

基于 DeepLabv3+ 的轻量化路面裂缝检测模型

夏晓华*, 苏建功, 王耀耀, 刘洋, 李明臻

长安大学工程机械学院, 陕西 西安 710000

摘要 裂缝是路面最主要的病害之一, 及时、有效地检测和评估裂缝对路面养护至关重要。为实现路面裂缝图像快速、准确的语义分割, 提出一种基于 DeepLabv3+ 模型的路面裂缝检测方法。为减小模型参数量、提高推理速度, 采用 MobileNetv3 作为模型的主干特征提取网络, 且在空洞空间金字塔池化模块中使用 Ghost 卷积代替普通卷积, 使模型更加轻量化。为避免替换主干网络降低模型精度: 首先, 在空洞空间金字塔池化模块中使用条形池化模块代替全局平均池化, 有效捕获裂缝结构的上下文信息, 避免无关区域噪声的干扰; 其次, 引入轻量级通道注意力机制 efficient channel attention (ECA) 模块, 增强特征的表达能力, 并设计浅层特征融合结构丰富图像的细节信息, 优化模型对裂缝的识别效果; 最后, 构造混合损失函数解决裂缝数据集类别不平衡而导致检测精度较低的问题, 利用迁移学习的训练方式提高模型的泛化能力。实验结果表明: 所提路面裂缝检测模型参数仅为 14.53 MB, 比原模型参数量减少 93.04%, 平均帧率达到 47.18, 满足实时检测的要求; 在精度方面, 该模型裂缝检测结果的交并比和 F1 值分别为 57.21% 和 72.76%, 优于经典的 DeepLabv3+、PSPNet、U-Net 模型和先进的 FPBHN、ACNet 等模型。所提方法可大幅减小模型参数量, 在保证路面裂缝检测精度的同时满足实时性, 为基于语义分割的路面裂缝在线检测奠定基础。

关键词 图像处理; 路面裂缝检测; 语义分割; DeepLabv3+; 轻量化; 检测精度

中图分类号 U418.6; TP391.4

文献标志码 A

DOI: 10.3788/L0P231323

Lightweight Pavement Crack Detection Model Based on DeepLabv3+

Xia Xiaohua*, Su Jianguo, Wang Yaoyao, Liu Yang, Li Mingzhen

College of Engineering Machinery, Chang'an University, Xi'an 710000, Shaanxi, China

Abstract Cracks are one of the main road surface diseases, and timely and effective crack detection and evaluation are crucial for road maintenance. To achieve fast and accurate semantic segmentation of road crack images, a road crack detection method based on the DeepLabv3+ model is proposed. To reduce the number of model parameters and improve inference speed, MobileNetv3 is used as the model's backbone feature extraction network, and Ghost convolution is used instead of ordinary convolution in the atrous spatial pyramid pooling module to make the model lightweight. To avoid degrading model accuracy by replacing the backbone network, the following measures are adopted. First, a strip pooling module is used in the atrous spatial pyramid pooling module to effectively capture the contextual information of crack structures while avoiding interference from irrelevant regional noise. Second, a lightweight channel attention mechanism, the effective channel attention (ECA) module, is introduced to enhance the feature expression ability, and a shallow feature fusion structure is designed to enrich the image's detailed information, optimizing the model's crack recognition effect. Finally, a mixed loss function is proposed to address the issue of low detection accuracy caused by imbalanced categories in the crack dataset, and transfer learning training is used to improve the model's generalization ability. The experimental results show that the proposed road crack detection model's parameters are only 14.53 MB, which is 93.04% less than the original model parameters, and the average frame rate reaches 47.18, meeting the requirements of real-time detection. In terms of accuracy, the intersection to union ratio and F1 value of this model's crack detection results are 57.21% and 72.76%, respectively, which are superior to classic DeepLabv3+, PSPNet, and U-Net models, as well as advanced FPBHN, ACNet, and other models. The proposed method can significantly reduce the number of model parameters while maintaining road crack detection accuracy and meeting real-time requirements, thus laying the foundation for online detection of road cracks based on semantic segmentation.

Key words image processing; road crack detection; semantic segmentation; DeepLabv3+; lightweight; accuracy of detection

收稿日期: 2023-05-15; 修回日期: 2023-07-04; 录用日期: 2023-07-24; 网络首发日期: 2023-08-15

基金项目: 国家自然科学基金项目(52205249)、陕西省重点研发计划项目(2019GY-116)

通信作者: *xhxia@chd.edu.cn

1 引言

裂缝作为常见的路面病害种类之一,对道路的性能、寿命及行车安全构成严重的潜在威胁。因此,及时、高效地检测路面裂缝并进行修复,对于保证交通运输质量、节约道路养护成本具有重要意义。早期的路面裂缝检测通过人工完成,费时费力、效率低,检测准确率依赖于专业检测人员的经验与知识。基于计算机视觉的路面裂缝自动检测方法可大幅度降低人力、时间成本,提高自动化水平。如何高效、准确、智能地检测路面裂缝已成为我国公路养护管理的热点问题^[1]。

早期的路面裂缝自动检测方法主要基于数字图像处理技术,如边缘检测法^[2]、阈值分割法^[3]、小波变换法^[4]及区域生长法^[5]等。这类方法利用机器完成裂缝检测,提高检测效率的同时避免产生人为误差,促进了路面裂缝检测技术的进步。但这些方法对阴影、光照及路面污渍等环境因素较敏感,鲁棒性不足,在实际应用中具有较大的局限性。

随着计算机视觉与深度学习的发展,卷积神经网络(CNN)在目标检测^[6-7]及图像分割^[8-9]领域中体现出巨大的优越性。CNN应用于裂缝检测时,主要分为两类,一类是以YOLO、SSD和Faster RCNN等为代表的目标检测方法,这类方法通过矩形检测框标记裂缝位置,并注明裂缝类别。罗晖等^[10]通过引入Focal损失函数,使用深度可分离卷积替代模型主干特征提取网络中的普通卷积,构建改进型YOLOv4轻量级神经网络模型,并借助迁移学习思想训练模型,在提高检测速度的同时检测准确率也有所提升。Hacıfendioğlu等^[11]利用两阶段(two stage)检测网络模型Faster RCNN识别混凝土路面裂缝,并研究不同拍摄方式、光照和天气条件对模型检测效果的影响。孙朝云等^[12]为提高SSD网络模型对路面裂缝分类和定位的准确率,在模型的特征提取结构中使用深度残差网络,且通过损失函数的收敛程度来优化模型中的超参数。基于目标检测的方法对形状简单、紧凑的裂缝有良好的效果,然而,对于复杂和狭小的裂缝,此类方法推理得到的矩形检测框面积过大,对裂缝的具体形状的检测不准确。此外,基于目标检测的方法只能确定裂缝的位置和相对大小,不能明确裂缝的形状、长度及宽度等详细信息,给道路精细化养护带来了困难。另一类方法是以fully convolutional network(FCN)、U-Net和DeepLabv3+等为代表的语义分割方法,这类方法对每个像素赋予合适的标签,实现图像端到端的像素级分类。彭磊等^[13]对U-Net模型的网络层数进行超参数优化,其在CRACK500测试集上的精确率为62%,召回率为59%,F1值为61%。王保宪等^[14]将U-Net++网络作为裂缝检测的主框架,通过融合更多的底层特征信息来提升检测精度。

