

基于多路径增强特征的早期烟雾检测算法<sup>\*</sup>

司盼召 何 丽 王宏伟 冉 腾

(新疆大学智能制造现代产业学院机械工程学院 乌鲁木齐 830017)

**摘 要:** 早期烟雾检测是及时消除火灾隐患的有效手段,然而火灾早期的烟雾尺度小且扩散形式复杂,这使得其检测极具挑战性。针对以上问题,提出了一种基于多路径增强特征的早期烟雾检测算法 MEF-YOLO。该算法采用 QA-ELAN 改进了骨干网络,实现了模型复杂度和精度兼顾优化;并设计了 FGCA 自主增强样本区域间的特征差异,以有效捕捉烟雾的空间信息;且通过 MEFAN 优化特征融合路径,实现了跨层次特征间的直接交互,有效缓解了细节信息损失;又引入 Wise-IOU 损失函数,通过权重调整机制全面考虑位置和尺度信息,进而提高其在复杂场景的鲁棒性。实验结果表明,在不同光照以及小尺度烟雾、烟雾扩散等实验场景中,本研究提出的算法对早期烟雾的检测准确率高达 92.5%,并且更具轻量化优势,参数量和 GFLOPs 分别下降了 27.5% 和 30.6%。

**关键词:** 早期烟雾;YOLOv5;多路径增强特征聚合网络;注意力机制

**中图分类号:** TP391.41;TN27 **文献标识码:** A **国家标准学科分类代码:** 520.2

## Early smoke detection algorithm based on multi-path enhanced features

Si Panzhao He Li Wang Hongwei Ran Teng

(School of Intelligent Manufacturing Modern Industry, School of Mechanical Engineering,  
Xinjiang University, Urumqi 830017, China)

**Abstract:** Early smoke detection is an effective means to eliminate fire hazards in a timely manner, but the small size and complex diffusion form of smoke in the early stage of a fire make its detection extremely difficult. To address the above problems, this paper proposes a multi-path enhanced feature-based YOLO (MEF-YOLO) early smoke detection algorithm, which adopts QA-ELAN to improve the backbone network and optimise the model complexity and accuracy, and develops FGCA to autonomously enhance the feature differences between the sampling regions to effectively capture the spatial information of the smoke. And the feature fusion path is optimised by the MEFAN, which realises the direct interaction between cross-level features and effectively mitigates the loss of detail information; and a Wise-IOU loss function is introduced, which comprehensively takes into account the position and scaling information through the weight adjustment mechanism to improve the robustness of the model in the complex scene. The experimental results show that the algorithm proposed in this paper has an accuracy of up to 92.5% for early smoke detection in experimental scenarios with different lighting and small-scale smoke and smoke diffusion, and has a lightweight advantage, with the number of parameters and GFLOPs reduced by 27.5% and 30.6%, respectively.

**Keywords:** early smoke;YOLOv5;multi-path enhanced feature aggregation network;attention mechanism

## 0 引 言

火灾初期,烟雾通常先于火焰出现,并具有较高的可观测性。因此,针对早期烟雾的检测技术研究,不仅能够抢占火灾预警与扑灭的先机,还为及时消除火灾隐患提供了重要支撑。然而,早期烟雾通常具有较小的尺度,并且在扩散后其可观测性显著下降<sup>[1]</sup>。这些因素为早期烟雾的检测带

来了巨大的挑战。

传统的图像处理方法依靠人工选取烟雾特征,如颜色、纹理、运动方向等特征,如李洪昌等<sup>[2]</sup>提出了一种通过比较帧间总有界变分值筛选烟雾区域的方法,提高了传统烟雾检测精度。然而,传统图像处理方法由于特征选择的不确定性<sup>[3]</sup>、计算阶段的复杂性,在实际场景中存在检测效果不稳定和滞后的问题,难以实现早期烟雾检测的目的。

近年来,深度学习驱动的目标检测算法逐渐成为了解决该问题的研究热点。为了有效利用烟雾密度信息,钱承山等<sup>[4]</sup>在常规的卷积中融合了 Transform 模型,提升了主干网络对于全局特征信息的感知能力,然而该模型对小尺度的烟雾检测性能不佳。针对小烟雾检测问题,郑远攀等<sup>[5]</sup>结合 BiFPN 改进 YOLOv5,引入混合注意力增强烟雾目标特征剔除冗余特征,但是该模型对于节省计算资源消耗方面的设计考虑较少。为缓解特征在网络传递过程中退化问题,王铮帅等<sup>[6]</sup>提出一种评分模块(scoring module),通过评分筛选关键特征进行融合,但是仍无法有效地应对烟雾复杂的扩散形式。为增强稀薄烟雾的特征,Masoom 等<sup>[7]</sup>通过将烟雾主成分分析 PCA 作为预处理模块去除冗余特征,提升了算法对烟雾的定位能力。Al-smadi 等<sup>[8]</sup>选取了 YOLO 系列、Fast R-CNN 系列等主流检测网络,从不同探测距离对烟雾进行检测实验对比,为本文选取最佳基线网络模型提供了重要参考。张立国等<sup>[9]</sup>为使模型轻量化,采用 GhostBottleneck 结合协调注意力的方式替换了 C3 模块,但是无法避免由参数量减少导致的检测精度下降问题。骈璐璐等<sup>[10]</sup>提出了极化自注意力机制,从空间维度和通道维度丰富特征的同时也增加了模型参数量,无法平衡模型精度和复杂度的优化。杜辰等<sup>[11]</sup>采用 RepConv 结构和 CBS 结构串联作为 PANet 的特征融合模块,增强了特征间的信息交互,但是增强方式过于单一,无法充分挖掘各自特征的特点。

此外,烟雾检测领域公开的数据集相对有限,例如 BoWFire<sup>[12]</sup>、VSD<sup>[13]</sup>、fire-dunnings-dataset<sup>[14]</sup>、USTC-RF<sup>[15]</sup>和 CVPR<sup>[16]</sup>等数据库,其中大多数烟雾样本是在大

火蔓延后采集的,烟雾浓度高、尺度大,并且混杂着火焰。这使得针对火灾早期的烟雾样本更加匮乏,同时这些样本还存在着图像质量欠佳,分辨率较低的问题。虽然 USTC-RF 数据集通过人工合成烟雾图像来扩充了数据集,但是真实性的差异也限制了算法性能的上限。

