly select a variable from the probability vector each time, sample the value of the current variable with the given value of other variables, and keep iterating until convergence, then output the parameters to be estimated.

After obtaining the topic distribution of all superpixels, the topic distribution vector of each superpixel is the semantic feature of that superpixel. Due to the spatial correlation between the ship targets and their backgrounds, although the superpixel segmentation strengthens the spatial correlation of the pixel-level features to a certain extent, at the semantic level, its spatial correlation still needs to be further strengthened.

For each superpixel, the superpixels adjacent to its boundary are its neighboring superpixels. Drawing on the LBP idea, the mean value of the difference between the topic distribution vector of a superpixel and the topic distribution vectors of all its neighboring superpixels is defined as the neighborhood topic distribution difference vector of the superpixel, which we call neighborhood semantic differences, and this value can be used as the neighborhood semantic feature of the superpixel. It is expressed as follows:

$$L(z_1, \dots, z_a) = \frac{1}{n} \sum_{i=1}^n \{ \theta_m(x_1, \dots, x_a) - \theta_{mi}(y_1, \dots, y_a) \}, \quad (17)$$

where $L(z_1, \dots, z_a)$, $\theta_m(x_1, \dots, x_a)$, and $\theta_{mi}(y_1, \dots, y_a)$ are the topic distribution vectors, and n is the number of neighborhood superpixels. For $n = 6$, a sketch map of the neighborhood structure is shown in Figure 8.

Replace the original semantic features of a superpixel with its neighborhood semantic features, making the extracted semantic features more spatially relevant. The superpixels with semantics of seawater and sea clutter previously obtained from prior knowledge need to be assigned corresponding feature vectors for easy calculation.

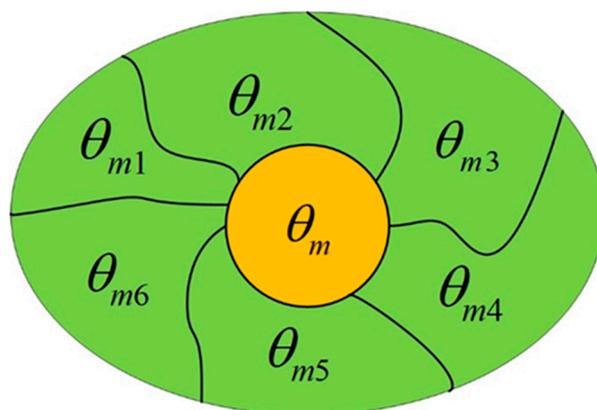


Figure 8. Sketch map of the neighborhood structure.

2.4. Ship Coarse Detection Based on SVM Classifier and Expert Knowledge Post-Processing

Nonlinear multi-classification of neighborhood semantic features uses Gaussian kernel function support vector machines (SVMs). Considering that in the polarimetric rotation domain feature map, the ship target has the highest pixel value and the seawater has the lowest pixel value, for a certain class of targets, the weighted values of the pixel values with the highest probability of the topic belonging to the first $(K' - 1)/K'$ of each component of the positive and negative parts of the neighborhood semantic feature vector of any of its superpixels are the positive and negative topic words of the class, respectively, where K' is the number of topics that have increased the number of prior semantic classes. After sorting the positive and negative topic words, if the positive topic word of the class is the highest value and not lower than 225, while the negative topic word of the class is the lowest value and not higher than 30, the superpixel with a linkage domain of not less than 20 pixels belonging to the class is a ship target; otherwise, the whole PolSAR image is considered to have no ship target.

When using the original semantic features of superpixels for classification, for a certain class of targets, the pixel value with the highest probability of the topic corresponding to the highest component of the original semantic feature vector of any of its superpixels is the topic word of the class. After sorting the topic words, if the topic word of the class is the highest value and not less than 250, the superpixel with a linkage domain of not less than 20 pixels belonging to the class is a ship target; otherwise, the whole PolSAR image is considered to have no ship target. The comparison results of classification using original semantic features and neighborhood semantic features are shown in Figures 9 and 10. The color of the superpixel in the figure indicates its semantics; red indicates the target whose semantic is the ship, and the color of the rest of the semantic targets is randomly assigned. As shown in Figure 9c, the superpixels in the black box mistakenly label the original land targets as ships when using the original semantic features for classification. As shown in Figure 9d, this misclassification is avoided when using neighborhood semantic features for classification.