袁嘉豪等^[15]对采用不同主干特征提取网络的DeepLabv3+模型进行对比实验,实验结果表明,该模型在裂缝检测中效果优异。李国燕等^[16]通过空间注意力机制、视觉atrous spatial pyramid pooling(ASPP)模块和多特征融合机制构造MFC-DeepLabv3+模型,该模型检测的裂缝分割图像区域更完整,边缘更清晰。然而,为提高路面裂缝检测的准确率,语义分割模型的网络深度不断增加、复杂程度不断提高,致使模型存在参数量巨大、计算复杂、推理速度慢及不适宜部署到嵌入式系统的缺点,影响检测效率,不能应用在实时场合中。因此,在保证路面裂缝检测精度的前提下,如何减少语义分割模型参数量使其更加轻量化并提高其推理速度是当前研究亟须解决的问题。

鉴于DeepLabv3+模型在基于语义分割的裂缝检测中具有良好的准确性,本文在其基础上开展路面裂缝检测方法研究,旨在构建具有良好检测精度和满足实时检测要求的轻量化路面裂缝检测模型。

2 研究方法

路面裂缝检测本质上是一种二分类的语义分割问题,语义分割方法对图像中的每个像素都进行分类,即为每个像素都分配一个适当的语义标签。本研究基于DeepLabv3+构建轻量化路面裂缝检测模型,包括以下步骤:

1)数据预处理。收集路面裂缝图像构建数据集,对数据集图像进行裁剪操作后划分训练集、验证集以及测试集。

2)模型改进与训练。从模型轻量化、精度保证和参数设置等方面改进DeepLabv3+模型,使用训练集和验证集对改进后的模型进行训练,获得路面裂缝检测模型。

3)模型验证。利用测试集样本检验改进后的模型,并对检测结果进行评价。

3 数据来源与预处理

本实验图像取自公开裂缝数据集CRACK500^[17],该数据集中的裂缝图像背景复杂,包含如阴影、光照变化和路面污渍等影响路面裂缝检测效果的干扰项,便于充分训练出泛化能力强的模型,是文献中常用的公开数据集。CRACK500数据集包括400张训练集、100张验证集以及200张测试集图像,分辨率均为2560 pixel×1440 pixel。为扩增图像数量,将每张图像无重叠地裁剪成16张大小相同的低分辨率图像,只保留包含超过1000个裂缝像素的图像。通过这种方式,获得1896张训练集、348张验证集和1124张测试集图像,分辨率均为640 pixel×360 pixel或360 pixel×640 pixel。最终将所有图像的分辨率均调整为512 pixel×512 pixel。

4 改进的 DeepLabv3+ 模型

4.1 传统 DeepLabv3+ 模型简介

DeepLabv3+ 是用于语义分割的典型模型之一,分为编码和解码两个部分。DeepLabv3+ 模型的特征图可视化结果如图 1 所示。

在编码部分,输入特征图经 Xception 网络进行特征提取后,共获得 5 个大小分别为 1/2、1/4、1/8、1/16、

1/16 的特征图,依次为图 1 中的 stage-1~stage-5,将 stage-5 通过 ASPP 模块进行特征提取并完成特征拼接后利用 1 个 1×1 卷积块压缩特征,拼接是指将多个大小相同的特征图在通道维度上进行连接,可增加通道数。在解码部分:首先采用 1×1 卷积块调整 stage-2 的通道数;然后将其与完成 4 倍上采样后的 1/16 大小特征图进行拼接;最后通过 3×3 卷积块及 4 倍上采样输出与原图分辨率相等的预测图像。

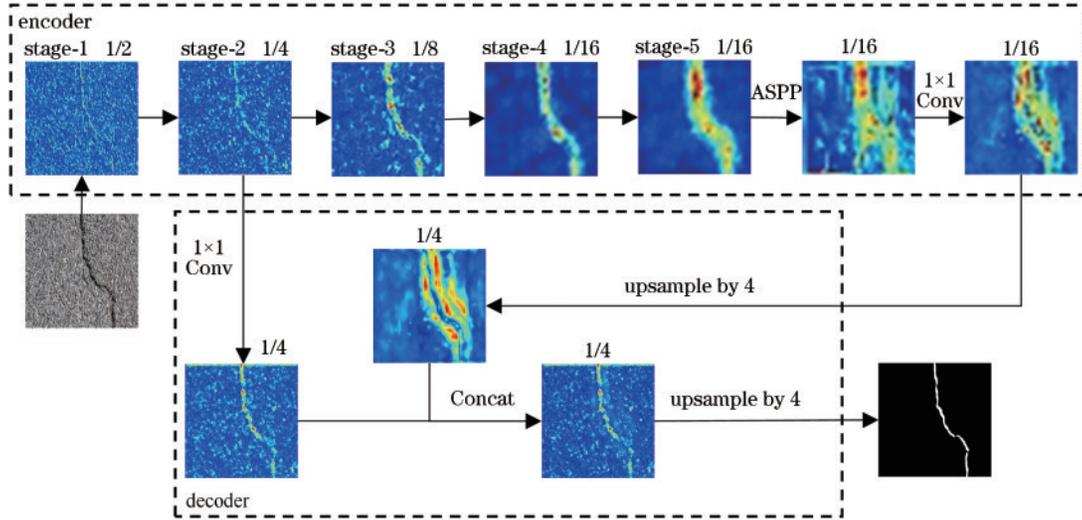


图 1 DeepLabv3+ 模型特征图可视化

Fig. 1 DeepLabv3+ model feature map visualization diagram

4.2 改进的 DeepLabv3+ 模型结构设计

DeepLabv3+ 模型采用参数量大、网络层数多的 Xception 作为主干特征提取网络,虽提高了分割精度,但模型的复杂度增加,对内存空间的要求更高。基于此,提出一种基于 DeepLabv3+ 的轻量级路面裂缝检测模型。首先,在传统 DeepLabv3+ 模型上将 MobileNetv3 作为主干特征提取网络,大幅度减小模型参数量,降低内存占用,提高模型推理速度。其次,将 ASPP 模块中的普通卷积替换为 Ghost 卷积,进一步轻量化模型。为保证轻量级 DeepLabv3+ 模型的分割精度,对其结构进行以下优化:1) 在 ASPP 中引入条形池化模块 (SPM); 2) 采用 ECA 注意力模块,并借助 ECA 设计浅层特征融合结构 (SFF); 3) 构造混合损失函数。改进后的 DeepLabv3+ 模型结构如图 2 所示。

