

DOI: 10.11835/j.issn.2096-6717.2022.145



开放科学(资源服务)标识码 OSID:



基于航拍图像与改进U-Net的建筑外墙裂缝检测方法

刘少华¹, 任宜春¹, 郑智雄², 牛孜颀²

(1. 长沙理工大学土木工程学院, 长沙 410114; 2. 中国建筑第五工程局有限公司, 长沙 410007)

摘要: 针对建筑外墙裂缝人工检测方法检测效率低、检测效果和安全性差的问题, 提出基于航拍图像与计算机视觉的裂缝检测方法。使用无人机绕建筑物航拍采集裂缝图像, 并构建裂缝数据集; 优化U-Net模型以解决细长裂缝分割不连续、复杂背景下裂缝漏检及背景误检的问题。将模型编码网络替换为预训练的ResNet50以提升模型特征表达能力; 添加改进的多孔空间金字塔池化(Atrous Spatial Pyramid Pooling, ASPP)模块, 获取多尺度上下文信息; 采用改进的损失函数处理裂缝图像正负样本分布极度不均的问题。结果表明: 改进的U-Net模型解决了原模型存在的问题, IoU指标和F1_score分别提升了3.53%、4.18%; 与经典分割模型相比, 改进模型的裂缝分割性能最优。与人工检测方法相比, 所提方法可以安全、高效且准确地进行建筑外墙裂缝检测。

关键词: 建筑外墙裂缝; 无人机; 计算机视觉; 语义分割; U-Net

中图分类号: TU17 **文献标志码:** A **文章编号:** 2096-6717(2024)01-0223-09

Building exterior wall crack detection based on aerial images and improved U-Net

LIU Shaohua¹, REN Yichun¹, ZHENG Zhixiong², NIU Ziyang²

(1. School of Civil Engineering, Changsha University of Science and Technology, Changsha 410114, P. R. China;

2. China Construction Fifth Engineering Bureau Co., Ltd, Changsha 410007, P. R. China)

Abstract: Aiming at the problems of low efficiency, unsatisfactory detection effect and poor safety of manual detection methods for building exterior wall cracks, a crack detection method based on aerial images and computer vision was proposed. Firstly, the Unmanned Aerial Vehicle (UAV) was used to collect the crack images through aerial photography around the buildings, and a crack dataset was constructed. Secondly, the U-Net was optimized to solve the problems of discontinuous segmentation of slender cracks as well as the missed and false detection under complex backgrounds. The encoder was replaced with pre-trained ResNet50 to improve the feature expression ability of the model. An improved Atrous Spatial Pyramid Pooling (ASPP)

收稿日期: 2022-09-27

基金项目: 湖南省自然科学基金(2021JJ30716); 湖南省高新技术产业科技创新引领计划(2020KG2026); 长沙理工大学土木工程优势特色重点学科创新性项目(16ZDXK05)

作者简介: 刘少华(1999-), 男, 主要从事计算机视觉与健康监测研究, E-mail: liushaohua2020@163.com。

Received: 2022-09-27

Foundation items: Natural Science Foundation of Hunan Province (No. 2021JJ30716); High-Tech Industry Science and Technology Innovation Leading Plan Project of Hunan Province (No. 2020KG2026); Civil Engineering Advantage Characteristic Key Discipline Innovation Project of Changsha University of Science and Technology (No. 16ZDXK05)

Author brief: LIU Shaohua (1999-), main research interests: computer vision and health monitoring, E-mail: liushaohua2020@163.com.

module was added to obtain multi-scale context information. The improved loss function was used to deal with the problem of extremely uneven distribution of positive and negative samples in crack images. Experiments show that the improved U-Net model solved the problems existing in the original model; the IoU and F1_score were increased by 3.53% and 4.18%, respectively. Compared with the classical segmentation model, the improved model has the best crack segmentation performance. Compared with manual detection methods, it can efficiently, accurately, and safely detect building exterior wall cracks.

Keywords: building exterior wall cracks; UAV; computer vision; semantic segmentation; U-Net

建筑外墙裂缝会使带腐蚀性的雨水渗入墙体,导致墙体涂料与防水材料等发霉、剥落,不仅影响建筑美观,甚至锈蚀墙内钢筋等材料,破坏墙体结构,带来安全隐患。准确检测出裂缝位置,可以为制定有效的维护方案提供依据。目前,建筑外墙裂缝主要采用人工检测,这是一种高空作业,不仅危险,而且检测效率和精度都很低。如何安全、高效地检测建筑外墙裂缝成为亟须解决的问题。

