



计算机科学

COMPUTER SCIENCE

改进YOLOv5的小型旋翼无人机目标检测算法

路琪, 于元强, 许道明, 张琦

引用本文

路琪, 于元强, 许道明, 张琦. 改进YOLOv5的小型旋翼无人机目标检测算法[J]. 计算机科学, 2023, 50(11A): 220900050-8.

LU Qi, YU Yuanqiang, XU Daoming, ZHANG Qi. Improved YOLOv5 Small Drones Target Detection Algorithm [J]. Computer Science, 2023, 50(11A): 220900050-8.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于边缘引导的多尺度医学影像分割方法](#)

Medical Image Segmentation Based on Multi-scale Edge Guidance

计算机科学, 2023, 50(11A): 220900059-7. <https://doi.org/10.11896/jsjcx.220900059>

[基于语义注意力的医学图像超分辨率方法](#)

Medical Image Super-resolution Method Based on Semantic Attention

计算机科学, 2023, 50(11A): 221200107-6. <https://doi.org/10.11896/jsjcx.221200107>

[一种基于因果推理的垃圾分类方法](#)

Novel Method for Trash Classification Based on Causal Inference

计算机科学, 2023, 50(11A): 220800218-6. <https://doi.org/10.11896/jsjcx.220800218>

[接诉即办智能派单业务调度算法研究](#)

Study on Scheduling Algorithm of Intelligent Order Dispatching

计算机科学, 2023, 50(11A): 230300029-7. <https://doi.org/10.11896/jsjcx.230300029>

[一种基于CutMix的增强联邦学习框架](#)

Enhanced Federated Learning Frameworks Based on CutMix

计算机科学, 2023, 50(11A): 220800021-8. <https://doi.org/10.11896/jsjcx.220800021>

改进 YOLOv5 的小型旋翼无人机目标检测算法

路琪¹ 于元强¹ 许道明¹ 张琦²

1 空军预警学院雷达士官学校 武汉 430345

2 中国西安卫星测控中心 西安 710000

(466729206@qq.com)

摘要 低空慢速小型目标检测一直是预警探测领域关注的重点和难点。目前,基于神经网络的主流目标检测算法在设计时主要考虑应用于 VOC 数据集或 COCO 数据集,在特定场景下检测精度不够理想。针对复杂背景下小型旋翼无人机目标检测的特定检测场景,提出一种基于改进 YOLOv5 的小型旋翼无人机目标检测算法。首先,增加小目标检测层以获得大尺寸的浅层特征图,从而提升算法对小目标的检测能力;其次,针对小型旋翼无人机尺寸不一的问题,利用 K-Means++ 聚类算法对先验框的尺寸进行优化并将其与各特征层进行匹配;最后,使用 Mosaic-SOD 方法进行数据增强以及改进损失函数,增强算法对小目标的感知能力以及提高网络训练效率。将改进后的算法应用在复杂背景下的小型旋翼无人机目标检测中,实验结果表明,相较于原始 YOLOv5 算法,该算法在小型旋翼无人机目标检测上具有更高的检测精度和特征提取能力,虽然检测速度有一定下降,但通过对可见光视频流进行检测可知其仍能够满足实时性的要求。

关键词: 低慢小目标;反无人机系统;深度学习;小目标检测;旋翼无人机检测;数据增强;特征融合;YOLOv5

中图分类号 TP391.4

Improved YOLOv5 Small Drones Target Detection Algorithm

LU Qi¹, YU Yuanqiang¹, XU Daoming¹ and ZHANG Qi²

1 School of Radar NCO, Air Force Early Warning Academy, Wuhan 430345, China

2 Xi'an Satellite Control Center, Xi'an 710000, China

Abstract The detection of low-altitude slow-speed small targets has always been the focus and difficulty in the field of early warning detection. At present, the mainstream target detection algorithms based on neural networks are mainly designed to be applied to VOC dataset or COCO dataset, and the detection accuracy is not ideal in specific scenarios. Aiming at the specific detection scene of small drones target detection in complex background, a small drones target detection algorithm based on improved YOLOv5 is proposed. First, a small target detection layer is added to obtain a large-sized shallow feature map, thereby improving the detection ability of the algorithm for small targets. Secondly, for the problem of different sizes of small drones, K-means++ clustering algorithm is used to detect the prior frame. The size of the inspection frame is optimized and matched with each feature layer. Finally, the Mosaic-SOD methods of data augmentation and improved loss function are used to enhance the algorithm's ability to perceive small targets and improve the efficiency of network training. The improved algorithm is applied to the target detection of small drones in complex background. Experimental results show that compared with the original YOLOv5 algorithm, the proposed algorithm has higher detection accuracy and characteristics in target detection of small rotor UAV. The extraction capability, although the detection speed has decreased to a certain extent, can still meet the real-time requirements by detecting the visible light video stream.

Keywords Low-Slow-Small target, Anti-UAV system, Deep learning, Small object detection, Small drones detection, Data augmentation, Feature fusion, YOLOv5 (You Only Look Only version 5)

低空慢速小型 (Low-altitude Slow-speed Small, LSS) 无人机是一种无人驾驶的小型飞行器,主要通过无线电设备或预编程进行操作控制,可作为载具完成特定的任务。随着技术的不断进步, LSS 无人机逐渐向体积小、重量轻的趋势发展,且可操作性、机动性越来越强。LSS 无人机,尤其是旋翼无人机,悬停精度最高可达到 0.1 m,可以在城市、山地、林地等复杂应用环境中随时起降。近年来,民用、军用等小型旋翼

无人机的广泛使用在给人们带来很多便利的同时,也带来了许多问题。例如,低空小型旋翼无人机可作为运载工具完成包裹的投送,但也可通过挂载枪支、炸药或其他危险物品,为公共安全或国防安全带来威胁;可通过挂载拍摄设备拍摄出具有视觉冲击力的影视作品,但同样也可用来拍摄个人活动或军事设施,从而危害个人隐私或造成军事秘密的泄露。近年来,国内外发生的诸多小型旋翼无人机非法入侵事件,对

基金项目:空军预警学院厚基工程(HJGC-2022-001)

This work was supported by the Consolidate Foundation Project of Airforce Early Warning Academy(HJGC-2022-001).

