



Article

Oil Tank Detection and Recognition via Monogenic Signal

Yunqing Fan ¹, Junjun Yin ^{1,*} and Jian Yang ²

¹ School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China; m202220868@xs.ustb.edu.cn

² Department of Electronic Engineering, Tsinghua University, Beijing 100084, China; yangjian_ee@mail.tsinghua.edu.cn

* Correspondence: junjun_yin@ustb.edu.cn

Abstract: With the rapid development of synthetic aperture radar (SAR) techniques, satellite systems' capabilities to acquire information are continually improving. The PAZ satellite, with its high resolution and wide scanning swath, can provide high-quality data support for SAR applications. Oil tanks serve as energy storage devices, and their identification holds significant value in both military and civilian fields. Challenges in the detection and recognition of oil tanks using classical methods include poor detection, slow computation speed, and multiple windows of correct recognition. This paper centers on the analysis of oil tanks using PAZ data. We employ a sliding-window approach to acquire candidate target windows, process the windows through Weibull distribution modeling and hole filling, and extract target features using the monogenic signal based on regional L2 norm. The results demonstrate that the proposed method effectively improves the accuracy, and the model exhibits strong generalization ability and robustness.

Keywords: oil tank detection; PAZ data; large-scale SAR images; monogenic signal



Citation: Fan, Y.; Yin, J.; Yang, J. Oil Tank Detection and Recognition via Monogenic Signal. *Remote Sens.* **2024**, *16*, 676. <https://doi.org/10.3390/rs16040676>

Academic Editors: José Fernández, Antonio Turiel, María José González Bonilla, Juan Manuel Cuerda Muñoz and Nuria Casal Vázquez

Received: 31 December 2023

Revised: 5 February 2024

Accepted: 8 February 2024

Published: 14 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Synthetic aperture radar (SAR) is a microwave sensor that actively observes the Earth. SAR emits electromagnetic waves directed towards targets on the Earth's surface. Upon reaching the ground, electromagnetic waves interact with the targets, generating echoes that are influenced by their individual characteristics. As electromagnetic wave transmission takes time, SAR has the capability to gather reflections at different positions and form synthetic apertures. These apertures are processed in both distance and azimuth to generate high-resolution ground images [1]. SAR exhibits a high degree of penetrating capability and the ability to conduct continuous monitoring in all weather conditions for a 24 h duration. Consequently, synthetic aperture radar (SAR) has been successfully applied in various fields.

SAR image target detection is the process of determining whether a given image contains one or more targets of interest and determining the position of each predicted target in the image. As the primary task of SAR image processing, target detection is a crucial step in rapidly and accurately extracting SAR image information. This holds significant research importance and offers broad application prospects. Since SAR images acquire a large scene in a single observation, the steps for SAR target detection are as follows: the first is to extract the regions of interest in the large-scale image; the second is to remove the regions that do not contain the target; finally, use precise algorithms to process regions that may contain targets. At present, SAR image target detection has become a prominent focus of global research.

Classic target detection algorithms have been developed to process SAR images with increasingly high resolution. Huertas et al. [2] assumed that buildings are rectangular or composed of rectangular components, and they designed shape models accordingly to detect buildings. Kim et al. [3] introduced the concept of local variance maps, which

reveal the spatial structure of images and help determine the optimal segmentation size for corresponding areas. Weber et al. [4] introduced a new definition of the hit-or-miss transform (HMT) for multivariate image analysis and used it as a template matching operator for coastline extraction and oil tank detection. Stankov et al. [5] generated gray-scale maps from multispectral images and applied HMT to achieve template matching for building detection.

The significance-based candidate region target detection method relies on feature extraction. Currently, commonly used features include not only basic geometric and texture information and the distribution of oil tanks, but also advanced operators such as histogram of oriented gradients (HOG) [6], scale-invariant feature transform (SIFT) [7], and Zernike moments [8]. After extracting features, various classifiers can be used for training to minimize classification errors. These classifiers include support vector machines (SVMs) [9], K-nearest neighbors (KNNs) [10], sparse representation classification (SRC) [11], etc. Although classic algorithms have achieved some results, they face challenges in target detection applications, such as poor detection, slow computation speed, and multiple windows of correct recognition.

As important energy storage devices with a special appearance, oil tanks often appear in areas with high risk [12]. The location information is valuable for conducting energy analysis, urban planning, and assessing damage from hazard sources. Meanwhile, oil tanks serve as directional or landmark objects and play an important role in providing navigation for individuals, vehicles, and ships. Therefore, research on oil tank target detection has significant practical value.

It is well known that the diverse types and densely packed arrangement of oil tanks bring more challenges for detection and recognition. In SAR images, the strong scattering characteristics of oil tanks are shown as multiple consecutive dense scattering points, leading to an unsmooth contour. The varied scattering distribution of diverse roofs results in significant differences in oil tanks. The distribution of small-size oil tanks is numerous and densely concentrated, resulting in a connected contour of different targets.

In order to better apply SAR in oil tank detection, it is crucial to accurately interpret SAR images. We focus on extracting monogenic signals for SAR oil tank detection. Felsberg and Sommer provided the formulation of the two-dimensional (2-D) Riesz transform and defined the monogenic signal [13]. The method has the capability to extract information in the spatial and frequency domains. Felsberg and Sommer [14] used a phase-based image processing method to generate the monogenic signal space. Huang et al. [15] utilized local binary encoding to encode monogenic signal components from different scale spaces, thereby generating histograms for biometric recognition. Dong et al. [16] introduced the monogenic signal into target recognition in SAR images and established an augmented monogenic feature vector by uniformly down-sampling, normalization, and concatenation of the monogenic components. Dong et al. [17] used log-Gabor filters to maintain characteristics of the Riesz transform, and the resulting feature vector was input for the classifiers to make inferences.

The monogenic signal usually adopts the method of down-sampling to reduce feature dimensionality, but this can cause information loss. We choose to calculate the L2 norm of each region to achieve the same dimensionality reduction effect while reducing information loss. This conjecture has been validated using the classified dataset of oil tanks. In this paper, we perform oil tank detection on Zhoushan data acquired by the PAZ satellite using the sliding windows approach. The training and test sets are manually designed based on the areas of interest and non-interest within a dense region of oil tanks in the image. Preprocessing methods involving Weibull distribution modeling and hole filling are applied. Subsequently, monogenic signal features are extracted using the regional L2 norm. Finally, they are input into the SRC for recognition. The results indicate that the proposed method has a high accuracy and low missed detection rate on this dataset.

The main contributions of this paper can be summarized as follows:

- (1) In order to minimize the redundancy of the monogenic signal and improve the algorithm's execution efficiency, we propose a feature dimension reduction method based on the regional L2 norm. This method aims to achieve a better balance between recognition accuracy and feature extraction time.
- (2) There are differences in the performance of various types of roofs on oil tanks, which makes model recognition difficult. To address the differences in strong scattering points between different roofs and backscenes, we propose using Weibull distribution modeling and the hole-filling method.
- (3) In order to enhance the model's adaptability for PAZ oil tank detection, we constructed a dataset for oil tank recognition using the Zhoushan data. The dataset includes large-size oil tanks, small-size oil tanks, connected oil tanks, and negative samples. The small-size and connected oil tanks serve as auxiliary positive samples for large-size oil tanks and contribute to the calculation of sparse matrices in SRC. The method can improve the detection rate and decrease the false detection rate.

The remaining sections of this paper are organized as follows. The description and improvement of the algorithm are introduced in Section 2. The schematic illustration of the proposed methodology is provided in Section 3. The datasets and experiments are displayed in Section 4, discussed in Section 5, and finally concluded in Section 6.

2. Materials and Methods

2.1. Weibull Distribution

The two-parameter Weibull distribution has the following probability density function [18]:

$$p(x) = \frac{\beta}{\theta^\beta} x^{\beta-1} e^{-\left(\frac{x}{\theta}\right)^\beta}, x > 0, \theta, \beta > 0 \quad (1)$$

where β is the shape parameter and θ is the scale parameter. When $\beta = 1$, the Weibull distribution is the exponential distribution or single-look gamma distribution. When $\beta = 2$, the Weibull distribution is the Rayleigh distribution or single-look Nakagami distribution. Therefore, the Weibull distribution can be used to fit the intensity data of single-look SAR images.

