

基于改进 YOLOv8 的工厂火灾烟雾检测方法研究

孙 杰¹

SUN Jie

摘 要

针对工厂复杂环境下火灾烟雾存在检测困难,无法及时准确进行预警,实际部署难度高等问题,文章提出了一种基于改进 YOLOv8 的工厂火灾烟雾检测方法。首先对 YOLOv8 原有的损失函数进行修改,使用 WIoU 函数优化网络损失并提高收敛速度;引入 EMA 注意力机制,抑制复杂环境干扰的同时提高网络对目标的特征关注,以此提高对火焰和烟雾目标的检测能力;使用 DualConv 双卷积轻量化设计,降低参数量和 GFLOPs 的同时提升精度。根据实验结果可知,在工厂火灾烟雾数据集上,改进后的检测模型其平均精度提高 3.2%,GFLOPs 降低 3%,模型参数量降低 7.3%。实验结果验证了改进后的检测模型同时满足了检测精度、检测速度及轻量化的需求。

关键词

YOLOv8; 工厂火灾烟雾检测; WIoU; EMA; DualConv

doi: 10.3969/j.issn.1672-9528.2024.12.022

0 引言

工厂内放置各种设备,同时往往伴随着大量的易燃易爆物品,各种因素影响下会产生极易燃环境,因此,在这一复杂且敏感的环境中,火灾发生概率大幅提升,其爆发时间、触发点位等都充满不确定性与随机性,时刻威胁着工厂的安全生产与人员生命财产安全。基于此,提前检测到火灾的发生并及时发出警报,对于保护人员生命及财产安全至关重要。

当前工厂的火灾烟雾检测多以传统的基于传感器的火灾烟雾检测方法为主。其主要通过对烟雾、温度及光线等物理量进行感知检测,在长期使用过程中发现,这一方法存在检测范围小、灵敏度低、部署复杂等问题。因此,近年来,关于工厂火灾烟雾检测的研究探索以基于视频图像的火灾烟雾自动检测算法为主。

基于深度学习的火灾烟雾检测算法优势显著,检测速度大幅提升,能更迅速地捕捉烟雾迹象,同时,准确率也更高,极大减少了误报、漏报情况的发生,展现出极强的鲁棒性,稳定性远超传统算法,为工业场景的火灾预警提供了坚实可靠的技术支撑。文献[1]构建了一个基于卷积神经网络的多尺度火灾预测框架,该框架将融合烟雾和火焰在不同尺度下的静态特征。这种融合可以在一定程度上提高模型的检测效果。然而该框架还未能融合烟雾和火焰的动态特征。文献[2]为了实现多场景的烟雾与火焰检测,改进了二阶段目标检测算法,其具有高精度和准确的定位能力,但模型相对复杂且检测速度较慢。因此面向端到端的一阶段目标检测算法更为

受到研究者的关注。

YOLO 算法是一阶段检测算法中的代表之作,Qian 等人^[3]尝试使用通道剪枝方法轻量化 YOLOv3,通过分析 BN 层权重的系数来判断通道的重要性,并移除非重要通道,从而减少卷积参数的数量,有效减小模型的体积,提高模型在火灾检测任务上的效果。马庆禄等人^[4]在 YOLOv5s 中引入了 CBAM 卷积注意力模块^[5],以提高对轮廓特征不明显的烟雾和初期火焰特征的检测准确率。此外,研究者们还对一些模块进行了替换,以提高检测效率。尚明鹏等人^[6]提出融合 BiFPN 与 YOLOv5 网络,通过融合加权双向特征金字塔(BiFPN)来增强特征信息,达到在复杂的工厂环境下对火灾检测的要求。

现存目标检测算法在火灾烟雾检测场景中已有一定成果,但仍存在诸如缺乏工厂复杂环境下的火灾烟雾检测、模型较大部署困难、检测准确性欠缺以及难以检测早期火灾等问题。本文提出一种基于改进 YOLOv8 的工厂火灾烟雾检测方法。

