姜浩, 崔璨, 常迪. 基于 YOLOv5s 的无人机航拍图像小目标检测算法研究[J]. 智能计算机与应用,2026,16(8):68-74.
DOI:10.20169/j.issn.2095-2163.25010801

基于 YOLOv5s 的无人机航拍图像小目标检测算法研究

姜 浩[1], 崔 璨[1], 常 迪[2]
(1 沈阳航空航天大学 民用航空学院, 沈阳 110136; 2 湖北国际物流机场有限公司, 湖北 鄂州 436055)

摘 要: 随着无人机的飞速发展,其应用场景越来越广泛、复杂。但在无人机执行任务过程中,极易与飞鸟、电线等小目标发生碰撞,对飞行安全造成威胁。提高无人机对小目标的检测能力,确保无人机在各类环境中都能稳定、安全地运行具有重要意义。针对无人机小目标检测中的漏检和误检的问题,本文以 YOLOv5s 为基准模型,对网络中的 Neck 端和 Head 端进行改进。首先在 Neck 端引入 BiFPN 网络和多层次特征融合模块 SDI,改善小目标语义丢失和信息传递不均衡问题。其次在 Head 端使用基于动态注意力机制的 DyHead 动态检测头,并引入可形变卷积核 DCNV4,提升特征识别能力。最后将损失函数更换为 EIoU,加快模型收敛速度。将改进的 YOLOv5s-BSDE 模型在 VisDrone2019 测试集进行验证,结果表明, 其精确率(P)、召回率(R)、mAP@0.5 和 mAP@0.50∶0.95 分别提高了 3.52%、5.02%、5.29%、4.61%,可显著提高无人机在执行任务中对小目标的检测性能。

关键词: YOLOv5; 小目标检测; BiFPN; Dyhead 动态检测头; EIoU 损失函数
中图分类号: V278 **文献标志码**: A **文章编号**: 2095-2163(2026)08-0068-07

Research on small target detection algorithm for UAV aerial images based on YOLOv5s

JIANG Hao[1], CUI Can[1], CHANG Di[2]
(1 College of Civil Aviation, Shenyang Aerospace University, Shenyang 110136, China;
2 Hubei International Logistics Airport Co., Ltd., Ezhou 436055, Hubei, China)

Abstract: With the rapid development of UAVs, their application scenarios have become increasingly extensive and complex. However, during mission execution, UAVs are prone to collisions with small targets such as birds and power lines, which poses a threat to flight safety. It is of great significance to improve the detection ability of UAVs for small targets and ensure their stable and safe operation in various environments. Aiming at the phenomena of missed detection and false detection in UAV small target detection, this paper takes YOLOv5s as the benchmark model, improvements to the Neck and Head ends of the network. Firstly, the BiFPN network and the multi-level feature fusion module SDI are introduced at the Neck end to address the issues of semantic loss and uneven information transmission for small targets. Secondly, the DyHead detection head based on dynamic attention mechanism is used at the Head end, and the deformable convolutional kernel DCNV4 is introduced to improve feature recognition ability. Finally, the loss function is replaced with EIoU to speed up model convergence. The improved YOLOv5s-BSDE model is verified on the VisDrone2019 test set, and the results show that its Precision(P), Recall(G), mAP@0.5, and mAP@0.50∶0.95 have been improved by 3.52%, 5.02%, 5.29%, and 4.61% respectively, significantly enhancing the UAV′s small-target detection performance during mission execution.

Key words: YOLOv5; small object detection; BiFPN; DyHead Dynamic Detection Head; EIoU loss function

0 引 言

无人机凭借高效性、便携性以及安全性广泛应用于各种场景,但随着应用场景的不断拓展,小目标检测技术的需求也日益增加[1]。尤其是在复杂飞行环境中,无人机需要精准识别并定位微小目标,以确保飞行安全。通过优化小目标识别系统,提升无人机的小目标检测能力,可以有效避免无人机与障

