

DOI: 10.13382/j.jemi.B2307019

融入 GhostNet 和 CBAM 的 YOLOv8 烟雾识别算法*

胡久松¹ 刘张驰¹ 余谦² 谷志茹¹ 钟皓¹

(1. 湖南工业大学轨道交通学院 株洲 412007; 2. 湖南汽车工程职业学院航海工程学院 株洲 412000)

摘要: 火灾防控在公共安全中至关重要,传统探测手段在特定环境下面临局限。计算机视觉技术能实时监控大范围,精准识别烟雾等火灾前兆。但烟雾形状、纹理和颜色的复杂性给机器视觉的烟雾精准识别带来了极大挑战。针对这个问题,设计了一种融入轻量级网络和卷积块注意力机制的 YOLOv8 烟雾分类算法,旨在提升烟雾分类的精度与效率。首先,算法采用了 GhostNet 架构,通过替换传统的卷积层,保持高性能的同时,极大减轻了模型的负担。其次,算法嵌入了 CBAM 注意力机制,能够自动调整对不同区域的关注程度,确保关键烟雾特征被优先处理和精细分析,增强了模型的鲁棒性。采用公开烟雾数据集和加入挑战样本的自制数据集进行了大量实验。实验结果证明,算法烟雾识别准确率在公开数据集上达到了 99.9%,在自制数据集上达到了 99.2%,优于同类方法。在实验电脑上,算法在 GPU 条件下帧率达到了 833 fps, CPU 条件下帧率达到了 28 fps,可以用于快速准确地进行早期火灾探测。

关键词: 早期火灾探测;烟雾识别;YOLOv8;GhostNet;CBAM

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.2060

YOLOv8 smoke detection algorithm integrated with GhostNet and CBAM

Hu Jiusong¹ Liu Zhangchi¹ Yu Qian² Gu Zhiru¹ Zhong Hao¹

(1. College of Railway Transportation, Hunan University of Technology, Zhuzhou 412007, China;

2. School of Marine engineering, Hunan Automotive Engineering Vocational University, Zhuzhou 412000, China)

Abstract: In the crucible of public safety, the imperative to guard against the scourge of fire is non-negotiable, yet conventional detection methodologies often falter when confronted with the complexities of specific environments. Herein lies the promise of computer vision technology, which offers the capability to monitor expansive territories in real-time and to identify the telltale signs of impending fires, most notably smoke. However, the intricate morphologies, textural variations, and chromatic subtleties of smoke present a significant challenge to the precision of its detection through machine vision. Addressing this exigency, we have conceived and developed an innovative smoke classification algorithm, seamlessly integrating a lightweight neural network and the convolutional block attention module (CBAM) within the YOLOv8 framework. This approach is designed to augment the accuracy and efficiency of smoke classification. Our algorithm leverages the GhostNet architecture, ingeniously replacing standard convolutional layers with a more efficient alternative, thereby maintaining high performance while drastically reducing the computational load on the model. Furthermore, the integration of CBAM imbues the algorithm with the ability to dynamically adjust its focus across different regions of the image, ensuring that salient smoke features are prioritized for detailed analysis. This feature enhances the model's robustness and adaptability to diverse scenarios. To validate the efficacy of our algorithm, we conducted extensive experiments using both a publicly available smoke dataset and a custom dataset augmented with challenging samples. Empirical results have demonstrated that our algorithm achieves a smoke recognition accuracy of 99.9% on the public dataset and 99.2% on the custom dataset, outperforming existing methods. On our experimental machine, the algorithm exhibited a frame rate of 833 fps under GPU-accelerated conditions and 28 fps under CPU-only operation, affirming its potential for rapid and accurate early fire detection.