1 YOLOv8 算法

2023 年 1 月 10 日,ultralytics 公司发布了 YOLOv5 的重大更新版本 YOLOv8,YOLOv8 由 Backbone、Neck 以及 Head 三部分构成。其提供了全新的 SOTA 模型以满足不同场景需求。Head 部分替换为主流的解耦头结构分离分类和检测头,同时也是无锚框模型,不易受到重叠物体的干扰。官方在 COCO Val 2017 数据集上测试的结果发现 YOLOv8 的精度相较于 YOLOv5 有大幅度提升。因此本文选取 YOLOv8 算法作为工厂火灾烟雾检测算法并将其进行改进。

1. 西安石油大学 陕西西安 710065

2 YOLOv8 目标检测网络改进

2.1 边界框回归损失函数优化

YOLOv8 使用的边界框回归损失为 Complete-IoU 即 CIoU。CIoU 虽考虑到重叠区域、长宽之比以及中心点距离，打破了 DIoU 出现负梯度抵消导致无法优化预测框的问题，针对边界框的回归有了更好的收敛率。但 CIoU 也存在预测框和真实框之间无法很好匹配的问题。文献 [7] 指出了 CIoU 公式中的 v 未反映预测框和真实框宽高的实际关系，同时也指出质量较差样本会对回归损失产生极大影响，同样质量较好样本也很难进一步优化，导致模型训练收敛速度和预测框准确性的受限。在本文中 Focal EIoUv1 被提出用来解决不同质量样本间 BBR（边界框回归）的平衡问题。但因其静态聚焦机制（focusing mechanism, FM），并未充分利用其非单调 FM 潜力。基于此，文献 [8] 提出了一种具有动态非单调 FM 的基于 IoU 的损失，将其命名为 Wise IoU（WIoU）。经综合考虑，本文选用 WIoUv3 作为模型的边界框回归损失函数。WIoUv3 具有在各个时刻都能及时调整梯度增益分配策略，更好地聚焦于一般质量之锚框，提升 YOLOv8 模型定位性能与收敛速度。

2.2 EMA 注意力机制

在工厂复杂的火灾烟雾场景检测中，光照变化可能导致图像亮度分布不均，影响火焰和烟雾目标与工厂背景的对比例。同时，火焰和烟雾目标的动态性使得目标边缘和轮廓模糊不清，降低了检测的准确性。

本文选用基于跨空间学习的高效多尺度注意力模块^[9]（efficient multi-scale attention, EMA）加入到 YOLOv8 网络中，其结构图如图 1 所示。

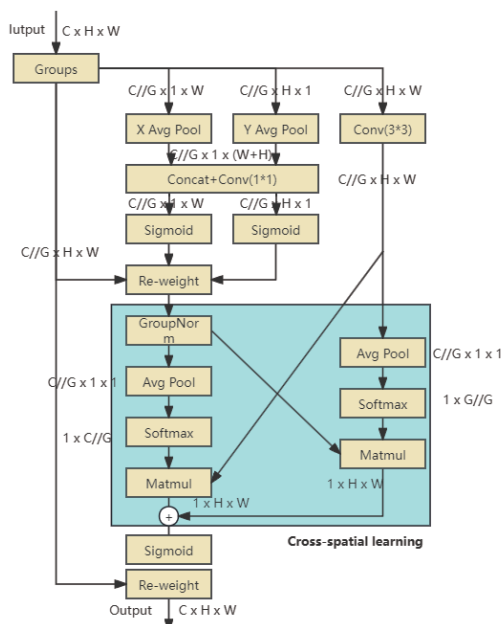


图 1 EMA 注意力结构图

该模块通过一系列策略来改进信息保留和计算开销。首先，将部分通道重塑为批量维度，并将通道维度分组为多个子特征，以实现更好的空间语义特征分布；其次，通过对每个并行分支中的通道权重进行重新校准，利用全局信息编码的方式来提高通道权重的准确性；最后，进一步聚合两个并行分支的输出特征，以捕获像素级的成对关系。

2.3 DualConv 双卷积

传统 YOLOv8 存在参数量大、检测速度较慢等问题，进而导致其无法适配部分硬件资源有限的嵌入式系统。为解决这一问题，本文引入 DualConv^[10] 双卷积构建轻量化深度神经网络。DualConv 是一种卷积神经网络结构，它结合了 3×3 和 1×1 卷积核，对相同的输入特征图通道进行处理，并采用组卷积技术优化卷积滤波器的排列方式。

