

DOI: 10.13382/j.jemi.B2306296

# 基于 FEF-DeepLabV3+ 的电力金具锈蚀分割方法\*

王凌云<sup>1</sup> 李婷宜<sup>1</sup> 李 阳<sup>2</sup> 万旭东<sup>1</sup> 童华敏<sup>2</sup>

(1. 三峡大学电气与新能源学院 宜昌 443002; 2. 国网宜昌供电公司 宜昌 443000)

**摘要:**金具锈蚀在输电线路航拍图像中细节丰富且分布不规则,为克服分割检测过程中局部信息丢失、精度低和速度慢等问题,提出基于 DeepLabV3+ 的金具锈蚀语义分割模型。替换其主干网络为轻量化改进 MobileNetV3 网络加快运算速度,并提出自适应特征金字塔(adaptive feature pyramid, AFP)结构融合多尺度。结合 FRN 层提出特征融合空洞空间金字塔池化(feature fusion atrous spatial pyramid pooling, FEF-ASPP)结构,能够在加强像素间联系的同时不降低分辨率。最后优化损失函数,提高算子的有效性。实验表明, mIoU 和 mPA 分别达到了 87.15%、96.64%, 相较于原模型提高了 3.09%、4.29%。参数量仅为原模型的 48%, 推理时间仅为 15.94 ms, 降低了对设备算力的要求, 实现高效高精度、轻量化的输电设备锈蚀缺陷分割检测。

**关键词:**深度学习; 架空输电线路巡检; 图像语义分割; 缺陷检测; 像素分类

**中图分类号:** TM755; TP391 **文献标识码:** A **国家标准学科分类代码:** 470.40

## Segmentation method of power armor clamp corrosion based on FEF-DeepLabV3+

Wang Lingyun<sup>1</sup> Li Tingyi<sup>1</sup> Li Yang<sup>2</sup> Wan Xudong<sup>1</sup> Tong Huamin<sup>2</sup>

(1. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China;

2. State Grid Yichang Power Supply Company, Yichang 443000, China)

**Abstract:** The proportion of armor clamp rust in aerial images of power transmission lines is rich in details and irregularly distributed. To overcome problems such as local information loss, low accuracy, and slow speed in the segmentation detection process, a DeepLabV3+-based semantic segmentation model for armor clamp rust is proposed. The backbone network is replaced with a lightweight improved MobileNetV3 network to speed up computation, and an adaptive feature pyramid (AFP) structure is proposed to merge multiple scales. A feature fusion atrous spatial pyramid pooling (FEF-ASPP) structure is proposed, combined with the FRN layer to strengthen pixel relationships without reducing resolution. Finally, the loss function is optimized to improve the effectiveness of the operator. Experiments show that the mIoU and mPA reach 87.15% and 96.64%, respectively, which is an improvement of 3.09% and 4.29% compared to the original model. The parameter quantity is only 48% of the original model, and the inference time is only 15.94 ms, reducing the requirement for device computing power and achieving high-efficiency, high-precision, and lightweight segmentation detection of armor clamp rust in power transmission equipment.

**Keywords:** deep learning; overhead transmission line inspection; image semantic segmentation; defect detection; pixel classification

## 0 引言

高压架空输电线路覆盖区域广阔,输电设备各部件易出现不同程度的损坏<sup>[1]</sup>。暴露于大气环境中的金属设备形成的锈蚀,以及高压直流输电系统中金具表面泄漏

电流导致的电解锈蚀<sup>[2]</sup>,均会导致金属有效面积减少、材料性能退化,不仅会缩短金属构件的使用寿命,还会给输电线路的正常运输带来安全风险与经济损失,因此需要对输电线路定期进行金具锈蚀检测<sup>[3-6]</sup>。

输电线路巡检方法经过人工、机器人到无人机巡检的转变,使线路的维护成本显著降低,安全可靠性能增

加<sup>[7]</sup>。传统锈蚀图像分割方法主要在实验室场景下,根据纹理和颜色的视觉特征结合人工设计分割模型<sup>[8-11]</sup>。然而其检测结果受遮挡、光影、涂装等外界随机因素影响,并且精度和效率不高。

近年来深度学习方法与缺陷检测相结合的方法逐渐成为主流<sup>[12-13]</sup>。其中卷积神经网络(convolutional neural network, CNN)用于故障识别,具有分类能力强及检测成本低等特点。锈蚀缺陷检测中,根据任务特点优化网络,在锈蚀区域检测任务中的变种包括但不限于 ANN、ECNN、Mask-RCNN 等网络<sup>[14-17]</sup>。

Long 等<sup>[18]</sup>在 2014 年提出全卷积神经网络(fully convolutional networks, FCN)架构,是首个端到端像素级的图像语义分割模型,使用上采样方式恢复图像尺寸<sup>[19]</sup>。在锈蚀分割任务中,Nash 等<sup>[20]</sup>利用 VGG-16<sup>[21]</sup>为主干网络的 FCN 模型。与 FCN 相似的 SegNet 在此基础上添加了解码器结构<sup>[22]</sup>。上述两种方法均采用了多次上采样加卷积的操作,会损失细粒度信息,不利于细节繁多的金具锈蚀分割。Katsamenis 等<sup>[23]</sup>利用 U-Net<sup>[24]</sup>模型用于分割金属结构腐蚀,其独特结构兼容不同尺寸输入,然而 U-Net 缺少对上层信息的融合,导致位置及边缘信息丢失。Han 等<sup>[25]</sup>采用线性迭代聚类对图像预分割后,结合 FPN 在 PANet 算法训练,分割钢结构表面腐蚀,在大目标任务中表现良好。Qian<sup>[26]</sup>对比 RefineNet<sup>[27]</sup>、PSPNet<sup>[28]</sup>、U-Net、DeepLab 系列<sup>[29-31]</sup>在金属表面锈蚀分割任务上的表现,其中 RefineNet 不适用于锈蚀检测,PSPNet 使用多尺度卷积,但细节表征能力不足,而 DeepLab 表现出性能最佳。Rahman 等<sup>[32]</sup>创建自动生成的数据集,通过 DeepLabV3+ 模型训练后,在锈蚀检测和分割任务上具有一定提升。然而 DeepLabV3+ 算法依旧存在局部细节丢失、信息相关性弱、参数量大及耗时长等问题。

针对上述问题,以及金具锈蚀在航拍图像中分布不均匀、细节信息丰富的特征,提出一种基于 FEF-DeepLabV3+ 模型的高效高精度轻量化电力金具分割模型。首先,替换主干网络为改进 MobileNetV3 网络,优化耗时层,加快训练速度、减少参数量及计算量大小,实现轻量化要求。其次,提出 AFP 结构融合多尺度,并且优化颈部网络为 FEF-ASPP,由于空洞卷积的采样输入信号稀疏,故增加卷积分支用于保留不同分辨率锈蚀区域边缘信息,并级联相邻分支,使得远距离卷积之间锈蚀像素相互依赖。同时替换归一化层以减少模型对 batch size 大小的依赖,优化损失函数以加快模型收敛。改进后模型具有较好的实时性和较高的精度,降低了边缘人工智能设备的性能要求,并且适用于复杂缺陷分割环境。