计算机视觉为裂缝检测提供了一种非接触式的检测技术^[1],包括图像处理和深度学习两类方法。基于图像处理的裂缝检测方法根据裂缝的纹理和边缘等特征,人为设计裂缝检测的依据,实现裂缝检测^[2]。这种方法能够在特定的数据集上取得良好的检测效果,但是裂缝的检测精度不高,且需要针对特定的图像设计专门的裂缝检测算法。近年来,基于深度学习的方法在裂缝检测领域取得了显著的成果,根据裂缝检测方式的不同可将其分为3类:图像分类、目标检测与语义分割。基于图像分类^[3]与目标检测^[4]的裂缝检测方式,只能识别并定位裂缝,不能在像素级别上标记裂缝,从而无法获得裂缝形状信息;基于语义分割的裂缝检测方法是对图像中的每个像素点进行分类,能将裂缝像素准确标记出来,既可以检测有无裂缝,又能得到裂缝的位置、形状等信息^[5]。经典语义分割模型 U-Net^[6]、DeepLabv3+^[7]、PSPNet^[8]及其改进模型在桥梁、路面及隧道等领域的裂缝检测中得到广泛应用。李良福等^[9]提出了改进的 PSPNet 模型,解决了桥梁裂缝分割细节信息丢失的问题;Li 等^[10]提出了一种可训练的上下文编码网络,实现了复杂条件下桥梁裂缝的检测;曹锦纲等^[11]提出了一种基于注意力机制的网络,实现了自动准确地路面裂缝检测;Nguyen 等^[12]提出了一种两级卷积神经网络来处理含噪声和低质量的道路图像;孟庆成等^[13]用改进的 MobileNet_v2 替换 U-Net 编码器,实现了对混凝土裂缝的实时检测;Dang 等^[14]提出了改进的 U-Net 模型,在满足裂缝分割性能要求的基础上支持裂缝自动测量。上述模型在裂缝检测中的成功应用,表明深度学习方法在裂缝检测领域具有广阔的应用

前景。

建筑外墙裂缝与桥梁、路面等混凝土裂缝拥有共同特征,然而建筑外墙因其材料多样性,且墙体包含复杂建筑构件,导致建筑外墙裂缝图像背景更加复杂,存在更多的背景噪声。因此,有必要为建筑外墙裂缝检测针对性地设计一个深度学习模型。基于深度学习的裂缝检测方法在训练和预测阶段都依赖原始图像数据,人工采集大型建筑的外墙图像需高空作业,存在安全隐患大、效率低等问题。近年来,无人机技术取得快速发展,无人机搭载传感器为高空采集图像提供了一种安全且高效的方式,在抗震救灾、农业植保等领域得到广泛应用^[15]。笔者使用无人机航拍采集建筑外墙图像,并根据裂缝图像特征设计了改进的 U-Net 模型用于裂缝检测。

1 裂缝数据集构建

1.1 图像采集与处理

图像采集包括裂缝数据集构建和裂缝检测两个阶段。在构建数据集时,为了保证训练后模型的鲁棒性,需要广泛采集各类型建筑外墙图像,包括不同光照条件、不同拍摄角度及不同航拍间距等。在裂缝检测阶段,需要根据裂缝检测精度要求来设计图像采集方案,规划采集路线、航拍间距等以确定裂缝位置及图像像素比例尺,从而指导裂缝修复与建筑维护。

广泛采集原始建筑外墙图像以构建建筑外墙裂缝数据集。共得到建筑外墙裂缝图像 566 张,分辨率大小为 $8\,000 \times 6\,000$ 。由于采集的图像分辨率过大,不利于模型训练,且裁剪后的裂缝图像细节信息更加丰富,于是对图像进行裁剪。利用滑动窗口方法将原图裁剪为 512×512 大小的子图,第 1 步:从图像左上角开始,沿着自左向右、自上而下的顺序依次裁剪,得到子图 165 张,保证子图覆盖原图,避免信息丢失;第 2 步:在图像左下角 $7\,488 \times 5\,488$ 范围内随机选择 35 个点作为子图左下角进行裁剪,得到子图 35 张,用于信息补充。在完成所有图像裁剪后,剔除非建筑外墙图像及不含裂缝的背

景图像后共得到裂缝图像 3 862 张。

为了确保模型的泛化性能和复杂背景下裂缝分割的鲁棒性,必须保证裂缝图像数据中裂缝样本及图像背景的多样性和复杂度。无人机航拍包括不同天气状况、一天的不同时间段,因光照条件不

同,得到明暗程度不同的裂缝图像;裂缝图像包含的裂缝类型全面,包括纵向、横向、斜向、交叉及任意分布的裂缝图像,且图像背景丰富,包括不同材料建筑外墙下砖缝、门窗、排水管、阴影及杂物等背景,部分裂缝图像如图 1 所示。

