

Article

Ship Detection for PolSAR Images via Task-Driven Discriminative Dictionary Learning

Huiping Lin ¹, Hang Chen ¹, Hongmiao Wang ¹, Junjun Yin ^{1,2} and Jian Yang ^{1,*}

¹ Department of Electronic Engineering, Tsinghua University, Beijing 100084, China; linhp15@mails.tsinghua.edu.cn (H.L.); h-chen15@mails.tsinghua.edu.cn (H.C.); wanghm17@mails.tsinghua.edu.cn (H.W.); yinjj07@gmail.com (J.Y.)

² School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

* Correspondence: yangjian_ee@mail.tsinghua.edu.cn; Tel.: +86-10-6279-4726

Received: 13 March 2019; Accepted: 27 March 2019; Published: 29 March 2019



Abstract: Ship detection with polarimetric synthetic aperture radar (PolSAR) has received increasing attention for its wide usage in maritime applications. However, extracting discriminative features to implement ship detection is still a challenging problem. In this paper, we propose a novel ship detection method for PolSAR images via task-driven discriminative dictionary learning (TDDDL). An assumption that ship and clutter information are sparsely coded under two separate dictionaries is made. Contextual information is considered by imposing superpixel-level joint sparsity constraints. In order to amplify the discrimination of the ship and clutter, we impose incoherence constraints between the two sub-dictionaries in the objective of feature coding. The discriminative dictionary is trained jointly with a linear classifier in task-driven dictionary learning (TDDL) framework. Based on the learnt dictionary and classifier, we extract discriminative features by sparse coding, and obtain robust detection results through binary classification. Different from previous methods, our ship detection cue is obtained through active learning strategies rather than artificially designed rules, and thus, is more adaptive, effective and robust. Experiments performed on synthetic images and two RADARSAT-2 images demonstrate that our method outperforms other comparative methods. In addition, the proposed method yields better shape-preserving ability and lower computation cost.

Keywords: ship detection; task-driven discriminative dictionary learning (TDDDL); contextual information; incoherence constraints; polarimetric synthetic aperture radar (PolSAR)

1. Introduction

Ship detection with synthetic aperture radar (SAR) images is one of the important applications in the field of maritime surveillance [1]. Recently, polarimetric synthetic aperture radar (PolSAR) ship detection has received increasing attention, as polarimetric information has proved to be of great benefit to improving the detection effect [2–12]. As a simple example, we can achieve satisfactory results at steep and middle (20° to 40°) incidence angles by using cross-polarization (HV) only [2]. However, co-polarization (HH or VV) may perform better at bigger incidence angles [2]. PolSAR images combine the advantages of all polarimetric channels, reveal the scattering characteristics differences of the ship and clutter, and help to improve detection effect.

In the most of the existing PolSAR ship detection methods, a scalar feature index is designed to discriminate the target and clutter at first, and then constant false alarm rate (CFAR) operation is conducted. The simplest feature index is the image span, which is the intensity of PolSAR image and defined as the square of the scattering matrix's Frobenius norm. With further research, more complicated features have been proposed, and these features can be roughly classified into

two types. The first is designed by enhancing the contrast between the interested targets and clutter. Novak et al. proposed the polarimetric whitening filter (PWF) to produce a speckle-reduced image by optimally combining all the elements of the scattering matrix [3]. Yang et al. presented the generalized optimization of polarimetric contrast enhancement (GOPCE) to maximize the signal-to-clutter ratio (SCR) in the image [4]. These methods work well in high SCR condition. However, when the SCR decreased, these methods may suffer from severe performance deterioration. The other type is designed by analyzing polarimetric scattering mechanism and introducing polarimetric parameters. Jeremy et al. implemented ship detection by using Cameron decomposition [5], while the symmetric scattering characterization method (SSCM) was developed by Touzi et al. [6]. Since the methods applied to single-look scattering matrix are generally more susceptible to the speckle and increase the probability of false alarms (PFAs) of small ship, multi-look covariance or coherency matrix-based methods have been explored further. Chen et al. introduced polarization cross entropy (PCE) based on the eigen-decomposition of polarimetric coherence matrix [7]. Moreover, degree of polarization [8] was fully investigated for ship detection. It is true that speckle is greatly reduced by spatial ensemble averaging in these methods. Nevertheless, although the scalar feature implicitly includes the contributions of all polarimetric channels, an explicit consideration of all the polarimetric channels should provide more information, which is not fully exploited. These design features are too simple to provide robust performance in complicated and changeable clutter conditions.

Among various detection algorithms, the constant false alarm rate (CFAR) detectors have been widely used for their simplicity and adaptive ability [9–11]. With the development of superpixel algorithms, some superpixel-based attempts have been carried out to achieve ship detection by combining the superpixel and CFAR detectors [12,13]. It was proved that the superpixel can help to retain the target outline and suppress speckle noise. However, only simple features of superpixels, such as entropy information and pixel intensity, were utilized in these methods. Simple features provide weak discrimination and depend on artificial design. In addition, the detection performance also depends on largely the accuracy of statistical modeling and parameter estimation in CFAR operation. The theoretical distributions of artificial features are usually analytically intractable, or the estimations of the distributions are extremely cumbersome.

With the development of deep learning, researchers have employed deep neural networks to achieve ship detection in PolSAR images. Zhou et al. [14] modified the faster region-based convolutional neural network (Faster-RCNN) and applied it to PolSAR ship detection. And Kang et al. [15] proposed contextual region-based convolutional neural network with multilayer fusion (CRCNN-MF) by combining contextual information, multi-scaling and region-based convolutional neural network (RCNN). However, these methods simply process each channel of the PolSAR images separately and finally fuse the results. And just like other deep neural network methods, the heavy computation burden, unstable convergence and lots of sensitive parameters are the bottleneck for the application of these methods above.

In this paper, we propose a novel ship detection method for PolSAR images via task-driven discriminative dictionary learning (TDDDL). The superpixel is utilized as the basic processing cell. Ship detection can be viewed as a binary classification problem at superpixel level. Task-driven dictionary learning (TDDL) methods have achieved a cynosure success in classification field [16,17]. To improve the discrimination between the ship and clutter, we propose to learn category-specific dictionaries for the ship and clutter. In this way, incoherence between sub-dictionaries is enhanced, producing more discriminative features. Contextual information is also considered by imposing joint sparsity prior. The complete dictionary is trained in TDDL framework. The proposed dictionary learning scheme is called TDDDL due to the strong discriminability of the learnt dictionary. Experimental results on synthetic images and two real-scene images show that our method outperforms all the comparative methods.

The main contributions of this paper can be summarized as follows: (1) We propose a novel dictionary learning algorithm to obtain more discriminative features and boost the detection

performance. Contextual information and incoherence constraints are all included in the algorithm. (2) We also describe an optimization procedure for solving sparse recovery problem with TDDDL. (3) Different from previous methods, the proposed ship detection method based on TDDDL employs active learning strategies rather than artificially designed rules, and thus, is more adaptive and effective. In addition, the strong discriminability of the learnt dictionary improves detection performance further.

The remainder of this paper is organized as follows. In Section 2, we propose TDDDL in detail, including the formulation and optimization. In Section 3, the complete scheme of the proposed ship detection method is given. We conduct extensive experiments to evaluate the proposed method in Section 4, and conclude our work and propose future work in Section 5.

2. Task-Driven Discriminative Dictionary Learning (TDDDL)

In this section, we first briefly revisit the task-driven dictionary learning (TDDL) [16]. Then, we will propose our TDDDL, including its formulation and optimization. We use $[Q_1; Q_2]$ to denote the vertical concatenation of two matrices with the same columns, and use $[Q_1, Q_2]$ to denote the horizontal concatenation of two matrices with the same rows.

2.1. Review of TDDL

In TDDL [16], signals are represented by their sparse codes, which are then fed into a linear regression. Consider a pair of training samples (x, y) , where $x \in \mathbb{R}^M$ is the sum feature extracted from PolSAR image, $y \in \mathbb{R}^K$ is a binary vector representation of corresponding label, M and K denote the dimensions of x and y , respectively. Given some dictionary $\mathbf{D} \in \mathbb{R}^{M \times P}$, where P is the number of atoms in the dictionary \mathbf{D} , x can be represented as a sparse vector $\alpha(x, \mathbf{D}) \in \mathbb{R}^P$, defined as the solution of an elastic-net problem [17]:

$$\alpha(x, \mathbf{D}) = \arg \min_{\alpha \in \mathbb{R}^P} \frac{1}{2} \|x - \mathbf{D}\alpha\|_2^2 + \lambda_1 \|\alpha\|_1 + \frac{\lambda_2}{2} \|\alpha\|_2^2 \quad (1)$$

where λ_1 and λ_2 are the regularization parameters.

For classification task, TDDL uses the sparse vector $\alpha(x, \mathbf{D})$ in a classical expected risk minimization formulation:

$$\min_{\mathbf{D}, \mathbf{W}} L(\mathbf{D}, \mathbf{W}, x) = \min_{\mathbf{D}, \mathbf{W}} f(\mathbf{D}, \mathbf{W}, x) + \frac{\mu}{2} \|\mathbf{W}\|_F^2 \quad (2)$$

where $L(\mathbf{D}, \mathbf{W}, x)$ is classification risk, \mathbf{W} is the parameter matrix of the classifier, μ is a classifier regularization parameter to avoid the overfitting of classifier [18], and $f(\mathbf{D}, \mathbf{W}, x)$ is a convex function defined as

$$f(\mathbf{D}, \mathbf{W}, x) = E_{y,x} [l_s(y, \mathbf{W}, \alpha(x, \mathbf{D}))]. \quad (3)$$

In this equation, $E_{y,x}$ denotes the expectation taken relative to the probability distribution $p(x, y)$, l_s is a convex loss function that measures how well one can predict y by observing $\alpha(x, \mathbf{D})$ given the parameter matrix \mathbf{W} , which can be the square, logistic, or hinge loss from SVM [19].

