

基于深度学习的电力杆塔缺陷检测^{*}

张关应¹ 杜云豪¹ 常宸铠² 候姝斌¹ 李庆武²

(1. 国网江苏省电力有限公司常州供电公司 常州 213000; 2. 河海大学信息科学与工程学院 常州 213000)

摘要: 针对目前无人机电力杆塔巡检过程中航拍图像背景复杂、杆塔组件尺寸差异大、缺陷种类多导致的巡检效率低和缺陷漏检率高等问题,本文提出一种动态位置查询引导的多尺度实例分割方法和基于图特征记忆的缺陷检测方法。所提实例分割方法通过提取多尺度航拍图像特征,选择特征中具有最高关注度分数的低分辨率像素,将其映射到高分辨率特征的相应位置,并添加边界框检测器以提高电力杆塔的分割精度。在缺陷检测算法中提出可学习的图特征描述符,构建了一个记忆库来提取关键元素,以获得更准确的样本特征,提高缺陷检测效率。本文方法在自建的两个缺陷检测数据集上与其他先进算法进行对比,实例分割 box_AP 和 mask_AP 相较于 Mask2Former 分别提升了 7.6% 和 0.5%,缺陷检测算法 AUROC 分别比次优算法提高 7.3% 和 1.6%,F1-Score 分别比次优算法提高 6.7% 和 6.9%,充分表明了本文算法出色的电力杆塔缺陷检测性能。

关键词: 杆塔; 电力设备; 无人机巡检; 实例分割; 缺陷检测

中图分类号: TP2 文献标识码: A 国家标准学科分类代码: 510.4050

Deep learning-based transmission tower defect detection

Zhang Guanying¹ Shu Yunhao¹ Chang Chenkai² Hou Shubin¹ Li Qingwu²

(1. State Grid Changzhou Power Supply Company, Changzhou 213000, China;

2. College of Information Science and Engineering, Hohai University, Changzhou 213000, China)

Abstract: In response to low efficiency and high leakage rate caused by different sizes of transmission tower components and more defects in the current defect detection of power line towers, this paper proposes a Dynamic Position Query-Guided Multi-Scale Instance Segmentation method and a Graph Feature Memory-based Defect Detection method. The proposed instance segmentation method extracts multi-scale aerial image features, selects low-resolution pixels with the highest attention scores from the features, maps them to the corresponding positions in high-resolution features, and incorporates a bounding box detector to enhance the segmentation accuracy of power transmission towers. In the defect detection algorithm, a learnable graph feature descriptor is introduced, a memory bank is constructed to extract key elements for more accurate sample feature extraction, thereby improving defect detection efficiency. The power transmission tower defect detection method presented in this paper is compared with other state-of-the-art algorithms on two self-constructed defect detection datasets, the box_AP and mask_AP of instance segmentation saw significant improvements of 7.6% and 0.5%, respectively, compared to Mask2Former. The AUROC indicator of defect algorithm was 7.3% and 1.6% higher than the second-best algorithm for the two datasets, and the F1-Score was improved by 6.7% and 6.9%, respectively. These results strongly demonstrate the outstanding performance of our algorithm in the transmission tower defect detection.

Keywords: transmission tower; power equipment; UAV inspection; instance segmentation; defect detection

0 引言

输电线路在确保国家电力系统正常运行方面发挥着至关重要的作用^[1]。电力杆塔通常位于偏远、高海拔的地区,

日常运行中存在各种潜在危险^[2]。近年来,无人机(unmanned aerial vehicles, UAVs)在输电线路检查中的应用日益增多,在减少巡检人员受伤风险的同时还能够获得高分辨率的图像和数据以供分析。

在无人机运用于输电线路检查的初期,获得的航拍图像主要依靠巡检人员后期进行人工分析。随着深度学习技术的快速发展,许多输电线路检测^[3]、电力设备检测^[4]、缺陷检测^[5]等计算机视觉算法^[6]被部署在巡检无人机上,以提升巡检的智能化^[7]。Bolya 等^[8]通过预测一组具有共享全局 mask 特征的特定实例 mask 加权系数来生成实例 mask,提升了分割效率。受 DETR (detection transformer)^[9]的启发,基于 Query 的实例分割方法将分割视为一个集合预测问题,利用查询表示感兴趣的对象,并共同对这些对象执行分类、检测和 mask 回归。Cheng 等^[10]进一步提出了一个通用于图像分割任务的方法 Mask2former,对 Transformer 解码器中的交叉注意进行改进和替换,构建多尺度的高分辨率特征金字塔,以轮询的方式进行解码,分割的 mask 具有准确的边缘。Roth 等^[11]和 Park 等^[12]将基于记忆库的方法应用到异常检测中,在不学习目标数据集的情况下,利用在大型数据集上预训练过的特征提取网络提取广义特征,将其存储到记忆库中,测试时将输入特征与记忆特征进行匹配,确定输入样本是否存在异常。

尽管上述方法在不同条件下取得了一定的成果,但现有电力杆塔缺陷检测方法仍存在一些问题,主要表现在:1)由于电力设备尺寸差异较大,现有的目标检测方法的准确性显著降低。此外,现有方法未在随后的缺陷检测过程中滤除背景干扰,容易出现漏检、错检等问题。2)由于缺陷和异物的类型众多^[13],训练集难以涵盖所有类型的缺陷,因此,现有的缺陷检测方法通常只能检测某些设备的特定缺陷,无法检测未经训练的缺陷类型。

针对前文提到的第 1 个问题,本文选择横向绝缘子、竖向绝缘子、连接金具和防振锤四类目标构建了短焦距成像

数据集(short focus imaging dataset, SFID)和长焦距成像数据集(long focus imaging dataset, LFID),提出了一种动态位置查询引导的多尺度实例分割方法。首先提取航拍图像多层次的特征,然后使用像素解码器获取多尺度的像素级特征。为了适当地保持跨关注层中关注区域之间的一致性,提出了动态位置查询,应用于 Transformer 解码器中,利用目标区域局部上下文信息减轻了电力杆塔关键组件的遮挡问题。此外,还添加了一个边界框检测头,进一步改善误识别实例目标的问题。

针对前文提出的第 2 个问题,本文提出了一种基于图特征记忆的缺陷检测方法。基于实例分割结果,定位设备位置并滤除背景干扰,获取电力杆塔部件前景图像。利用预训练网络提取杆塔部件前景图像不同尺度的特征,将第三层特征和第二层特征构建面向数据集的图特征描述符。在收集所有正常样本的图特征描述符后,通过核心子采样提取关键元素以初始化记忆库。通过在图特征描述符与记忆库上执行最近邻搜索,生成表示异常程度的热力图,实现电力杆塔常见多类型设备的缺陷检测。

1 本文方法

1.1 动态位置查询引导的多尺度实例分割方法

网络框架如图 1 所示。骨干网络提取图像特征,像素解码器逐步对图像特征进行上采样,以获得多尺度特征。Transformer 解码器从低分辨率的特征中选出关注度分数最高的前 K 个像素,以自上而下的方式将这些区域映射到高分辨率特征的相应位置,并仅对这些位置应用 Masked Attention。最终的预测结果由 3 个预测分支得出:类别预测分支、mask 预测分支和边界框预测分支。

