

文章编号 1004-924X(2024)02-0286-15

联合线性引导与网格优化的混凝土裂缝分割

刘光辉^{1,2*}, 陈 健^{1,2}, 孟月波^{1,2}, 徐胜军^{1,3}

- (1. 西安建筑科技大学 信息与控制工程学院, 陕西 西安 710055;
2. 建筑机器人陕西省高等学校重点实验室, 陕西 西安 710055;
3. 西安市建筑制造智能化技术重点实验室, 陕西 西安 710055)

摘要:针对混凝土表面裂缝分割过程中分割精度低、细微裂缝漏分和背景干扰等问题,提出一种联合线性引导和网格优化的裂缝分割模型。首先,在主干网络中引入多分支线性引导模块,通过自适应单维度池化增强网络对裂缝线性结构的表达能力,让不同区域的裂缝建立联系,增强全局上下文信息感知能力,提高网络分割精度;然后,提出网格细节优化模块,通过分区-优化-合并三步骤,将整个空间域划分为若干个空间网格,提取空间网格中的细微裂缝信息,防止细微裂缝漏分;最后,在主干网络的跳跃连接处嵌入混合注意力模块,在空间和通道双维度突出裂缝特征,减少背景干扰。在 Deepcrack537, Crack500 和 CFD 裂缝数据集上,所提模型的 IoU 值分别达到 77.07%, 58.96% 和 56.55%, F1-score 值分别达到 87.05%, 74.19% 和 72.24%, 明显优于大多数现有方法,具有更高的分割精度。

关键词:裂缝图像;线性引导;语义分割;网格优化;注意力机制

中图分类号: TP391 **文献标识码:** A **doi:** 10.37188/OPE.20243202.0286

Concrete crack segmentation combined with linear guidance and mesh optimization

LIU Guanghui^{1,2*}, CHEN Jian^{1,2}, MENG Yuebo^{1,2}, XU Shengjun^{1,3}

- (1. College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China;
2. Higher Education Key Laboratory of Construction Robot in Shaanxi Province, Xi'an 710055, China;
3. Xi'an Key Laboratory of Intelligent Technology for Building and Manufacturing, Xi'an 710055, China)

* Corresponding author, E-mail: guanghuil@163.com

Abstract: A model was proposed to address issues with low segmentation accuracy, leakage of tiny cracks, and background interference in the segmentation process of concrete surface cracks. The model combined linear guidance and mesh optimization for crack segmentation. Firstly, the backbone network was enriched with a multi-branch linear guidance module. The network's ability to represent the linear structure of cracks was boosted by adaptive single-dimensional pooling. This facilitated the establishment of connections between cracks in different areas, enhanced the capability to perceive global context data, and improved the network's segmentation accuracy. Then, a module for mesh detail optimization is pro-

收稿日期: 2023-06-02; 修订日期: 2023-07-13.

基金项目: 陕西省重点研发计划项目 (No. 2021SF-429); 陕西省自然科学基金基础研究计划项目 (No. 2023-JC-YB-532)

posed, which divides the entire spatial domain into several spatial meshes through the three steps of partitioning, optimization, and merging. The fine cracks' information in the spatial meshes was extracted to prevent the leakage of fine cracks. Finally, a mixed attention module was embedded in the skip connections of the backbone network, highlighting crack features in the two-dimensional space and channels while also reducing background interference. On the Deepcrack537, Crack500, and CFD crack datasets, the proposed model achieves IoU values of 77.07%, 58.96%, and 56.55%, respectively. The F1-score values also performs well, achieving 87.05%, 74.19%, and 72.24%, respectively. These results are significantly better than those of most existing methods, with superior segmentation accuracy.

Key words: crack image; linear guide; semantic segmentation; mesh optimization; attention mechanism

1 引言

裂缝作为混凝土表面常见的病害之一,过大的裂缝会降低混凝土工程结构的承载力、防水性能以及耐久性。如果混凝土表面的裂缝不能及时发现修补,可能会导致严重的安全事故发生。因此,定期对混凝土建筑表面进行裂缝检测至关重要。早期的裂缝检测依靠经验丰富的工作人员进行人工检测,速度慢、成本高,存在主观臆断性,无法满足现代建筑健康监测的需求。随着计算机视觉技术的日益发展,研究人员开始将其应用于裂缝检测中,提出大量的裂缝检测方法。这些方法主要分为两类:一种是基于传统方法的裂缝分割,另一种是基于深度学习的裂缝分割。

传统的裂缝分割方法通过人工或机器学习方法获取浅层特征来实现裂缝分割。文献[1]提出一种结合自适应Canny算法和迭代阈值分割算法的裂缝分割方法,使用自适应高斯滤波器平滑图像,并采用迭代阈值分割算法计算图像二值化阈值。文献[2]提出基于邻近差分直方图(NDHM)的裂缝图像阈值分割算法,通过构造目标像素与周围像素之间差异最大化的目标函数,形成裂缝图像的差分直方图,从而确定分割阈值。文献[3]通过小波分解去除图像噪声以及增强裂缝边缘,然后使用阈值分割裂缝。文献[4]利用随机化结构森林去生成裂缝检测模型,并且通过特征直方图去区分裂缝和噪声。虽然传统裂缝分割方法取得一定效果,但分割精度不佳,易受外界环境因素干扰,且模型鲁棒性差,难以满足实际场景的应用需求。

随着深度学习的快速发展,基于卷积神经网络

的深度学习方法被广泛应用于裂缝分割之中。如文献[5]提出使用全卷积网络(Fully Convolutional Networks, FCN)分割不同尺度的裂缝,但网络在面对背景复杂多变的裂缝图像时,其特征提取能力较弱,导致裂缝分割精度较低。文献[6]利用残差网络ResNet-34作为编码器,一方面借此加深网络层数来提高编码器的特征提取能力,另一方面利用残差结构避免因网络加深发生梯度消失的问题;文献[7]引入可变形卷积,通过其空间几何形变能力自适应地提取不同形态的裂缝特征,增强网络对裂缝特征的学习。这些方法增强了网络在复杂背景下裂缝特征的获取,但缺乏对裂缝结构信息的描述,导致分割出的裂缝总体结构模糊不清晰。为此,文献[8]通过多尺度监督学习融合不同卷积阶段的输出结果,提高网络对裂缝线性结构的捕捉能力;文献[9]依靠Transformer建模长距离依赖关系,从而获取长距离裂缝的结构信息。与此同时,为解决传统卷积感受野受限,无法纳入更大范围裂缝信息的缺陷。文献[10-13]在不降低特征图分辨率和减少细节信息丢失的情况下,利用空洞卷积扩大感受野,捕获更大范围的裂缝信息;为了让网络应对尺寸多变的裂缝形状,文献[14]提出串并联相结合的空洞卷积模块,通过不同扩张率的空洞卷积,去囊括不同尺度的裂缝特征。为提高网络的裂缝分割性能,研究人员尝试改变编解码端浅层和高层特征的融合方式。文献[15]设计特征融合模块将包含边缘细节信息的中层特征与包含大量语义信息的深层特征相融合,从而细化裂缝边界。并且通过不同扩张率的空洞卷积提取多尺度上下文信息,整合更多有价值信息,提高裂

缝分割的准确度;文献[16]通过多尺度特征融合模块,将编解码端的最深层特征逐次引入其他中浅层特征中进行特征融合,充分利用最深层特征中的高级语义信息,并且通过深度监督策略促进多尺度特征融合和模型收敛。

虽然上述基于深度学习的裂缝分割方法取得不错成效,但仍存在以下不足:编解码特征在提取过程中,网络专注挖掘最显著的裂缝特征信息,容易忽略局部细微裂缝信息;编码器浅层特征包含丰富空间信息,但同时带来了背景噪声,对后续编解码特征融合造成背景干扰;虽然空洞卷积能够扩大网络感受野,但无法有效捕捉到裂缝的线性结构。鉴于此,本文提出一种线性引导与网格优化联合网络(Linear Guidance and Mesh Optimization Joint Network, LGMO-Net),首先在模型编码阶段加入多分支线性引导模块(Multi-branch Linear Guidance Module, MLGM),通过自适应单维度池化提高网络对裂缝线性结构的捕获能力,加强空间不同区域之间的信息交流,增强模型全局信息感知能力,从而提高网络分割精度。其次通过设计网格细节优化模块(Mesh Detail Optimization Module, MDOM),利用分区-优化-合并三步骤,对特征图空间区域进行网格划分,学习网格中局部细节信息,避免细微裂缝特征信息丢失。与此同时采用混合注意力模块(Mixed Attention Module, MAM),过滤编码器浅层特征中的背景噪声,在空间与通道双维度突出裂缝特征信息,抑制背景信息对裂缝分割结果的干扰。

