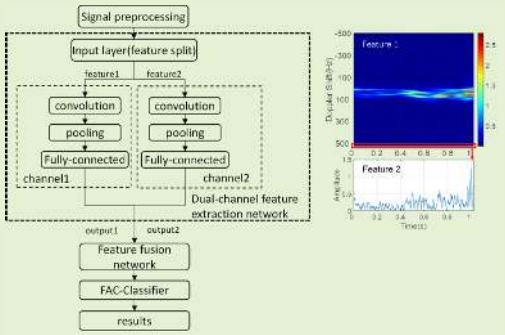
False-Alarm-Controllable Radar Detection for Marine Target Based on Multi Features Fusion via CNNs

Xiaolong Chen, *Member, IEEE*, Ningyuan Su, Yong Huang, and Jian Guan

***Abstract*—Due to the influence of the complex marine environment, the marine target detection based on statistical theory is difficult to achieve high-performance. Moreover, due to various targets' motion characteristics, only using a single feature for detection is unreliable. In this paper, from the perspective of feature extraction and classification, marine target and sea clutter are classified by deep learning methods. To achieve the required false alarm rate, the dual-channel convolutional neural networks (DCCNN) and false-alarm-controllable classifier (FACC)-based marine target detection method is proposed. Firstly, the measured sea clutter and the target signal are preprocessed to obtain the time-Doppler spectrum and amplitude information. The Marine-DCCNN (MDCCNN) is then constructed for features extraction and fusion, and the feature vectors of the signals are obtained. The performance of different feature extraction models is tested and compared. Finally, the FACC is used as a detector to classify the feature vectors into two categories and the control of the false alarm rate is realized. The detection performances were verified by two popular public radar datasets, i.e., IPIX radar dataset (floating target) and CSIR dataset (maneuvering marine target). The results show that compared with single-channel CNN and histogram of oriented gradient support vector machine (Hog-SVM) classification, a combination of MDCCNN feature extraction model and softmax classifier can achieve higher performance and controllable false alarm rate. Moreover, HH polarization and mixed training datasets under different sea states can help improve detection performance.**



***Index Terms*—Radar target detection, marine target, feature extraction, dual-channel convolutional neural network (DCCNN), false alarm controllable classifier (FACC).**

I. Introduction

SEA surface target detection is important in both military and civilian fields [1]. As the primary means, radar is widely used for maritime search and rescue, ocean surveillance, and national defense security [2]–[4], etc. However, sometimes the radar echo of marine target is weak due to the strong sea surface radar reflections, i.e., sea clutter [5], [6]. The marine target may exhibit low observability characteristics, which makes the radar detection performance difficult to meet the actual needs. Robust and adaptive detection of targets in sea clutter is a worldwide problem and a key point for radar signal processing [7], [8].

The traditional detection method is based on statistical theory [9], and the detection process is based on the amplitude distribution model hypothesis [10]. However, in practical applications, it is difficult to accurately describe the sea clutter distribution with the statistical models. Achieving high detection performance is difficult in complex marine environments [11]. Another kind of marine target detection method is from the perspective of characteristic differences [12], [13]. In addition to the radar amplitude, the speed variation is different for target and sea clutter, and the micro-Doppler can reflect the instantaneous velocity change of the target or the scattering points [14]–[16]. The time-Doppler spectrum using time-frequency analysis is useful, which stands for the frequency distribution over time [17], [18]. It is essentially the coherent integration, which can improve the signal-to-clutter ratio (SCR) of target's returns [19]. These methods include linear transform, such as short-time Fourier transform (STFT), fractional FT (FRFT), etc.; and nonlinear transform, such as Wigner-Ville distribution (WVD), fractional ambiguity

Manuscript received December 9, 2020; revised January 20, 2021; accepted January 23, 2021. Date of publication January 26, 2021; date of current version March 5, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant U1933135, Grant 61931021, and Grant 61871392 and in part by the Key Research and Development Program of Shandong under Grant 2019GSF111004. The associate editor coordinating the review of this article and approving it for publication was Dr. Michail Antoniou. *(Corresponding author: Xiaolong Chen.)*

The authors are with the Marine Target Detection Research Group, Naval Aviation University, Yantai 264001, China (e-mail: cxlcxl1209@163.com).

Digital Object Identifier 10.1109/JSEN.2021.3054744

function (FRAF), etc. However, they need to match with the target motion characteristics to achieve higher integration gain, and the detection performance cannot adapt to different motion status as well.

Recently deep learning methods are developed, which are quite efficient and effective for data expression [20], [21], especially for high-dimensional feature extraction [13]. As the most popular solution of deep learning, convolutional neural networks (CNN) models begin to be applied to radar signal processing, target detection, and recognition, such as SAR image processing [22], gesture recognition [23], clutter suppression [24], navigation radar image detection [25], etc. Reference [21] combined the preprocessed SAR image with the original image and classified the images with CNN, in order to achieve high recognition performance. Reference [26] used CNN to classify range-Doppler images of radar signals to achieve gesture recognition. In [24], the deep neural network (DNN) is applied for binary classification, i.e., target detection. In [25], CNN is employed to classify the segmented maritime radar image samples to achieve the classification of maritime targets, clutter, and coastline. Since the target detection is a binary classification of the signal, the binary classification of CNN can be applied to distinguish the target from the clutter.

Different from image detection and classification, several essential issues need to be considered when using CNN for radar signal detection and classification in clutter backgrounds. On the one hand, the radar detection environment is involved with different target characteristics. Only using a single feature is unreliable. For example, due to the influence of strong backscatters of the sea surface, e.g., sea spikes, sea clutter sometimes exhibits similar characteristics with the moving target on the time-Doppler spectrum [27], [28]. In [29], CNN was used to classify several types of micro-motion targets, achieving a classification accuracy of 99.46%. However, the target echo is discontinuous due to the influence of sea waves. Moreover, due to the low speed, the target's spectrum will be overlapped and covered by sea clutter on the time-Doppler domain, resulting in classification errors. Therefore, only using single-channel CNN (the time-frequency information) for maritime target detection is not satisfactory. In [30], multi-channel CNNs are used to extract and fuse lidar and hyperspectral images for ground objects classification. For radar signal detection under complex sea clutter background, how to select the multi-channel features and design proper CNN models still remains investigated.

The other important factor which should be considered for real applications is the controllable false alarm for stable detection performance [31]. The traditional constant false alarm rate (CFAR) detector is an adaptive threshold method determined by the statistical distribution characteristics of the background [32]. Because the prior probability information of the target is not available, the risk function of false alarm and missing alarm are different and cannot be quantified, and the minimum Bayes risk decision method cannot be adopted for CNN [33]. It is proved that the machine learning model using the minimum mean square error loss function could be used as the Neyman-Pearson (NP) criterion [34]. In [11], a cost-sensitive support vector machine (2C-SVM) is proposed, which is an SVM with different penalty factors for different classification categories. When processing the simulated signal of Swerling I and II target and additive white Gaussian noise (AWGN) background, this detector has good performances when detecting. However, the detection probability is lower when processing the real measured signal. Reference [35] extracted the radar signal sequence's statistical information to construct the feature vector, and used the SVM to classify the signal feature vectors to realize controllable false alarm. Existed CNN based detection methods mostly calculate false alarm based on the result itself, e.g., SAR image detection [36], while most works do not consider how to achieve better detection performance with controllable false alarms. Therefore, further research on the CNN-based detection method is needed to meet the actual radar false alarm requirement [37]–[39].

In this paper, a marine target detection method based on marine dual-channel CNN (MDCCNN) with a false-alarm-controllable classifier (FACC) is proposed for complex backgrounds. Two typical features, i.e., time-frequency information and amplitude information of radar returns are both subjected to feature extraction and fusion processing to improve radar target detection performance. In order to achieve the actual radar target detection requirements, two classifiers are constructed to realize controllable false alarm rates based on the NP criterion, i.e., variable threshold Softmax classifier and false alarm controllable SVM. Two types of real radar data under different conditions, i.e., polarizations (HH, HV, VH, VV), sea states (low and high), target characteristics (floating and maneuvering) are used for verification. Moreover, the proposed method is compared with the traditional histogram of oriented gradient (HOG)-SVM [40], single-channel CNN method, and CFAR detectors. The rest of the paper is organized as follows. The principle of the proposed MDCCNN and two types of FACC are introduced in section II. In section III, the detailed detection procedure for marine target is given and the datasets used in the paper are introduced. The effectiveness of the MDCCNN is demonstrated and validated by two real radar data in section IV. The last section concludes the paper and presents its future research direction.

II. Principle of MDCCNN and False Alarm Controllable Classifier

A. Principle of Target Detection Based on CNN

When using CNN for image processing, the CNN can be regarded as a function whose input is the original image and the output is the classification result [20].

$$\mathbf{y} = f_{\text{softmax}}\left(f_{\text{fc}}\left(f_{\text{conv}}\left(\mathbf{x}\right)\right)\right) \tag{1}$$

where $\mathbf{y}$ is the result of classification, $\mathbf{x}$ is the input image, f_{fc} is a fully-connected layer operation and f_{conv} is a convolution layer operation.

Compared with the traditional fully-connected network, the use of the convolutional layer in CNN improves the feature extraction performance significantly. Usually, a CNN consists of several layers. For each layer, the input of the current layer is the output of the previous layer. Using CNN to classify

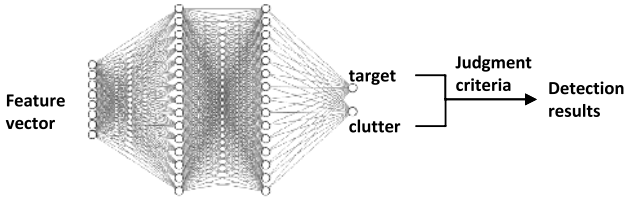


Fig. 1. Full connection layer structure for target detection.

the target and clutter signals include the procedure of feature extraction and classification, and the structure corresponds to the convolution-pooling and the fully-connected layer [41].

