

## 基于残差双通道注意力 U-Net 的古代壁画病害检测

赵辉荣, 余映\*, 陈安, 倪雪莹, 王信超

(云南大学信息学院 昆明 650091)  
(yuying.mail@163.com)

**摘 要:** 针对现有的古代壁画病害检测方法难以准确地检测壁画病害区域的问题, 提出一种基于残差双通道注意力 U-Net 的古代壁画病害检测模型. 首先设计残差双通道模块代替 U-Net 中的编码器和解码器, 构建具有多分辨率分析能力的网络检测复杂背景中不同尺度的壁画病害区域; 然后加入多尺度注意力门融合高层和低层的互补特征, 使网络能突出壁画病害区域的显著特征; 最后设计混合域注意力模块抑制壁画背景信息的干扰, 进一步准确地定位壁画病害区域; 此外, 采用多阶段损失相加的方式提高网络模型的性能. 实验结果表明, 在敦煌莫高窟壁画数据集和云南少数民族壁画数据集上, 所提模型的检测结果在视觉感受方面优于其他对比方法, 在  $F$ -score 指标上分别达到了 0.807 7 和 0.728 9, 均高于其他对比方法.

**关键词:** 古代壁画病害检测; U-Net; 残差双通道; 多尺度注意力门; 混合域注意力

**中图法分类号:** TP391.41 **DOI:** 10.3724/SP.J.1089.2023-00430

## Ancient Mural Disease Detection Based on Residual Dual Channel Attention U-Net

Zhao Huirong, Yu Ying\*, Chen An, Ni Xueying, and Wang Xinchao

(School of Information Science and Engineering, Yunnan University, Kunming 650091)

**Abstract:** To address the challenge of accurately detecting disease areas in ancient murals with existing methods, a novel ancient mural disease detection model based on a residual dual channel attention U-Net is proposed. Firstly, a residual dual channel module is designed to replace the encoder and decoder in U-Net, constructing a network capable of multi-resolution analysis to detect mural disease areas of varying scales against complex backgrounds. Then, a multi-scale attention gate is incorporated to fuse complementary features from both high and low levels, enabling the network to emphasize salient features of mural disease areas. Finally, a mixed domain attention module is designed to suppress the interference of mural background information, further accurately locating mural disease areas. Additionally, a multi-stage loss addition approach is employed to improve the performance of the network model. Experimental results on the Dunhuang Mogao Grottoes mural dataset and the Yunnan ethnic minority mural dataset demonstrate that the proposed model outperforms other comparative methods in visual perception, achieving  $F$ -score metrics of 0.807 7 and 0.728 9, respectively, both higher than those of other comparative methods.

**Key words:** ancient mural disease detection; U-Net; residual dual channel; multi-scale attention gate; mixed domain attention

收稿日期: 2023-07-12; 修回日期: 2024-04-11. 基金项目: 国家自然科学基金(62166048, 61263048); 云南省应用基础研究计划(2018FB102); 云南大学研究生实践创新项目(ZC-22222775); 云南大学中青年骨干教师培养计划(XT412003). 赵辉荣(1996—), 女, 硕士研究生, 主要研究方向为图像处理、计算机视觉、深度学习; 余映(1977—), 男, 博士, 副教授, 硕士生导师, 论文通信作者, 主要研究方向为图像与视觉、人工神经网络; 陈安(1999—), 男, 硕士研究生, 主要研究方向为图像处理、深度学习; 倪雪莹(1997—), 女, 硕士研究生, 主要研究方向为计算机视觉、神经网络; 王信超(1998—), 男, 硕士研究生, 主要研究方向为图像分割、深度学习.

中国古代壁画是珍贵的人类文化遗产,记录了大量各历史时期各民族的社会、艺术、文化、宗教活动等重要信息,对人类历史文化的研究起到了不可或缺的作用。然而,这些古代壁画艺术瑰宝长期受到人类活动及自然环境的影响,大多数壁画产生了起甲、脱落、霉菌、褪色、裂纹等病害类型,这些病害不但会加速壁画的消亡,还会影响其研究和观赏价值。在古代壁画数字化修复研究中,对壁画病害区域的标定是一个关键的技术环节,目前大多依靠人工方式标定壁画中的破损和病害区域,极其耗时耗力。利用现代计算机技术对壁画病害区域进行自动检测,可以极大地减少壁画修复人员的工作量。

近年来,国内外研究者对古代壁画病害的自动检测取得了初步的进展。Jaidilert 等<sup>[1]</sup>运用区域生长算法和形态学操作对傣族壁画进行半自动裂纹检测;张子迎等<sup>[2]</sup>通过局域最优层聚类方法提取壁画病害信息掩模;Cao 等<sup>[3]</sup>采用融合阈值分割的区域生长算法实现寺观壁画脱落病害的分割,可以提供很好的边界信息,但对于噪声和灰度不均的问题容易产生空洞和过分割;张豪远等<sup>[4]</sup>使用滚动引导滤波器进行滤波,采用 K-means 算法进行图像聚类并扩展到整幅图像的分割,其适用于病害区域明显的壁画;Gao 等<sup>[5]</sup>利用云边缘计算和数字化技术检测边缘连续和封闭的壁画裂缝病害。随着计算视觉和深度学习的发展,人工神经网络的方法被用于古代壁画病害检测。Lin 等<sup>[6]</sup>利用 BP 神经网络将壁画图像分为破损区域和正常区域;张叶娥等<sup>[7]</sup>通过改进 YOLOv5 网络检测壁画中的裂缝区域和脱落区域,并对病害区域进行分类;为了更准确地定位壁画病害区域,Wu 等<sup>[8]</sup>在 YOLOv5 网络中引入通道关注模块,调整不同特征通道的重要性。古代壁画病害检测的骨干网络广泛采用 U-Net 结构<sup>[9]</sup>,以获得病害的大小、位置,并逐像素地判断其是否属于病害。由于原始的 U-Net 不能完全满足病害检测任务的需求,因此需要在网络结构上进行改进。Wu 等<sup>[10]</sup>利用 ConvNet<sup>[11]</sup>作为 U-Net 的骨干,自上而下地提取病害特征。为了提升壁画图像脱落区域边缘的提取效果,吕书强等<sup>[12]</sup>在 U-Net 的编码部分引入空间金字塔池化层,在解码部分采用池化层索引上采样;Yu 等<sup>[13]</sup>将壁画的浅层特征输入 U-Net 编码器的不同阶段,以获取中国古寺壁画中油漆脱落病害的细节信息。壁画病害区域的检测可以视为一个图像分割问题。Wu 等<sup>[10]</sup>将 U-Net<sup>[9]</sup>,UNet++<sup>[14]</sup>,UCTransNet<sup>[15]</sup>等通用语义分割模型用作壁画病害检测的对比模型,验证所提

出的 TMCrack-Net 模型的性能。尽管上述基于 U-Net 的古代壁画病害检测方法取得了较好的结果,但它们都无法从原始图像中获取更为深入的高级语义特征,也不能有效地利用 U-Net 结构中的跳跃连接捕获病害图像在编码器中的低级特征,且存在计算复杂度高的问题。

针对目前古代壁画病害检测方法难以准确地检测不同尺度病害区域的问题,本文提出一种基于残差双通道注意力 U-Net(residual dual channel attention U-Net, RDCAU-Net)的壁画病害检测模型,适用于检测古代壁画中的多尺度病害区域。首先针对 U-Net 的编码器和解码器设计残差双通道(residual dual channel, RDC)模块,能够有效地提取图像空间特征和扩大感受野,显著地提升壁画病害检测的精度;为了高效地融合高层和低层的互补特征,构建多尺度注意力门(multi-scale attention gate, MSAG)模块,将其加入到 U-Net 的跳跃连接中,能够根据图像特征的重要程度对各个特征空间进行自适应加权;为了使网络有效地捕获跨通道交互信息并关注与目标相关像素信息等特征,设计了一种混合域注意力(mixed domain attention, MDA)模块,能够明显地提高壁画病害检测的精度;此外,通过计算解码器不同层次预测图的损失,对特征进行全面监督提高模型性能,以获得高质量的病害区域和清晰的边界。

## 1 RDCAU-Net 模型

古代壁画具有复杂的背景结构,为了充分地提取病害特征,有效地融合低层和高层语义信息,本文提出 RDCAU-Net,其结构如图 1 所示,包括 RDC 模块,MSAG 模块和 MDA 模块。输入图像经过编码器 RDC 模块充分地提取空间特征,同时还能有效地扩大网络感受野。使用 RDC 模块代替 U-Net 中的编码器和解码器,可以显著提升壁画复杂背景中病害检测的精度;MSAG 模块位于编码器与解码器之间的跳跃连接中,能够高效地融合编码器与解码器的特征,突出图像显著特征;轻量级的 MDA 模块由通道注意力和空间注意力(spatial attention, SA)级联构成,该模块位于解码器之后,既能使网络有效地捕获跨通道交互的信息,又能使网络更加关注与目标相关的像素信息等特征,明显地提高复杂图像中病害检测精度。每个编码器模块通道数分别减少到 32, 64, 128, 256 和 512 个,可以减少后续操作所需的模型参数量。编码器

使用  $2 \times 2$  的最大池化进行下采样, 解码器使用  $2 \times 2$  的转置卷积进行上采样, 以便 MSAG 模块融合编码器和解码器特征. 解码器中, 前 4 个阶段的输出特

征通过双线性插值上采样的方式生成 4 个辅助预测图, 最后一个阶段通过一个  $1 \times 1$  卷积和 Sigmoid 函数生成最终的预测图.

