



基于YOLOv8-ECW的井下人员行为实时检测算法

骆津津, 陈伟, 田子建, 张帆, 刘毅

Real-time detection algorithm of underground personnel behavior based on YOLOv8-ECW

引用本文:

骆津津, 陈伟, 田子建, 等. 基于YOLOv8-ECW的井下人员行为实时检测算法[J]. *矿业科学学报*, 2025, 10(2): 316–327.

LUO Jinjin, CHEN Wei, TIAN Zijian, et al. Real-time detection algorithm of underground personnel behavior based on YOLOv8-ECW[J]. *Journal of Mining Science and Technology*, 2025, 10(2): 316–327.

您可能感兴趣的其他文章

1. 基于改进HMM纠偏算法的露天矿车辆高精度定位方法

引用本文: 阮顺领, 李孟, 顾清华, 等. 基于改进HMM纠偏算法的露天矿车辆高精度定位方法[J]. *矿业科学学报*, 2023, 8(3): 381–389.

2. 基于多尺度卷积神经网络和LBP算法的浮选工况识别

引用本文: 蒋小平, 刘俊威, 王乐乐, 等. 基于多尺度卷积神经网络和LBP算法的浮选工况识别[J]. *矿业科学学报*, 2023, 8(2): 202–212.

3. 煤矸石粗集料理化性质和形状特征对混凝土强度的影响

引用本文: 王艳, 左震, 文波, 等. 煤矸石粗集料理化性质和形状特征对混凝土强度的影响[J]. *矿业科学学报*, 2022, 7(5): 554–564.

4. 煤矿采场围岩智能控制技术路径与设想

引用本文: 杨胜利, 王家臣, 李明. 煤矿采场围岩智能控制技术路径与设想[J]. *矿业科学学报*, 2022, 7(4): 403–416.

5. 基于地质统计学反演的透明化矿山岩性建模参数研究及应用

引用本文: 陈柏平, 崔凡, 刘波, 等. 基于地质统计学反演的透明化矿山岩性建模参数研究及应用[J]. *矿业科学学报*, 2022, 7(4): 427–436.

骆津津, 陈伟, 田子建, 等. 基于YOLOv8-ECW的井下人员行为实时检测算法[J]. 矿业科学学报, 2025, 10(2): 316-327. DOI: 10.19606/j.cnki.jmst.2024907

LUO Jinjin, CHEN Wei, TIAN Zijian, et al. Real-time detection algorithm of underground personnel behavior based on YOLOv8-ECW[J]. Journal of Mining Science and Technology, 2025, 10(2): 316-327. DOI: 10.19606/j.cnki.jmst.2024907

基于YOLOv8-ECW的井下人员行为实时检测算法

骆津津^{1, 2, 3}, 陈伟^{1, 3, 4}, 田子建^{1, 3}, 张帆^{1, 3}, 刘毅^{1, 3}

1. 中国矿业大学(北京)人工智能学院, 北京 100083;

2. 大唐滨州发电有限公司, 山东滨州 256651;

3. 煤矿智能化与机器人创新应用应急管理部重点实验室, 北京 100083;

4. 中国矿业大学计算机科学与技术学院, 江苏徐州 221116

摘要:针对现有煤矿井人员行为检测模型存在精度低、计算量大等问题,提出一种基于YOLOv8-ECW的井下人员行为实时检测算法。算法在YOLOv8n的基础上对骨干网络进行改进,提出多尺度卷积模块EMSC,再与C2f卷积相结合设计出C2f_EMSC模块,有效捕获目标的多尺度特征,减少模型的计算量、参数量;在网络中引入CGBlock下采样模块融合全局的上下文信息,引入WIoU损失函数提升检测框的定位精度和模型收敛速度。在矿井人员行为检测数据集上进行实验,结果表明:①相比于基线YOLOv8n模型,YOLOv8-ECW模型对各类目标平均精度均值mAP₅₀为92.4%,上升了2.1%;mAP₅₀₋₉₅为75.4%,上升了4.0%。②YOLOv8-ECW的检测速度为238 F/s,较YOLOv8n模型提高了5 F/s。③与YOLOv6、YOLOv7等主流网络模型相比,YOLOv8-ECW模型的检测性能最佳且具有较好的鲁棒性。

关键词:煤矿井下;YOLOv8;行为检测;C2f_EMSC;WIoU;特征融合

中图分类号:TD 76;TP 181

文献标志码:A 文章编号:2096-2193(2025)02-0316-12

Real-time detection algorithm of underground personnel behavior based on YOLOv8-ECW

LUO Jinjin^{1, 2, 3}, CHEN Wei^{1, 3, 4}, TIAN Zijian^{1, 3}, ZHANG Fan^{1, 3}, LIU Yi^{1, 3}

1. School of Artificial Intelligence, China University of Mining and Technology-Beijing, Beijing 100083, China;

2. Datang Binzhou Power Generation Co., Ltd., Binzhou Shandong 256651, China;

3. Key Laboratory of Intelligent Mining and Robotics, Ministry of Emergency Management, Beijing 100083, China;

4. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou Jiangsu 221116, China

Abstract: The existing models for detecting the behaviors of personnel in coal mine wells suffer from issues such as low accuracy and significant computational load. Therefore, a real-time detection algorithm for the behaviors of underground personnel based on YOLOv8-ECW is proposed. Based on YOLOv8n, the backbone network is enhanced by presenting the multi-scale convolution module EMSC. It is combined with the C2f convolution to design the C2f_EMSC module, effectively capturing the

收稿日期:2024-05-18 修回日期:2024-08-04

基金项目:国家自然科学基金(52274160, 51874300, 52074305, 52374165, 52121003)

作者简介:骆津津(1998—),男,湖北黄石人,硕士研究生,主要从事目标检测与识别方面的研究工作。E-mail: luojinjin@cumt.edu.cn

通信作者:陈伟(1978—),男,江苏徐州人,博士,教授,主要从事机器学习、图像处理和计算机网络等方面的研究工作。E-mail: chenwdavior@163.com

multi-scale features of the target and reducing the computational volume and parameter quantity of the model. The CGBlock downsampling module is introduced into the network to fuse the global context information. The WIoU loss function is incorporated to enhance the positioning accuracy of the detection box and the convergence speed of the model. Experiments conducted on the self-established dataset for detecting the behaviors of personnel in coal mines reveal the following results: ① Compared with the baseline YOLOv8n model, the average precision mean (mAP50) of the YOLOv8-ECW model for various targets is 92.4%, an increase of 2.1%; and the mAP50-95 is 75.4%, an increase of 4.0%. ② The detection speed of the YOLOv8-ECW is 238 frames per second, which is 5 frames per second higher than that of the YOLOv8n model. ③ Compared with the mainstream network models such as YOLOv6 and YOLOv7, the detection performance of the YOLOv8-ECW model is the best and it exhibits better robustness.

