



# 基于零样本深度学习的 码头表面剥落病害区域分割与定量计算\*

倪佳宁<sup>1</sup>, 王启明<sup>1</sup>, 朱瑞虎<sup>2</sup>, 胡艳<sup>1</sup>, 李成明<sup>1</sup>

(1. 河海大学理学院, 江苏南京 211100; 2. 河海大学港口海岸与近海工程学院, 江苏南京 211100)

**摘要:** 随着无人机与数字图像处理技术的发展, 基于机器视觉的表面病害识别方法因具有安全与快速性, 广泛应用于桥梁和道路等方面。由于码头剥落病害图像较复杂, 目前该方法难以实现码头表面剥落病害的精确分割与定量分析。提出一种基于零样本深度学习模型 SAM (Segment Anything Model) 与图像透射变换矫正等技术相结合的码头表面剥落病害区域分割与量化计算方法。SAM 算法能够有效克服混凝土剥落图像背景噪声多、灰度差异小的问题, 分割方法精度更高、受噪声影响更小; 进一步通过矫正、去噪、转换等图像处理操作, 实现了对剥落病害关键几何特征的量化计算。经实验室模型与现场图像验证表明, 新构建的方法泛化能力强、准确性高, 能够实现对码头混凝土建筑表面病害的准确和快速检测, 具有广泛的应用前景。

**关键词:** 结构健康监测; SAM 深度学习; 图像分割; 病害定量识别

中图分类号: U656

文献标志码: A

文章编号: 1002-4972(2024)02-0162-07

## Segmentation and quantitative calculation of wharf surface spalling disease regions based on zero-shot deep learning

NI Jianing<sup>1</sup>, WANG Qiming<sup>1</sup>, ZHU Ruihu<sup>2</sup>, HU Yan<sup>1</sup>, LI Chengming<sup>1</sup>

(1. College of Science, Hohai University, Nanjing 211100, China;

2. Ministry of Education College of Harbor, Coastal and Offshore Engineering, Hohai University, Nanjing 211100, China)

**Abstract:** With the development of unmanned aerial vehicle and digital image processing technology, surface disease recognition methods based on computer vision have been widely used in bridges and roads because of their safety and rapidity. However, due to the complexity of images, it is challenging to achieve accurate segmentation and quantitative analysis of wharf apparent diseases by current methods. This paper proposes a method for segmentation and quantitative calculation of wharf surface spalling disease regions based on zero-shot deep learning model SAM (Segment Anything Model) and image transmission transformation correction. SAM can effectively overcome the problems of concrete spalling images such as background noise, small gray difference and so on, thus being more accurate and less sensitive by noise. Further, the quantitative calculation of key geometric features of spalling diseases is realized through image processing operations such as correction, denoising and conversion. The experiments based on laboratory model and field images show that the new method has strong generalization ability and high accuracy, and can realize the accurate and rapid recognition of wharf concrete surface diseases and has a wide application prospect.

**Keywords:** structural health monitoring; SAM deep learning; image segmentation; quantitative disease recognition

收稿日期: 2023-05-18

\*基金项目: 国家重点研发计划项目(2022YFB3207400); 国家自然科学基金项目(51709093); 江苏省研究生科研与实践创新计划项目(422003263)

作者简介: 倪佳宁(2000—), 女, 硕士研究生, 从事码头表面病害识别研究。

通讯作者: 王启明(1980—), 男, 博士, 副教授, 从事港口结构安全监测与预警分析研究。E-mail: wqm@hhu.edu.cn

我国海港码头服役环境复杂, 大多已产生多种结构病害。为确保码头结构安全, 对其表面破损或病害状况的量化评估是重要任务之一<sup>[1]</sup>。目前, 对于码头桩基混凝土结构耐久性检查大多依赖于人工巡检, 需要人工划船近距离量测, 效率低、风险大、成本高。因此, 开发快速且安全的码头表观病害特征提取与非接触测量评估方案成为亟待解决的问题。

