

无人机视角轻量化红外与可见光图像检测方法

蒋行国^{1,2}, 王 峣¹, 林国军^{1,2}, 孙 潇¹, 刁豪杰¹, 黎 明¹

(1. 四川轻化工大学 自动化与信息工程学院, 四川 宜宾 644000; 2. 人工智能四川省重点实验室, 四川 宜宾 644001)

摘要:针对无人机航拍视角目标检测空间尺度变化大、物体像素占比少,算法部署边缘计算平台存储空间占比大等问题。本文基于 YOLOv8n 网络结构,提出一种改进航拍视角轻量化小目标检测方法 DSF-YOLO-P 算法。首先,主干网络 C2f 模块融合 FasterNet,组成 Faster-C2f 轻量化模块,以保证模型在不影响检测精确度的同时实现网络轻量化与提升检测速度。然后,新增 160×160 预测头并重构网络通道,以提升模型对小型目标检测的准确度及鲁棒性。改进后的 DSF-YOLO 算法在可见光数据集 VisDrone2019 与红外数据集 HIT-UAV 上准确度分别提升 2.5% 与 0.6%,参数量降低 10%。最后,将 DSF-YOLO 算法进行依赖图剪枝操作,减少模型冗余参数又不影响模型性能。剪枝后的 DSF-YOLO-P 算法在 VisDrone2019 数据集上相较于 DSF-YOLO 算法,精确度不变,计算量与参数量分别降低 45% 与 26%。实验结果充分证明了 DSF-YOLO-P 算法在无人机航拍视角小目标检测中的有效性。

关键词:边缘计算平台;无人机;FasterNet;预测头;剪枝

中图分类号:TP391.41;V279 文献标识码:A DOI:10.3969/j.issn.1001-5078.2025.03.020

Lightweight infrared and visible image detection methods for UAV perspectives

JIANG Xing-guo^{1,2}, WANG Yao¹, LIN Guo-jun^{1,2}, SUN Xiao¹, DIAO Hao-jie¹, LI Ming¹

(1. School of Automation and Information Engineering, Sichuan University of Science and Engineering, Yibin 644000, China;

2. Artificial Intelligence of Key Laboratory of Sichuan Province, Yibin 644001, China)

Abstract: Aiming at the UAV aerial photography viewpoint target detection spatial scale change is large, the object pixels account for a small proportion, and the algorithm deployment edge computing platform storage space occupies a large proportion of the problem. In this paper, based on the YOLOv8n network structure, an improved aerial photography viewpoint lightweight small target detection method DSF-YOLO-P algorithm is proposed. Firstly, the backbone network C2f module is integrated with FasterNet to form the Faster-C2f lightweight module to ensure that the model achieves network lightweighting and improves the detection speed without affecting the detection accuracy. Then, a new 160×160 prediction head is added and the network channels are reconfigured to improve the accuracy and robustness of the model for small target detection. The improved DSF-YOLO algorithm improves the accuracy by 2.5% and 0.6% on the visible dataset VisDrone2019 and infrared dataset HIT-UAV, respectively, and reduces the number of parameters by 10%. Finally, the DSF-YOLO algorithm is subjected to the dependency graph pruning operation to reduce the redundant parameters of the model without affecting the model performance. The pruned DSF-YOLO-P algorithm achieves the same accuracy and reduces the computational effort and number of parameters by 45% and 26%,

基金项目:四川轻化工大学人才引进项目(No. 2019RC12);四川省科技计划项目(No. 2024YFHZ0026)资助。

作者简介:蒋行国(1973-),男,博士,副教授,研究方向为智能信息处理。E-mail:tonny_jiang@suse.edu.cn

收稿日期:2024-06-19;修订日期:2024-07-15

测尺寸小,分辨率低,细节特征丢失严重的小尺度目标,更倾向于大中型尺度目标的检测识别。YOLOv8n 原模型中对输入图像下采样操作,生成 80×80 、 40×40 、 20×20 尺度特征图。三个不同维度的预测头,用以强化不同尺寸目标检测能力,方便对不同维度特征图检测。为更好捕获目标信息,提升小型物体语义识别能力。在 YOLOv8n 原模型基础上,增加第四个预测头并重构通道将 19 层及 22 层卷积核数量分别设为 256 与 512,以加强算法对中小型目标关注度,提高模型小目标检测能力。新增加预测头对应 160×160 大小特征图,虽然提升了模型计算成本,但却显著提高了微小目标检测精度及不同尺寸目标感知能力。小目标检测层如图 2 所示。

