基于机器视觉的重选摇床白钨矿带智能分割模型

阮顺领[1,2,3],陈思睿[1,3],南欣[1,3],孟令煊[1,3]

1.西安建筑科技大学资源工程学院,陕西 西安 710055;
2.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.西安建筑科技大学矿山系统工程研究所,陕西 西安 710055

摘 要:针对当前重选摇床选矿过程中人工识别矿带效率低、成本高,以及现有分割模型参数量大和精度不足的问题,提出了一种基于改进DeepLabV3+的轻量化矿带智能分割方法。模型以MobileNetV3为基础骨干网络,融入高效通道注意力机制,替代原Xception结构,有效降低了模型的复杂度和计算负担;在编码器部分引入串联连接结构并结合膨胀率分别为3和9的空洞卷积分支,增强了多尺度特征融合能力与感受野的连续性;解码器部分则集成可变形卷积和CBAM注意力机制,以应对矿带形态变化并抑制背景噪声干扰,从而提高分割精度。试验结果表明,基于自建紫外光矿带数据集,该模型取得了良好指标,*Dice*系数达到93.89%,*MIOU*达到81.04%,mAP@50:95达到91.97%,且参数量仅为18.32 M,计算复杂度为33.50 GFLOPs,模型在保持高分割精度的同时具有良好的实时性。与现有主流分割模型相比,本文方法在精度和轻量化方面均表现出明显优势,为重选摇床的无人化监控与智能控制提供了可行的技术方案。

关键词:选矿摇床;机器学习;重力选矿;语义分割;图像处理;矿带检测

中图分类号:TD928.9 **文献标志码:**A **文章编号:**1005-2518(2026)04-0939-15 DOI:10.11872/j.issn.1005-2518.2026.04.363

引用格式:阮顺领,陈思睿,南欣,等.基于机器视觉的重选摇床白钨矿带智能分割模型[J].黄金科学技术,2026,34(4):939-953.RUAN Shunling,CHEN Sirui,NAN Xin,et al.Intelligent Segmentation Model of Scheelite Belt in Gravity Separation Shaking Table Based on Machine Vision[J].Gold Science and Technology,2026,34(4):939-953.

当前,随着我国工业化进程的持续推进,对矿产资源的需求日益增长。推动机器视觉、大数据和智能机器人等数字技术与矿业生产深度融合,已成为建设智能矿山、实现行业高质量发展的核心路径(郑玉荣等,2025)。在此背景下,基于机器视觉的图像分割技术逐步应用于采矿工程的多个环节(田妞等,2023),为流程监控和工艺优化提供了新的技术手段。在重选摇床分选过程中,矿物在床面上按密度分层,形成精矿带、中矿带和尾矿带等呈扇形分布的区域。然而,由于给矿量、矿浆浓度和粒度等工况参数的波动,各矿带在颜色、位置和宽度上呈现动态变化,目前主要依赖人工观察和调整,导致分选效率和精矿品位稳定性难以保证。

国内外学者围绕摇床矿带的视觉分割开展了诸多研究。摇床分选过程的智能化研究,其技术范式经历了从机理模型到数据感知的演进。早期研究侧重于从物理原理出发进行理论推导和仿真,如:通过流体流动模拟来建立增强型摇床的配分函数模型(Kroll-Rabotin et al,2010)。然而,由于实际生产环境复杂多变,基于直接感知的视觉检测与数据驱动方法因其更强的适应性而成为近年来的主要研究方向,并沿着“图像处理—智能识别—系统

收稿日期:2025-12-08;**修订日期:**2026-05-11

基金项目:国家自然科学基金面上项目“基于融合学习的湿排尾矿堆积坝时空演化及失稳预警模型研究”(编号:52374160)和陕西省技术创新引导专项基金重点研发计划项目“露天矿高效智能开采关键技术及产业化应用”(编号:2023GXLH-062)联合资助

作者简介:阮顺领(1981—),男,河南周口人,教授,博士生导师,从事矿山智能科学与工程研究工作。E-mail:ruanshunling@163.com

***通信作者:**陈思睿(1999—),女,内蒙古包头人,硕士研究生,从事资源与环境智能矿山研究工作。E-mail:1559204745@qq.com

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

集成”的路径快速发展。在图像处理与特征提取阶段，研究人员采用多种算法从图像中直接定位矿物带边界，例如：融合深度学习边缘检测与 Hough 变换，提取了摇床矿带的重要边界（刘惠中等，2023）；采用 MSR 进行了图像预处理，并使用 OTSU 算法进行左侧矿带图像的分割（王锦煜，2022），还有磷虾优化算法（和丽芳等，2018）以及利用最小二乘法获得摇床分割带中心矢量的方法（杨文旺等，2020）。随着科技的进步，研究重点进入深度学习智能预测阶段，旨在建立端到端的映射模型，如：使用 VGG-16 网络预测接矿位置（赵玉华等，2023），设计更为复杂的 CNN-GRU-Attention 混合模型进行精矿边界的回归定位（刘惠中等，2025）。部分研究正逐步向系统集成和自动控制工程化深入，如：杨文龙等（2016）开发的集识别和调节于一体的自动选矿装置，标志着该项技术向闭环控制与工业落地迈出了关键一步。

综上所述，虽然已有研究在摇床矿带的分割上具有一定效果，但现有方法并不能满足企业的轻量化、精准高效和快速识别的需求。为实现摇床分选过程的精准感知与智能控制，本研究聚焦摇床矿带的实时识别问题，旨在研究一种兼具高精度与轻量化特征的图像分割模型。该模型需在保证分割性能的前提下，尽可能降低参数量和计算复杂度，以适应实际工业场景中对实时性和部署成本的要求。借鉴矿带分割方法和轻量级神经网络的优化策略，选择分割能力更强且运行效率更高的 DeepLabV3+ 模型，并对其骨干网络、解码器和解码器进行相应改进，旨在解决颜色相近矿带分割不清晰，以及分割速度较慢和算法模型较大等问题，经过实验检测，显著提升了 DeepLabV3+模型对摇床矿带分割的准确性和时效性，进一步减少了模型的参数量，为后续构造智能化选矿摇床提供了技术支撑。

1 改进DeepLabV3+轻量化矿带分割模型

DeepLabV3+模型是在 DeepLabV3 基础上进行的重要升级（Chen et al，2018），后者作为语义分割领域的代表性模型，已被广泛应用于各类语义分割任务中。DeepLabV3+模型在编码器部分延续并增强了空洞空间金字塔池化（ASPP）模块，通过多个不同空洞率的并行卷积和全局池化操作，有效融合多尺度上下文语义信息。在解码器设计中，模型引入轻量级上采样结构，将经过通道压缩的低层特征与富含语义的高层特征进行拼接融合，显著提升了目标边界的重建精度。为进一步实现效率与性能的平衡，DeepLabV3+模型支持通过调整输出步幅和相应层的空洞率，灵活控制感受野大小和计算复杂度。在损失函数方面，除了使用交叉熵作为基础监督，还可结合 Dice 损失或针对性设计的边界感知损失。DeepLabV3+模型的具体结构如图 1 所示。

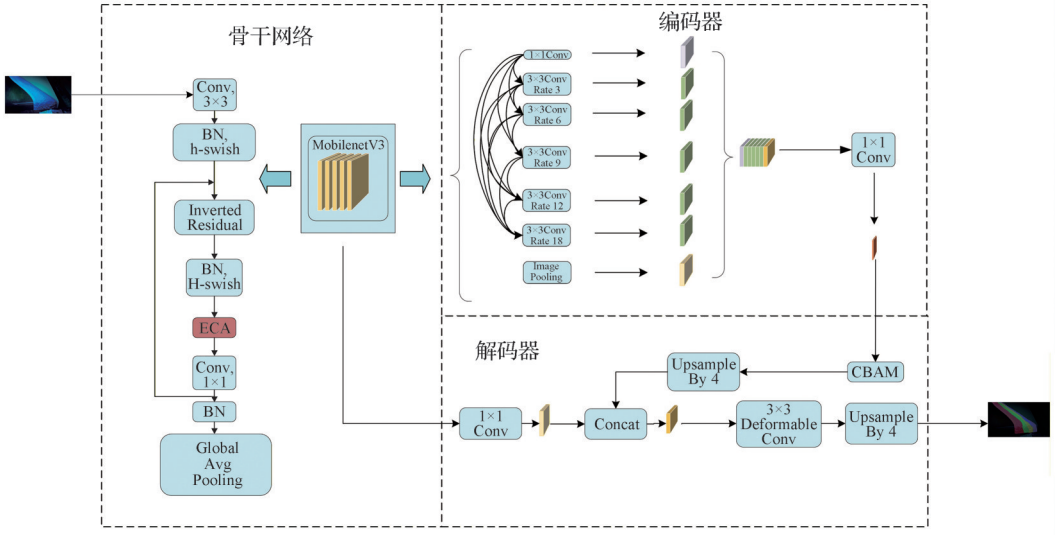


图 1 改进后的 DeeplabV3+模型网络

Fig. 1 Improved DeepLabV3+ model network

由上述可知,由于摇床钨矿带中品位相近的矿带特征较为复杂,针对标准DeepLabV3+模型在摇床矿带中存在的特征提取能力不足、分割精度不够以及无法有效识别矿带形变不足之处。为此,本文提出了一种融合多种基础创新点的深度神经网络架构,基于DeepLabv3+语义分割模型进行骨干网络、编码器和解码器改进,改进后的网络结构如图1所示。

