

DOI:10.19651/j.cnki.emt.2209295

一种基于特征重用金字塔的舰船检测算法*

卢鑫凯 王肖霞 杨凤暴 刘哲

(中北大学信息与通信工程学院 太原 030051)

摘要: 针对现有算法在 SAR 图像舰船目标检测场景中难以提取模糊目标特征的问题,提出一种基于特征重用金字塔的舰船目标检测算法。所提算法以 YOLOV4-tiny 为主体,首先将线性因子引入到 K-means 算法中整合初始锚框,加强网络对多尺度目标的适应性;其次在主干 CSPDarknet53-tiny 中添加注意力机制来抑制干扰信息,减弱复杂背景的影响;最后利用特征重用机制强化特征金字塔,提升网络对模糊目标特征的提取能力。实验结果表明,相较于 YOLOV4-tiny 网络,改进后的算法在 SSDD 数据集上的平均检测精度提升 11.79%,证明了改进后算法在舰船检测中的有效性。

关键词: 舰船目标;SAR 图像;K-means;注意力机制;特征重用;YOLOV4-tiny

中图分类号: TP751 **文献标识码:** A **国家标准学科分类代码:** 510.4010

A ship detection algorithm based on feature reuse pyramid

Lu Xinkai Wang Xiaoxia Yang Fengbao Liu Zhe

(School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

Abstract: Aiming at the problem that the existing algorithms are difficult to extract fuzzy target features in the SAR image ship target detection scene, a ship target detection algorithm based on feature reuse pyramid is proposed. The proposed algorithm takes YOLOV4-tiny as the main body. First, a linear factor is introduced into the K-means algorithm to integrate the initial anchor frame to enhance the adaptability of the network to multi-scale targets. Secondly, an attention mechanism is added to the backbone CSPDarknet53-tiny to suppress interference information, and weaken the influence of complex background; finally, the feature reuse mechanism is used to strengthen the feature pyramid and improve the network's ability to extract fuzzy target features. The experimental results show that, compared with the YOLOV4-tiny network, the average detection accuracy of the improved algorithm on the SSDD dataset is improved by 11.79%, which proves the effectiveness of the improved algorithm in ship detection.

Keywords: ship target;SAR image;K-means;attention mechanism;feature reuse;YOLOV4-tiny

0 引言

合成孔径雷达(synthetic aperture radar, SAR)因其成像受到外界环境影响较小,可以提供全天候、远距离观测的高分辨率的图像^[1]的特点,从而在海上舰船检测方面得到了广泛的应用。在民用领域上,有助于海上贸易运输管理;在军事领域上,有利战略防御策略制定,提高海上防御能力^[2]。然而 SAR 图像中部分舰船目标特征信息不丰富,难以在模型中充分表达,会导致当前流行的诸多目标检测算法无法准确提取目标特征,并且由于存在尺度变化大、易受复杂背景干扰等因素,严重影响后期的舰船分类与身份识别等工作。因此,提高 SAR 图像舰船检测算法的检测性能具有重要意义。

目前,深度学习在舰船目标上的检测已取得部分成果,主流的模型有 Faster R-CNN^[3]、SSD^[4]、YOLOV4^[5]等。文献[6]基于 Faster R-CNN,通过对主干网络特征层加入 FPN 结构进行多次融合,加强了各层特征的利用率,取得良好的检测效果;文献[7]在 SSD 模型引入线性修正单元,充分利用浅层特征的语义信息,实现了对小目标舰船的精准定位与识别;文献[8]通过改进非极大值抑制方法加强网络对密集目标的适应能力,提升了检测效果。文献[9]将改进的注意力机制加入 YOLOV4 模型中,增强了对图像中目标的特征表达能力,检测效果提升较大。但是目前的研究方法所用网络层次较深,忽略了目标特征信息不丰富时深层特征与浅层特征的融合,并且在轻量网络上进行的研

收稿日期:2022-03-16

* 基金项目:国家自然科学基金(61972363)项目资助

究很少,检测速度难以达到实际应用的要求。YOLOV4-tiny 算法因其超越了当前绝大部分网络的检测速度而被广泛应用,但是该网络的轻量设计导致其检测精度较低,在实际应用中仍存在提升空间。