基金项目: 辽宁省教育厅项目-基本科研业务费专项(纵 20240189)。
作者简介: 姜 浩(1999—),男,硕士。主要研究方向:深度学习,小目标检测。
通信作者: 崔 璨(1988—),女,博士,讲师。主要研究方向:图像处理,目标检测。Email:hvcc05@163.com。
收稿日期: 2025-01-08

碍物发生碰撞，降低事故风险。在此背景下，小目标检测技术作为无人机智能化、自主化发展的重要支撑，对保障无人机安全运行具有重要意义。

Yan Hui 等[2]针对遥感场景下小目标检测的难点，提出了一种基于 YOLOv5 的增强型 SEB-YOLO 算法。该方法首先对原网络结构进行重建，引入非步长卷积层模块（SPD-Conv）[3]以尽可能减少网络下采样导致的特征损失，同时在主干网络末端设计了一个结合注意力机制的池化模块以提高模型对不同分辨率图像的适应性。然后引入 BiFPN[4]，并采用双线性插值上采样操作解决原始颈部网络特征融合不足的问题。最后，使用解耦头替换模型中的检测头，进一步提高模型的检测性能。刘辉等[5]针对小目标漏检问题，提出一种融合注意力机制和多尺度上下文信息的改进 YOLOv5 目标检测算法，算法先构建了多尺度空洞可分离卷积模块（MDSCM）[6]，提升了模型对于不同尺度目标的自适应性。然后在通道注意力模块中引入位置注意力，增强感受野的特征获取能力，最后使用 Soft-NMS[7]代替原模型所使用的非极大值抑制（NMS），改进后的算法对无人机图像中小目标的检测效果有较好的性能提升。Xu 等[8]针对焊缝缺陷小目标提出了改进后的 YOLOv5-Improvement 算法：首先使用 CA 注意力机制[9]与原网络结构进行融合，提高其检测焊缝 X 射线胶片上小目标缺陷的能力，然后引入 SIoU 损失函数来改善模型对空间低灵敏度信息的捕获能力，最后通过 FReLU 激活函数增强全局优化能力来提高识别的准确性和速度。

尽管上述检测算法性能已经有了较大的提升，但对于小目标的检测效果依旧不够理想。随着网络层数的不断加深，模型对全局上下文和复杂背景的感知能力得到了提升，但小目标的特征信息在网络传播过程中容易被稀释或丢失，使得检测头难以准确识别到小目标的特征信息。针对上述问题，本文以 YOLOv5 作为基础进行研究，主要的研究内容如下：

（1）针对小目标特征信息在传递过程中丢失的问题，引入 BiFPN 网络结构优化小目标特征信息的传递，同时使用多层次特征融合模块 SDI 改善 BiFPN 网络中的双向连接可能会导致信息传递不均衡的问题。

（2）针对检测头难以准确识别到小目标的特征信息问题，引入动态检测头（DyHead）提升网络检测头对于特征信息的识别和提取能力，同时使用可形变卷积核 DCNV4 替换原检测头的卷积核，能够更好地捕捉网络中语义的细微变化。

1　算法介绍及改进

1.1　YOLOv5 目标检测算法

YOLOv5[10]，作为一阶段检测算法，其显著特点在于直接预测目标的边界框，无需先生成候选锚框。这种设计使其在目标检测任务中拥有更快的检测速度，因此被广泛应用于各种场景下的目标检测任务。在 YOLOv5 系列算法中，根据不同的特征图宽度和网络深度，模型规模从大到小可分为 YOLOv5x、YOLOv5l、YOLOv5m、YOLOv5s 四种模型[11]。本文为提高检测精度的同时，兼顾无人机的架构搭载情况，选择轻量化更好的 YOLOv5s 网络模型作为网络主体，该模型的具体结构如图 1 所示。