(1)骨干网络改进:针对原有DeepLabV3+模型骨干网络参数量大和特征提取不足的问题,本研究使用添加了ECA(Wang et al,2020)注意力机制的MobileNetV3骨干网络(Howard et al,2019)代替原有的Xception骨干网络(Chollet,2017)。MobileNetV3的轻量化特性使得模型分割精度在小幅下降的同时,极大地降低了模型的计算量,同时ECA增加了模型的表征能力和浅层特征提取能力,在降低参数量的同时提升了模型对矿带的浅层能力,弥补了替换骨干网络后导致的分割精度的缺失。

(2)编码器改进:使用DenseASPP(密集空洞空间金字塔)替代原有的池化金字塔,并在原有膨胀率为r=6,r=12,r=18的空洞卷积中添加了膨胀率为r=3和r=9的空洞卷积。密集链接使得图像特征从r=3的时候就开始特征融合,r=3和r=9的空洞卷积分支,增大了原有模型的感受野,强化了多尺度特征融合,减少特征损耗,从而提升了精矿带的分割精度和稳定性。

(3)解码器改进:在解码器中添加可变形卷积(Xia et al,2022)和CBAM(Woo et al,2018)注意力机制。可变形卷积通过为标准卷积核的每个采样点引入一个可学习的偏移量,使卷积核能够根据输入特征的内容自适应地调整其感受野的形状,从而在矿带形变时,能够更好地拟合矿带的形状,从而提升分割精度。CBAM的通道注意力机制放大了不同矿带之间的微小差异,强化了矿带的纹理特征和颜色特征,抑制背景噪声,提升相近矿带的区分度,最终让解码器输出更精细的分割结果。

2 摇床矿带分割模型优化设计

2.1 改进DeepLabV3+骨干网络模型

针对现有摇床选矿模型的浅层语义特征提取不完全和参数量大的问题,本研究使用添加了ECA模块的MobileNetV3替代原Xception骨干网络。MobileNetV3网络使用逆残差线性瓶颈结构的设计,首先使用1×1逐点卷积提升特征维度,随后利用3×3深度可分离卷积在更高维空间中进行高效的稀疏特征提取,最后通过线性1×1逐点卷积将特征投影回低维空间。其中,深度可分离卷积将标准卷积分解为深度卷积和逐点卷积两步,极大地减少了模型的参数总量和计算开销;而线性瓶颈的设计则移除了最后一个1×1卷积后的非线性激活函数,有效避免了低维空间下的特征信息损失。

本研究使用ECA模块替代MobileNetV3自带的SE模块(Hu et al,2018),ECA模块具体结构如图2所示。ECA模块通过全局平均池化获取通道特征的全局描述,并摒弃了传统的全连接层,转而采用一维卷积来直接计算通道间的局部交互,从而生成通道权重。这种方法在避免降维带来的信息损失的同时,以极低的计算开销实现了跨通道的信息交互,引导模型聚焦于关键特征通道。

在ECA注意力机制中,首先通过全局平均池化生成通道级统计描述向量,对于输入特征图$X \in R^{C\times H\times W}$,第$c$个通道的全局平均池化结果定义为

$$g_c=\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}X_c(i,j) \tag{1}$$

得到池化后的向量$\boldsymbol{g} \in \mathbb{R}^C$后,ECA模块利用一维卷积来捕获局部通道间的依赖关系,从而避免使用全连接层,减少计算量。注意力权重的生成过程如下:

$$M_c(F)=\sigma\left(ConvD_k\left(F_{\text{avg}}\right)\right) \tag{2}$$

式中:$M_c(F)$为基于通道的注意力权重;$\sigma(\cdot)$为Sigmoid激活函数;$Conv1D_k$表示核大小为k的一维卷积操作;F_{avg}为特征图在空间维度上的平均池化结果。

卷积核大小k并非固定不变,其计算公式为

$$k=\left|\frac{\log_2(C)}{\gamma}+b\right|_{\text{odd}} \tag{3}$$

式中:γ和b为超参数,默认取γ=2,b=1;$|\cdot|_{\text{odd}}$表示取最接近的奇数,以保证一维卷积的对称性;C为通道数。

随后,通过一维卷积和Sigmoid激活函数得到

注意力权重向量$\boldsymbol{\alpha}\in\mathbb{R}^C$。

$$\boldsymbol{\alpha}=\sigma\left(Conv1D(s,k)\right) \tag{4}$$

式中：s为输入到$Conv1D$卷积的特征向量；k为卷积核的大小；σ为通道注意力权重；α为注意力权重。

最终，将注意力权重与原始特征图逐通道相乘，完成特征重标定：

$$X^*=X\odot\alpha \tag{5}$$

式中：X为输入的特征图；α为注意力权重；$\odot$为逐通道的乘法操作。

ECA模块通过省略全连接层的降维与升维过程，将计算复杂度从$O(C^2)$降低至$O(Ck)$，在保持良好注意力效果的同时，显著减少了矿带分割模型参数量，提升了矿带分割模型的训练效率和分割准确性，尤其适用于强化对纹理细节信息的感知能力。

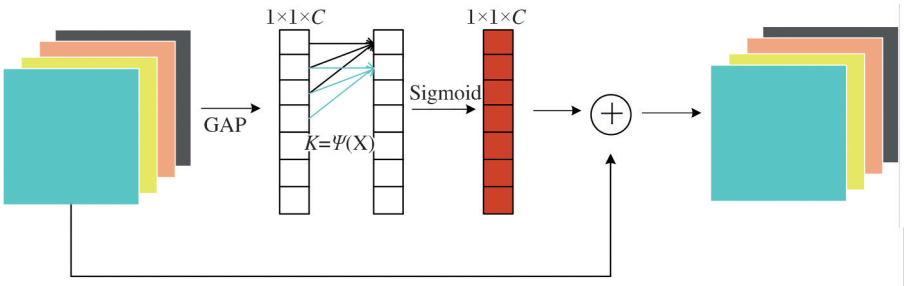


图2　ECA注意力机制结构

Fig. 2　Structure of ECA attention mechanism

2.2　矿带分割编码器优化

选矿摇床中的精矿带、中矿带和尾矿带存在形态各异、宽度动态变化的特性。这些矿带通常边界模糊、对比度低，尤其在光照不均或存在水膜反光时，容易与邻近矿带及背景区域混淆，难以被有效分割。此外，由于不同矿带间的视觉特征差异有限，语义区分度较低，容易导致分割边界不准或类别误判，特别是对于宽度较窄的精矿带，更容易出现断裂和漏分割的情况。

为了解决多尺度矿带特征捕获不连续的问题，本研究对DeepLabV3+模型中的ASPP模块进行了改进。传统的ASPP模块采用多分支并行结构，其各分支的感受野离散（DeepLabV3+自带的池化金字塔膨胀率为r=6，12，18），且分支间无信息交互，导致对不同宽度矿带的特征提取存在间隙，这是一种“迟融合”策略。为进一步增强模型对从细矿带到宽矿带的连续尺度特征的捕捉能力，本文设计了一种密集连接的空洞卷积（Yu et al，2016）金字塔池化模块，其结构如图3所示。

