

DOI: 10.11835/j.issn.2096-6717.2025.004



开放科学(资源服务)标识码 OSID:



基于改进 DeepLabV3+ 的既有钢结构表面锈蚀 图像分割方法

孟庆成¹, 张发¹, 熊爽¹, 齐欣²

(1. 西南石油大学 土木工程与测绘学院, 成都 610500; 2. 西南交通大学 土木工程学院, 成都 610031)

摘要: 表面锈蚀是钢结构中普遍存在的缺陷形式, 锈蚀会削弱构件截面, 降低其力学性能, 因此有必要对钢结构锈蚀损伤进行监测及状态评价。针对人工目测易出现视觉疲劳造成读数、纹理和颜色判断错误, 经典深度学习锈蚀识别网络分割精度低、参数量大等问题, 提出一种轻量化 HL-DeepLabV3+ 的锈蚀图像语义分割模型。首先, 将编码部分的主干网络替换成轻量化的 MobileNetV3 网络, 降低模型参数量; 其次, 在空洞空间金字塔池化模块引入条形池化模块和金字塔池化模块, 使网络能捕获孤立锈蚀区域的长距离依赖关系, 消除背景区域干扰; 最后, 引入注意力机制, 使网络更加关注图像中对锈蚀分类起决定作用的像素区域, 增强锈蚀的特征表达能力。结果表明: HL-DeepLabV3+ 模型相比原模型大小减小了 66.73%, 在 Accuracy、mIoU、MPA 和 F1 分数上分别提高了 2.54%、10.11%、6.79%、6.15%, 且效果优于经典的 UNet、FCN、SegNet 和 DeepLabV3+ 语义分割模型, 在实现模型轻量化的同时还提升了对锈蚀损伤的分割精度。

关键词: 钢结构; 锈蚀识别; 语义分割; 轻量化; 注意力机制

中图分类号: TU391; TP391.4 **文献标志码:** A **文章编号:** 2096-6717(XXXX)XX-0001-12

Image segmentation approach for surface rust of existing steel structures based on improved DeepLabV3+

MENG Qingcheng¹, ZHANG Fa¹, XIONG Shuang¹, QI Xin²

(1. School of Civil Engineering and Surveying, Southwest Petroleum University, Chengdu 610500, P. R. China;

2. School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, P. R. China)

Abstract: Surface corrosion constitutes a prevalent defect form in steel structures, which can undermine the section of components and deteriorate their mechanical properties. Thus, it is imperative to monitor and assess the corrosion damage of steel structures. A lightweight semantic segmentation model for rust images, based on HL-DeepLabV3+, has been proposed to address the challenges of reading errors, texture misjudgments, and color discrepancies caused by visual fatigue during manual inspections. This model aims to improve

收稿日期: 2024-09-28

基金项目: 国家自然科学基金(52078442); 四川省科技计划(2021YJ0038); 桥梁无损检测与工程计算四川省高校重点实验室
开放课题基金(2024QZJ03)

作者简介: 孟庆成(1980-), 男, 博士, 主要从事桥梁结构健康监测与损伤识别研究, E-mail: 214400395@qq.com。

Received: 2024-09-28

Foundation items: National Natural Science Foundation of China (No. 52078442); Science and Technology Program of Sichuan Province (No. 2021YJ0038); Opening Project of Sichuan Province University Key Laboratory of Bridge Non-destruction Detecting and Engineering Computing (No. 2024QZJ03)

Author brief: MENG Qingcheng (1980-), PhD, main research interests: bridge structure health monitoring and damage identification, E-mail: 214400395@qq.com.

segmentation accuracy while reducing the parameter count associated with traditional deep learning networks used in rust recognition. Firstly, the coding portion of the backbone network was replaced with a lightweight MobileNetV3 network to minimize the number of model parameters. Secondly, the strip pool module and pyramid pool module were incorporated into the void space pyramid pool module, enabling the network to capture the long-range dependence of isolated rust regions and eliminate the interference in the background region. Finally, the attention mechanism was introduced to make the network focus more on the pixel region that plays a decisive role in rust classification and enhance the feature expression ability of rust. The test results indicate that: Compared with the original model, the size of the HL-DeepLabV3+ model is reduced by 66.73%, and the Accuracy, mIoU, MPA, and F1 scores are increased by 2.54%, 10.11%, 6.79%, and 6.15% respectively. Moreover, the effect is superior to the classical UNet, FCN, SegNet, and DeepLabV3+ semantic segmentation models. This not only realizes the lightweight of the model but also enhances the segmentation accuracy of rust damage.

Keywords: steel structure; rust identification; semantic segmentation; lightweight; attention mechanism

钢结构作为土木工程领域重要的结构形式,由于其强度高、抗震性能优越、建设周期短及绿色低碳等优势^[1],在工程建设中得到广泛应用。但由于环境和使用条件的不同,钢结构表面往往会受到不同程度的损伤,损伤削弱了钢构件的力学性能和承载能力^[2],增加了结构的失效风险。钢结构常见的表面病害主要包括锈蚀、焊缝开裂及螺栓松动、缺失等。而表面锈蚀是钢结构中最主要的缺陷形式之一,锈斑和剥落不仅影响结构美观,而且会缩短结构设计使用寿命,严重锈蚀还会导致结构失稳或倒塌^[3],为了有效减少由锈蚀导致的工程事故发生,需要对钢结构表面锈蚀进行定期的病害检测,从而评估其服役性能^[4]。传统方法对于钢结构表面锈蚀的检测主要有物理检测、电化学检测和无损检测^[5],检测人员通常采用检测工具读取锈蚀深度、锈蚀面积等重要参数,同时结合人眼对锈蚀区域的纹理和颜色进行主观分析,较为灵活,但频繁的锈蚀参数测量和主观分析容易因视觉疲劳造成读数、纹理和颜色判断错误,这可能导致锈蚀病害检测结果不准确、结构状态评估结果误判以及后续更高的维修成本、设备更换费用和停工损失,甚至在使用过程中结构失效,引发安全事故,另有部分检测位置人员难以到达,使其在完成大量的锈蚀检测任务时具有一定的难度^[6]。

随着智能监测技术的快速发展,基于卷积神经网络(Convolutional Neural Networks, CNN)的语义分割技术取得了很大的进步^[7-9],因其具有高精度和实时性的优势,突破人工检测存在的局限性,成为研究的热点。目前经典的语义分割网络主要有UNet^[10]、FCN^[11]、SCNet^[12]、DeepLab系列^[13]等,这些语义分割网络也被越来越多的研究者们应用于包括锈蚀监测的检测领域。陈法法^[14]等设计的改进

UNet网络模型在复杂特征背景和无关区域等类似背景干扰下,具有较好的锈蚀图像分割效果。García Peña^[15]等基于机器学习方法提出了一种改进的FCN网络,用于使用点云数据检测金属零件表面缺陷,取得了较高的准确性,实现了快速的推理时间。为提高钢筋混凝土锈蚀裂缝检测分类的效率和精度,许颖等^[16]提出了基于CNN的钢筋混凝土锈蚀裂缝识别模型SCNet,取得了较高的识别率。为克服对航拍的金具锈蚀图片检测速度慢、信息缺失等问题,王凌云^[17]等提出了改进的DeepLabV3+网络模型,实现了对输电设备锈蚀缺陷的检测。