4.2.1 主干特征提取网络 MobileNetv3

MobileNet 是一种专注于嵌入式设备或移动端的轻量级卷积神经网络,有更准确、更高效的特点。MobileNetv3^[18] 是根据 MobileNet 改进的第 3 版,其网络结构主要分为 3 部分。第 1 部分采用 1 个 3×3 的卷积块完成特征提取;第 2 部分将得到的特征输入一系列 block 结构中,其由 9 个 3×3 和 6 个 5×5 的卷积块组成;最后 1 部分通过 1 个 1×1 卷积块对输出特征进行升维,再通过 1 个全局池化块将特征转为一维向量

特征。

本实验使用 MobileNetv3 的前 2 个部分作为主干特征提取网络,具体网络结构如表 1 所示。其中:粗体突出表示改进该行 bneck 结构;“—”表示该层不进行对应操作;在“Operator”中,bneck 为 MobileNetv3 的 block 结构,“ 3×3 ”及“ 5×5 ”表示该层使用的卷积核大小;“Exp size”是每一个 block 结构中第 1 层 1×1 卷积块的卷积核个数;“out”表示输出通道数;“SE”表示 squeeze and excitation 通道注意力机制;“√”表示使用 SE 注意力模块;“NL”表示每层使用的非线性激活函数的种类,“HS”为 h-swish 激活函数,“RE”为 ReLU 激活函数。深层特征图具有语义信息丰富的优点,为获得更深的 1/16 大小的特征图,将 MobileNetv3 中的第 14 个 bneck 结构的步长 (s) 设置为 1,并使用膨胀卷积增加其感受野。

4.2.2 Ghost 卷积

图片通过卷积操作后,可获得诸多特征图,但存在特征图冗余的现象,即有些特征图相似性很高,Han 等^[19] 将相似的特征图认为是彼此的 Ghost,并设计了更轻量的 Ghost 卷积来代替普通卷积。Ghost 卷积将普通卷积分为 2 部分,操作原理如图 3 所示。首先通过 1 个 1×1 普通卷积块把输出特征图的通道数减半;然后在第 1 部分的基础上逐层使用深度可分离卷积生成冗余的特征图;最后在通道维度上对 2 组特征

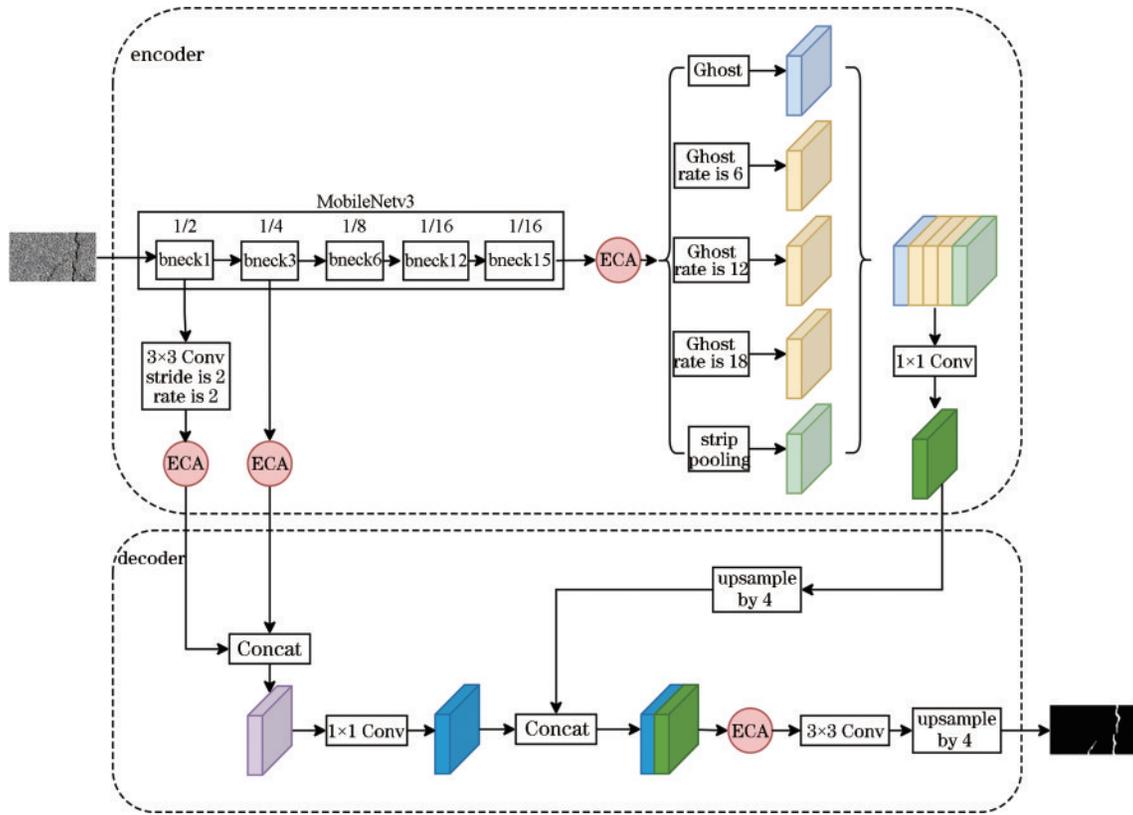


图 2 改进 DeepLabv3+ 模型结构图

Fig. 2 Improved DeepLabv3+ model structure diagram

表 1 MobileNetV3 结构

Table 1 MobileNetV3 structure

Input	Operator	Exp size	out	SE	NL	s
$224^2 \times 3$	Conv2d	—	16	—	HS	2
$112^2 \times 16$	bneck, 3×3	16	16	—	RE	1
$112^2 \times 16$	bneck, 3×3	64	24	—	RE	2
$56^2 \times 24$	bneck, 3×3	72	24	—	RE	1
$56^2 \times 24$	bneck, 5×5	72	40	✓	RE	2
$28^2 \times 40$	bneck, 5×5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 5×5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 3×3	240	80	—	HS	2
$14^2 \times 80$	bneck, 3×3	200	80	—	HS	1
$14^2 \times 80$	bneck, 3×3	184	80	—	HS	1
$14^2 \times 80$	bneck, 3×3	184	80	—	HS	1
$14^2 \times 80$	bneck, 3×3	480	112	✓	HS	1
$14^2 \times 112$	bneck, 3×3	672	112	✓	HS	1
$14^2 \times 112$	bneck, 5×5	672	160	✓	HS	2→1
$14^2 \times 160$	bneck, 5×5	960	160	✓	HS	1
$14^2 \times 160$	bneck, 5×5	960	160	✓	HS	1

