

DOI:10.11918/202502061

混凝土桥梁整体表面多缺陷图像精细分割方法

周勇军¹, 罗楠¹, 孙延晨¹, 尚嘉琪², 陈炽毅¹

(1. 长安大学公路学院, 西安 710064; 2. 上海市政交通设计研究院有限公司, 上海 200030)

摘要:为解决现有混凝土桥梁数字图像方法缺陷识别单一、分割精度较低等问题,本文提出了一种基于编解码架构的精细化语义分割模型 HDNet。编码器设计方面,采用层级化窗口自注意力机制,通过滑动窗口划分结合跨层残差连接增强梯度传播。引入核化注意力模块强化侵蚀、裂缝等局部缺陷的梯度响应,抑制桥梁背景纹理干扰。解码器设计像素-变形双路框架体现在像素路径通过逐点特征映射解析裂缝等形态细节,变形路径采用可变形卷积自适应匹配剥落区域的不规则几何轮廓。基于无人机采集的高清桥梁缺陷数据集(涵盖裂缝、侵蚀、露筋、剥落4类缺陷),开展与 DeepLabV3+、SegFormer 等主流模型的对比实验,随后进行消融实验分析、热力图分析和实桥测试。结果表明:HDNet 验证集交并比(mIoU)达 71.91%,较次优模型 SegFormer 提升了 7.86%;消融性实验验证了核化注意力(提升召回率 mRecall 5.83%)、层级化滑窗注意力(提升 mIoU 5.92%)与 Dice 损失函数协同设计的必要性;热力图分析证实 HDNet 能够精准捕捉缺陷纹理细节并解耦伴生缺陷的语义边界;实桥测试中,HDNet 将缺陷尺寸测量相对误差稳定控制在 $\pm 5\%$ 以内,验证了其在实际应用中的适用性。HDNet 通过编解码协同优化与跨分辨率层级化增强机制,有效提升复杂桥梁缺陷的识别精度与鲁棒性,可为桥梁表面病害智能化检测提供高精度技术手段。

关键词: 桥梁工程;混凝土表面缺陷;语义分割;深度学习;注意力机制;无人机

中图分类号: U446.2

文献标志码: A

文章编号: 0367-6234(2025)06-0103-13

Fine-grained image segmentation method for holistic surface multi-defects in concrete bridges

ZHOU Yongjun¹, LUO Nan¹, SUN Yanchen¹, SHANG Jiaqi², CHEN Chiyi¹

(1. School of Highway, Chang'an University, Xi'an 710064, China;

2. Shanghai Municipal Traffic Design and Research Institute Co., Ltd., Shanghai 200030, China)

Abstract: To address the issues of single-category defect identification and low segmentation accuracy in current digital image-based methods for concrete bridge defect detection, a refined semantic segmentation model named HDNet, which is built upon an encoder-decoder architecture, was introduced. In terms of encoder design, a hierarchical window-based self-attention mechanism was implemented, which combines sliding window partitioning and cross-layer residual connections to enhance gradient propagation. A kernelized attention module was incorporated to strengthen gradient responses for local defects, such as erosion and cracks, while simultaneously reducing interference from the background texture of the bridge. A pixel-deformation dual-path architecture was adopted in the decoder, in which the pixel path employs pointwise feature mapping to capture the morphological details of cracks and the deformation path utilizes deformable convolutions to adaptively match the irregular geometric contours of spalling regions. A series of experiments were carried out on a high-resolution dataset of bridge defects including four categories of defects: cracks, erosion, exposed rebar, and spalling, which was captured by unmanned aerial vehicle (UAV). Comparisons with those dominant models such as DeepLabV3+ and SegFormer were performed, and then ablation study analysis, heatmap analysis and real-bridge validation were carried out. The results indicate that HDNet attains a mean Intersection over Union (mIoU) of 71.91% on the validation set, surpassing the suboptimal model SegFormer by 7.86%. Ablation studies validate the necessity of kernelized attention (which improves mRecall by 5.83%), hierarchical sliding-window attention (which boosts mIoU by 5.92%), and the synergistic design with the Dice loss function. Heatmap analysis demonstrates HDNet's ability to accurately capture defect texture details and disentangle the semantic boundaries of co-occurring defects. In real-bridge testing, HDNet maintains the relative error of defect size measurement within $\pm 5\%$, which confirms

收稿日期: 2025-02-28; 录用日期: 2025-04-22; 网络首发日期: 2025-05-27

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20250526.1839.002>

基金项目: 国家重点研发计划(2021YFB2601000); 国家自然科学基金(52278138); 中央高校基本科研业务费资助项目(300102214301)

作者简介: 周勇军(1978—), 男, 教授, 博士生导师

通信作者: 周勇军, zyj@chd.edu.cn

its practical applicability. By integrating encoder-decoder co-optimization and cross-resolution hierarchical enhancement mechanisms, HDNet substantially enhances the recognition accuracy and robustness for complex bridge defects, thereby offering a high-precision technology for the intelligent detection of bridge surface deterioration.

Keywords: bridge engineering; concrete surface defects; semantic segmentation; deep learning; attention mechanism; unmanned aerial vehicle (UAV)

桥梁结构的定期检查是确保其安全运营的必要措施^[1]。在传统检测体系中,依托桥检车平台的人工目视检测占据主导地位,然而该方法存在检测效率低下、经济成本高昂、主观性较强等固有缺陷,已难以适应现代路网桥梁规模化检测需求^[2-3]。随着深度学习与计算机视觉技术的深度融合,混凝土桥梁表观缺陷智能检测技术逐渐得到应用并取得突破性进展^[4-5]。

Liu 等^[6]提出一种基于掩模区域的卷积神经网络(R-CNN),实现了裂缝像素级分割(mIoU 70.1%),但模型对剥落等非连续缺陷的适用性还有待验证。Wang 等^[7]通过双特征编码器耦合对比学习约束,构建无监督缺陷信号重构模型,显著提升了分割精度。余加勇等^[8]集成 YOLOv5 与 U-Net3+ 算法,经参数优化实现裂缝快速定位与像素级分割。杨炜等^[9]提出了一种基于改进 YOLOv8 的轻量化算法 YOLOv8-ALTE,通过在骨干 Backbone 特征提取模块的浅层网络中加入三元注意力机制,以此来提高模型对桥梁裂缝病害的识别及分割准确度。骆剑彬等^[10]对 Faster R-CNN 框架下 VGG16 网络模型进行改进,对水下桩墩孔洞和剥落等病害识别结果的总体准确率达 88.3%。Zhang 等^[11]研究了一种基于 U-NET 的像素级裂缝检测方法,分别实现了 91.45% 和 88.67% 的准确率和召回率,但其复杂背景下的性能稳定性不足。褚鸿鹄等^[12]采用级联 CA-TransUNet 对无人机拍摄的高分辨率裂缝图像实现了精细化分割。

尽管上述研究在单一缺陷检测中均取得了较好成效,但在多缺陷协同分割领域仍面临三重技术难题^[13-15]: 1) 裂缝(毫米级线状纹理)与剥落(厘米级块状特征)等缺陷在形态尺度与纹理特性上存在显著异质性,导致单任务模型难以实现跨类别特征特异性捕捉; 2) 无人机航拍图像因飞行高度波动产生的尺度多样性,使得传统金字塔池化方法在跨尺度特征融合时易丢失细节信息。3) 《混凝土结构设计规范》GB 50010—2010 中指出,裂缝渗透引发的氯离子扩散会加速钢筋锈蚀与保护层剥落,而露筋暴露将进一步加剧承载力衰减,这种多缺陷耦合作用机制亟需精细化分割技术支撑定量化评估。裂缝、侵蚀、露筋、剥落等表观缺陷是混凝土结构最直观的

性能退化标志,因其会极大影响桥梁服役的安全性与耐久性,始终是桥梁定期检查中的重点关注对象^[1]。因此,发展多缺陷精细化分割技术对桥梁智能检测具有十分重要的理论意义和工程价值。

