

DOI:10.19651/j.cnki.emt.2415911

基于像素强度顺序变换和 UNetFormer 的 裂缝分割模型*

姚成武¹ 常琛²

(1. 甘肃省经济研究院 兰州 730050; 2. 兰州交通大学电子与信息工程学院 兰州 730070)

摘要: 路面裂缝作为道路的最普遍病害之一,及时准确地识别和定位裂缝对道路的养护与持续健康运行具有重要意义。然而路面裂缝检测易受路面光照、路面阴影以及路面环境等复杂因素的影响,导致路面裂缝分割精度低并容易出现断裂等问题。为实现路面裂缝图像快速、准确的语义分割,本文提出了基于像素强度顺序变换和 UNetFormer 的路面裂缝分割模型。首先采用像素强度顺序变换算法对裂缝图片进行预处理,根据各像素与其附近像素之间的强度顺序,将图像沿着对角线 4 个方向转换为对比度更高的四通道图像,保留裂缝曲线结构固有特征的同时有效增强了裂缝与背景像素对比度;然后基于 UNet 和 Transformer 网络的结构特征,通过构建 UNetFormer 分割模型完成对路面裂缝的高精度分割,其中设计并引用了全局-局部注意力机制以充分捕获路面裂缝特征信息。最后,使用 CFD、Crack200 和 Crack500 三个开源数据集进行实例验证,实验结果表明,本文所提的裂缝分割模型 F_1 -score 分别达到 83.4%、82.6% 和 81.9%,模型参数仅为 UNet 网络模型的 37.7%,相较于现有的裂缝分割模型具有更高的分割精度以及更强的泛化能力。

关键词: 路面裂缝;语义分割;PIOT;UNetFormer;GLAM

中图分类号: TN98 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Pavement crack segmentation model based on pixel intensity order transform and UNetFormer

Yao Chengwu¹ Chang Chen²

(1. Gansu Economic Research Institute, Lanzhou 730050, China;

2. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: As one of the most common diseases of roads, the timely and accurate identification and localization of cracks is of great significance to the maintenance and continuous healthy operation of roads. However, the detection of pavement cracks is easily affected by complex factors such as road illumination, road shadows, and road environment, which leads to low segmentation accuracy of pavement cracks and prone to fracture and other problems. In order to realize fast and accurate semantic segmentation of pavement crack images, this paper proposes a pavement crack segmentation model based on pixel intensity order transform (PIOT) and UNetFormer. Firstly, the PIOT algorithm is used to preprocess the crack images, and according to the intensity order between each pixel and its neighboring pixels, the image is converted into a four-channel image with higher contrast along the four directions of the diagonal, which retains the intrinsic features of the crack curve structure and effectively enhances the contrast between the cracks and the background pixels. Then, based on the structural characteristics of UNet and Transformer networks, the high-precision segmentation of pavement cracks is accomplished by constructing the UNetFormer segmentation model, in which the global-local attention mechanism is designed and invoked to fully capture the pavement crack feature information. Finally, three open-source datasets, CFD, Crack200 and Crack500, are used for example validation, and the experimental results show that the F1-score of the crack segmentation model proposed in this paper reaches 83.4%, 82.6%, and 81.9%, respectively, and the model parameter is only 37.7% of that of the UNet network model, which provides higher segmentation accuracy compared to the existing crack segmentation model and stronger generalization ability than the existing crack segmentation models.

Keywords: pavement cracks; semantic segmentation; PIOT; UNetFormer; GLAM

0 引言

截至 2023 年底,中国公路总里程达到 550 万公里,过

去十年增加了 112 万公里稳居世界第一^[1]。公路建设飞速发展的同时,路面裂缝作为最常见的道路病害之一,成为公路养护任务中最关注的问题^[2]。对裂缝进行及时检测与防

收稿日期:2024-04-23

* 基金项目:国家自然科学基金(62262038)、甘肃省重点研发计划(22YF7GA145)项目资助

护,能够有效预防其发展成为大面积的龟裂甚至是坑槽等严重的路面病害,可有效提高道路的安全性并延长使用寿命。

传统的图像分割技术主要分为阈值化方法和边缘检测方法。其中阈值化方法是通过设定灰度值阈值,将图像的像素分为背景和对象两部分,目前使用最广泛的阈值化方法有类间方差、最优阈值、基于不变矩和梯度幅度等。边缘检测方法则是在相应的裂纹边界和细节等线性特征上,将属于不同灰度级的两个区域边界分隔开。然而传统的图像分割技术在分隔精度和效率上已很难满足对长距离路面检测的要求。

随着相关硬件设备的计算性能在不断提升,基于深度学习和计算机视觉的图像处理技术得到快速发展^[3]。其中语义分割作为计算机视觉领域的核心之一,不仅能够准确地检测裂缝的位置,还能提供对裂缝的语义理解。相较于传统的裂缝分割方法,可以有效地避免噪声、阴影、光照不均等因素对裂缝识别精度的影响。Cha 等^[4]提出了基于卷积神经网络(convolutional neural network,CNN)的滑动窗口裂缝检测方法,取得了较好的识别效果。廖延娜等^[5]针对裂缝形态复杂多样以及识别精度低的难题,提出了基于 Mask-RCNN(mask regionbased convolutional neural network)的桥梁裂缝检测算法,通过引入语义增强模块和结合特征金字塔网络来提高裂缝检测精度。张华卫等^[6]针对现有裂缝检测方法存在精度低、噪声多和细节丢失等问题,提出一种基于注意力融合网络和并行提取的裂缝检测方法,利用像素注意力机制融合裂缝的高低阶特征来消除噪声干扰并提高裂缝检测精度。Sun 等^[7]在 DeepLabv3+ 的解码器中加入一个注意力掩模,在多尺度注意力模块的帮助下,实现在低级和高级特征映射之间自动加权,有效提升了分割性能。Guo 等^[8]基于 U-Net 网络开发了一种边缘适应的细化网络(BARNet)来分割道路裂缝,有效地解决了在识别裂缝位置时边界不精确的问题。Wang 等^[9]将裂缝检测信息与分割信息进行融合,提出了 RDSNet 模型,从一定程度上提高了裂缝检测精度。虽然上述模型在路面裂缝分割效果方面都有不同程度的提升,但对边缘裂缝特征信息的提取,复杂噪声及路面阴影干扰下的裂缝分割,仍存在较大不足,难以快速准确地分割出裂缝区域。

