

文章编号: 1671-7449(2024)04-0354-09

改进YOLOv5的无人机小目标检测算法

李松林, 江剑*

(南京理工大学 机械工程学院, 江苏 南京 210094)

摘要: 小目标由于可用特征少、纹理模糊等因素, 一直是目标检测领域中的一个难点。针对无人机小目标检测中的误检与漏检问题, 提出了一种无人机小目标检测算法LASD-YOLOv5。设计了一种极化自注意力机制, 以更准确地提取微小特征; 引入加权双向特征金字塔网络, 替换路径聚合网络, 以加强对底层特征的利用, 对检测头进行解耦, 以提高模型的收敛速度。同时, 针对当前无人机小目标数据集中小目标占比少与场景不全面的问题, 贡献了一个多场景低慢小无人机目标数据集LASD-D。结果表明, 所提算法在LASD-D数据集上的平均精度为98.29%, 相比原网络提升了2.87%, 同时也优于YOLOv7, YOLOv8与QueryDet等主流算法, 完全满足无人机小目标检测领域的需求。

关键词: 小目标检测; 无人机; 注意力机制; 特征融合; YOLOv5

中图分类号: TP391.4

文献标识码: A

doi: 10.3969/j.issn.1671-7449.2024046

引用格式: 李松林, 江剑. 改进YOLOv5的无人机小目标检测算法[J]. 测试技术学报, 2024, 38(4): 354-362.

LI Songlin, JIANG Jian. Improved UAV small object detection algorithm based on YOLOv5[J]. Journal of Test and Measurement Technology, 2024, 38(4): 354-362.

Improved UAV Small Object Detection Algorithm Based on YOLOv5

LI Songlin, JIANG Jian*

(School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: The detection of small-scale targets, characterized by limited available features and unclear textures, has perennially posed a challenge in the field of object detection. To address issues related to false positives and false negatives unmanned aerial vehicle (UAV) targets, we propose an improved UAV small-target detection algorithm, termed LASD-YOLOv5. This algorithm introduces a polarized self-attention mechanism to more accurately extract minute features, incorporates a weighted bidirectional feature pyramid network to replace the path aggregation network, thus enhancing the utilization of low-level features. Furthermore, it decouples the detection head to expedite model convergence. Additionally, to tackle the scarcity of small targets and incomplete scene coverage in existing UAV small-target datasets, we contribute a multi-scene, low-speed, small UAV target dataset (LASD-D). The experimental results demonstrate that our proposed algorithm achieves an average precision of 98.29% in LASD-D, surpassing the baseline network by 2.87%. Notably, it outperforms mainstream algorithms such as YOLOv7, YOLOv8 and QueryDet, effectively meeting the demands of UAV small-target detection applications.

Key words: small object detection; UAV; attention mechanism; feature fusion; YOLOv5

近年来随着无人机领域蓬勃发展, 消费级无人机, 尤其是低空慢速小型无人机的数量也在迅速

收稿日期: 2023-09-23

作者简介: 李松林(1998-), 男, 硕士生, 主要从事目标检测与图像处理等研究。E-mail: 1720674172@qq.com。

*通信作者: 江剑(1973-), 男, 副教授, 硕士生导师, 主要从事目标检测与图像处理等研究。E-mail: 83905118@163.com。

速增加,因此,衍生出的暴力恐怖袭击、低空失密泄密、扰乱公共秩序等公共安全事故越来越多,时刻威胁着社会稳定和人民生命财产安全^[1]。因此,无人机防御问题尤其是低慢小无人机目标的检测受到广泛关注。低慢小无人机目标检测是小目标检测的一种,其目的是在无人机仅占图像面积约 1% 区域大小的早期阶段检测到目标并予以反制,该研究在低空反恐、智能监控等方面有着重要的意义和发展前景。

