

改进 DeepLabv3+ 的道路表面裂缝检测方法

杨萍, 张汐*

(陕西科技大学电子信息与人工智能学院, 陕西 西安 710021)

摘要: 有效的道路表面裂缝检测是维护道路安全、延长道路寿命的关键。针对传统道路表面裂缝检测方法存在的难以识别细小裂缝、分割断裂以及分割精度低等问题,提出了一种改进 DeepLabv3+ 的道路表面裂缝检测方法,旨在降低模型参数数量的同时提高裂缝检测的准确性。首先,使用优化后的 MobileNetv2 网络替换基础 DeepLabv3+ 模型的主干网络,以降低模型的参数数量和复杂度,提高运行速度;其次,将条形池化模块(SPM)融入空洞空间金字塔池化(ASPP)模块,使得网络能够捕获到更多的裂缝上下文信息,保留裂缝细小部分的特征;最后,引入卷积块注意力模块(CBAM),使网络更加关注图像中对裂缝检测起决定作用的像素区域,增强裂缝图像的特征表达能力。实验结果显示,改进 DeepLabv3+ 模型的平均像素准确率(MPA)为 87.85%,平均交并比(MIoU)为 80.53%,准确率为 97.51%,精确率为 88.65%,F1 值为 88.24%,相比于基础 DeepLabv3+ 模型分别提高了 1.77%、2.03%、0.30%、2.25%和 1.51%,且高于 U-Net、HR-Net 和 PSP-Net 模型。此外,改进 DeepLabv3+ 模型的参数数量为 6.382×10^6 ,是基础 DeepLabv3+ 模型的 88.3%,实时性更好,更适用于道路表面裂缝检测任务。

关键词: 裂缝检测;语义分割;卷积神经网络;条形池化模块;注意力机制

源代码链接: https://github.com/ZXiiii/Improved_DeepLabv3

中图分类号: TP391.41

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0069114

Improved DeepLabv3+ Road Surface Crack Detection Method

YANG Ping, ZHANG Xi*

(School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, Shaanxi, China)

【Abstract】 The effective detection of road surface cracks is key to maintaining road safety and prolonging road life. To address the problems of difficulty in identifying small cracks, segmentation fractures, and low segmentation accuracy for traditional road surface crack detection methods, an improved DeepLabv3+ road surface crack detection method is proposed to simultaneously reduce the number of model parameters and improve the accuracy of crack detection. First, the backbone of the DeepLabv3+ model is replaced with an optimized MobileNetv2 network to reduce the number of parameters and complexity of the model, which speeds up the operation. Second, the Strip Pooling Module (SPM) is integrated into the Atrous Spatial Pyramid Pooling (ASPP) module to enable the network to capture more crack context information and preserve the characteristics of small parts of the crack. Finally, a Convolutional Block Attention Module (CBAM) is introduced to make the network focus more on the pixel region that plays a decisive role in crack detection, which enhances the feature expression ability of crack images. According to the experimental results, the improved DeepLabv3+ model achieved a Mean Pixel Accuracy (MPA) of 87.85%, Mean Intersection over Union (MIoU) of 80.53%, accuracy of 97.51%, precision of 88.65%, and F1-Score of 88.24%; compared with the basic DeepLabv3+ model, the improvements are 1.77%, 2.03%, 0.30%, 2.25%, and 1.51%, respectively. These indices of the proposed model are higher than those of the U-Net, HR-Net, and PSP-Net models. In addition, the number of parameters of the improved model is 6.382×10^6 , which is 88.3% of that of the basis model, and the real-time performance is better, making it more suitable for road surface crack detection.

【Key words】 crack detection; semantic segmentation; convolutional neural network; strip pooling module; attention mechanism

0 引言

道路表面裂缝作为道路损坏的主要表现之一,不仅对交通安全构成威胁,还加速了城市道路的老

化和损耗^[1]。传统的人工检测方法耗时且效率低,检测精度取决于检测人员的经验,主观性较强^[2]。因此,开发高效准确的道路表面裂缝检测方法对于城市交通和管理至关重要。

收稿日期: 2023-12-27 修回日期: 2024-02-29

基金项目: 陕西省重点研发计划项目(2023-YBGY-208);陕西省教育厅服务地方专项计划项目(23JC016)。

通信作者 E-mail: * 2630440469@qq.com

传统的道路表面裂缝检测方法通常利用数字图像处理,通过图像预处理和一系列分类提取算法来识别裂缝,如阈值算法^[3]、边缘算法^[4]、区域算法^[5]以及匹配算法^[6]等。此类方法借助机器进行裂缝检测,不仅提高了检测效率,还避免了人为误差,推动了道路表面裂缝检测技术的发展。但此类方法对环境因素(如阴影、光照和路面污渍)较为敏感,稳定性低,在实际应用中存在较大局限。

随着计算机视觉和深度学习的不断发展,卷积神经网络(CNN)在目标检测和图像分割领域得到了广泛应用。在裂缝检测中,CNN的应用主要分为两类:一类是以YOLO^[7]、Faster-RCNN^[8]和SSD^[9]为代表的目标检测方法,这类方法主要通过矩形检测框标注裂缝位置,对形状简单、紧凑的裂缝表现良好,但在处理复杂和狭小裂缝时,矩形检测框的面积过大导致形状检测不准确,不利于道路养护。另一类方法是以FCN^[10]、U-Net^[11-12]、HR-Net^[13]、PSP-Net^[14]和DeepLabv3+为代表的语义分割方法,通过对裂缝图像每个像素进行分类标注,精确定位裂缝的具体位置、形状和大小等信息。其中,DeepLabv3+模型在道路表面裂缝检测方面的应用最为广泛。FU等^[15]提出了一种基于Dense-DeepLabv3+网络的裂缝图像分割算法,在DeepLabv3+中引入密集连接的空洞空间金字塔池化(ASPP)模块,使网络能够获得更密集的像素采样,改进后的算法对裂缝检测的效果优于原算法,具有更高的分割准确率。陈宇平等^[16]在DeepLabv3+架构的基础上提出了一种适用于道路裂缝检测的网络Crack-DeepLab,该网络虽然提升了裂缝识别能力,但相较于MobileNetV2作为主干网络的

DeepLabv3+至少多了300万个参数。黄荣霞等^[17]使用最大连通域协同的改进DeepLabv3+模型对裂缝进行提取,实现裂缝的准确分割,但该模型在某些裂缝特征细微的地方会出现分割断裂的情况。

