

基于可分离卷积与小波变换融合的道路裂缝检测

刘云清, 吴越, 张琼, 颜飞, 陈姗姗

引用本文

刘云清, 吴越, 张琼, 颜飞, 陈姗姗. 基于可分离卷积与小波变换融合的道路裂缝检测[J]. 计算机科学, 2024, 51(11A): 240100141-9.

LIU Yunqing, WU Yue, ZHANG Qiong, YAN Fei, CHEN Shanshan. Road Crack Detection Based on Separable Convolution and Wave Transform Fusion [J]. Computer Science, 2024, 51(11A): 240100141-9.

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)

[基于多模态融合的动态恶意软件检测方法](#)

Multimodal Fusion Based Dynamic Malware Detection

计算机科学, 2024, 51(11A): 240200098-7. <https://doi.org/10.11896/jsjcx.240200098>

[MB-ATMK:融合属性权重和时序元知识的多行为序列推荐模型](#)

MB-ATMK:Multi-behavior Sequential Recommendation Integrating Attribute Weights and Temporal Meta-knowledge

计算机科学, 2024, 51(11A): 231100047-9. <https://doi.org/10.11896/jsjcx.231100047>

[FCTNet:基于双域深度学习的公交车到站时间预测方法](#)

FCTNet:Bus Arrival Time Prediction Method Based on Dual Domain Deep Learning

计算机科学, 2024, 51(11A): 231000180-7. <https://doi.org/10.11896/jsjcx.231000180>

[基于季节分解的混合神经网络的时间序列预测](#)

Time Series Prediction of Hybrid Neural Networks Based on Seasonal Decomposition

计算机科学, 2024, 51(11A): 231200008-7. <https://doi.org/10.11896/jsjcx.231200008>

[基于相对位置编码转换器模块的深度步态识别网络](#)

Deep Gait Recognition Network Based on Relative Position Encoding Transformer

计算机科学, 2024, 51(11A): 240400064-6. <https://doi.org/10.11896/jsjcx.240400064>

基于可分离卷积与小波变换融合的道路裂缝检测

刘云清^{1,2} 吴越¹ 张琼^{1,2} 颜飞^{1,2} 陈姗姗¹

1 长春理工大学电子信息工程学院 长春 130000

2 吉林省智能感知与信息处理科技创新中心 长春 130000

摘要 针对目前对细小裂缝检测能力不强、分割精度低等问题,提出了一种改进的 U-Net 模型来检测路面裂缝,提高检测能力和分割精度。中文设计了新的模块 MSDWBlock(Multi-Scale Depthwise Separable Convolutional Block),应用在编码器和解码器部分,通过深度可分离卷积增强模型的能力,扩大模型感受野,在跳跃连接部分引入了 C2G 注意力机制模块,提升模型对裂缝特征的感知能力;并引入了 ASPP(Atrous Spatial Pyramid Pooling)和 DWT(Discrete Wavelet Transformation)。ASPP 通过在多个尺度上进行操作,有助于捕捉到裂缝的特征,而 DWT 能够减少卷积池化过程中的裂缝空间信息损失,保留裂缝边缘信息。这种结构设计使得网络更专注于裂缝的特征,从而提升了裂缝检测的准确性。通过实验证明所提模型显示出优于 U-Net,Segnet,U2net 等先进模型的精确性。在 CFD 数据集上 mIoU,F1 分别达到 78.51%,0.868。这些成果表明,所提方法能有效提升道路裂缝检测的性能。

关键词: 裂缝检测;U-Net 神经网络;深度可分离卷积;注意力机制;空间金字塔;小波变换

中图分类号 TP391

Road Crack Detection Based on Separable Convolution and Wave Transform Fusion

LIU Yunqing^{1,2}, WU Yue¹, ZHANG Qiong^{1,2}, YAN Fei^{1,2} and CHEN Shanshan¹

1 School of Electronics Information, Changchun University of Science and Technology, Changchun 130000, China

2 Jilin Provincial Science and Technology Innovation Center of Intelligent Perception and Information Processing, Changchun 130000, China

Abstract Aiming at the current problems of weak detection ability and low segmentation accuracy for small cracks, an improved U-Net model is proposed to detect road cracks and improve detection ability and segmentation accuracy. This paper designs a new module, multi scale depth separated convolutional block (MSDWBlock), which is applied in the encoder and decoder sections. Through its depthwise separable convolution, the model's ability is enhanced, the model's receptive field is expanded, and a C2G attention mechanism module is introduced in the skip connection section to enhance the model's perception of crack features. And atrous spatial pyramid pooling (ASPP) and discrete wavelet transformation (DWT) are introduced. ASPP helps to capture the characteristics of cracks by operating at multiple scales, while DWT can reduce the loss of crack spatial information during convolutional pooling and preserve crack edge information. This structural design makes the network more focused on the characteristics of cracks, thereby improving the accuracy of crack detection. It has been demonstrated through experiments that the accuracy of the proposed model is better than that of advanced models such as U-Net, Segnet, and U2net. On the CFD dataset, mIoU and F1 reaches 78.51% and 0.868 respectively. These results indicate that the proposed method can effectively improve the performance of road crack detection.

Keywords Crack detection, U-net neural network, Depthwise separable convolutional, Attention mechanism, Spatial pyramid, Wavelet transform

1 引言

裂缝作为道路表面破损的初期征兆,是最常见的道路损坏形式之一^[1]。它不仅削弱了道路的使用性能,更对交通安全和运输效率构成了严重威胁^[2]。及时检测道路裂缝不仅有助于改善道路性能,还可以确保道路的适用性^[3]。随着技术的飞速进步,道路裂缝的检测逐步从传统的人工方法过渡到基于图像处理的自动化技术。如 Shi 等^[4]使用积分

通道特征结合随机森林算法来识别裂缝特征;Zou 等^[5]则提出了削除路面阴影和加强裂缝对比度的算法。传统的裂缝检测方法大多是基于图像处理区域生长^[6]、基于边缘检测的方法^[7]和基于阈值的方法^[8]。这类方法的核心在于利用基本图像属性,例如图形的连续性、表面纹理和形状特征。然而,这些属性的简单性也导致了检测结果对环境噪声、影子效应以及光线变化等外界干扰的高度敏感性^[9]。这些传统技术在速度上优于人工检测,提高了效

基金项目:国家自然科学基金青年科学基金(42204144);吉林省自然科学基金(YDZJ202101ZYTS064)

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China(42204144) and Natural Science Foundation of Jilin Province, China(YDZJ202101ZYTS064).

