

# 基于改进 U-Net 的干式磁选矿带图像分割<sup>①</sup>

刘石梅, 肖晶峰, 刘洋, 黄勇, 肖盛旺, 张胜广

(长沙矿冶研究院有限责任公司, 湖南长沙 410012)

**摘要:** 为解决干式磁选过程中矿带不确定性问题, 采用机器视觉技术, 提出一种基于改进 U-Net 模型的图像分割方法。该模型利用 CBAM 注意力模块, 提高网络对目标区域的识别和关注能力, 有助于实现复杂背景下目标物体的图像分割; 采用深度可分离卷积, 降低计算复杂度的同时兼顾精度, 为获取分辨率较高的矿带图像提供有力支持, 从而适应磁选场景, 改善网络性能。该模型分割精度为 92.28%, 轮廓提取完整性和去噪能力优于经典 U-Net、DeepLabV3+ 和 PSPNet 模型。

**关键词:** 干式磁选; 图像识别; 图像分割; 机器视觉; U-Net; CBAM 注意力机制; 深度可分离卷积

中图分类号: TD924

文献标志码: A

doi: 10.3969/j.issn.0253-6099.2024.06.009

文章编号: 0253-6099(2024)06-0041-05

## Image Segmentation Based on Improved U-Net Model for Beneficiation Product Zone in Dry Magnetic Separation

LIU Shimei, XIAO Jingfeng, LIU Yang, HUANG Yong, XIAO Shengwang, ZHANG Shengguang

(Changsha Research Institute of Mining and Metallurgy Co., Ltd., Changsha 410012, Hunan, China)

**Abstract:** Aiming at uncertainty of beneficiation product zone in dry magnetic separation process, an image segmentation method based on an improved U-Net model was proposed by employing machine vision. In this improved model, convolutional block attention module (CBAM) is utilized to enhance the recognition and attention of the network for target areas, which is beneficial to the segmentation of target objects under complex backgrounds; depth-wise separable convolution is adopted to reduce computational complexity while maintaining accuracy, providing strong support for obtaining high-resolution images of beneficiation product zone. Thus, this model can be applied in magnetic separation and also improve network performance. It is found that this improved model can bring segmentation accuracy up to 92.28%, and also is superior to classic U-Net, DeepLabV3+ and PSPNet models in terms of contour extraction completeness and denoising capabilities.

**Key words:** dry magnetic separation; image recognition; image segmentation; machine vision; U-Net; convolutional block attention module (CBAM); depth-wise separable convolution

干式磁选是利用矿物之间磁性差异实现分离的重要矿物加工技术, 具有简单、环保、高效的特点, 在矿物处理中应用广泛<sup>[1]</sup>。为实现全过程监测与控制, 准确获取不同矿物分选矿带图像至关重要。但工业现场的复杂环境会导致获取的图像存在遮挡、模糊等问题, 给图像分割与分析带来困难<sup>[2-3]</sup>。针对这一问题, 本文提出一种基于改进 U-Net 的磁选矿带图像分割方法: 在经典 U-Net 模型<sup>[4]</sup>中添加 CBAM (convolutional block attention module) 注意力模块<sup>[5]</sup>, 显著提升模型对矿带目标区域识别和聚焦能力; 其次, 采用深度可分离卷积<sup>[6]</sup>技术, 使得网络能够在每一层实现特征密集连

接, 有效利用浅层与深层特征, 提高网络对目标区域的识别和关注能力, 大幅减少网络参数量和计算负担, 改善网络性能, 提高分割准确性<sup>[7]</sup>; 在此基础上, 构建磁选图像数据集, 采用有效数据增强策略扩充样本, 优化网络训练过程; 最后, 利用训练好的模型对复杂工业场景图像进行分割处理和参数分析, 为干式磁选智能化控制提供坚实的技术基础。

## 1 改进 U-Net 矿带识别网络

### 1.1 U-Net 网络

U-Net 是由 Ronneberger 等人在 2015 年提出的一

① 收稿日期: 2024-05-25

基金项目: 国家重点研发计划(2021YFC2902701); 湖南省科技创新计划(2022RC1053)

作者简介: 刘石梅(1985—), 男, 湖南新化人, 高级工程师, 主要研究方向为选矿装备研制及选冶过程自动化。E-mail: 358971237@qq.com

