基于YOLOv8-WBG的矿山井下作业钻头检测方法研究*

韩昊[1],蔡东升[2],黄琦[1,2,3],刘旭阳[2],张卓梵[3]

(1.成都理工大学 计算机与网络安全学院, 四川 成都 610059;
2.成都理工大学 核技术与自动化工程学院, 四川 成都 610059;
3.成都理工大学 能源学院, 四川 成都 610059)

摘要:矿山井下钻头检测是确保矿山井下作业安全的关键环节,但现有基于深度学习的检测模型在特征提取能力和检测效率上存在不足。为提升检测性能,提出了一种基于改进YOLOv8n的钻头检测模型YOLOv8-WGB。首先,采用WIoU损失函数替代了传统CIoU损失函数,强化了对关键特征的关注;其次,引入全局注意力机制(GAM),增强了YOLOv8骨干网络的全局信息交互能力,减少了信息损失并放大全局维度特征;最后,添加BiFPN结构,在融合不同尺度特征图时赋予其不同权重,以实现最优融合。试验结果表明,YOLOv8-WGB在精确率、召回率和*mAP*@0.5方面分别较YOLOv8n提高了7.1%、11.1%和8.3%,并且推理速度较YOLOv8s提高了173帧/s,表现出更快的检测速度。该模型在保持YOLOv8n模型复杂度和推理速度优势的基础上,检测精度显著优于YOLOv8s,有效平衡了检测的精度和速度,在矿山井下钻头检测任务中表现出较好的性能。

关键词:矿山井下;钻头检测;YOLOv8;WIoU;GAM;BiFPN

中图分类号:TD76 **文献标识码**:A **文章编号**:1005-2763(2025)12-0255-09

Research on the Drill Bit Detection Method for Underground Mine Operations Based on YOLOv8-WBG

HAN Hao[1], *CAI Dongsheng*[2], *HUANG Qi*[1,2,3], *LIU Xuyang*[2], *ZHANG Zhuofan*[3]

(1.College of Computer Science and Cyber Security, Chengdu University of Technology, Chengdu, Sichuan 610059, China;
2.College of Nuclear Technology and Automation Engineering,
Chengdu University of Technology, Chengdu, Sichuan 610059, China;
3.College of Energy, Chengdu University of Technology, Chengdu, Sichuan 610059, China)

Abstract: Underground mine drill bit detection is a critical link to ensure the safety of underground mine operations, but existing deep learning-based detection models have shortcomings in feature extraction capability and detection efficiency. To improve detection performance, a drill bit detection model named YOLOv8-WGB based on the improved YOLOv8n was proposed. Firstly, the traditional CIoU loss function was replaced with the WIoU loss function, which enhanced the focus on key features. Secondly, the global attention mechanism (GAM) was introduced, which enhanced the global information interaction capability of the YOLOv8 backbone network, reduced information loss, and amplified global dimensional features. Finally, the BiFPN structure was added, which assigned different weights to feature maps of different scales when fusing them, so as to achieve optimal feature fusion. The test results show that YOLOv8-WGB improves precision, recall, and *mAP*@0.5 by 7.1%, 11.1%, and 8.3% respectively compared with YOLOv8n, and its inference speed is 173 frames/s faster than that of YOLOv8s, exhibiting higher detection speed. On the basis of maintaining the model complexity and inference speed advantages of YOLOv8n, the proposed model has significantly superior detection accuracy to YOLOv8s. It effectively balances detection accuracy and speed, and exhibits excellent performance in the underground mine drill bit detection task.

Key words: Underground mine, Drill bit detection, YOLOv8, WIoU, GAM, BiFPN

* **收稿日期**:2024-12-10
基金项目:四川省自然科学基金项目(2024NSFSC0119)
作者简介:韩昊(2000—),男,河南开封人,硕士,主要从事图像检测方面的研究。E-mail:1986223321@qq.com
通信作者:蔡东升(1985—),男,安徽芜湖人,博士,教授,主要从事人工智能技术研究。E-mail:caidongsheng@cdut.edu.cn

0 引言

矿业的发展对于经济和就业都非常重要。矿业是国家的重要产业之一,为国家经济发展提供了重要的原材料和能源资源。现代科技是矿业行业发展中不可忽视的一环,通过现代化机械设备在矿井中进行开采和勘探,极大地提高了生产效率;但同发展一样重要的还有安全,矿井安全与机械设备运行状态密切相关,因此,钻头等设备的检测至关重要。

矿山井下具有环境复杂且人员较多的特点,在这种环境下,矿山井下钻头的识别与检测较为困难。目前,为了取代人工完成一系列井下检测工作,并达到或超过人工识别的准确性,矿山井下识别已被许多学者广泛研究,GAO 等[1]针对地下煤和脉石的识别任务,在特征提取阶段集成极化自注意力机制,对煤矸石目标的边缘信息进行优先级排序,通过结合加权双向特征金字塔网络的属性,高效融合多层煤矸石特征,实现对煤矸石目标的精确识别,但未能兼顾模型的速度。LEI 等[2]为了提升现有隧道面节理和裂隙识别的准确率,通过 Transformer 注意力机制和可变形卷积网络(Mask R-CNN-TD)关注重要区域,提高骨干网络提取图像特征的能力,使网络能够更精确地贴合隧道掌子面节理和裂缝的形态特征,进而提高检测的准确性,但未能考虑地下环境和兼顾模型的速度。矿山井下环境复杂,安全至关重要,矿山井下场景的检测任务对精度和及时性要求较高。双阶段的 R-CNN[3-4]系列方法精度高,但需经过两个阶段的计算,速度较慢,因此不适用矿山井下钻头作业场景的检测。

