WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 5, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v5n3.14
Journal homepage: https://wepub.org/index.php/IJCSIT/index

# Literature Review on Rotating Target Detection in SAR Images Based on Deep Learning in Complex Background

Jing Wang, Bin Wu *, Hongying Zhang

School of Information and Control Engineering, Southwest University of Science and Technology, Mianyang, China

*Corresponding Author: Jing Wang

## ABSTRACT

The detection of targets in Synthetic Aperture Radar (SAR) images holds paramount significance for both military and civilian applications, furnishing critical informational support for target surveillance. Deep learning, owing to its robust learning capabilities, has emerged as a favorable solution for addressing the intricacies of SAR image rotating target detection in complex backgrounds. However, numerous challenges persist. The mitigation of speckles in SAR images and the detection of rotating targets amidst dense and small objects within intricate backgrounds constitute pivotal and challenging issues in this domain. This paper systematically reviews algorithms pertaining to these challenges, contrasts the performance structures, advantages, and drawbacks of different algorithms, and ultimately delves into prospective research directions.

## KEYWORDS

Rotated object detection; Complex background; Synthetic Aperture Radar; Deep learning

## 1. INTRODUCTION

With the rapid development of synthetic aperture radar (SAR) technology, SAR images are increasingly used in the field of target detection, and its rotating target detection is particularly important in both military and civilian fields [1]. The presence of complex backgrounds makes the target detection task more complex and challenging. In this context, traditional target detection methods often face the problem of performance degradation. To solve this problem, researchers have turned to deep learning methods, because deep learning has shown excellent performance in processing complex scenes with its powerful learning ability. Deep learning methods can better adapt to the problem of target detection in SAR images by learning features from a large amount of data, especially when the target is rotated. By analyzing the application of various deep learning methods in this field in detail, we can understand their advantages and disadvantages. For the problem of coherent speckles in SAR images and the rotation detection of multiple and small targets in complex backgrounds, this will improve the accuracy and robustness of target detection, which has important theoretical and practical significance.

## 2. RESEARCH BACKGROUND AND SIGNIFICANCET

As an active microwave sensor, Synthetic Aperture Radar (SAR) has the characteristics of realizing all-day and all-weather ground observation without being restricted by light and climate conditions. It can even obtain information covered by the surface or vegetation, and can use satellite-borne, airborne (including drone-borne), missile-borne and other platforms to image high-resolution SAR images of a large range of observation areas at a long distance. Because of these characteristics, SAR is widely used in civil fields such as agriculture, surveying and mapping, water areas, geography, and disaster monitoring, and has very good development prospects. In the military field, it has even more unique advantages. It is the extension of the future battlefield space from the traditional sea, land, and air to the space field. As a reconnaissance means with unique advantages, synthetic aperture radar satellites have a decisive impact on the victory of war in order to seize the information control of the future battlefield. The detection of targets in SAR images can be used for battlefield reconnaissance, precision strikes, military deployment and other applications. Therefore, in recent years, using existing target detection technology to detect targets in SAR images has attracted widespread attention from scholars.

In addition, due to the radiation characteristics of SAR images, the coherent spots in SAR images are image noise caused by the coherence of the ground objects, which will reduce the resolution of SAR images and the ability to detect ground objects. In addition, the background of the target in the SAR images collected by SAR in real environments is usually changeable. The background may be a single and easy-to-distinguish background such as sand, lawn, ocean, etc. in large scenes, or it may be a complex and difficult-to-distinguish background such as urban suburbs, ordinary jungles, coastal ports, etc. in large scenes. Especially when the targets are numerous, dense and rotated in any direction, the general horizontal bounding box cannot accurately depict the target, and it is easy to miss or misdetect. The traditional SAR image detection algorithm requires human participation in the design of target features, is easily interfered by noise, and has a complex calculation amount, which is not conducive to the real-time requirements in real tasks. In recent years, due to the rapid development of computer hardware, computing resources have been greatly improved, and target detection algorithms based on deep learning have gradually emerged and have been widely used. At the same time, a large number of "plug-and-play" model algorithm plug-ins designed to solve various problems have emerged. However, due to the differences between natural images and SAR images, there are still some problems to be solved. Therefore, suppressing coherent speckle is an important part of improving the quality and application effect of SAR images. Reducing missed detection and false detection of targets in different backgrounds, especially in complex backgrounds, can improve the reliability of information acquisition. If you want to apply SAR image detection to real tasks, you must optimize the detection algorithm model. Therefore, it is meaningful to study the SAR image rotation target detection technology based on deep learning in complex backgrounds.

