

基于改进 RT-DETR 的玻璃绝缘子缺陷检测算法^{*}

张靖雯 孙 坚 徐红伟 许素安 付紫平

(中国计量大学机电工程学院 杭州 310018)

摘 要: 玻璃绝缘子作为输电线路中的关键组件,其缺陷直接影响电力系统的可靠性。针对其缺陷对比度低、多尺度等问题,本文提出了一种基于改进 RT-DETR 的轻量化玻璃绝缘子缺陷检测算法。该方法首先引入轻量化骨干网络 RE-FasterNet,通过创新性的部分重复卷积以及高效多尺度注意力机制提升特征提取效率和对小目标、低对比度缺陷的检测能力;其次在特征融合阶段,提出一种部分重复跨阶段特征融合模块,进一步提升网络对多尺度缺陷的检测能力;最后,在小目标检测头中嵌入注意力尺度序列融合框架,显著提升网络对微小缺陷的空间特征提取能力。实验结果表明:所提算法平均准确率相较于基准模型提升了 2.8%,模型大小下降了 23.6%,计算量下降了 13.1%。在玻璃绝缘子自动化缺陷检测领域,具有较强的实用性和推广价值。

关键词: 玻璃绝缘子;RT-DETR;缺陷检测;小目标;多尺度;轻量化

中图分类号: TN911.73; TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Defect detection algorithm for glass insulators based on improved RT-DETR

Zhang Jingwen Sun Jian Xu Hongwei Xu Su'an Fu Ziping

(College of Mechanical and Electrical Engineering, China Jiliang University, Hangzhou 310018, China)

Abstract: Glass insulators are critical components in transmission lines, and their defects can significantly impact the reliability of power systems. This paper proposes an improved defect detection algorithm for lightweight glass insulators based on enhanced RT-DETR, addressing issues related to low contrast and multi-scale defects. First, we introduce the lightweight backbone network RE-FasterNet, which enhances feature extraction efficiency and improves the detection of small targets and low-contrast defects through innovative partial duplication and an efficient multi-scale attention mechanism. Second, during the feature fusion stage, a partially repeated cross-stage feature fusion module is proposed to further enhance the detection capability for multi-scale defects. Finally, an attention scale sequence fusion framework is integrated into the small target detection head, significantly improving the network's spatial feature extraction ability for small defects. Experimental results demonstrate that the proposed algorithm increases the mean average precision by 2.8%, reduces the model size by 23.6%, and decreases computational requirements by 13.1% compared to the benchmark model. In the domain of automatic defect detection for glass insulators, this approach exhibits strong practicality and broad applicability.

Keywords: glass insulator; RT-DETR; defect detection; small target; multi-scale; lightweight

0 引 言

玻璃绝缘子是电力输电线路的重要组成部分,其绝缘性能直接影响着输电线路的安全运行^[1]。其生产过程包括料槽混合、熔制与澄清、压制成型等多个环节,工艺复杂,因此玻璃绝缘子无法避免会产生各种复杂的缺陷,这些缺陷会导致玻璃绝缘子使用性能大大下降。玻璃绝缘子由伞盘、芯体和颈部 3 个部分组成,其中伞盘部位常见的裂纹、

破损和气泡缺陷,最小面积仅占总面积的 1%,这对算法在检测小目标缺陷方面提出了极高要求,同时颈裂缺陷目标相对较大,网络也需适应多尺度的检测需求。传统检测方法以人工目视检测为主,存在效率低、成本高以及准确性差等问题。因此,开发高效、准确、自动化的玻璃绝缘子缺陷检测算法成为当前亟待解决的重要问题。

近年来,深度学习技术在缺陷检测领域展现出了巨大潜力。其中目标检测算法主要分为两阶段算法和单阶段算

法。前者包含特征提取和区域生成两个关键阶段,特征提取阶段利用卷积神经网络来提取输入数据的特征,而区域生成网络则使用回归器和分类器实现对目标的检测与定位,因算法受限于这两个阶段的处理方式,导致其检测速度慢、模型体积大,难以部署在边缘端设备和嵌入式系统上。其主流算法有 Faster R-CNN^[2],Mask R-CNN^[3],Mask R-FCN^[4]等。相比之下,后者通过一个神经网络模型同时完成目标位置和类别的检测,显著提升了检测速度并减少了计算量,因而更适合工业生产中对实时性要求较高的场景。其主流算法有 DETR^[5]、SSD^[6]、YOLO^[7]等。

目前,复杂玻璃类制品的缺陷检测多使用 YOLO 网络实现。尹玲等^[8]提出一种基于 YOLOv4 的缺陷检测方法,使用改进 K-means 聚类算法,采用交并比函数确定锚框的度量,解决原锚框大小不适用于玻璃缺陷小目标检测问题。张涛等^[9]提出结合 YOLOv4 和 MobileNetV3 的改进算法 YOLO-M,优化特征提取网络。Mao 等^[10]提出了一种新的网络 Dy-YOLOv5s,在残差结构中引入注意力模块,加入特征图的跨尺度和跨层连接,使用置信度传播簇和变焦损失函数对冗余锚框和正负样本的平衡进行重复数据删除。

尽管 YOLO 系列模型在目标检测领域取得了显著的成果,但其仍存在一定局限性,特别是在后处理阶段需要依赖非极大值抑制(non-maximum suppression, NMS)。虽然 NMS 能够有效地解决多个检测框之间的重复问题,但这一过程会引入额外的计算复杂度。此外,在某些情况下, NMS 可能会导致小目标的丢失或误分类,从而影响检测性能^[11]。

为此,本文选择了一种真正意义上实现端到端的实时目标检测模型(real-time detection transformer, RT-DETR)^[12]作为基础模型,与其他同样不依赖 NMS 后处理的 DETR 系列模型相比,RT-DETR 通过优化解码器设计、

内存使用和多尺度处理,在不显著牺牲精度的情况下,极大地提高了推理速度,满足了实时检测的需求。此外,RT-DETR 引入的不确定性最小化查询选择策略,进一步提升了模型的准确性,相比其他 DETR 系列模型具有明显优势。为进一步提高 RT-DETR 模型的检测效率,本文针对玻璃绝缘子缺陷的特征进行了适配性优化,具体的优化工作包括以下几个方面:

1)将 FasterNet^[13]引入骨干网络,在牺牲少量模型精度的条件下,提升检测速度。

2)提出一种部分重复卷积(partial repeat convolution, PRepConv)模块,替换 FasterNet 中原有的部分卷积(partial convolution, PConv),并将具有跨空间学习的高效多尺度注意力模块(efficient multi-scale attention, EMA)^[14]嵌入在 FasterNet 中,通过构建巨大的局部感受野来收集多尺度的空间信息,从而提升对于低对比度目标的特征提取能力,挽回模型轻量化所造成的精度损失。

3)在特征融合阶段,提出一种部分重复跨阶段特征融合模块 PRCFusion Block,该模块通过融合不同层级的特征图,能够高效地捕捉到不同尺度的细节特征,从而提高模型对于尺度变化较大缺陷的检测能力。

4)将注意力尺度序列融合框架(attentional scale sequence fusion, ASF)^[15]部署到 P3 预测头中。通过融合上下尺度的特征信息,提升模型对于小目标缺陷的空间特征和通道特征提取能力,增强网络对于小目标缺陷的识别能力。

1 RT-DETR 模型

RT-DETR 是一种基于 Transformer 架构的单阶段算法,其结构如图 1 所示,主要包括 3 个主要模块:骨干网络(backbone),高效混合编码器(efficient hybrid encoder),以及 Transformer 解码器(transformer decoder)。

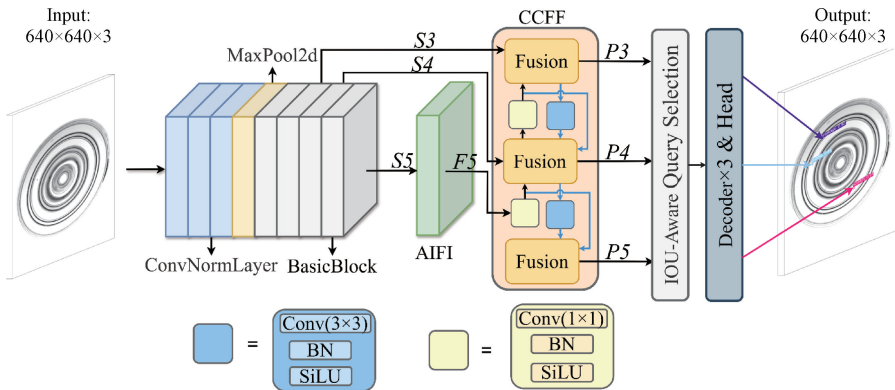


图 1 RT-DETR 模型结构

Fig. 1 RT-DETR model structure

RT-DETR 使用 ResNet 作为骨干网络,提取输入图像的多尺度特征,生成 3 个特征层 $\{S_3, S_4, S_5\}$, 分别对应不同的尺度。

高效混合编码器模块通过结合基于注意力的单尺度特征交互(attention-based intra-scale feature interaction, AIFI)和基于 CNN 的跨尺度特征融合(CNN-based cross-

scale feature fusion, CCFF)来优化多尺度特征处理。其计算公式如下:

$$Q = K = V = \text{Flatten}(S_5) \quad (1)$$

$$F5 = \text{Reshape}(\text{AIFI}(Q, K, V)) \quad (2)$$

$$O = \text{CCFF}(\{S_3, S_4, F5\}) \quad (3)$$

式中: S_3, S_4, S_5 表示多尺度特征, Flatten 和 Reshape 为特征形状变换操作, O 表示多尺度融合后的最终图像特征序列。AIFI 仅对高语义的 S_5 进行交互, 避免了低层次特征的冗余计算, 而 CCFF 则通过多卷积融合模块高效整合多尺度特征。

最后解码器利用辅助预测头和多层 Transformer 结构, 逐步优化从编码器输出的查询特征, 并生成目标的类别和边界框信息。

2 RT-DETR 模型改进

2.1 骨干网络的轻量化设计与优化

为了实现玻璃绝缘子缺陷检测的目标, 本文选择了 ResNet-18 残差网络作为骨干网络, 用于提取输入图像的

特征。其核心思想是残差学习, 通过引入跳跃连接, 将输入直接绕过中间层加到输出上, 解决了深度网络中的梯度消失和退化问题。

但在传统的残差块中, 特征通道之间往往存在显著冗余, 例如多个通道传递了重复或低效的信息。同时在 ResNet-18 中, 所有通道和空间位置的特征被一视同仁, 难以高效利用特征。

为减少冗余, 需采用轻量化操作, 主流轻量化网络包括: GhostNet^[16], EfficientNet^[17], VanillaNet^[18] 等。相比之下, FasterNet 通过减少特征映射冗余, 在提升计算效率的同时, 降低了参数量和内存占用。因此, 本文引入 RE-FasterNet 模块优化原骨干网络中的基本残差块, 改进后的 RN18-Faster-Resnet 和 RE-FasterNet-BasicBlock 的网络结构如图 2(a) 和 (b) 所示。

如图 2(c) 所示, 在优化后的 RE-FasterNet 结构中, 存在一个 3×3 的 PRepConv, 外加两个 1×1 卷积, 构建出倒置的残差块结构。此结构在减少层间冗余的同时, 确保了特征多样性不会出现下降的情况。

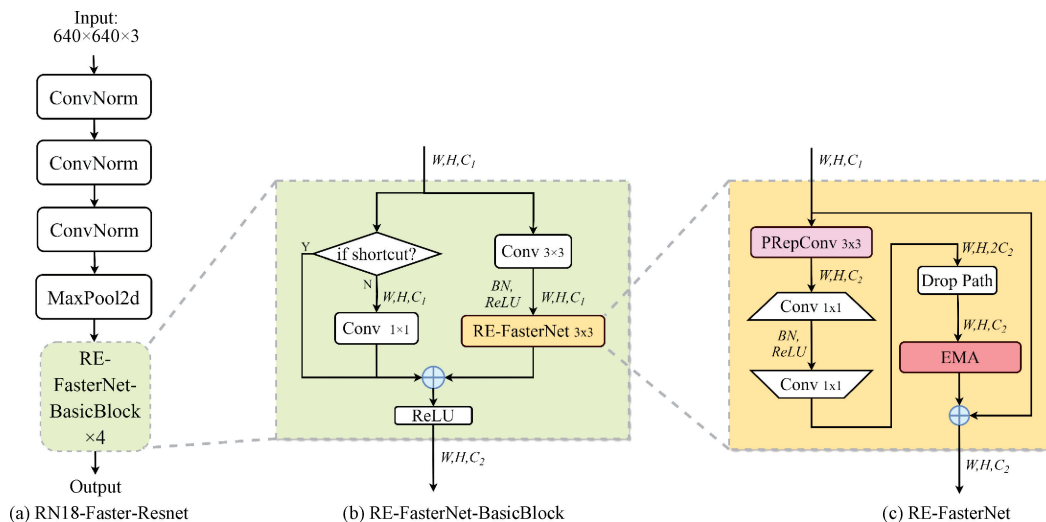


图 2 改进后的骨干网络结构

Fig. 2 Improved backbone network structur

标准卷积(convolution, Conv)对所有输入通道进行卷积处理, 如图 3(a) 所示, 其参数量与计算量大, 效率低。部分卷积相较标准卷积对输入特征进行分离处理, 仅对部分通道执行卷积操作, 而其余通道保持不变, 大大减少了降低计算冗余, 如图 3(b) 所示。然而部分卷积未参与卷积的通道特征直接传递, 可能导致部分细节特征丢失。