Key words: coal mine underground; YOLOv8; behavior detection; C2f_EMSC; WIoU; feature fusion

煤炭是我国的基础能源和工业原料,长期以来为经济社会发展和国家能源安全稳定供应提供了有力保障^[1-3]。近年来煤矿安全事故频发,造成严重的人员伤亡和财产损失。通过对大量煤矿事故分析,及时发现事故,迅速做出反应,就能避免人员伤亡^[4]。利用科技化和信息化手段提高煤矿智能化监控水平,可以有效降低人的不安全行为导致的煤矿事故^[5]。利用智能化综采工作面的视频监控技术,能够实现井下人员行为自动检测与识别,减少矿工不安全行为,对实现煤矿智能化^[6]运行、提升煤矿安全生产水平具有重要意义。

孙继平等^[7]基于压缩感知和归一化矩形特征的跟踪方法,在煤矿照明度低、目标尺度变化频繁环境下跟踪井下人员。张立亚^[8]提出基于动目标特征提取的井下高危区域动目标监测和管理系统,采用高度结构化匹配方法进行人员识别,改进尺度不变特征转换提取行为特征,实现对人员行为和车辆等的准确识别。温廷新等^[9]使用迁移学习和深度残差网络方法,实现对矿工安全行为、较安全行为、不安全行为图像的识别。张帆等^[10]提出基于SSD-LeNet的矿井移动目标检测与识别方法,有效解决井下雾尘、低照度环境中检测精度低、实时性差等问题。王卫东等^[11]针对煤中异物不规则形状和复杂特征,构建多尺度特征融合的语义分割网络,对现场采集的煤中异物图像进行端到端检测。赵学军等^[12]在YOLOv3模型基础上提出端到端煤矸石检测方法,采用深度可分离卷积和转置卷积缩减模型大小,提高模型运行速度,能够实时精准地检测出煤与矸石混合体中的矸石,提高煤炭分拣效率。刘浩等^[13]通过对人体关键点的特征提取,利

用MobileNetV3和ST-GCN网络实现煤矿井下人员不安全行为的智能识别。黄瀚等^[14]提出一种基于动态注意力与多层感知图卷积网络的煤矿井下人员行为识别方法。王宇等^[15]提出一种基于多模态特征融合的井下人员不安全行为识别方法,能够准确识别不安全行为。张春翔等^[16]基于深度卷积神经网络技术构建断层检测模型,提高煤矿断层的解释效率。

WEI等^[17]提出基于并行特征转移网络的井下人员检测器,可以在不引入额外参数和推理阶段计算成本的情况下提高检测速度。LI等^[18]针对地下环境存在光照不均、灰尘大、噪声大等因素干扰,提出基于YOLOv4的改进算法,通过超分辨率卷积神经网络和K-means++聚类解决检测精度低的问题。寇发荣等^[19]在YOLOv5算法的基础上提出轻量级Ucm-YOLOv5算法,在CPU端达到实时检测速度的标准。在文献[20]中,采用增强算法处理弱光图像,然后利用Dense-YOLO算法检测增强后的图像,从而解决矿井人员漏检的现象。为了准确检测复杂环境下的煤矿井下作业人员,SHAO等^[21]将井下图像特征与K-means++聚类锚点帧相结合,提出Rep-YOLO检测算法,在网络末端引入纵横交错与通道关注机制,以捕捉图像中的长距离依赖性。

在煤矿井下环境中,数字终端、监控等边缘设备种类多、部署方式差别大,云/边缘计算的能力受到限制,内存等计算资源相对紧张,无法满足煤矿井下人员检测模型稳定运行要求,对模型参数量、计算量和实时性等提出更高要求。YOLOv8n是YOLOv8的1个轻量级版本,参数量只有3.2M,在

边缘设备上部署更加友好。将其应用到煤矿井下场景中时,面临井下人员行为图像匮乏、亮度不均、采集图像存在大量背景相似性干扰因素、人员移动产生尺度变化会增加特征提取难度等难点。为此,提出一种基于 YOLOv8-ECW 的井下人员行为实时检测算法。经在 CMPBD 数据集上进行验证, YOLOv8-ECW 模型具有较高检测精度,满足实时性要求;经在 VOC 2012 数据集上进行验证, YOLOv8-ECW 模型具有良好鲁棒性。

1 YOLOv8 算法基本原理

2023 年 Ultralytics 公司^[22]发布了 YOLOv8 检测模型,提供 5 个缩放版本,即 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和 YOLOv8x,以满足不同场景的需求。在训练过程中, YOLOv8 使用 Mosaic 数据增强,在最后 10 个 epoch 关闭数据增强操作,提高检测精度。YOLOv8 的网络结构由输入端、主干网络、颈部、检测头构成,如图 1 所示。

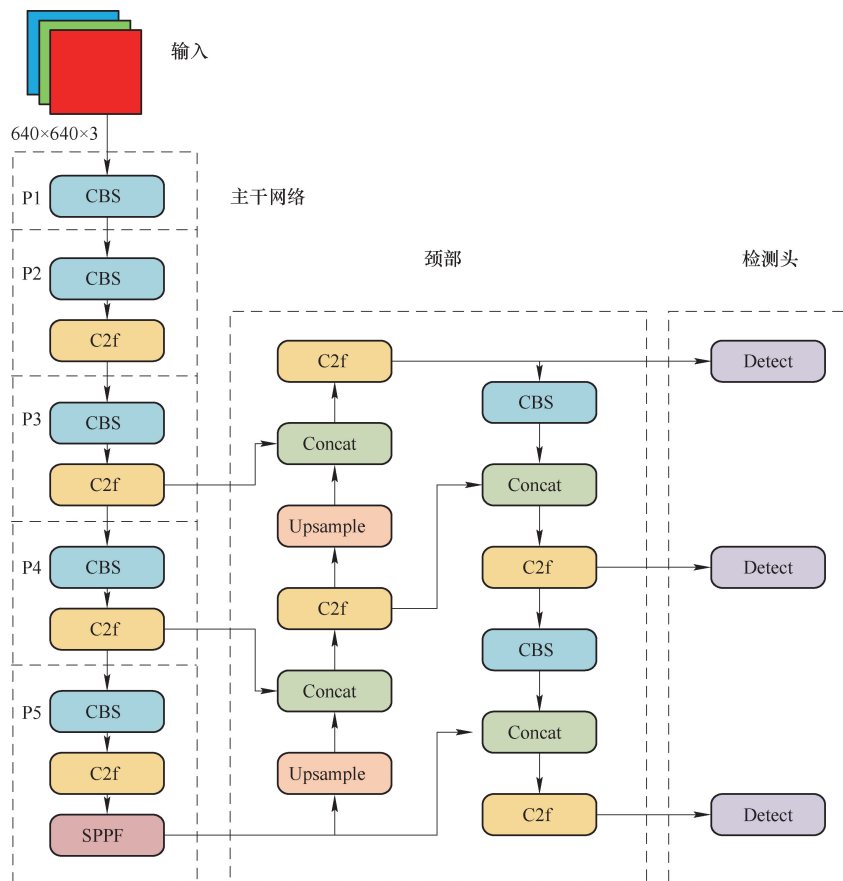


图 1 YOLOv8 网络结构
Fig. 1 YOLOv8 network architecture

1.1 输入端

YOLOv8 的输入端先对图像数据进行预处理操作,将输入图像的尺寸调整为 640 个像素×640 个像素,在不丢失目标全局特征的前提下,可以有效节约训练成本。YOLOv8 采用了马赛克增强、混合增强、空间扰动和颜色扰动的数据增强方法,增加数据量、引入噪声数据等提升模型的泛化性能和复杂环境中的鲁棒性,数据增强结果示例如图 2 所示。

