

基于改进 U-Net 模型的混凝土表面裂缝分割

杨 智, 杨 鹏, 覃 森, 朱邦盛

(宜昌三峡送变电工程有限责任公司, 宜昌 443000)

摘 要: 为了对混凝土表面裂缝进行自动分割, 通过搜集试验数据及公开数据集, 采用数据增强建立了混凝土表面裂缝数据集, 提出了 2 种基于 U-Net 的全监督裂缝分割模型: DA-CrackNet 和 TransUNet。前者在 U-Net 的基础上加入软注意力机制和空间与通道结合的卷积块注意力机制; 后者由 Transformer 与 U-Net 结合而成。利用上述数据集进行训练和测试, 对比不同改进 U-Net 模型的裂缝分割性能。结果表明: 与 U-Net 相比, DA-CrackNet 准确率、召回率、mIoU 及 F1 分数分别提高了 1.41%、1.35%、2.07% 及 1.53%; 与 TransUNet 相比, DA-CrackNet 各性能指标差别不到 1%, 但训练时间降为 50%。因此, 综合考虑计算精度和计算效率, DA-CrackNet 性能最优。研究成果可为相关工程提供参考。

关键词: 混凝土表面裂缝分割; 全监督; 注意力机制; 数据增强; U-Net

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1005-8249 (2026) 03-0162-07

DOI: 10.19860/j.cnki.issn1005-8249.2026.03.027

Concrete Surface Crack Segmentation based on Improved U-Net Model

YANG Zhi, YANG Peng, QIN Sen, ZHU Bangsheng

(Yichang Three Gorges Power Transmission & Transformation Engineering Co., Ltd., Yichang 443000, China)

Abstract: In order to automatically segment concrete surface cracks, this paper first collected experimental data and public datasets to establish concrete surface crack dataset through data augmentation. Then, two fully-supervised crack segmentation models based on U-Net were proposed, namely, DA-CrackNet and TransUNet. The former adds a soft attention mechanism and a convolutional block attention mechanism combining spatial and channel information to the U-Net structure, while the latter is established by combining Transformer with U-Net. Based on the above dataset, the two proposed improved U-Net models were trained and tested to compare their crack segmentation performance. The results show that, compared to U-Net, DA-CrackNet's accuracy, recall, mIoU and F1 score increase by 1.41%, 1.35%, 2.07% and 1.53% respectively. Moreover, the difference of each performance metric between DA-CrackNet and TransUNet is less than 1%, while the training time of DA-CrackNet is reduced to 50% compared to TransUNet. Therefore, DA-CrackNet has the best performance considering both computational accuracy and efficiency. The findings offer valuable insights for similar engineering applications.

Key words: concrete surface crack segmentation; fully-supervised; attention mechanism; data augmentation; U-Net

作者简介: 杨 智 (1975—), 男, 本科, 高级工程师, 研究方向: 建筑结构。

通信作者: 杨 鹏 (1992—), 男, 硕士研究生, 工程师, 研究方向: 建筑结构。

收稿日期: 2025-10-24

研究与应用

0 引言

在环境因素与外部荷载作用下,混凝土材料在服役期间不可避免地产生劣化,进而导致混凝土结构出现损伤。表面裂缝是混凝土结构最早可察觉的损伤形式,裂缝的发展可能导致结构失效,从而带来巨大的经济损失甚至人员伤亡。因此裂缝识别^[1-2]成为混凝土结构健康监测领域的重要研究方向。传统的裂缝检测方法既耗时又费力,且易出现主观错误,而计算机视觉技术^[3]具有自动化、高效性及准确性等特点,在裂缝识别方面得到了广泛的应用。

