

文章编号:1671-4229(2023)02-0043-09

基于改进 DeepLabV3 + 在复杂环境下的道路裂缝检测

陈宇平, 范高*

(广州大学 土木工程学院, 广东 广州 510006)

摘要:有效的道路裂缝检测是保障道路安全的关键。针对现有道路裂缝检测方法效率低,检测结果易受检测环境影响的问题,文章结合深度学习与计算机视觉技术,在 DeepLabV3 + 架构的基础上提出了一种适用于复杂道路场景下的道路裂缝检测网络 Crack-Deeplab。Crack-Deeplab 引入新颖的网络模块和结构设计,具有轻量化、强泛化的特点以及精细分割的能力。文章基于数据集 Crack500 进行试验,验证和测试数据集的裂缝交并比 (Intersection over Union, IoU) 分别达到了 0.67 和 0.58,比现有的网络有明显提高;另外,采用复杂环境下拍摄的广州大学校内道路图片对该网络进行实际工程验证,基于 Crack500 数据训练的 Crack-Deeplab 在无需新增附加训练数据的情况下,能精准识别和分割出不同场景及环境下校内道路的裂缝,证明了该方法的有效性和鲁棒性,以及在实际工程中的应用价值。

关键词: 裂缝检测; 深度学习; 语义分割; 计算机视觉

中图分类号: TP 391.41

文献标志码: A

Road crack detection based on improved DeepLabV3 + in complex environments

CHEN Yu-ping, FAN Gao*

(School of Civil Engineering, Guangzhou University, Guangzhou 510006, China)

Abstract: Effective road crack detection is the key to ensuring road safety. This paper combines deep learning and computer vision and proposes a road crack detection network Crack-Deeplab for complex road scenarios based on DeepLabV3 + architecture because of the low efficiency of existing road crack detection methods and the susceptibility of detection results to the detection environment. Crack-Deeplab introduces novel network modules and structural design with lightweight and strong generalization features as well as fine segmentation capability. In this paper, based on the dataset Crack500 for experimental validation, Crack-Deeplab achieves the crack Intersection over Union (IoU) ratios of 0.67 and 0.58 in the validation and test datasets, respectively, which is a significant improvement over the existing networks. The network was validated using images of roads on the campus of Guangzhou University taken under complex environments. Crack-Deeplab, trained on Crack500 data, accurately identified and segmented cracks on the campus roads in different scenes and environments without adding additional training data, proving the effectiveness and robustness of the method and its application value in practical engineering.

Key words: crack detection; deep learning; semantic segmentation; computer vision

收稿日期: 2022-12-05; 修回日期: 2023-01-05

基金项目: 广州市科技计划资助项目(202201010335)

作者简介: 陈宇平(1997—),男,硕士研究生. E-mail:2112016009@e.gzhu.edu.cn

*通信作者. E-mail:gao.fan@gzhu.edu.cn

引文格式: 陈宇平, 范高. 基于改进 DeepLabV3 + 在复杂环境下的道路裂缝检测[J]. 广州大学学报(自然科学版), 2023, 22(2): 43-51.

路面裂缝是公路交通的常见病害形式。由于不均匀的路基沉降、高温、暴雨等环境因素,以及车辆超载、冲击等人为因素,道路在长期使用中出现损坏并在表面产生裂缝。随着裂缝的发展,将导致路面径流渗入到内部结构,破坏道路内部结构的稳定,使路基失稳,严重时甚至可能引发路基的塌陷,从而造成人员伤亡和公共财产损失。为了保证道路的安全性及可靠性,运输维修部门必须定期检查路面的运营状态。早期的道路裂缝检测主要基于现场工程师的目测法,该方法不仅劳动强度大、效率低,而且由于检测人员经验参差,很可能出现漏检问题^[1]。

随着摄影设备和计算机性能的飞速发展,基于计算机视觉(Computer Vision, CV)的道路表面病害检测技术成为土木工程结构^[2]及交通^[3]领域重要的研究方向。相较于目测法,自动检测法不仅可以大幅降低人工成本,同时可以提高检测速度,降低漏检率,具有广泛的应用前景。目前,基于 CV 的裂缝识别主要由三类算法实现:第一类阈值法,假设裂缝像素比非裂缝像素更暗,通过设置适当的像素阈值分辨裂缝区域^[4]。采用阈值法检测裂缝速度快,但是由于光线变化及污点等噪声会使阈值失效,大大降低了识别准确率,因此,难以应用在复杂的环境中,一般在固定的数据集或固定的场景中可以得到良好的结果。第二类边缘检测算法,包括边缘检测器^[5]、数学形态学^[6]等。边缘检测算法比基于强度阈值的方法更具鲁棒性,但在实际环境中,由于路面的复杂性、光源的不均匀性和裂缝纹理的不规则性,识别结果稳定性差,只能作为人工检测的辅助手段。以上两类传统的计算机视觉检测方法自动化程度低,需要具有专业技术的工程师辅助调整参数,且对拍摄环境要求严格,无法作为通用的裂缝检测技术。