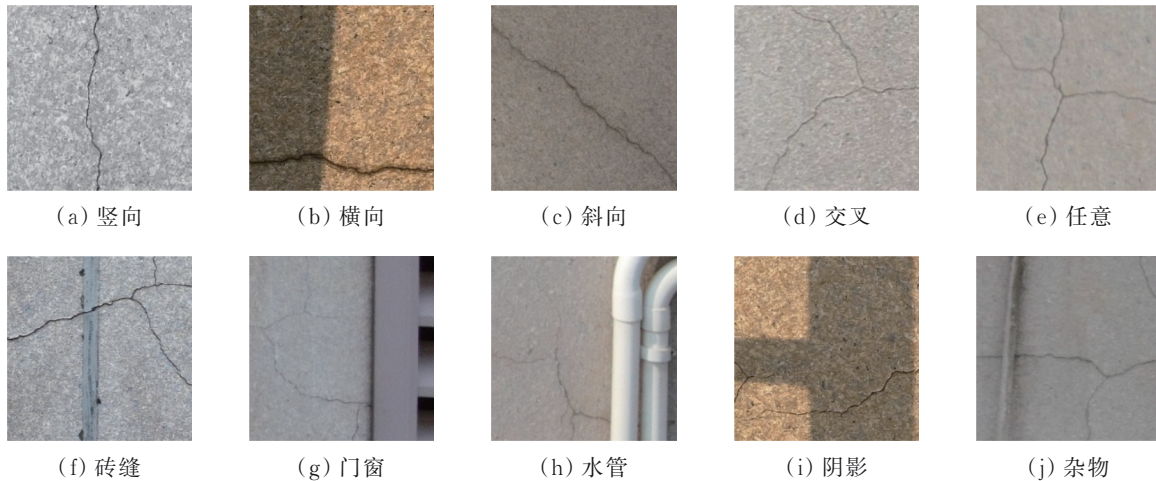


图 1 裂缝样本

Fig. 1 Crack samples

1.2 图像标注

为了保证数据集的可靠性,使用开源的分割任务标注软件 labelme 对裂缝进行人工标注。首先,沿着裂缝边缘逐点勾勒出裂缝的轮廓,生成一个包含裂缝位置坐标的 json 文件,然后用 python 代码将其转化为可供计算机学习的二值图像,标记结果如图 2 所示。

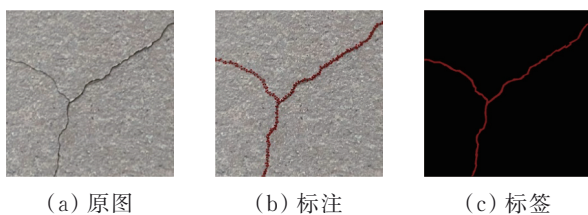


图 2 裂缝标注

Fig. 2 Crack labeling

在完成裂缝数据集构建后,按照 8:1:1 的比例将数据集划分为训练集、验证集和测试集,其中训练集 3 102 张,验证集与测试集各 380 张。

2 研究方法

2.1 U-Net 模型

U-Net 是在全卷积神经网络的基础上提出用于医学图像分割的网络模型。该模型是一个对称编解码结构,包括编码器、解码器两个部分。其独特之处在于将编码器中低分辨率特征图通过跳跃连接直接拼接到对应解码器上采样生成的特征图,从

而有效融合浅层纹理特征和深层语义特征,有利于图像分割。

编码器是由两个无填充的 3×3 卷积重复应用组成,每一个卷积后面有一个 ReLU 激活函数,两次卷积之后有一个 2×2 最大池化,用于下采样。编码器一共包含 4 个下采样,每一个下采样中,将特征通道的数量增加一倍,特征图大小减为一半。编码器主要负责提取输入图像的特征信息,并将多个尺度的特征图传入解码器。

解码器是模型的加强特征提取网络,利用转置卷积对主干网络获取的 5 个初步有效特征层进行上采样,并进行特征融合,最终得到一个融合了所有特征的有效特征层。

2.2 改进的 U-Net 模型

研究表明:U-Net 模型存在细长裂缝分割不连续、复杂背景下裂缝漏检及背景误检的问题。分析其可能原因为:1) 裂缝图像正负样本分布不均对模型性能造成影响;2) 模型特征提取能力不足导致裂缝漏检及误检;3) 缺少获取多尺度信息的能力导致裂缝分割不连续及背景误检。在分析原因基础上提出了基于 U-Net 的改进模型,相对于原模型做出了 3 点具有针对性的改进:1) 使用改进的损失函数处理裂缝图像正负样本分布极度不均衡的问题;2) 使用预训练的 ResNet50 网络替换原模型的编码网络,以提升模型特征表达能力;3) 添加改进的 ASPP 模块,获取多尺度上下文信息。网络结构如

图 3 所示。

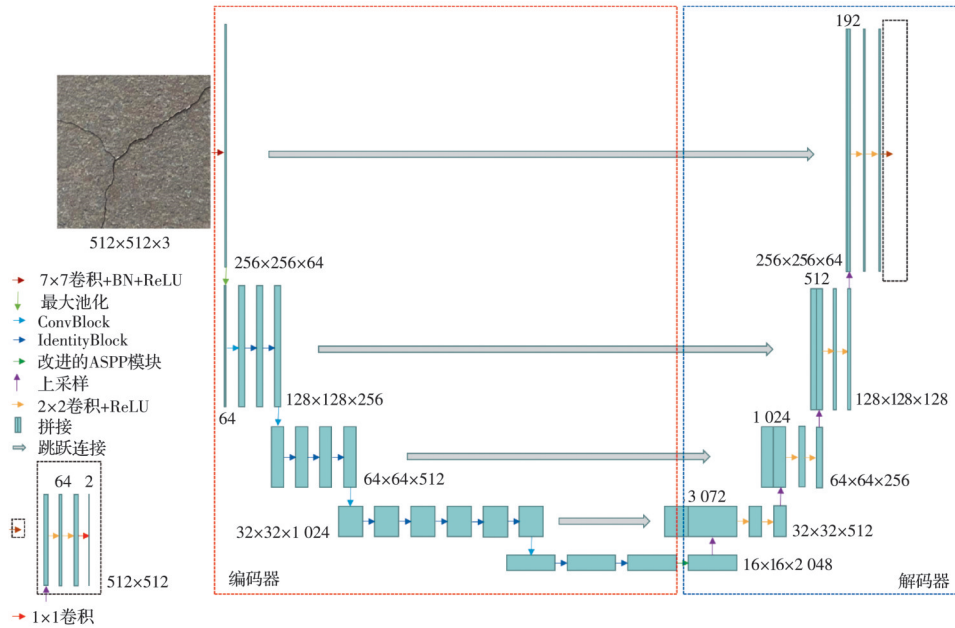
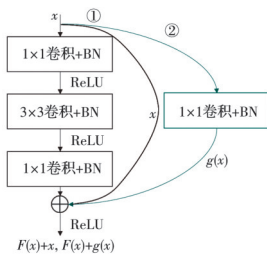