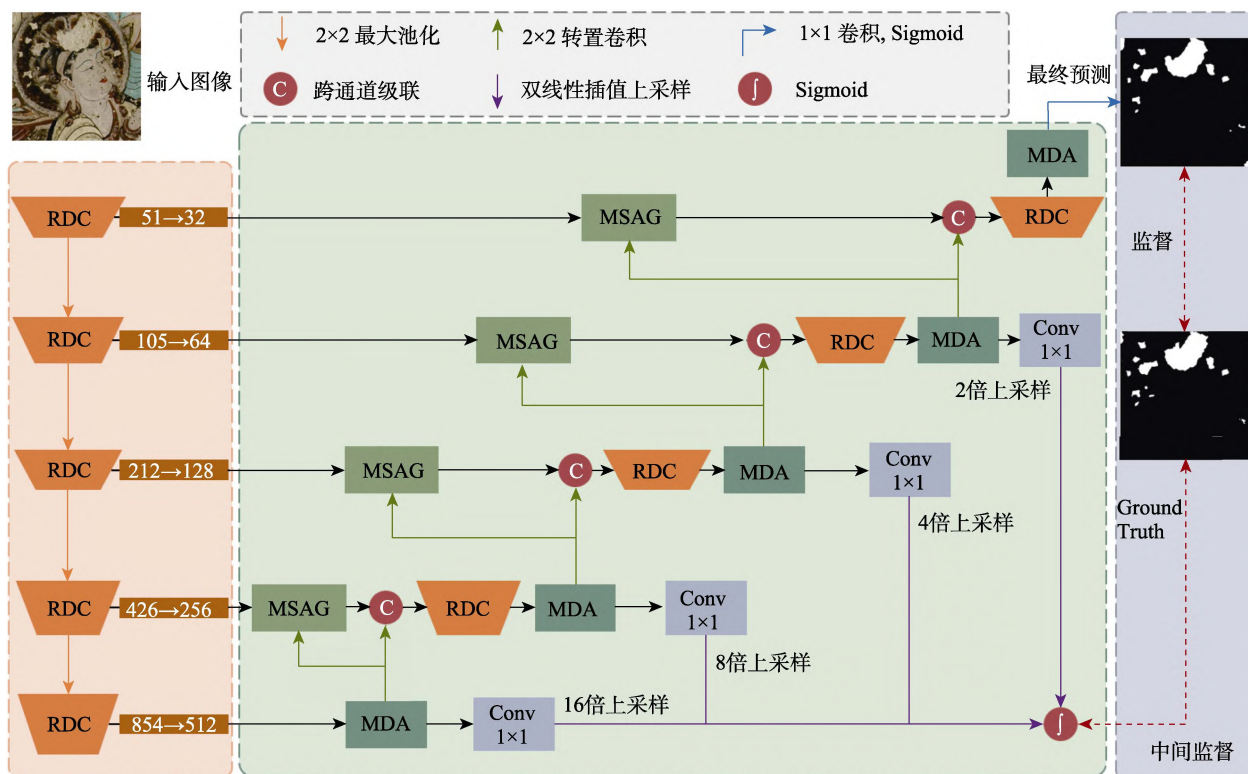


图 1 RDCAU-Net 的结构

### 1.1 RDC 模块

与自然图像相比, 古代壁画背景信息复杂, 同一幅壁画中可能存在不同大小的病害区域, 因此骨干网络需要具备学习不同尺度的图像特征能力. 在检测任务中, Szegedy 等<sup>[16]</sup>提出 Inception 模块, 利用  $3 \times 3$ ,  $5 \times 5$  和  $7 \times 7$  卷积并行运算, 使 U-Net 结构具有多分辨率分析能力. Ioffe 等<sup>[17]</sup>指出, 当 2 个  $3 \times 3$  卷积级联时, 其感受野范围相当于一个  $5 \times 5$  卷积; 当 3 个  $3 \times 3$  卷积级联时, 其感受野范围相当于一个  $7 \times 7$  卷积. 该策略有助于减少计算量和内存需求. 受到文献[16-17]工作的启发, 本文在编码器和解码器中, 利用一系列较小的  $3 \times 3$  卷积代替较大的  $5 \times 5$  和  $7 \times 7$  卷积; 然后将每个卷积层输出的不同尺度特征进行跨通道融合, 该过程被称为多尺度特征融合 (multi-scale feature fusion, MSFF). MSFF 模块可以减少参数量和计算时间, 扩大模型的感受野, 提升对多尺度目标的检测准确率; 此外, 基于 ResNet<sup>[18]</sup>的残差思想, 添加  $1 \times 1$  的残差连接卷积层构成多尺度残差 (multi-scale residuals, MSR) 模块, 该模块可以学习到一些额外的空间特征, 其结构如图 2 所示.

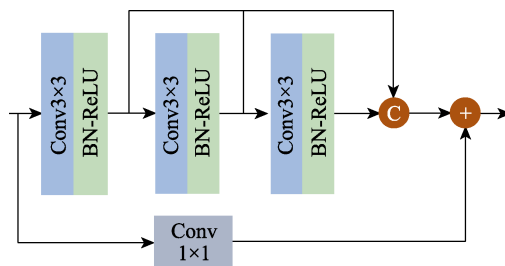


图 2 MSR 模块

针对病害区域与背景相似问题, 本文通过增加一个相同的 MSFF 模块克服空间特征不足的问题, 并将改造后的 MSR 模块称为 RDC 模块, 其结构如图 3 所示. 采用 RDC 模块作为网络的编码器和解码器, 输出特征中不仅包含多种不同尺度的感受野, 网络还提取了更多的空间特征, 有利于网络融合输入特征中不同尺度的空间信息, 提高检测模型对复杂背景病害特征的提取与表达能力.

RDC 模块特征提取步骤如下.

Step1. 将输入特征  $x_{in}$  连续输入 3 个相同的  $3 \times 3$  卷积层, 以扩大感受野范围. 卷积输出结果为

$$\{r_1, r_2, r_3\} = \{F(x_{in}), F(r_1), F(r_2)\}.$$

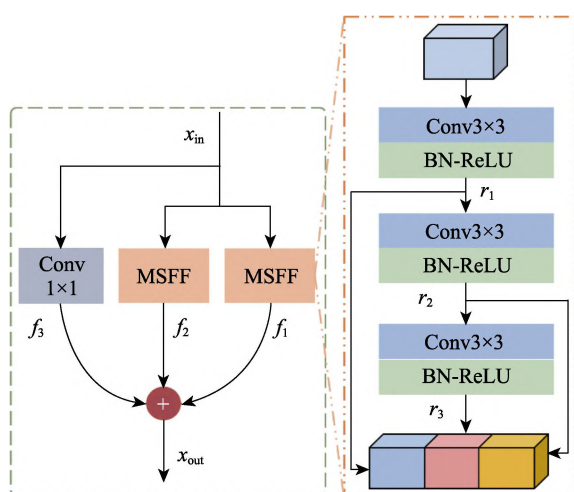


图 3 RDC 模块

其中,  $F(\cdot) = \text{ReLU}(\text{BN}(\text{Conv}_{3 \times 3}(\cdot)))$ ,  $\text{Conv}_{3 \times 3}(\cdot)$  表示输入特征通过大小为  $3 \times 3$  的卷积层,  $\text{BN}(\cdot)$  表示批量归一化,  $\text{ReLU}(\cdot)$  表示激活函数; 为了融合各个尺度的特征, 将  $r_i (i=1,2,3)$  沿着通道维度进行组合, 该操作进行 2 次, 计算公式为  $\{f_1, f_2\} = \text{Cat}(r_1, r_2, r_3)$ . 其中,  $\text{Cat}(\cdot)$  表示通道维度上的级联, 特征  $f_1$  和  $f_2$  包含多种尺度的感受野, 有利于 RDC 融合不同的空间信息, 提高检测模型提取病害图像多尺度特征的能力。

Step2. 将输入特征  $x_{in}$  通过  $1 \times 1$  卷积核改变通道数, 其结果被记为  $f_3$ . 为了改善模型在训练时梯度消失或爆炸的问题, 利用相加融合策略将  $f_1$ ,  $f_2$  和  $f_3$  的特征图相加, 获得 RDC 模块的最终输出特征  $x_{out} = f_1 + f_2 + f_3$ .

