

融合注意力和轻量级网络的井下人员检测算法

赵涛, 董立红, 秦映

(西安科技大学计算机科学与技术学院, 陕西 西安 710054)

摘要: 在煤矿智能化建设背景下, 煤矿井下人员实时监测对于保障矿井安全具有重要意义。然而, 目前常见检测模型在井下部署时由于参数过大, 难以满足实时监测的需求, 同时由于煤矿井下环境复杂, 人员检测通常存在漏检、误检等情况。为此, 在 YOLOv8 的基础上, 提出一种融合注意力和轻量级网络的井下人员检测算法。首先, 针对模型参数过大导致难以部署的问题, 采用重构轻量化模块 C2f_RepGhost 替换原模型主干网络的 C2f 模块; 其次, 为了提高模型检测准确率, 在主干网络中添加 EMA 注意力机制; 然后, 为了增强模型的小目标人员检测能力, 引入 DyHead 动态检测头; 最后, 使用 Inner-CIoU 替换原损失函数, 优化目标定位精度。在 PASCAL VOC 2012 数据集和自建煤矿井下数据集上进行对比实验, 结果表明: 在 PASCAL VOC 2012 数据集上, 改进模型相较于原模型精确率提升了 1.3 个百分点, 召回率提升了 1.2 个百分点; 在自建数据集上, 相较于原模型, 改进模型参数减少了 29.6%, 精确率和召回率分别提升了 2.4 和 3.5 个百分点, 分别达到了 95.3% 和 90%。改进模型在减少参数量的同时, 改善了原模型漏检、误检的情况, 能够满足煤矿井下人员检测的实际要求。

关键词: 井下人员检测; YOLOv8; 轻量级网络; 注意力机制; 检测头

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0070383

Underground Personnel Detection Algorithm Combining Attention and Lightweight Network

ZHAO Tao, DONG Lihong, QIN Yi

(College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, Shaanxi, China)

【Abstract】 In intelligent coal mine construction, the real-time monitoring of underground personnel in coal mines is important to ensure mine safety. However, several parameters are used for real-time monitoring, and commonly used detection models often struggle to meet these requirements when deployed underground. Moreover, given the complex environment of underground coal mines, issues such as missed or false detections frequently occur during personnel detection. To address these issues, this paper proposes a coal mine underground personnel detection algorithm based on YOLOv8 that integrates attention mechanisms and lightweight networks. First, to address deployment difficulties caused by excessive model parameters, the C2f module in the backbone network of the original model is replaced with a reconstructed lightweight module called C2f_RepGhost. Second, to enhance the detection accuracy of the model, an EMA attention mechanism is incorporated into the backbone network. Third, to strengthen the capability of the model to detect small-target personnel, the DyHead dynamic detection head is introduced. Finally, the original loss function is replaced with Inner-CIoU to optimize the target localization accuracy. Comparative experiments on the PASCAL VOC 2012 dataset and a self-built underground coal mine dataset are conducted. The results demonstrate that, compared to the original model, the improved model achieves a 1.3 percentage point increase in precision and a 1.2 percentage point increase in recall on the PASCAL VOC 2012 dataset. On the self-built dataset, the improved model reduces the parameters by 29.6% compared to the original model, and the precision and recall reach 95.3% and 90%, respectively, which are 2.4 and 3.5 percentage points, higher than those of the original model, respectively. The improved model not only reduces the number of parameters but also mitigates the issues of missed and false detections in the original model, meeting the practical requirements for personnel detection in underground coal mines.

【Key words】 underground personnel detection; YOLOv8; lightweight network; attention mechanism; detecting head

0 引言

我国煤矿地质条件复杂, 煤矿安全事故时有发生

发生, 因此, 保障煤矿井下工作人员的生命安全极为重要^[1]。近年来, 深度学习技术在煤矿安全领域的应用极大推动了煤矿智能化的进程^[2]。其

基金项目: 国家自然科学基金青年项目(62303375)。

作者简介: 赵涛, 男, 硕士研究生, 主研方向为深度学习、目标检测; 董立红, 教授; 秦映, 副教授。

收稿日期: 2024-09-18

修回日期: 2025-01-23

E-mail: 262771140@qq.com

中,井下人员检测是维护煤矿安全的重要举措,及时准确地检测井下人员,可以有效监测人员是否误入危险区域并及时预警联动控制,对于维护煤矿安全具有重要意义^[3]。然而,煤矿井下边缘设备算力有限,将现有人员检测模型部署到井下边缘设备时,存在参数过大的问题,因此,需要对现有模型进行轻量化处理。同时,煤矿井下环境复杂多变,导致人员目标特征提取困难,给煤矿井下人员检测带来诸多挑战。

目前,目标检测算法可以分为传统的目标检测算法以及基于深度学习的目标检测算法^[4]。传统目标检测算法在早期目标检测任务中发挥了重要作用,但随着场景复杂程度的提升,传统目标检测算法难以满足实际需求。而深度学习算法由于其强大的性能,逐渐取代传统的目标检测算法。文献[5]提出一种轻量级的 RT-DETR 目标检测算法,以轻量级模型 EMO 作为特征提取主干,并在 LEHE 中引入定位分类器,最后改进损失函数,在实现轻量化的同时提高了检测精度。文献[6]提出了一种以 SSD 为基础模型、MobileNetV2 为骨干网络的实时检测算法。为了提高特征融合能力,该算法在特征融合层添加 EFF 单元,最终实现了实时检测。文献[7]采用改进的 YOLOv5 进行井下人员检测,并采用 ShuffleNetV2 作为主干网络设计轻量化模型,在增强模型检测精度的同时满足了实时检测的要求。文献[8]提出一种改进的多尺度卷积神经网络(CNN)算法,以 YOLOv5 为基础网络,使用 PP-LCNet 替换主干网络,并采用 shuffle_block 替代 C3,最后改进损失函数,该算法实现了轻量化。文献[9]在 YOLOv5s 中引入 GhostConv 和 DWConv 模块,并在特征图中添加 ECA 注意力机制,提升了模型的整体检测精度。文献[10]提出通道重组和特征关注以改进 YOLOv7 网络,提高其在煤矿井下复杂场景中的性能,并在颈部网络添加 ACmix 模块,增强了模型的小目标检测能力。实验结果表明,上述方法往往专注于轻量化或检测性能的单方面提升,导致模型轻量化后检测效果不佳,而在增强检测性能时又会增加模型参数量,难以有效解决漏检、误检问题。

针对上述问题,本文提出一种融合注意力机制和轻量级网络的井下人员检测算法。首先,通过重构 C2f_RepGhost^[11]轻量化网络,减少特征图的冗余,降低模型参数,实现模型的轻量化;其次,添加 EMA^[12]注意力机制,增强模型的特征提取能力,提高模型的检测精度;然后,引入 DyHead 检测

头^[13],增强模型对小目标的关注和特征提取能力,提高模型的小目标检测能力;最后,采用 InnerCIoU^[14]作为损失函数,优化目标框的定位和回归性能,加快模型收敛速度,提升模型对目标的定位精度。

1 YOLOv8 模型

YOLOv8^[15]是基于 YOLO 系列^[16-18]算法改进的一阶段目标检测算法。相较于 YOLOv5^[19]和 YOLOv7^[20],YOLOv8 的检测精度和速度都有所提升,其主要由主干网络(Backbone)、颈部结构(Neck)和头部预测层(Head)3个组件组成。为了缓解分类和回归2个任务之间的内在冲突,YOLOv8 采用分类头和检测头分离的策略,从而提高模型的性能,YOLOv8 整体结构如图1所示。

YOLOv8 使用 CSPDarknet^[21]作为主干网络,其由卷积层和池化层构成,用于图像的特征提取。该网络采用 C2f 模块替代原主干网络的 CSP 模块,并通过梯度分流连接,丰富特征提取过程中的信息流。CSPDarknet 不仅提高了多尺度目标处理能力,还降低了计算复杂度,优化了特征表示。

