基于改进YOLOv10s的矿井火灾烟雾目标检测*

武晔秋[1],石泽来[2],朱慧敏[2],刘玥[1]

(1.山西大同大学 建筑与测绘工程学院, 山西 大同市 037000;
2.山西大同大学 煤炭工程学院, 山西 大同市 037000)

摘要:矿井火灾烟雾目标检测是防止矿井内发生火灾事故的重要预防措施之一。为了提高矿井内火灾烟雾检测的精度与速度,提出了一种基于改进YOLOv10s的矿井火灾烟雾检测方案。首先将收集的矿井火灾烟雾图像数据集通过在线平台完成标注,然后按7∶2∶1划分为训练集、验证集及测试集,在训练前对图片进行小波降噪及Auto-Msrcr算法增强处理;其次使用Focaler-IoU重构原始YOLOv10损失函数,以及添加强化卷积神经网络性能的注意力模块CBAM和使用跨阶段部分网络模块VoV-GSCSP替换颈部网络中的第三个C2f模块,从而提高模型的表征能力和决策性能,最终提高对矿井火灾烟雾识别的特征权重。试验结果表明,改进后的YOLOv10s模型与原模型相比准确率提升了3.5%,平均精度均值mAP@50和mAP@50-95分别提升了0.7%和2.2%,检测帧率提高了11.9帧/s;相较其他主流YOLO模型,改进后的YOLOv10s模型检测精度和速度更高,具有更好的综合检测性能,适用于高精度和高实时性的矿井火灾烟雾检测场景。

关键词:矿井火灾;烟雾检测;改进YOLOv10s;注意力模块;特征权重

中图分类号:TD75 **文献标识码**:A **文章编号**:1005-2763(2025)09-0200-08

Mine Fire Smoke Target Detection Based on Improved YOLOv10s

WU Yeqiu[1], *SHI Zelai*[2], *ZHU Huimin*[2], *LIU Yue*[1]

(1.College of Architecture and Geomatics Engineering, Shanxi Datong University, Datong, Shanxi 037000, China;
2.College of Coal Engineering, Shanxi Datong University, Datong, Shanxi 037000, China)

Abstract: Mine fire smoke target detection is one of the important preventive measures to prevent fire accidents in mines. In order to improve the accuracy and speed of fire smoke target detection in mines, a mine fire smoke target detection scheme based on improved YOLOv10s was proposed. Firstly, the collected mine fire smoke image dataset was divided into a training set, validation set and test set according to 7∶2∶1 after annotation on the online annotation platform. Before training, the images were subjected to wavelet denoising and Auto-Msrcr algorithm enhancement processing, then the original YOLOv10 loss function was reconstructed using Focaler-IoU. The attention module CBAM that strengthened the performance of convolutional neural networks was added, and the third C2f module in Neck was replaced by the cross-stage partial network module VoV-GSCSP, thereby improving the representation ability and decision-making performance of the model. Finally, the feature weights of mine fire smoke recognition were improved. The experimental results show that the improved YOLOv10s model has a 3.5% higher accuracy than the original model, the average precision mAP@50 and mAP@50-95 are increased by 0.7% and 2.2% respectively, and the detection frame rate is increased by 11.9 frames/s. Compared with other mainstream YOLO models, the improved YOLOv10s model has higher detection accuracy and speed, as well as better comprehensive detection performance, which is suitable for high-precision and high real-time mine fire smoke detection scenarios.

Key words: Mine fire, Smoke detection, Improved YOLOv10s, Attention module, Feature weight

---

* **收稿日期**:2024-10-10

基金项目:国家自然科学基金项目(51764044);山西省2024年度研究生教育创新计划实践创新项目(2024SJ379);山西大同大学2024年研究生实践创新项目(2024SJCX06)

作者简介:武晔秋(1986—),女,山西大同人,博士,教授,主要研究方向为矿井通风系统。E-mail:wuyeqiu@126.com

通信作者:石泽来(1996—),男,河南周口人,硕士研究生,主要研究方向为矿山安全工程。E-mail:s1931224921@outlook.com

0 引言

矿井火灾是矿井常见的五大灾害之一,若井下发生火灾事故,则会严重影响矿井的正常开采并威胁矿工的生命安全。如 2021 年山东省烟台招远市曹家洼金矿"2·17"较大火灾事故,2023 年河南三门峡河南大有能源股份有限公司耿村煤矿"5·9"较大火灾事故等均造成了巨大的经济损失和人员伤亡。在火源出现时如果能及时、精准地监测和定位,可有效采取措施,从而预防火灾的发生。

