

基于MEBP-YOLOv10的无人矿卡行进道路障碍物检测

顾清华^{1,2}, 陈永琪^{1,2}, 李学现^{2,3}, 王丹^{2,3}

1. 西安建筑科技大学资源工程学院, 陕西 西安 710055;
2. 西安建筑科技大学西安市智慧工业感知计算与决策重点实验室, 陕西 西安 710055;
3. 西安建筑科技大学管理学院, 陕西 西安 710055



摘要:随着露天矿无人驾驶技术的广泛应用,无人驾驶卡车在复杂矿区环境下的障碍物检测问题日益凸显。针对障碍物检测存在的多尺度目标精度差和小目标特征融合不足等问题,提出了一种基于MEBP-YOLOv10的矿卡前进道路上的障碍物检测模型。该模型首先设计了C2f-MSF模块,采用分组策略和多尺度卷积提高特征提取效率;其次在主干网络中嵌入ECA注意力机制,强化了模型对障碍物的特征提取能力;然后使用BiFPN特征融合模块,调整权重占比以提高小目标障碍物的特征融合能力;最后将预测框损失函数更换为PIoU损失函数,通过惩罚项加快模型收敛。试验结果表明,改进后的模型对露天矿区中障碍物的检测精度、召回率、mAP和实时检测速度分别达到89.8%、79.8%、83.5%和103.4FPS,显著提高了矿卡前进道路上障碍物的检测精度,同时兼顾了实时处理能力和模型轻量化需求,方便部署在矿卡上的边缘设备。

关键词:露天矿;无人矿卡;机器视觉;障碍物检测;特征融合

中图分类号:TD57 文献标志码:A 文章编号:1005-2518(2025)04-0888-12 DOI:10.11872/j.issn.1005-2518.2025.04.360
引用格式:顾清华,陈永琪,李学现,等.基于MEBP-YOLOv10的无人矿卡行进道路障碍物检测[J].黄金科学技术,2025,33(4):888-899. GU Qinghua, CHEN Yongqi, LI Xuexian, et al. Obstacle Detection on the Forward Path for Autonomous Mining Trucks Based on MEBP-YOLOv10[J]. Gold Science and Technology, 2025, 33(4): 888-899.

矿用卡车作为露天矿场的关键运输设备,在矿山生产中占据着重要的地位,随着智慧矿山理念(王国法等,2018)的深入推进,矿石运输从数字化、智能化向无人化发展,使用无人驾驶的矿山卡车运输矿石逐渐成为趋势。然而,在复杂多变的矿山环境中,无人矿卡面临着严峻的安全挑战,尤其是如何实现对前方障碍物的精准、实时检测,已成为确保无人驾驶安全的核心技术难题。因此针对矿用卡车前方障碍物实时检测技术的研究具有重要的现实意义。

针对露天矿无人矿卡行进道路上的障碍物检测问题,国内外学者采用不同设备开展了大量研究,这些设备可划分为激光雷达检测和视觉检测两大类。其中,激光雷达发射激光束至障碍物,得到障碍物表面具有深度信息的点云数据,通过点云预处理、点云聚类、边界框生成及匹配跟踪等步骤,实现障碍物的精确检测(张静,2022;高明宇等,2025)。然而,激光雷达虽然具有优秀的检测能力,但其存在成本较高以及对灰尘、雨雾等环境因素较为敏感等问题,导致其检测效果与使用场景受到制

收稿日期:2024-11-18;修订日期:2025-03-20

基金项目:国家自然科学基金项目“数据驱动下露天矿穿爆铲运破多工序协同采矿优化研究”(编号:52374135)和“金属露天矿无人驾驶多工序多目标协同智能调度方法研究”(编号:52074205)联合资助

作者简介:顾清华(1981—),男,山东诸城人,教授,博士生导师,从事矿业系统工程研究工作。E-mail:qinghuagu@126.com

© Editorial Department of Gold Science and Technology (CC BY-NC-ND)

约(李少博等,2025)。随着计算机视觉与深度学习技术的快速发展,且得益于相机设备成本低、适应能力强的优势,结合视觉与深度学习的障碍物检测方法成为研究热点。与激光雷达不同的是,通过视觉与深度学习进行障碍物检测时需首先通过摄像设备获取障碍物的图像,然后通过训练神经网络,回归预测到图像中障碍物的类别。当前主流方法有YOLO(Zhong et al,2024)、R-CNN(He et al,2017)和DETR(Carion et al,2020)等。视觉检测方法虽展现出显著优势,但也面临矿山复杂环境的严峻挑战,对此,国内外学者从不同层面对目标检测框架进行了优化(卢才武等,2020;白俊卿等,2022)。在数据层面,针对图像质量低和拍摄环境较暗等问题,雷杨等(2025)将自校正照明网格融入YOLOv8网络结构中,在训练过程中,逐步提升图像质量,成功检测出低质量图像中的障碍物。在算法层面,目标检测网络主要由特征提取、特征融合和回归预测3个步骤完成,因此国内外学者从网络结构、模块优化和损失函数等方面进行了改进(阮顺领等,2022;陈良,2025)。露天矿山中的障碍物检测虽然取得了较好的成果,但仍存在一些问题,如:当前露天矿障碍物检测模型中对于障碍物的划分较少,不能满足露天矿区道路上的所有障碍物检测;对于远距离小目标障碍物,当前模型难以识别;高精度障碍物检测模型的运行成本高,对于计算设备需求较为严格,难以部署至边缘设备。

基于以上分析,为应对露天矿复杂场景下障碍物检测存在的检测精度低、小目标检测存在偏差和模型部署不便等问题,提出了基于MEBP-YOLOv10的无人矿卡前进道路上的障碍物检测模型。通过对YOLOv10模型进行多种策略的改进,实现矿卡行进道路上的障碍物检测。

1 露天矿卡前进道路上的障碍物检测模型

1.1 MEBP-YOLOv10模型结构

针对露天矿区复杂的环境,模型从4个方面对YOLOv10进行改进,提出一种基于MEBP-YOLOv10(YOLOv10 Enhanced with Multi-scale Convolution, ECA, BiFPN, and PIoU)的障碍物检测模型。首先

将主干网络中的C2f模块替换为C2f-MS(C2f-Multi-Scale Conv)模块,然后在主干网络中添加高效通道注意力机制(Efficient Channel Attention, ECA)(Wang et al,2020),并将特征融合模块中的PANet(Liu et al,2018)替换为BiFPN(Tan et al,2020);最后,将预测框的回归损失函数由CIoU更换为收敛速度更快且精度更高的PIoU(Liu et al,2024)。MEBP-YOLOv10模型框架如图1所示。