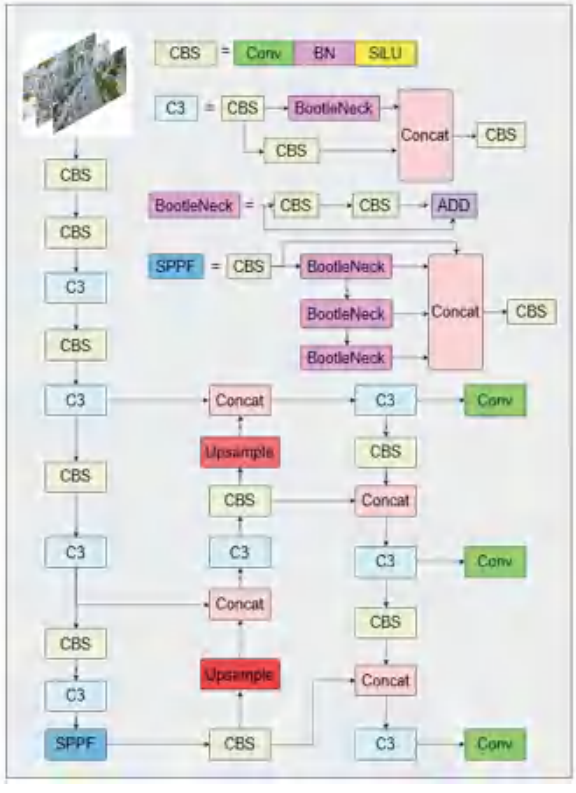


图 1　YOLOv5 网络结构模型图

Fig. 1　Diagram of YOLOv5 network structure model

1.2　引入改进后的 BiFPN-SDI 网络结构

在网络模型的设计中，随着网络层数的加深，网络能够捕获到更加丰富的语义信息，但同时也会造成部分浅层特征细节丢失。该问题对小目标检测任务尤为不利。而仅依赖浅层细节特征进行检测，又会因缺乏足够的语义信息而降低整体检测性能。为克服这一矛盾，通常采用深、浅层多尺度特征融合的方案。如图 2 所示，传统的 FPN 只有自上向下的单向特征信息传递路径[12]。YOLOv5s 原生模型中采

用的 PANet[13]，在 FPN[14] 基础上新增了自底向上的融合支路，保留了更多浅层细节特征。BiFPN 是一种加权的双向特征金字塔网络，其在 PANet 基础上增加了跳跃连接，将主干网络中所提取的特征直接与自底向上路径中对应尺度的特征相融合，增强了浅层特征信息与深层特征信息的融合。

BiFPN 的加权融合机制能够让网络在有效融合不同尺度特征的同时，更加注重信息量丰富的小目标特征。本文在引入网络架构的时候做了一定的修改，将原本的 P2 输出层改为单个卷积核，同时在 P4、P5 层添加了两个跳跃链接，并将引入方式从相加(Add)模式改为拼接(Concat)模式。但是 BiFPN 网络架构在信息传递过程中的双向连接可能会导致信息传递不均衡的问题，因此，本文引入多层次特征融合模块 SDI[15] 来改善这一问题。SDI 具有多尺度信息融合、自适应特征交互的特性，能更好适配小目标，SDI 模块结构如图 3 所示。

