Sea Surface Small Target Detection on One-Dimensional Sequential Signals

Xiang YIN, Wanhua LI, Liulin WANG, Yu ZHAO

Dept. of Information Engineering, Yangzhou University, Yangzhou, China

yinxiang@yzu.edu.cn, 2326427269@qq.com, wangliulin1019@163.com, zhaoyu@yzu.edu.cn

Submitted April 17, 2024 / Accepted June 27, 2024 / Online first July 31, 2024

Abstract. *Existing sea surface small target detection methods typically rely on intricate feature extraction techniques on transformed radar returns. However, these approaches suffer from issues of high computational complexity and low real-time performance. Temporal Convolutional Network (TCN) can enable direct processing of radar time-series echo data without the need for elaborate feature extraction, thus substantially improving computational efficiency. Building upon this, this paper presents a novel target detection algorithm based on Multi-layer Attention Temporal Convolutional Network (MA-TCN). The proposed algorithm processes the amplitude information in the original echo signals, and comprehensively extracts sequence feature information through the construction of stacked residual modules. Additionally, it integrates multi-layer attention mechanisms to adaptively adjust the output weights of each residual module, thereby further enhancing detection accuracy. Experimental results demonstrate that the proposed approach achieves significant improvements in both detection performance and efficiency compared to existing methods.*

Keywords

Sea surface small target detection, Temporal Convolutional Network (TCN), Multi-layer Attention (MA), residual network

1. Introduction

Sea clutter often exhibits non-Gaussian, non-linear, and non-stationary characteristics [1]. The electromagnetic waves scattered by sea clutter manifest strong backscattering, resulting in a low signal-to-clutter ratio between radar signals and clutter. This makes it easy for target signals to be submerged and difficult to detect. Therefore, small target detection in sea clutter background remains a crucial and challenging task in the field of maritime radar signal processing [2]. From the perspective of detection methodologies, current approaches mainly fall into two categories: traditional signal processing-based methods and machine learning and deep learning-based methods.

Extensive work has been conducted in the traditional field of sea surface target detection. In [3], Kelly proposed a generalized likelihood ratio detector based on the assumption that sea clutter conforms to a complex Gaussian distribution. Robey et al. [4] further proposed an effective matched filtering technique. However, the non-Gaussian nature commonly observed in sea clutter imposes constraints on the applicability of traditional distribution models. To gain a comprehensive understanding of the complexity of sea clutter, Haykin et al. [5] constructed a nonlinear prediction model based on the chaotic characteristics and short-term predictability of sea clutter. In [6], Lo et al. proposed a single-feature detector based on the Hurst exponent, leveraging the distinct fractal characteristics exhibited by targets and sea clutter. Shui et al. [7] proposed a detector based on three features. Shi et al. [8] proposed a three-feature detector based on the two-dimensional time-frequency domain. The majority of the aforementioned methods leverage the convex hull learning algorithm for target detection. However, the applicability of this algorithm becomes challenging when confronted with feature spaces comprising more than three dimensions due to the data dimension explosion issue associated with the convex hull algorithm. Xu et al. [9] proposed a sea surface small target detection algorithm based on joint features and the Support Vector Machine (SVM). Guo et al. [10] combined the K-Nearest Neighbor (KNN) algorithm with anomaly detection principles. They established a high-dimensional feature space with eight features. By adjusting parameters to control the false alarm rate, they avoided issues related to feature compression and dimensionality constraints, thus further bolstering the performance of the detector.

In recent years, deep learning techniques have garnered widespread application, showcasing remarkable performance across diverse domains. Su et al. [11] used a Convolutional Neural Network (CNN) to detect the micro-motion Doppler of offshore targets. Su et al. [12] utilized the Long Short-Term Memory (LSTM) network and CNN to detect sea surface targets in one-dimensional sequence signals. Furthermore, the team proposed an intelligent detection method based on the fusion of features using a dual-channel convolutional neural network, achieving superior detection performance [13]. Wang et al. [14] devised a dual-perspective attention-based maritime radar target detector employing CNN. Wan et al. [15] presented a detection method

DOI: 10.13164/re.2024.0463

based on the LSTM network, demonstrating favorable detection performance on real-world datasets. Qu et al. [16] proposed a method for small sea surface target detection based on enhanced CNN with controllable false alarm rates. Shi et al. [17] designed a detector based on dual-channel convolutional neural networks. Both methods achieved high-performance detection with globally controllable false alarm rates for high-dimensional features.

The above feature-based detection methods focus on improving detection probability and have demonstrated commendable performance. However, within the realm of small target detection on the sea surface, detection probability is not the only criterion for evaluating detectors. Real-time performance assumes equal significance. A detection system with strong real-time capabilities can rapidly respond to environmental changes and the appearance of targets, particularly in dynamically complex marine environments. This capability is imperative for promptly identifying and addressing potential security threats. Inadequate real-time performance not only delays information acquisition and processing but may also lead to erroneous decision-making, thereby impacting the overall performance of the monitoring system [18]. Therefore, the reduction of processing time, coupled with the decrease in computational complexity, stands as pivotal for improving the real-time performance of the detector, thereby streamlining its engineering applications.

However, most of the existing small target detection algorithms initially transform the radar echo signals before extracting features in the transformed domain, which entails considerable computation and time. Simultaneously, in order to extract finer features, it is necessary to continuously increase the depth of the network until the receptive field covers the entire signal. This results in a deep network structure and increased computational complexity. In contrast, TCN can directly utilize the raw echo sequences for detection without the need for transforming them. Leveraging a residual structure that incorporates operations such as dilated causal convolutions and skip connections, TCN can process the same input information with fewer network layers, thus mitigating information loss, reducing computational complexity, and achieving faster detection speeds.

This paper proposes a novel small target detection method for one-dimensional temporal signals from maritime radar, leveraging the deep learning principle by integrating TCN with multi-layer attention mechanisms. Firstly, the raw radar echo signals are preprocessed, followed by a proportional division of the dataset. Subsequently, an enhanced TCN architecture is employed to extract the temporal correlations within the echo signals, incorporating stacked residual modules to deepen the model for more comprehensive sequence feature extraction. Then, the multi-level attention mechanism is used to reasonably allocate attention when the signal is input. Additionally, a false alarm controllable method is adopted to enable target detection under preset false alarm rate constraints. Experimental results demonstrate the superior detection performance of the proposed detector, coupled with significantly reduced detection time, thereby facilitating its practical implementation.

The paper is organized as follows. Section 2 describes the problem of small target detection on the sea surface. Section 3 introduces a novel detector that integrates an enhanced Temporal Convolutional Network (TCN) with the attention mechanism. Section 4 presents the datasets used, namely IPIX and SDRDSP, along with the experimental findings, comparing them with the results obtained by other detectors. Section 5 concludes the paper.

2. Description of Detection Issues

Assuming there is a coherent radar system for sea surface search, which emits N consecutive coherent pulses per dwell, and employs I/Q channels for lossless reception during the receiving process. Through sampling processing of the echo data, a dataset containing an N-dimensional complex echo sequence is obtained where the echo data for the k-th range cell can be represented as:

$$x_k = \left[x(1),\cdots,x(n),\cdots,x(N)\right]^{\mathrm{T}} \tag{1}$$

where $[\cdot]^{\mathrm{T}}$ represents transpose.

In radar target detection, data is typically partitioned along the fast-time dimension into detection range cells, reference cells, and guard cells. Specifically, detection range cells constitute the primary region subject to sliding-window detection, while guard cells are regions proximal to the detection range cells utilized to prevent mutual interference between adjacent cells. Reference cells, on the other hand, serve as comparative reference regions [19]. As the sliding of the distance unit under inspection proceeds, the reference unit and the protection unit are reselected accordingly. According to the literature [7], under the premise of ignoring the noise component in the echo signal, the problem of detecting small targets on the sea surface can be summarized as the following binary hypothesis testing problem:

$$\begin{aligned} H_0 &: \begin{cases} z(n) = c(n), & n = 1,2,3,\cdots,N \\ z_p(n) = c_p(n), & p = 1,2,3,\cdots,P \end{cases}, \\ H_1 &: \begin{cases} z(n) = s(n) + c(n), & n = 1,2,3,\cdots,N \\ z_p(n) = c_p(n), & p = 1,2,3,\cdots,P \end{cases}. \end{aligned} \tag{2}$$

In the given equation, $z(n)$ and $z_p(n)$ represent the radar echo data received by the unit under test and the surrounding reference units, respectively. $c(n)$ and $c_p(n)$ represent the time series of pure clutter for the unit under test and the surrounding reference units, respectively, while $s(n)$ represents the time series of target echoes. P denotes the number of reference units, and N represents the pulse length. Under the null hypothesis H_0, the time series received by the unit under test consists solely of clutter. Under the alternative hypothesis H_1, the time series received by the unit under test contains echoes from the target [10].