通信作者:路琪(466729206@qq.com)

敏感地区安全构成了极大的威胁,例如军事基地、大型集会地、核电站、政府秘密部门的住所等。在军事领域,通过数十架甚至上百架无人机进行编队飞行,搭配红外相机、合成孔径雷达等探测设备,可以实现全天候的情报侦察、监视和打击功能。从近年来的多次局部战争以及目前仍在持续的俄乌冲突可知,小型无人机的使用,已经改变了现代的作战方式,是一支不可忽视的重要军事力量。因此,对小型旋翼无人机目标进行快速、准确地识别和检测有着极其重要的意义。通过对目标的识别和检测,可以完成目标的预警,及时采取有效的措施减少甚至消除威胁。

目前,世界各国对小型无人机防控的关注程度不断提高,对反无人机的研究也不算深入。早在2012年,美军就已经开展反无人机作战研究。Grant等^[1]提到,随着无人机在世界各国的扩散,美军在未来的战场上可能遭遇无人机的“蜂群”攻击。为此,美国空军进行了一系列演习和测试,以研究应对无人机的战术。2017年4月13日,美国陆军发布了《反无人机系统技术手册》,该手册旨在帮助美陆军更好地应对未来可能面临的无人机威胁,以及指导机动部队进行反无人机的战术训练和处置流程的规范化训练^[2-3]。2020年美军第11野炮团2营主任参谋Guelfi等^[4]认为“现有的防空系统无法探测或攻击慢速、低空飞行的无人机系统。如果不能在整个冲突范围内降低这种作战风险,美国陆军将很容易受到国家和非国家对手使用无人机的攻击。”2021年1月,为应对无人机小型化、集群化作战带来的挑战,美国国防部发布《反小型无人机系统战略》,由此可见美军已将反无人机问题提升到战略层面进行分析解决,而不仅仅停留在技术或战术层面。同年,A. B. Manbkob等提到近年来武装冲突的特点是广泛使用无人机。现在俄罗斯联邦军兵种领导人经常分析研究使用无人机的方法和对抗手段,以探索有效遂行作战方法并夺取胜利。总体来说,许多国家已经意识到无人机小型化、集群化、智能化带来的军事威胁,可见开展反无人机系统研究,尤其是反低空慢速小型无人机系统的研究的重要性和紧迫性。

目前采取的反无人机技术主要包括无人机侦察探测和无人机反制两类。无人机反制技术主要有软杀伤和硬杀伤两类,软杀伤主要通过无线电的手段破坏无人机的通信、定位等功能,从而使无人机无法完成其任务。硬杀伤是通过防空武器对无人机目标进行打击^[5-8]。而无人机侦察探测技术是通过测向和测距等方式对威胁目标进行识别和定位,是进行无人机反制的前提,根据探测所使用的介质不同,具体探测手段如下。

采用雷达系统完成目标的探测具有探测距离远、精度高、几乎不受天气影响、技术成熟度高的优点^[9-10],但是识别性能较差,主动辐射信号会影响探测系统的安全性;采用射频探测系统具有成本低、容易实现的优点,但是一旦无人机系统处于无线电静默状态就难以完成探测;采用热成像探测系统具有杂波影响小、夜间效果好的优点,但是大部分微小型无人机的热学特性不是十分明显,探测效果不是很好;采用声学探测系统具有成本低、安全性好的优点,但是探测精度难以保证且识别能力较差,容易受到风声的影响而造成虚警;而无源光学成像探测系统可利用紫外成像、可见光成像、近红外光谱成像等技术获取图像,并对图像进行分析,具有成本低、视场灵活性高的优点。

目前美军正致力于将人工智能技术应用于反无人机系统,如雷锡恩公司的Windshear反无人机系统,该系统利用人工智能手段可以实现对无人机目标的敌我识别,并且能为操作员提供操作选择。考虑到无源光学成像易受到杂波和天气的影响,故本文尝试利用人工智能技术解决无源光学成像探测系统在复杂背景条件下的小型旋翼无人机目标检测问题。

1 小目标检测技术

作为计算机视觉的基础技术,目标检测力求解决视觉应用中两个最基础的问题:物体是什么,或物体在何处?20世纪末人工智能技术进入第三次大发展,计算机性能的大幅度提升和大数据时代的到来,也给深度学习技术的发展提供了前提条件^[11-13]。将深度学习技术应用于目标检测领域后,目标检测技术取得显著性的突破,目前已被广泛应用于情报侦察、探测预警、视频监控等军用和民用领域,在无人机目标检测任务中也起到了不可或缺的作用。目前主流的目标检测算法大致分为两类:一类为基于区域推荐的目标检测方法,也称为两阶段检测算法,其基本思路是利用区域提出网络,替代选择性搜索算法进行候选区域提取,典型代表有R-CNN, Fast R-CNN, Faster R-CNN等系列算法;另一类称为单阶段检测算法,其基本思路是利用回归的方法来解决目标检测问题,与两阶段检测算法的不同之处在于,其无需区域推荐的阶段,直接进行区域的检测和识别,只需一个卷积神经网络就可直接得到不同目标的类别概率和位置坐标,直接回归目标信息,检测速度快,适合实时性要求高的任务。代表性的算法有YOLO(You Only Look Once), SSD, Retina-Net等^[14-15]。

虽然随着人工智能技术的飞速发展,目标检测技术有了重大突破,但由于可用特征少、目标占比少、样本不均衡等因素,如何快速、精准地检测小目标仍然是计算机视觉中的一个难点和研究热点。为了提高对小目标的检测精度与检测速度,小目标检测技术主要从以下几方面展开。

1) 数据增强。数据增强是提高小目标检测性能最简单、最有效的方法。数据增强策略在不增加额外训练数据集样本的情况下,通过组合、拼接等处理方式丰富数据集,提高样本的多样性,扩大训练数据集的规模,适用于小样本条件下的目标检测,从而提升增强检测模型的鲁棒性和泛化能力,但同时会带来计算成本的提高,对用于进行训练的系统能力有着较高的要求。目前,出现了一些适用于小目标检测的数据增强手段,如通过对图像中的小目标的复制与粘贴操作进行数据增强的复制增强、考虑上下文信息的自适应采样、通过尺度匹配策略对图像进行尺度变换用作额外的数据补充的尺度匹配、通过强化学习选择最佳数据增强策略自学习数据增强等。

