

融合双注意力机制和 U-Net 网络的锈蚀图像分割

陈法法，成孟腾，杨蕴鹏，陈保家，肖文荣，肖能齐
(三峡大学水电机械设备设计与维护湖北省重点实验室，443002，湖北宜昌)

摘要：针对传统方法难以精确分割出金属构件锈蚀区域特征的难题，构建了一种融合双注意力机制和 U-Net 深度学习网络的锈蚀图像区域分割模型。首先，基于 U-Net 网络的对称编解码架构搭建骨干网络，采用 VGG16 网络的预训练权重对模型参数进行初始化；其次，在下采样和上采样之间的跨层连接中融合双注意力机制使网络聚焦于局部锈蚀特征，同时在上采样中使用深度可分离卷积加速模型的运算效率；最后采用锈蚀图像数据集对该网络进行训练从而得到锈蚀图像分割模型。通过实采的金属构件锈蚀图像对所提模型进行验证，结果表明：所构建的锈蚀图像分割模型能够有效地从复杂背景图像中分割出锈蚀区域特征，锈蚀区域特征的识别准确率达到 95.5%，交并比指标为 81.4%；相较于传统 U-Net 方法，识别准确率和交并比指标分别提升了 3.3%和 9.2%。

关键词：图像分割；锈蚀；U-Net 网络；注意力机制；深度神经网络

中图分类号：TP319.3；TH132.2 **文献标志码：**A

DOI：10.7652/xjtuxb202112014 **文章编号：**0253-987X(2021)12-0119-10


OSID 码

A Segmentation Method Based on Dual Attention Mechanism
and U-Net for Corrosion Images

CHEN Fafa, CHENG Mengteng, YANG Yunpeng, CHEN Baojia, XIAO Wenrong, XIAO Nengqi
(Hubei Key Laboratory of Hydroelectric Machinery Design & Maintenance, China Three Gorges University, Yichang 443002, China)

Abstract: A novel segmentation model based on dual attention mechanism and U-Net deep learning network for corrosion images is constructed to solve the problem that traditional methods are difficult to accurately segment the features of corrosion areas from corrosion images of metal components. Firstly, the symmetrical code-decode architecture of U-Net is used to construct the backbone network, and the pre-training weights of VGG16 network are adopted to initialize parameters for the proposed model. Secondly, the dual attention mechanism is integrated in the cross layer connection between down-sampling and up-sampling to make the network focus on local corrosion features. At the same time, the deep separable convolution is used in up-sampling to improve the computational efficiency. Finally, the network is trained by using the corrosion image data set to obtain the corrosion image segmentation model. The effectiveness and practicability of the proposed model are verified by the actual corrosion images of metal components. Experimental results show that the proposed segmentation model effectively segments the corrosion area features from the complex background images. The recognition accuracy of corrosion area features reaches 95.5%, and the index of the intersection over union is

81.4%。Compared with the traditional U-Net model, the accuracy and the index of intersection over union of the proposed model are improved by 3.3% and 9.2%, respectively.

Keywords: image segmentation; corrosion; U-Net network; dual attention mechanism; deep neural networks

金属锈蚀是指金属材料在周围介质(水、空气、化学溶剂等)作用下产生物理体积与化学性质的损耗与破坏的过程。在工程上,锈蚀会直接引起金属构件断面面积减少、截面应力提高,由此导致构件承载能力、刚度和稳定性下降。金属锈蚀不仅会缩短金属构件的使用寿命,甚至会造成严重的灾难性事故。例如:我国广州海印桥在使用6.5年后、济南黄河公路大桥使用13年后,均因拉索锈蚀严重,不得不全部更换^[1];美国加利福尼亚州 Folsom 坝溢洪道弧形钢闸门由于锈蚀严重,支臂不能有效承载扭曲弯矩,闸门在关闭时突然发生垮塌^[2]。在已发生的事事故警示下,各主管部门已经充分认识到锈蚀检测、评估和防护加固的重要性。

金属构件产生锈蚀后,其表面呈现出凹凸不平、锈皮泛起等特征形貌,这些特征是进行锈蚀等级评测的重要数据来源。目前工程界对金属构件局部锈蚀特征的检测主要依赖于人工目测,然而由于金属构件表面在锈蚀形成初期特征微弱,在锈蚀动态发展过程中锈蚀分布不规则,利用人工目测很难对其锈蚀程度、锈蚀面积及锈蚀颜色等特征信息进行定量描述,检测结果的可靠性不易保证。

随着机器视觉和图像处理技术的快速发展,图像作为一种记录和描述信息的载体在金属锈蚀检测领域展现出强大的应用潜力^[3]。锈蚀区域检测可以理解为图像前景分割问题,通过图像采集设备获取金属表面的锈蚀图像,结合锈蚀区域的颜色和纹理特征建立像素分类准则,利用图像处理算法提取锈蚀区域特征,进而实现对金属构件锈蚀等级的定量评测。例如:郭建斌等通过对腐蚀图像进行灰度及二值化处理,对水工结构钢表面锈蚀特征及分布状况进行了定量描述^[4];卢树杰等根据锈蚀区域 HSV 空间的颜色特征,并结合单目视差原理对钢结构表面的锈蚀区域进行了检测与分割^[5];文献[6]融合机器视觉和数据挖掘方法,从金属表面提取出锈蚀区域数字特征。基于数字图像处理的锈蚀分割方法对于理想场景、简单背景、规则构件下的锈蚀图像分割具有一定的适用性。在实际工程应用中,不规则金属构件往往受限于复杂背景干扰、遮挡等特殊工作场景,传统数字图像处理方法难以客观准确的获取