1.2 多尺度特征融合模块C2f-MS

由于矿山环境复杂多变,障碍物种类繁多、大小不一,存在光照、灰尘等环境因素,导致模型的特征提取能力受到影响,且目标检测任务一般在矿卡卡车的边缘设备上,而边缘设备往往存在计算能力弱的问题。因此,为了提高模型的特征能力,使用C2f-MS模块替换部分C2f模块。C2f-MS参考GhostNet的思想,GhostConv采取将通道数分组的方式,将输出特征图划分为通道上相同的2个部分,一部分由简单卷积得到,另一部分由深度可分离卷积完成,然后拼接在一起,得到输出特征图。与普通卷积相比,GhostConv能够以更少的参数量得到效果更优的特征图。C2f-MS借鉴GhostNet的思想,提出MSConv,其结构如图2所示。

MSConv结构采取多尺度卷积提取特征,将输入特征图按照通道数平均划分为4组,每一组通过不同的卷积核进行并行处理,卷积完成后进行拼接,最后通过一个逐点卷积将特征进行整合。MSConv在提高特征提取能力的同时,检测速度不会受到太大的影响,适合部署在边缘设备上。

1.3 加权双向特征金字塔BiFPN

由于露天矿区中存在积水、行人和标识牌等目标,当拍摄位置与障碍物距离过远时,会使得拍摄到的障碍物过小,而小目标在检测时只会保留较小的特征,导致小目标障碍物特征在特征融合时易被忽略的情况。针对小目标检测的问题,使用BiFPN更换YOLOv10中原始的PANet,通过双向跨尺度连接与加权特征融合,提高模型对小目标障碍物特征融合的能力,BiFPN与FPN、PANet的结构对比如图3所示。

由图3可以看出,原始的FPN只有一条自上而下的路径,其在下采样过程中丢弃了一些特征信息,而YOLOv10中采用的PANet增加了一条自下而

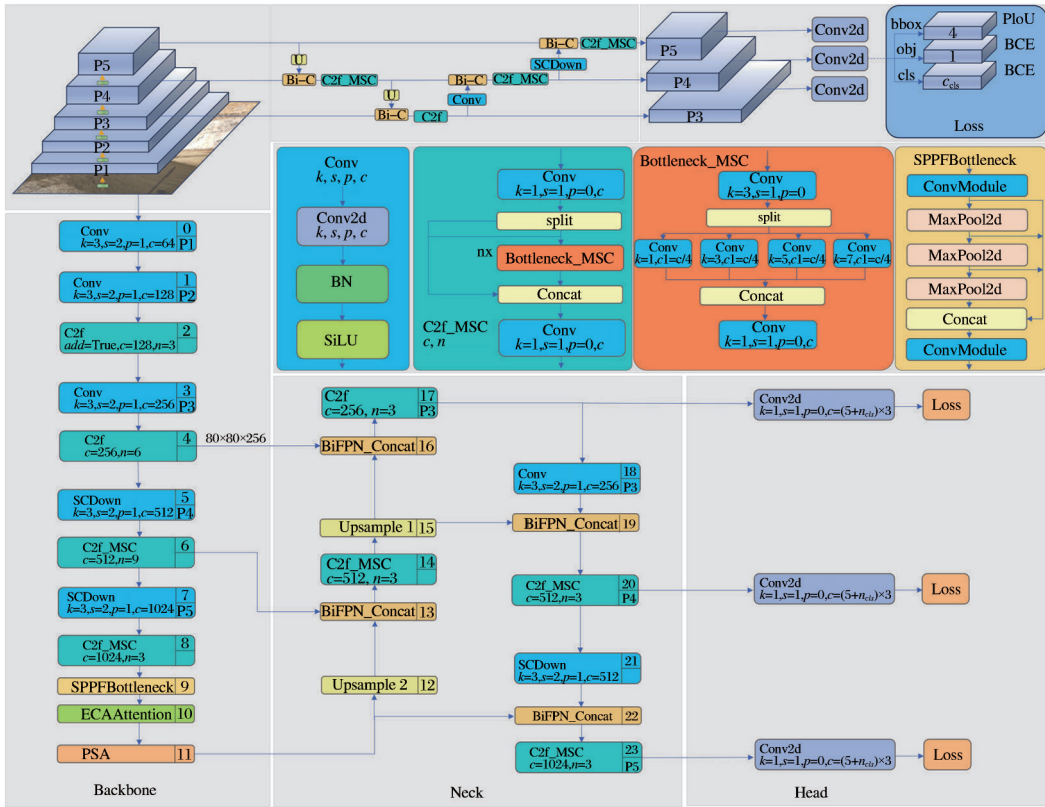


图1 MEBP-YOLOv10模型图

Fig. 1 Model diagram of MEBP-YOLOv10

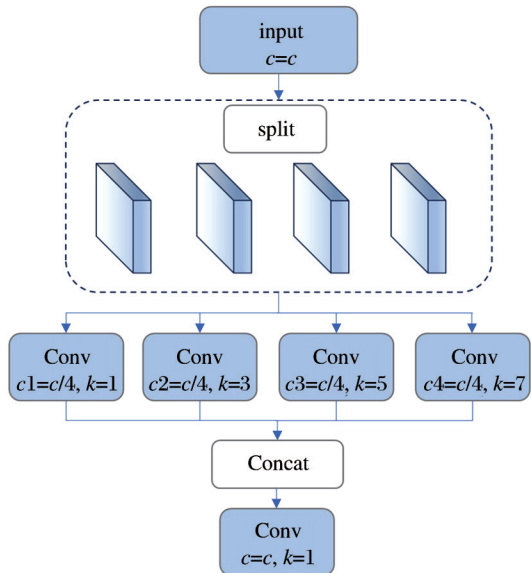


图2 MSConv结构图

Fig. 2 Structure diagram of MSConv

上的路径,将不同的特征层进行融合,从而获取更多特征信息。然而,由于PANet向下特征融合阶段

如 P_4^{out} 接收来自经过上采样特征融合之后 P_4^d 的特征,此时 P_4^d 得到的特征相对于原始输入的特征会有所损失。为了避免特征的损失, BiFPN采用残差结构,通过在同一层的原始输入与输出节点之间增加额外的连接边,确保特征信息的完整性。通过带权融合策略, BiFPN在每次融合时进行快速归一化,将可学习的权重缩放至 $[0, 1]$ 之间,并区分不同输入特征的重要性,实现有效融合。快速归一化融合的计算公式为