1) Convolutional Layer: The operation of the convolutional layer is as follows [20]

$$y_j^c = f\left(\sum_i^{M_j} x_i^c \otimes k + b_c\right) \tag{2}$$

where, y_j^c is the jth node in output of current convolutional layer, $f(\,)$ is a nonlinear activation function, such as sigmoid, hyperbolic tangent function, rectified linear unit, exponential linear unit, etc., x_i^c is the ith node in input of current convolutional layer, M_j is the value range of node x_i^c, i.e., the local receptive field, $\otimes$ indicating a convolution operation. And k and b_c are trainable parameters, which are convolution kernel and biases, respectively.

2) Pooling Layer: The pooling layer implements a down sampling operation, which is performed independently in each feature map. One node, which is in the output feature map, is the function of the nodes in the corresponding receptive field in the input. The function of the receptive field is

$$y_j = f_{\text{down}}\left(x_1, x_2, \cdots, x_N\right), \quad x_i \in \mathbf{P} \tag{3}$$

where N is the number of nodes in the receptive field $\mathbf{P}$, and f_{down} is the down sampling function, and the maximum pooling is usually selected (calculating the value of node with maximum value in the $\mathbf{P}$) or the average pooling (calculating the average value of each node in the $\mathbf{P}$).

3) Fully-Connected Layer: The convolutional layer and the pooled layer operation complete the image feature extraction, and the fully-connected layer calculates the category corresponding to the feature. Each node in the fully-connected layer input has a value link between each node in the output, as shown in Fig. 1.

The operation of the fully-connected layer is shown as

$$y_j^f = f\left(\sum_i^{o_f} x_i^f \cdot w + b_f\right) \tag{4}$$

where o_f is the number of the elements in the feature map output by last layer, w and b_f are the weights and biases of the fully-connected layer. x_i^f is the ith node in input of current fully-connected layer, and y_j^f is the jth node in output of current fully-connected layer.

Usually the first fully-connected layer converts multiple feature maps into vectors, and the last fully-connected layer acts as a softmax classification layer. When using CNN for image classification, the dimension of the network output, i.e., the number of nodes of the last fully-connected layer, is equal to the number of categories. The output vector is used to obtain the classification result according to a certain decision criterion.

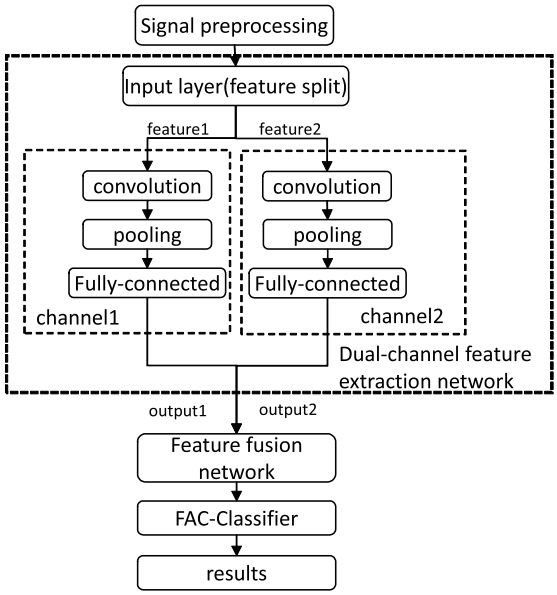


Fig. 2. The proposed DCCNN structure diagram.

B. Principle of the MDCCNN

Due to the similarity of frequency characteristics of some targets and clutter, the accuracy is usually low only using time-frequency graphs with CNN. The CNN classification via time series amplitude information also has a high error rate because the phase information of the signal is not utilized, i.e., Doppler information. Therefore, a novel dual-channel CNNs model is proposed, using both amplitude and Doppler characteristics of the signal. The MDCCNN structure proposed in this paper is shown in Fig. 2.

The MDCCNN includes two CNN channels, which respectively implement time-Doppler and amplitude features extraction. After extracting the two features, the fusion is performed with the fully-connected layer, and then the FACC is used for features classification. The decision threshold is determined according to the output data of the training set and the required false alarm rate. The network structures of two features extraction channels (taking LeNet and VGG19 as an example) are shown in Fig. 3.

The dual-channel feature extraction network includes an input layer, channel 1 and channel 2. The input layer separates the time-frequency graph and the amplitude vector in the sample matrix, and the two features are output to the extraction channels. One convolutional layer, a pooled layer, all or part of a fully-connected layer are employed for each channel CNN.

The m-th layer of convolution layer convolves the input feature map x_m by the convolution kernel k^m, and expands the convolution area with the offset coefficient b_m, which enables the edge feature more completely extracted and automatically extracts the characteristics of the signal in the feature map.

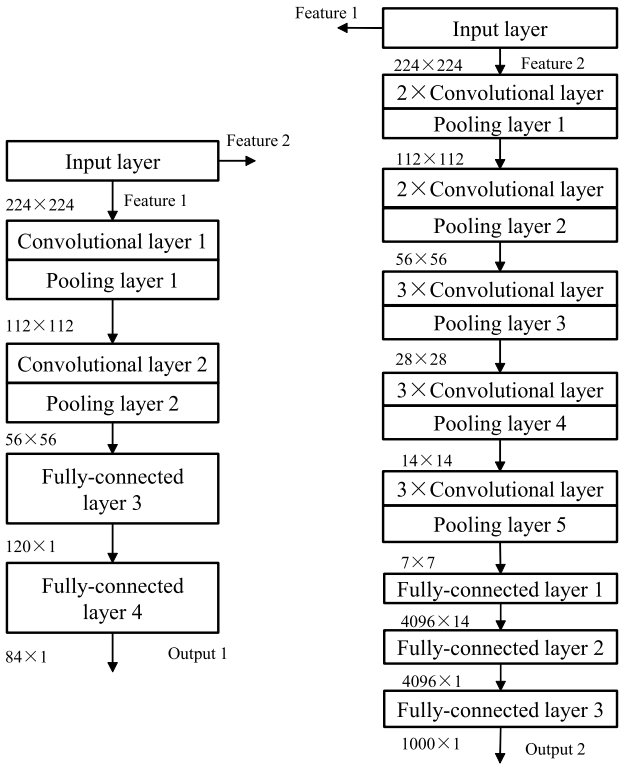


Fig. 3. Structure of two feature extraction channels: LeNet channel (left) and VGG16 channel (right).

Then nonlinear factor is introduced via the rectified linear unit (ReLU) activation function [42],

$$f(x) = \begin{cases} x, & x > 0 \\ 0, & x \le 0 \end{cases} \tag{5}$$

where x is input of the activation function. ReLU activation function converges quickly and robustly, and the problem of gradient vanishing can be solved.

Then the output feature maps,

$$y_j^m = f(\sum_{i \in L} x_i \otimes k_{ij}^m + b_j^m) \tag{6}$$

where, k_{ij}^m represents the j-th convolution kernel indicating the convolution operation of the i-th feature map in the m-th layer, b_j^m indicating the j-th bias coefficient corresponding to the j-th convolution kernel in the first layer, x_i is the input feature map, and y_j^m represents the j-th feature map of the first layer output.

The n-th layer of the pooling layer is down sampled by the pooling function y^n to reduce the data dimension and the computation burden. The j-th feature map y_j^n of the n-th layer output is obtained, and the pooling function adopts the average pooling.

$$y_j^n = \bar{A}_n(y_j^n)_{c \times c} \tag{7}$$

where the function $\bar{A}()_{c \times c}$ is the average pooling function, $\bar{A}(y_j^n)_{c \times c}$ means that the average value of the feature map matrix y_j^n in each $c \times c$ pooling kernel window, and c is the preset pooling kernel size. The value of c is in correspondence with that in the CNN type which the channel applies.

The number of output nodes of the p-th fully-connected layer is q, and the output feature vector of the $(p-1)$-th layer is multiplied by the weight matrix w_p, and an offset coefficient b_p is added. L_p is the output feature matrix,

$$L_p = f\left(w_p L_{p-1} + b_p\right) \tag{8}$$

The feature fusion network splices the two output feature vectors of the dual-channel feature extraction network by one or more layers of fully-connected layers. The number of layers and the number of nodes per layer are set according to the application scenario. Take the feature fusion network with three layers of fully-connected layers for example. The number of output nodes of the first fully-connected layer is 512, and the feature vector output from the feature extraction network is multiplied by the weight matrix $w_{3,1}$. The offset coefficient $b_{3,1}$ is added and $L_{3,1}$ is the output feature matrix,

$$L_{3,1} = f\left(w_{3,1}\left[L_{c1}, L_{c2}\right] + b_{3,1}\right) \tag{9}$$

The number of output nodes of the second fully connected layer is 128, and the output feature vector of the first fully-connected layer is multiplied by the weight matrix $w_{3,2}$, and the offset coefficient $b_{3,2}$ is added. $L_{3,2}$ is the output feature matrix,

$$L_{3,2} = f\left(w_{3,2} L_{3,1} + b_{3,2}\right) \tag{10}$$

The number of output nodes of the third fully-connected layer is 2, indicating a binary classification, and the feature vector outputted by the second layer is multiplied by a weight matrix $(2 \times n)$, and an offset coefficient $b_{3,3}$ is added. $L_{3,3}$ is the output feature matrix

$$L_{3,3} = f\left(w_{3,3} L_{3,2} + b_{3,3}\right) \tag{11}$$

C. False Alarm Controllable Classifier

Existed marine target detection is mainly based on CFAR detection. Under the premise of the clutter distribution characteristics, the CFAR detection can achieve a constant false alarm rate. However, CFAR detection has the disadvantage of being significantly affected by the environment. When the statistical characteristics of sea clutter are inconsistent with the thresholds, the method cannot achieve high-performance [43].