因此, 为了在保持 PConv 的优势前提下, 提高其捕获小目标或低对比度区域的局部特征的能力, 本文创新性地提出了一种 PRepConv 模块, 该模块通过引入并行结构, 将原始的一个 3×3 卷积核分解为两个 3×1 和 1×1 的卷积核, 使其在水平方向和垂直方向上能够更细致地提取特征, 能够更好地捕获小目标或低对比度区域的局部特征。

PRepConv 结构如图 3(c) 所示。

此外, 为加强网络对于玻璃绝缘子缺陷特征的学习效率, 本文在 FasterNet 主干网络中引进 EMA 注意力机制, 深入强化网络的特征提取能力。

EMA 是一种轻量化的注意力模块, 其核心思想是通过并行子网络结构, 捕获跨空间和跨通道的多尺度特征信息, 从而提升模型在小目标和低对比度目标检测中的能力。其主要通过两种关键机制实现: 多尺度特征提取与跨空间学习。当输入特征图为 $X \in R^{W \times H \times C}$, 其中 C 为通道数, H 和 W 为空间维度。EMA 首先将 X 沿通道维度划分为 G 个子特征组, 即 $X = \{X_0, X_i, \dots, X_{N-1}\}$, $X_i \in R^{C/G \times H \times W}$ 。划分后的子特征将作为全局特征编码与局部

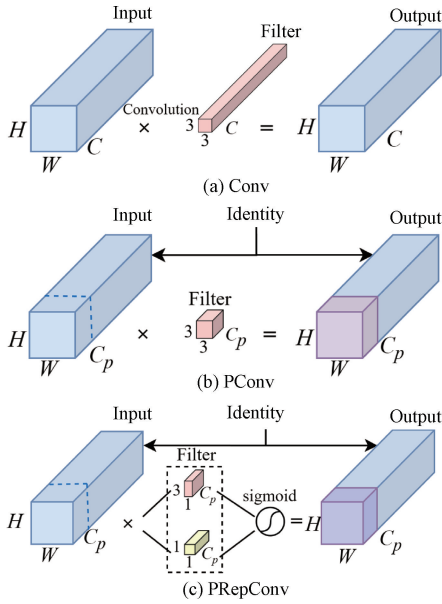


图 3 不同卷积模块结构

Fig. 3 Structures of different convolution modules

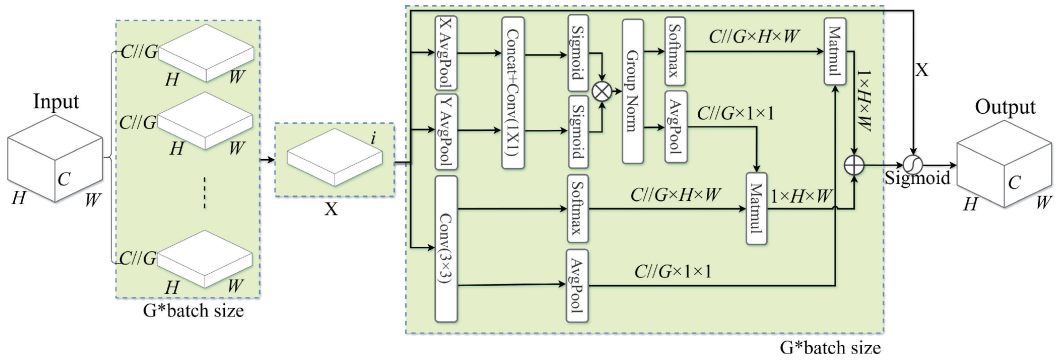


图 4 EMA 注意力机制结构

Fig. 4 Structure of EMA attention mechanism

但在玻璃绝缘子缺陷检测任务中,网络需要面对的是不同尺度的目标,传统的卷积层无法完全适应各种尺度的缺陷,不能进行灵活调整,故而难以准确地捕捉缺陷。在提取不同尺度的缺陷特征时,往往存在着有效信息的丢失。为了解决这一问题,本文将 RepConv 引入高效聚合网(efficient layer aggregation networks, ELAN)^[19],提出了一种 PRCFusion Block,其结构如图 5(b)所示。

相较于原始结构,PRCFusion Block 使用了多分支设计,这种设计更加重视不同层次之间的特征融合,通过层次化的特征提取方式,能够更好地捕获局部和全局的特征,适用于玻璃绝缘子缺陷检测任务中的多尺度目标的识别。同时,PRCFusion Block 的引入优化了计算复杂度,在保证高检测精度的同时,进一步减少了网络的计算量。

2.3 基于注意力尺度序列融合的网络改进

ASF 是一种专为提升多尺度特征提取与融合能力而设计的模块。本文通过将该模块引入 P3 预测头,进一步

特征编码的输入。前者对每个子特征图 X_i 分别沿水平和垂直方向进行全局平均池化得到 z_H 和 z_W ,将二者拼接后通过 1×1 卷积和非线性激活生成权重图 W_1 。后者通过 3×3 卷积直接对 X_i 进行处理,提取局部空间特征,生成权重图 W_2 。最后将两个分支生成的权重图进行融合,使用点积操作实现跨空间交互,并通过 Sigmoid 激活函数生成最终输出特征。EMA 注意力机制结构如图 4 所示。