单阶段方法在一个阶段同时进行物体位置估计和类别分类,计算成本相对较低,速度较快,代表算法有 YOLO[5-6]系列和 SSD[7-8]系列。ZHANG 等[9]针对森林火灾目标体积小导致的远距离探测场景漏检问题,通过修改 YOLOv5 的骨干网络,进而减少 YOLOv5 的参数量,提升了模型的运行效率,但未能兼顾精度,导致精确率有所下降。LAWAL 等[10]将颈部网络替换为 PANet,引入了自顶向下和自底向上的双向路径,通过横向连接聚合不同层次的特征信息,提高特征的表示能力,但未能考虑网络计算量和复杂度,导致时间成本增加。LI 等[11]针对成熟草莓的实时检测存在困难等问题,通过在 YOLOv5 中引入自适应空间特征融合(ASFF)模块,使网络可以自适应地学习每个尺度上草莓特征图的融合空间权重,从而充分获取草莓的图像特征信息,提升了模型检测的精度,但多级草莓检测算法增加了模型复杂性,因此在计算复杂度上可能更高。尽管上述两种方法提升了目标检测算法的精度,但仍存在无法兼顾模型速度和检测性能的问题,因此在矿山井下钻头检测方面并不适用。

本文提出一种基于改进 YOLOv8n 的新型轻量算法 YOLOv8-WGB,首先将损失函数 CIoU 替换为 WIoU 损失函数[12],对数据集中不同类别设置不同的权重,化解数据集中类别不平衡的问题,更有效地处理各种尺寸和形状的目标,从而提高模型的鲁棒性。引入全局注意力机制(GAM)[13],通过直接计算全局注意力来捕获序列中的全局信息,使得模型能够更好地理解整个序列的语义和结构,有效地整合不同特征图之间的信息,提高模型的表现能力。进一步融合 BiFPN 结构[14],该结构结合了高层特征向低层特征传递和低层特征向高层特征传递的信息传递机制,实现了不同层级的特征融合,同时 BiFPN 采用加权特征融合机制,为不同区域分配不同的权重,更加关注在任务中的关键信息。相较于其他研究,本研究在特征提取、特征融合和检测速度方面有一定优势。这些改进使得在矿山井下钻头检测上具有更好的适用性。

1 YOLOv8 算法

近年来,在目标检测领域涌现出许多优秀的算法,YOLO 系列算法作为这一领域的领先算法,持续推动着技术的发展。YOLOv8 是 YOLO 系列中较为优秀的版本,它保留了前几代算法的优势,同时在此基础上不断创新和改进,发展出了自己的优势;YOLOv8 在速度方面进行了优化,相比于前几代算法,YOLOv8 在保持相似准确率的情况下,能够实现更快的推理速度。图 1 为 YOLOv8 结构图。在 C2f 结构中的瓶颈层(Bottleneck),通过先降维、再升维的方式减少计算量,并保持特征的表达能力,捷径连接(Shortcut)根据值的真假,决定是否采用捷径连接的结构,用于缓解梯度消失问题,并加速模型的训练;此外,生成了 3 个检测头(Detect),基于特征图得出目标的最终预测。

