文章编号:1674-2974(2024)06-0063-10 DOI:10.16339/j.cnki.hdxbzkb.2024267

一种道路裂缝检测的变尺度 VS-UNet 模型

赵志宏[1,2†],何朋[2],郝子晔[2]
(1. 石家庄铁道大学 省部共建交通工程结构力学行为与系统安全国家重点实验室,河北 石家庄 050043;
2. 石家庄铁道大学 信息科学与技术学院,河北 石家庄 050043)

摘 要:为解决目前现有的图像分割算法存在检测精度低、对裂缝检测缺乏针对性等问题,采用多尺度特征融合方法,提出一种扩展 LG Block 模块 Extend-LG Block,其由多个并行不同膨胀率的空洞卷积组成.通过参数可调节分支数量和空洞卷积膨胀率,从而改变其感受野大小,进而提取和融合不同尺度的裂缝特征.对比在深层使用多尺度特征融合模块的网络以及使用固定尺度结构进行多尺度特征融合的网络的优劣,提出一种变尺度结构的 UNet 模型 VS-UNet,使用多个不同参数的 Extend-LG Block 替换 UNet 网络中的基本卷积块.该结构在网络浅层进行多尺度特征融合,多尺度特征融合模块提取的尺度随网络层加深逐渐减少.此结构在加强图像的细节特征提取能力的同时保持原有的抽象特征提取能力,还可避免网络参数的增加.在 DeepCrack 数据集以及 CFD 数据集上进行实验验证,结果表明,相较于其他两种结构和方法,提出的变尺度结构的网络在有更高检测精度的同时,在可视化实验对比上对各种大小的裂缝有更好的分割效果.最后与其他图像分割算法进行对比,各项指标与 UNet 相比均有一定程度提升,证明了网络改进的有效性.研究结果可为进一步提升道路裂缝检测效果提供参考.

关键词:U-Net;多尺度;裂缝检测;空洞卷积;深度学习
中图分类号:TP391.41 **文献标志码**:A

A Variable-scale VS-UNet Model for Road Crack Detection

ZHAO Zhihong[1,2†], HE Peng[2], HAO Ziye[2]
(1. State Key Laboratory of Mechanical Behavior and System Safety of Traffic Engineering Structures , Shijiazhuang Tiedao University, Shijiazhuang 050043, China;
2. School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: Existing image segmentation algorithms face challenges related to low detection accuracy and a lack of specificity in crack detection. To address these challenges, this paper proposes an extended LG-Block module Extend-LG Block, which leverages a multi-scale feature fusion method. This module consists of multiple parallel dilated convolutions with different expansion rates. The number of branches and the expansion rate of dilated convolutions can be adjusted by parameters to change the size of its receptive field, and then extract and fuse crack

* **收稿日期**:2023-05-05
基金项目:国家自然科学基金资助项目(11972236, 12172234),National Natural Science Foundation of China(11972236, 12172234)
作者简介:赵志宏(1972—),男,河北石家庄人,石家庄铁道大学教授,博士
† 通信联系人,E-mail:hb_zhaozhihong@126.com

features of different scales. By comparing the advantages and disadvantages of the network using a multi-scale feature fusion module in the deep layer and the network using a fixed scale structure for multi-scale feature fusion, a U-Net model with a variable scale structure named VS-UNet is proposed. The basic convolution Block in the UNet network is replaced by multiple Extend-LG blocks with different parameters. This structure performs multi-scale feature fusion in the shallow layer of the network, and the scale extracted by the multi-scale feature fusion module gradually decreases with the deepening of the network layer. This structure not only strengthens the detail feature extraction ability of the image while maintaining the original abstract feature extraction ability but also avoids the problem of increasing network parameters caused by the increase of convolution. Experiments are carried out on the DeepCrack dataset and CFD dataset. The results show that compared with the other two structures and methods, the proposed network with variable scale structure has higher detection accuracy and better segmentation effect for cracks of various sizes in visual experimental comparison. Finally, compared with other image segmentation algorithms, all indicators are improved to a certain extent compared with UNet, which proves the effectiveness of the improved network.

Key words: U-Net; multi-scale; crack detection; dilated convolution; deep learning

我国道路交通设施的快速发展为人们的出行提供了便利,但同时也出现了很多新的问题.我国道路多为混凝土结构,容易出现裂缝[1],随着时间的推移裂缝会逐渐加大,带来安全隐患[2].因此路面裂缝的及时检测是当前亟须解决的问题.

传统路面裂缝检测多采用人工检测的方法[3],这种方法对于裂缝的检测会比较全面,但同时会消耗较大的人力与物力.基于阈值的分割算法[4]对裂缝的识别率较高,然而在光照不均匀、背景图不一致的情况下识别率较低.基于边缘检测的分割算法[5]对于边缘明显的裂缝检测效果较好,但是对于边缘信息较弱的裂缝以及背景噪点较多的裂缝检测效果较差.传统的图像分割算法对于复杂情况下的图像分割效果仍然不佳[6].

近年来,基于深度学习的裂缝检测方法逐渐成为主流,取得了很好的效果[7].Shelhamer 等[8]提出使用 FCN(Fully Convolutional Networks)模型完成图像分割任务,在图像分割任务上具有不俗的效果.Yang 等[9]基于 FCN 网络完成了裂缝检测任务,实验结果表明 FCN 网络可应用于裂缝检测任务.Badrinarayanan 等[10]在 FCN 的基础上提出的 SegNet 网络使用编码器-解码器的结构,具有良好的分割性能.Ronneberger 等[11]提出的 UNet 网络使用特征拼接的方式将采样前后的特征进行融合,广泛应用在了医疗图像分割领域.