本文提出了基于层次化编码-解码架构的桥梁多缺陷分割模型 HDNet (hierarchical dual-decoder network)。重点研究如何通过优化编解码架构实现核化注意力、层次化滑窗注意力、交叉注意力与形状损失函数的有机统一。研究将从编码器架构优化、注意力机制强化、解码路径创新 3 个维度展开技术创新,重点解决传统方法在复杂缺陷分割中存在的特征捕捉特异性不足、尺度适应性有限及语义边界混杂等关键问题,以期为复杂桥梁缺陷的识别分析提供精细化分割技术支撑。

1 基于编解码架构的缺陷精细分割方法

1.1 编码器设计

编码器部分采用特征提取网络,模型前 5 层使用 3×3 卷积组和跨层残差连接,在卷积残差块后加入最大池化层,进行采样操作,有效获取图像的初级特征;随后通过 4 个层级化阶段(Stage)逐步减小输入特征图的分辨率,并逐层扩大感受野。在各 Stage 中,通过线性嵌入层将图像分割成小块并转换为嵌入向量,通过层次化窗口进行相对位置编码,通过自注意力模块输出 4 张多尺度特征图用于后续解码器的运算处理,如图 1(a)所示。为了使模型能够更精确地区分和识别不同缺陷与背景的纹理特征,本文通过核化注意力动态地调整模型的注意力焦点,首先,使用一个由初始内核 K_{i-1} 预测的掩模图像 M_{i-1} 来拼装每个像素组的特征 F_k ,由于每个组内部的内容有所区别,这些特征 F_k 将作为模型自适应更新其对应内核 K_{i-1} 的依据。随后,这些内核通过相互对比来对图像的上下文语义信息进行全局综合建模得到新的内核。最后,通过使用这些聚焦于组信息的内核对特征图进行卷积,获得更准确的掩模预测,如图 1 所示,其中 C 是 channel 通道, D 是 dimension 维度, N 是一组已学习卷积核的数量, K 是卷积核数, H, W 分别为图片的高度和宽度。

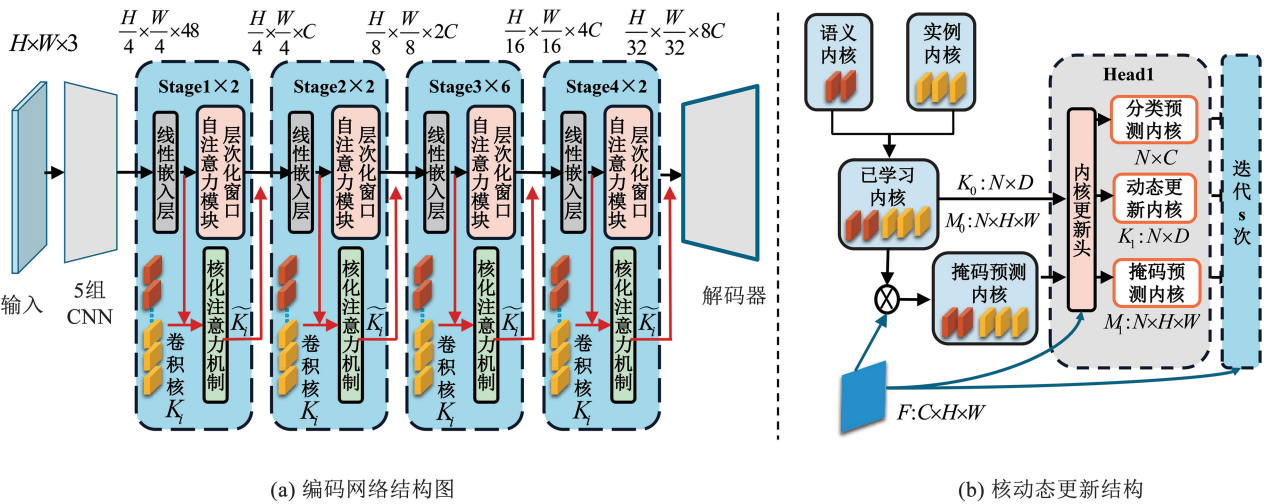


图1 编码网络结构与核动态更新结构

Fig. 1 Structure diagram of the encoding network and dynamic kernel update structure

1.2 解码器设计

本文采用像素解码器和变换解码器的并行结构,像素解码器利用分辨率由低到高的特征金字塔,将不同尺度特征分别馈送到对应的变换解码层,通过特征金字塔处理多尺度特征图^[16],并通过添加正弦位置嵌入和可学习尺度嵌入来捕捉纹理、边缘和

局部形状等细节信息,以实现精确的像素级预测,最终输出4张不同分辨率的特征图。如图2(a)所示,前3张图片(P1~3)作为掩膜特征图输入变形编码器进行进一步的空间变换学习,第4张图片(P4)经过归一化和全卷积层处理之后直接参与损失计算,该路径通过简化处理流程保持高频细节完整性^[17]。

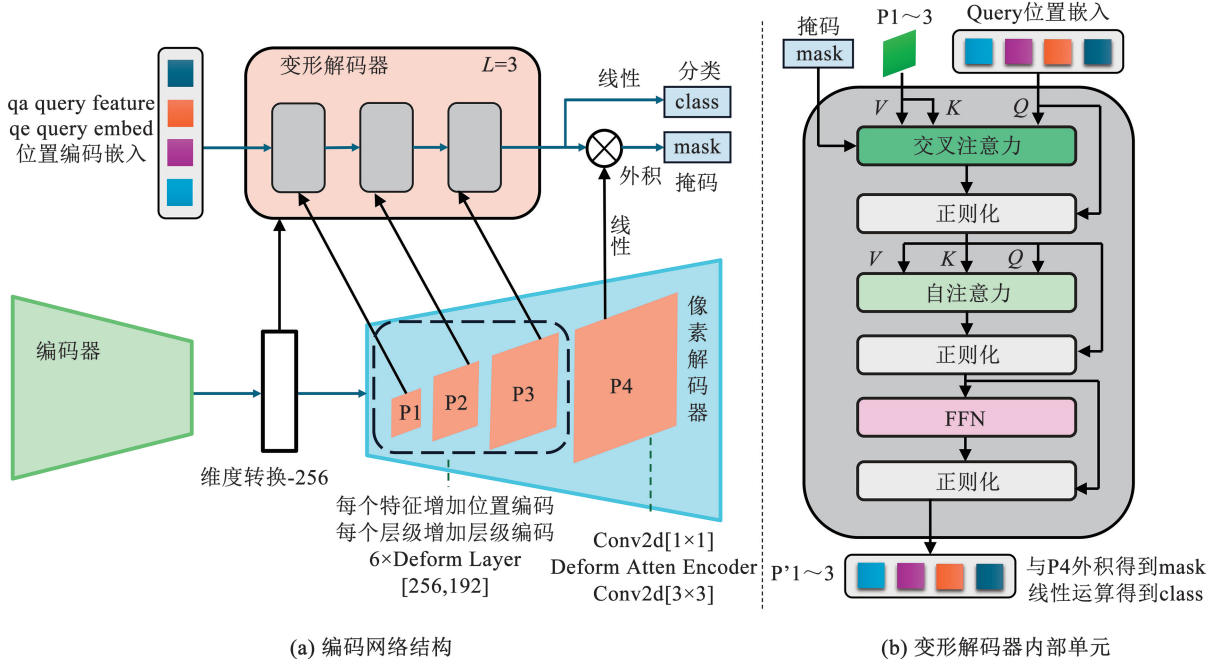


图2 解码网络结构与变形解码器内部基本单元

Fig. 2 Structure of the decoding network and basic internal unit of the deformable decoder

变形解码器在 Transformer 编码器基础上进行优化,基于交叉自注意力机制,由多个解码层组成,解码层负责接收像素解码器的不同尺度特征,每个解码器层包含多头注意力机制和前馈神经网络,如图2(b)所示,负责利用全局信息对图像进行整体的

语义理解,有助于处理前背景不易分辨的情况,并提高对尺度多样性的适应能力。

自注意力 (self-attention) 和交叉注意力 (cross-attention) 是变形解码网络中用于捕获数据长距离依赖的关键机制。自注意力机制使模型在序列内部直