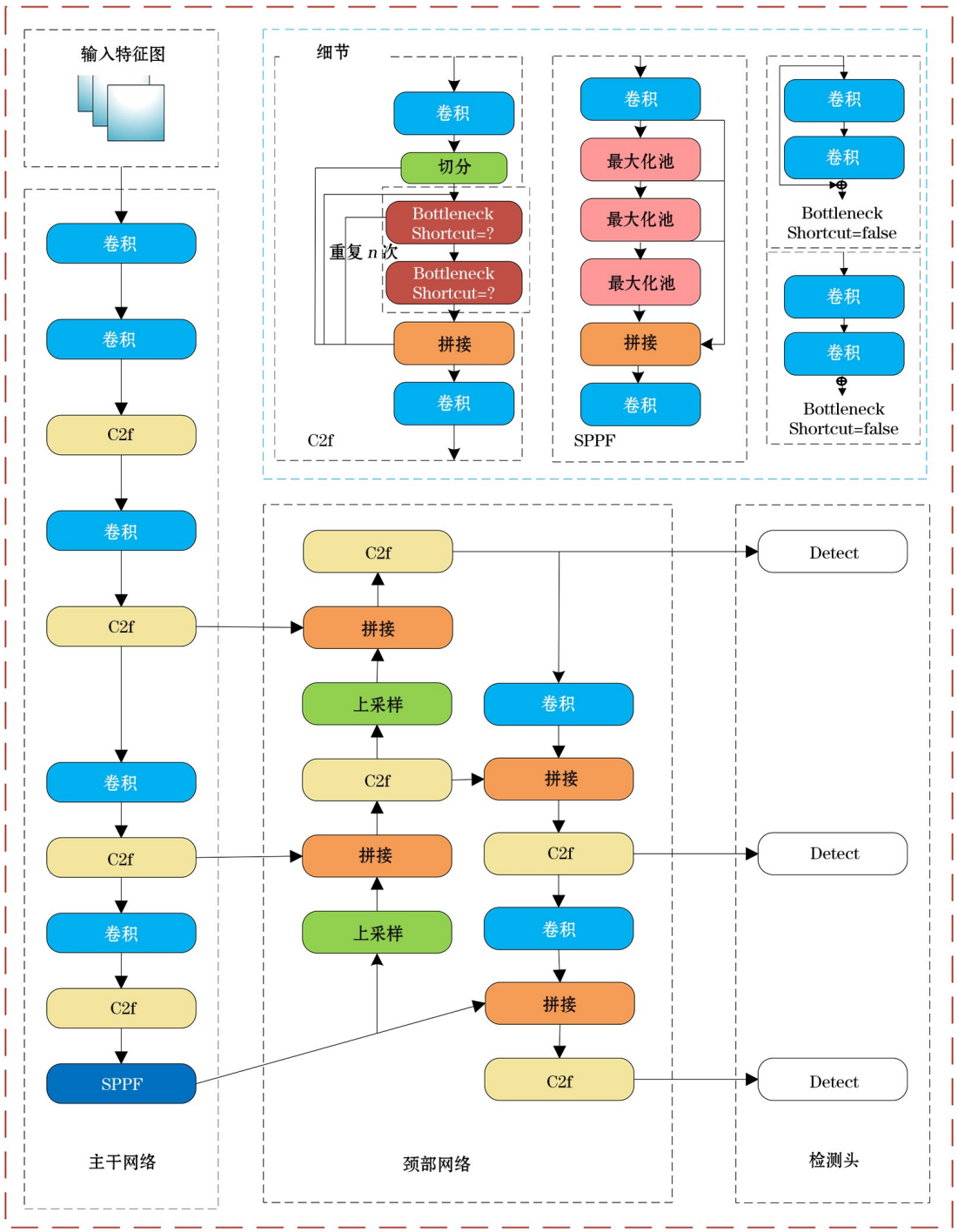


图 1　YOLOv8 的结构

Fig.1　Structure of YOLOv8

2　算法改进

2.1　损失函数

回归损失函数是目标检测损失函数的重要组成部分,对于提升目标检测性能至关重要。传统 IoU 损失仅以检测框与真实目标框的交并比为计算依据,而 YOLOv8 默认采用 CIoU 对此进行了改进,在几何方面引入了更多信息。

然而,训练数据无法确保全部为高质量数据,难以避免地包含低质量示例。大多数现有工作都假设训练数据中的示例是高质量数据,并且更注重提升损失函数的拟合能力。然而,如果盲目地

针对低质量样本强化损失函数,将会对低质量数据施加更高的权重,加强低质量样本对模型的影响,导致模型未能更多关注关键特征,从而导致模型性能下降。CIoU 及其他传统损失函数并没有考虑此类情况,也没有避免此情况对模型的负面影响。

WIoU 损失函数专注于解决目标检测任务中类别不平衡的问题,通过给高质量与低质量数据类别赋予不同的权重来解决这个问题,从而更有效地平衡各个类别之间的影响。

WIoU 首先计算 IoU 来评估两个框的重叠程度,然后通过对预测框和真实目标框各部分取最小权重后再进行加权,从而量化各部分的贡献。最后,WIoU 通过利用两个框各部分的最大权重来对加权后的 IoU 值进行归一化。

WIoU 损失函数的计算公式如下:

$$L_{\mathrm{WIoUv1}}=R_{\mathrm{WIoU}}L_{\mathrm{IoU}} \tag{1}$$

其中,R_{WIoU}和 L_{IoU}的计算公式如下:

$$\begin{cases}R_{\mathrm{WIoU}}=\exp\left(\dfrac{(x-x_{\mathrm{gt}})^2+(y-y_{\mathrm{gt}})^2}{(W_{\mathrm{gt}}^2+H_{\mathrm{gt}}^2)^*}\right)\\ L_{\mathrm{IoU}}=1-IoU\end{cases} \tag{2}$$

式中:x 和 y 表示预测框的中心坐标;x_{gt}和 y_{gt}表示真实框的中心坐标;W_{gt}和 H_{gt}分别表示真实框的宽度和高度。

2.2 注意力机制

为了解决深度学习模型在处理序列数据时所面临的一些挑战,让模型更关注某些特征,以及提高模型性能和效率,学者们引入了各种注意力机制,用于增加模型对关键特征的关注度。在过往的研究中,空间注意力机制注重不同空间位置特征的关联性,而通道注意力机制则关注特征在不同通道之间的关联性,以上两种注意力机制各自侧重单一维度(空间或通道)的交互性,忽视了特征在通道和空间两个维度之间的联系和交互。

GAM 采用顺序的通道-空间注意力机制,并对其子模块进行了重新设计,以减少信息损失并增强全局交互,从而提升深度神经网络的性能。其中,在空间注意力子模块中,使用了两个卷积层进行信息融合,通过分组卷积避免参数数量的大幅增加,以更有效地聚焦空间信息,并去除可能造成信息损失的最大池化操作。GAM 原理结构如图 2 所示。

对于一个给定的特征,中间状态特征图 F_2 和输出特征图 F_3 的定义如下:

$$\begin{cases}F_2=M_{\mathrm{c}}(F_1)\otimes F_1\\ F_3=M_{\mathrm{s}}(F_2)\otimes F_2\end{cases} \tag{3}$$

式中:M_{c}和 M_{s}分别是通道和空间的注意图;⊗表示按元素计算的乘法;F_1 为输入特征图。

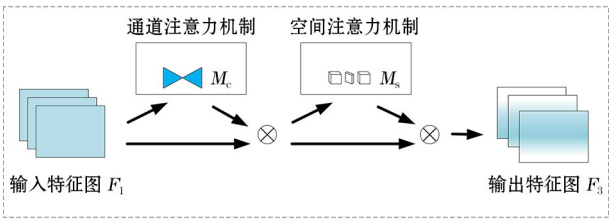
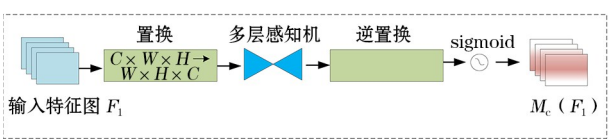