Stochastic gradient descent (SGD) algorithm is used to update the dictionary \mathbf{D} and the parameter matrix \mathbf{W} . The update rules are as follows.

$$\begin{cases} \mathbf{D}^{t+1} = \mathbf{D}^t - \rho^t \cdot \frac{\partial L^t}{\partial \mathbf{D}} \\ \mathbf{W}^{t+1} = \mathbf{W}^t - \rho^t \cdot \frac{\partial L^t}{\partial \mathbf{W}} \end{cases} \quad (4)$$

where t is the iteration index and ρ is the step size. The equation for updating \mathbf{W} is straightforward since $L(\mathbf{D}, \mathbf{W}, x)$ is both smooth and convex with respect to \mathbf{W} . We have

$$\frac{\partial L}{\partial \mathbf{W}} = (\mathbf{W}\alpha - y)\alpha^T + \mu\mathbf{W} \quad (5)$$

where T denotes the transposition. According to the chain rule, we have

$$\frac{\partial L}{\partial \mathbf{D}} = \frac{\partial L}{\partial \alpha} \cdot \frac{\partial \alpha}{\partial \mathbf{D}}. \tag{6}$$

The main difficulty comes from $\partial \alpha / \partial \mathbf{D}$, since the optimization problem in Equation (1) is not smooth [20]. Mairal et al. [16] use fixed point differentiation to solve the problem [21]. The detailed derivation of the algorithm can be found in the Appendix of Mairal et al. [16].

2.2. Formulation of TDDDL

The TDDL method provides us with a supervised dictionary learning framework to learn dictionaries adapted to various tasks instead of being only adapted to data reconstruction [16]. It works well for basic-level classification because the differences between categories are typically rather significant. The sparse codes of different categories are different and result in discriminative features for classification. However, when facing harsher classification tasks, where the signal to clutter ratio (SCR) is much lower and different categories show similar characteristics, the TDDL method would suffer from severe performance deterioration. The difference can be dominated by those similar sparse codes and even disappear at the feature encoding stage. Hence, it is desirable to find a dictionary that could encode the features of different categories with their own code-words. Such a dictionary would obviously boost the differences of the feature representations, and improve the consequent ship detection.

To this end, we propose learning a discriminative dictionary by using category-specific dictionary structure and imposing incoherence constraints between the sub-dictionaries. Since neighboring pixels often share the same label with high probability, contextual information is also considered via joint sparsity prior. The complete dictionary is trained jointly with a linear classifier in TDDL framework. We call the proposed dictionary learning scheme TDDDL, due to the strong discriminability of the learnt dictionary.

We denote training samples within the neighborhood as $\mathbf{X} = [x_1, x_2, \dots, x_i, \dots, x_N] \in \mathbb{R}^{M \times N}$. The number of categories is represented as k , the category-specific dictionary corresponding to l th category as $\mathbf{D}_l \in \mathbb{R}^{M \times P_l}$, where P_l is the number of atoms in the sub-dictionary \mathbf{D}_l . The complete dictionary is $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_l, \dots, \mathbf{D}_k] \in \mathbb{R}^{M \times P}$, with $P = P_1 + P_2 + \dots + P_k$. The samples \mathbf{X} can be represented by the sparse code $\mathbf{A}(\mathbf{x}, \mathbf{D}) \in \mathbb{R}^{P \times N}$ by solving the following Lasso problem:

$$\mathbf{A}(\mathbf{x}, \mathbf{D}) = \arg \min_{\mathbf{Z} \in \mathbb{R}^{P \times N}} \|\mathbf{X} - \mathbf{DZ}\|_F^2 + \lambda \|\mathbf{Z}\|_{1,2} \tag{7}$$

where λ is the regularization parameter, $\|\mathbf{Z}\|_{1,2} = \sum_{i=1}^P \|\mathbf{Z}_i\|_2$ is the $l_{1,2}$ -norm of \mathbf{Z} , and $\mathbf{Z}_i \in \mathbb{R}^{1 \times N}$ is the i th row of \mathbf{Z} . Since neighboring pixels often share the same label with high probability, joint sparsity is imposed to enforce the sparse codes to have a row sparsity pattern. The neighboring pixels are selected by superpixel segmentation, which will be described in detail in Section 3. Many sparse recovery techniques are able to solve Equation (7), such as the sparse reconstruction by separable approximation [22], alternating direction method of multipliers [23], and fast iterative shrinkage-thresholding algorithm [24].

Obviously, the effect of sparse coding in Equation (7) largely depends on the quality of dictionary \mathbf{D} . And the quality of dictionary \mathbf{D} depends on the defined loss function. In TDDL [16], Mairal et al. suggested defining the loss function by classification error, which fully utilized the label information. To improve the dictionary quality further, we impose incoherence constraints between the two sub-dictionaries of the ship and clutter. Denote the label information corresponding to \mathbf{X} as $\mathbf{Y} = [y_1, y_2, \dots, y_i, \dots, y_N] \in \mathbb{R}^{k \times N}$. Given the training data (\mathbf{X}, \mathbf{Y}) , the loss function can be formulated as follows:

$$\min_{\mathbf{D}, \mathbf{W}} L(\mathbf{D}, \mathbf{W}, \mathbf{X}) = \min_{\mathbf{D}, \mathbf{W}} \frac{1}{2} \|\mathbf{Y} - \mathbf{W}\mathbf{A}\|_F^2 + \frac{\mu}{2} \|\mathbf{W}\|_F^2 + \frac{\eta}{2} \sum_{l=1}^k \frac{1}{2P_l(P - P_l)} \|\mathbf{D}_l^T \mathbf{D}_{-l}\|_F^2 \tag{8}$$

where μ and η are the regularization parameters, $\mathbf{W} \in \mathbb{R}^{k \times P}$ is the parameters of the linear classifier, $\mathbf{A} \in \mathbb{R}^{P \times N}$ is given by Equation (7), \mathbf{D}_{-l} is denoted as the sub-dictionaries by removing \mathbf{D}_l from \mathbf{D} . In Equation (8), the term $\|\mathbf{Y} - \mathbf{W}\mathbf{A}\|_F^2$ describes the classification error, the term $\|\mathbf{W}\|_F^2$ is to avoid overfitting of the classifier, and the term $\|\mathbf{D}_l^T \mathbf{D}_{-l}\|_F^2$ as incoherence is to enforce category-specific sub-dictionaries incoherency. The coefficient $1/2P_l(P - P_l)$ is to reduce the influence of the sub-dictionary size and to make the learnt dictionary more stable for classification, which was introduced by Gao et al. [25].

2.3. Optimization Procedure

For convenience, the loss function $L(\mathbf{D}, \mathbf{W}, \mathbf{X})$ in Equation (8) can be further represented by two parts, L_1 and L_2 , which are defined as follows:

$$L_1 = \frac{1}{2} \|\mathbf{Y} - \mathbf{W}\mathbf{A}\|_F^2 + \frac{\mu}{2} \|\mathbf{W}\|_F^2 \tag{9}$$

$$L_2 = \frac{\eta}{2} \sum_{l=1}^k \frac{1}{2P_l(P - P_l)} \|\mathbf{D}_l^T \mathbf{D}_{-l}\|_F^2 \tag{10}$$

Following the derivations in [16], we can show that $L(\mathbf{D}, \mathbf{W}, \mathbf{X})$ is differentiable on $\mathbf{D} \times \mathbf{W}$. It is simple to obtain the gradient with respect to \mathbf{W} , i.e.,

$$\frac{\partial L}{\partial \mathbf{W}} = (\mathbf{W}\mathbf{A} - \mathbf{Y})\mathbf{A}^T + \mu\mathbf{W}. \tag{11}$$

Applying the chain rule, we can compute the gradient with respect to the dictionary \mathbf{D} :

$$\frac{\partial L}{\partial \mathbf{D}} = \frac{\partial L_1}{\partial \mathbf{A}} \frac{\partial \mathbf{A}}{\partial \mathbf{D}} + \frac{\partial L_2}{\partial \mathbf{D}}. \tag{12}$$

Obviously, the derivative $\partial L_1 / \partial \mathbf{A}$ can be computed in the same way as $\partial L / \partial \mathbf{W}$. The key point is to compute the derivative $\partial \mathbf{A} / \partial \mathbf{D}$, since there is no explicit expression of \mathbf{D} for the sparse codes \mathbf{A} . Applying the fixed point differentiation [21] to Equation (7), Sun et al. derived the explicit expression of $\partial \mathbf{A} / \partial \mathbf{D}$, which is illustrated in Appendix VII in [26]. Here, we give the vectorization form of the derivative of $\tilde{\mathbf{A}}$ with respect to D_{mn} :

$$\text{vec} \left(\frac{\partial \tilde{\mathbf{A}}^T}{\partial D_{mn}} \right) = \left(\tilde{\mathbf{D}}^T \tilde{\mathbf{D}} \otimes \mathbf{I}_N + \lambda \mathbf{\Gamma} \right)^{-1} \times \text{vec} \left(\tilde{\mathbf{A}}^T \frac{\partial \tilde{\mathbf{D}}^T \tilde{\mathbf{D}}}{\partial D_{mn}} + \frac{\partial \mathbf{X}^T \tilde{\mathbf{D}}}{D_{mn}} \right) \tag{13}$$

where $\tilde{\mathbf{A}} = \mathbf{A}_\Lambda \in \mathbb{R}^{P_\Lambda \times N}$ is denoted as the active rows of \mathbf{A} , $\tilde{\mathbf{D}} = \mathbf{D}_\Lambda \in \mathbb{R}^{M \times P_\Lambda}$ is denoted as the active atoms of \mathbf{D} , and Λ is the active set such that

$$\Lambda = \{i : \|\mathbf{A}_i\|_2 \neq 0, i \in \{1, \dots, P\}\} \tag{14}$$

where \mathbf{A}_i denotes the i th row of \mathbf{A} . And $\mathbf{\Gamma}$ is defined as

$$\mathbf{\Gamma} = \mathbf{\Gamma}_1 \oplus \dots \oplus \mathbf{\Gamma}_{P_\Lambda} \tag{15}$$

where \oplus is the direct sum of matrices, $\mathbf{\Gamma}_i = (\mathbf{I}_N / \|\tilde{\mathbf{A}}_i\|_2) - (\tilde{\mathbf{A}}_i^T \tilde{\mathbf{A}}_i / \|\tilde{\mathbf{A}}_i\|_2^3)$, $i = 1, \dots, P_\Lambda$. Combining Equations (13) and (14), the explicit form of $\partial L_1 / \partial \mathbf{A}$ can be easily computed.

