

基于IDOU-YOLO的红外图像无人机目标检测算法

陈海永 刘登斌 晏行伟

Infrared image UAV target detection algorithm based on IDOU-YOLO

CHEN Haiyong, LIU Dengbin, YAN Xingwei

引用本文:

陈海永, 刘登斌, 晏行伟. 基于IDOU-YOLO的红外图像无人机目标检测算法[J]. 应用光学, 2024, 45(4): 723-731. DOI: 10.5768/JAO202445.0402001

CHEN Haiyong, LIU Dengbin, YAN Xingwei. Infrared image UAV target detection algorithm based on IDOU-YOLO[J]. Journal of Applied Optics, 2024, 45(4): 723-731. DOI: 10.5768/JAO202445.0402001

在线阅读 View online: <https://doi.org/10.5768/JAO202445.0402001>

您可能感兴趣的其他文章

Articles you may be interested in

无人机对地目标自动检测与跟踪技术

Automatic target detecting and tracking technology based on UAV ground target images

应用光学. 2020, 41(6): 1153-1160 <https://doi.org/10.5768/JAO202041.0601003>

复杂背景下红外图像弱小目标检测

Dim-small targets detection of infrared images in complex backgrounds

应用光学. 2021, 42(4): 643-650 <https://doi.org/10.5768/JAO202142.0402002>

多尺度融合超分辨率算法在无人机探测中的应用

Application of multi-scale fusion super-resolution algorithm in UAV detection

应用光学. 2021, 42(3): 462-473 <https://doi.org/10.5768/JAO202142.0302003>

基于双路多尺度金字塔池化模型的显著目标检测算法

Salient target detection algorithm based on dual-channel multi-scale pyramid pooling model

应用光学. 2021, 42(6): 1056-1061 <https://doi.org/10.5768/JAO202142.0602007>

基于结构光的植保无人机障碍物在线检测系统

Obstacle detection system of plant protection UAVs based on structural light

应用光学. 2018, 39(3): 343-348 <https://doi.org/10.5768/JAO201839.0301008>

基于无人机载的光电与SAR图像融合技术研究

Fusion method of optical image and SAR based on UAV

应用光学. 2017, 38(2): 174-179 <https://doi.org/10.5768/JAO201738.0201004>



关注微信公众号, 获得更多资讯信息

文章编号: 1002-2082 (2024) 04-0723-09

基于 IDOU-YOLO 的红外图像无人机目标检测算法

陈海永¹, 刘登斌¹, 晏行伟^{2,3}

(1. 河北工业大学 人工智能与数据科学学院, 天津 300401; 2. 国防科技大学 电子科学学院, 湖南 长沙 410073;
3. 天津先进技术研究院, 天津 300459)

摘要: 低空小型无人机 (unmanned aerial vehicle, UAV) 侵扰敏感区域事件频发, 使国家和社会面临严重安全威胁。针对基于热成像的无人机目标检测存在漏检率高、检测精度不足的问题, 提出了 IDOU-YOLO (infrared detection of UAV-YOLO) 算法模型, 通过构建多尺度融合特征金字塔机制, 充分挖掘特征空间信息, 聚焦尺度的信息融合及丰富模型的信息表征能力, 增强目标检测能力; 同时引入了边界框损失函数 SIOU (Scylla IoU), 在训练过程中提高模型的检测精度, 加快模型的收敛速度。实验结果显示 IDOU-YOLO 模型的精确率、召回率、 F_1 分数、mAP@0.5 和 mAP@0.5:0.95 分别达到 99.2%、96.3%、97.7%、98.4% 和 70.2%, 表明 IDOU-YOLO 算法模型在红外无人机目标检测任务中具有显著优势和应用潜力。

关键词: 无人机探测; 目标检测; 红外图像; 特征金字塔网络

中图分类号: TN219; TP391.4

文献标志码: A

DOI: 10.5768/JAO202445.0402001

Infrared image UAV target detection algorithm based on IDOU-YOLO

CHEN Haiyong¹, LIU Dengbin¹, YAN Xingwei^{2,3}

(1. School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300401, China;
2. College of Electronic Science, National University of Defense Technology, Changsha 410073, China;
3. Tianjin Institute of Advanced Technology, Tianjin 300459, China)

Abstract: Low-altitude small unmanned aerial vehicle (UAV) frequently invade sensitive areas, which posing a serious threat to national and social security. Aiming at the problems such as high missed detection rate and insufficient detection accuracy for UAV target detection based on thermal imaging, the infrared detection of UAV-YOLO (IDOU-YOLO) algorithm model was proposed. A cross-scale fusion feature pyramid mechanism was constructed to fully explore the feature space information, focus on the cross-scale information fusion and the rich information representation ability of the model, and enhance the target recognition ability. At the same time, the bounding box loss function Scylla IoU (SIOU) was introduced to improve the detection accuracy and accelerate the convergence speed of the model in the training process. The experimental results show that the precision, recall, F_1 score, mAP@0.5 and mAP@0.5:0.95 reach 99.2%, 96.3%, 97.7%, 98.4% and 70.2%, respectively, which indicates that the IDOU-YOLO model improves the detection and recognition ability of UAV targets in various scenarios, and can better meet the application requirements of anti-UAV systems.