国内外的研究者们对矿井火灾烟雾检测的相关研究主要集中在传感器检测、红外图像检测、图像增强目标检测等方面。REIMANN P 等[1]研究开发了一个基于单个半导体气体传感器和相对湿度与温度传感器的系统,利用高动态范围的硬件控制和数据采集平台,实现了半导体气体传感器在矿井火灾中的应用。KUMAR S 等[2]采用电位法(SP)对地下煤火进行了研究,通过利用卫星热红外数据进行煤火成图,从而识别地表煤火位置和煤火的横向传播。王伟峰等[3]通过改进 YOLOv5 算法并结合 K-means 算法进行图像增强,提高了矿井火灾识别准确率。白宝军[4]通过分析烟雾特征和融合判定,采用改进 MSRCP 算法进行矿井图像增强,最后结合 SVM 矿井烟雾检测模型设计出了适合矿井的火灾烟雾检测算法。吕宗宝等[5]针对矿井火灾检测系统提出了多元数据融合算法,利用数字滤波技术,实现了对火灾灾情和位置的探测。孙继平等[6]提出了基于近红外和可见光图像的矿井外因火灾感知方法,融合近红外图像和可见光图像,提高了矿井外因火灾的识别率。王树斌等[7]提出了一种基于 Transformer 的矿井火灾时间序列预测方法:先利用 Hampel 滤波和拉格朗日插值方法检测异常值,然后基于 Transformer 的自注意力机制从时间序列数据中提取特征并预测火灾趋势。刘春霞等[8]提出了一种改进的 YOLOv5s 矿井烟火探测算法,利用 GhostNet 对 YOLOv5s 的颈部进行重构,并在头部位置加入 SA 注意机制使其能够满足矿井烟火检测的要求。

基于深度学习的目标检测算法具有强大的特征表达能力和泛化性,在复杂场景下表现较好的性能,故在矿井火灾烟雾检测这种实时性需求较高的场景中更有优势。本文选用单阶段(One-Stage)目标检测算法 YOLOv10s 作为检测的基础模型。为了使模型能精准检测和定位井下火灾烟雾,首先对数据集图像进行小波降噪和 Auto-Msrcr 算法增强,然后对轻量级目标检测模型 YOLOv10s 进行改进:通过使用 Focaler-IoU 替换原损失函数(Loss),以及添加强化卷积神经网络性能的注意力模块 CBAM 和使用跨阶段部分网络模块 VoV-GSCSP 替换 Neck 中的第三个 C2f 模块,从而提高模型的性能,使改进模型可以满足矿井火灾烟雾的检测。

1 YOLOv10s 算法原理

1.1 YOLOv10 概述

YOLOv10[9]是由清华大学研究团队开发,于 2024 年 5 月 29 日发布的全新 SOTA 模型。由于过往的 YOLO 版本对 NMS(Non-Maximum Suppression)技术过度的依赖和架构方面的低效率影响了最佳性能的实现,YOLOv10 通过引入双标签分配和以效率—准确性为导向的整体模型设计策略,其在训练阶段提供更多的正样本来丰富模型的训练,在推理阶段则通过梯度截断方式来切换一对一的检测头,使模型无需 NMS 后处理,同时在不损失性能的情况下减少了推理开销。本文选用 YOLOv10s 作为基础模型进行改进是因为其兼顾了速度和精度,对矿井火灾烟雾的检测具有显著的效果。

1.2 YOLOv10s 网络结构

YOLOv10s 网络主要由 4 部分组成:输入端、网络骨干、颈部网络和头部(输出端)。网络结构如图 1 所示。首先输入端输入 640 px×640 px 分辨率大小的图像,经过骨干网络卷积神经网络,逐步减小输入特征图的大小,并通过一系列的卷积和残差连接来学习提取图像的高级特征。颈部网络负责接收骨干网络卷积神经网络提取的特征,并继续通过一系列卷积和融合操作来学习图像特征。最终通过头部网络(输出端)进行目标检测。

在 YOLOv10s 的骨干网络部分 C2f 模块引入了增强版的 CSPNet[10](Cross Stage Partial Network)结构,该结构提高了 CNN 的学习能力,能在保持轻量化的同时保证准确性,以改善梯度流并减少计算冗余。SCDown 模块将空间和通道解耦,先通过 1×1 的逐点卷积调节通道数,再通过 3×3 的深度卷积做空间下采样,在降低计算成本的同时最大限度保留信息。C2fCIB 则是用 CIB 替换了原

本 C2f 中的 Bottleneck,CIB 则是将 Bottleneck 中的标准卷积用深度卷积加逐点卷积进行替换。SPPF 是空间金字塔池化,其作用是在不同尺度下对图像进行特征提取和编码,它可以将任意大小的输入图像重新缩放到固定大小,并生成固定长度的特征向量。PSA 是一种高效的局部自注意力模块,它将卷积后的特征一部分输入到由多头自注意力模块 MHSA 和前馈网络 FFN 组成的模块中,然后将两部分连接起来并通过卷积进行融合,增强全局建模能力,降低计算复杂度,其网络结构如图 2 所示。