## 1.2 MSAG 模块

为了高效地融合高层和低层的互补特征, 逐步抑制背景区域的干扰, 根据网格注意力<sup>[19]</sup>构建 MSAG 模块, 其结构如图 4 所示。

解码器使用  $2 \times 2$  的转置卷积对特征图进行上采样, 得到特征  $y$ . 在 MSAG 模块中, 编码器输出特征  $x$  和解码器上采样特征  $y$  首先被分别输入到卷积层以调整通道数. 然后对输出的低层和高层互补特征进行融合. 最后使用 ReLU 函数获得多尺度门控系数, 该过程可表示为

$$Q = \text{ReLU}(\text{BN}(\text{Conv}_{1 \times 1}(x)) + \text{BN}(\text{Conv}_{1 \times 1}(y)))$$

其中,  $\text{Conv}_{1 \times 1}(\cdot)$  表示  $1 \times 1$  的卷积运算; 得到  $Q$  后, 利用  $1 \times 1$  的卷积实现通道间的信息交互, 再通过 Sigmoid 函数  $\sigma(\cdot)$  编码非线性表达, 获得 SA 权重, 计算公式为  $w = \sigma(\text{BN}(\text{Conv}_{1 \times 1}(Q)))$ . 最终, 将权重  $w$  加权到对应的低层特征  $x$  上, 得到显著特征  $\hat{x} = w \otimes x$ ,  $\otimes$  表示 Hadamard 乘积。

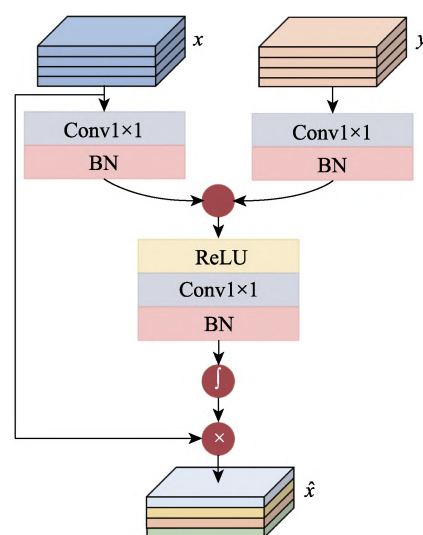


图 4 MSAG 模块

## 1.3 MDA 模块

近年来, 注意力已广泛应用于计算机视觉任务<sup>[20]</sup>. Hu 等<sup>[21]</sup>提出了一种通道注意力模块, 并取得了良好的性能; Gather-excite 使用深度卷积融合特征, 探索 SA<sup>[22]</sup>; DAN 同时考虑了通道注意力和 SA<sup>[23]</sup>; CBAM<sup>[24]</sup>和空间与通道联合注意力(spatial and channel SE, scSE)模块<sup>[25]</sup>使用  $k \times k$  的二维卷积计算 SA, 然后将其与通道关注相结合. 然而, 上述方法都侧重于开发复杂的注意力模块以获得更好的性能. 高效通道注意力(efficient channel attention, ECA)<sup>[26]</sup>模块使用少量参数就能显著地提升模型性能, 其本质是在重点关注的通道上设置高权重, 实现对更重要特征的有效关注; 但该模块忽略了每个通道内的局部信息以及对空间上的重要位置的关注. 考虑目标像素与非目标像素的不同, 本文采用 ECA<sup>[26]</sup>和 SA 进行级联构成 MDA 模块, 可以显著地降低模型的复杂性, 让网络在有效地捕获跨通道交互的信息的同时, 更加关注目标像素信息等特征, 进一步定位图像的目标区域. MDA 模块的结构如图 5 所示。

ECA 用于确定哪些通道需要重点关注, 其结构如图 6 所示. 设一个卷积块输出特征为  $x \in \mathbb{R}^{H \times W \times C}$ , 其中,  $W$ ,  $H$  和  $C$  分别为宽度、高度和通道数. MDA 模块中, 特征  $x$  首先被送入全局平均池化层中, 提取各通道的全局空间特征  $x_1 \in \mathbb{R}^C$ ; 再利用一维卷积实现局部跨通道交互范围, 产生新的全局空间特征  $x_2 = \text{Conv1D}_k(x_1)$ ,  $x_2 \in \mathbb{R}^C$ . 其中,  $\text{Conv1D}(\cdot)$  表示一维卷积; 卷积核的大小  $k$  根据通道维度  $C$  自适应选择, 即



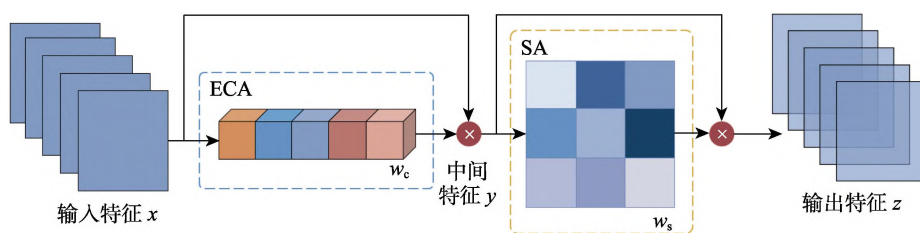


图 5 MDA 模块的整体结构

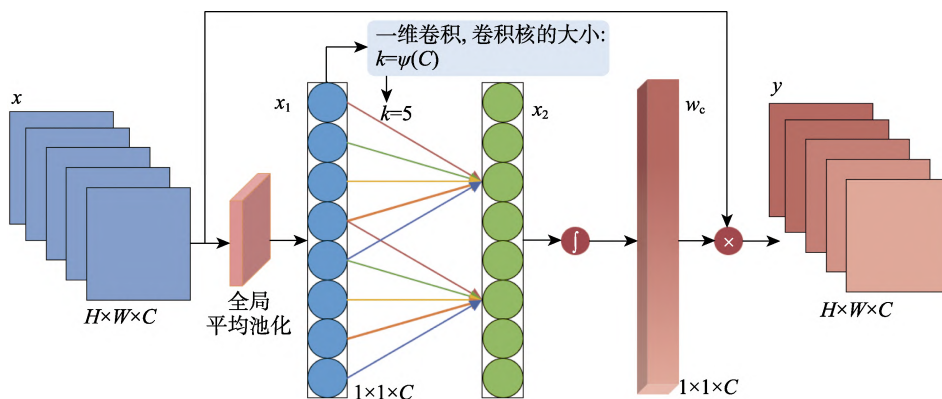


图 6 ECA 模块

$$k = \psi(C) = \left\lfloor \frac{\lfloor \log(C) \rfloor + b}{\gamma} + \frac{b}{\gamma_{\text{odd}}} \right\rfloor;$$

$|m|_{\text{odd}}$  表示离  $m$  最近的奇数; 本文实验中, 设  $\gamma = 2$ ,  $b = 1$ ; 通道维度  $C$  通过映射函数  $\psi$  后可知高维通道具有较大的交互范围, 低维通道具有较小的交互范围. 利用 Sigmoid 函数  $\sigma(\cdot)$  获得各通道的权重  $w_c = \sigma(x_2)$ , 再将该权重加权到相应的特征  $x$  上, 得到通道关注特征  $y = w_c \otimes x$ .

SA 用于决定在特征图中需要关注的位置, 其结

构如图 7 所示. 通道注意力输出特征  $y \in \mathbb{R}^{H \times W \times C}$ , 首先分别被输入全局平均池化层和全局最大池化层中, 提取特征  $y_1 \in \mathbb{R}^{H \times W \times 1}$  和  $y_2 \in \mathbb{R}^{H \times W \times 1}$ ; 然后将  $y_1$  和  $y_2$  特征进行通道拼接后, 通过  $7 \times 7$  的卷积(填充量为 3)运算增强空间上下文信息, 利用 Sigmoid 函数获得空间注意力权重  $w_s = \sigma(\text{Conv}_{7 \times 7}(\text{Cat}(y_1, y_2)))$ . 其中,  $\text{Conv}_{7 \times 7}(\cdot)$  表示  $7 \times 7$  的卷积运算; 最后将权重  $w_s$  和输入特征  $y$  执行逐元素的乘法运算, 得到空间关注特征  $z = w_s \otimes y$ .

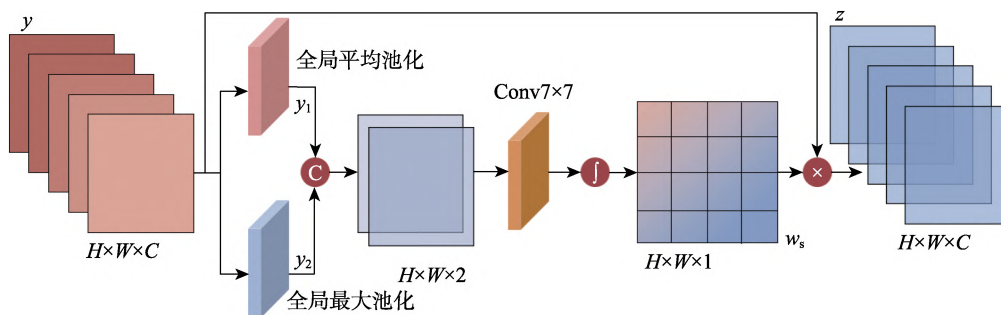


图 7 SA 模块

#### 1.4 损失函数

在分类和分割任务中广泛采用多种损失函数结合起来作为模型的训练损失, 其有效性已经得到验证. 使用与文献[27-28]中相同的定义, 将交并比(intersection over union, IoU)损失<sup>[29]</sup>  $L_{\text{IoU}}$  和二元交叉熵(binary cross entropy, BCE)损失<sup>[30]</sup>  $L_{\text{BCE}}$  相加

得到加权损失  $L_{\text{ADD}} = L_{\text{IoU}} + L_{\text{BCE}}$ .

为了获得高质量的病害区域和清晰的边界, 本文使用 4 个辅助预测分别计算加权损失, 然后将这些损失相加. 解码器最后一层生成最终的预测图, 采用一个  $L_{\text{BCE}}$  进一步对模型进行全面监督. 在 RDCAU-Net 模型中, 使用多阶段损失相加函数

$L = L_{\text{BCE}}(g, p) + \sum_{i=1}^4 L_{\text{ADD}}(g, p_i)$ . 其中,  $p$  为最终预测;  $g$  为 Ground Truth,  $p_i (i=1, 2, 3, 4)$  表示从不同层次的解码器中提取的辅助预测.