2 线性引导和网格优化联合网络

2.1 网络结构

线性引导和网格优化联合网络框架具体如图1所示,主要包括主干网络U-Net、多分支线性引导模块MLGM、网格细节优化模块MDOM、混合注意力模块MAM。首先,网络在编码器端添加MLGM模块,通过自适应单维度池化提高

网络对裂缝线性拓扑结构的捕捉能力,获取图像全局感受野。同时在编码端加入MDOM模块,避免因连续卷积和池化操作导致细节信息丢失,造成细微裂缝漏分。对于编码器输出的浅层特征包含了大量的空间信息,但同时也带来背景噪声。这对编码特征与解码特征在跳跃连接处进行特征融合带来背景干扰。为此,本文在跳跃连接处嵌入MAM模块,通过注意力机制过滤编码特征中的背景噪声。同时改变跳跃连接方式,先将经过背景噪声过滤后的浅层特征与深层特征进行特征融合,再将经过特征融合的浅层特征与深层特征进行通道合并,弥补之前跳跃连接方式在语义融合上的不足。

2.2 多分支线性引导模块

图2是裂缝图像,图中裂缝呈现出细长且不规则的结构。如图2(a)所示,当网络采用常规池化进行信息聚合时,因其池化核与裂缝形状不契合,无法捕捉裂缝的线性结构,并且将裂缝和背景的特征信息汇聚在一起,导致分割精度不高。同时裂缝分布于全图,通过常规池化操作,只能获取局部区域的裂缝信息,造成区域信息交流闭塞,区域之间的裂缝无法建立联系,难以获取完整的裂缝。

针对上述问题,借鉴文献[17]的思想,本文使用自适应单维度池化代替常规池化。与固定尺寸的常规池化不同,自适应单维度池化的形状会根据输入特征图的尺寸进行自适应调整,让其形状与目标尺度相匹配。并且自适应单维度池化只沿高度或宽度维度进行信息聚合。当输入特征图空间维度大小为 $H \times W$ 时,自适应单维度池化形状为 $(\alpha * H) \times 1$ 或 $1 \times (\alpha * W)$, α 为延伸率。如图2(b)所示,自适应单维度池化其形状更加契合裂缝的形态结构,能够更好捕获裂缝的线性结构。并且让图像不同区域的裂缝建立联系,避免区域之间孤立。基于上述分析,本文设计了基于自适应单维度池化的多分支线性引导模块,其结构如图3所示,通过对特

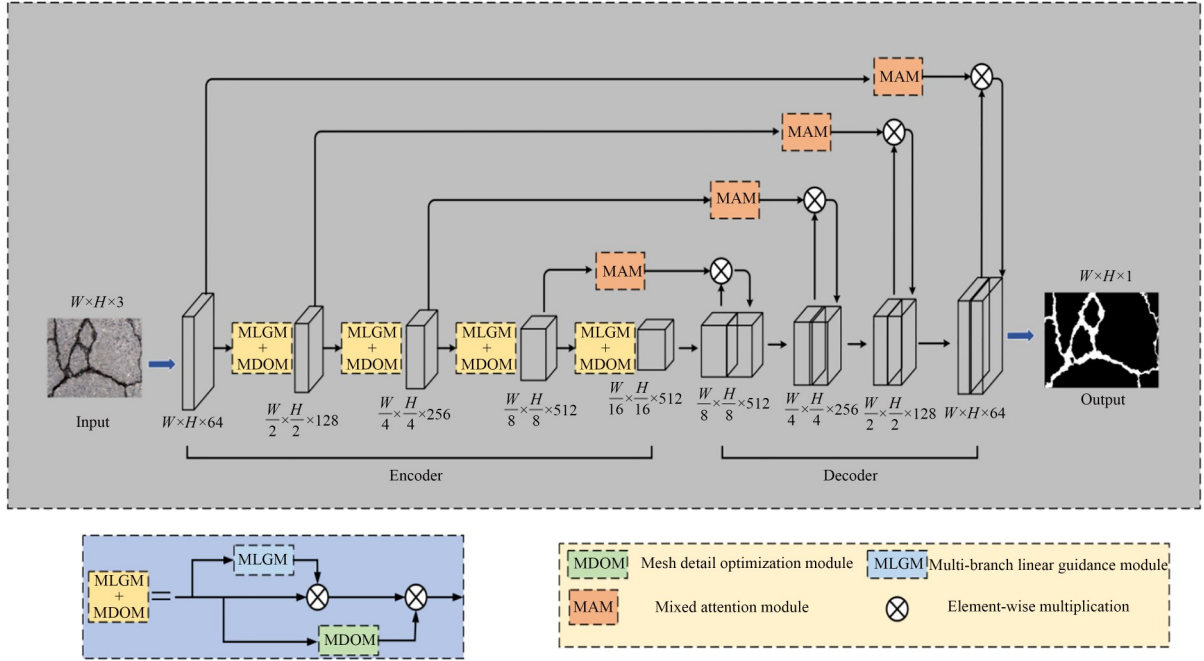


图 1 线性引导和网格优化联合网络框架

Fig. 1 Linear Guided and Mesh Optimization Joint Network Framework

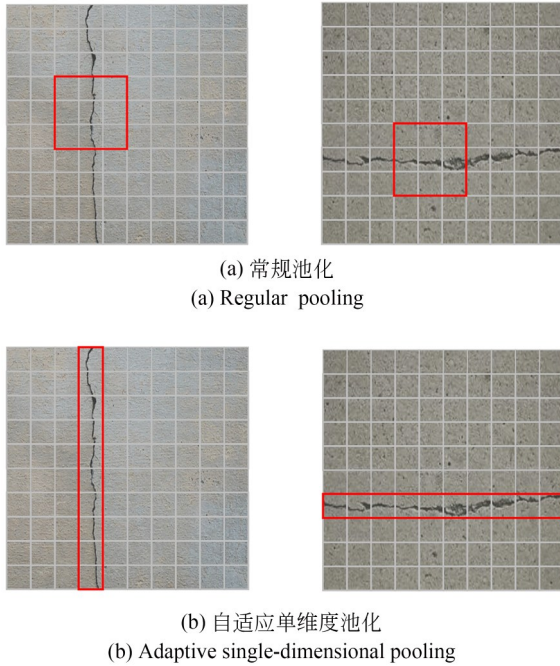


图 2 自适应单维度池化和常规池化对比

Fig. 2 Comparison of adaptive single-dimensional pooling and regular pooling

征图执行自适应单维度池化操作,学习到更丰富的裂缝信息。

具体的,多分支线性引导模块包含1条全局平均池化分支以及3条自适应单维度池化分支。对于全局平均池化分支,采用公式(4)对输入特征图 $F_{\text{input}} \in \mathbb{R}^{C \times H \times W}$ 执行全局平均池化操作,得到输出特征图 $F_{s4} \in \mathbb{R}^{C \times H \times W}$ 。通过上述方式,特征图 F_{s4} 获得输入特征图的全局上下文信息,提高模型对裂缝图像的整体感受能力。而对于自适应单维度池化分支,分别采用公式(1)~公式(3)对输入特征图 F_{input} 执行自适应单维度池化操作,得到输出特征图 $(F_{s1}, F_{s2}, F_{s3}) \in (\mathbb{R}^{C \times H \times W}, \mathbb{R}^{C \times H \times W}, \mathbb{R}^{C \times H \times W})$ 。为了应对裂缝尺寸多变的问题,三条自适应单维度池化分支的延伸率各不相同,延伸率 α 分别为1、0.5和0.25。并且每条自适应单维度池化分支还分别设置沿高度和宽度方向的自适应单维度池化子分支,去尽可能获取沿高度和宽度方向延伸的裂缝信息。

$$F_{s1} = \text{Conv}_{1 \times 1} \left(\text{Up} \left(\text{Avg}_{\alpha=1} (F_{\text{input}}) \right) \right), \quad (1)$$

$$F_{s2} = \text{Conv}_{1 \times 1} \left(\text{Up} \left(\text{Avg}_{\alpha=0.5} (F_{\text{input}}) \right) \right), \quad (2)$$

$$F_{s3} = \text{Conv}_{1 \times 1} \left(\text{Up} \left(\text{Avg}_{\alpha=0.25} (F_{\text{input}}) \right) \right), \quad (3)$$

$$F_{s4} = \text{Up} \left(\text{Conv}_{1 \times 1} \left(\text{GAP} (F_{\text{input}}) \right) \right), \quad (4)$$

