

面向地下矿山巷道裂缝智能检测的 UMTC-net 设计与工程化应用*

段雨静

(新乡职业技术学院, 河南 新乡市 453000)

摘要:针对地下矿山巷道复杂地质环境下裂缝检测精度不足、实时性受限等工程难题,提出一种融合多尺度特征感知与自适应注意力机制的地下矿山巷道裂缝分割网络(Underground Mine Tunnel Crack-segmentation network, UMTC-net)。该网络通过分层级联 Swin Transformer 模块群,实现裂缝图像从局部纹理到全局结构的跨尺度特征提取;同时,引入对数空间相对位置编码的缩放余弦注意力机制,有效抑制异常像素干扰;此外,构建基于补丁动态合并/扩展的编解码框架,解决了传统检测方法中细微裂缝边界定位模糊、复杂背景误检率高的问题。结果表明:UMTC-net 在 Crack500 数据集中准确率达 85.15%、平均交并比为 85.78%、 F_1 值达到 83.27%,在 MineTunnelCrack-2000 数据集中准确率高达 87.51%、平均交并比为 79.98%、 F_1 值达到 86.95%,均高于其他 4 种主流网络,尤其在低光照、高粉尘环境下表现出更强的鲁棒性;该网络在 RTX 3060 移动显卡上实现 38.9 ms 的推理速度,显存占用仅 5 230 MB,部署成本降低 40% 以上,满足便携式检测设备的实时性与低功耗需求,且适配性价比更高。现场测试中 UMTC-net 检测效率较人工巡检提升 8 倍,漏检率从 18% 降至 3.2%。研究结果为地下矿山巷道裂缝检测提供了高效精准的新方案,有助于及时发现安全隐患,保障矿山生产安全与设备稳定运行。

关键词:地下矿山巷道;裂缝检测;多尺度特征融合;自适应注意力机制;矿山图像分析

中图分类号:TD263;TP391

文献标识码:A

文章编号:1005-2763(2025)10-0152-08

UMTC-net Design and Engineering Application for Intelligent Detection of Underground Mine Tunnels Cracks

DUAN Yujing

(Xinxiang Vocational and Technical College, Xinxiang, Henan 453000, China)

Abstract: Aiming at the engineering problems such as insufficient crack detection accuracy and limited real-time performance in the complex geological environment of underground mine tunnels, an underground mine tunnel crack-segmentation network (UMTC-net) integrating multi-scale feature perception and adaptive attention mechanism was proposed. This network can realize cross-scale feature extraction of crack images from local texture to global structure through hierarchical cascading of Swin Transformer module groups. Meanwhile, a scaling cosine attention mechanism encoded by relative positions in logarithmic space was introduced to effectively suppress the interference of abnormal pixels. In addition, a codec framework based on dynamic patch merging/expansion was constructed, which solved the problems of ambiguous boundary positioning of fine cracks and high false detection rate in complex backgrounds in traditional methods. The results show that the UMTC-net has an accuracy of 85.15%, an average intersection-union ratio of 85.78%, and an F_1 value of 83.27% in the Crack500 dataset, and an accuracy of 87.51%, an average intersection-union ratio of 79.98%, and an F_1 value of 86.95% in the MineTunnelCrack-2000 dataset. It exhibits stronger robustness, especially in low light and high dust environments. This network achieves an inference speed of 38.9 ms on the RTX 3060 mobile graphics card, occupying only 5 230 MB of memory and reducing deployment costs by more than 40%. It meets the real-time and low-power requirements of portable detection devices, and has a higher cost-effectiveness for adaptation. In the field test, the detection efficiency of UMTC-net is 8 times higher than that of manual inspection, and the missed detection rate is reduced from 18% to 3.2%. The research results provide an efficient and accurate new scheme for crack detection in underground mine tunnels, which is helpful to find potential safety hazards in time and ensure the safety of mine production and stable operation of equipment.

Key words: Underground mine tunnel, Crack detection, Multi-scale feature fusion, Adaptive attention mechanism, Mine image analysis

* 收稿日期:2025-05-22

基金项目:河南软科学研究计划资助项目(132500710532)

作者简介:段雨静(1980—),女,河南新乡人,硕士,讲师,研究方向为计算机网络。E-mail:2674510508@qq.com

0 引言

地下矿山巷道作为矿石运输、人员通行以及通风等关键作业的通道,其结构完整性对矿山安全高效生产意义重大^[1]。然而,巷道长期受地质应力、地下水侵蚀以及开采活动影响,极易产生裂缝。及时、精准地检测这些裂缝,预防巷道坍塌等严重事故,是矿业工程领域亟待解决的工程难点^[2]。目前,矿山巷道裂缝检测方法主要包括人工检测和自动化检测两大类。人工检测方法主要依靠经验丰富的检测人员对巷道进行现场观察,该方法虽然直观,但效率低下、劳动强度大,且受主观因素影响较大。自动化检测方法主要包括基于图像处理、激光扫描和声波检测等技术^[3]。其中,基于图像处理的裂缝检测方法因其非接触、高效、易于实现等优点,在矿山巷道裂缝检测领域得到了广泛关注。随着人工智能与大数据技术蓬勃发展,将智能算法深度融入矿山安全检测成为热门研究方向。实现地下矿山巷道裂缝检测的智能化、精准化,不仅契合学术前沿趋势,更是推动矿山安全生产迈向新阶段的关键举措,具有重要的现实意义与研究价值^[4]。