该模块的核心是序列化的空洞卷积层。空洞卷积通过在标准卷积核的权重之间插入零值来扩大感受野，其输出特征图尺寸的计算公式为

$$H_{\text{out}}=\frac{H+2p-\left[k+(k-1)(r-1)\right]+s}{s} \tag{6}$$

$$W_{\text{out}}=\frac{W+2p-\left[k+(k-1)(r-1)\right]+s}{s} \tag{7}$$

式中：H和W分别为输入特征图的高度和宽度；H_{out}和W_{out}分别为经空洞卷积操作后输出特征图的高度和宽度；p为填充参数；k为卷积核大小；r为空洞卷积的扩张率；s为步长参数。

基于此，本文在原始ASPP（通常包含r=1，6，12，18的卷积层）的基础上，新增了膨胀率为r=3和r=9的2个空洞卷积层，从而构建了一个覆盖更细粒度和感受野连续增长的膨胀率序列（r=1，3，6，9，12，18）。模块中每个卷积层的操作不再独立，而是采用密集连接方式，其计算过程定义为

$$F_i=Conv_{r_i}\left(Concat\left(F_0,F_1,\cdots,F_{i-1}\right)\right) \tag{8}$$

式中：F_0为模块的输入特征图；$Conv_{r_i}$表示利用式（6）和式（7）所描述的、膨胀率为r_i的空洞卷积操作；$Concat\left(F_0,F_1,\cdots,F_{i-1}\right)$表示通道维拼接。这种结构使得任一卷积层均可直接访问并整合所有前置层的多尺度特征，实现了特征的持续复用和感受野的指数级扩展，是一种高效的“早融合”策略。

对于大面积且颜色相近的矿带，这种强大、连续的多尺度上下文信息有助于模型将整片矿带识

别为一个连贯的整体,有效抑制因内部轻微颜色波动而导致的分割噪声;对于狭窄的精矿带,其精细的局部特征得以在后续层中被持续利用,显著降低了断裂和漏分割的概率。

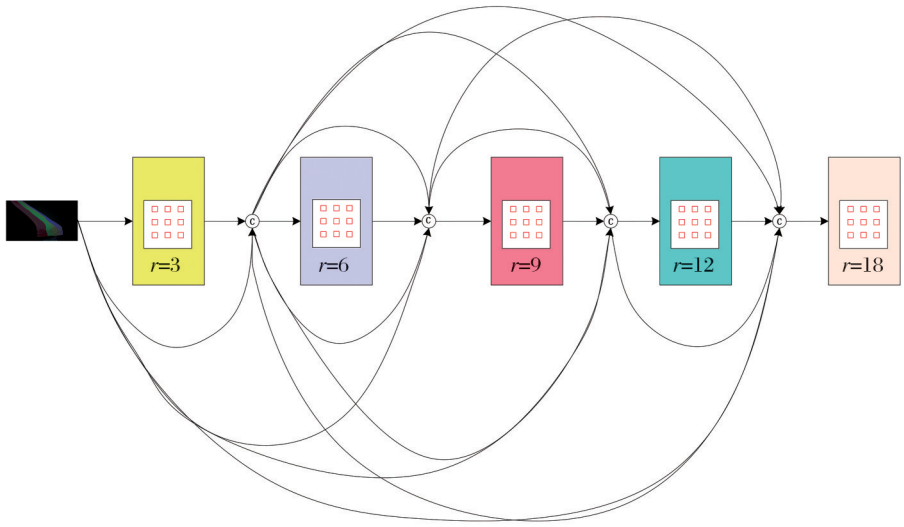


图3　密集链接结构

Fig. 3　Densely connected structure

2.3　矿带分割解码器优化

在选矿摇床的矿带分割任务中,精矿带、尾矿带和中矿带往往因矿物颗粒分布不均匀、边界模糊、光照条件多变以及噪声干扰而难以准确区分。传统的DeepLabV3+模型解码器在特征融合阶段可能无法充分关注矿带的关键区域,导致分割精度受限。因此,在解码器中引入CBAM(Convolutional Block Attention Module)模块和可变形卷积,能使得解码器自适应地优化特征表示,并扩大感受野,从而提高矿带分割的准确性。

CBAM表示卷积模块的注意力机制模块,是一种结合空间和通道的注意力机制模块。图4所示为CBAM模块的整体结构。CBAM主要由CAM和SAM这2个模块组成,分别代表通道注意力模块和空间注意力模块。输入的矿带特征图先后进入通道注意力和空间注意力,之后输出新的特征图。其中,⊗表示元素乘法,如图4所示。

通道注意力模块让通道维度不变,压缩空间维度,结构如图5所示。该模块关注矿带图中的矿带边界等具有独特特征意义的信息。模块将输入的矿带特征图经过2个并行的最大池化层和平均池化层,将特征图的高度和宽度变成1×1的大小,然后经过Share MLP模块。在该模块中,它先将通道数压缩为原来的1/r倍(r为通道缩减率,用于控制Share MLP的通道压缩比例),再扩张到原通道数。经过ReLU激活函数得到2个激活后的结果。将这

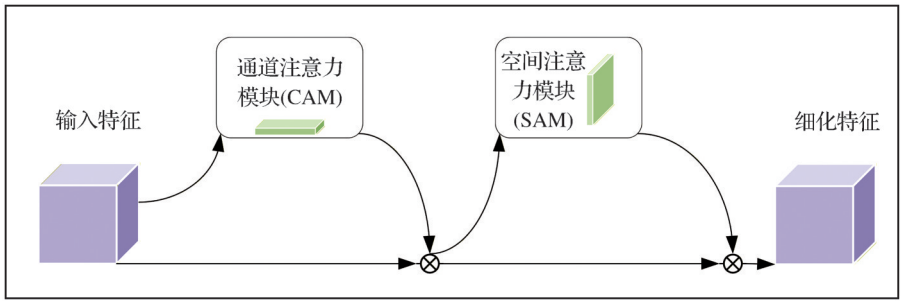


图4　CBAM注意力机制结构

Fig. 4　Architecture of the CBAM attention mechanism

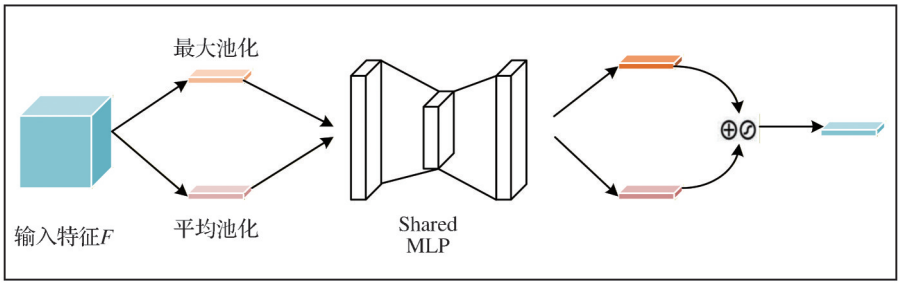


图5 CAM通道注意力模块结构

Fig. 5 Structure of CAM channel attention module

2个输出结果进行逐元素相加，然后通过一个*sigmoid*激活函数得到通道注意力的输出结果。再将这个输出结果乘原图，变回$C\times H\times W$的大小。其通道注意力计算公式为

$$\begin{cases} M_c(F)=sigmoid\left(MLP\left[AvgPool(F)\right]+MLP\left[MaxPool(F)\right]\right) \\ M_c(F)=sigmoid(W_1[W_0(F^c_{avg})]+W_1[W_0F^c_{max}]+\cdots \end{cases} \tag{9}$$

式中：F为输入的特征图；*sigmoid*为激活函数；$AvgPool(F)$和$MaxPool(F)$为特征图经过平均池化和最大池化。

W_0、W_1为Share MLP权重，W_0负责通道压缩，W_1负责通道恢复。M_c为经过通道的注意力；F^c_{avg}为经过通道注意力机制并采取平均池化的特征图；F^c_{max}为经过通道注意力机制并采取最大池化的特征图；$M_c(F)$为让特征图F进入通道注意力机制；MLP为多层感知器。

空间注意力模块让空间维度不变，压缩通道维度，结构如图6所示。将通道注意力的输出结果通过最大池化和平均池化得到2个$1\times H\times W$的特征图，然后对2个特征图进行拼接。通过7×7卷积变为1通道的特征图，之后经过*sigmoid*函数得到空间注意力的特征图。最后将输出结果乘通道注意力模块的输出特征图变回$C\times H\times W$大小。其空间注意力表示为

$$\begin{cases} M_s(F')=sigmoid(f^{7\times7}\left[AvgPool(F')\right];MaxPool(F')]) \\ M_s(F')=sigmoid\left(f^{7\times7}\left[F^s_{avg};F^s_{max}\right]\right) \end{cases} \tag{10}$$