## 1 电力金具锈蚀分割总体架构

### 1.1 锈蚀检测框架

首先通过装载高清摄像头的无人机设备巡检输电线路,获取含有金具的无人机巡检可见光图像。其次人工标注图像中锈蚀区域,制作为模型训练所需数据集。在特定的硬件设施及软件环境下,经过网络模型训练并保存最佳网络权重。使用嵌入式无人机巡检设备时,无人机采集图像及影像数据,利用已经学习到的权重数据分割锈蚀区域,完成无人机自动实时线路巡检。金具锈蚀分割技术路线如图 1 所示。

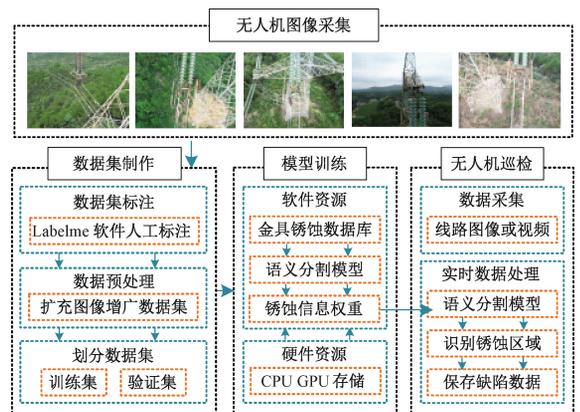


图 1 金具锈蚀区域图像分割框架

Fig. 1 Image segmentation frame for the corrosion area of the armor clamp

### 1.2 基础网络模型 DeeplabV3+

DeepLabV3+ 算法凭借其独特的空洞卷积结构,以及结构简洁的编码解码模型,实现了较高的图像分割精度。采用空洞卷积避免连续下采样和重复池化,导致最后特征图分辨率过低的情况。利用 ASPP 模块融合深层和浅层信息。不同空洞率的空洞卷积可以获取更多的上下文信息,并且能够让每个卷积输出都包含较大范围的信息,对于金具锈蚀的细目标具有较好的提取能力。其网络结构如图 2 所示。

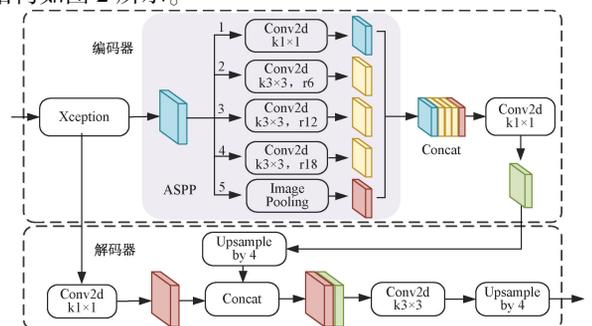


图 2 DeepLabV3+ 网络模型

Fig. 2 DeepLabV3+ network model

### 1.3 改进 FEF-DeepLabV3+模型

针对复杂环境背景条件下金具锈蚀区域的识别,考虑到 DeepLabV3+架构存在局部信息丢失、运算时间长和细节信息表征能力不足等问题,FEF-DeepLabV3+模型主

要针对金具锈蚀缺陷特征,通过替换主干网络、优化颈部结构和调整损失函数,达成优化模型、提高精度及效率、轻量化设计的目的,改进后模型结构如图 3 所示。

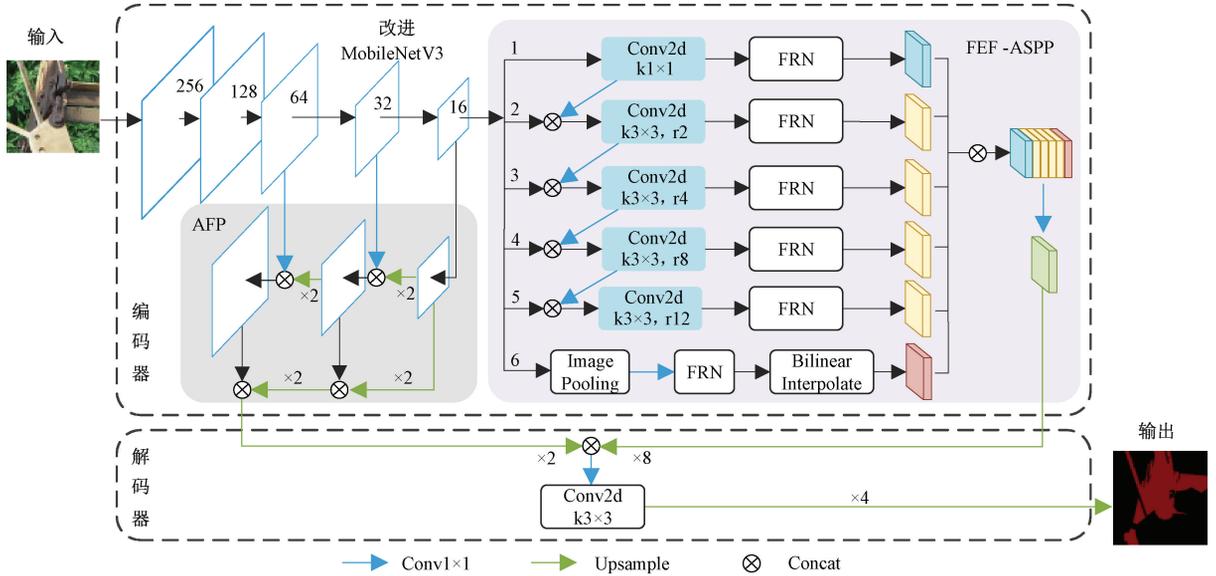


图 3 改进后 DeepLabV3+模型结构

Fig. 3 Improved DeepLabV3+ model structure

## 2 改进 FEF-DeepLabV3+锈蚀分割方法

### 2.1 替换骨干网络

#### 1) MobileNetV3 网络模块

MobileNetV3 网络作为轻量化网络适用于移动端及嵌入式设备,替换 DeepLabV3+模型原主干网络 Xception,图片提取有效特征信息的同时减少参数数量和计算量。MobileNetV3 网络运用深度可分离卷积代替基础卷积操作,能够降低参数数量和运算成本。深度可分离卷积采用倒残差结构,先将特征层通道数根据扩张率扩张到一定大小后,通过深度可分离卷积和 SENet 注意力网络,并压缩特征层通道数。当步距为 1 且特征层输入输出通道数相同时,采用 shortcut 连接输入输出特征层,倒残差结构如图 4 所示。

