

DOI:10.19651/j.cnki.emt.2208891

一种基于改进 YOLOv4 的 SAR 舰船检测算法*

陈 洋¹ 张 明^{1,2} 杨立东¹ 喻大华¹ 张宝华¹ 李建军¹

(1. 内蒙古科技大学信息工程学院 包头 014010; 2. 大连海事大学信息科学与技术学院 大连 116026)

摘要: 在大量数据支持的背景下,如何高效利用大量 SAR 图像,提升舰船目标的检测精度是当前舰船目标检测的难题。本文聚焦如何提升 YOLOv4 算法对 SAR 舰船目标的检测精度,提出了一种融合多尺度和注意力增强的 YOLOv4 增强算法。在原 YOLOv4 的 PANet 中加入注意力模块(CBAM),同时使用加强的 K-means 聚类算法对数据集中的舰船目标真实框进行聚类,并对锚框结果进行线性比例变换,让算法锚框更适合于训练集。实验证明本文提出的算法在 SAR 舰船检测中的平均准确率(mAP)达到了 94.05%,比原始 YOLOv4 精度提高了 0.7%。实验结果充分证明本文提出的算法能够提升 SAR 舰船图像检测精度,为海上活动判断精确化提供技术支持。

关键词: SAR 图像;舰船检测;深度学习;注意力机制;K-means

中图分类号: TP751 **文献标识码:** A **国家标准学科分类代码:** 510.4

A SAR ship detection algorithm based on improved YOLOv4

Chen Yang¹ Zhang Ming^{1,2} Yang Lidong¹ Yu Dahua¹ Zhang Baohua¹ Li Jianjun¹

(1. School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China;

2. Information Science and Technology College, Dalian Maritime University, Dalian 116026, China)

Abstract: Under the background of large amount of data support, how to use a large number of SAR images efficiently and improve the accuracy of ship target detection is the current problem of ship target detection. This paper focuses on how to improve the accuracy of YOLOv4 algorithm for SAR ship target detection, and presents a YOLOv4 enhancement algorithm that combines multiscale and attention enhancement. The Attention Module (CBAM) is added to the PANet of the original YOLOv4, and the enhanced K-means clustering algorithm is used to cluster the ship target real frame in the dataset, and the result of the anchor frame is transformed linearly to make the algorithm anchor frame more suitable for the training set. Experiments show that the average accuracy (mAP) of the proposed algorithm in SAR ship detection is 94.05%, which is 0.7% higher than that of the original YOLOv4. The experimental results fully demonstrate that the proposed algorithm can improve the accuracy of SAR ship image detection and provide technical support for the accuracy of sea activities judgment.

Keywords: SAR image; ship detection; deep learning; attention mechanism; K-means

0 引 言

合成孔径雷达(synthetic aperture radar, SAR)^[1]是一种主动式微波遥感成像传感器,是海上目标探测的重要手段之一。因其工作在微波波段,相较于其他遥感技术^[2],SAR 具有穿透特性,能够在能见度极低的天气条件下成像,具有对目标实施全天候、多角度、大尺度、远距离持续监测的能力^{-3-7]},基于以上优势,SAR 舰船图像更加适用于舰船目标检测与识别的研究。

传统的舰船目标检测方法^[8-9]主要是针对某类场景所

对应的相关特征进行建模^[10]和基于显著性的目标检测算法^[11],这类算法大多都依赖于特定目标场景的人工设计,泛化能力较差,鲁棒性不强。随着深度学习技术的不断发展,卷积神经网络(convolutional neural networks, CNN)^[12]在舰船目标检测方面的优势也不断凸显。随着 SAR 舰船图像的日益增加,基于深度学习的 SAR 目标检测技术越来越完善,检测精度不断攀升。当前,主流的深度学习方法目标检测技术大体上可分为两大类:一类是以 FasterR-CNN^[13]为代表的两阶段网络算法,两阶段网络大多有较好的检测精度,但相应的成本时间和硬件消耗都比