GroupConv 作为一种卷积神经网络中的技术，早在 AlexNet 中就有应用。该技术通过将卷积核和输入特征图通道分为若干组来降低计算代价。但由于每组卷积核只应用于相应的输入通道组，限制了不同组通道之间的信息流动，从而降低了 GroupConv 的特征提取能力。

HetConv 作为一种异构卷积方式，在一个卷积 Filter 中同时包含 3×3 卷积核和 1×1 卷积核，并且采用移位方式交替排列。这种卷积方式可以显著减少原始 3×3 标准卷积的计算复杂度，但会破坏跨通道信息集成的连续性，并且影响输入特征图的完整信息的保存，从而降低网络的准确性。

DualConv 将 GroupConv 和 HetConv 的优点进行结合，使一部分卷积核仅执行 1×1 卷积的同时，让另一部分卷积核同时执行 3×3 和 1×1 卷积操作，如图 2 所示。

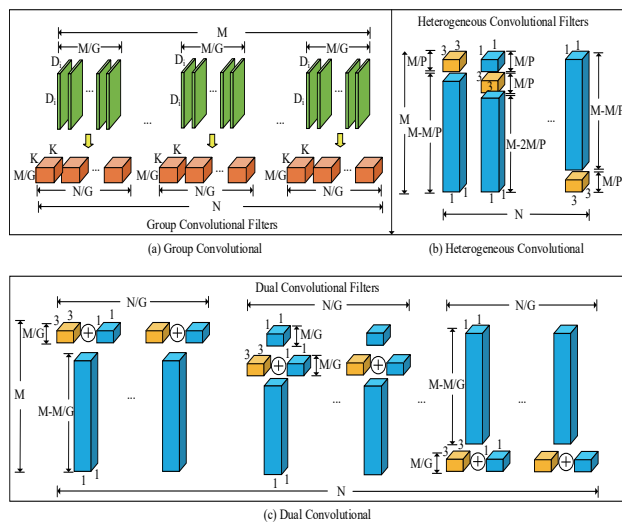


图 2 Group Conv(a)、HetConv(b) 和 DualConv(c) 原理对比图

该方案解决了跨通道信息传递的难题，并确保原始输入特征图内信息保存。DualConv 可以有效降低深度神经网络的参数量和计算成本。

3 实验及结果分析

3.1 构建工厂火灾烟雾数据集

本文实验数据集来源于网络视频图片及部分公开数据集。经过对所有图像数据的严格筛选，构建一个包含 3253 幅图像的数据集。其中训练集 2632 张，验证集 621 张。该数据集涵盖了不同明暗场景、不同尺度、不同角度以及不同距离的工厂烟雾与火焰图像，确保了数据集的差异性和多样性，为后续的研究工作提供了坚实的基础。

3.2 实验环境

本文进行的实验基于 Windows11 操作系统下，采用的硬件环境：AMD Ryzen 75800H，工作频率为 3.2 GHz，机带 RAM 为 16 GB，显卡为 NVIDIA GeForce RTX 3050 Ti，摄像头为 1080P HD 摄像头。采用的软件环境：Python 为 3.8.5，深度学习框架 PyTorch 版本为 1.8.0。构建好工厂火灾烟雾数据集和配置完对应实验环境后，本文采用 YOLOv8n 模型来进行相应的迭代训练。

3.3 评价指标

为准确分析和评估本文选用模型在工厂复杂环境下对火焰和烟雾的检测效果，本文采用平均准确率 AP（average precision）、平均精度均值 mAP（mean average precision）、参数量、GFLOPs 四大指标来对算法进行性能衡量评估。

3.4 实验结果及分析

3.4.1 损失函数对比试验

为了验证本文提出的 WIoU 损失函数在改进 YOLOv8n 模型上的实际效果和有效性，本文设计了一系列实验，将原 YOLOv8n 中使用的 CIoU 损失函数与基于 YOLOv8n 修改的 EIoU、DIoU、GIoU、SIoU 以及本文提出的 WIoU 五种不同的损失函数进行了对比，实验结果如表 1 所示。