目前, 基于数字图像处理的计算机视觉技术已运用于桥梁、隧道、公路等建筑工程安全监测中<sup>[2-6]</sup>, 也有少数学者开始对海港码头表观损伤进行研究<sup>[7-8]</sup>。然而, 上述数字图像处理方法大多针对与背景灰度差异较强的裂缝或墙面剥落等进行分割识别, 且需要大量预处理操作。对于海港码头建筑, 其图像背景复杂, 病害内外部灰度差异不明显, 传统的图像分割方法难以准确提取病害轮廓。近年来, 随着机器学习技术的发展, 基于深度学习的方法如 FCN<sup>[9]</sup>、Mask R-CNN<sup>[10]</sup>等语义分割算法在结构安全监测领域展现出优异的性能。但上述深度学习算法均是监督学习模型, 需要获取大量的病害图片, 且标注与训练时间较长<sup>[11]</sup>, 而适合进行分析的码头病害照片获取工作量巨大, 上述诸多算法难以实际开展。

针对上述问题, 本文提出一种基于零样本深度学习模型 SAM (Segment Anything Model) 与图像透射变换等技术相结合的码头表观剥落病害区域分割与量化计算方法。相比于传统分割算法, SAM 分割方法精度更高、泛化能力更强。在 SAM 算法分割基础上, 进一步通过矫正、去噪、转换等操作, 实现对剥落病害关键几何特征的量化计算。最后, 通过现场试验进行新算法的验证, 结果表明新方法无需现场码头数字图像作为新的训练集, 即能实现对码头混凝土建筑表观病害的简单、准确和快速检测, 具有广泛的应用前景。

## 1 基于 SAM 与数字图像处理技术的病害特征提取与定量计算方法

### 1.1 零样本学习 (zero-shot learning) 模型 SAM

Meta 团队于 2023 年发布了可提示的分割模型

SAM<sup>[12]</sup>, 并在一个超过 10 亿个掩膜 (Mask, 用选定的图像、图形或物体对处理的图像进行遮挡, 以提取感兴趣区域或者结构特征) 的数据集 SA-1B 上对其进行预训练, 结合分割提示, 可以零样本迁移到新的图像上进行分割并输出预测掩膜, 在常见图像上具有很强的泛化能力。

#### 1.1.1 模型整体结构

SAM 模型主要结构包括 3 个组件: 图像编码器、提示编码器与掩膜解码器, 整体结构如图 1a) 所示。图像编码器使用经预训练后的视觉变换器 (ViT, vision transformer)<sup>[13]</sup> 处理高分辨率的图像输入; 提示编码器用于输入提示, 如点、框、掩膜等; 掩膜解码器能够有效地将图像嵌入和提示嵌入映射到输出掩膜, 其结构如图 1b) 所示。该模块使用一个 2 层解码器, 每个解码器层执行 4 个步骤, 包括自注意力机制、多层感知机 mlp 以及 2 个交叉注意力机制。

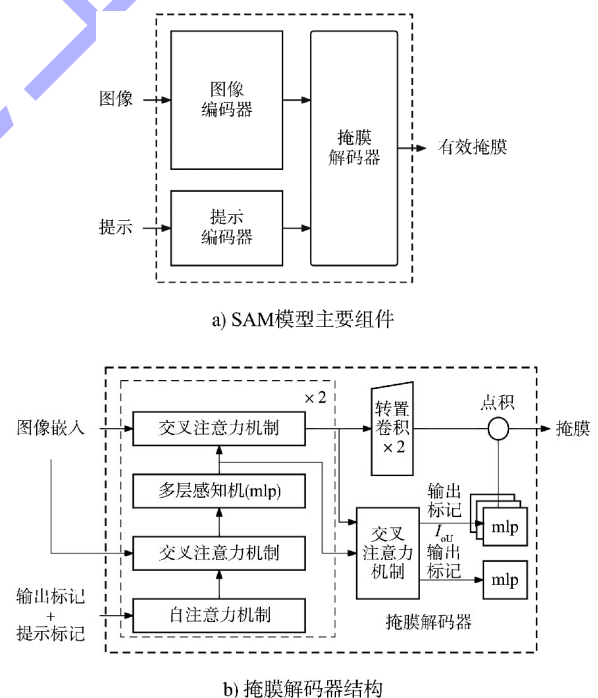


图 1 SAM 模型网络架构

#### 1.1.2 训练和损失

在预训练阶段, 模型会为每个训练样本模拟一系列提示, 并将预测掩膜与真实情况进行比较, 估计每个预测掩膜和它所覆盖对象之间的计数度量得分  $I_{ou}$ , 从而对输出的掩膜进行排序,

