

文章编号:1671-6833(2023)04-0016-06

# 改进 YOLOv5 算法在停车场火灾检测中的应用

张震<sup>1</sup>, 晋志华<sup>1</sup>, 陈可鑫<sup>2</sup>

(1. 郑州大学 电气与信息工程学院, 河南 郑州 450001; 2. 郑州大学 计算机与人工智能学院, 河南 郑州 450001)

**摘要:**针对传统传感器对于地下停车场火灾检测不及时、目标检测对小型火焰目标检测效果较差等问题,提出了一种改进的 YOLOv5 火灾检测算法。为了提高检测算法对小型火焰目标的检测效果,在 YOLOv5s 网络骨干中增加小目标检测层;为了增强火焰特征的表达,提出了一种基于 CA 注意力机制的间隔注意力结构;为了提升定位精度、降低目标漏检率,将 GIoU 替换为 CIoU。设计了 3 组消融实验以及 1 组对比实验用来验证所提算法的有效性。实验结果表明:所提算法在自定义数据集上的  $mAP_{0.5}$ 、召回率  $R$  分别为 92%、96.9%。与 YOLOv5s 模型相比,所提算法在自定义火焰数据集上的  $mAP_{0.5}$  提升了 1.8 个百分点, $R$  提升了 2.0 个百分点。所提算法权重大小仅为 16.4 MB,帧率能达到 113 帧/s,具有较小的模型体积以及较快的检测速度,且能够准确检出小型火焰目标,有效提升了地下停车场火灾防范能力。

**关键词:**地下停车场;火灾检测;YOLOv5;坐标注意力;CIoU 损失函数

**中图分类号:**TP391

**文献标志码:**A

**doi:**10.13705/j.issn.1671-6833.2023.04.001

近年来,随着中国经济的高度发展,居民汽车拥有量不断增加,地下车库成为汽车的主要停泊地。然而,由于电气短路、油箱漏油等原因,车辆在地下车库停放时容易发生火灾。停车场车辆密度大、火焰燃烧迅速、场所较为封闭、施救难度大,一旦发生火灾,经济损失巨大<sup>[1]</sup>。目前,大部分的公共场所都在使用传统的光电感烟探测器。由于地下停车场空间大,并配备一定的通风设备,使得传统的烟雾火灾检测装置效果大大降低,且光电感烟探测器工作需要一定的颗粒浓度,所以实时性较差。车库一般都安装有监控摄像头,因此将基于视频的目标检测用于地下停车场的火灾检测具有重要的研究意义。

当前目标检测算法主要分为由 R-CNN 代表的二阶段(two-stage)目标检测算法以及 YOLO(you only look once)代表的一阶段(one-stage)目标检测算法<sup>[2-3]</sup>。主流的二阶段目标检测算法有 R-CNN<sup>[4]</sup>、Fast R-CNN<sup>[5]</sup>等,这类算法首先生成一系列作为样本的候选框,再通过卷积神经网络进行样本分类。代表性的一阶段目标检测算法有 SSD<sup>[6]</sup>、YOLO 等。以 YOLO 为代表的一阶段目标检测算法

仅需处理一次图像就可以得到目标的位置以及分类结果,相较于二阶段目标检测算法,能够有效提升目标检测的速度,因此更适用于实时性要求高的地下停车场火灾检测问题。

YOLOv5<sup>[7]</sup>在目标检测领域有着广泛的应用。考虑到地下停车场空间较大、发生火灾初期火焰目标较小,原始的 YOLOv5s 模型在此条件下检测效果较差,本文通过增加小目标检测层、增加注意力机制以及修改损失函数来提升对小型火焰目标的检测效果。

## 1 YOLOv5 算法介绍

YOLOv5 网络结构分为输入端 Input、骨干网络 Backbone、颈部 Neck 和预测部分 Prediction。算法一共包含 4 个模型,按照模型由小到大分别是 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。其中,YOLOv5s 的模型最小,相同实验环境下检测速度最快、易于部署,且拥有较高的检测精度,能够满足对火灾火焰实时检测的要求。