基于计算机视觉的裂缝识别任务可分为两类:裂缝检测及裂缝分割。当前,机器学习和深度学习在裂缝识别任务中得到了最广泛的应用^[4-5]。在裂缝检测方面,Peng 等^[6]利用随机森林的三阈值路面裂缝检测方法,预测了局部裂缝补丁,生成了裂缝评分图,并从裂缝评分图中获得初步的裂缝检测结果;Liu 等^[7-8]提出了一种高效的隧道裂缝检测和识别方法,将原始图像转换为二值图像,将基于双阈值技术得到的三类物体边缘图像导入 SVM 中训练裂缝分类器;Shi 等^[9]应用积分通道特征重新定义构成裂缝的标记,引入随机森林算法生成高性能裂缝检测器,可识别任意复杂的裂缝;Liu 等^[10-11]提出了一种使用 VGG16 网络的二元分类方法,将图像中的混凝土裂缝分成结构性裂缝及非结构性裂缝,根据裂缝的特征对不同类型的裂缝进行分类。在裂缝分割方面,FCN 和 U-Net 模型得到了广泛应用。其中,Dung 等^[12]基于 FCN 提出了一种用于混凝土裂缝检测的语义分割网络,使用预训练好的 VGG16 分类网络作为 FCN 的编码器用于提取特征,将特征图经过解码器进行上采样得到裂缝的像素表示,测试结果表明该算法能够较好地检测出图像中的裂缝;Lau 等^[13]用预训练的 ResNet-34 替换 U-Net 的编码器部分,使用 Dice 系数损失作为损失函数,在 CFD 数据集^[14]和 Crack500 数据集^[15]进行训练和测试,结果表明其性能优于在这些数据集上训练和测试的其它模型。

综上所述,在混凝土表面裂缝识别(特别是裂缝分割)方面,深度学习方法得到了最为深入的研究,其中,U-Net 是其中应用较为成熟的深度学习模型,但其性能还有待进一步提升。基于此,首先搜

集试验数据及公开数据集,采用数据增强建立了混凝土表面裂缝数据集,然后分别基于注意力机制和 Transformer 架构,对 U-Net 进行改进,提出两种全监督裂缝分割模型:DA-CrackNet 和 TransUNet,并利用上述数据集进行训练和测试,对比不同模型的性能。

1 混凝土表面裂缝数据集

为建立混凝土表面裂缝数据集,先搜集原始混凝土表面裂缝图片,再进行数据增强。

原始图片主要来源于两部分:一是搜集相关文献中的试验数据^[16],即从文献中获取试验照片;二是利用公开数据集,如 Dorafshan 等^[17]。共收集了 500 张 512×512 的裂缝图像,并用 labelme 软件进行了标注,部分裂缝图像如图 1 所示。

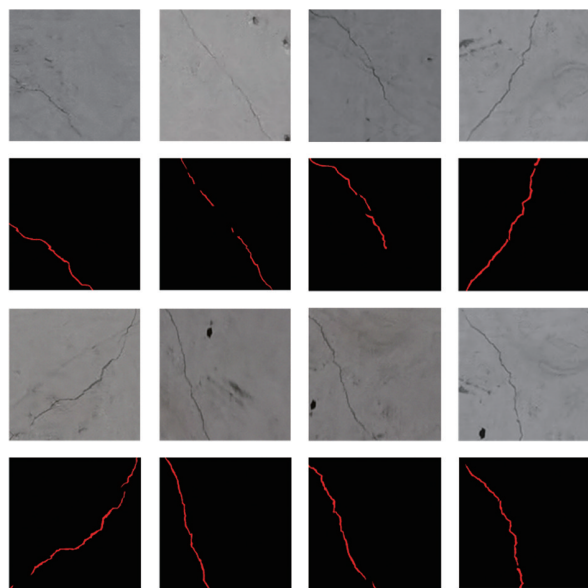


图 1 裂缝图像示例

Fig. 1 Examples of crack photos

为使数据集趋于平衡,除了常用的翻转和旋转等数据增强方法外,还提出了一种基于随机擦除算法的数据增强方法,即填充裂缝法。如图 2 所示,其操作步骤如下:(1)在裂纹掩模中随机选择裂缝,用黑色像素填充;(2)从裂纹图像中提取背景颜色;(3)用提取的背景颜色填充步骤(1)中选择的裂缝;(4)用内涂法使边缘平滑。在数据增强后,所建立的混凝土表面裂缝数据集中包含 2 000 个样本。根据留存法,以 8:1:1 的比例随机分为训练集、验证集和测试集。

