

基于改进的YOLOv10的焚烧目标检测算法

李林峰, 蒋玲凤, 田 淼, 冷震北

重庆对外经贸学院数学与计算机科学学院, 重庆

收稿日期: 2025年9月7日; 录用日期: 2025年10月8日; 发布日期: 2025年10月15日

摘 要

针对复杂焚烧场景下目标检测易受背景干扰、目标形变及光照突变影响的问题, 本研究提出一种基于YOLOv10的增强检测模型。通过引入Gather-and-Distribute机制重构多尺度特征融合路径, 以可变形卷积(DCNv4)提升对非刚性目标的几何表征能力, 并采用Wise-Outer-MPDIoU损失函数实现基于目标几何属性的边界框优化。在自建焚烧数据集上的实验表明, 所提出方法在保持高推理速度(359 FPS)的同时, 平均精度(mAP)达到85.7%, 较基线模型提升2.1%, 显著增强了对焚烧目标的鲁棒感知与定位性能, 具有良好的学术价值与工程应用潜力。

关键词

YOLOv10, 聚合 - 分配机制, 可变形卷积, Wise-Outer-MPDIoU损失函数

Improved Incineration Target Detection Algorithm of YOLOv10

Linfeng Li, Lingfeng Jiang, Miao Tian, Zhenbei Leng

School of Mathematics and Computer Science, Chongqing College of International Business and Economics, Chongqing

Received: September 7, 2025; accepted: October 8, 2025; published: October 15, 2025

Abstract

Addressing the challenges of target detection in complex incineration scenes, which are prone to background interference, target deformation, and sudden illumination changes, this study proposes an enhanced detection model based on YOLOv10. The model incorporates a Gather-and-Distribute mechanism to reconstruct multi-scale feature fusion paths, employs deformable convolution (DCNv4) to

enhance geometric representation of non-rigid targets, and adopts a Wise-Outer-MPDIoU loss function for geometry-aware bounding box optimization. Experimental results on a self-constructed incineration dataset demonstrate that the proposed method achieves a mean average precision (mAP) of 85.7% while maintaining high inference speed (359 FPS), representing a 2.1% improvement over the baseline model. The approach significantly enhances robust perception and localization of incineration targets, demonstrating considerable academic value and engineering application potential.

Keywords

YOLOv10, Gather-and-Distribute Mechanism, Deformable Convolution, Wise-Outer-MPDIoU Loss Function

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着工业化与城市化进程的加速, 焚烧活动产生的污染物(如颗粒物、SO₂、NO_x)对空气质量、人体健康及安全构成严重威胁, 亟需发展高效精准的监测技术。现有检测方法面临复杂背景、环境多变及目标多样性等挑战, 难以兼顾精度与实时性[1], 同时对系统稳定性提出更高要求[2]。当前目标检测范式主要包括两类: 基于区域提议的双阶段方法和端到端回归的单阶段方法。

学者们针对不同场景提出了改进策略, 如Xin等人[3]通过引入Hybrid Dilated CNN防止小目标信息丢失; Liu等人[4]利用YOLOv5检测遥感图像中的秸秆焚烧烟雾, 但限于分辨率, 难以识别稀疏目标; Mukhiddinov等人[5]结合无人机与YOLOv5实现野火实时监测, 通过锚框聚类与多尺度融合提升准确率。

然而, 现有研究多集中于大尺度火灾或工业烟雾, 针对秸秆焚烧场景的研究仍显不足。本文以YOLOv10为基线, 针对秸秆焚烧场景提出三项改进:

- 1) 在C2f模块中使用DCNv4替换CBS卷积, 提升形变目标的定位能力;
- 2) 加入Gather-and-Distribute机制, 加强多尺度信息融合, 抑制复杂背景、弱光及浑浊干扰, 提高检测精度;
- 3) 采用Wise-Outer-MPDIoU损失函数, 结合缩放因子与点距离惩罚, 增强模型在目标尺度多变和背景复杂场景中的适应性。