锈蚀分割特征。

近年来,深度学习的迅速发展为锈蚀区域分割提供了一种新的思路,它具备准确可靠地检测并提取复杂背景下锈蚀特征的潜质,能够为锈蚀特征量化评估提供理论数据支撑^[7]。Liu 等研究了海洋钢结构涂层破损与锈蚀检测问题,使用 VGG19 和 Faster R-CNN 对边缘锈蚀、表面锈蚀和未锈蚀 3 种类型的区域进行分类^[8];Papamarkou 等使用 ResNets 深度学习网络对核燃料的不锈钢罐中的腐蚀进行检测,能够快速推断出图像中的目标是否存在腐蚀迹象^[9]。但是上述方法只能对锈蚀状况定性分析,却无法准确识别并提取锈蚀区域。Ortiz 等通过飞行器采集被检金属表面图像,采用 FFNN 构建 CBC 检测模型,能够从图像中识别并提取出锈蚀区域^[10];王达磊等基于 U-Net 架构设计了深度神经网络,对自制的苏通大桥锈蚀图像进行语义分割,并对提取到的锈蚀区域进行了定量分析^[11];普渡大学 Qian Cheng 采用 RefineNet、PSPNet、U-Net 等深度网络对金属表面锈蚀图像进行语义分割,实验表明在小样本数据中 PSPNet 和 U-Net 均能够有效识别并提取锈蚀区域,而其中 U-Net 识别准确率较高且识别速度较快^[12]。

上述基于深度学习的锈蚀区域分割模型针对典型的金属构件锈蚀缺陷取得了一定的分割效果,但对于锈蚀程度较轻或是金属构件局部小区域的锈蚀特征,其分割精度往往偏离真实的锈蚀场景。同时,传统的深度学习方法在小样本数据集中也很难完整的聚焦于全部锈蚀特征像素,存在锈蚀特征误判的情形。

U-Net 深度学习网络作为经典的深度学习模型,它具备训练样本少、分割精度高等优势^[13]。在 U-Net 网络中嵌入注意力机制能够对锈蚀图像语义分割进行有效监督^[14-16],使网络模型聚焦于学习局部锈蚀语义特征,精确辨识伪锈蚀特征,从而进一步优化网络的图像分割效果,增强 U-Net 网络的鲁棒性和泛化推广能力。基于此,本文在 U-Net 深度学习网络中嵌入双注意力机制,构建集成的锈蚀图像分割模型,以实现复杂背景下锈蚀区域特征的定量检测和精确分割。

1 基础理论

基于深度学习语义分割的锈蚀区域检测方法,本质是一个像素级别的辨识问题^[17],基本原理如下:首先将锈蚀图像输入到全卷积神经网络,进行前向传播,输出与原图尺寸一致的概率图;再将其与手工标注的二值化标签图像进行比对,分别对每一个像素进行损失值计算,并据此进行反向传播,指导网络更新权重,如此迭代更新,直到网络模型学习到最优权重参数。训练好的模型能够区分图像上的锈蚀像素或非锈蚀像素,对锈蚀像素和非锈蚀像素分别涂上不同颜色,即可得到精准的锈蚀区域分割图。

1.1 传统 U-Net 模型

U-Net 网络是基于 FCN 优化改进得到的一种轻量级全卷积神经网络,其网络结构主要包括编码器和解码器两部分,分别对应于图像的下采样操作和上采样操作^[18-19]。编码器由 4 个相同的编码块组成,每组编码块包含 2 个卷积核为 3×3 的卷积层和 1 个 2×2 的最大池化层(下采样),激活函数采用 ReLU,每经过一次下采样操作,特征图尺寸减半,通道数翻倍;解码器由 4 个相同的解码块组成,每组解码块包含 2 个卷积核为 3×3 的卷积层和 1 个 2×2 的反卷积层(上采样),将反卷积之后的结果与来自对应下采样的特征图进行拼接,之后采用 2×2 卷积进行卷积运算,激活函数同样采用 ReLU,每经过一次上采样操作,特征图尺寸翻倍,通道数减半;在上采样的最后一层中,采用 1 个卷积核为 1×1 的卷积层将特征向量映射到网络的输出层。传统 U-Net 模型结构^[20-21]如图 1 所示。

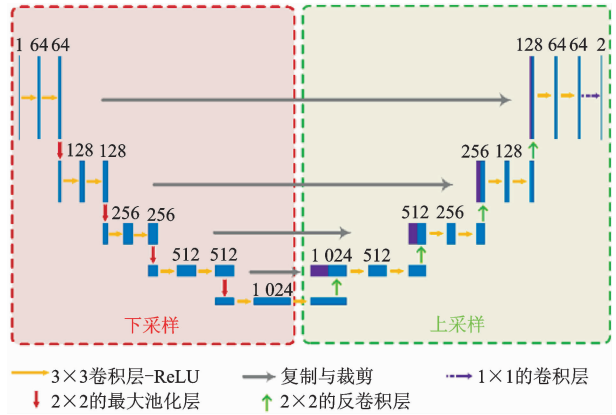


图 1 U-Net 网络结构图

Fig. 1 Network architecture of U-Net algorithm

1.2 双注意力机制

在锈蚀区域检测中,为了在复杂背景下实现

局部微小锈蚀区域、伪锈蚀区域进行准确区分,设计了双注意力特征融合模块,将其运用在下采样和上采样之间同尺度的跨层连接中,使网络聚焦于局部锈蚀特征的识别,改善分割效果。

双注意力特征融合模块如图 2 所示。首先来自上采样的深层特征图 H 包含丰富的分类信息,来自下采样的浅层特征图 L 包含丰富的位置信息,令前者通过通道注意力模块 CAM,后者通过空间注意力模块 SAM,实现特征重标定;然后将二者相加进行融合,送入通道注意力模块 CAM,形成结合位置和分类信息的注意力权重,将该权重与浅层特征图 L 进行相乘,增强对锈蚀特征的学习;最后将得到的特征图与深层特征图 H 相加,得到输出特征图 O 。

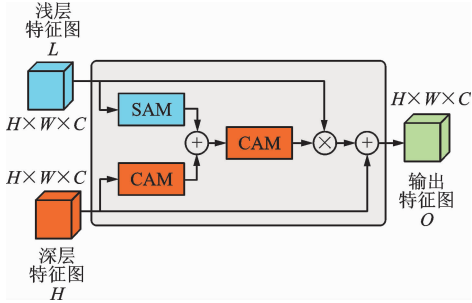


图 2 双注意力特征融合模块

Fig. 2 Dual attention feature fusion module

在 DAFFM 模块中,设计了两种注意力子模块,即通道注意力模块(CAM)和空间注意力模块(SAM)^[22-23],分别如图 3、图 4 所示。在通道注意力模块 CAM 中,首先对输入特征图 X 做全局平均池化(global average pooling, GAP)操作,得到每一个通道特征图的全局信息特征值,然后进行 1×1 卷积,最后通过 ReLU 激活和 Sigmoid 操作对特征进行非线性转换,将数值压缩到 0 到 1 之间,生成注意力权重。利用该权重与输入特征图 X 进行相乘得到输出特征图 Y ,从而实现通道方向上特征重标定。

