

中文引用格式:荣海,习洲勇,李锦程,等. SPDs-Conv 与 WIoU 优化的 YOLOv8n 矿井井下人员检测模型[J]. 中国安全科学学报, 2026, 36(5): 139–149.

英文引用格式:Rong Hai, Xi Zhouyong, Li Jincheng, et al. YOLOv8n-based personnel detection model for underground mines optimized with SPDs-Conv and WIoU[J]. China Safety Science Journal, 2026, 36(5): 139–149.

SPDs-Conv 与 WIoU 优化的 YOLOv8n 矿井井下人员检测模型*

荣海^{1,2}副教授, 习洲勇^{**1,2}, 李锦程¹, 潘相伊³, 张伟达¹, 韩明玉⁴

(1 辽宁工程技术大学 矿业学院, 辽宁 阜新 123000; 2 辽宁工程技术大学 鄂尔多斯研究院,

内蒙古 鄂尔多斯 017010; 3 中国建筑东北设计研究院有限公司, 辽宁 沈阳 110004;

4 鄂尔多斯市昊华精煤有限责任公司, 内蒙古 鄂尔多斯 017200)

中图分类号: X936

文献标志码: A

DOI: 10.16265/j.cnki.issn1003-3033.2026.05.1332

资助项目: 2024 年辽宁省引导地市科技发展专项项目(20240341); 辽宁工程技术大学鄂尔多斯研究院校地培育项目(YJY-XD-2024-A-014)。

【摘要】 为解决煤矿井下环境由于光照不足、人员尺度差异大及易被设备遮挡等问题而导致的现有检测算法精度低、鲁棒性弱, 且部分模型参数量和计算量偏高, 难以适配井下边缘设备问题, 提出一种改进的 YOLOv8n 模型, 优化矿井复杂环境下的人员检测任务。引入改进的空间到深度可分离卷积(SPDs-Conv)模块增强小目标特征提取能力, 提升对远景低像素人员的识别精度; 跨阶段部分特征融合+选择性注意力机制(C2f_SKAttention)强化模型对不同尺度目标的关注度, 应对井下人员尺度差异; 构建动态检测头以适应目标多样性与复杂性, 提高对遮挡等场景的适配性; 改进加权交并比(WIoU)损失函数提升边界框定位精度, 减少因低照度导致的定位偏差。结果表明: 提出的改进 YOLOv8n 模型在煤矿井下人员检测数据集上的平均精度均值(mAP)@0.5 为 83.5%、mAP@0.5:0.95 为 39.0%, 相较于原 YOLOv8n, 精确率提升 8.5%, 召回率提升 11.9%, mAP@0.5 提升 4.7%, mAP@0.5:0.95 提升 3.3%; 参数量从 3.1M 增至 3.2M, 浮点运算数(GFLOPs)从 14.0G 增至 14.4G, 在轻量化基础上提升了检测精度与鲁棒性, 有助于解决井下小目标漏检、多尺度适配不足及复杂环境抗干扰弱等问题, 适配井下边缘设备算力限制。

【关键词】 空间到深度可分离卷积(SPDs-Conv); 加权交并比(WIoU); YOLOv8n; 井下人员检测; 轻量化; 注意力机制; 损失函数

YOLOv8n-based personnel detection model for underground mines optimized with SPDs-Conv and WIoU

Rong Hai^{1,2}, Xi Zhouyong^{1,2}, Li Jincheng¹, Pan Xiangyin³, Zhang Weida¹, Han Mingyu³

(1 College of Mining, Liaoning Technical University, Fuxin Liaoning 123000, China;

2 Ordos Research Institute, Liaoning Technical University, Ordos Inner Mongolia 017010, China;

3 China Northeast Architectural Design & Research Institute Co., Ltd., Shenyang Liaoning 110004, China;

4 Ordos City Haohua Coking Coal Co., Ltd., Ordos Inner Mongolia 017200, China)

* 文章编号: 1003-3033(2026)05-0139-11; 收稿日期: 2025-12-01; 修稿日期: 2026-02-26

** 通信作者: 习洲勇(1999—), 男, 江西吉安人, 硕士研究生, 研究方向为矿山安全与矿井动力灾害防治。E-mail: xizhy2021@163.com。

Abstract: To address the issues of low accuracy and weak robustness in existing detection algorithms due to insufficient lighting, scale differences among personnel, and frequent obstruction by equipment in coal mine environments, as well as the challenges posed by high parameter and computational requirements of some models, which make them difficult to adapt to edge devices underground, an improved YOLOv8n model was proposed to optimize personnel detection tasks in complex mine environments. An enhanced SPDs-Conv module was introduced to enhance the extraction of small target features and improve the recognition accuracy of low-pixel personnel in distant views. Cross stage partial feature fusion + selective kernel attention (C2f_SKAttention) module was designed to strengthen the model's focus on targets of different scales and cope with the scale differences of underground personnel. A dynamic detection head was constructed to adapt to the diversity and complexity of targets, and to improve robustness to occlusion and other scenarios. The WIoU loss function was improved to increase the bounding box localization accuracy and reduce the localization deviation caused by low illumination. The results show that the proposed improved YOLOv8n model achieves an mean average precision (mAP) @ 0.5 of 83.5% and an mAP@ 0.5 : 0.95 of 39.0% on the mine personnel detection dataset. Compared with the original YOLOv8n, the P is improved by 8.5%, the R by 11.9%, the mAP@ 0.5 by 4.7%, and the mAP@ 0.5 : 0.95 by 3.3%. The number of parameters only increases from 3.1M to 3.2M, and the Giga Floating-point operations per second (GFLOPS) rises from 14.0G to 14.4G. The proposed model maintains a lightweight structure while improving detection accuracy and robustness. It effectively alleviates missed detection of small underground targets, insufficient multi-scale adaptation and weak anti-interference capability in complex environments, making it suitable for the limited computing power of underground edge equipment.

Keywords: space-to-depth separable convolution (SPDs-Conv); weighted intersection over union (WIoU); YOLOv8n; underground personnel detection; lightweighting; attention mechanism; loss function

0 引言

煤矿开采过程中,人员检测是保障煤矿安全生产、推进智慧矿山建设的关键环节^[1-2]。然而,煤矿井下环境的特殊性给人员检测带来诸多严峻挑战,如井下巷道亮度低、光照条件差,摄像机部署位置较远,导致采集到的图像质量低下,人员目标特征不明显,检测难度极大^[3]。同时,由于巷道纵深较大,近景人员与远景人员在图像中的尺度差异显著,远景人员往往以小目标形式存在,像素占比极低,容易出现漏检情况^[4]。因此,研究并提出适用于煤矿井下复杂环境的高效人员检测方法,提升人员检测的准确性与鲁棒性,对保障煤矿生产安全、推动智慧矿山建设具有重要意义^[5]。

目前,国内外学者已开展了一系列针对性研究,尝试从算法优化^[6]、特征增强^[7]、模型轻量化^[8]等角度突破现有技术瓶颈,如王帅等^[9]提出轻量级人员检测模型,通过嵌入坐标注意力、设计加权多尺度融合模块、引入尺度敏感交并比(Scale-Sensitive Intersection over Union, SIoU)损失及通用硬件导向系统转移模块,缓解煤矿井下低照度和小目标漏检问题,提升检测精度与速度,适配资源受限设备。徐飞翔等^[10]提出基于遮挡目标检测网络,借助多感受

