Journal of *Marine Science and Engineering*

Article

Sea-Surface Small Target Detection Based on Improved Markov Transition Fields

Ru Ye, Hongyan Xing * and Xing Zhou

School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 202212180010@nuist.edu.cn (R.Y.); 202212180033@nuist.edu.cn (X.Z.)
* Correspondence: xinghy@nuist.edu.cn

Abstract: Addressing the limitations of manually extracting features from small maritime target signals, this paper explores Markov transition fields and convolutional neural networks, proposing a detection method for small targets based on an improved Markov transition field. Initially, the raw data undergo a Fourier transform, feature fusion is performed on the series, and a spectrogram is generated using Markov transition fields to extract radar data features from both the time domain and frequency domain, providing a more comprehensive data representation for the detector. Then, the InceptionResnetV2 network is employed as a classifier, setting decision thresholds based on the softmax layer's output, thus achieving controllable false alarms in the detection of small maritime targets. Additionally, transfer learning is introduced to address the issue of sample imbalance. The IPIX dataset is used for experimental verification. The experimental results show that the proposed detection method can deeply mine the differences between targets and the maritime clutter background, demonstrating superior detection performance. When the observation time is set to 1.024 s, the IMIRV2 detector performs best. Cross-validation with different data preprocessing methods and classification models reveals a significant advantage in the performance of the IMIRV2 detector, especially at low signal-to-noise ratios. Finally, a comparison with the performance of existing detectors indicates that the proposed method offers certain improvements.

Keywords: sea clutter; target detection; MTF

Citation: Ye, R.; Xing, H.; Zhou, X. Sea-Surface Small Target Detection Based on Improved Markov Transition Fields. *J. Mar. Sci. Eng.* **2024**, *12*, 582. https://doi.org/10.3390/jmse12040582

Academic Editor: Zhiwei Xu

Received: 29 January 2024
Revised: 23 February 2024
Accepted: 27 March 2024
Published: 29 March 2024

1. Introduction

Sea clutter, also known as ocean surface clutter or wave clutter, is a type of interference signal received by radar systems when detecting targets at sea [1]. This signal is primarily caused by waves, sea surface fluctuations, and other oceanic phenomena. The characteristics of sea clutter typically include randomness, time variability, and non-uniformity [2]. Through research on the detection of small targets at sea, efficient monitoring and rapid response to maritime targets can be achieved, enabling the timely detection of anomalies and hazardous events, thereby ensuring the safety of maritime traffic and the protection of the marine environment. Due to the weak radar returns of small targets, they are difficult to detect. Traditional detection methods based on statistical characteristics or the energy of radar returns struggle to maintain high detection accuracy and a low false alarm rate [3], presenting many challenges to current detection technology.

In the exploration of research advancements in the field of small target detection at sea, scholars have adopted a variety of methods to overcome the challenges posed by sea clutter. These challenges include characteristics such as randomness, time variability, and non-uniformity, especially under low signal-to-noise ratio (SNR) conditions, where detection accuracy often leaves much to be desired.

In recent years, feature-based approaches have achieved significant breakthroughs in the field of detecting small targets at sea, such as using time–frequency domain features, short-time fractional Fourier transform, and the high-dimensional feature generalization capability of deep learning, all aimed at enhancing the detection ability of small targets against sea clutter. Relevant studies are shown in Table 1.

Table 1. Research on Sea Surface Small Target Detection Based on Feature Extraction Methods.

Author(s)	Method	Characteristics	Advantages	Disadvantages	Reference
Shui et al.	Three-feature detector in the time–frequency domain	Selects features including relative amplitude, relative Doppler peak height, and relative vector entropy	Significantly improves detection performance	The detection accuracy is not satisfactory under low signal-to-noise ratio conditions	[4]
Chen et al.	Detection and extraction of micro-motion targets based on short-time fractional Fourier transform	Verifies capability of short-time Fourier transform in representing time series features	Enhances the detection capability of micro-motion targets in sea clutter	Challenges in feature extraction	[5]
Shi et al.	Utilization of deep learning for radar dynamic target detection	High-dimensional feature generalization ability	Provides a new technical approach for target detection and identification against clutter background	Timeliness of target detection can be challenging	[6]
Zhao et al.	Improved four-feature extraction method using the FAST algorithm	Uses feature optimization to enhance feature distinctiveness	Achieves better detection results	Difficulties in feature extraction	[7]

As deep learning research continues to advance, convolutional neural networks [8,9] have been used to classify sea clutter and noise. Through convolutional neural networks and feature perception technologies, researchers have proposed various image-based encoding methods to enhance the accuracy and generalization capability of target detection, as shown in Table 2.

Table 2. Research on Sea Surface Small Target Detection Based on Image Encoding Methods.

Author(s)	Method	Characteristics	Advantages	Disadvantages	Reference
Xu et al.	Sea surface small target detection method based on feature perception of multi-channel graphs in the frequency domain	Exploring radar signal features in the frequency domain	Can improve the detection accuracy of small sea surface targets	Difficulties in feature extraction	[10]
Shi et al.	Sea surface small target detection method based on time–frequency images	Deeply mining the differences between targets and clutter	Effectively enhances the detection capability of small maritime targets at low signal-to-noise ratios	May require complex data processing and feature extraction steps	[11]

However, in these methods, feature extraction is difficult, and the information provided by image encoding methods for the classification model is limited. To extract features more effectively, this paper investigates Markov transition fields for processing time series. Markov transition fields (MTFs) [12] use Markov transition probabilities to preserve the information relationship between time domains. Zhao [13] chose MTFs for preprocessing fault sequences and inputting them into a convolutional neural network for classification, achieving good results. Wang Z [14] compared MTFs with Gram angular fields as two image encoding methods for time series, finding that MTFs had better representation ability for time series.

Based on this, this paper proposes a sea surface small target detection method based on Markov transition fields. Firstly, the raw data undergo a Fourier transform, and feature fusion is performed on the series. The MTF is used to encode the signal into a two-dimensional image, and a neural network is used to extract features from the image. Then, transfer learning is utilized to improve the performance of the deep learning network model, reduce training costs, and adjust the softmax threshold of the model to achieve a constant false alarm rate, establishing a classification model for sea clutter signals and weak target signals. Finally, using the IPIX radar signal as the experimental subject and comparing the performance of traditional detectors, this paper verifies that the proposed method can deeply mine the differences between targets and clutter, thereby having better detection performance.

2. The Theoretical Basis of Data Processing

The first consideration in target detection based on graph features is the conversion of data into a spectrogram. The effective representation of both time and frequency domain information in the spectrogram can significantly enhance detector performance. This paper optimizes the capture of information in the spectrogram through the effective integration of Markov transition fields and Fourier transform. The paper provides a detailed description of the optimization and transformation.