As the other part of $\partial L_2 / \partial \mathbf{D}$, the derivative $\partial L_2 / \partial \mathbf{D}$ can be rewritten as the following expression

$$\frac{\partial L_2}{\partial \mathbf{D}} = \left[\frac{\partial L_2}{\partial \mathbf{D}_1}, \dots, \frac{\partial L_2}{\partial \mathbf{D}_l}, \dots, \frac{\partial L_2}{\partial \mathbf{D}_k} \right]. \tag{16}$$

Therefore, we have

$$\frac{\partial L_2}{\partial \mathbf{D}} = \eta \left[\frac{\mathbf{D}_{-1} \mathbf{D}_{-1}^T \mathbf{D}_1}{P_1(P-P_1)}, \dots, \frac{\mathbf{D}_{-l} \mathbf{D}_{-l}^T \mathbf{D}_l}{P_l(P-P_l)}, \dots, \frac{\mathbf{D}_{-k} \mathbf{D}_{-k}^T \mathbf{D}_k}{P_k(P-P_k)} \right]. \quad (17)$$

Now, we conclude the derivation results as follows:

$$\frac{\partial L}{\partial \mathbf{D}} = -\mathbf{D} \boldsymbol{\beta} \mathbf{A}^T + (\mathbf{X} - \mathbf{D} \mathbf{A}) \boldsymbol{\beta}^T + \eta \mathbf{E} \quad (18)$$

where $\boldsymbol{\beta} \in \mathbb{R}^{P \times N}$ is defined as

$$\boldsymbol{\beta}_{\Lambda^c} = 0 \quad (19)$$

$$\text{vec}(\boldsymbol{\beta}_{\Lambda}^T) = \left(\tilde{\mathbf{D}}^T \tilde{\mathbf{D}} \otimes \mathbf{I}_N + \lambda \boldsymbol{\Gamma} \right)^{-T} \times \text{vec} \left(\left(\mathbf{W} \tilde{\mathbf{A}} - \tilde{\mathbf{Y}} \right)^T \tilde{\mathbf{W}} \right) \quad (20)$$

and $\mathbf{E} \in \mathbb{R}^{M \times P}$ is defined as

$$\mathbf{E} = \left[\frac{\mathbf{D}_{-1} \mathbf{D}_{-1}^T \mathbf{D}_1}{P_1(P-P_1)}, \dots, \frac{\mathbf{D}_{-l} \mathbf{D}_{-l}^T \mathbf{D}_l}{P_l(P-P_l)}, \dots, \frac{\mathbf{D}_{-k} \mathbf{D}_{-k}^T \mathbf{D}_k}{P_k(P-P_k)} \right]. \quad (21)$$

We summarize the overall optimization for TDDDL in Algorithm 1.

Algorithm 1 Stochastic gradient descent algorithm for TDDDL

Input:

The training samples: $\mathbf{X}_{training} \in \mathbb{R}^{M \times N_{training}}$, the corresponding labels: $\mathbf{Y}_{training} \in \mathbb{R}^{k \times N_{training}}$.

Initial dictionary $\mathbf{D} \in \mathbb{R}^{M \times P}$ and classifier $\mathbf{W} \in \mathbb{R}^{k \times P}$.

Regularization parameter $\lambda, \mu, \eta \in \mathbb{R}$.

Number of iterations T , parameter \mathbf{t}_0 , learning rate parameter ρ .

Repeat:

1: **for** $t=0$ to T **do**

2: Draw samples (\mathbf{X}, \mathbf{Y}) from training set.

Compute sparse code \mathbf{A} according to Equation (7).

3: Compute the active set Λ according to Equation (14).

4: Compute the matrix $\boldsymbol{\beta}$ and \mathbf{E} according to Equations (19), (20) and (21).

Choose the learning rate

5: $\rho_t \leftarrow \min(\rho, \rho \mathbf{t}_0 / t)$

normally, set $\mathbf{t}_0 = T/10$.

Update the dictionary \mathbf{D} and classifier \mathbf{W}

6: $\mathbf{W} \leftarrow \mathbf{W} - \rho_t ((\mathbf{W} \mathbf{A} - \mathbf{Y}) \mathbf{A}^T + \mu \mathbf{W})$

$\mathbf{D} \leftarrow \mathbf{D} - \rho_t (-\mathbf{D} \boldsymbol{\beta} \mathbf{A}^T + (\mathbf{X} - \mathbf{D} \mathbf{A}) \boldsymbol{\beta}^T + \eta \mathbf{E})$

and normalize each column of \mathbf{D} with respect to L_2 -norm.

7: **end for**

Output: \mathbf{D} and \mathbf{W}

3. The Proposed Ship Detection Method

Ship detection can be viewed as a binary classification problem. Conventional methods generally complete ship detection in an unsupervised way. Supervised methods, typically deep neural network methods, provide us a new vision for ship detection. Thus, we propose a novel ship detection method via TDDDL. Figure 1 shows the main framework of the proposed ship detection method. Given a PolSAR image, superpixel segmentation is performed following the filter operation with boxcar

filter. The superpixel is employed as the basic processing cell. Based on the superpixel segmentation result, we train a task-driven discriminative dictionary and a linear classifier jointly via TDDDL. Then, we encode the superpixels with the learnt dictionary. Finally, we achieve ship detection with binary classification.

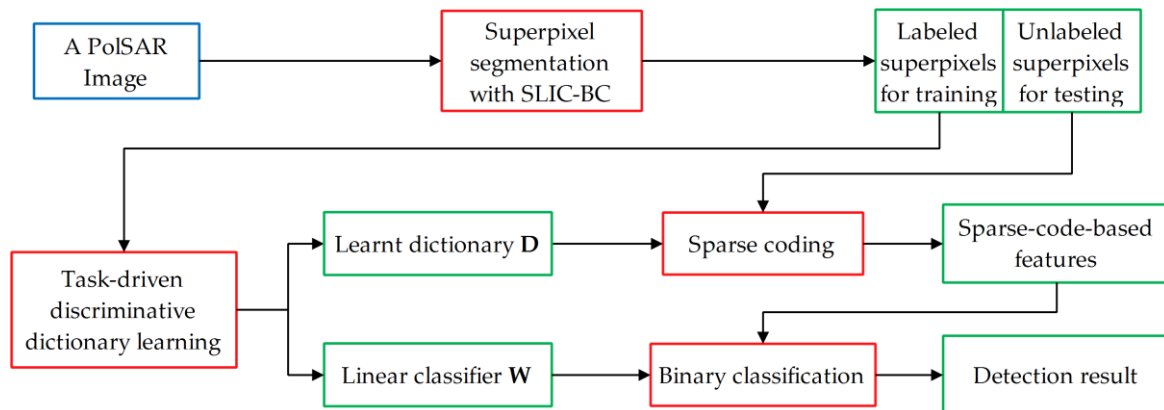


Figure 1. The flowchart of the proposed ship detection method. The blue box is the input, the red boxes are the key processing modules, and the green boxes are the outputs.

For the sake of completeness and readability of this article, we briefly introduce the PolSAR image data here. Considering a reciprocal target illuminated by a monostatic SAR, the polarimetric information can be described by a complex scattering vector

$$\mathbf{k} = \begin{bmatrix} S_{HH} & \sqrt{2}S_{HV} & S_{VV} \end{bmatrix}^T \quad (22)$$

where S_{HH} , S_{HV} and S_{VV} denote the complex scattering coefficients. The scattering vector can be multi-look processed for the purpose of speckle reduction, which can be expressed as

$$\mathbf{T} = \frac{1}{n} \sum_{i=1}^n \mathbf{k}_i \mathbf{k}_i^H \quad (23)$$

where H denotes the conjugate transpose operation, and n is the number of looks. The resulting matrix \mathbf{T} is called the n -look covariance matrix. We further represent the covariance matrix \mathbf{T} as a real vector $\mathbf{p} \in \mathbb{R}^{9 \times 1}$ i.e.,

$$\mathbf{p} = [T_{11}, T_{22}, T_{33}, Re(T_{12}), Im(T_{12}), Re(T_{13}), Im(T_{13}), Re(T_{23}), Im(T_{23})]^T \quad (24)$$

which is called pixel element in PolSAR images.

3.1. Superpixel Segmentation

Pixel-based methods utilize single pixel information only, but not the characteristics of local regions. With the improved resolution of PolSAR images, the ship target regions show detailed structure and texture. We can speculate that considering the cues of the region one pixel belongs to will benefit the decision of this pixel, since they may reveal the regional structural or textural difference between the target and clutter. Recent research has also shown that the superpixel can help to retain the target outline and suppress the speckle noise in target detection task [27,28]. On the other hand, since neighboring pixels often share the same label with high probability, joint sparsity prior can be adapted exactly at the superpixel-level.

Superpixel segmentation methods in optical images cannot be applied roughly in PolSAR images due to the influence of strong speckle noise. In this article, we use the simple iterative clustering

method with boundary constraints (SLIC-BC), which was proposed by Lin et al. [29]. SLIC-BC is an adaption of SLIC [30] with two modifications: (1) A new distance measure is proposed, providing control over boundary adherence, homogeneity and compactness of the superpixels simultaneously. (2) A new strategy to update the positions and intensities of superpixel seeds is proposed. Only reliable pixels within one superpixel can be used to update the superpixel seed. We give a brief introduction about SLIC-BC, and more implementation details can be found in Lin et al. [29].

The distance measurement in SLIC-BC consists of three parts: boundary term, homogeneity term and compactness term, which is defined as

$$d(x, l) = w_b d_b(x, l) + w_h d_h(x, l) + \alpha \times w_c d_c(x, l) \quad (25)$$

where $d(x, l)$ denotes the distance between pixel x and the l th superpixel, $d_b(x, l)$, $d_h(x, l)$ and $d_c(x, l)$ denote the boundary term, homogeneity term and compactness term of the measurement, respectively. The parameters w_b , w_h and w_c denote the weight coefficients, which are defined as $w_b = \frac{(d_h(x, l) + d_c(x, l))}{(2(d_b(x, l) + d_h(x, l) + d_c(x, l)))}$, $w_h = \frac{(d_b(x, l) + d_c(x, l))}{(2(d_b(x, l) + d_h(x, l) + d_c(x, l)))}$, and $w_c = \frac{(d_b(x, l) + d_h(x, l))}{(2(d_b(x, l) + d_h(x, l) + d_c(x, l)))}$. And α is a parameter to flexibly control the compactness of the resulting superpixels. The boundary term $d_b(x, l)$ is defined with the probability of a pixel lying on object boundary

$$d_b(x, l) = \exp\left(\frac{\sum_{x_i} |g(x_i) - g(x)|}{|C^{win}(x)|}\right), x_i \in C^{win}(x) \quad (26)$$

where $C^{win}(x)$ is the collection of pixels in a $win \times win$ window with pixel x as its center, $|C^{win}(x)|$ denotes the number of pixels in $C^{win}(x)$, and $g(x_i)$ denotes the gradient at pixel x_i . The homogeneity term $d_h(x, l)$ is defined with Wishart distance

$$d_h(x, l) = \exp\left(\ln(|C_l|) + \text{Tr}(C_l^{-1} Z_x)\right) \quad (27)$$

where $\text{Tr}(\cdot)$ denotes the trace of a matrix, Z_x and C_l is the covariance matrix of pixel x and l th superpixel seed, respectively. And the compactness term is defined with the Euclidean distance

$$d_c(x, l) = \exp\left(\sqrt{(r_x - r_l)^2 + (c_x - c_l)^2}\right) \quad (28)$$

where (r_x, c_x) and (r_l, c_l) denote the coordinates of pixel x and the l th superpixel seed in the xy plane, respectively.