由于混凝土表面具有不平整、光照不均和裂缝背景较为复杂等特点,并且裂缝像素所占比例较小[12],因此上述分割网络在裂缝检测任务上精度不佳.祝一帆等[13]使用密集连接结构改进 U-Net 网络,结合残差块以及扩张卷积提出 Crack UNet 网络,在裂缝检测任务上具有良好的效果.朱立学等[14]提出使用符合 HDC(Hybrid Dilated Convolution)设计原则的多尺度串联空洞卷积组合,有效避免了稀疏的串联空洞组合会出现的网格效应,在香蕉图像分割任务上取得较好的效果.石甜甜等[15]提出一种多尺度融合注意力模块,在特征融合部分增强网络获取不同尺度信息的能力.Chen 等[16]提出空洞卷积空间金字塔池化(ASPP)模块,使用不同扩张率的并行空洞卷积捕获多尺度信息.Song[17]提出一种多尺度的医学图像分割网络 PLU-Net,设计一种新的 LG Block(Local Guided Block)模块,通过并行的二分支空洞卷积结构来增加提取特征的尺度,实验结果表明,采用 LG Block 模块提高了分割的精度,增加了网络的性能.但是 PLU-Net 网络在浅层增加的尺度有限,模型的分割精度还有待提高.本文针对上述裂缝检测模型存在的问题,提出一种改进的道路裂缝检测模型,所做主要贡献有:

1)提出一种多尺度特征提取融合模块 Extend-LG Block 模块.针对裂缝细节特征提取比较困难,

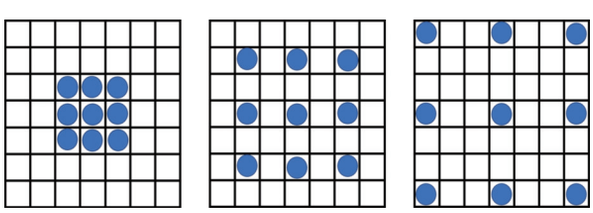
Extend-LG Block 模块可以更好地提取裂缝特征.模块通过参数调节分支数量以及空洞卷积的扩张率,采用 HDC 设计原则,在一定程度上避免了空洞卷积堆叠导致的网格效应,可以提取更丰富的多尺度特征信息.

2)提出一种采用变尺度结构设计的 U-Net 模型.考虑到深度卷积神经网络浅层提取细节特征信息,深层提取抽象特征信息的特点,将固定尺度结构改进为变尺度结构,具体地,网络浅层用多分支的 Extend-LG Block 提取多尺度特征,网络深层用少分支的 Extend-LG Block 提取多尺度特征,在最底层用普通卷积来提取抽象特征.

1　方 法

1.1　空洞卷积

在裂缝检测任务中,往往存在大小不一的各种裂缝,这就要求网络需要从各个尺度提取裂缝中的信息,而大尺度的信息需要大感受野的卷积来提取[18].传统卷积一般只能通过增加卷积核的大小或者增加网络的深度来增大感受野,这会增加网络的参数量.YU 等[19]提出的空洞卷积结构解决了该问题.空洞卷积如图 1 所示,空洞卷积在保持参数数量不变的情况下增大卷积核的感受野,使每个卷积输出都包含较大范围的信息;同时可以保证输出的特征映射的大小保持不变.



(a)膨胀率为 1　(b)膨胀率为 2　(c)膨胀率为 3

图 1　不同膨胀率的空洞卷积

Fig.1　Dilated convolutions with different dilation rates

传统卷积与空洞卷积的卷积核大小的对应关系如下:

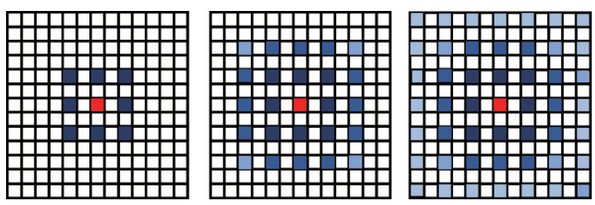
$$k' = d \times (k - 1) + 1 \tag{1}$$

空洞卷积感受野的计算公式如下:

$$F = 2(d - 1) \times (k - 1) + k \tag{2}$$

式中:k 表示输入的卷积核尺寸;d 表示扩张系数;k' 表示扩张后等效的卷积核尺寸.假设扩张系数为 7,卷积核尺寸为 3×3,通过上式(1)~(2)可以计算发现扩张之后等效的卷积核尺寸为 15.可以看到通过扩张卷积仅需要与 3×3 的普通卷积同样的参数量就可以获得与 15×15 的普通卷积相同的感受野大小.

空洞卷积在有效增加网络感受野的同时,也存在网格效应(the Gridding Effect)问题.网格效应如图 2 所示,网格效应会导致输入信号当中的一些像素得不到利用,即局部信息会丢失.



(a)膨胀率为 1　(b)膨胀率为 2　(c)膨胀率为 3

图 2　网格效应

Fig.2　The Gridding Effect

Wang 等[20]针对空洞卷积存在的网格效应问题,提出 HDC 设计结构.基于 HDC 结构的空洞卷积组合可以有效减少网格效应的产生.HDC 首先定义两个非零元素之间最大的距离公式,该公式如下所示:

$$M_i = \mathrm{Max}\left[M_{i+1} - 2r_i, M_{i+1} - 2\left(M_{i+1} - r_i\right), r_i\right] \tag{3}$$

式中:M_i 表示第 i 层两个非零元素之间的最大距离;r_i 表示第 i 层的膨胀系数.HDC 结构要求:1)假设卷积核大小为 $k \times k$,则 $M_2 \leqslant k$,即第二层的两个非零元素之间的最大距离小于或等于该层卷积核的大小.2)空洞卷积组合的扩张系数应为锯齿形状,如[1,3,5,1,3,5].3)空洞卷积组合的公约数不能大于 1.采用上述设计原则的空洞卷积组合可以有效减少网格效应的发生.

1.2　Extend-LG Block

为进一步提升网络对于裂缝特征的提取能力,设计一种多分支卷积结构对 UNet 模型的基本卷积块进行替换.

UNet 网络采用的基本卷积块如图 3 所示,由两个 3×3 普通卷积组成.这种结构的卷积块缺乏提取图像不同尺度特征的能力,不能同时很好地提取大裂缝与小裂缝的特征.

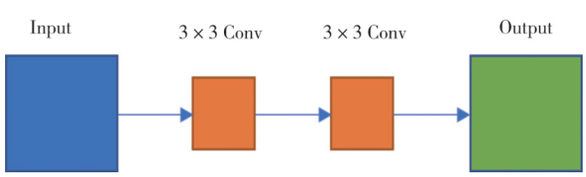


图 3　UNet 网络的基本卷积块

Fig.3　The basic convolutional block of the UNet network

针对UNet基本卷积块存在的问题,LG Block使用两个并行的不同扩张率的空洞卷积,其结构如图4所示,通过不同扩张率的空洞卷积提取不同尺度的特征.与UNet网络的基本卷积块相比,这种结构的卷积块拥有提取多尺度特征的能力,并且与其他采用串行结构的空洞卷积块相比,由于其设计符合HDC结构,可以很好地避免网格效应问题.