式中：$f^{7\times7}$为7×7的特征图；s为空间；M_s为基于空间的注意力；$M_s(F)$为特征图经过空间注意力机制；$[F^s_{avg};F^s_{max}]$为空间注意力中经过平均池化和最大池化的特征图。

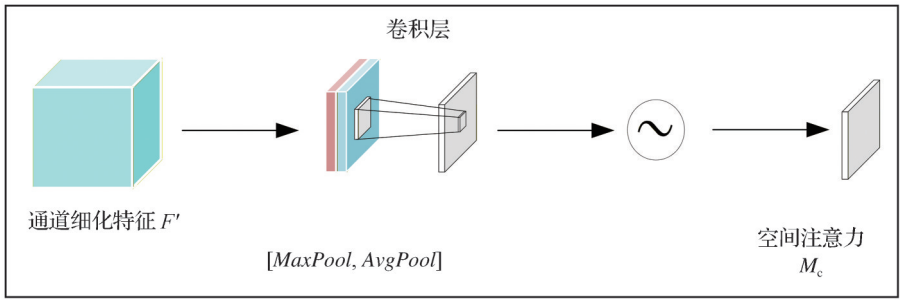


图6 SAM空间注意力机制模块结构

Fig. 6 Structure of SAM spatial attention mechanism module

CBAM通过通道和空间双注意力，精准解决精矿带与中矿带颜色相近的核心痛点。通道注意力能筛选关键区分特征，降低颜色通道权重，强化精矿带（细腻纹理）与中矿带（局部团聚纹理）的纹理、

边缘梯度通道权重,凸显二者微观差异;空间注意力可定位矿带区域,抑制摇床金属边框和冲洗水痕迹等背景干扰,减少矿带内部因灰度波动产生的分割空洞,让模型更聚焦有效矿带特征,避免因颜色混淆导致的分类错误。

可变形卷积(Deformable Convolution)是一种改进的传统卷积操作,其核心思想是通过引入可学习的空间采样偏移量使卷积核能够动态适应输入数据的几何形变,其结构如图7所示。可变形卷积可表示为

$$y\left(p^0\right)=\sum w\left(p_n\right)\cdot x\left(p^0+p_n+\Delta p_n\right) \tag{11}$$

式中:p^0为输出特征图上的当前位置;p_n为卷积核内相对中心的基准采样位置;Δp_n为可变形卷积学习到的额外偏移量。

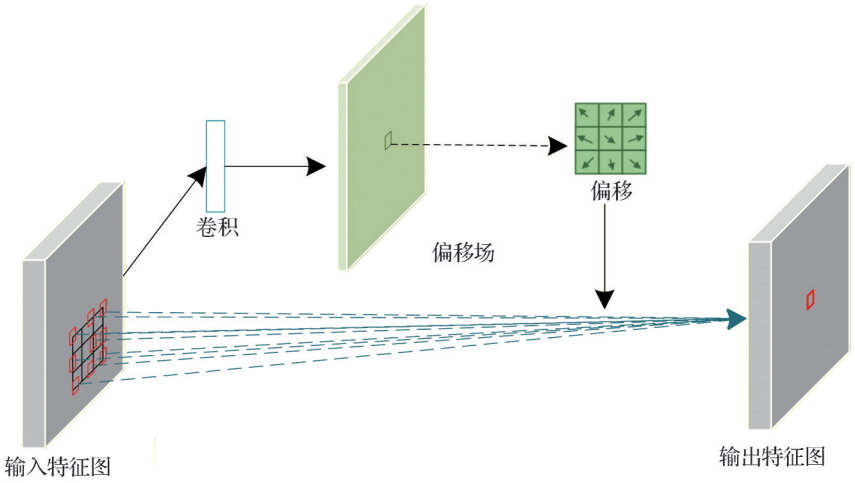


图7　可变形卷积结构

Fig. 7　Structure of deformable convolution network

可变形卷积的原理是通过额外的卷积层预测每个采样点的偏移量Δp_n,将传统卷积的固定采样位置p_n调整为可变形的$p_n'=p_0+p_n+\Delta p_n$,再通过双线性插值计算调整后的特征值。在可变形卷积中,首先由输入特征图通过独立卷积层生成$2N$维的偏移量场(N为采样点数量);然后利用这些偏移量将固定的采样网格调整为可变形的采样位置$p_n'=p^0+p_n+\Delta p_n$,由于偏移量通常为连续值,此时需通过双线性插值方法计算非整数坐标处的特征值$x\left(p_n'\right)$,确保梯度能够通过插值操作反向传播;最后对插值后的特征进行加权聚合得到输出。这种机制突破了传统卷积的刚性几何约束,使模型能够自动适应目标的形态变化和空间布局差异。后续提出的DCNv2进一步引入调制因子m_n,将公式扩展为

$$y\left(p^0\right)=\sum w\left(p_n\right)\cdot x\left(p^0+p_n+\Delta p_n\right)\cdot m_n \tag{12}$$

从而不仅能够调整采样位置,还能够动态调节各采样点的重要性权重。可变形卷积在保持与传统卷积相近计算效率的同时,显著提升了模型对几何形变的建模能力。可变形卷积通过动态偏移量调整采样位置,针对精矿带与中矿带颜色相近且边界不规则的问题,能够让卷积核主动向边界关键位置偏移,精准提取两侧差异特征,避免标准卷积的模糊过渡或锯齿状误判。同时,可自适应调整感受野,对小尺度精矿带用密集采样避免断裂,对大尺度尾矿带用宽范围采样减少内部空洞,有效适配精矿带、中矿带和尾矿带三类矿带尺度差异,提升整体分割完整性和边界精度。

3　试验与结果分析

3.1　矿带分割试验准备与数据集构建

在紫外线照射下,白钨矿会发出强烈荧光(刘衔宇等,2023),而绝大多数脉石矿物呈暗色,极大地增强了目标矿物与背景的对比度。这使得语义分割模型能够很容易学习到精矿带的视觉特征,从而实现对精矿带的精准提取和对中矿带的精细区分。

试验数据来自陕西省某矿业公司,数据均使用

海康威视摄像头在使用紫外光灯拍摄的白色钨矿带照片共542张，图片像素为512 pixel×512 pixel。紫外灯下的白钨矿带如图8所示。同时，使用Labelme工具对摇床矿带进行精矿带、中矿带和尾矿带的划分，并标注白钨矿带，标注结果如图9所示。

为了提升矿带分割模型的泛化能力，对初始数据进行5种方式的数据增强，包括水平和垂直方向的反转、镜像、改变HSV空间、加入高斯噪声、锐化和雾化，增加数据集至3 121张。通过数据扩增方式能够增加矿带图像的可变性，从而提高模型的泛化能力和稳健性，最后按照1.5∶1.5∶7划分了测试集、验证集和训练集。数据增强效果图如图10所示。

