

温秀兰¹ 焦良葆² 李子康¹ 姚波¹ 唐国寅¹

复杂环境下小尺度烟火目标检测研究

摘要

针对复杂环境下起火点目标尺寸较小、起火点特征易与实际场景混淆导致烟火检测效率及准确率低等问题,提出了一种基于改进 YOLOv5 的小尺度烟火目标检测方法.首先,在原始 YOLOv5 模型输出的第 3 个检测层上增加第 4 个检测层,以此获取更大的特征图对小目标进行检测,加强网络模型的特征提取能力.其次,为解决目标在被遮挡的场景中容易出现漏检的问题,将原网络中用于计算目标框损失函数的 GIoU_Loss 替换成 DIoU_Loss.最后,利用 TensorRT 对模型进行压缩和加速优化,并将其部署到 Jetson TX2 开发板上进行加速推理实验,通过复制增强方法扩充实际烟火场景数据.大量实验结果表明,本文所提方法用于复杂环境下的小尺度烟火目标检测不仅检测速度快而且精度高,适于推广应用.

关键词

烟火检测;改进 YOLOv5;DIoU_Loss;优化加速

中图分类号 TP391

文献标志码 A

收稿日期 2022-07-10

资助项目 国家自然科学基金(51675259);江苏省智能感知技术与装备工程研究中心开放基金(ITS202103);南京工程学院研究生科技创新基金(TB202217004)

作者简介

温秀兰,博士,教授,研究方向为精密检测、机器视觉及其应用、机器人标定与控制. zdhxwl@njit.edu.cn

¹ 南京工程学院 自动化学院,南京,211167

² 江苏省智能感知技术与装备工程研究中心,南京,211167

0 引言

火灾是威胁公共安全、危害人民生命财产的一种多发性灾害,能否快速有效地发现火情并及时预警,对维护日常生活生产秩序具有重大意义^[1].在烟火的实际检测中,由于起火点目标小,火点颜色易与车灯、路灯等颜色背景相近的目标混淆,极易误检,并且易受外部光线变化的影响,使得复杂环境下烟火检测难度大大增加^[2].目前市面上大多采用传感器进行烟火检测^[3],这种检测方法存在设备昂贵、检测范围相对较窄、延时和误报等问题.为解决传感器检测存在的弊端,有学者提出基于传统图像处理的方法^[4],主要是对烟火中易于辨识的颜色和边缘特征进行识别和检测,在一定程度上解决了不同光线下的干扰,但对于一些复杂环境背景中产生的烟火依然会受到较大的干扰,鲁棒性不强,不利于对火灾的实时检测.

随着深度学习的兴起,深度网络模型在视觉领域取得了突出进展,将深度网络模型应用于烟火检测成为近年来的研究热点^[5-9].在烟火的实际场景中进行检测时,烟火目标易受光照条件干扰,火焰的颜色特征易与一些环境背景颜色混淆,而要做到及时有效的火情预警,就必须在刚起火时检测出起火点,但这种小尺度的起火点检测难度极大,传统的基于图像处理的目标检测算法无法解决上述难点.基于深度学习的目标检测算法具有强大的特征提取能力,在室外远程拍摄和航空遥感场景中的小目标检测得到较成功应用^[10-11].

针对复杂环境下小尺度烟火目标检测难题,本文选择 YOLOv5 网络模型作为烟火目标检测与识别的基准网络,通过目标检测层设计、网络结构改进、目标框损失函数的选择等多种措施以解决小尺度烟火检测效率及准确率低等问题.

1 改进的 YOLOv5 模型

1.1 目标检测层的改进设计

在烟火检测任务中,要做到及时有效的火情预警,必须在刚起火时将起火点检测出来,然而现有的公开数据集都是以森林火灾为主,缺少小尺度的火焰数据,并且实际场景中的起火图像数据也很难收集.因此,本文首先采用一种复制增强的方法^[12-14]人为构造场景数据,并对 YOLOv5 网络模型的预测端进行重新设计,以提高模型对小目标的检测精度.

原始的YOLOv5模型中,只有3个检测层,对应3组初始化的锚框值.为了更好地检测小目标物体,在原始模型的3个检测层上增加第4个检测层.因此需要多增加一层锚框(第4组锚框)参数,增加后的锚框参数如表1所示.

