基于改进 YOLOv8s 的矸石流目标检测方法*

王国新[1,2],祝建东[1,2],汝洪芳[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.黑龙江科技大学 黑龙江省光学三维测量与检测重点实验室, 黑龙江 哈尔滨 150022)

摘要:煤矸分选是煤矿生产的必要环节,煤矸自动识别是实现智能化分选的前提。基于深度学习的煤矸石目标检测可以充分利用图像特征,其精度上限高,但存在模型计算量较大、难以部署实时分选现场等问题。针对上述问题,提出轻量级的 DVBS-YOLOv8s 煤矸石目标检测方法。为了提高模型的推理速度,在 YOLOv8s 的网络骨干部分引入 DSConv 和 Vanillanet 模块,在网络颈部引入 Slim-neck 实现轻量化;为了提高模型对皮带背景和煤矸石目标的特征抓取能力,引入 BoT 和 Slim-neck;最后利用 WIoU 和 Soft-NMS 算法增强训练效果。搭建图像采集试验平台,并自建 14 790 张煤矸石图像作为数据集进行试验,结果表明,在煤矸石目标检测任务中,与 YOLOv8s 相比,改进后的 DVBS-YOLOv8s-WS 网络模型更加关注全局特征,增大了感受野,同时推理速度加快了 32%;与 YOLOv5 系列、YOLOv7 系列、SSD、Faster-RCNN 和 YOLOv8s 相比,检测精度和速度均有所提高,为实现基于深度学习的煤矸石高精度实时分选提供了参考。

关键词:煤矸石;目标检测;改进 YOLOv8s;轻量化;实时分选

中图分类号:TD849+.5;TD94 **文献标识码**:A **文章编号**:1005-2763(2025)04-0229-09

An Object Detection Method for Gangue Flow Based on Improved YOLOv8s

WANG Guoxin[1,2], *ZHU Jiandong*[1,2], *RU Hongfang*[1]

(1.School of Electrical and Control Engineering, Heilongjiang University of Science and Technology, Harbin, Heilongjiang 150022, China;
2.Key Laboratory of Optical 3D Measurement and Testing in Heilongjiang Province, Heilongjiang University of Science and Technology, Harbin, Heilongjiang 150022, China)

Abstract: Coal gangue sorting is an essential process in coal mine production, and automatic identification of coal gangue is a prerequisite for intelligent sorting. Coal gangue object detection based on deep learning can make full use of image features, and the upper limit of accuracy is high, but there are problems such as large amount of model calculation and difficulty in deploying real-time sorting sites. Aiming at the above problems, a lightweight DVBS-YOLOv8s coal gangue object detection method was proposed. To enhance the inference speed of the model, DSConv and Vanillanet modules were introduced into the backbone of YOLOv8s, and Slim-neck was incorporated into the neck of the network to achieve lightweight design. Additionally, to improve the feature capture ability of the model for conveyor belt background and coal gangue objects, BoT and Slim-neck were introduced. Finally, WIoU and Soft-NMS algorithms were used to enhance the training effect. An image acquisition test platform was set up, and a dataset of 14 790 coal gangue images was self-constructed for testing. The results show that in the coal gangue object detection task, compared with YOLOv8s, the improved DVBS-YOLOv8s-WS network model pays more attention to global features, increases the receptive field, and inference speed increase by 32%. Compared to the YOLOv5 series, YOLOv7 series, SSD, Faster-RCNN, and YOLOv8s, both detection accuracy and speed are improved, providing reference for achieving high-precision real-time sorting of coal gangue based on deep learning.

Key words: Coal gangue, Object detection, Improved YOLOv8s, Lightweighting, Real-time sorting

---

* **收稿日期**:2024-06-12
基金项目:黑龙江省重点研发计划项目(GY2023ZB0056);2022年度黑龙江省省属高等学校基本科研业务费项目(2022-KYYWF-0545)
作者简介:王国新(1980—),男,辽宁建平人,硕士,副教授,主要从事深度学习和图像处理研究。E-mail:wanggx1980@126.com
通信作者:祝建东(2000—),男,山东德州人,硕士研究生,主要从事煤矸石分选和深度学习研究。E-mail:19953450530@163.com

0 引言

我国煤炭资源丰富,对煤炭进行清洁高效地开发利用是我国能源转型发展的立足点和首要任务[1]。矸石是煤炭开采过程中伴随的固体污染物,其密度大、发热量小,混入工业用煤中,严重影响燃烧质量和发热效率。因此,煤矸分选是保证煤矿清洁开采的必要环节。

目前煤矸分选方法主要有人工选煤、跳汰选煤、重介质选煤、射线法和图像识别法等[2]。人工选煤是传统的选煤方式,此方法成本高、效率低且危害人体健康。跳汰选煤和重介质选煤需要庞大复杂的设备,浪费水资源且运行成本高。射线法识别技术成熟且抗干扰能力强,但射线对人体危害大,防护成本高。图像识别法是工业相机与图像处理技术相结合的识别方法,作为人工智能的重要应用领域,图像识别法发展十分迅速,融合人工智能的机器视觉技术在煤矸石目标检测领域具有广阔的发展前景。