以上相关研究主要以解决钢结构表面锈蚀识别问题为主,对锈蚀等级的判断仍存在困难,且仍需进一步提高识别精度及效率。笔者提出一种HL-DeepLabV3+(Highprecision Lightweight, HL)的锈蚀图像语义分割模型。首先,采用MobileNetV3^[18]网络替换DeepLabV3+^[19]模型的主干网络Xception^[20],有效降低了模型参数量和复杂度;其次,在空洞空间金字塔池化模块(Atrous Spatial Pyramid Pooling, ASPP)中引入条形池化模块(Strip Pooling Module, SPM)^[21]和金字塔池化模块(Pyramid Pooling Module, PPM)^[22],有效地捕获裂缝结构的上下文信息,防止无关区域的污染信息干扰预测,提高裂缝的分割精度;最后,引入通道空间注意力机制(Convolutional Block Attention Module, CBAM)^[23],使神经网络更加关注图像中对锈蚀分类起决定作用的像素区域,而忽略无关的区域,增强锈蚀的特征表达能力,从而构建一个具有良好检测精度的轻量化钢结构表面锈蚀检测模型,助力钢结构表面锈蚀智慧监测发展。

1 DeepLabV3+模型

DeepLabV3+ 模型由编码器(Encoder)和解码器(Decoder)组合而成。网络结构如图 1 所示,在 Encoder 部分,输入经过 Xception 主干网络进行图片特征提取后得到两个输出:一个是浅层特征,直接提供给 Decoder;另一个是深层特征,通过 ASPP 模块的 4 个空洞卷积块和一个全局平均池化块处理得

到 5 组特征,这 5 个输出经过堆叠拼接和 1×1 卷积后送入 Decoder。在 Decoder 部分,浅层特征通过 1×1 卷积调整通道维度,深层特征进行 4 倍线性插值上采样,得到与浅层特征尺寸相同的特征,然后将这两个特征进行堆叠拼接、 3×3 卷积块处理和 4 倍线性插值上采样,得到和原图分辨率大小一样的预测图。

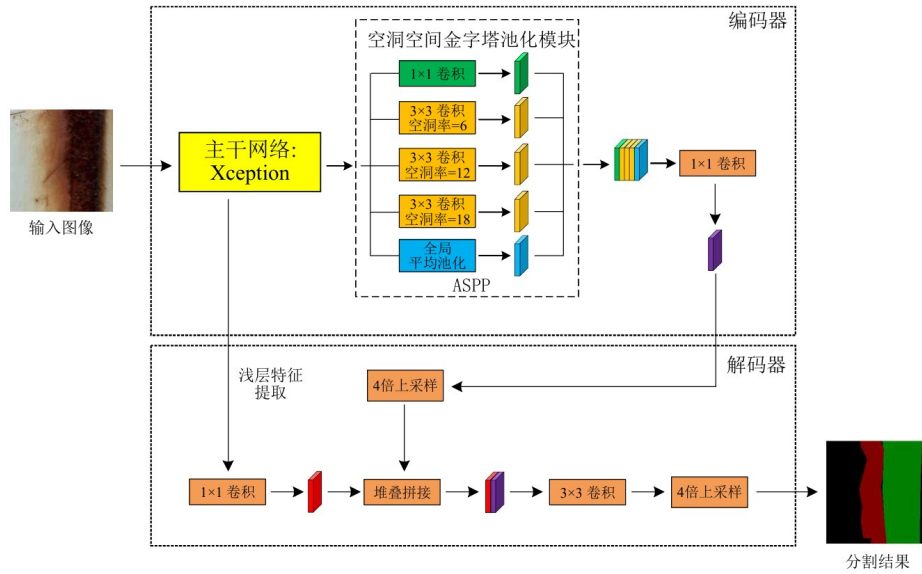


图 1 DeepLabV3+网络结构

Fig.1 DeepLabV3+ network structure

采用 DeepLabV3+ 模型作为基准模型是因为该模型具有高效的空洞卷积和多尺度信息融合能力,在不增加参数数量的前提下扩大锈蚀图片感受野,有效改善了传统卷积操作在捕捉长距离依赖关系时的局限性。在钢结构表面锈蚀病害的识别中,将锈蚀图片底层特征和高层特征进行融合,从而获得上下文信息,能够确保钢结构表面锈蚀分割结果的精细度和准确性。

2 HL-DeepLabV3+模型

DeepLabV3+ 模型的主干网络 Xception 网络结构较为复杂,网络层数多,模型参数量大,难以部署到嵌入式系统^[24],模型还需要捕获更加广泛的上下文信息,防止无关信息干扰预测,提升锈蚀分割结果的精细度和准确性,为了解决这些问题,提出用于钢结构表面锈蚀损伤检测的改进语义分割模型 HL-DeepLabV3+,具体网络结构如图 2 所示。

HL-DeepLabV3+ 模型 Encoder 部分,输入经过 MobileNetV3 主干网络进行图片特征提取后得到两个输出:一个是浅层特征,直接提供给 Decoder;另一个是深层特征,通过 CBAM 注意力机制捕获特征关键信息后送入 ASPP 模块,得到 5 个输出,其中 1

个条形池化串联 1 个金字塔池化,4 个不同膨胀率的空洞卷积,这 5 个输出经过堆叠拼接和 1×1 卷积后送入 Decoder。在 Decoder 部分,浅层特征通过 1×1 卷积调整通道维度,深层特征进行 4 倍线性插值上采样,得到与浅层特征尺寸相同的特征,然后将这两个特征进行堆叠拼接、两次 3×3 卷积处理和 4 倍线性插值上采样,得到和原图分辨率大小一样的预测图。

提出的 HL-DeepLabV3+ 网络,分别从轻量级主干网络 MobileNetV3、条形池化模块、金字塔池化模块和通道空间注意力机制等 4 方面进行结构改进,下面分别进行描述。

2.1 主干网络 MobileNetV3

原模型主干网络 Xception 结构复杂,网络层数多,训练和推理需要更多的计算资源,加上模型参数量大,需要更多的存储空间,难以部署到资源有限的设备上。MobileNetV3 是一种轻量化卷积神经网络,继承了 MobileNetV1 的深度可分离卷积和 MobileNetV2 的线性瓶颈的残差结构,显著提升了计算效率,并引入了 SE 通道注意力结构,提升了特征重用率和通道重要性的自适应调整,从而更好地捕获图像中的重要信息。在将 DeepLabV3+ 模型