2.2 跨阶段局部网络的改进

由于加工材料与生产工艺的限制,玻璃绝缘子不可避免地会出现大小不等的破损、裂纹和气泡等缺陷。为了更好地检测这些缺陷,针对不同的缺陷类型和目标尺度,必须设计更加精确的网络结构。在 RT-DETR 模型中,特征融合模块 Fusion Block 结构如图 5(a)所示,其由以下几部分构成:多个卷积分支、分支融合以及重参数化。主要使用标准卷积捕获局部特征,并增强通道间的交互引入恒等分支保持残差特性,各分支的输出特征通过加权相加或逐元素相加的方式融合,生成最终的输出特征。

提升模型在处理玻璃绝缘子小目标和密集目标时的敏感度。

ASF 的结构如图 6 所示,其主要由 3 个核心子模块组成:尺度序列特征融合模块(scale sequence feature fusion, SSFF)、三特征编码模块(triple feature encoder, TFE)以及通道与位置注意力机制(channel and position attention mechanism, CPAM)。ASF 的输入来自骨干网络输出的多尺度特征图,这些特征图包括了来自不同尺度和不同通道数的丰富信息,SSFF 模块通过多尺度的融合,整合了不同尺度特征图的信息,TFE 模块进一步通过特征拆分与调整、特征融合与拼接的策略,结合大、中、小 3 种尺度的特征,有效提升了网络对于小目标和密集目标的细节捕获能力。

在 CPAM 模块中,通道注意力机制使用全局平均池化对每个通道的特征进行降维,生成通道权重,引入 1D 卷积操作,捕获通道之间的依赖关系,对通道权重进行加权,优

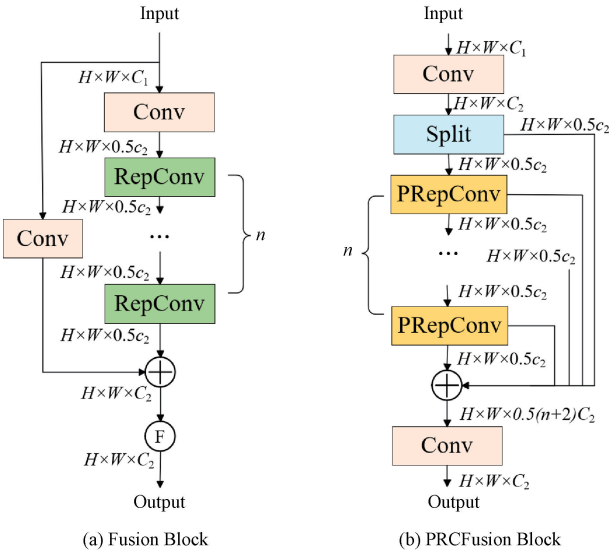


图 5 Fusion Block 和 PRCFusion Block 结构

Fig. 5 Fusion Block and PRCFusion Block structure

化通道特征的重要性;位置注意力机制在宽度和高度方向分别进行池化,提取空间位置信息,将水平和垂直位置的注意力信息合并,并通过卷积操作进一步编码,最终位置注意力权重如式(4)所示。

$$FCPAM = E \times s_w \times s_h \quad (4)$$

式中: E 为通道与位置注意力的权重矩阵, s_w 和 s_h 分别是

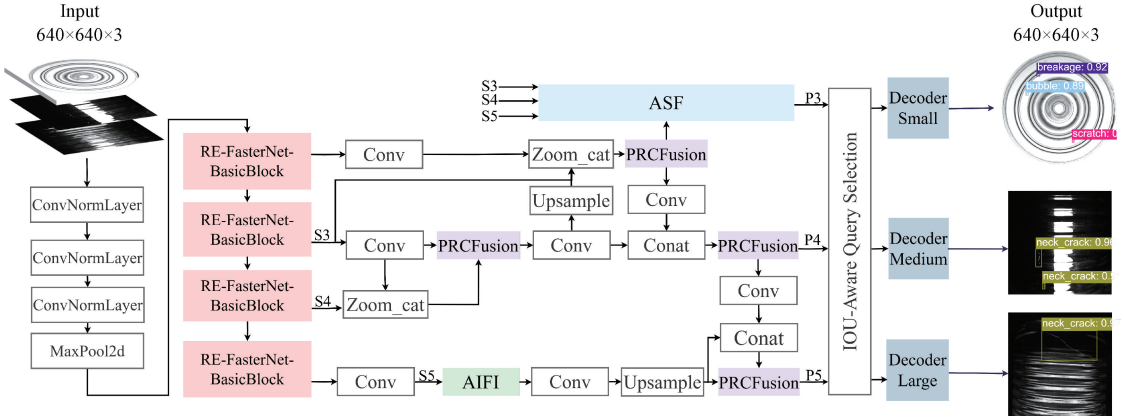


图 7 FAPC-RT-DETR 模型结构

Fig. 7 FAPC-RT-DETR model structure

3 结果与分析

3.1 相机选型与数据集构建

本研究中的实验数据源自自主采集的玻璃绝缘子缺陷数据集。图像采集装置如图 8 所示,所使用的相机为海康 MV-CH250-90GM 2500 万像素工业面阵相机,并配备 MVL-KF5040-25MP 镜头。

本文在实验过程中采集了 3 284 张存在缺陷的玻璃绝缘子图像,涵盖的缺陷类别包括裂纹、破损、气泡和颈裂,

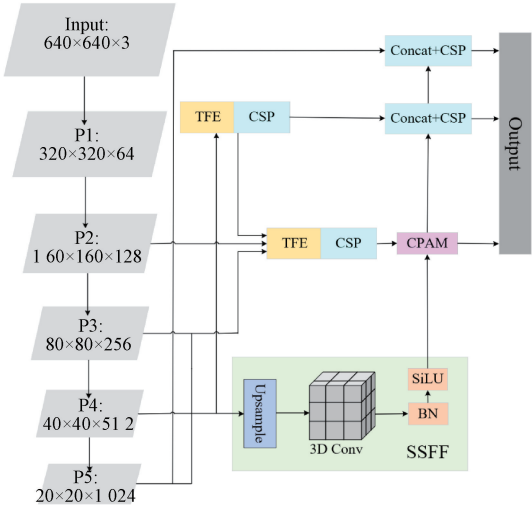


图 6 ASF 结构

Fig. 6 ASF structure

宽度和高度的权重。

2.4 改进模型结构设计与整合

本研究基于 RT-DETR 基础模型,针对玻璃绝缘子缺陷特征进行全面改进。具体包括在骨干网络中采用轻量化 RE-FasterNet 模块与 EMA 注意力机制,在特征融合阶段引入创新性的跨空间融合模块,在小目标检测头中嵌入注意力尺度序列融合框架。其模型结构如图 7 所示。