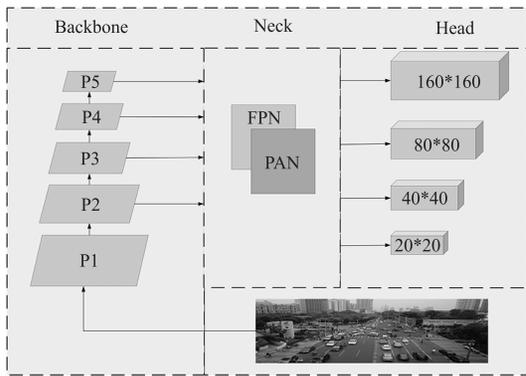


图 2 小目标检测层

Fig. 2 Small target detection layer

2.2 主干网络 C2f 融合 Fasternet

过去许多作者为设计快速神经网络,专注于减少浮点操作。然而,浮点操作的减少与延迟并非正相关,甚至会导致内存访问更加频繁。为此,基于 PConv 进一步提出新型神经网络 FasterNet^[16],通过削减频繁内存访问与冗余计算量,以在设备上实现相较其他网络更高的运行速度。FasterNet 由一个 PConv 及 2 个 PWConv 构成,设 w, h, k, c 分别为特征图宽、特征图高,卷积核尺寸以及通道数,PConv 计算量为普通卷积 $1/16$,公式为:

$$h \times w \times k^2 \times c_p^2 \quad (1)$$

其中, c_p 为 PConv 通道数,是 c 的 $1/4$ 。内存访问量是普通卷积 $1/4$,公式为:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (2)$$

如图 3、图 4 所示,FasterNet 作为一种更加节省资源的神经网络,与 C2f 模块通道到像素融合为轻量化模块 Faster-C2f,为模型的计算性能带来提升。

2.3 模型剪枝

模型剪枝通过修剪冗余参数以达到模型轻量化目的,在面向需要压缩深度学习网络的边缘计算平台时表现出显著的实用性与有效性。然而,需要进行剪枝操作的网络结构内部模块往往通过复杂的内部连接相耦合,如果不考虑其相互之间的依赖性,结束剪枝操作之后可能会得到一个错误的网络。

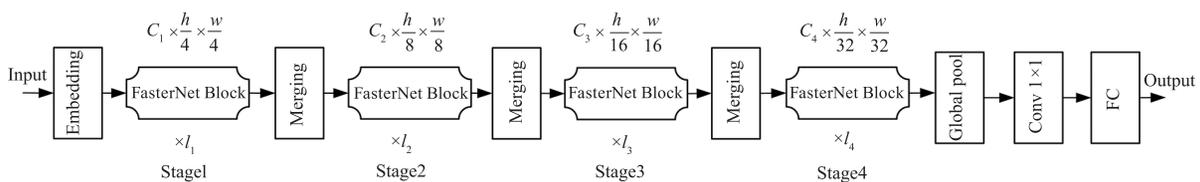


图 3 FasterNet 网络结构

Fig. 3 FasterNet network architecture

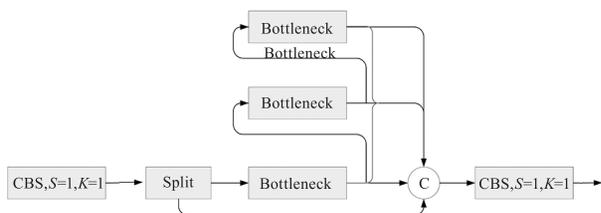


图 4 C2f 结构

Fig. 4 C2f structure

在复杂的网络结构中,各组参数之间可能存在依赖关系,需要同时对其进行修剪。如图 5(a)所示,该线性神经网络参数由 w_1, w_{l+1} 与 w_{l+2} 三个

二维权重矩阵构成,很简单就能发现参数之间的一些依赖关系。但如果神经网络是存在多种依赖关系的图 5(b)、5(c),并且通过嵌套组成更为复杂的结构,则难以分析所有依赖关系。为此,Dep-Graph 剪枝方法^[17]提供了一种面向任意结构的一般结构剪枝,可以显示的建模各个层的相关性。首先设定一个组 $g = \{w_1, w_2, w_3\}$,通过观察,可以从 $w_1 \leftrightarrow w_3$ 递归推导出 $w_1 \leftrightarrow w_2$ 与 $w_2 \leftrightarrow w_3$,这个递归过程最终以一个传递关系 $w_1 \leftrightarrow w_2 \leftrightarrow w_3$ 结束。从以上递归过程不难看出,只需要两个依赖关系就

