

DOI:10.19651/j.cnki.emt.2518804

# iMMI-YOLO: 基于残差模块的变压器缺陷检测算法<sup>\*</sup>

吴宇浩 朱文忠

(四川轻化工大学计算机科学与工程学院 宜宾 644000)

**摘要:** 针对当前变压器外观缺陷检测存在精度低、模型在复杂场景下泛化能力不足的问题,本文利用残差结构在特征融合与信息保留方面的优势,设计出3个模块对YOLOv11n进行改进。首先设计一种倒置残差注意力机制iEMA,其能够有效利用长距离依赖关系,旨在提升变压器缺陷检测的准确性;其次借助深度可分离卷积在多尺度特征提取方面的优势以及残差结构的特性,设计出MSCB结构,增强模型特征提取与融合能力;由于YOLOv11检测头对目标周围的上下文信息利用不够充分导致漏检,本文提出MR-Detect检测头,其设计理念融合了分组卷积和残差结构的想法,为后续的分类提供了丰富的特征表示。最后将非极大值抑制算法与Inner\_MPDIoU相结合,以解决传统损失函数在不规则物体和尺寸变化较大的物体上检测的局限性。实验结果表明,本文改进算法相比YOLOv11n,在保证实时检测的同时,mAP@0.5提升了5.9%,召回率提升2.8%,在复杂变压器工况检测场景下检测精度更高,能更有效地检测出各类缺陷。

**关键词:** 残差结构;变压器;缺陷检测;YOLOv11;iEMA

**中图分类号:** TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.60

## iMMI-YOLO: Transformer defect detection algorithm based on residual module

Wu Yuhao Zhu Wenzhong

(School of Computer Science &amp; Engineering, Sichuan University of Science &amp; Engineering, Yibin 644000, China)

**Abstract:** To address the issues of low accuracy and insufficient generalization ability of models in complex scenarios for current transformer appearance defect detection, this paper leverages the residual structure's merits to improve YOLOv11n with three modules. Firstly, an inverted residual attention mechanism, iEMA, is designed. It can effectively utilize the long-distance dependency and aims to improve the accuracy of transformer defect detection. Secondly, by leveraging the advantages of depthwise separable convolution in multi-scale feature extraction and the characteristics of the residual structure, an MSCB structure is designed to enhance the feature extraction and fusion capabilities of the model. Since, to address missed detections due to insufficient contextual utilization by YOLOv11's detection head, we design the MR-Detect head. It integrates grouped convolution and residual concepts, offering rich feature representations for classification. Finally, the non-maximum suppression algorithm is combined with Inner\_MPDIoU to address the limitations of traditional loss functions in detecting irregular objects and objects with large size variations. Experimental results show that compared with YOLOv11n, the improved algorithm in this paper, while ensuring real-time detection, increases the mAP@0.5 by 5.9% and the recall rate by 2.8%. It has higher detection accuracy in complex transformer operating condition detection scenarios and can more effectively detect various types of defects.

**Keywords:** residual structure; transformer; defect detection; YOLOv11; iEMA

## 0 引言

变压器作为电力系统中的关键设备,能够影响电力传

输和分配的效率和稳定性,因此对变压器的运行状态进行检测具有重要意义<sup>[1]</sup>。传统的巡检方式为设立电力安全督查制度,依靠人工巡视、手工纸介质记录的工作方式,然而

收稿日期:2025-05-13

<sup>\*</sup> 基金项目:四川省科技研发重点项目(2023YFS0371)、四川省智慧旅游研究基地(ZHYJ24-01)、企业信息化与物联网测控技术四川省高校重点实验室开放基金(2024WYJ03)项目资助

这几种方式受人为主观因素影响较大,也容易出现漏检、效率低下和管理成本高等问题,且变电站环境复杂,变压器周围存在高压电、强电磁干扰等影响,具有一定的安全风险<sup>[2]</sup>。

随着机器学习和深度学习技术的不断发展,智能化的变压器运行状态检测技术应运而生。国内外研究者重点关注利用图像处理 and 计算机视觉技术,结合深度学习模型来识别变压器的缺陷,如腐蚀、漏油、套管和硅胶破损等。这类方法通过捕捉变压器表面的视觉特征,在提升检测精度的同时实现了非接触式的检测方法,大大减少了人工巡检的工作量,更方便获取变压器的运行状况信息,在必要时再进行现场处理,从而保障运维人员的安全。在国内,杨昱宸等<sup>[3]</sup>利用变压器油的荧光特性,通过荧光实验检测变压器漏油情况,但此方法只适用于夜间检测,在户外日照情况时,阳光中的紫外线会显著增加环境的光亮度,导致荧光的反射效果被淹没,无法由肉眼直接看出。针对此问题,陆志欣等<sup>[4]</sup>使用 Mobile-ViT 替换 YOLOv4 原有的主干网络,并将卷积和 Transformer 相结合,以解决室外变压器漏油的检测问题,但此模型相比其他方法仍有较大的参数量。王铖浩等<sup>[5]</sup>对 YOLOv5s 进行轻量化处理,替换原有的损失函数为 WIoU,在牺牲小部分精度的情况下,大大降低了模型的参数量。田喆文<sup>[6]</sup>设计了一种新型金字塔结构,利用稳态学习的方式提高对电力设备缺陷检测的准确性。但其使用的数据集均为基本可见光范围内,且尚未考虑目标遮挡的情况。全欣宇等<sup>[7]</sup>选取 YOLOv8n 为基线模型,完成对变压器红外图像的组件识别。李坤等<sup>[8]</sup>在 YOLOv8 的基础上,为其添加了坐标注意力机制,并将网络结构中的普通卷积替换为可变形卷积,旨在提高对小目标的精度,并提出将 Slim-neck 代替原有的颈部网络,在保证检测精度的同时进一步降低计算量。位一鸣等<sup>[9]</sup>利用视觉几何组网络(visual geometry group network, VGG-Net)的图像迁移算法对变压器外部缺陷的数据集进行扩充,使得训练模型在实际巡检中具备较高的可行性。为了提高进一步提高模型在实际应用中的泛化性,刘冬<sup>[10]</sup>选取了 6 种变压器设备外部缺陷类型,基于残差特征对特征金字塔网络(feature pyramid network, FPN)增强,在网络中使用多尺度可变形注意力模块,但是对于变压器中表皮锈蚀的检测精度有待提升。