Key words: UAV detection; target detection; infrared images; feature pyramid networks

引言

近年来, 低空小型无人机凭借其体积小、机动

性强、易操作和价格便宜的优点, 被广泛应用于军事、农业、物流、测绘、救援等领域。为有效应对

收稿日期: 2023-05-04; 修回日期: 2023-11-06

基金项目: 国家自然科学基金项目 (U21A20482, 62073117, 62173124); 中央引导地方科技发展资金项目 (206Z1701G); 河北省自然科学基金 (F2022202064)

作者简介: 陈海永 (1980—), 男, 博士, 教授, 主要从事图像处理、机器视觉、模式识别研究。E-mail: haiyong.chen@hebut.edu.cn

通信作者: 晏行伟 (1985—), 男, 博士, 副教授, 主要从事多源信息融合、无源定位研究。E-mail: yanxingwei@nudt.edu.cn

无人机在军事战场的威胁以及社会安全问题,无人机探测技术已成为国内外研究者关注的热点。目前,无人机探测主要基于视觉、无线电频谱、雷达和声学探测技术^[1],其中视觉探测作为主要的无源探测方式被应用于各类反无人机系统中。但是基于可见光的视觉探测器在夜间或雾雪等恶劣天气中难以发挥作用,而红外热成像探测器利用热成像技术探测物体本身的红外特征,依据目标与背景之间的红外辐射差异来检测目标。虽然红外热成像探测器能够有效执行探测任务,但在应用中易受到目标特征弱小、杂波大、复杂背景噪声等因素的影响,造成红外热成像无人机检测困难、漏检率高的问题。

传统的无人机检测主要基于梯度、滤波和低秩稀疏恢复等方法,利用无人机目标的形状、大小和运动轨迹等先验知识,抑制背景和增强目标,以区分检测目标与背景。WANG W 等人^[2]提出一种基于空间多尺度分解滤波和时间多尺度差分处理的无人机目标检测算法;PANG D 等人^[3]利用视频的运动信息在空间域和时间域的映射分别得到空间显著图和时间显著图,并将二者融合以提升运动目标的检测性能。由于传统无人机检测方式依赖根据先验知识而精心设计的手工特征,在实际应用中的算法鲁棒性较差,检测性能不理想。随着深度学习技术的不断发展,研究人员开始使用卷积神经网络来提高检测性能。FANG H 等人^[4]针对红外无人机目标检测问题,提出了一种新的多尺度 U-Net 结构,融合不同尺度的特征图进行图像重建,通过生成背景残差图的方式区分无人机目标;夏铭江^[5]进一步提出了一种基于扩张残差全卷积自编码器网络的残差预测方法;欧洪璇^[6]基于区域建议的思想,设计了基于形状先验分割和多尺度特征的红外无人机目标检测方法;XU Z 等人^[7]提出了一种具有不对称注意力融合机制的模型算法,利用注意力机制实现深层语义信息和浅层位置信息的交互;MISBAH M 等人^[8]改进了一阶段目标检测算法 YOLOv5,有效提高了检测性能。基于深度学习的红外无人机目标检测方法,主要分为基于残差重建和目标特征学习两类方法。由于对标注数据的要求和计算资源的消耗,基于残差重建的方法逐渐被研究者摒弃。现有目标特征学习方法虽然通过注意力机制和多尺度卷积的方法改善小目标检测效果,但在一些诸如背景杂波信号复杂、运动模糊、目标与背景热交叉等实

际应用场景中,依然存在检测困难、漏检严重的问题。

在小尺寸无人机目标检测场景中,小目标本身存在因分辨率不足导致的细节信息不够准确,以及小目标容易与背景混淆的缺点,而神经网络的加深也会导致小目标特征丢失。针对无人机探测中存在检测困难、漏检严重的问题,本文以 YOLOv7^[9] 算法作为红外无人机目标检测的基准网络,聚焦特征多尺度信息,提出了 IDOU-YOLO (infrared detection of UAV-YOLO) 算法。通过设计多尺度融合金字塔网络结构,融合了丰富的浅层特征,使网络关注空间特征信息,改善了红外图像数据集中小尺寸目标检测召回率低的问题,引入 SIoU 边界框损失函数,提高算法模型精度和鲁棒性,实验证明了 IDOU-YOLO 有效提高对红外无人机目标的检测能力。

1 检测算法

1.1 IDOU-YOLO 网络模型

如图 1 所示, IDOU-YOLO 网络模型分为输入端(Input)、主干网络(Backbone)、颈部网络(Neck)以及预测头(Head)4 个部分。网络主要由卷积层、高效层聚合网络(efficient layer aggregation networks, ELAN)模块、最大池化卷积(max pooling convolution, MPC)模块和跨阶段部分通道空间金字塔池化(spatial pyramid pooling and cross-stage partial channel, SPPCSPC)模块组成。

输入端将图像映射成 $640 \times 640 \times 3$ 的特征向量,经过主干网络由浅层到深层提取特征。主干网络中的 $80 \times 80 \times 256$ 、 $40 \times 40 \times 512$ 和 $20 \times 20 \times 1024$ 的特征向量作为颈部网络的输入,通过特征金字塔结构增强颈部网络的表征能力。融合后的 $80 \times 80 \times 256$ 、 $40 \times 40 \times 512$ 和 $20 \times 20 \times 1024$ 特征向量输出到预测头,以预测不同尺寸大小的目标。预测头输出网络模型的检测结果,最后通过非极大值抑制(non-maximum suppression, NMS)筛选预测框,输出最终预测目标边界框和类型信息。

1.2 多尺度融合金字塔

YOLOv7 网络使用路径聚合网络(path aggregation network, PANet)^[10] 作为颈部网络用于特征融合,考虑到其在路径中仅聚合相同尺寸的特征信息,无法充分利用不同层特征信息,因此设计了多尺度融合金字塔网络(multi-scale merged feat-

