



基于CFS-YOLO算法的复杂工况环境下煤矸图像识别方法

李德永 王国法 郭永存 王爽 杨宇豪

引用本文:

李德永, 王国法, 郭永存, 等. 基于CFS-YOLO算法的复杂工况环境下煤矸图像识别方法[J]. 煤炭科学技术, 2024, 52(6): 226–237.

LI Deyong, WANG Guofa, GUO Yongcun. Image recognition method of coal gangue in complex working conditions based on CES-YOLO algorithm[J]. Coal Science and Technology, 2024, 52(6): 226–237.

在线阅读 View online: <https://doi.org/10.12438/cst.2023-1967>

您可能感兴趣的其他文章

Articles you may be interested in

改进YOLOv5s的采煤机滚筒与支架护帮板干涉状态智能识别

Interference state intelligent recognition method for shearer drum and hydraulic support guard plate of improved YOLOv5s algorithm
煤炭科学技术. 2024, 52(2): 253–263 <https://doi.org/10.12438/cst.2023-0757>

基于改进DeeplabV3+和迁移学习的煤岩界面图像识别方法

Coal-rock interface image recognition method based on improved DeeplabV3+ and transfer learning
煤炭科学技术. 2023, 51(S1): 429–439 <https://doi.org/10.13199/j.cnki.cst.2022-1392>

基于TW-RN优化CNN的煤矸识别方法研究

Research on coal and gangue recognition method based on TW-RN optimized CNN
煤炭科学技术. 2022, 50(1): 228–236 <http://www.mtkxjs.com.cn/article/id/317225f8-61ed-4bdd-a8c8-97c03d0bedac>

基于深度学习的煤矸石识别技术与实现

Coal and gangue recognition technology and implementation based on deep learning
煤炭科学技术. 2021, 49(12): 202–208 <http://www.mtkxjs.com.cn/article/id/5f975297-7e00-4ebe-aa8d-a4f4808d034b>

基于几何特征约束的煤矸DE-XRT精准识别方法

Accurate identification method of coal and gangue based on geometric feature constraints by DE-XRT
煤炭科学技术. 2024, 52(5): 262–275 <https://doi.org/10.12438/cst.2023-1276>

改进YOLOv5s和DeepSORT的井下人员检测及跟踪算法

Underground personnel detection and tracking based on improved YOLOv5s and DeepSORT
煤炭科学技术. 2023, 51(10): 291–301 <https://doi.org/10.13199/j.cnki.cst.2022-1933>



关注微信公众号, 获得更多资讯信息



移动扫码阅读

李德永, 王国法, 郭永存, 等. 基于 CFS-YOLO 算法的复杂工况环境下煤矸图像识别方法[J]. 煤炭科学技术, 2024, 52(6): 226–237.

LI Deyong, WANG Guofa, GUO Yongcun, *et al.* Image recognition method of coal gangue in complex working conditions based on CES-YOLO algorithm[J]. Coal Science and Technology, 2024, 52(6): 226–237.

基于 CFS-YOLO 算法的复杂工况环境下 煤矸图像识别方法

李德永^{1,2}, 王国法^{1,2,3}, 郭永存^{1,2}, 王 爽^{1,2}, 杨宇豪^{1,2}

(1. 安徽理工大学 深部煤矿采动响应与灾害防控国家重点实验室, 安徽 淮南 232001; 2. 安徽理工大学 矿山智能技术与装备省部共建协同创新中心, 安徽 淮南 232001; 3. 中国煤炭科工集团有限公司, 北京, 100013)

摘 要: 针对煤矿高噪声、低照度、运动模糊与大批量煤矸混杂等复杂工况环境因素导致煤矸识别存在误检、漏检以及检测精度低的问题, 提出一种基于 CFS-YOLO 算法的煤矸智能识别模型。采用 ConvNeXt V2(Convolutional Neural Network with NeXt Units Version 2)特征提取模块替换主干网络末端的 2 个 C3(Cross Stage Partial Bottle Neck Mudule)模块, 通过将掩码自动编码器(Masked Autoencoders, MAE)和全局响应归一化(Global Response Normalization, GRN)层添加到 ConvNeXt 架构中, 有效缓解特征崩溃问题以及保持特征在网络传递过程中的多样性; 采用 Focal-EIOU(Focal and Efficient Intersection Over Union)损失函数替换原 CIOU(Computer Intersection Over Union)损失函数, 通过其 Focal-Loss 机制和调整样本权重的方式优化边界框回归任务中的样本不平衡问题, 提高模型的收敛速度和定位精度; 添加无参注意力机制(Simple Attention Mechanism, SimAM)于主干网络每个 C3 模块的后端, 凭借其注意力权重自适应调整策略, 提升模型对尺度变化较大或低分辨率煤矸目标关键特征的提取能力。通过消融试验和对比试验验证所提 CFS-YOLO 模型的有效性与优越性。试验结果表明: CFS-YOLO 模型对于煤矸在煤矿高噪声、低照度、运动模糊与大批量煤矸混杂等复杂环境下的检测效果均得到有效提高, 模型的平均精度均值达到 90.2%, 相较于原 YOLOv5s 模型的平均精度均值提高了 3.7%, 平均检测速度达到 90.09 FPS, 可充分满足煤矸实时检测的需求。同时与 YOLOv5s、YOLOv7-tiny 与 YOLOv8n 等 6 种 YOLO 系列算法相比, CFS-YOLO 模型对煤矿复杂环境的适应性最强且综合检测性能最佳, 可为煤矸的智能高效分选提供技术支持。

关键词: YOLOv5s; 煤矸识别; 特征提取; 损失函数; 注意力机制

中图分类号: TD94; TP242.2 **文献标志码:** A **文章编号:** 0253-2336(2024)06-0226-12

Image recognition method of coal gangue in complex working conditions based on CES-YOLO algorithm

LI Deyong^{1,2}, WANG Guofa^{1,2,3}, GUO Yongcun^{1,2}, WANG Shuang^{1,2}, YANG Yuhao^{1,2}

(1. State Key Laboratory of Deep Coal Mining Response and Disaster Prevention and Control, Anhui University of Science and Technology, Huainan 232001, China; 2. Collaborative Innovation Center for Mining Intelligent Technology and Equipment, Huainan 232001, China; 3. China Coal Technology & Engineering Group, Beijing 100013, China)

Abstract: Aiming at the complex working conditions environmental factors such as high noise, low illumination, motion blur and mass gangue mixing in coal mines, which lead to the problems of misdetection, omission and low detection accuracy in gangue recognition, a gangue recognition model based on CFS-YOLO algorithm is proposed. The ConvNeXt V2(Convolutional Neural Network with NeXt Units Version 2)feature extraction module is adopted to replace the two C3(Cross stage partial bottle neck mudule)modules at the end of

收稿日期: 2023-12-23 责任编辑: 周子博 DOI: 10.12438/cst.2023-1967

基金项目: 安徽理工大学高层次引进人才科研启动基金资助项目(2023yjrc56); 安徽省自然科学基金资助项目(2308085Y37); 国家自然科学基金资助项目(52274152)

作者简介: 李德永(1991—), 男, 安徽萧县人, 讲师, 硕士生导师, 博士。E-mail: ldzyiou@126.com

the backbone network, which effectively mitigates the feature collapse problem as well as maintains the diversity of the features in the network delivery process by adding Masked Autoencoder and Global Response Normalization layers to the ConvNeXt architecture. The Focal-EIOU (Focal and Efficient Intersection Over Union) loss function is adopted to replace the original CIOU (Computer Intersection Over Union) loss function to optimize the sample imbalance problem in the bounding box regression task by means of its Focal-Loss mechanism and adjusting the sample weights, which improves the convergence speed and localization accuracy of the model. The parameter-free attention mechanism (Simple Attention Mechanism, SimAM) is added to the back-end of each C3 module of the backbone network to enhance the model's ability of extracting key features of coal gangue targets with large scale variation or low resolution by virtue of its attention weight adaptive adjustment strategy. The effectiveness and superiority of the proposed CFS-YOLO model is verified by ablation and comparison experiments. The experimental results show that the CFS-YOLO model can effectively improve the detection effect of coal gangue under the complex environment of high noise, low illumination, motion blur and large amount of mixed coal gangue in coal mines. The mean Average Precision (mAP) of the model reaches 90.2%, which is 3.7% higher than the mean Average Precision (mAP) of the original YOLOv5s model, and the average detection speed reaches 90.09 FPS (Frames Per Second), which can fully satisfy the demand of real-time detection of coal gangue. Meanwhile, compared with six YOLO algorithms such as YOLOv5s, YOLOv7-tiny and YOLOv8n, the CFS-YOLO model has the strongest adaptability to the complex environment of coal mines and the best comprehensive detection performance, which can provide technical support for intelligent and efficient sorting of coal gangue.