自从 2012 年 AlexNet 网络横空出世以来,神经网络一直迅速发展,并在图像检测领域得到广泛的应用^[2]。目前,基于深度学习的目标检测算法主要划分为一阶段算法与二阶段算法两类。一阶段算法可以通过神经网络计算直接输出目标的类别与位置,代表算法有 YOLO 系列算法^[3-5]、SSD 算法^[6]、CornerNet 算法^[7]以及 QueryDet 算法^[8]等。二阶段算法则是将目标检测划分为候选区域生成、分类与边界框调整两个阶段,代表算法为 RCNN 系列目标检测算法^[9-11]。近年来,YOLOv5 算法凭借其在精度和实时性方面的出色表现广泛应用于目标检测领域,但是在小目标检测方面,由于可用特征少、目标尺度分布不均衡、复杂背景干扰等问题导致的漏检、误检问题仍然有待国内外学者探索与解决。针对小尺寸无人机目标检测精度低、深层网络参数量大等问题,王建楠等^[12]裁剪掉 YOLOv5 冗余的网络层以减少网络参数量,引入多个不同采样率的空洞卷积,以增强小目标的多尺度特征提取能力。改进后的模型检测速度提高了 10.3%,内存开销节约了 65%,但其模型未很好地解决动态模糊等问题。Luo 等^[13]基于残差模型,将 Inception 块与多尺度卷积组相结合,并借助跳跃连接和直接级联构建了一个混合模型,用以识别低慢小无人机目标。该模型利用剩余连接,加速了训练且满足了实时要求,同时,利用放大的卷积核接受更多的特征,从而消除动态模糊中的噪声,较好地解决了动态模糊问题与尺度问题,但其训练数据集体量较小,且对于真实场景中复杂背景干扰等问题并未涉及。包文歧等^[14]基于 YOLOv5,对体积小、运动灵活的微小无人机进行实时探测,制作了一个包含多飞行姿态的无人机小目标数据集,该模型可在 30 m 范围内准确识别出特征尺寸为 200 mm 以上的无人机,但该研究并未针对小目标检测问题对 YOLOv5 进行优化,其模型精度与实时性还

有提升的空间。Wang 等^[15]设计了特征增强模块与空间感知模块,并将其与 YOLOv5 相结合,以解决小目标检测中的背景信息干扰问题,增大有效感受野。改进使模型的小目标检测性能得到了一定的提升,但也使模型的运算量增大。

针对无人机小目标检测中因可用特征少、目标尺度分布不均衡、复杂背景干扰而导致的误检与漏检问题,本文提出了改进 YOLOv5 的无人机小目标检测算法 LASD-YOLOv5(Low Attitude Slow Drone YOLOv5),在保证检测速度的前提下,提升了检测精度。主要贡献如下:

1) 针对小目标可用特征过少且容易被复杂背景干扰等问题,引入极化自注意力机制(Polarized Self Attention, PSA)改进骨干网络,将 C3(Concentrated Comprehensive Convolution)模块与 PSA 机制相结合,设计了 C3PP(Concentrated Comprehensive Convolution Parallel Polarized)模块,提升了模型对于微小特征的筛选提取能力。

2) 针对随着网络层数增加,小目标特征信息与位置信息逐渐丢失而导致的漏检问题,引入加权双向特征金字塔网络(Bidirectional Feature Pyramid Network, BiFPN),将浅层的表征信息和深层的语义信息进行多尺度特征融合,有效提升了模型的小目标检测性能。

3) 针对耦合头部中存在的分类与回归任务的冲突问题,将其替换为解耦的头部,以提高模型性能与收敛速度。

4) 针对现有低慢小无人机目标数据集场景与种类不全面、数据集中小目标占比少的现状,贡献了一个包含多场景的无人机小目标数据集 LASD-D(Low Attitude Slow Drone Dataset)。

1 LASD-YOLOv5 模型

1.1 LASD-YOLOv5 模型的提出

YOLOv5 是由 Ultralytics 于 2020 年 5 月出品的一款一阶段检测器,具有模型小、速度快、适用于移动端部署等特点^[16]。该模型有 4 个版本,分别是 YOLOv5s, YOLOv5m, YOLOv5l 和 YOLOv5x,它们的网络结构基本相同,都是由输入端、骨干网络(Backbone)、颈部网络(Neck)和预测层(Prediction) 4 个部分组成,仅在模型深度和模型宽度上有区别。YOLOv5 在输入端采用了 Mosaic 数据增强方法,从而丰富数据集、增强模型鲁棒性以及提升小目标检

测性能。骨干网络由Focus模块和跨阶段局域网络(Cross Stage Partial Network, CSPNet)组成。其中Focus模块将输入图片进行切片处理,然后将处理后的片段进行拼接,以提高模型的计算能力并避免信息丢失。CSPNet的残差结构使得梯度信息传递更加高效,从而有效减少了推理计算的负担。颈部网络采用了特征金字塔网络(Feature Pyramid Networks, FPN)与路径聚合网络(Path Aggregation Network, PAN)相结合。FPN将高分辨率特征图与低分辨率特征图融合,传递强语义信息。PAN将含有强位置信息的高分辨率特征图与强语义信息的低分辨率特征图融合,传递强位置信息。预测层使用CIoU(Complete Intersection over Union)损失函数,在3个尺度上进行检测,并通过非极大值抑制选取最优的预测框。

