

基于改进YOLOv5s的火焰烟雾检测方法

孙 剑*, 张数数

(信阳师范大学 计算机与信息技术学院, 河南 信阳 464000)

摘要: 针对传统火焰烟雾检测不及时、小目标烟火检测困难和检测精度低等问题,提出了一种改进YOLOv5s的火焰烟雾检测方法。首先,在YOLOv5s模型的Backbone层中引入了SE注意力机制,自适应地调整每个通道的特征权重,增强有用特征并抑制无用特征,从而提升网络对火焰烟雾特征的提取能力。其次,在YOLOv5s模型的Neck层中引入了BiFPN模块作为特征融合模块,通过BiFPN模块引入双向连接,结合自底向上和自顶向下的特征融合方式,能够充分利用不同层级的特征信息,提高特征的丰富性。最后,将改进YOLOv5s模型应用于实际火焰烟雾数据集上。实验结果表明:改进YOLOv5s模型的准确率、召回率和mAP值分别提升了1.8%、2.6%和1.5%,能够满足火焰烟雾检测的精度要求。

关键词: 火焰烟雾检测; YOLOv5s模型; SE注意力机制; BiFPN模块

中图分类号: TP391.4

文献标志码: A

开放科学(资源服务)标识码(OSID):



Flame smoke detection method based on the improved YOLOv5s

SUN Jian*, ZHANG Shushu

(College of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China)

Abstract: To solve the problems of untimely traditional flame smoke detection, difficult pyrotechnic detection and low detection accuracy of small targets, an improved flame smoke detection method of YOLOv5s was proposed. Firstly, the SE attention mechanism was introduced into the backbone layer of the YOLOv5s model, which can adaptively adjust the feature weight of each channel, enhance the useful features and suppresses the useless features, improve the network's ability to extract the features of flame and smoke. Secondly, BiFPN module was introduced as a feature fusion module in the Neck layer of the YOLOv5s model, and the bidirectional connection was introduced through BiFPN module, which can make full use of the feature information of different levels and improve the richness of features by combining the bottom-up and top-down feature fusion methods. Finally, the improved YOLOv5s model was applied to the actual flame smoke dataset, and the experimental results showed that the accuracy, recall rate and mAP value of the improved YOLOv5s model were increased by 1.8%, 2.6% and 1.5%, respectively, which can meet the accuracy requirements of flame smoke detection.

Key words: flame smoke detection; YOLOv5s model; SE attention mechanism; BiFPN module

0 引言

虽然人们对于火灾的防范意识在不断提高,但是近几年来火灾的发生数量依旧居高不下,仅在2023年全国接报处置各类警情213.8万起,直接

财产损失达39.4亿元^[1]。如何准确、快速地检测火焰和烟雾,提高火灾预警系统的性能,以便能够在第一时间采取措施,把火灾控制和消灭在初期阶段,成为当前亟待解决的问题^[2]。

随着计算机技术的更新和发展,火焰烟雾的

收稿日期:2024-05-06;修回日期:2024-07-04;*.通信联系人,E-mail:sunjian@xynu.edu.cn

基金项目:国家自然科学基金青年基金项目(62403405);河南省科技攻关项目(222102210300)

作者简介:孙剑(1989—),男,河南信阳人,讲师,博士,硕士生导师,主要从事机器学习与数据挖掘研究。

引用格式:孙剑,张数数.基于改进YOLOv5s的火焰烟雾检测方法[J].信阳师范大学学报(自然科学版),2025,38(2):152-158.

SUN Jian, ZHANG Shushu. Flame smoke detection method based on the improved YOLOv5s[J]. Journal of Xinyang Normal University (Natural Science Edition), 2025, 38(2):152-158.