Keywords: early fire detection; smoke recognition; YOLOv8; GhostNet; CBAM

收稿日期:2023-11-04 Received Date: 2023-11-04

* 基金项目:湖南省教育厅优秀青年项目(293832)、湖南省自然科学基金项目(2023JJ50198,2022JJ50005)资助

0 引言

消防安全预警系统,如石油化工领域的火炬黑烟监测系统^[1]和森林火灾预警系统^[2],对预防火灾至关重要。其中快速准确地进行烟雾识别是实现早期火灾探测的关键步骤,是快速预警和后续进一步进行烟雾火焰位置定位、智能灭火等操作的保障。当前烟雾识别主要依赖人工与传感器,但前者受限于人力与注意力,后者受制于滞后、环境干扰与覆盖不足,均难满足工业与安全预警需求。

基于图像处理和机器学习的烟雾识别技术因能克服这些局限而备受关注。例如,文献[3]对烟雾纹理采用了融合表面小波变换和隐马尔可夫树的方法来进行特征提取增强;文献[4]则利用局部二值模式和局部二值方差模式来识别动态纹理。这些方法虽然在某些应用场景中展现了不俗的性能,但由于其固有的依赖于手工设计特征的限制,其泛化能力和准确度仍然受到了一定约束,且通常实时性较差。相比之下,近几年深度学习技术在多个图像分类任务上展现了其超越传统方法的潜力,例如 Alex-Net^[5]、Dense-Net^[6]。针对烟雾识别,文献[7]提出的 DNCNN,在烟雾识别方法均取得了不错的准确率。而 Gu 等^[8]在 DNCNN 基础上提出的双网络结构 DCNN,在公开烟雾数据集上更是达到了 98.4% 的检测精确度。尽管这些算法获得了不错的效果,但在准确性、鲁棒性和实时性方面仍有待提高,特别是在复杂环境下的识别能力不足,如光照变化、遮挡情况以及背景干扰等条件下,识别性能下降明显。旨在进一步提高烟雾识别的准确性、鲁棒性和实时性,首先,在公共数据集的基础上构建了加入具有挑战性负样本的新数据集。然后,提出了一种融入 GhostNet 轻量级网络^[9]和卷积块注意力模块^[10](convolutional block attention module, CBAM)的 YOLOv8 烟雾识别算法—YOLOv8-smoke。YOLOv8-smoke 算法首先使用 GhostNet 轻量级网络代替原有的卷积操作生成特征映射从而减少计算成本,从而获得更高的帧率。然后,通过在主干网络中嵌入 CBAM 卷积注意力模块抑制背景信息干扰和强化烟雾特征以获得更高的准确率。使用公开数据集^[7]和自制数据集进行实验并与 Alex-Net^[5]、VGG-Net^[11]、Google-Net^[12]、Dense-Net^[6]、Res-Net^[13]、DNCNN^[7]、DCNN^[8]、YOLOv5^[14]、YOLOv8^[15]等进行了实验对比,结果表明 YOLOv8-smoke 烟雾分类准确率在公开数据集达到了 99.9%,在自制数据集达到了 99.2%,优于所有对比方法。在电脑配置 CPU Intel (R) Core i5-13490F 和 GPU NVIDIA GeForce GTX 4070Ti 硬件条件下,帧率可以达到 833 fps。在不使用 GPU 的情况下,帧率可以达到 28 fps,表明了 YOLOv8-smoke 算法的低时

耗性。

文章主要贡献:设计了适用烟雾二分类检测的 YOLOv8-smoke 深度神经网络,在公开数据集上准确率达到到了 99.9%,帧率达到了 833 fps;采用了 GhostNet 架构,通过替换传统的卷积层,极大减轻了模型的负担,使其成为一款轻量级且高性能的烟雾检测引擎;嵌入了 CBAM 注意力机制,能够自动调整对不同区域的关注程度,确保关键烟雾特征被优先处理和精细分析,从而大幅度提升模型的识别准确性和鲁棒性;在基础烟雾数据集之上,融入了一系列极具挑战性的图像样本,构建了一个全新的、更具代表性和多样性的数据集。

文章结构如下:第 1 节详述了 YOLOv8-smoke 网络框架基本原理和优势分析。第 2 节通过与其他烟雾识别算法进行对比实验以及自身模块的消融实验,证明了 YOLOv8-smoke 网络的优越性。第 3 节对全文内容进行总结。

1 YOLOv8-smoke 网络框架

YOLOv8-smoke 分类网络是在 YOLOv8 分类网络^[14]的基础上,如图 1 所示,融入 GhostNet 轻量级网络^[9]和 CBAM 卷积块注意力模块^[10]。