## 2 实验结果与分析

古代壁画因数据资源稀缺而受到保护,目前,古代壁画病害检测尚缺乏公共开放的数据集. 本文收集并筛选出 100 幅敦煌莫高窟壁画图像进行实验,为了增加样本数量和提高样本多样性,采用随机旋转、随机翻转、随机位置截取和随机加入高斯噪声的方式实现训练样本数据的扩增,最终形成 10 000 幅敦煌莫高窟壁画数据集. 实验中,选取其中的 8 594 幅壁画作为病害检测网络模型的训练集,剩下的 1 406 幅壁画作为验证集,以找到最优的超参数. 为了评估 RDCAU-Net 模型的性能,将独立于训练数据集的 248 幅敦煌莫高窟壁画作为测试集. 为了测试 RDCAU-Net 模型的泛化能力,本文收集了 832 幅云南各地区少数民族古代壁画数据,包括丽江纳西族地区白沙壁画、红河哈尼族彝族地区建水指林寺明代壁画、西双版纳地区的曼宰竜佛寺壁画和傣族佛寺壁画,这些壁画内容独特,风格迥异,特征多样,与敦煌莫高窟壁画在画风、地域和年代上具有较大的差异. 实验中,2 个壁画数据集均由壁画病害研究人员进行人工标注,标

注后可以得到对应壁画病害区域的 Ground Truth 图像;壁画图像的大小全部调整为  $512 \times 512$  像素. 本文实验的计算机硬件环境配置 Intel(R) @2.9 GHz CPU 和 NVIDIA GeForce RTX 2080 GPU;所有实验模型都在 PyTorch 1.11.0 中实现,采用 Adam 优化算法,参数为  $\beta_1 = 0.500$  和  $\beta_2 = 0.999$ ,初始学习率为 0.005,批量大小为 4.

### 2.1 定性评价

为了定性评价 RDCAU-Net 模型的性能,将其与 U-Net<sup>[13]</sup>以及相关壁画病害检测方法<sup>[1,3-4]</sup>进行对比. 通用语义分割模型 UNet++<sup>[14]</sup>和 UCTransNet<sup>[15]</sup>已被前人用作壁画病害检测的对比方法,因此本文也将它们进行对比;此外,本文还增加了通用语义分割模型 RAU-Net<sup>[20]</sup>, CMU-Net<sup>[31]</sup>, DC-UNet<sup>[32]</sup>,以及缺陷检测模型 SegDec-Net<sup>[33]</sup>进行对比. 选择上述深度神经网络模型作为对比方法的原因如下: (1) 这些通用语义分割模型和缺陷检测模型在壁画病害检测中均能获得良好的性能; (2) RDCAU-Net 模型和通用语义分割模型均采用 U-Net 结构; (3) 选择具有开源实现的模型作为对比实验. 所有对比模型都在本文数据集上重新训练.

图 8 所示为 RDCAU-Net 与 3 种传统方法对 4 幅敦煌莫高窟壁画图像的病害区域检测结果,图 9 所示为 RDCAU-Net 与 3 种传统方法对 4 幅云南少数民族壁画图像的病害区域检测结果. 可以看出,文献[1]能检测到部分病害区域,但会将背景区域

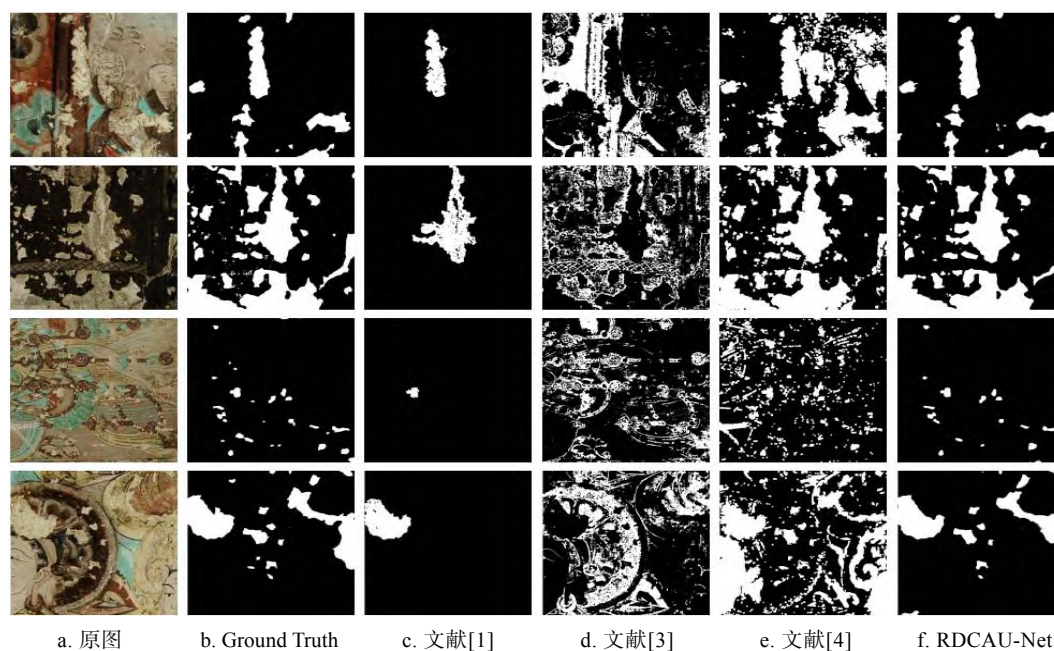


图 8 4 种方法对 4 幅敦煌莫高窟壁画图像的病害区域检测结果



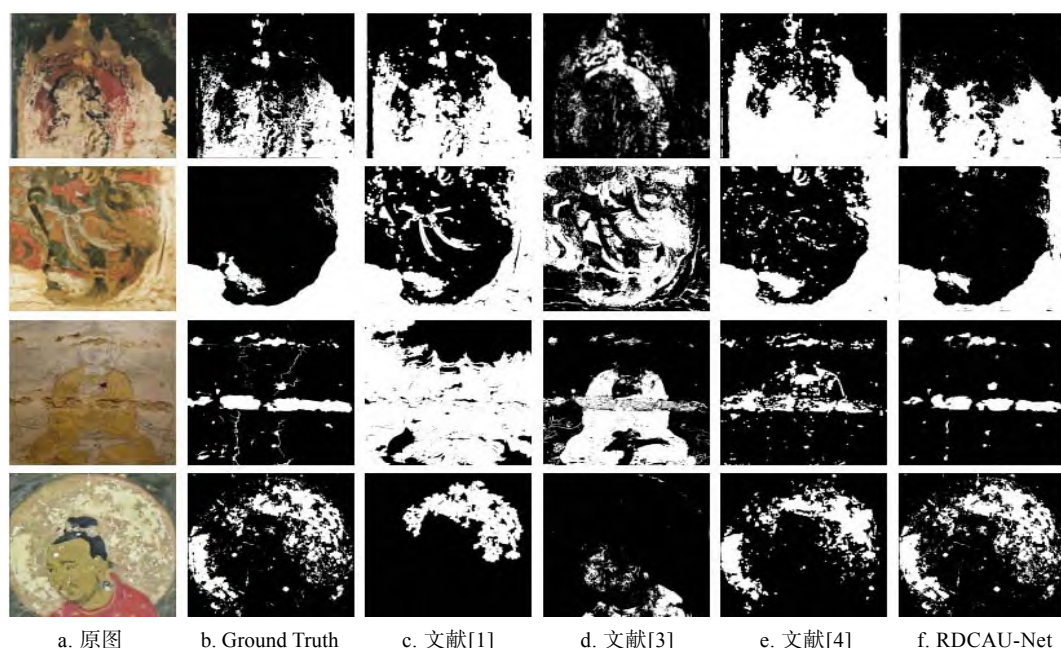


图 9 4 种方法对 4 幅云南少数民族壁画图像的病害区域检测结果

检测为病害区域,且需要手动选取种子点,使得在检测多幅壁画图像的病害区域时工作量加大;文献[3]在检测病害区域方面效果不佳,且需要对 3 个颜色通道进行区域生长,在处理病害严重的区域时耗时较长,计算速度慢;文献[4]虽然能大致定位病害位置,但同样面临误检问题,并且需要手动选择初始区域,增加了检测的复杂性;与文献[1]相比, RDCAU-Net 模型能更加完整地检测出病害区域;与文献[4]相比, RDCAU-Net 模型几乎没有将背景区域误检为病害区域。

图 10 所示为 8 种模型对 7 幅敦煌莫高窟壁画图像的病害区域检测结果。可以看出,第 1 幅图像的危害区域形状简单且明显,几乎所有模型都能较好地检测到不同尺度的病害区域;第 2 幅图像中, RDCAU-Net 模型采用 MSAG 和 MDA 模块有效地区分了病害区域与背景区域,而其他模型容易将与病害颜色相似的背景区域误检为病害区域;第 3 幅和第 4 幅图像存在病害区域模糊以及病害与背景极其相似的情况,与其他模型相比, RDCAU-Net 模型通过采用 RDC 模块进行复杂特征分析,更能准确地检测到病害区域;第 5 幅图像中存在大尺度的病害区域,7 种对比模型均不能较为完整地检测出大尺度的病害区域,特别是 UNet++<sup>[14]</sup>几乎检测不到, CMU-Net<sup>[31]</sup>和 RAU-Net<sup>[20]</sup>的检测结果存在不连续的情况;相比之下, RDCAU-Net 模型不仅能够检测出大尺度的病害区域,而且能获得病害区域清晰的边界;第 6 幅图像中不存在病害区域,只有