图 3 改进的 U-Net 模型

Fig. 3 Improved U-Net model

2.2.1 残差神经网络

理论上,神经网络的层数越多即深度学习模型越深,网络的特征表达能力越强。但神经网络达到一定深度时,继续加深会因为反向传播过程梯度消失导致模型性能变得更差。为了解决这个问题,He 等^[16]提出了用跳跃连接解决网络加深造成梯度消失的残差神经网络(ResNet),包括 Conv Block 和 Identity Block 两个基本模块,Conv Block 改变网络的维度,用来调整特征图大小和层数;Identity Block 不改变网络维度,用来加深网络,结构如图 4 所示。



注:①为 Identity Block;②为 Conv Block。

图 4 残差模块

Fig. 4 Residual Module

基于深度学习的方法对建筑外墙裂缝图像分割,然而裂缝的标注难度较大,自制大型数据集成本太高且不实际,为了克服深度学习方法依赖大量训练样本的问题,引入迁移学习方法。基于特征提取的迁移,用 ImageNet 数据集上预训练好的深度残差网络(删除最后的平均池化层、全连接层和 softmax 层),替换 U-Net 网络原来由两层卷积加一层池化组成的编码网络。

2.2.2 ASPP 模块

受空间金字塔池化(Spatial Pyramid Pooling, SPP)模块^[8]启发,Chen 等^[7]提出了 ASPP 模块,结构如图 5 所示。

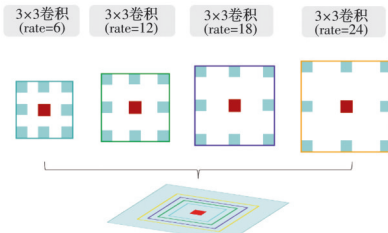


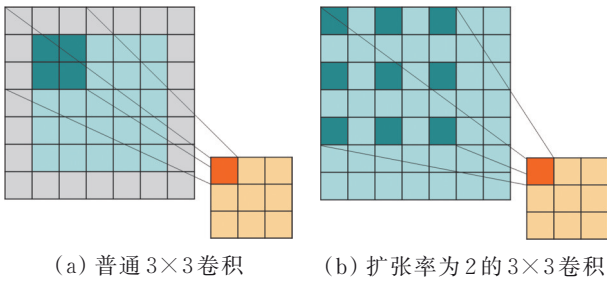
图 5 ASPP 模块

Fig. 5 ASPP module

采用并行的多个具有不同扩张率的扩张卷积,每个分支使用不同扩张率单独进行图像特征提取,并将各分支提取特征进行融合,从而提取得到多尺度的图像信息。该模块可以有效避免局部信息丢失,增加远距离信息的相关性。

扩张卷积是在普通卷积中增加“空洞”来扩大感受野的卷积,相对于普通卷积增加了“扩张率”这个超参数。通过设置扩张率的大小,可以实现相同参数量和计算量下更大的感受野。如图 6 所示,扩张率为 2 的 3x3 卷积核与普通 3x3 卷积核相比,感受野由 3x3 扩大为 5x5。

值得注意的是,组合扩张卷积的扩张率设置不当会产生“网格效应”。当卷积核大小为 3x3,组合扩张率设为[2,2,2]时,像素利用情况如图 7(a)所示(数字表示像素利用次数),此时,大量像素未被利

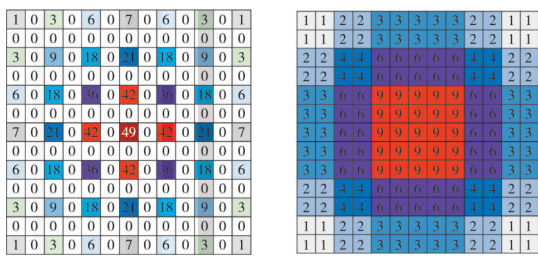


(a) 普通 3×3 卷积 (b) 扩张率为 2 的 3×3 卷积

图 6 感受野对比图

Fig. 6 Comparison of receptive field

用(数字 0 处),产生了“网格效应”;组合扩张率设为 [1, 2, 3] 时,像素利用情况如图 7(b) 所示,此时像素利用全面,未产生“网格效应”。



(a) 空洞率为 [2, 2, 2] (b) 空洞率为 [1, 2, 3]