1.2 特征提取网络

YOLOv8 的主干网络由 CBS 模块、C2f 模块和 SPPF 模块组成。CBS 模块由 1 个 Conv2d 层、Batch-

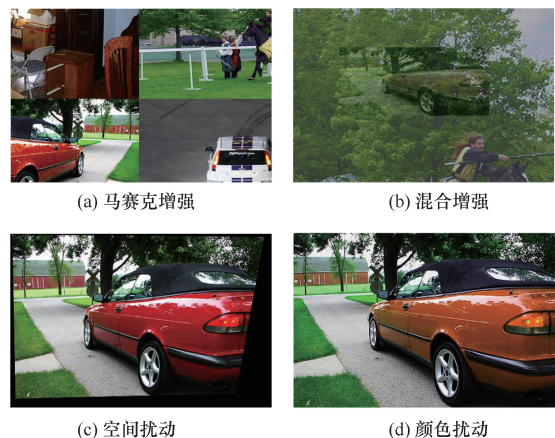


图 2 数据增强结果示例
Fig. 2 Example of data enhancement results

Norm2d 层和 SiLU 激活函数^[23]组成,用于提取原始特征,其结构如图 3 所示。

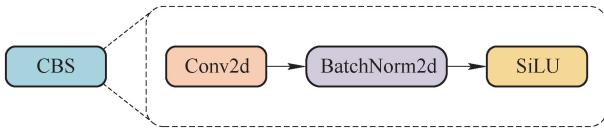


图 3 CBS 模块网络结构

Fig. 3 CBS module network structure

SiLU 函数是对 Sigmoid 和 ReLU 的改进,具有平滑、非单调的特性,定义为

$$f(x) = x\text{Sigmoid}(x) = \frac{x}{1 + e^{-x}} \quad (1)$$

SiLU 函数结合了线性函数和 Sigmoid 函数的特性。当 $x > 0$ 时, $f(x)$ 趋近于无穷大;当 $x < 0$ 时, $f(x)$ 的值会被降低,有助于抑制过拟合。

受到 ELAN 模块^[24]的启发, YOLOv8 设计了更轻量化的 C2f 模块,引入更多的分支结构丰富梯度回传的支流,替换了 YOLOv5 中的 C3 模块,使模型变得更轻量化,从而获得更加丰富的梯度流信息。C2f 和 C3 的网络结构如图 4 所示。

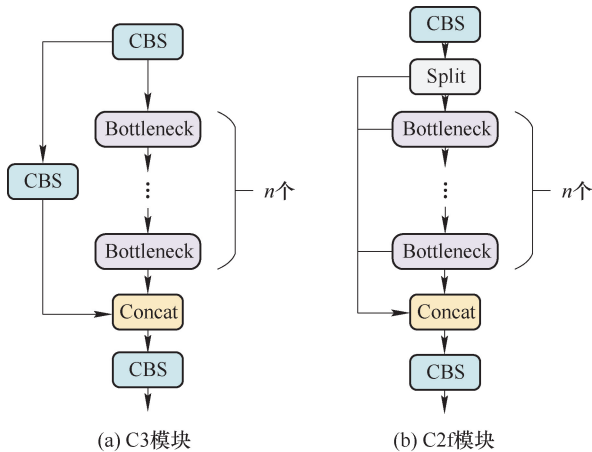


图 4 C2f 和 C3 模块网络结构

Fig. 4 C2f module and C3 module network structure

SPPF 模块为空间金字塔池化层,是在 SPP 的基础上的改进版本。SPP 模块与 SPPF 模块的网络结构如图 5 所示。SPPF 模块将并行输入的最大池化调整为多个小卷积核的串行输入,可以有效减小网络的计算量,提升计算效率。SPPF 模块还避免了对图像区域的裁剪,以及缩放操作导致的图像失真,实现局部特征和全局特征融合,完善目标的特征信息。

1.3 特征融合网络

YOLOv8 的特征融合网络采用 FPN+PAN 结构。FPN 自顶向下的结构可以将高层语义信息很

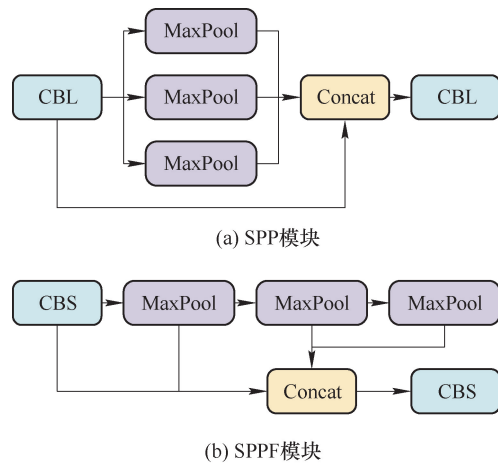


图 5 SPP 模块和 SPPF 模块网络结构

Fig. 5 SPP module and SPPF module network structure

好地传递给底层,使底层的特征表达更加丰富; PAN 可以将来自 FPN 的底层下采样的定位信息传递给顶层,增强顶层的定位能力。两者相结合能够实现不同尺度上的语义信息和定位信息的传递,使得模型可以处理不同尺度的目标,提高目标检测的准确性和鲁棒性。与 YOLOv5 相比, YOLOv8 删去了上采样层前面的 1×1 卷积,将 C3 模块替换为 C2f 模块,直接将来自骨干网络的不同阶段输出特征进行上采样操作,降低了模型大小和参数量。

1.4 检测头

YOLOv8 的检测头摒弃耦合设计,选择解耦头 (Decoupled Head) 结构,由 2 条并行的分支分别提取类别特征和位置特征,通过 1×1 卷积完成分类和定位任务。其网络结构如图 6 所示。

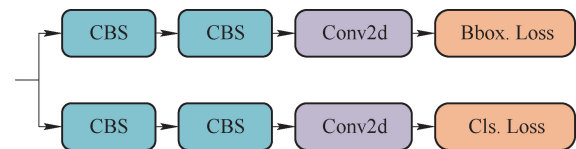


图 6 解耦头网络结构

Fig. 6 Decoupled head network structure

YOLOv8 采用 anchor-free 模型的思想,这使得网络可以直接预测目标的中心,而不是通过锚框的偏移量进行预测。该方法有效减少预测框的数量,从而提高非极大值抑制 (NMS) 的速度。

在正负样本分配策略上, YOLOv8 采用任务对齐分配的动态分配策略,对分类和回归分数进行加权,然后将加权分数进行排序后选择正样本。YOLOv8 在 loss 计算过程分为 2 个分支,即分类分

支和回归分支。前者采用二元交叉熵损失 (BCE loss); 后者采用分布焦点损失 (DFL), 同时还使用 CIoU loss 有效平衡锚框定位和类别分类之间的矛盾, 进一步提高目标检测性能。

2 基于改进 YOLOv8n 的人员行为检测

2.1 高效多尺度卷积模块

煤矿井下图像的对比度较低, 存在大量的背景相似性干扰, 井下终端设备计算能力有限, 以往的解决方案对硬件成本要求较高, 在应用部署人员检测模型时存在计算量大的问题。为了解决深度学习推理速度受限的问题, MobileNet 模型采用深度可分离卷积 (DSC) 代替传统的 CNN 结构。DSC 由深度卷积和逐点卷积组成, 相比于标准卷积, 可以显著减少模型计算量, 提高运行速度。ShuffleNet 模型采用逐点卷积和通道清洗构建网络结构, 逐点卷积对特征通道进行升维和降维操作, 通道清洗可以帮助信息在特征通道之间流动, 提高计算效率。GhostNet 模型将运算过程分为 2 个部分, 先利用普通卷积进行卷积计算得到通道数较少的特征图, 再进行廉价操作增加特征图和通道数, 然后将不同的特征图进行融合, 降低卷积层的计算成本, 用更少的参数生成更多的特征。