通信作者: 张胜广(1973—), 男, 山东曹县人, 正高级工程师, 主要研究方向为选冶自动化。E-mail: guang1098@163.com

种创新性网络结构,它的编码器-解码器架构能够在保留原始空间信息的同时,精确恢复图像细节,非常适合进行精细的像素级图像分割。其出色的特征提取能力和高效的信息传递机制,能够处理各种复杂的图像分割场景,包括但不限于产品质量检测、机器视觉检测等,在工业领域得到广泛应用和认可。

虽然 U-Net 在图像分割领域具有卓越表现和广泛适用性,但 U-Net 网络层数有限,提取的特征表达能力受到限制;而且随着网络深度增加,梯度消失问题会导致学习效率降低、准确率无法进一步提高。另外,直接跳跃连接无法突出目标区域信息,容易受到无关信息的干扰。因此,可以借鉴计算机视觉领域的注意力机制,通过学习数据内在关联来突出重要特征,忽略无关噪声信息。引入 CBAM 注意力模块改进 U-Net 网络,从理论上有望获得一种快速、轻量化且高精度的干式磁选矿带识别算法。

### 1.2 CBAM 注意力模块

CBAM 是一种用于前馈卷积神经网络的、先进的注意力机制模块。该模块在空间和通道两个独立维度上计算注意力图,通过与输入特征图相乘实现自适应特征优化。CBAM 由两个核心组件构成:通道注意力模块和空间注意力模块,如图 1 所示。通道注意力模块学习特征图的通道间关系,自适应调整各通道权重,增强网络对重要特征通道的感知能力。空间注意力模块学习特征图的空间维度关系,自适应调整不同位置的权重,提高网络对关键区域的关注度。实现过程中, CBAM 首先通过最大池化和平均池化获取空间显著区域特征和全局特征,结合这些信息生成空间注意力图。同时,利用多层感知机生成通道注意力,对不同语义特征进行加权。最后,将空间注意力和通道注意力相结合,实现对目标区域的精确聚焦。通过引入 CBAM,网络能更准确识别和分析关键特征,从而在诸如矿带目标区域分割等任务中显著提升性能。

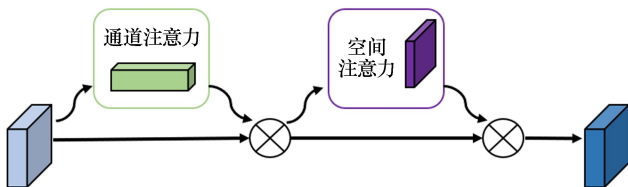


图 1 CBAM 注意力模块

给定一个中间特征图  $F$ ,以属性高  $H$ 、宽  $W$  和维度  $C$  作为输入,可表示为:

$$F \in R^{H \times W \times C} \quad (1)$$

CBAM 依次得到一个一维的通道注意力图和一个

二维的空间注意力图,整个流程可以概括为:

$$F' = M_c(F) \otimes F \quad (2)$$

$$F'' = M_s(F') \otimes F' \quad (3)$$

式中: $M_c \in R^{1 \times 1 \times C}$ ;  $M_s \in R^{H \times W \times 1}$ ;  $\otimes$  表示对应元素相乘操作。

CBAM 使网络能够同时自适应地关注重要的通道和空间位置,注意力机制使得网络能够自适应地分配计算资源,提高网络区分能力,增强其在视觉任务中的性能。

### 1.3 深度可分离卷积

深度可分离卷积是一种轻量级的卷积操作,通过将标准卷积分解为深度卷积和逐点卷积两个步骤,有效减少网络参数量和计算复杂度,同时保持良好的特征表示能力。深度可分离卷积原理如图 2 所示。在 U-Net 架构中应用深度可分离卷积可显著降低模型参数量、提高网络运行效率,并在某些情况下改善模型性能。具体而言,深度卷积通过独立处理每个输入通道,有效捕捉空间信息;逐点卷积利用  $1 \times 1$  卷积核整合不同通道的特征。这种结构能在保持特征表示能力的同时大幅降低计算负担。深度可分离卷积的应用使得 U-Net 模型能够在资源受限的环境中更加高效地运行,同时可能提升特征提取能力和模型的泛化性能。