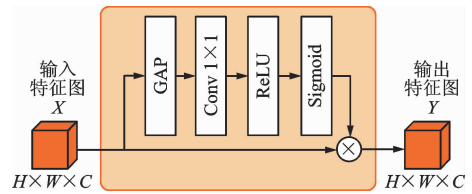


图 3 通道注意力模块

Fig. 3 Channel attention module

在空间注意力模块 SAM 中,利用 1×1 卷积核,将通道数降至 1,此时特征图中每一个像素都是该像素在全部通道方向上的一个特征值,对该特征

值进行 Sigmoid 操作,生成注意力权重,再与输入特征图 X 相乘,即可得到宽和高方向上重标定的特征图 Y 。

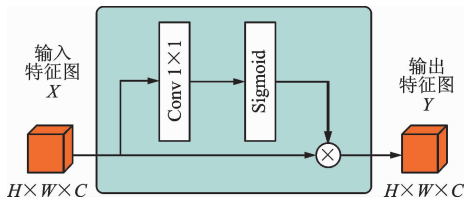


图 4 空间注意力模块
Fig. 4 Spatial attention module

1.3 深度可分离残差卷积

为了提升语义分割模型的运行速度,需要在保证识别精度的同时,尽可能地减小模型尺寸、降低计算复杂度^[24]。为此,本文借鉴 MobileNet 和 ResNet 设计了一种深度可分离残差卷积模块(DSRCM),将其运用在上采样阶段,同时引入残差结构,对浅层和深层特征进行融合,提高特征提取能力。此外,该模块中还使用了组归一化(group normalization, GN),来适应小批量数据的训练。深度可分离残差卷积模块如图 5 所示。

2 锈蚀图像分割模型与流程

2.1 锈蚀图像分割模型构架

本文基于 U-Net 设计了一种用于锈蚀区域检

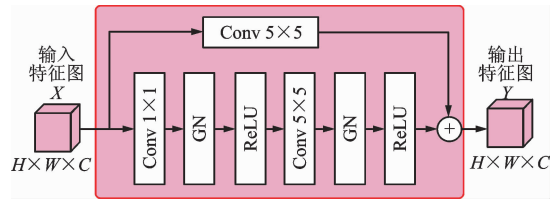


图 5 深度可分离残差卷积模块
Fig. 5 Depth wise separable residual convolution module

测的语义分割模型,其基本架构如图 6 所示。为了便于理解,将各层特征图的尺寸(记为 a)和通道数(记为 b)以 $a^2 \times b$ 的形式标注,各特征图所在的层数在侧边使用圆圈序号进行了标注,各符号的含义在图例中给出。

网络整体分为下采样、上采样两大部分。在下采样过程中,使用去掉全连接层的 VGG16 作为前端特征提取器。第 1 层使用由两组 3×3 卷积(步长为 1,填充边距为 1)、批归一化(batch normalization, BN)、ReLU 激活组成的标准卷积模块提取特征。第 2 层使用最大池化进行下采样后,执行前述标准卷积操作。第 3 层到第 5 层中,为了加强网络的特征提取能力,使用了 3 组上文中的卷积池化操作。在下采样过程中,每增加一层,特征图尺寸减半,通道数翻倍(其中第 5 层通道数不变)。

在上采样过程中,为了防止过拟合,首先将第 5

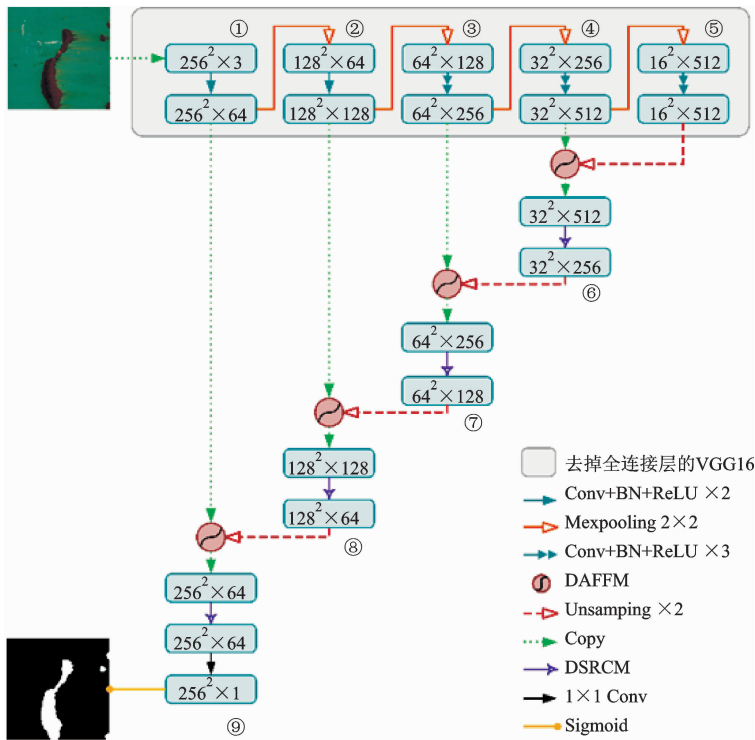


图 6 锈蚀图像语义分割网络结构

Fig. 6 Network structure of semantic segmentation for corrosion images

层以 0.5 的概率进行 Dropout 操作;再经过双线性插值上采样后,与来自下采样阶段第 4 层的同尺度特征图一并送入注意力模块中,输出加强语义信息的特征图。为了进一步提取特征,将特征图继续送入深度可分离残差卷积模块。在第 6 层到第 9 层均采取了上述操作,每增加一层,特征图尺寸翻倍,通道数减半。对第 9 层的特征图,使用 1×1 卷积,将特征图通道数降至 1,然后进行 Sigmoid 操作,从而得到预测概率图。在预测时,使用第 9 层的输出概率图,以 0.5 作为阈值生成二值化分割图像。