射线法和图像识别法等非接触式识别方法在煤矸石识别任务中已经取得了一定成果。李曼等[3]选用最小二乘支持向量机作为煤和矸石的图像识别分类器,输入煤矸的表面联合特征作为分类器的输入向量,经过训练和验证,煤和矸石分类准确率分别为90.3%和83.0%,平均耗时153 ms。杨慧刚等[4]利用X射线结合机器视觉对煤矸石图像进行采集,并通过图像处理方法获取煤矸石的灰度和纹理特征信息,从而实现煤矸石的识别。WATT等[5]利用双能γ射线对煤和矿物进行透射分析,进而对被测矿物进行分类。机器视觉与图像处理技术可以通过提取煤矸石的图像特征进行分类识别,在煤矸分选领域具有重要意义。MA等[6]利用小波变换的多尺度方法和粗糙集理论提取煤矸石边缘。LI等[7]利用灰度直方图、分形维数等作为识别特征,建立BP神经网络进行特征融合,进而识别煤矸石。众多研究者利用图像处理和机器学习的方法推动了煤矸分选技术的发展,但这些方法普遍存在数据预处理复杂、检测维度单一、识别准确率较低、泛化性差等问题[8-10]。

近年来,深度学习的快速发展带动了煤矸石目标检测方法的进步,基于深度学习的煤矸分选具有非接触、安全性高、精度上限高等优点,是未来煤矸石目标检测领域的重要发展方向。PU等[11]利用迁移学习方法在煤矸石数据集上进行训练,识别准确率仅有82.5%,模型性能有待提高。曹现刚等[12]利用迁移学习改进AlexNet特征提取网络,并结合RPN网络获取煤和矸石的分类信息和像素坐标,模型检测准确率为90.17%。来文豪等[13]提出了一种基于多光谱成像技术和目标检测的煤矸石分离方法,基于目标检测模型YOLOv4进行改进,成功检测出煤和矸石,但检测时间长,约为4.18 s。众多学者尝试利用深度学习方法进行煤矸分选,但上述方法受限于精度或速度,难以部署实时选矸现场[14-17]。

煤矸石检测任务具有目标和背景特征相似、实时性要求高等特点,深度学习模型可以安全准确地识别煤矸,但巨大的计算量使模型难以部署到实时的选矸任务中。在众多深度学习模型中,YOLO系列提供的YOLOv8s模型在精度和速度方面有良好的平衡,本文在其基础上进行改进:引入分布移位卷积DSConv、Vanillanet模块和Slim-neck结构,降低网络的计算量,提高推理速度,在煤矸石目标检测任务中有更好的实时性。针对煤矸石目标和传送带背景特征相似度高的问题,引入BoT模块,结合Slim-neck以提高模型对全局特征的分析能力,建立DSConv-Vanillanet-BoT-Slim-neck-YOLOv8s(DVBS-YOLOv8s)网络模型,利用WIoU和Soft-NMS算法增强训练效果,得到改进后的网络模型DVBS-YOLOv8s-WS,该模型有助于进一步提高检测精度。

1 网络模型改进

1.1 YOLOv8s网络结构

YOLOv5是一种基于深度学习的目标检测算法,它在速度和精度之间取得了良好的平衡[18]。YOLOv8作为YOLOv5的升级版本,在多个领域比YOLOv5具有更高的精度[19]。YOLOv8网络结构如图1所示。YOLOv8在YOLOv5的结构基础上进行改进,使用C2f模块代替C3模块,Head部分使用解耦检测头,将分类任务和检测任务进行解耦。

YOLOv8是一种高效的目标检测算法,是当前目标检测领域的研究热点之一。与YOLOv5相比,YOLOv8在众多领域有着更高的检测精度,但其较高的计算量却使它在实际工作应用中有一定困难。YOLOv8网络模型从小到大依次为YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x,网络模型层数越深,计算量越大,网络的精度越高,但推理速度会因为网络复杂度的提高而下降。综合考虑煤矸分选任务对速度和精度的要求,本文在YOLOv8s模型基础上进行优化,提出了基于DVBS-YOLOv8s的煤矸石目标检测方法。

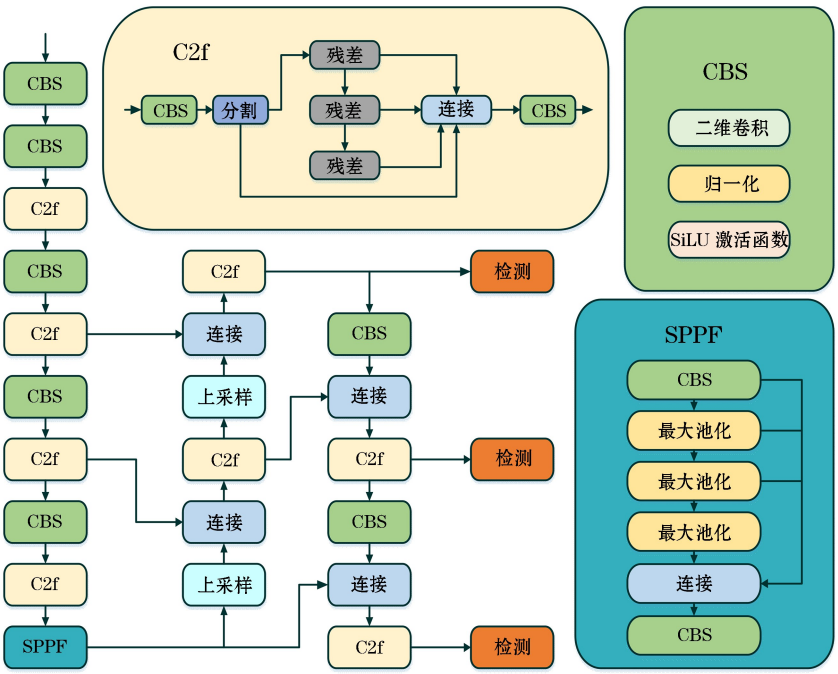


图 1 YOLOv8 网络结构

Fig.1 YOLOv8 network structure

1.2 分布移位卷积

卷积是网络模型中提取特征的重要操作,在资源受限的环境(如矿区现场)中,用 DSConv 代替传统卷积可以减少冗余计算,提高内存利用率,加快运行速度[20]。DSConv 原理如图 2 所示,DSConv 通过量化和分布移位模拟卷积层运算,将传统的卷积核分为两个部分:可变量化核 VQK 和分布偏移。

