



面向复杂场景的变电设备锈蚀检测方法

赵振兵, 席悦, 冯烁, 赵文清, 翟永杰, 李冰

引用本文:

赵振兵, 席悦, 冯烁, 等. 面向复杂场景的变电设备锈蚀检测方法[J]. *智能系统学报*, 2025, 20(3): 679-688.

ZHAO Zhenbing, XI Yue, FENG Shuo, et al. Corrosion detection method for complex scenario substation equipment[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(3): 679-688.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202403044>

您可能感兴趣的其他文章

结合模糊特征检测的鲁棒核相关滤波跟踪法

Robust KCF tracking algorithm combined with fuzzy feature detection

智能系统学报. 2021, 16(2): 323-329 <https://dx.doi.org/10.11992/tis.201912010>

基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152-1162 <https://dx.doi.org/10.11992/tis.201812003>

全钢子午线轮胎X光图像的缺陷检测研究现状

Research status of defect detection in X-ray images of all-steel radial tires

智能系统学报. 2019, 14(4): 793-803 <https://dx.doi.org/10.11992/tis.201806014>

BP神经网络和支持向量机相结合的电容器介损角辨识

Capacitor dielectric loss angle identification based on a BP neural network and SVM

智能系统学报. 2019, 14(1): 134-140 <https://dx.doi.org/10.11992/tis.201805034>

基于显著性检测的双目测距系统

Binocular distance measurement system based on saliency detection

智能系统学报. 2018, 13(6): 913-920 <https://dx.doi.org/10.11992/tis.201712005>

基于Object Proposals并集的显著性检测模型

Saliency detection model based on the union of Object Proposals

智能系统学报. 2018, 13(6): 946-951 <https://dx.doi.org/10.11992/tis.201801009>

DOI: 10.11992/tis.202403044

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250414.1644.002>

面向复杂场景的变电设备锈蚀检测方法

赵振兵^{1,2,3}, 席悦¹, 冯烁¹, 赵文清^{2,4}, 翟永杰⁴, 李冰⁴

(1. 华北电力大学 电子与通信工程系, 河北 保定 071003; 2. 华北电力大学 复杂能源系统智能计算教育部工程研究中心, 河北 保定 071003; 3. 华北电力大学 河北省电力物联网技术重点实验室, 河北 保定 071003; 4. 华北电力大学 控制与计算机工程学院, 河北 保定 071003)

摘要: 针对复杂场景下变电设备锈蚀检测中存在锈蚀形态差异大、尺度大小不一、特征显著性低等问题, 提出了一种面向复杂场景的变电设备锈蚀检测方法。引入了频率通道注意力机制, 通过更多的频率分量补充深层网络中的细节特征, 优化模型对锈蚀特征的提取; 在特征融合网络中使用多尺度特征增强模块重新构建 C2f 模块, 使网络可以更好地捕获不同大小的锈蚀区域; 引入附加检测头, 缓解模型在特征融合过程中因卷积层下采样造成的锈蚀关键信息丢失的情况, 从而提高变电设备锈蚀检测的精度。实验结果表明, 改进以后的网络模型相较于原始的 YOLOv8m 模型, 平均检测精度 (mAP50) 提升了 5.1%, 检测效果也优于其他主流目标检测模型, 为变电设备锈蚀检测提供了新的参考方法。

关键词: 变电设备; 不规则缺陷; 锈蚀检测; YOLOv8; 注意力机制; 多尺度特征; 检测头; 复杂场景; 电力视觉
中图分类号: TP183 **文献标志码:** A **文章编号:** 1673-4785(2025)03-0679-10

中文引用格式: 赵振兵, 席悦, 冯烁, 等. 面向复杂场景的变电设备锈蚀检测方法 [J]. 智能系统学报, 2025, 20(3): 679-688.

英文引用格式: ZHAO Zhenbing, XI Yue, FENG Shuo, et al. Corrosion detection method for complex scenario substation equipment[J]. CAAI transactions on intelligent systems, 2025, 20(3): 679-688.

Corrosion detection method for complex scenario substation equipment

ZHAO Zhenbing^{1,2,3}, XI Yue¹, FENG Shuo¹, ZHAO Wenqing^{2,4}, ZHAI Yongjie⁴, LI Bing⁴

(1. Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071003, China; 2. Engineering Research Center of Intelligent Computing for Complex Energy Systems of Ministry of Education, North China Electric Power University, Baoding 071003, China; 3. Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding 071003, China; 4. School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: To address the significant challenges posed by significant variations in corrosion morphology, varying scales, and low feature saliency in corrosion detection in substation equipment under complex scenarios, a corrosion detection method for complex scenario substation equipment is proposed. First, frequency channel attention networks are introduced, leveraging additional frequency components to complement the fine-grained features in deep networks, thus optimizing the extraction of corrosion-related features. Second, within the feature fusion network, the multi-scale module is employed to rebuild the C2f module, thereby allowing the network to better capture corrosion areas of different sizes. Finally, the additional detection head mitigates crucial corrosion-related information loss during feature fusion caused by downsampling, thereby enhancing corrosion detection accuracy. The relevant experimental results reveal a 5.1% improvement in the average detection accuracy (mAP50/%) compared with the results obtained by the original YOLOv8m model. Overall, this improved model outperforms other mainstream object detection models, thereby representing a novel approach for substation equipment corrosion detection.

Keywords: substation equipment; irregular defect; corrosion detection; YOLOv8; attention mechanism; multi-scale feature; detection head; complex scenarios; power computer vision

收稿日期: 2024-03-28. 网络出版日期: 2025-04-14.

基金项目: 国家自然科学基金项目 (U21A20486, 62373151, 62371188, 62303184); 中央高校基本科研业务费专项资金项目 (2023JC006); 河北省自然科学基金项目 (F2021502008, F2021502013).

通信作者: 赵振兵. E-mail: zhaozhenbing@ncepu.edu.cn.

变电站负责对电压进行转换并对电能进行分配, 站内设备的运行状态直接关系到变电站甚至整个电网的安全稳定^[1]. 由于大多数变电设备由金属材料制成, 且长期暴露在日晒雨淋的室外环

境中工作,其表面或者部件连接处极易产生锈蚀,进而引发变电设备故障。常规的检测方式为工作人员定期对户外设备进行巡检,但存在检测效率低、检测结果易受主观因素影响的问题^[2]。随着智能电网的建设以及人工智能技术的发展,基于深度学习与缺陷检测相结合的方法逐渐成为主流^[3]。但是,对于电力设备外表面锈蚀缺陷检测这类问题,依然缺少有效的检测方法。

传统的检测方法如尺度不变特征变换 (scale-invariant feature transform, SIFT)^[4]、方向梯度直方图 (histogram of oriented gradient, HOG)^[5] 等受复杂背景影响大,对于电力设备锈蚀缺陷形态差异显著且尺度变化较大的情况,往往仅基于特定特征实现检测,导致识别准确率不高,缺乏可扩展性,难以适用于真实的电力场景^[6]。