近年来,Transformer 模型因其优异的全局图像特征捕捉能力而受到广泛关注,相较于传统的深度 CNN,Transformer 具有更好的长距离依赖建模能力、并行计算能力、多尺度特征表示和位置编码的处理能力^[10-11]。可以通过自身的多头自注意力机制更加快速精确地定位和分割裂缝,从而提高分割的准确性和鲁棒性。其中 Liu 等^[12]针对裂缝背景中复杂的纹理噪声对分割的影响以及细小裂缝误识别的问题,提出了一种细粒度裂缝检测架构 CrackFormer;通过缩放注意块来抑制非语义特征,能够高效检测出细粒度裂缝。在此基础上,邝先验等^[13]针对如何

平衡裂缝检测的精度和实时性的问题,提出一种基于轻量级 Transformer 的裂缝分割方法 CrackViT,通过融合 CNN 与 Transformer 构建裂缝特征提取网络,提高对裂缝图像全局和局部特征信息提取能力的同时,有效减少了网络模型参数和计算量。

此外,由路面裂缝的特点可知,对裂缝的曲线结构分析是必要的。而现有的曲线对象分割方法通常将重点集中在具体的工程特征来进行设计^[14]。通过调整算法参数来解决某些特定环境下的路面裂缝问题,难以应对具有多种复杂曲线结构的路面裂缝的划分,特别是在裂缝像素强度不均匀以及嘈杂的背景像素之间对比度低的情况下效果很差。为此,本文提出了一种基于像素强度顺序变换(pixel intensity order transform,PIOT)和 UNetFormer 的路面裂缝分割模型 PIOT-UNetFormer,以提高不同路面的裂缝分割精度和鲁棒性。其中 PIOT 算法是一种增强路面裂缝曲线结构特征的方法,且经过裂缝分割网络 UNetFormer 模型可以更高效的捕捉裂缝特征,对道路裂缝进行精准分割。本文主要贡献如下:

1) 提出了一种图像局部强度顺序变换算法,通过转换像素强度的顺序关系来增强图像中的线性结构,同时抑制不相关的结构和噪声,可有效增强裂缝曲线结构的固有特征,并提高裂缝与背景像素的对比度,使分割模型更容易捕获裂缝的特征信息;

2) 设计了全局-局部注意力机制(global-local attention mechanism,GLAM)模块,以充分捕获路面裂缝全局与局部特征信息,增强对裂缝关键信息的关注并抑制无关信息,并将其引入到 UNetFormer 裂缝分割模型,可有效缓解边缘裂缝分割间断或识别不到的问题;

3) 基于 UNet 和 Transformer 网络构建裂缝分割模型 UNetFormer,充分融合并利用 CNN 与 Transformer 网络分割裂缝的优势,完成对路面裂缝的高精度分割的同时减少了模型参数和计算量,有效提高模型的效率和泛化能力;

4) 本文方法在数据集 CFD、Crack200 和 Crack500 上进行验证,在 3 个数据集上精度分别达到了 82.4%、83.5%和 80.4%。

1 路面裂缝分割模型

路面裂缝数据的采集易受路面光照、路面阴影以及路面环境等复杂因素的影响,且在整幅裂缝图像中裂缝本身占比较小,致使在图像中裂缝与路面背景对比不鲜明,特别是边缘或细小裂缝,某些细节信息在处理的过程中极易被忽略,从而导致路面裂缝分割精度低并容易出现断裂等问题。为实现路面裂缝的快速精准分割,本文提出了一种基于像素强度顺序变换和 UNetFormer 的路面裂缝分割模型,旨在提高路面裂缝的分割精度和增强模型的泛化能力。本文的总体框架流程如图 1 所示。其主要分为图像处理模