综上所述,针对早期烟雾尺度小和扩散不规则的问题,本文提出了一种基于多路径增强特征的早期烟雾检测算法 MEF-YOLO。首先引入了量化感知层聚合网络(quantization awareness-efficient layer aggregation network, QA-ELAN)提高模型表达能力,同时利用改进的重参数化策略降低对硬件的计算需求。改进特征引导协调注意力(feature guide coordinate attention, FGCA)增强样本间不同区域的特征差异,捕捉烟雾空间位置信息。提出了多路径增强特征聚合网络(multi-path enhanced feature aggregation network, MEFAN)和三重特征聚合模块(triple feature aggregation module, TFA)通过多路径、多层次的特征增强融合方式,提高对烟雾复杂纹理的表征能力,其中可变形卷积(deformable Conv, DConv)<sup>[17]</sup>对高分辨率特征进行自适应下采样,以适应烟雾的不规则扩散。最后,引入 Wise-IOU<sup>[18]</sup>利用权重调整更加全面的考虑目标位置和尺度信息,实现复杂场景下高精度、高实时性、高效率的烟雾检测。

## 1 MEF-YOLO 模型

### 1.1 网络整体结构

YOLOv5 由于其体积小、推理速度快,在目标检测领域获得了广泛应用。本文基于 YOLOv5 模型,提出了一种早期烟雾检测算法 MEF-YOLO,结构详如图 1 所示。

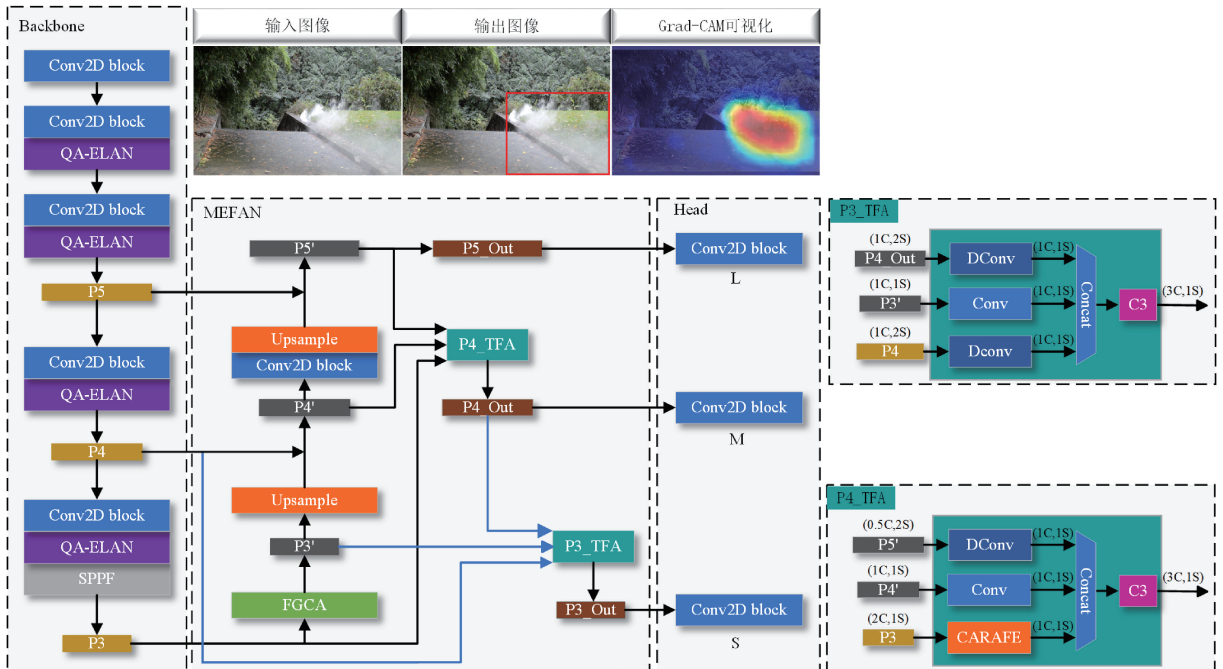


图 1 MEF-YOLO 网络结构

Fig. 1 MEF-YOLO network structure

MEF-YOLO 网络由 3 部分构成: 骨干网络 (backbone)、多路径增强特征聚合网络 (MEFAN) 和检测头 (head)。骨干网络包含卷积块 (Conv2D block)、量化感知层聚合网络 (QA-ELAN)、SPPF, Conv2D block 通过卷积操作提取出图像不同尺度的特征。QA-ELAN 替换了原来的 C3 模块, 旨在利用特征的分割重组策略, 丰富梯度流的同时保证轻量化。SPPF 的作用是处理不同尺度的输入, 增强模型的表达能力和计算效率。MEFAN 的改进主要为: 利用共享卷积的特性提升目标和背景区域差异, 并在相邻层间自顶向下和自底向上传递特征的基础上, 增加非相邻层间的特征路径, 丰富语义和位置信息。检测头使用损失函数 Wise-IOU 和非极大值抑制算法筛选出合格的锚框, 完成目标分类和定位。

### 1.2 QA-ELAN 量化感知层聚合网络改进

现有 YOLOv5 改进策略大多仅注重检测精度或模型复杂度的单一优化, 针对这一问题, 本文在广义高效层聚合网络 (generalized efficient layer aggregation network, GELAN)<sup>[19]</sup> 的基础上设计了 QA-ELAN, 并替代 YOLOv5 骨干网络中的 C3 模块, 实现了模型复杂度和精度兼顾优化。QA-ELAN 结构如图 2 所示, 利用分割重组理念, 将 CSPNet 与 ELAN 的结构优势整合在一起, 提升特征丰富性。然后, 又采用重参数化和模型量化的策略, 在分支中嵌入了 QARepConv 模块, 降低模型参数量。

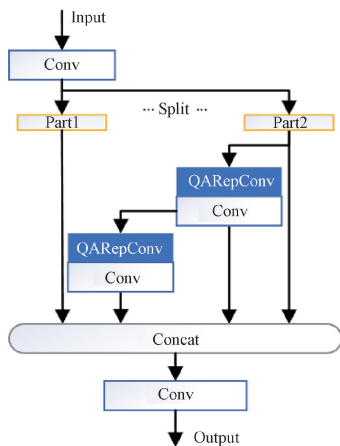


图 2 QA-ELAN 网络结构

Fig. 2 QA-ELAN network structure

QA-ELAN 利用 GELAN 框架深度整合了 QARepConv<sup>[20]</sup>, 输入特征  $X$  先经过一次常规卷积操作:

$$Y = \text{Conv}(X) \quad (1)$$

接着, 将  $Y$  通过 Split 操作分为两个部分  $Y_1$  和  $Y_2$ , 分支  $Y_2$ , 分别经过 3 个不同分支进行特征提取, 输出分别为  $Y_{21}, Y_{22}, Y_{23}$ 。最后, 将所有分支的特征进行组合:

$$\begin{cases} Y_{21} = Y_2 \\ Y_{22} = \text{Conv}(\text{QARepConv}(Y_2)) \\ Y_{23} = \text{Conv}(\text{QARepConv}(Y_{22})) \end{cases} \quad (2)$$

$$Y_{\text{output}} = \text{Conv}(\text{Concat}(Y_1, Y_{21}, Y_{22}, Y_{23})) \quad (3)$$