Key words: YOLOv5s; coal gangue recognition; feature extraction; loss function; attention mechanism

0 引 言

煤炭是当前社会的主体能源,是可以清洁高效利用的最经济安全的能源,而煤矿智能化是煤炭高质量发展的必由之路^[1-3]。原煤在开采过程中混有大量矸石,这不仅会提高选煤和筛煤的成本,而且燃烧排放的有害气体会污染环境,因此亟需将矸石从原煤中分选出来^[4-5]。随着人工智能技术的快速发展,我国选煤行业正向工艺简单化、设备智能化、分选精细化的方向发展^[6-7],煤矸智能分选的研究已成为煤矸分选的重要课题,其中煤矸的精准快速识别是实现煤矸分选的首要任务。近年来,国内外专家学者针对煤矸识别已经做出了大量的研究,煤矸识别常用的方法主要有红外热成像识别^[8-9]、多光谱识别^[10-11]、射线识别^[12-13]和图像识别^[14]等,红外热成像识别、多光谱识别和射线识别都存在一定局限性,未能大规模推广使用。而随着计算机视觉技术的显著进步,目前图像法^[15-16]因其具有非接触式识别、高效性、准确性和多样性的优点,在煤矸识别领域的应用更为广泛,其中基于深度学习^[17-18]的图像法具有无需特征选择、特征表达能力强、检测速度快、检测精度高等优势,从而逐渐成为了研究热点。徐志强等^[19]基于 ResNet 等经典网络与 SqueezeNet 等先进轻量级网络建立了煤矸图像识别模型,分析了各模型的训练收敛情况,实现了模型的压缩。郭永存等^[20]提出一种基于 TW-RN 优化深度卷积神经网络(CNN)的煤矸识别模型,通过迁移权重和简化神经元模型优化方法对 4 种模型进行改进,依据每种模型取不同超参数时的训练结果,进而确定每种网络的最优

超参数,最终得到效果最好的 Im-AlexNet 煤矸识别模型。XUE 等^[21]提出一种基于 ResNet18-YOLO 的煤矸石检测算法,通过特征尺度缩减和非结构化剪枝,在保持检测精度的同时进一步提高其轻量级和实时性。韦小龙等^[22]提出一种基于 CSPNet-YOLOv7 的煤矸检测模型,通过在 YOLOv7 主干网络引入跨阶段部分模块和颈部引入递归特征金字塔与可切换卷积,以提高煤矸的识别精度。WANG 等^[23]提出一种多目标煤矸石图像检测模型 SSD-BSP,通过将 BSP 模型与 SSD 模型相结合,可以实现煤矸石快速准确检测。曹现刚等^[24]利用迁移学习改进 AlexNet 特征提取网络,并结合 RPN(Region Proposal Network)网络获取煤和矸石的分类信息和像素坐标,检测精度达到 90% 以上。

综上所述,不少专家学者对煤矸识别做了深入研究,但是煤矿实际工况环境复杂,基于深度学习的无监督视频 AI 识别关键技术及理论方面尚不成熟,在识别精度和速度方面均达不到煤炭工业现场实际应用需求^[25]。针对煤矿高噪声、低照度、运动模糊与大批量煤矸混杂等恶劣环境导致煤矸识别存在误检、漏检以及识别精度低的问题,提出一种 CFS-YOLO 煤矸识别模型,可实现煤矿高噪声、低照度、运动模糊与大批量煤矸混杂等复杂工况下煤矸精准实时检测任务,从而为煤矸的智能高效分选提供技术支持。

1 CFS-YOLO 模型

1.1 ConvNeXt V2 模块

掩码自动编码器 MAE(Masked Autoencoders)^[26]是一种用于计算机视觉领域的自监督学习方法,在

预训练过程中会随机屏蔽部分像素,其中,编码器仅处理可见像素,解码器使用编码像素和掩码标记重建图像。预训练结束后,解码器停止工作,编码器处理未处理的图像,并将其学习到的关键特征用于识别图像。掩码自动编码器的整体结构如图1所示,它由一个基于稀疏卷积的 ConvNeXt(Convolutional Neural Network with NeXt Units)编码器和一个轻量级的 ConvNeXt 解码器组成,以构建一个非对称的编码器-解码器结构。

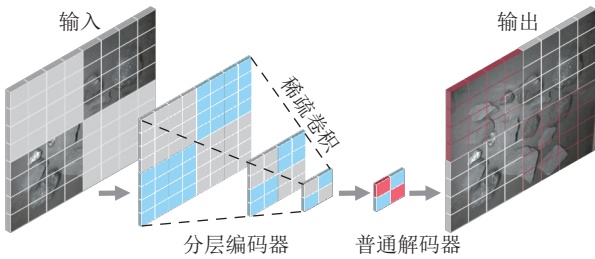


图1 掩码自动编码器结构
Fig.1 Masked autoencoder structure

采用 ConvNeXt V2 特征提取模块^[26],该模块是将掩码自动编码器 MAE 和全局响应归一化 GRN (Global Response Normalization) 层添加到 ConvNeXt 架构中,以增强通道间的特征竞争。ConvNeXt V1 和 ConvNeXt V2 模块的结构如图2所示,通过在维度扩展 MLP 层之后添加全局响应归一化 GRN 层,并删除 LayerScale^[27],以提高纯 ConvNets 在煤矸识别基准上的性能。

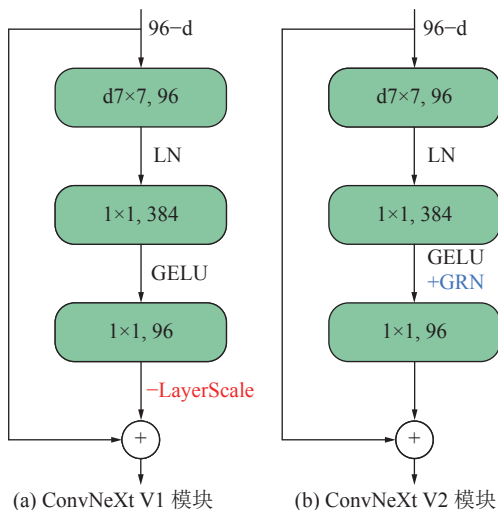


图2 ConvNeXt V1 和 ConvNeXt V2 结构
Fig.2 ConvNeXt V1 and ConvNeXt V2 structures

1.2 Focal-EIOU 损失函数

目前 YOLOv5 采用 CIOU(Computer Intersection Over Union)^[28] 作为边界框的定位损失函数,它

综合考虑了3个重要的几何因素:重叠面积、中心点距离和纵横比。CIOU 损失函数虽然充分考虑到预测框的纵横比,解决了预测框形状与真实框不一致问题,但是其惩罚项并不完全对应于边界框宽度和高度与其置信度之间的真实差值,导致收敛到预测框和真实框宽度和高度之间的线对线比率时,惩罚项会增加相似预测框和真实框间的损失,从而限制了模型对两框相似性的有效优化。