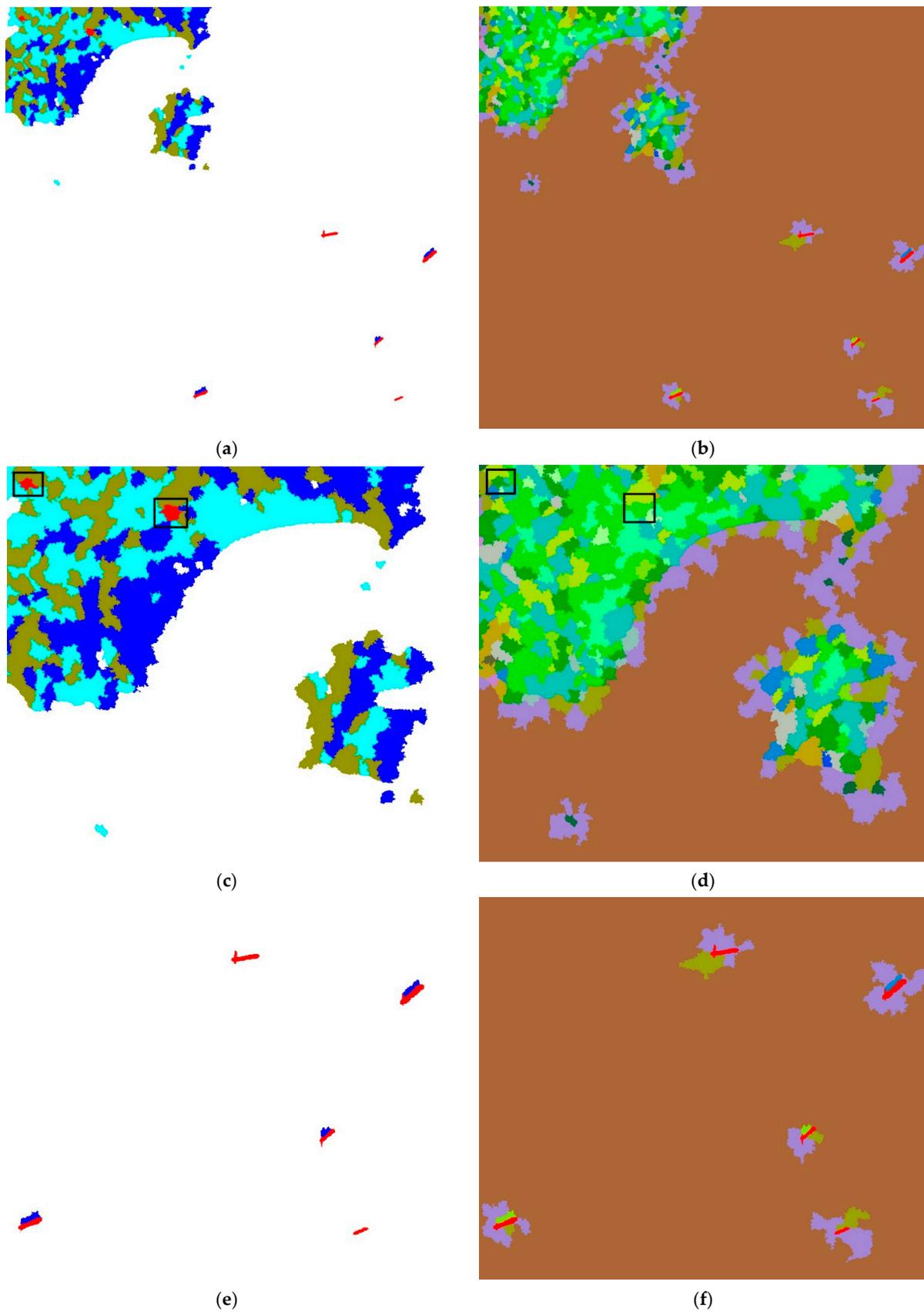


Figure 9. Comparison plot of classification results for nearshore feature map: (a,b) overall; (c,d) land area; (e,f) ship area; (a,c,e) using original semantic features; (b,d,f) using neighborhood semantic features.