收稿日期:2022-01-19

* 基金项目:国家自然科学基金(62161040,61771266,62066036,61962046)项目资助

较大,并且由于网络结构复杂,进行实时性检测存在较大难度。因此,单阶段网络应运而生。这类算法以 SSD^[14] (single shot multibox detector, SSD) 和 YOLO (you only look once, YOLO)^[15-17] 系列算法为代表。该类算法无需进行候选框的选定,直接通过回归的方式对目标进行检测。YOLOv4^[18] 兼顾网络的精度和效率,该算法融合了多种结构,如 SPP、PANet、CSP 等,在保证效率的情况下相较于 yolov3 在精度上进一步提升。虽然当前单阶段目标检测算法性能有一定的提升,但仍有较大的提升空间,为了进一步提升单阶段目标检测算法的检测精度,本文从两个维度对 YOLOv4 目标检测算法进行改进。首先我们使用了改进的 K-means 聚类算法对 YOLOv4 的锚框进行了调整,让算法的锚框更适配于数据集,其次在 YOLOv4 中集成注意力

机制,增加算法对舰船目标的权重分配,强化目标检测能力,提高抗干扰能力。实验结果证明本文的增强算法在 SAR 舰船检测中的平均准确率(mAP)达到了 94.05%,比原始 YOLOv4 算法提升了 0.7%。为海上目标侦察检测与识别提供更精准的基础技术条件。

1 方 法

针对原始 YOLOv4 算法中的锚框参数对检测 SAR 舰船目标尺寸存在不适应和对高背景噪声情况下对小目标存在误检、漏检等问题,本文分别通过使用改进的 K-means 聚类算法和加入注意力模块(CBAM)来提升算法精度,结构图如图 1 所示,接下来将分别介绍上述两部分内容。

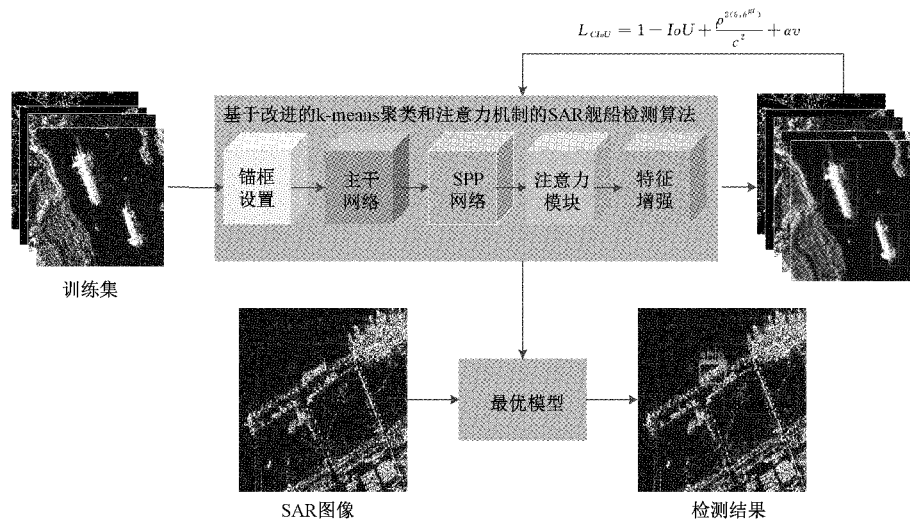


图 1 算法整体流程示意图

1.1 改进的 K-means 聚类算法

传统的 YOLOv4 算法的初始锚框尺寸设置并不适合于检测 SAR 舰船目标,为了使算法中的锚框更加接近数据集中目标的真实框尺寸,提升算法在回归过程中的速度与精确度,本文尝试改进了 K-means 聚类算法。起初,使用 K-means 聚类算法对数据集中的目标真实框的大小进行聚类,并使用聚类结果框的参数进行的实验,结果并不理想。实验后分析原因,推测是因为 SAR 舰船数据集中舰船目标的真实框尺寸相近,导致聚类出的锚框尺寸很集中,进而影响了算法的多尺度训练和输出。为了解决这一问题,在原始 K-means 聚类算法的基础上,对聚类结果进行线性放缩,提高聚类结果尺寸的丰富度。其实现过程如下:

1) 给定 m 个聚类中心点坐标 $(W_i, H_i), i \in \{1, 2, \dots, m\}$, W_i, H_i 表示锚框的宽和高。

2) 根据数据集的标签中的目标位置信息 $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$ 。

$$x_{\max} - x_{\min} = w_j \quad (1)$$

$$y_{\max} - y_{\min} = h_j \quad (2)$$

$$x_{\min} + \frac{w_j}{2} = x_j \quad (3)$$

$$y_{\min} + \frac{h_j}{2} = y_j \quad (4)$$

其中, $(x_j, y_j, w_j, h_j), j \in \{1, 2, \dots, n\}$ 其中 (x_j, y_j) 为真实框的中心点坐标, (w_j, h_j) 为真实框的宽和高, n 为真实框的数量。

