

基于语义分割数据增强与可变形卷积的 输变电线路缺陷检测

王译萱^{1,2,3}, 向思屿^{1,2}, 梁晖辉^{1,2}, 邝俊威⁴, 张菊玲^{1,2}, 刘松嘉⁵

(1. 国网四川省电力公司电力科学研究院, 四川 成都 610041; 2. 电力物联网四川省重点实验室, 四川 成都 610041; 3. 华北电力大学计算机系, 河北 保定 071003; 4. 国网四川省电力公司泸州供电公司, 四川 泸州 646000; 5. 国网四川省电力公司自贡供电公司, 四川 自贡 643000)

摘要:为提高输变电线路在长期运行和自然环境影响下产生的各类缺陷的检测精度和效率,提出了基于YOLOv8和可变形卷积的改进算法。针对各类缺陷样本数量差距较大导致的长尾分布问题,提出使用分割一切模型对数据进行扩充以提高数据平衡性;在主干网络中引入DCNv2结构,通过动态调整卷积核的形状适应各类样本包含的特征以提高泛化能力,并结合multi-CA注意力机制使网络注重各通道的融合信息;使用损失函数WIoU引导模型学习,以适应不同质量的锚框。将改进的算法与其他算法进行比较,结果表明改进算法增加了对输变电缺陷检测的精度。

关键词:DCNv2; YOLOv8; 分割一切模型; WIoU; multi-CA

中图分类号:TM 755 文献标志码:A 文章编号:1003-6954(2025)01-0032-09

DOI:10.16527/j.issn.1003-6954.20250105

Defect Detection for Transmission and Transformation Lines Based on SAM Data Augmentation and Deformable Convolution

WANG Yixuan^{1,2,3}, XIANG Siyu^{1,2}, LIANG Huihui^{1,2}, KUANG Junwei⁴, ZHANG Juling^{1,2}, LIU Songjia⁵

(1. State Grid Sichuan Electric Power Research Institute, Chengdu 610041, Sichuan, China; 2. Power Internet of Things Key Laboratory of Sichuan Province, Chengdu 610041, Sichuan, China; 3. Department of Computer Science, North China Electric Power University, Baoding 071003, Hebei, China; 4. State Grid Luzhou Electric Power Supply Company, Luzhou 646000, Sichuan, China; 5. State Grid Zigong Electric Power Supply Company, Zigong 643000, Sichuan, China)

Abstract:In order to enhance the detection precision and efficiency of various defects in transmission and transformation lines under the influence of long-term operation and natural environments, an improved algorithm based on YOLOv8 and deformable convolution is proposed. Aiming at long-tail distribution problem caused by the significant disparity in the quantity of various defect samples, the segment anything model (SAM) is proposed to augment the data so as to enhance data balance. Within the Backbone, DCNv2 is introduced to dynamically adjust the shape of convolutional kernels to adapt to the features contained in various samples, thus enhancing generalization capability. Moreover, the integration of multi-CA attention mechanism guides the network to focus on the fusion information of each channel, and WIoU is utilized to guide the model learning, which enables adaptation to anchor boxes of different qualities. Comparative analysis with other algorithms shows an increase in detection accuracy for the defects in transmission and transformation lines achieved by the proposed improved algorithm.

Key words:DCNv2; YOLOv8; segment anything model; WIoU; multi-CA

0 引言

输变电设备是电力系统中的重要组成部分,负责将高压电能转换为低压电能,并进行输送和分配。然而,长期无休运行和受自然环境因素的影响可能导致输变电设备产生各种缺陷,如绝缘子破损、线路异物和设备老化等。如若这些缺陷不能及时检测和修复,可能会导致设备故障、电力损失等不可估计的事故。

目前,输变电设备缺陷检测的方法主要包括人工巡检和基于技术手段的智能检测。传统的人工巡检方法主要由检测人员对输变电设备进行目视检查,通过排查设备可视状态、温度、是否有异响等特征来判断是否存在缺陷^[1]。然而,这种方法对检测人员的素质要求高,存在检测效率和结果受检测人员主观影响等不可控因素。

目前基于大数据和人工智能技术的快速发展,特征识别技术取得了巨大的进展,被应用于各个领域。在电力系统的智能化和信息化管理中,采用目标检测方法利用无人机技术进行巡检并结合深度学习算法对输电线路进行智能故障检测,已经成为不可或缺的重要技术手段^[2]。这不仅可提升检测准确性与效率,亦可降低人力成本。

目标检测通常分为两种范式:一阶段方法和二阶段方法。一阶段方法将目标检测当作一个任务完成,直接从采集的图像中识别并定位目标的位置和类别,同时执行定位和分类。典型的一阶段方法包括 YOLO(You only look once)算法和单步多框目标检测(single shot multiBox detector, SSD)算法等,这类方法通常具有较快的检测速度和较为简单的网络结构,适用于对目标数量较多且目标大小变化较大的场景。二阶段方法先是通过一种算法将包含检测目标的区域进行提取,再将这些候选区域送入另一个网络进行分类和精确定位^[3]。典型的二阶段方法包括快速基于区域的卷积神经网络(faster region-based convolutional neural network, faster R-CNN)和掩膜基于区域的卷积神经网络(mask region-based convolutional neural network, mask R-CNN)等。二阶段方法的准确性更高,但与此同时计算量也更大,网络结构更加复杂,适用于对目标定位和缺陷识别精度要求高的场景^[4]。

YOLO 算法在缺陷检测方面取得了一些研究进展,并在不同领域的缺陷检测任务中展示了较好的性能,如在绝缘子破损检测、金属结构缺陷检测、输变电设备缺陷检测上^[5]都有显著成绩。由于输变电线路中的故障种类多样且数据量较大,为了提高对输变电线路缺陷检测的准确度,并提升检测的效率,下面采用改进的 YOLOv8 算法对输变电线路上的故障进行诊断。