近年来,基于深度学习的检测方法相较传统检测方法取得了更好的性能,并被应用到电力设备缺陷检测领域中。为了解决设备缺陷尺度变化大的问题,文献 [7] 通过增加一个小物体检测头来提高小物体的检测精度,同时采用 CIoU (complete-IoU)^[8] 损失函数和 K 均值聚类得到边界框,提高检测精度,但是仅使用一个小物体检测头来解决缺陷多尺度问题过于局限。针对变电设备锈蚀图像质量存在显著差异的问题,文献 [9] 提出将注意力机制融入 Faster R-CNN (faster region-based convolutional neural network)^[10] 模型,并考虑背景亮度和伽马变换的锈蚀检测方法,其能够自适应地规范化图像,有效识别复杂背景、光亮差异大和受噪声影响的情况下的锈蚀区域。虽然这种方法解决了图像质量参差的问题,但没有根据锈蚀特点进行针对性检测。针对电力设备外表面缺陷具有特征不明显、形态多变的问题,文献 [11] 提出一种基于域适应网络的变电设备渗漏油分割模型,通过引入域适应、纹理提取和辅助损失等方法实现了高精度的渗漏油分割,但需要大量的像素级标注训练样本。针对跨设备检测覆冰、锈蚀、污秽等表面缺陷问题,文献 [12] 提出了一种自适应感受野网络和基于真实纹理的电力设备表面缺陷数据合成方法,解决了不同电力设备尺度不一以及电力设备缺陷数据集较少且场景单一的问题。文献 [13] 通过改进 YOLOv4^[14] 骨干网络,解决电力设备目标检测中定位困难的问题,同时采用焦点损失函数代替交叉熵损失函数,促进样本不平衡引起的目标检测精度低问题的解决。为了有效识别电力设备生锈区域,文献 [15] 在 SSD (single shot multibox detector)^[16] 算法的基础上,添

加注意力机制,同时使用 ResNet^[17] 网络代替原有骨干网络。文献 [18] 将 SE (squeeze-and-excitation)^[19] 模块嵌入深度残差网络,构建基于通道注意力的金属锈蚀检测网络,由学习到的通道注意力权重来突出不同锈蚀区域的重要特征,提高对不同锈蚀程度区域的检测精度。

上述文献虽然实现了多场景锈蚀检测,但锈蚀缺陷样本的特异性未被有效关注。由于锈蚀几乎不存在完全相同的两个样本,其特征难以提取,现有工作并未对这一问题进行深入研究。变电设备锈蚀检测的关键在于如何能更好地学习或提取到锈蚀缺陷共性描述特征。然而,变电场景环境复杂,锈蚀的形状、边缘等特征不明显,不同设备的尺寸、外观等又会令缺陷区域产生一定差异,且锈蚀区域的长宽比存在较大的随机性,这些因素导致缺陷区域的特征提取困难。

本文对上述问题进行了深入研究,针对锈蚀形态差异大、特征不明显的问题,考虑到不同设备表面缺陷在图像纹理分布上的共性及特殊性,通过在骨干网络和特征融合网络之间引入频率通道注意力机制 (frequency channel attention networks, FcaNet)^[20],以补充更多的频率信息,强化深层网络的锈蚀细节特征;针对锈蚀区域面积大小不统一的问题,在特征融合网络中使用多尺度特征增强模块 (multi-scale module, MS-Module) 重新构建 C2f 模块为 MS-C2f (multi-scale C2f module),以获得更为丰富的层次信息,使网络可以更好地捕获不同大小的锈蚀区域;针对锈蚀关键信息易在深层网络丢失的问题,引入附加检测头 (additional detection head, AD-Head),从主网络生成的特征映射中提取额外的特征,改善卷积层下采样导致的信息丢失的情况,获得更加完整的特征信息,同时通过损失函数实现浅层与深层特征图的信息交互,从而提升网络的特征表达能力。本文提出的面向复杂场景的变电设备锈蚀检测方法有效提升了变电设备锈蚀检测精度,可为电力设备智能运维提供新的研究方法。

1 基于改进 YOLOv8 模型的变电设备锈蚀检测方法

YOLOv8 模型主要由输入端 (Input)、骨干网络 (Backbone)、颈部网络 (Neck) 和预测输出网络 (Head) 组成。本文模型的实现过程如图 1 所示,变电设备锈蚀缺陷图像首先在输入端调整到合适尺寸,随后输入至骨干网络进行锈蚀特征的基本提取;然后在骨干网络和颈部网络之间使用 FcaNet

强化锈蚀的细节特征。随后的颈部网络部分采用特征金字塔网络 (feature pyramid network, FPN)^[21] 和路径聚合网络 (path aggregation network, PAN)^[22], 通过自顶向下与自下向顶的跨层连接使锈蚀特征更充分融合。在自下向顶的过程中, 使用特征增强模块重构的 MS-C2f 模块, 增大特征图感受野, 使更多锈蚀缺陷信息进入有效感受

野, 提高大小有所差异的锈蚀区域的检测能力。最后, 将融合后的特征图输入到预测输出网络中, 使用解耦检测头以及在颈部网络添加的 AD-Head 共同进行损失计算和网络参数更新, 实现深浅网络的信息交互, 获取更加完整的锈蚀特征信息, 指导模型更好地学习变电设备锈蚀缺陷特征。

