



HAL
open science

A Lightweight SAR Ship Detection Network Based on Superpixel Statistical Modeling

Zhengxi Guo, Biao Hou, Bo Ren

► **To cite this version:**

Zhengxi Guo, Biao Hou, Bo Ren. A Lightweight SAR Ship Detection Network Based on Superpixel Statistical Modeling. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.294-301, 10.1007/978-3-031-14903-0_31 . hal-04666454

HAL Id: hal-04666454

<https://hal.science/hal-04666454v1>

Submitted on 1 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

A Lightweight SAR Ship Detection Network Based on Superpixel Statistical Modeling

Zhengxi Guo, Biao Hou, and Bo Ren

Xidian University, Xi'an 710071, China
XiDianGZX@stu.xidian.edu.cn

Abstract. Deep learning technology has been widely used in SAR ship detection tasks. However, complex sea level backgrounds, such as sea clutter and shorelines, greatly interfere with the accuracy of the detection of ship targets. In addition, embedded devices need to deploy multiple detection models, such as FPGA, so model size and detection speed are also important indicators in practical applications. In order to solve these problems, we developed a method combining traditional detection methods with deep learning. In this paper, SAR image clutter distribution model is used to suppress SAR image clutter, and then the processed image is sent to the network for learning. Based on this idea, we have established a superpixel composite Gamma distribution model, which can obtain more accurate fitting results than pixel scale Gamma distribution and suppress background clutter more effectively. We also propose the YOLO-SX lightweight detection network, which significantly reduces model size, detection time, calculation parameters, and memory consumption. Its overall performance is superior to other detection methods.

Keywords: SAR Ship Detection · Lightweight Deep Learning · Superpixel Statistical Modeling · Gamma Distribution.

1 Introduction

Synthetic aperture radar (SAR) is an active sensor. Compared with optical and in-frared sensors, SAR has the characteristics of all-weather, all-day and penetration capability [1]. Due to these advantages, SAR technology has been widely used in military and civilian fields. With the development of SAR technology, SAR image interpretation faces many opportunities and challenges.

The field of SAR ship detection is broadly divided into two development stages: traditional methods and deep learning methods. Before the advent of deep learning, CFAR (Constant False Alarm Rate) algorithm was widely used in the detection of ground and marine targets in SAR images. It is a common adaptive algorithm for radar systems to detect target echoes in noisy, cluttered and interfering backgrounds [2]. It detects ship targets by modeling the statistical distribution of background clutter, but it is often difficult to model the clutter in a complex background with an appropriate probability density function [3]. Therefore, many previous researches have been devoted to the variation,

combination, and adjustment of parameter models to meet the desired accuracy of statistical modeling of clutter to determine detection thresholds [4]. However, in practical applications, the processing speed of the algorithm can not meet the requirements. Since AlexNet won the ImageNet image classification challenge in 2012, deep neural networks have shown very high accuracy and reliability in image detection. Traditional feature extraction methods are gradually replaced by convolutional neural networks [5].

The purpose of this paper is to combine SAR image statistical modeling with deep learning, and propose a SAR ship detection method based on SAR image statistical distribution and YOLOV5. Different from the traditional statistical modeling, we establish the clutter distribution model in the superpixel domain of the image, and send the processed image into the lightweight deep network for learning, so as to improve the detection speed without losing accuracy and improve the ship detection ability under complex backgrounds. In this paper, Section 2 introduces the method based on superpixel statistical modeling, Section 3 describes the construction of the lightweight detection network, Section 4 introduces the experiments and Section 5 is the conclusion.

2 Superpixel-Based Composite Gamma Distribution Modeling

2.1 Superpixel

We define a block of pixels consisting of adjacent pixels with similar texture, color and grayscale value in an image as a superpixel region [6]. Most of these regions retain the effective information for further image processing, and generally do not destroy the boundary information of objects in the image. Few superpixels are used to express image features instead of a large number of pixels to reduce the complexity of subsequent image processing. The segmentation effect as shown in Fig. 1(a).