基于以上研究现状,本文在 YOLOV4-tiny 的基础上提出一种改进的舰船目标检测算法。该算法主要通过特征重用金字塔结构来实现各层特征的融合,达到提升模糊目标的检测性能的目的。同时改进 K-means 算法^[10]计算出适合多尺度舰船数据集的锚框,并在 CSPdarknet53-tiny 主干网络中添加注意力模块抑制复杂背景,最后在 SSDD^[11]数据集上进行不同场景和不同尺度目标的检测实验,结果表明该方法能有效检测到模糊舰船,并且对复杂背景下的多尺度舰船目标有着较高的检测精度,为神经网络在 SAR 图像舰船目标检测提供了一定的研究基础。

1 YOLOV4-tiny 原理

YOLOV4-Tiny 是 YOLOV4 经过精简之后的轻量版本,计算量小,检测速度更快。YOLOV4-tiny 的主干网络 CSPDarknet53-tiny 中主要通过 CBL 模块和 CSPBlock 残差模块组成,CBL 模块由普通卷积层(Conv)、批量标准化(BN)和 Leaky Relu 激活函数组成,CSPBlock 模块将输入特征分为两部分,在进行残差块的堆叠的同时引入一个大的残差边,最后将两个通道组合后经过最大池化(Max Pool)操作完成下采样,这种结构可以降低计算量,提升网络运行速度。FPN 结构是将主干网络最后一个特征层进行上采样后与上一个 CSPBlock 模块输出的特征层进行 Concat 操作来实现,可以对高层特征和底层特征的信息进行融合,从而对网络的特征提取能力进行强化,利用最后得到的两个特征层进行预测。

2 改进的 YOLOV4-tiny 网络

针对 SAR 图像舰船目标检测场景,本文在 YOLOV4-Tiny 网络模型中采用基于线性因子改进的 K-means 算法降低算法在锚点选取的随机性带来的影响,以加强网络对多尺度目标的适应能力;另外在主干部分第一层卷积层和最后一层 CSPBlock 引入 CBAM 注意力模块提升网络对目标的关注程度,抑制复杂背景的影响。最后提出 R-FPN 的方法,采用特征重用机制将分辨率更高的低层特征和语义信息更丰富的高层特征进行融合,对 FPN 结构进行强化,增强网络的特征提取能力,提升检测精度。

2.1 引入线性因子优化的 K-means 算法

YOLOV4-tiny 网络原有的锚框是通过 VOC 数据集^[12]中用 K-means 算法聚类得来,VOC 数据集中的可见光目标与 SSDD 数据集中的舰船目标特征差异很大,虽然 SAR 图像舰船目标检测中只有一类目标,但是目标的尺度分布不均匀,因此原始锚框不再适用于当前场景。并且原始锚框计算方法对初始点的选取过于随机,会影响后续的

网络拟合速度,虽然网络在训练过程中可以逐渐整合锚框得到最优策略,但是合适的初始锚框能在一定程度上加速网络的拟合,提升网络对不同尺度目标的适应能力。因此本文引入线性因子改进的 K-means 算法对锚框进行新的初始化聚类,以下为改进后的 K-means 算法步骤:

步骤 1)分析数据集中的所有样本的 *box*, 随机选取 *K* 个作为聚类中心。

步骤 2)计算数据集中每个 *box* 到各个聚类中心的 IOU 距离。计算公式如下:

$$d(box, anchor) = 1 - IOU(box, anchor) \quad (1)$$

$$IOU(box, anchor) = \frac{\min(w_a, w_b) \min(h_a, h_b)}{w_a h_a + w_b h_b - \min(w_a, w_b) \min(h_a, h_b)} \quad (2)$$