YOLOv5在目标检测领域表现良好,但在小目标检测中的表现却不尽如人意,具体原因有:①骨干网络对于无人机小目标的特征提取能力较弱,漏检问题明显;②当颈部融合不同层级特征时,浅层特征一定程度上受到深层特征的干扰,导致定位不准确。本文针对上述不足对YOLOv5模型进行了改进,提出了LASD-YOLOv5模型,其具体结构如图1所示。首先,设计了C3PP模块代替原网络中的CSPNet,具体包括骨干网络中的CSP1_1, CSP1_3结构以及颈部网络中的CSP2_1结构,通过调整和加权通道注意力与空间注意力,显著增强了特征的特征能力;其次,在颈部网络使用BiFPN结构优化了PAN结构,将多层特征以双向交叉连接方式进行融合。最后,对耦合的检测头部位进行解耦,使得分类任务和回归任务相互独立。

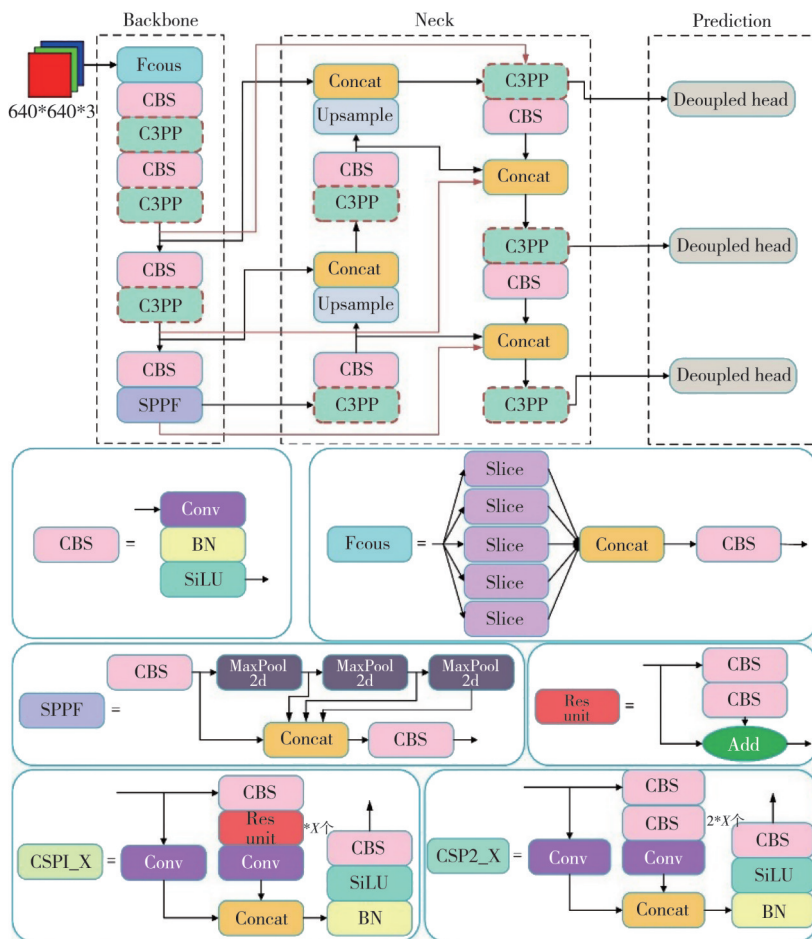


图1 LASD-YOLOv5网络结构图

Fig. 1 LASD-YOLOv5 network architecture

1.2 引入极化自注意力机制

无人机小目标存在尺寸小、特征分布稀疏、纹理模糊等问题,传统卷积操作对于此类微小特征的

捕获能力有限,难以提取相关特征。同时,当图像中存在复杂背景时,背景信息的干扰也会致使微小特征难以被准确定位。为了解决上述问题,本文引入了PSA机制^[17],并结合设计了C3PP模块。PSA

是一种双重注意力机制，最初用于人体姿态估计和语义分割任务。与其他注意力机制相比，PSA 机制有两个主要特点：首先，在通道和空间维度上保持着高分辨率，减少了信息损失；其次，通过组合非线性直接拟合典型细粒度回归的输出分布，适用于像素级回归任务。为了更好地理解和利用输入特征，本文将C3模块与PSA机制结合，设计了C3PP模块，具体结构如图2所示。C3PP模块通过1×1的卷积操作调整通道数，并使用深度可分离卷积操作降低计算量。然后，对特征图进行通道与空间注意力处理，再将处理后的特征图与输入特征图融合，得到两个输出特征图。最后，融合两个输出特征图得到最终的输出特征。

