

基于深度注意力模型的 YOLOv8 煤矸石检测算法

祖 少 杰

(山西潞安环保能源开发股份有限公司 王庄煤矿, 山西 长治 046031)

摘 要:针对煤矿复杂环境下煤矸石检测精度低的问题, 本文提出了一种基于深度注意力模型的 YOLOv8 煤矸石检测算法。首先, 通过超分辨率重构算法对待检测目标图像进行预处理, 提升图像对比度, 增强图像质量; 然后, 针对输送带高速运转易导致图像模糊进而造成煤矸石难以准确检测的问题, 在 YOLOv8 算法框架下通过改进网络的卷积层进而提高检测速度, 并通过优化网络的 NMS 来提高算法对遮挡目标的检测精度; 其次, 针对输送带待检测目标受矿井复杂环境影响造成显著性降低导致其难以准确检测的问题, 在上述改进的算法框架下通过融合卷积块注意模型来提高图像中煤矸石的显著度, 进而提高目标的检测精度; 最后, 利用煤矿输送带监控视频数据制备训练样本和测试样本, 并将提出的算法与 YOLOv8、SSD 和 Faster-RCNN 等经典的目标检测算法进行对比。实验结果表明, 本文提出的煤矸石检测算法在应对输送带高速运转、复杂环境干扰以及部分目标物存在遮挡的情况时具有较高的检测精度, 同时保持了较好的实时性。

关键词:煤矸石; YOLOv8; 超分辨率重构; 注意力机制

中图分类号:TD94; TD391.41 **文献标志码:**A **文章编号:**1009-0797(2025)02-0064-06

The algorithm of YOLOv8 coal gangue detection based on deep attention model

ZU Shaojie

(Shanxi Lu'an Environmental Protection Energy Development Co., LTD. Wangzhuang Coal Mine, Changzhi 046031, China)

Abstract: Aiming at the low accuracy of gangue detection in complicated coal mine environment, this paper proposes a YOLOv8 gangue detection algorithm based on deep attention model. First, preprocessing of the detected target image by a super-resolution reconstruction algorithm, improve the image contrast, Enhance the image quality. Then, in view of the problem that the conveyor belt is easy to lead to image blur and it is difficult to accurately detect coal gangue, improve the detection speed by improving the convolution layer of the network under the framework of the YOLOv8 algorithm, and by optimizing the network NMS to improve the algorithm to the occlusion target detection accuracy. Next, in view of the problem of the significance of the conveyor belt to be detected by the complex environment of the mine, to improve the significance of coal gangue in the image by incorporating the convolution block attention model under the above improved algorithm framework, then improve the detection accuracy of the target. Last, preparation of training samples and test samples by using surveillance video data of coal mine conveyor belt, the proposed algorithm is also compared with classical target detection algorithms such as YOLOv8, SSD and Faster-RCNN. The experimental results show that the coal gangue detection algorithm has high detection accuracy in dealing with the high-speed operation of the conveyor belt, the complex environmental interference and the occlusion of some targets.

Key words: Coal gangue; YOLOv8; Super-resolution reconstruction; Attention mechanism

煤矸石作为煤炭开采和选洗过程中排放的固体废物, 通过分拣去除煤矸石, 可以显著提高煤炭的纯度和品质^[1]。目前, 传统的煤矸石分拣方法主要包括人工分拣和机械筛分 2 种方式^[2]。采用人工分拣主要由工人凭借经验和肉眼观察, 从煤流中手工拣出煤矸石, 存在劳动强度大、效率低、成本高, 且工作环境恶劣, 对员工的身体健康有一定危害的问题; 机械筛分通过振动筛等设备, 利用煤和矸石的粒度差异进行分离, 然而该方法对于粒度相近的煤和矸石难以有效分离, 且设备维护成本较高。