近年来,伴随着计算机性能的不断提高,基于包括支持向量机^[7]、深度学习^[8-10]等在内的第三类 CV 裂缝识别方法,即机器学习,在裂缝检测领域取得了高速发展。作为深度学习^[11]的核心,卷积神经网络(Convolutional Neural Networks, CNN)通过参数共享和稀疏层连接设计,相较于其他机器学习算法大幅降低了模型中的参数数量,且更善于捕捉二维图像中的网格状特征,在解决 CV 任

务上具有独特的优势^[12-14]。CNN 一项重要的应用是语义分割,通过带有标签数据的有监督训练,建立不同语义像素与对应标签的映射,完成训练后,可对新图像每个像素赋予对应的语义标签,完成图像成分的分割。该原理非常适用于高精度裂缝检测,即通过分配给道路图像每个像素以背景或裂缝的标签,可以准确识别图片中的裂缝成分。Cha 等^[8]改进 CNN 模型进行混凝土表面裂缝检测,验证了基于深度学习的裂缝检测方法的可行性。Yang 等^[3]制作 Crack500 道路裂缝数据集,并且设计了一种特征金字塔和分层推进模块,用于提取道路图片丰富的语义信息,从而有效地分割裂缝。张海川等^[15]采用生成器是 UNet++ 和判别器为 PatchGAN 的条件对抗网络进行裂缝检测。Lau 等^[9]基于 UNet 卷积神经网络在 ImageNet 数据集上预训练后,再进行道路裂缝检测。宋榕榕等^[16]在 Deepcrack 网络基础上引用三重注意力模块,以获得更多互补性的裂缝特征,实现道路裂缝的检测。

尽管现有研究已证明 CNN 模型可以有效地识别和分割裂缝,但是目前研究中所使用的训练集和测试集的场景单一,网络的泛化性和鲁棒性未能有效验证。同时,裂缝具有细长的特征,在语义分割的时候非常考验网络精细化分割能力,然而现有方法识别出的裂缝边缘特征还不够清晰,有时甚至会忽略细小裂缝。因此,本文针对目前裂缝检测方法泛化能力弱,检测效果易受拍摄环境干扰,细小裂缝识别精细度不足等问题,基于深度学习与 CV 技术,提出了一种高效、轻量的裂缝检测模型 Crack-Deeplab。通过增加主干网络提取特征的维度以及在编码器部分引入多个并行的空洞空间卷积池化金字塔(Atrous Spatial Pyramid Pooling, ASPP)模块,获得更加丰富和可靠的裂缝图像特征,以实现复杂场景下高精度的裂缝检测。本文的主要内容归纳如下:第一节介绍 Crack-Deeplab 网络结构和模型参数;第二节根据 Crack500^[3]道路裂缝数据集训练、测试和评估 Crack-Deeplab,并与 DeepLabV3+ 比较;第三节在基于 Crack500 训练出的 Crack-Deeplab 网络中直接预测不同状况下现实道路图片中的裂缝;最后一节总结全文并提出未来改进的建议。

1 多尺度特征融合的 Crack-Deeplab

本文在 DeepLabV3 + 模型的基础上,通过主干网络多尺度的特征提取和添加 ASPP 模块,设计出性能优异的裂缝检测网络 Crack-Deeplab。其特殊的结构设计可以在处理图像时捕捉更加丰富和可靠的图像特征,提高裂缝检测的可靠性、鲁棒性以及裂缝分割的精细度。裂缝检测是一个二分类的语义分割任务,即将图片像素划分为裂缝和背景。在一张道路图像中,通常表现裂缝的像素数量非常少,占比小于 10%,而背景成分占比超过

90%。这样不均衡的样本分布使网络在训练的过程中更倾向于学习主要成分的数据特征,而忽略次要成分的特征学习,网络在预测时也有更高的概率将像素点标记为主要语义以取得较小的误差,从而造成大量裂缝像素的错误识别。为解决该问题, Crack-Deeplab 首次在裂缝检测任务中采用焦点损失 (focal loss)^[17] 和骰子系数损失 (dice coefficient loss)^[18] 作为损失函数,有效缓解了样本不均衡导致的网络误识别。