检测技术也在不断发展和完善,主要可分为基于图像视觉技术传统的火焰烟雾检测和基于深度学习的火焰烟雾检测两个方向^[3]。

基于图像视觉技术传统的火焰烟雾检测主要依赖于图像处理和计算机视觉技术,通过分析视频或图像中的特征来检测火焰和烟雾。CELIK^[4]提出一种基于计算机视觉的火灾探测算法,由火灾颜色建模和运动侦测两部分组成,实现火灾探测。GOMES等^[5]提出了一种基于颜色的火灾外观模型和一种基于小波的火灾频率特征模型,实现了固定监控智能摄像机火灾探测方法。杨雨卓^[6]采用多判断依据联合的方法,判断帧间图像差距,弥补单一特征的缺陷;并基于图像的静态和动态特征,实现火焰与烟雾的检测与识别。然而,基于图像视觉技术传统的火焰烟雾检测存在容易受到环境光照、阴影、颜色变化等因素干扰的问题,只适用于计算资源有限下的火焰烟雾检测,不能满足复杂应用场景的需求^[7]。

基于深度学习的火焰检测方法是使用深度神经网络模型进行训练,使其能够学习到火焰烟雾的深层次特征,这种方法具有更高的准确性和鲁棒性。CHAOXIA等^[8]提出一种基于改进Faster R-CNN模型的全局信息引导火焰检测方法,以引导方式执行火灾探测过程。颜洵等^[9]提出了基于改进YOLOv4网络模型的火焰和烟雾的目标检测方法,利用K-means算法计算火焰和烟雾的anchor值。RYU等^[10]提出一种基于计算机视觉检测和卷积神经网络的复杂火焰和烟雾检测方法,使用基于深度学习的卷积神经网络进行推理,以确定它是火焰还是烟雾。然而这种方法需要较高的计算资源和大量的训练数据,还常常受到光照条件、背景环境、燃烧物质等多种因素的影响,实际中往往难以准确捕捉到这些变化。

针对上述问题,本文提出了一种基于改进YOLOv5s模型的火焰烟雾检测方法。在YOLOv5s模型主干提取网络中,引入SE(Squeeze-and-Excitation)注意力机制^[11],通过对特征图进行通道级别的加权,使得模型在训练过程中能够自动学习到不同通道的重要性,提高模型的特征提取能力。在颈部结构中,引入BiFPN(Bidirectional Feature Pyramid Network)模块^[12],充分利用不同尺度的特征信息,提高模型对不同形态火焰烟雾的检测精度。基于实际火焰烟雾数据集验证了文

中所提出方法的有效性。

1 YOLOv5s网络模型

YOLO模型^[13]由REDMON等2016年提出,在目标检测领域得到了广泛的应用。2020年,YOLOv5模型^[14]由Ultralytics公司在GitHub上发布实现。该网络主要由输入层、特征提取网络(Backbone)、特征融合部分(Neck)以及检测头(Head)组成。输入层负责接收图像数据以及进行一些必要的预处理操作。

Backbone层负责从输入图像中提取高级特征,将图像中的物体信息通过卷积神经网络进行提取,形成特征金字塔,用于之后的特征融合操作,主要包括Conv模块、C3模块、SPPF模块。Conv模块由1个Conv2d、1个BatchNorm2d和SiLU激活函数构成。Conv2d主要作用是进行卷积操作,BatchNorm2d为批归一化层,对每批的数据做归一化,而SiLU激活函数是Sigmoid函数和ReLU函数的改进版。C3模块则是由3个CBS模块和1个BottleNeck模块组成。SPPF模块使用了1个CBS和多个相串联的最大池化,然后进行全连接,最后再经过一个CBS来实现局部特征和全局特征的融合。

Neck层由特征金字塔(Feature Pyramid Network, FPN)^[15]与路径聚合网络(Path Aggregation Network, PAN)^[16]两部分组成,是一段自底向上和自顶向下的链路。FPN结构通过进行上采样,向特征图中插值,使特征图的尺寸变大,使底层特征图包含更强的图像强语义信息,以便融合来自Backbone层的特征图。PAN结构通过进行下采样,传达位置特征信息。通过FPN+PAN的形式,使不同尺寸的特征图都包含更强的图像语义信息和图像定位信息,增强了网络结构的鲁棒性,加强了物体检测能力,并且将这些特征传递给Head层进行预测。

Head层也称为检测模块,主体部分是由3个 1×1 卷积构成。经过上述特征融合过程后,最后会得到3个加强特征,Head会根据加强后的特征进行最终的回归预测。YOLOv5s模型网络结构如图1所示。

YOLOv5s算法的优点在于其较高的速度和实时性,特别适用于视频监控、自动驾驶等对实时性要求较高的应用场景。因此选取YOLOv5s算法作为研究对象,对其模型进一步优化与改进以