3) 计算数据集中的目标真实框与 m 个聚类中心点的距离 d :

$$d = 1 - IOU[(x_j, y_j, w_j, h_j), (x_j, y_j, W_i, H_i)] \quad (5)$$

将真实框分配到离其最近的 (W_i, H_i) 点。

4) 对每一个分配完成的簇计算该簇中的真实框的平均值:

$$W'_i = \frac{1}{N_i} \sum w_i \quad (6)$$

$$H'_i = \frac{1}{N_i} \sum h_i \quad (7)$$

其中, N_i 表示第 i 个簇中真实框的个数。

5) 重复步骤 3)、4), 直至 (W'_i, H'_i) 几乎不再变化。

6) 对步骤 5) 的结果进行线性比例伸缩:

$$x'_1 = \alpha x_1, x'_9 = \beta x_9, y'_1 = \alpha y_1, y'_9 = \beta y_9 \quad (8)$$

$$x'_i = \frac{(x_i - x_1)}{(x_9 - x_1)}(x'_9 - x'_1) + x'_1 \quad (9)$$

$$y'_i = \frac{(y_i - y_1)}{(y_9 - y_1)}(y'_9 - y'_1) + y'_1 \quad (10)$$

α, β 分别为缩小比例和放大比例。

通过上述流程, 得到了聚类过后的 anchor。实验证明, 改进过后的 anchor 对算法精度有一定的提升, 且训练时间无明显增加。

1.2 注意力模块

在 SAR 成像的过程中, 噪声的掺杂是不可避免的, 这些噪声会影响算法对目标的检测。为了抑制背景噪声对算法的影响, 突出所要检测目标的特征, 我们在 PANet 中集成注意力模块(convolutional block attention module, CBAM), 如图 2 所示。

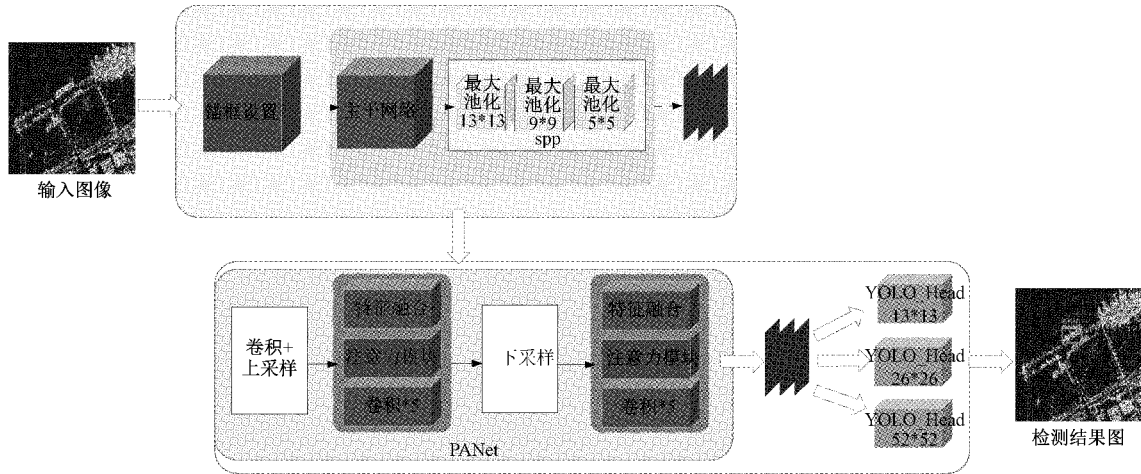


图 2 网络整体结构图

注意力机制是模拟人类视觉的大脑信号处理机制。通过快速扫描全局图像, 获得需要重点关注的目标区域, 然后对这一区域投入更多注意力资源, 以获取所需要关注的目标细节信息, 从而抑制其他无用信息。其本质是通过网络自主学习出的一组权重系数, 并以“动态加权”的方式来强调感兴趣的区域同时抑制不相关区域的机制。

CBAM(convolutional block attention module) 是一种结合了通道(channel)和空间(spatial)的注意力机制模块, 通过两部分的顺序操作对特征进行自适应细化, 以达到增强目标特征表达的目的。

其中, 通道注意力的作用是告诉我们目标“是什么”, 通过一系列的最大池化和平均池化, 给目标赋予较大的权重, 背景信息赋予较小的权重, 最后通过激活函数, 从而强调目标“是什么”。模型结构如图 3 所示。

