

DOI:10.19650/j.cnki.cjsi.J2514109

# 基于改进 DeepLabv3+的煤矿输送带异物分割模型 \*

刘光伟<sup>1</sup>, 张浩博<sup>1</sup>, 范忠胜<sup>2</sup>, 付恩三<sup>1,3</sup>, 雷 健<sup>1</sup>

(1. 辽宁工程技术大学矿业学院 阜新 123000; 2. 山西忻州神达梁家碛煤业有限公司 忻州 034000;  
3. 应急管理部信息研究院 北京 100029)

**摘要:**为准确检测煤矿带式输送机在复杂工况下的异物,构建基于改进 DeepLabv3+的煤矿输送带异物分割模型。针对煤矿高粉尘、光照不均、机械振动等干扰导致的异物检测难题,以及多尺度异物并存、边缘设备算力有限等实际需求,通过引入 MobileNetv3 轻量化主干网络,利用深度可分离卷积将计算量压缩至传统卷积的 1/9,并嵌入 SE 注意力模块增强异物的边缘、纹理等高频特征,抑制粉尘噪声对应的低频通道;采用 DASPP 模块替代传统 ASPP,通过串联不同膨胀率的空洞卷积层实现跨层特征密集交互,提升对多尺度异物的检测能力;集成 ECANet 通道注意力机制,通过免降维全局池化和动态一维卷积增强特征表达能力,进一步优化特征权重分配。实验结果表明,改进模型在 CUMT-Belt 数据集上实现了 87.1% 的平均交并比和 86.7% 的 F1 分数,参数数量仅为 9.8 M,浮点运算量为 5.1 G,推理速度达 38.6 fps,较原始 DeepLabv3+模型精度提升 4.6%、计算量降低 63.1%。与 PSPNet、U-Net 等主流模型相比,改进模型在小尺度异物漏检率、噪声鲁棒性及边缘设备适配性等关键指标上更优。该模型为解决复杂工况下异物与背景特征易混淆的难题提供了新途径,不仅为煤矿智能运输系统打造了兼具高分割精度与轻量化特性的异物检测方案,还有助于加速煤炭行业智能化与自动化的发展进程。

**关键词:** 煤矿输送带;机器视觉;图像处理;异物分割;深度学习

**中图分类号:** TH228      **文献标识码:** A      **国家标准学科分类代码:** 520. 2040

## Foreign object segmentation model for coal mine conveyor belts based on improved DeepLabv3+

Liu Guangwei<sup>1</sup>, Zhang Haobo<sup>1</sup>, Fan Zhongsheng<sup>2</sup>, Fu Ensan<sup>1,3</sup>, Lei Jian<sup>1</sup>

(1. School of Mining, Liaoning Technical University, Fuxin 123000, China; 2. Shanxi Xinzhou Shenda  
Liangjiaqi Coal Industry Co., Ltd., Xinzhou 034000, China; 3. Information Institute, Ministry of Emergency  
Management of the PRC, Beijing 100029, China)

**Abstract:** To accurately detect foreign objects on coal mine belt conveyors under complex working conditions, a coal mine conveyor belt foreign object segmentation model based on improved DeepLabv3+ was constructed. Aiming at the difficulties in foreign object detection caused by interference factors such as high dust, uneven illumination, and mechanical vibration in coal mines, as well as practical requirements including the coexistence of multi-scale foreign objects and limited computing power of edge devices, the following improvements were made: The MobileNetv3 lightweight backbone network was introduced, and depthwise separable convolution was used to compress the computational load to 1/9 of that of traditional convolution. Meanwhile, the SE attention module was embedded to enhance high-frequency features such as edges and textures of foreign objects and suppress low-frequency channels corresponding to dust noise. The DASPP module was adopted to replace the traditional ASPP, and cross-layer feature dense interaction was realized by concatenating atrous convolution layers with different dilation rates, thereby improving the detection for multi-scale foreign objects. The ECANet channel attention mechanism was integrated, which enhanced feature expression ability through dimension-reduction-free global pooling and dynamic 1D convolution, further optimizing the distribution of feature weights. Experimental results show that the improved

model achieved an average mean intersection over union of 87.1% and an F1-score of 86.7% on the CUMT-BelT dataset, with only 9.8 M parameters, 5.1 GFLOPs, and an inference speed of 38.6 fps. Compared with the original DeepLabv3+ model, the improved model increased accuracy by 4.6% and reduced computational load by 63.1%. In comparison with mainstream models such as PSPNet and U-Net, the improved model exhibits superior performance in key indicators including the missed detection rate of small-scale foreign objects, noise robustness, and adaptability to edge devices. This model provides a new approach to solving the problem of easy confusion between foreign objects and background features under complex working conditions. By simultaneously achieving high segmentation accuracy and low computational complexity, the proposed method supports real-time monitoring and contributes to advancing automation and intelligent development in the coal industry.

**Keywords:** coal mine conveyor belt; computer vision; image processing; foreign object segmentation; deeplearning

## 0 引言

作为全球最大的煤炭生产与消费国,我国煤炭产业在能源安全保障体系中占据基础性地位<sup>[1]</sup>。2023年我国规模以上工业原煤产量达46.6亿吨,同比增长2.9%,在能源消费结构中占比55.3%。在新能源发展阶段性受限的背景下,煤炭产业仍肩负着保障经济运行和民生用能的重要使命。带式输送机是煤矿运输系统的核心装备,通过摩擦驱动实现物料连续传输,具备输送容量大、适应复杂工况、长距离传输效率高等优势,是矿山物料运输的关键基础设施<sup>[2]</sup>。然而,高产高效开采需求下,设备长期满负荷运行,导致输送带撕裂、滚筒故障、异物卡阻等典型故障频发。因此,构建精准高效的异物检测与分割模型,是保障煤矿运输系统安全稳定的关键<sup>[3]</sup>。

在复杂恶劣的环境中,带式输送机异物识别面临多形态异物检测、环境干扰鲁棒性等挑战。众多学者长期致力于通过传统图像处理技术攻克这一难题。Siami等<sup>[4]</sup>提出基于热图像增强和卷积神经网络的皮带输送机托辊热缺陷语义分割方法,以少量手工标注热图像为基础,对比不同架构的性能差异。程健等<sup>[5]</sup>提出基于粒子群优化算法改进高斯混合模型的煤矸石视频检测方法,通过参数自整定实现复杂场景下煤矸石检测。Carvalho等<sup>[6]</sup>提出基于无人机集成热成像相机的带式输送机滚筒检测方法,利用无人机机动性与热成像技术,通过信号处理算法自动识别滚筒故障。张宁波等<sup>[7]</sup>提出基于自然射线的识别方法,分析矸石低水平自然射线的涨落规律,确定放射性识别阈值及精确判断阈值,并通过静态与动态实验验证煤矸辐射特性和探测器实时响应能力。Shiri等<sup>[8]</sup>提出基于声学信号的带式输送机托辊故障检测方法,通过局部均值分解降噪后,计算光谱自相关图并进行空间去噪,实现托辊轴承故障检测。尽管上述传统图像处理方法在算法处理速度、信息压缩效率等方面具有一定优势,但在煤矿复杂工况下仍暴露出局限性。一方面,对锚杆、矸石等多形态、多尺度异物的检测精度不足,尤

其难以处理密集小尺度异物的漏检问题;另一方面,算法对粉尘、光照不均、机械振动等环境干扰的鲁棒性较弱,常因背景噪声导致特征混淆,最终使得带式输送机异物检测的整体精度难以满足工业应用需求<sup>[9]</sup>。

随着深度学习技术发展,其自学习多层次特征的能力为解决上述问题提供新路径, YOLO<sup>[10]</sup>(you only look once)、U-Net<sup>[11]</sup>(convolutional networks for biomedical image segmentation)、DeepLab<sup>[12]</sup>(deep convolutional lab)系列、Transformer<sup>[13]</sup>等架构被引入输送带异物检测,通过自学习多层次特征提升复杂场景检测精度。程德强等<sup>[14]</sup>提出融合残差信息的轻量级网络,采用残差块为基本单元,通过交叉学习机制与特征拼接融合多尺度特征,提升运煤皮带异物分类的精度与计算速度。李永安等<sup>[15]</sup>提出多任务检测方法,通过扩张式残差模块、高效层聚合网络及损失函数优化,提升检测精度与效率。吕瑞宏等<sup>[16]</sup>提出基于改进(residual network, Res-Unet)的煤矸元素成分分析算法,在编码器中集成高效局部注意力模块,改进跳跃连接方式融合多尺度信息,实现CT图像元素区域精确分割及占比分析。马宏伟等<sup>[17]</sup>提出改进定向FAST和旋转BRIEF快速近似最近邻搜索库(oriented fast and rotated brief fast library for approximate nearest neighbors, ORB-FLANN)匹配方法,采用增强的高效局部图像特征描述符(boosted efficient binary local image descriptor, BEBLID)描述符改进ORB特征检测,融合渐近采样一致性(progressive sampling consensus, PROSAC)算法优化FLANN匹配,剔除错误点对,提升匹配精确率。Zhao等<sup>[18]</sup>提出基于RGB(red, green, blue)与红外图像融合的煤矸识别方法,利用多模态特征互补优势,结合ResNet网络的深层特征提取能力,提升煤矸识别的准确性与稳定性。相较于传统方法,深度学习模型在多目标检测、实时性和鲁棒性方面展现出优势,但在复杂环境下仍面临的挑战有:1)边缘设备算力有限,传统重型网络参数量大、推理延迟高,难以满足实时预警需求;2)输送带上多尺度异物并存,传统并行分支的特征融合方式缺乏跨层交互,导致小尺度异物特征易被噪声淹没;3)粉尘噪声与异物特征混叠问题突出,传统通道注意