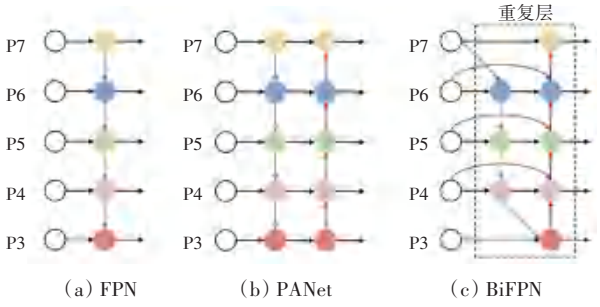


图 2 多尺度特征融合网络模型结构

Fig. 2 Structure of multi-scale feature fusion network model

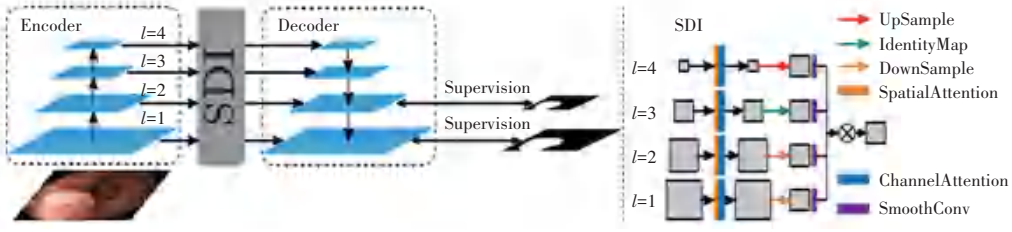


图 3 SDI 模块结构

Fig. 3 Structure of the SDI module

1.3 引入 DyHead-DCNV4 动态注意力机制检测头

DyHead[16] 是一种创新的动态检测头设计，其独特之处在于将目标检测头与注意力机制进行了动态集成。DyHead 在尺度感知的特征层次、空间感知的空间位置以及任务感知的输出通道等多个维度上，将注意力机制连贯地组合起来，实现了对检测场景的自适应调整。图 4 为 DyHead 动态检测头的模型结构图，其中 π_L、π_S、π_C 分别代表尺度感知注意力(Level-wise)、空间感知注意力(Spatial-wise)和任务感知注意力(Channel-wise)，使模型在不增加计算量的情况下，显著提升检测头的特征识别能力。

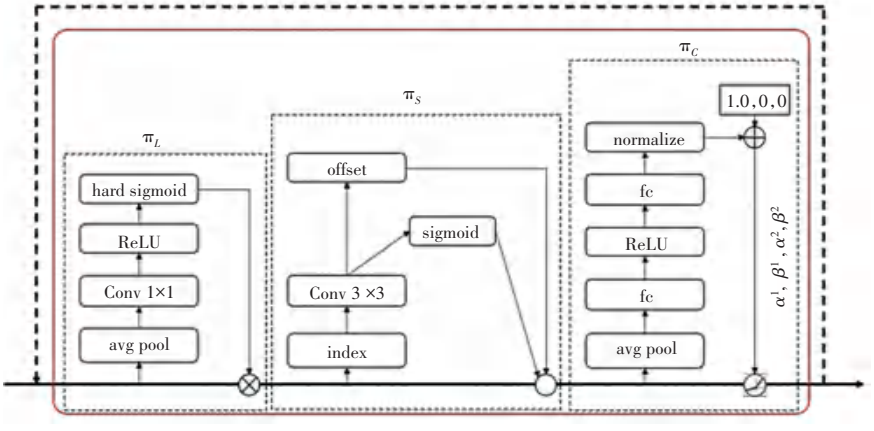


图 4 DyHead 动态检测头结构

Fig. 4 Structure of the DyHead dynamic detection head

在使用 Dyhead 动态检测头进行小目标检测时，由于原 YOLOv5 网络结构每一层的特征输出都是不同的，原结构中的 1 * 1、3 * 3 卷积核难以处理大量的参数计算，从而降低了 DyHead 检测头的特征识别能力。为优化上述问题，本文使用了可形变卷积核 DCNv4[17]，DCNv4 通过去除空间聚合中的 Softmax[18] 归一化，增强了其动态属性和表达力，同时通过优化内存访问以最小化冗余操作，显著提高

了处理速度和收敛速度。引入改进后的 DyHead 动态检测头，能够更好地捕捉网络中语义的细微变化，还可以减少大量的参数计算。

1.4　优化损失函数

YOLOv5 模型选择 CloU 作为损失函数，通过多维度考虑预测框与真实框之间的差异，但在部分场景下，尤其是在目标边界框不重叠或者两个框尺寸有较大差异时，效果并不理想。尤其当预测框与真实边界框的长宽比相同时，此时的长宽比的惩罚恒为 0，不符合实际检测情况。本文通过引入 EIoU[19] 损失函数来减少上述情况的发生，EIoU 是在 CIoU 的惩罚项基础上将预测框与真实框的纵横比的影响因子拆开，分别计算预测框与真实框的长和宽，避免上述情况的发生，同时加快网络模型收敛速度。EIoU 损失函数如下式：

$$L_{\mathrm{EIoU}} = L_{\mathrm{IoU}} + L_{\mathrm{dis}} + L_{\mathrm{asp}} = 1 - \mathrm{IoU} + \frac{\rho^2(b, b^{\mathrm{gt}})}{c^2} + \frac{\rho^2(w, w^{\mathrm{gt}})}{C_w^2} + \frac{\rho^2(h, h^{\mathrm{gt}})}{C_h^2} \tag{1}$$

