doi:10.3969/j.issn.1001-4616.2025.04.009

增强感受野特征的多尺度火灾检测方法

董　可,严云洋,耿嘉雯,于永涛,王盘龙,叶　翔

(淮阴工学院计算机与软件工程学院,江苏 淮安 223003)

[**摘要**]　针对当前火灾检测效果差和抗干扰能力弱等问题,提出一种增强感受野特征的多尺度火灾检测方法. 首先,引入感受野注意力卷积(receptive-field attention convolution,RFAConv),增强感受野空间特征的提取;其次,结合反向残差移动模块(inverted residual mobile block,iRMB)和通道先验卷积注意力(channel prior convolutional attention,CPCA)设计 C2fiC 模块,提高模型表达和融合不同尺度特征的能力;然后,采用共享参数结构,引入轻量卷积重构检测头,降低模型参数和计算复杂度;最后,引入 Focaler-GIoU 损失函数,平衡难易样本. 实验结果表明,改进模型参数量和计算量均有所降低,检测精度更高,能满足火灾场景中的检测要求.

[**关键词**]　火灾检测,感受野特征,注意力机制,损失函数,YOLOv8n
[**中图分类号**]TP391.4　[**文献标志码**]A　[**文章编号**]1001-4616(2025)04-0087-09

Multi-Scale Flame Detection Based on Enhanced Receptive Field Feature

Dong Ke,Yan Yunyang,Geng Jiawen,Yu Yongtao,Wang Panlong,Ye Xiang

(Faculty of Computer & Software Engineering,Huaiyin Institute of Technology,Huaian 223003,China)

Abstract:Aiming at the problems of poor fire detection effect and weak anti-interference ability,a multi-scale fire detection method based on enhanced receptive field feature is proposed. Firstly, the Receptive-Field Attention Convolution (RFAConv) is introduced to enhance the extraction of spatial features of receptive field. Secondly,the C2fiC module is designed by combining the Inverted Residual Mobile Block(iRMB) and the Channel Prior Convolutional Attention(CPCA) mechanism to improve the ability of the model to express and fuse different scale features. Then,the shared parameter structure is adopted,and the lightweight convolution reconstruction detector is introduced to reduce the model parameters and computational complexity. Finally,the Focaler-GIoU loss function is introduced to balance the difficulty samples. The experimental results show that the number of parameters and the amount of calculation of the improved model are reduced, and the detection accuracy is higher,which can meet the detection requirements in flame detection.

Key words:flame detection,receptive field feature,attention mechanism,loss function,YOLOv8n

火灾是发生率最高的灾害之一,严重威胁生命和社会公共安全,对易发生火灾的区域进行监控是一种有效减少火灾造成损失的方法,尽可能早地发现火灾并及时处理,能将损失降到最低. 早期的火灾检测技术利用各种传感器检测火灾初期产生的特征,如环境中烟雾浓度、环境温度以及光线的变化,进行预警. 但传感器检测火灾的方法受限于检测范围,需要部署大量的传感器,通常在火灾事故发生后才能检测到火灾,实时性较差. 此外,传感器检测方法容易受到环境干扰的影响,如烟雾传感器易受烟雾、尘埃等物质的干扰,温度传感器易受周围温度的影响;大型设备的运行、人员的行走等因素也会产生干扰,导致误报率较高[1].

随着人工智能的快速发展,深度学习在火灾检测领域取得了长足的进步,为火灾检测提供了新思路和新方法. 与早期的火灾检测技术相比,基于深度学习的火灾检测技术可以自动学习图像和视频数据中的高级和复杂特征,具有良好的鲁棒性和准确性. 皮骏等[2]采用 Shufflenetv2 对 YOLOv5s 进行轻量化操作,提高对森林火灾检测性能. 通过通道重组的方式提高模型的提取能力,在维持精度的同时提升检测速度. Zhang 等[3]提出一种多尺度收敛协调金字塔网络用于森林火灾烟雾检测,结合双注意力特征金字塔网络和协调收敛模块,提高对不同尺度目标的检测精度,设计混合注意力模块,增强对烟雾水平和垂直特征的提取. 龚成张等[4]提

收稿日期:2024-09-24.
基金项目:国家自然科学基金资助项目(62076107)、江苏省“六大人才高峰”资助项目(2013DZXX-023).
通讯作者:严云洋,博士,教授,研究方向:数字图像处理、模式识别. E-mail:yunyang@hyit.edu.cn

出一种基于 Fast-CAANet 的火焰检测方法,对 ConvNext 进行改进,有效提高了火焰特征的表达,在复杂场景下火焰检测效果更好. Wang 等[5]提出一种改进的 YOLOX 多尺度森林火灾检测方法,利用 HC-FPN 结构提高特征图的信息提取能力及其对多尺度目标的检测能力,采用小目标数据增强策略提高模型的训练效果和鲁棒性. Zheng 等[6]对 DETR 模型进行改进,通过在高级和低级特征之间交替更新优化编码器模块,简化检测过程,采用 Wise-IoU 损失函数提高模型在复杂环境下的鲁棒性及其对不同尺度火灾场景的适应性. 虽然基于深度学习的火灾检测技术在一定程度上取得了成效,但仍存在一些不足,如特征表示有限、对小目标的特征捕获能力差、局部特征处理不足等. 火灾场景通常更加混乱和复杂,随着火灾的发展,场景的复杂性会动态变化且高度不可控,包括强烈的光线、烟雾、火焰和周围结构的遮挡.

因此,提出一种增强感受野特征的多尺度火灾检测方法,用于解决火灾识别精度低和误检率高的问题.