YOLOv5 模型在数据输入端 Input 采用 Mosaic 数据增强算法对数据集进行训练前处理。随机读取

收稿日期:2022-10-07;修订日期:2022-11-05

基金项目:国家重点研发计划项目(2018\*\*\*\*\*);河南省重大公益专项(201300311200)

作者简介:张震(1966—),男,河南郑州人,郑州大学教授,博士,博士生导师,主要从事多媒体信息安全、图像处理、模式识别的研究,E-mail:zhangzhen66@126.com。

引用本文:张震,晋志华,陈可鑫.改进 YOLOv5 算法在停车场火灾检测中的应用[J].郑州大学学报(工学版),2023,44(4):16-21.(ZHANG Z, JIN Z H, CHEN K X. Application of improved YOLOv5 algorithm in parking lot fire detection[J]. Journal of Zhengzhou University (Engineering Science), 2023, 44(4): 16-21.)

数据集中的 4 张原始图片,对图片进行随机缩放、翻转、改变色域等操作,随后将处理后的图片进行拼接,得到一个拥有更高复杂度的样本。通过对数据集多次处理,可以大大增加数据集的复杂度,从而提升模型的泛化能力。

YOLOv5 模型的骨干网络 Backbone 采用了最新的 Focus 结构。输入矩阵经过间隔采样操作变为 4 个具有原矩阵信息的低维度矩阵,然后将 4 个矩阵在同一维度上进行堆叠,使得堆叠后的新矩阵能够得到更明显的原始矩阵信息,大大提高了后续的特征提取的效率。YOLOv5 在 CSP 模块<sup>[8]</sup>中加入 SPP 模块,有效地增加了主干特征的接收范围,即提高了感受野,显著地分离了最重要的上下文特征,从而提升了模型精度。

## 2 检测算法优化

本文的主要工作主要体现在以下 3 个方面:在原模型的基础上增加一个小目标检测层、加入改进坐标注意力结构以及使用 CIoU 替换 GIoU。

### 2.1 增加小目标检测层

针对火灾初期火源目标较小、原模型漏检的问题,在 YOLOv5s 模型的基础上增加一个小目标检测层,通过增加模型对细粒特征的分辨能力,从而增强模型对小目标的检测效果,虽然增加了计算成本,但是能够很好地检出小型火焰目标。

增加小目标检测层网络如图 1 所示,与原始 YOLOv5s 网络相比,改进后的网络从骨干网络的第 2 层就开始特征增强、提取浅层特征;将提取到的特征图与深层特征进行融合,继续进行上采样处理,特征图继续扩大,与骨干网络提取出的浅层特征进行融合,得到具有更多特征信息的特征图,并将富含细粒信息的特征图向下层传递。由此,模型能够提升对细粒信息的敏感度,从而提高对小目标的检测性能。

如图 2 所示,不同的特征层对小目标的感知是不同的,通过增加浅层的检测层,以及提供更小的 Anchor,使模型可以更好地关注小型目标的信息,能够提升模型对于小型火焰目标的检测能力。