式中: h_a 和 w_a 分别为一个 *anchor* 的长宽, h_b 和 w_b 分别为一个 *box* 的长宽。

步骤 3)找出每个 *box* 对应 IOU 距离最小的聚类中心,将其分配到离它最近的聚类部分。

步骤 4)再针对每个聚类部分中的 *box* 重新计算聚类中心。

步骤 5)重复步骤 3)和 4),直到聚类中心的位置不再变化。

步骤 6)引入线性因子 α 和 β , 对步骤 5)得到的 *anchor* 做线性变换。假设步骤 5)得到 *anchor* 的 *K* 个锚框为 $(r_1, t_1), \dots, (r_i, t_i), \dots, (r_k, t_k)$, 只对 r_1 和 r_k 引入 α 和 β 后结果为 $\alpha r_1, \beta r_k$, 则可以求得同一方向上的比例因子 θ , 再令同一方向上的所有 *r* 元素都遵循该比例因子进行比例变换,公式如下:

$$\theta = \frac{\beta r_k - \alpha r_1}{r_k - r_1} \quad (3)$$

$$\begin{cases} r'_i - \alpha r_1 = \theta \\ r'_i - r_1 \end{cases} \quad (4)$$

$$\begin{cases} t'_i = t_i \\ r'_i = r_i \end{cases}$$

式中: (r_i, t_i) 表示步骤 5)得到的第 *i* 个锚框值; (r'_i, t'_i) 表示 (r_i, t_i) 变换后的结果, α 和 β 为线性因子参数。

步骤 7)对上述 5)个步骤多次重复,并将每一次求得的 *anchor* 中的对应项求平均值,计算公式如下:

$$anchor_f = \frac{1}{n} \sum_{m=1}^{m=n} anchor_m \quad (5)$$

式中: $anchor_m$ 表示第 *m* 次计算所得到的 *anchor* 组合, $anchor_f$ 表示最终求得的 *anchor* 组合, *n* 为重复计算的次数。

本文令 $n=100$, 当 $\alpha=0.4$, $\beta=3.5$ 时检测效果最好, 6 个锚框尺寸分别为 (5, 7), (8, 25), (44, 128), (54, 50), (134, 335), (258, 129)

2.2 CBAM 注意力机制

由于 SAR 图像中的复杂背景信息较多,极大的影响网络对舰船的特征提取效果,因此引入 CBAM(convolutional block attention module)注意力机制模块^[13]来对复杂背景

进行抑制,其结构如图 1 所示。CBAM 注意力机制是一种轻量的网络模块,它同时结合了通道注意力模块^[14](channel attention module, CAM)和空间注意力模块^[15](spatial attention module, SAM),可以获得更多的有效信息,并且额外的运算量极小,经常被集成在各种模型中。

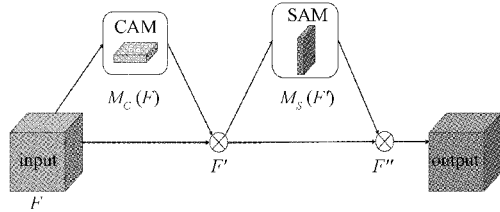


图 1 CBAM 模块图

通道注意力部分首先输入特征图进行池化操作和多层感知运算得到特征向量,将其激活后与原始的输入特征进行融合,得到新的特征 F' , 实现方式如式(6)所示。

$$F' = M_c(F) \otimes F \quad (6)$$

式中: F 为输入特征, $M_c(F)$ 为通道注意力操作, F' 为通道注意力模块的输出。

空间注意力部分将通道注意力模块的结果 F' 作为空间注意力模块的输入进行池化和通道数的整合,随后通过卷积运算和激活后与输入 F' 融合得到最终的输出特征 F'' 。实现方式如式(7)所示。

$$F'' = M_s(F') \otimes F' \quad (7)$$

式中: F' 为输入特征, $M_s(F')$ 为空间注意力操作, F'' 为空间注意力模块的输出。

由于添加注意力模块过多会增加网络复杂度,因此会通实验确定少量 CBAM 模块的最优添加位置。

2.3 特征重用金字塔网络

特征金字塔网络(feature pyramid network, FPN)^[16]部分对目标检测模型的检测精度有很大影响, YOLOV4-tiny 中的 FPN 结构较为简单,计算量小,检测速度快,图 2 为 YOLOV4-tiny 中的 FPN 结构。