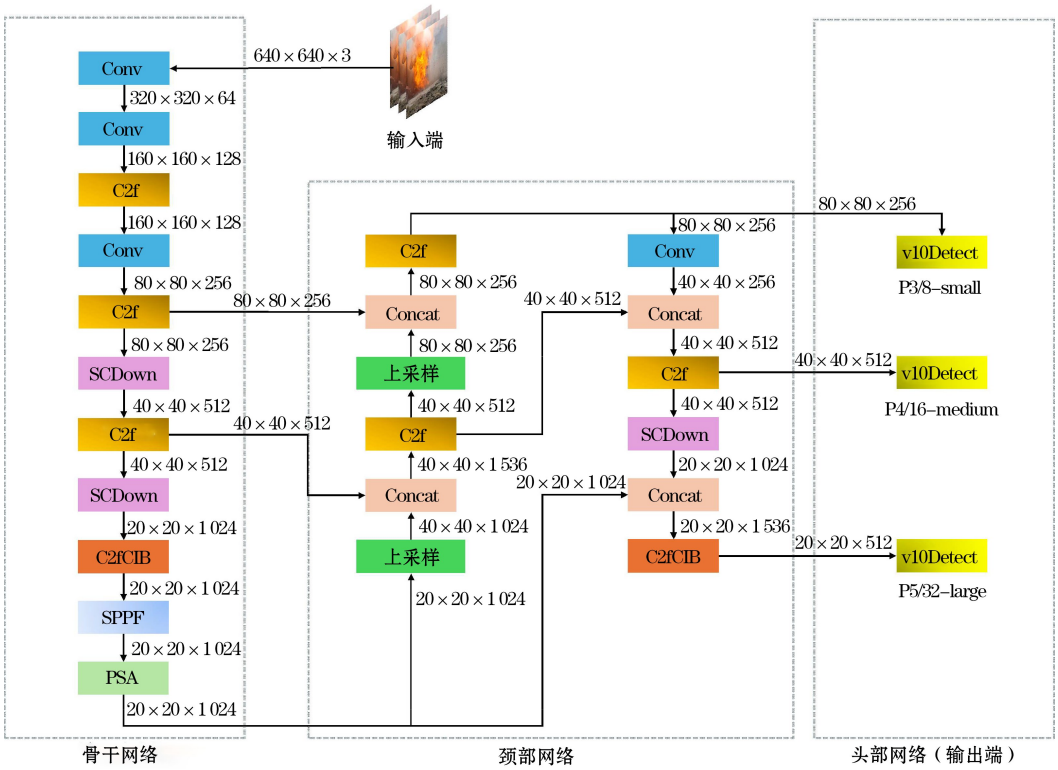


图 1 YOLOv10s 网络结构

Fig.1 YOLOv10s network structure

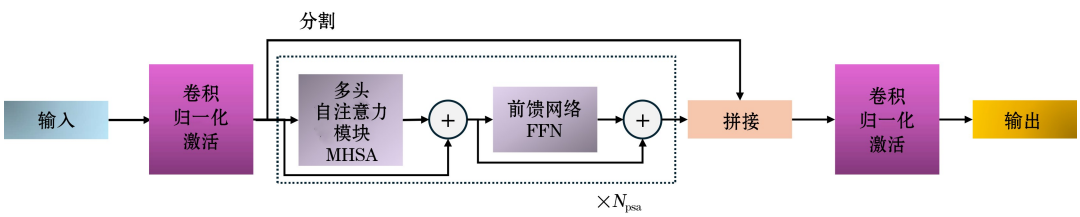


图 2 PSA 网络结构

Fig.2 PSA network structure

2 YOLOv10s 模型改进

2.1 损失函数改进

YOLOv10 模型采用 $CIoU$ 作为损失函数来评估检测中的边界框,其与 IoU 相比不仅考虑了重叠区域,还引入了中心点距离和长宽比,缓解了梯度消失问题,提升了检测的精度,公式如下:

$$IoU=\frac{|A\cap B|}{|A\cup B|} \tag{1}$$

$$CIoU=1-IoU+\frac{p^2(b,b_{gt})}{c^2}+\alpha v \tag{2}$$

式中:IoU 为交并比;$|A\cap B|$ 为边框交集面积;$|A\cup B|$ 为边框并集面积;$p^2(b,b_{gt})$ 为预测框 b 与真实框 b_{gt} 的中心点之间距离的平方;c 为最小外接框的对角线长度;α 是一个权重参数,其公式为:

$$\alpha=\frac{v}{1-IoU+v} \tag{3}$$

其中,v 可衡量宽高比一致性,其公式为:

$$v=\frac{4}{\pi^2}\left(\arctan\frac{w_{gt}}{h_{gt}}-\arctan\frac{w}{h}\right)^2 \tag{4}$$

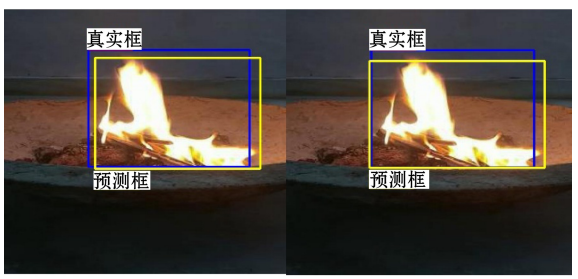
式中：w 和 h 是预测框的宽和高；w_{gt} 和 h_{gt} 是真实框的宽和高。由于 $CIoU$ 需要计算框的中心距离和长宽比一致性，其计算复杂度增加，从而收敛缓慢。当预测框和真实框的重叠面积非常小时，无法有效处理低重叠或不重叠的框。