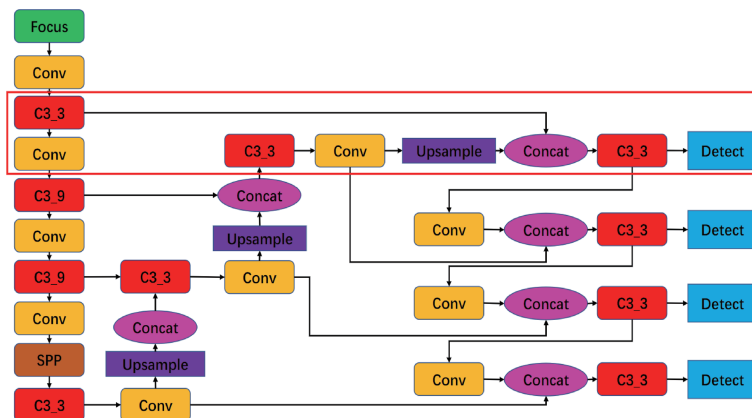


图 1 YOLOv5s 增加检测层结构图

Figure 1 YOLOv5s added detection layer structure diagram

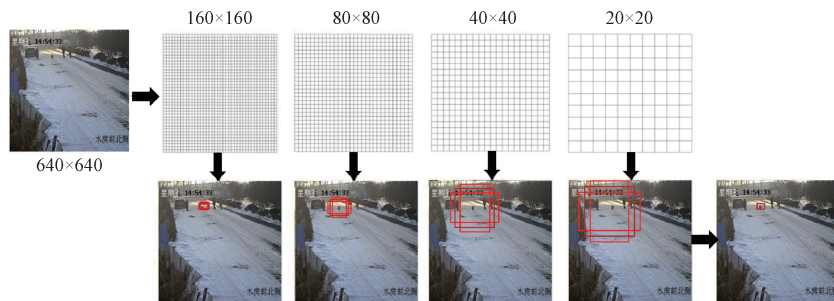


图 2 不同特征层检测不同大小目标示意图

Figure 2 Schematic diagram of objects of different sizes detected by different feature layers

### 2.2 改进的坐标注意力结构

坐标注意力机制<sup>[9]</sup>(coordinate attention, CA)在 SE 注意力机制<sup>[10]</sup>的基础上进行了针对性的改进,保留了图像的特征位置信息。相较于 SE 注意力模块,该模块不仅能获取到空间方向上的长程依赖,还

能增强特征的位置信息表达,同时增大网络的全局感受野。

如图 3 所示,通过将输入特征分别进行 X 轴与 Y 轴方向上的一维自适应平均池化,从而分别得到保留有 X 轴与 Y 轴信息的独立方向感知特征,其中

一个空间方向捕获长程依赖,另一个空间方向保留精准的位置信息。对得到的两个一维特征在  $W$  维度上进行拼接,再经过一个卷积以及非线性激活函数,紧接着将特征在通道维度进行拆分,通过卷积以及 Sigmoid 激活函数得到两个带有特定空间方向长程依赖的特征图,这两个特征图可以互补地应用到输入特征图来增强感兴趣的目标。通过与原特征进行特征融合,最终得到在宽度与高度方向上带有注意力权重的特征图。

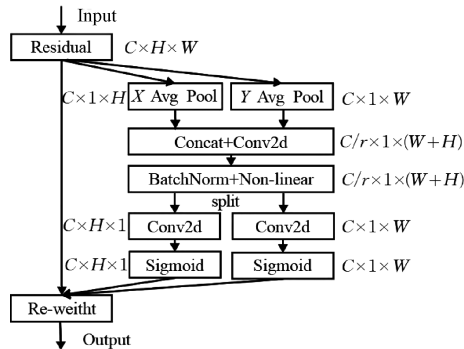


图 3 坐标注意力机制网络图  
Figure 3 network diagram of coordinate attention mechanism

目前,常见的 Backbone 注意力结构如图 4(a)、4(b)所示。其中,C3CA 结构是将注意力模块嵌入 YOLOv5 的 C3 模块中,即对每个特征层提取的特征均进行注意力机制的处理;CA 结构是只对深层次的特征图进行特征的增强。

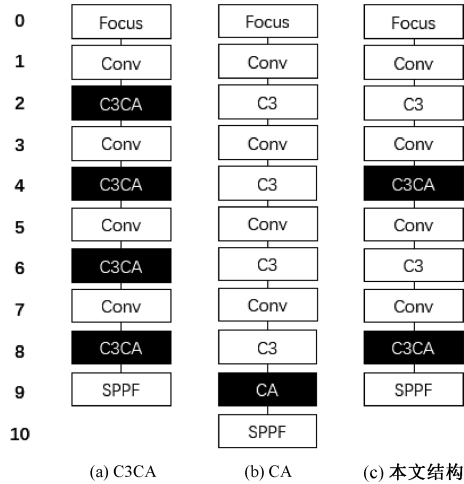


图 4 C3CA、CA 结构与本文结构对比  
Figure 4 Comparison among C3CA, CA structure and this structure

本文提出一种新的间隔注意力结构,如图 4(c)所示。图像经过一次提取得到的特征图经过注意力机制,得到一个小感受野的、具有独立方向感知特征的特征图;该带有位置信息的特征图继续进行一次特征提取,然后再一次通过坐标注意力机制,得到一

