

DOI: 10.11835/j.issn.2096-6717.2023.148



开放科学(资源服务)标识码 OSID:



基于无人机图像与深度学习的高原地区 隧道洞门墙病害检测方法

车博文¹, 包卫星¹, 郭强², 潘振华¹, 卢汉青¹, 尹严¹

(1. 长安大学公路学院, 西安 710064; 2. 新疆交通建设管理局, 乌鲁木齐 830002)

摘要:针对修建在高原恶劣环境下的隧道洞门墙病害频发及传统人工病害检测方法效率低下、危险性高等问题,提出基于无人机图像与深度学习的高原区隧道洞门墙病害检测方法。首先使用无人机航拍采集新疆某高原地区隧道洞门墙病害图像,建立多病害语义分割数据集;随后以语义分割模型 DeeplabV3+ 为基础,提出以 MobileNetV2 作为主干特征提取网络的改进模型 TP-DeeplabV3+,以减少模型参数;采用 Focal Loss 作为损失函数,以解决病害图像中的类别不平衡问题;添加 CA 注意力机制,以提升模型分割性能;最后提出病害量化方法。结果表明,TP-DeeplabV3+ 在测试集上可以达到 88.37% 和 94.93% 的 mIoU 和 mPA,模型体量压缩了 88.83%;提出的病害量化方法对于病害覆盖率的绝对误差不超过 0.3%,相对误差维持在 7.31% 以下。相比传统方法,该方法安全、准确地实现了高原恶劣环境下的隧道洞门墙智能化检测。

关键词:隧道洞门墙病害;深度学习;语义分割;无人机;高原地区

中图分类号:U457.2 **文献标志码:**A **文章编号:**2096-6717(2025)05-0086-11

Disease detection method of tunnel portals in plateau region based on UAV images and deep learning

CHE Bowen¹, BAO Weixing¹, GUO Qiang², PAN Zhenhua¹, LU Hanqing¹, YIN Yan¹

(1. School of Highway, Chang'an University, Xi'an 710064, P. R. China; 2. Xinjiang Transportation Construction Administration, Urumqi 830002, P. R. China)

Abstract: In light of the prevalent diseases of tunnel portals built in the harsh environment of the plateau region and the problems of low efficiency and high risk associated with traditional manual disease detection methods, a novel disease detection method for tunnel portals in the plateau region based on Unmanned Aerial Vehicle (UAV) image and deep learning was proposed. Firstly, an UAV was used to collect the disease images of a tunnel portal in the plateau region of Xinjiang, and a multi-disease semantic segmentation dataset was

收稿日期:2023-09-05

基金项目:新疆重大科技专项(2020A03003-7);陕西省自然科学基金基础研究计划(2021JM-180);中央高校基本科研业务费(领军人才计划)(300102211302)

作者简介:车博文(1999-),男,主要从事寒区结构健康监测研究,E-mail:cbwcbwcbw2@126.com。

包卫星(通信作者),男,博士,教授,E-mail:baowx@chd.edu.cn。

Received: 2023-09-05

Foundation items: Major Science and Technology Projects in Xinjiang (No. 2020A03003-7); Shaanxi Province Natural Science Basic Research Project (No. 2021JM-180); Basic Scientific Research Funds of Central Universities (Leading Talents Program) (No. 300102211302)

Author brief: CHE Bowen (1999-), main research interest: structural health monitoring in cold regions, E-mail: cbwcbwcbw2@126.com.

BAO Weixing (corresponding author), PhD, professor, E-mail: baowx@chd.edu.cn.

constructed. Then, based on DeeplabV3+, an improved model TP-DeeplabV3+ was proposed, which used MobileNetV2 as the backbone feature extraction network to reduce model parameters; Used Focal Loss as the loss function to solve the category imbalance problem in disease images; Used the CA attention mechanism to improve the segmentation performance; and proposed the disease quantification method. Experiment results show that TP-DeeplabV3+ attains 88.37% and 94.93% of mIoU and mPA on the test set, respectively. Furthermore, the model volume is reduced by 88.83%. The absolute error of the proposed disease quantification method for disease coverage rate is less than 0.3%, and the relative error is maintained below 7.31%. Compared with the traditional manual method, the proposed method facilitates the intelligent detection of tunnel portal safely and accurately in plateau region.

Keywords: tunnel portal disease; deep learning; semantic segmentation; unmanned aerial vehicle (UAV); plateau regions

公路隧道是最重要的道路交通基础设施之一。修建在高原地区的隧道,受特殊地理环境的影响,不可避免地遭遇各种病害的侵蚀,包括隧道衬砌开裂、渗漏水挂冰、洞门墙混凝土开裂和剥落等,这些病害会使隧道结构遭受不同程度的破坏,严重影响隧道结构的服役性能^[1]。因此,对高原地区隧道病害的日常检测至关重要。

目前,对隧道的病害检测还是以人工目视巡查为主。然而,在高原环境下,此方法存在很多缺陷,如人工决策中的失误,在高海拔寒冷、缺氧条件下的高劳动力成本等。近年来,随着深度学习技术的蓬勃发展,基于计算机视觉的缺陷检测技术迎来了新的机遇。作为深度学习中最具代表性的算法之一,卷积神经网络(CNN)能从各种类型的数据中学习并自动提取数据特征,这使得CNN在工程缺陷检测领域表现突出。研究者已经将基于CNN的深度学习技术应用于隧道病害智能化检测中,并取得了一定的成果^[2-4],但以上研究的主要关注点都是隧道内部衬砌病害,并未考虑到隧道洞门墙的健康状态。而长期暴露在高原恶劣环境下的隧道洞门墙,其产生各种病害的几率更大,对于整个隧道结构健康的影响不容忽视。

隧道洞门墙可以视作混凝土结构。目前,基于深度学习的混凝土结构病害检测方法可分为3大类:图像分类、目标检测及语义分割。相比前两者,语义分割可以在像素级尺度上对目标进行定位并精细地提取病害形状。在众多经典语义分割模型中,分割性能较好的U-net和DeeplabV3+及其改进模型已被广泛应用于工程领域^[5-6]。Fu等^[7]提出了一种改进的DeeplabV3+模型,其在结构中引入密集连接的空洞金字塔池化模块,实现了不同环境条件下的桥梁混凝土裂缝分割;考虑到仅针对混凝土裂缝进行检测具有一定局限性,饶勇成等^[8]建立了一个包含针对混凝土结构多病害的数据集,并提出了