1 YOLOv8 算法结构

1.1 网络结构

YOLOv8 模型的结构如图 1 所示,该模型由输入端开始,经过主干网络和特征融合网络最终到达输出端。在主干网络(Backbone)方面,借鉴了 YOLOv7 的高效层聚合网络(efficient layer aggregation networks, ELAN)结构的设计思想^[6],对 YOLOv5 的 C3 结构进行了改进,通过使用 C2f 结构使得模型梯度流变得更加丰富,同时针对不同尺度的模型对通道数也进行了不同的设计。

Head 部分相对于 YOLOv5 有两项重大改进:首先,区别于 YOLOv5 的单一目标检测头,YOLOv8 将分类和检测头分离开来,使用了解耦头结构;其次,不再依赖锚框(Anchor),而是使用 Anchor-Free 的方式,得以更灵活地处理目标的位置和形状,从而提高模型的性能和适应能力。YOLOv8 在损失函数方面采用了新的方法,摒弃了传统的交并比(intersection over union, IoU)匹配或单边比例的样本分配方式,采用了任务对齐分配器(task-aligned assigner)的正负样本匹配策略。此外,引入了分布焦点损失(distribution focal loss, DFL)函数来解决类别数量差距大和难易样本不平衡的问题^[7]。通过分布感知机制以及焦点损失机制的引入优化模型,达到提高目标检测性能的目的。在训练过程中,YOLOv8 借鉴了 YOLOX 的方法,在最后的 10 个轮次(epoch)关闭马赛克(Mosaic)增强操作。该改进有效地减少了过拟合的可能性,简化了训练过程,降低了数据相关性,并且提升了模型的泛化能力及精度。

1.2 C2f 模块

YOLOv8 中将主流 YOLOv5 模型使用的 C3 模块更改为更加轻量化的 C2f 模块。基于跨阶段部分网络(cross stage partial network, CSPNet)的提取分

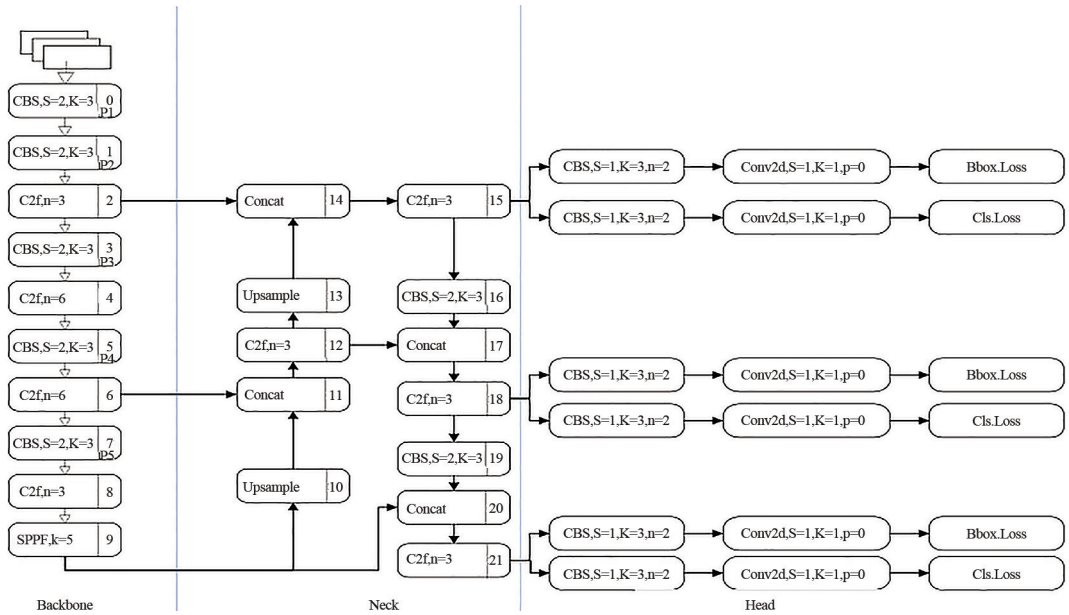


图 1 YOLOv8 结构

流思想,使用 n 个 Bottleneck 模块的堆叠(其中 n 是根据参数规模而自行变化的),再融合残差结构构造了 C3 模块。C3 模块由 3 个卷积层组成,如图 2 所示,图中 CBS 模块分别为卷积层(convolution, Conv)、批量归一化层(batch norm, BN)、SiLu 激活层,即 CBS = Conv + BN + SiLu,以及 n 个 Bottleneck。C2f 模块的设计是基于 C3 模块的思想以及 ELAN 的改进思想进行设计,使得 YOLOv8 在更加轻量化的同时拥有更全面的梯度流信息。C2f 模块如图 3 所示。

BCE) 损失函数来衡量模型的性能, BCE 损失函数的公式如式(1)所示。由于去除了对象损失,置信度分数最大值的选取也从 YOLOv5 的先取最大值再输出对应的最大类别,到 YOLOv8 的输出所有置信度的分数再选取最大值。对于回归分支,采用 DFL 函数中提出的积分形式表示法,优化和真实值最为相近的左右位置概率,使得网络可以更加快速地关注到目标附近,DFL 函数如式(2)所示。结合完全交并比(complete intersection over union, CIoU)损失函数,来捕捉目标位置的回归误差。这两种损失函数通过一定的权重比例加权,以平衡它们在训练过程中的影响。综上所述,通过这种损失函数的设计,可以更好地训练目标检测模型,提高其在检测任务中的性能表现。

