

基于视觉显著性的SAR遥感图像NanoDet舰船检测方法

刘方坚^{*①} 李 媛^②

^①(中国科学院空天信息创新研究院 北京 100190)

^②(北京理工大学机电学院 北京 100081)

摘 要: 在合成孔径雷达遥感图像中, 舰船由金属材质构成, 后向散射强; 海面平滑, 后向散射弱, 因此舰船是海面背景下的视觉显著目标。然而, SAR遥感影像幅宽大、海面背景复杂, 且不同舰船目标特征差异大, 导致舰船快速准确检测困难。为此, 该文提出一种基于视觉显著性的SAR遥感图像NanoDet舰船检测方法。该方法首先通过自动聚类算法划分图像样本为不同场景类别; 其次, 针对不同场景下的图像进行差异化的显著性检测; 最后, 使用优化后的轻量化网络模型NanoDet对加入显著性图的训练样本进行特征学习, 使系统模型能够实现快速和高精确度的舰船检测效果。该方法对SAR图像应用实时性具有一定的帮助, 且其轻量化模型利于未来实现硬件移植。该文利用公开数据集SSDD和AIR-SARship-2.0进行实验验证, 体现了该算法的有效性。

关键词: SAR图像; 舰船检测; 深度学习; 轻量化网络; 视觉显著性

中图分类号: TN957.52

文献标识码: A

文章编号: 2095-283X(2021)06-0885-10

DOI: 10.12000/JR21105

引用格式: 刘方坚, 李媛. 基于视觉显著性的SAR遥感图像NanoDet舰船检测方法[J]. 雷达学报, 2021, 10(6): 885–894. doi: 10.12000/JR21105.

Reference format: LIU Fangjian and LI Yuan. SAR remote sensing image ship detection method NanoDet based on visual saliency[J]. *Journal of Radars*, 2021, 10(6): 885–894. doi: 10.12000/JR21105.

SAR Remote Sensing Image Ship Detection Method NanoDet Based on Visual Saliency

LIU Fangjian^{*①} LI Yuan^②

^①(*Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China*)

^②(*School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China*)

Abstract: In the Synthetic Aperture Radar (SAR) remote sensing image, ships are visually significant targets on the sea surface. Because they are made of metal, thus the backscatter is strong, while the sea surface is smooth and the backscatter is weak. However, the large-width SAR remote sensing image has a complicated sea background, and the features of various ship targets are quite different. To solve this problem, a SAR remote sensing image ship detection model called NanoDet is proposed. NanoDet is based on visual saliency. First, the image samples are divided into various scene categories using an automatic clustering algorithm. Second, differentiated saliency detection is performed for images in various scenes. Finally, the optimized lightweight network model, NanoDet, is used to perform feature learning on the training samples added with the saliency maps, so that the system model can achieve fast and high-precision ship detection effects. This method is helpful for the real-time application of SAR images. The lightweight model is conducive to hardware transplantation in the future. This study conducts experiments based on the public data set SSDD and AIR-SARship-2.0, and the experiments results verify the effectiveness of our approach.

Key words: SAR image; Ship detection; Deep learning; Lightweight network; Visual saliency

收稿日期: 2021-07-22; 改回日期: 2021-09-18; 网络出版: 2021-10-19

*通信作者: 刘方坚 liufj@aircas.ac.cn

*Corresponding Author: LIU Fangjian, liufj@aircas.ac.cn

基金项目: 国家自然科学基金(61972021, 61672076)

Foundation Items: The National Natural Science Foundation of China (61972021, 61672076)

责任编辑: 孙显 Corresponding Editor: SUN Xian

1 引言

目前,合成孔径雷达(Synthetic Aperture Radar, SAR)图像受到越来越广泛的关注。因其具有全天时、全天候的工作特性,且不易受天气变化影响,特别对于变幻莫测的海洋场景具有良好的适应性^[1,2]。其应用范围广泛,在军事应用、海洋交通管理、油污勘探以及渔业管理等领域均发挥出良好作用^[3]。

然而,基于SAR图像的舰船目标检测依然面临很多困难。例如,海浪和杂波产生出较强的回波信号^[4],特别对于小型船只存在较大的干扰问题。并且由于SAR是相干系统,具有斑点噪声的固有特性。因此,如何减少干扰因素影响是提升检测精度的关键。同时,随着技术发展日益成熟,对图像处理任务的实时性要求也日渐增高^[5]。舰船检测一般包括图像预处理、特征提取和模型预测3个主要步骤^[6]。在研究早期阶段,大多数方法集中于提升网络的特征提取能力,旨在通过获取更多且更精细的特征提升网络对目标的检测性能。但不可忽略的是,高质量的图像是获取良好处理结果的基础。因此,图像预处理操作依然十分重要。

由于遥感图像具有复杂的空间结构和多样化场景,因此不同图像需要不同的预处理操作^[7,8],如阈值分割、图像增强和聚类处理等。而对于舰船目标检测任务,海面背景下的舰船目标具有显著特性。因此,降低背景复杂度,提升目标显著性能够获得更优质的图像样本,进而能够获取更准确的检测结果。2007年,侯小迪等人^[9]提出了谱残余(Spectral Residual, SR)显著性检测算法,通过将输入图像转换的频谱域进行残差计算,从而获取最大变化区域为显著性区域。该方法实现简单,检测速度快且效果良好。在文献^[10]中,作者设计了一种上下文感知显著性探测器(Context-Aware Saliency Detector, CASD)。其设计思想主要考虑到对象上下文特征与对象本身同样重要。即图像中的某一特征既要与邻域特征存在区别,也要与全局特征存在区别。该方法对于背景较为复杂的场景具有较好的显著性检测效果。同时,Liu等人^[11]提出一种新颖的显著性检测方法——显著树。该方法能够对区域和像素级显著性图进行分层表示,生成高质量的显著性图。各类方法有各自的优势,均能够对增强目标特征提供一定的帮助。

近些年,卷积神经网络(Convolutional Neural Network, CNN)以其强大的自主学习能力和特征捕获能力得到了越来越广泛的关注^[12]。现有基于神经网络的目标检测方法一般可以分为一阶段法和两阶