能描述组 g 内的关系。而依赖图可以看成是分组矩阵 G 的简化,保留相同信息的同时压缩得更加紧凑。

通过 DepGraph 分组能力全面稀疏化模型中的各个层,能够安全剪枝各层,在性能不至于降低太多的情况下,实现对改进模型的剪枝操作。

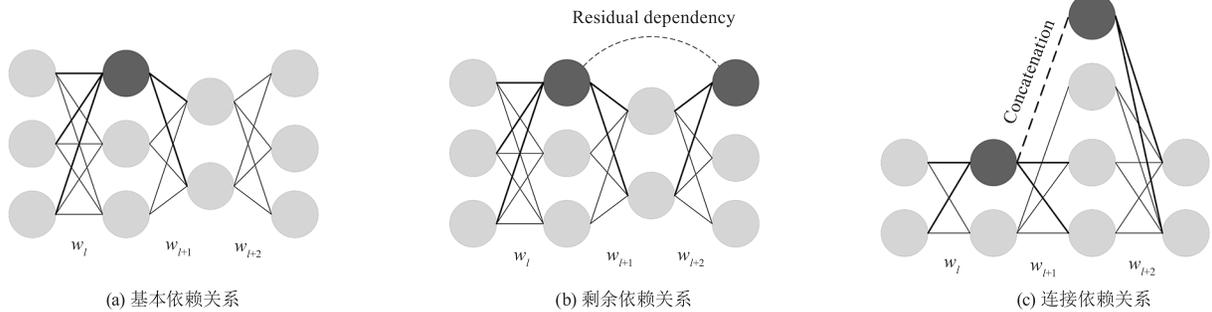


图5 神经网络之间依赖
Fig.5 Connection dependency

3 实验与结果

3.1 环境及参数设置

为保证实验的公平性及可比较性,所有实验训练、验证均使用相同硬件环境及参数设置。如表1所示。

表1 实验配置及训练参数

Tab.1 Experimental configuration and training parameters

名称	参数
操作系统	Windows10 64 位
显卡	RTX2070SUPER
CUDA	10.2
深度学习框架	PyTorch1.12.1
语言	Python3.8
学习率(lr0)	0.01
批量大小(batchsize)	16
图片尺寸	640 × 640

3.2 数据集

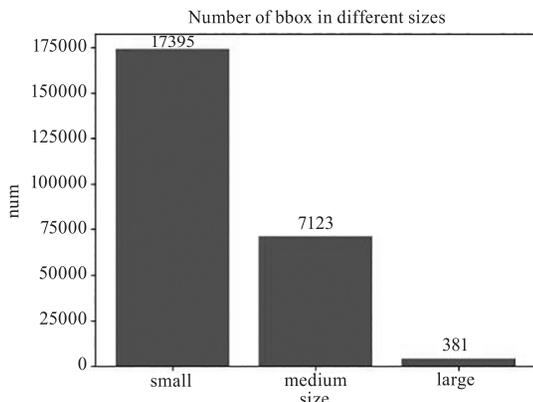
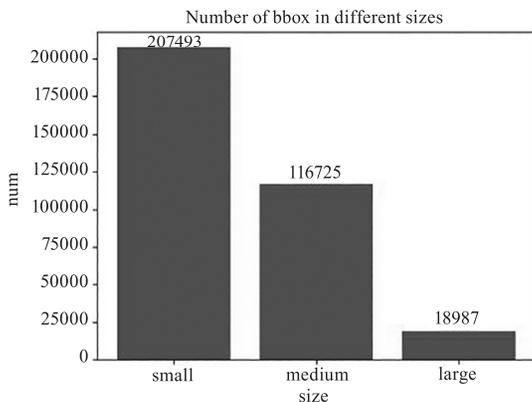
VisDrone2019 数据集是天津大学 AISKYEYE 团

队收集并开源的大型无人机数据集,有训练集 6741 张,验证集 548 张。共有十个类别,分别是行人、人、自行车、货车、汽车、三轮车、公共汽车、遮阳三轮车、卡车以及轿车。VisDrone2019 数据集分析图如图 6 (a)所示,小目标含量达到 60%,大部分图片每张只含有一类或两类目标。

