

基于改进 YOLOv8n 的电力设备表面缺陷检测研究

李金顺¹, 骆 斌¹, 吴抒源¹, 孙 雪¹, 张 革²,
张铎瀚², 王浩宇³

(1. 国网天津市电力公司 城南供电分公司, 天津 300000;

2. 国网天津市电力公司, 天津 300000;

3. 华北电力大学 计算机系, 河北 保定 071003)

摘要: 针对当前电力设备缺陷检测存在图像背景复杂、检测精确度低和识别效果差等问题, 提出一种基于改进 YOLOv8n 的电力设备表面缺陷检测方法; 该方法在 C2f 模块中引入 SaE 注意力机制, 增强主干网络对关键缺陷特征提取能力; 在颈部网络采用 BiFPN 优化特征融合层, 实现特征的跨尺度融合, 提升了模型多尺度缺陷检测性能; 设计融合 MSDA 注意力机制的 M-Detect 检测头, 强化模型对目标定位的精度; 使用 WIoU 作为损失函数, 提升模型对困难样本的检测性能; 实验结果表明, 改进后模型的 mAP 达到 83.8%, 较原始 YOLOv8n 模型提升 1.7%, 且满足实时检测要求, 证实了该方法在电力设备表面缺陷检测中的有效性。

关键词: 深度学习; YOLOv8n; 电力设备; 表面缺陷检测

Defect Detection Algorithm for Power Equipment Based on Improved YOLOv8n

LI Jinshun¹, LUO Bin¹, WU Shuyuan¹, SUN Xue¹, ZHANG Ge²,
ZHANG Zehan², WANG Haoyu³

(1. Chengnan District Power Supply Company, State Grid Tianjin Electric Power Company,
Tianjin 300000, China;

2. State Grid Tianjin Electric Power Company, Tianjin 300000, China;

3. Department of Computer, North China Electric Power University, Baoding 071003, China)

Abstract: To address the issues of complex image background, low detection accuracy, and poor recognition effect in current power equipment defect detection methods, a surface defect detection method for power equipment based on improved YOLOv8n is proposed. An SaE attention mechanism in the C2f module is introduced to enhance the ability of the backbone network to extract key defect features; The BiFPN in the neck network is used to optimize the feature fusion layer, which achieves the cross-scale fusion of features, thus improving the multi-scale defect detection performance of the model; The M-Detect detection head incorporating the MSDA attention mechanism is designed to strengthen the localization accuracy of the model on the targets; The WIoU is taken as a loss function to improve the detection performance of the model on difficult samples. Experimental results show that the mAP of the improved model reaches up to 83.8%, which is 1.7% higher than that of the original YOLOv8n model, which meets the real-time detection requirements, verifying the effectiveness of the method in the surface defect detection of power equipment.

Keywords: deep learning; YOLOv8; power equipment; surface defect detection

收稿日期:2025-01-22; 修回日期:2025-03-07。

基金项目:国网天津市电力公司科技项目(SGTJCN00YXJS2400817)。

作者简介:李金顺(1992-),男,大学本科,工程师。

通讯作者:王浩宇(2002-),男,硕士研究生。

引用格式:李金顺,骆 斌,吴抒源,等. 基于改进 YOLOv8n 的电力设备表面缺陷检测研究[J]. 计算机测量与控制,2026,34(2):39-45,79.

0 引言

在电力系统中, 电力设备的安全性和可靠性对于电力系统的稳定运行至关重要^[1]。然而, 对于长期暴露于自然环境之下的电力设备, 极易出现如破损、硅胶变色等各种表面缺陷^[2]。若这些缺陷不能被及时发现并加以处理, 极有可能导致设备故障, 甚至引发重大电力事故。因此对电力设备进行及时、准确的外部缺陷检测不可或缺。

传统的电力设备缺陷检测方法为人工巡检, 这种方法需要耗费大量人力物力, 而且检测效率低下、准确性难以保证。针对这些问题, 研究人员通过传统的图像识别方法, 使用基于具有梯度性的直方图特征^[3]、红外图像^[4]、基于声成像技术^[5]等方法对设备进行缺陷检测。但这类方法成本高, 而且成像效果容易受到环境因素影响, 无法准确定位缺陷位置。

