

基于改进 PSPNet 的桥梁裂缝图像分割算法

李良福, 王楠*, 武彪, 张晰

陕西师范大学计算机科学学院, 陕西 西安 710119

摘要 针对传统桥梁裂缝检测算法检测精度低和现有的主流语义分割算法容易丢失裂缝图像细节信息、结果不连续等问题,提出了一种基于改进 PSPNet 的桥梁裂缝图像分割算法。首先使用无人机采集桥梁图像,通过图像增强处理得到桥梁裂缝数据集;其次通过带有扩张卷积的残差网络初步提取裂缝特征;接着将提取到的特征送入到空间位置自注意力模块(SPAM)和金字塔池化模块的串联结构中,使其能够在空间维度上获得丰富的上下文信息。实验结果表明,与现有的主流语义分割算法相比,所提算法得到的裂缝细节更加丰富,各项分割指标都有较为显著的提升,平均交并比达到 84.31%,并能对细小桥梁裂缝进行准确、完整提取。

关键词 图像处理; 桥梁裂缝检测; 自注意力机制; 金字塔池化; 残差网络

中图分类号 TP391.9

文献标志码 A

doi: 10.3788/LOP202158.2210001

Segmentation Algorithm of Bridge Crack Image Based on Modified PSPNet

Li Liangfu, Wang Nan*, Wu Biao, Zhang Xi

School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China

Abstract This study proposes a bridge crack image segmentation algorithm based on modified PSPNet to resolve the problems such as the low detection accuracy of the traditional bridge crack detection algorithms, loss of details in crack images and discontinuous findings of the existing mainstream semantic segmentation algorithms. First, the bridge images are acquired using an unmanned aerial vehicle, and the bridge crack datasets are procured via image enhancement processing. Second, the crack features are initially extracted using the residual network with dilated convolution. Then, the extracted features are sent to the serial structure of the spatial position self-attention module (SPAM) and pyramid pooling module, enabling the features to achieve rich contextual information in spatial dimensions. The experimental results reveal that the proposed algorithm obtains more precise crack details compared with the existing mainstream semantic segmentation algorithms, with each segmentation index being greatly improved, reaching 84.31% on mean intersection over the union. The proposed algorithm can extract small bridge cracks accurately and completely.

Key words image processing; bridge crack detection; self-attention mechanism; pyramid pooling; residual network

OCIS codes 100.2000; 100.4996; 100.1830; 150.1135

1 引言

桥梁为人类扩大活动范围提供媒介,已成为人类扩大生存空间的重要渠道,极大地促进了社会发

展^[1]。改革开放以来,中国桥梁工程建设跨入黄金时期。据不完全统计,2019 年和 2020 年中国大跨度建成通车的桥梁接近 20 座。交通运输事业是我国经济的重要支撑,是社会发展的强力保障,随着交

收稿日期: 2020-12-19; 修回日期: 2021-01-10; 录用日期: 2021-01-21

基金项目: 国家自然科学基金(61573232, 61401263)

通信作者: *1763610462@qq.com

通运输业的不断发展,桥梁不仅承担着促进经济发展的重要功能,还关系着运输人员的人身安全。近年来,国内外大桥垮塌事故频繁发生,引起了我们对桥梁安全的高度重视。裂缝作为桥梁最常见的缺陷,是桥梁结构达到承载能力极限的标志,严重影响着桥梁的安全运营。因此针对桥梁裂缝检测方法的研究具有极为重要的意义。

随着计算机数字图像处理和计算机视觉的飞速发展,国内外科研人员开始将这些新兴的技术方法用于裂缝检测任务中,取得了不错的研究结果。在传统的图像处理领域, Li 等^[2]首先使用直方图均衡化来解决照明不均匀影响裂缝检测的问题,其次使用改进的 Canny 算法绘制裂缝轮廓。该方法对阴影条件下的裂缝图像具有很好的检测效果,但是当裂缝较为宽大时,检测效果降低。Qu 等^[3]基于改进的渗透模型,使用亮度特征和裂缝长度特征来去除噪声区域,用来检测不清晰的裂缝图像。Amhaz 等^[4]提出了一种基于最小代价路径搜索的二维路面图像裂缝检测算法,但是该算法需要对多个参数进行优化,从而增加了算法运行时间。近几年,深度学习成为计算机领域的研究热点,国内外学者开始将深度学习与裂缝检测结合在一起^[5-13]。Kang 等^[14]提出了一种基于 Faster R-CNN 的裂缝区域自动检测、定位算法,该算法使用改进的管状流场算法对裂缝区域进行定位,进而从检测到的裂缝区域中分割出裂缝像素。李良福等^[15]利用滑动窗口算法将裂缝图像切分成更小的图像块,再使用卷积神经网络进行分类,最后使用改进的滑动窗口算法进行检测,取得了较好的检测效果。Zhang 等^[16]使用卷积神经网络将图像切成若干小块,获得每块的裂缝概率图,但该模型忽视了各个像素之间的空间关系,使得结果图中裂缝宽度大,精确度仍需提高。Long 等^[17]将全卷积网络应用于像素级裂缝检测,使用反卷积进行上采样,弥补细节损失,但是没有考虑像素之间的关系,缺乏空间一致性。Zhao 等^[18]提出 PSPNet,设计了一个金字塔池模块来收集有效的上下文信息,包含不同尺度的信息。Badrinarayanan 等^[19]提出了 SegNet 模型,利用池化索引保证了高层信息的完整性,但是对小目标进行上采样时,仍会丢失部分边界信息。