通信作者:刘云清(234577142@qq.com)

率,但它们无法完全避免由环境因素干扰而引起的误判。

近年来研究人员通过神经网络的浅层卷积进行裂缝检测研究。Zhang 等^[10]的研究表明,相较于传统机器学习方法,CNN 对路面裂缝的检测更为准确。然而,这类方法可能会导致对裂缝宽度的高估,从而影响检测精度。为了解决这一问题,Zhang 等^[11]提出了一种基于 CNN 的高效网络。该网络架构称为 CrackNet,专门用于自动检测沥青表面裂缝。

U-Net^[12]作为一种全卷积神经网络模型,因其多尺度特征融合能力和即使在有限数据集上也能实现精确分割的优势,已被广泛应用于医学图像分割等领域。近年来,U-Net 亦被应用于桥梁、隧道等基础设施的裂缝检测^[13]中。U-Net 的优势是其能在数据稀缺的情况下实现对裂缝的精确像素级分割,为当前数据有限的问题提供了解决方案^[14]。基于 U-Net 的衍生模型也不断被开发并应用于道路表面缺陷的识别工作,Crack U-Net 模型便是其中之一,它能够精确地识别出由摄像头采集的图像中的道路裂缝^[15]。

以上研究在道路检测的过程中依然存在对裂缝的检测精度不足、多尺度裂缝识别能力有限等问题,针对以上问题本文提出了一种基于新模块的 U-Net 模型,并引入了 C2G 注意力机制与 ASPP-DWT Block 模块,以提升对裂缝检测的精度和鲁棒性。

本文的主要贡献包括:

1)提出了一种新型的针对裂缝特征的提取模块 MSDW-Block,通过结合深度可分离卷积和多尺度卷积核来获取更大

的感受野,提高了模型提取裂缝特征的能力。

2)针对细小裂缝识别的问题,引入了 C2G 注意力机制,该机制通过在特征图上加重聚焦于重要的特征,从而提高模型对细小裂缝的敏感性和识别能力。通过这种方式,即使在复杂的背景或低对比度的条件下,模型也能够有效地区分出裂缝的存在。

3)针对多尺度裂缝特征识别能力有限问题,采用 ASPP-DWT Block 模块。ASPP 能够在多个尺度上提取特征,扩大模型的感受野,从而帮助模型捕捉到不同大小的裂缝。同时,DWT 通过小波变换捕捉和强化裂缝的信息,并减少卷积和池化过程中的信息损失。

2 改进网络模型

2.1 改进的 U-net 网络模型整体结构

U-Net 模型通过结合下采样路径获得的上下文信息和上采样路径获得的精确定位信息,对于捕获丰富的空间信息以检测多尺度道路裂缝至关重要。本文基于 U-Net 网络模型的改进模型如图 1 所示。为了增强对路面裂纹识别的效能与准确性,本文在下采样(编码)阶段引入 MSDWBlock 模块,如图 2 所示。在编码过程中,使用 MSDWBlock 模块从输入图像中提取具有丰富特征通道信息的特征。MSDWBlock 可以通过其内部的深度可分离卷积有效地提取特征。这种卷积首先使用深度卷积(每个输入通道独立的卷积操作)来捕获空间特征,然后利用点卷积(1×1 卷积)聚合通道信息。

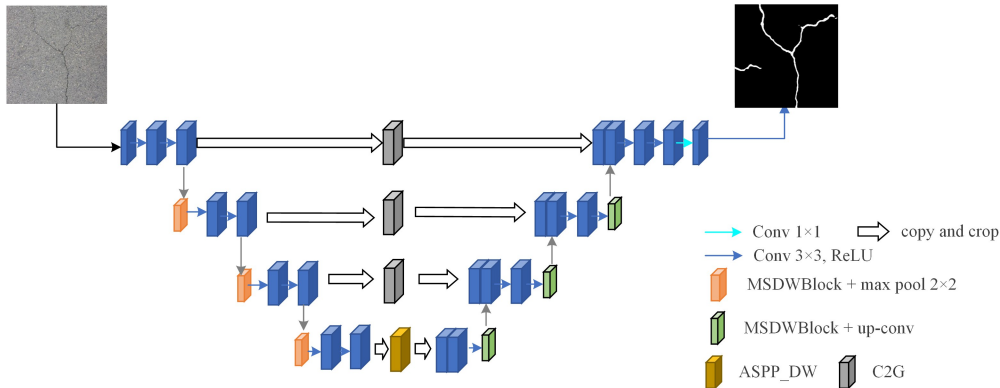


图 1 改进的 U-Net 网络模型整体结构

Fig. 1 Overall structure of the improved U-Net network model

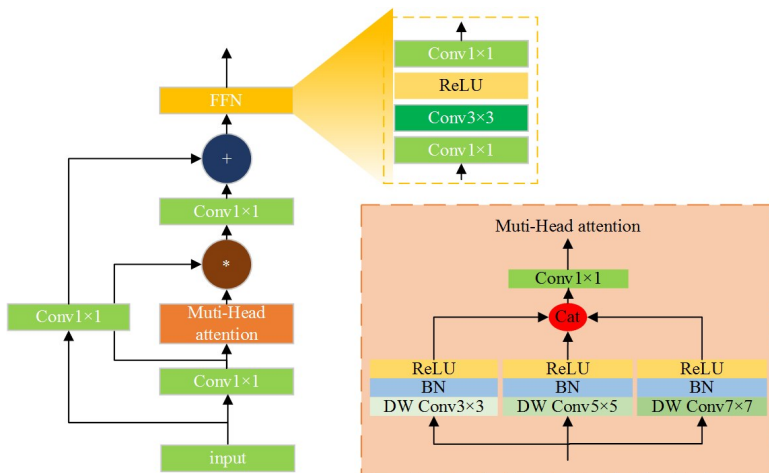


图 2 MSDWBlock 模块结构图

Fig. 2 MSDWBlock module structure diagram

在上采样(解码)阶段,每个上采样模块之后,将编码阶段的输出与上采样结果进行拼接,紧接着通过一个 MSDW-Block 块进行进一步的特征提取。MSDWBlock 可以帮助网络从低分辨率的特征映射中恢复出更多细节以及高分辨率的输出,这对于像素级别的任务(如裂缝检测)来说非常重要。通过逐步结合来自编码器的特征和逐层放大的特征图,MSDWBlock 可以帮助网络恢复出清晰的裂缝边界。

