An Incept-TextCNN Model for Ship Target Detection in SAR Range-Compressed Domain

HongCheng Zeng, *member, IEEE*, YuTong Song, Wei Yang, *member, IEEE*, Tian Miao, Wei Liu
WeiJie Wang and Jie Chen, *Senior Member, IEEE*

***Abstract*—Traditionally, SAR-based ship target detection is performed in the image domain, where SAR imaging processing has to be applied first. However, SAR imaging processing is complex and time-consuming, especially in the wide-swath working mode. Actually, for open sea scenes, most echoes are sea surface signals with no ship targets, and there is no need for imaging processing in those areas. Therefore, non-image domain ship target detection is studied in this paper, and a novel Incept-text convolutional neural network (TextCNN) model is proposed for ship target detection in the SAR range-compressed domain (RCD). In the proposed method, the SAR echo data is converted into a one-dimensional range profile signal firstly by range compression and mean pooling, and then, the Incept-TextCNN model is proposed and applied, and information about existence of ship targets in relevant range cells will be its output. Finally, the effectiveness and efficiency of the proposed method is testified by simulation and real spaceborne SAR data, and the results demonstrate that the proposed model can filter out the invalid range-compressed data of the sea surface area, which can significantly reduce the amount of data for subsequent SAR imaging and ship classification.**

***Index Terms*—SAR range-compressed domain, Data Filtering, Text convolutional neural network (TextCNN), Ship Target.**

I. INTRODUCTION

SHIP detection is an important application of Synthetic Aperture Radar (SAR). As SAR works in the microwave band, compared with optical sensors, SAR can obtain high-quality remote sensing images under complex weather conditions, and realize all-day, all-weather, and high-resolution wide swath imaging tasks. As a result, SAR has become an important tool for ocean imaging and monitoring.

Traditionally, widely used ship detection methods are based on the constant false alarm rate (CFAR) in SAR [1] . However, there are clear limitations for CFAR-based detectors: their detection ability is affected by surrounding buildings and ports in nearshore scenes; its characteristic pixel-by-pixel detection process leads to low processing efficiency. Recently, deep learning based techniques have been applied to SAR ship detection. Faster R-CNN model is combined with the CFAR detector in [2], while SAR image target detection based on SSD model is performed in [3]. In [4], a dense connection module is introduced in YOLOv3 to detect small targets. In [5], the large-size detection process is optimized, where the slices with potential targets are screened firstly, followed by further refined detection. Based on this, a new method is proposed in [6] to reduce the involved calculations through context information. These deep learning based algorithms significantly improve the detection speed and accuracy; however, those methods are performed in SAR image domain, which have two major drawbacks: (1) SAR image processing is time consuming, especially azimuth focusing; (2) for sea surface scenarios, most areas have no targets, and a lot of computing resources are wasted in these regions.

To overcome these shortcomings, some target detection methods in the non-image domain have been presented in recently years and one representative example is the range-compressed domain (RCD) as it does not require time-consuming azimuth focusing. In [7], a ship detector is proposed based on Faster R-CNN working in the RCD; in [8], a two-step detection method is presented with the first step using complex signal kurtosis in the RCD to screen possible ship areas coarsely, and the second step applying CNN to further detect the potential ship areas; an oriented ship detection strategy is designed in [9], which calculates the constant false alarm rate detection threshold in the range-Doppler domain; a supportive ship tracking concept is introduced in [10] in the range-Doppler domain using an airborne-based radar sensor; in addition, a method for ship detection from raw SAR echo data is proposed in [11]. However, most of them are based on two-dimensional data for detection, where the model size tends to be large and dependent on CPU resources, and detection based on one-dimensional data often doesn't take advantage of deep learning that can extract deep features.

In this paper, a novel Incept-TextCNN model is presented to detect ship targets. In the proposed method, the SAR echo signal is converted into the one-dimensional range profile firstly, and then the TextCNN model extracts the depth features of the amplitude information in the data and screens the range gates containing ship targets. As a result, the area of interest can be located fast, and the large non-target areas can be filtering

This work was supported by the Beijing Natural Science Foundation under Grant 4222006. (Corresponding author: *Wei Yang*.)

