基于MBE-YOLO的轻量化井下行人检测模型*

张娜[1],彭文韬[1],赵强[2],汝洪芳[1],吴信元[1]

(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.东北林业大学 机电工程学院, 黑龙江 哈尔滨 150040)

摘要:针对目前地下矿山行人检测算法计算量大、检测速度慢的问题,在YOLOv8n的基础上提出了一种MBE-YOLO的轻量化井下行人检测模型。首先,在加权双向特征金字塔网络基础上设计出一种更加轻量的颈部结构,能够简单高效地融合多尺度信息;其次,采用共享卷积的方式对原有网络的检测头进行重设计,并引入幻影卷积获取丰富有效的特征表示,解决了原始检测头参数量巨大的问题;最后,考虑到轻量化模型对精度的影响,添加了较为轻量的混合局部通道注意力机制,通过在全局和局部层面上自适应调节不同区域的重要性,进而提高精度。试验结果表明,相比于基准模型YOLOv8n,MBE-YOLO模型的$mAP@0.5:0.95$提升了0.2个百分点,计算量下降了44%,FPS上升了4.3帧/s,参数量下降了46%,模型大小减少了43%;与5种主流模型相比,该模型的检测精度和轻量化程度都具有较大的优势,更适合于实时检测场景。

关键词:地下矿山;行人检测;MBE-YOLO;轻量化

中图分类号:TD76 **文献标识码**:A **文章编号**:1005-2763(2025)05-0192-08

Lightweight Underground Pedestrian Detection Model Based on MBE-YOLO

ZHANG Na[1], *PENG Wentao*[1], *ZHAO Qiang*[2], *RU Hongfang*[1], *WU Xinyuan*[1]

(1.School of Electrical and Control Engineering, Heilongjiang University of Science and Technology, Harbin, Heilongjiang 150022, China;
2.School of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin, Heilongjiang 150040, China)

Abstract: In response to the problem of large computation and slow detection speed of the current pedestrian detection algorithm in underground mines, a lightweight underground pedestrian detection model named MBE-YOLO was proposed based on YOLOv8n. Firstly, a lighter neck structure was designed based on the weighted bidirectional feature pyramid network, with the goal of efficiently fusing multi-scale information. Secondly, the detection head of the original network was redesigned using shared convolution, and the phantom convolution was introduced to obtain rich and effective feature representations, thus solving the problem of the huge number of parameters in the original detection head. Finally, considering the impact of lightweight model on accuracy, a lightweight hybrid local channel attention mechanism was added, which adaptively adjusts the importance of different regions at the global and local levels to improve accuracy. Experimental results show that compared with the benchmark model YOLOv8n, the $mAP@0.5:0.95$ of MBE-YOLO model is improved by 0.2 percentage points, the computational cost is reduced by 44%, the FPS is increased by 4.3 frames/s, the number of parameters is reduced by 46%, and the model size is reduced by 43%. Compared with the five mainstream models, this model has significant advantages in detection accuracy and lightweight, making it more suitable for scenarios that require real-time detection.

Key words: Underground mine, Pedestrian detection, MBE-YOLO, Lightweight

0 引言

矿业是人类经济社会发展中不可替代的基础产业,矿山开发建设是矿业发展的中心任务,现代矿山建设需要高科技支撑,建设智能矿山是矿业发展的大趋势[1]。随着矿产资源开采的不断加速,矿山安

* **收稿日期**:2024-06-23

基金项目:黑龙江省普通高校青年创新人才培养计划项目(UNPYSCT-2020034);黑龙江省省属本科高校基本科研业务费青年创新人才培育计划项目(2022-KYYWF-0561);黑龙江科技大学博士后科研启动基金项目(2025BSH01);黑龙江省省属高等学校基本科研业务费项目(2023-KYYWF-0545)

作者简介:张娜(1988—),女,黑龙江哈尔滨人,博士,讲师,主要研究方向为车辆动力学与控制。E-mail:2018800587@usth.edu.cn

通信作者:彭文韬(2001—),男,四川南充人,硕士研究生,主要研究方向为深度学习。E-mail:1131771348@qq.com

全监测已经成为了一项至关重要的任务。在建设智慧矿山政策的不断推进下,深度学习技术在煤矿安全防护中具有巨大的发展潜力[2]。将行人检测技术运用到矿山井下对工作人员的安全可以起到极大的保障作用,因此具有重要的意义。