为了提高模型对裂缝识别的准确性,本文引入一种改进的注意力机制模型——C2G 注意力模块。该模块专注于更精细地提取和关注输入图像的重要特征,通过结合自适应最大池化和自适应平均池化策略,分别捕获裂缝图像的显著特征和全局上下文信息。模块内部的 CG 子模块和卷积层充当参数调整分支,通过改变卷积核尺寸和步长,以及采用 GELU 激活函数,实现多尺度特征的提取。C2G 模块还采用了空洞卷积的设计原则,避免了由空洞卷积堆叠引起的网格效应,使模型能够获取更丰富和精细的特征信息。通过 Sigmoid 函数,实现两种类型的注意力信息的有效融合。

采用了结合 ASPP 和 DWT 的模块。ASPP-DWT Block 模块通过在 ASPP 结构中新增一个池化分支和一个 DWT 分支来增强对裂缝特征的提取。新增的 pool 分支目的是获取全局上下文信息,而 DWT 分支则是一种特征提取手段,其垂直(v)、水平(h)和对角(d)分支有助于捕获裂缝图像在这 3 个方向上的特征。

2.2 改进的 U-Net 采样阶段

在多种检测任务中,不同深度和尺度的特征图携带的信息对检测结果有着显著影响。在卷积神经网络(CNN)体系结构中,初级层次的特征图倾向于包含丰富的局部纹理和边缘信息,呈现出更细致的结构细节,但这些特征通常伴随较低的语义水平和较高的噪声成分。相反,高级层次的特征图具备更强的语义表示能力,但由于重复的下采样操作,它们往往损失了一些空间细节,导致分辨率下降。

在道路裂缝检测的应用场景中,裂缝本身的微妙结构与其周围的视觉噪声具有一定程度的相似性,这即使在人工视觉检测中也是一个难题。低层次的卷积层能够捕捉到大量关于裂缝的信息。然而,当受到复杂道路环境背景的干扰时,这些特征图会包含许多噪声、暗斑和污点等。通过连续多层卷积和池化处理,高层次的卷积能够从这些初始特征图中抽象出更高级别的特征,减少噪声和其他干扰成分,但与此同时,网络深度的增加也可能导致诸多微小裂缝的细节信息丢失。这种情况下,高层提取的特征图将无法反映真实的路面状况,从而降低了检测精度。各层次的特征图尽管具有各自的优势和局限,但它们之间存在潜在的互补性。通过将低层次的细节特征图与高层次的语义特征图进行有效融合,可以综合利用各层特征的优点,以提升裂缝检测的整体精确度。针对该问题,设计了一种新的模块 MSDWBlock,如图 2 所示。该模块通过结合深度可分离卷积和多尺度卷积核,能够融合多尺度和多层次的特征,获取更大的感受野,并且使最终的预测图更接近路面真实状况。使用 MobileNet^[16]中的深度可分离卷积结构作为 attention 层,以便加强模型对感受野信息的获取能力。

在模型的编码阶段,MSDWBlock 的深度可分离卷积取代了传统的卷积层。这些卷积层之后分别接有 BN(Batch

Normalization)层和 ReLU 激活函数。下采样是通过最大池化层实现的,编码器输出的特征图将进一步在解码阶段进行特征融合。

解码阶段保留了 U-Net 中的跳跃连接和上采样结构,以逐通道方式将编码阶段的特征图与解码阶段的上采样特征图拼接,从而有效地传递细节信息,提高裂缝检测的精确度。每个跳跃连接合并后,继续应用深度可分离卷积层来提取特征,并通过 BN 层和 ReLU 激活函数进行进一步处理。之后,上采样层被用于逐步放大特征图,直至恢复至原始输入的尺寸。

在输出层,采用 1×1 卷积将特征图映射到指定的类别数量(文中设置数量为 2),然后通过 softmax 或 sigmoid 激活函数获得每个像素点的类别概率值。

2.3 MSDWBlock 模块结构设计

MSDWBlock 模块的核心是将输入特征图通过一个 1×1 卷积层进行通道数调整,接着应用 GELU 作为激活函数,以增强非线性表达能力。经过这一处理步骤得到的特征图称为 x_{att_in} ,它不仅可用于后续的深度可分离卷积,而且还可与原始输入进行特征融合。为了捕捉不同尺度的裂缝信息,本文设计了 3 个并行的深度可分离卷积块,分别使用 3×3 , 5×5 和 7×7 的卷积核。这些块的设计利用了分组卷积,先对各个输入通道进行独立的空间卷积,然后通过 1×1 卷积来聚合特征。每个卷积块都紧跟着一个批量归一化层和 ReLU 激活函数,以稳定训练过程并引入非线性。接下来,通过一个 1×1 卷积的通道混合层,将 x_{att_in} 与 3 个尺寸的特征图在通道维度进行拼接,实现特征的增强。这样的设计使得网络能够同时考虑局部细节和较大范围的上下文信息,对于裂缝这种具有多尺度的目标检测至关重要。

最终,采用了一个前馈网络(FFN),它包括 1×1 卷积、 3×3 卷积和 1×1 卷积,能够进一步增强特征表示。FFN 的输出是经过 MSDWBlock 模块处理后的特征图,确保经过模块处理的特征图能够有效地用于后续的网络结构中。

2.4 基于 C2G 注意力机制的特征融合网络

在背景干扰下,裂缝在视觉上特征并不显著。针对该问题,本文设计了 C2G 注意力模块,旨在针对网络对裂缝特征进行识别和获取,C2G 注意力机制如图 3 所示。C2G 模块利用通道注意力与空间注意力机制,不仅考虑通道间的全局统计信息,而且融合了局部及更广泛的空间上下文信息,实现了特征融合网络中对细微裂缝信息特征的增强。

通过自适应最大池化和自适应平均池化,C2G 模块有效地捕获了每个通道的全局最大值和全局平均值,这两种统计结果为模型提供了关于不同通道重要性的指标。

CG 子模块设计为一个卷积网络,包含两个卷积层及 GELU 激活函数。第一个卷积层(1×1 卷积核)维持通道维度不变,而第二个卷积层(3×3 卷积核,边缘填充为 1)则是为了捕捉局部上下文信息。这种结构不仅考虑了通道间的关系,还考虑了空间上的相邻关系。并且使用 5×5 卷积核进一步处理平均池化后的特征,目的是捕获更广泛的空间上下文信息,增加模型对裂缝空间细节的敏感度,从而有助于识别和分割裂缝区域。