近年来,深度学习技术在图像处理领域取得了显著的成果,许多基于深度学习的分割网络被应用于矿山巷道裂缝检测。例如,陈锡明等^[5]提出了一种双任务卷积神经网络,通过同时进行裂缝面状分割和线性踪迹提取,有效提高了裂缝检测的精度,该方法在自建和公共数据集上的性能均优于其他对比方法,且联合训练两个任务能显著提升模型效果。张华卫等^[6]设计了一种基于并行提取和注意力融合网络的裂缝检测算法,通过多尺度卷积并联神经网络提取特征,结合像素注意力机制进行特征融合,有效提高了检测精度和细节清晰度,精确率达到 85%,在消除噪声干扰方面取得了显著效果;但此算法对复杂场景下的裂缝检测仍存在局限性,且依赖大量标注数据进行监督学习,可能导致泛化能力不足。

综合来看,现有深度学习方法在矿山巷道裂缝检测中存在复杂背景下误检率高、细微裂缝边界定位模糊等问题,如传统编解码网络难以平衡局部纹理与全局结构的特征融合,注意力机制易受异常像素干扰。针对这些问题,本研究提出一种融合多尺度特征感知与自适应注意力机制的地下矿山巷道裂缝分割网络(Underground Mine Tunnel Crack-segmentation network, UMTC-net)。矿山图像具有低光照导致的对比度低、高粉尘引发的图像模糊、

潮湿表面反光造成的噪声干扰等特性,且裂缝多呈现细微、不连续的线性或网状结构,传统检测方法易受这些因素影响。相比之下,UMTC-net 的设计针对适配矿山场景:分层级联 Swin Transformer 模块群通过多尺度特征提取,能捕捉不同宽度裂缝的局部纹理与全局结构,解决粉尘模糊导致的特征丢失问题;对数空间相对位置编码的缩放余弦注意力机制,可抑制反光、异常像素等噪声干扰,增强对低光照下微弱裂缝特征的关注;补丁动态合并/扩展的编解码框架,配合跳跃连接,提升了细微裂缝边界定位精度,避免传统方法在复杂背景下的误检。程德强等^[7-8]的研究中,迭代压缩 U 型网络展现了对复杂结构的精细分割思路,其煤矿视频 AI 识别研究也提出了有效的抗噪声策略。UMTC-net 在特征融合与噪声抑制方面的设计,充分借鉴了这些思路与策略。因此,该网络更贴合矿山图像分析的实际需求,在极端环境下的鲁棒性也优于其他算法。

1 研究方法

1.1 UMTC-net 结构设计

地下矿山巷道裂缝分割网络(UMTC-net)的结构如图 1 所示。

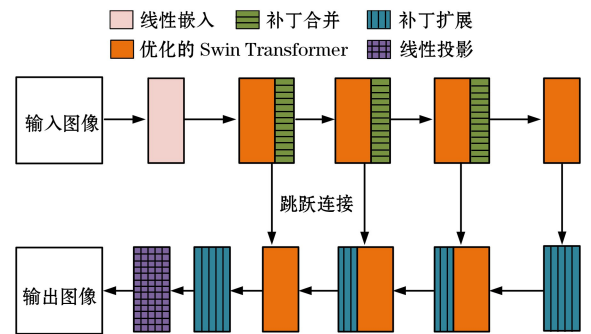


图 1 UMTC-net 结构

Fig.1 The UMTC-net structure

UMTC-net 结构主要由编码器、瓶颈部分、解码器以及跳跃连接组成。首先,当图像输入时,经过线性嵌入处理,将图像转化为适合网络处理的形式。接着,图像数据进入由多个优化的 Swin Transformer 模块和补丁合并层构成的编码器部分,实现不同尺度下的特征提取^[9-10]。在瓶颈部分,通过特定设置保持特征维度和分辨率。解码器则由多个自注意力模块和补丁扩展层组成,用于提升特征图分辨率,逐步恢复至原始图像尺寸^[11-12]。此外,跳跃连接将编码器和解码器的特征进行融合,增强网络对空间细节和上下文信息的整合能力。最终,经过线性投影层进行像素级

预测,输出包含裂缝检测结果的图像。

1.2 特征提取与优化流程

多个 Swin Transformer 模块与补丁合并层协同工作,完成图像不同尺度下的特征提取。在特征提取过程中,改进的注意力模块可增强特征学习能力,让网络敏锐地捕捉到图像中的各类关键信息。与此同时,残差连接可提升特征学习效果,并降低计算复杂度^[13-14]。在大规模视觉模型应用场景下,学习注意力图容易受到少量像素的影响。为此,研究采用 Swin Transformer V2 中的缩放余弦注意力来产生更为温和的注意力值,从而避免因少量异常像素而导致的注意力偏移。缩放余弦注意力为:

$$\text{sim}_{ij} = \frac{\cos\left(\frac{q_i^T k_j}{\sqrt{d_k}} + B_{ij}\right)}{\tau} \quad (1)$$