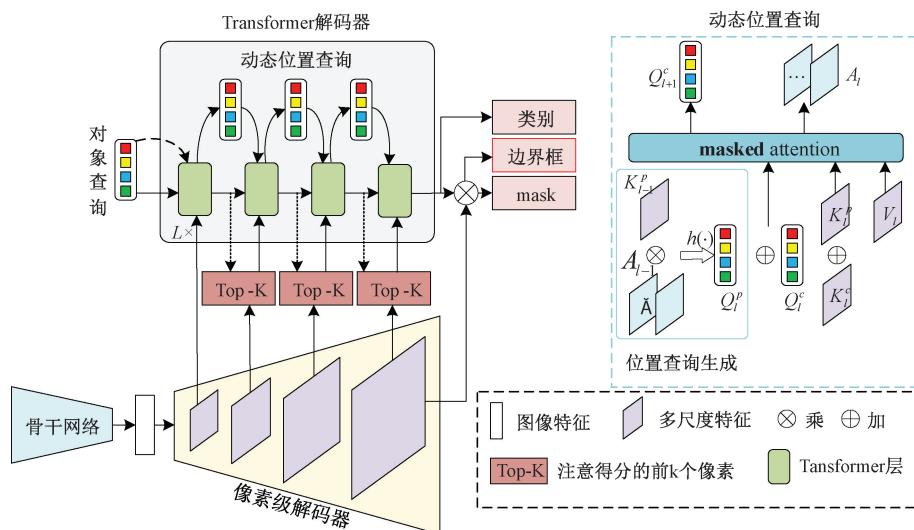


图 1 动态位置查询引导的多尺度实例分割网络的框架

1) 动态位置查询

动态位置查询图如图 1 中浅色虚线框所示。将编码器输出特征的位置编码 K_{l-1}^p 与注意映射 A_{l-1} 相乘,利用一

个投影函数 $h(\cdot)$ 进行处理,得到动态位置查询:

$$Q_l^p = h(A_{l-1} K_{l-1}^p + B) \quad (1)$$

其中, l 是 Transformer 解码器层的索引, $A_{l-1} \in$

$\mathbb{R}^{N \times HW}$ 是来自第 $(t-1)$ 层 Transformer 解码器块的 Masked Attention 分数, $B \in \mathbb{R}^{N \times D}$ 是可学习的网络参数, 注意偏置项 B 是原始随机初始化的可学习位置查询, h 是非线性投影-非线性 ReLU-非线性投影的三重结构, 其中 N, H, W 和 D 分别表示查询次数、图像特征的高度、宽度和隐藏维度。

2) 多尺度特征 Transformer 解码器

本文提出从像素解码器中的多尺度图像特征中额外引入最高分辨率特征, 将每个 Transformer 解码器组中包含的解码器层扩展为 4 个。在每个解码器层中, 首先从低分辨率的特征图中选择注意力得分最高的前 K 个像素, 以自上而下的方式将这些区域映射到高分辨率特征的对应位置, 并且只对这些位置进行 Masked Attention 操作。将之前的低分辨率注意图中关于注意得分的前 K 个像素纳入集合 Ω 中:

$$X_t = \text{softmax}(M_{t-1} + Q_t K_t'^T) V_t' + X_{t-1} \quad (2)$$

其中, $X_t \in \mathbb{R}^{N \times C}$ 指第 t 层 $Q_t = f_q(X_{t-1}) \in \mathbb{R}^{N \times C}$ 处的 NC 维度查询特征, X_0 为输入 Transformer 解码器初始查询特征, Masked attention M_{t-1} 在 (x, y) 特征处是:

$$M_{t-1}(x, y) = \begin{cases} 0, & M_{t-1}(x, y) = 1 \\ -\infty, & \text{其他} \end{cases} \quad (3)$$

其中, $M_{t-1} \in \{0, 1\}^{N \times H_t \times W_t}$ 为前一层第 $(t-1)$ 层 Transformer 解码器预测的二值化输出结果, $K_t' = e(K, g(\Omega)), V_t' = e(V, g(\Omega))$, $g(\cdot)$ 是将低分辨率映射到高分

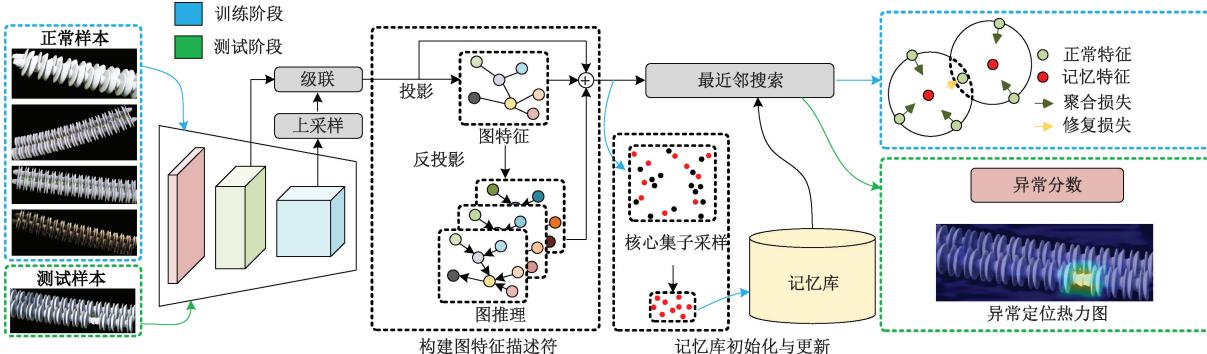


图 2 基于图特征记忆的缺陷检测网络的框架

在训练阶段, 通过正常特征与记忆特征的对比进行监督, 解决图特征描述符所提取的特征与目标数据集特性存在偏差的问题, 并逐渐更新记忆库。在测试阶段, 将从测试样本中提取的图特征描述符与记忆库进行最近邻搜索, 生成表示异常程度的热力图, 并通过特定函数计算异常分数, 显示异常特征的细化结果。

1) 图特征描述符

给定分辨率为 $H \times W$ 的输入样本, 采用 ResNet 18 提取图像特征: $F_1 \in \mathbb{R}^{64 \times \frac{H}{4} \times \frac{W}{4}}$, $F_2 \in \mathbb{R}^{128 \times \frac{H}{8} \times \frac{W}{8}}$, $F_3 \in \mathbb{R}^{256 \times \frac{H}{16} \times \frac{W}{16}}$, $F_4 \in \mathbb{R}^{512 \times \frac{H}{32} \times \frac{W}{32}}$ 。将第三层特征上采样到第二层

分辨率相应位置的映射操作, Ω 是索引操作。

3) 边界框检测分支与损失函数

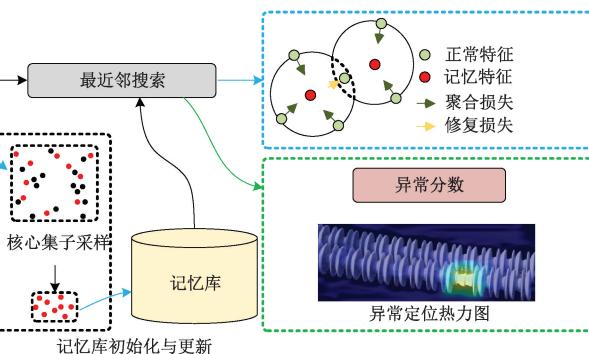
本文在 Mask2Former 的预测模块中添加边界框检测 head。与 mask 检测类似, 利用具有两个隐藏层的多层感知机将每段查询嵌入转换为 N 个边界框嵌入, 分别预测方框的标准化中心坐标、高度和宽度。相应地, 损失计算包括 3 个部分: 分类损失、mask 损失和边界框损失。对于分辨率为 $H \times W$ 的输入图像, $y = \{p_i, m_i, b_i\}_{i=1}^N$ 表示输出的类别、mask 和边界框预测集合, 给定真值 $y^{gt} = \{(c_i^{gt}, m_i^{gt}, b_i^{gt})\}_{i=1}^N, c_i^{gt} \in \{1, \dots, K\}, m_i^{gt} \in \{0, 1\}^{H \times W}, b_i^{gt} \in [0, 1]^4$, 并对 y 和 y^{gt} 进行匹配 s , 则总损失的计算方式如下:

$$l_{total} = l_{cls} + l_{mask} + l_{bbox} \quad (4)$$

其中, 分类损失 l_{mask} 为交叉熵损失, mask 损失 l_{mask} 由二值交叉熵损失和 dice 损失组成, 边界框损失 l_{bbox} 由 L1 损失和 GIoU(Generalized Intersection over Union)损失组成。