一种改进U-net模型,实现了混凝土结构多病害像素级分割。然而,以上方法仍需要检测人员去现场人工采集图像,效率较低且不确定因素多。鉴于此,基于无人机(Unmanned Aerial Vehicle)图像的结构智能检测方法研究受到重视。使用无人机进行结构检测具有轻巧灵活、自动化程度高等特点^[9],并能结合深度学习技术,极大降低人工成本。丁威等^[10]提出了一种多尺度特征融合结构(Crack-FPN),用于多种结构裂缝的图像分割,并提出了一种改进方法,对无人机云台相机的图像尺度场进行标定,以完成裂缝的定位和定量测量;余加勇等^[11]采用无人机航摄收集桥梁结构混凝土裂缝及锈蚀高清图像,并选用Mask R-CNN进行病害识别,实验结果表明,其方法可以实现较高的精度。以上研究在基于无人机图像的病害检测方面提出了思路方法,但大多基于常规算法进行检测,而高原地区的隧道洞门墙背景复杂,干扰因素多,仅使用常规算法无法保证检测精度。

笔者提出一种基于无人机图像与深度学习的高原地区隧道洞门墙病害检测方法,研究流程如图1所示。

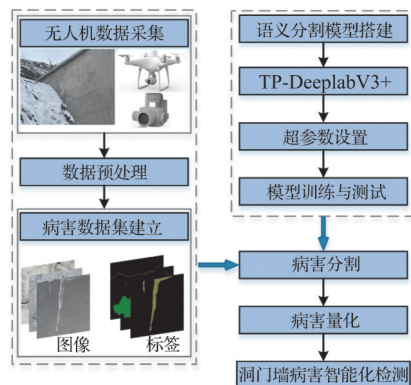


图1 隧道洞门墙病害智能检测化流程

Fig. 1 Disease intelligent detection process for tunnel portal

1 病害图像收集

1.1 依托工程概况

以G219线上的某隧道工程为依托,该隧道位于新疆喀什地区昆仑山腹地,海拔4 800余m,长2 731 m,线路穿越多年冻土,如图2所示。该隧址区属寒冷干旱高原气候,空气干燥,日照时间长,年温差及日温差大,年平均气温仅 -4°C 左右。

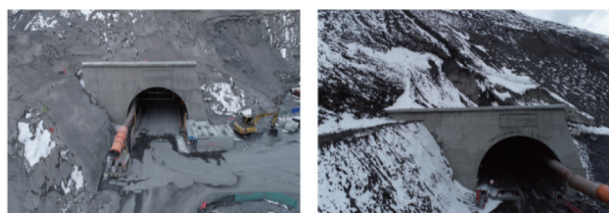


图2 隧道洞门所处环境

Fig. 2 Environment of the tunnel portal

此外,隧道洞口处浅埋段上部还存在松散坡堆积体,仰坡上部山体节理裂隙发育,有多处岩体碎落,如图3(a)、(b)所示。

基于以上恶劣的环境及地质条件,该隧道洞门结构的健康状态就显得尤为重要。然而,在冻融循环、低气压及外荷载等因素的耦合作用下,洞门墙混凝土很容易产生各种病害,如图3(c)、(d)所示,并且已产生的病害之间可能还会相互促进,造成进一步恶化,为整个隧道结构带来安全隐患。

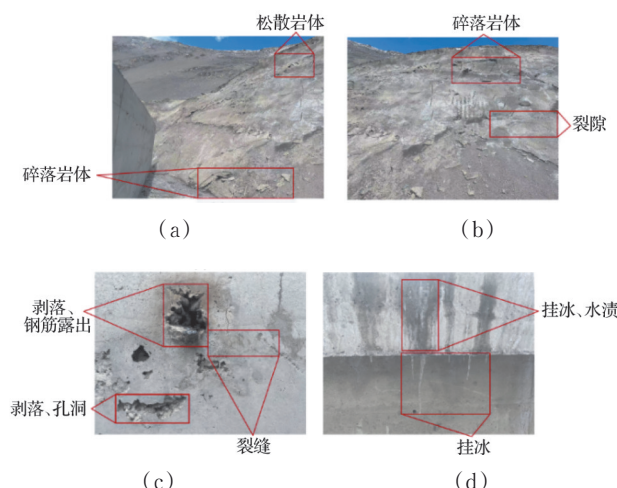


图3 洞门墙病害

Fig. 3 Tunnel portal wall diseases

1.2 无人机图像采集

用无人机收集洞门墙病害图像,如图4所示。采用大疆DJI PHANTOM 4 RTK作为飞行平台,下置可控制旋转角度的云台相机,转动范围为 -90° 至 $+30^{\circ}$,控制精度可达到 0.02° ,搭载1英寸2 000万像素的CMOS传感器。

用无人机收集了从2021年7月至2023年6月



图4 无人机图像采集

Fig. 4 UAV image acquisition

的洞门墙病害图像。在图像采集过程中,为保证较高的检测精度以及航拍安全,将无人机的飞行速度控制在 10 km/h 以下,维持拍摄角度始终垂直于洞门墙面,设置两者之间的拍摄距离为 $1\sim 3\text{ m}$ 。为了避免采集图像时受积雪覆盖影响,选择在该区域每日温度最高的 $15:00\sim 16:00$ 之间进行无人机航拍。共收集了375张分辨率为 $5\,472\times 3\,648\times 3$ 像素的洞门墙病害图像,其中主要包括3种病害:裂缝、剥落及挂冰。

1.3 数据集制作

图像采集完毕后,使用Labelme标注软件对图像中的病害进行像素级人工标注。标注时对于不同病害的分辨依据如下:

1)裂缝:包括狭长裂缝、交叉裂缝、宽裂缝、细微裂缝及不连续裂缝。

2)剥落:包括混凝土表面碎片脱落及较深的孔洞(不包含蜂窝状小孔)。

3)挂冰:包括表面产生的大范围结冰、水渍及悬挂于墙体边缘处的冰柱。

此外,考虑到使用大尺寸图像训练会给GPU带来较大负荷,影响训练效果,使用 512×512 大小的滑动窗口对原图像进行裁剪,在裁剪过程中剔除掉不包含病害的图片。图5展示了部分裁剪后的病害图像及标签。

