基于迁移学习和特征注意增强型PP_LiteSeg的矿井输送带区域分割算法

王媛彬[1,2],常文健[1,2],陈晓晶[3],王　旭[4],刘　佳[1,2]
(1. 西安科技大学 电气与控制工程学院,陕西 西安　710054;2. 西安市电气设备状态监测与供电安全重点实验室,陕西 西安　710054;3. 中煤科工集团常州研究院有限公司,江苏 苏州　213015;4. 兖矿能源集团,山东 邹城　237500)

摘要: 为了实现对矿井输送带区域违规交互行为的精准检测,针对井下图像样本稀缺、目标多尺度及边缘模糊等问题,提出一种融合迁移学习与特征注意力增强机制的改进PP_LiteSeg语义分割算法.设计了特征注意金字塔池化模块融合多尺度特征,增强对不同尺度目标细节的感知能力;构建轻量型多层次统一注意融合解码网络(LMUAFD),提升复杂背景下的目标判别能力;引入迁移学习预训练权重,优化模型在输送带数据集上的收敛速度.结果表明:该算法在输送带区域数据集上的准确率、平均交并比、相似系数及卡帕系数分别达到99.2%,90.4%,94.8%和95.4%,相较于主流语义分割网络Deeplabv3,Unet,EfficientNet,ShuffleNet,MobileNetV3以及原始PP_LiteSeg均有明显提升.该研究为实现矿井输送带区域违规行为的精确检测与智能安全监管提供了技术支撑.

关键词: 语义分割;PP_LiteSeg;特征注意金字塔;LMUAFD;迁移学习
中图分类号: TD 76;TP 391.41　**文献标志码:** A　**文章编号:** 1000-1964(2026)02-0435-12

Region segmentation algorithm for mine conveyor belt based on transfer learning and feature attention enhanced PP_LiteSeg

WANG Yuanbin[1,2], CHANG Wenjian[1,2], CHEN Xiaojing[3], WANG Xu[4], LIU Jia[1,2]
(1. College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an, Shaanxi 710054, China; 2. Xi'an Key Laboratory of Electrical Equipment Condition Monitoring and Power Supply Safety, Xi'an, Shaanxi 710054, China; 3. China Coal Research Institute Changzhou Automatization Research Institute, Suzhou, Jiangsu 213015, China; 4. Yankuang Energy Group Co., Ltd., Zoucheng, Shandong 237500, China)

Abstract: To achieve accurate detection of unauthorized interactions in mine conveyor belt areas, an improved PP_LiteSeg semantic segmentation algorithm integrating transfer learning and a feature attention enhancement mechanism was proposed to tackle the challenges of sample scarcity, multi-scale targets, and blurred edges in underground images. A Feature Attention Pyramid Pooling module was designed to fuse multi-scale features, thereby enhancing the perception of target details across various scales. A lightweight multi-level unified attention fusion decoder network was

收稿日期: 2025-09-23　**修回日期:** 2025-12-25　**DOI:** 10.13247/j.cnki.jcumt.20250385
基金项目: 国家自然科学基金项目(52174198);陕西省重点研发计划项目(2023YBSF-133);西安市科技计划项目(24GXFW0049)
通信作者: 王媛彬(1977—),女,河南省平顶山市人,副教授,硕士生导师.E-mail: wangyb998@163.com
引用格式: 王媛彬,常文健,陈晓晶,等.基于迁移学习和特征注意增强型PP_LiteSeg的矿井输送带区域分割算法[J].中国矿业大学学报,2026,55(2):435-446.
WANG Yuanbin, CHANG Wenjian, CHEN Xiaojing, et al. Region segmentation algorithm for mine conveyor belt based on transfer learning and feature attention enhanced PP_LiteSeg[J]. Journal of China University of Mining & Technology, 2026, 55(2): 435-446.

constructed to strengthen the discriminative ability between targets and complex backgrounds. Additionally, a transfer learning strategy using pre-trained weights was introduced to accelerate convergence on the conveyor belt dataset. The results show that the proposed algorithm achieves an accuracy of 99.2%, a mean intersection over union of 90.4%, a Dice similarity coefficient of 94.8%, and a kappa coefficient of 95.4%. These metrics show significant improvements compared with mainstream networks, including DeepLabv3, U-Net, EfficientNet, ShuffleNet, MobileNetV3, and the original PP_LiteSeg. This study provides high-precision technical support for unauthorized behavior detection and intelligent safety supervision in mine conveyor belt zones.

Key words: semantic segmentation; PP_LiteSeg; feature attention pyramid pooling; LMUAFD; transfer learning

煤矿开采作为能源供给的重要支柱,其高危属性长期受到社会关注[1]. 据统计,2016年至2022年,我国煤矿安全事故总量逾千起[2],其中运输类事故占比高达27%,成为发生率较高的井下事故类型之一. 研究表明,矿工与输送带、防护栏杆等区域关键设施之间的违规交互是引发事故的主要诱因[3]. 尽管井下安全规程已对设备操作和矿工行为进行明确约束,但受限于监管存在盲区、人员流动性大及个体安全意识薄弱等多重因素[4],传统人工巡查和监控视频抽查难以实现高危行为的实时预警与精准干预,亟需智能化技术手段提升风险预控能力[5].

近年来,计算机视觉技术的迅猛发展为井下安全监管智能化提供了新路径. 语义分割[6]通过像素级语义解析可精确勾勒输送带、防护栏杆等关键设备的轮廓细节,并实时量化矿工肢体与设备的空间关系. 这种细粒度的环境目标识别能力,为构建矿工、设备交互行为动态分析模型提供了关键技术支撑,有望突破传统监控系统对不安全行为[7]识别滞后、误判率高的局限性.

语义分割主要分为基于传统的方法以及基于深度学习的方法,传统方法主要有基于区域的语义分割方法[8]、基于图论的语义分割方法[9]以及基于马科夫随机场和条件随机场的语义分割方法[10]. 传统方法依赖于人工设计的特征提取器,这些特征通常难以充分捕捉图像的复杂信息,从而限制了分割精度. 在目标边界模糊的井下输送带区域场景中,传统方法的分割效果往往表现不佳.

随着深度学习在计算机视觉领域的不断发展,基于深度神经网络的语义分割算法逐渐成为研究热点[11]. 通过端到端的训练机制,实现了特征提取与像素分类的一体化学习,显著提高了图像的分割精度. 全卷积网络(Fully Convolutional Networks for Semantic Segmentation, FCN)[12]首次将深度学习应用到图像语义分割领域,该网络利用反卷积层对特征图上采样,恢复输入尺寸并保留空间信息,将图像分类从整体扩展到每个像素的级别. 文献[13]提出了一系列新型扩张卷积语义分割模型Deeplabv3,该类模型能够在保持参数量不变的同时有效扩大感受野,提高模型对不同尺度目标的分割效果. 文献[14]提出了Unet模型,该模型添加了收缩路径,用于提取上下文信息,同时构建了对称的扩展路径,以确保模型对边界分割的精确性. 文献[15]提出了新的语义分割算法,利用CNN和Efficient Transformer构建双编码器,解耦上下文信息和空间信息,提出特征融合模块加强编码器间的信息交互,有效融合全局上下文信息和局部细节信息. 然而,以上方法通过增加网络深度来提升精度,导致模型复杂度显著增加. 为了满足移动设备的部署需求,轻量化实时语义分割模型[16]应运而生. 文献[17]提出了Shufflenet模型,并利用通道混洗和逐点分组卷积对模型进行改进,在降低模型参数的同时保证分割精度. PaddlePaddle团队开发了轻量级图像语义分割算法PP_LiteSeg[18],利用柔性解码网络和简易金字塔实现分割精度与计算复杂度的平衡. 以上研究为轻量化语义分割技术的发展提供了重要的理论支撑和技术路径.

