

•智能交叉科学与工程•

DOI:10.12454/j.jsuese.202400242



本刊网刊

基于改进 YOLOv8 的煤矿皮带异物检测方法

赵小虎^{1,2}, 张狄^{1,2*}, 谢礼逊^{1,2}, 孙维青^{1,2}, 张景怡^{1,2}, 尤星懿^{1,2}

(1. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221008; 2. 中国矿业大学 矿山互联网应用技术国家地方联合工程实验室, 江苏 徐州 221008)

摘要:针对在煤炭输送过程中存在的大尺度煤矸石、小尺度锚杆等造成异物划伤、撕裂皮带和堵塞落煤口等安全隐患,本文提出了一种基于改进 YOLOv8 的煤矿皮带异物检测方法 YOLOv8-SPCD。首先,根据已有矿井图像制作煤矿皮带异物数据集;接着,利用空间到深度卷积层(SPD-Conv)代替 Backbone 中的部分普通卷积层,将输入特征图的空间块重新排列进入通道维度以增加通道数,同时减小空间分辨率,在特征提取阶段保留更丰富的信息;然后,引入部分卷积(Pconv)改进原网络中的 C2f 模块,通过只在输入通道的一部分上应用卷积,减少关于冗余特征图的计算量,同时保证仍能提取输入图像的空间特征;之后,利用轻量级的跨尺度特征融合模块(CCFM)改进原模型(YOLOv8)的 Neck 部分,增强模型对于不同尺度对象的检测能力;最后,为了消除原损失函数惩罚项对收敛速度的影响并获得更快、更有效的回归结果,使模型在训练时快速收敛并精确定位皮带异物,引入改进后的 Inner-DIoU 函数对网络的边界框回归损失进行优化。通过设计消融实验,分析了本文模型的相关性能:参数量和 GFLOPs 分别缩小为基线网络的 40% 和约 59%, mAP@0.5 提升了 4.3 个百分点, mAP@0.5:0.95 提升了 4.1 个百分点,且图片检测的每秒帧数(FPS)也有少量提升,说明本文模型和原模型相比在轻量化的同时还提升了精度。与其他主流检测模型相比,本文模型的 mAP@0.5 最多提升了 18.6 个百分点, mAP@0.5:0.95 最多提升了 29.8 个百分点,验证了本文模型在煤矿皮带异物检测方面的有效性,为矿井下的边缘端部署提供了先决条件。

关键词: YOLOv8; 异物识别; SPD-Conv; 部分卷积; 跨尺度特征融合模块

中图分类号: TP391.41; TD528.1

文献标志码: A

文章编号: 2096-3246(2026)02-0023-12

煤矿井下皮带机在实际生产中的应用包括煤矿的生产、转运、加工等,在井下煤矿运输中起重要作用,是煤矿运输中的关键设备^[1]。正常传输煤流的运输皮带上可能出现大块煤矸石、锚杆等异物,大块煤矸石等异物如果堆积在落煤口可能出现堆煤、堵煤等问题,锚杆等异物可能会与运输皮带的部件勾连,划伤甚至撕裂皮带,严重影响正常的煤流运输,不仅会造成不可估计的经济损失,更可能对工作系统中其他设备的运行带来安全风险。如果在运输过程中对运输皮带的监控视频进行分析,对异物进行准确识别,就可以提前预警,及时安排针对性处理,解决安全隐患,有效保证煤矿井下皮带机的安全生产。

针对煤矿井下皮带运输异物检测问题,国内外开展了广泛的研究。传统方法是使用基于传统机器学习的方法实现目标检测。Alfarzaei 等^[2]通过煤矸石的

3D 图像提取视觉感知值来表示物体的体积,使用支持向量机(SVM)分类器接收视觉感知值和质量进行煤矸石分类。Shukla 等^[3]提出一种基于视觉显著性的多类 SVM 检测系统,实现了煤矿复杂环境下传送带上不同类型皮带损伤的高精度检测。随着深度学习的不断发展,目标检测技术由基于传统机器学习的方法逐渐过渡到基于深度学习的方法。传统机器学习技术主要使用特征提取和分类器来检测目标,但特征选择的过程往往需要人工参与,而且难以应对复杂和变化的目标。

基于深度学习的主流目标检测方法分为两阶段和单阶段两种。典型的两阶段网络模型有 R-CNN^[4]、SPP Net^[5]、Fast R-CNN^[6]、Faster R-CNN^[7]等,典型的单阶段网络模型有 YOLO^[8]系列、SSD^[9]等。一般情况下,两阶段网络模型在准确度上有优势,而单阶段网

收稿日期:2024-04-08 修回日期:2024-09-21 网络出版日期:2024-12-17

作者简介:赵小虎(1976—),男,教授,博士。研究方向:矿山物联网、计算机网络和智能计算。E-mail: xiaohuzhaozx@163.com

*通信作者:张狄, E-mail: ts22060208p31@cumt.edu.cn

络模型在速度上有优势。Wang 等^[10]使用 Res2Net101 模型替换原 Faster R-CNN 模型特征来提取模块中的 VGG16 结构,在水下目标检测中获得更好的性能。Wang 等^[11]采用深度可分离卷积和替换损失函数等方法改进 SSD 算法提高了异物检测精度,但其检测速度较慢。因需要在煤矿井下中低性能边缘端部署皮带异物检测模型,故多选择使用单阶段网络模型。

YOLO 系列检测模型,如 YOLOv3、YOLOv4、YOLOv5s、YOLOv7 等都曾被应用于煤矿皮带异物检测且取得了不错的效果:Zheng 等^[12]提出了一种融入注意力机制和多特征融合的煤矿安全隐患检测算法,利用 YOLOv3 框架,集成 Gc Net 注意力模块,建立了反向特征融合路径,并构建了 3 尺度预测模块,增强了该方法在复杂情况下的鲁棒性;Chen 等^[13]利用 KinD++ 低照度图像增强算法,通过特征处理对采集的低质量图像进行质量提升,结合优化锚框的 YOLOv4 算法,实现带式输送机异物的高效检测;Li 等^[14]针对传统目标检测算法模型复杂度高、训练困难、复杂条件下煤矸石识别定位效果差等问题,基于 YOLOv4 模型提出了一种基于轻量级混合域注意力的煤矸石识别定位方法,该算法具有较高的检测精度和效率,能够在复杂条件下正确识别和定位煤矸石,置信度高;张磊等^[15]通过在 YOLOv5s 的主干网络中添加压缩和激励模块,利用深度可分离卷积替换普通卷积,改善小目标煤矸石检测效果,提升了检测速度;毛清华等^[16]运用限制对比度自适应直方图均衡化方法对采集的带式输送机监控图像进行增强,提高图像中物体轮廓的清晰度,通过在 YOLOv7 模型主干提取网络引入轻量化无参注意力机制,提高了模型对图像复杂背景的抗干扰能力和对异物特征的提取能力。以上研究成果说明了 YOLO 系列检测模型在本领域应用的可行性,同时为本文提供了借鉴。

但目前基于计算机视觉的皮带异物检测方法还没有采用精度更好的 YOLOv8 模型,实现模型在本领域的针对性和轻量化改进。因此,本文提出了一种基于改进 YOLOv8 模型的煤矿皮带异物检测方法 YOLOv8-SPCD,改进工作如下:

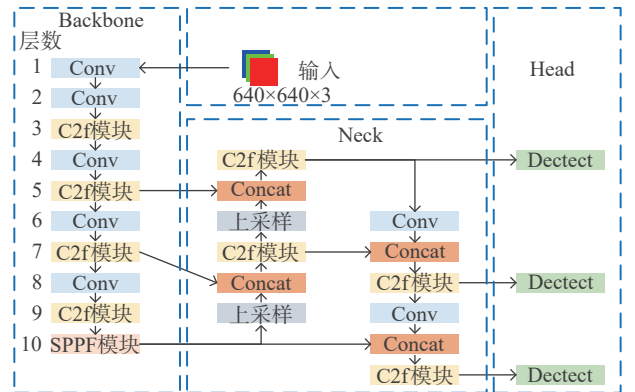
1)为了消除 CIoU Loss 损失函数中惩罚项对收敛速度的影响和获得更快、更有效的回归结果,结合 Inner 思想和 DIoU Loss 损失函数,改进 YOLOv8 模型的边界框回归损失函数,使本文模型在训练时能够更快地收敛并提升检测精度;