研究与应用

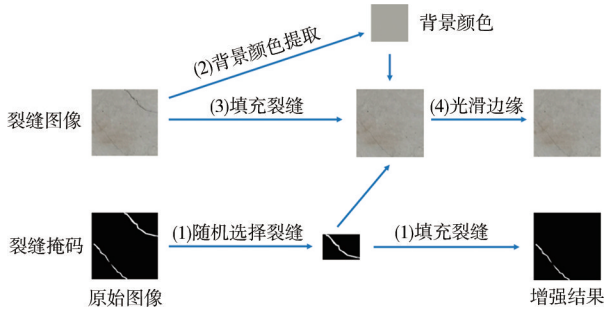


图2 裂缝填充算法

Fig. 2 The algorithm of crack filling

2 改进 U-Net 模型

基于 U-Net, 通过引入注意力机制, 并参考在大语言模型中得到广泛应用的 Transformer 架构, 提出两种改进的 U-Net 模型: DA-CrackNet 和 TransUNet。

2.1 U-Net

U-Net 由于其结构简单且具有较高的分割精度, 广泛应用于语义分割任务。其基本结构如图 3 所示, 主要包括编码器和解码器两个部分。其中, 编码器由卷积层和最大池化层组成, 在 5 个尺度上提取图像的特征, 通过融合编码器中的低层特征 (包含更多的空间信息) 和解码器中的高层特征 (包含更多的语义信息), 将拼接后的特征图经过四次上采样层并输入 1×1 的卷积层, 即可得到与输入尺寸相同的 $512 \times 512 \times 1$ 的裂缝分割结果。

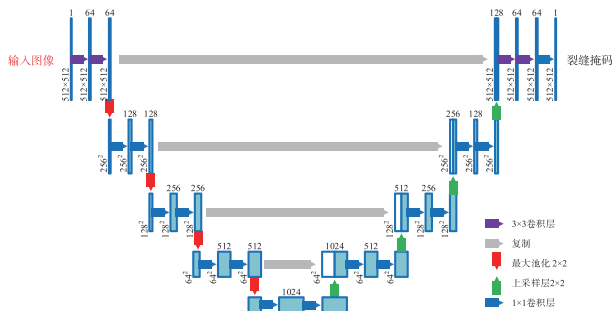


图3 U-Net 结构图

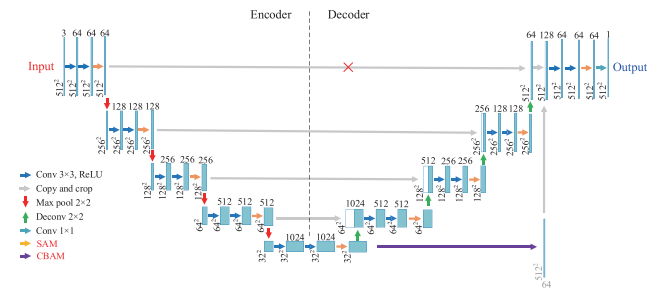
Fig. 3 The architecture of U-Net

2.2 DA-CrackNet

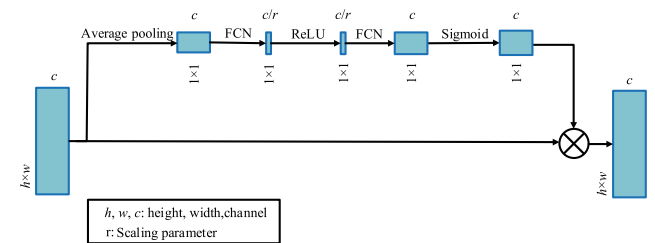
为了解决 U-Net 对小裂缝识别不足的问题, 对 U-Net 进行改进, 通过增加双注意力机制, 提出了用于裂缝分割的 DA-CrackNet 模型, 其结构如图 4 所示。