3. Design of a Detector Based on MA-TCN

3.1 Improved TCN Model

In the field of sea surface small target detection, current deep learning models predominantly suffer from high computational complexity and poor real-time performance. These hurdles greatly impede the practical application of deep learning models in this domain. The TCN architecture presents a promising solution to mitigate these challenges effectively [20], [21]. Nonetheless, TCN tends to treat the feature outputs of each residual block with equal significance in influencing the final classification outcome. Such a method may overlook valuable information, potentially compromising the performance of the detector. To address this limitation, the integration of a multi-layer attention mechanism becomes imperative. This mechanism dynamically enhances the representation of pertinent features for the classification task while suppressing those with minimal relevance. Consequently, this study proposes a novel one-dimensional temporal target signal detection algorithm based on MA-TCN. By augmenting the TCN architecture and incorporating a multi-layer attention mechanism, the proposed detector exhibits enhanced robustness and discriminative capability, thereby offering a viable solution for sea surface small target detection.

The detector employs an end-to-end approach which enables the model to autonomously and efficiently handle one-dimensional time series data. As illustrated in Fig. 1, the detector consists of four components, namely a data preprocessing module, stacked residual modules, a multi-layer attention mechanism module, and a classification module. Initially, the input I/Q complex data undergoes preprocessing to map the complex signals into real numbers, and the echo data is partitioned into training and testing sets. Subsequently, the model utilizes a residual module to extract temporal features of the echo data through convolutional operations. With a stacked structure, the network deepens its layers, facilitating the extraction of increasingly abstract and profound features to further enhance the representation capability of the model. Furthermore, the incorporation of a multi-layer attention mechanism enables adaptive scaling of the output weights of each residual module, thereby refining the model's attentional focus and enhancing its discriminative ability. Finally, a classification module is employed to classify the processed features, and a false alarm controllable algorithm is incorporated to obtain detection probabilities at a given false alarm rate.

3.1.1 Data Preprocessing Module

TCN is a neural network architecture rooted in convolutional operations, it primarily serves as a tool for processing temporal data. Since convolutional operations are essentially conducted in the real domain, the input data of TCN networks are generally assumed to be in real number form. However, as radar echo signals are typically represented as complex numbers (I/Q), a transformation of the data is imperative to align with the requirement of TCN. This paper focuses on analyzing the data from the amplitude perspective, with the calculation method as follows:

$$x(n) = |I + \mathrm{j}Q|, \quad n = 0,1,2,\cdots,N-1. \tag{3}$$

In the expression, $x(n)$ represents the magnitude of the echo signal; I and Q respectively represent the in-phase component and quadrature component of the received radar echo signal.

This article partitions data of unit length N into multiple small-scale vectors of length K to generate sufficient samples.

$$u_j = x\big(m(j-1)+1 : m(j-1)+M\big). \tag{4}$$

In the equation, $j = 1,2,\ldots,m$ is a constant used to adjust the length of overlap between adjacent vectors.

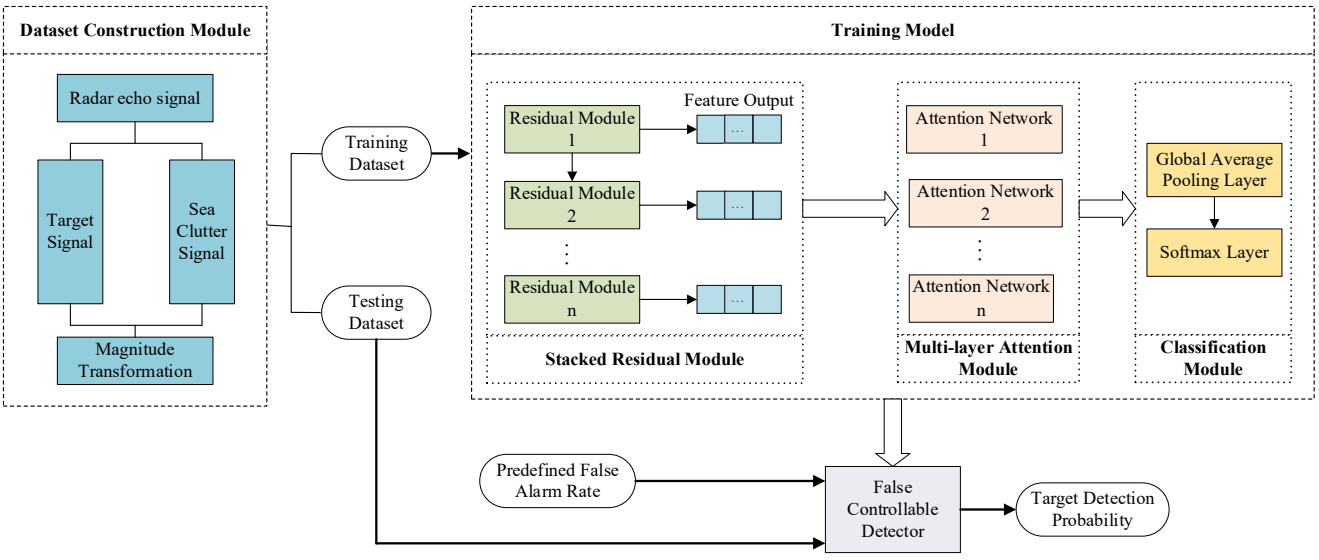


Fig. 1. The structural diagram of the MA-TCN model.

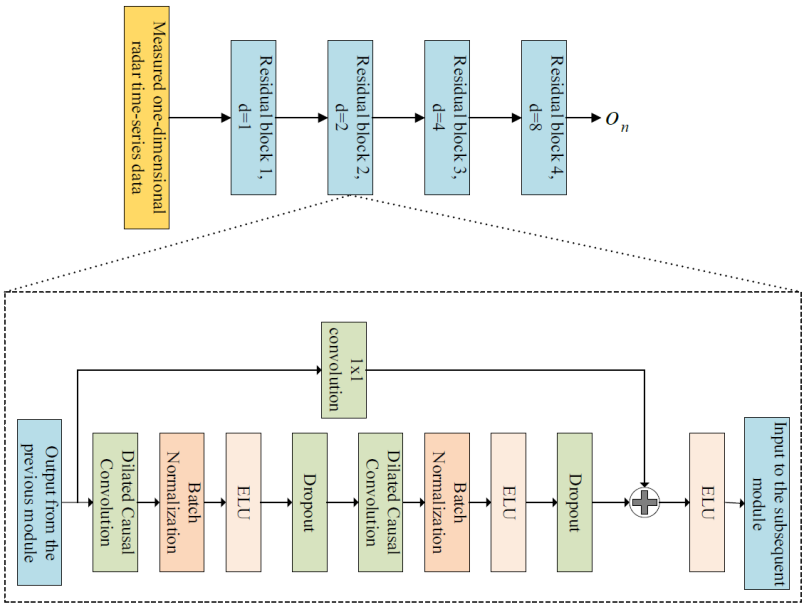


Fig. 2. Diagram of stacked residual module structure.

3.1.2 Stacked Residual Module

The paper enhances the capability of the model to extract deeper-level features from radar echo signals by stacking multiple residual blocks. The configuration is delineated in the upper segment of Fig. 2. Herein, the input of a certain residual module is the output of its previous one. Different residual modules employ dilated convolutions with varying dilation factors d, where d increases exponentially with the depth of the residual module.