段法^[13]。其中,两阶段法由两个级联网络构成,前者用于生成目标候选框,后者判断是否存在目标。2014年,基于区域的卷积神经网络方法^[13](Region-based CNN, R-CNN)被提出,首次将深度学习方法引入目标检测中。相较于Hog特征结合SVM分类器等典型方法,其结果有较大精度提升。空间金字塔池化网络^[14](Spatial Pyramid Pooling Net, SPPNet)引入了金字塔池化层,以使网络能够适应任意尺寸图像输入。此后,涌现出许多有效改进算法如Fast RCNN^[15], Faster RCNN^[16], Mask RCNN^[17]以及PANet^[18]等,加快了检测速度,并提升了检测精度。2020年,CBNet^[19]网络架构被提出,它通过复合连接多个骨干网络形成更加强大的复合主干网络。实验表明,将其集成到Mask RCNN等网络中能够显著提高检测性能。

此外,Redmon等人^[20]于2016年提出首个一阶段目标检测算法YOLO(You Only Look Once)。它通过将检测问题转化为回归分类问题,能够有效避免生成候选区域的耗时,大大提升了检测速度。但同时,放弃候选区域选取也使检测精确性在一定程度下降。因此,针对该问题,多种优化算法被提出。其中,SSD^[21]算法通过提取不同尺度特征图以及采用不同尺度的先验框等改进方法,实现了更高的目标检测性能。同时,改进的YOLOv2—YOLOv4^[22-24]系列算法也使目标检测方法进一步得到发展。除此之外,2020年Ultralytics等人在YOLOv4的基础上提出具有更高速度优势的YOLOv5,暂无论文发表。实验证明该方法具有与YOLOv4相当的检测效果,其检测速度令人惊艳。

在实际应用过程中,不仅需要高检测准确度还需要绝对的实时性。因此,轻量化网络模型成为一大研究热点。2016年,SqueezeNet^[25]小型CNN架构模型被提出,通过压缩再扩展模块,有效降低卷积计算量。MobileNet系列将原卷积操作优化为深度可分离卷积,提升了卷积计算效率。ShuffleNet系列则更进一步,提出了通道混洗方法,避免了大量冗余计算。在此基础上,NanoDet^[1]通过集成多种优秀轻量化方法得到一个兼顾精度、速度和体积的模型,能够满足应用实时性要求。

综上所述,本文提出一种基于NanoDet和显著性的SAR图像舰船检测方法,如图1所示。针对复杂背景干扰和检测耗时问题,提出相应的改进方法。首先,依据不同场景图像灰度直方图特征,通过自动聚类的方式划分出不同类别,以便可以通过不同操作实现不同类别的针对性预处理,降低方法泛化性要求,以获得更加精确的处理结果。其次,

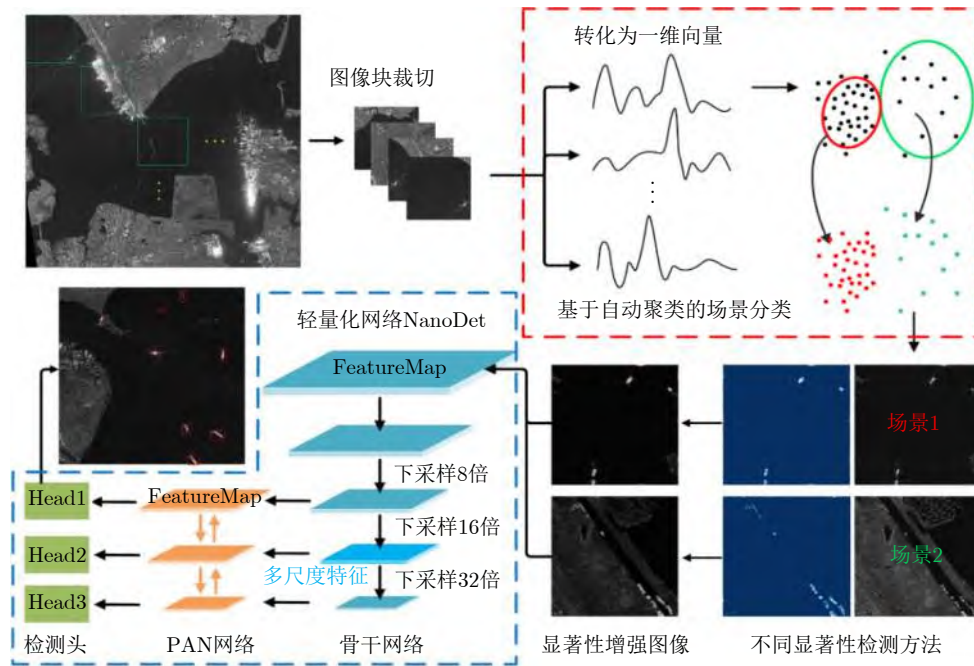


图1 系统整体流程示意图

Fig. 1 Schematic diagram of the overall system process

使用SR和CASD不同显著性区域检测方法获取显著性图像，并与原图叠加，以增加目标显著性。最后，将处理后的图像块送入调整后的NanoDet模型进行网络训练，以获取高精度度、近实时的目标检测模型。

2 预处理方法实现

本文的预处理模块主要分为两部分，依据各图像场景特征，完成基于自动聚类的图像场景类别划分；利用不同显著性检测方法检测不同类别场景显著性区域。通过对图像样本进行预处理，实现各场景类别针对性地显著目标增强，提高了用于舰船检测任务的图像样本质量。

2.1 基于密度峰值聚类的场景划分

自动聚类是一种可用于数据类别划分的简单、快速、有效的方式。对于SAR图像，其成像原理使其具有不受云雾干扰的特性，因此其主要场景分为海上场景和近岸场景。同时，SAR图像灰度信息直接反映了不同物体后向散射强度信息，通常来说表面越粗糙的物体反射强度越大，则图像中灰度值越高，颜色越亮。海上场景整体灰度值较低，背景灰暗，目标突出；近岸场景因存在较大范围地物，灰度值高且存在地物纹理特征。因此，将图像灰度值转化为一维向量并将其作为聚类元素。通过自动聚类算法，依据不同场景下SAR图像灰度特征的不同，完成场景类别的划分。在本文中，聚类处理之前通常将图像裁切成 512×512 大小，从而转化成相