YOLOv8 的颈部结构用于融合和传递主干网络提取的特征,以生成多层次的特征表示。它利用先进的特征融合技术整合来自不同层次的特征图,从而提升了对复杂场景中目标的检测能力。改进后的颈部结构能够更有效地处理目标尺度和背景变化。

在检测头方面,YOLOv8 将分类头和检测头分离,减少两者之间的冲突,并摒弃了传统的基于锚框思想,采用无锚框策略。通过引入 TAL(Task Alignment Learning)动态分配策略^[22],优化了模型分类和定位性能,使模型表现更加准确和稳定。

此外,YOLOv8 在检测头部分采用了无锚框匹配机制,无须预设锚点^[23],仅对不同尺度特征图上的目标中心点和宽高进行回归,显著减少了模型的计算量。通过利用不同尺度特征图上的丰富信息,模型能够精确检测并分类各尺度目标。

YOLOv8 模型由于其强大的检测能力,在多个领域得到广泛应用,如智慧交通^[24]、医学影像^[25]等。然而,在煤矿井下场景中,由于目标与背景相似、光照不均等原因,导致图像质量较差,目标特征提取困难,采用原模型进行检测容易产生漏检和误检。因此,在将 YOLOv8 模型应用到煤矿井下场景时,需要针对实际场景的具体问题重新构建模型。

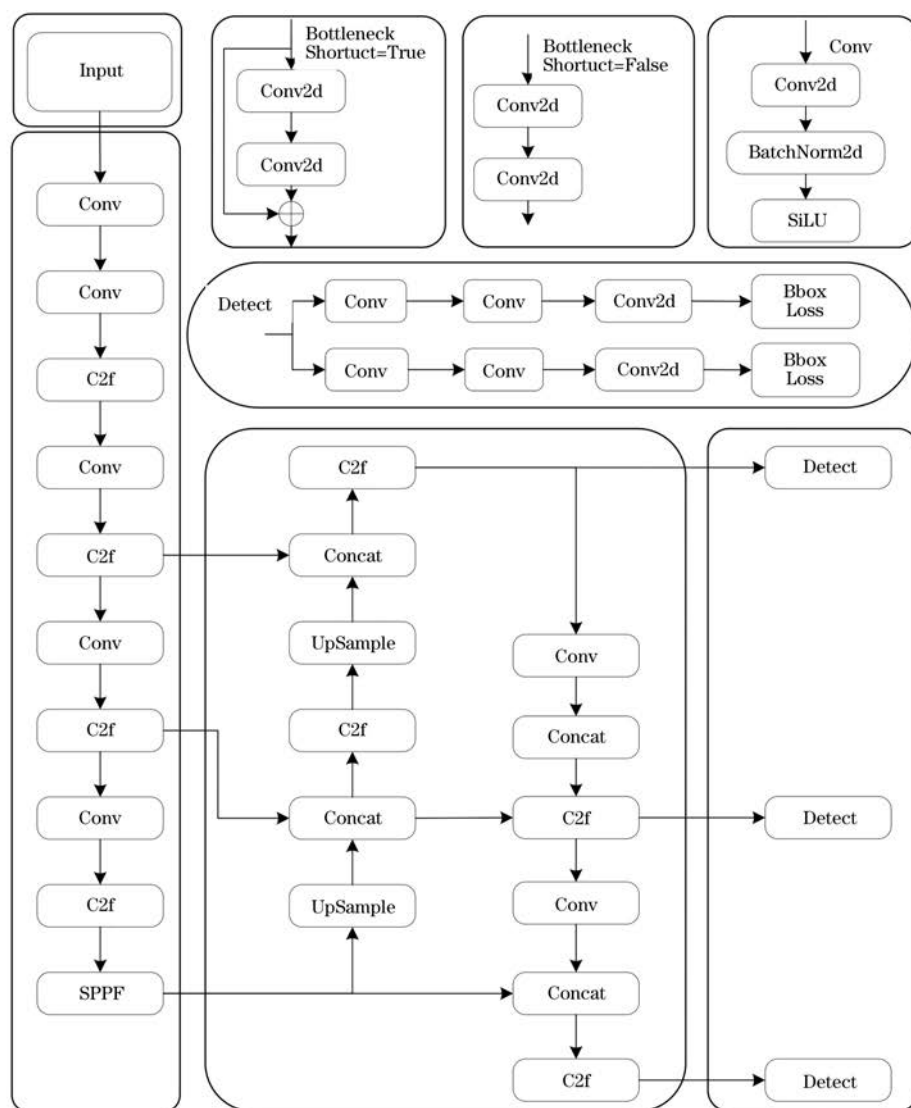


图 1 YOLOv8 网络结构

Fig.1 The structure of the YOLOv8 network

2 基于 YOLOv8 的改进策略

为满足井下边缘设备部署的需求,并针对当前煤矿井下人员检测存在的漏检、误检等问题,本文提出了一种融合注意力和轻量化网络的井下人员检测模型,其结构如图 2 所示。首先,在主干网络部分采用 C2f_RepGhost 轻量化网络替换 C2f 模块,从而降低模型的参数量,实现模型的轻量化。同时,在骨干网络中引入 EMA 注意力,平滑处理模型参数,增强模型的目标特征提取能力;其次,采用 DyHead 检测头替换原有检测头,通过注意力机制统一模型各项感知任务,增强模型的小目标检测能力;最后,引入 Inner-CIoU 损失函数,提升模型对边界框的定位精度并降低回归误差,加快模型收敛速度。

2.1 轻量化模块重构

在煤矿井下人员检测任务中,需要同时兼顾

检测性能和检测速度。由于井下边缘设备算力有限且需要实时响应,因此对算法的参数数量有严格要求。YOLOv8 原模型在井下部署时参数量相对较大,难以满足实际需求。通过引入轻量化模块降低模型参数量,可以有效平衡模型的检测性能和检测速度,从而实现模型在井下复杂环境中的应用。

在原模型中,随着模块的不断堆叠,会生成大量相似的特征图,导致计算速率下降。针对该问题,文献[26]提出了 GhostNet 网络,然而 GhostNet 将特征图拆分为多个子图时会降低模型对小目标的特征提取能力,影响整体检测效果。RepGhost^[11]在 GhostNet 的基础上进行改进,其模型结构如图 3 所示。虽然 RepGhost 在降低参数量的同时改善了模型的鲁棒性,但在处理煤矿复杂场景时检测效果不理想,需对其进一步改进。