图进行连接。使用 Ghost 卷积代替传统 DeepLabv3+ 模型编码部分 ASPP 模块中的普通卷积,减小模型参数量,使模型更轻量化。

4.2.3 条形池化模块

路面裂缝的结构呈带状分布,如图 4 所示,传统的

池化方法是在固定的正方形区域内收集信息,易丢失带状裂缝的远程上下文信息。针对此问题,在传统 DeepLabv3+ 模型编码部分的 ASPP 模块中利用条形池化模块^[20]进行池化操作,有效捕获裂缝结构的上下文信息,同时能避免无关区域的噪声干扰,进而更好地提升分割精度。

条形池化模块网络结构如图 5 所示。条形池化模块由 2 个途径组成,分别沿水平和垂直空间维度对远程上下文信息进行编码,可有效扩大网络的感受野范围。

在条形池化模块中,对输入尺寸为 $H \times W$ 的特征图,分别在水平和垂直空间维度上以 1 行或 1 列进行平均池化后得到 $1 \times W$ 和 $H \times 1$ 的特征图,如图 5(b)、(c) 所示,随后对 2 个输出特征图分别沿着水平和垂直方向进行扩容,如图 5(d)、(e) 所示,并将扩容后的特征图通过对应位置像素求和的方式进行融合,如图 5(f) 所示,最后通过 1×1 卷积块和 Sigmoid 函数获得权值,将权值与原输入特征图相乘得到输出特征图,如图 5(g) 所示。

4.2.4 ECA 注意力模块

注意力机制可让卷积神经网络模型在处理图像时自适应地注意重要、有效的目标信息,提升模型性能。ECA^[21]是一种新的捕捉局部跨通道交互信息的注意力模块,其高效且轻量。在 $1/16$ 大小特征图及其 4 倍上采样与 $1/4$ 大小特征图拼接后的位置加入 ECA,增

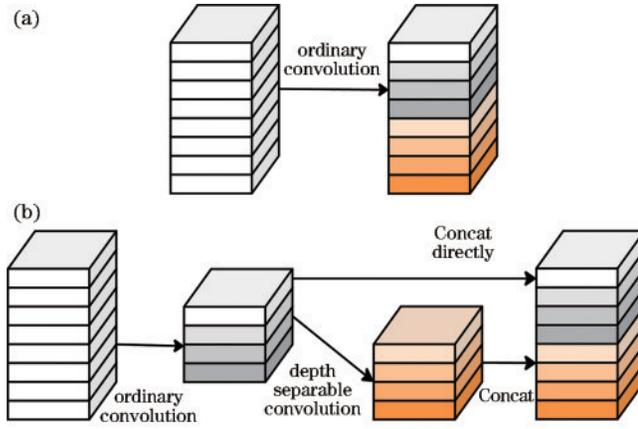


图3 普通卷积与Ghost卷积的对比。(a)普通卷积;(b) Ghost卷积

Fig. 3 Comparison of ordinary convolution and Ghost convolution. (a) Ordinary convolution; (b) Ghost convolution

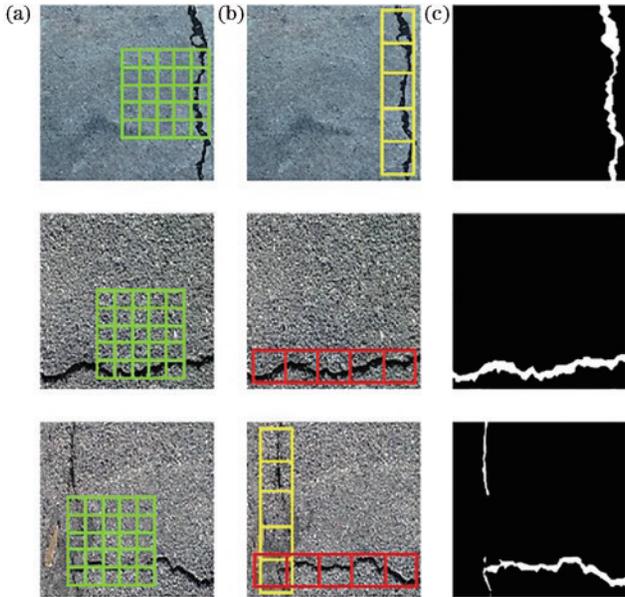


图4 普通池化、条形池化与对应标签。(a)普通池化;(b)条形池化;(c)对应标签

Fig. 4 Common pooling, strip pooling and corresponding label. (a) Common pooling; (b) strip pooling; (c) corresponding label

强模型特征提取能力。

传统DeepLabv3+模型在解码部分没有充分结合浅层特征图,提取的局部细节信息不够丰富,从而使整个模型的预测结果并不精确。为解决该问题,所提方法利用特征金字塔思想并借助ECA在应用1/4大小的特征图的基础上加入1/2大小的特征图,具体操作如图6所示。A和B分别为编码部分下采样时生成的特征图,大小分别为1/2和1/4,通道数分别为16和24。C为A通过1个卷积核大小为3、步长为2、空洞率为2的膨胀卷积块和ECA模块后得到的特征图,大小为1/4,使用膨胀卷积可扩大感受野,ECA模块可去除噪声,学习更有用的特征信息。D为通过ECA模块后

的特征图,大小不变。将C和D拼接后得到E,E使用1次1×1卷积块进行通道降维后得到F。ASPP模块输出的特征图通过1次1×1卷积块再进行4倍上采样后得到的特征图与F拼接,拼接后的特征图使用1次3×3卷积块后再进行4倍上采样得到的输出结果即为模型的预测结果。

相同条件下未引入和引入浅层特征融合结构时模型浅层特征图的对比结果如图7所示,可以看出,引入浅层特征融合结构后,裂缝特征的表达能力增强,局部信息更为丰富,而其余背景信息则得到了抑制。

4.2.5 混合损失函数

训练数据集中的类别不平衡导致深度学习模型在训练过程中过分关注像素数量多的样本,阻碍模型精度的提升。表2为训练集中各类别像素占总体像素的百分比。可以看到,背景类别的像素是目标类别裂缝像素的15倍之多,在训练过程中可能会将部分裂缝样本误判为背景,进而降低模型的精度。