1　YOLOv8n 模型

YOLOv8n 模型主要由三部分组成:Backbone,Neck 和 Head,如图 1 所示. Backbone 采用 Darknet-53 结构,包括 CBS、C2f 和 SPPF 三种模块. CBS 用于对输入图像或特征图进行卷积操作;C2f 模块通过划分(Split)和拼接(Concat)操作对输入进行降采样,提取特征,获取多尺度特征图;SPPF(Spatial Pyramid Pooling Fast)对输入特征图进行三次最大池化操作,捕获不同尺度的语义信息. Neck 采用 PAN-FPN,对 FPN 的特征图进行融合,从尺度较小的特征图向尺度较大的特征图进行特征融合,再从尺度较大的特征图向尺度较小的特征图进行特征融合,增强多尺度特征的表达能力. 这种策略将低尺度特征图和高尺度特征图相融合,将低层的强定位特征向上传递,将高层的强语义特征向下传递,同时增强了定位特征和语义特征. Head 采用 Decoupled Head 解耦头,较前代不同,这一代将分类头和检测头分离,特征图分别通过两条不一样的三次卷积得到最终结果. 在损失函数上,分类头和检测头采用不同的损失函数,其中分类头采用 BCELoss(Binary Cross-Entropy Loss),而检测头采用 DFL(Distribution Focal Loss).

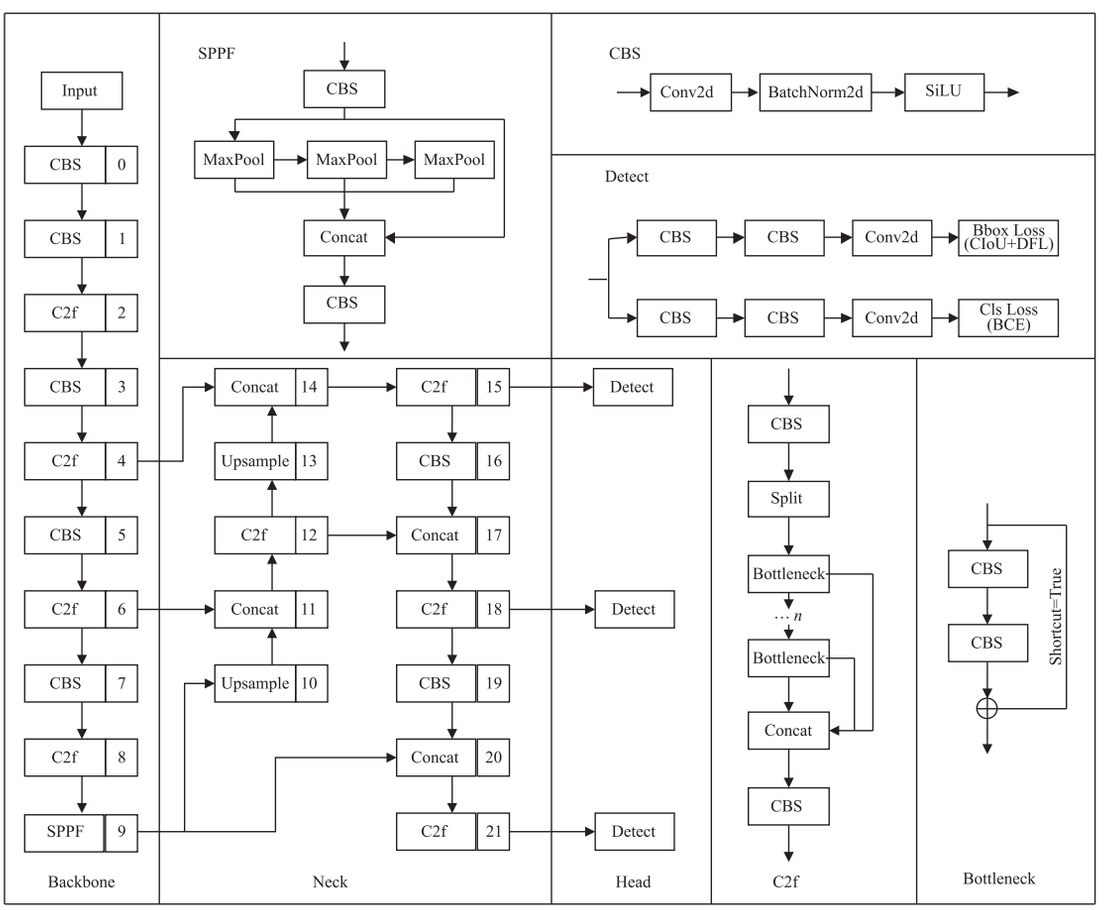


图 1　YOLOv8n 模型结构图

Fig. 1　YOLOv8n model structure diagram

2　增强感受野特征的多尺度火灾检测方法

2.1　感受野注意力卷积

感受野注意力卷积(receptive-field attention convolution,RFAConv)[7]动态生成感受野空间特征,根据卷积核的大小自适应调整感受野的形状和范围,为大尺寸卷积核提供良好的注意力权重,解决卷积核参数共享的问题. RFAConv 结构如图 2 所示.