个具有大感受野并具有独立方向感知的特征图。相较于常见的 C3CA、CA 结构,本文使用的间隔注意力结构在 Backbone 中间隔使用了两次注意力模块,将第 2、4 个 C3 层替换为具有注意力机制的 C3CA 层。本文结构能够让模型更加关注中小目标的特征,有助于对深层次目标的特征强化,使模型能够更精准地提取目标特征,从而增强目标检测的精度。

相比于 C3CA 结构,本文结构减少了两个注意力模块,模型的复杂度大大降低,使网络模型更加精简,有助于提升模型的检测速度;相比于 CA 结构,本文结构增加了在中间特征层的注意力模块,通过注意力模块提取的带有位置信息的特征经 FPN 结构融合,得到更多的特征语义信息,能够进一步提升模型的检测效果。

2.3 CIOU 损失函数

YOLOv5 中采用 GIoU Loss 作为 Bounding box 的损失函数。如果预测框与真实框出现非相交的情况,此时梯度将变为 0,神经网络将无法优化,预测框与真实框的相对位置无法区分。考虑到上述情况,本文使用 Complete IoU (CIOU) Loss<sup>[11]</sup>代替 GIoU Loss。CIOU 损失在考虑 GIoU 损失的基础上,考虑了 BBox 的重叠面积、中心点距离以及 BBox 长宽比的一致性,使模型得到更好的回归效果。

3 实验

3.1 创建自定义的地下停车场火灾数据集

当前并没有公开的地下停车场火灾的数据集,本文选择自定义数据集。其中图像来自于包括 CVPR 实验室采集的 KMU Fire and Smoke database 火焰图像<sup>[12]</sup>、停车场监控模拟火灾视频、室内火焰模拟视频以及网络视频等。为了更好地对小目标火焰进行检测,更快对地下停车场火灾做出反应,通过模拟实验采集一部分小型火焰图像数据。另外,地下停车场光线较差,普通监控摄像机容易过曝,本文对这种情况也进行了模拟。数据集包含夜晚、灯光、日光等多种场景,火焰目标大小不同,共计 3 420 张图片。本文通过 labelImg 标注工具对数据集图像进行标注,并保存为 YOLO 系列的 txt 格式。通过对数据进行乱序排列,随机选出训练集 2 394 张,测试集共 513 张,验证集共 513 张。数据集示例如图 5 所示。

训练集中所有标签框的大小分布如图 6 所示,其中横纵坐标分别代表标签框的宽度和高度。图中深蓝色表示尺寸在该区域的标签图较多,更符合火灾初期火焰目标较小的情况,所以该数据集适用于



图 5 数据集示例图

Figure 5 Dataset example graph

本文研究的火焰目标检测的问题。

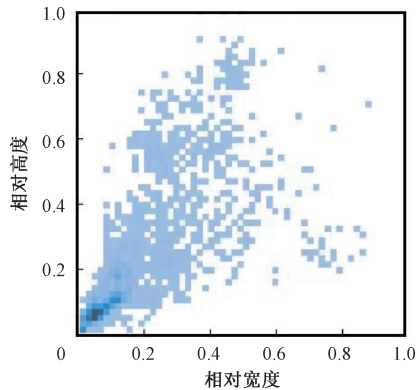


图 6 标签框大小分布图

Figure 6 Label box size distribution

### 3.2 实验环境与配置

本文通过搭建 PyTorch 深度学习框架对自定义数据集进行训练和测试。操作系统选用 Windos10 专业版 21H2, CPU 为 Ryzen 5800X, GPU 为 GeForceRTX3060Ti,显存为 8 GB,内存为 32 GB,CUDA 版本号为 11.1.1,PyTorch 版本号为 1.8.0。

实验中训练参数设置如下:输入图像尺寸为像素 640×640,batch size 为 16,epoch 设置为 100。

### 3.3 实验结果的评价指标

本实验中采用  $mAP_{0.5}$ 、参数量(parameters)、复杂度、精确率  $P$ (precision)、召回率  $R$ (recall) 和帧率等指标<sup>[13]</sup>作为模型性能的评价指标。其中复杂度用 GFLOPs 衡量,1 GFLOPs =  $10^9$  FLOPs, FLOPs (floating point operations) 为浮点运算数,可以衡量模型的复杂度。