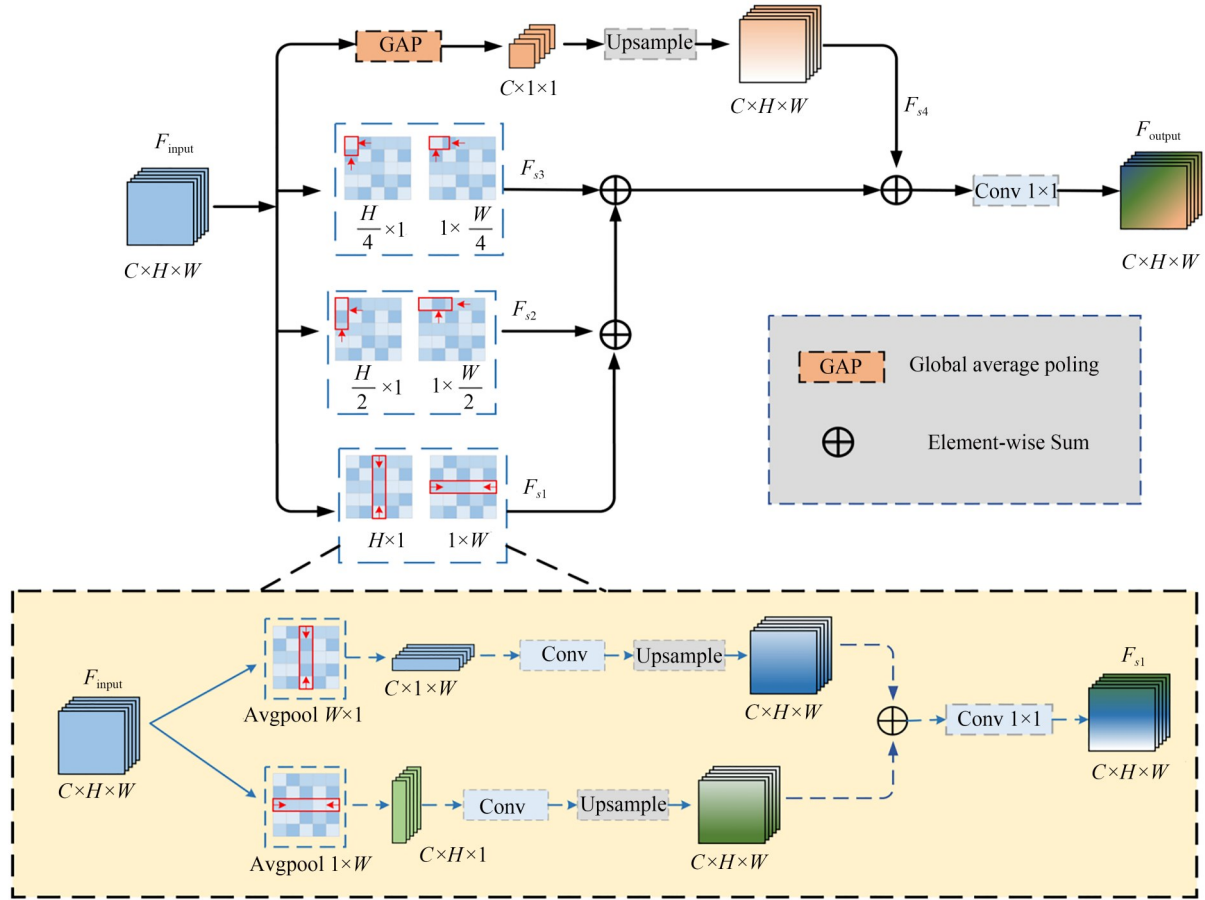


图 3 多分支线性引导模块

Fig. 3 Multi-branch Linear Guidance Module

其中: $Conv_{1 \times 1}$ 表示核大小为 1×1 的卷积; Up 表示上采样操作; GAP 表示全局平均池化操作; $Avg_{a=1}$ 表示核大小为 $H \times 1$ 和 $1 \times W$ 的自适应单维度池化; $Avg_{a=0.5}$ 表示核大小为 $H/2 \times 1$ 和 $1 \times W/2$ 的自适应单维度池化; $Avg_{a=0.25}$ 表示核大小为 $H/4 \times 1$ 和 $1 \times W/4$ 的自适应单维度池化。

最后, 将全局平均池化和自适应单维度池化分支输出的特征图进行逐元素特征融合。该步骤在保持特征图维度不变的情况下, 增加特征图空间与通道维度的信息量, 有效降低因特征融合所带来的参数和计算量, 生成含有丰富裂缝信息的特征图。之后通过 1×1 卷积操作, 实现特征图的跨通道信息交互, 获得最优特征图 F_{output} , 其过程如式(5)所示:

$$F_{output} = Conv_{1 \times 1}(F_{s1} \oplus F_{s2} \oplus F_{s3} \oplus F_{s4}), \quad (5)$$

式中, $\oplus$ 表示逐元素相加。

2.3 网络细节优化模块

由于网络在编码阶段不断地卷积和池化操作过程中, 图像细节信息丢失严重, 导致细微裂缝漏分。因此, 本文设计了网络细节优化模块, 以减少模型在下采样过程中细节信息的损失, 增强网络对小尺度裂缝的关注度。

网络细节优化模块结构如图4所示, 首先, 沿着通道维度对输入特征图 $F \in R^{C \times H \times W}$ 执行切片操作得到切片特征 $F_i (F_i \in R^{H \times W \times C/4}, i = 1, 2, 3, 4)$; 其次对每个切片特征 F_i 执行分区-优化-合并三步骤。

以切片特征 F_1 为例, 首先进行分区操作, 使用 3×3 平均池化对切片特征 F_1 进行空间区域划分, 将整个空间域划分为若干个网格, 得到网格描述特征, 其具体过程如式(6)所示:

$$z = f_{di}^{3 \times 3}(F_1) = \frac{1}{3 \times 3} \sum_0^2 \sum_0^2 (F_1), \quad (6)$$

其中: z 表示网格描述特征, $f_{di}^{3 \times 3}(\cdot)$ 表示分区操作。其次对网格描述特征 z 进行细节优化操作,利用卷积操作提取每个网格中的细节信息,其具体过程如公式(7)所示:

$$s = f_{ex}(z) = \delta(BN(Conv(z))), \quad (7)$$

其中: s 表示局部细节特征, $f_{ex}(\cdot)$ 表示细节优化操作, $Conv$ 表示卷积操作, BN 代表批归一化, δ 代表ReLU激活函数。

最后进行网格细节合并操作,通过双线性插值对局部细节特征 s 进行上采样,合并各个网格的信息,将细节特征非线性映射到整个空间域。其具体过程如下所示:

$$Y_1 = f_{bi}(s) = Bilinear(s), \quad (8)$$

其中: Y_1 是包含细节信息的分支特征, $f_{bi}(\cdot)$ 表示网格细节合并操作, $Bilinear$ 表示双线性插值操作。

对于四个切片特征 $F_i (F_i \in R^{H \times W \times C/4})$,使用

不同尺寸的池化执行分区-优化-合并三步骤。其计算过程如下:

$$Y_1 = f_{bi}(f_{ex}(f_{di}^{3 \times 3}(F_1))), \quad (9)$$

$$Y_2 = f_{bi}(f_{ex}(f_{di}^{4 \times 4}(F_2))), \quad (10)$$

$$Y_3 = f_{bi}(f_{ex}(f_{di}^{5 \times 5}(F_3))), \quad (11)$$

$$Y_4 = f_{bi}(f_{ex}(f_{di}^{6 \times 6}(F_4))), \quad (12)$$

式中: $Y_i (Y_i \in R^{H \times W \times C/4}, i=1,2,3,4)$ 表示第 i 条分支的分支特征, $f_{di}^{3 \times 3}, f_{di}^{4 \times 4}, f_{di}^{5 \times 5}, f_{di}^{6 \times 6}$ 表示分别使用 $3 \times 3, 4 \times 4, 5 \times 5, 6 \times 6$ 平均池化进行分区操作。之后将四条分支特征 Y_1, Y_2, Y_3 和 Y_4 ,按照通道维度进行拼接,并执行卷积操作,获得包含丰富局部细节信息的特征 Y_{output} 。其计算过程如式(13)所示:

$$Y_{output} = Conv(concat(Y_1, Y_2, Y_3, Y_4)), \quad (13)$$

其中: $Conv$ 表示卷积操作, $concat$ 表示通道拼接操作。

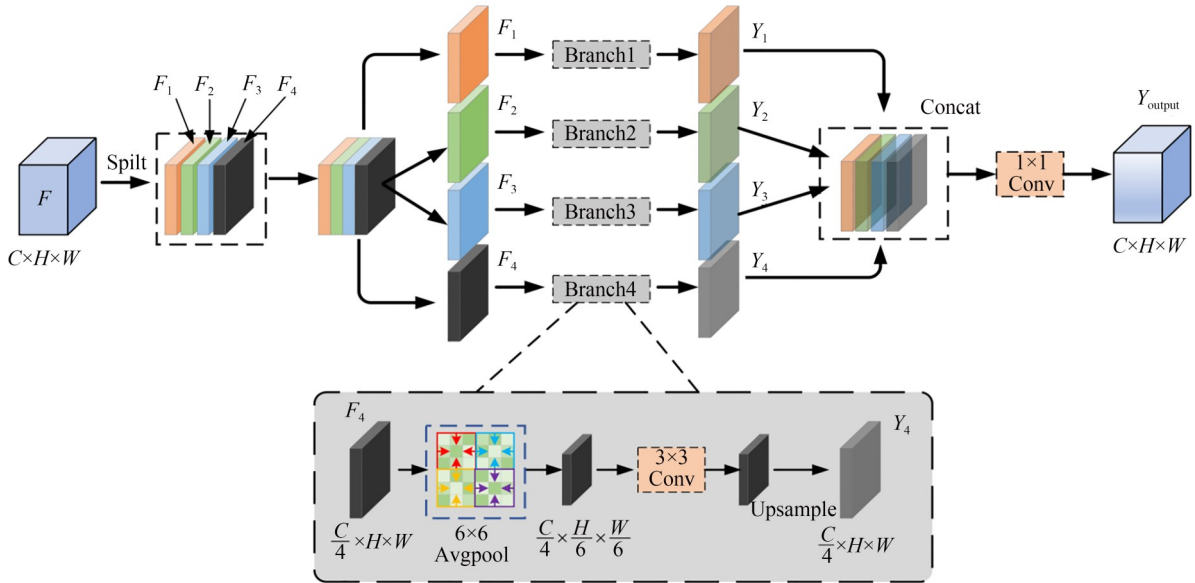


图4 网格细节优化模块

Fig. 4 Mesh Detail Optimization Module

2.4 混合注意力模块

在裂缝分割任务中,为了在复杂背景下实现对裂缝区域、背景区域进行准确分割,设计了空间-通道混合注意力模块,将其嵌入网络的跳跃连接处。通过混合注意力模块,特征图中裂缝区域被赋予较高权重,而背景区域被赋予较低权重,使网络聚焦于裂缝区域的分割,提高网络的

特征学习和表达能力,改善分割效果。

2.4.1 空间注意力模块

空间注意力机制通过对特征图不同位置的像素元素重新赋予权重,挖掘特征图空间信息,更好地聚焦裂缝区域。空间注意力模块结构如图5所示:

首先,空间注意力模块通过 1×1 卷积操作,

将特征图 $U \in R^{C \times H \times W}$ 沿着通道方向进行特征压缩, 去除特征对通道维度的依赖性, 得到特征图 $U_1 \in R^{1 \times H \times W}$ 。与此同时, 通过 3×3 池化操作, 将广泛的上下文信息聚集到局部特征中, 得到特征图 $U_2 \in R^{1 \times H/3 \times W/3}$ 。之后通过 3×3 卷积操作, 加强区域之间的信息交流, 建立空间关联性。然后上采样成大小为 $H \times W \times 1$ 特征图, 最后通过 ReLU 激活和 sigmoid 函数映射后生成空间特征权重矩阵 W_s 。利用该权重与输入特征图进行相乘, 即实现空间维度上特征重标定, 其计算过程如下:

$$U_1 = \text{Conv}_{1 \times 1}(U), \quad (14)$$

$$U_2 = \text{Avg}_{3 \times 3}(U_1), \quad (15)$$

$$W_s = \sigma(\text{Up}(\text{Conv}_{3 \times 3}(U_2))), \quad (16)$$

$$U_{\text{SFR}} = U \otimes W_s, \quad (17)$$

其中: σ 为 sigmoid 激活函数, Up 表示上采样, $\text{Conv}_{3 \times 3}$ 表示卷积核为 3×3 的卷积层, $\text{Conv}_{1 \times 1}$ 表示卷积核为 1×1 的卷积层, $\text{Avg}_{3 \times 3}$ 表示池化核为 3×3 的池化层, $\otimes$ 表示逐元素相乘, U_{SFR} 表示经过空间维度重标定的特征图。

2.4.2 通道注意力模块

通道注意力模块通过建模各个通道的重要程度, 自适应的建立特征通道间的依赖关系, 根据依赖关系对原特征图各通道进行加权处理, 使网络关注某些权重值大的通道。在通道注意力机制中, 通常采用全局池化方法将通道信息全局编码, 但容易忽略位置信息和空间结构, 造成分

割精度降低。为此, 本文设计了全新的通道注意力模块, 如图 5 所示:

首先, 通道注意力模块为了避免全局池化造成位置信息丢失, 将二维的全局池化分解为 2 个并行的自适应单维度池化。把特征图 $U \in R^{C \times H \times W}$, 沿着高度和宽度方向将进行特征聚合, 得到 2 个注意力特征图 $U_h \in R^{C \times 1 \times W}$ 和 $U_w \in R^{C \times H \times 1}$ 。该步骤让通道注意力模块在获取一个空间方向的远程依赖关系时, 同时保存沿另一个空间方向的精确位置信息。对两个注意力特征图进行降维融合, 得到每一个通道特征图的全局信息特征图 $U_g \in R^{C \times 1 \times 1}$ 。经过 1×1 卷积进行瓶颈操作, 将通道数由 C 变为 $C/4$ 。之后经过卷积和 ReLU 操作后, 将通道数恢复为 C 。最后, 经过 sigmoid 函数映射后生成通道特征权重矩阵 W_c , 与特征图 U_{SFR} 相乘, 从而实现通道方向上特征重标定。其计算过程如下:

$$U_h = \text{Avg}_{H \times 1}(U), \quad (18)$$

$$U_w = \text{Avg}_{1 \times W}(U), \quad (19)$$

$$U_g = U_h \oplus U_w, \quad (20)$$

$$W_c = \sigma(\text{Conv}_{1 \times 1}(\text{Conv}_{1 \times 1}(U_g))), \quad (21)$$

$$U_{\text{CFR}} = U_{\text{SFR}} \otimes W_c, \quad (22)$$

其中: σ 为 sigmoid 激活函数, $\text{Conv}_{1 \times 1}$ 表示卷积核为 1×1 的卷积层, GAP 表示全局平均池化, $\text{Avg}_{H \times 1}$ 和 $\text{Avg}_{1 \times W}$ 分别表示池化核尺寸为 $H \times 1$ 和 $1 \times W$ 的自适应单维度池化操作, U_{CFR} 表示经过通道维度重标定的特征图。