近年来,地面上行人检测算法发展迅速,主要分为一阶段算法和二阶段算法。一阶段算法直接在网络中同时预测类别和边界框的位置,无需单独生成候选区域,检测速度较快,但牺牲一定的准确度,如 YOLO 系列算法[3-6]、SSD 算法[7]。二阶段算法先生成候选区域,然后对这些候选区域进行分类和边界框回归,通常具有更高的准确度,但速度较慢,如 R-CNN 算法[8]、Fast R-CNN 算法[9] 和 Faster R-CNN算法[10]。由于井下环境昏暗,设备计算能力受限,传统检测算法的检测精度和速度均不能够满足井下检测要求,一些学者开始投入到井下行人检测的研究中。邹盛等[11]采用双尺度图像融合算法,将红外相机和深度相机采集的图像进行像素级融合,随后进行形态学处理,在改进网络结构的基础上,提高了行人检测的性能。WANG 等[12]针对矿山井下低照度的现象,将低照度图片通过 EnlightenGAN 网络进行图像增强,并且提出了一种 DK_YOLOv5 模型,两者结合能够明显地提高检测精度。周李兵等[13]提出了一种基于多注意力机制的可见光和红外图像融合算法,该算法能够很好地克服图像融合中细节信息损失的现象。以上模型存在网络较深、计算量较大的问题,对于检测的实时性和模型的部署还需要考虑到模型的轻量化。

基于真实煤矿井下环境,针对轻量化网络需求,本文在 YOLOv8n 模型基础上进行了改进。在加权双向特征金字塔网络(Bidirectional Feature Pyramid Network,BiFPN)的基础上,提出了一种更轻量的颈部网络结构。此外,通过共享卷积的方式来减少检测头的参数量,并引入了幻影卷积(GhostConv)来进一步轻量化模型。为了更准确地捕捉目标信息,引入了混合局部通道注意力机制(Mixed Local Channel Attention,MLCA),以此提高检测精度。

1 基于 YOLOv8n 改进的 MBE-YOLO 模型

1.1 YOLOv8 算法

YOLOv8 算法是 YOLO 系列的全新版本,由主干网络、颈部网络和头部网络组成,较以往版本有着较大的更新。YOLOv8 用梯度流更丰富的 C2f 模块替换了 C3 模块;颈部网络继续使用 PAN-FPN 结构,由特征金字塔网络(Feature Pyramid Network,FPN)和路径聚合网络(Path Aggregation Network,PAN)组合而成;头部网络换成了解耦头结构,将分类和定位分离,模型能够更好地理解目标的位置和类别,同时也从有锚框转向了无锚框,舍弃了对锚框的设置,使模型能够更灵活地适应不同尺寸、比例和形状的目标。

1.2 MBE-YOLO 模型

MBE-YOLO 模型是在 YOLOv8n 的基础上设计的一种轻量化井下行人检测模型,为了高效地融合多尺度信息且避免信息冗余,在 BiFPN 的基础上只保留了一处跨尺度连接,选择 P3、P4、P5 通道输出特征图,设计出一种轻量有效的颈部结构。对于原始检测头参数量巨大的问题,采用共享卷积的方式改变检测头结构,引入 GhostConv 进一步降低参数。为了使模型聚焦目标区域,在空间金字塔模块(SPPF)后添加了 MLCA 注意力机制,使模型能够动态地调整不同区域的重要程度,以此提高模型的精度。MBE-YOLO 结构如图 1 所示。

1.2.1 混合局部通道注意力机制

注意力机制被广泛用于改善神经网络的性能,但是传统的注意力机制通常具有较高的计算复杂度,这极大地影响了实时检测性能。混合局部通道注意力机制(MLCA)[14]是一种轻量级的注意力机制,通过在局部和全局层面上结合通道和空间注意力动态地调节不同区域的重要性,从而提高检测的精度。相较于传统的注意力机制,MLCA 避免了复杂的结构设置和计算流程,采用了简洁有效的特征提取和权重调整策略,以确保模型的高效性和实用性。该注意力机制的原理如图 2 所示,图中,C 为特征图通道数;W、H 分别为特征图的宽度和高度;Sigmoid 为激活函数。工作过程如下:

(1) 首先输入 $C\times W\times H$ 大小的特征图,通过局部平均池化(Local Average Pooling,LAP)得到尺寸为 $c\times k_s\times k_s$ 的特征图;

