

文章编号: 1671-7449(2026)02-0133-09

J. Test Measur. Technol., 2026, 40(2): 133-141.

基于深度学习的煤矿井下人员检测算法

郑爽^{1*}, 于海翔¹, 祝永涛²

(1. 黑龙江科技大学 电气与控制工程学院, 黑龙江 哈尔滨 150022;

2. 黑龙江龙煤双鸭山矿业有限责任公司, 黑龙江 双鸭山 155199)

摘要: 人员检测是保障煤矿安全生产和建设智慧矿山的重要内容, 针对煤矿井下环境复杂、对人员检测困难的问题, 提出一种改进的YOLOv8-F煤矿井下人员检测方法。在YOLOv8主干中引入MobileNetV4高效网络提升模型的精确度及高效性, 同时为保证特征的充分提取, 通过添加融合DCNV4的DAttention注意力机制, 保证特征信息的准确性和完整性。最后, 通过引入Shape-IoU边界损失函数提升定位框的精度, 实现对人员的精确检测。实验结果表明, 在特定场景的煤矿工人动作数据集上, 相比基线模型YOLOv8n, YOLOv8-F的准确率、召回率和 mAP 分别提升了3.2、2.7和2.3个百分点, 浮点运算量减少1.4 G。改进后的模型在检测精度和轻量化方面达到很好的平衡, 验证了新算法的有效性。

关键词: 人员检测; YOLOv8算法; MobileNetV4高效网络; 注意力机制; 损失函数

中图分类号: TD67

文献标识码: A

DOI: 10.62756/csjs.1671-7449.2026021

引用格式: 郑爽, 于海翔, 祝永涛. 基于深度学习的煤矿井下人员检测算法[J]. 测试技术学报, 2026, 40(2): 133-141.

ZHENG Shuang, YU Haixiang, ZHU Yongtao. Personnel detection algorithm in coal mine based on deep learning[J]. Journal of Test and Measurement Technology, 2026, 40(2): 133-141.

Personnel Detection Algorithm in Coal Mine Based on Deep Learning

ZHENG Shuang^{1*}, YU Haixiang¹, ZHU Yongtao²

(1. School of Electrical and Control Engineering, Heilongjiang University of Science and Technology, Harbin 150022, China;

2. Heilongjiang Longmei Shuangyashan Mining Co., Ltd., Shuangyashan 155199, China)

Abstract: Personnel detection is an important aspect of ensuring coal mine safety production and constructing intelligent mines. Given the complex underground environment of coal mines and the difficulties in personnel detection, an improved YOLOv8-F personnel detection method for coal mine underground environments is proposed. The MobileNetV4 efficient network is integrated into the YOLOv8 backbone to improve accuracy while reducing computational cost. To ensure thorough feature extraction, a DAttention mechanism fused with DCNV4 is added, improving the accuracy and completeness of feature information. Finally, by introducing a boundary loss function, the precision of the bounding boxes is improved, achieving accurate personnel detection. Experimental results show that, on a specific dataset of coal mine workers' actions, compared with the baseline model YOLOv8n, the YOLOv8-F model's accuracy, recall, and mAP improved by 3.2, 2.7, and 2.3, respectively.

【收稿日期】2025-03-31

【基金项目】黑龙江省省属高等学校基本科研业务资助项目(2024-KYYWF-1102)

*【通信作者】郑爽(1980—), 女, 教授, 硕士, 主要从事矿山智能化研究。E-mail: zhengshuang@usth.edu.cn.

and 2.3 percentage points, respectively, while the floating point operations (GFLOPs) decreased by 1.4 G. The improved model achieves a great balance between detection accuracy and model lightweightness, verifying the effectiveness of the new algorithm.

Key words: personnel detection; YOLOv8; MobileNetV4 efficient network; attention mechanisms; loss function

0 引言

我国作为全球最大的煤炭生产与消费国,煤矿安全生产至关重要。其中,人员安全是保障煤矿生产安全的关键环节。因此,在煤矿开采过程中,对作业人员进行实时检测已成为确保开采安全的重要手段。但由于煤矿井下环境复杂,影响因素较多导致对人员检测困难,常用的检测方法难以满足井下人员检测,因此,提出一种改进的井下人员检测方法十分必要。