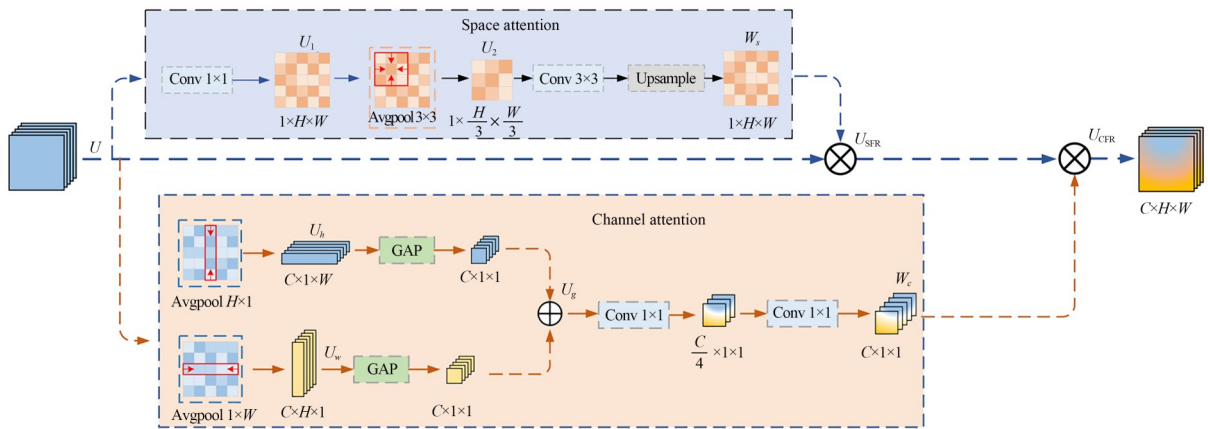


图 5 混合注意力模块

Fig. 5 Mixed Attention Module

2.5 损失函数

损失函数是用于衡量标签真实值与网络输出预测值之间差距的指标,通过不断减小损失值来训练网络模型,从而更好地拟合样本数据。本文采用二分类交叉熵损失函数, y_i 表示像素点的真实标签, y_p 表示像素点的预测值,二分类交叉熵损失函数的计算公式如式(23)所示:

$$Loss = -(y_i * \log(y_p) + (1 - y_i) \log(1 - y_p)). \quad (23)$$

3 实验分析与讨论

3.1 实验配置和数据集

本文实验是在Ubuntu 18.04.2系统下进行, GPU型号为GTX2080Ti。深度学习框架采用Pytorch-1.10.2,实验环境配置为CUDA 11.4+python3.6.8。在网络训练过程中,设置训练参数batch_size为8,epoch为400轮,初始学习率为0.0005,使用Adam优化器进行网络优化,学习率衰减策略采用余弦退火方法。

为了验证LGMO-Net的有效性,采用Crack500^[18]、Deepcrack537^[19]和CFD^[4]公共裂缝数据集进行实验对比。Deepcrack537裂缝数据集共有537张大小为544×384的裂缝图像,包括300张训练图像和237张测试图像以及对应的标签图。部分裂缝图像中含有较多的噪声干扰,如水坑、斑点、碎石、阴影等噪声,用以验证LGMO-Net在噪声干扰下的性能。Crack500裂缝数据集由3368张裂缝图像构成,图像中包含横向、纵向、龟裂等形状各异的裂缝,并且部分图像中裂缝与背景颜色相似,分割难度较大。该数据集按照训练集(1896张)、验证集(1124张)和测试集(348张)进行划分。CFD裂缝数据集由118张大小为480×320的裂缝图像构成,这些图像中的裂缝都是具有一定分割难度的狭长裂缝,裂缝占比小且难以辨认。该数据集按照8:2比例随机划分为训练集和测试集。

3.2 分割评价指标

为了更好地评价网络的分割性能,本文使用了4种常见的分割评价指标,包括准确率、召回率、F1-score和IoU。将裂缝定义为正样本,非裂缝定义为负样本。上述的评价指标可以通过混淆

矩阵进行计算,其中TP, FN, FP, TN的含义如表1所示:

表1 混淆矩阵

Tab. 1 Confusion matrix

样本	预测为裂缝	预测为非裂缝
实际为裂缝	TP	FN
实际为非裂缝	FP	TN

准确率(Precision)也称查准率,反映了预测为裂缝的样本中有多少实际为裂缝样本的概率,公式为:

$$Precision = \frac{TP}{TP + FP}. \quad (24)$$

召回率(Recall)也称查全率,表示实际为裂缝的样本中有多少被预测为裂缝样本的概率,公式为:

$$Recall = \frac{TP}{TP + FN}. \quad (25)$$

F1-score用于描述准确率和召回率之间关系,更好反映模型的分割性能,公式为:

$$F1-score = \frac{2 \times precision \times recall}{precision + recall}. \quad (26)$$

交并比(IoU)表示预测为裂缝区域与实际为裂缝区域之间交集与并集的比值,公式为:

$$IoU = \frac{TP}{TP + FP + FN}. \quad (27)$$

3.3 实验结果分析

为了证明LGMO-Net模型的优越性,与U-Net, PSPNet, DeepCrack18, LightCrackNet^[20], FFEDN^[21]等网络模型分别在Deepcrack537, Crack500和CFD裂缝数据集上进行对比实验。并且为了验证LGMO-Net模型各模块对混凝土裂缝分割任务的有效性,在Deepcrack537, Crack500和CFD裂缝数据集上进行消融实验。

3.3.1 Crack500裂缝数据集实验结果分析

图6是在Crack500数据集上, LGMO-Net模型与其他网络模型的分割结果对比(彩图见期刊电子版)。通过对比图6(a)和图6(b)的红框区域,可以看出FCN, SegNet等网络都出现了不同程度裂缝漏分现象,部分裂缝区域没有分割出来。而LGMO-Net所提出的多分支线性引导模块,能够精确关注到裂缝的线性结构,获取更加

丰富的裂缝特征信息,较为完整的分割出整个裂缝。通过图 6(c)的红框区域,可以看出 SegNet 等模型在裂缝分割过程中丢失了特征图的细节信息,导致细微裂缝分割效果不佳。而 LGMO-Net 利用网格细节优化模块,提取空间网格中的细节信息,较好分割出细微裂缝。在图 6(d)的红框中,由于裂缝图像中裂缝与背景特征相似,导致 DeeplabV3、DeepCrack18 等模型将背景错分割成裂缝。而 LGMO-Net 利用混合注意力模块在空间和通道两个维度强化裂缝特征,抑制背景噪声干扰,没有出现将背景错分为裂缝的现象。

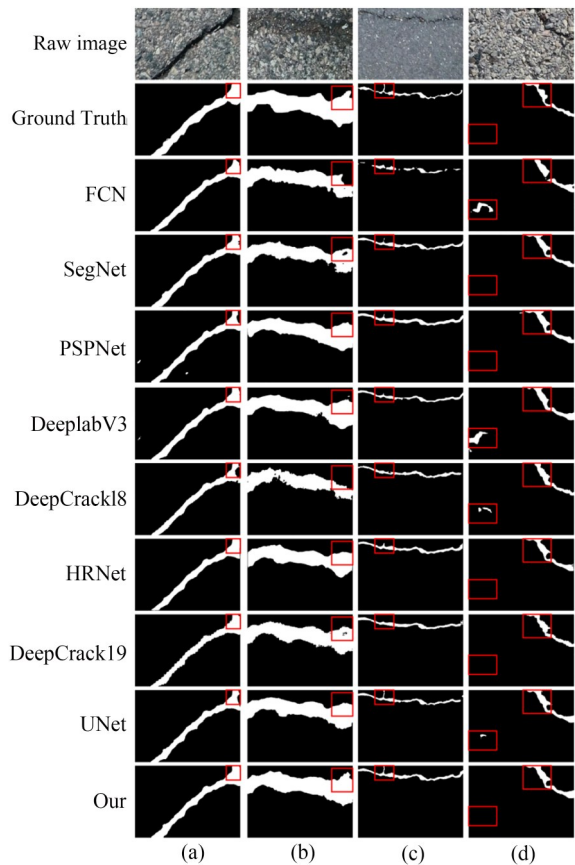


图 6 Crack500 数据集各种模型的裂缝分割结果
Fig. 6 Crack segmentation results of various models of Crack500 dataset

不同模型在 Crack500 数据集中分割结果的定量分析如表 2 所示,在 Crack500 数据集上,本文所提的 LGMO-Net 在准确率、召回率、F1-score 和 IoU 评价指标上,分别达到 78.11%, 70.64%, 74.19% 和 58.96%,其中准确率、F1-score 和 IoU 获得最高分数。FFEDN 网络在裂缝

表 2 Crack500 数据集定量分析实验结果

Tab. 2 Experimental results of quantitative analysis of Crack500 dataset

Methods	Metrics			
	Precision	Recall	F1-score	IoU
FCN ^[22]	66.63%	65.03%	65.82%	49.06%
DeeplabV3 ^[23]	70.69%	73.72%	72.17%	56.46%
SegNet ^[24]	69.79%	74.45%	72.04%	56.30%
PSPNet ^[25]	73.93%	71.23%	72.55%	56.93%
DeepCrack18 ^[26]	66.61%	69.56%	68.05%	51.58%
DeepCrack19 ^[19]	70.75%	70.78%	70.76%	54.75%
HRNet ^[27]	69.32%	74.99%	72.04%	56.30%
U-Net ^[28]	69.29%	73.73%	71.44%	55.57%
Qu et al ^[16]	65.40%	69.80%	67.50%	/
LightCrackNet ^[20]	75.30%	72.10%	73.70%	/
FFEDN ^[21]	71.01%	76.96%	73.87%	58.56%
Our model	78.11%	70.64%	74.19%	58.96%