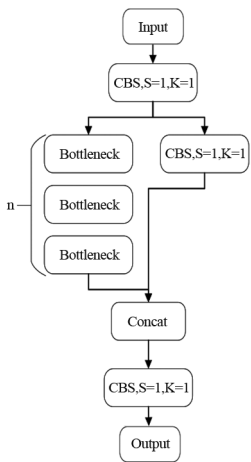


图 2 C3 模块

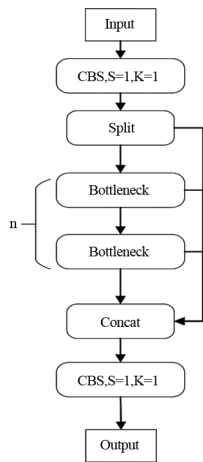


图 3 C2f 模块

1.3 损失计算

YOLOv8 的损失函数由两部分构成,分类损失和回归损失,不再包括之前的对象损失分支,由是否包含对象细致到是否包含此类别的对象。对于分类分支,仍然使用二元交叉熵(binary cross entropy,

$$L = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i - [y_i \lg p_i + (1 - y_i) \lg (1 - p_i)] \tag{1}$$

式中: L_i 为二元交叉熵损失函数; N 为模型预测数据总组数; y_i 为第 i 组数据真实的二元标签,即 0 或 1; p_i 为第 i 组数据预测标签。

$$L_{DFL}(S_i, S_{i+1}) = - [(y_{i+1} - y) \lg S_i + (y - y_i) \lg S_{i+1}] \tag{2}$$

式中: S_i 和 S_{i+1} 为网络聚焦位于标签附近的两个预测值 y_i 与 y_{i+1} 对应的概率; y 为类别信息真实值; y_i 为第 i 组数据单边预测值。

2 算法改进

2.1 可变形卷积网络 v2

传统的卷积神经网络 (convolutional neural network, CNN) 受限于其固定的几何结构,这使得它们主要应用于对模型需要处理复杂及预测结果多样且未知的输入数据进行建模。由于 CNN 模型在设计时其卷积单元只以固定的窗口大小和步长进行滑动,池化层的比例也通常不会发生变化导致空间分辨率变化单一^[8]。基于以上结构特点,使得 CNN 网络无法对于不同尺度和位置的变换进行鲁棒性建模。在可变形卷积中,基于两个新模块的引入,使得 CNN 可以对具有几何变换的模型适应力显著增强。可变形卷积允许卷积核根据输入特征图中的局部结构进行动态调整,从而能够更好地适应目标物体的非刚性形变和复杂结构^[9]。这样的设计使得 CNN 能够更灵活地捕捉图像中的几何变换,进而提高了模型的感知能力。可变形卷积,即在标准卷积中加入对常规网格采样区域的偏移,使采样网络能够更加自由地变形。可变形卷积通过为感兴趣区域 (region of interest, RoI) 池化中的每个空间箱 (bin) 位置添加偏移量,实现了对卷积核采样位置的自适应调整。这使得模型能够更好地适应不同形状的目标,并更精确地捕捉其局部特征。传统的 RoI 池化方法存在固定处理方式的局限性,无法很好地适应非矩形目标区域。可变形卷积通过引入偏移量,提供了更灵活、更精确的处理不同形状目标的方法,从而提高了模型在目标检测等任务中的性能和泛化能力。

基于具有偏移学习功能的卷积层的增加,可变形卷积网络 v2 (deformable convolution networks v2, DCNv2) 能够使卷积核根据局部结构动态调整位置,从而灵活适应复杂场景,提升了模型的性能和泛化能力,特别是在处理几何变换方面取得了显著的改进^[10]。基于调制机制的引入,使得网络对空间支持区域也具有更强大的控制能力。这个调制机制使得样本具有通过学习到的偏移或学习的特征幅度进行调整的能力。这说明网络模块不仅能够调整感知输入特征的位置偏移,还能够调整不同空间位置的输入特征的强度或振幅。当输入处于最为极端的情况时,该模块可以将特征的振幅调整为 0,这样就可以

有选择性地屏蔽特定位置的信号^[11]。因此,来自这些位置的图像内容对模块输出的影响将会被显著地减少甚至完全消除。这意味着调制机制为网络模块提供了额外的灵活性,使其能够调整其对空间支持区域的影响,如式(3)所示。

$$y(p) = \sum_{k=1}^K w_k \cdot x(p + p_k + \Delta p_k) \cdot \Delta m_k \quad (3)$$

式中: $y(\cdot)$ 为输出特征图; p 为当前输出位置; p_k 为卷积核的偏移量,通常表示卷积核相对于中心点的位置; w_k 为第 k 个位置的卷积核权重; K 为卷积核大小; x 为输入特征图; Δp_k 为第 k 个位置的可学习偏移; Δm_k 为学习到的权重因子。 Δp_k 和 Δm_k 均由将相同输入特征图应用于单独的卷积层来获得。

2.2 多通道注意力机制

通道注意力 (channel-wise attention, CA) 机制专注于调整神经网络中不同通道之间的关系,以使网络能够更有效地利用输入特征图中的信息。它的主要目标是根据输入数据动态地学习每个通道的重要性,并据此调整通道间的信息流^[12]。CA 机制通过全局信息捕捉、注意力权重计算、特征加权、缩放和归一化使得网络更加关注对当前任务重要的特征通道,从而提高网络的表征能力和性能。传统的 CA 机制仅注意单通道的数据,将 CA 与多通道进行结合再嵌入到 DCNv2 中,可以使得可变形卷积网络注意各通道的信息,以达到更好的信息融合的作用,加入正则化以防止模型过拟合并提高泛化能力。将 CA 机制融合到 DCNv2 结构中的结构图如图 4 所示。

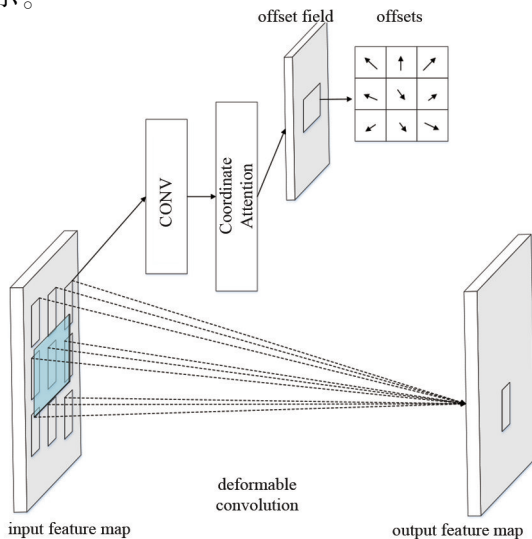


图 4 DCNv2-CA 结构

原始 CA 机制结构如图 5 所示。在原始 CA 机制的基础上,增加通道之间的融合,使各通道信息更好地交互。再将多通道注意力(multi channel-wise attention, multi-CA)机制加入到 DCNv2 中,可以使网络更加充分地提取各路特征,改进的 multi-CA 注意力机制如图 6 所示。