近年来语义分割技术有着较好的发展,但是在煤矿井下场景应用较少. 文献[19]对传统双边分割架构进行改进. 该研究引入多尺度金字塔注意力机制和通道注意力融合模块优化特征组合,实现井下复杂工况下的轨道识别. 文献[20]提出采用自适应加权和曲线内外区域像素的中值绝对差计算的改进CV模型,实现了煤矿火灾图像的分割. 文献[21]提出了一种轻量化的井下轨道实时分割方法,采用STDC骨干网络降低计算负担,同

时引入通道注意力特征模块，通过建模通道间相关性实现特征优化，显著提升了分割精度．文献[22]设计了基于超像素粒化及同质图像粒聚类的井下作业人员分割方法，实现对井下 4 种场景的人员图像精准分割．文献[23]提出双对齐网络模型．该模型从特征级和像素级两方面减少域间差异，将在合成数据集上训练的语义分割模型迁移到煤矿真实场景中，实现煤矿监控图像语义分割．现有研究虽然一定程度上能够实现井下场景的图像分割，但在输送带区域的应用仍存在显著挑战：在井下输送带区域图像数据获取困难且标注成本高，较少的样本使模型训练效果较差；输送带区域目标形状和尺寸差异大，模型难以捕捉复杂细节，多尺度目标分割表现不足；输送带区域光照条件差，粉尘和雾气干扰严重，导致分割模型难以区分前景目标与背景信息，从而影响模型的分割精度．

针对上述问题，提出基于迁移学习和特征注意增强型 PP_LiteSeg 的输送带区域分割算法，用于分析输送带区域矿工与环境目标的交互情况．首先，构建特征注意金字塔池化模块融合输入特征，提高模型对输送带区域各尺度分割目标细节的关注度．然后，构建多层次轻量型统一注意融合解码网络，使模型能更好地区分输送带区域的目标和背景．最后，使用迁移学习，利用预训练模型知识，提高模型在输送带区域图像数据集上的训练效果．

1　特征注意增强型 PP_LiteSeg 的输送带区域分割算法

1.1　算法框架

本文所提出的语义分割算法整体结构如图 1 所示，主要由编码网络、特征注意金字塔池化模块（Feature Attention Pyramid Pooling，FAPP）以及轻量型多层次统一注意融合解码网络（Lightweight Multi-level Unified Attention Fusion Decoding Network，LMUAFD）部分组成．

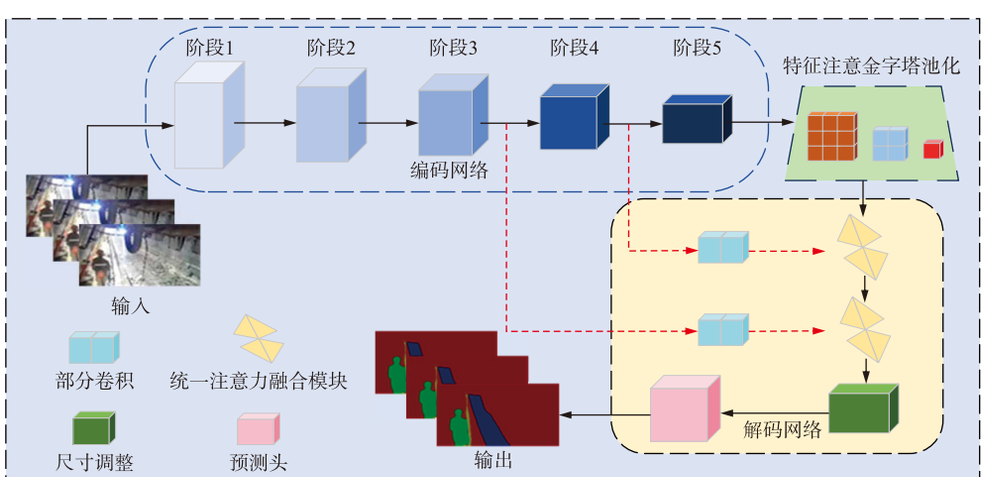


图 1　特征注意增强型 PP_LiteSeg 算法整体结构

Fig. 1　The overall structure of feature attention enhanced PP_LiteSeg algorithm

首先，编码网络负责对输入的输送带区域图像进行逐层特征提取．通过堆叠卷积层与下采样操作，逐步捕获从低层的边缘、纹理等细节信息，到高层的语义、目标结构等抽象特征，为后续分割提供多尺度特征基础．其次，特征注意金字塔池化模块（FAPP）用于增强模型对多尺度目标的感知能力．该模块融合不同尺度的平均池化操作以捕获图像中的全局上下文信息，并引入无参数注意力机制（Parameter-Free Average Attention，PFAA）[24]，提升模型对关键区域的关注度，适用于输送带区域中物体尺寸变化大、背景复杂的情况．最后，轻量型多层次统一注意融合解码网络（LMUAFD）将编码阶段提取的低分辨率高语义特征逐步恢复至原始图像的高分辨率．该模块通过引入部分卷积（Partial Convolution，PConv）[25]与统一注意力融合模块（Unified Attention Fusion Module，UAFM）[26]，实现不同层级特征之间的高效融合，有效提升边缘区域与小目标的分割精度，同时兼顾模型的轻量化设计，便于部署于资源受限的工业设备中．

1.2　编码网络

本文采用短期密集级联网络（Short Term Dense Connection 2，STDC2）[27]的主干作为编码网络对输送带区域图像进行特征提取，高效捕获输送带区域图像的空间和语义特征．STDC2 结构

如图2所示.STDC2主干网络由5个阶段构成,其中阶段1与阶段2主要由卷积层(ConvX1,ConvX2)、批归一化(BN)和ReLU激活函数组成,负责提取图像的浅层细节特征.阶段3至阶段5由多个短期密集级联模块构成,用于提取更深层的语义信息.短期密集拼接模块(Short-Term Dense Concatenate Module,STDCM)具有2种不同步长的结构,具体如图3所示.图3a展示了步长为1的STDCM-A.输入特征图先经过一个1×1的卷积以生成通道数为$N/2$的特征图,随后依次通过3个3×3卷积,分别输出通道数为$N/4$和2个$N/8$的特征图.最终通过短接连接将各分支特征图在通道维度拼接,得到输出通道数为N的融合特征图.该结构在浅层阶段保留更多通道,有助于细节与边缘信息的表达.图3b为步长为2的STDCM-B.该结构在首个3×3卷积中进行下采样,同时对首分支特征图应用平均池化(AVG Pool)以保持尺寸一致,便于后续的通道拼接.该模块通过通道压缩与下采样操作,提升了模型对语义信息的提取能力,减少冗余的同时聚焦关键区域.