基于上述模型设计思想, 提出一种高效多尺度卷积 (EMSC), 将其与 C2f 融合为 C2f_EMSC 模块, 网络结构如图 7 所示, 用来替换 YOLOv8n 中的部分 C2f 结构。

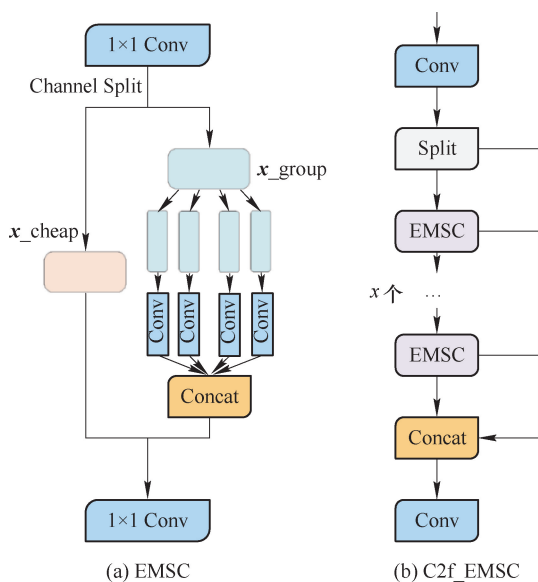


图 7 EMSC 和 C2f_EMSC 网络结构

Fig. 7 EMSC and C2f_EMSC network structure

EMSC 先利用 1×1 卷积调整输入特征通道数,

再经过通道切分 (Channel Split) 操作将输入张量 x 沿着通道维度拆分为 2 个子张量组 x_cheap 和 x_group , 每个子张量组分别具有特征通道 $1/2$ 的通道数, 接着使用相应的 Conv 模块对子张量组 x_group 进行卷积操作, 实现对图像特征的提取, 然后将提取的特征结果堆叠在一起, 最后将子张量组 x_cheap 和重新排列的子张量组 x_group 沿着通道维度连接在一起, 对连接后的张量进行 1×1 卷积操作。

相比于传统的卷积模块, EMSC 卷积模块使用 1×1 Conv 模块替换原先使用的 3×3 Conv 模块, 有效减小了卷积模块的参数数量, 进一步提升卷积处理的计算效率, 并有效利用特征通道。用 EMSC 卷积模型替换 C2f 中的部分卷积, 形成高效多尺度卷积 C2f_EMSC 模块。这种设计有效捕获不同尺度的特征, 不仅降低模型的复杂度和计算量, 而且进一步提升模型的性能和泛化能力。

2.2 下采样模块

在煤矿井下的环境中, 存在多种干扰因素, 如煤尘干扰、光照不均、背景复杂以及设备遮挡等, 导致 YOLOv8 模型的特征提取能力有所下降, 无法有效利用全局的上下文信息。因此, 通过 CGBlock 对 YOLOv8 的下采样结构进行改进, 增强其捕获全局上下文信息的能力, 从而提高检测精度。

CGBlock 的设计灵感来源于 WU 等^[25]提出的一种上下文引导网络 (CGNet), 是一种轻量级、高效的语义网络。CGNet 具备高效学习能力, 能够捕捉局部特征和上下文联合特征, 进而改善与全局上下文的特征表现。在 CGNet 的基础上, 设计了一种用于 YOLOv8 模型的 CGBlock 下采样模块, 其结构如图 8 所示。

CGBlock 主要由 4 个部分组成, 即局部特征提取器 $f_{loc} (*)$ 、周围上下文特征提取器 $f_{sur} (*)$ 、联合特征提取器 $f_{joi} (*)$ 和全局上下文特征提取器 $f_{glo} (*)$ 。CGBlock 的计算过程包含特征学习和特征改进 2 个步骤。

在特征学习步骤中, $f_{loc} (*)$ 被实例化为 1 个 3×3 标准卷积层, 从相邻的 8 个特征向量中学习局部特征。 $f_{sur} (*)$ 被实例化为 1 个 3×3 空洞/膨胀卷积层, 其具有相对较大的感受野, 可以有效地学习周围的上下文。然后, $f_{joi} (*)$ 将来自 $f_{loc} (*)$ 和 $f_{sur} (*)$ 输出的局部特征和周围上下文信息结合起来。为了降低计算复杂度, 将 $f_{joi} (*)$ 设计为 1 个连接层, 然后再经过批量归一化 (BN) 和 PReLU 算子处理。

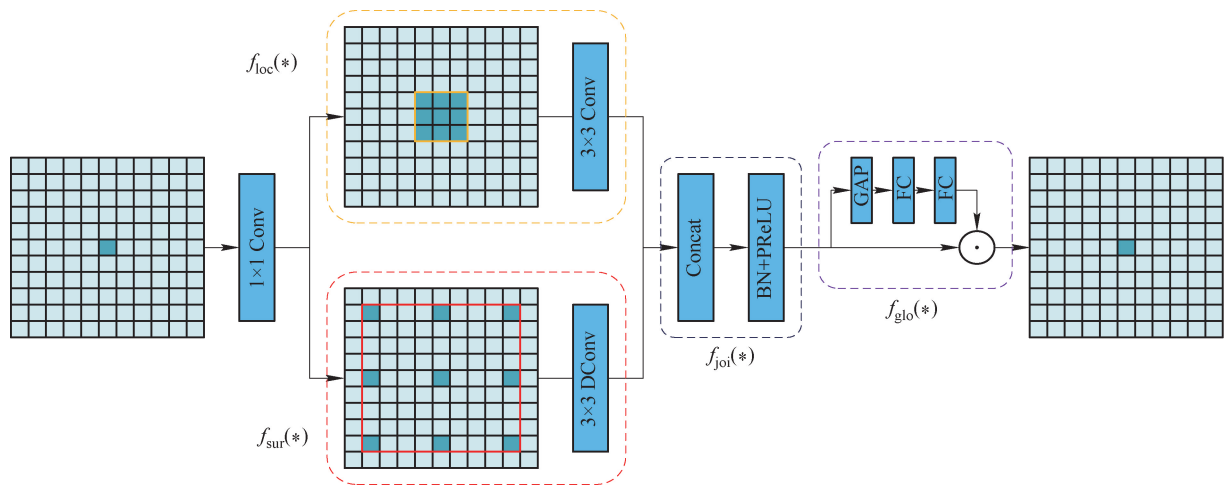


图 8 CGBlock 网络结构

Fig. 8 CGBlock network structure

在特征改进步骤中, $f_{glo}(*)$ 则负责提取全局上下文信息,进一步优化联合特征。全局上下文视为加权向量,并将其用于 Channel-Wisely 细化联合特征,突出有用的部分,抑制无用的部分。 $f_{glo}(*)$ 被实例化为 1 个全局平均池化层,后面是由 2 个全连接层组成的多层感知器,可以有效聚合场景图像的全局上下文信息。最后,使用尺度层对提取到的全局上下文信息进行联合特征加权。

