

基于多特征交叉融合及跨层级联的 航拍目标检测算法

高武奇¹, 杨婷², 李亮亮³

(1.西安工业大学 计算机科学与工程学院, 陕西 西安 710021;
2.西安工业大学 兵器科学与技术学院, 陕西 西安 710021;
3.西安工业大学 机电工程学院, 陕西 西安 710021)

摘要:复杂条件下特殊目标的精确检测是增强特定场景态势生成和预测能力的关键因素。目前的技术不能克服航拍视频中出现的烟雾和遮挡干扰、目标高度变化、尺度不一等问题。因此,提出一个多特征交叉融合及跨层级联的航拍特殊目标检测算法(YOLOv5-MFLC)。针对实际特殊目标保密性高、航拍图像资源匮乏的问题,构建了一个基于真实场景的航拍特殊目标数据集,并采用随机拼接和随机提取嵌入的方法进行数据增强以提高目标多样性和泛化性;针对复杂背景干扰问题,构建了多特征交叉融合注意力机制,增强了目标特征的可用信息;针对航拍图像中目标多尺度问题,设计了跨层级联多尺度特征融合金字塔,提高了跨尺度目标的检测准确率。实验结果表明,与现有的先进检测模型相比,所提算法的检测准确率有较大提升,算法平均准确率可达到81.0%,相比于原始网络提升了5.2%,特别是,在更小的目标类别“person”中达到了55.9%,提升了9.4%,进一步表明了所提改进算法对小目标检测的有用性。同时,所提算法的检测速率可以达56 frame/s,能够有效地实现实际复杂场景特殊目标的准确、快速检测,对特殊目标的识别具有一定的指导意义。

关键词:航拍图像;目标检测;YOLOv5;融合注意力机制;多尺度特征金字塔

中图分类号:TP391

文献标志码:A

文章编号:1000-2758(2023)06-1179-11

近年来,由于无人机规模的急剧增加和应用,通过无人机采集的航拍画面规模和品质也得到了飞跃级的提升^[1]。对特殊目标识别和定位是复杂场景态势感知中至关重要的因素^[2]。所以在大量无人机航拍图像上进行面向复杂环境特殊目标的检测技术研究对于态势的产生与分析有着重要的意义。

传统目标检测算法一般都是先手工提取目标特征,然后将其归类,它们依赖于领域专家对不同对象的纹理、形状、颜色等特征分析,设计出多类型特征描述子,从而检测出带有全局特征的对象^[3-4]。许多学者利用Hough变换进行角点提取,先提取出合作目标边缘,再求出边缘相交点就是期望合作目标角点^[5]。这种方法能够达到亚像素级的精度,但这

种方法需要很好的边缘提取效果且鲁棒性较差。曾接贤等^[6]在K均值算法的基础上改进了一个新的多尺度分形特征矢量,能够实现低对比度下红外坦克图像的准确分割,但是对具体场景目标的检测结果不够精确。郝帅等^[7]提出了基于SIFT分区双向匹配的角点检测算法,实现了图像发生大尺度和角度畸变时,协作目标角点的提取。王永学等^[8]将人工神经网络与蚁群优化相结合,以提供一种新的坦克目标检测算法,虽然这一结合方式具备了速度快、训练耗时较短的优势,但是在与目标相距较远时,检测准确度却较差。传统依靠机器学习的方式仍是利用机械特性来推测对象的重要部分,而不能摆脱对目标整体识别这一难题。因此在物体受到遮挡、出

收稿日期:2022-10-20

基金项目:国家自然科学基金(62171360)与陕西省研究生教育教学综合改革项目

(YJSYG2020073)资助

作者简介:高武奇(1975—),西安工业大学副教授,主要从事图像处理及机器学习研究。e-mail:gaowuqi@xatu.edu.cn

现形变、物体部位没有充分暴露特征等情况下,仍然不能很好地解决要害部位的检测问题。

伴随着近年来计算机技术以及人工智能的发展,深度卷积神经网络在各种目标探测领域^[9-10]中均得到了广泛应用,深度学习具有更强的非线性拟合能力,利用其替代传统手工提取方法对复杂目标深层特征进行分析,提高了算法精度和鲁棒性,更加适合实际复杂场景中航拍目标探测与识别。文献[11]以特征金字塔网络为切入点,通过向 FPN 中加入融合因子刻画相邻层耦合度以控制深层对浅层信息的传递,从而使 FPN 更适应小目标并改善其检测性能。文献[12]通过引入多尺度卷积模块对特征权重进行自适应优化,针对小目标的特点建立多尺度特征融合预测网络,选择多层次特征映射将其融合为高分辨率特征图以增强无人机影像目标检测精度。2020 年张瑞倩等^[13]利用多尺度空洞卷积对特征进行感受野扩大以改善复杂背景及存在遮挡时目标检测结果。在多任务旋转区域卷积神经网络检测模型中, Yang 等^[14]提出了构造稠密特征金字塔的方法来增强稠密航拍图像检测的准确性。王靖宇等^[15]为了解决低空飞行过程中其他视觉物体的干扰问题,提出了一种多隐含层无人机深度神经网络探测模型,通过深入刻画和提取无人机目标多尺度视觉特性,有效地增强探测模型泛化能力以实现长距离弱小无人机的精确探测。军事目标检测技术属于军事领域的基础技术,对于目标跟踪、导弹导航、精准打击都具有重要意义。2016 年, Neagoe 等^[16]提出了一种基于神经网络的合成孔径雷达 (synthetic aperture radar, SAR) 航空图像目标自动识别方法,总成功率为 97.36%。2018 年王全东等^[17]针对 Faster R-CNN 在小尺度坦克装甲车辆目标检