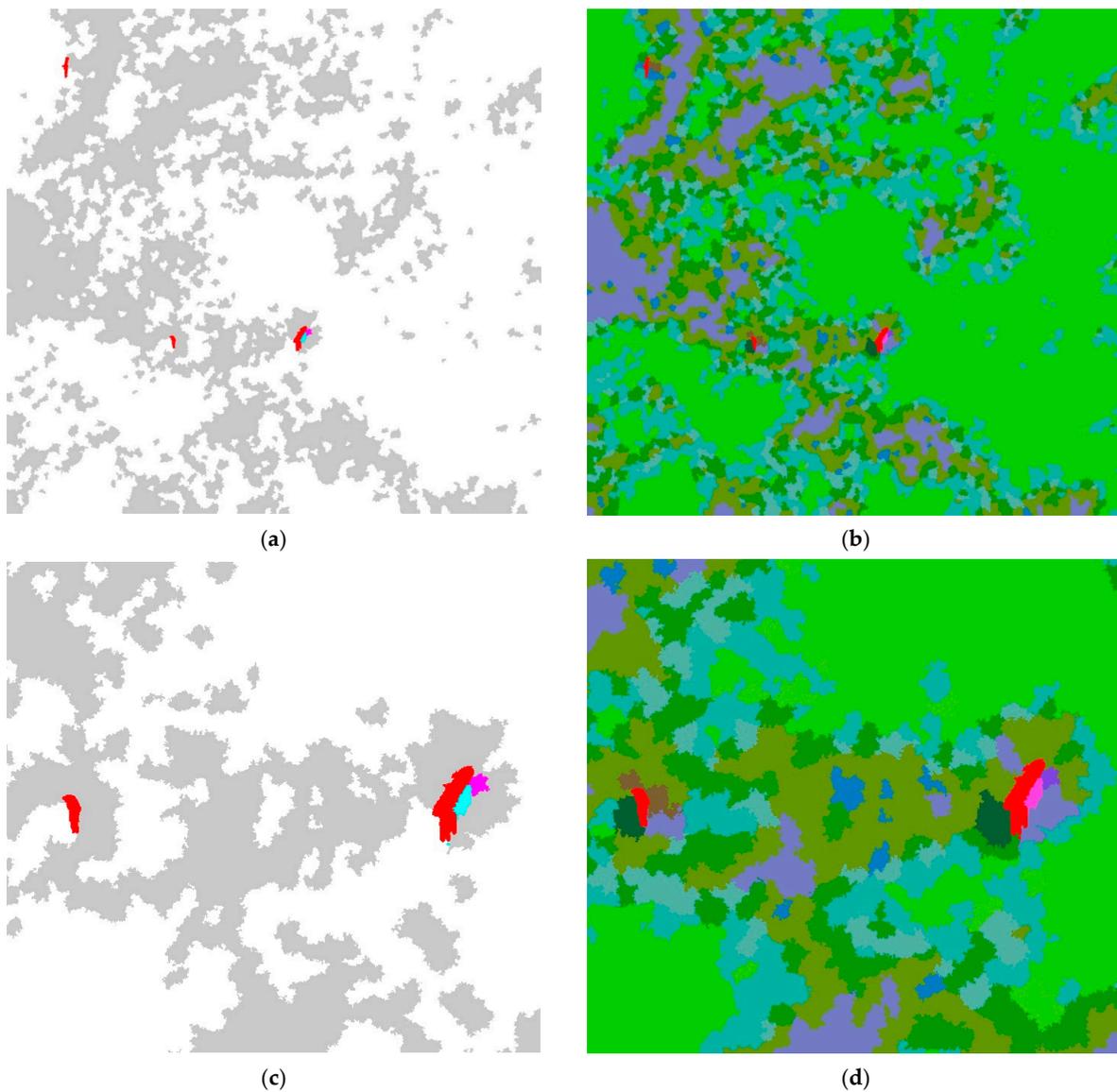


Figure 10. Comparison plot of classification results for distant ocean feature map: (a,b) overall; (c,d) ship area; (a,c) using original semantic features; (b,d) using neighborhood semantic features.

3. Results

In this section, comprehensive experiments are conducted to validate the effectiveness and demonstrate the superiority of the proposed method. Specifically speaking, the ship detection results of our proposed ship potential area extraction method are compared with those of two novel and one classical unsupervised ship detection methods.

3.1. Data Description

We use 10 full PolSAR images from the Chinese GF-3 satellite in nearshore and distant ocean scenes near Shanghai and Hong Kong for experiments, of which 5 are nearshore images and 5 are distant ocean images, to construct a dataset for ship potential area extraction. The GF-3 satellite is one of the civilian space-borne SAR systems, with 12 imaging modes, such as stripmap, spotlight, scanSAR, and so on, and the resolution can reach up to 1 m [33]. The 10 full PolSAR images used are obtained via the imaging mode of QPSI, and it has a spatial resolution of 8 m and observation swath of 30 km. The product level is L1A, which provides the complex data of images with HH, HV, VH, and VV polarizations.

3.2. Experimental Setup and Evaluation Index

We implement the proposed algorithm through python 3.6 and execute it on a 64-bit Ubuntu 20.04 workstation.

The evaluation indicator is the standard to measure the training effect of the model. In the process of our model's training and testing, *Precision* and *Recall* are mainly selected as the evaluation indicators. The combination of the sample real class and model prediction class is divided into four cases: true positive, false positive, true negative, and false negative. We denote them as *TP*, *FP*, *TN*, and *FN*, respectively. Obviously, $TP + FP + TN + FN = \text{total number of samples}$.

Precision means the ratio of actually positive examples in the examples divided into positive examples. *Precision* is defined as

$$Precision = \frac{TP}{TP + FP}. \quad (18)$$

Recall means the ratio of positive samples predicted as positive samples in the total positive samples, reflecting the comprehensiveness of the model's prediction of positive samples, which is defined as

$$Recall = \frac{TP}{TP + FN}. \quad (19)$$

For the ship potential area extraction task, we aim to improve the *Precision* as much as possible while ensuring a high *Recall*, so other evaluation indicators for target detection such as *mAP* are not used.

3.3. Comparison Experiments

We compare our method with two novel and one classical unsupervised PolSAR ship detection methods, including the adaptive polarimetric whitening filter truncated statistical CFAR (PWF-TS-CFAR) [5] method based on the background clutter distribution modeling, the scattering mechanism subspace projection (SP) [11] method based on polarimetric feature extraction, and the span polarimetric cross-entropy (SPCE) [34] method based on combining the polarimetric and span features. The recall and precision for nearshore and distant ocean images are shown in Table 4. Specifically, our method and the other three methods all achieve a recall of 1 in both the nearshore and distant ocean images. However, due to the influence of the complex background in the nearshore image, especially in the land area, the precision of the other three methods except ours drops significantly. The PWF-TS-CFAR method and the SP method still have a precision of only 0.086 and 0.131 after the addition of morphological filtering, and the precision of the SPCE method is 0.071. In the distant ocean image, the background is relatively simple, and the precision of the other three methods except ours is improved compared to the nearshore image, with 0.744, 0.865, and 0.865, respectively, but there is still a certain gap regarding the precision of our method. Our method achieves a precision of 0.96 and 0.97 in the nearshore and distant ocean images, respectively, demonstrating the applicability of our method to complex backgrounds.

Table 4. Ship detection results of different methods.