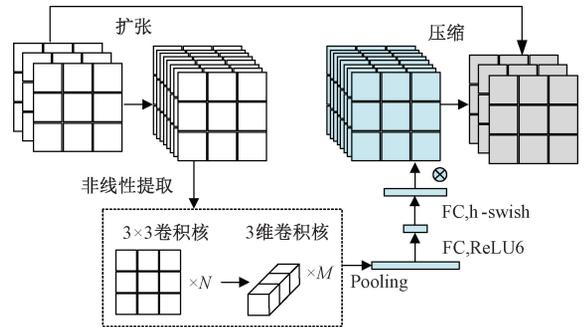


图 4 倒残差结构模块

Fig. 4 Inverted residual network block

失的信息,提高跨通道交互的覆盖率,如图 5(b) 所示。

加入 ECANet 模块后,删减主干网络中耗时层以显著减少模型参数及计算复杂度,使其在金具锈蚀分割任务中可以大幅降低训练及预测所需时间,表 1 为改进后 MobileNetV3 模型结构。

#### 3) AFP 自适应特征金字塔模块

金具锈蚀分割任务中,锈蚀区域在输电线路航拍图像中的占比小且细节信息丰富,因此需要保存图像像素的优势。语义分割不仅要关心语义信息,还要关注图像像素点包含的浅层信息。因此需要对网络进行融合,使其同时具有浅层纹理信息和深层语义信息。为防止损失锈蚀细节信息和整体锈蚀间的联系提出 AFP 模块。

#### 2) 替换注意力模块

模型中采用了通道注意力模块 SENet,如图 5(a) 所示,对输入特征层进行全局平均池化后,进行两次全连接,之后归一化权重与原输入特征层点乘,用于泛化模型。SENet 主要实现对通道注意力机制的预测,然而捕获所有通道的依赖关系是低效且不必要的。由于卷积具有良好的跨通道信息获取能力,因此 ECANet 模块中去除了 SENet 中的全连接层,改用 1x1 卷积可以避免降维损

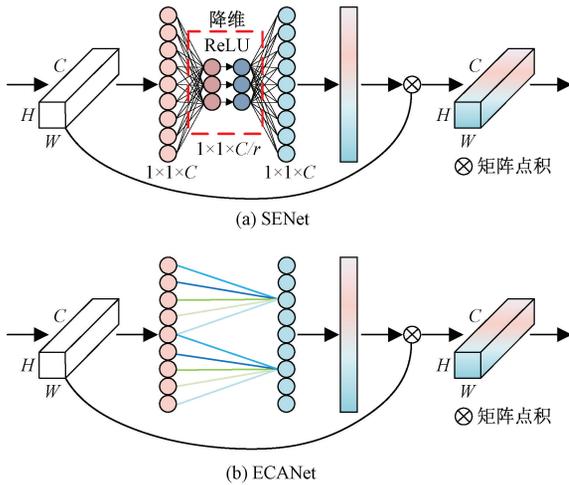


图 5 SENet 与 ECANet 结构对比

Fig. 5 Structure comparison of SENet and ECANet

表 1 改进 MobileNetV3 模型结构

Table 1 Improved MobileNetV3 model structure

编号	输入矩阵	序列操作	扩张率	输出通道数	步距
b0	$512^2 \times 3$	Conv, $3 \times 3$	-	16	2
b1	$256^2 \times 16$	bneck, $3 \times 3$	16	16	1
b2	$256^2 \times 16$	bneck, $3 \times 3$	64	24	2
	$128^2 \times 24$	bneck, $3 \times 3$	72	24	1
b3	$128^2 \times 24$	bneck, $5 \times 5$	72	40	2
	$64^2 \times 40$	bneck, $5 \times 5$	120	40	1
	$64^2 \times 40$	bneck, $3 \times 3$	240	80	1
	$64^2 \times 80$	bneck, $3 \times 3$	200	80	1
b4	$64^2 \times 80$	bneck, $3 \times 3$	184	80	2
	$32^2 \times 80$	bneck, $3 \times 3$	184	80	1
	$32^2 \times 80$	bneck, $3 \times 3$	480	112	1
	$32^2 \times 112$	bneck, $3 \times 3$	672	112	1
b5	$32^2 \times 112$	bneck, $5 \times 5$	672	160	2
	$16^2 \times 160$	bneck, $5 \times 5$	960	160	1
	$16^2 \times 160$	bneck, $5 \times 5$	960	160	1
	$16^2 \times 160$	Conv, $1 \times 1$	-	1 280	1

注: block 简写 b; bneck 表示深度可分离卷积模块; HS 表示 h-swish 激活函数, RE 表示 ReLU 激活函数

如图 3 所示, 将主干网络 block3、4、5 输出特征层上采样后与分别与 block2、3、4 输出特征层拼接, 生成的特征层经过上采样后拼接生成浅层特征, 与 FEF-ASPP 结构输出特征层一同作为解码器的输入。此外, 当下采样倍数设为 8 时, block4 不进行下采样并且特征层的大小不改变, 网络仅进行 4 下采样, 用于防止特征层分辨率降低、金具锈蚀细节像素信息的缺失, 缩短了低层纹理信息与顶层边缘细节特征之间的信息路径。

## 2.2 FEF-ASPP 模块

FEF-ASPP 结构结合多尺度和特征金字塔的结构, 使特征图中所有像素全局相互依赖。该模块设置不同空

洞率的空洞卷积捕捉不同距离信息, 可以提高远距离卷积算子的相关性。增加网络通道数不仅减少参数量, 而且比高、宽维度具有更强的表示能力, 可以提高细粒度信息分割的完整度。金具锈蚀区域像素量在整体像素中含有率低, 为了使特征图相邻像素及卷积后特征图之间联系更加紧密, 改造 ASPP 模块分支。如图 3 所示, 增加空洞卷积分支 5, 分支 2、3、4、5 的输入为各分支上层空洞卷积输出特征层与主干网络输出特征层在通道维度拼接, 可以在保留位置信息的同时, 语义信息保持不变。

为了更好地捕获分布不均匀且形状不规则的锈蚀区域, 使全局上下文信息联系紧密, 并实现跨尺度融合, 分支 6 首先通过全局平均池化层, 一维卷积调整通道数, 归一化后采用双线性插值方法恢复为输入特征图大小。最后将每个分支 concat 拼接后利用  $1 \times 1$  卷积调整特征层通道数大小。FEF-ASPP 结构能够在扩大感受野, 加强上下文联系的同时, 保证了细节像素信息的完整。

## 2.3 FEF-ASPP 模块优化

### 1) 替换归一化及激活函数

在计算机视觉领域深度学习网络中, 批归一化模块 BN 可以解决内部协变量偏移, 使各层分布相似, 平滑 loss 并加快模型训练速度, 故被各个检测模型广泛应用。然而在测试阶段的均值和方差计算可能与训练集的相差较大, 并且其性能受 batch size 大小影响, 随着 batch size 的增大, GPU 显存的占用率达峰值, 会导致训练中断并影响模型性能。