Weibull's parameter estimation can be done by using the maximum likelihood method. Assuming that $X = \{x_1, x_2, \dots, x_n\}$ are independently and identically distributed samples obeying Weibull's distribution, the shape parameter β can be derived from the following equation:

$$\frac{1}{\beta} = \frac{\sum_{i=1}^n x_i^\beta \ln(x_i)}{\sum_{i=1}^n x_i^\beta} - \frac{1}{n} \sum_{i=1}^n \ln(x_i) \quad (2)$$

The above equation can be solved numerically or given graphically. The estimate of the scale parameter θ is given by the following equation when $\hat{\beta}$ is found:

$$\hat{\theta} = \left\{ \frac{1}{n} \sum_{i=1}^n x_i^{\hat{\beta}} \right\}^{\frac{1}{\hat{\beta}}} \quad (3)$$

2.2. Analytic Signal

The analytic signal is derived from the Hilbert transform. For a given real signal $f(x)$, the expression of its Hilbert transform $f_H(x)$ in the time domain is [19]:

$$f_H(x) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(u)}{1-u} du = f(x) * \frac{1}{\pi x} \quad (4)$$

equivalent to convolving the original signal with a kernel of $h(x) = 1/\pi x$. Due to $1/\pi x \Leftrightarrow -j\text{sgn}(f)$, the expression of the Hilbert transform in the frequency domain is:

$$H(f) = -j\text{sgn}(f) = \begin{cases} -j & f > 0 \\ 0 & f = 0 \\ j & f < 0 \end{cases} \quad (5)$$

The Hilbert transform has the characteristics of antisymmetry and zero direct current (DC) components.

For a given real signal $f(x)$, the analytic signal $f_A(x)$ is defined as [20]:

$$f_A(x) = f(x) - if_H(x) \quad (6)$$

It can be concluded that the analytical signal has the following properties:

- (1) The energy of the analytical signal is doubled.

$$\int \|f_A(x)\|^2 dx = \int f^2(x) + \|f_H(x)\|^2 dx = 2 \int f^2(x) dx \quad (7)$$

- (2) It can be decomposed based on the real and imaginary parts in polar coordinates to represent local amplitude $A(\mathbf{x})$ and local phase $\varphi(\mathbf{x})$.

$$\begin{aligned} A(x) &= \|f_A(x)\| = \sqrt{f^2(x) + f_H^2(x)} \\ \varphi(x) &= \arctan2(f_H(x), f(x)), \varphi(x) \in [0, 2\pi) \end{aligned} \quad (8)$$

which contain the local energy and structure information.

2.3. Monogenic Signal

The 2-D Riesz transform [21] $f_R(\mathbf{x})$ is obtained by convolving the 2-D signal $f(\mathbf{x})$ with the 2-D Hilbert transform $h_2(x)$. The Riesz transform preserves the specifics of the Hilbert transform. The provided expression is:

$$f_R(\mathbf{x}) = -\frac{\mathbf{x}}{2\pi|\mathbf{x}|^3} * f(\mathbf{x}) \stackrel{\text{def}}{=} h_2(\mathbf{x}) * f(\mathbf{x}). \quad (9)$$

The monogenic signal $f_M(\mathbf{x})$ is defined as a linear combination of $f(\mathbf{x})$ and $f_R(\mathbf{x})$:

$$f_M(\mathbf{x}) = f(\mathbf{x}) - (i, j)f_R(\mathbf{x}) \quad (10)$$

where $[1, i, j]$ form an orthogonal basis in \mathbb{R}^3 space. The monogenic signal features can be represented in local amplitude $A(\mathbf{x})$, local phase $\varphi(\mathbf{x})$, and local orientation $\theta(\mathbf{x})$ in an ideal manner.

$$\begin{cases} A(\mathbf{x}) = \sqrt{f(\mathbf{x})^2 + |f_R(\mathbf{x})|^2} \\ \varphi(\mathbf{x}) = \arctan2(|f_R(\mathbf{x})|, f(\mathbf{x})) \in (-\pi, \pi] \\ \theta(\mathbf{x}) = \arctan2(f_{R2}(\mathbf{x})/f_{R1}(\mathbf{x})) \in (-\frac{\pi}{2}, \frac{\pi}{2}] \end{cases} \quad (11)$$

where $f_{R1}(\mathbf{x})$ and $f_{R2}(\mathbf{x})$ represent the imaginary components. $A(\mathbf{x})$ contains the local energy information. $\varphi(\mathbf{x})$ and $\theta(\mathbf{x})$ correspond to local structure and geometric information [22].

The monogenic signal retains the characteristics of the analytical signal. The information component is obtained through orthogonal decomposition, which possesses important characteristics such as being insensitive to size and direction. In applications for SAR target detection, it is possible to generate more comprehensive feature descriptions based on it.

2.4. Log-Gabor Bandpass Filter

Field [23] introduced a high-pass log-Gabor filter, which can integrate spatial and frequency domain information. The log-Gabor function maintains the antisymmetry and

zero DC components of the Riesz transform and can perform multi-scale representations of signals [24]. The function expression is:

$$G(\omega) = \exp(-(\log(\omega/\omega_0))^2/2(\log(\sigma/\omega_0))^2) \quad (12)$$

where ω_0 represents the central frequency and σ/ω_0 determines the shape. The monogenic signal convolves with the log-Gabor kernel h_{lg} to generate the multi-scale monogenic space:

$$f_M(\mathbf{x}) = (h_{lg} * f)(\mathbf{x}) - (i, j)(h_{lg} * f_R)(\mathbf{x}) \quad (13)$$

2.5. Regional L2 Norm

For a SAR image, its monogenic signal space of scale 3 is:

$$\left\{ \underbrace{A_1, \varphi_1, \theta_1}_{I_M^1}, \underbrace{A_2, \varphi_2, \theta_2}_{I_M^2}, \underbrace{A_3, \varphi_3, \theta_3}_{I_M^3} \right\}. \quad (14)$$

where I_M^1 , I_M^2 and I_M^3 represent the monogenic signal at various scales, with each component conveying three layers of information at the given scale.

The dimension of the feature at each scale is 3 times longer than that of the original signal. We choose to extract features from the monogenic signal space at a scale of 3. Using feature vectors that are 9 times longer will lead to a reduction in algorithmic efficiency. We propose calculating the L2 norm of each region [25]. This approach aims to decrease feature dimensionality while preserving the information present in the monogenic signal.

Similar to the down-sampling method of extracting information according to the sampling factor n , the feature vector is derived by calculating the L2 norm of $n \times n$ pixels in the region and subsequently combining the results from all regions. The regional L2 norm method has the same degree of dimensionality reduction as the down-sampling method while avoiding loss of information. The monogenic signal in a 12×12 SAR image has a feature vector length of $12 \times 12 \times 9$. When $n = 2, 3, 4$, as in the regional division results shown in Figure 1, the feature dimension of a single component will be reduced to 6×6 , 4×4 , and 3×3 .

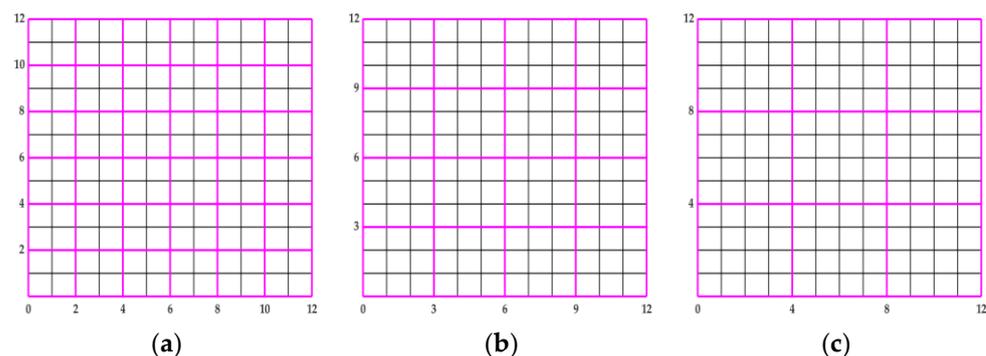


Figure 1. Region division of the single component at different sampling factors (n): (a) 6×6 regions at $n = 2$; (b) 4×4 regions at $n = 3$; (c) 3×3 regions at $n = 4$.

We represent the elements of a region in matrix form as A , and its L2 norm $\|A\|_2$ is the maximum eigenvalue. The L2 norm is also called the spectral norm of a matrix.

$$\|A\|_2 = \max(\text{svd}(A)) = \sqrt{V_{\max}(A^H A)} \quad (15)$$

where $\text{svd}()$ represents the process of singular value decomposition, A^H represents the conjugate transpose of A , and $V_{\max}()$ is used to calculate the maximum singular value.

3. Illustration of the Proposed Features

This section displays the schematic illustration of the proposed method and explains the principle. The oil tank shown in Figure 2A has strong intensity and uneven distribution of scattering points. The schematic diagram of the monogenic signal scale space represented in Equation (14) is shown in Figure 2B.