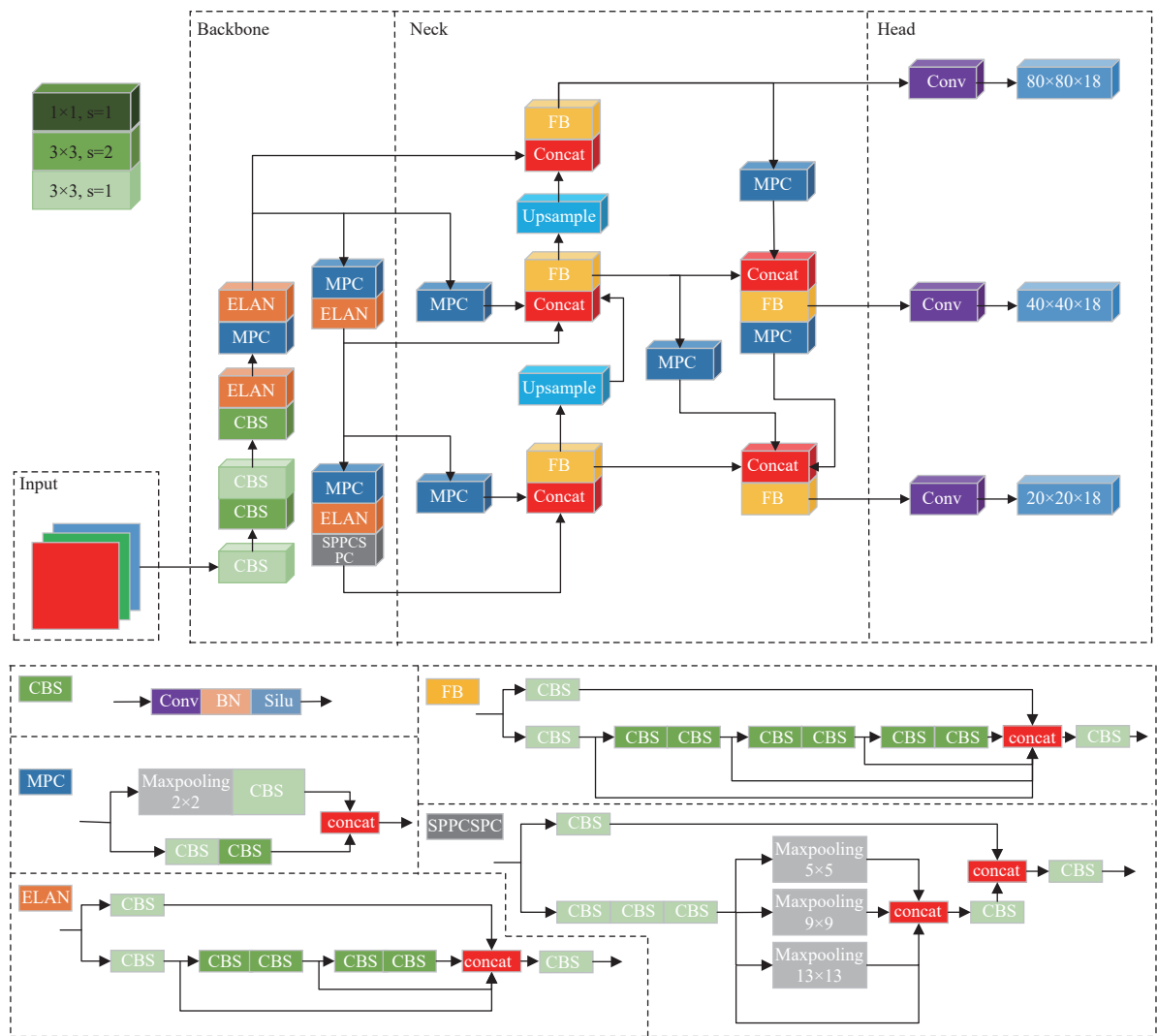


图 1 IDOU-YOLO 算法模型图

Fig. 1 IDOU-YOLO algorithm model

ure pyramid network, MMFPN)。在主干网络中提取不同分辨率特征,将语义信息更强的深层特征进行上采样,通过横向连接,将上采样的结果和在主干网络中提取的特征融合,在深层语义特征嵌

入图像的浅层信息,实现目标的特征融合和精准定位,在路径中聚合更加丰富的尺度信息,促进不同空间尺度之间的信息融合,其结构与常见颈部网络^[10-11]如图 2 所示。

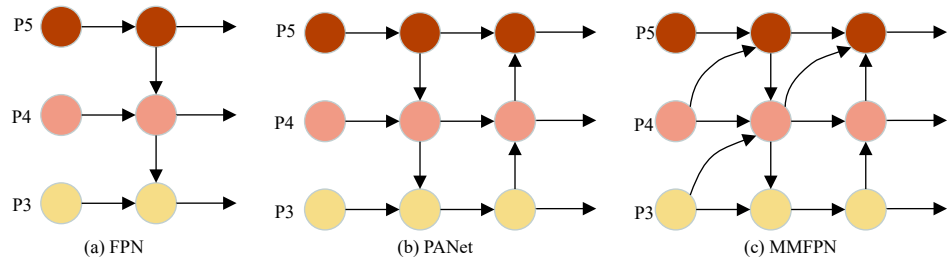


图 2 MMFPN 与常见颈部网络结构图

Fig. 2 Structure diagram of MMFPN and common neck network

特征金字塔网络能够以较低的计算代价近似图像金字塔,尽管通过自底向上的路径拓展增强

特征层次,利用浅层特征精确定位,但缺乏深层信息和浅层信息之间的交流。MMFPN 的中间节点

充分考虑了网络路径中的多层尺度信息,相比 PANet 融入了大尺寸层的特征信息。不同尺度的特征通过上采样和下采样算子对齐,促进模型的高级信息和低级信息交换,克服目标由于距离造成的尺度变化对检测的影响。

