

基于注意力机制和轻量级空洞卷积的混凝土路面裂缝检测

瞿中, 王彩云

引用本文

瞿中, 王彩云. [基于注意力机制和轻量级空洞卷积的混凝土路面裂缝检测](#)[J]. 计算机科学, 2023, 50(2): 231-236.

QU Zhong, WANG Caiyun. [Crack Detection of Concrete Pavement Based on Attention Mechanism and Lightweight Dilated Convolution](#) [J]. Computer Science, 2023, 50(2): 231-236.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于会话式机器阅读理解模型的事件抽取方法](#)

Event Extraction Method Based on Conversational Machine Reading Comprehension Model
计算机科学, 2023, 50(2): 275-284. <https://doi.org/10.11896/jsjcx.220400271>

[基于数据增强的自监督飞行航迹预测](#)

Self-supervised Flight Trajectory Prediction Based on Data Augmentation
计算机科学, 2023, 50(2): 130-137. <https://doi.org/10.11896/jsjcx.211200016>

[一种基于多模态深度特征融合的视觉问答模型](#)

Visual Question Answering Model Based on Multi-modal Deep Feature Fusion
计算机科学, 2023, 50(2): 123-129. <https://doi.org/10.11896/jsjcx.211200303>

[基于双向注意力机制和门控图卷积网络的文本分类方法](#)

Text Classification Method Based on Bidirectional Attention and Gated Graph Convolutional Networks
计算机科学, 2023, 50(1): 221-228. <https://doi.org/10.11896/jsjcx.211100095>

[残差注意力与多特征融合的图像去模糊](#)

Image Deblurring Based on Residual Attention and Multi-feature Fusion
计算机科学, 2023, 50(1): 147-155. <https://doi.org/10.11896/jsjcx.211100161>

基于注意力机制和轻量级空洞卷积的混凝土路面裂缝检测

瞿中 王彩云

重庆邮电大学计算机科学与技术学院 重庆 400065

摘要 混凝土路面上的裂缝会影响结构的安全性、适用性和耐久性,裂缝检测是一个充满挑战的研究热点。文中提出了由改进的全卷积网络和深监督网络组成的裂缝检测模型,以改进的 VGG-16 作为主干网络,首先将低层卷积特征聚合,通过空间注意力机制再次融合到主干网络;其次,将中高层卷积特征通过轻量级空洞卷积融合模块进行多尺度融合得到具有清晰边缘且分辨率较高的特征图像,所有的侧边特征图像相加产生最终的预测图像;最后,深监督网络为每个阶段的检测结果提供直接监督。该网络选择焦点损失函数作为评价函数,经过训练的网络模型能够在光照不均、背景复杂等各种条件下从输入的原始图像中高效地识别出裂缝位置。为验证所提方法的有效性和鲁棒性,在 DeepCrack,CFD,Crack500 这 3 个数据集上与 6 种方法进行了比较,所提算法表现出卓越的性能,F-score 值达到了 87.12%。

关键词: 裂缝检测;注意力机制;空洞卷积;多尺度融合;全卷积;深度监督网络

中图分类号 TP391.41

Crack Detection of Concrete Pavement Based on Attention Mechanism and Lightweight Dilated Convolution

QU Zhong and WANG Caiyun

College of Computer Science and Technology,Chongqing University of Posts and Telecommunications,Chongqing 400065,China

Abstract Cracks in the concrete pavement will affect the safety,applicability,and durability of the structure,and crack detection is a challenging research hotspot. This paper proposes a crack detection model composed of an improved full convolutional network and a deep supervision network,which uses the improved VGG-16 as the backbone network. Firstly,the low-level convolutional feature aggregation is fused to the backbone network again through the spatial attention mechanism. Secondly,the middle and high-level convolutional features are fused through the lightweight dilated convolution fusion module for multi-feature fusion to get the clear edge and high-resolution feature maps,all side feature maps are added to produce the final prediction map. Finally,the deep supervision network provides direct supervision for the detection results of each stage. In this paper,the focus loss function is selected as the evaluation function,and the trained network model can efficiently identify the crack location from the input original image under various conditions such as uneven illumination and complex background. To verify the effectiveness and robustness of the proposed method,it is compared with six methods on three datasets,DeepCrack,CFD,and Crack500,and the results show that it has excellent performance,and the F-score value reaches 87.12%.