HIT-UAV 数据集是以无人机视角在学校、停车场、道路和操场等各类场景下拍摄的高空红外数据集。本实验将其划分为训练集 2608 张,验证集 290 张。删除数据较少的几类,只保留人、自行车以及汽车三类。HIT-UAV 数据集分析图如图 6 (b)所示,小目标含量 70%,大尺寸目标多为长方形。

3.3 模型评价指标

本实验选取目标检测领域重要评价指标对模型进行性能评估。分别是 mAP (Mean of Average Precision)、FPS (Frames Per Second)、Params 以及 GFLOPs。其中 mAP 计算公式如下:



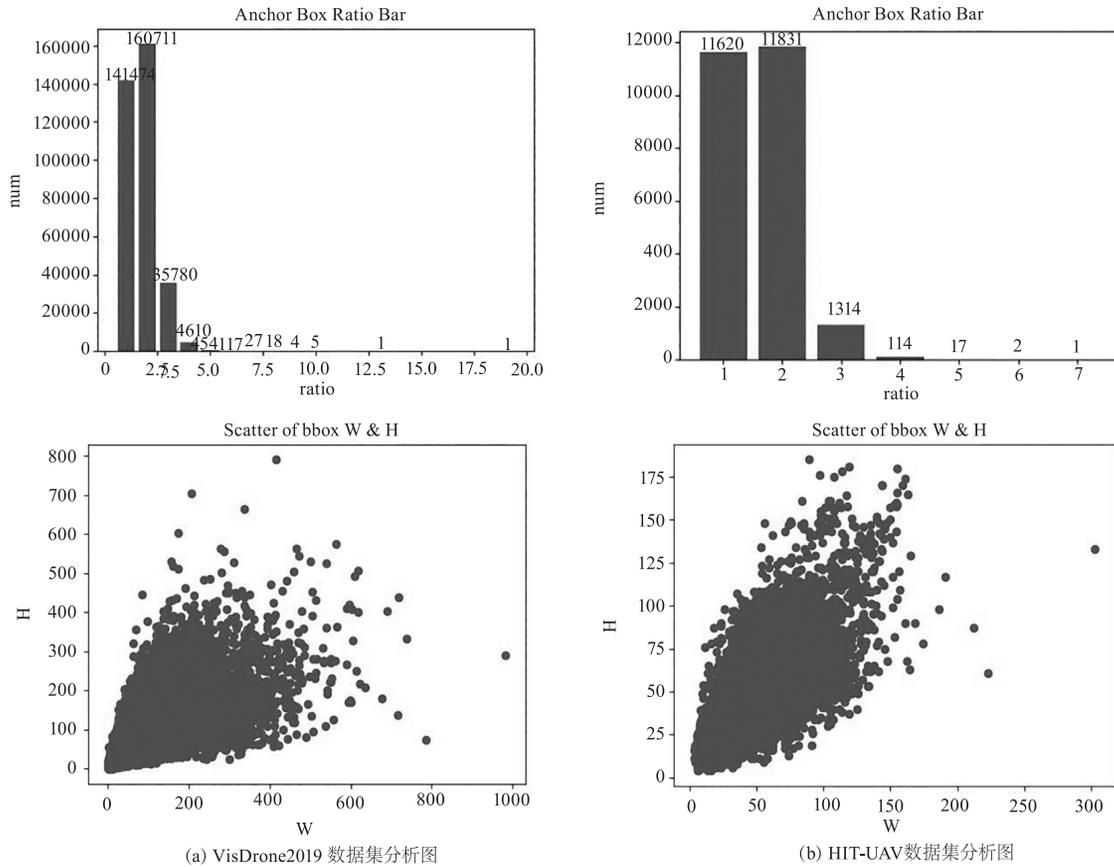


图 6 数据集分析图

Fig. 6 Data set analysis chart

$$AP = \int_0^1 p(r) d(r) \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (4)$$

式中, N 为类别的数量; AP 为每个类别的平均精度。在本实验检测任务中, VisDrone 2019 数据集 $N = 10$, HIT-UAV 数据集 $N = 3$ 。

3.4 Faster-C2f 性能对比实验

为验证 Faster-C2f 轻量化模块对航拍视角小型目标提升效果, 本实验于红外高热无人机数据集 HIT-UAV 上进行对比实验, 探究 Faster-C2f 模块添加数量对模型性能影响。对已添加新预测头的基线模型分别添加 0、1、2、3 个 Faster-C2f 模块, 实验结果如表 2 所示。