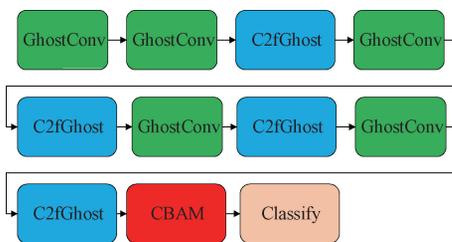


图 1 YOLOv8-smoke 的结构

Fig. 1 The structure of YOLOv8-smoke

1.1 YOLOv8 分类网络结构

如图 2 所示, YOLOv8 分类网络由 5 个卷积层 (Conv) 和 4 个 C2f 层 (coarse-to-fine, 粗细层) 交替组成。其中 Conv 是卷积神经网络中最基本的卷积操作,用于基本的特征提取。C2f 结构功能主要是通过结合粗糙和细节层次的特征来改善目标检测的准确性和效率。该结构使用多尺度特征融合技术,从粗糙到细致逐步提炼信息,使得模型能够在不同分辨率的输入图像上都实现高效的目标识别。这种自底向上和自顶向下的信息流动优化了特征的利用。

1.2 GhostNet

华为诺亚方舟实验室在 2020 年开发的 GhostNet 是一个轻量级网络^[9],旨在压缩网络并简化模型,同时确保模型具有一定的精确度,被广泛用于轻量化网络设

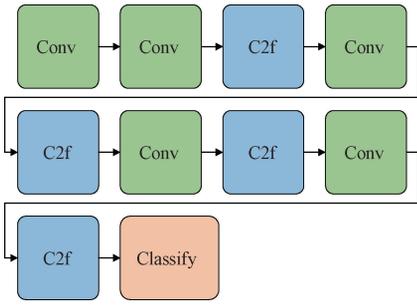


图 2 YOLOv8 分类网络结构图

Fig. 2 YOLOv8 classification network architecture diagram

计^[16-17]。如图 3 所示, GhostNet 模型在减少计算量和参数负载的同时, 保留了通道尺寸。它最初使用有限的普通卷积进行特征提取。之后, 对特征图进行线性变换操作, 这比常规卷积更加计算效率高。最终特征图通过张量拼接操作生成。

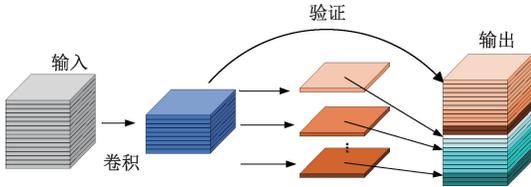


图 3 GhostNet 的结构

Fig. 3 The structure of GhostNet

如图 4 所示, GhostConv 作为 GhostNet 网络中的第一个卷积模块出现, 提供了传统卷积的有效替代。其中嵌入的“廉价操作”是一种经济的线性操作。与传统卷积不同, GhostConv 组织特征提取和成本效率高的线性操作的并行执行, 随后将两组结果特征图进行连接。这种组织有助于进一步减轻计算负担。因此, GhostConv 被有目的地构建为一个分阶段的卷积计算模块, 通过特征提取和成本效率高的线性操作的并行执行, 产生大量的特征图, 从而体现其计算效率。过程始于 GhostConv 使用半尺寸的普通卷积生成一半的特征图。

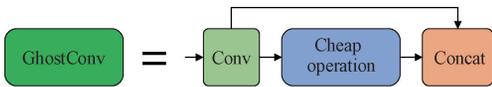


图 4 GhostNet 的组成结构

Fig. 4 Composition structure of GhostNet

图 5 为 GhostBottleneck 操作。首先, 通过利用初始的 GhostConv 层作为扩展层来增加通道数量。随后, 对特征图应用正则化和 Sigmoid 线性单元 (SiLU)。其次, 使用第 2 个 GhostConv 层将输出特征图中的通道数减少以匹配输入通道。最后, 将上一阶段获得的特征图添加到残差边缘以进行特征融合。