式中: i, j 为像素索引; B_{ij} 代表像素 i 和 j 之间的相对位置偏差; τ 代表可学习标量; d_k 表示缩放点积; q_i 表示查询向量 \mathbf{Q} 的元素; k_j 表示键向量 \mathbf{K} 的元素。最终的注意力计算为:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} + \mathbf{B}\right)\mathbf{V} \quad (2)$$

式中: \mathbf{V} 表示值向量; \mathbf{B} 表示相对位置偏差向量; softmax 表示激活函数。研究进一步优化,以平均池化对 \mathbf{Q} 和 \mathbf{K} 进行下采样,并引入残差连接结构,得到改进后的注意力机制为:

$$\text{Attention}(\mathbf{Q}', \mathbf{K}', \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}'\mathbf{K}'^T}{\sqrt{d_k}} + \mathbf{B}\right)\mathbf{V} + \mathbf{Q}' \quad (3)$$

式中: \mathbf{Q}' 表示平均池化后的查询向量; \mathbf{K}' 表示平均池化后的键向量。

相对位置偏差基于对数空间坐标计算,在计算注意力时其在窗口间传播外推量极小^[15-16]。改进后的注意力模块如图 2 所示。

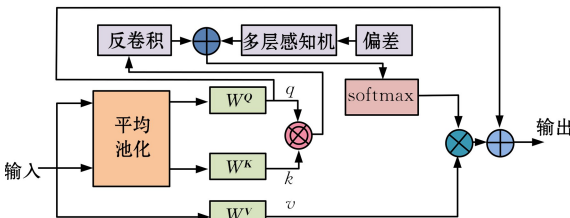


图 2 改进后的注意力模块

Fig.2 Improved attention module

研究通过整合补丁合并层与补丁扩展层,实现下采样与上采样的处理,从而构建出一个具备编码与解码功能的框架。该框架整合了自注意力机制的

优点,并充分发挥了编解码结构在多层次信息整合方面的特性。补丁合并层的工作过程如图 3 所示。在补丁合并层级中,每当一组小图块被编码完成后,这 4 个相邻的小图块便会融合成一个较大的图块,并沿着深度方向进行拼接^[17-19]。这一过程使图块的维度由 $H \times W \times C$ 转换为 $0.5H \times 0.5W \times 4C$ 。接下来,对合并后的图块执行层归一化处理,通过卷积操作将通道数减少至 $2C$ 。最终,得到尺寸为 $0.5H \times 0.5W \times 2C$ 的输出特征图。

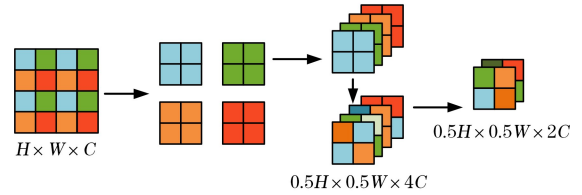


图 3 补丁合并层的工作过程

Fig.3 The working process of the patch merge layer

为解决结构收敛问题,瓶颈结构仅包含两个改进自注意力模块,以保持特征维度和分辨率。解码器由多个自注意力模块和补丁扩展层构成,通过逐级上采样,使特征图分辨率翻倍,最终恢复原始图像尺寸^[20]。然后,通过线性投影层进行像素级预测,判断像素是否为裂缝。同时,补丁扩展层先对特征增维、重排,实现分辨率扩展和维度降低。最后,跳跃连接结合编码器和解码器特征,融合空间细节与上下文信息,从而提高分割性能^[21]。

1.3 模型训练

在模型训练过程中,裂缝分割是一项二分类任务,因此本研究采用了 Dice Loss 与二元交叉熵损失函数(Binary Cross-Entropy Loss, BCE Loss)相结合的方式来进行损失函数的计算。Dice Loss 计算公式为:

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2} \quad (4)$$

式中: L_{Dice} 表示 Dice Loss 损失函数; p_i 表示模型预测值; g_i 表示标签值; N 表示样本数。BCE Loss 计算公式为:

$$L_{\text{BCE}} = -[a \cdot \ln(p_i) + (1 - a) \cdot \ln(1 - p_i)] \quad (5)$$

式中: L_{BCE} 表示二次交叉熵损失函数; a 表示真实标签; p_i 表示预测概率。联合损失函数计算公式为:

$$L = \alpha L_{\text{Dice}} + \beta L_{\text{BCE}} \quad (6)$$

式中: L 表示联合损失函数; α 表示 Dice Loss 的权重; β 表示 BCE Loss 的权重。模型参数的优化选择

了 Adam 优化算法,学习率设定为 0.000 1,种群规模设定为 45,迭代次数设定为 600。

2 试验分析

2.1 试验设置

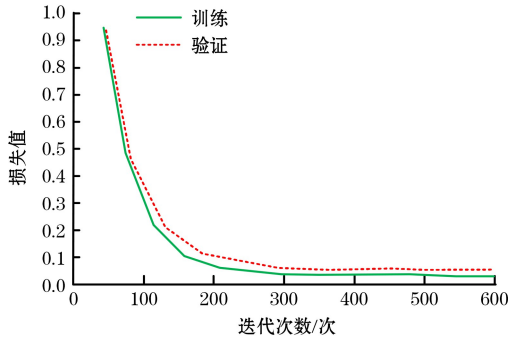
本研究选择 Linux 操作系统,采用 Python 3.8 作为主要的开发语言,同时使用 PyCharm 集成开发环境进行代码的编写与调试。试验所用的硬件设施包括 11 代 Intel Core i9-11900K 处理器,以及具备图形处理能力的 GeForce RTX 3090 GPU。在软件配置方面,采用 PyTorch 1.10.1 深度学习框架,并与 CUDA 11.4 和 cuDNN 8.2.4 相结合。