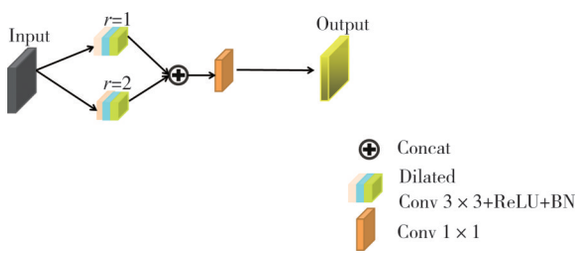


图4 LG Block结构

Fig.4 LG Block structure

LG Block由两个分支组成,每个分支有一个3×3的空洞卷积,其扩张率分别为1和3,然后将两个扩张卷积操作的结果连接起来,以增强特征传播.然后,采用1×1卷积运算,在不改变特征图大小的情况下,加入非线性特征,实现各个不同尺度特征的融合.

LG Block仅增加一个膨胀率为3的空洞卷积,虽然扩大了感受野,但增加的尺度有限,在浅层网络中提取到的特征仍不够丰富,故需要在浅层增加提取的尺度.而受制于UNet网络中特征图尺寸大小随着池化层减少的特点,在网络深层无须增加大感受野卷积来提取大尺度信息.因此网络需要多种不同感受野的卷积块调节网络不同层的提取特征的尺度.针对上述问题,对LG Block进行扩展,扩展之后的模块Extend-LG Block的结构如图5所示.

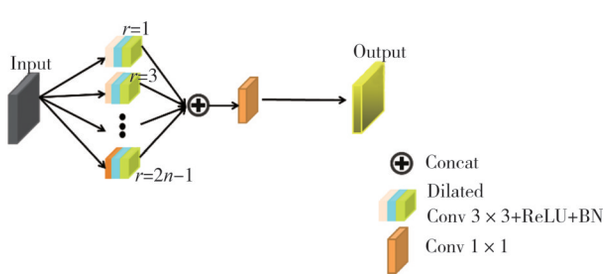


图5 Extend-LG Block结构

Fig.5 Extend-LG Block structure

扩展之后的Extend-LG Block与原有固定的两分支不同,可通过参数增减分支,网络增加的扩张卷积的膨胀系数为$2n-1$.特别地,LG Block是n=2时的Extend-LG Block. Extend-LG Block对于输入图像的处理过程与LG-Block类似,可描述为:

$$\begin{cases} F_1 = \sigma(\mathrm{BN}(f_{3\times3}^{1}(I))) \\ F_2 = \sigma(\mathrm{BN}(f_{3\times3}^{3}(I))) \\ \cdots \\ F_n = \sigma(\mathrm{BN}(f_{3\times3}^{2n-1}(I))) \\ F_o = f_{1\times1}(\mathrm{Cat}(F_1, F_2, \cdots, F_n)) \end{cases} \tag{4}$$

式中:I表示输入的图像;F_n表示经过扩张率为n的空洞卷积提取到的特征;$f_{k\times k}^{r}(\cdot)$表示扩张率为r的,卷积核大小为k的空洞卷积操作;$\sigma(\cdot)$表示ReLu激活函数;Cat$(\cdot)$表示特征拼接操作;BN$(\cdot)$表示归一化操作;$f_{1\times1}(\cdot)$表示一个卷积核大小为1的普通卷积操作;F_o表示Extend-LG Block模块输出的特征.

为防止使用多个扩张卷积导致网络出现网格效应,Extend-LG Block的组合采用HDC设计结构.网络的扩张系数设计为r=[1,3,5,⋯,2n-1],符合锯齿形状的扩张系数的要求,其次,扩张系数的最大公约数不大于1.需证明网络的M_2≤3,证明过程如下:

根据定义可知

$$M_n = r_n = 2n - 1$$

$$M_{n-1} = \mathrm{Max}\left[M_n - 2r_n, M_n - 2\left(M_n - r_n\right), r_n\right] = \mathrm{Max}\left[5 - 2n, 2n - 5, 2n - 3\right]$$

当n≤2时,即r=[1,3]或者r=[1]满足设计要求.

所以$M_{n-1} = 2n - 3$.

以下讨论n≥3的情况.

$$M_{n-2} = \mathrm{Max}\left[M_{n-1} - 2r_{n-2}, M_{n-1} - 2\left(M_{n-1} - r_{n-2}\right), r_{n-2}\right] = \mathrm{Max}\left[7 - 2n, 2n - 7, 2n - 5\right]$$

又因为n≥3,所以$M_{n-2} = 2n - 5$.

由上述可知:$M_{n-i} = 2n - 2i - 1$.

当i=n-2时,$M_2 = 2n - 2i - 1$.

$M_2 = 3 \leqslant k = 3$.

故扩张系数组合符合HDC设计结构.

为验证不同取值的Extend-LG Block提取的效果,将n=2、3、4时Extend-LG Block提取的特征图进行对比,不同n取值的Extend-LG Block提取的特征图如图6所示.

通过不同n取值的Extend-LG Block提取的特征图对比,可以发现n取值越大,Extend-LG Block对于裂缝的边缘信息提取效果越好,同时对于一些背景噪声的过滤效果也得到加强.综上所述,扩展之后的结构相较于LG Block能够提取更多尺度特征,对于各种裂缝的分割均具有较好的效果.