随着人工智能技术的快速发展, 诸多学者专

于利用深度学习算法研究煤矸石的分类问题, 并取得了显著进展^[3]。当前, 基于深度学习的目标检测方法主要分为两类: 第一类是 two-stage 目标检测模型, 代表性方法有 Fast R-CNN^[4] 和 Faster R-CNN^[5] 等。吴守鹏等^[6]提出了一种改进型 Faster-RCNN 模型, 用于识别运煤皮带上的异物。杨文斌等^[7]利用 Faster-RCNN 深度学习目标检测算法成功克服了在意外情况下准确辨别刮板输送机中的铁质异物(例如锚杆、角铁)所面临的挑战。这些新方法为相关领域的应用和发展提供了有益参考。另一种是 one-stage 目标检测模型, 在这类模型中, 代表性方法有

SSD^[8] 和 YOLO 系列等。这些模型以直接端到端训练为特点,具备简洁的模型架构和突出的实时性能特点,在实际应用中拥有明显优势。相比较于传统方法它们更为高效,因此在各个领域得到广泛应用并取得卓越成果。左明成等人^[9]基于改进的 SSD 算法实现煤矿井下作业场景中工人安全帽是否正确佩戴的识别;李珍峰等人^[10]提出一种基于轻量化的 YOLOv8 检测模型,可以在降低模型消耗的情况下实现对输送带上煤矸石的检测。因此,本文针对煤矿复杂环境造成的煤矸石检测精度低的问题,在 YOLOv8 算法基础上提出一种基于深度注意力模型的煤矸石检测算法。

1 传统的 YOLOv8 算法

YOLO 系列检测算法作为目前深度学习领域流行算法之一,经过近年来多次的更新迭代,发展出了许多经典的检测模型,如 YOLOv3^[11]、YOLOv4^[12]等。目前,YOLO 系列算法已经迭代到

YOLOv8,相较于之前的 YOLO 系列算法在检测精度和速度上都做了进一步优化。其结构主要由 backbone 部分、neck 部分和 head 部分组成。在 backbone 部分网络使用卷积层对输入图像进行初步的特征提取,与之前的版本相比,YOLOv8 对卷积核的大小、数量等进行了优化调整。并且 YOLOv8 骨干网络中的重要组成部分增加了结合 CSPNet (Cross Stage Partial Network) 思想的 C2f 模块。C2f 模块通过将特征图分成多个部分进行处理,然后再进行融合,有效地增强了模型的特征表达能力。在 neck 部分提出了增强版 FPN (Feature Pyramid Network),通过将骨干网络输出的不同层次的特征图进行上采样和下采样操作,然后进行融合,使得模型能够获取到不同尺度的特征信息。在 head 部分 YOLOv8 的检测头采用了解耦头结构,将分类和回归任务分离,可以使模型更加专注于不同的任务,提高分类和回归的准确性。整体的 YOLOv8 网络模型如图 1 所示。