# 3. THE STATUS OF RESEARCH DOMESTICALLY AND INTERNATIONALLY

With the continuous development of SAR technology and its in-depth application in various fields, a variety of traditional SAR image target detection technologies have been proposed. However, the requirements for the recognition rate and detection speed of targets in SAR images are constantly increasing, and the target detection technology for SAR images is constantly improving. In particular, with the continuous rise of artificial intelligence in recent years, target detection technology for SAR images based on deep learning has become increasingly popular among scholars. The following summarizes the existing methods from three aspects: (1) coherent speckle suppression of SAR images; (2) SAR image target detection; and (3) rotation target detection.

## 3.1. Speckle Suppression in SAR Images

Due to the coherent imaging principle of SAR, the collected SAR images will have inherent coherent speckle noise, which will cause edge blur and small-scale targets to be lost after multiple convolutions. Therefore, the suppression of coherent speckle should focus on edge, point and line protection, especially the protection of image texture features, feature details and key points, which is related to the later feature extraction stage.

### 3.1.1. Traditional speckle suppression algorithm.

The Boxcar Filter indiscriminately averages all pixels within a rectangular window, so both homogeneous areas and edge areas are smoothed, which will cause a decrease in image resolution. The larger the window, the more smoothed it is, and the more blurred the texture features are. Lee et al. [2] proposed the Lee Filter, which selects a window of a certain length as a local area and uses the local statistical characteristics of the image to perform speckle filtering on SAR images. It is a typical method based on a fully developed speckle noise model. Due to the defects of the Lee Filter, that is, it does not filter homogeneous area pixels close to the edge or point target sufficiently, Lee et al. [3] proposed the Refine Lee Filter, which is an adaptive filtering algorithm based on edge detection. It improves the accuracy of estimation by redefining the neighborhood of the center pixel, and usually uses a 7×7 sliding window. Then Lee et al. [4] proposed the Sigma Filter, which selects similar pixels based on the sigma parameter and adopts a strong point protection mechanism to retain target details. Vasile et al. [5] proposed the IDAN Filter, which is a region growing technology based on which candidate pixels are selected for filtering according to the pixel amplitude value information. Deledalle et al. [6] proposed the PPB Filter, proposed a concept of probability sub-block (PPB) based on the image noise distribution characteristics for similarity measurement, which has good applicability for SAR image speckle suppression; Xu et al. [7] proposed POTDF Filter, which is a speckle filtering method based on logarithmic-exponential transformation and block sorting; Dabov et al. [8] proposed a three-dimensional block matching filter algorithm BM3D Filter based on the NLM algorithm and combined with sparse representation. The algorithm is to form a three-dimensional array matrix of similar image sub-blocks, and then perform hard threshold filtering on the three-dimensional array to obtain a preliminary estimation result, and then perform the same operation on the preliminary estimation result, and then use the original image to perform Wiener filtering on the image obtained by the preliminary estimation, and finally obtain the filtered image. For the multiplicative speckle noise model of SAR images, Parrilli et al. [9] proposed SAR-BM3D Filter, which combines the non-domain filtering idea with wavelet edge, implements speckle filtering based on the linear minimum mean square error criterion, and obtains a good denoising effect on SAR images.

### 3.1.2. Deep learning-based speckle suppression algorithm.

In recent years, with the widespread application of deep learning methods in the field of image processing, some works have also used it to suppress SAR image speckle and achieved good results [10-12]. Chierchia et al. [13] proposed the SAR-CNN Filter, which uses 17 convolutional layers to estimate additive speckle noise in the logarithmic domain and perform filtering; Wang et al. [14] proposed the ID-CNN Filter, which is based on the multiplicative speckle model and uses an 8-layer convolutional neural network, batch normalization, rectified linear unit activation function and a component division residual layer to estimate speckle, and uses a combination of Euclidean loss and total change loss for training, which can automatically remove speckle noise from noisy images; Zhang et al. [15] proposed the SAR-DRN Filter, which uses an 8-layer convolutional neural network, batch normalization, rectified linear unit activation function and a component division residual layer to estimate speckle. Filter, which introduced the dilated residual network into the SAR image despeckle task. The network used dilated convolution to improve the image detail protection capability and skip connection to avoid the gradient vanishing problem during training. Cozzolino et al. [16] proposed a SAR image despeckle method combining CNN and NLM. This method used CNN to calculate the weights of similar image blocks in NLM, further improving the despeckle quality of NLM-based methods. Shen et al. [17] combined the optimization model with deep CNN denoising prior and proposed a recursive deep CNN method. Liu et al. [18] proposed a CNN despeckle method based on image spatial and transform domain features to further improve the despeckle performance. Mullissa et al. [19] proposed the DeSpeckNet Filter, which estimated the despeckle SAR image and the despeckle SAR image respectively, and used their mean square error as the loss function of the training process.