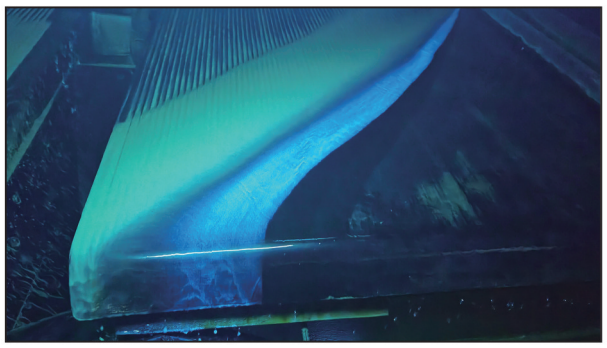


图8 紫外灯照射下的摇床矿带

Fig. 8 Ore zones on the shaking table under UV light

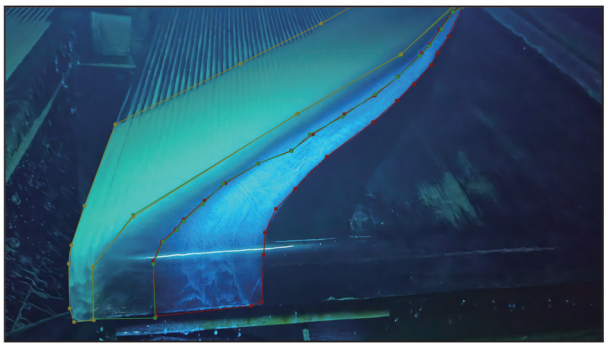


图9 使用Labelme对矿带进行标注

Fig. 9 Ore zones with an notations made using Labelme

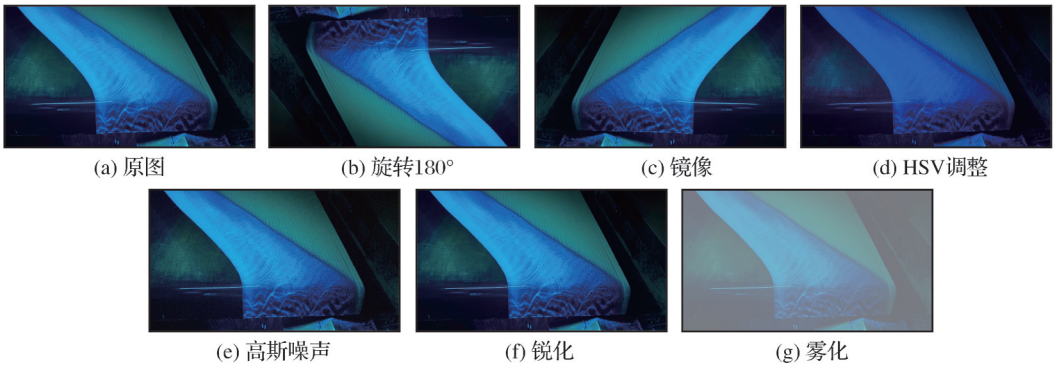


图10 矿带图像多种数据增强方式对比

Fig. 10 Comparison of multiple data enhancement methods for ore belt images

本试验的计算机配置为AMDRyzen74800HCPU，NVIDIARTX2060，操作系统为Windows10专业版，Python版本为3.11，CUDA版本为12.8，深度学习框架为pytorch2.2.0，使用Labelme版本为5.5.0。试验过程调用GPU进行训练，运行内存为16 GB，固态硬盘为500 GB，在所有实验训练参数设置中，输入图像大小为512 pixel×512 pixel，模型采用Adam优化器，优化器其他参数保持Pytorch默认值。损失函数采用多类别交叉熵损失（CrossEntropyLoss），对每个像素的预测类别与真实标签进行逐像素分类监督。训练批次大小设置为4，模型总共训练200个epoch。学习率采用多项式衰减策略，初始学习率为0.0012，最终学习率为1e-6，衰减因子power=0.9，学习率随当前epocht（从0开始）的变化公式为

$$lr_t=(lr_{\text{init}}-lr_{\text{end}})\times(1-\frac{t}{T})^{\text{power}}+lr_{\text{end}} \tag{13}$$

式中：lr_t为第t个epoch时的学习率；lr_{init}为初始学习率（本文取0.0012）；lr_{end}为最终学习率（本文取1×10^{-6}）；t为当前epoch编号（从0开始）；T为总训练epoch数（本文取200）；power为衰减因子（本文取0.9）。为防止过拟合，当验证集上的*MIOU*指标在连续30个epoch内未出现提升时，提前终止训练，并恢复至验证集最佳*MIOU*对应的模型权重；所有

可训练层均采用Xavier均匀初始化。

3.2 矿带分割模型评价指标

本文使用Dice、MIOU和GFLOPs作为评判模型的主要手段。

*Dice*系数是用于衡量2个样本相似度的统计指标,常用于图像分割任务。预测标签与实际标签之间的重叠程度越高,*Dice*系数越接近1,表示分割结果越准确。相反,如果重叠程度较低,则*Dice*系数将接近0。计算公式如下:

$$Dice_{c}officient=\frac{2|X\cap Y|}{|X|+|Y|} \tag{14}$$

式中:X和Y分别为实际前景和预测前景,而$|X\cap Y|$表示X中与Y相交的部分。MIOU(平均交并比)是语义分割任务中常用的评估指标,用于衡量预测分割结果与真实分割标签之间的平均重叠程度。计算公式如下:

$$MIOU=\frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{TP+FN+FP} \tag{15}$$

式中:*TP*表示样本的实际类为正,预测为正;*TN*表示样本的正确类别为负值,且模型预测负类别;*FP*表示其准确类别为负值,但是模型识别为正的样本;*FN*表示其准确类别为正的样本,但模型将其标识为负的样本。*IOU*是分割准确性的度量,用于计算预测的分割结果与实际标签之间的交集和串联比率;*MIOU*是通过对多个样本或类别的*IOU*求平均值而获得的值。*IOU*和*MIOU*的值越高,表示分割结果与实际标签之间的匹配性越好。

GFLOPs意为10亿次1 s的浮点运算数,代表模型运行时需要的运算量,其数值越大说明模型结构所需的运算数量越大,代表其所需参数量更多且模型结构越复杂。

3.3 矿带分割效果消融试验分析

为了深入探究各改进模块在钨矿摇床矿带分割任务中的具体贡献,并明确其对模型性能的优化效果,通过系统的消融试验对模型各组件进行深入分析。将各阶段改进后的模型分别命名为Method1、Method2、Method3、Method4和Ours。

如表1和图11所示,基准模型采用DeepLabV3+,在自制数据集上取得了85.71%的*Dice*系数、85.34%的*MIOU*以及77.40%的mAP@50:95,其模型大小为31.61 M,计算复杂度为53.47 GFLOPs,每秒帧数为20.19,该性能为后续改进提供了坚实基础。

表1 模型消融试验对比

Table 1 Comparison of ablation test of model

模型	MobileNetV3	ECA	DenseASPP和空洞卷积	CBAM	可变形卷积	*Dice*	mAP@50:95/%	*MIOU*/%	模型大小/M	GFLOPS/G	*FPS*/s
DeeplabV3+						85.71	85.34	77.40	31.6	53.47	20.14
Method1	√					82.02	79.15	71.38	13.24	18.72	16.88
Method2	√	√				83.74	81.53	74.16	15.49	24.78	22.36
Method3	√	√	√			87.59	87.23	76.36	15.83	26.90	23.70
Method4	√	√	√	√		92.46	88.96	79.84	17.21	32.13	26.46
Ours	√	√	√	√	√	93.89	91.97	81.27	18.32	33.50	27.14

本文以DeepLabV3+作为基线模型,通过逐步引入轻量化骨干网络和多种改进模块,开展了消融试验,各模型的性能对比结果如表1所示。

将DeepLabV3+模型中原有的Xception骨干网络替换为MobileNetV3轻量化骨干网络。Method1模型参数量由基线模型的31.6 M显著降低至13.24 M,GFLOPs由53.47 G降低至18.72 G,推理速度由20.19 FPS提升至16.88 FPS。但作为轻量化的代价,模型分割性能出现预期内下降:*Dice*系数由基线的85.71%降至82.02%,*MIOU*由85.34%降至79.15%,mAP@50:95由77.40%降至71.38%。上述结果表明,骨干网络替换是降低模型参数量、计算量并提升推理速度的关键手段,但也会导致原有模型特征提取能力与分割精度的下降。

Method2在Method1的轻量化骨干网络基础上,引入ECA注意力机制。模型参数量小幅提升至15.49 M,GFLOPs增至24.78 G,推理速度提升至22.36 FPS。同时分割性能实现全面提升:*Dice*系数回升至83.74%,*MIOU*提升至81.53%,mAP@50:95提升至74.16%,全面超越Method1。该结果证明,

