基于双分支 DB-TransUNet 模型的矿山地物精准识别研究*

姚奇[1,2],李少丹[1,2,3],薛彭帅[2,3],朱小龙[1,4],袁运许[1,4],胡引翠[1,2,3]

(1.河北省地质资源环境监测与保护重点实验室, 河北 石家庄 050022;
2.河北师范大学 地理科学学院, 河北 石家庄 050024;
3.河北省环境变化遥感识别技术创新中心, 河北 石家庄 050024;
4.河北省地质环境监测院, 河北 石家庄 050022)

摘要:针对矿山遥感影像中植被覆盖复杂、地物边界模糊导致传统分割模型地物识别精度不足的问题,提出了一种双分支 DB-TransUNet 模型。该模型创新性地构建了双分支编码结构,并通过自适应加权融合(AWF)模块整合双分支特征,结合通道自适应增强(CA)模块优化解码过程中的特征权重分配,从而实现了关键区域特征的聚焦与背景噪声的有效抑制。试验结果表明,引入 AWF 与 CA 模块,模型的 *mIoU* 分别提升了 4.16%和 78.63%;完整 DB-TransUNet 模型在复杂遥感场景中取得 *mIoU*=0.903 7、*mDice*=0.948 8 的优异分割效果;对比主流网络分割架构,DB-TransUNet 模型各项指标更为优异,在矿山复杂地物的识别与精细分割上精度更高、泛化能力更强。研究验证了多源特征互补与权重融合策略在缓解矿山地物识别受植被干扰问题方面的有效性,为露天矿区的高精度智能监测提供了可靠的技术方案。

关键词:矿山遥感影像;地物识别;特征融合;双分支网络;植被干扰

中图分类号:TD164 **文献标识码**:A **文章编号**:1005-2763(2026)04-0214-08

Study on Precise Identification of Mine Surface Features Based on the Dual-Branch DB-TransUNet Model

YAO Qi[1,2], *LI Shaodan*[1,2,3], *XUE Pengshuai*[2,3], *ZHU Xiaolong*[1,4], *YUAN Yunxu*[1,4], *HU Yincui*[1,2,3]

(1.Hebei Key Laboratory of Geological Resources, Environment Monitoring and Protection, Shijiazhuang, Hebei 050022, China;
2.School of Geographical Sciences, Hebei Normal University, Shijiazhuang, Hebei 050024, China;
3.Hebei Technological Innovation Center for Remote Sensing Identification of Environmental Changes, Shijiazhuang, Hebei 050024, China;
4.Hebei Institute of Geological Environment Monitoring, Shijiazhuang, Hebei 050022, China)

Abstract: To address the issue of insufficient surface feature recognition accuracy in traditional segmentation models caused by complex vegetation cover and blurred surface feature boundaries in mining remote sensing images, a dual-branch DB-TransUNet model was proposed. This model innovatively constructed a dual-branch encoding architecture and integrated features from both branches using an adaptive weighted fusion (AWF) module. By combining this with a channel adaptive enhancement (CA) module to optimize feature weight allocation during the decoding process, it achieved focused detection of features in key regions and effective suppression of background noise. The experimental results show that incorporating the AWF and CA modules improved the *mIoU* of the model by 4.16% and 78.63%, respectively. The full DB-TransUNet model achieves excellent segmentation results of *mIoU*=0.903 7 and *mDice*=0.948 8 in complex remote sensing scenarios. Compared to mainstream network segmentation architectures, the DB-TransUNet model demonstrates superior performance across all indicators, achieving higher accuracy and stronger generalization capabilities in the recognition and detailed segmentation of complex surface features in mining environments. The findings validates the effectiveness of a strategy combining complementary multi-source features with weighted fusion in mitigating vegetation-induced interference in the identification of

* **收稿日期**:2025-05-12

基金项目:自然资源部部省合作项目(2023ZRBSHZ006);河北省地质资源环境监测与保护重点实验室开放课题(JCYKP202310);河北省自然资源厅 2025 年度科技项目

作者简介:姚奇(2001—),男,黑龙江绥化人,硕士研究生,主要研究方向为遥感与矿山应用。E-mail:1518510115@qq.com

通信作者:李少丹(1989—),女,河北衡水人,博士,副教授,主要研究方向为遥感与矿山应用。E-mail:lishaodan@hebtu.edu.cn

mine surface features, thereby providing a reliable technical solution for high-precision intelligent monitoring in open-pit mining areas.
Key words: Mining remote sensing image, Surface feature recognition, Feature fusion, Dual-branch network, Vegetation interference