1.1 Crack-Deeplab 框架

Crack-Deeplab 网络包括编码器和解码器两个部分,如图 1 所示。

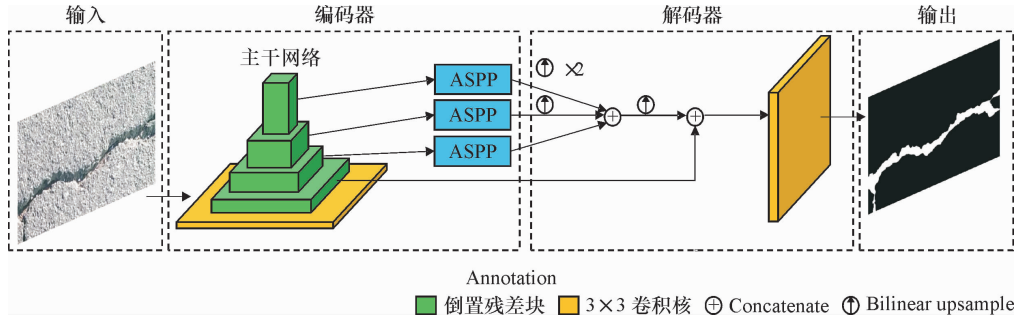


图 1 Crack-Deeplab 框架

Fig. 1 Architecture of Crack-Deeplab

图 1 中,编码器主要由主干网络和 ASPP 模块构成,编码器通过逐层的下采样连续压缩数据特征,促使网络挖掘数据代表性的高维特征。特征学习完成后,将主干网络不同深度提取出的特征图通过跳跃连接传递到 ASPP,然后再传递到解码器,通过双线性插值上采样改变特征图尺寸,再与主干网络提取的低维特征融合。通过对融合特征图进一步卷积和双线性插值上采样运算,将特征图尺寸升至原始输入图像大小,并通过 Softmax 层的概率运算预测各像素的输出语义。Crack-Deeplab 的特征提取通过堆叠的卷积层实现,既继承了卷积神经网络参数少的优点,又因其特殊的结构设计,在特征提取的效率上有很大提升。每个卷积层包含卷积、批标准化 (batch normalization) 和激活三步运算。通过在激活和卷积中间插入批标准化,可以防止梯度消失或梯度爆炸,并加快训练速度。Crack-Deeplab 中使用的激活函数包括 $ReLU$, $ReLU6$ 和 $Softmax$, 表达式如下:

$$ReLU(x) = \max(0, x), \quad (1)$$

$$ReLU6(x) = \min(\max(0, x), 6), \quad (2)$$

$$Softmax(p) = \frac{e^{y_j}}{\sum_i^N e^{y_i}}, j = 1, \dots, N, \quad (3)$$

公式(3)中, N 为分类数, y_j 为第 j 类的预测结果。

1.2 编码器的主干网络

本文采用 MobileNetV2^[19] 作为编码器的主干网络。MobileNetV2 通过大量使用包含逐点卷积 (pointwise convolution) 和逐通道卷积 (depthwise convolution) 的深度可分离卷积^[20], 具有预测精度高和轻量化的特点,在满足裂缝检测精度要求的同时还兼具优异的运算效率。MobileNetV2 由多个倒置残差块组成,如图 2 所示,图片特征图输入倒置残差块,经过逐点卷积进行升维,再通过逐通道卷积提取特征,最后再由逐点卷积进行特征压缩。与普通残差块相比,倒置残差块使用深度可分离卷积操作,可以大幅减少模型参数量,在不影响模型特征提取能力的前提下,使模型结构轻量化。此外,倒置残差块的最后一层不再进行激活计算,以避免对特征的损坏。图 2 中,倒置残差块有两种跨步不同的线性瓶颈结构,当跨步为 1 时,进行残差连接;当跨步为 2 时,则不进行残差的连接。

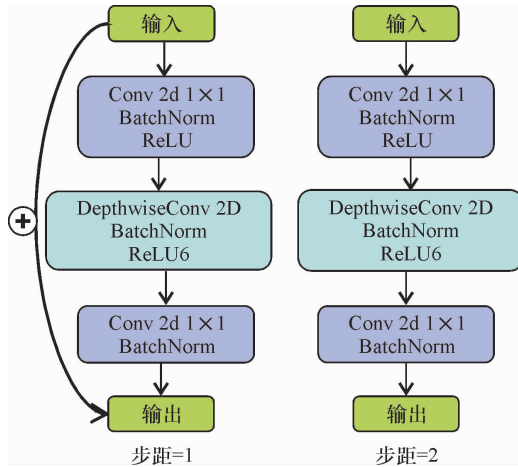


图 2 倒置残差块示意图

Fig. 2 The schematic illustration of the inverted residual block

1.3 编码器的 ASPP 模块

基于多个不同速率下并行的 ASPP 进行特征融合,可以在多个特征尺度上捕获图像的上下文信息,来更有效地恢复分割对象的整体和细节部分。ASPP 示意图如图 3 所示。

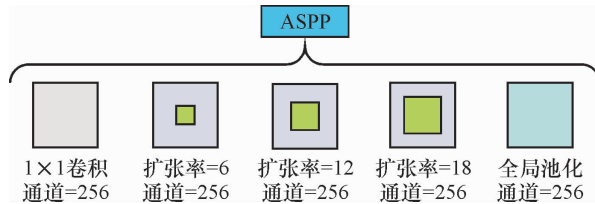
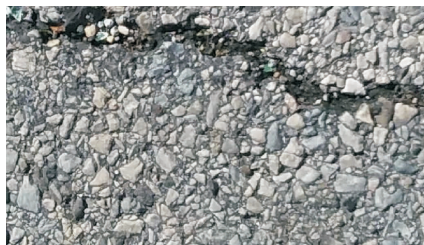


图 3 空洞空间卷积池化金字塔示意图 (ASPP)

Fig. 3 The schematic illustration of Atrous Spatial Pyramid Pooling (ASPP)

ASPP 包括一个 1×1 的卷积块,扩张率分别为 6、12 和 18 的 3 个空洞卷积和一个全局池化层。空洞卷积通过扩张率的大小增大感受野和长距离信息的联系,可以在扩大感受野的同时,保持空间分辨率和像素的相对空间位置。当使用多个不同扩张率空洞卷积提取的特征相融合时,能够有效加强数据信息的连续性,减少信息在特征传递时的丢失。