为此,构造一种结合二分类交叉熵损失(BCE)和Dice损失的混合损失函数。BCE是二分类任务中常用的损失函数,其在衡量标签与预测结果之间的相似性方面表现优异,因此,将BCE作为改进DeepLabv3+模型的损失函数之一。但该损失函数对类别不平衡问题较为敏感,会导致像素占比较多的类别出现过拟合,其表达式为

$$L_{BCE} = -\frac{1}{N} [y_i \lg p_i + (1 - y_i) \lg (1 - p_i)], \quad (1)$$

式中: N 表示图像的总像素数; y_i 和 p_i 分别表示第 i 个像素点的标签值和预测概率值。

Dice损失适用于训练类别不平衡的分割模型,在类别不平衡的情况下可以达到更好的效果。但其对噪声敏感,可能会将边界信息忽略,致使边界分割不佳,其表达式为

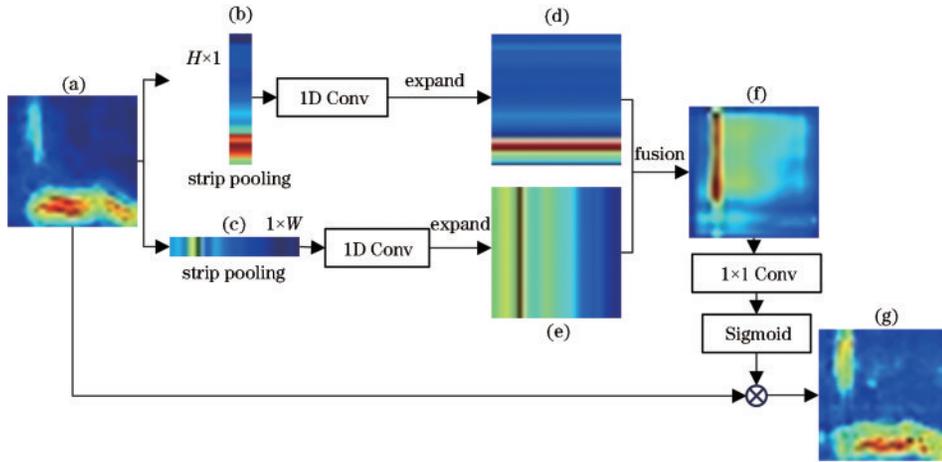


图 5 条形池化模块网络结构示意图。(a)输入特征图;(b)(c)池化特征图;(d)(e)扩容特征图;(f)融合特征图;(g)输出特征图
Fig. 5 Schematic diagram of the network structure of the strip pooling module. (a) Input feature map; (b) (c) pooling feature maps; (d)(e) expand feature maps; (f) fusion feature map; (g) output feature map

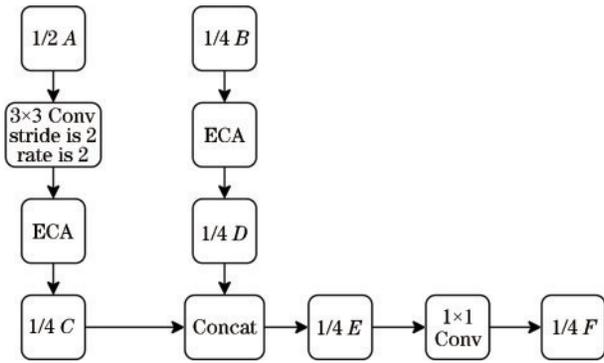


图 6 浅层特征融合结构图

Fig. 6 Diagram of shallow feature fusion structure

$$L_{Dice} = 1 - \frac{\sum_{i=1}^N p_i y_i + \epsilon}{\sum_{i=1}^N p_i + y_i + \epsilon} - \frac{\sum_{i=1}^N (1 - p_i)(1 - y_i) + \epsilon}{\sum_{i=1}^N (2 - p_i - y_i) + \epsilon}, \quad (2)$$

为加快收敛,减少过拟合, ϵ 值设置为 1。

因此,混合 BCE 和 Dice 损失对模型中的损失进行计算:

表 2 训练集中各类别像素占总体像素的比值

Table 2 The ratio of each pixel in the training set to the total pixel

Class	Background	Crack
Percentage / %	93.94	6.06

$$L_{Total} = L_{BCE} + L_{Dice} \quad (3)$$

混合损失函数可以解决裂缝图像中样本不均衡的问题,获得更精确的像素级检测效果,也能同时关注像素级别的分类准确率以及图像裂缝的分割效果,进而使模型更稳定训练。

5 模型的训练与测试

5.1 实验环境与模型训练

本实验在计算机上完成,利用 Python 及 PyTorch 框架来改进 DeepLabv3+ 模型。实验环境如表 3 所示。

为了提高模型的训练效率和性能,采用迁移学习的方法对模型进行训练。在模型的训练中,首先冻结主干特征提取网络层,以便将更多的计算资源用于训练后半部分的网络层,这个过程被称为冻结训练。训

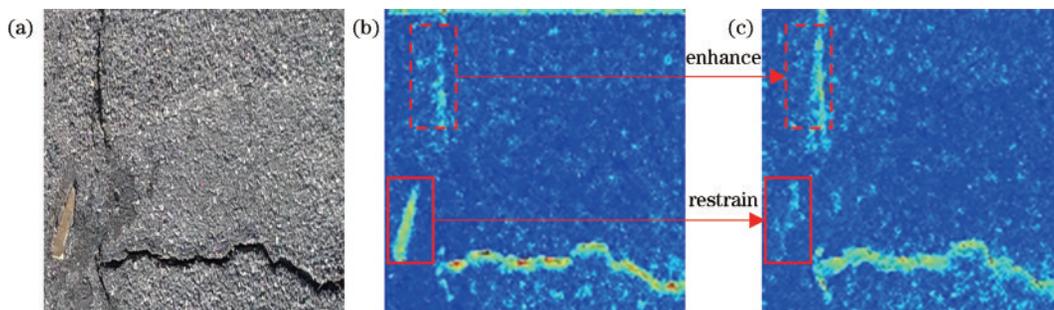


图 7 浅层特征图可视化。(a)原图像;(b)原特征图;(c)引入浅层特征融合结构后的特征图

Fig. 7 Shallow feature map visual diagram. (a) Original image; (b) original feature map; (c) feature map after introducing shallow feature fusion structure

表 3 实验环境

Table 3 Experimental environment

Configuration	Parameter
Operating environment	Ubuntu 16.04
Processor	7 Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40 GHz
Graphic processing unit	NVIDIA Gefore GTX Titan X
Accelerating environment	GPU
Development platform	Visual Studio Code

练一定代数后,解冻主干特征提取网络层,主干特征提取网络层的初始权重值采用 MobileNetv3 PASAL VOC2012 数据集上预训练好的模型参数,继续对整个网络进行训练,直至获得最终模型,即解冻训练。采用这种训练方式可以在有效保证权值的同时加快模型的收敛速度。