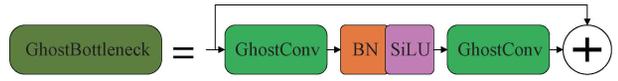


图 5 GhostBottleneck 结构

Fig. 5 The structure of GhostBottleneck

如图 6 所示, C2fGhost 模块用 GhostBottleneck 单元替换了原始网络 C2f 模块中的所有 Bottleneck 组件。这种结构整合了跨阶段特征融合策略和截断梯度流技术, 以增强不同网络层之间学习特征的多样性, 减少冗余梯度信息的影响, 提高学习能力。GhostConv 和 C2fGhost 模块的引入有效减少了对原始结构中众多 3×3 普通卷积的依赖。显著压缩了网络模型的尺寸, 减少了参数数量和计算负载, 并便于在移动设备上部署。

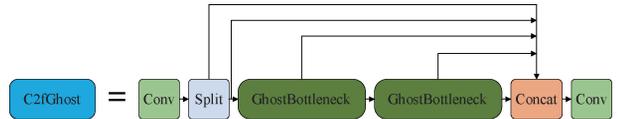


图 6 C2fGhost 的结构

Fig. 6 The structure of C2fGhost

1.3 CBAM 机制

注意力机制^[18-19]在捕获整个图像中的焦点区域方面发挥着重要的作用。通过空间和通道两个维度对烟雾图像进行特征提取, 这有利于提取烟雾的轮廓和纹理信息特征。

如图 7 所示, CBAM^[10] 由通道注意力机制 (channel attention module, CAM) 和空间注意力机制 (spatial attention module, SAM) 两个模块组成。图 8 和 9 分别详细描绘了 CAM 和 SAM 的基本结构^[20]。设特征图输入为 $A \in R^{C \times H \times W}$ 。

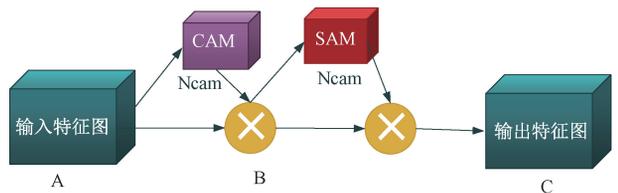


图 7 CBAM 注意力机制结构图

Fig. 7 The structure of CBAM

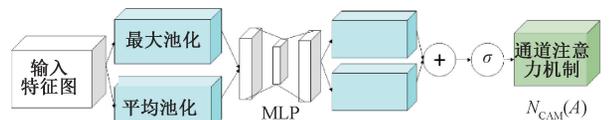


图 8 CAM 结构图

Fig. 8 The structure of CAM

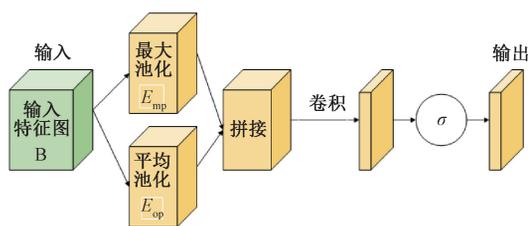


图 9 SAM 的结构

Fig. 9 The structure of SAM

在 CAM 结构中, A 先通过通道注意力处理得到 B , 再经空间注意力得到最终的激活特征图 C , 其公式表达如式(8)所示。

$$\begin{cases} B = N_{CAM}(A) \otimes A \\ C = N_{SAM}(B) \otimes B \end{cases} \quad (1)$$

其中, \otimes 代表逐元素乘法。当操作数维度不匹配时, 沿通道维度扩展空间注意力值, 而沿空间维度扩展通道注意力值。

2 实验

2.1 实验环境搭建

在 PC 上搭建了实验环境。实验环境的软硬件配置如表 1 所示。为评估 YOLOv8-smoke 相对于其他方法的效能, 选取了 3 个评价指标: 准确率 (accuracy rate, AR)、检测率 (detection rate, DR) 及假警报率 (false alarm rate, FAR)。相关指标定义如式(2)~(4)所示。

$$AR = \frac{P_1 + N_2}{T_1 + T_2} \times 100\% \quad (2)$$

$$DR = \frac{P_1}{T_1} \times 100\% \quad (3)$$

$$FAR = \frac{N_1}{T_2} \times 100\% \quad (4)$$

表 1 实验工具及环境

Table 1 Experimental tools and environment

工具及环境项目	详情
深度学习框架	Tensorflow, Keras
操作系统	Windows 11
CPU	Intel(R) Core i5-13490F
GPU	NVIDIA GeForce GTX 4070Ti
数据集	烟雾公共数据集 ^[7]
Python 版本	3.8