力机制因降维操作可能丢失高频细节,影响复杂背景下的特征判别力。

针对上述问题,提出基于改进 DeepLabv3+(deep convolutional lab v3+)的煤矿输送带异物分割模型。通过引入轻量化主干网络(mobile network v3, MobileNetv3),利用深度可分离卷积和通道注意力(squeeze-and-excitation, SE)压缩计算量。采用密集空洞空间金字塔池化模块(dense atrous spatial pyramid pooling, DASPP)替代传统空洞空间金字塔池化模块(atrous spatial pyramid pooling, ASPP),通过串联不同膨胀率的空洞卷积层实现跨层特征的密集交互,提升多尺度异物的感知能力。集成通道注意力机制高效通道注意力网络(efficient channel attention network, ECANet),通过免降维全局池化和动态一维卷积,增强特征表达能力,为煤矿带式输送机异物的精准识别提供解决方案。

## 1 煤矿带式输送机异物分割模型

### 1.1 煤矿带式输送机系统

煤矿带式输送机系统是矿井生产运输的核心环节,其运行状态直接影响煤炭开采的安全性与效率<sup>[19]</sup>。如图1所示,该系统主要由9部分关键部件构成,图1中数字1代表进料槽、2代表上托辊、3代表下托辊、4代表机架、5代表输送带、6代表改向滚筒、7代表传动滚筒、8代表张紧装置、9代表驱动装置。

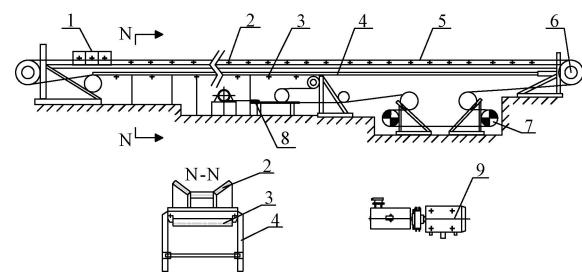


图1 煤矿带式输送机的结构组成

Fig. 1 Structural composition of coal mine belt conveyors

这些部件的协同工作实现了煤炭从开采面到仓储区的连续传输,其运行效率直接影响矿井的生产节奏。然而,煤矿环境高粉尘、潮湿、光照不足且机械振动剧烈,导致带式输送机运行面临噪声干扰严重、多尺度异物并存、实时性要求高等挑战,而现有依赖接触式传感器与人工巡检的传统监测技术存在检测维度单一、滞后性显著、人工成本高及多任务耦合不足等问题。随着煤矿智能化建设推进,带式输送机监测正向“视觉化、集成化、实时化”转型,需具备轻量化与鲁棒性、增强抗干扰能力。根据以上背景,提出一种基于机器视觉的矿用带式输送机异物分割模型,通过相关优化提升系统在复杂工况下对异物的检测性能,为煤矿智能运输系统提供关键技术支撑。

### 1.2 模型改进策略

为解决煤矿带式输送机复杂工况下异物分割面临的噪声干扰强、实时性要求高、多尺度特征捕捉难等问题,对原始DeepLabv3+模型进行改进,改进的网络结构如图2所示。

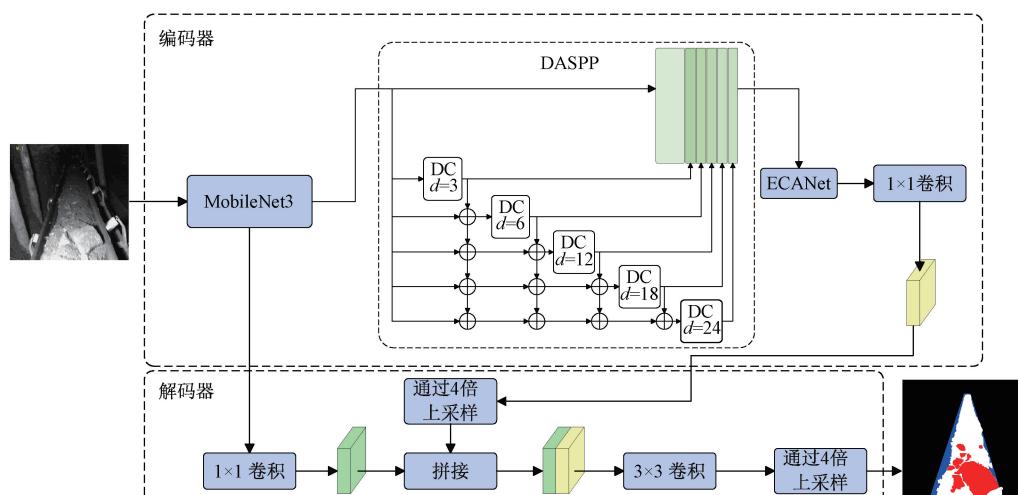


图2 改进 DeepLabv3+模型的网络结构

Fig. 2 Improved network architecture of deeplabv3+ model

图2中各模块说明为:

1)引入MobileNetv3作为轻量化主干网络,通过深度可分离卷积压缩计算量,利用线性瓶颈残差结构平衡性

能与效率,并嵌入SE通道注意力模块,增强异物边缘、纹理等高频特征通道,抑制粉尘噪声对应的低频通道,提升复杂背景下的特征判别力。

2)采用密集空洞空间金字塔池化模块 DASPP 替换原有的模块 ASPP, 将各空洞卷积构成的空洞层进行串联, 使每层空洞卷积后的输出馈送至后续每一个空洞层以实现信息共享, 同时各层空洞卷积的膨胀率逐层递增, 扩大感受野并融合多尺度特征, 提升网络对锚杆、矸石等多样形态异物的感知能力。

3)引入通道注意力机制 ECANet, 通过免降维全局池化保持通道维度完整, 避免降维导致的信息丢失, 采用核大小可自适应调整的一维卷积替代全连接层, 在几乎不增加计算成本的前提下捕获跨通道依赖, 提升特征表达能力, 以适配煤矿边缘设备的算力限制。

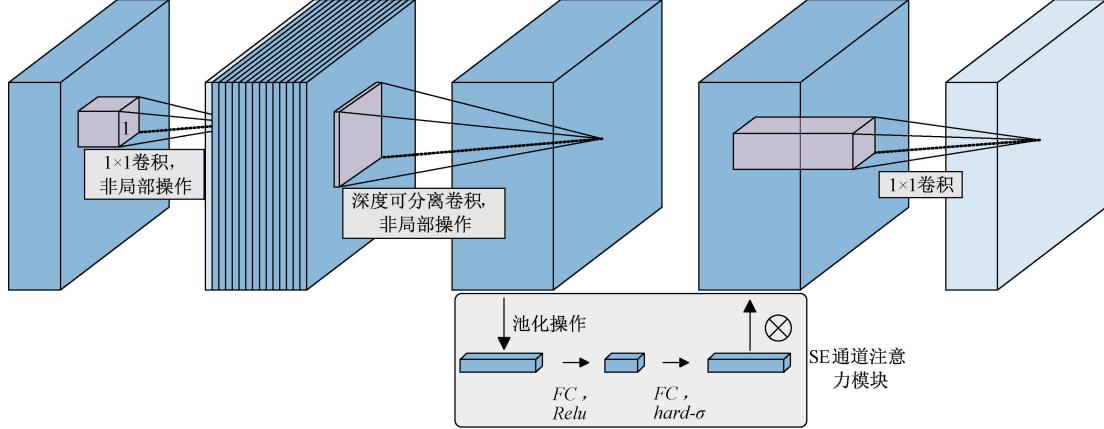


图 3 Mobilenetv3 模块

Fig. 3 Mobilenetv3 module

MobileNetv3 将标准卷积拆分为深度卷积与逐点卷积, 计算量压缩至标准卷积的  $1/K^2$ ,  $K$  为卷积核尺寸, 如  $3 \times 3$  卷积计算量减少至  $1/9$ , 大幅降低浮点运算次数。同时沿用 MobileNetv2 (mobile network v2) 的线性瓶颈残差结构, 通过扩张层、卷积层、投影层的设计, 在低维空间输入输出、高维空间提取特征, 平衡表达能力与计算效率, 其数学表达式如式(1)所示。

$$\mathbf{y} = \text{Linear}(\mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1 \cdot \mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2) \quad (1)$$