其公式为：

$$I_{ou} = \frac{T_p}{T_p + F_N + F_p} \quad (1)$$

式中： $T_p$  为前景被正确预测为前景的像素点个数； $T_N$  为背景被正确预测为背景的像素点个数； $F_p$  为表背景被错误预测为前景的像素点个数； $F_N$  为前景被错误预测为背景的像素点个数。 $I_{ou}$  反映预测掩膜与真实掩膜的交并比，能够有效评价图像分割的结果。

SAM 模型主要使用焦点损失  $L_F$ <sup>[14-15]</sup> 监督掩膜预测，它是一种动态缩放的交叉熵损失函数，通过调制因子  $(1-p_i)^\gamma$  可以动态降低训练过程中容易区分样本的权重，其公式为：

$$L_F(p_i) = -(1-p_i)^\gamma \lg p_i \quad (2)$$

$$\text{其中：} p_i = \begin{cases} p & (y=1) \\ 1-p & (\text{其他}) \end{cases} \quad (3)$$

式中： $y$  为前景或背景代表值，取值为 1 和 -1，分别代表前景和背景； $p$  为模型预测属于前景的概率，取值范围为 0~1； $p_i$  为模型预测某类别的概率； $\gamma \geq 0$  是可调节的聚焦参数， $\gamma$  值越大，模型的注意力就越聚焦于那些难分类的样本。

### 1.2 透射变换矫正图像

由于拍摄角度倾斜以及相机畸变等原因，所得到的病害图片会产生倾斜扭曲，难以实现精确定量计算，因此采用透射变换方法对图像进行矫正。透射变换是利用透视中心、像点、目标点三点共线的原理，按透视旋转定律使承影面绕迹线旋转某一角度，破坏原有的投影光线束后，仍能保持承影面上投影几何图形不变的变换，见图 2。

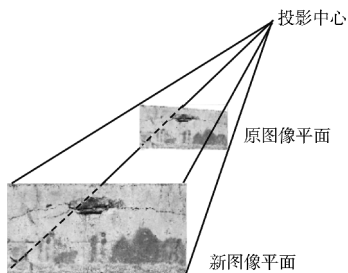


图 2 投影过程

透射变换公式为：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

式中： $X$ 、 $Y$ 、 $Z$  为透射变换后的三维坐标， $x$ 、 $y$  为透射变换前的二维坐标。在进行透射变换得到三维坐标后，将坐标值再次转换到二维坐标中，其转换公式为：

$$\begin{cases} x' = \frac{X}{Z} = \frac{a_1x + b_1y + c_1}{a_3x + b_3y + c_3} \\ y' = \frac{Y}{Z} = \frac{a_2x + b_2y + c_2}{a_3x + b_3y + c_3} \end{cases} \quad (5)$$

式中： $c_3=1$ ， $x'$  和  $y'$  为透射变换的最终计算结果。

### 1.3 中值滤波算法和闭运算

中值滤波是一种典型的非线性滤波技术，其原理是用像素点邻域灰度值的中值来代替该像素点的灰度值。中值滤波算法的公式可表示为：

$$g(x,y) = \text{median} \{f(x-i,y-i)\} \quad (i,j) \in S \quad (6)$$

式中： $g(x,y)$  和  $f(x,y)$  为像素的灰度值， $S$  为窗口。算法选用  $(2n+1) \times (2n+1)$  的窗口沿着图像的行和列方向滑动，并对其内的像素灰度值进行排序，用中间值代替窗口中心位置的像素灰度值。

开运算和闭运算均属于形态学操作，闭运算对图像先进行膨胀操作，即向外扩张图像边界点，其次进行腐蚀操作，即沿着图像边界向内收缩。本文使用中值滤波对复杂病害图像的二值掩膜的内外部噪点进行初步去噪，而闭运算操作能够进一步填充掩膜内部的黑色噪点，优化病害特征提取的效果，方便后续病害几何特征尤其是面积的计算。

### 1.4 基于转换系数的病害几何特征计算方法

为了确定图像中剥落部分的几何尺寸，需要使用图像中已知尺寸的参考物确定像素长度转换为真实物理长度单位的转换系数，该方法称为靶标法，其转换公式为：

$$L_r = k l_p \quad (7)$$

式中： $L_r$  为参照物真实的物理长度，mm； $l_p$  为对应的像素个数，pixels； $k$  为转换系数，即一个像素

点对应的真实物理长度, mm/pixel。

将每一粒像素点看成小正方形, 其对应的真实面积为  $k^2$ , 通过统计掩膜图像中病害区域白色像素点个数  $s_p$ , 可以粗略求得剥落部分的真实面积  $S_i = k^2 s_p$ , 从而达到对病害区域的定量计算。