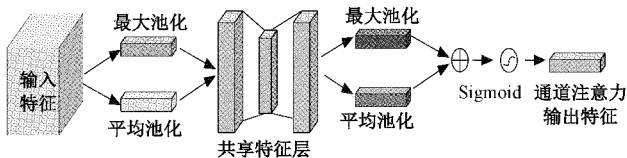


图 3 通道注意力模型结构图

空间注意力是着重于目标“在哪里”, 同样也是使用平均池化和最大池化, 强调目标的位置信息, 帮助网络确定目标在特征图中的位置, 提高算法在强噪声条件下目标检测的准确性。模型结构如图 4 所示。

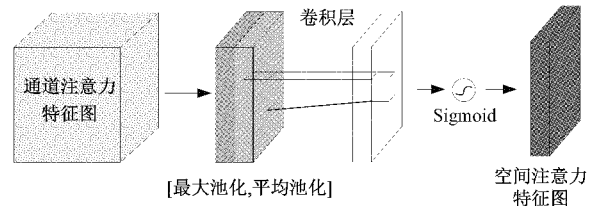


图 4 空间注意力模型结构图

将经过通道注意力和空间注意力后得到的特征图与原特征图进行融合作为下一层的输入。

1.3 损失函数

本文网络使用了 CIOU 损失函数, 该损失函数如下所示:

$$L_{CIOU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (11)$$

其中, IoU 为预测框和真实框的交集和并集的比值, $\rho(\cdot)$ 为欧氏距离, b 和 b^{gt} 表示 B 和 B^{gt} 的中心点, c 为真实框与预测框所形成的最小封闭框的对角线长度, α 为权重函数, v 为度量纵横比的一致性的参数。

IoU 表达如下:

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (12)$$

其中, $B^{gt} = (x^{gt}, y^{gt}, w^{gt}, h^{gt})$ 表示真实框, $B = (x, y, w, h)$ 表示预测框。

v 表示如下:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{st}}{h^{st}} - \arctan \frac{w}{h} \right)^2 \quad (13)$$

α 表示如下:

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (14)$$

2 实验与分析

本文中,所有实验都在 Tensorflow 框架中实现,并在具有 4 * NVIDIA Tesla V100 GPU 卡、128 GB 内存的 NVIDIA DGX 工作站服务器上进行。操作系统是带有 CUDA 10.1 的 Linux。权重衰减和动量分别为 0.000 1 和 0.9。训练回合数为 80,批量大小为 32,IoU 为 0.5。

2.1 数据集

本文实验的数据集为中科院空天信息研究院数字地球重点实验室王超研究员团队于 2019 年 5 月公开的 SAR 图像船舶检测数据集。该数据集以我国国产高分三号 SAR 数据和 Sentinel-1 SAR 数据为主数据源,共采用了 102 景高分三号和 108 景 Sentinel-1 SAR 图像构建高分辨率 SAR 船舶目标深度学习样本库。目前,该深度学习样本库包含 43 550 船舶切片,切片大小为 256 * 256。本次实验训练集,验证集,测试集比例划分为 8 : 1 : 1。表 1 为数据集配置信息,图 5 为部分数据集样本示例。

表 1 数据集信息

数据集	训练集	测试集	验证集	总计
SAR-Ship-Dataset	34 840	4 355	4 355	43 550

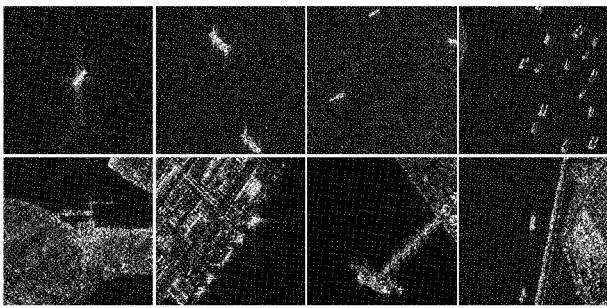


图 5 数据集示例图

2.2 评价指标

为了能更好的验证训练模型的有效性,本文从多个角度对所训练模型进行评价,分别从检测率 p_d , 虚警率 p_f 和平均准确率(mean average precision, mAP) 和每秒帧率(frame per second, FPS) 4 个指标对模型进行评价。4 个指标的公式分别为:

$$p_d = TP / GT \quad (15)$$

$$p_f = FP / (TP + FP) \quad (16)$$

$$mAP = \int_0^1 P(R) dR \quad (17)$$

其中, TP (true positive) 为实际中为正样本,被预测为正样本。 FP (false positive) 为实际中为负样本,被预测为正样本, GT (ground truth) 为实际中真实样本数。

$$FPS = 1/T \quad (18)$$

其中, T 为一幅图片检测的时间。

2.3 实验对比

首先,为了证明本文所提出方法的有效性,我们在相同数据集上进行了一组消融实验,实验内容如表 2 所示。实验结果显示使用原始 YOLOv4 算法对 SAR 舰船图像进行舰船目标检测时,最终的 mAP 值为 93.35%。在原始方法的基础上,我们使用了改进的锚框后, mAP 值提升了 0.47%,同时,虚警率也下降了 2.58%。在此基础上我们融入 CBAM 模块,利用 CBAM 模块对舰船特征的加权和抑制复杂背景噪声的优势,提升算法对 SAR 舰船的检测性能。实验结果也印证了 CBAM 的效果, mAP 值提升至 94.05%,虚警率也下降 2.55%,3 组实验中,检测率无明显的变化。实验结果表明,改进的 K-means 聚类算法配合 CBAM 模块能够提高 YOLO 算法对 SAR 舰船图像的目标检测性能。

表 2 基于 YOLOv4 的舰船算法检测性能

网络模型	改进 K-means	CBAM	检测率/ %	虚警率/ %	平均 准确率/%
	—	—	96.72	40.56	93.35
YOLOv4	✓	—	96.77	37.98	93.82
	✓	✓	96.70	35.43	94.05

除了客观评价之外,视觉检查也是一种有效的主观结果评价方法。因此,在实验中我们从数据集中随机选择几张图像测试模型的性能,实验结果如图 6 所示,蓝色框表示真实框,绿色框表示检测到的舰船目标,红色框表示错误检测目标。YOLOv4 与本文算法在 SAR-Ship-Dataset 中的比较结果。第 1 行为 YOLOv4 检测结果,第 2 行为改进 K-means + YOLOv4 算法检测结果,第 3 行为本文算

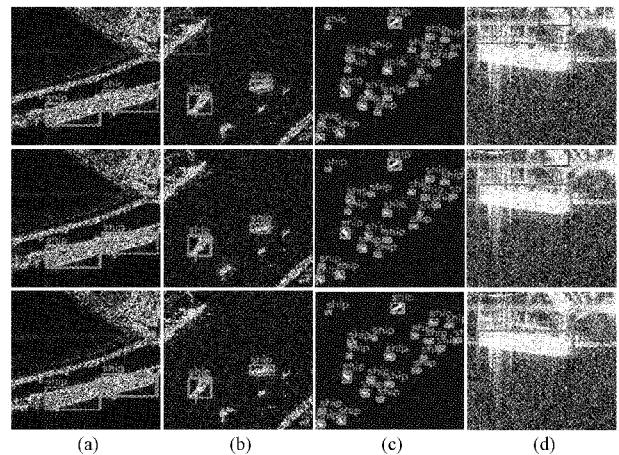


图 6 检测结果对比图

法。通过检测结果图可以看到,在近海环境,远海多目标和强噪声的情况下,原始 YOLOv4 的检测结果中都出现了较多的虚警目标,而本文方法与原始方法相比,具有相对更好的检测性能和检测效果。

最后,将本文的改进算法与主流的目标检测算法在 SAR 舰船图像下的检测结果进行对比,结果如表 3 所示。

表 3 不同算法对 SAR 舰船目标的检测性能对比

网络模型	输入图像大小	mAP/%
SSD-300 ^[13]	300×300	88.3
SSD-512 ^[13]	512×512	89.4
Faster-RCNN ^[12]	600×600	88.3
RetinaNet ^[19]	800×800	91.4
YOLOv3 ^[17]	256×256	90.6
YOLOv4 ^[18]	256×256	93.4
Ours	256×256	94.1

从实验结果中可以看出,本文提出的改进算法的 mAP 值高于原始 YOLO 系列等单阶段网络算法和 Faster-RCNN 等两阶段检测网络算法,究其原因,由于使用改进的 K-means 聚类算法对 SAR 舰船图像数据集聚类出来的锚框更适配于检测 SAR 舰船目标,并且配合 CBAM 模块,对舰船目标赋予较大权重,抑制复杂背景对舰船目标检测的干扰。同时我们也进行了与原始 YOLOv4 网络的 FPS 比较,结果表明两个网络模型的 FPS 均为 59。由此可知,本文所提出的改进算法在没有显著性检测速度损失的基础上,得到了更好的检测精度。