块和图像语义分割模块。首先采用 PIOT 算法对裂缝图片进行预处理,根据各像素与其附近像素之间的强度顺序,将图像沿着对角线 4 个方向转换为对比度更高的四通道图

像,保留裂缝曲线结构固有特征的并增强裂缝与背景像素对比度;然后通过 UNetFormer 图像分割模型完成对路面裂缝的快速与高精度分割。

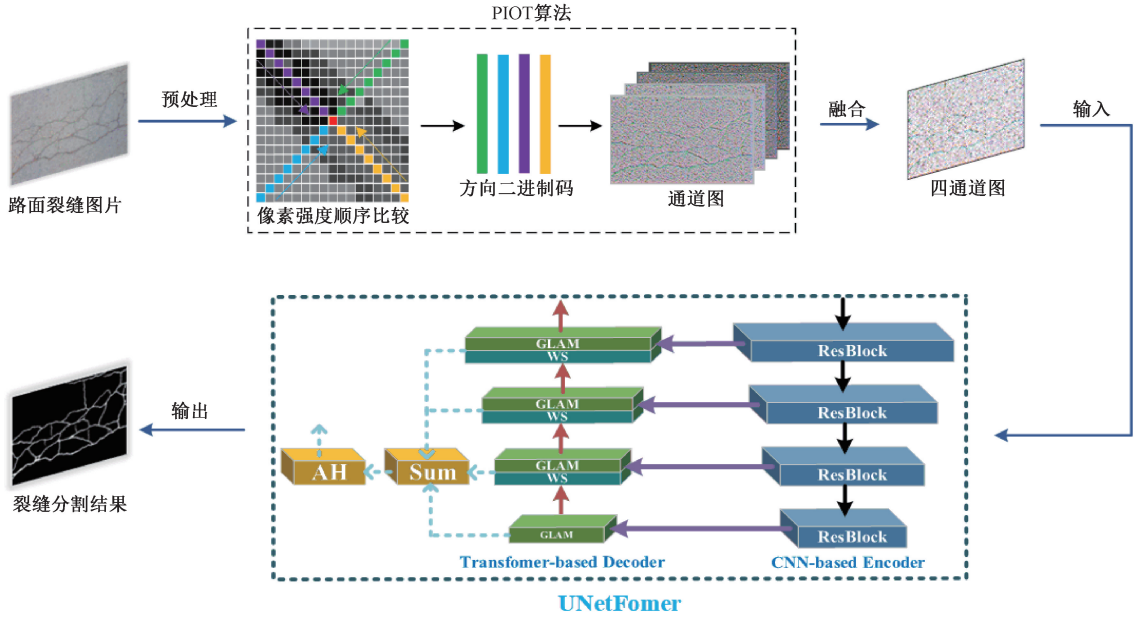


图 1 本文总体框架图

Fig. 1 Overall frame diagram of this paper

1.1 局部对角线灰度变换算法

PIOT 算法基于像素强度的相对顺序用于裂缝的特征描述,受 Census Transform(CT)^[15]算法启发。CT 算法是一种非参数局部变化,如图 2(a)所示,它的值仅取决于窗口中像素的相对灰度,所以该变换本质上不受失真的影响,从而保持相对灰度顺序。SparseCT^[16]在 CT 算法的基础上使用较大邻域区域内的像素子集,如图 2(b)所示。与其他基于 CT 的变体相比,本文中提出 PIOT 算法是专门从 4 个斜对角方向来捕捉裂缝结构,从中心像素的 4 个对角方向的邻域像素上依次由外向内进行像素强度顺序比较,如图 2(c)所示。

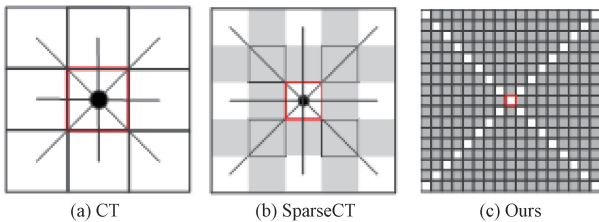


图 2 不同像素强度顺序变换算法对比

Fig. 2 Comparison of different pixel intensity order transform algorithms

尽管不同类型路面的裂缝图像的对比度和灰度分布都不同,但路面裂缝的共有特征是曲线结构比上下文更暗。PIOT 变换算法不依赖于每个像素的绝对值,因此像素之间对比度的变化更稳定。PIOT 算法工作过程如图 3

所示。首先将给定图像转换为灰度图像,根据局部灰度顺序计算 4 个方向二进制码,形成一个捕捉裂缝曲线结构特征的基于四通道灰度顺序的图像。准确来说该方法按照局部对角线方向的灰度顺序进行计算,对于一幅图 $f(x, y)$ 中的每个像素 p ,将 $f(p)$ 的值与 p 左上方 45°方向、左下方 45°方向、右上方 45°方向以及右下方 45°方向的每组 8 个相邻像素 $\{n'_i | i = 1, \dots, 8\}$ (p 的欧几里德距离为 $\sqrt{2} \sim 8\sqrt{2}$) 进行比较,这里像素间的像素强度顺序差依照从外向内的顺序对像素进行比较,其中, $s \in \{lu, ru, ld, rd\}$ 分别表示 p 的左上方 45°方向、右上方 45°方向、左下方 45°方向以及右下方 45°方向。通过 4 个方向上相应像素值的比较,分别得到 4 幅灰度图像。将这 4 个方向上形成的灰度图像串联在一起,构成一个基于四通道灰度顺序的图像 $F = [f'_{lu}, f'_{ru}, f'_{ld}, f'_{rd}]$ 。

对于每个方向 $s \in \{lu, ru, ld, rd\}$ 中,计算每个像素 p 对应各方向上的灰度图像 f'_s , 如式(1)所示。

$$f'_s(p) = \sum_{i=1}^8 [f(p) > f(n'_i)] \times 2^{i-1} \quad (1)$$

式中:若值 $f(p)$ 大于 $f(n'_i)$, 则 $[f(p) > f(n'_i)]$ 为 1, 否则为 0。PIOT 算法通过转换像素强度的顺序关系来增强图像中的线性结构,其变换不依赖于像素的绝对灰度值,而只取决于相对灰度值,重点着落于路面裂缝曲线结构中“比……暗”特性这个固有特点。相较于现有的裂缝图像处理方法,PIOT 算法可有效增强裂缝曲线结构的固有特

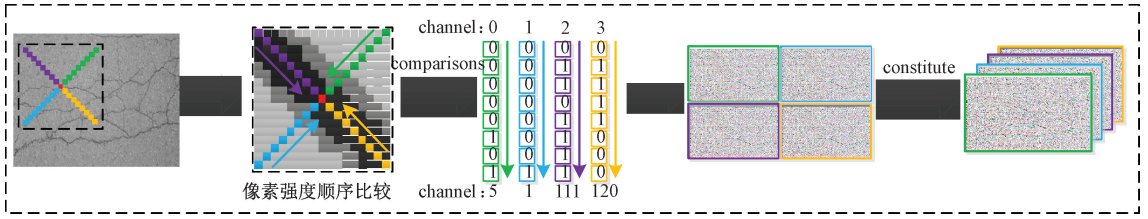


图 3 PIOT 算法变换流程图

Fig. 3 PIOT algorithm transform flowchart

征,同时抑制不相关的结构和噪声,并提高裂缝与背景像素的对比度,使分割模型更容易捕获裂缝的特征信息。

1.2 全局-局部注意力机制模块

针对边缘裂缝分割间断或识别不到的问题,本文引入了 GLAM 模块,以充分捕获路面裂缝全局与局部特征信息,增强对裂缝关键信息的关注并抑制无关信息的干扰。该模块由全局-局部注意力、多层感知器、两个批归一化层和两个附加操作组成,其结构如图 4 所示。模块中的全局-局部注意力层通过构造两个并行分支以分别提取全局和局部的裂缝特征信息,其中局部分支采用两个并行卷积层来捕获局部裂缝的细粒度特征信息,全局分支则基于窗口的多头自注意 (multi-head self-attention, MHSA) 来提取全局的裂缝特征信息,最后由全局分支提取的全局裂缝特征信息和局部分支提取的局部裂缝特征信息聚合为全局-局部裂缝特信息。