2.3 损失函数优化

YOLOv8 的边界框损失函数采用的是 DFL 和 CIoU loss,将两者结合作为回归损失。

DFL 采用该位置到边界框四周的相对偏移量作为回归目标。在真实的场景中,目标边界不是非常明确的,会出现多个目标遮挡或重叠情况,所以需要学习更宽的范围分布。因此,DFL 选择扩大最接近标签 y 的 y_i 和 y_{i+1} 的概率($y_i < y < y_{i+1}$),网络能够快速地聚焦到标注位置附近的数值。DFL 的定义如下:

$$DFL = (y - y_{i+1}) \log S_i + (y_i - y) \log S_{i+1} \quad (2)$$

式中, S_i, S_{i+1} 为概率输出值。

$$DFL \text{ 的全局最小解为 } S_i = \frac{y_{i+1} - y}{y_{i+1} - y_i}, S_{i+1} =$$

$\frac{y - y_i}{y_{i+1} - y_i}$,可以保证其作为回归损失的准确性。

交并比 (IoU) 主要用于评估目标检测和预测结果中位置信息的准确性,衡量的是预测边界框与原始图像中标注的真实框之间的重叠率, IoU 值越大,代表预测的准确性越高。CIoU 对目标的大小尺寸更加敏感,因此 YOLOv8 采用 CIoU,示意如图 9 所示。

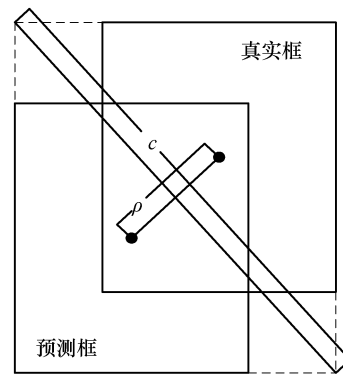


图 9 CIoU 示意

Fig. 9 Schematic diagram of CIoU

边界框回归损失 L_{CIoU} 的计算公式:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(p, p_{gt})}{c^2} + \alpha V \quad (3)$$

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

$$\alpha = \frac{V}{(1 - IoU) + V} \quad (5)$$

式中, ρ 为预测框和真实框的中心点之间的距离; p 和 p^{gt} 分别为预测框和真实框的中心点坐标; c 为预测框角点与真实框角点之间最远的距离; V 为纵横比,用来度量长宽比的相似性; α 为权衡参数; w 和 h 为预测框的宽和高; w^{gt} 和 h^{gt} 分别为真实框的宽和高。

训练数据会存在低质量样本,在模型训练过程中,加强低质量样本的边界框回归损失,会降低模型的泛化能力,而较少的干预则会提高模型的泛化能力。在此基础上, TONG 等^[26]提出了 L_{WIoUv1} , 计算公式为

$$L_{WIoUv1} = R_{WIoU} L_{IoU} \quad (6)$$

式中, $R_{WIoU} \in [1, e)$, $L_{IoU} \in [0, 1]$ 作为 2 层注意力机制, 平衡了模型对不同质量样本的关注。

焦点损失设计了单调聚焦机制, 使得模型能够聚焦于困难样本。WIoUv2 在其基础上, 提出了单调聚焦系数 L_{IoU}^* 。 L_{WIoUv2} 的计算公式为

$$L_{WIoUv2} = rL_{WIoUv1} \quad (7)$$

$$r = \left(\frac{L_{IoU}^*}{L_{IoU}} \right)^\gamma$$

式中, γ 为梯度增益; $\overline{}$ 为动量为 m 的滑动平均值, 有效解决训练后期收敛缓慢问题。

动态聚焦机制能够使模型在训练过程中动态调整梯度增益, 降低高质量样本的竞争力, 减小低质量样本产生的有害梯度, 提高模型的定位能力。WIoUv3 设计了离群度 β 来描述锚框质量, 定义为

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (8)$$

WIoUv3 利用离群度 β 构建了 1 个非单调聚焦系数 r , 定义为

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \quad (9)$$

当 $\beta = \delta$ 时, $r = 1$, 当离群度 β 为恒定值时, 其梯度增益最大。 L_{WIoUv3} 的计算公式为

$$L_{WIoUv3} = rL_{WIoUv1} \quad (10)$$

煤矿井下环境的特殊, 受照度低、粉尘多、水汽多和环境复杂等因素的影响, 获取的图像数据质量不一, 而且井下人员移动导致的尺度变化, 也会使训练样本的质量存在差异。因此, 将采用 DFL 和 WIoUv3 相结合的方法得到边界框回归损失。

2.4 YOLOv8-ECW 网络结构

在 YOLOv8n 的基础上, 设计了高效多尺度卷积 C2f_EMSC 模块, 可有效捕获目标的多尺度特征, 减少了模型的计算量、参数量; 引入 CGBlock 下采样模块融合全局的上下文信息; 采用 WIoUv3 作为边界框的回归损失, 弱化模型对尺度变化的敏感性, 加强边界框的回归效果, 提高模型的收敛速度, 进一步提升定位精度和模型的泛化能力。将上述改进后的模型命名为 YOLOv8-ECW, 其网络结构如图 10 所示。

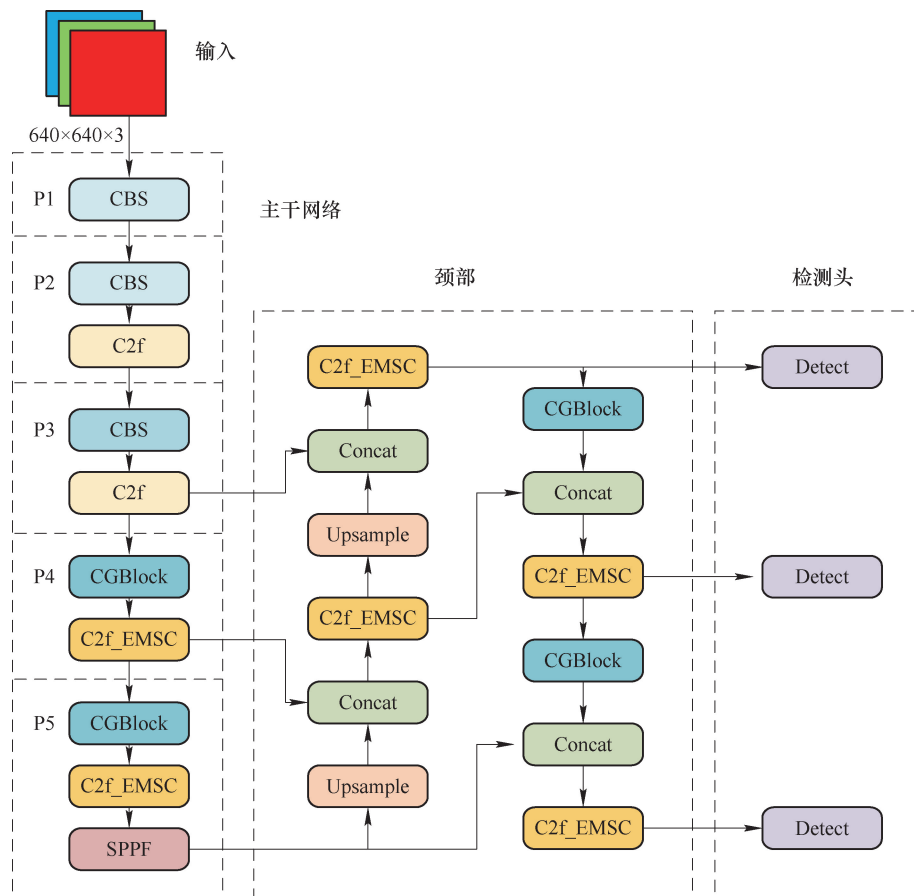


图 10 YOLOv8-ECW 网络结构

Fig. 10 YOLOv8-ECW network structure

3 实验与分析

3.1 实验环境

在PC端进行实验,操作系统为Windows 11。使用PyTorch 1.13.1作为网络学习框架,编程语言为Python 3.9.16,采用并行计算架构CUDA 11.7.1。在硬件配置方面,CPU选用Intel(R)Core(TM)i9-12900H@2.50GHz;GPU选用NVIDIA GeForce RTX 3060,6GB显存;RAM选用DDR5 16GB(4800MHz)。