UNet++<sup>[14]</sup>, DC-UNet<sup>[32]</sup>和 RDCAU-Net 模型没有误检;第 7 幅图像中存在大尺度的病害区域,且病害区域的颜色与背景相似,进一步增加了检测难度,虽然 RDCAU-Net 模型会漏检一部分病害区域,但 UNet++<sup>[14]</sup>和 DC-UNet<sup>[32]</sup>模型几乎检测不到,另外的 5 种对比模型错误地将背景检测为病害区域。实验结果表明,与其他模型相比, RDCAU-Net 模型通过采用 MSAG, MDA 和 RDC 模块,在准确识别各种病害区域,尤其是在大尺度和颜色与背景相似的病害区域检测方面,检测结果更接近 Ground Truth。

图 11 所示为 8 种模型对 7 幅云南少数民族壁画图像的病害区域检测结果。可以看出: (1) 第 1 幅和第 2 幅图像出现地仗层脱落的情况,8 种模型基本能够检测出壁画中的病害区域,检测结果互有优劣。(2) 第 3 幅和第 4 幅图像存在不同尺度的地仗层和颜料层脱落情况。只有 RDCAU-Net 模型能够较好地检测出第 3 幅图像中大尺度的病害区域;第 4 幅图像中的病害区域已经延伸到图像边界, RDCAU-Net 模型可以检测到大部分病害区域,而 7 种对比模型对病害区域的检测效果明显更差。(3) 第 5 幅和第 6 幅壁画图像破损严重,出现了起甲、颜料层和地仗层脱落等情况。 UNet++<sup>[14]</sup>, CMU-Net<sup>[31]</sup>, UCTransNet<sup>[15]</sup>和 RDCAU-Net 模型能检测到第 5 幅图像中大面积的脱落区域,但检测结果不连续, U-Net<sup>[13]</sup>, UCTransNet<sup>[15]</sup>, SegDec-Net<sup>[33]</sup>和 RDCAU-Net 模型可以检测到部分起甲区域;对于第 6 幅图像, CMU-Net<sup>[31]</sup>和 RDCAU-Net 模型可以

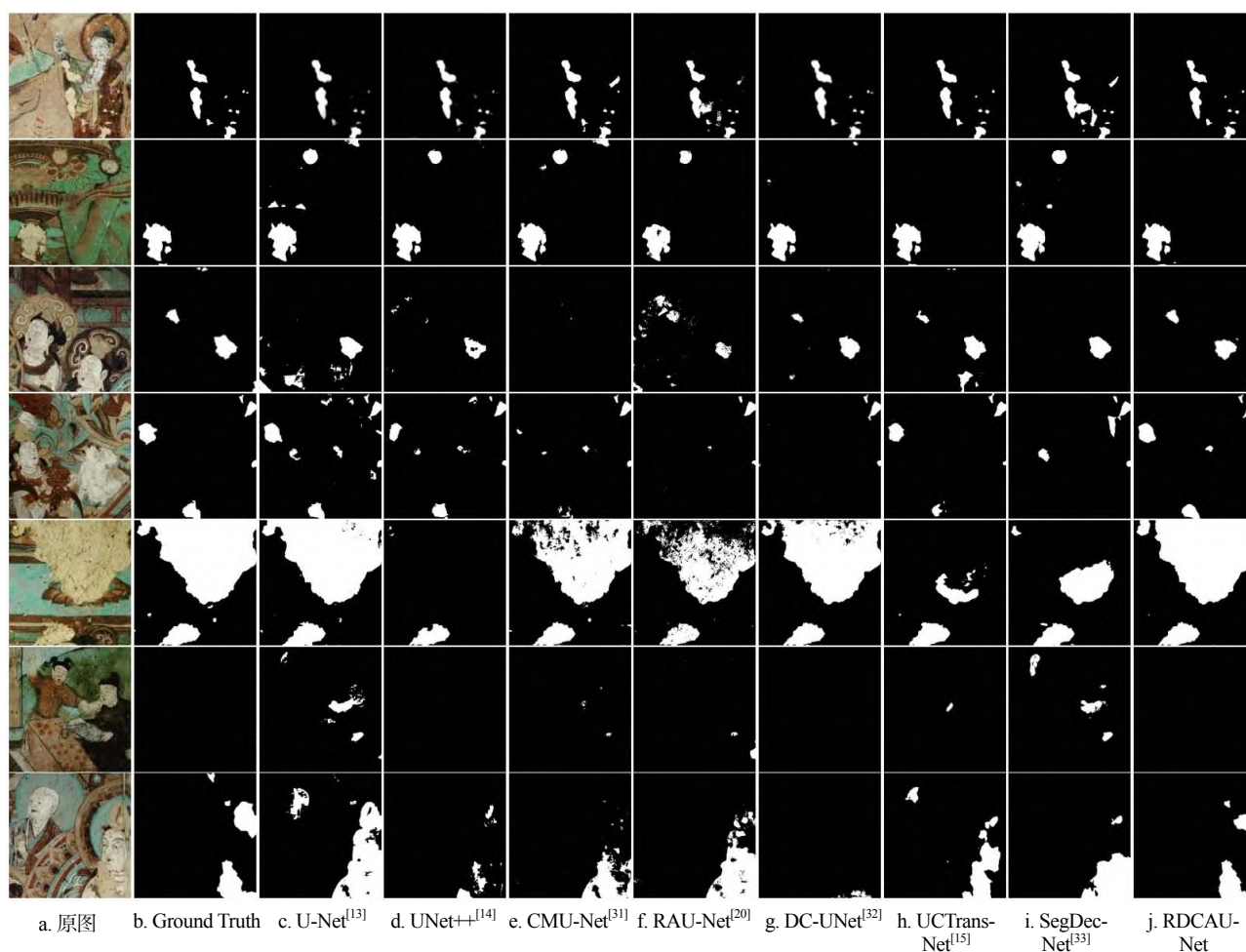


图 10 8 种模型对 7 幅敦煌莫高窟壁画图像の病害区域检测结果

检测到部分脱落区域。(4) 第 7 幅图像出现了许多小面积的颜料层脱落, RDCAU-Net 模型能检测到这些区域,但同时也会将背景区域误检为病害区域,原因可能在于背景与病害区域之间的高度相似性。实验结果表明,与其他模型相比, RDCAU-Net 模型在检测不同类型和不同尺度的病害区域(如地仗层脱落、颜料层脱落、起甲等)时,总体上具有更优的检测能力。

## 2.2 定量评价

采用精确率  $P$ , 召回率  $R$ ,  $F$ -score 值  $F_1$  和平均绝对误差(mean absolute error, MAE)这 4 种客观评价指标对 RDCAU-Net 模型的性能进行定量评价。各评价指标的计算公式分别为

$$P = \frac{TP}{TP + FP},$$

$$R = \frac{TP}{TP + FN},$$

$$F_1 = \frac{2 \times P \times R}{P + R},$$

$$MAE = \frac{1}{m} |\hat{y}_i - y_i|.$$

其中, 真阳性(true positive, TP)表示实际为正被预测为正的样本数量; 假阳性(false positive, FP)表示实际为负但被预测为正的样本的数量; 真阴性(true negative, TN)表示实际为负被预测为负的样本的数量; 假阴性(false negative, FN)表示实际为正而被预测为负的样本数量;  $P$  衡量正确预测为正的样本占全部预测为正的样本的比例;  $R$  衡量真正正确预测为正的样本占全部实际为正的样本的比例;  $F_1$  衡量同时考虑  $P$  和  $R$  的一个平衡值, 当两者同时达到最大时取最大值;  $\hat{y}$  和  $y$  分别表示预测值和观测值; MAE 可以反映预测值误差的实际情况。  $P$ ,  $R$  和  $F_1$  的数值越大, MAE 的数值越小, 表示检测效果越好; 反之, 效果越差。

表 1 和表 2 所示分别为 8 种模型在敦煌莫高窟壁画数据集和云南少数民族壁画数据集上的定量评价指标。可以看出, RDCAU-Net 模型在 2 个壁画数据集上的 4 个评价指标均明显优于 7 种对比模