2.2 锈蚀图像分割流程

本文模型中的模型训练与测试流程如图 7 所示。首先对锈蚀图像数据集进行预处理,划分为训练集、测试集和验证集;然后使用 VGG16 预训练权重对模型进行初始化,并结合训练集和验证集使用动态学习率衰减策略对模型进行迭代更新;最后将测试集输入到训练好的锈蚀图像分割模型上进行测试,从而实现锈蚀图像的区域分割。

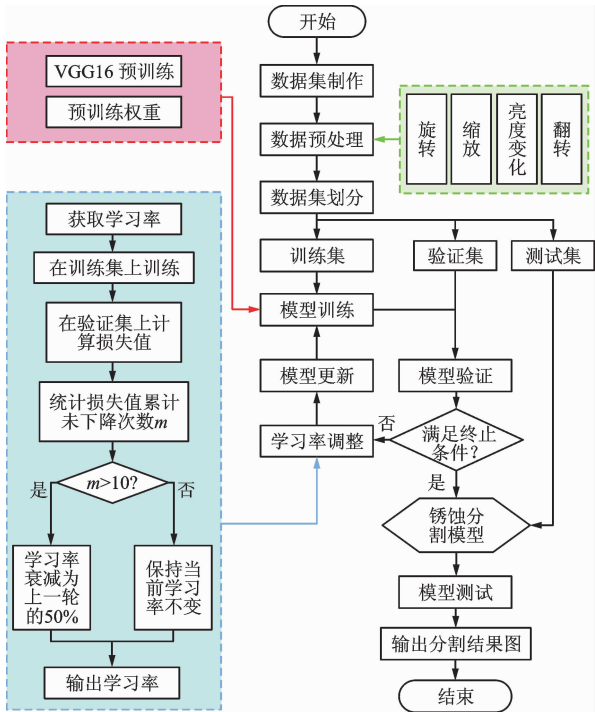


图 7 本文模型的训练与测试流程

Fig. 7 The training and testing process of the proposed model

2.3 损失函数

锈蚀图像语义分割本质上是一个像素级别的二分类任务,通常采用二进制交叉熵损失函数(binary cross entropy loss, BCE Loss)计算误差。然而,在锈蚀图像中,锈蚀像素数远少于非锈蚀像素数,存在

类不平衡问题,使用二进制交叉熵会导致所预测结果被非锈蚀像素所主导,造成识别能力下降。同时,在锈蚀图像中,锈蚀区域往往比非锈蚀区域更难识别,存在难易不平衡问题。因此,为了克服这些问题,本文采用焦点损失函数,其计算公式^[25]如下

$$F_L(p_i) = -\alpha_i(1 - p_i)^\gamma \log(p_i) \tag{1}$$

$$p_i = \begin{cases} p, & y = 1 \\ 1 - p, & y = 0 \end{cases} \tag{2}$$

式中: p 为模型输出的类别概率, $p \in [0, 1]$; α_i 为第 i 类样本的损失权重,所有类别的损失权重之和为 1; γ 控制难易样本损失的大小,且 $\gamma \geq 0$ 。当 γ 增大时,模型会更加关注难以区分的样本;当 $\gamma = 0$ 时,焦点损失函数退化成带 α_i 的普通交叉熵损失函数。

2.4 评价指标

为了定量评价本文所构建的网络模型在锈蚀图像语义分割中的性能,采用准确率(A_{cc})、精确率(P_{re})、召回率(R_{ec})、Dice 系数(D_{ice})和交并比(I_{ou})等作为评价指标,计算公式如下

$$A_{cc} = \frac{T_P + T_N}{T_P + T_N + F_N + F_P} \tag{3}$$

$$P_{re} = \frac{T_P}{T_P + F_P} \tag{4}$$

$$R_{ec} = \frac{T_P}{T_P + F_N} \tag{5}$$

$$D_{ice} = \frac{2T_P}{2T_P + F_P + F_N} \tag{6}$$

$$I_{ou} = \frac{T_P}{T_P + F_P + F_N} \tag{7}$$

式中: T_P 为实际是锈蚀且被准确识别为锈蚀的像素点数量; F_N 为实际是锈蚀却被识别为非锈蚀的像素点数量; T_N 为实际是非锈蚀且被准确识别为非锈蚀的像素点数量; F_P 为实际是非锈蚀却被识别为锈蚀的像素点数量。上述指标的取值范围均在 0 到 1 之间,数值越大,则表明模型性能越好。

3 实验与验证实例

3.1 数据集

由于目前没有公开的锈蚀图像数据集,本文所采用的数据集来源于网上搜索的 500 张不同尺寸的锈蚀图像。本文数据集图像涵盖了多种应用场景下的锈蚀特征,从带有局部锈蚀的钢架桥到整体发生锈蚀的钢制管道等,锈蚀大小与锈蚀程度覆盖范围较广。

锈蚀通常伴随着钢材表面的涂层破损,锈蚀部位往往会呈现红棕色,严重时甚至变为黑红色,而且

锈蚀表面较为粗糙。因此,在对原始数据集制作标签时,按照以上经验对图像中的锈蚀区域进行判定并标注,标注工具采用 Photoshop CS6。原始数据集中每一例数据都包括标签图像,图像标注如图 8 所示,其中图 8a 为原始的锈蚀图像,图 8b 为最终的标签图像,锈蚀区域用白色标注,其像素值设为 255,背景区域用黑色标注,其像素值设为 0。在完成所有图像数据的标注工作后,按照 8:1:1 的比例将数据集划分为训练集(400 张)、验证集(50 张)、测试集(50 张)。