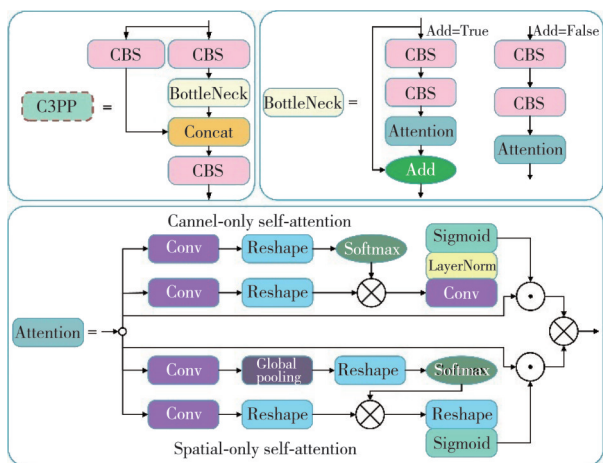


图2 C3PP模块结构图

Fig. 2 C3PP module architecture

C3PP模块的结构使得它可以关注到全局特征，而不像传统卷积仅能关注固定大小的感受野，同时，它会根据每个像素点的重要性对特征进行加权，从而捕捉到微小特征之间的相关性，弥补了传统卷积在微小特征提取方面的不足。同时，改进后的网络更注重不同层级目标的空间关系，可以适应不同尺度和分布的特征，使得多尺度特征能够更好地进行融合。此外，PSA机制使得网络更加关注小目标自身的特征，抑制了背景信息的干扰，提高了小目标检测的准确性。

1.3 融合BiFPN结构

由于图像中往往存在着多尺度无人机目标，其特征分布在不同的网络层级中。一方面，YOLOv5的网络结构对浅层特征的关注不够，使得浅层特征没有被充分利用；另一方面，在网络传播过程中，池化等操作也会导致部分特征信息丢失，使得网络无法充分保留小目标的细节信

息，进而产生漏检与误检。对此，本文引入了BiFPN结构使网络更合理有效地利用不同层级的特征信息，有效地解决了上述问题。FPN^[18]自被提出以来就被广泛应用于神经网络的多尺度特征融合。近年来，研究学者们也提出了许多多尺度特征融合网络结构，如PAN结构^[19]和神经架构搜索特征金字塔^[20]等。这些工作都对不同层次的输入特征进行了融合，但却并未充分区分不同输入特征的贡献，且忽略了不同输入特征的分辨率差异。针对上述问题，Tan等^[21]提出了BiFPN结构。该网络引入了可学习的权重因子来表征不同输入特征的重要程度，并通过自顶向下和自底向上的多尺度特征融合，实现了网络结构的优化。图3展示了BiFPN的结构图。

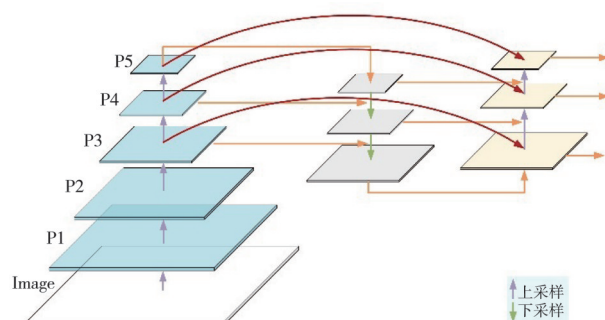


图3 BiFPN结构图

Fig. 3 BiFPN architecture

BiFPN通过跨连接解决了特征融合的问题，使底层特征更有效地传播到顶层，更好地保留目标的多尺度特征，充分利用了在小目标检测中尤为重要的浅层特征，避免了特征信息丢失，从而提升了模型的检测精度。

1.4 检测头的解耦

在目标检测中，分类任务主要关注对目标类别的准确分类，而回归任务主要关注目标位置和尺寸的精确预测，二者关注的准确性指标和优化目标不同，将它们强耦合在一起会导致相互干扰和冲突，进而影响模型的性能。而在YOLOv5的检测头中，分类任务和回归任务是共享权重的，这意味着它们共同参与了网络的前向传播和反向梯度传播过程，限制了分类和回归任务的独立优化能力。为了应对这一问题，YOLOX^[22]提出了解耦头的解决方案。该研究将检测头拆分为独立的分类头和回归头，使得分类任务和回归任务可以根据自身特点进行独立的优化和调整，从而提高模型的性能和灵活性。具体结构见图4。