因此,采用 Focal-EIOU(Focal and Efficient Intersection Over Union)^[29] 作为边界框的损失函数,Focal-EIOU 综合了 Focal Loss 和 EIOU(Efficient Intersection Over Union)的优势,在 CIOU 的基础上将纵横比拆开,并加入 Focal 聚类优质的锚框,其计算公式如式(1)、式(2)所示:

$$\text{Loss}_{\text{EIOU}} = 1 - \text{IOU} + \frac{\rho^2(b, b^{\text{gt}})}{C^2} + \frac{\rho^2(w, w^{\text{gt}})}{C_w^2} + \frac{\rho^2(h, h^{\text{gt}})}{C_h^2} \quad (1)$$

$$\text{Loss}_{\text{Focal-EIOU}} = \text{IOU}^\gamma \times \text{Loss}_{\text{EIOU}} \quad (2)$$

式中: γ 为控制异常值抑制程度的参数,一般为 3; C_w 和 C_h 为最小外接矩形框的宽度和高度。

Focal-EIOU 损失函数引入 Focal-Loss 优化了边界框回归任务中的样本不平衡问题,同时,引入 EIOU 使得预测框与真实框的宽度和高度差异最小化,从而进一步提高模型的收敛速度和定位精度。

1.3 SimAM 无参注意力机制

注意力机制可以使神经网络更加关注煤矸自身的特征信息,忽略背景等非重要信息。由于通道注意力机制生成的一维权重和空间注意力机制生成的二维权重对于每个通道或位置中神经元的关注较为均衡,从而限制模型学习更具区分性线索的能力。而 SimAM(Simple Attention Mechanism)无参注意力机制^[30]是具有全三维权重的注意力机制,通过发掘每个神经元的重要性来计算注意力权重,从而更好的细化特征,其整体结构如图3所示,图中 X 为特征向量, C 为通道数, W 为宽度, H 为高度。

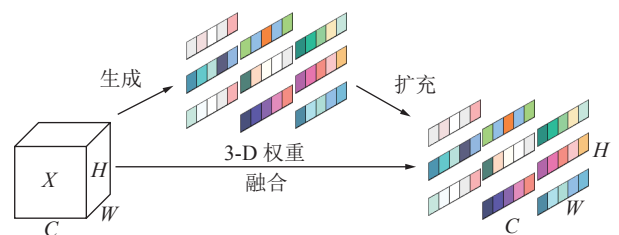


图3 SimAM 无参注意力机制结构
Fig.3 SimAM parameter-free attention mechanism structure

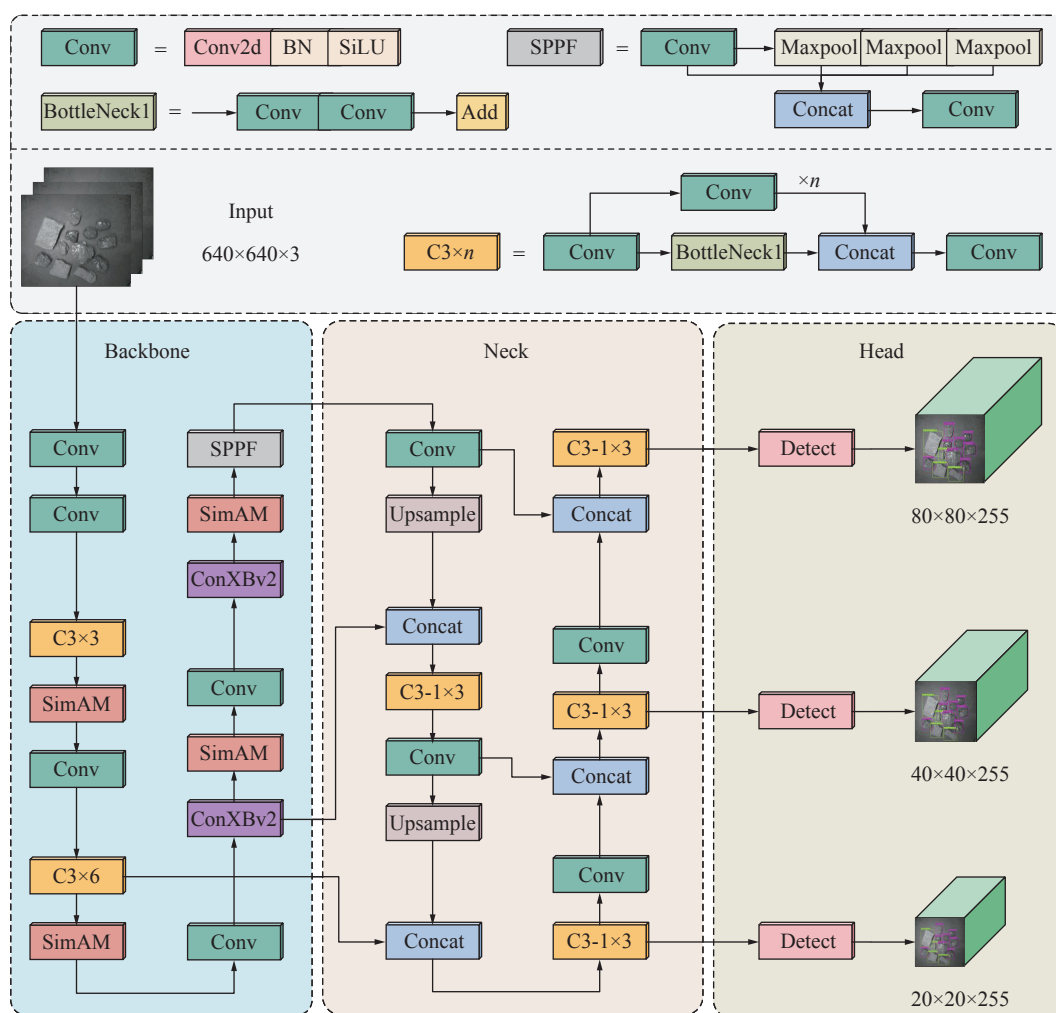
SimAM 无参注意力机制主要是基于视觉神经科学理论设计一种能量函数,表现出明显空间抑制效应的神经元应该被给予更高的优先级,通过度量目标神经元和其他神经元之间的线性可分性,以寻找重要神经元。同时能量越低,神经元与周围神经元的区别越大,重要性越高。

1.4 CFS-YOLO 模型结构

CFS-YOLO 模型主要是在 YOLOv5s 模型的基础上进行改进,其网络结构如图 4 所示,模型主要由输入端(Input)、主干网络(Backbone)、颈部(Neck)和头部(Head) 4 部分组成。

主干网络部分采用 ConvNeXt V2 特征提取模块来替换 YOLOv5s 主干网络末端的 2 个 C3 模块,

ConvNeXt V2 特征提取模块采用多分支结构和空间变换网络,多分支结构可以同时学习多种尺度和方向的特征,从而减少噪声的影响,提高特征提取的稳定性和准确性;空间变换网络可以对输入图像进行准确的空间变换,从而提高对运动模糊的鲁棒性。同时,在主干网络每个 C3 模块的后端嵌入 SimAM 无参注意力机制,在高噪声环境中,由于噪声的影响,图像中有效信息较少,而 SimAM 无参注意力机制通过自适应关注相似度较高的区域特征,从而提高特征提取的准确性。在低照度环境中,SimAM 无参注意力机制通过计算特征图中的相似度,对于特别黑暗的部分给予更大的注意力权重,以帮助网络更好地还原低照度图像的细节。