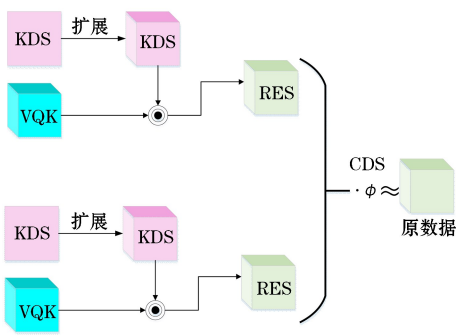


图 2 DSConv 原理

Fig.2 DSConv principle

VQK 将卷积核权重映射为离散的量化整数,通过在 VQK 中仅存储量化后的整数值,占用更低的内存,实现更高的计算速度,同时分别通过内核分布偏移 KDS 和通道分布偏移 CDS 模拟原始卷积核在位置和通道上的移动。KDS 和 CDS 两个张量通过可训练的参数改变 VQK 的分布,这种分布移位的方式在一定程度上模拟了传统卷积中卷积核的移动。

DSConv 通过量化权重实现了更高效的计算,分布移位通过可训练的参数代替卷积核,可以提高拟合能力。

1.3 Vanillanet 模块

煤矸石检测需要在复杂背景中识别出形态和尺寸各异的煤矸石,这种任务要求模型具有强大的特征提取能力和适应性。Vanillanet 是极简卷积神经网络,网络模型仅包含最简单的卷积、池化、全连接等基础算法[21],减少了模型复杂性和计算资源的需求。Vanillanet 结构如图 3 所示,以 6 个卷积层构成 Vanillanet-6 为例,网络结构通过简洁地堆叠而成,遵循传统卷积神经网络的常用方案:随着网络深度的增加,分辨率逐渐缩小,通道数逐渐增大。

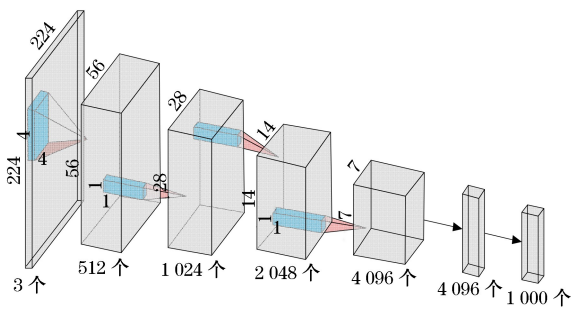


图 3 Vanillanet 结构(单位:px)

Fig.3 Vanillanet structure (Unit:px)

简单的网络结构导致弱非线性,为此 Vanillanet 中的一个卷积层可以拆分为两个卷积层,从而增加其非线性。为避免增加网络的计算复杂度,网络仅在训练过程中拆分网络层,推理时将网络层融合。被拆分的卷积层遵循的激活函数$A'(x)$可以表示为:

$$A'(x)=\left(1-\frac{e}{E}\right)A(x)+\frac{e}{E}x \tag{1}$$

式中:$A(x)$ 表示 ReLU 等非线性激活函数;e 表示训练任务中当前的迭代数;E 表示训练任务中总的迭代数。可以看出随着训练迭代数的增加,非线性激活函数逐渐变为线性恒等映射,当训练收敛时,$A(x)=x$,表示两个卷积层中没有激活函数,不具有非线性激活的两个卷积层就可以融为一层。

此外 Vanillanet 还利用了并发堆叠激活函数,为了避免简单累积,添加 a_i 和 b_i 作为每个激活函数叠加时的尺度和偏差,则 n 个激活函数的并发堆叠可以表示为:

$$As(x)=\sum_{i=1}^{n}a_iA(x+b_i) \tag{2}$$

通过并发堆叠激活函数可以显著提高函数的非线性,不仅使网络模型拥有深层网络的强特征提取能力,还有简洁的结构优势,推理速度和计算量得到了显著优化。DSConv 与 Vanillanet 模块相结合,可以实现网络模型的整个骨干部分的轻量化。

1.4 BoT 模块

卷积操作可以提取局部特征,YOLOv8 将不同维度提取的特征进行融合,但煤矸石检测任务中,目标与背景的颜色相近且占据视野较大,局部特征很难有效地产生依赖关系,难以集成全局信息,容易丢失部分特征,导致检测精度的下降。谷歌于 2020 年提出 Vision Transformer,并将 Transformer 应用在图像分类领域[22],此后很多研究者尝试将卷积与 Transformer 结合,让网络保留局部特征的同时也能够更好地学习上下文特征。

BoTNet 使用一种结合方法,在 BoT block 中使用 Transformer 中的多头自注意层 MHSA 结合卷积层,从而让网络更有效地关注全局信息[23]。将 BoT 模块插入 YOLOv8s 骨干尾部,可以让网络关注大目标的特征信息,提高网络对煤矸石检测的精度。

1.5 Slim-neck

GSConv 是一种轻量化的卷积,其结构如图 4 所示,通过卷积和深度卷积相结合的方式,减少了计算量和参数量,从而提高模型的训练和推理效率,这种结合的卷积操作使其很大程度上保持了与原始卷积相同的效果,提高了识别效率。