Radar target detection usually obeys NP criteria, which can been achieved by machine learning methods. The machine learning network trained by minimum mean square error function can achieve CFAR detection. The minimum mean square error function is:

$$J(F) = \int I\left(\mathbf{z}, F(\mathbf{z}), \frac{\partial F(\mathbf{z})}{\partial \mathbf{z}_1}, \cdots, \frac{\partial F(\mathbf{z})}{\partial \mathbf{z}_L}\right) \mathrm{d}\mathbf{z} \tag{12}$$

where, $\mathbf{z}$ is the input, $F(\mathbf{z})$ is the output, I represents the mean square error. L is the length of input vector $\mathbf{z}$. The $F(\mathbf{z})$ under the assumption the minimum $J(F)$ can be obtained by Lagrangian equation

$$\frac{\partial I}{\partial F} - \sum_{k=1}^{L} \frac{\partial}{\partial \mathbf{z}_k}\left(\frac{\partial I}{\partial F_k'}\right) = 0 \tag{13}$$

where $F_k' = \frac{\partial F}{\partial \mathbf{z}_k}$. Due to

$$I(\mathbf{z}, F(\mathbf{z})) = \sum_{i=0}^{1} P(H_i) \left(F(\mathbf{z}) - t_{H_i}\right)^2 \cdot f(\mathbf{z}, H_i) \quad (14)$$

where H_i represents categories, and t_{H_i} represents corresponding desired output. Then,

$$\frac{\partial}{\partial F}\left(\sum_{i=0}^{1} P(H_i) \left(F(\mathbf{z}) - t_{H_i}\right)^2 \cdot f(\mathbf{z}, H_i)\right) = 0 \quad (15)$$

The solution $F_0(\mathbf{z})$ can be obtained which satisfies the minimum mean square error condition

$$F_0(\mathbf{z}) = \frac{P(H_1) f(\mathbf{z}|H_1) t_{H_1} + P(H_0) f(\mathbf{z}|H_0) t_{H_0}}{P(H_1) f(\mathbf{z}|H_1) + P(H_0) f(\mathbf{z}|H_0)} \quad (16)$$

Since the softmax function of the CNN output layer can convert the classification result into a classification probability, it can be used as a NP detection.

$$S(x) = \frac{1}{1 + e^{-x}} \quad (17)$$

Then (16) can be rewritten as

$$F_0(\mathbf{z}) = \frac{P(H_1)\Lambda(\mathbf{z}) t_{H_1} + P(H_0) t_{H_0}}{P(H_1)\Lambda(\mathbf{z}) + P(H_0)} \quad (18)$$

where $\Lambda(\mathbf{z})$ is the likelihood ratio. Extracting $\Lambda(\mathbf{z})$ and comparing $F_0(\mathbf{z})$ to threshold η_0, (18) can be rewrite as

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)\left(\eta_0 - t_{H_0}\right)}{P(H_1)\left(t_{H_1} - \eta_0\right)} = \eta_{\text{lr}} \quad (19)$$

where η_{lr} is new threshold. The relationship between η_{lr} and η_0 is

$$\eta_0 = \frac{\eta_{\text{lr}} P(H_1) t_{H_1} + P(H_0) t_{H_0}}{P(H_0) + \eta_{\text{lr}} P(H_1)} \quad (20)$$

which does not depend on the input $\mathbf{z}$. Therefore, it can be used as a detector that satisfies the NP criteria.

In this paper, two kinds of false alarm controllable detectors are designed from the perspective of iterative training parameter optimization and statistical learning, i.e., variable threshold softmax classifier and false alarm controllable SVM.

1) Variable Threshold Softmax Classifier: The value of each node of the network output satisfies the interval [0, 1], and the sum is 1. The output vector of the output layer is a 1×2 array, and the value of the first bit indicates the probability that the sample is clutter, and the value of the second bit is the probability that the sample is target. By setting thresholds, different false alarm rate can be controlled,

$$\text{threshold} = O_1(i) \quad (21)$$

$$i = \text{Pfa}_{\text{desired}} \cdot N_{\text{clutter}} \quad (22)$$

where O_1 is the output of the clutter samples sorted by the value of the first bit from small to big, $\text{Pfa}_{\text{desired}}$ is the false alarm probability value to be controlled, and N_{clutter} means the number of clutter samples.

2) False Alarm Controllable Classifier SVM (FACC-SVM): 2C-SVM achieves false alarm rate control by setting different types of penalty factors [35]. Suppose F_i is the input vector of the softmax layer after the training samples with MDCCNN, and label $y_i \in \{-1, +1\}$ represents the clutter (-1) and target $(+1)$ samples, respectively. The training set of the SVM is $\{(F_i, y_i), i = 1, 2, \ldots, M\}$, where M is the number of training samples. The SVM maps the input vector to the multidimensional feature space via a kernel function, which is usually used by a radial basis function:

$$k(F_1, F_2) = \exp\left(-\frac{\|F_1 - F_2\|}{2\delta^2}\right) \quad (23)$$

where δ is an adjustable parameter which controls the radial range of the function. And it is set as 1 (default value) in this paper. $\|F_1 - F_2\|$ is the square of Euclidean distance.

At this time, the target and clutter samples data are linearly separable, and the hyperplane $\omega^T F\text{-}b = 0$ is determined according to the principle of maximum classification interval to distinguish the two kinds of data. The hyperplane in SVM, i.e., ω and b are determined by

$$\begin{aligned} &\min_{\omega,b,\xi} \frac{1}{2}\|\omega\|^2 + \beta \sum_{i=1}^{M_t} \xi_i \\ &\text{s.t. } y_i[k(\omega, F_i) - b] \geq 1 - \xi_i, \quad \xi_i \geq 0, \ i = 1, 2, \cdots, M \end{aligned} \quad (24)$$

where ξ_1 is a slack variable, which is used to enlarge the parameter solution ranges, β is the penalty parameter, which is used to set the tolerance for the error.

When training the SVM, the tolerance errors of different classifications are of little difference. For radar target detection, the false alarm rate, i.e., the misclassification ratio of the clutter signal, is strictly required; while the detection probability, i.e., the correct classification ratio of the target signal, allows appropriate compromise. Therefore, we propose a new 2C-SVM that different penalty parameters are set for the two types of targets, β_{target} and β_{clutter}. And by increasing β_{clutter}, the classification accuracy of the clutter signal is increased, i.e., reducing the false alarm rate [35].

$$\begin{aligned} &\min_{\omega,b,\xi} \frac{1}{2}\|\omega\|^2 + \sum_{i=1}^{M_t}\left(\frac{1+y_i}{2}\beta_{\text{target}} + \frac{1-y_i}{2}\beta_{\text{clutter}}\right)\xi_i \\ &= \min_{\omega,b,\xi} \frac{1}{2}\|\omega\|^2 + \beta_{\text{target}} \sum_{i=1}^{M_t}\left(\frac{1+y_i}{2} + \frac{1-y_i}{2}\beta_1\right)\xi_i \\ &\text{s.t. } y_i[i(k\omega, F_i) - b] \geq 1 - \xi_i, \quad \xi_i \geq 0, \ i = 1, 2, \cdots, M_t \end{aligned} \quad (25)$$

where $\beta_1 = \frac{\beta_{\text{clutter}}}{\beta_{\text{target}}}$. Determine the corresponding β_1 according to the required false alarm rate.

III. Marine Target Detection Method Based on MDCCNN-FACC

As shown in Fig. 4, the detection method includes three procedures: radar signal preprocessing, dataset construction, and model training. The data preprocessing phase extracts the time-frequency and amplitude information of the radar returns.

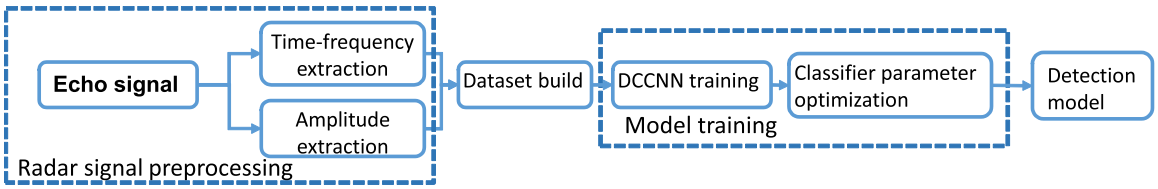


Fig. 4. Detection flowchart of the proposed DCCNN FACC.