### 1.5 病害特征提取与定量计算方法总体流程

本文提出的码头结构表观病害识别方法通过搭建深度学习模型 SAM 并结合数字图像处理技术来进行, 实现对码头结构病害分割等相关任务的系统性检测, 工作流程见图 3。

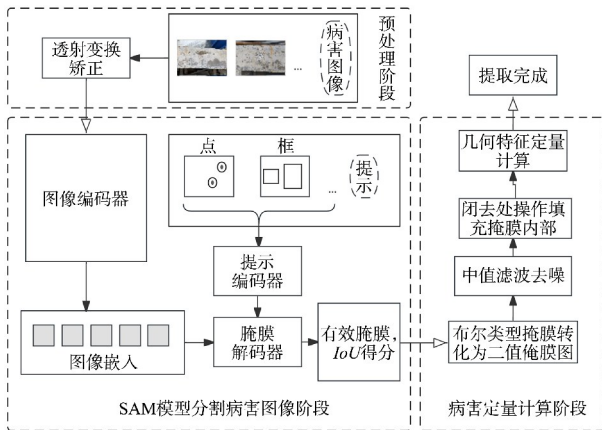


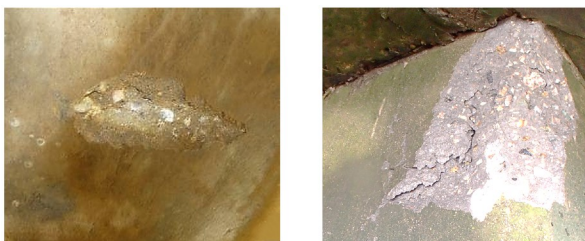
图 3 总体流程

## 2 试验结果与分析

### 2.1 典型剥落病害分割与算法对比

#### 2.1.1 典型剥落病害分割流程

本文采用福建某高桩码头采集的真实病害图像进行实例分割与算法对比, 选取 2 张典型的剥落病害图像进行试验。图 4a) 为面板剥落损伤, 图 4b) 为桩基剥落损伤。病害大小适中且处于图像中部, 以便于进行分割与算法对比, 使用数据集 SA-1B 预训练后的 SAM 模型对病害图像进行分割。



a) 图像A, 面板剥落损伤

b) 图像B, 桩基剥落损伤

图 4 典型剥落病害图像

为了验证 SAM 模型的分割效果, 对剥落图像进行手动标注, 并使用基于边缘的分割方法 Canny 算法和基于阈值的分割方法 Otsu 算法<sup>[16]</sup>与 SAM 模型分割得到的二值掩膜图像进行对比。采用  $I_{ou}$ 、精确率  $P$ 、召回率  $R$ 、综合  $F_1$  分数 4 个经典评价指标比较与人工标注掩膜的差异, 从而评估算法有效性, 其公式为:

$$P = \frac{T_p}{T_p + F_p}, R = \frac{T_p}{T_p + F_n}, F_1 = \frac{2PR}{P + R} \quad (8)$$

式中: 精确率  $P$  表示被正确预测的像素点占所有预测为前景的像素点的比例, 召回率  $R$  表示被正确预测的像素点占所有实际为前景的像素点的比例,  $F_1$  分数表示精确率和召回率的调和平均。

除上述经典评估指标外, 拟增加豪斯多夫距离  $H$  (Hausdorff distance) 衡量掩膜图像与人工标注图像的相似度, 通过测量度量空间的 2 个子集之间的距离描述 2 组点集之间的相似程度, 距离越小代表分割结果越准确。豪斯多夫距离公式如下:

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (9)$$

其中:

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|, h(B, A) = \max_{b \in B} \min_{a \in A} \|b - a\| \quad (10)$$

式中:  $\|a - b\|$  和  $\|b - a\|$  为点集  $A$ 、 $B$  间的范式距离, pixels。

为了使基于传统数字图像处理技术的分割方法达到更好的分割效果, 在分割前对图像进行灰度化操作, 对用 Canny 算法和 Otsu 算法进行提取后的二值图像进行开闭运算, 同时删除病害掩膜以外的连通区域, 填充内部孔洞, 从而得到完整的病害掩膜的二值图像。