$$O = \sum_i \frac{w_i \times I_i}{\sum_j w_j + \epsilon} \quad (1)$$

式中: O 为融合后的输出特征; w_i 为第 i 个输入特征图的可学习权重; I_i 为第 i 个输入特征; ϵ 设置为 0.001, 以避免分母为 0。

以图3中 P_4 层输出为例, BiFPN的计算公式为

$$P_4^{td} = Conv \left[\frac{w_1 \cdot P_4^{in} + w_2 \cdot Resize(P_5^{in})}{w_1 + w_2 + \epsilon} \right] \quad (2)$$

$$P_4^{out} = Conv \left[\frac{w_1' \cdot P_4^{in} + w_2' \cdot P_4^{td} + w_3' \cdot Resize(P_3^{out})}{w_1' + w_2' + w_3' + \epsilon} \right] \quad (3)$$

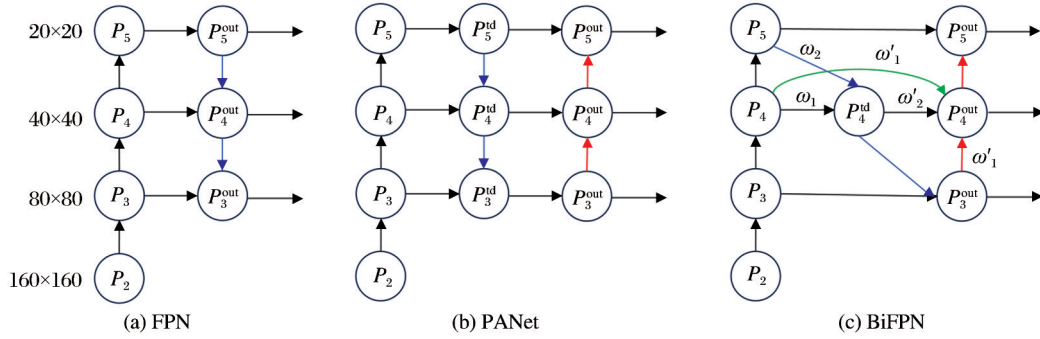


图3 特征融合模块结构对比

Fig. 3 Comparison of feature fusion module structure

式中: $Resize$ 操作通常是上采样或下采样操作; P_i^{in} 为第 i 个输入特征; P_i^{out} 为第 i 个输出特征; ϵ 设置为 0.0001, 以避免分母为 0, 造成数组不稳定; w 为学习到的参数, 用于区分特征融合过程中不同特征的重要程度。

1.4 轻量级高效通道注意力 ECA

为增强模型在通道维度的特征表示能力, 同时控制计算成本, 在主干网络中加入 ECA 注意力机制, 通过对每个通道进行不同程度的加权, 捕获通道间的关系, 提高障碍物特征的提取能力。ECA 注意力机制如图 4 所示。

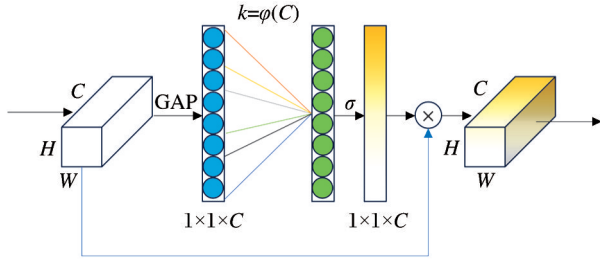


图4 ECA 注意力机制模型图

Fig. 4 Model diagram of ECA attention mechanism

ECA 注意力机制首先对输入的特征层进行全局平均池化(Global Average Pooling, GAP)操作, 从而获得每个通道的全局空间信息。全局平均池化计算公式为

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_{ijc} \quad (4)$$

式中: z_c 为第 c 个通道的输出值; H 为输入特征图的高度; W 为输入特征图的宽度; x_{ijc} 为特征图中第 c 个通道、第 i 行、第 j 列的像素值。

在提取全局特征后, 采用自适应核大小的一维

卷积实现跨通道信息交互, 其中卷积核尺寸通过学习函数动态调整, 确保高维特征通道能够获得更丰富的跨通道特征融合。首先设定一个超参数 γ 和 b , 其中 $\gamma=2, b=1$ 。然后对于给定的通道数 C , 卷积核大小 k 的计算公式为

$$k = \varphi(C) = \left\lfloor \frac{\log_2 C + b}{\gamma} \right\rfloor \quad (5)$$

当通道数 C 增加时, 卷积核大小 k 也会增加, 实现自适应的跨通道交互。在完成卷积操作之后, 将一维卷积后的结果通过 Sigmoid 激活函数, 得到每个通道的权重系数 s , 再将权重系数 s 与原始输入特征图进行相乘, 得到最终的输出特征图。

1.5 优化损失函数 PIoU

YOLOv10 的损失函数由 3 个部分组成, 分别为类别分类损失 cls_loss 、偏移量损失 dfl_loss 和预测框回归损失 box_loss 。在对露天矿卡前进道路上的障碍物检测模型进行实验时, 发现 box_loss 指标仍处于下降状态, 未能够按时收敛, 所以采取 PIoU (Powerful-IoU) 损失函数替换 YOLOv10 中的预测框回归损失函数。在 YOLOv10 中, 预测框回归损失函数采用的是 CIoU, 该函数定义如下:

$$CIoU = IoU - \left[\frac{\rho^2(b, b_{gt})}{c^2} + \alpha \times v \right] \quad (6)$$

$$c = \sqrt{c_w^2 + c_h^2} + \epsilon \quad (7)$$

$$\alpha = \frac{v}{1 - IoU + v + \epsilon} \quad (8)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2 \quad (9)$$

式中: IoU 为真实框与预测框的交并比; $\rho^2(b, b_{gt})$ 为真实框中心点与预测框中心点之间的距离; c 为包

围2个框的最小外接矩形框的对角线长度; c_w 和 c_h 分别为最小外接矩形框的宽和长; α 为权衡宽高比一致性的参数; v 为长宽比不一致性惩罚项, w_{gt} 和 h_{gt} 分别为真实框的宽度和高度; ϵ 是一个非常小的常数,用于避免除零错误。