在国外,利用计算机视觉技术仍为变压器检测的方向之一。其通过改进神经网络架构以适应复杂光照与遮挡场景,或融合多模态数据增强模型鲁棒性,解决检测过程中的实时性与高精度难以平衡等问题。例如, Ghorbani 等<sup>[11]</sup>通过迁移学习,利用 VGG16 检测图像中是否存在渗漏油,随后,使用语义分割网络进行像素级溢油检测。Chen 等<sup>[12]</sup>将残差网络的思想引入到 YOLOv3 中,将变压器漏水、漏油的精度提升至 98.2%,但是其 FPS 仅为 8,难以满足实际工程的检测需求。Peng 等<sup>[13]</sup>在选取 YOLOv5 为基线模型在网络中加入卷积块注意力模块<sup>[14]</sup>(convolutional block

attention module, CBAM)注意力机制,用于检测电厂内管道的漏油情况,然而这种方法只适用于室内地面光滑时的检测,对于室外的变压器漏油检测性能较差。Cheng 等<sup>[15]</sup>提出了检测变压器上套管破损的 AdIn-DETR 模型,并通过实验验证了该方法的泛化性。Li 等<sup>[16]</sup>提出了一种两阶段的变电设备缺陷检测模型,第一阶段利用分割算法提取目标设备,第二阶段通过缺陷识别方法来识别红外图像中的变电设备存在的缺陷。Chen 等<sup>[17]</sup>通过改进 YOLOv4 主干网络、引入焦点损失函数等方式有效解决了变压器等设备生锈、漏油等外部缺陷问题,但是检测效率较低。也有学者通过数据融合的方式对电力设备进行缺陷检测<sup>[18]</sup>:首先基于结构化时间序列数据构建递归图,其次利用卷积神经网络从递归图和现有图像数据中提取特征,从而产生不同的特征向量,然后使用自适应加权合并这些向量从而实现对电力设备的缺陷检测。该方法不仅提高了设备状态感知的准确性,还通过多模态数据特征的综合分析,增强了对电力设备整体状态的理解,为实际应用提供了新的思路。

尽管近年来基于深度学习的变压器缺陷检测研究成果颇多,但现有方法仍存在诸多亟待解决的问题。第一,数据集的局限性成为制约模型性能提升的关键因素。公开可用的变压器缺陷数据集不仅数量稀少,且数据分布不均衡,难以覆盖实际场景中复杂多变的缺陷类型与环境条件;第二,检测模型在复杂场景下的泛化能力不足,多数研究仅针对常规场景中的某类单一缺陷进行优化,在面对光照变化、视角差异、多缺陷等复杂工况时,检测精度和稳定性下降;第三,由于变压器长期处于潮湿的环境中,会不可避免地造成表皮锈蚀的情况,这些锈蚀颜色、形状、大小不一,模型对此类缺陷的检测能力较弱,难以实现缺陷的早期精准识别与预警。

同时,在目标检测领域, YOLO 系列因兼顾速度与精度成为主流选择。然而 YOLOv3 采用精简结构,但特征提取能力有限,对小目标和复杂特征检测效果不佳; YOLOv5 通过跨阶段局部连接和特征融合优化计算效率,但在复杂场景和小目标检测上仍有提升空间; YOLOv8 和 YOLOv10 引入新特征提取模块增强信息获取能力,但在处理密集目标时会出现漏检问题。而 YOLOv11 通过增强的特征提取方式更能准确地捕捉图像中的复杂细节。当前对 YOLOv11 的改进大多以特征提取与融合为核心,通过改进网络结构和模块设计,提升模型对目标特征的表达能力。例如包广斌等<sup>[19]</sup>通过改进 SPPF 以及 C2PSA 模块进一步增强特征提取能力,但其策略仅聚焦于局部特征聚合模块的结构优化,未引入长距离依赖建模机制,难以捕捉像素级全局依赖关系。殷旭鹏等<sup>[20]</sup>主要在颈部网络部分,采用自适应分层特征融合网络,通过自适应加权融合多尺度特征,以增强对主干网络信息的利用,但该方法侧重于权重分配,对细节特征的保留能力较弱。

综上所述,本文针对变压器缺陷检测中模型对复杂工

况场景检测的泛化性低、缺陷尺寸差异大造成的精度低等问题,构建出包含复杂工况的变压器缺陷数据集,选取 YOLOv11 作为基线模型,利用残差结构在信息保留方面的优势,提出 iMMI-YOLO 的算法。本文主要贡献如下:

1)设计倒置残差注意力 iEMA:该模块增强信息交互能力,捕捉像素级别关系,适应不同缺陷尺寸,动态计算特征权重,更好学习缺陷特征,进而提高变压器缺陷检测的准确性;

2)设计 MSCB 模块:其能够融合多尺度特征与通道信息,并将其与 C3k2 结合,进一步增强 YOLOv11 的特征提取与融合能力;

3)设计 MR-Detect 检测头:其设计理念融合了分组卷积和残差结构的思想,并利用 MLCA 双路径处理特征图的方式强化特征融合效果,为类别分类提供丰富特征表示;

4)改进损失函数:将非极大值抑制算法(Soft-NMS)与 Inner\_MPDIoU 相结合。Inner\_MPDIoU 能够准确衡量预测与目标框的差异,而 Soft-NMS 优化预测框筛选,将二者结合,通过优化边框回归和抑制重叠框的方式,解决传统损失函数对尺寸变化较大的物体不敏感的局限性,并使模型在复杂场景下的检测结果更加准确可靠。

### 1 改进的 YOLOv11n 算法

#### 1.1 iEMA 模块

倒置残差模块<sup>[21]</sup>(inverted residual mobile block, iRMB)结构结合了卷积神经网络(convolution neural networks,CNN)的轻量级特性和 Transformer 的动态处理能力,并通过倒置残差的设计改进了信息流的处理,允许在保持模型轻量的同时捕捉和利用长距离依赖。但它对于不同通道信息的关注程度较为平均,难以充分捕捉每个通道的关键信息,导致在处理复杂任务时会忽略重要特征,且在处理图像数据时跨维度交互能力较弱,影响模型在面对要求较高的任务时对细节的把控。针对此问题,本文重构 iRMB 模块,设计出 iEMA。输入特征先经过 BatchNorm2d,为后续处理提供稳定的数据分布。接着经过高效多尺度注意力机制<sup>[22]</sup>(efficient multi-scale attention,EMA)模块,在该模块中,会对特征图在不同尺度下进行分析,挖掘特征间的关联和重要程度。其将这两个经过全局平均池化的特征向量分别送入后续的卷积层进行处理,挖掘特征之间更复杂的关联;经过卷积处理后的特征向量,再通过卷积进行通道维度的调整 and 融合,然后将处理后的结果输入到 Sigmoid 函数中,而 Sigmoid 函数会将输出值映射到 0~1 之间,得到的输出值就是不同通道和空间位置上特征的重要性得分;而后对原始输入特征图的不同通道和空间位置上的特征进行加权操作,如果某个通道或空间位置的重要性得分较高,则在加权后,该位置的特征值就会相对增强,实现动态计算特征权重的功能。在 EMA 模块处理之后,使用普通的卷积进一步细化特征。中间的

残差连接将 EMA 输出特征与经过卷积处理的特征相加,这种方式融入了卷积提取的新特征,丰富了特征表示。在另一个残差分支中将原始输入特征与经过一系列处理的特征相加,使得网络能更好地学习特征差异。其结构如图 1 所示。

