

文章编号: 1002-2082 (2021) 04-0643-08

复杂背景下红外图像弱小目标检测

蔡 伟, 徐佩伟, 杨志勇, 蒋昕昊, 姜 波

(火箭军工程大学 兵器发射理论与技术国家重点学科实验室, 陕西 西安 710025)

摘 要: 红外弱小目标检测是安防监控、侦察探测、精确制导等领域的关键技术。为了提高复杂背景条件下红外弱小目标检测的准确性和实时性, 提出了一种基于深度学习的红外弱小目标检测算法 YOLO-FCSP。根据红外图像中弱小目标的特点, 在 YOLO 检测框架的基础上, 通过减少下采样次数, 结合跨阶段局部模块、Focus 结构和空间金字塔池化结构设计了特征提取网络。借鉴多路径聚合的思路优化特征融合网络, 同时调整检测输出层数量, 通过信息复用提高特征利用效率。实验结果表明, 本文提出的算法在检测红外弱小目标时具有较高的准确率和检测速度, 精度和召回率分别为 91.9% 和 94.6%, 平均准确率 (AP) 值达到 92.6%, 检测速度达到 170 f/s, 满足实际应用中实时检测的需求。

关键词: 深度学习; 红外弱小目标; 跨阶段局部网络; 多路径聚合; 实时检测

中图分类号: TN251; TP391.4

文献标志码: A

DOI: 10.5768/JAO202142.0402002

Dim-small targets detection of infrared images in complex backgrounds

CAI Wei, XU Peiwei, YANG Zhiyong, JIANG Xinhao, JIANG Bo

(National Key Discipline Laboratory of Armament Launch Theory and Technology,
Rocket Force University of Engineering, Xi'an 710025, China)

Abstract: The infrared dim-small target detection is a key technology in the fields of security surveillance, reconnaissance detection and precision guidance. To improve the accuracy and real-time performance of infrared dim-small target detection under complex background conditions, an infrared dim-small target detection algorithm YOLO-FCSP based on deep learning was proposed. Considering the characteristics of dim-small targets in infrared images, the feature extraction network was designed based on the YOLO detection framework by reducing the number of downsampling and combining cross-stage local modules, Focus and spatial pyramidal pooling structure. The feature fusion network was improved by utilizing the idea of multi-path aggregation. The number of detection output layers was adjusted to enhance the reuse of feature information. The experimental results show that the proposed method has higher accuracy and detection speed in detecting infrared dim-small targets, achieving 91.9% precision and 94.6% recall, the average precision (AP) value to 92.6%, and the detection speed as 170 fps, which meet the requirements of real-time detection in practical applications.

Key words: deep learning; infrared dim-small targets; cross-stage local networks; multi-path aggregation; real-time detection

引言

由于红外成像系统可以全天候提供清晰图像, 且能够穿透雾、烟和其他大气条件等障碍物, 在民

用与军事领域都有着广泛应用。红外弱小目标检测作为目标识别与跟踪领域的重要内容, 近年来已成为研究热点^[1]。根据国际光学工程学会

收稿日期: 2021-02-02; 修回日期: 2021-05-06

基金项目: 国家自然科学基金 (61505254)

作者简介: 蔡伟 (1974—), 男, 博士, 教授, 主要从事光电防护方面的研究。E-mail: clwn@163.com

通信作者: 徐佩伟 (1996—), 男, 硕士, 主要从事深度学习、红外图像处理方面的研究。E-mail: pwxu_0551@163.com

(SPIE)的定义,把面积不大于 9×9 像素的红外目标称为红外弱小目标^[2-3]。红外弱小目标检测的研究对象通常为距离较远的机动目标——红外弱小目标,往往占有像素少,缺少颜色、形状和纹理等细节信息,检测难度较大。而且红外图像获取受环境因素和设备的影响,会产生背景噪声和固有噪声,难以区分噪声和弱小目标,使得红外弱小目标检测更加困难^[4-6]。