## 3.2. SAR Image Target Detection

Early SAR image interpretation relied on experienced experts to extract intelligence. However, in the face of massive amounts of data, manual interpretation is inefficient, and the interpretation results vary due to the individual knowledge and experience of experts, resulting in unstable reliability. Therefore, SAR automatic target recognition (ATR) technology is an important part of SAR image automatic interpretation, and the Lincoln Laboratory has proposed a three-level SAR ATR processing flow, including detection, discrimination and identification, as shown in Figure 1. It uses the powerful processing power of computers to detect targets in SAR images without human intervention [16, 17]. Constant false alarm rate algorithms and their improved algorithms are widely used, but they are susceptible to noise interference and have complex calculations, which is not conducive to real-time requirements in real tasks. In recent years, with the development of computer hardware, the ability of computer processing and calculation has been greatly improved, so SAR image target detection algorithms based on deep learning have been widely studied and applied.

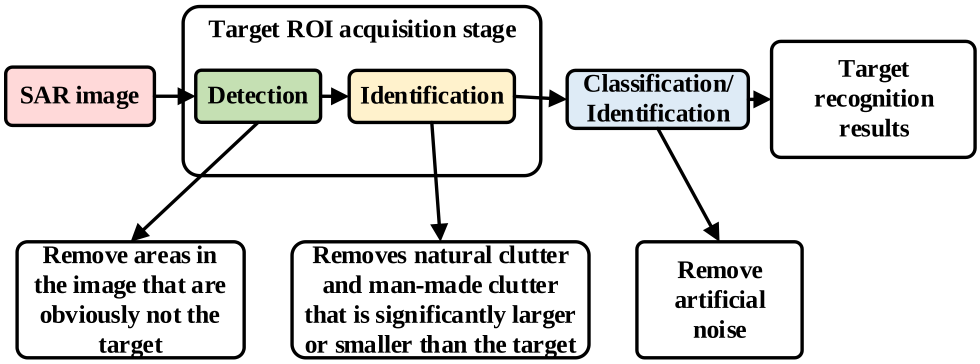


**Figure 1.** SAR ATR three-level processing flow chart proposed by Lincoln Laboratory

### 3.2.1. Traditional SAR image target detection algorithm.

Traditional SAR image detection methods require human participation in designing target features. These features generally include texture, shape, contrast, grayscale, etc., which can be used to distinguish land, sea areas, and targets. The earliest SAR ATR technology can be traced back to the SAR target detector based on constant false alarm rate (CFAR) proposed by Lincoln Laboratory of Massachusetts Institute of Technology in the late 1980s. In order to improve the robustness of the threshold segmentation algorithm, it adaptively selects the appropriate threshold while keeping the false alarm rate constant to detect the target [18, 19]. Compared with the global threshold segmentation algorithm, the dual-parameter CFAR can effectively suppress the false alarm caused by coherent speckle noise and natural clutter. This result was applied to the SAR ATR subsystem of the Semi-Automated Image Intelligence Processing (SAIP) system funded by DARPA [19, 20]. However, the Gaussian distribution model used in the dual-parameter CFAR is not suitable for all scenarios. Therefore, subsequent researchers have improved the detection effect of the CFAR detector according to different scenarios and the density of target distribution, such as GO-CFAR [25]. CFAR-type detectors such as SOCFAR [25], OS-CFAR [21, 22], AC-CFAR [27], CM-CFAR[28], and Superpixel-CFAR[29] have been proposed one after another.

