

基于多尺度语义网络的红外舰船目标检测

陈初侠¹, 丁勇²

(1. 巢湖学院 电子工程学院, 安徽 巢湖 238000; 2. 浙江大学 微纳电子学院, 浙江 杭州 310000)

摘要: 为了增强舰船检测的抗干扰性能, 本文提出了一种有效且稳定的单阶段舰船检测网络, 该网络主要由3个模块组成: 特征优化模块, 特征金字塔融合模块和上下文增强模块, 其中特征优化模块是提取多尺度上下文信息, 并进一步细化和增强顶层特征输入特性, 增强弱小目标检测性能; 特征金字塔融合模块能够生成表征能力更强的语义信息; 上下文增强模块则是整合局部和全局特征增强网络特征表达能力, 以降低复杂背景对检测性影响, 平衡前景和背景的不均衡差异, 消除鱼鳞波的影响。为了验证本文所提方法的有效性和鲁棒性, 本文对自建的舰船数据集进行了定性定量验证。实验结果表明, 相比现有最新基准对比模型, 本文所提网络在自建数据集上均达到了最优性能, 在不增加复杂度的情况下极大提升了检测精度。

关键词: 目标检测; 红外舰船; 单阶段网络; 金字塔池化; 上下文增强

中图分类号: TN219 **文献标识码:** A **文章编号:** 1001-8891(2022)05-0529-08

Infrared Ship Detection Based on Multi-scale Semantic Network

CHEN Chuxia¹, DING Yong²

(1. School of Electronic Engineering, Chaohu University, Chaohu 238000, China;

2. School of Micro Nano Electronics, Zhejiang University, Hangzhou 310000, China)

Abstract: To enhance the anti-jamming performance of ship detection, an effective and stable single-stage ship detection network is proposed in this study. The network is composed of three modules: feature optimization, feature pyramid fusion, and context enhancement modules. The feature optimization module extracts multi-scale context information and further refines the high-level feature input characteristics, to enhance the performance of dim-small object detection. The feature pyramid fusion module can generate semantic information with stronger representation ability. The context enhancement module integrates local and global features to enhance the network feature expression ability, reduce the impact of a complex background on detectability, adjust the imbalance between the foreground and background, and eliminate the impact of scale-wave. To verify the effectiveness and robustness of the proposed method, qualitative and quantitative verifications are performed on a self-built dataset. Experimental results show that the proposed network achieves optimal performance compared with the latest benchmark comparison model and considerably improves the detection accuracy without increasing complexity.

Key words: object detection, infrared ship, single-stage network, pyramid pooling, context enhancement

0 引言

众所周知, 红外反舰导弹具有对水面舰艇目标更高的抗干扰能力, 其核心功能是高精度的目标检测与跟踪^[1]。船舶检测是反舰制导武器的一项极具挑战的基础任务, 其目的是对视场内的舰船目标检测分类与定位^[2]。

传统的舰船目标检测算法大都是基于非平稳高斯噪声下的目标恒虚警率模型, 该模型集成了自适应阈值策略和海杂波统计模型, 对高分辨率的大尺度目标检测精度较高, 但对复杂背景低对比度目标的检测能力不足^[3]。近年来, 深度网络以其独特的特征表征能力, 大大提高了目标检测任务的精度。现有的深度网络检测方法大致可以分为单阶段检测网络与双阶

收稿日期: 2021-05-05; 修订日期: 2021-11-29.

作者简介: 陈初侠(1984-), 男, 实验师, 主要研究方向为数字图像处理。E-mail: feng84chen@163.com.

通信作者: 丁勇(1974-), 男, 教授, 博士生导师, 主要研究方向为图像深度分析与质量评价。E-mail: dingyong09@zju.edu.cn.

段检测网络^[4]。单阶段检测网络省略了候选区域生成，直接在原图上回归目标类别与位置，而双阶段检测网络专注于找出目标物体出现的位置，得到候选区域，并在其基础上进行分类，寻找更精确的位置。相比双阶段检测网络，单阶段检测网络的典型代表有 YOLO (You Only Look Once) 系列^[5]，SSD (Single Shot MultiBox Detector) 系列^[6]，RetinaNet^[7]，CornerNet^[8]，CenterNet^[9]，EfficientDet^[10]等。尽管这些网路在自然图像中取得了优异的性能，但很难直接迁移到红外舰船检测任务中。主要原因归咎于目标成像特性的差异，例如，自然图像中的目标大多分辨率较高，且目标尺度较大，而红外图像中的舰船目标受成像距离与鱼鳞波的影响较大。因此，国内外学者开始对深度神经网络进行改进，以适应红外舰船检测需求。西南技术物理研究所唐中和团队借鉴 YOLOv3^[5]的一些思想，设计了一种基于弹载轻量级深度学习网络，提升舰船目标检测性能的同时，大大降低了硬件资源需求^[11]。北航“新型成像探测技术与系统”蓝天创新团队^[12]对复杂场景下的舰船目标检测算法进行了分析，设计出鲁棒性好、准确度高与实时性强的深度学习模型，有效地解决了前景和背景的不均衡差异。UT Austin 大学崔屹教授^[9]等人将目标检测问题转变成关键点估计问题，提出了 CenterNet 网络，能够对高精度检测尺度大于 13×13 的清晰目标进行检测，但对弱小目标检测能力不足。