2)采用空间到深度卷积层 (SPD-Conv)、部分卷积 (PCConv)、轻量级的跨尺度特征融合模块 (module,

CCFM)对原模型进行改进,在实现轻量化的同时,提高对煤炭皮带异物检测的准确率,为矿井下的边缘端部署提供了先决条件。

1 YOLOv8网络介绍

YOLOv8 是 Ultralytics 公司发布的 YOLO 系列最新版本经典目标检测模型,和历史版本相比,模型检测速度更快、检测精度更高。其建立在 YOLO 系列历史版本的基础上,但引入了新的功能和改进点,以进一步提升性能和灵活性,已成为工业落地项目的主流选择^[17]。图 1 为 YOLOv8 网络结构。



注: Backbone 为主干网络, Neck 为颈部网络, Head 为头部网络, Conv 为卷积层, Concat 为拼接, Detect 为检测头, SPPF (spatial pyramid pooling-fast) 模块为快速空间金字塔池化模块。

图 1 YOLOv8 网络结构

Fig. 1 Network structure of YOLOv8

2 YOLOv8网络改进

本文对 YOLOv8 模型的改进主要有以下 4 点:

1)利用 SPD-Conv^[18]代替 Backbone 中的部分普通卷积层(第 2、4、6、8 层),在特征提取阶段保留更丰富的信息,从而提高模型对短小锚杆和低分辨率皮带异物的识别性能。

2)引入部分卷积改进原网络中的 C2f 模块,改进后的 C2f-PC 模块减少了对冗余特征图的计算,但仍能提取输入图像的空间特征,减少了原模型的参数和计算量。

3)利用轻量级的 CCFM 改进 Neck 部分,形成了 CF-Neck 结构,增强了模型对大尺度煤矸石的适应性和对小尺度锚杆等对象的检测能力,同时减少了模型参数。

4)为了获得更快、更有效的回归结果,使模型在训练时快速收敛并准确定位皮带异物,消除损失函数惩罚项对收敛速度的影响,引入改进后的 Inner-DIoU 函数对网络的边界框回归损失进行优化。

为便于描述,提取改进部分关键词英文首字母将改

进后的模型命名为YOLOv8-SPCD。图2为YOLOv8-SPCD网络结构。图2中, k 为卷积核尺寸(kernel size), s

为卷积核在特征上横滑动的步长(stride), p 为做卷积运算时在输入特征图四周补0的层数(padding)。

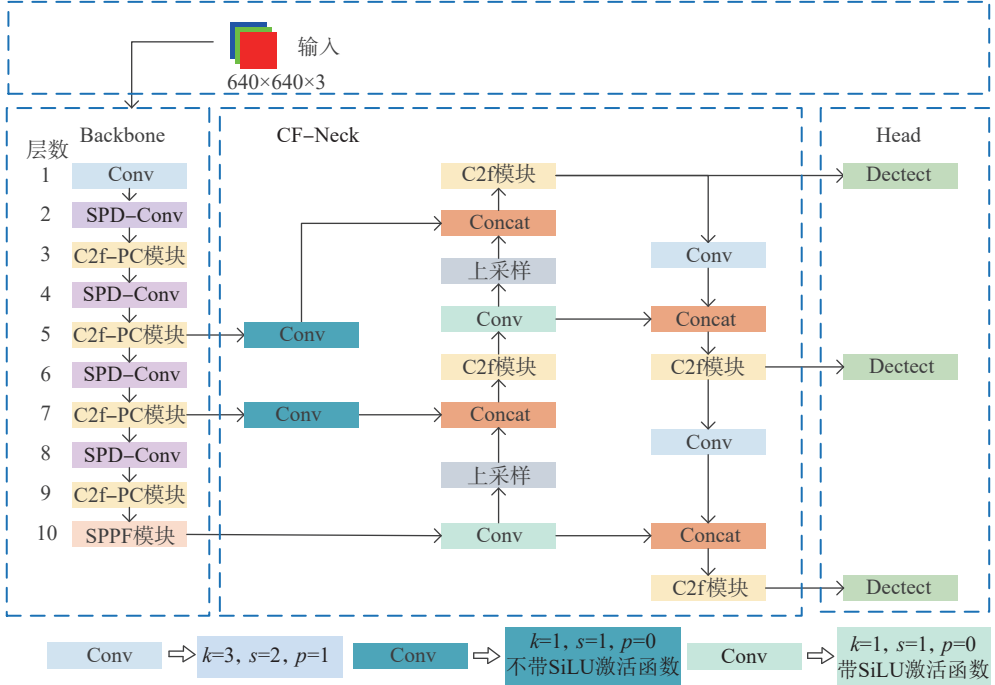


图2 YOLOv8-SPCD网络结构

Fig. 2 Network structure of YOLOv8-SPCD

2.1 SPD-Conv

SPD-Conv 模块由空间到深度(SPD)层和步长为1的普通卷积层组成。SPD层的作用是将输入特征图的空间块(像素块)重新排列进入深度(通道)维度,以增加通道数,同时减小空间分辨率,保留通道中的信息^[9]。普通卷积层执行步长为1的卷积操作,意味着在卷积过程中,卷积核会在输入特征图上逐像素移动,不跳过任何像素。这样可以确保在输入特征图的每个位置都能应用卷积核,最大限度地保留信息,并生成丰富的特征表示。

采用SPD-Conv可以在不丢失信息的情况下减小空间维度,在特征提取阶段保留更丰富的信息,从而提高模型对短小锚杆和低分辨率皮带异物的识别性能。根据Sunkara等^[18]的实验结果,采用比例因子scale(记为L)为2的SPD-Conv模型,scale为分割(Slice)操作的尺度,即特征图的行或列数能够整除的一个值。图3为SPD-Conv对YOLOv8 Backbone的改进。将任意原始特征图 $X \in \mathbb{R}^{S \times S \times C_1}$ (S和 C_1 分别为其边长和深度)分割为一系列子特征图:

$$\begin{aligned} f_{0,0} &= X[0:S:L, 0:S:L], \\ f_{1,0} &= X[1:S:L, 0:S:L], \dots, \\ f_{L-1,0} &= X[L-1:S:L, 0:S:L], \\ f_{0,1} &= X[0:S:L, 1:S:L], \\ f_{1,1} &= \dots, \\ &\vdots \end{aligned}$$

$$\begin{aligned} f_{L-1,1} &= X[L-1:S:L, 1:S:L], \\ &\vdots \\ f_{0,L-1} &= X[0:S:L, L-1:S:L], \\ f_{1,L-1} &= \dots, \\ f_{L-1,L-1} &= X[L-1:S:L, L-1:S:L]. \end{aligned}$$

其中, $X[\cdot]$ 为以指定参数执行Slice操作后的特征图。以 $f_{1,0} = X[1:S:L, 0:S:L]$ 为例, $f_{1,0}$ 为坐标(1,0)为起点的子特征图,1:S:L表示水平方向以特征图的横坐标1为开始、S为结束、L为步进进行下采样,0:S:L表示垂直方向以特征图的纵坐标0为开始、S为结束、L为步进进行下采样。

一般来说,给定任意 X ,子特征图 $f_{x,y}$ (x,y 分别为横、纵向起始坐标)由满足条件的特征点对应的特征图 $X(i,j)$ 组成, $X(\cdot)$ 为以括号内坐标为起点且保持不变的特征图, i 和 j 分别为分割前一点子特征图的横、纵坐标, $i+x$ 和 $j+y$ 都可以被L整除。因此,每个子特征图就是将 X 下采样一个比例因子得到的。图3、4均以L=2为例,得到4个大小为 $\frac{S}{2} \times \frac{S}{2} \times C_1$ 的子特征图 $f_{0,0}, f_{1,0}, f_{0,1}, f_{1,1}$ 。

沿通道维度将这些子特征图拼接起来,从而得到特征图 X' ,该特征图的空间维度变为原来的1/L,通道维度变为原来的 L^2 倍。也就是说,SPD层将特征图 $X \in \mathbb{R}^{S \times S \times C_1}$ 转化为 $X' \in \mathbb{R}^{\frac{S}{L} \times \frac{S}{L} \times L^2 C_1}$ 。