HongCheng Zeng, YuTong Song, Wei Yang and Jie Chen are with the School of Electronic and Information Engineering, Beihang University, Beijing 100191, China. (e-mail: zenghongcheng@buaa.edu.cn; songyutong@buaa.edu.cn;yangweigigi@sina.com; chenjie@buaa.edu.cn).

Tian Miao is with Key Laboratory of Network Information System Technology (NIST), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China (e-mail: miaotian@aircas.ac.cn)

Wei Liu is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, UK (e-mail: w.liu@qmul.ac.uk)

WeiJie Wang is with the Shanghai Aerospace Electronic Technology Institute, Shanghai 201108, China (e-mail: flyweijie@hotmail.com).

out effectively, and the consumption of subsequent imaging and detection of non-target areas can be reduced.

The remainder of this letter is organized as follows. Construction of the RCD data set is presented in Section II, and the Incept-TextCNN model is introduced in Section III for coarse detection of the ship target area. Experimental results are provided in Section IV, and conclusions are drawn in Section V.

II. The RCD Ship Target Data Set

To apply the deep learning based method, preparing the RCD ship target data set is an important step in the proposed method. The RCD data is the intermediate product of SAR imaging. In this paper, the ship target data set in RCD is constructed first, and corresponding one-dimensional range profile is also provided, which is obtained from range-compressed data and contains information about signal amplitude changes along the range direction.

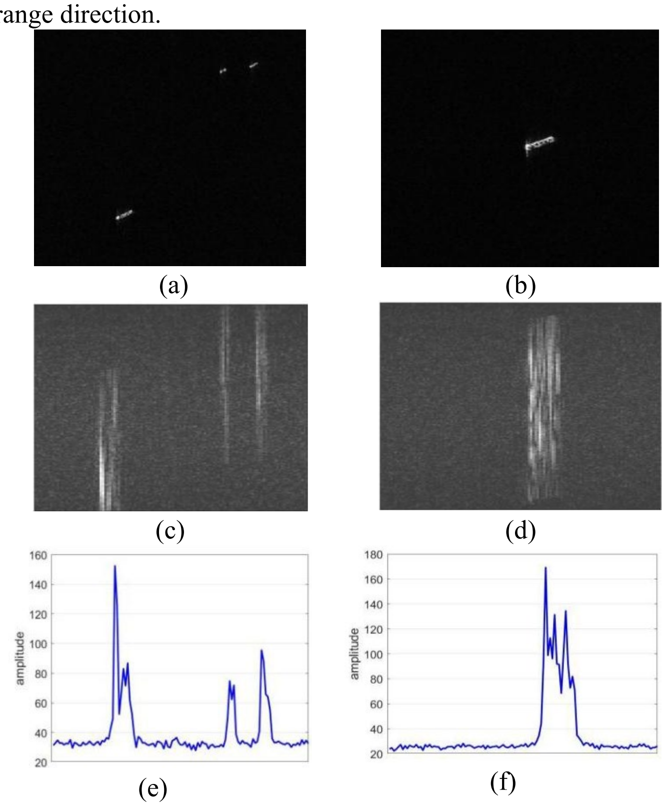

Fig.1. The simulated ship target data: (a) data A in SAR image; (b) data B in SAR image; (c) RCD data of A; (d) RCD data of B; (e) one-dimensional range profile of A; (f) one-dimensional range profile of B.

A. Simulated Ship Target data in RCD

Spaceborne SAR simulation is used to obtain sufficient range-compressed ship target data. As it is difficult to obtain the SAR echo data, the ship target in real SAR images is used to simulate the echo data which is then processed and transfomred into the range-compressed data. For the RCD data simulation, the target SAR image is used as an input, and the simulated echo data is obtained through parameter setting, SAR system simulation, random phase simulation and echo simulation. Then, range Fast Fourier Transform (FFT), range matched filtering and range inverse FFT (IFFT) are performed to obtain the range-compressed data. Furthermore, a mean pooling operation is carried out along the azimuth-direction to obtain the corresponding one-dimensional range profile signal, which is the input of the subsequent training model. Based on the presented simulation method, the simulated RCD ship target data and its corresponding one-dimensional range profile are presented in Fig.1, where the horizontal direction represents the range. Here, the range resolution is about 1m. As shown, the fluctuation of signal amplitude can be clearly observed, and this characteristic will be useful in the subsequent ship target detection.