2.1. Target Detection Problem Transformation

In practice, sea surface small target detection is commonly divided into single-class and dual-class problems [15,16]. A single classifier is trained based on pure sea clutter data, focusing on learning the clutter features, but often neglects the characteristics of the target signal, leading to limited applicability. In contrast, the use of dual classifiers becomes key in enhancing detection efficiency and accuracy, especially in scenarios where target echo samples are scarce and manual feature extraction is overly rigid. Exploring detection methods that autonomously learn features in an imbalanced sample environment is of significant value.

Based on this, a binary classification hypothesis is applied to the echoes $z(n)$ received from N consecutive pulses by the radar [17], which is:

$$\begin{cases} H_0 : z(n) = c(n) \\ H_1 : z(n) = c(n) + s(n) \end{cases}, \tag{1}$$

where $c(n)$ is the pure sea clutter, and $s(n)$ is the target returns. When the original observed data contain only pure sea clutter, that is, there is no observed target, they are judged as hypothesis H_0. When the original observed data include target echoes, indicating the possible presence of an observed target, they are judged as hypothesis H_1. Therefore, the problem of target detection is transformed into a binary classification.

2.2. The Theoretical Basis of the Improved Markov Transition Field

Markov transition fields (MTFs) [12] are a commonly used method to represent time series data, offering excellent descriptive capabilities for scale, magnitude, and shape changes in the data. They can reveal the inherent structure and dynamic changes in sea clutter data by describing state transition matrices, while also reducing the dimensionality of the original data and eliminating redundant information, thus enhancing computational efficiency and accuracy. Given the complexity of sea clutter data, traditional methods often fall short in capturing their subtle feature information. By integrating MTFs with Fourier transform, we have overcome this limitation. This approach not only comprehensively captures radar data characteristics from both the time and frequency domains but also optimizes the representation of data, significantly enhancing the identification capability for weak target signals within sea clutter. Through delicately mapping the dynamic relationship between state transitions and frequency changes, our method reveals target characteristics hidden in the complex maritime environment, providing a more efficient and precise solution for detecting small maritime targets.

Assume that the sea clutter time series is $X = \{x_1, x_2, \ldots, x_N\}$, where x_n is the nth ($n = 1, 2, \ldots, N$, N being the number of sampling points) sampled signal. First, the series is normalized:

$$X_{\text{norm}} = \frac{X - \min(X)}{\max(X) - \min(X)}, \tag{2}$$

and then the Fourier transform is calculated for the processed series:

$$F(x) = \sum\nolimits_{n=0}^{N-1} x_n \cdot e^{-\frac{i2\pi}{N}\text{kn}}, \tag{3}$$

$F(x)$ is the result of the Fourier transform.

Then, feature fusion is performed on the series, combining the original time series X with the magnitude of the Fourier transform $|F(X)|$:

$$X_{\text{combined}} = \text{concat}(X, |F(x)|), \tag{4}$$

based on the values of the fused time series, define Q regions q_j ($j = 1, 2, \ldots, Q$), such that each x_n can be mapped to one of the q_j. Then, calculate the transition probabilities between each q_j to obtain a Markov transition matrix with dimensions $Q \times Q$:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1Q} \\ w_{21} & w_{22} & \cdots & w_{2Q} \\ \vdots & & \ddots & \vdots \\ w_{Q1} & w_{Q2} & \cdots & w_{QQ} \end{bmatrix}$$
$$\begin{bmatrix} P(x_i \in q_1 | x_{i-1} \in q_1) & \cdots & P(x_i \in q_1 | x_{i-1} \in q_Q) \\ P(x_i \in q_2 | x_{i-1} \in q_1) & \cdots & P(x_i \in q_2 | x_{i-1} \in q_Q) \\ \vdots & \ddots & \vdots \\ P(x_i \in q_Q | x_{i-1} \in q_1) & \cdots & P(x_i \in q_Q | x_{i-1} \in q_Q) \end{bmatrix}, \tag{5}$$

where w_{jk}(k = 1, 2, ..., Q) is the probability of a sampled signal in region q_k being followed by a sampled signal in region q_j, with $w_{\text{jk}} = P(x_i \in q_j | x_{i-1} \in q_k)$. Finally, the Markov transition matrix is expanded by arranging each probability in temporal order, thus generating an $N \times N$ MTF matrix M:

$$M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{N1} & M_{N2} & \cdots & M_{NN} \end{bmatrix}$$
$$\begin{bmatrix} P(x_1 \in q_j \to x_1 \in q_k) & \cdots & P(x_1 \in q_j \to x_N \in q_k) \\ P(x_2 \in q_j \to x_1 \in q_k) & \cdots & P(x_2 \in q_j \to x_N \in q_k) \\ \vdots & \ddots & \vdots \\ P(x_N \in q_j \to x_1 \in q_k) & \cdots & P(x_n \in q_j \to x_N \in q_k) \end{bmatrix}, \tag{6}$$

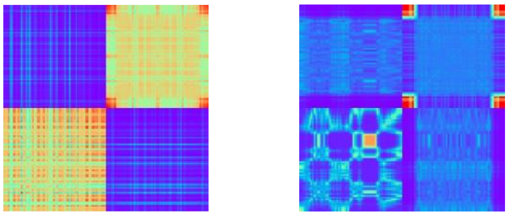
where M_{hm}(h, m = 1, 2, ..., n) is the transition probability from the region q_j, mapped from x_h, to the region q_k, mapped from x_m, with $M_{hm} = P(x_h \in q_j \to x_m \in q_k)$. The elements in the MTF matrix M range from [0, 1], and through Equation (6), each element's value in the matrix can be scaled from 0 to 255, aligning it with the pixel values in an image, thereby obtaining a two-dimensional image:

$$I(h,\ m) = \text{int}(255\, M_{hm}), \tag{7}$$

where $I(h,\ m)$ is the pixel value of the image in the hth row and mth column; $\mathrm{int}(\cdot)$ is the rounding function.

2.3. Data Trasformation

After the aforementioned transformation, radar signals can be converted into diagonally symmetric feature maps, as shown in Figure 1. This figure illustrates two-dimensional images formed from IPIX-measured radar data after FT-MTF transformations. As can be seen from Figure 1, after the transformation, due to the reflective properties of the target object, the radar echo images containing the target feature specific structural characteristics. In contrast, pure clutter images may lack such prominent structural features, exhibiting a more uniform chaotic image characteristic.



(**a**) (**b**)

Figure 1. Data transformation based on improved MTF. (**a**) Transformation of the pure clutter; (**b**) Transformation of the pure clutter with target returns.

The two-dimensional images obtained by the aforementioned method effectively reflect the differences in the time–frequency distribution between pure sea clutter and target returns, providing assistance for subsequent small target detection on the sea surface.