(2) 之后通过两个分支,第一个分支将特征图进行全局平均池化(Global Average Pooling,GAP)以捕捉全局特征,第二个分支将特征图重新排列成一维向量以获得局部特征;

(3) 第一条支路经过一维卷积后通过 Sigmoid 操作为各个通道生成注意力权值,然后通过反平均池化操作(UN-Average Pooling,UNAP)将一维向量恢复到 $c\times k_s\times k_s$ 的尺寸;

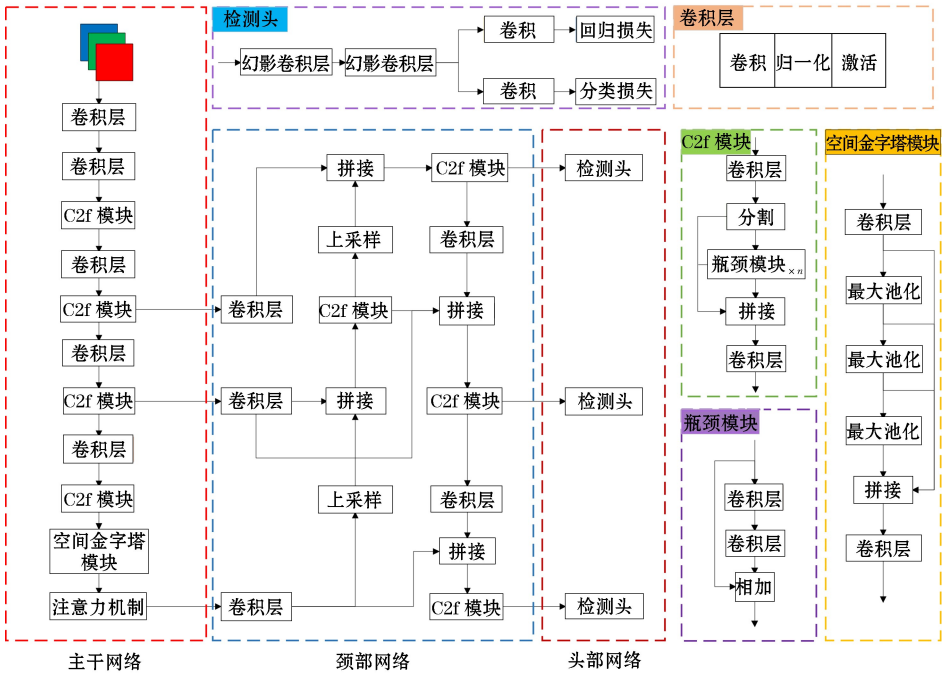


图 1 MBE-YOLO 结构

Fig.1 Structure of MBE-YOLO

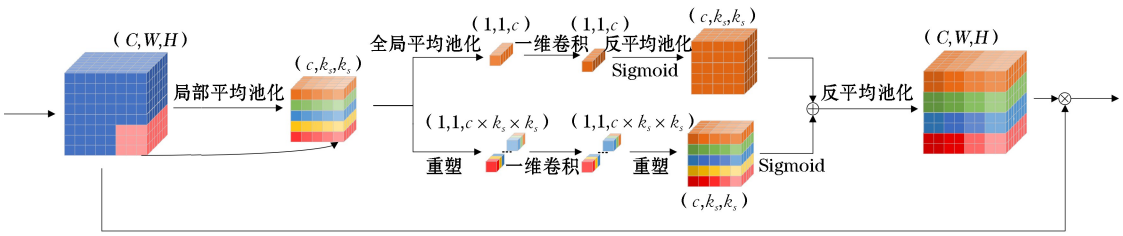


图 2 MLCA 算法的原理

Fig.2 The principle of MLCA algorithm

(4) 第二条支路经过一维卷积,再重新排列成 $c\times k_s\times k_s$ 的尺寸,最后通过 Sigmoid 为各个局部特征生成注意力权值;

(5) 最后将两条支路的特征信息进行融合,再反池化为 $C\times W\times H$ 的尺寸,并与输入特征图相结合,从而达到混合注意力的目的。

图中的一维卷积是通过输入特征图的通道数来自适应调整卷积核的大小,这可以更好地适应不同输入数据的特征,从而提高模型的泛化能力和效果。该卷积核大小 K 的计算公式为:

$$K=\varphi(c)=\left|\frac{\log_2(c)}{\beta}+\frac{\alpha}{\beta}\right|_{\text{odd}} \tag{1}$$