(a) 裂缝原始图像



(b) 裂缝标签

图 4 Crack500 数据集示意图

Fig. 4 Illustration of Crack500 dataset

1.4 损失函数

本文采用焦点损失和骰子系数损失作为网络的损失函数,来解决样本不均衡造成的裂缝检测失效。骰子系数损失是一个区域相关的损失值,即当前像素的损失值不仅和自身的预测值相关,也和附近其他像素的预测值相关,对正负样本不平衡的场景有较强的处理能力。焦点损失是另一种处理样本分布不均衡的损失函数,其根据判别样本的难易程度给对应的损失添加权重,使得模型在训练时更专注于难分类的样本特征学习。当执行二分类任务时,焦点损失函数和骰子系数损失函数表达式 L_F 和 L_D 分别为

$$L_F = -(1 - p_i)^\gamma \log(p_i), \quad (4)$$

$$L_D = 1 - \frac{2 \sum_i^N p_i g_i + \varepsilon}{\sum_i^N p_i^2 + \sum_i^N g_i^2 + \varepsilon}, \quad (5)$$

式中, L_F 表示焦点损失, L_D 表示骰子系数损失, N 代表样本的数量, p_i 代表预测的二进制分割体积, g_i 代表真正二进制分割体积, ε 是防止分母为零的最小值, p_i 表示预测为正样本的概率, $(1 - p_i)^\gamma$ 为调节系数。设置 $\varepsilon = 1e - 5$, $\gamma = 2$ 。网络整体的损失函数 L 表示为

$$L = L_D + L_F. \quad (6)$$

2 基于 Crack500 试验验证

Crack500 裂缝数据集示意图如图 4 所示。该图由 Yang 等^[3]在天普大学用手机收集的 500 张 2000×1500 像素的路面裂缝图像组成,原始图像为拍摄的 RGB 图片,标签为由背景和裂缝两种像素组成的语义标签图像。将每幅图像裁剪成 16 个不重叠的图像区域,只保留含有 1000 个以上裂缝像素的区域,最终形成包含训练集 1896 张图像,验证集 348 张图像,以及测试集 1124 张图像。

2.1 训练配置

在训练过程中,样本载入的批量大小(batch size)设置为16,训练共执行300个回合。将学习率设置在 $1e-3$ 和 $1e-6$ 之间,采用Adam优化器^[21]和余弦退火学习率^[22]调整学习率。Crack-Deeplab框架和MobileNetV2主干网络的代码由Python 3.8编写,并使用Tensorflow-GPU 2.6.0和Keras 2.6.0构建。用于训练拟议网络的工作站配置了NVIDIA Tesla A40、Intel Xeon Silver 4210R @ 2.40 GHz(X2)的处理器和128 GB RAM。

本文基于精度(Precision, Prec)和交并比(Intersection over Union, IoU)指标评估网络裂缝识别性能。这些指标可以表示为

$$Prec = \frac{TP}{TP + FP}, \quad (7)$$

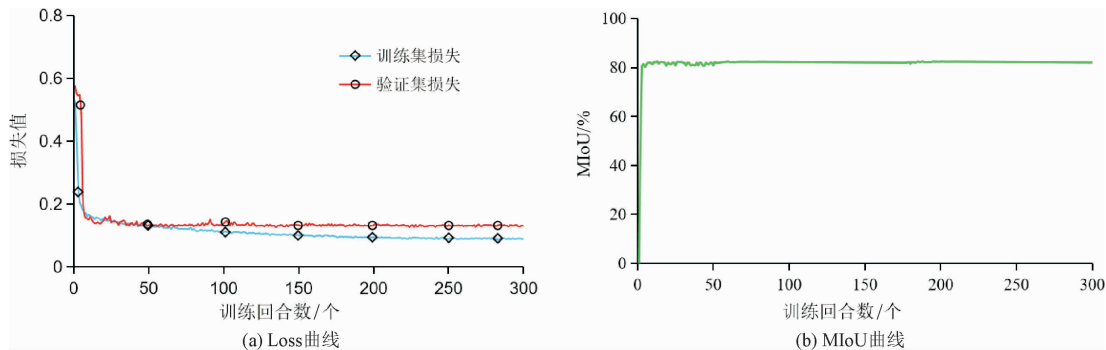


图5 Crack-Deeplab的Loss和MIoU曲线

Fig.5 Loss and MIoU curves of Crack-Deeplab

图5(a)展示了训练和验证数据集的损失曲线。训练和验证的损失在前30个回合迅速降低,达到0.14,而后,损失值缓慢下降,并在第220个回合趋于稳定,损失值最小的达到0.09。训练集和验证集的损失曲线在前80个训练回合中几乎重叠,这表明在大概80个训练回合时,网络已经拟合。图5(b)展示了验证集的MIoU曲线(背景和裂缝的平均IoU)。随着训练的推进,验证集的MIoU值迅速增长,在第15个回合,MIoU已经达到81.37%,后续至第78回合缓慢提高,并达到最大值82.22%。MIoU曲线表明Crack-Deeplab从训练数据集中快速地学习到了可靠的裂缝特征,网络在验证集上同样可以有效地识别和分割道路图像中的裂缝,具有优秀的泛化能力。网络训练完成后,最好的裂缝IoU达到0.67,背景IoU为0.98。