为解决上述问题, 将 FEF-ASPP 中 BN 及 ReLU 替换为 FRN 层。FRN 层沿宽、高维度去除噪声更新并减少梯度协方差, 所以没有对 batch 依赖的问题。若 FRN 替换 BN 后与使用 ReLU 激活函数, 则导致归一化后结果产生很多 0 值, 不利于模型训练。因此, 更改为与 FRN 兼容的 TLU 后性能有所提升。FRN 不仅消除了模型训练过程中对 batch 的依赖, 而且当 batch size 较小时性能优于 BN, 降低了网络精度对不同设备算力的依赖, 满足嵌入不同设备的需求。FRN 层结构如图 6 所示。

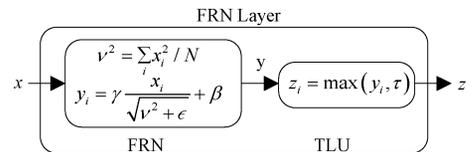


图 6 FRN 层结构

Fig. 6 FRN layer structure

FRN 去除了 BN 中减去均值的操作, 这种归一化方式类似 BN 可以用来消除卷积和非线性激活带来的尺度问题, 有助于模型训练。FRN 缺少去均值操作, 加入 TLU 防止 FRN 的结果任意偏移于 0, TLU 如式(1)所示。

$$z = \max(y, \tau) = \text{ReLU}(y - \tau) + \tau \quad (1)$$

式中:  $\tau$  是可学习的阈值。

## 2) 损失函数优化

Focal loss 主要用于平衡正负样本的权重, 以及控制容易分类样本和难分类样本的权重, 其表达式如式(2)所示。

$$L_{fl} = \begin{cases} -q(1-p)^\gamma \log(p) & , y = 1 \\ -(1-q)p^\gamma \log(1-p) & , y \neq 1 \end{cases} \quad (2)$$

式中:  $y$  代表前景或背景;  $p$  的取值范围为  $[0, 1]$ , 代表预测模型属于前景的概率; 权重因子  $q$  与  $1-q$  分别代表正负样本,  $q$  取 1 或 0; 调制因子  $(1-p)^\gamma$  用来降低容易分类样本的权重, 反之,  $p^\gamma$  则代表增大难以分类样本的权重;  $\gamma$  取值范围  $[0, 5]$ , 当  $\gamma=0$  时函数变为交叉熵损失函数 (cross entropy loss, CE loss)。

金具锈蚀分割任务中, 将识别种类分为背景和锈蚀区域, 而 Dice loss 也分为背景和物体两部分, 因此引入 Dice loss 解决样本极度不均的情况, 并且重点关注锈蚀区域的 loss。其中, Dice 系数是衡量集合相似度的度量函数, 表达式如式(3)所示。Dice loss 设计为 Dice 系数取反, 以防止对反向传播造成的负面影响, 即 Dice 系数越大, loss 越小, 两个样本越相近, Dice loss 表达式如式(4)所示。

$$\text{Dice} = 2 |A \cup B| / (|A| + |B|) \quad (3)$$

$$L_{dl} = 1 - \frac{\sum_{i=1}^k a_i b_i + \varepsilon}{\sum_{i=1}^k a_i + b_i + \varepsilon} - \frac{\sum_{i=1}^k (1 - a_i)(1 - b_i) + \varepsilon}{\sum_{i=1}^k 2 - a_i - b_i + \varepsilon} \quad (4)$$

式中:  $A, B$  表示两个轮廓区域所包含的点集;  $| \cdot |$  表示符号内元素个数;  $A \cup B$  表示两样本之间的交集;  $a_i, b_i$  指样本  $A, B$  中的点;  $k$  指集合内总数;  $\varepsilon$  防止分母为 0。由于分母计算实际上把交集重复计算两次, 故分子系数为 2。

将 Focal loss 以及 Dice loss 结合提出 Mixed loss, 由于两种 loss 的数量级不同, 故使用负对数函数放大 Dice loss, 使用  $\alpha$  缩小 Focal loss。Mixed loss 结合两种不同损失函数的特点, 同时符合分割任务目标, 其表达式如式(5)所示。

$$L_{ml} = \alpha L_{fl} - \log(L_{dl}) \quad (5)$$

式中:  $L_{ml}$  为 Mixed loss 函数;  $L_{fl}$  为 Focal loss 函数;  $L_{dl}$  为 Dice loss 函数;  $\alpha$  为函数缩小因子。

## 3 实验环境及评价指标

### 3.1 实验数据信息

实验数据集为某电网公司无人机航拍获取输电线路金具 RGB 可见光影像。使用 Labelme 标注软件像素级

标注数据集中锈蚀区域, 形成含标签的 JSON 文件, 转换为 PNG 格式作为训练所需标签文件。原始数据集中含  $4\ 864 \times 3\ 648$  pixels 96 dpi 图像共 360 张, 包含不同光线 (晴天 297 张、阴天 63 张)、角度 (正视角 86 张、俯视角 274 张)、地区 (包含沿输电线路 150 km 内的区域)、季节 (春 96 张、夏 76 张、秋 33 张、冬 155 张) 的输电线路航拍图像。金具锈蚀在不同图像中位于不同的位置以及大小不同, 因此其尺寸是不固定的, 但均具有形状不规则、分布不规则、细粒度信息丰富、高分辨率、高清晰度等特点, 可以在一定程度上提高模型算法的泛化能力。

由于语义分割为逐像素预测, 因此数据集在进行人工标注锈蚀区域时, 所遵循的判断标准主要为视觉特征, 这些特征通常包括纹理粗糙且颜色分布不均的表面, 并且锈蚀部位往往会呈现红棕色, 严重时甚至呈现黑红色。图 7 选取典型金具锈蚀图片, 包括放大锈蚀区域及其同比例放大标签。

采用 Python 内置库模块对图像扩充, 主要采用几何变换和颜色空间变换。前者包括垂直翻转、旋转 ( $90^\circ$ ,  $180^\circ$ ,  $-90^\circ$ ) 等操作, 用于模拟无人机在实际场景中不同角度的拍摄情况。后者主要包括对比度增强和亮度增强等操作, 为防止这两种操作扩充图片过多时, 可能引起的图像细节丢失或噪点增加问题, 这两种方法只进行一次扩充。最终基于上述 6 种扩充方式得到原数据集 6 倍的 2 160 张图片, 通过数据增强操作有效解决了数据类别不均衡问题。最后为了在保证训练数据充足的同时, 还可以提供足够的验证数据来评估模型性能, 以 9 : 1 比例随机划分训练集 1 944 张和验证集 216 张, 形成金具锈蚀分割 (armor clamp fittings corrosion segmentation, ACCS) 数据集。