为了避免模型在训练过程中产生过拟合,通过数据增强对数据集作进一步扩充。数据增强能使样本更加贴近真实世界的变化,图6展示了几个数

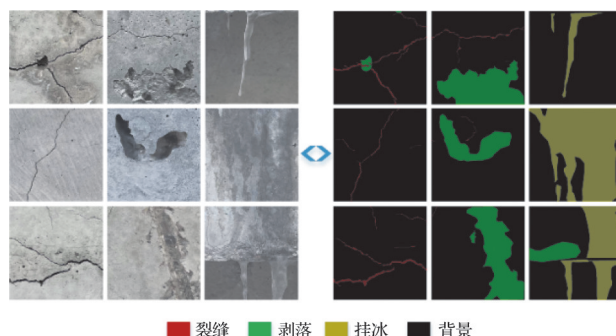


图5 洞门墙病害图像及标签

Fig. 5 Images and labels of tunnel portal wall diseases

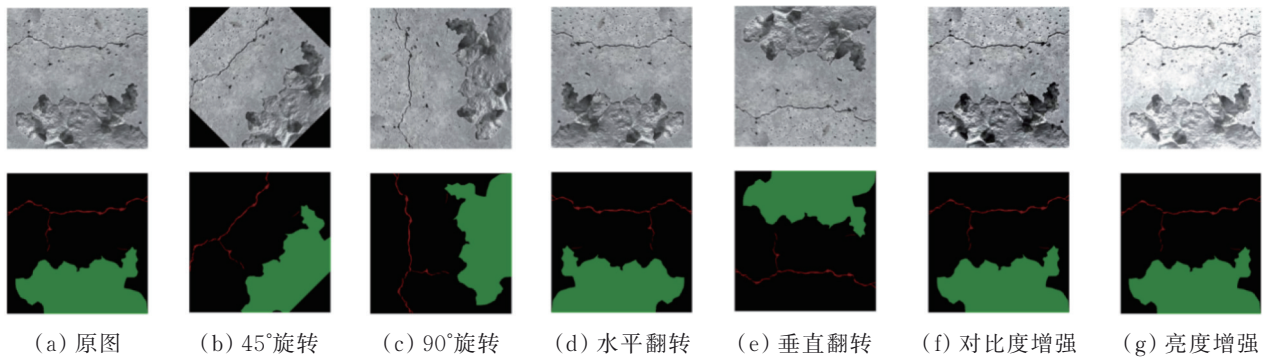


图6 数据增强

Fig. 6 Data augmentation

据增强后的病害样本图像及标签图像。最后得到4 800张病害图像及标签图像,随机选取400张病害图像作为测试集,并将剩下的数据集以4:1的比例随机划分为训练集和验证集。

2 洞门墙病害检测及量化方法

2.1 DeeplabV3+语义分割模型

DeeplabV3+[12]模型是Deeplab系列中的最新版本,其在模型架构中引入ASPP(空洞空间金字塔池化)模块进行多尺度特征提取,通过空洞卷积平衡模型的训练精度和耗时,已在多个领域中展现出优秀的分割性能。

DeeplabV3+整体呈编码-解码结构。在编码路径中,输入模型的图像首先采用主干特征提取网络,对原图像分别进行2次下采样和4次下采样,提取出两个有效特征层。经过4次下采样的有效特征层会通过一个ASPP模块,并行使用不同采样率(6、12、18)的空洞卷积对有效特征层进行多尺度语义特征提取,之后将不同尺度的特征拼接在一起,得到包含丰富语义信息的高级特征层(high_level_

features)。在解码路径中,经过2次下采样的有效特征层在经过一个 1×1 的卷积后,得到一个包含边界及纹理等细节信息的低级特征层(low_level_features),将其与经过4次双线性插值上采样后的高级特征层进行特征拼接,拼接特征层会通过 3×3 卷积与插值上采样逐渐恢复到原有的空间维度,实现像素级语义分割。

DeeplabV3+结构中的空洞卷积和ASPP模块提高了模型对于多尺度目标的特征提取及定位能力[13],可较好适用于背景复杂且目标尺度不一的高原地区洞门病害图像。但这些模块也可能导致精细局部信息的丢失,降低相邻像素的相关性[14],增加检测误差。故以DeeplabV3+作为基础框架,提出一种改进模型TP(Tunnel Portal)-DeeplabV3+,其结构如图7所示。

2.2 DeeplabV3+改进方法

2.2.1 替换主干特征提取网络

原始DeeplabV3+采用Xception作为主干特征提取网络。若想后续将算法嵌入无人机中,参数量过大的模型会带来诸多不便,故采用轻量化模型MobileNetV2[15]作为主干特征提取网络。