测方面存在的问题,提出了一种改进算法, mAP 可达 93.3%,但是检测速度较慢。2020 年,陈科名^[18]基于 SSD 算法,针对小目标问题,采用 ResNet18 代替 VGG16,同时改进边框尺寸,提升了检测速度。2021 年,舒朗等^[19]基于 YOLOv5 算法,针对尺度变化大、场景多变、可用特征不足的红外图像,提出了一种 Dende-Yolov5 的网络结构,提升了精度和召回率。从整个研究的历程可以看出,目标检测越来越依赖于深度学习算法的提升。

综上所述,现有常用的目标检测算法均不能很好地适应复杂航拍场景中的目标识别问题,当航拍图像出现复杂背景干扰、目标角度多变、尺度不一问题时算法均存在漏检和误检情况,并且检测速度较慢,无法达到较高的实时性要求,未能实现复杂场景目标的快速准确识别任务。本文在 YOLOv5 算法的基础上,提出了一种基于多特征交叉融合及跨层级联的航拍目标检测算法 YOLOv5-MFCL (multi-feature cross fusion and cross-layer concatenation, YOLOv5-MFCL),通过改进网络结构提高复杂背景下真实场景的目标检测性能,同时保证检测算法的实时性以便于实现移动端的部署。

1 YOLOv5-MFCL 算法分析与设计

1.1 YOLOv5 基本原理

YOLOv5 整体结构如图 1 所示。其中,输入端实现数据训练前预处理。主干网络 Backbone 主要用来实现目标特征的提取。颈部网络 Neck 主要实现目标特征的收集。输出检测层 Head 主要用于预测信息损失部分,从而提高对目标识别的准确度。

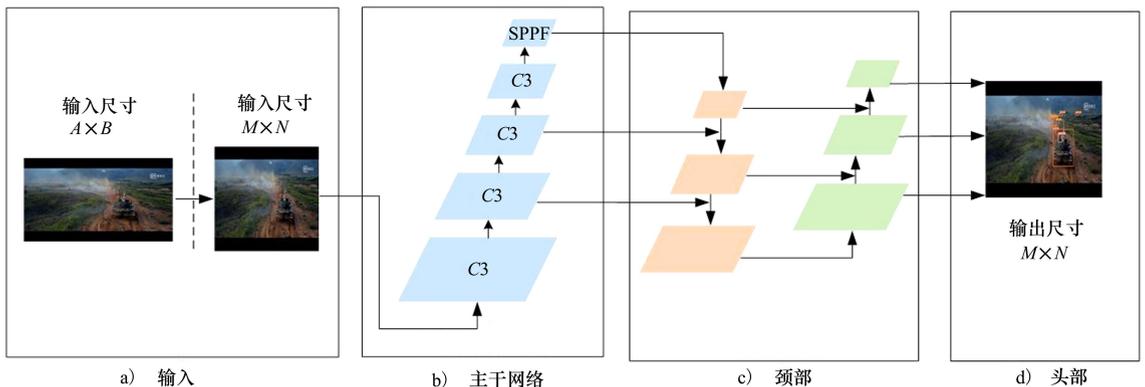


图 1 YOLOv5 算法检测结构

针对不同问题采用 2 种增强方法进行数据扩增。

2.1 基础数据集构建

本文采用无人机自主拍摄,爬虫网络搜索和影视资源截取方式获取基础数据,整理得到 5 500 张图片,采用 LabelImg 软件对其进行标注,最终得到本文初始数据集,如图 3 所示。



图 3 基础数据样本

通过分析基础数据集存在的问题,本文决定通过基于随机拼接的数据增强算法和基于目标提取嵌入的数据增强算法对原始数据进行增强以增强数据集中小目标占比和实现数据类别均衡,同时提高数据多样性和模型泛化能力。

2.2 基于随机拼接的数据增强算法设计

由于无人机航拍和目标的特殊性,无人机拍摄时摄像头探测范围较广,得到的图像视场很大,导致航拍角度的图像中包含的小目标数量较大,现有基础数据集包含较多的中目标和大目标,为了更进一步贴近航拍目标包含小目标多的特点,本文采用一种随机图像拼接方式增强图像中的小目标占比。具体方法流程如图 4 所示。

如图 4 所示,从基础数据集中随机选取 4 张样本作为待拼接图片,同时获取每张样本的标注信息,并进行一系列旋转、加噪、加雾、加雨等操作变换;最后,新建一个空的大图,将各个小图填充进去,同时更新拼接后图像每个目标的标注信息。