改进后的 YOLOv5s-BSDE 网络结构如图 5 所示，与原 YOLOv5s 网络相比，改进后的模型更注重网络中的小目标特征语义信息传输能力以及识别能力，有效地提高了模型对小目标的检测性能，并在小目标公开数据集 VisDrone2019 上取得了较为理想的效果。

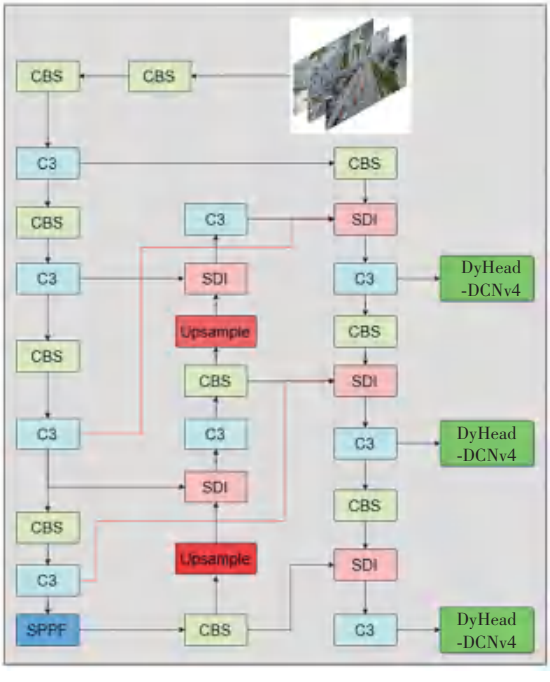


图 5　改进后的 YOLOv5s-BSDE 网络结构

Fig. 5　Structure of the improved YOLOv5s-BSDE network

2　实验结果与分析

2.1　实验环境

实验环境配置如下：操作系统采用 Windows 11，CPU 为 Intel（R）Core（TM） i5-12600KF，主频 3.7 GHz，GPU 为 NVIDIA GeForce RTX 4060，8G 显存，开发环境为 Python 3.8，深度学习框架选用 PyTorch 2.3.0，CUDA 版本为 12.1，CUDNN 版本为 8.9.4。

2.2　数据集选用以及数据预处理

为确保改进模型的泛化能力，本文选择了小目标公开数据集 VisDrone2019 作为实验数据集[20]。该数据集由中国天津大学的 AISKYEYE 团队提供，使用多种型号的无人机在中国的 14 个城市与乡村之间进行拍摄，在充分考虑了场景密度、天气状况等因素的同时，增加了数据的多样性与代表性。该数据集共包含 10 209 张静态图片，按照比例划分成了 6 471 张训练集、584 张验证集、1 610 张测试集和 1 580 张比赛专用测试集。该数据集对 10 种不同类别的物体进行标注，共有 200 多万个标签。

为提升网络模型的训练效果，防止训练过程中出现过拟合及部分异常的数据出现，将 VisDrone2019 数据集中部分存在拍摄模糊、曝光问题的异常图片筛除掉，并使用随机旋转、裁剪以及 Mosaic 数据增强[21]，将训练集扩充到 8 122 张，图 6 是扩充后的数据图片展示。



图 6　扩充后的数据图片

Fig. 6　Augmented data images

2.3　实验指标

YOLO 系列目标检测模型的评测通常是通过精确率（Precision，P）、召回率（Recall，R）、平均精度均值（mean Average Precision，mAP）等指标来衡量

模型的性能。首先,将预测结果分为以下 4 类:

TP:指样本实际为正例,预测结果为正例,即正确地预测为正例。

FP:指样本实际为负例,预测结果为正例,即错误地预测为正例。

FN:指样本实际为正例,预测结果为负例,即错误地预测为负例。

TN:指样本实际为负例,预测结果为负例,即正确地预测为负例。

(1)精确率(Precision, P):表示预测出的真实正例占所有预测为正例的比例。用以反映在预测结果中,正例被预测正确的概率。计算公式如下:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