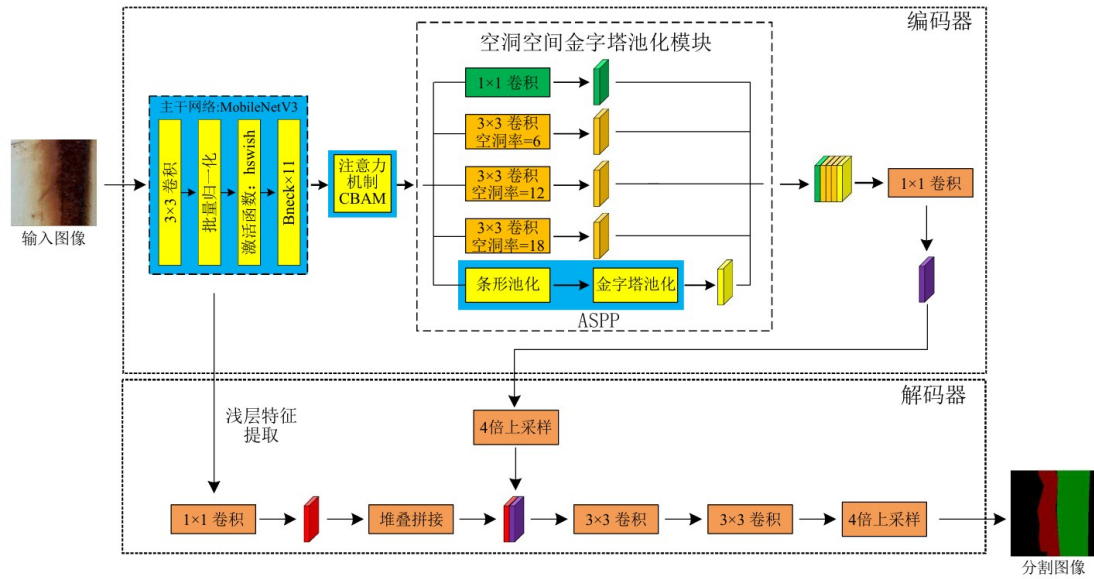


图2 HL-DeepLabV3+网络结构

Fig.2 HL-DeepLabV3+ network structure

的主干网络 Xception 替换为 MobileNetV3 后,网络层数变少、结构简单化、体积更小,模型推理时间缩短、参数和计算量减小,其性能和速度都表现得更加优异,更适用于存储空间和功耗受限的场景。整体网络结构如图 3 所示。

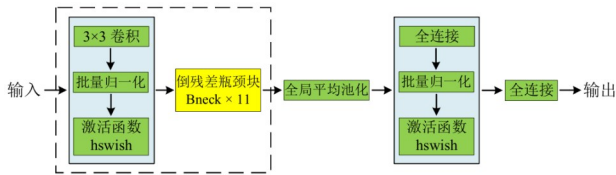


图3 MobileNetV3网络结构

Fig.3 MobileNetV3 network structure

其核心模块 Bneck,主要实现了通道可分离卷积,SE 通道注意力机制和残差连接,其结构如图 4 所示。

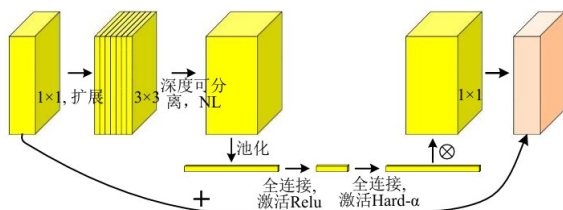


图4 Bneck网络结构

Fig.4 Bneck network structure

Encoder 部分中主干网络采用 MobileNetV3 的前 16 层,在经过 15 个 Bneck 模块处理后送入 ASPP 模块,已满足浅层特征提取,因此不再进行后续的池化、全连接处理,进一步减少计算量。所用到的部分 MobileNetV3 网络结构如表 1 所示。

表中,Input 表示输入,Operator 表示网络层,

表1 部分 MobileNetV3 网络结构

Table 1 Part of the MobileNetV3 network architecture

Input	Operator	Exp size	Out	SE	NL	S
$512^2 \times 3$	Conv2d		16		HS	2
$256^2 \times 16$	Bneck, 3×3	16	16		RE	1
$256^2 \times 16$	Bneck, 3×3	64	24		RE	2
$128^2 \times 24$	Bneck, 3×3	72	24		RE	1
$128^2 \times 24$	Bneck, 5×5	72	40	✓	RE	2
$64^2 \times 40$	Bneck, 5×5	120	40	✓	RE	1
$64^2 \times 40$	Bneck, 5×5	120	40	✓	RE	1
$64^2 \times 40$	Bneck, 3×3	240	80		HS	2
$32^2 \times 80$	Bneck, 3×3	200	80		HS	1
$32^2 \times 80$	Bneck, 3×3	184	80		HS	1
$32^2 \times 80$	Bneck, 3×3	184	80		HS	1
$32^2 \times 80$	Bneck, 3×3	480	112	✓	HS	1
$32^2 \times 112$	Bneck, 3×3	672	112	✓	HS	1
$32^2 \times 112$	Bneck, 5×5	672	160	✓	HS	2
$16^2 \times 160$	Bneck, 5×5	960	160	✓	HS	1
$16^2 \times 160$	Bneck, 5×5	960	160	✓	HS	1

Exp size 表示每一个 Bneck 块中第一层 1×1 conv2d 卷积核的个数, Out 表示输出特征通道数, SE 代表是否使用通道注意力机制, NL 表示激活函数的类型, 包括 HS (h-swish)、RE (ReLU), S 代表 stride 步长。

2.2 引入条形池化模块

池化是将一个区域中的信息压缩成一个值,完成信息的抽象。为了更有效地捕捉长距离依赖关系,基于空间池化层扩展卷积神经网络的感受野并获取上下文信息,提出了一种新的方法 SPM, SPM 的网络结构如图 5 所示,首先输入一个特征图,输入的特征图经过卷积核大小为 $(H, 1)$ 和 $(1, W)$, 按照

式(1)、(2)求平均池化的方法,分别进行竖直和水平条形池化后变为 $H \times 1$ 和 $1 \times W$,得到 y_i^h 和 y_j^v ,并以该值作为池化输出值,随后经过卷积对两个输出特征分别沿着左右和上下进行扩充,扩充后两个特征图尺寸相同,这两个尺寸相同扩充后的特征图对应相同位置进行逐像素求和得到 $H \times W$ 大小新的特征图,最后通过 1×1 的卷积与 sigmoid 激活,再将处理后与原输入图对应像素相乘得到了最终的输出结果。

出结果。

$$y_i^h = \frac{1}{W} \sum_{0 \leq j < W} x_{i,j} \quad (1)$$

$$y_j^v = \frac{1}{H} \sum_{0 \leq i < H} x_{i,j} \quad (2)$$

式中: y_i^h 为平均池化后输出的宽的特征张量, y_j^v 为平均池化后输出的高的特征张量, x 为输入的二维特征张量, W 为输入特征图的宽, H 为输入特征图的长。