3. Target Detection Model Based on Improved MTF-InceptionResnetV2

In this research, we have redefined the issue of detecting small targets on the sea surface, treating it as a binary classification challenge. To address the complex task of describing the characteristics of sea clutter signals, we introduce a novel detection framework known as the detector based on improved Markov transition field (IMIRV2). As illustrated in Figure 2, the architecture of the IMIRV2 detector uniquely integrates an online detection pathway with an offline training module. This architecture is meticulously designed, incorporating data preprocessing, a pretrained model, a classification network, and a strategic false alarm area decision making component.

Initially, to address the issue of imbalanced sample quantities between the two classes, the training process involves enriching the training dataset by adding simulated targets into a pure clutter environment to create synthetic target echoes. Subsequently, slicing and a sea-clutter-specific FT-MTF processing method are applied to transform the data into information-rich two-dimensional images. These images, representing clutter and target categories, are then used to train the classification network, ultimately developing a pretrained model.

In the detection phase, our method utilizes IPIX radar data, converting one-dimensional time series into feature-rich FT-MTF two-dimensional images. Before feature extraction and classification through the deep learning network, a pretrained model is integrated via transfer learning. This strategic integration aims to overcome the challenges of data scarcity and initial model performance limitations. The detection process is further refined by calculating the false alarm rate *Pfa* and comparing the network's output prediction value P with a dynamically adjusted decision threshold θ, ensuring precise optimization of the threshold. The final output of this process is highly accurate detection results. Each element of the IMIRV2 framework is meticulously designed, from initial data preparation to the final decision mechanism.

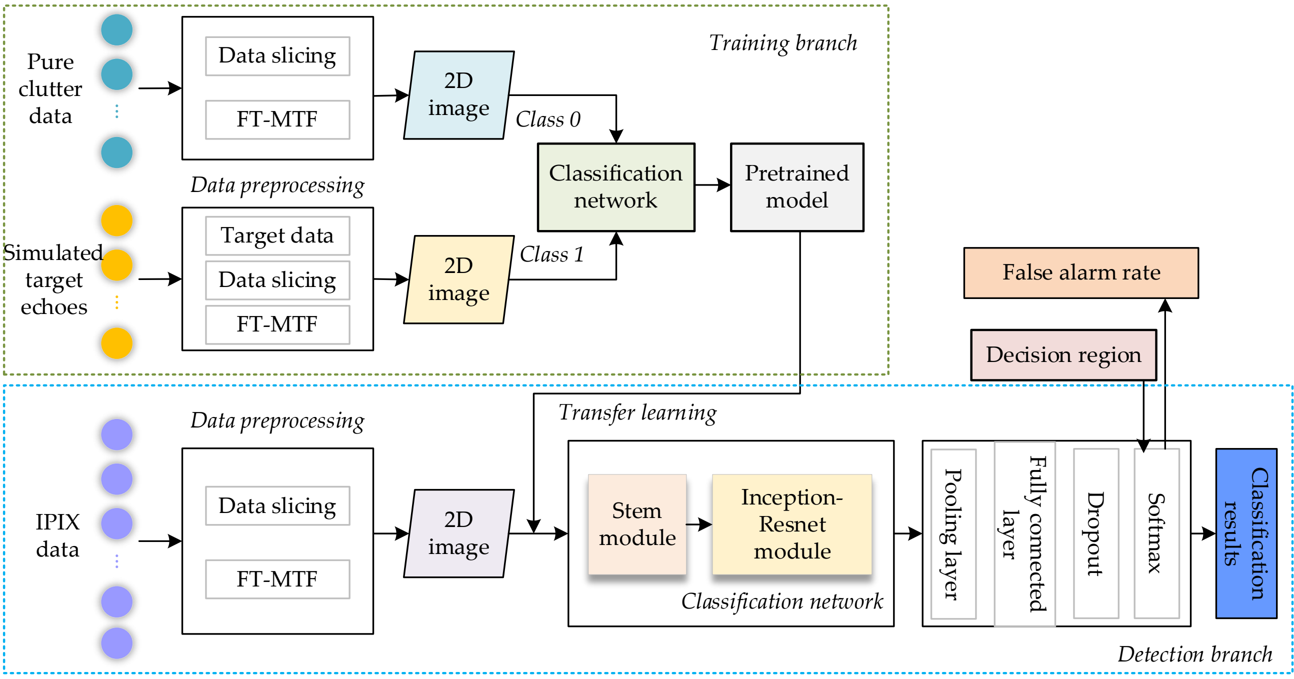


Figure 2. Target Detection Model Diagram Based on Improved MTF-InceptionResNetV2.

Below, we will delve into the intricacies of each component, showcasing the innovative techniques and methodological integration that set our work apart from traditional detection methods.

3.1. Data Preprocessing: Improved Markov Transition Field

In the field of sea surface small target detection, due to the high complexity of sea clutter, traditional methods face significant challenges in capturing key feature information. This study proposes a novel FT-MTF method, an improvement of the MTF method combined with Fourier transform technology, aimed at overcoming this difficulty. Our method thoroughly analyzes the combined effects of state transition probabilities and frequency features and efficiently expresses these features in the form of two-dimensional images, with the improved images generated shown in Figure 3.

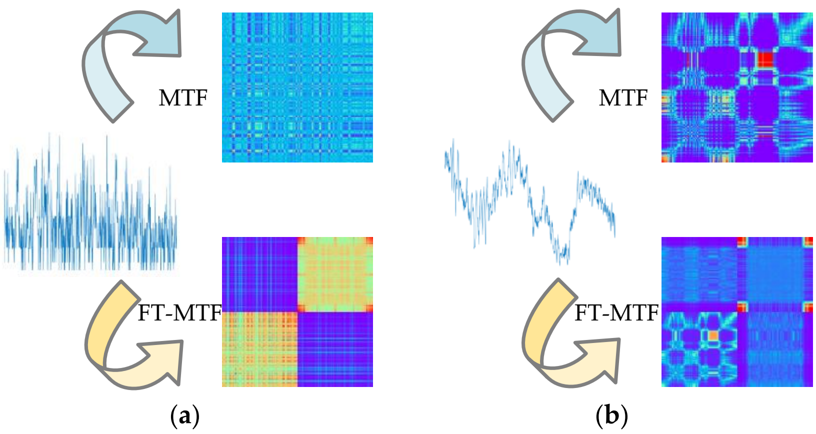


Figure 3. Data transformation based on MTF and improved MTF. (**a**) Transformation of the pure clutter; (**b**) Transformation of the pure clutter with target returns.

Specifically, the FT-MTF method significantly enhances the richness and effectiveness of feature representation through the following steps:

1. Expansion of the State Space: As can be seen from Figure 3, by introducing the Fourier transform, we have expanded the state space of the MTF image to include not only the state transition information of the original time series data but also the transition probabilities between these states and the Fourier transform features;
2. Quadrant Feature Analysis: The improved MTF image is divided into four quadrants, each representing different dynamics of state transitions. This decomposition method provides a new perspective for understanding and analyzing the behavior of small targets in sea clutter, significantly enhancing the expressiveness of features;
3. Integrated Feature Representation: The FT-MTF method effectively combines the frequency features of time series with state transition probabilities to generate two-dimensional images. This transformation not only preserves the essential information of the original data but also enriches the feature set for the IMIRV2 detector by incorporating the dynamic characteristics of state transitions, providing a more comprehensive and dynamic feature set;
4. Enhanced Model Recognition Capability: With the images generated by the FT-MTF method, the IMIRV2 model can more easily identify subtle changes in small targets within sea clutter. These changes are represented in the images through various patterns and structures, enabling the model to more effectively learn and distinguish between sea surface targets and background clutter.