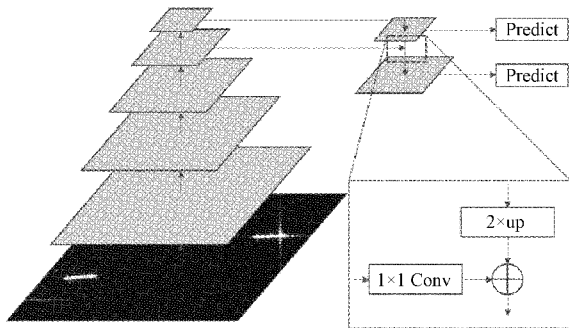


图 2 YOLOV4-tiny 中的 FPN 结构图

这种 FPN 结构难以提取分辨率较低的有效高层特征,因此本文基于密集连接^[17]的思想,提出了一种特征重用金字塔网络(reuse feature pyramid network, R-FPN),通过在 FPN 结构中加入特征重用机制对其进行优化,以提升网络

前后传播过程中对前层特征的利用率,从而增强对模糊目标的检测性能。图 3 为改进后的 R-FPN 结构,其中 up 表示二倍上采样, d 表示二倍下采样, C 表示 1×1 的卷积块。此模块首先对 CSPDarknet53-tiny 主干网络的特征层 C_5 进行一次 1×1 的卷积块进行通道变换得到 P_5 , 然后对特征层 C_4 实行下采样操作,将尺寸为 26×26 的特征图压缩到 13×13 大小,接着将 P_5 和经过下采样后的 P_4 进行 concat 操作进行融合,得到第一个 13×13 的特征层用来预测。同理,将主干网络的特征层 C_3 进行上采样,得到大小为 26×26 的特征层,该特征层与 P_4 进行通道数的整合后,和之前第一个预测头的结果再次进行 concat 操作,得到的 26×26 特征层作为新的预测。

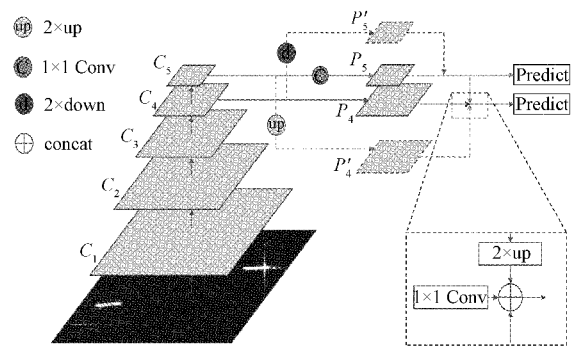


图 3 R-FPN 结构

R-FPN 可以使 YOLOV4-tiny 网络的当前层接收前层特征作为输入,提高了了特征利用率,并且可以通过少量的卷积核生成含有丰富位置信息和语义信息的特征图,从而减少模型的计算量,提升网络的特征提取能力。通过嵌入 R-FPN 和 CBAM 模块改进后的 YOLOV4-tiny 详细结构如图 4 所示。

3 实验结果与分析

3.1 SAR 图像舰船目标数据集

本文采用的数据集是目前舰船检测中流行的 SSDD 数据集,标注方式以 PASCAL VOC 格式为基准,共有 1 160 张图像,该数据集包含多种极化模式下采集到的 2 456 个舰船目标,并且场景多样,可以有效验证模型的性能。

3.2 评估指标

实验采用平均精确率 AP 作为主要评价指标来反映模型的性能。mAP 表示所有种类目标的 AP 值的均值, mAP 越大证明模型检测性能越好,因为实验只检测舰船目标一类,所以 AP 值等于 mAP。mAP 计算公式如下:

$$mAP = \frac{\sum_{K=1}^N AP_K}{N} \quad (8)$$

式中: N 表示目标类别数; AP 表示每个类别的平均精确率,由计算精确率(Precision)和召回率(Recall)形成的 PR 曲线围成的面积得来。

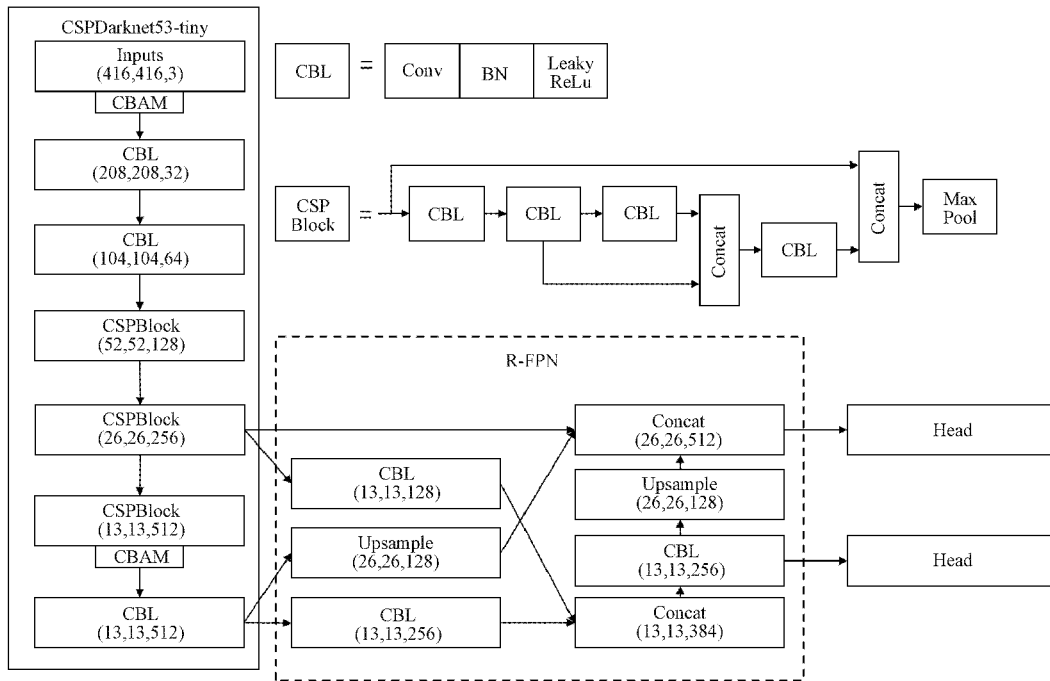


图 4 改进后的 YOLOV4-tiny

3.3 实验环境及参数设置

实验在 windows10 操作系统的 Pytorch1.6 框架下运行,并使用 CUDA10.0 和 cuDNN7.0 来加速训练,显卡为 NVIDIA GTX950M,显存为 2 G。将数据集按照 8 : 2 的比例随机分为训练集和测试集,采用随机梯度下降法进行优化,设置的初始学习率为 0.001,衰减系数设置为 0.0003,每个算法训练 100 个 epoch,训练的 batch size 取 32。训练过程中,保存验证结果最优的模型直至训练完成。所有模型训练使用的数据集均进行数据增强,数据增强的主要目的是增加训练数据,使数据多样化,有助于加强训练所得模型具的泛化能力。数据增强主要采用了翻转、旋转、缩放、裁剪、平移、加高斯噪声、改变亮度等方案。实验分为 3 部分,第 1 部分为改进后 K-means 算法中 α, β 参数值的选取、CBAM 注意力机制模块添加位置确认实验;第 2 部分为消融实验;第 3 部分为多种算法对比实验。

3.4 实验结果

首先对改进的 K-mean 聚类的参数选取做分析,采用 IOU 阈值为 0.5 时的 mAP 为判断基准。令重复次数 $n = 100$,采用 5 种参数选择策略进行对比,方法①为 $\alpha = 0.1, \beta = 2.0$;方法②为 $\alpha = 0.2, \beta = 2.5$;方法③为 $\alpha = 0.3, \beta = 3.0$;方法④为 $\alpha = 0.4, \beta = 3.5$;方法⑤为 $\alpha = 0.5, \beta = 4.0$ 。与原始 YOLOV4-tiny 对比,实验结果如图 5 所示。由图 5 可知,K-means 算法采用方法④设计的参数后相比原来算法的 mAP 值提升 0.45%,因此确认参数选择为 $\alpha = 0.4, \beta = 3.5$ 。