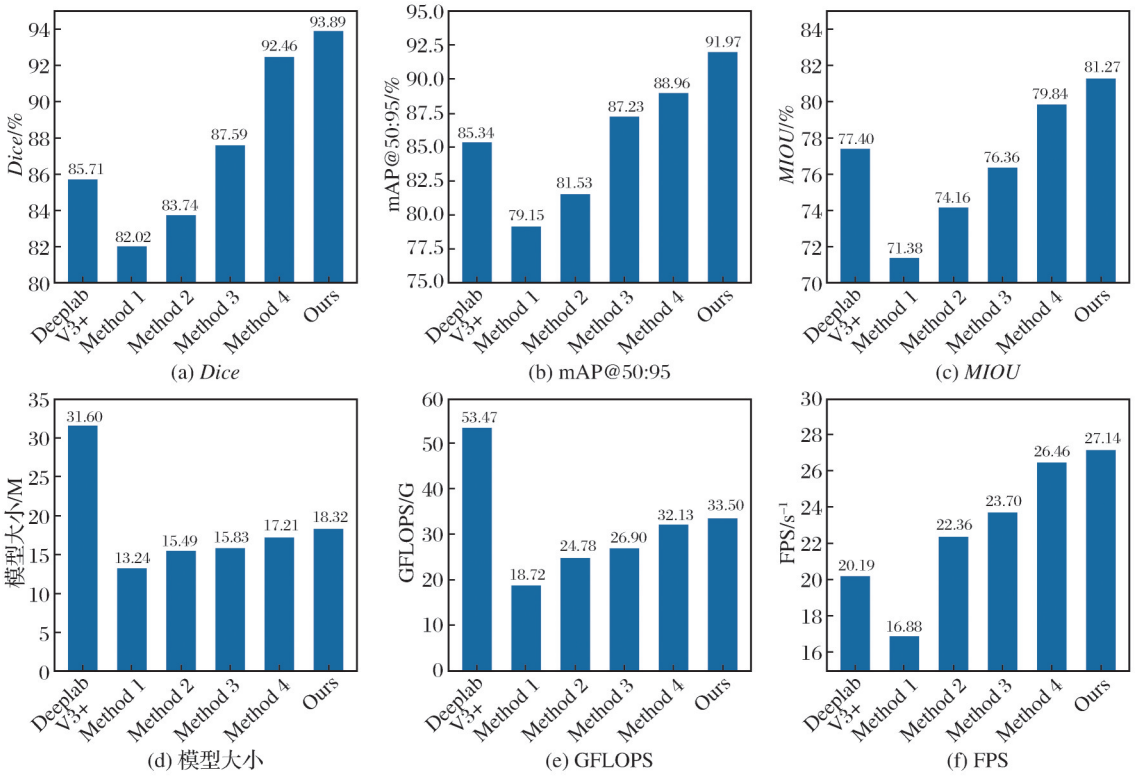


图11　消融试验对比直方图

Fig. 11　Histogram of ablation test constrast

ECA注意力机制能够有效增强模型对矿带关键特征的聚焦能力，在几乎不损失推理速度的前提下，弥补轻量化骨干网络带来的精度损失。

Method3在Method2的基础上，引入DenseASPP（密集空洞卷积）模块。模型参数量小幅提升至15.83 M，GFLOPs增加至26.90 G，推理速度提升至23.70 FPS。分割性能得到进一步提升：*Dice*系数提升至87.59%，*MIOU*提升至87.23%，mAP@50:95提升至76.36%，性能已接近且部分指标超越基线模型。该结果表明，密集空洞卷积所构建的多尺度特征金字塔，能够有效捕捉摇床矿带复杂多变的形态结构与边缘细节特征。

Method4在Method3的基础上，进一步引入CBAM注意力机制。模型参数量提升至17.21 M，GFLOPs增加至32.13 G，推理速度提升至26.46 FPS。分割性能得到显著跃升：*Dice*系数达到92.46%，*MIOU*达到88.96%，mAP@50:95提升至79.84%，已全面超越基线DeepLabV3+模型。该结果表明，结合通道与空间注意力的CBAM模块，能够进一步强化模型对矿带目标区域的特征响应，抑制背景噪声，实现精度与效率的平衡。

图12所示为DeepLabV3+原始结构和改进后的Method1至Ours网络在矿带检测中的Grad-CAM对比。从图12中可以看出，原始网络结构对矿带中心区域和边缘区域的关注度相对较低。相比之下，改进后的Method1、Method2、Method3、Method4和Ours对精矿带、中矿带和尾矿带的敏感程度整体高于原模型。结合改进模型的*mAP*、*MIOU*、*Dice*系数与注意力热力图分析，得出Ours模型更加适合矿带分割任务。

最终，集成全部改进模块（包括ECA、DenseASPP、CBAM和可变形卷积）的Ours模型实现了最优性能：*Dice*系数达到93.89%、*MIOU*达到91.97%、mAP@50:95达到81.27%，模型大小与计算成本分别为18.32 M和33.50 GFLOPs，FPS达到27.14。整套改进方案在显著提升分割精度的同时，保持了模型的轻量化特性，验证了各模块在复杂矿带选矿的有效性和互补性。

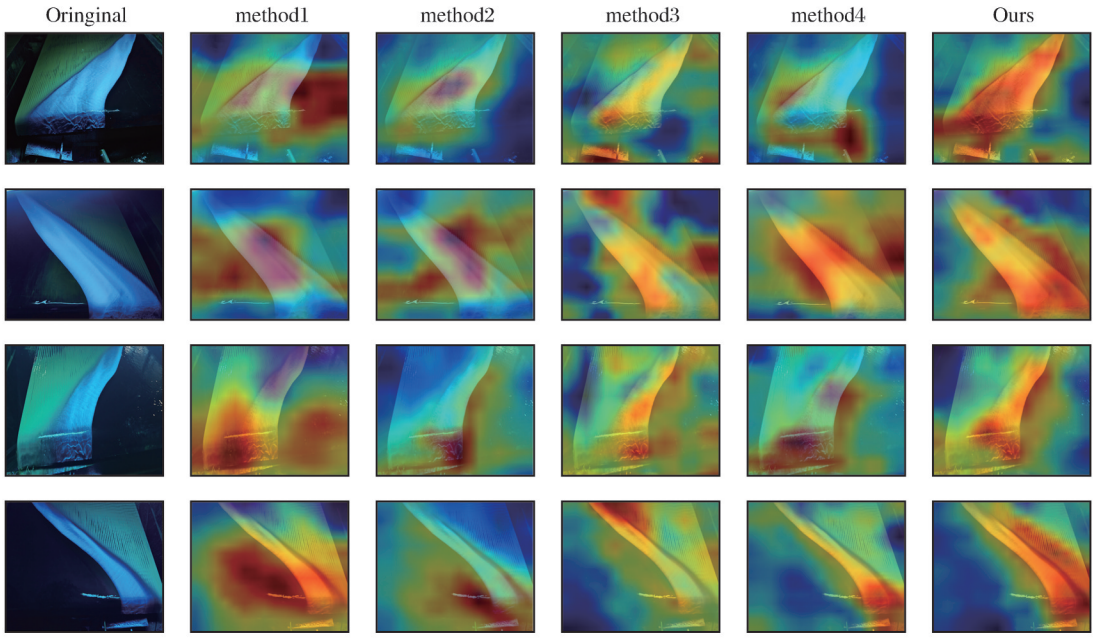


图 12　消融试验热力图

Fig. 12　Thermal diagram of ablation test

3.4　矿带分割对比试验分析

为了验证 Ours 模型在紫外灯下摇床矿带分割任务中的性能表现,开展了系统的对比试验。分别使用 U-NET(Ronneberger et al, 2015)、U-NET++(Zhou et al,2018)、HRNet(Wang et al,2021)、PSPNet(Zhao et al,2017)、DeepLabV3+和 Ours 共 6 种网络模型进行训练和评估。在试验过程中,所有模型均在相同的硬件环境下运行,且采用相同的训练数据集、迭代次数和超参数设置,以确保试验结果的公平性和可比性。图 13 所示为不同模型的分割效果图。其中,红色、绿色和黄色区域分别为精矿带、中矿带和尾矿带的语义分割结果。