图 7 像素利用情况

Fig. 7 Pixel utilization

Wang 等^[17]提出了混合扩张卷积 (Hybrid Dilated Convolution, HDC) 理论,用于解决“网格效应”问题。假设有 n 个卷积核大小为 $k \times k$, 扩张率为 $[r_1 \dots r_i \dots r_n]$ 的扩张卷积,定义两个非零值之间的距离为

$$M_i = \max [M_{i+1} - 2r_i, M_{i+1} - 2(M_{i+1} - r_i), r_i] \quad (1)$$

式中: $M_n = r_n, M_2 \leq k$ 。另外,还需保证扩张率的最大公约数为不能大于 1。

由于裂缝呈线性分布的特点,既有小目标物体的特点又有大目标物体的特点,且裂缝图像含有多尺度背景噪声。原 U-Net 模型对细长裂缝预测存在分割不连续及背景噪声误检的问题。

为了解决上述问题,在模型编解码之间添加 ASPP 模块。ASPP 模块扩张率设计基于 HDC 理论,并借鉴 DeepLabv3+ 中 ASPP 模块结构。设计了扩张率分别为 [2, 3, 5]、[2, 3, 7]、[2, 3, 9]、[2, 3, 12] 的 ASPP 模块进行消融实验。

2.2.3 损失函数

在建筑外墙裂缝图像中,裂缝像素所占比例极小,约为 1.5%~4%,造成正(裂缝)负(背景)样本分布极度不均衡,且样本中存在大量易分类的负样本。研究发现,经典的二分类损失函数在裂缝图像语义分割模型中,易分类的负样本会对损失起主要

贡献作用,导致模型对裂缝特征的学习被抑制,模型的性能和优化方向不稳定。

为了解决现有损失函数在裂缝分割模型中的不足,提出使用 Dice Loss^[18]作为正则项与 Focal Loss^[19]结合的新损失函数(FD Loss),如式(2)所示。

$$L_{FD} = wL_{Focal} + L_{Dice} \quad (2)$$

式中: w 为数量级校正因子,保证两者处在相同数量级上。

Focal Loss 是基于二分类交叉熵损失函数改进而来的处理正负样本分布不均,并让模型在训练时更专注于难分类样本的损失函数,如式(3)所示。

$$L_{Focal} = -\frac{1}{N} \sum_{n=1}^N [-y\alpha(1-\hat{y})^\gamma \log(\hat{y}) + (1-y)(1-\alpha)\hat{y}^\gamma \log(1-\hat{y})] \quad (3)$$

式中: N 为像素点总数; \hat{y} 与 y 分别表示第 n 个像素的预测值与真实值,通过设置 α 来控制正负样本对总损失的共享权重;通过设置 γ 来减少易分类样本的权重,使模型在训练时更专注于难分类的样本,本文 α 设为 0.25、 γ 设为 2。

Dice Loss 源于 Dice 系数,可以缓解正负样本在数量上不平衡的问题,并优化 F_{1_score} ,如式(4)所示。

$$L_{Dice} = \frac{1}{N} \sum_{n=1}^N 1 - \frac{2|\hat{y} \cap y|}{|\hat{y}| + |y|} \quad (4)$$

式中: N 为像素点总数; \hat{y} 与 y 分别表示第 n 个像素的预测值与真实值; $|\cdot|$ 表示其参数的所有矩阵元素的和; \cap 表示逐个元素的乘法运算。

3 实验设计与分析

实验环境配置如表 1 所示。

表 1 实验环境配置

Table 1 Experimental environment configuration

配置	参数
操作系统	Ubuntu 18.04
GPU	NVIDIA RTX A4000
实验环境	PyTorch 1.7.1, CUDA 11.0, cuDNN 8.0
算法语言	Python 3.8

模型训练 100 个 epoch,前 50 个 epoch 冻结编码器,对解码器进行训练。这样可以避免因为编码器预训练权重与未训练的解码器权重相差太大导致编码器预训练权重被破坏,同时还可以加快模型训练速度。批大小设为 16,初始学习率设为 0.000 1。后 50 个 epoch 对整个模型进行训练,批大小设为 8,初始学习率设为 0.000 01。各模型损失曲线与 mIoU 曲线如图 8 所示。总体上,模型大约在 85 个 epoch 后达到收敛。

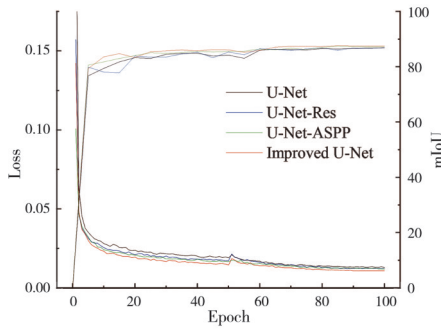


图 8 各模型损失曲线与 mIoU 曲线

Fig. 8 Loss curve and mIoU curve of each model

3.1 评价指标

对于二分类问题,在进行结果评价时,可将样例根据其真实类别与预测类别的组合划分为真正例(TP)、假正例(FP)、真反例(TN)、假反例(FN)4种情形。

使用交并比(IoU)、精准率(Precision)、召回率(Recall)和 F_{1_score} 这 4 个指标,对网络模型的分割性能进行定量评估。

- 1) 交并比: $IoU = \frac{TP}{TP + FP + FN}$
- 2) 精准率(即查准率): $Precision = \frac{TP}{TP + FP}$
- 3) 召回率(即查全率): $Recall = \frac{TP}{TP + FN}$
- 4) F_1 得分: $F_{1_score} = \frac{2 \times Precision \times Recall}{Precision + Recall}$