	Nearshore Images		Distant Ocean Images	
	Recall	Precision	Recall	Precision
PWF-TS-CFAR	1	0.086	1	0.744
SP	1	0.131	1	0.865
SPCE	1	0.071	1	0.865
Ours	1	0.96	1	0.97

The visualization of the detection results is shown in Figures 11–13, where Figure 11 shows the ground truth. Under the complex background interference, all three methods except ours experience a large number of false alarms. Specifically, in the nearshore image

(Figure 12a), the PWF-TS-CFAR method is sensitive to background targets, such as islands, defocusing, and azimuth ambiguity, and generates false alarms there. Due to the use of truncation to eliminate interference in the background window and the absence of a protective window, this method is suitable for dense ship detection, but it can also generate a large number of false alarms at the sea–land boundary. In the nearshore image (Figure 12b), the SP method detects the preset scattering mechanism through the subspace projection, and it is sensitive to the targets containing the preset scattering mechanism, generating a large number of false alarms in the land area, artificial targets on the land, and azimuth ambiguity, but it is not sensitive to defocusing. In the nearshore image (Figure 12c), the SPCE method increases the target and background differences by fusing the span feature, which is sensitive to high-energy backscattering and azimuth ambiguity, and generates false alarms, but it is not sensitive to defocusing. In the distant ocean images (Figure 13a–c), all three methods except ours cause false alarms due to azimuth ambiguity. In addition, our method uses numerical truncation to deal with the high-amplitude information in the polarimetric rotation domain feature maps, which results in very few cases where defocusing separated from ships is classified as ship targets during semantic extraction, resulting in false alarms. In summary, the comparison experiment results validate the effectiveness and demonstrate the superiority of our method in the PolSAR ship potential area extraction task with complex backgrounds.

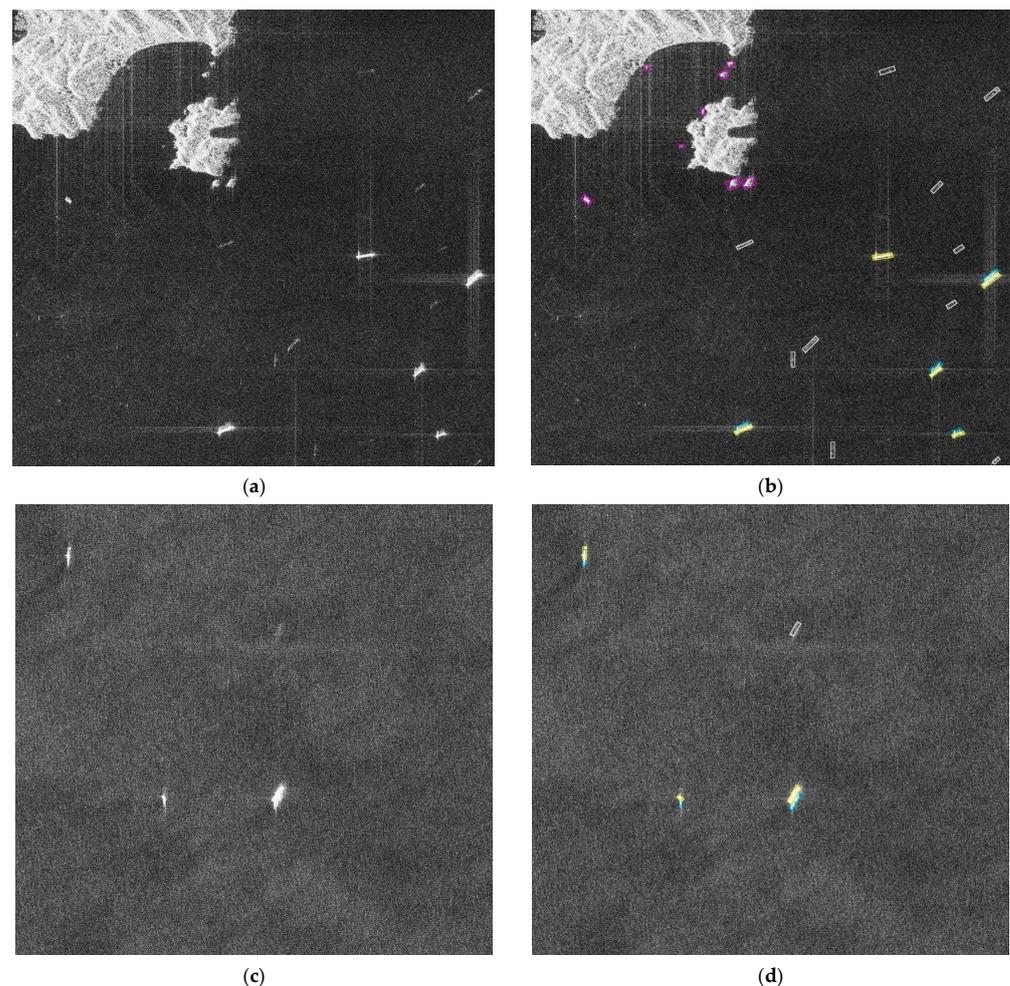


Figure 11. The ground-truth data, and the yellow rectangle represents a ship, the blue rectangle represents defocusing, the white rectangle represents azimuth ambiguity, and the purple rectangle represents an island: (a,b) the nearshore images; (c,d) the distant ocean images; (a,c) the HV channel of original PolSAR images; (b,d) the ground-truth maps.