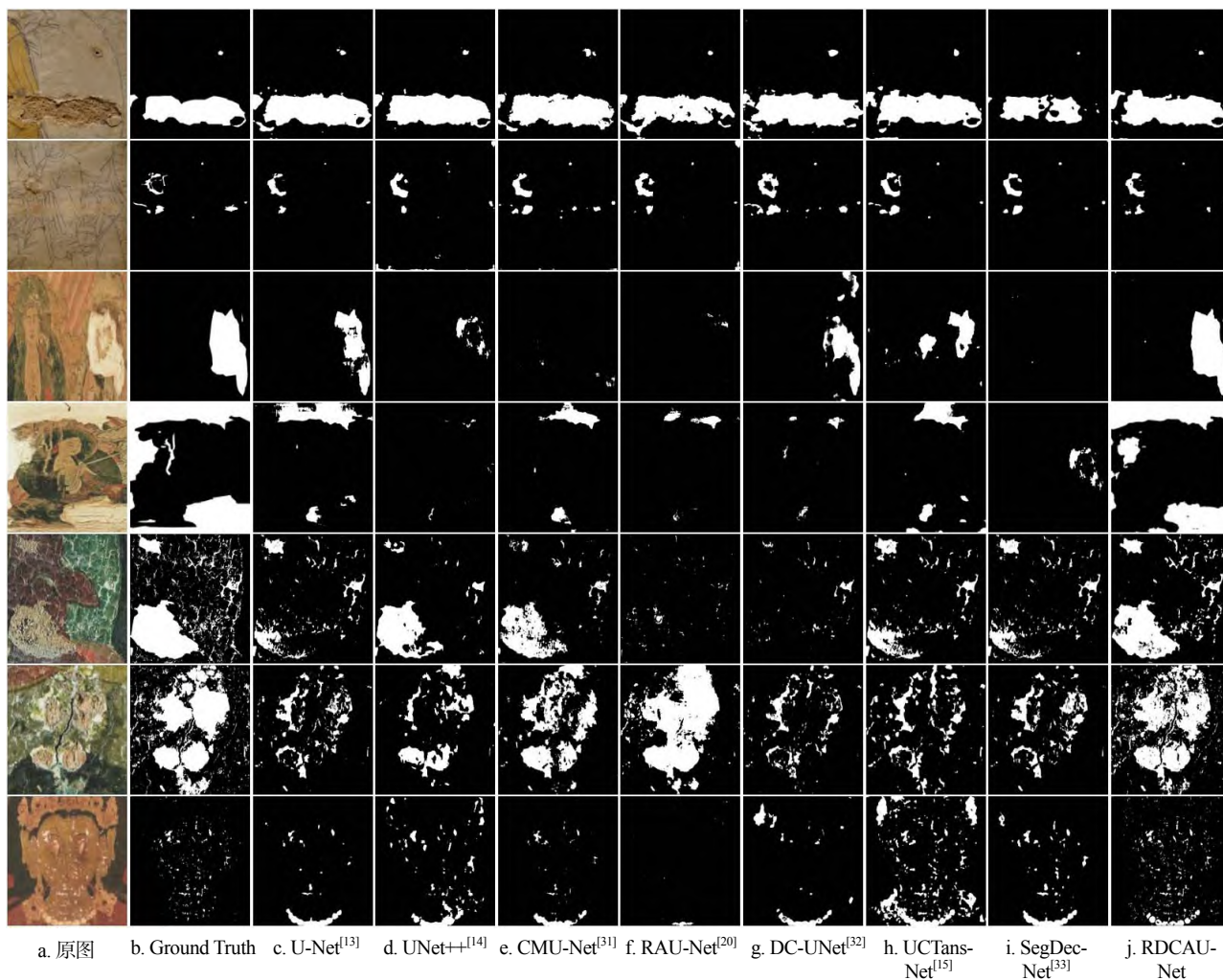


图 11 8 种模型对 7 幅云南少数民族壁画图像的病害区域检测结果

表 1 8 种模型在敦煌莫高窟壁画数据集上的对比

模型	$P \uparrow$	$R \uparrow$	$F_1 \uparrow$	$MAE \downarrow$
U-Net <sup>[13]</sup>	0.672 1	0.719 6	0.695 0	0.040 1
UNet++ <sup>[14]</sup>	0.534 9	0.760 4	0.628 0	0.040 5
CMU-Net <sup>[31]</sup>	0.732 6	0.601 1	0.660 3	0.060 5
RAU-Net <sup>[20]</sup>	0.580 2	0.652 4	0.614 1	0.066 6
DC-UNet <sup>[32]</sup>	0.670 0	0.799 8	0.729 2	0.030 2
UCTansNet <sup>[15]</sup>	0.686 6	0.693 0	0.689 8	0.051 4
SegDec-Net <sup>[33]</sup>	0.563 6	0.628 6	0.594 4	0.069 1
RDCAU-Net	<b>0.787 9</b>	<b>0.828 6</b>	<b>0.807 7</b>	<b>0.021 7</b>

注. 粗体表示各指标的最佳值.

表 2 8 种模型在云南少数民族壁画数据集上的对比

模型	$P \uparrow$	$R \uparrow$	$F_1 \uparrow$	$MAE \downarrow$
U-Net <sup>[13]</sup>	0.597 9	0.666 7	0.630 4	0.086 6
UNet++ <sup>[14]</sup>	0.641 3	0.629 8	0.635 5	0.101 4
CMU-Net <sup>[31]</sup>	0.663 0	0.670 8	0.666 9	0.095 0
RAU-Net <sup>[20]</sup>	0.668 8	0.534 8	0.594 1	0.158 4
DC-UNet <sup>[32]</sup>	0.584 3	0.721 0	0.645 5	0.077 5
UCTansNet <sup>[15]</sup>	0.678 7	0.556 2	0.611 4	0.109 8
SegDec-Net <sup>[33]</sup>	0.634 9	0.628 9	0.631 9	0.104 3
RDCAU-Net	<b>0.704 6</b>	<b>0.755 0</b>	<b>0.728 9</b>	<b>0.056 2</b>

注. 粗体表示各指标的最佳值.

型, 说明该模型对各类壁画病害区域具有更高的检测准确性.

表 3 和表 4 所示分别为 RDCAU-Net 与 3 种传统方法在敦煌莫高窟壁画数据集和云南少数民族

表 3 在敦煌莫高窟壁画数据集上与传统方法的对比

方法	$P \uparrow$	$R \uparrow$	$F_1 \uparrow$	$MAE \downarrow$
文献[1]	0.769 5	0.301 8	0.443 6	0.089 6
文献[3]	0.021 7	0.089 6	0.034 9	0.435 6
文献[4]	0.410 7	0.773 7	0.536 6	0.168 5
RDCAU-Net	<b>0.787 9</b>	<b>0.828 6</b>	<b>0.807 7</b>	<b>0.021 7</b>

注. 粗体表示各指标的最佳值.

表 4 在云南少数民族壁画数据集上与传统方法的对比

方法	$P \uparrow$	$R \uparrow$	$F_1 \uparrow$	$MAE \downarrow$
文献[1]	0.626 4	0.699 5	0.660 9	0.212 3
文献[3]	0.078 1	0.129 5	0.097 4	0.377 5
文献[4]	0.650 6	0.681 1	0.665 5	0.118 8
RDCAU-Net	<b>0.704 6</b>	<b>0.755 0</b>	<b>0.728 9</b>	<b>0.056 2</b>

注. 粗体表示各指标的最佳值.

壁画数据集上的定量评价指标对比。从表3可以看出, RDCAU-Net 在敦煌莫高窟壁画数据集上的  $F_1$  指标显著优于3种传统方法, 说明该模型在处理具有复杂背景的壁画图像时具有更好的病害检测效果。从表4可以看出, RDCAU-Net 在云南少数民族壁画数据集上的4个指标均优于传统方法, 说明该模型在未使用该数据集进行训练的情况下也能较为准确地预测病害区域。

### 2.3 模型复杂度

通常, 采用模型的计算量和参数量来评估网络模型复杂度。模型的计算量采用浮点乘法数量和加法数量 FLOPs 衡量, 反映模型训练和预测的时间复杂度; 模型参数量 Params 衡量使用的模块中涉及的参数数量, 其值越大, 则训练模型所需的数据量也越大。对于网络模型性能来说, 这2个指标越小越好。

表5所示为8种模型的 FLOPs 和 Params 的结果对比。可以看出, RDCAU-Net 模型的 FLOPs 比其他6种模型明显要小, 仅次于基础的 U-Net<sup>[13]</sup>; 此外, RDCAU-Net 模型的 Params 仅为 9.4M, 在所有对比模型中是最小的。因为 RDCAU-Net 模型使用轻量级通道信息交互, 采用多个  $1 \times 1$  卷积运算,

表5 8种模型的 FLOPs 和 Params 结果对比

模型	FLOPs	Params
U-Net <sup>[13]</sup>	<b>48.4G</b>	38.6M
UNet++ <sup>[14]</sup>	541.4G	36.6M
CMU-Net <sup>[31]</sup>	356.1G	49.9M
RAU-Net <sup>[20]</sup>	601.2G	39.4M
DC-UNet <sup>[32]</sup>	106.3G	11.6M
UCTansNet <sup>[15]</sup>	160.5G	66.2M
SegDec-Net <sup>[33]</sup>	91.4G	15.4M
RDCAU-Net	67.2G	<b>9.4M</b>

注: 粗体表示各指标的最佳值。

所以其在计算效率方面明显优于其他6种对比模型。

### 2.4 消融实验

为了验证 RDCAU-Net 模型中 RDC, MSAG 和 MDA 这3个模块的有效性, 采用 U-Net<sup>[13]</sup>作为基础模型(简称为 Baseline), 并通过在 Baseline 上添加各种模块结构作为对照模型。为了公平比较, 所有消融模型均以相同的设置进行训练。

图12和图13所示分别为 RDCAU-Net 模型与各种消融模型在敦煌莫高窟壁画数据集和云南少数民族壁画数据集上的病害区域检测结果。可以看出, 对比图12c和图12d, 以及图13c和图13d,