尽管学术界与工业界已经提出了许多红外舰船目标检测模型，但是仍然存在一些挑战亟需解决，如复杂背景下的目标检测和弱小目标检测。基于深度学习的目标检测算法主要针对自然图像，红外目标具有一些与自然图像明显不同的特性。具体来说，由于大气湍流和气溶胶散射等复杂成像环境的影响，导致红外舰船目标图像退化，目标与背景的红外辐射特性及其相似，鱼鳞波也呈现出忽明忽暗的纹理特征，增加了目标检测的难度^[13]。图1展示了红外场景下的船舶图像。船舶的位置距离红外传感器 20km，其图1(a)中目标特性较暗，且与背景中的鱼鳞波类似；图1(b)只有中间区域较亮、船身对比度弱，呈现小目标特性。现有的目标检测算法无法直接用于红外船舶目标检测。

为了提升复杂场景下船舶目标检测的精度，本文设计了一种有效且稳定的单阶段舰船检测网络，该网络由3个模块组成。首先，为了提取上下文特征信息，特征优化模块被引入来增强小目标的表征能力；然后，采用特征金字塔融合模块将获得的上下文特征与多尺度特征进行融合，增强特征的判别能力；最后，上下文增强模块则平衡前景和背景的不均衡差异，增

强目标的区分能力。实验结果表明，本文所提的网络对小目标有较准确的检测效果，具有较高的有效性和鲁棒性。



(a) 小目标

(a) Small object



(b) 低对比度

(b) Low-contrast

图1 复杂红外场景下的船舶示例

Fig.1 Examples of infrared ship in complex infrared scenes

1 CenterNet 网络

假定红外图像 $I \in R^{W \times H}$ 经过一系列的变换后能够

获得关键点热图 $H \in R^{\frac{W}{R} \times \frac{H}{R} \times C}$ ，其中 R 与 C 分别表示步幅与关键点类型数量。 $H_{x,y,c} = 1$ 表示图像检测到的关键点，而 $H_{x,y,c} = 0$ 则表示背景。为了有效地获得关键点热图 H ，一般采用全卷积编解码网络，如残差网络 (ResNet)、深层聚合网络 (deep layer aggregation, DLA)。在 CenterNet 网络中，对于第 c 类目标的基准关键点 $p \in R^2$ ，其对应的低分辨率图像的关键点位置为 $\hat{p} = \lfloor p/R \rfloor$ 。按照文献[14]所示，所有基准点可以通过高斯核投影到热图 H 上，即 $H_{x,y,c} = \exp(-((x - \hat{p}_x)^2 + (y - \hat{p}_y)^2) / 2\sigma_p^2)$ ，其中 σ_p 是与目标大小相关的标准差。如果相同类别的目标高斯函数重叠，则采用最大像素级融合。为了解决单阶段

检测网络中正负样本比例严重失衡的问题,采用基于像素级逻辑回归的交叉熵损失函数进行训练,其表达式如下所示:

$$L_k = \frac{-1}{N} \sum_{x,y,c} \begin{cases} (1-\hat{H}_{x,y,c})^\alpha \log(\hat{H}_{x,y,c}), & \text{if } H_{x,y,c} = 1 \\ (1-H_{x,y,c})^\beta (\hat{H}_{x,y,c})^\alpha \log(1-\hat{H}_{x,y,c}), & \text{others} \end{cases} \quad (1)$$

式中: α 与 β 是损失函数的超参数; N 是图像 I 中的关键点的个数,其目的是对损失函数归一化。由于基准关键点会因下采样导致偏差,CenterNet网络在目标中心点附近增加了局部自适应偏置项 $d_{x,y,c}$ 。所有类别 c 共享相同的偏置预测模型,并采用 l_1 损失函数进行训练。

$$L_{\text{offset}} = \frac{1}{N} \sum_p \left| d_{\hat{p}} - \left(\frac{P}{R} - \hat{p} \right) \right| \quad (2)$$

通过对关键点 \hat{p} 处损失项进行训练,最终可以获得不同关键点的最优偏置。目标检测过程中,假定目标 k 的边界框坐标是 $(x_1^k, y_1^k, x_2^k, y_2^k)$,其中心点可以表示为 $0.5 \times (x_1^k + x_2^k, y_1^k + y_2^k)$ 。若采用关键点来预测目标的中心点,其目标的大小 $s_k = (x_1^k - x_2^k, y_1^k - y_2^k)$ 可以通过回归获取。为了降低回归过程的计算复杂度,每种目标采用相同尺度 \hat{S} 进行预测,其尺度损失函数可以表示为:

$$L_s = \frac{1}{N} \sum_{k=1}^N \left| \hat{S}_{p_k} - s_k \right| \quad (3)$$

因此,整个网络的最终损失函数可以通过加权获取,即:

$$L = L_k + \lambda_1 L_{\text{offset}} + \lambda_2 L_s \quad (4)$$

式中: λ_1, λ_2 是不同损失项的权值。整个网络共同训练获得关键点 \hat{H} 、偏置 \hat{d}_x, \hat{d}_y 与尺度大小 \hat{S} 。因此,每个位置可以获得 $4 \times C$ 个输出值,所有输出共享一个全卷积主干网,并通过 3×3 卷积、ReLU和 1×1 卷积得到热力图。目标检测过程中,分别提取热图上每个目标类别的峰值点,最终得到目标检测结果。测试结果表明,CenterNet算法也存在不足:①在单个特征图上进行中心点检测,若多个舰船目标重叠的情况下,中心点容易挤到一起,导致最终目标无法区分;②生成的高斯权重只能检测尺度较大的目标,对弱小目标检测能力不足;③对遮挡严重的目标区域,其检测性能较差。