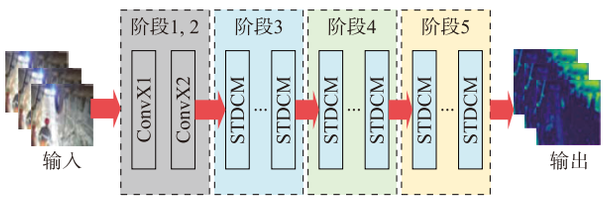


图2 STDC2通用主干网络架构

Fig. 2 Architecture of the STDC2 general-purpose backbone

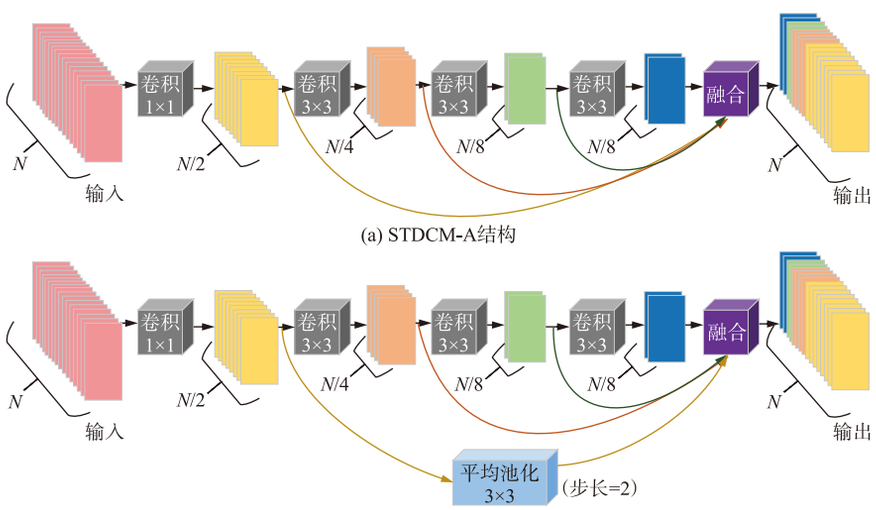


图3 短期密集级联块结构

Fig. 3 Structure of the short term dense cascade block

STDCM-B模块通过通道压缩与下采样的协同设计,从以下方面提升语义提取能力并聚焦关键区域:首先通过下采样操作,直接扩大模型感受野,使模块能捕捉更大范围的上下文信息,语义信息依赖全局上下文,扩大感受野后,模型可同时覆盖输送带、矿工及防护栏杆不同尺度的语义关联,避免局部特征导致的语义割裂.其次采用通道压缩操作,由于井下图像存在大量背景噪声,冗余通道会稀释语义特征权重,通道压缩可保留与目标语义强相关的通道,剔除噪声通道,间接实现关键区域的特征聚焦.最后使用平均池化:对首分支特征图应用平均池化以保持尺寸一致性,过程中进一步聚合局部特征和全局特征,强化语义关联性.

STDC2结构如表1所示.阶段1,2均由一个ConvX3×3模块所构成,阶段3由1个步长为2的STDCM-B和3个步长为1的STDCM-A组成,阶段4包括1个步长为2的STDCM-B和4个步长为1的STDCM-A,阶段5由1个步长为2的STDCM-B和2个步长为1的STDCM-A构成.

表1 STDC2组成结构

Table 1 STDC2 composition structure

阶段	模块	数量	步长
阶段1	ConvX3×3	1	2
阶段2	ConvX3×3	1	2
阶段3	STDCM-B	1	2
	STDCM-A	3	1
阶段4	STDCM-B	1	2
	STDCM-A	4	1
阶段5	STDCM-B	1	2
	STDCM-A	2	1

1.3　特征注意金字塔池化

井下输送带区域中的目标不仅形状复杂、尺度变化显著，且常伴随遮挡与背景干扰，这会给语义分割模型带来较大挑战．而金字塔池化作为一种经典的多尺度特征提取方法，能够有效增强模型对全局上下文的理解．传统金字塔池化结构如图 4 所示，通过对输入特征图进行不同尺度的池化（POOL）操作生成多组特征图，随后统一上采样（UPSAMPLE）至原始尺寸，并在通道维度进行拼接融合，获得包含丰富语义信息的复合特征图．然而，该方法存在 2 个主要问题：一是多尺度池化带来的计算负担较大，影响模型推理速度；二是池化过程可能造成细节信息丢失，尤其对小目标和边缘区域的分割效果不足．

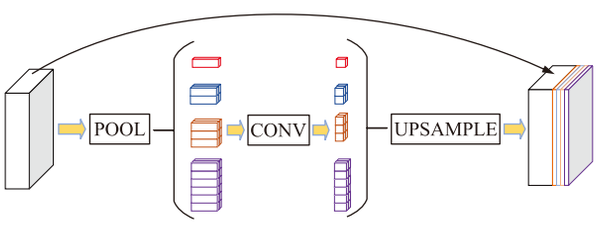


图 4　传统金字塔结构

Fig. 4　Structure of the traditional pyramid

为此，本文提出了特征注意金字塔池化模块（FAPP），其结构如图 5 所示．该模块在设计上采用轻量化的金字塔池化策略，减少多尺度分支数量，以降低整体计算复杂度．同时，引入无参数平均注意力机制（PFAA）以增强模型对关键区域与细节特征的关注，有效弥补传统金字塔结构在小目标识别与边缘感知方面的不足．FAPP 模块在保持高效性的同时，显著提升了模型在复杂场景中的分割精度与鲁棒性．

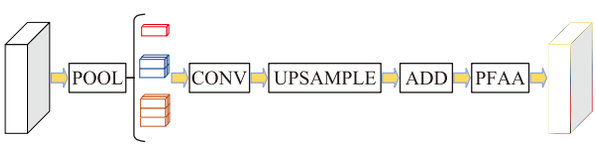


图 5　特征注意金字塔池化结构

Fig. 5　Feature attention pyramid pooling structure

针对输送带区域目标多尺度特点，FAPP 模块采用 3 个不同尺度的全局平均池化操作，池化窗口分别为 1×1，2×2 和 4×4，1×1 池化聚焦小目标的细节特征，4×4 池化捕捉大目标的全局上下文，用以捕获不同感受野下的上下文信息．各分支输出经 1×1 卷积进行通道压缩，以降低计算量．随后所有特征图上采样至输入尺度，并采用逐元素相加（ADD）方式完成多尺度特征融合，最终通过 PFAA 模块生成特征表示．PFAA 结构如图 6 所示，其中 H，W，C 分别表示特征图的高、宽以及通道数量，该模块分别在空间维度和通道维度上执行无参数平均池化．当输送带区域特征图输入该模块时，首先通过平均池化（Average Pooling）[28] 在空间方向和通道方向分别得到一个 $H\times W\times 1$ 和 $1\times 1\times C$ 大小的特征图，实现空间特征图与通道特征图的分离．随后，通过扩展函数（Expand）将空间注意图广播至通道维度，引导模型聚焦于图像中的细节区域；同时将通道注意图扩展至空间尺度，强化关键语义通道的表达能力．最终，将两部分注意图融合，并经 Sigmoid 函数激活后与原始特征图逐元素相乘（Elem wise multiply），生成细节感知增强的特征表示．