目标检测是对井下人员进行检测的重要步骤,目前基于深度学习的目标检测算法已日趋成熟,常用的目标检测算法主要分为单阶段检测模型(如Single Shot MultiBox Detector(SSD), YOLO系列算法)和双阶段检测模型(如Fast Region Convolutional Network(Faster R-CNN)^[1])。在工业领域应用中,双阶段检测算法虽然具有较好的鲁棒性和较高的精度,但对计算资源需求较大且计算速度较慢,难以保证实时性,在井下人员检测任务中并不实用。单阶段检测模型中,SSD检测速度较快但精度较差^[2],而YOLO系列算法因其高效的实时性和准确性备受关注,包括YOLOv3、YOLOv5、YOLOv7、YOLOv8等,能够在单一前向传递中直接预测图像中的目标位置和类别^[3-4],更加适用于井下人员检测任务,本文选用YOLO系列算法中的YOLOv8算法。芦碧波等^[5]提出一种轻量化的YOLOv8检测算法,以轻量化ShuffleNetv2作为主干网络的同时引入改进卷积块注意力模块(Focus-Convolutional Block Attention Module, F-CBAM),增强了对目标物体的关注度,对特征的学习更加准确,提升了目标检测精度。范伟强等^[6]提出在YOLOv7中融入相似性感知激活模块(Similarity-Aware Activation Module, SimAM)注意力机制提升模型对井下复杂环境的适应能力,同时引入深度学习简单在线实时跟踪(Deep Simple Online and Realtime Tracking, DeepSORT)算法,实现了对人员的跟踪。其中,SimAM注意力机制^[7]和DeepSORT算法^[8]在井下人员识别中也被

广泛应用且效果良好。倪云峰等^[9]将YOLOv8算法与光学字符识别技术相结合,通过添加卷积注意力模块和将卷积-批量归一化-激活函数(Cross Stage Binary, CBS)模块改进为控制障碍函数(Control Barrier Function, CBF)模块,增强了对不同尺度特征的提取能力,提高了目标检测的准确率。以上方法在一定程度上都能充分提取目标特征确保准确检测,但对于复杂环境下的人员检测,并不能很好保证检测精度,同时检测速度也会受到影响。

本文提出了一种基于YOLOv8的煤矿井下人员检测方法。在主干网络中将原有的C2f结构与MobileNetV4高效网络结合,同时采用融合DCNV4^[10]的DAttention注意力机制^[11],减少冗余计算的同时确保了特征信息的完整性,采用Shape-IoU^[12]损失函数替换原有的CIoU,增强了定位框的准确性,实现了井下人员的高精度检测。

1 基于YOLOv8的井下人员检测模型

1.1 改进的YOLOv8模型

YOLOv8模型是由Ultralytics在2023年发布的,其网络结构分为4个部分:输入端、骨干网络、颈部网络和预测端。输入端将图片输入主干网络,主干网络通过一系列卷积和反卷积层提取图片特征,同时使用残差连接和瓶颈结构来减少网络的大小并提高性能,颈部网络负责多尺度融合,将主干网络所提取到的不同维度的特征图进行特征融合,增强特征表示能力。预测端负责最终的目标检测和任务分类,包括检测头和分类头,检测头包含一系列卷积层和反卷积层,用于生成检测结果;分类头采用全局平均池化对每个特征图进行分类,通过减少特征图的维度,输出每个类别的概率分布。

本文针对井下环境复杂与人员检测困难的问题,运用MobileNetV4高效网络中UIB块操作与原有C2f进行结合形成新的C2f-M结构。在金字塔池化模块SPPF进行池化操作后添加融合DCNV4的DAttention注意力机制,增强对图片特征的提取能力,并采用Shape-IoU对原有损失函