In the updating strategy in SLIC-BC, Wishart distribution is used to measure the reliability of pixels

$$p(Z_x | C_l) = \frac{n^{nd} |Z_x|^{n-d} \exp\{-n \text{Tr}(C_l^{-1} Z_x)\}}{R(n, d) |C_l|^n}. \quad (29)$$

Only reliable pixels are used to update the position and intensities of the superpixel seeds. The position and intensities of the superpixel seeds are updated by

$$C_{l,i} = \frac{\sum_{x \in \Phi_{l,i-1}} Z_x}{|\Phi_{l,i-1}|}, \Phi_{l,i-1} = \{x | L(x) = l, p(Z_x | C_l) > \nu\} \quad (30)$$

where i is the number of iterations, $L(x)$ returns the superpixel label of pixel x , and ν is the reliability threshold.

Since we require target-contained superpixels with less background pixels, a relatively small superpixel size is preferred. In practice, the desired superpixel size is a key parameter determining the average size of superpixels. In this paper, we set the desired superpixel size to be half the ship target size.

3.2. Learning Dictionary with TDDDL

After obtaining the superpixel result, we train a dictionary and a linear classifier via TDDDL. In the training data, we denote the ship superpixel as $\mathbf{X}_s = [x_1^s, x_2^s, \dots, x_{T_s}^s] \in \mathbb{R}^{9 \times T_s}$, and the clutter superpixel as $\mathbf{X}_c = [x_1^c, x_2^c, \dots, x_{T_c}^c] \in \mathbb{R}^{9 \times T_c}$, where T_s and T_c are the corresponding superpixel size, respectively. The dictionary $\mathbf{D} = [\mathbf{D}_s, \mathbf{D}_c] \in \mathbb{R}^{9 \times P}$ consists of two sub-dictionaries, where $\mathbf{D}_s \in \mathbb{R}^{9 \times P_s}$ and $\mathbf{D}_c \in \mathbb{R}^{9 \times P_c}$ are the sub-dictionary corresponding to the ship and clutter. We initialize the dictionary \mathbf{D} by learning category-specific dictionary in an unsupervised way. Concretely, we compute the category-specific sub-dictionary by solving

$$\min_{\mathbf{D}_s, \mathbf{A}} \|\mathbf{X}_s - \mathbf{D}_s \mathbf{A}\|_F^2 + \lambda \|\mathbf{A}\|_{1,2} \text{ s.t. } \|\mathbf{D}_j^s\|_F^2 = 1, \forall j, 1 \leq j \leq P_s \quad (31)$$

and

$$\min_{\mathbf{D}_c, \mathbf{A}} \|\mathbf{X}_c - \mathbf{D}_c \mathbf{A}\|_F^2 + \lambda \|\mathbf{A}\|_{1,2} \text{ s.t. } \|\mathbf{D}_j^c\|_F^2 = 1, \forall j, 1 \leq j \leq P_c \quad (32)$$

where \mathbf{D}_j^s and \mathbf{D}_j^c are the j th column of \mathbf{D}_s and \mathbf{D}_c , respectively. The initial classifier parameter matrix \mathbf{W} is determined based on the label information \mathbf{Y} and the initial dictionary \mathbf{D} . We have

$$\min_{\mathbf{W}} \|\mathbf{Y}_s, \mathbf{Y}_c - \mathbf{W}[\mathbf{A}_s, \mathbf{A}_c]\|_2^2 + \frac{\mu}{2} \|\mathbf{W}\|_F^2 \quad (33)$$

where $\mathbf{Y}_s = [1, 1, \dots, 1] \in \mathbb{R}^{1 \times T_s}$ and $\mathbf{Y}_c = [-1, -1, \dots, -1] \in \mathbb{R}^{1 \times T_c}$ denote the label of the ship superpixels and clutter superpixels, and $[\mathbf{A}_s, \mathbf{A}_c]$ is the solution of Equation (7) by substituting $[\mathbf{X}_s, \mathbf{X}_c]$ for \mathbf{X} . With the initialization results of \mathbf{D} and \mathbf{W} above, we complete dictionary learning via Algorithm 1 and obtain the learnt dictionary \mathbf{D} and classifier \mathbf{W} .

3.3. Encoding with Learnt Dictionary

With the learnt dictionary \mathbf{D} , we encode an unlabeled superpixel $X_u = [x_1^u, x_2^u, \dots, x_{T_u}^u] \in \mathbb{R}^{9 \times T_u}$ by solving the following problem:

$$\arg \min_{\mathbf{A}_u} \|\mathbf{X}_u - \mathbf{D} \mathbf{A}_u\|_F^2 + \lambda \|\mathbf{A}_u\|_{1,2}. \quad (34)$$

The resulting feature \mathbf{A}_u is the sparse code of the sample \mathbf{X}_u on the learnt dictionary \mathbf{D} . In fact, it is constructed in virtual dictionary domain, while conventional features are constructed in the original image domain. Usually the features on the image domain do not achieve good enough results for harsh detection tasks, because the ship and clutter in the image domain are not sufficiently different. However, the difference can be amplified in transform domain. On the other hand, the features in most previous methods are designed artificially, while the feature in the proposed method is designed with active learning strategy, and thus, is more adaptive. The learnt feature in the proposed method includes more information about the data and can reveal the polarimetric structure difference between the ship and the clutter. In addition, the proposed method performs feature extraction and threshold determination jointly. It has been proven that feature and threshold joint learning is significantly better than their respective learning [26,31]. Therefore, we can conclude that the feature learnt in the proposed method theoretically offers significant advantages compared to previous methods.

3.4. Binary Classification

Once the feature \mathbf{A}_u is obtained, we identify the label of each pixel of \mathbf{X}_u based on the following rule

$$\text{identity}(x_i^u) = \text{sign}(\mathbf{W} \mathbf{a}_i^u) \quad (35)$$

where x_i^u denotes the i th pixel element of superpixel \mathbf{X}_u , \mathbf{a}_i^u denotes i th column of \mathbf{A}_u , and the function $\text{sign}(\cdot)$ returns the sign of a real number. Since pixels within a superpixel have similar characteristics

such as intensity, texture and polarimetric structure, we identify the label of superpixel X_u based on the pixel label

$$\text{identity}(X_u) = \text{sign} \left(\sum_{i=1}^{T_u} \text{sign}(\mathbf{W}\alpha_i^u) \right) \quad (36)$$

Finally, we obtain the binary ship detection result.

4. Experiments and Discussions

C-band RADARSAT-2 polarimetric SLC SAR images acquired over Tanggu port area and Dalian port area are used for experiments. The parameters of the images are tabulated in Table 1. The intensity images of single polarimetric channels, Pauli vector color-coded images and geographic locations are shown in Figures 2 and 3. The areas R1 and R3 in Figures 2e and 3e are used for testing, with training areas selected from R2 and R4, respectively.

Table 1. The parameters of the PolSAR images.

Location	Date	Area (Pixels) Slant Range \times Azimuth	Resolution (m) Slant Range \times Azimuth	Pixel Spacing (m) Slant Range \times Azimuth	Incidence Angle($^{\circ}$)
Tanggu port	23/06/2011	940 \times 1810	5.2 \times 7.6	4.7 \times 5.3	30
Dalian port	17/04/2012	1500 \times 870	5.2 \times 7.6	4.7 \times 5.3	21

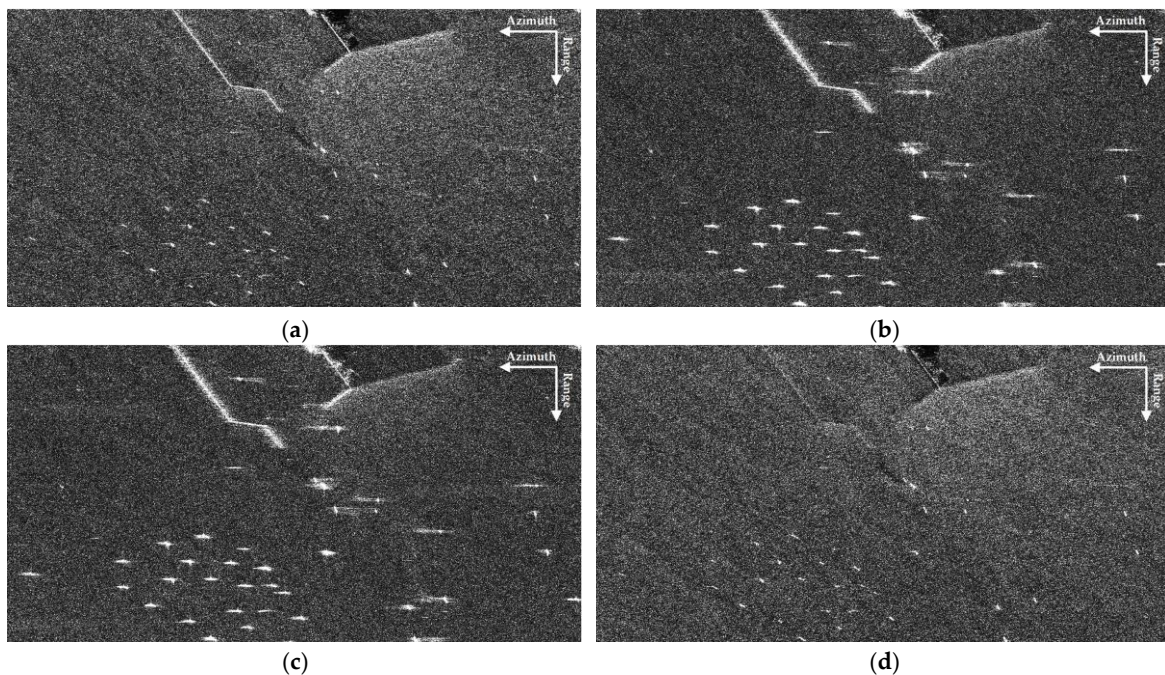


Figure 2. Cont.