(2)召回率(Recall, R):表示实验结果中预测正确的正例数量与实际正例数量的比值。用以衡量模型检测中的漏检情况,召回率越低,漏检率越高。计算公式如下:

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

(3)平均精度均值(mAP):AP 表示某一类物体的检测精度,AP 的值取自 PR 曲线与坐标轴围成的面积,计算公式如公式(4)所示。mAP 是体现目标检测任务中模型性能的重要指标,由于目标检测任务中通常会存在多个类别的检测目标,模型会计算每个类别的 AP 值,将全部类别的 AP 值相加后求均值,即可获得 mAP,公式如下:

$$\mathrm{AP} = \int_0^1 P(R)\,\mathrm{d}R \tag{4}$$

$$\mathrm{mAP} = \frac{\sum_{i=1}^{k} \mathrm{AP}_i}{k} \tag{5}$$

2.4 实验结果

为确保实验的可靠性与准确性,本文实验均采用相同的软硬件环境以及相同的超参数设置。其中部分超参数设置如下:输入图像的分辨率为 640×640,训练使用预训练权重为 YOLOv5s,训练时的训练轮次(epoch)为 200,批次大小(batch-size)为 8,其余超参数均使用默认参数。图 7 为原 YOLOv5s 模型与改进后的 YOLOv5s-BSDE 模型的 PR 曲线对比图, PR 曲线显示了在不同阈值下模型的精确率和召回率之间的关系。PR 曲线越靠近右上角表示模型的性能越好,从曲线图中可观察到改进后的模型各个类别的曲线均比原模型靠近右上角,且图中右上角各个类别的检测精度也均有提升。图 8 是模型对不同类别物体的检测图片,其中图片左侧是原模型,右侧是改进后的模型,图中红框标注的部分直观体现了改进后的模型在提升小目标检测能力方面所取得的显著成效。