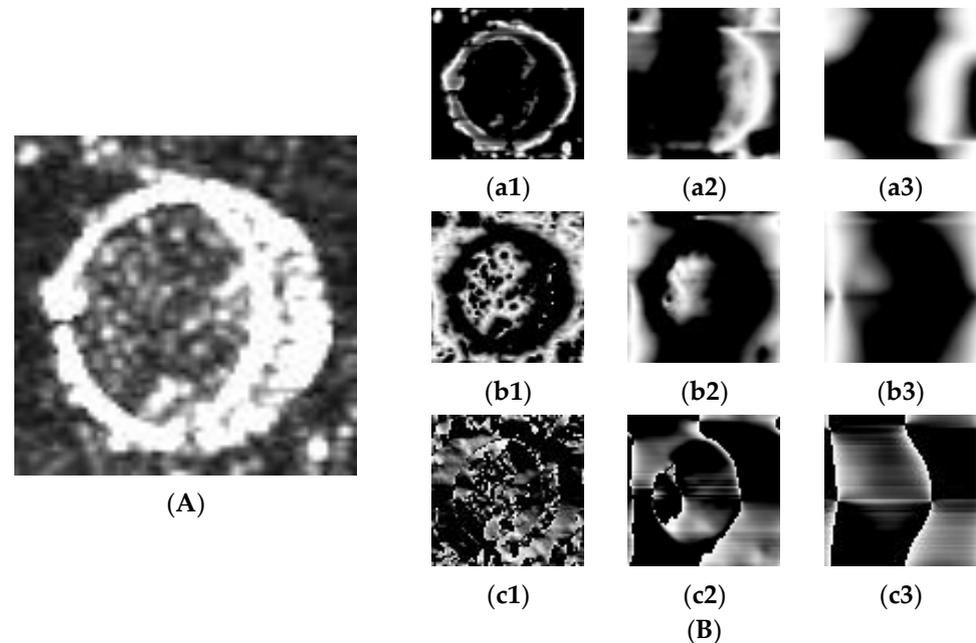


Figure 2. Schematic illustration of the monogenic signal. (A) An oil tank target; (B) the target monogenic space at scale 3; (a1,a2,a3) $A(x)$ at scale 1, 2, 3; (b1,b2,b3) $\varphi(x)$ at scale 1, 2, 3; (c1,c2,c3) $\theta(x)$ at scale 1, 2, 3.

The local amplitude is shown in Figure 2(a1) as an example, the dimensionality reduction results of the regional L2 norm and the down-sampling are shown in Figure 3. Under the same sampling factor n , the former retains more complete information in Figure 3b,c. As the degree of dimensionality reduction increases, the regional L2 norm can still maintain the shape of the target.

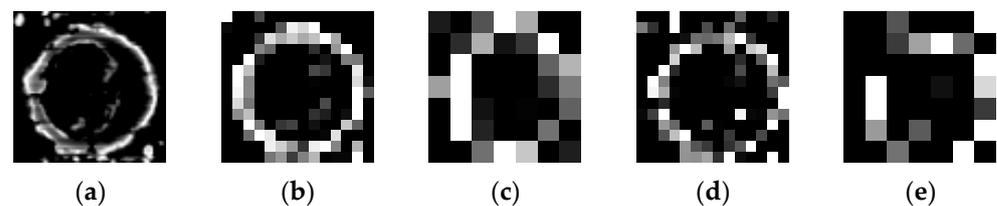


Figure 3. Schematic illustration of dimensionality reduction. (a) $A(x)$ at scale 1; (b) result of the regional L2 norm at $n = 5$; (c) result of the regional L2 norm at $n = 10$; (d) result of the down-sampling at $n = 5$; (e) result of the down-sampling at $n = 10$.

4. Experiments and Results

To evaluate the effectiveness of the regional L2 norm feature extraction method based on the monogenic signal for detecting oil tanks, recognition experiments were performed on the MSAR-1.0 dataset. Subsequently, a comparative analysis was conducted with other methods for dimensionality reduction. Then we manually constructed a training set using PAZ data to conduct detection experiments.

4.1. Oil Tank Recognition Experiments

4.1.1. MSAR-1.0 Dataset and Setting

The large-scale multi-class SAR image target detection dataset-1.0 (MSAR-1.0) originates from data collected by the Hisea-1 and Gaofen-3 satellites, including a total of 28,449 detection slices. Polarization methods encompass HH, HV, VH, and VV. The dataset encompasses various scenarios, such as airports, harbors, nearshore, islands, distant seas, and urban areas, and consists of 1851 bridges, 39,858 ships, 12,319 oil tanks, and 6368 aircraft. The majority of slices are 256×256 pixels, and some bridge slices are 2048×2048 pixels. According to the distribution characteristics of multi-class SAR targets in multi-scene in the dataset, we design recognition experiments to verify the effectiveness of the algorithm for oil tank targets.

Four class targets are segmented from the labeled samples in the dataset and are subsequently resized into 60×60 pixels using bilinear interpolation. The sliced images of oil tanks of different sizes and types are shown in Figure 4. The size of the oil tank is generally judged by its volume. In experiments, we use the diameter of the circular structure of the oil tank to determine its size, and the scattering characteristics exhibited by the oil tank to determine whether it is a floating-roof oil tank or a fixed-roof oil tank.

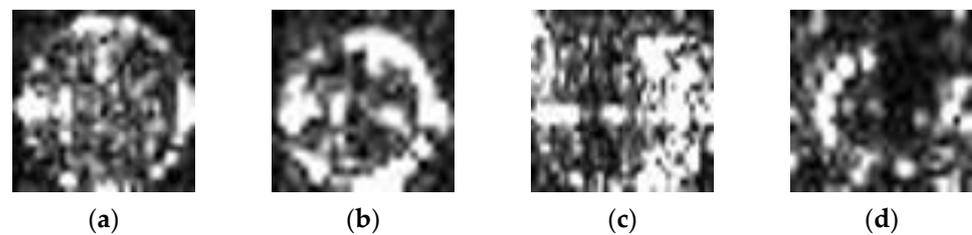


Figure 4. Oil tanks of different sizes and types in the MSAR-1.0 dataset: (a) large-size floating-roof tank; (b) small-size floating-roof tank; (c) large-size fixed-roof tank; (d) small-size fixed-roof tank.

As shown in Figure 4a,b, the circular structure of the target is presented in the image by the form of intensity, and the electromagnetic signals reflected from the top and bottom edges of the oil tanks form connected intensity information. As shown in Figure 4c,d, the presence of the fixed roof may obscure certain edge information, resulting in an incomplete circular structure of the target. The edge information of the oil tank in (c) appears more intense due to variations in the angles of satellite incidence.

Furthermore, the dataset exhibits instances of overexposure, darkness, and missing targets, as shown in Figure 5. The target in Figure 5a can still observe the structure of its oil tank, but the sample and the surrounding scattering intensity exhibit excessive strength. The target shown in Figure 5b exhibits limited scattering points and provides little information. To reduce the potential impact of these samples on the experiment, we set a threshold to filter them for comparative experiments.

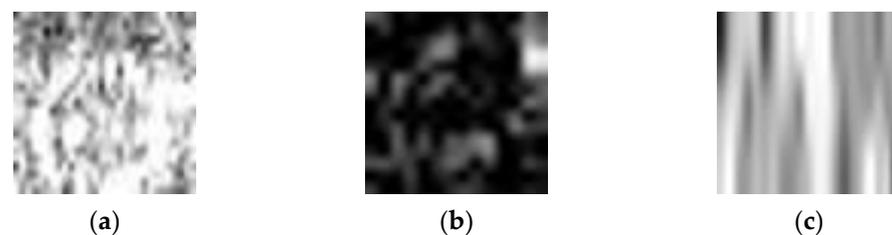


Figure 5. (a) Overexposure situation; (b) darkness situation; (c) missing target situation.

The dataset is divided into the training set and the test set in the ratio of 7:3. We set the parameters of the filter as follows: the minimum wavelength set to 12, the ratio of sigma to the center frequency set to 0.28, the multiplier factor set to 11.2, and the scale set to 3. We set the dimensionality reduction factor for the regional L2 norm to 10.

4.1.2. Evaluation Metric

We used widely used metrics, such as accuracy and precision, to evaluate the performance of the recognition model.

Accuracy is the most commonly used classification performance indicator. In general, a high accuracy of the model suggests excellent performance. The calculation is done by dividing the number of correctly recognized samples by the total number of samples.

$$Accuracy = \frac{TP}{Total} \quad (16)$$

where TP is the number of correctly predicted samples by the model and $Total$ is the overall number of test samples.