2) 多尺度学习。多尺度处理需要模型能检测并识别不同尺度的目标,对模型的鲁棒性提出了更高的要求。对大物体而言,由于高层特征图的感受野更大,更能观察到大物题的全貌,其语义信息(图像的纹理、颜色,或者目标的类别等信息)往往出现在高层次的特征图中。而对于小物体而言则恰恰相反,随着处理深度的不断加深,感受野逐渐扩大,小目标在高层次的特征图中的语义信息变得非常模糊,故其细节信息往往出现在低层特征图中,随着网络的加深和下采样的进行,其细节信息可能会完全消失,仅利用深层特征识别目标的类别,更是难上加难。为了更好地提升对小目标的检测效率,

有利于小型目标的特征提取,多尺度学习方法需要尽可能同时考虑多个特征图中的信息,将表征信息与语义信息进行融合。在提升检测性能的同时也带来了额外的计算量。常见的多尺度学习方法有图像金字塔(Image Pyramid)、特征金字塔(Feature Pyramid Network,FPN)等。

3)上下文学习。一个事物不是单独存在的。目标检测过程当中应综合考虑目标本身的信息以及目标周围的信息,包括背景、环境等,即上下文学习。比如,我们可以从正在行驶的汽车当中推测出驾驶室有一个人。有研究表明,通过对目标的上下文信息进行建模,可以提升神经网络的检测性能,尤其是对于外观特征不明显的小目标。目前,基于上下文学习的深度学习网络也有了一些发展,主要有基于隐式上下文特征学习的目标检测方法和基于显式上下文推理的目标检测两种。

4)生成对抗学习。生成对抗网络(Generative Adversarial Networks,GANs)是一种神经网络,它以随机噪声为输入并生成输出,主要包括生成器与判别器,通过训练输出尽可能与训练集分布相似的样本。利用生成对抗网络,可缩小小目标与中/大型目标之间的特征差异,使得小目标的特征表达更

加明显,更易被特征网络捕获到特征信息,进而提升对小目标的检测效果。相较于数据增强、多尺度学习、上下文学习3种通过增加计算成本来提升检测性能的手段,生成对抗学习的性能增益主要来源于生成对抗网络的训练效果,因此如何在生成器与判别器之间进行取舍和如何保证训练输出样本的多样性是提升检测性能的关键。

本文以主流单阶段检测算法YOLO系列的YOLOv5为基础,考虑到LSS无人机目标检测的实际应用场景,且优先保证目标检测的实时性,同时实现对LSS无人机目标检测的高速与高精度并存,通过改进网络结构、优化目标锚框、数据增强、优化误差函数等方式提升算法性能,使之更加适用于实际检测场景。

2 YOLOv5 网络结构

为了解决Faster-RCNN检测速度较慢的问题,2016年YOLO算法应运而生,历经7个版本。由于YOLOv6与YOLOv7于2022年7月刚刚发布,工程应用还不是十分成熟,YOLOv5是当前YOLO系列应用最广的目标检测算法。YOLOv5的网络结构如图1所示。

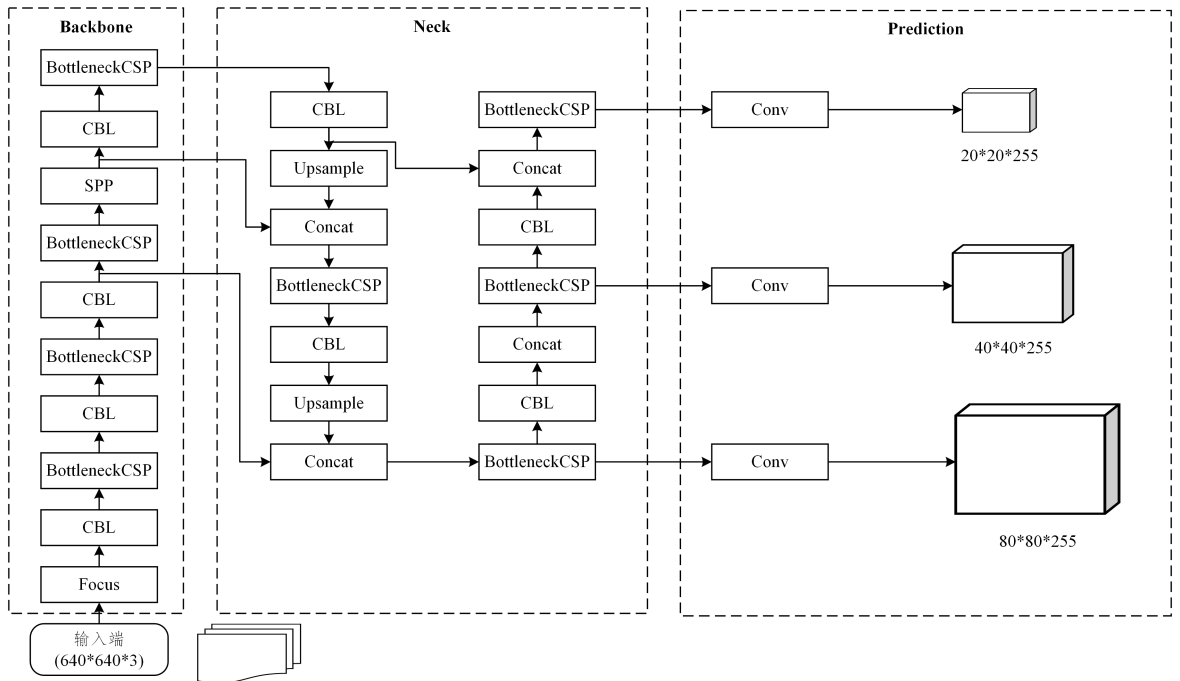


图1 YOLOv5网络结构示意图

Fig. 1 YOLOv5 network structure

2.1 输入端

YOLOv5网络的输入图像大小为 640×640 ,为了增强网络对多种类型图片的适应能力,在输入端通常会通过压缩、扩展、填充等预处理的方式,将输入图片归一化,从而便于后续网络的处理。在网络训练阶段,主要采用了以下手段:

1)Mosaic数据增强:通过随机选取、随机剪裁、随机组合的方式,将4张图片组成1张新的图片作为样本。通过此种数据增强方式,可提高数据集样本的丰富程度,还可以极大地提升训练速度。

2)自适应锚框计算:不同的检测目标长宽不同,长宽比也不同,在对模型进行训练时需要针对不同的训练数据设定大小不同、比例不同的特定锚框。在训练过程中,根据实际训练

样本的不同,会根据预测框与真实框之间的差距对预设锚框进行一定的修正,从而更新整个网络的参数,因此设定初始锚点框也是比较关键的一环。

3)自适应图片缩放:YOLOv5网络的输入格式为固定 $640 \times 640 \times 3$ 的图片,针对不同的数据集,通过缩放操作将图片格式进行转换,对于长宽比例不满足1:1的图片,还需进行padding操作,在填充过程中算法能够自适应地填充最少的空白数据,从而增强算法对数据格式的兼容性。

2.2 主干网络(Backbone)

YOLOv5的主干网络借鉴了一些其他检测算法的思路,使用的结构主要包括:Focus结构、CSP结构和SPP模块。

1)Focus结构:通常位于主干网络前端,在图像进入网络

前,对其进行切分操作,Focus层的主要作用就是为后续的特征提取保留更完整的图像下采样信息。对于一张图片的RGB 3个通道,在每个通道分别进行分割处理,在一张图片中每间隔一个像素取一个值组成一张小的新图片,以同样的方法对剩余像素进行处理,这样就可以将1张图片分割成为4张互补的图像,3个通道总共就可以得到12张图片,并且在处理过程当中没有信息丢失。如图2所示,输入图像的大小为 $640 \times 640 \times 3$,利用切分操作,可以得到 $320 \times 320 \times 12$ 的特征图,再经过一次卷积,最终变成 $320 \times 320 \times 32$ 的特征图。

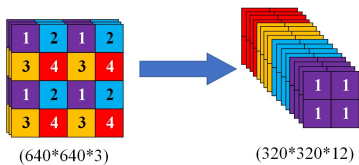


图2 Focus操作示意图

Fig. 2 Schematic of Focus operation

2) CSP结构:跨阶段局部网络(Cross Stage Partial)。神经网络在训练过程中往往计算量过大,一个很重要的原因是参数优化过程中梯度信息重复。而CSP结构将梯度变化与特征图相结合,在不降低准确率的同时大大减少了计算量,增强了模型的特征学习能力。此外,由于计算量的降低也减少了对内存的消耗,提高了资源利用率,减少了不必要的浪费。

3) SPP模块:空间金字塔池化(Spatial Pyramid Pooling, SPP)模块是一个通用的卷积神经网络框架,作为计算机视觉领域最成功的方法之一,SPP层常常出现在目标检测识别算法框架中,能将任意大小的特征图转换成固定大小的特征向量,消除了传统CNN对固定尺寸输入图像的限制,提升了网络的识别精度和速度。

2.3 Neck网络

Neck网络通常位于主干网络和预测网络的之间,对特征图进行一定的处理,可以进一步提升特征的多样性及鲁棒性。YOLOv5在Neck网络主要采用了FPN+PAN结构(见图3)。

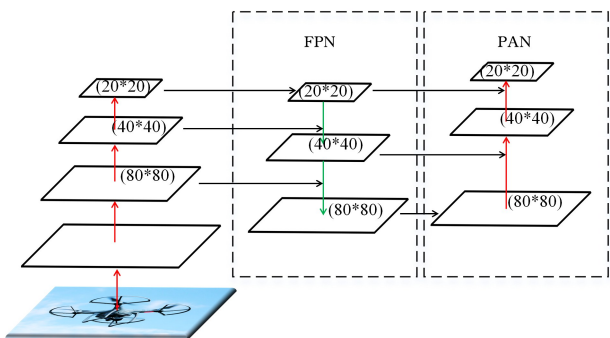


图3 FPN+PAN模块

Fig. 3 FPN+PAN module

特征金字塔网络(FPN),通过横向连接、自上而下的体系结构融合了语义信息强、低分辨率的特征信息与语义信息弱、高分辨率的特征信息。但是,其仅仅关注深层网络的语义特征,容易忽略掉浅层网络所包含的细节信息,会影响识别与定位的精度。

路径聚合网络(Path Aggregation Network, PAN)通过加强低层神经元的反应强度增强了特征金字塔,提高了网络对

图像细节信息的处理能力,使低层网络的精确定位信息更容易传导至高层网络,从而加强网络特征融合能力。

2.4 预测部分(Prediction)

将边界框的损失函数由完全交并比(CIoU)损失改进为广义交并比(GIoU)损失;在目标检测后处理过程中,YOLOv5采用加权非极大值抑制(Non-Maximum Suppression, NMS)运算对多个目标锚框进行筛选。

1) GIoU:GIoU来源于IoU,是一种对比预测边框与真实标注边框差异的损失计算方法。

2) 加权NMS运算:在目标检测的后期处理过程中,针对多目标框的筛选,通常需要进行NMS操作。结合不同损失函数的计算,选择合适的NMS算法,可以实现对遮蔽目标的检测。

3 YOLOv5网络结构改进

3.1 增加小目标检测层

本文在对YOLOv5网络结构进行深入分析的基础上,采用了增加小目标检测层的手段对网络结构进行优化,目的是使算法能够满足对微小型目标的检测要求,提高对小目标的检测精度。

为适应不同大小目标的检测,原始YOLOv5网络结构共设计3种不同尺寸的特征图。由卷积神经网络的结构分析可得出,随着网络深度的不断增加,所获得的特征图尺寸更小、更抽象,具有更大的感受野,包含更多的语义信息,因此小尺寸的特征图往往用来检测大目标;浅层卷积得出的特征图虽然语义信息不是十分丰富,但目标位置数据会更加精确,故浅层网络输出的特征图更加适合检测小目标。因此,为了增强网络结构对小目标的检测能力,需在FPN结构中添加更大尺寸的特征图,以提高对微小目标的检测能力,主干网络也应作出相应的改进。优化后的网络结构如图4所示,在Neck网络中增加小目标检测层,输出更大尺寸的特征图用于进行特征融合,从而提升对小目标的检测精度。