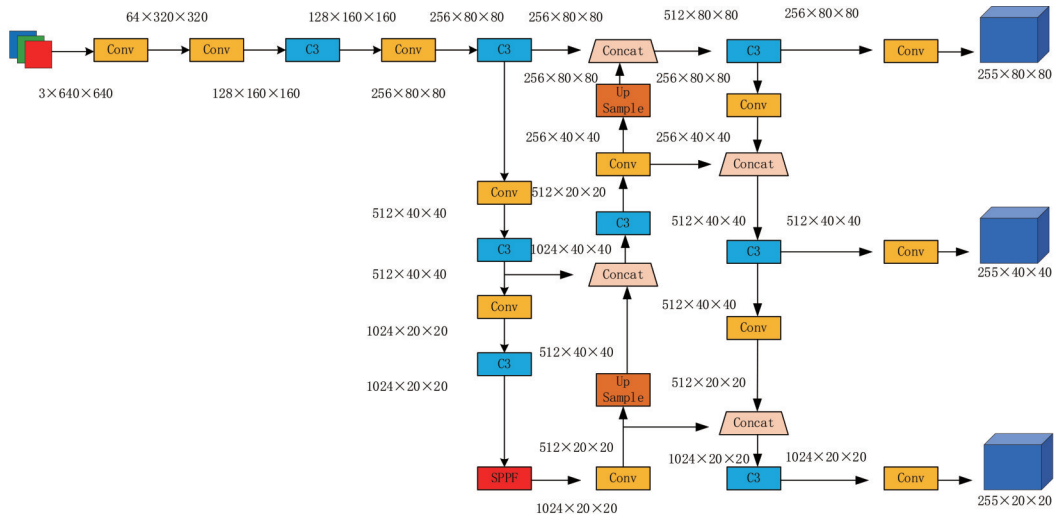


图 1 YOLOv5s 模型网络结构
Fig. 1 Network structure of YOLOv5s model

提高模型的性能,使其可以在火焰烟雾检测领域发挥其优势。

2 模型改进方法

2.1 改进 Backbone 层

注意力机制可以帮助模型在处理大量信息时,能够聚焦于关键部分,提高模型的性能和效率。在实际任务中,注意力机制通过对输入数据中的有用信息进行加权处理,使模型能够自动关注与任务相关的有用信息,从而提高模型的检测精度^[17]。

为了提升 YOLOv5s 模型的性能,在其 Backbone 层中加入 SE 注意力机制。SE 注意力机制是在通道维度上增加注意力机制^[18],能够获取特征图每个通道的重要程度。利用这个重要程度去给每个特征赋予一个权重值,从而让神经网络重点关注这些特征通道,提升对当前任务有重要影响的特征图通道,并抑制对当前任务作用不大的特征通道。SE 注意力机制实现流程如图 2 所示,在输入 SE 注意力机制之前(图 2 左侧图 C),特征图的每个通道的重要程度都是一样的,通过 SE-Net 之后(图 2 右侧图 C),每个特征通道的重要性就变得不一样了,图中不同颜色代表着不同的权重,神经网络会重点关注权重值大的通道。

SE 注意力机制的实现过程包括 3 个部分:

(1) 挤压(Squeeze): Squeeze 通过全局平均池化操作,将每个通道的二维特征($H \times W$)压缩为 1 个实数,将特征图从 $[H, W, C]$ 到 $[1, 1, C]$,整个输入特征图就被压缩成一个向量 z ,如式(1)所示。

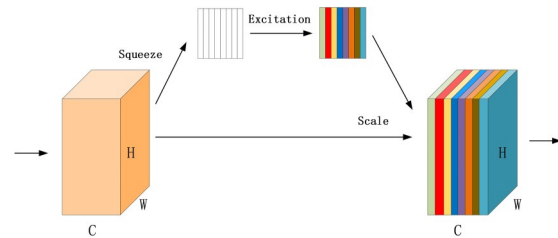


图 2 SE 模块实现流程
Fig. 2 Implementation process of SE module

$$z = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j). \quad (1)$$

(2) 激励(Excitation): Excitation 使用一个 Sigmoid 函数将压缩后的向量中的每个元素映射到 0~1 的范围,再通过两个全连接层对上一步得到的向量进行处理,得到通道权重 s ,如式(2)所示。