2 改进的单阶段舰船检测网络

众所周知,目标检测与目标分类所需要的主干网是不一样的,目标分类需要更深的特征以增强目标特征的区分度,而目标检测应该着重于浅层特征以获取丰富的位置信息。对于海洋环境中舰船目标检测,所设计主干网络需要同时兼顾浅层特征与深层特征,其中前者具有较大的感受野,位置信息丰富,而后的区分度更大。

由于远距离成像的红外图像含有大量小目标,固定尺度的感受野不能有效检测出小目标。不同于在主干网络增加反卷积层用于增大目标检测中的语义信息,本文采用特征精细模块从低分辨下采样图像中获取弱小目标的上下文信息,通过使用多尺度卷积层提取多尺度上下文信息,其网络结构如图2所示。首先,下采样图像经过大小为 3×3 与 1×1 的卷积核处理后可以得到多个不同尺度的特征图;为了提取用于小目标检测的多尺度上下文信息,3个并行的 1×1 卷积层将特征图的数量降低到 $C/3k$,然后在各自结果上增加一个大小为 3×3 的空洞卷积,可以增加感受野的范围,以便显著地改善目标检测的性能,并在小目标检测中发挥积极的作用;最后,引入一个特征级联算法对所有尺度的特征进行细化,并增加一个 3×3 卷积层来降低级联操作引起的特征混叠问题,从而确保提取的深度特征更稳定、表征能力更强。

本文采用的特征精细模块通过不同扩张率的并行分支来改变感受野的范围,从而提取多尺度上下文信息,有助于小目标检测。3个扩展卷积的参数分别是 $(k_s=3, p=1, d=0)$, $(k_s=3, p=2, d=2)$ 与 $(k_s=3, p=4, d=4)$,其中 k_s, p 和 d 分别表示卷积核尺度、填充范围与扩张率。不同的填充率和扩张率意味着不同的感受野。直观上,空洞卷积可以在不损失目标细节特征的情况下增加感受野,有利于小目标检测,3个并行的空洞卷积分支可以将低层次的特征信息转化为高表征能力的上下文信息,也有助于低分辨图像中缩小目标检测。

实际上,探测器在飞行过程中获得的红外图像极易被太阳光线、云层、风浪等干扰影响,仅仅靠提升弱小目标的检测能力,并不能增加目标的检测精度,反而会引入大量虚警。为了有效整合低层次特征与高层次特征,本文采用特征金字塔融合模块来耦合不同尺度特征信息,通过基于正则化的稠密连接特征金字塔可以将不同尺度的特征信息转换到相同尺度,最终融合得到高层次语义信息。

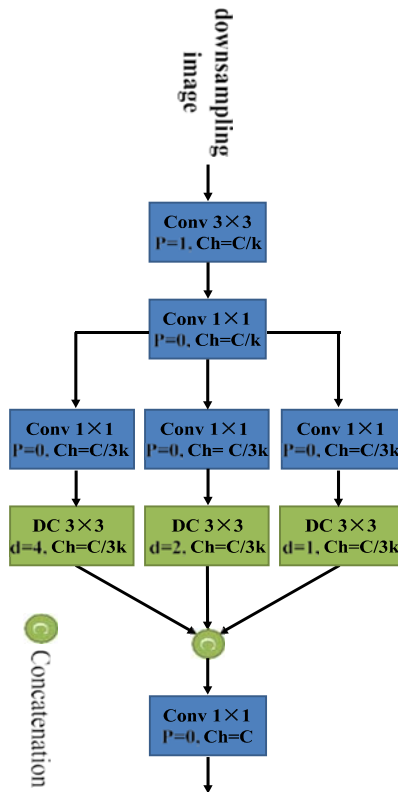


图2 多尺度上下文特征

Fig.2 Framework for multi-scale context feature

假定特征金字塔融合模块中多尺度金字塔特征表示为 $\{C_1, C_2, C_3, C_5\}$ ，图2展示了自顶向下特征金字塔融合过程。首先，深度特征金字塔分解可以采用可变形卷积提取多尺度特征；然后通过下采样操作将获得的多尺度特征正则化到与 C_5 相同的尺度，并与相应的特征映射进行融合，最后通过另一个可变形卷积来提升特征的表征；最终得到正则化的特征 $\{P_2, P_3, P_4, P_5\}$ 。为了实现对多尺度特征信息的融合，增强目标的表征能力，本文首先需要对特征映射 $\{P_2, P_3, P_4, P_5\}$ 进行级联操作，然后将此特征与精细化特征进行融合，作为单阶段目标回归模块的输入参数。此外，受残差学习形式的启发，本文将精细化特征与高层次特征 P_2 进行融合，并在其基础上进行回归与分类。