2.2 Composite Gamma Distribution Based on Superpixel

First, we calculate two key parameters based on the divided superpixels: 1) the number of pixels contained in each superpixel, called the superpixel domain; 2) the average gray value of the pixels contained in each superpixel, called the superpixel reflectivity [7].

Then we calculate the distribution law of each superpixel domain according to Equation (1), that is, the proportion of the number of pixels in each superpixel domain to the number of pixels in the whole image.

$$P(n) = \frac{N(n)}{\sum_{i=1}^S N(i)} \quad (1)$$

where n is the index of the superpixel domain, S is the number of superpixel domains of this image, and $N(i)$ denotes the number of pixels in the i -th superpixel domain.

Then calculate the probability density function of superpixel reflectivity as in (2).

$$f(Z) = \sum_{n=\min}^{\max} f_{z|n}(z | n)p(n) = \sum_{n=\min}^{\max} \frac{z^{\varphi_n-1} e^{-z/\sigma_n}}{\sigma_n^{\varphi_n} \Gamma(\varphi_n)} p(n) \quad (2)$$

where $p(n)$ denotes the distribution law of the superpixel domain, φ_n and σ_n the Gamma parameters of the superpixel reflectance, min denotes the minimum superpixel domain, and max denotes the maximum superpixel domain.

Then calculate the threshold value T as (3).

$$P_{fa} = \int_T^{\infty} f(z)dz \quad (3)$$

Where P_{fa} is the preset constant false alarm rate and $f(z)$ is the probability density function of the superpixel reflectivity.

Finally, we perform clutter suppression on the image based on the threshold values calculated in the above steps, setting the pixel values smaller than the threshold T to 0 and keeping the pixel values larger than the threshold T . The result as shown in Fig. 1(b).

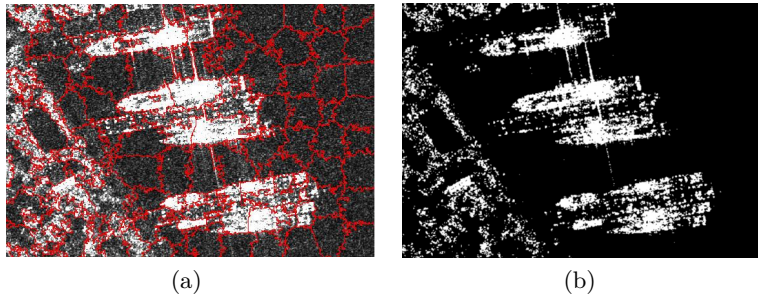


Fig. 1. (a) Superpixel segmentation result. (b) The result of clutter suppression.

3 Lightweight ship detection network

3.1 ShuffleNeck Module

To achieve the lightweight effect, we constructed a ShuffleNeck module (see Fig. 2) to replace the BottleNeck module in the YOLOV5 backbone network.

It is constructed by depth separable convolution (DWConv), channel shuffle and residual edges [8]. Firstly, the feature map is divided into two feature maps T1 and T2 with the same number of channels (half of the original number of channels) by a channel split operation. The first feature map T1 is passed through 1x1 convolution layer, 3x3 DWConv, and 1x1 convolution layer to obtain a deep feature map T3 containing a part of the original feature map. Then, the second

feature map T2 is spliced with T3 obtained by the first branch without any operation. Finally, the channel shuffle operation is used to obtain the deep feature maps with the same size and number of channels as the original feature maps.

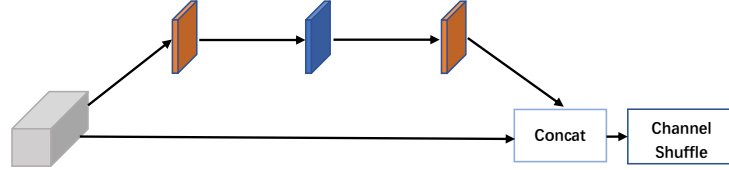


Fig. 2. The structure of ShuffleNeck module. In the figure, the orange cube represents 1x1 normal convolution and the blue cube represents 3x3 DWConv.