Keywords Crack detection, Attention mechanism,Dilated convolution, Multiscale fusion, Fully convolutional network, Deep supervision network

1 引言

裂缝是公路病害中最常见的一种缺陷,它们使腐蚀物到达路面内部的钢筋层,加速了腐蚀,使得公路的抗拉强度降低,不仅影响美观,更重要的是,情况严重时会影响行车安全,因此,定期检查道路裂缝的严重程度极其重要。早期裂缝检测采用手工检测,不仅成本高、耗时长、给作业人员带来安全隐患,其检测结果还依赖于不同人员的主观检测标准。随后,

不同的检测算法不断被应用到实践中,但受光照不均、油渍、水渍、阴影、道路背景纹理等大量干扰因素的影响,导致在检测过程中大量裂缝出现断裂、对比度低、误检等问题。深度学习快速发展,其在裂缝检测研究领域得到了广泛的应用。

本文提出了一种基于注意力机制和轻量级空洞卷积的混凝土路面裂缝检测网络模型。该网络模型以改进的 VGG-16 网络作为主干网络实现了裂缝特征的快速和精确提取,通过在低卷积层加入空间注意力机制,提高了特征提取能力。将

到稿日期:2021-12-27 返修日期:2022-07-02

基金项目:国家自然科学基金(62176034)

This work was supported by the National Natural Science Foundation of China(62176034).

通信作者:瞿中(quzhong@cqupt.edu.cn)

不同阶段的特征用并行的不同空洞率的空洞卷积融合,提升高层特征的分辨率;使用多尺度特征融合,提高检测的准确率。

2 相关工作

2.1 裂缝检测方法

裂缝检测从传统方法的检测发展到了深度学习检测,无论是检测精度还是鲁棒性都有了很大的提升。Cheng等^[1]采用了基于阈值的算法,该算法有一个前提条件是裂缝像素总是比背景像素更暗。Nejad等^[2]提出了基于Radon小波变换和动态神经网络阈值选择的自动裂缝检测系统。Amhaz等^[3]考虑到路面图像的光度和几何特征,提出了基于最小路径的裂缝检测方法。Yamaguchi等^[4]提出了基于渗流的快速检测裂缝的方法。

随着深度学习在图像分类、目标检测、语义分割等各个领域的广泛应用,卷积神经网络强大的特征提取和泛化能力在裂缝检测中也被广泛采用。Fan等^[5]以编码-解码架构为基础增添分层特征学习和不同空洞率的空洞卷积,实现了端到端的自动裂缝检测。Qu等^[6]提出了基于DeepLab的密集特征提取器,同时结合了多尺度特征融合和深监督网络的裂缝检测网络。Xu等^[7]基于改进U-Net网络建立HU-ResNet模型,引入scSE注意力机制模块在空间和通道重新标定输出特征。虽然这些网络模型相比传统算法在误检、检测速度、泛化能力上有了很大的提高,但是裂缝检测的精度还有待进一步提升,裂缝边缘还需要进一步精细,抗干扰能力还要进一步加强。

2.2 注意力机制

人们可以在复杂的图像中找到突出区域。在计算机视觉中,注意力机制可以模仿人类视觉系统,其被视为一个动态选择的过程,根据输入的重要性自适应地调整权重。注意力机制大致可以分为4种:通道注意力机制、空间注意力机制、时间注意力机制和分支通道注意力机制。Hu等^[8]对通道之间的依赖关系进行建模,来自适应地调整通道的特征响应值,首次提出了通道注意力机制,相应的网络模型命名为SENet。

Woo等^[9]沿着通道和空间两个维度依次推断出注意力系数,然后将注意力与输入特征相乘实现自适应特征细化,提出了轻量级通用模块CBAM。时间注意力机制用于决定何时注意动态时间抉择,通常用于视频处理。为了解决轻量级卷积神经网络的性能下降问题,Chen等^[10]提出了动态卷积,其核心思想是根据输入相关注意力动态聚合多个并行卷积核,组装多个内核不仅计算效率高而且具有更强的表现能力。

3 基于注意力机制和空洞卷积的混凝土路面裂缝检测

本文将裂缝检测视为像素级的二分类问题,用“0”和“1”分别表示非裂缝和裂缝。裂缝检测算法分为训练模型和测试模型两个阶段。在训练阶段,以损失函数为指引,寻找最优的权值参数,并将网络模型保存下来;在测试阶段,使用训练好的模型检测不同数据集,验证模型的有效性和鲁棒性。

3.1 模型结构

考虑到低层特征和高层特征各具特点以及感受野对特征提取的影响,本文提出的网络整体模型结构如图1所示,大致分为四大部分:1)改进了VGG-16网络并将其作为主干网络用于特征提取;2)低层嵌入的空间注意力机制用于为后续卷积提供多细节特征的高分辨率输入图像;3)高层采用并行轻量级空洞卷积,该模块能有效地融合多尺度特征;4)聚合所有特征图产生最终预测图的输出层。主干网络中VGG-16的前13层卷积深度较浅,Conv2阶段的侧边输出的裂缝图像有较清晰的裂缝边缘及细节信息,但同时也含有较多噪声。为减少噪声,从而为后续特征融合提供高质量裂缝图像,在Conv2阶段增加了一个 3×3 的卷积。丢弃全连接层和第五个池化层,因为第五个池化后产生的特征图分辨率过低,对产生精细的预测图像贡献较小。每一个卷积由3部分组成:卷积层、批归一化、激活层Relu。每一个卷积阶段都有一个侧边裂缝预测图像,都会根据损失函数计算出损失值,为形成最终的预测图像提供指导,构成深监督学习,如图2所示。所有的侧边特征图像通过一个 1×1 的卷积减少通道数,然后相加得到最终的预测图像。