#### 2.1.2 分割结果与算法对比分析

表 1 与图 5、6 展示了分割的最终结果, 可以看出, 在  $I_{ou}$  值、准确度、 $F_1$  得分、豪斯多夫距离指标方面, SAM 方法的效果明显优于基于数字图像处理的方法, 在召回率方面表现也比较好。而 Canny 算法和 Otsu 算法均会出现多识别的情况, 因此召回率会比较高, 且掩膜边缘模糊不连

续、噪点多，其原因为传统的数字图像处理方法主要识别图片中的灰度值、纹理、边缘等特征，但病害图片普遍存在分辨率低、背景复杂等问题，所以在分割时容易造成误识别，且在分割图像时需要做大量针对性的处理才能得到较好的结果，而SAM算法在效率和精度方面远超传统分割算法。

表1 不同算法分割效果对比

图像	分割方式	$I_{ou}$	$P$	$R$	$F_1$	$H/\text{pixels}$
A	Canny	0.798	0.810	0.982	0.888	92.574
	Otsu	0.265	0.295	0.719	0.418	107.331
	SAM	0.914	0.999	0.915	0.955	7.211
B	Canny	0.640	0.696	0.889	0.780	266.160
	Otsu	0.793	0.818	0.964	0.885	177.654
	SAM	0.974	0.999	0.976	0.987	49.406

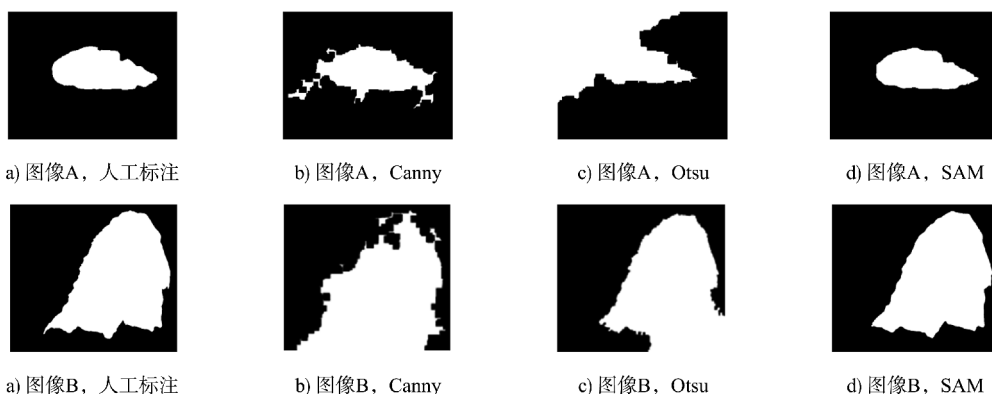
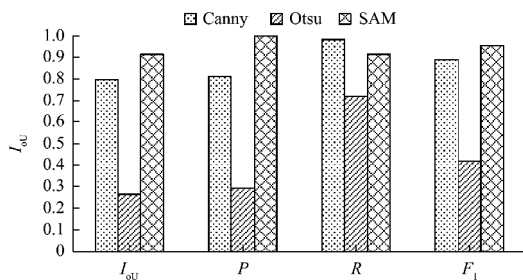
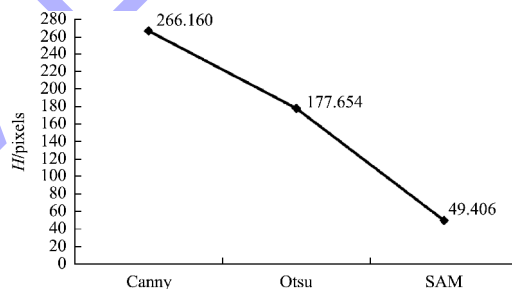