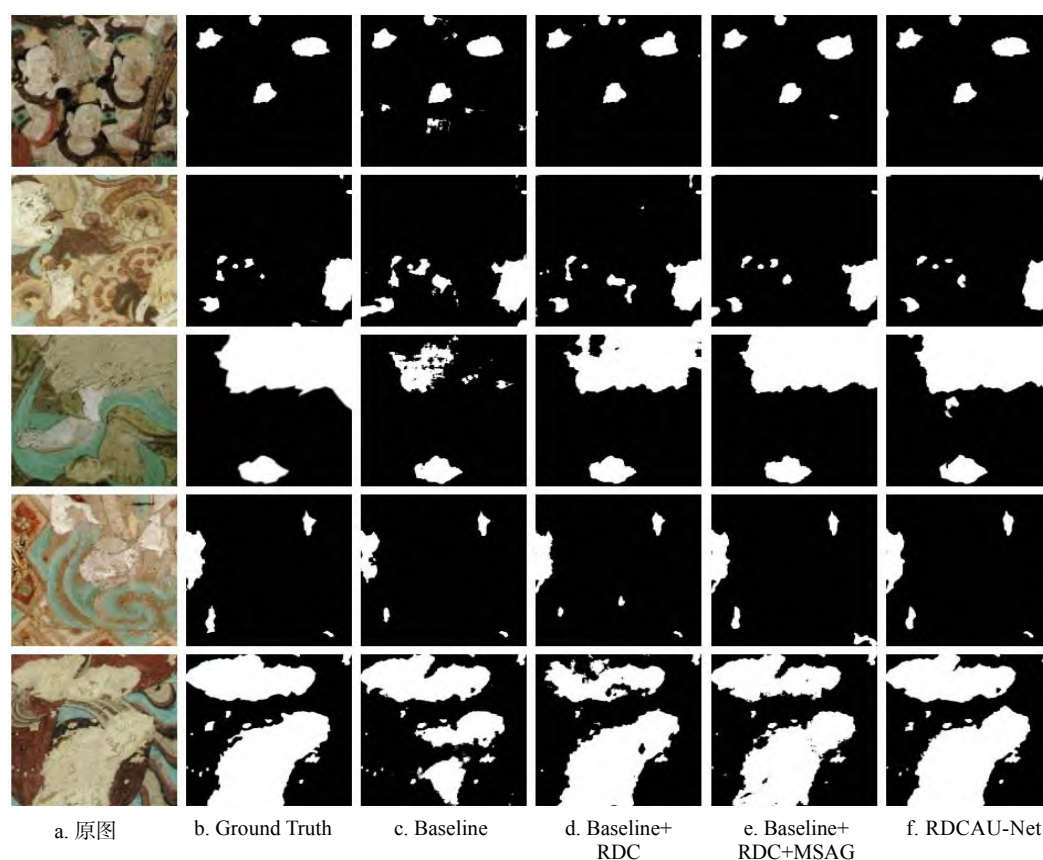


图12 各种消融模型在敦煌莫高窟壁画数据集上的病害区域检测结果

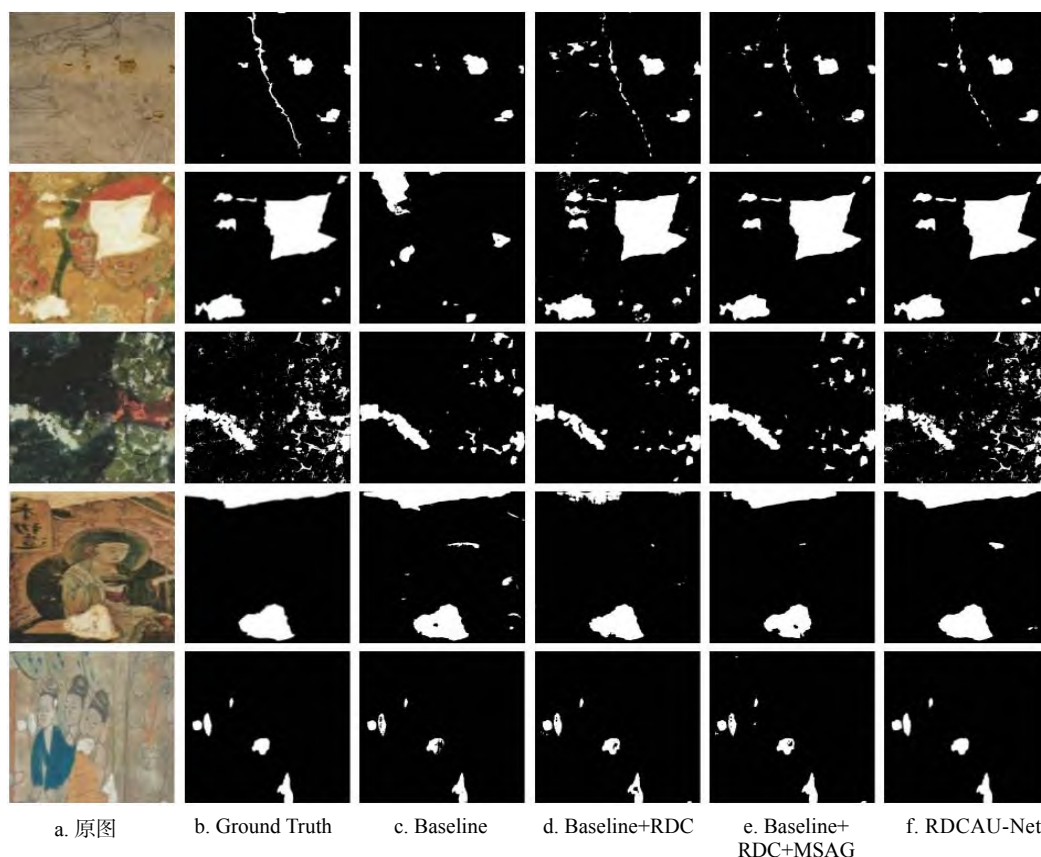


图 13 各种消融模型在云南少数民族壁画数据集上的病害区域检测结果

将 Baseline 模型中的编码器和解码器全部替换为 RDC 模块(简称为 Baseline+RDC)后,能较为完整地检测出不同尺度的病害区域,该模块能检测出图 13 的第 1 幅和第 2 幅图像中出现的裂缝;对比图 12d 和图 12e,以及图 13d 和图 13e,在 Baseline+RDC 的基础上加入 MSAG 模块(简称为 Baseline+RDC+MSAG)后,能抑制背景信息的干扰,并成功地检测出大尺度病害区域,进一步提高检测的准确性;对比图 12e 和图 12f,以及图 13e 和图 13f,在 Baseline+RDC+MSAG 的基础上加入 MDA 模块(简称为 Baseline+RDC+MSAG+MDA),即构成 RDCAU-Net 模型,使用 MDA 模块能更准确地定

位壁画中的病害区域,获得清晰的病害区域边界.实验结果表明,与 Baseline 模型相比,RDCAU-Net 模型能更好地检测到不同尺度的病害区域,获得病害清晰的边界.

表 6 所示为 RDCAU-Net 模型与各种消融模型在敦煌莫高窟壁画数据集和云南少数民族壁画数据集上的定量评价指标对比.其中,从 Baseline 模型中添加不同的模块,如 MDA 和 MSAG, \*表示使用的是  $L_{BCE}$ .可以看出:(1)仅加入 RDC 模块时,在 2 个数据集上较 Baseline 模型的  $F_1$  值分别提升 0.048 4 和 0.053 5,有效地提高了网络对多尺度特征的表达能力;(2)同时加入 RDC 和 MSAG

表 6 在 2 个数据集上消融实验结果对比

模型	敦煌莫高窟壁画图像数据集				云南少数民族壁画图像数据集			
	$P \uparrow$	$R \uparrow$	$F_1 \uparrow$	$MAE \downarrow$	$P \uparrow$	$R \uparrow$	$F_1 \uparrow$	$MAE \downarrow$
Baseline	0.672 1	0.719 6	0.695 0	0.040 1	0.597 9	0.666 7	0.630 4	0.086 6
Baseline+RDC	0.724 1	0.763 8	0.743 4	0.030 1	0.631 8	0.736 8	0.683 9	0.071 6
Baseline+RDC+MSAG	0.738 4	0.816 5	0.775 5	0.029 5	0.657 3	0.736 9	0.692 8	0.072 4
Baseline+RDC+MDA	0.729 5	0.779 6	0.753 7	0.029 8	0.641 7	<b>0.778 9</b>	0.703 6	0.064 5
RDCAU-Net*	0.719 2	0.753 5	0.736 0	0.045 1	0.592 0	0.742 7	0.658 8	0.068 2
RDCAU-Net	<b>0.787 9</b>	<b>0.828 6</b>	<b>0.807 7</b>	<b>0.021 7</b>	<b>0.704 6</b>	0.755 0	<b>0.728 9</b>	<b>0.056 2</b>

注. 粗体表示各指标的最佳值.