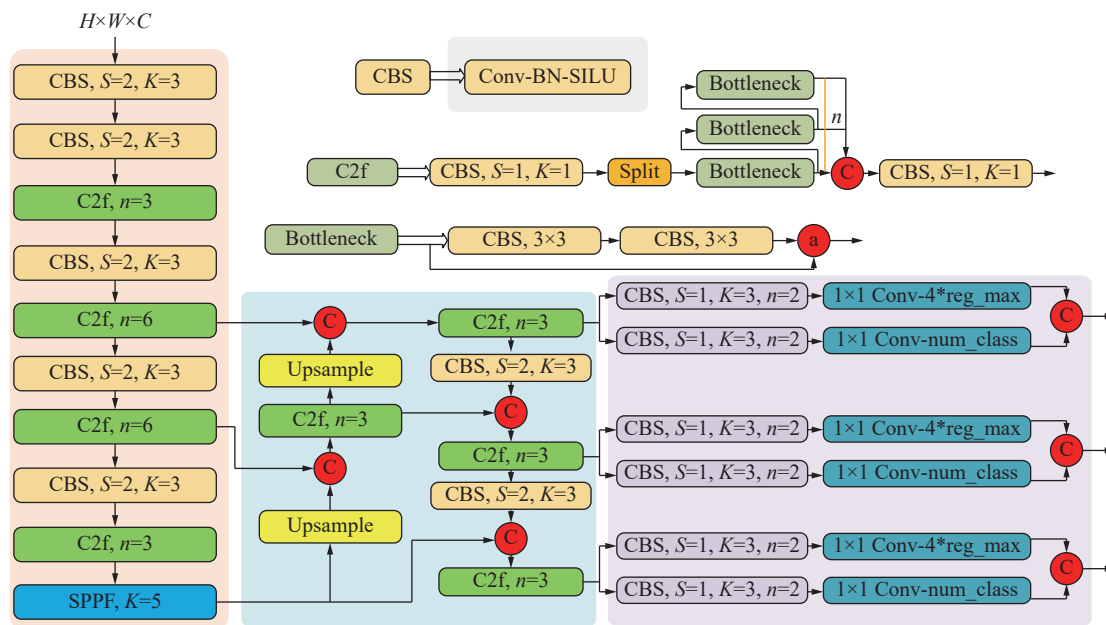


图 1 YOLOv8 网络模型

2 改进 YOLOv8 的煤矸石检测方法

本文针对矿井复杂环境下煤矸石检测精度低的问题,通过在传统的 YOLOv8 网络中融合注意力机制、改进网络卷积层及优化非极大值抑制等方法,提出了一种基于深度注意力模型的 YOLOv8 煤矸石检测算法。首先,需调整矿井传送带监控视频捕获的图像尺寸至 640×640,并对图像进行超分辨率处

理以重建;其次,将经过处理的图像输入预设计好的检测网络中进行训练,并得到相应的训练权重用于构建检测模型;最后,使用本文的测试数据集对新提出的检测网络进行验证。

2.1 融合注意力机制的改进网络

在人工智能领域,注意力模型 (AM) 已经成为神经网络结构的重要组成部分,并广泛应用于自然语言处理、统计学、图像及语音识别等不同的深

度学习任务中。在深度学习中,注意力机制的引入类似于人类视觉选择性关注的本质。无论是从智能系统还是生物角度看,在筛选所关心内容并抑制干扰信息方面都显示出相似之处,其目的在于提高效率和准确性。

因此,针对传统的目标检测算法不能有效解决复杂环境背景下的矿井输送带煤矸石检测问题,本文在原始的YOLOv8网络基础上融合注意力模型对其进行检测。该注意力模型是一种同时结合空间和通道的注意力机制模块CBAM(convolutional block attention module)^[13],通过在原始YOLOv8网

络的每个残差块后融合CBAM注意力模型用于对目标信息进行精准化的分配和处理,可有效提高原始网络的检测精度。

融合通道和空间注意力机制的网络模型如图2所示。其中Mc代表一维的通道注意力特征图,Ms代表二维的空间注意力特征图,并且与YOLOv8输入的特征图相乘以细化模型对特征的提取性能。由于CBAM是一个轻量级的模块,将其集成到CNN框架中,所占用的计算量可以忽略不计,并且可以融合到CNN网络中进行端到端的训练,实验结果表明该模块的加入提高了网络的检测性能。

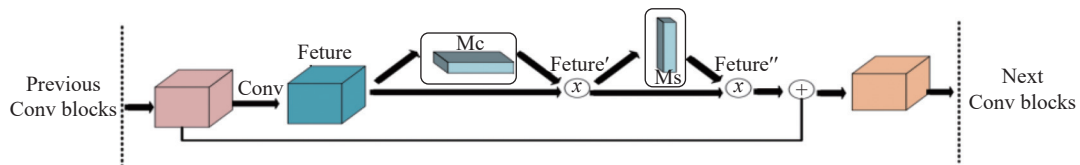


图2 融合通道和空间注意力机制的网络模型

2.2 软化非极大值抑制

在对目标图像进行检测时,边界框(bounding box)在经过特征提取和分类器分类识别后,很容易会产生重叠的现象,对边界框和相应的置信度,设定某一阈值 N_t ,当重叠度高于设定阈值时,将置信度分数最高的边界框作为最终的输出结果,而其他的边界框将被抑制,这种方法称为非极大值抑制(non-maximum suppression, NMS)^[14]。由于煤矿输送带上的煤和煤矸石在输送过程中存在相互重叠、遮挡等问题,而检测网络中的NMS很容易将重叠度较高的边界框分数自动置为0,造成待检测目标物被漏检的情况。