本文在验证集中随机选取3张图片,比较标

$$IoU = \frac{TP}{TP + FP + FN}, \quad (8)$$

TP (True Positive)表示模型预测的正样本为正类别, FN (False Negative)表示模型预测的正样本为负类别, FP (False Positive)表示模型预测的负样本为正类别。

网络首先基于开放的Pascal VOC^[23]数据集预训练。Pascal VOC数据集为图像语义分割数据集,具有20个分类。使用预训练可以调整初始模型的随机化参数,加快正式训练时网络梯度下降的速度,并且更易于获得一个低误差的模型。

2.2 评估和预测

本文使用包括旋转、水平翻转、高斯滤波及图像色域变换在内的数据增强技术增加训练数据量,提高网络的泛化能力。Crack-Deeplab的Loss和MIoU曲线如图5所示。

签图像和网络所预测图像(图6)。如图6中第一行原始图片所示,3张图的背景颜色各不相同,图6(a)的背景为赤红,图6(b)的背景为泥土黄,图6(c)背景为大理石白。网络对原始图片中的裂缝进行识别,由预测图像可以看出裂缝分割准确精细,边缘清晰,且与标签中的裂缝图形非常接近。同时,Crack-Deeplab面对不同背景的路面图像,没有受到背景变化的干扰而导致网络失准,体现出网络强大的泛化能力。

Crack500的测试数据集相较于验证集的图片内容更为复杂,有更多的细裂缝和多裂缝场景。对测试集裂缝检测非常考验网络对裂缝边界上和细节上精细分割的能力。本文在测试集中随机选取3张图片,比较标签图像和网络所预测图像,验证Crack-Deeplab网络精细分割能力,识别结果见图7。