为了将颈部网络中提取到的不同尺度特征图的特征信息融合,需要对 MMFPN 路径中不同层特征图经过缩放。上采样使用二次线性插值方法提高特征图的尺寸,下采样则通过最大池化和卷积层融合的 MPC 模块完成。

特征融合模块(fusion block, FB)对在通道维度上拼接后的特征进行融合,增强特征交互信息。如图 3 所示,特征融合模块采用高效的网络设计,同时权衡不同尺寸特征的通道数量的比例以及网络深度,可带来一定精度优势^[12]。

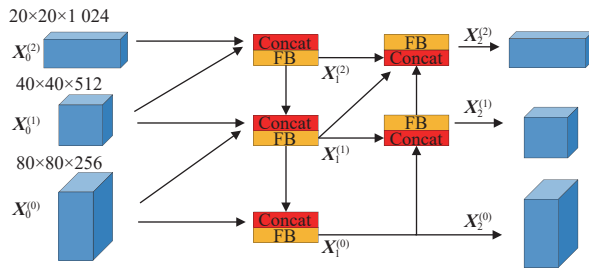


图 3 MMFPN 路径示意图

Fig. 3 Schematic diagram of MMFPN path

为说明多尺度融合金字塔网络特征融合路径,示意图省略了尺度特征图的缩放处理。如图 3 所示,定义 P3、P4、P5 这 3 个尺度的特征图为 $X_m^{(0)}$ 、 $X_m^{(1)}$ 、 $X_m^{(2)}$,其中 m 的取值为 0、1 和 2。主干网络中提取的 P3、P4、P5 层的特征图分别为 $X_0^{(0)}$ 、 $X_0^{(1)}$ 、 $X_0^{(2)}$,MMFPN 上采样路径中的 P3、P4、P5 层特征为 $X_1^{(0)}$ 、 $X_1^{(1)}$ 、 $X_1^{(2)}$,MMFPN 下采样路径中的 P3、P4、P5 层特征融合结果 $X_2^{(0)}$ 、 $X_2^{(1)}$ 、 $X_2^{(2)}$,特征图的计算方式如下文。

1) 将 $X_0^{(1)}$ 经过 MPC 模块下采样处理,得到与 P5 层尺度相同的特征图,与特征图 $X_0^{(2)}$ 在通道维度上进行拼接后,经过 FB 模块进行融合,得到特征图 $X_1^{(2)}$,如式(1):

$$X_1^{(2)} = \text{FB}(\text{concat}[X_0^{(2)}, \text{MPC}(X_0^{(1)})]) \quad (1)$$

2) 将 $X_1^{(2)}$ 经过线性插值上采样处理得到的特征图、 $X_0^{(0)}$ 经过 MPC 模块下采样处理得到的特征图和特征图 $X_0^{(1)}$,在通道维度上进行拼接后,经过 FB 模块进行融合,得到特征图 $X_1^{(1)}$,如式(2):

$$X_1^{(1)} = \text{FB}(\text{concat}[\text{Upsample}(X_1^{(2)}), X_0^{(1)}, \text{MPC}(X_0^{(0)})]) \quad (2)$$

3) 将 $X_1^{(1)}$ 经过线性插值上采样处理得到的特征图与特征图 $X_0^{(0)}$ 在通道维度上进行拼接后,经过 FB 模块进行融合,得到特征图 $X_1^{(0)}$,如式(3):

$$X_1^{(0)} = \text{FB}(\text{concat}[X_0^{(0)}, \text{Upsample}(X_1^{(1)})]) \quad (3)$$

4) 由于特征图 $X_0^{(2)}$ 未融合其他特征层特征,保留 $X_1^{(0)}$ 作为特征图 $X_2^{(0)}$,如式(4):

$$X_2^{(0)} = X_1^{(0)} \quad (4)$$

5) 将 $X_2^{(0)}$ 经过 MPC 模块下采样处理得到的特征图与特征图 $X_1^{(1)}$ 在通道维度上进行拼接后,经过 FB 模块进行融合,得到特征图 $X_2^{(1)}$,如式(5):

$$X_2^{(1)} = \text{FB}(\text{concat}[\text{MPC}(X_2^{(0)}), X_1^{(1)}]) \quad (5)$$

6) 将 $X_1^{(1)}$ 经过 MPC 模块下采样处理得到的特征图、 $X_2^{(1)}$ 经过 MPC 模块下采样处理得到的特征图和特征图 $X_1^{(0)}$ 在通道维度上进行拼接后,经过 FB 模块进行融合,得到特征图 $X_2^{(2)}$,如式(6):

$$X_2^{(2)} = \text{FB}(\text{concat}[\text{MPC}(X_1^{(1)}), X_1^{(0)}, \text{MPC}(X_2^{(1)})]) \quad (6)$$