数据采集于山西某铜矿地下巷道,该区域兼具低光照(100~1 000 Lux)、高粉尘(5~50 mg/m³)、潮湿等复杂作业环境,巷道表面裂缝形态多样,包括

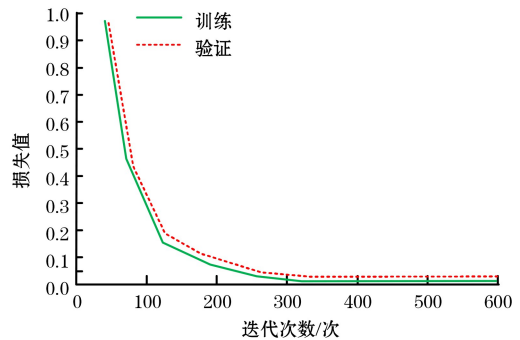
网状、线性、分叉状等。自制数据集命名为 MineTunnelCrack-2000,包含 2 000 组裂缝图像,由专业地质工程师依据矿山行业裂缝检测规范进行像素级标注,标注内容涵盖裂缝宽度(0.1~5 mm)、走向、连通性等特征,训练集、验证集、测试集按 7:2:1 比例划分,确保数据多样性与代表性。同时,将对公开数据集 Crack500 进行试验,该数据集包含了 500 张不同类型的裂缝图像。

2.2 试验结果

UMTC-net 分割网络的损失曲线如图 4 所示。图 4(a)为自制试验数据集 MineTunnelCrack-2000 下的损失曲线,随着迭代次数的增加,损失曲线先迅速下降,然后趋于稳定。最终,训练损失约为 0.04,验证损失约为 0.05。图 4(b)为 Crack500 数据集下的损失曲线,稳定后的训练损失约为 0.02,验证损失约为 0.03。



(a) MineTunnelCrack-2000 数据集

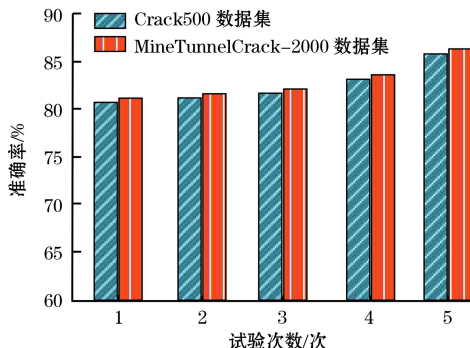


(b) Crack500 数据集

图 4 UMTC-net 的损失曲线

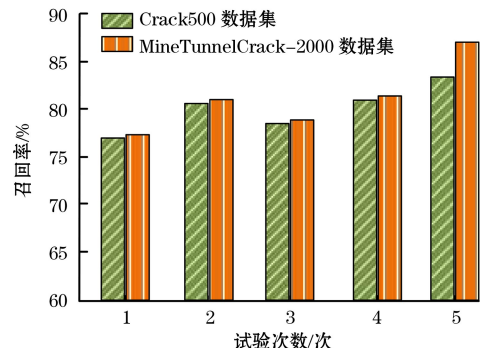
Fig.4 Loss curves of UMTC-net

UMTC-net 分割网络的消融试验结果如图 5 所示。图 5 中,试验次数 1~5 分别表示 Swin Transformer 模块、Swin Transformer V2 模块、Swin Transformer V2 模块+平均池化、Swin Transformer V2 模块+改进后的注意力模块,以及 Swin Transformer V2 模块+平均池化+改进后的注意力模块(即完整的 UMTC-net 分割网络)。由图 5 可知,随着试验次数的增加,准确率逐渐增



(a) 准确率变化

加,召回率、F1 值和平均交并比呈现先增加后减少再增加的趋势。可以看出,Swin Transformer V2 模块和改进后的注意力模块的加入,使模型的性能显著提升。Swin Transformer V2 模块在特征提取和表达上比原始的 Swin Transformer 模块更有效,注意力机制能够有效增强模型对于重要特征的识别和利用,进一步提高了模型的分割精度。