在这两个特征处理路径之后,C2G 模块通过逐元素乘法将处理后的特征与原始输入特征相结合,从而重新标定了通道和空间信息。随后,将这两个增强的特征图叠加,并通过 Sig-

moid 激活函数生成注意力图。该注意力图介于 0~1 之间，C2G 注意力模块生成的注意力图与输入特征相乘，有效地对

特征进行了重标定和强化。对于在视觉上不显著的裂缝图像，也能检测到相应的细节信息。

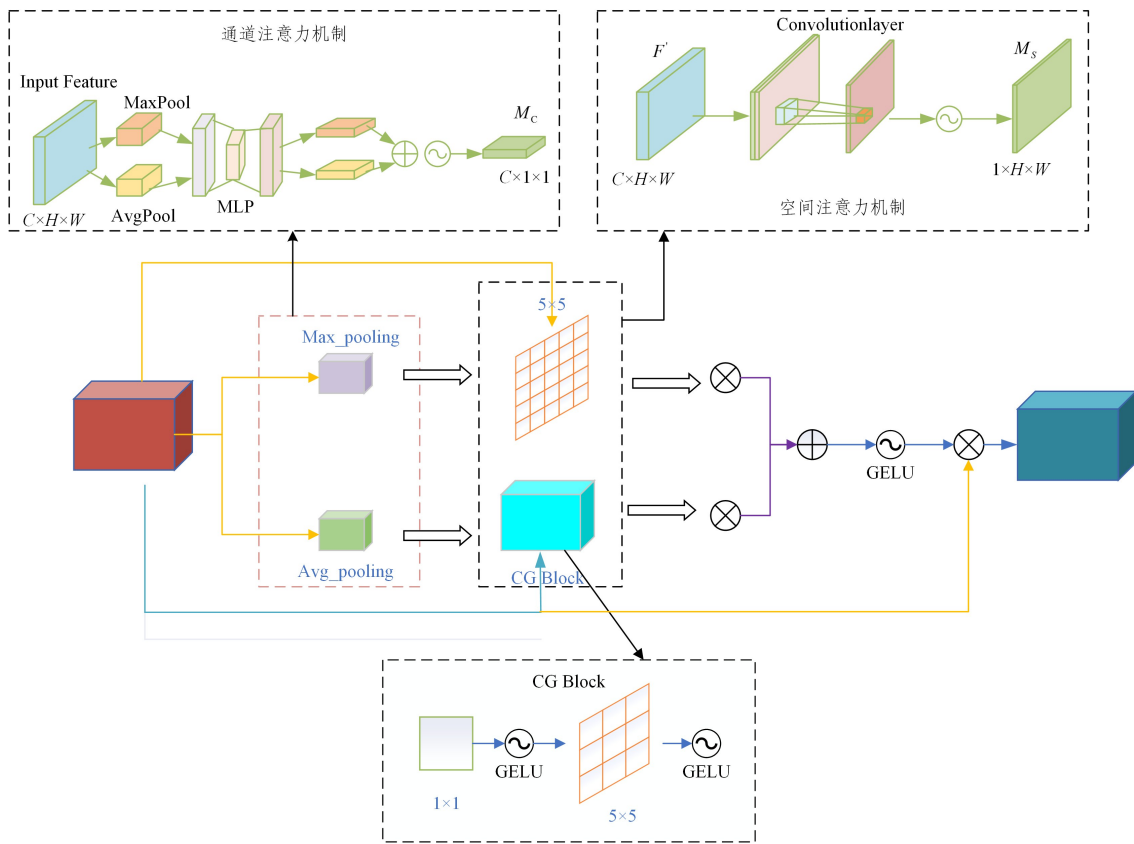


图3 C2G 注意力机制

Fig.3 C2G attention mechanism

2.5 改进的金字塔池化模块

2.5.1 小波变换的应用

在裂缝检测任务中，裂缝的空间细节信息对于检测精度至关重要。小波变换通过多级分解的特性，能够在不同频率范围内提取多个子频带，从而拥有在多分辨率和多尺度分析方面的能力。Guo 等^[17]利用了小波系数的稀疏性和结构性信息来提高中间层的稀疏性并丰富其结构性信息，使用 Haar 小波子带作为输入来恢复图像的缺失细节。

小波变换之所以适合于裂缝检测任务，是因为它可以减少池化和卷积操作中可能造成的空间信息丢失。通过连续应用高通和低通滤波器，小波变换使信号能够在不同的细节水平上进行分解。近年来，小波变换与深度学习的结合也开始展现出潜力，Bae 等^[19]结合小波变换与残差网络，研究发现小波变换的更多子带可以提高网络的学习效果。Li 等^[19]提出用小波变换代替神经网络中的池化操作，保留原始图像的高频信息以及边缘细节。

小波变换还可以嵌入到编码器-解码器的过程中。综上，本文采用了 2D Haar^[21] DWT 与 U-Net 相结合的方法，以保留不同的裂缝空间信息特征。采用 2D Haar 小波变换的 DWT 能够实现对原始图像的多尺度分解，执行一层小波变换。通过该变换，一幅原始图像可被分解为 4 个不同的子频带图像，分别对应原始图像在不同方向和尺度上的频率分量，4 个子带图像的尺寸均为原始图像的一半，因此总体分辨率降至原图的 1/4。在这一过程中，分别对应低通和高通滤波器的作用，相当于使用 4 个滤波器

$(f_{LL}, f_{LH}, f_{HL}, f_{HH})$ 对原始图像 x 进行分解，获得 $(x_{LL}, x_{LH}, x_{HL}, x_{HH})$ 4 个子带图像，其中 A 为低频图像， V 为竖直细节图像， H 为水平细节图像， D 为对角细节图像。滤波器的参数固定，不随网络训练的反向传播进行梯度下降。

2.5.2 ASPP 与小波变换结合

在深度网络模型中，频繁的下采样过程会减少输入图像的分辨率，随着网络深度的增加，图像的细节特征逐渐流失。现有的道路裂缝检测方法大多都基于卷积神经网络来提取特征，但是随着网络层数的堆叠，可提取的特征越来越有限，这不仅导致了模型参数数量的膨胀，而且对模型的准确率并没有带来预期的提升，反而可能会降低检测的精度和速度。为解决这一问题，一些裂缝检测算法开始采用空洞卷积技术。

通过空洞卷积，可以减少对下采样的依赖，从而保留更多的裂缝细节信息，减少感受野缩小导致的问题。在裂缝检测任务中，由于裂缝特征的背景可能非常复杂，且裂缝的形态多样，仅使用固定空洞率的连续卷积层虽然能够通过空洞卷积在初级层次上扩展感受野，以便捕获更多的裂缝特征，但当网络逐渐深入到更高层次时，最顶层的卷积层可能会变得太过稀疏。这种稀疏性可能会导致对细小裂缝的检测力度下降，同时宽裂缝的关键局部特征也可能遗失，进而大幅度削弱检测的准确性。

不仅要提取特征图中提取深层次的语义信息进行像素级分类，还要利用浅层特征图中的边缘和轮廓信息以增强定位精确度。利用金字塔池化可以在编解码过程中实现深层语义信息与细节轮廓信息的有效融合。此外，引入小波变换有助于