合理设定训练参数,是增强模型性能的重要手段。训练图像尺寸设为 $512 \text{ pixel} \times 512 \text{ pixel}$,在使用预训练参数时,设置模型的迭代次数(epoch)为 50 次,批处理的训练集样本数量(batch size)为 16。不使用预训练模型时,设置模型的迭代次数为 100 次,批处理的训练集样本数量为 8。基础学习率(base learning rate)为 0.0005,学习率采用余弦退火策略进行调节,基础学习率最小值设置为 0.000005。优化器选用 Adam 梯度优化器,其超参数 beta_1 设置为 0.9, beta_2 设置为 0.999。

5.2 评价指标

为客观评价训练所得模型对路面裂缝的检测效果,实验选取交并比(R_{IoU})、准确率($R_{precision}$)、召回率(R_{recall})、平均帧率(FPS)以及参数量(Params)作为评价指标。交并比、准确率和召回率评价指标计算公式分别为

$$R_{IoU} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}, \quad (4)$$

$$R_{precision} = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (5)$$

$$R_{recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (6)$$

式中: N_{TP} 表示分类正确的裂缝像素数; N_{FP} 表示将背景预测为裂缝的像素数; N_{FN} 表示将裂缝预测为背景的像素数。

交并比反映预测裂缝与实际裂缝的重合程度;准确率反映裂缝检测的可靠程度;召回率反映裂缝检测的完整程度;另外,平均帧率为模型 1 s 处理的图像数量,越高表示模型推理速度越快。参数量为评价模型体积的重要指标,越低表示模型的轻量化程度越高。

由于准确率和召回率是相互矛盾的,召回率高时,准确率偏低,为综合评价这两个指标,引入 F1 分数(S_{F1}),其值越大表示分割效果越佳,公式如下:

$$S_{F1} = \frac{2 \times R_{recall} \times R_{precision}}{R_{recall} + R_{precision}}. \quad (7)$$

5.3 评价指标

5.3.1 消融实验

为验证将传统 DeepLabv3+ 模型的主干特征提取网络 Xception 替换为轻量级的 MobileNetv3、ASPP 模块中的普通卷积改为 Ghost 卷积的必要性,以及采用条形池化模块、引入 ECA 注意力模块、设计浅层特征融合结构和构造混合损失函数操作的有效性,设置 6 组不同方案进行消融实验。实验结果如表 4 所示,粗体突出显示为最佳结果,方案 1 为传统 DeepLabv3+ 模型的实验结果。

表 4 消融实验结果

Table 4 Results of ablation experiment

Plan	MobileNetv3, Ghost	BCE+Dice	SPM	ECA	SFF	$R_{IoU} / \%$	$R_{precision} / \%$	$R_{recall} / \%$	F1 / %	FPS	Params / MB
1						56.47	74.47	70.03	72.18	21.48	208.70
2	✓					55.73	73.99	69.31	71.57	57.74	14.32
3	✓	✓				56.57	68.91	75.95	72.26	57.74	14.32
4	✓	✓	✓			56.84	69.95	75.20	72.48	52.20	14.46
5	✓	✓	✓	✓		57.00	70.50	74.87	72.62	50.22	14.46
6	✓	✓	✓	✓	✓	57.21	70.60	75.09	72.76	47.18	14.53

由表 4 消融实验结果可知,相较于传统 DeepLabv3+,方案 2 的参数量减小 194.38 MB,平均帧率提升 36.26,表明将传统 DeepLabv3+ 模型的 Xception 主干特征提取网络改为更轻量级的 MobileNetv3、将 ASPP 模块中的普通卷积改为 Ghost

卷积可明显减小模型参数量,显著提高模型推理速度。方案 6 相较于传统 DeepLabv3+ 和方案 2、3、4 及方案 5 在交并比、F1 值方面均有所提升,且比传统 DeepLabv3+ 的参数量减小 194.17 MB,减少 93.04%,平均帧率多 25.7,增加 1.20 倍,表明所提方

法可以大幅减小参数量,在提升路面裂缝检测精度的同时满足实时性。

5.3.2 不同模型性能对比分析

为进一步验证改进 DeepLabv3+ 模型的路面裂缝检测性能,将所提模型与 FPHBN^[17]、CNNs^[22] 和 ACNet^[23] 模型进行比较。FPHBN 是一种新颖的网络模型架构,其以特征金字塔的方式将上下文信息集成到低级特征中,用于路面裂缝检测,在不同数据集上均取得了较好的检测效果;CNNs 作为有效的裂缝检测模型,其稳定且可在不复杂的体系结构中快速运行;ACNet 是一种先进的基于注意力机制的裂缝检测模型,取得了理想的检测效果。

文献[23-24]中提供了 FPHBN、CNNs 和 ACNet 模型在 CRACK500 数据集上裂缝检测精度的实验结果,其训练及测试使用的数据集和预处理方式与本研究完全一致,对比结果如表 5 所示。由表 5 可知,所提

方法在交并比、召回率和 F1 值评价指标方面有很大的改进,说明其具有良好的检测精度。

表 5 在 CRACK500 数据集上使用不同模型的测试结果比较
Table 5 Comparison of test results using different models on the CRACK500 dataset unit: %

Model	R_{IoU}	$R_{precision}$	R_{recall}	F1
FPHBN	54.51	71.23	71.65	70.56
CNNs	52.61	69.54	67.44	68.95
ACNet	54.92	68.05	74.89	69.82
Proposed method	57.21	70.60	75.09	72.76

采用与本研究相同的数据集和训练方法对经典的 DeepLabv3+、M-PSPNet (主干网络为 MobileNetv2)、R-PSPNet (主干网络为 ResNet50) 和 U-Net 模型进行训练,并与所提方法进行比较,对比结果见表 6。

表 6 所提方法与不同语义分割模型的测试结果比较

Table 6 Comparison of test results between the proposed method and different semantic segmentation models

Model	backbone	$R_{IoU} / \%$	$R_{precision} / \%$	$R_{recall} / \%$	F1 / %	FPS	Params / MB
Traditional DeepLabv3+	Xception	56.47	74.47	70.03	72.18	21.48	208.70
M-PSPNet	MobileNetv2	53.86	73.48	66.86	70.01	86.82	9.06
R-PSPNet	ResNet50	55.37	71.74	70.81	71.27	47.09	178.17
U-Net	VGG16	55.59	71.22	71.70	71.50	26.91	94.95
Proposed method	MobileNetv3	57.21	70.60	75.09	72.76	47.18	14.53