TABLE I
DESCRIPTION OF IPIX DATASETS

Data ID	File name	Target rangebin	Guard rangebin	Wave height	Sea state (Dauglas)
IPIX_01#	19931108_220902_starea	7	6,8	0.9m	2
IPIX_02#	19931118_023604_stare	8	7,9-10	1.3m	3
IPIX_03#	19931107_135603_starea	9	8,10-11	2.2m	4

TABLE II
BRIEF INTRODUCTION TO CSIR DATASETS

Dataset	CSIR_01#	CSIR_02#
Date	04-Nov-2007	03-Aug-2006
Location	34°36′56.52″S 20°17′17.46″E	33°55′15.62″S 18°23′53.76″E
Altitude	67m	308m
Target Type	Rigid Inflatable Boat (5.7m)	
Start Time	11:44:45.855	14:02:42.609
Duration	49170 pats (49.17s)	524286 pats (104.857s)
Tx Frequency	8.8GHz	6.9GHz
PRF	1kHz	5kHz
Tracking Range	4672.76m	6600.00m
Range Extend	1499.0m(101 gates),15.0m res	720.0m(48 gates),15.0m res
Instant Wind	11.3kts, 320.0 N	15.3kts, 249.3 N
8hr Average Wind	3.65kts	12.2kts, 254.1 N
Wave	2.62m, 261.7 N	2.35m, 145.0 N
Grazing Angle	2.84 to 3.76 deg.	0.501 to 0.560 deg.
Antenna Azm	289.00 to 289.60 N	99.44 to 100.30 N
Antenna Elv	-3.252 to -3.159 deg.	-0.412 deg.
Sea state(Dauglas)	5	4
SCR	0.6dB	7dB

Then build the dataset using the information. The network parameters are trained and optimized and the CNN model that can be used for marine target detection is finally obtained.

A. Introduction of Radar Dataset

We use two types of internationally recognized radar sea clutter data for training and testing, i.e., intelligent PIXel processing radar (IPIX) data and radar data collected by council for scientific and industrial research (CSIR) [44]–[46]. Multiple datasets of measured data are used to reflect the detection performance of the proposed method for different observation conditions and target characteristics. The marine target in the IPIX dataset is a floating spherical block of styrofoam, wrapped with wire mesh, while in the CSIR dataset the target is maneuvering with different motions. Moreover, in these datasets, radar returns were selected under different sea state, i.e., low, medium and high sea states, and different polarizations, i.e., HH, VV, HV, and VH polarizations.

IPIX data was collected in 1993, by a fully coherent X-band radar [47], whose parameters and information can be found at http://soma.ece.mcmaster.ca/ipix/. The experiment field is at Osborne Head Gunnery Range (OHGR), Dartmouth, Nova Scotia, Canada, on a cliff facing the Atlantic Ocean, at the height of 100 feet above mean sea level and an open ocean view of about 130°. The target is a one-meter diameter spherical block of styrofoam, wrapped with wire mesh. The average signal to clutter ratio varies in the range 0-6 dB. Three datasets are selected as training and test data under different polarizations and sea states. Guard range bin is the range bin adjacent to the target range bin. The sea state in this paper is used as Douglas sea state (https://en.wikipedia.org/wiki/Douglas_sea_scale#cite_note-1). The IPIX data and CSIR data used in this paper are shown in Table I and Table II.

B. Data Preprocessing

In order to obtain input data for the two channels, the radar signal sequence is subjected to time-frequency analysis and modulo operation. In this paper, we mainly analyze and compare the outputs using different classical time-frequency analysis methods, e.g., STFT, WVD, and smooth pseudo WVD (SPWVD). The target and clutter signals are from IPIX data. WVD comes across severe cross-term effects, and the classification accuracy is low. For the signal of the same length, SPWVD is time-consuming, and therefore STFT is chosen for further analysis.

C. Dataset Construction

The data file IPIX_01# is used to construct the dataset, where the 7th range bin is the target bin, and the 1-5th, 9-14th range bins are clutter bins. The training set has 17000 clutter samples and 7872 target samples, and the test set has 44,033 clutter samples and 4000 target samples. The observation time is 131 s, and the sampling frequency is 1000 Hz. There are 10 datasets, each consisting of a time series of 131000 points. In order to balance the computational quantity of the time-frequency analysis and the output SCR, each sample sequence is set to the length of 1024. If all samples are entirely independent, about 100 target datasets and about 1000 clutter datasets can be obtained. Training the CNN based on such small samples may lead to over-fitting. Therefore, different samples have overlapped parts, and

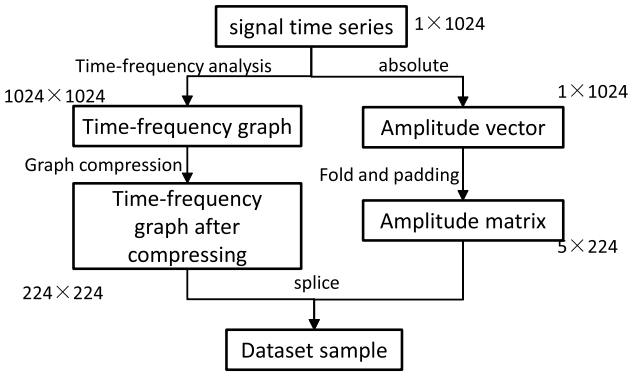


Fig. 5. Structure and dimension of training datasets.

the interval of 10-50 points is equally spaced according to the sampling frequency of different data. It is necessary to ensure that the training set and the test set are independent of each other. The training set is built with the first 90 s signal sequence, and the test set is built with the latter 41 s signal sequence.

During the training process, the dataset samples are input to the neural network randomly. The different features of a sequence are encapsulated as the input of the CNN. As shown in Fig. 5, each sample is a matrix of 229×224, which consists of two features of a signal sample. The compressed time-frequency graph occupies from row 1 to row 224, then row 225-228 and the first 128 points of row 229 are the amplitude vectors after folding. Examples of dataset samples are shown in Fig. 6. Finally, the samples are saved as images. And the amplitude values of each images are normalized to the interval [0, 1] at the input layer when processed by the network.

D. Model Training

1) MDCCNN Training: The training process is carried out in the Tensorflow 1.13 environment. The computer configuration includes CPU: i7-8700K, GPU: Nvidia 1080Ti, and computer memory 16 GB. When selecting the basic structure of MDC-CNN, the IPXI data of the second-degree sea state is used to test three classic CNNs: LeNet, VGG16, and ResNet. The batch size is 32, the learning rate is 0.01, the iteration number is 10000 times, and gradient descent optimizer is employed as the parameter optimization strategy [20],

$$\theta_i = \theta_i - \alpha \frac{\partial}{\partial \theta_i} J(\theta) \tag{26}$$

where i is the number of trainable variables. θ_i is model parameter to be optimized. α stands for learning rate, and $J(\theta)$ is the loss function. The cross-entropy is used as the loss function,

$$J(\theta) = -\sum_{i=1}^{M} y(\theta, \text{input})_i \log\left(y(\theta, \text{input})_{i\_}\right) \tag{27}$$

where $y(\theta, \text{input})_i$ is the desired output, and $y(\theta, \text{input})_{i\_}$ is the actual output.

Other parameters and settings of the MDCCNN model are as follows:

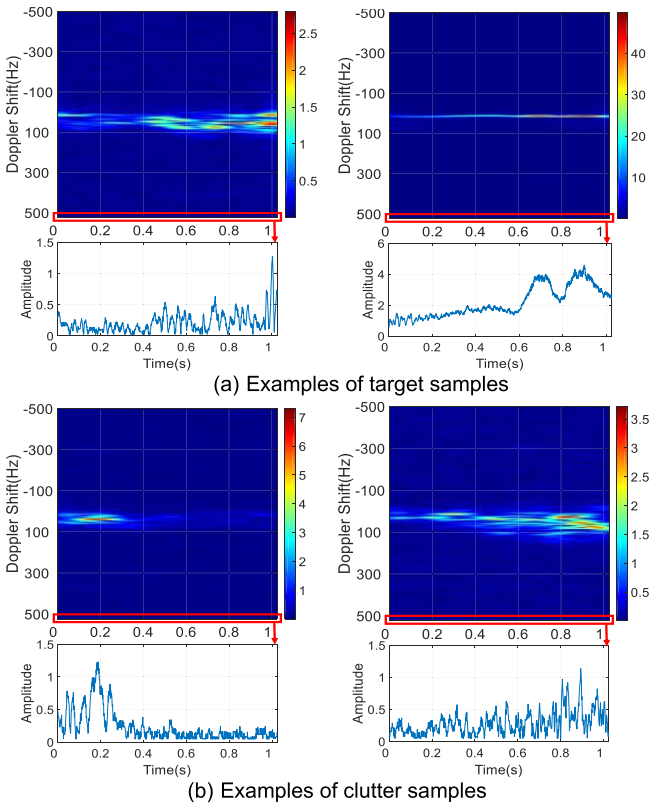


Fig. 6. Some examples of clutter and target data samples (IPIX01#).

a) Parameter initialization: The convolutional layer and the full connection layer weight are initialized by Xavier [20], and the offset is initialized to 0.

b) Pre-training: For the VGG16 gradient vanishing problem, the transfer learning method is used. Instead of the original initialization method, load the pre-training weight to initialize the weight and offset of the VGG16 convolution layer, and the gradient vanishing problem can be solved.

c) Model selection: ResNet may come across a serious over-fitting effect due to the low complexity of the samples and the small amount of data. We try to improve the method by increasing dropout and reducing the learning rate in the fully-connected layer. Finally, under the condition of the learning rate 105 and 50,000 iterations, only 59% (clutter) and 48% (target) classification accuracy are obtained. According to the test results, LeNet of one-dimensional convolution kernel is finally selected to process amplitude vector feature, and VGG16 and LeNet are used as the time-frequency graph feature extraction channels respectively. After constructing the MDCCNN structure, the corresponding weights of the previously trained single-channel CNN are loaded into the corresponding layers of the MDCCNN for initialization.