在实际应用中，对矿井火灾烟雾的精准定位要求极高，而目标检测定位精度很大程度上取决于边框回归损失函数，所以本文引用 Focaler-IoU[11] 方法，通过线性区间映射重构原始 IoU 损失函数，使边框回归效果提升。Focaler-IoU 与本文所使用重构后的公式如下：

$$IoU^{\text{Focaler}}=\begin{cases}0, IoU<d\\ \dfrac{IoU-d}{u-d}, d\ll IoU\ll u\\ 1, IoU>u\end{cases} \tag{5}$$

$$L_{\text{Focaler-}CIoU}=L_{CIoU}+IoU-IoU^{\text{Focaler}} \tag{6}$$

式中：IoU 为原始 IoU 值，$[d,u]\in[0,1]$，当 IoU 处于 d 和 u 之间时，损失函数是一个根据 IoU 值线性递增的函数。损失函数改进前后生成的真实框和预测框对比如图 3 所示。



(a) Loss 改进前　　(b) Loss 改进后

图 3　真实框及预测框对比

Fig.3　Comparison of ground truth boxes and predicted boxes

2.2　融合注意力机制

为了增强对矿井火灾烟雾特征的提取，本文融合“即插即用”的卷积块注意模块[12]（Convolutional Block Attention Module，CBAM），该模块是一种增强卷积神经网络性能的双重注意力机制，通过将通道注意力和空间注意力相结合，达到对火灾烟雾特征的双重精炼。CBAM 的流程为：输入特征先经过通道注意力（Channel Attention）模块进行加权处理，然后再经过空间注意力（Spatial Attention）模块进行最终加权处理，得到卷积输出的结果。

通道注意力模块结构如图 4 所示。其流程为：对输入特征图 F（维度为 $c\times h\times w$，c 为通道数，h 为高度，w 为宽度）进行全局平均池化（AvgPool）和最大池化操作（MaxPool），随后生成两个不同的空间上下文描述符 F^c_{avg}（平均池化后的特征）和 F^c_{max}（最大池化后的特征），然后这两个描述符被输入到一个共享的且包含隐藏层（多层感知器 MLP（Multi-Layer Perceptron））的网络中，其中隐藏层的激活尺寸设为 $R^{c/r\times1\times1}$，最终融合两个 MLP 的输出特征向量并相加，再通过激活函数 σ（Sigmoid）映射后生成通道注意力图。通道注意力模块公式如下：

$$M_cF=\sigma(\text{MLP}(\text{AP}(F))+\text{MLP}(\text{MP}(F)))=\sigma(W_1(W_0(F^c_{\text{avg}}))+W_1(W_0(F^c_{\text{max}}))) \tag{7}$$

式中：$M_c\in R^{c\times1\times1}$ 为权重系数；σ 为 sigmoid 函数；AP 为 AvgPool；MP 为 MaxPool；$W_0\in R^{c/r\times c}$，$W_1\in R^{c\times c/r}$，并且 MLP 权重 W_0 和 W_1 对于两个输入是共享的。

空间注意力模块结构如图 5 所示。其流程为：对输入的特征图 F 进行最大池化和平均池化操作，将其通道信息进行聚合，从而生成平均池化特征 $F^s_{\text{avg}}\in R^{1\times h\times w}$ 和最大池化特征 $F^s_{\text{max}}\in R^{1\times h\times w}$ 这两个 2D 特征图，然后将生成的特征图拼接并进行 7×7 卷积运算，再通过激活函数 σ（Sigmoid）映射后生成空间注意力图。空间注意力模块公式如下：

$$M_sF=\sigma(f^{7\times7}([\text{AP}(F);\text{MP}(F)]))=\sigma(f^{7\times7}([F^s_{\text{avg}};F^s_{\text{max}}])) \tag{8}$$