随着计算机视觉和图像处理技术的飞速发展, 自动化的缺陷检测方法开始得到应用。目前, 基于深度学习的电力设备表面缺陷检测技术受到了广泛的关注^[6]。基于深度学习的目标检测方法层出不穷, 这些方法从结构上主要分为两大类: 一类是以 Faster R-CNN^[7] (区域卷积神经网络) 和 SPPNet^[8] 为典型代表的“两阶段”网络; 另一类是以 SSD^[9] (Single Shot MultiBox Detector) 和 YOLO (You Only Look Once) 系列^[10-11] 为代表的“一阶段”网络。两者之间最根本的差异在于, “两阶段”网络需要首先生成一系列潜在的目标候选区域, 然后对这些区域进行深入的特征提取和识别分析, 而“一阶段”网络则通过一次性的神经网络运算直接对目标的分类和位置进行预测。文献 [12] 针对电力设备目标检测中的定位困难问题, 对 YOLOv4 主干网络进行改进, 并引入焦点损失函数来解决因正负样本不平衡导致的检测精度低的问题, 最后实现了改进的 YOLOv4 算法用于电力设备表面缺陷检测。文献 [13] 基于 Faster R-CNN, 通过对输入图像进行数据增强, 在网络中添加 SPP 结构以及改进特征融合方式, 对分类以及边界框回归损失函数进行改进的方式来提高缺陷的检测精度。

尽管深度学习算法在电力设备外部缺陷检测方面表现出色, 但在实际应用中仍然面临着缺陷复杂、小目标缺陷检测精度低以及检测精度上仍有所欠缺。本文通过使用 YOLOv8n 算法, 对电力设备缺陷进行检测, 为提升模型的性能, 针对图像背景复杂问题, 在原模型基础上引入挤压聚合激励 (SaE) 注意力机制, 聚焦关键缺陷特征, 抑制图像背景信息和无关特征的干扰; 同时借鉴双向特征金字塔网络 (BiFPN) 的思想, 对特征融合层进行优化, 使其可以在不同尺度的特征层之间进行高效的双向融合。YOLOv8n 会在不同层次的特征图上检

测, 为了更聚焦目标物体, 引入多尺度空洞注意力机制 (MSDA) 构建新的检测头, 提升模型目标定位精度。最后, 为了提升模型对电力设备表面缺陷困难样本的检测性能, 使用 WIoU 损失函数替换原 CIoU 损失函数。改进后的 YOLOv8n 算法, 在检测精度上有所提升, 并且可以满足实时性的要求。

1 YOLOv8n 模型

YOLOv8n^[14] 模型的核心构架采用了 CSPDarknet^[15] 作为骨干网络, 同时结合了 PAN 和 FPN 的融合策略作为颈部网络, 采用解耦合的输出头部 (Decoupled-Head) 进行最终的目标分类与定位。在骨干和颈部网络的构建上, YOLOv8n 对传统的 C3 模块进行了改良, 引入了 C2f 结构, 以促进更加充足的梯度流动, 针对不同规模的模型, C2f 结构还巧妙地调整了通道数, 进一步优化了网络。图像在进入主干网络后, 经历多重卷积 (Conv) 和 C2f 模块的处理, 为了扩大模型的感受野采用 SPPF (Spatial Pyramid Pooling-Fast)^[16] 模块, 以有效提取更广泛的图像特征, 最后将结果传递到颈部层。检测头则采用了当前广泛使用的解耦头部结构, 对分类头和检测头进行了分离, 并且从基于锚点的方法 (Anchor-Based) 转变为无锚点的方法 (Anchor-Free)^[17], 使得模型对小、中、大 3 种尺寸的目标都有着出色的检测能力。YOLOv8n 模型结构如图 1 所示。

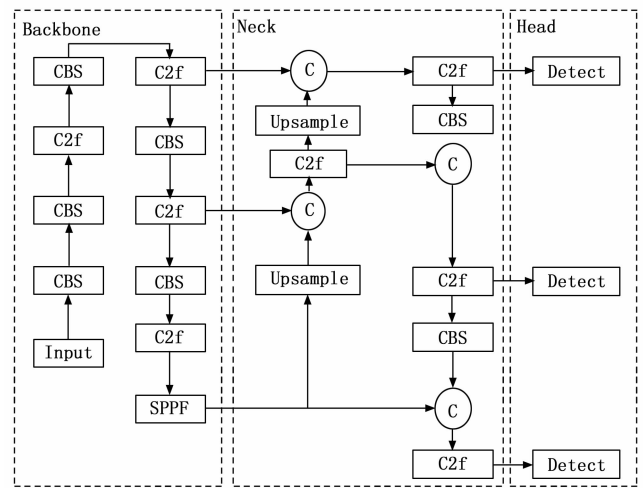


图 1 YOLOv8n 模型结构

2 改进 YOLOv8n 的电力设备缺陷检测模型

为有效检测电力设备的缺陷, 本文在 YOLOv8 的基础上提出一种改进模型, 结构如图 2 所示。图 2 中的 SaE-C2f 模块即为引入 SaE 后的 C2f 模块; M-Detect 为改进后的检测头。

