

文章编号:1001-7372(2024)02-0077-11

基于改进 Segformer 的混凝土桥梁表观病害 轻量级识别方法

蒋仕新^{*1}, 唐椿程¹, 杨建喜¹, 李昊², 熊元俊¹, 李韧¹, 刘新龙¹, 王笛¹

(1. 重庆交通大学 信息科学与工程学院, 重庆 400074; 2. 重庆交通大学 交通运输学院, 重庆 400074)

摘要:现有基于语义分割的混凝土桥梁表观病害识别方法具有模型参数量较大、特征提取不充分、分割精度不高等缺点。针对上述问题,提出基于改进 Segformer 的混凝土桥梁表观病害轻量级识别方法——Segformer-SP,该方法选取 MiT B0 作为编码器,并引入语义融合模块(Semantic Fusion Module, SFM)和极化自注意力机制(Polarized Self-Attention, PSA)。在 Segformer-SP 中, SFM 模块对低级特征和高级特征进行融合,以提高低级特征的语义信息量;同时,PSA 自注意力机制模块可以在分割过程中捕捉全局上下文信息,解决病害特征不够充分的问题,从而进一步提高分割效果。试验结果表明:Segformer-SP 的平均交并比(mIoU)和平均 F1 分数(mF1)相较于 Segformer-B0 分别提高了 2.41% 和 1.91%;此外,Segformer-SP 的 mIoU 和 mF1 均优于大部分主流语义分割算法;Segformer-SP 的参数量仅为 6.09×10^6 , FPS 为 56.54,更适合应用于终端检测设备。

关键词:桥梁工程;桥梁病害识别;Segformer;轻量级;特征融合

中图分类号:U446 文献标志码:A

Lightweight Detection Method of Surface Damage on Concrete Bridges Based on Improved Segformer

JIANG Shi-xin^{*1}, TANG Chun-cheng¹, YANG Jian-xi¹, LI Hao²,
XIONG Yuan-jun¹, LI Ren¹, LIU Xin-long¹, WANG Di¹

(1. School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China;
2. College of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: The existing methods for detecting surface damage on concrete bridges based on semantic segmentation suffer from several drawbacks, including large model parameters, inadequate feature extraction, and low segmentation accuracy. To address these challenges, this study proposes Segformer-SP, a lightweight surface-damage detection method for concrete bridges based on an enhanced Segformer architecture. Segformer-SP adopts the MiT B0 as the encoder and introduces two novel modules: the semantic fusion module (SFM) and polarized self-attention (PSA). In Segformer-SP, the SFM is employed to fuse low-level features with high-level features, thereby enhancing the semantic information captured by the low-level features. Additionally, PSA captures global contextual information during the segmentation process,

收稿日期:2023-06-11

基金项目:国家自然科学基金项目(62003063, 62103068);重庆市自然科学基金项目(CSTB2022NSCQ-MSX1599, cstc2020jcyj-msxmX0047);
重庆市教委科学技术研究计划项目(KJQN202000726, KJQN202100748, KJZD-M202000702)

*作者简介:蒋仕新(1992-),男,重庆大足人,副教授,工学博士,E-mail:shixinjiang@cqjtu.edu.cn。

effectively addressing the issue of insufficient feature representation for damages and further improving the segmentation accuracy. Experimental results demonstrate that Segformer-SP outperforms Segformer-B0, exhibiting a 2.41% increase in mean intersection over union (mIoU) and a 1.91% increase in mean F1-score (mF1). Moreover, Segformer-SP achieves better performance in terms of mIoU and mF1 than most state-of-the-art semantic segmentation models. Notably, Segformer-SP has a significantly reduced parameter count of only 6.09×10^6 , while maintaining a high frame-per-second (FPS) rate of 56.54, making it suitable for deployment on terminal detection equipment.

Keywords: bridge engineering; bridge damage detection; Segformer; lightweight; feature fusion

Received 11 June 2023

Funding: National Natural Science Foundation of China (62003063, 62103068); Natural Science Foundation of Chongqing (CSTB2022NSCQ-MSX1599, cstc2020jcyj-msxmX0047); Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN202000726, KJQN202100748, KJZD-M202000702)

*Corresponding author. *E-mail address:* shixinjiang@cqjtu.edu.cn.

<https://doi.org/10.19721/j.cnki.1001-7372.2024.02.007>

0 引言

随着中国基础设施建设和交通运输的不断发展,混凝土桥梁的建设规模也在不断扩大。截至2022年末,中国公路桥梁总数达到103.32万座,延长总量为8 576.49万米,其中特大桥梁8 816座,大桥15.96万座^[1]。然而,随着时间的推移,混凝土桥梁表观和内部可能会出现一定程度的损伤和缺陷,这些损伤可能导致混凝土桥梁发生坍塌等安全事故。因此,对混凝土桥梁的健康状态进行检测和评估具有至关重要的意义,能够及时发现混凝土桥梁中的损伤和缺陷,从而保障公路桥梁的安全运行。

混凝土桥梁表观病害识别是混凝土桥梁检测的重要环节,其中包括对破损、露筋和水侵蚀等方面检测^[2]。混凝土桥梁破损处可能会继续掉落混凝土,从而砸伤行人或者损坏汽车,而且桥梁破损处的混凝土进一步脱落可能会导致钢筋被空气腐蚀,从而影响桥梁结构的受力性能。长期的水侵蚀会破坏桥梁主体结构材料,扰动其承载力,缩短其使用年限。因此,对破损、露筋和水侵蚀等方面的检测是非常必要的。

传统的混凝土桥梁检测方法主要以人工巡检为主,使用搭手脚架或桥检车将检测人员送入检测区域。然而,这种方式存在着主观性强、工作环境危险、工作量大、效率低等缺点,已经逐渐不能满足人们的需求。随着电子技术的发展,通过无人机和爬壁机器人等技术可以实现桥梁表观巡检^[3],但仍需

要人工对机器所采集的照片或视频进行病害判别,效率依然较低。近年来,计算机视觉技术的高速发展为混凝土桥梁表观病害检测带来了新的解决方案。目前,基于计算机视觉的方法被广泛应用于表观病害检测^[4],主要包括表观病害分类^[5-6]、表观病害目标检测^[7-9]和表观病害语义分割^[10-11]。表观病害分类方法仅能对桥梁图片中的病害类型进行判定,无法对病害位置进行定位。表观病害目标检测方法可以识别和定位混凝土桥梁图片中的表观病害,与表观病害分类相比,不仅能够确定病害类型而且能够定位病害位置。表观病害语义分割方法不仅能够确定病害类别和位置,而且还能获取诸如面积等更加进一步的细化信息。虽然基于计算机视觉的混凝土桥梁表观病害识别方法在检测精度上具有一定优势,但是仍存在参数量大、推理速度慢等问题,而实际表观病害识别使用场景中需要快速对病害区域进行检测分析。混凝土桥梁表观病害轻量级识别方法指的是使用较少的参数和计算资源,但仍能保持较高性能和准确度方法,该方法旨在满足资源受限环境下快速、准确实现语义分割的需求,具有在嵌入式设备、移动端和实时系统等应用中的广泛潜力和前景。