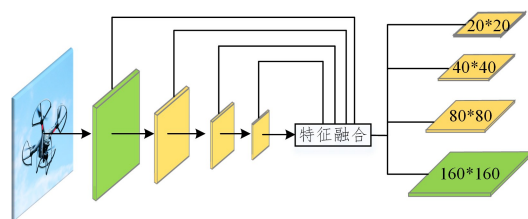


图4 增加小目标检测层

Fig. 4 Add small object detection layer

3.2 优化目标锚框

YOLOv5所设计的锚框分为大目标、中目标、小目标共三类,且每个类别下根据长宽比例的不同各设置3个比例的锚框,即YOLOv5共使用9个预设锚框进行目标的识别。但是,YOLOv5目标锚框设计的初衷主要是满足COCO数据集的目标检测。在COCO数据集中共包括有“人物”“雨伞”“汽车”“鸟”等80类,大小各有不同,长宽比例亦各不相同。而本文所研究的检测目标为小型旋翼无人机,针对这一特殊目标设计专门的锚框,可改善算法的检测精度。

为设计出适用于小型旋翼无人机目标检测的锚框,首先,准备自制数据集并完成数据标注,数据集共含有小型无人机的图片5361张;其次,采用无监督学习中的聚类算法进

行数据分类。 K 均值(K -mean)算法是经典的聚类算法之一,其基本原理是从众多数据中随机选取 k 个数据作为聚类中心,通过计算距离度量数值,不断更新迭代更新聚类中心,最后选出包括 k 个数据点的中心簇。但由于 K -mean 算法受初始点影响较大,而后出现了改进 K -mean 算法,即 K -mean++ 算法。 K -mean++ 算法对初始中心簇的选择进行了优化:对第一个点的选择仍然是随机的,从第二个点开始,第 n 个中心点选择与前 $n-1$ 个点相距最“远”的数据点。虽然 K -mean++ 算法带来了一定的计算复杂度,但是减小了初始点选择对算法精度的影响。从图 5 中可以看出,当 $2 \leq k \leq 19$ 时, K -mean++ 算法的精度明显高于 K -mean 算法。进一步,随着中心点个数 k 的增加,平均交并比(Mean Intersection over Union, Mean IoU)数据也相应增加,说明聚类效果逐渐提升。

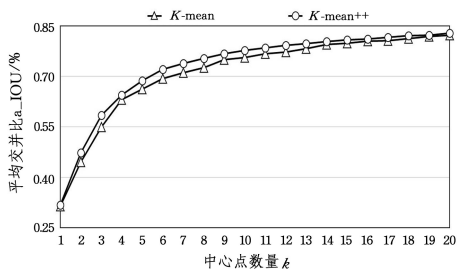


图 5 不同簇中心数量与平均交并比的关系

Fig. 5 Relationship between the number of centers of different clusters and mean IoU

为了验证目标锚框优化效果对目标检测精度和检测算法的影响,在不改变整个深度学习网络架构的基础上,选取与原 YOLOv5 算法相同的锚框数量,即选取中心点个数 k 为 9。通过 K -mean++ 算法计算并对应大、中、小 3 类锚框,通过归一化处理,得出优化后的目标锚框如表 1 所列。

表 1 改进后的预设先验框尺寸
Table 1 Improved anchor box size

特征图 尺寸	先验框尺寸		
	先验框 1	先验框 2	先验框 3
小尺寸	(20,13)	(27,17)	(38,23)
中尺寸	(57,34)	(90,51)	(131,84)
大尺寸	(255,120)	(368,202)	(586,290)

3.3 数据增强

目前公开的专用于进行小型旋翼无人机目标检测的训练集十分有限,且样本质量不高。故通过自制数据集,在小样本的条件下,采取数据增强的手段提升训练效果。

原始的 YOLOv5 网络本身就采用了数据增强手段,即 Mosaic 数据增强方法,其效果如图 6 所示。

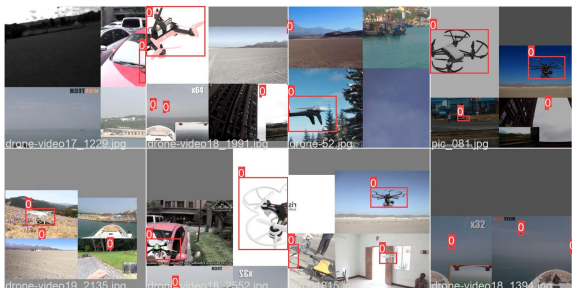


图 6 Mosaic 数据增强

Fig. 6 Mosaic data augmentation

原始的 Mosaic 数据增强方法存在较大的随机性,对训练效果的提升不是十分显著,考虑到所检测旋翼无人机的特点,本文提出一种适用于小目标检测网络训练的改进型 Mosaic 数据增强方法——Mosaic-SOD(Mosaic for Small Object Detection)。其基本思路如下:首先将目标背景进行分离,分设目标数据集和背景数据集,而后从背景数据集中任选 4 张进行拼接组成新的背景,从目标数据集中选取一定数量的无人机目标与拼接后的背景再进行拼接。在实际拼接过程中,背景可分为天空、森林、建筑等,通过随机拼接提升背景的复杂度。选取 3~5 个不同类型、不同姿态的无人机目标,通过随机复制粘贴的方式,将目标与拼接产生的背景进行结合。由于目标尺寸较小,可大大增加目标的数量(在实际训练过程中选择 10~20 个目标)进行训练,从而提升网络泛化能力。其具体流程如图 7 所示。