3.2 Shuffle-X(SX) Module

Unlike ShuffleNet, the SX module (see Fig. 3) processes the feature maps in two branches [9]. The first branch is a 1x1 DWConv. The second branch is a 1x1 DWConv and a ShuffleNeck module. Finally, the feature maps of the two branches are channel stitched to obtain the final deep features.

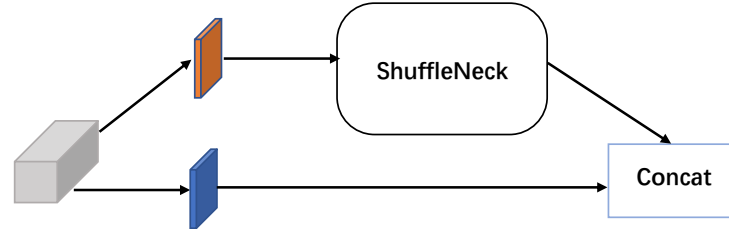


Fig. 3. The structure of SX module. In the figure, the orange cube represents 1x1 normal convolution and the blue cube represents 1x1 DWConv.

3.3 YOLO-SX Algorithm Description

In this paper, we use the proposed SX module to replace the C3 modules in YOLOV5. Specifically, our lightweight backbone network consists of DWConv and SX modules, which greatly reduces the number of parameters in the network. In the network, all four convolutional layers from P1-P4 have a step size of 2 and are used for down-sampling. The first SX module is used to quickly extract shallow features. The second and third SX modules are used to extract medium-scale feature maps. The last SX module is used to extract large scale features. In addition, our method makes extensive use of channel shuffle and residual

edges. This allows our network to retain enough shallow features for multi-scale detection tasks. The network structure as shown in Fig. 4.