B. Real Ship Target data in RCD

For real SAR data, only range FFT, range matched filtering and range IFFT operations are needed. Using the Pujiang-2 spaceborne SAR data, Fig. 2 presents the real data and corresponding one-dimensional range profile, including the good and bad sea conditions. Finally, based on the simulated and real ship target data, a one-dimensional range profile data set of ship target in RCD is generated, which contains 433 sets of training sample data, 110 sets of verification sample data, 252 sets of test sample data, with a positive and negative sample ratio of 1:1 and a resolution of about 1-3m. The overall ratio of simulated and real data in training and verification samples is 6:4.

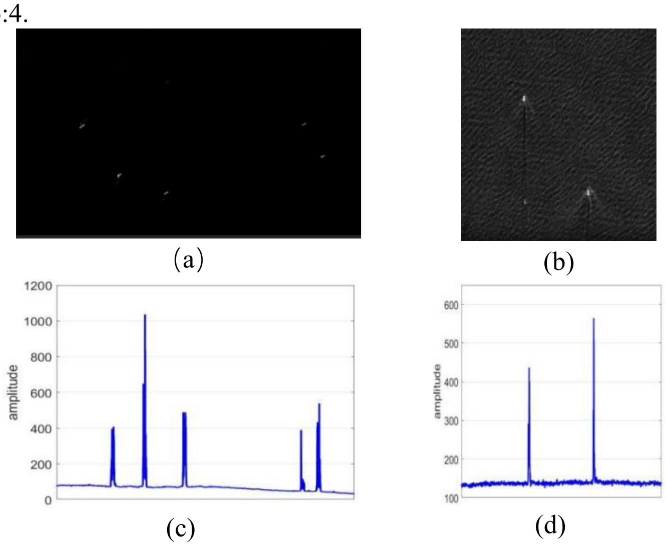

Fig.2. The real ship target data: (a) real data A in SAR image; (b) real data B in SAR image; (c) one-dimensional range profile of A; (d) one-dimensional range profile of B.

III. Ship Target Detection based on Incept-TextCNN

To filter out the non-target area data, a novel Incept-TextCNN model for ship target detection is proposed in this part. The general idea is shown in Fig. 3, where the one-dimensional range profile data is obtained from SAR echo first, and then TextCNN outputs which range gates contain targets and which range gates do not according to the different characteristics of amplitude in the background region and the region containing the ship targets in the data, thus giving the range in the range direction that requires further imaging processing, avoiding the need to image the entire data in the image domain refinement detection.

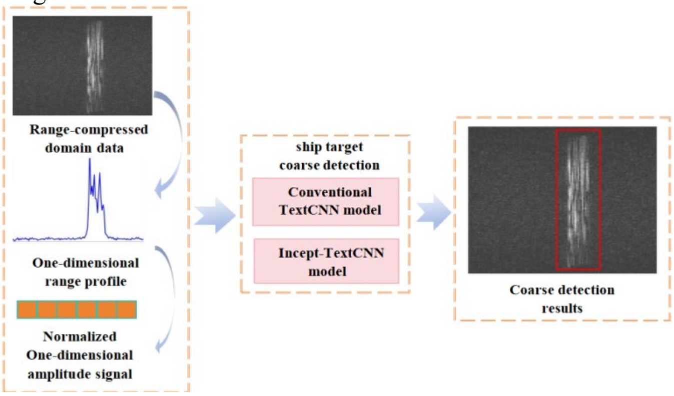

Fig. 3. The general idea for the proposed ship detection model based on the RCD data.

A. *Conventional TextCNN Model*

In TextCNN, output feature sequence of one-dimensional convolution has two dimensions: length and depth. The length depends on the size of convolution kernel, and the information of length dimension is similar to the information of each channel in image domain feature map. The depth is the number of channels, depends on the number of convolution kernels, and is similar to the number of channels in image domain feature map [12]. The relationship between the length of output feature sequence and the convolution kernel is shown in Fig. 4, and the calculation expression is given below:

$$o=\frac{i+2p-k}{s}+1 \tag{1}$$

where, o is the length of the output feature sequence, i is the length of the input feature sequence, p is the length of all zero pixels extended around the feature matrix during convolution, k is the size of the convolution kernel, and s is the step size, i.e., the span of each movement of the convolution kernel.