式中:  $\mathbf{x}$  表示输入到该结构的特征张量;  $\mathbf{W}_1$  是一个权重矩阵, 用于对输入  $\mathbf{x}$  进行线性变换;  $\mathbf{b}_1$  是对应的偏置向量;  $\mathbf{W}_1 \cdot \mathbf{x} + \mathbf{b}_1$  表示标准的线性层操作, 对输入特征进行初步变换, 改变通道数等维度信息;  $\text{ReLU}$  表示修正线性单元激活函数;  $\mathbf{W}_2$  是另一个权重矩阵;  $\mathbf{b}_2$  是偏置向量。  $\mathbf{W}_2$  对经过  $\text{ReLU}$  激活后的特征再次进行线性变换, 进一步调整特征维度, 使其符合输出要求,  $\mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1 \cdot \mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2$  得到中间结果。 $\text{Linear}$  表示线性投影, 对前面得到结果进行最后的线性变换, 将特征映射到期望的输出维度, 得到最终输出特征张量  $\mathbf{y}$ 。

对于 SE 通道注意力模块, 通过以下流程增强关键特征响应。先通过全局平均池化 (global average pooling,

### 1.3 MobileNetv3 模块

煤矿环境复杂, 粉尘、光照不均、机械振动等干扰会导致输送带图像对比度低、噪声大, 异物与背景区分困难。同时, 输送带高速运行要求分割算法低延迟响应, 且边缘设备算力有限, 需轻量化模型适配, 而异物形态多样又要求模型准确捕捉边缘细节。MobileNetv3<sup>[20]</sup> 作为轻量化网络, 继承了 v1 的深度可分离卷积以及 v2 的残差结构, 并引入 SE 通道注意力结构, 在性能与速度上表现优良, 具有参数少、计算量小的优点, 适用于内存与功耗受限的场景, 其网络结构如图 3 所示。

GAP) 将空间维度压缩为  $1 \times 1$ , 保留通道维度信息, 提取各通道全局特征, GAP 的具体数学表达如式(2)所示。

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (2)$$

式中:  $z_c$  表示经过全局平均池化后第  $c$  个通道的输出值;  $x_c(i, j)$  代表输入特征图中第  $c$  个通道、坐标为  $(i, j)$  位置的像素值;  $H$  和  $W$  分别表示输入特征图的高度和宽度;  $1/(H \times W)$  是归一化系数。

然后通过两层全连接层与激活函数, 生成通道注意力权重  $s_c$ , 自适应调整各通道特征重要性。通道注意力权重生成的数学表达式如式(3)所示。

$$s_c = \sigma(FC_2(\text{ReLU}(FC_1(z)))) \quad (3)$$

式中:  $s_c$  表示第  $c$  个通道的注意力权重;  $z$  是经过全局平均池化后的结果;  $FC_1$  和  $FC_2$  分别表示 2 个全连接层;  $\sigma$  代表 Sigmoid 函数。

最后将权重与原始特征相乘, 增强异物边缘、纹理等高频特征通道, 抑制粉尘噪声对应的低频通道, 缓解复杂背景下的特征混淆问题。该模块通过轻量化设计降低计算成本, 适配边缘设备算力限制, 同时 SE 注意力机制通过全局平均池化提取各通道的全局特征, 经全连接层生

成通道注意力权重,对异物边缘、纹理对应的高频特征通道赋予更高权重,同时抑制粉尘噪声主导的低频通道,从而提升特征纯净度。初步解决煤矿场景中边缘设备算力不足与噪声干扰的问题,为后续多尺度特征融合奠定基础。

#### 1.4 密集空洞空间金字塔池化

带式输送机作为煤矿智能化建设的核心运输装备,其安全运行面临多尺度异物并存、实时性要求高等挑战。传统 DeepLabv3+语义分割模型虽采用空洞空间金字塔池化 ASPP 模块提取多尺度特征,但各分支独立运算的架构导致跨层特征交互缺失,导致对形态差异显著的异物检测能力不足。因此引入密集空洞空间金字塔池化 DASPP 模块,通过串联不同膨胀率的空洞卷积层,使每层输出馈送至后续所有层,实现跨层特征的密集信息共享<sup>[21]</sup>。具体而言,设第  $l$  层为输入特征图  $y_{l-1}$ ,对应膨胀率为  $d_l$ ,则该层输出  $y_l$  由空洞卷积操作  $H_{d_l}(\cdot)$  与输入特征级联生成,如式(4)所示。

$$y_l = H_{d_l}(y_{l-1}) \oplus y_{l-1} \quad (4)$$

式中: $\oplus$  表示特征级联操作。通过这种递归连接机制,浅层细节特征(如异物边缘纹理)与深层语义特征(如煤流全局结构)逐步融合,形成包含多尺度上下文的复合特征图。

与 ASPP 的并行独立分支不同,DASPP 的密集连接架构突破了特征交互的局限性,实现了不同膨胀率下特征的逐层渐进式融合,解决传统模块对多形态异物的特征割裂问题。如图 4 所示,DASPP 模块的密集连接架构突破了传统并行分支的局限。每次空洞卷积后的特征输出均作为后续各层的输入,形成多层次特征交互网络,增强网络对上下文关系的感知能力,从而获取更丰富的多尺度特征信息。同时,模块内各层空洞卷积的膨胀率采用逐层递增设计,在不降低特征图分辨率的前提下渐进式扩大感受野,使模型能够兼顾小尺度异物局部细节捕捉与大尺度异物全局轮廓感知,进一步提升复杂工况下的分割精度。

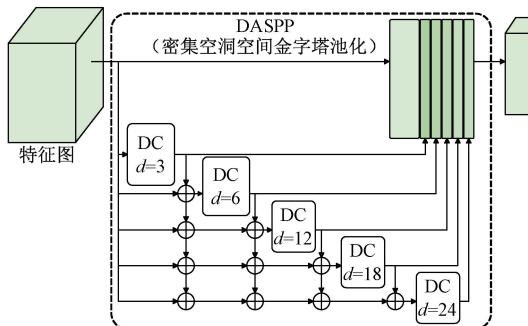


图 4 DASPP 模块

Fig. 4 DASPP module

空洞卷积的目的是保证特征图分辨率不变的情况下增加感受野。数学表达如式(5)所示。

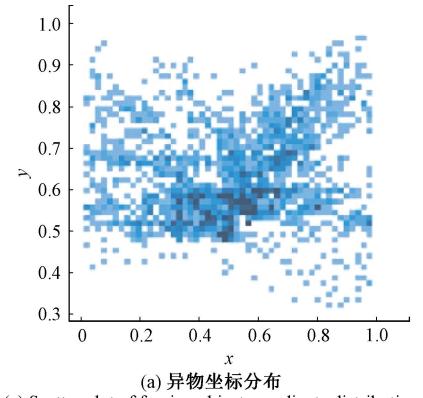
$$y[i] = \sum_{k=1}^K x[i + d \times k] \times w[k] \quad (5)$$

式中: $y[i]$  表示输出信号; $x[i]$  表示输入信号; $d$  表示膨胀率; $w[k]$  表示第  $k$  个过滤器的参数; $K$  表示过滤器(卷积核)的尺寸。空洞卷积是对输入  $x$  和在两个过滤器值之间插入  $d-1$  个 0 而产生的上采样滤波器进行卷积,因此膨胀率越大则感受野越大。DASPP 就是将不同膨胀率的空洞卷积进行密集连接增大感受野。 $H_{k,d}(x)$  表示空洞卷积,则 DASPP 的数学表达如式(6)所示。

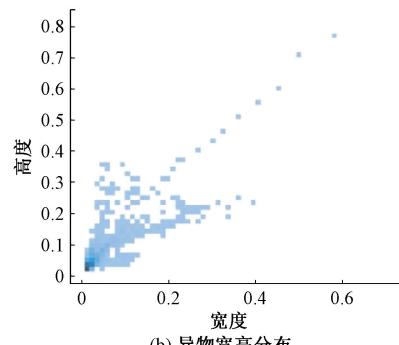
$$y_l = H_{K,d_l} \{ [y_{l-1}, y_{l-2}, \dots, y_0] \} \quad (6)$$

式中: $d_l$  表示第  $l$  层的膨胀率;  $[\dots]$  表示级联操作; $[y_{l-1}, \dots, y_0]$  表示特征图是通过对上层输出进行级联操作得到的。

如图 5 所示为标签分布情况图,呈现了煤矿输送带异物的标签分布特征,清晰反映出异物与背景的像素占比、空间位置及尺寸形态等关键信息。



(a) 异物坐标分布  
(a) Scatter plot of foreign object coordinate distribution



(b) 异物宽高分布  
(b) Scatter plot of foreign object width-height distribution

图 5 标签分布

Fig. 5 Label distribution

从整体像素占比来看,输送带与煤流在图像中占比超过 60%,对应大面积低灰度区域,凸显出背景复杂度高且异物前景占比小的特点。相比之下,锚杆、矸石等异物对应高灰度区域,二者像素占比合计不足 20%,且呈现碎

片化、多尺度分布特征,这直接加剧了小尺度异物检测的难度。此外,图像边缘因机械结构遮挡与光照干扰,常出现输送带边缘、滚筒部件与煤流的像素混叠,形成灰度过渡带,进一步增加了边界分割的复杂性。