挂飞获取的红外舰船图像异常复杂，海面上漂流的杂物、鱼鳞波等干扰极易增加检测的虚警。一般来说，特征金字塔融合模块提取了具有较强鉴别能力的特征，有助于提升复杂背景中目标的表征能力；同时，本网络采用的可变形卷积模块可以处理任何几何变换问题，具有较高的泛化能力。然而，这些方法对于简单背景下的弱小目标具有较好的性能，但不能有效增强复杂背景的抗干扰能力。因此，本文所提网络在特征金字塔融合模块的基础上引入了上下文增强模块，有助于平衡前景和背景的不均衡差异，增强目标

的区分能力。

本文引用的上下文增强模块仅仅使用目标分类特征，中心点与边界框尺度参数进行训练，该结构本质上是单级检测器。假定上下文增强模块得到的分类特征为 f ，目标中心点与边界框的偏移参数为 $d_c=(\sigma_x, \sigma_y)$ 与 $d_b=(w, h)$ 。当分类特征 f 中 $H_{x,y,c}=1$ 表示像素 (x,y) 为目标，而 $H_{x,y,c}=0$ 则表示背景。也就是说，像素点 (x,y) 为目标中心点偏差为 (σ_x, σ_y) ，边界框的半径为 $(w/2, h/2)$ 。那么目标的边界框坐标可以表示为 $(x + \sigma_x - w_x/2, y + \sigma_y - w_y/2, x + \sigma_x + w_x/2, y + \sigma_y + w_y/2)$ 。受非局部网络启发，局部信息聚合有助于获取更大范围的相关信息。因此，本文也通过聚合操作，将目标语义特征 S 耦合到分类特征 f 中，得到目标的特征表示 $S \otimes f + S$ ；然后利用 1×1 卷积 $C_{1 \times 1}(x,y)$ 操作压缩位置点 (x,y) 的通道维数，并对获得的特征进行融合，并采用可变形卷积获取任意不规则目标的特征。因此，上下文增强模块获得的最优特征可以表示为 $DCN(S \otimes f + S + C_{1 \times 1}(x,y))$ 。

3 训练函数

与文献[15]相似，本文所设计网络损失函数仍然采用等式(4)所示的加权策略。分类损失项 L_k 采用改进的交叉熵损失函数，其表达式如下：

$$L_k = -(\text{conf}(x, y) - 1)^\alpha \log(\text{conf}(x, y)) \quad (5)$$

式中： $\text{conf}(x, y)$ 表示像素点 (x,y) 处的置信度； α 是控制正负样本对总体样本损失的共享权重，一般设置为2。

由于本文的应用场景是对舰船目标进行检测，目标类别单一。与等式(2)类似，中心点偏差损失项与尺度损失项采用L1损失函数。假定目标基准边界框的4个参数可以表示 (x_1, y_1, x_2, y_2) ，其中 (x_1, y_1) 与 (x_2, y_2) 分别表示左上与右下坐标点，即：目标的长宽分别可以表示为 $w=(x_2-x_1)$ 与 $h=(y_2-y_1)$ 。因此，中心点偏差损失项 L_s 可以改写为：

$$L_s = \frac{1}{N} \sum_{i=1}^N (|\hat{w}_i - w| + |\hat{h}_i - h_i|) \quad (6)$$

式中： \hat{w} 与 \hat{h} 分别表示目标尺度的预测值； N 是图像中选择的疑似目标数量。众所周知，卷积层在前向传播过程中下采样会产生特征偏移。为了解决这个问题，本文首先将学习到的锚框与特征图中对应尺度进行计算获得中心点偏差，然后采用可变形卷积进行对齐，并回归分类。因此，中心点偏差损失项 L_{offset} 可以表示为：

$$L_{\text{offset}} = \frac{1}{N} \sum_{i=1}^N \left(\Delta p - \left(\frac{p}{R} - \left\lfloor \frac{p}{R} \right\rfloor \right) \right) \quad (7)$$

与等式(2)类似, R 表示下采样系数, 一般设置为 2。

4 实验结果及其分析

4.1 数据集

目前, 国际上对船舶的检测研究主要集中在 SAR 图像领域。例如, 中科院公开的 SAR 图像船舶检测数据集, 包含多源、多模式 SAR 图像。红外探测与成像技术广泛地应用于海洋舰船监测与制导, 积累了大量的红外船舶目标图像。虽然大数据时代背景下数据共享逐渐成为趋势, 然而大多数红外船舶图像数据集并没有开源, 形成信息孤岛。本文所采用的数据集是在横向课题的资助下, 利用浙江大立生产的制冷热像仪采集了大量港口、近海、远海的红外数字视频, 红外图像大小为 640×512 , 并对其中的舰船目标进行了标注, 部分数据如图 3 所示。目标的大小从 10×10 到 540×248 不等。为了便于分析, 所有的标注数据集都归一化相同的尺度, 并采用随机裁剪、旋转等数据增强进行数据扩增。最终本文所选择的数据集包含 23750 幅图像, 随机选取其中 90% 作为训练集, 10% 作为测试集。