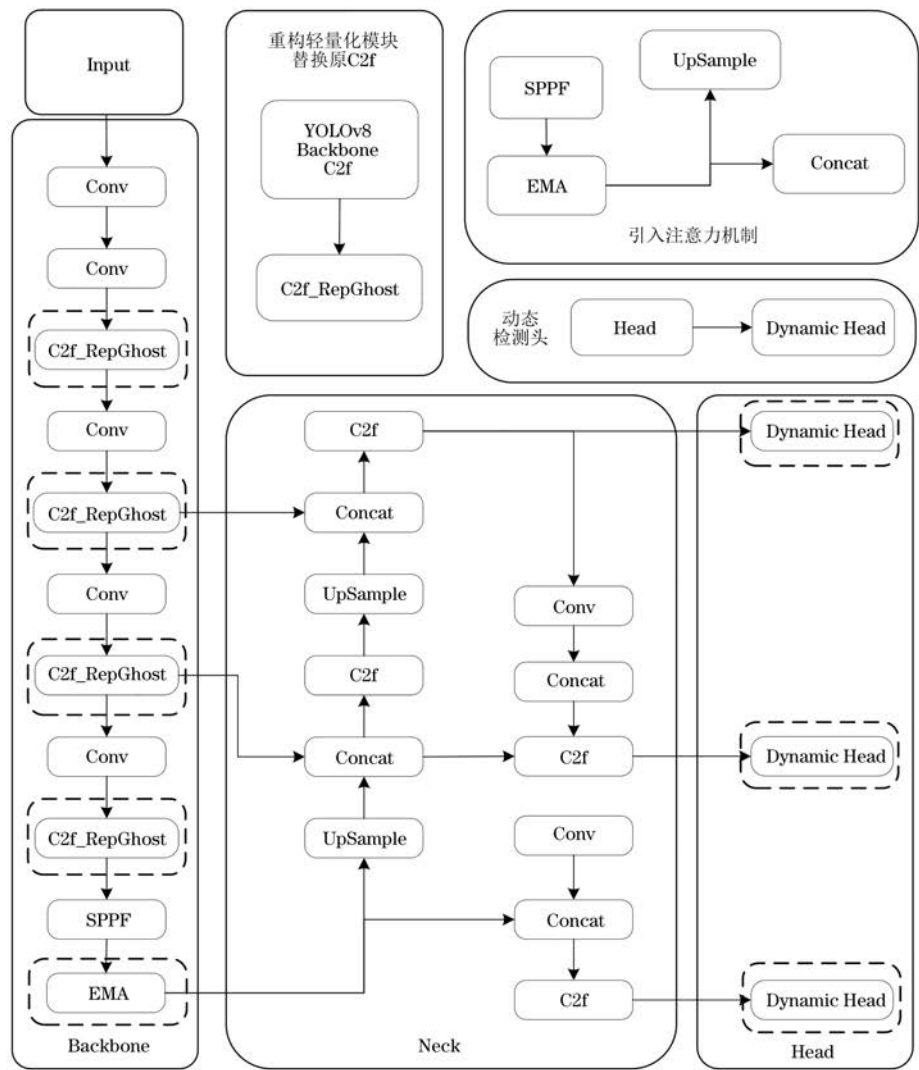


图 2 改进的 YOLOv8 网络结构

Fig.2 The structure of the improved YOLOv8 network

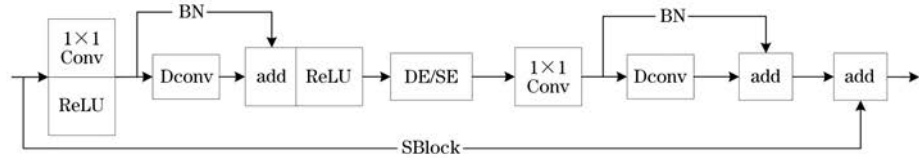


图 3 RepGhost 模块结构

Fig.3 RepGhost module structure

C2f 是 YOLOv8 中的一个模块,主要用于特征转换、分支处理和特征融合。该模块首先对输入进行降维,然后采用瓶颈结构进行连续卷积,最后升维至最初通道数。通过这种设计,实现了浅层特征和

深层特征之间的融合,提升了模型的准确率。然而,由于 C2f 采用瓶颈操作,导致计算量较大,难以满足煤矿井下实际需求。为解决这一问题,需要对其进行改进,C2f 结构如图 4 所示。

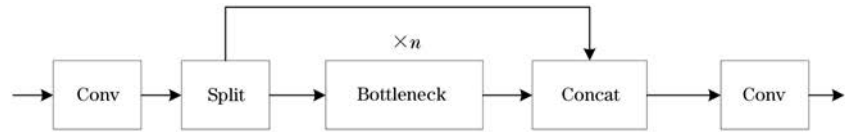


图 4 C2f 模块结构

Fig.4 C2f module structure

通过将 C2f 的 Bottleneck 模块替换为 RepGhost 模块,重构新的 C2f_RepGhost 轻量化

模块,其改进后的结构如图 5 所示。改进后的模块通过结合 C2f 和 RepGhost,保留 C2f 的输入降

维操作,并采用 RepGhost 的特征重用技术替换传统的瓶颈卷积,从而减少模型的计算量。此外,改进后的模块在进行特征融合时,通过

RepGhost 的重参数化技术将特征图升维至原始通道数,增强特征融合效果,可以有效提升模型的检测精度。

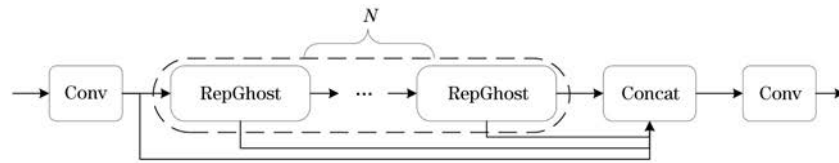


图 5 C2f_RepGhost 模块结构

Fig. 5 C2f_RepGhost module structure

综上可知,重构的轻量化模块不仅可以有效降低模型的参数量,提升模型的检测速度,还能通过结合特征重用与重参数化技术,保持较好的检测性能,实现检测准确率和检测速度之间的平衡。

2.2 融合通道重构与注意力的复杂场景检测

YOLOv8n 在进行特征提取时,虽然采用了深度卷积网络和跨层特征融合等技术,但在处理复杂环境时,特别是煤矿井下复杂环境,目标与背景具有相似的纹理、颜色和形态,模型容易将背景信息错误识别为目标,从而导致误检和漏检的情况。

为解决这一问题,本文在 YOLOv8n 基准模型的骨干网络中添加多尺度注意力 EMA 机制^[12],以提升模型在井下复杂环境中的适应能力,减少漏检和误检现象。EMA 通过动态调整不同维度特征的权重,增强模型对目标特征的关注,尤其是在煤矿井下复杂环境下,可以有效避免背景信息的干扰。此外,通过多尺度特征的融合,使得模型可以在不同尺度下提取目标特征,从而增强模型在煤矿井下复杂场景中的检测精度,减少漏检和误检的现象,更好地应对煤矿井下复杂环境中的人员检测挑战,EMA 结构如图 6 所示。

EMA 采用 3 条并发路径提取不同注意力权重特征,利用这些特征对特征图进行分组。在 2 条 1×1 的分支中采用 2 个 1D 全局平均池化操作,沿 2 个方向对通道进行编码,并将单个 3×3 内核堆叠到 3×3 分支中,以捕获多尺度特征。然后,采用 2D 全局平均池化对 1×1 分支输出的全局空间信息进行编码,从而获取全局空间信息。池化操作的公式为:

$$Z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i, j) \quad (1)$$

式中: Z 是输入的特征图,维度为 $H \times W \times C$, H 和 W 分别是特征图的高度和宽度, C 是通道数。

EMA 首先对部分通道维度进行重构,将其划分为多个子特征,以此减少通道间的冗余信息;然后通过多尺度注意力机制捕获不同尺度的特征信息,

并通过跨空间学习的方法融合并行子网络的输出映射,提升模型的检测性能。

通过在 YOLOv8n 中引入 EMA 注意力,使得模型能够在复杂背景干扰情况下准确提取目标人员特征,降低背景等冗余信息的干扰,从而增强模型在煤矿井下复杂环境中的人员检测能力,降低漏检和误检现象,进一步提升模型的检测准确率和鲁棒性,更好地满足煤矿井下人员安全监测的实际需求。

2.3 基于动态路由与融合注意力的检测头

YOLOv8 检测头采用基于特征金字塔的多尺度融合策略。在融合过程中,由于复杂背景、光照不均等环境干扰,使得小目标人员特征容易被忽视,从而导致小目标人员漏检情况的产生。

针对上述问题,本文引入了 DyHead(Dynamic Head)动态检测头^[13],该检测头通过动态路由机制,能够根据人员目标特征动态调整特征层权重,使模型在井下环境中能更准确地捕获小目标人员区域。同时,得益于动态感受野设计,该检测头还能自适应地调整卷积核大小,使得该检测头可以根据人员目标尺寸调整卷积核大小,更精确地关注和强化小目标人员特征,增强模型对小目标人员的检测能力。