卷积运算会产生局部感受野,相同标签像素对应的特征可能会有所不同,这些差异会导致类内不一致,影响识别的准确性,同时经多次池化下采样操作,图像的细节信息也会逐步丢失。为了解决这些

问题,本文提出了一种基于 PSPNet 改进的桥梁裂缝检测模型。该方法能够充分考虑各像素之间的关系,在特征与注意力之间建立关联来探索全局上下文信息,在空间维度上对裂缝语义信息进行处理,提升了桥梁裂缝检测的精度。

2 数据集的采集和扩增

2.1 图像采集

所采用的桥梁裂缝数据集是大疆 Mavic 2 行业双光版无人机自带的 1200 万像素相机采集的。相机距离桥梁表面的高度越高,采集到的范围就越大,就会导致裂缝图像分辨率下降,以至于细小裂缝不能清晰地展示在图片中。经过多次测试,最终确定无人机镜头距离裂缝表面 40 cm 时采集到的图像质量最佳。首先将无人机悬停在裂缝位置的上方,调整云台位置,使得相机镜头平行于裂缝表面;接着将无人机从悬停状态转为沿着裂缝区域进行平稳飞行状态,每隔 1 s 拍摄一次图像,一共采集 20000 张图像,其中包含了单根裂缝、交叉裂缝、网状裂缝、无桥梁裂缝等类型的图片。对采集到的图像进行归一化,得到分辨率大小为 512×512 的图像。经过筛选,最终选出 2000 张图像作为原始数据集。

为了保证数据集图像的质量,需要对阴天、雨后等特殊天气拍摄的低质量图像进行预处理。图 1(a)为光照不足的情况下采集到的裂缝图像。使用拉普拉斯锐化^[20]来提升裂缝图像对比度和裂缝边缘细节,如图 1(b)所示,可以看到,经过处理后的图像与原始图像相比视觉效果明显提升。

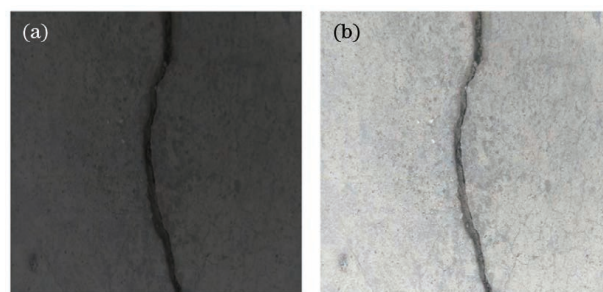


图 1 低质量图像预处理。(a)原图;(b)拉普拉斯锐化后的结果

Fig. 1 Low-quality image preprocessing. (a) Original image; (b) result obtained by Laplacian sharpening

2.2 数据集的扩增

使用深度学习进行桥梁裂缝分割时需要大量带有标签的数据作为数据集,如果直接使用采集来的

2000 张桥梁裂缝图像去训练网络模型,将会导致网络模型过拟合。若是完全采用人工的方式采集大量的桥梁裂缝图像组建数据集,不仅任务繁重,而且效率低下。针对这个问题,使用数字图像处理的方法

对原始桥梁裂缝数据集进行数量规模的扩增。对已采集到的桥梁裂缝图像数据集进行图像几何变换的扩充,如旋转 180° 和 270° 、水平翻转、垂直翻转等,如图 2 所示。