注:Conv-卷积操作; Concat-特征融合; BN-批归一化操作; SiLU-损失函数; SPPF-空间金字塔池化结构; Maxpool-最大池化处理; BottlenNeck-残差结构; Upsample-上采样处理

图 4 CFS-YOLO 网络模型结构

Fig.4 CFS-YOLO network model structure

颈部部分主要采用特征金字塔 (Feature Pyramid Network, FPN) 结构和路径聚合网络 (Path Aggregation Network, PANet), 通过特征金字塔结构提供多

尺度的特征表达使模型获取更加丰富的特征信息,通过路径聚合网络聚合不同特征层间的信息路径从而实现特征的有效融合以及保持特征的完整性和

多样性。

头部部分采用 3 个不同大小的卷积模块作为检测层,然后应用锚框输出目标的置信度、类别和预测帧位置,最后采用 Focal-EIOU 损失函数替换原 CIOU 损失函数作为锚框的损失函数,其通过 Focal-Loss 优化边界框回归任务中的样本不平衡问题,通过 EIOU 使得预测框与真实框的宽度和高度差异最小化,以提高模型的收敛速度和定位精度。

2 数据采集与处理

2.1 煤矸识别分拣装置

在实验室搭建煤矸识别装置如图 5 所示,该试验装置主要包括振动给料机、输送机、CMOS 工业面阵相机、可调光源、照度计、调速装置、计算机和光源控制器。利用该装置的图像采集系统采集煤矸图像,CMOS 工业面阵相机型号为 Havel VisionM-VCA050-11 UM/UC,数据采集帧数为 35FPS,分辨率

为 2 448×2 048。通过调节光源控制器外加 2 条 LED 灯条作为辅助光源可将采集区域光强稳定在(3 600±20)Lux。采集系统与相机之间采用 USB3.0 接口通信,实时显示采集画面并保存煤矸图像。

2.2 数据采集与预处理

采集数据集分为 2 种,分别是井下工作面煤矸数据集和实验室自制煤矸数据集。井下工作面煤矸数据集采样于安徽省淮南市张集煤矿某工作面,共采集煤矸图像 1 200 张。实验室采集的煤矸样本来源于安徽省淮南市张集煤矿,通过调节光源照度、改变煤和矸石摆放位置、不同大小、输送带速度等组合采集煤矸图像 1 532 张。由于实际现场环境复杂,存在多因素干扰,通过图像增强软件 ImgAug3.2 对采集到的所有煤矸图像进行高斯模糊、运动模糊、高斯噪声、椒盐噪声和亮度调节等操作,模拟煤矿高噪声、低照度与运动模糊等实际情况,最终将数据集扩充为 11 856 张,图像变换效果对比如图 6 所示。

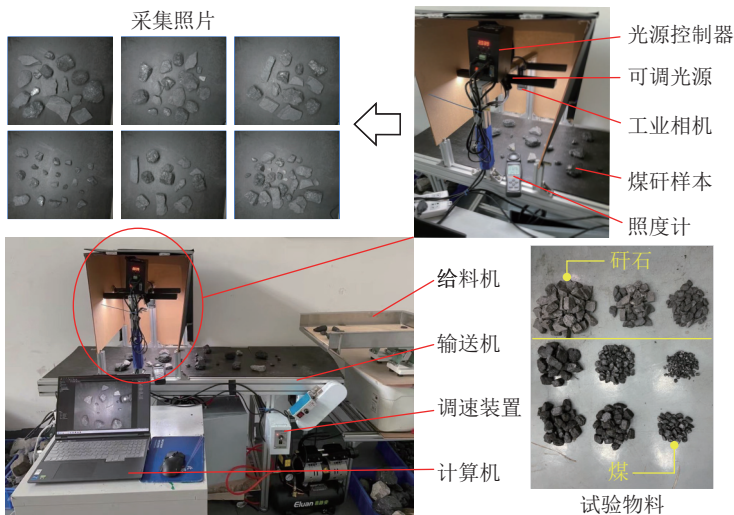


图 5 煤矸识别装置

Fig.5 Coal gangue identification device

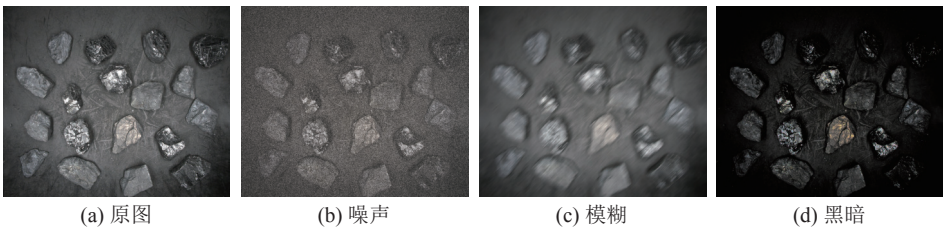


图 6 图像变换效果对比

Fig.6 Comparison of image transformation effect

将获取到的 11 856 张图片采用 Labellmg 图像标注工具进行手工标注,并采用“Coal”和“Gangue”作为煤和矸石 2 类检测目标的标签。11 856 张图片经过标注后共得到 125 976 个标签,其中包括 65 116

个煤和 60 860 个矸石。将标注好的数据集保存为 Pascal VOC 格式的 XML 文件,最后将标注完成的数据集按 8 : 2 的比例随机划分为训练集和验证集,并从额外拍摄的图像中选取未经任何处理的 300 张图

像作为测试集,用于测试 CFS-YOLO 模型的检测性能。

3 试验与结果分析

3.1 模型训练

试验所有算法的训练以及测试均基于 AutoDL 云服务器平台进行,网络模型训练环境具体配置见表 1。

表 1 网络模型训练环境

Table 1 Network model training environment

配置参数	类型
中央处理器CPU	Intel(R) Xeon(R) Platinum 8255c
图形处理器GPU	NVIDIA GeForce RTX 3 090(24 GB)
加速环境	CUDA11.3
语言环境	Python3.8
深度学习框架	PyTorch1.11.0

网络模型训练之前,优化模型配置文件中的相关参数,以获取最优训练模型。图像输入尺寸设置为 640×640;训练批量大小设置为 32;线程数设置为 16;初始学习率设置为 0.01;训练迭代次数设置为 300。

3.2 模型评价指标

在该试验中,主要评估指标^[31]为:精确率(Precision)、召回率(Recall)、精度均值(Average Precision)、平均精度均值(Mean Average Presicion)平均检测速度(Average Detection Speed)和 F_1 值(F_1 score),部分评价指标公式如下:

$$P = \frac{TP}{N_{pred}} = \frac{TP}{TP + FP} \quad (3)$$

表 2 不同特征提取模块结果对比

Table 2 Comparison of results of different feature extraction modules

模型	层数	参数量	浮点运算次数	平均检测速度/FPS	平均精度均值/%
YOLOv5s	270	7 025 023	16.0	101.15	86.5
Swim Transformer	296	7 264 659	84.8	76.92	85.5
DenseNet	405	7 675 391	18.3	58.52	85.2
ConvNeXt V2	258	7 064 063	16.0	93.46	87.6

3.4 不同损失函数试验结果

原 YOLOv5s 模型采用的是 CIOU 作为预测框和真实框之间的损失函数,但由于 CIOU 损失函数在训练过程中没有分别考虑边界框宽度和高度与其置信度的真实差异,从而导致收敛速度缓慢。因此

$$R = \frac{TP}{N_{GT}} = \frac{TP}{TP + FN} \quad (4)$$

$$AP = \int_0^1 P(r)dr \quad (5)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_{(i)} \quad (6)$$

$$FPS = \frac{N_f}{T_t} \quad (7)$$

$$F_1 = \frac{2PR}{P + R} \times 100\% \quad (8)$$

式中: TP(True Positive)为将正样本预测为正确的数量; FP(False Positive)为将负样本预测为正确的数量; FN(False Negative)为将正样本预测为错误的数量; N_{pred} 为预测出的所有检测框的数量; N_{GT} 为所有真实框的数量; N_f 为检测到的图像总数量; T_t 为总检测时间。