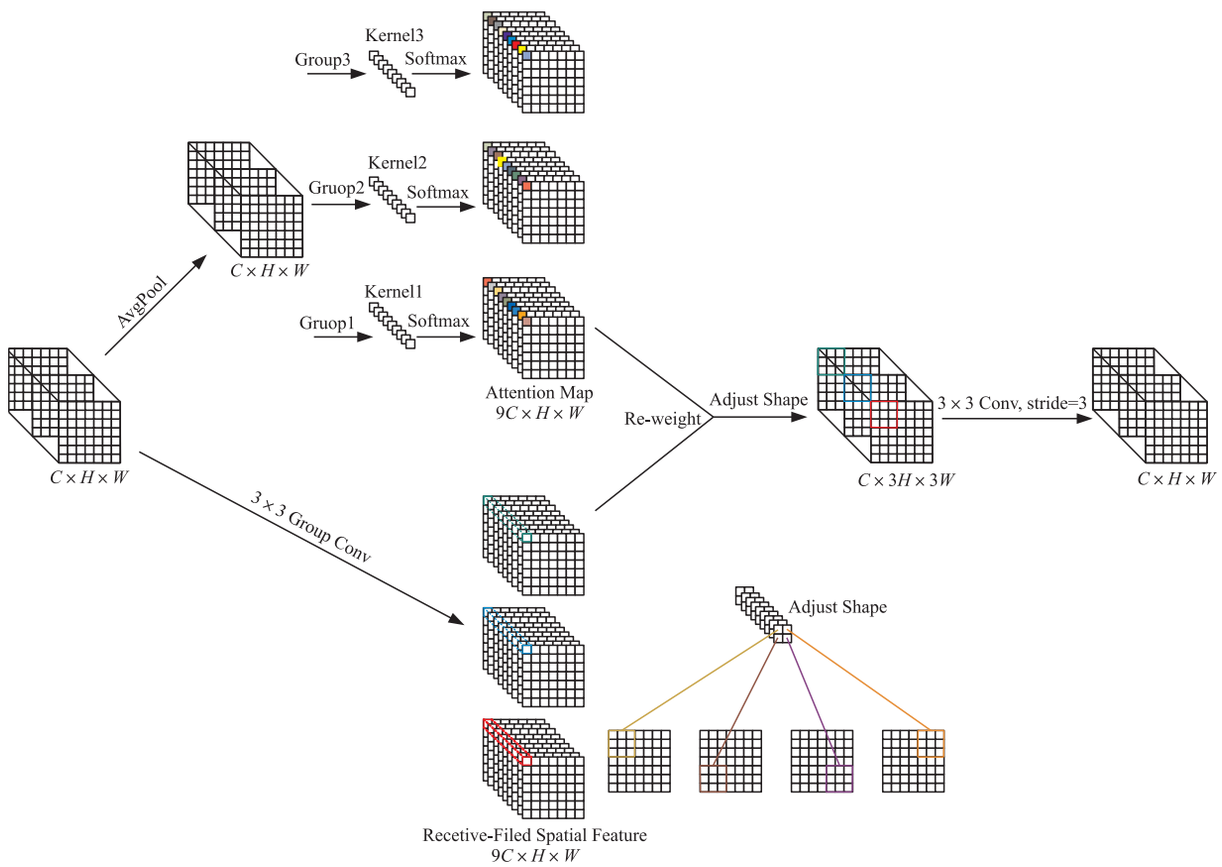


图 2　RFAConv 结构

Fig. 2　RFAConv structure

RFAConv 的引入使模型能够强调感受野滑块内各种特征的重要性,并关注到感受野空间特征. 其实现过程分为两部分:注意力图的生成和感受野空间特征的提取.

第一部分首先使用 AvgPool 池化从特征图中提取全局信息,聚合每个感受野特征;然后通过 1×1 组卷积运算促进感受野内信息的交互;最后利用 Softmax 突出感受野特征的重要性,为每个感受野特征图生成相应的注意力图.

第二部分根据感受野的大小,利用 3×3 组卷积提取空间特征,该特征通过非重叠滑动窗口变换原始特征图获得. 每个滑动窗口代表一个感受野滑块,其大小可以根据卷积核的大小动态调整. 通过这种方式,模型能够根据卷积核的大小生成适当大小的感受野空间特征. 将感受野信息丰富的注意力图与得到的加权特征图融合在一起,使用感受野权重矩阵为每个感受野位置和特征通道分配不同的权重,以突出重要的详细特征,并调整输出特征图的大小,得到最终的输出结果. RFAConv 计算过程如式(1)所示.

$$F=\mathrm{Softmax}(g^{1\times1}(AvgPool(X)))\times ReLu(Norm(g^{k\times k}(X)))=A_{rf}\times F_{rf},\tag{1}$$

式中,Softmax 表示向量激活函数,g 表示分组卷积,$AvgPool$ 表示平均池化,$ReLu$ 表示非线性激活函数,$Norm$ 表示规范化,k 表示卷积核大小,X 表示输入特征图,F 是通过注意力图 A_{rf} 与变换的感受野空间特征 F_{rf} 相乘而获得的输出特征图.

2.2　C2fiC 模块

反向残差移动模块(inverted residual mobile block,iRMB)[8]是一种结合 IRB(inverted residual block)、

MHSA(multi-head self-attention)和 FFN(feedforward network)三者优势的轻量化模块. IRB 处理图像中的尺度变化和细节信息,MHSA 捕获输入特征之间的全局远程依赖关系,而 FFN 执行非线性转换. 通过将 MHSA 和 DWConv 进行级联,兼顾动态全局建模和静态局部信息融合的优势,有效增强模型的感受野. 此外,iRMB 有两条分支,第一个分支基于 CNN(convolutional neural network)利用局部连通性、权重共享和池化从图像中提取局部特征,捕捉底层细节;第二个分支基于 Transformer 依赖神经网络结构中的自注意力机制,动态学习元素之间的远程依赖序列,促进图像的高级语义构建. 通过跨尺度连接实现两个分支之间的特征融合,提高模型的泛化能力和鲁棒性.

通道先验卷积注意力(channel prior convolutional attention,CPCA)[9]通过平均池化和最大池化等操作对特征图的空间信息进行聚合,然后通过一个共享的多层感知器(multilayer perceptron,MLP)进行处理,并将处理后的特征图添加到通道注意图中. 通道先验特征图通过输入特征和通道注意图进行逐元素乘法得到并将其输入到深度卷积模块中,生成空间注意图. 卷积模块接收用于通道混合的空间注意映射. 最后,通过对通道混合结果和通道先验特征图进行逐元素相乘,得到细化特征作为输出. 计算方法如式(2)~(5)所示.

$$CA(F)=\sigma(MLP(AvgPool(F))+MLP(MaxPool(F))), \tag{2}$$

$$SA(F)=Conv_{1\times1}\Big(\sum_{i=0}^{3}Branch_i(DwConv(F))\Big), \tag{3}$$

$$F_C=CA(F)\otimes F, \tag{4}$$

$$\hat{F}=SA(F_C)\otimes F_C, \tag{5}$$

式中,CA 表示通道注意力,F 表示特征图,σ 表示 Sigmoid 函数,$AvgPool$ 表示平均池化,$MaxPool$ 表示最大池化,SA 表示空间注意力,$Conv_{1\times1}$表示 1×1 卷积,B_i,$i\in\{0,1,2,3\}$ 表示第 i 个分支,$DwConv$ 表示深度卷积,F_C 表示通道特征图,⊗表示逐元素乘法,$\hat{F}$ 表示最终输出特征图.