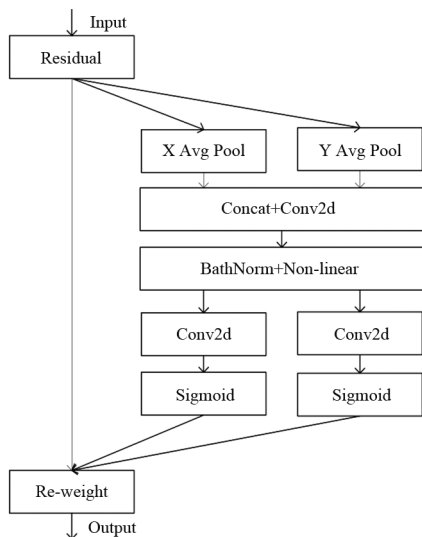


图 5 CA 机制结构

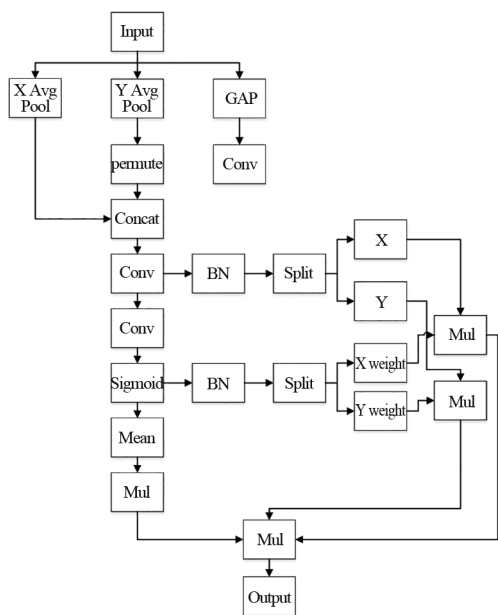


图 6 multi-CA 机制

2.3 加权交并比损失函数

YOLOv8 的边界框损失函数为 CIoU,公式如式(4)所示。

$$C_{IoU-loss} = 1 - I_{ou} + \frac{\rho^2(b, b_{gt})}{c_1^2} + \alpha v \quad (4)$$

式中: I_{ou} 为交并比; b 为预测中心坐标的参数; b_{gt} 为真实目标边界框中心的参数; ρ 为 b 和 b_{gt} 之间的欧氏距离; c_1 为两个矩形的最小外接矩形对角线长

度; α 为权重函数; v 为度量长宽比相似性函数^[13]。 α 和 v 计算分别如式(5)和式(6)所示。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2 \quad (5)$$

$$\alpha = \frac{v}{(1 - I_{ou}) + v} \quad (6)$$

式中: w 、 h 分别为预测框的宽度和高度; w_{gt} 、 h_{gt} 分别为真实框的宽度和高度。

尽管 CIoU 在一些情况下相对于传统的 IoU 和 GIoU 有一些改进,但它仍然存在一些缺点:计算复杂度高,对目标框的中心点、宽度和高度的归一化等操作也会增加也会导致算量随之提高;CIoU 对于边界框的调整非常敏感,这意味着即使较小的边界框调整也会对损失函数产生较大的影响。这可能导致模型在训练过程中过度关注边界框的微小变化,而忽略了目标区域的重要性。这种情况下,模型可能会在边界框调整方面过度拟合,而忽略了更重要的目标检测任务;对小目标不敏感,在处理小目标时可能不如其他损失函数表现好。由于绝缘子缺陷一般都属于小目标检测,所以使用 CIoU 损失函数的效果并不十分理想。

下面使用基于动态非单调的聚焦机制的边界框损失函数加权交并比(weighted IoU, WIoU)代替 CIoU。动态非单调聚焦机制通过使用“离群度(outlierness)”代替传统的 IoU 来评估锚框的质量。更少的人为干预模型训练,可以使模型的泛化能力得到有效提高。这种方法使得模型更有效地找到最具信息量的锚框,增加目标检测准确性和效率。离群度可以更加综合地考虑传统 IoU 较容易忽略的如目标大小、目标所处位置、目标是否受到遮挡等因素。此外,该机制还采用了一种更有效的梯度增益分配策略,以便更好地调整模型的注意力。基于这种策略的使用,使得模型可以更精确地学习到不同类型、不同质量的锚框,使得目标检测模型得以更加有效率的检测。

结合了离群度评估和梯度增益分配的动态非单调聚焦机制,能够更灵活地适应各类复杂场景下的目标检测任务,并取得更为显著的检测成果。在保留更多高质量锚框的同时还可以做到削减低质量锚框太多而生成有害梯度。通过使用 WIoU 可以使模型更加关注普通质量的锚框即绝大部分正常的锚框,最终可以达到提高模型整体能力的作用。WIoU 分为 3 个版本,其中 v1 版本是基于注意力的边界框损失,v2 和 v3 版本附加了聚焦机制。这里使用 WIoU-v3 并嵌套最小点距离交并比(minimum point

distance intersection over union, MPD-IoU), 通过动态调整梯度增益, 更有效地引导模型学习, 以适应不同质量的锚框, 从而提高目标检测的性能。WIoU-v1 计算方式如式(7)所示, 其 WIoU 的总体损失函数如式(8)所示。离群度的大小代表锚框的质量, 离群度越小代表锚框的质量越高, 离群度 β 的计算公式如式(9)所示。式(10)为 WIoU-v3 的计算方式^[14], 式(11)为 MPD-IoU 的计算公式, 基于 L_{IoU} 的动态变化使得锚框的质量划分标准也在动态变化, 从而可以帮助 WIoU-v3 在每个时刻都能得到最符合当前状态的梯度增益分配策略。

$$L_{WIoUv1} = R_{WIoU} L_{IoU} \quad (7)$$

$$R_{WIoU} = \exp \left[\frac{(a - a_{gt})^2 + (b - b_{gt})^2}{(W_g^2 + H_g^2)^*} \right] \quad (8)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (9)$$

$$L_{WIoUv3} = r L_{WIoUv1}, \quad r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (10)$$