接计算注意力并聚焦于关键信息^[18],对于输入序列 X ,其包含 n 个像素,每个元素维度为 d ,通过 3 个线性变换分别得到查询预测(query, Q)、键(key, K)和值(value, V)矩阵。

$$\begin{aligned} Q &= XW^Q \\ K &= XW^K \\ V &= XW^V \end{aligned} \quad (1)$$

式中 $W^{Q,K,V}$ 是可学习的参数矩阵。

而后,计算每个元素对于其他所有元素的注意力权重

$$A = (Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

式中: A 为注意力权重 Attention; QK^T 为序列元素间的相似度; $\sqrt{d_k}$ 为缩放因子,以稳定梯度;softmax 归一化函数是将相似度分数转换为概率分布,用于对各位置分配注意力权重。

交叉注意力机制则用于处理两个独立输入序

列,其中 Q 源自像素解码器输出,而 K 和 V 序列源自编码器输出,同样通过对 V 加权求和,实现特征信息的筛选和注意力聚焦。

1.3 前向传播与损失函数

本文提出的框架基于查询-预测路径的前向传播机制,该框架将图像分割任务解耦为并行的形状查询与类别预测两个核心过程,通过独立输出缺陷掩码和类别标签,以灵活处理复杂场景和重叠缺陷的问题。模型分割形状由查询过程输出,输入维度为 $[B, Q, H, W]$, B 代表本批次输入的图像数量, Q 代表查询预测(query, Q),对应图像中独立缺陷的实例数,每个 Q 代表一个缺陷实例的形状,因此查询形状的输出为一张仅包含缺陷形状的掩码图片;缺陷类别则由预测过程输出,输出格式为 $[B, Q, C]$, C 代表不同的类别号,其中 0 为背景,裂缝、露筋、侵蚀、剥落分别为缺陷类别 1~4,其查询形状与类别相互对应,如图 3 所示。

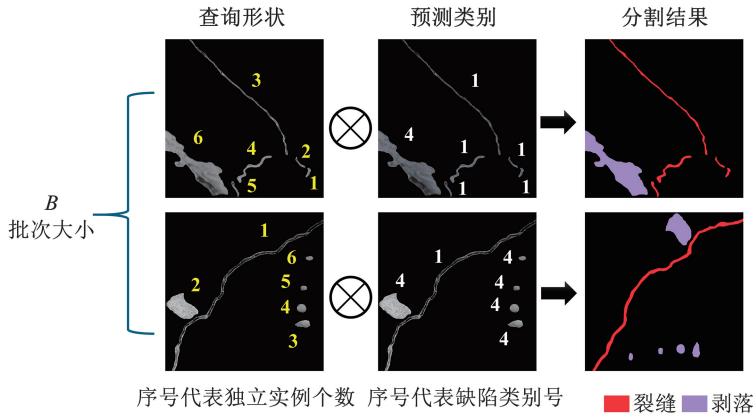


图 3 模型的查询-预测框架

Fig. 3 Query-prediction framework of the model

形状损失函数用于计算每一对匹配的预测掩码和真实掩码形状之间的差异,本文基于 Dice 系数(Dice coefficient)选用 Dice Loss 作为形状损失的度量以强化对少数类别的关注:

$$L_{\text{Dice}} = 1 - \text{Dice} = 1 - \frac{2 \times \sum_{i=1}^N p_i \times g_i}{\sum_{i=1}^N p_i + \sum_{i=1}^N g_i} \quad (3)$$

式中: p_i 为第 i 个像素的预测概率; g_i 为第 i 个像素的真实标签,通过直接计算预测掩码和真实掩码之间的像素级别重叠,作为衡量二者相似度的指标,即鼓励模型生成最大化 Dice 系数的预测。

为了增加少数类别的重要性,本文在标准交叉

熵损失的基础上引入对数归一化倍率作为每个类别的权重系数,计算公式为

$$L_{\text{WCE}} = - \sum_{i=0}^C w_i \cdot y_i \cdot \log(p_i) \quad (4)$$

式中: C 为类别总数,本文中 $C = 5$; y_i 为一个指示向量,若样本属于 C ,则 $y_i = 1$,否则 $y_i = 0$; w_i 为类别 C 的权重,即对数归一化倍率。

本文采用编码-解码网络实现桥梁表观缺陷的像素级分割,网络整体结构图如图 4 所示,其中 ε 为各类掩码的编码。编码器通过多阶段卷积与注意力机制提取多尺度缺陷特征,解码器利用并行路径重建缺陷形态,结合形状-语义联合优化损失函数提升分割精度。

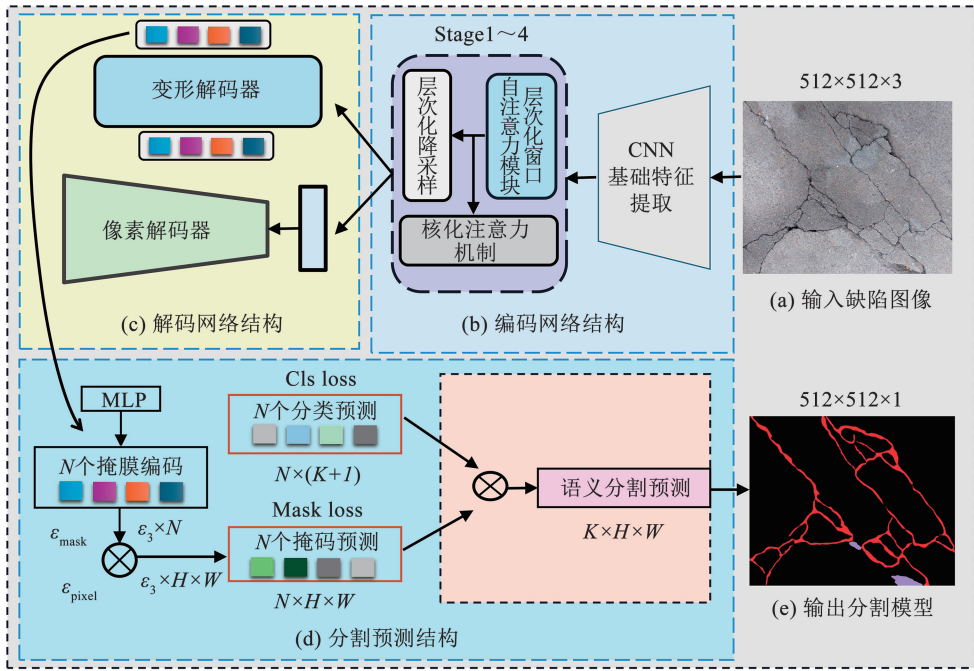


图4 网络整体结构

Fig. 4 Overall structure diagram of the network

2 试验数据采集

2.1 数据采集设备及流程

本文所使用的图像数据来源于由无人机进行采集的桥梁结构高清图像。所采用的无人机型号为大疆 M350 RTK。无人机最大信号有效距离 NCC/FCC 为 15 km,最大飞行速度为 23 m/s,最大载重为 2.7 kg。无人机搭载具有 4 000 万像素 3 轴稳定系统的 H30T 多传感器相机,其中影像传感器为 1/1.8 英寸的 CMOS,实际焦距为 7.1 ~ 172 mm,确保航拍图像分辨率可达 0.1 mm/pixel。

数据采集覆盖江苏、陕西等地多座混凝土桥梁(桥梁缺陷图像数据采集如图 5 所示)。采集作业遵循以下规范:无人机飞行高度动态调整于 20 ~ 50 m,航速控制在 5 m/s 以内,航向重叠率 80%、旁向重叠率 70%,通过带状航路规划规避桥塔、拉索等遮挡物。采集过程中实时监测图像质量,剔除模糊或畸变帧,最终获取 3 023 张高清缺陷图像,地面分辨率 0.1 ~ 0.2 m/pixel,满足规范 DB34/T 310022—2024《长大桥梁无人机巡检作业技术规程》。为确保拍摄后从二维图像中准确提取缺陷尺寸信息,本文采用以下技术策略:基于多视角图像的空间连续