### 1.3 FGCA 特征引导-协调注意力模块设计

目前, 在 YOLOv5 中引入常规注意力机制的改进策略, 往往由于空间和通道维度上的信息交互不足, 导致其关注点局限于特征的局部信息, 从而限制检测性能的提升。协调注意力 (coordinate attention, CA)<sup>[21]</sup> 弥补了这一缺陷, 通过水平方向和垂直方向分别编码空间信息到通道权重中, 不仅捕获了全局长距离依赖, 又保留了细粒度的空间信息。本文在 CA 的基础上提出了特征引导-协调注意力 FGCA, 其主要改进为: 引导模块先在空间维度上对输入特征  $X$  进行语义相似区域划分, 在不同区域采用不同卷积核进行二维卷积, 突出不同区域间的信息差异。随后输入到 CA 注意力中, 提升其对烟雾区域的关注度。

首先, 将特征图语义相似的区域划分为单独的共享区域, 具体过程如图 3 所示。输入特征  $X$  经过  $3 \times 3$  的卷积输出引导特征  $F$ , 如式 (4) 所示。其中,  $K_{3,3,c}$  代表  $c$  个通道数  $3 \times 3$  的卷积核,  $j$  代表引导特征  $F$  的通道。

$$F_{u,v,j} = \sum_{c=1}^C X_{u,v,c} * K_{3,3,c,j}, j \in \{0,1\} \quad (4)$$

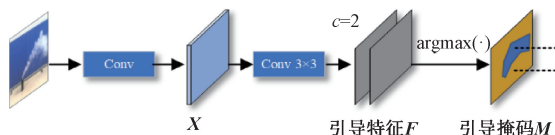


图 3 共享区域划分过程

Fig. 3 Shared area division process

引导特征  $F$  在通道维度上进行数值比较, 输出最大值的通道索引, 得到引导掩码  $M$ , 对于整个空间维度中每个位置  $(u, v)$ , 引导掩码  $M$  的计算公式为:

$$M_{u,v} = \text{argmax}(F_{u,v,0}, F_{u,v,1}) \quad (5)$$

式中:  $\text{argmax}()$  表示输出最大值的索引,  $F_{u,v,0}$  和  $F_{u,v,1}$  分别表示位置  $(u, v)$  处通道 0 和通道 1 上引导特征的值。因此, 引导掩码  $M$  被分为值为 0 和值为 1 的两个区域, 在共享区域 0 中, 共享使用卷积核  $\text{Kernel } 0$  进行正常的二维卷积操作, 在共享区域 1 中, 共享使用卷积核  $\text{Kernel } 1$  进行正常卷积操作, 特征引导过程如图 4 所示。

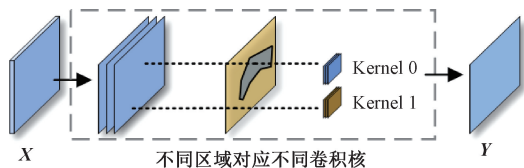


图 4 特征引导原理图

Fig. 4 Schematic diagram of feature guidance

从图像语义角度来看, 语义相似的特征会被分配到同一区域, 输出引导后特征  $Y$  的第  $n$  个通道特征计算公式如式 (6) 所示, 其中  $t$  为共享区域,  $(u, v)$  是区域  $t$  中的点之一,  $K_{t,c}$  表示区域  $t$  中的共享卷积核  $K_t$  的第  $c$  通道。

$$Y_{u,v,n} = \sum_{c=1}^C X_{u,v,c} * K_{t,c}, (u,v) \in t, t \in \{0,1\} \quad (6)$$

最后,设计特征引导-协调注意力 FGCA 的网络结构,结构示意图如图 5 所示,FG 为特征引导。经过引导的特征图  $Y$  分别进行水平和垂直方向上的平均池化,在空间方向上聚合特征,以此捕获长距离的依赖关系,其中特征图高为  $H$ ,宽为  $W$ ,所以输出的第  $c$  通道的高  $h$  可以表示为一维特征,如式(7)所示。

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} Y_c(h,i) \quad (7)$$

第  $c$  通道的宽  $w$  可以表示为一维特征,如式(8)所示。

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} Y_c(j,w) \quad (8)$$

假如输入的特征层形状为  $[C, H, W]$ ,经过水平方向平均池化得到  $[C, H, 1]$  的特征层,经过垂直方向平均池化的到  $[C, 1, W]$  的特征层。再进行串联和卷积操作,使得位置方向信息交互融合,其中系数  $r$  为缩减比率。卷积后进行批归一化和非线性激活函数,其计算式为:

$$f = \delta(F_1([Z^h, Z^w])) \quad (9)$$

式中:  $f \in \mathbf{R}^{c/r \times (H+W)}$ ,  $[\cdot]$  为串联操作,  $F_1$  为卷积函数,  $\delta$  为非线性激活函数。接着将特征图  $f$  分割,分别进行卷积和激活得到两个向量  $g_h$  和  $g_w$ ,分别提取水平方向和垂直方向的全局上下文信息,定位目标对象的位置,最后将两个方向上的特征图相乘,再乘上自适应池化后的输入特征重新计算输出。

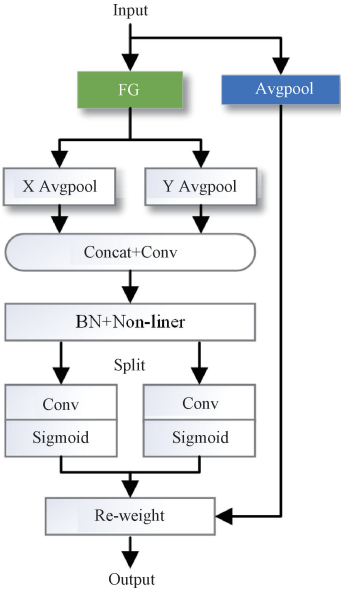


图 5 特征引导-协调注意力

Fig. 5 Characteristic guidance-coordinated attention

#### 1.4 MEFAN 多路径增强特征聚合网络设计

传统的特征金字塔网络,如 PANet、BiFPN、NAS-FPN 等结构的主要改进在于通过优化信息流动路径来改善梯度消失问题,对不同路径中的特征进行针对性增强的考虑