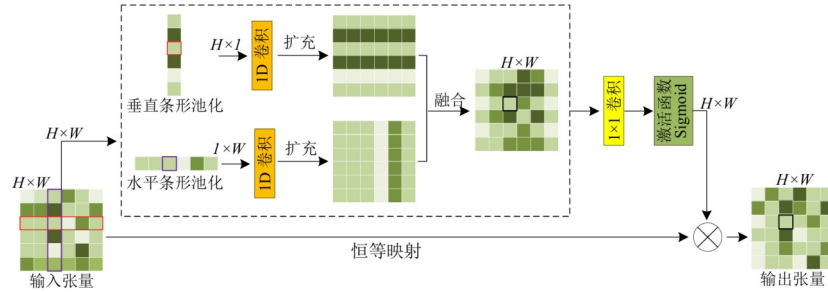


图5 SPM 网络结构

Fig.5 SPM network structure

DeepLabV3+ 网络中 ASPP 分支 5 使用的池化方式为最大值池化,该池化使用方形框对输入的锈蚀特征进行池化,获取了大量的锈蚀区域和非锈蚀区域的特征信息,但无法捕获孤立区域的长距离依赖关系,其池化形式如图 6(a) 所示。而 HL-DeepLabV3+ 网络中引入 SPM,它沿着一个空间维度部署成一个长条的池化形状,沿其他空间维度保持成一个狭窄的池化形状,使网络能够捕获孤立锈蚀区域的长距离依赖关系,将锈蚀区域和背景的特征信息区别开来,以消除背景区域干扰,其池化形式如图 6(b) 所示。因此,HL-DeepLabV3+ 网络 ASPP 中将最大值池化调整成条形池化,在钢梁、钢柱、钢桁架等条形构件表面锈蚀损伤识别中,能更有效地捕获锈蚀结构的上下文信息,更好地分割出钢结构表面锈蚀,分割结果如图 6(c) 所示。

2.3 引入金字塔池化模块

PPM 最先出现在 PSPNet 网络中,针对 FCN 网络在场景分析数据上存在的错误匹配,小目标的丢失问题,该网络做出一系列的改进,提升场景分析中对于相似颜色、形状的物体的检测精度,该网络关键的结构就是 PPM,其网络结构如图 7 所示,首先对原始特征图进行不同尺度的池化,得到 4 个不同尺寸的特征图,对得到的特征图进行上采样操作,将 4 个特征图恢复至原始相同的尺寸大小,最后在通道维度上进行 Concat 拼接,得到最终的复合特征图。

HL-DeepLabV3+ 网络利用 PPM 进一步聚合

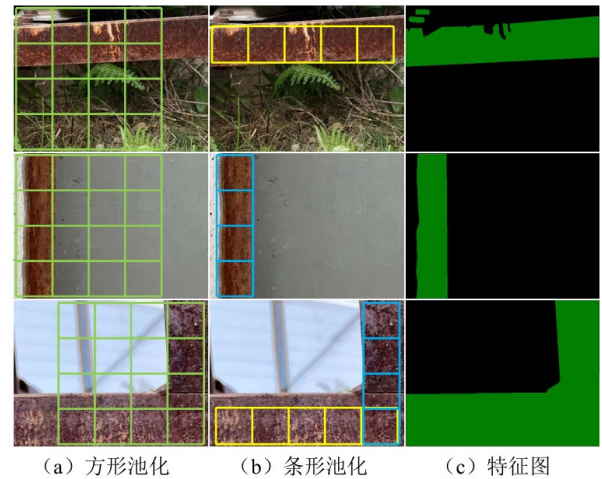


图6 传统池化和条形池化对比

Fig.6 Comparison between traditional pooling and strip pooling

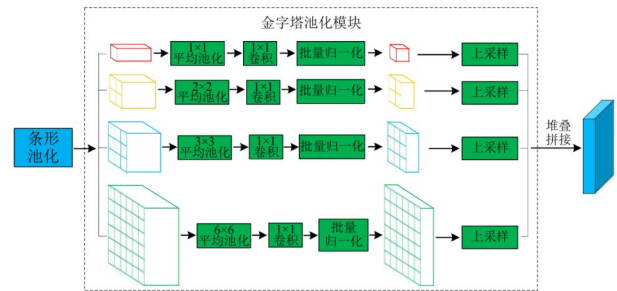


图7 PPM 网络结构

Fig.7 PPM network structure

不同区域的全局上下文信息,通过对不同层级的锈蚀图像进行池化操作,使网络能够处理复杂钢结构表面锈蚀,并对每个像素进行精确地分类预测,有

效提高了锈蚀场景解析的准确性,通过消融实验证明,这种串并联相结合的模式有效提升了 HL-DeepLabV3+网络模型锈蚀识别的精确率。

2.4 通道空间注意力机制

CBAM 模块是为了提升前馈卷积神经网络性能而提出的一种简单而有效的注意力模块,由通道

注意力模块和空间注意力模块构成,首先将在通道注意力模块得到的通道注意力权重乘以输入特征图,再将在空间注意力模块得到的空间注意力权重乘以通过通道注意力得到的特征图,最终得到输出结果,网络结构如图 8 所示。

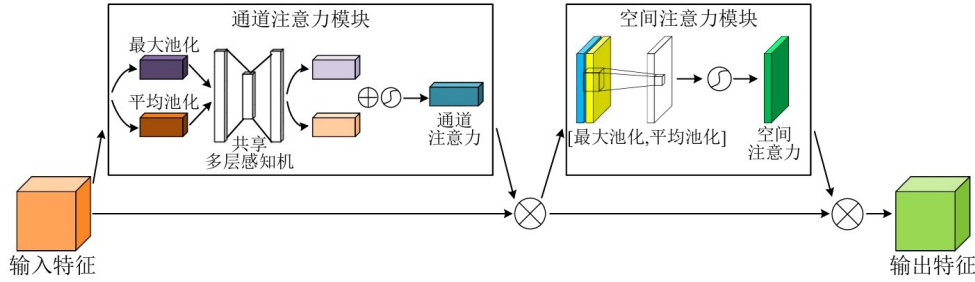


图 8 CBAM 网络结构

Fig.8 CBAM network structure

在通道注意力模块中,对输入进来的特征进行最大值池化和平均池化操作,将每个通道在空间维

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \quad (3)$$

式中: F 为输入特征图; $M_c(F)$ 为输出通道注意力值; σ 为 Sigmoid 激活函数; W_0 和 W_1 为权重参数; F_{avg}^c 为在单个通道的平均池化特征; F_{max}^c 为在单个通道的最大池化特征。

在空间注意力模块中,对输入特征图沿通道轴应用平均池化和最大池化,将输入的通道维度压缩为 1,从而将全局通道信息整合为一个单一的通道特征图,其计算过程如式(4)所示。

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\ = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (4)$$

式中: $M_s(F)$ 为输出的空间注意力值; $f^{7 \times 7}$ 为卷积大小为 7×7 的卷积操作; F_{avg}^s 为沿通道轴的平均池化特征; F_{max}^s 为沿通道轴的最大值池化特征。

HL-DeepLabV3+ 网络将 CBAM 模块设置在编码部分主干网络和改进后的 ASPP 模块中间,首先通过通道注意力,提高通道对钢材锈蚀的关注度,然后通过空间注意力,进一步加强锈蚀区域边

度上进行压缩,从而实现对全局空间信息的压缩和提取,其计算过程如式(3)所示。

缘的识别能力,二者相互结合,使 CBAM 能够自适应地调整锈蚀特征图中每个通道和空间位置的重要性,从而提高模型的表征能力和决策准确性。

3 实验与结果

3.1 数据集

实验数据集有 TraData 和 VerData。TraData 数据集采集自某工业钢结构厂房现场,采用 iPhone13 手机采集图像,共采集 700 张像素为 $1\,702 \times 1\,276$ 的锈蚀照片,并将其切割成 512×512 像素的子图像,最终得到 3 510 张图像,按 9:1 比例随机划分得训练集 3 160 张,验证集 350 张,以测试模型识别效果的有效性。VerData 数据集采集自某高校内五座小型中承式钢拱桥现场,采用相机和其他型号智能手机采集图像,共采集 200 张,并将其细分切割,最终得到 1 200 张图像,按 9:1 比例随机划分得训练集 1 080 张,验证集 120 张,以测试模型识别效果的通用性。详细信息如表 2 所示。