0 引言

露天矿山作为资源开发的重要载体,其生态环境动态监测与精准管理对协调资源开发与环境保护具有重要战略意义[1]。在长期的开采过程中,露天矿区常伴随地表扰动、地貌变化频繁等现象,给环境恢复与治理工作带来了极大挑战。随着高分辨率遥感技术的快速发展,基于无人机与卫星影像的地物自动化提取已成为矿山监测的核心技术手段[2]。传统矿山地物提取方法主要依赖人工设计特征(如纹理、形状、光谱指数等)与机器学习算法(如支持向量机、随机森林等)相结合[3-5]。例如,康俊锋等[6]通过构建多尺度分割规则集实现稀土矿区地物分类,但该方法依赖先验知识,致使其泛化能力受限;PENG 等[7]基于视觉注意模型改进 GrabCut 算法,虽提升了矿区识别效率,但对高分辨率影像的细节捕捉能力不足,并且在特征表达能力与复杂场景适应性方面存在固有局限,难以满足矿山动态监测的高精度需求。近年来,深度学习技术为遥感影像解译提供了新范式。以 U-Net[8]、DeepLabV3+[9]为代表的卷积神经网络(CNN)通过端到端的特征学习能力,在矿山地物分类任务中展现出显著优势[10]。例如,SAPUTRA 等[11]提出了一种多模态深度学习方法,利用多光谱遥感影像实现了矿区边界的高精度语义分割,有效提升了模型在复杂矿区场景下的分割表现;LIU 等[12]提出了一种结合注意力机制的多通道 CNN 框架,用于矿区地物分类,显著提升了堆场与道路等复杂类别的识别精度;CHEN 等[13]提出了一种基于改进双边分割网络的语义分割模型,用于高分辨率卫星影像中露天矿与周围植被的识别。然而,CNN 的局部感受野特性使其对长距离空间依赖关系的建模能力不足,导致在矿山中光谱相似的地物区域易出现误分割问题。相比之下,Transformer 架构通过自注意力机制能够有效捕捉全局上下文关联[14]。WANG 等[15]将视觉 Transformer 模型(Vision Transformer,ViT)引入遥感地物分类任务,借助自注意力机制显著增强了模型的全局感知能力;CHEN 等[16]提出的 TransUNet 模型则进一步实现了将 CNN 与 Transformer 架构融合,并在 U-Net 编码器中嵌入 ViT 模块,利用 CNN 提取局部纹理特征,同时通过 Transformer 建模远距离空间依赖关系,最终达成局部细节与全局语境的联合建模。然而,在矿山遥感场景中,复垦绿化和自然植被等覆盖现象普遍存在。受可见光波段限制,RGB 影像光谱分辨率较低,使植被与矿山地物在反射特性上存在较大重叠,易引发特征混叠,进而造成语义混淆与误分类现象,这一现象在植被分布密集或覆盖复杂的区域表现更为突出。因此,如何有效抑制植被干扰,提升复杂矿山场景下的语义分割精度,已成为当前遥感智能解译领域亟需突破的关键技术问题。

针对上述挑战,本研究提出一种双分支 DB-TransUNet 模型,其创新性体现在三个方面:(1) 构建双输入分支,主分支用于提取 RGB 三通道影像的全局特征,辅助分支作为主分支的补充,主要用于提取经过归一化植被指数(Normalized Difference Vegetation Index,NDVI)掩膜处理后的影像特征,以减弱植被区域对分割精度的干扰;(2) 设计自适应加权融合(AWF)模块,将双分支中的特征融合,实现多尺度语义信息的优化整合;(3) 在解码阶段引入通道自适应增强(CA)模块,强化矿山地物的关键特征响应。试验结果表明,该模型在植被覆盖条件下具有优越的地物识别能力,为高分辨率遥感影像的矿山智能监测提供了可靠的技术方案。

1 模型及方法

1.1 DB-TransUNet 模型结构

本文提出的 DB-TransUNet 模型旨在提升矿山遥感影像中地物分割的精度,整体结构如图 1 所示。模型主要由双分支编码结构、AWF 模块和 CA 模块构成。其中,双分支编码结构包括主分支与辅助分支两条路径:主分支输入包含完整地物信息的全要素影像,辅助分支输入经 NDVI 掩膜处理后去除植被干扰的影响,以减弱植被区域对分割结果的干扰。AWF 模块用于融合双分支特征,通过动态调整权重并将辅助分支特征按比例补充到主分支特征,有效抑制植被干扰,增强矿山区域的语义响应;CA 模块应用于解码阶段,通过自适应调整通道权重,强化与矿山地物目标高度相关的特征通道,实现多尺度语义信息与边缘结构的高效融合。