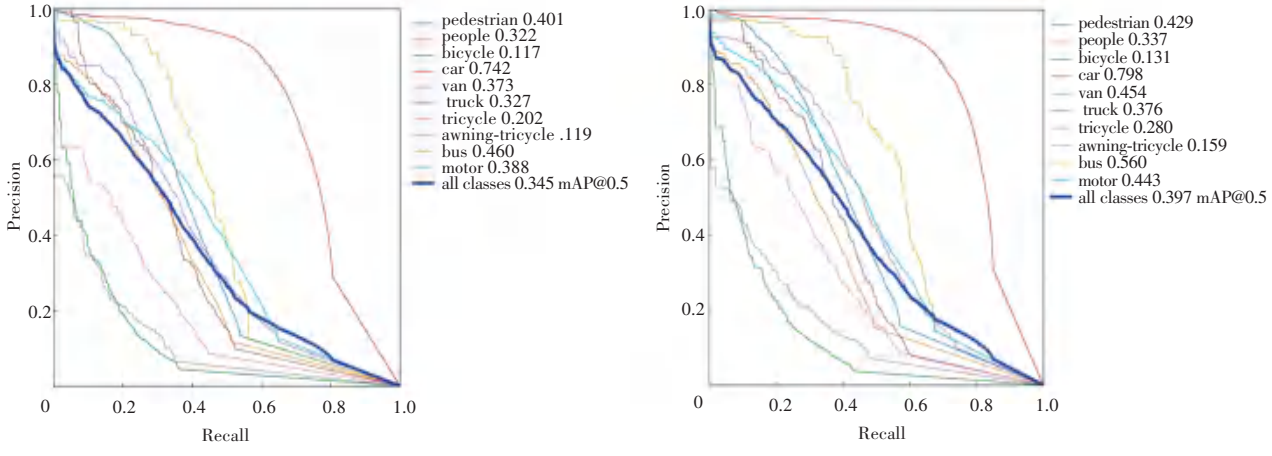


图 7 YOLOv5s(左)与 YOLOv5s-BSDE(右)模型 P-R 曲线图

Fig. 7 P-R curves of YOLOv5s (left) and YOLOv5s-BSDE (right) models

2.4.1 消融实验

为验证改进后的 BiFPN-SDI 网络、DyHead-DCNv4 动态检测头以及 EIoU 损失函数给模型带来的性能增益,本文设计了 4 组消融实验。其中将基础的 YOLOv5s 网络模型记录为 YOLOv5s,加入 DyHead-DCNv4 动态检测头的网络模型记录为 YOLOv5s-1,将加入 DyHead-DCNv4 动态检测头以及 BiFPN-SDI 网络记录为 YOLOv5s-2,将加入了 DyHead-DCNv4 动态检测头、BiFPN-SDI 网络以及 EIoU 损失函数最终的改进模型记录为 YOLOv5s-3。实验结果见表 1。



图 8　YOLOv5s(左)与 YOLOv5s-BSDE(右)模型检测效果图

Fig. 8　Detection effect comparison between YOLOv5s (left) and YOLOv5s-BSDE (right) models

表 1　不同改进模块的性能对比

Table 1　Performance comparison of different improved modules

	P	*R*	mAP@0.5	mAP@0.5∶0.9
YOLOv5s	0.465 1	0.339 4	0.343 4	0.188 7
YOLOv5s-1	0.489 2	0.373 2	0.373 1	0.220 1
YOLOv5s-2	0.492 5	0.378 8	0.383 1	0.232 1
YOLOv5s-3	0.500 3	0.389 6	0.396 3	0.234 8

从实验结果中可以观察到，改进后的模块在 Visdrone2019 数据集上都有一定的性能提升。从实验数据当中，可以发现 DyHead-DCNv4 动态检测头提升较大，准确率和召回率分别提升了 2.41%和 3.38%，mAP@0.5 和 mAP@0.50∶0.95 分别提升了 2.97%和 3.14%。后续实验每添加一个模块，其性能都有一定程度的提升，有效地证明了改进方法的可行性。

2.4.2　对比实验

在保持原实验环境不变的前提下，VisDrone2019 公开数据集上分别复现了 TPH-YOLOv5、YOLOv6、YOLOv7 和 YOLOX 四个模型的轻量化版本以及 SEB-YOLO、AC-YOLO、YOLO-Improvement，最终的实验结果见表 2。由表中数据可知，虽然改进后 YOLOv5s-BSDE 的召回率要低于 YOLOv6 和 YOLOX，但是 mAP 和准确度都是要高于其他的 YOLO 系列算法，可以验证改进后的模型在小目标检测方面对比其它模型更具优势。

表 2　不同算法的性能对比

Table 2　Performance comparison of different algorithms

	P	*R*	mAP@0.5	mAP@0.5∶0.9
TPH-YOLOv5	0.460 1	0.361 4	0.357 2	0.191 4
YOLOv6	0.472 1	0.401 2	0.374 2	0.229 4
YOLOv7	0.441 5	0.377 6	0.346 9	0.185 3
YOLOX	0.498 5	0.396 5	0.392 2	0.221 7
YOLOv5s-BSDE	0.500 3	0.389 6	0.396 3	0.234 8
SEB-YOLO	0.486 3	0.374 3	0.383 2	0.231 6
AC-YOLO	0.482 2	0.375 6	0.393 6	0.225 7
YOLO-Improvement	0.472 5	0.382 1	0.389 4	0.220 4

3　结束语

本文针对无人机在执行任务过程中对小目标检测能力低，误检及漏检率高等问题，提出一种基于 YOLOv5s 网络模型改进的算法。首先在 YOLOv5 的 Neck 端引入了 BiFPN 网络结构，改善网络中特征信

息丢失情况,同时引入了具备多尺度信息融合、自适应特征交互的多层次特征融合模块 SDI,改善 BiFPN 网络中的双向连接可能会导致信息传递不均衡的问题。然后在 Head 端引入了基于动态注意力机制的 DyHead 检测头,并将模块中的卷积核替换为可形变卷积核 DCNv4,提高网络模型对融合后的特征信息识别能力。最后将损失函数更换为 EIoU 损失函数,解决原本的 CIoU 损失函数在目标边界框不重叠或者有较大的大小差异时带来的负面效果。实验结果表明,改进后的算法在各项指标上均优于原网络模型和其他主流算法,显著提高了无人机在执行任务中对小目标的检测能力,保证了安全性。

参考文献

[1] 刘芳, 吴志威, 杨安喆, 等. 基于多尺度特征融合的自适应无人机目标检测[J]. 光学学报, 2020, 40(10):133-142.