随着计算机视觉的不断发展,基于深度学习的目标检测算法由于其突出的学习和表征能力,已经广泛应用于各类目标检测任务中^[7]。李小宁等人^[8]提出一种用于检测车辆小目标的方法,采取浅层与深层特征融合的策略来提高检测效果,但在复杂背景下算法性能有待提高。受到可见光目标检测领域的启发,深度学习也被一些学者用于红外目标检测。Zhao 等人^[9]设计了一种用于红外弱小目标检测卷积神经网络 TBC-Net,使用语义约束提取目标特征,解决了背景和目標不平衡导致的特征学习困难问题。总的来说,基于深度学习的检测算法有更好的适应性和更高的准确性,但在复杂背景下弱小目标检测仍然存在虚警的情况。

本文提出一种基于深度学习的网络模型 YOLO-FCSP,对于不同复杂背景下的红外弱小目标检测,具有良好的准确性和实时性。首先,对 YOLOv3 网络采取减少下采样次数的操作,并借鉴残差学习和跨阶段局部网络构建特征提取网络,以保证弱小目标特征不丢失。其次,利用多路径聚合网络融合浅层和深层的特征,充分利用重要的目标细节信息,同时使用两个检测输出层负责检测小目标,提高了红外弱小目标的检测效果。

1 有关方法

为了提高目标检测的速度,J. Redmon 等人提出将目标检测任务视作回归问题的端到端检测算法 YOLO。由于 YOLO 集成了检测的全流程,运行速度快,但检测精度有待提高。原作者受锚框策略的启发提出了 YOLOv2,有效地平衡了检测速度和精度,但难以探测一些小目标。为有效解决小目标的检测问题,又进一步提出了改进版本 YOLOv3。该网络包含多个残差模块,能够减少梯度消失的问题,提高了检测精度。YOLOv3 与前一版本不同,有 9 个预测矩形框,并被分成了 3 种不

同尺度。此外,为了检测到更多的细节特征,用上采样的方式使特征图尺寸变大获取更多的深层语义信息。再对不同层融合,得到的特征图包含更丰富的特征信息,提高了对小目标的检测准确率。对于特征提取的骨干网络,借鉴残差网络设计了 Darknet-53 代替 Darknet-19,加深网络的层数,增强了特征提取能力。

由于 YOLOv3 网络层数较深和参数量巨大,检测速度有待提升,因此轻量级网络 YOLOv3-tiny 被提出。相较于原始 YOLOv3 网络,特征提取部分没有采用复杂的 Darknet-53,而是使用了池化操作和卷积简化网络结构,使检测速度得到大幅提升。特征融合部分将 3 个不同尺度的特征层减少至 2 个尺度,分别用于大、小目标的独立检测。YOLOv3-tiny 通过减少网络层数和检测分支,减少网络参数量和计算量,使得检测速度大幅提升,适合对实时性要求比较高的场景。但由于其网络深度较浅,高层次语义信息提取能力不足,网络检测精度受到一定影响。

2 YOLO-FCSP 网络

YOLO-FCSP 网络模型的设计主要从特征提取网络和特征融合网络 2 个方面进行考虑,网络结构如图 1 所示。在特征提取阶段,考虑弱小目标信息的完整性而使用 8 倍下采样,并通过使用 CSPRBlock 模块增强信息传递,以更好地适应红外弱小目标的检测任务。在特征融合阶段,使用多路径聚合^[10]对浅层空间位置信息和深层语义信息进行融合,进一步降低虚警,提高网络检测性能。