$$s = \sigma(g(z, w)) = \sigma(W_2 \sigma(W_1 z)). \quad (2)$$

(3) Scale: s 中不同的数值表示不同通道的权重信息,这个权重向量与原始输入特征图相乘,就得到加权后的特征图,如式(3)所示。这个步骤允许模型根据每个特征的重要性对其进行自适应的加权操作。

$$\bar{X} = F(u_c, S_c) = S_c * U_c. \quad (3)$$

SE 注意力机制的主要优点在于其能够自适应调整通道权重。此外,SE 注意力机制还可以降低模型的复杂性,因为它只关注重要的特征,而忽略那些对结果作用不大的特征,为模型提供了更强大的特征学习和数据处理能力。

2.2 改进 Neck 层

首先提出基于特征融合思想的 FPN,在 FPN

中建立一条自顶向下的通路,进行特征融合,用融合后具有更多语义信息的特征图进行预测,这种方法可以提高一定的精度。然而,这种自顶向下的FPN受到单向信息流的限制,精度不高。为了解决不同的输入特征具有不同的分辨率,对融合后输出特征做出不平等的贡献这一问题,提出BiFPN模块。这种模块结构解决了传统特征金字塔网络在提取不同尺度特征时存在的信息丢失和冗余的问题,通过引入双向流动的特征信息,可以实现多尺度特征的有效融合。BiFPN模块通过逐层融合上下采样的特征图,同时引入横向连接和纵向连接,使得不同尺度的特征能够更好地融合和利用。这种设计使BiFPN模块在处理尺度变化以及复杂场景时具有很强的鲁棒性。

2.3 改进后YOLOv5s网络结构

通过在YOLOv5s模型的Backbone层中引入SE模块,在Neck层中引入BiFPN模块,旨在提升YOLOv5s模型对火焰烟雾特征的提取能力以及能够有效地捕获多尺度特征,进而提高火焰烟雾检测的性能。虽然改进后的YOLOv5s网络结构

增加了一些模块,但可以融合更多的特征信息。在Backbone部分中共有11层结构,主要是Conv模块、C3模块、SE模块和SPPF模块,用于提取图像特征。

SE注意力机制加在主干网络SPPF前一层使模型能够更准确地识别不同尺度下的目标,提高目标检测的准确性和鲁棒性。此外,SE注意力机制还能够缓解在深度神经网络中增加深度带来的梯度消失问题,提高模型的训练效率和稳定性。在Neck部分中共有14层结构,主要由上采样、下采样以及可以结合双向流动的特征信息的全连接构成,用于特征融合。BiFPN模块加在Neck网络中,能够有效地融合不同尺度的特征信息,保留更多的空间信息和上下文信息,同时减少网络参数和运算量,从而提高模型的效率和性能。

相比于现有研究,本研究能更快速地响应并准确地检测火灾,尤其是在复杂背景和多尺度目标检测方面,为火焰烟雾检测领域带来了新的发展机遇,为公共安全提供了更有效的保障。改进后的网络结构如图3所示。