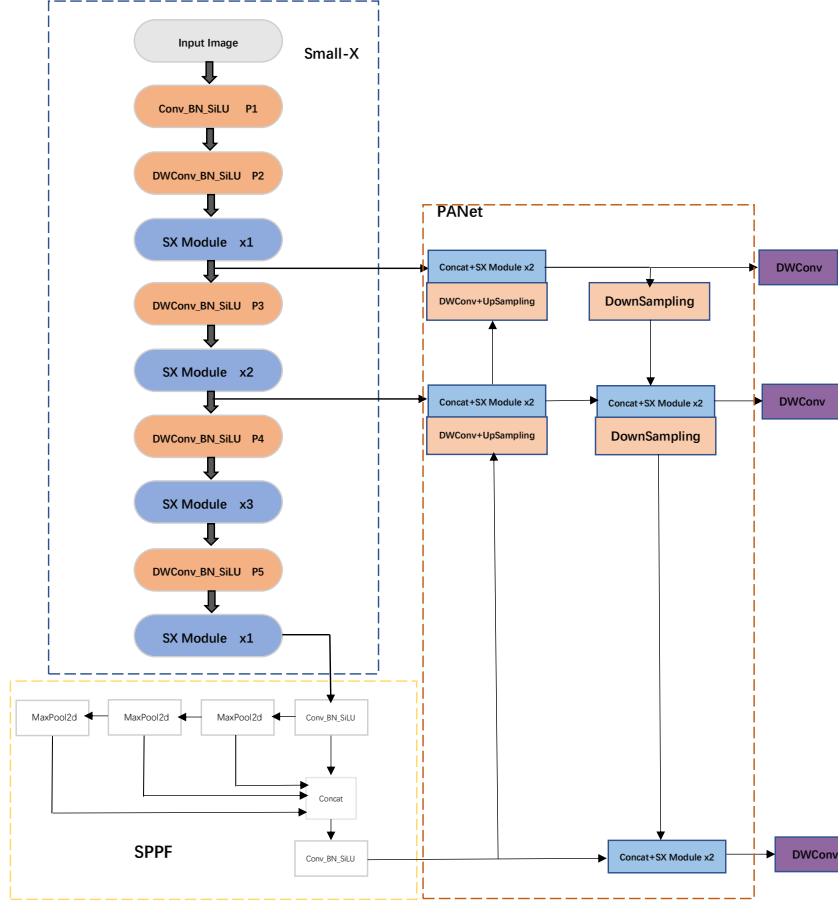


Fig. 4. The structure of YOLO-SX network.

4 Experiment

In this section, we will discuss the experimental dataset, process and results.

4.1 Experiment Dataset and Details

The experiment uses the SSDD (SAR-Ship Detection Dataset) which is publicly available and widely used for SAR ship detection algorithm performance measurement. The dataset has 1160 images with 2456 ships, and the average

number of ships per image is 2.12. The dataset has small ships with insignificant features, ships densely parallel moored in ports, large ships, severe ships under speckle noise and in complex background. Four polarization modes HH, HV, VV and VH, are also included [10].

We first perform clutter suppression in the superpixel domain for all images in the dataset, and then divide the training set, validation set and test set in the ratio of 7:2:1. The GPU used for the experiments is a 24G GeForce GTX3090. And the experimental platform is Ubuntu 16.04, using the Pytorch deep learning framework, implemented in python language.

4.2 Ablation Experiment

To verify the effectiveness of the clutter suppression method in this paper, we conducted ablation experiments using complex background and nearshore target data in SSDD. The images processed with clutter suppression and the images without processing are detected respectively. The results as shown in Fig. 6. We can clearly see that the missed detection rate of the images without clutter suppression is higher than the processed images, and the confidence level is also relatively lower.

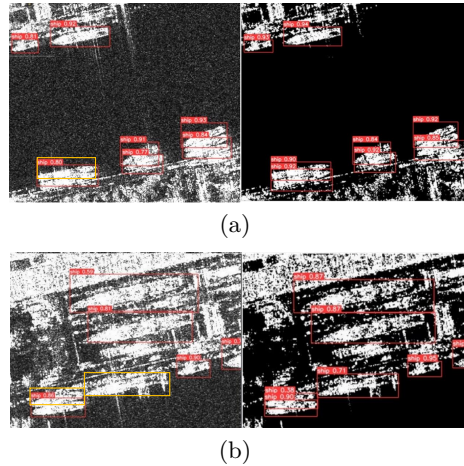


Fig. 5. The left side of (a) and (b) shows the ship images with complex backgrounds and near-shore targets without clutter suppression processing, and the right side shows the processed image. The red box is the detection result, and the yellow box is the missed ship target.

We also conducted experiments on the tiny target data on the sea surface to verify the effectiveness of our method, and the results as shown in Fig. 9. Obviously, the unprocessed detection results in missed detection of small targets with lower confidence than the images processed with clutter suppression.

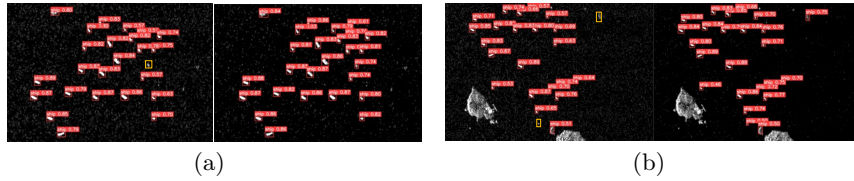


Fig. 6. The left side of (a) and (b) shows the image of tiny small targets on the sea surface without clutter suppression processing, and the right side shows the processed image. The red box is the detection result, and the yellow box is the missed ship target.

In summary, our method not only performs excellently in detecting ship targets with complex backgrounds, but is also applicable to cross-scale detection tasks.