IoU 突出预测结果与标签的重合度,是常见的图像分割评价指标;Precision 突出预测结果中误检像素的比例;Recall 突出漏检的比例; F_{1_score} 综合考虑了 Precision 与 Recall, F_{1_score} 值越大代表模型性能越好。

3.2 实验分析

3.2.1 损失函数对比实验

分别使用 Focal Loss、Dice Loss 以及 FD Loss 作为模型的损失函数,在自制数据集上训练,并在测试集上测试,得到实验结果如图 9 所示。

由图 9 可知, Focal Loss 从局部进行考察,其 Precision 较高, Recall 相对较低; Dice Loss 从全局进

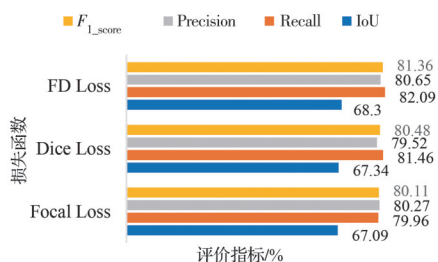


图 9 损失函数实验结果

Fig. 9 Experimental results of loss function

行考察, Recall 较高, Precision 较低。FD Loss 结合两者共同优化模型, 引导模型更加注重对难分类的正(裂缝)样本进行学习, 并提高预测裂缝区域与标注裂缝区域的重合度。相较于单独使用 Focal Loss 和 Dice Loss, FD Loss 在裂缝数据上表现出更优异的效果, 且训练过程更加稳定。

3.2.2 编码网络对比实验

在 U-Net 模型基础上, 将模型编码网络分别替换为预训练的 ResNet18、ResNet34、ResNet50 及 ResNet101 进行实验(模型使用 FD Loss 作为损失函数), 实验结果如表 2 所示。

表 2 编码网络实验结果

Table 2 Experimental results of encoder

残差网络	IoU/%	Recall/%	Precision/%	$F_{1_score}/%$
ResNet18	66.95	82.87	78.01	80.37
ResNet34	68.35	83.39	78.92	81.09
ResNet50	68.46	84.46	79.52	81.92
ResNet101	65.95	78.91	77.68	78.29

由表 2 可知, 随着编码网络的加深, 各评价指标都逐步提升, 说明加深编码网络提高了模型特征提取能力, 模型对裂缝的分割性能得到优化; 然而编码网络使用 ResNet101 时, 各评价指标反而大幅下降, 说明此时的模型复杂度与数据集大小及裂缝的分割难度不匹配, 模型出现了过拟合现象。选择 ResNet50 作为迁移网络, 在增强模型特征提取能力的同时避免网络过深导致模型与分割任务复杂度不匹配, 造成过拟合。

3.2.3 ASPP 模块扩张率对比实验

在 U-Net 模型基础上, 分别在编解码网络间添加 3 种不同扩张率的 ASPP 模块进行实验(模型使用 FD Loss 作为损失函数), 实验结果如表 3 所示。

表 3 扩张率实验结果

Table 3 Experimental results of dilated rate

扩张率	IoU/%	Recall/%	Precision/%	$F_{1_score}/%$
[2,3,5]	67.28	78.94	81.26	80.08
[2,3,7]	68.42	79.14	83.32	81.18
[2,3,9]	68.62	81.66	85.63	83.60
[2,3,12]	66.10	77.95	81.45	79.66

由表 3 可知, 为了避免“网格效应”, 根据 HDC 理论设计的 ASPP 模块为两个较小扩张率加一个较大扩张率的组合, 随着较大扩张率由 5、7、9 逐渐增大时, 各项评价指标均逐渐上升, 说明组合的较大扩张率增大有助于提升模型的裂缝分割性能。当扩张率设置为 [2, 3, 12] 时, 评价指标反而减小。这是因为其 $M_2=6>3$, 不满足 HDC 理论, 导致模型性能下降。当扩张率设为 [2, 3, 9] 时, 各项评价指标均

达到最高,模型分割性能最优。基于此,改进的 ASPP 模块选择扩张率为[2,3,9]的一组扩张卷积,结构如图 10 所示。

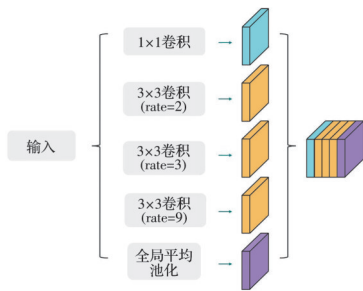


图 10 改进的 ASPP 模块

Fig. 10 Improved ASPP module

改进的 ASPP 模块首先使用一个 1×1 的普通卷积作为第 1 层,保留原特征图以补充空间信息;然后使用扩张率为[2,3,9]的扩张卷积作为 2 到 4 层,提取多尺度物体的特征;最后使用全局平均池化作为第 5 层,获取全局特征。