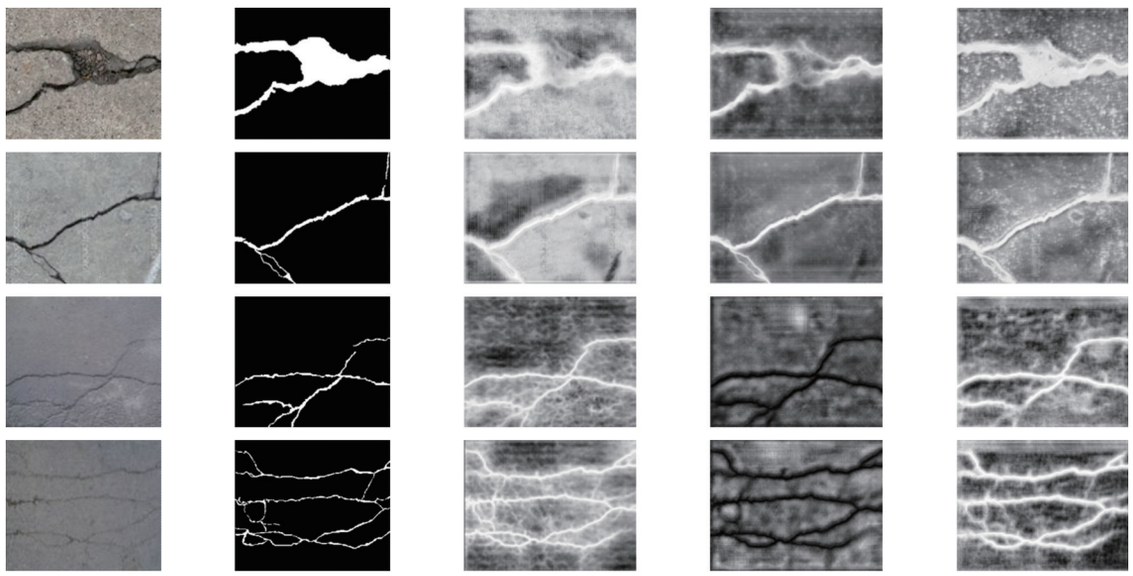


图 6　不同 n 取值的 Extend-LG Block 提取特征图对比

Fig.6　Comparison of feature maps extracted from Extend-LG Block with different n values

1.3　VS-UNet

UNet 起初应用于医疗图像分割领域.近几年有学者将其引入裂缝检测领域.UNet 结构如图 7 所示,其采用下采样(编码器),上采样(解码器),跳跃连接的设计方法.浅层可以提取图像细节特征,如边界、颜色、纹理等,深层会提取图像抽象特征[21].UNet 采用跳跃连接的方法在特征传播阶段进行多尺度特征融合,缺乏在特征提取阶段进行多尺度特征融合的能力.

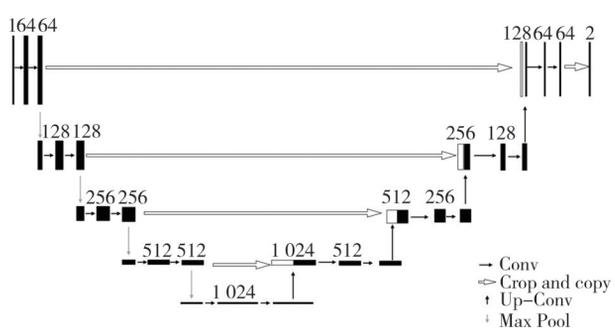


图 7　UNet 结构图

Fig.7　Diagram of the UNet structure

为了进一步提高 UNet 网络的多尺度特征提取能力,李国燕等[21]在网络的编码器与解码器之间加入并行多尺度特征融合模块,如 ASPP 模块,增强网络多尺度特征提取能力,其结构如图 8 所示.在编码器与解码器之间加入多尺度特征融合模块虽然可以增强网络对于不同大小裂缝的提取能力,但是其放置位置在网络的深层中,仅针对网络提取的抽象特征进行多尺度特征融合,浅层提取的细节信息经过最大池化层会有一定损失,使得这种模块不能很好地融合网络浅层的特征.

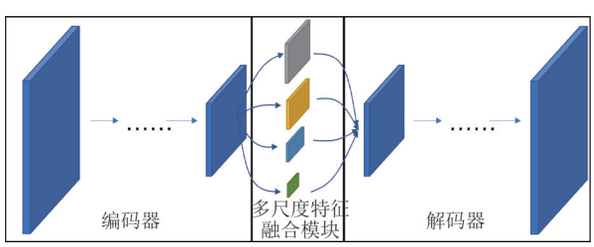


图 8　加入多尺度特征融合模块的网络

Fig.8　Network with multi-scale feature fusion module

越来越多的研究者使用多尺度特征融合模块作为网络的基本网络单元替换 UNet 网络卷积块的方法,来提升网络多尺度特征提取的能力.Song[17]使用固定尺度结构提取特征,如图 9 所示.该结构在网络的每层都使用相同大小和扩张率的扩张卷积块提取特征,这会同时保持网络对于浅层特征以及深层特征的多尺度特征提取能力,并且由于 UNet 下采样机制,深层网络拥有较大的感受野可提取大尺度信息,但是会导致网络过于关注全局信息,从而影响小裂缝的分割效果,除此之外该结构由于增加卷积数量会导致网络参数量增加.