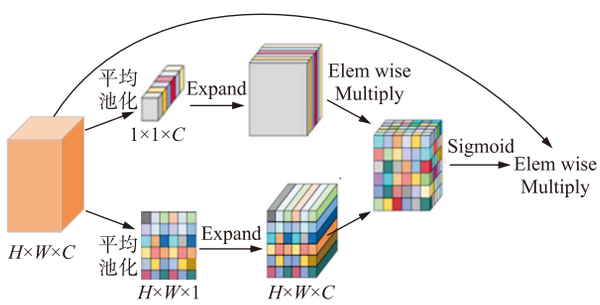


图 6　PFAA 结构

Fig. 6　PFAA structure

相较于传统金字塔池化模块，FAPP 在结构与计算效率方面进行了多项优化．首先，通过压缩中间与输出特征的通道数，有效降低了计算复杂度；其次，移除原有短路连接结构，简化了特征流路径；最后，将通道拼接操作替换为逐元素相加，进一步减少参数量与内存开销．这些改进显著提升了模块的运行效率，更契合输送带区域分割对轻量化的实际需求．此外，FAPP 融入无参数的 PFAA 模块，PFAA 注意力机制通过空间与通道双维度无参数池化，可自动聚焦井下图像的关键区域：例如粉尘覆盖的输送带区域，空间注意力会强化未被覆盖的输送带纹理，通道注意力会突出输送带颜色，避免粉尘区域对目标分割的干扰，增强了模型对关键细节区域的关注能力，提升了整体分割精度与鲁棒性．

1.4　轻量型多层次统一注意融合解码网络

井下输送带区域光照条件匮乏，环境昏暗，在复杂且充满干扰的输送带区域环境中，模型难以精确区分输送带区域的前景目标与错综复杂的背景

信息，导致模型对输送带区域的分割效果不理想．因此，本文构建了轻量型多层次统一注意融合的解码网络(LMUAFD)，旨在解决井下低照度和粉尘导致前景、背景边界模糊的问题，其结构如图 7 所示．

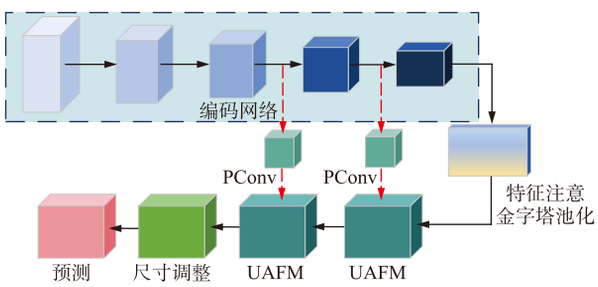


图 7 轻量型多层次统一注意融合的解码网络

Fig. 7 Lightweight decoding network with multi-level unified attention fusion

LMUAFD 首先接受由特征注意金字塔池化处理后的高层次特征图，该特征图包含丰富的语义信息，有利于模型区分输送带区域的前景目标与背景信息．然后利用部分卷积(PConv)对包含更多井下细节信息的较低层次语义信息进行初步特征提取，再利用统一注意融合模块(UAFM)将高层次特征图和低层次特征图进行融合，最后将融合后的特征图上采样后进行预测，得到输送带区域的分割结果．

由于传统卷积易将模糊边界的像素误判为背景．所以采用 PConv 将其代替．PConv 通过动态忽略缺失像素，仅对有效目标像素进行卷积，可减少粉尘噪声对边界特征的污染，PConv 会抑制噪声像素的权重，保留边缘的连续特征，避免分割结果出现边缘锯齿．PConv 的结构如图 8 所示，该结构通过减少冗余计算和内存访问，能够更高效地提取空间特征，从而提升计算效率和模型性能．

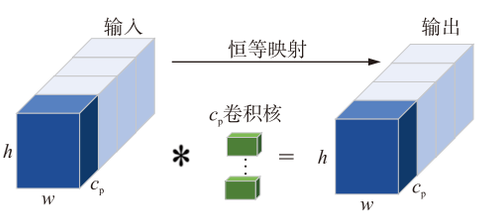


图 8 PConv 结构

Fig. 8 PConv structure

常规卷积和 PConv 的浮点运算次数(F_{normal}，F_{pconv})计算公式如公式(1)，(2)所示．

$$F_{normal}=h\times w\times k^2\times c^2/s, \tag{1}$$

$$F_{pconv}=h\times w\times k^2\times c_p^2, \tag{2}$$

式中：h 和 w 分别为特征图的高度和宽度；k 为卷积核尺寸；c 为输入特征图的通道数量；s 为卷积步长；c_p 为 PConv 中经过通道压缩后的输入特征图通道数量．当压缩比率 $r=c_p/c=1/4$ 时，PConv 的计算量仅为标准卷积操作的 1/16．

因此，采用 PConv 对低层次特征图进行特征提取，可以在少量参数增加的情况下获得低层次特征图所包含的输送带区域细节信息．

为了充分整合高层次与低层次特征图中的视觉信息，本文采用统一注意力融合模块(UAFM)实现多层次特征的有效融合，其结构如图 9 所示．

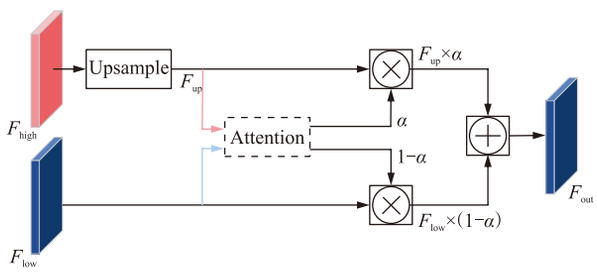


图 9 统一注意融合模块结构

Fig. 9 Structure of the unified attention fusion module

当高层次特征输入 UAFM 时会先采用双线性插值法[29]对其进行上采样(Upsample)，使其与低层次特征的大小一致．过程如公式(3)所示．

$$F_{up}=\text{Upsample}(F_{high}), \tag{3}$$

式中：F_{high} 表示高层次特征；F_{up} 表示上采样后的高层次特征．然后将 F_{up} 和低层次特征 F_{low} 送入空间注意力模块(Attention)以生成相应的特征权重 α，其过程如公式(4)所示．

$$\alpha=\text{Attention}(F_{up},F_{low}), \tag{4}$$

空间注意力模块结构如图 10 所示．

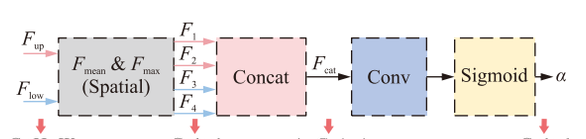


图 10 空间注意力结构

Fig. 10 Structure of the spatial attention module

当大小为 $C\times H\times W$ 的 F_{low} 和 F_{up} 输入空间注意力(Spatial)时，首先，沿着通道方向分别计算 F_{low} 和 F_{up} 的均值 F_{mean} 和最大值 F_{max}，生成 F_1，F_2，F_3，F_4 以上 4 个特征，这些特征的维度为 $C\times1\times1$．然后，将这 4 个特征在通道维度进行拼接(Concat)，生成大小为 $4\times C\times1\times1$ 的连接特征 F_{cat}，最后通过卷积和 Sigmoid 函数得到权重 α．为了获得 F_{low} 和 F_{up} 的注意力加权特征，分别对 F_{low} 和 F_{up} 与 $1-\alpha$

和α进行逐元素乘法操作，然后再对注意力加权特征进行逐元素加法操作，具体过程如公式(5)所示.

$$F_{out}=\alpha\odot F_{up}+(1-\alpha)\odot F_{low}, \tag{5}$$

式中：F_{out}表示最后的输出注意力加权融合特征. UAFM通过空间注意力生成F_{low}和F_{up}的注意力权重，使得F_{high}中的输送带区域目标信息和F_{low}中的细节信息在F_{out}中更加显著，提高模型对输送带区域前景目标和背景信息的区分能力.