式中:α、β 为超参数,默认值为 2;odd 表示卷积核大小只能为奇数,若为偶数则加 1。

GAP、LAP 计算公式可表示为:

$$Z=\frac{1}{h\times w}\sum_{j}^{h}\sum_{i}^{w}X(i,j) \tag{2}$$

式中:X 为特征图;h、w 分别为池化核的高度和宽度;Z 为池化结果。

GAP、LAP 和 UNAP 原理如图 3 所示。

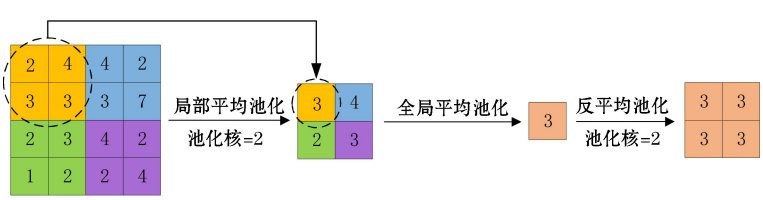


图 3 GAP、LAP 和 UNAP 原理

Fig.3 The principle of GAP, LAP and UNAP

1.2.2 改进颈部网络结构

为了简单高效地整合多尺度信息,引入了加权双向特征金字塔网络(BiFPN)[15]结构来代替 YOLOv8 原始的 PANet 颈部结构。

如图 4(a)所示,PANet 由自顶向下的特征提取和自底向上的特征汇聚两个部分组成,能够充分融合多尺度信息,但是有着较多的参数量和计算量。BiFPN 在 PANet 的基础上移除了对特征融合贡献较少的节点,以简化网络模型;引入了双向跨尺度连接的方式,以促进更好的信息传递和融合更多的信息特征,如图 4(b)所示。

为了轻量化结构,选择以 P3、P4、P5 通道的输出特征图,并且仅保留一条跨尺度连接。改进后的颈部结构如图 4(c)所示。

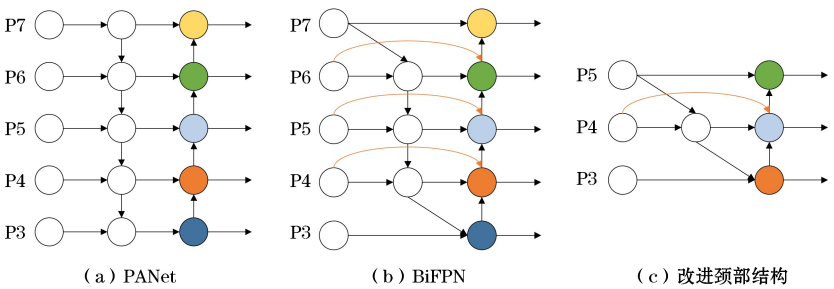


图 4 PANet、BiFPN 和改进的颈部结构

Fig.4 PANet, BiFPN and improved neck structure

1.2.3 改进检测头

如图 5 所示,YOLOv8 检测头采用解耦头结构,将分类任务和定位任务相互独立,分别预测边界框坐标和目标类别。相比于耦合头,解耦头使得每个部分可以单独优化,有助于提高模型的精度和泛化能力,但是参数量和计算量也极大地增加。

为了减少检测头的模型大小,使检测头更加轻量,将检测头上下分支中的前两个卷积层结合在一起,采用共享卷积的方式减少模型参数量。为了进一步简洁高效地提取特征信息,共享卷积层选择由两个幻影卷积组成,如图 6 所示。

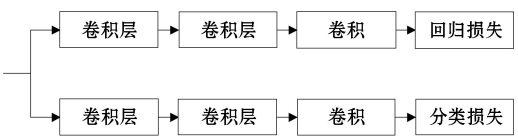


图 5 YOLOv8 检测头结构

Fig.5 The structure of detection head of YOLOv8

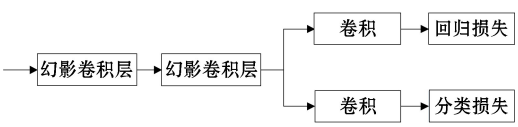


图 6 改进检测头结构

Fig.6 The structure of the improved detection head

GhostConv 是一种轻量级的卷积操作,最初由华为提出,并在 GhostNet[16]中得到了应用。GhostConv 的设计旨在降低参数量和减少计算成本,同时保持模型的性能。如图 7 所示,GhostConv 由普通卷积和分组卷积组成。首先,利用传统的卷积操作生成通道较少的特征图,以降低计算量;接着,在这些特征图的基础上,通过分组卷积进一步减少计算量,生成新的特征图;最后,将这两组特征图拼接到一起,以获取更丰富和更有效的特征参数。