在卷积和池化操作中减少信息丢失。

本文引入了 ASPP 模块,该模块使用具有不同扩张率的空洞卷积并行处理输入特征图,从而在不增加额外参数的情况下扩大感受野。还可通过捕获多尺度信息,使模型适应不同尺寸的裂缝,保证了对大面积和微小裂缝都有良好的识别率,并且与小波变换相结合以改善道路裂缝分割的性能。

在本文的方法中,ASPP 模块被改进,通过并行的空洞卷积,引入了 3 个不同尺度的空洞卷积。改进的 ASPP-DWT 模块图如图 4 所示。小波变换被用于分解图像,得到 3 个高频子带,这些子带分别捕获图像水平、垂直和对角线方向的细节信息。在改进的 ASPP 模块中,DWT 被用于分解输入图像,得到一个低分辨率的近似图像和 3 个细节系数图像;LL 表示低频近似系数,它包含了图像的主要分布,有助于捕捉裂缝的全局信息;LH,HL 和 HH 分别代表水平、垂直和对角线方向的高频细节系数,它们可突出裂缝相关的细节和纹理。

通过不同膨胀率的卷积核来捕获裂缝特征信息,并结合 ASPP 和 DWT 的方法使模型能够同时不同的尺度和方向上工作。

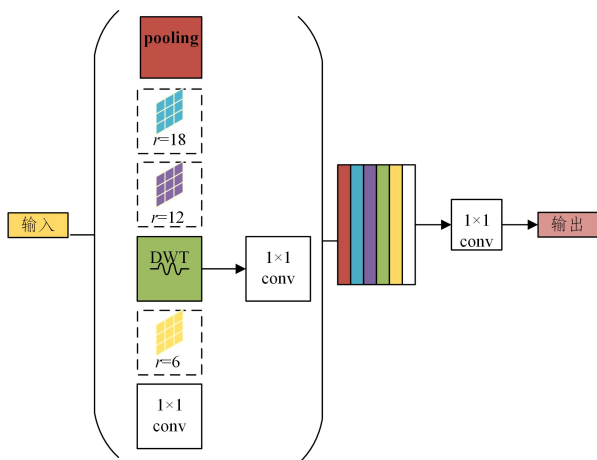


图 4 ASPP-DWT 模块图

Fig. 4 ASPP-DWT module

3 实验

3.1 实验环境配置

本文实验环境硬件配置:使用 Windows 11 操作系统,

运行内存 16 GB,处理器 AMD Ryzen 7-5800H,显卡为 NVIDIA RTX 3090,程序语言为 Python 3.8,本文所有模型基于 PyTorch 深度学习框架实现,使用 GPU 加速。为了防止训练过程中模型陷入局部最优解,采用 Adam 优化器对训练进行优化。

3.2 数据集

本文实验使用的数据集是 Crack500 和 CFD(Crack Forest)数据集。Crack 500 数据集经过处理后包含 3368 张裂缝图像,其中训练数据集 3000 张图像,验证数据集 368 张图像。所有图像均配备有像素级人工标注的裂缝真值图像。CFD 数据集包含 118 幅大小为 480×320 像素的沥青路面图像,在数据集中有白色标线、阴影等背景噪声,标签图像全部由手工标记。为了提高模型的泛化能力,采用随机翻转、裁剪等数据增强方法扩充 CFD 数据集。

3.3 评价指标

为了评估不同的分割模型,本文选取的分割精度评估指标包括调和平均数(F1-score)、召回率(Recall, R)、精确率(Precision, P),以及平均交并比(mean Intersection over Union, mIoU)。其中 F1 值能够同时衡量精准率和召回率,在裂缝检测遗漏少的前提下误判越少,F1 值也会越高;R 表示模型正确检测的裂缝像素占有所有裂缝像素的百分比;P 表示正确检测的裂缝像素原本是正确像素的比例。

$$F1 = \frac{2 \times P \times R}{P + R} \quad (1)$$

$$R = \frac{T_{TP}}{T_{TP} + F_{FN}} \quad (2)$$

$$P = \frac{T_{TP}}{T_{TP} + F_{FP}} \quad (3)$$

$$mIoU = \frac{1}{2} \sum_{i=0}^1 \frac{TP}{TP + FP + FN} \quad (4)$$

其中, T_{TP} 表示被正确分为裂缝像素的数量; F_{FP} 表示被错误分为裂缝像素的数量。

3.4 改进模型的性能评估

为了验证所提出网络的有效性,将其与 U-Net, Segnet, linknet, U2net, Atten-unet 置于同一硬件环境及相同的参数初始化条件下,分别在 CFD 和 Crack500 数据集上进行训练和测试。在这两个数据集上不同网络的检测结果图像如图 5 和图 6 所示。