Through these innovations, this study not only provides a new perspective and method for sea surface small target detection but also significantly improves the accuracy and efficiency of detection. These technological advancements offer powerful tools for solving the challenges of detecting small targets in sea clutter and are expected to drive the development of related technologies and applications.

3.2. Classification Network: InceptionResNetV2 Classifier

To capture multi-scale features from fine to coarse granularity in the improved MTF image data, this study employs Inception-ResNetV2 (IRV2) [18] as the classification network. It extracts image features through convolutional kernels of various sizes and residual connections. This network structure allows for the extraction and fusion of features at multiple levels, enhancing the model's ability to recognize objects of different sizes in images.

InceptionResNetV2, proposed by the Google Brain team in 2016, is a deep convolutional neural network architecture that, compared to previous Inception [19] and ResNet [20] models, possesses a deeper network structure. Deep networks are typically capable of learning more abstract and complex features, enhancing the model's representational capability. Combining features of both Inception and ResNet, InceptionResNetV2 can capture features at multiple scales (Inception module) and maintain good gradient propagation (ResNet module), thereby improving the model's feature extraction capabilities. The specific structure is illustrated in Figure 4.

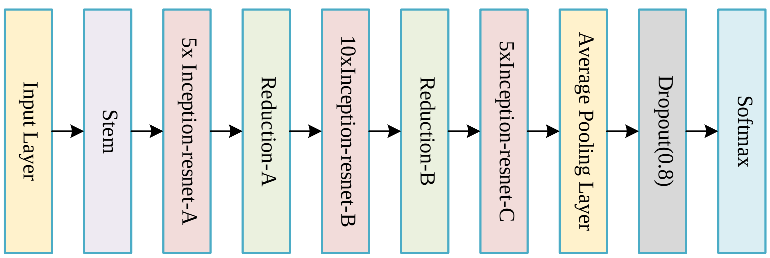


Figure 4. InceptionResNetV2 Model Diagram.

Considering that most of the data are correlated, introducing transfer learning can accelerate and optimize the model's classification accuracy. Based on the FT-MTF images of sea clutter and target echoes, transfer learning [21] training can extract features from the images. By fine-tuning the training to adjust the weights of the InceptionResNetV2 model, the learning rate is reduced. A validation set from a portion of the training data is used for further weight adjustment. Transfer learning is trained on the basis of existing deep network models, reducing training costs and suitable for small datasets. When the data volume is sufficient and training costs are not a concern, retraining a new model based on the data can be chosen.

3.3. Controllable False Alarm in the Decision Region

In the context of small target detection on the sea surface, the detection probability of a classifier while maintaining a consistently low false alarm rate is the criterion for evaluating the effectiveness of the classifier's target detection. The detection probability is the proportion of correctly identified samples among all actual target samples. The false alarm rate is the proportion of incorrectly identified samples among all actual non-target samples. To prevent an imbalance in network judgment, the target detection requirement generally does not exceed a false alarm rate of 10^{-3}.

To address this issue, this paper proposes a detection strategy based on softmax probability output and employs the Monte Carlo method to determine the decision boundary for a controllable false alarm rate. In a binary classification scenario, the output of the softmax reflects the probabilities of belonging to two categories, as shown in Equation (8).

$$\begin{array}{c} \text{Class 0 : Softmax}(\mathrm{Z})_0 = \frac{e^{Z_0}}{e^{Z_0}+e^{Z_1}} \\ \text{Class 1 : Softmax}(\mathrm{Z})_1 = \frac{e^{Z_1}}{e^{Z_0}+e^{Z_1}} \end{array} \quad (8)$$

Typically, the classification network makes a judgment based on a preset threshold θ (0.5), as shown by the black line in Figure 5. To handle classification bias in imbalanced datasets, a different threshold θ (where $0 < \theta < 1$) is set. The classification decision rule can be expressed as:

$$\begin{cases} \textit{if } P(\text{class } 1) \geq \theta, \text{ the sample is classified as Class 1} \\ \textit{if } P(\text{class } 1) < \theta, \text{ the sample is classified as Class 0}' \end{cases} \quad (9)$$

where P(class 1) is the probability predicted by the model that a sample belongs to Class 1. As shown in Figure 5, by adjusting the decision threshold to the red line, it is possible to control the number of data classified as "Class 1", thereby achieving controllable false alarms in the detection of small targets at sea.

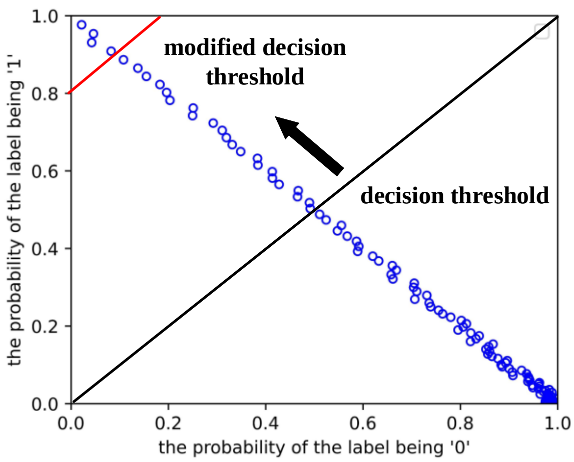


Figure 5. False Alarm Controllable Decision Area.

4. Experimental System and Performance Analysis

This study uses the IPIX [22] database. The IPIX radar data, provided by McMaster University in Canada, offer valuable resources for the study of sea clutter characteristics and the detection of weak signals. The IPIX radar operates at a sampling frequency of

1000 Hz, with a sampling time of 131.072 s, and each range gate is 15 m long. It operates at grazing angles less than or equal to 1 degree. The experiment utilizes six sets of data, with specific parameters shown in Table 3.

Table 3. Detailed Description of IPIX Radar Data.