1.3 边界框损失函数

YOLOv7 中采用边界框定位损失函数 CIOU (Complete-IoU)^[13],CIOU 的设计考虑了预测框和真实框的重叠面积、中心点距离和长宽比 3 种几何参数。SIOU^[14] 引入角度向量回归,重新定义惩罚项损失函数,角度损失约束使预测框中心位置加速回归到最近的坐标轴,之后只需要在一维方向上进行坐标 X 或 Y 的回归,降低了回归角度的自由度,加快了损失函数的收敛速度和回归精度。SIOU 损失函数由 4 个代价函数组成:角度损失、距离损失、形状损失和 IoU 损失。图 4 为 SIOU 参数示意图,其中红色细框代表预测目标框,绿色粗框代表真实目标框。

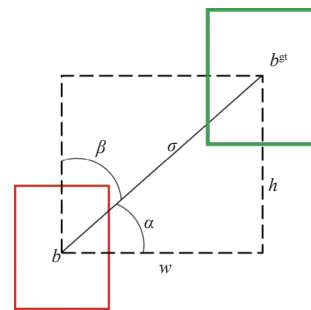


图 4 SIOU 损失参数示意图

Fig. 4 Schematic diagram of SIOU loss parameters

1) 角度损失, 其设计思想是最大限度地减少与距离相关的变量的数量, 计算公式如式(7):

$$\Lambda = 1 - 2\sin^2\left(\arcsin\left(\frac{h}{\sigma}\right) - \frac{\pi}{4}\right) \quad (7)$$

式中: h 为预测框与真实框的中心点高度差; σ 为预测框与真实框中心点之间的欧式距离。

2) 距离损失, 由上述角度损失重新定义距离损失, 计算公式如式(8):

$$\Delta = \sum_{l=x,y} (1 - e^{-\gamma \rho_l}) \quad (8)$$

其中 ρ_x 、 ρ_y 、 γ 分别为

$$\rho_x = \left(\frac{b_{cx}^{gt} - b_{cx}}{c_w}\right)^2 \quad (9)$$

$$\rho_y = \left(\frac{b_{cy}^{gt} - b_{cy}}{c_h}\right)^2 \quad (10)$$

$$\gamma = 2 - \Lambda \quad (11)$$

式中: $(b_{cx}^{gt}, b_{cy}^{gt})$ 表示真实框中心点的位置; (b_{cx}, b_{cy}) 表示预测框中心点的位置; (c_w, c_h) 表示真实框和预测框外接矩形的宽和高。当 α 越接近 0° 或 90° 时, 对距离损失的贡献越小; 越接近 45° 时, 对距离损失的贡献越大。

3) 形状损失, 定义为

$$\Omega = \sum_{l=w,h} (1 - e^{-\omega_l})^\theta \quad (12)$$

$$\omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad (13)$$

$$\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (14)$$

式中: w 、 h 、 w^{gt} 、 h^{gt} 分别是预测框和真实框的宽、高; θ 是控制形状损失的超参数。

4) IoU 损失 L_{IoU} , 为预测边界框和真实边界框的交集和并集之比。

综合上述 4 种损失, SIOU 的损失函数定义为

$$L_{SIOU} = 1 - L_{IoU} - \frac{\Delta + \Omega}{2} \quad (15)$$

2 实验及结果分析

2.1 数据集

实验选用 ICCV 2021 挑战赛发布的 anti-UAV 红外无人机跟踪^[15]数据集。按照 4 : 1 的比例将数据集划分为训练集与测试集, 最终得到 5 770 张训练集和 1 201 张测试集图像。

2.2 实验评估指标

本文以精确率 (precision, P)、召回率 (recall,

R), F_1 分数 (F_1 -score, F_1)、平均精度 (average precision, AP, P_{AP}), 均值平均精度 (mean average precision, mAP, P_{mAP}) 作为评估模型指标。精确率、召回率、 F_1 分数的计算公式分别为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (16)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (17)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (18)$$

式中: N_{TP} 表示预测为正样本的真实正样本数量; N_{FP} 表示预测为正样本的真实负样本数量; N_{FN} 表示预测为负样本的正样本数量。

AP 和 mAP 表示单个类别和全部类别的平均精度。由于本文研究的检测目标只有单一类别, AP 和 mAP 在数值和意义上相同。实验选用 $P_{mAP}@0.5$ 和 $P_{mAP}@0.5 : 0.95$ 作为评估指标, 分别表示真实框与预测框交并比 (IoU) 阈值为 0.5 时的 mAP 值和在不同交并比阈值 (从 0.5 到 0.95, 步长为 0.05) 上的平均 mAP。计算公式为

$$P_{AP} = \int_0^1 P(R) dR \quad (19)$$

$$P_{mAP} = \frac{1}{C} \sum_{i=1}^C P_{AP_i} \quad (20)$$

$$P_{mAP}@0.5 : 0.95 = \frac{1}{9} \sum_{IoU=0.5}^{0.95} P_{mAP}^{IoU} \quad (21)$$

式中: P 表示精确率; R 表示召回率; C 表示目标类别总数; P_{AP_i} 表示第 i 类目标 AP 值; P_{mAP}^{IoU} 表示该交并比阈值下的 mAP 值。

2.3 实验环境及参数设置

本文所进行的实验均基于操作系统为 Ubuntu 18.04 的服务器平台, 搭载 Intel Xeon Gold 5120 CPU @ 2.2 GHz 和 NVIDIA GeForce RTX 2080ti。算法模型基于 Python 3.10 编程语言, 通过 Pytorch 1.12 框架搭建深度学习网络。