DyHead 的核心部分是动态注意力机制,该机制通过在不同维度上自适应调整注意力来优化输入特征,从而提升特征表示的准确性。动态注意力机制主要包括尺度感知注意力、空间感知注意力和任务感知注意力。通过统一上述 3 种注意力,可以实现对不同特征层的动态调整,从而提高模型对小目标的检测性能。注意力计算公式如下所示:

$$W(F) = \pi(F) \cdot F = \pi_C(\pi_S(\pi_L(F) \cdot F) \cdot F) \cdot F \quad (2)$$

式中: F 为给定的特征张量, $F \in \mathbb{R}^{L \times S \times C}$; $\pi(\cdot)$ 是注意力函数; π_L 、 π_S 、 π_C 分别是应用于维度 L 、 S 、 C 的尺度感知注意力、空间感知注意力和任务感知注意力。3 种注意力的详细结构如图 7 所示。

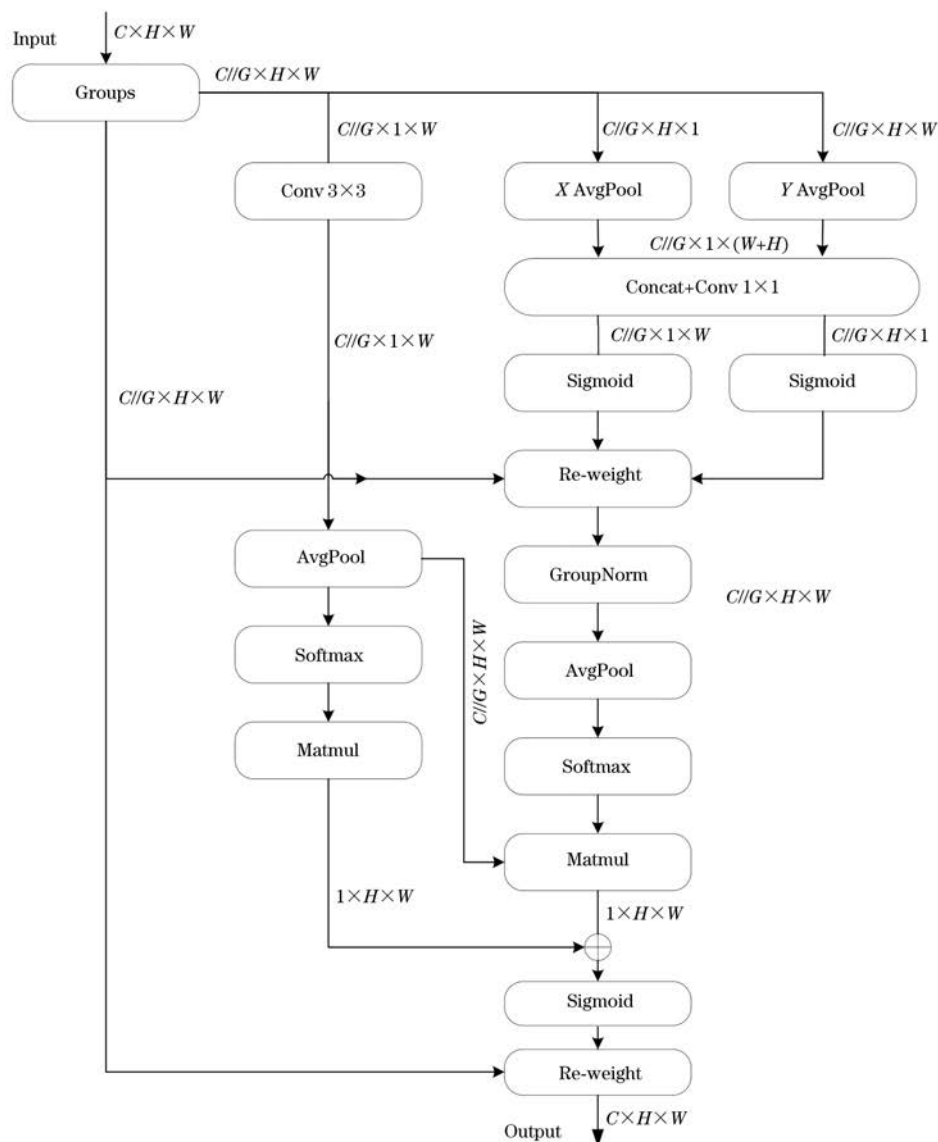


图 6 EMA 模型结构

Fig.6 EMA model structure

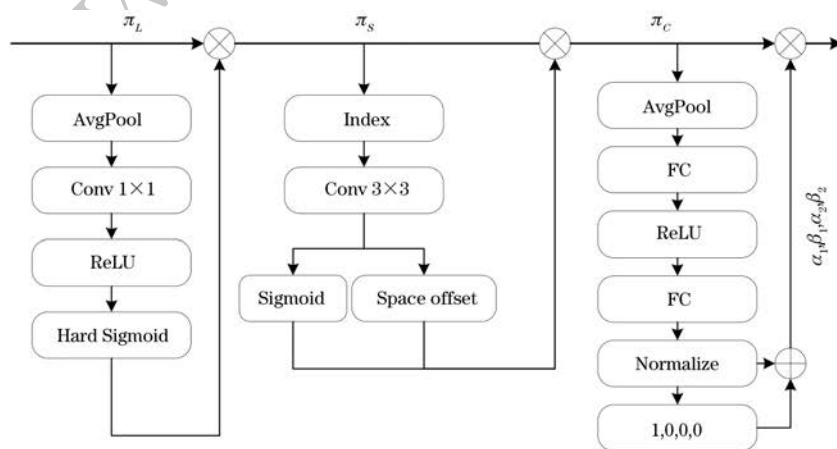


图 7 Dynamic Head 结构

Fig.7 Dynamic Head structure

通过采用 DyHead 动态检测头对 YOLOv8n 进行改进,使模型能够动态关注不同目标的特征信息,

从而提升对小目标的特征提取效率,增强模型对小目标人员的检测能力,降低模型检测的漏检率。

2.4 损失函数改进

YOLOv8 结合 DFL 和 CIoU 来计算边界框回归,其计算公式如下:

$$L_{\text{CIoU}} = 1 - I_{\text{IoU}} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \quad (3)$$

式中: I_{IoU} 表示预测框与真实框的交并比; b, b^{gt} 分别表示预测框和真实框的中心点; $\rho(b, b^{\text{gt}})$ 表示真实框与预测框之间的欧氏距离; c 表示预测框与真实框最小外接矩形的对角线长度; α 表示权重系数; v 表示预测框与真实框之间的长宽比。