由式(9)可知,当真实框与预测框的宽高一致时, α 和 v 均为0,从而导致惩罚项 v 失去了作用,损失函数不能达到预期的结果,使得收敛速度变慢,检测精度变差。因此,将CIoU损失函数更换为PIoU损失函数,其示意图如图5所示。

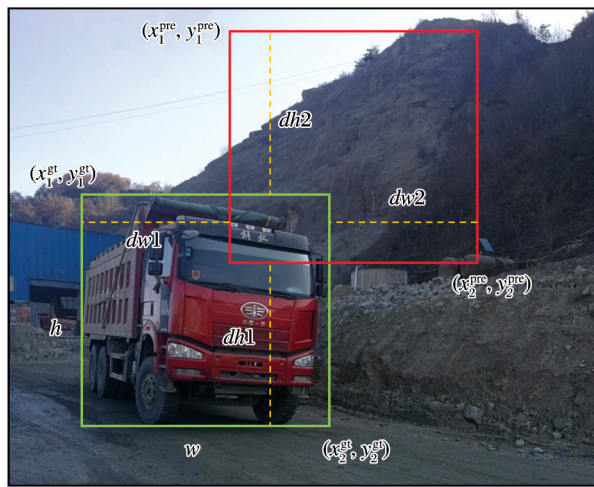


图5 PIoU示意图

Fig. 5 Schematic diagram of PIoU

如图5所示,绿色框为真实框,红色框为预测框, (x_1^{gt}, y_1^{gt}) 和 (x_2^{gt}, y_2^{gt}) 分别为真实框左上角与右下角坐标, (x_1^{pre}, y_1^{pre}) 和 (x_2^{pre}, y_2^{pre}) 分别为预测框左上角与右下角坐标, w 和 h 分别为真实框的宽和高, $dw1$ 和 $dh2$ 分别为2个框在水平方向上左侧边与右侧边

之差, $dh1$ 和 $dh2$ 分别为2个框在竖直方向顶部边与底部边之差。PIoU损失函数的定义如下:

$$PIoU = 3 \times x \times e^{-x^2} \times (IoU - e^{-P^2} + 1) \quad (10)$$

$$x = e^{-P} \times \alpha \quad (11)$$

$$P = \left(\frac{dw1 + dw2}{w} + \frac{dh1 + dh2}{h} \right) \div 4 \quad (12)$$

$$w = x_2^{gt} - x_1^{gt}, h = y_2^{gt} - y_1^{gt} \quad (13)$$

式中: IoU 为真实框与预测框的交并比; P 为惩罚项; α 为人为设定的参数,通常设定为1.3。

2 试验与结果分析

2.1 障碍物数据集构建与试验准备

试验使用的数据集来自国内3个大型露天矿山的非结构化道路,共有782张图片。为了更好地检测和分类露天矿山中的障碍物,将障碍物划分为7类:卡车、行人、汽车、挖掘机、积水、标识牌和其他车辆。由于初始训练集数据有限,神经网络难以理想地拟合。因此,采用随机数据增强策略,引入噪声和变化,增强模型的健壮性和鲁棒性。共采用7种数据增强方式,包括裁剪、平移、亮度调整、高斯噪声、旋转、镜像和随机矩形擦除。将这些操作随机融合,每张图片随机增强5次,生成与原始数据相似但不完全相同的新数据。通过随机数据增强,增强了露天矿障碍物数据集的样本丰富度,显著提高了检测模型的精度和泛化能力。数据增强后的图片共计4692张,图像分辨率为1420 pixel×1080 pixel,输入网络的图片尺寸为640 pixel×640 pixel,将数据集按7:3的比例划分为训练集与验证集,划分完成后的数据集样本数据量如表1所示。

表1 数据增强前后的数据集样本数量

Table 1 Number of dataset samples before and after data augmentation

类别	训练集数据增强前	训练集数据增强后	验证集数据增强前	验证集数据增强后
卡车	856	5 136	377	2 262
行人	153	918	57	342
汽车	114	684	44	264
挖机	230	1 380	95	570
积水	165	990	89	534
标识牌	204	1 224	87	522
其他车辆	13	78	6	36
图片数	547	3 282	235	1 410

本试验对模型的训练与验证均采用 NVIDIA RTX3090 显卡, 24 GB 显存, 对图片与视频的推理采用 NVIDIA GTX1060 显卡, 6 G 显存, 并基于 python3.8 与 pytorch1.9 框架。试验参数如表 2 所示。

表 2 试验参数设置
Table 2 Setting of experimental parameter

参数名称	数值	参数名称	数值
图片输入	640×640	权重衰减系数	0.0005
批次大小	24	学习率调整策略	余弦退火策略
优化器	SGD	预热批次	3
初始学习率	0.01	动量	0.937

试验的精度评价指标为准确率 P (Precision)、召回率 R (Recall)、平均精度 AP (Average Precision) 和平均精度均值 mAP (Mean Average Precision)。

2.2 障碍物检测模型有效性分析

露天矿障碍物检测结果如图 6 所示。经过试

验可得, 障碍物检测模型在目标障碍物距离远、目标小且太阳光照强的条件下依然能够保持精确的检测效果, 卡车和挖掘机等较大目标的置信度能够保持在 0.9 以上, 行人和标识牌等小目标的置信度也保持在 0.8 之上, 满足障碍物模型在矿山场景中的检测需求。

本文将提出的障碍物检测模型与目前主流的目标检测模型进行对比试验, 试验结果如表 3 所示。为了优化模型的训练收敛速度, 所有模型均采用迁移学习的方式进行训练, YOLO 系列使用经过 COCO 数据集训练的预训练权重进行迁移学习, 输入图片大小为 640×640, 训练 300 轮次, YOLOv8 (Quan et al, 2023) 采用 n 模型, YOLOv9 (Bakirci et al, 2024) 采用原模型, EfficientDet 基于 ImageNet 数据集进行迁移学习, 输入图片大小为 512×512, 训练 100 轮次。以上模型均使用训练得到效果最好的权重文件进行障碍物检测, 检测部分结果如图 7 所示。