表1 锚框参数

Table 1 Anchor box parameter table

组别	锚框尺寸	输出特征图尺度
第1组锚框	[10,13,16,30,33,23]	80×80
第2组锚框	[30,61,62,45,59,119]	40×40
第3组锚框	[116,90,156,198,373,326]	20×20
第4组锚框	[5,6,8,14,15,11]	160×160

当输入的图像尺寸为640×640时,网络的第1个输出层输出的特征图尺寸为(80×80),用以检测(8×8)以上的目标;网络的第2个输出层输出的特征图尺寸为(40×40),用以检测(16×16)以上的目标;网络的第3个输出层,输出的特征图尺寸为(20×20),用以检测(32×32)以上的目标.在增加了一个检测层之后,输出第4个尺度的特征图为(160×160),理论上可以检测到最小的目标为(4×4),提高了对小目标的检测精度.

1.2 改进的YOLOv5网络结构

针对烟火检测任务中起火点目标小、易与环境背景相混淆,且存在光照变化、目标遮挡和目标尺度变化等问题,本文对YOLOv5的网络结构进行了改进,改进后YOLOv5(Improved-YOLOv5)网络结构如图1所示.骨干网络由1组Focus单元、4组CBL单元和3组C3T_X组成.C3T_X就是经典的CSP结构,由多个残差组件和卷积层张量拼接组成,其中,X代表的是使用残差组件的个数,这种C3结构是整个网络中特征提取的重要环节,它不仅增强了CNN的学习能力,还极大地降低了计算瓶颈和计算的成本,使得在轻量化的同时保持较高的准确性.在Neck端采用FPN+PAN的结构,该结构不仅加强了语义信息的传递,同时还加强了特征信息的定位,因采用了C3结构,进一步加强了网络特征融合的能力.因此,改进后的网络继续对第3层输出的(80×80)的特征图进行上采样等处理,得到尺寸为(160×160)的特征图,将此特征图与骨干网络中的第1层输出的特征图进行融合操作,进行特征信息的补偿,从而提高对小目标的检测精度.

1.3 目标框损失函数的选择

模型的好坏通常采用损失函数来衡量,目标框

损失函数用来评价模型的预测框与目标框之间的一致程度.常见的损失函数有IoU_Loss,定义为

$$\text{IoU_Loss} = 1 - \text{IoU} = 1 - \frac{A}{B}, \quad (1)$$

式(1)中,A,B分别代表预测框与目标框之间的交集与并集.

YOLOv5原网络中用GIoU_Loss计算目标框损失函数:

$$\text{GIoU_Loss} = 1 - \left(\text{IoU} - \frac{D}{C} \right), \quad (2)$$

式(2)中,C为预测框与目标框的最小外接矩形,D为C与B之差.尽管GIoU_Loss解决了IoU_Loss存在的弊端,但当预测框小于目标框,位于目标框内多个不同位置时,因目标框与预测框差集相同,使得GIoU_Loss的值相同,无法区分相对应的位置关系.

因此,本文选择DIoU_Loss作为目标框损失函数:

$$\text{DIoU_Loss} = 1 - \left(\text{IoU} - \frac{R_2^2}{R_1^2} \right), \quad (3)$$

式(3)中,R₁代表最小外接矩形C的对角线距离,R₂代表目标框与预测框2个中心点之间的欧式距离.DIoU_Loss同时考虑了重叠面积和中心点的距离,因此,当目标框处于预测框内部时,通过直接度量2个框的距离,不仅使得模型训练时网络收敛快,而且当目标在被遮挡的场景中也可以将预测框快速回归出来.

2 实验及结果分析

2.1 实验数据集

数据集的丰富与否极大地影响着深度学习模型的训练效果.目前网络上还没有公开的火灾数据集,能用于研究的烟火数据图像更是匮乏,因此,通过视频抽帧、Baidu、Github以及数据复制增强等方法,建立自制的烟火数据集FDD,并用于烟火目标的检测与识别.该自制数据集包含不同场景下的烟火图像共10668张,FDD的部分图像样本如图2所示.

2.2 评价指标

在目标检测领域常用准确率(Precision)和召回率(Recall)来评价一个算法的优劣.为了更好地评价模型的准确性^[15-17],通常在评价算法性能优劣时还会引入平均准确率(Average Precision, AP)和平均准确率均值(Mean Average Precision, mAP)2个指标^[18].本文将采用Precision、Recall、AP和mAP作为