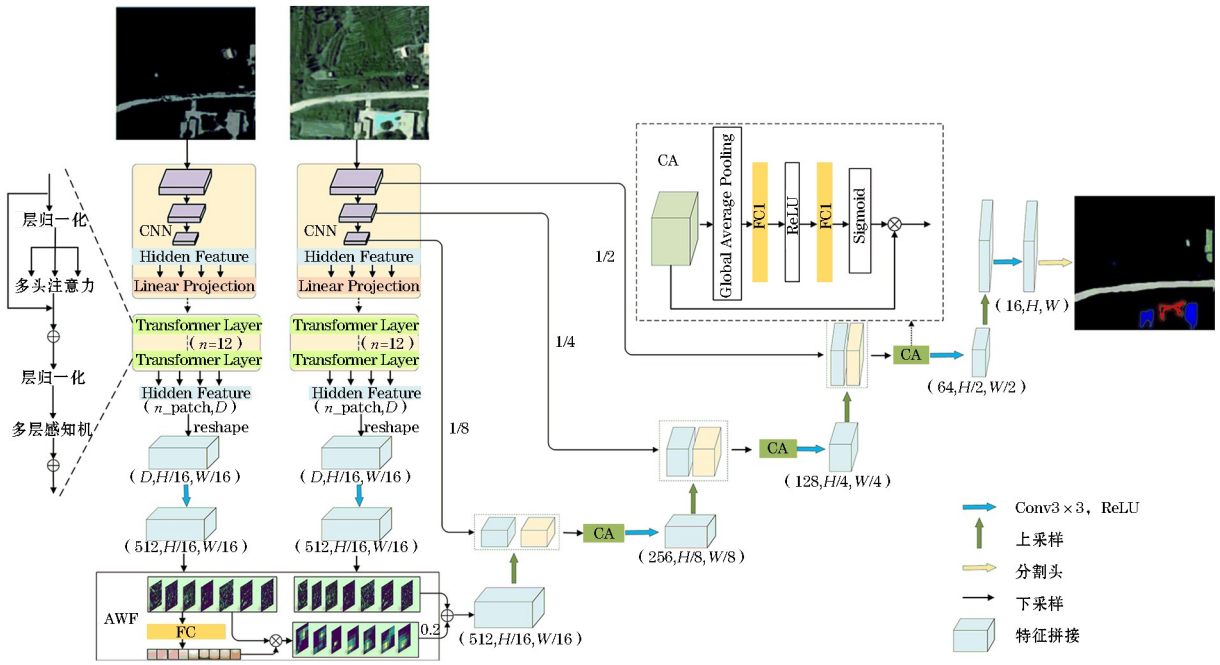


图 1 双分支 DB-TransUNet 结构

Fig.1 Dual-branch DB-TransUNet architecture

1.2 双分支网络

本文所提出的 DB-TransUNet 模型设计了两条并行的信息处理路径,分别接收包含完整地物信息的全要素影像与去除植被干扰的影像。两条路径采用一致的 CNN-ViT 混合结构,首先通过 CNN 模块提取图像的局部细节特征,如纹理、边缘信息,然后利用分块嵌入方式将特征图序列化,输入至 ViT 模块进行全局上下文建模。其中,ViT 部分由多个 Transformer 模块堆叠而成,每个模块包含多头自注意力机制与多层感知机,可有效捕捉长距离依赖关系并生成多尺度语义特征。

为了增强模型对植被干扰的鲁棒性,主分支接收包含完整地物信息的全要素影像作为输入,辅助分支则对影像进行了 NDVI 预处理。通过近红外波段计算 NDVI 生成植被掩膜,从原始 RGB 影像中剔除植被区域,仅保留非植被区域作为模型输入,从而提高矿山地物的显著性。随后,将辅助分支提取的特征与主分支提取的特征在后续阶段进行融合,通过上下文增强机制提升模型对复杂背景中矿区地物的识别能力。

1.3 自适应加权融合模块

AWF 模块是模型的关键部分,负责融合主分支和辅助分支的特征,其核心思想是通过计算辅助分支特征的均值生成加权系数,并利用该系数对辅助分支特征进行加权调整。加权后的辅助分支特征被适量补充到主分支特征中,从而增强模型对矿山区域目标的表征能力。

假设F_{m}为主分支的特征图,F_{a}为辅助分支的特征图,AWF 模块首先计算辅助分支特征图的全局均值μ_{a}:

$$\mu_{\mathrm{a}}=\frac{1}{H\times W\times C}\sum_{i=1}^{H}\sum_{j=1}^{W}\sum_{k=1}^{C}F_{\mathrm{a}(i,j,k)} \tag{1}$$

式中,H、W、C 分别为特征图的高度、宽度和通道数。

随后,将计算得到的特征均值通过一个 Sigmoid 激活函数进行归一化处理,生成加权系数α,该系数的取值范围为[0,1],具体公式如下:

$$\alpha=\sigma(\mu_{\mathrm{a}}) \tag{2}$$

式中,σ 表示 Sigmoid 函数。利用生成的加权系数α对辅助分支F_{a}的特征进行加权调整,得到加权后的辅助分支的特征:

$$F'_{\mathrm{a}}=\alpha\times F_{\mathrm{a}} \tag{3}$$

式中,$\times$ 表示逐元素乘法,通过逐元素乘法将加权系数应用于辅助分支的每个特征值,从而调整辅助分支特征的整体强度。

最后,将主分支的特征图F_{m}与加权后的辅助分支特征图 F'_{a} 进行融合,得到最终的融合特征图F_{f}。具体公式为:

$$F_f = F_m + \beta \times F'_a \quad (4)$$

式中,β 为超参数,用于控制辅助分支在融合过程中的贡献度。经过参数调整和性能评估发现,当 β 取 0.2 时,模型能够达到最佳的分割性能。

1.4 通道自适应增强模块

CA 模块的核心目标是通过自适应地调整通道权重,强化矿山目标的关键特征响应,从而提高模型对复杂矿山场景的分割精度。该模块通过全局信息聚合、特征映射学习、加权融合 3 个关键步骤实现通道特征的自适应增强。

全局信息聚合是对输入的特征图进行全局平均池化并对特征图的空间维度进行压缩,得到每个通道的全局信息表示。其数学表达式如下:

$$z_c = \frac{1}{H \times W} \sum_{h=1}^{H} \sum_{w=1}^{W} X_{h,w,c}, c = 1,2,\cdots,C \quad (5)$$