3.2.4 消融实验

为了验证改进策略对提高模型分割性能的有效性,在自制数据集上进行消融实验。分别验证进行迁移学习、添加改进的 ASPP 模块以及二者组合对模型分割效果的影响(模型都使用 FD Loss 作为损失函数),实验结果如表 4 所示。

表 4 消融实验结果

Table 4 Ablation experimental results

U-Net	ResNet50	ASPP 模块	IoU/%	Recall/%	Precision/%	$F_{1_score}/\%$
✓			65.53	77.68	82.91	80.21
✓	✓		68.46	84.46	79.52	81.92
✓		✓	68.62	81.66	85.63	83.60
✓	✓	✓	69.06	82.57	86.30	84.39

比较实验结果可知,在改进损失函数的基础上用 ResNet50 替换编码网络,Recall 提升了 6.78%,同时 IoU、Precision 和 F_{1_score} 都有相应的提升,说明引入迁移学习的 U-Net 模型可以提取更丰富的裂缝特征,减少了部分裂缝不能被准确分割的状况;在改进损失函数的基础上添加改进的 ASPP 模块, Precision 提升了 2.72%,同时 IoU、Recall 和 F_{1_score} 都有相应的提升,说明改进的 ASPP 模块获取多尺度上下文信息的能力在 U-Net 模型中得到利用,降低了将背景误检为裂缝的状况,保证了裂缝分割的连续性;改进的 U-Net 模型是在改进损失函数的基础上引入迁移学习方法并加入改进的 ASPP 模块,与单个改进策略相比,以上评价指标都有一定的提升,说明用 ResNet50 替代编码网络与添加改进的

ASPP 模块可以共同优化模型,并提升模型性能。与原模型相比, IoU 指标和 F_{1_score} 分别提升了 3.53%、4.18%,说明所提改进策略有效提升了模型的裂缝分割性能。

3.2.5 经典语义分割模型对比

为了验证改进模型的先进性和鲁棒性,在自制建筑外墙裂缝数据集上复现了经典语义分割模型 U-Net、DeepLabv3+ 及 PSPNet,并与模型进行对比。为了确保实验有效性,各模型在相同实验参数下进行训练。通过评价指标与分割结果进行定性与定量地对比模型性能。不同模型在建筑外墙裂缝数据集上实验结果如表 5 所示。

表 5 模型对比结果

Table 5 Model comparison results

模型	IoU/%	Recall/%	Precision/%	$F_{1_score}/\%$
PSPNet	57.10	76.36	65.70	70.63
DeepLabv3+	64.04	79.56	78.44	79.00
U-Net	65.53	77.68	82.91	80.21
本文模型	69.06	82.57	86.30	84.39

由表 5 可知,改进的 U-Net 模型各项评价指标分别为 69.06%、82.57%、86.3%、84.39%,均高于其他语义分割模型,说明模型对建筑外墙裂缝分割具有最优的性能。

为了更直观地对比经典模型与本文模型对建筑外墙裂缝图像的分割效果,选取 3 张分别包含细长裂缝、任意分布裂缝及复杂背景噪声的裂缝图像进行预测,得到各模型的分割结果如图 11 所示。

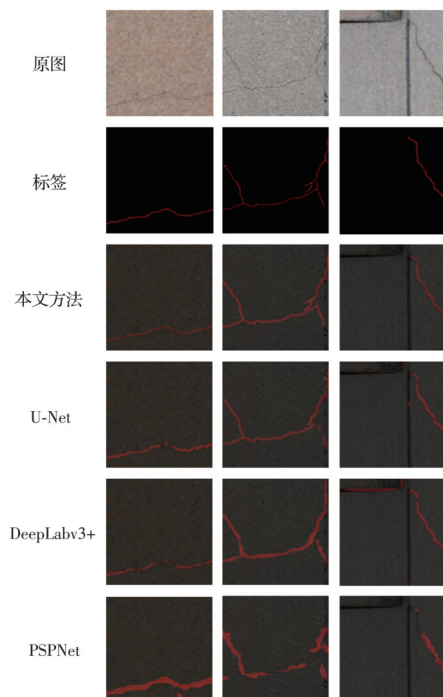


图 11 裂缝分割结果对比

Fig. 11 Crack segmentation results comparison

由图11可知,经典语义分割模型在裂缝数据集上训练后均能进行裂缝分割,但由于原模型并非针对裂缝特征进行设计,故裂缝分割效果较差,存在大量漏检和误检的情况。对比经典模型分割效果可知,U-Net模型对裂缝的分割效果较理想,存在少量误检和漏检的情况;Deeplabv3+模型对裂缝的分割效果较U-Net要差,漏检和误检情况更加明显;PSPNet模型对裂缝的分割效果较差,不仅分割粗糙,且存在大量漏检和误检。与经典模型相比,针对裂缝特征改进的U-Net模型对裂缝分割不仅准确,而且分割精细,与标签图像基本一致;与U-Net模型相比,改进模型解决了原模型细长裂缝分割不连续、复杂背景下裂缝漏检及背景误检的问题。