1.5　迁移学习训练步骤

迁移学习[30]的核心在于将大规模数据集上预训练得到的模型参数作为目标任务的初始化权重，从而在数据有限、计算资源受限的条件下实现高效建模. 由于井下输送带区域图像获取困难且标注成本较高，直接使用小规模数据进行训练容易导致模型性能不稳定. 通过迁移学习策略，能够有效提升模型的收敛速度与泛化能力.

本研究采用两阶段的迁移学习训练流程，如图11所示. 第一阶段，使用Cityscapes[31]、ADE20K[32]等主流公开语义分割数据集对PP_LiteSeg模型进行预训练，以学习通用的图像语义表征并生成预训练权重. 这些数据集包含道路、行人等，与输送带、矿工具有相似的目标类型，预训练模型可复用边缘提取、类别区分等基础能力，避免模型从零学习导致的收敛缓慢；第二阶段，在第一阶段模型基础上，对金字塔池化模块与解码网络结构进行优化. 训练过程中，编码网络参数直接加载预训练权重进行初始化，以保留通用语义信息；改进后的金字塔池化模块与解码分支则采用随机初始化，并在井下输送带区域数据集上进行端到端训练与微调. 该训练策略在利用大规模先验知识的同时，结合目标场景的特征进行局部优化，有效提升了模型在井下输送带场景下的分割性能.

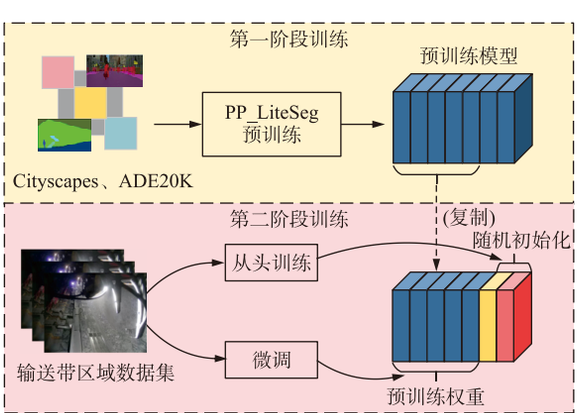


图11　迁移学习训练步骤

Fig. 11　Transfer learning training steps

尽管主流公开语义分割数据集与矿井下输送带区域数据集差异较大，会导致一定的负迁移风险，以及迁移初期模型分割精度骤降. 但通过微调阶段大量学习低质量图像的特征增强逻辑可以得到改善. 并且预训练模型的低级特征提取能力仍可复用，为微调提供高质量初始化权重，避免模型从零学习导致的收敛缓慢. 这样既复用了预训练模型的基础特征提取能力，又通过局部微调快速适配输送带场景，降低了领域差异导致的负迁移风险.

2　试验结果与分析

2.1　试验设置与评价指标

软硬件试验环境：处理器为AMD Ryzen 76800H；16 GB系统内存；操作系统为windows10；显卡为NVIDIA RTX 3060，6 GB显存容量；CUDA(统一计算设备架构)为11.3；深度学习框架为Pytorch1.13；编程语言为Python3.9.

本文基于Paddle框架搭建井下输送带区域语义分割算法，采用SGD优化器，初始学习率设置为0.01，采用多项式衰减对学习率进行调度，batch_size(批尺寸)设置为4，迭代10 000次. 本文所用数据集为自制输送带区域图像分割数据集，以DSLMF+(综采工作面异常状态公开数据集)为基础构建. DSLMF+原始数据覆盖综采工作面全区域，本文针对输送带区域分割任务进行定向筛选，保留输送带在图像中占比≥30%的样本，确保目标区域为图像核心内容，剔除仅含局部设备或纯背景的图像，聚焦输送带运行相关场景，剔除与输送带无关的掘进、支护等场景图像，共包含600张图像，用Labelme工具进行二次标注，其中包括待分割的矿工、防护栏杆和输送带目标. 将标注好的数据集按8∶1∶1化分为训练集、验证集和测试集，用于输送带区域分割模型的训练和验证.

本文采用平均交并比(mIoU)、准确率(Accuracy)、相似系数(Dice)、卡帕系数(Kappa)和F1作为评价指标对模型性能进行评估. mIoU计算过程如公式(6)，Accuracy计算过程如公式(7).

$$\mathrm{mIoU}=\frac{1}{k+1}\sum_{i=0}^{k}\frac{\mathrm{TP}}{\mathrm{FN}+\mathrm{FP}+\mathrm{TP}}, \tag{6}$$

$$\mathrm{Accuracy}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}, \tag{7}$$

式中：k代表目标的类别数；TP表示真正；FN表示

假负;FP表示假正;TN表示真负.

Dice用于衡量2个集合的相似度,取值范围为[0,1],其计算过程如公式(8).

$$\mathrm{Dice}=\frac{2\mathrm{TP}}{2\mathrm{TP}++\mathrm{FP}+\mathrm{FN}}. \tag{8}$$

F1指标是衡量模型分割精度的核心评价指标之一,P为检测精度,R为召回率,F1计算过程如公式(11).

$$\mathrm{F1}=\frac{2\times P\times R}{P+R}. \tag{11}$$

2.2 不同训练条件下的结果对比

图12展示了在不同训练策略下,模型在输送带区域图像分割任务中,评价指标mIoU与Accuracy随迭代次数变化的曲线对比.

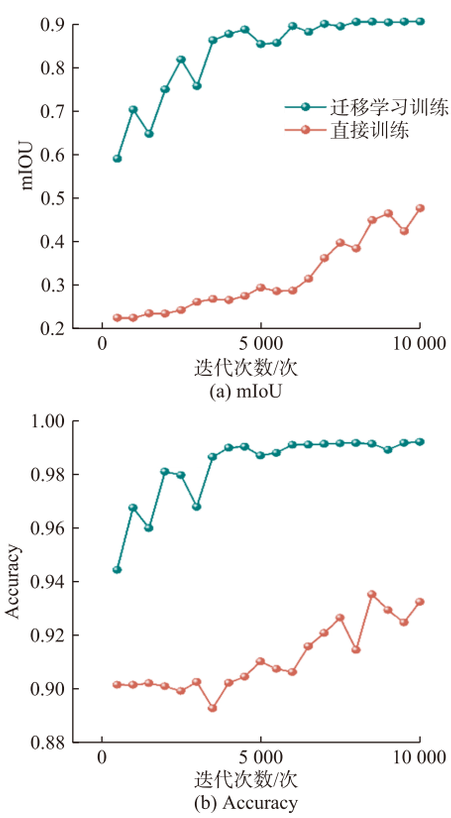


图12 mIoU和Accuracy随迭代次数变化曲线

Fig. 12 Curves of mIoU and Accuracy versus iteration count

由图可见,未使用迁移学习策略时,模型在初始阶段收敛缓慢,且最优精度较低,说明在数据稀缺条件下直接训练难以获得理想效果.采用迁移学习后,模型在较少训练轮次内即可快速提升性能,表明预训练权重已包含目标边缘提取、类别语义关联等通用能力,在输送带数据集上训练时,模型无需从零学习基础特征,仅需优化适配井下场景的参数,如FAPP的注意力权重、LMUAFD的边界融合系数,因此在2 000次迭代时,mIoU已比直接训练高12.9%;5 000次迭代时,迁移学习模型接近稳定,证明预训练过程显著加速了模型收敛过程.