实验中设置模型训练轮次为 400, batch size 为 8, 图像大小设置为 640×640 像素, 优化器选择随机梯度下降法 (stochastic gradient descent, SGD), 初始学习率设为 0.01, 动量设置为 0.937, 衰减系数为 0.000 5。为了提高模型的鲁棒性, 对训练集图像采用 Mosaic 数据增强方法。

2.4 实验结果与分析

为验证本文提出的 IDOU-YOLO 模型的检测

性能, 首先将 IDOU-YOLO 和 YOLOv7 在 anti-UAV 数据集上的实际结果进行对比, 实验结果如表 1 所示。

表 1 IDOU-YOLO 和 YOLOv7 在 anti-UAV 数据集上整体性能

Table 1 Overall properties of IDOU-YOLO and YOLOv7 on anti-UAV dataset

Model	P/%	R/%	$P_{mAP}@0.5/\%$	$P_{mAP}@0.5:0.95/\%$	$F_1/\%$	FPS
YOLOv7	98.9	94.2	97.0	67.6	96.5	93.5
IDOU-YOLO	99.2	96.3	98.4	70.2	97.7	79.4

表 1 的结果表明, IDOU-YOLO 的 P 、 R 、 $P_{mAP}@0.5$ 、 $P_{mAP}@0.5:0.95$ 、 F_1 指标分别为 99.2%、96.3%、98.4%、70.2%、97.7%, 相比于 YOLOv7 分别提升 0.3%、2.1%、1.4%、2.6%、1.2%, 充分证明 IDOU-YOLO 算法在以无人机为目标红外图像检测任务中的优越性。受益于一阶段目标检测算法的高效性, 即便 MMFPN 网络结构的设计带来轻微的检测效率损失, IDOU-YOLO 算法仍然能够满足实时检测的需求。

2.4.1 颈部网络对比实验

为更好地说明 MMFPN 的优势, 在输入端、主干网络和预测头的网络结构和损失函数等设置完全相同的情况下, 颈部网络分别采用 FPN、PANet、BiFPN^[16] 和 MMFPN 进行对比实验, 实验结果如表 2 所示。

表 2 颈部网络对比实验结果

Table 2 Experimental results of performance of neck network comparison

Neck	P/%	R/%	$P_{mAP}@0.5/\%$	$P_{mAP}@0.5:0.95/\%$	$F_1/\%$	FPS
FPN	97.6	89.2	95.6	66.2	93.2	123.5
PANet	98.6	96.0	98.0	69.4	97.3	95.2
BiFPN	98.7	96.2	98.4	70.1	97.4	95.2
MMFPN	99.2	96.3	98.4	70.2	97.7	79.4

表 2 的实验结果表明, 颈部网络采用 MMFPN 后的网络在精确率、召回率、 F_1 分数、mAP 性能指标上表现最好。在红外图像的无人机目标检测任务中, MMFPN 的性能优于其他 3 种特征金字塔网络结构, MMFPN 作为颈部网络丰富了模型的表征信息, 有效增强了神经网络对红外无人机目标的检测能力。

2.4.2 边界框损失函数对比实验

为验证边界框损失函数采用 SIoU 在 anti-UAV 红外无人机目标数据集上的检测性能的提升, 在 IDOU-YOLO 算法模型下, 分别采用 SIoU 与 DIoU、GIoU、CIoU 几种边界框损失函数进行对比实验, 实验结果见表 3 所示。

表 3 边界框损失函数对比实验结果

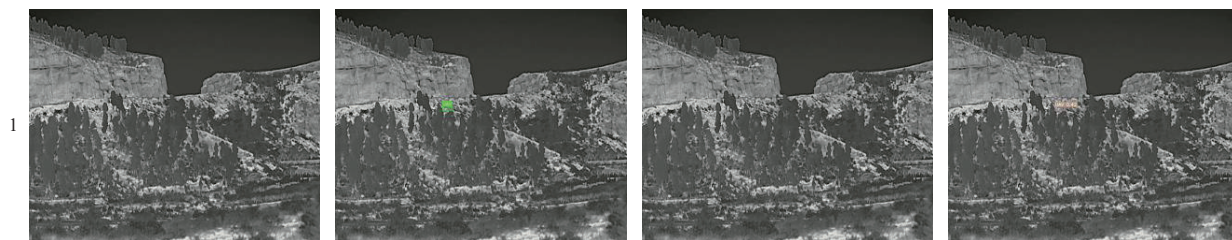
Table 3 Experimental results of bounding box loss function comparison

Loss	P/%	R/%	$P_{mAP}@0.5/\%$	$P_{mAP}@0.5:0.95/\%$	$F_1/\%$
DIoU	99.1	96.0	98.0	69.7	97.5
GIoU	98.7	96.3	97.9	69.7	97.5
CIoU	99.6	95.4	98.4	69.9	97.5
SIoU	99.2	96.3	98.4	70.2	97.7

SIoU 与 DIoU、GIoU、CIoU 边界框损失函数均保证模型训练过程最终收敛。表 3 中的实验结果表明, SIoU 除了精确率指标仅次于 CIoU 以外, 其他指标均优于其他 3 种边界框损失函数, 综合各项性能指标优于其他 3 种边界框。边界框损失函数仅在训练过程中引导网络参数优化, 不影响网络结构和推理时间, 在不改变网络结构的前提下提升检测性能。

2.5 IDOU-YOLO 算法检测结果可视化

本文提出的 IDOU-YOLO 算法以 YOLOv7 为基础, 相比 YOLOv7 对红外无人机目标检测效果有比较显著的提升, 图 5 为选取几类检测难度比较大的无人机场景对比两种算法的实际效果。