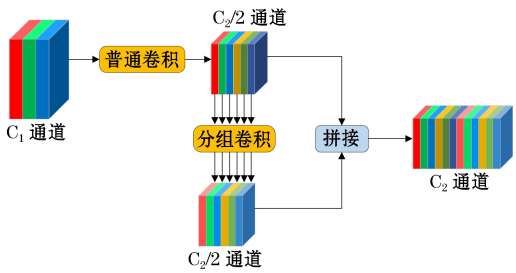


图 7 GhostConv 的结构

Fig.7 The structure of GhostConv

假设输入一个尺寸为$C_1\times W_1\times H_1$的特征图,输出一个尺寸为$C_2\times W_2\times H_2$的特征图,常规卷积核大小为 K,分组卷积核大小为 D。

普通卷积操作的参数量 N_1 为:

$$N_1=K\times K\times C_1\times C_2 \tag{3}$$

GhostConv 的参数量 N_2 为:

$$N_2=K\times K\times C_1\times\frac{C_2}{2}+D\times D\times\frac{C_2}{2} \tag{4}$$

GhostConv 和普通卷积操作的参数量比值:

$$\frac{N_2}{N_1}\approx\frac{C_1+1}{2C_1}\approx\frac{1}{2} \tag{5}$$

2 试验与结果分析

2.1 井下行人数据集制作

为了验证改进网络的可行性,采用自制数据集

进行试验。数据来源于鹤岗富力矿、七台河新建煤矿的监控视频,少量数据来源于网络收集,含有泵房、采掘工作面、机头、煤矿运行区等多个场景。

对视频每 25 帧进行一次截取,最终筛选获得 3 000 张可用图片,按照 7∶2∶1 的比例随机分为训练集、验证集和测试集,用标注软件对图片中的行人进行标注获得标签,制作出井下行人数据集。

2.2 试验环境和试验设定

试验的操作系统为 Ubuntu 20.04,深度学习框架采用 PyTorch,版本为 1.11.0,python 3.8 版本,CUDA 11.3 版本,GPU 为 NVIDIA RTX 3090 (24GB),CPU 为 AMD EPYC 9754 128-Core Processor。

模型训练选用 SGD 优化器,初始学习率设为 0.01,动量设为 0.937,权重衰减系数设为 0.000 5,训练批次大小为 64,一共迭代 300 个周期。为了确保试验结果的可靠性,模型训练过程中保持一致的参数设置和环境条件。

2.3 评价指标

本文对试验结果的精度分析采用精确率 P、召回率 R 和平均精度均值 mAP。模型的轻量化分析采用计算量、参数量、推理速度、模型大小。相关指标的运算公式如下:

$$P=\frac{TP}{TP+FP} \tag{6}$$

$$R=\frac{TP}{TP+FN} \tag{7}$$

$$AP=\int_0^1 P(R)\mathrm{d}R \tag{8}$$

$$mAP=\frac{\sum_{i=1}^{n}AP(i)}{n} \tag{9}$$

式中:TP 表示预测值为正样本,真实值也为正样本的数量;FP 表示预测值为正样本,真实值为负样本的数量;FN 表示预测值为负样本,真实值为正样本的数量;AP 为预测单个类别的平均准确率;mAP 为所有类别 AP 的均值,其中 $mAP@0.5$ 表示当置信度阈值为 0.5 时的 mAP,$mAP@0.5{:}0.95$ 表示置信度阈值从 0.5 到 0.95,步长取 0.05 的平均 mAP;n 为类别个数。

2.4 消融试验与分析

为了体现不同模块的提升效果,在 YOLOv8n 的基础上进行了 8 组消融试验,见表 1。表中“√”表示使用改进点,“×”表示不使用改进点,FPS 在 CPU 上测得。

