Radar Maritime Target Detection Method Based on Decision Fusion and Attention Mechanism

Jurong Hu, Yongruo Sun, Mohammed Mutahar Abduljalil Shujaa Aldeen, and Ning Cao

Abstract This paper presents a decision fusion model based on two-channel convolutional neural network (DF-TCNN) as a way to solve the problems of insufficient feature representation, low performance detection, and poor fault tolerance in radar target detection based on deep learning. In the pro-posed model, a mean strategy is incorporated on each branch's predictions, and a decision fusion algorithm is applied to refine the classification results. Moreover, a dual-channel network structure with an attention mechanism is embedded for feature enhancement and learning adaptation. Verification of the radar data shows that the method offers a high fault tolerance rate and strong anti-interference ability, which can significantly improve radar target detection in the background of complex sea clutter.

Keywords Radar target detection · Decision fusion · Attention mechanisms · Convolutional neural network

1 Introduction

Target detection technology under the background of sea clutter is widely used in military and civilian settings, and it is a radar research hotspot [1]. Sea clutter has non-Gaussian, nonlinear, and non-stationary characteristics [2–4], which lead to mismatches of the statistical model which reduce the target detection performance of the constant false alarm rate (CFAR) method [5]. Contrary to CFAR detections, convolutional neural networks (CNN) [6–9] are data-driven and construct models using deep perceptual networks, which overcom1es the limitation of relying solely on statistical characteristics to simulate clutter distributions [10].

In Ref. [11], radar target detection is defined as a binary classification problem between targets and clutter, which is realized by using the Doppler domain information of the echo signal. To classify maritime targets, clutter, and coastline, CNN is used to classify the segmented maritime radar image samples in [12]. Since the

J. Hu (✉) · Y. Sun · M. M. A. S. Aldeen · N. Cao
Hohai University, Nanjing, Jiangsu Province, China
e-mail: hujurong@sina.com

target has a variety of motion characteristics, the Doppler velocity of the clutter unit tends to be large in low sea conditions, and the Doppler spectrum of the target and the clutter overlaps partially, detecting it with only a single feature is unreliable.

In this paper time frequency and amplitude characteristics of sea clutter were extracted using two-channel convolutional neural networks. Based on the viewpoints of feature fusion and feature extraction, the feature vector layer fusion model and the decision layer fusion model are then developed. Similarly, an attention mechanism is simultaneously implemented in the detection of sea clutter objects to improve the model's ability to extract features. For the proposed strategy, simulation analysis is used to determine the most suitable model, taking into account detection performance, parameter amounts, calculation complexity, and detection time.

2   Proposed Method Descriptions

In this section, a two-channel feature extraction network structure was developed, along with a convolutional attention module. The predicted results from each branch were fused at the decision layer, and the target detection problem was transformed into a binary classification problem using the Softmax classifier. The decision threshold was adjusted to effectively control the system's false alarm rate.

2.1   Fusion Attention Mechanism for Feature Extraction Networks

Convolutional attention module (CBAM) [13, 14] is integrated into the feature extraction network VGG16, and end-to-end training is performed together with VGG16 to form an improved feature extraction network. Figure 1 illustrates the network structure. There are two types of CBAM: Channel Attention Modules (CAM) and Spatial Attention Modules (SAM). In order to adapt the features, the two modules infer the attention weights in turn along the channel and space dimensions.

As shown in Fig. 2, an input feature map F is compressed to reduce its spatial dimension, and information about its spatial location is consolidated using average pooling and maximum pooling to create a $1 \times 1 \times C$ feature map: $F_{\max}^c$ and F_{avg}^c. As a result, the multilayer perceptron (MLP) of the hidden layer is applied to the two feature maps, and the output feature vectors are combined element-by-element. As a final step, the Sigmoid activation operation is used to produce a detailed feature map of channel attention A_c.