共计 5 552 例缺陷。图像缺陷的标注使用了 LabelImg 工具。具体而言,裂纹常见于光滑面及其与环形接线的连接处,呈直线或曲线形;破损主要出现在玻璃件的环形接线处,呈弧状;气泡则主要出现在玻璃件的光滑面上,形状为椭圆形;颈裂则出现在玻璃绝缘子的颈部,形态为直线或曲线。

在标注过程中,缺陷的边界框应尽可能准确地覆盖缺陷的所有部分,确保不遗漏缺陷区域,同时避免过度标注,尽量贴合缺陷边缘。各类缺陷及其标注示例如图 9 所示。

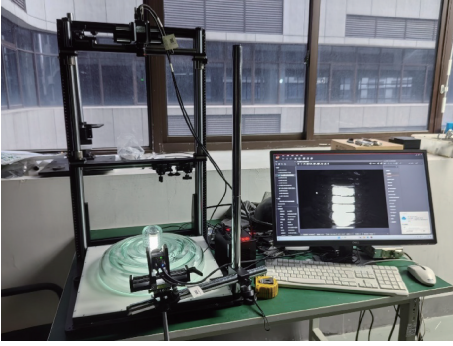


图 8 图像采集装置

Fig. 8 Image acquisition device

最终,所有数据被随机划分为训练集(4 441 例)、验证集(555 例)和测试集(556 例),其比例为 8:1:1。

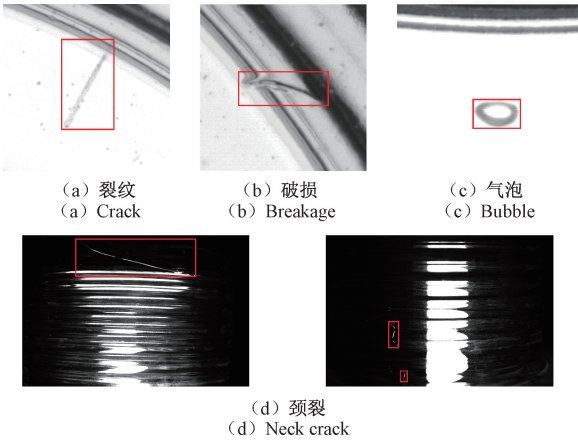


图 9 玻璃绝缘子缺陷类型及标注示例

Fig. 9 Defect types and labeling examples of glass insulators

图 10 为数据集中不同缺陷类别的数量图,原始的图像数据存在类别不平衡问题。为了进一步提升模型的泛化性,通过旋转、镜像、明暗度调整、噪声引入等数据增强方法,以模拟不同光照条件、噪声干扰。将各类缺陷的数量扩充平衡,最终将总缺陷数扩充至 13 222 个。

3.2 实验环境构建与模型评价指标

本文的实验硬件配置为:AMD Ryzen 7 6800H 处理器,GeForce RTX3060 显卡,32 GB 运行内存,1 TB 固态硬盘;系统平台为 64 位 Windows11 专业版,使用 Pycharm 集成开发环境,配置的 CUDA 版本为 12.3,Pytorch 版本为 2.1.2。实验中每个批次处理图片数量为 8,总共进行 200 轮训练;采用动量为 0.937 的 AdamW 优化器,初始学习率为 0.01,权重衰减系数为 5×10^{-4} 。

本文选取平均精度均值(mean average precision, mAP)、每秒 10 亿次的浮点运算数(giga floating-point operations per second, GFLOPs)、模型大小以及每秒传输帧数(frame per second, FPS)评估模型性能,其中 mAP 计算公式为:

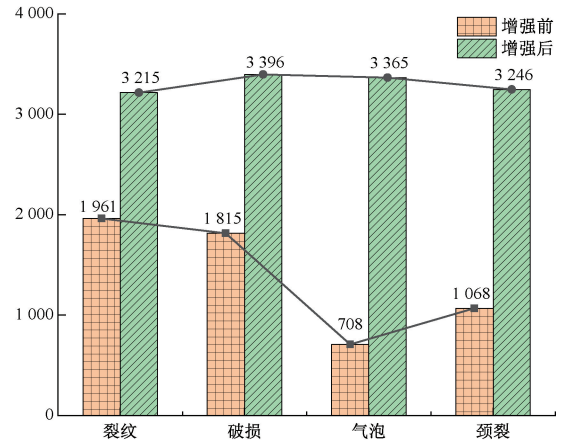


图 10 各类缺陷数量图

Fig. 10 Quantity diagram of various defects

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \frac{\sum P}{TP + FP} \quad (7)$$

$$mAP = \frac{\sum AP}{n} \quad (8)$$

式中: TP 为真正例; FP 为假正例; FN 为假负例; AP 为平均精度; P 为准确率; R 为召回率; n 为缺陷类别数。

3.3 实验结果对比与分析

1) 消融实验

为验证所提创新性算法的优越性,在相同的实验环境下进行消融实验,使用 mAP50、mAP50-95、GFLOPs、模型大小和 FPS 进行模型评估,结果如表 1 所示。

在骨干网络引入 RE-FasterNet 模块后的模型 IV 相较于基准模型 I, mAP 与 mAP50-95 分别提升 0.2%、3.4%,计算量与模型大小分别下降了 8.3%、13.1%,说明 RE-FasterNet 模块通过并行网络构建的巨大局部感受野,能够有效提升模型的空间特征提取能力,同时内部的分离式卷积结构能够进一步降低骨干网络内的层间冗余,实现了模型的轻量化。

在特征融合阶段引入 PRCFusion 模块得到模型 V,对比基础模型 I, mAP 与 mAP50-95 分别提升了 0.9%、5.0%,计算量与模型大小分别缩小了 8.8%、7.1%,这是因为 PRCFusion 模块采用合并不同处理路径特征的方式,增强了网络的特征表达能力和性能。

最终改进后的 RT-DETR 模型,虽然检测速度降低了 6 fps,但 mAP50 与 mAP50-95 分别提升了 2.8%、8.0%,计算量与模型大小分别下降了 13.1%、23.6%。综上,引入各个改进模块使得模型的检测精度、计算量、模型大小与检测速度得到了较好的平衡,模型的综合性能得到