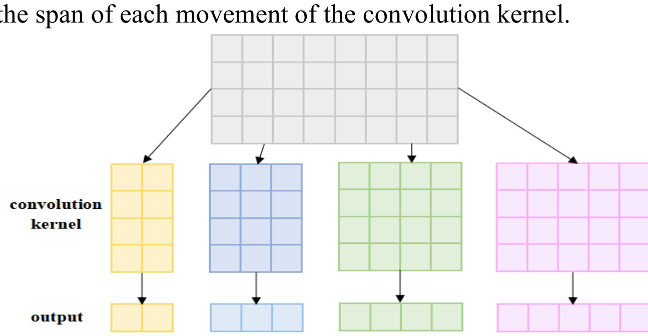

Fig. 4. Description of the one-dimensional convolution operation.

The structure of the conventional text convolutional neural network (TextCNN) model for the RCD is shown in Fig. 5. After the one-dimensional RCD data is fed into the model, the feature is extracted by five convolutional blocks successively. The length of the feature sequence gradually decreases and the dimension gradually increases. Each convolutional block adopts the form of one-dimensional convolution plus batch normalization plus activation function, and carries out feature extraction, normalization, and nonlinear assignment processing to improve the characterization ability of the model. After the fifth convolutional block outputs the feature sequence, the sequence is transformed into a one-dimensional sequence through the "flattening" operation. After that, the sequence dimension is reduced through three fully connected layers, normalized and activated through the batch normalization (BN) layer and ReLU activation function, and finally, the confidence of each item is obtained through the Softmax function.

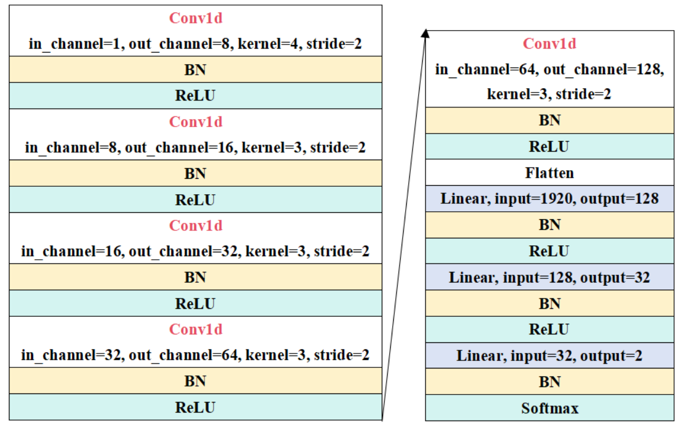

Fig. 5. Structure of the conventional TextCNN model.

B. *Proposed Incept-TextCNN Model*

Based on the structure of the conventional TextCNN model, by analogy with GoogLeNet's Inception [13] module in the image domain, the concatenation operation of different convolutional feature sequences is introduced and in this way, TextCNN based on the Inception module (Incept-TextCNN) for RCD is constructed.

The Inception structure is shown in Fig. 6, where Fig. 6(a) is the original version. The feature extraction from the input feature map is carried out through different convolution kernels. The convolution kernel of multiple sizes is used respectively to carry out convolution operations on the same input feature map so that the feature map of different scales can be obtained. After that, all feature maps are spliced to obtain the output feature map containing information of different scales. Fig. 6(b) shows an improved version of inception. On the basis of the original structure, 1×1 convolution is used to reduce the number of channels in the feature map, thus reducing the accumulation of parameters.

As shown in Fig. 7, the Incept-TextCNN model is based on the structure of TextCNN. When the one-dimensional RCD signal is fed into the model, first of all, a one-dimensional convolution operation is performed by three convolution kernels of different sizes in block1, block2, and block3, and the three output feature sequences are kept the same length by adding 0s and adjusting the step size. Then, in order to avoid a too large model size caused by too many feature sequence channels after concatenation, referring to the operation of the Inception module in GoogLeNet, the convolution kernel of size 1 is used to reduce the depth of feature sequence. Finally, output sequences representing three different scale features are concatenated along the depth dimension. The output of the first

Inception module is obtained and sent to block4. The feature sequence output by block4 is then sent to the next Inception module, and features are further extracted by three different convolution kernels. Then convolution, flattening, full connection layer processing, and activation function processing are carried out successively. Finally, the confidence of each item is obtained through the Softmax function.