异物坐标分布散点图(图 5(a))中,横轴和纵轴分别为归一化后的图像宽度与高度,散点分布直观呈现了异物的空间位置特征。其中,锚杆类异物的散点多沿输送带运行方向呈线状排列,与实际场景中锚杆沿输送带侧边或特定位置固定安装的特点一致;矸石类异物的散点则较为分散,在图像各区域均有分布,体现出矸石在煤流中随机出现的特性。同时,小尺度异物的散点相对密集且分布范围广,大尺度异物的散点则较稀疏,主要集中在煤流堆积较厚或存在大块矸石的区域。

异物宽高分布散点图(图 5(b))中,横轴为归一化后的异物宽度,纵轴为归一化后的异物高度,清晰反映了异物的尺寸形态差异。小尺度异物的宽高值较小,散点主要集中在图像左下角,多为细小矸石碎片或短小锚杆残留;大尺度异物的宽高值较大,散点分布在图像右上角,以大块矸石和完整长锚杆为主。此外,不同类型异物的宽高比例差异显著:细长锚杆宽度数值小、高度数值相对较大,呈现长条状特征;块状矸石的宽高比例则较均衡,接近方形分布。

上述类别不平衡、多尺度异物及背景干扰等问题,要求模型具备更强的多尺度特征融合能力与抗噪声干扰机制。DASPP 模块通过密集连接架构与逐层递增膨胀率设计,使模型在利用 MobileNetv3 轻量化主干压缩计算量的同时,能够通过跨层特征交互捕获多尺度异物细节,有效弥补了轻量化主干因参数精简可能导致的特征表达不足问题,二者形成协同平衡,针对性解决传统并行分支结构对小尺度异物漏检的核心挑战。

## 1.5 ECANet 注意力机制模块

矿用带式输送机运行中面临的高粉尘、低光照及机械振动等干扰,导致输送带图像质量较差,异物与背景特征区分度低,传统通道注意力模块因降维操作易丢失高频细节,且全连接层增加计算成本,难以满足煤矿边缘设备轻量化需求。基于此,在模型中引入 ECANet 注意力机制模块<sup>[22]</sup>,结构如图 6 所示。

其核心在于通过免降维全局池化与局部跨通道交互的高效建模,在几乎不增加计算成本的前提下提升特征表达能力,解决传统通道注意力模块,其理论推导过程为:

首先通过 GAP 直接聚合空间特征,其中,GAP 表示在不降维的情况下,对输入特征图执行全局平均池化,将通道维度的数量保持不变,而空间维度压缩为 1; $k$  是卷积核大小。具体流程为先通过全局平均池化层,将每个通道大的二维特征压缩为一个实数,特征维度由

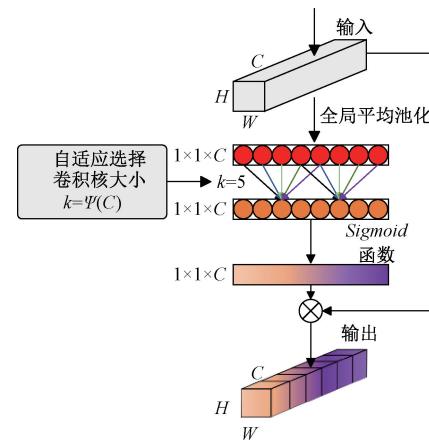


图 6 ECANet 网络结构

Fig. 6 ECANet network structure

$(C, H, W)$  变为  $(C, 1, 1)$ , 最后将归一化的权重加权乘以输入特征图,即  $(C, H, W) \times (C, 1, 1)$ , 得到输出的维度  $(C, H, W)$ 。具体表达如式(7)所示。

$$\omega = \sigma(Conv1D_k(z)) \quad (7)$$

式中: $\sigma$  代表 Sigmoid 函数; $k$  定义局部交互范围。

核大小  $k$  通过对对数函数与通道数  $C$  动态关联,使模型能根据通道数自适应调整交互范围,在轻量级的同时捕获跨通道依赖关系。具体表达如式(8)所示。

$$k = \psi(C) = \left\lfloor \frac{\log_2(C) + b}{\gamma} \right\rfloor_{odd}, \gamma = 2, b = 1 \quad (8)$$

式中: $k$  表示一维卷积的核大小; $\psi(C)$  是关于通道数  $C$  的函数; $\log_2(C)$  是以 2 为底,通道数  $C$  的对数; $\lfloor \cdot \rfloor_{odd}$  表示取最近的奇数; $b$  和  $\gamma$  是式(8)中固定的参数,根据经验设置为  $b=1, \gamma=2$ , 以平衡不同通道数下的局部交互范围。例如当通道数  $C=32$  时, $k=4$ , 经  $\lfloor \cdot \rfloor_{odd}$  操作取最近奇数后,核大小为 5。这种动态调整机制确保模型在浅层和深层均能自适应捕获合适范围的跨通道依赖,避免固定核大小导致的特征建模不足或计算冗余。核大小  $k$  取最近奇数时选择 5 而非 3, 是因模型中层需中等范围跨通道依赖,5 的一维卷积覆盖更广交互范围,且原文参数使结果趋较大奇数,同时奇数核可中心对称卷积以均衡通道权重建模。

ECANet 通过免降维全局池化与动态一维卷积的设计,在轻量化框架下实现高效的通道注意力建模,与 MobileNetv3 轻量化主干和 DASPP 多尺度特征提取模块协同作用时,ECANet 进一步优化了模型的整体性能。MobileNetv3 通过深度可分离卷积降低计算量,ECANet 在此基础上以极低代价增强特征质量,避免轻量化导致的精度下降。DASPP 解决多尺度异物的特征融合问题,ECANet 则对各尺度特征的通道权重进行精细化调整,三者共同构建完整的技术体系,满足煤矿带

式输送机异物精准分割在实时性、鲁棒性和集成化方面的应用需求。

## 2 实验结果与分析

### 2.1 数据集构建

此次实验使用的煤矿带式输送机异物数据集源自中国矿业大学智能检测与模式识别研究中心研发的煤矿专用视频AI分析数据集(China University of Mining and Technology-Belt Conveyor, CUMT-BelT)。该数据集采集于真实矿井下皮带运输场景,涵盖多种典型工况下的带式输送机运行图像,部分示例如图7所示。

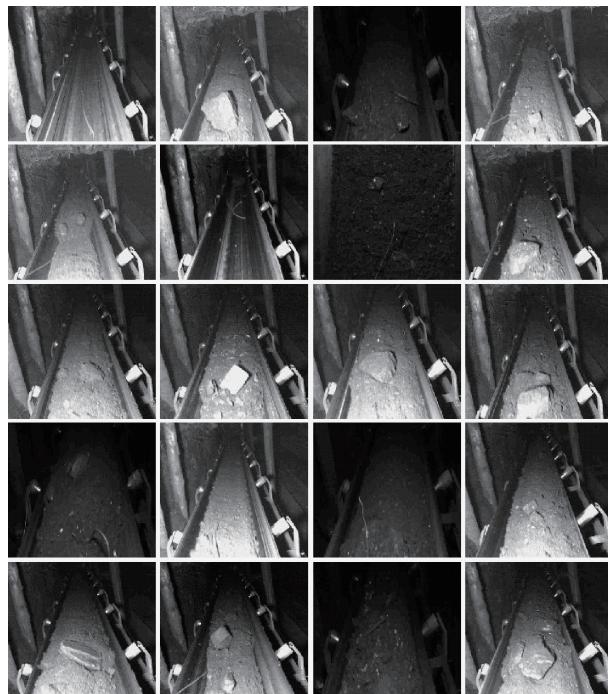


图7 部分数据集展示

Fig. 7 Partial dataset display

为进一步提升模型在复杂环境中的鲁棒性与泛化能力,数据集构建过程中引入了多样化的数据增强策略,如图8所示。

通过水平/垂直旋转、平移及旋转扭曲等操作,模拟输送带不同运行角度、物料堆积形态及粉尘干扰程度下的图像特征,将原始数据集规模扩展3倍,有效缓解样本类别不平衡问题。经数据增强处理后,数据集扩张至5 417张,分辨率为 $640\times 640$ 。然后将数据集划分为训练集、验证集和测试集,比例为7:2:1。训练集用于模型参数的学习,验证集用于在训练过程中调整模型的超参数,而测试集则用于评估最终模型的性能。

目标标注工作采用基于Python开发的开源工具

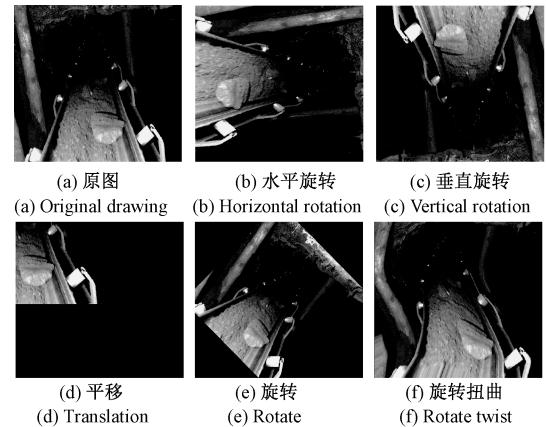


图8 数据增强策略

Fig. 8 Data augmentation processing

Labelme完成,实现对输送带、煤流、锚杆及矸石等关键目标的精细化标注,构建像素级语义分割真值图,示例如图9所示。标注过程中遵循“语义优先”原则处理重叠区域,如锚杆覆盖煤流时优先标注锚杆,并通过人工二次校验修正模糊边界,确保标注精度。