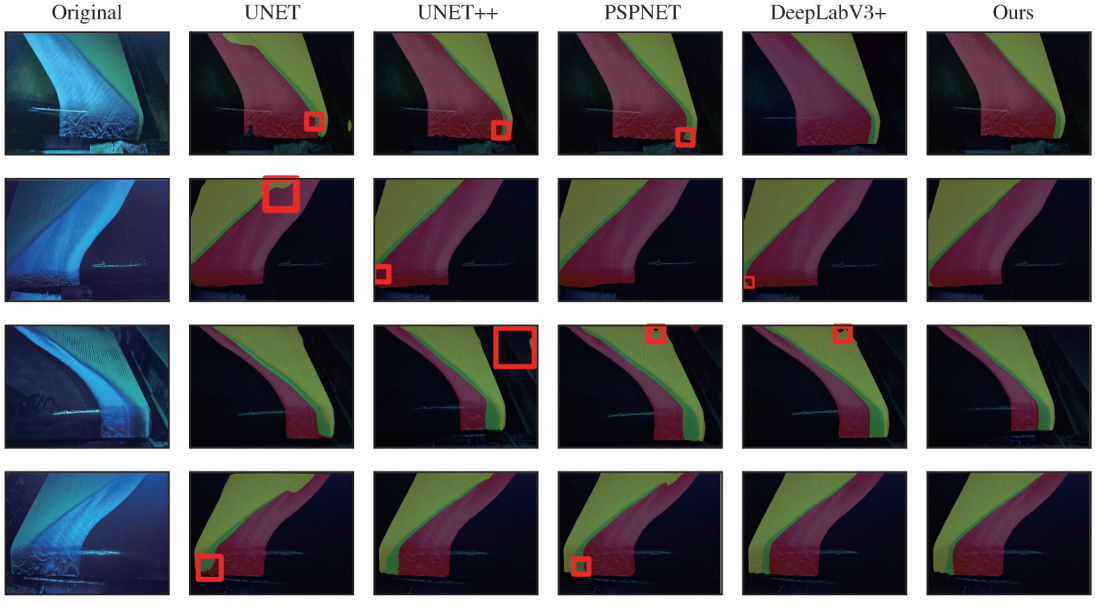


图 13　主流分割模型分割效果对比图

Fig.13　Performance comparison of mainstream segmentation models

图13所示为不同模型在典型摇床矿带工况下的分割结果对比。从整体上看，传统语义分割模型（UNET、UNET++、PSPNet和DeepLabV3+）在矿带边缘分割、噪声抑制和类别区分上均存在明显缺陷，而本文提出的模型能够精准还原矿带轮廓，无明显误分割与漏检现象，分割结果更贴合真实矿带分布。

具体来看，在矿带末端边缘细节分割中，UNET、UNET++、PSPNet和DeepLabV3+均出现边缘欠分割问题，矿带末端轮廓存在缺口，而本文模型能够平滑、连续地还原矿带末端边界；在矿带上下部区域，UNET出现了跨类别误分割，UNET++和DeepLabV3+则存在边缘过分割，而本文模型的矿带边界清晰、无噪声；在背景噪声抑制方面，UNET++、PSPNet和DeepLabV3+均出现假阳性误检，本文模型则无此类问题；在矿带起始端，UNET和PSPNet存在类别混淆导致的漏检，本文模型则完整保留了不同矿带类别的边界。

上述结果表明，本文模型在边缘细节刻画、背景噪声抑制和多类别边界区分上均具有显著优势，对复杂工况下的摇床矿带具有更强的适应性和稳健性。试验结果如表2和图14所示，Ours模型以33.50 GFLOPs的计算复杂度实现了93.89的*Dice*系

表2　主流分割网络模型性能对比

Table 2　Performance comparison of mainstream segmentation network models

模型	*Dice*/%	mAP@50:95/%	*MIOU*/%	模型大小/M	GFLOPS/G	FPS/s
U-NET	84.13	82.43	76.65	16.6	25.31	17.53
U-NET++	86.38	84.20	75.24	21.8	28.42	18.27
Hrnet	83.52	78.75	77.58	38.5	43.78	13.61
PSPnet	82.32	79.61	68.07	26.7	58.45	11.89
DeepLabV3+	85.71	85.34	77.40	31.6	53.47	20.14
Ours	93.89	91.97	81.04	18.3	33.50	27.14

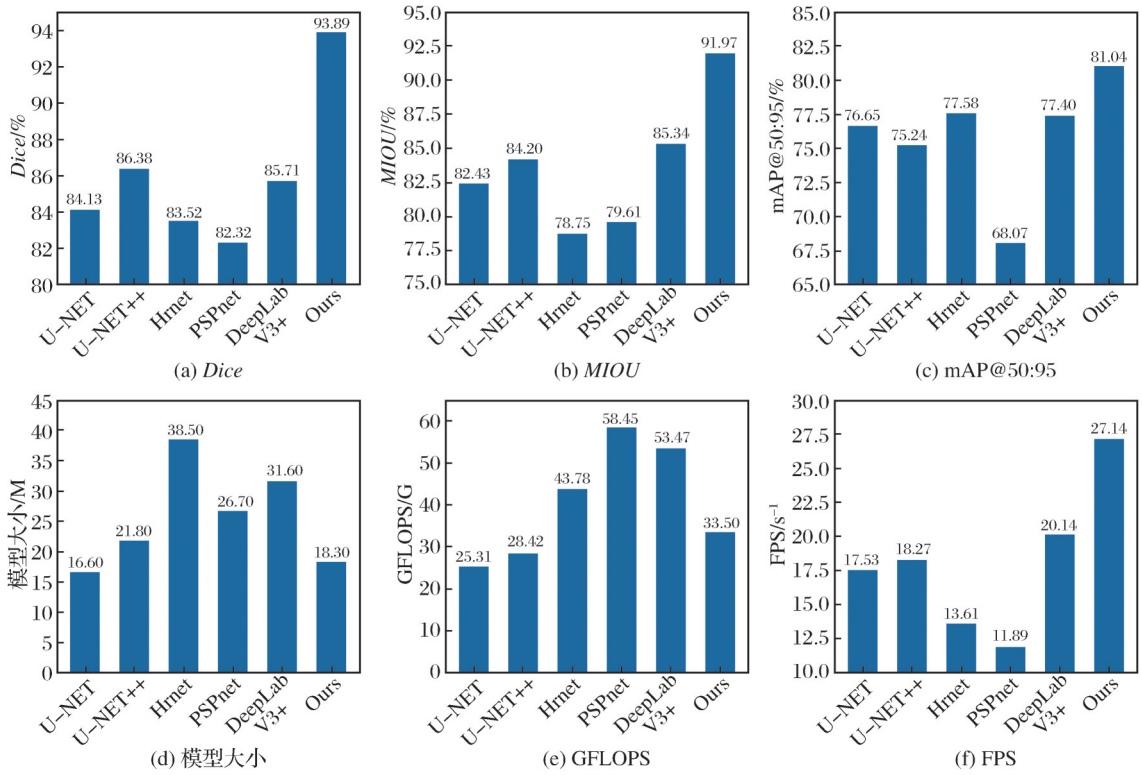


图14　主流网络分割参数对比

Fig. 14　Comparison of mainstream network segmentation parameters

数、91.97的*MIOU*、81.04的mAP@50:95和27.14的FPS。尽管在模型大小与计算量上略高于U-NET和U-NET++,但其各项指标均显著优于这2种模型,展现出更优的分割精度与实时性。

相比之下,Hrnet模型虽以43.78的GFLOPS实现了83.52的*Dice*系数和77.58的mAP@50:95,但在各项指标上均逊于Ours模型。PSPNet和DeepLabV3+分别以58.45和53.47的GFLOPS实现了85.32和86.71的*Dice*系数,但其mAP@50:95指标仍低于Ours模型,且模型大小分别达到26.7 M和31.6 M,远大于Ours模型的18.3 M。

相比基础U-NET,U-NET++的*Dice*系数和*MIOU*指标有所提升,分别达到89.38和84.20,但mAP@50:95(75.24)仍明显低于Ours模型,且模型大小(21.8 M)和计算复杂度(28.42)并未体现出轻量化优势。

Ours模型在特征提取与多尺度融合方面展现出显著优势,其mAP@50:95指标领先所有对比模型,且模型大小更具优势,表明Ours模型通过高效的网络结构设计,能够更精准地捕捉图像细节特征,从而在保持较低模型大小的同时实现更高精度的分割效果。

4 结论

(1)针对现有摇床矿带识别模型不够精确的问题,提出了基于紫外线的矿带分割模型。该模型在原有网络基础上,以带有ECA模块的轻量型MobileNetv3网络替换原始的Xception作为骨干网络,降低了模型的复杂度;在ASPP中加入密集链接和2条空洞卷积分支,以提升矿带的多尺度融合能力,在解码器中加入可变形卷积和CBAM,满足了摇床矿带在复杂光源下对矿带的精准识别要求,为实现钨矿摇床选矿的智能化提供技术支撑。

(2)试验验证表明,改进后的模型在自制数据集上取得了卓越的识别性能,*MIOU*达到91.97%,*Dice*系数达到93.89%,mAP@0.5:95达到81.04%,FPS达到21.17/s。

(3)本文优化方法虽然能够有效解决摇床精矿带、中矿带和尾矿带的分割问题,但仍存在自制数据集规模受限和复杂工况下分割能力不足等问题,需进一步扩充不同工况的数据集,并优化改进后的DeepLabV3+模型,以适应硬件资源受限的环境,并投入工业化场景进行精矿回收率验证,进一步展开后续研究。

参考文献(References):

Chen L C, Zhu Y K, Papandreou G, et al, 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Computer Vision-ECCV 2018. Cham: Springer, 833-851.