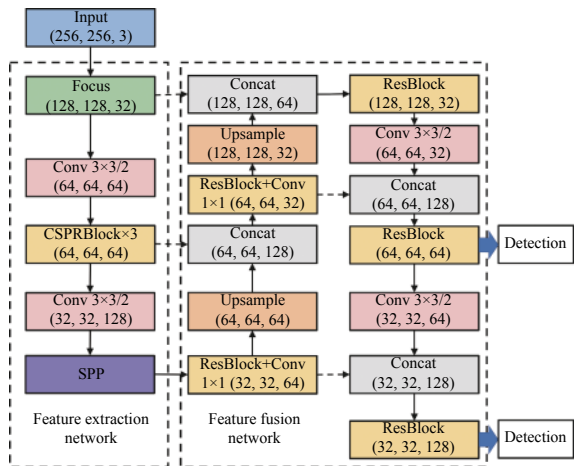


图 1 YOLO-FCSP 网络结构

Fig. 1 Structure of YOLO-FCSP network

2.1 特征提取网络

由于弱小目标在红外图像中像素少, 特征细节不明显, 过深的网络容易丢失特征信息, 加之网络模型参数量大而造成推理速度变慢, 因此相对较浅的网络更适合红外弱小目标的实时检测。此外, 下采样会使特征图缩小, 无法判定目标与原图像素的对应关系, 从而影响定位精度。为提高检测精度和速度, 将YOLOv3^[11]的32倍下采样调整至8倍, 避免弱小目标空间位置信息过度丢失。

考虑到下采样会导致目标部分信息丢失的问题, 在网络的第一层引入Focus结构, 如图2所示, 该结构对数据进行切片操作, 间隔采样得到4个特征图, 输入通道数扩充了4倍。然后将得到的特征图再经过卷积操作进行通道信息融合, 最终得到没有信息丢失的2倍下采样特征图。不仅减少了网络模型的参数量和计算量, 还将图像的宽度和高度信息整合到通道空间中去, 为后续的特征提取保留了更完整的图像信息, 能够进一步提高网络模型的检测性能。

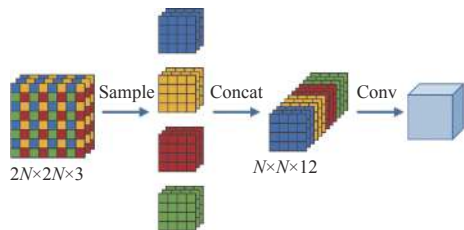


图2 Focus 结构
Fig. 2 Structure of Focus

受到跨阶段局部网络^[12]启发, 设计了网络组件 CSPRBlock 模块, 能够增强网络特征提取能力, 并减少模型参数量和推理时间。CSPRBlock 模块结构如图3所示, 由跳跃连接和主体路径两部分构

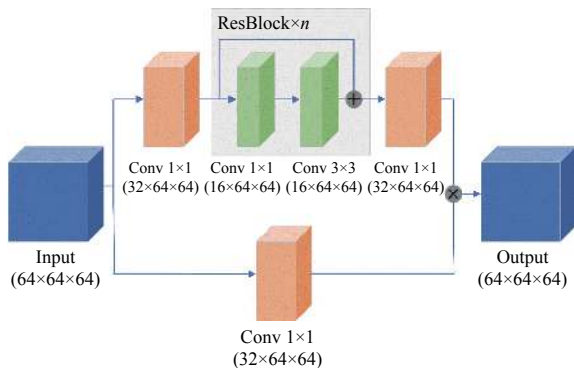


图3 CSPRBlock 模块
Fig. 3 CSPRBlock module

成, 主体路由由1×1卷积、n个堆叠的残差学习模块和1×1卷积组成, 其中2个1×1卷积作用是调整通道数实现降维, 降低网络参数量。跳跃连接部分经过一次点卷积跨层连接到网络的拼接层, 与主体部分拼接, 然后采用点卷积调整通道数与输入一致, 实现通道信息的融合。包含CSPResen的特征提取网络能够保证检测性能不下降, 且比YOLOv3的骨干网络Darknet拥有更少的参数量和计算量。

将空间金字塔池化 (spatial pyramid pooling, SPP) 作为特征提取网络的最后一层, 主要目的是分离上下文特征, 增加感受野, 便于后续融合全局特征信息, 改进后的SPP如图4所示。具体过程如下: 首先对输入特征进行卷积操作, 实现通道数减半; 然后分别进行不同大小的最大池化 (maxpooling), 池化核大小分别为5、9和13, 步长均为1; 最后将3个池化层和上一层的输出特征图合并, 输入到下一个卷积层进行特征学习, 得到更为丰富的局部特征。SPP结构几乎不会降低网络推理速度, 同时还能提高网络检测精度。