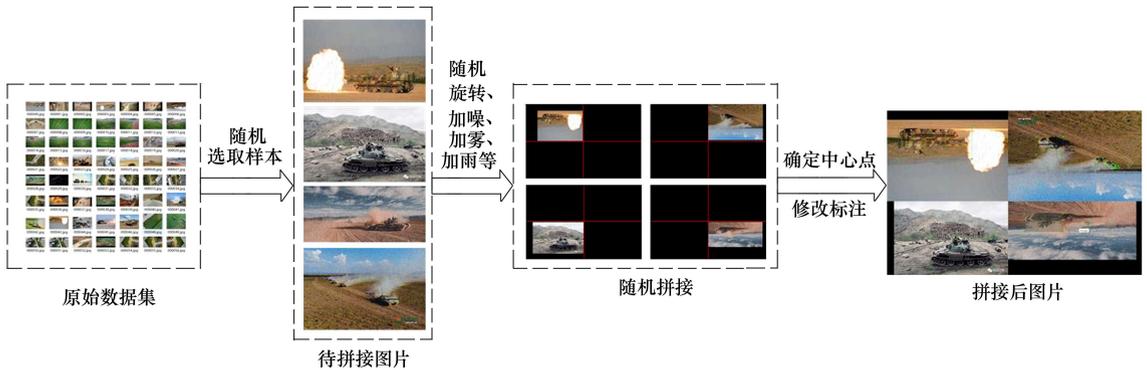


图 4 随机拼接算法示意图

2.3 基于随机目标提取嵌入的目标增强算法设计

对于基础数据集 person 类目标占比较大,tank 和 military-car 占比较小,数据集分布不平衡的问题,本文构建了一种基于随机目标提取嵌入的小目标增强方法来模拟航拍目标。其增强流程如图 5 所示。



图 5 目标提取嵌入示意图

首先运用 Sobel 算子将图片中的目标从背景中分离出来,并进行一系列操作,然后随机选择不同的背景图像,两者结合得到新的模拟航拍图像,实现对罕见数据的增强。

3 YOLOv5-MFCL 航拍目标检测模型构建

3.1 多特征交叉融合注意力机制构建

航拍图像中的小目标所占像素较小,同时也容易受到高空摄影的复杂背景条件限制,原 YOLOv5 模型对小目标的检测作用较差,本文通过分析 GAM 和 CA 存的缺陷与不足,采用残差模块的思想构建

了一个多特征交叉融合注意力机制,以进一步提高复杂背景下算法的检测性能。其结构图如图 6 所示。

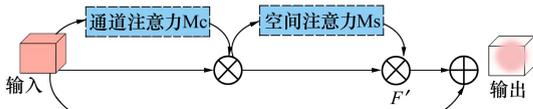


图 6 多特征交叉融合注意力机制结构

输入特征首先进入 GAM 模块中的通道注意力模块来保留三维信息并放大跨维通道-空间依赖关系,再经过空间注意力模块进行空间信息的聚焦。对于全局注意力和坐标注意力的连接方式,本文采用引入残差结构的方式进行组合。此操作不仅能综合 GAM 对空间通道信息的跨维交互,而且能够结合 CA 对位置信息的精确定位,进一步增强有用的特征信息,聚焦并选择复杂背景中对任务有效的信息,强化特征输入到后续网络中,从而提高整体网络的性能,提升网络模型的运行效率。

3.2 跨层级联的多尺度特征融合金字塔设计

无人机的飞行高度不定,空中拍摄角度多变,拍摄范围广的特点导致不同目标之间尺度差异较大。YOLOv5 采用 PANet 结构来实现多尺度特征融合模块,通过简单的双向特征提取网络直接融合不同密度的信息,忽略了不同尺度特征之间冲突信息的存在,限制多尺度特征的表达。本文在 PANet 的基础上融合 BIFPN 的思想,并建立了一个全新的跨层级联的多尺度特征融合金字塔模型,具体结构如图 7 所示,主要包括下面几个部分。

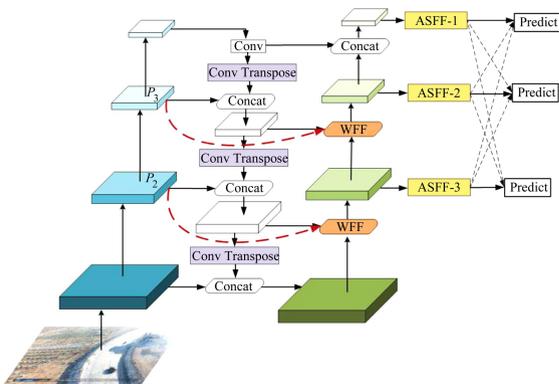


图 7 本文设计金字塔结构图

3.2.1 跨层级联特征增强路径

在 3 层 FPN 结构的基础上基于 BIFPN 的网络

结构设计了一种跨层级联的特征融合金字塔,将其深度变为 4 层,向上加深了金字塔的深度,如图 8 所示。

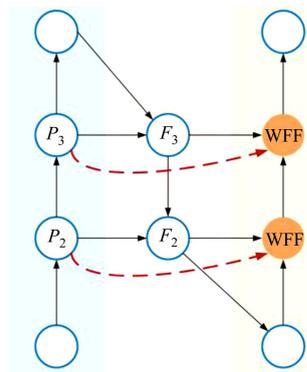


图 8 跨层级联特征增强路径

图中红色虚线所示即为跨层级联路径。在自底向上的通路中进一步结合基础特征图中的丰富信息,可以使深层网络的语义信息和浅层网络的细节信息进一步融合。另外采用加权特征融合的方式 WFF (wighted feature fusion, WFF) 来平衡不同特征的权重,判断不同输入的重要性,不仅可以保留更多的特征信息,还可以简化融合计算,提高计算速度。