式中:假设输入特征图为 $X \in \boldsymbol{R}^{H \times W \times C}$;$z_c$ 表示 C 个通道的全局平均值,它能够捕获每个通道在整个图像上的平均响应,为后续的通道权重学习提供全局上下文信息。

特征映射学习是通过全连接层(FC)生成通道注意力权重,用于表示每个通道在当前任务中的重要性。其数学表达式如下:

$$W_c = \sigma(FC(ReLU(FC(z_c)))) \quad (6)$$

式中:FC 是两个全连接层;$ReLU$ 是激活函数,用于引入非线性;W_c 是通道注意力权重,反映了该通道在当前任务中的重要性。通过这种方式,模型能够自适应地学习每个通道的重要性,并为重要的通道分配更高的权重,同时抑制不相关或冗余的通道。

加权融合是将学习到的通道注意力权重乘回特征图,得到加权后的特征图。其数学表达式如下:

$$F_{\text{enhanced}} = F \times W_c \quad (7)$$

式中:F_{enhanced} 为加权后的特征图;F 为原始特征图。通过加权融合,模型能够突出关键通道特征,增强特征图的表达能力,提升对矿山目标的识别精度。

2 试验配置

2.1 数据集介绍及预处理

本研究选取河北省石家庄市鹿泉区 4 个具有代表性的矿山区域作为研究对象,所选矿区涵盖正在开采与生态修复阶段等不同类型的矿山,具备较强的代表性。遥感数据由无人机获取,影像包含红(R)、绿(G)、蓝(B)、红边(Red Edge)和近红外(NIR)5 个波段,具备良好的地物光谱分辨能力,能够支持复杂矿区地表类型识别及生态恢复评估。

为了适配深度学习网络的输入要求,原始矿山影像被裁剪为 1 024×1 024 像素的子图,共获得 366 张符合标准的影像样本。随后,依据植被占比筛选数据,去除植被覆盖比例较高的影像,最终保留 248 张影像作为试验数据。该数据集涵盖矿区中常见的 7 类典型地物,分别为:(1) 废渣(矿石)堆放区,用于存储开采过程中产生的废弃矿石或岩石;(2) 矿区建筑物,主要为矿区内的居住或加工建筑,以玻璃钢板房为主;(3) 露天矿,进行剥离和采矿作业的核心区域;(4) 采矿道路,矿区内部的运输通道;(5) 废弃矿坑,已停止作业的开采区域;(6) 已开采区域,指无明显矿体但具有开采痕迹的区域;(7) 矿区施工设备,主要包括矿区中的罐体等施工设施。

为防止模型在训练过程中出现过拟合现象,提升其鲁棒性和泛化能力,对影像数据及其对应标签进行数据增强处理,包括随机角度旋转、亮度调整、镜像变换等操作,将数据集扩充至 1 238 张影像。最终,按照 8∶1∶1 的比例将影像数据集划分为训练集、验证集和测试集,以验证模型在不同数据集上的泛化性能。

2.2 试验环境及网络配置

试验采用 PyTorch 深度学习框架,编程语言为 Python,并采用 NumPy 和 OpenCV 进行数据处理与图像预处理。试验运行于 Intel Core i5-14600KF CPU 和 NVIDIA RTX 4070 Super(12 GB 显存)GPU 环境,内存 32 GB,硬盘 1 TB。训练过程中,模型迭代 200 次,采用 Adam 优化器,损失函数采用交叉熵损失(Cross Entropy Loss)与 Dice Loss 结合,以提高模型的分割性能。批处理大小设为 8,初始学习率为 0.01,并在每 5 轮迭代后保存一次权重,最终选取最优模型参数进行测试与评估。

2.3 评价指标

采用 5 种常用评价标准来衡量模型的分割性能,分别为平均交并比($mIoU$)、平均 Dice 系数($mDice$)、精确率(P)、召回率(R)和 F_1 系数。这些指标从不同角度反映了模型在矿山地物识别任务中

的表现:$mIoU$ 用于衡量模型预测与真实标签之间的重叠度,反映了模型在不同类别上的分割质量;$mDice$ 通过计算预测和真实标注区域的相似性,提供了一种衡量分割精度的方式;P 与 R 分别评估了模型的精确率和召回率;F_1 系数则为精确率和召回率的调和平均值,用于评估模型的综合性能。通过综合上述 5 个指标,能够全面评估模型在矿山地物识别中的表现。指标计算公式如下。

$$mIoU=\frac{1}{N}\sum_{i=1}^{N}\frac{TP}{TP+FP+FN} \tag{8}$$

$$mDice=\frac{1}{N}\sum_{i=1}^{N}\frac{2\times TP}{2\times TP+FP+FN} \tag{9}$$

$$P=\frac{TP}{TP+FP} \tag{10}$$

$$R=\frac{TP}{TP+FN} \tag{11}$$

$$F_1=2\times\frac{P\times R}{P+R} \tag{12}$$

式中:TP 表示模型将实际为正类的样本正确预测为正类的数量;FP 表示模型将实际为负类的样本错误预测为正类的数量;FN 表示模型将实际为正类的样本错误预测为负类的数量;N 表示类别总数。