2) FACC Training: The purpose of the training process of the false alarm controllable SVM is to determine the penalty factor β_0 that satisfies the false alarm rate control condition. A set of vectors output from the trained MDCCNN output layer after inputting training data is used as training set during this process. The softmax classifier determines different thresholds

TABLE III
CLASSIFICATION ACCURACY OF DIFFERENT POLARIZATIONS

Polarizations	HH	HV	VH	VV
Accuracy rate (Target)	89.57%	87.04%	87.18%	88.70%
Accuracy rate (Clutter)	96.74%	96.15%	97.28%	97.08%
Average accuracy rate	93.25%	91.60%	92.23%	92.89%

TABLE IV
MDCCNN STRUCTURE AND FUSING MODES

	Channel 1 (Time-frequency graph)	Channel 2 (Amplitude vector)	Fusion mode
Model 1	LeNet	LeNet	Decision fusion
Model 2	LeNet	LeNet	Feature fusion
Model 3	VGG16	LeNet	Decision fusion
Model 4	VGG16	LeNet	Feature fusion

based on training set samples, labels, and given false alarm rates.

IV. EXPERIMENTAL TEST AND RESULTS ANALYSIS

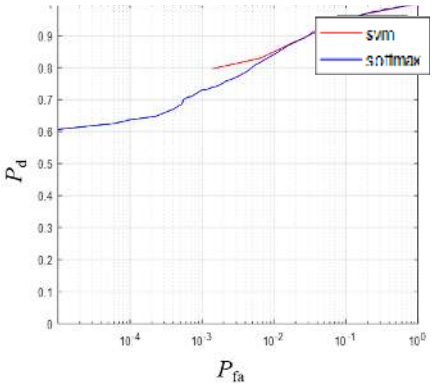
A. Performance Analysis of Target Detection Under Different Polarizations

The radar echo characteristics obtained by different polarizations have obvious differences. In order to analyze the influence of polarization modes on CNN models, the LeNet is first used for the analysis of the four polarization modes of IPIX data, i.e., HH, HV, VH, and VV. Each data file in IPIX data contains echo signal data obtained by four radar polarizations under the same conditions. The data IPIX_01# is used to train and test the model. The training set consists of 4500 clutter samples and 4500 target samples, and the test set includes 1500 clutter samples and 1500 target samples. The results are shown in Table III, which shows that under HH polarization [48], the characteristics difference of target and clutter are most apparent. Some works also suggested that generally, the HH polarization can achieve lower clutter scattering intensity [49]–[51], which is consistent with our results. Therefore, in the following analysis, data with HH polarization is employed or model training and testing to obtain higher classification accuracy.

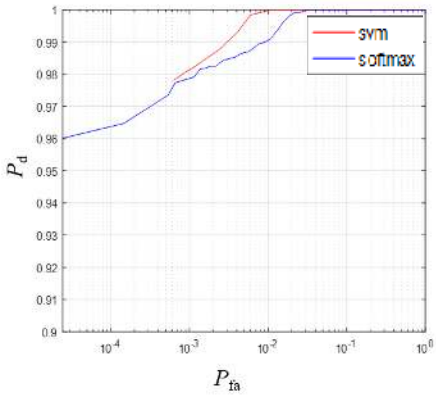
B. Performance Analysis of Different MDCCNN Models

In this part, we will analyze and compare several MDCCNN models. Considering the computational cost and performance, two CNN models, i.e., LeNet and VGG, are employed as the dual-channel feature extraction networks. The feature fusion network adopts two fusion modes: decision layer fusion and feature vector layer fusion. Therefore, the following four network structures are employed, which is shown in Table IV.

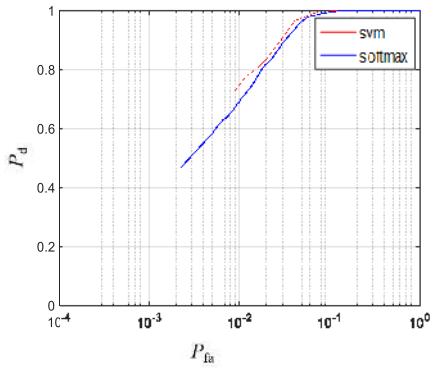
1) Performance Analysis of Feature Extraction Models: The IPIX_01# is used for building datasets, and different feature extraction methods are trained and tested. The performances of feature extraction models, the single-channel CNN, and traditional HOG-SVM image processing methods are compared, which are shown in Table V.



(a) IPIX_01# data



(b) IPIX_02# data

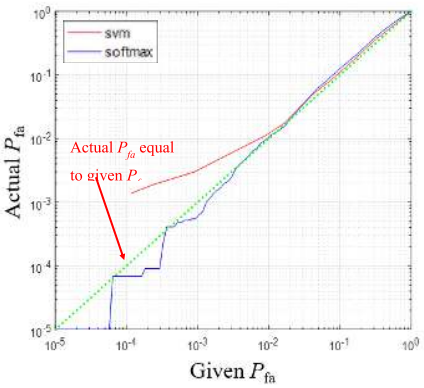


(c) IPIX_03# data

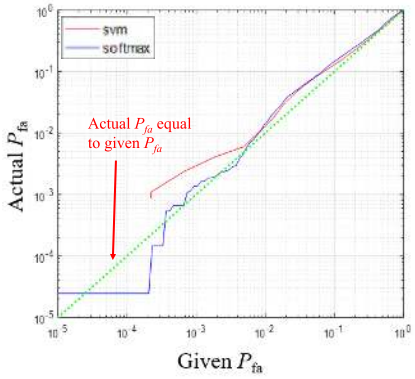
Fig. 7. ROC curves of different classifiers using IPIX data.

Compared with the Hog-SVM image classification method [40], CNN has apparent advantages in classification accuracy of target and clutter time-frequency graphs with faster classification speed. Compared with single-channel CNN, MDCCNN employs both amplitude and time-frequency information, which can significantly improve the target classification accuracy. Model 4 using VGG16 and LeNet to extract the time-frequency and amplitude characteristics can achieve an accuracy of 90.0% for the target samples. The calculation cost is reduced accordingly because the number of fully-connected layers in the feature extraction channel is reduced. Therefore, the classification time of MDCCNN is shorter than the single-channel VGG16 operation time.

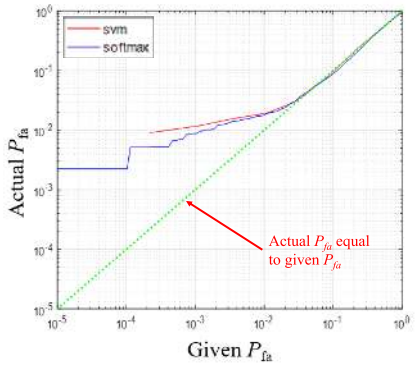
2) FACC Detection Performance Analysis: Model 4 is selected for feature extraction and fusion. The variable threshold softmax classifier and false alarm controllable SVM are



(a) IPIX_01# data

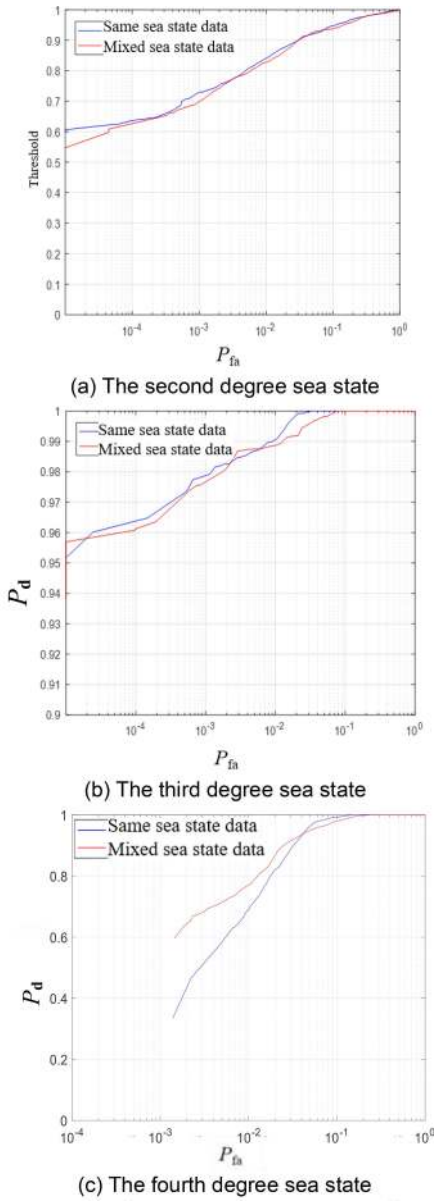


(b) IPIX_02# data



(c) IPIX_03# data

Fig. 8. False alarm loss curves of different classifiers using IPIX data.



(a) The second degree sea state

(b) The third degree sea state

(c) The fourth degree sea state

Fig. 9. ROC curves of different training sets under different sea states.

used as classifiers to classify the test set. Then receiver operating characteristic (ROC) curves of the test results (IPIX_01#, IPIX_02#, IPIX_03#) are given in Fig. 7, which indicates the relation between the false alarm probability (P_{fa}) and the detection probability (P_{d}).

Under the same P_{fa}, FACC-SVM can achieve higher P_{d}, but it has a problem of excessive P_{fa}. The P_{fa} on the test set is higher than that of the training set. Moreover, during the training process, the P_{fa} of the training set can reduce to 0. Since the test set data is different from the training set data characteristics, it would result in a mismatch between the model training parameters and the test data. When testing the model with the test set, the P_{fa} is still high and cannot be lowered further, which is called false alarm loss, as shown in Fig. 8. The x-axis is the desired P_{fa}, and the y-axis represents the actual P_{fa} while the variable threshold softmax classifier can achieve a lower P_{fa}.

The green dotted line in the Fig. 8 indicates that the actual P_{fa} is equal to the set P_{fa}. It can be seen that due to the difference between the characteristics of training data and test data, the actual P_{fa} will be higher than the P_{fa} during training. Compared to SVM, the actual P_{fa} of the softmax classifier is close to the training set P_{fa}, i.e., smaller false alarm loss. Therefore, the softmax classifier is more suitable for practical target detection.