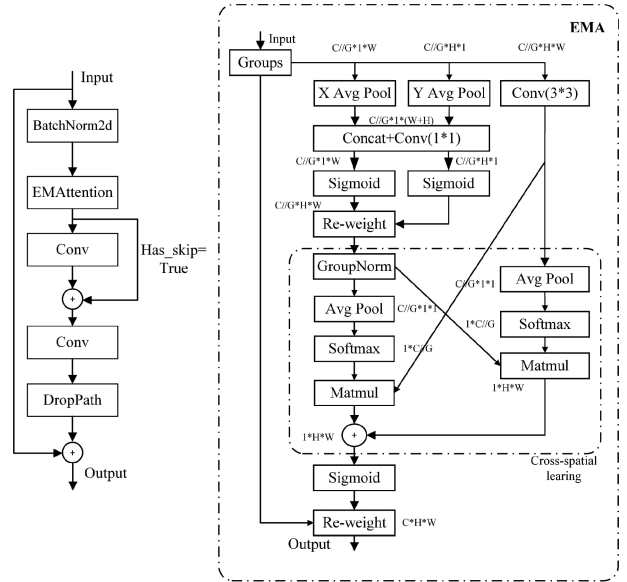


图 1 iEMA 结构图

Fig. 1 iEMA structure diagram

iEMA 不仅继承了 iRMB 结构优势,还能够通过全局信息编码来重新校准每个并行分支的权重,这种全局视野的引入使模型能够在不丢失局部细节的同时,确保每个并行分支的权重与全局上下文保持一致,进而提升检测的鲁棒性。iEMA 还增强了信息交互能力,其通过跨维度交互来捕捉像素级别关系,这种交互方式使模型能够适应不同缺陷的尺寸,动态地计算特征权重,从而更好地学习每种缺陷的特征。

#### 1.2 MSCB 模块

传统卷积使用固定大小的卷积提取特征,这种方式无法全面捕捉不同大小的缺陷特征。多尺度卷积块(multi-scale convolutional block, MSCB)是利用<sup>[23]</sup>高效多尺度卷积注意力解码(efficient multi-scale convolutional attention decoding, EMSDC)设计的模块,主要用于增强特征表示。它通过多尺度深度可分离卷积,利用不同大小的卷积核并行提取多尺度特征,较小的卷积能捕捉图像中的局部细节特征,而较大的卷积通过大的感受野获取大尺度目标的整体信息,以此做到把握细节和整体信息之间的平衡。此外, MSCB 在 EMSDC 之前引入逐点卷积,其通过对每个像素点在所有通道上进行加权求和,实现了各个通道信息的融合。在图像特征中,不同通道往往承载着不同类型的信息,有的通道侧重于颜色信息,有的侧重于纹理信息等。逐点卷积能够将这些分散在不同通道的信息整合起来,使得后续的多尺度深度可分离卷积操作能够基于更综合、更具代

代表性的特征进行处理。MSCB结构则在这二者的基础上,利用残差结构实现特征融合与增强。通过不同尺度的深度可分离卷积获取多尺度特征后,残差连接将原始输入特征与经过卷积处理的特征相加,这样既保留了原始特征中的重要信息,又融入了卷积提取的新特征,丰富了特征表示。其结构如图2所示。

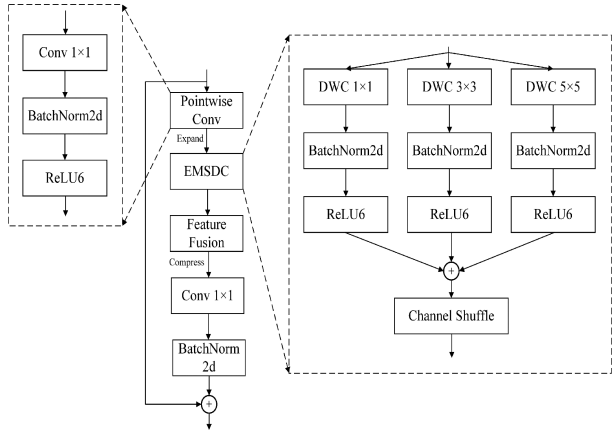


图2 MSCB结构图

Fig. 2 MSCB structure diagram

将MSCB与C3k2进行结合,进一步增强YOLOv11的特征提取与融合能力,其结构如图3所示。

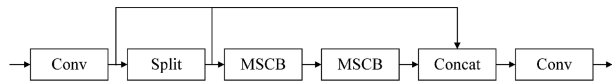


图3 C3k2\_MSCB结构图

Fig. 3 C3k2\_MSCB structure diagram

1.3 MR-Detect

针对YOLOv11原检测头在检测目标时,对目标周围的上下文信息利用不够充分导致漏检等问题,本文设计了MR-Detect检测头。在MR-Detect中,每个Block都包含了混合局部通道注意力(mixed local channel attention, MLCA)模块。MLCA作为一个高效的注意力模块,能够结合局部和全局特征以及通道和空间特征的信息输入特征

进行检测。MLCA模块通过双路径处理特征图:其一,先经局部平均池化(local average pooling, LAP)和全局平均池化(global average pooling, GAP),分别聚焦局部区域与捕捉整体统计信息,再经1D卷积、上采样,生成局部通道注意力特征;其二,通过平均池化聚合全局信息,经1D卷积和重塑后,用残差连接的方式与前者融合。最终将融合特征上采样,以此融合局部与全局、通道与空间特征,增强特征表达。MLCA结构如图4所示。

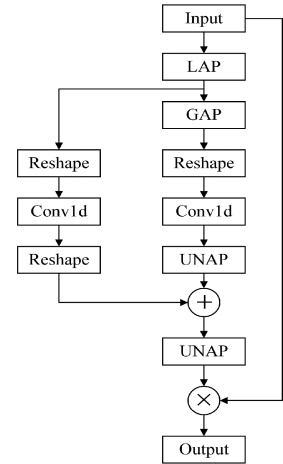


图4 MLCA结构图

Fig. 4 MLCA structure diagram

对于MR-Detect的设计还采用了分组卷积的思想,将输入的特征图分成多个组,每个组独立地进行卷积操作,最后将所有组的输出拼接起来。在这种结构中多分支的分组卷积能够同时提不同类型的特征,而后将这些不同的特征组合起来,模型能够获得更丰富、更全面的特征表示,从而提高目标检测的精度。此外,还利用了残差结构允许原始输入特征直接参与到后续的特征计算中的特点,保留了原始输入特征中的重要信息,又融入了经过卷积和通道注意力调整后的新特征,实现了信息更为全面的特征融合方式。综上所述,本文所设计的MR-Detect结构图如图5所示。