注意力机制位置确定实验采用 IOU 阈值为 0.5 时的 AP 为判断基准,选取 5 种方法,方法 A 为 C_1 层后添加;方

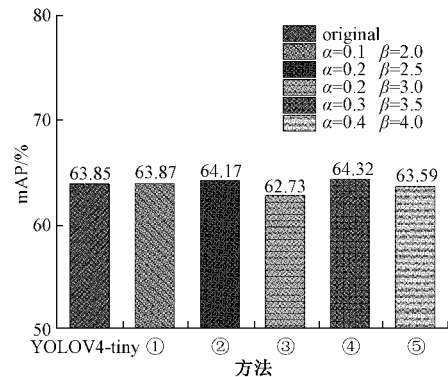


图 5 改进 K-means 的参数选取对比图

法 B 为在 C_3 层后添加;方法 C 为在 C_5 层后添加;方法 D 为同时在 C_1 和 C_5 后添加;方法 E 为同时在 C_1 和 C_3 后添加,与原始 YOLOV4-tiny 对比,实验结果如图 6 所示;由图 6 可知,采用方法 D 中的位置添加注意力模块后,模型的 AP 上升 0.76%,提升最大,因此选择在 C_1 层和 C_5 层后同时串联 CBAM 模块。

网络的训练过程中的 Loss 值随着 epoch 轮数变化如图 7 所示,可以看出改进网络相比于原始 YOLOV4-tiny 的 Loss 值下降的更快,可以更好的收敛。表 1 中消融实验的 AP0.3、AP0.4、AP0.5、AP0.6、AP0.7 分别表示 IOU 阈值在 0.3、0.4、0.5、0.6、0.7 时的检测精度,AP 均值为对 AP0.3、AP0.4、AP0.5、AP0.6、AP0.7 求平均。从消融实验可以看出,本文的 3 个改进点均可以提升模型性能,改进 K-means、添加 CBAM 模块对模型性能提升分别为 0.96%、1.93%;改进 K-means 同时添加 CBAM 模块后,

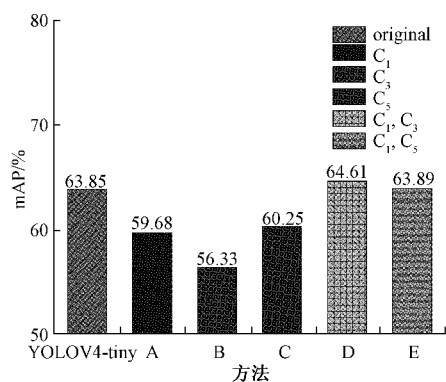


图 6 不同位置下注意力模块效果对比图

模型性能提升 2.7%，尤其是在加入 R-FPN 改进后的模型相比于原始 YOLOV4-tiny 模型，在检测时间增加极少的

情况下，AP 均值提升 11.79%，证明了本文几种改进方法的有效性。

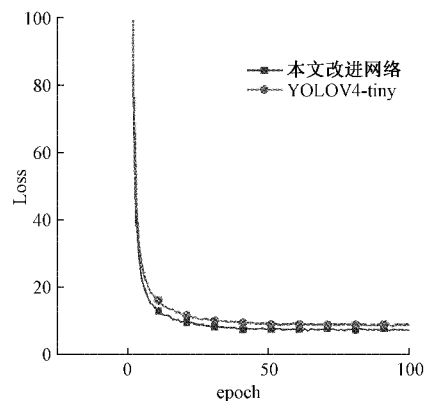


图 7 训练过程中 Loss 值变化图

表 1 消融实验结果表

网络	AP0.3	AP0.4	AP0.5	AP0.6	AP0.7	AP 均值	时间/ms
YOLOV4-Tiny	62.65	63.87	63.85	59.88	51.66	60.38	44.2
改进 K-means	62.88	63.82	64.32	60.27	55.40	61.34	44.1
CBAM	62.53	63.79	64.61	63.61	57.04	62.31	45.6
改进 K-means+CBAM	64.11	64.72	64.73	62.50	59.36	63.08	45.2
R-FPN+改进 K-means+CBAM	74.18	74.56	73.86	71.55	66.69	72.17	52.7