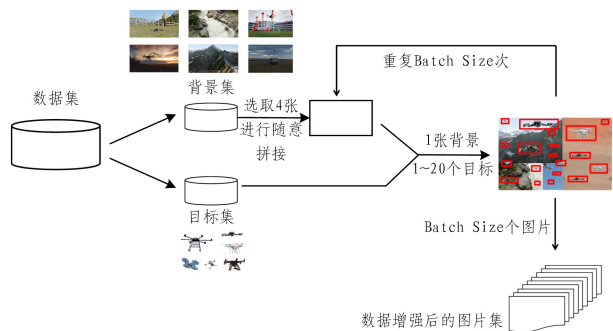


图 7 Mosaic-SOD 数据增强处理流程

Fig. 7 Data processing of Mosaic-SOD data augmentation

3.4 优化损失函数

损失函数的调整和计算对于神经网络的稳定性和输出结果的准确性至关重要。目标检测模型的损失函数通常由 3 部分组成,分别是分类损失、置信度损失和定位损失。本文分别选用 BCEWithLogits Loss、Focal Loss 和 SIOU Loss 来评价 3 种损失。总体损失定义如式(1)所示:

$$L = L_{\text{Class}} + L_{\text{conf}} + L_{\text{SIOU}} \quad (1)$$

分类损失常用二元交叉熵损失函数来进行计算, BCE-WithLogits Loss 包括 Sigmoid 层和 BCE Loss 层,适用于多标签分类任务。首先对预测输出作 sigmoid 变换,然后求变换后的结果与真实值的交叉熵,其定义如式(2)所示:

$$L_{\text{class}} = - \sum_{i=1}^N y_{(i)} \log \hat{y}_{(i)} + (1 - y_{(i)}) \log (1 - \hat{y}_{(i)}) \quad (2)$$

其中, N 表示一个批次中图片的数量, $y_{(i)}$ 表示样本 i 的类别标签, $\hat{y}_{(i)}$ 表示样本 i 预测为正的的概率。

置信度损失主要解决一阶段目标检测方法中正负样本比例不均衡问题,本文算法中的置信度损失使用的是 Focal Loss,定义如式(3)所示。

$$L_{\text{conf}} = - \alpha_i (1 - p_i)^{\tau} \log(p_i) \quad (3)$$

其中, p_i 是不同类别的概率, α_i 主要是解决正负样本的不平衡, τ 主要是解决难易样本的不平衡,分别被设置为 0.25, 2。

定位损失往往采用 IoU 损失来进行计算。IoU 是一种对比预测边框与真实标注边框差异的损失计算方法。当预测边框与真实标注边框不重叠时, IoU 值为 0, 不能度量样本距离的远近程度。为弥补 IoU 的不足,计算定位损失时应同时考虑重叠区域, 纵横比等, 相继出现了 GIoU, CIoU 等。然而, 迄今为止提出和使用的方法都没有考虑到所需真实框与预测框

之间不匹配的情况。这种不足导致收敛速度较慢且效率较低,因为预测框可能在训练过程中“四处游荡”并最终产生更差的模型。针对以上问题,采用 SIOU 损失函数替换原损失函数。SIOU 损失函数由 4 个 cost 函数组成:Angle cost, Distance cost, Shape cost 和 IoU cost。

Angle cost 添加了角度感知组件,当计算出真实框与预测框之间的差距时,能够感知真实框相对于预测框的方位,从而在后续优化过程中更具有针对性,模型将尝试将预测框沿着 X 轴或 Y 轴(选最近的一个)向真实框的方向移动,如图 8 所示。

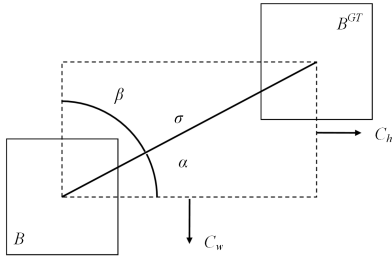


图 8 Angle Loss 的计算示意图

Fig. 8 Calculation schematic of Angle Loss

如果 $\alpha \leq \frac{\pi}{4}$, 将优先对 α 进行最小化, 否则优先对 β 进行最小化。

$$\beta = \frac{\pi}{2} - \alpha$$

角度感知组件通过式(4)定义:

$$\Lambda = 1 - 2 * \sin^2 \left(\arcsin(x) - \frac{\pi}{4} \right) \quad (4)$$

其中,

$$x = \frac{C_h}{\sigma} = \sin \alpha$$

$$\sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2}$$

$$C_h = \max(b_{c_x}^{gt}, b_{c_x}) - \min(b_{c_x}^{gt}, b_{c_x})$$

考虑到 Angle cost 的变化, Distance cost 被重新定义, 如图 9 所示。

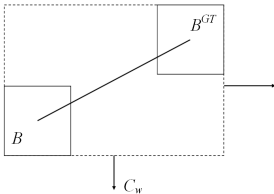


图 9 Distance Loss 的计算示意图

Fig. 9 Calculation schematic of Distance Loss

Distance cost 如式(5)所示:

$$\Delta = \sum_{l=x,y} (1 - e^{-\rho_l}) \quad (5)$$

其中,

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{C_w} \right)^2$$

$$\rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{C_h} \right)^2$$

$$\gamma = 2 - \Lambda$$

Shape cost 的定义如式(6)所示:

$$\Omega = \sum_{l=w,h} (1 - e^{-\omega_l})^\theta \quad (6)$$

其中,

$$\omega_w = \frac{|\tau w - \tau w^{gt}|}{\max(w, \tau w^{gt})}$$

$$\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})}$$

θ 的值定义了 Shape cost 的大小, 并且 θ 的值对于每个数据集都是唯一的。它决定了 Shape cost 的受关注程度, 如果 θ 设置为 1, 将立即优化形状, 从而使形状的移动受到影响。通过实验计算可知, θ 值接近于 4, 故往往在 2~6 之间选取。如图 10 所示, 灰色部分为真实锚框与预测锚框的重叠部分, 重叠部分占比越大代表 IoU cost 越小, 此时的预测锚框更加准确。

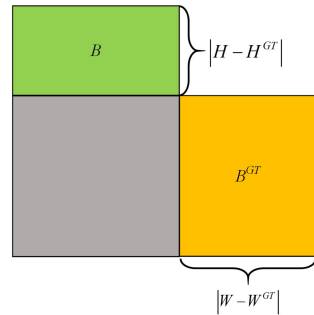


图 10 IoU 组件示意图(电子版为彩图)