3.2.2 基于反卷积的上采样方式

原网络最邻近插值法会造成很严重的特征损失,降低小目标的检测精度。针对此问题,本文采用反卷积的方式进行上采样,以减少采样过程带来的信息丢失。其结构图如图 9 所示。

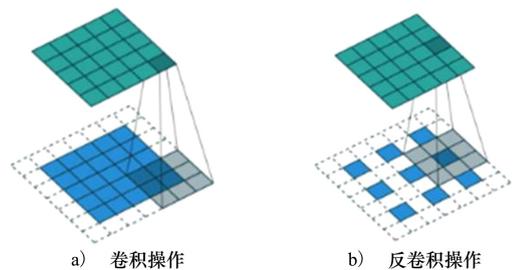


图 9 卷积与反卷积示意图

如图 9 所示,3×3 特征图作为转置卷积的基础输入,对 3×3 特征图上每间隔一行与一列进行插 0 操作,得到的填充后的特征图作为新的输入,最后用 3×3 的卷积核进行步长为 1 的反卷积操作,得到最终的 5×5 特征图。

3.2.3 自适应空间特征融合 ASFF 模块设计

在目前的目标检测任务预测分支当中, 低层特征适合检测图片中的小物体, 高层特征适合检测图片中的大物体, 中等大小的目标则适用中间的特征层。一般 FPN 均采用直接连接或者按元素相加这种特征拼接的方式, 但是这种连接方式不能充分学习到不同尺度之间的特征信息, 从而不能有效地进行多尺度特征融合。因此, 本文提出利用一种自适应空间特征融合 ASFF (adaptively spatial feature fusion, ASFF) 的方式来进行跨尺度特征融合。其网络结构如图 10 所示。

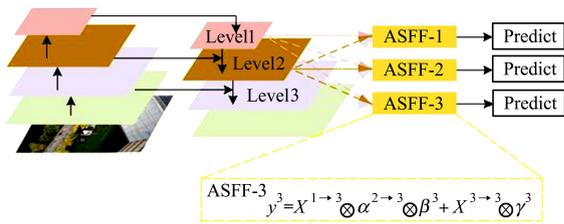


图 10 ASFF 结构示意图

ASFF 通过利用权重参数调整特征融合时不同尺寸特征的贡献大小, 避免了小尺度语义信息和大尺度细节信息的混乱和丢失, 进一步增强了多尺度之间的信息交互融合, 提高了检测精度。

4 实验分析

4.1 实验环境及参数配置

本文实验的操作系统为 Ubuntu16.04, 深度学习框架为 PyTorch。所有对比实验采用相同的数据集和硬件平台, 除了引入不同模块之外, 其他所有训练参数均保持相同, 模型训练参数设置如表 1 所示。

表 1 模型训练参数设置

参数名	参数值
动量	0.937
权重衰减	0.000 5
批量大小	8
学习率	0.01
迭代轮数	150

4.2 数据增强效果分析

4.2.1 随机拼接算法增强结果

经过统计, 采用随机拼接算法得到的图片数量为 5 820 张, 通过统计图片中不同大小目标的具体

个数, 图 11 给出了随机拼接算法增强前后数据集中目标大小分布图, 横坐标和纵坐标分别代表物体的宽和高, 由对比图可观察到增强之前总目标数量较少, 且小目标占比较小, 中目标占比较多; 增强之后色块向坐标原点位置移动, 说明小目标占比增多, 且总体数据量也得到提升。

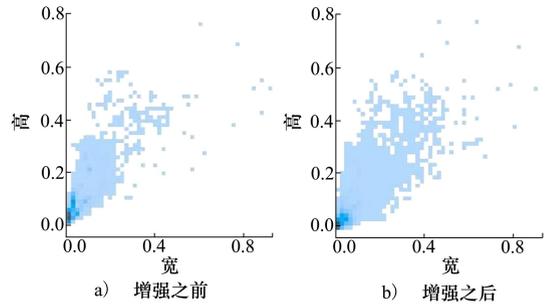


图 11 数据增强前后目标大小分布对比

4.2.2 随机提取嵌入算法增强结果

经过统计, 采用随机提取嵌入算法之后得到的图片数量为 4 963 张, 图 12 给出了随机提取嵌入方法增强前后 military-car 和 tank 这两类目标的数量分布情况。如图所示, 蓝色表示增强之前基础数据, 红色表示增强之后数据, 柱形图代表 military-car, 折线图代表 tank, 横坐标为小目标、中目标和大目标, 纵坐标为目标数量, 可以看出采用本文方法进行增强之后 2 类目标的数量明显有了大幅度提升, 并且其中的小目标占比也得到较大提升。