2.1 SaE 改进 C2f 模块

变电站中复杂背景会对电力设备外部缺陷检测造成

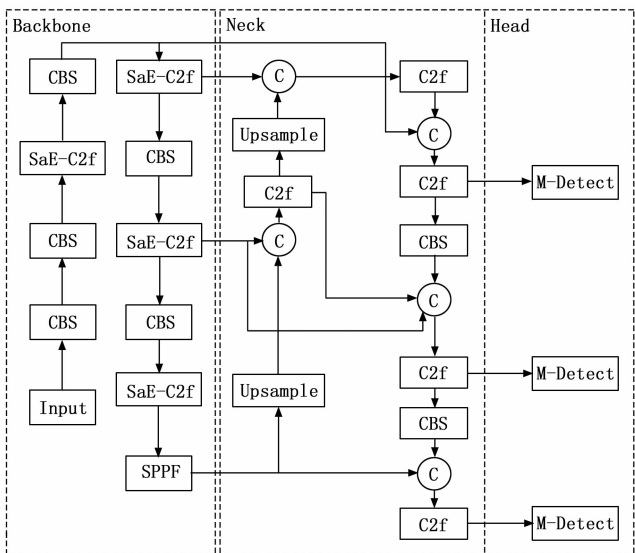


图 2 改进的 YOLOv8n 网络结构

一定干扰, 为让 YOLOv8 模型更加专注于包含电力设备缺陷的关键区域和关键特征, 抑制背景信息和无关特征的干扰, 在 C2f 模块中引入 SENetv2 网络中的 SaE 注意力机制, 提升模型性能。

SENetv2^[18] 是改进的 SENet^[19] 网络, 其中引入了一个升级的挤压和激励 (SE) 模块, 即 SaE 模块。与 SE 模块相比, SaE 模块利用多分支的稠密层来增强网络的特征表示能力, 从而提升模型的性能。聚合残差 (ResNeXt) 模块、SE 模块以及 SaE 模块之间的比较如图 3 所示。该图说明 SE 模块和 SaE 模块都选择性地传输关键特征。然而, SaE 模块通过增加层之间的基数来优化此阶段, 使得不同层次、不同类型的特征能够更有效地被传输和利用。

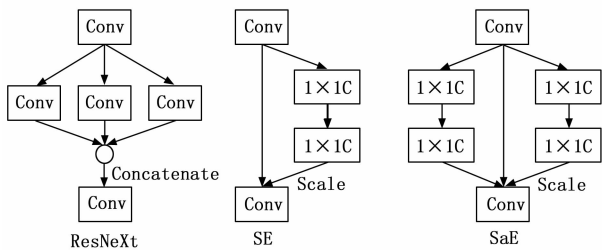


图 3 ResNeXt 模块、SE 模块以及 SaE 模块

在 SaE 模块中输入经过标准卷积操作后, 首先利用全局平均池化操作来挤压特征, 然后通过多分支全连接层和激活函数来获取通道权重, 最后对卷积特征进行缩放, 将其恢复到原始形式。然后将该缩放后的输出与输入连接起来, 具体表示如下:

$$SEnetV2 = x + F\{x \cdot Ex[\sum Sq(X)]\} \quad (1)$$

其中: x 表示为输入, $F\{\}$ 表示批量归一化和 dropout 操作, Sq 函数表示挤压操作, 包括一个 FC 层。在

Sq 操作之后, 执行 Ex 操作, 表示激励操作。在此之后, 将执行缩放操作, 也就是将 Ex 操作后的输出与模块的输入 x 相乘, 从而产生维度恢复。最后, 此缩放后的输出与输入连接。工作机制如图 4 所示。这样网络能够更有效地学习到通道之间的复杂关系, 增强了网络对关键特征的选择性传输能力。

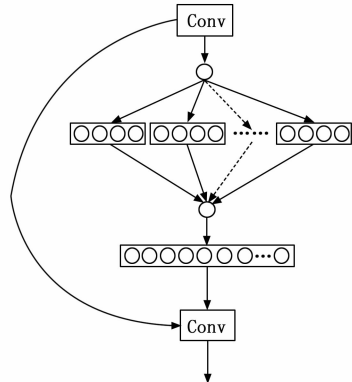


图 4 SaE 模块内部工作机制

在挤压操作中使用多分支的全连接层, 但基数设为 4, 同时, 缩减尺寸保持为 32, 限制了每个分支的参数数量, 从而 SaE 模块整体参数量增加有限。

为了提升模型对复杂多样缺陷的识别能力, 通过引入 SaE 对 YOLOv8 的 C2f 模块进行改进, SaE-C2f 模块结构如图 5 所示, 具体是在 C2f 模块的 Bottleneck 模块的第二个 Conv 后添加一个 SaE 模块。