鉴于此,本文提出一种基于改进DeepLabv3+的道路表面裂缝检测方法。首先,通过将基础DeepLabv3+模型的主干网络Xception替换为优化后的MobileNetv2网络,降低模型参数数量和复杂度;其次,在ASPP模块中引入条形池化模块(SPM),以有效地捕获裂缝结构的上下文信息,防止无关区域的污染信息干扰预测,提高裂缝的分割精度;最后,添加卷积块注意力模块(CBAM),使神经网络更加关注图像中对分类起决定作用的像素区域,忽略无关紧要的区域,增强裂缝的特征表达能力,从而构建一个具有较高检测精度的轻量化道路表面裂缝检测模型。

1 DeepLabv3+模型

DeepLabv3+模型^[18]在DeepLabv3的基础上添加了解码器,构成编码器-解码器结构。在编码器部分,主干网络Xception提取原始图像的特征信息,将浅特征层传入解码器,将深层特征传入ASPP结构中,采用不同的膨胀率进行膨胀卷积和平均池化,对得到的特征图进行拼接融合,再通过 1×1 卷积进行通道数的调整。在解码器部分,对在编码器中获得的深层特征进行4倍上采样,将其与主干网络提取到的浅层特征进行融合,再经过 3×3 卷积和4倍上采样后,输出模型分割后的图像,即输入图像每个像素点的类别。DeepLabv3+模型的网络结构如图1所示(彩色效果见《计算机工程》官网HTML版,下同)。

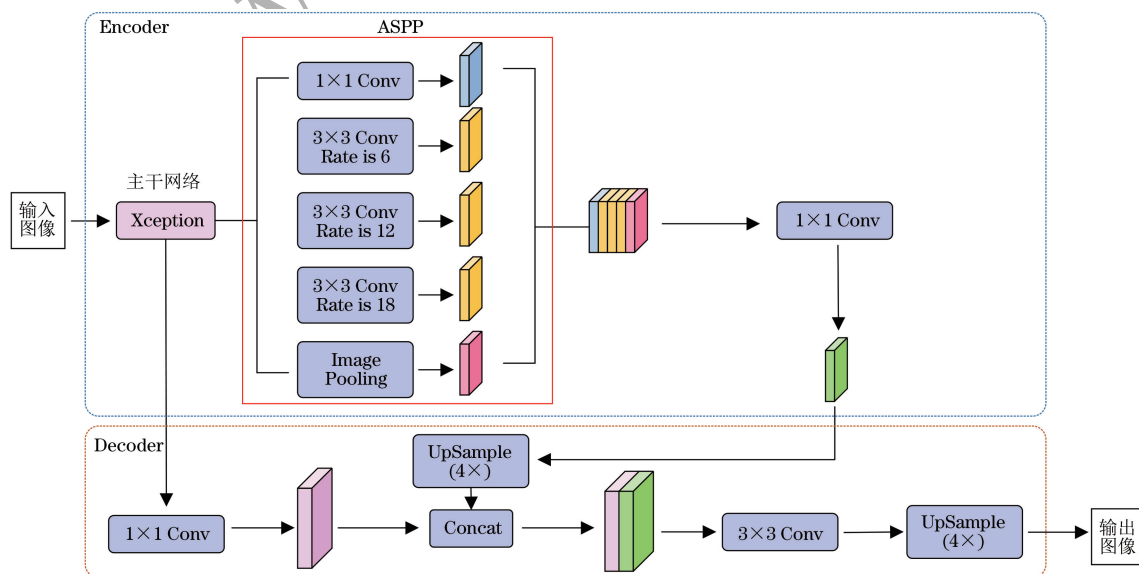


图1 DeepLabv3+模型的网络结构

Fig.1 Network structure of DeepLabv3+ model

2 改进 DeepLabv3+模型

DeepLabv3+模型的主干网络即 Xception 网络结构较为复杂,网络层数多,模型参数量大,很难部署到嵌入式系统;同时,在 DeepLabv3+基础模型的 ASPP 结构中,全局平均池化将整个特征图压缩成单一的值,导致细节信息丢失,获取的裂缝轮廓细节信息少,出现漏分割、多分割、分割断裂现象,造成识别精度低的问题。

为了解决这些问题,本文从以下三方面对

DeepLabv3+基础模型进行了改进。首先,将原 DeepLabv3+模型的主干网络 Xception 替换为优化后的 MobileNetv2 网络,仅采用 MobileNetv2 网络的前 8 层,使模型参数量更少,更适合部署在嵌入式设备上;其次,在 ASPP 模块中引入 SPM,使得网络能够捕获到更多的上下文信息,防止无关区域的污染信息干扰预测,从而提升分割精度;最后,添加 CBAM,使神经网络更加关注图像中对分类起决定作用的像素区域,而忽略无关紧要的区域,增强裂缝的特征表达能力。改进 DeepLabv3+模型的网络结构如图 2 所示。