2. 相关原理

YOLOv10 算法

YOLOv10沿袭YOLO单阶段设计, 整体由输入、主干、颈部与检测头四大组件构成, 并提供N/S/M/L/X五种缩放规格, 以适应不同算力场景。实验选用YOLOv10架构(见图1)。输入端集成最新数据增强策略, 依据目标尺度与位置在线修正锚框, 并自适应缩放输入分辨率; 主干网络引入Conv、SCDown与PSA三大核心模块, 强化特征提取能力。各模型的基础卷积单元由CBS (Conv-BN-SiLU)模块构建, 其依次由卷积层、批归一化层及SiLU激活函数串联而成, 可在保持梯度稳定性的同时实现深层图像特征的高效抽取。输入图像经Backbone完成多层次特征提取后, 传递至Neck区域。本文提出一种新颖的聚合-分配机制

(Gather-and-Distribute, GD), 通过统一模块对来自所有层级的特征进行全局汇聚与自适应再分配。该机制在摒弃传统FPN固有的信息衰减问题的同时, 以极低的计算延迟显著增强了颈部网络的局部融合能力。经GD精炼后的多尺度特征图被送入检测头(Head)。Head采用大、中、小三级解耦式预测头, 各级预测头独立负责对应尺度目标框的回归与分类任务。每个预测头由若干卷积层堆叠而成, 用于提取不同分辨率下的目标判别特征; 随后将特征解码为类别概率与边界框偏移量, 并借助非极大值抑制(NMS)剔除冗余框, 最终输出高精度、低重叠的检测结果。