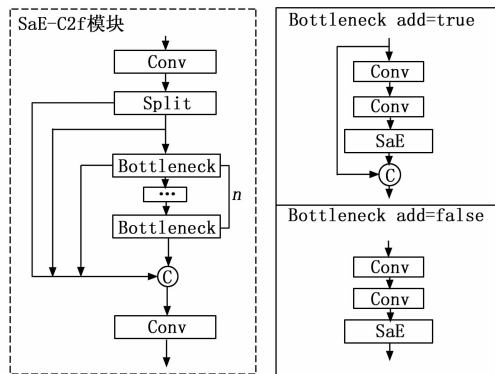


图 5 改进 C2f 模块结构

2.2 BiFPN 优化特征融合层

在电力设备缺陷检测中, 电力设备的部件大小各异, 缺陷特征也存在不同的尺度, YOLOv8n 本身的特征融合方式无法充分利用这些多尺度的特征, 通过引入 BiFPN, 利用其双向的跨尺度连接可以将高层语义特征和低层细节特征进行更有效的融合。

BiFPN^[20] 是一种高效的多尺度特征融合网络, 它在传统特征金字塔网络 (FPN) 的基础上进行优化。FPN、PANet、NAS-FPN 与 BiFPN 的网络结构的对比

如图 6 所示,从图中可以看出 FPN 引入了自顶向下的路径来融合多尺度特征,丰富低层次特征语义,整合了全局和局部信息;YOLOv8n 中使用的 PANet 网络则在 FPN 的基础上增加了自底向上的额外路径,将低层的细节信息反向传递到高层,与自顶向下的语义信息形成互补;NAS-FPN 则使用神经架构搜索 (NAS) 来找到不规则的特征网络拓扑,能够自动探索和发现更适合特定任务和数据集的特征网络拓扑结构,可以适应不同尺度特征,同时重复应用相同的高效块,减少模型的参数数量,降低计算复杂度,同时提高模型的训练和推理速度,但在搜索时耗时较长。

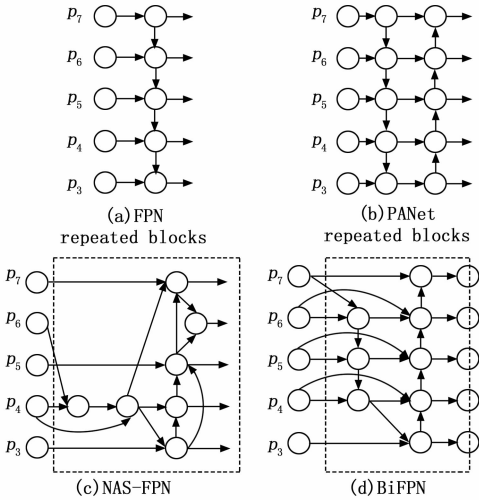


图 6 FPN、PANet、NAS-FPN 和 BiFPN 结构对比

BiFPN 则通过高效的双向跨尺度连接和重复的块结构,改进了准确度和效率之间的权衡。双向特征融合在 BiFPN 中指的是一种机制,它允许在特征网络层中的信息在自顶向下和自底向上两个方向上流动和融合,如图 6 所示。它增强了网络对特征融合的能力,使网络能够更有效地利用不同尺度的信息,从而提高目标检测的性能。

在 YOLOv8n 中引入 BiFPN 实现高效的多尺度特征融合与传递,有效地提升了目标检测模型的性能和表达能力。如图 2 所示,为了融合不同尺度的特征,首先在颈部网络中的第二个 C2f 模块后增加了一个特征融合模块 (Concat) 和 C2f 模块,然后将主干网络中的第二个和第三个 SaE-C2f 模块的输出分别输入到颈部网络中第三个和第四个 Concat 模块,实现不同尺度特征的融合。YOLOv8 结构与引入 BiFPN 后结构对比如图 7 所示。

2.3 M-Detect 检测头

YOLOv8n 模型在不同层次的特征图上进行检测,不同层次的特征图都会包含不同尺度的信息,其中有些信息会干扰对目标物体的检测,为了更好地融合不同尺度的信息,在检测时让模型聚焦于目标物体所在的区域,