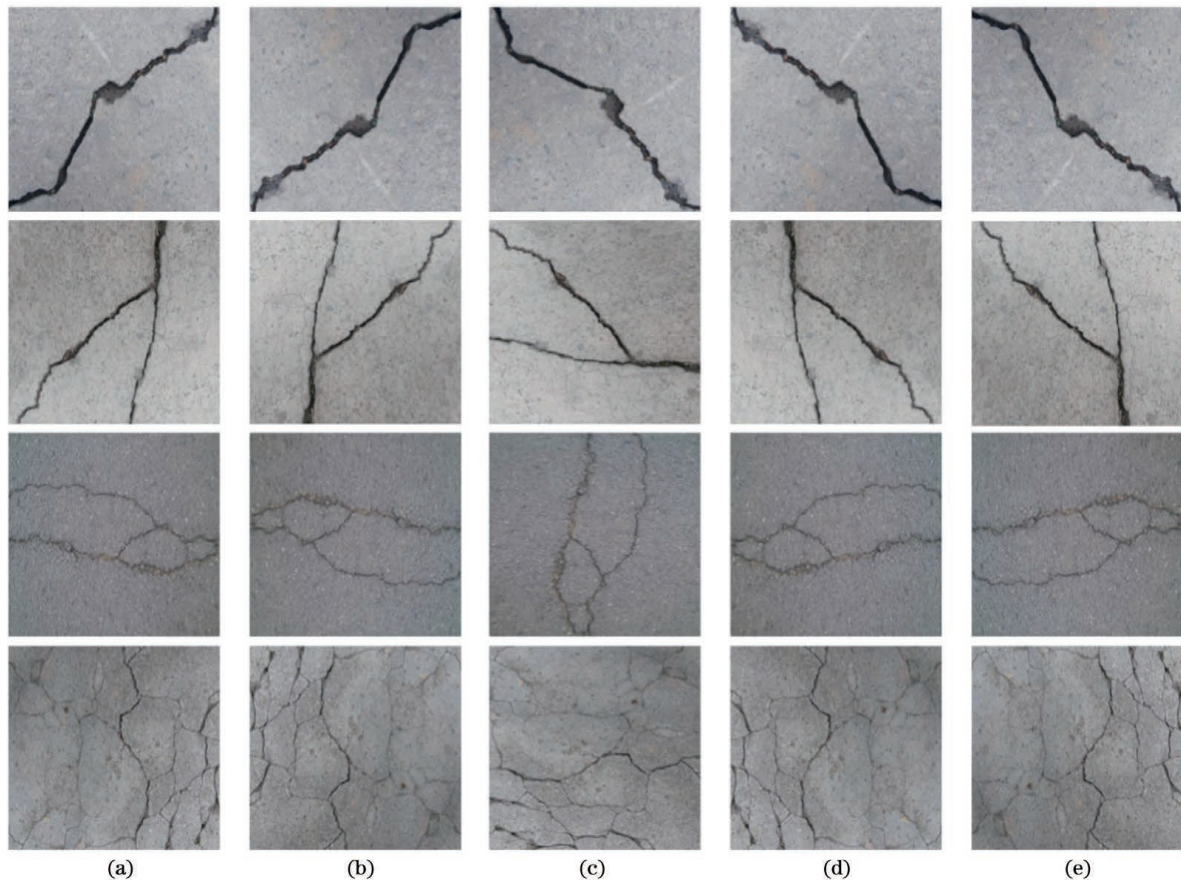


图 2 桥梁裂缝数据集扩增示意图。(a)原图;(b)旋转 180° ;(c)旋转 270° ;(d)水平翻转;(e)竖直翻转

Fig. 2 Schematic of bridge crack dataset amplification. (a) Original image; (b) rotate 180° ; (c) rotate 270° ; (d) horizontal flip; (e) vertical flip

经过数据集扩充方法处理后,桥梁裂缝图像最终的数量达 10000 张。将扩增后的数据按照一定的比例划分为训练集、测试集和验证集。为了确保数据集的多样性,在划分时考虑到每个集合都应该包含单根裂缝、交叉裂缝、网状裂缝、无桥梁裂缝等类型的图像,经过划分后,桥梁裂缝检测模型的训练集有 8000 张图像,验证集和测试集各有 1000 张图像。

3 基于改进 PSPNet 网络的桥梁裂缝检测方法

3.1 PSPNet 网络

随着全卷积网络应用于语义分割中,一系列基于深度学习的语义分割网络得到了飞速发展,例如 U-Net 和 DeepLab 系列的网络。但是这些网络经

过多次卷积和池化等下采样操作,图像的细节信息会逐步丢失,进而影响识别的准确性。而 PSPNet 模型采用残差网络作为特征提取网络,并使用金字塔池化模块收集不同区域、不同尺度的信息,在空间精度上有了较好的提升。整体结构如图 3 所示。

虽然 PSPNet 使用金字塔池化层,在一定程度上缓解了不同子区域间上下文信息丢失的问题,但该网络仅对分辨率较小的裂缝高层特征图进行解码,没有充分考虑像素之间的关系,忽略了底层空间细节特征对裂缝图像分割产生的影响。而桥梁裂缝背景复杂多样,噪声种类繁多,分布及种类毫无规律,若是将 PSPNet 模型直接用于桥梁裂缝检测任务中,将会出现细小裂缝和复杂裂缝检测不完整等问题。对语义分割任务来说,判别特征十分重要。

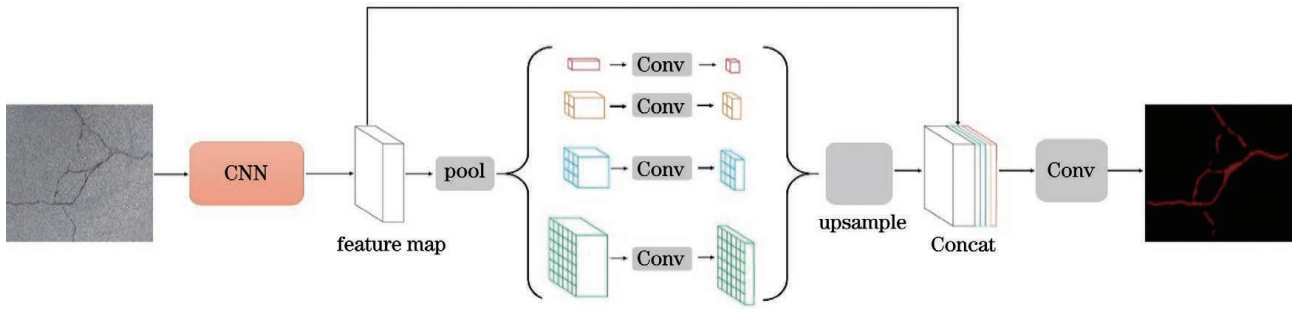


图 3 PSPNet 结构

Fig. 3 PSPNet structure

目前主流的语义分割网络提取特征时多以局部特征为主,这不利于长期上下文信息获取,容易造成目标和物体类内的分类错误,影响检测精度。

3.2 基于改进 PSPNet 的桥梁裂缝检测模型

在 PSPNet 网络上引入自注意力机制^[21]来获

得丰富的上下文信息,使得不同位置的相似特征可以相互增强,整体结构如图 4 所示。将一种基于自注意力机制的空间位置自注意力模块 (SPAM) 嵌入到残差特征提取模块和金字塔池化模块之间,形成串联结构。