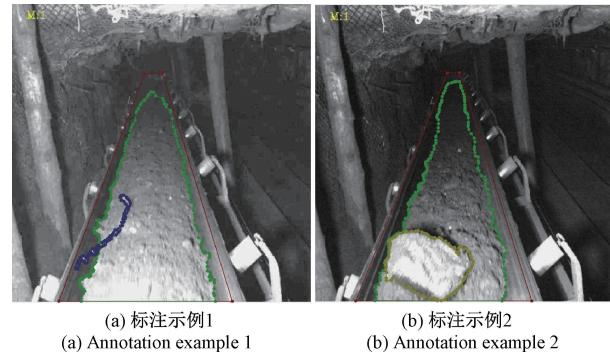


图9 Labelme 标注示例

Fig. 9 Labelme annotation example

### 2.2 实验设置与评价指标

#### 1) 实验环境配置

实验使用Pycharm作为编程软件,CPU为12th Gen Intel(R) Core(TM) i9-12900H,操作系统为Windows 11,内存为64 Gib, GPU为NVIDIA GeForce RTX 3080Ti,CUDA版本为12.2,PyTorch版本为2.5.0,Python版本为3.10.0。为保证模型检测性能的公平性和可比性,所有实验在训练时统一采用 $512\text{ pixels}\times 512\text{ pixels}$ 的输入尺寸,没有使用任何预训练权值。

在训练过程中,优化器<sup>[23]</sup>选择(adaptive moment estimation, Adam),其自适应调整学习率的特性能够在训练初期快速收敛,同时在接近最优解时避免过度振荡。将学习率定为0.001,权重衰减系数设为0.0001,以此来避免模型出现过拟合的情况。同时,训练时的批处理大

小(batchsize)被设置为 16, 经过多次试验, 该值在模型训练速度和内存占用之间达到了较好的平衡。训练的总轮数(epoch)设定为 100, 通过观察验证集上的性能指标, 判断模型是否收敛, 确保模型能够充分学习数据特征, 又不至于出现过拟合现象。

针对 MobileNetv3 轻量化主干网络, 本实验采用 MobileNetv3-Large 版本。该版本通过深度可分离卷积将标准卷积计算量压缩至 1/9, 并利用线性瓶颈残差结构在低维空间处理输入输出、高维空间提取特征, 平衡表达能力与计算效率。此外, MobileNetv3-Large 嵌入 SE 通道注意力模块, 通过全局平均池化与两层全连接层生成通道权重, 增强异物边缘、纹理等高频特征通道, 抑制粉尘噪声对应的低频通道, 较 MobileNetv3-Small 版本更适用于煤矿复杂工况下的多尺度特征提取。

密集空洞空间金字塔池化 DASPP 模块通过串联不同膨胀率的空洞卷积层实现跨层特征交互, 各层膨胀率参数依据煤矿异物多尺度分布特性设计为:

第 1 层为膨胀率  $d_1=6$ , 对应  $3\times 3$  空洞卷积, 捕获锚杆边缘、矸石碎片等小尺度异物的局部细节特征; 第 2 层为膨胀率  $d_2=12$ , 通过扩大感受野融合中等尺度异物(如块状矸石)的轮廓信息; 第 3 层为膨胀率  $d_3=18$ , 进一步扩展感受野以感知煤流与输送带的中等范围上下文关系; 第 4 层为膨胀率  $d_4=24$ , 通过大尺度空洞卷积捕获输送带全局形态与煤流分布等语义特征。

各层输出特征按式  $y_i = H_{dl}(y_{i-1}) \oplus y_{i-1}$  级联, 使浅层细节特征与深层语义特征渐进式融合, 形成包含多尺度上下文的复合特征图, 解决传统 ASPP 并行分支结构对小尺度异物的漏检问题。

## 2) 评价指标

为全面、客观地评估改进后的煤矿带式输送机异物分割模型性能, 选用以下 3 种评价指标。

(1) 平均交并比<sup>[24]</sup> (mean intersection over union, mIoU), 平均交并比是衡量模型预测分割结果与真实标注的重叠程度, 计算每个类别预测区域与真实区域的交并比, 再取平均值, 计算 mIoU 的公式如式(9)所示。

$$mIoU = \frac{1}{n} \sum_{i=1}^n \frac{|A_i \cap B_i|}{|A_i \cup B_i|} \quad (9)$$

式中:  $n$  表示数据集中的类别数;  $A_i$  表示模型预测的第  $i$  类目标的像素集合;  $B_i$  表示真实标注的第  $i$  类目标的像素集合;  $|A_i \cap B_i|$  表示第  $i$  类预测结果与真实标注的交集像素数;  $|A_i \cup B_i|$  表示第  $i$  类预测结果与真实标注的并集像素数。 $mIoU$  越趋近于 1, 说明模型对异物区域的分割效果越好, 能够精准定位异物位置、完整勾勒其范围, 并实现异物与背景的有效辨识。

(2) 精确率(precision,  $P$ )是指模型正确预测为异物的像素数占所有预测为阳性像素数的比例, 反映模型预

测阳性结果的可靠性, 计算公式如式(10)所示。

$$P = \frac{TP}{TP + FP} \quad (10)$$

式中: 真正例(true positive,  $TP$ )表示被正确预测的像素数; 假正例(false positive,  $FP$ )表示被错误预测的像素数。

(3)  $F1$  分数( $F1$ -Score)表示分割模型精确率  $P$  和召回率(recall,  $R$ )的一种调和平均, 如式(11)所示。

$$F1 = \frac{2 \times P \times R}{P + R} \quad (11)$$

$F1$  分数通过均衡考量精度与召回率为模型性能提供了更具全面性的评估框架, 较高的  $F1$  分数表明模型既能实现对带式输送机异物的精准识别, 又可最大限度完成潜在异物的检测覆盖, 有效降低漏检与误检概率; 在模型复杂度评估方面, 采用参数数量(parameters)、浮点运算次数(floating point operations, FLOPs) 2 项核心指标, 通过多维度指标综合分析, 可准确了解模型在结构复杂度与工程部署效率间的平衡特性, 为煤矿带式输送机异物分割模型的优化设计提供关键理论依据。

## 2.3 实验结果对比与讨论

### 1) 消融实验

为验证改进 DeepLabv3+ 模型中各模块的有效性, 设计消融实验逐步移除关键组件, MobileNetv3 轻量化主干、DASPP 多尺度特征提取模块、ECANet 通道注意力机制, 对比分析各配置下模型的性能表现与复杂度指标。实验采用相同的训练参数与数据集划分方式, 以平均交并比、 $F1$  分数、参数数量、FLOPs 及推理速度为评估基准。在消融实验的实验配置中, 基准模型为原始 DeepLabv3+ 模型, 配备 ASPP 模块与基础通道注意力机制; 模型 A 是将基准模型的主干替换为 MobileNetv3, 同时保留 ASPP 与基础注意力; 模型 B 在模型 A 的基础上, 把 ASPP 替换为 DASPP 模块, 并移除注意力机制; 模型 C 则是在模型 B 的基础上集成 ECANet 通道注意力机制, 此为完整改进模型。其实验结果如表 1 所示。

表 1 消融实验

Table 1 Ablation experiment

| 模型配置 | 平均交并比<br>(mIoU)/% | $F1$ 分数<br>( $F1$ )/% | 参数数量<br>/M | FLOPs<br>/G | 推理速<br>度/fps |
|------|-------------------|-----------------------|------------|-------------|--------------|
| 基准模型 | 82.5              | 81.2                  | 27.3       | 12.8        | 18.7         |
| 模型 A | 83.8              | 83.1                  | <b>9.1</b> | <b>4.5</b>  | <b>42.3</b>  |
| 模型 B | 85.2              | 84.9                  | 9.4        | 4.8         | 40.1         |
| 模型 C | <b>87.1</b>       | <b>86.7</b>           | 9.8        | 5.1         | 38.6         |