1.2 基于图特征记忆的缺陷检测方法

如图 2 所示, 本文提出了一种基于图特征记忆的电力杆塔部件缺陷检测方法。在初始化阶段, 提取目标数据集的不同尺度特征, 将第三层特征上采样后与第二层特征级联, 并将其投影为具有图结构的特征, 利用图卷积层-图注意力层-图卷积层的形式对其进行推理, 反投影后与原特征相加, 构建数据集导向的图特征描述符。通过核心集子采样提取正常样本图特征描述符中的关键元素, 从而初始化记忆库。



特征的大小后获得 $F'_3 \in \mathbb{R}^{256 \times \frac{H}{8} \times \frac{W}{8}}$, 再通过级联操作将两个特征进行结合, 获得中级特征表示 F , 利用模块 $\varphi(\cdot)$ 构建图特征描述符。

将 F 变形为 $F' \in \mathbb{R}^{C \times N}$, 其中 $C = 384$ 表示通道数, $N = \frac{H}{8} \times \frac{W}{8}$ 表示节点数。

利用 $W \in \mathbb{R}^{K \times C}$ 和 $\sum \in \mathbb{R}^{K \times C}$ 对投影操作进行参数化^[14], W 的每列 w_k 都为第 k 个节点制定一个可学习的聚类中心, 每个节点表示的计算方式如下:

$$v_k = \frac{v'_k}{\|v'_k\|_2}, v'_k = \frac{1}{\sum_i q_k^i} \sum_i q_k^i (f_i - w_k) / \sigma_k \quad (5)$$

其中, $f_i \in \mathbb{R}^c, i \in [1, \dots, N]$ 为特征向量, σ_k 是 \sum 的列向量, v'_k 是特征向量 f_i 和 w_k 之间残差的加权平均值。 v_k 代表第 k 个节点的表示。 q_k^i 是特征向量 f_i 对 w_k 的软分配。

然后利用一层图卷积层增强图特征的表示, 图卷积层的描述公式如下:

$$\begin{aligned} V_1 &= f_{g\text{-conv}}(V; W_g^1) = (A_1 \times V^T \times W_g^1)^T \\ A_1 &= \text{softmax}(V^T \times V) \end{aligned} \quad (6)$$

其中, $W_g^1 \in \mathbb{R}^{C \times C}$ 为图卷积层的可学习参数, $V^T \times W_g^1$ 体现属性信息的仿射变换, 而邻接矩阵 A_1 通过计算节点间的相似性进行构造。

之后, 图注意力层在具有相似特征的节点之间建立联系。首先通过测量节点表示之间的亲和度计算图的相似度矩阵 S , 并设置一个阈值 th 来过滤相似度较低的节点联系, 从而构建邻接矩阵 $A(x, y)$, 然后将 V_1 视为节点的集合 $V_1 = \{v_1, v_2, \dots, v_K\}$, 利用线性变换处理每个节点 $v_i \in \mathbb{R}^c$ 以学习属性信息, 并计算节点与邻域其他节点之间的权重系数:

$$\omega_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^T [v'_i \| v'_j]))}{\sum_{v_j \in N_{v_i}} \exp(\text{LeakyReLU}(\mathbf{a}^T [v'_i \| v'_j]))} \quad (7)$$

其中, ω_{ij} 衡量节点 v_j 对于节点 v_i 的重要性, $\mathbf{a} \in \mathbb{R}^{2C \times 1}$ 为可学习参数。 N_{v_i} 表示满足 $A(x, y) = 1$ 的节点, 即节点 v_i 的邻域。计算权重系数后, 得到节点的新特征向量 v''_i , 而输出的图特征也可以表示为 $V_2 = \{v''_1, v''_2, \dots, v''_K\}$, 最后再利用一层图卷积层进行特征学习得到 V_3 , 再经过反投影映射回原始坐标空间, 将此前计算的软分配组合为分配矩阵 Q , 则反投影后的图特征描述符 $\hat{F} \in \mathbb{R}^{N \times C}$ 可以通过下式获得:

$$\hat{F} = Q \times V_3^T + F'^T, Q \in \mathbb{R}^{N \times K} \quad (8)$$

2) 记忆库构建

本文引入贪婪核心集子采样算法来减小记忆库 \mathbf{M} , 选用基于 minimax facility location 的核心集子采样方法^[15], 保证采样后的核心集记忆库 \mathbf{M}_c 在描述符特征空间中的大致覆盖范围与原记忆库相似, 即:

$$\mathbf{M}_c^* = \underset{\mathbf{M}_c \in \mathbf{M}}{\text{argminmax}} \underset{m \in \mathbf{M}}{\min} \underset{n \in \mathbf{M}_c}{\min} \|m - n\|_2 \quad (9)$$

为减少核心集子采样消耗的时间, 通过随机线 $\psi: \mathbb{R}^d \rightarrow \mathbb{R}^{d^*} (d^* < d)$ 降低记忆特征的维数。记忆库构建方法的伪代码如算法 1 所示, 其中核心集大小为记忆库大小的 $n\%$, 在本文算法中设为 10%。

3) 异常定位的分数地图计算

训练阶段, 通过图特征描述符 \hat{F} 和记忆库 \mathbf{M} 的最近邻搜索获得 j 个最近邻 $f_n^j, n \in [1, \dots, N]$, 以 f_n^j 为中心构建半径为 r 的超球面^[16], 监督并将其嵌入到超球面。则聚合损失函数通过将罚值加上, 直到半径 r 远离 f_n^j :

算法 1: 记忆库算法

```

输入: 正常样本数据集  $Z$ , 核心集大小  $g$ , 随机线性投影
 $\psi$ 
输出: 记忆库  $\mathbf{M}$ 
方法:
 $\mathbf{M} \leftarrow \{\}$ 
for  $z_i \in Z$  do
     $\mathbf{M} \leftarrow \mathbf{M} \cup \varphi(z_i);$ 
end
 $\mathbf{M}_c \leftarrow \{\}$ 
for  $i \in [0, \dots, g - 1]$  do
     $m_i \leftarrow \underset{m \in \mathbf{M} - \mathbf{M}_c}{\text{argmax}} \underset{n \in \mathbf{M}_c}{\min} \|\psi(m) - \psi(n)\|_2$ 
     $\mathbf{M}_c \leftarrow \mathbf{M}_c \cup \{m_i\};$ 
end
 $\mathbf{M} \leftarrow \mathbf{M}_c$ 

```

$$\ell_{att} = \frac{1}{NJ} \sum_{n=1}^N \sum_{j=1}^J \max\{0, D(\hat{F}, f_n^j) - r^2\} \quad (10)$$

其中, J 表示与 \hat{F} 的最近邻的数量, $D(\cdot)$ 是一个预定的距离度量, 采用的是欧式距离。

将 \hat{F} 易被错估计为异常特征的特征定义为 hard negative 特征。将 hard negative 特征定义为通过最近邻搜索与记忆库 \mathbf{M} 匹配的 \hat{F} 的 $J+1$ 个最近邻 f_n^l , 定义修复损失函数进行对比监督:

$$\ell_{rep} = \frac{1}{NL} \sum_{n=1}^N \sum_{l=1}^L \max\{0, r^2 - D(\hat{F}, f_n^l) - a\} \quad (11)$$