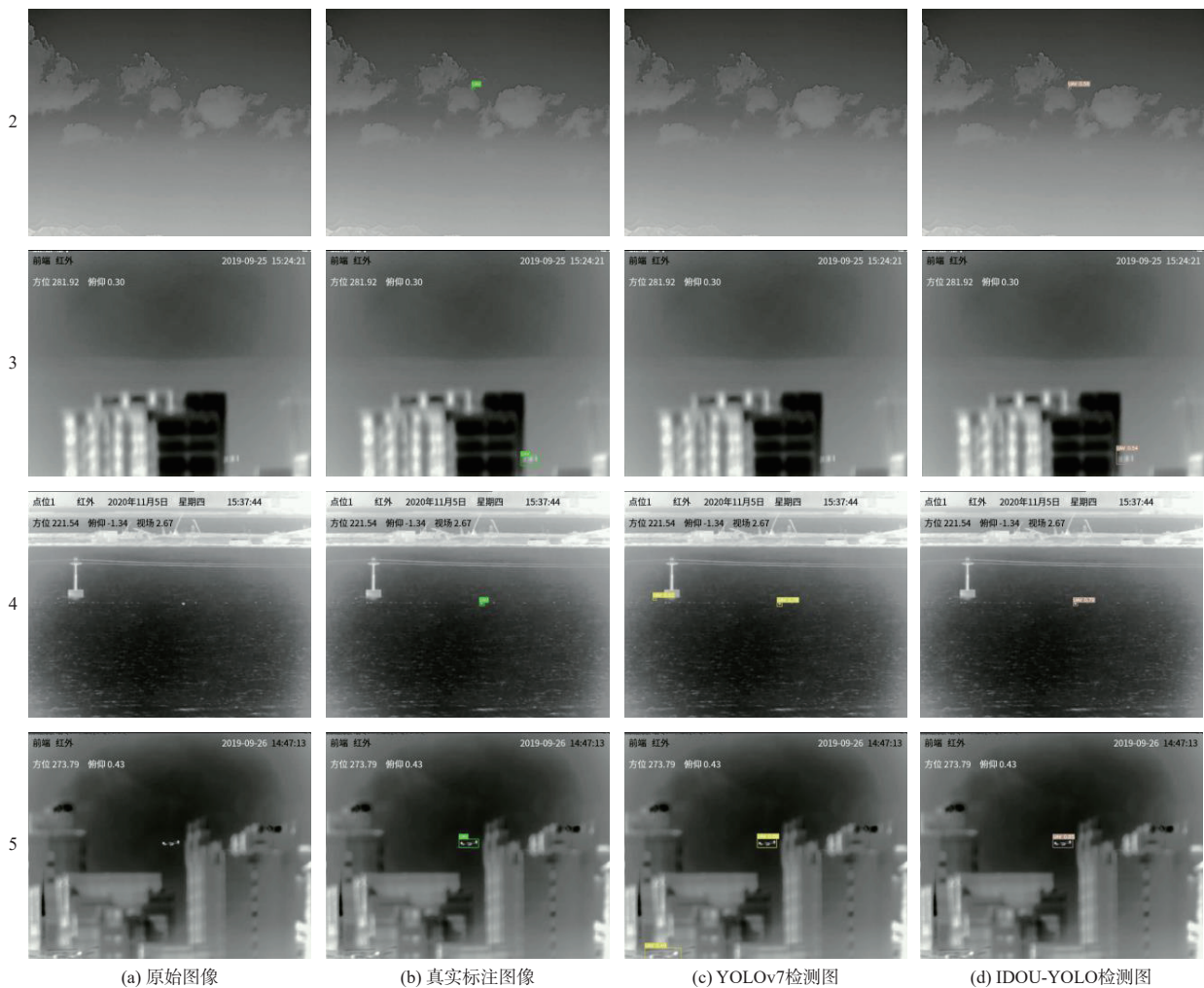


图 5 算法检测结果图

Fig. 5 Images of detection results for algorithm

如图 5 所示, 场景 1、场景 2 和场景 3 分别是复杂植被背景、小目标云层背景和运动模糊的检测场景, 其中 YOLOv7 算法无法识别目标造成了漏检, 而 IDOU-YOLO 算法能够有效从背景中提取目标特征并检测无人机目标。高虚警率是红外图像目标检测领域的一大难点, 场景 4 和场景 5 背景中存在与检测目标特征相似的区域中, YOLOv7 算法因无法准确区别目标特征而产生了误检, IDOU-YOLO 算法有效抑制了背景信息的干扰, 对小目标也有良好的检测能力。IDOU-YOLO 算法在复杂背景、小目标的检测、失焦及运动造成的特征模糊和相似背景区域干扰的情况下, 均展现了更强的鲁棒性和更高的检测精度。

2.6 IDOU-YOLO 算法对比其他常见目标检测算法

为了进一步验证 IDOU-YOLO 算法的性能, 选取目前目标检测领域中几种主流的算法进行对

比, 实验结果见表 4。

表 4 IDOU-YOLO 与主流目标检测算法的对比实验结果
Table 4 Comparative experimental results of IDOU-YOLO and mainstream target detection algorithms