分割过程中,通过注意力和深度监督机制减少裂缝样本漏分现象发生,使其召回率指标获得最高分数。LGMO-Net 相比于 FFEDN 网络,更专注于获取裂缝特征信息,提高裂缝分割准确率,其召回率有待提高。与其他算法相比,虽然本文算法在 Recall 指标没有获得最高分数,但 Precision 指标远超其他算法,并且在综合考虑准确率和召回率的 F1-score 指标上取得最高分数,表明 LGMO-Net 分割性能更佳。与主干网络 U-Net 相比,本文方法的准确率提高 8.82%, F1-score 提高 2.75%, IoU 提高 3.39%。与最佳对比算法 FFEDN 相比, LGMO-Net 的评价指标得到一定程度的提升,准确率提高 7.1%, IoU 提高 0.4%, F1-score 提高 0.32%,说明 LGMO-Net 通过多分支线性引导和网格细节优化模块提高网络对裂缝线性结构的捕获能力,在一定程度上缓解裂缝图像分割过程中细微裂缝漏分问题,并通过混合注意力机制抑制背景因素干扰,提高网络分割精度。

3.3.2 Deepcrack537 裂缝数据集实验结果分析

LGMO-Net 与其他网络在 Deepcrack537 数据集上的可视化结果如图 7 所示(彩图见期刊电子版)。从图 7(a)红框区域可以看出,HRNet、U-

Net等模型的分割效果较差,裂缝漏分现象严重,整体分割效果不理想。此外在图 7(b)的红框区域中,SegNet 和 U-Net 错分现象较为严重,把背景区域错分为裂缝。而本文方法在上述的红框区域都能进行精确的裂缝分割,表明 LGMO-Net 与其他方法相比能够获得更加丰富的裂缝特征信息,并且对于复杂背景下各种强干扰具有较好的鲁棒性。在图 7(c)~7(e)中,其他网络对于大面积裂缝区域分割效果较好,但容易出现细微裂缝漏分。而 LGMO-Net 明显改善了细微裂缝漏分问题,对于小尺度裂缝分割效果更为突出。

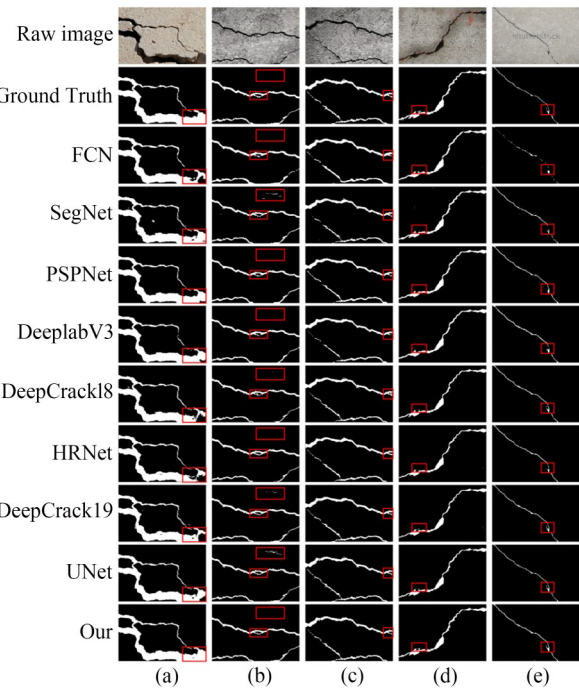


图 7 Deepcrack537数据集各种模型的裂缝分割结果
Fig. 7 Crack segmentation results of various models of Deepcrack537 dataset

在 Deepcrack537 裂缝数据集上,对各种网络模型的分割结果进行定量分析,结果如表 3 所示。本文所提出的 LGMO-Net 在 Precision, Recall, F1-score 和 IoU 指标上,分别达到 84.61%, 89.64%, 87.05% 和 77.07%。与其他网络模型相比,本文算法的 Recall, F1-score 和 IoU 指标获得最高分数。对比主干网络 U-Net, LGMO-Net 各项评价指标都得到大幅度提升,准确率提高 3.28%,召回率提高 1.50%,IoU 提高 2.45%,以及 F1-score 提高 3.76%。与对比算法 FFEDN 相

表 3 Deepcrack537数据集定量分析实验结果
Tab. 3 Experimental results of quantitative analysis of Deepcrack537 dataset

Methods	Metrics			
	Precision	Recall	F1-score	IoU
FCN ^[22]	74.95%	82.22%	78.42%	64.49%
DeeplabV3 ^[23]	82.31%	88.49%	85.29%	74.35%
SegNet ^[24]	83.12%	87.82%	85.41%	74.53%
PSPNet ^[25]	82.67%	88.26%	85.37%	74.48%
DeepCrack18 ^[26]	82.21%	89.08%	85.51%	74.69%
DeepCrack19 ^[19]	81.10%	88.45%	84.62%	73.34%
HRNet ^[27]	80.70%	86.88%	83.68%	71.94%
U-Net ^[28]	81.33%	88.14%	84.60%	73.31%
Qu et al ^[16]	86.70%	85.00%	85.80%	/
LightCrackNet ^[20]	85.50%	87.90%	86.70%	/
FFEDN ^[21]	86.73%	85.62%	86.17%	75.70%
Our model	84.61%	89.64%	87.05%	77.07%

比,召回率、F1-score 和 IoU 分别提高 4.02%, 0.88% 和 1.37%。

3.3.3 CFD 裂缝数据集实验结果分析

在 CFD 数据集中,裂缝形状细长且难以辨认,分割难度较大。为进一步验证 LGMO-Net 模型有效性,各模型在 CFD 裂缝数据集进行裂缝分割对比实验,对比结果如图 8 所示。在图 8(c) 的红框区域中,FCN、DeeplabV3 等网络模型对细微裂缝进行分割时,出现漏分现象。而 LGMO-Net 通过网格细节优化模块保留裂缝细节信息,能够完整分割出细微裂缝。在图 8(b)中,裂缝图像左下角存在阴影区域,干扰模型分割结果。如 DeepCrack18 网络受阴影区域干扰,将阴影区域错分为裂缝。并且 U-Net, HRNet 等网络模型在阴影区域发生裂缝漏分问题,分割效果表现不佳。而 LGMO-Net 能够克服阴影干扰,成功分割出裂缝。从图 8 的整体分割结果可以看出, LGMO-Net 相比于其他网络模型分割效果更佳,说明本文算法对于细长结构的裂缝,具有更好的特征捕获能力。

不同模型在 CFD 裂缝数据集分割结果的定量分析如表 4 所示,在 CFD 数据集上,本文所提的 LGMO-Net 在准确率、召回率、F1-score 和 IoU 评价指标上,分别达到 71.07%, 73.46%, 72.24% 和 56.55%,其中准确率、F1-score 和 IoU