因此,本文提出了一种基于改进Segformer的混凝土桥梁表观病害轻量级识别方法——Segformer-SP,该方法选取Segformer-B0作为方法基线,并在此基础上引入语义融合模块(Semantic Fusion Module, SFM)和极化自注意力机制(Polarized

Self-Attention, PSA), 用以解决低级特征中含有较少的语义信息和语义分割任务中编码器提取到的混凝土桥梁表观病害特征可能不够充分的问题。

1 基于计算机视觉的混凝土桥梁表观病害识别方法

近年来,随着人工智能的飞速发展,深度学习等机器学习算法在图像处理领域取得了显著的成果,基于计算机视觉的混凝土桥梁表观病害识别方法也得到了广泛的研究。根据检测任务的不同,混凝土桥梁表观病害识别可分为表观病害分类、表观病害目标检测和表观病害语义分割3种类型。

1.1 混凝土桥梁表观病害分类

在混凝土桥梁表观病害分类中,通常将采集到的图像划分为相同大小的图像块,然后通过机器学习等算法进行分类,以确定其中是否包含桥梁病害。Aliyari等^[12]使用基于迁移学习的卷积神经网络(Convolutional Neural Network, CNN)进行无人机桥梁裂缝检测性能的研究,试验结果显示,CNN模型能够以81%的准确率识别出裂缝元素和非裂缝元素。李良福等^[6]将CNN应用于检测无人机拍摄的桥梁裂缝图片,并结合改进的滑动窗口算法对桥梁裂缝进行检测,获得了超过97%的裂缝分类准确率。然而,将桥梁裂缝图像划分成 16×16 的单元过于细小,也会增加检测所需的时间。

1.2 混凝土桥梁表观病害目标检测

混凝土桥梁表观病害目标检测是一种利用深度学习算法在图像中识别和定位桥梁表观病害的方法。与表观病害分类任务相比,表观病害目标检测不仅能确定图像中包含的混凝土桥梁病害类型,还能够使用识别框对混凝土桥梁表观病害部分进行自主定位。Yang等^[13]在YOLOV3中引入空间金字塔池化(SPP)模块改进了YOLOV3,同时采用迁移学习训练策略,对混凝土桥梁表面病害进行识别,试验结果表明SPP模块使平均分类精度(mAP)提升了1.3%。Cha等^[14]使用Faster R-CNN(Region CNN)方法,对混凝土裂缝和不同程度的钢腐蚀、螺栓腐蚀、钢板分层等5种病害类型进行识别和定位,结果显示,5种病害类型的平均分类精度(mAP)达到了87.8%。邹俊志等^[15]提出一种基于改进YOLOV3算法的桥梁表观病害识别方法,在YOLOV3检测层中嵌入SE注意力机制模块和空间金字塔池化模块,加强了语义特征提取能力,再通过K-means算法对桥梁表观病害缺陷数据集进行

锚框聚类,其试验结果表明,改进后的YOLOV3模型相比原有模型在精确率、召回率、平均分类精度等指标上均有较大提升,其总体平均精度均值提高了5.5%。

1.3 混凝土桥梁表观病害语义分割

混凝土桥梁表观病害分类方法和混凝土桥梁表观病害目标检测方法可以为混凝土桥梁表面病害的快速检测提供技术支持。但这2种类型的混凝土桥梁表观病害识别方法并不能提供诸如面积等更加进一步细化的信息。因此,为了获得混凝土桥梁表观病害面积等定量参数,语义分割技术被应用于混凝土桥梁表观病害识别领域。Rubio等^[16]提出了一种基于FCN的混凝土桥梁分层和钢筋暴露像素级识别方法,FCN使用预训练过的VGG作为特征提取器,该方法使得分层和钢筋暴露的平均准确率分别达到89.7%、78.4%。Deng等^[17]在LinkNet中引入空洞空间金字塔池化(Atrous Spatial Pyramid Pooling, ASPP)模块,并使用预训练的ResNet34作为特征提取器,并采用加权平衡联合交集(IoU)损失函数对高度不平衡的小数据集进行精确分割,其试验结果表明该方法使得mIoU达到61.95%,明显优于其他模型。沙爱民等^[18]使用3个CNN模型级联对路面裂缝与坑槽进行检测,其中CNN1用于病害识别,CNN2与CNN3分别以CNN1输出结果进行裂缝与坑槽提取,结果表明,与传统图像处理方法相比,CNN抗噪性能与场景适应能力更好。

2 基于改进 Segformer 的混凝土桥梁表观病害轻量级识别方法

本文提出基于改进Segformer的混凝土桥梁表观病害轻量级识别方法——Segformer-SP,其结构如图1所示。为了使得算法能够部署到边缘检测设备上,选取MiT B0作为算法的编码器。

Segformer^[19]将Transformer与轻量级多层感知器(Multilayer Perceptron, MLP)解码器结合在一起。Segformer具有2个特点:①Segformer包括一个新颖的分层结构Transformer编码器——MiT(Mix Transformer),该编码器不需要位置编码,避免了位置代码的插值,同时可以输出多尺度特征;②Segformer避免了复杂的解码器,MLP解码器从不同层聚合信息,结合了局部关注和全局关注,具有更加强大的表示能力。编码器MiT仿照CNN结构,通过在不同阶段进行下采样,生成多尺度特征。

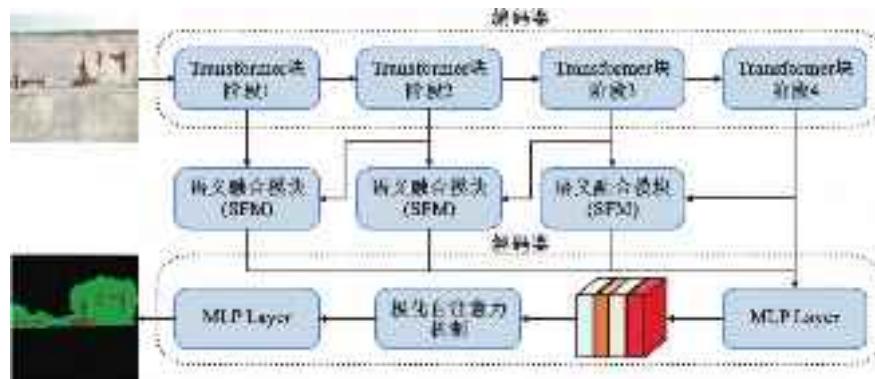


图 1 基于改进 Segformer 的混凝土桥梁表观病害轻量级识别方法

Fig. 1 Lightweight Detection Method of Surface Damage of Concrete Bridges Based on Improved Segformer