图3 数据集示例

Fig.3 Example for self-built dataset

4.2 参数设置及其评价指标

本实验采用 PyTorch 框架来搭建本文所提的检测网络, 硬件环境为: i5 8250@3.10GHz (X4), 32 GB (DDR3 1600 MHz), Nvidia GeForce GTX 1080 Ti, Ubuntu16.04, 64 位操作系统。损失函数中 λ_1 与 λ_2 分别设置为 0.1 与 1; 下采样系数 R 设置为 2; 指数因子 α 一般设置为 2。总损失函数采用 Adam 算法进行优化, 其中动量为 0.9, 权值衰减为 0.0001, 批量大小 (batch_size) 为 32。训练过程中, 学习速率初始设置为 1.25×10^{-4} , 然后在训练到第 50 个 Epoch 时, 学习速率改为 1.25×10^{-5} ; 若 100 个 Epoch 后, 整个网络

的损失值不超过迭代阈值则停止训练。

为了评估所提网络的检测性能, 本实验沿用了图像分类任务中精确率 (Precision, P)、召回率 (Recall, R)、 F_1 、mAP 指标进行评价, 并根据检测边框与基准框的交并比重新定义了正负样本的划分标准, 所有指标的计算方法参见文献[16]。

4.3 消融分析

本文所提的网络是在 CenterNet 的基础上进行改进。首先, 提取多尺度上下文信息 (multi-scale context information, MCI), 进一步细化和增强顶层特征输入特性, 增强弱小目标检测性能; 然后利用特征金字塔融合模块生成表征能力更强的语义信息 (semantic information, SI); 最后将多尺度上下文信息与语义信息进行融合 (Fusion), 增强网络特征表达能力, 以降低复杂背景对检测性影响, 平衡前景和背景的不均衡差异, 消除鱼鳞波的影响。

为了分析不同改进模块对整体检测性能的影响, 表 1 展示了不同改进模块的消融分析结果, 其中基于深层聚合的 CenterNet 网络作为对比基线。与基线结果相比, MCI 特征提取模块细化和增强顶层特征输入特性, 有助于检测小型船舶。对比结果也表明, MCI 特征将 P , mAP, R 与 F_1 分别提高了 3.4%, 0.6%, 0.5% 与 0.1%; SI 特征提取模块将 P , mAP, R 与 F_1 分别提高了 3.7%, 1.3%, 0.3% 与 0.6%, 这表明该模块将低层次特征与高层次语义特征进行了融合, 更加有助于表征具有丰富语义信息的目标。上下文增强融合模块将 P , mAP, R 与 F_1 分别提升了 2.3%, 1%, 1.5% 与 1.6%, 这表明该模块增强了目标的表征能力, 并削弱了背景的影响, 有助于更准确地检测目标。

表1 不同模块的消融结果

MCI	SI	Fusion	P	mAP	R	F_1
			71.1	76.3	82.7	86.5
√			74.5	76.9	83.2	86.6
√	√		78.2	78.2	83.5	87.2
√	√	√	80.5	79.2	85.0	88.8

4.4 定性定量实验结果

为了对本文所提出的单阶段舰船检测网络进行定性定量性能分析, 实验选择了 5 个最新的目标检测算法和本文算法进行对比分析, 如 YOLO-V3^[5], RetinaNet^[17], RefineNet^[18], CenterNet^[9] 与 FCOS (fully convolutional one-stage object detection)^[19], 其中 CenterNet 是整个对比实验的基线。与基线模型相比, 本文改进的网络主要是对特征优化模块, 特征金字塔

融合模块和上下文增强模块进行改进，但其引入的计算开销可以忽略不计。

RetinaNet 是一种单阶段网络，使用 FPN (feature pyramid networks) 结构对多个降维特征层进行处理，并对分类支路和回归支路并行优化；RefineDet 是一种仿二段式检测器，通过启发式方法来避免类别不平衡问题，并通过两个级联的检测器使预测框更加准确；FCOS 算法是无锚框检测算法，借鉴了 FCN 的思想，对每个像素进行直接预测。

表 2 展示了不同对比算法下的定量性能对比。从整体上看，本文改进的网络性能稳定，对不同场景的适应性较好。与基线 CenterNet 的结果相比，改进的算法具有明显提升，其中 P , mAP , R 与 F_1 分别提高了 3.4%, 0.6%, 0.5% 与 0.1%。红外探测器在 25 km 发现的目标呈现出弱小特征，本文改进方法利用多尺度上下文信息进一步细化和增强顶层特征输入特性，有助于小型舰船目标的检测。

表 2 自建数据集上的检测结果对比

Table 2 Comparison of results on non-public data sets

Models	P	mAP	R	F_1
YOLOv3	75.5	74.2	81.3	83.9
RetinaNet	77.3	80.6	78.9	77.4
RefineNet	78.4	83.1	79.3	81.1
CenterNet	77.1	78.6	84.5	88.7
FCOS	78.7	85.1	76.6	86.5
Ours	80.5	79.2	85.0	88.8