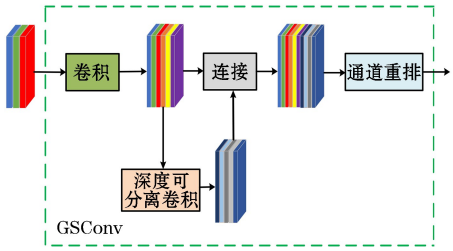


图 4 GSConv 结构

Fig.4 GSConv structure

以煤矸石为前景、传送带为背景的数据集图像,因前景与背景特征相近,增大感受野时容易因特征信息不明显,导致精度下降。Slim-neck 参考 VoVnet、GSConv 和 CSPnet,设计了 VoV-GSCSP 结构,如图 5 所示,通过利用多个卷积和 GSConv 进行特征融合,能够更全面地提取图像特征[24]。

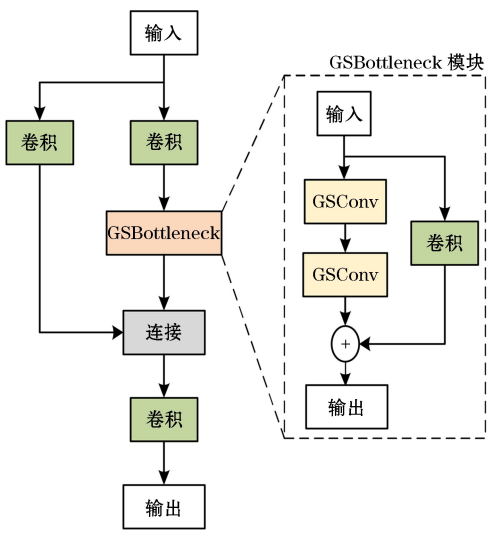


图 5 VoV-GSCSP 结构

Fig.5 VoV-GSCSP structure

将 VoV-GSCSP 结构优秀的特征融合能力和 BoT 全局特征抓取能力相结合,可以充分利用特征信息,有效避免局部特征之间联系性差异而导致精度降低的现象,同时 Slim-neck 的低参数量和低计算量使网络拥有更高的检测速度,实现模型颈部的轻量化。

1.6 DVBS-YOLOv8s 网络结构

本文在 YOLOv8s 的基础上进行改进,引入轻量化卷积 DSConv 代替主干网络的传统卷积,引入极简卷积网络 Vanillanet 模块代替主干网络的 3 个 C2f 模块,在主干网络的尾部引入 BoT 模块,在网络颈部引入 Slim-neck 结构,得到 DVBS-YOLOv8s,网络结构如图 6 所示。

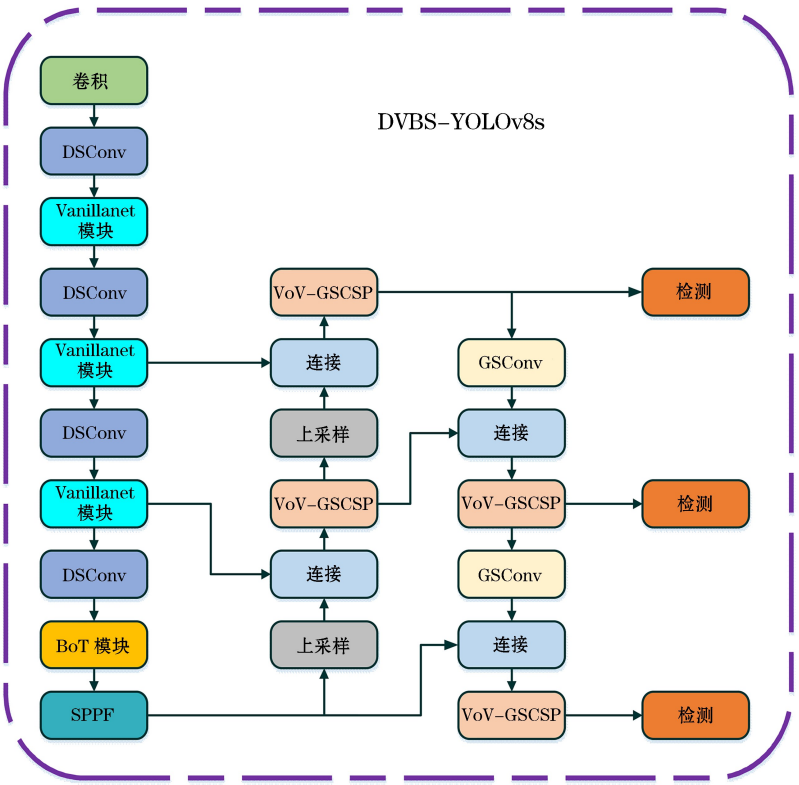


图 6 DVBS-YOLOv8s 结构

Fig.6 DVBS-YOLOv8s structure

1.7 损失算法改进

YOLOv8 训练过程中使用 CIoU 计算损失[25],并用 NMS 算法剔除效果差的检测框[26]。

传统的 NMS 算法是将检测框按得分排序,保留得分最高的框,并删除与该框 IoU 大于一定比例的其他框,然而这样的操作可能会误删一些效果良好的检测框。Soft-NMS 是一种贪婪算法,为避免草率地删除所有 IoU 大于阈值的框,算法会对检测框 IoU 大小进行评分,局部高分的检测框不会被删去,Soft-NMS 引入评分机制代替删除机制,有利于根据评分寻找全局最优检测框[27]。