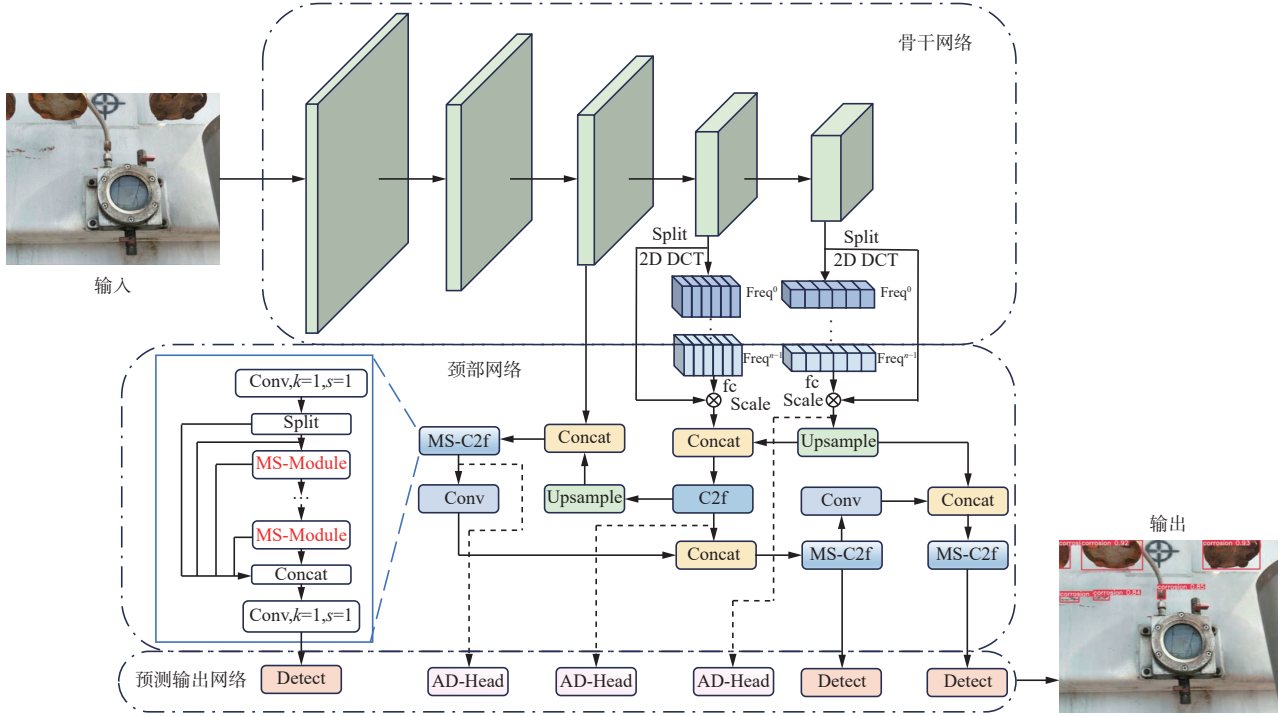


图 1 本文模型网络结构

Fig. 1 Model network structure of this paper

1.1 频率通道注意力机制

传统通道注意力机制使用标量来表示通道, 会造成大量的信息丢失, 无法充分获取每个通道的多样性的信息。频率通道注意力机制将通道表示问题视为基于频率分析的压缩过程, 通过引入了更多的频率分量, 可以得到更加丰富的信息。虽然电力设备上的锈蚀区域形态多变且特征不明

显, 但是锈蚀区域与非锈区域的粗糙程度通常存在差别。随着网络的加深, 特征图上丰富的细节信息会逐渐损失, 进而导致模型对锈蚀识别能力降低。为增强优化模型对锈蚀特征的提取能力, 本文引入了频率通道注意力机制 FcaNet, 引入更多的频率分量, 以补充锈蚀细节特征, 其结构如图 2 所示。

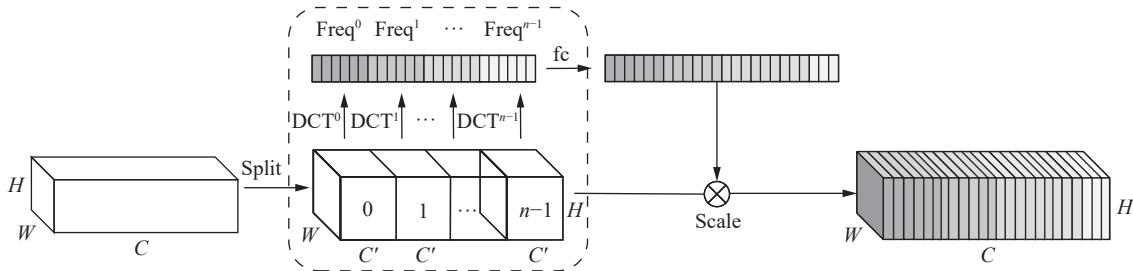


图 2 频率通道注意力机制

Fig. 2 Frequency channel attention networks

FcaNet 将通道表示问题视为基于频率分析的压缩问题, 从数学角度证明了全局平均池化 (global average pooling, GAP)^[23] 是二维离散余弦变

换 (discrete cosine transform, DCT) 的一个特例。离散余弦变换是离散傅里叶变换 (DFT) 的一种特殊形式, 相比 DFT 变换具有更好的频域能量聚集

度, 2D DCT^[24] 可以表示为

$$f_{h,w}^{2D} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} x_{i,j}^{2D} B_{h,w}^{i,j} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} x_{i,j}^{2D} \cos\left(\frac{\pi h}{H}\left(i + \frac{1}{2}\right)\right) \cos\left(\frac{\pi w}{W}\left(j + \frac{1}{2}\right)\right) \quad (1)$$

式中: $f^{2D} \in \mathbf{R}^{H \times W}$ 代表二维离散余弦变化的频谱, 输入特征图用 $x^{2D} \in \mathbf{R}^{H \times W}$ 表示, H 代表输入特征图的高, W 代表输入特征图的宽, $B_{h,w}^{i,j}$ 表示对输入特征图进行 2D DCT 的基函数, $h \in \{0, 1, 2, \dots, H-1\}$, $w \in \{0, 1, 2, \dots, W-1\}$ 。

GAP 仅保留最低频率分量, 而丢弃其他频率分量信息, 导致频域信息利用不充分, 不足以表征所有的特征信息。为融入更多的频率分量, 将输入 X 沿通道维度划分成 n 个部分, 用 $[X^0 \ X^1 \ \dots \ X^{n-1}]$ 表示, 其中 $X^i \in \mathbf{R}^{C \times H \times W}$, $i \in \{0, 1, \dots, n-1\}$, 每个通道为 $C' = C/n$, C 应为 n 的倍数。每个部分都分配一个对应的 2D DCT 频率分量, 对应的 2D DCT 频率分量可表示为

$$F^i = 2D \text{ DCT}^{u_i, v_i}(X^i) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X_{h,w}^i B_{h,w}^{u_i, v_i} \quad (2)$$

式中: X^i 表示输入特征图的第 i 个通道部分, $i \in \{0, 1, \dots, n-1\}$, $[u_i, v_i]$ 是对应 X^i 的频率分量的二

维索引, F^i 表示 X^i 对应的 2D DCT 频率分量, 每个频率分量 $F^i \in \mathbf{R}^C$ 变为 C' 维的向量, $X_{:,h,w}^i$ 表示高宽为 h, w 输入特征图的第 i 个通道部分。 $B_{h,w}^{u_i, v_i}$ 表示输入的第 i 个通道部分的二维离散余弦变化基函数。整个注意力框架可以写成

$$A = \sigma(\text{fc}(\text{cat}([F^0, F^1, \dots, F^{n-1}]))) \quad (3)$$

式中: A 表示经过注意力调整后的输出特征图, $\sigma(\cdot)$ 表示 Sigmoid 激活函数, $\text{cat}(\cdot)$ 表示对 n 个通道对应的频率分量进行拼接操作, $\text{fc}(\cdot)$ 表示如全连接层或一维卷积的映射函数。

1.2 多尺度特征增强模块

锈蚀会发生在设备的任何部位, 有些范围较大, 如设备表面、连接处, 有的范围相对较小, 如设备上的部件。由于锈蚀区域大小不统一, 可能其不同尺度的特征图上表现出的特征具有差异。本文在特征融合网络中引入多尺度特征增强模块 (MS-module) 重新构建 C2f 模块, 得到 MS-C2f (multi-scale C2f module, MS-C2f) 模块, 先采用自校准卷积 (self-calibrated convolutions, SCConv)^[25], 更精确地定位锈蚀区域, 再结合不同扩张率的并行卷积让网络充分利用不同尺度上的特征信息, 从而更全面地捕获锈蚀缺陷特征, 最终提高锈蚀缺陷的检测精度, 具体结构如图 3 所示。