Data Name	Target Unit	Sub-Target Unit	Wind Speed/km·h^{-1}	Wave Height/m	Angle/°
#17	9	8~11	9	2.2	9
#26	7	6~8	9	1.1	97
#30	7	6~8	19	0.9	98
#54	8	7~10	20	0.7	8
#310	7	6~9	33	0.9	30
#311	7	6~9	33	0.9	40

Table 3 records the parameters of six types of data. To quantitatively measure data characteristics, the average signal-to-clutter ratio (ASCR) is introduced as an indicator to distinguish performance indicators under different sea conditions. ASCR is defined as the ratio of the average power of the target echoes to the power of the clutter echoes. Based on radar echo data [23,24], the estimation formula is as follows:

$$\text{ASCR} = 10\lg\left[\frac{\sum_{i=1}^{N}\left|z\left(\frac{1}{H_1}\right)\right|^2 - \sum_{i=1}^{N}\left|z\left(\frac{1}{H_0}\right)\right|^2}{\sum_{i=1}^{N}\left|z\left(\frac{1}{H_0}\right)\right|^2}\right]. \tag{10}$$

Figure 6 shows a comparison chart of the average signal-to-clutter ratio (SCR) for the six datasets under four types of polarization. It can be observed from the figure that the average SCR under two cross-polarization methods is generally higher than that under two co-polarization methods. Additionally, the VV polarization method produces more Bragg scattering compared to the HH polarization method, resulting in a lower average SCR for the former. It is clearly evident from Figure 2 that datasets #17, #54, and #311 have strong target abilities, while datasets #26, #30, and #310 have weaker target echoes, which are easily drowned out by clutter signals, increasing the difficulty of detection.

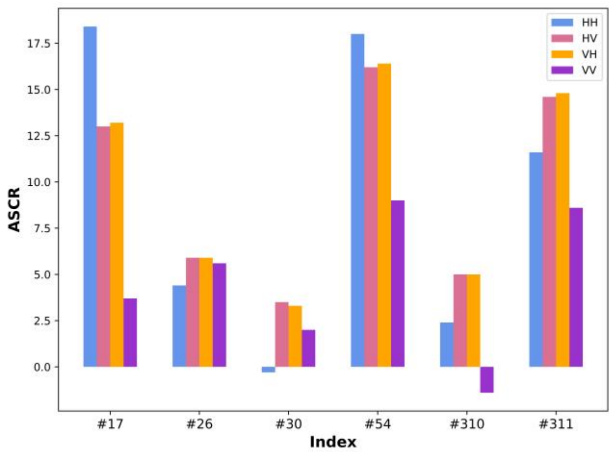


Figure 6. ASCR under The Four Types of Polarization.

4.1. Data Dimensionality Experiment

In practical tests, when converting radar data into two-dimensional images, selecting different data dimensions affects the features contained in the images and thus the detection performance. To explore the effect of the IMIRV2 detector on sea surface small target detection with different data dimensions, dataset #54 is selected, and comparative experiments are conducted with dimensions of 256, 512, 1024, and 2048, respectively. As shown in Figure 7, these are the detection probabilities obtained by training the IMIRV2 model with four different data dimensions.

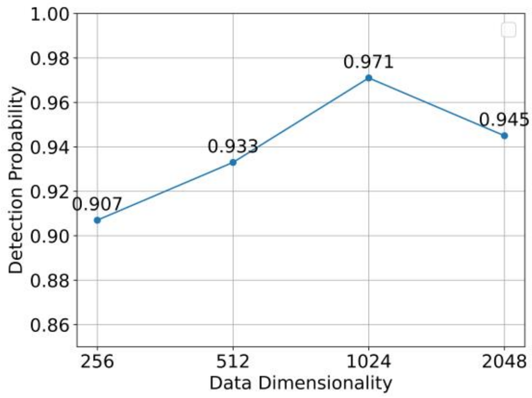


Figure 7. Detection Probability under Different Data Dimensionalities.

From Figure 7, it can be seen that at 256 dimensions, the detection probability is about 0.907; at 512 dimensions, it increases to about 0.933; at 1024 dimensions, it reaches a peak of about 0.971; however, at 2048 dimensions, the detection probability slightly decreases to about 0.945. This indicates that after reaching a certain number of dimensions, adding more features does not improve detection performance. Therefore, 1024 is chosen as the feature dimension for subsequent experiments.

4.2. Experiments on the Impact of Improved MTF on Detector Performance under Constant False Alarm Rate

To further investigate the specific impact of the improved MTF on the performance of sea surface small target detection, this study conducted experiments on measured radar data under different sea conditions while maintaining a constant false alarm rate. The classification confusion matrix used in the experiments reflects the relationship between the actual categories of the samples and the categories predicted by the classifier, including four key indicators: true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). In this study, the positive class is defined as the category containing target echoes, while the negative class is defined as sea clutter.

Given that a key objective in sea surface small target detection is to improve detection precision while maintaining a false alarm rate of 10^{-3}, this experiment adopted the method described in Section 3.3 to set FN to zero. Therefore, the performance evaluation of the detector primarily relies on precision ($P = \frac{TP}{TP+FP}$).

4.2.1. The Impact Analysis of Different Preprocessing Methods on Detector Performance

To explore the impact of different data preprocessing methods on enhancing detector performance, the experiment introduced two commonly used methods for encoding time series data into images, recurrence plots (RPs) [25] and the Gram angular field (GAF) [26], for comparative experiments. They were combined with the Fourier transform (FT) to investigate the superiority of the proposed data processing methods.

An RP is a visualization tool that reveals repeating patterns in time series data. The process of combining RPs with FT is shown in Figure 8a. The experiment found that, due to the significant difference in magnitude between FT features and the original data, simply adding FT features to the original data was not sufficient to change the structure of the similarity matrix. Therefore, FT features were standardized to the same range as the original data, ensuring that the FT features were similar in magnitude to the original data, allowing for better integration of the features. The resulting images are shown in Figure 8b.

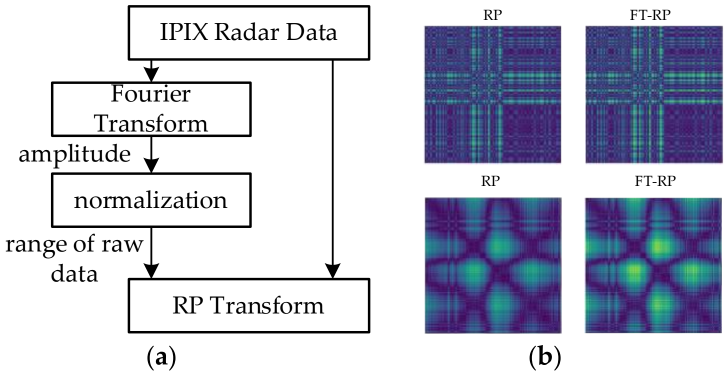


Figure 8. Improved RP: (**a**) Specific steps of the improved RP method, (**b**) Comparative images before and after the improvement of RP.

The GAF is a method for converting time series data into images by calculating the relative angles between points in the time series, thus generating images that express the dynamics of the time series. Therefore, when combining Fourier transform with GAFs, it is necessary to convert the amplitude of FFT into relative strength and calculate its angle. The converted FFT features are then merged with the original time series data to generate image data. The FT-GAF conversion process and the resulting images are shown in Figure 9.