3.3 不同主干网络试验结果

为验证 ConvNeXt V2 特征提取模块的优势,同时引入 Swim Transformer、DenseNet、ConvNeXt V2 三种特征提取模块对原 YOLOv5s 模型进行改进和训练,并将训练的结果进行分析,验证 ConvNeXt V2 特征提取模块的可行性,试验结果见表 2。

由表 2 可知,当采用 ConvNeXt V2 特征提取模块时,模型的平均精度均值相较于原 YOLOv5s 提高 1.1%,在参数量增加 39 040 的情况下,其平均检测速度仅降低了 7.69 FPS。而当采用 Swim Transformer 和 DenseNet 特征提取模块时,模型的平均精度均值相较于原 YOLOv5s 分别降低 1.0% 与 1.3% 参数量分别增加 239 636 和 650 368,平均检测速度分别降低 19.23 FPS 和 37.63 FPS,效果不佳。综上所述,ConvNeXt V2 特征提取模块的综合性能最佳。

本试验通过对比不同的 IOU 损失函数,分析各损失函数对模型性能的影响,试验结果见表 3,不同损失函数在训练集上的定位损失曲线和总损失曲线如图 7a 和图 7b 所示。

由表 3 可知,SIOU 损失函数和 Alpha-IOU 损失

函数相较于 CIOU 损失函数平均精度均值保持不变为 86.5%，而 EIOU 损失函数和 Focal-EIOU 损失函数相较于 CIOU 损失函数平均精度均值由 86.5% 提高到了 87.0%，层数和权重分别为 270 和 13.7，均保持不变。由图 7a 和图 7b 可知，Focal-EIOU 损失函数在定位损失和总损失上整体波动较小，相较于 CIOU、EIOU、SIOU 和 Alpha-IOU 四种损失函数，其损失下降更快且拥有更小的收敛损失。综上所述，Focal-EIOU 损失函数的收敛效果最好，具有更快的收敛速度和更高的定位精度。在目标检测任务中，通常会有 2 种类型的错误：漏检和错检。漏检是指模型没有检测到真实存在的目标，错检是指模型错误地将背景或其他不是目标的物体识别为目标。而 Focal-EIOU 损失函数通过 Focal 权重对错误进行加权处理，当预测错误时，Focal 权重会增大以放大错

误的影响，当预测正确时，Focal 权重会减小以降低正确预测的影响，从而使得模型更加关注难以检测的目标，并减轻背景对于误检的影响。通过平衡漏检和错检之间的权衡，模型可以更好地进行目标检测任务。

表 3 不同 IOU 损失函数试验结果对比

Table 3 Comparison of experimental results with different IOU loss functions

损失函数	精度均值/%		F_1	平均精度均值/%
	煤	矸石		
CIOU	87.2	85.7	0.83	86.5
EIOU	87.7	86.2	0.82	87.0
SIOU	87.6	85.4	0.82	86.5
Alpha-IOU	86.8	86.1	0.82	86.5
Focal-EIOU	87.9	86.1	0.83	87.0

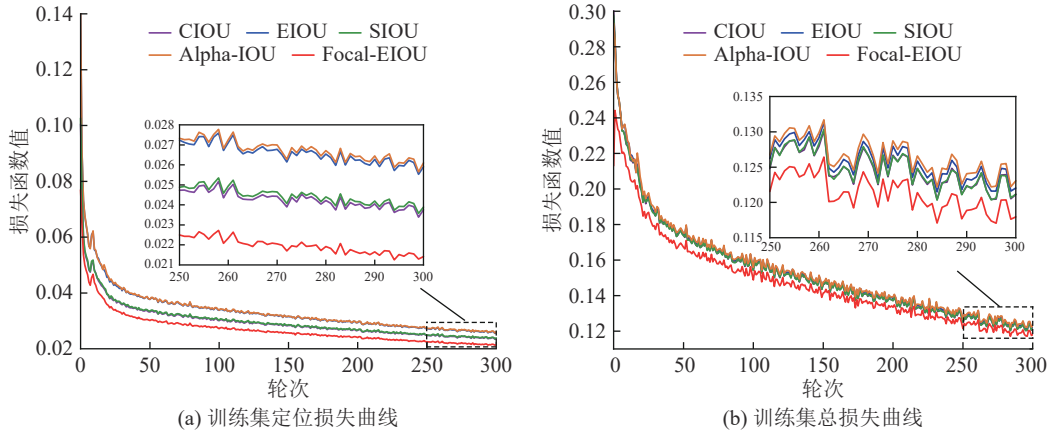


图 7 不同 IOU 损失函数曲线
Fig.7 Loss function curves for different IOU

3.5 添加注意力机制试验结果

通过在主干网络每个 C3 模块的后端嵌入 SimAM 无参注意力机制，凭借其注意力权重自适应调整策略，提升模型对尺度变化较大或低分辨率煤矸目标关键特征的提取能力。为了验证 SimAM 无参注意力机制的性能，在模型主干网络的相同位置处还

分别加入了 CA 注意力机制、CBAM 注意力机制、SE 注意力机制和 ECA 注意力机制进行对比试验，试验结果见表 4。从试验结果可以看出，SimAM 无参注意力机制在保持较低参数量和层数的同时，其平均精度均值最高达 88.7%。因此，添加 SimAM 无参注意力机制后模型效果最优。

表 4 不同注意力机制对比试验

Table 4 Comparative experiments of different attention mechanisms

注意力机制	精度均值/%		参数量	浮点运算次数	层数	平均精度均值/%
	煤	矸石				
CA	89.0	88.2	8 074 303	17.6	270	88.5
CBAM	89.1	87.8	7 721 735	17.1	314	88.5
SE	88.3	87.9	7 721 343	16.5	294	88.1
ECA	87.8	86.1	7 025 027	16.0	286	86.9
SimAM	89.2	88.1	7 025 023	16.0	278	88.7

同时,为更加直观地展现添加不同注意力机制的区别,采用 Grad-CAM(Gradient-weighted Class Activation Mapping)分别获取模型在最后 3 层的热力图,分别对应最终的检测层。Grad-CAM 是神经网络可视化工具,可以根据不同部位的权重生成热力图,从而更为具体地描述模型的检测效果,其热力图检测效果如图 8 所示。

热力图中红色部分越深表示该区域对于最终预测结果的贡献越大,黄色部分关注度次之,蓝色部分

影响最小。由图 8 可知,YOLOv5s 模型和使用 ECA 注意力机制的模型对于煤矸图像的背景也给予了较大的权重,从而导致识别结果出错造成误检,而使用 CA 注意力机制、CBAM 注意力机制、SE 注意力机制和 SimAM 无参注意力机制的模型更加关注于煤炭区域的特征提取,对于其他部分则较少关注。但其对比发现,使用 SimAM 无参注意力机制的模型效果最优,进而证明了 SimAM 无参注意力机制的有效性。