多种算法对比实验结果如表 2 所示，可以看出本文改进的模型相比于 YOLOV4-tiny、YOLOV3-tiny 两种轻量

模型的 AP 均值提升较大，并且与当前流行的舰船检测算法 SSD、CenterNet 的效果相当。

表 2 多种算法比较实验结果表

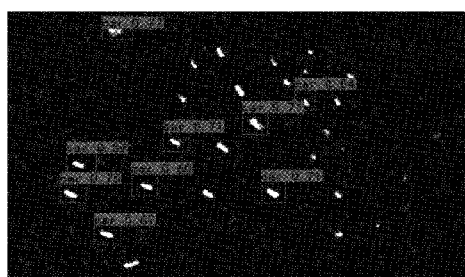
网络	AP0.3	AP0.4	AP0.5	AP0.6	AP0.7	AP 均值
YOLOV4-tiny	62.65	63.87	63.85	59.88	51.66	60.38
YOLOV3-tiny	57.62	59.20	60.31	56.82	53.45	57.48
SSD	72.26	73.33	74.16	73.29	65.85	71.78
CenterNet	68.21	69.76	69.87	66.63	64.24	67.74
Our method	74.18	74.56	73.86	71.55	66.69	72.17

考虑到实验的严谨性，在图 8 和 9 中分别对简单背景和复杂背景下不同尺度目标进行了对照检测，算法采用原始的 YOLOV4-tiny 算法和本文改进后的算法。由图 8 得出，改进后的算法在简单背景下可以检测到更多的小目标，有效降低漏检率。由图 9 的对比结果可得，改进后的模型对复杂背景下目标的检测置信度有了提升，尤其是目

标尺寸较大时，可以检测到 YOLOV4-tiny 检测不到的目标，检测效果略有提升，但依然存在漏检的情况。总体来看改进后的 YOLOV4-tiny 对多种场景下多尺度目标的检测更加准确，表明改进后的特征重用型 YOLOV4-tiny 网络更充分的利用浅层特征，对于分辨率低、特征信息模糊、有大面积复杂背景的多尺度舰船目标检测效果更好。