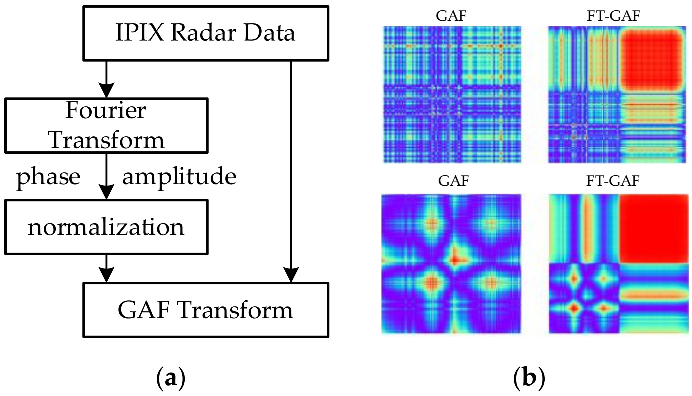


Figure 9. Improved GAF: (**a**) Specific steps of the improved GAF method, (**b**) Comparative images before and after the improvement of GAF.

Under a constant false alarm rate, comparative experiments were conducted using the IRV2 model in combination with the previously mentioned preprocessing methods and the preprocessing method proposed in this paper. To more clearly analyze the experimental results, two sets of high sea conditions (#54 and #311) and two sets of low sea conditions (#30 and #310) were selected for the experiments. The results are recorded in Figure 10.

It can be clearly seen from the figure that the methods combined with FT perform better across all detectors, especially the FT-MTF+IRV2 method, which achieved higher scores on data from different sea conditions. In low sea conditions, there was a significant improvement in performance, showing a notable performance advantage.

Before integrating FT features, the best detection performance was achieved using the MTF preprocessing method, with the detection performance using RP preprocessing slightly higher than that using GAF. However, after integrating FT features, detectors using MTF and GAF preprocessing methods saw significant improvements in detection performance in low sea conditions, while the improvement in detection performance using RP preprocessing was limited, with detection performance lower than that of the other two preprocessing methods in all sea conditions. Although the detection performance of GAF combined with FT features significantly improved, it still did not match the MTF preprocessing method combined with FT features. The integration of MTF and FT can study state changes in dynamic systems, while the combination of GAF and FT is more suitable for the analysis of periodic features. For the detection of small targets on the sea surface, the MTF preprocessing method combined with FT features is superior to other preprocessing methods.

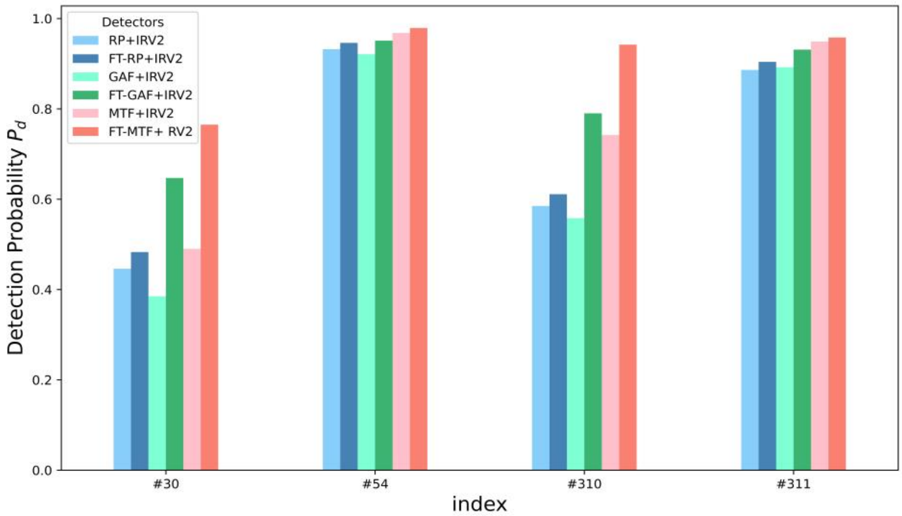


Figure 10. Detector performance under different preprocessing methods.

4.2.2. Analysis of the Impact of Improved MTF on Detector Performance under Different Classification Networks

To further explore the impact of the improved MTF on detection performance, we conducted experiments under a constant false alarm rate, employing both MTF and FT-MTF preprocessing methods with different models such as ResNet, InceptionV3, and IRV2. The results are recorded in Table 4.

Table 4. Detector Performance Under Different Detectors.

Detectors	#30	#54	#310	#311
MTF + ResNet	0.426	0.929	0.558	0.892
FT-MTF + ResNet	0.712	0.951	0.799	0.932
MTF + InceptionV3	0.433	0.932	0.585	0.886
FT-MTF + InceptionV3	0.735	0.956	0.811	0.941
MTF + IRV2	0.490	0.968	0.742	0.949
FT-MTF + IRV2	0.765	0.979	0.942	0.958

Table 4 displays the detection performance under different detectors. By comparing the performance of each model on different datasets (#30, #54, #310, #311), it was found that the improvement effect of FT-MTF preprocessing is significant: regardless of the combination with ResNet, InceptionV3, or IRV2, detectors using FT-MTF preprocessing significantly outperform those using only MTF across all datasets. Thus, FT-MTF preprocessing effectively enhances the model's detection capability. Regarding the superiority of the IRV2 model, compared to ResNet and InceptionV3, the IRV2 model performs better across all test datasets, whether combined with MTF or FT-MTF. IRV2 can more effectively process the complex data involved in sea surface small target detection tasks. Regarding performance differences under different sea states, the performance on datasets #30 and #310 is generally lower compared to #54 and #311, reflecting the impact of different sea states on detection performance. More challenging sea states may make target detection more difficult, but the FT-MTF preprocessing method combined with efficient deep learning models can significantly mitigate this issue.

In summary, the analysis of Figure 10 and Table 4 indicates that the method proposed in this paper can significantly improve the performance of sea surface small target detection, especially under complex sea conditions.

4.3. Detection Performance Comparison

This section of the experiment uses the IPIX dataset shown in Table 3 to compare the detection performance of the proposed IMIRV2 detector against three other detectors: the tri-feature detector [4], the support vector machine (SVM) detector [27], and the Bi-LSTM detector [28].

At a preset false alarm rate of 10^{-3}, the detection performance of the four detectors on the IPIX dataset is illustrated in Figure 11. From Figure 11, it can be observed that the proposed IMIRV2 detector demonstrates good detection performance across all subsets of the IPIX dataset, while the other three detectors show a significant drop in detection probability on certain subsets of the IPIX dataset. The experimental results indicate that the detection performance of the proposed IMIRV2 detector surpasses that of the other three detectors.