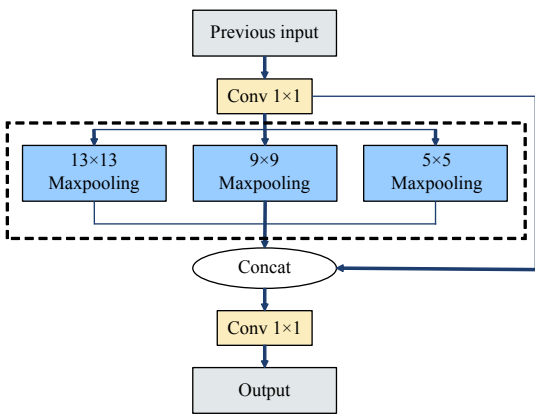


图4 改进的SPP模块
Fig. 4 Improved SPP module

2.2 特征融合网络

浅层特征图具有分辨率高、空间位置信息丰富的特点, 而较深的网络层特征图的细节信息少, 且存在像素错位的问题, 但后者能够高效地提取丰富的语义信息。为充分利用深层特征图语义信息, 同时不损失浅层特征图的空间位置信息, 需要将浅层与深层进行特征融合, 融合后的特征图兼具丰富的语义信息和位置信息。鉴于红外弱小目标的特殊性, 结合PANet优化特征融合网络, 充分融合浅层和深层的特征图, 增强各个网络层的信

息传递,形成自下而上和自上而下的网络路径,实现特征信息的重复利用,结构如图 1 所示。

特征融合网络的输入来自 3 个部分,其中 2 个来自特征提取网络的 2 倍和 4 倍下采样的特征图,另一个是经过 SPP 处理的 8 倍下采样特征图。特征融合网络的输出,又称检测头,是目标检测的输入。由于弱小目标特征细节少,特征图感受野越大,越容易将噪点和强干扰误认为目标。通过对比实验分析,去除 128×128 像素大小的检测输出层,保留 32×32 像素和 64×64 像素大小的检测输出层。

锚框(anchor boxes)是一组宽高固定的初始先验框,对目标的检测精度和速度有影响,故重新设置了检测头对应的初始锚框。根据红外弱小目标数据集集中的先验框,利用 K-means 聚类算法进行聚类,表 1 为检测输出层特征图的尺度参数。

表 1 检测输出层的尺度

Table 1 Scale of detected output layer

Scale size/pixel	Receptive fields	Anchor boxes
32×32	big	[2, 2], [3, 3], [4, 3]
64×64	small	[4, 4], [5, 5], [6, 6]

3 实验结果与分析

为验证本文提出的检测网络在复杂背景下红外弱小目标检测的检测效果,实验在不同背景下的红外弱小目标的数据集进行测试,设计消融实验证明网络模型所采用策略的有效性,并与其他典型的网络模型检测结果进行对比。

3.1 数据集

实验所采用的是红外序列图像中弱小飞机目标检测跟踪数据集^[13],由国防科技大学和中国航天科工二院合作制作。从中选取 5 个复杂背景下的红外序列图像作为实验数据集,图像尺寸大小为 256×256 像素,共 3299 张,利用 Labelme 软件对图像进行矩形框标注,将数据集按照 9:1 的比例分为训练集和测试集。图 5 展示了部分数据集图像,可以看出目标尺寸均小于 9×9 像素,符合红外弱小目标定义。表 2 为 5 个红外弱小目标序列详细情况,涵盖了地面、天空和地空交界多种复杂环境背景。

3.2 网络训练

选择 Python 作为实验编程语言,实验基于

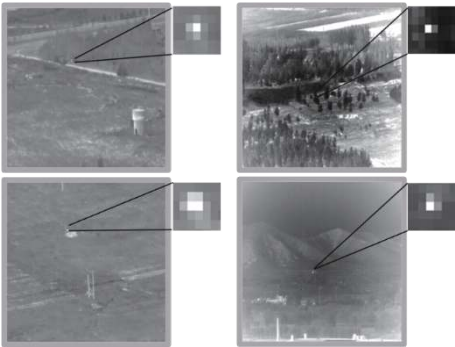


图 5 红外弱小目标数据集

Fig. 5 Infrared dim-small target dataset

表 2 5 个红外弱小目标序列详细情况

Table 2 Detailed information of 5 infrared dim-small target sequences

Sequences	Frame	Image size/pixel	Details of background
1	399	256×256	Ground background
2	1 500	256×256	Ground background
3	500	256×256	Ground background
4	400	256×256	Sky and ground background
5	500	256×256	Woodland background