### 3.2.2. SAR image target detection algorithm based on deep learning.

Deep learning methods have also achieved good performance in SAR target detection [30-34]. However, due to the great differences between natural images and SAR images, we cannot simply apply the deep learning framework directly to SAR image detection. SAR image target detection algorithms based on deep learning can be roughly divided into one-stage detection and two-stage detection according to the detection stage. a) The two-stage detection algorithm lays a series of anchor boxes on the original image, uses a CNN network, performs two classifications and two regressions on these anchor boxes, and obtains the detection results [35]. Girshick et al. [36] first used selective search to generate candidate regions; He et al. [37] proposed the SPP-Net model based on R-CNN, which removed the limitation of CNN on input size and greatly accelerated the running speed; Due to the multi-scale and multi-feature characteristics of SAR images, relying solely on the fixed-size convolution kernel commonly used in CNN to perform convolution operations, the feature extraction process may only extract local features of the target after multiple convolution operations, resulting in incomplete extraction of the target's feature system, thus leading to missed detection or misdetection of the target. Ai et al. [38] proposed a multi-kernel feature fusion convolutional neural network to ensure that the target in the SAR image maintains complete feature information during the feature extraction process; Wang et al. [39] proposed a method based on visual attention, which can improve the detection rate of targets in high-resolution complex SAR images. b) Single-stage detection algorithm lays a series of anchor boxes on the original image, uses a full convolutional network to perform one classification and one regression on these anchor boxes to obtain the detection result [35]. Scholars have proposed a series of related detectors such as YOLO [40], SSD [41], DSOD [42], RetinaDe[43] and CornerNe [44]. Due to the uniqueness of this type of algorithm, it can greatly improve the network's detection speed of the target. Therefore, this type of algorithm is favored and studied by more and more scholars, especially the YOLO series of algorithms. Long et al. [45] proposed a lightweight model of Lira-YOLO, which combines the advantages of YOLOv3 [46] and RetinaDe [43] models, and can reduce the number of model parameters and its computational complexity while maintaining its detection accuracy. Chang et al. [47] proposed a YOLOv2-reduced model, studied the use of high-performance GPU in SAR ship target detection applications, and achieved significant improvements compared with YOLOv2. Zhang et al. [48] proposed a DS-CNN network, which combined the deep convolution D-Conv2D and point-by-point convolution P-Conv2D to form a deep separable convolution DS-CNN to replace the traditional convolutional neural network C-CNN. The detection speed was greatly improved with only a slight sacrifice in detection accuracy. Gao et al. [49] proposed an enhanced ship feature extraction method, which improved the RetinaNet by combining the split convolution block that enhances the feature representation of small targets and the spatial attention block that reduces the spatial information loss caused by the dimensionality reduction process. The detection accuracy is improved, but the overall detection time is slightly increased; Zhang et al. [50] designed a lightweight feature optimization network LFO-Net to improve SSD, which has significant advantages in speed and accuracy, but the model seems to miss some small offshore ships. Du et al. [51] proposed a new semi-supervised SAR ship detection network scene feature learning, which focuses on using scene-level annotations of SAR images to improve detection performance with limited target-level annotations.

### 3.3. Rotating Target Detection

Due to the difference between SAR images and natural images, SAR images have complex backgrounds, densely distributed targets and arbitrary orientations. If a horizontal box is used to select the target like a natural image, a large amount of background will be introduced into the generated candidate box, and there will be a large amount of stacking of adjacent candidate boxes. This is prone to misdetection or omission of targets after the subsequent non-maximum suppression processing, resulting in a low recall rate. Therefore, rotated target detection in SAR images is particularly important. Jiang et al. [52] proposed a rotated region CNN based on Faster R-CNN, namely R2CNN, for detecting arbitrarily oriented text in natural images; Ma et al. [53] proposed RRPN based on the improvement of ordinary RPN, using the rotating RPN to generate angled anchor boxes, and proposed the calculation of oblique box IoU, and also proposed RRoI pooling layer; Qian et al. [54] changed the commonly used five-parameter regression method for the rotated box to an eight-parameter regression method, which may cause training instability and performance degradation, and proposed a new loss modulated rotation loss; Yang et al. [55] proposed SCRDet, a multi-category rotation detector for small, cluttered and rotating objects, which improved the sensitivity of small target detection; Xu et al. [56] solved the problem that the parameters of the eight-parameter regression method are not labeled according to the rules. They first detected the horizontal box and then slid the four fixed points of the horizontal box to obtain the rotated box; Chen et al. [57] proposed a new loss function PIoU to solve the problem that the loss is equal when the loss is equal when using the Somth-L1 loss function, which results in the loss being unable to truly measure the regression accuracy of the box; Li et al. [58] proposed anchor-free rotation target detection, which uses adaptive point representation to capture the geometric information of any oriented instance; Zhang et al. [59] proposed a refined rotation target detector, the foreground refinement network (FoRDet), which first uses a coarse stage to distinguish the foreground and background, and then a fine stage to output specific classification and prediction box, and designed a foreground anchor weighted loss (FRW) loss. Zheng et al. [60] proposed a loss function based on the circumscribed circle of the rotation bounding box and the intersection-over-union ratio. This function not only considers the distance between the center point of the predicted box and the true box, but also considers the correlation between the parameters of the rotation box, which has a good effect. Du et al. [61] proposed a SAR target detection method based on the Faster R-CNN detection model and combined it with reinforcement learning to adaptively select candidate boxes. This method can adaptively search for areas in the feature map that may contain targets through reinforcement learning, and select candidate boxes in the search area for further classification and regression.

# 4. DATASET SUMMARY AND EVALUATION METRICS

## 4.1. Dataset Summary

Appropriate dataset selection can improve the performance of the network model during training. The following public SAR image datasets are used:MSTAR, RBox-SSDD, RSDD-SAR and SAR-AIRcraft-1.0. The datasets are summarized in Table 1.