其中, L 为对比监督的 hard negative 特征的总数, 超参数 a 为协调两个损失函数的比例, 设为 0.1。最后结合两个损失函数得到 ℓ , 共同优化 CNN (Convolutional Neural Networks) 和 $\varphi(\cdot)$ 的参数。

测试阶段, 首先计算输入图像 x^{test} 的图特征描述符 $\varphi(x^{test})$, 然后基于已构建的图特征描述符记忆库 \mathbf{M} , 通过最大化图特征 $m^{test} \in \varphi(x^{test})$ 到 \mathbf{M} 中每个最近邻 m^* 之间的距离分数 s^* 来估计图像级别的异常分数:

$$m^{test, *}, m^* = \underset{m^{test} \in \varphi(x^{test})}{\text{argmax}} \underset{m \in \mathbf{M}}{\text{argmin}} \|m^{test} - m\|_2 \quad (12)$$

为获得 s , 使用重新加权方法替代直接计算最大距离的方法, 若记忆库特征最接近已经远离邻近样本的异常候选样本 $m^{test, *}, m^*$, 那么则增加异常分数:

$$s = \left(1 - \frac{\exp \|m^{test, *} - m^*\|_2}{\sum_{m \in \mathbb{N}_b(m^*)} \exp \|m^{test, *} - m\|_2}\right) \cdot s^* \quad (13)$$

其中, $\mathbb{N}_b(m^*)$ 为测试图特征 m^* 在 \mathbf{M} 中的 b 个最近邻特征。通过 argmax 操作获取图像级的异常分数, 基于每个特征的空间位置重新排列, 基于此分割异常检测结果。最后通过双线性插值对结果进行上采样。

2 实验结果与分析

2.1 实例分割方法性能

1) 实验目的

输电线路航拍图像的实例分割不仅能对电力部件进行定位与分类,还能够提取出设备的实例信息,从而能够在后续任务中完成逐设备的缺陷检测。因此,利用无人机在输电线路进行多次飞行,采集不同焦距可见光图像,人工标注构建多组不同的实例分割数据集,与现有多款实例分割算法进行比较,从定性与定量两个方面充分验证本文所提出的设备实例分割性能。

2) 评估数据集和实施细节

(1) 数据集构建与评价指标

本文使用自建电力杆塔部件实例分割数据集评估所提实例分割方法,数据集中包括 452 张远距离杆塔图像和 206 张的近距离杆塔图像。通过 LabelMe 软件人工标注,并按照 COCO 实例分割^[17]的格式制成长焦成像数据集(LFID)

和短焦成像数据集(SFID)。采用边界框平均精度(average precision, AP)和 mask 的交并比(intersection over union, IoU)分别评估预测边界框和 mask 的平均精度。

(2) 实施细节

本文算法使用 PyTorch 框架实现,并在配备英特尔酷睿 i7、英伟达 RTX-2080 GPU 和 32 GB 内存的计算机上运行。采用 AdamW 优化器和分布学习率策略,迭代次数为 360 000 次。初始学习率和权重衰减分别设置为 0.000 1 和 0.05。使用了大规模抖动增强^[18]进行数据增强。

3) 消融实验

为了验证所提出的 3 个改进点的有效性,通过定量分析来评估各改进点的性能,实验结果如表 1 所示。在 SFID 和 LFID 数据集中,3 个改进点都使边界框 AP 和 mask AP 得到了增强,其中第 3 个改进点,即边界框检测分支和损耗优化使边界框 AP 得到了显著改善,在 SFID 数据集和 LFID 数据集中边界框 AP 分别提升了 3.2% 和 2%。实验结果证明本文提出的创新模块能有效改善实例轮廓的分割结果。

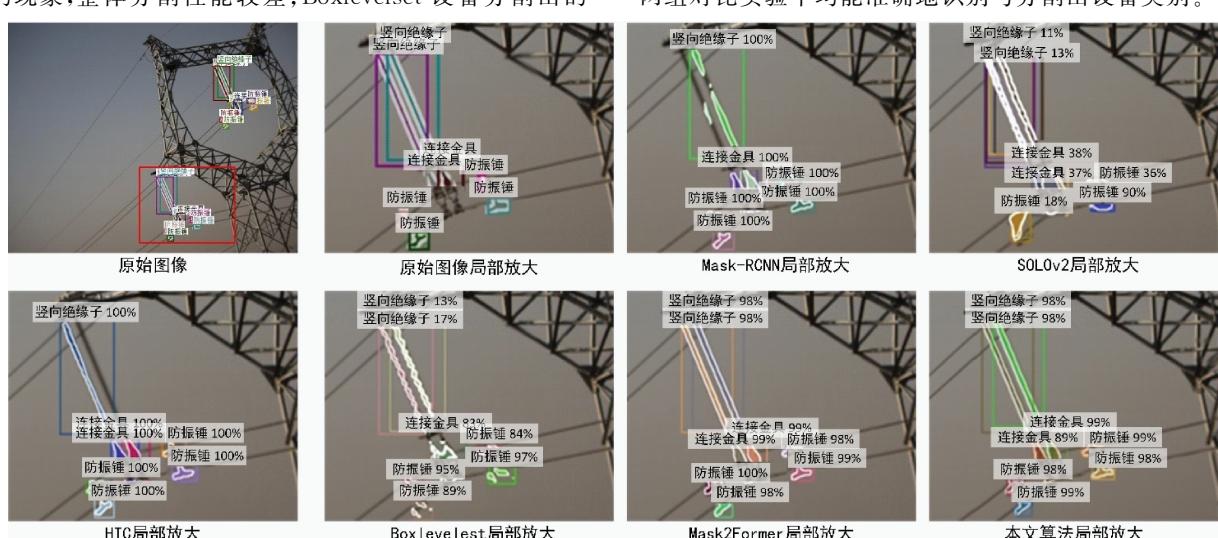
表 1 消融实验

Method	改进点 1	改进点 2	改进点 3	SFID		LFID	
				bbox AP	mask AP	bbox AP	mask AP
Mask2Former	√	√	√	59.9	58.9	20.9	18.6
				62.3	62.1	22.3	23.5
		√	√	61.7	61.3	21.8	22.6
				63.1	60.8	22.9	21.7
本文算法	√	√	√	65.6	63.4	25.0	26.0

4) 对比试验分析

从图 3 和 4 SFID 数据集中 A、B 两组图像的可视化对比结果可以看出,A 组对比实验中,Mask-Rcnn 在分割两组竖向绝缘子时都存在严重的漏分割和分割不全的问题;SOLOv2 漏分割了一个防振锤,并且识别和分割均存在重叠的现象,整体分割性能较差;Boxlevelset 设备分割出的

设备边缘不平滑,无法分割出完整的防振锤类别。B 组对比实验中 SOLOv2 的检测框识别存在严重的偏离和重叠现象;Boxlevelset 分割出的横向绝缘子的边缘轮廓不平滑,且漏识别了一个防振锤;Mask2Former 则出现将吊锤电线误识别为防振锤的问题。本文所提实例分割算法在两组对比实验中均能准确地识别与分割出设备类别。