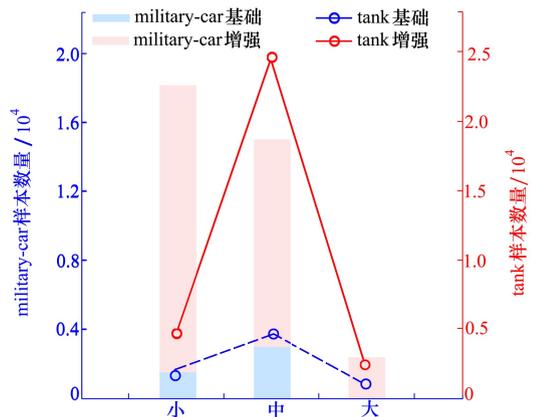


图 12 随机提取嵌入数据增强前后目标数量分析

综上所述, 本文的数据增强算法解决了基础数据集小目标占比小和数据种类分布不平衡的问题, 同时针对本文研究背景实现了各种真实场景下的图像构建, 共计得到了 15 783 张图片, 并对其标注文

件进行整理,训练集和测试集的划分比例为 7 : 3,为后续航拍目标检测提供了丰富可靠的数据支持。

4.3 消融实验

4.3.1 多特征交叉融合注意力机制实验结果对比

为了验证本文构建的融合注意力机制的有效

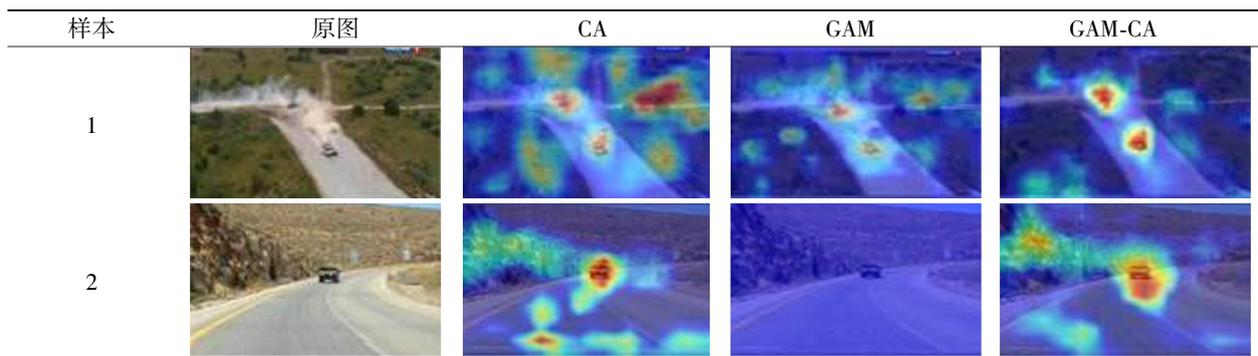
表 2 改进注意力实验方法及结果对比

模型	注意力	召回率/%	阈值为 0.5 的平均准确率/%	阈值为 0.5 : 0.95 的平均准确率/%
YOLOv5s		67.9	75.8	45.2
YOLOv5s-CA	CA	68.2	75.8	45.3
YOLOv5s-GAM	GAM	68.1	76.6	45.4
YOLOv5s-GC1	GAM-CA	68.3	76.4	45.9

由表 2 可知,加入注意力机制普遍能够提升网络的检测精度,本节实验通过对比单独 GAM、CA 注意力和多特征交叉融合注意力机制 GAM-CA 发现,引入 GAM-CA 之后阈值为 0.5 的平均准确率提升了 0.6%,阈值为 0.5 : 0.95) 的平均准确率提升了 0.7%。因此本文最终决定引入 GAM-CA 来进行主干网络的特征融合,从而进一步提高航拍复杂场景下算法的检测性能。

对引入 GAM-CA 前后加权重叠图进行了对比,

表 3 不同注意力的热力图对比



4.3.2 跨层级联多尺度特征融合金字塔

为了验证本文所设计金字塔的有效性,本节选择以上述引入本文设计注意力机制的模型 YOLOv5s-GC1 作为基线,接着在网络中逐步引入多

性,本节对分别引入本文设计注意力机制和单一注意力之后的检测结果进行对比,通过分析实验结果说明本文构建注意力机制的有效性,其中,具体对比方法和实验结果如表 2 所示。

结果如表 3 所示。从表中可以看出,和其他 2 种主流注意力机制相比,加入本文设计注意力机制后,网络对检测目标区域的覆盖度和关注程度都获得了提升,证明本文设计的注意力机制能够帮助深度卷积网络提取到更加关键的特征信息,在复杂的航拍图像中能够更快地捕获其中的关键信息,从而使检测器对目标进行“有区别”检测,提升网络模型运行效率。

尺度金字塔的不同模块,通过对比引入前后的精度变化来说明本文改进机制的有效性。各模块消融实验对比如表 4 所示。

表 4 改进金字塔实验结果对比

YOLOv5s-GC1	BIFPN	跨层级联	Head-GAM-CA	反卷积	ASFF	召回率/%	阈值为 0.5 的平均准确率/%	阈值为 0.5 : 0.95 的平均准确率/%
✓						68.3	76.4	45.9
✓	✓					69.3	77.2	47.0
✓		✓				72.0	79.7	49.1
✓		✓	✓			73.0	80.2	50.0
✓		✓	✓	✓		73.3	80.6	50.9
✓		✓	✓	✓	✓	73.3	81.0	51.3