表 2 数据集信息

Table 2 Data set information

数据集	采集设备	型号	原始图像数量/张	原始图像尺寸/像素	裁剪后图像数量/张	裁剪后图像尺寸/像素
TraData	智能手机	iPhone13	700	$1\,702 \times 1\,276$	3510	512×512
	相机	尼康 Z30	50	$5\,568 \times 3\,712$	300	694×464
VerData	智能手机	OPPO Find X50	50	$4\,096 \times 2\,303$	300	409×460
	智能手机	Vivo X27	50	$4\,000 \times 1\,856$	300	500×464
	智能手机	荣耀 X40	50	$4\,096 \times 3\,072$	300	512×614

3.2 锈蚀等级分类标准

在钢结构锈蚀识别研究中,只有明确结构的锈蚀程度,才能在后续管养中采取合适的锈蚀处理措施,目前中国现行钢结构锈蚀等级分类标准^[25]是将钢材锈蚀等级分为4个等级,分别以A、B、C、D表示。A:大面积覆盖着氧化皮而几乎没有铁锈的钢材表面;B:已发生锈蚀,并且氧化皮已开始剥落的钢材表面;C:氧化皮已因锈蚀而剥落,或者可以刮除,并且在正常视力观察下可见到轻微点蚀的钢材表面;D:氧化皮已因锈蚀而剥离,并且在正常视力观察下可见普遍发生点蚀的钢材表面。实验图像TraData和VerData数据集基于中国现行规范提出了具体的钢结构表面锈蚀分类标准,如表3所示,该标准在划分锈蚀类别时考虑了锈蚀颜色的影响。

表3 钢结构表面锈蚀等级分类标准

Table 3 Standard for classification of surface corrosion of steel structures

锈蚀等级	判断标准
无锈蚀	表面干净,没有锈蚀
浅层锈蚀	表面有少量锈迹、锈斑较小且分散,锈蚀的颜色浅,不影响结构的整体性能
深层锈蚀	涂层剥落,表面可见普遍锈迹,锈蚀的颜色深,可能需要进行部分处理

注:锈蚀等级难以界定时,按更严重一级判定。

在标注软件Labelme对图像进行语义分割标注中,浅层锈蚀的图像类别设定为light_corrosion,深层锈蚀的图像类别设定为deep_corrosion,无锈蚀的图像类别设定为background。部分锈蚀图像如图9所示。

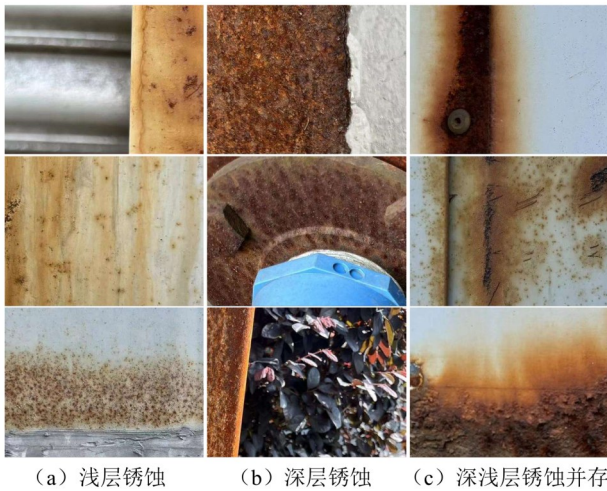


图9 部分锈蚀图

Fig.9 Partial corrosion diagram

3.3 实验配置及参数设置

实验环境和配置如下:硬件配置CPU为14th Gen Intel(R) Core(TM)i5-14600KF,显卡为

NVIDIA GeForce RTX4070 Ti SUPER,显存16 GB,运行内存64 GB。软件配置操作系统为Windows11(64位),编程语言为Python3.8.19,深度学习框架为Pytorch2.3.1,在编写代码的过程中使用了utils、torchvision、numpy等一系列依赖模块。

HL-DeepLabV3+网络模型设置训练迭代次数为100次,批量大小为16,初始学习率为0.0001,学习率下降方式选用余弦退火衰退,优化器使用Adam自适应调整学习率。损失函数使用戴斯损失函数(Dice Loss)。

3.4 评价指标

为评价网络模型在锈蚀识别中的性能,采用的算法评价指标有准确率(Accuracy)、平均交并比(mIoU)、平均像素准确率(MPA)、F1分数(F1_score)和参数量(Parameters)。

1)准确率(Accuracy)衡量的是模型在所有像素上的整体分类性能,即模型正确分类的像素数占总像素数的比例。它提供了一个总体的性能概述。计算公式如式(5)所示。

$$A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (5)$$

式中:A表示准确率, T_p 表示正确分类为正类的像素数, T_n 表示正确分类为负类的像素数, F_p 表示错误分类为正类的像素数, F_n 表示错误分类为负类的像素数。

2)交并比(IoU)衡量的是预测结果与真实标签的重叠程度。平均交并比(mIoU)是对所有类别的IoU取平均值,反映了模型在不同类别上的分割质量。mIoU是语义分割中最常用的评价指标之一,它能够较好地反映模型在各个类别上的表现,计算式如式(6)、(7)所示。

$$I = \frac{T_p}{T_p + F_p + F_n} \quad (6)$$

$$m = \frac{1}{N} \sum_{i=1}^N I_i \quad (7)$$

式中:I表示交并比,m表示平均交并比,N为类别的数量。

3)MPA是针对每个类别计算的全像素精度的平均值。计算式如式(8)所示。

$$M = \frac{1}{N} \sum_{i=1}^N \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (8)$$

式中:M表示平均像素准确率。

4)F1_score是精确度和召回率的调和平均值,它在精度和召回率之间取得平衡,高F1_score意味着模型在正类预测的准确性和覆盖率上都表现良好。计算式如式(9)所示。

$$F=\frac{2\cdot P_r\cdot R_e}{P_r+R_e}\tag{9}$$

式中： F 表示F1分数， P_r 表示精确度， R_e 表示召回率。

3.5 实验结果分析

3.5.1 不同主干网络对比实验

根据钢结构表面锈蚀损伤检测的研究情况,选

取了深度卷积神经网络架构的Xception、残差神经网络ResNet101、轻量化神经网络MobileNetV2和MobileNetV3四种不同的网络作为DeepLabV3+模型的主干网络,对比分析在不同主干网络(Xception、ResNet101、MobileNetV2、MobileNetV3)下,DeepLabV3+模型在TraData数据集锈蚀识别中的表现,其结果如表4所示。