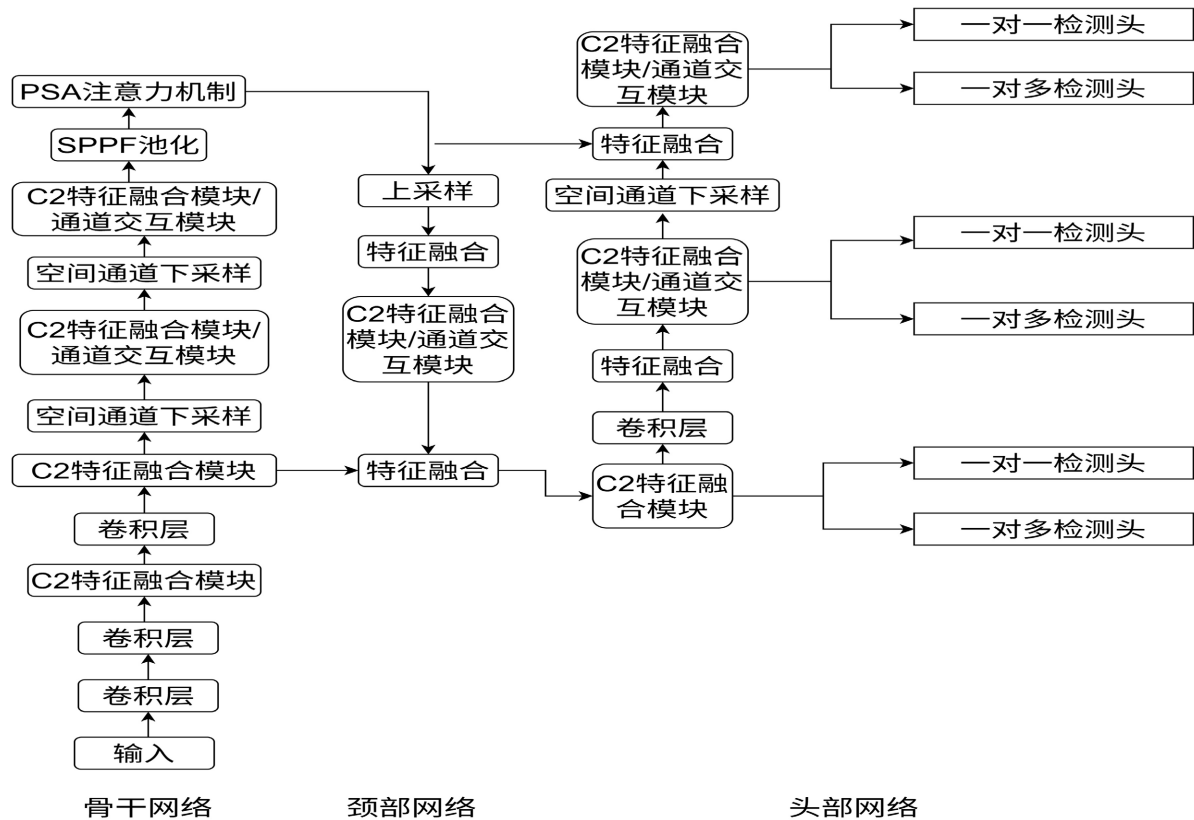


Figure 1. YOLOv10 algorithm architecture diagram

图 1. YOLOv10 的算法结构图

3. YOLOv10 算法的改进

尽管YOLOv10已在众多通用检测任务中表现优异, 但面对焚烧场景时, 复杂地形、多变天气及人为扰动使其误检、漏检频发; 若单纯堆叠模块提升精度, 又会带来参数量激增、推理耗时增加的新问题。为此, 本文提出轻量化改进方案(见图2): 在骨干C2f中嵌入DCNv4可变形卷积提升形变目标感知, 于Neck引入FAM + IFM双模块精炼多尺度特征, 并用Wise-Outer-MPDIoU替换原损失, 兼顾精度与速度。

3.1. DCNv4 模块

受焚烧场景复杂背景及目标形变影响, 原始YOLOv10的C2f模块难以灵活捕捉几何变化。为此, 本文把C2f中的常规卷积升级为DCNv4可变形卷积(见图3), 使卷积核具备自适应采样能力, 从而显著提升对形变目标的感知精度。