式中：$M_s\in R^{1\times h\times w}$ 为权重系数；$f^{7\times7}$ 表示卷积核为 7×7 的运算。

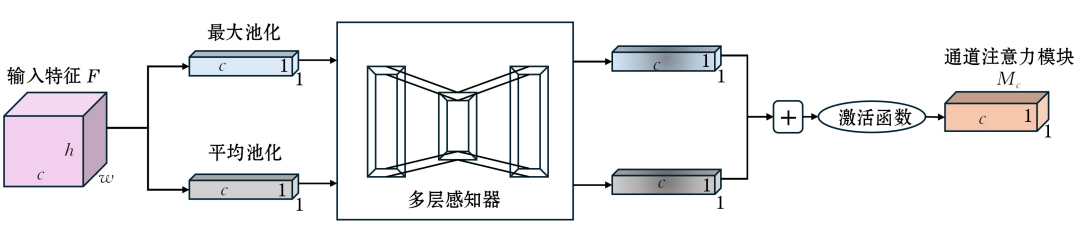


图 4　通道注意力模块结构

Fig.4　Structure of the channel attention module

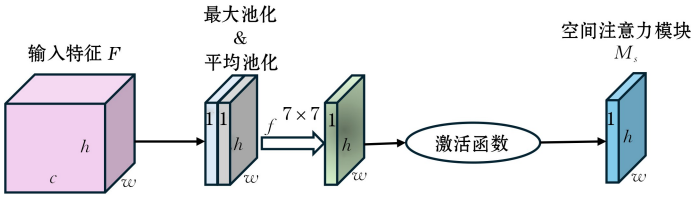


图 5 空间注意力模块结构

Fig.5 Structure of the spatial attention module

2.3 替换神经网络模块

YOLOv10s 的颈部网络结构中第三个 C2f 模块负责处理中等尺寸的特征图,该层对模型检测不同大小的火灾烟雾场景至关重要。本次使用 Slim-neck[13] 轻量级框架中的跨阶段部分网络 VoV-GSCSP 模块来替换原 C2f 模块。该模块结合了 VoVNet[14] 的思想和 CSPNet 的特点,同时加入了全局上下文感知(Global Spatial Context, GSC)的机制。通过 VoVNet 的多层次特征聚合方式,VoV-GSCSP 模块能够有效融合不同层次的特征;结合 CSPNet 能减少参数量和计算量;GSC 机制可以增强模型对全局空间信息的感知。该模块非常适合在火灾烟雾检测这种复杂场景中使用,能够显著提升检测精度和效率。VoV-GSCSP 模块结构如图 6 所示。其流程为:将输入的特征图通过残差连接(SC)进行初步特征提取,接着进入 GS Bottleneck 进行深层次特征压缩,然后将输出特征与初始特征拼接(Concat)后,再经过两个残差连接处理,最终生成输出结果。整体改进后的 YOLOv10s 模型网络结构如图 7 所示。

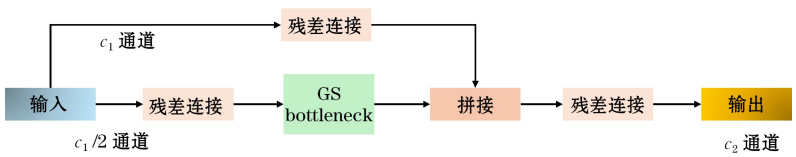


图 6 VoV-GSCSP 模块结构

Fig.6 Structure of the VoV-GSCSP module

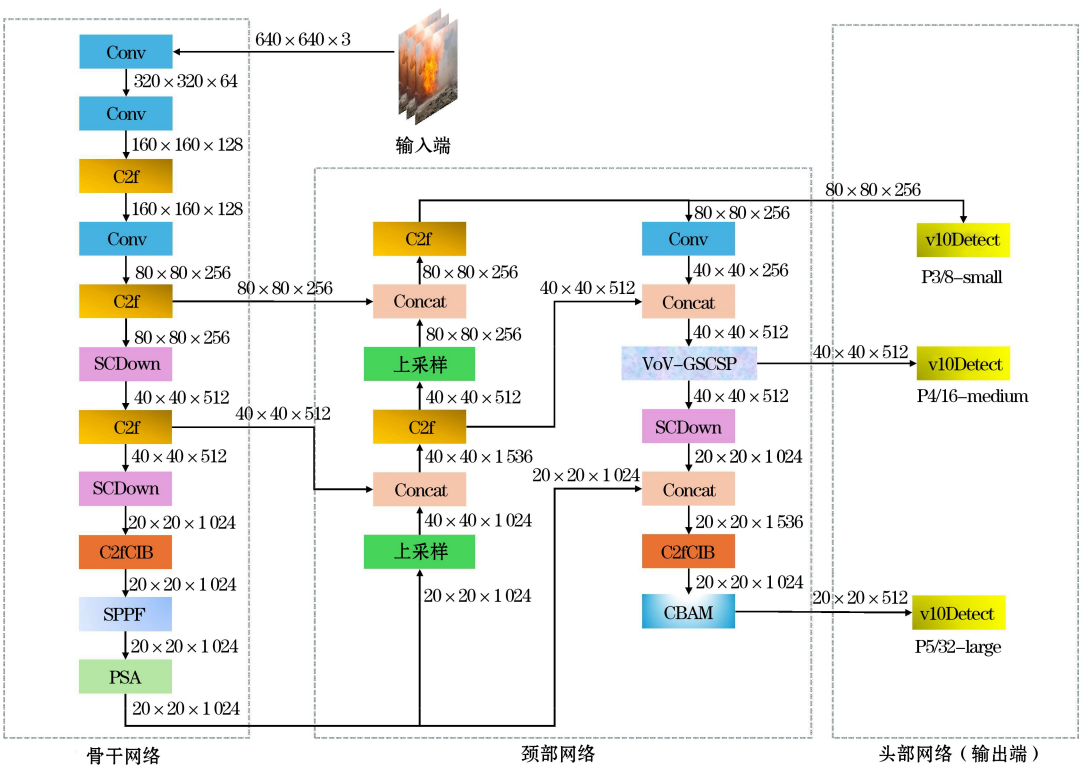


图 7 改进后的 YOLOv10s 模型网络结构

Fig.7 Improved network structure of YOLOv10s model

3　试验结果及分析

3.1　数据集来源

数据集的质量和规模直接影响研究和应用效果，因此，保证数据集的准确性和完整性是非常重要的。由于目前矿井火灾烟雾图像资料较少，且在井下试验具有很大的危险性，本文的图片数据来源于网络上与矿井下火灾烟雾场景相似的图片以及各种场景下的火灾烟雾图片，共计 4 011 张。为了减少图像中的干扰信息，对图像进行小波降噪[15]和 Auto-Msrcr 算法[16]图像增强处理。小波降噪技术通过对图像进行卷积操作，能够有效去除图像中的噪声，从而提升图像的质量；Auto-Msrcr 算法能够有效地增强图像中的细节，改善在低照度条件下拍摄的图像清晰度。图片降噪和增强后的效果如图 8 所示。