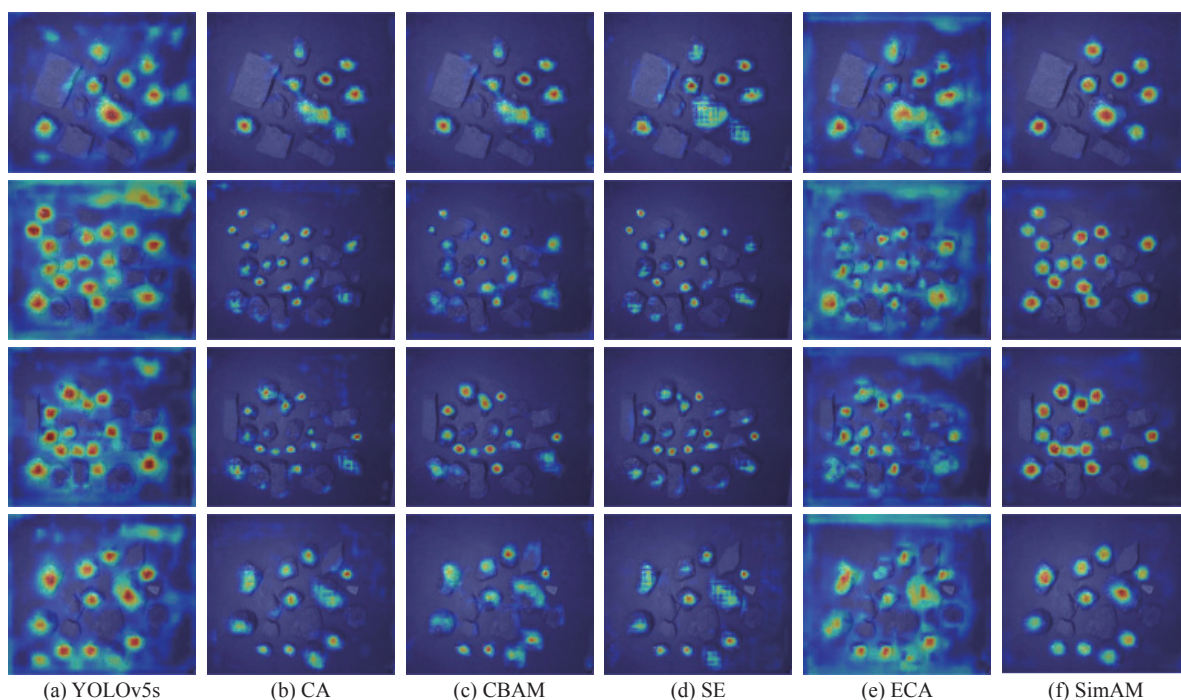


图 8 加入不同注意力机制热力图

Fig.8 Heat maps of different attention mechanisms added

3.6 消融试验

为了进一步提高 YOLOv5s 模型在煤矿复杂工况中的检测性能,主要采用 ConvNeXt V2 特征提取模块、Focal-EIOU 损失函数和 SimAM 无参注意力机制 3 种改进策略。为了验证各项改进策略的有效性,通过消融试验分析了各项改进策略对 YOLOv5s 模型检测性能的影响,其试验结果如图 9 和表 5 所示。

模型 A 是无任何改进策略的 YOLOv5s 模型,其平均精度均值为 86.5%,平均检测速度达到 103.15 FPS。模型 B 至模型 D 是分别加入 ConvNeXt V2 特征提取模块、Focal-EIOU 损失函数和 SimAM 无参注意力机制的模型,其平均精度均值相较于模型 A 分别提高了 1.1%、0.5% 和 2.2%,平均检测速度模型 C 最高达 101.01 FPS。模型 E 至模型 G 是依次加入两种改进策略的模型,其平均精度均值相较于模

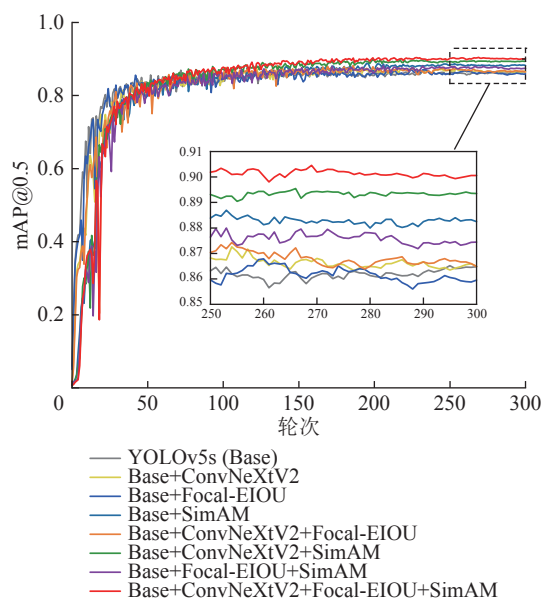


图 9 消融试验 mAP 曲线

Fig.9 Ablation experiment mAP curve

表 5 消融试验结果比较
Table 5 Comparison of results of ablation experiments

模型	方法			精度均值/%		精确率/%	召回率/%	F_1	平均检测速度/FPS	平均精度均值/%
	ConvNeXt V2	Focal—EIOU	SimAM	煤	矸石					
A	—	—	—	87.2	85.7	86.1	79.7	0.82	103.15	86.5
B	√	—	—	88.0	87.3	86.7	80.0	0.83	93.46	87.6
C	—	√	—	87.9	86.1	86.3	79.8	0.83	101.01	87.0
D	—	—	√	89.2	88.1	86.6	80.9	0.83	83.33	88.7
E	√	√	—	87.7	87.2	86.4	79.8	0.83	90.91	87.4
F	√	—	√	90.1	89.0	87.8	82.1	0.84	86.96	89.6
G	—	√	√	88.6	87.4	87.0	81.1	0.83	86.26	88.0
H	√	√	√	90.8	89.5	88.8	83.2	0.85	90.09	90.2

型 A 分别提高了 0.9%、3.1% 和 1.5%。平均检测速度相较于模型 A 均有所下降。模型 H 是同时加入 3 种改进策略的模型,其平均精度均值相较于模型 A 提高了 3.7%,效果最佳,平均检测速度为 90.09 FPS,仍满足实时检测的需求。

3.7 对比试验

为了客观评估提出的 CFS-YOLO 煤矸检测模型的综合检测性能,基于同一煤矸检测数据集和网络模型训练环境,将 CFS-YOLO 模型与其他经典检测模型(YOLO 系列模型、Faster R-CNN 模型、SSD 模型和 CenterNet 模型)进行训练和测试,其测试结果见表 6。

在表 6 中,分别从平均精度均值、模型复杂程度和平均检测速度 3 个方面对 7 种模型进行对比,具体分析如下:

1)平均精度均值: CFS-YOLO 模型的平均精度均值为 90.2%,在 7 种模型中最高,相较于 YOLOv5s、YOLOv7-tiny 和 YOLOv8s 分别高出 3.7%、2.1% 与 1.5%。相较于 Faster R-CNN、SSD 和 CenterNet 分别高出 10.2%、4.1% 与 3.2%。上述结果表明所提出

的 CFS-YOLO 模型对检测精度提升效果最为明显。

2)模型复杂程度: CFS-YOLO 模型的参数量高于 YOLOv5s 和 YOLOv7-tiny,但远低于 YOLOv8s、Faster R-CNN、SSD 和 CenterNet。相对适中的模型复杂程度可以在维持较高平均检测精度的同时不急剧降低其平均检测速度,更加适用于煤矿井下的煤矸精准实时检测任务。

3)平均检测速度: CFS-YOLO 模型的平均检测速度为 90.09 FPS,与 YOLOv5s、YOLOv7-tiny 和 SSD 相比分别低于 13.06、25.56、2.13 FPS,与 YOLOv8s、Faster R-CNN 和 CenterNet 相比分别高出 12.57、62.68、4.01 FPS。但参考实际检测速度标准(60 FPS),但仍充分满足煤矸实时检测的需求,此外,虽然其他 3 种模型的平均检测速度较快,但相应的平均精度均值低于 CFS-YOLO 模型。

3.8 检测结果对比

为了验证 CFS-YOLO 模型对煤矸在煤矿复杂环境中的检测效果,选取运动模糊、高噪声、低照度和大批量煤矸 4 种煤矿实际场景对传统 YOLOv5s 和 CFS-YOLO 进行对比检测,测试结果如图 10 所示。