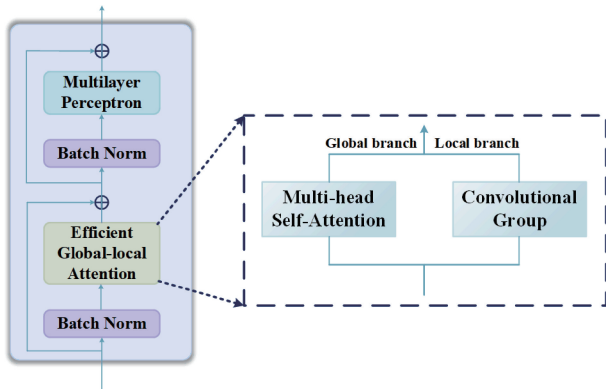


图 4 GLAM 模块结构图

Fig. 4 GLAM module structure diagram

MHSA 作为 GLAM 模块的核心,首先将原始数据通过不同的可学习线性映射生成查询矩阵 Q 、关键矩阵 K 和价值矩阵 V ,以投影到更高维的空间中,使得后续的计算能够考虑更多的信息。然后,计算查询和所有键的点积,通过缩放因子进行归一化,并通过 softmax 函数应用以获得注意力权重。最后,将这些权重乘以 V 以形成加权特征图。其具体实现公式如下:

$$Q_h = QW_h^Q, K_h = KW_h^K, V_h = VW_h^V \quad (2)$$

$$A_h = \text{SoftMax}\left(\frac{Q_h K_h^T}{\sqrt{C_H}}\right) \quad (3)$$

$$MHSA = \text{Concat}(A_1 V_1, A_2 V_2, \dots, A_H V_H) W^x \quad (4)$$

其中, x 代表输入数据, $Q_h, K_h, V_h (h = 1, 2, \dots, H)$ 分别表示每个头中的查询矩阵、键矩阵和价值矩阵, W_h^Q, W_h^K, W_h^V 分别为是相应转换的权重, A_h 代表第 h 个头部的注意力计算的结果。

1.3 基于 UNetFormer 的路面裂缝分割模型

为对经过 PIOT 算法处理的裂缝图像数据进行快速准确的分割,本文充分融合并利用 CNN 与 Transformer 网络分割裂缝的优势,提出了一种基于 UNetFormer 的路面裂缝分割模型,该模型由 CNN 的编码器和基于 Transformer 的解码器来组成,其网络结构如图 5 所示。完成对路面裂缝的高精度分割的同时减少了模型参数和计算量,有效提高模型的效率和泛化能力。

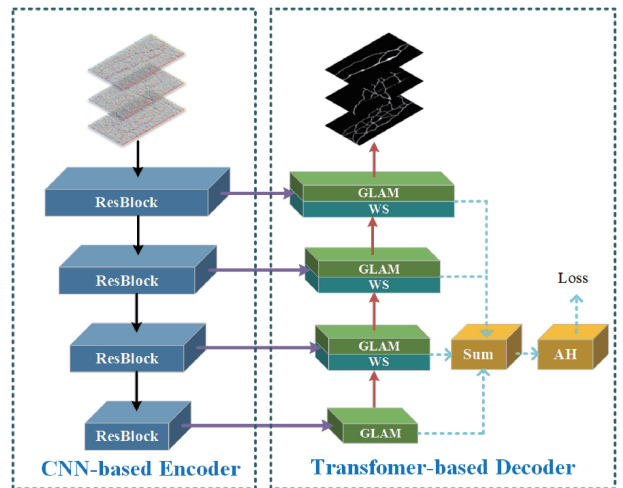


图 5 UNetFormer 网络结构图

Fig. 5 UNetFormer network structure diagram

由于 ResNet18^[17] 在实时语义分割任务中的优异表现,本文选择 ResNet18 作为编码器,以显著低的计算成本提取多尺度语义特征。其中编码器部分由 4 个阶段的 Resblocks 组成,每个阶段都以一定的比例因子对特征图进行下采样。然后将每一阶段生成的特征图都与解码器对应的特征图进行 1×1 卷积(通道数为 64),即跳跃连接,通过加权求和将残差块产生的特征与解码器 GLAM 模块产生的特征进行聚合。同时根据两个特征对分割精度的贡献有选择地对它们加权,从而学习更广泛的融合特征。

加权求和的运算公式为:

$$FF = \alpha \cdot RF + (1 - \alpha) \cdot GLF \quad (5)$$

其中, FF 代表融合特征, RF 表示 Resblocks 提取的特征, GLF 表示 GLAM 模块捕获的特征。

解码器部分则是由 4 个 GLAM 块组成, 构建出一个轻量级的基于 Transformer 的解码器。通过这种分层和轻量级的设计, 解码器能够在多个尺度上捕获全局和局部上下文的裂缝特征, 同时保持高效率。

UNetFormer 模型训练时, 本文采用使用 Sigmoid 交叉熵损失对网络参数进行优化, 输出损失为 L_i , 定义为:

$$L_i = -y_i \log(p_i) - (1 - y_i) \log(1 - p_i) \quad (6)$$

式(4)中 y_i 表示像素 i 的标签是否正确, p_i 为像素 i 的预测概率, 然后使用权重求和, 如式(7)所示。

$$L = \sum_i L_i \quad (7)$$

2 实验结果及分析

2.1 数据集

为验证本文方法的有效性, 本文使用了 3 个数据集, 分别为 CFD、Crack200 和 Crack500。

1) CFD 数据集

CFD 是第一个公开的路面裂缝集^[18], 它包含 118 张图像, 图像分辨率为 480×320 像素, 将原始数据按 7 : 3 比例划分为训练集和测试集。该数据集包含的图像光照变化、表面污渍、油腻污垢和复杂的背景纹理等多种噪声类型, 裂缝分割具有一定的挑战性。