表 1 消融实验结果
Table 1 Results of ablation experiment

模型	RE-FasterNet			PRC-Fusion	ASF	mAP50/%	mAP50-95/%	GFLOPs/G	模型	FPS/ fps
	Faster-Net	PR-Conv	EMA						大小/MB	
I						93.7	62.6	58.9	40.6	54.2
II	✓					92.6	60.7	50.2	34.5	59.3
III	✓	✓				93.3	65.4	51.5	34.9	56.4
IV	✓	✓	✓			93.9	66.0	54.0	35.3	50.4
V				✓		94.6	67.6	53.7	37.7	56.9
VI					✓	94.9	65.9	64.1	41.0	48.4
VII	✓	✓	✓	✓		95.4	68.4	67.8	32.4	49.1
VIII	✓	✓	✓	✓	✓	96.5	70.6	51.2	31.0	48.2

了极大的提升。

为了更全面评估模型的整体性能,结合模型 VIII 的精确率-召回率曲线和 F1 分数曲线,如图 11、图 12 所示,可以看出,模型在所有缺陷类别上的表现较为均衡,尤其是在气泡和颈裂缺陷上的表现突出。但模型对裂纹的识别精度略低,主要是由于该类缺陷特征相较于背景对比度较低,导致模型对该类别的召回率有所牺牲。

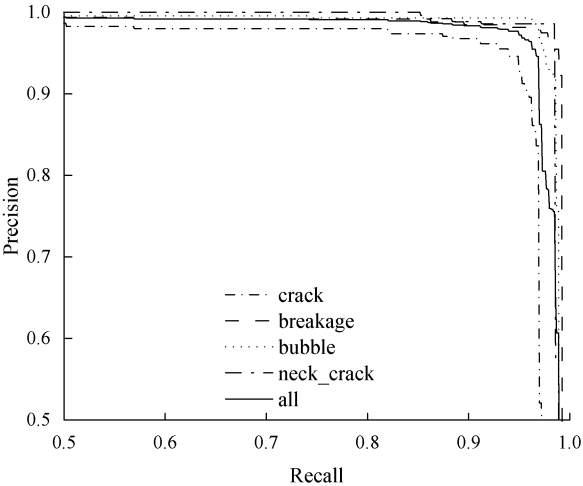


图 11 精确率-召回率曲线
Fig. 11 Precision-Recall curve

图 13 是模型通过结合预测的类别及其边界框反向传播所得 Grad-CAM++^[20] 热力图。通过对比图 13(a) 和 (b),在未添加 EMA 注意力机制时,热点未完全对齐目标区域,存在噪声点;在添加 EMA 注意力机制后,热点主亮点位置精确、靠近中心且与目标区域特征更为契合,这得益于 EMA 注意力机制通过并行子网络构建的巨大局部感受野,使得该模型在处理背景干扰与噪声更具优势。对比图 13(c) 和 (d),在 P3 引入 ASF 框架后,热点更加集中,分布形状更紧凑,表明模型能够有效提升模型对于小目标缺陷的敏感度。

2) 对比实验

为进一步验证本文所提算法的有效性,将 FAPC-RT-

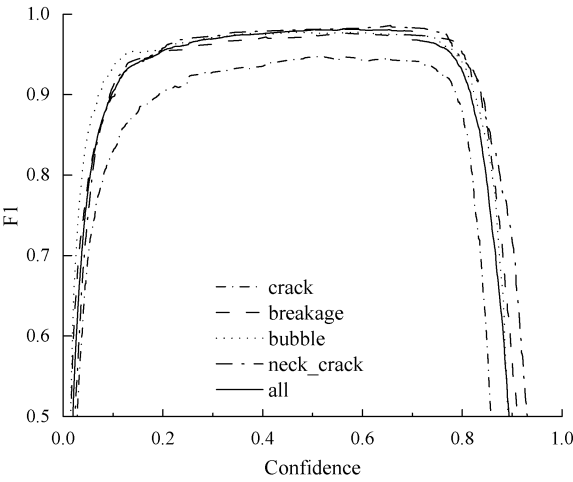


图 12 F1 分数曲线
Fig. 12 F1 score curve

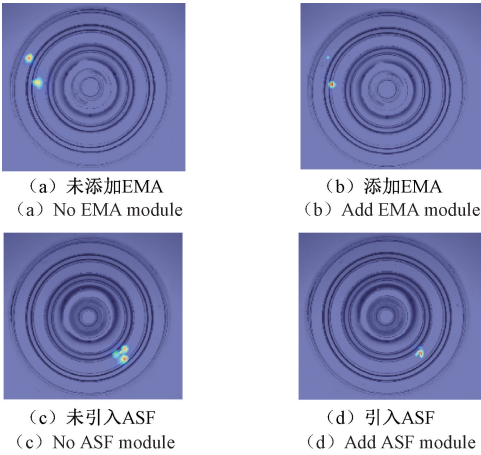


图 13 Grad-CAM++ 热力图
Fig. 13 Grad-CAM++ thermal map

DETR 与 5 种主流的检测算法 Faster R-CNN、DETR、Deformable-DETR、YOLOv5m、YOLOv8m 进行对比,将计算量、参数量、检测精度和检测时间作为性能指标,所得实验结果如表 2 所示。

表 2 不同模型对比实验

Table 2 Comparative experiments of different models

模型	GFLOPs/G	Parameters/M	AP/%				mAP/%	检测时间/ms
			crack	breakage	bubble	nect-crack		
Faster R-CNN	207.0	40.76	87.4	90.5	91.2	90.5	89.9	81.7
DETR	89.4	41.01	89.6	92.4	93.0	92.1	91.8	61.5
Deformable-DETR	78.7	34.68	90.1	92.8	93.6	93.2	92.4	55.1
YOLOv5m	80.2	26.15	91.8	93.2	93.7	93.4	93.1	25.4
YOLOv8m	79.1	25.86	92.6	95.6	96.0	95.4	94.9	18.2
Ours	51.4	17.77	94.2	97.2	97.6	97.0	96.5	20.8

由表 2 可知,本文算法不仅模型的计算量与参数量均为最低,同时在不同缺陷类别下检测精度均优于所列模型,平均检测精度分别提升了 6.6%、4.7%、4.1%、3.4% 和 1.6%,显示出其在内存使用和性能方面的高效性。检测时间较 Faster R-CNN 与 DETR 具有绝对的优势,但与 Yolov8m 算法相比增加了 2.6 ms,这是由于 RT-DETR 本质上属于 Transformer 系列模型,它相较于纯 CNN 网络的 YOLO 系列模型在检测速度上具有先天的劣势,但是完全满足了工业上玻璃绝缘子检测实时性的需求。此外,与同样属于 Transformer 框架的 Deformable DETR 相比,本文算法在计算量、检测精度和检测速度等方面均表现出更优的性能。因此,当需要综合考虑模型准确性、硬件资源限制以及检测速度时,FAPC-RT-DETR 算法是一个极佳的选择。