由表 4 可知,引入金字塔的不同模块对模型精度均有一定增长作用,最终的模型阈值为 0.5 的平均准确率达到 81.0%,阈值为 0.5 : 0.95 的平均准确率达到 51.3%,相比于基线网络分别提升了 4.6% 和 5.5%,召回率达到了 73.3%,提升了 5.0%。证明了本文跨层级联多尺度金字塔结构能进一步提高网络对边界的回归精度。

4.4 目标检测性能分析

为了更加直观地说明本文算法改进前后的对比情况,图 13 给出了对比曲线。通过图 13a) 精度对比可知改进算法优于原算法,通过图 13b) 损失函数对比可知改进算法损失低于原算法,具有更好的性能。

4.5 先进性对比

为了进一步说明本文改进算法的有效性,表 5 给出了本文提出的算法与目前主流的其他目标检测

算法的结果对比。

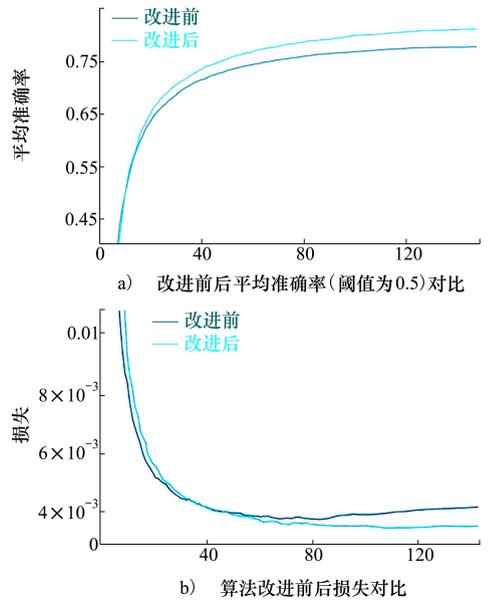


图 13 算法改进前后对比

表 5 不同算法检测精度对比

算法	阈值为 0.5 的平均准确率/%	准确率/%				模型体积/MB	帧率
		military-car	person	tank	car		
YOLOv5s-本文	81.0	87.4	55.9	97.2	83.4	47.1	56
YOLOv5s	75.8	82.4	46.5	95.5	78.6	14.8	140
YOLOv5m	79.6	86.2	52.8	97.0	82.4	42.2	74
YOLOv5l	81.0	88.0	55.2	97.4	83.6	92.8	45
YOLOv4	55.2	53.0	14.0	79.0	35.0	256.2	44
YOLOv4-tiny	34.3	41.0	16.0	53.0	28.0	23.6	46
YOLOv6	80.1					38.0	300