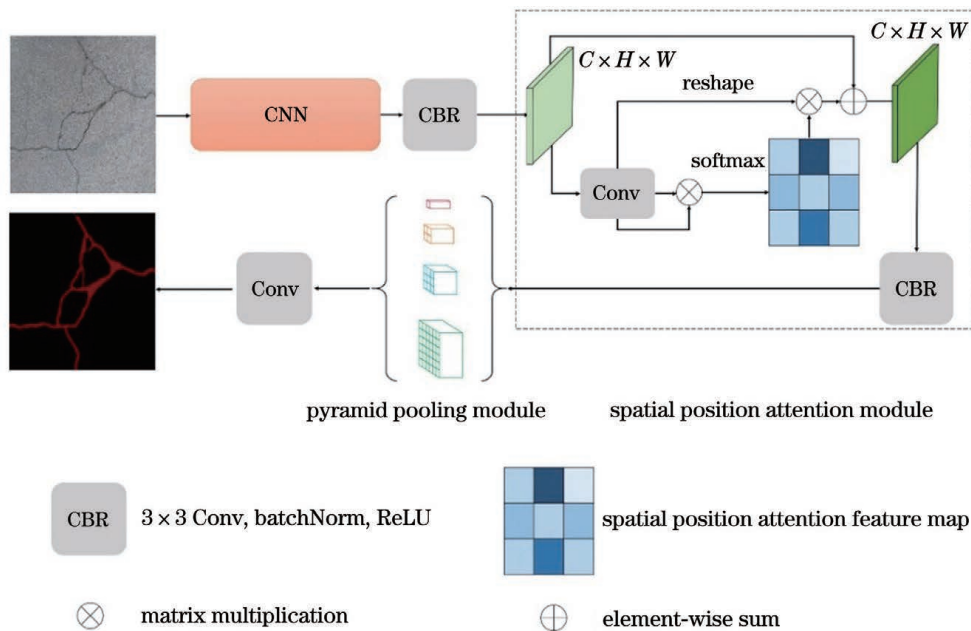


图 4 所提网络的结构

Fig. 4 Proposed network structure

首先使用 ResNet-50 作为特征提取网络。为了减少计算量,将输入部分的 7×7 卷积核换成了 3 个 3×3 卷积核。并且针对残差网络不断的池化和下采样导致特征图分辨率下降,网络丢失敏感细节的问题,采用扩张卷积来替换残差网络内部的下采样层,保持特征图的空间分辨率,如图 5 所示。扩张卷积就是在基础卷积核中加入间隔,在保持参数个数不变的情况下增大卷积核的感受野,同时它可以保证输出的特征映射大小保持不变,提高了检测结果的准确率。在标准残差网络中,输入大小为 512×512 的图像时,输出特征图的大小只有 16×16 。这

里通过引入扩张卷积,对于输入大小为 512×512 的图像,输出特征图的大小可以达到 64×64 ,能够保留更多的细节信息。

将残差网络提取到的裂缝特征图送入到 SPAM 中获取更加丰富的上下文信息,使得不同位置的相似特征可以相互增强,从而增加了类内的紧凑性和语义一致性。计算机视觉自注意力机制和人类的注意力机制非常类似,基本思想就是让模型能重点关注一些信息而忽略掉大量无关的信息。SPAM 具体结构如图 6 所示。

首先将具有扩张卷积的残差网络提取出的局部

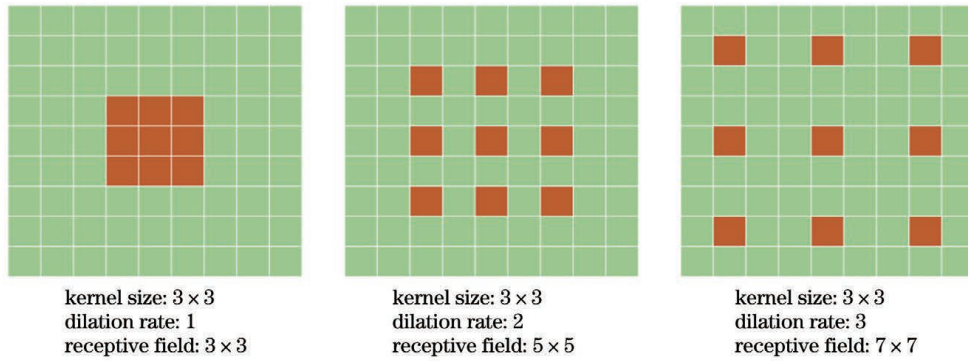
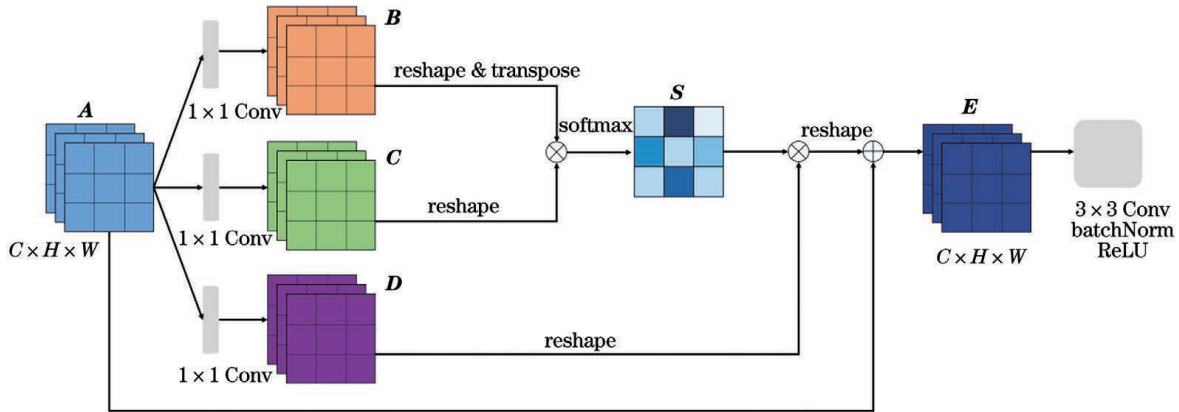
图 5 扩张率为 1,2,3 的 3×3 卷积核感受野示意图Fig. 5 Schematic of the 3×3 convolution kernel receptive field with expansion rates of 1, 2, 3