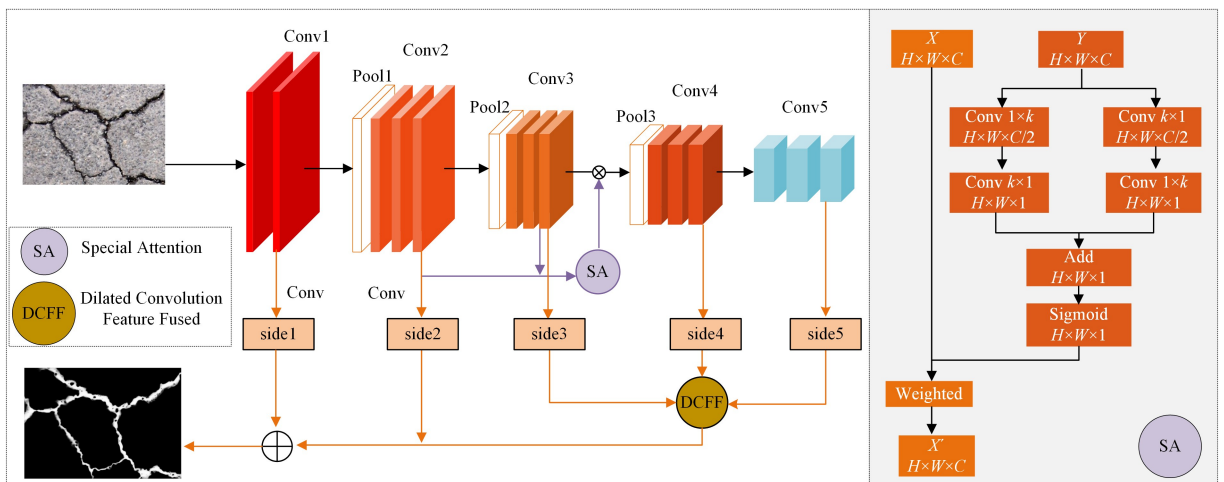


图1 网络模型结构图

Fig. 1 Architecture of the proposed network

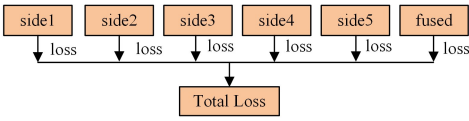


图 2 深监督网络结构图

Fig. 2 Architecture of deep supervised network

3.2 注意力机制

随着卷积层的不断加深,裂缝在特征提取过程中的边缘信息和细节信息会不断丢失。从图 3 可以看出,特征图 side4 分辨率较低,为了使深层卷积仍然有丰富的细节信息,将有清晰边缘信息的噪声较少的浅层特征图并入到主干网络的想法应运而生。图像通常包含关注的前景细节和复杂的背景信息,在低层特征中裂缝区域应该被重点关注,使用空间注意力机制可以为有效特征区域带来更大的关注度。

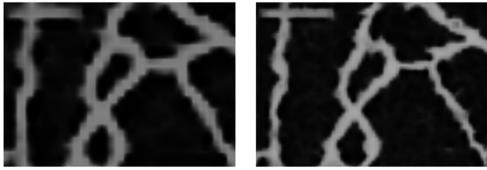


图 3 特征图 side 4 经 SA 模块处理的对比图

Fig. 3 Contrast images of side 4 after SA processing

本文使用的空间注意力机制^[11]原理如图 1 右侧框中的 SA 所示。首先为了增大感受野并捕获全局信息, $1 \times k$ 和 $k \times 1$ 的这两种卷积被使用,这样可以提取被关注区域的特征信息。随后将提取的特征相加,利用 Sigmoid 运算将结果映射到 $[0, 1]$ 的范围,得到相应的权重。将经过空间注意力机制处理过的特征图与主干网络融合,为后续卷积层提供更高分辨率且噪音较少的输入。

3.3 并联的轻量级空洞卷积

从全局视图来看,裂缝通常呈长且不规则的曲状,拓扑结构复杂。传统卷积由于感受野的限制,没有完美契合裂缝的结构特征,而空洞卷积通过不同空洞率改变感受野的大小,提取了更多细节信息。因此,本文提出了并行的轻量级空洞卷积(Dilated Convolution Feature Fusion, DCFE),如图 4 所示。提出的模块有两大特点:

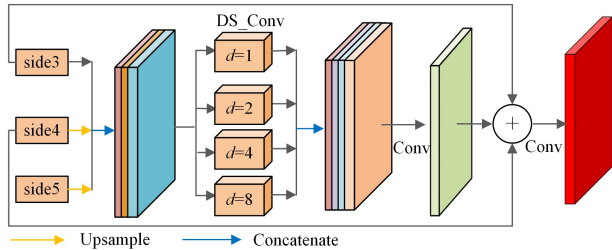


图 4 并行轻量级空洞卷积融合模块展开结构

Fig. 4 Unrolled structure of parallel lightweight dilated convolution fusion module

(1)同时将高层特征和低层特征输入到 DCFE 模块中获取具有细节信息的高分辨率特征图。图 1 中特征图 side 4 和 side 5 作为高层特征具有大量的语义信息,但是其结构信息丢失,特征图 side 2 分辨率高且边缘信息精确,有别于

ASPP^[12] 只对最后的特征图像做处理,该模块集成了多侧边特征图像的有价值的信息,侧边特征被渐进式地多次利用。

(2)普通的空洞卷积会增加模型参数量,降低模型训练速度,为了在提升性能的同时不忽略时间成本,采用不同空洞率的深度可分离卷^[13]构成轻量级空洞卷积模块,不仅没有消耗更多的时间和内存,还提升了性能。

3.4 损失函数

深监督网络的核心是在每个卷积阶段产生预测图,并根据损失函数计算损失值,而不是只在输出层进行监督。深监督网络达到了对卷积中间层和最终结果综合监督的目的。本文使用的损失函数由两部分构成:侧边网络的损失函数和融合层的损失函数。

侧边网络主要由 5 个侧边特征图构成,每个侧边特征图都计算损失值,最后加权求和得到侧边的总损失值,计算式如下:

$$L_{side}(X, Y, W, \tau) = \sum_{m=1}^M \partial_m l_{side}^{(m)}(X, Y, W, \tau^{(m)}) \quad (1)$$

由于每个侧边输出特征图对最终预测图的贡献不尽相同, ∂_m 为每个侧边输出特征图的权重^[14]。 X 和 Y 分别对应原图像和标签图像的总像素, W 表示网络模型的总参数,每一个侧边输出层都可以视为带权 $\tau^{(m)}$ 的分类器, m 表示第 m 个侧边层。 $l_{side}^{(m)}$ 代表第 m 个侧边层的损失函数,其表达式如下:

$$l_{side}^{(m)}(X, Y, W, \tau^{(m)}) = - \sum_{i \in Y_+} \tau_{i_0} \log P_i(y_i = 1 | X, W, \tau^{(m)}) - \sum_{i \in Y_-} \tau_{i_1} \log P_i(y_i = 0 | X, W, \tau^{(m)}) \quad (2)$$

裂缝图像中裂缝像素和背景像素的比例相差甚远,正负样本不平衡,成为了训练高性能模型的障碍。为了解决这一问题,本文通过给正负分配权重来降低样本不平衡的影响, τ_{i_0} 和 τ_{i_1} 即为裂缝像素和非裂缝像素的权重,标签图像中裂缝与非裂缝像素的总数表示为 Y_+ 和 Y_- 。

融合层的损失计算方式 (L_{fuse}) 和侧边输出层的损失函数 (见式(2)) 相同,最终的损失函数为:

$$L = L_{side}(X, Y, W, \tau) + L_{fuse}(X, Y, W) \quad (3)$$

4 实验结果分析

4.1 实验平台及参数设置

本文提出的网络模型及所有的对比方法均使用深度学习的框架 PyTorch 编程实现,在模型训练阶段,用随机梯度下降(SGD)优化模型,用“Xavier”方法初始化卷积层参数,偏差设置为 0,初始化全局学习率为 1×10^{-4} ,动量为 0.9, epoch 设置为 500,每训练 50 个 epoch 就会保存一次模型参数。本文提出的网络模型及所有对比方法的训练和测试均在 Ubuntu16.04 的操作系统、多个高性能的 Tesla-V100-SXM2-32 GB GPU 集群,以及充足的 CPU 资源的平台上进行。

4.2 数据集

(1)DeepCrack:该数据集来源于文献^[14],被称为 DeepCrack,包含了 537 幅不同类型的裂缝图像,部分图像中含有较多的噪声,如水坑、斑点、污渍、遮挡物、阴影。这些图像主要分为两组:一组包含 300 幅图像的训练集和另一组包含

237 幅图像的测试集。所有图像的分辨率固定为 544×384 。

(2)CFD: Shi 等^[15]提出了 CFD 带标签图像的路面裂缝数据集,包含 118 幅分辨率为 480×320 的图像。所有图像都有手工标记的标签图像,可直接用于指标评估。这些图像中有车道线和光影的噪声,并且大多数裂缝都是具有检测难度的细裂缝。