然而,采用 YOLOv8 原损失函数处理井下复杂场景人员检测任务时,具有一定的局限性。CIoU 难以平衡难易样本之间的权重,特别是在处理背景复杂的场景时,无法准确评估预测框与真实框的重叠程度,从而导致漏检和误检现象频繁发生。为解决 CIoU 的局限性,本文在 CIoU 的基础上引入 Inner-IoU 损失函数^[14],其优势在于通过引入尺度因子 ratio,动态调整辅助边框的大小,可以更加精准地捕获小目标特征,从而提高模型在复杂场景中的检测能力和实用性。将 Inner-IoU 应用到 CIoU 中,计算方式如下所示:

$$b_l^{\text{gt}} = x_c^{\text{gt}} - \frac{w^{\text{gt}} \cdot r_{\text{ratio}}}{2}, b_r^{\text{gt}} = x_c^{\text{gt}} + \frac{w^{\text{gt}} \cdot r_{\text{ratio}}}{2} \quad (4)$$

$$b_t^{\text{gt}} = y_c^{\text{gt}} - \frac{h^{\text{gt}} \cdot r_{\text{ratio}}}{2}, b_b^{\text{gt}} = y_c^{\text{gt}} + \frac{h^{\text{gt}} \cdot r_{\text{ratio}}}{2} \quad (5)$$

$$b_l = x_c - \frac{w \cdot r_{\text{ratio}}}{2}, b_r = x_c + \frac{w \cdot r_{\text{ratio}}}{2} \quad (6)$$

$$b_t = y_c - \frac{h \cdot r_{\text{ratio}}}{2}, b_b = y_c + \frac{h \cdot r_{\text{ratio}}}{2} \quad (7)$$

$$i_{\text{inter}} = \min(b_r^{\text{gt}}, b_r) - \max(b_l^{\text{gt}}, b_l) \cdot \min(b_b^{\text{gt}}, b_b) - \max(b_t^{\text{gt}}, b_t) \quad (8)$$

$$u_{\text{union}} = (w^{\text{gt}} \cdot h^{\text{gt}}) \cdot r_{\text{ratio}}^2 + (w \cdot h) \cdot r_{\text{ratio}}^2 - i_{\text{inter}} \quad (9)$$

$$I_{\text{IoU}}^{\text{inner}} = \frac{i_{\text{inter}}}{u_{\text{union}}} \quad (10)$$

$$L_{\text{Inner-CIoU}} = L_{\text{CIoU}} + I_{\text{IoU}} - I_{\text{IoU}}^{\text{inner}} \quad (11)$$

式中: x_c^{gt} 和 y_c^{gt} 为真实框的中心点; x_c 和 y_c 为预测框的中心点; $b_l^{\text{gt}}, b_r^{\text{gt}}, b_t^{\text{gt}}, b_b^{\text{gt}}$ 为真实框的左、右、上、下边界; b_l, b_r, b_t, b_b 为预测框的左、右、上、下边界; i_{inter} 表示真实框与预测框的交集部分; u_{union} 表示真实框与预测框的并集部分。

Inner-CIoU 通过融合 CIoU 和 Inner-IoU,并结合尺度因子 ratio 动态调整辅助边框大小,优化了定位损失的计算方式,可以有效解决煤矿井下人员检测中的困难样本检测问题。

3 实验验证

3.1 实验准备

3.1.1 实验环境

在本文实验中,模型训练配置、数据集以及训练参数设置均保持一致。实验环境配置如表 1 所示。其中,在模型训练过程中,批次大小为 8,初始学习率为 0.001,动量为 0.937,迭代次数为 100。

表 1 实验参数配置

配置环境	版本型号
操作系统	Windows 11
计算机架构	CUDA 12.0
内存/GB	16
语言	Python 3.9
CPU	12th Gen Intel® Core™ i7-12700H@2.30 GHz
处理器型号	NVIDIA GeForce RTX 3060

3.1.2 数据集

为验证本文模型的可行性和有效性,采用 PASCAL VOC 2012 公开数据集^[27]和自建煤矿井下人员检测数据集进行实验。

PASCAL VOC 2012 是一个经典且具有挑战性的目标检测数据集,该数据集包含多个场景,共有 17 215 张图片。数据集具有多种类别的目标,从中清洗出标注有人员类别的数据,共计 7 116 张图片,作为本文实验数据集。其中,所有图片都经过精确标注,确保了其标注数据的准确性和一致性。

自建煤矿井下人员检测数据集的数据来源为某煤矿井下视频监控系统。对采集到的视频,每 15 帧截取一帧图像,最终得到 46 262 张原始图片。由于煤矿井下存在大量人工光源,导致摄像头经常捕获到高曝光图片,这些图片严重失真且质量差,无法提供有效信息。为此,对原始图像进行筛选和清洗。具体地,剔除所有空白以及高曝光图片,经过筛选,最终得到共计 11 734 张图片,这些图片涵盖了多个井下典型场景,包括皮带机头、采煤工作面、井下运输车场、井下巷道、井下乘车点、井下等候室等,如图 8 所示,保证了自建数据集的多样性。

3.1.3 评价指标

在本文实验中,采用精确率(Precision, P)、召回率(Recall, R)、mAP@0.5($R_{\text{mAP}@0.5}$)来评估模型的性能,计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (12)$$

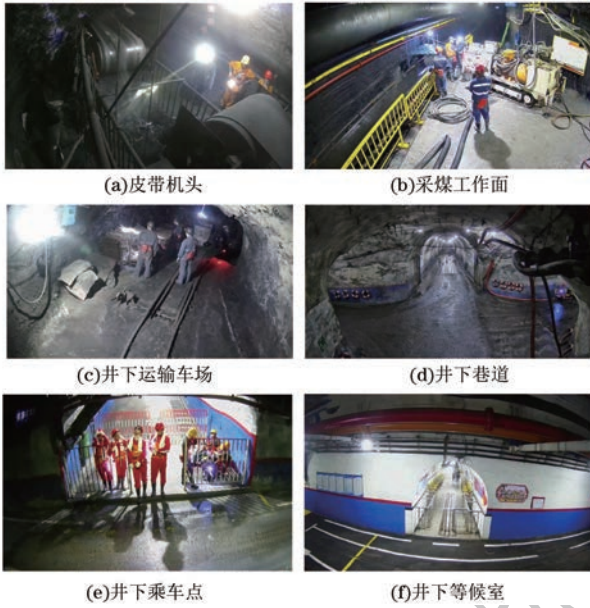


图 8 井下不同场景示例

Fig. 8 Examples of different underground scenarios

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (13)$$

表 2 消融实验结果

Table 2 Ablation experiment results

实验组别	C2f_RepGhost	EMA	DyHead	Inner-CIoU	Precision	Recall	mAP@0.5	GFLOPs	Param/10 ⁶
Baseline					0.929	0.865	0.937	8.1	3.00
1	✓				0.924	0.859	0.928	6.9	2.52
2		✓			0.947	0.878	0.943	8.3	3.02
3			✓		0.942	0.886	0.949	7.8	2.51
4				✓	0.933	0.870	0.939	8.1	3.00
5	✓	✓			0.941	0.876	0.947	7.0	2.60
6	✓		✓		0.939	0.887	0.955	6.7	2.09
7	✓	✓	✓		0.952	0.893	0.954	6.8	2.11
8	✓	✓	✓	✓	0.953	0.900	0.957	6.8	2.11

由表 2 可知:在实验 1 中,在基准模型中引入 C2f_RepGhost 后,模型参数量减少了 16.0%,模型精确率和召回率分别下降了 0.5 和 0.6 百分点,说明该模块有效降低了模型的参数量,实现了检测准确率和检测速度之间的平衡;在实验 2 中,添加 EMA 注意力后,模型精确率提升了 1.8 百分点,召回率提升了 1.3 百分点,模型参数量增加了 0.7%,有效降低了漏检和误检的情况;在实验 3 中,采用 DyHead 替换基准模型检测头后,模型精确率提升了 1.3 百分点,召回率提升了 2.1 百分点,参数量减少了 16.3%,召回率的提升说明 DyHead 有效缓解了原模型的小目标人员漏检情况;在实验 4 中,采用 Inner-CIoU 损失函数后,模型精确率和召回率分别提升了 0.4 和 0.5 百分点;在实验 5 中,同时引入 C2f_RepGhost 和 EMA 注意力,

$$R_{mAP} = \frac{1}{N} \sum_{n=0}^N \int_0^1 P_n(R) dR \quad (14)$$

式中: T_p 表示模型正确检测的人员数量; F_p 表示模型错误检测的人员数量; F_n 表示模型未检测的人员数量; N 代表类别数; P_n 代表某一类的平均精度 (AP); $mAP@0.5$ 表示预测框与真实框面积的交并比大于 0.5 时的平均精度; mAP 是评价模型全局检测性能的指标。