图 6 SPAM 结构

Fig. 6 SPAM structure

特征 $A \in \mathbf{R}^{C \times H \times W}$ 分别送入到三个 1×1 卷积层中, 减少了通道数和网络的计算量。产生的 3 个新的特征图 B, C, D 的大小均为 $\mathbf{R}^{C \times N}$ 。将 B 和 C 重塑为 $\mathbf{R}^{N \times C}$, 其中像素点的数目 $N = H \times W$ 。同时对 C 的转置 $\mathbf{R}^{N \times C}$ 和 B 进行矩阵乘法, 得到 $S \in \mathbf{R}^{N \times N}$, 再进行 softmax 运算, 得到空间注意力映射矩阵。

$$s_{ji} = \frac{\exp(B_i \cdot C_j)}{\sum_{i=1}^N \exp(B_i \cdot C_j)}, \quad (1)$$

式中: B_i 代表矩阵 B 中的第 i 个元素; C_j 代表矩阵 C 中第 j 个元素; s_{ji} 代表第 i 个位置的特征对第 j 个位置的影响, 如果这两个位置恰好都是属于同一类的像素点, 那么 s_{ji} 就会产生很大的值, 这样就会突出相似特征之间的联系。接着将 D 也重塑为 $\mathbf{R}^{C \times N}$, 与 S 的转置矩阵 $\mathbf{R}^{N \times N}$ 进行矩阵乘法, 再乘上自注意力系数 α , 将其结果重塑为 $\mathbf{R}^{C \times H \times W}$ 。最后与特征图 $A \in \mathbf{R}^{C \times H \times W}$ 进行逐元素相加, 得到输出 $E \in \mathbf{R}^{C \times H \times W}$, 再经过一个卷积层, 得到最终输出。

$$E_j = \alpha \sum_{i=1}^N (s_{ji} D_i) + A_j. \quad (2)$$

将 α 初始化为 0, 通过网络训练逐渐学习到更

大的权重。由(2)式可知, 得到的每个位置的特征 E 是所有位置的特征与原始特征的加权和, 因此它具有全局语境观, 根据空间注意力图选择性地聚合语境。

最后将经 SPAM 计算后的特征图送入到金字塔池化模块中。该模块包含 4 个池化层, 其大小分别为 $1 \times 1, 2 \times 2, 3 \times 3$ 和 6×6 , 然后分别在这些子区域内使用全局平均池化, 提取局部特征。对每层池化使用 1×1 卷积降维后, 通过双线性插值对低维的特征图进行上采样, 得到与输入特征图相同尺寸的特征图, 并且在通道上对不同层级的特征图进行拼接, 得到最后的裂缝分割结果。

4 实验结果与分析

所提算法使用 Python, C++ 语言实现, 深度学习框架采用 Pytorch 1.6.0。操作系统为 Ubuntu20.04 LTS, 实验平台硬件条件为 Ryzen 5 3600 处理器, 主频为 3.6 GHz, RAM16G, NVIDIA GeForce GTX 2080Ti 显卡, 显存 11 G。数据集图像分辨率设置为 512×512 。

4.1 训练细节和实验设置

在训练过程中,选择随机梯度下降(SGD)算法进行优化, Batch-size 设置为 16, 初始学习率设为 0.01, 动量设置为 0.9, 权重衰减为 0.0005。训练时采用 Ploy 学习策略, 其中初始学习率在每 50 次迭代之后乘以 $(1 - N_{iter} / N_{totaliter})^{0.9}$, 其中 $N_{totaliter}$ 为总的迭代次数, 最小学习率设为 1×10^{-4} 。最大迭代次数设置为 40000, 每迭代 4000 次保存一次训练结果, 同时使用验证集进行验证。

由于桥梁裂缝语义分割是一个二分类的问题,

只有裂缝和背景两个类别, 所以采用二分类交叉熵作为分割网络的损失函数。二分类交叉熵损失函数为

$$L_{bce} = \frac{1}{M} \sum_m - [r_m \cdot \log(p_m) + (1 - r_m) \cdot \log(1 - p_m)], \quad (3)$$

式中: r_m 表示像素点的真实值; p_m 表示像素点 m 预测值的概率; M 为图像像素点总数目。训练过程中的损失曲线如图 7 所示, 结果表明, 所提算法可使数据集的训练损失迅速收敛。