野边缘感知、边缘引导特征增强等模块,解决低光环境下目标大面积遮挡、多尺度特征互扰问题,提升检测精度与鲁棒性。然而,因报道不足等原因,目前较多研究未综合考虑煤矿井下光照条件差和模型部署困难的问题,无法满足井下人员检测对精度、速度和轻量化的实际需求。

因此,针对煤矿井下的特殊环境,笔者拟提出一种改进的 YOLOv8n 模型,用以专门优化井工矿复杂环境下的人员检测任务,并将空间到深度可分离卷积(Space-to-Depth separable Convolution, SPDs-Conv)模块、跨阶段部分特征融合+选择性注意力机制(Cross Stage Partial Feature Fusion + Selective Kernel Attention, C2f_SKAttention)模型模块、动态检测头、加权交并比(Weighted Intersection over Union, WIoU)损失函数 4 个模块融合后,提升模型的综合性能,以期有效应对小目标、多尺度及复杂环境等检测难点,满足井下人员检测对实时性与准确性的需求,为矿工的生命安全提供更可靠的保障。

1 矿井井下人员检测模型

YOLOv8n 作为 YOLOv8 系列轻量版本,以低参数量、低计算量适配资源受限边缘设备,通过跨阶段部分特征融合模块(Cross Stage Partial Feature

Fusion, C2f) 提取特征、路径聚合网络 (Path Aggregation Network, PANet) 融合特征、解耦头分离分类与回归任务,兼具检测速度与精度优势。但应用于矿井井下人员检测时,存在小目标检测能力薄弱等适配性问题。井下场景对模型要求严苛,人员多为

远距离小目标、近远景尺度差异大、环境干扰强,且边缘设备算力有限。虽 YOLOv8n 具备轻量化基础,但精度与效率的平衡仍需优化,故需针对性改进以提升井下人员检测性能。改进后的模型结构如图 1 所示。

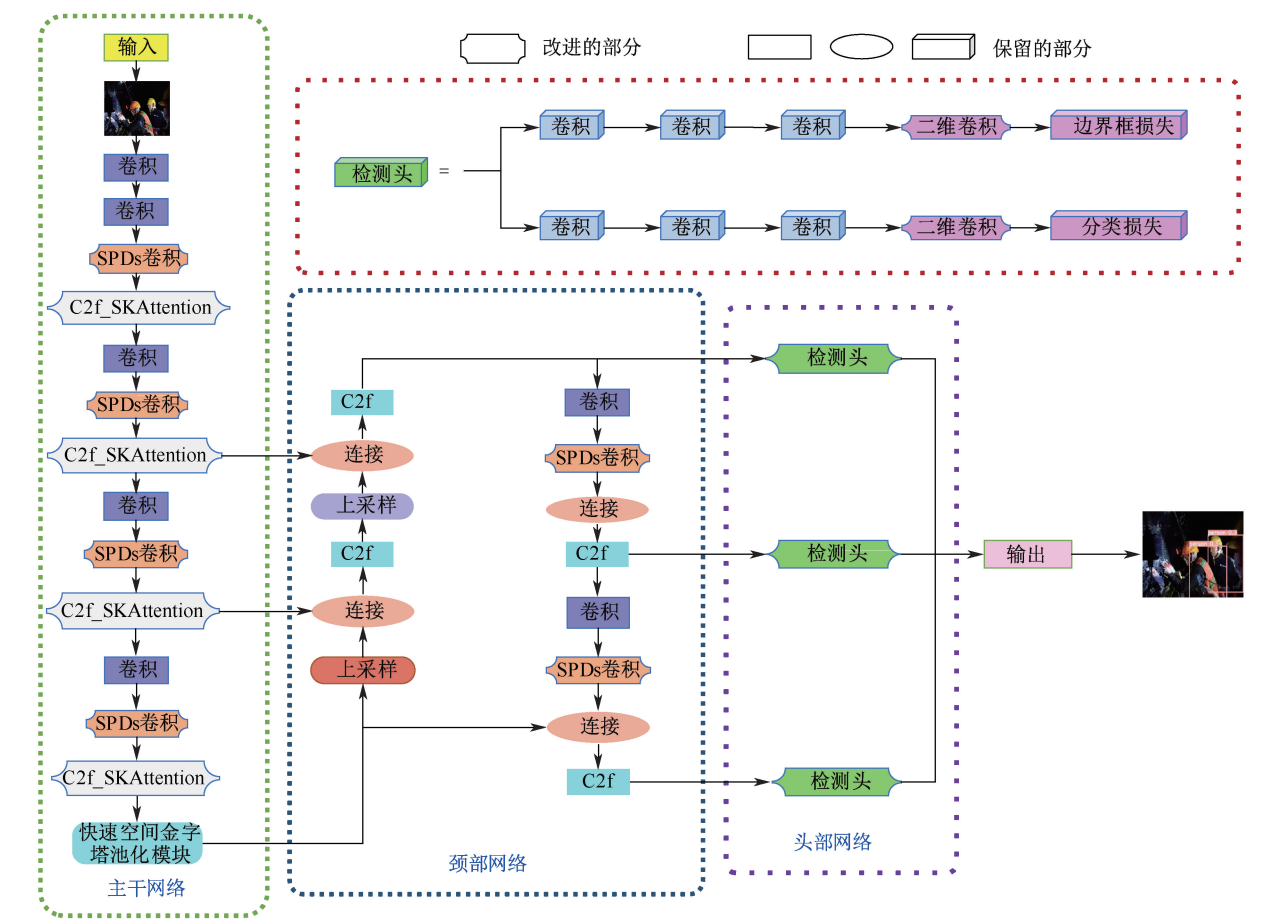


图 1 改进后的模型结构
Fig. 1 Improved model structure

2 改进的 YOLOv8n 井下人员检测模型

2.1 引入 SPDs-Conv 模块

SPDs-Conv 是 SPD-Conv 的轻量化改进版,核心是保留小目标特征提取能力的同时降低计算量与参数量,适配井下边缘设备等资源受限场景。针对巷道纵深导致的远距离人员小目标检测难题,该模块通过空间到深度转换保留小目标精细特征,结合 SPDs-Conv 减少参数冗余,既增强远距离小目标捕捉能力,又满足井下设备算力要求,有效降低巷道深处人员漏检率,精准匹配井下远景人员检测需求。

结构上,SPDs-Conv 由空间到深度层和非跨步卷积层组成。特征处理流程为:对通道数 C_1 、空间尺寸为 $S \times S$ 的输入特征图,按空间下采样系数 2 分割为 4 张 $S/2 \times S/2$ 的子特征图,经 $\oplus$ 深度维度叠加后通道数变为 $4C_1$;之后进入 SPDs-Conv,深度卷积 (Depthwise Convolution, DWConv) 先对各通道独立卷积提取空间特征,再经一次 DWConv 变换信息,最终输出通道数 C_2 、尺寸 $S/2 \times S/2$ 的特征图,如图 2b 所示;原始 SPD-Conv 的结构如图 2a 所示,两者差异在于 SPDs-Conv 引入了 SPDs-Conv 来实现轻量化优化。具体改进过程及处理逻辑如图 2 所示。