同时,本文还使用浮点运算量 (GFLOPs) 和模型参数量 (Param) 来评价模型的计算效率。模型参数量和浮点运算量是衡量深度学习算法计算效率的重要指标。模型参数量直观地反映了模型的网络复杂度,而浮点计算量则是指运行一次网络模型需要进行浮点运算的次数。

3.2 消融实验

为了评估本文引入的各模块对实验效果的影响及其有效性,本文以 YOLOv8n 为基准模型,逐步添加各模块,其中,“✓”表示在该组实验中引入了该模块,实验结果如表 2 所示,结果的最优项用加粗表示。

模型精确率提升了 1.2 百分点,召回率提升了 1.1 百分点,两者相结合,在降低参数的情况下提升了模型的性能;在实验 6 中,同时引入 C2f_RepGhost 和 DyHead,模型精确率提升了 1.0 百分点,召回率提升了 2.2 百分点,参数量降低了 30.3%;在实验 7 中,引入了 C2f_RepGhost、EMA 和 DyHead,模型精确率提升了 2.3 百分点,召回率提升了 2.8 百分点,同时参数量减少了 29.7%,结合使用 C2f_RepGhost、EMA 注意力和 DyHead 检测头,比单独引入上述模块效果更好;在实验 8 中,同时引入 C2f_RepGhost、EMA 注意力和 DyHead 检测头,并改进损失函数,模型精确率提升了 2.4 百分点,召回率提升了 3.5 百分点, $mAP@0.5$ 提升了 2.0 百分点,参数量下降了 29.7%,有效平衡了模型参数量和模型性能。由实验结果可

知,添加的各个模块对模型检测性能均有一定的提升效果,当共同作用时,模型效果最佳。

实验结果证明,本文改进的模型相比于基准模型在自建数据集上具有更好的检测效果,不仅降低了模型的参数量,还提高了模型的检测性能,能够改善原模型漏检、误检的情况,相比于原模型,本文模型更适用于部署到煤矿井下实际场景中。

在井下 5 种不同场景中,对基准模型和本文模型进行定性评价,效果如图 9 所示。在第 1 列图片中,目标存在部分遮挡且由于矿灯导致曝光等情况,基准模型在检测时出现了漏检现象,而本文模型可以准确检测出基准模型漏检的目标;在第 2 列图片中,目标人员尺寸较小,基准模型检测时存在漏检的

情况,本文模型可以将基准模型漏检的小目标人员检测出来,即本文模型增强了对小目标人员的检测能力;在第 3 列图片中,由于人员姿态变化较大,且存在部分遮挡情况,导致基准模型出现了漏检的情况,本文模型并未出现漏检的情况;在第 4 列图片中,由于图片存在曝光以及目标之间相互遮挡,且图片尺寸较小,基准模型仅检测到其中一个目标,导致了漏检的情况,而本文模型可以在这种复杂的情况下检测出人员目标,没有出现漏检、误检的情况;在第 5 列图片中,由于目标和背景相似,且存在一定的遮挡和曝光,基准模型在检测时出现了漏检和误检的情况,本文模型在检测时可以将原模型未检测到的人员目标准确无误地检测出来,且未出现误检的情况。

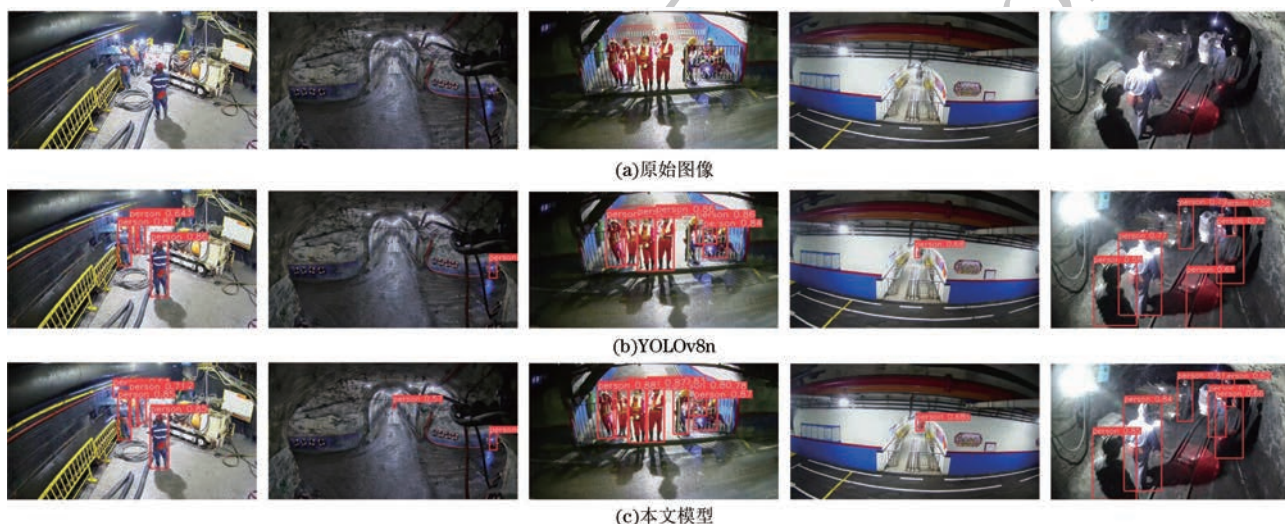


图 9 本文模型和基准模型的实验结果对比

Fig.9 Comparison of experimental results between the model in this paper and the benchmark model

综上所述,本文提出的改进模型在煤矿井下复杂场景中的检测能力优于 YOLOv8n 模型,显著增强了对小目标人员的检测能力。本文模型的实际检测效果比原模型更出色,有效改善了原模型存在的漏检和误检情况,提升了模型的可靠性和准确性,更适合部署到井下复杂场景中。

3.3 对比实验

本文选取 YOLOv8n 作为基准模型,并在自建数据集上对 5 种不同的经典模型 YOLOv5n、YOLOv5s、YOLOv7-tiny^[20]、Faster R-CNN^[28]、YOLOv8n 进行训练,对比分析实验结果,以证明本文基准模型选择的合理性,实验结果如表 3 所示。

由表 3 实验结果可知,在对比的 5 个经典模型中,两阶段检测算法 Faster R-CNN 由于其较大的运算量,难以满足煤矿井下人员检测的实时性要求。相比之下,虽然 YOLOv8n 的参数量和运算量略高于 YOLOv5n,但 YOLOv8n 的精确率和召回率均

表 3 常见检测模型实验结果

Table 3 Experimental results of common detection models

Model	Precision	Recall	GFLOPs	Param/ 10^6
YOLOv5n	0.918	0.857	4.1	1.76
YOLOv5s	0.924	0.861	15.9	7.02
YOLOv7-tiny ^[20]	0.923	0.864	13.2	6.00
Faster R-CNN ^[28]	0.916	0.837	63.9	37.80
YOLOv8n	0.929	0.865	8.1	3.00

优于 YOLOv5n、YOLOv5s、YOLOv7-tiny 等 YOLO 系列的轻量化模型,且 YOLOv8n 的参数量也初步满足井下边缘设备的需求。煤矿井下人员检测模型需要兼顾模型的检测速度和检测性能,综合分析,YOLOv8n 模型在本文数据集上展现的综合性能最优。因此,本文选择 YOLOv8n 作为基准模型,在此基础上进行改进。

为进一步验证本文模型的准确性和可行性,

将本文模型和近三年煤矿井下人员检测模型以及其他领域最新目标检测模型进行对比分析,并在公开数据集 PASCAL VOC 2012 和自建煤矿井下数据集上进行实验验证,实验结果如表 4、表 5 所示(表中模型顺序按照年份排列)。由表 4 可知,在 PASCAL VOC 2012 数据集上,本文模型的检测精确率和召回率分别达到了 92.7%和 89.1%,相比于基准模型分别提升 1.3 和 1.2 百分点。虽然本文模型在 PASCAL VOC 2012 数据集上的检测精确率低于模型 RT-DETR+EMO+LEHE+