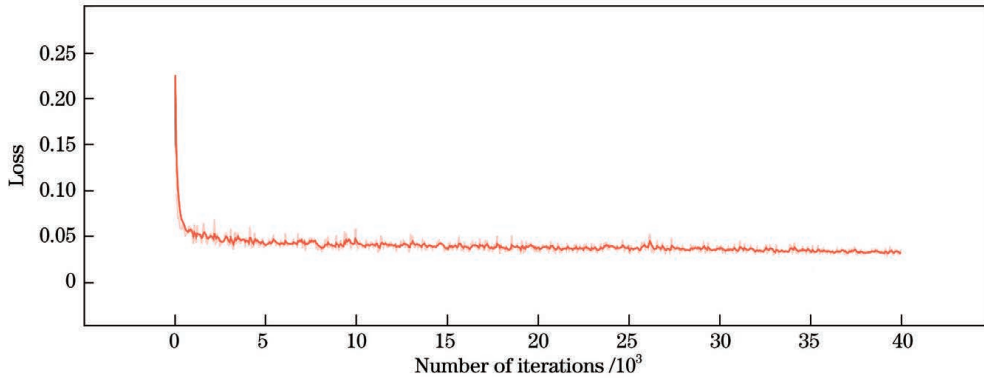


图 7 训练过程中损失曲线

Fig. 7 Loss curve during training

4.2 评估标准

针对桥梁裂缝检测结果, 使用总体像素准确率 (PA)、平均交并比 (mIoU) 和 F1_score 进行量化评价。

总体像素准确率和平均交并比计算公式分别为

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

$$P_{mIoU} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}, \quad (5)$$

式中: N_{TP} 表示预测为正类别, 标签也为正类别的像素; N_{TN} 表示预测为负类别, 标签也为负类别的像素; N_{FP} 表示预测为正类别而标签为负类别的像素; N_{FN} 表示预测为负类别但标签是正类别的像素。F1_score 可以看作是模型精确率 (precision) 和召回率 (recall) 的一种加权平均, 兼顾了精确率的查全和召回率的查准作用, 计算公式为

$$P_{precision} = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (6)$$

$$R_{recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (7)$$

$$F_{F1_score} = 2 \times \frac{P_{precision} \times R_{recall}}{P_{precision} + R_{recall}}. \quad (8)$$

4.3 对比实验

为了验证所提算法的准确性, 设置了两组对比实验。第一组, 验证 SPAM 的有效性; 第二组, 对所提算法与现有的一些主流语义分割算法进行对比, 不仅对比了各个算法的检测准确性, 还进行了时间效率方面的对比。

第一组实验, 对 PSPNet 和带有 SPAM 的 PSPNet 进行对比, 结果如表 1 所示, SPAM 显著提升了总体像素准确率和平均交并比。与 PSPNet 相比, 将 SPAM 与金字塔池化模块串联的网络的平均交并比达到了 84.31%, 提高了 7.97 个百分点。时

表 1 有无 SPAM 的实验结果对比

Table 1 Comparison of experiment results without and with SPAM

SPAM	Precision / %	Recall / %	PA / %	F1_score / %	mIoU / %	Time / s
-	86.14	82.25	98.28	84.15	76.34	0.46
✓	95.34	90.27	99.37	92.74	84.31	0.60

间效率方面,在同样的训练集下设置批处理大小为 16,原 PSPNet 迭代一次需要 0.46 s,所提算法迭代一次需要 0.60 s,虽然引入了 SPAM 导致算法的计

算复杂度提高,但是所耗时间并没有大幅提高。

如图 8 所示,带有 SPAM 的网络使结果图中的裂缝细节和边界更加清晰,细小裂缝分割结果更好。

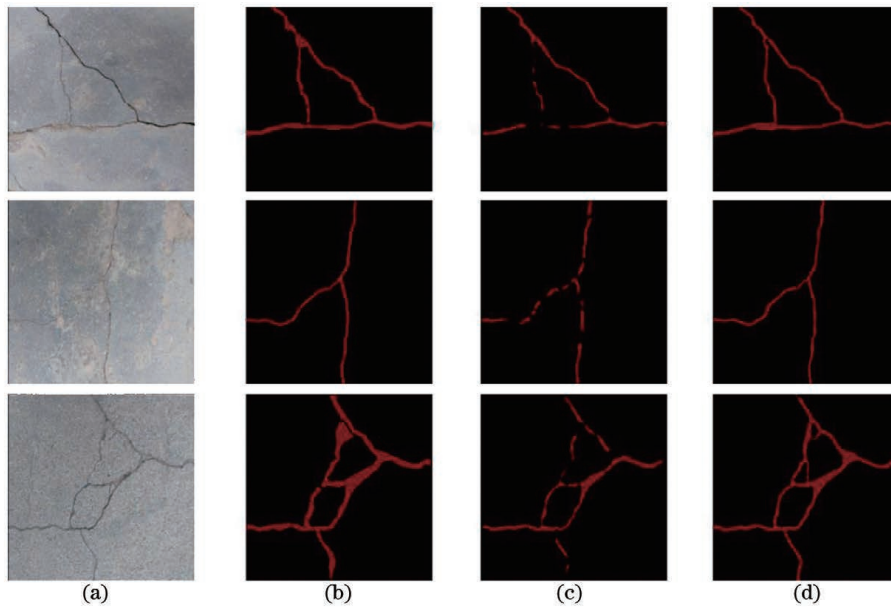


图 8 有无 SPAM 对比结果图。(a)原图;(b)标签;(c)无 SPAM;(d)有 SPAM

Fig. 8 Comparison results without and with SPAM. (a) Original image; (b) label; (c) without SPAM; (d) with SPAM