数进行替换,提升定位框的定位能力。改进后的 YOLOv8模型结构如图 1 所示。

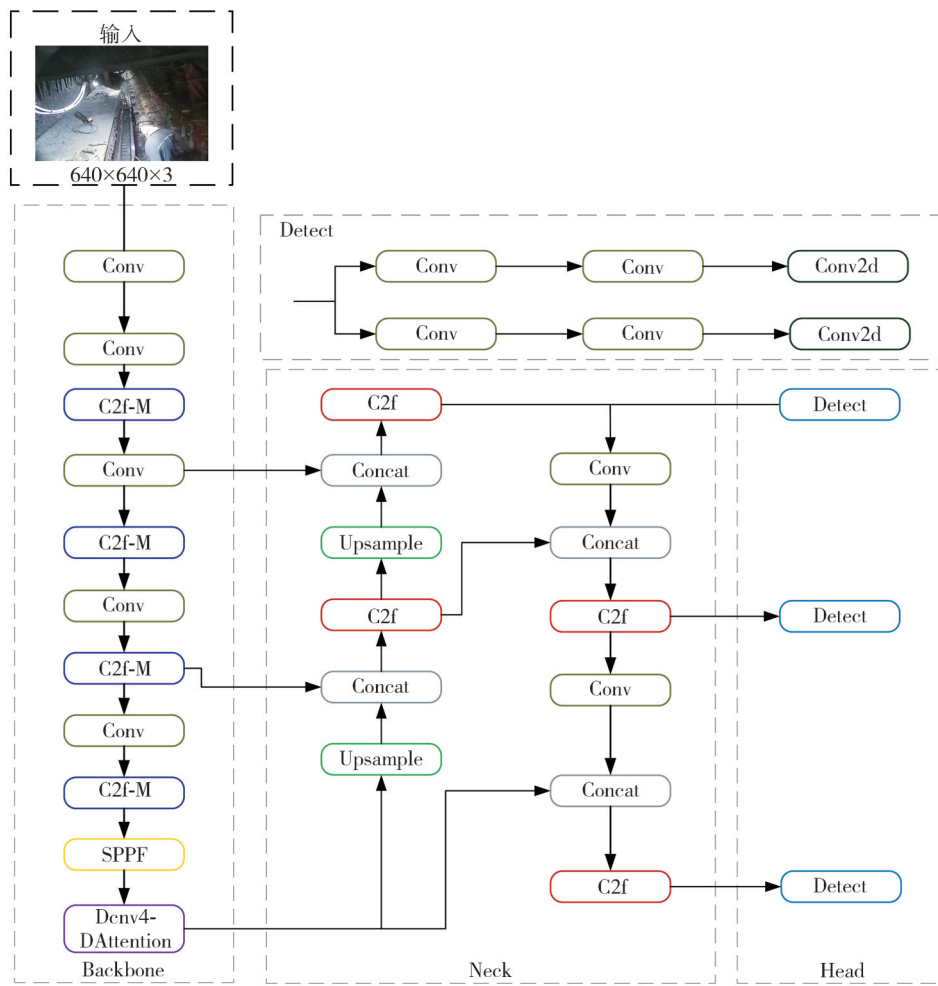


图 1 改进的 YOLOv8 模型

Fig. 1 Improved YOLOv8 model

1.2 MobileNetV4 高效网络

MobileNetV4 是最新一代的 MobileNet 模型,由通用倒置瓶颈块(Universal Inverted Bottleneck, UIB)、移动注意力块(Multi-Query Attention, MQA)和两阶段神经架构搜索(Neural Architecture Search, NAS)组成。其中,UIB 结构结合了倒置瓶颈、ConvNext 前馈神经网络(Feed-Forward Network, FFN)和一种新的 Extra Depth-wise 变体,旨在提升提取特征的灵活性和准确性,其包括 2 个通用 IB 块和 2 个可选择深度卷积模块。UIB 结构图及可能的实例化形式见图 2。

UIB 块中 2 个可选卷积依赖有 4 种可能的实例化,通过不同的卷积操作使 UIB 模块更加高效和准确,同样也使得 Mobile-NetV4 网络更加可靠,提升了模型表达能力,在减少参数计算量的同时有较高性能。

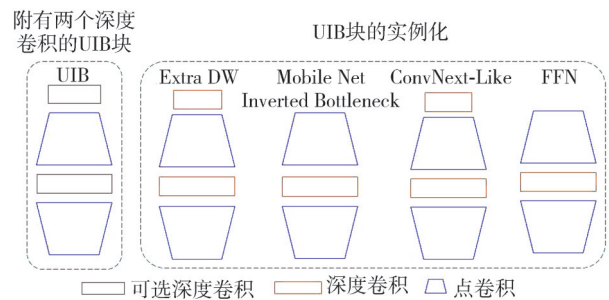


图 2 UIB 块结构图和实例化图

Fig. 2 UIB block and instantiation structure diagram

1.3 DAttention 注意力机制

可变形注意力机制 DAttention 引入了额外的可学习参数,用于计算注意力权重,同时,在确定偏移量值的过程中,为使计算结果更加准确,采用可变形卷积网络 DCNV4 对原计算网络进行替换。在进行模型训练过程中,通过反向传播操作对参数进行学习,使得模型能够调整注意力模式以更好地适应

输入数据。

对于一个给定为 $x \in R^{H \times W \times C}$ 的特征图 (H 和 W 为特征图的高度和宽度, C 为通道数), 生成 $p \in R^{H_G \times W_G \times 2}$ 的网格作为参考点, 其中, 作为参考点的网格需要均匀同一, 其具体大小是根据输入特征图下采样所得到的大小因子 r 、 $H_G = H/r$ 、 $W_G = W/r$ 所确定的。然后需对其进行归一化操作, 将它们归一化到 $[-1, +1]$ 范围内。

每个参考点所对应的偏移量均不相同, 为获取

偏移量, 首先通过查询标记对特征图进行线性投影计算得出结果, 通过一个轻量级的子网络 $\theta_{\text{offset}}(\cdot)$ 计算出偏移量 $\Delta p = \theta_{\text{offset}}(q)$; 为得到更加准确的偏移量值, 保证参考点的位置准确, 采用可变形卷积网络 DCNV4 代替原轻量级网络, 得到新的子网络 $\theta_{\text{offset1}}(q)$, 再运用自适应卷积核来计算更加精准的偏移量, 将注意力区域准确聚焦目标区域。

DCNV4 可变卷积网络详细结构如图 3 所示。

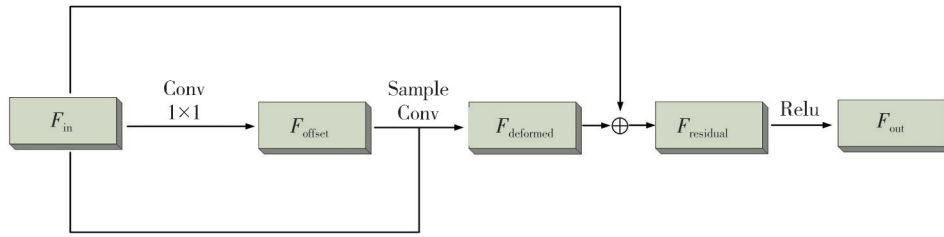


图 3 DCNV4 可变卷积网络结构

Fig. 3 DCNV4 variable convolutional network structure

通过以上计算, 得出偏移量来确定变形点位置, 再进行采样, 将得到的结果作为键和值。其中, 进行投影操作时的投影矩阵为

$$q = xW_q, \tilde{k} = \tilde{x}W_k, \tilde{v} = \tilde{x}W_v, \quad (1)$$

$$\text{with } \Delta p = \theta_{\text{offset1}}(q), \tilde{x} = \phi(x; p + \Delta p), \quad (2)$$