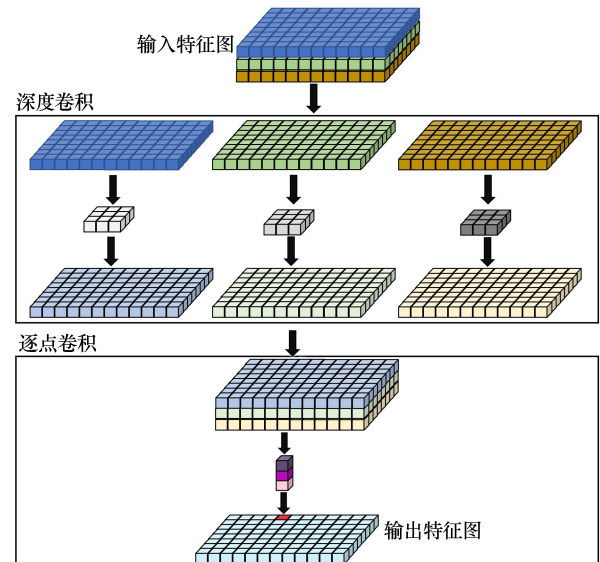


图 2 深度可分离卷积原理图

### 1.4 改进 U-Net 网络结构与训练策略

针对原始 U-Net 模型中有限的上采样层特征深度导致复杂矿带图像分割精度较低的问题<sup>[8]</sup>,将所有下采样的特征图沿通道维度连接起来,并将它们作为输入馈送到修改后的模型中每个相应的上采样层,从而实现更大的特征深度<sup>[9]</sup>。此外,集成深度可分离卷积,以

减少模型大小和计算成本。

改进 U-Net 网络模型结合 CBAM 注意力模块和深度可分离卷积,提高模型关注能力并减少模型大小和计

算成本,不同层之间进行特征连接,通过有效组合和利用串接特征图,实现对磁选矿带更精确的分割。改进 U-Net 网络架构如图 3 所示。

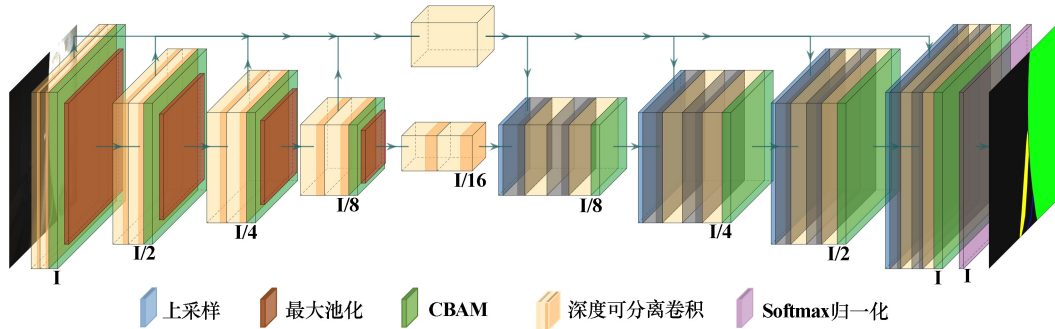


图 3 改进 U-Net 网络架构图

通过深度可分离卷积将 U-Net 每层的特征水平密集连接起来,自适应地有效利用浅层和深层特征。在编码器和解码器中引入密集连接,融合来自较高层和较低层的多分辨率特征图,以加强当前层中的特征传播,从而改善编码器和解码器之间的信息流。编码器的每层逐层向下层连接与解码器的每层逐层向上层连接形成密集连接,获得不同级别的图像组合。

U-Net 对处理小样本数据集的效果非常突出,但图像中存在颗粒粘连导致欠分割的情况。图像中矿带分布并不总是均匀的,可能部分图像中矿带较少而背景较多,因此采用 dice 损失函数和交叉熵损失函数共同训练,同时在训练时赋予颗粒更多权重,使得模型更加关注矿带边界,改善颗粒粘连导致的欠分割情况。总损失函数公式如下:

$$L_{\text{loss}} = \alpha L_{\text{bec}} + \beta L_{\text{dice}} \quad (4)$$

式中: $L_{\text{loss}}$  为总损失函数; $L_{\text{bec}}$  为交叉熵损失函数; $L_{\text{dice}}$  为 dice 损失函数; $\alpha$  和  $\beta$  为两个损失函数的权重,皆设置为 1。