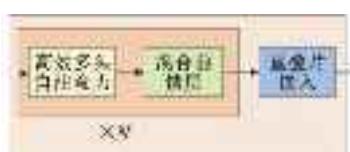


图 2 Transformer 模块

Fig. 2 Transformer Module

Transformer 模块结构如图 2 所示, 每个 Transformer 模块为了保证所提取特征的局部连续性, 采用重叠片嵌入(Overlapped Patch Merging)对输入数据进行特征提取与下采样。然后将经过重叠片嵌入的特征输入到高效多头自注意力(Efficient Multihead Self-Attention, EMSA)层和混合前馈(Mix Feed Forward, Mix-FFN)层。Segformer 解码器部分采用了简单的结构, 该轻量级解码器仅由 MLP 层(全连接层)组成, 减少了手动设计、计算需求高等问题。通过 MLP 层对多尺度特征进行通道维度变换, 统一维度, 再通过插值上采样对多尺度特征进行空间维度变换, 统一尺寸, 最后通过 MLP 层进行特征拼接与通道压缩, 再利用 1×1 卷积进行分类预测。

在原始的 Segformer 中, 引入语义融合模块(Semantic Fusion Module, SFM)将来自高级特征的更多语义信息引入到低级特征中, 从而使得低级特征包含更加丰富的语义信息, 能够在解码阶段使得低级特征与高级特征实现语义对齐, 提升分割语义一致性。同时在经过 MLP 使得多尺度特征具有相同大小之后加入极化自注意力机制(Polarized Self-attention, PSA)模块, 用以解决混凝土桥梁表观病害语义分割中编码器提取到的特征在像素级回归任务中可能不够充分的问题, 在保证计算量的情况下, 实现一种非常有效的长距离建模方式, 提高混凝土表观病害分割性能。

2.1 语义融合模块(Semantic Fusion Module, SFM)

编码器 MiT 输出多尺度特征, 包含低级特征和高级特征。低级特征包含图像的底层细节和空间信息, 如边缘、纹理等, 但其缺乏高级语义信息, 如物体的形状、类别等; 高级特征具有更抽象、语化的信息, 如物体的语义边界、上下文信息等, 但其缺乏低级细节和空间信息。然而, 由于语义层次和空间分辨率的差异, 编码器简单地将低级特征和高级特征融合, 可能导致分割效果不佳, 这是因为低级特征噪声较大, 无法为分割结果提供足够的高分辨率语义引导。为了解决该问题, 笔者借鉴了 Exfuse^[20]思路, 引入了语义融合模块(Semantic Fusion Module, SFM), 该模块结构如图 3 所示; 其将来自高级特征的更多语义信息引入到低级特征中, 从而使得低级特征包含更加丰富的语义信息。融合了高级特征的低级特征具有相对清晰的语义信息, 在解码器中该低级特征与高级特征融合时, 能够对高级特征与低级特征中的语义信息进行对齐, 提升了分割结果的语义一致性, 从而获得更精细的分割结果。

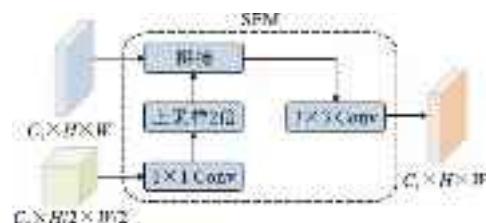


图 3 语义融合模块

Fig. 3 Semantic Fusion Module

在 SFM 模块中, 首先需要将高级特征进行线性插值, 使得高级特征和低级特征在长、宽上具有相同大小。接着, 为了避免融合所得特征过度依赖高级特征, 使用 1×1 卷积对高级特征进行降维, 以减少高级特征所占用的通道数。然后, 将降维所得的

高级特征与低级特征进行拼接,生成一个新的特征表示,该特征表示既包含了低级特征的空间细节信息,也融合了高级特征的语义信息。最后,使用 3×3 卷积对拼接所得特征进行特征融合,以进一步增强特征表示的表达能力。通过这种方式,可以有效地利用高级特征和低级特征之间的关系,获得更加准确和具有语义信息的特征表示,从而提高混凝土桥梁表观病害识别方法的检测精度。

2.2 极化自注意力机制 (Polarized Self-Attention, PSA)

语义分割是一种像素级视觉任务,需要充分考虑每个像素的回归。相较于目标检测等粗粒度视觉任务,细粒度图像分割需要在较低的计算开销下,对高分辨率输入或输出的长程依赖性进行建模,以估计像素级的语义。然而,利用 MiT 作为分割任务的编码器提取到的特征在细粒度像素级回归上可能不够充分。因此,如何获取适用于细粒度图像分割的特征非常重要。

为了解决混凝土桥梁表观病害语义分割中编码器提取到的特征在像素级回归任务中可能不够充分的问题,笔者引入了极化自注意力机制 (Polarized Self-Attention, PSA)^[21]。PSA 用于解决混凝土桥梁表观病害语义分割的像素级回归任务,相较于其他注意力机制,PSA 在通道和空间维度上保持了较高的分辨率。极化自注意力机制在保持正交方向压

缩信息损耗较小的同时,将另一个方向上的特征完全折叠,然后在注意力模块中的最小张量上用 Softmax 函数来增加注意力范围并进行动态映射。在保证计算量的情况下,实现了一种非常有效的长距离建模方式,达到提高分割性能的目的。PSA 包含仅通道自注意力和仅空间自注意力 2 个部分,如图 4 所示。在通道维度上,PSA 保持原始特征的一半维度;在空间维度上,PSA 保持原始特征的完整维度,这种设计能够减少由降维所带来的信息损失。通道注意力部分首先使用 1×1 卷积将输入特征 X 转换成 $Q(1 \times H \times W)$ 和 $V(C/2 \times H \times W)$,其中 Q 的通道维度被压缩,而 V 的通道维度保持在较高水平。由于 Q 的通道维度被压缩,因此使用 Softmax 对 Q 的信息进行加强。然后将 Q 和 K 进行矩阵相乘,并在之后添加 1×1 卷积和归一化层,将通道维度 $C/2$ 提升至 C 。最后,使用 Sigmoid 函数将所有参数限制在 0~1 之间。在空间注意力部分,首先使用 1×1 卷积将输入特征转换为 Q 和 V 。对于 Q 特征,使用全局池化压缩其空间维度;而 V 特征的空间维度则保持在较高水平 ($H \times W$)。由于 Q 的空间维度被压缩,因此使用 Softmax 对 Q 的信息进行加强。然后将 Q 和 K 进行矩阵相乘,并在之后进行特征尺寸变换和 Sigmoid 函数处理,以将所有参数限制在 0~1 之间。