As illustrated in the lower portion of Fig. 2, each residual module comprises a series of mapping transformations denoted as F, which consist of dilated causal convolutions, batch normalization, the non-linear mapping function, dropout, and residual connections. This framework forms a multi-layer feature extraction and temporal modeling structure. Such a design enables parallel computation and allows adjustment of the receptive field size of the convolutional layer through dilation factors, facilitating the processing of radar echo data with fewer convolutional layers.

The output of each residual module not only serves as the input for the subsequent residual module but also contributes to a multi-layer attention mechanism. The output of the first residual block is:

$$o = \text{Activation}\left(x + F\left(x\right)\right) = T\left(x\right). \tag{5}$$

To better handle complex temporal data, this paper proposes improvements within the framework of the fundamental TCN. The following section will detail the modifications made to the residual module.

a) Dilated causal convolution

The dilated causal convolution designed in this paper is shown in Fig. 3, with a convolution kernel size of 3 and an expansion coefficient of 2. Assuming the input samples are denoted as $X = \{x_1, x_2, \ldots, x_n\}$, and the convolution kernel as $F = \{f_1, f_2, \ldots, f_K\}$. Then, combining with reference [20], the output at position x_t after dilated causal convolution can be defined as:

$$\left(F * X\right)\left(x_t\right) = \sum_{k=1}^{K} x_{t-d(K-k)} f_k \ . \tag{6}$$

b) Batch normalization, non-linear mapping, and regularization techniques

The original TCN uses Weight Normalization (WN) to standardize network parameters, ensuring consistent weight scales and facilitating gradient descent, but neglects to standardize the input features, which may lead to instability in network parameters. Therefore, this paper adopts Batch Normalization (BN) instead of WN when designing the detector. BN normalizes the inputs within each training batch, ensuring that the mean of each feature approaches 0 and the standard deviation approaches 1. This process alleviates gradient vanishing and exploding while enhancing the training speed and stability of the network. Let the normalized features be denoted as $Cov_{\text{BN}} = [\tilde{C}_1, \ldots, \tilde{C}_M]$.

Normalization itself is a linear operation and cannot provide sufficient nonlinearity for effective feature representation. Therefore, to meet the demands for intricate and nonlinear feature extraction, this paper introduces the nonlinear activation function ELU after the batch normalization operation, which is linear on the positive interval and has values in the negative interval. In this case, the nonlinear transformation features are:

$$\hat{C}_m\left(i,j\right) = \text{ELU}\left(\hat{C}_m\left(i,j\right)\right). \tag{7}$$

Introducing non-linear mapping after batch normalization can enhance the adaptability of TCN networks to complex patterns and non-linear relationships within the data.

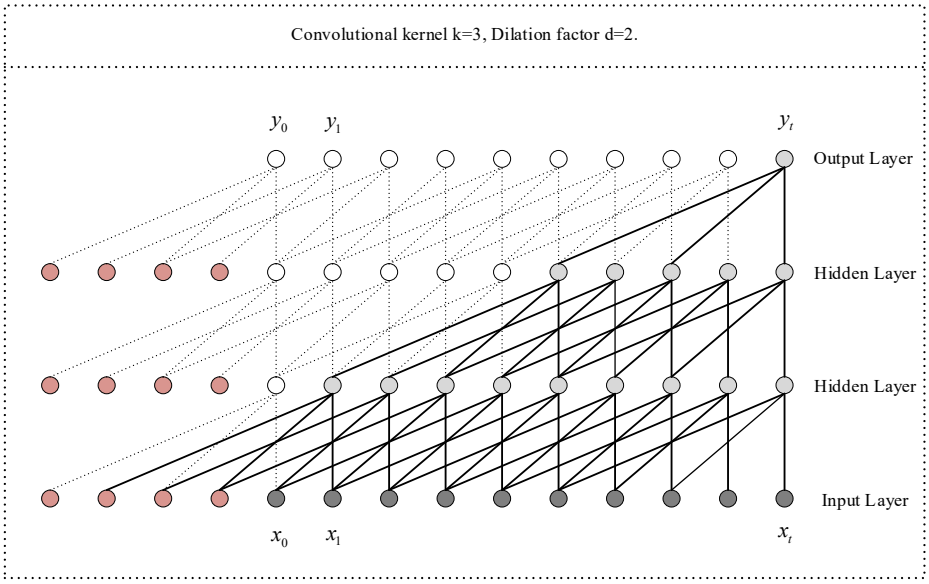


Fig. 3. The schematic of dilated causal convolution in the MA-TCN model.

c) Residual connection

Residual connections, also known as skip connections, establish direct connections in the network that span layers, adding the output of the previous layer to the output of subsequent layers, thereby constituting a "residual block". The principle is shown in Fig. 4.

In this paper, during the feature extraction of time series data, a skip connection mechanism is employed to add the input x of the module to its mapping transformation output $F(x)$. Additionally, a 1×1 convolution operation is utilized to adjust the channel dimension of the input x, ensuring it matches the channel number of the mapping transformation output $F(x)$. The output of each residual module unit, denoted as formula (5), is preserved as an intermediate result, from which subsequent multi-layer attention modules derive their inputs.

3.1.3 Multi-layer Attention Module

The attention mechanism is a resource allocation strategy aimed at focusing computational resources on the most informative parts of the signal. When dealing with one-dimensional time-series data, different segments contribute disparately to classification results. Some segments may exhibit similar patterns across different sequences, making it challenging to effectively differentiate sequences based on these segment features. On the other hand, certain segments may be unique to a particular sequence, playing a significant role in classification [22]. Based on this viewpoint, this paper combines TCN with attention mechanisms, empowering the detector to adaptively scale the feature weights involved in the classification process. This dynamical scaling highlights features beneficial for the classification while suppressing the impact of task-irrelevant features.

Various scenarios require different attention mechanisms, and the Self-Attention mechanism is the one that can associate distinct features within a single sequence. By simultaneously attending to all features in the same sequence, it calculates the response to each feature in the sequence. This implies a focus on the input sequence itself to discover the inherent connections within the sequence. Notably, the Self-Attention mechanism exhibits parallel computation capabilities, enabling efficient computation of correlations between representations at each position in the sequence. This attribute renders it well-suited for handling time-series data with complex spatiotemporal correlations such as radar echo data. Through this mechanism, comprehensive capture of correlations between different time steps and different positions can be achieved, aiding in more accurately modeling the dependency relationships among radar echo data. The principle of the Self-Attention mechanism is illustrated in Fig. 5.

In addition, to capture features with varying degrees of abstraction, this paper integrates an attention mechanism calculation on the output of each residual module, namely, a multi-layer attention module, as illustrated in Fig. 6. Assuming the output of each residual module is $T_i(x) \in R^{C \times n}$, where C denotes the number of channels and n represents the sequence length.

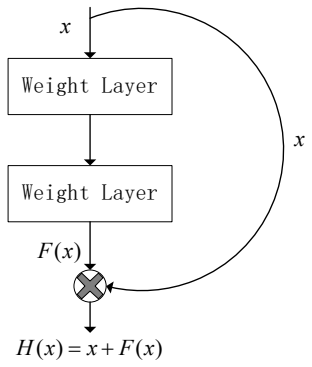


Fig. 4. Residual connection schematic diagram.

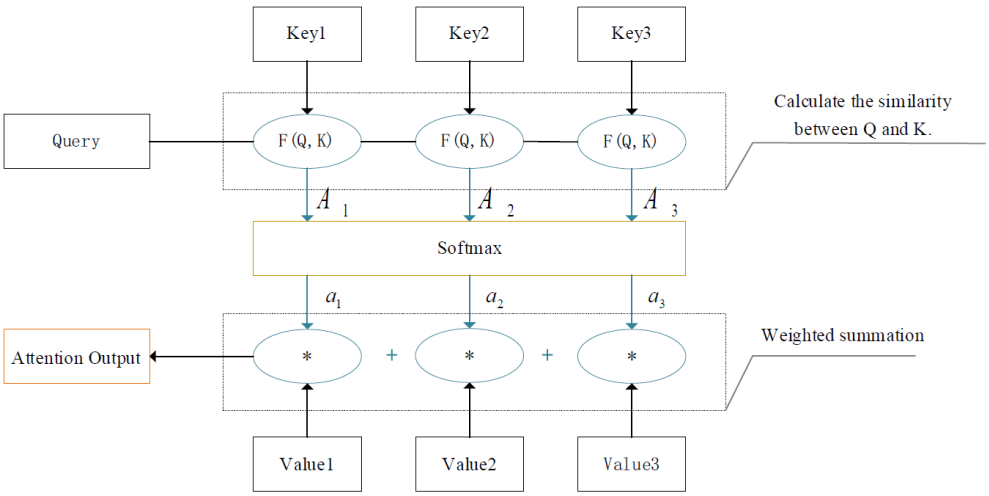


Fig. 5. The schematic diagram of self-attention mechanism.