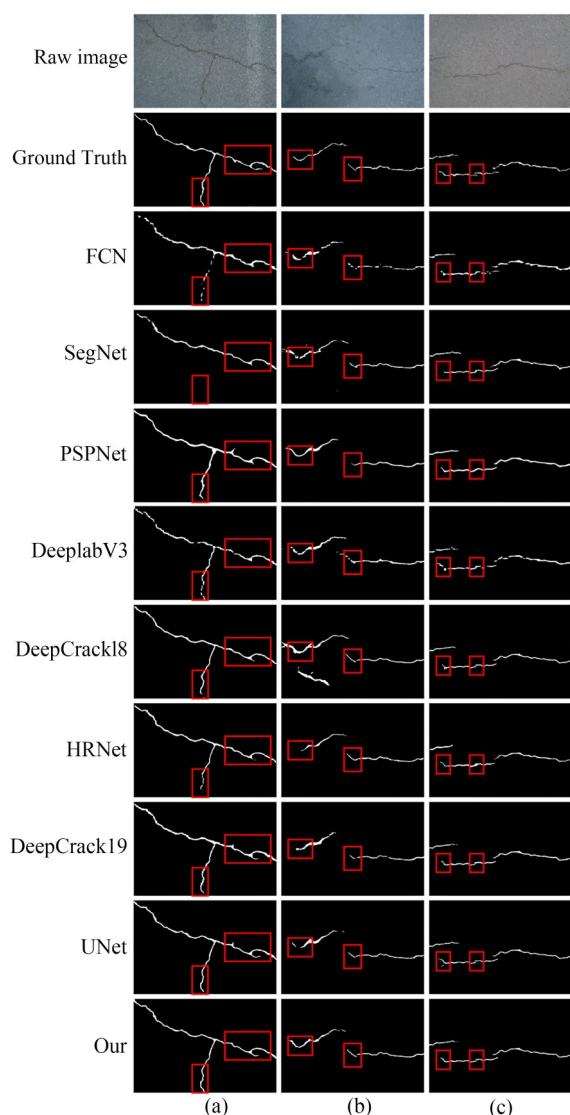


图8 CFD数据集各种模型的裂缝分割结果

Fig. 8 Crack segmentation results of various models of CFD dataset

获得最高分数。与 U-Net 相比,本文方法的准确率提高 3.09%,召回率提高 0.33%,IoU 提高 2.16%,F1-score 提高 1.78%。

综上所述,从 DeepCrack537, Crack500 和 CFD 裂缝数据集对比实验的可视化结果和定量分析上,可以看出 LGMO-Net 模型的裂缝分割效果十分显著,有效地解决了网络分割精度低、细微裂缝漏分、背景干扰等问题,能准确分割出混凝土的裂缝区域。

3.3.4 消融实验

为了更好的验证所提模型以及各个模块对混凝土裂缝分割的有效性,在 Deepcrack537、

表 4 CFD数据集定量分析实验结果

Tab. 4 Experimental results of quantitative analysis of CFD dataset

Methods	Metrics			
	Precision	Recall	F1-Score	IoU
FCN ^[22]	54.64%	56.65%	55.63%	38.53%
DeeplabV3 ^[23]	60.57%	61.67%	61.11%	44.00%
SegNet ^[24]	61.51%	72.99%	66.76%	50.10%
PSPNet ^[25]	62.67%	74.38%	68.03%	51.54%
DeepCrack18 ^[26]	70.27%	69.26%	69.76%	53.56%
DeepCrack19 ^[19]	69.95%	70.89%	70.41%	54.34%
HRNet ^[27]	65.32%	70.95%	68.02%	51.53%
U-Net ^[28]	67.98%	73.13%	70.46%	54.39%
Our model	71.07%	73.46%	72.24%	56.55%

Crack500 和 CFD 裂缝数据集上进行消融实验。所提模型是以 U-Net 为主干网络,MLGM 表示多分支线性引导模块,MDOM 表示网格细节优化模块,MAM 表示混合注意力模块。表 5 给出了在 Deepcrack537 数据集上的消融实验结果,U-Net 的 Precision, Recall, F1-score 和 IoU 的结果分别是 81.33%, 88.14%, 84.60% 和 73.31%; 添加 MLGM 模块后,指标分别上升 1.19%, 0.81%, 1.04% 和 1.57%; 添加 MDOM 模块后,指标相比主干网络 U-Net 分别上升 1.21%, 0.87%, 1.05% 和 1.60%; 添加 MAM 模块后, Precision, F1-score 和 IoU 三个指标分别上升 2.97%, 0.71% 以及 1.07%。表 6 展示了在 Crack500 数据集上的消融实验结果,当主干网络 U-Net 添加 MLGM 模块后, Precision, F1-score 和 IoU 分别提高 7.8%, 2.56% 和 3.16%; 添加 MDOM 模块后,指标相比 U-Net 分别上升 3.89%, 1.81% 和 2.22%; 添加 MAM 模块后, Precision, F1-score 和 IoU 三个指标分别上升 5.75%, 2.00% 以及 2.46%。CFD 数据集的消融实验结果如表 7 所示,U-Net 添加 MLGM 模块后, Precision, F1-score 和 IoU 分别提高 2.24%, 0.23% 和 0.28%; 添加 MDOM 模块后, Recall, F1-score 和 IoU 指标相比 U-Net 网络分别上升 2.41%, 0.15% 以及 0.18%; 添加混合注意力模块后, Recall, F1-score 和 IoU 三个指标分别上升 4.29%, 1.36% 和 1.64%。

表 5 Deepcrack537数据集消融实验结果

Tab. 5 Deepcrack537 dataset ablation experiment results

Methods	Metrics			
	Precision	Recall	F1-score	IoU
U-Net	81.33%	88.14%	84.60%	73.31%
U-Net+MLGM	82.52%	89.00%	85.64%	74.88%
U-Net+MDOM	82.54%	89.01%	85.65%	74.91%
U-Net+MAM	84.30%	86.35%	85.31%	74.38%
Our model	84.61%	89.64%	87.05%	77.07%

表 6 Crack500数据集消融实验结果

Tab. 6 Crack500 dataset ablation experiment results

Methods	Metrics			
	Precision	Recall	F1-score	IoU
U-Net	69.29%	73.73%	71.44%	55.57%
U-Net+MLGM	77.09%	71.14%	74.00%	58.73%
U-Net+MDOM	73.18%	73.32%	73.25%	57.79%
U-Net+MAM	75.04%	71.91%	73.44%	58.03%
Our model	78.11%	70.64%	74.19%	58.96%

表 7 CFD数据集消融实验结果

Tab. 7 CFD dataset ablation experiment results

Methods	Metrics			
	Precision	Recall	F1-score	IoU
U-Net	67.98%	73.13%	70.46%	54.39%
U-Net+MLGM	70.22%	71.17%	70.69%	54.67%
U-Net+MDOM	66.29%	75.54%	70.61%	54.57%
U-Net+MAM	66.98%	77.42%	71.82%	56.03%
Our model	71.07%	73.46%	72.24%	56.55%

图 9 为消融实验的可视化结果对比图(彩图见期刊电子版),在图 9(c)的红框区域中,U-Net 网络分割出的裂缝形状粗糙,并且出现将背景错分为裂缝以及细微裂缝漏分问题。而图 9(d)红框区域内,当 U-Net 添加 MLGM 和 MDOM 模块后,网络分割出的裂缝整体更加完整,没有发生细微裂缝漏分现象。在图 9(e)红框区域内,当 U-Net 添加 MAM 模块后,网络通过注意力机制抑制背景区域对分割结果干扰,未将背景区域错分为裂缝。在图 9(f)红框中,LGMO-Net 相比于 U-Net 网络在细节和背景干扰方面表现良好,极大地减少了漏分、误分问题。

消融实验结果充分表明,本文所提多分支线性引导模块 MLGM、网格细节优化模块 MDOM

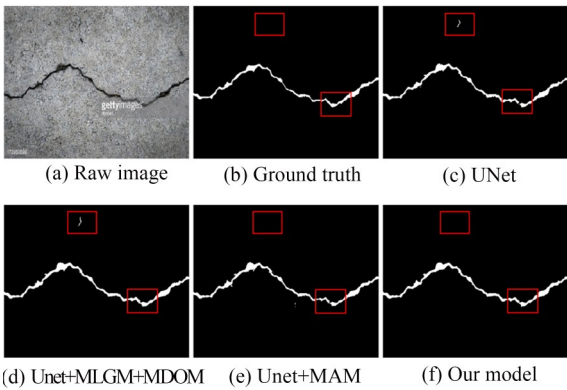


图 9 消融实验的可视化结果

Fig. 9 Visualization results of ablation experiments

以及混合注意力模块 MAM,对模型的整体分割精度提升较为显著,在不同数据集上的分割效果均十分优异,说明了本文方法在面对不同裂缝分割任务时均具有一定的适应性和鲁棒性,整体性能表现良好。

3.3.5 不同延伸率组合的对比实验

在实验过程中,采用本文算法 LGMO-Net 进行实验。通过改变 LMGO-Net 中 MLGM 模块的延伸率 α 组合,验证不同延伸率组合对 LGMO-Net 性能的影响,并在 Deepcrack537,CFD 数据集上进行对比实验。

表 8 给出了在 Deepcrack537 数据集上的对比实验结果,当 MLGM 的延伸率选择 1,0.5 和 0.25 组合时,LGMO-Net 网络的 Precision, Recall, F1-score 和 IoU 指标分别为 84.61%, 89.64%, 87.05% 和 77.07%。而当 MLGM 的延伸率为 1,0.75,0.5 和 0.25 的组合时,LGMO-Net 网络指标分别下降 3.25%,0.71%,2.07%,3.19%。表 9 给出了在 CFD 数据集上对比实验结果,当 MLGM 的延伸率选择 1,0.5 和 0.25 组合时,LGMO-Net 的 Precision, Recall, F1-score 和 IoU 指标分别为 71.07%, 73.46%, 72.24% 和 56.55%。而当 MLGM 的延伸率为 1,0.75,0.5 和 0.25 组合时,LGMO-Net 的 Precision, F1-score 和 IoU 指标分别下降 5.29%,2.34% 和 2.82%。