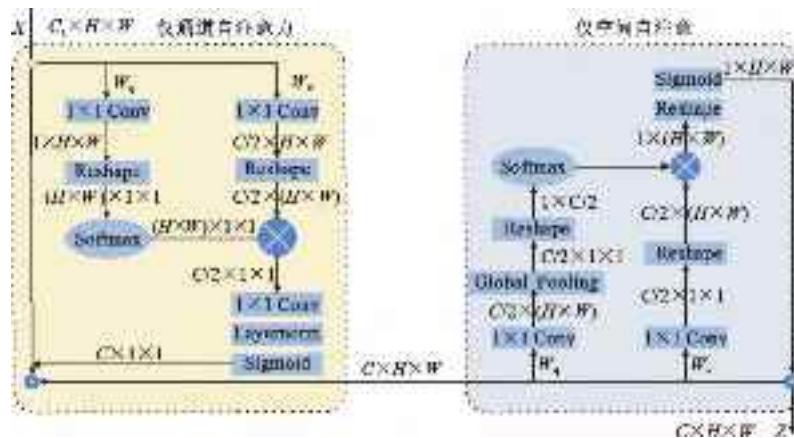


图 4 极化自注意力机制

Fig. 4 Polarized Self-attention

3 试验结果及分析

3.1 混凝土桥梁表观病害数据集

由于目前并没有公开的桥梁表观病害数据集,故首先需要人工采集混凝土桥梁表观病害图像以构建数据集。为了保证数据的多样性与泛化性,桥梁

病害图像采集自不同的桥梁巡检人员,既没有限制固定的拍摄条件,如焦距、物距、拍摄设备等,也没有固定病害图像的分辨率大小,筛选图片的原则是病害区域清晰且图像分辨率较高。此外,这些图像中还包含了不同的光照条件、遮挡阴影和拍摄角度,可以保证模型对于不同复杂环境的适应性。最终我们

手动筛选了包含破损、露筋和水侵蚀等桥梁表观病害的图片1348张。2名经过培训的学生采用通用

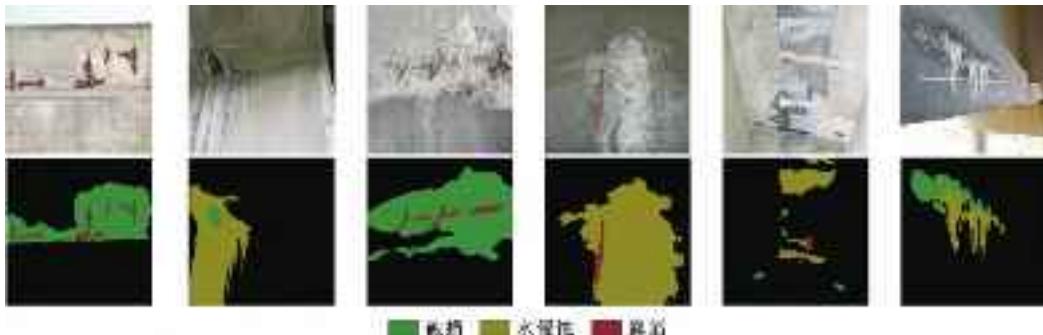


图5 混凝土桥梁表观病害图像以及其对应的标签

Fig. 5 Concrete Bridge Surface Damage Images and Their Corresponding Labels

为了防止过拟合,对数据集进行了旋转、翻转、裁剪等方式的增强。由于发现露筋与其他类别相差过大,对含有露筋的图片进行了进一步的增强,选择含有露筋的图片进行裁剪、旋转,以增加露筋类别的数量。经过预处理后,数据集包含2712张图片,最后按照8:2的比例划分数据集为训练集和验证集。

3.2 试验参数设定

试验基于pytorch深度学习框架,使用Intel(R) Xeon(R) Gold 6338 CPU@2.00 GHz、251Gi RAM、NVIDIA A40 GPU的服务器。使用Adamw优化器,动量参数为0.9,权值衰减0.01。最大学习率 5×10^{-5} ,最小学习率 5×10^{-7} ,采用Cosine Annealing作为学习率下降方式,迭代训练轮数200轮。本文中所有试验数据均是4次试验数据的平均值。

3.3 评价指标

本试验中,采用F1得分(F1-score, F1)、平均F1得分(mean F1-score, mF1)、交集联合比(Intersection Over Union, IoU)、平均交集联合比(Mean Intersection Over Union, mIoU)以及每秒帧数(Frame Per Second, FPS)作为评价指标。

查准率(Precision)是指模型预测为正例的所有样本中,预测正确(真实标签为正)样本的占比,其公式如下

$$P_r = \frac{T_p}{T_p + F_p} \quad (1)$$

式中: P_r 为查准率; T_p 和 F_p 分别为真正例和假正例的数量。

查全率(Recall)是指模型预测正样本占所有占样本比例,其公式如下

$$R_e = \frac{T_p}{T_p + F_n} \quad (2)$$

标注工具Labelme对表观病害所在区域进行标注。混凝土桥梁表观病害图像及其对应的标签见图5。



图5 混凝土桥梁表观病害图像以及其对应的标签

Fig. 5 Concrete Bridge Surface Damage Images and Their Corresponding Labels

式中: R_e 为查全率。

F1得分(F1-score, F1)是查准率和查全率的调和平均数,用于衡量模型的综合性能,其公式如下

$$F_1 = \frac{2P_r R_e}{P_r + R_e} \quad (3)$$

式中: F_1 为F1得分。

平均F1得分(Mean F1-Score, mF1)是所有类别F1得分的平均值,其公式如下

$$F_m = \frac{1}{k+1} \sum_{i=0}^k F_k \quad (4)$$

式中: F_m 为平均F1得分; F_k 为第 k 类的F1得分。

交并比(Intersection Over Union, IoU)是指预测结果与真实标签之间交集与并集之间的比例,用于衡量模型的分割能力,其公式如下

$$I = \frac{T_p}{T_p + F_n + F_p} \quad (5)$$

式中: I 为交并比; T_p 、 F_n 和 F_p 分别为真正例、假负例和假正例。

平均交并比(Mean Intersection Over Union, mIoU)是所有类别交并比的平均值,其公式如下

$$I_m = \frac{1}{k+1} \sum_{i=0}^k I_k \quad (6)$$

式中: I_m 为平均交并比; I_k 为第 k 类的交并比。

每秒帧数(Frame Per Second, FPS)是指模型每秒可以处理的图像数量。

3.4 消融试验结果分析

本文将Segformer-B0定义为基线方法。为了验证SFM模块和PSA自注意力机制模块的有效性,进行了消融试验。在消融试验中,每5轮在验证集上计算一次mIoU和mF1。根据图6所示,可以发现,在训练开始时,基线模型的评价指标相比其他模型更高。然而,随着训练的进行,基线模型在后期