Pytorch1.8, 训练和测试在 Pycharm 中实现,硬件平台配置见表 3。实验中,模型输入图像大小均为 256×256 像素,不进行压缩、扩大和裁剪等处理,采用马赛克增强来增强训练数据^[14],以保证网络模型对弱小目标的检测效果。

表 3 实验硬件平台配置

Table 3 Configuration of experimental hardware platform

Component	Hardware
CPU	intel i5-10600KF CPU
GPU	NVIDIA GTX 1080Ti
OS	Ubuntu20.02
Memory	16 GB

网络训练使用 Adam 作为优化器,将衰减权重设置为 0.0005,批次大小为 16,初始学习率设置为 0.01。采用余弦函数衰减法,学习率衰减至 0.002 左右,学习率曲线如图 6 所示。图 7 为模型训练的损失下降曲线,约在 500 次迭代后损失值稳定。图 8 为平均准确率曲线,检测精度随训练迭代次数增加而提升,最后基本保持稳定,其中 YOLO-FCSP

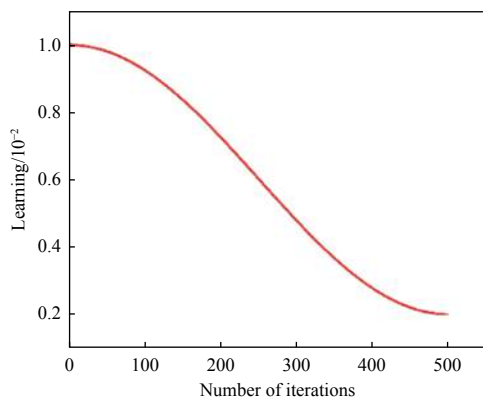


图6 学习率衰减曲线

Fig. 6 Decay curve of learning rate

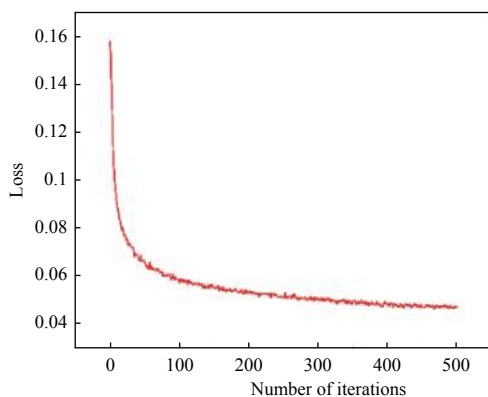


图7 损失值下降曲线

Fig. 7 Descent curve of loss value

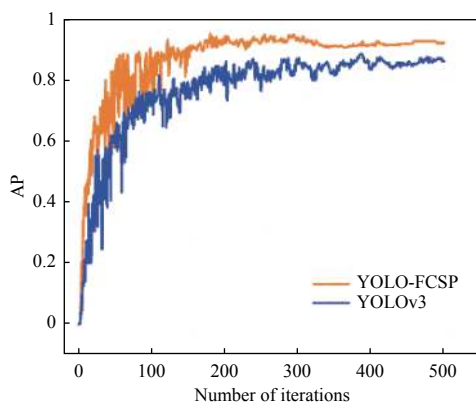


图8 平均准确率曲线

Fig. 8 Average precision curve

的平均准确率增幅比 YOLOv3 更快。根据训练曲线和检测精度情况,可以看出训练结果比较理想,证明了网络模型结构的有效性。

3.3 评估指标

为了评估红外弱小目标的检测效果,使用以下评估指标:

1) 精度、召回率和平均准确率

精度 (precision, P) 是被预测为正样本中实际为正样本的比例, 召回率 (recall, R), 也称查全率, 是实际的正样本被预测为正样本的比例, 两者公式如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