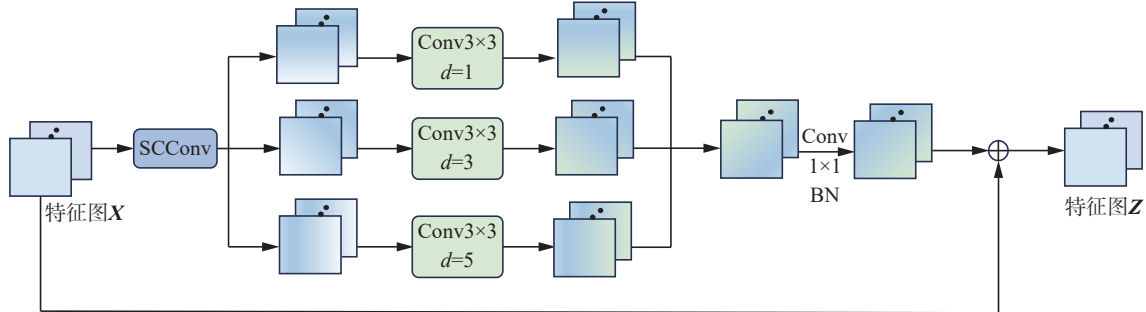


图 3 多尺度特征增强模块

Fig. 3 Multi-scale feature enhancement module

多尺度特征增强模块先采用自校准卷积, 能够自适应地在每个空间位置构建远距离空间和通道的依

赖关系, 从而生成更显著的锈蚀特征, 实现更精确和完整的锈蚀目标定位。其自校准卷积结构如图 4 所示。

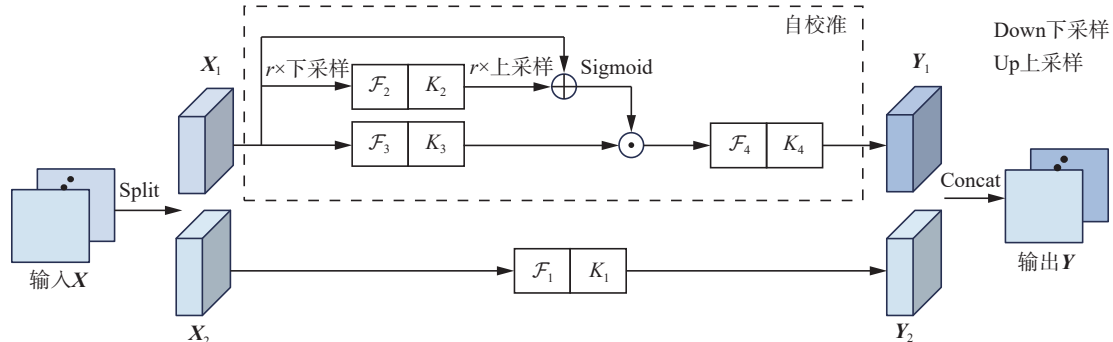


图 4 自校准卷积结构

Fig. 4 Structure of self-calibrated convolutions

首先将输入特征图 X 分成两部分 $\{X_1, X_2\}$, 将给定的一组卷积核 (C, C, k_h, k_w) 分成 4 部分, 其中输入通道与输出通道相同, 用 C 表示, k_h, k_w 分别表示空间高度和宽度。4 个部分别负责不同的功能, 记为 $\{K_1, K_2, K_3, K_4\}$, 维度均为 $(C/2, C/2, k_h, k_w)$ 。为高效收集每个空间位置的上下文信息, 分别在原始尺度空间 (输入共享相同的分辨率) 和下采样后的具有较小分辨率的潜在空间 (用于自校准) 进行卷积特征变换。对于 X_1 通路进行自校准操作, X_2 通路进行简单的卷积运算, $Y_2 = \mathcal{F}_1(X_2) = X_2 * K_1$ 。 X_1 首先进行平均池化, 然后再使用 K_2 进行特征变换, 校准操作可表示为

$$Y'_1 = \mathcal{F}_3(X_1) \cdot \sigma(X_1 + \text{Up}(\mathcal{F}_2(\text{AvgpPool}_r(X_1)))) \quad (4)$$

式中: Y'_1 表示校准后操作后的特征图; AvgpPool_r 表示池化核大小为 $r \times r$ 、步长为 r 的平均池化操作; $\mathcal{F}_2(\text{AvgpPool}_r(X_1)) = \text{AvgpPool}_r(X_1) * K_2$; $\text{Up}(\cdot)$ 表示双线性插值算子, 用于将中间参考量从小比例空间映射到原始特征空间; ‘ \cdot ’ 表示逐元素乘运算; $\mathcal{F}_3(X_1) = X_1 * K_3$ 。最终校准后的输出为

$$Y_1 = \mathcal{F}_4(Y'_1) = Y'_1 * K_4 \quad (5)$$

最终 SCConv 的输出为

$$Y = \text{cat}([Y_1, Y_2]) \quad (6)$$

式中: Y_1, Y_2 表示 X_1 通路和 X_2 通路的最终输出, $\text{cat}(\cdot)$ 表示拼接操作。多尺度特征增强模块进一

步采用多分支结构以扩展感受野, 通过扩张率 d 为 1、3、5 的并行卷积提取多尺度特征信息, 将得到的输出特征经过 BN 层处理后, 再采用逐点卷积对不同尺度的特征进行整体压缩和聚合。最后, 将输入特征与聚合后特征拼接, 以构建更全面的特征表示:

$$C_1(Y, d) = D_d \text{DCConv}(Y) \quad (7)$$

$$C(Y) = \text{PConv}(\text{BN}(\Gamma_d\{C_1(Y, d)\})) + X \quad (8)$$

式中: $D_d \text{DCConv}(\cdot)$ 表示大小为 3×3 、扩张率为 d 的扩张卷积, $C_1(Y, d)$ 表示 SCConv 输出特征 Y 经过扩张率为 d 的扩张卷积的输出, $C(Y)$ 表示多尺度特征增强模块的最终输出, X 表示输入特征图, $\text{PConv}(\cdot)$ 表示逐点卷积, $\Gamma_d\{\cdot\}$ 表示所有扩张卷积的级联运算。

1.3 附加检测头

电力场景图像背景复杂, 且锈蚀的显著度低, 其关键信息易在特征融合的下采样过程中丢失, 进而导致网络检测困难。为此, 本文采取深度监督^[26]方法, 通过在深度神经网络的某些中间层增加辅助分类器作为一种网络分支, 对主干网络进行监督, 从而减少锈蚀特征信息丢失, 提高对锈蚀缺陷的检测能力。在训练过程中, 增加附加检测头 (AD-Head) 以提高模型识别能力, 在推理时, 去掉 AD-Head 不影响模型推理速度。该附加检测头结构如图 5 所示。