C. Performance Analysis Under Different Sea States

In practical applications, the radar working environment is complex and changeable. The trained network needs to adapt to various sea conditions. The radar returns collected in different sea states are used for model training in order to improve the generalization ability of the model. Four different training datasets are used, including the second, third, and fourth-degree sea state, and samples of mixed sea states. Model 4 was trained with the four training sets, and the variable threshold softmax classifier is used as the output layer. The target detection performance under different sea states is shown in Fig. 9.

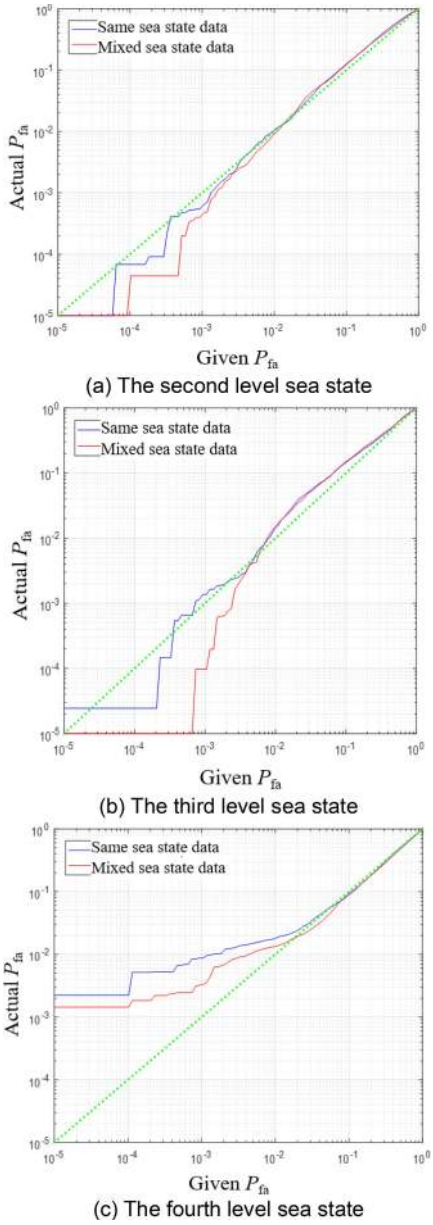
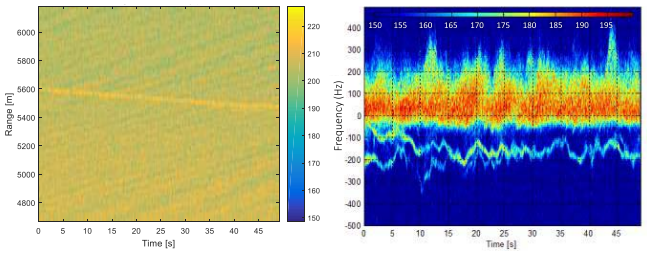


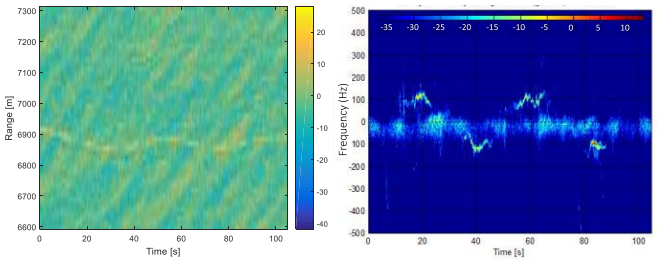
Fig. 10. False alarm loss curves of different training sets under different sea states using IPIX data.



(a) Range-Time intensity plot (b)Time-frequency of target (range bin 51)

Fig. 11. Description of CSIR_01# data.

In terms of P_d, the performances of the two training sets with the 2nd and 3rd degree sea states are quite similar. The performance of the model obtained by training the mixed sea state as the training set is better. Further analysis of the false alarm losses of the two methods is shown in Fig. 10. The test results show that training the model with a variety of sea states data can effectively reduce the false alarm loss. And



(a) Range-Time intensity plot (b)Time-frequency of target (range bin 19)

Fig. 12. Description of CSIR_02# data.

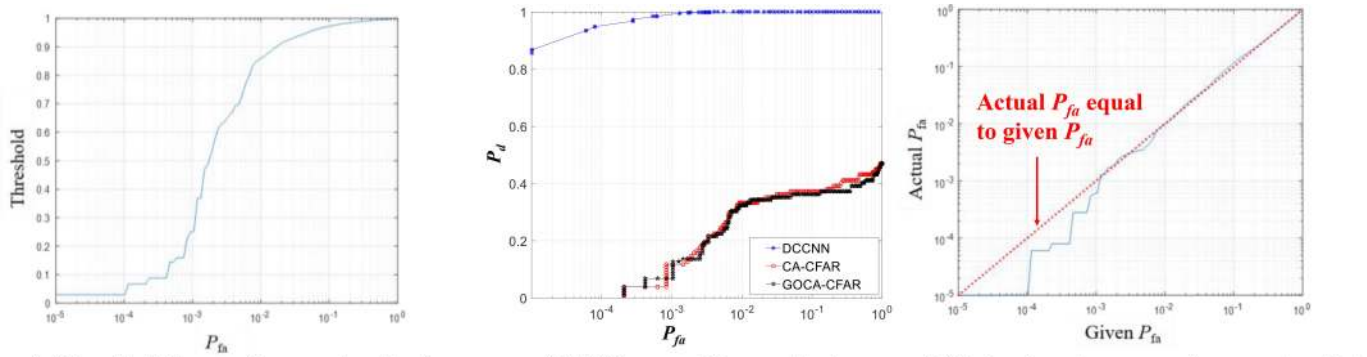
TABLE V
FEATURE EXTRACTION PERFORMANCE OF DIFFERENT MODELS

Model	Preprocessing	Feature extraction model	Fusion mode	Accuracy rate (target)	Time (4000 samples) (s)	Accuracy rate (clutter)
1	Ch1:STFT	LeNet	Decision fusion	85.250%	14.76	97.76%
	Ch2:amp	LeNet				
2	Ch1:STFT	LeNet	Feature fusion	84.625%	9.38	97.73%
	Ch2:amp	LeNet				
3	Ch1:STFT	VGG	Decision fusion	87.92%	19.22	98.00%
	Ch2:amp	LeNet				
4	Ch1:STFT	VGG	Feature fusion	90.00%	18.98	97.02%
	Ch2:amp	LeNet				
5	STFT	HOG-SVM		71.76%	181.72	95.06%
6	STFT	VGG		73.22%	21.19	99.94%
7	STFT	LeNet		83.78%	10.12	98.95%
8	amplitudes	LeNet		78.65%	11.06	90.95%

under higher sea state, the detection probability can increase by 25%. It should be noted that if the Doppler spectrum of sea clutter and target is largely different, i.e., not overlapped with each other, the P_d is usually higher. While for the second-degree sea state, the target and clutter Doppler spectrum are overlapped of and under the 4th degree sea state, the clutter is strong. Hence the detection probabilities of the two conditions are lower. But it does not change the conclusion, i.e., using mixed data training under different sea states would bring in lower false alarm losses.

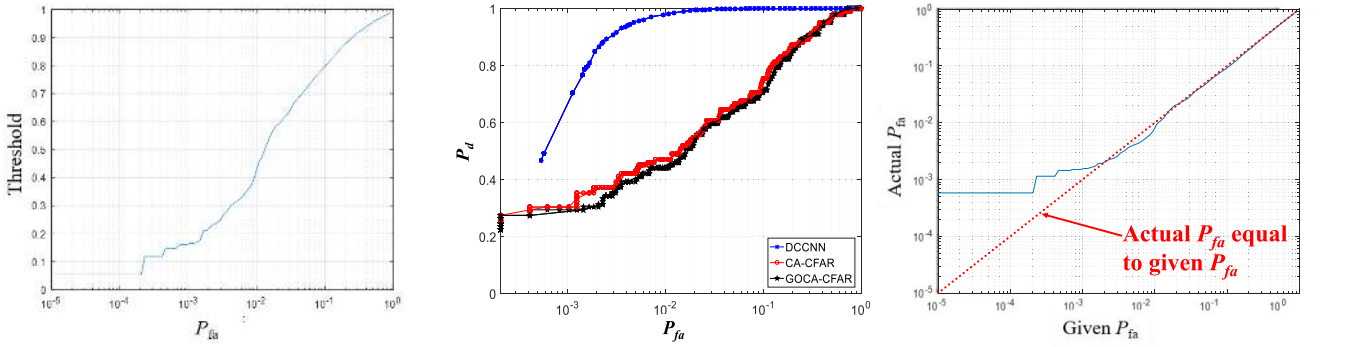
D. Maneuvering Target Detection Performance Analysis With X-Band CSIR Data

The target in CSIR_01# moves uniformly, while in CSIR_02#, the marine target maneuvers with high-order motion. The target time versus range bin image is shown in Fig. 11 and Fig. 12. The CSIR data can obtain the target real-time position data using GPS, and therefore the true target location can be confirmed as well. In this part, the detection results are compared with traditional CFAR methods, e.g., cell-averaging CFAR (CA-CFAR) [52] and greatest of CFAR (GO-CFAR) [53]. In most simple CFAR detection schemes, the threshold is calculated by estimating the noise level around the cell under test (CUT). This can be found by taking a block of cells around the CUT and calculating the average power level. A target is declared present in the CUT if it is both greater than all its adjacent cells and greater than the local average power level.