双注意力机制包括软注意力机制 (SAM) 和卷积块注意力机制 (CBAM)。如图 4 (a) 所示, SAM

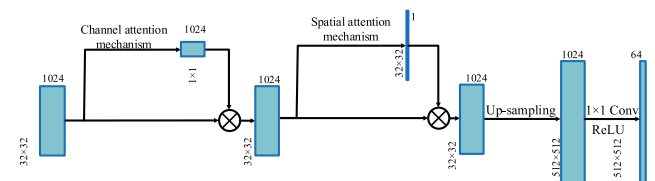
嵌入到原始 U-Net 的主干中, 用于恢复微小裂缝中的低级语义信息, 然后由解码器将其集中到恢复的高级语义特征中, 以提高微小裂缝分割的性能。图 4 (b) 显示了 SAM 的细节, 它主要通过对图像的不同通道进行加权来工作。具体来说, 高度相关的信息通道被赋予较大的权重, 而不相关的信息通道则被赋予较小的权重。与 SAM 不同, CBAM 连接到网络的最深层, 以更有效提取必要的信息, 如图 4 (a) 所示。CBAM 的结果被连接到离输出最近的卷积块上, 因此第一级的跳跃连接被相应取消。CBAM 由通道注意力机制和空间注意力机制组成, 如图 4 (d)、(e) 所示, 前者自适应地重新校准每个通道的权重, 后者是一种自适应的空间区域选择机制, 聚焦于重要的位置。



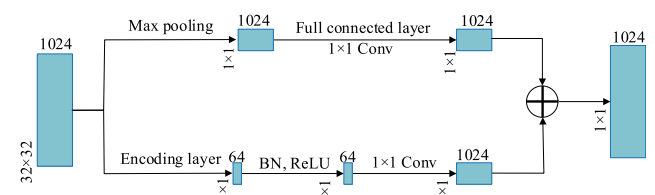
(a) DA-CrackNet 架构图



(b) 软注意力机制

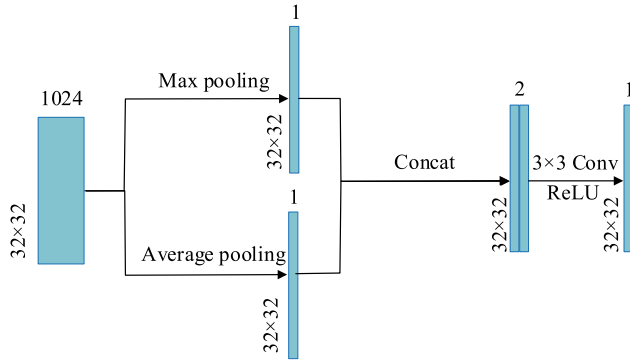


(c) CBAM 模块



(d) 通道注意力机制

研究与应用



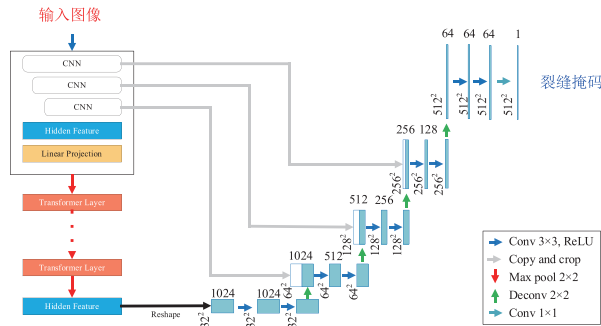
(e) 空间注意力机制

图 4 DA-CrackNet 结构图

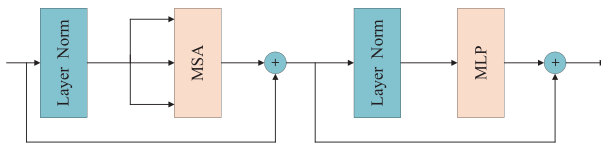
Fig. 4 The architecture of DA-CrackNet

2.3 TransUNet

Transformer 是当今应用最为广泛的大语言模型。将 Transformer 与 U-Net 结合起来,提出了用于裂缝分割的 TransUnet 模型,结构如图 5 所示。整体架构同样分为编码器与解码器,其中前者采用混合编码器,由两个关键模块组成:一个是 CNN 骨干特征提取网络,用于提取多尺度局部特征图;另一个是分层 Transformer 模块,用于提取全局特征。