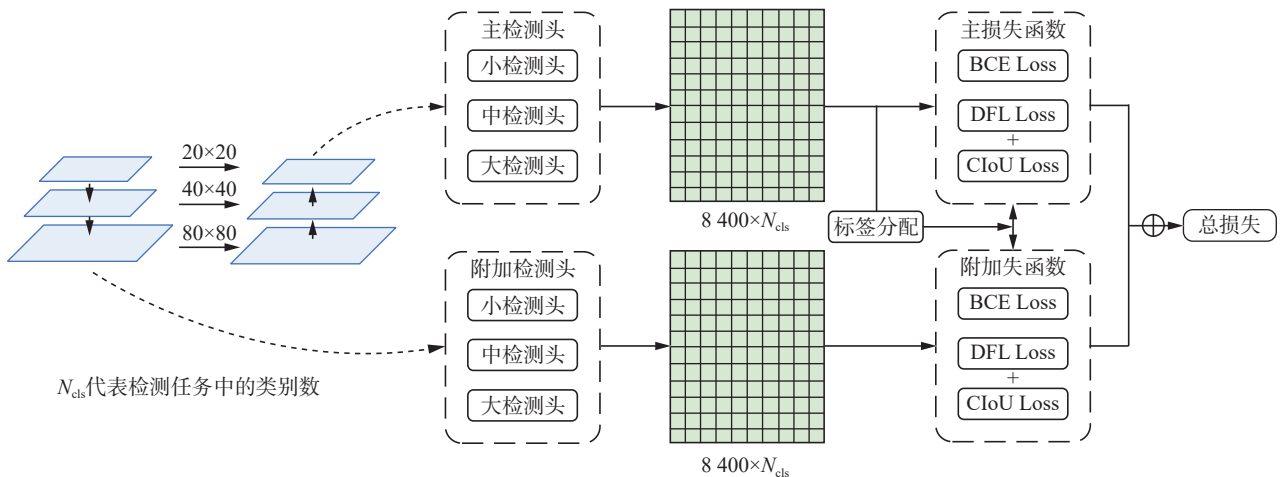


图 5 附加检测头结构

Fig. 5 Structure of additional detection head additional detection head

YOLOv8 采用的是 Task-Aligned Assigner^[27] 动态标签分配策略, 该策略根据分类与回归分数的加权分数选择正样本。其计算公式为

$$t = s^\alpha \times u^\beta \quad (9)$$

式中: t 为加权分数, 通过训练可以引导网络动态地关注于高质量的正样本。 s 代表标注类别的分类分数; u 代表预测框与真实框 (GTbox) 的交并

比 I_{ou} (intersection over union, IoU), 即回归分数; α 和 β 为权重超参数。该方法使用主检测头的预测结果作为指导, 生成从粗到细的层次标签, 分别用于附加检测头和主检测头的学习。主检测头选择加权分数最高的 top- k 个 (默认 $k=10$) 预测框作为正样本。附加检测头则通过放宽对正样本分配约束, 纳入更多候选框作为正样本。

YOLOv8 类别损失采用的是交叉熵损失 (binary cross-entropy loss, BCE Loss), 用 L_{BCE} 表示:

$$L_{BCE} = -((1 - t_{mb}) \log(1 - p_{yc}) + t_{mb} \log(p_{yc})) \quad (10)$$

式中: t_{mb} 为模型预测值, p_{yc} 为正样本处的预测框与目标框的 IoU。YOLOv8 的边框回归损失采用的是 DFL Loss^[28] 和 CIoU Loss 的组合, 分别用 L_{DFL} 和 L_{CIoU} 表示:

$$L_{DFL}(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (11)$$

式中: S_i, S_{i+1} 为 Softmax 函数输出, y_i 和 y_{i+1} 表示最接近目标 y 的周围区间值, $y_i \leq y \leq y_{i+1}$ 。

$$L_{CIoU} = 1 - I_{ou} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha \nu \quad (12)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{\eta^{gt}} - \arctan \frac{\omega}{\eta} \right)^2 \quad (13)$$

$$\alpha = \frac{\nu}{1 - I_{ou} + \nu} \quad (14)$$

式中: α 是权重函数, ν 用来度量长宽比的相似性, b 和 b^{gt} 分别代表预测框和真实框的中心点, ρ 代表两个中心点的欧氏距离, c 代表能同时包含预测框和真实框的最小闭包区域的对角线距离, ω^{gt} 、 η^{gt} 分别表示真实框的宽和高, ω 、 η 分别代表预测框的宽和高。同理, 代入 AD-Head 参数可以得到 L_{BCE-AD} 、 L_{DEL-AD} 、 $L_{CIoU-AD}$ 。Main-Lead Loss 和 AD-Head Loss 最终按 1:0.25 融合得到 $L_{BCE-last}$ 、 $L_{DEL-last}$ 、 $L_{CIoU-last}$ 。最终的损失函数为 DFL、CIoU 及 CLS 3 个损失函数的加权求和。

2 实验结果与分析

2.1 实验数据与环境

本文选取变电站锈蚀缺陷图像作为实验对象, 由变电站实际巡检拍摄得到, 共 1 124 张图片, 通过 LabelImg 标注工具进行标注。对标注完的数据按 8:2 的比例随机划分为训练集和测试集, 划分后的训练集和测试集图片数量分别为 900 张和 224 张。本文的实验环境如表 1 所示, 模型参数设置: 训练轮次 (epoch) 为 200 轮, 批大小 (batchsize) 为 16。

表 1 实验环境

Table 1 Experimental environment

| 名称 | 型号 |
|--------|--------------|
| 操作系统 | Ubuntu16.04 |
| CPU型号 | i7-5930K |
| GPU型号 | 1080Ti |
| 深度学习框架 | PyTorch1.9.0 |
| 编程语言 | Python3.8 |
| CUDA | 10.2 |

2.2 评价指标

本文选取准确率 P (Precision)、召回率 R (Recall) 和平均精度均值 I_{mAP} (mean average precision, mAP) 作为模型评价指标。准确率又称为查准率, 即正确预测为正的样本占全部预测为正的样本的比例。召回率又称查全率, 即正确预测为正的样本占所有实际为正的样本的比例。mAP50 表示 IoU 阈值在 0.5 时的平均检测精度。计算公式分别为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (15)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (16)$$

$$I_{AP} = \int_0^1 P(R) dR \quad (17)$$

$$I_{mAP} = \frac{1}{N} \sum_{i=1}^N \int_0^1 P_i(R_i) dR \quad (18)$$

式中: N_{TP} 表示正确预测为锈蚀的样本数, N_{FP} 表示错误预测为锈蚀的样本数, N_{FN} 表示未被预测为锈蚀的样本数, N 为检测的类别数。

在评价模型性能方面, 选取模型参数量 (Params) 以及每秒帧数 (FPS) 分别作为模型大小和检测实时性的评价指标。

2.3 频率通道注意力模块位置分析

为研究 FcaNet 的位置与数量对锈蚀缺陷检测效果的影响, 在保证模型其他结构参数相同的情况下进行了对比实验。模型 1~5 依次表示将 FcaNet 分别添加在 Backbone 倒数第 2 个 C2f 与 Neck 连接处, Backbone 的 SPPF 后, Backbone 倒数第 2 个 C2f 与 Neck 连接处以及 SPPF 后, Backbone 倒数第 2 个、第 3 个 C2f 与 Neck 连接处以及 SPPF 后, Neck 的最后两个 C2f 后, 实验结果见表 2。