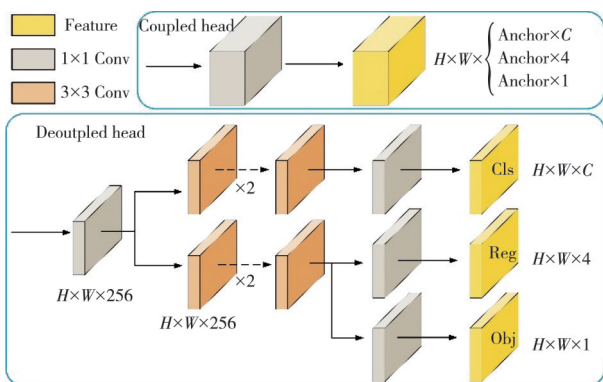


图4 耦合检测头与解耦检测头结构图

Fig. 4 Coupled and decoupled detection head architecture

检测头解耦后的YOLOv5模型使得目标的分类和定位分支可以分别优化而不会相互干扰,从而加速模型的收敛过程,特别是在处理小目标时,可以更快地调整网络参数以适应小目标的特征。此外,小目标的定位信息相对更加敏感,通过解耦合的检测头,可以更好地优化定位分支,从而提升小目标检测的定位准确性。

2 实验验证

实验使用自行构建的无人机小目标数据集LASD-D来评估模型的性能,并分别从数据集、训练设置、评价指3个部分对实验的前期准备工作进行了详细的介绍。

2.1 实验准备

2.1.1 数据集的建立

深度学习目标检测算法本质上是由数据驱动的,因此,数据集的质量直接影响网络性能。当数据集中样本的多样性与普适性不足时,往往很难训练出优秀的网络,因此,建立一个样本覆盖范围大、样本标注准确性高的数据集往往是搭建一个优秀模型的第一步。针对当前无人机小目标数据集中小目标占比少与场景不全面的问题,建立了低慢小无人机数据集LASD-D,为了确保数据集的多样性和高质量,数据集考虑了多种天气条件,如晴天、阴天、多云以及黄昏,这使得模型在各种天气条件下都能表现出良好的性能。此外,数据集还包含了天空、树木、近地面等多种背景场景,这确保了模型能够在包含复杂背景的检测中有着良好的表现。数据集制作过程中共拍摄了20段视频,视频的分辨率为 $1\,920 \times 1\,080$,以每秒2帧的频率进行截取,从中挑选出2\,300张图片。同时,从开源无人机小目标数据集tib-net^[23]中选择了700张质量达标的图像作为补充,最终构建了包含3\,000张图片的数据集。所有图像都经过了手动标注和复核,以确保标注的准确性,数据集中部分图片如图5所示。



图5 数据集的部分图片

Fig. 5 Partial images of the dataset

为了确保数据集适用于小目标检测模型的训练,对该数据集中的目标尺寸进行了统计分析。小目标根据定义不同一般可分为相对小目标和绝对小目标。绝对小目标是指像素点总数小于 32×32 的目标,相对小目标是指在原图中占比小于10%的目标。本文结合绝对小目标的定义与相对小目标的定义,以原图占比不大于0.2%的目标为小目标,对其分布情况进行描述。将目标的边界框(Bounding Box)面积视为目标的面积,对数据集中无人机目标的尺寸进行统计,结果见图6。

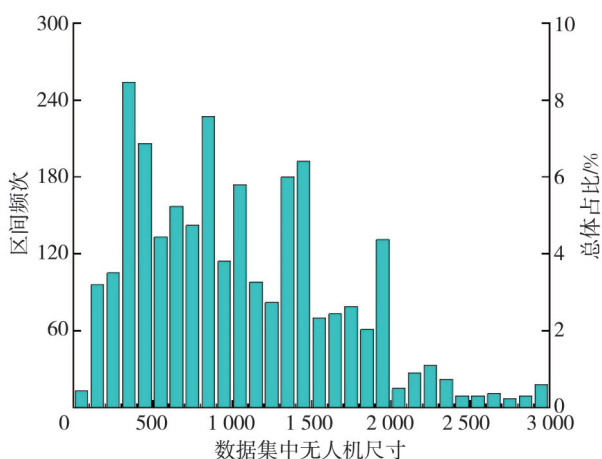


图6 数据集中无人机尺寸统计柱状图

Fig. 6 Bar chart of UAV size statistics in the dataset

统计结果表明,在LASD-D数据集中,绝对小目标占目标总数的57.3%,占比不超过图片尺寸0.2%的小目标占目标总数的91%。这说明本数据集包含了丰富多样的小尺寸无人机目标,适用于小目标检测算法的训练与性能验证,为研究小目标检测算法以及评估其性能提供重要的数据支持。

2.1.2 训练设置

1) 实验环境配置

本实验在Windows10操作系统中进行,使用显存为8GB的NVIDIA GeForce RTX 3070 Ti显卡进行训练推理,采用的深度学习框架为Pytorch 1.13.1, CUDA版本为11.7,所使用的编程语言为Python,其版本为3.9.16。