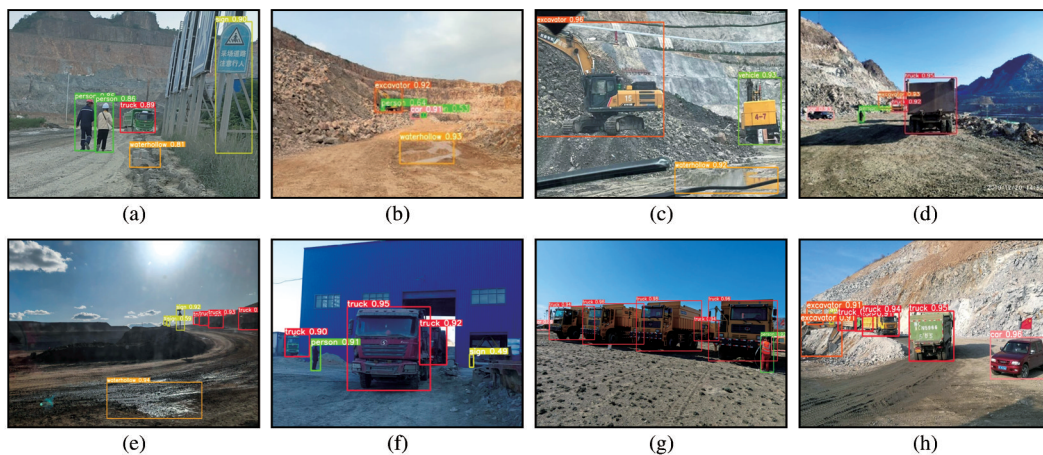


图 6 露天矿障碍物检测结果
Fig. 6 Detection results of open-pit mine obstacle

表 3 主流神经网络模型性能对比

Table 3 Performance comparison of mainstream neural network models

模型	精度 P /%	召回率 R /%	mAP/%	模型大小/MB	速度/FPS	浮点运算数/GFLOPs
EfficientDetd4	69.7	56.1	63.8	79.6	20.1	92.6
YOLOv8n	79.9	67.3	74.6	5.33	120.3	7.5
YOLOv9	82.1	70.8	76.4	83.6	50.6	164.8
YOLOv11n	85.4	75.5	79.6	5.79	119.3	8.4
MEBP-YOLOv10	89.8	79.8	83.5	6.87	103.4	10.3

本文提出的障碍物检测模型在表现上优于其他主流目标检测模型。在 300 轮次时, 各项损失及

评价指标均达到动态平衡, 模型达到了收敛, 精度 P 为 89.8%, 召回率 R 为 79.8%, 平均精度均值 mAP



图7 不同模型障碍物检测结果

Fig. 7 Obstacle detection results of different models

为 83.5%，检测速度为 103.4FPS，满足实时需求。EfficientDetd4 训练速度较慢，难以收敛，最终检测效果不佳。图7中显示EfficientDet在障碍物检测中普遍出现大规模漏检和错检，且对小目标的检测效果较差，不适用于露天矿障碍物检测。在光照差的情况下，YOLOv8出现了误检，YOLOv9在检测过程中均存在误检、漏检，YOLOv11在检测过程中均存在漏检情况。

2.3 障碍物检测模型优化分析

(1) C2f-MSConv效果分析

本文将 YOLOv10 框架中的 C2f 更换为 C2f-MSConv，同时对 MSConv 中 4 个卷积核大小 k 的取值进行试验，试验数据与原始模型比较结果如表 4 所示。从表 4 中可得当 k 取 [1, 3, 5, 7] 时，模型的性能提升最为显著，MSConv 通过分组操作与多尺度卷积提高了模型的特征提取能力，使得模型的推理速度不受影响。通过更换 C2f 模块的试验，得到 C2f 模块与 C2f-MSConv 模块的 mAP 对比图如图 8 所示。

由图 8 可得，将 C2f 模块更换为 C2f-MSConv 模块后，模型的 mAP 提升了 1.2%，同时，通过对 200 张矿山道路障碍物图片进行推理，得到推理时间为 30.45 s，相较于 YOLOv10n 模型 (29.83 s)，慢了 0.62 s。在批次大小为 24 的前提下，推理所占 GPU 内存为 1 929 MB，相较于 YOLOv10n 模型的 1 911 MB，多占用了 18 MB。证明了 C2f-MSConv 模块在只增加较低资源占用的情况下，特征提取能力得到了提升。

(2) ECA 注意力机制效果分析

在主干网络下采样的过程中，加入 ECA 注意力

表 4 卷积核大小 k 取值的试验结果对比

Table 4 Comparison of experimental results for different values of convolution kernel size k

k 的取值	mAP/%	模型大小/MB	速度/FPS
[1, 3, 5, 7]	+1.2	+0.18	-4.8
[2, 4, 6, 8]	+1.2	+0.21	-5.5
[3, 5, 7, 9]	+0.9	+0.32	-7.4
[1, 2, 3, 4]	+0.5	+0.06	-1.5

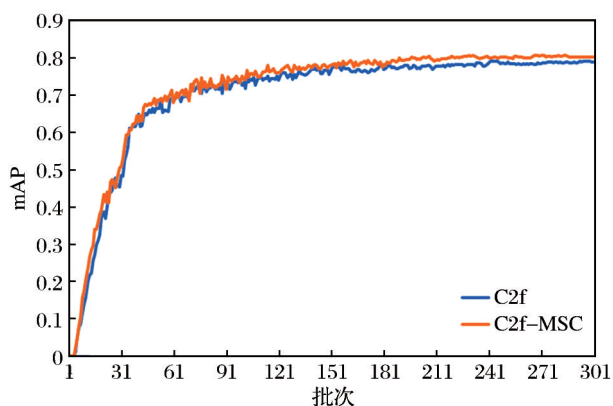


图8 mAP对比图

Fig. 8 Comparison diagram of mAP

机制,进一步增强特征的提取能力,并将ECA注意力机制与另外几种主流的注意力机制进行试验对比,试验中,mAP和浮点运算量的变化是相对于加入了C2f-MSA模块的YOLOv10模型而言的,试验结果如表5所示。

由表5可知,EMA(Ouyang et al, 2023)(Efficient

表5 注意力效果对比
Table 5 Comparison of attention effect

注意力机制	mAP变化/%	浮点运算数变化/GFLOPs
SimAM	+0.1	+0
SE	+0.6	+0.1
EMA	+0.3	+1.1
CBAM	+0.8	+0.2
ECA	+1.2	+0