Formulas (1)–(3) show the calculation process:

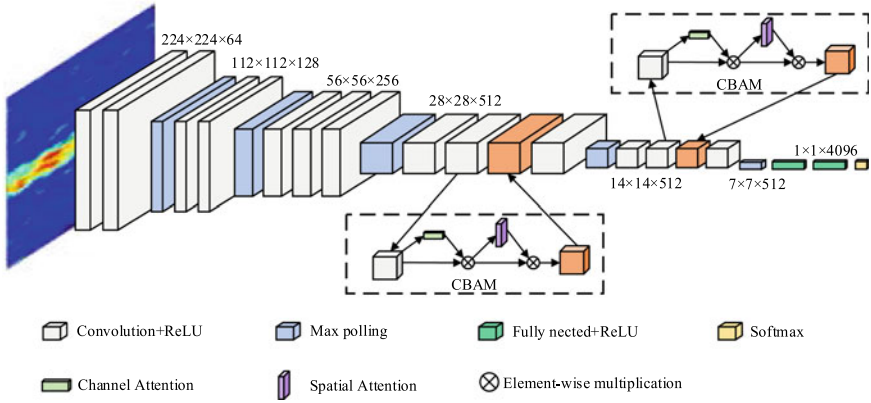$$A_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$
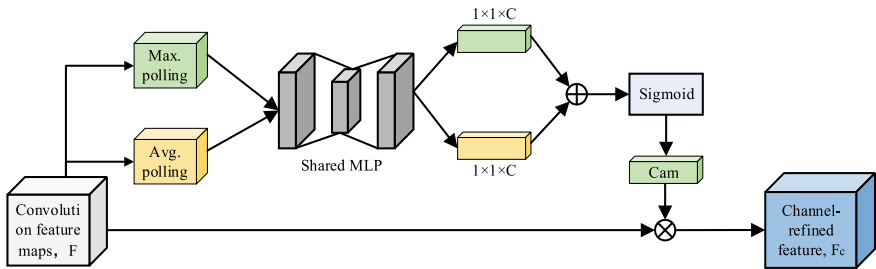
Fig. 1 Network of VGG16 feature extraction with CBAM

Fig. 2 Process of CAM operation

$$A_c(F) = \sigma\left(W_1\left(W_0\left(F_{\max}^c\right)\right) + W_1\left(W_0\left(F_{avg}^c\right)\right)\right) \tag{2}$$

$$F_c = A_c(F) \otimes F \tag{3}$$

where $Maxpool$ and $Avgpool$ represent average pooling and maximum pooling, W_0 and W_1 represent the two weights of MLP, σ represents sigmoid function, and $\otimes$ represents element-wise multiplication.

Figure 3 shows the operation process, which is the most informative part of the SAM. As a first step, the average pooling and maximum pooling operations will be performed in the channel dimension to obtain two feature maps of $H \times W \times 1$. As a result, two features are spliced together into a feature map of $H \times W \times 2$, and a 7×7 convolution kernel is used to reduce the dimension again into a feature map of $H \times W \times 1$. As a final step, the spatial attention map A_s is generated via the Sigmoid activation method and the final salient feature map F_s is obtained through element-wise multiplication.

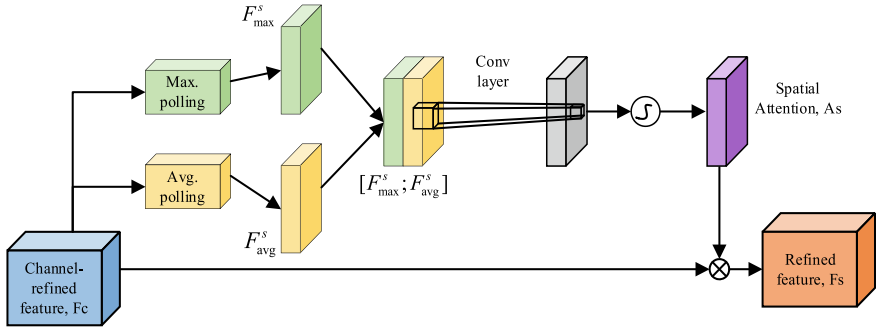Formulas (4)–(6) show the calculation process:

Fig. 3 Process of SAM operation

$$A_s(F_c) = \sigma(Conv_{7\times7}([MaxPool(F_c); AvgPool(A_c)])) \qquad (4)$$

$$A_s(F_c) = \sigma\left(Conv_{7\times7}\left(\left[F^s_{\max}; F^s_{avg}\right]\right)\right) \qquad (5)$$

$$F_R = A_s(A_c) \otimes F_c \qquad (6)$$

2.2 Algorithm for Decision Fusion

In this framework, decision-level feature fusion is implemented through two modules: classification and decision fusion. LeNet-5 and VGG16 are used as feature extraction channels in the classification module; the mean fusion algorithm is used in the decision fusion module (Fig. 4).