Precision is an important measure of a model's ability to recognize a particular class of targets. This metric is computed by dividing the number of correctly identified samples in a class by the total number of samples in the class.

$$Precision = \frac{TP_x}{Total_x} \quad (17)$$

where x is a specific category of focus and TP_x and $Total_x$ are the number of correctly predicted samples and test samples belonging to this category.

The average time required for feature extraction of an image is a crucial factor for evaluating the effectiveness of methods. The duration is from read-in to output features.

4.1.3. Recognition Result

Experiments were conducted on three classifiers: KNN, SRC, and SVM. The comparison results with several classic methods are presented in Table 1. The comparison results after threshold filtering are shown in Table 2. The classical algorithm results under specific classifiers given in Tables 1 and 2 are the best.

Table 1. Comparison of accuracy in the MSAR1.0 dataset on different methods.

MSAR1.0	Classifier	Accuracy (%)	Oil Tank Precision (%)
Proposed method	KNN	99.17	98.92
	SRC	98.77	98.21
	SVM	92.81	92.29
Monogenic signal	SRC	98.27	97.19
Zernike moment	SRC	97.97	97.05
PCA	KNN	97.78	97.97

Table 2. Comparison of accuracy in the threshold-filtering MSAR1.0 dataset on different methods.

MSAR1.0	Classifier	Accuracy (%)	Oil Tank Precision (%)
Proposed method	KNN	99.44	99.33
	SRC	98.75	97.81
	SVM	96.72	97.14
Monogenic signal	SRC	98.35	97.90
Zernike moment	SRC	98.56	97.90
PCA	KNN	98.17	98.57

The proposed method gains high accuracy. Although the improvement on the threshold-filtered dataset was not significant, the model has great enhancement in the precision of the oil tank we are concerned about. The presented outcome demonstrates the effectiveness of the threshold-filtering method in the recognition of oil tanks. In the

proposed method, KNN and SRC have high classification accuracy for recognizing oil tanks, making them the preferred classifiers for detection experiments.

To assess our feature dimensionality reduction method, we designed comparative experiments with down-sampling [26] and random projection [27] algorithms under SRC in the threshold-filtering MSAR1.0 dataset. The comparative experimental results are shown in Table 3.

Table 3. Comparative experimental results of feature dimensionality reduction methods under SRC.

Dimensionality Reduction Methods	Feature Dimension	Accuracy (%)	Average Time (s)
Regional L2 norm	900	99.02	0.0061
Down-sampling	900	97.56	0.0022
Random projection	1039	98.37	0.4021

In the experiment, the monogenic signal feature dimension of an image is 32,400. At close feature dimensions, the regional L2 norm expresses the highest accuracy. Down-sampling has the lowest average computation time but causes information loss. Random projection utilizes all information but requires a long time. The regional L2 norm reduces information loss, while also ensuring a shorter time. The results indicate that the regional L2 norm dimensionality reduction method based on monogenic signal features can effectively balance accuracy and computation average time.

4.2. Oil Tank Detection Experiments

4.2.1. PAZ Data and Constructed Dataset

The PAZ satellite is a high-resolution X-band SAR. It operates in the same orbit as the twin satellites TerraSAR-X and TanDEM-X, and the three satellites work together as a constellation. The PAZ satellite adopts beam focusing and strip imaging modes and polarization modes of HH and VV. The data were obtained on 30 November 2019 in the Zhoushan Port area of China, the pixel size is 1.5×1.5 m, the range resolution is 3.1 m, the azimuth resolution is 3.5 m, and the incidence angle is 51.1° . The original image size of this data is $12,333 \times 10,666$, and it encompasses intensity information acquired under HH and VV polarization. The PAZ products adopt geocoded ellipsoid corrected (GEC) to represent intensity information. The products are typically used to apply and analysis of processed images for target detection with position information in Earth coordinates [28].

The PAZ satellite acquired other data on 9 November 2019, in the Zhoushan Port area of China. The pixel size is 2.75×2.75 m, the range resolution is 6.0 m, the azimuth resolution is 6.1 m, and the incidence angle is 43.2° . The original image size is $19,818 \times 10,363$, and it also contains intensity information acquired under HH and VV polarizations. Both images have been rectified through geocoding, with invalid black edges around the data. The effective information is extracted through mask processing and then multiplied by $5E5$ to generate the experimental data. The sizes of the two images after information enhancement are 9845×8988 and $18,898 \times 8897$, respectively.

In order to realize oil tank detection and evaluate the effectiveness of the proposed method, we intend to construct the dataset on images of the Zhoushan data. The Zhoushan data include strong scattering targets such as oil tanks, harbors, ships, containers, and buildings with metal structures. The dataset contains regions with dense oil tanks, as shown in Figure 6.

Oil tanks with diameters equal to or exceeding 80 m are considered large-size positive samples. Oil tanks with diameters ranging from 30 to 80 m are deemed small-size positive samples. Oil tanks with diameters below 30 m, which generally have connected, indistinct structures showing a cluster of irregularly scattering points, are regarded as connected positive samples. Other regions with high scattering intensity are classified as negative samples.



Figure 6. Oil tank dense area in the Zhoushan data.

In order to enhance the model's ability to locate oil tanks, we only preserve the areas that encompass complete oil tanks in both large and small-size samples. The method of judgment is to determine whether the center of the connected domain is at the central part of the image. For connected oil tanks, we manually judge and process regions containing at least one oil tank as positive samples. In the case of negative samples, we choose regions containing strong scattering distributions, such as ports, ships, walls, buildings, and roads. The number of these samples is 910, 1827, 1440, and 5176 with a size of 70×70 pixels.

The dataset contains samples from diverse and complex backgrounds to improve the model's generalization capability. The samples under each class are displayed in the red box in Figure 7. Several sample images are shown in Figure 8. In addition, to enhance the robustness of the model, strong scattering targets from other regions are added as negative samples to participate in the learning process. The total number of negative samples reaches 7316.

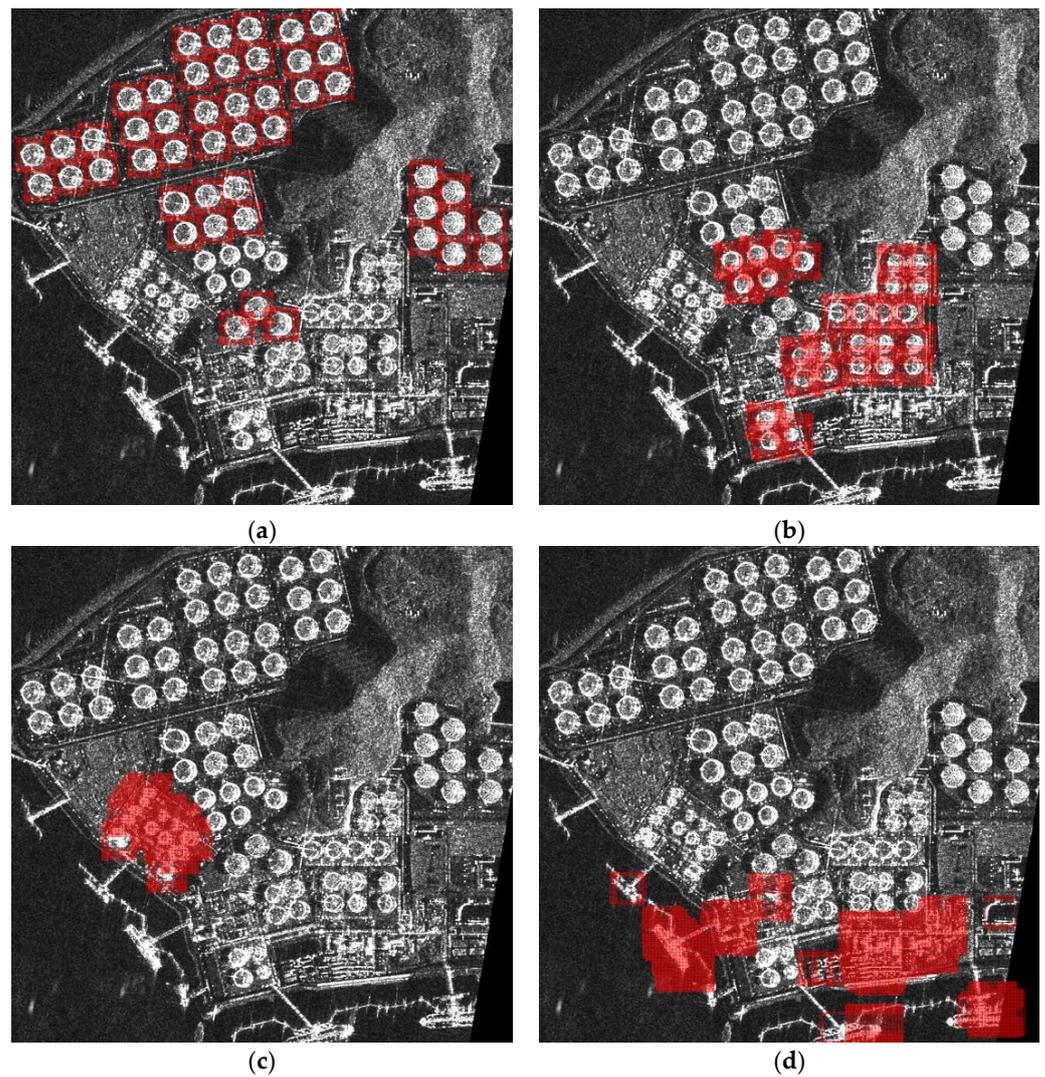


Figure 7. Sample regions under different classes: (a) large-size positive sample region; (b) small-size positive sample region; (c) connected positive sample region; (d) negative sample region.