首先,针对输入特征图 $X \in (S, S, C_1)$,其中, C_1 为通道数,SPDs-Conv 会依据比例因子 $Q=2$ 划分

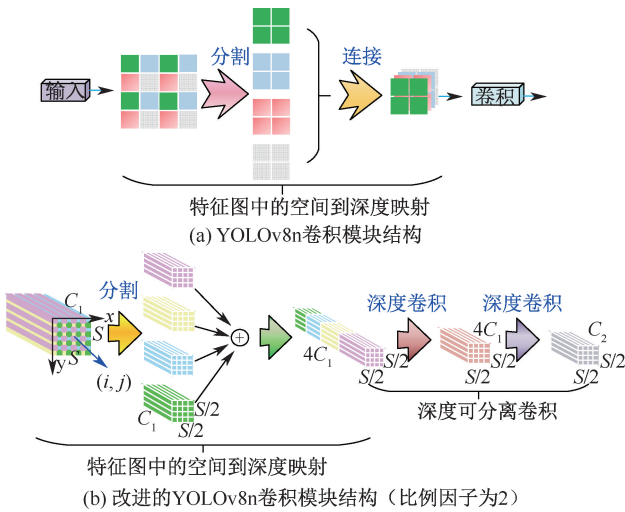


图2 卷积模块结构

Fig. 2 Structure of convolution module

子特征映射。划分过程通过一系列索引操作实现,见下式。

以此类推,得到4个子映射 $f_{0,0}, f_{1,0}, f_{0,1}, f_{1,1}$,每个尺寸均为 $S/2 \times S/2 \times 4C_1$ 。首先,通过等比例下采样保留局部细节,避免传统跨步卷积丢失小目标关键信息,为后续特征提取奠基。然后,将子映射沿通道维度连接,得到新特征图 $X'(S/2 \times S/2 \times 4C_1)$,实现空间信息到深度信息的转换,把空间维度的小目标精细特征编码到通道维度,减少空间维度的同时最大限度保留小目标特征,解决传统卷积细粒度信息易丢失的问题。最后,SPDs-Conv的改进在于采用SPDs-Conv处理特征图 X' ,这是与原始SPD-Conv形成差异的核心所在。

$$\begin{aligned}
 f_{0,0} &= X[0:S:Q, 0:S:Q], \\
 f_{1,0} &= X[1:S:Q, 0:S:Q], \\
 &\vdots, \\
 f_{Q,0} &= X[Q-1:S:Q, 0:S:Q]; \\
 f_{0,1} &= X[0:S:Q, 1:S:Q], \\
 f_{1,1} &= X[1:S:Q, 1:S:Q], \\
 &\vdots, \\
 f_{Q-1,1} &= X[Q-1:S:Q, 1:S:Q]; \\
 &\vdots, \\
 f_{0,Q-1} &= X[0:S:Q, Q-1:S:Q], \\
 f_{1,Q-1} &= X[1:S:Q, Q-1:S:Q], \\
 &\vdots, \\
 f_{Q-1,Q-1} &= X[Q-1:S:Q, Q-1:S:Q] \quad (1)
 \end{aligned}$$

式中: X 为输入的特征图张量; S 为分割的步长/间隔; i 和 j 为特征图的空间索引序号, $i \in (0, S), j \in (0, S); Q$ 为比例因子。

SPDs-Conv操作分为2个环节:深度卷积对每个通道单独卷积,减少跨通道参数冗余,得到同输入通道数的特征图;逐点卷积用 $1 \times 1 \times C_2$ 卷积核处理,将特征图维度从 $X'(S/2 \times S/2 \times 4C_1)$ 变换为 $X''(S/2 \times S/2 \times C_2)$ 。该方式在调整通道数、增强小目标特征表达的同时,减少跨通道参数,降低计算复杂度,使得改进后的模块更适合在算力有限的移动平台上部署。

子特征映射通过精确索引步长划分(如 $Q=2$),确保均匀覆盖原始特征图,为信息转换提供完整输入;空间转深度操作通过通道扩展弥补空间压缩的信息损失,有效保留小目标特征;深度可分离卷积拆分操作,在保持特征提取能力的同时降低参数规模和计算量,实现轻量化改进,使SPDs-Conv在增强小目标检测性能的同时,满足了移动平台对资源消耗的限制要求。

2.2 尺寸自适应注意力模块的改进

C2f_SKAttention是在YOLOv8n原有C2f模块基础上,通过嵌入SKAttention模块改进而来,其结构如图3a所示,核心是实现了对多尺度信息的动态调整,以增强模型对不同尺度目标的特征捕捉能力。该模块首先对输入特征图 $X_i \in R^{H \times W \times C}$ (i 为特征图索引序号, H 和 W 分别为特征图的高和宽),执行切片操作,之后分别利用 3×3 和 5×5 卷积核将其转换为 X_{i1} 与 X_{i2} ,再通过元素求和的方式融合得到 X'_i ;随后借助全局平均池化融入全局信息,生成通道级统计数据 $T(T \in R^C)$,并通过全连接层创建特征 $z = \mathcal{T}_{fc}(T)$ 以提供选择指导;然后利用跨通道的软关注力机制,经Softmax运算得到概率 a_c 和 b_c ,进而通过下式选择性强调或抑制融合特征;最后将结果与输入特征图 X_i 相加构成残差网络,输出 X_{i+1} ,并通过颈部多个串联的C2f模块,结合跨层连接强化多尺度特征提取与融合。改进后的C2f模块结构如图3所示。

$$T_c = \mathcal{T}_{gp}(X'_{ic}) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X'_{ic}(i, j) \quad (2)$$

$$\begin{aligned}
 a_c &= \frac{\exp(A_c z)}{\exp(A_c z) + \exp(B_c z)} \\
 b_c &= \frac{\exp(B_c z)}{\exp(A_c z) + \exp(B_c z)} \quad (3)
 \end{aligned}$$