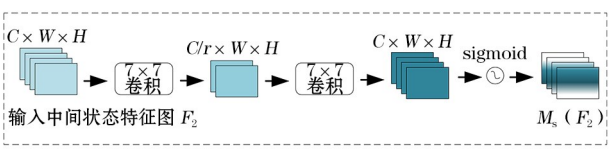
图 2 GAM 原理结构

Fig.2 Principle structure of GAM

GAM 包含两个子模块:通道注意力机制(CAM)和空间注意力机制(SAM)。CAM 和 SAM 的原理结构如图 3 所示。在 CAM 中,首先对输入特征图进行维度转换,然后将转换后的特征图送入多层感知器(MLP),最后再转换回原始维度并经过 Sigmoid 处理得到输出特征图。在 SAM 中,首先通过一个 7×7 的卷积核减少输入特征图的通道数量,从而降低计算量。接着,再经过另一个 7×7 的卷积操作,保持通道数与原输入一致,最后通过 Sigmoid 函数生成输出特征图。其中,C 表示通道数;W 表示特征图的宽度;H 表示特征图的高度;Sigmoid 为一种常用的激活函数,用于计算注意力权重,将权重限制在(0,1)之间;r 表示通道压缩比。



(a) CAM 模块



(b) SAM 模块

图 3 CAM 和 SAM 原理结构

Fig.3 Principle structure of CAM and SAM

2.3 特征融合模块

在神经网络中,信息传播方式非常重要。传统的特征金字塔网络 FPN[15] 采用自顶向下的传播模式,将高层特征传递到低层。然而,这种方法使得底层特征无法向高层反向传递,同时 FPN 顶部信息逐层传递的计算量较大。为了解决这一问题,学者提出了改良的 FPN——PANet[16]。PANet 通过引入自底向上的路径,为底层信息向高层的传递提供了通道,并且为了减少特征图数量、降低目标检测时的计算负担,PANet 还执行自下向上的特征融合。然而,无论是 FPN 还是 PANet,它们在特征融合过程中都对不同尺度的特征图进行相同权重的融合处理,这就导致未能有效考虑输入特征图的解析力差异,而这又会对输出特征图的均衡性造成影响。而 BiFPN 根据输入特征的重要性分配权重,使模型能够更关注关键特征,从而提升模型性能。同时,它还具有高效的双向跨尺度连接功能,即同时包含自顶向下和自底向上的通道结构,通过双向通道连接特征,在尽可能不损失深层语义信息的同时,将特征提取网络输出的特征与自下而上通道中的特征直接融合,并反复采用这种加强特征融合的结构,在保留较浅语义信息的同时,提高模型的效率和性能。FPN、PANet 和 BiFPN 的结构对比如图 4 所示。

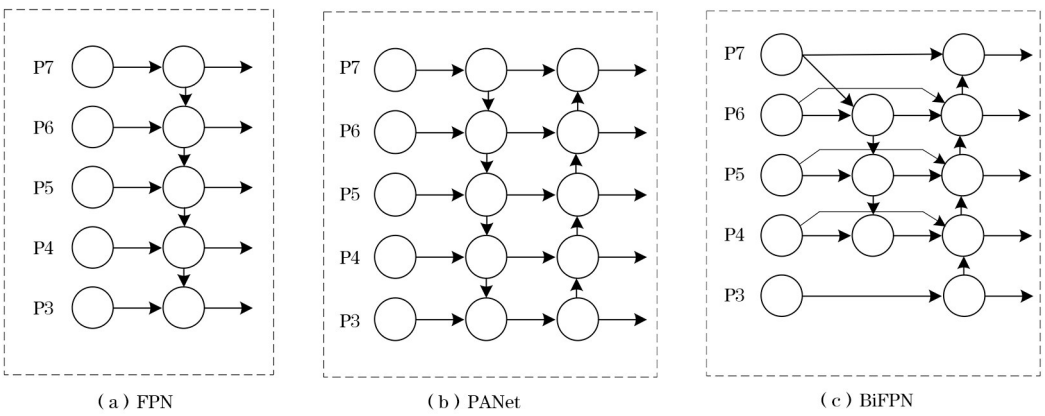


图 4 FPN、PANet 和 BiFPN 的结构对比

Fig.4 Structure comparison of FPN, PANet and BiFPN

2.4 YOLOv8-WBG 算法

首先将 YOLOv8 默认 CIoU 损失函数更换为 WIoU 损失函数,并引入权重因子。进一步在 YOLOv8 的特征提取网络(backbone)中增加 GAM 注意力机制,使模型捕捉更全面、更复杂的序列结构和语义信息。进一步优化特征融合模块,引入 BiFPN 网络进行优化,实现了双向跨尺度连接和加权特征融合[17]。通过以上 3 种方法的改进,有效地整合了不同尺度的特征,提升了模型的检测性能,在准确性和效率之间达到了更好的平衡。YOLOv8-WBG 的结构如图 5 所示。