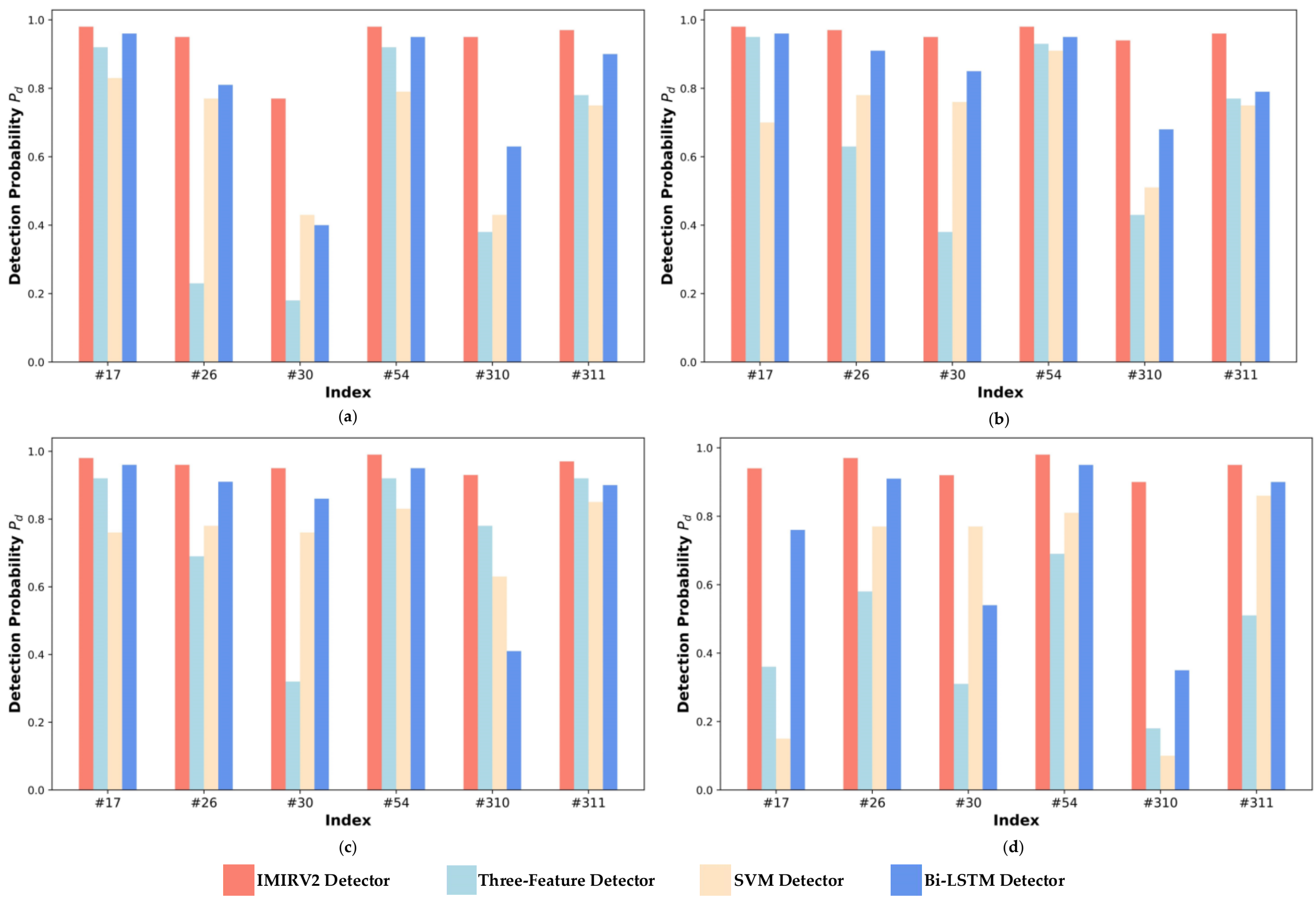


Figure 11. Comparison of detection performance of four types of detectors: (**a**) Detection performance under HH polarization condition, (**b**) Detection performance under HV polarization condition, (**c**) Detection performance under VH polarization condition, (**d**) Detection performance under VV polarization condition.

4.3.1. Analysis of the Impact of Different Classifiers on Performance

Deep learning detectors can map data into a high-dimensional space to learn and extract the distinctive features between targets and clutter. Compared to clutter, there is a smaller difference in target features between the training and test sets, allowing deep learning detectors to better learn target features. Therefore, compared to the tri-feature detector and the support vector machine (SVM) detector that utilize handcrafted features, the other two deep-learning-based detectors overall achieved a higher average detection probability. However, this is not absolute; for some subsets, the Bi-LSTM detector's detection probability is lower than that of the tri-feature detector and SVM detector. Due to the strong time variability of clutter, the clutter distribution in the training and test sets may not be entirely consistent, leading to a slightly higher average actual false alarm rate of the detectors than the preset false alarm rate. The proposed IMIRV2 detector can efficiently integrate features extracted from various data sources. Compared to other detectors, the IMIRV2 detector has a broader range of detection feature sources and extracts target features and clutter features with stronger distinction. Therefore, the proposed IMIRV2 detector achieves a higher average detection probability while having a lower average actual false alarm rate.

4.3.2. Analysis of Detection Performance under Four Polarization Conditions (HH, HV, VH, VV)

Performance tests were conducted on six selected sets of data under different polarization conditions. There are changes in the average signal-to-clutter ratio (SCR) under the four polarization types, resulting in differences in detection probabilities. HV and VH polarization methods have more advantages than HH and VV methods. Figure 11 shows a comparison chart of the detection performance of ten sets of data under four polarization types. In this figure, the false alarm rate is set to 10^{-3} and the cumulative number of pulses N is 1024. It can be seen that the detection performance of HH, HV, and VH polarization is significantly higher than that of VV polarization. The detection method proposed in this experiment performs particularly well on HH polarization data and even slightly exceeds HV and VH polarization in some aspects. Under each polarization method, compared with the three-feature detector and SVM detector, the IMIRV2 detector shows significant detection performance improvement. Compared with the Bi-LSTM detector, the performance of the IMIRV2 detector also has a certain improvement.

4.3.3. Detection Performance Analysis under Different Environments

Comparing the ASCR across various datasets, datasets #17, #54, and #311 are considered as high sea states, while datasets #26, #30, and #310 are viewed as low sea states. It can be observed from Figure 11 that, under high sea states, all detectors demonstrate good detection performance, but the detection performance of IMIRV2 is still superior to other detectors. Under low sea states, especially when the SCR is low and the Doppler distance is short (#30), since the target echo is completely masked within the main clutter frequency, traditional detectors struggle to detect the target, rendering the triple-feature detector almost inoperative. Meanwhile, the SVM detector and the Bi-LSTM detector also perform poorly, whereas the proposed detector shows about a 50% performance improvement over the triple-feature detector and about a 30% improvement over the other two detectors. This indicates that the proposed IMIRV2 detector surpasses three classic constant false alarm rate (CFAR) detectors in detection performance under complex sea surface environments.