(a) Threshold-P_{fa} curve(proposed method) (b) ROC curve of three methods (c) False alarm loss curve of proposed method

Fig. 13. Detection results of CSIR_01# using different methods.



(a) Threshold-P_{fa} curve(proposed method) (b) ROC curve of three methods (c) False alarm loss curve of proposed method

Fig. 14. Detection results of CSIR_02# using different methods.

TABLE VI
DETECTION PERFORMANCE OF DIFFERENT METHODS FOR CSIR_01# AND CSIR_02# DATASETS

CSIR_01#				
Method		MDCCNN-FACC	CA-CFAR	GO-CFAR
Detection performance (P_d)	$P_{fa}=10^{-3}$	99.60%	11.76%	6.86%
	$P_{fa}=10^{-2}$	100%	33.33%	32.35%
Time Consuming (s)[1]		57.49	15.43	15.97
Time Consuming (s)[2]		2.94	1.52	1.61
CSIR_02#				
Method		MDCCNN-FACC	CA-CFAR	GO-CFAR
Detection performance (P_d)	$P_{fa}=10^{-3}$	70.37%	30.39%	29.41%
	$P_{fa}=10^{-2}$	97.83%	47.06%	44.12%
Time Consuming (s)[1]		58.70	15.56	15.81
Time Consuming (s)[2]		3.07	1.63	1.70

1: Calculation time including data preprocessing in section IV A.
2: Calculation time of the MDCCNN-FACC only.

The time series of 1024 lengths in the data matrix is divided into one sample, and all samples are binary-classified by the MDCCNN. The detection result of each range unit is compared with the target prior GPS information, and then the number and position of false alarms and missing samples are obtained for P_{fa} and P_d, which are shown in Fig. 13 and Fig. 14. Fig 13(a) and Fig 14(a) show the relationship between the classifier threshold and P_{fa} of CSIR01# and CSIR02#. Fig 13(b) and Fig 14(b) show the ROC curves of the three detection methods, i.e., MDCCNN, CA-CFAR, GO-CFAR. Fig 13(c) and Fig 14(c) show the false alarm loss of the two datasets. The detail detection performance of different methods for CSIR_01# and CSIR_02# datasets is shown in Table VI. It can be seen that the proposed MDCCNN-FACC can achieve a higher P_d under the same P_{fa}. Due to the data preprocessing in section III A (time-frequency analysis and amplitude segmentation), the proposed method is more time-consuming. However, only considering the detection process, MDCCNN-FACC, CA-CFAR, and GO-CFAR are almost the same.

V. CONCLUSION

In this paper, using deep learning, marine target detection is carried out by binary classification. The performance of different feature extraction models and classifiers is tested by using measured sea clutter and target signal data. The MDC-CNN detection method with the FACC is proposed, which realizes the application of the deep learning method in marine target detection. The time-Doppler spectrum and amplitudes information are both utilized, and the features are extracted by VGG16 and LeNet networks, respectively. Sea clutter and marine target are classified by the false alarm controllable SVM and the variable threshold softmax classifier, which can achieve more practical detection results. The following conclusions can be summarized: 1) The selection of training set greatly influences the performances of deep learning networks. Compared with single-channel CNN and traditional image processing methods, the dual-channel CNN has better feature extraction ability and higher classification accuracy for target and clutter samples. The feature layer fusion MDCCNN based on VGG19 and LeNet can achieve 90.00% target sample classification accuracy and 97.02% clutter sample classifica-

tion accuracy. 2) FACC-SVM can achieve higher detection probability. However, the false alarm rate is also higher, and the false alarm loss is severe. The softmax classifier is suitable for target detection applications with a lower false alarm rate. 3) Under high sea state, the model trained by the mixed sea states data can achieve a 25% increment for detection probability (false alarm rate 10^{-5}). Training the model with data collected under different sea states can help improve the detection performance and enhance the generalization ability. 4) Compared with traditional CFAR detectors, the proposed method achieves a higher detection probability under the same false alarm rate. And the computational quantity is almost the same, which indicates promising application prospects. In the future, more data will be collected to verify the algorithm under multiple observation conditions. Other signal features will be analyzed to improve the model performance to adapt complex maritime environment.

REFERENCES

[1] K. D. Ward and S. Watts, "Use of sea clutter models in radar design and development," *IET Radar, Sonar Navigat.*, vol. 4, no. 2, p. 146, 2010.

[2] A. Patel *et al.*, "Fine-resolution radar altimeter measurements on land and sea ice," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2547–2564, May 2015.

[3] A. Damini, M. McDonald, and G. E. Haslam, "X-band wideband experimental airborne radar for SAR, GMTI and maritime surveillance," *IEE Proc.-Radar, Sonar Navigat.*, vol. 150, no. 4, p. 305, 2003.

[4] X. Chen, J. Guan, N. Liu, and Y. He, "Maneuvering target detection via radon-fractional Fourier transform-based long-time coherent integration," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 939–953, Feb. 2014.

[5] E. R. Stofan *et al.*, "Overview of results of spaceborne imaging radar-C, X-band synthetic aperture radar (SIR-C/X-SAR)," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 817–828, Jul. 1995.

[6] X. Chen, J. Guan, N. Liu, W. Zhou, and Y. He, "Detection of a low observable sea-surface target with micromotion via the radon-linear canonical transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1225–1229, Jul. 2014.

[7] X.-K. Xu, "Low observable targets detection by joint fractal properties of sea clutter: An experimental study of IPIX OHGR datasets," *IEEE Trans. Antennas Propag.*, vol. 58, no. 4, pp. 1425–1429, Apr. 2010.

[8] Y. Yang, S.-P. Xiao, and X.-S. Wang, "Radar detection of small target in sea clutter using orthogonal projection," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 3, pp. 382–386, Mar. 2019.

[9] Y. Liu, S. Zhang, J. Suo, J. Zhang, and T. Yao, "Research on a new comprehensive CFAR (comp-CFAR) processing method," *IEEE Access*, vol. 7, pp. 19401–19413, 2019.

[10] J. Hu, W.-W. Tung, and J. Gao, "A new way to model nonstationary sea clutter," *IEEE Signal Process. Lett.*, vol. 16, no. 2, pp. 129–132, Feb. 2009.

[11] D. D. L. Mata-Moya, M. P. Jarabo-Amores, J. Martín de Nicolás, and M. Rosa-Zurera, "Approximating the Neyman–Pearson detector with 2C-SVMs. Application to radar detection," *Signal Process.*, vol. 131, pp. 364–375, Feb. 2017.

[12] Y. Gao, F. Gao, J. Dong, and S. Wang, "Transferred deep learning for sea ice change detection from synthetic-aperture radar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 10, pp. 1655–1659, Oct. 2019.

[13] S. Xu, J. Zheng, J. Pu, and P. Shui, "Sea-surface floating small target detection based on polarization features," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 10, pp. 1505–1509, Oct. 2018.

[14] X. Chen, J. Guan, Z. Bao, and Y. He, "Detection and extraction of target with micromotion in spiky sea clutter via short-time fractional Fourier transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1002–1018, Feb. 2014.

[15] X. Chen, G. Wang, Y. Dong, and J. Guan, "Sea clutter suppression and micromotion marine target detection via radon-linear canonical ambiguity function," *IET Radar, Sonar Navigat.*, vol. 9, no. 6, pp. 622–631, Jul. 2015.

[16] N. Lei, Y. Jin, L. Kang-Le, J. Wei-dong, and Z. Zhao-wen, "The inversion of sea state based on the micro-Doppler analysis of sea-clutter," in *Proc. IEEE 10th Int. Conf. Signal Process.*, Beijing, China, Oct. 2010, pp. 2206–2209.

[17] T. Thayaparan and S. Kennedy, "Detection of a manoeuvring air target in sea-clutter using joint time-frequency analysis techniques," *IEE Proc.–Radar, Sonar Navigat.*, vol. 151, no. 1, pp. 19–30, Feb. 2004.

[18] S.-N. Shi and P.-L. Shui, "Sea-surface floating small target detection by one-class classifier in time-frequency feature space," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6395–6411, Nov. 2018.

[19] X. Chen, J. Guan, Y. He, and J. Zhang, "Detection of low observable moving target in sea clutter via fractal characteristics in fractional Fourier transform domain," *IET Radar, Sonar Navigat.*, vol. 7, no. 6, pp. 635–651, Jul. 2013.

[20] M. Z. Alom *et al.*, "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, p. 292, Mar. 2019.

[21] J. Ding, B. Chen, H. Liu, and M. Huang, "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 364–368, Mar. 2016.

[22] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, Dec. 2017.

[23] Y. Yang, C. Hou, Y. Lang, G. Yue, Y. He, and W. Xiang, "Person identification using micro-Doppler signatures of human motions and UWB radar," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 5, pp. 366–368, May 2019.

[24] L. Wang, J. Tang, and Q. Liao, "A study on radar target detection based on deep neural networks," *IEEE Sensors Lett.*, vol. 3, no. 3, pp. 1–4, Mar. 2019.

[25] C. Hu, "Target recognition for marine radar using deep learning methods," *Inf. Res.*, vol. 44, no. 2, pp. 63–67, 2018.

[26] J. Wang, T. Zheng, P. Lei, Y. Zhang, and M. Qiao, "Hand gesture recognition method by radar based on convolutional neural network," *J. Beijing Univ. Aeronaut. Astronaut.*, vol. 44, no. 6, pp. 1117–1123, 2018.

[27] L. Zuo, M. Li, X. Zhang, Y. Wang, and Y. Wu, "An efficient method for detecting slow-moving weak targets in sea clutter based on time–frequency iteration decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3659–3672, Jun. 2013.