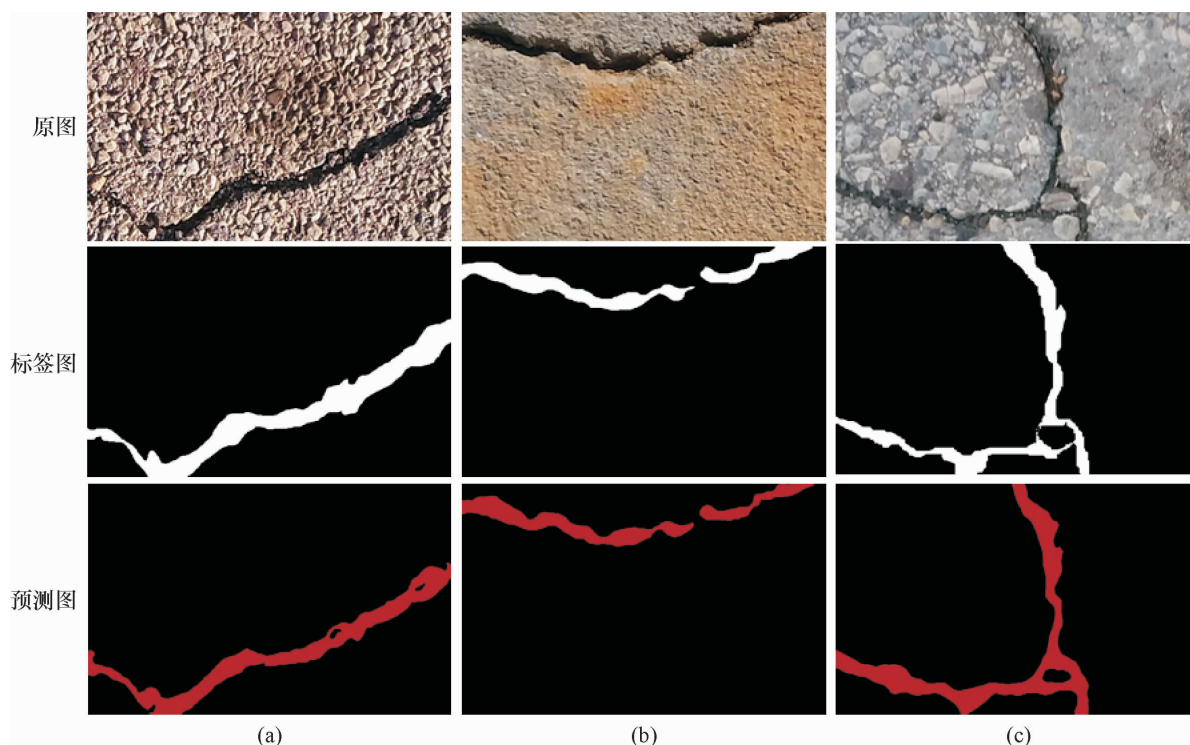


图 6 Crack500 验证集裂缝识别结果

Fig. 6 The crack detection results of samples from Crack500 validation dataset

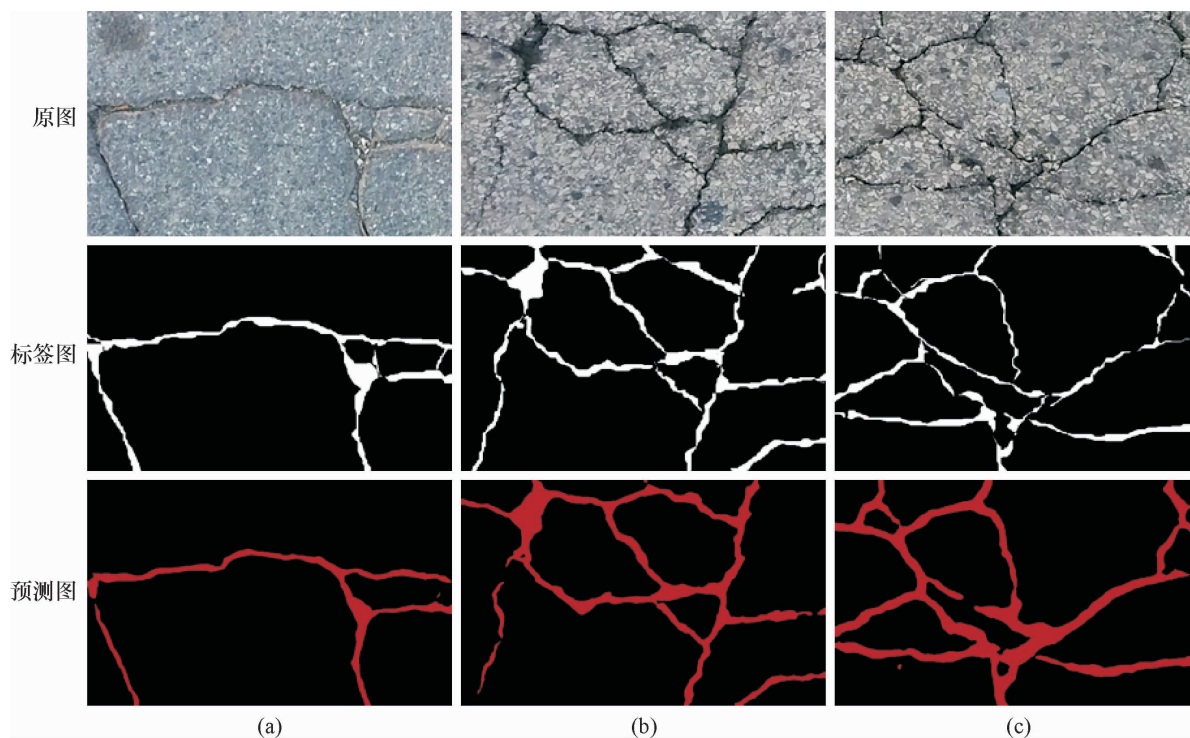


图 7 Crack500 验证集测试集识别结果

Fig. 7 The crack detection results of samples from Crack500 test dataset

在更为复杂的测试集上,裂缝的 IoU 值仍然可以达到 0.58。如图 7 所示,本文提出的 Crack-

Deeplab 在面对裂缝更多、更细的复杂场景时依然有着优秀的表现。由图 7(a) 和图 7(b) 可以

看出,面对狭长的裂缝,网络做到精细的分割。而图 7(c) 由于裂缝太多且过于密集,网络虽然无法将细小裂缝像素完全准确地分割,但依然有很高的准确度。

为了证明 Crack-Deeplab 网络结构设计的优越性,将主干网络为 MobileNetV2 的 Crack-Deeplab 与主干网络分别为 MobileNetV2 和 Xception^[24] 的 DeepLabV3 + 基于 Crack500 数据集进行对比。值得注意的是,原始的 DeepLabV3 + 和 Xception 同样是非常优秀的网络。在经过相同条件的预训练和训练之后,各网络表现如表 1 所示。

表 1 各网络 Crack500 测试集裂缝识别结果

Table 1 Performance of three networks with Crack500 test set

方法	主干网络	Precision	IoU /%	参数量 /百万
DeepLabV3 +	MobileNetV2	69.22	57.85	2.75
DeepLabV3 +	Xception	71.14	58.18	41.25
Crack-Deeplab	MobileNetV2	72.83	58.35	6.13

从表 1 可以看出,本文提出的 Crack-Deeplab 识别裂缝的 IoU 结果最佳,达到 58.35%,甚至超过参数多了接近 4 000 万的 DeepLabV3 + (Xception 作为主干网络)。另外,Crack-Deeplab 虽然参数量相较于 MobileNetV2 作为主干网络的 DeepLabV3 + 多了 300 多万个参数,但其裂缝识别能力有明显提高。裂缝分割任务更需要在细节和边界上做到精准分割。因此,Crack-Deeplab 通过空间金字塔丰富了裂缝边界信息,可以更精准地识别和分割裂缝,使得在网络参数少量提升的前提下,有效提高识别精度。

3 工程实际图像验证

Crack-Deeplab 在 Crack500 数据集上验证了其泛化和精细分割的能力。尽管训练集、验证集和测试集图片不同,但是,其拍摄环境和图像背景较为一致。而实际的检测场景和裂缝背景复杂多样,一个通用的自动检测技术需要能应对各种现实场景。因此,本节将第二节训练的 Crack-Deeplab 在未增加任何附加训练数据的情况下,应用于实际道路检测,验证方法的适用性。用于检测的图像是广州大学附近的不同道路,用手机距离地