本文在使用多尺度特征融合模块作为网络的基本卷积块替换原有网络的基本卷积块的基础上,将固定尺度的多尺度特征融合模块替换为可变尺度的

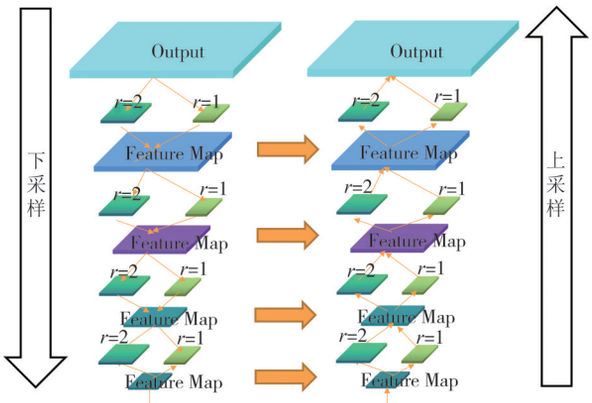


图 9 采用固定尺度结构设计的编码器与解码器

Fig. 9 The encoder and decoder designed with fixed scale structure

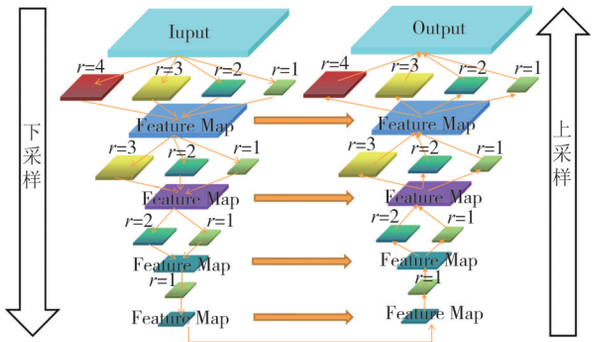


图 10 采用变尺度结构设计的编码器与解码器

Fig.10 The encoder and decoder designed with variable scale structure

多尺度特征融合模块,进而提出一种变尺度的多尺度特征融合结构,如图 10 所示.具体表现为在浅层使用大尺度卷积块提取细节特征,在深层使用小尺度卷积提取抽象特征.这种变尺度结构可以提升网络对于细节特征的提取能力,保持网络对于抽象特征的提取能力,在避免网络参数量增加的情况下,仍保持有较高的精度.除此之外,同时兼顾提取到的全局信息与局部信息,避免网络偏向全局信息而影响小裂缝的检测.

本文设计一种采用变尺度结构的多尺度裂缝检测网络 VS-UNet(Variable Scale-UNet),如图 11 所示,由 n=4,3,2,1 的 Extend-LG Block 组合而成,n 根据下采样的深度依次递减.在浅层使用较大的数值 n 可以提取到更多尺度的细节特征,在深层使用较小数值的 n 可以避免网络尺度过大导致的对细小裂缝分割能力减弱的问题,还会减少网络的参数量增加.值得注意的是为保持深层网络提取抽象特征的能力,在最低层保持了 UNet 中的两个 3×3 的普通卷积.上采样的过程与下采样相反,Extend-LG Block 依据网络的深度 n 依次增加.

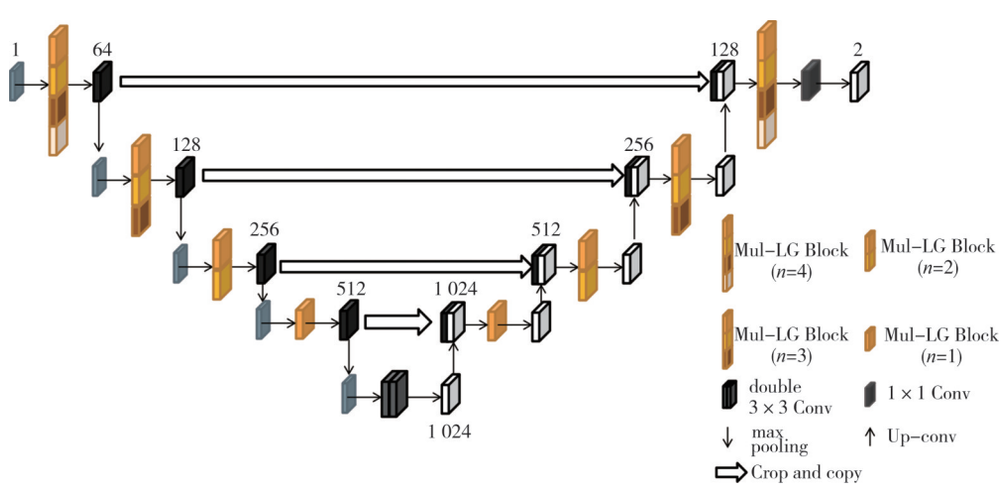


图 11 VS-UNet 网络结构模型图

Fig.11 Diagram of the VS-UNet network structure model

使用上述的结构,网络一定程度上增加了每层的宽度,相较于 UNet 层数没有增加,还可以避免因层数增加而导致的梯度消失问题.

VS-UNet 网络主要分为编码器与解码器两个部分.分别为 4 次下采样与 4 次上采样.每次下采样前特征图会经过不同参数 n 的 Extend-LG Block 模块,此时模块会进一步提取出 n 个尺度的特征进行融合,此外会使特征图的通道数翻倍.使用最大池化层对模型进行下采样,采样之后特征图的大小会缩小到原来的 1/2.解码器部分与编码器部分相反,特征图先由转置卷积进行上采样,之后与下采样提取到的相同大小特征图进行拼接操作,融合浅层提取到的特征,再由 Extend-LG Block 模块进一步进行特征提取.最后由一个 1×1 的卷积对像素进行预测.

由于网络输出的结果为单个像素的概率预测,故需要对网络预测结果进行后处理.本文使用阈值来进行分割结果的二值化[23],即通过设定的阈值来划分目标对象(白色)和背景对象(黑色),从而得到最终的分割图像.

2　实验与分析

2.1　数据集与参数设置

为验证改进的有效性,实验采用 DeepCrack 数据集[24]以及 CFD(Crack Forest Dataset)数据集[25].其中,DeepCrack 数据集包含 573 张大小为 544×384 的混凝土路面图像,CFD 数据集包含 118 张分辨率为 480×320 的已标注裂缝图片.实验中以 8∶2 的比例划分训练集与测试集,同时在训练集中再以 8∶2 的比例划分出验证集.考虑到 CFD 数据集数据量较少,使用翻转、镜像等数据增强技术进行扩充.

实验环境使用 Windows11 操作系统,运行内存 16GB,处理器为 Intel Core I5 8400,GPU 为 NVIDIA GTX1060,编程语言为 Python3.8,模型基于 Pytorch1.13.1 深度学习框架实现.

训练过程中选用 Adam 优化器作为模型的优化器,损失函数为均方误差损失函数,BatchSize 大小设置为 2.学习率设置为 3×10^{-3}.训练的轮数为 50 轮.为避免网络陷入局部最优解,实验采用多组随机初始化参数.使用阈值 0.5 对预测结果进行二值化处理.

2.2　评价指标

实验采用查准率 P(Precision)、查全率 R(Recall)、F_1(F_1_Score)作为评价指标.P 是正确检测的裂缝像素与原本正确像素的比值.R 表示正确检测的裂缝像素占所有裂缝像素的百分比.F_1(F_1_Score)表示同时衡量 P 和 R 的值,尽可能保证在裂缝像素遗漏少的前提下误判也少.以上几个评价指标均与模型的性能成正比.上述指标计算公式如下所示:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{5}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

$$F_1 = \frac{2 \times P \times R}{P + R} \tag{7}$$

式中:TP(True Positive)表示被正确分为裂缝像素的数量;FP(False Positive)为错误分为裂缝像素的数量;FN(False Negative)表示目标被错误分割为背景的像素数量.