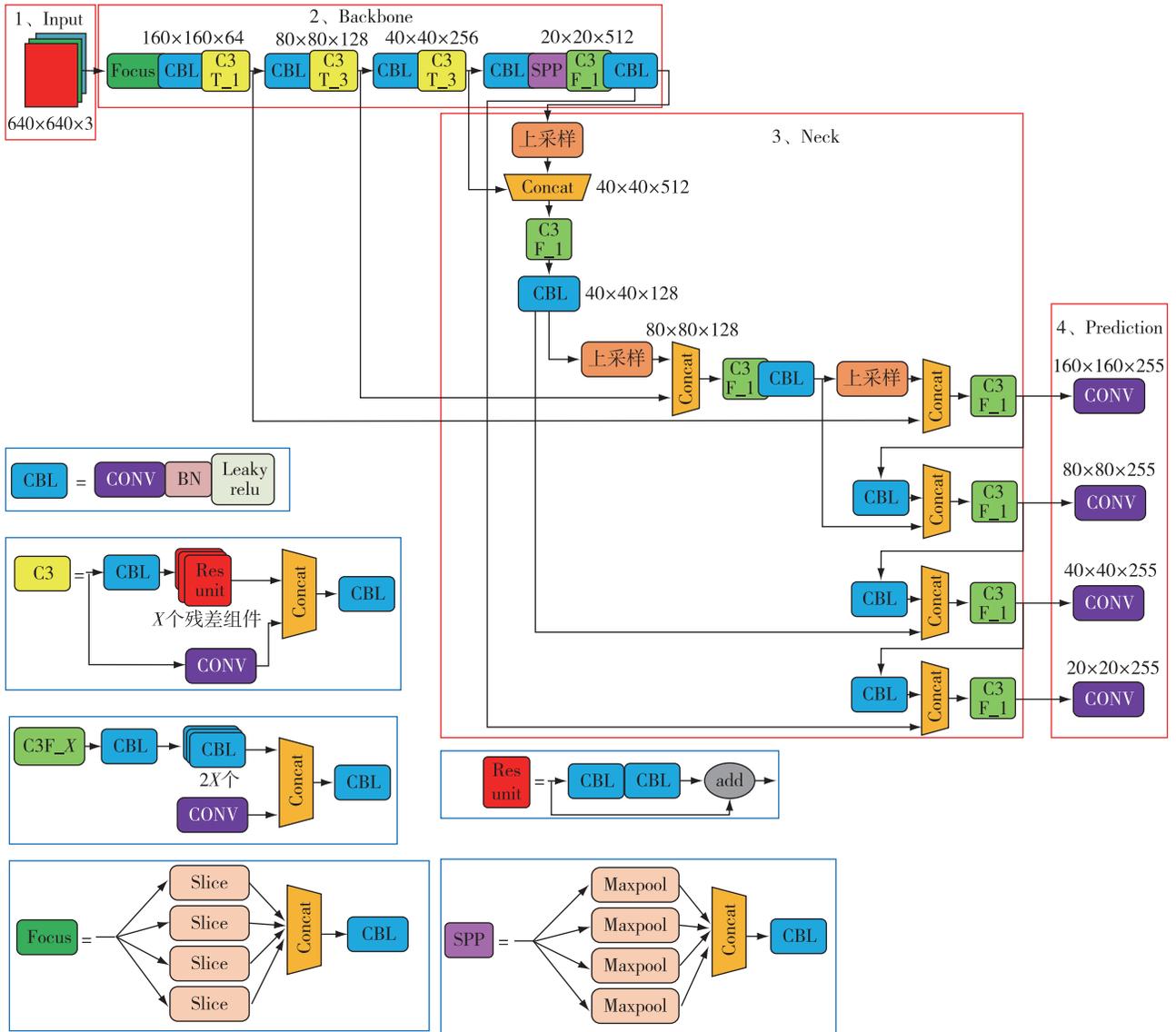


图1 改进后的网络结构
Fig. 1 Improved network structure

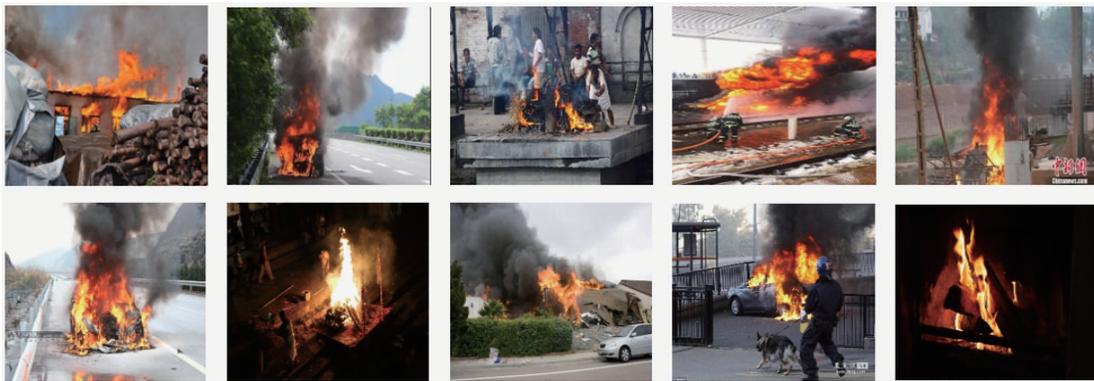


图2 FDD数据集部分图像样本示例
Fig. 2 Example of some image samples from the FDD dataset

模型的训练评估指标,并且利用每秒帧数(Frame Per Second,FPS)来衡量模型的检测速度.