检测框的宽、高和中心点坐标分别用 w、h、(x,y) 表示,真实框的宽、高和中心坐标为 w_{gt}、h_{gt}、(x_{gt},y_{gt}),真实框和检测框的最小外接矩形的宽和高为 W_g、H_g,IoU 为交并比,则 CIoU 损失为:

$$L_{CIoU}=1-\left(IoU-\alpha v-\frac{(x-x_{gt})^2+(y-y_{gt})^2}{W_g^2+H_g^2}\right) \tag{3}$$

其中,α 和 v 的计算公式如下:

$$\alpha=\frac{v}{1-IoU+v} \tag{4}$$

$$v=\frac{4}{\pi^2}\left(\tan^{-1}\frac{w}{h}-\tan^{-1}\frac{w_{gt}}{h_{gt}}\right)^2 \tag{5}$$

式中:v 为检测框和真实框宽高比例的差异程度;α 为权重系数。

EIOU 损失算法使惩罚项直接应用于预测框,可以有效提高回归精度和收敛速度[28],则 EIOU 损失为:

$$L_{EIoU}=1-(IoU-\frac{(x-x_{gt})^2+(y-y_{gt})^2}{W_g^2+H_g^2}-\frac{(x-x_{gt})^2}{W_g^2}-\frac{(y-y_{gt})^2}{H_g^2}) \tag{6}$$

为解决样本质量不均衡的问题,引入控制异常值抑制程度的算子 γ,Focal-EIoU 的损失为:

$$L_{Focal\text{-}EIoU}=IoU^{\gamma}L_{EIoU} \tag{7}$$

Focal-EIoU 的聚焦机制是静态的,无法充分发挥非单调聚焦机制的潜能。WIoU 补充了动态非单调聚焦机制,使用离群度对锚框质量进行评估,并引入了梯度增益[29],这不仅可以抑制低质量实例产生的劣质锚框,还能降低高质量锚框的竞争力,从而降低对标注监督的依赖性,提高模型生成检测框的鲁棒性,使 WIoU 可以更注重普通质量的锚框,增强训练过程中锚框的有效评估,则 WIoU 的损失为:

$$L_{WIoU}=rR_{WIoU}(1-IoU) \tag{8}$$

$$R_{WIoU}=\exp\left(\frac{(x-x_{gt})^2+(y-y_{gt})^2}{W_g^2+H_g^2}\right) \tag{9}$$

$$r=\frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{10}$$

式中:R_{WIoU}是一种用于衡量检测框与真实框之间位置关系的权重系数;β 用于定义描述锚框质量的离群度,利用 β 构建梯度增益系数,由超参数 α、δ 控制的离群度 β 与梯度增益 r 之间的映射,当离群度满足阈值时,锚框获得最高的梯度增益。锚框的质量评估随着 IoU 动态变化,WIoU 与 Soft-NMS 算法相结合,得出本研究改进的模型 DVBS-YOLOv8s-WS,该模型可以更有效地训练网络。

2 试验结果与分析

2.1 试验平台

图像采集试验平台由 4 个条形光源、工业相机、运输皮带和遮光板组成,工业相机型号为海康 MV-CA050-20UC。

将煤和矸石样本放置在传送带上,在暗箱中进行图像采集与识别,提取煤和矸石的分类信息以及矸石的坐标信息,并发送到分选机构进行抓取。试验软硬件环境见表 1。

表 1 试验软硬件环境

Table 1 Software and hardware environment of the test

名称	信息
操作系统	Ubuntu 20.04 LTS
深度学习框架	Pytorch2.0.1
处理器	Intel(R) Xeon(R) Gold 6226R * 2
内存	196 GB
显卡	NVIDIA Tesla T4 * 4
运行环境	Anaconda3
Python 版本	Python 3.9

2.2 数据集建立

在实验室搭建了长为 7 m 的皮带输送机以及光源系统,使用工业相机在暗箱内采集传送带上的煤和矸石的图像数据,在集贤煤矿的选矸现场拍摄矸石流图像。经筛选和数据增强后,得到总数为 14 790 的煤和矸石图像数据集。将数据集以 8∶1∶1 的比例分为训练集、验证集和测试集。

2.3 评估指标

在目标检测算法中,召回率、F_1 分数、mAP 以及推理速度常被选作网络模型的评价指标。在指定软硬件条件下,推理速度为网络模型每秒处理的图像数目。

$$P=\frac{N_{TP}}{N_{TP}+N_{FP}} \tag{11}$$

$$R=\frac{N_{TP}}{N_{TP}+N_{FN}} \tag{12}$$

$$F_1=\frac{2PR}{P+R} \tag{13}$$

$$mAP=\frac{\sum_{k=1}^{C}P_{ek}}{C} \tag{14}$$

$$P_{ek}=\int P\,\mathrm{d}R \tag{15}$$

式中:P 为准确率,表示检测正确部分占检测正样本数的比例;R 为召回率,表示检测正确部分占全体正样本数的比例;F_1 分数为评价二分类问题的重要指标,其值越高,检测效果越好;mAP 为全类别平均准确率的平均值,mAP@0.5:0.95 即 IoU 从 0.5 均匀变动到 0.95 的各类别平均 P_{ek} 值;N_{TP} 为模型检测正确的正样本数;N_{FP} 为误检的样本数;N_{FN} 为漏检的样本数;P_{ek} 为平均准确率;C 为类别数。

2.4 试验结果与分析

训练过程中,网络模型在训练集和验证集上的损失函数曲线分别如图 7 和图 8 所示。DVBS-YOLOv8s-WS 改进模型损失函数曲线波动小,收敛速度快,说明 WIoU 和 SoftNMS 算法有利于网络模型的训练。