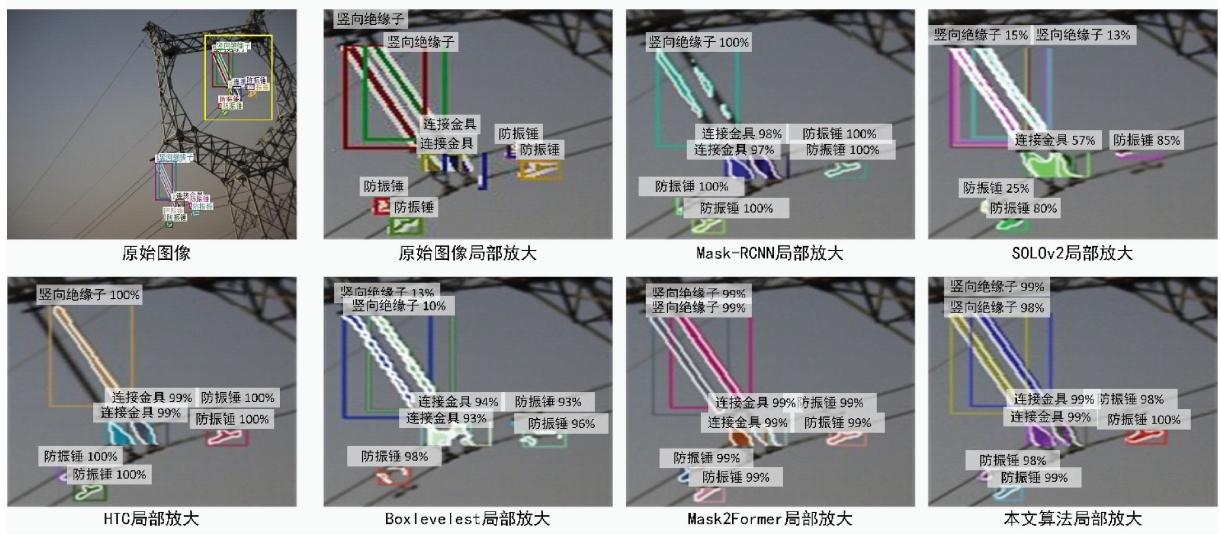


图3 SFID数据集A组对比结果

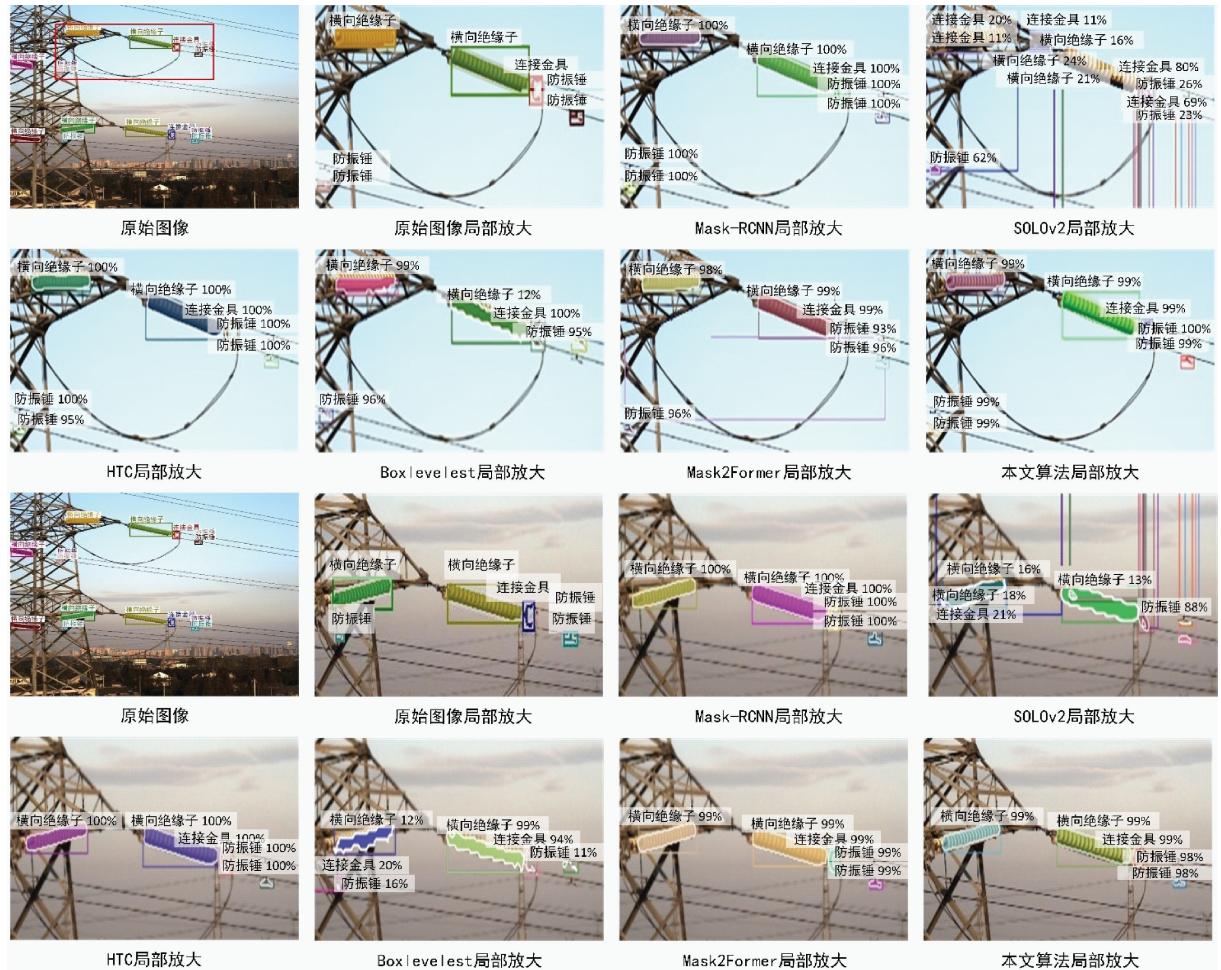


图4 SFID数据集B组对比结果

表2显示了所提方法与其他实例分割方法在SFID数据集上的定量实验结果。在竖向绝缘子类别中,本文算法的边界框AP和mask AP相对于Mask2Former分别提升了7.6%和0.5%。对于横向绝缘子、连接金具和防振锤3个类别,由于HTC网络包含区域建议网络,提取预选边界框更好,因此本文方法的边界框AP略低于

提升了7.6%和0.5%。对于横向绝缘子、连接金具和防振锤3个类别,由于HTC网络包含区域建议网络,提取预选边界框更好,因此本文方法的边界框AP略低于

HTC 方法。表 2 中列出了各算法的 AP_{50} 、 AP_{75} 、 AP_s 、 AP_m 和 AP_l 五项指标的对比结果。本文算法在所有指标的 mask AP 上都胜出。在 AP_{50} 的 mask AP 上, 本文算法仅比次优算法提升 1.0%, 但在 AP_{75} 上却提高了 5.3%, 这表明在对分割精度的要求较低时, 其他算法与本文算

法没有较大差距, 但是当提高分割精度的要求时, 本文算法有较大优势。在 AP_s 、 AP_m 和 AP_l 三个尺度的指标下, 本文算法分别领先次优算法 4.4%、3.8% 和 4.5%, 显示出本文实例分割方法所提出的多尺度 Transformer 解码器的优越性。

表 2 SFID 数据集定量实验指标对比结果

对比方法	横向绝缘子	竖向绝缘子	连接金具	防振锤	AP_{50}	AP_{75}	AP_s	AP_m	AP_l
bbox mask									
Mask-RCNN	67.8	70.9	55.7	38.0	60.3	58.3	60.8	55.0	91.1
SOLov2	45.3	49.1	17.4	20.6	10.1	13.3	6.6	8.1	55.3
HTC	70.3	71.5	61.8	40.1	61.2	57.3	64.9	56.8	91.3
Boxlevelset	70.0	43.5	62.0	15.3	47.8	18.6	32.3	1.5	79.4
Transfiner	67.5	62.0	60.0	36.1	60.9	46.7	63.6	42.1	90.7
Mask2former	65.5	69.5	72.7	61.7	54.8	54.5	46.3	50.1	84.3
本文算法	68.8	72.2	80.3	62.2	57.4	61.1	56.0	58.1	89.7
									91.6
									76.4
									76.0
									68.4
									66.0
									61.2
									56.4
									61.9
									58.8