5. Conclusions

This paper addresses the challenge of characterizing data features against a background of sea clutter by first applying the MTF method combined with Fourier transform to the raw data to generate spectrograms. This approach captures the characteristics of radar data from both the time and frequency domains, thereby providing a more comprehensive data representation. The InceptionResNetV2 network is employed as the classifier, and a predefined decision threshold is set based on the softmax layer's output to achieve controlled false alarms in the detection of small targets on the sea surface. To address the issue of sample imbalance, transfer learning is introduced by training on simulated data to generate a pretrained model. The IMIRV2 detection method is proposed and validated using the IPIX dataset. Experimental results demonstrate that the proposed detection method exhibits excellent detection performance. When N = 1024, the IMIRV2 detector performs optimally. Cross-validation with different data preprocessing methods and classification models reveals that the method proposed in this paper significantly enhances the detection performance of small targets on the sea surface. Compared to the detection performance of existing detectors, improvements are observed under all sea conditions, especially in low sea states, where the proposed detection method performs approximately 30% better than several existing detection methods. In summary, the IMIRV2 detection method provides strong detection performance, a simple model, and fast computation, making it suitable for detecting small targets on the sea surface.

However, there are still some limitations to the proposed detection method in this article. For example, the spectrograms generated using FT-MTF contain rich spatiotemporal features. Future research may focus on improving the classification model by incorporating attention mechanisms to make the model pay more attention to important parts of the input data, thus enhancing sensitivity to crucial information and improving classification accuracy.

Author Contributions: Conceptualization, R.Y. and H.X.; methodology, H.X.; software, R.Y.; validation, X.Z.; formal analysis, R.Y.; resources, R.Y.; data curation, R.Y.; writing—original draft preparation, R.Y.; writing—review and editing, H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by National Natural Science Foundation of China (Grant No. 62171228).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data were downloaded from the following website: http://soma.ece.mcmaster.ca/ipix/index.html (accessed on 27 May 2021). The data were measured with the McMaster IPIX Radar, a fully coherent X-band radar, with advanced features such as dual transmit/receive polarization, frequency agility, and stare/surveillance mode.

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

References

1. Haykin, S.; Bakker, R.; Currie, B.W. Uncovering Nonlinear dynamics-the case study of sea clutter. *Proc. IEEE* **2002**, *90*, 860–881. [CrossRef]
2. Zhou, W.; Xie, J.; Li, G.; Du, Y. Robust CFAR Detector with Weighted Amplitude Iteration in Nonhomogeneous sea clutter. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 1520–1535. [CrossRef]
3. Xing, H.-Y.; Zhu, Q.-Q. The Sea Clutter denosing Based on Ensemble Empirical Mode Decomposition. *Acta Electron. Sin.* **2016**, *44*, 1–7.
4. Shui, P.-L.; Li, D.-C.; Xu, S.-W. Tri-feature-based detection of floating small targets in sea clutter. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 1416–1430. [CrossRef]
5. Chen, X.; Guan, J.; Bao, Z.; He, Y. Detection and extraction of target with micromotion in spiky sea clutter via short-time fractional Fourier transform. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 1002–1018. [CrossRef]
6. Shi, S.-N.; Shui, P.-L. Sea-surface floating small target detection by one-class classifier in time-frequency feature space. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6395–6411. [CrossRef]
7. Zhao, D.; Xing, H.; Wang, H.; Zhang, H.; Liang, X.; Li, H. Sea-Surface Small Target Detection Based on Four Features Extracted by FAST Algorithm. *J. Mar. Sci. Eng.* **2023**, *11*, 339. [CrossRef]
8. Wagner, S.A. SAR ATR by a combination of convolutional neural network and support vector machines. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2861–2872. [CrossRef]
9. Yan, Y.; Xing, H. Small floating target detection method based on chaotic long short-term memory network. *J. Mar. Sci. Eng.* **2021**, *9*, 651. [CrossRef]
10. Xu, S.-W.; Jiao, Y.-P.; Bai, X.-H.; Jiang, H.-Z. Detection of Small Sea Surface Targets Based on Frequency Domain Multi-Channel Image Feature Perception. *J. Electron. Inf. Technol.* **2023**, *45*, 1567.
11. Shi, S.-N.; Dong, Z.-Y.; Yang, J.; Yang, C.-J. Detection of Small Sea Surface Targets Based on Self-Learning Time-Frequency Images. *J. Syst. Eng. Electron.* **2021**, *43*, 33.
12. Koupai, A.K.; Bocus, M.J.; Santos-Rodriguez, R.; Piechocki, R.J.; McConville, R. Self-supervised multimodal fusion transformer for passive activity recognition. *IET Wirel. Sens. Syst.* **2022**, *12*, 149–160. [CrossRef]
13. Zhao, Z.-H.; Li, C.-X.; Dou, G.-J.; Yang, S.-P. Research on bearing fault diagnosis based on MTF-CNN. *J. Vib. Shock.* **2023**, *42*, 1076.

14. Wang, Z.; Oates, T. Imaging time-series to improve classification and imputation. *arXiv* **2015**, arXiv:1506.00327.
15. Zhao, W.; Jin, M.; Cui, G.; Wang, Y. Eigenvalues-based detector design for radar small floating target detection in sea clutter. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [CrossRef]
16. Xu, S.; Zhu, J.; Jiang, J.; Shui, P. Sea-surface floating small target detection by multifeature detector based on isolation forest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *14*, 704–715. [CrossRef]
17. Li, D.; Shui, P. Floating small target detection in sea clutter via normalised Hurst exponent. *Electron. Lett.* **2014**, *50*, 1240–1242. [CrossRef]
18. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
19. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
20. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
21. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [CrossRef]
22. IPIX Radar. The McMaster IPIX Radar Sea Clutter Database. 2021. Available online: http://soma.ece.mcmaster.ca/ipix/index.html (accessed on 27 May 2021).
23. Shui, P.; Guo, Z.; Shi, S. Feature-compression-based detection of sea-surface small targets. *IEEE Access* **2019**, *8*, 8371–8385. [CrossRef]
24. Xu, S.; Zheng, J.; Pu, J.; Shui, P. Sea-surface floating small target detection based on polarization features. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1505–1509. [CrossRef]
25. Dias, D.; Dias, U.; Menini, N.; Lamparelli, R.; Torres, R. Pixelwise Remote Sensing Image Classification Based on Recurrence Plot Deep Features. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019.
26. Yang, C.-L.; Chen, Z.-X.; Yang, C.-Y. Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images. *Sensors* **2019**, *20*, 168. [CrossRef] [PubMed]
27. Li, Y.; Xie, P.; Tang, Z.; Jiang, T.; Qi, P. SVM-based sea-surface small target detection: A false-alarm-rate-controllable approach. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1225–1229. [CrossRef]
28. Wan, H.; Tian, X.; Liang, J.; Shen, X. Sequence-feature detection of small targets in sea clutter based on Bi-LSTM. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–11. [CrossRef]