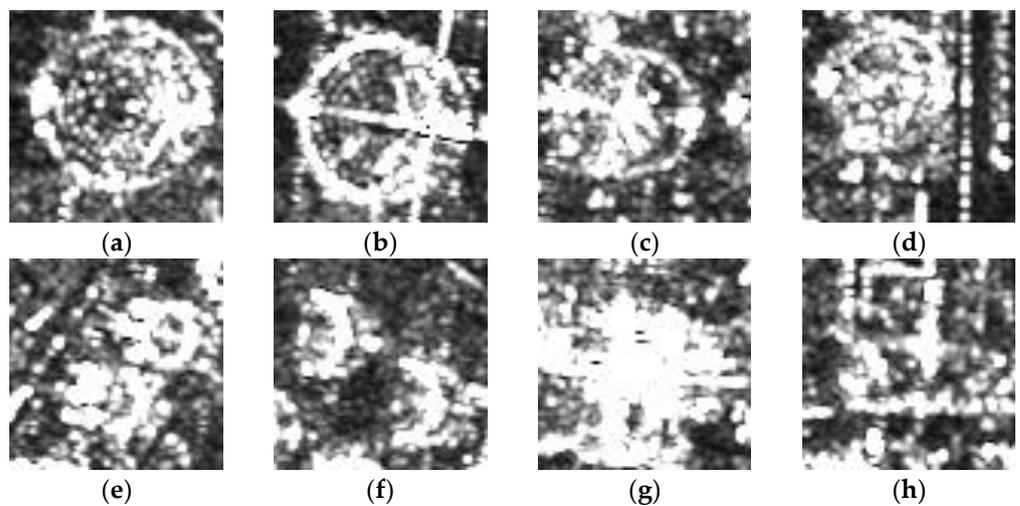


Figure 8. Sample examples of the constructed dataset: (a,b) large-size positive samples; (c,d) small-size positive samples; (e,f) connected positive samples; (g,h) negative samples.

Due to the difference in scattering intensity of the roof material of the oil tank, some oil tanks only exhibit scattering characteristics through the circular structure of the tank top and bottom; other oil tanks' floating-roof structures also exhibit scattering characteristics. This discrepancy can have an adverse impact on the model's training and recognition performance. Figure 8d,f depict fixed-roof or inner floating-roof oil tanks, which exhibit similar appearances. The structural information of these samples is often incomplete. Consequently, we adopt a method based on Weibull distribution modeling and hole filling to narrow the difference between the two situations. Eventually, most oil tanks exhibit similar scattering characteristics.

4.2.2. Dataset Evaluation

To evaluate the validity of the dataset, the data are partitioned into training and testing sets using a 7:3 ratio. The experimental parameter settings were kept consistent with those of the MSAR 1.0 recognition experiment, except that the dimensionality reduction factor of the regional L2 norm was adjusted to 5. The confusion matrices resulting from the experiments conducted on KNN and SRC classifiers are shown in Tables 4 and 5.

Table 4. Confusion matrix of the proposed method under the KNN classifier in the constructed dataset.

Constructed Dataset	Large	Small	Connected	Negative
Large	273	0	0	24
Small	0	430	0	6
Connected	0	0	546	15
Negative	0	2	3	2150
Total	273	432	549	2195
Precision (%)	100.00	99.54	99.45	97.95

Table 5. Confusion matrix of the proposed method under the SRC classifier in the constructed dataset.

Constructed Dataset	Large	Small	Connected	Negative
Large	272	0	0	0
Small	0	432	0	2
Connected	0	0	542	5
Negative	1	0	7	2188
Total	273	432	549	2195
Precision (%)	99.63	100.00	98.72	99.68

The proposed method behaves well in the constructed dataset. The precision of large-size positive samples is high and the false positive rate for negative samples is 0 under SRC. Therefore, we try to conduct experiments for oil tank detection using sliding windows on the Zhoushan dataset under SRC.

4.2.3. Evaluation Metric

We choose to use statistical indicators to evaluate the performance of the detection model, such as the number of true detected, missed detected, and false detected samples.

The count of true detected samples is determined by calculating the number of detection windows containing complete positive samples. The definition of detection rate is:

$$\text{Detection rate} = \frac{TD_x}{Total_x} \quad (18)$$

where x is a category, TD_x denotes the number of true detected samples, and $Total_x$ is the number of samples within this category.

We determine the count of missed detected samples by calculating the number of positive samples that were not detected. The missed detection rate is defined as follows:

$$\text{Missed detection rate} = \frac{MD_x}{Total_x} \quad (19)$$

where MD_x denotes the number of missed detected samples.

The count of false detected samples is the number of negative samples that the model recognizes as positive samples. This metric measures the extent to which the model incorrectly identifies negative samples.

4.2.4. Detection Results

The Zhoushan dataset is a large-scale SAR image. In order to improve the execution efficiency of the sliding window, the threshold-filtering method only considers the window with a mean value of 1.5 or higher. The size of the sliding window is set to 70×70 pixels with a step size of 10 pixels.

The method obtained 30,606 detection windows. Nevertheless, for the positive samples of the constructed dataset, which are complete oil tanks, it is less effective in detecting small-size and connected oil tanks. We propose utilizing these two classes as auxiliary positive samples, which are only involved in computing the sparse matrices of the SRC and not participating in the final classification of the results.

The sliding window detection results suffer from the problem of having multiple correct detection windows for the same target. We use points to represent the position of the candidate window, and a 30×30 region surrounding each point is weighted to obtain the unique point within this region. The windows reconstructed based on the points of each region are the correct detected results. The final result on the Zhoushan dataset is shown in Figure 9. The red box is the positive sample obtained from the model.

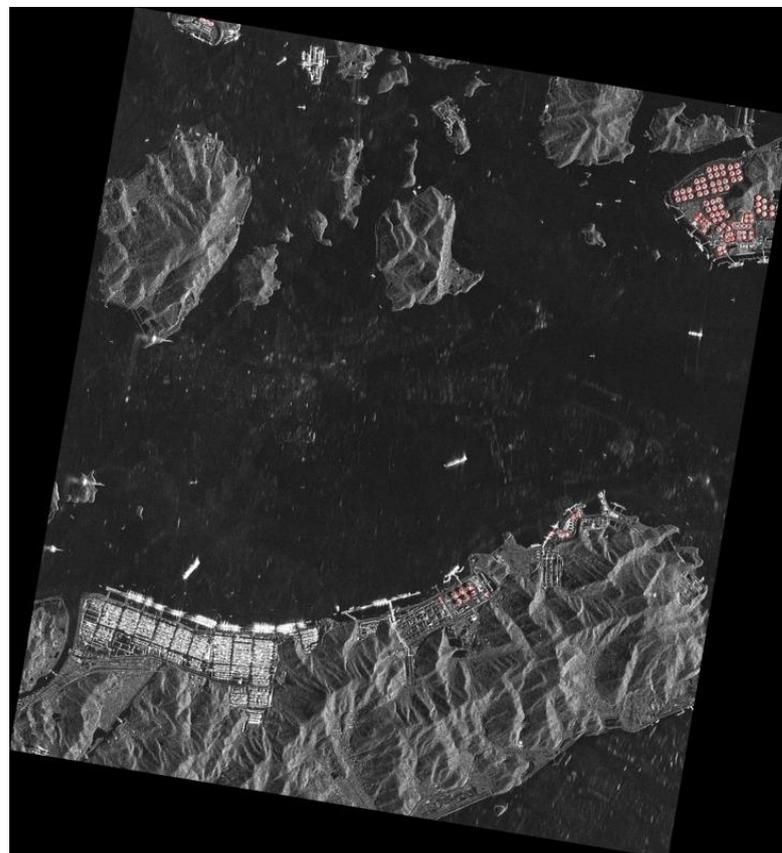


Figure 9. The detection result in a large-scale SAR image of Zhoushan data.