（a）图片降噪和增强前



（b）图片降噪和增强后

图 8　部分图片降噪和增强前后对比

Fig.8　Some images before and after reducing noise and enhancement

3.2　试验环境及步骤

本次试验采用的操作系统为 ubuntu22.04，处理器为 12 vCPU Intel（R） Xeon（R） Platinum 8352V CPU @ 2.10 GHz，内存为 90 G，显卡为 NVIDIA GeForce RTX 4090 Gpu，显存为 24 G，编程语言为 Python3.12，深度学习框架版本为 PyTorch2.3.0+Cuda12.1。具体操作步骤如下：

（1）数据集在 Roboflow 在线平台上完成标注及生成分辨率好 640 px×640 px 的图片，将生成的图片按 7∶2∶1 的比例分为训练集，验证集以及测试集；

（2）通过将训练集的图片进行小波降噪及 Auto-Msrcr 算法增强后使用改进的 YOLOv10s 模型进行训练，设置训练轮次为 300，批处理尺寸为 32，初始学习率为 0.001，动量因子为 0.9；

（3）与各个模型对比结果。

3.3　结果分析

3.3.1　评估指标

试验的结果采用深度学习传统的评价指标来评估，即 P（Precision，准确率）、R（Recall，召回率）、AP（Average Precision，平均精度）、mAP（mean Average Precision，平均精度均值），计算公式如下：

$$P=\frac{TP}{TP+FP} \tag{9}$$

$$R=\frac{TP}{TP+FN} \tag{10}$$

$$AP=\int_0^1 P(r)dr \tag{11}$$

$$mAP=\frac{1}{n}\sum_{i=1}^{n} A\,P_i \tag{12}$$

式中：TP（True Positives）为真正例，即正确检验框数；FP（False Positives）为假正例，即错误检验框数；FN（False Negatives）为假负例，即漏检框数；$P(r)$为召回率 r 对应的精确率；AP_i 为第 i 个类别的平均精度；n 为总类别数（本次试验只检测火灾与烟雾，故取 $n=2$）。

3.3.2　损失函数结果分析

在目标检测任务中，损失函数一般由分类损失（cls_loss）、边界框损失（box_loss）和置信度损失（dfl_loss）组成，评价标准为越低越好。YOLOv10s 采用双标签分配，在对齐时采用了 oo（one2one）和 om（one2many）两种策略方式，在训练时两种损失函数策略同时计算，而推理时仅使用 oo 分支。本次试验在改进损失函数方面只针对边界框损失进行提升，所以分类损失和置信度损失结果曲线几乎完全重合，如图 9 和图 10 所示。而边界框损失有明显降低，如图 11 所示，较低的边界框损失可以使模型在矿井火灾烟雾检测任务中的定位精度较高，可以更准确地预测目标的位置和大小。