(3)Crack500: Yang 等^[16]共享了带有标签的数据集 Crack500,包含 3792 幅训练图像、697 幅测试图像。不同场景的裂缝图像展现了横向、纵向、龟裂等各式裂缝,有着不同程度的检测难度。

4.3 对比实验结果分析

将本文提出的方法与基于深度学习的裂缝检测方法进行了比较,这些裂缝检测方法本文提出的方法使用相同的网络参数在 DeepCrack 上进行训练,并在 3 个数据集上进行测试。

(1)FPHBN^[16]: FPHBN 是在 HED 网络上的提升,它应用了特征金字塔和层次增强网络。本文使用 PyTorch 框架,所有的参数设置与文献^[16]网络模型相同。

(2)DeepCrack^[7]: DeepCrack 建立在 SegNet 的编码器-解码器架构上,在编码和解码的阶段相同尺度的特征会成对融合。

(3)DeepCrack^[14]: DeepCrack 网络的主干网路是 VGG-16 的前 13 层,每一个卷积层次都会产生一个侧边输出,所有的侧边相加可得到最终的预测图像。

(4)CrackSegNet^[17]: CrackSegNet 是一种由主干网络、扩展卷积、空间金字塔池和跳过连接模块组成的改进的深度卷积神经网络模型。

(5)U-HDN^[5]: U-HDN 在编码-解码体系结构中嵌入了多个扩展模块,并设计了分层特征学习来综合特征信息。

(6)CENet^[18]: Gu 等基于 UNet 提出的 CENet 网络。主要由特征编码器、上下文提取器和特征解码器 3 部分组成。CENet 用于医学图像处理,本文经过少量修改之后将其用于裂缝检测。

4.4 实验评价指标

在实验对比中需要评估不同算法的性能,本文使用了 4 个常见的语义分割评估指标:精度 Precision(P)、召回率 Recall(R)、F 值 F-score(F)、平均交并比(MIoU)。其中 F 值作为性能评估的总体指标。根据标记情况和实际预测情况,将像素分为 4 种类型:真阳性(TP)、假阳性(FP)、真阴性(TN)、

假阴性(FN)。P, R, F-score 的定义如式(1)~式(3)所示。

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F-score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

用 n_{ij} 表示类别 i 中被预测为类别 j 的总像素数, n_{cls} 表示像素类别的总类别数, $t_i = \sum_j n_{ij}$ 代表类别 i 的总像素数,这里的类别 i 包括真阳性和假阳性,平均交并比(MIoU)表示如下:

$$MIoU = \frac{1}{n_{cls}} \cdot \frac{\sum_i n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}} \quad (7)$$

4.5 实验结果与分析

本文将 FPHBN, DeepCrack^[7], DeepCrack^[14], CrackSegNet, U-HDN, CENet 和本文提出的网络在 DeepCrack, Crack500, CFD 数据集上进行对比,验证所提网络的有效性。

(1)各方法在 DeepCrack 数据集上的检测结果分析:在图 5 所示的 P-R 曲线中,本文提出的方法是所有方法中最优的。由图 6 可以看出,本文提出的方法在 DeepCrack 数据集上的检测效果最佳,其检测的裂缝边缘清晰,噪声较少,对复杂拓扑结构的裂缝检测最完整,且能辨别出障碍物。图 7(a)给出了本文算法在该数据集上的更多检测结果。从表 1 可以看出,本文提出的方法的 P, R, F-score 值,以及 MIoU 都是最高的,其中 MIoU 为 88.02%,其 F-score 比 U-HDN, CrackSegNet, FPHBN, DeepCrack^[7], DeepCrack^[14], CENet 的 F-score 值分别提高了 2.39%, 4.59%, 6.26%, 1.74%, 1.35% 和 7.84%。本文算法的参数总量为 25.8×10^6 , 由于使用了多个模块,因此相应的参数量相比其他算法有所增加。