### 3.4 实验结果和分析

在训练阶段,选用 Ultralytics6.0 版本的 YOLOv5s 模型作为基准模型,使用上述数据集对 YOLOv5s 模型与本文算法进行重新训练,以下所示结果均为采用同一数据集得出。

为了提升对火灾初期小的火焰目标的检测效果,加入小目标检测层模型 A 与 YOLOv5s 对比如表 1 所示。加入小目标检测层后  $mAP_{0.5}$  提升 0.9 百分点, $R$  提升 1.2 百分点,由于增加了网络的深度,导致模型复杂度上升,帧率有所下降。

表 1 消融实验 1 数据表				
Table 1 Data sheet of ablation experiment 1				
模型	$mAP_{0.5}/\%$	$R/\%$	复杂度/GFLOPs	帧率/(帧·s <sup>-1</sup> )
YOLOv5s	90.2	94.9	15.8	138
模型 A	91.1	96.1	26.3	116

加入小目标检测层后,置信度损失有较大的下降,置信度损失曲线如图 7 所示。随着迭代次数的增加,本文算法的置信度损失曲线明显收敛更快、损失值更小。

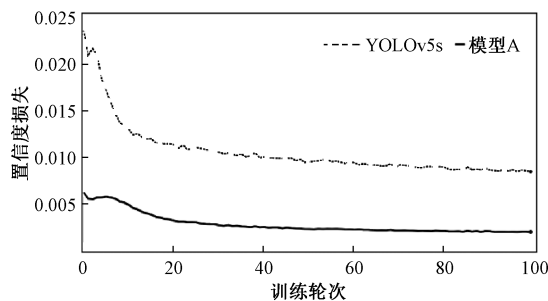


图 7 置信度损失曲线

Figure 7 Confidence loss curve

通过以上实验,说明采用小目标检测层虽然帧率有所降低,但是对火焰目标的置信度误差大大降低,对于小的火焰目标的检测置信度和对网络的检测精度有所提升。因此,通过增加 FPN+PAN 的网络深度,以及增加小目标检测层,能够有效提升网络对小目标的检测效果。

为了加强检测网络对火焰目标的特征提取能力,在网络加入改进的注意力模块,并与主流的 SE 注意力机制、CBAM 注意力机制、原 CA 注意力机制进行对比,以验证本文改进结构的性能,结果如表 2 所示。其中,本文模型 B 的  $mAP_{0.5}$  最高,相较于 YOLOv5s 网络, $mAP_{0.5}$  有 1.8 百分点的提升, $R$  提升 1.4 百分点;与主流的注意力机制 SE、CBAM 机制相对比,模型 B 检测精度及检测速度均最高;与 C3CA、CA 结构相对比,模型 B 的  $mAP_{0.5}$  分别提升 0.9 百分点、1.8 百分点, $R$  分别提升 2.8 百分点、0.5 百分点。

为了验证 CIoU 替换 YOLOv5s 中 GIoU 后的效果,实验 3 对比了 YOLOv5s 替换目标框损失函数前后的表现,如表 3 所示。使用 CIoU 损失函数后, $mAP_{0.5}$  有 0.6 百分点的提升, $R$  提升 2.3 百分点,可



见 CIoU 能够显著提升模型的召回率。并且可以优化模型的检测速度。

表 2 消融实验 2 数据表

Table 2 Data sheet of ablation experiment 2

模型	$mAP_{0.5}/\%$	$R/\%$	复杂度/ GFLOPs	帧率/ (帧·s <sup>-1</sup> )
YOLOv5s	90.2	94.9	15.8	138
YOLOv5s+SE	90.5	93.5	15.9	133
YOLOv5s+CBAM	91.3	96.1	19.7	123
YOLOv5s+C3CA	91.1	93.5	17.5	91.7
YOLOv5s+CA	90.2	95.8	16.7	130
模型 B	92.0	96.3	16.9	118

表 3 消融实验 3 数据表

Table 3 Data sheet of ablation experiment 3

模型	$mAP_{0.5}/\%$	$R/\%$	复杂度/GFLOPs	帧率/(帧·s <sup>-1</sup> )
YOLOv5s	90.2	94.9	15.8	138
模型 C	90.8	97.2	15.8	139