从上述实验结果的定量分析中,可以发现当 LGMO-Net 中 MLGM 模块的延伸率选择 1,0.75,0.5,0.25 组合时,模型的各项评价指标有一定幅度下降,模型的分割性能不佳,并且随着

表 8 Deepcrack537数据集对比实验结果

Tab. 8 Comparative experimental results of Deepcrack537 dataset

α	Metrics			
	Precision	Recall	F1-score	IoU
1, 0.5, 0.25	84.61%	89.64%	87.05%	77.07%
1, 0.75, 0.5, 0.25	81.36%	88.93%	84.98%	73.88%

表 9 CFD 数据集对比实验结果

Tab. 9 Comparative experimental results of CFD dataset

α	Metrics			
	Precision	Recall	F1-score	IoU
1, 0.5, 0.25	71.07%	73.46%	72.24%	56.55%
1, 0.75, 0.5, 0.25	65.78%	74.57%	69.90%	53.73%

MLGM 模块分支数增多会导致其参数量上升。因此,本文 MLGM 模块选择延伸率为 1, 0.5, 0.25 组合。

3.4 复杂性分析

表 10 展示了不同模型的复杂度比较结果,表中 Params 表示模型参数量,能够衡量模型空间复杂度;表中的 FLOPs (Floating-point Operations) 表示浮点运算次数,能够衡量模型时间复杂度。为了统一衡量各个模型复杂度,所有模型的输入都是大小为 $512 \times 512 \times 3$ 的张量。通过表 10 的比较结果可见,本文方法的模型参数量相较 U-Net 增加 10.98 M,但相较 PSPNet 减少 25.08 M。与 U-Net 相比, LGMO-Net 在增加少数参数数量的情况下实现更优越的裂缝分割性能。而与 PSPNet 相比, LGMO-Net 在模型参数量更小的情况下表现出更出色的分割性能。综合而言, LGMO-Net 在模型复杂度和模型精度之间实现了良好的平衡。

表 10 参数量和浮点运算次数

Tab. 10 Number of parameters and number of floating-point operations

Methods	FLOPs/G	Params/M
FCN ^[22]	80.51	15.31
DeeplabV3 ^[23]	161.50	39.04
SegNet ^[24]	160.56	29.44
PSPNet ^[25]	201.30	53.32
DeepCrack18 ^[26]	548.01	30.91
DeepCrack19 ^[19]	80.46	14.72
HRNet ^[27]	90.74	29.53
U-Net ^[28]	160.76	17.26
Our model	170.49	28.24

4 结 论

针对混凝土裂缝分割精度低以及容易出现复杂背景干扰、细微裂缝漏分等问题,本文提出一种 LGMO-Net 模型用于混凝土裂缝分割。所提模型在 U-Net 模型基础上,在网络的编码阶段运用多分支线性引导模块,来引导网络捕获裂缝线性结构,提高裂缝分割精度。同时利用网格细节优化模块,提取特征网格中的细节信息,避免细微裂缝特征信息丢失,进一步提高分割精度。并且改变编码阶段和解码阶段的跳跃连接方式,在跳跃连接上运用空间-通道混合注意力模块过滤编码器特征中的背景噪声,减少背景因素干扰。本文所提出的网络在 Deepcrack537, Crack500 和 CFD 三个公开裂缝数据集上的 F1-score 指标为 87.05%, 74.19% 和 72.24%, IoU 指标为 77.07%, 58.96% 和 56.55%, 两个指标均获得最高分数。实验结果表明,所提网络具有较好的裂缝分割能力。

参考文献:

- [1] SONG Q, LIN G Y, MA J Q, *et al.* An edge-detection method based on adaptive canny algorithm and iterative segmentation threshold [C]. 2016 2nd International Conference on Control Science and Systems Engineering (ICCSSE). Singapore. IEEE, 2016: 64-67.
- [2] LI Q Q, LIU X L. Novel approach to pavement image segmentation based on neighboring difference histogram method [C]. 2008 Congress on Image and Signal Processing. Sanya, China. IEEE, 2008: 792-796.
- [3] WU S G, LIU Y K. A segment algorithm for crack detection [C]. 2012 IEEE Symposium on Electrical & Electronics Engineering (EEESYM). Kuala Lumpur. IEEE, 2012: 674-677.
- [4] 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.
- [5] 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.
- [6] LAU S L H, CHONG E K P, YANG X, *et al.* Automated Pavement Crack Segmentation Using U-Net-Based Convolutional Neural Network [EB/OL]. 2020: *arXiv*: 2001.01912. <http://arxiv.org/abs/2001.01912.pdf>
- [7] 李海丰, 景攀, 韩红阳. 基于可变形卷积与特征融合的机场道面裂缝检测算法[J]. *南京航空航天大学学报*, 2021, 53(6): 981-988.
- LI H F, JING P, HAN H Y. Airport pavement crack detection algorithm based on deformable convolution and feature fusion[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2021, 53 (6): 981-988. (in Chinese)
- [8] 徐胜军, 郝明, 孟月波, 等. 基于特征增强整体嵌套网络裂缝检测方法[J]. *激光与光电子学进展*, 2022, 59(10): 1010003.
- XU S J, HAO M, MENG Y B, *et al.* Crack detection method of holistically-nested network based on feature enhancement [J]. *Laser & Optoelectronics Progress*, 2022, 59(10): 1010003. (in Chinese)
- [9] XU Z S, GUAN H Y, KANG J, *et al.* Pavement crack detection from CCD images with a locally enhanced transformer network[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 110: 102825.
- [10] JI W P, ZHANG Y Z, HUANG P F, *et al.* A neural network with spatial attention for pixel-level crack detection on concrete bridges [C]. 2022 *IEEE 11th Data Driven Control and Learning Systems Conference (DDCLS)*. Chengdu, China. IEEE, 2022: 481-486.
- [11] 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 & Optoelectronics Progress*, 2021, 58(22): 2210001. (in Chinese)
- [12] 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)
- [13] 瞿中, 陈雯. 基于空洞卷积和多特征融合的混凝土路面裂缝检测[J]. *计算机科学*, 2022, 49(3): 192-196.
- QU Z, CHEN W. Concrete pavement crack detection based on dilated convolution and multi-features fusion[J]. *Computer Science*, 2022, 49(3): 192-196. (in Chinese)
- [14] QU Z, CHEN W, WANG S Y, *et al.* A crack detection algorithm for concrete pavement based on attention mechanism and multi-features fusion [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(8): 11710-11719.
- [15] QU Z, WANG C Y, WANG S Y, *et al.* A method of hierarchical feature fusion and connected attention architecture for pavement crack detection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(9): 16038-16047.
- [16] 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*, 2022, 33(9): 4890-4899.
- [17] HOU Q B, ZHANG L, CHENG M M, *et al.* Strip pooling: rethinking spatial pooling for scene parsing[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 4002-4011.
- [18] 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.
- [19] 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.
- [20] ZHOU Q, QU Z, JU F R. A lightweight network for crack detection with split exchange convolution and multi-scale features fusion[J]. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(3): 2296-2306.

- [21] LIU C Q, ZHU C G, XIA X, *et al.* FFEDN: feature fusion encoder decoder network for crack detection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23 (9): 15546-15557.
- [22] 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.
- [23] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [24] 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.
- [25] 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)*. Honolulu, HI, USA. IEEE, 2017: 6230-6239.
- [26] ZOU Q, ZHANG Z, LI Q Q, *et al.* DeepCrack: learning hierarchical convolutional features for crack detection [J]. *IEEE Transactions on Image Processing*, 2018: 1498-1512.
- [27] SUN K, ZHAO Y, JIANG B R, *et al.* High-Resolution Representations for Labeling Pixels and Regions [EB/OL]. 2019: *arXiv*: 1904.04514. <http://arxiv.org/abs/1904.04514.pdf>
- [28] 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.

作者简介:



刘光辉(1976—),男,副教授,硕士生导师,2016年于西安建筑科技大学获得工学博士学位,主要从事计算机视觉理解、建筑智能化技术、人机交互与多模态感知方面的研究。

E-mail: guanghuail@163.com