性,利用运动恢复结构生成稀疏 3D 点云,并结合无人机惯性测量单元和激光雷达获取的高度、距离与姿态参数计算全局尺度基准。针对目标缺陷区域,通过至少 3 个不同视角的图像采集,构建局部稠密点云,运用三角测量原理解算缺陷实际尺寸,包括长度、宽度及面积。严格遵循跨视角尺寸校验机制,当同一缺陷在不同视角下的尺寸差异超过 5% 时,立即进行图像重拍。

2.2 数据集处理

为确保数据集的质量,所有图像在收集后均经过严格的筛选和预处理步骤。原始图像经对比度增强与几何校正后,使用 EISeg 软件进行像素级标注。标注标准严格区分 4 类缺陷:裂缝定义为宽度 ≥ 0.1 mm 的连续线状损伤;剥落为面积 ≥ 1 cm² 的混凝土剥蚀区域;露筋需暴露长度 ≥ 3 cm 且钢筋轮廓清晰;侵蚀表现为混凝土表面粉化或蜂窝状孔洞。

标注结果以 MS COCO 格式存储,生成每张图像的独立 JSON 文件及全局标注索引文件。数据集按 8:1:1 比例划分为训练集(2 418 张)、验证集(302 张)与测试集(303 张),4 类缺陷标注示例如图 6 所示。

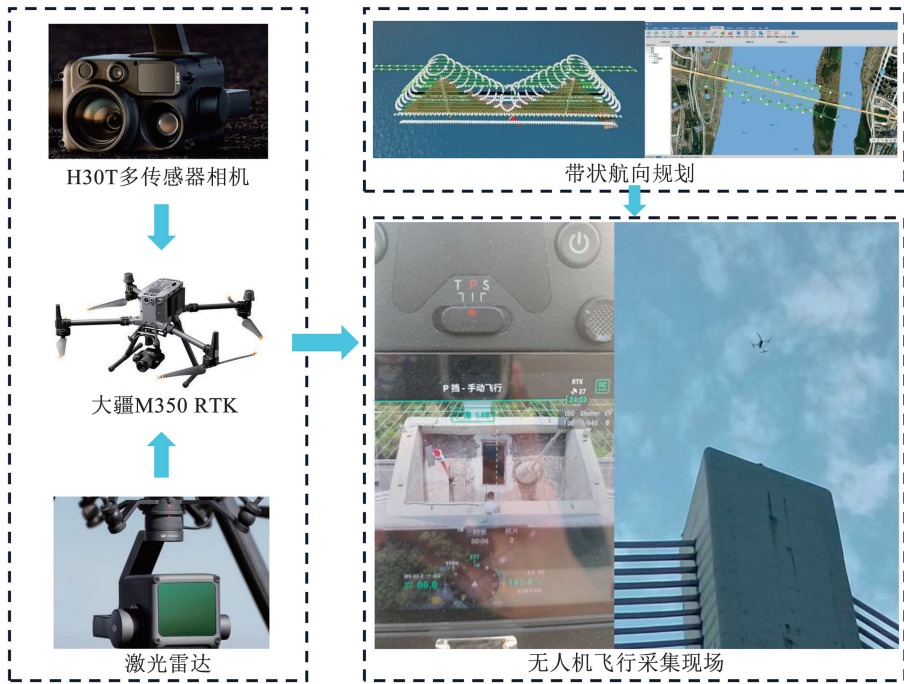


图 5 桥梁缺陷图像数据采集

Fig. 5 Bridge defect image data acquisition

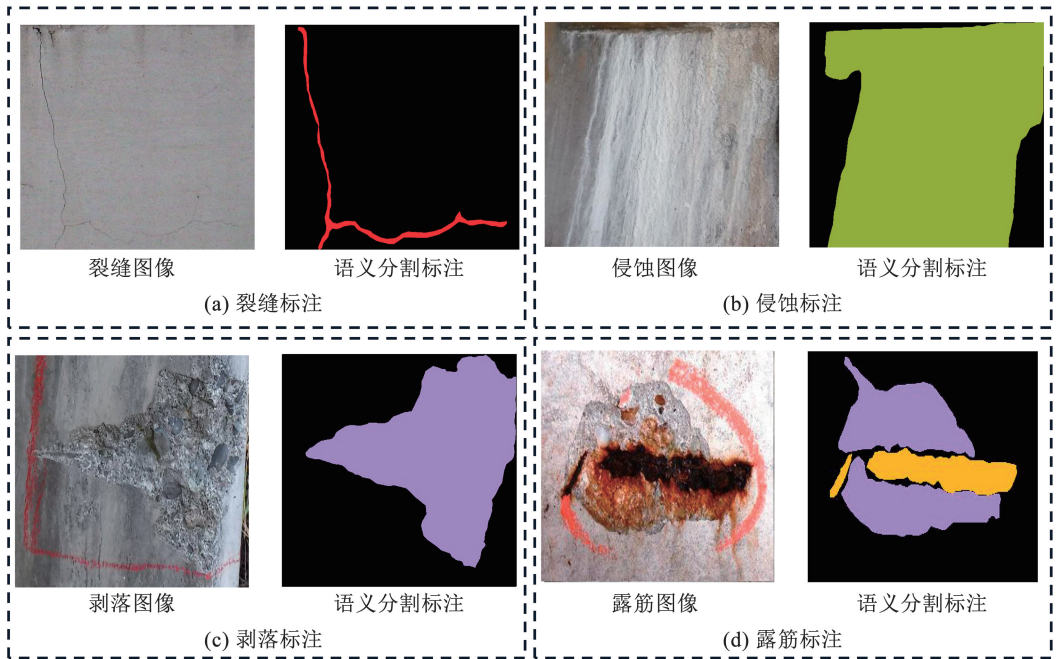


图 6 4 种缺陷标注处理示例图

Fig. 6 Example images of annotation processing for four types of defects

3 试验结果分析

3.1 语义分割任务评价指标和训练环境配置

为对混凝土桥梁表面缺陷精细化分割模型进行全面评估,本文将采用以下因素作为评价指标:像素准确率(pixel accuracy, PA)、均值像素准确率(mean pixel accuracy, mPA)、交并比(intersection over union, IoU)、平均交并比(mean intersection over union, mIoU)、Dice 系数(Dice coefficient)、平均 Dice

系数(mean Dice coefficient)、召回率(Recall)、平均召回率(mean Recall),具体如表 1 所示。

模型训练基于以下软硬件环境:硬件平台搭载 Intel Xeon Platinum 8352V 处理器(主频 2.10 GHz)与 NVIDIA RTX 4090 显卡(24 GB 显存),操作系统为 Ubuntu 20.04 LTS;软件环境采用 PyCharm Community Edition 2022.3 作为集成开发工具,深度学习框架为 PyTorch 1.11.0, CUDA 11.3 加速库支持并行计算。训练过程中启用混合精度训练

(AMP)与分布式数据并行(DDP),单批次最大输入尺寸设定为 512 × 512 像素。

表 1 深度学习视觉任务评价指标

Tab. 1 Evaluation metrics for deep learning vision tasks

编号	指标名称	特点及意义
1	PA, mPA	分割模型的查准率。
2	IoU, mIoU	真实形状与预测形状的相似程度。
3	Dice, mDice	真实形状与预测形状的相似程度,更倾向于小目标。
4	Recall, mRecall	查全率,模型对缺陷的捕捉能力,越高代表模型漏掉的缺陷越少。

3.2 整体性能对比研究分析

为验证本文提出的精细化分割模型的性能优势,将数据增强后的桥梁缺陷数据集分别输入 DeeplabV3 +、PSPNet、Unet、FastSCNN、SegFormer 等主流分割模型及本文模型进行对比试验。试验采用最大批量样本原则,在 24 GB 显存约束下动态调整各模型的批量尺寸,以确保计算资源的高效利用。训练过程中系统记录各模型的损失与精度变化,并以 500 次迭代为间隔采集验证集性能指标,试验结果如图 7 与表 2 所示。