(a) TransUNet 架构



(b) Transformer 层

图 5 TransUNet 结构图

Fig. 5 The architecture of TransUNet

3 模型训练

3.1 评价指标

在语义分割领域,评估模型性能的常用指标包括准确率 (Precision)、交并比 (Mean Intersection over

Union)、召回率 (Recall) 以及 F1 分数,其定义如下:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1a)$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (1b)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (1c)$$

$$\text{F1} = \frac{2TP}{2TP + FP + FN} \quad (1d)$$

式中: TP 、 TN 、 FP 和 FN 分别表示真阳性、真阴性、假阳性和假阴性的个数。

3.2 模型训练

对于裂缝分割,背景像素的数量大于裂缝区域的像素数量。传统的二值交叉熵 (BCE) 由于具有较好的抗噪声性能而被广泛用作损失函数。此外,骰子损失函数 (Dice) 为非平衡数据集的分割提供了更准确的结果。考虑到 BCE 损失函数的泛化能力和 Dice 损失函数的准确性,采用如下损失函数:

$$\text{Loss}_{\text{BCE} + \text{Dice}} = 0.5\text{Loss}_{\text{BCE}} + \text{Loss}_{\text{Dice}} \quad (2a)$$

$$\text{Loss}_{\text{BCE}} = \sum_{i=1}^N y_i \log p_i \quad (2b)$$

$$\text{Loss}_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N y_i p_i}{\sum_{i=1}^N (y_i + p_i)} \quad (2c)$$

式中: y_i 是实际值; p_i 是预测值; N 是像素数量。

模型训练和测试的环境配置见表 1。经过反复试算,最优超参数设置如下:批量大小设置为 8, epoch 次数设置为 300;使用 SGD 优化器,初始学习率设置为 10^{-3} ,动量设置为 0.9 以加速收敛;此外,采用余弦退火策略调整学习率以抑制过拟合,最小学习率设为 10^{-5} 。

表 1 环境配置

Table 1 Environmental configurations

环境	配置
CPU	AMD Ryzen9 5900X
GPU	NVIDIA GeForce RTX 3090/24GB
操作系统	Windows Server 2022
Python	3.8.8
CUDA	11.4
PyTorch	1.10.0

使用相同的超参数值,分别训练 U-Net、DA-CrackNet 以及 TransUnet 模型。三种模型的训练集和验证集的损失曲线如图 6 所示。对于 U-Net 模型,

研究与应用

当 epoch 数达到 274 时, 从验证集得到的损失值最小, 说明已经达到了最优网络; 对于 DA-CrackNet 模型, 最优网络对应的 epoch 数为 286; 对于 TransU-net 模型, 当 epoch 数达到 290 时, 从验证集得到的

损失值最小, 说明达到了最优网络。图 7 给出了 U-Net、DA-CrackNet 及 TransU-net 模型在验证集上的准确率、平均交并比、召回率及 F1 分数随训练迭代次数增加的变化曲线。