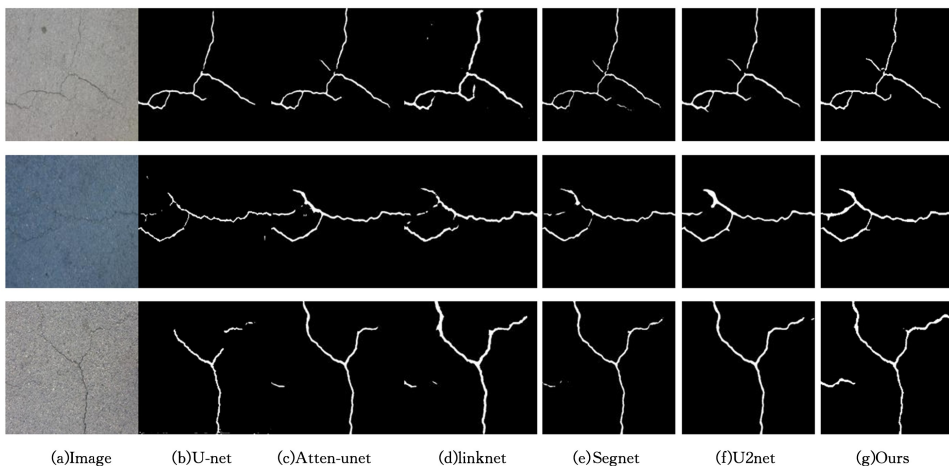


图 5 不同模型在 CFD 数据集上的裂缝检测效果

Fig. 5 Crack detection effects of different models on CFD dataset

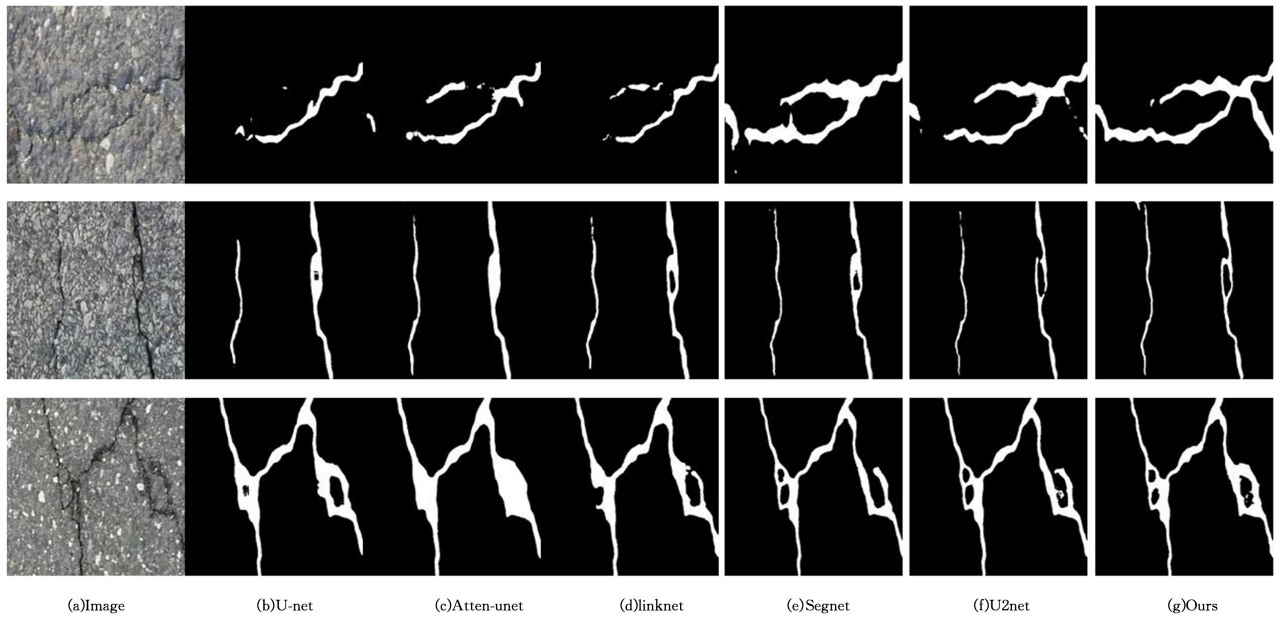


图6 不同模型在Crack500数据集上的裂缝检测效果

Fig. 6 Crack detection effects of different models on Crack500 dataset

此外,本文使用所提出的评估指标对不同的模型结果进行分析评估,对比实验的评估结果中黑体加粗数据为最优数据。

表1 不同模型在CFD数据集上的评价指标

Table 1 Evaluation indicators of different models on CFD dataset

模型	<i>mIoU</i>	<i>P</i>	<i>R</i>	<i>F1</i>
U-net	69.32	0.751	0.802	0.776
Atten-unet	72.65	0.789	0.837	0.812
linknet	74.38	0.805	0.843	0.824
Segnet	76.28	0.827	0.866	0.846
U2net	77.21	0.84	0.876	0.858
本文模型	78.51	0.851	0.881	0.868

表2 不同模型在Crack500数据集上的评价指标

Table 2 Evaluation indicators of different models on Crack500 dataset

模型	<i>mIoU</i>	<i>P</i>	<i>R</i>	<i>F1</i>
U-net	73.39	0.804	0.836	0.820
Atten-unet	77.62	0.846	0.872	0.859
linknet	78.85	0.851	0.878	0.864
Segnet	79.74	0.858	0.884	0.871
U2net	80.71	0.864	0.891	0.877
本文模型	81.24	0.873	0.897	0.885

由表1可知,本文模型在CFD数据集上的*mIoU*,*P*,*R*,*F1*的数值均高于其他5种模型,在该数据集上本文模型表现最好,4项指标比原始U-Net分别高出9.19%,10%,7.9%,9.2%。由表2可知,本文模型在Crack500数据集上的*mIoU*,*P*,*R*,*F1*的数值均高于其他5种模型,在该数据集上本模型表现最好,4项指标比原始U-Net分别高出7.85%,6.9%,6.1%,6.5%;引入注意力机制以后,网络模型能对裂缝赋予更多的权重,使得网络更加关注裂缝信息。而MSD-WBlock模块和ASPP-DWT结构可以获取感受野和捕获多尺度信息的能力,并且能够减少卷积和池化过程中的信息损失,提供更加丰富的空间位置信息。从图5的第2行和3行能够明显看出,改进后的模型相比原始网络U-Net能够更加关注裂缝区域,抑制背景干扰的能力更强,以此来提高模型检测的准确率。在图6的第三行能比较清楚地看出,改进后的

模型更加关注裂缝的细节信息,与本实验其他对比模型相比,本文模型可以更加精确地检测到细节信息。最终的改进模型提取微小裂缝的完整度较高。

3.5 改进模型的消融实验

在本文中,通过一系列的消融实验对逐步构建的模型架构进行了验证与分析;研究重点集中于探讨U-Net网络与不同模块融合时的多种配置情况,并分别评估了以下配置对道路裂缝检测精度的影响:单独融合MSDWBlock模块、单独融合ASPP-DWT模块、同时融合ASPP-DWT和MSDWBlock模块、同时融合ASPP-DWT和C2G模块,以及ASPP-DWT,MSDWBlock和C2G模块三者联合融合。

在CFD数据集上的消融实验结果如表3所列,加粗字体表示最优数据,可以看到本文最终的改进模型在所有指标上均优于其余阶段的结果。从表3中可以看出,在CFD数据集上将U-Net模型分别加入ASPP-DWT,MSDWBlock和C2G模块后,相较于原始网络*mIoU*分别提升2.52%,2.93%,4.55%;相同地融入ASPP-DWT以后分别加入MSDWBlock和C2G模块后,后者相较于前者*mIoU*提升1.63%,在相同情况下,C2G模块对于提升模型的抗干扰能力和检测性能更强;最后在全融合后,相较于原始U-Net网络*mIoU*提升9.19%。从图7的消融图像中可以清晰地看到,本文模型在检测裂缝细节上均优于其他的实验对比模型。