2) 网络训练设置

实验选用YOLOv5s作为基础网络,并采用官方预训练权重以加速网络训练。图像输入尺寸设置为 640×640 ,采用SGD优化器进行训练,初始学习率为0.01。为了调整学习率,采用了余弦退火衰减策略,并设置动量大小为0.937,权重衰减系数为0.0005。同时,将Batchsize设为8,并进行了100个Epoch的迭代训练。

在损失函数方面,采用了与YOLOv5相同的损失函数,包括坐标回归损失、类别损失以及置信度损失。坐标回归损失采用CIoU计算公式。相比于IoU,其考虑了重叠区域、中心距离以及宽高比等信息,使得模型对不同尺度和形状的物体具有更好的鲁棒性。类别损失采用类别交叉熵损失作为计算公式。其考虑了预测类别与真实类别之间的概率分布差异,提高了模型识别的准确性。置信度损失使用二进制交叉熵损失,其能够判断某个区域是否包含目标,有助于消除误检和漏检。总体而言,该损失函数使模型具有更好的鲁棒性,同时因采用类别交叉熵损失和二进制交叉熵损失,使得模型更容易收敛,减少了过拟合的风险。

2.1.3 评价指标的选取

本文采用精确率(Precision, P)、召回率(Recall, R)以及平均精度(Mean Average Precision, mAP)作为评估模型检测性能的指标。

1) 精确率和召回率

在模型性能评估时,使用精确率衡量预测正确的正样本数占总正样本数的比例,使用召回率衡量预测正确的正样本数占总标签数的比例。计算公式为

$$P = \frac{TP}{TP + FP} \cdot 100\%, \quad (1)$$

$$R = \frac{TP}{TP + FN} \cdot 100\%, \quad (2)$$

式中:TP(True Positives),FP(False Positives),FN(False Negatives)分别表示被正确检测的目标数量、误检为目标背景数量、误检为背景的目标数量。

2) 误检率与漏检率:

使用误检率(False Alarm, F)来衡量模型正确预测正样本纯度的能力,使用漏检率(Miss Rate, M)衡量模型正确预测负样本纯度的能力。其计算公式为

$$F = \frac{FP}{TN + FP} \times 100\%, \quad (3)$$

$$M = \frac{FN}{TP + FN} \times 100\%。 \quad (4)$$

3) 平均精度与平均精度均值

在模型性能评估时,平均精度均值(mAP, $P_{m\Delta}$)综合考虑了各个类别的平均精度(Average Precision, P_A)以及类别数量(Number of Class, N_C),其计算公式为

$$P_A = \int_0^1 P(R) dR, \quad (5)$$

$$P_{mA} = \frac{\sum P_A}{N_C} \quad (6)$$

2.2 实验结果与分析

2.2.1 综合对比实验

为了评估改进后模型的检测效果,将LASD-YOLOv5与当下主流的目标检测模型YOLOv3, CornerNet, YOLOv7, YOLOv8以及小目标检测算法QueryDet进行了性能对比实验,实验在LASD-D数据集上进行,实验结果见表1。

表1 模型性能对比实验结果

Tab. 1 Experimental results of model performance comparison

模型	P/%	R/%	$P_{mA}/\%$	延时/(帧·s ⁻¹)
YOLOv3	88.45	85.11	87.15	53.7
CornerNet	84.74	90.23	89.32	67.4
YOLOv5	96.90	93.33	95.42	6.34
QueryDet	89.32	92.15	90.73	6.12
YOLOv7	88.49	89.34	89.48	8.34
YOLOv8	93.37	87.65	91.22	6.45
Ours	97.32	96.99	98.29	6.71

由表1可知,改进后模型在LASD-D数据集上的mAP指标上比YOLOv5提高了2.87%,召回率提高了6.66%,精确率提高了0.43%,表明改进显著提升了模型性能。改进后模型在精确率、召回率以及mAP指标上均优于YOLOv3, CornerNet, YOLOv7, YOLOv8与QueryDet,表明LASD-YOLOv5在无人机小目标检测方面的性能强于主流模型。在检测速度上,改进后模型的检测速度为6.71帧/ms,略慢于原模型的6.34帧/ms与QueryDet模型的6.12帧/ms,完全满足实时性要求。

本文对改进前后模型的复杂度进行了对比实验,对比结果见表2。

表2 改进前后模型复杂度对比

Tab. 2 Comparison of model complexity before and after improvement

模型	参数量/M	FLOPS/G
YOLOv5	7.01	15.80
Ours	13.99	56.66

由表2可知,LASD-YOLOv5的参数量为13.99 M,相比于原模型的7.01 M的参数量略有增加。FLOPS为56.6 G,也高于原模型15.80 G。结合表1数据可知,模型复杂度的增加对其检测速度的影响较小,模型仍能很好地实现实时检测。

2.2.2 消融实验

为了验证每个组件在模型中的贡献,本文通过逐个添加组件的方式进行了消融实验,得出的

实验结果见表3。

表3 LASD-YOLOv5模型消融实验结果

Tab. 3 Experimental results of the ablation study for the LASD-YOLOv5 model.