2.3　Extend-LG Block 不同放置位置实验

为验证 Extend-LG Block 放置位置对网络的影响,分别设计仅放置在解码器部分的网络模型、仅放置在编码器部分的网络模型以及同时放置在编码器与解码器部分的网络模型,实验结果如表 1 所示.

表 1　Extend-LG Block 放置位置对比结果

Tab.1　Comparison results of Extend-LG Block placement positions

数据集	下采样	上采样	P/%	R/%	F_1/%
DeepCrack	√		85.12	74.18	79.27
		√	78.51	79.93	79.21
	√	√	88.06	78.02	82.74
CFD	√		73.28	59.09	65.43
		√	77.99	58.36	66.76
	√	√	82.71	64.83	72.68

在 DeepCrack 数据集中,放置在上采样中的 P 值比放置在下采样过程中的 P 值低 6.61%,而放置在上采样中的 R 值比放置在下采样过程中的 R 值高 5.75%.在 F_1 指标对比上,下采样比上采样高 0.06%.在 CFD 数据集中,放置在上采样中的 P 值比放置在下采样过程中的 P 值高 4.71%,而放置在上采样中的 R 值比放置在下采样过程中的 R 值低 0.73%.在 F_1 指标的对比上,上采样比下采样高出 1.33%.

根据上述分析可知同时放置在下采样与上采样过程中的 F_1 指标在两个数据集上均优于放置在单一位置上的 F_1 指标,故同时放置的效果比任意单一放置的效果更好.

2.4　Extend-LG Block 不同参数取值对网络性能的影响

为验证使用不同尺度的 Extend-LG Block 对网络的影响,分别对只采用 n=4、3、2 的 Extend-LG Block 的 U-Net 网络以及 VS-UNet 进行实验,实验结果如表 2 所示.

从表中可以看到,在 DeepCrack 数据集中,n=3 时 Extend-LG Block 的 F_1 指标与 R 指标最高,F_1 指标比 n=4 时高出 0.64%,参数量低 8.52M,F1 指标比 n=2 时高出 2.08%,比 VS-UNet 高出 1.52%.通过 n=4 与 n=3 的对比实验发现,提升 n 的取值并不一定会取得更好的效果,反而会增加参数量.VS-UNet 相比于 n=2 的 Extend-LG Block,在 F1 指标上高出 0.56%,虽然在 F_1 指标上略低于 n=4,n=3 的情况,但是参数量明显低于其他几种方法.在 CFD 数据集中可以发现,VS-UNet 的 F_1 指标比 n=4 的 Extend-LG Block 高出 2.43% 的 Extend-LG Block,比 n=3 高出 1.5%,比 n=

2 高出 2.43%. 注意到 Extend-LG Block n=3 模型在 DeepCrack 上优于 VS-UNet 而 VS-UNet 在 CFD 数据集上优于 Extend-LG Block,这是因为 CFD 数据集多为细小裂缝,深层网络感受野的过度增加使得网络过度关注于全局特征,从而对小裂缝的分割效果产生了负作用.

表 2 使用不同参数的 Extend-LG Block 实验对比结果

Tab.2 Comparison results of Extend-LG Block experiments with different parameters

数据集	方法	n	P/%	R/%	F_1/%	Parameters/M
DeepCrack	Extend-LG Block	4	88.58	79.19	83.62	51.03
	Extend-LG Block	3	87.95	80.87	84.26	42.51
	Extend-LG Block	2	87.61	77.38	82.18	33.99
	VS-UNet	—	88.06	78.02	82.74	28.13
CFD	Extend-LG Block	4	78.14	63.81	70.25	51.03
	Extend-LG Block	3	78.29	65.25	71.18	42.51
	Extend-LG Block	2	81.86	61.60	70.30	33.99
	VS-UNet	—	82.71	64.83	72.68	28.13

上述结果表明,增加 Extend-LG Block 的尺度可以提升网络的特征提取效果,但过多增加尺度会使得网络偏向全局特征,降低了网络的分割效果并增加网络参数量,故网络的整体尺度不应过大.同时,与采用固定尺度的 Extend-LG Block 的网络相比,变尺度结构的网络在有更小参数量的同时,对于细小裂缝数据集分割效果更好.

2.5 采用不同多尺度特征提取融合方法的网络对比

为验证本文提出的变尺度结构的有效性,本节分别设计了两个网络与 VS-UNet 进行对比.ASPP-UNet 为在编码器与解码器之间加入并行多尺度特征融合模块的 UNet 网络,将其编码器与解码器之间的普通卷积块替换为 ASPP 模块.FS-UNet(Fixed scale-UNet)为采用固定尺度结构的 UNet 网络,将其基本卷积块全部替换为 LG Block.

采用不同多尺度特征提取方法的网络对比结果如表 3 所示,在 P 指标和 R 指标的对比上,3 种多尺度网络结构在不同数据集上各有优劣,而在综合 P 指标与 R 指标的 F_1 指标的对比上,本文提出的 VS-UNet 在两个数据集中均取得三者之中最高值.在 DeepCrack 数据集中,VS-UNet 的 F_1 指标比 ASPP-UNet 与 FS-UNet 分别高出 3.68%,2.34%. 在 CFD 数据集中,VS-UNet 的 F_1 指标比 ASPP-UNet 与 FS-UNet 分别高出 1.20%,3.51%. 在参数的对比上,VS-UNet 比 ASPP-UNet 高 1.02 M,比 FS-UNet 低 3.24 M.

本文提出的变尺度结构在裂缝检测的效果上优于其他两种方法,在参数的对比上接近使用 ASPP 模块的 UNet 网络,优于使用固定尺度结构的 UNet 网络.实验结果表明了本文提出的变尺度结构的有效性.