图 5、6 为 LFID 中两组图像的可视化对比结果。A 组对比实验结果中, SOLOv2 和 Boxlevelset 将多个实例识别为同一实例, 且分割出的设备轮廓边缘较为粗糙; Mask RCNN 难以识别和分割被遮挡的部件; Mask2Former 受到部件遮挡的影响, 边界框定位不准确, 部件分割不完整。本文算法分割结果更完整, 部件轮廓边缘更精细。LFID

数据集的定量实验结果如表 3 所示, 本文所提方法在横向绝缘子、竖向绝缘子、连接金具、防振锤四个类别的 mask AP 指标上相较于次优算法分别有 2.8%、5%、5.9% 和 1.7% 的提升, 边界框 AP 仅在横向绝缘子类别上低于 HTC, 在其他设备类别上均取得了明显的优势, 说明本文算法在对分割精度要求更高的场景下具有很大的优势。

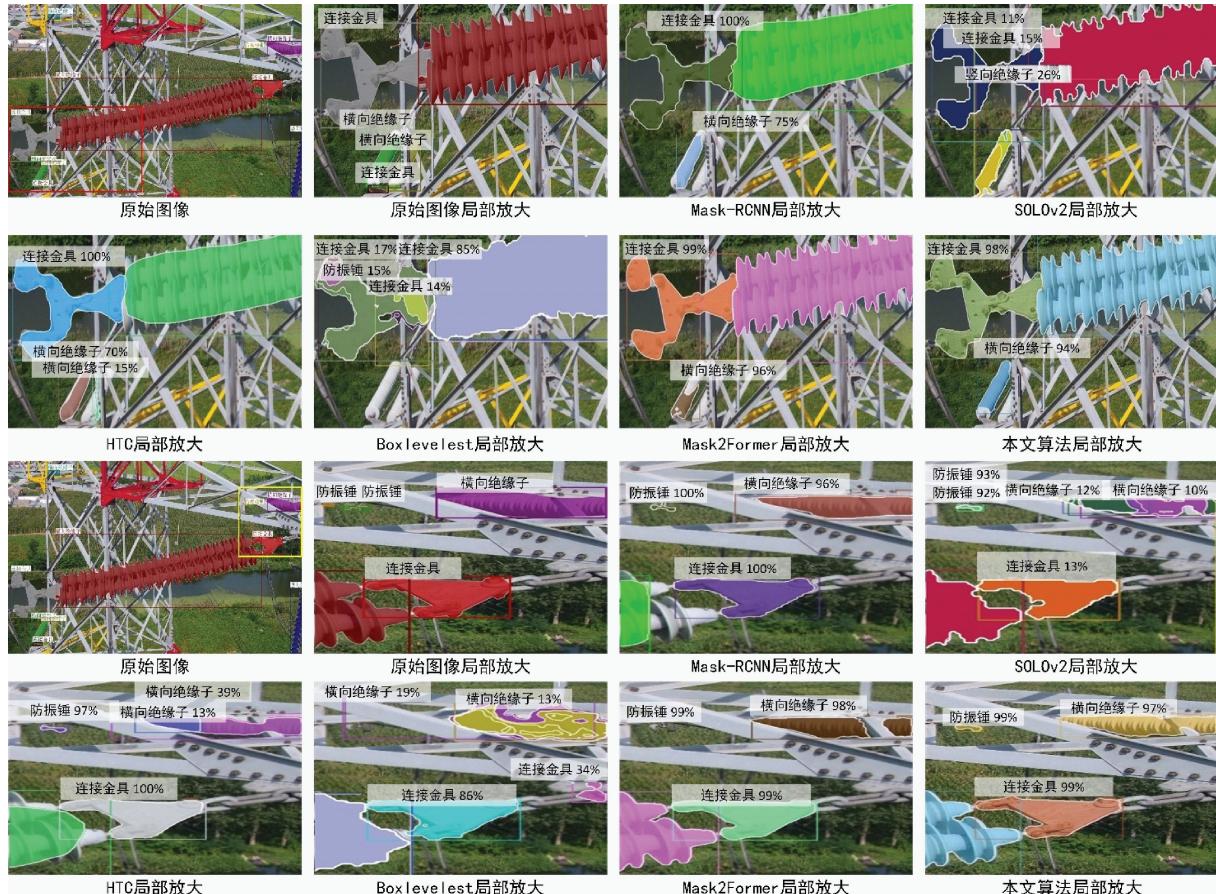


图 5 LFID 数据集 A 组对比结果



图 6 LFID 数据集 B 组对比结果

表 3 LFID 数据集定量实验指标对比结果

对比方法	横向绝缘子	竖向绝缘子	连接金具	防振锤	AP ₅₀	AP ₇₅	AP _S	AP _M	AP _L									
	bbox	mask	bbox	mask	bbox	mask	bbox	mask	bbox									
MaskRCNN	25.0	32.6	14.0	15.8	8.1	8.1	19.1	20.2	49.2	42.1	36.7	36.1	17.9	11.5	16.1	7.8	7.8	56.9
SOLOv2	20.4	24.0	15.2	10.7	8.0	8.0	13.9	14.3	21.1	29.0	28.2	13.4	11.1	7.4	3.1	1.8	1.8	28.3
HTC	31.4	32.4	10.9	20.5	5.1	4.4	22.6	22.0	54.4	42.8	35.5	37.7	19.9	12.9	20.7	9.6	9.6	59.5
Boxlevelset	17.0	13.0	1.1	5.0	1.7	1.3	10.5	5.0	16.5	16.1	9.1	10.0	2.7	0.8	1.8	0.1	0.1	28.9
Transfiner	24.8	20.9	6.8	13.8	1.3	1.3	12.0	6.1	33.2	21.5	15.6	16.8	5.7	2.1	7.9	2.7	2.6	44.0
Mask2former	20.2	25.4	18.7	17.8	13.0	12.9	18.9	18.2	29.8	38.4	38.0	21.4	14.6	13.7	9.1	4.5	4.5	39.8
本文算法	22.0	35.4	29.8	25.5	18.8	18.8	25.2	23.7	34.1	43.1	42.5	25.1	23.4	22.6	17.1	10.1	10.1	47.4

2.2 缺陷检测方法性能

1) 实验目的

基于上述实例分割结果,排除复杂背景像素干扰,定位设备本身的缺陷。电力设备种类众多且形态各异,因此,本节中首先利用实例分割结果提取所有设备实例,针对每种设备分别构建缺陷检测数据集,通过与现有各类型缺陷检测算法的效果对比,验证提出缺陷检测算法的泛化性与优异性能。

2) 评估数据集和实施细节

(1) 数据集构建与评价指标

本文设置横向绝缘子、竖向绝缘子、连接金具和防振锤四类作为缺陷检测类别。其中,横向绝缘子、竖向绝缘子和防振锤的正常样本数量为 194 张、210 张和 297 张,异

常样本数量分别为 58 张、70 张和 116 张。连接金具含有 374 张正常样本和 116 张异常样本。本文采用 AUROC 和 F1-Score 作为评估指标。