YOLOv5	C3PP	BiFPN	解耦头	$P_{mA}/\%$
✓				95.42
✓	✓			97.13
✓		✓		95.53
✓			✓	96.23
✓	✓	✓	✓	98.29

注: ✓表示模型中包含该组件

从表3中可以得知,本文设计的组件都在一定程度上提升了模型对于无人机小目标的检测性能,特别是C3PP模块的引入对mAP指标的提升效果最为显著,提高了1.71%,表明C3PP模块能够有效地捕捉目标的微小特征,从而提升检测效果。此外,BiFPN结构的加入使模型mAP指标有0.11%的提升,增强了模型对于不同尺度的目标的适应性。最后,检测头的解耦使mAP指标提升了0.81%。表明解耦检测头可以消除分类任务和回归任务之间的冲突,从而提高目标检测的精度。这些改进使得LASD-YOLOv5模型的mAP指标相比于YOLOv5提升了2.87%,达到了98.29%,表明改进后的模型在无人机小目标检测方面的性能出众。

2.2.3 可视化分析

为了更直观地体现模型的检测效果,本文进行了可视化分析,从单帧检测与视频序列检测两方面对优化后的网络模型进行性能验证。

1) 单帧检测效果对比

图7展示了LASD-D数据集上的部分检测结果。

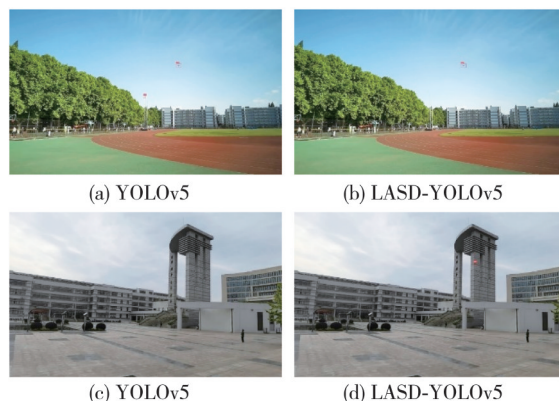


图7 单帧可视化检测结果对比图

Fig. 7 Visual comparison of detection results

在图7(a)中,YOLOv5模型错把路灯顶端识别为无人机目标,在图7(c)中,YOLOv5模型忽略了与无人机具有相似背景的目标,而图7(b)、

图 7(d)的 LASD-YOLOv5 均准确识别出了无人机目标。可以清晰地看出,相比于原算法,改进后的算法在处理复杂背景下的目标检测与特征稀疏的极小目标检测时的漏检、误检现象有所改善,在无人机小目标检测上表现出了十分优异的检测性能。

2) 视频序列检测效果

为了更全面地确认优化后模型的性能,使用 1 655 帧包含无人机小目标的视频序列进行了检测,并计算改进前后模型在该视频中的检测精准率、误检率及漏检率,计算结果见表 4。

表 4 视频序列综合识别效果

Tab. 4 Comprehensive recognition of video sequences

模型	P/%	F/%	M/%
YOLOv5	93.1	3.7	3.9
LASD-YOLOv5	95.3	3.5	1.8

由表 4 可知,改进后模型的漏检率降低了 2.1%,误检现象降低了 0.2%,误检与漏检问题得到了有效改善。改进后的模型对于无人机小目标的检测精确率可达 94.7%,可以满足实际场景中的任务需求。

图 8 展示了视频序列中的部分检测结果,其中图 8(a)、图 8(b)正确检测出了目标,图 8(c)将右下角的斑点误检为目标,图 8(d)漏检了目标,结合更多的检测结果进行分析,发现改进后模型的漏检、误检现象主要集中于复杂背景下的极小目标与移动的极小目标上。