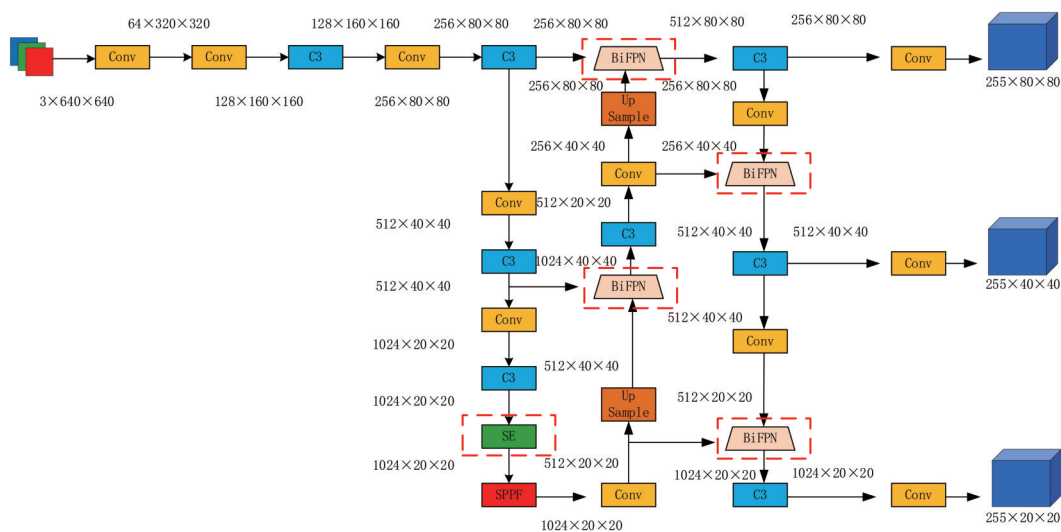


图3 改进后YOLOv5s网络结构

Fig. 3 The improved YOLOv5s network structure

3 实验结果与分析

3.1 实验环境

基于NVIDIA Geforce RTX 3080 Ti显卡,使用PyTorch 1.8.1深度学习框架进行实验,操作系统为Ubuntu 18.04,使用Python 3.8语言开发,并配置了对应的CUDA加速训练。

3.2 数据预处理

数据集是从Github、Baidu以及CSDN等多个网络平台所搜集到的火焰烟雾图片构建的1个火焰烟雾数据集,共计6000多张图片。按照6:2:2的比例划分为训练集、验证集和测试集,包括不同场景(如室内、室外、森林、建筑等)、不同火焰烟雾形态(如小火苗、大火、白烟、黑烟等)、不同光照条

件(如白天、夜晚、阴晴等),如图 4 所示。

使用 labelingm 图像标注软件对这些图片进行标注和分类,如图 5 所示。图 5 共有两类标签,分别是 fire 和 smoke。图片标注完成后会生成 VOC

格式文件,再转化为 YOLO 所需的 txt 格式文件并做归一化处理,这样,每一张火焰烟雾图像都对应有一份 txt 文件。



图 4 数据集部分图片

Fig. 4 Part of the dataset

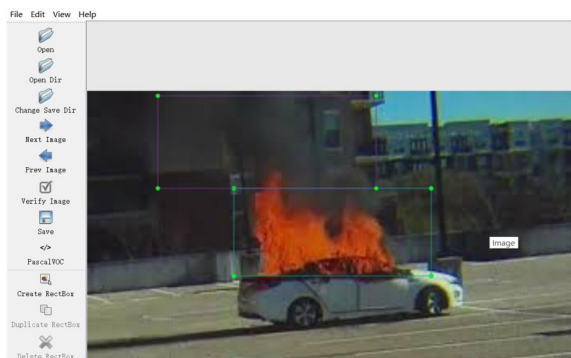


图 5 标注实例

Fig. 5 Annotated examples

3.3 参数设置

在模型的训练阶段,初始学习率设置为 0.01,优化算法选择随机梯度下降法(Stochastic Gradient Descent, SGD),轮次(epoch)设置为 100,迭代批次(batch size)设置为 16,输入图片尺寸设置为 640×640 。

3.4 评价指标

评价指标采用准确率(Precision, P),召回率(Recall, R)以及置信度 0.5 下的平均精度均值(mean Average Precision, mAP)。具体计算公式为:

$$P = \frac{TP}{TP + FP}, \quad (4)$$

$$R = \frac{TP}{TP + FN}, \quad (5)$$

$$AP = \int_0^1 P(R) dR, \quad (6)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (7)$$

式中:TP 代表预测正确的火焰烟雾目标数量,FP 代表预测错误的火焰烟雾目标数量,FN 代表遗漏的火焰烟雾目标数量。

3.5 实验与结果分析

为了验证所提出改进模型中使用每个模块的有效性,进行了消融实验。通过将 YOLOv5s、加入 SE 注意力机制模块后的模型(YOLOv5s-SE)、加入 BiFPN 模块后的模型(YOLOv5s-BiFPN)和同时加入 SE 注意力机制模块与 BiFPN 模块后的模型(YOLOv5s-SE-BiFPN),在实际火焰烟雾数据集上进行训练,对比不同模块对原始模型的影响效果,得到 4 组模型测试结果,如表 1 所示。