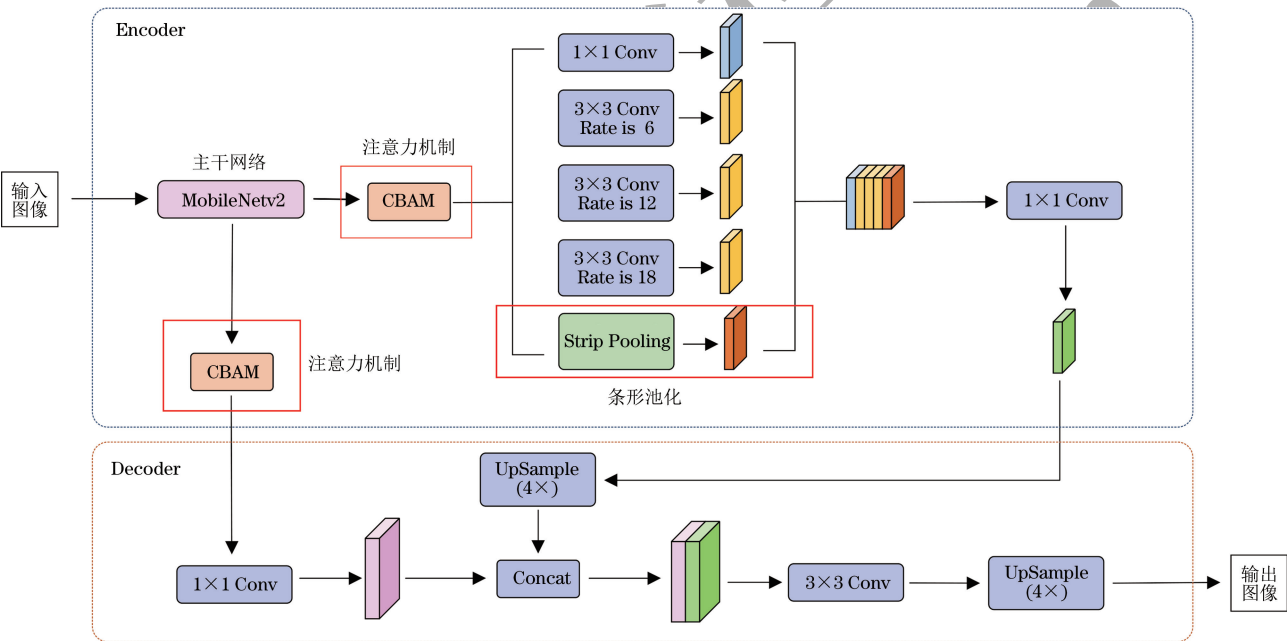


图 2 改进 DeepLabv3+模型的网络结构
Fig. 2 Network structure of improved DeepLabv3+ model

2.1 MobileNetv2 优化网络

为了使模型轻量化,本文将 DeepLabv3+模型的主干网络替换为 MobileNetv2 网络,并对其进行了优化。MobileNetv2^[19]是一种轻量级卷积神经网络,它在 MobileNetv1^[20]的基础上进行了改进,保留了深度可分离卷积(DSC),解决了 MobileNetv1 中 DSC 在输入层卷积核数量固定的瓶颈问题,并且带来了更小的模型体积和更高的计算效率。

本文仅采用 MobileNetv2 的前 8 层,减少计算资源的消耗。此外,第 7 层和第 8 层采用空洞卷积提取特征,并将第 7 层的步长设为 1,以提高 MobileNetv2 网络分割的准确性,优化的 MobileNetv2 网络结构如表 1 所示,表 1 中 Input 表示输入,Operator 表示网络层, t 表示通道倍数, c 表示输出通道的大小, n 表示循环次数, s 表示步长, r 表示膨胀系数。

表 1 优化的 MobileNetv2 网络结构

Table 1 Network structure of optimized MobileNetv2						
Input	Operator	t	c	n	s	r
$512 \times 512 \times 3$	Conv2d	—	32	1	2	1
$256 \times 256 \times 32$	Bottleneck	1	16	1	1	1
$256 \times 256 \times 16$	Bottleneck	6	24	2	2	1
$128 \times 128 \times 24$	Bottleneck	6	32	3	2	1
$64 \times 64 \times 32$	Bottleneck	6	64	4	2	1
$32 \times 32 \times 64$	Bottleneck	6	96	3	1	1
$32 \times 32 \times 96$	Bottleneck	6	160	3	1	2
$16 \times 16 \times 160$	Bottleneck	6	320	1	1	4

2.2 引入条形池化模块

池化是卷积神经网络特征提取的关键步骤,在现有的各种卷积神经网络模型中,常见的池化为平均值池化和最大值池化。这类池化都是使用方形框对输入特征图进行池化操作,只能获取到局部的特征信息,限制了获取上下文信息的灵活性;此外,方形池化核形状单一,会不可避免地将检测物体与背

景之间的特征信息合并在一起,从而造成特征污染问题。如图 3(a)所示,方形池化核的传统池化导致裂缝与背景之间的特征信息混淆在一起,造成了特征污染。而从图 3(b)可以看出,条形池化凭借其独特的内核形状,使网络能够在有效利用长距离关系的同时,消除无关噪声,从而更好地对各种形状的道路表面裂缝进行分割,避免了传统池化存在的问题。图 3(c)是对应的标签图,用于标注裂缝的精确位置。因此,本文在 ASPP 模块中使用条形池化进行池化操作,有效地捕获裂缝结构的上下文信息,防止无关区域的污染信息干扰预测^[21],提高裂缝的分割精度。

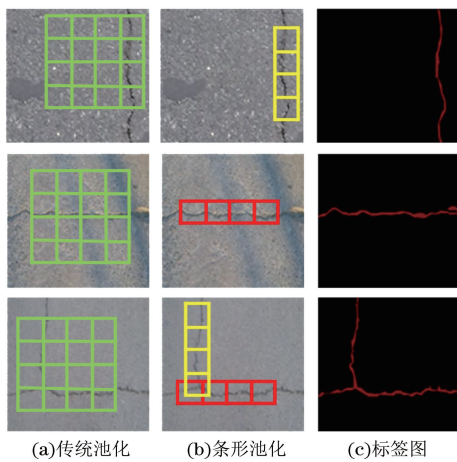


图 3 传统池化和条形池化对比图

Fig.3 Comparison between conventional pooling and strip pooling

图 4(a)为输入特征。以平均值池化为例,在一个平均值池化层中,将输入的图像划分为若干个矩形区域,对每个子区域输出平均值。平均池化的公式为:

$$y_{i_0, j_0} = \frac{1}{h \times w} \sum_{0 \leq i < h} \sum_{0 \leq j < w} x_{i_0+h+i, j_0+w+j} \quad (1)$$

式中: $x_{i_0+h+i, j_0+w+j} \in \mathbb{R}^{H \times W}$ 表示一个输入的二维特征张量; $H \times W$ 分别表示输入特征图的长和宽; $h \times w$ 表示池化核大小; y_{i_0, j_0} 是平均池化后输出的二维特征张量,它的每个空间位置对应一个大小为 $h \times w$ 的池化窗口。

从平均值池化的公式可以看出,此类池化操作会不可避免地在池化过程中涉及到无关区域的信息,而条形池化的出现在一定程度上解决了该问题。条形池化的池化核大小为 $(H, 1)$ 或 $(1, W)$,由水平和垂直两个方向组成,分别通过 $H \times 1$ 和 $1 \times W$ 的池化窗口进行均值池化操作, $x_{i, j} \in \mathbb{R}^{C \times H \times W}$ 表示输入张量,其中 C 为信道数。水平条形池化和垂直条形池化的操作公式分别为:

$$y_i^{(h)} = \frac{1}{W} \sum_{j=0}^W x_{i, j}, y_j^{(v)} = \frac{1}{H} \sum_{i=0}^H x_{i, j} \quad (2)$$