Muliti-Scale Attention)引入指数移动平均,将先前时刻的注意力信息累积,使模型更注重历史上下文,但因此需要维护和更新历史信息,增加了计算开销。CBAM(Woo et al, 2018)(Convolutional Block Attention Module)对通道和空间维度分别加权,使得网络可以更加准确地捕捉重要的特征,从而提升模型的整体性能。相较于其他主流注意力机制,ECA注意力机制在不增加模型复杂度的情况下,检测的精度也得到明显的提高。使用GradCAM(Selvaraju et al, 2020)方法得出不同的注意力机制的热力图,如图9所示。

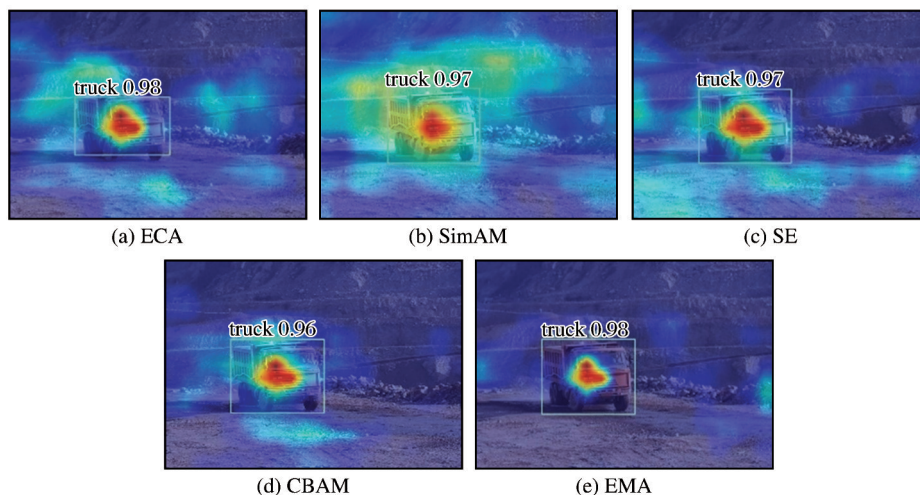


图9 GradCAM热力对比图

Fig. 9 Comparison diagram of GradCAM heat mAP

由图9可知,SimAM(Yang et al, 2021)(Simple Attention Module)在计算注意力时,引入相似性度量,通过度量特征之间的相似性来调整注意力分配,导致容易受到噪声与不相关信息的干扰,使其得到的热力图中特征过于分散,并不集中于待检测的目标上。相比之下,EMA因过于集中于待检测的目标上,而丢失了周围的其他特征信息,ECA、CBAM和SE(Tang et al, 2020)(Squeeze-and-Excitation)能

够有效对目标特征进行采样,得到较好的效果。

(3)BiFPN效果分析

为了验证本文提出的使用BiFPN替换YOLO中原有FPN的有效性,将YOLOv10采用的PANet替换为BiFPN,并进行对比试验,试验得到的PR曲线图如图10所示。从图10中可以看出,BiFPN所有类别的平均精度均值得到提升,相较于PANet,BiFPN通过为每个输入添加一个额外的权重,使网

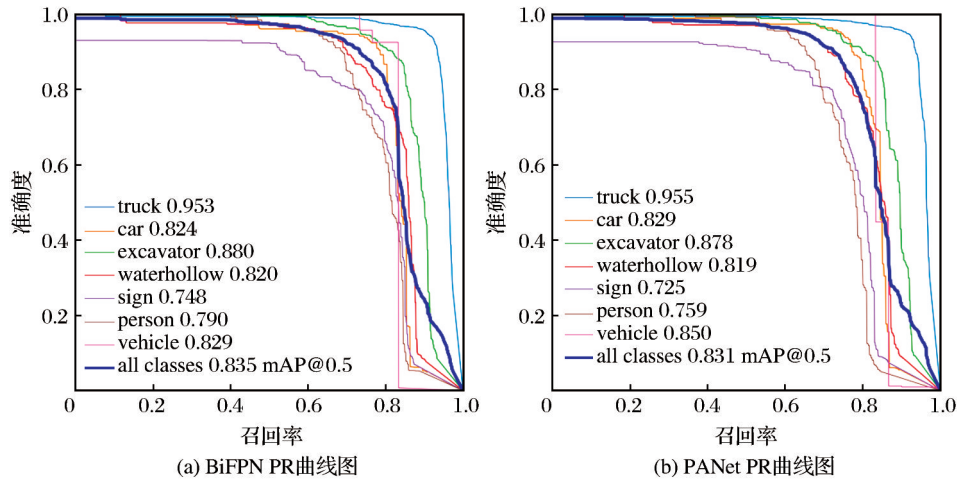


图 10 PR 曲线对比图

Fig. 10 Comparison diagram of PR curves

络得到每个特征的重要性,进而优化不同特征的训练比重,从而使 mAP 值提高了 0.4%,并且小目标障碍物的 mAP 值也有所提升,行人和标识牌的 mAP 值分别提高了 3.1% 和 2.3%,在保证小目标特征融合的同时,也保证了模型总体的检测性能。

(4)PIoU 效果分析

为了验证以 PIoU 损失函数替换 CIoU 的有效性,将 CIoU 与 PIoU 损失函数进行对比,得到损失函数对比如图 11 所示。从图 11 中可得,CIoU 在 300 批次时仍处于下降趋势,未达到收敛。PIoU 损失函数在 300 批次时趋于平缓。PIoU 的损失下降速度快于 CIoU,且 PIoU 最终的 box_loss 值也低于 CIoU。PIoU 在真实框与预测框的宽高一致时,惩罚项不会和 CIoU 一样失去作用,从而保证了损失函数的下降速度。

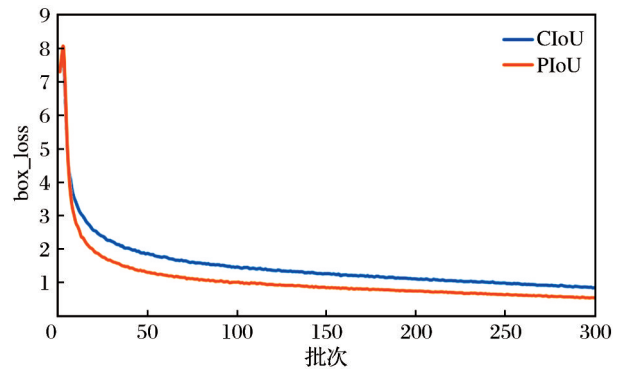


图 11 损失函数对比图

Fig. 11 Comparison diagram of loss functions

(5)消融试验

为了验证本文提出的 C2f-MSc 模块替换 C2f 模块、引入 ECA 注意力机制、BiFPN 替换 PANet、预测框损失函数 PIoU,设计了消融试验,如表 6 所示。