面约 130 cm 的高度进行拍摄。拍摄得到的裂缝图片用 labelme 制作一份标签用于对比和评估。采集到的图片以 512 × 512 像素进行切割后输入 Crack-Deeplab 网络进行检测,预测结果如图 8 所示。图 8(a) 是在夜晚拍摄的道路图像,图 8(b) 和图 8(c) 拍摄于白天,场景都是沥青道路;图 8(d) 的场景是人行跑道。不同图片的裂缝和背景像素都不相同,甚至图片裂缝中还夹杂有落叶或花絮。观察预测图可以看出,尽管落叶和花絮与裂缝像素相差甚远,但是网络依然准确地识别和分割出道路裂缝,说明网络不靠像素大小识别和分割,而是基于裂缝的图像特征。在这样的复杂场景下,阈值法将完全失效。另外,如图 8(f) 所示,红色圈所示区域内有一段树根,其颜色和纹理都与裂缝非常接近,肉眼观测难以区分。在这种严苛的测试条件下,Crack-Deeplab 仍然正确地将树根与裂缝区别,并标记为背景。而阈值法或边缘检测法在这种场景下很难准确地识别裂缝。试验结果表明,Crack-Deeplab 可以准确识别不同光线、颜色、背景下的裂缝,且不会错误识别划痕、坑洼、树根等近似图像,不易受背景纹理的影响。

将制作的标签评估图 8 中 6 个图的网络预测图,得到总体裂缝的 IoU 值为 0.72。其中,图 8(a) ~ 图 8(f) 的 IoU 值分别为 0.68, 0.68, 0.62, 0.72, 0.78 和 0.76。结果表明,Crack-Deeplab 具有优秀的特征提取和泛化能力,有广泛的工程应用前景。

4 结 论

本文基于深度学习和 CV 技术,提出一种参数量少、精度高的裂缝检测模型 Crack-Deeplab。归由于其优秀的结构设计,Crack-Deeplab 在 Crack500 的验证和测试数据集测试中的裂缝 IoU 值分别达到了 0.67 和 0.58,且只有 613 万的参数量,表明网络具有轻量化和精细分割能力。在实际工程中展示出了出众的特征提取和泛化能力。对广州大学校园内随机拍摄的路面图片进行裂缝识别,总体裂缝 IoU 值达到了 0.72,具有优异的工程应用前景。未来,基于 Crack-Deeplab 模型高效、准确的特点,拟开展基于视频检测路面裂缝的研究。