表 1 消融试验

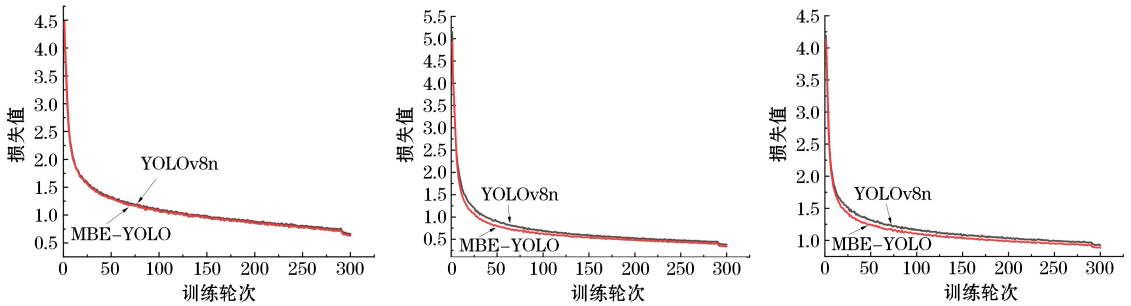
Table1 Ablation experiments

编号	注意力机制	改进颈部结构	改进检测头	$mAP@0.5$/%	$mAP@0.5{:}0.95$/%	计算量/G	推理速度/(帧/s)	参数量/M	模型大小/MB
1	×	×	×	94.4	70.5	8.1	37.6	3.01	6.0
2	√	×	×	94.7	71.5	8.1	35.9	3.01	6.0
3	×	√	×	93.9	70.2	6.6	40.4	2.00	4.1
4	×	×	√	94.4	70.3	6.0	42.7	2.73	5.5
5	√	√	×	94.7	71.3	6.6	39.5	2.00	4.1
6	×	√	√	93.8	69.9	4.5	43.4	1.63	3.4
7	√	×	√	94.4	71.1	6.0	38.7	2.73	5.5
8	√	√	√	94.4	70.7	4.5	41.9	1.63	3.4

以 YOLOv8n 为基准模型,试验 2 引入了 MLCA 注意力机制,使得 $mAP@0.5$ 提升了 0.3 个百分点,$mAP@0.5{:}0.95$ 提升了 1 个百分点,参数量、计算量和模型大小没有变化,FPS 下降了 1.7 帧/s。试验 3 将原始的 PANet 替换为改进颈部结构,计算量下降了 1.5 G,FPS 上升了 2.8 帧/s,参数量下降了 1.01 M,模型大小减少了 1.9 MB,平均精度略微下降。试验 4 对检测头进行改进,计算量下降了 2.1 G,FPS 上升了 5.1 帧/s,参数量下降了 0.28 M,模型大小减少了 0.5 M,平均精度略微下降。

对比试验 1 和试验 2、试验 3 和试验 5、试验 4 和试验 7、试验 6 和试验 8,可以充分说明 MLCA 注意力机制能够提高精度,并且不影响模型轻量化程度;对比试验 1 和试验 3、试验 2 和试验 5、试验 4 和试验 6、试验 7 和试验 8,可以表明改进颈部结构可以在较小的平均精度损失下获得较大的轻量化;对比试验 1 和试验 4、试验 3 和试验 6、试验 2 和试验 7、试验 5 和试验 8,说明改进检测头可以使模型更加轻量化,并且模型的检测性能几乎不受影响。最终

的改进模型在原始模型的基础上,$mAP@0.5:0.95$ 提升了 0.2 个百分点,计算量下降了 44%,FPS 上升了 4.3 帧/s,参数量下降了 46%,模型大小减少了 43%,改进模型轻量化的提升使得模型在实际应用中更具实用性。图 8 展示了 MBE-YOLO 改进模型和 YOLOv8n 的训练损失曲线,可以看出模型收敛时改进模型拥有更小的损失值,表明该模型的预测结果更为准确。



(a) 边界框损失 (b) 分类损失 (c) 特征点损失

图 8 训练损失曲线

Fig.8 Training loss curves

2.5 对比试验与分析

为了比较网络的改进效果,将改进后的网络与 SSD、YOLOv3-tiny、YOLOv5n、YOLOv7-tiny、YOLOv8n 进行对比试验,试验均在相同条件下进行,试验结果见表 2。