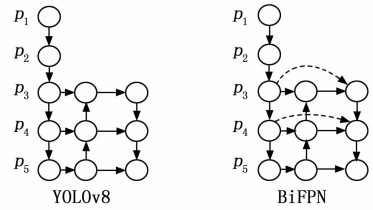


图 7 YOLOv8 结构与引入 BiFPN 后结构对比

减少不相关信息的干扰,引入 MSDA 改进原检测头^[21]。

不同的头部使用不同的空洞率执行 SWDA (Sliding Window Dilated Attention, 滑动窗口膨胀注意力) 模块,以实现特征图中有效的长程依赖建模。从形式上讲,SWDA 的描述如下:

$$X = \text{SWDA}(Q, K, V, r) \quad (2)$$

其中: Q 、 K 和 V 分别表示查询、键和值矩阵。3 个矩阵的每一行都表示一个查询/键/值特征向量。对于原始特征图中位置 (I, j) 处的查询,SWDA 稀疏地选择键和值,以便在大小为 $w \times w$ 的滑动窗口中以 (I, j) 为中心进行自我关注。此外,定义了膨胀 $r \in N^+$ 来控制稀疏度。具体而言,对于位置 (I, j) ,SWDA 操作的输出 X 的相应分量 x_{ij} 定义如下:

$$x_{ij} = \text{Attention}(q_{ij}, K_r, V_r) = \text{Softmax}\left(\frac{q_{ij} K_r^T}{\sqrt{d_k}}\right) V_r, \quad 1 \leq i \leq W, 1 \leq j \leq H \quad (3)$$

其中: H 和 W 是特征图的高度和宽度。 K_r 和 V_r 表示从特征图 K 和 V 中选择的键合值。

MSDA 的工作原理如图 8 所示。给定一个特征图 X ,通过线性投影获得相应的查询、键和值 (Q 、 K 和 V),之后,将特征图的通道划分为个不同的头,并在不同的头中以不同的膨胀速率进行多尺度 SWDA。具体而言,MSDA 可表示为:

$$h_i = \text{SWDA}(Q_i, K_i, V_i, r_i), 1 \leq i \leq n \quad (4)$$

$$X = \text{Linear}[\text{Concat}(h_1, \dots, h_n)] \quad (5)$$

其中: r_i 是第 i 个头的膨胀速率,膨胀率 $r=1、2$ 和 3 ,不同头部的参与接受野的大小分别为 $3 \times 3、5 \times 5$ 和 7×7 。 Q_i 、 K_i 和 V_i 则表示输入第 i 个头的特征图。

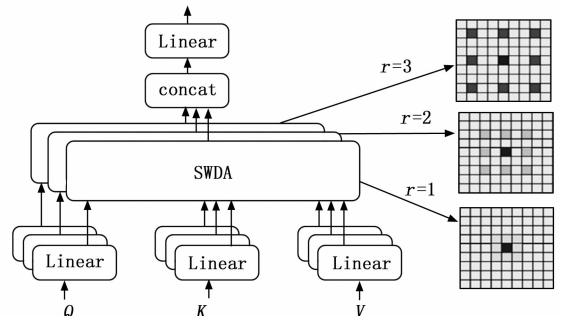


图 8 MSDA 的工作原理

输出 $\{h_i\}_{i=1}^n$ 被连接在一起, 然后发送到线性层进行特征聚合。

把 MSDA 与原模型中的检测头相结合, 组成一个新的检测头模块 M-Detect, 可以让模型检测时聚焦于目标物体所在的区域, 提高检测精度。M-Detect 模块结构如图 9 所示。

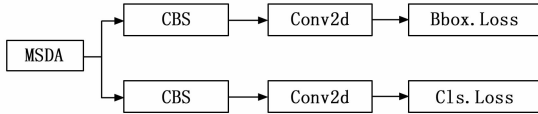


图 9 M-Detect 模块结构

2.4 改进损失函数

YOLOv8n 模型使用的是 $CIoU$ 损失函数, 计算公式如下:

$$CIoU = IoU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \quad (6)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \left(\frac{\arctan w^{gt}}{h^{gt}} - \frac{\arctan w}{h} \right)^2 \quad (8)$$

式中, b 和 b^{gt} 分别表示预测框与真实框的中心点坐标, $\rho^2(b, b^{gt})$ 表示两个中心点的欧式距离, c 代表两个图像的最小外接矩形的对角线长度, α 是权重函数, v 用来度量长宽比的相似性, w 和 h 分别表示框的高度和宽度。