其中, T_1 和 T_2 分别表示正样本数量和负样本数量; P_1 代表被正确检测到的真正例样本数量; N_1 被错误分类为正样本的负样本数量; N_2 被正确检测到的真负例样本数量。理想的模型在 AR 和 DR 指标上应显示高值, 而在

FAR 指标上则应有低值。

YOLOv8-smoke 实验方案及参数如表 2 所示, 均是经过多次实验验证获取的经验操作和经验值。其中, 为降低图像块的方差并增强稳健性, 采用了归一化和数据增强策略。使用像素级的最小最大值归一化, 如式(5)所示。

$$d_n = \frac{d_r - d_{\min}}{d_{\max} - d_{\min}} \quad (5)$$

其中, d_n 是像素的归一化值, d_r 表示像素的强度值, d_{\min} 和 d_{\max} 分别是图像块中像素的最小值和最大值。此外, 通过如随机裁剪、旋转和色彩抖动等数据增强手段, 增大了模型的训练样本多样性, 提高其泛化能力。

表 2 YOLOv8-smoke 实验方案及参数

Table 2 Experimental scheme and parameters for YOLOv8-smoke

方案及参数	详情
初始化网络参数方法	Glorot 均匀方法
优化方法	随机梯度下降法
Momentum 系数	0.9
初始学习率	0.01
学习率衰减系数	0.0001
损失函数	交叉熵
编码方式	One-hot 编码
Batch size	32
训练周期 (epoch)	300

公开烟雾数据集具有局限性。如图 10(a) 和 (b) 所示, 其分类的正负样本的差异性太大, 这不利于对图 10(c) 所示具有挑战性的差异性较小的样本进行识别。因此, 在公开数据集的基础上, 加入了具有挑战性的样本做成自制数据集。

在实验过程中, 首先, 对 Ghost-Net 和 CBAM 分别进行了消融实验, 对比了精度指标和速率。其次, 对提出的 YOLOv8-smoke、经典网络和其他烟雾分类网络在公共数据集和自制数据集进行了实验对比。

2.2 GhostNet 消融实验

为了证明 GhostNet 的优势, 以 GhostNet 为单一变量, 分别对使用了 GhostNet 和没有使用 GhostNet 的算法进行训练, 然后进行指标的比较。如表 3 中消融实验 2 所示, 其结果表明, 使用了 GhostNet 和没有使用 GhostNet 算法的 AR、DR 和 FAR 分别为 99.90%、99.93%、0.70% 和 98.62%、98.77%、2.33%, 显然使用了 GhostNet 的算法指标更佳。

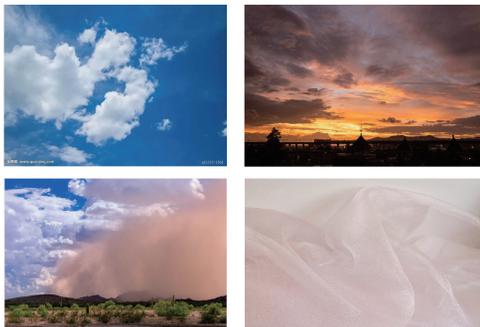
为更明确地呈现 GhostNet 的优势, 分别对使用了 GhostNet 和没有使用 GhostNet 的网络训练出来的模型进行推理, 结果如表 4 所示。显然, 使用了 GhostNet 的算法推理速度更快, 模型也更加轻盈。在硬件环境为 CPU



(a) 正样本
(a) Positive sample



(b) 负样本
(b) Negative sample



(c) 挑战性样本
(c) Challenging divergent samples

图 10 公开烟雾数据集的局限性

Fig. 10 Limitations of the public smoke dataset

Intel (R) Core i5-13490F, GPU NVIDIA GeForce GTX 4070Ti 时,使用了 GhostNet 的算法帧率达到了 833 fps,而未使用 GhostNet 的 YOLO8 算法只有 333 fps。而在未使用 GPU 的情况下,使用了 GhostNet 的算法帧率也可以达到 28 fps。

表 3 YOLOv8-smoke 的消融实验

Table 3 Ablation experiment of YOLOv8-smoke (%)