模型训练采用相同的参数设置,输入图像的尺寸统一设置为640×640;批量大小batch设置为16;迭代周期epoch设置为300;初始学习率为0.01;动量参数设置为0.937;衰减系数为0.005;学习率优化器使用随机梯度下降(SGD)。除特别说明外,所有实验均采用上述设置。

3.2 数据集

实验使用自制的煤矿人员行为检测CMPBD数据集对模型进行训练和测试,从而验证改进算法的有效性。

CMPBD数据集图像一部分来自煤矿井下综采工作面异常状态图像数据集;另一部分来自马泰壕煤矿、潘三矿井下监控视频,通过视频抽帧、筛选得到图像数据,涵盖井下巷道、工作面、主斜井、副井底等场景。部分数据集图片如图11所示。



图11 CMPBD数据集图片示例

Fig. 11 Example of CMPBD dataset image

CMPBD数据集有19200张图像,通过LabelImg软件对数据集图像进行标注,生成具有YOLO格式的txt标注文件。该数据集包含8种行为标签,包括行走、站立、坐、操作、弯腰、倚靠、跌倒和攀爬。

数据集统计的小(0~32×32像素)、中(32×32~96×96像素)、大(96×96~640×640像素)真实框的分布情况如图12所示。低于32×32像素的小目

标真实框数量约为总数的28%,客观反映了修改边界框回归损失的必要性。

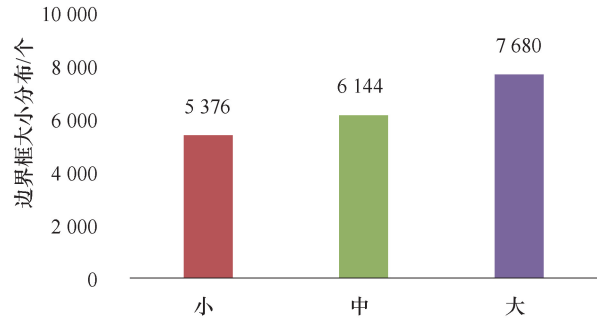


图12 边界框大小分布

Fig. 12 Bounding box size distribution

图13统计了CMPBD数据集的8种行为标签的样本数量。

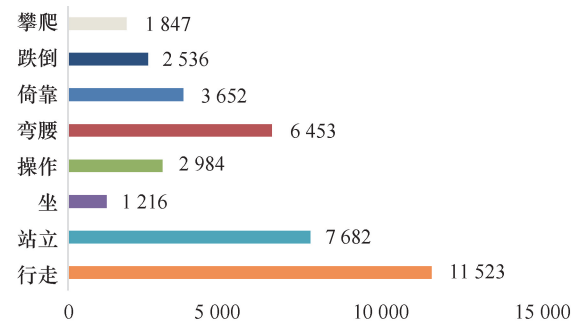


图13 8种行为标签数量

Fig. 13 Number of eight behavior labels

实验将CMPBD数据集图像按照7:1:2的比例进行划分,共得到13440张训练集图像、1920张验证集图像和3840张测试集图像。

3.3 评价指标

评价指标包括平均精度均值(mAP)、检测速度(帧率)、参数量(Params)、每秒执行10亿次浮点运算数(GFLOPs)。

AP是PR曲线中精度值的均值,用来表示在不同召回率情况下的平均检测准确度,为评价目标检测器在某一类别的效果提供依据。mAP即每一类AP的均值,特指不同召回率下的均值精度值,用来度量模型对各类别的检测性能:

$$AP = \int_0^1 P(r) dr \quad (11)$$

$$mAP = \frac{\sum_{n=1}^n AP(n)}{n} \quad (12)$$

式中, $P(r)$ 为召回率为 r 时准确率 P 的值; n 为检测类别总数; $AP(n)$ 为模型在第 n 类上的AP值。

FPS表示检测器在1s内可以处理的图像帧

数,能够衡量算法的检测速度。FPS 值越大,意味着算法具有更快的处理速度:

$$\text{FPS} = \frac{1}{t_1 + t_2 + t_3} \quad (13)$$

式中, t_1 为图像预处理耗费的时间; t_2 为训练好的模型对预处理后的图像进行预测耗费的时间; t_3 为在模型推理后,对模型输出进一步优化和调整耗费的时间。

Params 是衡量模型复杂度和性能的重要因素,参数量越小,模型越轻量化。GFLOPs 用于衡量模型的计算复杂度关键指标,该指标越低,模型所需计算量越小,模型效率越高。

3.4 消融实验

将现有主流损失函数,如 SIoU、EIoU 和 AlphaIoU 等,与 WIoUv3 在 CMPBD 数据集上进行实验,结果见表 1, WIoUv3 实现了最佳的性能增益。

表 1 不同损失函数实验结果

Table 1 Experimental results of different loss functions

损失函数	SIoU	EIoU	AlphaIoU	WIoUv1	WIoUv2	WIoUv3
mAP ₅₀ /%	90.5	90.6	88.4	90.8	90.9	91.2
mAP ₅₀₋₉₅ /%	73.0	73.2	73.3	73.5	73.5	73.6

为了验证对 YOLOv8 基线模型的优化效果以及改进方法的有效性,设置了 8 组消融实验(N1~N8)在自建的 CMPBD 数据集上进行实验,验证每

个改进策略对于模型性能的影响。每组消融实验的环境配置和参数设置均保持一致,消融实验的结果见表 2。

表 2 消融实验结果

Table 2 Results of ablation experiments

模型	C2f_EMSC	CGBlock	WIoUv3	mAP ₅₀ /%	mAP ₅₀₋₉₅ /%	参数量/M	计算量/GFLOPs	检测速度/(F·s ⁻¹)
N1	-	-	-	90.3	71.4	3.01	8.1	233
N2	√	-	-	91.4	73.2	2.72	7.6	256
N3	-	√	-	90.7	71.8	2.14	6.5	263
N4	-	-	√	91.2	73.6	3.01	8.1	244
N5	√	√	-	91.6	74.3	3.31	8.4	233
N6	√	-	√	91.3	73.5	2.72	7.6	250
N7	-	√	√	91.5	73.8	3.6	8.8	233
N8	√	√	√	92.4	75.4	3.31	8.4	238

注:“√”表示该实验组使用了相应的改进策略;“-”表示未使用相应的改进策略。

表 2 中, N1 为基线 YOLOv8n 模型的实验结果, mAP₅₀ 为 90.3%, mAP₅₀₋₉₅ 为 71.4%, 参数量为 3.01M, 计算量为 8.1 GFLOPs, FPS 为 233 F/s, 可作为后几组实验的对比基准。在基线模型中, N2、N3、N4 分别使用 C2f_EMSC 卷积、引入 CGBlock 下采样模块、采用 WIoUv3 损失函数, 模型的平均检测精度 mAP₅₀ 分别提升了 1.1%、0.4% 和 0.9%, mAP₅₀₋₉₅ 分别提升了 1.8%、0.4% 和 2.2%, 参数量和计算量略有降低, 检测速度分别提高了 23、30、11F/s。

N5 为同时使用 C2f_EMSC 卷积和 CGBlock 下采样模块的实验结果, mAP₅₀ 提升了 1.3%, mAP₅₀₋₉₅ 提升了 2.9%, 参数量增加了 0.3M, 计算量增加了 0.3 GFLOPs, 检测速度为 233 F/s。N6