由于原网络 YOLOv5 在 COCO 数据集上训练好的权重 yolov5m.pt 已经具备良好的特征信息的提取能力,利于做迁移学习,所以将该训练权重文件用于 Improved-YOLOv5 训练.训练的数据集采用自制的 FDD 数据集.对训练好的模型进行参数的微调,训练完成后烟雾和火焰的准确率结果如图 3 所示.在置信度为 0.6 的情况下,火焰的准确率为 91.4%,烟雾的准确率为 89.1%.

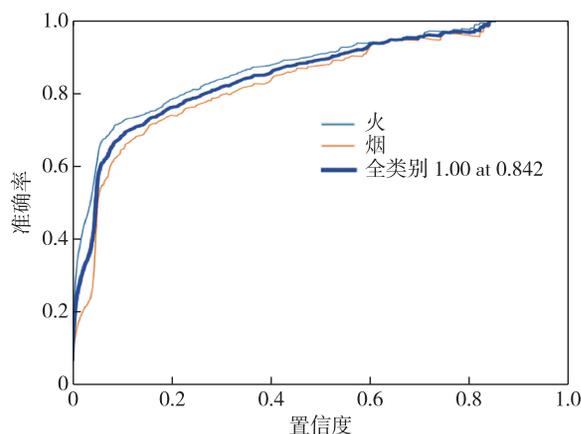


图3 准确率

Fig. 3 Precision

为了更好地评价模型的准确性,在评价算法性能优劣时还会引入 mAP 指标.mAP@0.5 是指 IoU 为 0.5 时烟雾和火焰 2 个类别的平均 AP.召回率和准确率的计算结果如图 4 所示,火焰的 AP 为 94%,烟雾的 AP 为 87.6%,mAP 可根据曲线下方的面积大小来计算.

2.3 与已有算法的对比实验

在网络模型训练过程中,可以通过 Loss 曲线观察网络训练的状态.为了验证替换后预测框损失函数 DIoU_Loss 是否可以提高网络的收敛速度,在同一个数据集上分别对改进前后损失函数的网络进行 200 个 epoch 数量的训练,其 Loss 曲线如图 5a 所示,红色曲线和蓝色曲线分别表示原 YOLOv5 网络损失曲线和替换损失函数后的网络曲线,而绿色曲线代表使用经典的 YOLOv4 算法在该数据集下的训练情况.由图 5a 可见,因 DIoU_Loss 考虑了目标框的中心点能够更快地将目标回归出来,模型的收敛速度也快于原网络,说明将 GIoU_Loss 替换成 DIoU_Loss 后提高了网络的收敛能力,损失函数值明显小于原 YOLOv5 和 YOLOv4 算法结果,表明改进后的网络模

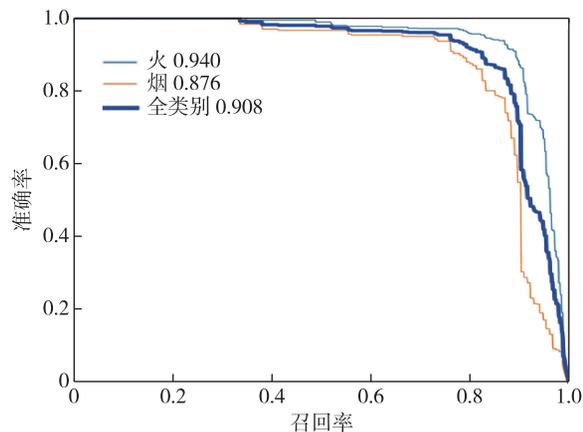


图4 平均准确率均值

Fig. 4 Mean average precision

型训练结果更佳,可以作为一个合适的网络模型进行后续的实际部署.

同时从准确率(Precision)、召回率(Recall)及 mAP 指标对改进前后网络整体的性能进行评价,结果如图 5b—d 所示.可见:相比于原 YOLOv5 网络,改进后的网络模型平均准确率均值 mAP 从 83.5% 上升到 90.8%,提升了 7.3 个百分点,但由于增加了检测层导致推理速度从 60.32 FPS 下降到 58.94 FPS;与 YOLOv4 相比,改进后 YOLOv5 网络的 mAP 提升了 14.6 个百分点.