较少。针对这一问题,本文在设计了 MEFAN 结构时,如图 6 所示。针对高分辨率的细节特征提取和小尺度特征重建进行了增强处理,此外,还增加了跨层次的特征传递路径,使得特征信息更加完整丰富,例如 P3 和 P4 的特征可以分别直接传递到 P4\_TFA 和 P3\_TFA,同时 P5' 的特征也传递到 P4\_TFA。

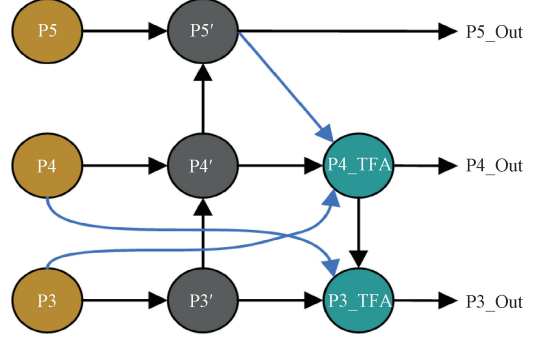


图 6 多路径增强特征聚合网络

Fig. 6 Multi-path enhanced feature aggregation network

三重特征聚合模块 TFA 的主要作用是对 MEFAN 中不同分支的特征进行增强和融合,使模型能够更好地整合各层次地信息, P4\_TFA 和 P3\_TFA 结构如图 7 所示。

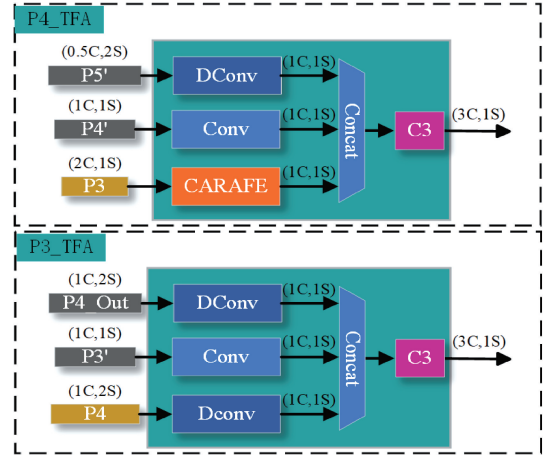


图 7 P4\_TFA 和 P3\_TFA 模块

Fig. 7 P4\_TFA and P3\_TFA modules

由于烟雾作为非刚体检测目标,所以 TFA 采用 DConv 来连接高分辨率的特征图,通过动态地调整卷积核的位置适应扩散后边缘和纹理的变化。相比传统的卷积或池化方法, DConv 能够保留底层更多的高分辨率特征,以此捕捉到图像中小尺度烟雾的细节信息。TFA 模块的 DConv 分支的输出特征  $Y_{DC}$  公式为:

$$Y_{DC} = \sum_{j=1}^N \alpha_j(X) \times (X * W_j) \quad (10)$$

式中:  $X * W_j$  表示输入  $X$  与第  $j$  个卷积核  $W_j$  的卷积操作。  $N$  是动态生成卷积核个数,  $\alpha_j(X)$  是输入自适应地权重表示输入  $X$  对第  $j$  个卷积核  $W_j$  的加权系数。



对于中层特征, TFA 采用  $1 \times 1$  常规卷积的分支来改变其通道维度, 进而控制三层特征拼接时的信息贡献度, 常规卷积分支输出表示为  $\mathbf{Y}_{Conv}$ 。

而对于小尺寸的深层特征图, 传统的 Upsample 上采样层依赖于最近邻像素值, 难以捕捉图像的细微变化和密集语义信息, 增加了漏检风险。相比之下, CARAFE<sup>[22]</sup> 通过内容感知的重组方式生成上采样核, 能够对特征图进行更精细的插值重建, 不仅考虑了局部区域的特征, 而且可以整合更大范围的上下文信息, 从而更精准的还原原图的细节。这种上采样机制可以有效地减少烟雾由于特征细微导致的空间信息损失, 降低漏检情况。TFA 模块的 CARAFE 分支的处理过程为:

$$\mathbf{Y}_{CARAFE}^l = \sum_{n=-\frac{k_{up}}{2}}^{\frac{k_{up}}{2}} \sum_{m=-\frac{k_{up}}{2}}^{\frac{k_{up}}{2}} \mathbf{W}_l(n, m) \times \mathbf{X}(i+n, j+m) \quad (11)$$

式中:  $\mathbf{Y}_{CARAFE}^l$  表示上采样后特征图位置  $l$  的特征值;  $\mathbf{W}_l(n, m)$  是上采样核在位置  $(n, m)$  的权重;  $\mathbf{X}(i+n, j+m)$  是输入特征图中对应位置的像素值;  $k_{up}$  表示上采样核的大小。

最后将尺寸调整后的 3 个特征图进行一次卷积, 在通道尺寸上进行拼接, 公式为:

$$\mathbf{F}_{TFA} = \text{Concat}(\mathbf{F}_l, \mathbf{F}_m, \mathbf{F}_s) \quad (12)$$

式中:  $\mathbf{F}_{TFA}$  表示 TFA 模块输出的特征图,  $\mathbf{F}_l$ 、 $\mathbf{F}_m$  和  $\mathbf{F}_s$  是代表大、中、小不同尺寸的特征图, 即分别对应  $\mathbf{Y}_{DC}$ 、 $\mathbf{Y}_{Conv}$  和  $\mathbf{Y}_{CARAFE}$ 。

### 1.5 Wise-IoU 损失函数

传统的损失函数主要侧重于精确匹配目标边界, 在处理模糊和不规则边界的目标(如烟雾)时, 较大的回归误差会导致样本不平衡问题, 即回归误差大的低质量样本远多于误差小的低质量样本。原 YOLOv5 使用 CIoU 未考虑训练样本的不平衡问题, 这会导致低质量样本在梯度计算中占主导地位, 造成损失函数的剧烈波动。为了降低低质量样本对梯度的影响, 并提高模型对扩散烟雾和小尺度烟雾等困难样本的学习能力, 本文引入了动态非单调聚焦机制的 WIoU 来替换原损失函数, 计算如式(13)所示。

$$\begin{cases} L_{WIoU} = rR_{WIoU}L_{IoU} \\ R_{WIoU} \in [1, e) \\ L_{IoU} \in [0, 1] \end{cases} \quad (13)$$

式中: 距离聚焦机制  $R_{WIoU}$  用于放大普通质量锚框的  $L_{IoU}$  非单调聚焦系数  $r$  用于聚焦普通质量锚框。  $R_{WIoU}$  定义如式(14)所示,  $r$  定义如式(15)所示。