为同时使用 C2f_EMSC 卷积和 WIoUv3 损失函数的实验结果, mAP₅₀ 提升了 1.0%, mAP₅₀₋₉₅ 提升了 2.1%, 参数量降低了 0.29M, 计算量降低了 0.5GFLOPs, 检测速度提高了 17 F/s。N7 为引入 CGBlock 下采样模块和 WIoUv3 损失函数的实验结果, mAP₅₀ 提升了 1.2%, mAP₅₀₋₉₅ 提升了 2.4%, 参数量增加了 0.59M, 计算量增加了 0.7GFLOPs, 检测速度为 233 F/s。N8 为最终的改进模型 YOLOv8-ECW 的实验结果, mAP₅₀ 提升了 2.1%, mAP₅₀₋₉₅ 提升了 4.0%, 参数量增加了 0.3 M, 计算量增加了 0.3 GFLOPs, 检测速度提高了 5 F/s。

实验结果表明, 提出的改进策略能明显提升 YOLOv8 模型性能, 有效解决了煤矿井下人员行为检测的问题, 满足煤矿井下高速实时检测的要求。

3.5 对比实验

为进一步验证YOLOv8-ECW算法的先进性,将改进模型与Fast-RCNN、SSD、YOLOv5s、YOLOv6s等主流模型在CMPBD数据集上进行对比实验,结果见表3。

表3 实验结果对比

Table 3 Comparison of experimental results

模型	mAP ₅₀ / %	mAP ₅₀₋₉₅ / %	参数量/ M	检测速度/ (F · s ⁻¹)
Fast-RCNN	82.8	59.4	134.7	26
SSD	84.6	61.5	57.5	84
YOLOv5s	88.4	67.8	7.2	192
YOLOX-S	87.8	65.3	9.0	116
YOLOv6s	89.1	70.3	18.5	123
YOLOv7-tiny	88.7	71.5	6.02	208
YOLOv8-ECW	92.4	75.4	3.31	238

由表3可以看出,YOLOv8-ECW模型在精度、参数量、速度等指标上优于其他主流模型,在煤矿井下人员行为检测任务中具有更好的检测性能和效率。

3.6 可视化分析

Grad-CAM是一种用于解释卷积神经网络决策的方法,通过可视化模型对于给定输入的关注区域来提供洞察。相比于传统的CAM方法,Grad-CAM能够处理任意种类的神经网络,因为它不需要修改网络结构或使用特定的层结构。Grad-CAM定义为

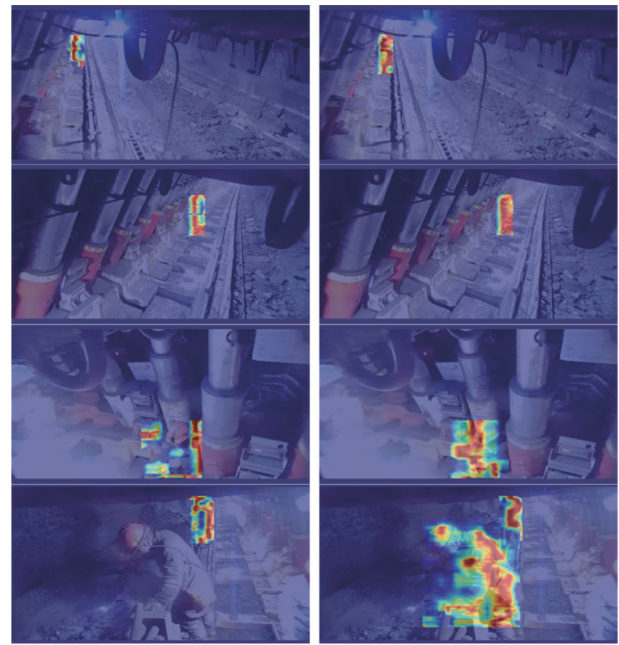
$$L_{\text{Grad-CAM}} = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right) \quad (14)$$

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (15)$$

式中, A^k 为特征层A中的第k个通道;Z为 A^k 的元素总数(宽和高的乘积); y^c 为类别c的预测分数; A_{ij}^k 为特征层A在通道k中坐标位置为(i,j)的数据。

基于Grad-CAM对改进模型的关键特征层进行热力图可视化,将标注8种行为标签的图像输入模型网络指定检测层中,进行梯度求导,经过全连接层和ReLU激活函数操作,最终得到热力图可视化的特征图,检测效果对比如图14所示。

由图14可视化结果分析可知,相比于基线模型,YOLOv8-ECW模型解决了漏检问题,能够有效提取图像特征信息,检测网络更加专注于井下人员。这进一步验证了YOLOv8-ECW模型检测结



(a) YOLOv8n

(b) YOLOv8-ECW

图14 热力图可视化效果

Fig. 14 Heat map visualization effect

果的可信度高。

3.7 鲁棒性验证

为了进一步验证YOLOv8-ECW模型的鲁棒性,将基线YOLOv8n与YOLOv8-ECW在VOC2012数据集上进行实验。VOC2012数据集共有17125张图片,按照7:1:2的比例进行划分,得到11987张训练集图像、1713张验证集图像和3425张测试集图像。基线YOLOv8n与YOLOv8-ECW在VOC2012数据集上的实验结果见表4。相较于基线模型,YOLOv8-ECW模型的mAP₅₀为60.6%,提高了0.5%;mAP₅₀₋₉₅为44.5%,提高了0.8%;参数量为3.31M,略微增加0.3M;检测速度为238 F/s。实验结果表明,YOLOv8-ECW模型具有良好的鲁棒性。

表4 VOC2012数据集上的实验结果

Table 4 Experimental results on the VOC 2012 dataset

模型	mAP ₅₀ / %	mAP ₅₀₋₉₅ / %	Params/ M	检测速度/ (F · s ⁻¹)
YOLOv8n	60.1	43.7	3.01	233
YOLOv8-ECW	60.6	44.5	3.31	238

4 结论

(1) 基于YOLOv8网络模型,结合煤矿井下影响视觉检测的因素,通过高效多尺度卷积C2f_EMSC和CGBlock下采样来提升对图像中人员区域的特

征学习效果,采用 WIoUv3 作为边界框损失函数,加强边界框的回归效果,提高模型收敛能力,进一步提升定位精度和模型泛化能力。

(2)CMPBD 数据集上的消融实验结果验证了改进方案的有效性,相较基线模型, mAP_{50} 提高 2.1%, mAP_{50-95} 提高 4.0%, 参数量略微增加 0.3M;模型计算量略增 0.3 GFLOPs;检测速度提高 $5 F \cdot s^{-1}$ 。对比实验结果证明 YOLOv8-ECW 模型的优越性能,满足实时性要求。YOLOv8-ECW 模型在 VOC 2012 数据集上的实验结果与基线模型相比, mAP_{50} 提高 0.5%, mAP_{50-95} 提高 0.8%, 参数量略增 0.3M;检测速度为 238 F/s,具有良好的鲁棒性。

(3)随着 YOLO 系列算法以及其他更先进、更高效、更轻量化检测算法的发展,将其应用于煤矿井下场景,并针对具体实际应用需求不断完善,实现井下关键目标的自动识别、场景实时动态智能感知,可为煤矿智能化安全生产提供助力。