**Table 1.** Dataset Summary

| Dataset name | Total number of images | Total number of targets | Target Category | Data characteristics |
|---|---|---|---|---|
| MSTAR | 5,173 | 2,740 | Military targets (10 categories) | Multi-view, multi-scene, high resolution, classic benchmark |
| RBox-SSDD | 1,160 | 2,456 | Ships | Precise rotation frame annotation, near-shore and offshore scenes |
| RSDD-SAR | 7,000 | 10,263 | Ships | Multi-mode, multi-polarization, random direction, high proportion of small targets |
| SAR-AIRcraft-1.0 | 4,368 | 16,463 | Aircraft (7 categories) | Rich categories, multi-scale, complex background |

The Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset is provided by the Defense Advanced Research Projects Agency and the Air Force Research Laboratory of the United States. It contains 5,173 SAR images, covering 10 types of stationary and moving military targets, such as tanks, armored vehicles, trucks, etc., and the targets are all at different perspectives and azimuths (). In addition, it also contains SAR images of complex scenes such as forests, ground, and buildings. It is an important benchmark dataset in the field of target recognition.

Box-SSDD (Rotatable Bounding Box SSDD) is a dataset expanded on the basis of SSDD (SAR Ship Detection Dataset). It converts the original horizontal box annotations into rotating box annotations and is dedicated to the task of ship target detection in SAR images. The dataset contains 1,160 images and 2,456 ship targets, with an average of about 2.12 ship instances per image. Ship targets are of various sizes and directions, covering complex scenes such as near-shore and offshore, and are suitable for the research of ship detection and rotating box detection tasks in SAR images.

The Rotated Ship Detection Dataset in SAR Images (RSDD-SAR) is constructed based on remote sensing data from the domestic Gaofen-3 satellite and the European Space Agency's TerraSAR-X satellite. It contains 127 SAR images, generates 7,000 slices, and covers 10,263 ship targets. The dataset has multiple imaging modes, polarization modes, and different resolutions. It has the characteristics of complex scenes, diverse target shapes, significant changes in aspect ratios, and a high proportion of small targets. It is suitable for rotation direction detection tasks.

SAR Aircraft Detection and Recognition Dataset-1.0 (SAR-AIRcraft-1.0) has a total of 4,368 images, a total of 4 high-resolution datasets of different image sizes, and a total of 16,463 aircraft target instances, including 7 categories of A220, A320/321, A330, ARJ21, Boeing737, Boeing787, and other, supporting detection tasks, recognition tasks, and integrated detection and recognition tasks. This dataset has the characteristics of complex scenes, rich categories, dense targets, noise interference, diverse tasks, and multi-scale.

## 4.2. Evaluation Metrics

### 4.2.1. Evaluation Indicators for Speckle Suppression.

In the process of suppressing speckle noise, the quality of SAR images may suffer a certain degree of loss or distortion. In order to objectively evaluate the performance of different speckle suppression algorithms, it is crucial to evaluate the quality of denoised SAR images. This not only helps to measure the denoising effect of the algorithm, but also optimizes the processing method to achieve a better balance between noise suppression and image detail preservation. The evaluation indicators of our speckle suppression model are mainly evaluated based on the different characteristics of synthetic images and real images. For synthetic images, the structural similarity index (SSIM) and peak signal to noise ratio (PSNR) are used as the main indicators . The peak signal to noise ratio (PSNR) mainly uses the difference between the grayscale values of the pixels of the two images to give a quantitative score, and then evaluates the image quality based on the score. A higher PSNR value means a lower noise residue. The formula of PSNR is defined as follows:

$$PSNR = 10 \cdot log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{1}$$

Where *MAX* represents the maximum value of the image pixel value, represents the mean square error, which is defined as:

$$MSE = \frac{1}{m \cdot n}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(I\left(i,j\right) - K\left(i,j\right)\right)^2 \tag{2}$$

Where $I(i,j)$ and $K(i,j)$ represent the pixel values of the reference image and the processed image respectively, $m$ and $n$ represent the number of rows and columns of the image.

Structural similarity SSIM is an indicator used to evaluate the similarity between two images. The higher the value, the better the image quality. This indicator comprehensively considers factors such as brightness, contrast and structural information. The evaluation of image quality is more in line with human visual perception characteristics. Therefore, it is widely used in image denoising, compression and enhancement. The formula of SSIM is:

$$SSIM(I,K)=\frac{(2\mu_I\mu_K+C_1)(2\sigma_{IK}+C_2)}{(\mu_I^2+\mu_K^2+C_1)(\sigma_I^2+\sigma_K^2+C_2)} \tag{3}$$

Where, $\mu_I$ and $\mu_K$ represents the average value of two images; $\sigma_I^2$ and $\sigma_K^2$ represents the variance of $C_2$.

#### 4.2.2. Evaluation Metrics for Rotational Target Detection

By conducting comparative experiments on the rotation object detection model, we can intuitively see the performance of the model. In general, it can be evaluated by precision, recall rate, F1 score, and mean average precision (mAP) .