The detection results include 96 windows with positive samples predicted. Out of all the windows, 88 windows contain complete oil tanks, while the other 8 windows are false detected. The overall accuracy of the model is 91.67% obtained from the percentage of windows correctly detected. We classified the results into large-size, small-size, and connected oil tanks. The statistical indicators of the result are shown in Table 6.

Table 6. The statistical indicators of detection result in the Zhoushan dataset.

Zhoushan Data	Large	Small	Connected
Positive	53	31	14
Ture detected	53	26	10
Missed detected	0	5	4
Detection rate (%)	100.00	83.87	71.43
Missed detection rate (%)	0.00	16.13	28.57

We selected four areas containing oil tank targets to show the oil tank detection results more clearly, as shown in Figure 10.

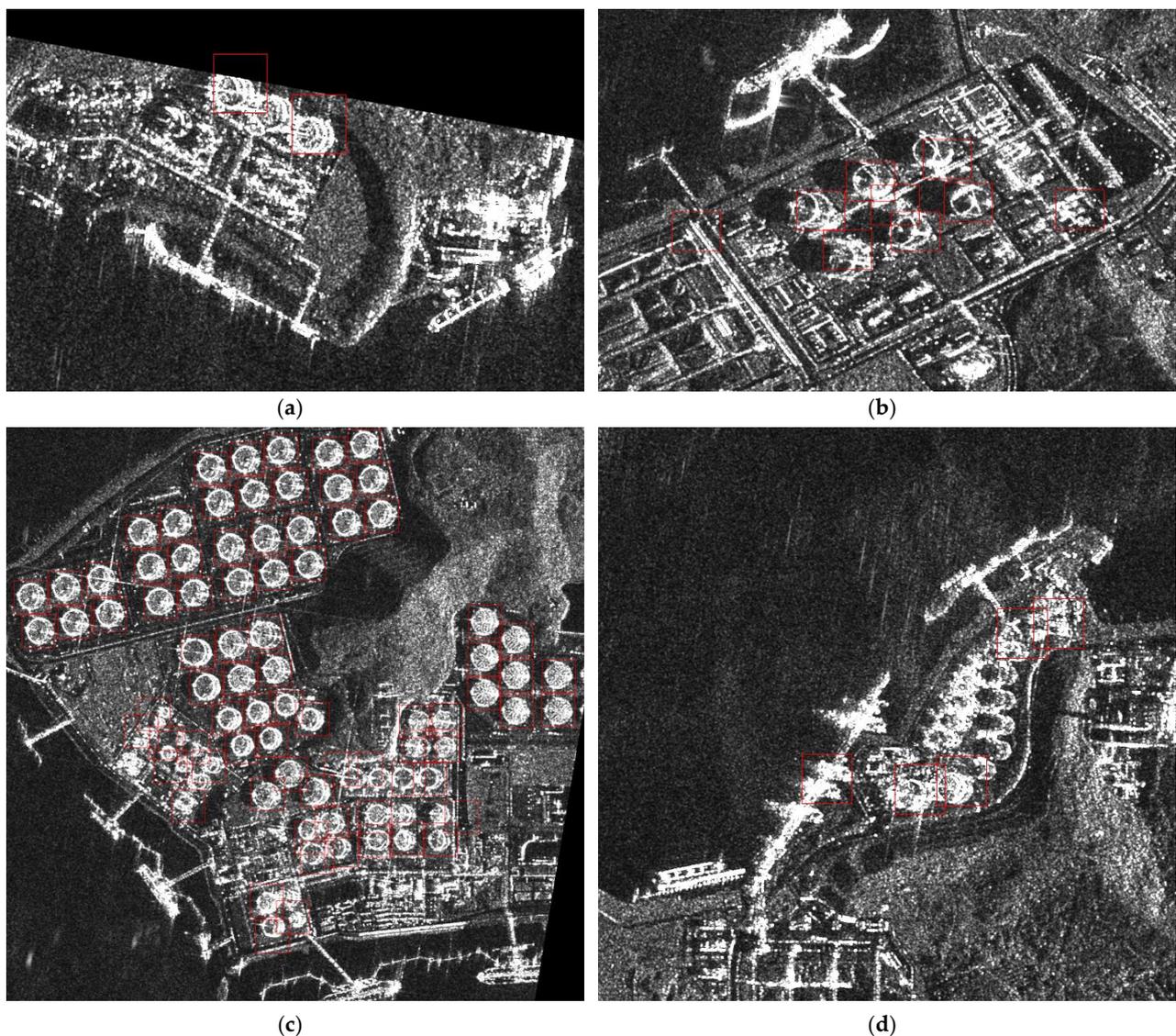


Figure 10. The detection result of four areas in Zhoushan data: (a) data edge oil tank area; (b) fixed-roof oil tank area; (c) densely packed oil tank area; (d) near port oil tank area.

We used other classic methods for sliding window detection, and the most effective approaches were HOG and Zernike moments under SRC. The area shown in Figure 10c has multiple detection windows, including 44 large-size oil tanks, 26 small-size oil tanks, and 10 connected oil tanks. The detection results of HOG and Zernike moments in this area are shown in Figure 11.

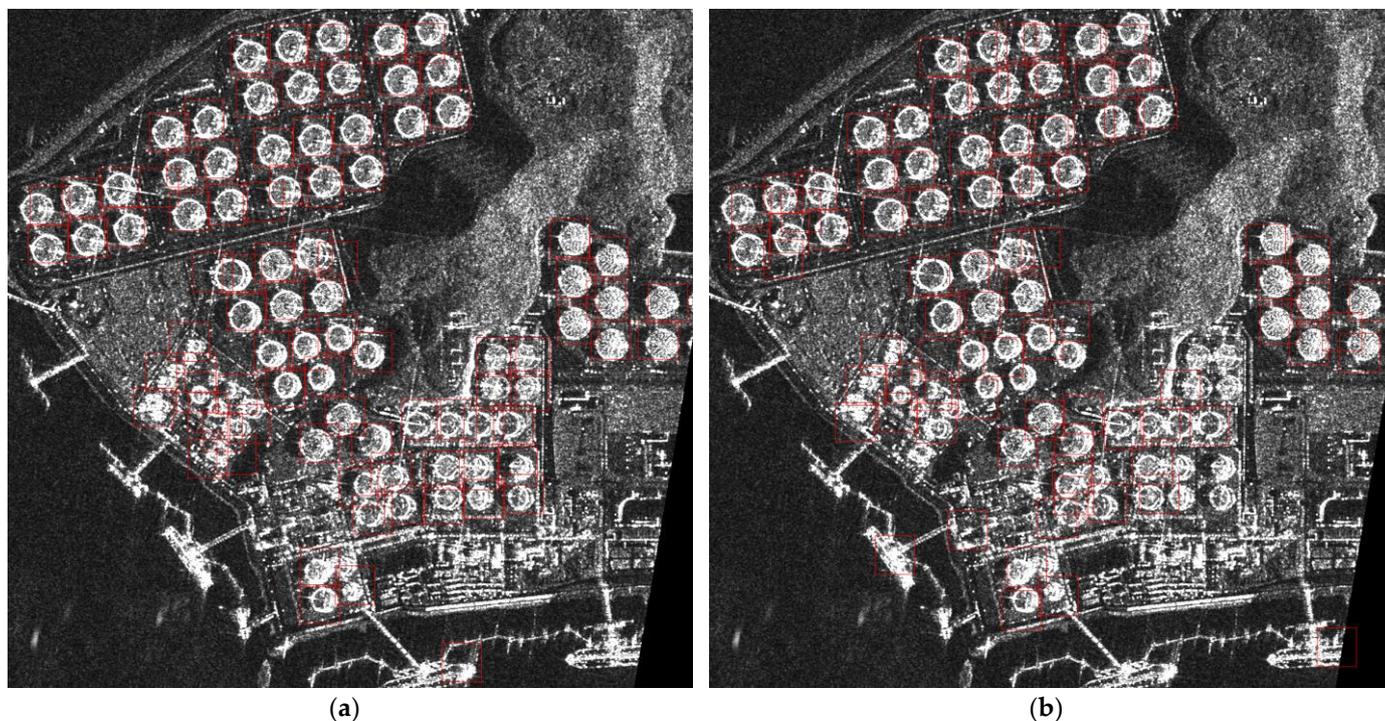


Figure 11. The comparison results in the densely packed oil tank area. (a) HOG detection result; (b) Zernike moment detection result.