$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (14)$$

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (15)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (16)$$

式中:  $\beta$  表示离群度, 其中  $\overline{L_{IoU}}$  表示动态滑动平均值, 在值大或值小时都分配较小的梯度增益, 降低对边界框回归的影响。  $\alpha, \delta$  表示超参数。  $r$  通过降低高质量样本对损失值贡献, 动态地为边界框赋予梯度增益, 并在训练后期减少低质量锚框产生的有害梯度, 聚焦于普通质量的锚框, 提高模型的定位能力。 WIoU 去除了 CIoU 中的纵横比惩罚项, 同时平衡了高质量和低质量锚框对模型回归的影响, 增强了模型的泛化能力, 提高了整体性能。 因此, 本文采用 Wise-IoU 替换原始模型的边界损失函数, 以优化模型的损失函数。

## 2 实验环境与数据

### 2.1 实验数据准备

通过对公开数据集进行筛选, 从烟雾样本较为丰富的 USTC-RF 和 CVPR 数据库中筛选了一部分优质早期烟雾样本, 作为早期烟雾数据集的数据来源之一。 在此基础上, 补充了通过自行实地拍摄和网络渠道获取的 1 246 张图片, 并使用矩形框和标签进行手动注释。

构建的早期烟雾数据集如图 8 所示, 共采集场景 23 个, 图片 2 561 张, 共包括目标对象小尺度烟雾图像 1 187 张, 烟雾扩散图像 1 062 张。 采用 8:2 的比例划分数据集, 1 920 张图片作为训练集, 641 张图片作为验证集。



图 8 早期烟雾数据集部分图片

Fig. 8 Partial image of the early smoke dataset

### 2.2 试验训练环境与训练方法

本实验训练环境采用的计算机硬件配置包括: 中央处理器 Intel Core i5-12400F, 内存 16 GB, 显卡为 GeForce RTX 3060 12 GB 显存, 采用 Pytorch 深度学习框架。

训练参数: 彼此大小  $\text{batch\_sizes}$  设为 16, 迭代次数 epoch 为 250, 其余参数为默认值。

### 2.3 性能评价指标

为综合考虑烟雾检测网络的性能, 本文采用精确率(P)、召回率(R)、平均精度(AP)等指标衡量模型的检测精度, 采用浮点运算次数(GFLOPs)和参数量(params)来衡量模型的复杂度, 其计算公式为:

$$P = \frac{TP}{TP + FP} \quad (17)$$

$$R = \frac{TP}{TP + FN} \quad (18)$$

$$AP = \int_0^1 p(r) dr \quad (19)$$

式中:  $TP$  表示真正例,  $FP$  表示假正例,  $FN$  表示假负例。 以召回率为横轴、精确率为纵轴就可以绘制  $P(R)$  曲线, 对

曲线进行积分可得到 AP,  $AP_{0.5}$  表示交并比取值为 0.5 时的 AP 值。

3 实验结果与分析

3.1 不同模块对比实验

为验证 Q-ELAN 改进骨干网络的有效性,选取 DarkNet53、MobileNet、GELAN 作为对比对象,实验结果如表 1 所示。

表 1 不同骨干网络检测性能对比

Table 1 Comparison of detection performance of different backbone networks

骨干网络	$AP_{0.5}$	params/M	FPS/fps
DarkNet53	0.891	7.2	53.3
MobileNet	0.847	6.2	61.5
GELAN	0.895	4.7	47.4
Q-ELAN(本文)	0.901	4.3	62.8

由表 1 数据分析可得,QA-ELAN 引入模型后  $AP_{0.5}$  为 0.901、params 为 4.3 M、FPS 为 62.8,综合效果最好,实现了模型的检测效果和模型参数量同时优化。

为验证 FGCA 注意力、MEFAN 及损失函数 Wise-IOU 的有效性,分别选取不同注意力机制、不同 Neck 改进及不同损失函数进行了检测性能对比,实验结果分别为表 2~4 所示。

表 2 不同注意力检测性能对比

Table 2 Comparison of different attention detection performance

注意力	$AP_{0.5}$	$P$	$R$
SE	0.872	0.842	0.811
CBAM	0.883	0.856	0.833
CA	0.895	0.870	0.842
FGCA(本文)	0.903	0.884	0.849

由表 2 数据分析可得,FGCA 引入后使模型的检测精度和召回率达到了最优,并且相较于原始 CA 指标提升也十分明显。这表明,预先对特征进行引导的策略赋予了 FGCA 模块更高效的烟雾检测能力。

表 3 不同 Neck 改进检测性能对比

Table 3 Comparison of different Neck improvement detection performance

Neck	$AP_{0.5}$	$P$	$R$
PANet	0.891	0.867	0.836
BiFPN	0.879	0.841	0.827
NAS-FPN	0.892	0.849	0.846
MEFAN(本文)	0.911	0.904	0.868

表 4 不同损失函数检测性能对比

Table 4 Comparison of detection performance of different loss functions

损失函数	$AP_{0.5}$	$P$	$R$
CIOU	0.891	0.867	0.836
SIOU	0.879	0.869	0.841
Inner-IOU	0.872	0.851	0.818
Focal Loss	0.895	0.876	0.843
Wise-IOU(本文)	0.909	0.877	0.854

由表 3 可知,MEFAN 的针对性特征增强效果最好,相较于 YOLOv5 的 PANet 在指标  $AP_{0.5}$ 、 $P$  和  $R$  上分别增长了 2%、3.7%和 3.2%。由表 4 可知,引入 Wise-IOU 后,模型的综合指标最好,其中召回率的显著提升可有效环节小烟雾的漏检问题。

3.2 消融实验

根据消融实验结果表 5 所示,引入 QA-ELAN 改进的骨干网络极大地降低了模型参数量,同时检测精度和召回率也得到了提升。分析原因在于 QA-ELAN 实现了不同层特征的有效融合,增强了特征表示能力,同时利用 CSP 结构的跨阶段连接和 QARepConv 的量化感知训练策略,减少了计算成本并提高了模型的鲁棒性。

为了更直观的展示添加模块的效果,本文使用 GradeCAM<sup>[23]</sup>生成热力图来显示模型关注区域的变化。如图 9 所示,FGCA 模块使得热点覆盖区域更加集中,降低了模型对背景的关注度,这是因为烟雾特征经过引导模块预处理后,划分并增强了不同的语义区域,提升了目标区域和背景区域的差异度,使得模型获得了更强的空间位置定位能力。