使用 $CIoU$ 损失会通过距离、长宽比等几何因数加重困难样本的惩罚, 降低模型的泛化能力, 为了降低高质量锚框的重视, 同时也减少低质量样本所产生的有害梯度, 提高模型检测的整体性能, 将预测框回归损失中的 $CIoU$ 损失替换为 $WIoU$ 损失^[22]。 $WIoU$ 的计算公式如下:

$$L_{IoU} = 1 - IoU \quad (9)$$

$$R_{WIoU} = \exp \left[\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2) * } \right] \quad (10)$$

$$L_{WIoU} = R_{WIoU} L_{IoU}$$

$$L_{IoU} \in [0, 1], R_{WIoU} \in [1, e] \quad (11)$$

式中, (x, y) 为预测框的中心点坐标; (x_{gt}, y_{gt}) 为真实框的中心点坐标; W_g 和 H_g 为包含预测框和真实框的最小外接矩形的宽和高。 L_{IoU} 表示预测框和真实标注框的交并比损失, 改进低质量的锚框; R_{WIoU} 为距离加权交并比, 当预测框与标注框重合度较高时, 将显著减少高质量锚框的 R_{WIoU} , 并减少其对预测框和标注框中心点之间距离的关注。

3 实验结果与分析

3.1 实验环境

本文实验环境操作系统为 Linux Ubuntu 20.04,

CPU 为 13th Gen Intel (R) Core (TM) i9-13900KF, GPU 为 NVIDIA GeForce RTX 4090 24 GB 显存, 使用 12.0 版本的 CUDA 加速模型训练, 深度学习框架为 Pytorch2.1.1, Python 版本为 3.8.18。训练时, 超参数设置批大小 (Batch Size) 为 8, 训练周期 (Epochs) 为 300, 输入图像大小为 640×640 , 初始学习率 (Learning Rate) 为 0.01, 不采用预训练权重, 使用 SGD 优化器。

3.2 数据集

本实验使用自制的数据集, 选取某变电站电力设备中最常见的两种, 分别为表计和呼吸器, 不同区域的设备型号不同, 拍摄的角度也不同, 一共 3 320 张带有外部缺陷的图片进行标注作为实验数据集, 样例如图 10 所示。其中各类外部缺陷标签数量分布如下: 表盘模糊 725 个、表盘破损 704 个、外壳破损 443 个、硅胶变色 1 012 个、硅胶筒破损 91 个和油封油位异常 278 个。按 8:1:1 的比例将原始数据划分为训练集、验证集和测试集, 划分过程中保持各类别占比一致。

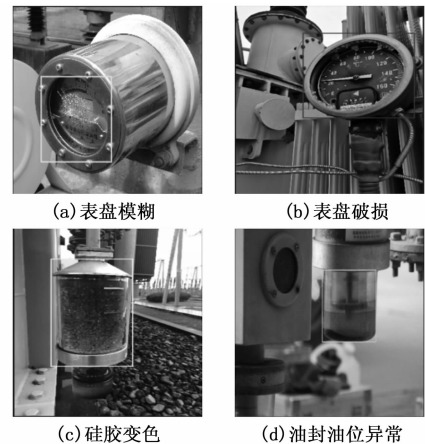


图 10 数据集样例

3.3 评价指标

为了验证改进模型对电力设备缺陷检测的提升, 使用目标检测中常用的评价指标对模型的性能进行全面评估, 本文使用精度 (Precision)、召回率 (Recall)、平均精度均值 (mAP, mean average precision)、模型参数量 (Params)、计算量 (GFLOPs)、Size (模型大小) 以及每秒处理帧数 (FPS, frames per second) 作为评估指标。各个评价指标的计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

$$AP = \int_0^1 PdR \quad (14)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (15)$$

$$FPS = \frac{Framenum}{ElapsedTime} \quad (16)$$

其中： TP 、 FP 和 FN 分别表示正确检测的数量、错误地将目标标记为正例的数量和没有检测出的目标数量； AP 表示 P - R 曲线与坐标轴围成的面积； N 表示待检测目标的类别数； $Framenum$ 代表检测图片总数； $ElapsedTime$ 表示检测所需总时间。

FPS 衡量的是模型每秒可以处理的图像帧数。 FPS 值较高则反映了模型在处理图像时的效率更高。在实际应用中， FPS 达到 30 帧/秒便可满足大多数实时检测的需求。

$GFLOPs$ 作为衡量模型或算法计算复杂性的指标，反映了模型在处理数据时的计算密度。而参数量则直接与模型的规模所关联，它表示构成模型的所有参数的总数。以上两个指标的值越低，则表明模型在运行时对计算资源的需求越小，对硬件性能的依赖也越少。

3.4 实验结果分析

3.4.1 对比实验

为了证明提出的电力设备缺陷检测模型的有效性，将本文提出的模型与 YOLO 系列模型、RT-DETR 模型、SSD 模型以及文献 [23] 提出的模型在实验数据集上进行了对比，实验环境和参数均相同，实验结果如表 1 所示。