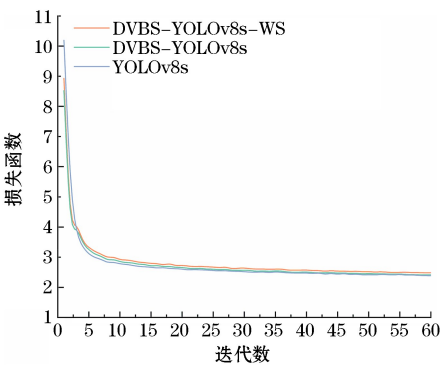


图 7 训练集损失函数曲线

Fig.7 Loss function curves of training set

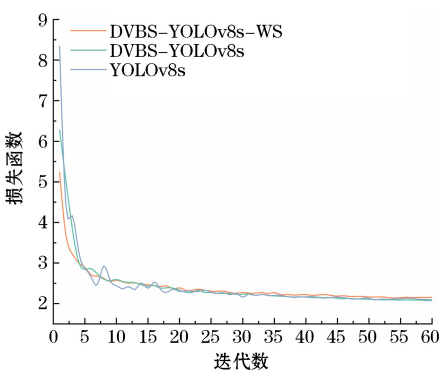


图 8 验证集损失函数曲线

Fig.8 Loss function curves of validation set

为验证 DVBS-YOLOv8s-WS 网络模型的推理性能,在验证集上与单阶段深度学习网络 YOLOv5s、YOLOv5x、YOLOv7、YOLOv7x、YOLOv8s、SSD[30] 和两阶段深度学习网络 Faster-RCNN[31] 的精度进行比较,各网络模型在训练过程中,精度随迭代数变化的曲线如图 9 所示。试验参数统一设置为:无预

训练权重,样本数量为 16,迭代数为 60,置信度阈值为 0.25,IoU 阈值为 0.7。

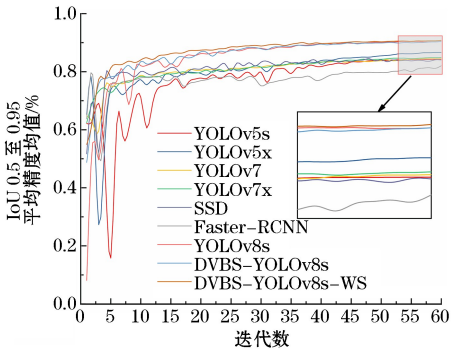


图 9 各网络模型精度随迭代数的变化曲线

Fig.9 Variation curves of the accuracy of each network model with the number of iterations

YOLOv8s 与改进网络 DVBS-YOLOv8s-WS 在验证集上的平均精度均值分别为 90.3% 和 90.8%,改进后的网络精度有少量提升。用测试集对各模型进行测试,试验结果见表 2。

表 2 各网络模型试验结果对比

Table 2 Comparison of test results of each network model

模型	F_1 分数/%	召回率/%	帧率/(帧/s)	IoU0.5 至 0.95 平均精度均值/%
YOLOv5s	98.8	98.8	83.3	84.2
YOLOv5x	99.2	99.3	22.3	86.7
YOLOv7	98.8	99.1	28.0	84.5
YOLOv7x	99.1	99.3	21.1	85.1
SSD	99.2	99.0	58.6	84.2
Faster-RCNN	98.0	99.4	7.9	82.4
YOLOv8s	99.6	99.5	80.6	92.5
DVBS-YOLOv8s-WS	99.6	99.5	106.4	92.6

与 YOLOv8s 网络模型相比,改进后的 DVBS-YOLOv8s-WS 模型的帧率提高了 32%,即推理速度提高了 32%,精度和推理速度是煤矸石实时分选任务的重要指标。网络模型中使用更多网络层可以得到更强的特征提取能力,但随着计算量的增加,推理速度会降低,因此需要权衡网络模型的复杂度和速度,使其拥有更好的实时分选性能。各网络模型精度随帧率的变化曲线如图 10 所示,横轴表示帧率,纵轴表示精度,圆形半径表示模型的大小。

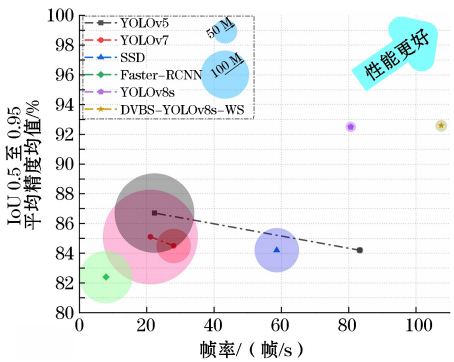


图 10 各网络模型精度随帧率的变化曲线

Fig.10 Variation curves of the accuracy of each network model with the frame rate

由图 10 可知,在众多经典深度学习网络中,YOLOv5 系列在煤矸石目标检测任务中的性能优于 YOLOv7 系列,YOLOv8s 和 DVBS-YOLOv8s-WS 模型拥有更高的精度和推理速度。改进后的模型 DVBS-YOLOv8s-WS 保持了相同大小,同时精度有少量提升,推理速度有显著提升。为了探究各改进部分的效果,用测试集对各个网络模块的改进部分进行测试,得到的消融试验结果见表 3。