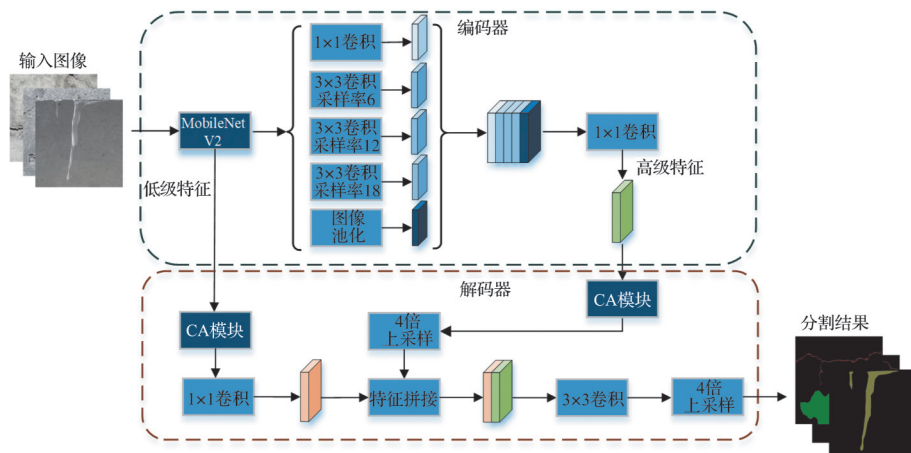


图7 TP-DeeplabV3+模型结构

Fig. 7 Overall structure of TP-DeeplabV3+

MobileNet 模型是 Google 针对手机等嵌入式设备提出的一种轻量级深层神经网络。MobileNetV2 是 MobileNet 的升级版,其主要创新点是在结构中引入了 Inverted Resblock 和 Linear Bottleneck。Inverted Resblock 可以使网络结构学习到如何更好地扩展数据和压缩数据,整个 MobileNetV2 结构都由 Inverted Resblock 组成,其结构如图 8 所示:左边为主干部分,首先采用 1×1 卷积对特征层通道升维,然后采用 3×3 深度可分离卷积进行特征提取,最后再用 1×1 卷积对特征层通道降维;右边为残差连接部分,将输入和输出直接相连,以解决梯度消失问题。此外,为了避免非线性激活层导致的特征信息损失,每一个 Inverted Resblock 的最后一层都使用了 Linear Bottleneck,即用线性激活层代替原来的 ReLU 非线性激活层。



图 8 Inverted Resblock 结构

Fig. 8 Overall structure of Inverted Resblock

2.2.2 添加注意力机制

注意力机制^[16-17]是深度学习中常用的一个技巧,已被广泛应用于自然语言处理及计算机视觉领域。CA (coordinate attention) 是一种轻量级注意力模块^[18],于 2021 年提出,通过将位置信息嵌入到通道注意力中,解决了常见注意力机制(如 CBAM、SE)存在的空间信息损失问题。CA 整体结构主要包括 2 部分,如图 9 所示。

第 1 部分为坐标信息嵌入,该过程有助于模型更好地感知和处理不同位置的信息。假设输入进来的特征层形状为 $[C, H, W]$,分别以尺寸为 $[H, 1]$ 及 $[1, W]$ 的池化核沿水平方向及垂直方向对每个通道进行编码。故位于第 c 个通道在高度 h 处的输出可以表示为

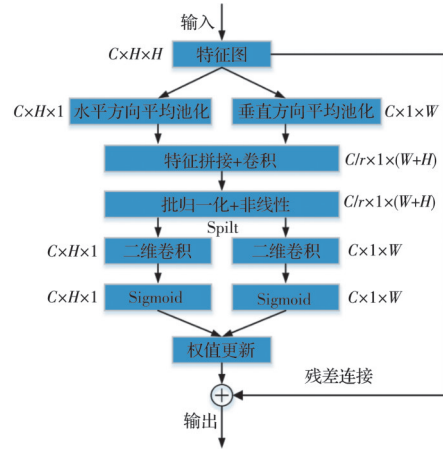


图 9 CA 结构图

Fig. 9 Structure of the CA module

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (1)$$

同样,第 c 个通道在宽度 w 处的输出可以表为

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \quad (2)$$

第 2 部分为坐标注意力生成,首先对第 1 部分提取出的尺寸分别为 $[C, H, 1]$ 和 $[C, 1, W]$ 的特征层进行特征拼接,将其输入一个共享的 1×1 卷积变换函数 F_1 中,计算过程表达为

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

式中: $[\cdot, \cdot]$ 表示沿空间维度的拼接操作; δ 为非线性激活函数; $f \in R^{C/r \times (H+W)}$ 为在水平方向和垂直方向对空间信息编码生成的中间特征; r 为通道下采样率。

此时获得的特征层尺寸为 $[C/r, 1, W+H]$,随后将 f 分解为 $f^h \in R^{C/r \times H}$ 和 $f^w \in R^{C/r \times W}$,再使用两个 1×1 卷积变换 F_h 和 F_w ,将 f^h 和 f^w 变换为具有相同通道数的张量,其表达式为

$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

式中: g^h 和 g^w 即为计算出的注意力权重; σ 为 sigmoid 激活函数。

故 CA 的最终输出 y 可表示为

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

在由主干特征提取网络后产生的低级特征层以及经过 ASPP 产生的高级特征层后分别嵌入一个 CA 模块,将注意力权重初始化为一个较小的随机值,并在训练过程中不断迭代优化,直至得到最优权重。随后利用最优注意力权重对特征层进行更新,使其包含更加丰富的病害边界细节信息及主体信息,提高模型在复杂背景干扰下的病害分割精度。

2.3 病害量化

TP-DeeplabV3+模型实现了对洞门墙病害的自动化识别和定位,在实际检测过程中,还需对洞门墙病害进行量化。

目前对工程结构病害的量化研究多以计算裂缝宽度为主^[19],但考虑到高原地区洞门墙病害种类多且形态复杂,测量裂缝宽度的实际价值有限。因此,选取病害面积及病害覆盖率作为量化指标,其中,病害面积可以直观判断出不同病害的发育情况;病害覆盖率 R 定义为该区域病害面积与图像面积之比,其可以对洞门墙健康状况判断提供重要依据。

在完成对病害的分割后,TP-DeeplabV3+模型能自动计算出图像中不同病害所占的掩膜像素数,对于实际病害面积的计算,可由式(7)表示^[20]。

$$S_i = \frac{T_i}{T} S \quad (7)$$

式中: T_i 为病害的像素数; T 为图像的总像素数; S_i 为洞门墙病害的实际面积; S 为待测图像的实际面积。

在计算出病害实际面积后,实际病害覆盖率 R 表示为

$$R = \sum_{i=1,2,3} \frac{S_i}{S} \quad (8)$$

3 模型训练

所有实验都在一台内存为42 G的Ubuntu 20.04 Linux主机上进行,使用一块显存为24 G的NVIDIA GeForce RTX 3090显卡进行GPU加速。基于编程语言Python和深度学习框架Pytorch搭建模型,选用Pycharm作为集成开发环境。

3.1 训练参数设置

将TD-DeeplabV3+的初始学习率设置为 7×10^{-3} ,选用SGD作为优化器,设置动量因子及权重衰减,并使用StepLR策略对学习率进行动态调整。模型一共训练250轮,每一轮都会在验证集上进行评估,每批次数量设为32。此外,在主干网络中采用了预训练权重,其可以缩短模型的训练时间,提高训练精度^[21]。