第二组对比实验,使用了 Deeplab-V3+[22]、U-Net++[23]、EncNet[24] 模型进行了桥梁裂缝检测。

图 9 给出了部分裂缝图像检测结果。可以看到,对于裂缝宽度较大、噪声干扰少的图像,三种模型均能

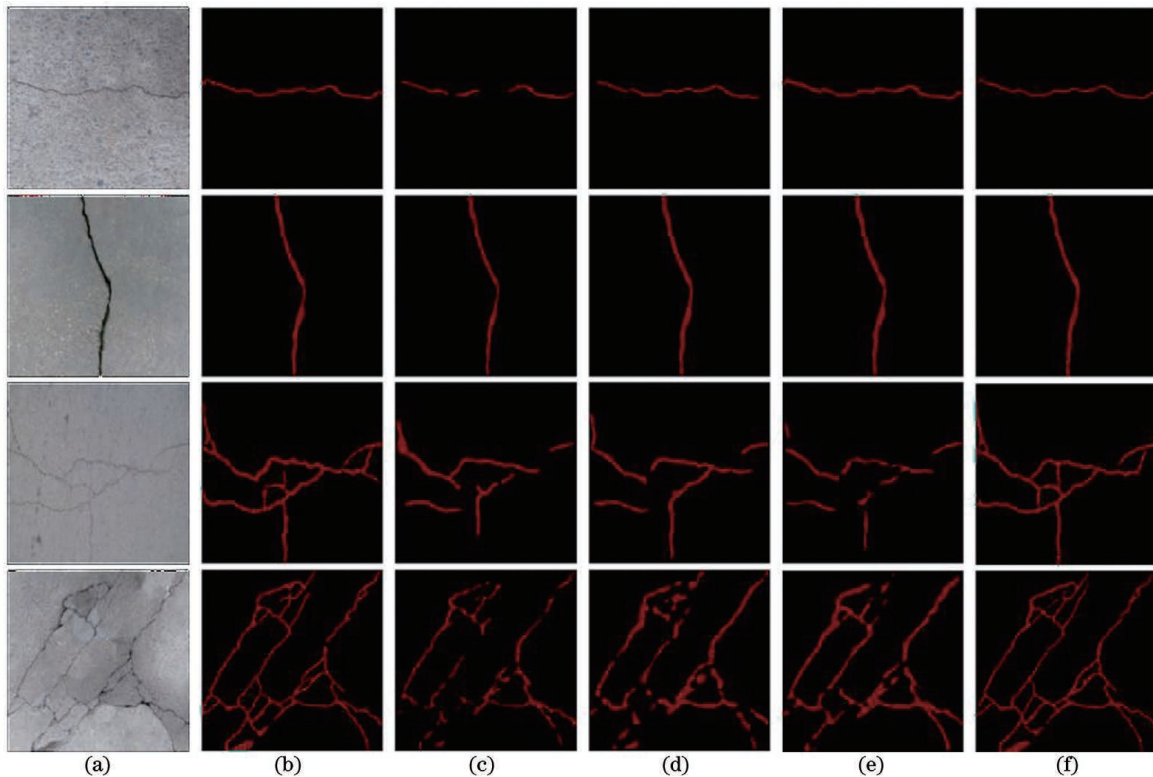


图 9 不同算法的结果对比。(a)原图;(b)标签;(c)Deeplab-V3+;(d)U-Net++;(e)EncNet;(f)所提算法

Fig. 9 Comparison of results between different algorithms. (a) Original image; (b) label; (c) Deeplab-V3+; (d) U-Net++; (e) EncNet; (f) proposed algorithm

够将裂缝信息完整检测出来,表现效果较好;但是当图像存在细小裂缝或者图像裂缝较为复杂时,三种模型均会出现误检,检测结果不连续等问题。图 9 第 1 行,图像中存在一根细小裂缝,可以看到 Deeplab-V3+ 会出现漏检现象。第 3 行,图像存在细小裂缝并且交叉时,各检测模型都出现了不同程度的漏检,Deeplab-V3+ 和 U-Net++ 漏检较为严重,而 EncNet 和所提算法则能够更多地体现裂缝细节。针对第 4 行细小复杂的网状裂缝,Deeplab-V3+、U-Net++、EncNet 模型表现效果较差,检测的裂缝图像不连续,其中 Deeplab-V3+,U-Net++

漏检较为严重。整体而言,所提算法检测结果更准确,细节表现更好,能够将裂缝准确完整检测出来。

继续使用评价指标对分割结果进行更为直观的评估,评估结果如表 2 所示。结果表明,对于不同类型的桥梁裂缝,所提算法在总体像素准确率、平均交并比、F1_score 指标上明显高于其他几种主流的语义分割模型。在时间效率方面,在同样的训练集下设置批处理大小为 16,Deeplab-V3+ 迭代一次需要 0.52 s, U-Net++ 需要 0.63 s, EncNet 需要 0.42 s,所提算法需要 0.60 s,进一步从数据角度论证了所提算法对桥梁裂缝检测的有效性。

表 2 所提算法与主流语义分割模型对比

Table 2 Comparison between the proposed algorithm and mainstream semantic segmentation models

Algorithm	Precision /%	Recall /%	PA /%	F1_score /%	mIoU /%	Time /s
U-Net++	84.52	80.13	97.14	82.27	75.18	0.63
Deeplab-V3+	84.76	81.42	98.93	83.06	77.95	0.52
EncNet	91.38	83.79	99.06	87.42	80.75	0.42
Proposed algorithm	95.34	90.27	99.37	92.74	84.31	0.60