表 4 DeepLabv3+模型不同主干网络的实验结果

Table 4 Experimental results of different backbone networks of DeepLabv3+ model

主干网络	Accuracy/%	mIoU/%	MPA/%	F1_score/%	Parameters/M
Xception	95.61	84.33	91.04	91.08	37.87
ResNet101	97.69	90.96	95.59	95.12	42.50
MobileNetV2	97.17	88.78	94.36	93.80	2.23
MobileNetV3	97.77	91.88	94.95	95.67	2.97

从实验结果可以看出,MobileNetV3主干网络在准确率、平均交并比和F1分数上均为最优,且参数量只有2.97 M,在Accuracy上分别比Xception、ResNet101、MobileNetV2高2.16%、0.08%、0.60%;在mIoU上分别比Xception、ResNet101、MobileNetV2高7.55%、0.92%、3.10%;在F1分数上分别比Xception、ResNet101、MobileNetV2高4.59%、0.55%、1.87%;在MPA上分别比Xception、MobileNetV2高3.91%、0.59%,比ResNet101低0.64%,虽然ResNet101主干网络在平均全像素精度上表现最优,但其参数量远高于MobileNet系列网络,约为MobileNet系列网络的14倍,难以满足工业现场的实时性要求。因此,采用轻量级的MobileNetV3作为HL-DeepLabV3+模型

的主干网络,在确保图像特征提取能力的同时,有效地降低了模型参数量,提高了算法的实用性。

3.5.2 注意力机制选取及存放位置对比实验

基于MobileNetV3主干网络、SPM模块和PPM模块,讨论HL-DeepLabV3+网络设置不同注意力机制以及注意力机制存放的位置对钢结构表面锈蚀检测性能的影响,比较SE、ECA、CBAM、SIMAM四种不同注意力机制分别设置在主干网络Mobilenetv3提取特征之后、Encoder部分1×1卷积之前、主干网络浅层特征提取之后和Decoder部分4倍上采样之前4处位置,HL-DeepLabV3+网络在TraData数据集中的锈蚀检测性能,实验结果如表5所示。

表 5 不同注意力机制在不同位置处网络的实验结果

Table 5 Experimental results of different attention mechanisms at different locations

位置	注意力机制	Accuracy/%	mIoU/%	MPA/%	F1_score/%	Parameters/M
After MobileNetV3		97.99	93.85	97.43	96.80	18.08
	SE	97.98	93.96	97.27	96.86	18.20
	ECA	97.96	93.69	96.90	96.71	18.08
	CBAM	98.15	94.44	97.83	97.23	18.20
	SIMAM	97.80	92.39	95.17	95.97	18.08
Before 1×1Conv	SE	97.89	92.51	95.42	96.03	18.10
	ECA	97.99	93.20	96.00	96.43	18.08
	CBAM	97.98	93.77	97.11	96.75	18.09
	SIMAM	97.98	93.83	97.01	96.79	18.08
After Low-level features	SE	97.68	91.20	94.23	95.26	18.10
	ECA	98.02	94.04	97.54	96.90	18.08
	CBAM	98.00	93.65	97.22	96.68	18.09
	SIMAM	98.01	93.28	95.84	96.47	18.08
Before Upsample By 4	SE	97.91	93.70	96.72	96.72	18.08
	ECA	98.03	94.05	97.18	96.90	18.08
	CBAM	97.95	93.84	97.16	96.80	18.08
	SIMAM	98.02	94.03	97.33	96.90	18.08

从实验结果可以看出,引入注意力机制在改变少量参数情况下,可以不同程度地改变模型的检测精度,模型性能的提升与注意力机制的类型、具体添加的位置和方法等因素都有很大关系。针对钢结构锈蚀数据的检测,选择了表现最佳的 CBAM 注意力机制,存放位置在主干网络 Mobilenetv3 提取特征之后处,所构建的 HL-DeepLabV3+ 锈蚀识别网络在 Accuracy、mIoU、MPA、F1_score 分别达到了

98.15%、94.44%、97.83%、97.23%,相比未引入注意力机制,在 Accuracy、mIoU、MPA、F1_score 分别提高了 0.16%、0.59%、0.40%、0.43%。

3.5.3 不同语义分割算法的对比实验

将搭建的 HL-DeepLabV3+ 网络和 UNet、FCN、SegNet、DeepLabV3+ 等四种主流语义分割网络分别放在 TraData 数据集上进行钢结构表面锈蚀训练测试,其实验结果如表 6 所示。

表 6 不同语义分割算法的实验结果

Table 6 Experimental results of different semantic segmentation algorithms

网络模型	Accuracy/%	mIoU/%	MPA/%	F1_score/%	Parameters/M
UNet(vgg16)	97.86	90.79	94.97	94.98	24.89
FCN(vgg16)	96.74	87.42	94.26	92.99	33.61
SegNet(vgg16)	97.76	91.03	95.22	95.14	29.44
DeepLabV3+(Xception)	95.61	84.33	91.04	91.08	54.70
HL-DeepLabV3+	98.15	94.44	97.83	97.23	18.20

从实验结果可以看出,HL-DeepLabV3+ 网络模型在进行钢结构表面锈蚀分割时,参数量最少,且 HL-DeepLabV3+ 模型的准确率、平均交并比、平均全像素精度和 F1 分数均优于其他 4 种模型,在 Accuracy 上分别比 UNet、FCN、SegNet、DeepLabV3+ 高 0.29%、1.41%、0.39%、2.54%;在 mIoU 上分别比 UNet、FCN、SegNet、DeepLabV3+ 高 3.65%、7.02%、3.41%、10.11%;在 MPA 上分别比 UNet、FCN、SegNet、DeepLabV3+ 高 2.86%、3.57%、2.61%、6.79%;在 F1_score 上分别比 UNet、FCN、SegNet、DeepLabV3+ 高 2.25%、4.24%、2.09%、6.15%。通过对比以上五种网络,HL-DeepLabV3+ 网络在实现模型轻量化的同时还提升了对锈蚀损伤的分割精度,验证了改进方法的有效性。

同时给出了以上 5 种网络模型检测钢结构表面锈蚀的可视化效果对比,随机选取 6 张图片进行可视化对比,可视化结果如图 10 所示。第 1 行为 6 张原始图片,第 2 行为原始图片分割标签,从锈蚀预测可视化效果图中可以看出,UNet 网络出现了对图像

边缘的锈蚀分割不精确、多分割和较少误分割的情况,FCN 网络出现了边缘分割锯齿状、较多的漏分割和误分割的情况,SegNet 网络在锈蚀预测上主要是深浅层锈蚀边界分割不精确、少量误分割和漏分割的情况,DeepLabV3+ 网络出现了边界分割不精确、误分割和严重漏分割的情况。而提出的 HL-DeepLabV3+ 网络,在锈蚀预测上仅有极少量的噪声干扰,能更好地提取出锈蚀的特征,且对锈蚀的分割更为准确、连续、识别错误率更低。

3.5.4 消融实验

为了验证 HL-DeepLabV3+ 检测算法的时效性与准确性,选择与 DeepLabV3+ 模型在 TraData 数据集上进行测试比较。分析使用 MobileNetV3、SPM、PPM、CBAM 对模型产生的影响,验证改进的锈蚀识别检测模型的有效性。其实验结果如表 7 所示。