3 试验分析

3.1 消融试验

为了验证双分支结构中的 AWF 模块以及 CA 模块在模型中的实际作用,基于矿山数据集设计并开展了消融试验。试验共设置 4 种不同的输入与模块组合,分别为:单分支(去除植被干扰影像)+CA、单分支(全要素影像)+CA、双分支(全要素影像+去除植被干扰影像)+AWF,以及完整的 DB-TransUNet 模型(双分支+AWF+CA)。所有试验采用相同的训练条件,确保结果具有可比性与可靠性。

图 2 展示了 4 种模块配置下的模型分割效果,可以直观对比各模块的作用及不同配置对分割性能的影响。图 2(c)为“单分支(去除植被干扰影像)+CA”配置,该模型利用去植被处理减弱了部分背景干扰,在地物提取方面表现较好,但由于缺乏全要素信息支撑,模型整体特征表达能力有限,存在一定程度的误判,特别是在地物与植被交界区域,分割边界模糊,细节不够清晰。图 2(d)为“单分支(全要素影像)+CA”配置,该模型在整体地物结构的识别方面表现优异,能够有效提取主要地物轮廓。然而,由于未对植被区域进行显式剔除处理,仍存在将植被误识别为目标地物的问题,尤其在背景复杂或纹理特征与目标相似的区域,易产生较为明显的假阳性分割,影响分割结果的准确性与可靠性。图 2(e)为“双分支(全要素影像+去除植被干扰影像)+AWF”配置,该模型引入了双分支输入,有助于增强特征表达,但缺乏通道层面的显著性建模,导致部分矿区地物边缘模糊,背景干扰较为严重,尤其在高混杂区域出现了较多误判和漏分,整体分割效果不理想。图 2(f)为本文提出的完整 DB-TransUNet 模型,即“双分支(全要素影像+去除植被干扰影像)+AWF+CA”配置,该模型通过 AWF 模块有效融合多源特征,并利用 CA 强化关键通道响应,显著抑制背景噪声。在矿山地物识别方面,尤其在植被交界区域和边缘细节上,该模型分割边界清晰,目标轮廓完整,表现出最优的分割效果。相比其他模型,该配置有效降低了误判、漏分和边界模糊等问题,体现了各模块协同优化的优势。

表 1 为 4 种不同模型配置在矿山地物分割任务中的消融试验结果。对比引入 CA 模块的两种单分支结构,当输入为去除植被干扰影像时(单分支+CA),模型的 $mIoU$ 为 0.825 0,$mDice$ 为 0.902 5;当输入为全要素影像时,$mIoU$ 提升至 0.867 6,$mDice$ 为 0.927 6。结果表明,全要素影像有助于增强模型对地物整体结构的表达能力,从而提高了模型分割精度。相比之下,去除植被干扰影像虽然有效抑制了背景干扰,但可能伴随部分关键信息丢失,从而影响模型对细节和边界的精确捕捉。在完整的 DB-TransUNet 模型中,$mIoU$ 达到 0.903 7,$mDice$ 为 0.948 8,相比上述两种单分支+CA 配置,完整模型在 $mIoU$ 指标上分别提升了 9.54%和 4.16%。这一结果充分表明,AWF 模块在双分支特征融合中能够有效弥补单一输入所导致的信息缺失,实现多源特征的互补增强,从而显著提升模型在复杂场景下的分割性能。更为显著的是,当仅引入 AWF 模块而未引入 CA 模块时(双分支+AWF),模型 $mIoU$ 为 0.505 9,在引入 CA 模块后提升至 0.903 7,提升了 78.63%。这一结果证实,CA 模块在优化通道权重分布、突出关键语义通道、抑制背景噪声方面发挥了重要作用。