为防止卷积迭代导致的特征细节丢失,将 CPCA 集成到 iRMB 模块中并设计了 iC 模块,采用多分支捕获多尺度信息融合到空间注意力图中,对特征图的空间信息进行聚合,增强多尺度特征融合能力,优化小目标的特征学习. iC 模块如图 3 所示.

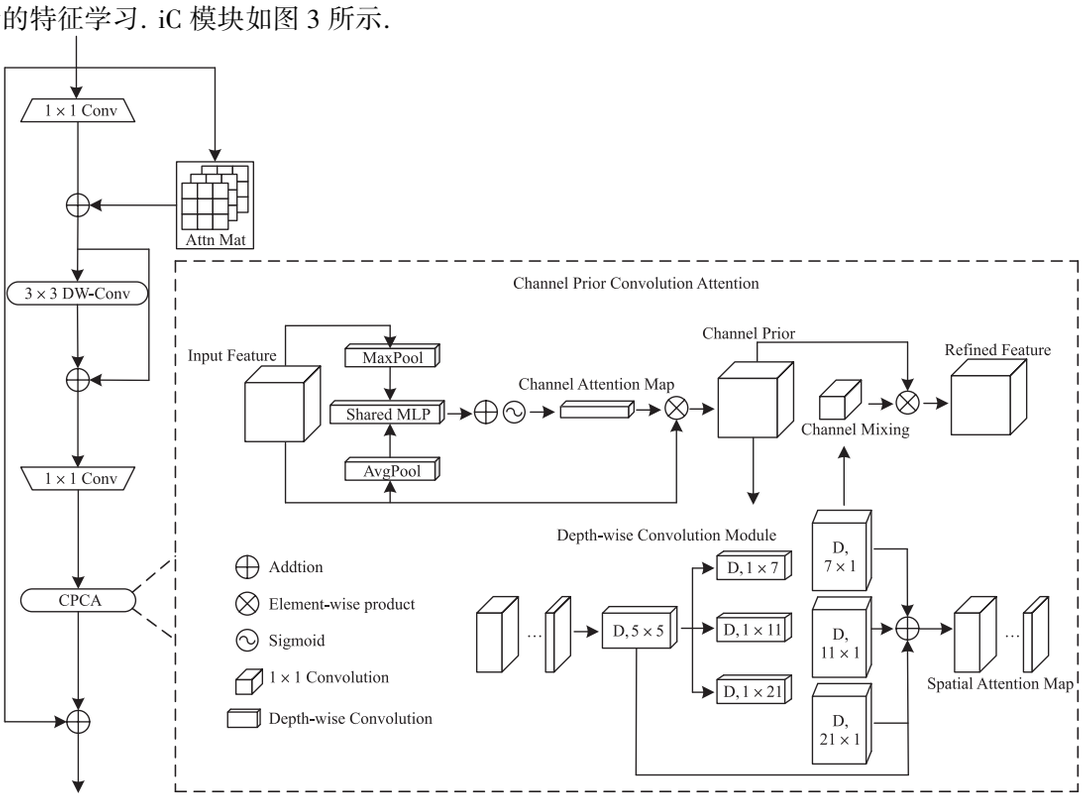


图 3　iC 模块

Fig. 3　iC module

在 iC 模块的基础上进一步构建了 BottleneckiC 模块和 C2fiC 模块,如图 4 中(a)和(b)所示,在颈部网络使用 C2fiC 模块,通过多个分支的并行处理和注意力机制的引入,C2fiC 模块能够充分利用多尺度的特征信息,增强低维和高维特征之间的相关性,从而更好地捕捉火灾区域的细微特征.

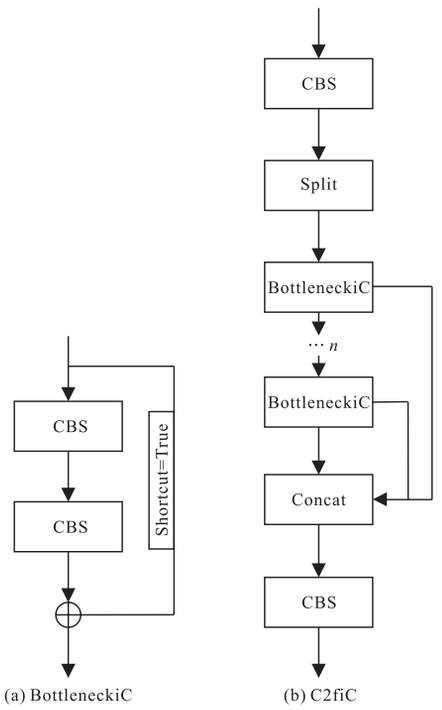


图 4　BottleneckiC 模块和 C2fiC 模块

Fig. 4　BottleneckiC module and C2fiC module

2.3　轻量检测头

YOLOv8n 采用了 Decoupled Head 解耦头,将分类头和检测头进行分离,特征图分别通过两个 3×3 卷积和一个 1×1 卷积来提取信息得到最终结果,这导致了检测头占了近一半的参数和计算负载. 考虑到实验主要研究火灾检测,实现火灾预警,计算复杂度要求较低,模型较轻,有利于实际部署. 基于此,采用共享参数的思想,重新设计检测头中的卷积部分,在 Input 之后添加了两个共享参数模块并引入了 GSConv[10] 卷积,然后继续进行解耦分类和回归分支.