A VGG16 channel contains two fully connected layers with an output vector size of 4096×1, while a LeNet-5 channel has an output vector size of 120×1. In the last fully connected layer of each network branch, vector probability is calculated using
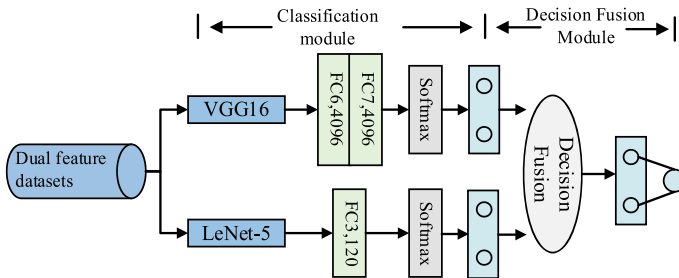

Fig. 4 Model based on fusion of decision layers

the Softmax function. During feature extraction, the weights and parameters of each layer of the trained network are loaded using the transfer learning method.

Each branch of the classification module's prediction results is fused at the decision level in the fusion module, denotes the prediction result of the i-th branch, then there are two branches, so $i \in \{1, 2\}$ is obtained. By using different fusion rules, $P_f = (p_{f,1}, p_{f,2})$ can be predicted for the final fusion module. In this paper, the inter-element mean strategy is used for fusion. The calculation rules for the fusion of the j-th element $P_{f,j}$ through the element mean strategy are as follows:

$$P_{f,j} = \frac{1}{2} \sum_{i=1}^{2} p_{i,j} \tag{7}$$

3 Experimentals and Analysis

The proposed method is comprised of three parts: data preprocessing, dataset construction, and model training. During the forward propagation of the dataset in the network model, preliminary prediction results are obtained. Subsequently, in the backpropagation process, the model weights are adjusted by computing the error between the predicted and expected values. This adjustment enables obtaining the optimal network parameter model, facilitating binary classification of targets and clutter.

3.1 Data Set Description and Settings

In this paper, a dual-feature dataset that combines sea clutter amplitude and Doppler velocity is produced for the parallel dual-channel feature network structure described in Sect. 2. IPIX sea clutter data is used for training and testing.

A sea clutter dual-feature image sample is made by splicing and packaging the data obtained from the same signal sequence after different preprocessing methods to ensure that the time–frequency map and amplitude map are identical. Figure 5 displays a double-featured image of sea clutter. The time–frequency features are represented by lines 1–224, while the compressed amplitude features are represented by lines 225–229.
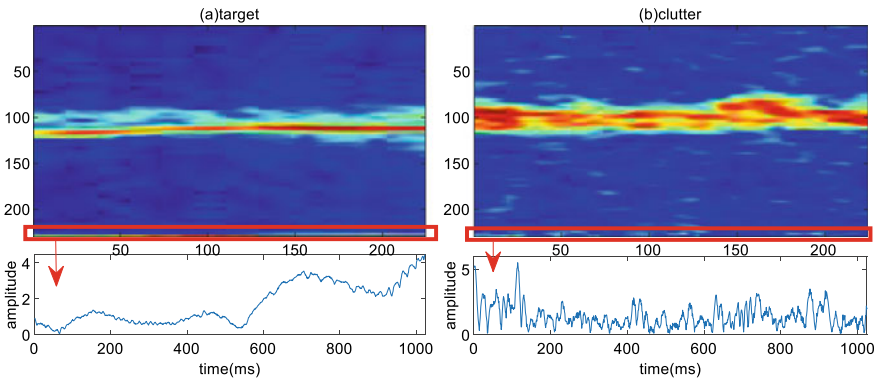
Fig. 5 Sea clutter dual feature dataset

3.2 Evaluations of the Proposed Method

As shown in Table 1, compared with models 1–4, VGG16 achieves higher clutter classification accuracy, while LeNet shows higher target classification accuracy. Moreover, compared with single-channel models 5–7, dual-channel models significantly improve target sample classification accuracy.