模块时,在2个数据集上较 Baseline 模型的  $F_1$  值分别提升 0.080 5 和 0.062 4,验证了多尺度特征融合策略的可行性和有效性;(3)同时加入 RDC 和 MDA 模块时,在2个数据集上较 Baseline 模型的  $F_1$  值分别提升 0.058 7 和 0.073 2,验证了 MDA 模块能增强模型对关键特征的关注;(4)同时加入 RDC, MSAG 和 MDA 模块构成 RDCAU-Net 模型时,在2个数据集上较 Baseline 模型的  $F_1$  值分别提升 0.112 7 和 0.098 5;(5)当 RDCAU-Net 模型中使用的  $L$  替换为  $L_{BCE}$  时,各项评价指标均出现退化,验证了网络模型中使用中间监督策略的有效性。

### 3 结 语

本文提出一种 RDCAU-Net 模型,从4个方面提高壁画病害检测的性能。(1)使用 RDC 模块提升模型对多尺度特征提取和表达的能力,有效地检测壁画复杂背景中不同尺度的病害区域;(2)在编码器与解码器之间加入 MSAG 模块,以实现高效地跳转连接,使得网络融合低阶和高阶的跨层互补信息,突出特定区域的显著特征;(3)设计新的 MDA 模块抑制背景信息的干扰,提高复杂背景中病害的检测精度;(4)构造多阶段损失函数监督不同特征层的预测图。本文收集敦煌莫高窟壁画和云南各地区少数民族古代壁画,通过数据增强的方式制作了敦煌莫高窟壁画数据集和云南少数民族壁画数据集,在这2个数据集上验证 RDCAU-Net 模型中每个模块的有效性。在2个数据集上的实验结果表明,与 Baseline 模型相比, RDCAU-Net 模型的  $F_1$  值分别提升 0.112 7 和 0.098 5;与10种方法对比的实验结果表明, RDCAU-Net 模型在视觉效果和客观评价指标上均取得了更优的结果。

由于中国各地区的古代壁画风格类型迥异,病害种类很多,特征差异很大,因此本文以收集到的敦煌莫高窟壁画、丽江白沙壁画、建水指林寺壁画、西双版纳曼宰竜佛寺壁画和傣族佛寺壁画为研究对象。由于收集到的古代壁画数据数量有限,虽然通过数据扩充方法来弥补训练数据集的不足,但这些扩充的数据存在信息冗余,在一定程度上影响了网络模型的性能以及泛化能力。RDCAU-Net 模型为中国各地古代壁画病害区域的智能检测提供了一种新的思路和参考,是采用深度学习方法自动检测古代壁画病害区域的一种新的尝试和探索,可为壁画文物病害数据的归档和破损修复提供指

导,有助于保护和传承云南乃至全国各地区的壁画文化遗产。未来将不断跟进国内外深度神经网络的发展前沿,改进模型结构,同时,进一步收集和采集中国各地区古代壁画图像,扩充各类型壁画数据集,探索更具普适性的古代壁画病害检测网络模型。

### 参考文献(References):

- [1] Jaidilert S, Farooque G. Crack detection and images inpainting method for Thai mural painting images[C] //Proceedings of the IEEE 3rd International Conference on Image, Vision and Computing. Los Alamitos: IEEE Computer Society Press, 2018: 143-148
- [2] Zhang Ziyang, Shui Wuyang, Zhou Mingquan, *et al.* Research on disease extraction and inpainting algorithm of digital grotto murals[J]. Application Research of Computers, 2021, 38(8): 2495-2498+2504(in Chinese)  
(张子迎, 税午阳, 周明全, 等. 数字壁画病害提取与修复算法研究[J]. 计算机应用研究, 2021, 38(8): 2495-2498+2504)
- [3] Cao J F, Li Y F, Cui H Y, *et al.* Improved region growing algorithm for the calibration of flaking deterioration in ancient temple murals[J]. Heritage Science, 2018, 6(1): Article No.67
- [4] Zhang Haoyuan, Xu Dan, Luo Haini, *et al.* Multi-scale mural restoration method based on edge reconstruction[J]. Journal of Graphics, 2021, 42(4): 590-598(in Chinese)  
(张豪远, 徐丹, 罗海妮, 等. 基于边缘重建的多尺度壁画修复方法[J]. 图学学报, 2021, 42(4): 590-598)
- [5] Gao Y L, Zhou Z J. Automatic recognition and repair system of mural image cracks based on cloud edge computing and digitization[J]. Mobile Information Systems, 2022, 2022: Article No.1534596
- [6] Lin Y X, Xu C Q, Lyu S Q. Disease regions recognition on mural hyperspectral images combined by MNF and BP neural network[J]. Journal of Physics: Conference Series, 2019, 1325: Article No.012095
- [7] Zhang Ye'e, Wu Ligang. The detection of cave mural damage based on deep learning[J]. Yungang Research, 2022, 2(1): 85-90(in Chinese)  
(张叶娥, 吴利刚. 基于深度学习的石窟壁画破损检测[J]. 云冈研究, 2022, 2(1): 85-90)
- [8] Wu L G, Zhang L, Shi J H, *et al.* Damage detection of grotto murals based on lightweight neural network[J]. Computers and Electrical Engineering, 2022, 102: Article No.108237
- [9] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C] //Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Heidelberg: Springer, 2015: 234-241
- [10] Wu M, Jia M, Wang J. TMCrack-Net: a U-shaped network with a feature pyramid and transformer for mural crack segmentation[J]. Applied Sciences, 2022, 12(21): Article No.10940
- [11] Liu Z, Mao H, Wu C Y, *et al.* A ConvNet for the 2020s[C] //Proceedings of the IEEE/CVF Conference on Computer Vi-

- sion and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2022: 11966-11976
- [12] Lyu Shuqiang, Wang Shihan, Hou Miaole, *et al.* Extraction of mural paint loss diseases based on improved U-Net[J]. *Geomatics World*, 2022, 29(1): 69-74(in Chinese)  
(吕书强, 王诗涵, 侯妙乐, 等. 基于改进 U-Net 的壁画颜料层脱落病害区域提取[J]. *地理信息世界*, 2022, 29(1): 69-74)
- [13] Yu K, Li Y H, Yan J, *et al.* Intelligent labeling of areas of wall painting with paint loss disease based on multi-scale detail injection U-Net[C] //Proceedings of the SPIE, Optics for Arts, Architecture, and Archaeology VIII. Bellingham: Society of Photo-Optical Instrumentation Engineers, 2021: Article No.1178409
- [14] Zhou Z W, Siddiquee M M R, Tajbakhsh N, *et al.* UNet++: re-designing skip connections to exploit multiscale features in image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1856-1867
- [15] Wang H N, Cao P, Wang J Q, *et al.* UCTransNet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer[C] //Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2022: 2441-2449
- [16] Szegedy C, Vanhoucke V, Ioffe S, *et al.* Rethinking the inception architecture for computer vision[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 2818-2826
- [17] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C] //Proceedings of the 32nd International Conference on International Conference on Machine Learning. Lille: JMLR.org, 2015: 448-456
- [18] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 770-778
- [19] Oktay O, Schlemper J, Le Folgoc L, *et al.* Attention U-Net: learning where to look for the pancreas[OL]. [2023-07-12]. <https://arxiv.org/abs/1804.03999>
- [20] Ghaznavi A, Rychtáriková R, Saberioon M, *et al.* Cell segmentation from telecentric bright-field transmitted light microscopy images using a Residual Attention U-Net: a case study on HeLa line[J]. *Computers in Biology and Medicine*, 2022, 147: Article No.105805
- [21] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 7132-7141
- [22] Hu J, Shen L, Albanie S, *et al.* Gather-excite: exploiting feature context in convolutional neural networks[C] //Proceedings of the 32nd International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc, 2018: 9423-9433
- [23] Fu J, Liu J, Tian H J, *et al.* Dual attention network for scene segmentation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 3141-3149
- [24] Woo S, Park J, Lee J Y, *et al.* CBAM: convolutional block attention module[C] //Proceedings of the 15th European Conference on Computer Vision. Heidelberg: Springer, 2018: 3-19
- [25] Roy A G, Navab N, Wachinger C. Recalibrating fully convolutional networks with spatial and channel “squeeze and excitation” blocks[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(2): 540-549
- [26] Wang Q L, Wu B G, Zhu P F, *et al.* ECA-Net: efficient channel attention for deep convolutional neural networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 11531-11539
- [27] Wei J, Wang S H, Huang Q M. F<sup>3</sup>Net: fusion, feedback and focus for salient object detection[C] //Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020: 12321-12328
- [28] Qin X B, Zhang Z C, Huang C Y, *et al.* BASNet: boundary-aware salient object detection[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 7471-7481
- [29] de Boer P T, Kroese D P, Mannor S, *et al.* A tutorial on the cross-entropy method[J]. *Annals of Operations Research*, 2005, 134(1): 19-67
- [30] Mátyus G, Luo W J, Urtasun R. DeepRoadMapper: extracting road topology from aerial images[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 3458-3466
- [31] Tang F F, Wang L T, Ning C P, *et al.* CMU-Net: a strong ConvMixer-based medical ultrasound image segmentation network[OL]. [2023-07-12]. <https://arxiv.org/abs/2210.13012>
- [32] Lou A, Guan S Y, Loew M. DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation[C] //Proceedings of the SPIE 11596, Medical Imaging 2021: Image Processing. Bellingham: Society of Photo-Optical Instrumentation Engineers, 2021: Article No.115962T
- [33] Tabernik D, Šela S, Skvarč J, *et al.* Segmentation-based deep-learning approach for surface-defect detection[J]. *Journal of Intelligent Manufacturing*, 2020, 31(3): 759-776