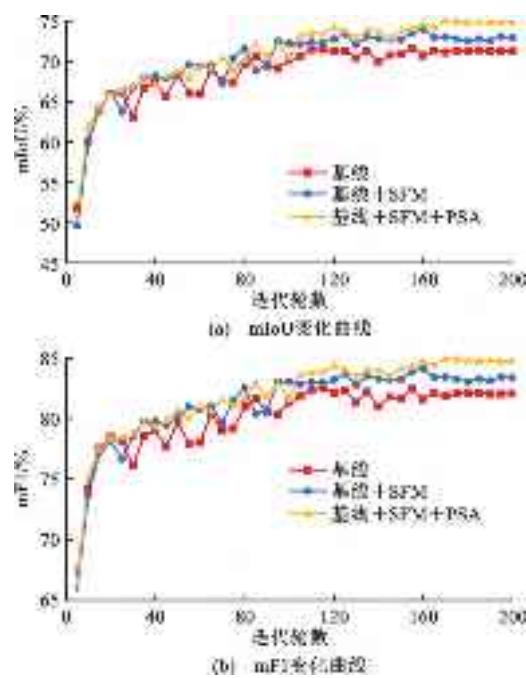


图 6 消融试验训练过程中验证集的评价指标变化曲线

Fig. 6 Change Curves of Evaluation Metrics on the Validation Set During the Ablation Experiment Training Process

的评价指标逐渐低于其他模型。由于 SFM 和 PSA 自注意力机制在开始时是随机初始化参数,没有学习到足够的特征知识,因此它们的评价指标较基线模型低。然而,在训练后期,SFM 模块和 PSA 自注意力机制模块已经学习到足够的特征知识,因此它们的评价指标优于基线模型。由于初始模型参数的随机性,模型训练开始时的评价指标会有些许振荡,但是当 epoch 达到 160 时,模型已经接近拟合,评价指标趋近于平缓。

通过表 1 分析可知,相较于基线模型,加入 SFM 模块的模型在各个评价指标上均有提升,其中 mIoU 提升了 1.43%,mF1 提升了 1.16%。这表明将来自高级特征的更多语义信息引入到低级特征中,从而使得低级特征包含更加丰富的语义信息是有效的。相较于基线模型,加入 SFM 模块和 PSA 自注意力机制模块的模型在各个评价指标上均有提升,其中 mIoU 提升了 2.41%,mF1 提升了 1.91%。这表明 PSA 自注意力机制模块,能够有效解决混凝土桥梁表观病害语义分割中编码器提取到的特征在像素级回归任务中可能不够充分的问题。分割效果消融试验可视化如图 7 所示,验证了 SFM 模块与 PSA 自注意力机制对混凝土桥梁表观病害识别的有效性。

表 1 消融试验评价指标统计

Table 1 Evaluation Metrics Statistics for Ablation Experiment

算法	水侵蚀		破损		露筋		平均	
	IoU/ %	F1/ %	IoU/ %	F1/ %	IoU/ %	F1/ %	IoU/ %	F1/ %
基线	65.14	78.90	74.94	85.67	49.01	65.77	71.59	82.24
基线+SFM	65.27	78.98	75.84	86.26	53.70	69.86	73.02	83.40
基线+SFM+ PSA	66.48	79.86	76.92	86.95	55.18	71.11	74.00	84.15

3.5 对比试验结果分析

为了验证基于改进 Segformer 的混凝土桥梁表观病害轻量级识别方法的有效性,将其与主流的语义分割算法 UNet^[22]、Deeplabv3+^[23]、HRNet^[24]、Segmenter^[25]、Segformer-B1^[19]、LinkNet^[26]、PSPNet^[27]、UNet++^[28]以及 FPN^[29]进行对比。在对比试验中,在每 5 个训练轮次后,在验证集上计算一次 mIoU 和 mF1。根据图 8 所示,可以发现:在训练初始阶段,Segformer-SP 的评价指标低于大部分模型,这是因为 Segformer-SP 的 SFM 和 PSA 模型在开始时是随机初始化参数,还未学习到足够的特征知识;虽然 Segformer-SP 在训练后期表现良好,但 Segformer-B1 模型的评价指标仍然高于 Segformer-SP。这是因为 Segformer-B1 的编码器 MiT-B1 具有更大的参数数量和更好的性能,使其在学习特征知识方面具有优势。

通过表 2 分析可知,对比 Segformer-B1,Segformer-SP 虽然 mIoU 和 mF1 分别低 0.46% 和 0.35%,但是在模型参数量和 FPS 方面更优。Segformer-SP 模型相较于 Segformer-B1 模型,参数量低 7.36×10^6 且 FPS 高 13.53。Segformer-SP 模型在实际应用中需要的计算资源更少,更加轻量级,可以更好地适应一些资源受限的场景。Segformer-SP 的 mIoU 和 mF1 均优先于主流的语义分割算法 UNet、Deeplabv3+、HRNet、Segmenter、LinkNet、PSPNet、UNet++^[25] 以及 FPN;且 Segformer-SP 模型的 FPS 更高,说明 Segformer-SP 模型可以更快地处理图像,对于实时应用来说表现更好。对比试验效果可视化如图 9 所示,可以发现:Segformer-SP 与对比方法相比,在对混凝土桥梁表观病害进行检测时不仅整体的分割精度更高,对小目标的分割效果更好,不同类别的分割边界也更加清晰。