交叉熵损失函数计算公式为:

$$L_{\text{bec}} = -y \ln y' - (1 - y) \ln(1 - y') \quad (5)$$

式中: $y$  为真实值; $y'$  为预测值。该损失函数能够公正地对待每一个像素点,但对于背景占比较多的图像,其效果不够理想。

dice 损失函数的计算公式为:

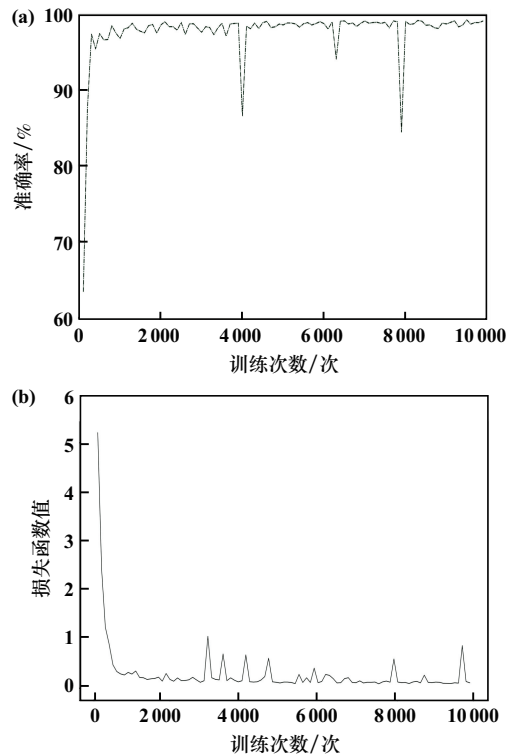
$$L_{\text{dice}} = 1 - 2 \times \frac{A \cap B}{A \cup B} \quad (6)$$

式中  $A$  和  $B$  分别为预测图与标签图。通过衡量两张图像的相似性,可计算该损失函数。

对比式(5)和式(6)可以看出, $L_{\text{bec}}$  损失函数侧重每一个像素点的预测,平等看待前景和背景; $L_{\text{dice}}$  损失函数注重区域之间的相似性,能够从总体上评估预测

结果,理论上会更加关注前景。两者结合能兼顾像素点和整体目标的训练,采用这两个损失函数共同训练 U-Net 网络模型,通过引入轻量化、注意力机制和多尺度融合提升整个模型的精度。

采用 Adam 优化器,学习率为 1%,批量大小为 8,采用 L2 正则化,权重衰减为  $1 \times 10^{-4}$ ,学习率下降策略为按步长下降。改进模型采用 NVIDIA RTX 4090 GPU 进行训练,经历了 10 000 次的训练过程。训练过程中的准确率和损失函数值如图 4 所示,模型损失函数值稳定时约为 0.07,同时网络训练过程的损失曲线光滑,网络收敛且准确率较高。



(a) 训练集准确率; (b) 训练集损失函数值

图 4 改进 U-Net 训练过程准确率和损失函数值

## 2 实验结果与分析

### 2.1 实验数据集

在不同环境中采集磁选分选图像 8 000 张,同时对难以注释的图像进行等分及标记处理,使用 Label-Me 软件对图像进行人工标注,精确绘制磁性颗粒和非磁性颗粒矿带,利用旋转、翻转、缩放、平移和颜色抖动等数据增强方法获得有效数据集。

### 2.2 评价指标

对图像分割性能进行评估,将分割结果与人工标注结果对比,根据真实标签将检测结果划分为真正例 (TP, true positive)、真反例 (TN, true negative)、假正例 (FP, false positive) 和假反例 (FN, false negative) 共 4 类。评估指标选择图像分割常用的准确率  $A$ 、精确率  $P$ 、召回率  $R$  和  $F_1$  值对矿带提取结果进行精度评价为:

$$A = \frac{n(\text{TP}) + n(\text{TN})}{n(\text{TP}) + n(\text{FP}) + n(\text{TN}) + n(\text{FN})} \quad (7)$$

$$P = \frac{n(\text{TP})}{n(\text{TP}) + n(\text{FP})} \quad (8)$$

$$R = \frac{n(\text{TP})}{n(\text{TP}) + n(\text{FN})} \quad (9)$$

$$F_1 = 2 \frac{P \times R}{P + R} \quad (10)$$

式中  $n$  为相应的标签个数。

### 2.3 消融实验

训练了 4 个不同网络模型进行消融实验,以验证本文所构建的网络中各模块的有效性,结果如表 1 所示。其中模型 1 为原始 U-Net 模型;模型 2 在模型 1 基础上引入深度可分离卷积,将 U-Net 网络每层的特征水平密集,实现多尺度特征提取;模型 3 为改进的 Attention U-Net 模型,在每一个卷积块后添加 CBAM 注意力模块;模型 4 为本文提出的模型。消融实验结果表明,本文提出的模型能够有效提高网络的精确率和召回率。