$$V_c = a_c X_{i1c} + b_c X_{i2c} \quad (4)$$

式中: C 为通道数; T_c 为通道级统计数据向量的第 C 个元素; $\mathcal{T}_{gp}$ 为全局平均池化; a_c, b_c 分别为最终计算

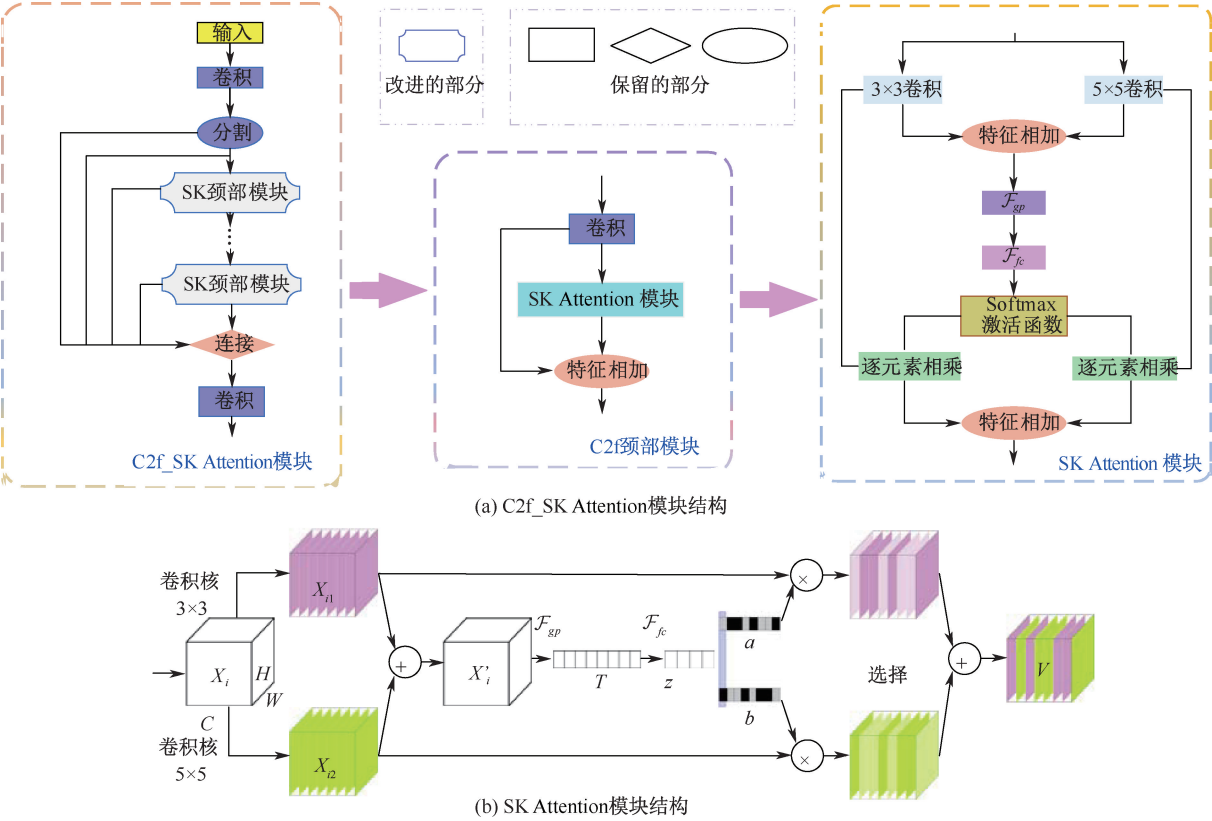


图 3 改进前后 C2f 模块结构
Fig. 3 Structure of C2f module before and after improvement

的通道权重, $a_c + b_c = 1$; $\mathbf{A}$ 、 $\mathbf{B}$ 为注意力向量; $\exp(\mathbf{A}_c \mathbf{z})$ 、 $\exp(\mathbf{B}_c \mathbf{z})$ 分别为通道特征的指数变换结果; $\mathbf{z}$ 为通道选择指导特征向量; X_c 为空间位置 (i, j) 的特征图; V 为最终特征图。

C2f_SKAttention 模块结构如图 3a 所示, 该模块通过 3×3 和 5×5 卷积核分别捕捉不同尺度特征, 并借助注意力机制动态调整权重, 能差异化关注近景人员细节特征和远景人员微弱特征; 同时残差连接和多模块串联确保特征不丢失, 增强模型对井下复杂尺度分布人员的适应能力, 减少因尺度变化导致的漏检或误检。井下近景人员与远景小目标人员尺度差异为 $5 \sim 10$ 倍, 3×3 卷积核可捕捉近景人员衣物纹理、姿态等细节特征, 5×5 卷积核能完整包裹远景小目标整体轮廓, 注意力机制动态分配权重, 使模型对近景、远景人员的识别精度得以提升, 解决巷道纵深带来的尺度适配难题。感知野公式见下, 二者在 C2f 模块感知野为 $11 \sim 43$ 、 $19 \sim 75$ 像素, 覆盖 90% 人员尺寸。测试集上该组合平均精度均值 (Mean Average Precision, mAP) @ 0.5 更高, 且帧率 (Frames Per Second, FPS) 降低, 但参数量仅些许增加。该模块双分支结构通过不同尺寸卷积核提取局部细节与全局上下文信息, 注意力机制动态融合二者, 从而增

强模型对尺度变化的适应能力。

$$F_n = F_{n-1} + (p - 1) \prod_{i=1}^{n-1} S_i \quad (5)$$

式中: F_n 为第 n 层网络的感知野大小; F_{n-1} 为第 $n-1$ 层的感知野大小; p 为当前层卷积核大小; S_i 为第 i 层的下采样步长。

SK 注意力模块结构如图 3b 所示, 改进前的 SKAttention 模型通过多尺度卷积拆分输入特征、融合后动态加权输出; 改进后的 C2f_SKAttention 模型在此基础上, 前置流程经输入模块初步变换、多分支 SK 颈部模块做通道加权, 借跨阶段连接整合特征, 还细化 SKAttention 流程。改进后强化多尺度特征提取与整合, 解决巷道纵深远景人员小目标漏检问题; 精准抑制冗余通道, 应对井下粉尘等干扰, 提升检测精度; 以空间拆分、跨阶段连接优化计算, 适配井下边缘设备, 解决部署难、推理慢问题; 残差结构助力梯度回传, 增强模型训练稳定性, 保障复杂环境运行可靠, 让检测更适配井下场景。

2.3 动态检测头的构建

基于注意力机制的动态检测头, 整合尺度、空间、任务 3 种注意力机制改进目标检测头, 核心是增

强目标特征表达、降低背景干扰与遮挡影响。该检测头围绕特征张量 $F \in \mathbb{R}^{L \times U \times C}$ 处理,通过分解注意力函数为3个连续单维度:尺度感知注意力动态融合多尺度特征,空间感知注意力借可变形卷积聚合空间特征,任务感知注意力动态调整通道重要性($\theta = [\alpha_1, \alpha_2, \beta_1, \beta_2]$, θ 为学习超函数),原理见下式。3种注意力嵌套迭代,逐步提取丰富区分性特征。

$$W(F) = \pi_c(\pi_u(\pi_L(F)F)F)F \quad (6)$$

$$\pi_L(F)F = \sigma\left(f\left(\frac{1}{U \cdot C} \sum_{u,c} F\right)\right)F \quad (7)$$

$$\pi_u(F)F = \frac{1}{L} \sum_{l=1}^L \sum_{k=1}^K w_{l,k} F(l; p_k + \Delta p_k; c) \Delta m_k \quad (8)$$

$$\pi_c(F)F = \max(\alpha^1(F)F_c + \beta^1(F), \alpha^2(F)F_c + \beta^2(F)) \quad (9)$$

式中: $W(F)$ 为输出的加权特征; $\pi_L(\cdot)$ 、 $\pi_u(\cdot)$ 、 $\pi_c(\cdot)$ 分别为尺度(L)、空间(U)和任务(C)维度的

感知注意力函数; σx 为归一化函数; U 为空间尺寸; L 为尺度总数; l 为尺度索引; K 为子尺度总数; k 为单个尺度内的子尺度; $w_{l,k}$ 为尺度的权重系数; Δm_k 为尺度调整因子; $\alpha(F)$ 、 $\beta(F)$ 为特征变换函数。