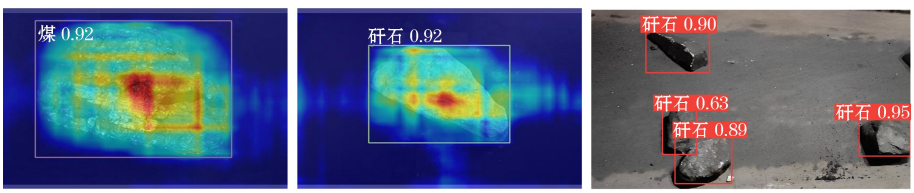
表 3 消融试验结果

Table 3 Ablation test results

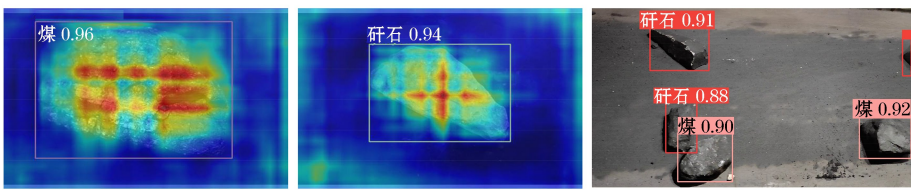
DSConv	Vanillanet 模块	BoT 模块	Slim-neck	WIoU+Soft-NMS	F_1 分数/%	十亿次浮点运算/s	帧率/(帧/s)	IoU0.5 至 0.95 平均精度均值/%
					99.6	28.4	80.6	92.5
√					99.5	24.7	90.9	92.2
	√				99.7	22.6	84.7	92.6
		√			99.6	28.6	71.9	92.5
			√		99.6	23.9	92.8	92.8
√	√				99.6	18.8	97.1	92.2
√		√			99.5	24.9	82.6	91.3
√			√		99.7	20.9	97.1	92.6
	√	√			99.6	22.8	83.3	92.4
	√		√		99.7	18.8	103.1	92.8
		√	√		99.7	24.8	78.1	93.0
√	√	√			99.4	19.0	99.0	92.3
√	√		√		99.5	15.0	111.1	92.4
√		√	√		99.8	21.1	94.3	92.7
	√	√	√		99.7	19.0	90.9	92.7
√	√	√	√		99.5	15.2	107.5	92.5
√	√	√	√	√	99.6	15.2	106.4	92.6

由表 3 可知,Slim-neck 与 BoT 模块在煤矸石目标检测任务中有良好的兼容性,Vanillanet 模块、DSConv 和 Slim-neck 有降低计算量、轻量化网络模型的效果。

为了更直观地观察改进模块的效果,使用 Grad-CAM 热力图算法。进行算法特征图的可视化[32],如图 11 所示。从图 11 可以看出,原网络模型 YOLOv8s 更聚焦于目标的中心位置,缺乏目标的整体特征,在现场环境下,进行检测任务时存在误检情况。改进后的网络模型 DVBS-YOLOv8s-WS 更关注全局特征,且增大了感受野,能够检测出视野边缘的目标,提高了煤矸石目标检测的精度。



(a) YOLOv8s



(b) DVBS-YOLOv8s-WS

图 11 热力图及现场检测结果

Fig.11 Heat maps and field test results

3 结论

本文利用深度学习方法对煤矸石进行目标检测,搭建图像采集平台,制作煤矸数据集,提出基于 DVBS-YOLOv8s 的煤矸石目标检测方法,并利用 WIoU 和 Soft-NMS 算法增强训练效果。

(1) 试验结果表明,与网络模型 YOLOv8s 相比,改进后的 DVBS-YOLOv8s-WS 模型的推理速度提高了 32%,增大了感受野,提高了煤矸石目标检测的精度。

(2) DVBS-YOLOv8s-WS 网络模型有助于推动深度学习方法应用于煤矸分选领域,高精度和高推理速度有利于满足煤矸分选的实时性,低计算量有利于实现模型的部署。

(3) 利用深度学习方法进行煤矸石识别比传统方法更安全,发展空间更广阔。在错误的检测样本中存在难以定义类别的煤矸混合样本,因此扩增数据集并添加新的分类项是后续研究重点。

参考文献(References):

[1] 孙超,姜琳,袁广玉."十四五"期间我国煤炭供需趋势分析[J].煤炭工程,2021,53(5):193-196.
SUN Chao, JIANG Lin, YUAN Guangyu. Trend analysis of China's coal supply and demand during the 14th Five-Year Plan[J]. Coal Engineering,2021,53(5):193-196.

[2] 曹现刚,李莹,王鹏,等.煤矸石识别方法研究现状与展望[J].工矿自动化,2020,46(1):38-43.
CAO Xian'gang, LI Ying, WANG Peng, et al. Research status of coal-gangue identification method and its prospect[J].Journal of Mine Automation,2020,46(1):38-43.

[3] 李曼,段雍,曹现刚,等.煤矸分选机器人图像识别方法和系统[J].煤炭学报,2020,45(10):3636-3644.
LI Man, DUAN Yong, CAO Xian'gang, et al. Image identification method and system for coal and gangue sorting robot[J]. Journal of China Coal Society, 2020, 45(10): 3636-3644.

[4] 杨慧刚,乔志敏.基于 X 射线和机器视觉的煤与矸石分选系统设计[J].工矿自动化,2017,43(3):85-89.
YANG Huigang, QIAO Zhimin. Design of separation system of coal and gangue based on X-ray and machine vision[J]. Journal of Mine Automation,2017,43(3):85-89.

[5] WATT J S, STERFFNER E J. Dual energy gamma-ray transmission techniques applied to on-line analysis in the coal and mineral industries[J]. The International Journal of Applied Radiation and Isotopes,1985,36(11):867-877.