在 SPD 层进行特征变换后,在 $C_2 < L^2 C_1$ (C_2 为所需的输出特征图深度) 的条件下增加一个非跨步卷积层(步长为 1),并进一步将 $X' \in \mathbb{R}^{\frac{S}{L} \times \frac{S}{L} \times L^2 C_1}$ 变换为 $X'' \in \mathbb{R}^{\frac{S}{L} \times \frac{S}{L} \times C_2}$ 。使用非跨步卷积层(non-strided Conv)是为了尽可能多地保留所有的判别特征信息。如果使用步长为 3、核大小为 3×3 的卷积层,特征图虽然会被缩小,但每个像素只能被采样一次;而如果步长为 2,将发生非对称采样,即偶数和奇数行将在不同时间被采样。可见,步长大于 1 的跨步卷积层会导致信息的非判别性损失,尽管形式上也可将特征图 $X \in \mathbb{R}^{S \times S \times C_1}$ 转换为 $X'' \in \mathbb{R}^{\frac{S}{L} \times \frac{S}{L} \times C_2}$ 。

使用 SPD-Conv 对图 3 中标号 1~4 处的 Conv 模块进行替换,图 4 为 SPD-Conv 模块处理步骤。首先,来自上一层的输入特征图尺寸为 $320 \times 320 \times 64$,经过 Slice 操作后得到 4 个尺寸为 $160 \times 160 \times 64$ 的特征图;然后,将不同的特征图在通道维度上进行 Concat 操作,

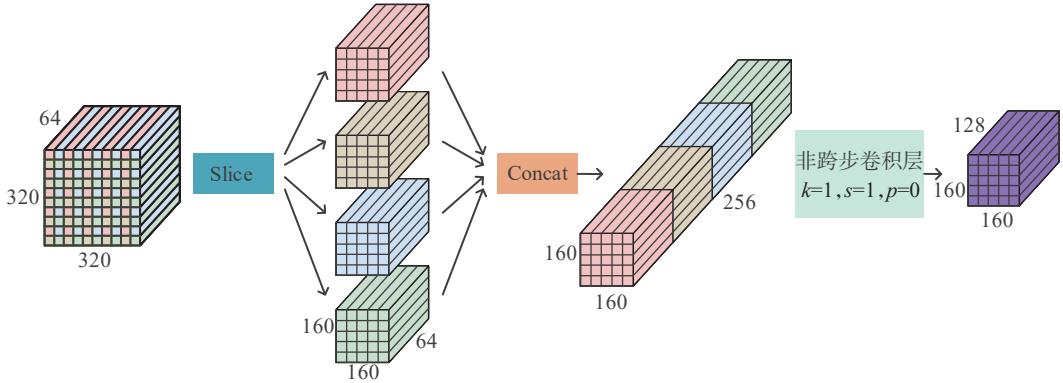


图 4 SPD-Conv 模块处理步骤
Fig. 4 Processing steps of the SPD-Conv module

2.2 C2f-PC

在卷积神经网络的特征图中,不同通道间存在大量相似或重复的信息,称为特征图冗余^[20]。在许多情况下,特征图的某些通道可能包含与其他通道高度相似的特征,这意味着在进行网络的前向传播时,对这些特征的多次处理并没有提供额外的有用信息,反而增加了计算量和内存访问的开销。在实际应用中,这种冗余可能导致计算资源的浪费,因为神经网络会在所有通道上执行卷积运算,包括那些冗余或不会对网络性能产生显著影响的通道。为了解决这个问题,本文引入 PConv^[21],图 5 为 PConv 设计原理。

内存访问量 M_{AC} 的计算方法如下:

$$r = \frac{c_p}{c} \tag{1}$$

$$F = h \cdot w \cdot k^2 \cdot c_p^2 \tag{2}$$

$$M_{AC} = h \cdot w \cdot 2c_p + k^2 \cdot c_p^2 \approx h \cdot w \cdot 2c_p \tag{3}$$

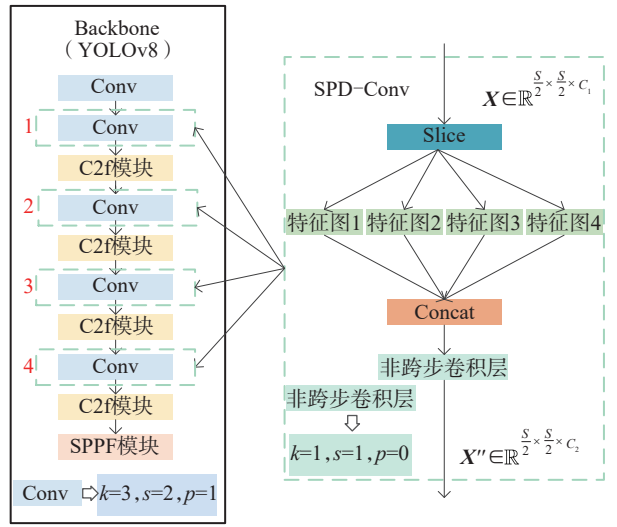


图 3 SPD-Conv 对 YOLOv8 Backbone 的改进
Fig. 3 Improvements of SPD-Conv to Backbone of YOLOv8

得到 $160 \times 160 \times 256$ 的特征图;最后,通过一个步长为 1 的非跨步卷积层得到输出特征图。

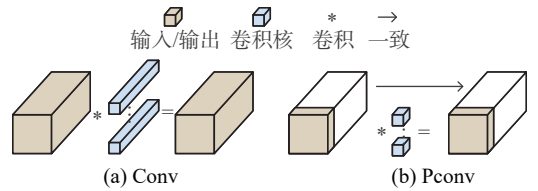


图 5 PConv 设计原理

Fig. 5 Design principle of PConv

式(1)~(3)中, h 、 w 、 c 分别为输入特征图的长、宽、通道数, c_p 为参与卷积的通道数, r 为部分卷积比例, F 为每秒浮点运算次数(FLOPs)。

采用 Pconv 时,只有通道数 c_p 参与空间特征提取,在后续特征通道信息未丢失的同时减少了计算量和内存访问量。一般典型的部分卷积比例 $r = 1/4$,则 PConv 的计算量仅为常规卷积的 $1/16$,内存访问量约为常规卷积的 $1/4$ 。

在实际应用中,只在输入通道的一部分应用卷积,减少了计算上的冗余和内存访问^[22],同时仍能有

效提取输入的大尺度煤矸石和小尺度锚杆图像的空间特征。将PConv结合到原网络的C2f模块,形成C2f-

PC模块。图6为C2f-PC模块结构。图6中, c_{in} 和 c_{out} 分别为本模块输入特征图和设定的输出特征图通道数。

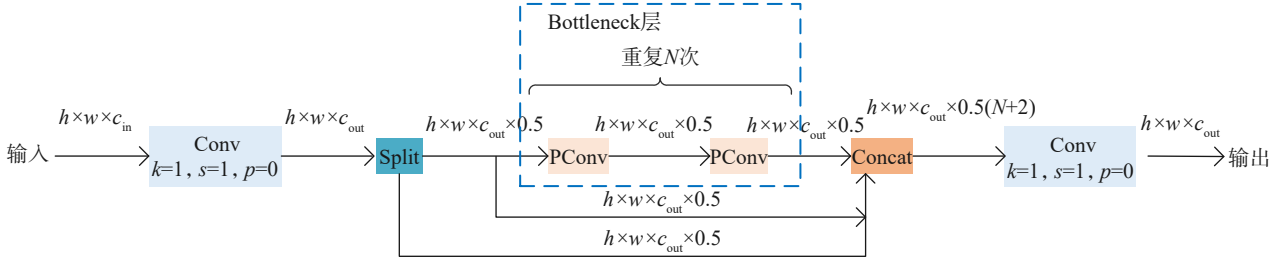


图6 C2f-PC模块结构

Fig. 6 Structure of C2F-PC module

2.3 CF-Neck

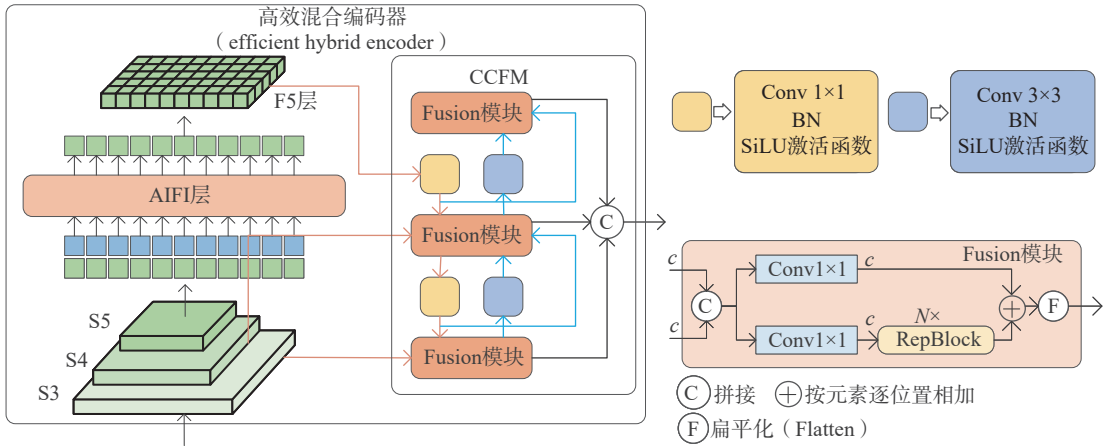
为了进一步轻量化模型大小和参数量,对YOLOv8模型的Neck部分进行改进。通过学习结构如图7所示的CCFM^[23],设计了CF-Neck(cross-scale feature-fusion neck)结构,如图2中CF-Neck部分所示。