(2) 实施细节

采用 AdamW 优化器和 AMS-Grad 优化方法对网络训练了 30 个 epoch。初始学习率和权重衰减分别设置为 0.001 和 0.0005。最近邻数 J 和损失函数中的最近邻数 L 均设为 3,超球半径 r 为 1×10^{-5} ,超参数 a 设置为 0.1,输入数据的比例统一为 256×256 。

3) 实验比较

(1) 可视化比较

图 7、8、9 和 10 分别显示了横向绝缘子、竖向绝缘子、连接金具和防振锤的 4 组缺陷检测对比结果。在横向绝

缘子、竖向绝缘子、防振锤缺陷检测结果中,其他方法存在误检测的情况。当存在如图 8(b)组图像所示的绝缘子遮挡且背景干扰时,STFPM、FastFlow 和 DraemSS 将背景

区域视为异常特征,存在误检和漏检现象。在连接金具结果对比中,其他方法存在过度分割等问题。本文提出的方法定位更准确,更关注异常区域,分割结果更精确、完整。

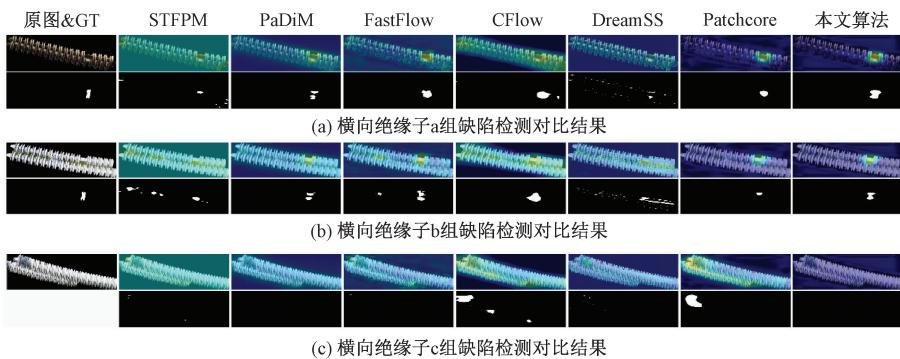


图 7 横向绝缘子缺陷检测结果对比

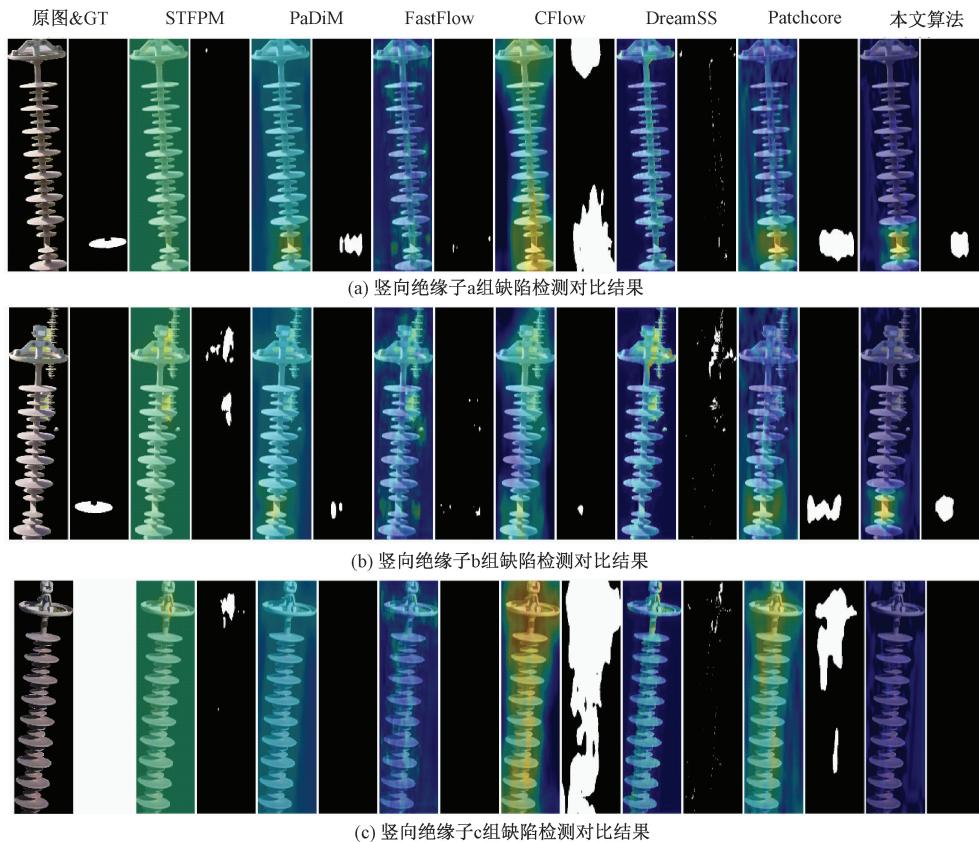
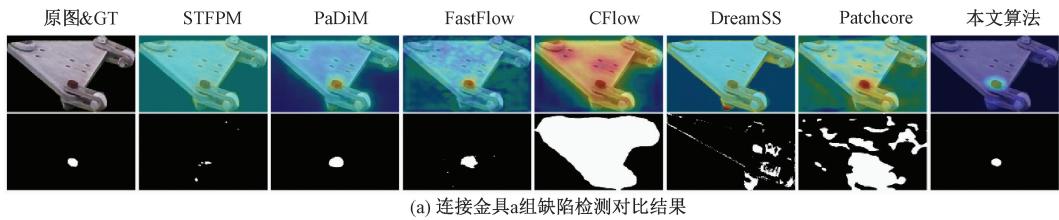
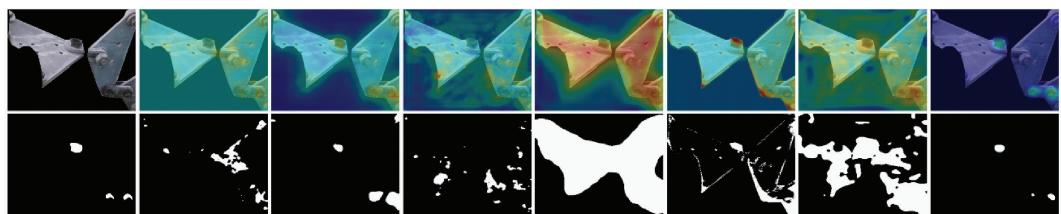
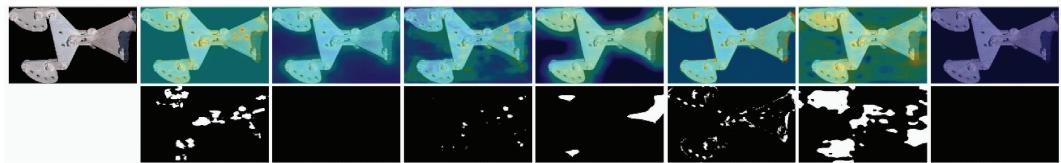