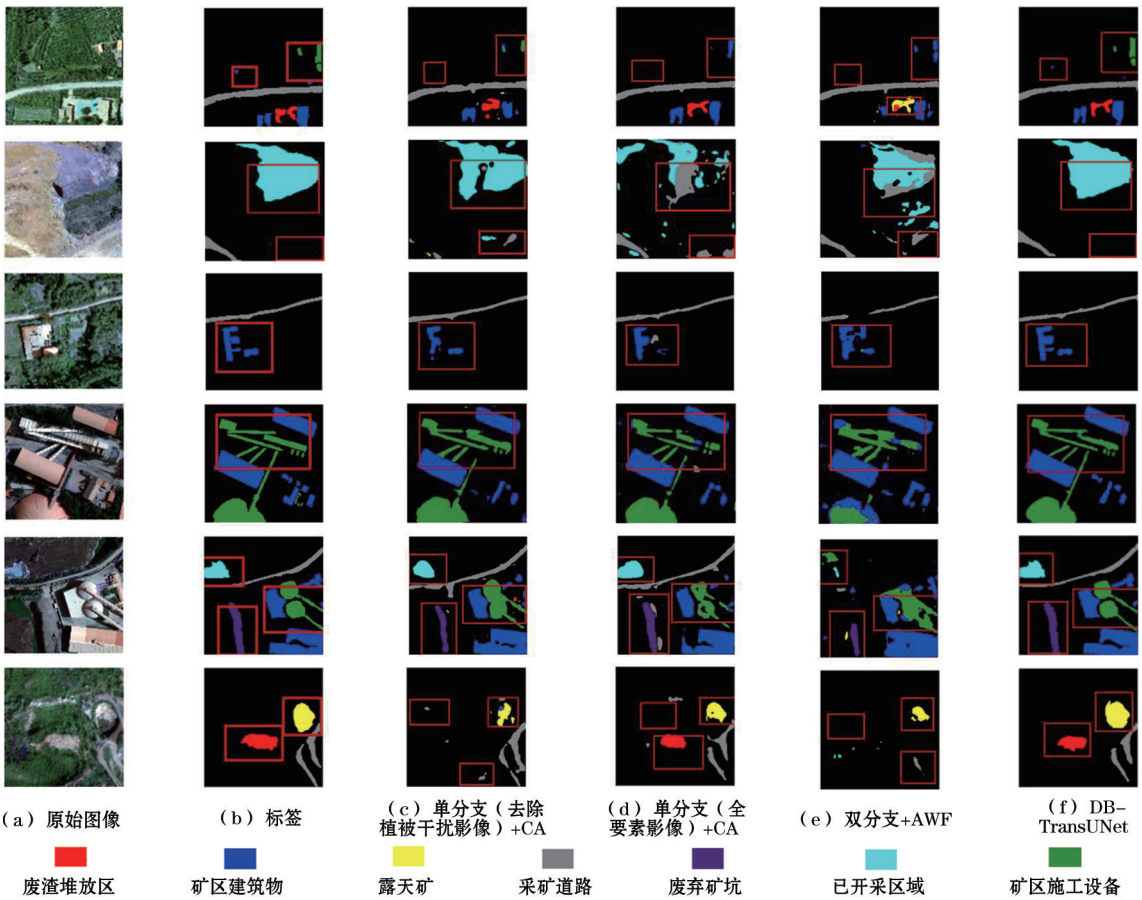


图 2 消融试验可视化结果

Fig.2 Visualization results of the ablation experiments

表 1 消融试验结果

Table 1 Results of the ablation experiments

去除植被特征	全要素特征	AWF	CA	*mIoU*	*R*	*P*	*mDice*	F_1 系数
√			√	0.825 0	0.888 9	0.918 2	0.902 5	0.903 3
	√		√	0.867 6	0.933 9	0.922 0	0.927 6	0.926 8
√	√	√		0.505 9	0.608 4	0.711 5	0.655 2	0.656 0
√	√	√	√	0.903 7	0.956 2	0.941 6	0.948 8	0.938 3

综上所述，消融试验验证了各模块在模型性能提升中的实际贡献。去除植被干扰的影像有助于抑制背景中的冗余信息，从而减少误分割现象；全要素遥感影像则保留了丰富的地物信息，有助于强化地物结构的表达能力与类别判别性；AWF 模块实现了跨分支的特征动态融合，提升了整体语义一致性；CA 模块则在通道层面优化特征表达，进一步强化了目标区域的判别能力。总而言之，完整的 DB-TransUNet 模型在复杂遥感场景中展现了优越的地物分割性能（$mIoU=0.903\ 7$，$mDice=0.948\ 8$），充分体现了多源特征融合与通道增强机制在遥感语义分割中的应用潜力与实用价值。

3.2 对比试验

为了验证所提出的双分支 DB-TransUNet 模型在矿山地物分割任务中的性能优势，将该模型与当前几种主流的分割网络架构进行对比试验，包括 U-Net、SegFormer、DeepLabV3＋和 TransUNet。模型均在相同的矿山数据集上进行训练和测试，以确保试验结果的公平性和可比性。

图 3 为对比试验可视化结果。从图 3 可以看出，U-Net 模型在矿山地物分割中能够较好地识别出矿山中主要地物，但在植被与地物交界处的分割边界不够清晰，存在一定的误分割现象；SegFormer 模型虽然在召回率上表现较好，但其分割结果存在噪声点，

导致分割精度有所下降;DeepLabV3+模型在处理矿山地物分割时,对边缘细节的捕捉能力不足,影响分割结果的完整性;TransUNet 模型受单分支结构的局限性,在植被干扰严重的区域分割效果较差,难以准确区分植被与矿山地物;相比之下,双分支 DB-TransUNet 模型的分割结果更加准确、清晰,该模型通过双分支结构和特征融合,有效抑制了植被干扰,增强了矿山地物的特征表征。在植被与矿山中地物交界处,分割边界清晰准确,能够很好地区分不同地物类别。同时,模型在解码阶段引入的 CA 模块进一步优化了特征权重分配,强化了关键区域的特征响应,使得分割结果更加精细,细节提取更加准确,明显减少了误分割现象。