Model	$P/\%$	$R/\%$	$P_{mAP}@0.5/\%$	$F_1/\%$	FPS
Faster-RCNN	34.7	66.0	52.3	45.5	19.2
SSD	98.8	60.0	86.1	74.7	74.1
Retinanet	99.7	60.9	65.8	75.6	39.5
Centernet	98.6	85.9	91.1	91.8	66.2
FCOS	98.7	81.9	91.9	89.5	41.6
YOLOX	98.8	89.8	93.1	94.1	37.9
YOLOv5	99.1	95.7	98.2	97.4	63.3
IDOU-YOLO	99.2	96.3	98.4	97.7	79.4

如表 4 所示, 两阶段目标检测算法中, Faster-RCNN^[17] 算法在精确率、召回率、mAP 和 F_1 分数指标均处于较低水平, 推理速度也无法满足实时

检测的应用需求。Retinanet^[18]作为早期的一阶段检测算法,召回率比较低,检测性能不够理想。Centernet^[19]、FCOS^[20]和YOLOX^[21]都是无锚框的检测算法,与基于锚框的YOLOv5^[22]和IDOU-YOLO算法相比,虽然精确率普遍较高,但召回率相对较低。本文提出的IDOU-YOLO目标检测算法相较于目标检测中的主流算法有着更高的精确率、召回率、mAP和 F_1 分数,能够有效应对复杂场景下的无人机检测任务,算法检测速度也满足反无人机系统的应用需求。

3 结论

本文针对小型低空无人机的红外热成像检测场景中检测困难、漏检严重的问题,提出IDOU-YOLO网络模型,通过多尺度融合金字塔网络MMFPN深度融合多层特征信息,增强模型表征能力,提高检测性能。在IDOU-YOLO模型基础上,进一步优化边界框损失函数,引入SIoU损失函数,增强模型的定位能力,从而提升了网络模型的精度。实验结果表明,在红外无人机图像数据集中,相比于主流目标检测算法模型,IDOU-YOLO算法对无人机的检测效果具有一定优势,所提算法具备很高的适用性。未来将从目标跟踪等领域对无人机目标探测任务展开研究和实验,改善无人机目标探测系统的可靠性和实用性。

参考文献:

- [1] 蒋榕圻,白若楷,彭月平.低慢小无人机目标探测技术综述[J].飞航导弹,2020(9):100-105.
JIANG Rongqi, BAI Ruokai, PENG Yueping. Overview of low slow small unmanned aerial vehicle target detection technology[J]. Winged Missiles Journal, 2020(9): 100-105.
- [2] WANG W, WANG P, NIU Z. A real-time detection algorithm for unmanned aerial vehicle target in infrared search system[C]//2018 IEEE International Conference on Signal Processing, Communications and Computing. New York: IEEE, 2018: 1-5.
- [3] PANG D, SHAN T, MA P, et al. A novel spatiotemporal saliency method for low-altitude slow small infrared target detection[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1-5.
- [4] FANG H, DING L, WANG L, et al. Infrared small UAV target detection based on depthwise separable residual dense network and multiscale feature fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-20.
- [5] 夏铭江.复杂动态场景下红外无人机目标检测法研究[D].西安:西安电子科技大学,2022.
XIA Mingjiang. Research on target detection method of infrared UAV in complex dynamic scene [D]. Xi'an: Xidan University, 2022.
- [6] 欧洪璇.基于深度学习的红外无人机目标监视技术研究[D].西安:西安电子科技大学,2022.
OU Hongxuan. Research on infrared small target monitoring method based on deep learning[D]. Xi'an: Xidan University, 2022.
- [7] XU Z, SU J, HUANG K. A-RetinaNet: a novel RetinaNet with an asymmetric attention fusion mechanism for dim and small drone detection in infrared images[J]. Mathematical Biosciences and Engineering, 2023, 20(4): 6630-6651.
- [8] MISBAH M, KHAN M U, YANG Z, et al. Tf-net: deep learning empowered tiny feature network for night-time uav detection[C]//International Conference on Wireless and Satellite Systems. Cham: Springer Nature Switzerland, 2023: 3-18.
- [9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2023: 7464-7475.
- [10] 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. New York: IEEE, 2018: 8759-8768.
- [11] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 2117-2125.
- [12] XU X, JIANG Y, CHEN W, et al. DAMO-YOLO: a report on real-time object detection design[EB/OL]. (2022-12-15) [2023-03-14]. <https://arxiv.org/abs/2211.15444>.
- [13] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

-
- [14] GEVORGYAN Z. Siou loss: more powerful learning for bounding box regression [EB/OL]. (2022-5-25) [2023-03-14]. <https://arxiv.org/abs/2205.12740>.
- [15] JIANG N, WANG K, PENG X, et al. Anti-UAV: a large multi-modal benchmark for UAV tracking[EB/OL]. (2021-02-08) [2023-07-31]. <https://arxiv.org/abs/2101.08466>.
- [16] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New York : IEEE, 2020: 10781-10790.
- [17] 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.
- [18] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. New York : IEEE, 2017: 2980-2988.
- [19] DUAN K, BAI S, XIE L, et al. Centernet: keypoint triplets for object detection[C]//Proceedings of the IEEE/CVF international conference on computer vision. New York : IEEE, 2019: 6569-6578.
- [20] TIAN Z, SHEN C, CHEN H, et al. Fcos: fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF international conference on computer vision. New York : IEEE, 2019: 9627-9636.
- [21] GE Z, LIU S, WANG F, et al. Yolox: exceeding yolo series in 2021[EB/OL]. (2021-08-06) [2022-12-25]. <https://arxiv.org/abs/2107.08430>.
- [22] ULTRALYTICS.YOLOV5[EB/OL]. [2023-8-2]. <https://github.com/ultralytics/yolov5>.