表 1 损失函数实验结果对比

损失函数	AP ^{fire}	AP ^{smoke}	mAP
CIoU	0.645	0.637	0.641
DIoU	0.648	0.641	0.644
GIoU	0.652	0.637	0.645
SIoU	0.653	0.645	0.649
EIoU	0.656	0.644	0.650
WIoU	0.654	0.672	0.663

通过表 1 中不同损失函数的实验结果，可以看出，本文所采用的 WIoU 损失函数在性能上显著优于其他几种损失函数。相较于原 YOLOv8n 所使用的 CIoU 损失函数，WIoU 损失函数在 AP^{fire} 指标上从 0.645 提升至 0.654，在 AP^{smoke} 指标上则从 0.637 提升至 0.672，而 mAP 更是实现了 3.4% 的显著

提升。这一实验结果充分证明了 WIoU 损失函数在 YOLOv8n 模型中的实际应用效果和有效性，为提升模型性能提供了新的有效方法。

3.4.2 注意力机制对比实验

为验证本文提出的在原 YOLOv8n 模型中引入 EMA 注意力机制的合理性与有效性，进行了详尽的注意力机制对比实验。在实验中，选取市面上主流的注意力机制，包括 CoTAttention、CBAM、SE、SA、CA、GAM、SimAM、SKAttention、ECA 以及本文提出的 EMA，并将它们分别添加至 YOLOv8n 模型中，实验结果如表 2 所示。

表 2 注意力机制对比实验结果

注意力机制	AP ^{fire}	AP ^{smoke}	mAP	参数量 /MB	GFLOPs
YOLOv8n	0.645	0.637	0.641	3.01	8.2
YOLOv8n+CoTAttention	0.650	0.626	0.638	3.59	8.7
YOLOv8n+CBAM	0.639	0.643	0.641	3.05	8.2
YOLOv8n+SE	0.648	0.645	0.646	3.02	8.2
YOLOv8n+SA	0.645	0.649	0.647	3.01	8.2
YOLOv8n+CA	0.653	0.659	0.656	3.04	8.2
YOLOv8n+GAM	0.656	0.659	0.657	4.43	12.4
YOLOv8n+SimAM	0.656	0.658	0.657	3.01	8.2
YOLOv8n+SKAttention	0.653	0.665	0.659	8.56	12.6
YOLOv8n+ECA	0.651	0.667	0.659	3.01	8.2
YOLOv8n+EMA	0.657	0.669	0.662	3.02	8.4

由实验结果可知，在使用的 10 种不同注意力机制中，添加 EMA 注意力机制后的检测精度最高，相较于原 YOLOv8n 网络，其 AP^{fire} 由 0.645 提高到了 0.657，AP^{smoke} 由 0.637 提升到了 0.666，检测精度提高了 3.3%，参数量提高了 0.1 MB，GFLOPs 提高了 0.2。牺牲了少量的参数量和检测速度，有效提高了检测精度，验证了本文添加的 EMA 注意力机制的合理性和有效性。

3.4.3 主流算法对比试验

为进一步验证本文提出改进方法的优越性，本文尝试通过将改进后的 YOLOv8n 与目前主流的几种目标检测算法进行了对比实验，分别选取了 Faster-RCNN、SSD、YOLOv5、YOLOv7-tiny、YOLOv8n 作为实验对象，实验对比结果如表 3 所示。根据实验结果可知，在相同的训练参数及训练集的条件下，Faster-RCNN 算法有着不俗的检测精度，但是其参数量和检测速度都差强人意，很难满足火灾烟雾检测实时性的要求；YOLOv7-tiny 算法在参数量上有着其他模型难以企及的大小，但是其检测精度较低；而本文改进后的 YOLOv8n 算法在检测速度和检测精度上都远远优于其他检测算法模型，同时具有仅次于 YOLOv7-tiny 的参数量，从各个指标上

进行综合考虑可以总结本文所提出的改进算法更适用于火灾烟雾的检测。

表 3 主流算法对比实验结果

主流算法	AP ^{fire}	AP ^{smoke}	mAP	参数量 /MB	GFLOPs
SSD	0.645	0.637	0.641	3.01	8.2
Faster-RCNN	0.654	0.672	0.663	3.01	8.2
YOLOv5	0.657	0.666	0.662	3.02	8.4
YOLOv7-tiny	0.652	0.640	0.646	2.78	7.8
YOLOv8n	0.664	0.671	0.667	3.02	8.4
Ours	0.678	0.668	0.673	2.79	7.8