参考图7中对F5层的处理,对YOLOv8模型中SPPF模块的输出特征图应用核大小为1x1卷积(Conv 1x1),在不改变输入尺寸的情况下,将输出特征图的通道数由原来的1024减少为256,在降低模型参数量的同时,实现通道间信息的线性组合变化,达到跨通道信息交互的目的。

参考图7中S3、S4层输入Fusion模块中进行核大

小为1x1的卷积降维处理,对YOLOv8模型主干网络中第4、6层的输出特征进行卷积降维,再与经处理后的其余尺度特征进行融合。以此实现的CF-Neck结构在极大减少模型参数的同时,还融合了不同尺度特征图像的信息,增强了模型对大尺度煤矸石的适应性。

通过拼接层将深层特征图上采样后与浅层特征图进行拼接,这种跨层连接不仅增加了模型的感受野,还使模型能够同时利用高分辨率的细节信息和低分辨率的语义信息。这对检测小尺度目标尤其重要,因为小尺度目标在高分辨率特征图中可能更容易被识别,所以提升了模型对锚杆等小尺度对象的检测能力。



注:AIFI (intrascale feature interaction)为跨尺度特征交互层,Fusion模块为特征融合模块,BN为批归一化,RepBlock (reparameterized block)为重参数化块,N为RepBlock个数。

图7 CCFM结构

Fig. 7 Structure of CCFM

2.4 Inner-DIoU

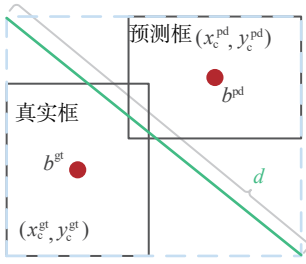
YOLOv8采用的边界框回归损失函数为CIoU损失函数(CIoU Loss)^[24],记为 \mathcal{L}_{CIoU} ,其表达式如下:

$$\mathcal{L}_{CIoU} = 1 - U + \frac{\rho^2(b^{pd}, b^{gt})}{d^2} + \alpha v \tag{4}$$

$$\alpha = \frac{v}{(1 - U) + v} \tag{5}$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^{pd}}{h^{pd}} \right)^2 \tag{6}$$

式(4)~(6)中: U 为交并比(intersection over union, IoU),表示预测框(predicted box)与真实框(ground truth box)的交集和并集面积的比值;函数 $\rho(\cdot)$ 用于计算两点间的欧氏距离; b^{pd} 和 b^{gt} 分别为预测框和真实框的中心点坐标(x_c^{pd}, y_c^{pd})和(x_c^{gt}, y_c^{gt}); d 为能够同时包含预测框和真实框的最小边界框对角线长度,如图8所示; α 为权衡参数; v 为衡量预测框与真实框长宽比一致性的惩罚项; w^{gt} 和 h^{gt} 分别为宽度和高度; w^{pd} 和 h^{pd} 分别为预测框的宽度和高度。

图 8 参数 d 示意Fig. 8 Schematic of parameter d

CIoU 损失函数在 DIoU 损失函数 (DIoU loss)^[24] 的基础上进一步添加了惩罚项 αv , 综合考虑了目标检测模型在检测目标位置和大小方面的准确性, 能更好地处理目标框间的重叠和错位, 适用于复杂背景等场景。但在 CIoU 损失函数的定义中, 衡量长宽比的 αv 惩罚项过于复杂, 从两个方面减缓了收敛速度^[25]:

1) v 仅反映了真实框与预测框纵横比的差异, 而不是 w^{gt} 与 w^{pd} 或 h^{gt} 与 h^{pd} 间的实际关系, 即当 $w^{\text{pd}} = Kw^{\text{gt}}$, $h^{\text{pd}} = Kh^{\text{gt}}$ (K 为任意实数) 时, $v = 0$, 与实际需求不符。

2) v 对 w^{pd} 和 h^{pd} 的偏导数分别为:

$$\frac{\partial v}{\partial w^{\text{pd}}} = \frac{8}{\pi^2} \left(\arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w^{\text{pd}}}{h^{\text{pd}}} \right) \times \frac{h^{\text{pd}}}{(w^{\text{pd}})^2 + (h^{\text{pd}})^2} \quad (7)$$

$$\frac{\partial v}{\partial h^{\text{pd}}} = -\frac{8}{\pi^2} \left(\arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w^{\text{pd}}}{h^{\text{pd}}} \right) \times \frac{w^{\text{pd}}}{(w^{\text{pd}})^2 + (h^{\text{pd}})^2} \quad (8)$$

由此可得 $\frac{\partial v}{\partial w^{\text{pd}}} = -\frac{h^{\text{pd}}}{w^{\text{pd}}} \cdot \frac{\partial v}{\partial h^{\text{pd}}}$, 可见 $\frac{\partial v}{\partial w^{\text{pd}}}$ 与 $\frac{\partial v}{\partial h^{\text{pd}}}$ 具有相反的符号, 因此在任何时候, 如果 w^{pd} 和 h^{pd} 中一个增加, 另一个将减少, 这是不合理的, 尤其当 $w^{\text{pd}} < w^{\text{gt}}$ 且 $h^{\text{pd}} < h^{\text{gt}}$ 或 $w^{\text{pd}} > w^{\text{gt}}$ 且 $h^{\text{pd}} > h^{\text{gt}}$ 时。所以 CIoU 损失函数存在一定的局限性。

为了消除惩罚项 αv 对收敛速度的影响, 选择将其去除, 采用 DIoU 损失函数。同时, 为了获得更快、更有效的回归结果, 结合 Inner^[26] 思想和 DIoU 损失函数, 引入 Inner-DIoU 作为本文模型的边界框回归损失函数。

Inner 思想的实质是引入一个尺度因子 ratio (记为 T), 其值一般在 $[0.5, 1.5]$ 区间内^[27]。指定 ratio 值后, 根据真实框和预测框的相关参数计算出覆盖真实框与预测框的最小边框, 即辅助框的参数, 相关计算如下:

$$b_l^{\text{gt}} = x_c^{\text{gt}} - \frac{w^{\text{gt}} \times T}{2}, b_r^{\text{gt}} = x_c^{\text{gt}} + \frac{w^{\text{gt}} \times T}{2} \quad (9)$$

$$b_t^{\text{gt}} = y_c^{\text{gt}} - \frac{h^{\text{gt}} \times T}{2}, b_b^{\text{gt}} = y_c^{\text{gt}} + \frac{h^{\text{gt}} \times T}{2} \quad (10)$$

$$b_l^{\text{pd}} = x_c^{\text{pd}} - \frac{w^{\text{pd}} \times T}{2}, b_r^{\text{pd}} = x_c^{\text{pd}} + \frac{w^{\text{pd}} \times T}{2} \quad (11)$$

$$b_t^{\text{pd}} = y_c^{\text{pd}} - \frac{h^{\text{pd}} \times T}{2}, b_b^{\text{pd}} = y_c^{\text{pd}} + \frac{h^{\text{pd}} \times T}{2} \quad (12)$$

$$S_{\text{intersection}} = (\min(b_r^{\text{gt}}, b_r^{\text{pd}}) - \max(b_l^{\text{gt}}, b_l^{\text{pd}})) \times (\min(b_b^{\text{gt}}, b_b^{\text{pd}}) - \max(b_t^{\text{gt}}, b_t^{\text{pd}})) \quad (13)$$

$$S_{\text{union}} = (w^{\text{gt}} \times h^{\text{gt}}) \times T^2 + (w^{\text{pd}} \times h^{\text{pd}}) \times T^2 - S_{\text{intersection}} \quad (14)$$

$$U^{\text{Inner}} = \frac{S_{\text{intersection}}}{S_{\text{union}}} \quad (15)$$