式中: $y_i^{(h)}$ 和 $y_j^{(v)}$ 分别为水平池化的结果和垂直池化的结果, $y_i^{(h)} \in \mathbb{R}^{C \times H}$ 和 $y_j^{(v)} \in \mathbb{R}^{C \times W}$,如图 4(b)、图 4(c)所示;然后分别对 $y_i^{(h)}$ 和 $y_j^{(v)}$ 进行横向和竖向扩充,得到两个 $W \times H$ 的方阵 $y_{c, i}^{(h)}$ 和 $y_{c, j}^{(v)}$,如图 4(d)、图 4(e)所示;最后,将两个方阵进行加法融合,得到如图 4(f)所示的池化结果:

$$y_{c, i, j} = y_{c, i}^{(h)} + y_{c, j}^{(v)} \quad (3)$$

SPM 在条形池化的基础上,进行一次 1×1 卷积和 Sigmoid 激活,得到如图 4(g)所示的 SPM 输出结果 Z :

$$Z = \text{scala}\{x_{i, j}, \sigma[f(y_{c, i, j})]\} \quad (4)$$

式中: scala 是数值乘法; σ 是 Sigmoid 函数。

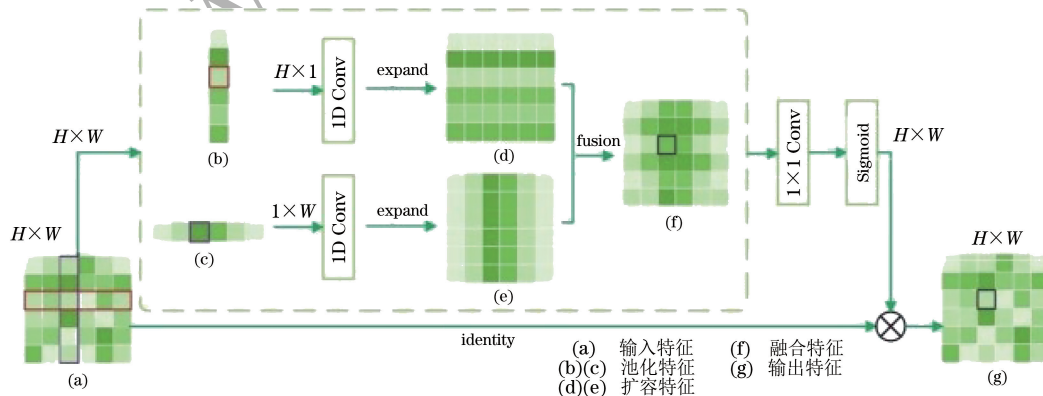


图 4 SPM 结构图

Fig.4 SPM structure diagram

条形池化具有长条形的内核形状,很好地规避了传统池化存在的问题,能够在建立区域间长范围依赖关系的同时,消除不必要的无关噪声,从而降低对目标分割的影响。

2.3 添加 CBAM

DeepLabv3+模型在解码过程中,没有充分融合浅层特征图,导致对局部细节信息提取不足。针对此问题,本文在主干网络之后添加 CBAM,提高

了整体模型的预测准确性。

CBAM 包含 2 个独立的子模块,即通道注意力模块(CAM)和空间注意力模块(SAM),可分别推

断出通道与空间上的注意力权重,将其与输入特征相乘,可对特征进行自适应调整^[22],CBAM 结构如图 5 所示。

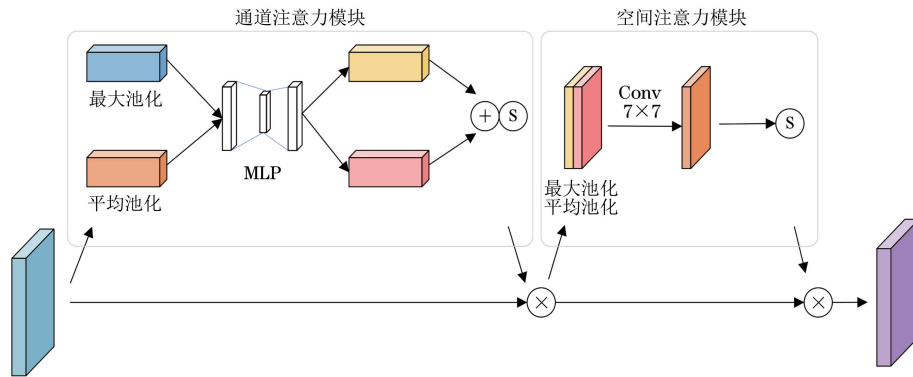


图 5 CBAM 结构图

Fig. 5 CBAM structure diagram

CAM 是对输入的特征 $F(H \times W \times C)$ 分别进行平均池化和最大池化,得到两个 $1 \times 1 \times C$ 的特征,即 $F_{\text{avg}}^{(C)}$ 和 $F_{\text{max}}^{(C)}$ 。然后利用多层感知器(MLP)进行处理,将 MLP 输出的特征进行相加,再经过 Sigmoid 激活操作,获得输入特征层每一个特征点的权值(在 0~1 之间),即 M_c 。最后,将 M_c 和输入特征 F 相乘,生成 SAM 需要的输入特征 F' 。CAM 的计算公式为:

$$M_c(F) = \sigma \{ \text{MLP}[\text{AvgPool}(F)] + \text{MLP}[\text{MaxPool}(F)] \} = \sigma [W_1 W_0 F_{\text{avg}}^{(C)} + W_1 W_0 F_{\text{max}}^{(C)}] \quad (5)$$

式中:AvgPool 表示平均池化;MaxPool 表示最大池化; W_0 和 W_1 表示 MLP 中的两层参数。

SAM 是将 CAM 输出的特征 F' 作为本模块的输入特征。首先进行基于通道维度的平均池化和最大池化,得到两个 $H \times W \times 1$ 的特征 $F_{\text{avg}}^{(s)}$ 和 $F_{\text{max}}^{(s)}$,之后将其进行堆叠。然后经过 7×7 卷积调整通道数($H \times W \times 1$),再经过 Sigmoid 生成权重系数 M_s 。最后将 M_s 和 F' 相乘,得到最终生成的特征。SAM 的计算公式为:

$$M_s(F) = \sigma \{ f^{7 \times 7} [\text{AvgPool}(F'); \text{MaxPool}(F')] \} = \sigma [f^{7 \times 7} (F_{\text{avg}}^{(s)} F_{\text{max}}^{(s)})] \quad (6)$$