5 结 论

研究了图像语义分割在桥梁表面裂缝分割提取中的应用。针对 PSPNet 网络在桥梁表面裂缝分割提取任务上存在的细小裂缝检测不完全、图像背景比例较大等问题,在 PSPNet 网络基础上,引入了一种基于自注意力机制的空间位置模块,将其与金字塔池化模块串联起来,共同捕捉上下文信息。实验结果表明,与原始 PSPNet 和其他主流的语义分割模型相比,所提算法在细小裂缝和复杂裂缝等方面取得了良好的分割结果,能够更好地捕捉到裂缝细节信息。

参 考 文 献

- [1] Zhou X H, Zhang X G. Thoughts on the development of bridge technology in China[J]. Engineering, 2019, 5(6): 1120-1130, 1245-1256. 周绪红, 张喜刚. 关于中国桥梁技术发展的思考[J]. 工程(英文), 2019, 5(6): 1120-1130, 1245-1256.
- [2] Li W J, Zhang M, Shen Z H, et al. Track crack detection method in complex environment[C]//2018 11th International Symposium on Computational Intelligence and Design (ISCID), December 8-9, 2018, Hangzhou, China. New York: IEEE Press, 2018: 356-359.
- [3] Qu Z, Lin L D, Guo Y, et al. An improved algorithm for image crack detection based on percolation model[J]. IEEE Transactions on Electrical and Electronic Engineering, 2015, 10(2): 214-221.
- [4] Amhaz R, Chambon S, Idier J, et al. Automatic crack detection on two-dimensional pavement images: an algorithm based on minimal path selection [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10): 2718-2729.
- [5] Xing Y F, Zhong L, Zhong X. An encoder-decoder network based FCN architecture for semantic segmentation[J]. Wireless Communications and Mobile Computing, 2020, 2020: 1-9.
- [6] Mou L C, Hua Y S, Zhu X X. Relation matters: relational context-aware fully convolutional network for semantic segmentation of high-resolution aerial images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(11): 7557-7569.
- [7] Li L F, Hu M. Method for small-bridge-crack segmentation based on generative adversarial network [J]. Laser & Optoelectronics Progress, 2019, 56(10): 101004. 李良福, 胡敏. 基于生成式对抗网络的细小桥梁裂缝分割方法[J]. 激光与光电子学进展, 2019, 56(10): 101004.
- [8] Cheng J, Ye L, Guo Y N, et al. Ground crack recognition based on fully convolutional network with multi-scale input[J]. IEEE Access, 2020, 8: 53034-53048.

- [9] Zou Q, Zhang Z, Li Q Q, et al. DeepCrack: learning hierarchical convolutional features for crack detection [J]. *IEEE Transactions on Image Processing*, 2019, 28(3): 1498-1512.
- [10] Zhang C, Chen Y. Object detection based on hard examples mining using residual network[J]. *Laser & Optoelectronics Progress*, 2018, 55(10): 101003.
张超, 陈莹. 残差网络下基于困难样本挖掘的目标检测[J]. *激光与光电子学进展*, 2018, 55(10): 101003.
- [11] Xia K J, Yin H S, Qian P J, et al. Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images[J]. *IEEE Access*, 2019, 7: 96349-96358.
- [12] Li G, Gao Z Y, Zhang X C, et al. Improved global convolutional network for pavement crack detection [J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 081011.
李刚, 高振阳, 张新春, 等. 改进的全局卷积网络在路面裂缝检测中的应用[J]. *激光与光电子学进展*, 2020, 57(8): 081011.
- [13] Sun M Y, Guo R H, Zhu J H, et al. Roadway crack segmentation based on an encoder-decoder deep network with multi-scale convolutional blocks [C] // 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), January 6-8, 2020, Las Vegas, NV, USA. New York: IEEE Press, 2020: 0869-0874.
- [14] Kang D, Benipal S S, Gopal D L, et al. Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning [J]. *Automation in Construction*, 2020, 118: 103291.
- [15] Li L F, Ma W F, Li L, et al. Research on detection algorithm for bridge cracks based on deep learning [J]. *Acta Automatica Sinica*, 2019, 45(9): 1727-1742.
李良福, 马卫飞, 李丽, 等. 基于深度学习的桥梁裂缝检测算法研究[J]. *自动化学报*, 2019, 45(9): 1727-1742.
- [16] Zhang L, Yang F, Daniel Z Y, et al. Road crack detection using deep convolutional neural network [C] // 2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 3708-3712.
- [17] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 3431-3440.
- [18] 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 (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6230-6239.
- [19] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [20] Yao H B, Jiang J G, Qi M B, et al. Image deblurring algorithm using Laplacian prior and bilateral filtering approach [J]. *Transducer and Microsystem Technologies*, 2017, 36(1): 139-142.
姚海波, 蒋建国, 齐美彬, 等. 拉普拉斯与双边滤波的图像去运动模糊算法 [J]. *传感器与微系统*, 2017, 36(1): 139-142.
- [21] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3141-3149.
- [22] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [M] // Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 833-851.
- [23] Zhou Z W, Siddiquee M M R, Tajbakhsh N, et al. UNet++: a nested U-net architecture for medical image segmentation [M] // Stoyanov D, Taylor Z, Carneiro G, et al. *Deep learning in medical image analysis and multimodal learning for clinical decision support. Lecture notes in computer science*. Cham: Springer, 2018, 11045: 3-11.
- [24] Zhang H, Dana K, Shi J P, et al. Context encoding for semantic segmentation [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake, UT, USA. New York: IEEE Press, 2018: 7151-7160.