GSConv 是一种由卷积(convolution,Conv)、深度可分离卷积(depthwise separable convolution,DSConv)和混洗(Shuffle)操作组成的轻量化卷积模块. 首先,对输入特征图进行卷积下采样;其次,使用 DSConv,执行 Concat 操作连接两个卷积的输出结果;然后,对通道维度上的特征映射进行分组,重组两个卷积的输出通道;最后,执行 Shuffle 操作,融合两个卷积生成的特征信息. GSConv 卷积操作在不影响检测精度的前提下,降低计算复杂度,提高模型泛化能力,其结构如图 5 所示.

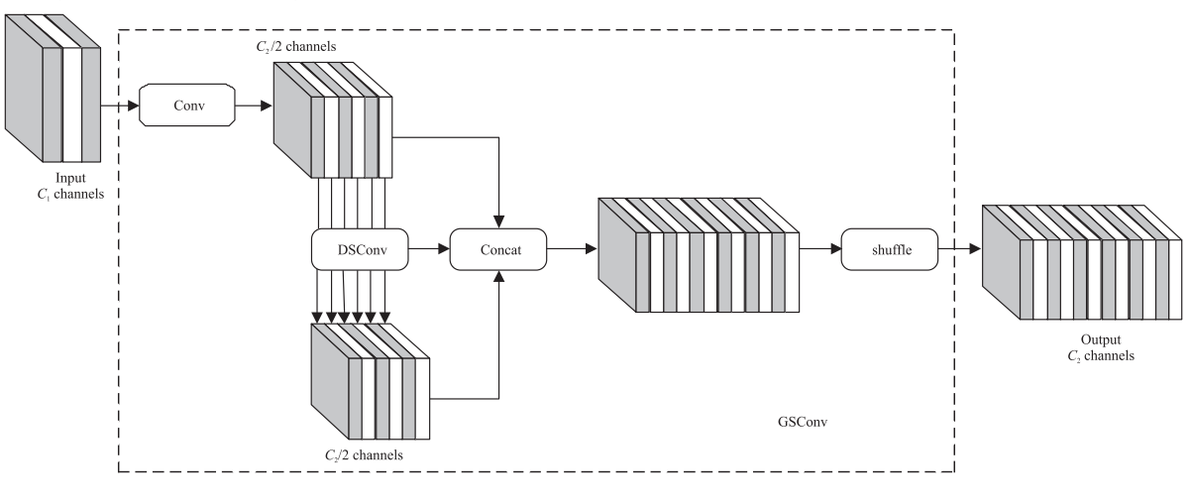


图 5　GSConv 结构图

Fig. 5　GSConv structural diagram

为了取得速度和准确性之间的平衡,在检测头使用 GSConv 卷积,实现降低模型复杂度的同时保持准确性. 改进后的轻量检测头如图 6 所示.

2.4　Focaler-GIoU 损失函数

YOLOv8n 采用 CIoU Loss[11] 作为损失函数,解决边界框回归问题. CIoU Loss 考虑了位置偏移、比例差异和纵横比等因素,从而更准确地评估预测框和真实框之间的相似性. 虽然 CIoU Loss 在解决目标检测中的边界框偏移和纵横比不平衡等

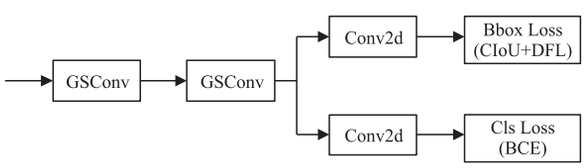


图 6　轻量检测头

Fig. 6　Lightweight detection head

问题方面优于传统的 IoU 计算,但是未考虑难易样本的平衡. 基于此,采用 Focaler-GIoU Loss 替换 CIoU Loss,GIoU Loss[12]优化网络学习策略,同时结合 Focaler-IoU Loss[13]关注不同的回归样本. GIoU Loss 如式(6)~(7)所示.

$$L_{\mathrm{IoU}}=1-\frac{|B\cap B^{gt}|}{|B\cup B^{gt}|}, \tag{6}$$

$$L_{\mathrm{GIoU}}=L_{\mathrm{IoU}}-\frac{|C-B\cap B^{gt}|}{|C|}, \tag{7}$$

式中,B 表示预测边界框,B^{gt}表示真实目标框 C,表示包含预测框和真实框的最小外接矩形框.

将 Focaler-IoU Loss 应用到基于 GIoU Loss 的边界框回归损失函数中,Focaler-GIoU Loss 如式(8)~(9)所示.

$$L_{\mathrm{Focaler\text{-}IoU}}=\begin{cases}0, & IoU<d\\ \dfrac{IoU-d}{u-d} & d\leqslant IoU\leqslant u\\ 1, & IoU>u\end{cases} \tag{8}$$

$$L_{\mathrm{Focaler\text{-}GIoU}}=L_{\mathrm{GIoU}}+L_{\mathrm{IoU}}-L_{\mathrm{Focaler\text{-}IoU}}, \tag{9}$$

式中,$L_{\mathrm{Focaler\text{-}IoU}}$表示 Focaler-IoU Loss,$[d,u]\in[0,1]$,通过调整 d 和 u 的值,可以使 $L_{\mathrm{Focaler\text{-}IoU}}$关注不同的回归样本.

2.5　改进的 YOLOv8n 模型结构

改进的 YOLOv8n 模型结构如图 7 所示.