In order to verify the detection performance of Model 6 based on the decision fusion algorithm, we designed Model 5, which uses different feature fusion strategies but the same feature extraction channel model. The experimental results show that Model 6 has a target accuracy of 91.27% and a clutter accuracy of 98.33%. Target classification performance improved by 4.45%. Afterwards, the detection probability increased by 1.54% and 0.86% after the CBAM module was introduced in VGG16.

As shown in Fig. 6, Model 7 outperforms Models 5 and 6 if a variable threshold Softmax classifier is used. With false alarm probabilities greater than 10^{-2}, models 5

Table 1 Extraction model performance

| Model	Preprocessing	Feature extraction channel	Fusion method	Accuracy (Target)/%	Accuracy (Clutter)/%
1	AMP	LeNet-5	/	78.89	91.12
2	STFT	LeNet-5	/	85.12	94.34
3	AMP	VGG	/	76.52	92.86
4	STFT	VGG	/	84.23	97.67
5	Ch1:AMP Ch2:STFT	LeNet-5 VGG	Feature vector	86.82	98.69
6	Ch1:AMP Ch2:STFT	LeNet-5 VGG	Decision-making layer	91.27	98.33
7	Ch1:AMP Ch2:STFT	LeNet VGG (CABM)	Decision-making layer	92.81	99.19

Fig. 6 ROC curves for
model 5–7
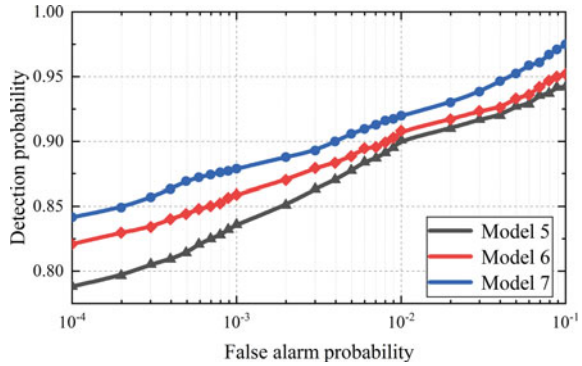
and 6 display similar detection performance, but both perform worse than model 7. When the false alarm rate is less than 10^{-3}, model 7 can achieve a higher detection probability, and when the false alarm rate is 10^{-4}, model 7 can still achieve a detection accuracy of 84.16%.

3.3 Analysis of Influencing Factors

In practical detection tasks, environmental factors such as wind speed, temperature, and weather can contribute to a complex and variable radar operating environment. Moreover, high-speed acquisition of sea surface information is necessary during detection, requiring control of radar dwell time. The shorter the dwell time, the shorter the observation time. Therefore, this section conducts experimental tests to verify the detection performance of the proposed detector under different observation durations and sea conditions.

Figure 7a presents a comparison of the detector's performance for different observation periods under the HH polarization. The results indicate that an increase in observation time significantly improves the detection probability. Notably, under high false alarm rates, increasing the observation time from 2048 to 4096 ms slightly improves the detector's detection performance. Moreover, when the observation time is only 256 ms and the false alarm probability is 10^{-3}, the detector's detection accuracy can exceed 80%, which can meet the requirements of low observation time, low false alarm probability, and high detection accuracy in practical applications.