式中: q 为查询向量; $\tilde{k}$ 和 $\tilde{v}$ 分别为经过计算变形后的键和值的嵌入; W_q , W_k 和 W_v 分别为可变投影矩阵; with Δp 为经过新的偏移子网络计算得到的新偏移量; $\tilde{x}$ 为经过计算得到的新特征值。将采样函数 $\phi(\cdot)$ 设置为双线性插值^[13]使其可微, 可得

$$\phi(z; (p_x, p_y)) = \sum_{(r_x, r_y)} g(p_x, r_x) g(p_y, r_y) z(r_x, r_y), \quad (3)$$

式中: $g(a, b) = \max(0, 1 - |a - b|)$ 为计算两个位置的相关性; (r_x, r_y) 为特征图 $z \in R^{H \times W \times C}$ 上将获取所有位置的索引; $z(r_y, r_x)$ 为输入在 (r_y, r_x) 的位置的对应值; 最后, 通过多头注意力来获得相对位置偏移向量 R 。其中, 注意力头的计算输出为

$$z^{(m)} = \sigma \left(\frac{q^{(m)} \tilde{k}^{(m)}}{\sqrt{d}} + \phi(\hat{B}; R) \right) \tilde{v}^{(m)}, \quad (4)$$

式中: $q^{(m)}$ 、 $\tilde{k}^{(m)}$ 和 $\tilde{v}^{(m)}$ 为第 m 个注意力头的查询向量、键向量和值向量; $\sqrt{d}$ 为缩放因子; $\hat{B}$ 为对应位置参考点; σ 为 Softmax 激活函数; $\phi(\hat{B}; R) \in R^{HW \times H_G W_G}$ 。为得到输出结果, 需将所得不同变形点的特征进行连接后, 通过 W_o 进行投影, 投影结果即为输出结果 z 。

DAttention 可变注意力模块信息流如图 4 所示。

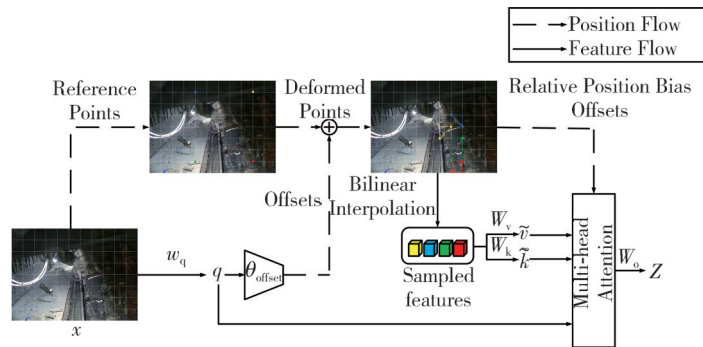


图 4 可变注意力模块信息流

Fig. 4 Variable attention module information flow

由于环境复杂, 影响因素较多会导致井下人员检测的准确性降低, 所以引入可变注意力机制。融

合 DCNV4 的 DAttention 注意力机制能够更好地捕捉不同尺度下的特征, 同时使注意力不再固定的分

布,可以使注意力机制随数据特性进行自适应变化,从而将最关键的特征充分提取,聚焦重要位置,同时减少了不必要的计算,性能表现优异。

1.4 Shape-IoU 损失函数

IoU 损失通常用来衡量预测边界框与真实边界框之间的相似度,计算公式为

$$L_{IoU} = 1 - IoU = 1 - \frac{|A \cap B|}{|A \cup B|}. \quad (5)$$

原始 IoU 损失函数对边界框的形状和尺度方面不够敏感, Shape-IoU 是通过关注边界框自身形状和尺寸的边界框回归方法。

Shape-IoU 的计算公式为

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|}, \quad (6)$$

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}}, \quad (7)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}}, \quad (8)$$

$$L_{IoU} = 1 - IoU = 1 - \frac{|A \cap B|}{|A \cup B|}, \quad (9)$$

$$distance^{shape} = \frac{hh \times (x_c - x_c^{gt})^2}{c^2} + \frac{ww \times (y_c - y_c^{gt})^2}{c^2}, \quad (10)$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta, \theta = 4, \quad (11)$$

$$\begin{cases} \omega_w = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})}, \\ \omega_h = ww \times \frac{|h - h^{gt}|}{\max(h, h^{gt})}, \end{cases}$$