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2$$

$$L_{MPD-IoU} = I_{oU} - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (11)$$

式中: L_{WIoUv1} 为 WIoU-v1 版损失函数; R_{WIoU} 为重新加权交并比; L_{IoU} 为位置交并比; a 、 b 为锚框预测值; a_{gt} 、 b_{gt} 分为锚框真实值; W_g 和 H_g 为最小框的宽度和高度; L_{IoU}^* 为 L_{IoU} 的单调聚焦系数; L_{WIoUv3} 为 WIoU-v3 版损失函数; r 为非单调聚焦系数; α 和 δ 为超参数; x_1 和 y_1 分别为真实框和预测框左上顶点的坐标; x_2 和 y_2 对应真实框和预测框框右下顶点的坐标; d_1 为真实框与预测框左上顶点的欧几里得距离; d_2 为真实框与预测框右下顶点的欧几里得距离; prd 表示预测框; gt 表示真实框。

3 实验

3.1 实验环境配置与数据处理

3.1.1 实验环境配置

实验软件环境为 python 3.8、torch 1.7.1、cuda 10.1 等, 硬件环境为 NVIDIA GeForce RTX 3090 显卡。所有实验都在相同超参数下进行。以下是对所提供名词的替代定义: mAP@0.5 表示 IoU 为 0.5 时所有类别待检测目标的平均检测精度; GFLOPs 为每秒进行的十亿次浮点运算数, 用于衡量算法的复杂度。

3.1.2 数据预处理

所用输电线路数据主要由无人机对现场状态的

采集构成, 全部数据由人工进行标注和清洗。由于原始数据量庞大、数据类型繁多且各类数据数量差异较大^[15], 前期的数据处理十分必要。原始数据中绝缘子故障只有一种分类, 在对数据集进行清洗过程中发现图片间缺陷差别较大, 于是将其细分为绝缘子自爆、瓷质绝缘子破裂、绝缘子弯曲, 并使用分割一切模型(segment anything model, SAM) 将样本数量较少的后两类缺陷进行数据扩充。数据处理步骤为:

- 1) 无人机图像采集;
- 2) 利用 SAM 模型提取缺陷部分掩码;
- 3) 对缺陷部分使用缩放、变形、裁减、旋转以及添加高斯噪声等方式进行数据加强;
- 4) 结合数据集中其他背景部分进行数据扩充;
- 5) 使用 Labellmg 进行数据标注。

下面基于 SAM 模型对数据进行扩充。SAM 模型作为图像编码器, 主要功能为生成图像嵌入。该模型由两种信息源组成输入数据, 一种是用于将图像信息嵌入到模型中的编码器, 另一种则是文字信息或输入框信息嵌入的编码器。这两类信息源由一种轻量级的掩码解码器组合在一起, 最终达到预测分割掩码的目的^[16]。SAM 模型结构如图 7 所示。在预训练阶段, 模型先接受关于点、边框和掩码的信息, 然后输出相应的分割结果, 并将其和实际数据加以对比。相较于传统方法, SAM 模型采用了交互式分割, 能够一次性提供多种可能的合理结果, 无需多轮交互或多次标注选择^[17]。此外, SAM 模型属于零样本学习的一种形式, 通过在不同的下游任务中应用不同的提示, 展现出卓越的泛化性能。与多任务分割方法相比, SAM 模型通过单一模型同时处理多个任务, 而不是将多个模型组合在一起处理各自的子任务^[18]。输入图像由掩码自编码器(masked auto encoders, MAE) 预训练的视觉变换器(vision transformer, ViT)^[19] 图像编码器进行处理。ViT 模型首先分割图像为固定大小的块, 并将每个块进行位置嵌入先行处理后, 将输出的向量序列输入进 Transformer 编码器; 然后, 将图像映射至特征空间, 生成一次性图像嵌入, 这些图像嵌入可以被视为图像的紧凑特征表示, 能够有效地保留输入图像的关键信息; 接着, 输入的提示(例如点提示和方框提示使用位置编码的形式, 文本提示则直接使用 CLIP 的文本编码器^[20]) 经过提示编码器处理后, 生成对应的提示嵌入; 最后, 将图像嵌入与提示嵌入结合,

并输入至轻量级的掩码解码器,从而生成相应的分割掩码。通过不断使用绝缘子数据对 SAM 模型进行微调,使得该模型能够更好地适应当前的应用场景。

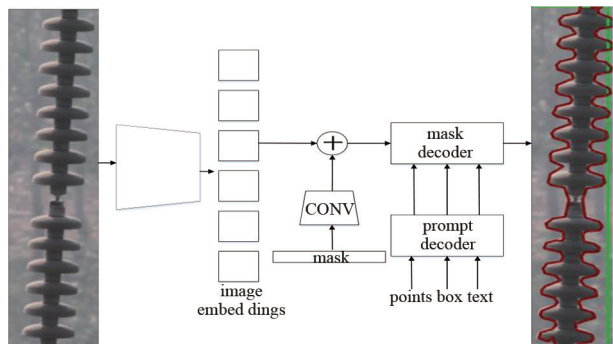
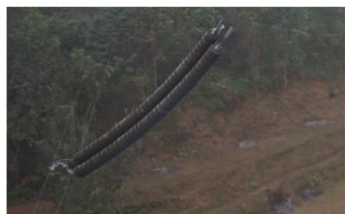