表 1 消融实验结果

模型编号	准确率/%	精确率/%	召回率/%	$F_1$ /%	时间/ms
1	88.32	88.10	84.24	86.12	50.85
2	87.70	83.01	83.29	86.75	27.28
3	91.65	90.44	90.95	90.69	58.88
4	92.28	91.36	92.18	91.43	29.32

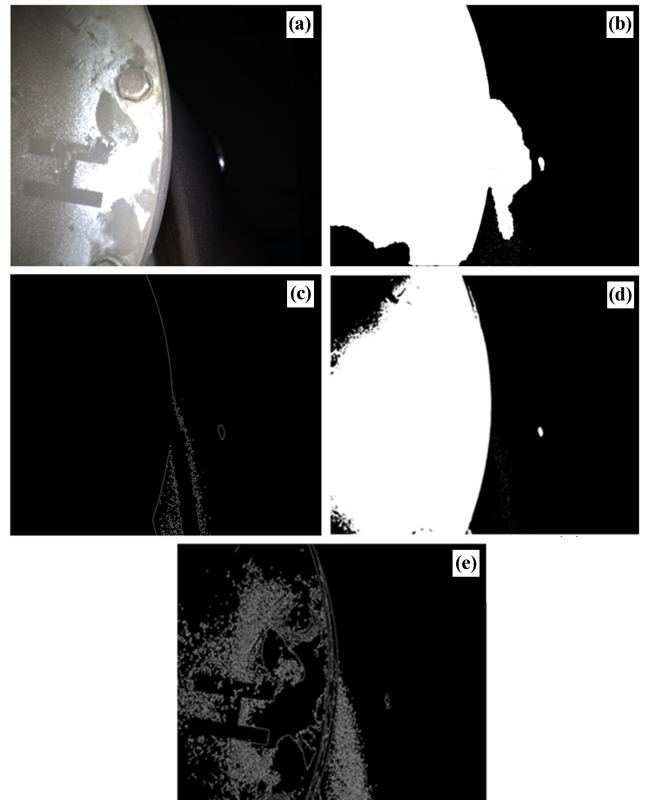
CBAM 对 U-Net 网络架构的影响主要通过增强 U-Net 在分割任务中的性能和表达能力来实现。CBAM 的集成能够改善 U-Net 在分割任务中的特征表达能

力、上下文理解能力和准确性,同时保持较低的计算成本。这使得 U-Net 能够更好地适应磁选过程中的图像分割和分析任务,提高干式磁选的效率 and 准确性。

深度可分离卷积在 U-Net 网络架构中具有减少参数量和计算量、增强特征提取能力以及提高模型泛化能力的优点,它的引入使得 U-Net 模型更加轻量化、高效,并具备更好的特征表达和分割能力,适用于资源受限的场景和大规模图像分割任务。

### 2.4 性能对比

传统分割方法在磁选分选图像处理中的效果如图 5 所示。传统图像分割方法主要依赖灰度、纹理和形状等低级特征,在处理复杂磁选分选图像时显现出多方面局限性。这些方法往往难以准确捕捉目标和背景间的本质差异,需要手动调整参数,对噪声和复杂背景敏感,导致分割结果不稳定且质量差异大。特别是对于具有高对比度、复杂纹理和多样形状特征的磁选分选图像,传统方法表现欠佳。相比之下,基于深度学习的方法能更好地克服这些挑战,提供更准确、稳健的分割结果。