式 (9) ~ (15) 中, $b_l^{\text{gt}}, b_r^{\text{gt}}, b_t^{\text{gt}}, b_b^{\text{gt}}$ 分别为真实框的左 (left)、右 (right)、上 (top)、下 (bottom) 边界与坐标轴的交点值, $b_l^{\text{pd}}, b_r^{\text{pd}}, b_t^{\text{pd}}, b_b^{\text{pd}}$ 分别为预测框的左、右、上、下边界与坐标轴的交点值, $S_{\text{intersection}}$ 为真实框和预测框交集 (intersection) 的面积, S_{union} 为真实框和预测框并集的面积, U^{Inner} 为前二者的比值。

图 9 为真实框与预测框的内交集 (inner-intersection) 示意图。当 $T \in [0.5, 1.0]$ 时, 使用较小尺度的辅助框计算 IoU 损失, 能够加速高 IoU 样本回归, 达到加速收敛的效果^[28], 如图 9(a) 所示; 当 $T = 1.0$ 时, 改

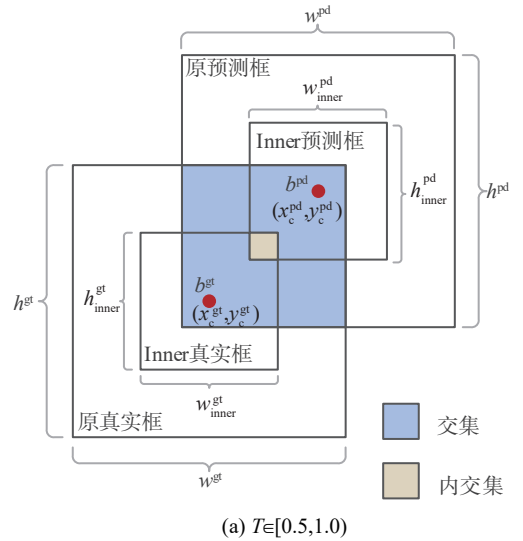
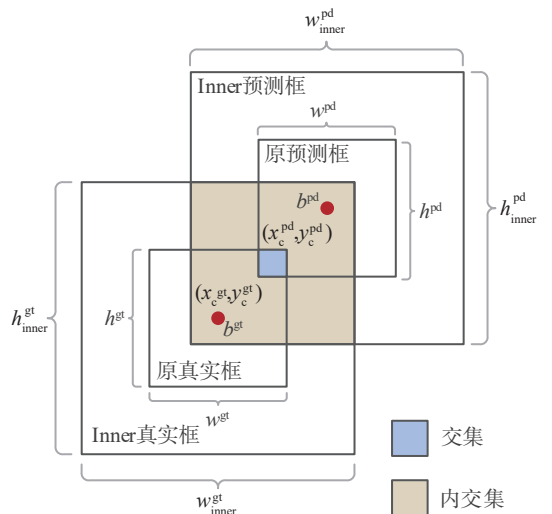
(a) $T \in [0.5, 1.0]$ (b) $T \in (1.0, 1.5]$

图 9 真实框与预测框内交集示意图

Fig. 9 Schematic diagram of inner-intersection between ground truth box and predicted box

进后的 Inner-DIoU 与原始的 IoU 相等;当 $T \in (1.0, 1.5]$ 时,使用较大尺度的辅助框计算 IoU 损失,能够加速低 IoU 样本回归,如图 9(b)所示。

为了直接最小化预测框与真实框间的归一化距离以实现更快的收敛速度,引入惩罚项 $\mathcal{R}_{\text{DIOU}}^{[24]}$:

$$\mathcal{R}_{\text{DIOU}} = \frac{\rho^2(b^{\text{pd}}, b^{\text{gt}})}{d^2} \quad (16)$$

结合 Inner 思想, Inner-DIoU 损失函数 $\mathcal{L}_{\text{Inner-DIoU}}$ 可定义为:

$$\mathcal{L}_{\text{Inner-DIoU}} = 1 - U^{\text{Inner}} + \mathcal{R}_{\text{DIOU}} \quad (17)$$

3 实验设计与结果分析

3.1 实验环境

实验使用的服务器配置和软件环境如下:训练模型平台硬件配置 CPU 为 Intel[®] Xeon[®] Gold 6430,显卡为 NVIDIA GeForce RTX 4090,显存为 24 GB;软件环境为 Ubuntu-22.04 系统,Python 3.8.10,Pytorch 2.0.0 框架,CUDA 11.8。训练参数如表 1 所示。

表 1 训练参数

Tab. 1 Training parameters

训练参数	数值
初始学习率	0.01
优化器	SGD(stochastic gradient descent)
优化器动量	0.937
优化器权重衰减系数	0.000 5
批大小	64
迭代周期数	250

3.2 数据集获取

本文制作数据集的图像来自公开的矿井皮带图像^[29],图像采集自矿井下皮带的运输环境,包括运输皮带上的大块煤矸石和锚杆两类图像,每类 1 600 张,包含训练图像 1 300 张和测试图像 300 张。其中,存在一些严重模糊和分类错误的图像,经筛选和标注后,共获得 2 755 张图像。图 10 为数据集图像示例。

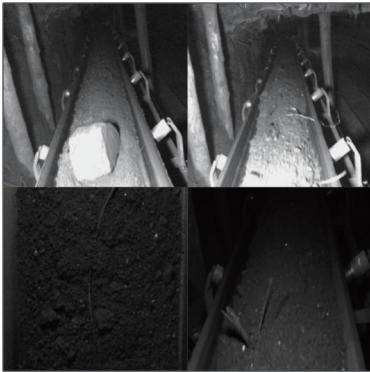


图 10 数据集图像示例

Fig. 10 Examples of the dataset image

本文使用 labelme 软件进行图像数据标注,将 2 755 张图像按 8:1:1 划分为训练集(train)、验证集(val)和测试集(test)。

3.3 评价指标

采用精确率(Precision, 记为 P)、召回率(Recall, 记为 R)、mAP@0.5、mAP@0.5:0.95 作为模型的评价指标。mAP@0.5 表示将预测框与真实边界框的交并比(IoU)设置为 0.5 时每个类别的平均精度(average precision, AP, 记为 P_{AP})的均值。mAP@0.5:0.95 表示 IoU 从 0.5 取到 0.95,间隔为 0.05 时 AP 的均值。 P 和 R 均在 IoU 阈值为 0.50 时统计。

精确率 P 表示在所有被预测为正的样本中实际为正样本的概率,计算方法如下:

$$P = \frac{Q_{\text{TP}}}{Q_{\text{TP}} + Q_{\text{FP}}} \quad (18)$$

式中: Q_{TP} 为真阳性(true positive)样本数,即预测为正样本,实际也为正样本的样本数; Q_{FP} 为假阳性(false positive)样本数,即预测为正样本,实际为负样本的样本数。

召回率 R 表示实际为正的样本中被预测为正样本的概率,计算方法如下:

$$R = \frac{Q_{\text{TP}}}{Q_{\text{TP}} + Q_{\text{FN}}} \quad (19)$$

式中, Q_{FN} 假阴性(false negative)样本数,即预测为负样本,实际为正样本的样本数。

平均精度 P_{AP} 通过计算每个类别的 P - R 曲线下的面积来获得,衡量了模型在各个阈值下的平均精度^[30]。 P - R 曲线通过在不同的阈值下绘制精度和召回率间的关系得到。平均精度的计算方法如下:

$$P_{\text{AP}} = \int_0^1 P(R) dR \quad (20)$$

均值平均精度(mean average precision, mAP, 记为 P_{mAP})是所有类别的平均精度,mAP 值越高,表示模型在不同阈值下的表现越好,计算方法如下:

$$P_{\text{mAP}} = \frac{1}{M} \sum_{i=1}^M P_{\text{AP},i} \quad (21)$$

式中, M 为类别总数,下标 i 为 P_{AP} 的类别序号。

每秒 10 亿次的浮点运算数(giga floating-point operations per second, GFLOPs)用于表示模型的计算量,是衡量模型计算复杂度的指标。每秒帧数(frames per second, FPS)表示每秒内可以检测的图片数量,是衡量模型检测速度的指标。设备和数据相同的情况下,FPS 值越大代表检测算法运行速度越快,也意味着目标检测模型算法复杂度越低。