式中: $scale$ 为缩放因子,与数据集中目标的大小相关; ww 和 hh 分别为水平方向和垂直方向上的权重系数,其值与 GT 框的形状有关; $shape$ 为对应边框的形状参数。相应的边界框回归损失为

$$L_{Shape-IoU} = 1 - IoU + distance^{shape} + 0.5 \times \Omega^{shape}. \quad (12)$$

Shape-IoU 参数的示意图如图 5 所示。

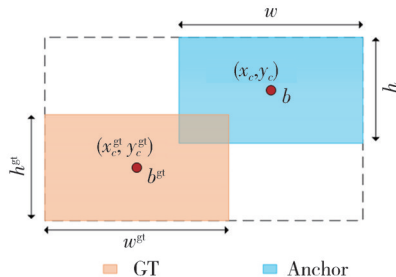


图 5 Shape-IoU 参数示意图

Fig. 5 Schematic diagram of Shape-IoU parameters

通过以上公式及示意图可以分析出 Shape-IoU 在 IoU 的基础上添加新的几何约束来使回归更准确,加强了对边界框自身形状和尺寸的考虑,引入了新的权重系数,优化了原有损失函数的不足。

2 井下人员数据集和模型训练

2.1 数据集

实验数据使用公开发表的地下长壁开采工作面图像数据集(DsLMF+)^[14],其中包含来自煤矿井下综采面的矿井人员、液压支护板、大型煤、拖链、矿工行为和矿山安全帽的图像。本文主要针对煤矿井下人员进行检测,故只选取数据集中综采面的矿山人员图片。对 DsLMF+ 数据集中的矿井人员图片数据进行筛选,选取来自不同场景的图片 2 600 张。在此基础上,运用图像增强技术(旋转、缩放和高斯模糊)对原始图像进行扩增,生成 2 000 张增强图片。经过整合共选取 3 297 个训练样本,其中 80% 为训练集,包含 2 637 个样本,剩余 20% 为测试集,包含 660 个样本。为保证模型对多尺度目标的适应性,对所选样本的尺度进行统计,其中小目标占比 19.26%,中目标占比 14.60%,大目标占比 66.14%。

2.2 改进 YOLOv8-F 模型训练

为保证实验的公平合理,所有实验的实验环境与参数配置均保持相同,操作系统为 Windows10, GPU 为 NVIDIA RTX4060ti,具体参数设置:初始学习率为 0.01,输入图像大小为 640×640 像素,优化器使用随机梯度下降算法,训练轮数为 200 轮,为防止过拟合和充分利用进程增加训练速度,批次大小设置为 16,线程数设为 4,在最后 50 个训练周期中禁用马赛克数据增强,增加模型对图像的鲁棒性和泛化能力。

为了验证本文模型对井下人员检测时可靠有效,对结果进行评价,本文选择评价指标精确度、召回率、每秒次浮点运算数(FLOPs)、平均精度均值(mAP)来综合评估模型性能。其中,召回率、精确度、mAP 的公式为

$$R = \frac{TP}{(FN + TP)}, \quad (13)$$

$$P = \frac{TP}{(TP + FP)}, \quad (14)$$

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k, \quad (15)$$

式中： TP 为检测到的井下人员目标个数； FP 为检测到的环境非人员的目标个数； n 为类别总数； k 为检测次数； AP_k 为第 k 个类别的平均准确率。

对YOLOv8-F模型进行200轮训练，并对所得结果绘制曲线，通过所得训练过程中的损失函数曲线对模型性能进行评估，损失值越小，模型性能越好，预测结果更加准确。图6为训练过程中的损失函数曲线。由图6可以看出，在训练至155轮时，各损失函数值趋于稳定，表明其已收敛，反映了良好的训练效果。