同维数的一维向量完成类别划分。实现过程如图1红色虚线框所示。

假设图像数据块转化的一维向量表示为 $\mathbf{x} = [x_0, x_1, \dots, x_{n-1}]^T$ ，其中 n 表示图像块大小。首先通过曼哈顿距离^[26]等距离计算公式完成向量 \mathbf{x}_i 和 \mathbf{x}_j 的相似度 s_{ij} 计算，表示为

$$d_{ij} = \sqrt{|\mathbf{x}_i - \mathbf{x}_j|' \sum^{-1} |\mathbf{x}_i - \mathbf{x}_j|} \quad (1)$$

$$s_{ij} = d_{ij}^{-1} \quad (2)$$

再依据以上结果计算两个关键因子：各聚类元素的局部密度 ρ 和到聚类中心距离 δ 。其中，第 i 个数据点的局部密度 ρ_i 为

$$\rho_i = \sum_j \varphi(s_{ij}^{-1} - c) \quad (3)$$

其中，当 $m < 0$ 时， $\varphi(m) = 1$ ；当 $m = 0$ 时， $\varphi(m) = 0$ ， c 是相似性截止阈值。局部密度 ρ_i 被视为与点 \mathbf{x}_i 相似的，在 c 相似度范围内点的数量。而到聚类中心距离 δ_i 表示为点 \mathbf{x}_i 与其他密度更高点 \mathbf{x}_j 之间的最小距离。

$$\delta_i = \min_{j: \rho_j > \rho_i} (s_{ij}^{-1}) \quad (4)$$

循环计算所有样本点向量的局部密度和距离构造决策图，越靠近右上角的孤立的点被认为是聚类中心点。计算过程中也可依据密度与距离的乘积大于设定阈值 $\rho * \delta > k$ 的判定条件完成聚类中心查找。其

中, k 值仅用于划分出决策图中的独立中心点, 在某范围内变化不会对结果产生影响。最后, 通过计算各样本点到各聚类中心的距离远近, 完成类别划分。

2.2 不同场景下的显著性检测方法

不同场景图像包含不同特征, 因此需要针对性的方法以达到良好显著性检测结果。对于海上场景图像, 基于谱残差的显著性检测方法容易使大型船只出现边缘显著性, 而非整个舰船区域的显著表示。而由于近岸场景存在复杂的地物纹理信息, 上下文感知显著性检测器难以获取最佳显著性区域。因此, 本文选择上下文感知方法用于海上场景, 而选用谱残差方法用于近岸场景。

谱残差方法主要在频域完成计算。首先, 将图像通过傅里叶变换映射到频域中, 并提取其幅度和相位响应。

$$a = A(\mathfrak{F}(x)) \quad (5)$$

$$p = P(\mathfrak{F}(x)) \quad (6)$$

其中, \mathfrak{F} 表示傅里叶变换, A 和 P 分别表示所获取的幅度谱和相位谱。然后对其做 \log 变换, 得到 \log 频谱。并计算两者之差, 即为频域下的显著性目标。

$$l = \log(a) \quad (7)$$

$$r = l - h \times l \quad (8)$$

其中, h 表示均值滤波器, r 表示所得残差谱。最后, 将残差谱 r 和相位谱 p 通过逆傅里叶变换转化回时域, 再通过非线性滤波器得到显著性图。

$$s = h \times \mathfrak{F}^{-1}[\exp(r + p)]^2 \quad (9)$$

上下文感知方法不同于其他显著性检测方法, 能够检测出代表场景的显著性区域。它主要依据4个原则完成显著性检测: 局部低层次特征、全局特征、视觉特征和高层次特征。首先, 将图像分成多个子块, 将目标块 p_i 与Lab颜色空间中其他块 p_j 进行比较。若 p_i 块与其他块的差异越大, 显著性越明显。计算公式为

$$d(p_i, p_j) = \frac{d_c(p_i, p_j)}{1 + c \cdot d_p(p_i, p_j)} \quad (10)$$

其中, d_c 是欧式颜色距离, d_p 是欧式空间距离, c 为常数参量。因此, 可以获得与 p_i 距离最小的前 N (通常取 N 为65)个图像块。且通过计算其与图像块 p_i 的差异, 显著性公式定义为

$$s_i = 1 - \exp\left\{-\frac{1}{n} \sum_{n=1}^N d(p_i, q_n)\right\} \quad (11)$$

另一方面, N 能够调节显著性区域范围, 适当调整

能够有效提升检测效果。同时, 多尺度结果平均值能够增强显著性图像对比度。使 $R = \{r_1, r_2, r_3, \dots, r_M\}$ 表示图像块 p_i 所需尺度集合。则多尺度显著性结果增强方法可表示为

$$\bar{s}_i = \frac{1}{M} \sum_{r \in R} s_i^r \quad (12)$$

此外, 还需对以上结果进行上下文校正。设置一个阈值, 若某像素显著性超过该阈值, 则认为该像素属于显著性区域。显著性区域外像素的显著性值则通过计算其与最近显著区像素的加权欧氏距离获得, 表示为

$$\hat{s}_i = \bar{s}_i (1 - d_f(i)) \quad (13)$$

其中, $d_f(i)$ 为像素 i 与最近显著像素之间的欧式距离。该步能增加显著目标邻域显著值, 降低背景区域显著值, 从而提升显著性检测效果。

3 优化的NanoDet模型

NanoDet模型是一个开源的实时检测模型, 它能够提供更不亚于YOLO系列目标检测算法的检测性能, 实现更快处理。同时, 其网络简单, 模型参数少, 更方便训练和模型移植。

3.1 标准NanoDet网络结构

该模型基于一阶段的检测算法完成模型搭建, 主要包含3部分: 骨干网络backbone, Neck模块和检测头Head, 如图1蓝色虚线框所示。骨干网络为ShuffleNetV2 1.0x, 并删除最后卷积层, 抽取8, 16和32倍下采样特征作为下层输入。Neck模块由优化后的极小PAN构成, 只包含自上而下和自下而上两条通路。其中, 仅包含由骨干网络特征提取后的 1×1 卷积完成通道维度对齐, 采用插值算法完成上、下采样, 并通过相加实现多尺度特征图融合, 如图2所示。

检测头选取FCOS系列检测头并进行优化, 放弃共享权重, 依然使用不同卷积组完成各层特征提取。同时, 将原有的组归一化(Group Normalization)