表3 在CFD数据集上的消融实验结果

Table 3 Ablation experimental results on CFD dataset

编号	U-net	ASPP-DWT	MSDW-Block	C2G	<i>mIoU</i>	<i>P</i>	<i>R</i>	<i>F1</i>
a	✓	—	—	—	69.32	0.751	0.802	0.776
b	✓	✓	—	—	71.84	0.775	0.824	0.799
c	✓	—	✓	—	72.13	0.782	0.829	0.805
d	✓	—	—	✓	73.87	0.794	0.838	0.815
e	✓	✓	✓	—	74.35	0.802	0.841	0.821
f	✓	✓	—	✓	75.98	0.824	0.859	0.841
g	✓	—	✓	✓	75.14	0.821	0.853	0.837
h	✓	✓	✓	✓	78.51	0.851	0.881	0.868

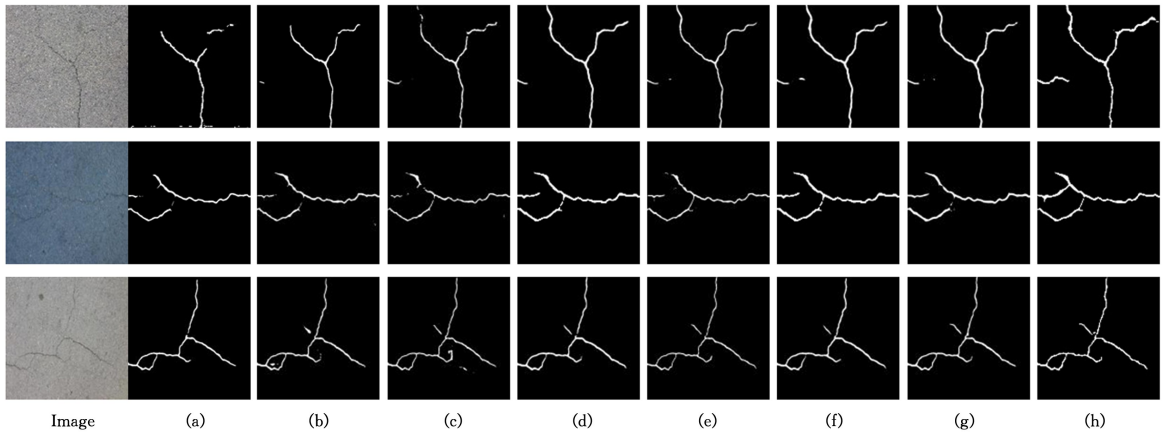


图 7 CFD数据集上消融实验图像对比

Fig. 7 Comparison of ablation experimental images in CFD dataset

从表 4 中可以看出,在 Crack500 数据集上:将 U-Net 模型分别加入 ASPP-DWT,MSDWBlock 和 C2G 模块后相比原始网络 mIoU 分别提升 1.18%,3.37%,3.92%;再加入改进方法后模型对细节信息的检测效果比原始 U-Net 模型更优。在图 8 第三行,除了能检测出裂缝的位置,还可以对裂缝的细节形状信息进行检测还原。

实验结果表明,相比原始 U-Net 模型,本文融合 ASPP-DWT,MSDWBlock 和 C2G 模块的模型检测能力较优,可以有效提升 U-Net 对裂缝的检测性能。

表 4 在 Crack500 数据集上的消融实验结果

Table 4 Ablation experimental results on Crack500 dataset

编号	U-net	ASPP-DWT	MSDW-Block	C2G	mIoU	P	R	F1
a	✓	—	—	—	73.39	0.804	0.836	0.820
b	✓	✓	—	—	74.57	0.817	0.842	0.829
c	✓	—	✓	—	76.76	0.835	0.863	0.849
d	✓	—	—	✓	77.31	0.843	0.871	0.857
e	✓	✓	✓	—	79.74	0.856	0.882	0.869
f	✓	✓	—	✓	80.35	0.862	0.889	0.875
g	✓	—	✓	✓	80.12	0.860	0.886	0.873
h	✓	✓	✓	✓	81.24	0.873	0.897	0.885

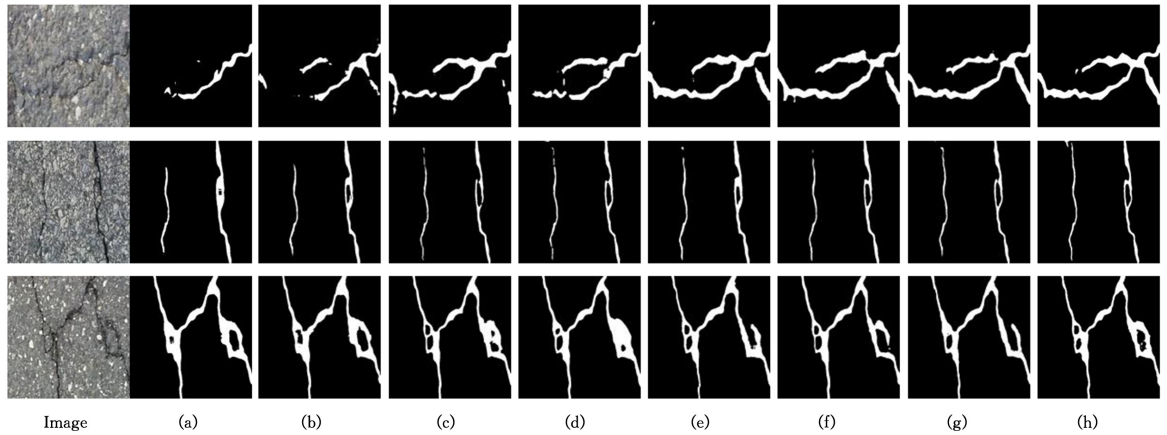


图 8 Crack 500数据集上消融实验图像对比

Fig. 8 Comparison of ablation experiment images in Crack 500 dataset

3.6 ASPP 与 ASPP-DWT 验证

针对 ASPP 与 ASPP-DWT 再进行分类实验。实验将通过在 CFD 数据集上进行 ASPP 与 ASPP-DWT 模块自身,以及分别融合其他两种模块进行对比,来验证 ASPP 加入 DWT 后,模型检测精度是否提升。

从表 5 可以看出,ASPP-DWT 相比 ASPP 在 mIoU, F1 等值上均有提高,并且从图 9 的可视化图像也能看出,融入 DWT 后,相比原 ASPP,模型对裂缝细节的检测上较好,说明 DWT 在加入模型后,能通过减少卷积和池化过程中的空间信息损失来提升模型总体精度。

表 5 在 CFD 数据集上的对比实验

Table 5 Comparative experiments on CFD dataset

模型	ASPP				ASPP-DWT			
	mIoU	F1	P	R	mIoU	F1	P	R
—(a)	70.86	0.789	0.763	0.817	71.84	0.799	0.775	0.824
C2G(b)	73.42	0.818	0.797	0.840	75.98	0.841	0.824	0.859
MSDWBlock(c)	74.12	0.829	0.816	0.840	74.35	0.821	0.802	0.841
C2G and MSDWBlock(d)	77.19	0.861	0.844	0.879	78.51	0.868	0.851	0.881