式中: $f^{7 \times 7}$ 表示 7×7 的卷积操作。

CBAM 不同于当下流行的其他注意力模块(如 SENet、ECA),它结合了卷积和注意力机制,实现了对图像的空间和通道两个方面的关注。CAM 使网络能够重点学习对整体网络性能具有重大影响的通道信息;SAM 使网络能够更好地捕捉位置关系信息,从而提高网络学习能力。在主干网络之后加入

CBAM,网络可以更好地学习道路表面裂缝图像的特征,降低噪声的影响,减少不相关的信息对模型的干扰;同时,CBAM 的自适应性允许模型在不同的情况下适应道路表面裂缝的变化。

3 实验与结果

3.1 数据集准备

本文使用的是公开裂缝数据集 CFD^[23]、CRACK500^[24]、Cracktree200^[25] 和 DeepCrack^[26],CFD 包含 118 张分辨率为 480×320 像素的裂缝图片,CRACK500 包含 476 张分辨率为 2560×1440 像素的裂缝图片,Cracktree200 包含 206 张分辨率为 800×600 像素的裂缝图片,DeepCrack 包含 537 张分辨率为 384×544 像素的裂缝图片。将 4 个数据集进行整合后,得到 1337 张图像,再使用 OpenCV 库将图像分辨率批量修改为 512×512 像素,目的是在不增加模型复杂度的前提下,提高裂缝检测模型对细节的敏感性。为了避免模型过拟合、提高网络的泛化能力,在这些原裂缝图像的基础上通过裁剪、旋转、翻转、亮度调整和灰度调整等方法对道路表面裂缝数据集进行扩充,并使用开源标注软件 labelme 对每张图像需要识别的裂缝区域进行标注,最后得到含 8000 张图像的增强数据集,按照 8:1:1 的比例随机划分训练集、验证集和测试集。

3.2 实验环境配置及参数设置

实验所用硬件平台为 13th Gen Intel® Core™ i7-13650HX 2.60 GHz 处理器,NVIDIA GeForce RTX 4060 Laptop GPU 显卡,16 GB 内存,软件环境为 Windows 11+cuda 11.6+PyTorch 1.12.0+anaconda3+PyThon 3.7。

合理设置训练参数是提高模型性能的重要手

段。本文设置模型训练迭代次数为 100 次,批处理的训练集样本数量为 4,最大学习率为 0.007,最小学习率为 0.000 07,学习率下降方式选用余弦退火衰减,优化器选用随机梯度下降(SGD)。

3.3 评价指标

本文使用平均像素准确率(A_{MPA})、平均交并比(M_{IoU})、准确率(A)、精准率(P)和 F_1 值(F_1)来评价模型在道路裂缝检测中的性能。

1) A_{MPA} 是一种用于评估图像分割模型性能的度量标准,它衡量了每个类别的分类准确性,并考虑了每个类别的像素数量。 A_{MPA} 的计算公式为:

$$A_{\text{MPA}} = \frac{1}{n+1} \times \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \times 100\% \quad (7)$$

式中: n 表示分类数; N_{TP} 表示作为裂缝被准确地预测为裂缝的数量; N_{FP} 表示作为背景像素被错误地预测为裂缝像素的数量。

2) M_{IoU} 是一种用于评估图像分割模型性能的指标,通过将标签图像和预测图像视为两个集合,计算它们的交集和并集的比值,再对这些比值进行累加并取平均值来计算。 M_{IoU} 的计算公式为

$$M_{\text{IoU}} = \frac{1}{n+1} \times \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}} + N_{\text{FP}}} \times 100\% \quad (8)$$

式中: N_{FN} 表示作为裂缝像素被错误地预测为背景像素的数量。

3) A 是针对所有样本的统计量,指分类正确的样本数占总样本数的比例,是最直接、最易理解的性能指标。 A 的计算公式为:

$$A = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{TP}} + N_{\text{TN}} + N_{\text{FN}} + N_{\text{FP}}} \times 100\% \quad (9)$$

式中: N_{TN} 表示作为背景像素且被正确预测为背景像素的数量。

4) F_1 是 P (P 是指分类正确的正样本数占被判别为正样本的样本个数的比例)和 R (R 是指分类正确的正样本数占实际正样本数的比例)的调和平均

均值,通常用于评估模型的分割性能。 F_1 越大,代表模型的性能越好。 P 、 R 和 F_1 的计算公式分别如下:

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \times 100\% \quad (10)$$

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \times 100\% \quad (11)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \times 100\% \quad (12)$$

3.4 改进模型性能评估

3.4.1 不同模型性能对比

为了验证改进 DeepLabv3+ 模型相对于其他模型的优势,使用相同的参数在 U-Net 模型、HR-Net 模型、PSP-Net 模型和 DeepLabv3+ 模型上进行训练,得到损失函数的变化曲线如图 6 所示。

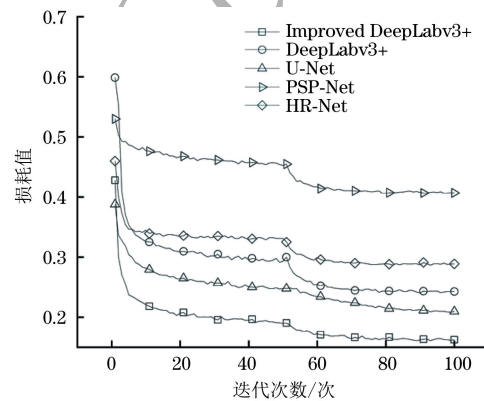


图 6 不同语义分割模型的 Loss 曲线

Fig.6 Loss curves of different semantic segmentation models

总的来看,5 个模型在训练过程中的损失值都随着迭代次数的增加而逐渐稳定,改进 DeepLabv3+ 模型的 Loss 值最小,收敛性更强。

为比较不同语义分割模型对道路表面裂缝识别的精度、参数量和推理时间,在测试集上进行测试后得到 A_{MPA} 、 M_{IoU} 、 A 、 P 、 F_1 、参数量和推理时间,将改进 DeepLabv3+ 模型与 U-Net 模型、HR-Net 模型、PSP-Net 模型和 DeepLabv3+ 模型进行比较,对比结果如表 2 所示,其中粗体显示为最佳结果。

表 2 不同模型的精度、参数量和推理时间对比

Table 2 Comparison of precision, number of parameters, and inference time of different models