表 2 Faster-C2f 性能对比实验

Tab. 2 Faster-C2f performance comparison experiment

个数	mAP@0.5/%	FPS	Params/M
0	95.80	91	2.98
1	95.80	92	2.70
2	95.60	93	2.65
3	95.30	89	2.62

由表 2 可以看出, 随着 Faster-C2f 模块的不断添加, 参数量也随之减少。但相应, 自添加一个轻量化模型后, 模型精确度也开始开始小幅度降低。综合考虑, 选择将一个 Faster-C2f 模块加入骨干网络, 此时模型精确度不变, 参数量降低 0.28 M。

3.5 消融实验

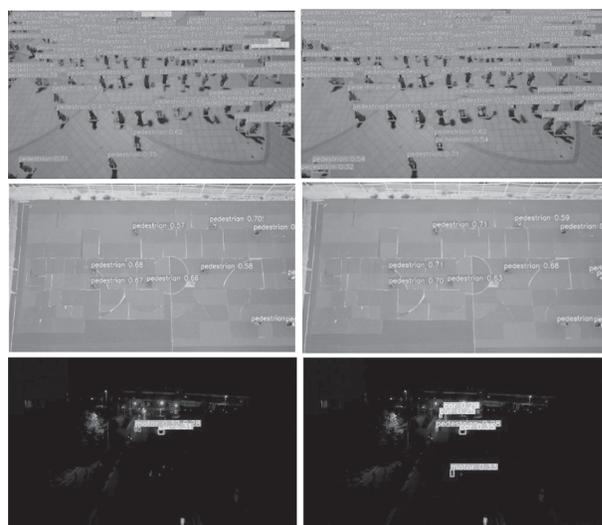
为验证改进模块对改进模型性能提升, 实验以 YOLOv8n 为基线模型, 将不同改进模块依次加入改进模型进行性能评估。训练使用预训练权重, 训练次数 200 轮。

由表 3 可知, 基线模型加入 Faster-C2f 模块后, 精确度并未有太大变化, 但参数量降低 7%, FPS 也有了明显提升。说明 Faster-C2f 模块在未影响模型识别精度的情况下, 削减了算法冗余计算量并有效提升了运行速度; 将 160×160 尺度预测头加入基线模型, 虽然 FPS 下滑了 14, 但精确度提升了 2.5%。说明新预测头成功加强了模型对小目标物体关注度; 最后, 在基线模型中将两个改进点依次加入得到 DSF-YOLO 算法, 模型精确度提升 2.5%, 参数量下降 10%, 而 FPS 未有太大下滑。由此, 证明改进模型面向航拍视角小目标检测领域的有效性。

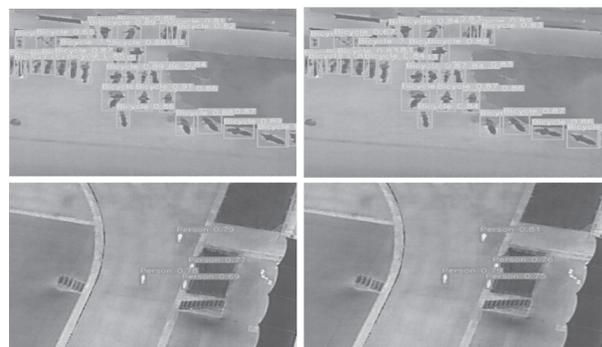
表 3 消融实验

Tab.3 Ablation experiment

模型	mAP@0.5		FPS (HIT-UAV)	Params (HIT-UAV)/M
	VisDrone2019/%	HIT-UAV/%		
基线模型	34.80	95.20	105	3.01
+ Faster-C2f	35.00	95.30	109	2.79
+ 新预测头	37.30	95.80	91	2.98
Ours	37.30	95.80	92	2.70