表 3 采用不同多尺度特征提取方法的网络对比结果

Tab.3 Comparison results using different Multi-Scale feature extraction methods

数据集	模型	P/%	R/%	F_1/%	Parameters/M
DeepCrack	ASPP-UNet	90.06	70.45	79.06	27.11
	FS-UNet	91.37	71.78	80.40	31.37
	VS-UNet	88.06	78.02	82.74	28.13
CFD	ASPP-UNet	78.02	65.95	71.48	27.11
	FS-UNet	80.72	60.51	69.17	31.37
	VS-UNet	82.71	64.83	72.68	28.13

2.6 可视化分析

网络分割结果的可视化是验证改进算法有效性的途径之一,本节对 UNet、ASPP-UNet、FS-UNet 以及 VS-UNet 分别进行实验,将实验结果进行可视化对比.

在 DeepCrack 数据集以及 CFD 数据集上的可视化对比结果如图 12 所示,通过与其他算法对比可以发现,采用多尺度特征提取方法设计的 UNet 网络在裂缝识别的完整性上均优于 UNet 网络.在第一列中,ASPP-UNet 与 VS-UNet 均完整分割出裂缝.在第二列中,ASPP-UNet 表现了较高的误检率,FS-UNet 在裂缝边缘的分割上表现较差.在 CFD 数据集中,VS-UNet 对于细小裂缝分割效果优于其他两种方法.

通过上述分析可知,VS-UNet 对于大小裂缝均具有较好的分割效果,对于裂缝分割的完整性上优于其他几种方法.

2.7 与其他网络模型对比

为进一步验证本文网络的优越性,与 SegNet 网络、DeepCrack 网络[23]、RCF 网络[26]、UNet 网络进行对比,实验结果如表 4 所示.在 DeepCrack 数据集中,本文提出的 VS-UNet 相较于上述几个网络,在 F_1 指标上分别提高 5.7%、28.2%、17.5%、15.4%. 在 CFD 数据集的对比之中,VS-UNet 相较于上述几个网络在 F1 指标上分别提高 23.2%、19.7%、27.5%、17.8%.

根据上述对比可以证明本文提出 VS-UNet 网络在裂缝检测上均优于其他几个同类型网络.

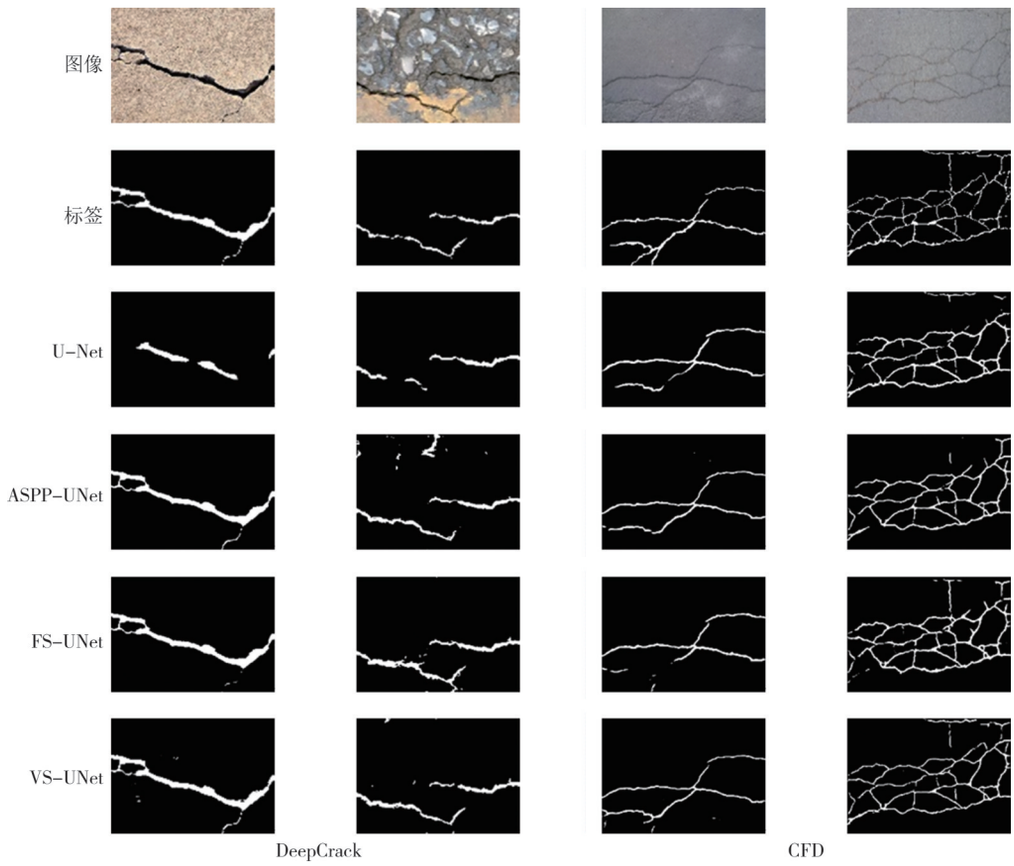


图 12　可视化分析对比结果

Fig. 12　Visual analysis and comparison results

表 4　与其他网络模型对比

Tab.4　Comparison with other network models

数据集	模型	P/%	R/%	F_1/%
DeepCrack	SegNet[6]	73.2	81.2	77.0
	DeepCrack[23]	53.5	55.5	54.5
	RCF[26]	60.1	71.3	65.2
	UNet[6]	64.2	70.6	67.3
	VS-UNet	88.0	78.0	82.7
CFD	SegNet[6]	42.0	60.2	49.5
	DeepCrack[23]	46.7	61.5	53.0
	RCF[26]	41.5	49.5	45.2
	UNet[6]	61.7	54.9	54.9
	VS-UNet	82.7	64.8	72.7

3　结 论

本文针对目前图像分割网络存在的问题,对 UNet 网络结构进行改进,得出以下的结论:

1)本文首先提出一种扩展的 LG-Block 模块,使用多分支结构代替原有的二分支结构,增大模型的感受野,增加特征提取的尺度.其次提出一种变尺度结构的 U-Net 模型,该模型由多个增强后的 LG-Block 模块组成,相较于 U-Net 拥有更大的感受野,对裂缝的分割效果更好.

2)采用变尺度结构可有效地提取裂缝特征.在可视化对比实验中,VS-UNet 可有效地分割出大小不一的裂缝.与其他几个网络模型进行对比,结果表明 VS-UNet 拥有更高的 F1-Score,其在 DeepCrack 数据集以及 CFD 数据集上分别较 UNet 提升 15.4%,证明其对于裂缝的分割效果更好.