3.2 损失函数

由于高原地区隧道洞门墙病害图像背景通常较为复杂,在对一些微小病害(如纤细裂缝)进行检测时,图像中的病害像素数量会远低于背景像素数量,因此,有必要采取措施解决正负样本不平衡问题。FocalLoss是一种处理样本分类不均衡的损失函数,通过减少易分类样本(背景)的权重,使得模型在训练时更专注于难分类(病害)的样本。其原

理如式(9)、式(10)所示。

$$FL(p_i) = -a_i(1 - p_i)^\gamma \log(p_i) \quad (9)$$

$$FL(p) = \begin{cases} -\alpha(1 - p)^\gamma \log(p), & y = 1 \\ -(1 - \alpha)p^\gamma \log(1 - p), & y = -1 \end{cases} \quad (10)$$

式中: $y = 1$ 表示为正样本,即病害; $y = -1$ 表示为负样本,即背景; $p \in [0, 1]$ 表示模型对 $y = 1$ 的类的估计概率; α 表示平衡因子,用于平衡正样本和负样本的权重;聚焦参数 γ 用于降低易分类样本的损失比例,当 $\gamma = 0$ 时,FocalLoss就是传统的交叉熵损失。 α 和 γ 分别被设置为0.25和2;log运算的底默认为自然常数e。

3.3 语义分割评估指标

对TP-DeeplabV3+模型分割性能的评估,主要从模型体量和分割精度两方面进行分析。

模型体量采用单张图片检测速度(ms)和模型权重大小(MB)来衡量;分割精度采用平均像素精确度(mean pixel accuracy, m_{PA})、平均交并比(mean intersection over union, m_{IoU})、精度(precision, P)和召回率(recall, R)来衡量。其中, m_{PA} 表示每个类中正确分类的像素数量占像素总数比例的平均; m_{IoU} 表示图像中每一个类交并比的平均; P 和 R 则用于衡量模型的误检率和漏检率。上述精度衡量指标均为语义分割常用评价指标,越靠近1则表示性能越好,其定义为

$$m_{PA} = \frac{1}{k+1} \sum_{i=0}^k \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (11)$$

$$m_{IoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{T_P}{T_P + F_P + F_N} \quad (12)$$

$$P = \frac{T_P}{T_P + F_P} \quad (13)$$

$$R = \frac{T_P}{T_P + F_N} \quad (14)$$

式中:假设有 $k+1$ 类标签, T_P (真阳性)表示一个病害像素被正确地识别为病害; F_P (假阳性)表示一个背景像素被错误地识别为病害; F_N (假阴性)表示一个病害像素被错误地识别为背景; T_N (真阴性)表示一个背景像素被正确地识别为背景。

4 结果与讨论

4.1 模型训练结果

为了验证TD-DeeplabV3+模型的优越性,将其与DeeplabV3+、U-Net和PSPNet进行了对比。上述模型都在同一数据集上训练和验证,采用相同的训练策略,其在验证集上的损失函数及 m_{IoU} 变化如图10、图11所示。

随着训练的进行,4种语义分割模型在验证集

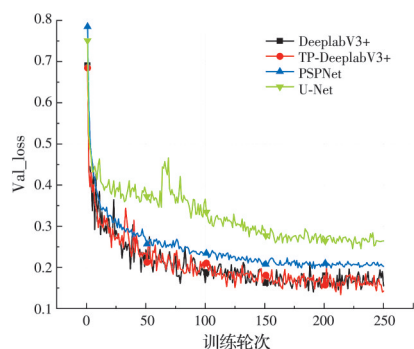


图 10 模型训练验证损失曲线

Fig. 10 Model training validation loss curves

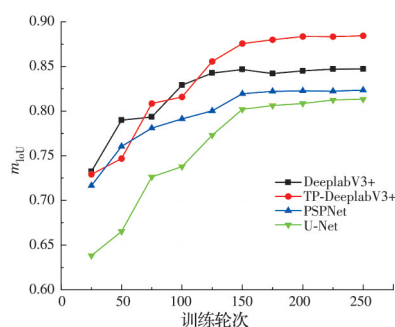
图 11 模型训练 m_{IoU} 曲线

Fig. 11 Model training process

上的验证损失均逐渐降低,下降趋势在前 50 个轮次尤为明显,并在约 150 个轮次后逐渐趋于收敛和稳定,仅在小范围内震荡,没有发生过拟合。TP-DeeplabV3+ 模型的验证损失最终收敛在 0.14 左右,具有较好的泛化性,而 U-Net 仅能达到 0.26 左右,为 4 种模型中最差。

由图 11 可知,在训练的前 150 个轮次中,4 种模型的 m_{IoU} 基本都保持上升趋势,随后逐渐趋于稳定,这也与验证损失曲线的变化趋势相对应。TP-DeeplabV3+ 在验证集上的 m_{IoU} 可达到 88.44%,高于 DeeplabV3+ (84.71%),而 U-Net 和 PSPNet 的 m_{IoU} 仅为 81.35% 和 82.34%。

4.2 不同模型测试结果对比

训练完成后,用训练过程中验证损失最小的模型在测试集上进行最终评估,结果如图 12、表 1 和表 2 所示。

表 1 不同模型体量评估结果

Table 1 Evaluation results of different model sizes

模型	参数量/MB	检测速度/(ms/张)
DeeplabV3+	199.82	59.18
PSPNet	178.52	68.96
U-Net	95.14	86.95
TP-DeeplabV3+	22.31	45.61

表 2 不同模型分割性能评估结果

Table 2 Segmentation performance evaluation results of different models

语义分割模型	$m_{IoU}/\%$	$m_{PA}/\%$	$P/\%$			$R/\%$		
			裂缝	剥落	挂冰	裂缝	剥落	挂冰
DeeplabV3+	83.57	91.95	74.24	94.15	90.43	83.41	93.02	92.42
PSPNet	79.56	90.32	63.13	90.29	89.71	77.11	94.41	92.19
U-Net	76.70	85.98	78.37	88.19	81.46	82.31	92.49	72.19
TP-DeeplabV3+	88.37	94.93	79.31	96.27	95.05	88.17	95.03	97.36