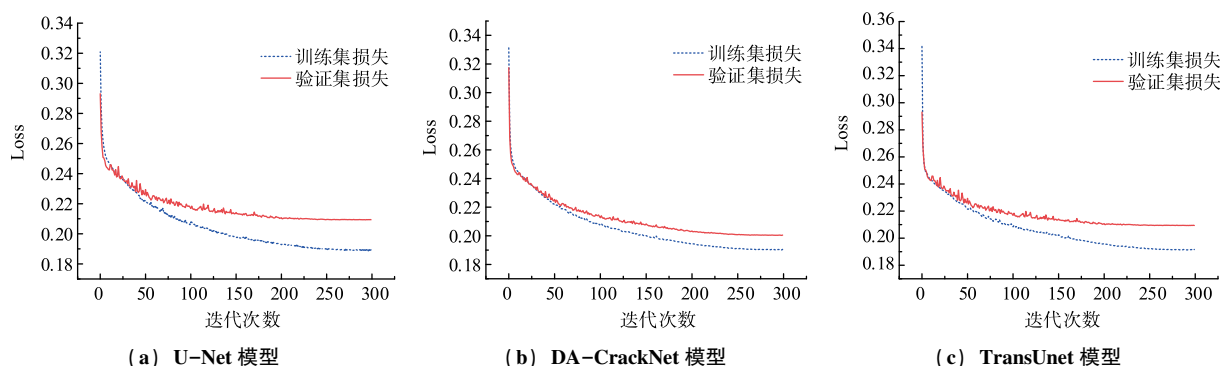


图 6 训练过程损失函数曲线
Fig. 6 The loss of the training process

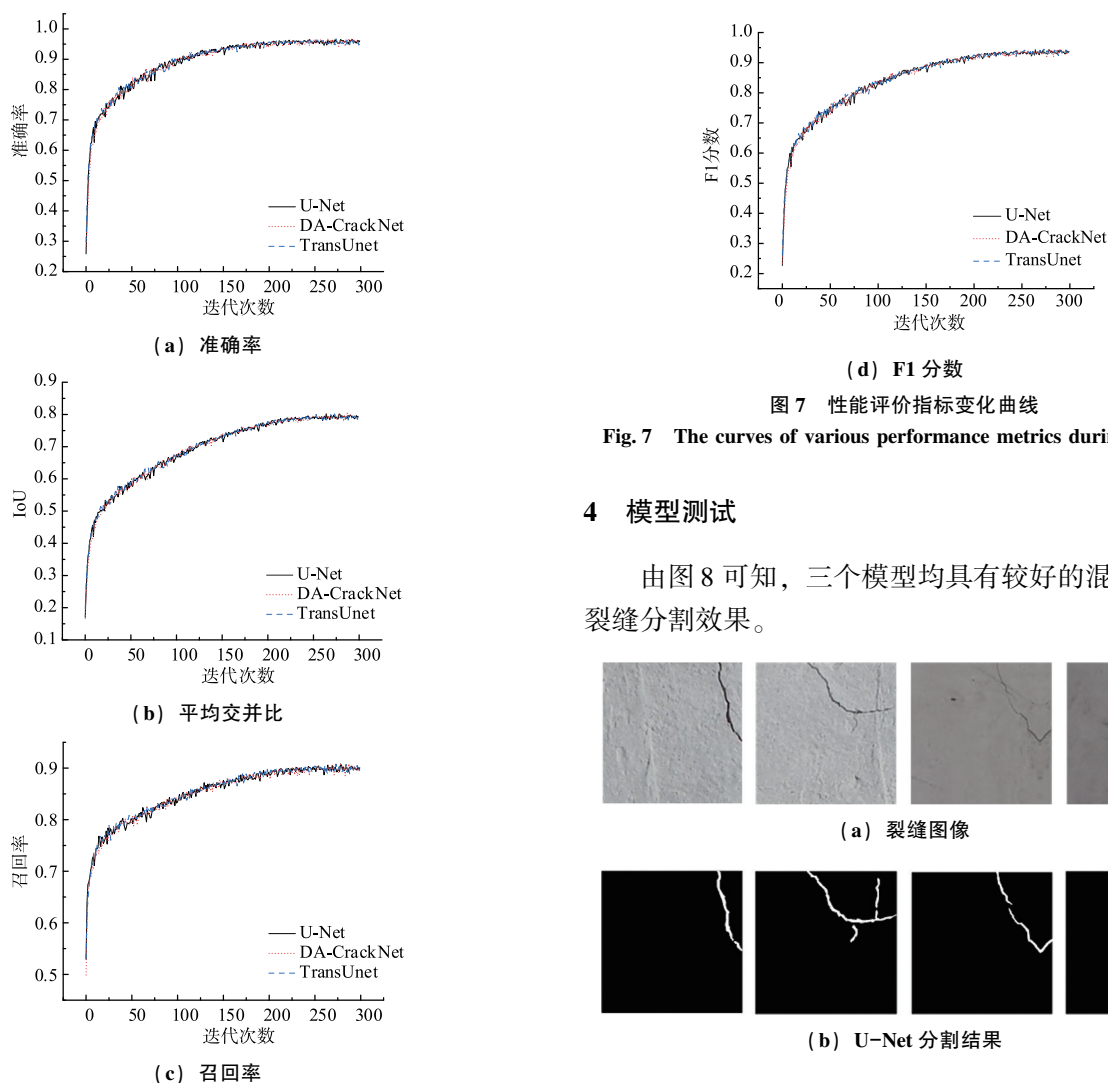
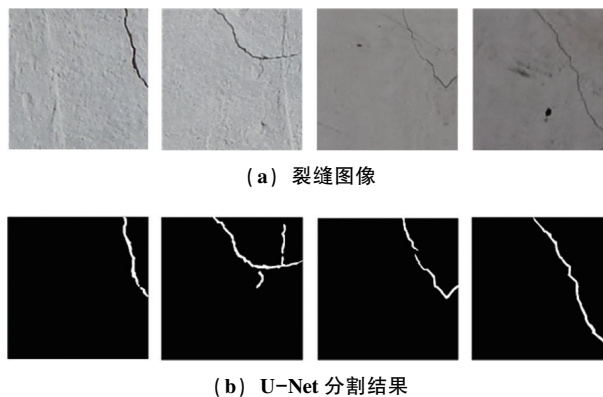