式中: TP 为目标被正确预测为正样本的数量; FP 为目标被错误预测为正样本的数量, 即误检; FN 为目标被错误预测为负样本的数量, 即漏检。

平均准确率 (average precision, AP) 指的是精度的平均值, 越大检测精度越高。其值是 P-R 曲线同坐标轴所围成的面积, 最大值为 1, 公式如下:

$$AP = \int p(r) dr \quad (3)$$

2) FPS

每秒帧数 (frame per second, FPS), 指的是每秒可以检测到的图像数量, 是用来评估检测速度的常用指标, FPS 值越大检测速度越快。

3) 参数量和计算量

参数量 (parameters) 对应于模型的空间复杂度, 占用显存的量, 包括模型总参数量和各层输出特征图大小。模型的参数量越多, 需要训练的数据量越大, 训练时容易导致过拟合。计算量 (GFLOPS) 又称为时间复杂度, 指网络模型完成一次前向传播所需要的总计算量, 通常用每秒浮点运算次数来表示。计算量越大, 训练和检测的时间越长。

3.4 结果分析

3.4.1 消融实验

为验证特征提取网络改进的有效性和可靠性, 将本文算法模型拆分成 5 组进行训练, 分别在 5 个不同背景下的序列图像上进行测试, 测试结果如表 4 所示。第 1 组为原始 YOLOv3 算法; 第 2 组为将 YOLOv3 的特征提取网络下采样次数由 32 倍改为 8 倍; 第 3 组在上一组的基础上加入 CSPRBlock 组件; 第 4 组在第 3 组的基础上加入改进的 Focus 结构; 第 5 组在第 4 组的基础上, 将改进的 SPP 模块加入特征提取网络的最后阶段。

由表 4 可以看出: 第 1 组各个序列的 AP 分别为 89.8%、87.6%、90.9%、81.8% 和 91.5%, FPS 均在 80 左右。由于 YOLOv3 网络模型层数较多, 参数量和计算量十分巨大, 分别为 61 M 和 154 GFLOPS。

表 4 消融实验结果

Table 4 Results of ablation experiments

Group	1/8	CSPRen	Focus	SPP	Para/M	GFLOPS	Sequence 1		Sequence 2		Sequence 3		Sequence 4		Sequence 5	
							AP/%	FPS	AP/%	FPS	AP/%	FPS	AP/%	FPS	AP/%	FPS
1					61	154.7	89.8	81	87.6	79	90.9	79	81.8	80	91.5	79
2	✓				0.55	22.8	87.6	153	84.5	151	86.4	152	80.2	158	89.6	152
3	✓	✓			0.41	13.1	91.3	154	90.5	153	88.8	154	86.1	156	91.9	155
4	✓	✓	✓		0.40	11.5	91.4	174	92.6	176	89.5	172	87.0	176	92.1	174
5	✓	✓	✓	✓	0.45	12.0	93.5	163	93.0	172	91.5	162	89.7	173	93.7	173

第 2 组考虑到红外弱小目标信号弱特征少, 调整网络 32 倍下采样至 8 倍, 模型参数量和计算量大幅降低, 且检测速度提升较大。与第 1 组相比, 5 个背景下检测的平均准确率均有所下降, 这与网络的深度变浅有关。第 3 组引入借鉴跨阶段局部网络设计的 CSPRBlock 模块, 各个序列的 AP 值均大幅上升, 网络模型参数量与第 2 组相比降低 20%, 计算量减少一半, 这是因为 CSPRBlock 能够增强网络学习能力, 有效消除网络存在的计算瓶颈和减少内存占用, 从而提高了检测网络的推理速度和准确性。第 4 组将第一层卷积下采样替换为 Focus 结构, 实验结果表明, 针对占有少量像素的小目标, Focus 结构可以在一定程度上减少模型参数量和计算量, 提升网络检测精度和速度。第 5 组是在第 4 组的基础上添加了改进的 SPP 结构, 网络参数量和计算量分别增加了 0.05 M 和 0.5 GFLOPS, FPS 小幅降低, 但各个序列的 AP 值分别提升了 2.0%、0.4%、1.9%、2.6% 和 1.5%。

为验证多路径聚合的特征融合对网络模型检测性能的提升, 将使用 FPN 融合策略的网络模型 YOLO-FCSF 和结合 PANet 的 YOLO-FCSP 算法进行对比, 实验结果如表 5 所示。