图 8 竖向绝缘子缺陷检测结果对比





(b) 连接金具b组缺陷检测对比结果



(c) 连接金具c组缺陷检测对比结果

图9 连接金具缺陷检测结果对比

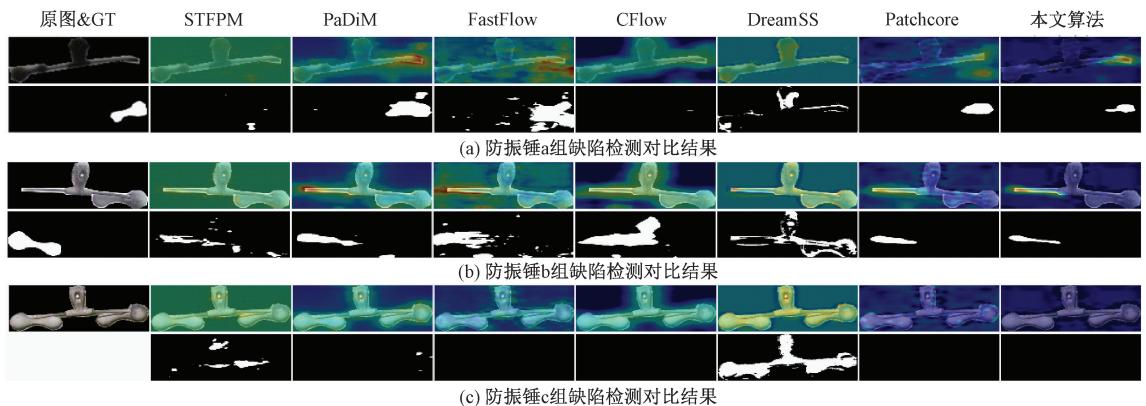


图10 防振锤缺失检测结果对比

(2) 定量比较

如表4所示,在横向绝缘子、竖向绝缘子和连接金具3个类别的异常检测中,本文方法在像素AUROC和F1-

Score上都有明显优势,但在防振锤类别中分别比FastFlow低9.7%和25.8%。原因在于分割异常时, FastFlow倾向于将异常值不高的区域也分割出来。

表4 缺陷检测数据集AUROC指标和F1-Score指标(%)对比实验

指标类别	AUROC指标								F1-Score指标							
	横向绝缘子		竖向绝缘子		连接金具		防振锤		横向绝缘子		竖向绝缘子		连接金具		防振锤	
对比方法	image	pixel														
STFPM	73.6	87.4	36.6	73.5	65.7	89.9	89.0	66.6	52.9	10.1	80.4	1.6	81.8	11.3	90.1	8.4
PaDiM	90.1	95.3	89.2	90.9	84.4	90.3	96.8	74.0	90.1	9.4	97.2	5.1	84.7	16.7	91.3	22.4
FastFlow	85.0	93.4	59.0	73.2	49.9	77.9	99.4	90.3	59.2	10.6	80.5	3.7	81.3	7.8	97.4	42.9
CFlow	91.2	95.8	67.2	78.0	44.6	81.4	79.3	74.7	59.8	22.7	81.4	3.2	81.7	7.2	89.2	18.3
DraemSS	63.3	75.2	54.3	55.9	56.4	90.0	69.6	49.0	55.6	5.7	80.4	2.1	81.1	14.5	79.4	3.1
Patchcore	95.1	98.3	85.3	92.1	76.1	89.6	98.5	80.3	91.4	33.4	90.9	7.0	82.1	6.1	93.6	19.3
本文算法	99.1	99.0	95.6	93.1	92.7	95.1	99.8	80.6	96.6	39.9	98.6	10.5	95.8	17.7	98.2	17.1

从图10可以看出, FastFlow虽然分割结果中包含缺陷的真实位置,但是也将正常区域错判为异常。本文算法的AUROC和F1-Score指标在横向绝缘子、竖向绝缘子、连接金具四个类别上相较于次优算法取得了明显的优势。综上,对于电力杆塔关键部件缺陷检测数据集,本文所提

出的算法具有更好的性能。

3 结论

本文提出了一种电力杆塔组件缺陷检测算法,将动态位置查询引导的多尺度实例分割方法与基于图特征记忆

的缺陷检测方法相结合,完成电力杆塔组件的分割和缺陷检测。具体来说,提出的动态位置查询引导的多尺度实例分割方法利用动态位置查询、多尺度特征提取和损失函数优化来识别不同的实例,根据分割掩码区分预测帧中的零件和背景区域,消除干扰,为缺陷检测提供高质量的输入源。提出的一种基于图特征记忆的缺陷检测方法利用图特征描述器、软边界回归的损失函数和构建的记忆库得到异常区域的具体位置,解决提取特征的然而,在实际的应用过程中,可能受到天气、光照或遮挡等因素的影响,导致分割设备区域的区分度较低,缺陷检测算法容易将不同设备的遮挡误检为缺陷。因此,后续研究需要考虑改进对遮挡区域的表示方案,解决该类型误检问题。

参考文献

- [1] 龙珊珊,信瑞山. 基于人工智能图像识别的输电线路巡检研究[J]. 电子测量技术,2023,46(6):116-121.
- [2] QIU L, HE L, LU H, et al. Systematic potential analysis on renewable energy centralized co-development at high altitude: A case study in Qinghai-Tibet plateau [J]. Energy Conversion and Management, 2022, 267: 115879.
- [3] XU C, LI Q, ZHOU Q, et al. Power line-guided automatic electric transmission line inspection system[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-18.
- [4] YU Q, LIU H, HE Q, et al. Research on pressure-based detection technology for partial overheat insulation defect of oil-less power equipment[J]. IOP Publishing, 2021, 632(4): 042009.
- [5] 赵博,马宏忠,张潇,等. 定向识别航拍绝缘子及其缺陷检测方法研究[J]. 电子测量与仪器学报,2023,37(5): 240-251.
- [6] YANG H, LI X, GUO Y, et al. Discretization-filtering-reconstruction: Railway detection in images for navigation of inspection UAV [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-13.
- [7] 隋宇,宁平凡,牛萍娟,等. 面向架空输电线路的挂载无人机电力巡检技术研究综述[J]. 电网技术,2021,45(9): 3636-3648.
- [8] BOLYA D, ZHOU C, XIAO F, et al. Yolact: Real-time instance segmentation [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision(ICCV), 2019: 9157-9166.
- [9] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]. Proceedings of the European Conference on Computer Vision(ECCV), 2020: 213-229.
- [10] CHENG B, MISRA I, SCHWING A G, et al. Masked-attention mask transformer for universal image segmentation [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2022: 1290-1299.
- [11] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 14318-14328.
- [12] PARK H, NOH J, HAM B. Learning memory-guided normality for anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 14372-14381.
- [13] 刘行谋,田浩,杨永明,等. 复杂环境背景下绝缘子缺陷图像检测方法研究[J]. 电子测量与仪器学报,2022,36(2): 57-67.
- [14] LI Y, GUPTA A. Beyond grids: Learning graph representations for visual recognition[C]. Proceedings of the International Conference on Neural Information Processing Systems(NeurIPS), 2018: 9245-9255.
- [15] SINHA S, ZHANG H, GOYAL A, et al. Small-gan: Speeding up gan training using core-sets [C]. Proceedings of the International Conference on Machine Learning(ICML), 2020: 9005-9015.
- [16] YI J, YOON S. Patch svdd: Patch-level svdd for anomaly detection and segmentation[C]. Proceedings of the Asian Conference on Computer Vision(ACCV), 2020: 375-390.
- [17] LIN T, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context[C]. Proceedings of the European Conference on Computer Vision(ECCV), 2014: 740-755.
- [18] GHIASI G, CUI Y, SRINIVAS A, et al. Simple copy-paste is a strong data augmentation method for instance segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2021: 2918-2928.

作者简介

- 张关应,高级工程师,主要研究方向为工程项目、高压电力运行与维护。
E-mail: dwzhangguanying@163.com
- 束云豪,工程师,主要研究方向为电网工程项目、高压电力运行与维护。
E-mail: shuyunhaodw@163.com
- 常宸铠,硕士研究生,主要研究方向为输配电智能感知。
E-mail: hhucck@hhu.edu.cn
- 候姝斌,工程师,主要研究方向为电网工程项目、高压电力运行与维护。
E-mail: Houshubinnn@163.com
- 李庆武(通信作者),教授,主要研究方向为视觉感知与人工智能、输配电智能感知、水下成像探测、传感网及其应用。
E-mail: li_qingwu@163.com