(b) 召回率变化

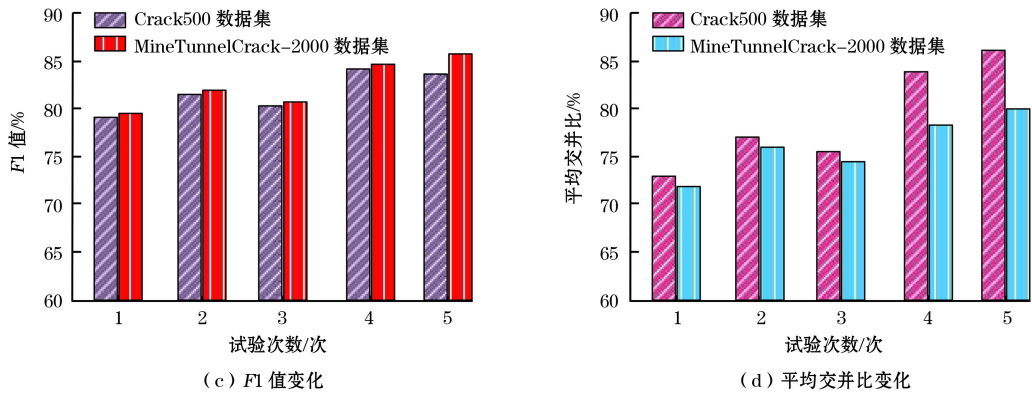


图 5 UMTC-net 的消融试验结果

Fig.5 Ablation experimental results of UMTC-net

参数的消融试验结果见表 1。参数消融试验结果显示,当 $\alpha=0.6, \beta=0.4$ 时,联合损失函数效果最优,较比例 $\alpha=0.5, \beta=0.5$ 的准确率提升 1.23~1.38 个百分点,说明该参数设置更适配裂缝二分类任务。

表 1 参数的消融试验结果

Table 1 The ablation experiment results of the parameters

α	β	准确率/%		平均交并比/%	
		Crack500	MineTunnelCrack-2000	Crack500	MineTunnelCrack-2000
0.4	0.6	82.18	85.78	83.47	78.24
0.5	0.5	83.92	86.13	84.21	78.65
0.6	0.4	85.15	87.51	85.78	79.98
0.7	0.3	84.56	86.89	85.02	79.12

2.3 与其他主流网络对比分析

为了验证本研究提出的 UMTC-net 分割网络的优势,将改进 Deep Labv3+ 网络^[21]、U-Net^[10]、SegNet^[3] 和基于 Transformer 的全卷积网络

(Transformer based Fully Convolutional Network, TFCN)^[14] 作为对比。Crack500 数据集中 5 种网络的性能对比结果见表 2。由表 2 可知,UMTC-net 准确率达 85.15%,远高于 U-Net 的 77.21%;平均交并比为 85.78%,优于改进 Deep Labv3+ 网络的 83.56%;F1 值达到 83.27%,也领先于其他 4 种网络。

表 2 Crack500 数据集中 5 种网络的性能对比结果

Table 2 Performance comparison results of five networks in the Crack500 dataset

网络	准确率/%	召回率/%	F1 值/%	平均交并比/%
UMTC-net	85.15	80.13	83.27	85.78
改进 Deep Labv3+	83.45	80.09	82.95	83.56
U-Net	77.21	71.18	74.23	77.51
SegNet	79.41	77.82	78.54	81.55
TFCN	84.65	79.89	83.77	84.52

Crack500 数据集中不同网络的分割结果如图 6 所示。

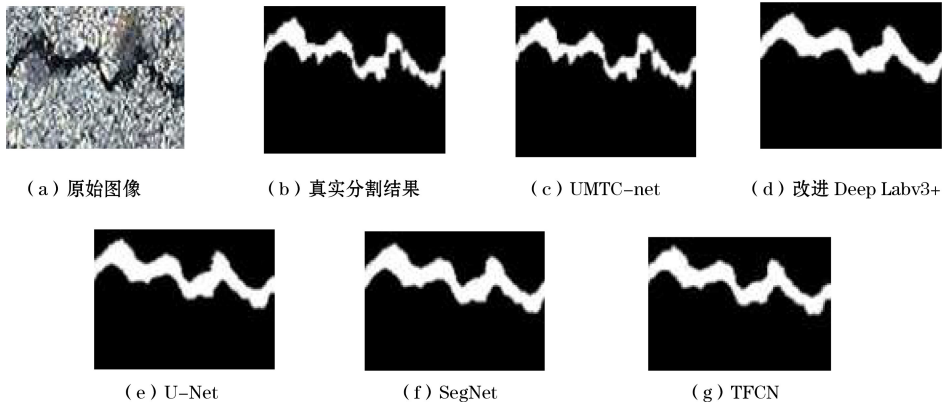


图 6 Crack500 数据集中不同网络的分割结果

Fig.6 The segmentation results of different networks in the Crack500 dataset

由图 6 可知:U-Net 分割结果出现了一定程度的偏差,改进 Deep Labv3+ 的分割结果则在细节上

有所欠缺,对一些细小裂缝的捕捉能力不足,使得裂缝的整体轮廓不够清晰;TFCN 的分割效果相对较

好,能够清晰地勾勒出裂缝的边缘,但在某些区域仍存在漏判的情况;SegNet 的分割结果则呈现出了较多的噪点,使得分割结果看起来较为粗糙;UMTC-net 模型能够精确地分割出裂缝的边界,而且对细小裂缝的处理也十分到位。可以看出,UMTC-net 模型所呈现出的分割结果与真实分割结果几乎一致,其对裂缝的完整分割程度远远超过了其他算法。