2) Crack200 数据集

Crack200 数据集由 Wang 等^[19]提供, 其中包括 206 张 RGB 图像并和其对应的真值图, 图像分辨率为 800×600 像素。在实验过程采用垂直翻转、水平翻转、垂直水平翻转、对比度增强等方法对数据集进行了扩充, 将数据集扩增至 1 030 张, 并按 6 : 4 比例划分为训练集和测试集。

3) Crack500 数据集

Crack500 路面裂缝数据集由 Yang 等^[20]提出。数据集中包括 500 张原始的高分辨率图像 ($2\,000 \times 1\,500$ 像素), 将原始数据按 6 : 4 比例划分为训练集和测试集, 每张图像的大小调整为 320×640 像素。不同场景的裂缝图像展现了横向、纵向、龟裂等各式裂缝, 有着不同程度的分割难度。

2.2 实验环境及参数设置

实验环境为 Intel(R) Xeon(R) Gold 5115 CPU@2.40 GHz 2.39 GHz, NVIDIA Quadro P5000 GPU 16 GB, window 10 操作系统, 本文提出的方法使用 PyTorch 框架实现。在 GPU 上进行模型训练, 其中, Batch Size 设置为 16, 最大训练轮数设置为 150, 采用动态学习率以避免模型的震荡与过拟合, 并提高模型收敛速度, 其初始学习率为 1×10^{-2} , 优化器选用 RAdam。

2.3 评价指标

为了更精确地评估本文方法, 本文选取 Precision (Pr)、Recall (Re) 和 F1-score (F_1) 这 3 个指标进行衡量。其中, Pr 代表精度, Re 代表召回率, 两者衡量了模型分割的准确性^[21], 其定义公式为:

$$Pr = \frac{TP}{TP + FP} \quad (8)$$

$$Re = \frac{TP}{TP + FN} \quad (9)$$

式中: TP 、 FP 、 FN 分别代表真阳性、假阳性和假阴性的数量。对于裂缝与背景来说, TP 是正确预测为裂缝的像素点数量, FP 是错误判定为裂缝的像素点数量, FN 是错误判定为背景的裂缝像素点数量。为了获得更好的效果, Pr 和 Re 的值应尽可能的高。由于本文分割的目标为裂缝, 而裂缝在图像中的占比较小, 导致 Pr 和 Re 会出现一定的矛盾, 所以为了更好地评判模型性能引入了综合评价指标 F_1 -score^[22] 用来衡量二分类模型精确度, 同时兼顾了分类模型的 Pr 和 Re , 可看作是模型精确率和召回率的一种调和平均。其计算公式如下:

$$F_1 = \frac{2 \times Pr \times Re}{Re + Pr} \quad (10)$$

2.4 实验结果与分析

本文经计算将 8 个对角相邻像素按照计算结果从外向内进行两两比较, 形成的四通道图像。发现针对裂缝图像, 裂缝区域与非裂缝区域之间的差异性更强。与普通四通道图相较, 结果如图 6 所示, PIOT 算法在主观视觉上保留了裂缝的更多细节, 红框中为两种算法保留同一细节的对比展示, 裂缝痕迹保留更完整更清晰, 说明其鲁棒性更好。

为验证本文方法的有效性, 将本文与 UNet、FPHBN、Deeplabv3+ 进行对比, 以此验证所提出网络的有效性。

由表 1 可看出, 从本文方法在 CFD 数据集上的表现比较优秀, F_1 值为 83.4%。根据图 7, 从整体数据可看出, 所有方法在 CFD 数据集上的分割性能都不太理想, 这是因为 CFD 数据集中有大量的细裂缝图像, 增加了分割难度, 细裂缝是所有方法都面临的挑战, 而本文的方法针对细裂缝识别更准确, 更接近真值图像, 精度更高, 实验结果图如图 8(a) 所示。

由表 2 可看出, 基于 Pr 、 Re 和 F_1 分数的定量比较, 将本文方法与表中的基于 Crack200 数据集的其他两种方法进行比较。据表 2 可知本文方法在 Crack200 数据集上表现较好, Pr 的值达到了 83.5%。 F_1 的值显著高于其他两种方法, 为 82.6%。值得一提的是此数据集中的图像全部为细裂缝, 本文算法针对细小裂缝表现优秀, 实验结果如图 8(b) 所示。