3 试验与分析

3.1 试验数据

试验数据为来源于真实矿山井下作业场景的图像,试验训练集包含 2 599 张图片,验证集包含 626 张图片,测试集包含 1 375 张图片。所有模型预设训练 200 轮次。对于该数据集,考虑到样本角度单一化,在图像中的特征位置较为固定,可能导致模型在训练过程中过度依赖数据的单一特征,故在本文中对样本采用翻转、亮度调整和镜像反转等数据增强手段[18],通过在样本中引入这些变化,模型能够更好地适应各种实际场景,从而提升其性能和泛化能力。试验软硬件环境配置见表 1。数据增强示例如图 6 所示。

表 1 试验软硬件环境配置

Table 1 Hardware and software environment configuration for test

配置名称	版本/参数
操作系统	Windows 11
CPU	i5-12400f
GPU	NVIDIA GeForce RTX 4 060
内存	16G
Python	3.8
深度学习框架	PyTorch

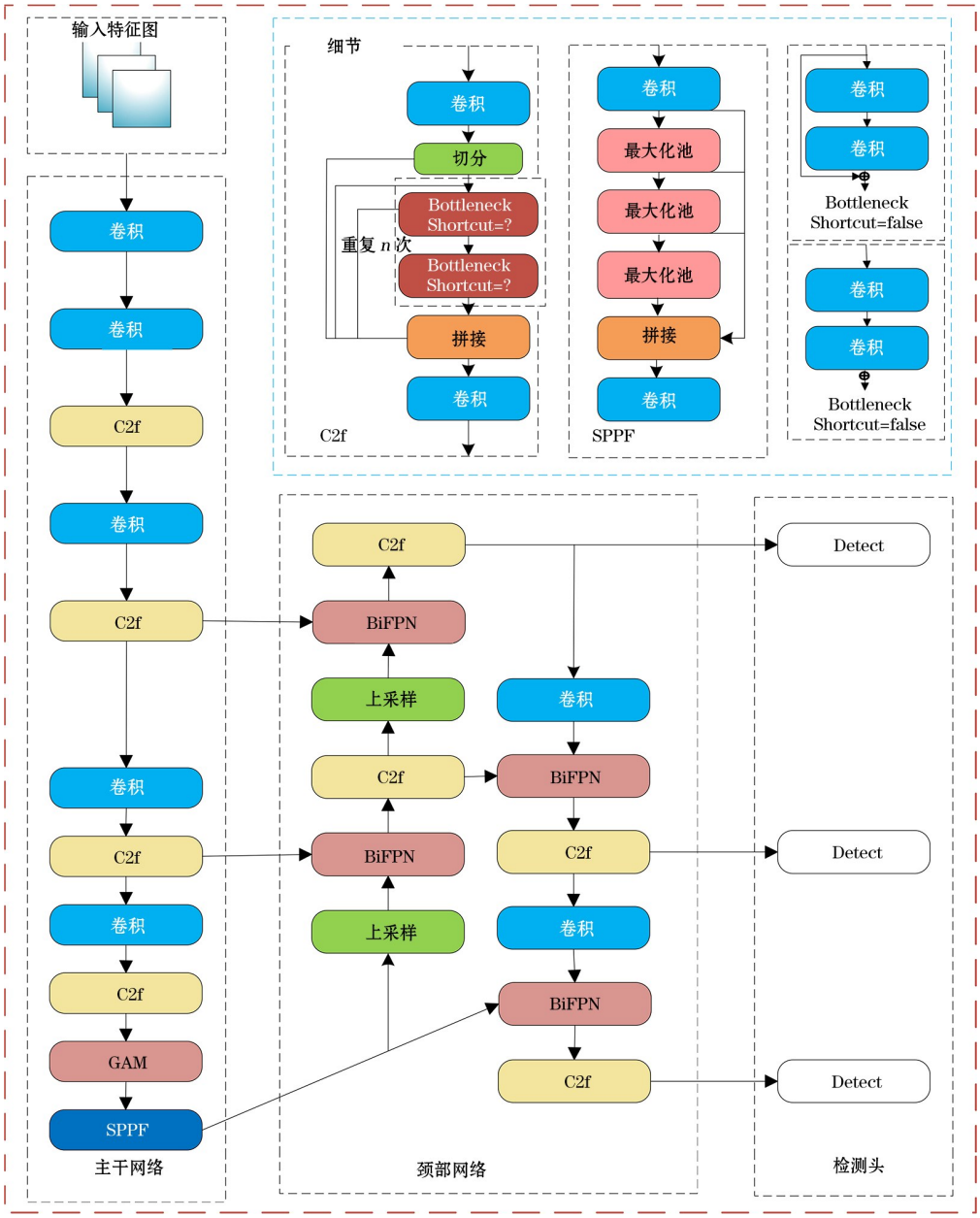
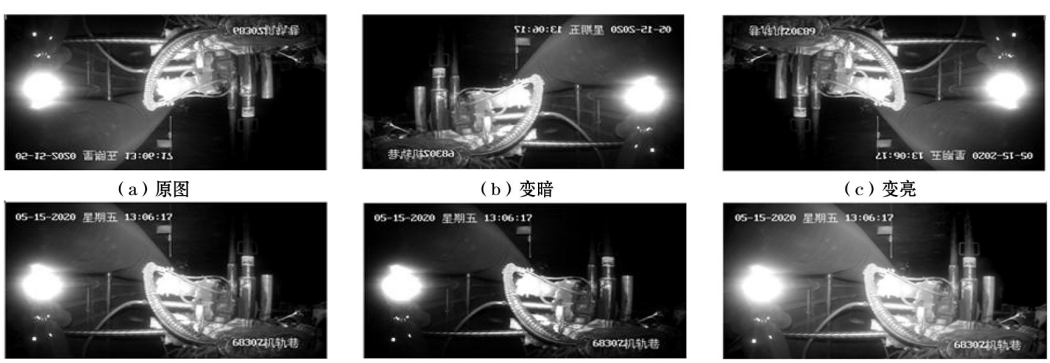