模型	$A_{\text{MPA}}/\%$	$M_{\text{IoU}}/\%$	$A/\%$	$P/\%$	$F_1/\%$	参数量/ 10^6	推理时间/ms
U-Net	87.19	78.82	97.12	87.88	86.99	43.933	5.74
HR-Net	81.23	75.82	97.02	87.90	84.59	29.538	6.60
PSP-Net	75.44	70.80	96.45	86.75	80.23	2.376	3.80
DeepLabv3+	86.08	78.50	97.21	86.40	86.73	54.709	4.42
Improved DeepLabv3+	87.85	80.53	97.51	88.65	88.24	6.382	3.25

由表 2 中结果可知:改进 DeepLabv3+ 模型在进行道路表面裂缝分割时,推理时间最短,且

A_{MPA} 、 M_{IoU} 、 A 、 P 和 F_1 均优于其他 4 种模型;改进 DeepLabv3+ 模型的 A_{MPA} 比 U-Net、HR-Net、

PSP-Net、DeepLabv3+ 模型分别提高了 0.66、6.62、12.41、1.77 百分点, M_{IoU} 分别提高了 1.71、4.71、9.73、2.03 百分点, A 分别提高了 0.39、0.49、1.06、0.30 百分点, P 分别提高了 0.77、0.75、1.90、2.25 百分点, F_1 分别提高了 1.25、3.65、8.01、1.51 百分点;同时,改进 DeepLabv3+ 模型的参数量相较于原 DeepLabv3+ 模型减少了 88.3%,相较于 U-Net 减少了 85.5%,相较于 HR-Net 减少了 78.4%;PSP-Net 模型虽然参数量

最少,但精确度较低,不适合进行道路表面裂缝的分割。

3.4.2 消融实验

为验证改进 DeepLabv3+ 模型的不同模块对分割性能的影响,分别在基础网络上修改主干网络,引入 SPM,添加 CBAM,然后进行训练,实验结果见表 3,表中粗体表示最佳结果,Network 代表修改主干网络,+SPM 代表引入 SPM,+CBAM 代表添加 CBAM。

表 3 模型消融实验

Table 3 Model ablation experiments

模型	$A_{MPA}/\%$	$M_{IoU}/\%$	$A/\%$	$P/\%$	$F_1/\%$	参数量/ 10^6	推理时间/ms
DeepLabv3+	86.08	78.50	97.21	86.40	86.73	54.709	4.42
Network	86.17	78.90	97.15	87.15	86.95	5.813	3.97
+SPM	86.13	75.52	96.87	86.67	86.87	61.270	6.83
+CBAM	86.62	79.59	96.88	86.82	86.98	54.709	6.20
Network+SPM	87.43	79.01	96.10	87.09	87.10	6.087	4.12
Network+CBAM	87.29	78.37	96.79	87.59	87.09	6.152	3.84
+SPM+CBAM	87.64	79.22	97.48	88.04	87.48	61.270	5.04
Improved DeepLabv3+	87.85	80.53	97.51	88.65	88.24	6.382	3.25

从表 3 的消融实验结果可以看出,修改主干网络、引入 SPM 和添加 CBAM 均提升了网络性能。其中,修改主干网络为优化的 MobileNetv2 网络后,模型的参数量降低了 89.4%,参数量达到最低。SPM 通过长条形的内核聚合全局和局部上下文信息,避免无关区域干扰预测,使 ASPP 结构保留了更多的道路表面裂缝尾端和纤细部分的特征,只增加了 0.274×10^6 的参数量,对网络的分割性能提升贡献较大。CBAM 内部无大量卷积结构,模块复杂度低,参数量小;同时,CBAM 使神经网络更加关注图像中对分类起决定作用的像素区域,而忽略无关紧要的区域,使得 A_{MPA} 、 M_{IoU} 、 A 、 P 和 F_1 相比于基础 DeepLabv3+ 模型均有提高。

3.4.3 不同模型检测裂缝的可视化效果对比

为了能够更加直观地观察不同模型的分割性能,本文选取了 6 张具有代表性的图片进行可视化对比。改进 DeepLabv3+ 模型与 U-Net、HR-Net、PSP-Net 和 DeepLabv3+ 模型在道路表面裂缝检测方面的效果如图 7 所示。

从图 7 中可以看出,除改进 DeepLabv3+ 模型外的其他 4 种语义分割模型在道路表面裂缝图像分割方面都存在一些漏分割、多分割和分割断裂的问题,且对图像边缘的裂缝分割效果不够理想。从图 7 第 1 行可以看出:当裂缝边缘清晰、干扰较少时,除 PSP-Net 模型外的其他 5 种模型均能检测出

裂缝;但当裂缝较窄、边缘不够清晰以及存在少许干扰时,各模型表现出较大差异。从图 7 第 2~4 行可以看出,模型出现了漏分割的情况,PSP-Net 模型最为严重,而改进 DeepLabv3+ 模型漏检最少,更好地保留了裂缝的细节特征。从图 7 第 5 行可以看出,在检测到裂缝的同时,除改进 DeepLabv3+ 模型外的其他模型均出现了多分割的情况,改进 DeepLabv3+ 模型检测的裂缝更加清晰、准确。从图 7 第 6 行可以看出,模型出现了分割断裂的情况,HR-Net、PSP-Net 及 DeepLabv3+ 模型的断裂情况较多,U-Net 出现少许断裂,改进 DeepLabv3+ 模型几乎没有断裂,基本能够保证裂缝的完整性。

为了验证改进 DeepLabv3+ 模型的鲁棒性,本文在测试集中选取了 4 张含有干扰噪声的图片进行测试。改进 DeepLabv3+ 模型与 U-Net、HR-Net、PSP-Net 和 DeepLabv3+ 模型在复杂背景下对道路表面裂缝检测的效果如图 8 所示(矩形框区域表示有误检区域)。

从图 8 可以看出,图像中的光照阴影、车道线和污渍等道路表面的特性噪声干扰信息会对裂缝的分割能力产生影响。从图 8 可以看出,在复杂背景下,U-Net、HR-Net、PSP-Net 和 DeepLabv3+ 模型会出现多分割、漏分割的情况,本文的改进 DeepLabv3+ 模型对裂缝的分割更为准确、连续,且误检更少,具有一定的鲁棒性。