为了改善NMS在应对目标物相互遮挡情况时的不适应性,本文采用高斯加权函数对原公式进行改进,用软化非极大值抑制(Soft-NMS)的方法替换非极大值抑制(NMS)。该得分衰减机制的引入有助于在模型训练期间逐渐降低某些初始得分较高的边界框权重,这种调整能够增加检测阶段对可能存在互相遮挡目标的识别准确性。

改进后的重置函数如式(1)所示。

$$s_i = s_i e^{-\frac{IoU(M, b_i)^2}{\sigma}} \quad (1)$$

式中: σ 为常数, s_i 表示边界框的得分, M 表示当前得分最高的框, b_i 表示第 i 个边界框,IoU(inter section over union)表示重叠率,实验结果表明,采用高斯加权函数后,网络检测精度有所提

高,对具有相互遮挡的目标有很强的适应性。

2.3 改进网络卷积层

为了更好地平衡改进后网络在融合注意力及优化NMS后造成的检测速度低的问题,本文在主干网络的原有卷积中引入深度可分离卷积(Depthwise Separable Convolution)^[15]以提升整个网络的检测速度,其结构如图3所示。

图中, M 表示网络输入的通道个数, N 表示网络输出的通道个数; D_x 表示网络输入长度, D_y 表示网络输入宽度; D_k 表示网络的卷积核大小; D_w 表示网络输出长度, D_h 表示网络输出宽度。其计算量用 Q_1 表示,整个过程如下:

$$Q_1 = D_k^2 \cdot M \cdot N \cdot D_w \cdot D_h \quad (2)$$

假设输入的特征图 F 为 $D_x \times D_y \times M$,经过深度卷积操作后得到 $D_h \times D_w \times N$ 的特征图 G ,其计算量 Q_2 为:

$$Q_2 = D_k^2 \cdot M \cdot D_w \cdot D_h + M \cdot N \cdot D_w \cdot D_h \quad (3)$$

由公式(2)和(3)可以推出,深度可分离卷积与标准卷积之间的计算量比值为:

$$\frac{Q_2}{Q_1} = \frac{1}{N} + \frac{1}{D_k^2} \quad (4)$$

在网络模型中,设计了大小为 3×3 的卷积核,其中输入通道数为3,输出通道数量达到了256。经过对卷积的改进不仅有助于提高模型的效率,还减

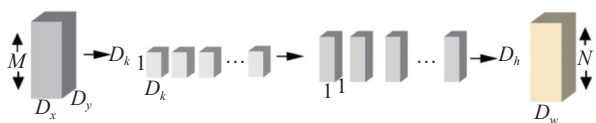


图 3 深度可分离卷积模型结构

少了网络运行过程中所需的参数,使得模型能够更快地完成目标识别与处理任务。

2.4 待检测图像超分辨率重建

为了有效解决图像质量不佳可能导致网络难以有效检测煤矸石的问题,本文采用 SRCNN^[16] 技术对图像进行超分辨率重建处理,以提高待检测图像质量,网络模型如图 4 所示。

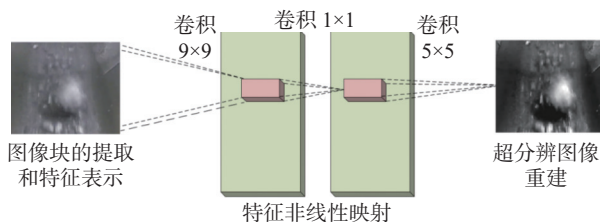


图 4 基于 SRCNN 的超分辨率重建

从图 4 可看出,整个系统架构由 3 层卷积神经网络(CNN)组成,分别采用 \$9 \times 9\$、\$1 \times 1\$ 和 \$5 \times 5\$ 大小的卷积核,具体处理步骤如下。