Kappa衡量分类效果,其取值范围为[-1,1].Kappa计算过程如公式(9).其中p_0为Accuracy,p_e为偶然准确度,其计算过程如公式(10).

$$\mathrm{Kappa}=\frac{p_0-p_e}{1-p_e}, \tag{9}$$

$$p_e=\frac{(\mathrm{TP}+\mathrm{FP})\times(\mathrm{TP}+\mathrm{FN})\times(\mathrm{FN}+\mathrm{TN})\times(\mathrm{FP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})^2}. \tag{10}$$

2.3 不同优化策略的消融试验

为验证FAPP和LMUAFD两项优化策略对模型性能的影响,设计消融试验分别对不同的优化策略组合进行验证,其结果如表2所示.从表中数据可以看出,模型B在单独加入FAPP后,四项客观评价指标均有提升,这得益于FAPP模块中加入了PFAA模块,该模块加强了模型对输送带区域不同尺度特征信息的关注度,从而提升模型的分割效果.模型C在单独加入LMUAFD优化策略时,四项优化指标也都得到了提升,这得益于LMUAFD采用PConv对低层次特征图进行了有效的特征提取,同时UAFM可以有效融合高、低层次的特征信息,提高了模型对输送带环境的理解能力.相较于模型A,模型D在两项优化策略同时加入后,四项评价指标均达到最优,分别提升了2.03%,0.41%,1.28%和1.81%,说明两项优化策略可以很好地相互融合,共同提高模型在输送带区域的分割能力.

表2 不同优化策略的消融试验

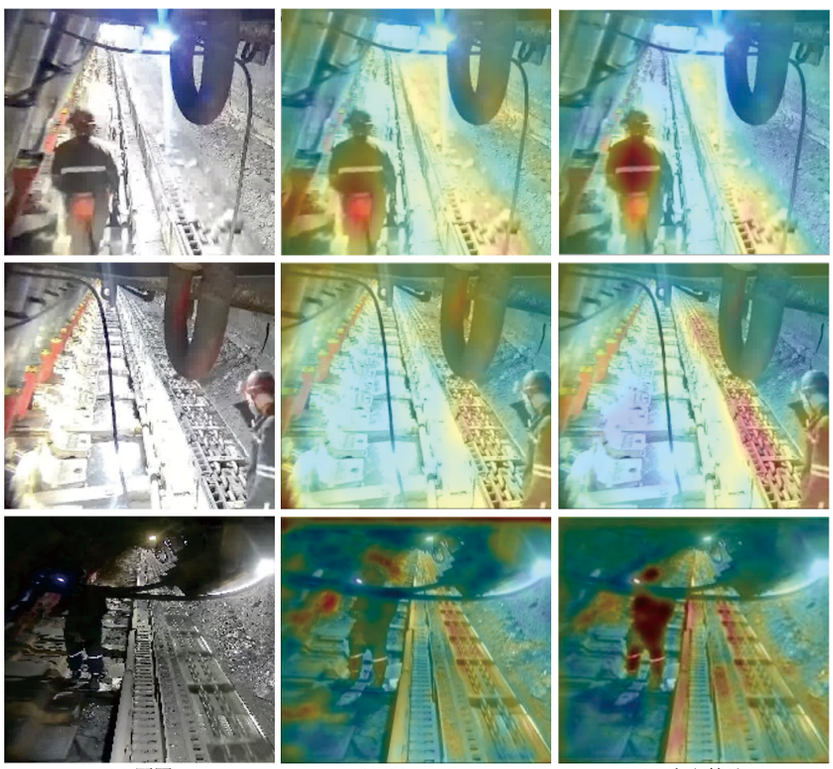
Table 2 Experiments on ablation with different optimization strategies

模型	FAPP	LMUAFD	mIoU	Accuracy	DICE	Kappa
A			0.886	0.988	0.936	0.937
B	√		0.887	0.991	0.947	0.943
C		√	0.891	0.989	0.944	0.943
D	√	√	0.904	0.992	0.948	0.954

2.4 热力图可视化

为验证本文方法较原PP_LiteSeg算法对输送带目标理解能力的优越性,选取3个场景进行热力图可视化分析,其结果如图13所示.从图中可以看出,PP_LiteSeg过多地关注无关的背景信息,而

本文算法对矿工、防护栏杆和输送带等前景目标的关注度更高，这得益于LMUAFD显著提升模型对输送带区域前景目标与背景信息的区分能力，PConv通过对低层次特征图进行高效的特征提取，充分挖掘了细节信息，而UAFM则通过有效整合高层次语义特征与低层次细节特征，增强了模型对输送带区域动态特征的多层次感知能力，从而提升了整体性能.



(a) 原图　(b) PP_LiteSeg　(c) 本文算法

图 13　特征图可视化

Fig. 13　Visualization of feature maps

2.5　训练阶段模型精度指标变化趋势对比

试验数据表明，本研究提出的算法从训练初期就在两项评价指标上持续保持领先优势，并在迭代后期快速收敛至稳定平台. 图 14 展示了改进算法与原始 PP_LiteSeg 在训练过程中 mIoU 和 Accuracy 指标的变化对比. 该性能提升主要源于两方面协同优化：首先，新构建的特征注意金字塔模块有效提升了模型对输送带场景多尺度特征的捕获能力；其次，创新设计的轻量化多层次统一注意融合机制，在解码阶段强化了目标特征的判别性表达. 这两项关键技术通过互补作用，最终实现了输送带复杂场景下多类目标分割精度的显著提升.

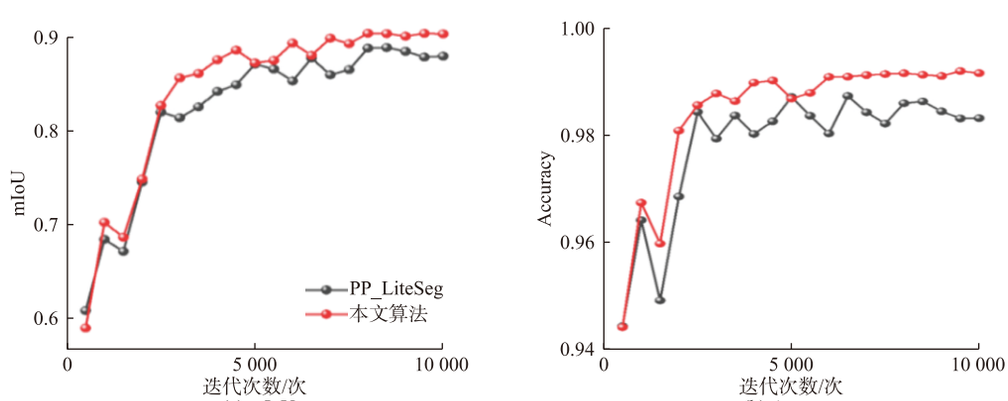


图 14　训练过程中精度变化对比

Fig. 14　Comparison of accuracy changes in training process

2.6 不同算法的对比试验

为进一步验证本文算法的优越性，将本文算法与主流语义分割网络 Deeplabv3，Unet[33]，Efficientnet，Shufflenet，Mobilenet3 以及原始 PP_LiteSeg 算法进行对比，其结果如表 3 所示.