由表 3 可看出, 在 Crack500 数据集上, 本文提出的方法 Pr 、 Re 和 F_1 值是最高的, 分别为 80.4%、82.5% 以及

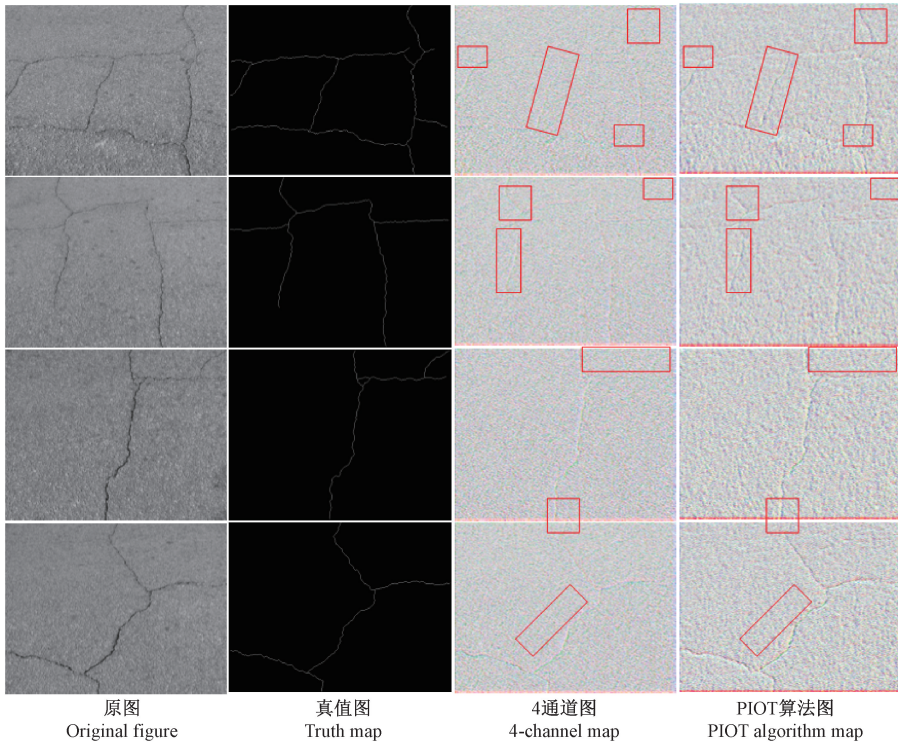


图 6 PIOT 算法结果对比图

Fig. 6 Comparison of PIOT algorithm results

表 1 不同方法在 CFD 数据集上对比实验结果

Table 1 Comparison of experimental results of different methods on CFD datasets

方法	P_r	R_e	F_1
U-Net ^[23]	0.740	0.776	0.781
FPHBN ^[20]	0.599	0.753	0.667
Deeplabv3 ^[24]	0.619	0.735	0.672
Ours	0.824	0.818	0.834

表 2 不同方法在 Crack200 数据集上对比实验结果

Table 2 Comparison of experimental results of different methods on Crack200 datasets

方法	P_r	R_e	F_1
U-Net ^[23]	0.781	0.811	0.796
FPHBN ^[24]	0.601	0.631	0.611
Ours	0.835	0.819	0.826

表 3 不同方法在 Crack500 数据集上对比实验结果

Table 3 Comparison of experimental results of different methods on Crack500 datasets

方法	P_r	R_e	F_1
U-Net ^[23]	0.760	0.762	0.767
FPHBN ^[24]	0.784	0.807	0.774
Ours	0.804	0.825	0.819

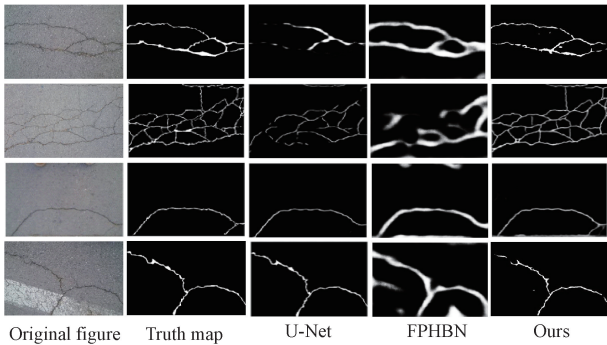


图 7 不同方法在 CFD 数据集的结果展示图

Fig. 7 Presentation of results of different methods on CFD dataset

81.9%。与 U-Net 和 FPHBN 相比较,在 F_1 上分别提升了 5.2% 和 4.5%。观察图 8(c)可知,对于龟裂的、结构复杂、裂缝较大的图像,本文算法具有更好的表现。

2.5 消融实验

为了比较本文中所提出方法中各个模块的特征提取的性能并验证各模块的性能,本文在 Crack200 数据集上进行消融实验,结果如表 4 所示。

在表 4 中,RGB 表示原图,U-NetFormer 表示原始网络,U-NetFormer(GLAM)表示加入了 GLAM 模块。由实验结果显示,相较于 RGB 图像,PIOT 算法预处理后的四通道图像经过 U-NetFormer 分割网络后,实验的各项指标