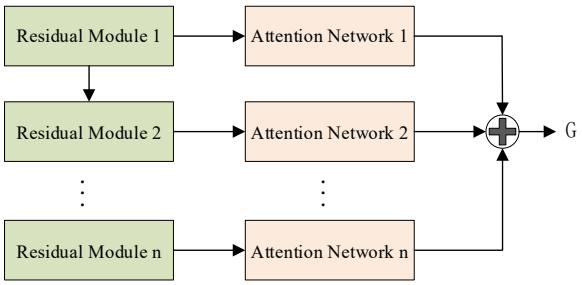


Fig. 6. Multi-layer attention module.

The output of the multi-layer attention module in this model is the weighted sum of the output features of K residual modules after passing through the attention mechanism. This can be expressed as:

$$G = \sum_{k=1}^{K} \sum_{j=1}^{C} a_{kj} H_{kj} \tag{8}$$

where a_{kj} and H_{kj} respectively represent the weight coefficient and output feature of the j-th channel in the k-th residual module. The formula for calculating the weight coefficient is:

$$a_{kj} = \text{softmax}\left(\tanh\left(\omega_k H_{kj}\right)\right) = \frac{\exp\left(\tanh\left(\omega_k H_{kj}\right)\right)}{\sum_{k=1}^{K} \sum_{j=1}^{C} \exp\left(\tanh\left(\omega_k H_{kj}\right)\right)} \tag{9}$$

where ω_k represents the trainable parameters of the k-th residual block. Based on this, the model can assign varying weights to different H_{kj}, thereby accentuating features conducive to the classification task while attenuating irrelevant features. Finally, the G calculated by the module serves as the final extracted signal feature for the classification task.

3.1.4 Classification Module

In the classical TCN model, fully connected layers are commonly integrated for extracting global features, which are subsequently fed into a softmax classifier for classification. However, in the context of small target detection on the sea surface, the substantial presence of sea clutter leads to a large number of parameters in the fully connected layers, thus posing a risk of overfitting and increasing computational complexity. Recent studies have indicated that employing substituting fully connected layers with Global Average Pooling (GAP) layers shows promising results in addressing this issue [23]. The primary advantages of the GAP layer lie in its reduced parameter count and enhanced computational efficiency, particularly beneficial for small-scale datasets. Furthermore, they possess a smoothing property, which aids in mitigating overfitting more effectively, thus contributing to the improvement of model generalization capabilities.

Therefore, this paper adopts the GAP layer to extract global information for each sample and perform feature compression. It controls the dimensionality of the output feature vectors from the GAP layer and computes the average of each feature vector, obtaining the average feature values of each sample across different feature dimensions. Finally, the softmax function is employed to transform the average feature values into class confidences, thus fulfilling the classification objective, as depicted in Fig. 7.

Assuming the output of the multi-layer attention module is G, and the result after GAP is denoted as F_{avg}, its computation formula is:

$$F_{\text{avg}} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} G(i, j). \tag{10}$$

In the expression, $F_{\text{avg}} \in R^{C \times 1}$; H and W represent the height and width of the sample, respectively. The result is input into the softmax function for binary classification.

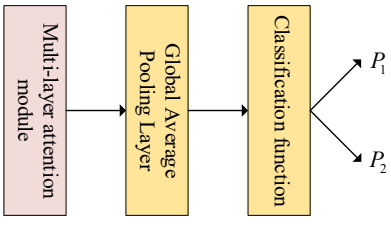


Fig. 7. Classification module.

3.2 Design of a Controllable False Alarm Detector

In the field of small target detection on the sea surface, the sea spike effect is closely related to the occurrence of false alarms. It directly reflects the non-stationary characteristics of sea clutter, which is caused by changes in wind speed and waves. Especially when the wind speed is high, the sea surface surge will become violent, resulting in the generation of broken waves, which will cause strong echoes and form echo signals with larger amplitudes. Due to the prominent manifestation of sea spike effects in radar signals, they are susceptible to detection, consequently triggering false alarms. False alarm rate refers to the ratio of the number of sea clutter misjudged as targets to the total number of sea clutter samples in the dataset, which is an important metric for assessing target detection performance. In the realm of sea-based radar target detection, once a potential target is identified, significant human and material resources are required for further confirmation and identification. Therefore, the occurrence of false alarms may lead to substantial resource wastage. Consequently, this paper proposes a false alarm controllable method for making final classification decisions, achieving target classification under a preset false alarm rate.

The detector proposed in this paper uses a softmax classifier to determine distinct thresholds to control the false alarm rate. According to reference [12], the softmax layer can convert the classification results into classification probabilities, as follows:

$$S\left(z_i\right)=\frac{\mathrm{e}^{z_i}}{\sum_{c=1}^{C}\mathrm{e}^{z_c}} \tag{11}$$

where z_i denotes the output of the i-th node; C represents the number of classes for classification. Therefore, the output of this detector is a two-dimensional vector, with results ranging between 0 and 1, and sums up to 1. The first element represents the probability of the sample being a target signal, while the second element represents the probability of the sample being clutter. Thresholds can be set using the following method to achieve control over varying false alarm rates.

$$T = Y_i\left(1\right), \tag{12}$$

$$i = p_{\mathrm{fa}} \times N_{\mathrm{clutter}}. \tag{13}$$

In the equations, Y represents the set of clutter samples after being sorted in descending order; p_{fa} denotes the false alarm rate to be controlled; and N_{clutter} stands for the number of clutter samples.

3.3 Training and Testing Procedures

The training process of the proposed MA-TCN model mainly consists of two stages: data preprocessing and model training, as depicted in Algorithm 1. During the data preprocessing stage, the raw measured radar signals are processed, and the training set and testing set are partitioned. Subsequently, all parameters of the model are randomly initialized. Following this, the model is iteratively trained using the training set to optimize network parameters. Ultimately, by adjusting the detection threshold to meet the requirements for false alarm rate, the model parameters are saved for use in the testing phase, thus achieving controlled false alarm target detection.

During the testing phase, the testing dataset is fed into the trained model to obtain the maximum detection probability under a controllable false alarm rate. This process utilizes the parameters learned by the MA-TCN model during training to effectively detect new data.

Algorithm. 1. Training process of the MA-TCN model.

Input: The actual radar echo data, randomly initialize the parameter set $\phi = \{\theta_{\mathrm{ten}}, \theta_{\mathrm{attention}}\}$, along with the learning rate lr, where $\{\theta_{\mathrm{ten}}, \theta_{\mathrm{attention}}\}$ represent the parameters of the stacked residual module and the multi-level attention module, respectively.
Output: MA-TCN model (with trained model parameters ϕ).

Perform magnitude transformation and equally spaced partitioning on the data according to formulas (3) and (4), and split it into training and testing sets in a 7:3 ratio.
For epoch = 1 to E (where E denotes the total training batches)
Shuffle the dataset and divide it into K mini-batches.
For batchsize = 1 to K
For n = 1 to N (where N is the number of residual modules)
Calculate the output features of each residual module according to formulas (5)–(7).
End
Calculate the output features through the multi-layer attention module according to formulas (8) and (9).
Calculate the probabilities belonging to each category according to formula (10).
Calculate the loss function and compute the random gradients for each parameter, then perform backpropagation using gradient descent.
End
End
Utilize the false alarm controllable method in Sec. 3.3 to obtain the optimal classifier under the current preset false alarm rate.