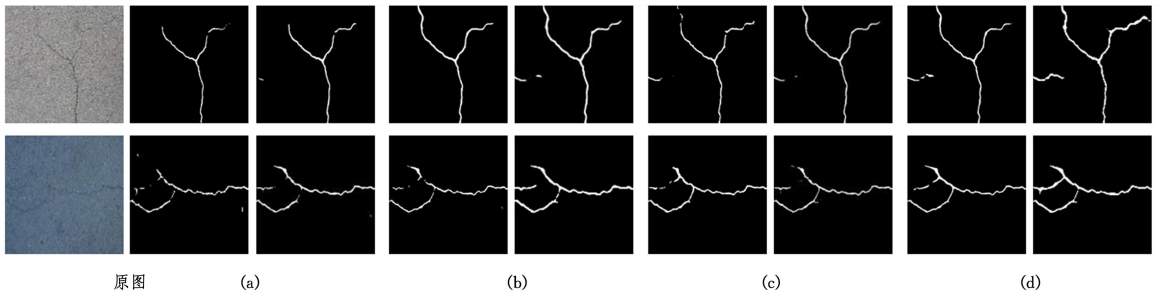


图9 CFD数据集上的对比实验

Fig. 9 Comparative experiments on CFD dataset

结束语 本文提出了一种改进的道路裂缝检测模型,此方法在原始的 U-Net 模型的基础上进行了创新,通过融入 ASPP-DWT 模块、MSDWBblock 模块,以及 C2G 注意力机制,成功提升了模型识别裂缝的性能。ASPP-DWT 模块的引入,使网络能够在卷积和池化的过程中减少位置信息的损失。这一改进使得模型在检测裂缝时,能够保留更多的结构细节,而且在一定程度上抑制了背景噪声的干扰,从而优化了检测结果。

同时加入 MSDWBblock 模块有助于增强感受野,进一步强化了对关键特征的提取,提高了模型的泛化性能。C2G 注意力机制的融入,进一步提升了模型对裂缝细节的感知能力。这一机制通过精细调整特征权重,让网络更加专注于裂缝特征的学习,从而有效地增强了模型在复杂场景中的裂缝检测能力。本文通过一系列消融实验验证了这些改进措施的有效性。

实验结果表明,与原始 U-Net 模型相比,加入 ASPP-DWT 模块的模型在减少位置信息损失方面相比 ASPP 在裂缝检测的精度上有所提升,加入 MSDWBblock 模块的模型在精度上相比原始网络在 mIoU, F1 等指标均有提升,而引入 C2G 注意力机制后的模型则在裂缝特征的细节捕捉上更为精准。

尽管模型在道路裂缝检测任务上取得了一定的成果,但仍存在一些需要改进和探索的地方。例如,本模型在处理一些特殊类型的裂缝(如图 8)时,在部分细节上的检测,离原始图像仍有一些差别。因此,未来的研究方向将致力于改善模型在处理这类复杂裂缝上的性能。此外,也希望进一步优化模型的计算效率,以便于在实际应用。

参考文献

- [1] 2020 Statistical Bulletin on the Development of the Transportation Industry [J]. Communications Finance and Accounting, 2021(6):92-97.
- [2] MA L, LI J. SD-GCN: Saliency-based dilated graph convolution network for pavement crack extraction from 3D point clouds [J]. International Journal of Applied Earth Observation and Geoinformation, 2022, 111: 102836.
- [3] GUPTA P, DIXIT M. Image-based crack detection approaches: a comprehensive survey [J]. Multimedia Tools and Applications, 2022, 81(28): 40181-40229.
- [4] 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.
- [5] ZOU Q, CAO Y, LI Q, et al. CrackTree: Automatic crack detection from pavement images [J]. Pattern Recognition Letters, 2012, 33(3): 227-238.
- [6] LI Q, ZOU Q, ZHANG D, et al. FoSA: F* seed-growing approach for crack-line detection from pavement images [J]. Image and Vision Computing, 2011, 29(12): 861-872.
- [7] DORAFSHAN S, THOMAS R J, MAGUIREM. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete [J]. Construction and Building Materials, 2018, 186: 1031-1045.
- [8] KAMALIARDAKANI M, SUN L, ARDAKANI M K. Sealed-crack detection algorithm using heuristic thresholding approach [J]. Journal of Computing in Civil Engineering, 2016, 30(1): 04014110.
- [9] WANG J, LIU F, YANG W, et al. Pavement crack detection using attention u-net with multiple sources [C] // Pattern Recognition and Computer Vision: Third Chinese Conference (PRCV 2020). Nanjing, China, Part II 3. Springer International Publishing, 2020: 664-672.
- [10] ZHANG L, YANG F, ZHANG Y D, et al. Road crack detection using deep convolutional neural network [C] // 2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016: 3708-3712.
- [11] ZHANG A, WANG K C P, LI B, et al. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network [J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(10): 805-819.
- [12] LI G, MA B, HE S, et al. Automatic tunnel crack detection based on u-net and a convolutional neural network with alternately updated clique [J]. Sensors, 2020, 20(3): 717.
- [13] ZHU S, DU J, LI Y, et al. Method for bridge crack detection based on the U-Net convolutional networks [J]. Journal of Xidian University, 2019, 46(4): 35-42.
- [14] LIU Z, CAO Y, WANG Y, et al. Computer vision-based concrete crack detection using U-net fully convolutional networks [J]. Automation in Construction, 2019, 104: 129-139.
- [15] HUYAN J, LI W, TIGHE S, et al. CrackU-net: A novel deep convolutional neural network for pixelwise pavement crack detection [J]. Structural Control and Health Monitoring, 2020, 27(8): e2551.
- [16] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J].

arXiv:1704.04861,2017.

- [17] GUO T, MOUSAVI H S, VU T H, et al. Deep wavelet prediction for image super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017:104-113.
- [18] BAE W, YOO J, CHULYE J. Beyond deep residual learning for image restoration; Persistent homology-guided manifold simplification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017:145-153.
- [19] LI Q, SHEN L, GUO S, et al. Wavelet integrated CNNs for noise-robust image classification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

2020:7245-7254.

- [20] MNIH V, BADIA A P, MIRZA M, et al. Asynchronous methods for deep reinforcement learning[C]// International Conference on Machine Learning. PMLR, 2016:1928-1937.



LIU Yunqing, born in 1970, professor, Ph.D supervisor. His main research interests include computer vision and radar signal processing and laser communication.