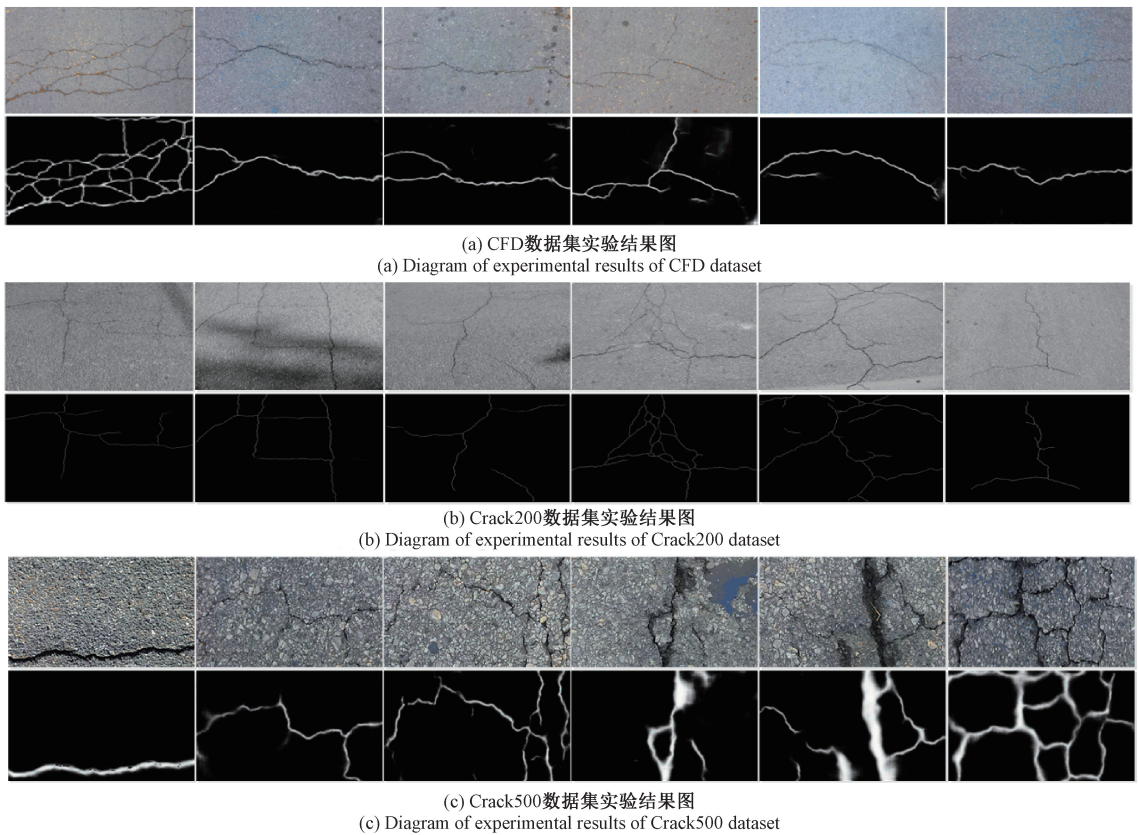


图 8 本文方法在不同数据集上的实验结果

Fig. 8 Experimental results of this paper's method on different datasets

表 4 在 Crack200 数据集上的消融实验结果

Table 4 Results of ablation experiments on the Crack200 datasets

方法	Pr	Re	F_1
RGB-UNetFormer	0.748	0.754	0.738
PIOT-UNetFormer	0.807	0.821	0.804
RGB-UNetFormer(GLAM)	0.773	0.785	0.764
PIOT-UNetFormer(GLAM)	0.835	0.819	0.826

均有所提升,其中,在 Pr 、 Re 和 F_1 上分别提升了 5.9%、6.9% 和 6.6%,说明本文所提的 PIOT 图像预处理算法能有效提高裂缝图像分割的精度。为进一步提高模型的分割性能,本文在 UNetFormer 网络中引入了 GLAM 模块,根据实验结果可看出,添加 GLAM 模块的 UNetFormer 网络要比单一的 UNetFormer 网络具有更好的分割效果,各项指标也有一定的提升,其中,RGB-UNetFormer (GLAM)相较于 RGB-UNetFormer 在 Pr 、 Re 和 F_1 上分别提升了 2.5%、3.1% 和 2.6%,PIOT-UNetFormer (GLAM)相较于 PIOT-UNetFormer 在 Pr 和 F_1 上分别提升了 2.8% 和 2.2%,只有在召回率 Re 降低了 0.2%,基本持平。综上所述,本文所提的 PIOT 图像预处理算法以及 UNetFormer(GLAM)网络模型都能有效提升裂缝分割的

精度,两者的结合最终将 F_1 提升到 82.6%。

2.6 模型参数量比较

随着网络模型的层数增多,传统的 UNet 网络模型训练所需的参数量也会越来越多,使得模型训练所需的时间也逐渐变长,降低了模型的时效性和泛化能力。将 UNet 和 Transformer 进行有效融合,在提高分割精度的同时会极大的减少计算量,将本文方法分别与 UNet、SwinUNet、GhostUNet 和 UNetFormer 模型进行参数量对比,对比结果如表 5 所示。

表 5 模型参数量对比结果

Table 5 Comparison results of model parameters

方法	参数量/MB
UNet ^[25]	31.0
SwinUNet ^[26]	41.4
GhostUNet ^[27]	13.4
UNetFormer	9.2
UNetFormer(GLAM)	11.7

由表 5 可知,本文所提模型参数量为 11.7 MB,比参数量最少的 UNetFormer 模型增加了 2.5 MB,这是由于本文模型中引入了 GLAM 模块,相较于其他模型,本

文模型的参数量仍具有较大优势,只有 UNet 网络模型参数的 37.7%,在完成路面裂缝高精度分割的同时减少了模型参数和计算量,有效提高了模型的效率和泛化能力。

3 结 论

针对裂缝像素强度的不均匀以及嘈杂的背景像素之间对比度低而导致的微小裂缝难以识别的问题,本文提出的基于像素强度顺序变换和 UNetFormer 的路面裂缝分割模型,使用 PIOT 算法对裂缝图像进行像素强度顺序转换,有效抑制噪声的同时有效增强了图像中的裂缝特征信息,为 UNetFormer 分割模型提供更多的裂缝细节特征信息;此外,通过充分融合并利用 CNN 与 Transformer 网络分割裂缝的优势,构建出 UNetFormer 裂缝分割模型,并引入 GLAM 模块,有效地增强了对细小及边缘裂缝特征的关注度,在完成对路面裂缝的高精度分割的同时减少了模型参数和计算量。最后根据实验结果表明本文所提的路面裂缝分割模型的分割精度更高,分割出的裂缝细节信息更丰富,模型预测效果稳定,具有较高的分割效率和泛化性能。

虽然本文方法在路面裂缝分割方面表现优异,但仍存在一些不足,相较于三通道的图像数据,分割模型在处理四通道的数据时消耗的时间更长,对硬件的计算能力要求更高。未来将对分割模型进行改进和优化,进一步实现路面裂缝分割网络的轻量化。