Overall, the MA-TCN model achieves efficient detection of sea surface targets through two phases: training and testing. During training, an improved TCN is employed for direct feature extraction from one-dimensional temporal signals, obviating the necessity for manually designed complex signal transformation procedures. Integration of multiple layers of attention modules enables adaptive adjustment of the weights of residual blocks, thereby minimizing information loss. Finally, a controllable false alarm detector is utilized to accomplish target detection under a certain false alarm rate. During testing, the trained model accurately classifies new data samples, providing a reliable solution for the automatic recognition of targets and sea clutter.

4. Experimentation and Results Analysis

4.1 Experimental Environment

The operating environment of this study is Windows 10, utilizing a 64-bit operating system, and based on an x64 architecture processor. The selected graphics card is the NVIDIA GeForce RTX 3080Ti with a memory capacity of 32 GB. Python 3.7 and MATLAB 2019 are employed as the development languages for all experiments. The training parameters of the model are set as follows: 300 training epochs, with a batch size of 64. The loss function employed is cross-entropy, and the optimization algorithm used is Adam.

The code for this paper has been publicly shared on GitHub at https://github.com/XiangYinyzu/MA-TCN.git.

4.2 Introduction to the Dataset

In order to obtain a sufficient number of training samples, this study adopts (4) to segment the radar echo data during the construction process of the marine radar dataset. In the experiments, d and D are set to 10 and 1024, respectively.

4.2.1 The IPIX Dataset

In this paper, an experiment was conducted using the IPIX database collected in 1993 by a team led by Simon Haykin of McMaster University [24]. The database comprises 10 sets of data in staring mode, with the target being a foam plastic spherical block wrapped in metallic wire mesh, possessing a diameter of 1 meter. Each set of IPIX data consists of 14 data units, including 1 target unit, 9 clutter units, and 4 guard units. Each dataset contains data in four polarization modes: HH, HV, VH, and VV, with each polarization mode's echo data consisting of 14 range bins and matrices of 131,072 pulses, among which only one range bin contains the target. Additionally, the IPIX radar operates in the X-band with a working frequency of 9.3 GHz, a pulse repetition frequency of 1000 Hz, a beamwidth of 0.9 degrees, and an antenna gain of 45.7 dB. For ease of reference in the experiment, the 10 datasets are numerically labeled, and specific information regarding file names, wind speed, target units, and affected units is provided in Tab. 1.

4.2.2 SDRDSP

Although the IPIX database has been widely adopted as a benchmark for sea surface small target detection, some of its physical parameters are relatively outdated. In response to the advancements in advanced radar systems, the Naval Aviation University Information Fusion Research Institute has proposed the Sea Detection Radar Data Sharing Program (SDRDSP) to meet the requirements of current radar data [25], [26]. The SDRDSP database is available for research on target detection, sea clutter characteristics, and clutter suppression. The database is collected by an X-band radar with two pulse modes: single frequency mode and linear frequency modulation mode. Each dataset in the SDRDSP database consists of two sets of measurement data, both of which can be utilized for target detection. To further validate the performance of the proposed detector, three sets of measurement data under the staring mode from 2020 were selected for testing, and Table 2 provides detailed descriptions of them.

Number	Filename	Wind speed (km/h)	Wave height (m)	Units under test	Units affected
1	19931107_135603_starea17	9	2.2	9	8,10,11
2	19931108_220902_starea26	9	1.1	7	6,8
3	19931109_191449_starea30	19	0.9	7	6,8
4	19931109_202217_starea31	19	0.9	7	6,8,9
5	19931110_001635_starea40	9	1.0	7	5,6,8
6	19931111_163625_starea54	20	0.7	8	7,9,10
7	19931118_023604_stareaC0000280	10	1.6	8	7,9,10
8	19931118_162155_stareaC0000310	33	0.9	7	6,8,9
9	19931118_162658_stareaC0000311	33	0.9	7	6,8,9
10	19931118_174259_stareaC0000320	25	0.9	7	6,8,9

Tab. 1. An explanation of the 1993 measured IPIX data.

Filename	Sea condition level	Number of pulses	Information description
20210106150614_01 20210106150614_02 20210106150614_03	Level 3-4	$>10^4$	Target: Navigational buoy Distance: 4.84 km
20210106160919_01 20210106160919_02	Level 3-4	$>10^4$	Target: Navigational buoy and ship Distance: 2.778 km and 4.115 km
20210106172511_01 20210106172511_02	Level 3-4	$>10^4$	Target: Two anchored ships Distance: 2.81 km and 4.16 km

Tab. 2. Detailed information of the selected SDRDSP dataset.

4.3 Detector Performance Analysis

To validate the performance of the detector proposed in this paper, it is compared against four other typical and state-of-the-art methods, namely, a detector based on three features, an Isolation Forest-based detector [27], a CNN-based detector, and a SAE-GA-XGBoost-based detector [28].

4.3.1 Efficiency Analysis Detection

Given the significance of real-time performance in actual radar systems, this paper conducts experiments on the time taken by each detector in the detection process under the HH polarization mode of the #54 dataset. The specific results are presented in Tab. 3.

From the table, it can be observed that the previous four detectors require significant time for data preprocessing. The primary reason for this lies in the need to acquire the requisite features for detection, which often involves computationally intensive operations such as frequency transformation and spectral analysis of the raw radar echo signals. Subsequently, constructing feature spaces based on time domain, frequency domain, and time-frequency domain further contributes to the computational load. In comparison, the detector proposed in this paper operates on one-dimensional temporal radar data using an improved TCN without the need for complex signal transformation processes. By introducing dilation factors, the network model is simplified, thereby reducing the overall computational complexity of the algorithm. Consequently, the proposed method demonstrates outstanding real-time performance, requiring only 20 seconds for data preprocessing, and the entire training process takes only 300 seconds. Experimental results demonstrate that, under equivalent conditions, the detector proposed in this paper exhibits a significant improvement in the overall time required for small sea surface target detection compared to other detectors. This indicates a notable enhancement in response speed, highlighting clear advantages in reducing computational complexity and improving real-time performance.

4.3.2 Performance Analysis Detection

In this paper, a certain number of test sets were selected, and the confusion matrix under HH polarization for dataset #30 is shown in Fig. 8. Additionally, confusion matrices under VV, HV, and VH polarizations as well as other datasets were tested but are not presented here due to space limitations. According to the analysis of the confusion matrix shown in Fig. 8, when the preset false alarm rate is 0.001, the model demonstrates excellent performance in recognizing target samples in this dataset, achieving an accuracy of 89.48%. The actual false alarm rate is 0.0018, which shows a relatively small deviation from the preset false alarm rate, indicating effective control over the false alarm rate.

In Fig. 9, comparative results of detection probabilities across various polarization modes for multiple detectors are depicted for 10 datasets in the IPIX database, with a false alarm rate of 0.001. Through a comprehensive comparison across 40 sets of data, it is evident that under these conditions, the proposed detector consistently exhibits the highest detection probability across all polarization modes, as compared to four other detectors. This outcome strongly corroborates the effectiveness of the proposed detector in handling temporal signals.

Furthermore, it can be observed from Fig. 9 that the polarization mode significantly affects the detection results. Taking the #17 dataset as an example, the detection probability is 0.8879 when the polarization mode is HH; for VV polarization, the detection probability is 0.7426; while for HV and VH modes, the detection results are 0.8891 and 0.8696, respectively. It can be seen that the detection results under VV polarization mode are the lowest among all files. On the other hand, under the same polarization mode, there is a significant difference in detection probabilities among different datasets. For instance, under HH polarization, the detection probability of #17 is approximately 19% higher than that of #30. These discrepancies primarily stem from variations in the average signal-to-clutter ratio of the data.