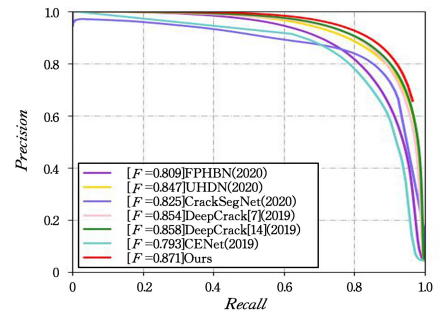
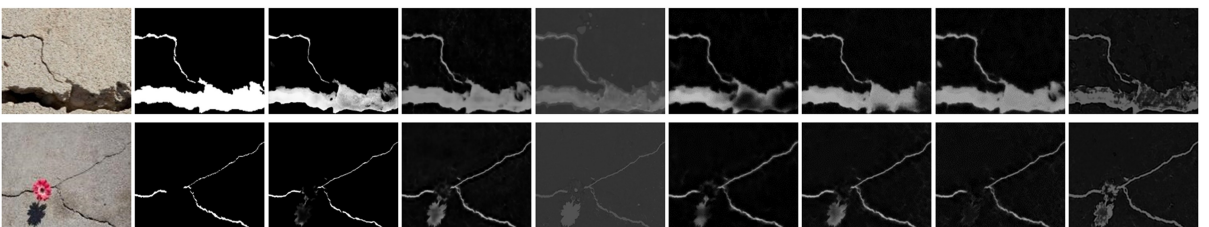


图 5 在 DeepCrack 数据集上的 P-R 曲线

Fig. 5 P-R curves on DeepCrack dataset



(a)原图 (b)标签图像 (c)Ours (d)U-HDN (e)CrackSegNet (f)FPHBN (g)DeepCrack^[7] (h)DeepCrack^[14] (i)CENet