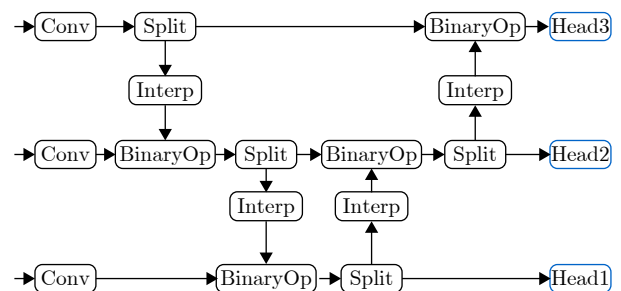


图2 极小PAN结构

Fig. 2 Structure diagram of minimize PAN

方式改回批量归一化(Batch Normalization), 选择深度可分离卷积替换普通卷积, 并减少了回归分类过程中卷积核数量和卷积通道数。最大程度降低计算复杂度, 实现轻量化网络模型, 其结构如图3所示。

最后, 使用GFocalloss去点FCOS中的中心分支, 解决了其在轻量化模型上难以收敛的问题, 并省去该部分卷积计算开销。本文在此基础上对NanoDet进行了参数调整, 增加了数据增强方法, 使检测准确性进一步提升。这些优化策略使得NanoDet不仅能够实现良好的检测效果, 还能够有效减少模型体积, 节约运行时间。

3.2 NanoDet网络优化

由于实际应用场景中往往存在不同尺寸目标, 特别是小尺寸目标容易受噪声和其他地物干扰, 从而使得检测模型容易出现误差。针对该问题, 本文对NanoDet网络模型进行了优化调整。借鉴了TridentNet^[27]的三叉戟结构, 将NanoDet中下采样第4层特征图通过多尺度空洞卷积实现。空洞卷积计算空洞卷积计算时在卷积核元素中间加入不同数量的空洞, 空洞率不同卷积核的尺寸也就不同。因此, 加入空洞卷积能够扩大网络的感受野, 捕获多尺度的上下文信息, 且不需要引入额外的参数。3个分支其他结构相同, 差别在于卷积空洞率不同, 其结构如图4所示。不同的卷积操作使得网络能够捕获不同尺度的特征信息, 因此能够更精确的检测不同尺度大小的目标。

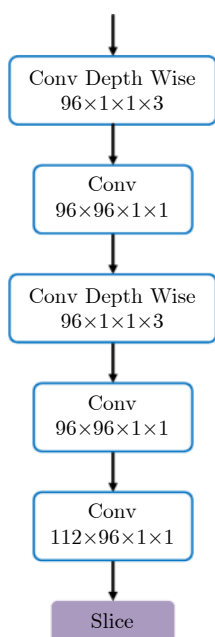


图3 FCOS检测头结构图

Fig. 3 Structure diagram of FCOS detection head

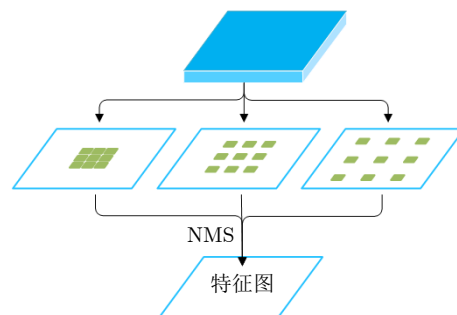


图4 多尺度特征提取示意图

Fig. 4 Schematic diagram of multi-scale feature extraction

4 实验验证

本文实验硬件平台为Intel®Core™ i7-97 CPU @3.60 GHz×8, 64 G内存计算机, 使用GeForce RTX 2080Ti GPU高性能图形处理器完成加速运算。软件环境为Linux平台, 操作系统为Ubuntu18.04.5 LTS。软件环境基于Anaconda平台安装使用, 并基于深度学习框架Pytorch^[28]构建网络模型, 使用Python完成程序编写。同时, 使用CUDA 10软件包实现程序加速。在本实验中, 设置检测阈值IoU为0.5, 即当检测框与真值框重叠区域超过两者并集区域的50%时, 认为该框检测正确。

4.1 实验数据

本文实验基于国内公开SAR数据集SSDD和AIR-SARship-2.0实现。其中, SSDD数据集共包含1160张不同尺寸的SAR图像, 共有2358个舰船目标。其中, 舰船的最大尺寸为211×298像素, 最小尺寸为7×7像素。该数据集已包含正确标注信息^[29,30], 且可以通过其中多分辨率、多场景图像实现对舰船检测模型性能的有效验证。AIR-SARship-2.0数据集来源于高分三号卫星, 共包含300幅图像, 分辨率包括1 m和3 m, 图像尺寸为1000×1000像素。同时, 该数据集中包括不同的成像模式和极化模式, 不同尺寸和形状的舰船目标, 且覆盖多种场景图像, 为验证舰船检测方法性能提供了可靠保证。

4.2 训练策略与评价指标

实验过程中, 按照6:2:2的比例对训练集、验证集和测试集进行划分。使用SGD优化器完成网络参数修正, 共训练70个epoch, 并采用动态学习率调整训练速度。同时, 在训练过程中使用Tensorboard工具监控训练结果参数, 查看训练情况, 还可查看中间特征图。每训练10个epoch保存一次网络模型, 最终根据训练结果选出最优模型。

本文使用两种评价指标完成实验, 包括传统评价指标检测率 P_D 和虚警率 P_F , 深度学习评价指标检测精确度(Precision, P_r), 召回率(Recall, R_e)和

平均精确度(mean Average Precision, mAP)值、曲线。评价指标中的常用概念表示如下:

(1) 真实的正例(True Positives, TP): 被正确检测出的舰船数目;

(2) 错误的正例(False Positives, FP): 检测出的虚警数目;

(3) 真实的负例(True Negatives, TN): 真值为负例且被分为负例;