从表 6 可知:所提方法相较于传统 DeepLabv3+、R-PSPNet 和 U-Net 模型,路面裂缝检测的交并比分别提高 0.74 百分点、1.84 百分点和 1.62 百分点,F1 分数分别提高 0.58 百分点、1.49 百分点和 1.26 百分点;模型的参数量分别减小 194.17 MB、163.64 MB 和 80.42 MB;平均帧率分别提高 25.7、0.09 和 20.27。所提方法与 M-PSPNet 模型相比,参数量增加 5.47 MB,平均帧率降低 39.64,M-PSPNet 模型虽效率比所提方法高,但是以牺牲精度为代价,其交并比和 F1 分数比传统 DeepLabv3+ 还低,检测效果不佳。值得注意的是,所提方法牺牲了一些准确率来换取更高的召回率,较于传统 DeepLabv3+、M-PSPNet、R-PSPNet 和 U-Net 模型,准确率分别降低 3.87 百分点、2.88 百分点、1.14 百分点和 0.62 百分点,召回率分别提高 5.06 百分点、8.23 百分点、4.28 百分点和 3.39 百分点。召回率的提高有利于裂缝的检出,裂缝漏判少,利于制定准确的路面养护决策。

综上所述,所提方法能有效提升路面裂缝的检测精度,相较于经典的 DeepLabv3+、M-PSPNet、R-PSPNet 及 U-Net 模型有更高交并比和 F1 分数。虽然所提方法的参数量及平均帧率评价指标不及 M-PSPNet 模型,但其以 57.21% 的交并比、72.76% 的 F1 分数、14.53 MB 的参数量和 47.18 的平均帧率在

保证路面裂缝检测精度的同时满足实时检测要求,取得最优的检测效果,且可有效避免阴影、光照变化和路面污渍等因素的干扰,有较强的鲁棒性。上述检测精度和速度的实验验证了所提方法的有效性。

5.3.3 不同模型检测裂缝的可视化效果对比

所提方法与不同语义分割模型的可视化效果如图 8 所示,图中红色实线方框区域为漏检区域,红色虚线方框区域为误检区域。从图 8 第 1 行和第 2 行可以看出,当图像中裂缝清晰且干扰较少时,这 5 种模型均能较好地检测出裂缝。当图像清晰度不高或图像中裂缝的裂缝较为细小、不够清晰或存在其他干扰时,各个模型检测结果有较大差距。图 8 第 3 行中,由于图像中的裂缝难以辨识,各检测模型都出现了不同程度的漏检,其中,M-PSPNet、R-PSPNet 及 U-Net 漏检较为严重,而所提方法漏检较少,能更多地体现裂缝细节,检测结果与 ground truth 相近。图 8 第 4 行中,M-PSPNet 和 R-PSPNet 漏检较多,检测的裂缝不连续,其中,M-PSPNet 漏检最多,传统 DeepLabv3+、U-Net 和所提方法存在误检,U-Net 误检较严重,所提方法虽有一定程度的误检,但整体检测效果与 ground truth 最相似,检测的裂缝准确、连续。从图 8 第 5~6 行中可以看出,所提方法检测的裂缝细节信息更丰富,与 ground truth 的重合程度最高,基本能够保持裂缝的完整性。

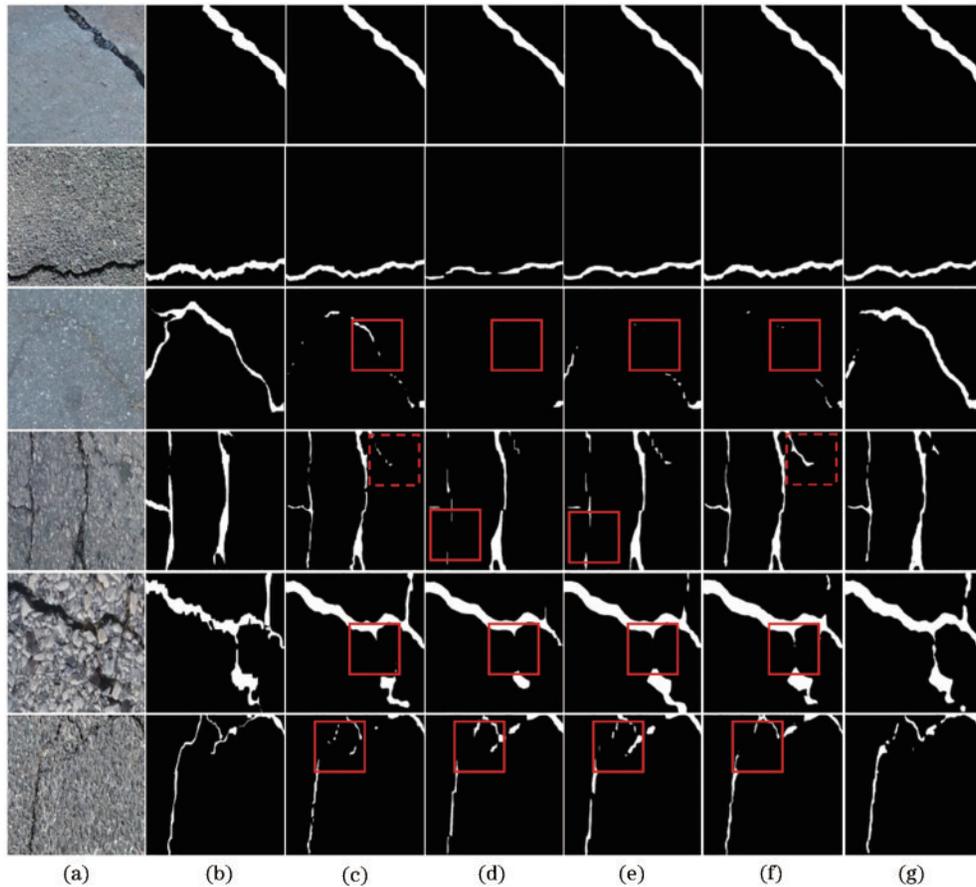


图 8 5 种模型的可视化结果。(a)原图像;(b) ground truth;(c)传统 DeepLabv3+;(d) M-PSPNet;(e) R-PSPNet;(f) U-Net;(g)所提方法

Fig. 8 Visualized results of 5 models. (a) Original images; (b) ground truth; (c) traditional DeepLabv3+; (d) M-PSPNet; (e) R-PSPNet; (f) U-Net; (g) proposed method

6 结 论

检测路面裂缝是修复裂缝的前提,能有效预防裂缝向更严重病害发展,对提高路面的安全性及养护经济性具有重要意义。为实现快速、准确的路面裂缝自动检测,提出基于改进 DeepLabv3+ 模型的路面裂缝自动检测方法,有效避免了路面阴影、污渍及其他因素的干扰,鲁棒性强。

在 1124 张相同测试集上交并比和 F1 分数为判断依据,通过对比实验验证所提方法的有效性与可行性。较于传统 DeepLabv3+ 模型,所提方法的交并比提高 0.74 百分点, F1 分数提高 0.58 百分点,且以 14.53 MB 的参数量和 47.18 的平均帧率在保证路面裂缝检测精度的同时满足实时检测要求,取得最优的检测效果。