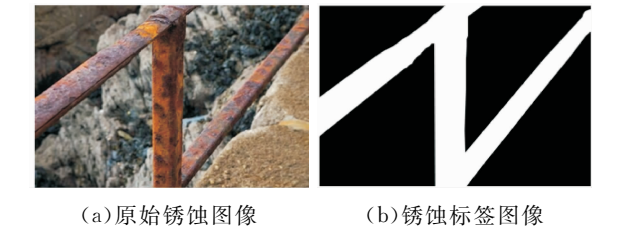


图 8 锈蚀图像标注
Fig. 8 Marking of corrosion image

3.2 数据预处理

为了增加数据的多样性,避免过拟合发生,采用水平翻转、垂直翻转、放大 1.4 倍、旋转 45°并放大 1.4 倍、亮度增强至 1.2 倍、亮度降低至 0.8 倍以及高斯模糊等操作对训练集和验证集进行离线增强。以上增强操作可将训练集数据扩充至 3 200 张,验证集扩充至 400 张。对于测试集,则不采取增强操作,直接用于模型评估从而保证测试集数据的真实性。

3.3 模型训练

为了对所提方法进行合理评估,所有模型均在 Windows10+Python3.6+Pytorch1.0 环境下训练和测试,使用内存为 24 GB 的 GeForce RTX 3090 显卡进行加速。

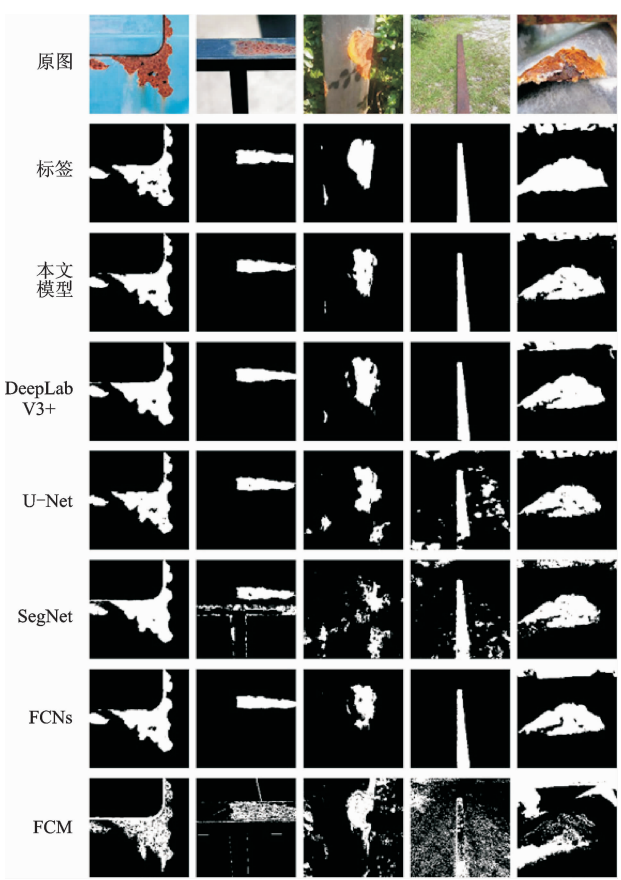
通过预训练权重来改善模型的性能,使用在大规模数据集 ImageNet 上训练得到的 VGG16 模型参数初始化所提模型的编码部分,然后与解码部分结合进一步训练和微调,将 VGG16 强大的特征提取能力迁移到锈蚀图像的锈蚀区域分割中。

通过选择自适应矩估计优化器(Adam)来进行网络优化更新,设置一阶矩估计的指数衰减率为 0.9,二阶矩估计的指数衰减率为 0.999,权重衰减系数为 0.000 01。设置批大小为 4,训练轮数为 50 个 epoch,初始学习率为 0.001,使用 ReduceLROn-Plateau 动态学习率调整策略,完成学习率自动衰减。即在每个 epoch 训练完成后,就在验证集上进

行测试,记录其损失值。当检测到 2 个 epoch 下验证集的损失值未发生下降时,将学习率调整为原来的 50%。学习率最小值设置为 0.000 001。

3.4 实例验证

3.4.1 不同算法效果对比 为了验证本文所提方法在锈蚀语义分割中的有效性,选择传统图像分割模糊 C 均值法(FCM)、传统的深度语义分割模型 FCN、SegNet、U-Net、DeepLab V3+ 作对比测试,对同一组测试图像进行锈蚀分割,其检测结果如图 9 所示。



(a)铁板 (b)护栏 (c)金属支柱 (d)铁架 (e)铁片
图 9 不同锈蚀分割方法的效果对比

Fig. 9 Performance comparison of different methods for corrosion images segmentation

图 9 中每一列分别由锈蚀图像原图、人工标注的真实图像以及不同模型的输出图像组成。通过对比图 9 可以发现:当图像中前景与背景差异较大、纹理较简单时,传统 FCM 图像分割模型借助人工筛选结果,能够粗略分割出锈蚀区域(如图 9a 所示),但在复杂背景下,几乎无法辨识出锈蚀区域(如图 9d、图 9e 所示);深度学习模型总体上比传统图像分割效果更好;SegNet、U-Net 对非锈蚀区域的区分

能力较差,分割出的锈蚀区域出现孔洞、轮廓断开的情况较多,边缘不够平滑(如图 9b、图 9d 所示);FC-Ns 在总体轮廓的识别上效果较好,但细节上无法聚焦,因此出现了一些误判(如图 9d、图 9e 所示);DeepLab V3+整体上识别效果较好,但是部分图像中出现较多孔洞(如图 9b、9c);本文模型所设计的注意力机制以及预训练策略,使得模型更加聚焦于锈蚀区域与非锈蚀区域的细节区分,得到的锈蚀区域轮廓连续,且对于面积小的局部锈蚀区域也能准确识别(如图 9b 所示)。

采用 3.1 节中的评价指标基于测试集分别对不同深度学习模型进行定量分析,采用单张图片的测试时间 t 分析模型时间复杂度,各算法在 50 张测试集图像上评价指标的平均值如表 1 所示。从表 1 可以看出:本文模型在准确率、召回率、Dice 系数和 IoU 系数上性能明显优于传统深度学习方法,在精确率上也较传统方法有较大优势;单张图片的平均测试时间中,相较于 FCNs 等经典网络,本文模型由于结构更为复杂,因此测试时间略有增加,而语义分割常用网络 DeepLab V3+由于骨干网络层数更深导致测试时间相对较长。

表 1 不同深度学习模型的锈蚀图像分割性能对比
Table 1 Performance comparison of different deep learning methods for corrosion images segmentation