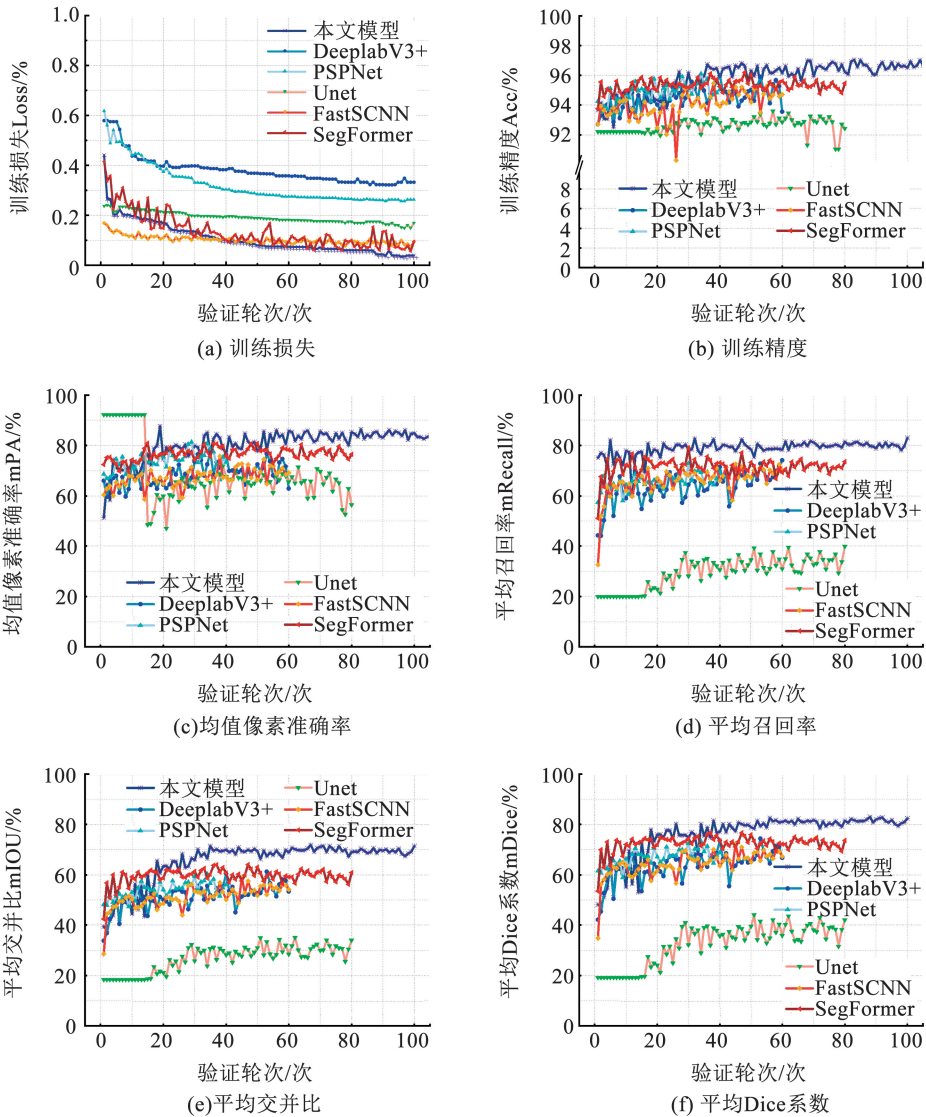


图 7 不同模型的整体性能评估指标随验证轮次演变曲线

Fig. 7 Evolution curves of overall performance evaluation metrics for different models over validation epochs

如图 7(a)所示,本文模型的训练损失曲线呈现显著优势,模型表现出快速收敛特性(本文模型损失值较 DeepLabV3 + 降低 71%)。训练初期损失值迅速下降,并在迭代后期稳定于较低水平。对比模型中,DeeplabV3 +、PSPNet、Unet 等模型的损失值下

降速率相对缓慢,且最终收敛值高于本文模型。FastSCNN 的损失曲线在迭代后期出现明显震荡,表明其优化过程存在不稳定性。在训练精度方面(图 7(b)),本文模型的精度值在训练早期即快速上升,并在迭代后期维持在较高水平,波动幅度显著

小于其他模型。DeeplabV3 + 与 SegFormer 的精度曲线虽呈上升趋势,但整体提升幅度低于本文模型,而 FastSCNN 在训练后期出现精度下降现象。

如表 2 所示,本文模型在验证集上取得最优性能:召回率 (Recall) 达 82.95%,较 DeeplabV3 + (74.08%)、PSPNet (72.82%)、Unet (39.91%)、

FastSCNN(72.68%)和 SegFormer(79.02%)分别提升 8.87%、10.13%、43.04%、10.27%和 3.93%。在平均交并比(mIoU)与平均 Dice 系数(mDice)指标上,本文模型分别以 71.91%和 83.02%的表现显著优于所有对比模型,其中 mIoU 指标较次优模型 SegFormer(64.05%)提升了 7.86%。

表 2 模型验证指标对比
Tab.2 Comparison of model validation metrics

模型	mIoU	mDice	mPA	mRecall
HDNet(本文模型)	71.91	83.02	88.08	82.95
DeeplabV3 +	61.18	74.30	77.22	74.08
PSPNet	60.33	73.30	81.28	72.82
Unet	35.14	44.05	92.21	39.91
FastSCNN	58.97	72.43	75.67	72.68
SegFormer	64.05	76.67	80.83	79.02

对比实验表明,本文模型在训练稳定性与泛化能力方面具有显著优势。其验证集精度曲线在迭代后期保持平稳,未出现对比模型中常见的过拟合波动。

3.3 消融实验分析

通过逐步退化核化注意力、层次化滑窗自注意力、交叉自注意力及 Dice 损失模块,验证各组件对性能的贡献,结果如图 8 所示。

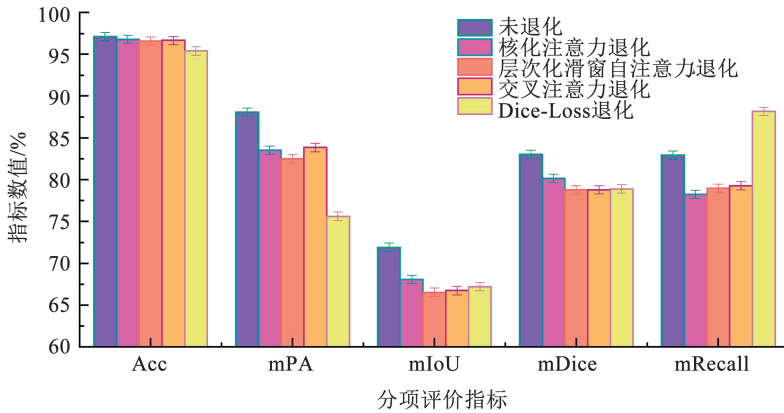


图 8 网络关键模块退化验证后指标变化情况

Fig. 8 Changes in metrics after degradation verification of key network modules

未经过退化的模型在大部分指标上达到了峰值 (Acc:96.71%, mPA:88.08%, mIoU:72.41%, mDice:84.02%, mRecall:83.95%),验证了多模块协同工作的有效性。核化注意力退化后 mRecall 下降幅度最大(5.83%),而 mDice 仅降低 2.15%,表明该模块通过特征核重构显著提升了缺陷区域的查全能力,但对形状相似度指标的直接影响有限。层次化

滑窗自注意力退化后各指标出现全域性显著下降,其中 mIoU 与 mDice 分别下降 5.92%与4.81%,证明该模块通过多尺度位置编码机制,对缺陷形状识别具有显著作用。交叉自注意力退化后性能小幅下降(mPA 下降 2.31%,mRecall 下降 2.96%),说明跨尺度特征交互虽能提升模型鲁棒性,但并非核心性能来源。Dice-loss 退化后出现指标分化现象:mRecall

异常升高 4.27%, 而 mPA 与 mIoU 分别下降了 11.15% 与 4.73%。这表明损失函数退化导致模型偏向背景预测, 印证了 Dice 损失对数据不平衡问题的校正作用。