图 6 不同方法的检测结果对比

Fig. 6 Comparison of detection results of different methods

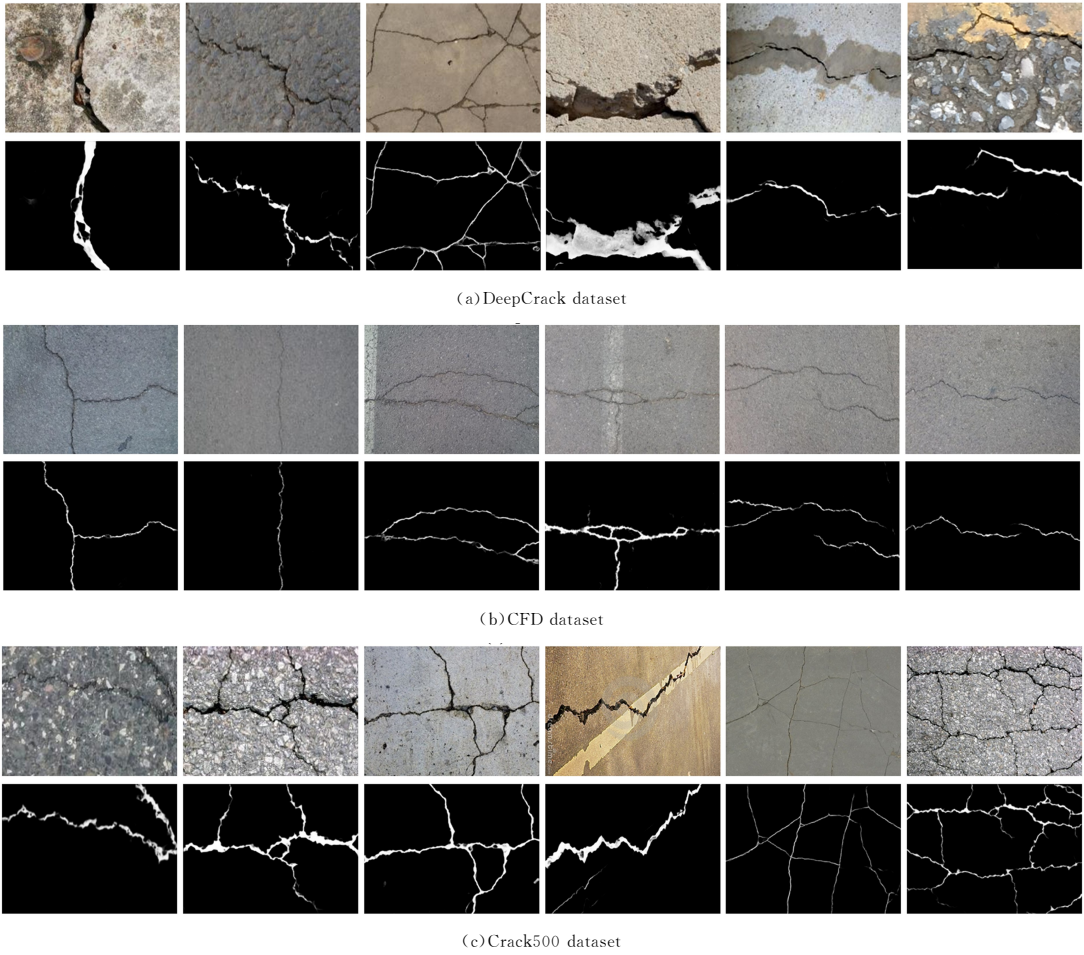


图 7 DeepCrack,CFD,Crack500 数据集上的检测结果

Fig. 7 Detection results on DeepCrack,CFD and Crack500 datasets

表 1 在 DeepCrack 上的对比实验结果

Table 1 Comparative experimental results on DeepCrack

Methods	<i>P</i>	<i>R</i>	<i>F-score</i>	<i>MIoU</i>	<i>Params</i>
U-HDN ^[5]	0.8429	0.8518	0.8473	0.8606	34.5×10^6
CrackSegNet ^[17]	0.8128	0.8383	0.8253	0.8433	14.9×10^6
FPHBN ^[16]	0.8293	0.7888	0.8086	0.8446	15.5×10^6
DeepCrack ^[7]	0.8642	0.8437	0.8538	0.8666	29.4×10^6
DeepCrack ^[14]	0.8608	0.8537	0.8577	0.8693	14.7×10^6
CENet ^[18]	0.8135	0.7731	0.7928	0.8103	29.0×10^6
Ours	0.8807	0.8619	0.8712	0.8802	25.8×10^6

(2)各方法在 CFD 数据集上的检测结果分析:从表 2 和图 8 可以看出,本文提出的方法在 CFD 数据集上的 *MIoU* 是最高的,*F-score* 值为 0.7040,比最低值高出 7.83%。从整体数据可看出,所有方法在 CFD 上的检测性能都不太理想,这是因为 CFD 数据集中有大量的细裂缝图像,增加了检测难度,细裂缝是所有方法都面临的挑战。

表 2 在 CFD 上的对比实验结果

Table 2 Comparative experimental results on CFD

Methods	<i>P</i>	<i>R</i>	<i>F-score</i>	<i>MIoU</i>
U-HDN ^[5]	0.6395	0.7755	0.7010	0.7651
CrackSegNet ^[17]	0.6369	0.6148	0.6257	0.7224
FPHBN ^[16]	0.5987	0.7528	0.6670	0.7080
DeepCrack ^[7]	0.5967	0.7327	0.6577	0.7394
DeepCrack ^[14]	0.6244	0.7341	0.6581	0.7494
CENet ^[18]	0.6191	0.7347	0.6720	0.7479
Ours	0.6569	0.7584	0.7040	0.7671

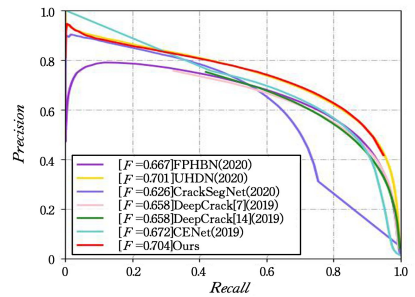


图 8 在 CFD 数据集上的 P-R 曲线

Fig. 8 P-R curves on CFD dataset

(3)各方法在 Crack500 数据集上的检测结果分析。

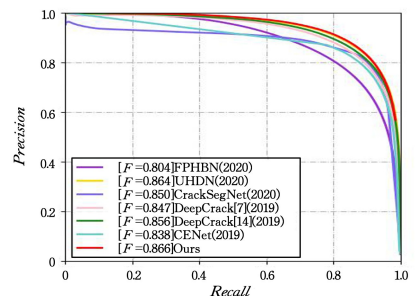


图 9 在 Crack500 数据集上的 P-R 曲线

Fig. 9 P-R curves on Crack500 dataset

从图 9 中的 *P-R* 曲线图可以看出,本文提出的方法位于

坐标最右上方,在表 3 中,本文提出的方法 F -score 值和 $MIoU$ 也是最高的,分别为 0.8659 和 0.8778。与 U-HDN, CrackSegNet, FPHBN, DeepCrack^[7], DeepCrack^[14] 和 CENet 相比, F -score 值分别提高了 0.2%, 1.57%, 6.22%, 1.86%, 1.04% 和 2.79%。观察图 7(c)可知,对于龟裂的、结构复杂、受车道线干扰的裂缝,本文算法的检测结果接近原裂缝,精度较高。

表 3 在 Crack500 上的对比实验结果

Table 3 Comparative experimental results on Crack500

Methods	P	R	F -score	$MIoU$
U-HDN ^[5]	0.8623	0.8655	0.8639	0.8761
CrackSegNet ^[17]	0.8245	0.8776	0.8502	0.8651
FPHBN ^[16]	0.7918	0.8158	0.8037	0.8306
DeepCrack ^[7]	0.8356	0.8593	0.8473	0.8629
DeepCrack ^[14]	0.8407	0.8708	0.8555	0.8694
CENet ^[18]	0.8244	0.8521	0.8380	0.8557
Ours	0.8619	0.8701	0.8659	0.8778