(4) 错误的负例(False Negatives, FN): 漏检的舰船数目。

因此, 检测率表示如下:

$$P_D = \frac{TP}{GT} \tag{14}$$

其中, GT (Ground Truth)为真实舰船总数。

虚警率 P_F 定义为

$$P_F = \frac{FP}{TP + FP} \tag{15}$$

检测精确度Precision定义为

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

召回率Recall定义为

$$Recall = \frac{TP}{TP + FN} \tag{17}$$

平均精确度mAP定义为

$$mAP = \int_0^1 P_r(R_e) dR_e \tag{18}$$

其中, $P_r(R_e)$ 为在不同的置信度阈值下, 由检测精确度和召回率所构造的曲线。

4.3 实验结果分析

本文基于公开SAR图像数据集SSDD和AIR-SARship-2.0完成实验。图5展示了显著目标增强预

处理操作效果图。与原图对比, 由于在SAR图像中舰船相对海面背景具有显著特征, 因此在添加预处理操作后, 能够明显增强目标与背景对比度。尽管在某些图像中也包含了显著的建筑目标, 但并不影响突出显示舰船目标, 依然能够对提升舰船检测性能提供一定帮助。同时, 表1中最后两行为是否加入预处理操作的数值效果度量。可以看出, 本文提出方法无论是检测率、虚警率或平均精确度均有更

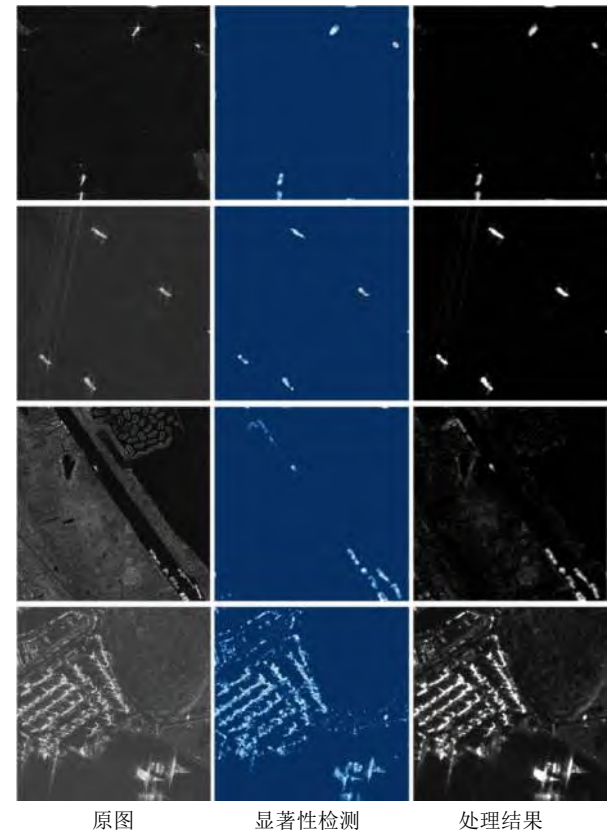


图 5 添加预处理效果图

Fig. 5 The result diagram of the preprocessing module

表 1 基于全部实验数据的不同方法检测性能对比

Tab. 1 Comparison of detection performance of different methods based on all dataset

方法	P_D (%)	P_F (%)	Precision (%)	Recall (%)	mAP (%)	Time (ms)
Yolov2-tiny	53.10	14.97	75.36	53.10	55.20	9.37
Yolov3-tiny	71.03	13.55	76.15	71.03	61.42	11.65
NanoDet	83.56	10.03	83.64	83.56	76.33	4.96
Faster R-CNN	84.98	9.62	82.37	84.98	74.94	313.55
SSD	94.36	7.96	85.83	92.36	82.52	53.98
Yolov4	94.98	4.66	89.34	93.98	88.62	120.34
Yolov5	95.12	5.43	94.68	95.12	90.56	37.52
PANet	93.22	4.52	94.96	93.22	90.77	183.68
未加预处理方法	92.65	5.53	86.71	92.65	85.89	5.39
本文方法	95.64	3.48	95.47	95.64	92.49	5.22

好表现。图6表示是否经过图像预处理的检测方法对比结果。其中,绿色框表示真值,橘色框表示检测结果,蓝色框表示漏警目标。可以看出,添加预处理方法后,能够实现目标与背景的对比度增强并提升图像质量,从而对于较小、较弱目标及密集排列目标等,能够有效提升检测性能。

图7和图8分别为两数据集部分样本舰船检测结果。由此可以直观看出,本文所提出的方法在不同

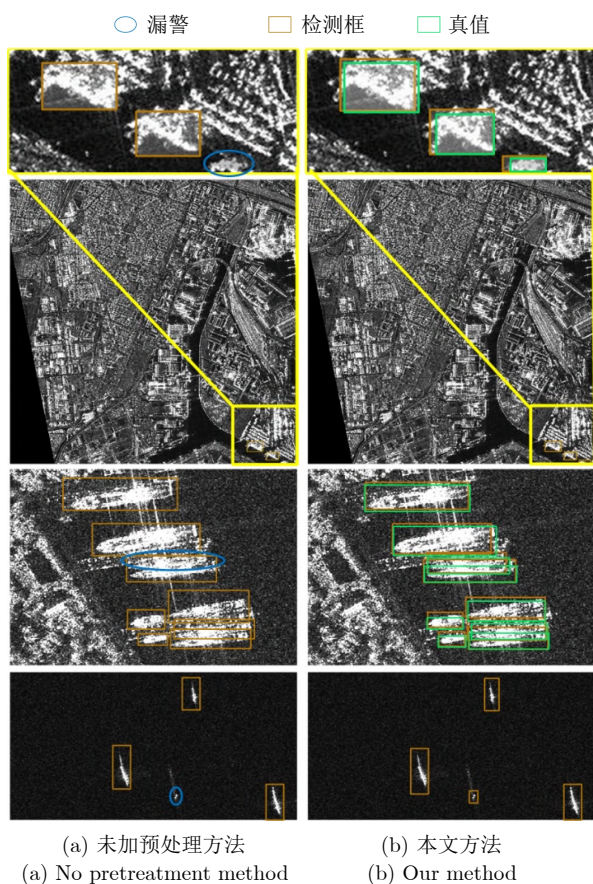


图6 是否加入本文预处理方法检测结果对比图

Fig. 6 The result diagram of the ship detection method with the preprocessing module added or not