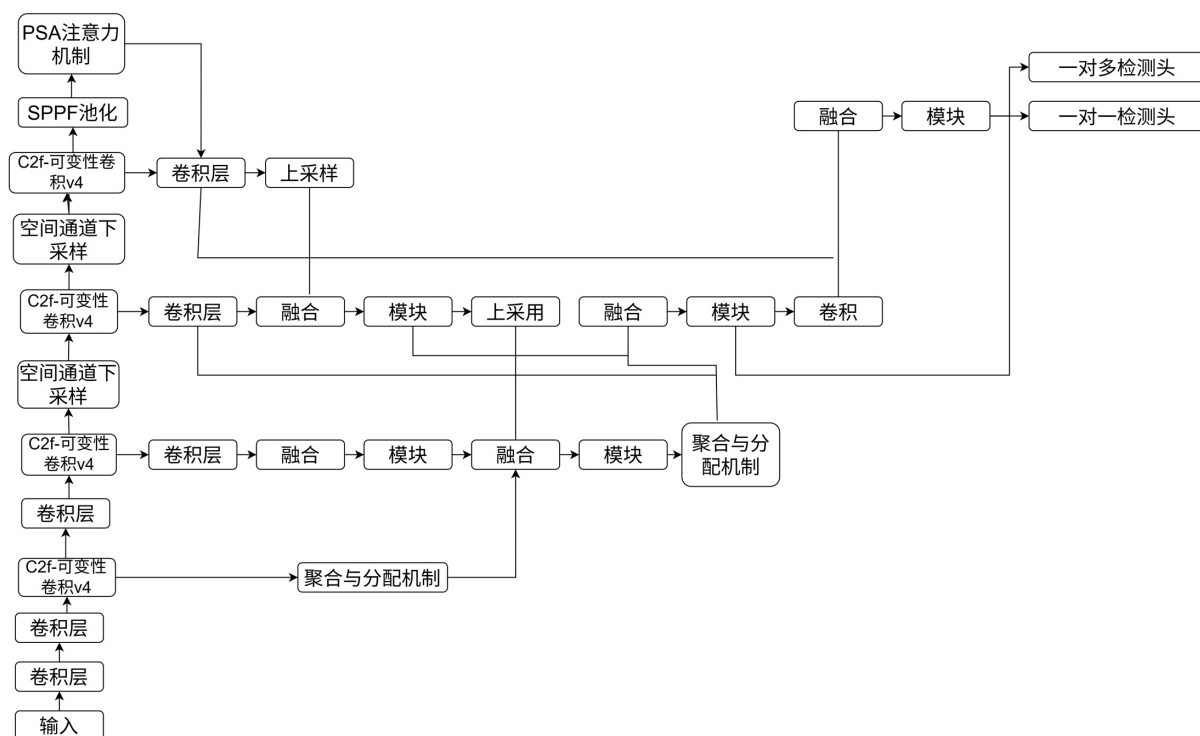


Figure 2. Improved YOLOv10 algorithm architecture diagram

图 2. 改进后的 YOLOv10 算法结构图

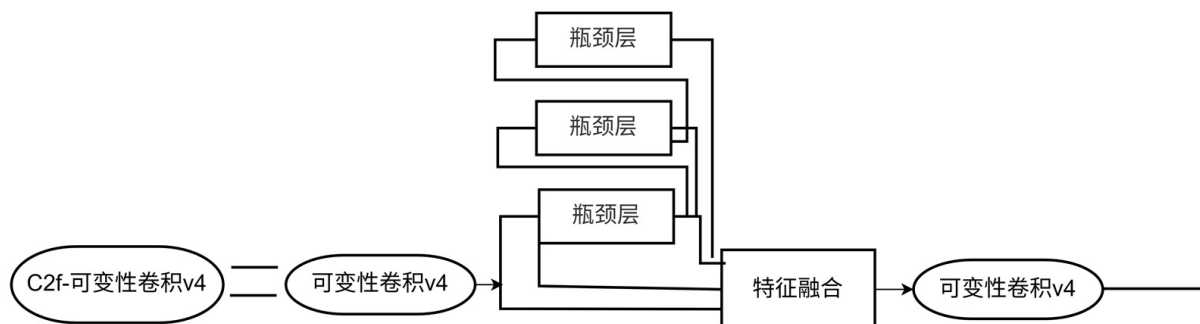


Figure 3. C2f-DCNv4 network architecture

图 3. C2f-DCNv4 网络结构

改进后的C2f单元在维持高效特征提取的同时，为卷积核引入可学习的空间形变参数，使网络对不同形态的焚烧目标具备更强的几何适应能力。2024年，孙己龙等人[6]针对DCNv3在速度与收敛性上的不足，提出DCNv4，其核心是在空间聚合阶段舍弃softmax归一化，转而采用无界动态权重，从而在减少内存访问开销的同时提升动态表达能力。具体而言，对于输入张量 $X \in R^{H \times W \times C}$ ，任意位置 P 的DCNv3输出可重新参数化为：

$$y(P) = \sum_{k=1}^K w_k \cdot X(P + \Delta P_k)$$

其中， K 为采样点数， ΔP_k 为网络预测的偏移量， w_k 为第 k 个采样位置的权重，DCNv4通过移除通道维度的归一化约束，使 w_k 直接由输入特征动态生成，无需额外softmax处理，进而加速收敛并降低冗余

计算。

DCNv3采用受限的动态权重，在统一窗口内完成空间聚合；DCNv4继承其自适应窗口思想，同时引入无界权重，使每个位置拥有独立滑窗，形成更灵活的动态聚合策略。

3.2. GD 模块

GD模块[7]增强了多尺度特征融合能力，并且在所有模型尺度下，都能实现在延迟与精度之间理想的平衡，不仅增强了颈部网络的信息融合能力，还提升了模型在不同物体尺寸下的性能，其结构图如图4所示。