(a) YOLOV4-tiny 密集目标检测结果



(b) 本文算法密集目标检测结果

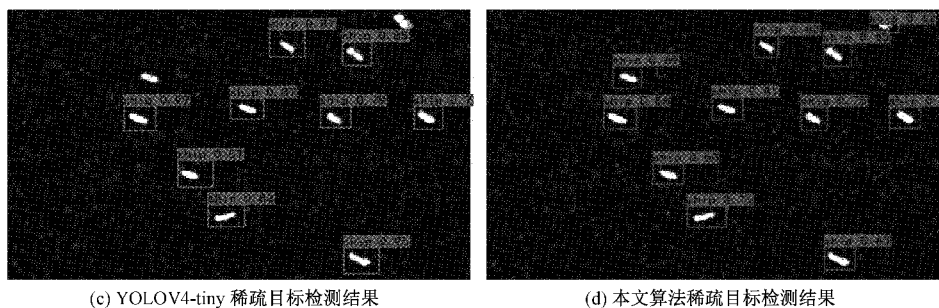


图 8 简单背景下检测结果对比图

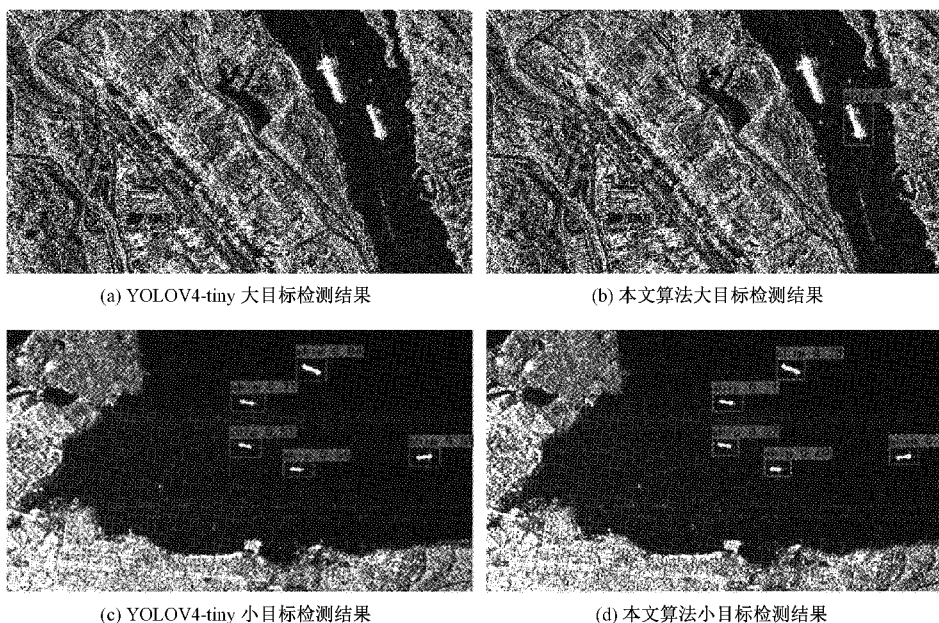


图 9 复杂背景下检测结果对比图

4 结 论

本文提出了一种基于特征重用金字塔的 SAR 图像舰船检测算法,从舰船的多尺度分布、存在复杂背景干扰和模糊特征信息难以利用三个角度对模型分别进行创新。通过 SSDD 数据集验证验证可知,相比于 YOLOV4-tiny 网络,在 K-means 算法中引入线性因子后平均精度提升 0.96%;添加注意力模块到主干网络后平均精度提升 1.93%;结合以上两种方式和特征重用金字塔后平均精度提升 11.79%,效果最为明显,证明了所提特征重用金字塔方法在对目标特征提取上的有效性,并且检测精度高于目前的主流算法,具有一定的应用价值。

参考文献

[1] VILLANO M, KRIEGER G, STEINBRECHER U, et al. Simultaneous single-/dual-and quad-pol SAR imaging over swaths of different widths[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 58(3): 2096-2103.

[2] 常佳慧,赵建辉,李宁.一种改进的 2P-CFAR SAR 舰船检测方法[J].国外电子测量技术,2021,40(11):7-12.

[3] 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.

[4] LIU W, ANGUOLOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision Springer, Cham, 2016: 21-37.

[5] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint,2020, ArXiv:2004.10934.

[6] 韩子硕,王春平,付强.基于深层次特征增强网络的 SAR 图像舰船检测[J].北京理工大学学报,2021,41(9):1006-1014.

[7] 李晖晖,周康鹏,韩太初.基于 CReLU 和 FPN 改进的

- SSD 舰船目标检测[J]. 仪器仪表学报, 2020, 41(4): 183-190.
- [8] 张宏群, 班勇苗, 郭玲玲, 等. 基于 YOLOv5 的遥感图像舰船的检测方法[J]. 电子测量技术, 2021, 44(8): 87-92.
- [9] 崔江波, 侯兴松. 基于注意力机制的 YOLOv4 输电线路故障检测算法[J]. 国外电子测量技术, 2021, 40(7): 24-29.
- [10] SINAGA K P, YANG M S. Unsupervised K-means clustering algorithm[J]. IEEE access, 2020, 8, DOI: 10.1109/ACCESS.2020.2988796.
- [11] LI J, QU C, SHAO J. Ship detection in SAR images based on an improved faster R-CNN[C]. 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA), IEEE, 2017: 1-6.
- [12] EVERINGHAM M, ESLAMI S M, VAN GOOL L, et al. The pascal visual object classes challenge: A retrospective[J]. International Journal of Computer Vision, 2015, 111(1): 98-136.
- [13] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [14] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.
- [15] ZHU X, CHENG D, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6688-6697.
- [16] 刘英杰, 杨风暴, 胡鹏. 基于 Cascade R-CNN 的并行特征金字塔网络无人机航拍图像目标检测算法[J]. 激光与光电子学进展, 2020, 57(20): 302-309.
- [17] 许敏, 李博涵, 王凯, 等. 特定辐射源个体识别算法研究[J]. 电子测量与仪器学报, 2021, 35(10): 116-123.

作者简介

卢鑫凯, 硕士研究生, 主要研究方向为信号处理、遥感目标检测技术。

E-mail: 786063184@qq.com

王肖霞(通信作者), 工学博士, 副教授, 主要研究方向为信号处理、关联成像、可能性分布理论研究。

E-mail: wangxiaoxia@nuc.edu.cn