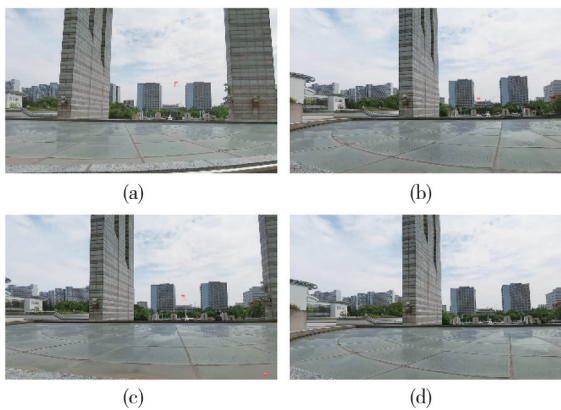


图 8 部分视频序列检测结果

Fig. 8 Partial video sequence detection results

3 结论与展望

为了解决无人机小目标检测中的漏检与误检问题,本文提出了一种无人机小目标检测算法 LASD-YOLOv5,并贡献了一个适用于该领域的多场景低慢小无人机目标数据集 LASD-D。首先,引入 C3PP

模块,增强了网络的特征提取能力;接着,使用 BiFPN 结构替换 PAN 结构,强化了对底层特征的利用;最后,将原本耦合的检测头替换为解耦的头部,提高了模型的收敛速度和检测性能。在 LASD-D 数据集上分别进行了消融实验与综合对比试验,前者的实验结果验证了上述改进的有效性。后者的实验结果表明了改进后算法优于主流算法,能够有效地检测无人机小目标。

在未来的研究中,将继续努力改进模型的网络结构,在不影响模型性能的前提下尽可能减小模型大小。同时,还将进一步扩充数据集,使其包含更多种类的无人机目标和更全面的场景信息,为相关研究的进行提供更好的研究基础。

参考文献:

- [1] HADI H J, CAO Y, NISA K U, et al. A comprehensive survey on security, privacy issues and emerging defence technologies for UAVs [J]. *Journal of Network and Computer Applications*, 2023, 213: 103607.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//*IEEE Conference on Computer Vision & Pattern Recognition*, 2016: 779-788.
- [4] REDMON J, FARHADI A. YOLOv3: an incremental improvement [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 6517-6525.
- [5] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2020. <http://arxiv.org/abs/2004.10934v1>.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [C]//*European Conference on Computer Vision*. Cham: Springer, 2016: 21-37.
- [7] LAW H, DENG J. CornerNet: detecting objects as paired keypoints [J]. *International Journal of Computer Vision*, 2020, 128(3): 642-656.
- [8] YANG C, HUANG Z, WANG N. QueryDet: Cascaded sparse query for accelerating high-resolution small object detection [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 13668-13677.

- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [10] GIRSHICK R. Fast R-CNN[C]// 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 1440-1448.
- [11] 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.
- [12] 王建楠, 吕胜涛, 牛健. 基于改进YOLOv5的无人机检测方法[J]. 光学与光电技术, 2022, 20(5): 48-56.
WANG Jiannan, LÜ Shengtao, NIU Jian. Drone detection method based on improved YOLOv5 [J]. Optics & Optoelectronic Technology, 2022, 20(5): 48-56. (in Chinese)
- [13] LUO K, ZHU G, LI Y, et al. Low altitude and low speed uav identification based on hybrid model [C]// Proceedings of the 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019), 2019: 76-87.
- [14] 包文歧, 谢立强, 徐才华, 等. 基于YOLOv5的微小无人机实时探测方法[J]. 兵器装备工程学报, 2022, 43(5): 232-237.
BAO Wenqi, XIE Liqiang, XU Caihua, et al. Real time detection method of micro UAV based on YOLOv5[J]. Journal of Ordnance Equipment Engineering, 2022, 43(5): 232-237. (in Chinese)
- [15] WANG M, YANG W, WANG L, et al. FE-YOLOv5: Feature enhancement network based on YOLOv5 for small object detection [J]. Journal of Visual Communication and Image Representation, 2023, 90: 103752.
- [16] ZHU X, LYU S, WANG X, et al. TPH-YOLOv5: Improved yolov5 based on transformer prediction head for object detection on drone-captured scenarios [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2778-2788.
- [17] LIU H, LIU F, FAN X, et al. Polarized self-attention: towards high-quality pixel-wise mapping [J]. Neurocomputing, 2022, 506: 158-167.
- [18] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 936-944.
- [19] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [20] GHIASI G, LIN T Y, LE Q V, et al. NAS-FPN: learning scalable feature pyramid architecture for object detection [C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 7029-7038.
- [21] TAN M, PANG R, LE Q V. EfficientDet: scalable and efficient object detection [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 10778-10787.
- [22] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021 [EB/OL]. <http://arxiv.org/abs/2017.08430v2>.
- [23] SUN H, YANG J, SHEN J, et al. TIB-Net: drone detection network with tiny iterative backbone [J]. IEEE Access, 2020, 8: 130697-130707.