[2] Yan H , You S J, Hu X, et al. SEB-YOLO: An improved YOLOv5 model for remote sensing small target detection[J]. Sensors, 2024, 24(7): 2193.

[3] Gu Z, Zhu K, You S. YOLO-SSFS: A method combining spd-conv/stdl/im-fpn/siou for outdoor small target vehicle detection[J]. Electronics, 2023, 12(18): 3744.

[4] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 10781-10790.

[5] 刘辉, 张琳玉, 王复港, 等. 基于注意力机制和上下文信息的目标检测算法[J]. 计算机应用, 2023, 43(5):1557-1564.

[6] Zhao T, Mao Z, Karniadakis G E. Multi-domain spectral collocation method for variable-order nonlinear fractional differential equations [J]. Computer Methods in Applied Mechanics and Engineering, 2019, 348: 377-395.

[7] Bodla N, Singh B, Chellappa R, et al. Soft-NMS-improving object detection with one line of code[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway, NJ:IEEE, 2017: 5561-5569.

[8] Xu L, Dong S, Wei H, et al. Defect signal intelligent recognition of weld radiographs based on YOLO V5-improvement[J]. Journal of Manufacturing Processes, 2023, 99: 373-381.

[9] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]// Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 13713-13722.

[10] Jocher G, Chaurasia A, Stoken A, et al. Ultralytics/YOLOv5: v7.0-YOLOv5 SOTA Realtime Instance Segmentation[SW/OL]. Zenodo. (2022-11-23)[2026-08-01]. http://doi.org/10.5281/zenodo.7347926.

[11] Yang R, Li W, Shang X, et al. KPE-YOLOv5: An improved small target detection algorithm based on YOLOv5 [J]. Electronics, 2023, 12(4): 817.

[12] 王宏志, 宋明轩, 程超, 等. 基于改进 YOLOv5 算法的道路目标检测方法[J/OL]. 吉林大学学报(工学版), 2024, 54(9): 2658-2667. http://kns.cnki.net/kcms/detail/22.1341.T.20230315.1018.002.html.

[13] Cheng Q Q, Li X, Zhu B, et al. Drone detection methodbased on MobileViT and CA-PANet[J]. Electronics, 2023, 12(1):1-8.

[14] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 2117-2125.

[15] Peng Y, Sonka M, Chen D Z. U-Net v2: Rethinking the skip connections of U-Net for medical image segmentation[J]. arXiv preprint arXiv, 2311.17791, 2023.

[16] Dai X, Chen Y, Xiao B, et al. Dynamic head: Unifying object detection heads with attentions[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2021: 7373-7382.

[17] Xiong Y, Li Z, Chen Y, et al. Efficient deformable convnets: Rethinking dynamic and sparse operator for vision applications [C]// Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition . Piscataway, NJ: IEEE, 2024: 5652-5661.

[18] Jang E, Gu S, Poole B. Categorical reparameterization with gumbel-softmax[J]. arXiv preprint arXiv, 1611.01144, 2016.

[19] Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.

[20] Du D, Zhu P, Wen L, et al. VisDrone-DET 2019: The vision meets drone object detection in image challenge results[C]// Proceedings of ICCV Visdrone Workshop. Piscataway, NJ:IEEE, 2019:213-226.

[21] Hao W, Song Zhili, Wang Hao. Improved mosaic: Algorithms for more complex images[J]. Journal of Physics: Conference Series,2020,1684(1):012094. DOI:10.1088/1742-6596/1684/1/012094.