Figure 10 depicts a comparative analysis of detection probabilities from various detectors across four polarization modes for ten datasets within the IPIX database, with a false alarm rate of 0.01. Upon comparison, it was observed that

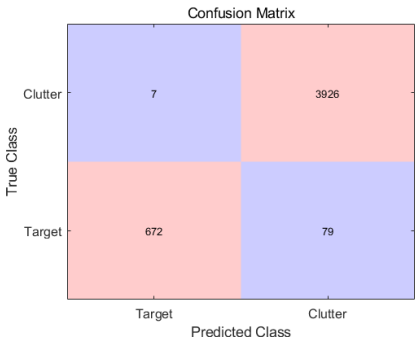


Fig. 8. Confusion matrix for HH polarization under the #30 dataset.

Time \ Detector	Tri-feature-based detector	Isolation Forest-based detector	CNN-based detector	SAE-GA-XGBoost -based detector	Proposed detector
Data preprocessing time	150 min	1740 min	150 min	90 min	20 s
Training time	10 s	200 s	405 s	140 s	300 s

Tab. 3. Detection efficiency of different detectors.

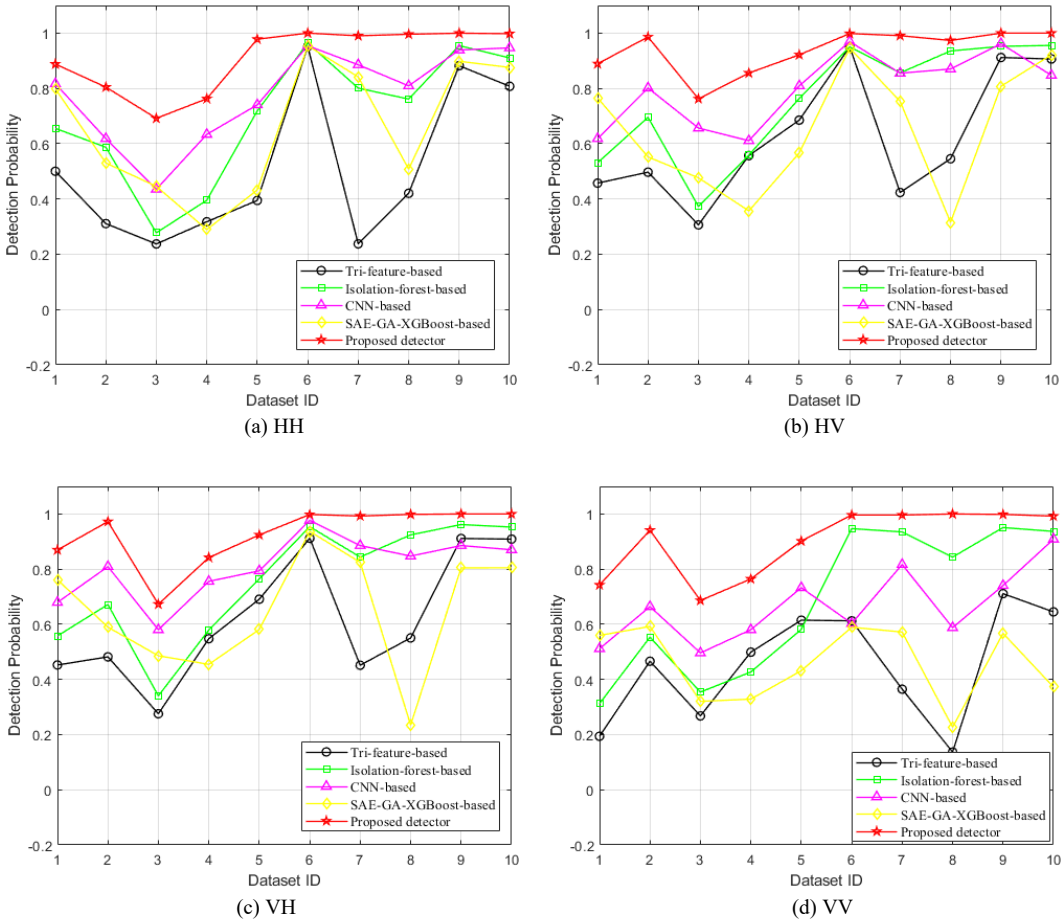


Fig. 9. Comparison of detection probabilities of different detectors at a false alarm rate of 0.001.

among the 40 datasets, the detector proposed in this study exhibits superior detection performance when compared to four other detectors.

For an overall evaluation of the IPIX database performance, Table 4 presents the average detection probabilities and time consumption of various detectors under different false alarm rates across four polarization modes. From the table, it can be observed that the average detection probabilities for the three-feature detector are 0.5668 and 0.5905; for the Isolation Forest detector, they reach 0.7246 and 0.7676; for the CNN-based detector, they stand at 0.7645 and 0.8179; while for the SAE-GA-XGBoost-based detector, they stand at 0.6017 and 0.6625. Notably, the average detection performance of the proposed detectors is 0.9194 and 0.9927, respectively. Compared to the other four detectors, the proposed detector, which directly detects one-dimensional radar temporal signals using an improved TCN and adapts the weights of each residual module through a multi-layer attention mechanism, achieves a detection probability improvement of 15%–40%.

The observations from Tab. 4 further reveal that with the increasing of false alarm rate, the detection probabilities of various detectors demonstrate an upward trend. This phenomenon arises from the adjustment of corresponding classification thresholds during the classification process as a result of increasing the false alarm rate, leading to more samples being identified as targets. In practical applications, detection must be conducted under specific false alarm rate conditions. The experimental results in this section demonstrate that under low false alarm rates, the MA-TCN detector can also achieve high detection probabilities. In summary, among the five detectors discussed, the detector proposed in this paper requires the least amount of time for detecting small targets on the sea surface, thereby not only enhancing detection performance but also improving detection efficiency, which is conducive to engineering implementation.

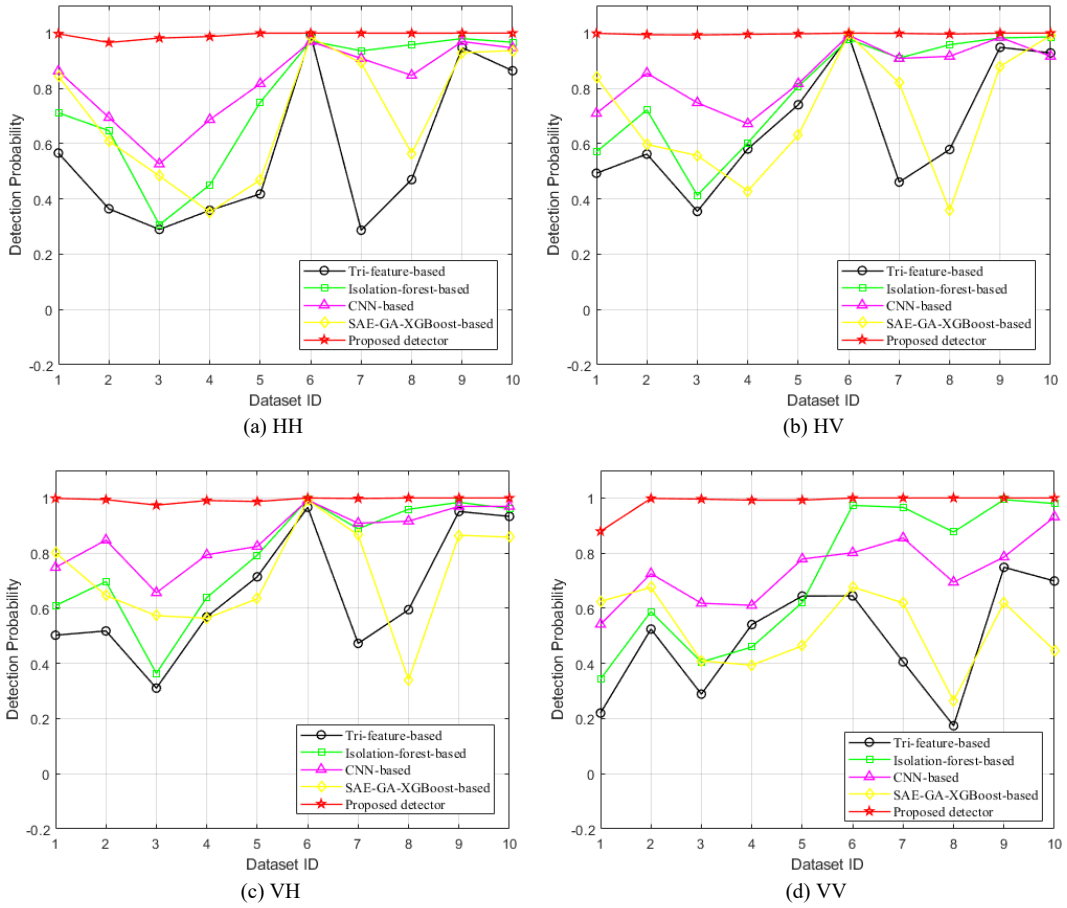


Fig. 10. Comparison of detection probabilities of different detectors at a false alarm rate of 0.01.