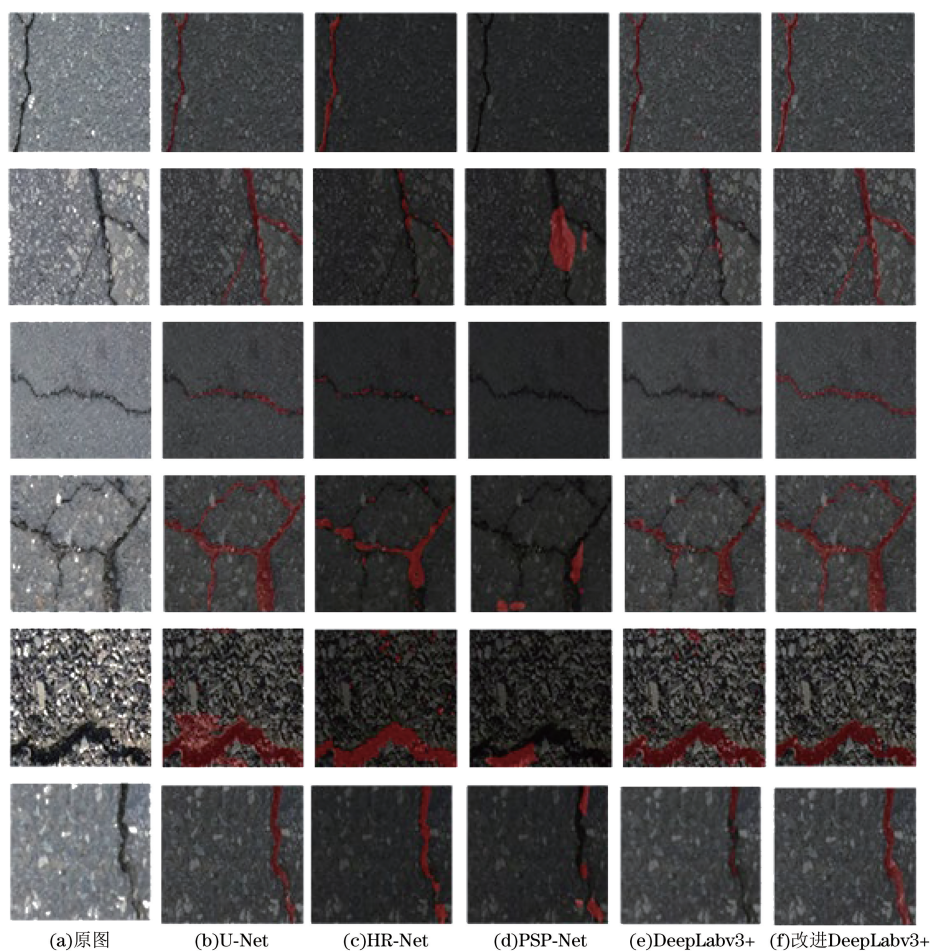


图 7 不同模型检测裂缝的效果图对比

Fig.7 Comparison of effects of different models for crack detection

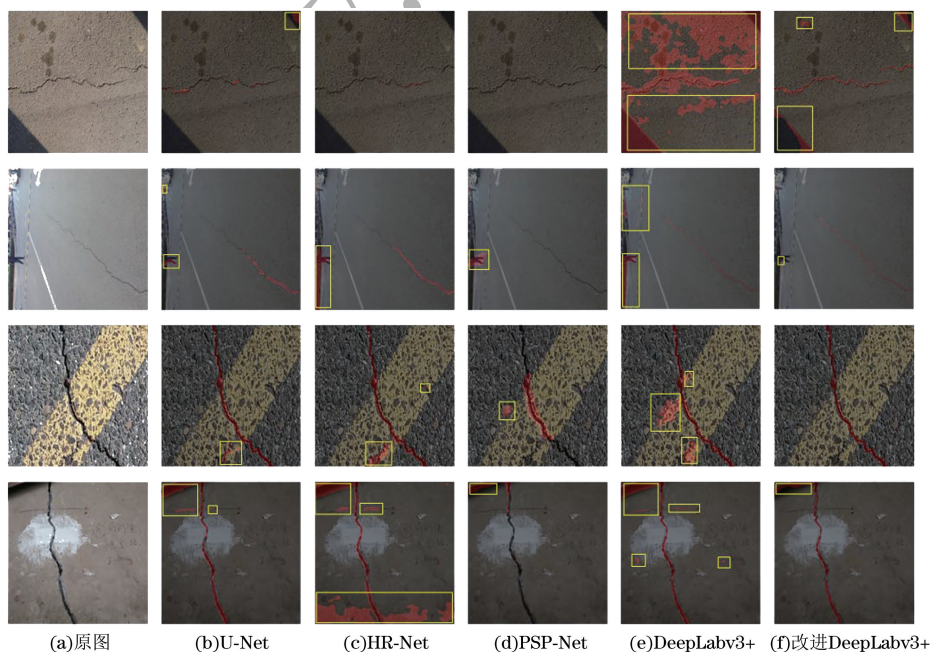


图 8 不同模型在复杂背景下检测裂缝的效果图对比

Fig.8 Comparison of effects of different models in crack detection under complex background

4 结束语

针对目前道路表面裂缝检测存在的难以识别细

小裂缝、分割断裂以及分割精度低等问题,本文提出了一种改进 DeepLabv3+模型的道路表面裂缝检测方法。首先,通过修改主干网络来减少模型参数量,

加快运行速度,让模型能够部署于嵌入式系统;其次,针对分割断裂的问题优化 ASPP 结构,引入 SPM,使得网络能够捕获到更多的裂缝上下文信息,防止无关区域的污染信息干扰预测;最后,通过添加 CBAM,使神经网络更加关注图像中对分类起决定作用的像素区域。基于此,本文改进模型解决了难以识别细小裂缝、边缘检测效果差的问题,提高了道路表面裂缝检测的性能。实验结果显示,改进 DeepLabv3+ 模型在 A_{MPA} 、 M_{IoU} 、 A 、 P 和 F1 值等方面都优于对比模型,并且其参数量和平均每张图片的网络推理时间最短,具有更高的检测精度,有助于推动路面裂缝检测任务的信息化和智能化发展。后续的工作将继续改进 DeepLabv3+ 模型,扩充及改善道路表面裂缝数据集,以进一步提升在不同环境下模型对道路表面裂缝检测的准确性。