(a) 基线算法检测可见光图像 (b) DSF-YOLO算法检测可见光图像



(c) 基线算法检测红外图像 (d) DSF-YOLO算法检测红外图像

图 7 检测效果对比

Fig.7 Comparison of detection effect

图 7 是算法对不同场景小目标检测效果对比。可以看出在极端昏暗条件下,由于目标特征丢失严重导致基线算法识别困难,即使是表现相对较好的 DSF-YOLO 算法也存在大量漏检现象。而当检测对象转变为红外图像时,两类算法稳定检测目标则无需考虑光照因素。无论是处于红外还是可见光条件下,DSF-YOLO 算法检测稀疏小目标均能更精确识别到目标。但当检测对象更加密集,基线算法便无法准确识别到目标。在可见光图像稀疏小目标检测任务中,基线算法识别出 162 个目标,DSF-YOLO 算法则检测出 202 个目标。展现出 DSF-YOLO 算法优异的鲁棒性。

3.6 模型剪枝对比

为验证不同剪枝比例下改进模型性能,寻找最优剪枝比例。本实验通过对模型剪枝比例进行调整,对比基线模型与剪枝模型性能指标,实验结果如表 4 所示,剪枝后算法与基线算法通道对比如图 5 所示。

由表 4 可知,随着对 Seed_up 的不断加大,模型的参数量与 GFLOPs 也相应不断减少。然而,剪枝率过高可能会导致模型性能下降,进而导致算法准确度降低,在剪枝比例与算法准确度之间寻找一个平衡点尤为关键。综合表 3 实验数据可知,当 Seed_up 为 1.25 时,模型能够在精确度不被影响的情况下最大比例剪除冗余网络通道,故最终选定 Seed_up 为 1.25 作为 DSF-YOLO-P 算法的剪枝比例。

由图 8 可知,剪枝的通道主要集中在 7~9 层与 22~28 层。并且大部分层都被剪除过一部分,说明了剪枝方法对 DSF-YOLO 算法有效。

表 4 模型剪枝对比实验

Tab.4 Model pruning comparison experiment

模型	Seed_up	mAP@0.5/%	GFLOPs	Params/M	模型大小/M
基线模型	/	34.80	8.1	3.01	6.11
DSF-YOLO	/	37.30	12.0	2.70	5.64
DSF-YOLO-P	1.25	37.30	6.6	1.99	4.23
DSF-YOLO-P	1.26	37.20	6.5	1.96	4.17
DSF-YOLO-P	1.27	37.10	6.5	1.90	4.01
DSF-YOLO-P	1.29	37.00	6.4	1.81	3.88
DSF-YOLO-P	1.35	30.10	4.1	1.03	2.35