锈蚀分割模型	A_{cc}	P_{re}	R_{ec}	D_{ice}	I_{oU}	t/s
FCNs	0.945	0.871	0.865	0.849	0.771	0.253
SegNet	0.892	0.799	0.816	0.779	0.666	0.201
U-Net	0.922	0.885	0.793	0.817	0.722	0.205
DeepLab V3+	0.926	0.869	0.826	0.829	0.738	4.912
本文模型	0.955	0.875	0.906	0.883	0.814	0.596

3.4.2 各模块效果验证 为了分析本文所构建的网络中各个模块的作用,训练了 4 个不同的网络用于对比分析。第 1 个模型为原始的 UNet;第 2 个模型为 VGG-UNet,在第 1 个模型的基础上,前端特征提取部分采用了去掉全连接层的 VGG16 结构,上采样采用了深度可分离残差卷积;第 3 个模型为 VGG-UNet-Att,在第 2 个模型的基础上加上了注意力机制;第 4 个模型为 P-VGG-UNet-Att,在第 3 个模型上加上了预训练权重,这也是本文所采用的最终模型。以上 4 个模型均采用同一套数据集,使用相同的损失函数和训练策略,并保证相关参数一致,分别训练以上模型,其结果如表 2 所示。

从表 2 中可以看出,当逐步引入所改进的模块

时,模型在准确率、召回率、Dice 系数和交并比系数上均得到了不同程度的提升。在原始 UNet 上使用 VGG 作为前端特征提取,并采用深度可分离残差卷积后,模型的交并比上升了 2.6%;引入注意力机制后,上升了 1.6%;进一步使用预训练权重后,交并比上升了 5%。

表 2 各模型效果的定量评价表
Table 2 Quantitative evaluation of each model

锈蚀分割模型	A_{cc}	P_{re}	R_{ec}	D_{ice}	I_{iu}
U-Net	0.923	0.885	0.794	0.817	0.722
VGG-UNet	0.940	0.839	0.855	0.834	0.748
VGG-UNet-Att	0.941	0.843	0.889	0.851	0.764
本文模型	0.955	0.875	0.907	0.884	0.814

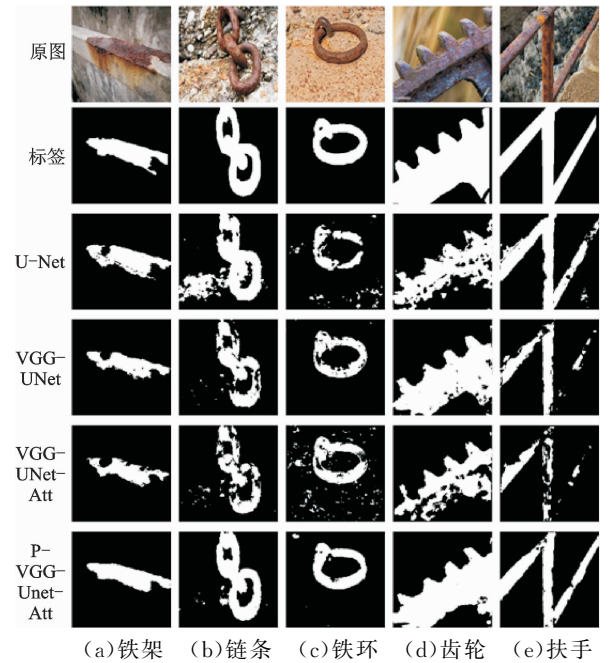


图 10 各模块效果的对比图
Fig. 10 Effect comparison of different modules

此外,对比分析一些复杂场景下的锈蚀图像,分别测试这 4 种模型的锈蚀分割效果,如图 10 所示。从图 10 可以看出,引入注意力机制以及预训练权重后,能够加强模型对锈蚀区域的辨识能力,所得到的锈蚀分割图像具有更清晰的边缘轮廓,并且对不同锈蚀程度下的同一锈蚀区域具有好的包容性,几乎没有出现孔洞和不连续的情况。

不同深度学习网络模型在训练集中 50 个 epoch 的准确率和损失函数值变化如图 11 所示。从图 11a 可以看出,FCNs 模型的收敛速度较慢,其在 30 个训练周期后才接近收敛,其他模型在 20 个训练周期即已接近收敛,由此可知本文模型收敛速度

快且波动幅度小;从图 11b 可以看出,SegNet 模型的准确率最低,而 U-Net 和 DeepLab V3+ 模型的较为接近,本文模型的准确率经过 50 个训练周期后在训练集中达到 97.8%,相比其他模型有较大优势。

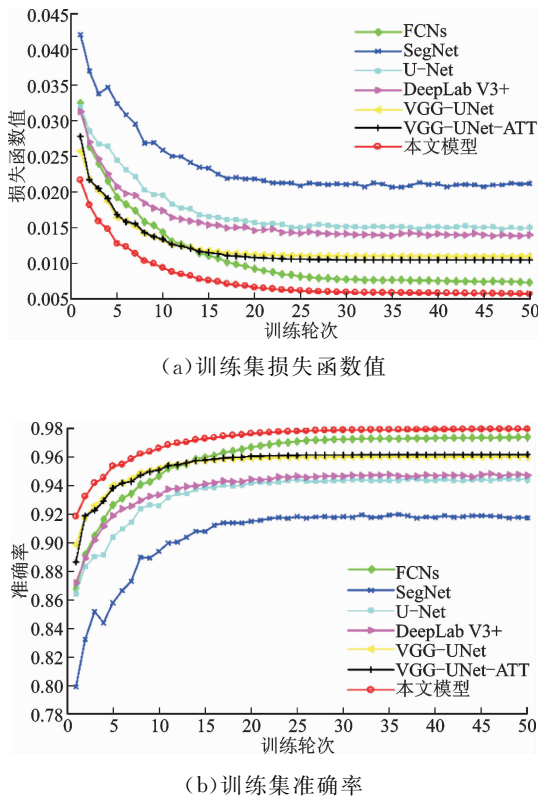


图 11 各模型在训练集上的对比曲线

Fig. 11 Comparison curve of each model on training sets

3.4.3 工程测试验证 为了验证本文网络结构的鲁棒性和有效性,对实地复杂场景采集到带有锈蚀的图像进行分割,如图 12 所示。由图 12a 可以看出,对于几种不同尺寸和形状的锈蚀钢材图像,本文