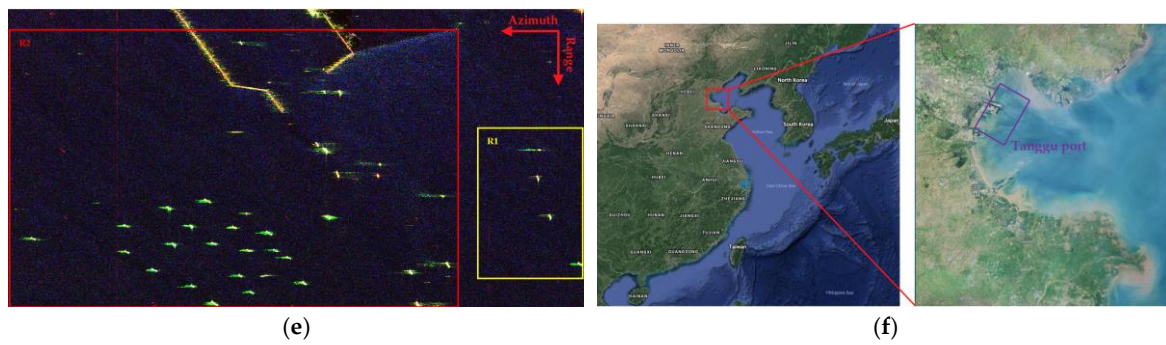


Figure 2. RADARSAT-2 PolSAR image over Tanggu port area. (a–d) Intensity images of HH, HV, VH and VV channels. (e) Pauli vector color-coded image, using $|HH - VV|$, $|HV|$, and $|HH + VV|$ as red, green, and blue, respectively. The areas R1 and R2 are used for training and testing, respectively. (f) The geographic location of Tanggu port area.

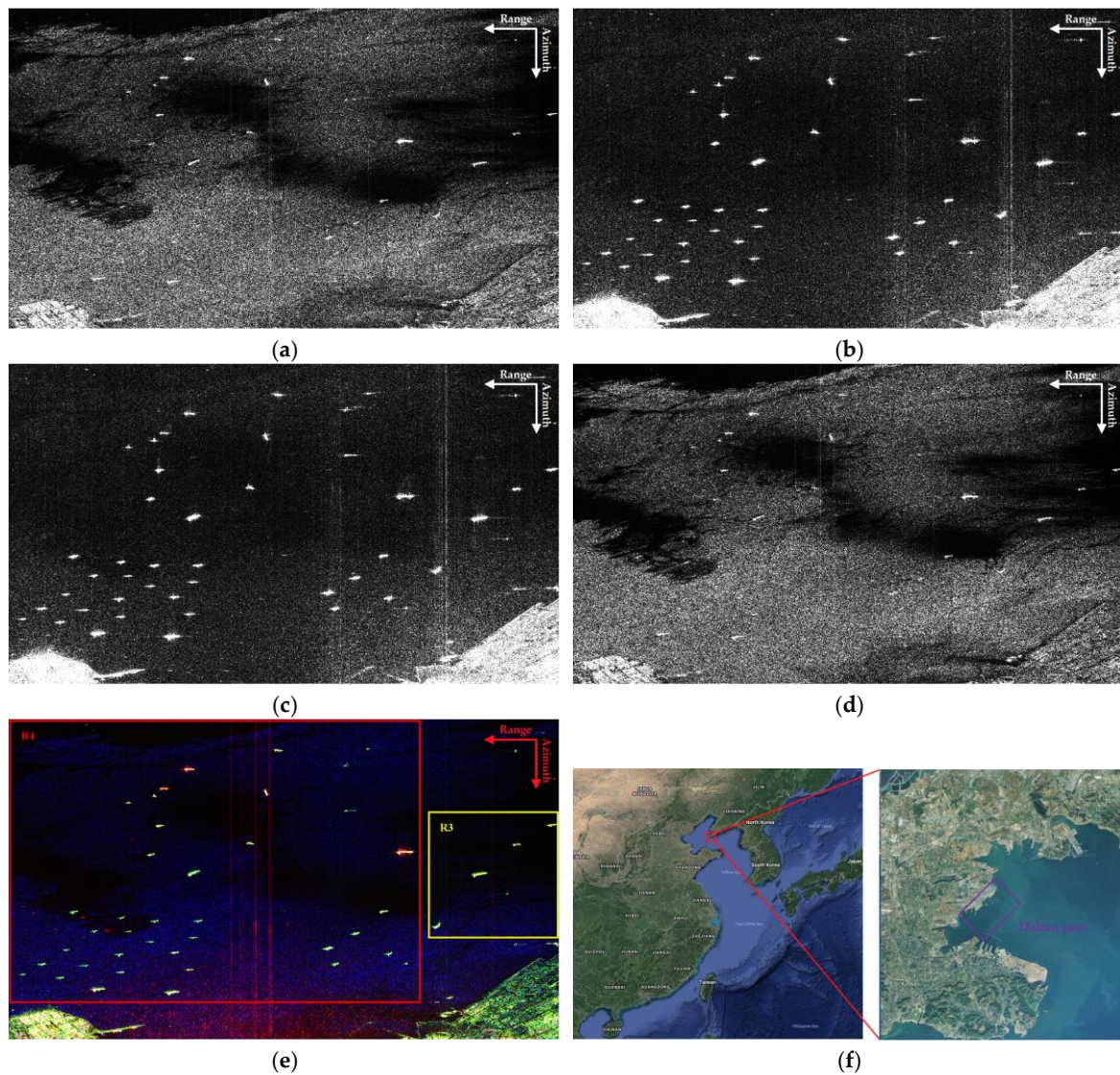


Figure 3. RADARSAT-2 PolSAR image over Dalian port area. (a–d) Intensity images of HH, HV, VH and VV channels. (e) Pauli vector color-coded image, using $|HH - VV|$, $|HV|$, and $|HH + VV|$ as red, green, and blue, respectively. The areas R3 and R4 are used for training and testing, respectively. (f) The geographic location of Dalian port area.

The ground truth of the testing areas R1 and R3 are shown with Pauli vector color-coded in Figure 4. The strong (group S) and weak (group W) targets are marked with green rectangles and yellow circles, respectively. The ground truth definition is based on the previous work by Song et al. [31] and He et al. [32]. We know that it is easy to detect strong targets and difficult to detect weak targets. And in real-scene SAR images, strong and weak targets often appear together. In order to demonstrate the performance comparison, especially the weak target detection performance comparison, we separate the targets into strong (group S) and weak (group W) targets according to the relative magnitude of the target average intensity and its surrounding clutter average intensity. If the target average intensity is higher than its surrounding clutter average intensity, the target is grouped as the strong target; otherwise, the target is grouped as the weak target. The surrounding clutter pixels are chosen in a window with the target as its center and the window size is set as twice the average target size. Meanwhile, adjacent target pixels in the window are excluded as outliers for clutter estimation based on ground truth.

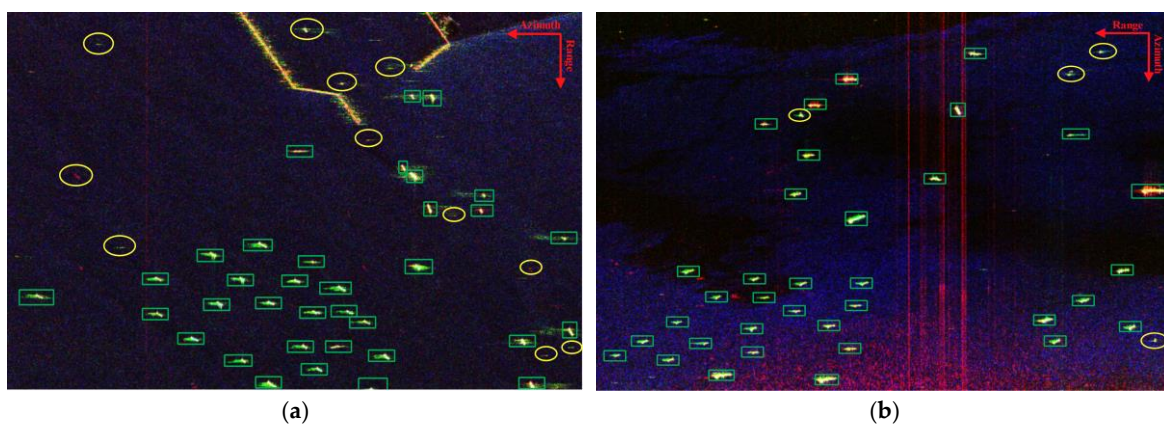


Figure 4. RADARSAT-2 PolSAR images for testing. Pauli vector color-coded images over (a) R2 and (b) R4 areas. In each image, the strong (group S) and weak (group W) targets are marked with green rectangles and yellow circles, respectively.

In the following, we first describe the parameter setting of the proposed method, and then give the performance evaluation on synthetic data. Finally, the performance evaluation on real-scene data is also presented. We compare the proposed method with the iterative censoring CFAR (IC-CFAR) detector [10], Variational Bayesian Inference (VBI) [31], Superpixel-level local information measurement (SLIM) detector [32], and contextual region-based convolutional neural network with multilayer fusion (CRCNN-MF) [15].

4.1. Parameter Setting

The parameters of the comparative methods are set up according to the original papers. For instance, all the hyperparameters involved in the VBI method are set in noninformative manner to reduce their impact on the estimation of posterior distributions. Thus, both hyperparameters β_1 and β_2 in VBI are set to be 10^{-6} . For the IC-CFAR method, the confidence level of a pixel being target is set as 0.02 [10]. Thus, an index matrix can be obtained to label whether each pixel of the image is a potential target pixel or not. More parameter setting details can be found in [9,15,31,32].