表 2 频率通道注意力模块不同位置对模型检测性能的影响
Table 2 Effects of different positions of the FcaNet module on model detection %

| 模型 | 添加位置 | mAP50 |
|----|-----------------------------------|-------|
| 1 | Backbone倒数第2个C2f与Neck连接处 | 73.6 |
| 2 | Backbone的SPPF后 | 73.8 |
| 3 | Backbone倒数第2个C2f与Neck连接处以及SPPF后 | 74.6 |
| 4 | Backbone倒数第2、3个C2f与Neck连接处以及SPPF后 | 73.0 |
| 5 | Neck的最后两个C2f后 | 72.4 |

实验结果表明, 与模型 1 相比, 模型 2 取得了更高的检测精度, 这是因为随着网络的加深, 细节信息会逐渐损失, 在越深层的网络补充锈蚀的细节信息, 网络对锈蚀的识别能力越强。与其他

模型相比, 模型 3 提升的效果最显著, 得益于补充的锈蚀细节特征在深层网络得到了更加精细的调整以及在 Neck 网络中得到了充分的融合, 使得模型能够更好地关注与锈蚀相关的特征信息。模型 4 的检测精度低于模型 1 与模型 2, 因为在 Backbone 倒数第 3 个 C2f 与 Neck 连接处加入频率通道注意力机制, 此时网络层数较浅, 提取的无关信息较多, 随着网络的加深, 反而会对模型造成干扰。另外, 与其他模型相比, 模型 4 的检测效果最差, 因为频率通道注意力机制添加的位置靠近网络输出, 导致锈蚀的细节特征没有得到及时的补充与充分的融合。

2.4 多尺度模块效果分析

为探究多尺度模块对不同大小的锈蚀区域的检测效果, 本研究使用多尺度特征增强模块重构 C2f 模块, 将检测头前的 C2f 模块替换为重构的 MS-C2f 模块, 并与 YOLOv8m 模型进行对比实验, 其结果如表 3 示。采用 AP_s 、 AP_m 和 AP_l 作为评

估指标, 分别表示像素面积小于 32×32 、 $32 \times 32 \sim 96 \times 96$ 以及大于 96×96 目标的 $AP_{50:95}$ 值。由结果可知, 相较于原始的 YOLOv8m 模型, 添加 MS-C2f 以后的 YOLOv8m 模型对不同大小的锈蚀区域的检测准确率有一定的提升, 其中锈蚀小目标、中等目标以及大目标的检测精度分别提升了 3.4%、1.8% 和 0.6%, 证明了多尺度特征增强模块的有效性。

表 3 多尺度特征增强模块的效果
Table 3 Effect of multi-scale modules %

| 模型 | AP_s | AP_m | AP_l |
|----------------|--------|--------|--------|
| YOLOv8m | 28.8 | 31.5 | 46.0 |
| YOLOv8m+MS-C2f | 32.2 | 33.3 | 46.6 |

2.5 消融实验

为了验证本文方法中所提出的频率通道注意力机制、多尺度特征增强模块以及附加检测头模块的有效性, 设计了 8 组消融实验, 如表 4 所示。

表 4 消融实验结果
Table 4 Ablation experimental results

| 模型 | YOLOv8m | FcaNet | MS-C2f | AD-Head | $P/\%$ | $R/\%$ | $mAP_{50}/\%$ | 参数量/ 10^6 | 检测速度/(f/s) |
|----|---------|--------|--------|---------|--------|--------|---------------|-------------|------------|
| A | √ | | | | 81.0 | 67.4 | 73.0 | 25.8 | 91.7 |
| B | √ | √ | | | 82.6 | 67.5 | 74.6 | 25.9 | 91.7 |
| C | √ | | √ | | 83.1 | 66.0 | 75.2 | 26.8 | 87.0 |
| D | √ | | | √ | 83.7 | 68.5 | 75.8 | 25.8 | 92.6 |
| E | √ | √ | √ | | 81.8 | 68.3 | 76.1 | 26.9 | 86.2 |
| F | √ | √ | | √ | 83.3 | 68.5 | 76.4 | 25.9 | 90.9 |
| G | √ | | √ | √ | 81.6 | 69.9 | 77.4 | 26.8 | 87.7 |
| H | √ | √ | √ | √ | 86.5 | 69.0 | 78.1 | 26.9 | 86.2 |

注: “√”表示添加此模块。

从表 4 中可以看出, 在 YOLOv8m 模型中使用本文方法均能提高模型检测的准确率。其中模型 A 代表使用的基线模型, 即 YOLOv8m 模型, 其余实验均在基线模型上进行改进。模型 B 表示在模型 A 的基础上添加了 FcaNet, FcaNet 引入了更多的频率信息, 补充了深层网络受轮廓变化影响较小的纹理等细节特征, 优化了模型对锈蚀特征的提取, 模型 B 的 mAP_{50} 提升 1.6%, 证明了 FcaNet 的有效性。模型 C 在模型 A 的基础上引入利用多尺度特征增强模块重构的 MS-C2f 模块, SCConv 可以更精确和完整地定位锈蚀目标, 扩展卷积可以在卷积核大小不变的情况下扩大网络感受野, 有利于捕获多尺度的上下文信息, 使网络更全面地捕获锈蚀缺陷的特征, 使模型的 mAP_{50} 提升了 2.2%, 提高了模型对锈蚀的检测精

度。模型 D 在模型 A 的基础上添加了 AD-Head, 缓解锈蚀特征易在下采样过程中丢失关键信息的情况, 模型的 mAP_{50} 提升了 2.8%, 证明了 AD-Head 的有效性。模型 E、F、G 表示将本文所提的 3 个模块分别两两组合加入, 相较于模型 B、C、D, 即将所提模块单独添加的模型, 网络的平均精度均有提高, 证明了所提模块两两添加的有效性。模型 H(本文方法) 表示对 YOLOv8m 添加本文所有的改进措施, 改进后的网络准确率提升了 5.5%, 召回率提升了 1.6%, mAP_{50} 提升了 5.1%。添加的 3 个模块使得网络在多个层级中进行了调整 and 计算, 导致了参数量和检测时间的增加。改进后的模型虽增加了参数量和计算时间, 但提高了模型的检测精度。综上所述, 本文提出的 3 种改进方法对变电设备锈蚀图像检测存在优势。