DIoU,但本文模型在其他性能上,特别是模型的参数量上,优于模型 RT-DETR+EMO+LEHE+DIoU。综合分析,本文模型的整体性能要优于其他模型,证明本文模型具有较好的泛化性。由表 5 可知,在自建煤矿井下人员检测数据集上,本文模型的检测精确率和召回率相较原模型分别提升了 2.4 和 3.5 百分点,达到了 95.3%和 90.0%,性能优于其他模型,能够兼顾检测性能和检测速度的要求,实现了两者之间的平衡,满足煤矿井下人员检测的实际需求。

表 4 PASCAL VOC 2012 行人数据集上各模型的对比实验结果

Table 4 Comparative experimental results of various models on the PASCAL VOC 2012 pedestrian dataset

Model	Year	Precision	Recall	mAP@0.5	GFLOPs	Param/10 ⁶
YOLOv5s+GhostNet+ECA+DWConv ^[9]	2022	0.908	0.879	0.890	9.4	4.19
YOLOv5s+CBAM+ShuffleNetV2 ^[7]	2023	0.918	0.882	0.886	8.0	3.83
YOLOv5s+PP-LCNet+ShuffleNet ^[8]	2023	0.916	0.884	0.890	8.1	3.82
YOLOv7+ACmix+ShuffleNetV2 ^[10]	2024	0.921	0.880	0.881	8.7	4.55
RT-DETR+EMO+LEHE+DIoU ^[5]	2024	0.929	0.886	0.894	16.4	6.43
EFSSD-MV2 ^[6]	2024	0.890	0.875	0.879	7.1	2.36
YOLOv8n(Baseline)	2023	0.914	0.879	0.887	8.1	3.00
Ours	—	0.927	0.891	0.896	6.8	2.11

表 5 自建煤矿井下数据集上各模型的对比实验结果

Table 5 Comparative experimental results of various models on self-built coal mine underground dataset

Model	Year	Precision	Recall	mAP@0.5	GFLOPs	Param/10 ⁶
YOLOv5s+GhostNet+ECA+DWConv ^[9]	2022	0.931	0.872	0.937	9.4	4.19
YOLOv5s+CBAM+ShuffleNetV2 ^[7]	2023	0.941	0.850	0.927	8.0	3.83
YOLOv5s+PP-LCNet+ShuffleNet ^[8]	2023	0.942	0.862	0.933	8.1	3.82
YOLOv7+ACmix+ShuffleNetV2 ^[10]	2024	0.936	0.865	0.917	8.7	4.55
RT-DETR+EMO+LEHE+DIoU ^[5]	2024	0.946	0.883	0.947	16.4	6.43
EFSSD-MV2 ^[6]	2024	0.932	0.869	0.938	7.1	2.36
YOLOv8n(Baseline)	2023	0.929	0.865	0.937	8.1	3.00
Ours	—	0.953	0.900	0.957	6.8	2.11

为展示本文模型的检测效果,在自建井下数据集上选取 4 组经典场景图片,对各模型的检测效果进行定性分析,结果如图 10 所示。在第 1 列图片中,人员目标较多且有目标被遮挡,对比的 4 组煤矿井下人员检测模型均出现了漏检、误检的情况,难以检测遮挡目标,但是本文模型可以准确检测出所有目标。在第 2 列图片中,存在小目标人员,YOLOv5s+CBAM+ShuffleNetV2 和 YOLOv5s+GhostNet+ECA+DWConv 分别出现了误检和漏检的情况,本文模型成功检测出了小目标人员且未出现误检的情况,充分说明本文模型增强了小目标人员的检测能力。在第 3 列图片中,人员目标之间

存在严重的遮挡现象,而本文模型和 YOLOv5s+GhostNet+ECA+DWConv 模型检测出了被遮挡目标,其他 3 个模型均出现了漏检的情况。在第 4 列图片中,人员特征和背景相似,并且存在遮挡和曝光,本文模型和 YOLOv5s+PP-LCNet+ShuffleNet 模型检测出了所有目标,但本文模型定位更加精确,而另外 3 组模型均出现了漏检和误检的情况。实验结果证明,本文模型在煤矿井下复杂场景中的检测效果相较于对比模型更好,能够有效缓解实际场景检测时的漏检和误检情况,与当前主流目标检测算法相比,本文模型更适用于井下复杂场景中的人员检测任务。

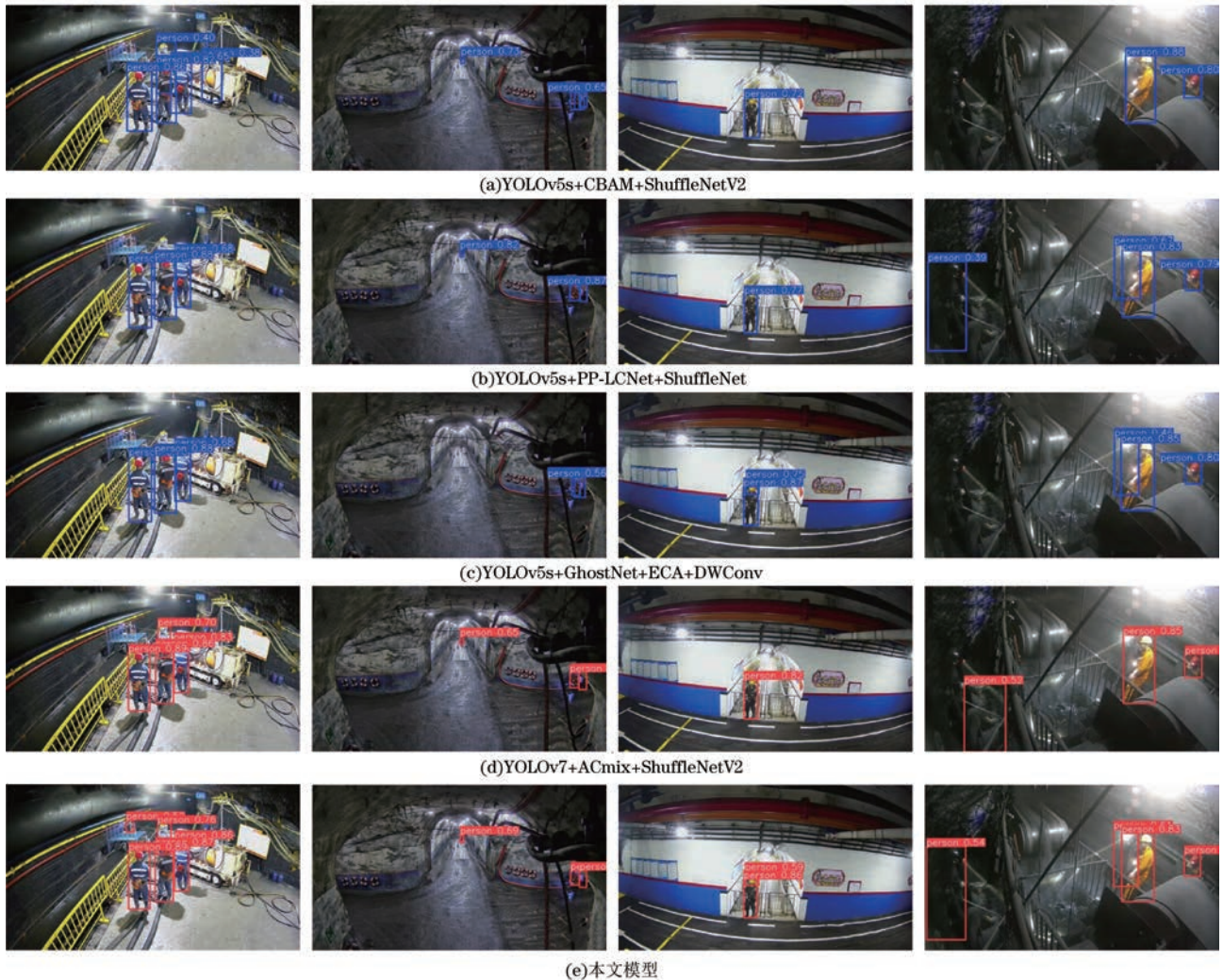


图 10 自建煤矿井下数据集上各模型的实验结果

Fig.10 Experimental results of various models on the self-built coal mine underground dataset