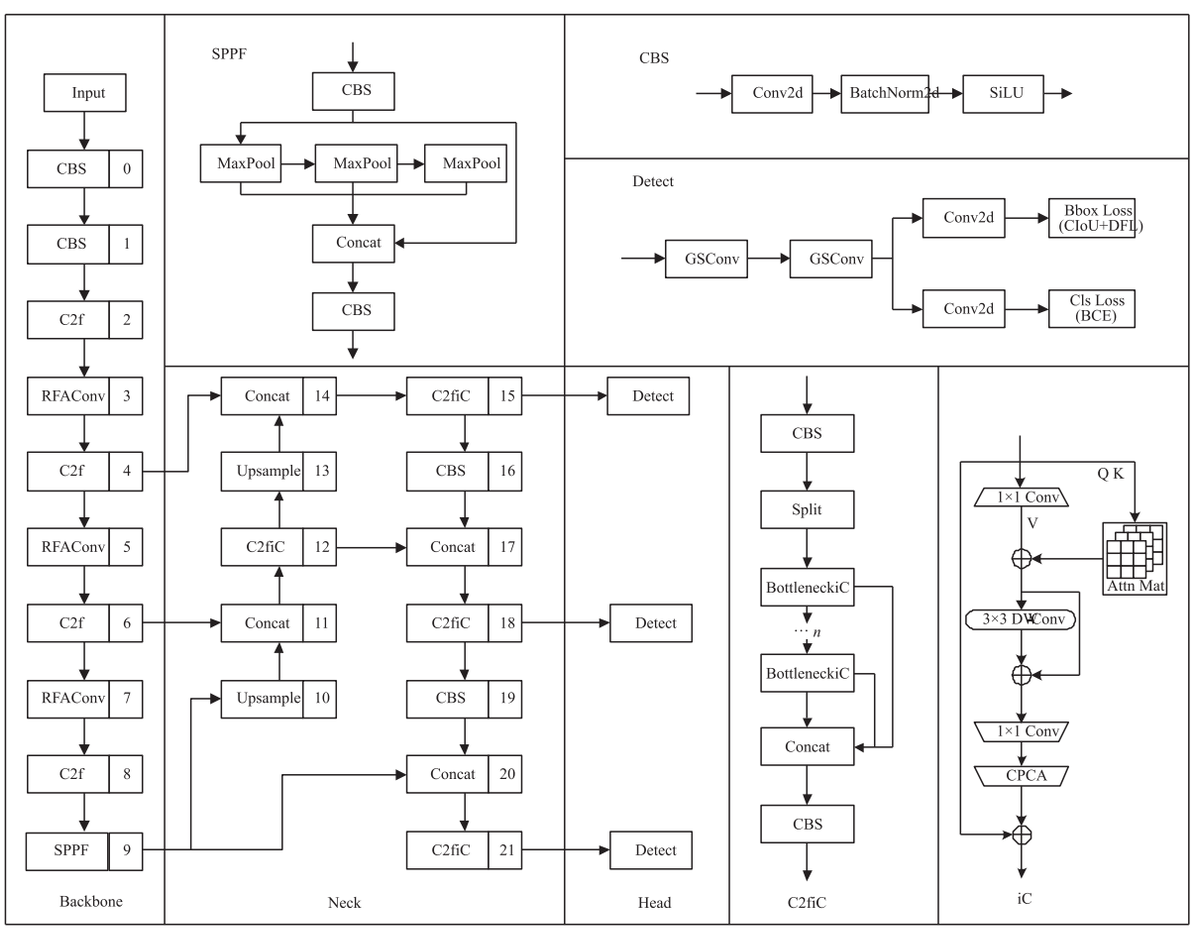


图 7　改进的 YOLOv8n 模型结构

Fig. 7　Improved YOLOv8n model structure

3　实验结果与分析

3.1　实验环境与参数配置

操作系统为 Windows11,CPU 为 Intel(R)Core(TM)i5-12400F,GPU 为 NVIDIA GeForce RTX 3060 Ti,编程语言为 Python3.9,采用 Pytorch 深度学习框架. 以 YOLOv8n 作为基准模型,初始学习率设置为 0.01,使用 SGD 优化器,batch size 设置为 8,epoch 设置为 200.

3.2　数据集

实验使用 Fire-Smoke-Dataset、Forest-Fire-Dataset 和 Huai-Gong-Fire 三个数据集. Fire-Smoke-Dataset(https://github.com/DeepQuestAI/Fire-Smoke-Dataset)为公开数据集,包含大规模和小规模火灾,总计 3 000 张图像. Forest-Fire-Dataset(https://www.kaggle.com/datasets/irdiansyah/forest-fire-dataset)为公开数据集,包括 7 167 张森林场景下的火灾图片. 鉴于 Fire-Smoke-Dataset 和 Forest-Fire-Dataset 数据集场景单一,不满足火灾场景的多样性,建立了 Huai-Gong-Fire 数据集. Huai-Gong-Fire 由 51 717 张火灾图像组成,涵盖了森林、工厂、隧道和公园等不同场景下的火灾图像. 从中随机选取 5500 张不同场景的火灾图像进行实验,火灾数据样本如图 8 所示,其中(a)~(c)为正样本,(d)~(f)为负样本. 将 Fire-Smoke-Dataset、Forest-Fire-Dataset 和 Huai-Gong-Fire 三个数据集分别按 8:1:1 的比例随机划分为训练集、验证集和测试集.