通过分析不同类型的检测误差,能够更加深入了解模型在检测任务上的表现。因此,本文采用 TIDE 框架^[21]将检测错误分为 6 种类型,分别为:分类错误(Cls)、定位错误(Loc)、分类和定位错误(Both)、重复检测错误(duplicate, Dupe)、背景误检(background, Bkg)以及漏检错误(missed GT, Miss),如图 14 所示。该框架将待求错误类型中所有的错误检测结果均视为真值,并将计算得到修复后的 AP。与真实 AP 作差,从而得到相应的检测错误值 E,各类错误的计算表达式为:

$$E = \Delta AP = AP_o - AP \tag{9}$$

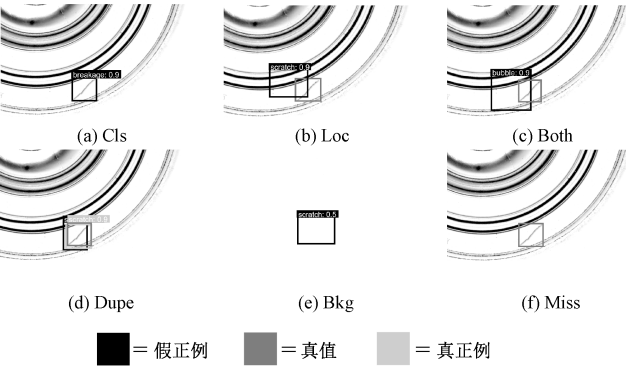


图 14 错误类型定义
Fig. 14 Error type definition

图 15 和 16 显示了不同模型在 coco 评估指标下错误类型比较的权重分布,可以看出本文算法在缺陷分类的准确性与缺陷检测全面性上均优于其他模型,对正负类缺陷样本的区分能力更强,从而能够有效的降低模型的误检与漏检率。

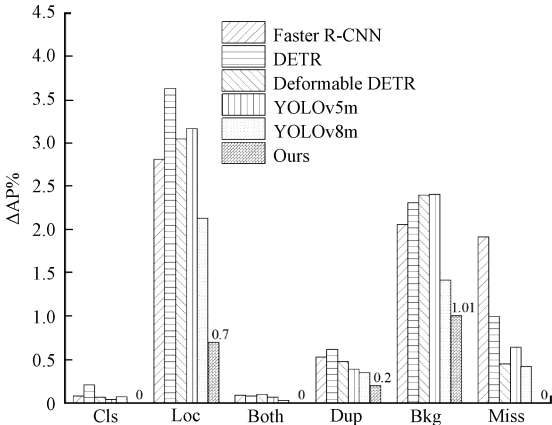


图 15 不同模型检测误差对比图
Fig. 15 Comparison of detection errors of different models

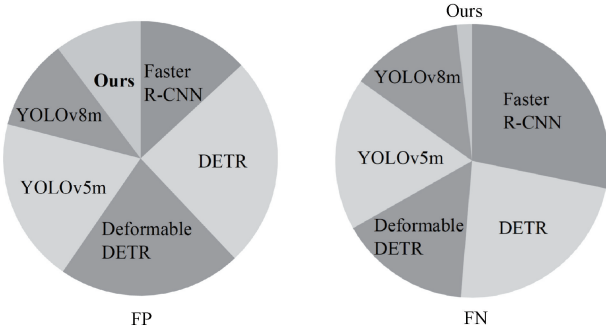


图 16 不同模型误分类对比图
Fig. 16 Comparison of misclassification of different models

为了更直观的体现所提算法检测结果的泛化性,给出了 YOLOv8m 和 FAPC-RT-DETR 算法在不同光照、噪声干扰以及遮挡下的检测结果对比,如图 17 所示。当出现小目标缺陷时,所提算法出现误检与漏检的概率更低,检测准确率更高;当出现大尺寸缺陷时,所提算法对缺陷的

定位精度更高,并且能更准确的贴合缺陷外部轮廓。总体来看,相较于其他模型,FAPC-RT-DETR 在玻璃绝缘子缺

陷检测任务中的性能表现更为突出。

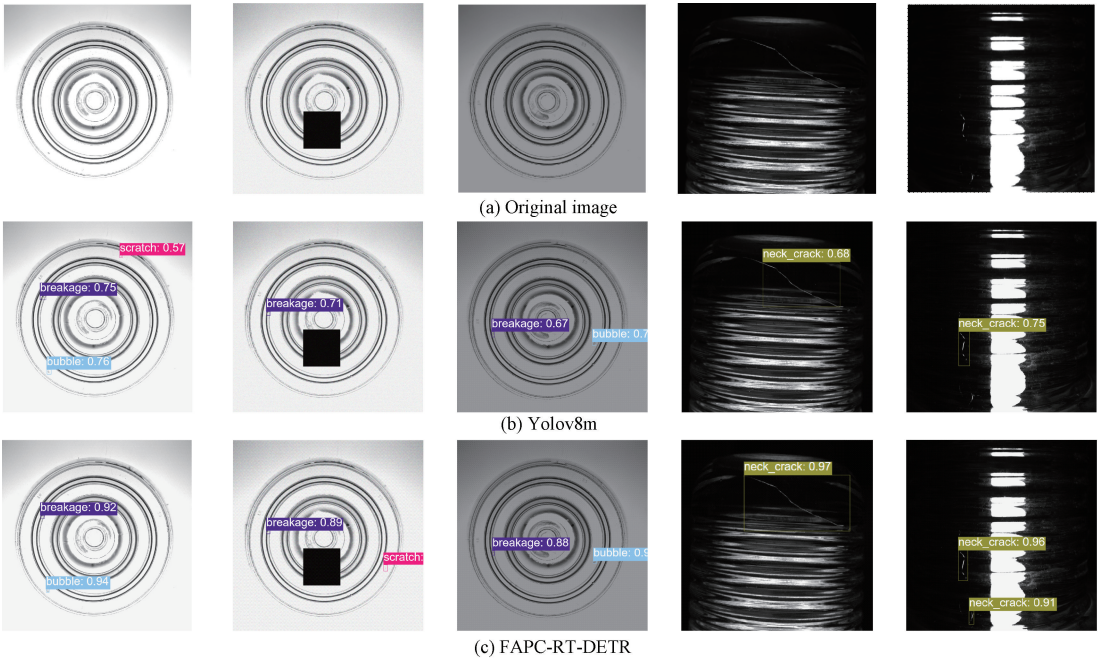


图 17 检测结果对比图
Fig. 17 Comparison of test results

4 结 论

本文提出了一种基于改进 RT-DETR 的轻量化玻璃绝缘子缺陷检测算法 FAPC-RT-DETR,重点解决了传统检测方法在低对比度缺陷、小目标缺陷及多尺度特征提取等方面的不足。通过理论分析与实验验证,得出以下主要结论:

轻量化网络设计的有效性:引入 RE-FasterNet 模块优化骨干网络结构,大幅降低了网络参数量和计算复杂度,同时提升了特征提取效率,为工业部署中的资源受限场景提供了可行性。

多尺度特征融合能力的增强:通过 PRCFusion 模块和 ASF 的应用,进一步增强了网络的多尺度特征捕获能力,显著提升了对于复杂玻璃绝缘子缺陷的检测精度,尤其是在小目标缺陷和低对比度缺陷的检测任务中表现出色。

算法性能的显著提升:相比其他主流目标检测算法,本文提出的 FAPC-RT-DETR 模型在平均检测精度、模型轻量化和计算效率方面均实现了全面优化,平均检测精度提高了 6.6%,且检测时间完全满足工业实时性要求。

参考文献

[1] 熊炜,黄玉谦,孟圣哲. 基于改进 YOLOv8 算法的绝缘子缺陷检测模型[J]. 电子测量技术, 2024, 47(12): 132-139.
XIONG W, HUANG Y Q, MENG SH ZH. Insulator defect detection model based on improved YOLOv8

algorithm[J]. Electronic Measurement Technology, 2024, 47(12): 132-139.
[2] CAO CH Q, WANG B, ZHANG W R, et al. An improved faster R-CNN for small object detection[J]. IEEE Access, 2019, 7: 106838-106846.
[3] BHARATI P, PRAMANIK A. Deep learning techniques— R-CNN to mask R-CNN: A survey[J]. Computational Intelligence in Pattern Recognition; Proceedings of CIPR 2019, 2020: 657-668.
[4] ZHANG Y F, CHI M M. Mask-R-FCN: A deep fusion network for semantic segmentation[J]. IEEE Access, 2020, 8: 155753-155765.
[5] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]. European conference on Computer Vision, Cham: Springer International Publishing, 2020: 213-229.
[6] KANG S H, PARK J S. Aligned matching: Improving small object detection in SSD[J]. Sensors, 2023, 23(5): 2589.
[7] JIANG P Y, ERGU D J, LIU F Y, et al. A review of Yolo algorithm developments[J]. Procedia Computer Science, 2022, 199: 1066-1073.
[8] 尹玲,叶正伟,陈新度,等. 基于深度学习的曲面玻璃表面缺陷检测方法[J]. 机床与液压,2023, 51(16): 120-125,132.

- YIN L, YE ZH W, CHEN X D, et al. A detection method for surface defects of curved glass based on deep learning[J]. Machine Tool & Hydraulics, 2023, 51(16): 120-125,132.
- [9] 张涛, 谢探阳, 李玉梅, 等. 基于改进 YOLOv4 算法的玻璃杯缺陷识别方法研究[J]. 电子测量技术, 2023, 46(2): 46-51.
- ZHANG T, XIE T Y, LI Y M, et al. Research on glass defect recognition method based on improved YOLOv4 [J]. Electronic Measurement Technology, 2023, 46(2): 46-51.
- [10] MAO Y CH, YUAN J L, ZHU Y J, et al. Surface defect detection of smartphone glass based on deep learning[J]. The International Journal of Advanced Manufacturing Technology, 2023, 127: 5817-5829.
- [11] 徐仟祥, 曾勇, 卢倩, 等. 基于改进 RT-DETR 的车门内拉手表面缺陷检测方法[J]. 电子测量技术, 2024, 47(18): 172-181.
- XU Q X, ZENG Y, LU Q, et al. Surface defect detection method for inner handle of car door based on improved RT-DETR [J]. Electronic Measurement Technology, 2024, 47(18): 172-181.
- [12] ZHAO Y AN, LYU W Y, XU SH L, et al. DETRs beat yolos on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [13] CHEN J, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [14] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP). IEEE, 2023: 1-5.
- [15] KANG M, TING C M, TING F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation[J]. Image and Vision Computing, 2024, 147: 105057.
- [16] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations[C]. IEEE/CVF Conference on computer Vision and Pattern Recognition, 2020: 1580-1589.
- [17] 付军军, 陈林, 童华敏, 等. 基于改进 EfficientNet 的海上风机叶片早期缺陷检测及分类[J]. 电子测量技术, 2024, 47(16): 110-119.
- FU J J, CHEN L, TONG H M, et al. Early defect detection and classification of offshore wind turbine blades based on improved EfficientNet[J]. Electronic Measurement Technology, 2024, 47(16): 110-119.
- [18] CHEN H T, WANG Y H, GUO J Y, et al. Vanillanet: The power of minimalism in deep learning[J]. Advances in Neural Information Processing Systems, 2023, 36: 7050-7064.
- [19] WANG C Y, LIAO H Y M, YEH I H. Designing network design strategies through gradient path analysis[J]. ArXiv preprint arXiv: 2211.04800, 2022.
- [20] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks[C]. 2018 IEEE Winter Conference on Applications of Computer Vision(WACV). IEEE, 2018: 839-847.
- [21] BOLYA D, FOLEY S, HAYS J, et al. Tide: A general toolbox for identifying object detection errors [C]. Computer Vision-ECCV 2020: 16th European Conference, 2020: 558-573.

作者简介

张靖雯, 硕士研究生, 主要研究方向为机器视觉。

E-mail: zjwen7304@163.com

孙坚(通信作者), 教授, 主要研究方向为自动化。

E-mail: sunjian@cjlu.edu.cn

徐红伟, 副教授, 主要研究方向为检测技术与自动化装置。

E-mail: xhw@cjlu.edu.cn

许素安, 教授, 主要研究方向为检测技术及其自动化。

E-mail: xusuan@cjlu.edu.cn

付紫平, 硕士研究生, 主要研究方向为机器视觉。

E-mail: 598943484@qq.com