实验结果与桥梁缺陷特性高度契合: 即层次化滑窗自注意力模块的有效性 (mIoU 贡献度达 7.7%) 源于其对桥梁表面多尺度纹理特征的适应性提取; 核化注意力模块在 mRecall 上的主导作用 (贡献度 7.2%) 验证了局部特征增强对细小缺陷检测的重要性; Dice 损失函数通过平衡类别权重, 缓解了数据分布不平衡问题。

3.4 不同类别缺陷分割效果对比分析

本实验基于训练日志 (总迭代次数 1.2×10^5 , 即 120 个训练周期) 提取类间评估数据, 以 500 次迭代为间隔记录验证集指标。图 9 展示了裂缝、露筋、

侵蚀、剥落 4 类缺陷的独立验证曲线。

经过对这些不同缺陷间验证指标的得分、一致性、稳定性和关联性的分析可以得到模型对不同类别缺陷的学习效果。如图 9(a) 所示, HDNet 在 IoU 与 Dice 指标上保持稳定, 验证了模型对裂缝缺陷这类线性纹理特征的精确识别能力。如图 9(c) 所示, 在侵蚀类别缺陷方面, HDNet 训练初期指标波动显著, 随着训练进程逐渐稳定, 表明模型通过自适应学习有效提升了不规则区域的分割鲁棒性。如图 9(b) 和 (d) 所示, HDNet 在露筋缺陷上呈现高像素准确率 (PA 均值 97.21%) 与低召回率 (Recall 均值 85.40%) 的特性, 反映模型对此类结构简单缺陷的保守预测倾向。而剥落缺陷部分 HDNet 的召回率高达 81.91%, 但 PA 仅 75.94%, 表明存在因纹理相似导致的背景误判现象。

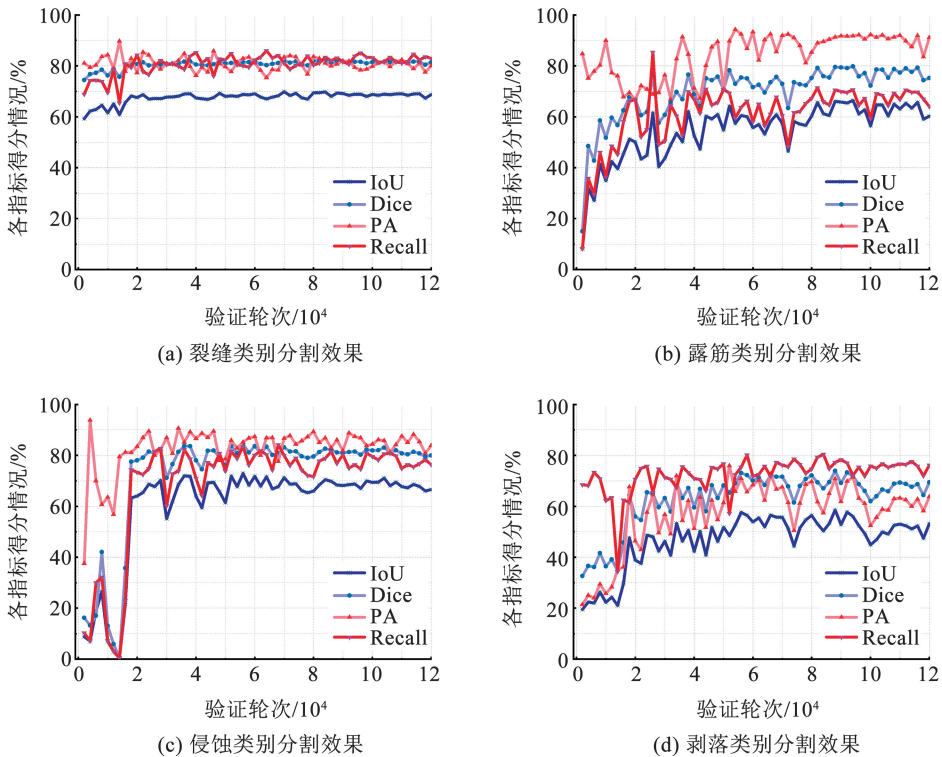


图 9 模型不同缺陷间性能得分情况

Fig. 9 Performance scores across different defect types

提取本文与其他 5 种模型训练过程中各类别验证指标的最佳值, 将本文模型的不同缺陷间评估指标与其他 5 种模型的结果进行了对比研究, 目的是了解不同网络在分割不同类别缺陷时的性能差异, 以及本文模型对 4 种缺陷识别能力的具体情况。

通过对表 3 中各种模型验证指标的横向对比分析发现, 本文所提出的模型在 4 种缺陷分割任务中均展现出最佳的性能, 特别是相较其他模型在分割精度上实现了显著提升。

表 3 不同模型对 4 种缺陷的分割验证指标得分对比

Tab. 3 Comparison of segmentation validation metrics scores for four types of defects among different models %

缺陷类别	模型	IoU	Dice	PA	Recall
裂缝	HDNet(本文模型)	70.82	82.23	89.69	86.47
	DeepLabV3+	64.36	78.32	87.80	80.03
	PSPNet	65.51	79.16	89.43	75.84
	FastSCNN	60.15	75.11	81.74	79.05
	SegFormer	66.41	79.82	90.42	82.84
	Unet	49.24	65.99	93.29	55.86
	露筋	HDNet(本文模型)	68.38	79.80	97.21
DeepLabV3+		45.26	62.31	87.07	68.04
PSPNet		46.80	63.76	91.05	57.29
FastSCNN		47.07	64.01	88.31	58.23
SegFormer		60.77	75.60	93.49	91.75
Unet		31.91	48.38	77.19	45.55
侵蚀		HDNet(本文模型)	75.44	85.35	95.34
	DeepLabV3+	63.00	77.30	96.23	85.10
	PSPNet	65.06	78.83	86.72	81.60
	FastSCNN	60.82	75.64	96.81	93.26
	SegFormer	67.75	80.78	99.34	85.91
	Unet	0	0	0	0
	剥落	HDNet(本文模型)	61.66	73.94	75.94
DeepLabV3+		42.47	59.62	68.73	69.41
PSPNet		44.09	61.20	61.03	72.18
FastSCNN		40.20	57.34	53.55	71.94
SegFormer		45.88	62.90	71.26	75.39
Unet		15.74	27.21	34.76	33.01

表 3 结果表明:1) 裂缝分割综合表现最优秀,性能显著超过其他模型,本文模型 IoU 达 70.82%,较 DeepLabV3+ (64.36%) 提升 6.46%,Dice 系数(82.23%) 领先次优模型 SegFormer (79.82%) 2.41%,验证了层次化滑窗结构对细长裂缝边缘的增强作用。

2) 露筋缺陷方面,本文模型 IoU (68.38%) 显著优于 SegFormer (60.77%) 与 DeepLabV3+ (45.26%),且 PA 达 97.21%,证明核化注意力机制有效提升了裸露特征的定位精度。

3) 侵蚀识别方面,本文模型以 75.44% 的 IoU 领先 SegFormer (67.75%) 7.69%,Dice 系数(85.35%) 较 FastSCNN (75.64%) 提升 9.71%,显示多尺度特征融合对不规则区域的表征优势。

4) 由于客观上识别难度较高,在剥落类别上所有模型得分均偏低,但本模型在该类别中所有指标均取得了最高,IoU (61.66%) 较 SegFormer (45.88%) 提升 15.78%,Dice 系数(73.94%) 领先 11.04%,表明动态核更新机制有效缓解了类间混淆问题。

3.5 网络的缺陷感知行为可视化分析

本文选取 4 类典型桥梁表观缺陷图像作为测试样本,基于网络感知热力图可视化技术,对 SegFormer 基准模型与本文提出的改进模型 HDNet 进行特征感知机制对比研究。通过解析模型内部特征响应模式,揭示不同网络架构对复杂缺陷特征的认知差异,热力图示意图 10。