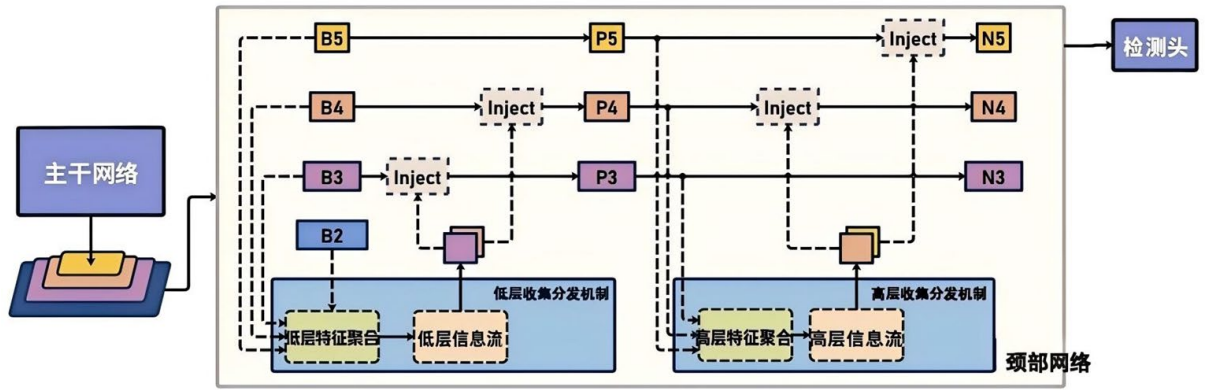


Figure 4. Gold-YOLO network architecture

图 4. Gold-YOLO 网络结构

如图4所示，neck的输入包括backbone提取的特征图B2、B3、B4、B5，其中 $B_i \in \mathbb{R}^{N \times C_{B_i} \times R_{B_i}}$ ， N 是batch size， C 是通道数， $R = H \times W$ 。P3、P4、P5代表来自Backbone并通过自上而下路径聚合的多尺度特征图，用于检测不同尺度的目标(P3分辨率高用于小目标，P5分辨率低用于大目标)。N3、N4、N5代表Neck中进一步聚合或增强后的多尺度特征输出，这些特征会送入检测头(Head)进行预测。在实际部署中，Gather-and-Distribute流程由三个轻量级模块完成：FAM负责跨层特征对齐，IFM执行全局信息融合，Inject则完成精炼特征的层级回注。gather过程包括两步。首先，FAM从不同层收集和对齐特征。然后，IFM通过融合对齐的特征得到全局信息。在获得全局信息后，inject模块将这些信息distribute到每个level中，并使用简单的注意力操作进行注入，从而提高分支检测能力。

为了增加模型检测不同大小对象的能力，GD模块提出了两个分支，low-stage GD和high-stage GD。

在low-stage GD分支中，Low-FAM和Low-IFM分别是低阶段特征对齐模块和低阶段信息融合模块。在Low-FAM中，用average pooling下采样得到一个统一大小的 F_{align} ，选择 $R_{B4} = \frac{1}{4} R$ 目标大小。而Low-IFM包括

多层重参数化卷积Block (RepBlock) 和一个split操作。具体来说，RepBlock取 F_{align} ($channel = \text{sum}(C_{B2}, C_{B3}, C_{B4}, C_{B5})$) 作为输入得到 F_{fuse} ($channel = C_{B4} + C_{B5}$)，然后沿通道维度split成 F_{inj_P3} 和 F_{inj_P4} 。如下：

$$F_{align} = \text{Low_FAM}([B2, B3, B4, B5]), \quad (1)$$

$$F_{fuse} = \text{RepBlock}(F_{align}), \quad (2)$$

$$F_{inj_P3}, F_{inj_P4} = \text{Split}(F_{fuse}). \quad (3)$$

在high-stage GD分支中, High-FAM和High-IFM分别是高阶特征对齐模块和高阶信息融合模块。High-FAM和Low-FAM的操作一样, 通过全局平均池化下采样来对齐大小, 目标大小为 $R_{p5} = \frac{1}{8}R$ 。High-IFM包括多个transformer block和一个split操作。具体包括三步:

- 1) High-FAM的输出 F_{align} 通过transformer block融合得到 F_{fuse} ;
- 2) F_{fuse} 通过 1×1 卷积通道降维得到 $sum(C_{p4}, C_{p5})$;
- 3) 沿通道进行split操作得到 F_{inj_N4} 和 F_{inj_N5} 。

具体如下:

$$F_{align} = \text{High_FAM}([P3, P4, P5]) \quad (4)$$

$$F_{fuse} = \text{Transformer}(F_{align}) \quad (5)$$

$$F_{inj_N4}, F_{inj_N5} = \text{Split}(\text{Conv} \times 1(F_{fuse}))$$

上式中的transformer融合模块包括多个堆叠的transformer block, 每个block包含一个multi-head attention block、ffn、residual connection。

3.3. Wise-Outer-MPDIoU

原版YOLOv10选用CIoU Loss, 该损失专为解决边界框重叠与偏移而设, 可同步优化定位精度与模型评估指标。在焚烧目标检测的实际应用中, 数据集的质量参差不齐, 因而, 本文将Wise-IoUv3 [8]、Outer-IoU [9]两种损失函数改进作为一种新损失函数并使用, 有效提高了模型的检测准确率。

3.3.1. Wise-IoU

Wise-IoU在经典IoU的基础上引入类别权重, 缓解类别失衡; 其v3版利用离群度 β 计算非单调聚焦系数 r , 通过式(6)~(8)动态调整梯度关注区域, 从而提升边界框回归的鲁棒性。

$$r = \frac{\beta}{\delta \cdot \alpha^{\beta-\delta}} \quad (6)$$

$$\beta = \frac{\mathcal{L}_{bU}}{\mathcal{L}_{bU}[0, +\infty)} \quad (7)$$

$$\mathcal{L}_{wbU_{v3}} = r \cdot \mathcal{L}_{wbv1} \quad (8)$$

其中, β 反映框的偏离程度, 其滑动均值 $\overline{\mathcal{L}_{bU}}$ 抑制训练末期的梯度衰减。当 β 触及预设阈值时, 模型即可获得峰值梯度收益。随着 $\overline{\mathcal{L}_{bU}}$ 实时刷新, 质量分级准则同步自适应, 帮助网络在迭代中精准调整边界框位置。

3.3.2. Outer-IoU

本文通过Outer-IoU损失函数辅助边框计算IoU损失。Outer-IoU通过引入可缩放辅助框计算IoU损失: 当辅助框尺寸较小时, 优先优化高IoU样本以加速收敛; 尺寸较大时, 则侧重低IoU样本以加快其回归。缩放因子scale动态调整辅助框大小, 使模型在保持精度的同时快速获得边界框修正结果。

Outer-IoU 的计算如下所示:

$$\text{预测框: } B_p = (x_1^p, y_1^p, x_2^p, y_2^p) \quad (9)$$

$$\text{真实框: } B_g = (x_1^g, y_1^g, x_2^g, y_2^g) \quad (10)$$

$$\begin{cases} x_1^{inter} = \max(x_1^p, x_1^g) \\ y_1^{inter} = \max(y_1^p, y_1^g) \\ x_2^{inter} = \min(x_2^p, x_2^g) \\ y_2^{inter} = \min(y_2^p, y_2^g) \end{cases} \quad (11)$$