本文所提网络主要针对舰船目标进行检测，但统

计整体性能并不能对算法性能进行全面分析。因此，项目组将 4.1 节介绍的自建数据集分成了小目标数据子集 (small object sub-set, SOS)、复杂背景数据子集 (complex background sub-set, CBS) 与其他数据子集 (Others)，其中 CBS 子集主要是太阳迎头带来的光线干扰、天地线干扰、海岸线干扰、鱼鳞波干扰等。深度网络是一种模拟人类视觉的推理网络^[20]，其检测的弱小目标并不是点目标，而是具备一些轮廓特征。从实验结果可以看出，所提网络的 mAP 比 RetinaNet 高 4.6%，比 RefineNet 和 FCOS 高 2.9% 与 8.9%，主要是由于改进网络采用自顶向下的特征连接，削弱了复杂背景的权重，而其他模型不可避免会遭受噪声的干扰。

YOLOv3 是目前工程上应用较多的单级网络，但其不能有效地检测小型目标。对于 CBC 数据子集，YOLOv3 通过采用跨连接和多尺度预测功能实现了 80.6% 的 mAP ，本文的方法有最好的准确性，比 YOLOv3 高 4.5%。从表 3 的结果也可以看出，相比于基线模型 CenterNet，所提网络在 Others 数据子集上的 P , mAP , R 与 F_1 分别提高了 4.2%, 6.5%, 7.5% 和 0.5%。

本文网络在红外 CBC 数据集上获得了最佳的检测精度，85.2% 的 mAP 和 90.1% 的 F_1 评分。这表明该方法能够提取出强大的特征表示，用于复杂背景下的舰船检测。值得注意的是单级检测网络性能优于两级检测网络，可以采用各种策略来提高准确性^[19]。

表 3 不同数据子集上的检测结果对比

Table 3 Comparison results for different sub-set

Models	SOS				CBC				Others			
	P	mAP	R	F_1	P	mAP	R	F_1	P	mAP	R	F_1
YOLOv3	67.3	67.4	70.6	68.9	72.1	80.6	83.0	88.3	76.4	85.1	81.5	76.0
RetinaNet	66.6	70.3	72.5	69.4	75.4	81.1	83.1	83.3	78.5	84.5	85.8	76.6
RefineNet	64.8	78.8	78.3	70.9	73.7	82.3	85.0	89.1	77.2	89.6	86.4	75.7
CenterNet	67.8	74.6	79.6	73.2	73.6	77.1	81.9	93.0	79.3	78.9	80.1	85.5
FCOS	64.8	80.8	78.3	70.9	72.5	76.3	82.4	86.2	78.7	77.7	78.5	85.9
Ours	68.0	83.3	83.6	74.9	73.9	85.2	84.9	90.1	83.5	85.4	87.6	86.0

图 4 展示了对比算法检测结果。本文选用了 5 张比较有代表性的图像进行检测。第一行是一张背景复杂目标对比度较低的图像，其中水波呈现的高亮区域也容易造成虚检。从实验结果可以看出，对比算法很容易检测出近处的大目标，而远处的小目标则出现漏检。虽然 FCOS 检测出小目标，但以边界框与目标实际尺寸相差较大。从实验结果分析得知，FCOS 将每个像素看作训练样本，图像中的虚假目标容易干扰检测性能，最终是将小目标附近的高亮背景当作了目

标。第二行是一张带有海岸线背景的红外图像，其目标大多停靠在岸边。由于本文算法与对比算法都不具备海岸线检测的能力，无法排除陆地上的干扰。所提的网络：RetinaNet, RefineNet 与 CenterNet 都找到了陆地上的假目标，但对于海洋上的目标，本文算法性能明显优于对比算法最优网络。FCOS 网络不是只将边界框看作样本，与语义分割的全卷积相似，通过预测每一个像素类别，进而实现目标检测。YOLO-v3 是所有对比算法中结果最差的，但在这幅图像中 YOLO-

v3 能够检测出所有海面船舶目标。第三行是图像中天际线附近存在目标,由于目标较远,天际线附近目标很难发现。从实验结果可以看出,对比算法都没能找到目标,本文算法虽然找到了目标,但也找到假目标。从图上结果也可以看出,若没有先验知识,虚假目标确实容易造成误判,这就需要长期检测并增强目标的识别概率。第四行中图像存在多个目标且相互重叠,只能靠智能推理实现目标检测与区分,然而实际结果中所有算法都没能正确找到目标,本文算法找到了一个大目标,但遮挡目标只标示出来一小部分。

4.5 挂飞数据分析

为了验证算法的性能与适应性,本小节将对实际

挂飞数据检测结果进行分析。在挂飞验证视频中,目标在 28 km 处出现在视场内时,目标很弱,本文算法不能检测出疑似目标。当目标的距离增加时,目标的轮廓明显,最终检测出目标。由于目标机动运行,探测器获得的目标在转弯时,目标的尺度缩小,此时本文算法出现了漏检,但最终目标重现时,目标又被重新检测出来。图 5 展示了本文算法对一段红外挂飞视频进行目标检测的结果。值得注意的是,图 5 右下角图形受光照影响较大,本文算法也能检测到目标。整个项目在工程化研制阶段,目标的检测性能较稳定。充分验证了本文算法在工程应用中的有效性。