表 3 不同算法对比结果

Table 3 Comparison results of different algorithms

算法	mIoU	Accuracy	Dice	Kappa	F1	参数量
Deeplabv3	0.835	0.957	0.850	0.853	0.845	27.2
Unet	0.843	0.946	0.855	0.815	0.852	25.5
Efficientnet	0.871	0.978	0.856	0.884	0.850	18.7
Shufflenet	0.842	0.986	0.908	0.924	0.905	10.3
Mobilenetv3	0.818	0.983	0.892	0.911	0.889	9.6
PP_LiteSeg	0.886	0.988	0.936	0.937	0.932	14.1
本文算法	0.904	0.992	0.948	0.954	0.947	12.1

从表 3 可以看出，相较于其他对比算法，本文所提算法在前五项指标中均达到了最优，且参数量相比于原模型还略有下降. 在输送带区域分割任务上，本文所提算法的平均精度达到了 99.2%，相较于表中主流分割算法分别提高了 3.66%，4.86%，1.42%，0.61%，0.90% 和 0.41%. 得益于特征注意金字塔池化结构可以提高模型对输送带区域图像细节信息的关注度，同时轻量型多层次统一注意力融合编码网络可以更好地融合高、低层次特征，使模型能更好地区分分割目标和背景信息，有效提高模型对输送带区域目标的分割精度.

具体而言，基础模型 PP_LiteSeg 在轻量化方面具有一定优势，但在复杂背景下表现有限. 通过引入 FAPP 模块，模型在多尺度上下文融合能力方面显著增强，mIoU 和 F1 提升明显，尤其在处理边缘细节与小目标区域时表现更优. 进一步结合无参数注意力机制 PFAA 后，模型对关键区域的响应更加集中，有效抑制了背景干扰，整体分割精度达到较高水平. 同时，本文采用迁移学习策略进行预训练，有效缓解了数据样本较少带来的训练困难，使得模型在较少迭代中快速收敛，进一步提升了整体性能. 综合来看，本文提出的“轻量化主干＋FAPP 多尺度感知＋PFAA 注意力增强＋迁移学习”策略，形成了一种高效、精确、适用于煤矿输送带场景的分割框架，优于表中所列的多个对比方法.

为了进一步验证本文算法在输送带区域分割任务上的有效性，选取 4 个场景进行试验分析，其结果如图 15 示.

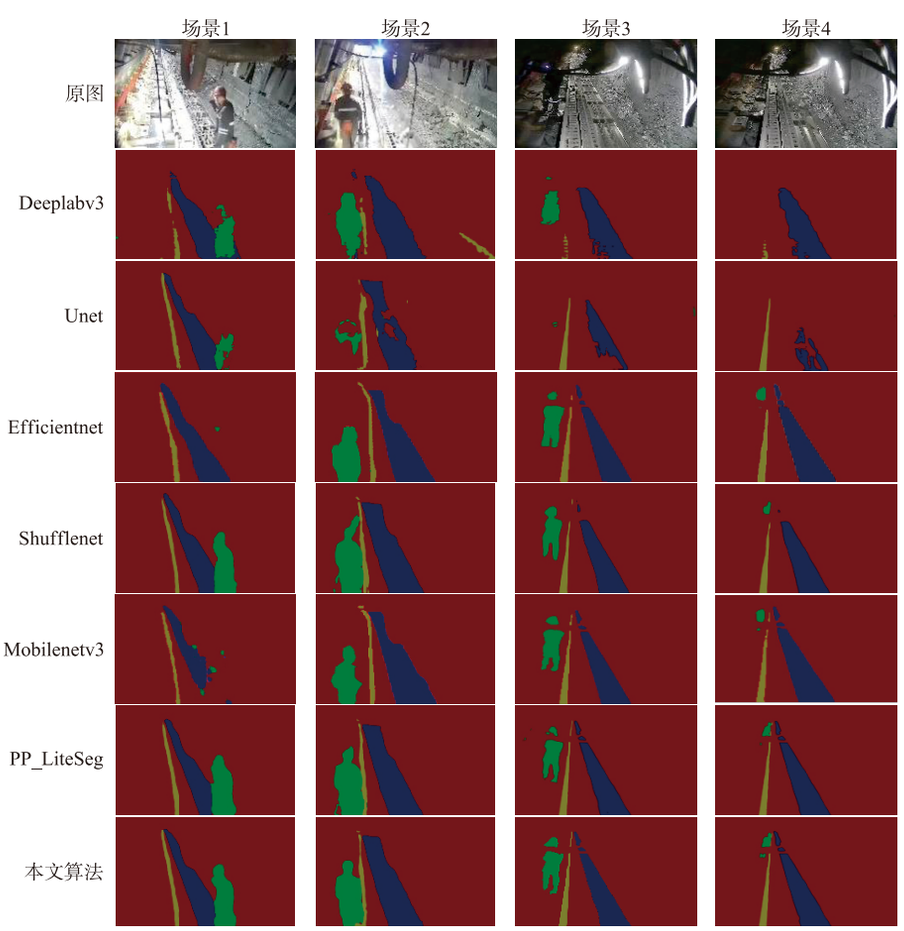


图 15 不同算法的分割结果对比

Fig. 15 Comparison of segmentation results of different algorithms

从图15可以看出，在场景1和2中，Deeplabv3，Unet等算法对矿工边缘和细长栏杆处理模糊，而本文算法得益于特征注意金字塔对细节信息的增强，实现了更精细的边缘勾勒，有效分离了目标与复杂背景．在场景3和4的遮挡与远景任务中，传统轻量化网络如Shufflenet，Mobilenetv3易出现目标漏检或像素误判．本文算法通过多层次统一注意力融合解码网络，强化了对关键特征的提取，对遮挡目标和小目标的分割完整度明显优于对比算法．

3　结　论

1)对输送带区域目标形状和尺寸差异大，模型难以捕捉复杂细节，多尺度目标分割表现不足的问题，构建FAPP模块融合输入特征，提高模型对输送带区域各尺度分割目标细节的关注度，有效提高了多目标分割的精确度．

2)由于输送带区域光照条件差，粉尘和雾气干扰严重，导致分割模型难以区分前景目标与背景信息．为了使模型能更好地区分输送带区域目标和背景，构建了多层次轻量型统一注意融合解码网络，显著提升了模型对复杂背景的区分性能．

3)为了缓解输送带区域图像数据获取困难且标注成本高、较少的样本使模型训练效果较差的问题，本文引入迁移学习策略，在通用目标检测数据集上进行模型预训练，再迁移至输送带区域图像分割任务中，显著提升了模型在样本较少场景下的分割性能．

4)结果表明，在自建井下输送带数据集上所提算法准确率达到了99.2%，平均交并比达到了90.4%，较原始PPLiteSeg模型分别提升2.03%和0.41%，验证了所提方法的有效性与实用性．本文算法为井下输送带区域人机交互行为的实时分析与预防违规行为提供了可靠的方法支撑．

参考文献：

[1] 卞正富，雷少刚，韩晓彤，等．干旱半干旱区露天煤矿关键生态要素异质性及其环境效应[J]．中国矿业大学学报，2025，54(6)：1194-1205．
BIAN Zhengfu，LEI Shaogang，HAN Xiaotong，et al. Heterogeneity of key ecological elements in open-pit coal mines in arid and semi-arid regions and their environmental effects [J]. Journal of China University of Mining & Technology，2025，54(6)：1194-1205.