3大注意力模块遵循“从粗到细、逐层聚焦”的协同逻辑,构建递进式特征优化链: π_L 先对特征张量 $F \in \mathbb{R}^{L \times U \times C}$ 进行尺度优化,通过计算全局均值动态分配权重,突出人员适配尺度特征、过滤无关干扰,为空间聚焦奠基。 π_u 在尺度优化后的特征上,通过可变形卷积学习自适应性偏移量($p_k + \Delta p_k$),聚合人员关键空间特征、抑制背景干扰,遮挡场景下可通过偏移量避开遮挡区域。 π_c 针对分类与回归两大任务,动态调整通道重要性——对分类任务,增强纹理特征通道权重。其中。对回归任务,增强轮廓特征通道权重,通过 $\max(\alpha_1 F_c + \beta_1, \alpha_2 F_c + \beta_2)$ 的双分支选择,适配井下低光照导致的特征模糊问题,提升任务适配性。其结构设计如图4所示。

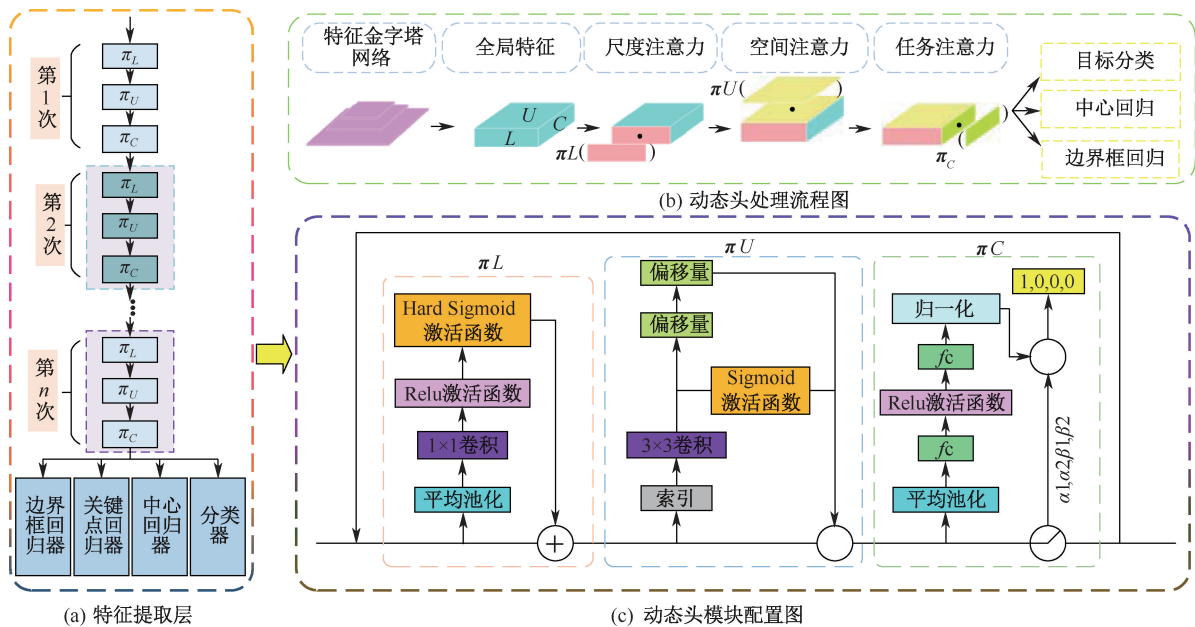


图4 动态检测头结构设计

Fig. 4 Dynamic detection head structure design

2.4 损失函数的改进

在对矿井井下人员检测任务中,提出 WIoU 损失函数,包括 WIoUv1、WIoUv2 和 WIoUv3,核心是通过动态调整损失权重,提升模型对目标的定位精度和鲁棒性。WIoUv1 通过距离注意力放大普通锚框的损失,见下式:

$$L_{WIoUv1} = R_{WIoU} L_{IoU} \quad (10)$$

$$L_{IoU} = 1 - I_{ou} \quad (11)$$

$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (12)$$

式中: L_{IoU} 为损失函数; I_{ou} 为重叠程度; R_{WIoU} 用于度量预测框与真实框的中心距离; W_g^2 和 H_g^2 为并集框的宽和高; L_{WIoUv1} 为加权后的总损失; x 、 y 和 x_{gt} 与 y_{gt} 分别为预测框和真实框的中心坐标。

WIoUv2 引入单调聚焦系数,减弱简单样本对损失的影响,加速收敛,原理见下式:

$$L_{\text{WIoUv2}} = \left(\frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \right)^\gamma L_{\text{WIoUv1}} \quad (13)$$

式中: L_{IoU}^* 为目标阈值距离损失; γ 为自定义的超参数, $\gamma > 0$ 。

WIoUv3 则通过离群度 β 构建非单调聚焦系数, 通过动态调整梯度增益, 聚焦普通质量锚框, 避免低质量锚框产生有害梯度。 r 为动态计算的加权系数, α, δ 为超参数, 原理见下式:

$$L_{\text{WIoUv3}} = r L_{\text{WIoUv1}} \left(r = \frac{\beta}{\delta \alpha^{\beta-\delta}}, \beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \right) \quad (14)$$

将 WIoU 系列应用于矿井井下人员检测, 推理如下: 井下低光照易致人员特征模糊、定位精度下降, 且存在近远景多尺度场景, 设备遮挡还会产生低质量锚框。 WIoUv1 通过距离注意力机制增强中心偏移敏感性, 解决低光照/粉尘导致的定位偏差; WIoUv3 以非单调聚焦机制平衡样本贡献, 降低遮挡低质量锚框的有害梯度影响; WIoUv2 则通过单调聚焦系数缓解样本不均衡、加速收敛, 三者分别适配模糊、遮挡、样本不均衡场景, 全面覆盖井下定位难点。超参数 $\gamma = 1.2, \delta = 0.3$ 由网格搜索确定, WIoU 系列不增加额外计算量, 适配井下边缘设备, 与特征提取模块协同可减少漏检、提升检测精度。

3 检测模型效果与特征可视化分析

3.1 井下人员检测数据集构建

试验数据集来自内蒙古、山西等煤炭主产区典型矿井, 覆盖主流煤矿类型与井下人员检测核心场景, 共 3 676 张图像。数据集按“环境干扰类型”与“人员尺度特征”分类, 完整呈现井下人员多种状态, 按 4 : 1 比例划分为训练集 (2 941 张) 与测试集 (735 张), 并借助 Labelimg 标注工具, 标注图像中的人员目标, 为模型训练和性能评估提供了丰富且真实的样本支撑, 部分数据集如图 5 所示。

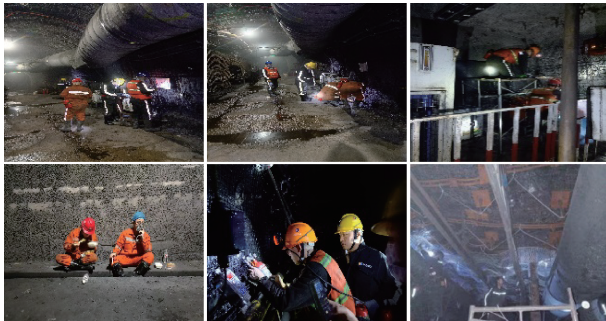


图5 部分数据集

Fig. 5 Partial datasets

3.2 井下人员检测评价指标

在矿井井下人员检测中, 综合评价指标的构建必须紧密贴合井下特殊场景的复杂性与严苛性, 需从检测精度、模型效率及环境适应性 3 个核心维度进行系统设计, 以全面、客观地衡量模型的实际应用价值, 其中主要包括精确率 P 、召回率 R 、mAP (mAP @ 0.5、mAP @ 0.5 : 0.95) 和浮点运算量 (Giga Floating-point Operations Per Second, GFLOPs) 等评价指标, 具体公式^[11-12]见下式:

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (15)$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (16)$$

$$\text{AP} = \int_0^1 P(R) dR \quad (17)$$

$$\text{mAP} = \frac{1}{N} \sum_{m=1}^N A_p(m) \quad (18)$$

式中: TP (True Positive) 为正确检测的样本; FP (False Positive) 为误检的样本; FN (False Negative) 为模型漏检的样本; AP (Average Precision) 为单类别 Precision-Recall 曲线下面积; N 为卷积核总数; m 为卷积核索引。

将 YOLOv8n 模型和改进后的 YOLOv8n 模型放入数据集中测试, YOLOv8n 模型改进前后的各类评价指标对比如图 6 所示。

由图 6a、图 6b 可知: YOLOv8n 模型在迭代 150 次前边界框损失更小, 迭代 150 次之后改进后模型反超并趋近于 0, 收敛更优; 改进后模型的分类损失始终更低, 小目标等场景精度更突出。图 6c、图 6d 验证了改进策略的有效性。由 2 图可知: 改进后模型训练前期增速更快、后期更平稳高精度, 表明优化模块提升了收敛性与鲁棒性, 适配矿井检测需求。

3.3 损失函数对比

在矿井井下人员检测中, 对比 WIoU 系列 (Ours) 与交并比 (Intersection over Union, IoU)、完全交并比 (Complete Intersection over Union, CIoU)、广义交并比 (Generalized Intersection over Union, GIoU)、距离交并比 (Distance Intersection over Union, DIoU) 等损失函数的性能对比, 对比结果见表 1。由表 1 可知: 在矿井井下人员检测任务中, WIoU 系列 (WIoUv1、WIoUv2、WIoUv3) 在检测精度上全面优于 IoU、SIoU 等传统损失函数, 且计算效率无额外损耗: 在精确率 P 上, WIoUv3 以 90.3% 领

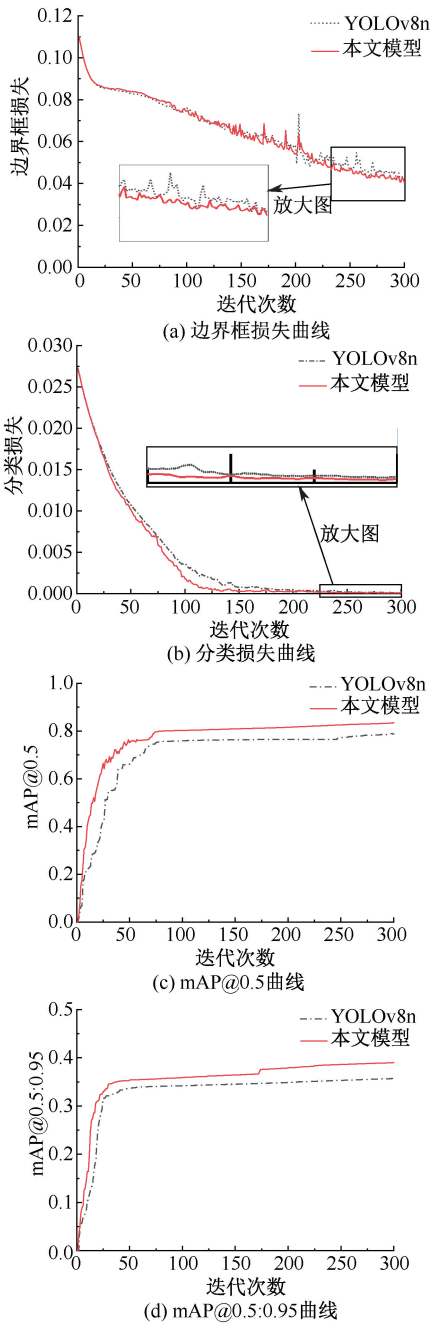


图6 模型损失和性能指标曲线对比

Fig. 6 Comparison of model loss and performance indicator curves

先,较IoU(86.2%)、CIoU(87.5%)分别提升4.1%、2.8%;召回率 R 方面,WIoUv3为89.0%,较IoU(84.8%)、DIoU(85.9%)分别提升4.2%、3.1%;核心定位精度指标 $mAP@0.5$ 中,WIoUv3的89.7%较IoU(85.5%)、CIoU(86.8%)分别提升4.2%、2.9%;多阈值鲁棒性指标 $mAP@0.5:0.95$ 上,WIoUv3以88.3%优于IoU(83.9%)、SIoU(84.9%)等,且WIoU系列内部呈现 $v3>v2>v1$ 的递进优化趋势,同时所有

损失函数的计算量均为14.2,证明WIoU系列在不增加计算量的前提下,提升了井下人员检测的精准度与鲁棒性,尤其适配井下复杂环境中远距离小目标、遮挡人员的检测需求。

3.4 消融试验

试验以原始检测模型为基准,通过逐一添加改进高效注意力机制、轻量化特征融合模块、优化损失函数等模块,以此判断各模块的有效性及其组合优化的合理性,试验结果见表2。由表2可知:基准模型精确率 P 为67.6%、召回率 R 为76.4%、 $mAP@0.5$ 为78.8%、 $mAP@0.5:0.95$ 为35.7%,而4个模块全启用后,精确率 P 提升至76.1%、召回率 R 提升至88.3%、 $mAP@0.5$ 提升至83.5%、 $mAP@0.5:0.95$ 提升至39.0%,计算量仅从14.0增至14.4,增幅2.9%,计算量控制合理;单一模块中,WIoU提升效果最突出,单独启用时精确率 P 为72.3%、召回率 R 为80.1%、 $mAP@0.5$ 为81.8%,较基准分别提升4.7%、3.7%、3.0%;动态检测头模块次之,单独启用时 R 为79.5%、 $mAP@0.5:0.95$ 为37.1%,较基准分别提升3.1%、1.4%;单独使用SPDs-Conv模块时 $mAP@0.5$ 为80.1%、 $mAP@0.5:0.95$ 为36.5%,较基准分别提升1.3%、0.8%,参数量从3.1降低至2.8,降低了9.6%,FPS从205.6降至155.1,降幅24.6%,体现了该模块的轻量化,尽管计算量从14.0 G微增至14.1 G,但参数量的降低更契合井下边缘设备对存储空间的严苛限制,且该模块与其他3个模块组合后,参数量仅回升至3.2 M,充分证明SPDs-Conv模块在不显著增加整体计算负担的前提下,通过降低核心参数量实现了轻量化设计目标,适配井下资源受限设备的部署需求。尺寸自适应注意力模块单独启用时 $mAP@0.5$ 为79.2%,较基准提升0.4%,也均有正向贡献;且模块组合性能优于单一模块,如SPDs-Conv模块+尺寸自适应注意力模块组合时 $mAP@0.5$ 为81.0%,较2模块单独使用最高值80.1%提升0.9%,4模块全组合各项指标均为最优,证明各单一模块对矿井人员检测性能有明确增益,协同作用更能放大精度提升效果,充分验证改进策略与井下人员检测特点的适配性。