Figure 7b shows detection of Class 2, Class 3 and Class 4 sea states with 1024 ms observation time and a false alarm probability of 10^{-3}. Due to the difference between the sea clutter and the Doppler spectrum of the target unit in the third sea state, the network is better able to extract and learn image features based on the differences between those two factors; Nevertheless, in level 2 sea states, the target Doppler spectrum overlaps with the clutter Doppler spectrum, which limits the performance of target detection; Similarly, when the sea state is level 4, backward electromagnetic
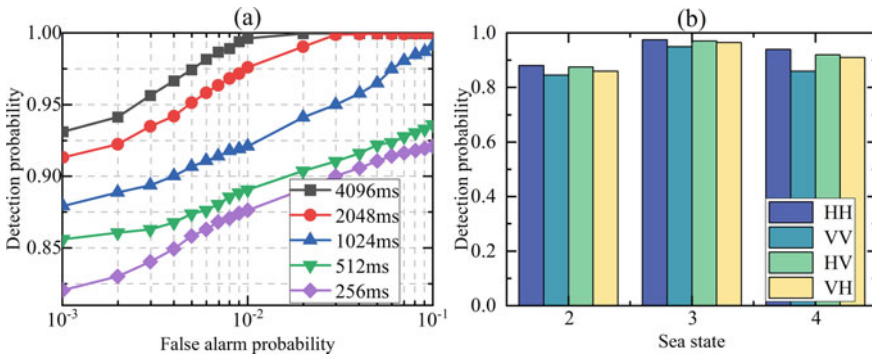
Fig. 7 Model detection figures under different observation durations and sea states

scattering characteristics of sea clutter are strong, and sea peaks due to waves and swells on the sea surface are similar to the target echo and have a high amplitude, making it easier to cover up the target. At this time, the detection probability decreases slightly.

4   Conclusion

As a part of this paper, radar signal target detection is converted into a binary classification problem, and the measured sea clutter and target radar signal data are used to test the performance of different feature extraction models, as well as an attention mechanism-based method.

We process radar signals by short-time Fourier transform and modulo method, VGG16 and LeNet networks are used for feature extraction, and CBAM is fused into the VGG16 network in the decision layer. The detection probability is 87.88% under the conditions of 10^{-3} false alarm rate, and 84.16% under the conditions of 10^{-4} false alarm rate.

References

1. Liu X, Xu S, Tang S (2020) CFAR strategy formulation and evaluation based on Fox's H-function in positive Alpha-Stable Sea clutter. Remote Sens 12(8)
2. Fiche A, Angelliaume S, Rosenberg L, Khenchaf L (2015) Analysis of X-band SAR Sea-clutter distributions at different grazing angles. IEEE Trans Geosci Remote Sens 53(8)
3. Zhihui et al (2017) Analysis of distribution using graphical goodness of fit for airborne SAR Sea-clutter data. IEEE Trans Geosci Remote Sens
4. Conte E, Maio AD, Galdi C (2004) Analysis of distribution using graphical goodness of fit for airborne SAR Sea-clutter data. IEEE Trans Aerosp Electron Syst 40(3):903–918

5. Yuan J, Ma X, Zhang C (2001) Real data based analysis of sea clutter characteristics. J Airforce Radar Acad
6. Oquab M, Bottou L, Laptev I (2014) Learning and transferring mid-level image representations using convolutional neural networks. Comput Vision Pattern Recogn IEEE 1717–1724
7. Chen Y, Jiang H, Li C et al (2016) Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans Geosci Remote Sens 54(10):6232–6251
8. Xu C, Hong X, Yao Y et al (2020) Multi-scale region-based fully convolutional networks. In: Intelligent computing and systems (ICPICS), pp 500–505
9. Carrera EV, Lara F, Ortiz M et al (2020) Target detection using radar processors based on machine learning. In: The technical and scientific conference of the Andean Council of the IEEE
10. Oquab M, Bottou L, Laptev I et al (2014) Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1717–1724
11. Wang L, Tang J, Liao Q (2019) A study on radar target detection based on deep neural networks. IEEE Sens Lett 3(3):1–4
12. Hu C (2018) Target recognition for marine radar using deep learning methods. Informatization Res 44(2):63–67
13. Woo S, Park J, Lee JY et al (2018) Cbam: convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV), pp 3–19
14. Yan Y, Kai J, Zhiyuan G et al (2021) Attention-based deep learning system for automated diagnoses of age-related macular degeneration in optical coherence tomography images. Med Phys 48(9):4926–4934