图 7 SAM 模型结构

数据的扩充步骤为:首先使用 SAM 将图片中目标对象的掩码自动进行分割,使用数据集中其他图片的背景,将分割出的包含缺陷的绝缘子串掩码进行缩放、变形、裁剪、旋转^[21]后再嵌入其中。扩充结果为:由原始的约 600 张图片扩展到 900 张;对绝缘子部分数据集进行处理前 mAP@0.5 为 0.492、mAP@0.95 为 0.285,处理后 mAP@0.5 提升到 0.781、mAP@0.95 提升到 0.462,相较处理前有显著的提升。原图和处理后的图片示例如图 8 所示,分别展示旋转、缩放处理后的结果。



(a) 原图



(b) 旋转扩充图



(c) 缩放扩充图

图 8 瓷质绝缘子破裂图片处理

除绝缘子缺陷外,数据集中还包括防震锤锈蚀、重锤锈蚀和塔身异物。原始数据集中存在数据量过大且有重复的现象,对数据进行清洗后保留了总计 5711 张图片,按照 8:1:1 分为训练集、测试集和验证集。

3.2 消融实验

为证明所提出各改进模块均具有有效性,设计了消融实验。实验结果如表 1 所示,模型 A 为基础 YOLOv8-n 模型,模型 B 为在骨干网络中加入 DCNv2;模型 C 为将 multi-CA 机制加入 DCNv2 中;模型 D 为将 CIoU 更改为 WIoU。表 1 数据均为各模型在同一环境下训练 100 轮次的结果。由表 1 可见,在参数量并未大幅度提高的前提下,相较于初始的 YOLOv8 模型,所提方法的 mAP@0.5 和召回率都有显著提升,其中 mAP@0.5 提高了 5.92%,召回率提高了 7.5%,结果如图 9 所示。改进后的热力图对比如图 10 所示,由图可知:YOLOv8 基础模型对缺陷产生位置的关注度仍有偏差且无法关注到大部分缺陷;加入 DCNv2 后,虽能检测出大多数缺陷但会出现误检状况,将 multi-CA 机制加入 DCNv2 中误检情况减少;加入 WIoU 后模型可基本识别到图片中的所有缺陷且关注度较为集中在缺陷区域。

表 1 消融实验结果

模型	mAP@0.5	参数量/ 10^6	GFLOPs	Recall
A	0.808	3.00	8.2	0.760
B	0.825	3.03	8.0	0.791
C	0.842	3.04	8.0	0.803
D	0.860	3.04	8.0	0.835

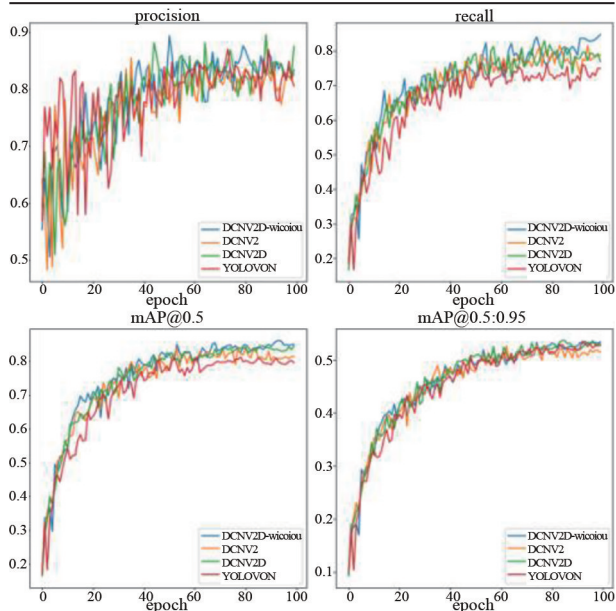
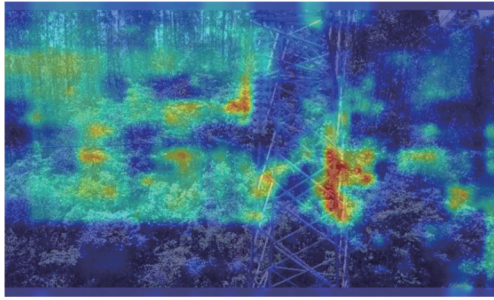
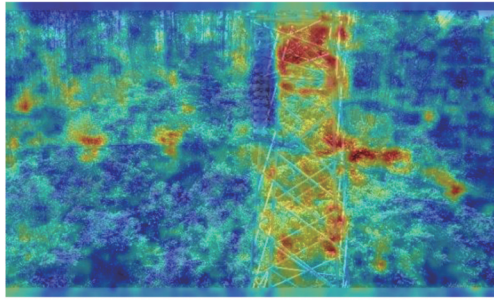