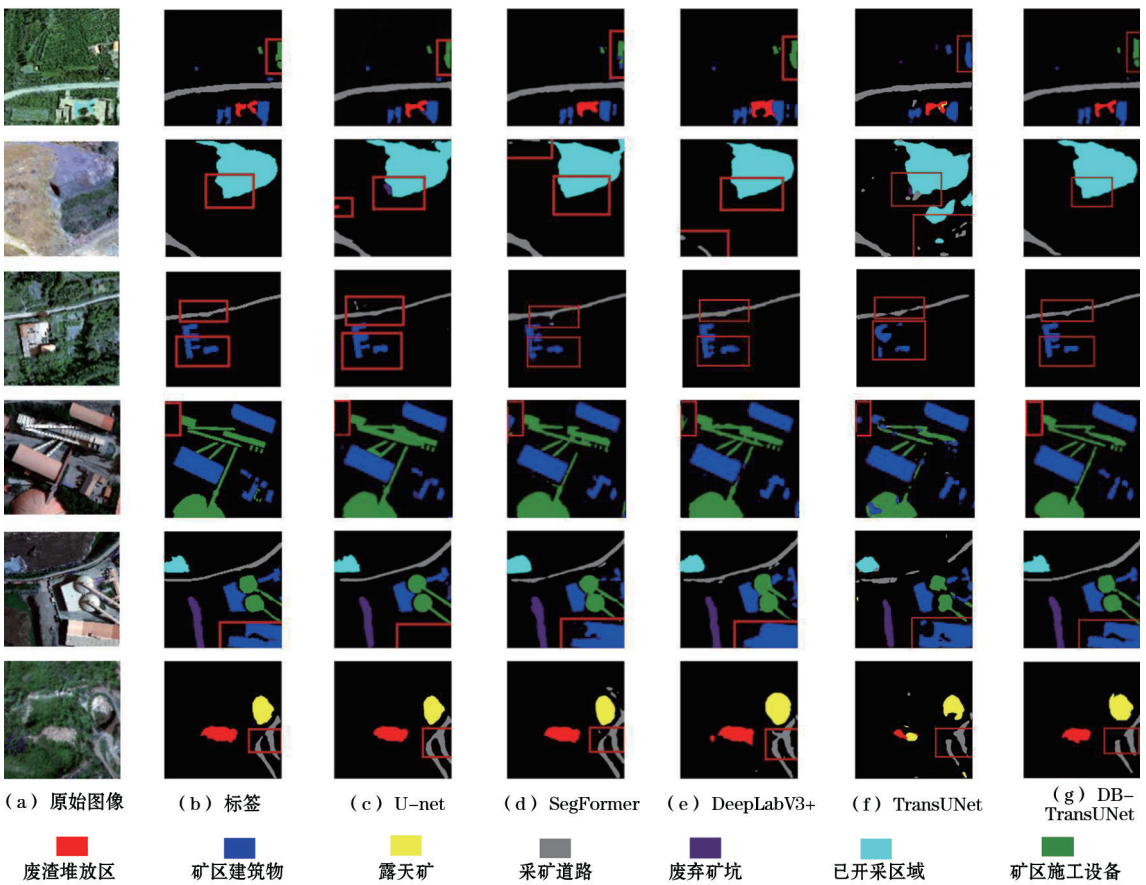


图 3 对比试验可视化结果

Fig.3 Visualization results of the comparative experiments

从表 2 的评估结果可以看出,双分支 DB-TransUNet 模型在各项指标上均显著优于其他主流对比模型。具体而言,双分支 DB-TransUNet 模型的 *mIoU* 值达 0.903 7,相比于 U-Net、SegFormer、DeepLabV3+ 和 TransUNet 模型分别提升了 6.8%、8.79%、16.29%、40.61%。

表 2 对比试验结果

Table 2 Results of the comparative experiments

方法	*mIoU*	*R*	*P*	*mDice*	F_1 系数
U-Net	0.846 2	0.896 0	0.936 8	0.929 1	0.915 8
SegFormer	0.830 7	0.913 0	0.898 7	0.932 1	0.906 8
DeeplabV3+	0.777 1	0.863 2	0.882 3	0.899 4	0.873 5
TransUNet	0.642 7	0.731 4	0.821 3	0.775 3	0.773 5
DB-TransUNet	0.903 7	0.956 2	0.941 6	0.948 8	0.938 3

对比试验结果表明,双分支 DB-TransUNet 在矿山地物分割精度方面具有显著优势,能够更准确地识别和区分矿区中结构复杂、分布多样的地物类别。此外,*mDice* 作为衡量预测结果与真实标注之间相似度的重要指标,双分支 DB-TransUNet 同样表现优异,达到了 0.948 8,显著高于所有对比模型,进一步验证了所提方法在处理矿山遥感影像中多类别复杂地物的精细分割任务时,具备更强的泛化能力与表达能力,具有良好的应用潜力。

4 结论与展望

4.1 结论

针对矿山遥感影像中因植被干扰导致地物分割精度下降的问题,提出了一种改进的双分支语义分

割模型 DB-TransUNet。该模型创新性地构建了双分支编码结构,并通过自适应加权融合(AWF)模块整合双分支特征,引入自适应增强(CA)模块优化解码过程中的特征权重分配,得到以下主要结论。

(1) AWF 与 CA 模块分别实现了特征动态融合与通道特征优化,消融试验结果表明,引入 AWF 与 CA 模块,$mIoU$ 增幅为 4.16% 和 78.63%。完整 DB-TransUNet模型在复杂遥感场景中取得 $mIoU=0.903\ 7$、$mDice=0.948\ 8$ 的优异分割效果,验证了多源特征融合与通道增强机制的实用价值。