The parameters in the proposed method include the parameters of superpixel segmentation and the parameters of TDDDL. The parameter α in superpixel segmentation largely depends on our needs. When α is large, spatial proximity is more important and the resulting superpixels are more compact. When α is small, the resulting superpixels adhere more tightly to image boundaries, but have less regular shape. For PolSAR images, α can be in the range [0.5, 20]. In this paper, we prefer superpixels with more boundary adherence. Thus, smaller α should be employed. Figure 5 shows the superpixel

segmentation results with different α . We can find that boundary adherence is not guaranteed when $\alpha = 3.0$, and the boundary is too complicated when $\alpha = 0.7$. Empirically, the range $[0.8, 1.5]$ is preferred. In this paper, we set the parameter α to be 1.0.

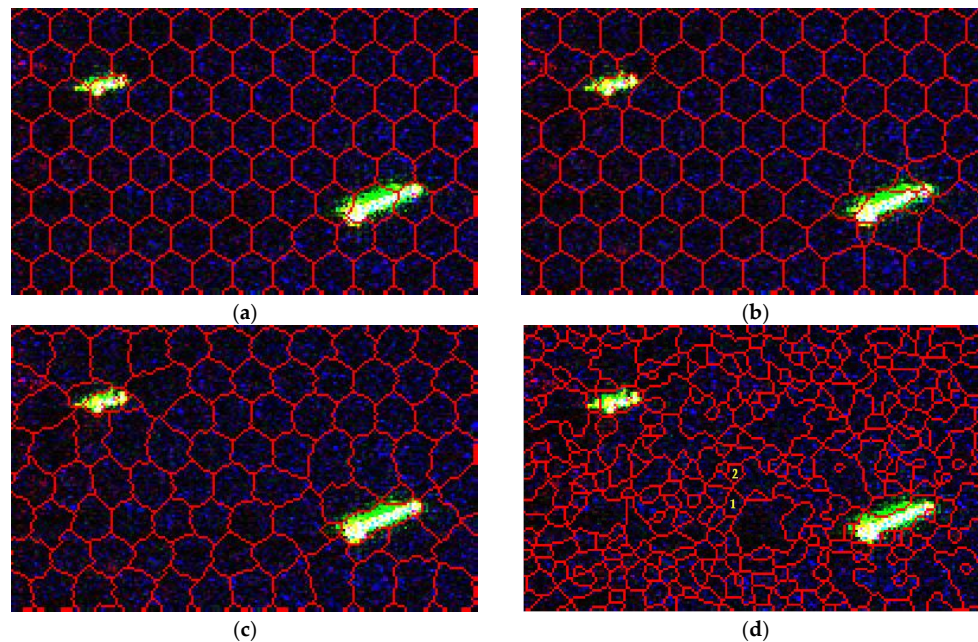


Figure 5. Superpixel segmentation results with different α . (a) $\alpha = 3.0$. (b) $\alpha = 1.5$. (c) $\alpha = 1.0$. (d) $\alpha = 0.7$. As α decreases, the resulting superpixels adhere more tightly to image boundaries. But an insufficient α produces overly complicated superpixel boundaries, where even the connectivity of superpixels cannot be guaranteed. For instance, region 1 and region 2 in sub-figure (d) belong to the same superpixel.

For the parameters in TDDDL, we use a few simple heuristics to reduce the search space, which are used in many dictionary learning methods [16,33,34]. The regularization parameter μ is fixed to be 10^{-3} [34]. And we try parameters $\lambda = 0.35 + 0.05j$, with $j \in \{-3, -2, \dots, 2, 3\}$. The candidate parameters of η are $\{0, 0.05, \dots, 0.25, 0.3\}$. The detection performance versus the regularization parameter λ and η is demonstrated in Figure 6. Based on these figures, we obtain optimal parameters in TDDDL. Here we list the parameters used in the proposed method in Table 2.

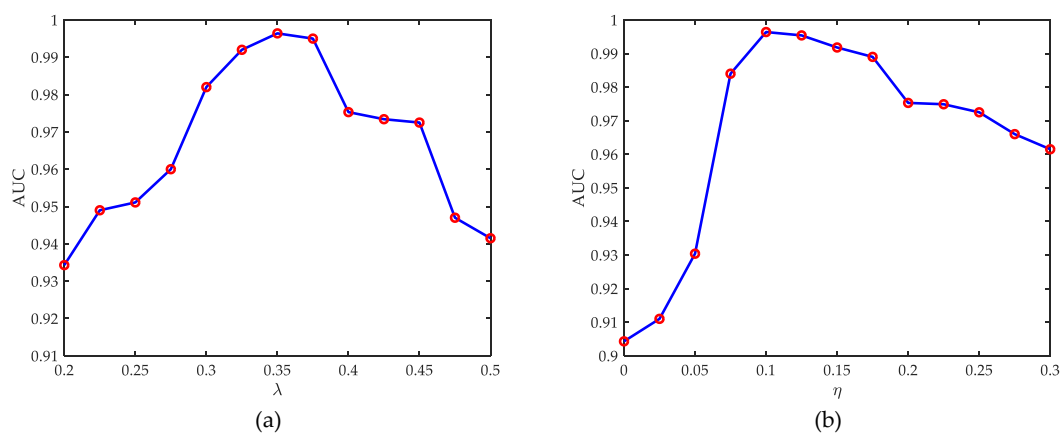


Figure 6. The area under the receiver operation characteristic (ROC) curve (AUC) versus the regularization parameters λ and η : (a) the detection probability versus λ , where we set λ to be 0.2, 0.225, 0.25, 0.275, 0.3, 0.325, 0.35, 0.375, 0.4, 0.425, 0.45, 0.475, and 0.5; and (b) the detection probability versus η , where we set η to be 0, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3.

Table 2. The parameters used in the proposed method.

Parameter	α	λ	μ	η
Values	1.0	0.35	10^{-3}	0.1

4.2. Performance Evaluation on Synthetic Data

First, we evaluate the performance of the test methods quantitatively using synthetic data. In the synthetic data, sea clutter is modeled by K distribution with SCR ranging from 0 dB to 12 dB. Figure 7 shows the span images of the synthetic data and the corresponding ground truth image where white pixels denote target pixels. In order to evaluate the ship detection results of the test methods in a quantitative way, the actual detection possibility P_d and figure of merit FoM are defined as follows:

$$P_d = \frac{n_{dt}}{n_{target}}, FoM = \frac{n_{dt}}{n_{target} + n_{dc}} \tag{37}$$

where n_{target} denotes the total number of true target pixels, n_{dt} and n_{dc} are the number of correctly detected target pixels and that of clutter pixels detected as target pixels, respectively. Higher detection probability and figure of merit implies better detection methods. Figure 8 presents the detection performances of the five test methods under different SCR conditions. We can find that all the test methods show satisfactory performances in detection probability at relatively high SCR conditions, while the proposed method and SLIM achieve higher detection probability at low SCR conditions. Obviously, the figure of merit of the proposed method is higher than those of the comparative methods, which implies that the proposed method produces less false alarms. Therefore, we can conclude that the proposed method outperforms the comparative methods on synthetic data.

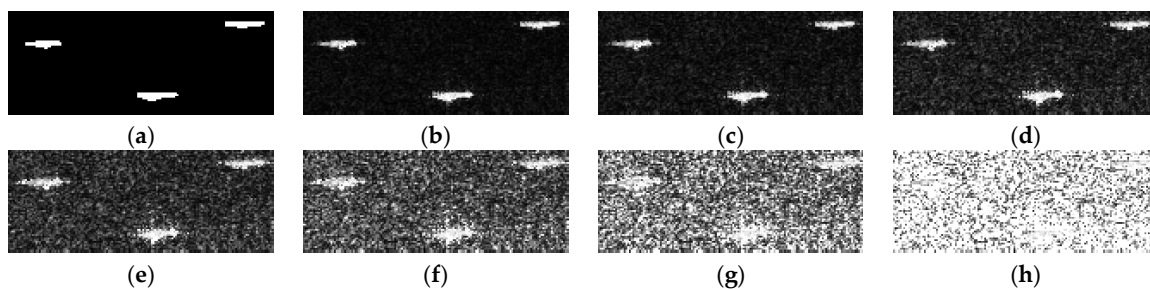


Figure 7. Illustration over synthetic data. (a) Ground truth image. (b–h) Synthetic span image with SCR = 12 dB, 10 dB, 8 dB, 6 dB, 4 dB, 2 dB and 0 dB.

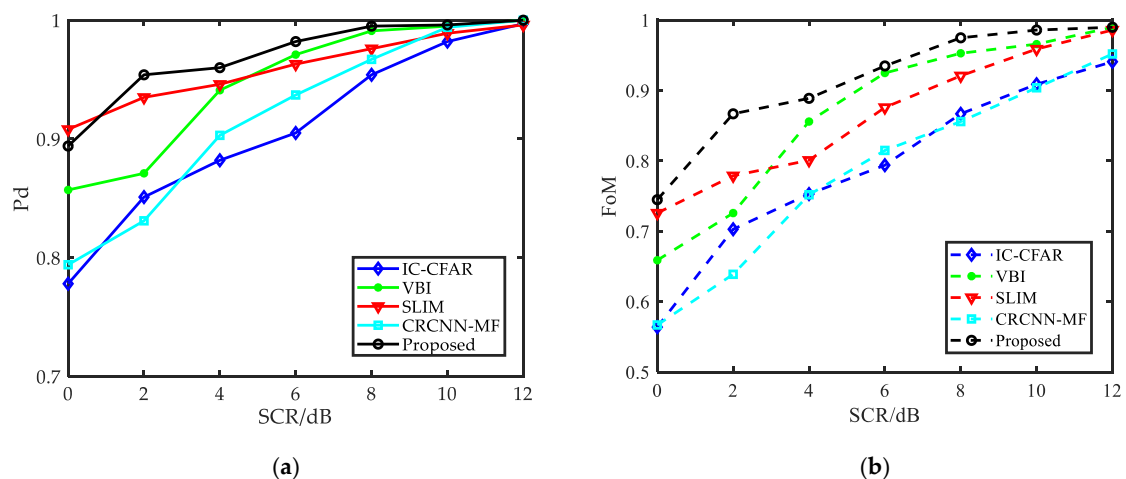


Figure 8. Performance evaluation on synthetic data. (a) The detection possibility. (b) The figure of merit.