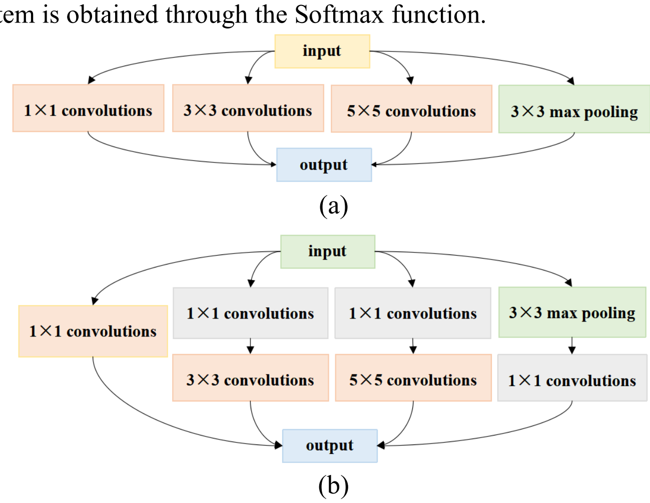

Fig. 6. Inception module diagram: (a) original version; (b) improved version.

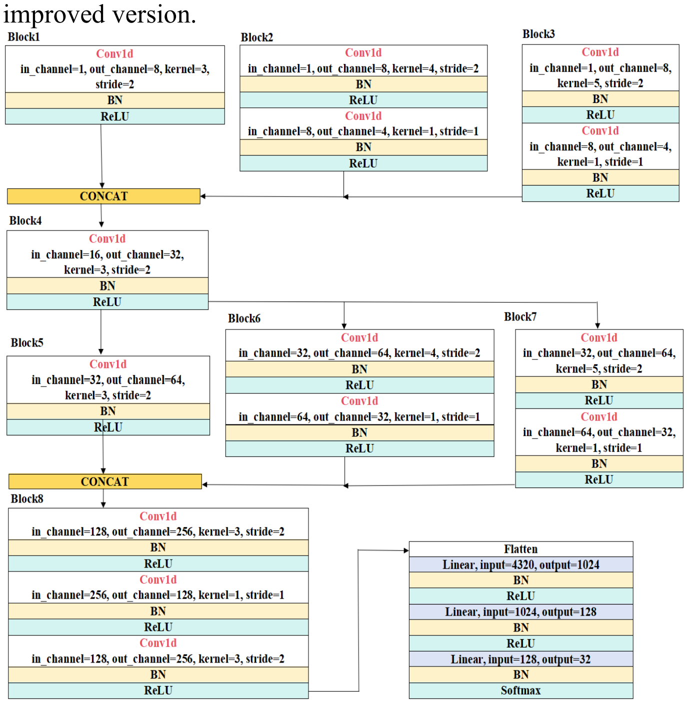

Fig. 7. Structure of the proposed Incept-TextCNN model .

IV. Experimental Results

In this section, the experimental results are presented and discussed using the RCD data set described in Section II. TextCNN and Incept-TextCNN are tested and the detection performance of the two models are compared. Moreover, the Incept-TextCNN model is used to test two real large scene images, and it is demonstrated that the proposed model can detect potential targets present in the range gate effectively.

A. Specific experimental setup

The models are trained on the constructed simulated and real RCD data set. The training parameters of the coarse detection model based on TextCNN and Incept-TextCNN are set as follows: the initial learning rate is set to 0.0001, the batch size to 16, and the number of training rounds is 200. The AdaMax optimizer is used for training, in which β_1 is set to 0.9 and β_2 to 0.999.