参 考 文 献

- [1] 沈思远,华蓓,黄汝维.改进 YOLOv5 的路面裂缝检测模型研究[J].电子测量技术,2023,46(21):132-142.
SHEN S Y, HUA B, HUANG R W. Research on improved YOLOv5 pavement crack detection model[J]. Electronic Measurement Technology, 2023, 46(21): 132-142.
- [2] 李可心,王钧,戚大伟.基于 G-S-G 的混凝土结构裂缝识别及监测方法[J].振动与冲击,2020,39(11):101-108.
LI K X, WANG J, QI D W. Research on crack identification and monitoring method of concrete structure based on G-S-G[J]. Journal of Vibration and Shock, 2020, 39(11): 101-108.
- [3] 翟军治,孙朝云,裴莉莉,等.多尺度特征增强的路面裂缝检测方法[J].交通运输工程学报,2023,23(1):291-308.
ZHAI J ZH, SUN ZH Y, PEI L L, et al. Pavement crack detection method based on multi-scale feature enhancement [J]. Journal of Traffic and Transportation Engineering, 2023, 23(1): 291-308.
- [4] CHA Y J, CHOI W, BÜYÜKÖZTÜK O. Deep learning-based crack damage detection using convolutional neural networks [J]. Computer-Aided Civil And Infrastructure Engineering 2017, 32(5): 361-378.
- [5] 廖延娜,豆丹阳.基于 Mask RCNN 的桥梁裂缝检测方法设计及研究[J].应用光学,2022,43(1):100-105.
LIAO Y N, DOU D Y. Design and research of bridge cracks detection method based on Mask RCNN[J]. Journal of Applied Optics, 2022, 43(1): 100-105.
- [6] 张华卫,贾文娟,张金龙,等.并联提取与特征融合注意力网络下的裂缝检测[J].电子测量技术,2022,45(10):102-111.
ZHANG H W, JIA W J, ZHANG J L, et al. Crack detection under parallel extraction and feature fusion attention network [J]. Electronic Measurement Technology, 2022, 45(10): 102-111.
- [7] SUN X Z, XIE Y C, JIANG L M. DMA-Net: DeepLab with multi-scale attention for pavement crack segmentation [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10): 18392-18403.
- [8] GUO J M, MARKONI H, LEE J D. BARNet: Boundary aware refinement network for crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 7343-7358.
- [9] WANG S W, GONG Y C, XING J L, et al. RDSNet: A new deep architecture for reciprocal object detection and instance segmentation [C]. AAAI 2020-34th AAAI Conference on Artificial Intelligence. New York: AAAI press, 2020.
- [10] 褚鸿鸽,袁华青,龙砺芝,等.基于 Transformer 的高分辨率桥梁裂缝图像级联分割方法[J].中国公路学报,2024,37(2):65-76.
CHU H H, YUAN H Q, LONG L ZH, et al. A transformer-based cascade method for segmenting bridge cracks from high-resolution images[J]. China Journal of Highway and Transport, 2024, 37(2): 65-76.
- [11] 丁威,夏哲,舒江鹏,等.基于负压吸附爬壁机器人和 Transformer 的混凝土桥塔裂缝识别检测[J].中国公路学报,2024,37(2):53-64.
DING W, XIA ZH, SHU J P, et al. Recognition and detection of concrete bridge tower cracks using a negative pressure adhesion Wall-climbing Robot and Transformer [J]. China Journal of Highway and Transport, 2024, 37(2): 53-64.
- [12] LIU H J, MIAO X, MERTZ C, et al. CrackFormer: Transformer network for fine-grained crack detection[C]. 2021 IEEE/CVF International Conference on

- Computer Vision(ICCV),2021.
- [13] 邝先验,徐姚明,雷卉,等.基于轻量级 Transformer 的隧道裂缝分割[J].铁道科学与工程学报,2024, DOI: 10.19713/j.cnki.43-1423/u.T20231768.
KUANG X Y, XU Y M, LEI H, et al. Tunnel crack segmentation based on lightweight Transformer[J]. Journal of Railway Science and Engineering, 2024, DOI: 10.19713/j.cnki.43-1423/u.T20231768.
- [14] SHI T Y, BOUTRY N, XU Y C, et al. Local intensity order transformation for robust curvilinear object segmentation[J]. IEEE Transactions on Image Processing, 2022, 31: 2557-2569.
- [15] XI D, YANG H Z, TAN B. Stereo matching algorithm based on improved census transform and minimum spanning tree cost aggregation[J]. Journal of Visual Communication and Image Representation, 2024, 98, DOI: 10.1016/j.jvcir.2023.104023.
- [16] HUANG C H, YANG J F. Improved quadruple sparse census transform and adaptive multi-shape aggregation algorithms for precise stereo matching[J]. IET Computer Vision, 2022, 16(2):159-179.
- [17] LIANG C M, LI Y W, LIU Y H, et al. Segmentation and weight prediction of grape ear based on SFNet-ResNet18 [J]. Systems Science & Control Engineering, 2022, 10(1): 722-732.
- [18] YONG S, CUI L, 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.
- [19] WANG R D, WANG H, HE Z H, et al. A novel real-time pixel-level road crack segmentation network [J]. Journal of Real-Time Image Processing, 2024,21(3).
- [20] YANG F, ZHANG L, YU S, et al. Feature pyramid and hierarchical boosting network for pavement crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(4): 1525-1535.
- [21] 胡成雪,何莉,陶健,等.邻域与梯度显著特征融合的沥青路面裂缝检测方法[J].计算机辅助设计与图形学学报,2022,34(2):245-253.
HU CH X, HE L, TAO J, et al. Asphalt pavement crack detection based on fusion of neighborhood and gradient salient features [J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(2): 245-253.
- [22] 朱代先,齐蜻蜓,车路行,等.基于亚像素级的墙面裂缝宽度检测方法研究[J].电子测量技术,2023,46(6): 166-172.
ZHU D X, QI Q T, CHE L X, et al. Research on detection method of wall crack width based on sub-pixel level[J]. Electronic Measurement Technology, 2023, 46(6): 166-172.
- [23] LIU Z Q, CAO Y W, WANG Y Z. Computer vision-based concrete crack detection using U-net fully convolutional networks [J]. Automation in Construction, 2019, 104: 129-139.
- [24] JI A, XUE X L, WANG Y, et al. An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement[J]. Automation in Construction, 2020, 114: 103176.
- [25] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. CoRR, 2015: 234-241.
- [26] CAO H, WANG Y, CHEN J, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation[C]. Computer Vision-ECCV 2022 Workshops, 2023: 205-218.
- [27] 赵志宏,郝子晔,何朋.融合注意力机制与 GhostUNet 的路面裂缝检测方法[J].电子测量技术,2023, 46(24):164-171.
ZHAO ZH H, HAO Z Y, HE P. Combining attention mechanism with GhostUNet method for pavement crack detection[J]. Electronic Measurement Technology, 2023, 46(24): 164-171.

作者简介

姚成武,学士,高级工程师,主要研究方向为政务信息化、人工智能等。

E-mail:386971864@qq.com

常琛,博士研究生,主要研究方向为计算机视觉、图像处理和风电机组故障诊断。

E-mail:13230072@stu.lzjtu.edu.cn