对各个模块的有效性 & 性能进行验证后,与目前主流的目标检测算法 YOLOv3、YOLOv5s 以及 YOLOv5m 进行对比,对比结果如表 4 所示。

表 4 主流目标检测算法对比

Table 4 Comparison of mainstream target detection algorithms

模型	$mAP_{0.5}/\%$	$R/\%$	权重/MB	帧率(帧·s <sup>-1</sup> )
YOLOv3	83.7	85.5	234.1	83
YOLOv5s	90.2	94.9	13.7	138
YOLOv5m	92.1	96.5	41.5	92
本文算法	92.0	96.9	16.4	113

实验表明,本文算法在地下停车场火焰目标检测任务上可达到  $mAP_{0.5}$  为 92.0% 的检测精度。相较于未改进前的 YOLOv5s 算法,本文算法的  $mAP_{0.5}$  有 1.8 个百分点的提升, $R$  有 2.0 个百分点的提升,虽然牺牲了一部分的检测速度,但是得到了具有更好检测精度的模型。相较于 YOLOv5m 算法,虽然两种算法在精度上相当,但是本文算法具有较高的召回率  $R$  以及较高的帧率,具有较快的检测速度,且本文算法权重仅为 16.4 MB,远远小于 YOLOv5m 的 41.5 MB,对检测设备的配置要求更低。

为了更好地验证本文算法对于小型火焰目标的检测性能,从测试集中选取部分图像数据进行测试,如图 8 所示。在图 8(a)中,本文算法在检测精度上明显高于原始模型。在图 8(b)中,YOLOv5s 出现了漏检,无法可靠地检测出目标,而本文算法能够准确地检测到火焰目标,主要是因为本文模型增加了小目标的检测层,从而具有更强的分辨力,能够更准确地检测到小型火焰目标。在图 8(c)中,YOLOv5s 模型没有正确框选出火焰目标,而本文算法能够准确

地框选出火焰目标,这归结于坐标注意力机制带来了较大的感受野,增加了模型对火焰的感知能力,并且 CIoU 损失函数能够提升预测时的定位精度。

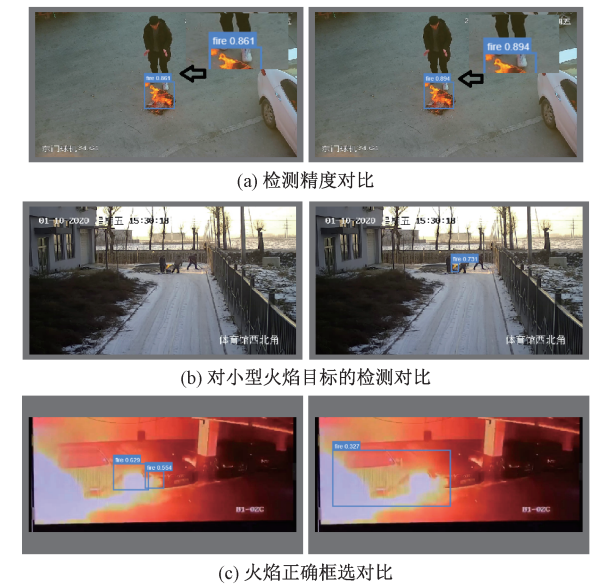


图 8 YOLOv5s(左)与本文算法(右)火焰检测对比  
Figure 8 Comparison between YOLOv5s (left) and the proposed algorithm (right) for flame detection

4 结论

目前地下停车场缺乏成熟的视频火灾检测算法,已有的技术依赖传感器等设备,且无法针对火灾初期小型火焰目标做出快速反应,本文提出一种改进的 YOLOv5 火灾检测算法。