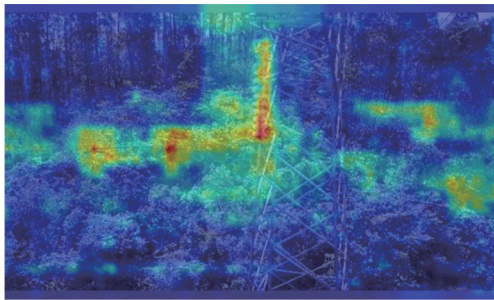
图 9 各项结果



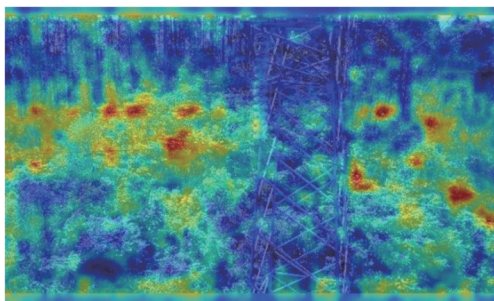
(a) YOLOv8



(b) YOLOv8+DCNv2



(c) DCNv2+multi-CA



(d) DCNv2+multi-CA+WIoU



(e) 标注结果

图 10 消融实验热力图对比

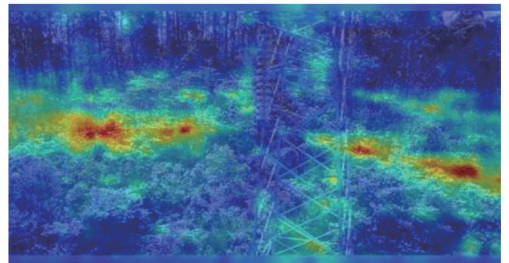
3.3 对比实验

为了体现所提算法在检测精度及参数量上的优

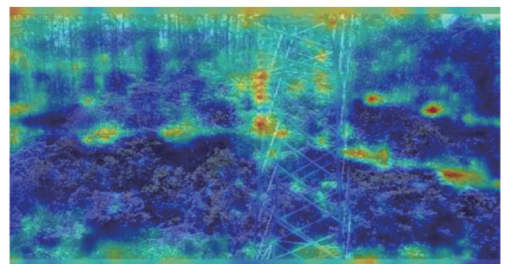
越性, 和主流可变形卷积网络进行对比, 例如 DCNv2-dyhead、DCNv2、DCNv3 以及可变形卷积注意力机制 Dattention^[22], 在相同数据集和软硬件条件下进行比较, 验证所提算法模型所作改进的有效性。对比结果如表 2 所示, 模型 E 为 Dattention; 模型 F 为 DCNv2-dyhead; 模型 G 为 DCNv3; 模型 H 为 DCNv2; 模型 I 为所提出的 DCNv2+Multi-CA。由表中数据可得, 相较于其他相似算法, 所提出的方法在 mAP@0.5 及召回率指标上皆处于领先位置。用于对比的其他模型热力图对比如图 11 所示, Dattention 对缺陷的检测有遗漏部分, DCNv2 存在少量误检现象, DCNv3 对大量位置关注度过高。

表 2 对比实验结果

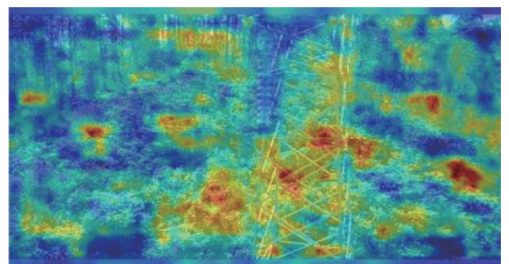
模型	mAP@0.5	参数量/10 ⁶	GFLOPs	Recall
E	0.825	3.07	8.1	0.758
F	0.827	3.03	8.0	0.791
G	0.819	2.89	8.0	0.749
H	0.825	3.03	8.0	0.719
I	0.842	3.04	8.0	0.803



(a) Dattention



(b) DCNv2-dyhead



(c) DCNv3

图 11 对比实验热力图对比

参考文献

- [1] 蒋玮,王铭华,陈锦铭,等.基于 Neo4j 图数据库的配电网供电可靠性计算 [J].电力系统自动化,2022,46(15):104-111.
- [2] 郑博文,杨隽,杨承辰,等.基于风险性指标的主动配电网理论可靠性评估 [J].供用电,2019,36(85):53-58.
- [3] 张伯明,陈寿孙,严正,等.高等电力网络分析[M].北京:清华大学出版社,2007.
- [4] 李子韵,成乐祥,王自桢,等.考配电自动化终端布局规划方法[J].电网技术,2016,40(4):1271-1276.
- [5] LZADI M, SATDARTAN A. A MIP model for risk constrained switch placement in distribution networks[J]. IEEE Transactions on Smart Grid, 2019, 10(4): 4543-4553.
- [6] 张笑弟,高强,潘弘.面向态势全感知的高弹性配电网一二次协同规划问题研究 [J].供用电,2022,39(3):3-8.
- [7] 郑玲玲,王铮,杨丽徙.不同网络结构及可靠性要求环境下 FTU 的最优配置 [J].电力系统保护与控制,2015,43(24):69-74.
- [8] BEZERRA J R, BARROSO G C, LEAO R P S, et al. Multi-objective optimization algorithm for switch placement in radial power distribution networks[J]. IEEE Transactions on Power Delivery, 2015, 30(2): 545-552.
- [9] CHEN C S, LIN C H, CHUANG H J, et al. Optimal placement of line switches for distribution automation systems using Immune algorithm[J]. IEEE Transactions on Power Systems, 2006, 21(3): 1209-1217.
- [10] 于腾凯,李晓军.基于 LCC 和改进 BDE 法的配电网开关优化配置 [J].电测与仪表,2015,52(13):96-100.
- [11] HEIDARI A, AGELIDIS V G, KIA M. Considerations of sectionalizing switches in distribution networks with distributed generation [J]. IEEE Transactions on Power Delivery, 2015, 30(3): 1401-1409.
- [12] 袁月,王守相,蒋菱,等.适应自愈要求的配电网开关优化配置 [J].中国电力,2016,49(3):76-82.
- [13] 孙磊,杨贺钧,丁明.配电系统开关优化配置的混合整数线性规划模型 [J].电力系统自动化,2018,42(16):87-95.
- [14] LZADI M, SAFDARIAN A. Financial risk evaluation of RCS deployment in distribution systems [J]. IEEE Systems Journal, 2019, 13(1): 692-701.
- [15] ABIRL-JAHROMI A, FOTUHI-FIRUZABAD M, PARVANLA M, et al. Optimized sectionalizing switch placement strategy in distribution systems [J]. IEEE Transactions on Power Delivery, 2012, 27(1): 362-370.
- [16] 王杰,王维庆,王海云,等.主动配电网中考虑条件风险价值的智能软开关的规划方法 [J].电力系统保护与控制,2022,50(2):1-11.
- [17] 郑益,朱俊澎,袁越,等.基于条件风险价值的风柴储孤岛微网经济风险评估 [J].电力自动化设备,2019,39(11):57-63.
- [18] 孙昊.考虑差异可靠性与故障后负荷响应的主动配电网供电能力评估[D].天津:天津大学,2018.
- [19] GROWE-KUSKA N, HEITSCH H, ROMISCH W. Scenario reduction and scenario tree construction for power management problems [C]//2003 IEEE Bologna Power Technology Conference Proceedings, June 23-26, 2003, Bologna, Italy. IEEE, 2003: 1-7.
- [20] 张成昊.考虑微能源网支撑作用的配电网韧性评估方法和提升策略研究[D].天津:天津大学,2019.
- [21] ALLAL R N, BILLINTON R, SJARIEF I, et al. A reliability test system for educational purposes-basic distribution system data and results [J]. IEEE Transactions on Power System, 1991, 6(2): 813-820.