图 5 YOLOv8-WBG 结构

Fig.5 Structure of YOLOv8-WBG



(a) 原图 (b) 变暗 (c) 变亮

(d) 垂直镜像 (e) 水平镜像 (f) 上下翻转

图 6 数据增强示例

Fig.6 Examples of data enhancement

3.2 评价指标

为了验证算法改进的合理性和优越性,该试验采用精确率(P)、召回率(R)、模型复杂度、$mAP@0.5$、推理时间(PIP)和每秒帧数(FPS)等 6 个参数作为评价指标。

P 和 R 的计算公式如下:

$$\begin{cases} P = \dfrac{TP}{TP + FP} \\ R = \dfrac{TP}{TP + FN} \end{cases} \tag{4}$$

式中:TP 代表成功检测的目标数量;FP 代表错误检测的目标数量;FN 代表未被检测出的正确目标数量。

$mAP@0.5$ 表示在所有类别中,IoU 阈值设定为 0.5 时的平均检测精度。平均精度均值(mAP)的计算公式如下:

$$mAP = \frac{\sum_{i=1}^{k} AP_i}{k} \tag{5}$$

其中,AP 的计算公式如下:

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i) P_{\text{inter}} (r_i + 1)^k \tag{6}$$

式中:r_1, r_2, …, r_n 是按升序排列的 P 插值段第一个插值点对应的召回值;k 代表数据集中目标类别的数量。

PIP 包括预处理(P_1)、推理(I)和后处理(P_2)三部分,PIP 的计算公式如下:

$$PIP = P_1 + I + P_2 \tag{7}$$

FPS 是用来评估模型推理计算速度的重要指标,反映了模型在 1 s 内能够处理的图像数量,FPS 值的高低反映模型处理图像的速度快慢。FPS 的计算公式如下:

$$FPS = \frac{1}{PIP} \tag{8}$$

3.3 模块消融试验

模块消融参数对比见表 2,其中"√"代表改进该模块,"×"代表对该模块不做改进。试验数据表明,相较于原始 YOLOv8n,当 YOLOv8n 分别改进损失函数、注意力机制和特征融合模块时,模型的准确率分别是下降、上升和下降,而召回率正好相反,分别为上升、下降和上升,很难将一个模型同时保证"找得准"和"找的全"。通过进一步的两两消融试验,即对注意力机制+特征融合模块进行组合改进的试验,该组合改进显著提高了模型的准确率,但同时召回率出现了明显下降,模型"找得准",却"找不全"。这表明当设备发生故障时,漏检将会大幅度增加,而每次漏检都可能造成安全事故,这是不可取的。另外两组(损失函数+注意力机制、损失函数+特征融合模块)则实现了准确率提升,同时召回率同步提高。

经过三个方面改进之后获得模型 YOLOv8-WGB。在模型性能方面,YOLOv8-WGB 在钻头检测任务中,P、R 和 $mAP@0.5$ 指标均取得较为理想的提升。相较于原模型,YOLOv8-WGB 的 P、R 和 $mAP@0.5$ 分别提高了 7.1%、11.1%和 8.3%。

在模型推理速度和复杂度方面,YOLOv8-WGB 在钻头检测任务中的 PIP 增加了 0.2 ms,稍高于原模型 YOLOv8n,FPS 减少了 38 帧/s,稍低于原模型 YOLOv8n,模型复杂度增加了 0.6 GFLOPs。综合来看,相较于 YOLOv8n,YOLOv8-WGB 的推理时间和模型复杂度基本不变,准确率和召回率实现全面提升。

表 2 模块消融参数对比

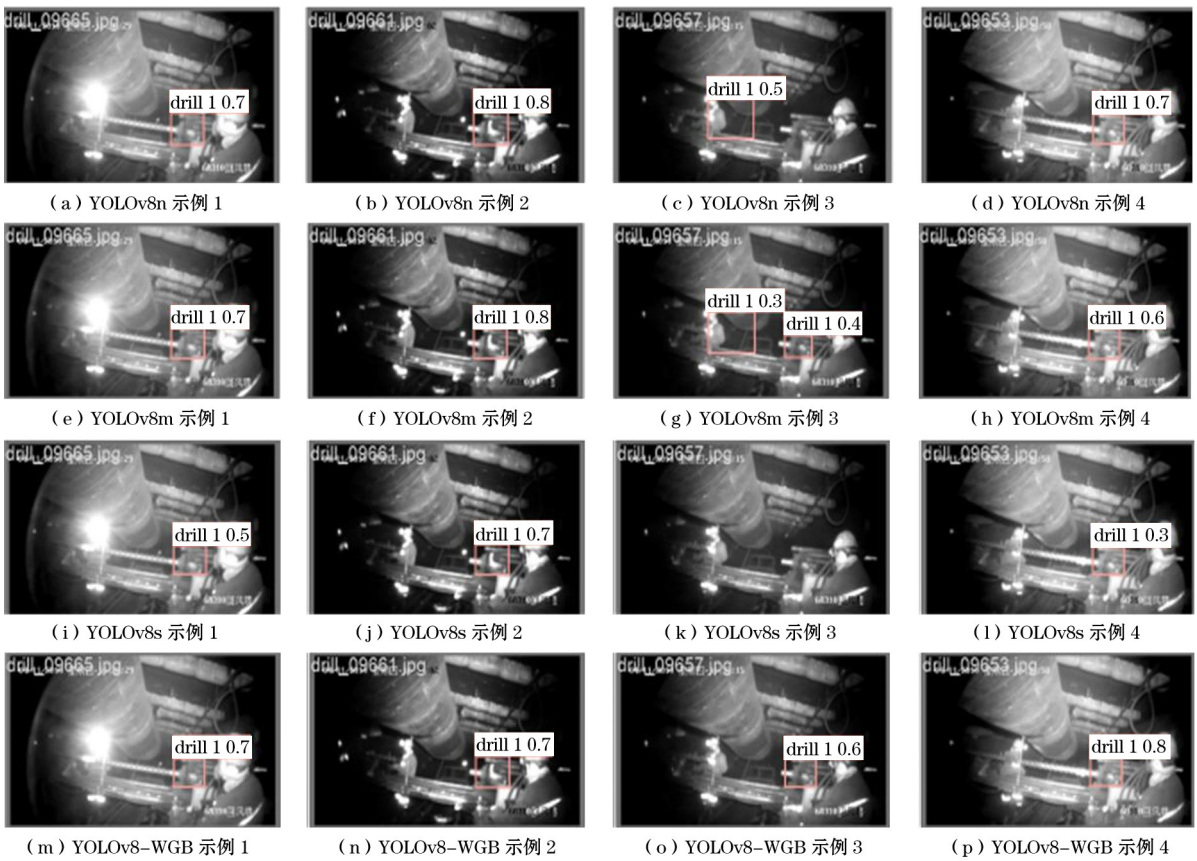
Table 2 Comparison of module ablation parameters