(2) 对比主流网络分割架构,DB-TransUNet 模型各项指标更优异,在矿山复杂地物的识别与精细分割上精度更高、泛化能力更强,具备突出的应用潜力。

4.2 展望

本研究所提出的方法为高分辨率遥感影像的矿山智能监测提供了高精度的技术支撑,具有重要的工程应用价值。未来将围绕模型结构进一步优化、数据集规模扩展、分割性能提升与推理效率提高等方向展开研究,以持续提升模型的泛化能力与实用性,满足矿山动态监测对高精度和高效率的双重需求。

参考文献(References):

[1] 谢武平,何益君,曾蔚,等.露天煤矿矿山地质环境因子提取与环境恢复治理分区综合评价研究[J].中国矿业,2024,33(8):85-91.
XIE Wuping, HE Yijun, ZENG Wei, et al. Research on the extraction of geological environmental factors and comprehensive evaluation of environmental restoration and management zoning in open-pit coal mines[J]. China Mining Magazine,2024,33(8):85-91.

[2] 张增祥,汪潇,温庆可,等.土地资源遥感应用研究进展[J].遥感学报,2016,20(5):1243-1258.
ZHANG Zengxiang, WANG Xiao, WEN Qingke, et al. Research progress of remote sensing application in land resources[J]. National Remote Sensing Bulletin, 2016, 20(5):1243-1258.

[3] 张国敏,殷建平,祝恩,等.遥感图像中基于视觉显著性的分层目标检测[J].吉林大学学报(工学版),2007,37(3):625-629.
ZHANG Guomin, YIN Jianping, ZHU En, et al. Hierarchical object detection in remote sensing image based on vision saliency [J]. Journal of Jilin University (Engineering and Technology Edition),2007,37(3):625-629.

[4] MAXWELL A E, WARNER T A, STRAGER M P, et al. Assessing machine-learning algorithms and image- and lidar-derived variables for GEOBIA classification of mining and mine reclamation [J]. International Journal of Remote Sensing,2014,35(9):3332-3355.

[5] 陈华丽,陈植华,丁国平.用基于知识的决策树方法分层提取矿区土地类型:以湖北大冶为例[J].国土资源遥感,2004(3):49-53.
CHEN Huali, CHEN Zhihua, DING Guoping. The application of the knowledge-based decision tree classification method to the extraction of land types in mining areas: a case study of Daye area, Hubei Province[J]. Remote Sensing for Land & Resources,2004(3):49-53.

[6] 康俊锋,周琳,赵得军,等.基于区域合并多尺度分割的稀土矿区典型地物识别方法[J].稀土,2020,41(2):32-40.
KANG Junfeng, ZHOU Lin, ZHAO Dejun, et al. High resolution remote sensing image object recognition method based on histogram feature knowledge base[J]. Chinese Rare Earths,2020,41(2):32-40.

[7] PENG Yan, ZHANG Zhaoming, HE Guojin, et al. An improved grab cut method based on a visual attention model for rare-earth ore mining area recognition with high-resolution remote sensing images[J]. Remote Sensing, 2019,11(8):987.

[8] ZHAO Qiuquan, LI Jianyuan. SPS-UNet: a super-pixel sampling UNet for extracting buildings from high-resolution satellite images[J]. The Visual Computer,2025,41(8):1-14.

[9] YAO Xudong, GUO Qing, LI An. Light-weight cloud detection network for optical remote sensing images with attention-based DeeplabV3 + architecture [J]. Remote Sensing,2021,13(18):3617.

[10] WANG Jiawen, LI Dewei, LONG Qihang, et al. Real-timesemantic segmentation for underground mine tunnel[J]. Engineering Applications of Artificial Intelligence,2024,133(Part C):108269.

[11] SAPUTRA M R U, BHASWARA I D, NASUTION B I, et al. Multi-modal deep learning approaches to semantic segmentation of mining footprints with multispectral satelliteimagery[J]. Remote Sensing of Environment,2025,318:114584.

[12] LIU Yang, ZHANG Zelin, LIU Xiang, et al. Deep learning based mineral image classification combined with visual attention mechanism[J]. IEEE Access,2021(9):98091-98109.

[13] CHEN Mian, YANG Bin, WANG Feng, et al. Identification of open-pit mines and surrounding vegetation on high-resolution satellite images based on improved bilateral segmentation network semantic segmentation model[J]. Journal of Applied Remote Sensing,2023,17(4):044518.

[14] ALEISSAEE A A, KUMAR A, ANWER R M, et al. Transformers in remote sensing: a survey[J]. Remote Sensing,2023,15(7):1860.

[15] WANG Hongmiao, XING Cheng, YIN Junjun, et al. Land cover classification for polarimetric SAR images based on vision transformer[J]. Remote Sensing,2022,14(18):4656.

[16] CHEN Jieneng, MEI Jieru, LI Xianhang, et al. TransUNet: rethinking the U-Net architecture design for medical image segmentation through the lens of transformers[J]. Medical Image Analysis,2024(97):103280.