作者简介:

江均(1981),男,硕士,高级工程师、高级经济师,研究方向为电网规划建设与项目管理等。

(收稿日期:2024-03-26)

(上接第 40 页)

- [21] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision [EB/OL]. [2024-03-14]. <https://arxiv.org/pdf/2103.00020>.
- [22] XIA Z F, PAN X R, SONG S J, et al. Vision transformer with deformable attention [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 4784-4793.

作者简介:

王译莹(2000),女,硕士研究生,从事输变电缺陷检测、

人工智能技术研究;

向思屿(1992),女,硕士,工程师,从事人工智能技术研究;

梁晖辉(1982),男,博士,高级工程师,从事电力信息安全;

邝俊威(1995),男,硕士,工程师,从事人工智能技术研究;
张菊玲(1990),女,高级工程师,从事网络安全技术研究;

刘松嘉(1996),男,硕士,助理工程师,从事电力人工智能、网络安全与数据管理研究。

(收稿日期:2024-04-28)

4 结 论

针对输变电线路上故障各类数量悬殊的问题,提出使用 SAM 模型和各图片变化方式相结合的形式以增加样本数量和质量,针对目标大小不一且检测正确率低的问题提出使用可变形卷积网络 DCNv2,由于原始的 DCNv2 只考虑单通道特征,在其中加入 multi-CA 注意力机制以达到特征更好融合的作用,将 CIoU 损失函数改为 WIoU 达到引导模型学习以适应不同质量的锚框,从而提高目标检测的性能。最终模型的 mAP@0.5 达到 0.86,相较于扩充数据前增加了 36.80%,相较于基础 YOLOv8 模型提高了 5.92%,与其他同类型算法相比具有优越性。

下一步的主要研究方向有两方面,一是增加缺陷的类型,使模型对更多缺陷类型都能达到较好的识别效果,适应于各类不同大小的样本;二是将模型进行轻量化处理包括但不限于嵌入轻量化模块、更改轻量化的主干网络等,使模型在参数量减少的情况下保持准确率。

参考文献

[1] 李冰,赵宽,白云山,等.融合注意力的多尺度光伏板缺陷检测方法[J].电力科学与工程,2023,39(8):1-10.

[2] 陈伟.基于无人机巡检图像的电力系统故障智能检测技术研究[D].哈尔滨:哈尔滨工业大学,2021.

[3] 邓亚平,李迎江.YOLO 算法及其在自动驾驶场景中目标检测综述[J].计算机应用,2024,44(6):1949-1958.

[4] 张明.单阶段无锚框目标检测算法研究[D].太原:山西大学,2023.

[5] REDMON J, DIWALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016:779-788.

[6] BLOICE M D, ROTH P M, HOLZINGER A. Biomedical image augmentation using Augmentor [J]. Bioinformatics, 2019,35(21):4522-4524.

[7] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023:7464-7475.

[8] 陈鹏,包倍源,陈旭.基于多尺度特征融合与重要性排序损失的遥感图像匹配网络[J].激光与光电子学进展,2024,61(18):276-284.

[9] 王永生,姬嗣愚.基于深度学习的目标检测算法综述[J].计算机与数字工程,2023,51(6):1231-1237.

[10] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional Networks [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. IEEE, 2017:764-773.

[11] ZHU X Z, HU H, LIN S, et al. Deformable ConvNets V2: More deformable, better results [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. IEEE, 2019:9300-9308.

[12] 刘向举,刘洋,蒋社想.基于 SimAM 注意力机制的 DCN-YOLOv5 水下目标检测[J/OL].重庆工商大学学报(自然科学版),1-9[2024-03-06].http://kns.cnki.net/kcms/detail/50.1155.n.20231020.1542.004.html.

[13] 赵继达,甄国涌,储成群.基于 YOLOv8 的无人机图像目标检测算法[J].计算机工程,2024,50(4):113-120.

[14] REZATOFIGHI H, TSOI N, GWAK J. Generalized intersection over union: A metric and a loss for bounding box regression [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. IEEE, 2019:658-666.

[15] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. IEEE, 2021:13708-13717.

[16] 康泰安,王宝莉,刘少航,等.输电线路金具及其缺陷深度学习检测方法综述[J].电力信息与通信技术,2022,20(11):1-12.

[17] 苏海军,薛任谦.基于图像分割及修复的数据生成[J].邮电设计技术,2023(7):49-53.

[18] 赵振兵,冯烁,席悦,等.大模型时代:电力视觉技术新起点[J].高电压技术,2024,50(5):1813-1825.

[19] KIRILLOV A, MINTUN E, RAVI N, et al. Segment Anything [C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), June 18-22, 2023, Paris, France. IEEE, 2023:3992-4003.

[20] DOSOVITSKIY A, BEYER L, KOLESNIKOV A. An image is worth 16x16 words: Transformers for image recognition at scale [EB/OL]. [2024-03-14]. https://doi.org/10.485501/arXiv.2010.11929. (下转第 84 页)