目前注意力机制使用相对广泛,可以提高网络的分割效果,可以作为网络改进的下一步研究方向.

参考文献

[1]　蒋文波,罗秋容,张晓华. 基于数字图像的混凝土道路裂缝检测方法综述[J]. 西华大学学报(自然科学版),2018,37(1):75-84.
JIANG W B,LUO Q R,ZHANG X H. A review of concrete roads crack detection methods based on digital image[J]. Journal of Xihua University (Natural Science Edition),2018,37(1):75-84. (in Chinese)

[2] 史梦圆,高俊钗. 改进U-Net算法的路面裂缝检测研究[J]. 自动化与仪表,2022,37(10):52-55.
SHI M Y,GAO J C. Research on pavement crack detection based on improved U-Net algorithm[J]. Automation & Instrumentation,2022,37(10):52-55.(in Chinese)

[3] 邓露,褚鸿鹄,龙砺芝,等. 基于深度学习的土木基础设施裂缝检测综述[J]. 中国公路学报,2023,36(2):1-21.
DENG L,CHU H H,LONG L Z,et al. Review of deep learning-based crack detection for civil infrastructures[J]. China Journal of Highway and Transport,2023,36(2):1-21.(in Chinese)

[4] KIRSCHKE K R,VELINSKY S A. Histogram-based approach for automated pavement-crack sensing[J]. Journal of Transportation Engineering,1992,118(5):700-710.

[5] 张爱华,余胜生,周敬利. 一种基于边缘检测的局部阈值分割算法[J]. 小型微型计算机系统,2003,24(4):661-663.
ZHANG A H,YU S S,ZHOU J L. A local-threshold segment algorithm based on edge-detection [J]. Mini-micro Systems,2003,24(4):661-663.(in Chinese)

[6] ZHANG L X,SHEN J K,ZHU B J. A research on an improved Unet-based concrete crack detection algorithm [J]. Structural Health Monitoring,2021,20(4):1864-1879.

[7] 黄鹏,郑淇,梁超. 图像分割方法综述[J]. 武汉大学学报(理学版),2020,66(6):519-531.
HUANG P,ZHENG Q,LIANG C. Overview of image segmentation methods[J]. Journal of Wuhan University (Natural Science Edition),2020,66(6):519-531.(in Chinese)

[8] SHELHAMER E,LONG J,DARRELL T. Fully convolutional networks for semantic segmentation [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE,2017:640-651.

[9] YANG X C,LI H,YU Y T,et al. Automatic pixel-level crack detection and measurement using fully convolutional network[J]. Computer-Aided Civil and Infrastructure Engineering,2018,33(12):1090-1109.

[10] 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.

[11] RONNEBERGER O,FISCHER P,BROX T. U-Net:convolutional networks for biomedical image segmentation[M]//Lecture Notes in Computer Science. Cham:Springer International Publishing,2015:234-241.

[12] 甘霖,谢爱荣,燕阳,等. 基于改进U-Net网络的混凝土表面裂缝分割[J]. 重庆邮电大学学报(自然科学版),2021,33(4):645-652.
GAN L,XIE A R,YAN Y,et al. Crack segmentation of concrete surface based on improved U-Net [J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition),2021,33(4):645-652.(in Chinese)

[13] 祝一帆,王海涛,李可,等. 一种高精度路面裂缝检测网络结构:Crack U-Net[J]. 计算机科学,2022,49(1):204-211.
ZHU Y F,WANG H T,LI K,et al. Crack U-Net:towards high quality pavement crack detection[J]. Computer Science,2022,49(1):204-211.(in Chinese)

[14] 朱立学,伍荣达,付根平,等. 基于多尺度串联空洞卷积的轻量化UNet香蕉图像分割[J]. 农业工程学报,2022,38(13):194-201.
ZHU L X,WU R D,FU G P,et al. Segmenting banana images using the lightweight UNet of multi-scale serial dilated convolution [J]. Transactions of the Chinese Society of Agricultural Engineering,2022,38(13):194-201.(in Chinese)

[15] 石甜甜,郭中华,闫翔,等. 基于多尺度融合注意力改进UNet的遥感图像水体分割[J]. 液晶与显示,2023,38(3):397-408.
SHI T T,GUO Z H,YAN X,et al. Water body segmentation in remote sensing images based on multi-scale fusion attention module improved UNet[J]. Chinese Journal of Liquid Crystals and Displays,2023,38(3):397-408.(in Chinese)

[16] CHEN L C,PAPANDREOU G,SCHROFF F,et al. Rethinking atrous convolution for semantic image segmentation [J]. arXiv prepriut arXiv:1706.05587,2017.

[17] SONG W. PLU-Net: Extraction of multi-scale feature fusion[J]. 2023.

[18] ARAUJO A,NORRIS W,SIM J. Computing receptive fields of convolutional neural networks[J]. Distill,2019,4(11):e21.

[19] VAN HASSELT H,GUEZ A,SILVER D. Deep reinforcement learning with double Q-learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence,2016,30(1):129-144.

[20] WANG P Q,CHEN P F,YUAN Y,et al. Understanding convolution for semantic segmentation [C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe,NV,USA: IEEE,2018:1451-1460.

[21] 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.(in Chinese)

[22] 许薛军,张肖宁. 基于数字图像的混凝土桥梁裂缝检测技术[J]. 湖南大学学报(自然科学版),2013,40(7):34-40.
XU X J,ZHANG X N. Crack detection of concrete bridges based digital image [J]. Journal of Hunan University (Natural Sciences),2013,40(7):34-40.(in Chinese)

[23] LIU Y H,YAO J,LU X H,et al. DeepCrack:a deep hierarchical feature learning architecture for crack segmentation [J]. Neurocomputing,2019,338(03):139-153.

[24] SHI Y,CUI L M,QI Z Q,et al. Automatic road crack detection using random structured forests [J]. IEEE Transactions on Intelligent Transportation Systems,2016,17(12):3434-3445.

[25] LIU Y,CHENG M M,HU X W,et al. Richer convolutional features for edge detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu,HI,USA: IEEE,2017:5872-5881.

[26] JING P,YU H Y,HUA Z H,et al. Road crack detection using deep neural network based on attention mechanism and residual structure[J]. IEEE Access,2022,11:919-929.