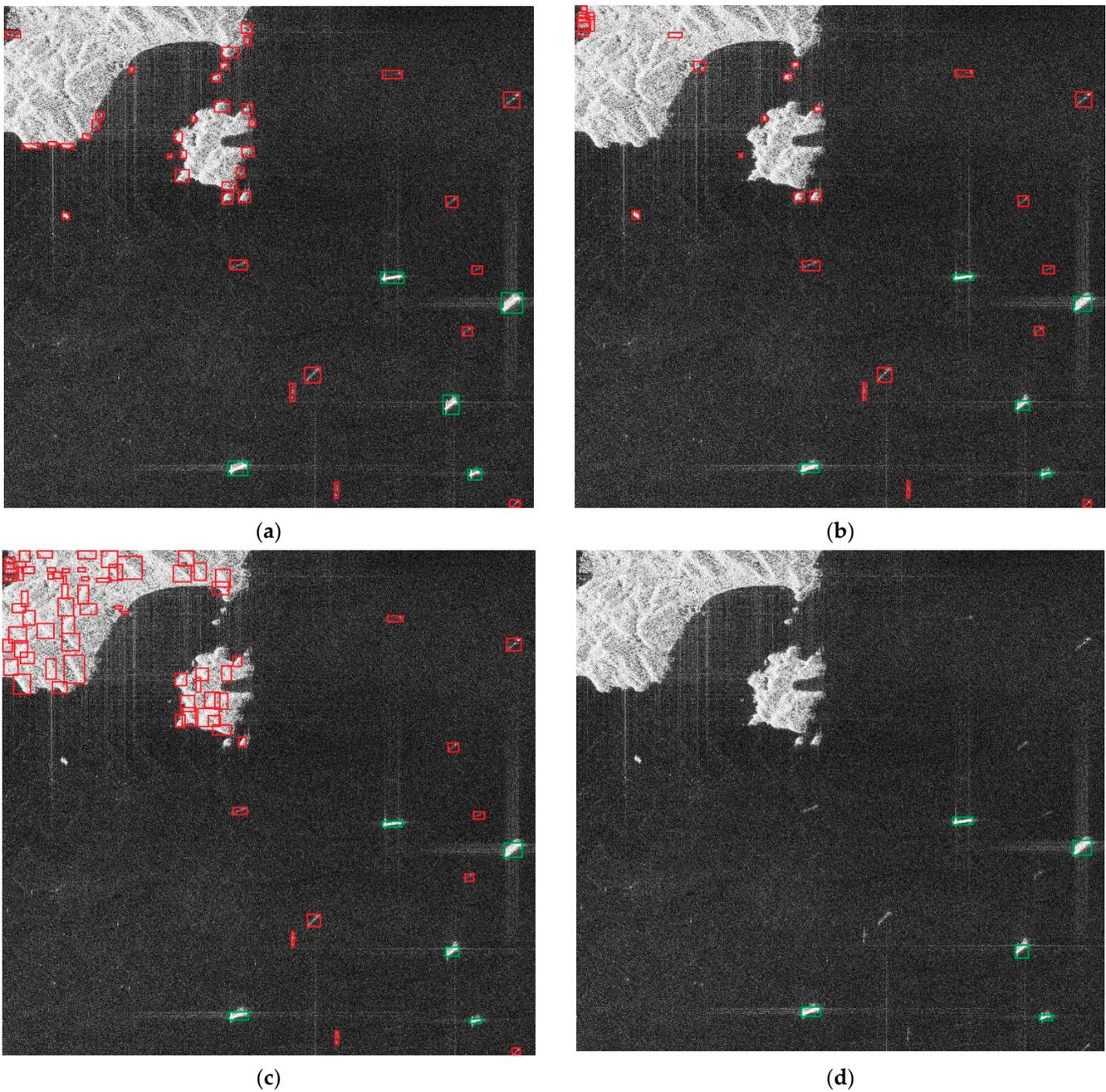


Figure 12. Comparison plot of ship detection results for nearshore image, and the green rectangle represents true positive, and the red rectangle represents false alarm: (a) PWF-TS-CFAR method; (b) SP method; (c) SPCE method; (d) our method.