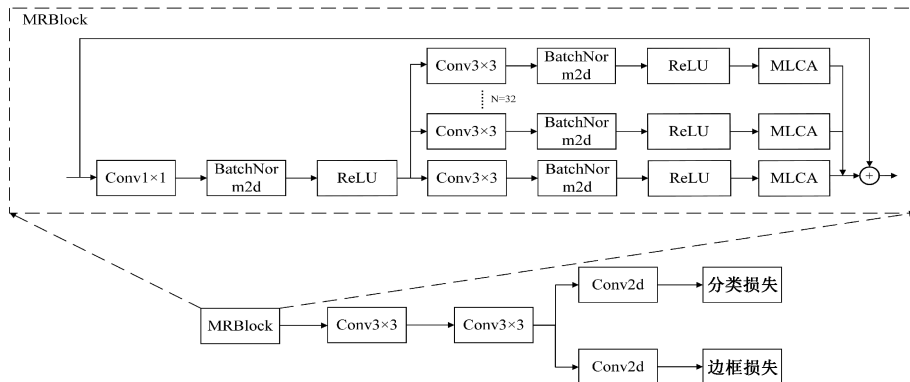


图5 MR-Detect结构图

Fig. 5 MR-Detect structure diagram



### 1.4 改进的 Inner\_MPDIoU

为了解决 CIoU 不能适应不规则物体的检测以及对尺寸变化较大的物体不敏感的局限性,本文设计了一种新的边界框回归优化方法,以改进模型的性能。这种方法是在 Inner\_IoU 的基础上,加入 MPDIoU 的思想并结合 Soft\_NMS 形成的一种损失函数。

Inner\_MPDIoU 是由 Inner\_IoU 和 MPDIoU 结合形成的损失函数。Inner\_IoU 引入了比例因子 ratio 比,控制辅助边界框的比例大小, ratio 通常在  $[0.5, 1.5]$  之间。当 ratio 较大时,计算差异的边框尺度相对较大,在检测大尺寸缺陷时能更好地捕捉整体特征;当 ratio 较小时,使用较小尺度的辅助边框计算差异,对小尺寸缺陷的检测更加敏感,能够突出小缺陷的特征,更快学习到小缺陷的特征表示,加速收敛,使得模型能够在收敛速度和泛化能力之间取得较好的平衡。其定义如式(1)~(7)所示,式(8)为 Inner\_IoU 的损失公式。

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} \times ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} \times ratio}{2} \quad (1)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} \times ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} \times ratio}{2} \quad (2)$$

$$b_l = x_c - \frac{w \times ratio}{2}, b_r = x_c + \frac{w \times ratio}{2} \quad (3)$$

$$b_t = y_c - \frac{h \times ratio}{2}, b_b = y_c + \frac{h \times ratio}{2} \quad (4)$$

$$inter = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) \times (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)) \quad (5)$$

$$union = (w^{gt} \times h^{gt}) \times (ratio)^2 + (w \times h) \times (ratio)^2 - inter \quad (6)$$

$$IoU^{inner} = \frac{inter}{union} \quad (7)$$

$$L_{Inner\_IoU} = 1 - IoU^{inner} \quad (8)$$

式中:  $b^{gt}$  代表目标框,  $(x_c^{gt}, y_c^{gt})$  表示目标框的中心点,  $w^{gt}$  和  $h^{gt}$  分别为目标框的宽和高,  $b, x_c, y_c, w$  和  $h$  分别表示预测框以及该框的中心点横纵坐标的值、宽、高等属性,  $inter$  代表预测框和目标框的交集面积。

MPDIoU 是一种基于水平矩形的最小距离进行计算、能够考虑重叠区域、中心点距离以及宽高的偏差的方法,包含现有损失函数考虑的所有相关因素,通过直接计算预测框与目标框关键点距离的方式,提供了更精确的度量方法。其计算公式如式(9)~(11)所示。

$$d_1^2 = (x_1^B - x_1^A)^2 + (y_1^B - y_1^A)^2 \quad (9)$$

$$d_2^2 = (x_2^B - x_2^A)^2 + (y_2^B - y_2^A)^2 \quad (10)$$

$$MPDIoU = \frac{A \cap B}{A \cup B} - \frac{d_1^2}{W^2 + H^2} - \frac{d_2^2}{W^2 + H^2} \quad (11)$$

其损失定义如式(12)所示。

$$L_{MPDIoU} = 1 - MPDIoU \quad (12)$$

式中:  $A, B$  为任意的两种凸形,  $(x_1^A, y_1^A), (x_2^A, y_2^A)$  分别是

$A$  左上角和右下角的坐标,同理可得  $B$  左上角及右下角的坐标。 $W, H$  表示输入图像的宽高。

将 Inner-IoU 损失应用于 MPDIoU 边界框回归损失函数中,其损失定义如式(13)所示。

$$L_{Inner-MPDIoU} = L_{MPDIoU} + IoU - IoU^{inner} \quad (13)$$

Inner\_MPDIoU 将 Inner\_IoU 的损失融入 MPDIoU 边界框回归损失函数,通过综合二者计算结果,以更精准地反映目标框和预测框的差距。在模型训练时,它能为模型参数更新提供更精确的梯度信息,帮助模型更快收敛到更优解。由于兼顾了不同尺寸和形状目标的特征,能提升模型对不规则物体和尺寸变化较大物体的检测精度。

而 Soft-NMS 通过合理抑制重叠框竞争,让模型专注学习准确框特征,避免被大量重叠框干扰,引导损失函数在训练过程中进一步优化预测框,使训练更有效,减少模型因预测框重叠造成的误差。它避免了传统 NMS 直接删除重叠框可能导致的真实目标被误删问题,提高了检测的精度。并且通过置信度的调整,对这些框的重要性进行了合理排序和保留,使模型的预测结果更加准确和可靠。利用 Soft\_NMS 改进的 Inner\_MPDIoU 提供了一种新的视角来优化 YOLOv11 的边界框回归问题,相比于其他 IoU 能克服其他 IoU 泛化能力弱、收敛慢等问题,并更精准的反映目标框和预测框的差距。

### 1.5 改进的 YOLOv11n 算法模型

改进之后的 YOLOv11 在特征提取层添加了 iEMA 注意力机制。由于过多的 C3k2\_MSCB 模块会过度强调特征融合与上下文信息捕捉,导致细节信息丢失,不利于小目标检测,因此使用 C3k2\_MSCB 替换部分原有的 C3k2,最后用 MR-Detect 替换原有检测头,其结构如图 6 所示。

## 2 实验结果分析

### 2.1 实验环境及参数设置

实验使用 Ubuntu20.04 操作系统,采用 PyTorch2.0.0 深度学习框架,CPU 为 AMD EPYC 9754 128-Core Processor,GPU 型号为 NVIDIA RTX4090D,显卡内存 24 GB,编程语言为 Python3.8。基线模型选取 YOLOv11n,采用 SGD 优化器,批量大小为 32,初始学习率 0.001,迭代 200 次。

### 2.2 数据介绍

由于电力行业的特殊性以及国家电网的保密性,导致数据样本获取难度较大,现阶段变压器外部缺陷公开数据集相对较少,因此本文通过网上收集等方式,获取 2 798 张图片,按照变压器常见缺陷分为 7 类。由于部分数据集存在多种缺陷的情况,所以实际缺陷样本数量为 5 050 个。各缺陷种类及对应的样本数量和标签名称如表 1 所示。