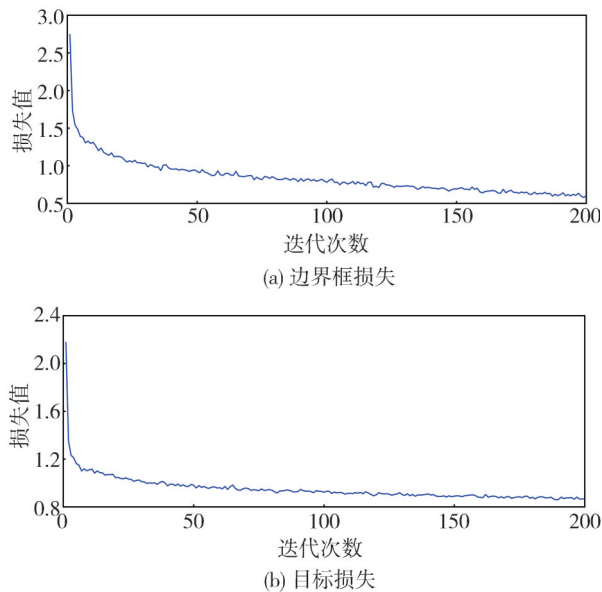


图6 损失值-迭代次数曲线图

Fig. 6 Loss value-number of iterations

3 井下人员检测实验及结果分析

3.1 模块性能对比分析

3.1.1 主干网络试验及对比分析

为检验本文所提主干网络改进方案的有效性，将当前广泛使用的主流主干网络架构与本文改进后的主干网络进行对比。实验结果如表1所示。

表1 主干网络对比

Tab. 1 Test comparison of different backbone networks

主干网络	FLOPs/G	mAP@0.5/%	P/%	R/%
YOLOv8n(baseline)	8.3	88.6	86.4	83.4
YOLOv8n-EfficientViT	8.1	88.9	85.9	76.7
YOLOv8n-EfficientFormV2	6.1	79.5	80.3	68.1
YOLOv8n-Vanillanet	7.9	89.2	86.6	83.8
YOLOv8n-MobileNetV4	7.1	88.3	87.5	84.5

EfficientViT^[15]的FLOPs有小幅下降的同时使mAP@0.5略有提升，但精确度和召回率都有一定的

下降。EfficientFormV2^[16]虽然参数量显著减少且提升了检测速度，但精确度和mAP@0.5均有明显下降，检测精度损失较多，这是因为其特有的Transformer架构在CPU上的运行效率较低导致提升速度的同时损失精度。Vanillanet^[17]是由华为诺亚提出的一种全新骨干架构，主要针对移动设备和嵌入式系统进行了优化，但在各个方面的提升均不明显，表现平庸。实验结果充分显示，MobileNetV4高效网络在井下人员检测的任务上，各项评价指标均表现优异，具有较高的检测精度和检测速度，表现出模型具有良好的检测性能。

3.1.2 注意力模块试验及对比分析

为了验证本文引进的注意力机制的改进性能，在改进过程中将现阶段主流注意力机制分别添加到YOLOv8的主干网络中。SimAM^[18]是一种轻量级、无参数的卷积神经网络注意力机制，通过计算特征图的局部自相似性来生成注意力权重，无需引入额外参数使模型运算量减少，模型变得轻量化，但在复杂环境下特征较多时，其检测精度明显下降。MLCA^[19]在传统通道注意力机制基础上融合空间信息，主要应用Local SE方法对空间信息进行处理，但对于煤矿井下复杂环境，由于在处理空间信息时采用了过多的计算，导致模型性能明显下降。CPCA^[19]是一种用于增强特征表示和动态分配注意力权重的注意力机制，结合了通道注意力和空间注意力，具备聚焦信息渠道和重要区域的能力，能够显著提升检测精度，但浮点运算量显著增加，收敛速度减慢。本文所使用的DCNV4-DAttention注意力机制能够在提升检测精度的基础上，减少计算量，提升收敛速度，达到对人员的精确检测。具体实验结果如表2所示。

表2 注意力机制对比

Tab. 2 Comparative tests of attention mechanism

注意力机制	FLOPs/G	mAP@0.5/%	P/%	R/%
YOLOv8n(baseline)	8.3	88.6	86.4	83.4
YOLOv8n-SimAM	6.8	78.6	69.5	72.5
YOLOv8n-MLCA	8.9	82.5	80.6	80.8
YOLOv8n-CPCA	9.1	90.7	88.5	73.5
YOLOv8n-DCNV4-DAttention	7.4	89.6	88.3	84.3

3.2 模型有效性分析

3.2.1 消融实验

为检测本文所提模型在煤矿井下人员检测任

务中的性能及改进模块的有效性,借助消融实验进行进一步分析。本实验选用最轻量化的n模型进行测试,实验结果如表 3 所示。采用 MobileNetV4 主干网络的 YOLOv8 模型,减少了模型的运算量,提升了精度,在此基础上,主干网络中融

合 Dcnv4-DAttention 注意力机制,计算量显著减小的同时能够提升模型的精度。结合 Shape-IoU 损失函数,通过加强对边界框自身形状和尺寸的关注,使得模型的检测结果更加精确,实验结果见表 3。

表 3 消融实验
Tab. 3 Ablation experiments

MobileNetV4	DCNV4-DAttention	Shape-IoU	FLOPs/G	mAP@0.5/%	P/%	R/%
			8.3	88.6	86.4	83.4
✓			7.1	88.3	87.5	84.5
	✓		7.4	89.6	88.3	84.3
		✓	8.0	89.3	87.6	83.9
✓	✓		7.2	90.2	88.8	85.7
	✓	✓	7.8	90.6	89.2	86.3
✓		✓	7.1	90.3	88.4	85.6
✓	✓	✓	6.9	90.9	89.6	86.1

注:“✓”表示在该组实验中引入了相应模块。

3.2.2 不同模型对比实验

为了验证本文模型的性能,在保持各种实验环境一致的前提下,将 YOLOv5n、YOLOv8s、YOLOv7tiny、YOLOv10 和本文模型进行训练,具体结果见表 4。

表 4 不同模型对比
Tab. 4 Comparison of different models

模型	FLOPs/G	mAP@0.5/%	P/%	R/%
YOLOv5n	12.6	89.5	86.4	80.3
YOLOv7tiny	13.8	90.1	88.2	82.6
YOLOv8s	28.6	88.6	83.2	85.4
YOLOv10	6.8	82.8	81.8	78.6
YOLOv8-F (本文)	6.9	90.9	89.6	86.1