WIoU	GAM	BiFPN	P	R	$mAP@0.5$	模型复杂度/GFLOPs	PIP/ms	FPS/(帧/s)
×	×	×	0.857	0.432	0.533	8.9	2.2	455
√	×	×	0.855	0.441	0.536	8.1	2.3	435
×	√	×	0.954	0.418	0.581	9.4	2.2	455
×	×	√	0.802	0.458	0.558	8.9	2.3	435
√	√	×	0.893	0.46	0.582	9.4	2.4	417
√	×	√	0.896	0.474	0.587	8.3	2.4	417
×	√	√	0.974	0.419	0.585	9.5	4.8	208
√	√	√	0.918	0.480	0.577	9.5	2.4	417

3.4 模型对比试验

为了验证本文提出的改进算法的有效性,选择 YOLOv8n、YOLOv8m 和 YOLOv8s 网络模型作为对比模型进行对比试验。将各模型在数据集上的检测效果进行对比,并进一步整理试验数据的平均精度均值。检测效果示例如图 7 所示。由图 7 可知,改进后的模型漏检和误检减少。各模型的平均精度均值对比如图 8 所示。由图 8 可知,改进模型 YOLOv8-WGB 的平均精度均值优于其他 3 个对比模型,在平均精度均值上有效提升了模型性能;为了

更加直观地展示模型性能的提升效果,将各模型的评价指标进行整理汇总,具体数据见表 3。



(a) YOLOv8n 示例 1 (b) YOLOv8n 示例 2 (c) YOLOv8n 示例 3 (d) YOLOv8n 示例 4

(e) YOLOv8m 示例 1 (f) YOLOv8m 示例 2 (g) YOLOv8m 示例 3 (h) YOLOv8m 示例 4

(i) YOLOv8s 示例 1 (j) YOLOv8s 示例 2 (k) YOLOv8s 示例 3 (l) YOLOv8s 示例 4

(m) YOLOv8-WGB 示例 1 (n) YOLOv8-WGB 示例 2 (o) YOLOv8-WGB 示例 3 (p) YOLOv8-WGB 示例 4

图 7 检测效果示例

Fig.7 Examples of detection effect

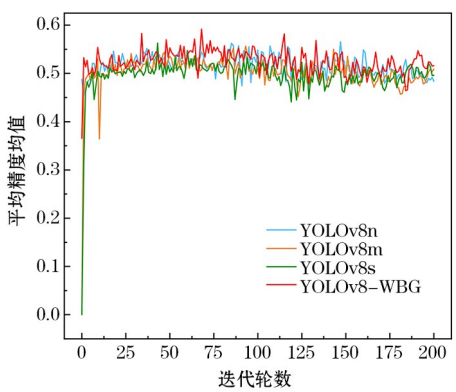


图 8 各模型的平均精度均值对比

Fig.8 Comparison of mean average precision of each model

表 3 试验模型参数对比

Table 3 Parameters comparison of test models

模型	P	R	$mAP@0.5$	模型复杂度/GFLOPs	PIP/ms	FPS/(帧/s)
YOLOv8n	0.857	0.432	0.533	8.9	2.2	455
YOLOv8m	0.895	0.453	0.561	78.7	9.2	109
YOLOv8s	0.918	0.475	0.588	28.4	4.1	244
YOLOv8-WGB	0.918	0.480	0.577	9.5	2.4	417

由表 3 可知,改进后的模型在各项评估指标上均表现出良好的效果。在模型精度方面,YOLOv8-WGB 的 P 值相较于 YOLOv8n 和 YOLOv8m 分别提高了 7.1%和 2.6%,与 YOLOv8s 持平;R 值相较于 YOLOv8n、YOLOv8m 和 YOLOv8s 分别提高了 11.1%、5.96%和 1.05%;mAP@0.5 相较于 YOLOv8n 和 YOLOv8m 分别提高了 8.3%和2.9%,稍低于 YOLOv8s。综上,YOLOv8-WGB 模型在精确率、召回率和 mAP@0.5 方面基本达到了 YOLOv8s 水平。