从消融实验的结果来看, 仅替换为 MobileNetv3 主干的模型 A, 凭借深度可分离卷积将计算量压缩至传统卷

积的 1/9,结合 SE 注意力抑制粉尘噪声、增强异物高频特征,使 mIoU 从 82.5% 提升至 83.8%,F1 分数从 81.2% 提升至 83.1%,参数减少 66.7% 至 9.1 M,验证了轻量化主干对效率与基础特征提取能力的提升。但由于保留了传统 ASPP 的并行分支结构,缺乏跨层特征交互机制,模型对锚杆、矸石等多尺度异物的上下文关联感知不足,导致小尺度异物特征易被噪声淹没,因此 mIoU 和 F1 分数提升幅度有限。进一步替换 ASPP 为 DASPP 的模型 B,通过密集连接实现跨层特征交互,逐层递增的膨胀率扩大感受野以融合多尺度信息,mIoU 提升至 85.2%,F1 分数提升至 84.9%,表明多尺度特征融合对解决小尺度异物漏检和复杂形态异物检测的关键作用。然而,由于移除了注意力机制,模型缺乏对通道间特征权重的动态优化,粉尘噪声对应的低频通道未被有效抑制,导致异物与背景的特征区分力不足,因此 mIoU 和 F1 分数提升幅度低于集成注意力机制的模型 C。最终集成 ECANet 的模型 C,利用免降维池化保留通道完整信息,动态一维卷积通过建模跨通道依赖,精准抑制粉尘噪声对应的低频特征,增强异物边缘、纹理等高频特征的响应强度,使改进模型的 mIoU 提升至 87.1%,F1 分数提升至 86.7%,且参数仅增加 0.4~9.8 M。这一优化效果源于三大技术路径的协同。首先,MobileNetv3 轻量化主干通过深度可分离卷积将计算量压缩至传统卷积的 1/9,结合线性瓶颈残差结构与 SE 注意力,使参数从 27.3 M 降至 9.1 M;其次,DASPP 模块通过密集跨层连接与逐层递增的膨胀率设计,在仅增加 0.3 M 参数的情况下实现多尺度特征高效融合;最后,ECANet 通过免降维池化与动态一维卷积,以 0.4 M 参数增量强化通道特征权重优化。三者协同使改进模型计算量较基准模型降低 63.1% (12.8 G 降到 5.1 G),参数精简至 9.8 M。这一结果表明,通道注意力机制通过优化多尺度特征的权重分配,弥补轻量化和多尺度融合过程中可能损失的特征判别力,提升复杂背景下的分割精度。

从推理速度来看,模型 A 凭借 MobileNetv3 轻量化主干的深度可分离卷积设计,将推理速度提升至 42.3 fps,较基准模型的 18.7 fps 实现翻倍增长,验证了轻量化架构对边缘设备实时性的关键提升作用。模型 B 因 DASPP 模块的密集跨层连接,虽计算量略有增加导致推理速度小幅下降至 40.1 fps,但 mIoU 提升 1.4%,说明多尺度特征融合对精度的增益可抵消部分速度损耗。模型 C 集成 ECANet 后,推理速度进一步降至 38.6 fps,这是由于通道注意力机制引入的动态一维卷积带来额外计算开销,但 mIoU 提升 1.9% 且 F1 分数突破 86%,表明轻量级注意力模块在精度与速度间实现了良好平衡。

实验数据表明,MobileNetv3 的轻量化设计、DASPP 的多尺度特征融合机制与 ECANet 的通道优化能力形成

递进协同,在改进模型参数仅为基准模型 35.9% 的情况下,实现 4.6% 的 mIoU 提升和 5.5% 的 F1 分数提升,解决了煤矿场景中噪声干扰强、多尺度异物并存的检测难题。各模块的协同优化并非简单的性能叠加,而是通过架构创新实现效率与精度的协同进化,MobileNetv3 的线性瓶颈残差结构在压缩计算量的同时保留关键特征通道,为 DASPP 的多尺度融合提供优质特征输入;DASPP 通过渐进式膨胀率设计扩大感受野,使 ECANet 能更精准地定位异物高频特征通道;而 ECANet 的免降维池化策略则避免了轻量化过程中特征维度丢失,该模型以 38.6 fps 的推理速度和 87.1% 的 mIoU,既满足了输送带高速运行下的实时预警需求,又实现了对锚杆、矸石等多尺度异物的精准分割,为工业现场的智能化改造提供了可落地的技术方案。

## 2) 不同模型对比实验

为验证改进模型在煤矿带式输送机异物分割任务中的优越性,选取 PSPNet (pyramid scene parsing network)、DeepLabv3+、U-Net、HRNet (high-resolution network)、SegFormer、MaskR-CNN (mask regional convolutional neural network)、Swin Transformer (shifted window transformer)、MobileSAM (mobile segment anything model)、Fast-SCNN (fast semantic segmentation network)、SeaFormer (seaformer vision transformer) 等主流分割模型开展对比实验。所有模型均采用默认超参数,未使用预训练权重,与改进模型保持相同训练轮次和输入尺寸,并基于相同数据集与实验环境进行训练和测试,采用平均交并比、F1 分数、参数数量、FLOPs 作为评估指标,结果如表 2 所示。

表 2 不同分割算法下对比结果

Table 2 Comparison results under different segmentation algorithms

| 模型名称             | 平均交并比<br>(mIoU)/% | F1 分数<br>(F1)/% | 参数数量<br>/M | FLOPs<br>/G |
|------------------|-------------------|-----------------|------------|-------------|
| PSPNet           | 78.3              | 77.1            | 49.2       | 28.5        |
| DeepLabV3+       | 82.5              | 81.2            | 27.3       | 12.8        |
| U-Net            | 80.1              | 79.5            | 34.8       | 18.7        |
| HRNet            | 83.2              | 82.3            | 63.5       | 35.6        |
| SegFormer        | 84.5              | 83.8            | 18.2       | 9.6         |
| Mask R-CNN       | 79.4              | 78.2            | 41.7       | 22.1        |
| Swin Transformer | 85.0              | 84.1            | 56.8       | 32.4        |
| MobileSAM        | 76.8              | 75.9            | 35.4       | 18.3        |
| Fast-SCNN        | 81.5              | 80.7            | 5.7        | 2.3         |
| SeaFormer        | 78.9              | 84.2            | 11.4       | 7.2         |
| 改进模型             | <b>87.1</b>       | <b>86.7</b>     | <b>9.8</b> | <b>5.1</b>  |

从表 2 不同分割算法对比结果可见,改进模型在多项关键指标上展现出优势。基准模型 DeepLabV3+虽具备一定分割能力,mIoU 为 82.5%、F1 分数为 81.2%,但其参数数量达 27.3 M,在轻量化上存在局限。PSPNet 的 mIoU 为 78.3%、F1 分数为 77.1%,U-Net 的 mIoU 为 80.1%、F1 分数为 79.5%,二者因缺乏高效多尺度特征融合机制,对煤矿场景中锚杆、矸石等多形态异物的检测精度较低,且参数规模均超过 30 M,计算成本较高。HRNet 的 mIoU 为 83.2%、F1 分数为 82.3%,虽通过高分辨率网络提升特征精度,但参数数量高达 63.5 M,计算复杂度较高。MaskR-CNN 的 mIoU 为 79.4%、F1 分数为 78.2%,作为目标检测与分割结合的模型,在语义分割任务中表现出特征区分度不足的问题。Swin Transformer 的 mIoU 为 85.0%、F1 分数为 84.1%,通过 Transformer 架构增强全局依赖建模,但其参数数量 56.8 M,计算复杂度较高,不适用于低功耗场景。MobileSAM 的 mIoU 为 76.8%、F1 分数为 75.9%,其在煤矿复杂工况下对异物与背景的特征区分能力较弱,导致分割精度相对较低,且参数数量为 35.4 M,在轻量化方面也不具优势。Fast-SCNN 的 mIoU 为 81.5%、F1 分数为 80.7%,虽然参数数量仅为 5.7 M,在轻量化上表现突出,但由于其网络结构相对简单,对多尺度异物的特征捕捉能力有限,精度提升空间较小。SeaFormer 的 mIoU 为 78.9%、F1 分数为 84.2%,其在 F1 分数上有一定表现,但 mIoU 较低,说明在异物分割的整体准确性上还有不足,且参数数量为 11.4 M,综合性能有待提高。相比之下,改进模型以 9.8 M 参数实现了 87.1% 的 mIoU 和 86.7% 的 F1 分数,其核心优势在于,通过 MobileNetv3 轻量化主干压缩计算量,结合 DASPP 密集连接的多尺度特征融合与 ECANet 免降维通道注意力机制,在几乎不增加计算成本的前提下,提升复杂背景下异物与背景的特征区分度。实验结果表明,改进模型在多尺度异物分割精度与轻量化之间实现了最优平衡。

另外,改进模型在小尺度异物检测中具有优势,Transformer 架构模型在此场景下性能表现与改进模型存在差异。改进模型通过引入 DASPP 模块,以串联不同膨胀率的空洞卷积层实现跨层特征密集交互,结合逐层递增的膨胀率设计,在不降低特征图分辨率的前提下渐进式扩大感受野,使模型能够兼顾小尺度异物局部细节捕捉与大尺度异物全局轮廓感知,提升对小尺度异物的检测能力。同时,ECANet 通道注意力机制通过免降维全局池化和动态一维卷积,增强异物高频特征通道、抑制粉尘噪声对应的低频通道,进一步提升复杂背景下小尺度异物与背景的特征区分度,从而降低小尺度异物的漏检率。以 Swin Transformer 为代表的 Transformer 架构模型,虽然通过 Transformer 架构增强