选取典型金具锈蚀类型作为数据集展示, 图 7 中可以看出无人机航拍图像背景复杂, 包含草木植被、塔基、农田等。此外, 金具在图像中占比较小, 金具锈蚀分割任务需要提高对小目标的特征提取。图 7(a) 放大图像中包含典型连接金具 U 型挂环、螺栓螺母、联板等金具的锈蚀; 图 7(b) 放大图像中包含扇形联板、螺栓螺母等金具锈蚀; 图 7(c) 放大图像中包含防震锤锈蚀; 图 7(d) 放大图像中包含螺栓螺母、平行联板的不规则锈蚀; 图 7(e) 放大图像中包含锚固金具悬垂线夹的锈蚀。

### 3.2 实验环境及参数设置

参数设置中, 输入图像的大小统一设为  $512 \times 512$ , 批处理大小 (Batch size) 设为 4, 学习率范围设为  $[0.000\ 01, 0.001]$ , 下采样倍数设为 8, 优化器采用动量随机梯度下降 (momentum stochastic gradient descent, MSGD), 动量参数设为 0.9, 选用 cos 策略作为学习率下降方式, 取 0.000 1 作为权值衰减系数, 训练世代数 epoch 设为 100。表 2 为实验平台硬件及软件环境参数信息。



图7 ACCS数据集中典型金具展示

Fig. 7 Typical armor clamp display in the ACCS dataset

表2 实验平台及环境配置参数

Table 2 Experimental platform and environment configuration parameters

实验环境	类别	参数
硬件设备	CPU	Intel I5-11400F@ 2.60GHz
	GPU	NVIDIA GeForce RTX3060
	RAM	16 GB
	VRAM	12 GB
	操作系统	Windows11-64位
软件环境	深度学习框架	Pytorch1.11
	CUDA	cuda11.3
	CUDNN	8200
	开发环境	Pycharm2021
	编程语言	Python3.8

### 3.3 评价指标

在图像语义分割领域,常用的性能评价指标主要包括平均像素准确率(mean pixel accuracy, mPA)、平均交并比(mean intersection over union, mIoU)。其中,mIoU表示分割结果与其真值的重合度,是目前使用频率最高和最常见的评价指标。

1) mPA 用于计算正确分割的像素数量与图像像素总量的比例,其具体计算方法如式(6)所示。

$$mPA = \frac{1}{N} \sum_{i=1}^N X_{ii} / \sum_{i=1}^N T_i \quad (6)$$

2) mIoU 表示分割结果与原始图像真值的重合程度即交并比,其具体计算方法如式(7)所示。

$$mIoU = \frac{1}{N} \sum_{i=1}^n \frac{X_{ii}}{T_i + \sum_{j=1}^N (X_{ji} - X_{ii})} \quad (7)$$

其中, $N$ 代表图像像素的类别数量; $T_i$ 代表第*i*类的像素总数; $X_{ii}$ 代表实际类型为*i*、预测类型为*i*的像素总数; $X_{ji}$ 代表实际类型为*i*、预测类型为*j*的像素总数。

## 4 实验结果与分析

为了验证改进模型在 ACCS 数据集上的有效性,在同一实验环境即参数设置下,分别对主干网络替换、瓶颈结构模块以及各部分优化进行消融实验。

### 4.1 主干网络替换对比

选用 DeepLabV3+模型作为实验的基础模型,使用 ACCS 数据集作为训练数据集,训练过程中,选用 MobileNetV3 作为基础主干网络,在同一环境参数下,使用不同注意力机制模块进行训练。利用 mIoU、推理时间、帧率(frames per second, FPS)作为评价指标,其中,推理时间指处理每幅图片所需要的时间, FPS 指每秒处理帧数。

从表3可以看出,在原始 MobileNetV3 网络上,不添加注意力机制模块的推理时间最低,但评价指标 mIoU 仅有 78.46%;引入 SENet 模块后评价指标比初始模型提升了 2.79%。ECANet 模块相较于 SENet 模块推理时间加快 2.24 ms,帧率增加了 9.52 fps, mIoU 提升了 2.06%,故 ECANet 对于网络的性能提升优于 SENet。当 AFP 结构和 ECANet 共同优化主干提取网络时,与原始网络相比,评价指标推理时间和帧率仅相差 0.34 ms 与 2.87 fps,但 mIoU 提高了 5.81%。由此可得,加入 AFP 结构和 ECANet 注意力模块的主干提取网络,对锈蚀区域的分割效果具有较好的提升。

表3 消融实验对比

Table 3 Comparison of ablation experiments

序号	SENet	ECANet	AFP	mIoU/%	推理时间/ ms	帧率/fps
1	-	-	-	78.46	<b>14.08</b>	<b>72.41</b>
2	√	-	-	80.95	16.49	60.63
3	-	√	-	82.01	14.25	70.15
4	-	√	√	<b>84.27</b>	14.42	69.54

为验证改进 MobileNetV3 网络的有效性,选用不同主

干网络进行对比实验。

如表 4 所示,可以看出改进 MobileNet 对网络的优化主要体现在 4 个评价指标的综合提升,相较于 DeepLab 系列的主干网络 VGG16、ResNet101 以及 InceptionV3, mIoU 分别提高了 8.57%、5.85%、3.81%, mPA 分别提高了 12.57%、9.96%、6.72%, 且推理时间最短。与 Xception 的对比中, mIoU 与 mPA 分别提高了 0.21%、0.79%, 并且推理时间减少了 20.81 ms, FPS 提高了 2.45 倍, 综合 4 种评价指标, 改进 MobileNetV3 网络在 ACCS 数据集上有效。

表 4 不同主干网络在 ACCS 数据集上结果对比

Table 4 Comparison of results of different backbone networks on the ACCS dataset

主干网络	mIoU/%	mPA/%	推理时间/ms	帧率/fps
VGG16	75.70	80.57	14.88	67.21
ResNet101	78.42	83.18	18.92	52.85
InceptionV3	80.46	86.42	42.18	16.42
Xception	84.06	92.35	35.23	28.38
改进 MobileNetV3	<b>84.27</b>	<b>93.14</b>	<b>14.42</b>	<b>69.54</b>

#### 4.2 瓶颈模块优化对比实验

网络瓶颈模块主要包括 FEF-ASPP 模块, 对比实验使用 ACCS 数据集作为训练数据集, 使用改进 MobileNetV3 主干网络, 对 FEF-ASPP 模块改进部分进行分析对比。

##### 1) FEF-ASPP 模块参数对比

FEF-ASPP 模块在 ASPP 结构基础上优化, 在金具锈蚀分割任务中, 需要根据分割目标调整膨胀率, 以获得最佳性能。为验证最佳膨胀率以及改进后性能, 以 DeepLabV3+ 为基础模型进行对比实验。实验结果如表 5 所示。