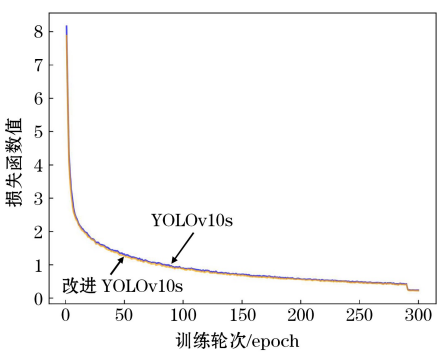


图 9　分类损失函数曲线对比

Fig.9　Comparison of classification loss function curves

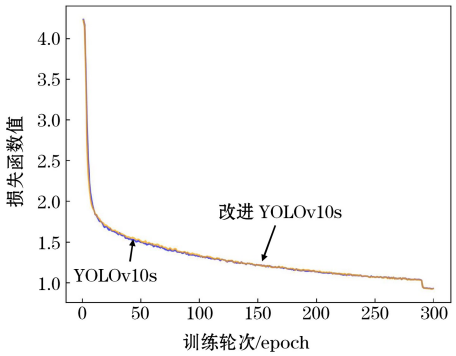


图 10 置信度损失函数曲线对比

Fig.10 Comparison of confidence loss function curves

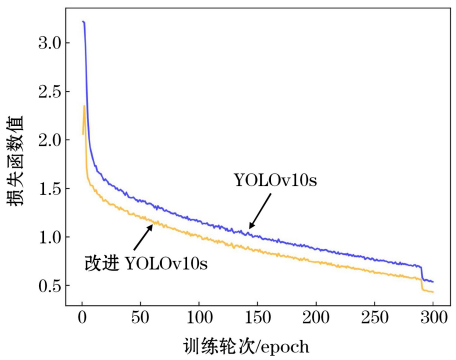


图 11 边界框损失函数曲线对比

Fig.11 Comparison of boundary frame loss function curves

3.3.3 模型对比结果分析

为了验证改进模型的有效性,与原模型及先前版本模型进行性能指标消融对比,试验结果见表 1。

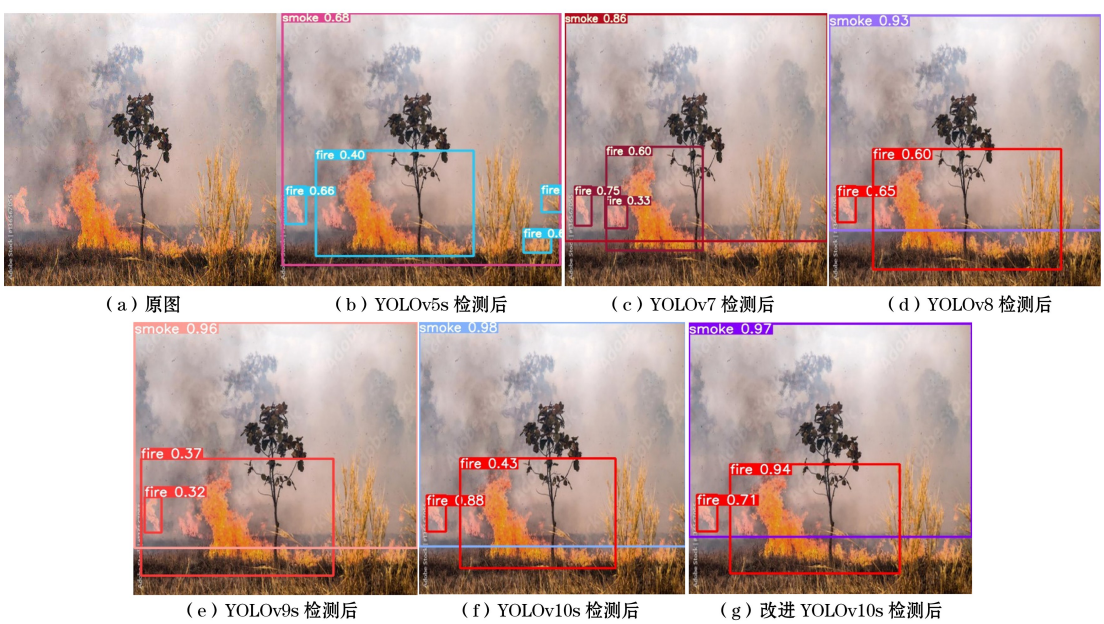
表 1 中 $mAP@50$ 是在 $IoU=0.50$ 的阈值下计算的所有类别的平均精度,该指标的衡量标准为:只要预测框与真实框的重叠超过 50%就会算作正确的检测,容错率较高,可以更容易得到较高的精度值。$mAP@50\text{-}95$ 为 $IoU=0.50\sim0.95$ 之间的阈值下计算得出的平均精度的平均值,该指标考虑了多种 IoU 阈值,能够更严格地全面反映模型的检测精度和鲁棒性。FPS(每秒帧数)为检测速度,该指标反映模型每秒处理图像的帧数,用来评估模型的实时性能。从表 1 的数据可以看出,本文改进模型召回率 R 对比 YOLOv5s、YOLOv7、YOLOv10s 有所下降,平均精度均值 $mAP@50$ 对比 YOLOv7 有所下降,但其他数据均有提升,尤其是 $mAP@50\text{-}95$ 这个重要指标,较 YOLOv5s、YOLOv7、YOLOv8、YOLOv9s、YOLOv10s 分别提升 10.5%、7.5%、9.5%、7.3%、2.2%,对于检测矿井火灾烟雾这种需要高精度和高实时性的场景来说,改进后的 YOLO10s 模型可以满足。使用改进后的 YOLOv10s 模型对火灾烟雾场景进行检测和其他各模型的对比效果如图 12 所示。从图 12 中可以看出改进后的模型综合检测精度最好。

表 1 各模型性能数据

Table 1 Performance data of each model

模型	P/%	R/%	$mAP@50$/%	$mAP@50\text{-}95$/%	FPS/(帧·s^{-1})
YOLOv5s	89.3	86.9	89.5	62.6	285.7
YOLOv7	89.9	**87.3**	**90.8**	64.4	238.1
YOLOv8	89.6	82.3	88.1	63.2	243.9
YOLOv9s	89.2	82.7	88.8	64.5	322.5
YOLOv10s	89.4	86.0	89.5	67.7	324.6
本文方法	**92.5**	84.7	90.1	**69.2**	**336.5**