表 6 不同算法对比试验
Table 6 Comparative Experiments of Different Algorithms

模型	精度均值/%		精确率/%	召回率/%	F_1	参数量	浮点运算次数	平均检测速度/FPS	平均精度均值/%
	煤	矸石							
YOLOv5s	87.2	85.7	86.1	79.7	0.82	7 025 023	16.0	103.15	86.5
YOLOv7-tiny	89.3	87.0	85.1	80.0	0.82	6 017 694	13.2	115.65	88.1
YOLOv8s	88.8	88.6	87.6	81.4	0.84	11 136 374	28.6	77.52	88.7
Faster R-CNN	81.2	78.9	70.7	78.5	0.74	28 286 000	940.0	27.41	80.0
SSD	86.7	85.5	81.7	78.6	0.80	23 745 000	60.0	92.22	86.1
CenterNet	88.0	86.0	88.2	70.2	0.84	32 665 000	70.0	86.08	87.0
CFS-YOLO	90.8	89.5	88.8	83.2	0.85	7 064 063	16.0	90.09	90.2

图中, A1-A8 为 CFS-YOLO 的检测结果, B1-B8 为传统 YOLOv5s 的检测结果, 其中图 A1-A2、B1-B2 为运动模糊场景, 图 A3-A4、B3-B4 为高噪声场景, 图 A5-A6、B5-B6 为低照度场景, 图 A7-A8、B7-B8 为大批量煤矸场景。CFS-YOLO 在不同场景下均能够

准确识别出煤和矸石, 未出现煤和矸石误检和漏检的情况, 而传统 YOLOv5s 出现了较多的漏检和误检问题。因此, 所提出的 CFS-YOLO 模型具有较高的鲁棒性和优良的检测性能。

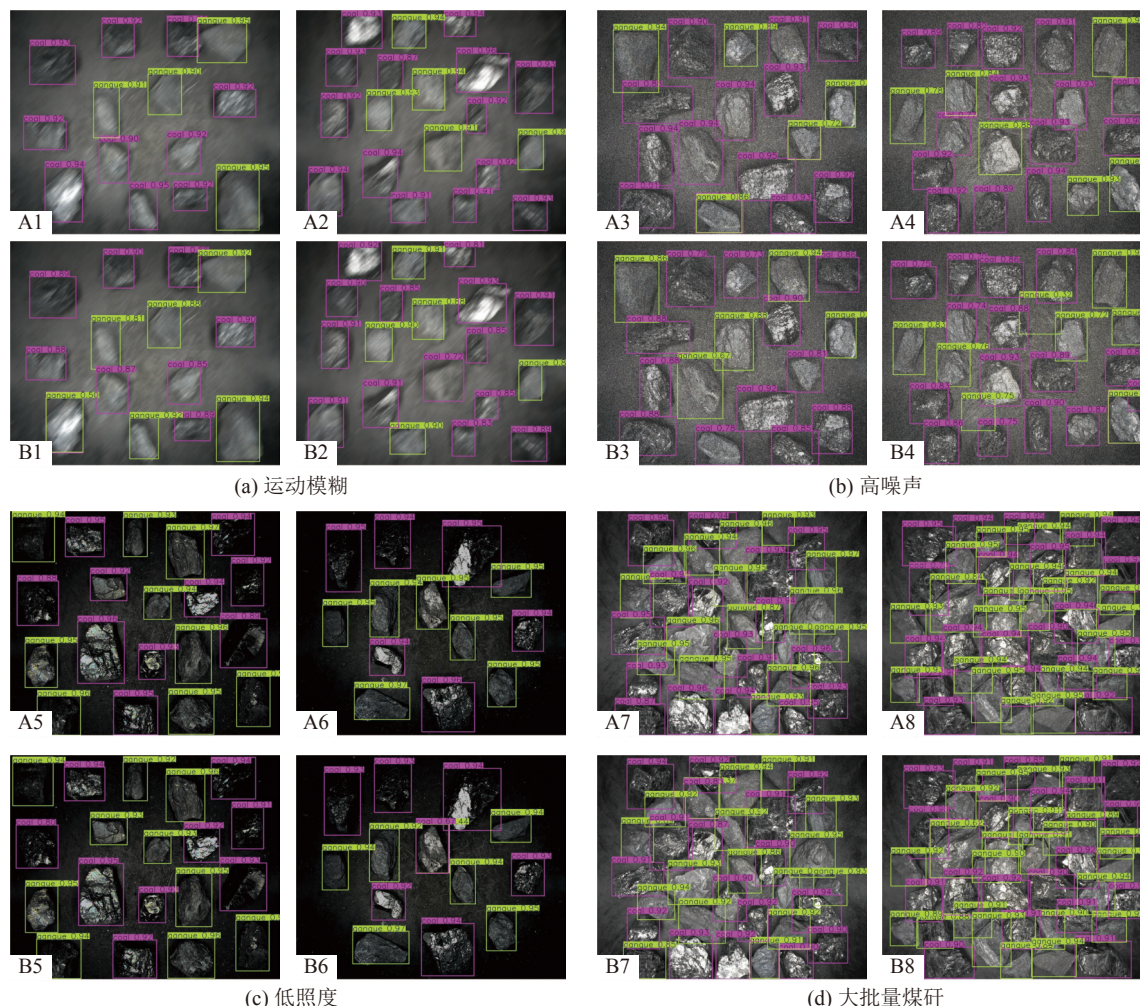


图 10 CFS-YOLO 与 YOLOv5s 检测结果对比

Fig.10 Comparison of CFS-YOLO and YOLOv5s test results

4 结 论

1)提出了一种 CFS-YOLO 目标检测模型, 试验结果表明, 其平均检测精度达 90.09%, 检测速度为 90.09 FPS, 该模型能够快速准确的对煤矸进行识别, 满足实时识别要求。

2)采用 ConvNeXt V2 特征提取模块替换主干网络末端的 2 个 C3 模块, 构建 Focal-EIOU 损失函数替换原 CIOU 损失函数, 同时添加无参注意力机制, 改进后的 CFS-YOLO 模型相较于原 YOLOv5s 模型提升了 3.7 个百分点, 网络参数体积和识别速度基本保持不变。与其他算法相比, 综合性能表现最佳。

3)提出的 CFS-YOLO 目标检测模型对于煤矸在

煤矿高噪声、低照度、运动模糊与大批量煤矸混杂等复杂工况环境下的检测效果均得到有效提高, 为复杂工况条件下煤矸智能识别提供了技术支撑。

参考文献(References):

- [1] 王国法, 孟令宇. 煤矿智能化及其技术装备发展[J]. 中国煤炭, 2023, 49(7): 1-13.
WANG Guofa, MENG Lingyu. Development of coal mine intelligence and its technical equipment[J]. China Coal, 2023, 49(7): 1-13.
- [2] 王国法, 任怀伟, 马宏伟, 等. 煤矿智能化基础理论体系研究[J]. 智能矿山, 2023, 4(2): 2-8.
WANG Guofa, REN Huaiwei, MA Hongwei, *et al.* Research on basic theoretical system of coal mine intelligentization[J]. Journal