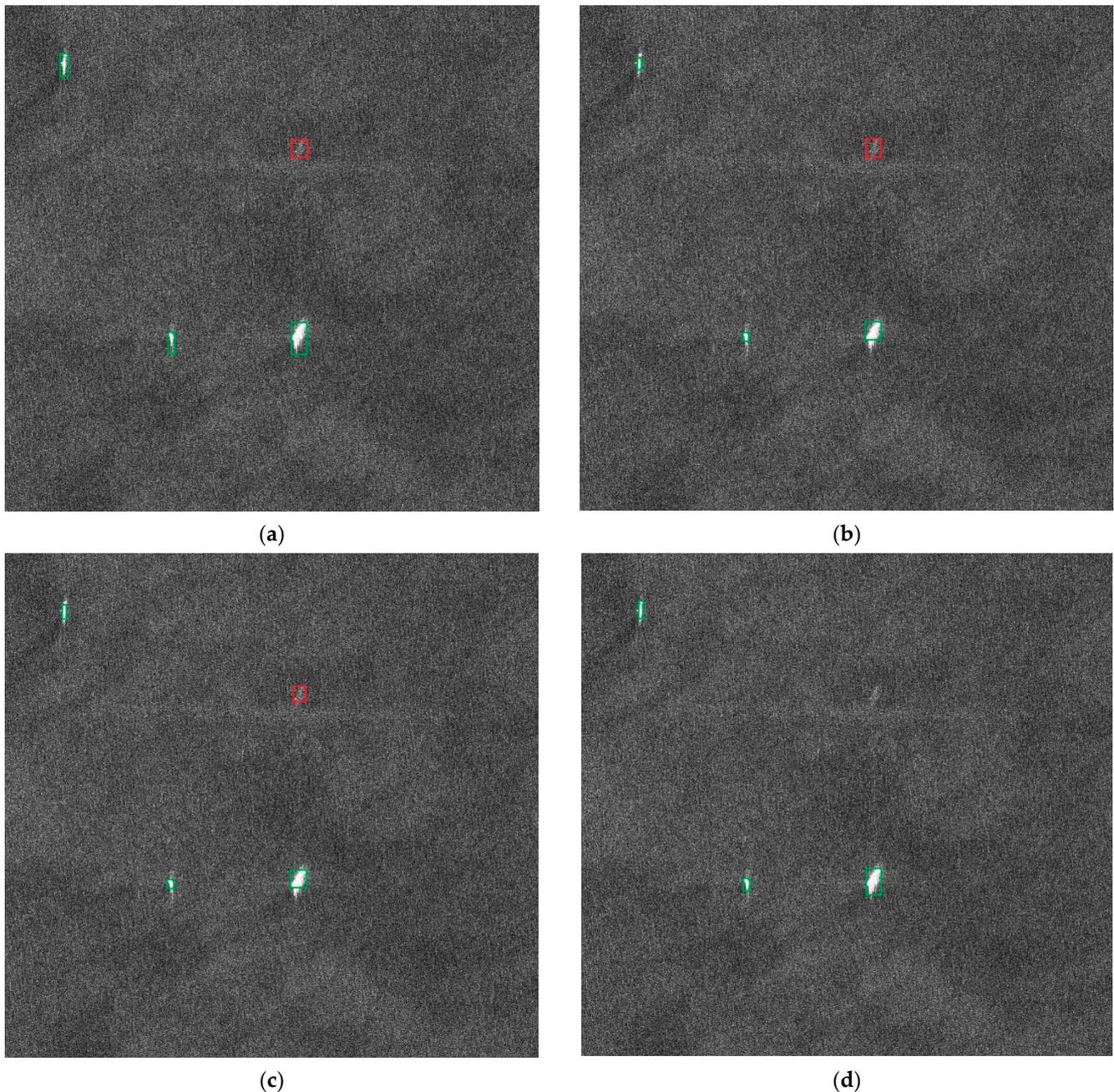


Figure 13. Comparison plot of ship detection results for distant ocean image, and the green rectangle represents true positive, and the red rectangle represents false alarm: (a) PWF-TS-CFAR method; (b) SP method; (c) SPCE method; (d) our method.

4. Discussion

Compared with the other three unsupervised PolSAR ship detection methods, our proposed method improves the detection ability of ships in complex backgrounds, but further work is still needed to improve detection performance. In addition, we analyze and discuss the effectiveness of several steps of our method.

Firstly, we replace the polarimetric feature we selected in the polarimetric rotation domain feature selection step with three other polarimetric features in order of decreasing classification weights, and we construct polarimetric feature maps to test the impact of polarimetric feature selection on the detection results, which are shown in Table 5. The method for constructing the correlation feature map is described in Section 2.1.2. When con-

structuring coherence feature maps, the original coherence features are multiplied by 256 and then rounded down and stretched to construct 8-bit feature maps for easy comparison. It can be observed that feature $|\hat{\gamma}_{HH-HV}(0)|$, which has the second-highest classification weights after the feature we selected, has exactly the same detection results as the feature we selected, whereas features $\max\{|\gamma_{HH-VV}(\theta)|\}$ and $\max\{|\gamma_{(HH-VV)-HV}(\theta)|\}$ have a lower precision in their detection results compared to the precision of the feature we selected. Therefore, both features $|\hat{\gamma}_{(HH-VV)-HV}(0)|$ and $|\hat{\gamma}_{HH-HV}(0)|$ can be used for polarimetric feature map construction.

Table 5. The effect of polarimetric rotation domain features selection on ship detection results.

	Nearshore Images		Distant Ocean Images	
	Recall	Precision	Recall	Precision
$\max\{ \gamma_{HH-VV}(\theta) \}$	1	0.848	1	0.914
$\max\{ \gamma_{(HH-VV)-HV}(\theta) \}$	1	0.805	1	0.897
$ \hat{\gamma}_{HH-HV}(0) $	1	0.96	1	0.97
$ \hat{\gamma}_{(HH-VV)-HV}(0) $ (Ours)	1	0.96	1	0.97

Secondly, we replace the superpixel segmentation method in the bag-of-words generation step based on superpixel segmentation with watershed and SLIC methods in order to test the effect of bag-of-words generation methods on the detection results, which are shown in Table 6. It can be observed that when the watershed and SLIC methods are used for superpixel segmentation, the recall in the detection results is significantly reduced. The core of ship potential area extraction is to improve the precision as much as possible while ensuring a high recall; therefore, adopting our proposed superpixel segmentation method applicable to the constructed polarimetric feature maps is a key step in ship potential area extraction.

Table 6. The effect of bag-of-words generation based on superpixel segmentation on ship detection results.