在此基础上,模型 Neck 部分通过 MEFAN 进行特征路径优化和 TFA 加权增强,使得模型对烟雾深层语义信息和浅层细节信息的融合利用更高效,提升了模型对烟雾复杂纹理的表征能力,大幅提升了模型的  $AP_{0.5}$ 、 $P$  和  $R$ ,分别为 1.7%、1.0%和 1.9%,有效缓解了模型的精度较差和漏检问题。最后,损失函数 Wise-IOU 的引入,使得模型训练后期减少了小尺度烟雾和烟雾产生扩散后的低质量样本产生的有害梯度,同时平衡了高质量和低质量锚框对于模型回归的影响,提高了模型的整体性能,相较引入之前  $AP_{0.5}$ 、 $P$  和  $R$  分别提升了 0.6%、0.4%和 1.7%。

3.3 模型对比试验

为验证本文算法的优越性,将本文算法 MEF-YOLO 与 YOLO 系列、Faster R-CNN<sup>[24]</sup>、SSD<sup>[25]</sup>、FCOS<sup>[26]</sup>等通用目标检测算法,以及与文献[5]、文献[10]和文献[27]提出的最新烟雾检测算法,在早期烟雾数据集上分别进行对比试验。实验以精确度  $P$ 、召回率  $R$ 、 $AP_{0.5}$ 、模型参数量、GFLOPs、FPS 和模型体积作为评价指标,实验结果如表 6 所示。

表 5 MEF-YOLO 的消融实验结果

Table 5 Results of ablation experiments with MEF-YOLO

QA-ELAN	FGCA	MEFAN	WIoU	$AP_{0.5}$	$P$	$R$	params/M	GFLOPs/G
—	—	—	—	0.891	0.867	0.836	7.2	16.0
✓	—	—	—	0.901	0.892	0.845	<b>4.3</b>	<b>10.7</b>
✓	✓	—	—	0.908	0.898	0.841	<b>4.3</b>	<b>10.7</b>
✓	✓	✓	—	0.919	0.908	0.857	5.1	11.1
✓	✓	✓	✓	<b>0.925</b>	<b>0.912</b>	<b>0.873</b>	5.1	11.1

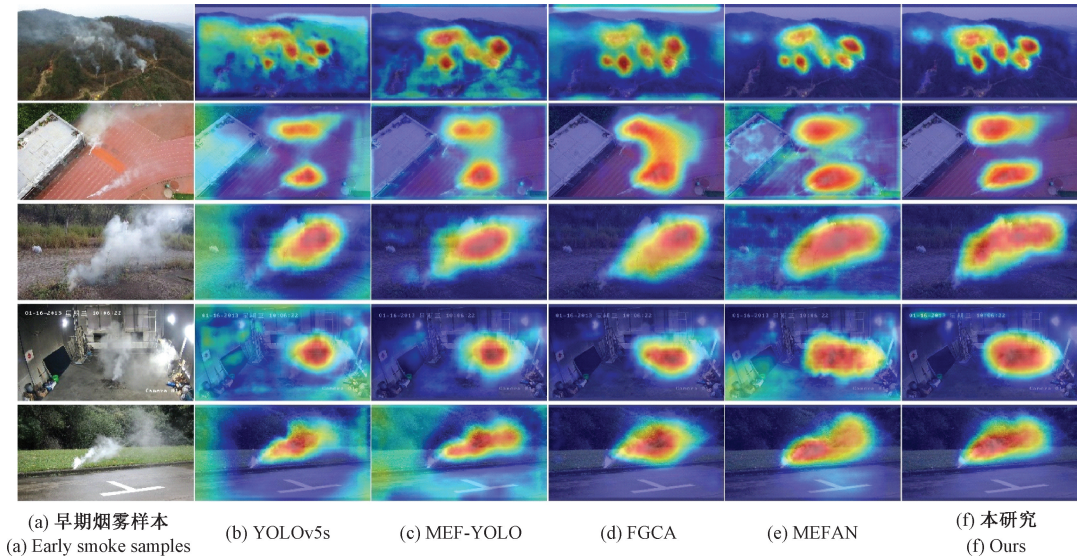


图 9 消融实验热力图可视化

Fig. 9 Visualisation of the ablation experiment heat map

表 6 不同模型的实验结果对比

Table 6 Comparison of experimental results for different models

类别	模型	$AP_{0.5}$	$P$	$R$	params/M	GFLOPs/G	FPS/fps	Size/MB
通用模型	Faster R-CNN	0.772	0.739	0.551	60.3	195.3	16.4	108.0
	SSD	0.246	0.243	0.254	25.4	215.3	58.5	91.1
	FCOS	0.844	0.869	0.77	32.12	125.6	19.6	122.5
	YOLOv5s	0.891	0.867	0.836	7.2	15.8	53.3	13.7
	YOLOX-s	0.909	0.872	0.854	9.0	26.8	<b>78.5</b>	34.4
	YOLOv8s	0.895	0.895	0.837	11.1	28.6	70.3	23.5
	RT-DETR- <sup>[28]</sup>	0.897	0.875	0.786	31.9	110	42.9	92.0
	YOLOv9s	0.893	0.904	0.844	7.2	26.8	62.2	14.3
	YOLOv10s <sup>[29]</sup>	0.909	0.892	0.851	8.0	24.4	69.7	15.8
烟雾模型	文献[5]	0.891	0.908	0.897	8.57	23.2	46.5	17.2
	文献[10]	0.866	0.843	0.825	8.3	19.4	51.6	15.8
	文献[21]	0.923	0.901	0.817	8.9	26.7	64.2	69.5
	MEF-YOLO(本文)	<b>0.925</b>	<b>0.912</b>	<b>0.873</b>	<b>5.1</b>	<b>11.1</b>	<b>55.4</b>	<b>11.2</b>

由表 6 可知,通用算法之间相比,SSD、FCOS、Faster R-CNN 等模型在检测效果和计算资源消耗方面表现均较差,YOLO 系列检测框架在资源利用效率上更具优势,是更理想的选择。其中本文提出的 MEF-YOLO 综合检测性



能最佳,在  $AP_{0.5}$ 、 $P$ 、 $R$  等指标上均超越了最先进的通用目标检测模型,甚至是 YOLOv10。

与最新且具有代表性的烟雾检测模型相比,MEF-YOLO 在检测精度和模型复杂度之间实现了兼顾优化。本文提出的方法  $AP_{0.5}$  为 0.925、精确度( $P$ )为 0.912、召回率( $R$ )为 0.873,达到最优水平;同时,其参数数量和 GFLOPs 相比文献[21]提出的先进方法分别减少了 3.8 M 和 15.6 G,模型体积最小。这使得 MEF-YOLO 在早期烟雾预警和监测服务方面展现出更显著的优势。