(a) 原始图像; (b) 基于区域的分割; (c) 基于边缘的分割; (d) Otsu 分割; (e) 基于轮廓的分割

图 5 传统分割方法在磁选分选图像处理中的效果

训练 4 个不同的网络模型,在自建图像数据集上对 4 个网络模型进行训练和测试,用于处理磁选分选

图像,结果如图 6 和表 2 所示。

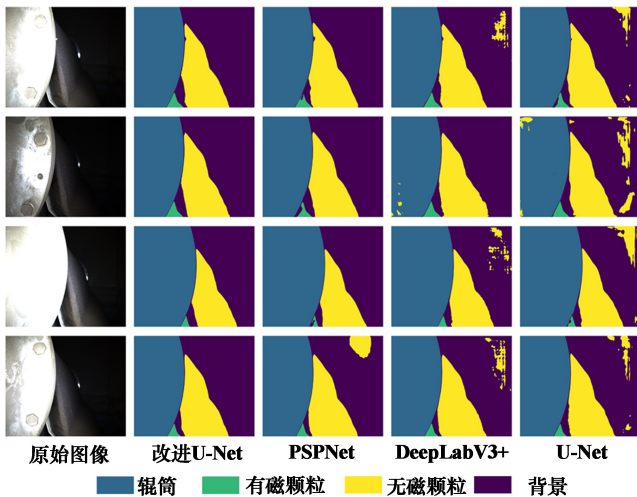


图 6 不同网络模型分割效果对比

表 2 网络模型分割结果

模型名称	准确率/%	精确率/%	召回率/%	$F_1$ 值/%	时间/ms
U-Net	88.32	88.10	84.24	86.12	50.85
PSPNet <sup>[10]</sup>	90.91	90.65	81.44	85.90	59.81
DeepLabV3+ <sup>[9]</sup>	90.04	88.95	88.55	87.42	88.29
改进 U-Net	92.28	91.36	92.18	91.43	29.32

U-Net 模型在本次实验中的各项指标均落于其他 3 个模型之后,准确率(88.32%)、精确率(88.10%)和召回率(84.24%)较低,从可视化结果可以看出,其分割质量存在明显不足,如轮廓残缺、噪声较多等。PSPNet 模型在准确率和精确率方面表现不错,分别达到 90.91%和 90.65%,仅次于改进 U-Net 模型,但该模型召回率仅 81.44%,导致其  $F_1$  值只有 85.90%,存在一定的性能缺陷;从可视化结果来看,其分割的目标轮廓整体较为完整,但在一些细节区域仍存在分割残留和噪声。DeepLabV3+模型准确率(90.04%)和精确率(88.95%)表现中等,略于 PSPNet 模型, $F_1$  值为 87.42%,但该模型推理时间长(88.29 ms);从可视化效果上看,其存在一些轮廓不完整、目标分离不够彻底的问题。改进 U-Net 模型在实验中展现出了极为优异的综合能力,不仅在准确率、精确率、召回率和  $F_1$  值等评估指标上全面领先,且推理时间短(29.32 ms);从可视化结果来看,改进 U-Net 分割的目标轮廓更完整、边缘更锐利、目标与背景分离度更高、细节保留也更加充分,进一步验证了深度可分离卷积和 CBAM 注意力模块相结合可显著改进分割效果。

### 3 结论

本文所建立的模型融合 CBAM 注意力模块和深度可分离卷积,通过对每层特征水平密集连接克服 U-Net 模型各上采样层特征深度输入局限性,模型尺寸及计算成本得以减小和降低,对矿带的分割能力增强,同时网络对目标区域的关注和识别能力进一步提高。自制数据集验证实验结果表明,本文所建立的模型平均分割准确率达到 92.28%,其效果显著优于 PSPNet 以及 DeepLabV3+模型,可更准确地对目标实现分割。

### 参考文献:

- [1] 吕波,骆振福. 国内磁选机的发展现状和趋势[J]. 矿山机械, 2011,39(7):1-4.
- [2] XU C H. The development of preconcentration technology in low-grade magnetite's beneficiation[J]. Advanced Materials Research, 2012, 361:324-327.
- [3] 马淮湘. 超贫磁铁矿选矿技术新进展与思考[J]. 现代矿业, 2011,27(4):33-34.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention. 2015:234-241.
- [5] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018:3-19.
- [6] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:1251-1258.
- [7] YANG Z, WU H, DING H, et al. Enhanced U-Net model for rock pile segmentation and particle size analysis[J]. Minerals Engineering, 2023,203:108352.
- [8] ZHANG D J, LI Q Q, CHEN Y, et al. An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection[J]. Image and Vision Computing, 2017,57:130-146.
- [9] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018:801-818.
- [10] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:2881-2890.

引用本文: 刘石梅,肖晶峰,刘洋,等. 基于改进 U-Net 的干式磁选矿带图像分割[J]. 矿冶工程, 2024,44(6):41-45.