Chollet F, 2017. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 1800-1807.

Howard A, Sandler M, Chen B, et al, 2019. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 1314-1324.

Hu J, Shen L, Sun G, 2018. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 7132-7141.

Kroll-Rabotin J S, Bourgeois F, Climent É, 2010. Fluid dynamics based modelling of the Falcon concentrator for ultrafine particle beneficiation [J]. Minerals Engineering, 23(4): 313-320.

Ronneberger O, Fischer P, Brox T, 2015. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer, 234-241.

Wang J D, Sun K, Cheng T H, et al, 2021. Deep high-resolution representation learning for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(10): 3349-3364.

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, 11531-11539.

Woo S, Park J, Lee J Y, et al, 2018. CBAM: convolutional block attention module[M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 3-19.

Xia Z F, Pan X R, Song S J, et al, 2022. Vision transformer with deformable attention[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 4784-4793.

Yu F, Koltun V, 2016. Multi-scale context aggregation by dilated convolutions[C]//Bengio Yoshua, LeCun Yann. 4th International Conference on Learning Representations (ICLR) 2016. JuanSan, Puerto Rico: International Conference on Learning Representations.

Zhao H S, Shi J P, Qi X J, et al, 2017. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 6230-6239.

Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al, 2018. UNet++: a nested U-Net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 3-11.

和丽芳, 郭思哲, 黄宋魏, 等, 2018. 一种基于磷虾优化算法的摇床矿带分带图像分割法: CN105844628B[P]. 2018-06-15. He Lifang, Guo Sizhe, Huang Songwei, et al, 2018. An image segmentation method based on krill optimization algorithm for zoning of rock bed ore zone: CN105844628B[P]. 2018 06 15.

刘惠中, 刘建业, 黄翱, 等, 2025. 摇床"控制参数—精矿边界坐标"关系模型构建研究[J]. 黄金科学技术, 33(3): 648-658. Liu Huizhong, Liu Jianye, Huang Ao, et al, 2025. Research on the construction of the relationship model of "control parameter-concentrate boundary coordinates" in shaking table[J]. Gold Science and Technology, 33(3): 648-658.

刘惠中, 游科顺, 2023. 基于深度学习边缘检测和Hough变换的摇床矿物带边界提取和特征表示[J]. 制造业自动化, 45(6): 180-183, 188. Liu Huizhong, You Keshun, 2023. Boundary extraction and feature representation of rocking mineral belt through deep learning edge detection and Hough transform[J]. Manufacturing Automation, 45(6): 180-183, 188.

刘衔宇, 杨九昌, 涂彩, 等, 2023. 中国四川平武雪宝顶白钨矿的光谱学表征[J]. 光谱学与光谱分析, 43(8): 2550-2556. Liu Xianyu, Yang Jiuchang, Tu Cai, et al, 2023. Study on spectral characteristics of scheelite from Xuebaoding, Pingwu County, Sichuan Province, China[J]. Spectroscopy and Spectral Analysis, 43(8): 2550-2556.

田妞, 黄宋魏, 和丽芳, 等, 2023. 深度学习在选矿领域的应用进展[J]. 化工矿物与加工, 52(8): 69-74, 82. Tian Niu, Huang Songwei, He Lifang, et al, 2023. Application of deep learning in mineral processing field[J]. Industrial Minerals & Processing, 52(8): 69-74, 82.

王锦煜, 2022. 摇床矿带识别算法及其应用[D]. 湘潭: 湖南科技大学. Wang Jinyu, 2022. Algorithm and application of shaking ore belt identification[D]. Xiangtan: Hunan University of Science and Technology.

杨文龙, 吴富姬, 罗璇, 等, 2016. 一种选矿摇床矿带识别及接矿板自动调节装置: CN105903557A[P]. 2016-08-31. Yang Wenlong, Wu Fuji, Luo Xuan, et al, 2016. An ore belt identification and automatic adjustment device for ore receiving plate of mineral processing shaking table: CN105903557A[P]. 2016 08 31.

杨文旺, 何庆浪, 兰希雄, 等, 2020. 选矿摇床智能巡检机器人开发与应用[J]. 有色金属(选矿部分), (5): 102-106. Yang Wenwang, He Qinglang, Lan Xixiong, et al, 2020. Development and application of intelligent inspection robot for table concentrator[J]. Nonferrous Metals Mieral Processing Section, (5): 102-106.

赵玉华, 杨文旺, 武涛, 等, 2023. 基于深度学习的摇床接矿位置预测方法[J]. 有色金属(选矿部分), (2): 142-146, 172. Zhao Yuhua, Yang Wenwang, Wu Tao, et al, 2023. Prediction method of shaking table contacting position based on deep learning[J]. Nonferrous Metals (Mineral Processing Section), (2): 142-146, 172.

郑玉荣, 靳军宝, 吴新年, 等, 2025. 智能矿山关键技术研发与应用进展[J]. 黄金科学技术, 33(2): 404-416. Zheng Yurong, Jin Junbao, Wu Xinnian, et al, 2025. Research and application progress of key technologies in smart mine[J]. Gold Science and Technology, 33(2): 404-416.

Intelligent Segmentation Model of Scheelite Belt in Gravity Separation Shaking Table Based on Machine Vision

RUAN Shunling[1,2,3], CHEN Sirui[1,3], NAN Xin[1,3], MENG Lingxuan[1,3]

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

2. *Xi'an Key Laboratory of Perceptual Computing and Decision Making for Smart Industry, Xi'an* 710055, *Shaanxi, China*;

3. *Institute of Mine System Egineering, Xi'an University of Architecture and Technology, Xi'an* 710055, *Shaanxi, China*

Abstract: At present, there are some problems in the process of gravity separation shaking table beneficiation, such as low efficiency and high cost of manual identification of ore belts, as well as large number of parameters and insufficient accuracy of existing segmentation models. This study addresses these challenges by developing a model that is both computationally efficient and highly accurate for the segmentation of scheelite zones under ultraviolet light, a condition that enhances the natural fluorescence of the target mineral. Central to this research is a significantly enhanced DeepLabV3+ model, which integrates several critical architectural modifications. Initially, the computationally intensive Xception backbone was substituted with the more efficient MobileNetV3 network, which was further enhanced by integrating an efficient channel attention module. This modification aimed to reduce the parameter count while enhancing feature representation. Subsequently, the encoder's atrous spatial pyramid pooling module was restructured into a DenseASPP configuration. This redesign incorporated additional atrous convolutional branches to establish a more continuous range of receptive fields, thereby facilitating superior multi-scale feature extraction from mineral bands of varying widths. Furthermore, the decoder was augmented with a convolutional block attention module to highlight discriminative features and suppress noise, complemented by deformable convolutions to accommodate the irregular shapes and boundaries of the mineral zones. The model was trained and evaluated on a custom-developed dataset of ultraviolet-lit scheelite images, which was expanded using data augmentation techniques. A series of ablation studies confirmed the contribution of each proposed enhancement. In comparative experiments, the final model exhibited superior performance, achieving a Dice coefficient of 93.89%, a mean intersection over union of 91.97%, and an mAP@50:95 of 81.04%. The study demonstrates that the proposed model achieves a high level of accuracy with a notably reduced model size of 18.32 million parameters and a computational cost of 33.50 GFLOPs. This indicates an optimal balance between precision and efficiency, making it well-suited for industrial applications. The research conclusively shows that the model effectively fulfills the dual requirements of high accuracy and a lightweight architecture for segmenting shaking table mineral zones. By incorporating specific advancements in network architecture, the model offers a robust technical foundation for automating shaking table control systems. This development has the potential to significantly improve beneficiation efficiency and consistency, marking a substantial advancement in intelligent mineral processing. Future research will focus on further validating the model's generalization capabilities across a broader range of industrial operating scenarios.

Key words: mineral processing shaking table; machine learning; gravity separation; semantic segmentation; image processing; ore belt detection