[2] 张　超，赵瑞军，高　艳，等．煤矿安全生产事故分析及控制策略[J]．内蒙古煤炭经济，2024(5)：103-106．
ZHANG Chao，ZHAO Ruijun，GAO Yan，et al.Analysis and control strategies for coal mine safety production accidents [J]. Inner Mongolia Coal Economy，2024(5)：103-106.

[3] 张洋杰，周　瑶．基于集成DEMATEL-ISM方法的煤矿运输事故致因因素分析与评价[J]．中国煤炭，2024，50(10)：40-47．
ZHANG Yangjie，HOU Yao. Analysis and evaluation of causes of coal mine transportation accidents based on integrated DEMATEL-ISM method [J]. China Coal，2024，50(10)：40-47.

[4] 李瑞群．煤矿员工行为安全管控研究 [J]．中国安全科学学报，2022，32(增2)：26-31．
LI Ruiqun. Research on behavioral safety management of coal mine workers [J]. China Safety Science Journal，2022，32(Sup 2)：26-31

[5] 张盈盈，杨　云，张　景，等．2021—2023年我国煤矿事故统计分析及致因研究[J]．矿业研究与开发，2024，44(11)：224-231．
ZHANG Yingying，YANG Yun，ZHANG Jing，et al. Statisti-cal analysis and cause study of coal mine accidents in China from 2021 to 2023 [J]. Mining Research and Development，2024，44(11)：224-231.

[6] 陈善娟，于云龙，李英明．小样本语义分割研究现状与分析[J]．计算机学报，2024，47(10)：2417-2451．
CHEN Shanjuan，YU Yunlong，LI Yingming. The current status and analysis of semantic segmentation research in small samples[J]. Chinese Journal of Computers，2024，47(10)：2417-2451.

[7] 陈庆峰．矿井皮带区域矿工不安全行为识别方法的研究[D]．徐州：中国矿业大学，2019．
CHEN Qingfeng. Research on the identification method of unsafe behaviors of miners in mining belt areas [D]. Xuzhou：China University of Mining and Technology，2019.

[8] 谢　昭，高　隽．一种基于粗糙集区域分割和语义分类的方法[J]．模式识别与人工智能，2007，20(2)：287-294．
XIE Zhao，GAO Jun. The current status and analysis of semantic segmentation research in small samples[J]. Pattern Recognition and Artificial Intelligence，2007，20(2)：287-294.

[9] 姜　枫，顾　庆，郝慧珍，等．基于内容的图像分割方法综述[J]．软件学报，2017，28(1)：160-183．
JIANG Feng，GU Qing，HAO Huizhen，et al. Overview of content based image segmentation methods [J]. Journal of Software，2017，28(1)：160-183.

[10] 肖春姣，李　宇，张洪群，等．深度融合网结合条件随机场的遥感图像语义分割[J]．遥感学报，2020，24(3)：254-264．

XIAO Chunjiao, LI Yu, ZHANG Hongqun, et al. Deep fusion net-work combined with conditional random field for emantic segmentation of remote sensing images[J]. National Remote Sensing Bulletin, 2020, 24(3):254-264.

[11] 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 Xuan, WANG Liang, DING Qi. Overview of image semantic segmentation methods based on deep learning [J]. Journal of Software, 2019, 30 (2): 440-468.

[12] BENAHMED A, KHEMIS K, ALIM L. A whole 3D liver reconst-ruction for personalised preoperative surgery based on FCN U-net model segmentation [J]. International Journal of Biomedical Engineering and Technology, 2024,44 (2):103-119.

[13] CHEN L C, ZHU Y , PAPANDREOU G, al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//FERRARI V. European Conference on Computer Vision. Cham: Springer,2018:1-18.

[14] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [J]. Medical Image Computing and Computer Assisted Intervention, 2015, 56(9): 234-241.

[15] 张振利,胡新凯,李 凡,等.基于CNN和Efficient Transformer多尺度遥感图像语义分割算法[J].浙江大学学报,2025,59(4):778-786.
ZHANG Zhenli, HU Xinkai, LI Fan, et al. Multi scale remo-te sensing image semantic segmentation algorithm based on CNN and Efficient Transformer [J]. Journal of Zhejiang University, 2025, 59(4): 778-786.

[16] 董荣胜,刘 意,马雨琪,等.轻量化卷积注意力特征融合网络的实时语义分割[J].计算机辅助设计与图形学学报,2023,35(6):935-943.
DONG Rongsheng, LIU Yi, MA Yuqi, et al. Real time semantic segmentation of lightweight convolutional attention feature fusion network[J]. Computer Aided Design, 2023, 35(6): 935-943.

[17] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//ZISSERMAN A. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 6848-6856.

[18] ZHU Q C, JUN Z, YONG H Z, et al. Research on tobacco field semantic segmentation method based on multispectral unmanned aerial vehicle data and improved PP_LiteSeg model[J]. Agronomy, 2024, 14 (7): 1502-1521.

[19] 卫 星,刘邵凡,杨国强,等.基于改进双边分割网络的井下轨道检测算法[J].计算机应用研究,2020,37(增1):348-350.
WEI Xing, LIU Shaofan, YANG Guoqiang, et al. Undergrou-nd orbit detection algorithm based on improved bilateral segmentation network [J]. Application Research of Computers, 2020, 37(Sup 1): 348-350.

[20] 韩 斌,吴一全,宋 昱.基于改进CV模型的煤矿井下早期火灾图像分割[J].煤炭学报,2017,42(6):1620-1627.
HAN Bin, WU Yiquan, SONG Yu. Early fire image segmentation in coal mines based on improved CV model[J]. Coal Journal, 2017, 42(6): 1620-1627.

[21] 马 天,李凡卉,杨嘉怡,等.基于改进STDC的井下轨道区域实时分割方法[J].工矿自动化,2023,49(11):107-114.
MA Tian, LI Fanhui, YANG Jiayi, et al. Real time segment-ation method for underground track area based on improved STDC[J]. Journal of Mine Automation, 2023, 49(11): 107-114.

[22] 李晓宇,杨 维,刘 斌,等.基于超像素粒化与同质图像粒聚类的矿井人员图像分割方法[J].煤炭学报,2021,46(4):1341-1354.
LI Xiaoyu, YANG Wei, LIU Bin, et al. Mining personnel image segmentation method based on superpixel graining andhomogeneous image graining clustering[J]. Coal Journal, 2021, 46(4): 1341-1354.

[23] 杨 潇,陈 伟,任 鹏,等.基于域适应的煤矿环境监控图像语义分割[J].煤炭学报,2021,46(10):3386-3396.
YANG Xiao, CHEN Wei, REN Peng, et al. Semantic segme-ntation of coal mine environmental monitoring images basedon domain adaptation [J]. Coal Journal, 2021, 46(10): 3386-3396.

[24] KORBER N. Parameter-free average attention improves convolutional neural network performance (almost) free of charge[J]. ArXiv, 2022: 1-11.

[25] 郭 艳,王智文,赵润星.YOLO-POD:基于多维注意力机制的高精度PCB微小缺陷检测算法[J].电子学报,2024,52(7):2515-2528.
GUO Yan, WANG Zhiwen, ZHAO Runxing. YOLO-POD: High precision PCB micro defect detection algorithm based on multidimensional attention mechanism[J]. Acta Electronica Sinica, 2024, 52(7): 2515-2528.

(下转第545页)