为了更直观展示出以上缺陷的情况,图 7(a)~(b)展示出各类缺陷经典图片。该数据集涵盖了光照变化、视角差异、多缺陷并存等复杂工况检测场景。

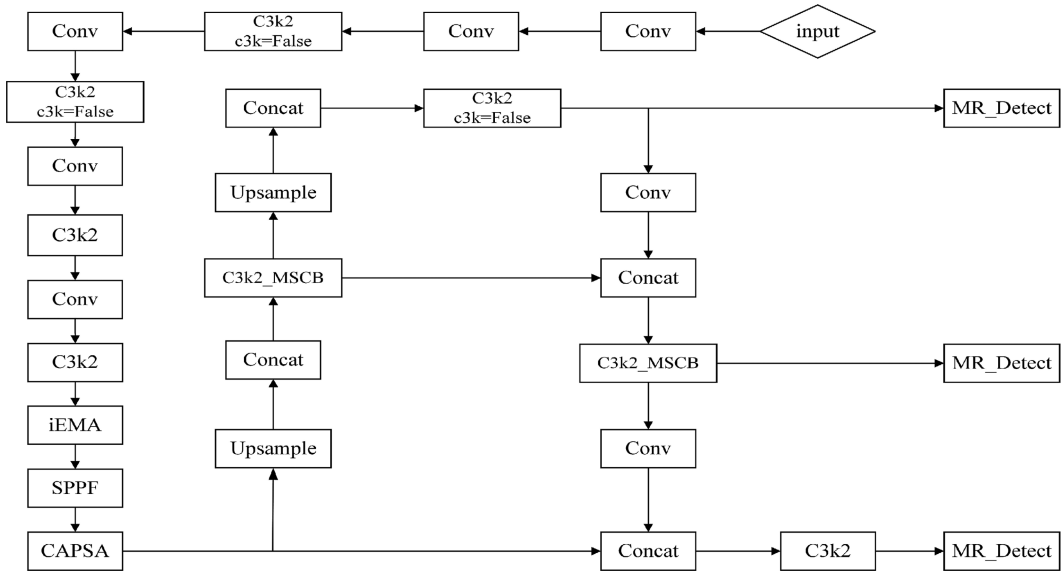


图6 改进后的YOLOv11  
Fig. 6 Improved YOLOv11

表1 数据集分类  
Table 1 Dataset classification

类别	样本数量	标签名称
渗漏油污	671	slyw
硅胶变色	743	gjbs
硅胶破损	590	gjps
油位异常	564	ywyc
油变色	669	ybs
套管破损	832	jyz_pl
金属锈蚀	981	jsxs

2.3 评价指标

为了能够更好的体现改进后的模型检测效果,本文按照7:2:1的比例对现有数据集划分为训练集、验证集、测试集,采用准确率(P)、召回率(R)、mAP@0.5计算公式如式(14)~(17)所示,并用每秒十亿次浮点运算(giga floating point operations per second,GFLOPs)反映训练过程中所需的计算量,每秒处理帧数(FPS)来评价模型的检测效率。

$$P = \frac{TP}{TP + FP} \tag{14}$$

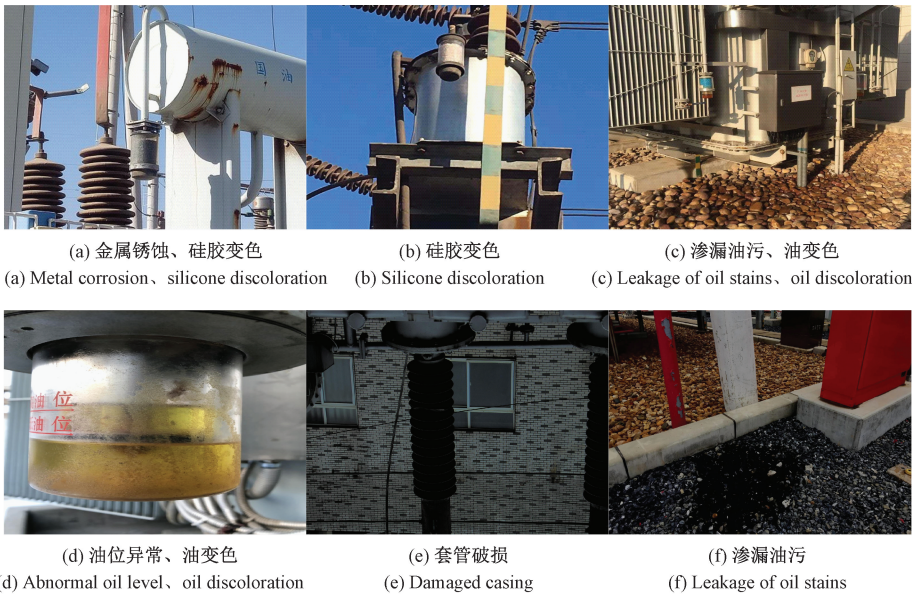


图7 各类缺陷展示

Fig. 7 Display of various defects

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$AP = \int_0^1 Precision(R) dR \quad (16)$$

$$mAP@0.5 = \frac{\sum_{i=1}^N AP_i}{N} \quad (17)$$

式中:  $N$  代表类别数量;  $TP$  表示模型正确预测的正类样本;

$FP$ 、 $FN$  分别表示模型错误地预测为正例的样本个数和模型错误地预测为反例的样本数量。因此召回率越高,意味着模型对正样本的区分能力越强,也不容易出现漏检的情况。

### 2.4 实验结果分析

#### 1) 消融实验

为了验证本文改进模型有效性,在预设输入图像为  $640 \times 640$  的大小下进行消融实验,结果如表 2 所示。

表 2 消融实验结果

Table 2 Results of ablation experiment

实验	iEMA	MSCB	MR-Detect	损失函数	mAP@0.5/%	R/%	P/%	GFLOPs	FPS
Base					76.3	72.4	83.9	<b>6.4</b>	<b>84.3</b>
1	✓				77.5	72.8	88.1	6.5	76.8
2		✓			77.3	72.3	85.6	<b>6.4</b>	77.2
3			✓		76.5	73.3	86.7	8.2	68.7
4				✓	79.2	73.2	85.7	<b>6.4</b>	82.9
5	✓	✓			78.4	73.4	83.8	<b>6.4</b>	70.6
6	✓	✓	✓		79.4	74.5	85.4	8.2	61.0
7	✓	✓	✓	✓	<b>82.4</b>	<b>75.2</b>	<b>86.9</b>	8.2	59.9

由表 2 可知,在添加 iEMA 后 mAP@0.5 相比基线模型高出 1.2%,表明此结构有一定的优化特征提取能力;在此基础上(实验 1),继续使用 MSCB 模块(实验 5),计算量降低 0.1,而 mAP@0.5 和召回率分别提升 0.9% 和 0.6%,这表明该模块在保证轻量的同时还能提高检测的准确率。之后加入 MR-Detect 检测头,由于其采用了一定数量的并行卷积残差结构导致特征处理路径增多,计算量有所上升,但召回率也有一定提高,意味着漏检的目标更少,有效减少漏检的现象。而在使用 Soft\_NMS 改进的 Inner\_MPDIoU 后,检测模型能够在训练中有效抑制重叠框的影响,进一步提升精度。