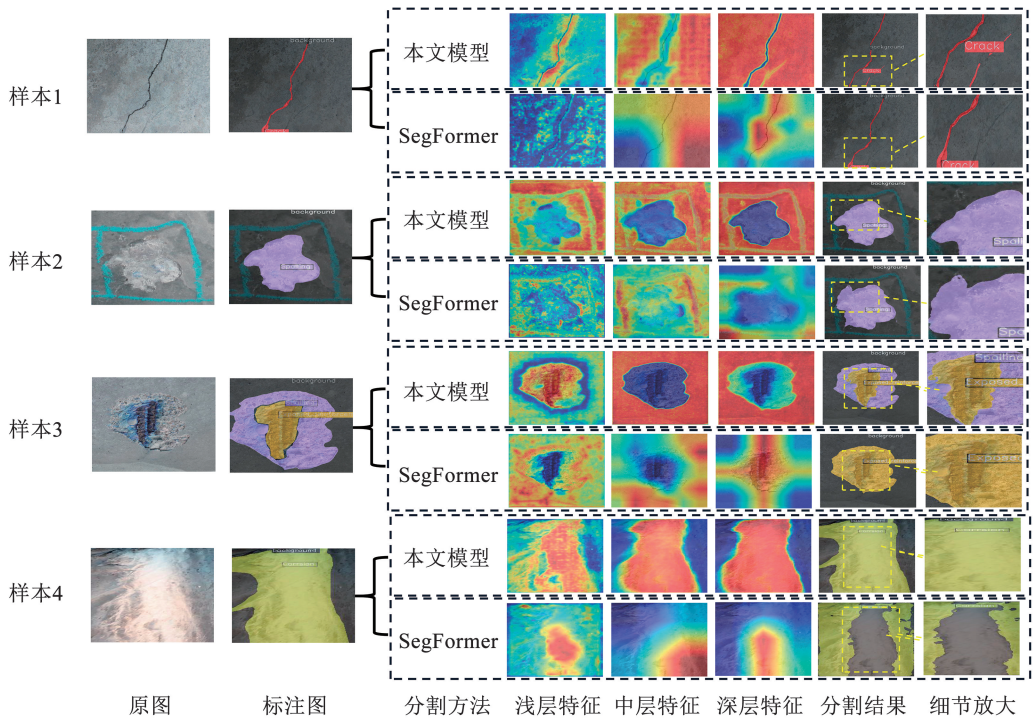


图 10 本文模型与 SegFormer 热力图对比

Fig. 10 Comparison of heatmaps between the proposed model and SegFormer

由图 10 可以看出:本文模型的编码网络通过桥梁材料纹理先验学习,在缺陷内部及边缘区域生成密集连续的热力响应,完整覆盖裂缝延伸趋势与剥落轮廓。而 SegFormer 模型仅在诸如严重侵蚀部位这类纹理对比度显著的局部区域产生零散响应,且最大激活区域面积不足本文模型的 40%,表明其全局特征整合能力受限,说明 HDNet 具有良好的纹理特征感知能力。同时,HDNet 边缘抗干扰性能明显强于 SegFormer,针对剥落区域的人工笔迹干扰,本文模型通过像素解码器与变形解码器的协同优化,在干扰区域维持稳定热力值 < 0.2 的低响应水平,准确锁定真实缺陷边界。SegFormer 模型因缺乏物理边缘判别机制,在笔迹区域产生与真实缺陷相近的响应强度,导致误判现象。在标注未区分的露筋-剥落伴生缺陷样本中,本文模型通过跨层级特征交互机制,建立两类缺陷的关联响应,准确解耦复合缺陷的语义边界。SegFormer 模型因注意力机制局限

于单样本特征提取,未能迁移其他样本的露筋特征权重,导致对伴生关系的认知缺失。实验结果验证,针对桥梁检测场景的特性优化网络结构,可有效提升模型对复杂表观缺陷的认知精度与稳定性。

3.6 实桥测试结果分析

实桥测试选取西安某大桥作为测试对象,测试过程中,采用双模传感器:12.80 mm × 9.60 mm 传感器用于裂缝检测,5.76 mm × 4.29 mm 传感器用于其他缺陷采集。在采集主梁、桥墩等部位的 4 096 × 3 072 像素高分辨率图像时同步记录物距与焦距参数。通过手持式激光测距仪、裂缝宽度观测仪及数字刻度尺对缺陷尺寸进行 5 次重复测量,取均值作为真值基准。实桥测试设备及过程见图 11。

表观裂缝尺寸测量结果如表 4 所示,表中宽度指表观裂缝最大宽度,长度指表观裂缝最大长度,其他 3 种缺陷实测结果如表 5 所示,表中宽度指缺陷横向尺寸,长度指缺陷纵向尺寸。



图 11 实桥测试设备及过程

Fig. 11 Actual bridge testing equipment and process

表 4 表观裂缝尺寸识别值与实测值对比

Tab. 4 Comparison between identified and measured values of apparent crack dimensions

编号	物距/ mm	宽度识别值/ mm	宽度实测值/ mm	宽度相对误差/ %	长度识别值/ mm	长度实测值/ mm	长度相对误差/ %
1	164	0.406	0.40	1.50	40.764	39.6	2.94
2	165	0.542	0.52	4.23	53.755	51.3	4.79
3	150	0.697	0.68	2.50	37.338	35.8	4.38
4	160	1.295	1.26	2.78	66.954	70.1	-4.49
5	150	0.943	0.94	0.32	44.935	47.0	-4.36
6	150	0.712	0.68	4.71	43.642	41.7	4.76
7	130	0.386	0.38	1.58	15.643	16.2	-3.59
8	135	0.283	0.28	1.07	28.367	27.1	4.49
9	125	0.387	0.38	1.84	17.915	17.2	4.04
10	130	0.348	0.34	2.35	14.335	13.7	4.50

表 5 其他缺陷尺寸识别值与实测值对比

Tab. 5 Comparison between identified and measured values of other defect dimensions

缺陷	编号	焦距/ mm	物距/ mm	长度识别值/ mm	长度实测值/ mm	长度相对误差/ %	宽度识别值/ mm	宽度实测值/ mm	宽度相对误差/ %
剥落	1	19.4	2 500	196.091	188.79	3.87	697.452	670.92	3.95
	2	8.7	210	60.367	62.37	-3.21	11.329	10.93	3.61
	3	9.9	830	63.845	61.93	3.09	65.332	68.62	-4.79
	4	8.7	130	43.128	41.48	3.98	26.945	25.90	4.03
侵蚀	1	19.4	3 700	124.860	130.85	-4.57	139.236	144.10	-3.37
	2	19.4	4 000	86.802	84.10	3.21	163.505	168.23	-2.81
露筋	1	19.4	1 700	19.267	18.80	2.50	24.431	23.67	3.23
	2	19.4	1 700	110.463	105.96	4.25	25.169	24.37	3.27
	3	19.4	1 700	25.785	26.55	-2.86	39.951	39.39	1.43
	4	19.4	1 700	34.546	33.52	3.06	34.546	33.64	2.68

测试结果显示,表观裂缝宽度识别值与实测值的平均相对误差为 2.28%,长度平均相对误差为 4.23%,其中最小检测宽度为 0.28 mm。对于剥落、侵蚀、露筋等大尺寸缺陷,长宽识别误差均低于 5%。HDNet 识别值长度、宽度多数大于实际值,误差主要源于分割阈值保守导致边缘轻微外扩约 2%~5% 及非垂直拍摄的镜头畸变,但系统性偏大的识别结果为桥梁安全评估提供了天然冗余,符合工程检测“从严控制”的原则。HDNet 将缺陷尺寸测量相对误差稳定控制在 $\pm 5\%$ 以内,满足规范 JTG/T H21—2011《公路桥梁技术状况评定标准》桥梁检测的高精度要求,特别在裂缝宽度测量中实现了 ± 0.03 mm 的绝对精度,显著优于传统检测方法。

在后续研究中,建议针对 HDNet 模型开展结构化剪枝策略与自适应混合精度量化方法,旨在通过优化参数冗余与降低计算复杂度,缓解模型对硬件资源的依赖性,以推动 HDNet 在边缘计算设备上的轻量化部署,最终构建低功耗、高精度的桥梁病害智能检测系统。