通过主客观评价对比,本文所提出算法在抑制背景噪声干扰和适应目标多尺度变化方面比原算法有一定增强,减少了舰船目标漏检率和误检率,整个模型的检测性能有一定的提升。因此,本文所提方法在速度几乎没有损失的基础上,还能实现比原算法更好的检测性能,具有较高的实用价值,为后续开发海上舰船目标检测系统奠定基础。

3 结 论

随着 SAR 成像技术的发展,越来越多的 SAR 舰船数据可应用于舰船目标检测方面的研究。为了有效提升 SAR 舰船检测算法的性能,本文基于 YOLOv4 提出了一种融合多尺度和注意力机制的 YOLOv4 舰船检测算法,使用改进的聚类算法对舰船数据集进行聚类,提升算法锚框的丰富度,同时结合注意力机制进一步提升算法检测性能。实验结果表明,相较于原始 YOLOv4,舰船目标的检测速度无明显变化,舰船检测平均准确率提升了 0.7%,达到了 94.05%。同时,通过大量实验结果可以看出,我们所有提出的模型与当前主流的目标检测模型相比,本文模型具有较好的检测性能。该方法能够有效检测各种海况下的舰船目标检测任务,在国家海洋强国建设的背景下,具

有较高的研究意义,为未来各海况下目标检测提供技术支撑。

参考文献

- [1] 邓云凯,禹卫东,张衡,等. 未来星载 SAR 技术发展趋势[J]. 雷达学报, 2020, 9(1):1-33.
- [2] 吴东,郝明. 基于图像对比度的舰船目标成像算法[J]. 电子测量技术, 2017, 40(12):110-116.
- [3] EIDHUSET K. An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions[J]. IEEE Transactions on Geoscience Remote Sensing, 1996, 34(4):1010-1019.
- [4] CHEN S Z, WANG H P, XU F, et al. Target classification using the deep convolutional networks for SAR images[J]. IEEE Transactions on Geoscience Remote Sensing, 2016, 54(8):4806-4817.
- [5] STASOLLA M, MALLORQUÍ J J, MARGARIT G, et al. A comparative study of operational vessel detectors for maritime surveillance using satellite-borne synthetic aperture radar[J]. IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing, 2019, 9(6):2687-2701.
- [6] WANG Y, WANG C, ZHANG H, et al. A SAR dataset of ship detection for deep learning under complex backgrounds[J]. Remote Sensing, 2019, 11(7):765, DOI:10.3390/rs11070765.
- [7] ZHANG T, ZHANG X. High-speed ship detection in SAR images based on a grid convolutional neural network[J]. Remote Sensing, 2019, 11(10):1206, DOI:10.3390/rs11101206.
- [8] GUI G. A parzen-window-kernel-based CFAR algorithm for ship detection in SAR images[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(3):557-561.
- [9] HOU B, CHEN X, JIAO L. Multilayer CFAR detection of ship targets in very high resolution SAR images[J]. IEEE Geoscience Remote Sensing Letters, 2014, 12(4):811-815.
- [10] HUI D, LAN D, YAN W, et al. A modified CFAR algorithm based on object proposals for ship target detection in SAR images[J]. IEEE Geoscience Remote Sensing Letters, 2016, PP(99):1-5.
- [11] TAO D, DING G, ANFINSEN S N, et al. Robust CFAR detector based on truncated statistics in multiple-target situations[J]. IEEE Transactions on Geoscience Remote Sensing, 2016, DOI: 10.1109/TGRS.2015.2451311.
- [12] 周晓彦,王珂,李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术, 2017, 40(11):89-93.

- [13] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2017, 39 (6): 1137-1149.
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. Springer, Cham, 2016:21-37.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [16] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [17] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv Preprint, 2018, ArXiv. 1804. 02767.
- [18] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv. 2004. 10934.
- [19] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99): 2980-2988.

作者简介

张明,副教授,主要研究方向为遥感图像处理,模式识别和深度学习等。

E-mail:nkd_zm@imust.edu.cn

陈洋,硕士研究生,主要从事遥感图像处理工作。

E-mail:cyang980323@163.com