综上所述,本文所设计的 4 个改进方案均有一定的效果,其构成的 iMMI-YOLO 模型相较基线模型实现了精度和召回率的双重优化。

#### 2) 损失函数 ratio 对比实验

在损失函数中, ratio 的取值至关重要,其直接影响着模型的性能表现。为研究 ratio 对模型的影响,选取最合适的取值,本文在表 2 实验 7 的基础上,并保持环境及其他参数不变的情况下,对 ratio 分别设置不同的值进行对比实验,结果如表 3 所示。

由表 3 可知,当 ratio > 1 和 < 1 时,模型性能呈现出类似抛物线的变化状态。当 ratio > 1 时,随着其值从 1.2 逐渐减小至 1.05,模型性能各项指标整体呈现先上升后下降的趋势。以 mAP@0.5 为例,当 ratio = 1.1 时, mAP@0.5 为 81.5%, 而 ratio 取 1.2 和 1.05 时, mAP@0.5 分别为 80.4% 和 80.0%, P、R 和 mAP@0.5:0.95% 也呈现出相似的变化趋势。当 ratio < 1 时,模型性能同样呈现这样的

表 3 ratio 对比实验

Table 3 Ratio comparison experiment

ratio	P/%	mAP@0.5/%	R/%	mAP@0.5:0.95/%
1.2	84.8	80.4	74.0	52.9
1.1	84.9	81.5	<b>75.4</b>	<b>54.8</b>
1.05	84.6	80.0	74.2	52.8
0.95	<b>87.5</b>	80.8	73.6	53.9
0.9	86.9	<b>82.4</b>	75.2	<b>54.8</b>
0.8	82.9	80.7	75.2	53.3
0.7	85.1	80.6	74.4	54.5

特征。在这一区间, 0.9 左右的 ratio 值使得模型在检测精度和召回率等方面达到相对较优的平衡状态, 整体性能较好, 而偏离该值时, 模型性能会有所下降。

#### 3) 损失函数对比实验

为了验证 Inner\_MPDIoU 更适合检测此类缺陷尺寸差异较大的检测任务, 本文将其他 Inner 系列损失函数的 ratio 值固定为 0.9, 并采用 Soft\_NMS 算法进行后处理。实验结果如表 4 所示。

由表 4 可以看出, 倘若只使用单一类型的边框损失函数 DIoU、SIoU、MPDIoU 等, 其得到的结果交叉, 但是 Inner\_IoU 的出现以及其与各类损失函数相结合的结果改变了这一现状。将两个不同的损失函数的思想结合起来, 可以提高模型的性能。这是因为每个损失函数都强调不同的方面, 结合它们可以弥补单一损失函数的缺陷, 使模型更全面地学习目标检测任务。Inner\_MPDIoU 这种新的

表4 损失函数对比实验

Table 4 Loss function comparison experiment

损失函数	P/	mAP@0.5/	R/	mAP@0.5;0.95/
	%	%	%	%
DIoU	84.9	79.2	74.3	53.6
SIoU	84.7	79.7	75.0	53.4
MPDIoU	85.3	80.0	75.1	53.8
Inner_IoU	86.4	81.1	73.6	54.0
Inner_DIoU	85.2	80.3	74.7	54.2
Inner_SIoU	86.2	80.8	<b>75.5</b>	54.0
Shape_IoU	85.9	79.8	73.7	53.5
Inner_MPDIoU	<b>86.9</b>	<b>82.4</b>	75.2	<b>54.8</b>

组合提供了更全面的目标函数,更好地反映了任务的复杂性。在这种情况下,使得模型更好地适应了任务的需求。

## 4)不同检测模型对比实验

为了进一步证明本文改进方法对变压器外观缺陷检测的效果,现将一些目标检测主流模型与电力行业相关的检测模型在相同环境配置下进行对比实验。结果如表5所示。

由表5可知,SSD作为早期的单阶段目标检测算法,在精度和召回率方面均落后于YOLO系列的小型模型。Faster-RCNN作为two-stage目标检测算法的代表,具有独特的Anchor机制,召回率相对来说较高,但准确度低,总体的mAP@0.5不高,且训练时间长,并具有参数规模

表5 对比实验结果

Table 5 Comparison of experimental results

模型	mAP@0.5/%	Recall/%	Precision/%	GFLOPs	FPS
YOLOv3-tiny	75.9	70.2	81.8	13.0	<b>133.4</b>
YOLOv5n	72.8	70.2	78.3	<b>4.3</b>	128.3
SSD	68.2	60.2	85.1	82.3	21.1
Faster-RCNN	72.5	<b>78.0</b>	48.6	138.1	22.0
YOLOv8n	76.0	72.4	78.9	8.1	118.3
YOLOv10n	73.5	66.4	82.7	8.4	87.2
UDD-YOLO <sup>[24]</sup>	73.2	70.1	79.8	5.4	67.4
文献[25]	77.0	72.7	84.5	16.3	56.2
文献[26]	76.2	73.0	83.7	5.1	85.1
YOLOv11n	76.3	72.4	83.9	6.4	84.3
本文方法	<b>82.4</b>	75.2	<b>86.9</b>	8.2	59.9

大的缺点。相比其他YOLO系列的模型,本文方法利用残差结构在信息保留方面的优势以及设计新的边界框回归方法,形成更有效的特征提取网络,具有较高的检测精度及召回率,且相比基线模型,mAP@0.5和R分别高出了5.9%和2.8%。UDD-YOLO与文献[26]所提的模型都对电力设备缺陷检测优化,其计算量低于其他常见模型,

且检测效率较高,为边缘部署提供一个较优的选择,但其在复杂的变压器检测场景中表现不佳,检测精度等指标低于基线模型。相比此类模型,本文方法检测速度有所降低,但就变压器缺陷检测任务而言,其更注重提高精度减少漏检,在此基础上满足实时检测(FPS $\geq$ 30)即可。改进前后各类缺陷精度如表6所示。

表6 本文方法与基线模型各缺陷数据对比

Table 6 Comparison of defect data between the method and baseline model in this paper

类别	基线模型 mAP@0.5/%	基线模型 Recall/%	改进模型 mAP@0.5/%	改进模型 Recall/%
渗漏油污	66.8	61.8	71.9	59.8
硅胶变色	95.4	92.5	95.9	92.7
硅胶破损	98.4	96.2	97.5	96.2
油位异常	87.4	84.0	94.8	89.3
油变色	91.4	89.9	94.9	91.9
套管破损	54.7	48.9	67.0	53.5
表面锈蚀	42.3	33.8	54.9	43.2
总计	76.3	72.4	82.4	75.2



从各类缺陷来看,改进模型表现较为突出,尤其是在套管破损和表面锈蚀上,其精度分别提高 12.3%和 12.6%,召回率也有相应提升。虽然在硅胶破损这一缺陷中,出现了精度降低的情况,但其降低的值也在可接受范围之内,且从总体来看,改进模型的 mAP@0.5 从基线模型的 76.3% 提升至 82.4%,召回率从 72.4% 提升到 75.2%。而这表明改进后的