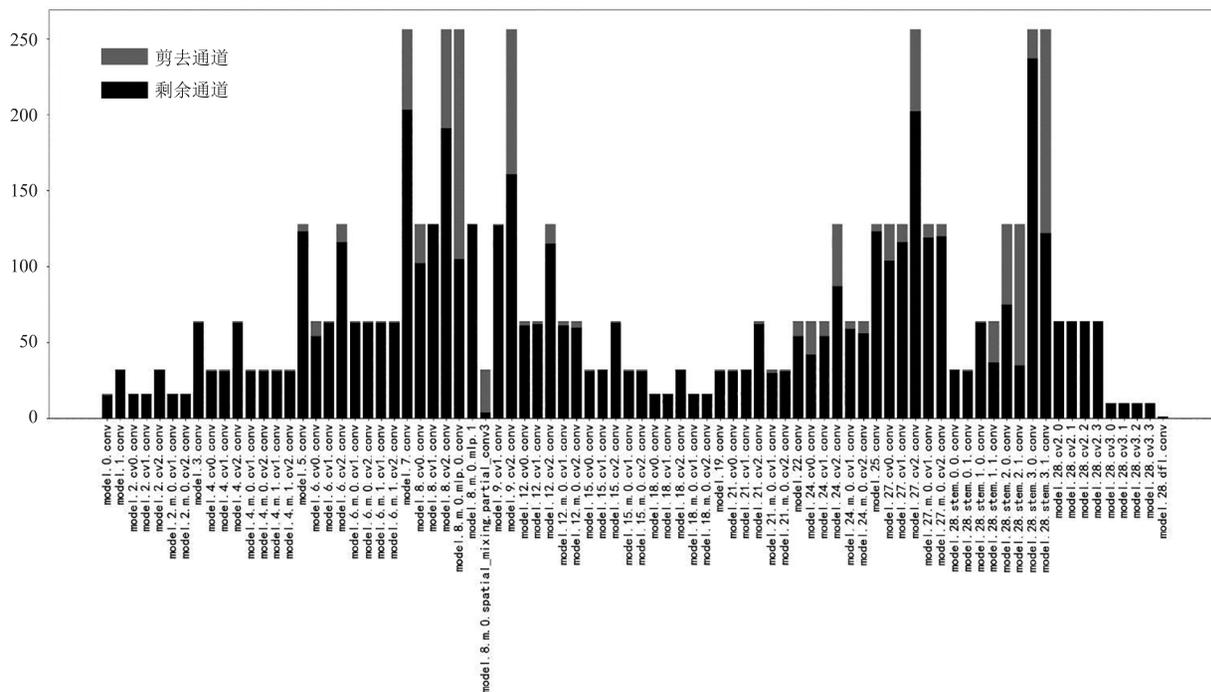


图 8 剪枝通道对比图

Fig. 8 Pruning channel comparison chart

3.7 不同模型对比实验

为进一步验证 DSF-YOLO-P 算法的有效性,本实验于 VisDrone2019 数据集上将 DSF-YOLO-P 算法与各主流检测算法相比较。比较算法包括一阶段算法与二阶段算法,比较结果如表 5 所示。

表 5 不同算法对比实验

Tab. 5 Comparative experiments of different algorithms

模型	mAP@0.5/%	GFLOPs	Params/M
YOLOv3-tiny	16.60	12.9	8.69
YOLOv5n	24.90	4.2	1.77
YOLOv5s	34.30	15.9	7.04
YOLOv7-tiny	34.30	13.1	6.03
YOLOv8n	34.80	8.1	3.01
Faster R-CNN ^[18]	31.00	118.8	41.2
Cascade R-CNN ^[18]	31.10	146.6	69.0
FSAF ^[19]	36.50	/	/
DSF-YOLO-P	37.30	6.6	1.99

表 5 比较结果可以看出, DSF-YOLO-P 算法在保证轻量化的情况下,算法检测精度相比主流算法达到最高。一阶段算法中, YOLOv5n 实现了

模型轻量化,但丢失了一大部分检测精度。YOLOv5s、YOLOv7-tiny 相较 YOLOv5n 检测精度有了大的提升,但相较 DSF-YOLO-P 算法仍有较大差距,模型轻量化方面也相差甚远,综合性能较差。二阶段算法中, Faster R-CNN 与 Cascade R-CNN 骨干网络提取特征图分辨率小,在小目标检测中表现偏差,而 FSAF 虽然达到了 36.5% 的精度,但仍低于 DSF-YOLO-P 算法 0.8%,并且二阶段算法计算量普遍偏大,在无人机边缘计算平台上并不适用。因此, DSF-YOLO-P 算法在提升检测精度时还考虑到了资源消耗,在无人机平台具有更大优势。

3.8 可视化结果与分析

为展示模型语义感知方面优劣,对不同尺度热力图像进行比较,如图 9 所示。从不同尺度热力图对比可知,剪枝后改进模型划分物体边界更加清晰,提取特征更为准确。对于大、中型目标,剪枝后改进模型与基线模型差距较小,只是改进模型对物体边界轮廓划分更为精确。但对于小型目标,改进热力图分布特征更加一致、稳定,能够更好拟合小尺度目标的分布空间,进一步验证了剪枝后改进算法对小型目标检测的显著优势。