网络	消融实验 1 (去除 CBAM)	消融实验 2 (去除 GhostNet)	YOLOv8	YOLOv8-smoke
公共数据集				
AR	98.82	98.62	97.10	99.90
DR	98.35	98.77	97.41	99.93
FAR	1.84	2.33	3.56	0.70
自制数据集				
AR	97.68	97.60	95.63	99.20
DR	97.64	96.89	94.34	98.70
FAR	2.10	3.50	4.56	1.40

表 4 GhostNet 实验效果对比

Table 4 Comparison of experimental results for GhostNet

硬件配置		算法	帧率/ fps	时间/ms
CPU Intel(R) Core i5-13490F	GPU NVIDIA GeForce GTX 4070Ti			
GPU	YOLOv8-smoke		833	1.2
CPU	YOLOv8-smoke		28	35.4
GPU	YOLOv8 (不使用 GhostNet)		333	3
CPU	YOLOv8 (不使用 GhostNet)		15	65.4
GPU	Res-Net		588	1.7
CPU	Res-Net		19	50.1
GPU	Dense-Net		714	1.4
CPU	Dense-Net		24	41.3

2.3 CBAM 消融实验

为了证明 CBAM 的优势,以 CBAM 为单一变量,分别对融入 CBAM 和没有融入 CBAM 的算法进行训练后的指标比较。如表 3 中消融实验 1 所示,其结果表明,融入了 CBAM 和没有融入 CBAM 的算法的 AR、DR 和 FAR 分别为 99.90%、99.93%、0.70% 和 98.82%、98.35%、1.84%。为更明确地呈现 CBAM 注意力机制的优势,可视化了特征图。在 backbone 的最后一层随机提取了 3 个通道的特征图,结果如图 11 所示。对比 YOLOv8-smoke 加入 CBAM 之前与之后的特征图,引入 CBAM 后,烟雾特征的提取明显优于原始网络,验证了 CBAM 能够增强模型对于关键特征的捕获。

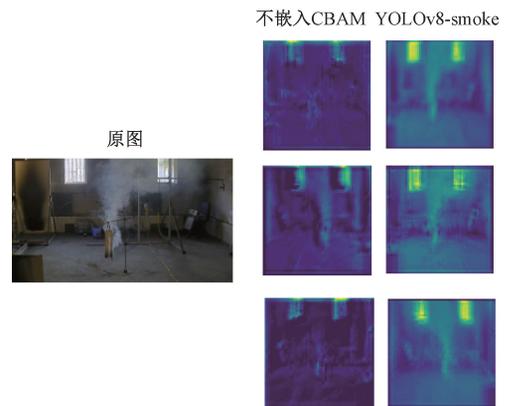


图 11 YOLOv8-smoke 在嵌入 CBAM 前后特征图对比

Fig. 11 Comparison of feature maps before and after CBAM integration in YOLOv8-Smoke

2.4 与经典网络进行比较

如表 5 所示,将 YOLOv8-smoke 与其他先进网络进行了对比,其中包括 Alex-Net、VGG-Net、Google-Net、Res-Net、Dense-Net、DNCNN、DCNN、YOLOv5 和 YOLOv8。结果显示,YOLOv8-smoke 在 AR 上,公开数据集可以达到 99.9%,自制数据集可以达到 99.2%,优于所有对比方法,在 DR 和 FAR 上也表现较佳。

表 5 YOLOv8-smoke 与其他分类网络的比较

Table 5 Comparison of YOLOv8-Smoke with nine other classification networks

(%)

网络	VGG-Net [11]	Google-Net [12]	Res-Net [13]	Dense-Net [6]	Alex-Net [5]	DNCNN [7]	DCNN [8]	YOLOv5 [14]	YOLOv8 [15]	YOLOv8-smoke 本文提出
公共数据集										
AR	96.96	94.76	97.04	94.25	93.27	95.51	97.22	98.56	97.10	99.90
DR	97.03	95.62	97.58	93.47	94.25	96.77	97.20	95.62	97.41	99.93
FAR	0.30	5.20	2.90	5.70	4.30	4.40	2.80	2.36	3.56	0.70
自制数据集										
AR	92.21	91.3	91.85	90.26	92.75	95.61	96.42	96.35	95.63	99.20
DR	91.51	91.25	89.91	88.45	90.21	95.23	96.96	95.64	94.34	98.70
FAR	2.37	4.68	5.36	4.73	3.52	2.36	1.92	1.65	4.56	1.40