在分割精度方面,相比 PSPNet 和 U-Net,TP-DeeplabV3+ 和 DeeplabV3+ 的各项评估指标均比较优秀。TP-DeeplabV3+ 在测试集上的 m_{IoU} 和 m_{PA} 可以达到 88.37% 和 94.93%,为 4 种模型中的最高值,相比 DeeplabV3+ ($m_{IoU}=83.57\%$, $m_{PA}=91.95\%$) 分别提高了 4.8% 和 2.98%,而表现最差的 U-Net 只有 76.70% 和 85.98%。当病害特征清晰且背景干扰因素少时,如图 12(a)、(b),4 种模型都能较精确地识别出图像中的病害。而在面对包含复杂交叉裂缝及微小裂缝的图 12(c)、(d) 时, PSPNet 和 DeeplabV3+ 对图像中纤细裂缝的分割连续性不好,这是因为其结构中的空洞卷积和空间金字塔池模块会导致精细特征的丢失。TP-DeeplabV3+ 和 U-Net 则更善于捕捉精细目标的边界,但 U-Net 缺乏对图像语义信息之间依赖关系的关注^[22],导致其对挂冰及剥落的分割效果不佳,对于

挂冰仅有 81.46% 的 Precision 和 72.19% 的 Recall,远低于其他模型。

当病害形态复杂且背景干扰因素多时,如图 12(e)、(f) 所示。此时 DeeplabV3+、PSPNet 和 U-Net 的分割效果都显得不尽人意,存在明显的漏检及误检,而 TP-DeeplabV3+ 依旧可以实现较精确的分割,表现出较好的鲁棒性。

在模型体量方面,TP-DeeplabV3+ 模型的权重大小仅为 22.31 MB,相比原始 DeeplabV3+,压缩了 88.83%,相比 PSPNet 和 U-Net,分别小 87.50% 和 76.55%,并且 45.61 ms 的单张图像检测速度也为 4 种模型中的最低值。综合上述定量及定性分析可以看出,TP-DeeplabV3+ 有效解决了原模型对于精细目标分割不够精细的问题,并具备了轻量化的特点,有一定优越性。

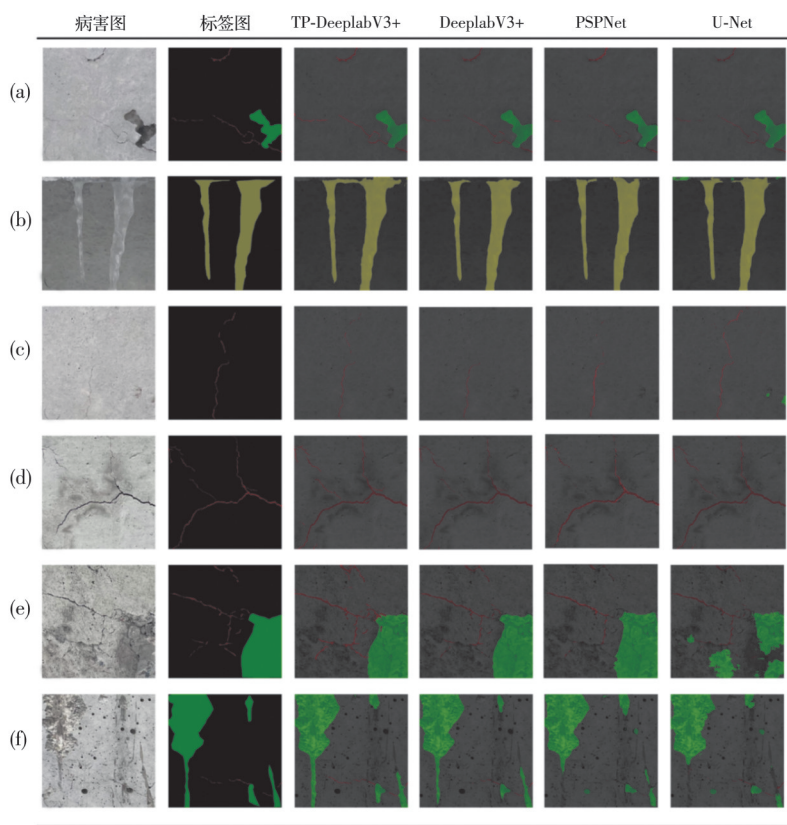


图 12 不同模型分割结果对比

Fig. 12 Comparison of segmentation results among different models

4.3 消融实验

为进一步探究对模型不同改进措施的有效性,设计了消融实验,如表 3、表 4 所示。DeeplabV3+(CA)表示仅在 DeeplabV3+基础模型上添加 CA 注意力机制,DeeplabV3+(MobileNet)表示采用 MobileNetV2 作为主干特征提取网络。上述模型均使用 Focalloss 作为损失函数。消融实验结果表明:在使用 FocalLoss 的基础上将主干网络替换为 MobileNetV2 后,模型在各项指标上仍有相应提升,

对于裂缝的 Precision 和 Recall 分别提升了 4.07% 和 3.76%,并且单张图片检测速度提升了 13.57 ms,参数量降低了 89.5%,提升效果显著;当在使用 Focalloss 的基础上添加 CA 注意力机制时,模型的 m_{IoU} 和 m_{PA} 分别提升了 4.68% 和 3.03%,而检测速度与参数量仅提高了 0.84 ms 和 1.32 MB,这也体现出了 CA 轻量级的特点。而 TP-DeeplabV3+在 FocalLoss 的基础上以 MobileNetV2 为主干网络并添加了注意力机制,与仅使用单一改进策略相比,兼顾了模型体量与分割精度。由此可见,提出的各项改进措施有效:CA 注意力机制可以使模型更好地捕捉不同病害的语义信息,更精细地提取病害形态;以 Focalloss 作为损失函数有效地解决了像素类别不平衡问题,提高了模型在面对复杂背景干扰下的鲁棒性;使用 MobileNetV2 作为主干模型有效减少了参数量,提升了检测速度。

表 3 消融实验模型体量对比

Table 3 Comparison of model sizes in ablation experiments

模型	参数量/MB	检测速度/(ms/张)
DeeplabV3+	199.82	59.18
DeeplabV3+(MobileNet)	20.97	44.95
DeeplabV3+(CA)	201.14	60.02
TP-DeeplabV3+	22.31	45.61