从实验结果可见,采用轻量级 MobileNetV3 主干网络相较于原主干网络 Xception,模型的参数量降低了 92.2%,在 Accuracy、mIoU、MPA、F1 分数上分别提高了 2.16%、7.55%、3.91%、4.59%;在

表 7 消融实验结果

Table 7 Results of ablation experiment

No.	Mobilenetv3	SPM	PPM	CBAM	Accuracy/%	mIoU/%	MPA/%	F1_score/%	Parameters/M
①	×	×	×	×	95.61	84.33	91.04	91.08	37.87
②	√	×	×	×	97.77	91.88	94.95	95.67	2.97
③	√	√	×	×	97.85	93.77	97.31	96.76	15.59
④	√	×	√	×	97.91	93.73	97.33	96.73	14.22
⑤	√	×	×	√	97.91	93.54	97.38	96.63	11.84
⑥	√	√	√	×	97.99	93.85	97.43	96.80	18.20
⑦	√	√	√	√	98.15	94.44	97.83	97.23	18.20

注:√表示引入模块;×表示未引入模块。

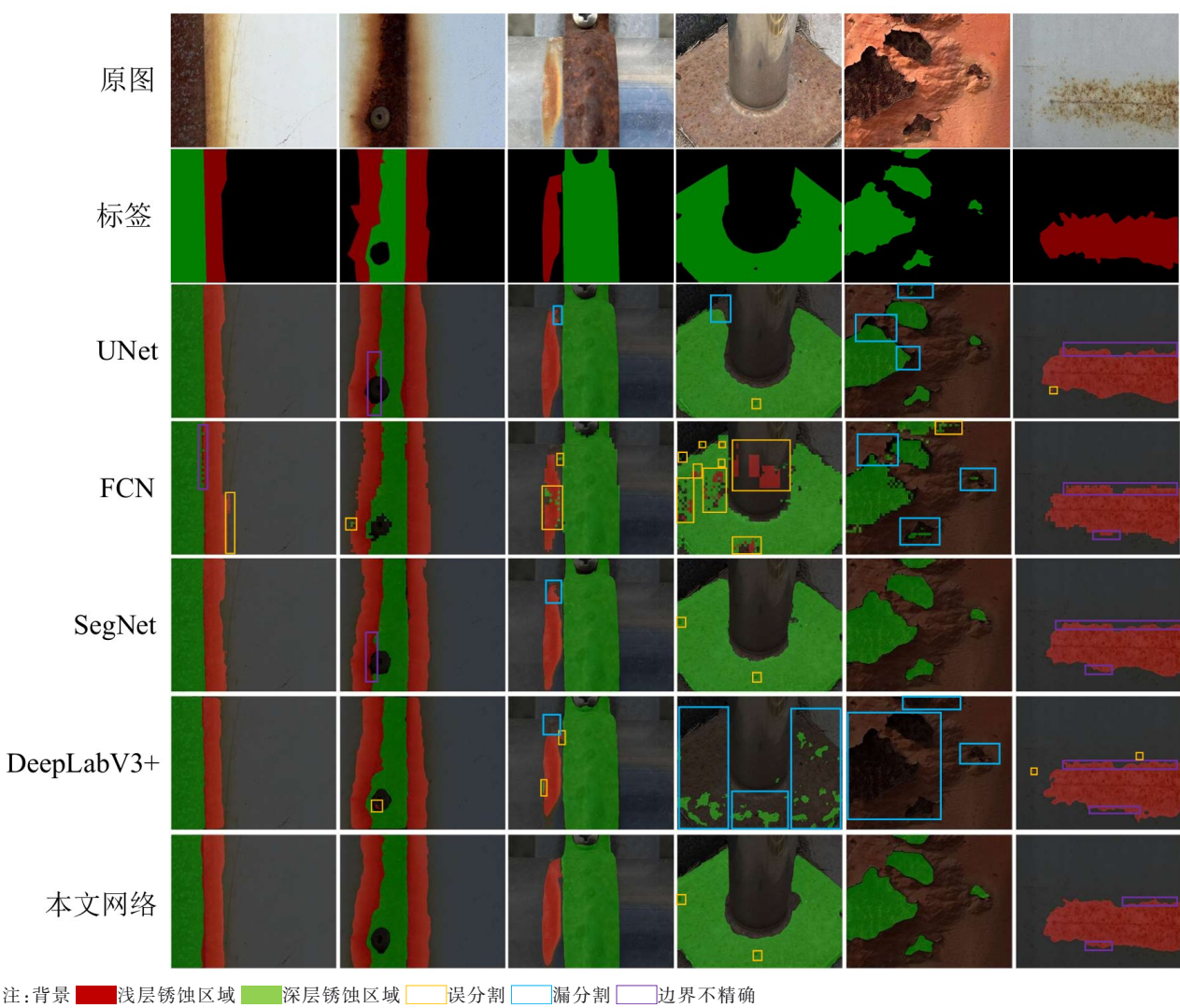


图 10 可视化效果对比

Fig. 10 Visualization effect comparison

此基础上引入 SPM 模块,网络在 Accuracy、mIoU、MPA、F1 分数上分别提高了 0.08%、1.89%、2.36%、1.09%;再引入 PPM 模块,网络在 Accuracy、mIoU、MPA、F1 分数上分别提高了 0.14%、0.08%、0.12%、0.04%;最后设置注意力机制 CBAM 模块,在 Accuracy、mIoU、MPA、F1 分数上分别提高了 0.16%、0.59%、0.40%、0.43%。验证了替换轻量级 MobileNetV3 主干网络、ASPP 模块分支 5 引入 SPM 模块和 PPM 模块、设置

CBAM 注意力机制 4 方面改进设计均对网络性能有贡献。说明提出的 HL-DeepLabV3+ 网络在钢结构表面锈蚀损伤检测中能够取得良好效果。

3.5.5 泛化实验

为进一步验证 HL-DeepLabV3+ 网络模型对不同场景相似锈蚀识别能力的泛化性能,选择在不同的拍摄设备采集的不同分辨率图像的 VerData 数据集上进行验证,其实验结果如表 8 所示。

表 8 HL-DeepLabV3+ 在 VerData 数据集中的实验结果

Table 8 Experimental results of HL-DeepLabV3+ in VerData dataset

网络模型	Accuracy/%	mIoU/%	MPA/%	F1_score/%	Parameters/M
HL- DeepLabV3+	98.08	94.27	97.61	97.08	18.20

从实验结果可以得出,HL-DeepLabV3+ 在 VerData 数据集上得到评价网络的指标值均较高,其中 Accuracy 达到了 98.08%、mIoU 达到了 94.27%、MPA 达到了 97.61%、F1 分数达到了 97.08%,证明 HL-DeepLabV3+ 网络在不同设备及

不同分辨率图片类型数据集上对钢结构表面锈蚀识别任务中的识别结果具有可靠性,实验测试结果表明该网络模型具有良好的泛化性能。

4 结论

针对钢结构表面锈蚀损伤人工目测易出现视觉疲劳造成读数、纹理和颜色判断错误,经典深度学习钢结构锈蚀识别网络分割精度低、参数量大等问题,提出的 HL-DeepLabV3+ 轻量级的语义分割网络,实现了对钢结构表面锈蚀损伤的识别,通过多组测试实验,得出以下结论:

1) 通过将 DeepLabV3+ 进行轻量级 MobileNetV3 主干网络替换,模型参数量由 37.87 M 减小为 2.97 M,降低 92.2%;引入 SPM、PPM 模块后,增强了网络捕获孤立锈蚀区域的长距离依赖关系,减少了非锈蚀区域的干扰。引入 CBAM 注意力机制,使网络更关注对锈蚀分类重要的像素区域,增强了锈蚀特征表达能力。得到的新模型比原模型在 Accuracy、mIoU、MPA、F1 分数上分别提高了 2.54%、10.11%、6.79% 和 6.15%,实验结果表明这四方面结构改进提升了网络性能。

2) HL-DeepLabV3+ 方法在准确率上达到 98.15%、平均交并比达到 94.44%、平均像素准确率达到 97.83%、F1 分数达到 97.23%,在各项评价指标上均优于经典语义分割模型 UNet、FCN、SegNet 和 DeepLabV3+,实现了钢结构表面损伤图像的精确分割,取得了良好效果。

参考文献

- [1] 钢结构推广应用和建筑领域用钢需求调查工作组. 中国钢结构建筑发展现状调研及未来思考建议[N]. 中国冶金报, 2024-08-31(002).
Steel Structure Promotion and Application and Construction Sector Steel Demand Research Working Group. Investigation on the development status and future recommendations for steel structure buildings in China [N]. China Metallurgical News, 2024-08-31(002). (in Chinese)
- [2] 秦荣杰. 基于深度学习 Transformer 网络的钢板锈蚀类别识别方法研究 [D]. 西安: 西安建筑科技大学, 2023.
QIN R J. Research on steel plate rust category identification method based on deep learning Transformer network [D]. Xi'an: Xi'an University of Architecture and Technology, 2023. (in Chinese)
- [3] 黄莺, 李嘉晨, 朱彦飞, 等. 考虑腐蚀作用的钢框架结构连续性倒塌分析[J]. 土木与环境工程学报(中英文), 2023, 45(6): 21-28.
HUANG Y, LI J C, ZHU Y F, et al. Progressive collapse analysis of steel frame structures with corrosion effect [J]. Journal of Civil and Environmental Engineering, 2023, 45(6): 21-28. (in Chinese)
- [4] 钱偲偲. 土木工程智能化监测评估系统的理论应用[J]. 建筑结构, 2023, 53(11): 164.
QIAN S S. Theoretical application of intelligent monitoring and evaluation system for civil engineering [J]. Building Structure, 2023, 53(11): 164. (in Chinese)
- [5] 成孟腾. 基于机器视觉的水工钢结构表面锈蚀区域分割研究 [D]. 湖北 宜昌: 三峡大学, 2022.
CHENG M T. Research on corrosion region segmentation of hydraulic steel structural surface based on machine vision [D]. Yichang, Hubei: China Three Gorges University, 2022. (in Chinese)
- [6] 倪有豪, 陆欢, 季超, 等. 基于语义分割的桥梁锈蚀病害识别对比分析[J]. 东南大学学报(自然科学版), 2023, 53(2): 201-209.
NI Y H, LU H, JI C, et al. Comparative analysis on bridge corrosion damage detection based on semantic segmentation [J]. Journal of Southeast University (Natural Science Edition), 2023, 53(2): 201-209. (in Chinese)
- [7] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy. IEEE, 2017: 2980-2988.
- [8] YU L P, VIJAY M M, SUNIL J, et al. Hybrid deep learning model based smart IOT based monitoring system for Covid-19 [J]. Heliyon, 2023, 9(11): e21150.
- [9] CHEN X Y, MA Q H, HE Z Z, et al. Real-time detection and localization method for weld seam of narrow butt joint based on semantic segmentation [J]. Measurement Science and Technology, 2024, 35(3): 035205.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [11] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 7-12, 2015, Boston, MA, USA. IEEE, 2015: 3431-3440.
- [12] LIU J J, HOU Q B, CHENG M M, et al. Improving convolutional networks with self-calibrated convolutions [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 10093-10102.
- [13] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.

- [14] 陈法法, 董海飞, 何向阳, 等. 改进 U_Net 网络的钢结构表面锈蚀图像分割方法[J]. 电子测量与仪器学报, 2024, 38(2): 49-57.
CHEN F F, DONG H F, HE X Y, et al. Improved steel structure surface rust image segmentation method for U_Net network [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(2): 49-57. (in Chinese)
- [15] GARCÍA PEÑA D, GARCÍA PÉREZ D, DÍAZ BLANCO I, et al. Exploring deep fully convolutional neural networks for surface defect detection in complex geometries [J]. The International Journal of Advanced Manufacturing Technology, 2024, 134(1): 97-111.
- [16] 许颖, 张天瑞, 金淦. 基于深度学习 SCNet 的钢筋混凝土锈蚀裂缝识别[J]. 湖南大学学报(自然科学版), 2022, 49(3): 101-110.
XU Y, ZHANG T R, JIN G. Identification of corroded cracks in reinforced concrete based on deep learning SCNet model [J]. Journal of Hunan University (Natural Sciences), 2022, 49(3): 101-110. (in Chinese)
- [17] 王凌云, 李婷宜, 李阳, 等. 基于 FEF-DeepLabV³⁺ 的电力金具锈蚀分割方法[J]. 电子测量与仪器学报, 2023, 37(7): 166-176.
WANG L Y, LI T Y, LI Y, et al. Segmentation method of power armor clamp corrosion based on FEF-DeepLabV³⁺ [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(7): 166-176. (in Chinese)
- [18] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27-November 2, 2019. Seoul, Korea. IEEE, 2019: 1314-1324.
- [19] 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. ACM, 2018: 833-851.
- [20] CHOLLET F. Xception: Deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 1800-1807.
- [21] HOU Q B, ZHANG L, CHENG M M, et al. Strip pooling: Rethinking spatial pooling for scene parsing [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020. Seattle, WA, USA. IEEE, 2020: 4002-4011.
- [22] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6230-6239.
- [23] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//Computer Vision-ECCV 2018. ACM, 2018: 3-19.
- [24] 杨萍, 张汐. 改进 DeepLabv3+ 的道路表面裂缝检测方法[J/OL]. 计算机工程. <https://doi.org/10.19678/j.issn.1000-3428.0069114> YANG P, ZHANG X, et al. Improved DeepLabv3+ road surface crack detection method [J/OL]. Computer Engineering. <https://doi.org/10.19678/j.issn.1000-3428.0069114>. (in Chinese)
- [25] 涂覆涂料前钢材表面处理 表面清洁度的目视评定 第 1 部分: 未涂覆过的钢材表面和全面清除原有涂层后的钢材表面的锈蚀等级和处理等级: GB/T 8923.1—2011 [S]. 北京: 中国标准出版社, 2013.
Preparation of steel substrates before application of paints and related products—Visual assessment of surface cleanliness: Part 1: Rust grades and preparation grades of uncoated steel substrates and of steel substrates after overall removal of previous: GB/T 8923.1—2011 [S]. Beijing: Standards Press of China, 2013. (in Chinese)

(编辑 XXX)