最后,图 12 展示了训练的准确率曲线。与其他领先网络如 Alex-Net^[5]、Google-Net^[12]、VGG-Net^[11]、Dense-Net^[6]、Res-Net^[13]、YOLOv5^[14]和 YOLOv8^[15]等进行对比, YOLOv8-smoke 的收敛速度明显更快。根据综合的准确率评估,各网络的性能排名为:YOLOv8-smoke>YOLOv5>YOLOv8>DCNN>DNCNN>Res-Net>VGG-Net>Google-Net>Alex-Net。

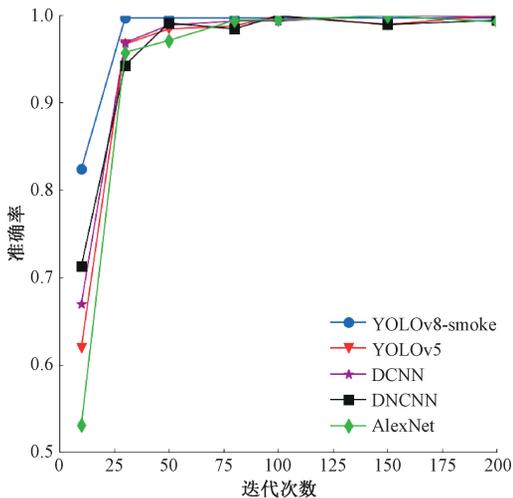


图 12 各模型在训练过程中准确率变化曲线图

Fig. 12 Plots of accuracy curves of each-model during the training process

3 结论

针对图像中烟雾的精准识别与分类,提出了融合 CBAM 注意力机制与 GhostNet 的 YOLOv8-smoke 轻量级网络架构。通过在 YOLOv8 主干网络中嵌入 CBAM 模块,不仅在空间和通道维度上增强了特征提取的精度,还确保了网络能够智能地聚焦于烟雾的关键细节,从而提升了模型的识别能力。采用 GhostNet 替代传统卷积层,以及 C2Fghost 替代 C2f 层,大幅优化了 YOLOv8-smoke 的运算效率,使其在保证高性能的同时,显著降低了计算

资源的需求。在公开烟雾数据集及自制数据集上的实验验证中, YOLOv8-smoke 表现出了卓越的性能,平均准确率超过 99.9%,且在实验环境下实现了高达 833 fps 的帧率,这不仅超越了现有模型,也充分证明了本研究在烟雾识别领域的先进性和必要性。下一步工作,拟研究烟雾火焰位置定位技术。通过对烟雾识别和烟雾火焰位置定位的研究,促进早期火灾探测技术的发展进程。

参考文献

- [1] HAMEED H, AZAM N, USMAN M, et al. RF sensing for smoking detection at oil fields [C]. 2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI), Denver, CO, USA, 2022: 944-945.
- [2] ZHANG L, LU C, XU H, et al. MMFNet: Forest fire smoke detection using multiscale convergence coordinated pyramid network with mixed attention and fast-robust NMS [J]. IEEE Internet of Things Journal, 2023, 10(20): 18168-18180.
- [3] YE W, ZHAO J H, WANG S, et al. Dynamic texture based smoke detection using surfacelet transform and HMT model [J]. Fire Safety Journal, 2015, 73: 91-101.
- [4] LIN G H, ZHANG Y M, ZHANG Q X, et al. Smoke detection in video sequences based on dynamic texture using volume local binary patterns [J]. KSII Transactions on Internet and Information Systems (TIIS), 2017, 11(11): 5522-5536.
- [5] KRIZHEVSKY A, SUTSKEVER I Y, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Advances in neural information processing systems, 2012, 25(2): 1097-1105.
- [6] HUANG G, LIU Z, VAN L, et al. Densely connected convolutional networks [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA: IEEE, 2017: 2261-2269.
- [7] YIN Z, WAN B, YUAN F, et al. A deep normalization