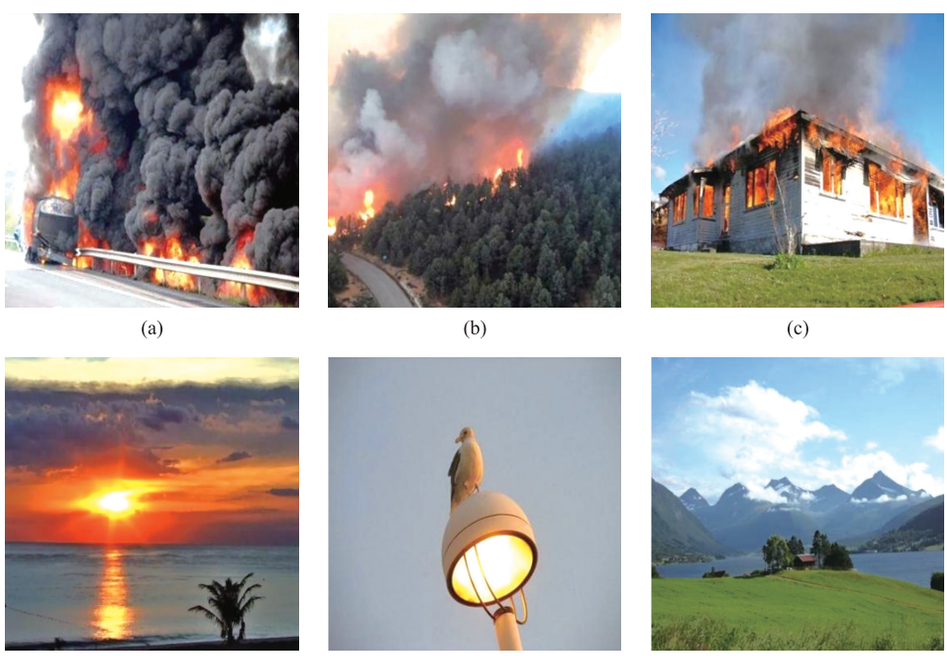


图 8　火灾数据样本

Fig. 8　Fire data samples

3.3　评价指标

采用平均精度均值(mean Average Precision,mAP)、参数量(Params)、计算量(GFLOPs)和帧速率(frames per second,FPS)对模型进行评价.

mAP 是指不同召回率下的精度均值,衡量模型在所有类别上的好坏. 计算方法如式(10)所示.

$$mAP=\frac{\sum_{i=1}^{N}AP_i}{N}, \tag{10}$$

式中,AP 表示 Precision 与 Recall 曲线与以 Precision 为纵轴和 Recall 为横轴围成的面积,N 表示类别总数.

帧速率是指模型每秒可以检测图像的数量,衡量模型的检测速度. 计算方法如式(11)所示.

$$FPS=\frac{1\ 000}{Preprocess+Inference+NMS}, \tag{11}$$

式中,$Preprocess$ 表示图像预处理时间,$Inference$ 表示推理时间,NMS 表示非极大值抑制处理时间.

3.4 消融实验

为验证改进模型对火灾的检测性能,在相同的实验环境下对 Huai-Gong-Fire 数据集进行测试,将改进模块进行不同的组合,评估各个模块对检测性能的影响,消融实验如表 1 所示. 在 YOLOv8n 主干网络使用 RFAConv 模块替换部分卷积使 mAP 提高了 1.3%,提升了模型的特征提取能力;颈部网络引入 C2fiC 模块,mAP 提高了 1.7%,参数量和计算量小幅度降低;重构检测头,设计轻量检测头,在提高 1.5%mAP 的同时,参数量和计算量分别减少了 20%和 33.3%,降低了检测头参数冗余,进一步使模型轻量化;优化损失函数,mAP 提高了 0.8%,实现难易样本的平衡. 不同模块的组合也实现了模型精度的提高以及参数量和计算量的下降,最终改进模型相比 YOLOv8n 模型 mAP 提高了 3%,参数量和计算量分别减少了 33.3%和 39.5%,证明了改进模型的有效性.

表 1 消融实验

Table 1 Ablation experiment

RFAConv	C2fiC	轻量检测头	Focaler-GIoU	mAP@0.5/%	Params/M	GFLOPs	FPS
				84.2	3.0	8.1	119
✓				85.5	3.0	8.2	108
	✓			85.9	2.6	7.5	90
		✓		85.7	2.4	5.4	**123**
			✓	85.0	3.0	8.1	120
✓	✓			86.3	2.7	7.6	75
✓	✓	✓		86.6	2.0	4.9	88
✓	✓	✓	✓	**87.2**	**2.0**	**4.9**	90

3.5 对比实验

将改进模型与目前主流模型 Faster R-CNN[14]、RT-DETR[15]、EfficientDet[16]、YOLOv5、YOLOv7[17]、YOLOv8 和 YOLOv10[18]进行对比实验,实验采用相同的数据集和训练参数,对比实验结果如表 2-表 4 所示. 改进模型在 Fire-Smoke-Dataset、Forest-Fire-Dataset 和 Huai-Gong-Fire 三个数据集上 mAP 值均达到最优,分别为 76.6%、81.0%和 87.2%,且计算复杂度最小,计算负载仅次于 EfficientDet,检测速度分别达到了每秒检测帧率 88、每秒检测帧率 84 和每秒检测帧率 90. 综上所述,改进模型实现了检测精度和检测速度的平衡,且实现了模型的轻量化,能够很好地应用于复杂场景下的火灾检测.

表 2 Fire-Smoke-Dataset 对比实验

Table 2 Comparison experiment on the Fire-Smoke-Dataset

Models	mAP@0.5/%	Params/M	GFLOPs	FPS
Faster R-CNN	69.8	136.7	369.7	50
RT-DETR-l	67.3	28.4	100.6	38
EfficientDet	70.7	3.8	**4.7**	78
YOLOv5s	70.7	7.0	15.8	**131**
YOLOv7-tiny	69.2	6.0	13.0	99
YOLOv8n	73.3	3.0	8.1	103
YOLOv10n	72.5	2.7	8.2	66
improved	**76.6**	**2.0**	4.9	88

表 3 Forest-Fire-Dataset 对比实验

Table 3 Comparison experiment on the Forest-Fire-Dataset

Models	mAP@0.5/%	Params/M	GFLOPs	FPS
Faster R-CNN	37.2	136.7	369.7	7
RT-DETR-l	77.8	28.4	100.6	38
EfficientDet	52.3	3.8	**4.7**	9
YOLOv5s	78.7	7.0	15.8	**129**
YOLOv7-tiny	73.0	6.0	13.0	123
YOLOv8n	78.5	3.0	8.1	106
YOLOv10n	76.8	2.7	8.2	125
improved	**81.0**	**2.0**	4.9	84

表 4　Huai-Gong-Fire 对比实验

Table 4　Comparison experiment on the Huai-Gong-Fire

Models	mAP@0.5/%	Params/M	GFLOPs	FPS
Faster R-CNN	61.6	136.7	369.7	10
RT-DETR-l	82.6	28.4	100.6	42
EfficientDet	79.8	3.8	**4.7**	11
YOLOv5s	82.5	7.0	15.8	**130**
YOLOv7-tiny	83.1	6.0	13.0	127
YOLOv8n	84.2	3.0	8.1	119
YOLOv10n	84.5	2.7	8.2	105
improved	**87.2**	**2.0**	4.9	90

4　结论

针对目前火灾检测效果不佳,无法平衡速度和精度的问题,提出一种增强感受野特征的多尺度火灾检测方法,提高了特征提取能力,促进信息之间的全局交互,防止过拟合,实现了检测精度和检测速度的平衡,提高了检测性能. 在未来的研究工作中,计划将改进模型部署到边缘设备上,进一步优化模型实现火灾场景中的实际应用.