4 结束语

本文自建一个涵盖井下多个典型场景的数据集,针对井下人员检测模型无法兼顾检测速度和检测精度且难以实际应用的问题,对 YOLOv8n 进行改进。首先,采用重构网络 C2f_RepGhost 轻量化主干网络,降低模型参数;然后,将 EMA 注意力机制嵌入主干网络,增强模型的检测准确率;随后,采用 DyHead 改进原检测头,提升模型的小目标检测能力;最后,引入 Inner-CIoU,使目标定位更加精确,加快模型收敛速度。在 PASCAL VOC 2012 数据集和自建煤矿井下人员检测数据集上进行实验,结果表明,本文模型实现了检测性能和检测速度间的平衡,改善了原模型在煤矿井下人员检测中的漏检、误检现象,可以满足煤矿井下人员检测的实际需求。然而,煤矿井下场景中由于存在大型设备以及人员之间的相互遮挡,使得遮挡人员目标检测始终是一项挑战,因此,下一步研究工作将聚焦于遮挡人

员目标检测问题。

参考文献

- [1] ZHANG K X, KANG L, CHEN X X, et al. A review of intelligent unmanned mining current situation and development trend[J]. Energies, 2022, 15(2): 513.
- [2] 王国法. 煤矿智能化最新技术进展与问题探讨[J]. 煤炭科学技术, 2022, 50(1): 1-27.
WANG G F. New technological progress of coal mine intelligence and its problems [J]. Coal Science and Technology, 2022, 50(1): 1-27. (in Chinese)
- [3] 张翼翔, 林松, 李雪. 基于 CenterNet-GhostNet 的选煤厂危险区域人员检测[J]. 工矿自动化, 2022(4): 66-71.
ZHANG Y X, LIN S, LI X. Personnel detection in dangerous area of coal preparation plant based on CenterNet-GhostNet[J]. Journal of Mine Automation, 2022(4): 66-71. (in Chinese)
- [4] 李姝, 李思远, 刘国庆. 基于 YOLOv8 无人机航拍图像的小目标检测算法研究[J]. 小型微型计算机系统, 2024, 45(9): 2165-2174.
LI S, LI S Y, LIU G Q. Research on small target detection algorithm based on YOLOv8 UAV aerial images[J]. Journal of Chinese Computer Systems, 2024, 45(9): 2165-2174. (in Chinese)
- [5] 姜香菊, 王瑞彤, 马彦鸿. 基于轻量级改进 RT-DETR 边缘

- 部署算法的绝缘子缺陷检测[J]. 电工技术学报, 2025, 40(3): 842-854.
- JIANG X J, WANG R T, MA Y H. Insulator defect detection based on lightweight improved RT-DETR edge deployment algorithm [J]. Transactions of China Electrotechnical Society, 2025, 40 (3): 842-854. (in Chinese)
- [6] 林珊玲, 彭雪玲, 王栋, 等. 多尺度增强特征融合的钢表面缺陷目标检测[J]. 光学精密工程, 2024, 32(7): 1075-1086.
- LIN S L, PENG X L, WANG D, et al. Object detection of steel surface defect based on multi-scale enhanced feature fusion[J]. Optics and Precision Engineering, 2024, 32(7): 1075-1086. (in Chinese)
- [7] 芦碧波, 周允, 李小军, 等. 融合注意力机制的 YOLOv5 轻量化煤矿井下人员检测算法[J]. 煤炭技术, 2023, 42(10): 200-203.
- LU B B, ZHOU Y, LI X J, et al. YOLOv5 lightweight coal mine underground personnel detection algorithm base on attention mechanism[J]. Coal Technology, 2023, 42(10): 200-203. (in Chinese)
- [8] 寇发荣, 肖伟, 何海洋, 等. 基于改进 YOLOv5 的煤矿井下目标检测研究[J]. 电子与信息学报, 2023, 45(7): 2642-2649.
- KOU F R, XIAO W, HE H Y, et al. Research on target detection in underground coal mines based on improved YOLOv5 [J]. Journal of Electronics & Information Technology, 2023, 45(7): 2642-2649. (in Chinese)
- [9] 周孟然, 李学松, 朱梓伟, 等. 井下矿工多目标检测与跟踪联合算法[J]. 工矿自动化, 2022, 48(10): 40-47.
- ZHOU M R, LI X S, ZHU Z W, et al. A joint algorithm of multi-target detection and tracking for underground miners[J]. Journal of Mine Automation, 2022, 48(10): 40-47. (in Chinese)
- [10] 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 X Q, LI X, YANG Y D, 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. (in Chinese)
- [11] CHEN C P, GUO Z C, ZENG H E, et al. RepGhost: a hardware-efficient ghost module via re-parameterization[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2211.06088>.
- [12] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning [C] // Proceedings of the 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Washington D.C., USA: IEEE Press, 2023: 1-5.
- [13] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic Head: unifying object detection heads with attentions [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Washington D.C., USA: IEEE Press, 2021: 7369-7378.
- [14] ZHANG H, XU C, ZHANG S J. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2311.02877>.
- [15] HUSSAIN M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection[J]. Machines, 2023, 11(7): 677.
- [16] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/1804.02767>.
- [17] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2004.10934>.
- [18] LI C Y, LI L L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2209.02976>.
- [19] ZHU X K, LÜ S C, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on Transformer prediction head for object detection on drone-captured scenarios[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Washington D.C., USA: IEEE Press, 2021: 2778-2788.
- [20] WANG C Y, BOCHKOVSKIY A, LIAO H 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 (CVPR). Washington D.C., USA: IEEE Press, 2023: 7464-7475.
- [21] WANG C Y, MARK L H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Washington D.C., USA: IEEE Press, 2020: 1571-1580.
- [22] FENG C J, ZHONG Y J, GAO Y, et al. TOOD: task-aligned one-stage object detection [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Washington D.C., USA: IEEE Press, 2022: 3490-3499.
- [23] KONG T, SUN F C, LIU H P, et al. FoveaBox: beyond anchor-based object detection [J]. IEEE Transactions on Image Processing, 2020, 29: 7389-7398.
- [24] 陈梓延, 王晓龙, 何迪, 等. 基于改进 YOLOv8 的轻量化车辆检测网络[J]. 计算机工程, 2025, 51(5): 314-325.
- CHEN Z Y, WANG X L, HE D, et al. Lightweight vehicle detection network based on improved YOLOv8 [J]. Computer Engineering, 2025, 51(5): 314-325. (in Chinese)
- [25] PANDEY S, CHEN K F, DAM E B. Comprehensive multimodal segmentation in medical imaging: combining YOLOv8 with SAM and HQ-SAM models[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Washington D.C., USA: IEEE Press, 2023: 2584-2590.
- [26] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Washington D.C., USA: IEEE Press, 2020: 1577-1586.
- [27] EVERINGHAM M, ALI E S M, VAN G L, et al. The pascal visual object classes challenge: a retrospective[J]. International Journal of Computer Vision, 2015, 111(1): 98-136.
- [28] REN S 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, 2017, 39(6): 1137-1149.

文字编辑 吴云芳

栏目编辑 宋 圆