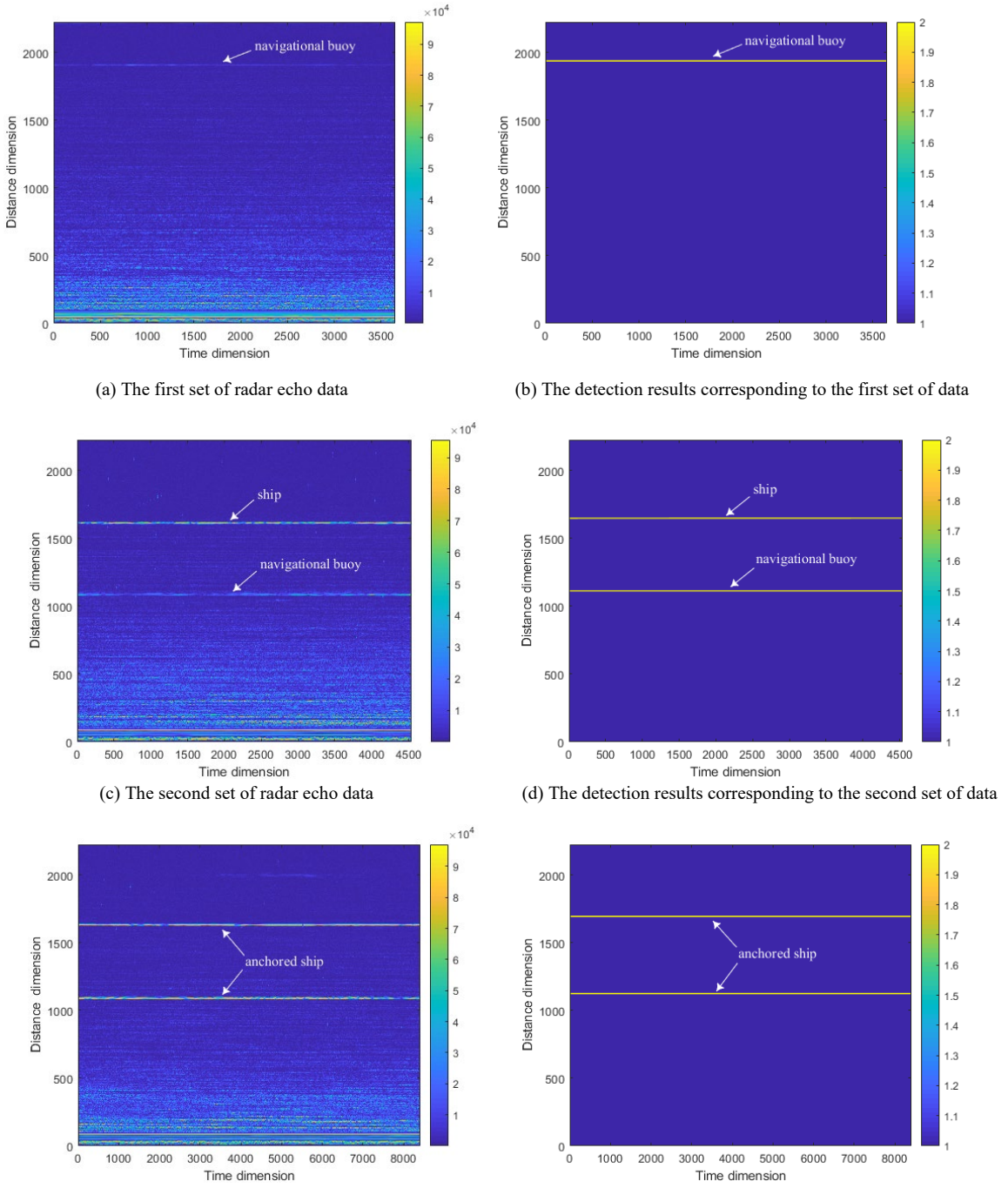
Detector	PF	Average detection probability under different polarization modes				Average detection probability	Average detection time (s)
		HH	HV	VH	VV		
Tri-feature-based detector	0.001	0.5200	0.6309	0.6277	0.4885	0.5668	9010
	0.01	0.5557	0.6639	0.6532	0.4892	0.5905	
Isolation Forest-based detector	0.001	0.7025	0.7571	0.7549	0.6841	0.7246	104600
	0.01	0.7678	0.7930	0.7890	0.7206	0.7676	
CNN-based detector	0.001	0.7779	0.8000	0.8160	0.6641	0.7645	9405
	0.01	0.8229	0.8519	0.8626	0.7344	0.8179	
SAE-GA-XGBoost-based detector	0.001	0.6573	0.6460	0.6478	0.4560	0.6017	5540
	0.01	0.7065	0.7099	0.7140	0.5196	0.6625	
Proposed detector	0.001	0.9108	0.9377	0.9270	0.9019	0.9194	320
	0.01	0.9932	0.9975	0.9943	0.9858	0.9927	

Tab. 4. Average detection performance and detection efficiency of different detectors.

To further validate the effectiveness of the proposed detector, three sets of data from the SDRDSP dataset were utilized for validation analysis. In the first dataset, the target is a navigational buoy located at a distance of 4.84 km from the radar (corresponding to the 1937th and 1938th distance units). In the second dataset, the targets consist of a naviga-

tional buoy and a ship, positioned at distances of 2.778 km (at the 1112th distance unit) and 4.115 km (at the 1647th and 1648th distance units) from the radar, respectively. In the third dataset, the targets are two anchored vessels situated at distances of 2.81 km (at the 1124th and 1125th distance units) and 4.16 km (at the 1694th and 1695th distance units) from the radar, respectively. When the false alarm rate is 0.001, the detection results of the proposed detector on three sets of data are shown in Fig. 11. Figures 11(a) to (f) represent the original radar signal and the detection result for the first, second, and third sets of data, respectively. A comparison reveals that the proposed detector can effectively distinguish between the target and sea clutter, demonstrating the effectiveness of the detector.

Table 5 presents the detection performance of the proposed detector on the three sets of data from the SDRDSP dataset at preset false alarm rates of 0.001 and 0.01. As shown in the table, the difference between the actual false alarm rates and the preset false alarm rates is relatively small across the three sets of data, demonstrating a controlled false alarm rate. Additionally, all three sets of data achieve high detection probabilities.



(a) The first set of radar echo data

(b) The detection results corresponding to the first set of data

(c) The second set of radar echo data

(d) The detection results corresponding to the second set of data

(e) The third set of radar echo data

(f) The detection results corresponding to the third group of data

Fig. 11. When the false alarm rate is 0.001, the detection results of three sets of SDRDSP data.

Data	Prescribed false alarm rate	Actual false alarm rate	Detection probability	Detection time (s)
First group	0.001	0.003	1	244
	0.01	0.01	1	244
Second group	0.001	0.001	1	404
	0.01	0.02	1	404
Third group	0.001	0.001	1	599
	0.01	0.01	1	599

Tab. 5. Performance of detectors on three datasets of SDRDSP dataset.

5. Summary

This paper presents a sea surface small target detection method based on MA-TCN. Addressing the issues of lengthy feature extraction time and low real-time performance of existing detectors, an improved TCN is utilized to directly extract features from one-dimensional temporal signals. Furthermore, a multi-layer attention module is employed to adaptively adjust the weights of residual blocks, aiming to shorten data preprocessing time and minimize information loss. Additionally, different thresholds are determined based on the softmax classifier to control the false alarm rate. Experimental results on IPIX and SDRDSP datasets demonstrate that compared to existing methods, the proposed detector achieves higher detection probability and shorter processing time, showcasing its superiority in sea surface small target detection tasks. This indicates the potential practical application value of MA-TCN in handling radar time-series data in complex environments. Further work needs to obtain or construct more real radar data under different sea states to verify the universality of the proposed method.