4 结 论

本文基于深度学习技术,从使用深度学习模型执行桥梁表观缺陷的自动化检测任务出发,以无人机采集的多座桥梁的表观缺陷图像作为试验对象,针对混凝土桥梁表观缺陷检测问题设计了一个基于编码器-解码器架构的精细化语义分割模型 HDNet,并将 HDNet 与 SegFormer 等主流的语义分割模型进行了全面的对比,得出以下结论。

1) 编码器采用层级化窗口自注意力与跨层残差连接策略,结合核化注意力模块增强局部缺陷响应;解码器通过像素-变形双路径并行架构实现全

局语义解析与形态重建。实验表明,多尺度特征融合策略显著提升了模型对复杂缺陷的建模能力。

2) 构建形状查询-类别预测双路径前向传播机制,联合优化语义分割与边缘检测任务。该设计通过动态平衡不同类别特征权重,有效缓解了数据分布不平衡问题,提升了对少数类缺陷的识别鲁棒性。

3) 在 1.2×10^5 次迭代训练中,HDNet 表现出快速收敛特性(损失值较 DeepLabV3+ 降低 71%),最终验证集指标达到 mIoU 71.91%、mDice 83.02%、mPA 88.08%、mRecall 82.95%,证明模型具有综合性能优势。消融实验验证了核化注意力(mRecall 提升 5.83%)、层次化滑窗注意力(mIoU 提升 5.92%)与 Dice 损失函数协同设计的必要性。

4) 与 SegFormer 的感知热力图对比结果显示,模型能够准确捕捉缺陷内部纹理与边缘细节特征,同时有效解耦露筋-剥落等伴生缺陷的语义边界,在复杂背景干扰下仍保持稳定的形状解析能力。实桥尺寸测试中,HDNet 将缺陷尺寸测量相对误差稳定控制在 $\pm 5\%$ 以内,验证了其在实际应用中的适用性。

参 考 文 献

- [1]《中国公路学报》编辑部. 中国桥梁工程学术研究综述·2024[J]. 中国公路学报, 2024, 37(12): 1
Editorial department of china journal of highway and transport. China bridge engineering academic research review · 2024 [J]. China Journal of Highway and Transport, 2024, 37(12): 1. DOI: 10.19721/j.cnki.1001-7372.2024.12.001
- [2] BAO Yuequan, CHEN Zhicheng, WEI Shiyin, et al. The state of the art of data science and engineering in structural health monitoring [J]. Engineering, 2019, 5(2): 234. DOI:10.1016/j.eng.2018.11.027
- [3] 董一庆, 王达磊, 潘玥, 等. 结构表观病害人工标记数据标准化批处理方法 [J]. 华南理工大学学报(自然科学版), 2021, 49(12): 113

- DONG Yiqing, WANG Dalei, PAN Yue, et al. Standardized batch processing method for manual labeling data of structural surface defects [J]. Journal of South China University of Technology (Natural Science Edition), 2021, 49(12): 113. DOI:10.12141/j.issn.1000-565X.210016
- [4] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436. DOI: 10.1038/nature14539
- [5] ZHU Jinsong, SONG Jinbo. Weakly supervised network based intelligent identification of cracks in asphalt concrete bridge deck [J]. Alexandria Engineering Journal, 2020, 59(3): 1307. DOI:10.1016/j.aej.2020.02.027
- [6] LIU Zhen, YEOH J K W, GU Xingyu, et al. Automatic pixel-level detection of vertical cracks in asphalt pavement based on GPR investigation and improved mask R-CNN [J]. Automation in Construction, 2023, 146: 104688. DOI: 10.1016/j.autcon.2022.104689
- [7] WANG Zhengfang, WANG Jing, CHEN Kefu, et al. Unsupervised learning method for rebar signal suppression and defect signal reconstruction and detection in ground penetrating radar images [J]. Measurement, 2023, 211: 112652. DOI:10.1016/j.autcon.2022.104689
- [8] 余加勇, 刘宝麟, 尹东, 等. 基于 YOLOv5 和 U-Net3+ 的桥梁裂缝智能识别与测量 [J]. 湖南大学学报(自然科学版), 2023, 50(5): 65
YU Jiayong, LIU Baolin, YIN Dong, et al. Intelligent identification and measurement of bridge cracks based on YOLOv5 and U-Net3+ [J]. Journal of Hunan University (Natural Sciences), 2023, 50(5): 65. DOI:10.16339/j.cnki.hdxzbk.2023056
- [9] 杨伟, 方虹苏, 唐湘松, 等. 桥梁裂缝病害检测的轻量化 YOLOv8-ALTE 算法 [J/OL]. 交通运输工程学报, 1-17 [2024-12-17]
YANG Wei, FANG Hongsu, TANG Xiangsong, et al. Lightweight YOLOv8-ALTE algorithm for bridge crack defect detection [J/OL]. Journal of Transportation Engineering, 1-17 [2024-12-17]. <http://kns.cnki.net/kcms/detail/61.1369.U.20241124.1901.002.html>
- [10] 骆剑彬, 姜绍飞, 沈圣, 等. 基于声呐成像的水下桩墩表观病害深度学习与智能检测 [J]. 土木工程学报, 2021, 54(7): 90
LUO Jianbin, JIANG Shaofei, SHEN Sheng, et al. Deep learning and intelligent detection of underwater pile pier surface defects based on sonar imaging [J]. China Civil Engineering Journal, 2021, 54(7): 90. DOI:10.15951/j.tmgxcb.2021.07.008
- [11] ZHANG Lingxin, SHEN Junkai, ZHU Baijie. A research on an improved Unet-based concrete crack detection algorithm [J]. Structural Health Monitoring, 2021, 20(4): 1864. DOI:10.1177/1475921720940068
- [12] 褚鸿鹄, 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
CHU Honghu, YUAN Huaqing, LONG Lizhi, et al. Cascade segmentation method for high-resolution bridge crack images based on transformer [J]. China Journal of Highway and Transport, 2024, 37(2): 65. DOI:10.19721/j.cnki.1001-7372.2024.02.006
- [13] SUN Hang, LU Dewen, LI Xi, et al. Research on multi-apparent defects detection of concrete bridges based on YOLOR [J]. Structures, 2024, 65: 106735. DOI: 10.1016/j.istruc.2024.106735
- [14] ELLENBERG A, KONTOSOS A, MOON F, et al. Bridge related damage quantification using unmanned aerial vehicle imagery [J]. Structural Control and Health Monitoring, 2016, 23(7): 1168. DOI: 10.1002/stc.1831
- [15] 中华人民共和国住房和城乡建设部. 混凝土结构设计规范: GB 50010—2010(2015 年版) [S]. 北京: 中国建筑工业出版社, 2015
Ministry of Housing and Urban-Rural Development of the People's Republic of China. Code for design of concrete structures: GB 50010—2010 (2015 Edition) [S]. Beijing: China Architecture & Building Press, 2015
- [16] 张富财, 许建龙, 包晓安. 基于稠密扩张卷积的图像语义分割模型 [J]. 计算机系统应用, 2022, 31(3): 19
ZHANG Fucui, XU Jianlong, BAO Xiaolan, et al. Image semantic segmentation model based on dense dilated convolution [J]. Computer Systems Applications, 2022, 31(3): 19. DOI:10.15888/j.cnki.csa.008376
- [17] FANG Jie, YANG Chen, SHI Yuetian, et al. External attention-based TransUNet and label expansion strategy for crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10): 19054. DOI: 10.1109/TITS.2022.3154407
- [18] 李林娟, 贺赞, 谢刚, 等. 跨层细节感知和分组注意力引导的遥感图像语义分割 [J]. 中国图象图形学报, 2024, 29(5): 1277
LI Linjuan, HE Yun, XIE Gang, et al. Cross-layer detail perception and grouped attention guided semantic segmentation of remote sensing images [J]. Journal of Image and Graphics, 2024, 29(5): 1277. DOI:10.11834/jig.230653

(编辑 吕雪梅)