3.4 消融实验

从两个方向设计消融实验:方向 1,在原有 YO-

LOv8 模型基础上增加改进模块验证其对基准模型的影响(消融实验中组别 0~8);方向 2,在最终改进模型 YOLOv8-SPCD 基础上,仅改变损失函数并验证其对最终改进模型的影响(消融实验中组别 8、9)。表 2 为消融实验结果。

选用 YOLOv8n 作为基线网络,其为 YOLOv8 系列中最小、最轻量化(通过大幅缩减网络深度与通道数实现)、速度最快、资源占用最少的版本(Nano)。

第 0 组实验为基线网络(YOLOv8n)的实验结果,后续实验组皆在基线网络上改进后得到。实验中的 FPS 值通过使用测试集测试 12 次,去除极大和极小值后取均值得到。

第 1~4 组实验是在基线网络上单独增加改进后

表 2 消融实验结果

Tab. 2 Results of ablation experiments

组别	SPD-Conv	C2f-PC	CF-Neck	Inner-DIoU	mAP@0.5/%	mAP@0.5:0.95/%	参数量/ 10^6	GFLOPs	FPS
0					89.4	59.9	30	8.1	140
1	√				90.6	61.3	27	7.6	135
2		√			91.1	60.5	25	6.8	137
3			√		90.3	60.6	19	6.6	138
4				√	91.3	62.7	30	8.1	151
5	√	√			91.0	62.0	23	6.3	135
6	√		√		91.4	62.3	17	6.1	141
7		√	√		92.5	63.8	15	5.3	130
8	√	√	√		93.3	62.5	12	4.8	139
9	√	√	√	√	93.7	64.0	12	4.8	147

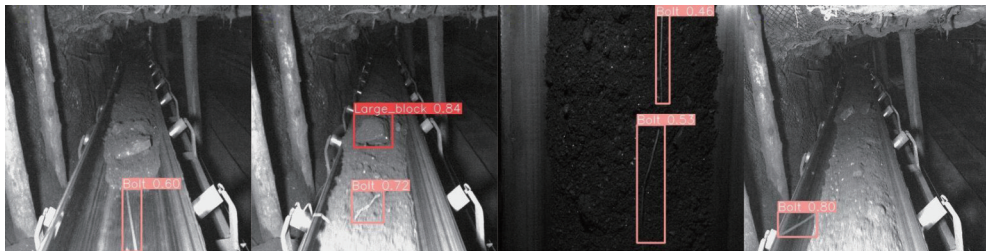
注:“√”表示采用了相应的改进模块。

为了直观地对比改进前后模型的性能,选出一部分测试图像进行对比,图 11 为 YOLOv8n 模型改进前后可视化预测结果。

模块的独立实验,由表 2 可见单一模块对基线网络的改进情况。第 3 组实验使用 CF-Neck 结构替换原 Neck 部分,在模型参数量降低约 37% 的情况下,仍能保持 mAP@0.5 和 mAP@0.5:0.95 值不降低,表明 CF-Neck 可以增强模型对于不同尺度对象的检测能力。第 4 组实验使用 Inner-DIoU 替换了 CIoU 损失函数,可以看到 mAP@0.5、FPS 等指标均有提升,说明 Inner-DIoU 损失函数能够很好地提高模型的拟合程度。

第 9 组实验即为本文模型。由表 2 可见,其参数量缩小为基线网络的 40%,GFLOPs 减小为原来的约 59%,mAP@0.5 提升 4.3 个百分点,mAP@0.5:0.95 提升了 4.1 个百分点,且 FPS 也有少量提升。

图 12 为改进前后模型训练图。设定当 50 轮训练后 mAP@0.5 未高于历史最高值或验证集的边界框损失(val/box_loss)小于历史最低值,即触发训练早停。



(a) YOLOv8n



(b) 本文模型

注:Large_Block 表示识别为大尺度煤矸石,Bolt 表示识别为小尺度锚杆,其后数字为置信度。

图 11 YOLOv8n 模型改进前后可视化预测结果

Fig. 11 Visual prediction results before and after the improvement of the YOLOv8n model

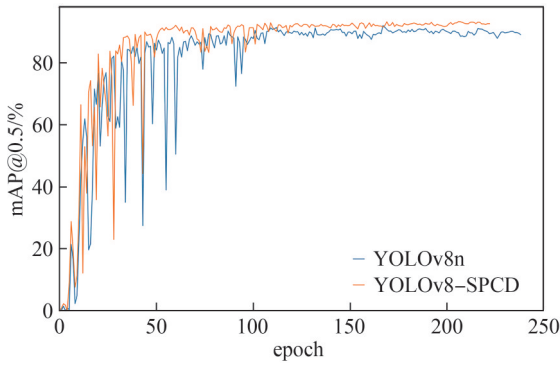


图 12 改进前后模型训练图

Fig. 12 Model training graph before and after improvements

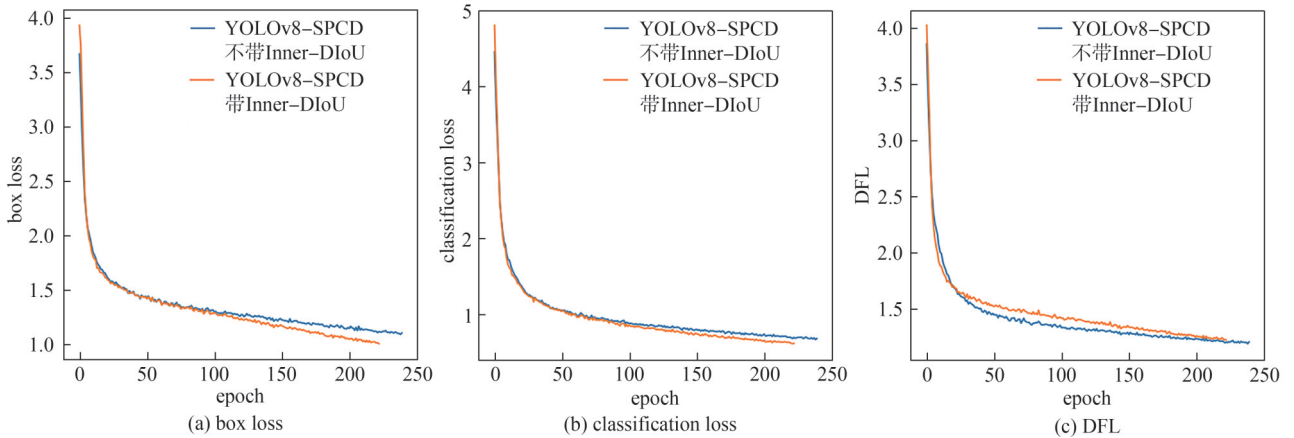


图 13 带有和不带有 Inner-DIoU 的 YOLOv8-SPCD 模型训练损失

Fig. 13 Training loss of the YOLOv8-SPCD model with and without Inner-DIoU

3.5 主流检测模型对比实验

将本文模型与 YOLOv3-tiny、YOLOv5n、YOLOv6n、SSD、Faster R-CNN 等主流目标检测模型进行实验对比^[31]。图 14 为本文模型与主流模型精度对比。实验中:SSD 模型的主干特征提取网络为 VGG,Faster

R-CNN 的主干特征提取网络为 Resnet50,两者都通过冻结训练 50 轮以加快模型的训练速度,随后解冻训练 250 轮,参与对比的 SSD 和 Faster R-CNN 模型数据皆为解冻训练数据;YOLOv3-tiny、YOLOv5n、YOLOv6n 模型的训练轮数设定为 250 轮。