To demonstrate the detection performance of the methods, Table 7 shows the comparative results in the densely packed oil tank area.

Table 7. The statistical indicators of detection result in the Zhoushan dataset.

Densely Packed Oil Tanks	Number of Correct Windows	Accuracy (%)	Missed Detection Rate (%)
Proposed method	80	95.00	7.50
HOG	89	85.39	7.50
Zernike moment	83	73.49	26.25

The proposed detection model obtains the highest accuracy with the least number of correct windows. The true detection windows shown in Figure 10c all contain complete targets. There is a problem in Figure 11 that some windows contain incomplete targets, especially evident in Figure 11b. The results indicate that the proposed method has stronger localization ability compared to HOG and Zernike moments.

4.2.5. Robustness Validation Experiments

Since some of the testing data in our detection experiments were taken from the constructed dataset, it is essential to conduct the robustness validation of the model by using other data. We select other data from the Zhoushan Port area of China. The areas contain dense oil tanks, and the detection results shown in Figure 12 indicate the prediction of 116 windows are positive samples, in which 98 windows contain complete oil tanks,

while the other 18 windows are false alarms. The model demonstrates an overall accuracy of 84.48% and a missed detection rate of 16.97%, which indicates that the method is also robust for data acquired over different areas.



Figure 12. The detection results in the robustness validation experiments.

5. Discussion

The results of the recognition experiments on the MSAR1.0 dataset show that the proposed method can effectively recognize oil tanks. The model trained using the threshold-filtering method demonstrates improved recognition of oil tank targets, effectively solving the problem of poor-quality samples. In comparison to other advanced classical algorithms, our proposed method demonstrates superior recognition results. In contrast to other methods for reducing feature dimensionality, the regional L2 norm effectively achieves a balance between model accuracy and the time required for feature extraction in the monogenic signal.

The detection experiment results show that the method of Weibull distribution modeling and hole filling can effectively reduce the differences exhibited by the scattering characteristics of diverse oil tank roofs in SAR images. However, for oil tanks with an unclear distribution of scattering points, such as those depicted in Figure 13a, this method cannot change the wrong results. As depicted in Figure 13c, the scattering points in certain densely packed oil tanks exhibit a high degree of continuity, posing a challenge for the model to differentiate between these tanks and consequently leading to subpar recognition performance.

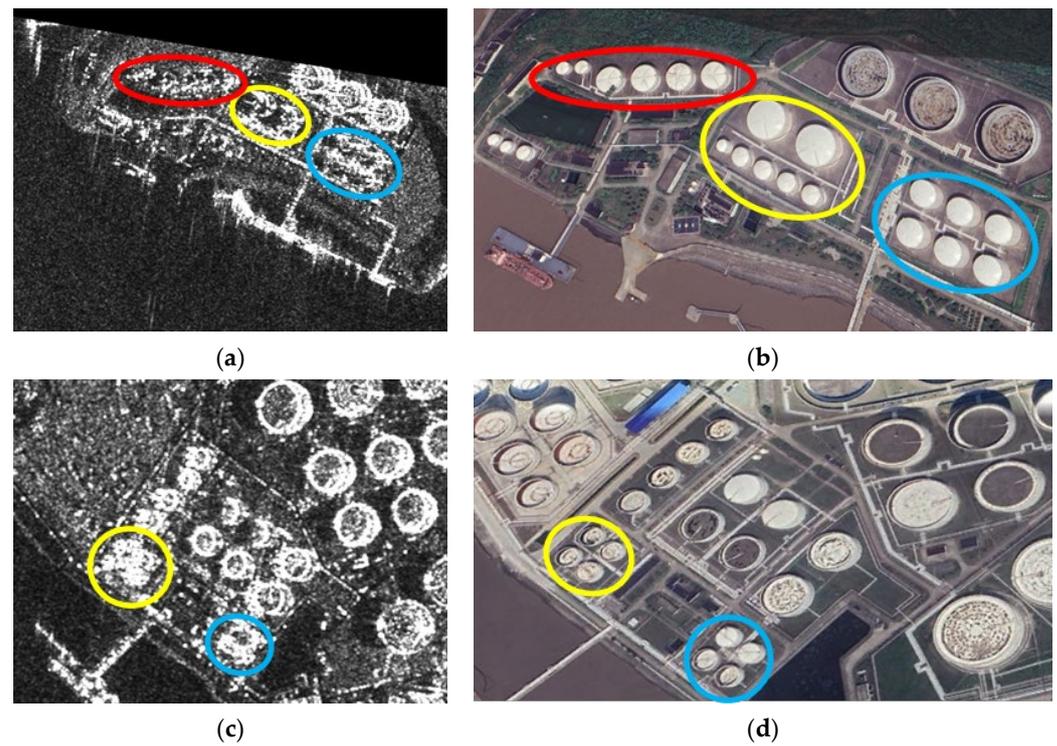


Figure 13. (a) The oil tanks with unclear scattering in the SAR image. (b) The same oil tanks with unclear scattering samples in the optical image. (c) The oil tanks with overlapped scattering in the SAR image. (d) The oil tanks in the optical image corresponding to the samples in (c).

In general, the model has better detection results with an accuracy of 91.67%. The proposed method achieves full recognition of large-size oil tanks and high precision recognition of small-size and connected oil tanks, indicating a generalization in the model.

In the robustness validation experiments, the model was evaluated using data obtained from different collection conditions. Variations in the incident angle and other environmental conditions can influence the scattering characteristics of oil tanks, but the achieved accuracy of 84.48% demonstrates the robustness of the results.

The training samples are constructed in the dense oil tank area shown in Figure 6, and the positive samples in this area are floating-roof oil tanks. The oil tanks appearing in the area shown in Figure 14a are fixed-roof oil tanks. The model indicates its capacity for generalization by correctly identifying the oil tanks shown in Figure 10b.

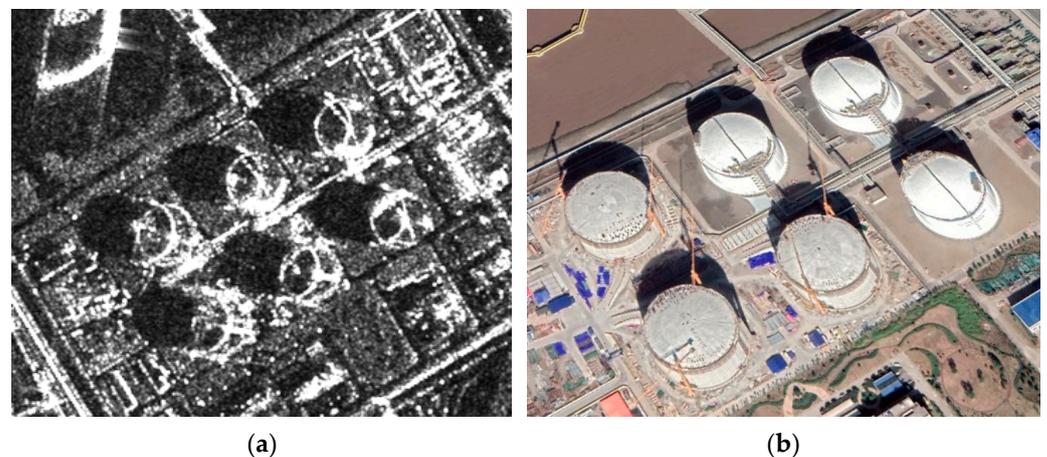


Figure 14. (a) The fixed-roof oil tanks in the SAR image. (b) The same oil tanks in the optical image.

While promising experimental results have been achieved in oil tank detection, there is space for improvement in performance on small-size and connected oil tanks. We use the two classes of small-size and connected oil tanks as auxiliary positive samples to participate in the calculation of sparse matrices, and cannot recognize oil tanks at multiple scales. In the dataset we constructed, negative samples contain many strong scattering targets, which greatly reduces the false detection rate and correspondingly increases the missed detection rate during detection. The balance between positive and negative samples is a problem worth thinking about. Although samples with poor scattering characteristics exhibit weak oil tank features in SAR images, we need to consider incorporating these positive samples to provide information about the capacity and location of the oil tanks in practical applications.

6. Conclusions

Oil tank detection and recognition have important application value. The variations in the conditions for obtaining high-resolution images in SAR satellites result in differences in the scattering characteristics represented by targets in different data. The process of SAR side imaging and multi-part scattering has led to overlapping and geometric distortion of the structure of oil tanks. In regions with many oil tanks, particularly those of small-size and connected samples, there is a great challenge in segmentation and recognition. These challenges are of great importance in practical applications. In this study, for oil tank detection, we made the following achievements.