3.4.4 消融实验

本次消融实验以原 YOLOv8 模型为基线，依次将其损失函数修改为 WIoU，引入 EMA 注意力机制，使用 DualConv 双卷积。通过表 4 结果可知，当将损失函数修改为 WIoU 时，mAP 提高了 3.4%；在此基础上引入 EMA 注意力机制后，其模型参数量由原来的 3.01 MB 提升至 3.02 MB，GFLOPs 由 8.2 提升至 8.4，不过其检测精度提高至 0.667。在使用 DualConv 双卷积后，模型的参数量下降至 2.79 MB，模型的 GFLOPs 降低到 7.8，同时精度也有少量提升。总而言之，经过本文改进后的模型可以满足高精度、高检测速度及低参数量的要求，证明该改进方法的优越性。

表 4 消融实验结果

	AP ^{fire}	AP ^{smoke}	mAP	参数量 /MB	GFLOPs
YOLOv8n	0.645	0.637	0.641	3.01	8.2
YOLOv8n+Wiou	0.654	0.672	0.663	3.01	8.2
YOLOv8n+EMA	0.657	0.666	0.662	3.02	8.4
YOLOv8n+DualConv	0.652	0.640	0.646	2.78	7.7
YOLOv8+Wiou+EMA	0.664	0.671	0.667	3.02	8.4
YOLOv8n+Wiou+EMA+DualConv	0.678	0.668	0.673	2.79	7.8

4 结语

针对工厂复杂环境下的火灾烟雾检测任务，本文提出了一种基于改进 YOLOv8 的工厂火灾烟雾检测方法，通过修改损失函数进一步优化网络损失；添加 EMA 注意力机制，提高对火焰和烟雾目标的检测能力；引入 DualConv 双卷积，减小模型参数量，降低模型实际部署难度。

本文对工厂火灾烟雾数据集进行了充分验证，结果表明，本文改进的检测方法参数量小、检测速度快，同时在此基础上平均精度也有所提高，整体改进效果明显。但因各类工厂具体环境有所差别，因此，数据集可能依然存在部分问题，需要进一步改善。

参考文献：

[1] YUAN F N, ZHANG L, WAN B Y, et al. Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition[J]. Machine vision and applications, 2018, 30(11): 345-358.

[2] 张坚鑫, 郭四稳, 张国兰, 等. 基于多尺度特征融合的火灾检测模型 [J]. 郑州大学学报工学版, 2021, 42(5): 13-18.

[3] QIAN H M, SHI F, CHEN W, et al. A fire monitoring and alarm system based on channel-wise pruned YOLOv3[J]. Multimedia tools and applications, 2021, 81(10): 1833-1851.

[4] 马庆禄, 鲁佳萍, 唐小焱, 等. 改进 YOLOv5s 的公路隧道烟火检测方法 [J]. 浙江大学学报 (工学版), 2023, 57(4): 1-12.

[5] WOO S, PARK J H, LEE J Y, et al. CBAM:convolutional block attention module[C]//European Conference on Computer Vision(ECCV).Berlin:Springer, 2018: 3-19.

[6] 尚明鹏, 周敏, 陈燕军, 等. 融合 BiFPN 与 YOLO v5 网络的工厂火灾检测 [J]. 智能计算机与应用, 2023, 13(10): 137-141+146.

[7] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[DB/OL].(2019-11-19)[2024-06-11].<https://doi.org/10.48550/arXiv.1911.08287>.

[8] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[DB/OL].(2022-07-16)[2024-06-13].<https://doi.org/10.48550/arXiv.2101.08158>.

[9] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[DB/OL].(2023-06-06)[2024-06-15].<https://doi.org/10.48550/arXiv.2305.13563>.

[10] ZHONG J C, CHEN J Y, MIAN A. DualConv: dual convolutional kernels for lightweight deep neural networks[J]. IEEE transactions on neural networks and learning systems, 2023, 34(11): 9528-9535.

【作者简介】

孙杰(1999—), 男, 浙江宁波人, 硕士研究生, 研究方向: 目标检测。

(收稿日期: 2024-09-19)