4.6 检测效果分析

为展示本文改进算法在实际场景中检测的有效性,选取了部分真实场景图片进行测试,通过可视化分析对比不同场景下的检测效果。

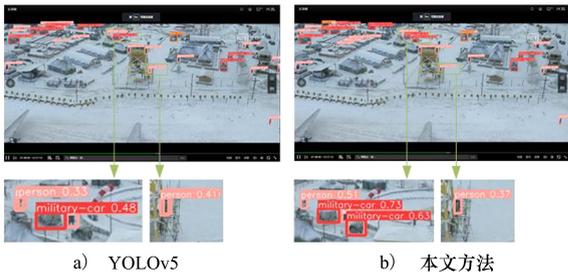


图 14 雪天环境结果对比

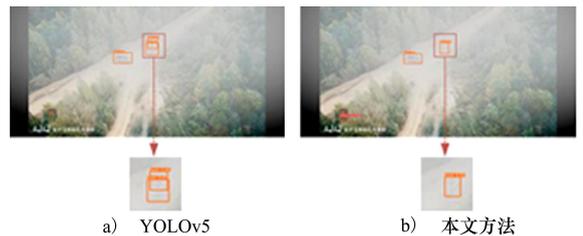


图 15 雾天环境结果对比

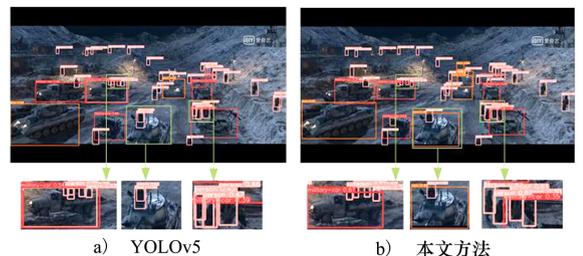


图 16 夜晚环境结果对比

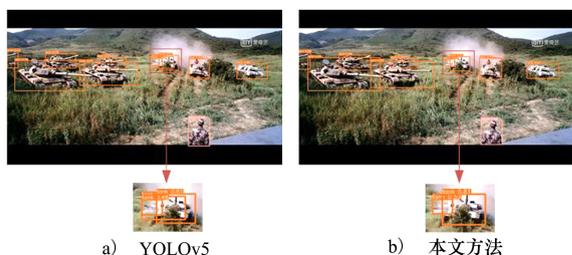


图 17 遮挡条件结果对比

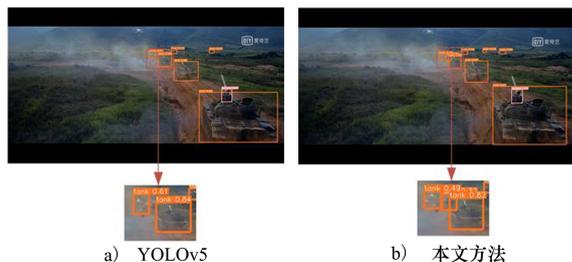


图 18 多尺度条件下结果对比

综上所述,不论是从客观还是主观方面分析,本文提出的改进算法均具有较好的表现,检测精度高于原始算法,并且成功地降低了航拍图像在复杂背

景、目标密集和多尺度情况下网络的误检和漏检率,具有较好的性能。

5 结 论

本文通过随机拼接、随机提取嵌入等多个数据增强算法构建并扩充了基于真实场景的航拍特殊目标数据集,提出了一种基于多特征交叉融合和跨层级联的多尺度航拍目标检测算法(YOLOv5-MFCL),解决了真实场景下由于复杂背景干扰、目标角度多变、尺度不一带来的检测正确率低、误检率高等问题,本文提出的改进检测算法达到了 81.0% 的准确率,相比于原网络提高了 5.2%,检测速度可达到 56 frame/s,可以快速、准确地检测到多种实际场景中的航拍特殊目标,为实际态势把握提供准确、高效的理论支持。在未来的研究中,为了实现实际特殊场景移动端目标检测的部署,本文将考虑对网络进行剪枝、蒸馏等处理,进一步降低模型参数和复杂度,实现模型的轻量化。

参考文献:

- [1] HUI Pengyifan. Battlefield image situational awareness application based on deep learning[J]. IEEE Intelligent Systems, 2020, 35(1): 36-43
- [2] 戴文君,常天庆,褚凯轩,等. 基于时空卷积特征记忆模型的坦克火控系统视频目标检测方法[J]. 兵工学报, 2020, 41(9): 1708-1718
DAI Wenjun, CHANG Tianqing, CHU Kaixuan, et al. Video target detection method for Tank fire control system based on spatio-temporal convolution feature memory model[J]. Acta Armamentarii, 2020, 41(9): 1708-1718 (in Chinese)
- [3] 张奔,徐锋,李晓婷,等. 针对军事集群目标的 YOLOv3 改进算法研究[J]. 火力与指挥控制, 2021, 46(5): 81-85
ZHANG Ben, XU Feng, LI Xiaoting, et al. Research on improved YOLOv3 algorithm for military cluster target[J]. Fire Control & Command Control, 2021, 46(5): 81-85 (in Chinese)
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision & Pattern Recognition, 2014: 580-587
- [5] 汪烈兵,姜雄飞,石春光,等. 基于图像滤波与 Hough 变换的红外弱小目标检测[J]. 红外技术, 2020, 42(7): 5
WANG Liebing, JIANG Xiongfei, SHI Chunguang, et al. Infrared dim small target detection based on image filtering and Hough transform[J]. Infrared Technology, 2020, 42(7): 5 (in Chinese)
- [6] 曾接贤,王军婷,符祥. K 均值聚类分割的多特征图像检索方法[J]. 计算机工程与应用, 2013, 49(2): 226-230
ZENG Jiexian, WANG Junting, FU Xiang. Multi-feature image retrieval method based on K-means clustering segmentation[J]. Computer Engineering and Applications, 2013, 49(2): 226-230 (in Chinese)
- [7] 郝帅,程咏梅,马旭,等. 无人直升机视觉着舰中合作目标角点检测算法[J]. 西北工业大学学报, 2013, 31(4): 653-659
HAO Shuai, CHENG Yongmei, MA Xu, et al. Corner detection algorithm for cooperative target with visual landing of unmanned helicopter[J]. Journal of Northwestern Polytechnical University, 2013, 31(4): 653-659 (in Chinese)