- of Intelligent Mine, 2023, 4(2): 2–8.
- [3] 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 thinking of mine intelligent construction[J]. Journal of Intelligent Mine, 2022, 3(10): 2–15.
- [4] 石焕, 程宏志, 刘万超. 我国选煤技术现状及发展趋势[J]. 煤炭科学技术, 2016, 44(6): 169–174.
- SHI Huan, CHENG Hongzhi, LIU Wanchao. Present status and development trend of China's coal preparation technology[J]. Coal Science and Technology, 2016, 44(6): 169–174.
- [5] 张袁浩, 潘祥生, 陈晓晶, 等. 智能选矸机器人关键技术研究[J]. 工矿自动化, 2022, 48(6): 69–76, 111.
- ZHANG Yuanhao, PAN Xiangsheng, CHEN Xiaojing, *et al.* Research on key technologies of intelligent gangue sorting robot[J]. Journal of Mine Automation, 2022, 48(6): 69–76, 111.
- [6] 葛世荣, 张晞, 薛光辉, 等. 我国煤矿煤机智能技术与装备发展研究[J]. 中国工程科学, 2023, 25(5): 146–156.
- GE Shirong, ZHANG Xi, XUE Guanghui, *et al.* Development of intelligent technologies and machinery for coal mining in China's underground coal mines[J]. Strategic Study of CAE, 2023, 25(5): 146–156.
- [7] 李德永. 选矸机器人煤矸图像识别算法与分拣方法研究[D]. 淮南: 安徽理工大学, 2024.
- LI Deyong. Research on image recognition algorithm and sorting method of coal gangue in gangue sorting robot[D]. Huainan: Anhui University of Science & Technology, 2024. method of coal gangue in gangue sorting robot [D]. Anhui University of Science and Technology, 2023.
- [8] 张锦旺, 王家臣, 何庚, 等. 液体介入提升煤矸识别效率的试验研究[J]. 煤炭学报, 2021, 46(S2): 681–691.
- ZHANG Jinwang, WANG Jiachen, HE Geng, *et al.* Experimental study on improving the identification efficiency of coal gangue by liquid intervention[J]. Journal of China Coal Society, 2021, 46(S2): 681–691.
- [9] 李廉洁. 基于可见/近红外光谱及成像技术的煤和矸石分类方法研究[D]. 太原: 太原理工大学, 2023.
- LI Lianjie. Research on coal and gangue classification method based on vis/near infrared spectroscopy and hyperspectral imaging technology[D]. Taiyuan: Taiyuan University of Technology, 2023.
- [10] 胡锋. 基于多光谱成像和深度学习的煤矸识别研究[D]. 淮南: 安徽理工大学, 2021.
- HU Feng. Research on recognition of coal and gangue based on multispectral imaging and deep learning[D]. Huainan: Anhui University of Science & Technology, 2021.
- [11] 李瑞, 李博, 王学文, 等. 基于XGBoost与可见-近红外光谱的煤矸识别方法[J]. 光谱学与光谱分析, 2022, 42(9): 2947–2955.
- LI Rui, LI Bo, WANG Xuewen, *et al.* A classification method of coal and gangue based on XGBoost and visible-near infrared spectroscopy[J]. Spectroscopy and Spectral Analysis, 2022, 42(9): 2947–2955.
- [12] 郭永存, 何磊, 刘普壮, 等. 煤矸双能X射线图像多维度分析识别方法[J]. 煤炭学报, 2021, 46(1): 300–309.
- GUO Yongcun, HE Lei, LIU Puzhuang, *et al.* Multi-dimensional analysis and recognition method of coal and gangue dual-energy X-ray images[J]. Journal of China Coal Society, 2021, 46(1): 300–309.
- [13] 司垒, 谭超, 朱嘉皓, 等. 基于X射线图像和激光点云的煤矸识别方法[J]. 仪器仪表学报, 2022, 43(9): 193–205.
- SI Lei, TAN Chao, ZHU Jiahao, *et al.* A coal-gangue recognition method based on X-ray image and laser point cloud[J]. Chinese Journal of Scientific Instrument, 2022, 43(9): 193–205.
- [14] 王家臣, 潘卫东, 张国英, 等. 图像识别智能放煤技术原理与应用[J]. 煤炭学报, 2022, 47(1): 87–101.
- WANG Jiachen, PAN Weidong, ZHANG Guoying, *et al.* Principles and applications of image-based recognition of withdrawn coal and intelligent control of draw opening in longwall top coal caving face[J]. Journal of China Coal Society, 2022, 47(1): 87–101.
- [15] IWASZENKO S, RÓG L. Application of deep learning in petrographic coal images segmentation[J]. Minerals, 2021, 11(11): 1265.
- [16] 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.
- 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.
- [17] 杜京义, 史志芒, 郝乐, 等. 轻量化煤矸目标检测方法研究[J]. 工矿自动化, 2021, 47(11): 119–125.
- DU Jingyi, SHI Zhimang, HAO Le, *et al.* Research on lightweight coal and gangue target detection method[J]. Industry and Mine Automation, 2021, 47(11): 119–125.
- [18] LI D Y, REN H W, WANG G F, *et al.* Coal gangue detection and recognition method based on multiscale fusion lightweight network SMS-YOLOv3[J]. Energy Science & Engineering, 2023, 11(5): 1783–1797.
- [19] 徐志强, 吕子奇, 王卫东, 等. 煤矸智能分选的机器视觉识别方法与优化[J]. 煤炭学报, 2020, 45(6): 2207–2216.
- XU Zhiqiang, LYU Ziqi, WANG Weidong, *et al.* Machine vision recognition method and optimization for intelligent separation of coal and gangue[J]. Journal of China Coal Society, 2020, 45(6): 2207–2216.
- [20] 郭永存, 王希, 何磊, 等. 基于TW-RN优化CNN的煤矸识别方法研究[J]. 煤炭科学技术, 2022, 50(1): 228–236.
- GUO Yongcun, WANG Xi, HE Lei, *et al.* Research on coal and gangue recognition method based on TW-RN optimized CNN[J]. Coal Science and Technology, 2022, 50(1): 228–236.
- [21] XUE G H, LI S X, HOU P, *et al.* Research on lightweight Yolo coal gangue detection algorithm based on resnet18 backbone feature network[J]. Internet of Things, 2023, 22: 100762.
- [22] 韦小龙, 王方田, 刘超, 等. 基于CSPNet-YOLOv7目标检测算法的煤矸图像识别模型[J/OL]. 煤炭科学技术, 1–13[2023-11-13].
- WEI Xiaolong, WANG Fangtian, LIU Chao, *et al.* Coal gangue image recognition model based on CSPNet-YOLOv7 target detec-

- tion algorithm [J/OL]. *Coal Science and Technology*, 1–13[2023-11-13].
- [23] WANG L Y, WANG X W, LI B. Data-driven model SSD-BSP for multi-target coal-gangue detection[J]. *Measurement*, 2023, 219: 113244.
- [24] 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 Xiangang, 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.
- [25] 程德强, 钱建生, 郭星歌, 等. 煤矿安全生产视频 AI 识别关键技术研究综述[J]. *煤炭科学技术*, 2023, 51(2): 349–365.
CHENG Deqiang, QIAN Jiansheng, GUO Xingge, *et al.* Review on key technologies of AI recognition for videos in coal mine[J]. *Coal Science and Technology*, 2023, 51(2): 349–365.
- [26] WOO S, DEBNATH S, HU R H, *et al.* ConvNeXt V2: co-designing and scaling ConvNets with masked autoencoders[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada. IEEE, 2023: 16133–16142.
- [27] TOUVRON H, CORD M, SABLAYROLLES A, *et al.* Going deeper with image transformers[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 32–42.
- [28] ZHENG Z H, WANG P, LIU W, *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.
- [29] ZHANG Y F, REN W Q, ZHANG Z, *et al.* Focal and efficient IOU loss for accurate bounding box regression[EB/OL]. 2021: 2101.08158. <https://arxiv.org/abs/2101.08158v2>
- [30] YANG Lingxiao, ZHANG Ruyuan, LI Lida, *et al.* SimAM: a simple, parameter-free attention module for convolutional neural networks (supplementary materials)[C]//Proceedings of the 38th International Conference on Machine Learning, PMLR 139: 11863–11874, 2021.
- [31] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 779–788.