References

[1] ZHU, J. N. *Research on Multi-feature Anomaly Detection Technology for Small Targets in Sea Clutter Background*. Shaanxi: Xidian University, 2020.

[2] TEPECIK, C., NAVRUZ, I. Atmospheric refractivity estimation from radar sea clutter using novel hybrid model of genetic algorithm and artificial neural networks. *Radioengineering*, 2020, vol. 29, no. 3, p. 512–520. DOI: 10.13164/re.2020.0512

[3] KELLY, E. J. An adaptive detection algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 1986, vol. AES-22, no. 2, p. 115–127. DOI: 10.1109/TAES.1986.310745

[4] ROBEY, F. C., FUHRMANN, D. R., KELLY, E. J., et al. A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 1992, vol. 28, no. 1, p. 208–216. DOI: 10.1109/7.135446

[5] HAYKIN, S., LI, X. B. Detection of signals in chaos. *Proceedings of the IEEE*, 1995, vol. 83, no. 1, p. 95–122. DOI: 10.1109/5.362751

[6] LO, T., LEUNG, H., LITVA, J., et al. Fractal characterisation of sea-scattered signals and detection of sea-surface targets. *IEE Proceedings F (Radar and Signal Processing)*, 1993, vol. 140, no. 4, p. 243–250. DOI: 10.1049/IP-F-2.1993.0034

[7] SHUI, P. L., LI, D. C., XU, S. W. Tri-feature-based detection of floating small targets in sea clutter. *IEEE Transactions on Aerospace and Electronic Systems,* 2014, vol. 50, no. 2, p. 1416 to 1430. DOI: 10.1109/TAES.2014.120657

[8] SHI, S. N., SHUI, P. L. Sea-surface floating small target detection by one-class classifier in time-frequency feature space. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, vol. 56, no. 11, p. 6395–6411. DOI: 10.1109/TGRS.2018.2838260

[9] XU, Y., JIN, L. Sea-surface floating small target detection based on SVM and multidimensional features. In *2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*. Chongqing (China), 2020, p. 453–457. DOI: 10.1109/ICIBA50161.2020.9276826

[10] GUO, Z. X., SHUI, P. L. Anomaly based sea-surface small target detection using K-nearest neighbor classification. *IEEE Transactions on Aerospace and Electronic Systems,* 2020, vol. 56, no. 6, p. 4947–4964. DOI: 10.1109/TAES.2020.3011868

[11] SU, N., CHEN, X., GUAN, J., et al. Detection and classification of marine target with micro-motion based on CNNs (in Chinese). *Journal of Radars*, 2018, vol. 7, no. 5, p. 565–574. DOI: 10.12000/JR18077

[12] SU, N. Y., CHEN, X. L., GUAN, J. One-dimensional sequence signal detection method for marine target based on deep learning (in Chinese). *Journal of Signal Processing*, 2020, vol. 36, no. 12, p. 1987–1997. DOI: 10.16798/j.issn.1003-0530.2020.12.004

[13] SU, N. Y., CHEN, X. L., CHEN, B. X., et al. Dual-channel convolutional neural networks feature fusion method for radar maritime target intelligent detection (in Chinese). *Modern Radar*, 2019, vol. 41, no. 10, p. 47–52. ISSN: 1004-7859

[14] WANG, J., LI, S. Maritime radar target detection in sea clutter based on CNN with dual-perspective attention. *IEEE Geoscience and Remote Sensing Letters*, 2023, vol. 20, p. 1–5. DOI: 10.1109/LGRS.2022.3230443

[15] WAN, H., TIAN, X., LIANG, J., et al. Sequence-feature detection of small targets in sea clutter based on Bi-LSTM. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, vol. 60, p. 1–11. DOI: 10.1109/TGRS.2022.3198124

[16] LIU, T., ZHANG, L., ZENG, Z. G., et al. Study on the composite electro-magnetic scattering from 3D conductor multi-objects above the rough surface. *Radioengineering*, 2021, vol. 30, no. 2, p. 323 to 334. DOI: 10.13164/RE.2021.0323

[17] QU, Q., LIU, W., WANG, J., et al. Enhanced CNN-based small target detection in sea clutter with controllable false alarm. *IEEE Sensors Journal*, 2023, vol. 23, no. 9, p. 10193–10205. DOI: 10.1109/JSEN.2023.3259953

[18] SHI, Y., CHEN, W. Sea surface target detection using global false alarm controllable adaptive boosting base on correlation features. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, vol. 61, p. 1–14. DOI: 10.1109/TGRS.2023.3265480

[19] SHI, S. N., YANG, J. Detection of sea-surface small target based on multi-domain and multi-dimensional feature fusion (in Chinese). *Journal of Signal Processing*, 2020, vol. 36, no. 12, p. 2099–2106. DOI: 10.16798/j.issn.1003-0530.2020.12.016

[20] LEA, C., FLYNN, M. D., VIDAL, R., et al. Temporal convolutional networks for action segmentation and detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, (HI, USA), 2017, p. 1003–1012. DOI: 10.1109/CVPR.2017.113

[21] LIU, X. *Research and Application of Time Series Data Feature Extraction Based on TCN.* Liaoning: Dalian University of Technology, 2021.

[22] KHATUN, M. A. Deep CNN-LSTM with self-attention model for human activity recognition using wearable sensor. *IEEE Journal of Translational Engineering in Health and Medicine*, 2022, vol. 10, p. 1–16. DOI: 10.1109/JTEHM.2022.3177710

[23] PANG, Y., SUN, M., JIANG, X., et al. Convolution in convolution for network in network. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, vol. 29, no. 5, p. 1587–1597. DOI: 10.1109/TNNLS.2017.2676130

[24] HAYKIN, S. *The McMaster IPIX Radar Sea Clutter Database in 1993*. [Online] Available at: http://soma.ece.mcmaster.ca/ipix/

[25] LIU, N. B., DING, H., HUANG, Y., et al. Annual progress of the sea-detecting X-band radar and data acquisition program (in Chinese). *Journal of Radars*, 2021, vol. 10, no. 1, p. 173–182. DOI: 10.12000/JR21011

[26] LIU, N. B., DONG, Y. L., WANG, G. Q., et al. Sea-detecting X-band radar and data acquisition program (in Chinese). *Journal of Radars*, 2019, vol. 8, no. 5, p. 656–667. DOI: 10.12000/JR19089

[27] XU, S., ZHU, J., JIANG, J., et al. Sea-surface floating small target detection by multi-feature detector based on isolation forest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, vol. 14, p. 704–715. DOI: 10.1109/JSTARS.2020.3033063

[28] ZHAO, D., HANG, H. Y., WANG, H. F., et al. Sea-surface small target detection based on SAE-GA-XGBoost algorithm (in Chinese). *Radar Science and Technology*, 2023, vol. 21, no. 1, p. 88–96. DOI: 10.3969/j.issn.1672-2337.2023.01.011

About the Authors …

Xiang YIN received his Ph.D. degree in Signal and Communication Engineering from Hefei University of Technology, China, in 2008. He is currently working as an Associate Professor at the School of Information Engineering, Yangzhou University, Yangzhou, China. His research interests include intelligent signal processing and radar target detection.

Wanhua LI is a M.S. student at the School of Information Engineering, Yangzhou University, Yangzhou, China. Her research interests include radar signal processing and machine learning.

Liulin WANG is a M.S. student at the School of Information Engineering, Yangzhou University, Yangzhou, China. Her research interests include radar signal processing and target detection.

Yu ZHAO (corresponding author) received his Ph.D. degree in Information and Communication Engineering from Chungbuk National University in 2018. In 2019, he joined the School of Information Engineering at Yangzhou University, where he currently serves as an Associate Professor. His key research areas include digital signal processing and deep learning.