B. Comparison between TextCNN and Incept-TextCNN models

Simulated and real data sets are used to train TextCNN and Incept-TextCNN models respectively, where the input data size of each group is 1×512. The training efficiency results of the verification set are shown in Fig. 8, where blue represents the TextCNN model and red represents the Incept-TextCNN model. Both models can realize ship target detection after a period of iterative training. The detection rate by the Incept-TextCNN model on the verification set improves faster, and its accuracy stabilizes above 90% after only a few rounds of training; while the convergence speed of the TextCNN model is slower, it reaches the steady state after about 60 rounds of training. In the verification set, the accuracy of the Incept-TextCNN model can reach 97%, while the accuracy of the TextCNN model is relatively poor, and the average detection rate of the former after stabilization is 2.4% higher than that of the latter, where the accuracy is calculated by the ratio of correctly judged samples to the total samples. Tests on the sliced test data set resulted in an accuracy of 93.2%, a recall [3] of 88.4%, and a F1-score [3] of 90.7% based on TextCNN, and an accuracy of 94.5%, a recall of 90.7%, and a F1-score of 92.6% based on Incept-TextCNN. It can be seen that the detection accuracy, recall and F1-score of Incept-TextCNN model are all higher than TextCNN.

In summary, Incept-TextCNN outperforms TextCNN in detecting, and its model size is 16.58M, although it is larger compared to TextCNN's 1.10M, but because it is a method that use one-dimensional convolution for one-dimensional data processing, it is lighter than the model of two-dimensional convolution for two-dimensional data processing used in image and non-image domains commonly, and has an advantage in model size and detection efficiency.

C. Real data coarse detection based on Incept-TextCNN model

The validity of the Incept-TextCNN model is further verified by using real large image SAR data. For the input of large image, the whole one-dimensional range profile signal is obtained first, and then the long one-dimensional data is divided into many groups of 1×512 data by the method of overlapping sliding window, which is input in sequence for model detection, and finally the groups of data with potential targets are determined, that is, the range gates containing potential targets are the corresponding output. Fig. 9 shows the coarse detection results of SAR images of the Taiwan Strait taken by the Pujiang-2 satellite, with the stride of 200 range gates when sliding the window on data A and 20 range gates on data B.

As can be seen from Fig. 9, the Incept-TextCNN model has successfully detected the positions in range direction of five ship targets in image A and two in image B , and 73.09% non-target sea area in image A and 70.34% non-target sea area in image B are excluded, demonstrating the effectiveness of using one dimensional convolutional network for deep feature extraction on the one dimensional RCD data which only contains amplitude features but lacks ship outline and size features. Note that the vertical length of the red box in Fig. 9(a)(b) is consistent and covers all points in the azimuth-direction, meaning that the exact position of target in the azimuth-direction cannot be given.

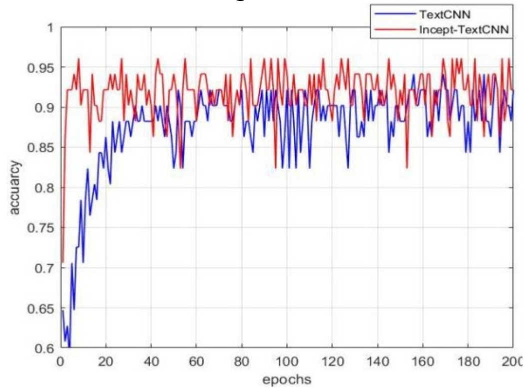

Fig. 8. Accuracy curves of the TextCNN model and the Incept-TextCNN model.

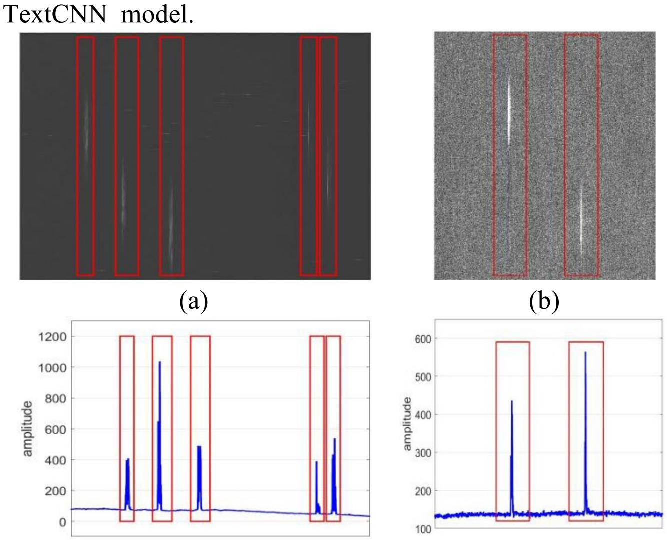

Fig. 9. Coarse detection results using real data: (a) detection results using real data A displayed in RCD; (b) detection results using real data B displayed in RCD; (c) detection results using real data A displayed on one-dimensional range profile signal; (d) detection results using real data B displayed on one-dimensional range profile signal.