MineTunnelCrack-2000 数据集中 5 种网络的性能对比结果见表 3。UMTC-net 准确率高达 87.51%, 远超改进 Deep Labv3+ 的 74.46%、U-Net 的 73.18% 以及 SegNet 的 74.56%;UMTC-net 的平均交并比达到 79.98%, 而其他 4 种网络中表现相对较好的 TFCN 仅为 78.25%;UMTC-net 的召回率为 86.17%, F1 值达到 86.95%, 均显著高于其他 4 种网络。结果表明,UMTC-net 在准确率、平均交并比、召回率和

F1 值等关键指标上表现出色,具有更优的分割性能。

表 3 MineTunnelCrack-2000 数据集中 5 种网络的性能对比结果

Table 3 Performance comparison results of five networks in the MineTunnelCrack-2000 dataset

网络	准确率/%	召回率/%	F1 值/%	平均交并比/%
UMTC-net	87.51	86.17	86.95	79.98
改进 Deep Labv3+	74.46	77.15	75.56	74.06
U-Net	73.18	70.13	71.17	70.59
SegNet	74.56	72.51	73.52	71.65
TFCN	85.61	84.12	83.62	78.25

MineTunnelCrack-2000 数据集中不同网络的分割结果如图 7 所示。由图 7 同样可以看出,UMTC-net 分割结果更接近真实分割结果,显示出该网络在细节处理和整体分割性能上的优势。

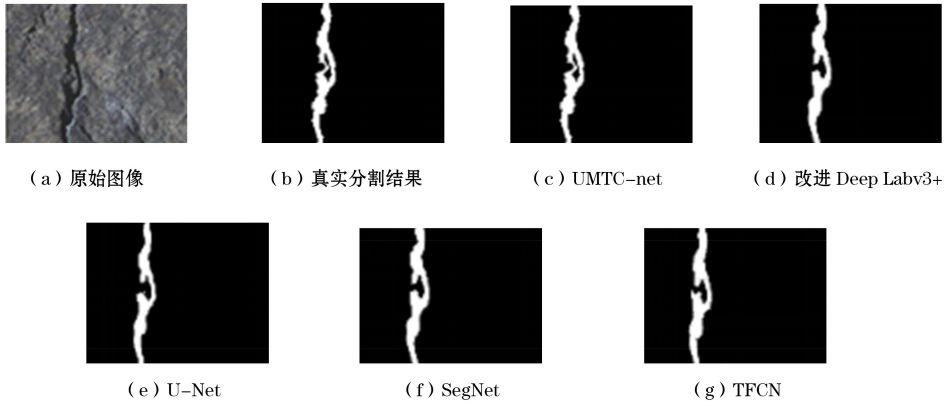


图 7 MineTunnelCrack-2000 数据集中不同网络的分割结果

Fig.7 Segmentation results of different networks in the MineTunnelCrack-2000 dataset

2.4 实际部署效果评估

在实时性测试中,选取主流 GPU 型号(RTX 3090、RTX 4080、Tesla V100)作为硬件平台,测试输入图像尺寸为 512×512 时的推理速度,并对比不同网络的性能,结果见表 4。

表 4 不同网络模型的推理速度与参数规模对比

Table 4 Comparison of inference speed and parameter scale of different network models

网络	推理速度/(帧/s)	参数数量/M	计算量/FLOPs	单张图像推理时间/ms
UMTC-net	48.7	68.2	12.5 G	20.5
改进 Deep Labv3+	32.2	52.1	8.3 G	31.1
U-Net	55.3	34.7	6.2 G	18.1
SegNet	42.6	45.8	9.1 G	23.5
TFCN	45.9	65.3	11.8 G	21.8

由表 4 可知,UMTC-net 的推理速度为 48.7

帧/s,满足实时检测需求(≥ 30 帧/s),尽管参数规模略高于其他 4 种网络,但其优化的注意力机制和编解码架构有效平衡了精度与速度。U-Net 推理速度最快,为 55.3 帧/s,但参数规模较小导致精度较低。TFCN 参数规模与 UMTC-net 接近,但推理速度稍慢。

测试了不同 GPU 型号下 UMTC-net 的运行时间和内存占用,评估其硬件适配性,结果见表 5。在高端 GPU(如 RTX 4080)上,UMTC-net 推理时间缩短至 15.2 ms,适合对速度要求极高的场景;在笔记本级 GPU(RTX 3060 Laptop)上仍能达到 38.9 ms,满足移动检测设备的需求。显存占用和功耗表现均衡,RTX 3060 Laptop 的低功耗特性表明 UMTC-net 可适配矿山现场的便携式检测设备,降低能耗成本。

表5 UMTC-net 在不同 GPU 上的性能指标

Table 5 The performance index of UMTC-net on different GPU

GPU 型号	推理时间/ms	显存占用/MB	峰值功耗/W
RTX 3090	20.5	4 820	280
RTX 4080	15.2	3 980	250
Tesla V100	32.7	6 510	300
RTX 3060 Laptop	38.9	5 230	115