Precision is the ratio of correctly detected positive samples to all predicted positive samples, which is used to measure the accuracy of model prediction. The higher the precision, the stronger the model's ability to distinguish positive samples. The formula is:

$$Precision=\frac{TP}{TP+FP} \tag{4}$$

In the formula, TP represents the number of positive samples detected correctly; FP represents the number of samples detected incorrectly as positive samples.

Recall Rate is the ratio of correctly detected positive samples to all actual positive samples, which is used to measure the detection coverage of the model. The higher the recall rate, the more positive samples the model can capture. The formula is:

$$Recall=\frac{TP}{TP+FN} \tag{5}$$

Mean Average Precision is the mean of the AP values of all categories and is the core indicator for measuring the overall performance of the detection model. The formula is:

$$mAP=\frac{1}{N}\sum_{i=1}^{N}AP_i \tag{6}$$

Where N represents the number of detection categories; APi represents the average precision of the th category.

F1-Score is one of the important evaluation indicators for measuring the performance of target detection models. It is the harmonic mean of precision and recall. The formula is:

$$F1=2\cdot\frac{Precision+Recall}{Precision\cdot Recall} \tag{7}$$

The value range of F1 is [0, 1] , and the closer the value is to 1, the better the model performance.

# 5. SUMMARY AND OUTLOOK

## 5.1. Summary

In summary, we can understand that the deep learning-based SAR image rotation target detection method under complex background needs to construct a suitable deep network structure, support the contextual information, and combine image enhancement and loss function design, so as to have better processing capabilities for SAR image coherent speckle suppression and complex background with numerous, dense and arbitrary rotating targets. Its shortcomings are: the complex deep network structure and too many network parameters will lead to a large amount of calculation for rotating target detection under complex background, increase the difficulty of training, and require high computing power. In addition, the rotating box test data set related to relevant SAR images needs to be expanded urgently.

(1) Extracting the key features of targets in SAR images with complex backgrounds is the key to target detection, and whether the features are complete is very important. However, the inherent coherent speckle noise in SAR images with complex backgrounds leads to blurred edges and small-scale objects are easily lost after multiple convolutions, which makes efficient and accurate target detection difficult. How to effectively suppress the coherent speckle noise in SAR images and reduce the loss of key detail features of the target after denoising.

(2) There is a lack of datasets with rotation box annotations of targets in SAR images with complex backgrounds to optimize model training and improve model accuracy, robustness and real-time performance. How to form a complete and rich dataset and study the boundary problems faced by rotation target detection to enhance the generalization ability of the model.

(3) How to strengthen the innovation and optimization of the basic model framework based on deep learning. Design a lightweight network architecture and develop an efficient training algorithm so that it can be deployed on low-cost, low-power and low-computation mobile devices and embedded devices processing platforms to reduce the requirements for hardware devices.

## 5.2. Outlook

(1) Speckle suppression of SAR images: To address the inherent speckle noise of SAR images, images need to be denoised. To suppress speckle, it is necessary to protect edges, points, lines, and other targets, image texture features, and feature details and key points. This prevents the loss of key features of the image after denoising.

(2) Problems with SAR image datasets: To address the problem of insufficient datasets of rotated targets in current SAR images, a rotation label annotation tool is used to organize and improve the datasets of rotated annotations and enhance the generalization ability of the model.

(3) Problems in the rotated target detection algorithm: The boundary problem, square-like detection problem, loss and evaluation inconsistency problems faced by rotated target detection are solved by adding solutions, and then the rotation target detection loss function with high-precision detection and scale invariance is combined to improve the performance of the model.

(4) Through the study of the attention mechanism, an attention mechanism is designed that can adaptively focus on target features according to different targets. In addition, the detection accuracy of the model for multiple and small targets in complex backgrounds is kept basically unchanged while improving the detection speed of the model.

# REFERENCES


[1] Sun Z, Dai M, et al. Fast Detection of Ship Targets for Complex Large-scene SAR Images Based on a Cascade Network [J]. Signal Processing, 2021, 37(6): 941–951.

[2] Lee J-S. Digital Image Enhancement and Noise Filtering by Use of Local Statistics [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1980, PAMI-2(2): 165–168.

[3] Lee J-S. Refined filtering of image noise using local statistics [J]. Computer Graphics and Image Processing, 1981, 15(4): 380–389.

[4] Lee J-S. A simple speckle smoothing algorithm for synthetic aperture radar images [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1983, SMC-13(1): 85–89.

[5] Vasile G, Trouve E, Jong-Sen Lee, et al. Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(6): 1609–1621.

[6] Deledalle C-A, Denis L, Tupin F. Iterative Weighted Maximum Likelihood Denoising With Probabilistic Patch-Based Weights [J]. IEEE Transactions on Image Processing, 2009, 18(12): 2661–2672.