### 3.4 检测结果对比

为对比模型改进前后的检测效果,使用原始 YOLOv5s 和改进 MEF-YOLO 进行整体实验测试,如图 10 所示。实验场景包括:图 10(a)、(b)正常光照、图 10(c)光照强烈、图 10(d)光线昏暗。

图 10 中,(a)组实验结果:YOLOv5s 算法存在明显的漏检问题,MEF-YOLO 不仅能够稳定地定位到烟雾的位置,而且在不同的环境和条件下都能保持较高的检测准确率。(b)组实验结果:两种算法都正常检测到烟雾,MEF-YOLO 不仅定位框更加精准而且置信度更高,分别为 0.88、0.9、0.77 和 0.84。(c)组实验结果:光线过强的区域会出现白色亮斑,会掩盖烟雾的边缘和内部结构。在第 1 幅图中,图像过曝后,YOLOv5s 将光源错误地识别为烟雾,在第 2、3 幅图中,YOLOv5s 受到了光源的强烈干扰,无法捕捉到烟雾的边缘信息,定位精度大幅下降。(d)组图像的对比较低,这会削弱烟雾的可见性,使其纹理细节隐藏在背景中。在第 1、3 幅图中,由于图像亮度整体偏低,YOLOv5s 难以捕捉到烟雾的细微纹理信息,出现了漏检现象。

综上,MEF-YOLO 无论是应用于光线强烈还是昏暗的场景,检测框的准确度已经有了明显的改善,对烟雾细微特征的提取具有更强的鲁棒性和稳定性。

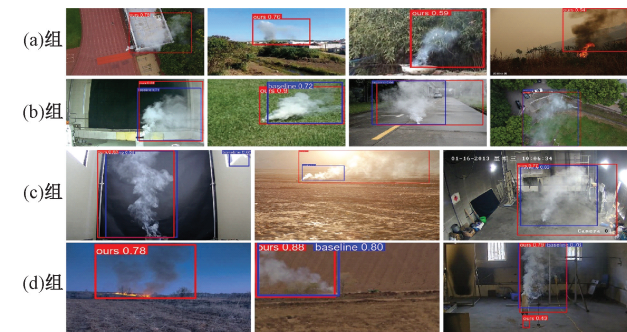


图 10 YOLOv5 和 MEF-YOLO 在复杂场景检测效果对比

Fig. 10 Comparison of YOLOv5 and MEF-YOLO in complex scene detection effect

### 3.5 可视化对比

为直观展示 MEF-YOLO 在应对早期烟雾尺度较小及扩散形式复杂场景中的能力,借助 Grad-CAM 算法[18]生成

热力图。图 11 所示为扩散烟雾的热力图,场景 A 为两处烟雾扩散的场景、场景 B 为近处视角、场景 C 为光线干扰的远处视角。

从图 11 中可以看出:在场景 A 中,YOLOv5 受到相近颜色的背景干扰,对灰色房顶区域表现出了很强的注意力,而 MEF-YOLO 可以精准地关注到左下角和右上角两处产生扩散的早期烟雾;在场景 B 中,MEF-YOLO 对烟雾轮廓更加敏感,能够关注到烟雾十分稀薄的区域,YOLOv5s 对烟雾扩散区域表现出的关注力较弱;在场景 C 中,在逆光情况下 YOLOv5 受到了严重干扰,模型注意区域比较分散,而 MEF-YOLO 的热力图在烟雾区域高亮且集中。

图 12 所示为小尺度烟雾热力图,共选取 A、B 和 C 三个场景,场景 A 为多处小尺度烟雾的远视角,场景 B 为近处视角,场景 C 为分辨率较差的模糊视角场景。

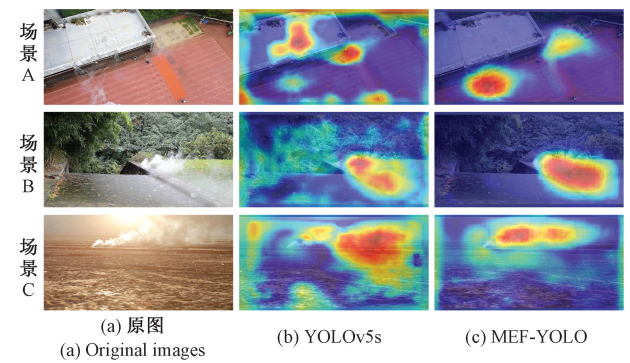


图 11 扩散烟雾热力图可视化图

Fig. 11 Heat map visualisation of diffuse smoke

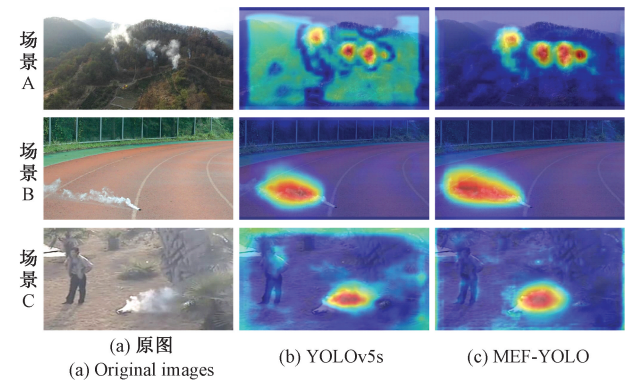


图 12 小尺度烟雾热力图可视化

Fig. 12 Visualisation of small-scale smoke thermograms

从图 12 中可以看出:在场景 A 中,MEF-YOLO 的关注区域精准覆盖到了每个烟雾的产生位置,YOLOv5 在全局范围内无法获取到关键性特征,在多处非目标区域都表现出了注意力。在场景 B 和 C 中,MEF-YOLO 相较于 YOLOv5s 提取的烟雾特征更加精确,热力图高亮颜色更深,覆盖到的烟雾区域也更加完整。



## 4 结 论

为应对早期烟雾尺度小、扩散不规则的问题,本文提出了一种基于多路径增强特征聚合的早期烟雾检测算法 MEF-YOLO,能够同时提升模型的精度和效率。其中,FGCA 模块通过学习样本区域间的特征差异,增强了模型的空间定位能力。MEFAN 模块采用多路径、多层次的增强融合方式,提升了对烟雾边缘的适应性和复杂纹理的表征能力。模型对比实验表明,改进后的方法在占用最少计算资源的情况下,对早期烟雾的检测精度最高, $AP_{0.5}$  为 92.5%。可视化对比实验表明,MEF-YOLO 对复杂场景下小尺度烟雾及扩散后烟雾的检测效果有显著提升,为火灾预警领域提供了一种高效、可靠的解决方案。