- (1) To enhance the model's ability to recognize oil tanks, we propose the regional L2 norm dimensionality reduction method based on a monogenic signal. The method gains higher accuracy in oil tank target recognition experiments using the MSAR1.0 dataset compared to other advanced algorithms.
- (2) We conducted a comparative experiment to evaluate various dimensionality reduction methods. The feature extraction method based on the regional L2 norm has higher accuracy compared to the down-sampling method, and shorter computation time compared to the random projection method. The results indicate that the regional L2 norm dimensionality reduction method based on monogenic signal features can effectively strike a balance between accuracy and computation time.
- (3) To reduce the variations in scattering characteristics in SAR images of different oil tank roofs, we use Weibull distribution modeling and hole filling to process images, which improves the detection rate of oil tanks. In the detection experiment, the model successfully recognized fixed-roof oil tanks that were not included in the training set, indicating a high generalization ability in the model.
- (4) To evaluate the robustness of the model, we validate it on data with different parameters, and the result shows little variation. The model achieves an accuracy of 91.67% on the image for constructed dataset, and 84.48% on another image at different collection conditions.

The results indicate that the proposed method exhibits high accuracy, and the model has strong generalization and robustness. The application of this method in oil tank detection and recognition may achieve great results in the research of targets such as airplanes and ships.

The capacity of the model to learn from multiple scattering types of oil tanks is crucial for enhancing the results. The constructed dataset should encompass a wider variety of dispersed oil tank types and ensure a balanced distribution of quantities. The classification of negative samples is necessary to enhance the model's recognition ability. In our future research, we will explore the above direction and the multi-scale detection and segmentation of connected targets.

Author Contributions: Conceptualization, Y.F. and J.Y. (Junjun Yin); methodology, Y.F.; validation, Y.F., J.Y. (Junjun Yin), and J.Y. (Jian Yang); formal analysis, Y.F.; investigation, Y.F.; resources, J.Y. (Junjun Yin); data curation, J.Y. (Junjun Yin); writing—original draft preparation, Y.F.; writing—review

and editing, Y.F. and J.Y. (Junjun Yin); visualization, Y.F.; supervision, J.Y. (Junjun Yin); project administration, J.Y. (Jian Yang); funding acquisition, J.Y. (Junjun Yin). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSFC under Grant no. 62222102, NSFC no. 62171023, and the Fundamental Research Funds for the Central Universities under Grant no. FRF-TP-22-005C1.

Data Availability Statement: The storage URL for the MSAR1.0 dataset used for oil tank recognition experiments is <https://radars.ac.cn/web/data/getData?dataType=MSAR> (accessed on 18 May 2023).

Acknowledgments: We sincerely thank Centro Espacial INTA Torrejón for providing the high-resolution PAZ data, and also thank the CICG, School of Electronic and Information Engineering, Anhui University, for providing the large-scale multi-class SAR target detection dataset 1.0.

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

References

- Bamler, R. Principles of Synthetic Aperture Radar. *Surv. Geophys.* **2000**, *21*, 147–157. [[CrossRef](#)]
- Huertas, A.; Nevatia, R. Detecting buildings in aerial images. *Comput. Vis. Graph. Image Process.* **1988**, *41*, 131–152. [[CrossRef](#)]
- Kim, M.; Madden, M.; Warner, T.A. Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery. In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*; Springer: Berlin/Heidelberg, Germany, 2008. [[CrossRef](#)]
- Weber, J.; Lefèvre, S. Spatial and spectral morphological template matching. *Image Vis. Comput.* **2012**, *30*, 934–945. [[CrossRef](#)]
- Stankov, K.; He, D.-C. Detection of Buildings in Multispectral Very High Spatial Resolution Images Using the Percentage Occupancy Hit-or-Miss Transform. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4069–4080. [[CrossRef](#)]
- Tian, S.; Bhattacharya, U.; Lu, S.; Su, B.; Wang, Q.; Wei, X.; Lu, Y.; Tan, C.L. Multilingual scene character recognition with co-occurrence of histogram of oriented gradients. *Pattern Recognit.* **2016**, *51*, 125–134. [[CrossRef](#)]
- Lindeberg, T. Scale Invariant Feature Transform. *Scholarpedia* **2012**, *7*, 10491. [[CrossRef](#)]
- Khotanzad, A.; Hong, Y.H. Invariant Image Recognition by Zernike Moments. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 489–497. [[CrossRef](#)]
- Steinwart, I.; Christmann, A. Support vector machines. *Wiley Interdiscip. Rev. Comput. Stat.* **2008**, *1*, 49. [[CrossRef](#)]
- Kramer, O. K-Nearest Neighbors. In *Dimensionality Reduction with Unsupervised Nearest Neighbors*; Springer: Berlin/Heidelberg, Germany, 2013. [[CrossRef](#)]
- Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.; Ma, Y. Robust Face Recognition via Sparse Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 210–227. [[CrossRef](#)]
- Koohi-Fayegh, S.; Rosen, M.A. A review of energy storage types, applications and recent developments. *J. Energy Storage* **2020**, *27*, 101047. [[CrossRef](#)]
- Felsberg, M.; Sommer, G. The monogenic signal. *IEEE Trans. Signal Process* **2001**, *49*, 3136–3144. [[CrossRef](#)]
- Felsberg, M.; Sommer, G. The Monogenic Scale-Space: A Unifying Approach to Phase-Based Image Processing in Scale-Space. *J. Math. Imaging Vis.* **2004**, *21*, 5–26. [[CrossRef](#)]
- Huang, X.; Zhao, G.; Zheng, W.; Pietikäinen, M. Spatiotemporal Local Monogenic Binary Patterns for Facial Expression Recognition. *IEEE Signal Process. Lett.* **2012**, *19*, 243–246. [[CrossRef](#)]
- Dong, G.; Wang, N.; Kuang, G. Sparse Representation of Monogenic Signal: With Application to Target Recognition in SAR Images. *IEEE Signal Process. Lett.* **2014**, *21*, 952–956. [[CrossRef](#)]
- Dong, G.; Kuang, G. SAR Target Recognition Via Sparse Representation of Monogenic Signal on Grassmann Manifolds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1308–1319. [[CrossRef](#)]
- Watkins, A.J. On expectations associated with maximum likelihood estimation in the Weibull distribution. *Stat. Methods Appl.* **1998**, *7*, 15–26. [[CrossRef](#)]
- Hahn, S.L. *Hilbert Transforms in Signal Processing*; Artech House Inc.: Norwood, MA, USA, 1996.
- Bülow, T.; Sommer, G. Hypercomplex signals—A novel extension of the analytic signal to the multidimensional case. *IEEE Trans. Signal Process.* **2001**, *49*, 2844–2852. [[CrossRef](#)]
- Auscher, P.; Coulhon, T.; Duong, X.T.; Hofmann, S. Riesz transform on manifolds and heat kernel regularity. *Ann. Sci. Ec. Norm. Super.* **2004**, *37*, 911–957. [[CrossRef](#)]
- Li, F.; Yi, M.; Zhang, C.; Yao, W.; Hu, X.; Liu, F. POLSAR Target Recognition Using a Feature Fusion Framework Based on Monogenic Signal and Complex-Valued Nonlocal Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 7859–7872. [[CrossRef](#)]
- Field, D.J. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A Opt. Image Sci.* **1987**, *4 Pt 12*, 2379–2394. [[CrossRef](#)]
- Dong, G.; Kuang, G.; Wang, N.; Zhao, L.; Lu, J. SAR Target Recognition via Joint Sparse Representation of Monogenic Signal. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3316–3328. [[CrossRef](#)]

25. Fan, Y.; Yin, J.; Yang, J. SAR Target Recognition via Features Extracted from Monogenic Signal. In Proceedings of the IGARSS 2023—2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023; pp. 7471–7474. [[CrossRef](#)]
26. Kamran; Khan, A.; Malik, S.A. A high capacity reversible watermarking approach for authenticating images: Exploiting down-sampling, histogram processing, and block selection. *Inf. Sci.* **2014**, *256*, 162–183. [[CrossRef](#)]
27. Bingham, E.; Mannila, H. Random projection in dimensionality reduction: Applications to image and text data. In Proceedings of the Knowledge Discovery and Data Mining, San Francisco, CA, USA, 26–29 August 2001. [[CrossRef](#)]
28. Chaturvedi, S.K. Study of synthetic aperture radar and automatic identification system for ship target detection. *J. Ocean Eng. Sci.* **2019**, *4*, 173–182. [[CrossRef](#)]

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.