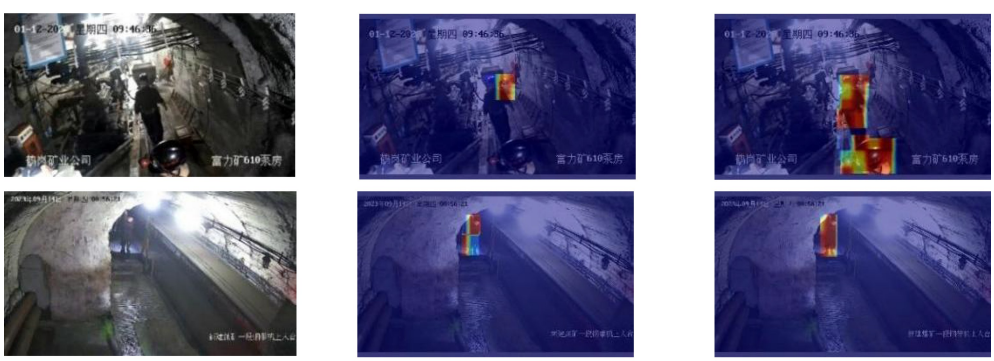
表 2 不同检测算法对比试验

Table 2 Comparative experiment of different detection models

模型	$mAP@0.5$/%	计算量/G	参数量/M	模型大小/MB
SSD	77.4	282.0	26.29	90.6
YOLOv3-tiny	94.0	12.9	8.67	16.6
YOLOv5n	94.0	4.1	1.76	3.7
YOLOv7-tiny	94.1	13.0	6.01	11.7
YOLOv8n	94.4	8.1	3.01	6.0
MBE-YOLO	94.4	4.5	1.63	3.4

由表 2 可以看出,改进算法的平均精度优于各种主流检测模型且轻量化程度达到最优,参数量和模型大小均优于其他模型。这表明 MBE-YOLO 在保持高精度的同时,显著优化了计算效率,适合于需要实时检测的应用场景。

为了体现 MLCA 注意力机制对模型的提升效果,运用 Grad-CAM++[17] 方法绘制出热力图,能够直观地展示注意力权重的分布情况。Grad-CAM++是一种用于解释深度神经网络决策的改进方法,通过基于梯度的权重计算、全局平均池化和 ReLU 操作,并引入双线性插值和注意力增强机制,能够更精准和准确地定位图像中的重要区域,提高解释性能,帮助人们理解模型的决策过程和提高模型的解释性。图 9 展示了添加 MLCA 注意力机制前后的热力图可视化。由图 9 可以看出,在添加了注意力机制后,目标区域的颜色加深,面积加大,这表明注意力机制成功地增强了模型对于目标区域的关注度,使网络能够更加准确地捕捉目标特征。



(a) 原图 (b) 未添加 MLCA 注意力机制 (c) 添加 MLCA 注意力机制

图 9 热力图可视化

Fig.9 Heat map visualization

图 10 为 YOLOv8n 算法和 MBE-YOLO 算法在自制数据集上的检测结果对比。第一组和第二组试验中存在目标遮挡和小目标的情况,YOLOv8n 算法存在一定的的漏检现象,MBE-YOLO 算法能够更好地检测小目标和处理目标遮挡问题;第三组试验中存在目标与环境融为一体,特征不明显的现象,MBE-YOLO 算法的检测效果优于 YOLOv8n 算法;第四组试验中存在目标与非目标之间遮挡的情况,YOLOv8n 算法出现误检现象,MBE-YOLO 算法能够有效检测出目标。

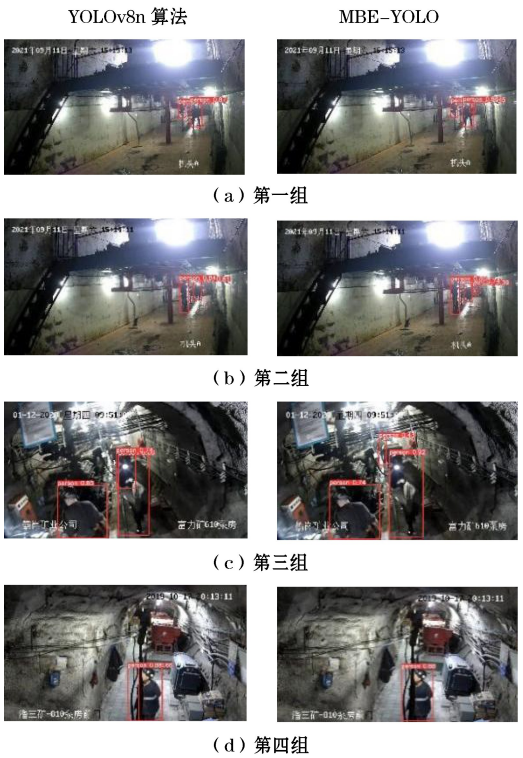


图 10 模型检测效果对比

Fig.10 Comparison of model detection results

3 结论

(1) 针对井下环境中由于设备计算能力受限,传统检测网络难以实现实时检测行人的问题,本文设计出了一种轻量化井下行人检测模型——MBE-YOLO 模型。该模型在 YOLOv8n 的基础上 *mAP*@0.5:0.95 提升了 0.2 个百分点,计算量下降了 44%,FPS 上升了 4.3 帧/s,参数量下降了 46%,模型大小减少了 43%,与 SSD、YOLOv3-tiny、YOLOv5n、YOLOv7-tiny 检测网络相比,在检测精度和模型轻量化程度方面都具有较大的优势。