3.5 不同算法的试验结果对比

为进一步验证文中文模型的优越性,将文中算法与目前主流算法进行对比试验,训练参数和训练环境统一,对比试验结果见表3。

表 1 损失函数对比

Table 1 Loss function comparison

模块	$P/\%$	$R/\%$	mAP@ 0. 5/%	mAP@ 0. 5;0. 95/%	GFLOPs
IoU	86. 2	84. 8	85. 5	83. 9	14. 2
CIoU	87. 5	86. 1	86. 8	85. 3	14. 2
SIoU	87. 1	85. 7	86. 4	84. 9	14. 2
GIoU	86. 7	85. 3	86. 0	84. 5	14. 2
DIoU	87. 3	85. 9	86. 6	85. 1	14. 2
WIoUv1	88. 8	87. 6	88. 2	86. 8	14. 2
WIoUv2	89. 5	88. 2	88. 9	87. 5	14. 2
WIoUv3	90. 3	89. 0	89. 7	88. 3	14. 2

表 2 消融试验结果

Table 2 Ablation test results

SPDs-Conv	尺寸自适应注意力	动态检测头	WIoU	$P/\%$	$R/\%$	mAP@ 0. 5/%	mAP@ 0. 5; 0. 95/%	参数量/M	GFLOPs/G	FPS/(帧/s)
×	×	×	×	67. 6	76. 4	78. 8	35. 7	3. 1	14. 0	205. 6
√	×	×	×	70. 2	78. 1	80. 1	36. 5	2. 8	14. 1	155. 1
×	√	×	×	69. 5	77. 3	79. 2	36. 2	3. 5	14. 1	176. 5
×	×	√	×	71. 8	79. 5	80. 3	37. 1	3. 1	14. 2	145. 2
×	×	×	√	72. 3	80. 1	81. 8	37. 5	3. 3	14. 2	179. 4
√	√	×	×	73. 5	81. 2	81. 0	38. 0	3. 0	14. 2	85. 1
√	√	√	×	75. 2	82. 7	81. 8	38. 7	3. 1	14. 4	61. 5
√	√	√	√	76. 1	88. 3	83. 5	39. 0	3. 2	14. 4	45. 2

表 3 不同算法对比试验结果

Table 3 Comparison test results of different algorithms

模型	$P/\%$	$R/\%$	mAP@ 0. 5/%	mAP@ 0. 5;0. 95/%	GFLOPs/G	参数量/M	FPS/(帧/s)
YOLOv4	59. 5	68. 6	70. 8	26. 2	70. 8	64. 0	39. 5
YOLOv5s	74. 5	84. 6	62. 3	17. 6	16. 0	7. 2	95. 2
YOLOv5l	64. 2	72. 5	70. 8	23. 9	108. 4	44. 0	50. 6
YOLOv7-tiny	70. 5	78. 3	65. 0	18. 5	13. 2	6. 0	126. 1
YOLOv8s	68. 1	75. 6	69. 1	23. 6	28. 8	11. 1	165. 2
YOLOv8n	67. 6	76. 4	78. 8	35. 7	14. 0	3. 1	205. 6
YOLOv10s	71. 5	81. 5	79. 1	23. 5	24. 5	8. 1	143. 0
YOLOv11n	74. 9	85. 6	73. 3	19. 4	6. 3	2. 6	188. 1
Ours	76. 1	88. 3	83. 5	39. 0	14. 4	3. 2	45. 2

由表 3 可知:文中改进的 YOLOv8n 模型(Ours)在矿井人员检测任务中,精度与轻量化特性均展现显著优势:在核心精度指标上,精确率 P 为 76. 1%、 R 为 88. 3%、mAP@ 0. 5 为 83. 5%、mAP@ 0. 5;0. 95 为 39. 0%,较原 YOLOv8n(精确率 P 为 67. 6%、召回率为 76. 4%、mAP@ 0. 5 为 78. 8%、mAP@ 0. 5;0. 95 为 35. 7%)分别提升 8. 5%、11. 9%、4. 7%、3. 3%,较轻量化模型 YOLOv5s(精确率 P 为 74. 5%、召回率 R 为 84. 6%、mAP@ 0. 5 为 62. 3%)分别提升 1. 6%、3. 7%、21. 2%,较 YOLOv11n(精确率 P 为 74. 9%、召回率 R 为 85. 6%、mAP@ 0. 5 为 73. 3%)分别提升

1. 2%、2. 7%、10. 2%;在轻量化与效率指标上,其参数量为 3. 2 M、计算量为 14. 4 G,仅比原 YOLOv8n(3. 1 M/14. 0 G)微增 3. 2%/2. 9%,远低于 YOLOv5s、YOLOv8s 等模型,虽 FPS 为 45. 2 低于部分轻量化模型,但仍能满足井下边缘设备实时检测需求,整体实现“精度更高、轻量化适配性更强”的优化目标。

3. 6 可视化分析

为直观验证改进模型在煤矿井下复杂环境中的检测效果,通过可视化手段展示不同场景下的人员

检测结果,模型改进效果可视化如图 7 所示。

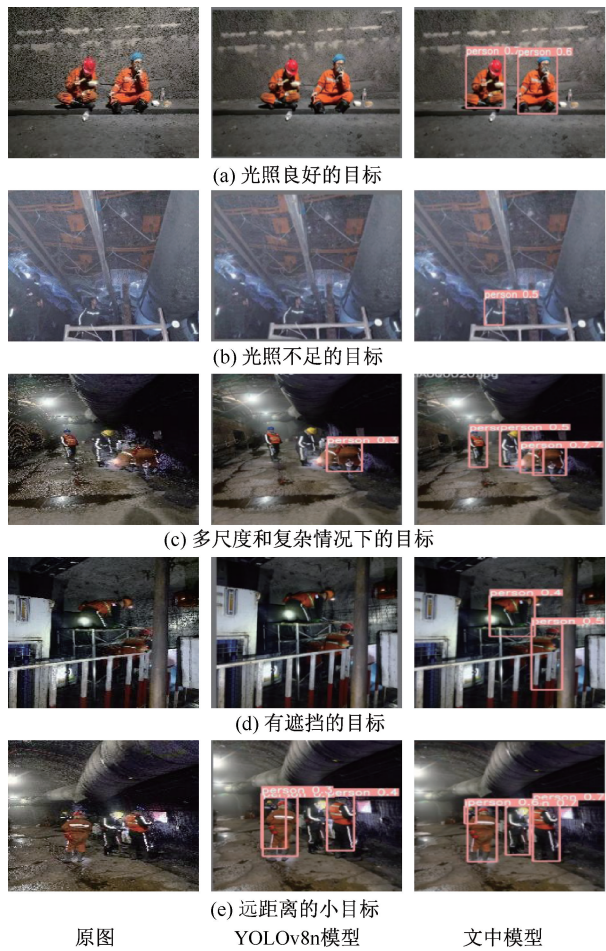


图 7 模型改进效果可视化

Fig. 7 Visualization of model improvement effects

煤矿井下普遍存在光线明暗不均、目标尺度多变、遮挡频繁及远距离小目标难识别等问题,影响检测模型的实际应用效果。由图 7 可知:第 1 列是原始图像,第 2 列是改进前模型检测结果,第 3 列是文中模型检测结果;图 7a、图 7b 中,原模型对光照明暗目标存在漏检,文中模型均能检测到,其中光照不足场景(图 7b)漏检率下降;图 7c、图 7d 中,原模型在遮挡及多尺度场景漏检明显,文中模型检测效果更优;图 7e 针对远距离小目标,原模型漏检突出,其精确率 68.7%、召回率 65.4%,文中模型提升至