表 6 消融试验分析

Table 6 Analysis of ablation experiments

C2f-MSc	ECA	PIoU	BiFPN	精度 P/%	召回率 R/%	mAP/%	模型大小/MB	速度/FPS
				85.1	75.6	79.3	5.86	117.9
√				86.4	76.6	80.5	6.04	113.1
√	√			87.9	77.8	81.7	6.04	113.1
√	√	√		89.4	79.1	83.1	6.04	113.1
√	√	√	√	89.8	79.8	83.5	6.87	103.4

从表 6 可知,以 C2f-MSc 模块替换 C2f 模块,以较小的代价提高了模块的特征提取能力,mAP 值提高了 1.2%,模型大小只增加了 0.18 MB。在主干网

络中加入 ECA 注意力机制,提高网络对关键信息的感知能力,进一步增强感受野,使得后续步骤能够更好地理解输入特征图的上下文信息。模型对

YOLOv10的特征融合部分重新进行设计,使用BiFPN替代PANet,使模型在特征融合阶段融合多个尺度的特征,提高小目标障碍物识别精度,从而改善了整个模型的检测性能。通过将预测框损失函数由CIoU更换为PIoU,使得模型的box_loss收敛速度提高,收敛效果更好。

3 结论

(1)提出了MEBP-YOLOv10露天矿区道路障碍物检测模型,该模型的平均精度和mAP分别为89.8%和83.5%,分别比YOLOv10模型提高了4.7%和4.2%。模型大小为6.87 MB,检测速度为103.4 FPS。

(2)模型通过使用C2f-MSB模块替换部分C2f模块,提高了特征提取能力;在主干网络的最后添加ECA注意力机制;使用BiFPN结构替换原有的PANet结构;最后通过将CIoU损失函数更换为PIoU损失函数,加速了box_loss的收敛。

(3)试验结果表明,本文提出的MEBP-YOLOv10障碍物检测模型在多个矿区环境中均展现出优异的检测性能,能够有效部署于矿卡的边缘计算设备,实现对矿卡前进路径上障碍物的快速精准检测。然而,尽管该模型在常规光照条件下表现出色,但在夜间及光线不足的矿区环境中,其检测效果仍有待提升。在今后的工作中,应结合红外相机采集夜间矿区障碍物数据,构建相应的数据集,以进一步优化模型在低光照环境下的检测能力,提升其在实际应用中的鲁棒性和适应性。

参考文献(References):

Bakirci M, Dmytrovych P, Bayraktar I, et al, 2024. Challenges and advances in UAV-Based vehicle detection using YOLOv9 and YOLOv10 [C]//7th IEEE International Conference on Actual Problems of Unmanned Aerial Vehicles Development. Ukraine: Institute of Electrical and Electronics Engineers Inc.

Carion N, Massa F, Synnaeve G, et al, 2020. End-to-end object detection with transformers [M]//Computer Vision-ECCV 2020. Cham: Springer International Publishing.

He K M, Gkioxari G, Dollár P, et al, 2017. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE.

Liu C, Wang K G, Li Q, et al, 2024. Powerful-IoU: More straightforward and faster bounding box regression loss with a nonmonotonic focusing mechanism [J]. *Neural Networks*, 170: 276-284.

Liu S, Qi L, Qin H F, et al, 2018. Path aggregation network for instance segmentation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE.

Ouyang D L, He S, Zhang G Z, et al, 2023. Efficient multi-scale attention module with cross-spatial learning [C]//2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, Greece: IEEE.

Quan Y, Wang P, Wang Y, et al, 2023. GUI-Based YOLOv8 license plate detection system design [C]//5th International Conference on Control and Robotics, ICCR 2023. Tokyo: IEEE.

Selvaraju R R, Cogswell M, Das A, et al, 2020. Grad-CAM: visual explanations from deep networks via gradient-based localization [J]. *International Journal of Computer Vision*, 128(2): 336-359.

Tan M X, Pang R M, Le Q V, 2020. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE.

Tang Z, Liu X, Li Y, et al, 2020. Multi-atlas brain parcellation using squeeze-and-excitation fully convolutional networks [J]. *IEEE Transactions on Image Processing*, 29: 6864-6872.

Wang Q L, Wu B G, Zhu P F, et al, 2020. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE.

Woo S, Park J, Lee J Y, et al, 2018. CBAM: convolutional block attention module [C]//Computer Vision-ECCV 2018. Cham: Springer.

Yang L, Zhang R Y, Li L, et al, 2021. SimAM: a simple, parameter-free attention module for convolutional neural networks [C]//38th International Conference on Machine Learning, ICML 2021. Virtual, Online: ML Research Press.

Zhong H, Chen Z, Shang Q, et al, 2024. YOLOv10 improvement approach based on ros robot visual obstacle avoidance system [C]//9th International Conference on Cyber Security and Information Engineering, ICCSIE 2024. Hybrid, Kuala Lumpur, Malaysia: Association for Computing Machinery.