表 1 模型测试结果

Tab. 1 Model training results

模型	mAP/%	P/%	R/%
YOLOv5s	92.6	92.6	86.4
YOLOv5s-SE	93.2	93.9	86.9
YOLOv5s-BiFPN	93.6	94.0	87.0
YOLOv5-SE-BiFPN	94.1	94.4	89.0

从表 1 中可以看出,不同模块对模型检测精度提高均能发挥一定的作用。通过引入 SE 注意力

机制对YOLOv5改进的模型,在mAP、Precision和Recall指标上均得到了一定程度上的提升,分别相较于原始YOLOv5模型提高了0.6、1.3和0.5个百分点。通过引入BiFPN模块对YOLOv5改进的模型,在mAP、Precision和Recall指标上也得到了一定程度上的提升,分别相较于原始YOLOv5模型提高了1.0、1.4和0.6个百分点。以上说明了加入SE注意力机制模块或BiFPN模块改进了YOLOv5模型的可行性与有效性。对于同时加入两种模块的模型,检测精度也得到了明显的提升,这种提升相较于添加1种模块的提升效果更加明显,在mAP、Precision和Recall指标上,相比于原始YOLOv5模型提高了1.5、1.8和2.6个百分点,这也恰恰证明了YOLOv5-SE-BiFPN模型显著的优越性。实验结果表明,所提出的改进火焰烟雾检测模型在检测中精度更高,具有较好的性能表现。

YOLOv5s-SE-BiFPN模型的混淆矩阵如图6所示。混淆矩阵把模型检测性能可视化,矩阵的每1行代表1个真实的类别,每1列代表的是预测的类别。在矩阵中主对角线上的元素属于正确检测的样本概率;左下三角区域属于模型漏检的样本概率,表示模型未能正确检测到真实存在的样本,导致该目标未能被识别或错误分类成其他的类别;右上三角区域属于模型误检的样本数量,表示模型将背景或其他类别分类成当前类别。从图6可以看出,fire类别正确检测率为96%,smoke类别正确检测率为85%,检测准确率较高,基本上可以满足火焰烟雾检测的实时性与准确性的要求。

YOLOv5s-SE-BiFPN模型的检测效果如图7所示。可以看出,该模型在火焰烟雾图片上的表现较好,基本可以不遗漏地检测出图片中的火焰和烟雾,并且具有较高的置信度。

4 结束语

针对目前火焰烟雾检测算法存在的一些问题,提出了一种改进YOLOv5s的火焰烟雾检测方

参考文献:

- [1] 宋凯,赵碧. 回眸2023中国消防这一年[J]. 中国消防, 2024(1): 10-19.
SONG Kai, ZHAO Bi. Looking back on the year 2023 of China's fire department[J]. China Fire, 2024(1): 10-19.
- [2] BU Fengju, GHARAJEH M S. Intelligent and vision-based fire detection systems: A survey[J]. Image and Vision Computing, 2019, 91: 103803.
- [3] 曹康壮,焦双健. 融合注意力机制的轻量级火灾检测模型[J]. 消防科学与技术, 2024, 43(3): 378-383.

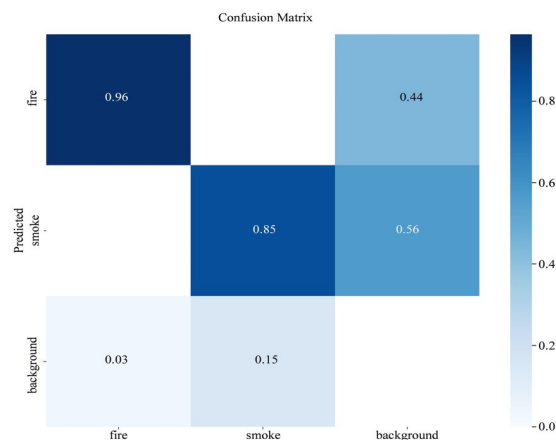


图6 YOLOv5s-SE-BiFPN混淆矩阵

Fig. 6 YOLOv5s-SE-BiFPN confusion matrix



图7 YOLOv5s-SE-BiFPN模型检测效果

Fig. 7 Detection effect of YOLOv5s-SE-BiFPN model

法。首先,在Backbone层中引入SE注意力机制,通过自适应调整通道权重,提升网络对火焰烟雾特征的提取能力。其次,在Neck层中引入BiFPN模块,使用双向连接提升模型多尺度特征融合能力。实验结果表明:改进后的YOLOv5s火焰烟雾检测方法能够对各种场景下的火焰烟雾目标进行有效地检测,并且检测精度较高,模型的泛化能力更好。在未来可以考虑进一步优化模型网络结构,以增强模型泛化能力;还可以尝试利用模型剪枝、量化等技术,降低模型训练的复杂度和计算成本。