Fig. 10 Schematic of IoU components

回归损失函数如式(7)所示:

$$L_{\text{SIOU}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \quad (7)$$

其中,

$$\text{IoU} = \frac{|B \cap B^{GT}|}{|B \cup B^{GT}|}$$

4 实验结果与分析

将改进后的算法应用于小型旋翼无人机目标的检测场景下, 并与原始算法进行比较。本文主要以小型旋翼无人机目标为研究对象进行检测, 以满足空天领域识别低空慢速小型目标的现实需求。

4.1 数据集

本文数据集来源于 Kaggle 公开数据集、网络和自制数据集, 通过人工标注的方式制作不同背景下含有无人机目标和未含有无人机目标的图片, 最终得到 5876 张图片, 其中训练数据集 4500 张, 测试数据集 1376 张。

利用 Labelimg 软件对数据集进行标注, 并将所得 .xml 格式的文本文件转换为适用于 YOLOv5 算法的 .txt 文件。每一个 .txt 文件中信息如图 11 所示。

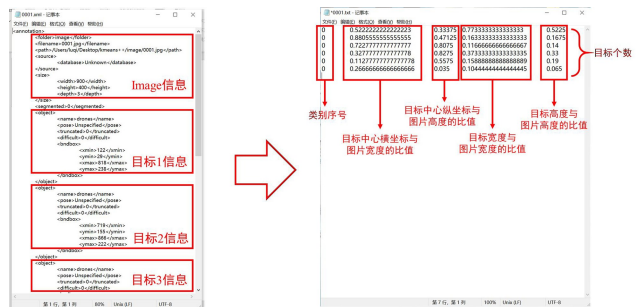


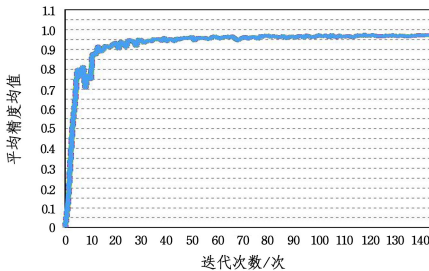
图 11 标签文件格式转换

Fig. 11 Transformation of label file

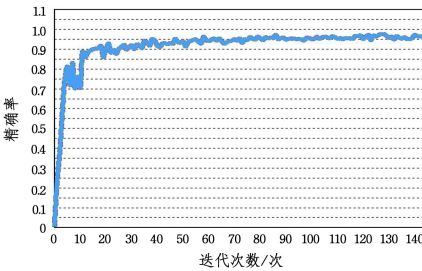
4.2 实验环境与模型训练

实验环境使用 macOS Monterey 12.4 操作系统,处理器为 Apple M1 Pro 芯片,python 语言版本为 3.9.11,Pytorch 版本为 1.11.0,原始算法采用 YOLOv5 v6.0 版本。训练过程总迭代次数为 150 次,迭代批量大小设置为 32。

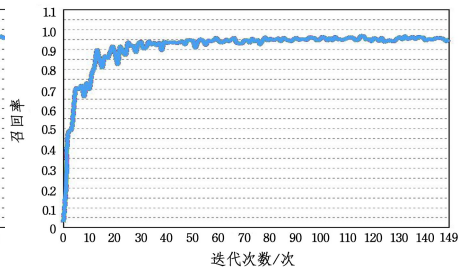
在同样的环境下,设置对照组来比较算法的优化效果。第一组为原始 YOLOv5 模型,不进行优化;第二组利用增加小目标检测层的方法进行优化;第三组利用优化锚框的方法进行优化;第四组利用改进损失函数的方法进行优化;第五组



(a)mAP 变化曲线



(b)Precision 变化曲线



(c)Recall 变化曲线

图 12 模型训练过程

Fig. 12 Process of model training

如表 2 所列,为验证本文算法的有效性,利用上文所提到的增加小目标检测层、优化预设锚框、数据增强、优化损失函数方法改进后的算法与原始 YOLOv5 算法在同一测试集上进行测试的结果。

表 2 改进前后 YOLOv5 算法性能对比结果

Table 2 Performance comparison results of YOLOv5 algorithm before and after improvement

算法	小目标层	锚框优化	数据增强	损失函数	mAP	FPS
YOLOv5					0.923	42.2
YOLOv5_1	✓				0.948	33.3
YOLOv5_2		✓			0.941	40.2
YOLOv5_3			✓		0.934	39.1
YOLOv5_4				✓	0.926	38.8
YOLOv5_5	✓	✓	✓	✓	0.964	31.4

从表 2 中可以看出,改进后的算法对小型旋翼无人机目标的平均精确率 mAP 值达到了 96.4%,相较于原始 YOLOv5 算法提高了 4.1 个百分点。但是由于运算量的增加,本文算法的检测速度略有下降,FPS 为 31.4,下降了 10.8,检测单张图片的时间由 23.7 ms 增加到了 31.8 ms。常用的可见光视频帧率为 25 fps,故为满足检测实时性的要求(大于 25 fps),要求算法检测速率大于 25 fps,可见改进后的算法仍能够满足检测实时性的要求。

为验证算法性能,首先将本文改进后的 YOLOv5 算法与主流两阶段检测算法 Faster R-CNN、基于 Transformer 的检测算法 DETR 进行比较,重点从检测精度与检测速度两个方面对各算法性能进行比较,结果如表 3 所列。可见,目前改进后的 YOLOv5 单阶段检测算法在检测精度和检测速度上均优于 Faster R-CNN 与 DETR 算法。

如图 13 所示,横坐标为训练迭代次数,纵坐标为预测边框的损失值,可以看出改进后的算法 Loss 值下降更快,训练效率更高。

综合利用增加小目标检测层、优化锚框、改进损失函数的方法进行优化。

4.3 评估指标与实验结果分析

本文实验使用深度学习评分指标,例如平均精确率(mean Average Precision, mAP)、精确率(Precision)和召回率(Recall)。

图 12 分别展示了算法在训练过程中的平均精度均值、精确率曲线和召回率曲线。准确率指预测为正样本的数据中有多少个是真正的正样本。召回率从真实结果角度出发,描述了真实的正样本有多少被该分类器召回。