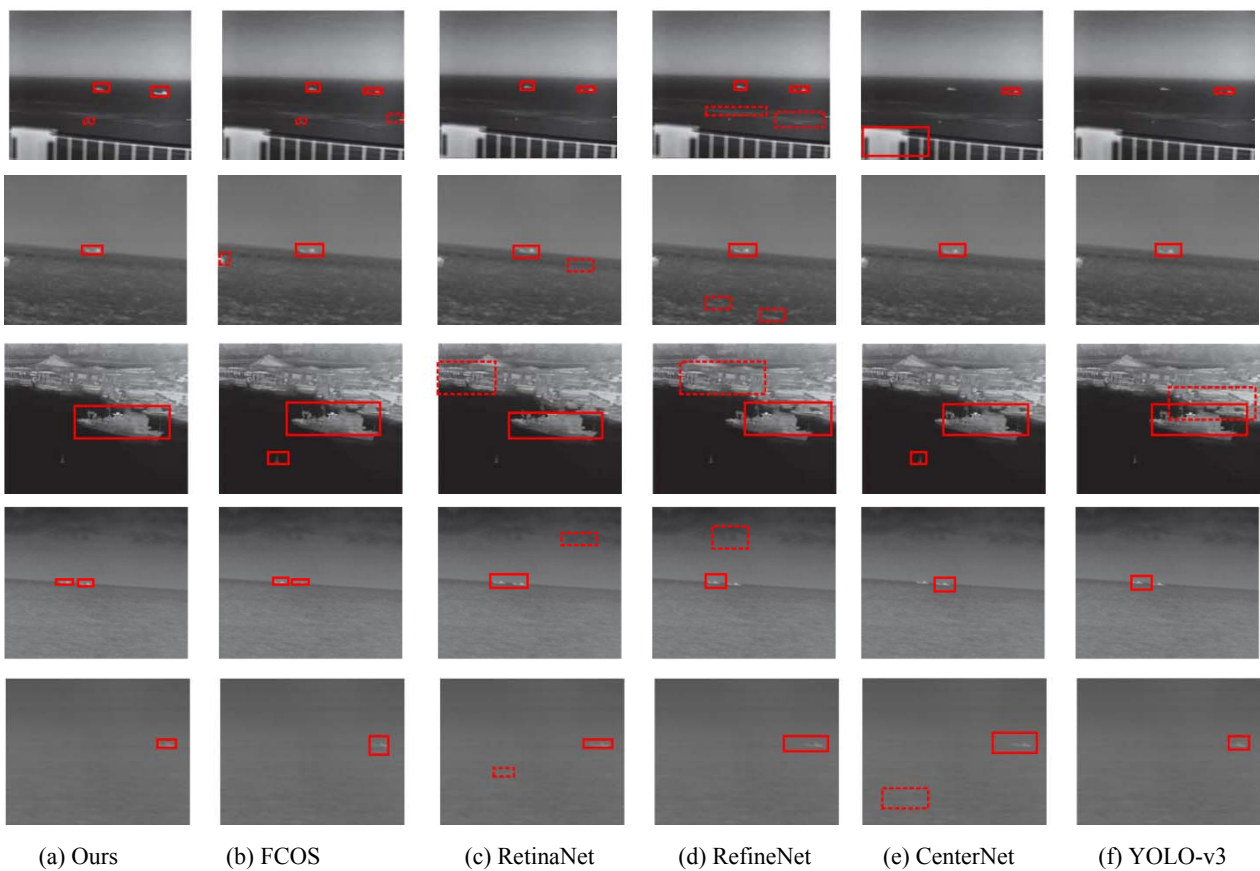


图 4 不同算法定性结果对比

Fig.4 Comparison of qualitative results for different algorithms

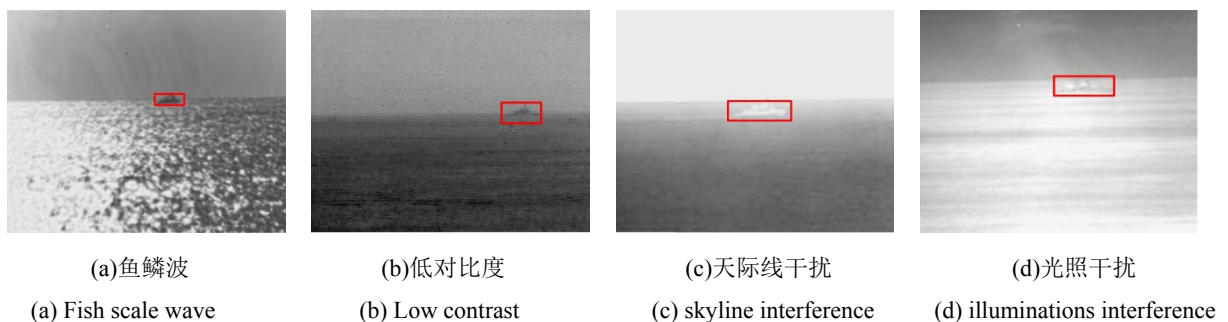


图 5 不同场景下的工程验证结果

Fig.5 Analysis of engineering results for different scenarios

5 结论

本文提出了一种改进的单阶段舰船检测网络,该网络主要由特征优化模块,特征金字塔融合模块和上下文增强模块组成。为了验证本文所提方法的有效性和鲁棒性,本文对自建的舰船数据集进行了定性定量验证,结果表明,所提网络在自建数据集上均达到了最优性能,具有较强抗干扰能力,但对于类别差异较大的船舶,所提方法的泛化性能仍然存在不足。同时,本文所提算法不具备海岸线检测的能力,无法排除陆地上的干扰。这些不足将是项目组下一步工作计划。