4.3. Performance Evaluation on Real-Scene Data

Two RADARSAT-2 images acquired over Tanggu port and Dalian port are used for real-scene validation. The ship number of different groups is tabulated in Table 3. Table 4 and Figures 9–13 report the ship detection results of the proposed method and the comparative methods. In Table 4, $N_{dt} = N_{dt}^S + N_{dt}^W$ denote the total number of detected ships, where N_{dt}^S and N_{dt}^W denote the number of detected ships belonging to group S and group W, respectively. The number of false alarms is denoted as N_{dc} . In Figures 9–13, the false alarms and missed targets are marked with red circles and white rectangles, respectively. In Figure 9a over Tanggu port area, we can see that the proposed method can detect all the ship targets, including 35 strong targets and 11 weak targets, with only 1 false alarm. In Figure 9b over Dalian port area, the proposed method can detect all 35 strong targets and 4 weak targets, with no false alarms. All comparative methods can detect almost all the strong targets. The performance gaps lie in false alarms and weak target detection. For Tanggu port area, the IC-CFAR detector detect 2 weak targets with 11 false alarms, and CRCNN-MF detect 5 weak targets with 9 false alarms. The VBI and SLIM perform better than the IC-CFAR detector and CRCNN-MF in the image of Tanggu port area. The VBI detect 6 weak targets with 5 false alarms, and SLIM detect 10 weak targets with 4 false alarms. For Dalian port area, the proposed method, VBI and CRCNN-MF outperform the IC-CFAR detector and SLIM, producing fewer false alarms. In summary, the proposed method achieves the highest detection accuracy and lowest false alarm rate. Therefore, we can conclude that the proposed method is superior to the comparative methods.

Table 3. Ground truth of ships.

Location	Group	Number
Tanggu port	group S	35
	group W	11
Dalian port	group S	35
	group W	4

Table 4. Performance comparison of the test methods.

Method	Proposed	IC-CFAR	VBI	SLIM	CRCNN-MF
Tanggu port	N_{dt}^S	35	35	35	35
	N_{dt}^W	11	2	6	10
	N_{dt}	46	37	40	45
	N_{dc}	1	>11	5	4
Dalian port	N_{dt}^S	35	34	35	35
	N_{dt}^W	4	4	4	4
	N_{dt}	39	38	39	39
	N_{dc}	0	3	0	2

N_{dt}^S and N_{dt}^W denote the number of detected ships belonging to group S and group W, respectively. N_{dt}^W denotes the total number of detected ship, and N_{dc} denotes the number of false alarms.

Besides detection accuracy and false alarms, the shape preserving of detection results is another issue. First, we require that the detected targets are complete. Moreover, the detected targets should not be redundant. In the detection results, we mark the broken detected targets with blue rectangles. By comparing the results in Figures 9–13, we can see that the proposed method, SLIM and CRCNN-MF produce much less broken detected targets than the other two methods. However, the sea clutter pixels around the ships influenced by ship tend to be detected by SLIM and CRCNN-MF, and the detected targets produced by SLIM and CRCNN-MF are fatter than those produced by the proposed method, IC-CFAR detector and VBI. The proposed method mainly captures the dominant scatters on ships and maintains the best shape preserving ability.

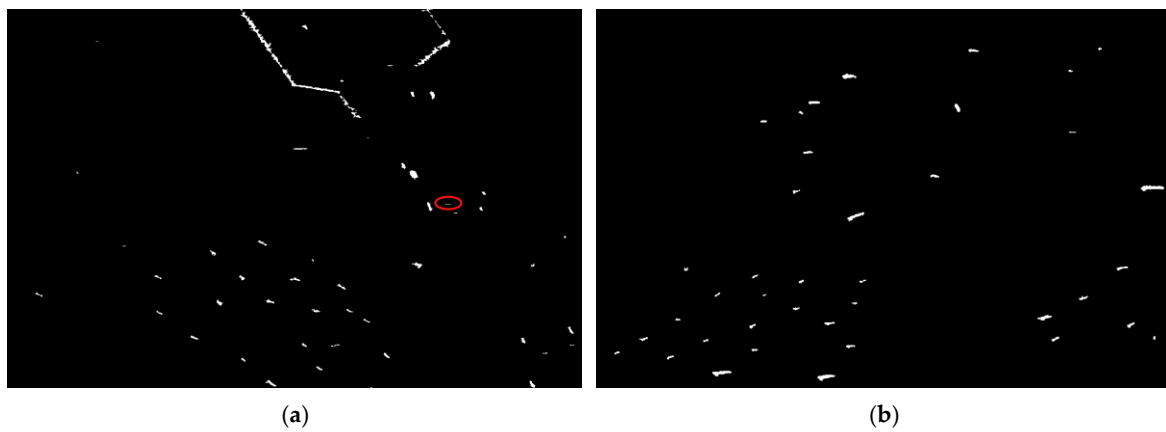


Figure 9. Ship detection results of the proposed method. (a) Tanggu port. (b) Dalian port. The false alarms are marked with red circles. No missed targets or broken detected targets.

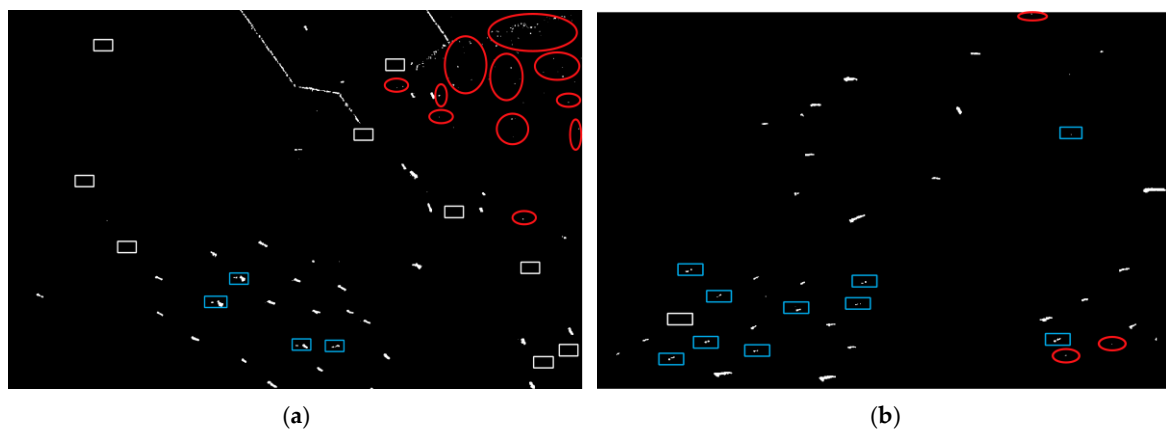


Figure 10. Ship detection results of IC-CFAR. (a) Tanggu port. (b) Dalian port. The false alarms, missed targets and broken detected targets are marked with red circles, white rectangles and blue rectangles, respectively.

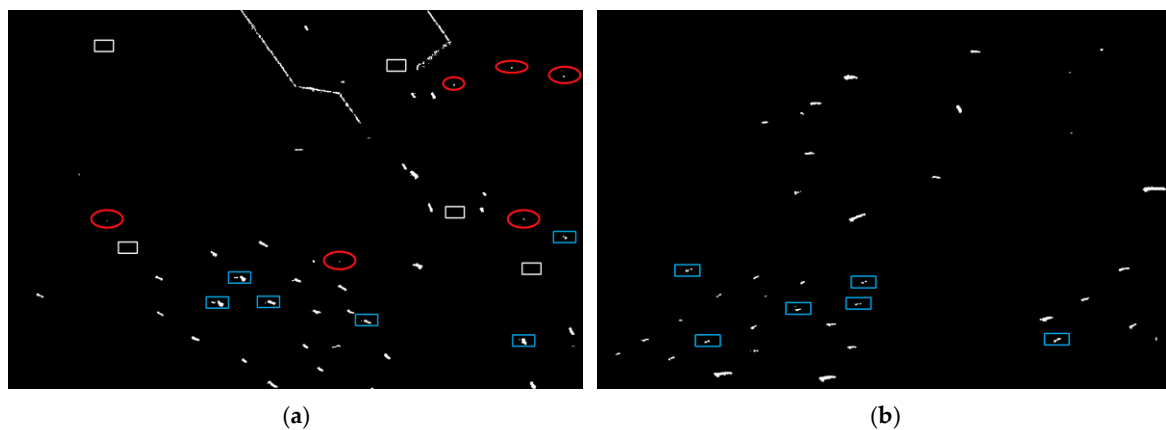


Figure 11. Ship detection results of VBI. (a) Tanggu port. (b) Dalian port. The false alarms, missed targets and broken detected targets are marked with red circles, white rectangles and blue rectangles, respectively.

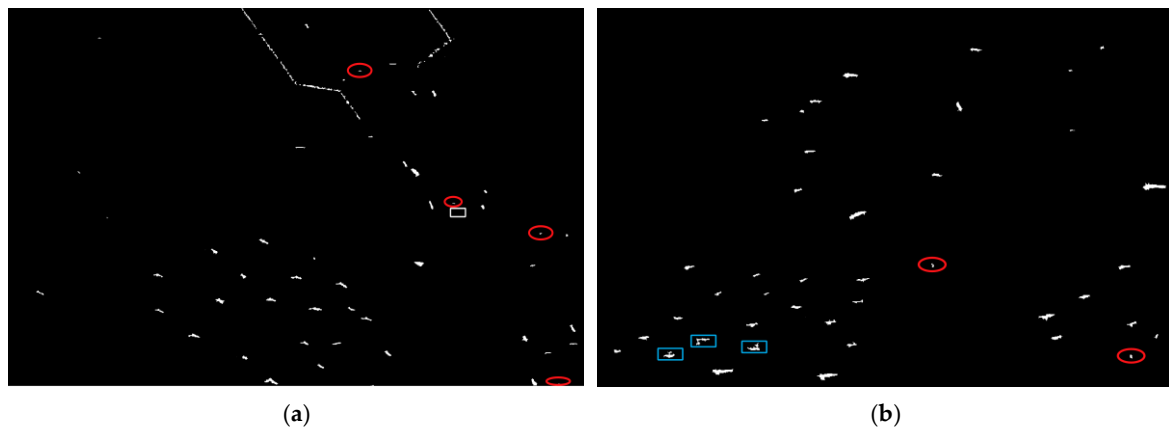


Figure 12. Ship detection results of SLIM. (a) Tanggu port. (b) Dalian port. The false alarms, missed targets and broken detected targets are marked with red circles, white rectangles and blue rectangles, respectively.