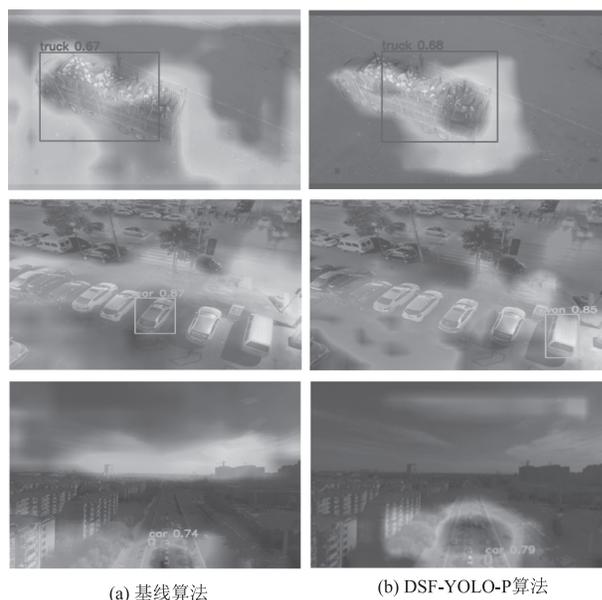


图9 不同尺度热力图对比图

Fig.9 Comparison of heat maps at different scales

4 结 语

为解决无人机航拍检测过程中算法复杂度大,小目标检测难的问题。本文基于 YOLOv8n 提出一种同时兼顾精确度以及轻量化的检测算法 DSF-YOLO-P 算法。并在 VisDrone2019 数据集与 HIT-UAV 数据集上进行 200 轮次训练。首先,在基线模型加入轻量化模块 Faster-C2f 模块,使算法既不被影响检测精度,又能减少参数量与提升 FPS;然后,针对模型检测小目标难的问题,为模型加入新预测头并重构通道,加大小目标感知度及关注度;最后,利用依赖图剪枝删除冗余通道,进一步解决模型计算成本问题。实验结果表明,DSF-YOLO-P 算法在极大降低算法复杂度的同时于可见光与红外数据集上检测精度分别提升 2.5% 与 0.6%,于无人机航拍检测过程中显著改善了模型性能,能为无人机领域的实际应用提供参考。

参考文献:

- [1] Song Jiale, Yang Dezhen, Liu Tong, et al. Downward looking infrared ship recognition algorithm for composite wing UAV platform[J]. *Laser & Infrared*, 2022, 52(11): 1649 - 1656. (in Chinese)
宋嘉乐, 杨德振, 刘彤, 等. 复合翼无人机平台的下视红外船舶识别算法[J]. *激光与红外*, 2022, 52(11): 1649 - 1656.
- [2] Yang Xue, XiuJihong, LiuXiaojia, et al. Autonomous vehicle target detection and recognition from airborne infrared imagery[J]. *Laser & Infrared*, 2023, 53(12): 1877 - 1884. (in Chinese)
杨雪, 修吉宏, 刘小嘉, 等. 航空机载红外图像的车辆目标自主检测识别[J]. *激光与红外*, 2023, 53(12): 1877 - 1884.
- [3] Souza B J, Stefenon S F, Singh G, et al. Hybrid-YOLO for classification of insulators defects in transmission lines based on UAV[J]. *International Journal of Electrical Power & Energy Systems*, 2023, 148: 108982.
- [4] Pu H, Chen X, Yang Y, et al. Tassel-YOLO: a new high-precision and real-time method for maize tassel detection and counting based on UAV aerial images[J]. *Drones*, 2023, 7(8): 492.
- [5] Shao Y, Yang Z, Li Z, et al. Aero-YOLO: an efficient vehicle and pedestrian detection algorithm based on unmanned aerial imagery[J]. *Electronics*, 2024, 13(7): 1190.
- [6] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multi-box detector[C]//Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 - 14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21 - 37.
- [7] Redmon J, Farhadi A. Yolov3: an incremental improvement[J/OL]. arXiv preprint 2018. arXiv:1804.02767.
- [8] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: optimal speed and accuracy of object detection[J/OL]. arXiv preprint 2020. arXiv:2004.10934.
- [9] Ge Z, Liu S, Wang F, et al. YOLOX: exceeding YOLO series in 2021 [J/OL]. 2021. arXiv preprint arXiv: 2107.08430.
- [10] 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, 2023: 7464 - 7475.
- [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 and Machine Intelligence*, 2016, 39(6): 1137 - 1149.
- [12] Cai Z, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154 - 6162.
- [13] Liu M, Wang X, Zhou A, et al. Uav-yolo: small object detection on unmanned aerial vehicle perspective[J]. *Sen-*

- sors, 2020, 20(8):2238.
- [14] Zhang T, Zhang Y, Xin M, et al. A light-weight network for small insulator and defect detection using UAV imaging based on improved YOLOv5 [J]. *Sensors*, 2023, 23(11):5249.
- [15] Li Y, Fan Q, Huang H, et al. A modified YOLOv8 detection network for UAV aerial image recognition [J]. *Drones*, 2023, 7(5):304.
- [16] Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:12021 – 12031.
- [17] Fang G, Ma X, Song M, et al. Depgraph: towards any structural pruning [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:16091 – 16101.
- [18] Zhang R, Shao Z, Huang X, et al. Object detection in UAV images via global density fused convolutional network [J]. *Remote Sensing*, 2020, 12(19):3140.
- [19] Zhu, C. He, Y. Savvides, M. Feature selective anchor-free module for single-shot object detection [C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16 – 20 June 2019:840 – 849.