参考文献:

- [1] 邢莎,吉林,雍杨,等.复杂场景下的自动红外舰船目标检测[J].*红外技术*,2014,**36**(4):320-325.
XING Sha,JI Lin,YONG Yang,et al. Automatic infrared warship target detection under complicated background[J].*Infrared Technology*,2014,**36**(4):320-325.
- [2] 赵文涛,曹昕鸫,田志勇.基于自适应阈值区域生长的红外舰船目标分割方法[J].*红外技术*,2018,**40**(2):158-163.
ZHAO Wentao,CAO Xinzhong,TIAN Zhiyong. An infrared ship target segmentation method based on adaptive threshold region growth[J].*Infrared Technology*,2018,**40**(2):158-163.
- [3] 丁荣莉,韩传钊,谢宝蓉,等.红外遥感图像舰船目标检测[J].*红外技术*,2019,**41**(2):127-133.
DING Rongli,HAN Chuanzhao,XIE Baorong,et al. Ship target detection in infrared remote-sensing image[J].*Infrared Technology*,2019,**41**(2):127-133.
- [4] Kumar D,ZHANG X. Ship detection based on faster R-CNN in SAR imagery by anchor box optimization[C]//2019 *International Conference on Control, Automation and Information Sciences (ICCAIS)*,2019:309-313.
- [5] Redmon J,Farhadi A. YOLOv3: an incremental improvement[J]. arXiv:1804.02767,2018.
- [6] LIU W,Anguelov D,Erhan D,et al. SSD: single shot multibox detector[C]//*European Conference on Computer Vision*,2016:21-37.
- [7] LIN T Y,Goyal P,Girshick R,et al. Focal loss for dense object detection[C]//*International Conference on Computer Vision*,2017:2999-3007.
- [8] LAW H,DENG J. CornerNet: detecting objects as paired keypoints[C]//*European Conference on Computer Vision*,2018:765-781.
- [9] DUAN K,BAI S,XIE L,et al. CenterNet: keypoint triplets for object detection[J/OL]. arXiv:1904.08189,2019.
- [10] TAN M,PANG R,LE Q V,et al. EfficientDet: scalable and efficient object detection[J/OL]. arXiv:1911.09070,2019.
- [11] 杨涛,戴军,吴钟建,等.基于深度学习的红外舰船目标识别[J].*红外技术*,2020,**42**(5):426-433.
YANG Tao,DAI Jun,WU Zhongjian,et al. Target recognition of infrared ship based on deep learning[J].*Infrared Technology*,2020,**42**(5):426-433.
- [12] 黄洁,姜志国,张浩鹏,等.基于卷积神经网络的遥感图像舰船目标检测[J].*北京航空航天大学学报*,2017(9):132-139.
HUANG Jie,JIANG Zhiguo,ZHANG Haopeng,et al. Ship target detection in remote sensing image based on convolutional neural network[J].*Journal of Beijing University of Aeronautics and Astronautics*,2017(9):132-139.
- [13] 顾佼佼,李炳臻,刘克,等.基于改进 Faster R-CNN 的红外舰船目标检测算法[J].*红外技术*,2021,**43**(2):170-178.
GU Jiaojiao,LI Bingzhen,LIU Ke,et al. Infrared ship target detection algorithm based on improved faster R-CNN[J].*Infrared Technology*,2021,**43**(2):170-178.
- [14] GUO Haoyuan,YANG Xi,WANG Nannan,et al. A CenterNet++ model for ship detection in SAR images[J].*Pattern Recognition*,2021,**112**(88):25-34.
- [15] FAN S,ZHU F,CHEN S,et al. FII-CenterNet: an anchor-free detector with foreground attention for traffic object detection[J].*IEEE Transactions on Vehicular Technology*,2021,**120**(99):1-14.
- [16] Barrios J M,Bustos B. Competitive content-based video copy detection using global descriptors[J].*Multimedia Tools and Applications*,2013,**62**(1):75-110.
- [17] ZHANG Xiang,YANG Wei,TANG Xiaolin,et al. A fast learning method for accurate and robust lane detection using two-stage feature extraction with YOLOv3[J].*Sensors*,2018,**18**(12):4308-4315.
- [18] LIN G,Milan A,SHEN C,et al. RefineNet: multi-path refinement networks for high-resolution semantic segmentation[C]//2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,2017:1289-1301.
- [19] TIAN Z,SHEN C,CHEN H,et al. FCOS: fully convolutional one-stage object detection[C]//2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*,2020:1099e3233.
- [20] LIU Z,ZHANG X,JIANG T,et al. Infrared salient object detection based on global guided lightweight non-local deep features[J].*Infrared Physics & Technology*,2021,**12**(3):2309-2315.