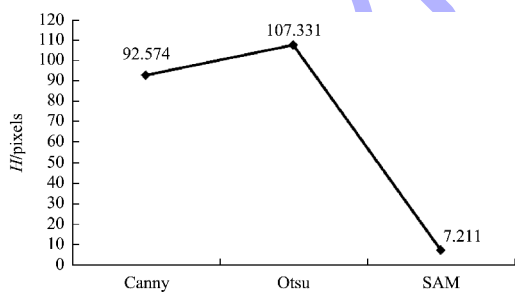
图5 各算法分割图像对比



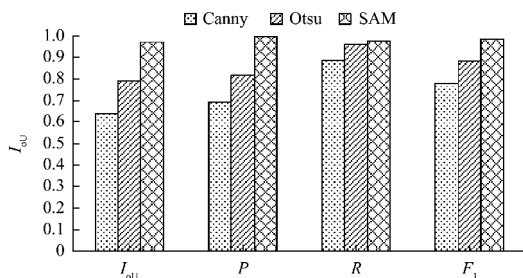
a) 图像A经典评估指标对比



d) 图像B豪斯多夫距离对比



b) 图像A豪斯多夫距离对比



c) 图像B经典评估指标对比

图6 各算法评估指标对比

## 2.2 复杂剥落病害分割与定量计算

### 2.2.1 基于高桩码头结构试验模型的剥落病害分割与定量计算流程

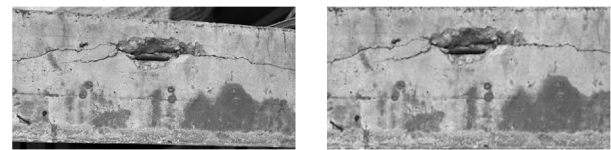
采用实验室等比例高桩码头结构试验模型(图7)模拟真实场景下的病害分割提取验证方法的定量准确性。试验模型为混凝土梁板结构,采用C30混凝土整体浇筑,由3跨4个排架组成,长1.44 m、宽0.658 m,码头面板厚0.11 m、横梁高0.12 m、纵梁高0.08 m,基桩采用双直桩、单桩和叉桩3种形式,模型表面人为建立剥落、露筋、裂缝等多种复杂病害。图像矫正结果如图8所示,框取矫正后图像的中间部分作为SAM模型的提示,其效

果如图 9a) 所示, 此时 SAM 模型对剥落部分的分割的效果最好, 其二值掩膜如图 9b) 所示。进一步使用中值滤波算法和闭运算对掩膜进行去噪, 见图 9c)。在实际应用中, 需要根据掩膜预测结果选择是否需要去噪以及选择合适的去噪方式。

使用图像中现成的参照物即面板的垂直厚度求取转换系数  $k$ 。在真实工况下, 面板、桩基与梁等构件的真实物理尺寸均可通过查阅相关图纸手册得到。已知面板的真实厚度  $D$  在图像中对应的像素长度为  $d$ , 则转换系数  $k = D/d$ 。根据转换系数, 通过计算病害掩膜某处高度的像素点个数  $h_p$  和某处宽度的像素点个数  $w_p$ , 求得真实高度  $H_r$  和真实宽度  $W_r$ 。



图 7 高桩码头结构试验模型



a) 矫正前 b) 矫正后

图 8 图像局部矫正效果



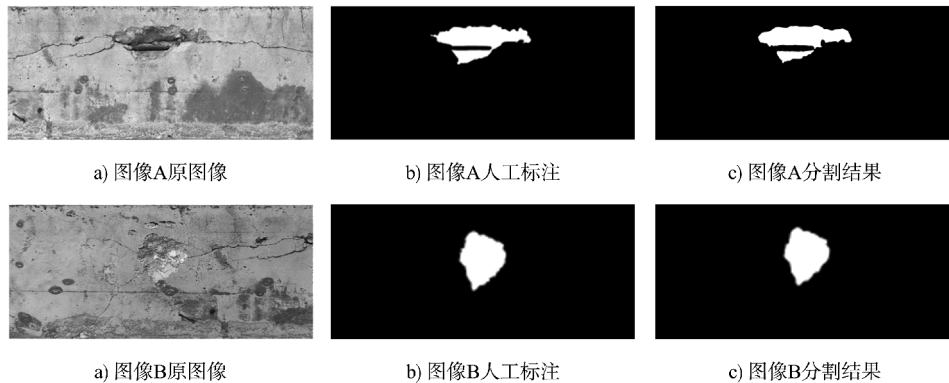
a) 病害分割 b) 预测二值掩膜 c) 去噪后二值掩膜

图 9 病害分割与去噪效果

### 2.2.2 剥落病害分割与几何特征计算结果分析

剥落病害分割与其特征计算结果如图 10 与表 2 所示, 通过与真实长度的比较可知, 病害某处高度  $H_r$  与宽度  $W_r$  误差较小, 均在可接受范围内, 证明该定量计算方法的精度能够达到实际工