了全局依赖建模, mIoU 达到 85.0%、F1 分数为 84.1%,但在小尺度异物检测方面表现不如改进模型。Transformer 架构更擅长处理全局语义信息和长距离依赖关系,但对局部细节特征的捕捉能力相对较弱,在面对煤矿场景中碎片化、多尺度分布的小尺度异物时,难以精准定位和分割其边缘细节,导致小尺度异物的分割精度不足。此外,Transformer 架构通常参数量较大,计算复杂度高,如 Swin Transformer 参数数量达 56.8 M,这在一定程度上影响了模型在边缘设备上的实时性和适配性,而改进模型通过 MobileNetv3 轻量化主干压缩计算量,参数数量仅为 9.8 M,在小尺度异物检测精度和模型轻量化方面实现了更好的平衡。

如图 10 所示呈现了不同模型在煤矿带式输送机异物分割任务中的可视化效果,其中最左侧列为原始图像,第 2~9 列依次为各模型的分割结果。图 10 中分割结果的灰度语义定义为灰度值 0 对应输送带以外的背景,灰度值 100 对应输送带,灰度值 150 对应煤矸石,灰度值 200 对应锚杆,灰度值 250 对应煤流。该定义适用于图 10(a) 不同模型效果对比图中各模型分割结果的可视化表征,并与图 10(b) 分割结果颜色编码释义图对应。

从可视化结果可见,各模型在复杂工况下的分割性能差异显著。PSPNet 对煤流与输送带的基础分割尚可,但对锚杆、矸石等小尺度异物漏检严重,大量细碎矸石未被标注,锚杆的标注也存在缺失,这与其缺乏跨层特征交互的多尺度融合机制有关。DeepLabV3+作为基准模型,能较好地分割输送带和煤流的整体区域,但对小尺度煤矸石的识别不完整,锚杆的分割存在断裂,体现出传统 ASPP 模块在多尺度特征融合上的局限性。U-Net 虽能捕捉部分中等尺度矸石的轮廓,但对细长锚杆的分割出现细微断裂,且受粉尘噪声干扰,分割结果呈现区域融合过度的特征,难以区分异物与噪声的细微差异。HRNet 凭借高分辨率特征保留优势,对输送带边缘的分割较平滑,但参数冗余导致对小尺度煤矸石的特征响应不足,整体分割精度提升有限。SegFormer 对大尺度煤流区域的分割较为完整,但小尺度异物的分割连续性差,如锚杆被分割为多段短线条且特征并不完整,煤矸石的边缘也存在较多缺失,这与其特征融合过程中对高频细节的保留不足相关。MaskR-CNN 作为检测与分割结合的模型,在异物定位上有一定能力,但分割边界粗糙,尤其对锚杆特征分割出现断裂情况,不符合语义分割的像素级精度要求。Swin Transformer 对全局语义的捕捉较好,输送带与煤流的整体轮廓准确,但对小尺度异物的局部细节刻画不足,细小煤矸石被遗漏,锚杆分割出现断裂。MobileSAM 的分割效果较差,煤矸石仅能识别极少数

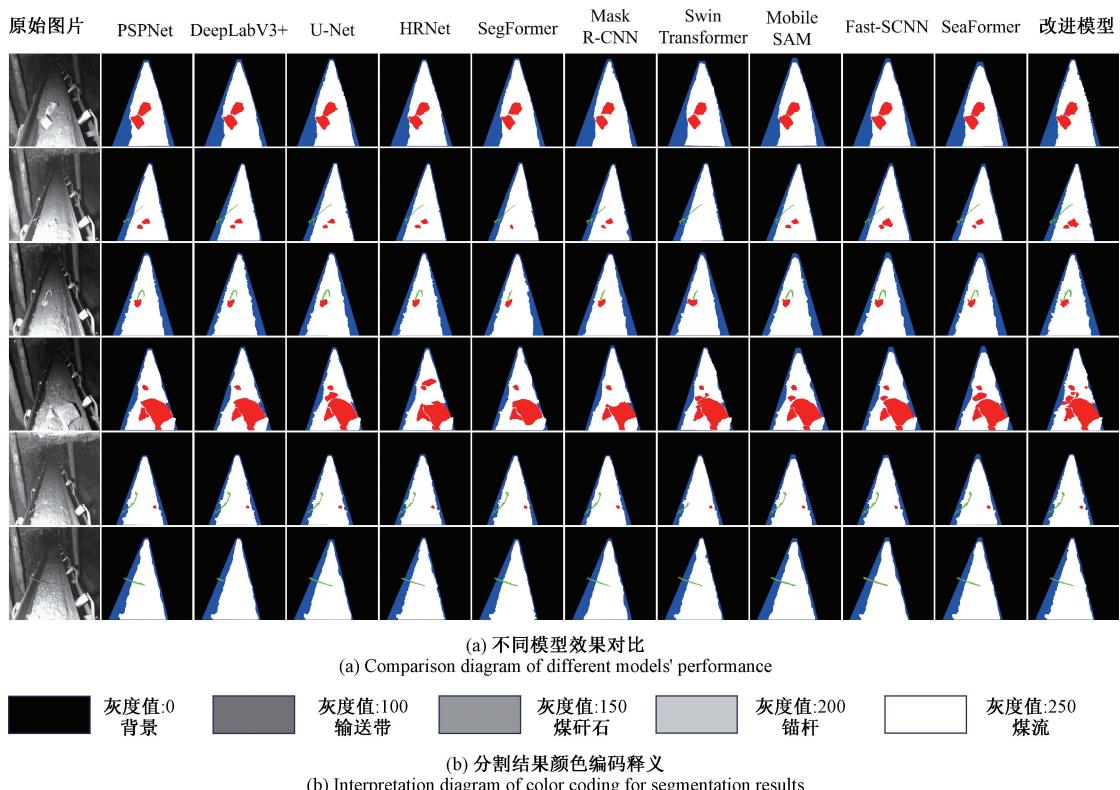


图 10 不同模型分割效果对比

Fig. 10 Comparison of segmentation effects among different models

大尺寸目标煤流与输送带的边界模糊,整体分割结果粗糙,难以适应煤矿复杂工况。Fast-SCNN 虽参数轻量化,但分割结果粗糙,煤矸石仅能识别大尺寸块状目标,煤流与输送带的边界也不清晰,无法应对煤矿场景中多尺度异物并存的复杂情况,体现出简单网络结构在特征表达能力上的局限性。SeaFormer 对煤矸石的部分区域分割较完整,小尺度煤矸石漏检较多。煤流与背景的边界处理上,分割结果整体偏于平滑,难以精准呈现异物的精细化特征。这些表现充分暴露了其在多尺度特征捕捉能力以及抗噪声干扰性能上的不足。

相比之下,改进模型的分割效果较优。锚杆被完整标注为连续线条,即使在被煤流部分遮挡的情况下仍能保持完整性。煤矸石无论大小均被精准分割,边缘平滑且与背景界限清晰,无过多漏检。输送带与煤流的边界分割准确,有效避免了粉尘噪声导致的特征混淆,整体分割结果与真值图的重叠度最高。这一优势源于 DASPP 模块的跨层特征密集交互强化了多尺度异物的特征关联,ECA-Net 对高频特征通道的增强提升了异物与噪声的区分度,而 MobileNetv3 的轻量化设计在压缩计算量的同时,通过 SE 注意力保留了关键细节特征,实现了高精度与抗干扰能力的平衡。

### 3 结 论

针对煤矿带式输送机复杂工况下异物分割面临的噪声干扰强、多尺度特征捕捉难及模型轻量化与精度难以平衡等问题,提出基于改进 DeepLabv3+的煤矿输送带异物分割模型,通过引入 MobileNetv3 轻量化主干、DASPP 多尺度特征融合模块及 ECA-Net 通道注意力机制,有效提升模型性能。实验结果表明,改进模型的平均交并比达 87.1%,参数数量仅 9.8 M;较原始模型精度提升 4.6%、计算量降低 63.1%,性能优于主流模型。为煤矿带式输送机异物精准分割提供了可靠的技术支撑,未来可进一步提升模型泛化性与实时性。

### 参考文献

- [1] 王鹏超,卢冰,刘超,等.煤炭行业现状及发展格局趋势预测[J].现代工业经济和信息化,2025,15(1):35-36,39.  
WANG P CH, LU B, LIU CH, et al. Currentsituation and development trend prediction of the coal industry [J]. Modern Industrial Economy and Informatization, 2025, 15(1): 35-36,39.
- [2] 毛清华,郭文瑾,翟姣,等.煤矿带式输送机异常状态