表 5 加入 PANet 的效果对比

Table 5 Effect comparison of add PANet

Model	P/%	R/%	AP/%	FPS
YOLO-FCSF	37.5	89.9	83.9	200
YOLO-FCSP	91.9	94.6	92.6	170

可以看出, 使用结合多路径聚合网络的方法有效地融合了不同层的特征, 精度、召回率和平均准确率分别提升了 54.4%、4.7% 和 8.7%, 其 FPS 小幅下降了 30。综合以上实验结果, 本文设计的网络

模型可以有效提升复杂背景下的红外弱小目标检测效果, 且有着良好的实时性。

3.4.2 检测网络的对比实验

为进一步验证本文提出算法的检测性能, 与经典的 Faster R-CNN^[15]、SSD^[16]、YOLO-V3 和 YOLO-Tiny 进行了对比, 测试结果如表 6 所示。

表 6 算法对比实验结果

Table 6 Experimental results of methods comparison

Method	P/%	R/%	AP/%	FPS
Faster R-CNN	50.6	63.9	58.3	12
SSD	84.6	34.1	64.6	69
YOLOv3	86.1	85.3	90.9	76
YOLOv3-Tiny	80.3	83.6	85.3	256
YOLO-FCSP	91.9	94.6	92.6	170

从表 6 可以看出, 与经典的深度学习算法 YOLOv3、Faster R-CNN 和 SSD 相比, YOLO-FCSP 检测准确率、召回率和平均准确率最高, 分别达到 91.9%、94.6% 和 92.6%, 检测速度大幅提高至 170, 红外弱小目标检测效果提升显著。在复杂背景条件下, YOLO-FCSP 各项指标均能够取得较好的结果, 这表明相对较浅的网络更适合特征不明显的红外弱小目标。与轻量化网络 YOLO-Tiny 相比, 检测速度有小幅下降, 但检测性能提升较为明显。尽管 YOLO-FCSP 检测结果比较突出, 但检测精度也并未达到 100%, 这主要是由于图像中存在高亮边缘、噪声和大幅明暗变化的背景, 与小目标极其类似而产生少量虚警。

为了能够更加直观地分析模型检测性能, 对复杂背景下 5 种不同序列进行红外弱小目标检测, 结果如图 9 所示。结果表明, 在背景亮度剧烈变化的复杂背景下, YOLO-FCSP 可在一定程度

上区分较强迷惑性的斑点状背景和目标,对红外弱小目标具有较好的检测性能,且能够做到较低的虚警率。

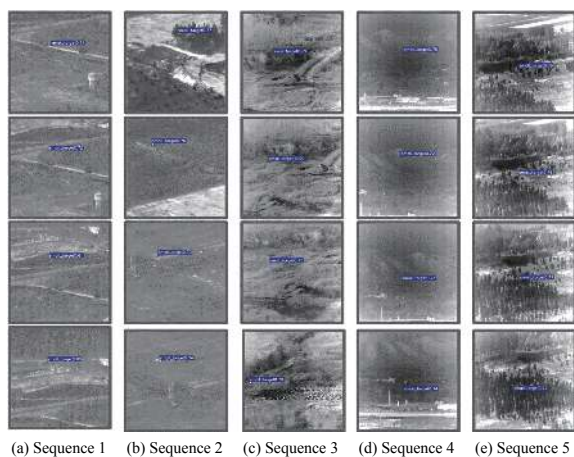


图9 5种不同序列中YOLO-FCSP算法的检测结果

Fig.9 Detection results of YOLO-FCSP in 5 different sequences

4 结论

为解决复杂条件下红外弱小目标的检测问题,本文设计了一种基于深度学习的目标检测算法,利用结合跨阶段局部模块的网络高效提取弱小目标特征,通过自下而上和自上而下的融合路径将浅层细节和位置特征传递到高层。以红外序列图像中的弱小飞机目标为研究对象进行实验,结果表明,本文提出的算法能够有效解决检测过程中弱小目标细节丢失和虚警率高的问题,提高了复杂背景下检测红外弱小目标的准确性和实时性。此外,网络模型复杂度大幅降低,能够又快又准地检测出红外弱小目标,满足了部署到移动硬件平台和嵌入式设备的条件。在后续工作中,将考虑加入注意机制来增强检测算法的特征学习能力,继续优化网络结构,使用生成对抗网络对红外弱小目标的数据集进行数据扩充。