从模型训练时间、数据标注成本和部署成本 3 个方面对 5 种网络的效率成本关键指标进行对比分析,结果见表 6。由表 6 可知:UMTC-net 训练时间比改进 Deep Labv3+ 和 TFCN 少,得益于优化的注意力机制和残差连接,减少了计算冗余;标注数据量与其他 4 种主流网络相当,但精度优势显著,说明其数据利用效率更高;部署成本低于 TFCN、改进 Deep Labv3+ 和 SegNet,适配性价比更高。研究成果已在山西某铜矿完成现场测试,采用 UMTC-net 网络的检测效率较人工巡检提升 8 倍,裂缝漏检率从 18% 降至 3.2%。该模型在复杂地质条件下的高鲁棒性,使其在矿山巷道日常安全监测、灾害预警等场景中具有显著应用潜力,可有效助力矿山智能化安全生产管理体系的构建。

表6 效率成本关键指标对比

Table 6 Comparison of key indicators of efficiency and cost

网络	训练时间/h	标注数据量/张	部署成本/元
UMTC-net	12.5	2 000(自制)+500(公开)	10 780(RTX 3 060)
改进 Deep Labv3+	15.3	2 500+800	14 380(RTX 3 070)
U-Net	8.2	1 500+500	8 630(RTX 3 050)
SegNet	10.8	2 000+600	12 940(RTX 3 080)
TFCN	14.1	2 200+700	17 970(RTX 4 090)

3 结论

(1) 本研究构建了适用于地下矿山复杂场景的裂缝智能检测网络 UMTC-net,其创新点在于:通过分层级联 Swin Transformer 模块实现多尺度特征提取,结合对数空间位置编码的缩放余弦注意力机制抑制噪声干扰,设计补丁动态合并/扩展的编解码框架提升边界定位精度,解决了传统检测方法中细微裂缝边界定位模糊、复杂背景误检率高的问题。

(2) UMTC-net 在 Crack500 数据集中准确率达 85.15%、平均交并比为 85.78%、F1 值达到 83.27%,在 MineTunnelCrack-2000 数据集中准确率高达 87.51%、平均交并比为 79.98%、F1 值达到

86.95%,均高于其他 4 种主流网络,尤其在低光照、高粉尘环境下表现出更强的鲁棒性。

(3) 工程化部署显示,模型在 RTX 3060 移动设备上实现 38.9 ms 实时检测,显存占用 5 230 MB,部署成本降低 40% 以上,满足便携式检测设备的实时性与低功耗需求,且适配性价比更高。

(4) 现场测试中 UMTC-net 检测效率较人工巡检提升 8 倍,漏检率从 18% 降至 3.2%。未来将聚焦多传感器融合与轻量化模型压缩,引入知识蒸馏技术将参数压缩至 30 MB 以下,提升边缘设备适配性,为全矿井智能检测提供技术支撑。

参考文献(References):

- [1] 苏宇,汤伏全,李景祥,等.矿山采动地表裂缝智能识别的 YOLOv7 模型改进研究[J].煤矿安全,2024,55(4):169-176. SU Yu, TANG Fuquan, LI Jingxiang, et al. Improvement of YOLOv7 model for intelligent recognition of mining surface cracks[J]. Safety in Coal Mines,2024,55(4):169-176.
- [2] 王田龙,陈从新,夏开宗,等.基于监测数据的陡倾结构金属矿山下盘岩层移动机制研究[J].岩石力学与工程学报,2022,41(7):1409-1426. WANG Tianlong, CHEN Congxin, XIA Kaizong, et al. Study of footwall ground movement mechanism in the iron mine with steeply dipping discontinuities based on monitoring data[J]. Chinese Journal of Rock Mechanics and Engineering,2022,41(7):1409-1426.
- [3] 刘超,吴纪曙.基于矢量量化变分自编码器的混凝土表面裂缝检测算法[J].同济大学学报(自然科学版),2024,52(11):1699-1705. LIU Chao, WU Jishu. Detection algorithm of concrete structural apparent cracks based on VQ-VAE-2[J]. Journal of Tongji University (Natural Science), 2024, 52(11): 1699-1705.
- [4] 刘恩良,袁俊亮,霍守东,等.基于深度学习的页岩地质露头裂缝特征自动识别[J].科学技术与工程,2024,24(30):12853-12863. LIU Enliang, YUAN Junliang, HUO Shoudong, et al. Automatic recognition of fracture characteristics from shale geological outcrop based on deep learning [J]. Science Technology and Engineering,2024,24(30):12853-12863.
- [5] 陈锡明,姚鑫,任开瑀,等.煤矿采空区地表裂缝双任务检测方法研究[J].遥感学报,2024,28(12):3271-3286. CHEN Ximing, YAO Xin, REN Kaiyu, et al. Dual-task model for ground crack detection in the goaf of coal mines[J]. National Remote Sensing Bulletin,2024,28(12):3271-3286.
- [6] 张华卫,贾文娟,张金龙,等.并联提取与特征融合注意力网络下的裂缝检测[J].电子测量技术,2022,45(10):102-111. ZHANG Huawei, JIA Wenjuan, ZHANG Jinlong, et al. Crack detection under parallel extraction and feature fusion attention network[J]. Electronic Measurement Technology, 2022,45(10):102-111.