[7] Xu B, Cui Y, Li Z, et al. Patch Ordering-Based SAR Image Despeckling Via Transform-Domain Filtering [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(4): 1682–1695.

[8] Dabov K, Foi A, Katkovnik V, et al. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering [J]. IEEE Transactions on Image Processing, 2007, 16(8): 2080–2095.

[9] Parrilli S, Poderico M, Angelino C V, et al. A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage [J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(2): 606–616.

[10] Molini A B, Valsesia D, Fracastoro G, et al. Speckle2Void: Deep Self-Supervised SAR Despeckling With Blind-Spot Convolutional Neural Networks [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1–17.

[11] Xiong K, Zhao G, Wang Y, et al. SPB-Net: A Deep Network for SAR Imaging and Despeckling With Downsampled Data [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(11): 9238–9256.

[12] Dalsasso E, Yang X, Denis L, et al. SAR Image Despeckling by Deep Neural Networks: from a Pre-Trained Model to an End-to-End Training Strategy [J]. Remote Sensing, 2020, 12(16): 2636.

[13] Chierchia G, Cozzolino D, Poggi G, et al. SAR image despeckling through convolutional neural networks [J]. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX: IEEE, 2017: 5438–5441.

[14] Wang P, Zhang H, Patel V M. SAR Image Despeckling Using a Convolutional Neural Network [J]. IEEE Signal Processing Letters, 2017, 24(12): 1763–1767.

[15] Zhang Q, Yuan Q, Li J, et al. Learning a Dilated Residual Network for SAR Image Despeckling [J]. Remote Sensing, 2018, 10(2): 196.

[16] Cozzolino D, Verdoliva L, Scarpa G, et al. Nonlocal Sar Image Despeckling by Convolutional Neural Networks [J]. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan: IEEE, 2019: 5117–5120.

[17] Shen H, Zhou C, Li J, et al. SAR Image Despeckling Employing a Recursive Deep CNN Prior [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(1): 273–286.

[18] Liu Z, Lai R, Guan J. Spatial and Transform Domain CNN for SAR Image Despeckling [J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1–5.

[19] Mullissa A G, Marcos D, Tuia D, et al. deSpeckNet: Generalizing Deep Learning-Based SAR Image Despeckling [J]. IEEE Transactions on Geoscience and Remote Sensing, Institute of Electrical and Electronics Engineers, 2022, 60: 1–15.

[20] Liu S. Theoretical research on target detection and recognition in low detection rate SAR images [D]. University of Electronic Science and Technology of China, 2017.

[21] Novak L, Owirka G, Brower W, et al. The Automatic Target- Recognition System in SAIP [A]. 1997.

[22] Novak L M, Halversen S D, Owirka G, et al. Effects of polarization and resolution on SAR ATR [J]. IEEE Transactions on Aerospace and Electronic Systems, 1997, 33(1): 102–116.

[23] Novak L M, Owirka G J, Netishen C M. Radar target identification using spatial matched filters [J]. Pattern Recognition, 1994, 27(4): 607–617.

[24] Shaw G A, Zuerndorfer B W, Ford R A, et al. RASSP Benchmark 2 Technical Description. [A]. Fort Belvoir, VA: Defense Technical Information Center, 1995.

[25] He Y, Guan J, et al. Radar target detection and constant false alarm processing [M]. Tsinghua University Press, 2011.

[26] Rohling H. Radar CFAR Thresholding in Clutter and Multiple Target Situations [J]. IEEE Transactions on Aerospace and Electronic Systems, 1983, AES-19(4): 608–621.

[27] Gui Gao, Li Liu, Lingjun Zhao, et al. An Adaptive and Fast CFAR Algorithm Based on Automatic Censoring for Target Detection in High-Resolution SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(6): 1685–1697.

[28] Rickard J, Dillard G. Adaptive Detection Algorithms for Multiple-Target Situations [J]. IEEE Transactions on Aerospace and Electronic Systems, 1977, AES-13(4): 338–343.

[29] Yu W, Wang Y, Liu H, et al. Superpixel-Based CFAR Target Detection for High-Resolution SAR Images [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(5): 730–734.

[30] Cui Z, Li Q, Cao Z, et al. Dense Attention Pyramid Networks for Multi-Scale Ship Detection in SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 8983–8997.

[31] Xu X, Zhang X, Zhang T. Lite-YOLOv5: A Lightweight Deep Learning Detector for On-Board Ship Detection in Large-Scene Sentinel-1 SAR Images [J]. Remote Sensing, 2022, 14(4): 1018.

[32] 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, 2019, 16(1): 150–154.

[33] Fu K, Fu J, Wang Z, et al. Scattering-Keypoint-Guided Network for Oriented Ship Detection in High-Resolution and Large-Scale SAR Images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 11162–11178.