参考文献:

- [1] WU Shuangchen, ZUO Zhengrong. Small target detection in infrared images using deep convolutional neural networks[J]. *Journal of Infrared and Millimeter Waves*, 2019, 38(3): 371.
吴双忱, 左峥嵘. 基于深度卷积神经网络的红外小目标检测[J]. *红外与毫米波学报*, 2019, 38(3): 371.
- [2] REN Xiangyang, WANG Jie, MA Tianlei, et al. Review on infrared dim and small target detection technology[J]. *Journal of Zhengzhou University(Natural Science Edition)*, 2020, 52(2): 1-21.
任向阳, 王杰, 马天磊, 等. 红外弱小目标检测技术综述[J]. *郑州大学学报(理学版)*, 2020, 52(2): 1-21.
- [3] FAN Pengcheng, ZHANG Weiguo, LIU Wangang, et al. Infrared weak small target detection algorithm based on embedded GPU[J]. *Journal of Applied Optics*, 2020, 41(5): 1089-1095.
范鹏程, 张卫国, 刘万刚, 等. 基于嵌入式GPU的红外弱小目标检测算法[J]. *应用光学*, 2020, 41(5): 1089-1095.
- [4] LIU D, Cao L, LI Z, et al. Infrared small target detection based on flux density and direction diversity in gradient vector field[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(7): 2528-2564.
- [5] DENG H, SUN X, LIU M, et al. Infrared small-target detection using multiscale gray difference weighted image entropy[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2016, 52(1): 60-72.
- [6] HAN J, LIANG K, ZHOU B, et al. Infrared small target detection utilizing the multiscale relative local contrast measure[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(4): 612-616.
- [7] PEI Jiaxin, SUN Shaoyuan, WANG Yulan, et al. Night-time environment perception of driverless vehicles based on improved YOLOv3 network[J]. *Journal of Applied Optics*, 2019, 40(3): 380-386.
裴嘉欣, 孙韶媛, 王宇岚, 等. 基于改进YOLOv3网络的无人车夜间环境感知[J]. *应用光学*, 2019, 40(3): 380-386.
- [8] LI Xiaoning, LEI Tao, ZHONG Jiandan, et al. Detecting method of small vehicle targets based on improved SSD[J]. *Journal of Applied Optics*, 2020, 41(1): 150-155.
李小宁, 雷涛, 钟剑丹, 等. 基于改进SSD的车辆小目标检测方法[J]. *应用光学*, 2020, 41(1): 150-155.
- [9] ZHAO M, hCHENG L, YANG X, et al. TBC-Net: a real-time detector for infrared small target detection using semantic constraint[EB/OL]. [2019-12-27]. <https://arxiv.org/abs/2001.05852v1>.
- [10] GONG Y, YU X, DING Y, et al. Effective fusion factor in FPN for tiny object detection[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Com-

- puter Vision. USA: IEEE, 2020: 1160-1168.
- [11] REDMON J, FARHAD A. YOLOv3: an incremental improvement[EB/OL]. [2018-04-08]. <https://arxiv.org/abs/1804.02767>
- [12] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. USA: IEEE, 2020: 390-391.
- [13] HUI Bingwei, SONG Zhiyong, FAN Hongqi, et al. A dataset for dim-small target detection and tracking of aircraft in infrared image sequences[DB/OL]. Science Data Bank, 2019 [2019-10-28]. DOI: [10.11922/csdata.2019.0074.zh](https://doi.org/10.11922/csdata.2019.0074.zh).
- 回丙伟, 宋志勇, 范红旗, 等. 红外序列图像中弱小飞机目标检测跟踪数据集[DB/OL]. Science Data Bank, 2019 [2019-10-28]. <http://10.11922/csdata.2019.0074.zh>.
- [14] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. USA: IEEE, 2019: 6023-6032.
- [15] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: towards real-time object detection with region proposal networks[J]. arXiv preprint arXiv, 2015: 1506.01497.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Switzerland: Springer, Cham, 2016: 21-37.