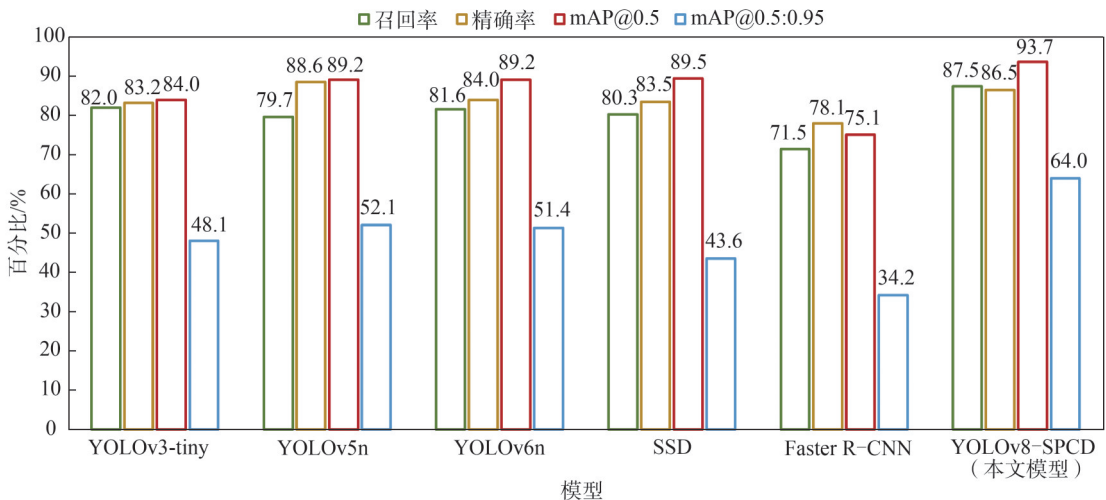


图 14 本文模型与主流模型精度对比

Fig. 14 Accuracy comparison between the proposed model and mainstream models

由图 14 可见, YOLOv8-SPCD 模型的召回率、平均精确度均大于其他模型,其 mAP@0.5 比 YOLOv3-

tiny、YOLOv5n、YOLOv6n、SSD、Faster R-CNN 分别高出 9.7、4.5、4.5、4.2、18.6 个百分点,说明本文模型和

现有其他检测模型相比具有明显优势。

图 15 为本文模型和主流模型训练图。实验中, YOLOv3-tiny、YOLOv6n、YOLOv8-SPCD 模型训练时,因最近 50 次训练后未有精度提升,故迭代轮次未到设定的 250 轮便停止训练。

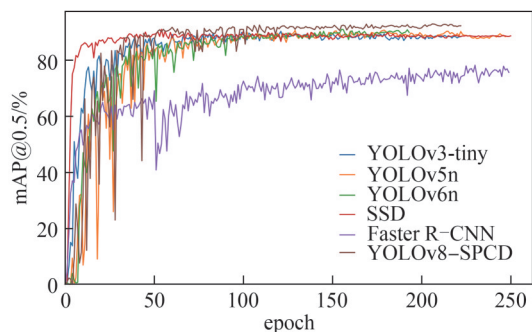


图 15 本文模型和主流模型训练图

Fig. 15 Training graph of the proposed model and dominant model

4 结 论

针对皮带在煤炭输送过程中存在大块煤矸石、锚杆等异物划伤、撕裂皮带和堵塞落煤口等安全隐患,本文提出了一种基于改进 YOLOv8 模型的煤矿皮带异物检测方法 YOLOv8-SPCD。本文模型参数量缩小为基线网络的 40%,GFLOPs 减小为原来的 59%,mAP@0.5 提升了 4.3 个百分点,mAP@0.5:0.95 提升了 4.1 个百分点,且 FPS 也有少量提升,验证了本方法在煤矿皮带异物检测方面的有效性,与 YOLOv3-tiny、YOLOv5n、YOLOv6n、SSD、Faster R-CNN 等主流目标检测模型相比具有一定优势,为矿井下的边缘端部署提供了先决条件。

未来的研究目标是将改进后的模型部署到矿井下的边缘端嵌入式设备中,实现算法的应用落地,并在这个过程中进一步优化模型。

参考文献:

[1] Xu Peng.Study on the key technology of foreign object detection of coal mine belt based on edge computing[D].Xuzhou: China University of Mining and Technology, 2021. [许鹏.基于边缘计算的煤矿井下皮带异物检测关键技术研究[D].徐州:中国矿业大学,2021.]

[2] Alfarzaei M S,Hu Eryi,Peng Wang,et al.Coal gangue classification based on the feature extraction of the volume visual perception ExM-SVM[J].Energies,2023,16(4): 2064.

[3] Shukla R K,Tiwari A K,Jha A K.An efficient approach of face detection and prediction of drowsiness using SVM[J].Mathematical Problems in Engineering,2023,2023:2168361.

[4] Girshick R,Donahue J,Darrell T,et al.Rich feature hierar-

chies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition.Columbus:IEEE,2014: 580-587.

- [5] He Kaiming,Zhang Xiangyu,Ren Shaoqing,et al.Spatial pyramid pooling in deep convolutional networks for visual recognition[C]//Computer Vision-ECCV 2014.Cham: Springer,2014:346-361.
- [6] Girshick R.Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision(ICCV).Santiago:IEEE,2015: 1440-1448.
- [7] 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,2017,39(6):1137-1149.
- [8] Redmon J,Divvala S,Girshick R,et al.You only look once: Unified,real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR).Las Vegas:IEEE,2016:779-788.
- [9] Liu Wei,Anguelov D,Erhan D,et al.SSD:Single shot Multi-Box detector[C]//Computer Vision-ECCV 2016.Cham: Springer International Publishing,2016:21-37.
- [10] Wang Hao,Xiao Nanfeng.Underwater object detection method based on improved faster RCNN[J].Applied Sciences,2023, 13(4):2746.
- [11] Wang Yuanbin,Wang Yujing,Dang Langfei.Video detection of foreign objects on the surface of belt conveyor underground coal mine based on improved SSD[J].Journal of Ambient Intelligence and Humanized Computing,2023, 14(5):5507-5516.
- [12] Zheng Junhui,Wang Deyong,Geng Zexun.Real-time detection of safety hazards in coal mines utilizing an enhanced YOLOv3 algorithm[J].Traitement Du Signal, 2023,40(4): 1565-1572.
- [13] Chen Yiming,Sun Xu,Xu Liang,et al.Application of YOLOv4 algorithm for foreign object detection on a belt conveyor in a low-illumination environment[J].Sensors,2022, 22(18):6851.
- [14] Li Deyong,Wang Guofa,Guo Yongcun,et al.An identification and positioning method for coal gangue based on lightweight mixed domain attention[J].International Journal of Coal Preparation and Utilization,2023,43(9):1542-1560.
- [15] Zhang Lei,Wang Haosheng,Lei Weiqiang,et al.Coal gangue target detection of belt conveyor based on YOLOv5s-SDE [J].Journal of Mine Automation,2023,49(4):106-112.[张磊,王浩盛,雷伟强,等.基于YOLOv5s-SDE的带式输送机煤矸目标检测[J].工矿自动化,2023,49(4):106-112.]
- [16] Mao Qinghua,Li Shikun,Hu Xin,et al.Foreign object rec-