- and convolutional neural network for image smoke detection[J]. IEEE Access, 2017,5: 18429-18438.
- [8] GU K, XIA Z, QIAO J, et al. Deep dual-channel neural network for image-based smoke detection [J]. IEEE Transactions on Multimedia, 2020,22(2): 311-323.
- [9] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020: 1580-1589.
- [10] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[J]. Springer, Cham, 2018, DOI: 10.1007/978-3-030-01234-2_1.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]. International Conference on Learning Representations (ICLR), San Diego, CA, USA, 2015: 1-14.
- [12] SZEGED Y C, LIU W, JIA Y Q, et al. Going deeper with convolutions [C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA: IEEE, 2015: 1-9.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [14] YAO J, QI J, ZHANG J, et al. A real-time detection algorithm for kiwifruit defects based on YOLOv5 [J]. Electronics, 2021, 10(14): 1711-1711.
- [15] TERVEN J, ESPARZA C M D, GONZÁLEZ R A J. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS [J]. Machine Learning and Knowledge Extraction, 2023, 5(4): 1680-1716.
- [16] 林德钰, 周卓彤, 过斌, 等. 高斯混合模型与 GhostNet 结合的 YOLO-G 遗留物检测方法 [J]. 计算机辅助设计与图形学学报, 2023, 35 (1): 99-107.
LIN D Y, ZHOU ZH T, GUO B, et al. YOLO-G abandoned object detection method combined with Gaussian mixture model and GhostNet [J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35 (1): 99-107.
- [17] 叶卓勋, 刘妹琴, 张森林. 基于轻量化深度学习网络的工业环境小目标缺陷检测 [J]. 控制与决策, 2023, 38 (5): 1231-1238.
YE ZH X, LIU M Q, ZHANG S L, et al. Small-scale defect detection in industrial environment based on lightweight deep learning network [J]. Control and Decision, 2023, 38 (5): 1231-1238.
- [18] 张银胜, 杨宇龙, 吉茹, 等. 改进 YOLOv5s 的风力涡轮机表面缺陷检测 [J]. 电子测量与仪器学报, 2023, 37(1): 40-49.
ZHANG Y SH, YANG Y L, JI R, et al. Surface defect detection of wind turbine based on YOLOv5s [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 40-49.
- [19] 陈彦蓉, 高刃, 吴文欢, 等. 改进 YOLOv5 的新能源电池集流盘缺陷检测方法 [J]. 电子测量与仪器学报, 2023, 37(5): 58-67.
CHEN Y R, GAO R, WU W H, et al. Defect detection method for new energy battery collector disc based on improved YOLOv5 network [J]. Journal of Electronic measurement and Instrumentation, 2023, 37(5): 58-67.
- [20] WANG SH H, FERNANDES S L, ZHU Z Q, et al. AVNC: Attention-based VGG-Style network for COVID-19 diagnosis by CBAM [J]. IEEE Sensors Journal, 2022, 22(18): 17431-17438.

作者简介



胡久松(通信作者), 2019年于湖南大学获得博士学位, 现为湖南工业大学轨道交通学院讲师, 主要研究方向为智能感知与信息处理。

E-mail: hujiusong2008@163.com

Hu Jiusong (Corresponding author) received the Ph. D. degree from Hunan University in 2019. He is now a lecturer at the School of Rail Transit, Hunan University of Technology. His main research interests include intelligent sensing and information processing.



刘张驰, 2018年于湖南科技大学信息与电气工程学院获得学士学位, 现为湖南工业大学硕士研究生, 主要研究方向为深度学习、图像处理。

E-mail: liuzc_hn@163.com

Liu Zhangchi received his B. Sc. degree from Hunan University of science and technology, School of Information and Electrical Engineering in 2018. Now he is a M. Sc. candidate in Hunan University of technology. His main research interests include deep learning and image processing.



余谦, 2012年于湖南大学获得硕士学位, 湖南汽车工程职业大学电子学副教授, 主要研究方向为计算机视觉技术、森林经理学。

E-mail: yuqian_hunan@163.com

Yu Qian received his M. Sc. degree from Hunan University in 2012. Now he is an Associate Professor of Electronics at Hunan Automotive Engineering Vocational University. His main research interests include computer vision technology and forest management.