3.6 鲁棒性试验结果分析

本文使用的混凝土桥梁表观病害数据集仅包括水侵蚀、破损和露筋等 3 种常见缺陷,而裂缝也是混

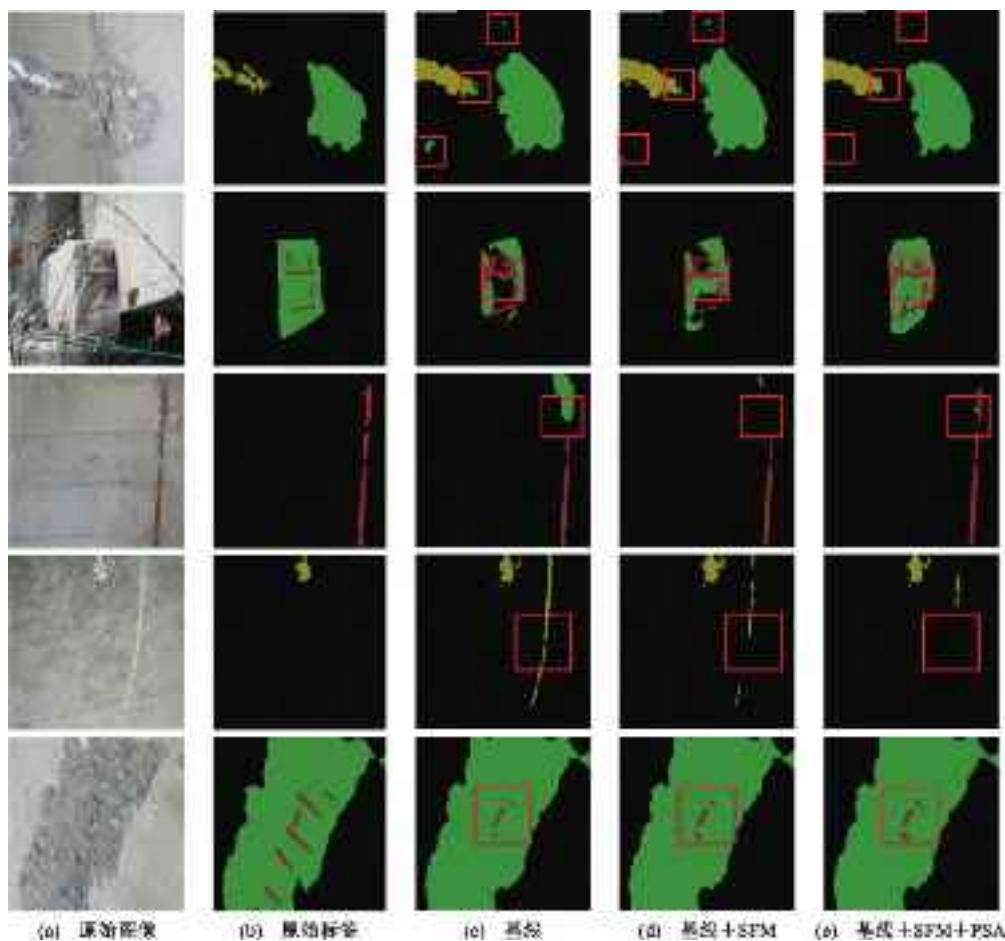


图 7 消融试验效果可视化

Fig. 7 Visualization of Ablation Experiment Results

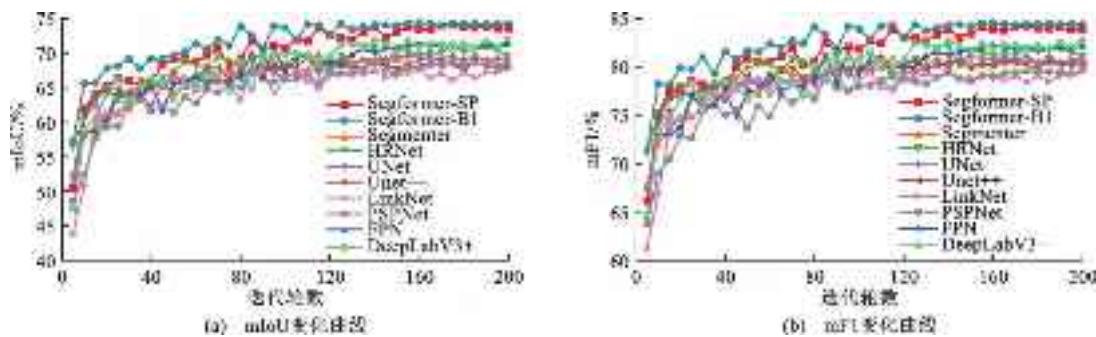


图 8 对比试验训练过程中验证集的评价指标变化曲线

Fig. 8 Change Curves of Evaluation Metrics on the Validation Set During the Comparative Experiment Training Process

混凝土桥梁表观病害中常见的类型之一。为了验证 Segformer-SP 具有较强的鲁棒性和对裂缝也具有良好的检测性能,笔者在公开的混凝土裂缝数据集 GAPs384^[30]上对 Segformer-SP 进行了验证,该数据集包含裂缝图像 509 张。通过表 3 分析可知,Segformer-SP 的裂缝 IoU 比 SegFormer-B1、HRNet、Deeplabv3+、UNet 分别高 0.17%、0.05%、0.75%、0.26%。Segformer-SP 的裂缝 F1 比 SegFormer-B1、HRNet、Deeplabv3+、UNet 分别高

0.18%、0.05%、0.75%、0.27%。可见,Segformer-SP 对裂缝的检测精度优于目前主流的语义分割方法,说明 Segformer-SP 不仅适用于对常见的水侵蚀、破损和露筋的表观病害识别,还适用于对裂缝的识别。然而,分析表 3 数据也可以发现,Segformer-SP 相比于其他模型在裂缝数据集上的提升并不是很大。这是因为该数据集相对较小,无法提供足够多样化的样本,无法涵盖整个裂缝数据空间的各种变化和模式以及模型在小数据集上学习过程中过度

表2 对比试验评价指标统计

Table 2 Evaluation Metrics Statistics for Comparative Experiment

算法	主干网络	mIoU/ %	mF1/ %	FPS	参数量/ 10^6
Segformer-SP	Mit_B0	74.00	84.15	56.54	6.09
Segformer-B1	Mit_B1	74.46	84.50	43.01	13.66
Segmenter	Vit_Small	69.82	80.50	46.41	25.26
HRNet	HRNetV2-W18	70.51	81.23	14.95	9.64
UNet	ResNet50	68.36	79.88	38.74	32.52
UNet++	ResNet50	68.46	80.14	13.31	48.99
LinkNet	ResNet50	67.57	78.53	31.69	31.18
PSPNet	ResNet50	68.99	79.85	35.59	46.71
FPN	ResNet50	70.99	81.75	37.46	26.12
Deeplabv3+	ResNet50	70.57	81.43	38.43	26.68

依赖于嘈杂的数据,也导致不同模型之间的差异可能在该数据集中无法得到充分的展现^[31]。鲁棒性试

表3 鲁棒性试验评价指标统计

Table 3 Evaluation Metrics Statistics for Robustness Experiment

算法	主干网络	IoU/%	F1/%
Segformer-SP	Mit_B0	41.67	58.83
SegFormer	Mit_B1	41.50	58.65
HRNet	HRNetV2-W18	41.62	58.78
Deeplabv3+	ResNet50	40.92	58.08
UNet	ResNet50	41.41	58.56

验可视化如图10所示,可以发现:Segformer-SP与对比方法相比,在对裂缝分割效果上更加准确。

4 结语

(1)为了实现混凝土桥梁表观病害高精度、轻量化识别,本文在Segformer-B0的基础上,提出了基于改进Segformer的混凝土桥梁表观病害轻量级识别方法——Segformer-SP。