表 4 结果显示,本文 YOLOv8-F 模型在对井下人员进行检测时,拥有更高的召回率、精度和更快的模型推理速度。其中,YOLOv8s 的计算量显著增加,是由于其模型深度对比 YOLOv8n 增加一倍所致。

由表 4 可知,相较于 YOLO 系列的其它模型,本文改进的 YOLOv8-F 模型准确率、mAP 和其他评价指标均优于其他模型,在井下人员检测任务中表现优异,具有更好的检测效果,能够满足检测需求。但在实验过程中也存在少许极端环境,如低光照影响下,所提模型可能会出现识别精度大幅降低的情况,可以在此模型基础上结合图像增强算法来解决,例如结合自校准照明学习框架提升模型对极端环境的适应能力。

为进一步直观地展示模型对人员进行检测时所重点关注的图像信息,本文采用 HiResCAM 热力图^[20]对所提模型进行可视化分析。HiResCAM 热力图可

视化策略是在 Grad-CAM 基础上,提出的一种新的输出级、基于梯度的卷积神经网络的解释方法,能够在对人员进行检测时解决因 Grad-CAM 平均步骤产生的限制,更加有效地聚焦到每次预测的位置,突出显示模型预测所关注的区域。图 7 所示为热力图可视化策略所得到的结果。

对图 7 中的可视化结果进行分析,可知改进后的 YOLOv8-F 模型在检测过程中能够有效聚焦于人员特征,对比原模型性能提升明显,有效解决了模型在复杂环境下对人员检测精度低的问题。

4 结 论

1) 针对煤矿井下环境复杂,干扰较多对人员检测难度较大的问题,本文提出一种改进的 YOLOv8-F 煤矿井下人员检测模型,其结合了 DCNV4-Dattention 注意力机制和 MobileNetV4 高效网络,在提升模型检测速度的同时满足了对人员特征信息的充分提取,Shape-IoU 损失函数的应用使定位框精度提升,实现了对处于煤矿井下复杂环境中人员的精准检测。

2) 与 YOLOv8n 的基线模型对比,经过改进的 YOLOv8-F 模型在浮点运算量、准确率、召回率以及平均精度 4 个关键评价指标上均取得显著提升。同时,与 YOLOv5s、YOLOv8s、YOLOv7tiny 和 YOLOv10 模型对比,本文提出的模型拥有最小的运算量和最高的平均精度值,浮点运算量仅为 6.9 G,平均精度值为 90.9%,充分显示出模型具有良好的检测性能。