[28] A. Yasotharan and T. Thayaparan, "Time-frequency method for detecting an accelerating target in sea clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 4, pp. 1289–1310, Oct. 2006.

[29] N. Sun and X. Chen, "Detection and classification of maritime target with micro-motion based on CNNs," *J. Radars*, vol. 7, no. 5, pp. 565–574, 2018.

[30] P. Ghamisi *et al.*, "Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 1, pp. 6–39, Mar. 2019.

[31] H. Zhou and T. Jiang, "Decision tree based sea-surface weak target detection with false alarm rate controllable," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 793–797, Jun. 2019.

[32] X. Yang, G. Wen, C. Ma, B. Hui, B. Ding, and Y. Zhang, "CFAR detection of moving range-spread target in white Gaussian noise using waveform contrast," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 282–286, Feb. 2016.

[33] S. Uhlich, "Bayes risk reduction of estimators using artificial observation noise," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5535–5545, Oct. 2015.

[34] M.-P. Jarabo-Amores, M. Rosa-Zurera, R. Gil-Pita, and F. Lopez-Ferreras, "Study of two error functions to approximate the Neyman–Pearson detector using supervised learning machines," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4175–4181, Nov. 2009.

[35] Y. Li, P. Xie, Z. Tang, T. Jiang, and P. Qi, "SVM-based sea-surface small target detection: A false-alarm-rate-controllable approach," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1225–1229, Aug. 2019.

[36] A. B. Rosenkrantz *et al.*, "Prostate magnetic resonance imaging and magnetic resonance imaging targeted biopsy in patients with a prior negative biopsy: A consensus statement by AUA and SAR," *J. Urology*, vol. 196, no. 6, pp. 1613–1618, Dec. 2016.

[37] A. J. X. Guo and F. Zhu, "A CNN-based spatial feature fusion algorithm for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7170–7181, Sep. 2019.

[38] X. Li, M. Ding, and A. Pizurica, "Deep feature fusion via two-stream convolutional neural network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2615–2629, Apr. 2020, doi: 10.1109/TGRS.2019.2952758.

[39] J. Ai, R. Tian, Q. Luo, J. Jin, and B. Tang, "Multi-scale rotation-invariant Haar-like feature integrated CNN-based ship detection algorithm of multiple-target environment in SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10070–10087, Dec. 2019.
[40] W. Xing, N. Deng, B. Xin, Y. Liu, Y. Chen, and Z. Zhang, "Identification of extremely similar animal fibers based on matched filter and HOG-SVM," *IEEE Access*, vol. 7, pp. 98603–98617, Jan. 2019.
[41] S.-H. Kong, M. Kim, L. M. Hoang, and E. Kim, "Automatic LPI radar waveform recognition using CNN," *IEEE Access*, vol. 6, pp. 4207–4219, 2018.
[42] G. Wang, G. B. Giannakis, and J. Chen, "Learning ReLU networks on linearly separable data: Algorithm, optimality, and generalization," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2357–2370, May 2019.
[43] J. Zuk, S. Bocquet, and L. Rosenberg, "New saddle-point technique for non-coherent radar detection with application to correlated targets in uncorrelated clutter speckle," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2221–2233, Apr. 2019.
[44] P. L. Herselman, C. J. Baker, and H. J. de Wind, "An analysis of X-band calibrated sea clutter and small boat reflectivity at medium-to-low grazing angles," *Int. J. Navigat. Observ.*, vol. 2008, pp. 1–14, Nov. 2008.
[45] H. de Wind, J. Cilliers, and P. Herselman, "DataWare: Sea clutter and small boat radar reflectivity databases [best of the Web]," *IEEE Signal Process. Mag.*, vol. 27, no. 2, pp. 145–148, Mar. 2010.
[46] P. Herselman and C. Baker, "Analysis of calibrated sea clutter and boat reflectivity data at C- and X-band in South African coastal waters," in *Proc. IET Int. Conf. Radar Syst.*, Edinburgh, U.K., Oct. 2007.
[47] A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, "High resolution sea clutter data: Statistical analysis of recorded live data," *IEE Proc.-Radar, Sonar Navigat.*, vol. 144, no. 3, pp. 121–130, 1997.
[48] M. Belmonte Rivas, A. Stoffelen, and G.-J. van Zadelhoff, "The benefit of HH and VV polarizations in retrieving extreme wind speeds for an ASCAT-type scatterometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4273–4280, Jul. 2014.
[49] H. W. Melief, H. Greidanus, P. van Genderen, and P. Hoogeboom, "Analysis of sea spikes in radar sea clutter data," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 4, pp. 985–993, Apr. 2006.
[50] Y. Liu, S. J. Frasier, and R. E. McIntosh, "Measurement and classification of low-grazing-angle radar sea spikes," *IEEE Trans. Antennas Propag.*, vol. 46, no. 1, pp. 27–40, Jan. 1998.
[51] M. Fingas, *Oil Spill Science and Technology*, 2nd ed. Cambridge, MA, USA: Gulf Professional Publishing, 2016.
[52] A. Maali, A. Mesloub, M. Djeddou, H. Mimoun, G. Baudoin, and A. Ouldali, "Adaptive CA-CFAR threshold for non-coherent IR-UWB energy detector receivers," *IEEE Commun. Lett.*, vol. 13, no. 12, pp. 959–961, Dec. 2009.
[53] P. E. Pace and L. L. Taylor, "False alarm analysis of the envelope detection GO-CFAR processor," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 3, pp. 848–864, Jul. 1994.

Xiaolong Chen (Member, IEEE) was born in Yantai, Shandong, China, in 1985. He received the bachelor's and master's degrees in signal and information processing and the Ph.D. degree in radar signal processing from Naval Aviation University (NAU) in 2010 and 2014, respectively.

From 2015 to 2017, he had been a Lecturer with the Marine Target Detection Research Group, NAU, where he lectures "Radar Principle." He is an Associate Professor with NAU. He has published more than 80 academic articles, two books, and holds 39 national invention patents. He has given more than 20 presentations on radar signal processing, especially marine targets. He is a reviewer of IEEE TRANSACTIONS ON SIGNAL PROCESSING (TSP), IEEE SIGNAL PROCESSING LETTERS (SPL), IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (GRSL), IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATION AND REMOTE SENSING (J-STARS), *IET Radar, Sonar and Navigation (RSN)*, *IET Signal Processing (SP)*, *IET Electronics Letters (EL)*, *Digital Signal Processing (DSP)*, and many international conferences. His current research interests include radar signal processing, especially for marine target detection, moving target detection, micro-Doppler, and clutter suppression.

Dr. Chen has been serving as the Committee Member for CIE Youth Commission, and the Vice Executive Secretary of Radar and Information System Committee for CIE Young Scientist Club since 2018. He was the TPC member of the 2015 and 2018 IET International Radar Conference, and the 2019 IEEE International Conference on Signal, In-formation and Data Processing Conference (ICSIDP). He was selected in the Young Talents Program of China Association for Science and Technology (CAST) in 2016, and received the Excellent Doctor Dissertation of CIE. In 2017, he also received the Chinese Patent Award. He won four excellent paper awards at the 2016 International Radar Conference, the 2017 EAI International Conference on Machine Learning and Intelligent Communications (MLICOM 2017), the 14th National Radar Conference, and the 2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT 2019), respectively. In 2017 and 2018, his articles were selected as highly cited articles of *Journal of Radars*, and F5000 Top Articles from Outstanding S&T Journals of China. In 2019, he won the Civil-Military Integration Award of China Industry-University-Research Institute Collaboration Association (CIUR). He was selected for the Young Scientist Award both at 2019 URSI Asia-Pacific Radio Science Conference and 2019 International Applied Computational Electromagnetics Society Symposium (ACES), China. He was the Section Chair of the 2016 International Conference on Mathematical Characterization, Analysis and Applications of Complex Information, 2017 MLICOM, the 2017 International Conference on Radar Systems (U.K.), 2019 ICEICT, and 2019 ACES. He is also the Organizer of the special session "Recent Development on Radar Signal Processing" of 2019 ICEICT, special session "Radar Marine Target Detection and Recognition" of ACES 2019, and workshop "Advances in Radar Signal Processing and Target Recognition" of International Conferences on Communications Signal Processing and Systems (CSPS 2019). He has also been on the Editorial Board of the *Journal of Radars* since 2019. He has also been serving as an Associate Editor for IEEE ACCESS since 2018.

Ningyuan Su received the bachelor's degree in telecommunication engineering from the Huazhong University of Science and Technology (HUST) in 2017. He is currently pursuing the doctor's degree with Naval Aviation University (NAU).

His research interest includes deep learning-based maritime target detection.

Yong Huang was born in Hunan, China, in 1979. He received the M.S. and Ph.D. degrees in information and communication engineering from NAU in 2005 and 2010, respectively.

From 2011 to 2016, he had been a Lecturer with the Department of Electronic and Information Engineering, NAU. He is an Associate Professor with NAU. His current research interests include radar signal processing and clutter rejection.

Jian Guan received the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2000.

He is currently a Professor with NAU. His research interests include radar target detection and tracking, image processing, and information fusion. He has authored numerous articles in his areas of expertise and holds 21 national invention patents.

Prof. Guan is also a Senior Member of CIE and a Committee Member of the Radio Positioning Technology Branch in CIE. He has won the prize of the National Excellent Doctoral Dissertation, "Realistic Outstanding Youth Practical Engineering Award" of CAST, and was selected for the National Talents Engineering of Ministry of Personnel of China. He has served in the technical committee for many international conferences on radar.