模型在综合检测精度和对缺陷的识别能力上有了全面提升,能够更有效地检测出变压器的各类缺陷。

### 2.5 缺陷结果可视化

为了更加直观展示检测模型改进后的结果,本文将置信度设置为 0.25,选择测试集中部分图片进行不同模型的检测效果展示,具体效果如图 8 所示。

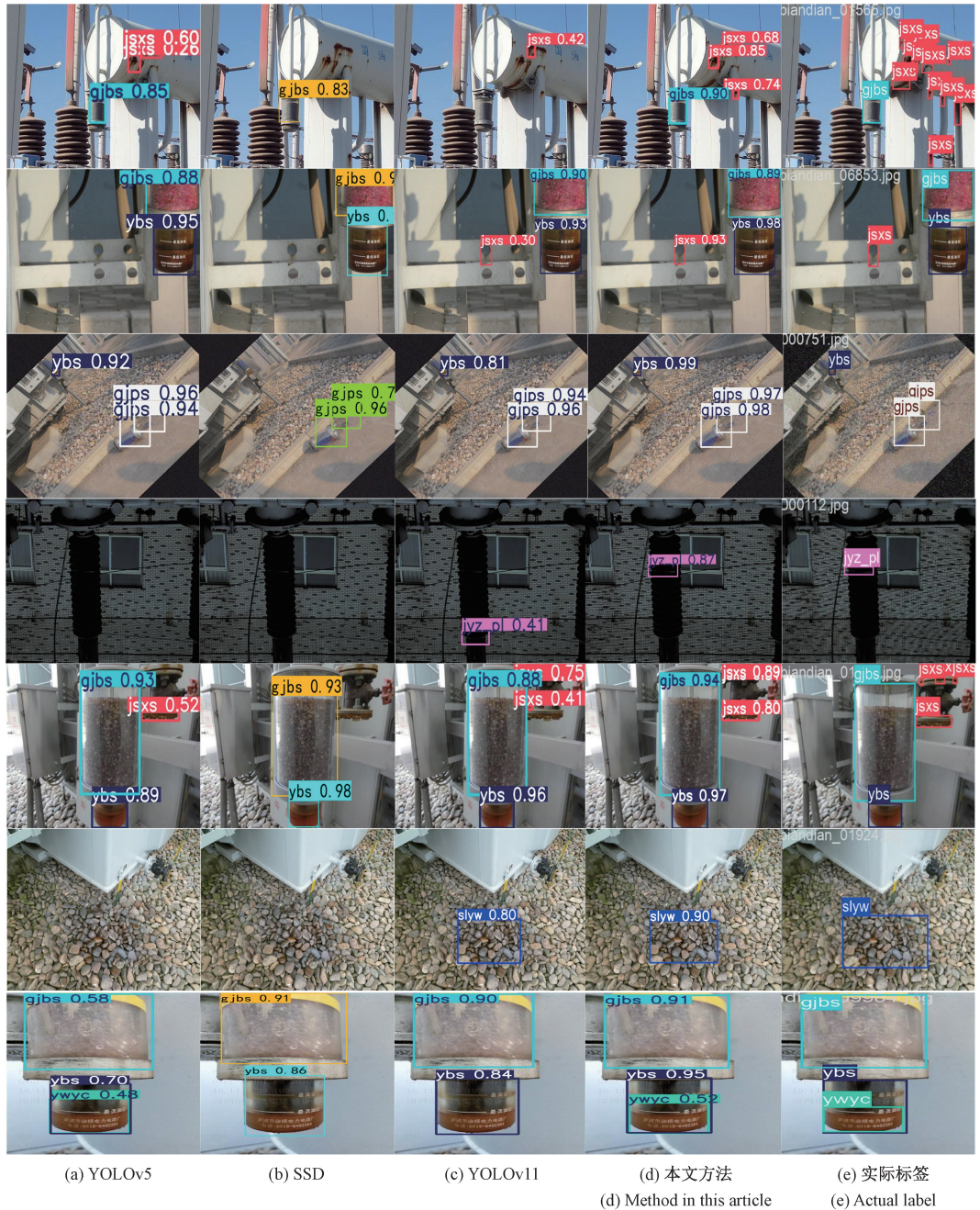


图 8 检测结果可视化

Fig. 8 Visualization of test results

图 8 展示了不同模型在不同缺陷中的检测结果。在特征不明显和光照强度较低的情况下,如渗漏油污、套管破损等,相比于 SSD 和 YOLOv11,改进后的模型对目标的

边界定位更加准确,并给出较高的置信度作为判断依据。在多目标检测场景中,本文模型也展现出较强的检测能力,能够同时检测出多个不同类型的缺陷,例如在第二组

和第七组图片中,其他模型更注重于硅胶变色这类较为明显的缺陷,遗漏了细小的金属锈蚀和轻微的油位异常情况。而改进后的模型不仅精准识别出较大缺陷的区域,还捕捉到其他模型忽视的细节,对其边界框进行标注。虽然改进模型在金属锈蚀这一方面效果不佳,仍然漏检较多目标,但就其他模型而言,本文模型具有更好的检测性能,其实验结果可作为对复杂场景下模型泛化能力的基础参考。

### 3 结 论

由于变压器的缺陷种类繁多、尺寸差异较大,且往往伴随复杂的背景和光照条件,导致传统模型在缺陷检测中的表现较差。针对这些难点,本文提出一种 iMMI-YOLO 的变压器缺陷检测方法。该方法基于残差结构在特征融合与信息保留方面的优势,设计出 iEMA、MSCB、MR-Detect 以此来改进 YOLOv11 模型,旨在解决 YOLOv11 特征提取能力不足以及对上下文信息利用不够充分的问题。此外通过对 Inner\_MPDIoU 的改进以达到解决传统损失函数对尺寸变化较大的物体不敏感的局限性。实验结果表明,本文提出的 iMMI-YOLO 在变压器缺陷检测任务中 mAP@0.5 为 82.4%, Recall 为 75.2%, FPS 为 59.9,相比其他模型,在保证实时检测的同时,具有更高的检测精度,适用于复杂的变压器工况检测场景。

然而改进模型在金属锈蚀等方面仍存在提升空间。模型在处理锈蚀与其他部件相互遮挡产生的复杂特征时,现有的特征提取和融合机制仍存在局限性,无法准确分离和识别锈蚀区域。在后续工作中应深入优化模型结构,例如改进 MR-Detect 检测头中局部与全局特征融合的方式,使其更适应复杂场景下的特征提取。此外,引入更先进的注意力机制,增强模型对金属锈蚀这类小目标和特征不明显目标的关注度,有望进一步提高检测精度。最后在未来的工作中应当将该模型投入到实际的应用中,在实践中不断优化模型,以在真实场景中达到更好的泛化能力。

### 参考文献

[1] 栗磊,王廷涛,赫嘉楠,等.考虑过采样器与分类器参数优化的变压器故障诊断策略[J].电力自动化设备,2023,43(1):209-217.  
LI L, WANG Y T, HE J N, et al. Transformer fault diagnosis strategy considering parameter optimization of oversampler and classifier [J]. Electric Power Automation Equipment, 2023, 43(1): 209-217.