表 5 FEF-ASPP 模块性能对比实验

Table 5 FEF-ASPP ASPP module performance comparison experiment

序号	下采样倍数	FEF-ASPP	膨胀率	mIoU/%	mPA/%
1	16	-	(6, 12, 18)	84.27	93.14
2	16	√	(6, 12, 18, 24)	86.13	94.53
3	8	√	(6, 12, 18, 24)	86.20	93.87
4	8	√	(5, 7, 11, 15)	86.36	95.47
5	8	√	(3, 6, 9, 12)	86.41	95.58
6	8	√	(4, 8, 10, 12)	86.45	95.65
7	16	-	(2, 5, 7)	85.35	93.52
8	16	√	(2, 5, 7, 11)	86.61	95.48
9	8	√	(2, 5, 7, 11)	<b>86.62</b>	<b>95.87</b>

根据实验结果可以得出, 使用 FEF-ASPP 结构时, 评

价指标数值均优于原 ASPP 结构, 并且当下采样系数为 8、膨胀率设为 (2, 5, 7, 11) 时性能最佳。与原颈部网络 ASPP 模块相比, FEF-ASPP 模块针对金具锈蚀任务设置最佳膨胀率时, 评价指标 mIoU 和 mPA 分别提高了 2.35%、2.73%。

##### 2) 不同归一化层对比

由于特征量纲差异引起的梯度震荡和收敛速度过慢, 训练过程中使用 FRN Layer 为归一化层。为验证 FRN 在网络中的作用, 采用不同归一化方法在同一环境下进行验证, 对比不同批处理大小时的结果。实验中除了 FRN 外均配合使用 ReLU 激活函数, 数值稳定性均设为  $1 \times 10^{-5}$ , BN 和 IN 动量设为 0.1。

如图 8 所示, BN 在 batch size 较小时效果不佳, 但随着 batch size 增大, 其性能会大幅提升。BRN 重新分配深层网络, 因此训练深层网络模型时会更加稳定。LN 对中间层的所有神经元归一化, 对其分配不同的均值和方差, 因此在序列长度不一致的 RNN 上有良好的表现, 但在语义分割任务中效果不佳。GN 而在特征图通道维度上分组, 以每组为单位求均值和方差并进行归一化, 能够解除对 batch size 的依赖。在不同硬件资源的情况下, 使用 FRN 层可以使网络在不同 batch size 上均有良好的表现, 当 batch size 设为 4 时达到最优性能, 此时 mIoU 为 87.15%。

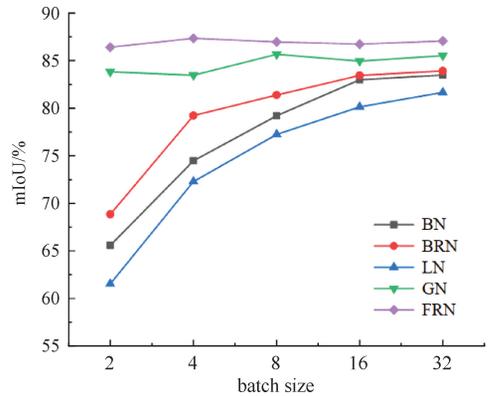


图 8 不同归一化方法下 mIoU 性能对比

Fig. 8 Comparison of mIoU performance under different normalization methods

#### 4.3 损失函数优化对比

为验证改进后的 Mixed loss 在金具锈蚀分割任务上收敛并优于原损失函数, 对比原网络中使用的交叉熵损失 CE loss 和 Mixed loss 损失函数的性能。

损失函数改进前后在训练过程中的训练集和验证集的对比如图 9 所示。可以看出, Mixed loss 相对于 CE loss 来说收敛速度更快, loss 值更小, 且 Mixed loss 每个 epoch 的 loss 比 CE loss 平均低 0.15。

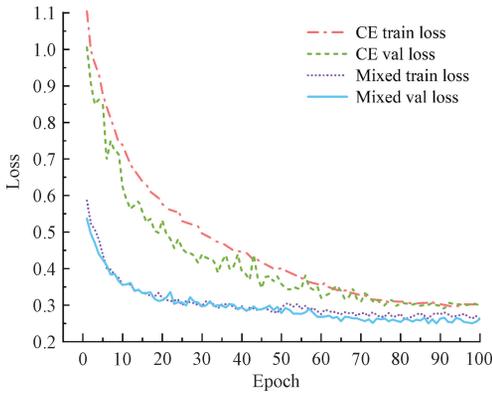


图 9 损失函数收敛曲线对比

Fig. 9 Comparison of loss function convergence curves

### 4.4 不同语义分割模型对比

在同一环境下,对 U-Net、HRNet、SegNet、PSPNet、DeepLabV3+5 种经典语义分割模型,与本文算法在 ACCS 数据集上进行训练,选用综合考虑模型精度、体积、速度的评价指标,对比结果如表 6 所示。

表 6 不同架构对比结果

Table 6 Compare the results of different architectures

主干网络	mIoU/	mPA/	推理时	帧率/	参数量/	FLOPs/
	%	%	间/ms	fps	M	G
U-Net	67.53	71.78	<b>15.52</b>	<b>64.43</b>	26.89	452.31
HRNet	76.90	81.18	70.16	14.25	29.54	81.17
SegNet	80.38	85.34	41.64	24.01	27.36	113.90
PSPNet	83.44	92.87	43.52	21.68	46.71	118.44
DeepLabV3+	84.06	92.35	35.23	28.38	54.71	486.65
本文算法	<b>87.15</b>	<b>96.64</b>	15.94	62.07	<b>26.53</b>	<b>51.14</b>

从表 6 中可以看出,基于改进 DeepLabV3+ 的本文算法在 mIoU、mPA 评价指标上的精度最高,与原方法分别提高了 3.09% 和 4.29%,与 U-Net 相比分别提高了 19.62% 和 24.86%。推理时间和 FPS 等表示运算速度的评价指标与 U-Net 仅差 0.42 ms 和 2.36 fps,推理时间和帧率分别是原 DeepLabV3+ 模型 45% 和 2.2 倍,并且优于其他网络模型。其中,帧率高达 62.07 fps,能对 60 帧以内的视频影像进行检测。参数量表示模型的大小以及占用显存的量,本文算法参数量仅占 26.53 M,进一步降低训练网络设备的算力要求,使嵌入在移动端成为可能。计算量(floating-point of operations, FLOPs)用来衡量模型复杂度,可以看作网络执行时间的长短。本文算法参数量和 FLOPs 均优于其他网络模型,仅为原 DeepLabV3 模型的 48% 和 10%。

如图 10 所示,在 ACCS 数据集上,分别和 U-Net、PSPNet 以及基础模型 DeepLabV3+ 与本文算法分割结果

进行可视化对比。图 10 中每一行包括原图、标签图片和语义分割后可视化图像,将分割结果与原图融合在一起以便于观察。具体实验结果分析如下:

图 10 第 1 列中,除本文模型外均未检测到电力线的锈蚀区域,U-Net 右侧悬垂线夹锈蚀未检出,PSPNet 与基础模型将部分杆塔误检为锈蚀;图 10 第 2 列中 PSPNet 与基础模型在右侧出现部分误检,右下方锈蚀扇形联板漏检,图像上方的防护金具仅在本文模型中被正确检出,然而绝缘子部分连接处阴影与金具锈蚀颜色相近,导致本文算法产生细小误检;图 10 第 3 列中其他 3 种算法在屏蔽环内联板以及周围连接金具区域检测界限模糊,检测结果与背景区域混同,左侧远距离锈蚀区域漏检,细小螺栓锈蚀也并未检出,然而本文模型在细小锈蚀中仅有一小部分未检出;图 10 第 4 列中 U-Net 将均压环细节部分未检出,PSPNet 与原始模型分割结果在绝缘子左侧出现误检现象,将屋舍土地识别为锈蚀,并均有误检未锈蚀金具区域的情况。

图 10 中可以看出通过本文模型训练后,得出的分割结果更接近标签真实值,能更完整并精准的识别和分割锈蚀区域,边缘像素细节检测优于其他模型,分割轮廓清晰明显,具备较强的鲁棒性。此外本文改进模型能检测出细小螺栓锈蚀,证明该算法在细粒度特征检测效果良好。

## 5 结 论

本文以 DeepLabV3+ 分割模型为基础,提出一种基于深度学习的金具锈蚀区域分割方法,替换轻量级改进 MobileNetV3 算法作为特征提取网络,减少了下采样次数,在保证分割精度的同时提升了效率,降低了分割任务对设备算力的要求。采用 FEF-ASPP 结构平衡了大范围感受野和高像素的需求,防止割裂像素间的联系,有效提取锈蚀细节信息,当膨胀率设为 [2, 5, 7, 11] 时,模型性能最佳。最后优化损失函数加快模型收敛。

实验结果得出,改进 DeepLabV3+ 模型在 ACCS 数据集上的 mIoU 为 87.15%,准确率为 96.64%,优于其他网络模型,与基础模型相比分别提高了 3.09% 和 4.29%。参数量与计算量均优于其他网络。推理时间为 15.94 ms,比基础网络快 19.29 ms,相当于减少 1/2 的时间。此外,高于 60 的 FPS 能够兼容大多数视频检测任务需求,具有小体积、高精度的实时智能输电线路巡检应用前景。

### 参考文献

[ 1 ] 胡毅,刘凯,吴田,等. 输电线路运行安全影响因素分析及防治措施 [ J ]. 高电压技术, 2014, 40 ( 11 ): 3491-3499.

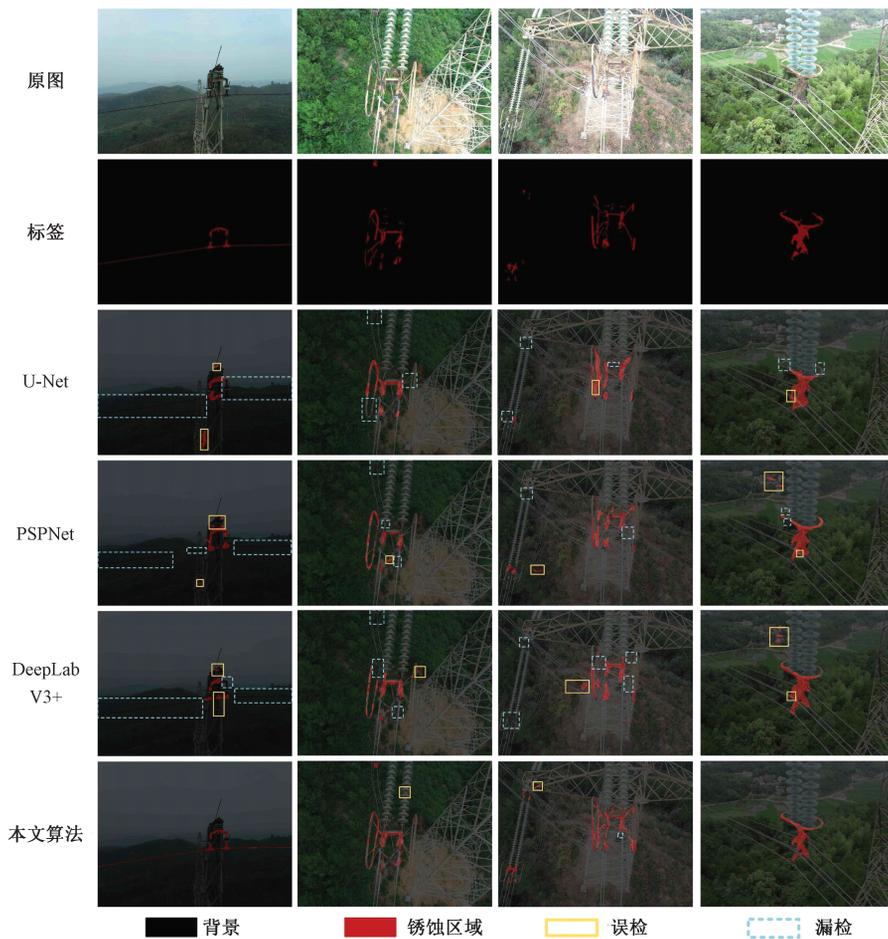


图 10 分割结果可视化对比

Fig. 10 Visualization comparison of segmentation results

factors on operation safety of transmission line and countermeasures [ J ]. High Voltage Engineering, 2014,40(11):3491-3499.

- [ 2 ] 吴向东,程绳,金哲,等. 通过降低泄漏电流提高 UHVDC 盘型瓷绝缘子的抗金具腐蚀能力[J]. 电网技术,2019,43(2):746-752.

WU X D, CHENG SH, JIN ZH, et al. Improvement of metal-fitting corrosion resistance of UHVDC disc porcelain insulators by reducing leakage current [ J ]. Power System Technology, 2019,43(2):746-752.

- [ 3 ] 邵瑰玮,刘壮,付晶,等. 架空输电线路无人机巡检技术研究进展[J]. 高电压技术,2020,46(1):14-22.

SHAO G W, LIU ZH, FU J, et al. Research progress in unmanned aerial vehicle inspection technology on overhead transmission lines [ J ]. High Voltage Engineering, 2020,46(1):14-22.

- [ 4 ] 刘壮,杜勇,陈怡,等.  $\pm 500$  kV 直流输电线路直线塔无人机巡检安全距离仿真与试验[J]. 高电压技术,2019,45(2):426-432.

LIU ZH, DU Y, CHEN Y, et al. Simulation and experiment on the safety distance of typical  $\mu 500$  kV DC transmission lines and towers for UAV inspection [ J ]. High Voltage Engineering,2019,45(2):426-432.