2.4.1 图像块特征提取

该层网络主要作用是提取输入图像的低层次局部特征,将图 4 中输入的低分辨率图像记为 \$Y\$,通过训练得到相应的映射图像记为 \$F(Y)\$,使用卷积操作来实现,通过一定数量的卷积核对经过双 3 次插值处理后的低分辨率图像进行卷积运算,得到一系列特征图。如式(5)所示。

$$F_1(Y) = \max(0, W_1 \times Y + B_1) \quad (5)$$

式中: \$W_1\$ 和 \$B_1\$ 分别代表网络模型第 1 层卷积核权重和偏置。

2.4.2 网络非线性映射

将第 1 层提取到的低层次特征转换为高层次特征。该层使用 \$1 \times 1\$ 的卷积操作来实现向量维数的变换,其本质是对输入的特征图进行非线性变换,增强特征的表达能力。经过这一层的处理,特征图能够更准确地表示高分辨率图像的特征信息。如式(6)所示。

$$F_2(Y) = \max(0, W_2 \times F_1(Y) + B_2) \quad (6)$$

式中: \$W_2\$ 和 \$B_2\$ 分别代表网络模型第 2 层卷积核权重和偏置。

2.4.3 高分辨率图像重建

网络第 3 层是重构层,用于将经过非线性映射后的特征图组合起来,重建出高分辨率图像。该层同样使用卷积操作,将前 2 层得到的高层次特征转换为最终的高分辨率图像输出。在训练过程中,该层的卷积核权重和偏置也会不断调整,以提高重建图像的质量,如式(7)所示。

$$F(Y) = W_3 \times F_2(Y) + B_3 \quad (7)$$

式中: \$W_3\$ 和 \$B_3\$ 分别代表网络模型第 3 层卷积核权重和偏置。

3 模型训练及结果分析

软硬件平台配置:处理器为 Intel® Core™ i5-9400F CPU @ 2.90GHz ×6,操作系统为 ubuntu 16.04 LTS 版本;深度学习框架为 darknet 网络框架,上述设备的组合构成了实验所需的基本环境。

3.1 数据的采集与处理

使用的数据集均来自于王庄煤矿输送带运输过程中的实时监控视频,并借助 labelImg 标注工具对截取的样本图片进行标注,最后共标注了 2000 张图片的数据集,从中任选 1500 张目标图片作为训练样本,其余 500 张目标图片作为测试样本。并且为了直观的看出待检测的目标,用该标注工具将目标物的名称进行自定义命名。

3.2 网络模型训练

为了验证本文提出的煤矸石检测算法的优越性,将 YOLOv8 和本文算法对比训练,网络输入的图像分辨率都设置为 \$640 \times 640\$,将训练中的学习率设置为 0.0001,衰减系数设置为 0.0005,并且使模型每迭代 500 步保存一次权重。

在训练过程中网络模型损失值函数逐渐趋向于收敛的态势。曲线变化呈现出模型性能的持续改善,可以明显看出本文模型对数据集的学习和拟合是有效的,平均损失值函数曲线如图 5 所示。

经过大约 2000 批次的迭代训练后,模型在验证集上实现超过 90% 的平均准确率。并且随着迭代次数的逐步增加网络检测精确度也呈稳步增加的趋势。当训练批次为 5000 左右时,平均准确率达到 95% 左右。因此,将训练批次在 5000 时的权重作为本文实验的检测权重,平均准确率变化曲线如图 6 所示。