表 1 模型对比实验结果

模型	$mAP/\%$	$GFLOPs$	参数量/ $\times 10^6$	FPS
YOLOv5	82.7	7.1	2.50	463.1
SSD	74.98	68.8	24	87.5
YOLOv6	82.0	11.8	4.23	475.4
YOLOv8n	82.1	8.1	3.0	528.6
AFL-YOLOv8n	82.5	11.5	1.8	341.0
YOLOv9c	82.9	102.3	25.32	201.5
RT-DETR	79.0	100.6	28.45	58.8
Ours	83.8	8.9	3.39	196.3

从表 1 可看出，本文提出的模型的 mAP 值为 83.8%，比 YOLOv5、YOLOv6、原模型、YOLOv9c、RT-DETR、SSD 模型以及 AFL-YOLOv8n 分别高 1.1%、1.8%、1.7%、0.9%、4.8%、8.82% 和 1.3%， mAP 作为检测算法的重要指标，表明本文改进的 YOLOv8n 模型在检测精度上具有一定优势，能够提供更可靠的检测结果；在参数量上，本文提出的模型比 YOLOv5 模型高 0.89×10^6 、比 YOLOv9c 模型低 21.93×10^6 、比 AFL-YOLOv8n 高 1.59×10^6 ；同时，由于加深了网络深度，在计算量上有所增加，比 v5 和 v8n 分别多 1.8 和 0.8 $GFLOPs$ ，但比 v6、v9c、RT-DETR、SSD 以及 AFL-YOLOv8n 分别低 2.9、93.4、91.7、59.9 和 2.6 $GFLOPs$ ；虽然本文提出的模型在帧率上略低一些，

只有 196.3 FPS，但考虑到其在检测性能和计算效率方面的综合优势，在实际应用中的实时性表现仍然较为出色，可以满足实时检测的要求。

为了进一步验证改进模型的有效性，在北京大学人工智能实验室公开的印刷电路板 (PCB) 瑕疵数据集进行了对比实验，其中一共 1 386 张图像，缺陷种类包含缺失孔、鼠标咬伤、开路、短路、杂散和伪铜，实验结果如表 2 所示。

表 2 PCB 数据集实验结果

模型	$Precision/\%$	$Recall/\%$	$mAP/\%$
YOLOv8n	92.5	87.9	92.0
Ours	94.0	88.3	92.8

从表 2 中可以看出，改进后的模型在 PCB 数据集上表现更加出色， mAP 值比原模型提升了 0.8%，证明了本文提出模型的有效性。

3.4.2 消融实验

为了验证本文改进方法的有效性，本节将以 YOLOv8n 模型作为基准模型，依次将 SaE 改进的 c2f 模块、BiFPN 网络结构、M-Detect 检测头和 WIoU 添加到基准模型中，分别训练和验证模型性能，对比实验的模型组成如表 3 所示，其中打对号表示添加了该模块，对比实验结果如表 4 所示。

表 3 模型改进模块

模型编号	SaE	BiFPN	M-Detect	WIoU
1				
2	✓			
3		✓		
4			✓	
5				✓
6	✓	✓		
7	✓	✓	✓	
8	✓	✓	✓	✓

表 4 消融实验结果

模型	$mAP/\%$	$GFLOPs$	参数量/ $\times 10^6$	Size/MB
1	82.1	8.1	3.0	6
2	82.4	8.1	3.02	6
3	83.0	8.2	3.03	6
4	82.7	8.7	3.35	6.6
5	82.6	8.1	3.0	6
6	83.3	8.3	3.04	6.1
7	83.5	8.9	3.39	6.8
8	83.8	8.9	3.39	6.8

从表 4 中可知，每个改进模块均实现了不同平均精度的提升，模型 2、模型 3、模型 4 和模型 5 与原模型相比， mAP 值提升 0.3%、0.9%、0.6% 和 0.5%；模