图 7 性能评价指标变化曲线

Fig. 7 The curves of various performance metrics during training

4 模型测试

由图 8 可知, 三个模型均具有较好的混凝土表面裂缝分割效果。





(c) DA-CrackNet 分割结果



(d) TransUnet 分割结果

图 8 不同模型的分割结果对比

Fig. 8 Comparative analysis of segmentation results of different models

进一步, 利用上述 4 个指标和一个 epoch 的训练时间, 对 U-Net、DA-CrackNet 和 TransUnet 模型的裂缝分割性能进行了综合对比, 结果见表 2。可以看出: 三种模型都具有优异的裂缝分割性能; 与 U-Net 相比, DA-CrackNet 和 TransUnet 的准确率分别提高了 1.41%、1.44%, mIoU 分别提高了 2.07%、2.92%, 召回率分别提高了 1.35%、1.39%, F1 分数分别提高了 1.53%、2.39%; DA-CrackNet 和 TransUnet 模型之间各性能指标差别不到 1%, 但 TransUnet 训练一次花费的时间为 104.86 s, 约为 U-Net 和 DA-CrackNet 训练时间的两倍。因此, 综合考虑计算精度和计算效率, DA-CrackNet 为最优模型。

表 2 模型性能对比

Table 2 Performance comparison between different models

模型	准确率/%	召回率/%	F1/%	交并比/%	时间/s
DA-CrackNet	96.25	79.22	77.47	72.34	53.97
TransUnet	96.28	79.26	78.33	73.19	104.86
U-Net	94.84	77.87	75.94	70.27	47.56

5 结论

基于 U-Net 模型, 分别引入注意力机制和参考 Transformer 架构, 提出了两种用于混凝土表面裂缝分割的改进 U-Net 模型: DA-CrackNet 和 TransUnet, 并通过混凝土表面裂缝数据集进行了训练和测试, 综合对比了不同模型的性能, 主要结论如下:

(1) 提出了一种基于随机擦除算法的数据增强方法, 即填充裂缝法, 建立了混凝土表面裂缝数据集。

(2) 通过引入软注意力机制和卷积块注意力机制, 提出了用于裂缝分割的 DA-CrackNet 模型, 以

解决 U-Net 对小裂缝识别不足的问题。

(3) 将 Transformer 与 U-Net 结合起来, 提出了用于裂缝分割的 TransUnet 模型, 整体架构可分为编码器与解码器, 其中前者采用混合编码器, 包括 CNN 骨干特征提取网络和分层 Transformer 模块。

(4) 在测试集上, DA-CrackNet 及 TransUnet 的裂缝分割性能均优于 U-Net。DA-CrackNet 和 TransUnet 之间各性能指标差别不大, 但前者计算效率更高。综合考虑计算精度和计算效率, DA-CrackNet 性能最优。