(2) 引入改进颈部网络结构,简单高效地对多尺度信息进行融合,极大地减少了网络的参数和计算量;采用共享卷积的方式改进 YOLOv8 检测头,在几乎不影响检测精度的同时,解决了原始模型中检测头参数量巨大的问题;引入较为轻量的 MLCA 注意力机制,几乎不增加参数和计算量的同时,提高模型检测精度。

(3) MBE-YOLO 改进模型极具轻量化,适合嵌入便携式检测设备和硬件资源受限的矿井中,对于边缘设备的部署具有较大优势。由于模型的检测精度还有一定提升空间,未来的研究将聚焦于提高检测算法的性能,力求在不显著增加计算成本和模型复杂度的前提下,实现更高的检测精确度。

参考文献(References):

[1] 王国法,庞义辉,任怀伟,等.矿山智能化建设的挑战与思考[J].智能矿山,2022,3(10):2-15.
WANG Guofa, PANG Yihui, REN Huaiwei, et al. Challenges and reflections on the intelligent construction of mines[J].Journal of Intelligent Mine,2022,3(10):2-15.

[2] 谢斌红,袁帅,龚大立.基于 RDB-YOLOv4 的煤矿井下有遮挡行人检测[J].计算机工程与应用,2022,58(5):200-207.
XIE Binhong, YUAN Shuai, GONG Dali. Detection of blocked pedestrians based on RDB-YOLOv4 in coal mine[J]. Computer Engineering and Applications, 2022, 58(5): 200-207.

[3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]// Proceedings of the IEEE conference on computer vision and pattern recognition.San Diego:IEEE,2016:779-788.

[4] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu: IEEE, 2017:7263-7271.

[5] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv,2018:02767.

[6] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv,2020:10934.

[7] LIU Wei, ANGUELOV D, ERHAN D, et al. Ssd:Single shot multibox detector [J]. Springer International Publishing,2016,873:21-37.

[8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[J]. IEEE Computer Society,2014,81:580-587.

[9] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision.Sydney:IEEE, 2015:1440-1448.

[10] REN Shaoqing, HE Kaiming, 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, 2016, 39(6):1137-1149.

[11] 邹盛,周李兵,季亮,等.基于图像融合和改进 CornerNet-Squeeze 的煤矿井下行人检测方法[J].工矿自动化,2023,49(2):77-84.
ZOU Sheng, ZHOU Libing, JI Liang, et al. A pedestrian target detection method for underground coal mine based on image fusion and improved CornerNet-Squeeze[J]. Industry and Mine Automation,2023,49(2):77-84.

[12] WANG Jing, YANG Peng, LIU Yuansheng, et al. Research on improved yolov5 for low-light environment object detection[J]. Electronics,2023,12(14):3089.

[13] 周李兵,陈晓晶,贾文琪,等.用于井下行人检测的可见光和红外图像融合算法[J].工矿自动化,2023,49(9):73-83.
ZHOU Libing, CHEN Xiaojing, JIA Wenqi, et al. Visible and infrared image fusion algorithm for underground personnel detection[J]. Industry and Mine Automation, 2023,49(9):73-83.

[14] WAN Dahang, LU Rongsheng, SHEN Siyuan, et al. Mixed local channel attention for object detection[J]. Engineering Applications of Artificial Intelligence,2023,123:106442.

[15] TAN Mingxing, PANG Ruoming, LE Quoc V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Seattle:IEEE,2020:10778-10787.

[16] HAN Kai, WANG Yunhe, TIAN Qi, et al. Ghostnet:more features from cheap operations[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Seattle:IEEE,2020:1580-1589.

[17] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-cam ++: Generalized gradient-based visual explanations for deep convolutional networks[C]//2018 IEEE winter conference on applications of computer vision (WACV).California:IEEE,2018:839-847.