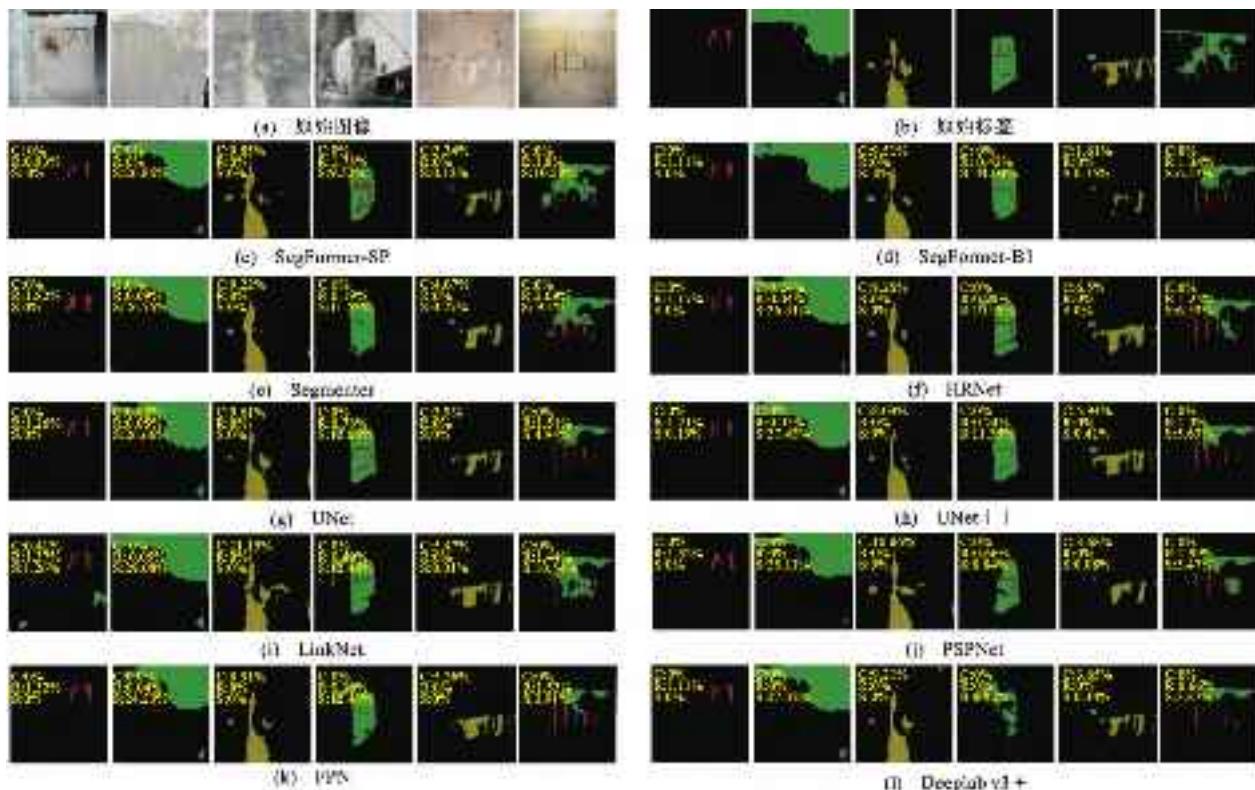


图9 对比试验效果可视化

Fig. 9 Visualization of Comparative Experiment Results

(2)针对现有语义分割模型存在编码器简单地将低级特征和高级特征融合,而低级特征噪声较大无法为分割结果提供足够的高分辨率语义引导,导致分割精度较低的问题,本文在Segformer-B0基础上引入了SFM模块,该模块能够通过将来自高级特征的更多语义信息引入低级特征,使得低级特征

包含更丰富的语义信息,在解码器阶段,高级特征与含有丰富的语义信息的低级特征进行语义对齐,从而提升分割精度。

(3)针对语义分割模型中解码器对其提取到的特征在像素级回归任务中可能不够充分的问题,本文在Segformer-B0基础上引入PSA自注意力机制,

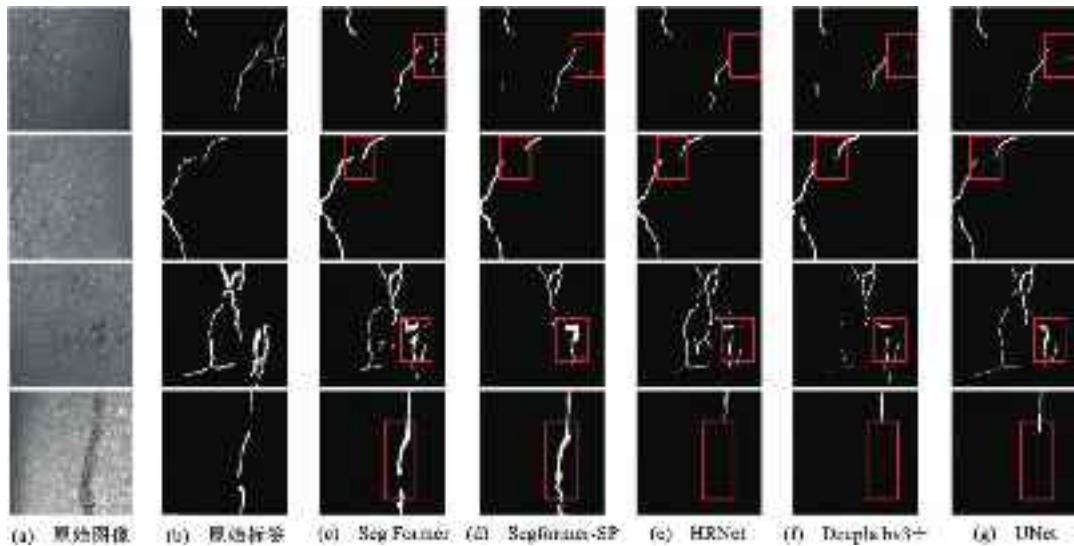


图 10 鲁棒性试验效果可视化

Fig. 10 Visualization of Robustness Experiment Results

该自注意力机制能在保证计算量的情况下,实现了
一种非常有效的长距离建模方式,达到提高分割性
能的目的。

(4)在消融试验中对SFM模块以及PSA自
注意力的性能进行了验证与分析;并且在自建混凝土
桥梁表观病害语义分割数据集以及公开裂缝数据集
上,将Segformer-SP分别与Segformer、Segmenter、
HRNet、UNet++、LinkNet、PSPNet、FPN、Deep-
labv3+进行对比与分析。结果表明Segformer-SP
在混凝土桥梁表观病害识别方面具有较强的鲁棒性
和较高的准确性,能在较小参数量的情况下实现桥
梁表观病害的高效快速检测。

(5)基于改进Segformer的混凝土桥梁表观病
害轻量级识别方法,可以快速、准确地检测和识别桥
梁表观病害,包括破损、水侵蚀和露筋等。这有助于
及早发现结构问题并采取适当的维修和保护措施,
延长混凝土桥梁的使用寿命。

(6)下一步计划将对Segformer-SP进行剪枝,
使得该方法能够在较小参数量的情况下达到更高的
准确率;然后将其部署到混凝土桥梁健康检测无人
机上,使其可以提高安全性、检测效率和准确性,及
早发现结构问题,为混凝土桥梁维护和管理提供科
学依据,保证混凝土桥梁的可靠性和持久性。

参考文献:

References:

- [1] 交通运输部.2022年交通运输行业发展统计公报[N].中国交
通报,2023-06-16(2).
Ministry of Transport. Statistical bulletin of transport indus-

try development in 2022 [N]. China Communications News, 2023-06-16 (2).