表 4 消融实验分割精度对比

Table 4 Comparison of segmentation accuracy in ablation experiments

语义分割模型	$m_{IoU}/\%$	$m_{PA}/\%$	$P/\%$			$R/\%$		
			裂缝	剥落	挂冰	裂缝	剥落	挂冰
DeeplabV3+	83.57	91.95	74.24	94.15	90.43	83.41	93.02	92.42
DeeplabV3+(MobileNet)	85.05	91.82	78.31	95.27	95.05	87.17	84.03	97.36
DeeplabV3+(CA)	88.25	94.98	79.29	95.98	95.16	87.41	95.19	97.39
TP-DeeplabV3+	88.37	94.93	79.31	96.27	95.05	88.17	95.03	97.36

考虑到仅用裁剪后的图像对检测效果进行评价不够全面,使用 TP-DeeplabV3+ 对高分辨率原图直接进行检测,结果如图 13 所示。模型可以较准确地识别出图像中的不同病害,但会对少量精细目标产生误检,总体可以满足实际工程应用的需求。

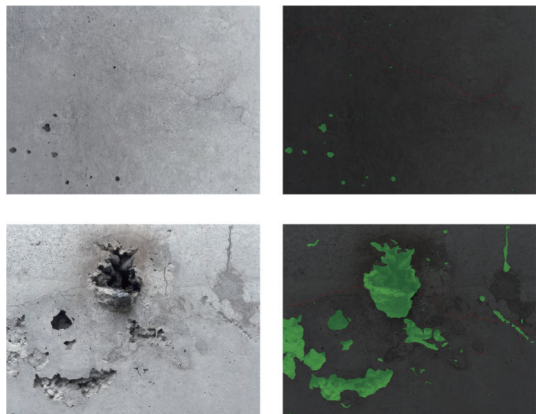


图 13 高分辨率图像分割结果
Fig. 13 Segmentation results for high-resolution images

4.4 病害量化结果分析

人工实测值由工程现场实测及数字图像处理两种方法测得,工程现场实测法由现场检测人员借助测量工具对病害进行测量,主要针对特征明显、面积较大的病害,而形态不规则和细小病害则由数字图像处理法进行测量,其步骤如下:

- 1)使用图像标注软件 Labelme 对病害进行像素级标注;
- 2)使用 Canny 边缘检测算法提取出病害轮廓,并通过形态学腐蚀操作去除孤立点;
- 3)对不同病害轮廓进行像素填充,以计算出病害的掩膜像素数。

为保证实际检测中病害量化结果的精度,以尺寸为 512×512 像素、分辨率为 96dpi 的裁剪窗口为基准,沿图像边缘裁剪,对每个裁剪窗口进行单独计算,之后将每个裁剪窗口中的病害面积相加,即可实现对原图像病害的量化。

依据真实尺寸与图像尺寸的转换关系,96dpi 表示 1 英寸(2.54 cm)包含 96 个像素点,故每个裁剪窗口中 $1\text{ cm}\approx 37.795\text{ px}$,真实面积约为 183.50 cm^2 。随机选取一张未参与过训练的现场图像,先对其进行初步裁剪,保留病害最集中的核心区域,随后对其进行量化。如图 14 所示,经过窗口裁剪后,该图像共包含 10 个病害窗口(剔除只包含背景的窗口),由式(7)、式(8)计算出每个裁剪窗口中的病害面积及覆盖率,结果如表 5 所示。与人工实测值相比,绝对误差不超过 0.3%,相对误差控制在 7.31% 以下。为进一步验证方法的有效性,在原图像上直接计算病害覆盖率,其绝对误差为 0.67%,相对误差高达

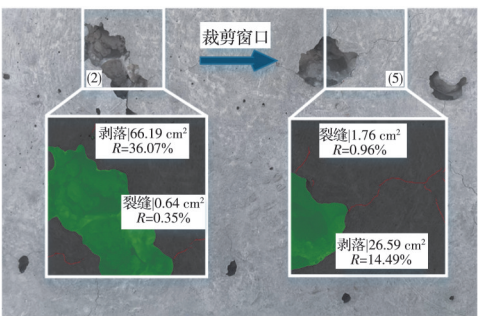


图 14 病害覆盖率计算
Fig. 14 Disease coverage calculation

表 5 病害覆盖率计算结果验证

Table 5 Verification of disease coverage calculation results					
区域 编号	计算面积/ cm ²	计算覆盖 率/%	实测覆盖 率/%	绝对误 差/%	相对误 差/%
1	1.523	0.830	0.840	0.010	1.190
2	66.831	36.420	36.132	0.288	0.798
3	1.853	1.010	0.982	0.028	2.901
4	31.819	17.341	17.171	0.169	0.981
5	28.351	15.453	15.411	0.039	0.251
6	9.799	5.341	5.237	0.103	1.970
7	2.312	1.260	1.200	0.060	5.024
8	3.028	1.650	1.538	0.112	7.303
9	0.752	0.411	0.400	0.010	2.556
10	1.817	0.990	1.015	0.025	2.472
总计	148.085	3.650	3.616	0.034	0.940
原图	174.031	4.290	3.616	0.674	18.639

18.64%,准确性较低。

5 结论

高原地区隧道洞门墙通常存在所处自然环境恶劣、病害种类繁多、结构高耸不易检测等特点,考虑到无人机具有轻巧灵活、自动化程度高的特点,可以在高原寒冷、缺氧的环境下代替人工巡查,实现对隧道洞门墙病害图像的智能化采集。提出了基于无人机图像与深度学习的高原地区隧道洞门墙病害检测方法,得出以下结论:

1)提出的改进模型 TP-DeepLabV3+ 能精细地提取不同病害图像的语义特征,有效解决了 DeepLabv3+ 对精细目标存在分割结果粗糙及洞门墙病害图像中的类别不平衡等问题,并具备轻量化特点。实验结果表明,TP-DeepLabV3+ 在测试集上的各项评价指标均优于其他模型,可以达到 88.37% 的 m_{IoU} 及 94.93% 的 m_{PA} ,在面对病害形态复杂及背景干扰严重的图像时,展现出了较好的鲁棒性。