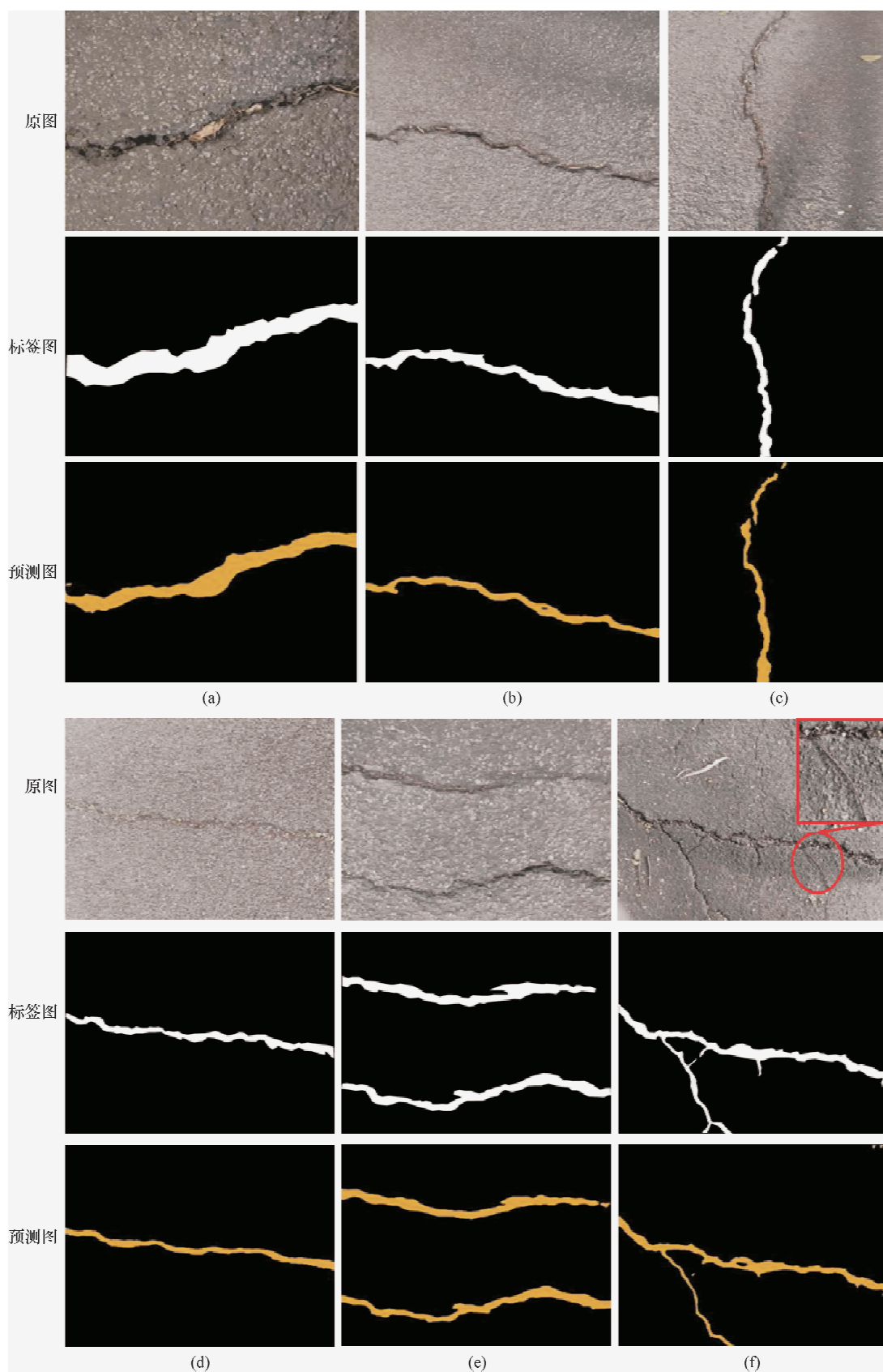


图 8 工程实际图像预测结果
Fig. 8 Prediction graphs for practical scenarios

参考文献:

- [1] Kaseko M S. A neural network-based methodology for automated distress classification of highway pavement images[D]. Irvine: University of California, Irvine, 1992.
- [2] Prasanna P, Dana K J, Gucunski N, et al. Automated crack detection on concrete bridges[J]. IEEE Transactions on Automation Science and Engineering, 2016, 13(2): 591-599.
- [3] Yang F, Zhang L, Yu S, et al. Feature pyramid and hierarchical boosting network for pavement crack detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(4): 1525-1535.
- [4] 邵闯,王生怀,邹春龙,等. 基于混合粒子群优化的2维 Otsu 路面裂缝图像阈值分割方法[J]. 湖北汽车工业学院学报, 2019, 33(3): 53-57.
- [5] 赵芳,周旺辉,陈岳涛,等. 改进的 Canny 算子在裂缝检测中的应用[J]. 电子测量技术, 2018, 41(20): 107-111.
- [6] Tanaka N, Uematsu K. A crack detection method in road surface images using morphology[C]//Proceedings of IAPR Workshop on Machine Vision Applications. San Diego: DBLP, 1998: 154-157.
- [7] Hoang N D, Nguyen Q L. A novel method for asphalt pavement crack classification based on image processing and machine learning[J]. Engineering with Computers, 2019, 35(2): 487-498.
- [8] Cha Y J, Choi W, Büyükköztürk O. Deep learning-based crack damage detection using convolutional neural networks[J]. Computer-aided Civil and Infrastructure Engineering, 2017, 32(5): 361-378.
- [9] Lau S L H, Chong E K P, Yang X, et al. Automated pavement crack segmentation using U-Net-based convolutional neural network[J]. IEEE Access, 2020, 8: 114892-114899.
- [10] Wang W, Su C. Convolutional neural network-based pavement crack segmentation using pyramid attention network[J]. IEEE Access, 2020, 8: 206548-206558.
- [11] Lecun Y, Bengio Y, Hinton G E. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [12] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2015: 234-241.
- [13] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017:6230-6239.
- [14] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//15th European Conference on Computer Vision. Berlin: Springer, 2018:833-851.
- [15] 张海川,彭博,许伟强. 基于 UNet ++ 及条件生成对抗网络的道路裂缝检测[J]. 计算机应用, 2020,40(2):158-161.
- [16] 宋榕榕,王财勇,田启川,等. 多层次特征融合和注意力机制的道路裂缝模型[J/OL]. 计算机工程与应用:1-11(2022-06-14). <https://kns.cnki.net/kcms/detail/11.2127.tp.20220611.1409.004.html>.
- [17] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [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. Piscataway: IEEE, 2016:565-571.
- [19] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018:4510-4520.
- [20] Howard A, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2022-11-20]. <https://arxiv.org/pdf/1704.04861.pdf>.
- [21] Kingma D, Ba J. Adam: A method for stochastic optimization[EB/OL]. (2014-12-22)[2022-11-20]. <https://arxiv.org/pdf/1412.6980v1.pdf>.
- [22] Loshchilov I, Hutter F. SGDR: Stochastic gradient descent with warm restarts[EB/OL]. (2016-08-13)[2022-11-20]. <https://arxiv.org/pdf/1608.03983v1.pdf>.
- [23] Everingham M, Eslami S M A, Gool L V, et al. The PASCAL visual object classes challenge: A retrospective[J]. International Journal of Computer Vision, 2015, 111(1): 98-136.
- [24] Chollet F. Xception: Deep learning with depthwise separable convolutions[EB/OL]. (2016-08-13)[2022-11-20]. <https://arxiv.org/pdf/1610.02357v1.pdf>.