模型能够精准识别出锈蚀区域;由图 12b 可以看出,钢材在不同部位出现了不同程度的锈蚀,在左下角还存在零散的小面积点蚀,对于这种锈蚀图像,本文模型也能够定位到锈蚀区域,给出准确的锈蚀轮廓;由图 12c 可以看出,即使钢材出现了十分微弱的粉状锈蚀,模型也能识别出主要的锈蚀区域,包容性较强;由图 12d 可以看出,圆桶内壁的锈蚀区域处于不同的光照条件下,一部分被阳光直射,另一部分则处于阴影之中,对于这种情况,模型也能兼容,将二者同样识别为同一锈蚀区域;由图 12e 可以看出,钢材处于复杂的山地环境,而枯黄的叶子等由于与锈蚀颜色较为接近通常会对传统网络模型的识别造成较大干扰,但在图 12e 中其对本文模型干扰较小,模型仍能准确识别该锈蚀钢材的主体轮廓;由图 12f 可以看出,当拍摄的锈蚀图像存在景深效果时,不会对模型造成严重干扰,说明模型已经学习到锈蚀的高级语义信息,具有较强的抗干扰能力。

综上所述,本文所提模型能处理各种复杂环境下的锈蚀图像,具有较强的鲁棒性和实用性。

4 结 论

本文针对复杂多变背景图像中锈蚀区域难以精确分割的问题,提出一种融合双注意力机制和 U-Net 网络的锈蚀图像区域分割方法,得出如下结论:

(1)采用 VGG16 预训练权重对 U-Net 模型进行初始化,并采用深度可分离残差卷积模块减少模型尺寸,不仅能够有效降低小样本导致的网络过拟合问题,而且可以提高网络特征学习能力,加快收敛速度;

(2)在改进的网络结构中引入双注意力机制并

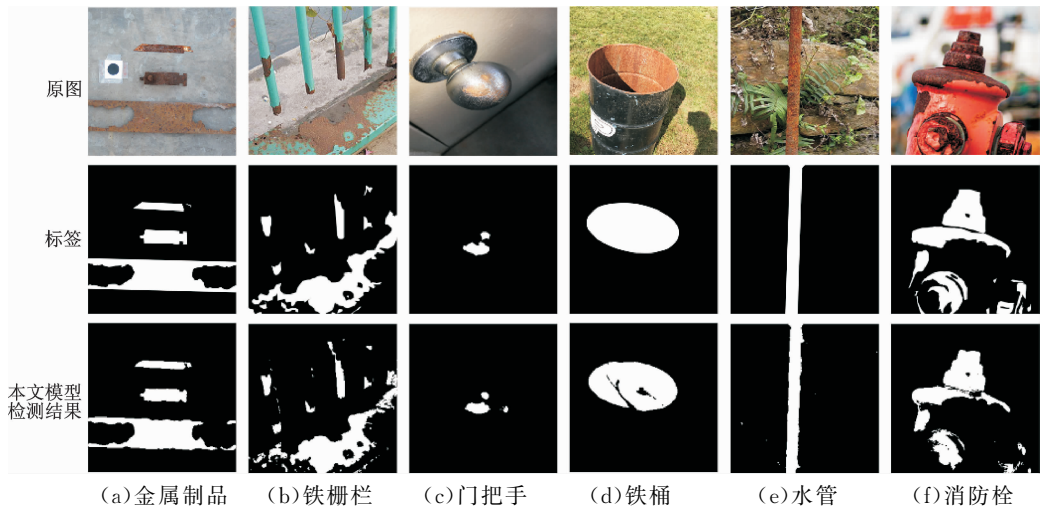


图 12 工程图像的锈蚀分割效果

Fig. 12 Effects of corrosion segmentation on engineering images

使用动态学习率衰减策略对模型进行迭代更新,加强对于锈蚀目标特征的学习,抑制复杂多变背景的干扰,实现了在复杂多变背景中的锈蚀特征检测,提高了锈蚀区域分割的精度和效率;

(3)分别从不同算法效果对比、各模块效果验证和工程测试验证多个角度,从实例上对本文算法进行验证,准确率达到 95.5%,交并比指标为 81.4%,单张图片的测试时间为 0.596 s,多个评价指标均显示出所提方法在图像锈蚀区域分割中整体性能更优,具备一定的有效性和实用性。

参考文献:

- [1] 刘红波,刘东宇,徐杰. 天津新港船闸桥锈蚀检测与结构性能评估 [J]. 天津大学学报(自然科学与工程技术版), 2015, 48(7): 147-150.
LIU Hongbo, LIU Dongyu, XU Jie. Corrosion detection and structural evaluation of lock bridge at Tianjin Newport [J]. Journal of Tianjin University (Science and Technology), 2015, 48(7): 147-150.
- [2] ANAMI K, ISHII N, KNISELY C W. Added mass and wave radiation damping for flow-induced rotational vibrations of skin plates of hydraulic gates [J]. Journal of Fluids and Structures, 2012, 35: 213-228.
- [3] 宋伟,左丹,邓邦飞,等. 高压输电线防震锤锈蚀缺陷检测 [J]. 仪器仪表学报, 2016, 37(S1): 113-117.
SONG Wei, ZUO Dan, DENG Bangfei, et al. Corrosion defect detection of earthquake hammer for high voltage transmission line [J]. Chinese Journal of Scientific Instrument, 2016, 37(S1): 113-117.
- [4] 郭建斌,王楠,王泽民. 水工钢结构腐蚀的图像识别技术 [J]. 河海大学学报(自然科学版), 2012, 40(5): 539-543.
GUO Jianbin, WANG Nan, WANG Zemin. Image identification technology of corrosion in hydraulic steel structure [J]. Journal of Hohai University (Natural Sciences), 2012, 40(5): 539-543.
- [5] 卢树杰,陈庆诚,文祥,等. 基于无人机的钢结构锈蚀视觉检测 [J]. 工业控制计算机, 2019, 32(7): 79-81.
LU Shujie, CHEN Qingcheng, WEN Xiang, et al. Visual inspection of steel structure corrosion based on UAV [J]. Industrial Control Computer, 2019, 32(7): 79-81.
- [6] HOANG N D. Image processing-based pitting corrosion detection using metaheuristic optimized multilevel image thresholding and machine-learning approaches [J]. Mathematical Problems in Engineering, 2020,

2020: 6765274.