V. Conclusion

In this paper, a novel Incept-TextCNN model has been proposed for ship target detection in SAR RCD. It employs one-dimensional convolution to extract the features of one-dimensional RCD data, and then uses activation functions, full connection layers, and other structures to classify the target features. And the Inception structure in GoogLeNet is introduced to fuse the feature information of different scales together, which can improve the detection accuracy.

A key feature of the proposed method is to output the range gates containing potential ship targets (including strong clutters, artificial platforms, or islands areas) to achieve coarse detection of ship targets with high detection rate and high false alarm rate. After the detection by the proposed method, only the selected areas should be imaged and input into the subsequent image domain fine detection model. Thus, the advantages of time saving of the overall framework are mainly reflected in: (1) the coarse detection is for one-dimension, which is faster than the detection speed of two-dimension; (2) there is no need to image all the echo data, but only the part containing targets; (3) since a large number of non-target areas are not imaged, especially for open sea, there is no need to carry out image domain fine detection on non-target areas. Therefore, the proposed model will be very useful in improving the processing efficiency of ship target detection. In the future, we will study nearshore ship targets and the scene with more clutter, improve the existing model to make it suitable for more complex scenes, and explore the performance of using other models such as NPL for detection.

References

[1] Tao X, Mingxing L, Mingjiang Z, et al. Ship detection based on a superpixel-level CFAR detector for SAR imagery[J]. International Journal of Remote Sensing,2022,43(9).

[2] Kang M, Leng X, Lin Z, et al. A modified faster R-CNN based on CFAR algorithm for SAR ship detection[C]//2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP). IEEE, 2017: 1-4.

[3] Wang Z, Du L, Mao J, et al. SAR target detection based on SSD with data augmentation and transfer learning[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 16(1): 150-154.

[4] Wang Z. Research on intelligent detection algorithm of SAR ship target in complex environment[J]. Beijing University of Aeronautics and Astronautics, 2021.

[5] Xiaoya F, Zhaocheng W. SAR Ship Target Rapid Detection Method Combined with Scene Classification in the Inshore Region [J]. Journal of Signal Processing, 2020, 36(12): 2123-2130. DOI: 10. 16798 /j. issn. 1003-0530. 2020. 12. 019.

[6] Zhai L, Li Y, Su Y. Inshore Ship Detection via Saliency and Context Information in High-Resolution SAR Images[J]. IEEE Geoence and Remote Sensing Letters, 2016:1870-1874.

[7] Loran T, A. Barros Cardoso da Silva, Joshi S. K, Baumgartner S. V, and Krieger G. Ship Detection Based on Faster R-CNN Using Range-Compressed Airborne Radar Data[J], IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 3500205, doi: 10.1109/LGRS.2022.3229141.

[8] Leng X, Wang J, Ji K, Kuang G. Ship Detection in Range-Compressed SAR Data[C], IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022, pp. 2135-2138, doi: 10.1109/IGARSS46834.2022.9884909.

[9] Joshi S. K, Baumgartner S. V, Silva D C B A , et al. Range-Doppler Based CFAR Ship Detection with Automatic Training Data Selection[J].Remote Sensing,2019,11(11):1270.

[10]Joshi S. K, Baumgartner S. V, and Krieger G. Tracking and Track Management of Extended Targets in Range-Doppler Using Range-Compressed Airborne Radar Data[J], IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-20, 2022, Art no. 5102720, doi: 10.1109/TGRS.2021.3084862.

[11]Leng X, Ji K and Kuang G, Ship Detection From Raw SAR Echo Data[J], IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-11, 2023, Art no. 5207811, doi: 10.1109/TGRS.2023.3271905.

[12]Kim Y .Convolutional Neural Networks for Sentence Classification.[J]. CoRR,2014,abs/1408.5882.

[13]Szegedy C,0015 L W,Jia Y, et al. Going Deeper with Convolutions.[J]. CoRR,2014,abs/1409.4842.