在模型复杂度方面,YOLOv8-WGB 相较于 YOLOv8n 增加了0.6 GFLOPs,相较于 YOLOv8m 和 YOLOv8s 分别减少了 18.9 GFLOPs 和 69.2 GFLOPs。在推理时间方面,PIP 相较于 YOLOv8n 增加了 0.2 ms,相较于 YOLOv8m 和 YOLOv8s 分别减少了 6.8 ms 和 1.7 ms。在推理速度方面,FPS 相较于 YOLOv8n 增加了 38 帧/s,相较于 YOLOv8m 和 YOLOv8s 分别增加了 308 帧/s

和 173 帧/s。YOLOv8-WGB 模型在推理速度上达到 YOLOv8n 水平。

4 结论

鉴于矿山井下环境复杂多变,给钻头的识别与检测带来了较大的挑战。为此,提出了一种改进的目标检测算法,即基于 YOLOv8n 优化的 YOLOv8-WGB 模型。通过试验验证,该改进模型在性能上表现出显著优势,模型复杂度为 9.5 GFLOPs,推理速度达到 417 帧/s,P 值、R 值及 $mAP@0.5$ 分别为 91.8%、48.0%和 57.7%。与原始 YOLOv8n 模型相比,YOLOv8-WGB 的模型复杂度和推理速度与之相当,P 值、R 值及 $mAP@0.5$ 分别提高了 7.1%、11.1%和 8.3%。YOLOv8-WGB 在保持 YOLOv8n 模型复杂度和推理速度优势的基础上,检测精度显著优于 YOLOv8s。在矿山井下钻头检测任务中,该模型有效平衡了检测速度与准确性,展现出较强的适用性与潜在优势。

参考文献(References):

[1] GAO Jihong, LI Bo, WANG Xuewen, et al. A coal and gangue detection method for low light and dusty environments[J]. Measurement Science and Technology, 2023,35(3):035402.

[2] LEI Mingfeng, ZHANG Yunbo, DENG E, et al. Intelligent recognition of joints and fissures in tunnel faces using an improved mask region-based convolutional neural network algorithm[J]. Computer-Aided Civil and Infrastructure Engineering,2024,39(8):1123-1142.

[3] XIE Xingxing, CHENG Gong, WANG Jiabo, et al. Oriented R-CNN and beyond[J]. International Journal of Computer Vision,2024,132:2420-2442.

[4] HE Deqiang, MA Rui, JIN Zhenzhen, et al. Welding quality detection of metro train body based on ABC mask R-CNN [J]. Measurement,2023,216:112969.

[5] ORESKI Goran. YOLO*C:adding context improves YOLO performance[J]. Neurocomputing,2023,555:126655.

[6] CHEN Youliang, XU Hanli, ZHANG Xiangjun, et al. An object detection method for bayberry trees based on an improved YOLO algorithm[J]. International Journal of Digital Earth,2023,16(1):781-805.

[7] WANG Haijian, MO Han, LU Shilin, et al. Electrolytic capacitor surface defect detection based on deep convolution neural network[J]. Journal of King Saud University-Computer and Information Sciences,2024,36(2):101935.

[8] GONG Lixiong, HUANG Xiao, CHAO Yinkang, et al. An enhanced SSD with feature cross-reinforcement for small-object detection[J]. Applied Intelligence, 2023, 53(16): 19449-19465.

[9] ZHANG Long, LI Jiaming, ZHANG Fuquan. An efficient forest fire target detection model based on improved YOLOv5[J]. Fire,2023,6(8):6080291.

[10] LAWAL Olarewaju Mubashiru. YOLOv5-LiNet: a lightweight network for fruits instance segmentation[J]. Plos ONE,2023,18(3):e0282297.

[11] LI Yaodi, XUE Jianxin, ZHANG Mingyue, et al. YOLOv5-ASFF:a multistage strawberry detection algorithm based on improved YOLOv5[J]. Agronomy,2023,13(7):1901.

[12] JIN Li, YU Yanqi, ZHOU Jianing, et al. SWVR: a lightweight deep learning algorithm for forest fire detection and recognition[J]. Forests,2024,15(1):15010204.

[13] TANG Chaoshen, LI Bin, SUN Junding, et al. GAM-SpCaNet: gradient awareness minimization-based spinal convolution attention network for brain tumor classification [J]. Journal of King Saud University-Computer and Information Sciences,2023,35(2):560-575.

[14] SU Pengyan, LI Hao, WANG Xiaoyun, et al. Improvement of the YOLOv5 model in the optimization of the brown spot disease recognition algorithm of kidney bean[J]. Plants, 2023,12(21):3765.

[15] WU Yun, JIANG Jianyong, HUANG Zimeng, et al. FPANet: feature pyramid aggregation network for real-time semantic segmentation[J]. Applied Intelligence,2022,52(3):3319-3336.

[16] LI Rui, ZHANG Baopeng, LIU Wei, et al. PANet:an end-to-end network based on relative motion for online multi-object tracking[J]. ACM Transactions on Multimedia Computing, Communications and Applications,2023,19(6):1-21.

[17] WU Xianjun, JIANG Xian, DONG Ligang. Gated weighted normative feature fusion for multispectral object detection [J]. The Visual Computer,2024,40:6409-6419.

[18] CHEN Chao, LIANG Rui, SONG Mengyao, et al. Noise-assisted data enhancement promoting image classification of municipal solid waste[J]. Resources, Conservation and Recycling,2024,209:107790.