4.3 Comparison Experiment

To evaluate the performance of the model, we use mAP (average accuracy), FPS (frames per second) and model size as evaluation metrics for ship detection. mAP is the average of the average accuracy for detecting each category. Since we have only ship targets in our experiments, mAP is equal to the average accuracy (AP).

We compare the performance of YOLOV4-tiny, YOLOV5x, SSD, Faster R-CNN and YOLO-SX by evaluating their performance in the above three aspects. The results as shown in Table 1. It is obvious that SSD and Faster R-CNN, which are two-stage detection networks, perform worse than YOLO. And although YOLOV5x achieves the best performance in terms of mAP, our method is much superior to YOLOV5x and other models in terms of FPS and model size.

Table 1. Comparison of different methods in mAP, FPS and model size

Model	mAP	FPS	Model Size(MB)
YOLOV4-Tiny	93.2%	43	22.4
YOLOV5x	98.8%	37	165.2
SSD	85.4%	36	93.2
Faster R-CNN	88.7%	27	527.8
YOLOV-SX	97.1%	50	1.4

5 Conclusion

This paper introduces a traditional SAR image statistical modeling method combined with deep learning. The method divides the image into several superpixel regions composed of adjacent pixels with similar texture, color, and grayscale values. These superpixels are then statistically modeled and clutter suppression using composite Gamma distributions. The experiments were conducted on

SSDD, a widely accepted dataset that is superior in accuracy and speed to existing algorithms, both for small ship targets in large areas of the ocean and for ship targets in complex settings. At the same time, a lightweight backbone was introduced to replace the existing backbone based on YOLOV5, significantly reducing the model's size and number of parameters and improving detection speed. It also facilitates the deployment of the model in embedded devices and is suitable for real-time tasks. It is hoped that this paper can provide a solution idea for future deep learning lightweight work and SAR image ship detection work.

Acknowledgements The authors would like to thank the Chinese Academy of Sciences for providing the SSDD dataset.

References

1. Leng, X.; Ji, K.; Yang, K.; Zou, H. A Bilateral CFAR Algorithm for Ship Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* 2015, 12, 1536–1540.
2. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.E.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In *Proceedings of the European Conference on Computer Vision (ECCV)*, Amsterdam, The Netherlands, 8–16 October 2016; pp. 21–37.
3. Liao, M.; Wang, C.; Wang, Y.; Jiang, L. Using SAR Images to Detect Ships from Sea Clut-ter. *IEEE Geosci. Remote Sens. Lett.* 2008, 5, 194–198.
4. X. Zhou, X. Liu, Q. Chen and Z. Zhang, "Power transmission tower CFAR detection algo-rithm based on integrated superpixel window and adaptive statistical model", *Proc. IGARSS - IEEE Int. Geosci. Remote Sens. Symp.*, pp. 2326-2329, Jul. 2019.
5. Krizhevsky, A.; Sutskever, I.; Hinton, G. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; ACM: New York, NY, USA, 2012; pp. 1097–1105.
6. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
7. M. Li, G. Wen, X. Huang, K. Li, and S. Lin, A Lightweight Detection Model for SAR Air-craft in a Complex Environment, *Remote Sensing*, vol. 13, no. 24, p. 5020, Dec. 2021, doi: 10.3390/rs13245020.
8. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems*; IEEE: New York, NY, USA, 2015; pp. 91–99.
9. Chen, Z.; Gao, X. An Improved Algorithm for Ship Target Detection in SAR Images Based on Faster R-CNN. In *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, Wanzhou, China, 9–11 November 2018; pp. 39–43.
10. K. J. Jiang, X. Fu, R. Qin, X. Wang, and Z. Ma, High-Speed Lightweight Ship Detection Algorithm Based on YOLO-V4 for Three-Channels RGB SAR Image, *Remote Sensing*, vol. 13, no. 10, p. 1909, May 2021, doi: 10.3390/rs13101909.