3) 通过热力图可视化分析表明, 相比于原模型, 改进的 YOLOv8-F 模型能够更加准确地聚焦

人员特征, 有效检测井下人员。

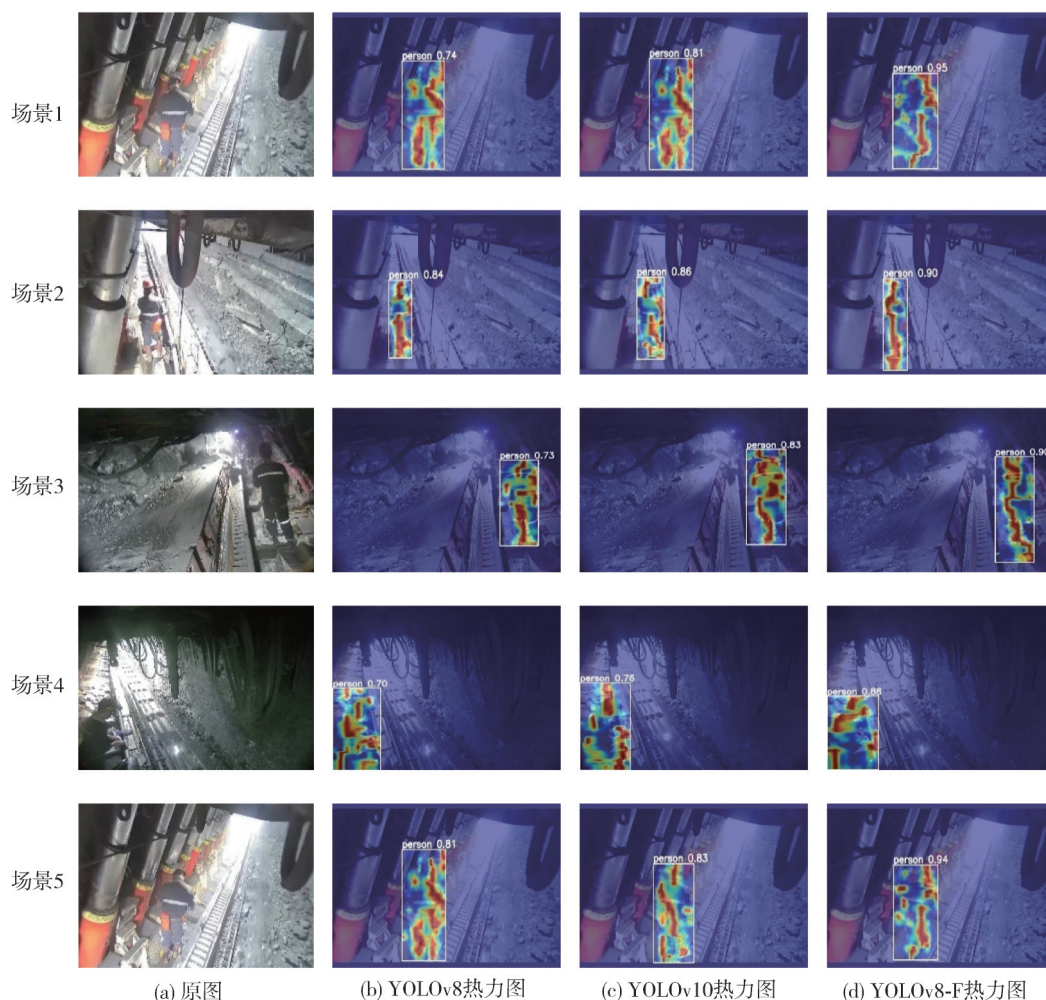


图 7 不同算法结果在多场景下的可视化对比

Fig. 7 Visual comparison of results of different algorithm

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors declare no relevant conflict of interests.

参考文献

- [1] REN S, HE K, 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.
- [2] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [EB/OL]. [2025-01-30]. <https://arxiv.org/abs/1512.02325v5>.
- [3] ZHENG S, WU J, DUAN S, et al. An improved crowd counting method based on YOLOv3 [J]. Mobile Networks and Applications, 2024, 29(4):

1181-1189.

- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [5] 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 Bibo, ZHOU Yun, LI Xiaojun, et al. YOLOv5 lightweight coal mine underground personnel detection algorithm base on attention mechanism[J]. Coal Technology, 2023, 42(10): 200-203. (in Chinese)
- [6] 范伟强, 王雪瑾, 张颖慧, 等. 改进 YOLOv7 和 DeepSORT 的井下人员检测与跟踪算法[J]. 煤炭科学技术, 2024, 52(S2): 343-355.
FAN Weiqiang, WANG Xuejin, ZHANG Yinghui, et al. Underground personnel detection and tracking

- using improved YOLOv7 and DeepSORT [J]. Coal Science and Technology, 2024, 52 (S2): 343-355. (in Chinese)
- [7] 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]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 7464-7475.
- [8] PUJARA A, BHAMARE M. DeepSORT: real time & multi-object detection and tracking with YOLO and TensorFlow [C]//2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAIS), 2022: 456-460.
- [9] 倪云峰, 霍洁, 侯颖, 等. 基于 YOLOv8-OCR 的井下人员检测算法[J]. 无线电工程, 2024, 54(8): 1847-1853.
- NI Yunfeng, HUO Jie, HOU Ying, et al. Underground personnel detection algorithm based on YOLOv8-OCR [J]. Radio Engineering, 2024, 54(8): 1847-1853. (in Chinese)
- [10] XIONG Y, LI Z, CHEN Y, et al. Efficient deformable ConvNets: rethinking dynamic and sparse operator for vision applications[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 5652-5661.
- [11] XIA Z, PAN X, SONG S, et al. Vision transformer with deformable attention[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 4784-4793.
- [12] ZHANG H, ZHANG S. Shape-IoU: more accurate metric considering bounding box shape and scale[EB/OL]. [2025-01-30]. <https://doi.org/10.48550/arXiv.2312.17663>
- [13] KIM K H, SHIM P S, SHIN S. An alternative bilinear interpolation method between spherical grids [J]. Atmosphere, 2019, 10(3): 123.
- [14] YANG W, ZHANG X, MA B, et al. An open dataset for intelligent recognition and classification of abnormal condition in longwall mining [J]. Scientific Data, 2023, 10(1): 416.
- [15] LIU X, PENG H, ZHENG N, et al. EfficientViT: memory efficient vision transformer with cascaded group attention[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 14420-14430.
- [16] LI Y, HU J, WEN Y, et al. Rethinking vision transformers for MobileNet size and speed [C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 16843-16854.
- [17] CHEN H, WANG Y, GUO J, et al. Vanillanet: the power of minimalism in deep learning[J]. Advances in Neural Information Processing Systems, 2023, 36: 7050-7064.
- [18] YANG L, ZHANG R Y, LI L, et al. Simam: a simple, parameter-free attention module for convolutional neural networks [C]//International Conference on Machine Learning, PMLR, 2021: 11863-11874.
- [19] HUANG H, CHEN Z, ZOU Y, et al. Channel prior convolutional attention for medical image segmentation [J]. Computers in Biology and Medicine, 2024, 178: 108784.
- [20] DRAELOS R L, CARIN L. Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks [EB/OL]. [2025-02-10]. <https://doi.org/10.48550/arXiv.2011.08891>.