[34] Zhang T, Zhang X, Liu C, et al. Balance learning for ship detection from synthetic aperture radar remote sensing imagery [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 182: 190–207.

[35] Hou X, Jin G, et al. Review of ship target detection in SAR images based on deep learning [J]. Laser & Optoelectronics Progress, 2021, 58(4): 0400005.

[36] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [J]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA: IEEE, 2014: 580–587.

[37] He K, Zhang X, Ren S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916.

[38] Ai J, Mao Y, Luo Q, et al. SAR Target Classification Using the Multikernel-Size Feature Fusion-Based Convolutional Neural Network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1–13.

[39] Visual Attention-Based Target Detection and Discrimination for High-Resolution SAR Images in Complex Scenes | Semantic Scholar [EB/OL]. /2023-11-01. https://www.semanticscholar.org/paper/Visual-Attention-Based-Target-Detection-and-for-SAR-Wang-Du/8df684a757894b0025b2116e234dcaadd06cff11.

[40] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection [J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, 2016: 779–788.

[41] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector [A]. B. Leibe, J. Matas, N. Sebe, et al. Cham: Springer International Publishing, 2016, 9905: 21–37.

[42] Shen Z, Liu Z, Li J, et al. DSOD: Learning Deeply Supervised Object Detectors from Scratch [J]. 2017 IEEE International Conference on Computer Vision (ICCV), Venice: IEEE, 2017: 1937–1945.

[43] Zhang S, Wen L, Bian X, et al. Single-Shot Refinement Neural Network for Object Detection [J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT: IEEE, 2018: 4203–4212.

[44] Law H, Deng J. CornerNet: Detecting Objects as Paired Keypoints [J]. International Journal of Computer Vision, 2020, 128(3): 642–656.

[45] Long Z, Suyuan W, Zhongma C, et al. Lira-YOLO: A lightweight model for ship detection in radar images [J]. Journal of Systems Engineering and Electronics, 2020, 31(5): 950–956.

[46] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement [J]. ArXiv, 2018.

[47] Chang Y-L, Anagaw A, Chang L, et al. Ship Detection Based on YOLOv2 for SAR Imagery [J]. Remote Sensing, 2019, 11(7): 786.

[48] Zhang, Zhang, Shi, et al. Depthwise Separable Convolution Neural Network for High-Speed SAR Ship Detection [J]. Remote Sensing, 2019, 11(21): 2483.

[49] Gao F, Shi W, Wang J, et al. Enhanced Feature Extraction for Ship Detection from Multi-Resolution and Multi-Scene Synthetic Aperture Radar (SAR) Images [J]. Remote Sensing, 2019, 11(22): 2694.

[50] Zhang X, Wang H, Xu C, et al. A Lightweight Feature Optimizing Network for Ship Detection in SAR Image [J]. IEEE Access, 2019, 7: 141662–141678.

[51] Du Y, Du L, Guo Y, et al. Semi-Supervised SAR Ship Detection Network via Scene Characteristic Learning [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, PP: 1–1.

[52] Jiang Y, Zhu X, Wang X, et al. R2CNN: Rotational Region CNN for Orientation Robust Scene Text Detection [J]. ArXiv, 2017.

[53] Ma J, Shao W, Ye H, et al. Arbitrary-Oriented Scene Text Detection via Rotation Proposals [J]. IEEE Transactions on Multimedia, 2018, 20(11): 3111–3122.

[54] Qian W, Yang X, Peng S, et al. Learning Modulated Loss for Rotated Object Detection [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(3): 2458–2466.

[55] Yang X, Yang J, Yan J, et al. SCRDet: Towards More Robust Detection for Small, Cluttered and Rotated Objects [A]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) [C]. Seoul, Korea (South): IEEE, 2019: 8231–8240.

[56] Xu Y, Fu M, Wang Q, et al. Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(4): 1452–1459.

[57] Chen Z, Chen K, Lin W, et al. PIoU Loss: Towards Accurate Oriented Object Detection in Complex Environments [J]. A. Vedaldi, H. Bischof, T. Brox, et al. Cham: Springer International Publishing, 2020, 12350: 195–211.

[58] Li W, Chen Y, Hu K, et al. Oriented RepPoints for Aerial Object Detection [A]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2022: 1819–1828.

[59] Zhang T, Zhang X, Zhu P, et al. Foreground Refinement Network for Rotated Object Detection in Remote Sensing Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1–13.

[60] Zheng Z, Zhang T, et al. Ship target detection method in SAR images based on rotation loss function RCIoU [J]. Journal of Shandong University (Engineering Science Edition), 2022, 52(2): 15-22.

[61] Du Lan, Wang Zilin, Guo Yuchen, et al. SAR target detection method based on adaptive candidate box selection combined with reinforcement learning [J]. Journal of Radars, 2022, 11(5): 884–896.