- [7] 程德强,张瑞,谢同喜,等.基于迭代压缩 U 型网络的煤颗粒分割与粒度分析方法[J].煤炭学报,2025,50(2):1362-1375. CHENG Deqiang, ZHANG Rui, XIE Tongxi, et al. Segmentation and particle size analysis of coal particles based on ISUNet[J]. Journal of China Coal Society,2025,50(2):1362-1375.
- [8] 程德强,钱建生,郭星歌,等.煤矿安全生产视频 AI 识别关键技术研究综述[J].煤炭科学技术,2023,51(2):349-365. CHENG Deqiang, QIAN Jiansheng, GUO Xingge, et al. Review on key technologies of AI recognition for videos in coal mine[J]. Coal Science and Technology,2023,51(2):349-365.
- [9] 李元海,徐晓华,朱鸿鹄,等.基于计算机视觉的岩石裂隙识别表征与软件研制[J].岩土工程学报,2024,46(3):459-469. LI Yuanhai, XU Xiaohua, ZHU Honghu, et al. Identification and characterization of rock fractures based on computer vision and software development [J]. Chinese Journal of Geotechnical Engineering,2024,46(3):459-469.
- [10] 韩凤岩,李慧臻,杨少君,等.基于 FC-ResNet 网络的隧道衬砌裂缝像素级分割方法[J].现代隧道技术,2024,61(5):111-119. HAN Fengyan, LI Huizhen, YANG Shaojun, et al. Pixel-level segmentation method for tunnel lining cracks based on FC-ResNet network[J]. Modern Tunnelling Technology,2024,61(5):111-119.
- [11] DOUGLAS A, LANGENDERFER M, JOHNSON C. Road condition monitoring utilizing UAV photogrammetry aligned to principal curve of mine haul truck path [J]. Mining, Metallurgy & Exploration,2024,41(1):61-72.
- [12] 刘磊,丁勇,李登华.基于单目视觉下的裂缝三向变化检测[J].激光与光电子学进展,2024,61(18):177-186. LIU Lei, DING Yong, LI Denghua. Three-dimensional crack-change detection based on monocular vision [J]. Laser & Optoelectronics Progress,2024,61(18):177-186.
- [13] 蓝章礼,徐元通,赵胜薇,等.基于 Sobel 算子桥接的双编码器路面裂缝检测网络[J].重庆交通大学学报(自然科学版),2024,43(9):18-24+33. LAN Zhangli, XU Yuantong, ZHAO Shengwei, et al. Dual encoder pavement crack detection network based on Sobel operator bridging [J]. Journal of Chongqing Jiaotong University (Natural Science),2024,43(9):18-24+33.
- [14] ZHOU Z, ZHANG J J, GONG C J. Hybrid semantic segmentation for tunnel lining cracks based on Swin Transformer and convolutional neural network[J]. Computer-Aided Civil and Infrastructure Engineering,2023,38(17):2491-2510.
- [15] MOHANTY S, PANI S K. Empowering structural integrity: YOLO-based crack detection and ML-driven concrete strength prediction of critical infrastructure caused due to mining operation[J]. Journal of Electrical Systems,2024,20(5):2705-2721.
- [16] 胡波,陈翰新,任松,等.一种基于分割掩码的隧道裂缝病害自动识别后处理算法[J].测绘学报,2024,53(9):1715-1724. HU Bo, CHEN Hanxin, REN Song, et al. A post-processing algorithm for automatic recognition of tunnel crack diseases based on segmentation masks [J]. Acta Geodaetica et Cartographica Sinica,2024,53(9):1715-1724.
- [17] 赵佰亭,程瑞丰,贾晓芬.融合多尺度特征的 YOLOv8 裂缝缺陷检测算法[J].计算机工程与应用,2024,60(22):261-270. ZHAO Baiting, CHENG Ruifeng, JIA Xiaofen. YOLOv8 crack defect detection algorithm based on multi-scale features[J]. Computer Engineering and Applications,2024,60(22):261-270.
- [18] 徐雷,杜时贵,朱益军.露天矿山岩体力学参数估算法的应用研究[J].地下空间与工程学报,2022,18(3):925-934+954. XU Lei, DU Shigui, ZHU Yijun. Application research of rock mass mechanics parameter estimation in open-pit mines [J]. Chinese Journal of Underground Space and Engineering,2022,18(3):925-934+954.
- [19] YAMAGUCHI T, MIZUTANI T. Road crack detection interpreting background images by convolutional neural networks and a self-organizing map [J]. Computer-Aided Civil and Infrastructure Engineering,2024,39(11):1616-1640.
- [20] MUTAZ E, SERATI M, WILLIAMS D J. Crack mode-changing stress level in porous rocks under polyaxial stress conditions[J]. Acta Geotechnica,2024,19(2):783-803.
- [21] 金磊,杨晓伟,张浩,等.基于改进 DeepLabv3+与 SE 注意力机制融合的非结构化道路识别方法[J].煤炭工程,2024,56(7):200-204. JIN Lei, YANG Xiaowei, ZHANG Hao, et al. An unstructured road recognition method based on the fusion of improved DeepLabv3+ and SE attention mechanism [J]. Coal Engineering,2024,56(7):200-204.