背景下,对不同尺度目标均能实现良好检测。其中,测试集共包含459个目标,439个目标被准确测出,也同时产生了16个虚警目标。因此其检测率为95.64%,虚警率为3.48%。深度学习领域中的评价指标,IoU阈值范围在0.5~1.0时,mAP为92.49%。当IoU为0.5时,检测精确度为95.47%。本质上,召回率应等于传统评价指标中的检测率,为95.64%。以上结果表明,本文所提出的方法能够实现较好的检测效果。

为了进一步验证检测算法能力,本文将所提出方法与最新目标检测方法:单阶段的Yolov5, Yolov4^[24], SSD^[21]等,双阶段的Faster R-CNN^[16], PANet^[18],以及轻量化方法Yolov3-tiny^[23], Yolov2-tiny^[22]等方法进行了对比。图9展示了本文方法与其他几种经典方法的检测对比结果。其中,绿框和红框分别表示真值目标框和检测结果框,而红圈和蓝圈分别表示虚警和漏警。可以看出,轻量化算法Yolov3-tiny的检测结果中出现较多的虚警和漏警目标,而本文方法与Yolov5和PANet相比,具有相对更好的检测性能,能够实现良好的检测效果。

此外,表1展示了实验统计结果,对比了基于前述评价指标的各方法检测性能数值。由此可得,在本文所提方法相较于NanoDet有较高的性能提升,因而文中所述优化方法确实能够实现提升目标检测方法的性能。同时,本文所提方法检测平均精确度比Yolov4高约4个百分点,比Yolov5和PANet提高约2个百分点,且相较于经典目标检测方法Faster R-CNN, SSD以及其他轻量化方法等,本文提出方法能够获得更好的检测效果。同时,表1中最后一列为512×512像素大小的图像样本检测时间。对于SSDD数据集,每张图像的尺寸大小不同,但均在512×512左右。本文实验过程中,测试一张653×469大小的图像需要6.19 ms,测试一张500×505大小的图像需要5.04 ms。而对于AIR-SARship数据集,每张图像大小为1000×1000,测

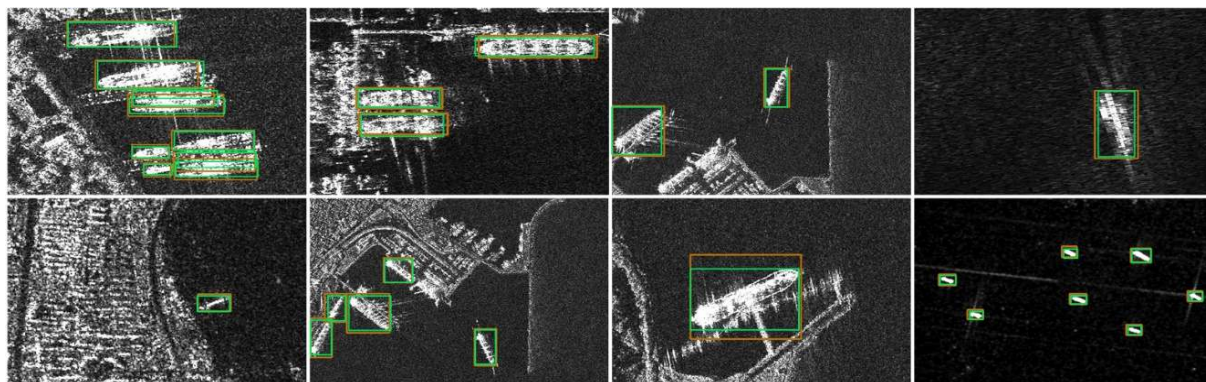


图7 本文方法基于SSDD数据集的检测结果图

Fig. 7 The ship inspection result diagram based on SSDD dataset of the method proposed in this paper