4 结论

针对建筑外墙裂缝人工检测方法效率低、检测效果和安全性差的问题,提出基于无人机航拍与计算机视觉的裂缝检测方法。针对U-Net模型存在的问题进行改进,并进行对比实验,得出以下结论:

1)无人机灵活性强,可以绕建筑飞行并快速获取高清的外墙图像。在构建裂缝数据集和建筑外墙裂缝定期检测阶段,可以使用无人机绕建筑航拍实现安全、高效地采集建筑外墙图像。

2)提出了组合Focal Loss与Dice Loss的FD Loss,用于模型优化。新损失函数具有更强的处理正负样本分布不均的能力,保证了模型训练过程的稳定性,并提高了模型对裂缝的分割性能。

3)所提改进的U-Net模型增强了模型编码网络的特征提取能力,以及模型获取多尺度上下文信息的能力,解决了原模型细长裂缝分割不连续、复杂背景下裂缝漏检及背景误检的问题,IoU指标和 F_1 _score分别提升了3.53%、4.18%。实验表明,改进的U-Net模型可以对复杂背景下的裂缝进行准确识别,具有较强的鲁棒性。

4)所提基于无人机航拍与计算机视觉的建筑外墙裂缝检测方法,使用无人机获取裂缝图像,并利用改进的U-Net模型对裂缝进行检测。该方法解决了传统建筑外墙裂缝检测方法需要人工高空作业带来的安全问题,并提高了裂缝检测的效率和精度。

参考文献

- [1] 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),2019,47(9):1277-1285.
ZHOU Y, LIU T. Computer vision-based crack detection and measurement on concrete structure [J]. Journal of Tongji University (Natural Science), 2019, 47(9): 1277-1285. (in Chinese)
- [2] MUNAWAR H S, HAMMAD A W A, HADDAD A, et al. Image-based crack detection methods: A review [J]. Infrastructures, 2021, 6(8): 115.
- [3] DAIS D, BAL I E, SMYROU E, et al. Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning [J]. Automation in Construction, 2021, 125: 103606.
- [4] 宋立博,费燕琼.两类YOLOv4-tiny简化网络及其裂缝检测性能比较[J].同济大学学报(自然科学版),2022,50(1):129-137.
SONG L B, FEI Y Q. Comparison of two types YOLOv4-tiny simplified networks and their crack detection performance [J]. Journal of Tongji University (Natural Science), 2022, 50(1): 129-137. (in Chinese)
- [5] JU H Y, MA T, LI W, et al. Pixelwise asphalt concrete pavement crack detection via deep learning-based semantic segmentation method [J]. Structural Control and Health Monitoring, 2022, 29(8): e2974.
- [6] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany. 2015: 234-241.
- [7] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany. 2018: 833-851.
- [8] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. IEEE, 2017: 6230-6239.
- [9] 李良福,王楠,武彪,等.基于改进PSPNet的桥梁裂缝图像分割算法[J].激光与光电子学进展,2021,58(22):101-109.
LI L F, WANG N, WU B, et al. Segmentation algorithm of bridge crack image based on modified PSPNet [J]. Laser & Optoelectronics Progress, 2021, 58(22): 101-109. (in Chinese)
- [10] LI G, LI X, ZHOU J, et al. Pixel-level bridge crack detection using a deep fusion about recurrent residual convolution and context encoder network [J]. Measurement, 2021, 176: 109171.
- [11] 曹锦纲,杨国田,杨锡运.基于注意力机制的深度学习路面裂缝检测[J].计算机辅助设计与图形学学报,2020,32(8):1324-1333.
CAO J G, YANG G T, YANG X Y. Pavement crack

- detection with deep learning based on attention mechanism [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(8): 1324-1333. (in Chinese)
- [12] NGUYEN N H T, PERRY S, BONE D, et al. Two-stage convolutional neural network for road crack detection and segmentation [J]. *Expert Systems with Applications*, 2021, 186: 115718.
- [13] 孟庆成,李明健,万达,等.基于 M-Unet 的混凝土裂缝实时分割算法[J]. *土木与环境工程学报(中英文)*, 2024, 46(1): 215-222.
- MENG Q C, LI M J, WAN D, et al. Real-time concrete crack segmentation algorithm based on M-Unet [J]. *Journal of Civil and Environmental Engineering*, 2024, 46(1): 215-222. (in Chinese)
- [14] DANG L M, WANG H, LI Y, et al. Automatic tunnel lining crack evaluation and measurement using deep learning [J]. *Tunnelling and Underground Space Technology*, 2022, 124: 104472.
- [15] LI Y, ZHANG J H, GUAN Z G, et al. Experimental study on the correlation between crack width and crack depth of RC beams [J]. *Materials*, 2021, 14(20): 5950.
- [16] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [17] WANG P Q, CHEN P F, YUAN Y, et al. Understanding convolution for semantic segmentation [C]//2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe, NV, USA. IEEE, 2018: 1451-1460.
- [18] MILLETARI F, NAVAB N, AHMADI S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation [C]//2016 Fourth International Conference on 3D Vision (3DV). Stanford, CA, USA. IEEE, 2016: 565-571.
- [19] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//2017 IEEE International Conference on Computer Vision. Venice, Italy. IEEE, 2017: 2999-3007.

(编辑 胡玲)