程的需求。图像 1 的  $I_{oU}$  值为 0.874, 原因是该处病害比较复杂, 且 SAM 模型对中间露筋部分造成误识别。通过换算得到 2 幅图剥落病害真实面积分别为 1 154.42、880.19  $\text{mm}^2$ , 从而为后续评估构件的损伤程度提供依据。



a) 图像A原图像 b) 图像A人工标注 c) 图像A分割结果  
a) 图像B原图像 b) 图像B人工标注 c) 图像B分割结果

图 10 病害分割效果

表2 病害几何特征定量计算结果

图像	$I_{ou}$	$k/(mm \cdot pixel^{-1})$	像素/pixels	换算值/mm	实测值/mm	绝对误差/mm	相对误差/%	
A	高度 $H$	0.874	0.060 076	453	27.21	26.50	0.71	2.68
	宽度 $W$			1 276	76.67	81.20	4.53	5.58
B	高度 $H$	0.965	0.063 510	658	41.79	41.90	0.11	0.26
	宽度 $W$			511	32.45	34.00	1.55	4.56

### 3 结论

1) 为码头建筑外观病害的自动定量识别提出一种基于深度学习与数字图像处理技术相结合的表观病害区域分割与几何特征定量计算方法,对码头安全监测和维护工作具有重要的工程应用价值。

2) 分割模型 SAM 无需加入新的数据进行训练即可零样本迁移到新的领域,适合用于码头病害监测工作中。面对背景复杂、灰度差异小的图片,该方法  $I_{ou}$  能够达到 0.9 以上,具有广泛的研究前景。此外,提出一种剥落病害的定量识别方法,无需近距离测量和复杂的标定流程即可计算病害高度、宽度以及面积等关键几何特征,相对误差在 6% 以内,是一种安全且便捷的测量方法。

3) 现场结构病害常为多种类型病害的综合,即拍摄的 1 张图片中可能存在多种病害类型,后续研究将结合多目标检测方法,突出不同病害区域,综合评价码头结构的健康状态。

### 参考文献:

- [1] 魏明晖,严锋,关战伟.海港高桩码头结构维修技术应用[J].水运工程,2016(3):167-171.
- [2] PANERU S, JEELANI I. Computer vision applications in construction: Current state, opportunities & challenges[J]. Automation in construction. 2021, 132 (12): 103940. 1-103940. 17.
- [3] 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.
- [4] SHAO C, CHEN Y H, XU F H, et al. A kind of pavement crack detection method based on digital image processing [C]//IEEE. 2019 IEEE 4th advanced information technology, electronic and automation control conference. Chengdu: IEEE, 2019: 397-401.
- [5] 何昊,贺福强,谢丹,等.基于SVM和改进区域生长法的桥梁裂缝分割算法[J].公路交通科技,2022,39(11):115-123.
- [6] 夏子祺,马临原,单伽程,等.基于计算机视觉的建筑外墙剥落和裂缝两阶段检测方法[J].建筑结构学报,2023,44(2):207-216.
- [7] 李嘉民,应宗权,杨帅,等.基于分形理论和二次分割的图像裂缝特征提取方法与应用[J].水运工程,2023(4):166-173.
- [8] 陈秀瑛.港口码头的老化机理与病害评估[J].水运工程,2010(2):95-98.
- [9] LI S, ZHAO X, ZHOU G. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network [J]. Computer-aided civil and infrastructure engineering. 2019, 34(7): 616-634.
- [10] KIM B, CHO S. Automated multiple concrete damage detection using instance segmentation deep learning model[J]. Applied sciences. 2020, 10(22): 8008.
- [11] 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(S02):1439-1445.
- [12] ALEXANDER K, ERIC M, NIKHILA R, et al. Segment anything[EB/OL]. (2023-04-05) [2023-05-02]. <https://arxiv.org/abs/2304.02643>.
- [13] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: transformers for image recognition at scale[EB/OL]. (2020-11-22) [2023-05-04]. <https://arxiv.org/abs/2010.11929>.
- [14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[EB/OL]. (2017-08-07) [2023-05-04]. <https://arxiv.org/abs/1708.02002>.
- [15] MILLETARI F, NAVAB N, AHMADI S A: Fully convolutional neural networks for volumetric medical image segmentation [C]// IEEE. 2016 Fourth international conference on 3D vision (3DV). San Francisco: IEEE, 2016: 565-571.
- [16] 杨陶,田怀文,刘晓敏,等.基于边缘检测与Otsu的图像分割算法研究[J].计算机工程,2016,42(11):255-260,266.

(本文编辑 王传瑜)