2.6 检测结果可视化分析

为进一步地展示本文方法对锈蚀检测效果

的提升,对比分析了基线模型和改进模型(本文方法)的可视化预测结果,如图 6 所示。

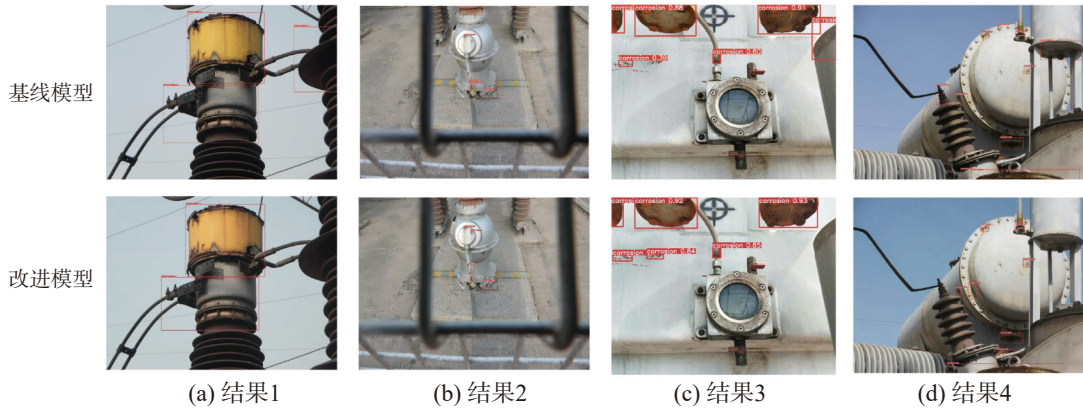


图 6 测试结果可视化

Fig. 6 Visualization of test results

由结果 1 和结果 2 分析可知,锈蚀的颜色一般呈红色或者红褐色,基线模型容易将颜色相近的其他区域误判为锈蚀,而改进模型通过添加 FcaNet,引入更多的频率分量,学习到锈蚀更多的细节特征,从而提高了对锈蚀的识别能力。由结果 3 分析可知,锈蚀区域面积大小有所差异,基线模型对面积较小的锈蚀区域产生了漏检,而改进模型通过添加经过多尺度特征增强模块重构的 MS-C2f 模块,使更多锈蚀缺陷信息进入有效感受野,改善了对不同大小的锈蚀区域的识别情况。

由结果 2 和结果 4 分析可知,基线模型出现了漏检与检测框重叠的情况,改进后的模型通过添加 AD-Head,与主检测头共同计算损失,实现深浅网络的信息交互,获取了更加完整的锈蚀特征信息,减少了漏检与检测框重合的情况。

综上所述,改进模型关注的锈蚀区域更加准确,减少了对锈蚀区域的误判、漏检以及检测框重叠的情况,验证了本文方法的可行性与有效性。

2.7 模型对比实验

为进一步验证本文方法的优越性,在变电设备锈蚀缺陷数据集上与目前先进的目标检测算法进行对比,对比模型分别是 Faster R-CNN、YOLOv5m、YOLOv6m^[29]、YOLOv7^[30]、RT-DETR^[31]以及 YOLOv9c^[32],各项测试指标如表 5 所示。由表 5 可知,本文方法的平均检测精度达到 78.1%,mAP50:95 达到 48.2%,明显高于其他模型,在参数量方面仅次于 YOLOv5m 和 YOLOv8m,浮点计算量方面仅次于 YOLOv5m,本文方法检测速度优于 YOLOv7、RT-DETR、YOLOv9c 模型。

表 5 与先进检测算法性能对比

Table 5 Performance comparison with advanced detect algorithms

| 模型 | Backbone | mAP50/% | mAP50:95/% | 参数量/10 ⁶ | 浮点计算量/10 ¹² | 检测速度/(f/s) |
|--------------|--------------|---------|------------|---------------------|------------------------|------------|
| Faster R-CNN | Resnet50 | 40.5 | 17.0 | 137.1 | 370.2 | 11.9 |
| YOLOv5m | Darknet53 | 73.1 | 36.9 | 20.8 | 47.9 | 108.7 |
| YOLOv6m | EfficientRep | 68.0 | 37.0 | 34.8 | 85.62 | 91.6 |
| YOLOv7 | E-ELAN | 72.5 | 39.1 | 36.4 | 103.2 | 53.2 |
| YOLOv8m | CSPDarkNet | 73.0 | 43.0 | 25.8 | 84.4 | 91.7 |
| RT-DETR | HGNetv2 | 62.4 | 31.7 | 65.4 | 222.5 | 28.5 |
| YOLOv9c | GELAN | 71.2 | 41.9 | 50.7 | 236.6 | 36.0 |
| 本文方法 | CSPDarkNet | 78.1 | 48.2 | 26.9 | 79.5 | 86.2 |

图 7 为不同模型测试结果:由组①可以看出,YOLOv5m、YOLOv6m、YOLOv8m 以及 YOLOv9c 都将下面的红色牌子误检为锈蚀,YOLOv7 误将背景识别为锈蚀,RT-DETR 出现了检测框重合的情况,本文方法没有出现误检且检测框更贴合目

标;由组②可以看出,除 YOLOv8m 和本文方法外的其他模型都将其他区域误检为锈蚀,本文方法相较于 YOLOv8m 模型的检测精度更高。由此可以看出,本文算法对于锈蚀缺陷特征的捕捉更为准确,验证了本文方法的先进性。

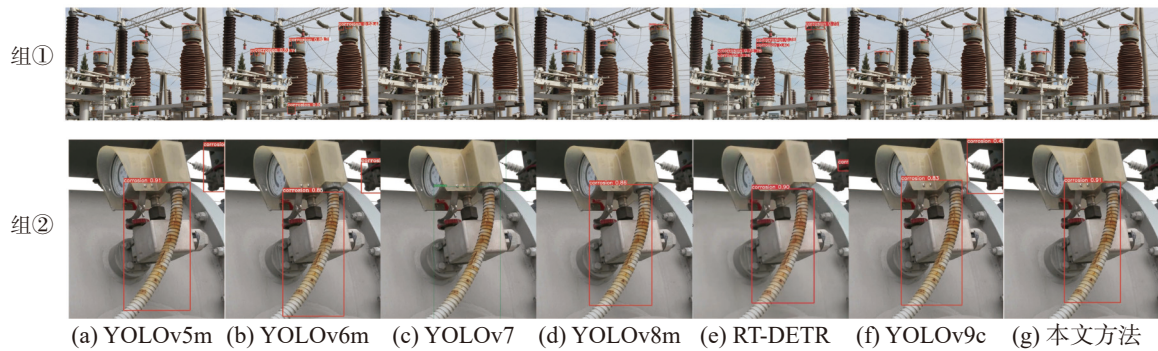


图 7 不同模型测试结果

Fig. 7 Test results of different models

3 结束语

针对复杂场景下变电设备锈蚀缺陷在检测过程中形态差异大、长宽比随机、关键特征易丢失的问题, 本文提出一种面向复杂场景的变电设备锈蚀检测方法。在 YOLOv8m 的特征提取网络与特征融合网络之间加入 FcaNet, 加强了模型对锈蚀细节特征的学习能力, 提高了模型的判别能力, 降低了误判的概率; 在特征融合网络中添加基于多尺度特征增强模块重构的 MS-C2f 模块, 提高了模型对不同大小锈蚀区域的识别能力; 通过引入 AD-Head 从主网络生成的特征映射中提取额外的特征, 改善卷积层下采样导致的信息丢失的情况, 并与主检测头共同计算损失, 实现深浅网络的信息交互, 改善了漏检的情况。实验结果表明, 本文提出的变电设备锈蚀检测方法的平均准确率达到 78.1%, 较基线模型提升了 5.1%, 有效地提升了对变电设备锈蚀缺陷的检测精度, 同时检测效果也优于其他主流目标检测模型。虽然改进模型精度有所提升, 但仍不能满足实际巡检的高精度要求, 同时改进模型相较于原 YOLOv8m 模型参数量有所增加、检测速度有所下降, 同样制约了模型在实际中的应用。未来研究将聚焦于进一步提高对变电设备锈蚀缺陷区域的检测精度, 尽可能进行模型的轻量化, 希望将所提出的网络应用到实际的缺陷检测场景中。