	Nearshore Images		Distant Ocean Images	
	Recall	Precision	Recall	Precision
Watershed method	0.684	0.942	0.698	0.957
SLIC method	0.432	0.953	0.458	0.978
Our method	1	0.96	1	0.97

Thirdly, we replace our method with the original LDA method in the semantic features extraction step to test the effect of semantic extraction on the detection results, which are shown in Table 7. It can be observed that our method outperforms the original LDA method in terms of precision in complex backgrounds, but in simple backgrounds, our method is consistent with the detection results of the original LDA method, proving the effectiveness of our method in complex backgrounds.

Table 7. The effect of semantic features extraction on ship detection results.

	Nearshore Images		Distant Ocean Images	
	Recall	Precision	Recall	Precision
Original LDA	1	0.792	1	0.97
NSD-LDA (Ours)	1	0.96	1	0.97

In order to better demonstrate the effect of our proposed three steps on the overall performance improvement, we demonstrate the superiority and effectiveness of our proposed method by verifying the effect of different combinations of steps on the ship potential

area extraction results in Table 8. In this case, Baseline consists of the polarimetric feature $\max\left\{\left|\gamma_{(HH-VV)-HV}(\theta)\right|\right\}$, SLIC superpixel segmentation, and original LDA topic model. It can be observed that the detection effect improved by selecting polarimetric features that maximize the difference between ships and backgrounds is related to the representation ability of the selected feature. The improved superpixel segmentation method suitable for low-amplitude and speckle noise areas is of great significance in ensuring high recall. The NSD-LDA method improves the detection results of complex backgrounds more than simple backgrounds.

Table 8. The effect of different step combinations on ship detection results.

Method	Features Selection (FS)	Improved Superpixel Segmentation (ISS)	NSD-LDA (NL)	Nearshore Images		Distant Ocean Images	
				Recall	Precision	Recall	Precision
Baseline				0.392	0.826	0.424	0.985
Baseline + FS	✓			0.432	0.807	0.458	0.978
Baseline + ISS		✓		1	0.653	1	0.897
Baseline + NL			✓	0.392	0.958	0.424	0.985
Baseline + FS + ISS	✓	✓		1	0.792	1	0.97
Baseline + FS + NL	✓		✓	0.432	0.953	0.458	0.978
Baseline + ISS + NL		✓	✓	1	0.805	1	0.897
Our method	✓	✓	✓	1	0.96	1	0.97

Finally, in the LDA bag-of-words topic model, the preset hyperparameter K has a certain effect on the semantic extraction results, and after experiments, the value of K in our method is set to 10. The LDA topic model obtains the latent topic distribution information, and the detection effect is the best when the number of preset topics matches the scene. Otherwise, fewer topics bring about the phenomenon of synonymy, and more topics bring about the phenomenon of polysemy, which both reduce the accuracy of the semantic extraction. The hyperparameters need to be adjusted manually and cannot be given automatically, which is a limitation of our method. In addition, our proposed superpixel method differs from SLIC and watershed methods in that it does not pre-set the number of superpixels, and the generated superpixel scale is determined with the constructed polarimetric rotation domain feature map. When the number of pixels of an initial superpixel is less than a threshold, it will be merged with the neighborhood superpixels with the smallest dissimilarity, making it difficult to avoid small ships with scales smaller than the threshold being submerged. This is another limitation of our method. The core idea of this work is to extract the ship potential areas without relying on the labeled samples and using unsupervised methods under the scattering diversity and complex background conditions, so as to reduce the computational complexity of polarimetric and spatial feature extraction of the subsequent deep learning-based PolSAR ship fine detection.

5. Conclusions

In this article, a PolSAR ship potential area extraction (coarse detection) method based on neighborhood semantic differences of an LDA bag-of-words topic model is proposed. Based on polarimetric rotation domain feature selection and feature map construction, superpixel segmentation-based bag-of-words generation applicable to polarimetric rotation domain feature maps and high-level semantic feature extraction based on neighborhood semantic differences of the LDA bag-of-words topic model are applied to achieve ship potential area extraction capability under complex background conditions. Comparison experiments were conducted on the GF-3 dataset. The experimental results show the superiority of our proposed method over other unsupervised PolSAR ship detection methods under complex background conditions. The effectiveness of our proposed method was verified by discussing the effects of each step of our proposed method on the detection results.

We will continue to study how to better perform PolSAR ship potential area extraction in our future work.

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

Funding: This research received no funding.

Data Availability Statement: Data are contained within the article.