- 白俊卿, 张文静, 2022. 基于YOLOv4优化的轻量级无人机障碍物检测方法[J]. 电子测量技术, 45(22): 87-91. Bai Junqing, Zhang Wenjing, 2022. Lightweight UAV obstacle detection method optimized based on YOLOv4 [J]. Electronic Measurement Technology, 45(22): 87-91.
- 陈良, 2025. 基于RT-DETR的露天矿区路面障碍物检测[J]. 黄金科学技术, 33(1): 202-213. Chen Liang, 2025. Detection of road obstacles in open-pit mining areas based on RT-DETR [J]. Gold Science and Technology, 33(1): 202-213.
- 高明宇, 鲍久圣, 阴妍, 等, 2025. 基于优化蚁群算法的露天矿无人矿卡绕跨并行类三维路径规划[J]. 煤炭科学技术, 53(增刊1): 399-411. Gao Mingyu, Bao Jiusheng, Yin Yan, et al, 2025. Detour-Straddle 3D-like path planning of unmanned mining truck in open pit mines based on optimized ant colony algorithm [J]. Coal Science and Technology, 53(Suppl.1): 399-411.
- 雷杨, 何江, 秦丽杰, 等, 2025. 基于SCI-YOLOv8的井下无人电机车低照度目标检测技术[J]. 金属矿山, 54(2): 172-179. Lei Yang, He Jiang, Qin Lijie, et al, 2025. Low light target detection technology for underground unmanned electric vehicles based on SCI-YOLOv8 [J]. Metal Mine, 54(2): 172-179.
- 李少博, 顾清华, 阮顺领, 等, 2025. 融合小波变换的露天矿无人车视觉多任务感知方法[J/OL]. 煤炭学报: 1-13. (2025-03-18). <https://link.cnki.net/doi/10.13225/j.cnki.jccs.2024.1627>. Li Shaobo, Gu Qinghua, Ruan Shunling, et al, 2025. A visual multi-task perception method for unmanned vehicles in open pit mines incorporating wavelet transforms [J/OL]. Journal of China Coal Society: 1-13. (2025-03-18). <https://link.cnki.net/doi/10.13225/j.cnki.jccs.2024.1627>.
- 卢才武, 齐凡, 阮顺领, 2020. 融合目标检测与距离阈值模型的露天矿行车障碍预警[J]. 光电工程, 47(1): 190161. Lu Caiwu, Qi Fan, Ruan Shunling, 2020. An open-pit mine roadway obstacle warning method integrating the object detection and distance threshold model [J]. Opto-Electronic Engineering, 47(1): 190161.
- 阮顺领, 董莉娟, 卢才武, 等, 2022. 基于RCR_YOLOv4的矿井巷道红外障碍检测研究[J]. 黄金科学技术, 30(4): 603-611. Ruan Shunling, Dong Lijuan, Lu Caiwu, et al, 2022. Research on infrared obstacle detection of mine roadway based on RCR_YOLOv4 [J]. Gold Science and Technology, 30(4): 603-611.
- 王国法, 王虹, 任怀伟, 等, 2018. 智慧煤矿 2025 情景目标和发展路径[J]. 煤炭学报, 43(2): 295-305. Wang Guofa, Wang Hong, Ren Huaiwei, et al, 2018. 2025 scenarios and development path of intelligent coal mine [J]. Journal of China Coal Society, 43(2): 295-305.
- 张静, 2022. 基于多传感器融合的露天矿山障碍物检测方法研究[D]. 长沙: 湖南大学. Zhang Jing, 2022. Research on obstacle detection methods for open-pit mines based on multi-sensor fusion [D]. Changsha: Hunan University.

Obstacle Detection on the Forward Path for Autonomous Mining Trucks Based on MEBP-YOLOv10

GU Qinghua^{1,2}, CHEN Yongqi^{1,2}, LI Xuexian^{2,3}, WANG Dan^{2,3}

1. School of Resource Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China;

2. Xi'an Key Laboratory of Intelligent Industrial Perception Computing and Decision-Making, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China;

3. School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China

Abstract: With the extensive implementation of unmanned driving technology in open-pit mines, the challenge of obstacle detection for autonomous trucks operating in complex mining environments has become increasingly significant. To address issues such as low accuracy in multi-scale target detection and inadequate feature fusion for small targets, an obstacle detection model based on MEBP-YOLOv10 is proposed for the forward path of mining trucks. Initially, to improve the model's feature extraction capabilities cost-effectively, certain C2f modules are substituted with the C2f-MS module. Drawing inspiration from GhostNet, the C2f-MS module

achieves superior feature maps with a reduced number of parameters. Furthermore, the ECA attention mechanism is integrated into the backbone network. By assigning varying weights to each channel, it captures inter-channel relationships, enhances the extraction of obstacle features, and manages computational costs. Subsequently, to tackle the challenge of detecting small targets, the original PANet in YOLOv10 is replaced with BiFPN. Through bidirectional cross-scale connections and weighted feature fusion, the model's capacity to integrate features of small obstacle targets is improved. The BiFPN employs a residual architecture, incorporating additional connections between the original input and output nodes within the same layer to preserve the integrity of feature information. Subsequently, the PIoU (Powerful-IoU) loss function is introduced to supplant the bounding box regression loss function in YOLOv10, addressing the limitations of ineffective penalty terms and thereby enhancing the model's convergence rate and obstacle detection accuracy. Experimental results demonstrate that the algorithm achieves an average detection accuracy of 89.8%, a recall rate of 79.8%, and a mean Average Precision (mAP) of 83.5%. In comparison to the original YOLOv10 model, accuracy and mAP are improved by 4.7% and 4.2%, respectively. Moreover, the model surpasses current mainstream object detection networks in terms of accuracy, recall rate, and other performance metrics. Additionally, the detection speed of this model reaches 103.4 frames per second (FPS), satisfying the real-time requirements for obstacle detection on mining truck paths. The model's size is merely 6.87 MB, rendering it suitable for deployment on edge devices. Therefore, MEBP-YOLOv10 enables real-time and accurate obstacle detection on mining truck paths in mountainous mining areas, ensuring safe driving for unmanned trucks.

Key words: open-pit mine; autonomous mining truck; machine vision; obstacle detection; feature fusion

中国科研团队发现新稀土矿物“钕黄河矿”

记者7月17日从中国地质大学(武汉)获悉,该校联合内蒙古自治区地质调查研究院,在内蒙古白云鄂博矿床主矿段的矿体中部,发现一种新稀土矿物,命名为“钕黄河矿”。

据悉,“钕黄河矿”命名已获得国际矿物学协会新矿物命名及分类委员会全票批准通过。

研究团队负责人、中国地质大学(武汉)地质过程与成矿预测全国重点实验室研究员赵来时介绍,钕黄河矿隶属于稀土碳酸盐矿物家族,其化学成分具有独特的钕元素富集特征。钕作为高性能永磁材料的核心成分,在新能源汽车、风力发电、电子信息等领域需求旺盛。

白云鄂博矿是世界最大的稀土矿床,也是中国矿物资源宝库,此前已发现210余种矿物。赵来时表示,钕黄河矿的发现,进一步彰显了该矿床的复杂性与资源多样性。

赵来时介绍,团队在开展项目研究时,通过系统野外地质调查及地球化学分析,揭示了白云鄂博矿床的稀土元素配分模式及赋存规律。在此基础上,已建立白云鄂博矿床的矿体分带特征及矿床成因模型。

钕黄河矿便是在该成矿模型理论指导下,通过矿物微区分析取得的新发现。此前,团队还发现另外两种白云鄂博新稀土矿物——氟碳钕矿、钕独居石。

(来源:中新社)