2)与人工实测值相比,TP-DeepLabV3+ 模型计算出的病害量化结果对于病害覆盖率的绝对误差不超过 0.3%,相对误差维持在 7.31% 以下,可以

满足实际工程应用的需求。

3)提出的方法实现了高原恶劣环境下的隧道洞门墙病害智能化检测,在施工和运营阶段,使用无人机对洞门墙进行定期巡查能节省人工成本。

该方法主要以采集图像的方式对病害进行检测,后续研究重点在于开发可嵌入在无人机上的病害实时检测算法。

参考文献

- [1] 刘海京,夏才初,朱合华,等.隧道病害研究现状与进展[J].地下空间与工程学报,2007,3(5):947-953.
LIU H J, XIA C C, ZHU H H, et al. Studies on tunnel damage [J]. Chinese Journal of Underground Space and Engineering, 2007, 3(5): 947-953. (in Chinese)
- [2] HUANG H W, LI Q T, ZHANG D M. Deep learning based image recognition for crack and leakage defects of metro shield tunnel [J]. Tunnelling and Underground Space Technology, 2018, 77: 166-176.
- [3] 任松,朱倩雯,涂歆玥,等.基于深度学习的公路隧道衬砌病害识别方法[J].浙江大学学报(工学版),2022,56(1):92-99.
REN S, ZHU Q W, TU X Y, et al. Lining disease identification of highway tunnel based on deep learning [J]. Journal of Zhejiang University (Engineering Science), 2022, 56(1): 92-99. (in Chinese)
- [4] 陈莹莹,刘新根,黄永亮,等.基于神经网络与边缘修正的隧道衬砌裂缝识别[J].现代隧道技术,2022,5906:24-34.
CHEN Y Y, LIU X G, HUANG Y L, et al. Crack identification of tunnel lining based on neural network and edge correction [J]. Modern Tunnelling Technology, 2022, 59(6): 24-34. (in Chinese)
- [5] 孟庆成,李明健,万达,等.基于M-Unet的混凝土裂缝实时分割算法[J].土木与环境工程学报(中英文),2024,46(1):215-222.
MENG Q C, LI M J, WAN D, et al. Real-time segmentation algorithm of concrete cracks based on M-Unet [J]. Journal of Civil and Environmental Engineering, 2024, 46(1):215-222. (in Chinese)
- [6] 张修杰,袁嘉豪,岳学军,等.基于优化DeepLabv3+的混凝土梁裂缝分割及特征量化[J].科学技术与工程,2023,23(9):3794-3803.
ZHANG X J, YUAN J H, YUE X J, et al. Crack segmentation and feature quantification of concrete beams based on optimized DeepLabv3+ [J]. Science Technology and Engineering, 2023, 23(9): 3794-3803. (in Chinese)
- [7] FU H X, MENG D, LI W H, et al. Bridge crack semantic segmentation based on improved Deeplabv3+ [J]. Journal of Marine Science and Engineering, 2021, 9(6): 671.
- [8] 饶勇成,韩晓健,肖飞,等.基于深度学习的混凝土结构多病害检测[J].建筑结构,2021,51(增刊2):1439-1445.
RAO Y C, HAN X J, XIAO F, et al. Multiple disease detection of concrete structures based on deep learning [J]. Building Structure, 2021, 51(Sup 2): 1439-1445. (in Chinese)
- [9] 刘少华,任宜春,郑智雄,等.基于航拍图像与改进U-Net的建筑外墙裂缝检测方法[J].土木与环境工程学报(中英文),2024,46(1):223-231.
LIU S H, REN Y C, ZHENG Z X, et al. Building exterior wall crack detection based on aerial images and improved U-Net [J]. Journal of Civil and Environmental Engineering, 2024, 46(1):223-231. (in Chinese)
- [10] 丁威,俞珂,舒江鹏.基于深度学习和无人机的混凝土结构裂缝检测方法[J].土木工程学报,2021,54(增刊1):1-12.
DING W, YU K, SHU J P. Method for detecting cracks in concrete structures based on deep learning and UAV [J]. China Civil Engineering Journal, 2021, 54(Sup 1): 1-12.
- [11] 余加勇,李锋,薛现凯,等.基于无人机及Mask R-CNN的桥梁结构裂缝智能识别[J].中国公路学报,2021,34(12):80-90.
YU J Y, LI F, XUE X K, et al. Intelligent identification of bridge structural cracks based on unmanned aerial vehicle and mask R-CNN [J]. China Journal of Highway and Transport, 2021, 34(12): 80-90. (in Chinese)
- [12] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Computer Vision-ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII. ACM, 2018: 833-851.
- [13] ASADI SHAMSABADI E, XU C, RAO A S, et al. Vision transformer-based autonomous crack detection on asphalt and concrete surfaces [J]. Automation in Construction, 2022, 140: 104316.
- [14] LI Z Q, CHEN X, JIANG J, et al. Cascaded multiscale structure with self-smoothing atrous convolution for semantic segmentation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5605713.
- [15] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 4510-4520.
- [16] WANG F, JIANG M Q, QIAN C, et al. Residual attention network for image classification [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE,

- 2017: 6450-6458.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [18] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA. IEEE, 2021: 13708-13717.
- [19] 刘宇飞, 樊健生, 聂建国, 等. 结构表面裂缝数字图像法识别研究综述与前景展望[J]. 土木工程学报, 2021, 54(6): 79-98.
- LIU Y F, FAN J S, NIE J G, et al. Review and prospect of digital-image-based crack detection of structure surface [J]. China Civil Engineering Journal, 2021, 54(6): 79-98. (in Chinese)
- [20] 朱劲松, 李欢. 基于深度学习的钢桥病害分割与量化[J]. 东南大学学报(自然科学版), 2022, 53(3): 516-522.
- ZHU J S, LI H. Segmentation and quantification of steel bridge defects based on deep learning [J]. Journal of Southeast University (Natural Science Edition), 2022, 53(3): 516-522. (in Chinese)
- [21] ZHENG Z J, ZHANG Z Q, PAN W. Virtual prototyping- and transfer learning-enabled module detection for modular integrated construction [J]. Automation in Construction, 2020, 120: 103387.
- [22] CHEN J N, LU Y Y, YU Q H, et al. TransUNet: Transformers make strong encoders for medical image segmentation [J/OL]. arXiv: 2102.04306[cs.CV]. <http://arxiv.org/abs/2102.04306.pdf>

(编辑 胡英奎)