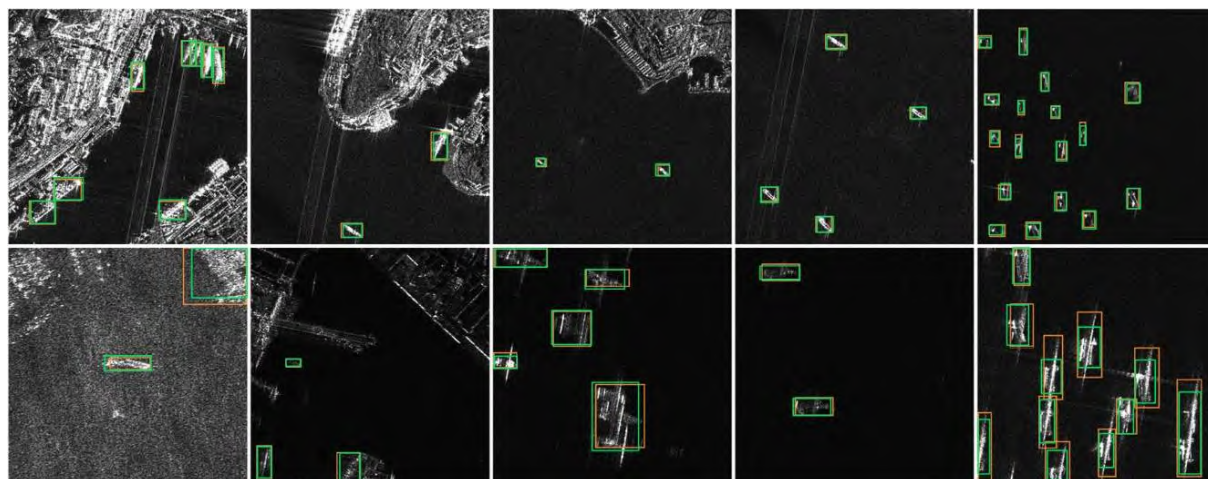


图 8 本文方法基于AIR-SARship数据集的检测结果图

Fig. 8 The detection result diagram based on AIR-SARship dataset of our approach

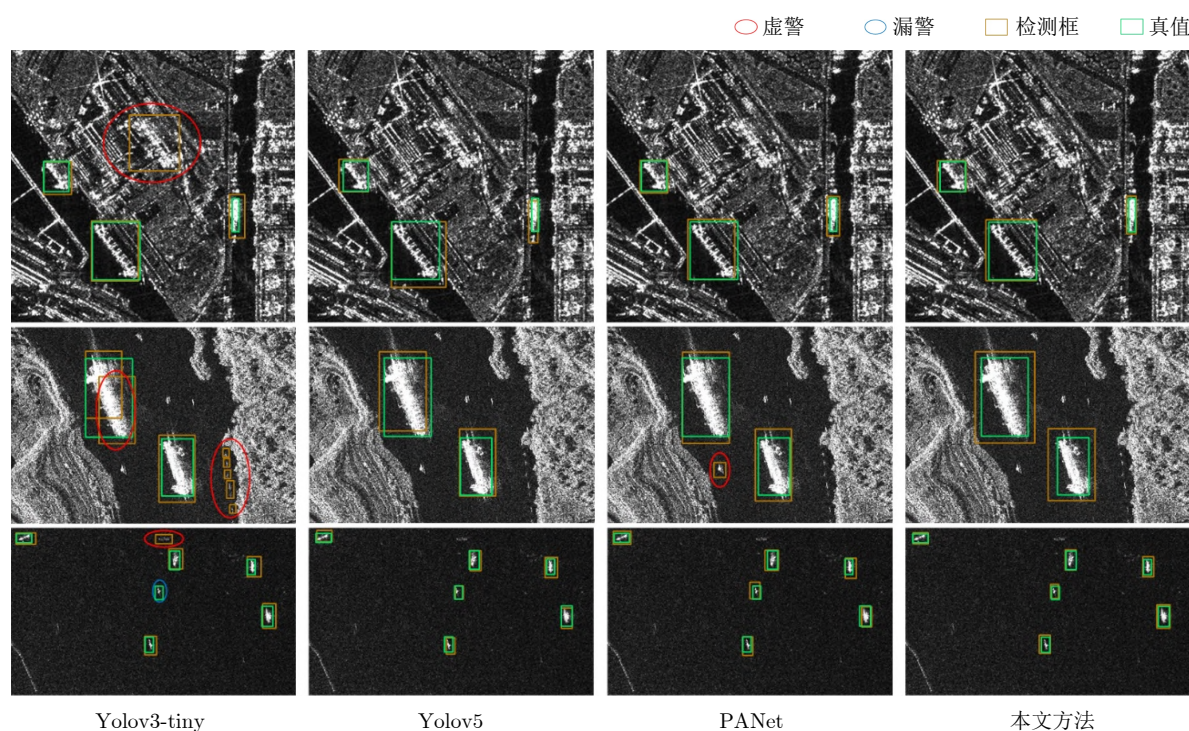


图 9 不同方法检测结果对比图

Fig. 9 Comparison chart of detection results of different methods

试一张图像的时间为19.53 ms。因此,通过像素转换,平均完成处理一张 512×512 像素的图像需要5.22 ms。相较于目前优秀检测方法Yolov5,其检测速度能够达到其7倍左右。因此,本文所提方法不仅能够实现良好检测,且具有较高处理速度,能够满足实际应用中所需实时性要求,因此具有较高的实用价值。

图10为基于SSDD数据集和AIR-SARship-2.0数据集训练所得实验结果。图中线条为各方法的精度-召回率曲线,表示了整体网络模型的检测性能。

该曲线覆盖的面积越大,表示模型性能越好。由此可以看出,本文提出的方法具有与最新方法相当的检测性能,但能获得更快的处理速度。

5 结论

本文基于NanoDet与显著性提出一种实时SAR图像舰船检测方法。在公开舰船检测数据集上实验表明,相较于其他方法,本文算法在保证具有较高检测精确度的同时,也能有效降低检测时间。且对于不同尺寸、形态的目标及不同环境背景下均能实现较好的检测效果。该方法能够适应各种实时性高

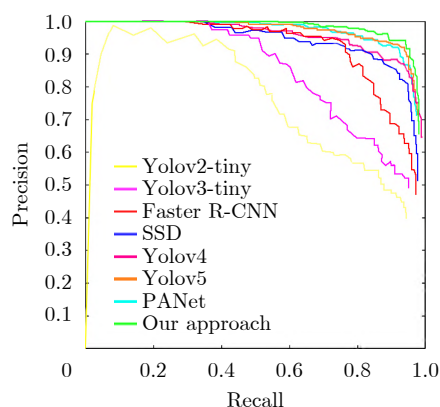


图 10 不同方法精度-召回率曲线

Fig. 10 Precision-Recall curves of different methods.