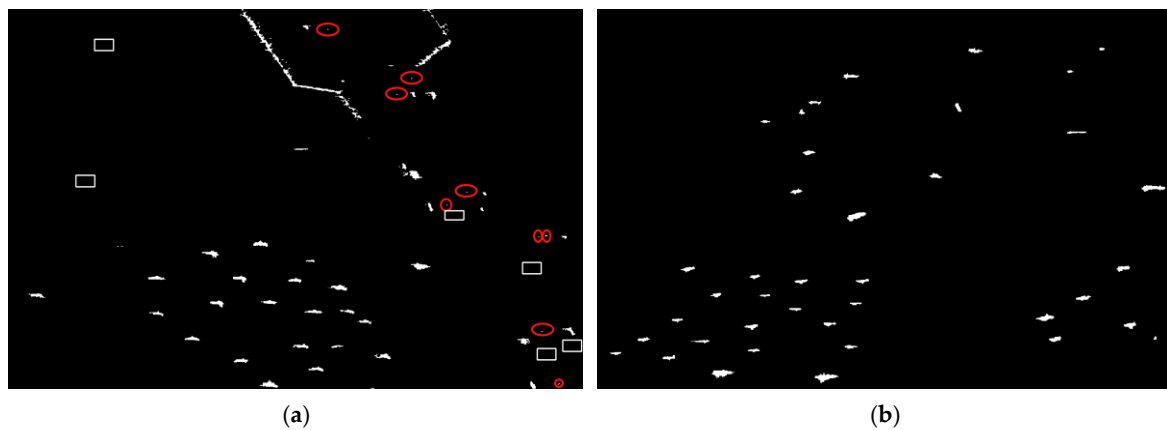


Figure 13. Ship detection results of CRCNN-MF. (a) Tanggu port. (b) Dalian port. The false alarms and missed targets are marked with red circles and white rectangles, respectively. No broken detected targets.

4.4. Computational Cost

The proposed approach and the comparative methods are written in MATLAB code. All the experiments are performed through MATLAB 2017b in 64-bit Windows system with a hardware environment of Intel Core i7 8700 processor and 16-GB RAM. The time consumption of different methods is listed in Table 5. We can see that the proposed method maintains comparable testing time with the CRCMM-MF, while the training time of the proposed method is greatly reduced. The VBI is much slower than the other methods. The main reason for this may depend on the special structure of algorithm, where the updates of the latent variables' expectations are highly coupled. The IC-CFAR detector is the fastest method. The efficiency of the IC-CFAR detector is mainly due to the initial detector applied to the entire cross-polarization image without sliding window. In the SLIM, the sliding window is applied on the superpixel level for a fast processing. And the multiscale superpixel segmentation and local information computation are time consuming. Thus, its computation complexity is moderate. In summary, the proposed method maintains low computation complexity, while it has satisfactory detection probability and figure of merit, as well as low false alarms and good shape preserving ability.

Table 5. Time consumption of different methods.

Location	Method	Training Time/Testing Time
Tanggu port	The proposed method	67.2 s/6.4 s
	IC-CFAR	-/4.9 s
	VBI	-/478.1 s
	SLIM	-/48.6 s
	CRCNN-MF	3080.5 s/5.3 s
Dalian port	The proposed method	71.6 s/8.5 s
	IC-CFAR	-/5.4 s
	VBI	-/515.3 s
	SLIM	-/60.1 s
	CRCNN-MF	3141.5 s/7.8 s

5. Conclusions

In this paper, we proposed a novel ship detection method for PolSAR images via TDDDL. The method is performed at superpixel level to retain target outlines and suppress speckle noise. We extract discriminative features between the ship and clutter by sparse encoding with learnt task-driven discriminative dictionary. A linear classifier is obtained in the training processing. With the learnt features and classifier, we achieve ship detection. Different from previous methods, our ship detection cue is obtained through active learning strategies rather than artificially designed rules, and thus, is more adaptive and effective. The effectiveness and superiority of the proposed method is demonstrated by experimental results performed on two RADARSAT-2 images. In the future, we would like to deeply analyze the polarimetric information hidden behind TDDDL and apply the proposed method to general target detection task for PolSAR images, such as vehicle detection and bridge detection.

Author Contributions: Conceptualization, H.C.; Data curation, H.W.; Formal analysis, H.L. and H.W.; Funding acquisition, J.Y. (Jian Yang); Investigation, H.L.; Methodology, H.L.; Project administration, J.Y. (Junjun Yin); Software, J.Y. (Junjun Yin); Supervision, J.Y. (Jian Yang); Validation, H.L. and H.C.; Writing—original draft, H.L.; Writing—review & editing, H.L. and J.Y. (Jian Yang).

Funding: The work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant No. 61490693 and 61771043, in part by the Key Project of NSFC under Grant No. 61132008 and in part by the CAST Innovation Fund.

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

References

1. Bruschi, S.; Lehner, S.; Fritz, T.; Soccorsi, M.; Soloviev, A.; van Schie, B. Ship Surveillance with TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1092–1103. [[CrossRef](#)]
2. Touzi, R.; Hurley, J.; Vachon, P.W. Optimization of the degree of polarization for enhanced ship detection using polarimetric RADARSAT-2. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5403–5424. [[CrossRef](#)]
3. Novak, L.M.; Sechtin, M.B.; Cardullo, M.J. Studies of target detection algorithms that use polarimetric radar data. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *25*, 150–165. [[CrossRef](#)]
4. Yang, J.; Zhang, H.; Yamaguchi, Y. GOPCE-based approach to ship detection. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 1089–1093. [[CrossRef](#)]
5. Yerey, M.; Campbell, J.W.M.; Mattar, K.; Potter, T. Ocean surveillance with polarimetric SAR. *Can. J. Remote Sens.* **2001**, *27*, 328–344. [[CrossRef](#)]
6. Touzi, R.; Raney, R.K.; Charbonneau, F. On the use of symmetric scatterers for ship characterization. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2039–2045. [[CrossRef](#)]
7. Chen, J.; Chen, Y.L.; Yang, J. Ship detection using polarization crossentropy. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 723–727. [[CrossRef](#)]
8. Touzi, R.; Charbonneau, F. Characterization of target symmetric scattering using polarimetric SARs. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2507–2516. [[CrossRef](#)]

9. Cui, Y.; Zhou, G.; Yang, J.; Yamaguchi, Y. On the iterative censoring for target detection in SAR images. *IEEE Trans. Geosci. Remote Sens.* **2011**, *8*, 641–645. [[CrossRef](#)]
10. An, W.; Xie, C.; Yuan, X. An improved iterative censoring scheme for CFAR ship detection with SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4585–4595.
11. Tao, D.; Anfinson, S.; Brekke, C. Robust CFAR detector based on truncated statistics in multiple-target situations. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 117–134. [[CrossRef](#)]
12. Pappas, O.; Achim, A.; Bull, D. Superpixel-Level CFAR Detectors for Ship Detection in SAR Imagery. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1397–1401. [[CrossRef](#)]
13. Li, T.; Liu, Z.; Xie, R.; Ran, L. 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**, *11*, 184–194. [[CrossRef](#)]
14. Zhou, F.; Fan, W.; Sheng, Q. Ship Detection Based on Deep Convolutional Neural Networks for PolSAR Images. In Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 681–684.
15. Kang, M.; Ji, K.; Leng, X.; Lin, Z. Contextual Region-Based Convolutional Neural Network with Multilayer Fusion for SAR Ship Detection. *Remote Sens.* **2017**, *9*, 860. [[CrossRef](#)]
16. Mairal, J.; Bach, F.; Ponce, J. Task-driven dictionary learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 791–804. [[CrossRef](#)]
17. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [[CrossRef](#)]
18. Alpaydm, E. *Introduction to Machine Learning*; Pitman: London, UK, 1988.
19. Shawe-Taylor, J.; Cristianini, N. *Kernel Methods for Pattern Analysis*; China Machine Press: Beijing, China, 2005.
20. Bradley, D.M.; Bagnell, J.A. Differentiable sparse coding. In Proceedings of the International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–11 December 2008; Curran Associates Inc.: Red Hook, NY, USA, 2008; pp. 113–120.
21. Yang, J.; Yu, K.; Huang, T. Supervised translation-invariant sparse coding. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 3517–3524.
22. Wright, S.; Nowak, R.; Figueiredo, M.A.T. Sparse reconstruction by separable approximation. *IEEE Trans. Signal Process.* **2009**, *57*, 2479–2493.
23. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2011**, *3*, 1–122. [[CrossRef](#)]
24. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202. [[CrossRef](#)]
25. Gao, S.; Tsang, I.W.H.; Ma, Y. Learning category-specific dictionary and shared dictionary for fine-grained image categorization. *IEEE Trans. Image Process.* **2014**, *23*, 623–634. [[PubMed](#)]
26. Sun, X.; Nasrabadi, N.M.; Tran, T.D. Task-driven dictionary learning for hyperspectral image classification with structured sparsity constraints. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4457–4471. [[CrossRef](#)]
27. Liu, S.; Cao, Z.; Yang, H. Information theory-based target detection for high-resolution SAR image. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 404–408. [[CrossRef](#)]
28. Yu, W.; Wang, Y.; Liu, H.; He, J. Superpixel-based CFAR target detection for high-resolution SAR images. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 730–734. [[CrossRef](#)]
29. Lin, H.; Bao, J.; Yin, J.; Yang, J. Superpixel segmentation method with boundary constraints for polarimetric SAR images. In Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 6195–6198.
30. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)] [[PubMed](#)]
31. Song, S.; Xu, B.; Yang, J. Ship Detection in Polarimetric SAR Images via Variational Bayesian Inference. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2819–2829. [[CrossRef](#)]
32. He, J.; Wang, Y.; Liu, H.; Wang, N.; Wang, J. A Novel Automatic PolSAR Ship Detection Method Based on Superpixel-Level Local Information Measurement. *IEEE Trans. Geosci. Remote Sens.* **2018**, *15*, 384–388. [[CrossRef](#)]

33. Lin, H.; Song, S.; Yang, J. Ship Classification Based on MSHOG Feature and Task-Driven Dictionary Learning with Structured Incoherent Constraints in SAR Images. *Remote Sens.* **2018**, *10*, 190. [[CrossRef](#)]
34. Wang, Z.; Nasrabadi, N.M.; Huang, T.S. Semisupervised hyperspectral classification using task-driven dictionary learning with Laplacian regularization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1161–1173. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).