- [7] ATHA D J, JAHANSHAHI M R. Evaluation of deep learning approaches based on convolutional neural network for corrosion detection [J]. Structural Health Monitoring, 2017, 17(5): 1110-1128.
- [8] LIU L, TAN E, YIN X J, et al. Deep learning for coating condition assessment with active perception [C]// ACM International Conference Proceedings Series. New York, USA: ACM, 2019: 75-80.
- [9] PAPAMARKOU T, GUY H, KROENCKE B, et al. Automated detection of corrosion in used nuclear fuel dry storage canisters using residual neural networks [J]. Nuclear Engineering and Technology, 2020, 53(2): 657-665.
- [10] ORTIZ A, BONNIN-PASCUAL F, GARCIA-FIDALGO E, et al. Vision-based corrosion detection assisted by a micro-aerial vehicle in a vessel inspection application [J]. Sensors, 2016, 16(12): 2118.
- [11] 王达磊,彭博,潘玥,等. 基于深度神经网络的锈蚀图像分割与定量分析 [J]. 华南理工大学学报(自然科学版), 2018, 46(12): 121-127.
WANG Dalei, PENG Bo, PAN Yue, et al. Segmentation and quantitative analysis of corrosion images based on deep neural networks [J]. Journal of South China University of Technology (Natural Science Edition), 2018, 46(12): 121-127.
- [12] QIAN Cheng. Evaluation of deep learning-based semantic segmentation approaches for autonomous corrosion detection on metallic surfaces [D]. West Lafayette, IN, USA: Purdue University, 2019: 8-10, 41-49.
- [13] ZHANG Yuhang, REN Hongshuai, YANG Wensi, et al. The strong substructure and feature attention mechanism for image semantic segmentation [EB/OL]. [2021-01-07]. <https://onlinelibrary.wiley.com/share/FKBQCBEPGRSFWBVPZPQU?target=10.1002/cpe.5920>.
- [14] HUANG G, ZHU J, LI J, et al. Channel-attention U-Net: channel attention mechanism for semantic segmentation of esophagus and esophageal cancer [J]. IEEE Access, 2020, 8: 122798-122810.
- [15] 董月,冯华君,徐之海,等. Attention Res-Unet: 一种高效阴影检测算法 [J]. 浙江大学学报(工学版), 2019, 53(2): 373-381.
DONG Yue, FENG Huajun, XU Zhihai, et al. Attention Res-Unet: an efficient shadow detection algorithm [J]. Journal of Zhejiang University (Engineering Science), 2019, 53(2): 373-381.

- [16] JIN Yuwei, XU Wenbo, HU Zhongwen, et al. GSCA-UNet: towards automatic shadow detection in urban aerial imagery with global-spatial-context attention module [J]. *Remote Sensing*, 2020, 12(17): 1-23.
- [17] 殷晓航, 王永才, 李德英. 基于 U-Net 结构改进的医学影像分割技术综述 [J]. *软件学报*, 2021, 32(2): 519-550.
- YIN Xiaohang, WANG Yongcai, LI Deying. Survey of medical image segmentation technology based on U-Net structure improvement [J]. *Journal of Software*, 2021, 32(2): 519-550.
- [18] 侯向丹, 赵一浩, 刘洪普, 等. 融合残差注意力机制的 UNet 视盘分割 [J]. *中国图象图形学报*, 2020, 25(9): 1915-1929.
- HOU Xiangdan, ZHAO Yihao, LIU Hongpu, et al. Optic disk segmentation by combining UNet and residual attention mechanism [J]. *Journal of Image and Graphics*, 2020, 25(9): 1915-1929.
- [19] LIU Zhenqing, CAO Yiwen, WANG Yize, et al. Computer vision-based concrete crack detection using U-net fully convolutional networks [J]. *Automation in Construction*, 2019, 104: 129-139.
- [20] PAN Gang, ZHENG Yaoxian, GUO Shuai, et al. Automatic sewer pipe defect semantic segmentation based on improved U-Net [J]. *Automation in Construction*, 2020, 119: 103383.
- [21] 柳小波, 张育维. 基于 U-Net 和 Res-UNet 模型的传送带矿石图像分割方法 [J]. *东北大学学报(自然科学版)*, 2019, 40(11): 1623-1629.
- LIU Xiaobo, ZHANG Yuwei. Ore image segmentation method of conveyor belt based on U-Net and Res-UNet models [J]. *Journal of Northeastern University (Natural Science)*, 2019, 40(11): 1623-1629.
- [22] 凌兴宏, 李杰, 朱斐, 等. 基于双重注意力机制的异步优势行动者评论家算法 [J]. *计算机学报*, 2020, 43(1): 93-106.
- LING Xinghong, LI Jie, ZHU Fei, et al. Asynchronous advantage actor-critic with double attention mechanisms [J]. *Chinese Journal of Computers*, 2020, 43(1): 93-106.
- [23] LUO Haifeng, CHEN Chongcheng, FANG Lina, et al. High-resolution aerial images semantic segmentation using deep fully convolutional network with channel attention mechanism [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(9): 3492-3507.
- [24] 刘媛媛, 张硕, 于海业, 等. 基于语义分割的复杂场景下的秸秆检测 [J]. *光学精密工程*, 2020, 28(1): 200-211.
- LIU Yuanyuan, ZHANG Shuo, YU Haiye, et al. Straw detection algorithm based on semantic segmentation in complex farm scenarios [J]. *Optics and Precision Engineering*, 2020, 28(1): 200-211.
- [25] 杨威, 张建林, 徐智勇, 等. 一种改进的 Focal Loss 在语义分割上的应用 [J]. *半导体光电*, 2019, 40(4): 555-559.
- YANG Wei, ZHANG Jianlin, XU Zhiyong, et al. An improved Focal Loss function for semantic segmentation [J]. *Semiconductor Optoelectronics*, 2019, 40(4): 555-559.

(编辑 刘杨)