参 考 文 献

- [1] 赵发宾. 基于计算机方法的混凝土识别与应用研究 [J]. 粉煤灰综合利用, 2020, 34 (3): 29-32.
- [2] 陈军, 王青, 周辉阳. 装配式工程易变形结构裂缝红外微弱目标识别方法 [J]. 粉煤灰综合利用, 2023, 37 (4): 122-127.
- [3] 章毓晋. 计算机视觉教程 [M]. 北京: 人民邮电出版社, 2017.
- [4] 沈俊凯. 基于计算机视觉的混凝土裂缝检测算法研究 [D]. 哈尔滨: 中国地震局力学工程研究所, 2019.
- [5] 周颖, 刘彤. 基于计算机视觉的混凝土裂缝识别 [J]. 同济大学学报 (自然科学版), 2019 (9): 1277-1286.
- [6] PENG C, YANG M Q, ZHENG Q H, et al. A triple-thresholds pavement crack detection method leveraging random structured forest [J]. Construction and Building Materials, 2020, 263.
- [7] LIU X F, HONG Z L, SHI W, et al. Image-processing-based subway tunnel crack detection system [J]. Sensors, 2023, 23 (13).
- [8] LIU X, BO L, LIU Y, et al. Detection of micro-cracks using nonlinear lamb waves based on the Duffing-Holmes system [J]. Journal of Sound and Vibration, 2017, 405: 175-186.
- [9] SHI Y, CUI L M, QI Z Q, et al. Automatic road crack detection using random structured forests [J]. Ieee Transactions on Intelligent Transportation Systems, 2016, 17 (12): 3434-3445.
- [10] LIU Y Q, YEOH J. Automated crack pattern recognition from images for condition assessment of concrete structures [J]. Automation in Construction, 2021, 128 (2): 103765.
- [11] LIU Y, YEOH J K W. Robust pixel-wise concrete crack segmentation and properties retrieval using image patches [J]. Automation in Construction, 2020, 123: 103535.
- [12] DUNG C V, ANH L D. Autonomous concrete crack detection using deep fully convolutional neural network [J]. Automation in Construction, 2019, 99: 52-58.
- [13] LAU S, CHONG E, YANG X, et al. Automated pavement crack segmentation using U-Net-based convolutional neural network [J]. Ieee Access, 2020, 8: 114892-114899.

研究与应用

- [14] REDMON J, FARHADIA. YOLOv3: An Incremental Improvement [Z]. arXiv e-prints, 2018.
- [15] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [Z]. arXiv, 2017.
- [16] 陈俊玮, 刘汉赋, 袁焕鑫, 等. 地铁车辆段检修库装配式混凝土空心柱抗震性能研究 [J]. 粉煤灰综合利用, 2024, 38 (3): 113-119.
- [17] DORAFSHAN S, THOMAS R J, MAGUIRE M. SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks [J]. Data in Brief, 2018, 21: 1664.

(上接第 156 页)

- [10] 史乐贤, 张永辉, 周大兴, 等. 新型玄武岩纤维增强混凝土的性能研究 [J]. 混凝土世界, 2026 (199): 33-36.
- [11] 王伟, 洪雅璐, 姜屏, 等. 玄武岩纤维增强混凝土的研究进展: 力学性能, 耐久性能, 微观机制 [J/OL]. 复合材料学报, 1-27 [2026-05-25]. <https://doi.org/10.13801/j.cnki.fhclxb.20260114.002>.
- [12] 吴闻秀, 田浩辰, 张洪光, 等. 玄武岩纤维混凝土力学强度与多尺度增韧机理研究 [J]. 高科技纤维与应用, 2025 (6): 71-82.
- [13] 戴博玄. 玄武岩纤维混凝土在路桥工程伸缩缝施工中的应用 [J]. 汽车画刊, 2025 (11): 226-228.
- [14] 刘欣鑫, 陈庆, 韩非, 等. 玄武岩纤维混凝土抗硫酸盐侵蚀性能及侵蚀深度预测模型 [J]. 混凝土, 2025 (12): 11-15.
- [15] 谢福君, 张家生. 基于实用效应的桥梁加固或拆除重建方案比选 [J]. 铁道科学与工程学报, 2019, 16 (7): 1714-1718.

E-mail: fmhzhly@163.com