- [8] 王永学,王双进. 基于蚁群优化与神经网络的坦克目标识别算法[J]. 物理实验, 2010, 30(5): 8-11
WANG Yongxue, WANG Shuangjin. Tank target recognition algorithm based on ant colony optimization and neural network[J]. Physical Experiment, 2010, 30(5): 8-11 (in Chinese)
- [9] 段仲静,李少波,胡建军,等. 深度学习目标检测方法及其主流框架综述[J]. 激光与光电子学进展, 2020, 57(12): 59-74
DUAN Zhongjing, LI Shaobo, HU Jianjun, et al. Review of deep learning based object detection methods and their mainstream frameworks[J]. Laser & Optoelectronics Progress, 2020, 57(12): 59-74 (in Chinese)
- [10] 戴文君,常天庆,张雷,等. 图像目标检测技术在坦克火控系统中的应用[J]. 火力与指挥控制, 2020, 45(7): 147-152
DAI Wenjun, CHANG Tianqing, ZHANG Lei, et al. Application of image target detection technology in tank fire control system [J]. Fire Control & Command Control, 2020, 45(7): 147-152 (in Chinese)
- [11] GONG Y, YU X, DING Y, et al. Effective fusion factor in FPN for tiny object detection[C]//IEEE Winter Conference on Applications of Computer Vision, 2021: 1159-1167
- [12] 刘芳,韩笑. 基于多尺度深度学习的自适应航拍目标检测[J]. 航空学报, 2022, 43(5): 325270
LIU Fang, HAN Xiao. Adaptive aerial object detection based on multi-scale deep learning[J]. Acta Aeronauticae Astronautica Sinica, 2022, 43(5): 325270 (in Chinese)
- [13] 张瑞倩,邵振峰, ALEKSEI Portnov, 等. 多尺度空洞卷积的无人机影像目标检测方法[J]. 武汉大学学报(信息科学版), 2020, 45(6): 108-116
ZHANG Ruiqian, SHAO Zhrngfeng, ALEKSEI Portnov, et al. Multi-scale cavity convolution for UAV image target detection[J]. Geomatics and Information Science of Wuhan University, 2020, 45(6): 108-116 (in Chinese)
- [14] YANG X, SUN H, SUN X, et al. Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network[J]. IEEE Access, 2018, 6: 50839-50849
- [15] 王靖宇,王霰禹,张科,等. 基于深度神经网络的低空弱小无人机目标检测研究[J]. 西北工业大学学报, 2018, 36(2): 258-263
WANG Jingyu, WANG Xianyu, ZHANG Ke, et al. Research on low altitude dim small UAV target detection based on deep neural network[J]. Journal of Northwestern Polytechnical University, 2018, 36(2): 258-263 (in Chinese)
- [16] NEAGOE V E, CARATA S V, CIOTEC A D. An advanced neural network-based approach for military ground vehicle recognition in SAR aerial imagery[J]. International Scientific Committee, 2016, 18: 41-48
- [17] 王全东,常天庆,张雷,等. 面向多尺度坦克装甲车辆目标检测的改进 Faster R-CNN 算法[J]. 计算机辅助设计与图形学学报, 2018(12): 2278-2291
WANG Quandong, CHANG Tianqing, ZHANG Lei, et al. Improved Faster R-CNN algorithm for multi-scale tank and armored vehicle target detection[J]. Journal of Computer Aided Design and Graphics, 2018, 30(12): 2278-2291 (in Chinese)
- [18] 陈科名. 基于 SSD 的多尺度目标检测算法研究[D]. 北京:北京交通大学, 2020: 61-62
CHEN Keming. Research on multi-scale target detection algorithm based on SSD[D]. Beijing: Beijing Jiaotong University, 2020: 61-62 (in Chinese)
- [19] 舒朗,张智杰,雷波. 一种针对红外目标检测的 Dense-Yolov5 算法研究[J]. 光学与光电技术, 2021, 19(1): 69-75
SHU Lang, ZHANG Zhijie, LEI Bo. Research on Dense-Yolov5 algorithm for infrared target detection[J]. Optics and Optoelectronic Technology, 201,19(1): 69-75 (in Chinese)

Aerial military target detection algorithm based on multi-feature cross fusion and cross-layer concatenation

GAO Wuqi¹, YANG Ting², LI Liangliang³

(1.School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710021, China;
2.School of Weapon Science and Technology, Xi'an University of Technology, Xi'an 710021, China;
3.School of Mechanical and Electrical Engineering, Xi'an University of Technology, Xi'an 710021, China)

Abstract: The precise detection of military targets under complex conditions is a key factor to enhance the ability of war situation generation and prediction. The current technology can not overcome the problems of smoke and occlusion interference, target height change, and different scales in aerial video. In this paper, a multi feature cross fusion and cross layer cascade aerial military target detection algorithm (YOLOv5-MFLC) is proposed. Firstly, aiming at the high confidentiality of the military targets and the shortage of battlefield aerial image resources, a real scene based aerial military target dataset is constructed, and the methods of random splicing and random extraction embedding are used for data enhancement in order to improve the diversity and generalization of targets. Secondly, aiming at the problem of complex background interference, a multi feature cross fusion attention mechanism is constructed to enhance the available information of target features. Finally, for the multi-scale problem of targets in aerial images, a cross layer cascaded multi-scale feature fusion pyramid is designed to improve the detection accuracy of cross scale targets. The experimental results show that, comparing with the existing advanced detection models, the detection accuracy of the algorithm in this paper has been greatly improved. The average accuracy of the algorithm can reach 81.0%, which is 5.2% higher than the original network. In particular, it has reached 55.9% in the smaller target category "person", which is 9.4% higher. And the experimental results further show the usefulness of the improved algorithm for small target detection. At the same time, the detection rate of this algorithm can reach 56 frame/s, which can effectively achieve accurate and fast detection of battlefield targets, and has certain experience value for guiding complex modern wars.

Keywords: aerial image; target detection; YOLOv5; fusion attention mechanism; multiscale characteristic pyramid

引用格式:高武奇, 杨婷, 李亮亮. 基于多特征交叉融合及跨层级联的航拍目标检测算法[J]. 西北工业大学学报, 2023, 41(6): 1179-1189

GAO Wuqi, YANG Ting, LI Liangliang. Aerial military target detection algorithm based on multi-feature cross fusion and cross-layer concatenation[J]. *Journal of Northwestern Polytechnical University*, 2023, 41(6): 1179-1189 (in Chinese)