82.1%、79.8%,漏检率从 24.6%降至 10.2%。可见文中模型对上述场景的检测效果均有优化,提升了矿井人员检测的精度与鲁棒性。

4 结 论

1) 提出的改进模型通过引入 SPDs-Conv 模块等,可有效提升检测性能,mAP@ 0.5 为 83.5%、mAP@ 0.5:0.95 为 39.0%,优于主流对比模型。

2) SPDs-Conv 模块能够减少小目标漏检,C2f_SKAttention 模块缓解尺度差异导致的性能波动,动态检测头提升定位精度,WIoU 损失函数增强低质量目标定位鲁棒性。各模块均针对井下人员检测特点定向设计:SPDs-Conv 匹配远景小目标特征提取需求,C2f_SKAttention 适配近远景尺度差异,动态检测头应对遮挡干扰,WIoU 损失函数解决模糊目标定位偏差,四者协同形成小目标提取-多尺度适配-抗干扰检测-精准定位的全链路优化,全面覆盖井下人员检测的核心痛点。

3) 消融试验验证了各模块有效性,其中,动态检测头和 WIoU 损失函数提升效果突出,且模块组合后性能优于单一模块,体现改进策略在特征增强、注意力机制及损失优化等方面的协同合理性。

4) 改进模型在轻量化基础上能够实现精度与效率的平衡,参数量 3.2 M、GFLOPs14.4 G,适配井下边缘设备算力限制,可视化结果也证明其在光线不均、遮挡、远距离小目标等复杂场景的检测效果更优。

5) 数据集虽涵盖薄煤层、厚煤层等矿井,但未涉及南方高湿度矿井,高湿度导致的图像雾化问题未纳入研究,模型在该场景下的泛化性需进一步验证。未来需采集高湿度、高水雾等矿井数据,引入图像去雾化算法与 GFLOPs 适配的特征增强模块,提升模型在特殊湿度场景的泛化性。

5 致 谢

特别感谢辽宁省阜新市科技局、内蒙古自治区鄂尔多斯市政府的资助与支持。

参 考 文 献

[1] Diwan T, Anirudh G, Tembhurne J V. Object detection using YOLO: challenges, architectural successors, datasets and applications[J]. Multimedia Tools and Applications, 2023, 82(6): 9 243-9 275.

[2] Rong Hai, He Longyue, Li Nannan, et al. Technical research on surrounding rock control of roadways crossing collapse columns in strong mine pressure working faces[J]. Scientific Reports, 2025, 15(1): DOI:10.1038/s41598-025-15850-8.

- [3] 邵小强, 李鑫, 杨永德, 等. 基于改进 YOLOv7 的矿井人员检测算法[J]. 电子科技大学学报, 2024, 53(3): 414–423.
- Shao Xiaoqiang, Li Xin, Yang Yongde, et al. Mine personnel detection algorithm based on improved YOLOv7[J]. Journal of University of Electronic Science and Technology of China, 2024, 53(3): 414–423.
- [4] 贾桂敏, 程羽, 齐孟飞. 多尺度注意力特征增强融合的红外小目标检测新网络[J]. 中国安全科学学报, 2024, 34(6): 90–98.
- Jia Guimin, Cheng Yu, Qi Mengfei. Multi-scale attention feature-enhanced fusion of a new network for infrared small object detection[J]. China Safety Science Journal, 2024, 34(6): 90–98.
- [5] 樊红卫, 闫昕山, 曹现刚, 等. 一种 DSP-YOLO 煤矿带式输送机人员不安全状态智能检测算法[J/OL]. 煤炭科学技术: 1–18. [2025–08–06]. <https://link.cnki.net/urlid/11.2402.TD.2025.0605.1557.008>.
- Fan Hongwei, Yan Xinshan, Cao Xiangang, et al. A DSP-YOLO intelligent detection algorithm of unsafe personnel state for coal mine belt conveyor[J/OL]. Coal Science and Technology: 1–18. [2025–08–06]. <https://link.cnki.net/urlid/11.2402.TD.2025.0605.1557.008>.
- [6] Alif M A R, Hussain M. YOLOv1 to YOLOv10: a comprehensive review of YOLO variants and their application in the agricultural domain[EB/OL]. arXiv preprint arXiv: 2024; DOI:10.48550/arXiv.2406.10139.
- [7] Yang Wenjuan, Wang Yanqun, Zhang Xuhui, et al. MSS-YOLO: multi-scale edge-enhanced lightweight net-work for personnel detection and location in coal mines[J]. Applied Sciences, 2025, 15(6): DOI:10.3390/app15063238.
- [8] Wang Zheng, Hua Zhixin, Wen Yuchen, et al. E-YOLO: recognition of estrus cow based on improved YOLOv8n model[J]. Expert Systems with Applications, 2024, 238; DOI:10.1016/j.eswa.2023.122212.
- [9] 王帅, 杨伟, 李宇翔, 等. 基于注意力机制和特征融合的井下轻量级人员检测方法[J]. 煤炭科学技术, 2025, 53(4): 383–392.
- Wang Shuai, Yang Wei, Li Yuxiang, et al. A lightweight personnel detection method for underground coal mines[J]. Coal Science and Technology, 2025, 53(4): 383–392.
- [10] 徐飞翔, 张端, 程德强, 等. LECODNet: 基于边缘引导与上下文感知的低光环境下无人矿卡遮挡目标检测网络[J/OL]. 煤炭学报: 1–14. [2025–08–06]. <https://doi.org/10.13225/j.cnki.jccs.2025.0511>.
- Xu Feixiang, Zhang Rui, Cheng Deqiang, et al. LECODNet: edge-guided and context-aware occlusion detection network for autonomous mining trucks under low-light condition[J/OL]. Journal of China Coal Society: 1–14. [2025–08–06]. <https://doi.org/10.13225/j.cnki.jccs.2025.0511>.
- [11] 赵涛, 董立红, 秦昶. 融合注意力和轻量级网络的井下人员检测算法[J/OL]. 计算机工程: 1–10. [2025–08–19]. <https://doi.org/10.19678/j.issn.1000-3428.0070383>.
- Zhao Tao, Dong Lihong, Qin Yi. Subsurface personnel detection algorithm combining attention and lightweight network[J/OL]. Computer Engineering: 1–10. [2025–08–19]. <https://doi.org/10.19678/j.issn.1000-3428.0070383>.
- [12] Yue Min, Zhang Liqiang, Huang Juan, et al. Lightweight and efficient tiny-object detection based on improved YOLOv8n for UAV aerial images[J]. Drones, 2024, 8(7): DOI:10.3390/drones8070276.



作者简介: 荣海 (1988—),男,辽宁沈阳人,博士,副教授,主要从事矿山安全与矿井动力灾害防治等方面的研究。E-mail: ronghai1988@163.com。