$$Area_{inter} = \max(0, x_2^{inter} - x_1^{inter}) \times \max(0, y_2^{inter} - y_1^{inter}) \quad \begin{cases} x_1^{outer} = \min(x_1^p, x_1^g) \\ y_1^{outer} = \min(y_1^p, y_1^g) \\ x_2^{outer} = \max(x_2^p, x_2^g) \\ y_2^{outer} = \max(y_2^p, y_2^g) \end{cases} \quad (12)$$

$$Area_{outer} = (0, x_2^{outer} - x_1^{outer}) \times (0, y_2^{outer} - y_1^{outer})$$

$$Outer-Iou = \frac{Area_{inter}}{Area_{outer}} \quad (12)$$

$$\mathcal{L}_{Outer-Iou} = 1 - Outer-Iou \quad (13)$$

4. 实验对比及结果展示

4.1. 实验详情

4.1.1. 实验环境及参数设置

硬件端采用RTX 4090 (24 GB GDDR6X)与i9-12700KF组合，系统为Windows 10；软件栈选用PyTorch 1.13.0 + CUDA + cuDNN，Python 3.8。训练时以SGD为优化器，Batch_Size为16，初始学习率0.01，权重衰减0.0005，动量0.937，输入尺寸固定640 × 640，共迭代200 epoch。

4.1.2. 模型评价指标

本文评估体系兼顾准确度与实时性，选取mAP、计算量(GFLOPs)、参数量(M)及帧率(FPS)四项指标，其计算公式如下：

$$\begin{aligned} P &= \frac{TP}{TP + FP} \\ R &= \frac{TP}{TP + FN} \\ AP &= \int_0^1 P(R) dR \\ mAP &= \frac{1}{n} \sum_{i=1}^n AP_i \end{aligned} \quad (14)$$

$$FPS = \frac{\text{Number}}{\text{time}} \quad (15)$$

4.2. 对比实验

4.2.1. 注意力机制比较实验

为验证GD注意力模块的有效性，以YOLOv10为基线进行对照实验。表1结果显示：引入GD后，mAP达85.7%，较基线提升2.1%，而SE、CBAM、Biformer的提升均低于1.5%；同时GD仅带来2.52 M额外参数与28.6 G FLOPs，兼顾精度与实时性，更契合焚烧检测场景。

Table 1. Experimental comparison results
表 1. 实验对比结果

Network	平均精度/%	参数量/M	浮点运算量
YOLOv10M	83.6	37.5	29.1 G
YOLOv10M + SE	84.6	37.3	28.7 G
YOLOv10M + CBAM	84.2	37.4	28.2 G
YOLOv10M + Biformer	84.4	37.3	27.9 G
YOLOv10M + GD	85.7	2.52	28.6 G

4.2.2. 消融实验

为验证所提模块的有效性，在YOLOv10基线上展开消融实验(实验参数与表1完全一致)。表2结果显示：融合DCNv4、GD注意力及Wise-Outer-MPDIoU后，参数量与计算量均小幅下降，mAP提升2.1 pct，满足焚烧检测对精度与实时性的双重要求。

Table 2. Ablation study results
表 2. 消融实验结果

Index	Model	DCNv4	GD	BiFPN	Wise-Outer-MPDIoU	帧率/F·s ⁻¹	参数量(M)	平均精度(%)
Exp.1	YOLOv10M					355	2.37	83.6
Exp.2	YOLOv10M	✓				347	2.41	84.6
Exp.3	YOLOv10M		✓			359	2.52	85.7
Exp.4	YOLOv10M			✓		364	2.48	85.1
Exp.5	YOLOv10M				✓	352	2.51	85.3
Exp.6	YOLOv10M	✓	✓			332	2.54	85.4
Exp.7	YOLOv10M			✓	✓	313	2.56	86.0
Exp.8	YOLOv10M	✓	✓	✓		324	2.61	86.6
Exp.9	YOLOv10M	✓	✓		✓	330	2.58	85.8
Exp.10	YOLOv10M	✓	✓	✓	✓	332	2.63	87.3