改进前后网络模型在自制的视频序列上烟火目标检测的定性对比实验结果如图 6a—f 所示.

对比图 6a 和 6b 可以发现,原 YOLOv5 对于存在遮挡的目标图像检测存在漏检现象,而改进后的 YOLOv5 在目标被部分遮挡的情况下依然可以精准地检测出目标,说明改进后的 YOLOv5 在替换损失函数后,可以保持对图像特征提取的能力.对比图 6c 和 6d,原 YOLOv5 对于小目标的检测效果较差,不能够将 30×30 像素点以下的小目标检测出来,而改进后的 YOLOv5 依旧能很好地检测出小尺寸的着火点,说明改进后的 YOLOv5 在增加检测层之后,网络对图像特征信息的提取能力进一步加强,使得漏检问题得到有效的改善.对比图 6e 和 6f,原 YOLOv5 在特征信息相似的物体上存在误检的情况,而改进后的 YOLOv5 有效避免了误检,说明复制粘贴的数据增强方式利于网络对特征的提取和分析,减少误检率,提高了模型的泛化能力^[19-20].

3 模型加速及实时检测

为了满足实际场景中的烟火检测需求,需要对

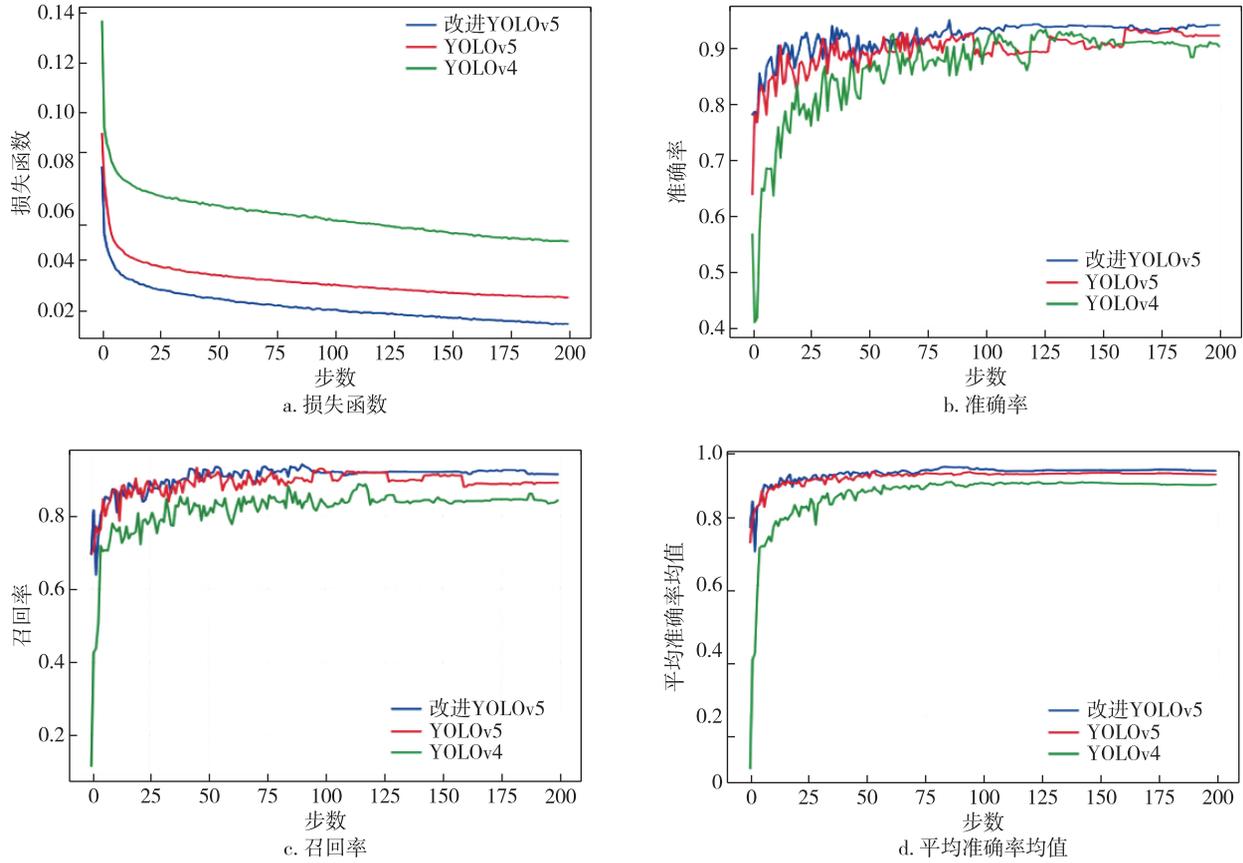


图 5 不同算法评价指标变化曲线对比

Fig. 5 Comparison chart of evaluation index changes curves of different algorithms