- [2] JTG/H11—2004, 公路桥涵养护规范[S].
JTG/H11—2004, Code for Maintenance of Highway Bridges
and Culvers [S].
- [3] 唐文. 钢筋混凝土拱桥外观智能检测与评估技术研究[D].
重庆:重庆交通大学,2022.
TANG Wen. Research on intelligent detection and evaluation
technology of reinforced concrete arch bridge appearance
[D]. Chongqing: Chongqing Jiaotong University, 2022.
- [4] 邓露,褚鸿鹄,龙丽芝,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 Lu, CHU Hong-hu, LONG Li-zhi, et al. Review of
deep learning-based crack detection for civil infrastructures
[J]. China Journal of Highway and Transport, 2023, 36
(2): 1-21.
- [5] XU H Y, SU X, WANG Y, et al. Automatic bridge crack
detection using a convolutional neural network [J]. Applied
Sciences, 2019, 9 (14): 2867.
- [6] 李良福,马卫飞,李丽,等.基于深度学习的桥梁裂缝检测算
法研究[J].自动化学报,2019,45(9):1727-1742.
LI Liang-fu, MA Wei-fei, LI Li, et al. Research on detection
algorithm for bridge cracks based on deep learning [J]. Acta
Automatica Sinica, 2019, 45 (9): 1727-1742.
- [7] SONG W Y, CAI S, GUO H X, et al. Bridge apparent dam-
age detection system based on deep learning [C] //
TALLÓN-BALLESTEROS A J. Fuzzy Systems and Data
Mining V -Proceedings of FSDM 2019. Kitakyushu City, Ja-
pan: IOS Press, 2019: 475-480.
- [8] ZHANG C B, CHANG C C, JAMSHIDI M. Concrete bridge
surface damage detection using a single-stage detector [J].
Computer-Aided Civil and Infrastructure Engineering, 2020,
35 (4): 389-409.
- [9] 晏班夫,徐观亚,栾健,等.基于Faster R-CNN与形态法的

- 路面病害识别[J]. 中国公路学报, 2021, 34(9): 181-193.
- YAN Ban-fu, XU Guan-ya, LUAN Jian, et al. Pavement distress detection based on faster R-CNN and morphological operations [J]. China Journal of Highway and Transport, 2021, 34 (9): 181-193.
- [10] NARAZAKI Y, HOSKERE V, YOSHIDA K, et al. Synthetic environments for vision-based structural condition assessment of Japanese high-speed railway viaducts [J]. Mechanical Systems and Signal Processing, 2021, 160: 107850.
- [11] TANG C C, JIANG S X, LI H, et al. Lightweight concrete crack segmentation method based on Deeplabv3+ [C] // SHEN L, ZHONG G. Proc SPIE 12754, Third International Conference on Computer Vision and Pattern Analysis. Kuala Lumpur: ICCPA, 2023: 399-404.
- [12] ALIYARI M, DROGUETT E L, AYELE Y Z. UAV-based bridge inspection via transfer learning [J]. Sustainability, 2021, 13 (20): 11359.
- [13] YANG J X, LI H, HUANG D E, et al. Concrete bridge damage detection based on transfer learning with small training samples [C] // IEEE. 2021 7th International Conference on Systems and Informatics (ICSAI). Chongqing, China: IEEE, 2021: 1-6.
- [14] CHA Y J, CHOI W, SUH G, et al. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33 (9): 731-747.
- [15] 邹俊志, 杨建喜, 李昊, 等. 复杂背景下基于改进 YOLO v3 算法的桥梁表观病害识别[J]. 铁道科学与工程学报, 2021, 18(12): 3257-3266.
- ZOU Jun-zhi, YANG Jian-xi, LI Hao, et al. Bridge apparent damage detection based on the improved YOLO v3 in complex background [J]. Journal of Railway Science and Engineering, 2021, 18(12): 3257-3266.
- [16] RUBIO J J, KASHIWA T, LAITEERAPONG T, et al. Multi-class structural damage segmentation using fully convolutional networks [J]. Computers in Industry, 2019, 112: 103121.
- [17] DENG W L, MOU Y L, KASHIWA T, et al. Vision based pixel-level bridge structural damage detection using a link ASPP network [J]. Automation in Construction, 2020, 110: 102973.
- [18] 沙爱民, 童峥, 高杰. 基于卷积神经网络的路表病害识别与测量[J]. 中国公路学报, 2018, 31(1): 1-10.
- SHA Ai-min, TONG Zheng, GAO Jie. Recognition and measurement of pavement disasters based on convolutional neural networks [J]. China Journal of Highway and Transport, 2018, 31 (1): 1-10.
- [19] XIE E Z, WANG W H, YU Z D, et al. SegFormer: simple and efficient design for semantic segmentation with transformers [EB/OL]. 2021, arXiv: 2105.15203. <https://arxiv.org/abs/2105.15203.pdf>
- [20] ZHANG Z L, ZHANG X Y, PENG C, et al. ExFuse: Enhancing feature fusion for semantic segmentation [C] // LAURA L, STEFAN R. Computer Vision-ECCV 2018. Munich, Germany: Springer International Publishing, 2018: 273-288.
- [21] LIU H J, LIU F Q, FAN X Y, et al. Polarized self-attention: Towards high-quality pixel-wise mapping [J]. Neurocomputing, 2022, 506: 158-167.
- [22] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C] // NAVAB N, HORNEGGER J, WELLS W M, et al. Lecture Notes in Computer Science. Munich, Germany: Springer International Publishing, 2015: 234-241.
- [23] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C] // FERRARI V, HEBERT M. Computer Vision-ECCV 2018. Munich, Germany: Springer International Publishing, 2018: 833-851.
- [24] SUN K, ZHAO Y, JIANG B R, et al. High-resolution representations for labeling pixels and regions [EB/OL]. 2019, arXiv: 1904.04514. <https://arxiv.org/abs/1904.04514.pdf>
- [25] STRUDEL R, GARCIA R, LAPTEV I, et al. Segmenter: Transformer for semantic segmentation [C] // IEEE. Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE, 2021: 7262-7272.
- [26] CHAURASIA A, CULURIELLO E. LinkNet: exploiting encoder representations for efficient semantic segmentation [C] // IEEE. 2017 IEEE Visual Communications and Image Processing (VCIP). St. Petersburg, FL: IEEE, 2017: 1-4.
- [27] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C] // IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 2881-2890.
- [28] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHSH N, et al. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation [J]. IEEE Transactions on Medical Imaging, 2020, 39 (6): 1856-1867.
- [29] MARTINSSON J, MOGREN O. Semantic segmentation of fashion images using feature pyramid networks [C] // IEEE. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul: IEEE, 2019: 1-4.
- [30] EISENBACH M, STRICKER R, SEICHTER D, et al. How to get pavement distress detection ready for deep learning? A systematic approach [C] // IEEE. 2017 International Joint Conference on Neural Networks (IJCNN). Anchorage, USA: IEEE, 2017: 2039-2047.
- [31] YING X E. An overview of overfitting and its solutions [J]. Journal of Physics: Conference Series, 2019, 1168: 022022.