[参考文献]

[1] 李倩,岳亮. 吸气式感烟火灾探测器设计改进研究[J]. 消防科学与技术,2021,40(11):1644-1647.

[2] 皮骏,刘宇恒,李久昊. 基于 YOLOv5s 的轻量化森林火灾检测算法研究[J]. 图学学报,2023,44(1):26-32.

[3] 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.

[4] 龚成张,严云洋,卞苏阳,等. 基于 Fast-CAANet 的火焰检测方法[J]. 南京师大学报(自然科学版),2024,47(2):109-116.

[5] WANG T,WANG J,WANG C,et al. Improving YOLOX network for multi-scale fire detection[J]. The visual computer,2023:1-13.

[6] ZHENG H,WANG G,XIAO D,et al. FTA-DETR:An efficient and precise fire detection framework based on an end-to-end architecture applicable to embedded platforms[J]. Expert systems with applications,2024,248:123394.1.

[7] ZHANG X,LIU C,YANG D,et al. RFAConv:Innovating spatital attention and standard convolutional operation[EB/OL]. (2024-3-28)[2024-5-18]. https://arxiv.org/abs/2304.03198.

[8] ZHANG J,LI X,LI J,et al. Rethinking mobile block for efficient attention-based models[C]//2023 IEEE/CVF International Conference on Computer Vision(ICCV). Paris,France:IEEE Computer Society,2023:1389-1400.

[9] HUANG H,CHEN Z,ZOU Y,et al. Channel prior convolutional attention for medical image segmentation[J]. Computers in biology and medicine,2024,178:108784.

[10] LI H,LI J,WEI H,et al. Slim-neck by GSConv:a lightweight-design for real-time detector architectures[J]. Journal of real-time image processing,2024,21(3):62.

[11] ZHENG Z,WANG P,LIU W,et al. Distance-IoU loss:Faster and better learning for bounding box regression[C]// Proceedings of the AAAI conference on artificial intelligence. AAAI:New York,2020,34(7):12993-13000.

[12] REZATOFIGHI H,TSOI N,GWAK J Y,et al. Generalized intersection over union:A metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach,CA,USA:IEEE,2019:658-666.

[13] ZHANG H,ZHANG S. Focaler-IoU:More Focused Intersection over Union Loss[EB/OL]. (2024-1-19)[2024-5-18]. https://arxiv.org/abs/2401.10525.

[14] REN S,HE K,GIRSHICK R,et al. Faster R-CNN:Towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence,2016,39(6):1137-1149.

(下转第 105 页)

[17] ZHENG X, ZHANG L, YAN L. Sparse discriminant twin support vector machine for binary classification[J]. Neural computing and applications,2022,34(19):16173-16198.
[18] WANG J, GUO S, NIE F, LI X. Local-global fuzzy clustering with anchor graph[J]. IEEE transactions on fuzzy systems, 2024,32(1):188-202.
[19] ZHU J, CHEN X, YANG H, et al. Unsupervised adaptive bipartite graph embedding[J]. IEEE transactions on knowledge and data engineering,2023,35(10):10514-10525.
[20] CHEN L, ZHONG Z. Adaptive and structured graph learning for semi-supervised clustering[J]. Information processing & management,2022,59(4):102949.
[21] 李静,陈秀宏. 基于竞争性协同表示的局部判别投影特征提取[J]. 智能系统学报,2019,14(5):974-981.
[22] SMOLA A J, SCHOLKOPF B. A tutorial on support vector regression[J]. Statistics and computing,2004,14:199-222.
[23] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science,2000,290(5500): 2323-2326.
[24] 陈聪. L1 正则化与 pinball 损失函数的极限学习机[J]. 信息技术与信息化,2023(3):37-40.
[25] KWAK N. Principal component analysis based on L1-norm maximization[J]. IEEE transactions on pattern analysis and machine intelligence,2008,30(9):1672-1680.
[26] WANG H, LU X, HU Z, ZHENG W. Fisher discriminant analysis with L1-norm[J]. IEEE transactions on cybernetics,2013, 44(6):828-842.
[27] 杨红鑫,杨绪兵,寇振宇,等. 基于 L1 范数的 k 平面聚类算法设计[J]. 南京航空航天大学学报,2019,51(5):681-686.
[28] WANG X, MA L, WANG X. Apply semi-supervised support vector regression for remote sensing water quality retrieving[C]// Proceedings of IEEE International Geoscience and Remote Sensing Symposium, Piscataway, NJ:IEEE,2010:2757-2760.
[29] HUA Z, YANG Y, QIU H. Node influence-based label propagation algorithm for semi-supervised learning[J]. Neural computing and applications,2020,33(7):2753-2768.
[30] ZHANG L, ZHOU W, CHANG P, et al. Kernel sparse representation-based classifier[J]. IEEE transactions on signal processing,2011,60(4):1684-1695.

[责任编辑:陆炳新]

(上接第 95 页)

[15] ZHAO Y, LV W, XU S, et al. Detrs beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA:IEEE,2024:16965-16974.
[16] TAN M, PANG R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. ,Seattle, WA, USA:IEEE,2020:10781-10790.
[17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, BC, Canda:IEEE,2023:7464-7475.
[18] WANG A, CHEN H, LIU L, et al. YOLOv10:Real-Time End-to-End Object Detection[EB/OL]. (2024-5-23)[2024-5-18]. https://arxiv.org/abs/2405.14458.

[责任编辑:陆炳新]