注:加粗的数据为最佳数据。



(a) 原图 (b) YOLOv5s 检测后 (c) YOLOv7 检测后 (d) YOLOv8 检测后

(e) YOLOv9s 检测后 (f) YOLOv10s 检测后 (g) 改进 YOLOv10s 检测后

图 12 各模型检测效果图对比

Fig.12 Detection effects of different models

4 结论

为了在矿井火灾发生的初期高效检测和精准定位火源，提出了使用改进的 YOLOv10s 模型作为检测工具，旨在预防矿井火灾发生。

(1) 首先将收集的火灾烟雾图片在 Roboflow 在线平台上完成标注并按 7∶2∶1 划分成训练集、验证集以及测试集，在训练前对图片进行小波降噪及 Auto-Msrcr 算法增强处理；然后提出了使用 Focaler-IoU 方法重构 YOLOv10 原始损失函数，并融合强化卷积神经网络性能的注意力模块 CBAM，以及使用跨阶段部分网络模块 VoV-GSCSP 替换颈部网络中的第三个 C2f 模块，通过这 3 种改进方法提高了 YOLOv10s 检测性能和精度。

(2) 改进后的 YOLOv10s 模型与原模型相比，准确率提升了 3.5%，召回率下降了 1.5%、平均精度值 mAP@50 提升了 0.7%，平均精度均值 mAP@50-95 提升了 2.2%、FPS 达到了 336.5 帧/s，提高了 11.9 帧/s。

(3) 召回率微降 1.5%，但精度显著提升，虽然对整体检测效果影响很小，但可以满足矿井火灾烟雾检测的实时性需求，未来可针对此问题加强试验研究，以期优化小目标漏检。

参考文献(References)：

[1] REIMANN P, SCHUTZE A. Fire detection in coal mines based on semiconductor gas sensors[J]. Sensor Review, 2012,32(1):47-58.

[2] KUMAR S, PAL K S. Underground coal fire mapping using analysis of self-potential (SP) data collected from Akashkinaree Colliery, Jharia Coalfield, India[J].Journal of the Geological Society of India,2020,95(4):350-358.

[3] 王伟峰，张宝宝，王志强，等.基于 YOLOv5 的矿井火灾视频图像智能识别方法[J].工矿自动化，2021，47(9)：53-57.
WANG Weifeng, ZHANG Baobao, WANG Zhiqiang, et al. Intelligent identification method of mine fire video images based on YOLOv5[J]. Journal of Mine Automation, 2021, 47(9):53-57

[4] 白宝军.基于图像增强与多特征的煤矿井下火灾烟雾检测方法研究[D].西安：西安建筑科技大学，2021，20-46.
BAI Baojun. Research on smoke detection method of coal mine fire based on image enhancement and multiple features [D]. Xi'an: Xi'an University of Architecture and Technology,2021,20-46.

[5] 吕宗宝，董军，汝洪芳.改进的多元数据融合算法在矿井火灾检测系统中的应用[J].工业仪表与自动化装置，2010(4)：94-97.
LYU Zongbao, DONG Jun, RU Hongfang. Application of mine fire detection system based on improved multi-data fusion algorithm [J]. Industrial Instrumentation & Automation,2010(4):94-97.

[6] 孙继平，崔佳伟.矿井外因火灾感知方法[J].工矿自动化，2021，47(4)：1-5+38.
SUN Jiping, CUI Jiawei. Mine external fire sensing method [J]. Journal of Mine Automation,2021,47(4):1-5+38.

[7] 王树斌，王旭，闫世平，等.基于 Transformer 的矿井内因火灾时间序列预测方法[J].工矿自动化，2024，50(3)：65-70+91.
WANG Shubin, WANG Xu, YAN Shiping, et al. Transformer based time series prediction method for mine internal caused fire[J]. Journal of Mine Automation, 2024, 50(3):65-70+91.

[8] 刘春霞，李超，潘理虎，等.改进 YOLOv5s 的煤矿烟火检测算法[J].计算机工程与应用，2023，59(17)：286-294.
LIU Chunxia, LI Chao, PAN Lihu, et al. Improved coal mine smoke and fire detection algorithm of YOLOv5s[J]. Journal of Computer Engineering & Applications, 2019, 59 (17):286-294.

[9] WANG A, CHEN H, LIU L H, et al. Yolov10: Real-time end-to-end object detection[J]. arXiv preprint arXiv: 2405. 14458, 2024.

[10] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops [C]. Seattle, USA: IEEE, 2020: 390-391.

[11] ZHANG H, ZHANG S. Focaler-IoU: more focused intersection over union loss[J]. arXiv preprint arXiv: 2401. 10525, 2024.

[12] 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.

[13] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles [J]. arXiv preprint arXiv: 2206. 02424, 2022.

[14] LEE Y, HWANG J, LEE S, et al. An energy and GPU-computation efficient backbone network for real-time object detection: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops [C]. Long Beach, USA: IEEE, 2019: 752-760.

[15] CHEN G, BUI T, KRZYZAK A. Image denoising using neighbouring wavelet coefficients[J]. Integrated Computer-Aided Engineering, 2005, 12(1): 99-107.

[16] PARTHASARATHY S, SANKARAN P. An automated multi scale retinex with color restoration for image enhancement: 2012 National Conference on Communications (NCC)[C]. Kharagpur, India: IEEE, 2012: 1-5.