## 参考文献

- [1] 朱弥雪, 刘志强, 张旭, 等. 林火视频烟雾检测算法综述[J]. 计算机工程与应用, 2022, 58(14):16-26.  
ZHU M X, LIU ZH Q, ZHANG X, et al. Review of research on video-based smoke detection algorithms [J]. Journal of Computer Engineering & Applications, 2022, 58(14): 16-26.
- [2] 李洪昌, 安明伟. 基于总有界变分的森林火灾烟雾图像检测方法[J]. 电子测量与仪器学报, 2020, 34(11): 211-217.  
LI H CH, AN M W, Smoke image detection method of the forest fire based on total bounded variation[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(11): 211-217.
- [3] 高洁, 王战红, 刘纲. 基于 FSSD 的微光烟雾检测方法[J]. 电子测量技术, 2021, 44(5):123-128.  
GAO J, WANG ZH H, LIU G, Low light level smoke detection method based on FSSD[J]. Electronic Measurement Technology, 2021, 44(5): 123-128.
- [4] 钱承山, 沈有为, 孙宁, 等. 基于 Transformer 改进 YOLOv5 的山火检测方法研究[J]. 电子测量技术, 2023, 46(16):46-56.  
QIAN CH SH, SHEN Y W, SUN N, et al. Research on improved YOLOv5 forest fire detection method based on Transformer[J]. Electronic Measurement Technology, 2023, 46(16): 46-56.
- [5] 郑远攀, 许博阳, 王振宇. 改进的 YOLOv5 烟雾检测模型[J]. 计算机工程与应用, 2023, 59(7):214-221.  
ZHENG Y P, XU B Y, WANG ZH Y. Improved YOLOv5 smoke detection model [J]. Computer Engineering and Applications, 2023, 59(7): 214-221.
- [6] 王铮帅, 邱联奎, 李迎港. 复杂环境下的 YOLOv5s 烟火检测方法[J]. 电子测量技术, 2023, 46(24): 149-156.

- WANG ZH SH, QIU L K, LI Y G. Smoke and flame detection method with YOLOv5s in complex environment [J]. Electronic Measurement Technology, 2023, 46(24): 149-156.
- [7] MASOOM S M, ZHANG Q X, DAI P W, et al. Early smoke detection based on improved YOLO-PCA network[J]. Fire, 2022, 5(2): 40.
- [8] AL-SMADI Y, ALAUTHMAN M, AL-QEREM A, et al. Early wildfire smoke detection using different yolo models[J]. Machines, 2023, 11(2): 246.
- [9] 张立国, 张琦, 金梅, 等. 一种改进 YOLOv5s 的森林火灾烟雾检测算法[J]. 计量学报, 2024, 45(9): 1314-1323.  
ZHANG L G, ZHANG Q, JIN M, et al. An improved YOLOv5s forest fire smoke detection algorithm[J]. Acta Metrologica Sinica, 2024, 45(9): 1314-1323.
- [10] 骞璐璐, 裴焕斗, 张宇璇. 多场景烟雾环境下改进的 YOLOv5s 烟雾检测算法[J]. 工业仪表与自动化装置, 2024(2):101-105.  
PIAN L L, PEI H D, ZHANG Y X. Improved YOLOv5s smoke detection algorithm in multi-scenario smoke environment[J]. Industrial Instrumentation & Automation, 2024(2): 101-105.
- [11] 杜辰, 王兴, 董增寿, 等. 改进 YOLOv5s 的地下车库火焰烟雾检测方法[J]. 计算机工程与应用, 2024, 60(11):298-308.  
DU CH, WANG X, DONG Z SH, et al. Improved YOLOv5s flame and smoke detection method for underground garage [J]. Computer Engineering and Applications, 2024, 60(11): 298-308.
- [12] CHINO D Y T, AVALHAIS L P S, RODRIGUES J F, et al. BoWFire: Detection of fire in still images by integrating pixel color and texture analysis[C]. 2015 28th SIBGRAPI Conference on Graphics, Patterns and Images. IEEE, 2015: 95-102.
- [13] GEETHA S, ABHISHEK C S, AKSHAYANAT C S. Machine vision based fire detection techniques: A survey[J]. Fire Technology, 2021, 57(2): 591-623.
- [14] DUNNINGS A J, BRECKON T P. Experimentally defined convolutional neural network architecture variants for non-temporal real-time fire detection[C]. 2018 25th IEEE International Conference on Image Processing(ICIP). IEEE, 2018: 1558-1562.
- [15] ZHANG Q X, LIN G H, ZHANG Y M, et al. Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images [J]. Procedia Engineering, 2018, 211: 441-446.

- [16] KO B C, PARK J O, NAM J Y. Spatiotemporal bag-of-features for early wildfire smoke detection[J]. Image and Vision Computing, 2013, 31(10): 786-795.
- [17] DAI J F, QI H, XIONG Y W, et al. Deformable convolutional networks [C]. IEEE International Conference on Computer Vision, 2017: 764-773.
- [18] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. ArXiv preprint arXiv:2301.10051, 2023.
- [19] WANG C Y, YE H I, MARK L H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision, Springer, Cham, 2025: 1-21.
- [20] CHU X X, LI L, ZHANG B. Make repvgg greater again: A quantization-aware approach [C]. AAAI Conference on Artificial Intelligence, 2024, 38(10): 11624-11632.
- [21] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [22] WANG J Q, CHEN K CH, XU R, et al. Carafe: Content-aware reassembly of features[C]. IEEE/CVF International Conference on Computer Vision, 2019: 3007-3016.
- [23] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]. IEEE International Conference on Computer Vision, 2017: 618-626.
- [24] REN SH Q, HE K M, 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.
- [25] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, Springer International Publishing, 2016: 21-37.
- [26] TIAN ZH, SHEN CH H, CHEN H, et al. FCOS: A simple and strong anchor-free object detector [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(4): 1922-1933.
- [27] WANG J J, ZHANG X M, JING K L, et al. Learning precise feature via self-attention and self-cooperation YOLOX for smoke detection[J]. Expert Systems with Applications, 2023, 228: 120330.
- [28] ZHAO Y AN, LYU W Y, XU SH L, et al. Detsr beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [29] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.

## 作者简介

司盼召,硕士研究生,主要研究方向为深度学习、计算机视觉。

E-mail: spz@stu.xju.edu.cn

何丽(通信作者),博士,副教授,主要研究方向为模式识别与智能控制、移动机器人自主导航。

E-mail: xju\_heli@163.com

王宏伟,博士,副教授,主要研究方向为计算机视觉、智能故障诊断。

E-mail: wanghongwei\_xju@126.com

冉腾,博士,副教授,主要研究方向为计算机视觉、移动机器人定位与导航。

E-mail: ranteng@xju.edu.cn