表 1 是网络在加入不同策略后的试验结果对比。

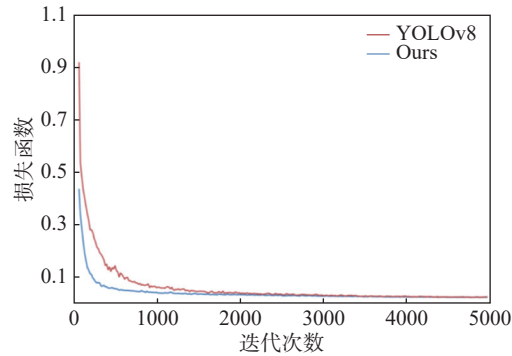


图 5 平均损失值函数曲线

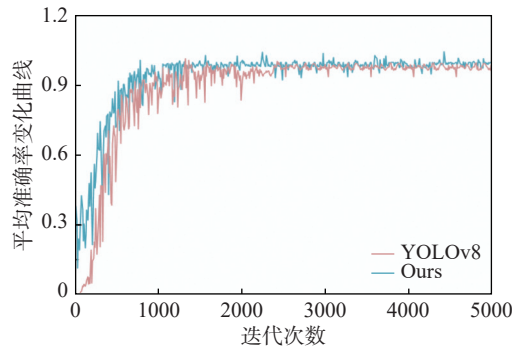


图 6 平均准确率变化曲线

从表中可以看出传统的 YOLOv8 网络检测精度为 92.3%，本文在传统算法的基础上经过 4 个方面的优化，检测精度得到明显的提升，证明了改进的煤矸石检测方法的有效性。

表 1 本文加入的各策略的效果对比

NMS	CBAM	DSC	Soft NMS	SRCNN	mAP/%
√					92.3
√	√				93.6
√		√			92.5
			√		92.8
√				√	93.1
	√	√	√	√	94.3

3.3 实验结果及分析

召回率和准确率是目标检测领域经常使用的评价指标。本文选取了准确率 M_p ，召回 M_f 以及运算速度 M_o 作为本文实验的客观评价指标。各指标定义如公式（8）-（10）所示。

$$M_p = \frac{TP}{TP + FP} \tag{8}$$

$$M_f = \frac{TP}{TP + FN} \tag{9}$$

$$M_o = S_k^s - S_k^t \tag{10}$$

式中： TP 表示能够正确检测出的目标， FN 表示没有被检测出的目标， FP 表示被误检的目标， S_k^s, S_k^t 分别表示网络对煤矸石检测的开始和结束的时间。不同网络的检测结果见表 2。

表 2 不同网络检测结果对比

模型	$M_p/\%$	$M_f/\%$	$M_o/(\text{ms/frame})$
SSD	85.3	85.2	3.1309
Faster-RCNN	88.7	88.3	0.0539
YOLOv8	92.3	95.3	0.026
Ours	94.3	95.8	0.028

从上述表格的对比算法检测结果可以看出，本文算法在检测精度及召回率上均高于其他 3 种网络，检测速度虽然略低于 YOLOv8，但相比于其他对比算法具有更高的检测速度。

将本文算法在测试集上进行测试，并选取复杂背景环境、皮带高速运动下及存在遮挡情况等场景，其检测结果如图 7（a）-（c）所示。

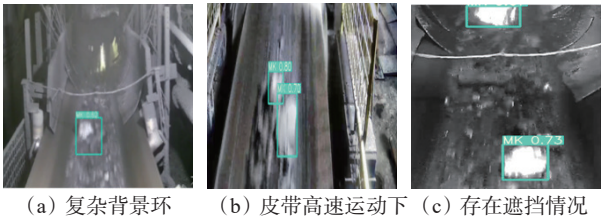


图 7 本文算法检测结果

从实验结果可看出，在煤尘较大的复杂背景环境下，本文算法可以准确的识别出待检测目标；在高速运动的皮带上，检测图像存在模糊失真的现象，而本文所提出的检测网络在针对高速运动的目标时也保持了较高的检测精度。

4 结 论

针对煤矿输送带高速运转、复杂环境干扰以及部分目标物存在遮挡等情况造成煤矸石难以准确检测的问题，本在 YOLOv8 网络模型的基础上，通过融合注意力机制模块、优化锚点参数和网络卷积层等方法，极大地提高了网络模型在煤矸石检测上的鲁棒性和实时性，可以有效解决智能选矸过程中存在的检测精度低的问题。