为更多地检出裂缝,防止漏检,所提方法在保证综合指标 F1 分数的前提下牺牲了一些准确率来换取更高的召回率。

下一步研究可从模型剪枝和知识蒸馏等方法入手,进一步优化检测模型的体积与精度,提升在不同环境状态下对路面裂缝检测的实时性与准确性。

参 考 文 献

- [1] 马建, 赵祥模, 贺拴海, 等. 路面检测技术综述[J]. 交通运输工程学报, 2017, 17(5): 121-137.
Ma J, Zhao X M, He S H, et al. Review of pavement detection technology[J]. Journal of Traffic and Transportation Engineering, 2017, 17(5): 121-137.
- [2] 王博, 王霞, 陈飞, 等. 航拍图像的路面裂缝识别[J]. 光学学报, 2017, 37(8): 0810004.
Wang B, Wang X, Chen F, et al. Pavement crack recognition based on aerial image[J]. Acta Optica Sinica, 2017, 37(8): 0810004.
- [3] Li P, Wang C, Li S M, et al. Research on crack detection method of airport runway based on twice-threshold segmentation[C]//2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), September 18-20, 2015, Qinhuangdao, China. New York: IEEE Press, 2016: 1716-1720.
- [4] 李鹏, 赵芬芬, 杜敏. 基于双树复小波的直方图路面裂缝检测算法[J]. 安徽大学学报(自然科学版), 2018, 42(1): 38-44.
Li P, Zhao F F, Du M. The cracks detection algorithm of pavement based on histogram of dual-tree complex

- wavelet[J]. Journal of Anhui University (Natural Science Edition), 2018, 42(1): 38-44.
- [5] 韩锬, 韩洪飞. 基于区域级和像素级特征的路面裂缝检测方法[J]. 铁道科学与工程学报, 2018, 15(5): 1178-1186.
- Han K, Han H F. Pavement crack detection method based on region-level and pixel-level features[J]. Journal of Railway Science and Engineering, 2018, 15(5): 1178-1186.
- [6] 赵菲, 邓英捷. 融合多异构滤波器的轻型弱小目标检测网络[J]. 光学学报, 2023, 43(9): 0915001.
- Zhao F, Deng Y J. Light dim small target detection network with multi-heterogeneous filters[J]. Acta Optica Sinica, 2023, 43(9): 0915001.
- [7] Lin J P, Haberstroh F, Karsch S, et al. Applications of object detection networks in high-power laser systems and experiments[J]. High Power Laser Science and Engineering, 2023, 11(1): e7.
- [8] 陈兵, 贺晟, 刘坚, 等. 基于轻量化 DeepLab v3+ 网络的焊缝结构光图像分割[J]. 中国激光, 2023, 50(8): 0802105.
- Chen B, He S, Liu J, et al. Weld structured light image segmentation based on lightweight DeepLab v3+ network[J]. Chinese Journal of Lasers, 2023, 50(8): 0802105.
- [9] 高慧, 阎晓东, 张衡, 等. 基于 Res2Net 的多尺度遥感影像海陆分割方法[J]. 光学学报, 2022, 42(18): 1828004.
- Gao H, Yan X D, Zhang H, et al. Multi-scale sea-land segmentation method for remote sensing images based on Res2Net[J]. Acta Optica Sinica, 2022, 42(18): 1828004.
- [10] 罗晖, 贾晨, 李健. 基于改进 YOLOv4 的公路路面病害检测算法[J]. 激光与光电子学进展, 2021, 58(14): 1410025.
- Luo H, Jia C, Li J. Road surface disease detection algorithm based on improved YOLOv4[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1410025.
- [11] Haciefendioğlu K, Başağa H B. Concrete road crack detection using deep learning-based faster R-CNN method[J]. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2022, 46(2): 1621-1633.
- [12] 孙朝云, 马志丹, 李伟, 等. 基于深度卷积神经网络融合模型的路面裂缝识别方法[J]. 长安大学学报(自然科学版), 2020, 40(4): 1-13.
- Sun Z Y, Ma Z D, Li W, et al. Pavement crack identification method based on deep convolutional neural network fusion model[J]. Journal of Chang'an University (Natural Science Edition), 2020, 40(4): 1-13.
- [13] 彭磊, 张辉. 基于 U-Net 的道路缺陷检测[J]. 计算机科学, 2021, 48(S2): 616-619.
- Peng L, Zhang H. U-Net for pavement crack detection[J]. Computer Science, 2021, 48(S2): 616-619.
- [14] 王保亮, 白少雄, 赵维刚. 基于特征增强学习的路面裂缝病害视觉检测方法[J]. 铁道科学与工程学报, 2022, 19(7): 1927-1935.
- Wang B X, Bai S X, Zhao W G. Pavement crack damage visual detection method based on feature reinforcement learning[J]. Journal of Railway Science and Engineering, 2022, 19(7): 1927-1935.
- [15] 袁嘉豪, 张伟锋, 岳学军, 等. 多种主干网络下 DeepLabv3+ 的混凝土梁裂缝语义分割研究[C]//中国水利学会 2021 学术年会论文集第五分册. 北京: 中国水利学会, 2021: 160-164.
- Yuan J H, Zhang W F, Yue X J, et al. Semantic Segmentation of concrete beam cracks based on DeepLabv3+ in Multiple Backbone Networks [C]//Chinese Hydraulic Society 2021 Annual Conference Proceedings Volume 5. Beijing: Chinese Hydraulic Society, 2021: 160-164.
- [16] 李国燕, 梁家栋, 刘毅, 等. MFC-DeepLabV³⁺: 一种多特征级联融合裂缝缺陷检测网络模型[J]. 铁道科学与工程学报, 2023, 20(4): 1370-1381.
- Li G Y, Liang J D, Liu Y, et al. MFC-DeepLabV³⁺: a multi feature cascade fusion crack defect detection network model[J]. Journal of Railway Science and Engineering, 2023, 20(4): 1370-1381.
- [17] Yang F, Zhang L, Yu S J, et al. Feature pyramid and hierarchical boosting network for pavement crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(4): 1525-1535.
- [18] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Republic of Korea. New York: IEEE Press, 2020: 1314-1324.
- [19] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1577-1586.
- [20] Hou Q B, Zhang L, Cheng M M, et al. Strip pooling: rethinking spatial pooling for scene parsing[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 4002-4011.
- [21] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [22] Nguyen N T H, Le T H, Perry S, et al. Pavement crack detection using convolutional neural network[C]//Proceedings of the 9th International Symposium on Information and Communication Technology, December 6-7, 2018, Danang City, Viet Nam. New York: ACM Press, 2018: 251-256.
- [23] 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.
- [24] Wang W J, Su C. Convolutional neural network-based pavement crack segmentation using pyramid attention network[J]. IEEE Access, 2020, 8: 206548-206558.