为增强对小型火焰目标的检测性能,增加小目标检测层,提升了检测网络对小型火焰目标的检测效果;为提高模型对火焰特征的提取以及获得更大的感受野,提出了一种新的间隔注意力结构,进一步提升了网络对火焰目标的特征提取能力,且相较于 C3CA、CA 结构,本文改进模型具有复杂度较低、精度高的特点;为了提升定位精度、降低目标漏检率,使用 CIoU 损失函数,提升了网络的回归性能。在实验设备不变的情况下,本文算法与原 YOLOv5s 相比, $mAP_{0.5}$  提升了 1.8 百分点,召回率  $R$  提升了 2.0 百分点,本文算法权重大小仅为 16.4 MB,帧率在本文实验环境下能达到 113 帧/s。

实验表明,本文算法对火焰目标较小的情况也能够正确检出,具有较高的精确率和召回率、较小的模型体积,易于部署,基本能够满足地下停车场对火灾实时检测的要求。

参考文献:

[1] 盛原. 汽车及停车场库火灾起因、特点及消防对策

- [J]. 交通与运输, 1997, 13(4): 21.
- SHENG Y. Causes, characteristics and fire control countermeasures of automobile and parking garage fires [J]. Traffic & Transportation, 1997, 13(4): 21.
- [2] 周晓彦, 王珂, 李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术, 2017, 40(11): 89-93.
- ZHOU X Y, WANG K, LI L Y. Review of object detection based on deep learning[J]. Electronic Measurement Technology, 2017, 40(11): 89-93.
- [3] 李柯泉, 陈燕, 刘佳晨, 等. 基于深度学习的目标检测算法综述[J]. 计算机工程, 2022, 48(7): 1-12.
- LI K Q, CHEN Y, LIU J C, et al. Survey of deep learning-based object detection algorithms[J]. Computer Engineering, 2022, 48(7): 1-12.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.
- [5] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440-1448.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[EB/OL]. (2015-12-08) [2022-10-01]. <https://arxiv.org/abs/1512.02325>.
- [7] Ultralytics. YOLOv5[EB/OL]. [2022-10-01]. <https://github.com/ultralytics/yolov5>.
- [8] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSP-Net: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2020: 1571-1580.
- [9] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 13708-13717.
- [10] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [11] ZHENG Z H, 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.
- [12] 李浩方. 电力重点区域火灾监测技术与系统开发[D]. 郑州: 郑州大学, 2021.
- LI H F. Research on fire monitoring technology and system development in key power areas[D]. Zhengzhou: Zhengzhou University, 2021.
- [13] 贾云飞, 郑红木, 刘闪亮. 基于YOLOv5s的金属制品表面缺陷的轻量化算法研究[J]. 郑州大学学报(工学版), 2022, 43(5): 31-38.
- JIA Y F, ZHENG H M, LIU S L. Lightweight surface defect detection method of metal products based on YOLOv5s[J]. Journal of Zhengzhou University (Engineering Science), 2022, 43(5): 31-38.

## Application of Improved YOLOv5 Algorithm in Parking Lot Fire Detection

ZHANG Zhen<sup>1</sup>, JIN Zhihua<sup>1</sup>, CHEN Kexin<sup>2</sup>

(1. School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China; 2. School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China)

**Abstract:** Aiming to solve the problem that the traditional sensors could not timely in detecting the fire of underground parking lot and the object detection was not effective in detecting small flame targets, an improved YOLOv5 fire detection algorithm was proposed. In order to improve the detection effect of the detection algorithm on small flame targets, small target detection layer was added to YOLOv5s network backbone. In order to enhance the expression of flame features, an interval attention structure based on CA attention mechanism was proposed. In order to improve the positioning accuracy and reduce the rate of missed target detection, GIoU was replaced by CIoU. Three groups of ablation experiments and one group of contrast experiment were designed to verify the effectiveness of this improvement. The experimental results showed that the  $mAP_{0.5}$  and  $R$  of the algorithm on the user-defined dataset were 92% and 96.9%, respectively. Compared with YOLOv5s model, the  $mAP_{0.5}$  of the proposed algorithm on the customized flame data set was increased by 1.8 percentage, and  $R$  was increased by 2.0 percentage. The weight size of the proposed algorithm was only 16.4 MB, and the frame rate could reach 113 frames per second. It had smaller model volume and faster detection speed, and could accurately detect small flame targets, could effectively improve the fire prevention capability of the underground parking lot.

**Keywords:** underground parking; fire detection; YOLOv5; coordinate attention; CIoU loss function