图 6 改进前后算法的实际检测效果

Fig. 6 The actual detection effect algorithm before and after the improvement

训练好的模型进行优化加速后部署到嵌入式设备中.TensorRT 是一个高性能的深度学习推断的优化

器和运行的引擎,其核心是一个 C++库,可以促进对 NVIDIA 图形处理单元(GPU)的高性能推理,旨在与

TensorFlow、Pytorch、MXNet 等训练框架以互补的方式工作^[20]。

3.1 TensorRT 的推理实验

由于 RTX 3080 服务器目前还不支持 FP16 的低精度推理,因此,在 Pytorch 框架和 TensorRT 框架下对改进的 YOLOv5 模型和 YOLOv5 基准模型进行 FP32 的高精度推理对比实验,其对比结果如表 2 所示。

表 2 不同框架下的推理速度对比
Table 2 Comparison of inference speed under different frameworks

模型	推理平台	时间/ms	FPS
YOLOv5	Pytorch FP32	15	75
YOLOv5	TensorRT FP32	14	81
改进 YOLOv5	Pytorch FP32	17	61
改进 YOLOv5	TensorRT FP32	15	75

如表 2 所示,在输入尺寸相同的情况下,同一种模型在不同的框架下的推理时间有所差异, TensorRT 框架下的推理速度高于 Pytorch 框架下的推理速度.虽然 TensorRT 采用了算子融合的优化方式,一些网络结构层进行了合并,减少了大量的参数,模型的推理速度也有所提升,但在推理时依旧采用 FP32 高精度推理方式,导致 TensorRT 的推理性能没有完全释放^[21]。

而 Jetson TX2 平台既支持 FP32 的高精度推理也支持 FP16 的低精度推理,因此,在 Jetson TX2 上进行 TensorRT 加速推理实验时,分别采用 FP16 的低精度和 FP32 的高精度对改进的 YOLOv5 模型进行推理对比试验,结果如表 3 所示。

表 3 TensorRT 框架下 FP16 和 FP32 的推理速度对比
Table 3 Comparison of inference speed between FP16 and FP32 in TensorRT framework

推理方式	图片尺寸	时间/ms	FPS
FP32	640×640	56	21
FP32	416×416	29	39
FP16	640×640	38	28
FP16	416×416	19	58

如表 3 所示:当输入图片尺寸为(416×416)时,改进的 YOLOv5 模型在 FP16 下的推理速度为 58 FPS,相比于 FP32 高精度推理提升了 47.7%;当输入图片尺寸为(640×640)时,改进的 YOLOv5 模型在 FP16 下的推理速度为 28 FPS,相比于 FP32 高精度

推理提升了 33.3%。可以发现,虽然 RTX 3080 服务器的算力远强于 Jetson TX2,但是当网络的输入尺寸为 416×416 时,在 Jetson TX2 开发板上执行 FP16 低精度的 TensorRT 加速推理时,其推理速度和 RTX 3080 服务器上执行 FP32 高精度的推理速度相当,这主要得益于 TensorRT 中的算子融合以及 FP16 低精度转化.因此,改进后的 YOLOv5 模型在 Jetson TX2 开发板上能够保持较高的推理速度,满足实时性的要求。

3.2 实际场景测试

将改进后的 YOLOv5 模型部署到 Jetson TX2 开发板上,同时进行 TensorRT 的加速推理,然后选择自制的一段烟火视频作为测试的视频序列,测试时常约为 20 min,最终的测试结果如图 7 所示。