4.6 消融实验

为了证明本文提出的方法中的各个模块对特征提取性能都有一定的提升,本文在 DeepCrack 数据集上进行了消融实验,结果如表 4 所列。

表 4 在 DeepCrack 上的消融实验结果

Table 4 Ablation experimental results on DeepCrack

Methods	P	R	F -score
VGG ⁺	0.8660	0.8562	0.8611
VGG ⁺ _SA	0.8745	0.8576	0.8660
VGG ⁺ _DCFF	0.8761	0.8619	0.8690
VGG ⁺ _SA_DCFF	0.8807	0.8619	0.8712
VGG_SA_DCFF	0.8810	0.8596	0.8702

由表 4 前 4 行的数据可以看出,注意力机制的加入和并行轻量级空洞卷积融合模块(DCFF)都对指标提升做出了较大的贡献,其中 DCFF 模块对性能的提升更加明显,两者组合在一起将 F -score 提升到了 0.8712。比较表 4 最后两行可知,VGG⁺ 给检测指标带来了微小的提升(VGG⁺ 表示在图 1 网络结构中 Conv2 卷积阶段增加了一层卷积)。

结束语 本文提出了一种基于空洞卷积和注意力机制的混凝土路面裂缝检测方法。在特征提取阶段,低层特征通过空间注意力机制融入主干网络,为后续卷积提供了更多的细节信息;在特征融合阶段,将中高层特征经过并联的轻量级空洞卷积模块,在增大感受野的同时提高了特征图的分辨率,并且侧边特征渐进式地被多次使用。最后将并联的轻量级空洞卷积模块的输出与低层侧边输出相加作为最终的预测图像。本文提出的方法提高了裂缝检测的精确性,也具有较弱的鲁棒性。

未来将继续研究裂缝检测的问题并进一步提高算法的精度和模型的泛化能力,尝试将目标检测的方法用于裂缝检测。

参考文献

[1] CHENG H, SHI X, GLAZIER C. Real-Time Image Thresholding Based on Sample Space Reduction and Interpolation Approach [J]. Journal of Computing in Civil Engineering, 2003, 17(4):264-272.

[2] NEJAD F M, ZAKERI H. An Optimum Feature Extraction Method based on Wavelet-Radon Transform and Dynamic Neural Network for Pavement Distress Classification [J]. Expert Systems with Applications, 2011, 38(8):9442-9460.

[3] 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 Intelligence Transport System, 2016, 17(10):2718-2729.

[4] YAMAGUCHI T, HASHIMOTO S. Fast crack detection method for large-size concrete surface images using percolation-based image processing [J]. Machine Vision and Applications, 2010, 21(5):797-809.

[5] FAN Z, LI C, CHEN Y, et al. Automatic Crack Detection on Road Pavements Using Encoder-Decoder Architecture [J]. Materials, 2020, 13(13):2960.

[6] QU Z, CAO C, LIU L, et al. A Deeply Supervised Convolutional Neural Network for Pavement Crack Detection with Multiscale Feature Fusion [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(9):4890-4899.

[7] XU G Z, LIAO C G, C J J, et al. Extraction of Apparent Crack Information of Concrete Based on HU-ResNet [J]. Computer Engineering, 2020, 46(11):279-285.

[8] HU J, LI S, GANG S, et al. Squeeze-and-Excitation Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(8):2011-2023.

[9] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [C]//European Conference on Computer Vision. Berlin: Springer, 2018:3-19.

[10] CHEN Y, DAI X, LIU M, et al. Dynamic Convolution: Attention over convolution kernels [C]//IEEE Computer Vision and Pattern Recognition. 2020:11030-11039.

[11] ZHAO T, WU X Q. Pyramid Feature Attention Network for Sa-liency detection [C]//IEEE Computer Vision and Pattern Recognition. 2019:3080-3089.

[12] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation [J]. arXiv:1706.05587, 2017.

[13] CHOLLET F. Xception: Deep Learning with Depthwise Separable Convolutions [C]//IEEE Computer Vision and Pattern Recognition. IEEE, 2017:1063-6919.

[14] LIU Y H, YAO J, LU X H, et al. DeepCrack: A Deep Hierarchical Feature Learning Architecture for Crack Segmentation [J]. Neurocomputing, 2019, 338:139-153.

[15] SHI Y, CUI L, QI Z, et al. Automatic Road Crack Detection using Random Structured Forests [J]. IEEE Transactions Intelligence Transport System, 2016, 17(12):3434-3445.

[16] 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.

[17] REN Y P, HUANG J S, HONG Z Y, et al. Image-based Concrete Crack Detection in Tunnels Using Deep Fully Convolutional Networks [J]. Construction and Building Materials, 2020, 234(20):117367.

[18] GU Z W, CHENG J, FU H Z, et al. CE-Net: Context Encoder Network for 2D Medical Image Segmentation [J]. IEEE Transactions on Medical Imaging, 2019(10):2281-2292.



QU Zhong, born in 1972, Ph.D, professor. His main research interests include digital image processing, machine learning, neural networks and artificial intelligence.