参考文献:

- [1] 赵卫.井下煤矸石分选系统可行性研究分析[J]. 煤, 2021, 30(2): 100-101.
- [2] 申龙, 单浩然, 王永利, 等. 智能煤矸石分选机器人的研究与应用[J]. 中国煤炭, 2023, 49(10): 74-78.
- [3] 石亦琨, 李峥, 李润田, 等. 基于机器视觉的煤矸石分选方法研究[J]. 中国矿业, 2024, 33(1): 114-121.
- [4] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1440-1448.
- [5] 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.
- [6] 吴守鹏, 丁恩杰, 俞啸. 基于改进 FPN 的输送带异物识别方法[J]. 煤矿安全, 2019, 50(12): 127-130.
- [7] 杨文斌. 基于 Faster-RCNN 算法的刮板输送机异物识别技术研究[J]. 煤矿机械, 2022, 43(11): 54-56.
- [8] Wei L, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of the 14th European Conference. Netherlands: Springer, 2016: 21-37.
- [9] 左明成, 焦文华. 面向煤矿井下作业场景的安全帽佩戴识别算法[J]. 中国安全科学学报, 2024, 34(3): 237-246.
- [10] 李珍峰, 符世琛, 徐乐, 等. 基于 MBI-YOLOv8 的煤矸石目标检测算法研究[J/OL]. 图学学报, 1-12[2024-09-30]<http://kns.cnki.net/kcms/detail/10.1034.T.20240906.1856.004.html>.
- [11] REDMON J, ALI Farhadi. YOLOv3: An incremental improvement[EB/OL][2018-04-08]. <https://arxiv.org/pdf/1804.02767.pdf>.
- [12] Alexey B, Chien-Yao W, Hong-Yuan Mark L. YOLOv4: Optimal Speed and Accuracy of Object Detection [EB/OL]. [2018-09-30]. <https://arxiv.org/abs/2004.1093>.
- [13] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//The 15th European Conference. Munich, Germany: Springer, 2018: 3-19.
- [14] NEUBECK A, VAN GOOL L. Efficient non-maximum suppression[C]//Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06). Hong Kong: IEEE, 2006: 850-855.
- [15] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C] //Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [16] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution[C]//13th European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014: 184-199.

作者简介:

组少杰 (1982-), 男, 辽宁彰武人, 硕士, 工程师, 研究领域: 煤矿机电管理、电力系统状态监测与故障诊断、基于图像的目标检测与识别。E-mail: 1473251806@qq.com

(收稿日期: 2024-11-11)

(上接第 63 页)

2) 通过观测覆岩垮落形态和裂隙带发育高度, 发现 2 种不同顶板岩性煤层的采动覆岩运动规律相似, 顶板沿工作面顺槽方向呈马鞍形垮落, 在切眼处裂隙带发育高度最高, 顶板在截水孔上呈现平行于分层路径的梯形垮落。

参考文献:

- [1] 赵潇辉. 马兰矿 18301 工作面“两带”研究分析[J]. 江西煤炭科技, 2023(1): 96-99.
- [2] 李志华, 耿倩, 杨科, 等. 综采工作面冒落带注浆充填开采覆岩采动裂隙定量表征试验研究[J]. 中国矿业, 2024, 33(2):

159-167.

- [3] 侯晓松, 梁超, 刘士磊, 等. 厚坚硬顶板工作面覆岩运动规律与采动应力分布研究[J]. 煤, 2023, 32(12): 15-19.
- [4] 鲍银辉. 复合煤层开采覆岩裂隙带发育特征[J]. 陕西煤炭, 2024, 43(3): 74-78.
- [5] 陈维池. 煤层开采上覆巨厚白垩系离层动态发育及含水层响应机制[D]. 中国矿业大学, 2023.

作者简介:

张敏 (1989-), 女, 山西太原人, 高级工程师, 从事矿山地质工作。E-mail: zhangminminbao@163.com

(收稿日期: 2024-4-7)