- ognition of belt conveyor in coal mine based on improved YOLOv7[J]. *Journal of Mine Automation*,2022,48(12):26–32.[毛清华,李世坤,胡鑫,等.基于改进YOLOv7的煤矿带式输送机异物识别[J]. *工矿自动化*,2022,48(12):26–32.]
- [17] Luo Bingxin,Kou Ziming,Han Cong, et al.A “hardware-friendly” foreign object identification method for belt conveyors based on improved YOLOv8[J]. *Applied Sciences*, 2023,13(20):11464.
- [18] Sunkara R,Luo Tie.No more strided convolutions or pooling:A new CNN building block for low-resolution images and small objects[M]//*Machine Learning and Knowledge Discovery in Databases*. Cham: Springer Nature Switzerland,2023:443–459.
- [19] Peng Yanfei, Ji Yue.Road crack detection algorithm based on improved YOLOv8[C]//*Proceedings of the 2023 5th International Conference on Artificial Intelligence and Computer Applications(ICAICA)*. Dalian:IEEE,2024:28–32.
- [20] Wang Yuanyuan,Jiang Feilong,Li Yazhou, et al.Safety helmet detection algorithm for complex scenarios based on PConv-YOLOv8[C]//*Proceedings of the 2023 International Conference on the Cognitive Computing and Complex Data(ICCD)*.Huaian:IEEE,2024:90–94.
- [21] Chen Jierun,Kao S H,He Hao, et al.Run, don’t walk: Chasing higher FLOPS for faster neural networks[EB/OL].(2023–03–07)[2024–04–08].<https://arxiv.org/abs/2303.03667>.
- [22] Li Mao, Xiao Yangyi, Zong Wangyuan, et al. Detecting chestnuts using improved lightweight YOLOv8[J]. *Transactions of the Chinese Society of Agricultural Engineering*,2024,40(1):201–209.[李茂,肖洋轶,宗望远,等.基于改进YOLOv8模型的轻量化板栗果实识别方法[J]. *农业工程学报*,2024,40(1):201–209.]
- [23] Zhao Yian, Lv Wenyu, Xu Shangliang, et al. DETRs beat YOLOs on real-time object detection[EB/OL].(2023–04–17)[2024–04–08].<https://arxiv.org/abs/2304.08069>.
- [24] Zheng Zhaohui, Wang Ping, Liu Wei, et al. Distance – IoU loss:Faster and better learning for bounding box regression [J].*Proceedings of the AAAI Conference on Artificial Intelligence*,2020,34(7):12993–13000.
- [25] Zhang Yifan, Ren Weiqiang, Zhang Zhang, et al.Focal and efficient IOU loss for accurate bounding box regression[J]. *Neurocomputing*,2022,506:146–157.
- [26] Zhang Hao, Xu Cong, Zhang Shuaijie. Inner-IoU: More effective intersection over union loss with auxiliary bounding box[EB/OL].(2023–11–06)[2024–04–08].<https://arxiv.org/abs/2311.02877>.
- [27] Pan Wei, Wei Chao, Qian Chunyu, et al.Improved YOLOv8s model for small object detection from perspective of drones[J].*Computer Engineering and Applications*,2024, 60(9):142–150.[潘玮,韦超,钱春雨,等.面向无人机视角下小目标检测的YOLOv8s改进模型[J]. *计算机工程与应用*,2024,60(9):142–150.]
- [28] Tian Peng, Mao Li.Improved YOLOv8 object detection algorithm for traffic sign target[J]. *Computer Engineering and Applications*,2024,60(8):202–212.[田鹏,毛力.改进YOLOv8的道路交通标志目标检测算法[J]. *计算机工程与应用*,2024,60(8):202–212.]
- [29] Cheng Deqiang, Xu Jinyang, Kou Qiqi, et al. Lightweight network based on residual information for foreign body classification on coal conveyor belt[J]. *Journal of China Coal Society*,2022,47(3):1361–1369.[程德强,徐进洋,寇旗旗,等.融合残差信息轻量级网络的运煤皮带异物分类[J]. *煤炭学报*,2022,47(3):1361–1369.]
- [30] Wu Boyang, Mao Shengke, Lin Teyu, et al.Surface damage detection and identification method of wind turbine blades based on improved YOLOv8[J]. *Mechanical & Electrical Engineering Magazine*,2024,41(7):1260–1268.[吴博阳,毛胜轲,林特宇,等.基于改进YOLOv8的风电叶片表面损伤检测与识别方法[J]. *机电工程*,2024,41(7):1260–1268.]
- [31] Liu Qinghua, Yang Xinyi, Jie Hao, et al. Rice grain detection based on YOLO v7 fusing of GhostNetV2[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023,54(12):253–260.[刘庆华,杨欣仪,接浩,等.基于融合GhostNetV2的YOLO v7水稻籽粒检测[J]. *农业机械学报*,2023,54(12):253–260.]

A Foreign Body Detection Method for Coal Mine Belt Based on Improved YOLOv8

ZHAO Xiaohu^{1,2}, ZHANG Di^{1,2*}, XIE Lixun^{1,2}, SUN Weiqing^{1,2}, ZHANG Jingyi^{1,2}, YOU Xingyi^{1,2}

(1.School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221008, China;

2.The National Joint Engineering Laboratory of Internet Applied Technology of Mines, China University of Mining and Technology, Xuzhou 221008, China)

Abstract:

Objective On the transport belt used for normal coal flow, large coal gangue, anchor rods, and other foreign objects can be present. When large coal gangue or other foreign objects accumulate at the coal drop port, issues such as coal stacking and coal blockage occur. Anchor rods and other foreign objects can become entangled with transport belt components, causing surface scratches or even severe belt tearing, which seriously affects the normal coal flow transport. Deep learning methods previously applied demonstrate inferior baseline network performance compared to the YOLOv8 (You Only Look Once) model and fail to incorporate targeted lightweight optimization for edge deployment scenarios. Currently,

computer vision-based detection methods do not achieve performance improvements over the YOLOv8 model in coal mine target detection tasks. Therefore, this study proposes a foreign object detection method for coal mine conveyor belts, YOLOv8-SPCD, which is developed based on an improved YOLOv8 framework.

Methods The YOLOv8-SPCD model introduced several key improvements to enhance the detection performance of the original YOLOv8 model. First, the coal belt foreign body dataset was constructed based on existing mine images. The labelme tool was utilized to annotate the image data, and the images were divided into the training set (train), validation set (val), and test set (test) based on a ratio of 8:1:1. Then, SPD-Conv was utilized to replace the convolutional component in the Backbone, and the spatial blocks of the input feature map were rearranged into the channel dimension to increase the number of channels, reduce the spatial resolution, and retain richer information during the feature extraction stage. Next, partial convolution was introduced to improve the C2f structure in the original network. The computation of redundant feature maps was reduced, while the spatial features of the input images were still effectively extracted by applying convolution only to part of the input channels. Then, a lightweight cross-scale feature fusion module (CCFM), was utilized to improve the Neck component and enhance the detection capability of the model for objects at different scales. Finally, to eliminate the adverse effect of the penalty term in the original loss function on convergence speed and to obtain faster and more effective regression results, the improved Inner-DIoU function was introduced to optimize the bounding box regression loss of the network, enabling faster convergence and more accurate localization of belt foreign bodies during training.

Results and Discussions Groups 1 to 4 experiments were independent experiments in which the improved modules were modified separately on the baseline network, allowing the impact of each individual module on the baseline network to be clearly observed. In the third group of experiments, the CF-Neck structure was utilized to replace the original Neck component, and the mAP value remained unchanged even though the number of model parameters was reduced by 37%, indicating that CF-Neck enhanced the detection capability of the model for objects at different scales. In the fourth group of experiments, Inner-DIoU was utilized to replace the CIoU loss function, and the experimental indicators, such as mAP@0.5 and FPS, were improved, indicating that Inner-DIoU effectively enhanced the fitting performance of the model. The ninth group of experiments corresponded to the YOLOv8-SPCD model proposed in this study. The model weight was reduced to 43% of the baseline network, GFLOPs was reduced to 59% of the original value, mAP@0.5 was increased by 4.3 percentage points, mAP@0.5:0.95 was increased by 4.1 percentage points, and FPS was slightly improved. The effectiveness of the proposed method for detecting foreign objects on coal mine belts was thus verified. The training loss curves of the YOLOv8-SPCD model with Inner-DIoU and without Inner-DIoU were compared in this study, and the results showed that the convergence speed of the YOLOv8-SPCD model with Inner-DIoU was significantly faster than that of the model without Inner-DIoU. The Box Loss, which measured the discrepancy between the actual boundary box and the predicted boundary box of the target object, and the Classification Loss, which measured the accuracy of the model in predicting each target category, were both significantly reduced. The distribution focal loss (DFL), which was utilized to correct errors in predicting object boundary frames, remained similar to that before modification during training, indicating that the fitting performance of the proposed model on the mine image dataset was superior to that of the original model. The proposed model was also compared to mainstream target detection models such as YOLOv3-tiny, YOLOv5n, YOLOv6n, SSD, and Faster R-CNN. The comparison results showed that the proposed model exhibited clear advantages.

Conclusions The YOLOv8 model provides a feasible technical solution for detecting the presence of coal gangue, bolts, and other foreign matter during the coal conveying process on conveyor belts. The improved model integrates a series of enhancement strategies, including SPD-Conv, PConv, the CCFM, and the Inner idea, demonstrating the broad application potential of the YOLOv8 model in coal mine target detection. This work provides a prerequisite for deployment at the mine edge. Then, the research objective is to deploy the improved model on embedded equipment at the mine edge end, realize practical algorithm application, and further optimize the model during the deployment process.

Key words: YOLOv8; foreign body identification; SPD-Conv; partial convolution; cross-scale feature fusion module

(编辑 李轶楠)

引用格式: Zhao Xiaohu, Zhang Di, Xie Lixun, et al. A foreign body detection method for coal mine belt based on improved YOLOv8[J]. *Advanced Engineering Sciences*, 2026, 58(2): 23-34. [赵小虎, 张狄, 谢礼逊, 等. 基于改进 YOLOv8 的煤矿皮带异物检测方法[J]. *工程科学与技术*, 2026, 58(2): 23-34.]