- 视频AI识别技术研究[J]. 工矿自动化, 2023, 49(9): 36-46.
- MAO Q H, GUO W J, ZHAI J, et al. Research on video AI recognition technology for abnormal state of coal mine belt conveyors [J]. Journal of Mine Automation, 2023, 49(9): 36-46.
- [3] 周坪, 马国庆, 周公博, 等. 智能化带式输送机健康监测技术研究综述[J]. 仪器仪表学报, 2023, 44(12): 1-21.
- ZHOU P, MA G Q, ZHOU G B, et al. A review of research on health monitoring technology for intelligent belt conveyors [J]. Chinese Journal of Scientific Instrument, 2023, 44(12): 1-21.
- [4] SIAMI M, BARSZCZ T, WODECKI J, et al. Semantic segmentation of thermal defects in belt conveyor idlers using thermal image augmentation and U-Net-based convolutional neural networks [J]. Scientific Reports, 2024, 14(1): 5748-5763.
- [5] 程健, 王东伟, 杨凌凯, 等. 一种改进的高斯混合模型煤矸石视频检测方法[J]. 中南大学学报(自然科学版), 2018, 49(1): 118-123.
- CHENG J, WANG D W, YANG L K, et al. An improved Gaussian mixture model for coal gangue video detection [J]. Journal of Central South University (Science and Technology), 2018, 49(1): 118-123.
- [6] CARVALHO R, NASCIMENTO R, D'ANGELO T, et al. A UAV-based framework for semi-automated thermographic inspection of belt conveyors in the mining industry [J]. Sensors, 2020, 20(8): 2243-2262.
- [7] 张宁波, 刘长友, 陈现辉, 等. 综放煤矸低水平自然射线的涨落规律及测量识别分析[J]. 煤炭学报, 2015, 40(5): 988-993.
- ZHANG N B, LIU CH Y, CHEN X H, et al. Measurement analysis on the fluctuation characteristics of low level natural radiation from gangue [J]. Journal of China Coal Society, 2015, 40(5): 988-993.
- [8] SHIRI H, WODECKI J, ZIETEK B, et al. Inspection robotic UGV platform and the procedure for an acoustic signal-based fault detection in belt conveyor idler [J]. Energies, 2021, 14(22): 7646-7663.
- [9] 李嘉豪, 司垒, 王忠宾, 等. 综放工作面煤矸识别技术及其应用[J]. 仪器仪表学报, 2024, 45(1): 1-15.
- LI J H, SI L, WANG ZH B, et al. Coal gangue identification technology and its application in fully-mechanized coal mining face [J]. Chinese Journal of Scientific Instrument, 2024, 45(1): 1-15.
- [10] BUTT M, GLAS N, MONSuur J, et al. Application of YOLOv8 and detectron2 for bullet hole detection and score calculation from shooting cards [J]. AI, 2024, 5(1): 72-90.
- [11] KOMATSU R, GONSALVES T. Comparing U-Net based models for denoising color images [J]. AI, 2020, 1(4): 465-486.
- [12] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2): 440-468.
- TIAN X, WANG L, DING Q. Review of image semantic segmentation based on deep learning [J]. Journal of Software, 2019, 30(2): 440-468.
- [13] 陈雷, 杨吉斌, 曹铁勇, 等. 一种基于Transformer特征金字塔的自蒸馏目标分割方法[J]. 电子与信息学报, 2025, 47(2): 551-560.
- CHEN L, YANG J B, CAO T Y, et al. A self-distillation object segmentation method based on Transformer feature pyramid [J]. Journal of Electronics & Information Technology, 2025, 47(2): 551-560.
- [14] 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 D Q, XU J Y, KOU Q Q, 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.
- [15] 李永安, 陈腾杰, 王宏伟, 等. 基于DR-YOLOM的带式输送机运行状态多任务检测方法[J]. 煤炭科学技术, 2025, 53(S1): 454-467.
- LI Y AN, CHEN T J, WANG H W, et al. Multi task detection method for operating status of belt conveyor based on DR-YOLOM [J]. Coal Science and Technology, 2025, 53(S1): 454-467.
- [16] 吕瑞宏, 李大玮, 沈红博. 基于CT图像的煤矸元素成分分析算法研究[J]. 仪器仪表学报, 2025, 46(3): 170-179.
- LYU R H, LI D W, SHEN H B. Research on coal gangue elemental composition analysis algorithm [J]. Chinese Journal of Scientific Instrument, 2025, 46(3): 170-179.
- [17] 马宏伟, 周文剑, 王鹏, 等. 改进的ORB-FLANN煤矸石图像高效匹配方法[J]. 煤炭科学技术, 2024, 52(1): 288-296.
- MA H W, ZHOU W J, WANG P, et al. Improved ORB-FLANN efficient matching method for coal gangue image [J]. Coal Science and Technology, 2024, 52(1): 288-296.
- [18] ZHAO L J, HAN L G, ZHANG H N, et al. Study on recognition of coal and gangue based on multimode

- feature and image fusion[J]. Plos One, 2023, 18(2): 0281397.
- [19] 梁运涛,张鹏成,孙勇,等.带式输送机回程托辊卡死故障摩擦升温实验研究[J].工矿自动化,2025,51(1): 95-103.  
LIANG Y T, ZHANG P CH, SUN Y, et al. Experimental study on frictional heating due to return idler seizure failure in belt conveyors[J]. Journal of Mine Automation, 2025, 51(1): 95-103.
- [20] 陈双叶,徐雷桁,黄成意,等.基于MobileNetV3卷积神经网络的供水管道漏损音频分类[J].北京工业大学学报,2024,50(7): 797-804.  
CHEN SH Y, XU L H, HUANG CH Y, et al. Research on audio classification of water supply pipeline leakage based on MobileNetV3 convolutional neural network[J]. Journal of Beijing University of Technology, 2024, 50(7): 797-804.
- [21] 唐雪瑾,杨卫华,于晋伟.基于多尺度池化与特征融合的轻量级语义分割算法[J].微电子学与计算机,2024,41(12):1-9.  
TANG X J, YANG W H, YU J W. Lightweight semantic segmentation algorithm based on multi-scale pooling and featurefusion[J]. Microelectronics & Computer, 2024, 41(12):1-9.
- [22] 胡程喜,谭立新,王文胤,等.基于改进DeepLabV3+的轻量化茶叶嫩芽采摘点识别模型[J].智慧农业(中英文),2024,6(5):119-127.  
HU CH X, TAN L X, WANG W Y, et al. Lightweight tea shoot picking point recognition model based on improved DeepLabV3+[J]. Smart Agriculture, 2024, 6(5):119-127.
- [23] 蔡玉芳,王涵,李琦,等.联合自然梯度和AdamW算法的RSF图像分割模型[J].仪器仪表学报,2023,44(3): 261-270.  
CAI Y F, WANG H, LI Q, et al. RSF image segmentation model joint natural gradient and AdamW algorithm[J]. Chinese Journal of Scientific Instrument, 2023, 44(3): 261-270.
- [24] 李天放,孙一宸,于明鑫,等.结合语义分割与跨模态差分特征补偿的红外与可见光图像融合方法[J].电子测量与仪器学报,2024,38(7): 34-45.  
LI T F, SUN Y CH, YU M X, et al. Infrared and visible image fusion method integrating semantic segmentation and cross-modality differential feature compensation[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 34-45.

## 作者简介



**刘光伟**,2004年于辽宁工程技术大学获得学士学位,2007年于辽宁工程技术大学获得硕士学位,2010年于辽宁工程技术大学获得博士学位,现为辽宁工程技术大学教授,主要研究方向为露天开采理论与技术、矿业系统工程等。

E-mail:liuguangwei@lntu.edu.cn

**Liu Guangwei** received his B. Sc., M. Sc., and Ph. D. degrees from Liaoning Technical University in 2004, 2007, and 2010, respectively. Now he is a professor at Liaoning Technical University. His main research interests include open-pit mining theory and technology, mining systems engineering, etc.



**张浩博**(通信作者),2022年于广西科技大学获得学士学位,现为辽宁工程技术大学在读研究生,主要研究方向为机器视觉和露天开采理论与技术。

E-mail:871407626@qq.com

**Zhang Haobo** (Corresponding author)

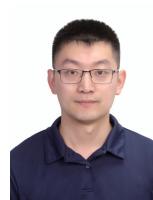
received his B. Sc. degree from Guangxi University of Science and Technology in 2022. Now he is a graduate student at Liaoning Technical University. His main research interests include computer vision, open-pit mining theory and technology.



**范忠胜**,2004年于辽宁工程技术大学获得学士学位,现为山西忻州神达梁家碛煤业有限公司高管,主要研究方向为露天开采理论与技术。

E-mail:2572442156@qq.com

**Fan Zhongsheng** received his B. Sc. degree from Liaoning Technical University in 2004. Now he is a senior executive in Shanxi Xinzhou Shenda Liangjiaqi Coal Industry Co., Ltd. His main research interests include open-pit mining theory and technology.



**付恩三**,2012年于辽宁工程技术大学大学获得学士学位,2015年于辽宁工程技术大学获得硕士学位,2025年于辽宁工程技术大学获得博士学位,现为应急管理部信息研究院高级工程师。主要研究方向为矿山智能化、矿山监测预警以及应急管理等。

E-mail:429001709@qq.com

**Fu Ensan** received his B. Sc., M. Sc., and Ph. D. degrees from Liaoning Technical University in 2012, 2015, and 2025, respectively. Now he is a senior engineer in Information Research Institute of the Ministry of Emergency Management. His main research interests include mine intellectualization, mine monitoring and early warning, emergency management, etc.