参考文献

- [1] XIANG X, ZHANG Y, EL SADDIK A. Pavement crack detection network based on pyramid structure and attention mechanism[J]. IET Image Processing, 2020, 14(8): 1580-1586.
- [2] KANG D H, CHA Y J. Efficient attention-based deep encoder and decoder for automatic crack segmentation[J]. Structural Health Monitoring, 2022, 21(5): 2190-2205.
- [3] 陈浩瀚, 谢仁平, 魏文红. 基于区域生长和梯度阈值分割的路面裂缝提取算法[J]. 东莞理工学院学报, 2022, 29(3): 64-68.
CHEN H H, XIE R P, WEI W H. Crack extraction algorithm for road image based on region growing algorithm and gradient segmentation threshold[J]. Journal of Dongguan University of Technology, 2022, 29(3): 64-68. (in Chinese)
- [4] YANG C, GENG M. The crack detection algorithm of pavement image based on edge information[C]//Proceedings of 6th International Conference on Computer-Aided Design, Manufacturing, Modeling and Simulation. USA, New York: AIP Publishing, 2018: 040023.
- [5] 李鹏, 李强, 马味敏, 等. 基于 K-means 聚类的路面裂缝分割算法[J]. 计算机工程与设计, 2020, 41(11): 3143-3147.
LI P, LI Q, MA W M, et al. Pavement crack segmentation based on K-means clustering[J]. Computer Engineering and Design, 2020, 41(11): 3143-3147. (in Chinese)
- [6] 陶健, 田霖, 张德津, 等. 基于局部纹理特征的沥青路面裂缝检测方法[J]. 计算机工程与设计, 2022, 43(2): 517-524.
TAO J, TIAN L, ZHANG D J, et al. Asphalt pavement crack detection method based on local texture features[J]. Computer Engineering and Design, 2022, 43(2): 517-524. (in Chinese)
- [7] 郝巨鸣, 杨景玉, 韩淑梅, 等. 引入 Ghost 模块和 ECA 的 YOLOv4 公路路面裂缝检测方法[J]. 计算机应用, 2023, 43(4): 1284-1290.
HAO J M, YANG J Y, HAN S M, et al. YOLOv4 highway pavement crack detection method using Ghost module and ECA[J]. Journal of Computer Applications, 2023, 43(4): 1284-1290. (in Chinese)
- [8] 付强, 卜凡民, 任洪鹏, 等. 基于深度学习的路面裂缝目标检测[J]. 公路, 2023, 68(9): 395-405.
FU Q, PU F M, REN H P, et al. Pavement crack target detection based on deep learning method[J]. Highway, 2023, 68(9): 395-405. (in Chinese)
- [9] 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. (in Chinese)
- [10] 翁飘, 陆彦辉, 齐宪标, 等. 基于改进的全卷积神经网络的路面裂缝分割技术[J]. 计算机工程与应用, 2019, 55(16): 235-239, 245.
WONG P, LU Y H, QI X B, et al. Pavement crack segmentation technology based on improved fully convolutional networks[J]. Computer Engineering and Applications, 2019, 55(16): 235-239, 245. (in Chinese)
- [11] 杨秋媛, 李宁, 石林, 等. 基于空洞卷积与动态多核融合池化的裂缝检测[J]. 计算机工程与设计, 2022, 43(12): 3529-3537.
YANG Q Y, LI N, SHI L, et al. Crack detection based on dilated convolution and dynamic multi-kernel fusion pooling module[J]. Computer Engineering and Design, 2022, 43(12): 3529-3537. (in Chinese)
- [12] 于海洋, 景鹏, 张文涛, 等. 基于残差与注意力机制的道路裂缝检测 U-Net 改进模型[J]. 计算机工程, 2023, 49(6): 265-273.
YU H Y, JING P, ZHANG W T, et al. Improved U-Net model for road crack detection based on residual and attention mechanism[J]. Computer Engineering, 2023, 49(6): 265-273. (in Chinese)
- [13] 张伯树, 张志华, 张洋. 改进的 HRNet 应用于路面裂缝分割与检测[J]. 测绘通报, 2022(3): 83-89.
ZHANG B S, ZHANG Z H, ZHANG Y. Improved HRNet applied to segmentation and detection of pavement crack[J]. Bulletin of Surveying and Mapping, 2022(3): 83-89. (in Chinese)
- [14] 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): 2210001.
LI L F, WANG N, WU B, et al. Segmentation algorithm of bridge crack image based on modified PSPNet[J]. Laser and Optoelectronics Progress, 2021, 58(22): 2210001. (in Chinese)
- [15] FU H, MENG D, LI W, et al. Bridge crack semantic segmentation based on improved Deeplabv3+[J]. Journal of Marine Science and Engineering, 2021, 9(6): 671.
- [16] 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.
CHEN Y P, FAN G. Road crack detection based on improved DeepLabV3+ in complex environments[J]. Journal of Guangzhou University (Natural Science Edition), 2023, 22(2): 43-51. (in Chinese)
- [17] 黄荣霞, 刘德儿. 最大连通域协同的改进 Deeplabv3+ 路面裂缝检测[J]. 计算机仿真, 2023, 40(5): 182-186.
HUANG R X, LIU D E. Improved Deeplabv3+ pavement crack detection based on maximum connection region collaboration[J]. Computer Simulation, 2023, 40(5): 182-186. (in Chinese)
- [18] LIU R, HE D. Semantic segmentation based on Deeplabv3+ and attention mechanism[C]//Proceedings of 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference. Washington D. C., USA: IEEE Press, 2021: 255-259.
- [19] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottle-necks[C]//Proceedings of IEEE Conference on Computer Vision and

- Pattern Recognition. Washington D. C. , USA: IEEE Press, 2018; 4510-4520.
- [20] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2023-11-27]. <https://arxiv.org/abs/1704.04861>.
- [21] 李强龙, 周新文, 位梦恩, 等. 基于条形池化和注意力机制的街道场景红外目标检测算法[J]. 计算机工程, 2023, 49(8): 310-320.
- LI Q L, ZHOU X W, WEI M E, et al. Infrared target detection algorithm based on strip pooling and attention mechanism in street scenes [J]. Computer Engineering, 2023, 49(8): 310-320. (in Chinese)
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[C]//Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer, 2018; 3-19.
- [23] SHI Y, CUI L, QI Z, et al. Automatic road crack detection using random structured forests[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(12): 3434-3445.
- [24] 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, 2020, 21(4): 1525-1535.
- [25] ZOU Q, CAO Y, LI Q, et al. CrackTree: automatic crack detection from pavement images [J]. Pattern Recognition Letters, 2012, 33(3): 227-238.
- [26] LIU Y, YAO J, LU X, et al. DeepCrack: a deep hierarchical feature learning architecture for crack segmentation [J]. Neurocomputing, 2019, 338: 139-153.

编辑 陆燕菲