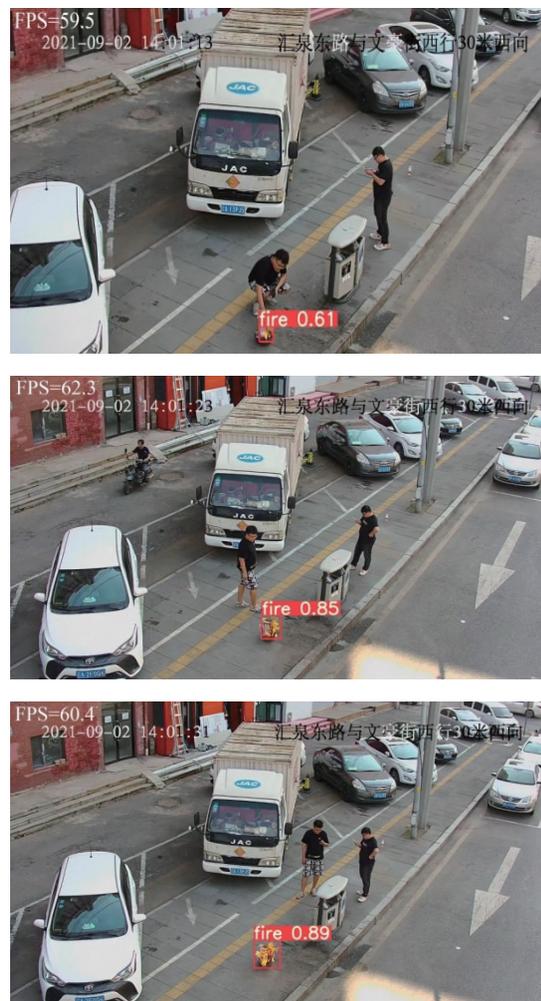


图 7 自制视频序列测试结果

Fig. 7 Test results on self-made video sequence

图 7 中 3 张结果图均为测试视频序列中的随机截取,图中左上角显示的是进行 TensorRT 加速推理

时的实时 FPS. 可以发现, 在输入图像尺寸为 416×416 时, 在 Jetson TX2 开发板上进行 TensorRT 加速推理的烟火目标检测方法能够满足实时性的要求, 其平均的帧率可以达到 60 FPS. 通过自制的烟火视频序列的测试实验, 验证了本文所提方法的有效性.

4 结论

针对小尺度烟火目标检测效率及准确率低的问题, 提出一种改进的 YOLOv5 烟火检测方法. 首先在 YOLOv5 原始的 3 个检测层上增加第 4 个检测层, 加强对小尺寸目标的检测; 然后将原网络中用于计算目标框损失函数的 GIOU_Loss 改成 DIOU_Loss, 使得目标在被遮挡的场景中也可以快速回归出预测框. 对于实际场景数据少、误检率高的情况, 采用一种复制粘贴的数据增强方式, 在丰富场景数据的同时, 大大减少了误检率, 从而得到更适用于烟火检测的模型. 最后对改进后的模型进行优化加速部署, 通过对实际场景进行测试, 结果表明改进后的 YOLOv5 模型不仅具有较高的准确率而且能够满足实时性要求, 易于在实际场景中推广应用.

参考文献

References

- [1] 李继超, 郭圣煜, 孔刘林, 等. 施工现场火焰检测和预警机器人设计及应用[J]. 中国安全科学学报, 2021, 31(4): 141-146
LI Jichao, GUO Shengyu, KONG Liulin, et al. Design and application of a fire detection and early warning robot in construction site[J]. China Safety Science Journal, 2021, 31(4): 141-146
- [2] 胡凯, 郑翡, 卢飞宇, 等. 基于深度学习的行为识别算法综述[J]. 南京信息工程大学学报(自然科学版), 2021, 13(6): 730-743
HU Kai, ZHENG Fei, LU Feiyu, et al. A survey of action recognition algorithms based on deep learning[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(6): 730-743
- [3] 张学颖, 杨遂军, 傅琳, 等. 基于组合传感器阵列的广谱火焰高度检测[J]. 仪表技术与传感器, 2013(6): 93-95
ZHANG Xueying, YANG Suijun, FU Lin, et al. Height detection of broad-spectrum flame based on combination sensor array[J]. Instrument Technique and Sensor, 2013(6): 93-95
- [4] 张驰, 孟庆浩, 井涛. 基于改进 GMM 和多特征融合的视频火焰检测算法[J]. 激光与光电子学进展, 2021, 58(4): 136-145
ZHANG Chi, MENG Qinghao, JING Tao. Video flame detection algorithm based on improved GMM and multi-feature fusion[J]. Laser & Optoelectronics Progress, 2021, 58(4): 136-145
- [5] 陈西江, 安庆, 班亚. 优化 EfficientDet 深度学习的车辆检测[J]. 南京信息工程大学学报(自然科学版), 2021, 13(6): 653-660
CHEN Xijiang, AN Qing, BAN Ya. Optimized EfficientDet deep learning model for vehicle detection[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(6): 653-660
- [6] 陈浩霖, 高尚兵, 相林, 等. FIRE-DET: 一种高效的火焰检测模型[J]. 南京信息工程大学学报(自然科学版), 2023, 15(1): 76-84
CHEN Haolin, GAO Shangbing, XIANG Lin, et al. FIRE DET: an efficient flame detection model [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(1): 76-84
- [7] 石磊, 张海刚, 杨金锋. 基于改进型 SSD 的视频烟火检测算法[J]. 计算机应用与软件, 2021, 38(12): 161-167, 173
SHI Lei, ZHANG Haigang, YANG Jinfeng. Video-based fire and smoke detection based on improved SSD [J]. Computer Applications and Software, 2021, 38(12): 161-167, 173
- [8] 李欣健, 张大胜, 孙利雷, 等. 复杂场景下基于 CNN 的轻量火焰检测方法[J]. 模式识别与人工智能, 2021, 34(5): 415-422
LI Xinjian, ZHANG Dasheng, SUN Lilei, et al. CNN-based lightweight flame detection method in complex scenes [J]. Pattern Recognition and Artificial Intelligence, 2021, 34(5): 415-422
- [9] 赵媛媛, 朱军, 谢亚坤, 等. 改进 Yolo-v3 的视频图像火焰实时检测算法[J]. 武汉大学学报·信息科学版, 2021, 46(3): 326-334
ZHAO Yuanyuan, ZHU Jun, XIE Yakun, et al. A real-time video flame detection algorithm based on improved Yolo-v3 [J]. Geomatics and Information Science of Wuhan University, 2021, 46(3): 326-334
- [10] Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 1-48
- [11] 李红光, 于若男, 丁文锐. 基于深度学习的小目标检测研究进展[J]. 航空学报, 2021, 42(7): 024691
LI Hongguang, YU Ruonan, DING Wenrui. Research development of small object tracking based on deep learning [J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(7): 024691
- [12] Hu C, Tang P, Jin W D, et al. Real-time fire detection based on deep convolutional long-recurrent networks and optical flow method [C]//2018 37th Chinese Control Conference (CCC). Wuhan, China. IEEE, 2018: 9061-9066
- [13] Leng J X, Ren Y H, Jiang W, et al. Realize your surroundings: exploiting context information for small object detection[J]. Neurocomputing, 2021, 433: 287-299
- [14] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. IEEE, 2017: 936-944