[2] 钟志明.基于改进 YOLOv7 的变电站套管红外图像智能诊断模型构建及应用[J].国外电子测量技术,2025,44(1):140-147.  
ZHONG ZH M. Construction and application of intelligent diagnosis model of infrared image of substation casing based on improved YOLOv7 [J]. Foreign Electronic Measurement Technology, 2025,

44(1): 140-147.

[3] 杨旻宸,张彦,蔡磊,等.基于变压器油荧光特性的渗漏油检测方法[J].电世界,2018,59(3):32-34.  
YANG M CH, ZHANG Y, CAI L, et al. Leakage oil detection method based on the fluorescence characteristics of transformer oil [J]. Electric World, 2018, 59(3): 32-34.

[4] 陆志欣,田涵宁,郭国伟.基于改进 YOLOv4 的电网变压器油液渗漏检测方法[J].计算机测量与控制,2024,32(2):85-92.  
LU ZH X, TIAN H N, GUO G W. Oil leakage detection method for power grid transformers based on improved YOLOv4 [J]. Computer Measurement & Control, 2024, 32(2): 85-92.

[5] 王铨浩,骆忠强,漆梓渊.基于改进 YOLOv5 的变压器漏油检测[J].国外电子测量技术,2023,42(9):169-176.  
WANG CH H, LUO ZH Q, QI Z Y. Oil leakage detection of YOLOv5 transformer [J]. Foreign Electronic Measurement Technology, 2023, 42(9): 169-176.

[6] 田喆文.基于多特征融合和深度稳态学习的电力设备缺陷检测研究[D].西安:长安大学,2023.  
TIAN ZH W. Research on defect detection of power equipment based on multi feature fusion and deep steady state learning [D]. Xi'an: Chang'an University, 2023.

[7] 全欣宇,彭诗媛,钱宏杰.基于红外图像处理的箱式变压器组件检测研究[J].科学技术创新,2024(14):51-54.  
QUAN X Y, PENG SH Y, QIAN H J. Research on detection of box transformer components based on infrared image processing [J]. Science and Technology Innovation, 2024(14): 51-54.

[8] 李坤,杜科,李舟演,等.一种改进 YOLOv8n 的电力设备红外图像识别网络[J].无线电工程,2024,54(10):2362-2370.  
LI SH, DU K, LI ZH Y, et al. An improved YOLOv8n infrared image recognition network for power equipment [J]. Radio Engineering, 2024, 54(10): 2362-2370.

[9] 位一鸣,童力,罗麟,等.基于卷积神经网络的主变压器外观缺陷检测方法[J].浙江电力,2019,38(4):61-68.  
WEI Y M, TONG L, LUO L, et al. A method for detecting appearance defects of main transformers based on convolutional neural networks [J]. Zhejiang Electric Power, 2019, 38(4): 61-68.

[10] 刘冬.基于深度学习的几种变压器缺陷检测研究[D].



- 保定:华北电力大学, 2022.
- LIU D. Research on several transformer Defect detection methods based on deep learning [D]. Baoding: North China Electric Power University, 2022.
- [11] GHORBANI Z, BAHZADAN A H. Monitoring offshore oil pollution using multi-class convolutional neural networks[J]. Environmental Pollution, 2021, 289: 117884.
- [12] CHEN L, XIONG W, YANG S, et al. Research on recognition technology of transformer oil leakage based on improved YOLOv3 [C]. 2020 International Conference on Computer Information and Big Data Applications(CIBDA). IEEE, 2020: 454-458.
- [13] PENG D G, PAN J ZH, WANG D H, et al. Research on oil leakage detection in power plant oil depot pipeline based on improved YOLOv5[C]. 2022 7th International Conference on Power and Renewable Energy(ICPRE). IEEE, 2022: 683-688.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [15] CHENG Y, LIU D M. AdIn-DETR: Adapting detection transformer for end-to-end real-time power line insulator defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73:1-11.
- [16] LI B, WANG T, HU ZH D, et al. Two-level model for detecting substation defects from infrared images [J]. Sensors, 2022, 22(18): 6861.
- [17] CHEN X J, AN ZH Y, HUANG L SH, 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(ICPES). IEEE, 2020: 256-261.
- [18] XU H F, XU J N, LIU L Q, et al. Family defects detection method of power equipment based on datafusion[C]. 2023 3rd International Conference on New Energy and Power Engineering (ICNEPE). IEEE, 2023: 332-336.
- [19] 包广斌, 范超林, 罗瞳, 等. 基于改进 YOLOv11 的太阳能电池板缺陷检测 [J]. 电子测量技术, 2025, 48(17):16-25.
- BAO G B, FAN CH L, LUO T, et al. Solar panel defect detection based on improved YOLOv11 [J]. Electronic Measurement Technology, 2025, 48 (17): 16-25.
- [20] 殷旭鹏, 赵兴强. YOLOv11-MAS: 一种高效的 PCB 缺陷检测算法[J]. 计算机工程与应用, 2025, 61(17): 102-111.
- YIN X P, ZHAO X Q. Improved YOLOv11-based algorithm for PCB defect detection [J]. Computer Engineering and Applications, 2025, 61(17):102-111.
- [21] ZHANG J N, LI X T, LI J, et al. Rethinking mobile block for efficient attention-based models [C]. 2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2023: 1389-1400.
- [22] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. 2023 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP), 2023: 1-5.
- [23] RAHMAN M M, MUNIR M, MARCULESCU R. Emcad: Efficient multi-scale convolutional attention decoding for medical image segmentation[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 11769-11779.
- [24] 杨阳, 杨帅, 闫敏, 等. 基于 UDD-YOLO 的边缘端绝缘子放电严重程度评估算法[J]. 电子测量与仪器学报, 2024, 38(1): 219-227.
- YANG Y, YANG SH, YAN M, et al. UDD-YOLO based edge-end insulator discharge severity assessment algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(1): 219-227.
- [25] 吕越, 薛念明, 邓昊, 等. 基于金字塔分割注意力机制的输电线路设备检测方法研究[J]. 电力大数据, 2025, 28(3): 47-57.
- LYU Y, XUE N M, DENG H, et al. Research on the target detection method of transmission line equipment based on the pyramid split attention Mechanism[J]. Electric Power Big Data, 2025, 28(3): 47-57.
- [26] 周运磊, 董效杰, 刘三军, 等. 基于改进 YOLOv11n 的轻量级电力设备过热故障红外图像检测算法[J]. 湖北民族大学学报(自然科学版), 2025, 43(1): 114-118, 140.
- ZHOU Y L, DONG X J, LIU S J, et al. The lightweight power equipment overheating fault infrared images detection algorithm based on the improved YOLOv11n [J]. Journal of Hubei Minzu University (Natural Science Edition), 2025, 43(1): 114-118, 140.

## 作者简介

吴宇浩, 硕士研究生, 主要研究方向为深度学习。

E-mail: 1291343034@qq.com

朱文忠(通信作者), 教授, 主要研究方向为物联网技术, 工业缺陷检测和深度学习。

E-mail: zwz@suse.edu.cn