参考文献

- [1] 王国法,赵国瑞,任怀伟. 智慧煤矿与智能化开采关键核心技术分析[J]. 煤炭学报,2019,44(1): 34-41.
WANG Guofa,ZHAO Guorui,REN Huaiwei. Analysis on key technologies of intelligent coal mine and intelligent mining[J]. Journal of China Coal Society,2019,44(1): 34-41.
- [2] 丁恩杰,俞啸,夏冰,等. 矿山信息化发展及以数字孪生为核心的智慧矿山关键技术[J]. 煤炭学报,2022,47(1): 564-578.
DING Enjie,YU Xiao,XIA Bing,et al. Development of mine informatization and key technologies of intelligent mines[J]. Journal of China Coal Society,2022,47(1): 564-578.
- [3] 王国法,张建中,薛国华,等. 煤矿回采工作面智能地质保障技术进展与思考[J]. 煤田地质与勘探,2023,51(2): 12-26.
WANG Guofa,ZHANG Jianzhong,XUE Guohua,et al. Progress and reflection of intelligent geological guarantee technology in coal mining face[J]. Coal Geology & Exploration,2023,51(2): 12-26.
- [4] 孙继平. 煤矿事故分析与煤矿大数据和物联网[J]. 工矿自动化,2015,41(3): 1-5.
SUN Jiping. Accident analysis and big data and Internet of Things in coal mine[J]. Industry and Mine Automation,2015,41(3): 1-5.
- [5] 程德强,寇旗旗,江鹤,等. 全矿井智能视频分析关键技术综述[J]. 工矿自动化,2023,49(11): 1-21.
CHENG Deqiang,KOU Qiqi,JIANG He,et al. Overview of key technologies for mine-wide intelligent video analysis[J]. Journal of Mine Automation,2023,49(11): 1-21.
- [6] 苗丙,葛世荣,郭一楠,等. 煤矿数字孪生智采工作面系统构建[J]. 矿业科学学报,2022,7(2): 143-153.
MIAO Bing,GE Shirong,GUO Yinan,et al. Construction of digital twin system for intelligent mining in coal mines[J]. Journal of Mining Science and Technology,2022,7(2): 143-153.
- [7] 孙继平,贾倪. 矿井视频图像中人员目标匹配与跟踪方法[J]. 中国矿业大学学报,2015,44(3): 540-548.
SUN Jiping,JIA Ni. Human target matching and tracking method in coal mine video[J]. Journal of China University of Mining & Technology,2015,44(3): 540-548.
- [8] 张立亚. 基于动目标特征提取的矿井目标监测[J]. 煤炭学报,2017,42(S2): 603-610.
ZHANG Liya. Mine target monitoring based on moving target feature extraction[J]. Journal of China Coal Society,2017,42(S2): 603-610.
- [9] 温廷新,王贵通,孔祥博,等. 基于迁移学习与残差网络的矿工不安全行为识别[J]. 中国安全科学学报,2020,30(3): 41-46.
WEN Tingxin,WANG Guitong,KONG Xiangbo,et al. Identification of miners' unsafe behaviors based on transfer learning and residual network[J]. China Safety Science Journal,2020,30(3): 41-46.
- [10] 张帆,栾佳星,崔东林,等. 基于 SSD-LeNet 的矿井移动目标检测与识别方法[J]. 矿业科学学报,2021,6(1): 100-108.
ZHANG Fan,LUAN Jiaying,CUI Donglin,et al. SSD-LeNet based method of mine moving target detection and recognition [J]. Journal of Mining Science and Technology,2021,6(1): 100-108.
- [11] 王卫东,张康辉,吕子奇,等. 基于深度学习的煤中异物机器视觉检测[J]. 矿业科学学报,2021,6(1): 115-123.
WANG Weidong,ZHANG Kanghui,LÜ Ziqi,et al. Machine vision detection of foreign objects in coal using deep learning[J]. Journal of Mining Science and Technology,2021,6(1): 115-123.
- [12] 赵学军,李建. 一种基于深度学习的煤矸石检测方法[J]. 矿业科学学报,2021,6(6): 730-736.
ZHAO Xuejun,LI Jian. A method of coal gangue detection based on deep learning[J]. Journal of Mining Sci-

- ence and Technology, 2021, 6(6): 730-736.
- [13] 刘浩, 刘海滨, 孙宇, 等. 煤矿井下员工不安全行为智能识别系统[J]. 煤炭学报, 2021, 46(S2): 1159-1169.
- LIU Hao, LIU Haibin, SUN Yu, et al. Intelligent identification system of unsafe behavior of employees in coal mine[J]. Journal of China Coal Society, 2021, 46(S2): 1159-1169.
- [14] 黄瀚, 程小舟, 云霄, 等. 基于DA-GCN的煤矿人员行为识别方法[J]. 工矿自动化, 2021, 47(4): 62-66.
- HUANG Han, CHENG Xiaozhou, YUN Xiao, et al. DA-GCN-based coal mine personnel action recognition method[J]. Industry and Mine Automation, 2021, 47(4): 62-66.
- [15] 王宇, 于春华, 陈晓青, 等. 基于多模态特征融合的井下人员不安全行为识别[J]. 工矿自动化, 2023, 49(11): 138-144.
- WANG Yu, YU Chunhua, CHEN Xiaoqing, et al. Recognition of unsafe behaviors of underground personnel based on multi modal feature fusion[J]. Industry and Mine Automation, 2023, 49(11): 138-144.
- [16] 张春翔, 唐焯锈, 邹冠贵, 等. 深度卷积神经网络目标检测算法在煤矿断层检测上的应用[J]. 矿业科学学报, 2023, 8(6): 733-743.
- ZHANG Chunxiang, TANG Yexiu, ZOU Guangui, et al. Deep convolutional neural network target detection algorithm for coal mine fault detection[J]. Journal of Mining Science and Technology, 2023, 8(6): 733-743.
- [17] WEI X, ZHANG H T, LIU S F, et al. Pedestrian detection in underground mines via parallel feature transfer network[J]. Pattern Recognition, 2020, 103: 107195.
- [18] LI X Y, WANG S, LIU B, et al. Improved YOLOv4 network using infrared images for personnel detection in coal mines[J]. Journal of Electronic Imaging, 2022, 31(1): 013017.
- [19] 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 Farong, XIAO Wei, HE Haiyang, 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.
- [20] 张明臻. 基于Dense-YOLO网络的井下行人检测模型[J]. 工矿自动化, 2022, 48(3): 86-90.
- ZHANG Mingzhen. Underground pedestrian detection model based on Dense-YOLO network[J]. Journal of Mine Automation, 2022, 48(3): 86-90.
- [21] SHAO X Q, LIU S B, LI X, et al. Rep-YOLO: an efficient detection method for mine personnel[J]. Journal of Real-Time Image Processing, 2024, 21(2): 28.
- [22] JOCHER Glenn, CHAURASIA Ayush, QIU Jing. Ultralytics YOLOv8 (Version 8.0.0). 2023.
- [23] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning[J]. Neural Networks, 2018, 107: 3-11.
- [24] 王尧, 廖宏鹏, 叶宜豪. 通过梯度路径分析设计网络设计策略[J]. 电子与信息学报, 2023, 39(3): 975-995.
- WANG Yao, LIAO Hongpeng, YE Yihao. Designing network design strategies through gradient path analysis[J]. Journal of Information Science and Engineering, 2023, 39(3): 975-995.
- [25] WU T Y, TANG S, ZHANG R, et al. CGNet: a lightweight context guided network for semantic segmentation[J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [26] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism[J]. arXiv, 2023. <https://arxiv.org/abs/2301.10051>.

(责任编辑:陈骏)