型 2 和模型 5 的计算量与模型 1 相同, 模型 3 与模型 1 相比计算量仅提高了 0.1 GFLOPs, 模型 4 与模型 1 相比计算量提高了 0.6 GFLOPs, 说明添加 SaE 模块、BiFPN 和更换 WIoU 并不会增加过多计算量, 但添加 M-Detect 检测头会增加计算量, 但更好地提升模型检测能力和效果; 添加 SaE 注意力机制让模型更加专注于包含电力设备缺陷的关键区域和关键特征, 抑制背景信息和无关特征的干扰, 提升模型的精度; 引入 BiFPN 可以有效优化网络结构, 提升模型性能; 更换 WIoU 使模型可以专注于普通质量的标注框, 提升检测精度。模型 6 和模型 2 相比 mAP 提高了 0.9%, 说明在添加 SaE 模块的基础上添加 BiFPN 可以有效提升模型性能。模型 1、模型 2、模型 3 和模型 5 的 $Size$ 均为 6 MB, 模型 4、模型 7 和模型 8 的 $Size$ 比模型 1 的 $Size$ 分别大 0.6、0.8 和 0.8 MB, 说明单独添加 SaE 模块、BiFPN 和更换 WIoU 对模型大小影响较小, 而添加 M-Detect 会使模型尺寸增大。由模型 8 可知, 将 4 种改进同时引入原模型, mAP 值提升 1.7%, 参数量提升 0.39×10^6 , 计算量提高 0.7 GFLOPs, 在增加少量计算资源的需求的同时, 有效的提升了模型的检测精度。

3.5 实验结果对比

为了更直观地了解改进后的检测效果, 选取缺陷样本进行检测, 检测结果如图 11 所示, 其中, 左边为原模型检测结果, 右边为改进模型检测结果。

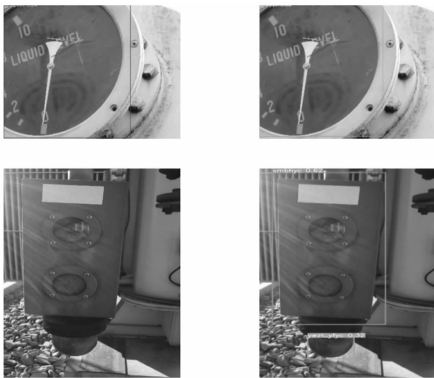


图 11 部分检测结果

从图 11 可以看出原模型有漏检和误检的情况, 而改进后的模型准确的识别出了缺陷种类, 比如第一张图像原模型错误的将表盘破损检测为表盘模糊, 而改进模型正确识别; 第二张图像原模型出现漏检, 改进模型准确识别出缺陷。同时经过改进后的模型在目标的定位和准确度都有所提升, 证明了本文提出算法的有效性。

4 结束语

电力设备多部署于户外复杂场景, 尤其变电站设备常与金属架构混合, 设备缺陷检测的场景复杂, 检测精度存在瓶颈。本文针对这些问题, 基于 YOLOv8n 算法

进行改进, 首先构建了融合注意力机制与跨尺度特征融合的协同优化方法, 通过 SaE 模块增强关键缺陷特征选择能力, 结合 BiFPN 实现多层次语义信息互补, 有效提升了复杂背景下的目标区分度; 其次, 设计基于 MSDA 注意力机制的检测头结构, 通过动态权重分配强化了微小缺陷的定位精度; 最后是采用了 WIoU 损失函数。实验结果表明, 所提算法在增加少量参数量和计算量的基础上, 检测精度提升 1.7%, 综合优于对比模型, 适合对检测精度有要求且计算资源有限的场景。未来, 我们将对模型进一步优化, 以进一步提升检测精度和速度, 以适应电力设备智能巡检的需要。

参考文献:

- [1] HUANG Q, JING S, LI J, et al. Smart substation: state of the art and future development [J]. IEEE Transactions on Power Delivery, 2016, 32 (2): 1098 - 1105.
- [2] TAO X, ZHANG D, WANG Z, et al. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, 50 (4): 1486 - 1498.
- [3] 林克正, 张元铭, 李昊天. 分块 Gabor 结合梯度直方图的特征提取算法 [J]. 小型微型计算机系统, 2019, 40 (12): 2662 - 2666.
- [4] WANG K, ZHANG J, NI H, et al. Thermal defect detection for substation equipment based on infrared image using convolutional neural network [J]. Electronics, 2021, 10 (16): 1986.
- [5] 邵宇鹰, 王 泉, 彭 鹏. 基于声成像技术的电力设备缺陷检测方法研究 [J]. 中国测试, 2021, 47 (7): 42 - 48.
- [6] CUI Y, HUANG X, ZHANG X, et al. A defects detection system for substation based on YOLOX [C] //2022 IEEE 5th International Electrical and Energy Conference (CIEEC). New York: IEEE, 2022: 4703 - 4707.
- [7] REN S, HE K, 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, 2016, 39 (6): 1137 - 1149.
- [8] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9): 1904 - 1916.
- [9] LIU W, ANGELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] //Computer Vision-ECCV 2016, Cham: Springer International Publishing, 2016: 21 - 37.

(下转第 79 页)