- CAO Kangzhuang, JIAO Shuangjian. A light weight fire detection model integrating attention mechanism[J]. Fire Science and Technology, 2024, 43(3): 378-383.
- [4] CELIK T. Fast and efficient method for fire detection using image processing[J]. ETRI Journal, 2010, 32(6): 881-890.
- [5] GOMES P, SANTANA P, BARATA J. A vision-based approach to fire detection[J]. International Journal of Advanced Robotic Systems, 2014, 11(9): 149.
- [6] 杨雨卓. 基于图像处理的森林火险检测系统[J]. 电子世界, 2018(16): 19-20, 23.
YANG Yuzhuo. Forest fire risk detection system based on image processing[J]. Electronics World, 2018(16): 19-20, 23.
- [7] 谭勇, 谢林柏, 冯宏伟, 等. 基于图像的火焰检测算法[J]. 激光与光电子学进展, 2019, 56(16): 161012.
TAN Yong, XIE Linbo, FENG Hongwei, et al. Flame detection algorithm based on image processing technology[J]. Laser & Optoelectronics Progress, 2019, 56(16): 161012.
- [8] CHAOXIA Chenyu, SHANG Weiwei, ZHANG Fei. Information-guided flame detection based on faster R-CNN[J]. IEEE Access, 2020, 8: 58923-58932.
- [9] 颜洵, 吴正平, 雷帮军. 基于深度学习的视频火焰烟雾检测方法[J]. 信息通信, 2020(9): 70-72.
YAN Xun, WU Zhengping, LEI Bangjun. Research on video flame smoke detection based on deep learning[J]. Information & Communications, 2020(9): 70-72.
- [10] RYU J, KWAK D. A study on a complex flame and smoke detection method using computer vision detection and convolutional neural network[J]. Fire, 2022, 5(4): 108.
- [11] XIAO Pengcheng, XU Wenguang, ZHANG Yan, et al. Research on scrap classification and rating method based on SE attention mechanism[J]. Chinese Journal of Engineering, 2023, 45(8): 1342-1352.
- [12] HE Lei, WEI Haijun, WANG Qixuan. A new target detection method of ferrography wear particle images based on ECAM-YOLOv5-BiFPN network[J]. Sensors, 2023, 23(14): 6477.
- [13] SHAFIEE M J, CHYWL B, LI F, et al. Fast YOLO: A fast you only look once system for real-time embedded object detection in video[EB/OL]. (2017-09-18) [2024-05-01]. <https://arxiv.org/abs/1709.05943>.
- [14] WU Wentong, LIU Han, LI Lingling, et al. Application of local fully convolutional neural network combined with YOLOv5 algorithm in small target detection of remote sensing image[J]. PLoS One, 2021, 16(10): e0259283.
- [15] DENG Chunfang, WANG Mengmeng, LIU Liang, et al. Extended feature pyramid network for small object detection[J]. IEEE Transactions on Multimedia, 2022, 24: 1968-1979.
- [16] YU Haifeng, LI Xinbin, FENG Yankai, et al. Multiple attentional path aggregation network for marine object detection[J]. Applied Intelligence, 2023, 53(2): 2434-2451.
- [17] 孙艳歌, 吴飞, 周棋赢. 一种茶叶病害的深度学习检测算法[J]. 信阳师范学院学报(自然科学版), 2024, 37(2): 246-251.
SUN Yange, WU Fei, ZHOU Qiyong. A deep learning detection algorithm for tea diseases[J]. Journal of Xinyang Normal University (Natural Science Edition), 2024, 37(2): 246-251.
- [18] HUANG Guoheng, ZHU Junwen, LI Jianian, et al. Channel-attention U-Net: Channel attention mechanism for semantic segmentation of esophagus and esophageal cancer[J]. IEEE Access, 2020, 8: 122798-122810.

责任编辑: 郭红建