- [ 5 ] 周孝信,陈树勇,鲁宗相. 电网和电网技术发展的回顾与展望——试论三代电网[J]. 中国电机工程学报,2013,33(22):1-11.

ZHOU X X, CHEN SH Y, LU Z X. Review and prospect for power system development and related technologies: A concept of three-generation power systems[J]. Proceedings of the CSEE, 2013,33(22):1-11.

- [ 6 ] 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):121-127,146.

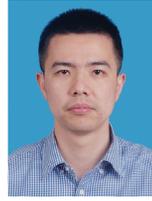
WANG D L, PENG B, PAN Y, et al. Segmentation and quantitative analysis of corrosion images based on deep neural networks[J]. Journal of South China University of Technology (Natural Science Edition), 2018,46(12):121-127,146.

- [ 7 ] 赵振兵,蒋志钢,李延旭,等. 输电线路部件视觉缺陷检测综述[J]. 中国图象图形学报, 2021, 26(11): 2545-2560.
- ZHAO ZH B, JIANG ZH G, LI Y X, et al. Overview of visual defect detection of transmission line components[J]. *Journal of Image and Graphics*. 2021, 26(11): 2545-2560.
- [ 8 ] LIAO K W, LEE Y T. Detection of rust defects on steel bridge coatings via digital image recognition [ J ]. *Automation in Construction*, 2016, 71: 294-306.
- [ 9 ] 马云鹏,李庆武,何飞佳,等. 金属表面缺陷自适应分割算法[J]. 仪器仪表学报, 2017, 38(1): 245-251.
- MA Y P, LI Q W, HE F J, et al. Adaptive segmentation algorithm for metal surface defects[J]. *Chinese Journal of Scientific Instrument*, 2017, 38(1): 245-251.
- [ 10 ] SHEN H K, CHEN P H, CHANG L M. Automated steel bridge coating rust defect recognition method based on color and texture feature [ J ]. *Automation in Construction*, 2013, 31: 338-356.
- [ 11 ] KHAYATAZAD M, DE PUE L, DE WAELE W. Detection of corrosion on steel structures using automated image processing [ J ]. *Developments in the Built Environment*, 2020, 3: 100022.
- [ 12 ] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1-9.
- [ 13 ] 刘传洋,吴一全. 基于深度学习的输电线路视觉检测方法研究进展[J/OL]. 中国电机工程学报: 1-24 [2023-02-23].
- LIU CH Y, WU Y Q. Research progress of vision detection methods based on deep learning for transmission lines[J/OL]. *Proceedings of the CSEE*: 1-24 [2023-02-23].
- [ 14 ] ORTIZ A, FRANCISCO B P, EMILIO G F, et al. Vision-based corrosion detection assisted by a micro-aerial vehicle in a vessel inspection application [ J ]. *Sensors*, 2016, 16(12): 2118.
- [ 15 ] FORKAN A R M, KANG Y B, JAYARAMAN P P, et al. CorrDetector: A framework for structural corrosion detection from drone images using ensemble deep learning [ J ]. *Expert Systems with Applications*, 2022, 193: 116461.
- [ 16 ] XU J, GUI C, HAN Q. Recognition of rust grade and rust ratio of steel structures based on ensembled convolutional neural network [ J ]. *Computer - Aided Civil and Infrastructure Engineering*, 2020, 35(10): 1160-1174.
- [ 17 ] TA Q, HUYNH T C, PHAM Q Q, et al. Corroded bolt identification using mask region-based deep learning trained on synthesized data [ J ]. *Sensors*, 2022, 22(9): 3340.
- [ 18 ] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [ C ]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 3431-3440.
- [ 19 ] 罗会兰,张云. 基于深度网络的图像语义分割综述[J]. 电子学报, 2019, 47(10): 2211-2220.
- LUO H L, ZHANG Y. A survey of image semantic segmentation based on deep network [ J ]. *Acta Electronica Sinica*, 2019, 47(10): 2211-2220.
- [ 20 ] NASH W, DRUMMOND T, BIRBILIS N. Quantity beats quality for semantic segmentation of corrosion in images[J]. *arXiv preprint arXiv:1807.03138*, 2018.
- [ 21 ] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [ C ]. *International Conference on Learning Representations*, 2015: 1-14.
- [ 22 ] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation [ J ]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [ 23 ] KATSAMENIS I, DOULAMIS N, DOULAMIS A, et al. Simultaneous precise localization and classification of metal rust defects for robotic-driven maintenance and prefabrication using residual attention U-Net [ J ]. *Automation in Construction*, 2022, 137: 104182.
- [ 24 ] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. *International Conference on Medical Image Computing and Computer-assisted Intervention*, 2015: 234-241.
- [ 25 ] HAN Q, ZHAO N, XU J. Recognition and location of steel structure surface corrosion based on unmanned aerial vehicle images [ J ]. *Journal of Civil Structural Health Monitoring*, 2021, 11(5): 1375-1392.
- [ 26 ] QIAN C. Evaluation of deep learning-based semantic segmentation approaches for autonomous corrosion detection on metallic surfaces[D]. West Lafayette, IN, USA: Purdue University, 2019.
- [ 27 ] LIN G, MILAN A, SHEN C, et al. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation [ C ]. *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE, 2017. 5168-5177.
- [ 28 ] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [ C ]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017:

2881-2890.

- [29] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs [ C ]. Computer Science, 2014: 357-361.
- [30] 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, 2017, 40 ( 4 ): 834-848.
- [31] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder decoder with atrous separable convolution for semantic image segmentation [ C ]. Proceedings of the European Conference on Computer Vision ( ECCV ), 2018: 801-818.
- [32] RAHMAN A, WU Z Y, KALFARISI R. Semantic deep learning integrated with RGB feature-based rule optimization for facility surface corrosion detection and evaluation [ J ]. Journal of Computing in Civil Engineering, 2021, 35(6): 04021018.

## 作者简介



王凌云, 2003 年于三峡大学获得学士学位, 2009 年于中南大学获得博士学位, 现为三峡大学副教授, 主要研究方向为电力设备智能检测。

E-mail: wly@ctgu.edu.cn

**Wang Lingyun** received his B. Sc. degree from China Three Gorges University in 2003 and Ph. D. degree from Central South University in 2009, respectively. Now he is an associate professor in China Three Gorges University. His main research interest includes the intelligent detection of electrical power equipment.



李婷宜(通信作者), 2021 年于三峡大学获得学士学位, 现为三峡大学硕士生, 主要研究方向为输电线路智能缺陷检测。

E-mail: 815387220@qq.com

**Li Tingyi** ( Corresponding author ) received her B. Sc. degree from China Three Gorges University in 2021. She is now a M. Sc. candidate in China Three Gorges University. Her main research interest includes intelligent defect detection of transmission lines.