表 3 实验结果对比

Table 3 Comparison of experimental results

算法	Precision	Recall	mAP	FPS
Faster R-CNN	0.819	0.856	0.826	5.6
DETR	0.902	0.898	0.918	9.8
YOLOv5	0.918	0.921	0.923	42.2
本文算法	0.961	0.958	0.964	31.4

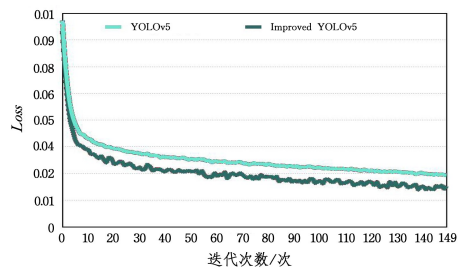


图 13 Loss 值下降曲线

Fig. 13 Decline curve of Loss value

图 14 为 mAP 指标曲线图,其中 mAP_0.5 表示 IoU 的阈值为 0.5 时对应的 mAP,改进后的算法精度有了一定的提升。

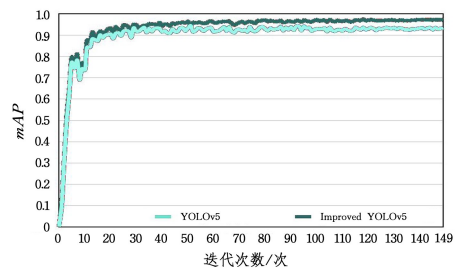


图 14 mAP 曲线图

Fig. 14 mAP curve

图 15 给出了改进前后算法在测试集上的小型旋翼无人机检测结果对比图。可以看出改进后的算法对于密集分布的

小型目标检测效果更好,漏检率更低。

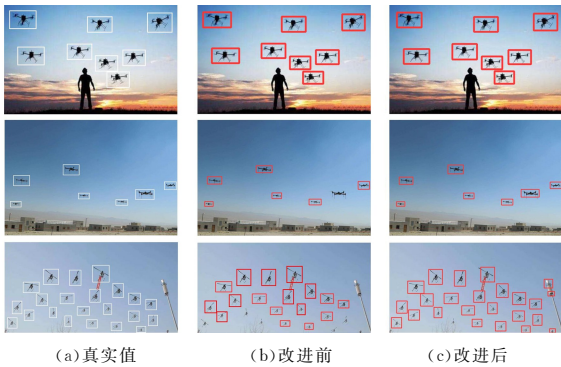


图 15 改进 YOLOv5 的测试结果

Fig. 15 Test results of improved YOLOv5

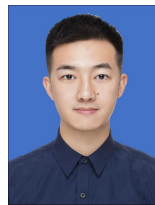
结束语 本文在原始 YOLOv5 算法的基础上,分别从网络结构检测层、目标锚框优化、数据增强和损失函数优化等方面进行改进,改进后的算法性能提升主要体现在两个方面:训练效率和小型旋翼无人机特定目标的检测精度。改进后的算法在计算复杂度上有所增加,导致对单张图片的处理速度有所下降,但考虑到小型旋翼无人机目标实际检测背景下,FPS 还在可接受的范围内,能够满足检测实时性的要求,可直接用于对小型旋翼无人机目标的检测,促进反无人机系统的智能化发展。

参 考 文 献

- [1] GRANT R. RPAs FOR ALL [J]. Air Force Magazine, 2012, 95(8): 54-57
- [2] XIE X. US Army releases anti-drone technology manual (Part 1)[J]. Modern Military, 2017(7): 91-100.
- [3] XIE X. US Army releases anti-drone technology manual (Part 2)[J]. Modern Military, 2017(8): 85-89.
- [4] GUELF E A, BUDDHIKA J, ROBISON T. The Imperative for the U. S. Military to Develop a Counter-UAS Strategy[J]. JFQ: Joint Force Quarterly, 2020(97): 4-12.
- [5] MA W, CHIGAN X X. Research on Development of Anti-UAV Technology[J]. Aero Weaponry, 2020, 27(6): 19-24.
- [6] JIE C, MIAO Z, YE T T. Research on the development of anti-

UAV systems in U. S. military service[J]. Airborne Missile, 2020(12): 36-42.

- [7] CHEN W. A Survey on Low-Slow-Small UAV Countermeasure Equipment and Theoretical[J]. Henan Science and Technology, 2022, 41(2): 10-13.
- [8] ZHU M Z, CHEN X, LIU X, et al. Situation and key technology of tactical laser anti-UAV[J]. Infrared and Laser Engineering, 2021, 50: 188-200.
- [9] WU Y F. Research on LSS-Target(the Low altitude, Slow speed and Small target) in complex background [D]. Changchun: Changchun university of Chinese Academy of Sciences, China(Chinese Academy of Sciences Institute of Optical Precision Machinery and Physical), 2020.
- [10] XU D M, ZHANG H w. Overview of Radar LSS Target Detection Technology[J]. Modern Defence Technology, 2018, 46(1): 148-155.
- [11] MAO T. Research on Small Object Detection Algorithm Based on YOLOv5[D]. Anhui University of Science and Technology, 2021.
- [12] REN J, WANG Z J, ZHANG Y F, et al. YOLOv5-R: lightweight real-time detection based on improved YOLOv5[J]. Journal of Electronic Imaging, 2022, 31(3): 033033.
- [13] XIANG X Z, WANG Z Y, QIAO Y L. An Improved YOLOv5 Crack Detection Method Combined with Transformer[J]. IEEE Sensors Journal, 2022, 22(14): 1
- [14] JIANG X K, HU H C, LIU X, et al. A smoking behavior detection method based on the YOLOv5 network [J]. Journal of Physics: Conference Series, 2022, 2232: 012001.
- [15] DONG X D, YAN S, DUAN C Q. A lightweight vehicles detection network model based on YOLOv5[J]. Engineering Applications of Artificial Intelligence, 2022, 113: 104914.



LU Qi, born in 1994, master. His main research interests include intelligent detection and automatic target recognition.