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

References

1. Moreira, A.; Iraola, P.; Younis, M.; Krieger, G.; Hajnsek, I.; Papathanassiou, K.P. A Tutorial on Synthetic Aperture Radar. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–43. [[CrossRef](#)]
2. Touzi, R.; Boerner, W.M.; Lee, J.S.; Lueneburg, E. A Review of Polarimetry in the Context of Synthetic Aperture Radar: Concepts and Information Extraction. *Can. J. Remote Sens.* **2004**, *30*, 380–407. [[CrossRef](#)]
3. Qiu, W.; Pan, Z.; Yang, J. Few-Shot PolSAR Ship Detection Based on Polarimetric Features Selection and Improved Contrastive Self-Supervised Learning. *Remote Sens.* **2023**, *15*, 1874. [[CrossRef](#)]
4. Leng, X.; Ji, K.; Yang, K.; Zou, H. A Bilateral CFAR Algorithm for Ship Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *7*, 1536–1540. [[CrossRef](#)]
5. Liu, T.; Yang, Z.; Marino, A.; Gao, G.; Yang, J. Robust CFAR Detector Based on Truncated Statistics for Polarimetric Synthetic Aperture Radar. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6731–6747. [[CrossRef](#)]
6. Cloude, S.R.; Pottier, E. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 68–78. [[CrossRef](#)]
7. Freeman, A.; Durden, S.L. A three-component scattering model for polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 963–973. [[CrossRef](#)]
8. Huynen, J.R. The stokes matrix parameters and their interpretation in terms of physical target properties. *Proc. SPIE* **1990**, *1317*, 195–207.
9. Ringrose, R.; Harris, N. Ship Detection Using Polarimetric SAR Data. In Proceedings of the CEOS SAR Workshop, Toulouse, France, 26–29 October 1999.
10. Touzi, R.; Charbonneau, F.; Hawkins, R.K.; Murnaghan, K.; Kavoun, X. Ship-Sea Contrast Optimization When Using Polarimetric SARs. In Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium (IGARSS), Sydney, Australia, 9–13 July 2001.
11. Bordbari, R.; Maghsoudi, Y. A New Target Detector Based on Subspace Projections Using Polarimetric SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 3025–3039. [[CrossRef](#)]
12. De Grandi, G.D.; Lee, J.; Schuler, D.L. Target Detection and Texture Segmentation in Polarimetric SAR Images Using a Wavelet Frame: Theoretical Aspects. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3437–3453. [[CrossRef](#)]
13. Chen, S.; Li, Y.; Wang, X.; Xiao, S.; Sato, M. Modeling and Interpretation of Scattering Mechanisms in Polarimetric Synthetic Aperture Radar: Advances and perspectives. *IEEE Signal Proc. Mag.* **2014**, *31*, 79–89. [[CrossRef](#)]
14. Chen, S.; Wang, X.; Xiao, S.; Sato, M. *Target Scattering Mechanism in Polarimetric Synthetic Aperture Radar: Interpretation and Application*; Springer: Singapore, 2018; pp. 1–225.
15. Bahmanyar, R.; Cui, S.; Datcu, M. A Comparative Study of Bag-of-Words and Bag-of-Topics Models of EO Image Patches. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1357–1361. [[CrossRef](#)]
16. Sivic, J.; Zisserman, A. Video Google: A text retrieval approach to object matching in videos. In Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV), Nice, France, 13–16 October 2003.
17. Yuan, J.; Wu, Y.; Yang, M. Discovery of Collocation Patterns: From Visual Words to Visual Phrases. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, MN, USA, 17–22 June 2007.
18. Deerwester, S.; Dumais, S.T.; Furnas, G.W.; Landauer, T.K.; Harshman, R. Indexing by latent semantic analysis. *J. Assoc. Inf. Sci. Technol.* **1990**, *41*, 391–407. [[CrossRef](#)]
19. Hofmann, T. Probabilistic latent semantic indexing. In Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Berkeley, CA, USA, 15–19 August 1999; pp. 50–57.
20. Bosch, A.; Zisserman, A.; Muñoz, X. Scene Classification Via pLSA. In Proceedings of the 9th European Conference on Computer Vision (ECCV), Graz, Austria, 7–13 May 2006.
21. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
22. Li, F.; Perona, P. A Bayesian hierarchical model for learning natural scene categories. In Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–25 June 2005.

23. Zhong, Y.; Huang, R.; Zhao, J.; Zhao, B.; Liu, T. Aurora Image Classification Based on Multi-Feature Latent Dirichlet Allocation. *Remote Sens.* **2018**, *10*, 233. [[CrossRef](#)]
24. Chen, S.; Wang, X.; Sato, M. Uniform polarimetric matrix rotation theory and its applications. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4756–4770. [[CrossRef](#)]
25. Chen, S. Polarimetric Coherence Pattern: A Visualization and Characterization Tool for PolSAR Data Investigation. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 286–297. [[CrossRef](#)]
26. Cui, X.C.; Tao, C.S.; Su, Y.; Chen, S.W. PolSAR Ship Detection Based on Polarimetric Correlation Pattern. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 471–475. [[CrossRef](#)]
27. Kononenko, F. Estimating attributes: Analysis and extensions of RELIEF. In Proceedings of the European Conference on Machine Learning (ECML), Catania, Italy, 6–8 April 2021.
28. Nghiem, S.V.; Yueh, S.H.; Kwok, R.; Li, F.K. Symmetry Properties in Polarimetric Remote Sensing. *Radio Sci.* **1992**, *27*, 693–711. [[CrossRef](#)]
29. Meyer, I. Estimating attributes: Color image segmentation. In Proceedings of the 1992 International Conference on Image Processing and Its Applications, Maastricht, The Netherlands, 7–9 April 1992.
30. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Susstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)]
31. Shui, P.; Cheng, D. Edge Detector of SAR Images Using Gaussian-Gamma-Shaped Bi-Windows. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 846–850. [[CrossRef](#)]
32. Geman, S.; Geman, D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 721–741. [[CrossRef](#)] [[PubMed](#)]
33. Pang, D.; Pan, C.; Zi, X. GF-3: The Watcher of the Vast Territory. *Aerosp. China* **2016**, *9*, 8–12.
34. You, B.; Yang, J.; Yeh, C.; Song, J. Improved ship detection method based on span polarimetric cross entropy. *J. Tsinghua Univ. Sci. Technol.* **2014**, *54*, 453–457.

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