- [15] Du S J, Zhang P, Zhang B F, et al. Weak and occluded vehicle detection in complex infrared environment based on improved YOLOv4 [J]. IEEE Access, 2021, 9: 25671-25680
- [16] Zhao Z Q, Zheng P, Xu S T, et al. Object detection with deep learning: a review [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30 (11): 3212-3232
- [17] Pan S J, Yang Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359
- [18] Tao L, Hong T, Guo Y C, et al. Drone identification based on CenterNet-TensorRT [C]//2020 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting. Paris, France. IEEE, 2020: 1-5
- [19] Min J, Lee Y. An experimental view on fairness between HTTP/1.1 and HTTP/2 [C]//2019 International Conference on Information Networking (ICOIN). Kuala Lumpur, Malaysia. IEEE, 2019: 399-401
- [20] Wang X H, Yue X B, Li H Y, et al. A high-efficiency dirty-egg detection system based on YOLOv4 and TensorRT [C]//2021 International Conference on Advanced Mechatronic Systems (ICAMEchS). Tokyo, Japan. IEEE, 2021: 75-80
- [21] Zhang Q, Wang Y Y, Zhu L, et al. Research on real-time reasoning based on Jetson TX2 heterogeneous acceleration YOLOv4 [C]//2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics. Chengdu, China. IEEE, 2021: 455-459

Small scale smoke & fire target detection in complex environment

WEN Xiulan¹ JIAO Liangbao² LI Zikang¹ YAO Bo¹ TANG Guoyin¹

¹ School of Automation, Nanjing Institute of Technology, Nanjing 211167, China

² Jiangsu Intelligent Perception Technology and Equipment Engineering Research Center, Nanjing 211167, China

Abstract To address the low efficiency and accuracy of smoke & fire detection due to the small size of target and the confusion of fire feature with actual scene in complex environment, a small scale smoke & fire target detection method based on improved YOLOv5 is proposed. First, a fourth detection layer is added to the third detection layer output in the original YOLOv5 model, so as to obtain a larger feature map for small target detection and strengthen the feature extraction capability of the network model. Second, to solve the easy missing detection of target in shielded scene, DIoU_Loss is used to replace the GIoU_Loss in calculating the regression loss function of the target frame. Finally, TensorRT is used to compress and accelerate the optimization of the model, and then deployed to the Jetson TX2 development board for accelerated inference experiments. In addition, more smoke & fire scene data are constructed by replication enhancement. Experimental results show that the proposed method has fast convergence speed and high accuracy for small scale smoke & fire detection, possessing the prospect for popularization and application.

Key words smoke & fire detection; improved YOLOv5; DIoU_Loss; optimization and acceleration