的检测应用任务, 如紧急救援、实时目标监控和军事部署等。其具有较高的研究意义, 并为未来的轻量化模型硬件移植打下基础。

参 考 文 献

- [1] MOREIRA A, PRATS-IRAOLA P, YOUNIS M, *et al.* A tutorial on synthetic aperture radar[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2013, 1(1): 6–43. doi: [10.1109/MGRS.2013.2248301](#).
- [2] 郭倩, 王海鹏, 徐丰. SAR图像飞机目标检测识别进展[J]. *雷达学报*, 2020, 9(3): 497–513. doi: [10.12000/JR20020](#).
GUO Qian, WANG Haipeng, and XU Feng. Research progress on aircraft detection and recognition in SAR imagery[J]. *Journal of Radars*, 2020, 9(3): 497–513. doi: [10.12000/JR20020](#).
- [3] 张杰, 张晰, 范陈清, 等. 极化SAR在海洋探测中的应用与探讨[J]. *雷达学报*, 2016, 5(6): 596–606. doi: [10.12000/JR16124](#).
ZHANG jie, ZHANG Xi, FAN Chenqing, *et al.* Discussion on application of polarimetric synthetic aperture radar in marine surveillance[J]. *Journal of Radars*, 2016, 5(6): 596–606. doi: [10.12000/JR16124](#).
- [4] 牟效乾, 陈小龙, 关键, 等. 基于INet的雷达图像杂波抑制和目标检测方法[J]. *雷达学报*, 2020, 9(4): 640–653. doi: [10.12000/JR20090](#).
MOU Xiaoqian, CHEN Xiaolong, GUAN Jian, *et al.* Clutter suppression and marine target detection for radar images based on INet[J]. *Journal of Radars*, 2020, 9(4): 640–653. doi: [10.12000/JR20090](#).
- [5] WACKERMAN C C, FRIEDMAN K S, PICHEL W G, *et al.* Automatic detection of ships in RADARSAT-1 SAR imagery[J]. *Canadian Journal of Remote Sensing*, 2001, 27(5): 568–577. doi: [10.1080/07038992.2001.10854896](#).
- [6] 陈慧元, 刘泽宇, 郭炜炜, 等. 基于级联卷积神经网络的大场景遥感图像舰船目标快速检测方法[J]. *雷达学报*, 2019, 8(3): 413–424. doi: [10.12000/JR19041](#).
CHEN Huiyuan, LIU Zeyu, GUO Weiwei, *et al.* Fast detection of ship targets for large-scale remote sensing image based on a cascade convolutional neural network[J]. *Journal of Radars*, 2019, 8(3): 413–424. doi: [10.12000/JR19041](#).
- [7] AUDEBERT N, LE SAUX B, and LEFÈVRE S. How useful is region-based classification of remote sensing images in a deep learning framework?[C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 2016: 5091–5094. doi: [10.1109/IGARSS.2016.7730327](#).
- [8] WANG Wenxiu, FU Yutian, DONG Feng, *et al.* Remote sensing ship detection technology based on DoG preprocessing and shape features[C]. The 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2017: 1702–1706. doi: [10.1109/CompComm.2017.8322830](#).
- [9] HOU Xiaodi and ZHANG Liqing. Saliency detection: A spectral residual approach[C]. 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, USA, 2007: 1–8. doi: [10.1109/CVPR.2007.383267](#).
- [10] GOFERMAN S, ZELNIK-MANOR L, and TAL A. Context-aware saliency detection[C]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 2376–2383. doi: [10.1109/CVPR.2010.5539929](#).
- [11] LIU Zhi, ZOU Wenbin, and LE MEUR O. Saliency tree: A novel saliency detection framework[J]. *IEEE Transactions on Image Processing*, 2014, 23(5): 1937–1952. doi: [10.1109/TIP.2014.2307434](#).
- [12] CHENG Gong, ZHOU Peicheng, and HAN Junwei. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(12): 7405–7415. doi: [10.1109/TGRS.2016.2601622](#).
- [13] LIU Li, OUYANG Wanli, WANG Xiaogang, *et al.* Deep learning for generic object detection: A survey[J]. *International Journal of Computer Vision*, 2020, 128(2): 261–318. doi: [10.1007/s11263-019-01247-4](#).
- [14] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *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. doi: [10.1109/TPAMI.2015.2389824](#).
- [15] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA,

- 2014: 580–587.
- [16] REN Shaoqing, HE Kaiming, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
- [17] NIE Xuan, DUAN Mengyang, DING Haoxuan, *et al.* Attention mask R-CNN for ship detection and segmentation from remote sensing images[J]. *IEEE Access*, 2020, 8: 9325–9334. doi: [10.1109/ACCESS.2020.2964540](https://doi.org/10.1109/ACCESS.2020.2964540).
- [18] LIU Shu, QI Lu, QIN Haifang, *et al.* Path aggregation network for instance segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 8759–8768. doi: [10.1109/CVPR.2018.00913](https://doi.org/10.1109/CVPR.2018.00913).
- [19] LIU Yudong, WANG Yongtao, WANG Siwei, *et al.* Cbnet: A novel composite backbone network architecture for object detection[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 11653–11660. doi: [10.1609/aaai.v34i07.6834](https://doi.org/10.1609/aaai.v34i07.6834).
- [20] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: Unified, real-time object detection[C]. The IEEE conference on computer vision and pattern recognition, Las Vegas, USA, 2016: 779–788.
- [21] LIU Wei, ANGUELOV D, ERHAN D, *et al.* SSD: Single shot multibox detector[C]. The 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016: 21–37. doi: [10.1007/978-3-319-46448-0_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [22] REDMON J and FARHADI A. YOLO9000: Better, faster, stronger[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 6517–6525. doi: [10.1109/CVPR.2017.690](https://doi.org/10.1109/CVPR.2017.690).
- [23] REDMON J and FARHADI A. Yolov3: An incremental improvement[C]. arXiv: 1804.02767, 2018.
- [24] BOCHKOVSKIY A, WANG C Y, and LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[C]. arXiv: 2004.10934, 2020.
- [25] IANDOLA F N, HAN Song, MOSKEWICZ M W, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[C]. arXiv: 1602.07360, 2016.
- [26] MITTELMANN H and PENG Jiming. Estimating bounds for quadratic assignment problems associated with hamming and Manhattan distance matrices based on semidefinite programming[J]. *SIAM Journal on Optimization*, 2010, 20(6): 3408–3426. doi: [10.1137/090748834](https://doi.org/10.1137/090748834).
- [27] LI Yanghao, CHEN Yuntao, WANG Naiyan, *et al.* Scale-aware trident networks for object detection[C]. 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 6053–6062.
- [28] PASZKE A, GROSS S, MASSA F, *et al.* Pytorch: An imperative style, high-performance deep learning library[C]. The 33rd Conference on Neural Information Processing Systems, Vancouver, Canada, 2019: 8026–8037.
- [29] LI Jianwei, QU Changwen, and SHAO Jiaqi. Ship detection in SAR images based on an improved faster R-CNN[C]. 2017 SAR in Big Data Era: Models, Methods and Applications, Beijing, China, 2017: 1–6. doi: [10.1109/BIGSDATA.2017.8124934](https://doi.org/10.1109/BIGSDATA.2017.8124934).
- [30] 李健伟, 曲长文, 彭书娟. 基于级联CNN的SAR图像舰船目标检测算法[J]. 控制与决策, 2019, 34(10): 2191–2197. doi: [10.13195/j.kzyjc.2018.0168](https://doi.org/10.13195/j.kzyjc.2018.0168).
- LI Jianwei, QU Changwen, and PENG Shujuan. A ship detection method based on cascade CNN in SAR images[J]. *Control and Decision*, 2019, 34(10): 2191–2197. doi: [10.13195/j.kzyjc.2018.0168](https://doi.org/10.13195/j.kzyjc.2018.0168).

作者简介



刘方坚(1979–), 男, 山东临沂人, 中国科学院空天信息创新研究院副研究员, 主要研究方向为遥感卫星地面处理系统技术研究等。



李媛(1996–), 女, 河北石家庄人, 于北京化工大学获得学士、硕士学位, 现为北京理工大学博士生, 主要研究方向为遥感图像分类、目标检测等。