4.2.3. 不同算法对比实验

实验表明(表3)，改进后的YOLOv10在焚烧检测任务中表现优异：在统一参数与数据集下，其mAP达到87.3%，较YOLOv8-M提升3.7 pct，显著优于Faster-RCNN、SSD等常用模型，验证了本方法的有效性。

Table 3. Algorithm comparison experiment
表 3. 算法对比实验

模型	帧率	参数量(M)	平均精度(%)
Faster-RCNN	52	28.3	56.1
SSD	82	33.4	57.4
YOLOv5l	153	23.2	83.5
YOLOv8M	66.2	11.1	83.6
Ours	332	2.63	87.3

5. 结论

本研究基于YOLOv10M提出面向焚烧场景的目标检测框架：一方面将DCNv4可变形卷积融入主干网络，增强模型对形变及非刚性目标的建模能力；另一方面嵌入GD注意力模块，通过自适应捕获相似特征，强化网络对关键区域相关性的判别。其次，将YOLOv10M颈部原有的PAFPN升级为带权双向特征金字塔BiFPN，利用跨尺度双向信息流动和自适应加权融合机制，在不显著增加计算量的前提下提升多尺度表达能力。最后，在预测端采用Wise-Outer-MPDIoU损失，兼顾高、低质量样本的梯度贡献，使边界框回归在训练初期即可快速收敛并保持稳定，从而增强模型泛化性能，显著抑制焚烧场景中的误检与漏检现象。

基金项目

重庆对外经贸学院校级科研项目(KYZK2024015)、重庆对外经贸学院校级科研项目(KYZK2024003)。

参考文献

- [1] Ye, X., Yu, F., Zhou, C. and Jiang, M. (2020) A Blurry Image Recognition Method for Straw Burning Detection. 2020 *IEEE 5th International Conference on Signal and Image Processing (ICSIP)*, Nanjing, 23-25 October 2020, 323-327. <https://doi.org/10.1109/icsip49896.2020.9339292>
- [2] Wu, X.H., Liu, T., Cheng, Y.Z., Wang, L.G., Guo, Y., Zhang, Y. and He, J. (2017) Dynamic Monitoring of Straw Burned Area Using Multi-Source Satellite Remote Sensing Data. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, **33**, 153-159.
- [3] Xin, F., Zhang, H. and Pan, H. (2023) Hybrid Dilated Multilayer Faster RCNN for Object Detection. *The Visual Computer*, **40**, 393-406. <https://doi.org/10.1007/s00371-023-02789-y>
- [4] Liu, H., Li, J., Du, J., Zhao, B., Hu, Y., Li, D., *et al.* (2022) Identification of Smoke from Straw Burning in Remote Sensing Images with the Improved Yolov5s Algorithm. *Atmosphere*, **13**, Article 925. <https://doi.org/10.3390/atmos13060925>
- [5] Mukhiddinov, M., Abdusalomov, A.B. and Cho, J. (2022) A Wildfire Smoke Detection System Using Unmanned Aerial Vehicle Images Based on the Optimized YOLOv5. *Sensors*, **22**, Article 9384. <https://doi.org/10.3390/s22239384>
- [6] 孙己龙, 刘勇, 路鑫, 等. 基于可变形卷积网络和 YOLOv8 的衬砌裂缝检测模型研究[J]. 中国安全生产科学技术, 2024, 20(8): 181-189.
- [7] Wang, C.C., He, W., *et al.* (2023) Gold-YOLO: Efficient Object Detector via Gather-and-Distribute Mechanism. arXiv: 2309.11331.
- [8] Tong, Z., Chen, Y. and Xu, Z. (2023) Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. arXiv: 2301.10051.
- [9] 韩强. 面向小目标检测的改进 YOLOv8 算法研究[D]: [硕士学位论文]. 长春: 吉林大学, 2023.