[6] MA Xianmin, LIANG Che. Application of rough set theory in coal gangue image process[C]//2009 Fifth International Conference on Information Assurance and Security. New York: IEEE,2009:87-90.

[7] LI Wenhui, WANG Ying, FU Bo, et al. Coal and coal gangue separation based on computer vision[C]//2010 Fifth International Conference on Frontier of Computer Science and Technology. New York: IEEE, 2010:467-472.

[8] DOU Dongyang, WU Wenze, YANG Jianguo, et al. Classification of coal and gangue under multiple surface conditions via machine vision and relief-SVM[J]. Powder Technology,2019,356:1024-1028.

[9] 于国防.煤矸区分中的间隔灰度压缩扩阶共生矩阵[J].中国图象图形学报,2012,17(8):966-970.
YU Guofang. Expanded order co-occurrence matrix to differentiate between coal and gangue based on interval grayscale compression[J]. Journal of Image and Graphics, 2012,17(8):966-970.

[10] SUN Zhaoqing, LU Wu, XUAN Pengcheng, et al. Separation of gangue from coal based on supplementary texture by morphology[J]. International Journal of Coal Preparation and Utilization,2019,42:221-237.

[11] PU Y Y, APEL D B, SZMIGIEL A, et al. Image recognition of coal and coal gangue using a convolutional neural network and transfer learning[J]. Energies,2019,12(9):1735.

[12] 曹现刚,刘思颖,王鹏,等.面向煤矸分拣机器人的煤矸识别定位系统研究[J].煤炭科学技术,2022,50(1):237-246.
CAO Xian'gang, LIU Siying, WANG Peng, et al. Research on coal gangue identification and positioning system based on coal-gangue sorting robot[J]. Coal Science and Technology, 2022,50(1):237-246.

[13] 来文豪,周孟然,胡锋,等.基于多光谱成像和改进 YOLO v4 的煤矸石检测[J].光学学报,2020,40(24):72-80.
LAI Wenhao, ZHOU Mengran, HU Feng, et al. Coal gangue detection based on multi-spectral imaging and improved YOLO v4[J]. Acta Optica Sinica,2020,40(24): 72-80.

[14] GAO Rong, SUN Zhaoyun, LI Wei, et al. Automatic coal and gangue segmentation using U-Net based fully convolutional networks[J]. Energies,2020,13(4):829.

[15] 李莹.基于深度学习的煤矸石目标检测方法研究[D].西安:西安科技大学,2020.
LI Ying. Research on coal gangue detection based on deep learning[D]. Xi'an: Xi'an University of Science and Technology,2020.

[16] LIU Qiang, LI Jin'gao, LI Yusheng, et al. Recognition methods for coal and coal gangue based on deep learning[J]. IEEE Access,2021,99:77599-77610.

[17] ALFARZAEAI M S, NIU Q, ZHAO J S, et al. Coal/gangue recognition using convolutional neural networks and thermal images[J]. IEEE Access,2020,8:76780-76789.

[18] 顾清华,何鑫鑫,王倩,等.基于改进 YOLOv5 的煤矿井下暗环境矿工安全穿戴智能识别[J].矿业研究与开发,2024,44(3):201-208.
GU Qinghua, HE Xinxin, WANG Qian, et al. Research on intelligent recognition of safety wearing of miners in dark enviroment of coal mine based on improved YOLOv5[J]. Mining Research and Development,2024,44(3):201-208.

[19] 杨焜,王宏伟,张夫净,等.多尺度复杂环境下的锚孔定位方法[J].矿业研究与开发,2025,45(1):126-134.
YANG Kun, WANG Hongwei, ZHANG Fujing, et al. Method of anchor hole positioning in multi-scale complex environment[J].Mining Research and Development,2025,45(1):126-134.

[20] NASCIMENTO Marcelo Gennari Do, PRISACARIU Victor, FAWCETT Roger. DSConv: efficient convolution operator[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV).New York:IEEE,2019.

[21] CHEN Hanting, WANG Yunhe, GUO Jianyuan, et al. VanillaNet: the power of minimalism in deep learning[J]. Advances in Neural Information Processing Systems,2023, 36:7050-7064.

[22] DOSOVITSKIY Alexey, BEYER Lucas, KOLESNIKOV Alexander, et al. An image is worth 16 × 16 words: transformers for image recognition at scale[J]. ArXiv, 2020:11929.

[23] SRINIVAS Aravind, LIN Tsungyi, PARMAR Niki, et al. Bottleneck transformers for visual recognition[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE,2021.

[24] LI Hulin, LI Jun, WEI Hanbing, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles[J]. Arxiv,2022:02424.

[25] 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.

[26] XU Haoran, TANG Shibin, WANG Jia, et al. Rock fracture identification algorithm based on the confidence score and non-maximum suppression[J]. Bulletin of Engineering Geology & the Environment,2024,83(6):213.

[27] BODLA Navaneeth, SINGH Bharat, CHELLAPPA Rama, et al. Soft-NMS:improving object detection with one line of code[C]//2017 IEEE International Conference on Computer Vision (ICCV). New York: IEEE,2017.

[28] ZHANG Yifan, REN Wei, ZHANG Zhang, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing,2022,506:146-157.

[29] TONG Zanjia, CHEN Yuhuang, XU Zewei, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. Arxiv,2023:10051.

[30] YI J R, WU P X, METAXAS D N. ASSD: attentive single shot multibox detector[J]. Computer Vision and Image Understanding,2019,189:102827.

[31] REN Shaoqing, HE Kaiming, GIRSHICK Ross, 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.

[32] SELVARAJU R R, COGSWELL MDAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision,2019,128:336-359.