参考文献:

- [1] 赵振兵, 冯烁, 赵文清, 等. 融合知识迁移和改进 YOLOv6 的变电设备热像检测方法[J]. *智能系统学报*, 2023, 18(6): 1213–1222.
ZHAO Zhenbing, FENG Shuo, ZHAO Wenqing, et al. Thermd image detection method for substation equipment by incorporating knowledge migration and improved YOLOv6[J]. *CAAI transactions on intelligent systems*, 2023, 18(6): 1213–1222.
- [2] 吴添权, 戴美胜, 杨刚, 等. 基于改进 Mask-RCNN 模型的电力设备锈蚀识别[J]. *电力信息与通信技术*, 2021, 19(4): 25–30.
WU Tianquan, DAI Meisheng, YANG Gang, et al. Corrosion detection of power equipment based on improved mask-RCNN model[J]. *Electric power information and communication technology*, 2021, 19(4): 25–30.
- [3] 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 Lingyun, LI Tingyi, LI Yang, 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.
- [4] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International journal of computer vision*, 2004, 60(2): 91–110.
- [5] ZHU Qiang, YEH M C, CHENG K T, et al. Fast human detection using a cascade of histograms of oriented gradients[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2006: 1491–1498.
- [6] 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]. *中国电机工程学报*, 2023, 43(19): 7423–7446.
LIU Chuanyang, WU Yiquan. Research progress of vision detection methods based on deep learning for transmission lines[J]. *Proceedings of the CSEE*, 2023, 43(19): 7423–7446.
- [7] ZHANG Yu, LI Weixing, ZHOU Jian, et al. Substation equipment abnormality detection based on improved YOLOv5 [C]//2022 41st Chinese Control Conference. Hefei: IEEE, 2022: 6654–6659.
- [8] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020: 12993–13000.
- [9] XU Bo, DING Yuan. Research on the application of deep learning object detection in rust defect detection of power equipment[C]//2023 3rd International Conference on Electrical Engineering and Mechatronics Technology. Nanjing: IEEE, 2023: 770–775.
- [10] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(6): 1137–1149.
- [11] 张迎晨, 王波, 马富齐, 等. 基于域适应网络的电力设备

- 不规则外表面缺陷高精度检测方法[J]. 高电压技术, 2022, 48(11): 4516–4526.
- ZHANG Yingchen, WANG Bo, MA Fuqi, et al. High-precision detection method of irregular outer surface defects of power equipment based on domain adaptation network[J]. High voltage engineering, 2022, 48(11): 4516–4526.
- [12] 于豪, 蒋锦霞, 赖晓翰, 等. 基于自适应感受野的电力设备表面缺陷检测方法[J]. 系统仿真学报, 2023, 35(7): 1572–1580.
- YU Hao, JIANG Jinxia, LAI Xiaohan, et al. Surface defect detection of power equipment using adaptive receptive field network[J]. Journal of system simulation, 2023, 35(7): 1572–1580.
- [13] CHEN Xiaojiao, AN Zhenyu, HUANG Liansheng, et al. Surface defect detection of electric power equipment in substation based on improved YOLOV4 algorithm[C]//2020 10th International Conference on Power and Energy Systems. Chengdu: IEEE, 2020: 256–261.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020–04–23)[2024–01–01]. <https://arxiv.org/abs/2004.10934v1>.
- [15] DAI Yan, HAN Rui, LIU Li, et al. Research on substation equipment rust detection method based on improved SSD[C]//2021 International Conference on Advanced Electrical Equipment and Reliable Operation. Beijing: IEEE, 2021: 1–3.
- [16] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[M]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21–37.
- [17] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770–778.
- [18] ZHANG Sheng, DENG Xinling, LU Yumin, et al. A channel attention based deep neural network for automatic metallic corrosion detection[J]. Journal of building engineering, 2021, 42: 103046.
- [19] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [20] QIN Zequn, ZHANG Pengyi, WU Fei, et al. FcaNet: frequency channel attention networks[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 763–772.
- [21] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 936–944.
- [22] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759–8768.
- [23] LIN Min, CHEN Qiang, YAN Shuicheng. Network in network[EB/OL]. (2014–03–04)[2024–01–01]. <https://arxiv.org/abs/1312.4400v3>.
- [24] AHMED N, NATARAJAN T, RAO K R. Discrete cosine transform[J]. IEEE transactions on computers, 1974, C-23(1): 90–93.
- [25] LIU Jiangjiang, HOU Qibin, CHENG Mingming, et al. Improving convolutional networks with self-calibrated convolutions[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10093–10102.
- [26] LEE C Y, XIE S, GALLAGHER P, et al. Deeply-supervised nets[C]//Artificial intelligence and statistics. San Diego: PMLR, 2015: 562–570.
- [27] FENG Chengjian, ZHONG Yujie, GAO Yu, et al. TOOD: task-aligned one-stage object detection[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 3490–3499.
- [28] LI Xiang, WANG Wenhai, WU Lijun, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection[J]. Advances in neural information processing systems, 2020, 33: 21002–21012.
- [29] LI Chuyi, LI Lulu, JIANG Hongliang, et al. YOLOv6: A single-stage object detection framework for industrial applications[EB/OL]. (2022–09–07)[2024–01–01]. <https://arxiv.org/abs/2209.02976>.
- [30] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[EB/OL]. (2022–07–06)[2024–01–01]. <https://arxiv.org/abs/2207.02696v1>.
- [31] ZHAO Yian, LYU Wenyu, XU Shangliang, ZHAO Y, et al. Detsr beat yolos on real-time object detection[EB/OL]. (2023–07–06)[2024–01–01]. <https://arxiv.org/abs/2304.08069>.
- [32] WANG C Y, YEH I H, LIAO H M. YOLOv9: learning what you want to learn using programmable gradient information[EB/OL]. (2024–02–29)[2024–03–28]. <https://arxiv.org/abs/2402.13616v2>.

作者简介:



赵振兵, 教授, 博士, 主要研究方向为电力视觉。主持国家自然科学基金项目等纵向课题 10 项; 获省科技进步奖一等奖 2 项; 以第一完成人获得国家专利授权 16 项; 以第一作者出版专著 2 部, 发表学术论文 50 余篇。E-mail: zhaozhenbing@ncepu.edu.cn。



席悦, 硕士研究生, 主要研究方向为电力视觉。E-mail: 17325207795@163.com。



冯烁, 博士研究生, 主要研究方向为电力视觉。E-mail: fs_ncepu@163.com。