

# 基于 YOLOv8 的铝电解电容外观缺陷检测方法

李泽沁, 赵子荣, 盛磊, 曾良涛, 姜丽, 谭德立

(广东东软学院 计算机学院, 广东 佛山 528225)

**摘要:** 传统铝电解电容质检依赖人工目视, 以识别外壳划痕、破损等缺陷, 这种检测方法存在检测准确性和效率的问题。为解决这一问题, 提出并实现了一种基于 YOLOv8 的电容缺陷检测方案。通过构建电容缺陷数据库, 并在 YOLOv8 模型中进行凹陷、划痕、破损等缺陷的训练, 方案成功实现了电容外观缺陷检测功能。实验数据结果显示, 该模型的 mAP@50 超过 87%。与传统检测方法相比, 基于 YOLOv8 的电容缺陷检测方案具有更高的准确性, 效率更高。进一步构建电容外观缺陷数据库, 可以提高检测准确性和效率, 为电容工业生产的缺陷检测提供了可行的解决方案。

**关键词:** 缺陷检测; YOLOv8; 深度学习; 卷积神经网络

中图分类号: TP391

文献标识号: A

文章编号: 2096-4706 (2024) 20-0055-06

## Appearance Defect Detection Method of Aluminum Electrolytic Capacitor Based on YOLOv8

LI Zeqin, ZHAO Zirong, SHENG Lei, ZENG Liangtao, JIANG Li, TAN Deli

(School of Computer Science, Neusoft Institute Guangdong, Foshan 528225, China)

**Abstract:** Traditional aluminum electrolytic capacitor quality inspection relies on manual visual inspection to detect defects such as scratches and damages on the casing, and this detection method encounters problems with accuracy and efficiency. To address the problems, this paper proposes and implements a capacitor defect detection scheme based on YOLOv8. By constructing a capacitor defect database and training the YOLOv8 model on defects such as dents, scratches, and damages, the scheme successfully achieves capacitor appearance defect detection function. Experimental data results show that the model's mAP@50 exceeds 87%. Compared to traditional detection methods, the capacitor defect detection scheme based on YOLOv8 has higher accuracy and efficiency. Further construction of a capacitor appearance defect database can enhance detection accuracy and efficiency, providing a feasible solution for defect detection in capacitor industrial production.

**Keywords:** defect detection; YOLOv8; Deep Learning; Convolutional Neural Networks

## 0 引言

电容作为电子器件中的关键组成部分, 其质量检测在生产流程中至关重要。然而, 传统的电容缺陷检测方式往往依赖于人工目视, 不仅效率低下, 且易受到人为因素的影响, 导致检测结果的准确性和稳定性受到质疑。因此, 如何高效、准确地实现电容缺陷的自动化检测, 成为当前研究的热点与难点。近年来, 深度学习技术在图像识别领域取得了显著的进展。本研究旨在利用目标检测技术, 以 YOLOv8 算法为实验模型, 设计并实现一款电容缺陷检测系统。通过构建电容缺陷数据库, 并训练 YOLOv8 模型以识别凹陷、划痕、破损等缺陷, 实验期望实现电容外观缺陷的自

动化检测, 提高检测的效率和准确性。本研究不仅解决了传统检测方法存在的问题, 还通过实际应用验证了基于 YOLOv8 的电容缺陷检测方案的可行性和有效性。

## 1 相关研究

近年来, 随着人工智能技术的发展, 缺陷检测系统<sup>[1]</sup>在工业生产中的应用越来越广泛。深度学习作为图像识别和目标检测领域的核心动力, 相较于传统缺陷检测方法展现出明显的优势。其自动学习和提取图像特征的能力消除了烦琐的手工设计和特征选择过程。此外, 深度学习具备强大的鲁棒性, 能有效处理各类图像变换和噪声, 从而提升了检测的准确性和稳定性。在电容生产中, 表面缺陷检测作为至关重要的工作, 通过基于深度学习的目标检测技术实现电容表面缺陷的检测变得可行。然而, 深度学习模型在训练和测试过程中需要大量计算资源, 而终端设备的计算能力有限, 这为电容表面缺陷检测的实际应用带来了

收稿日期: 2024-03-26

基金项目: 广东省大学生创新创业训练项目 (S202312574017); 广东大学生科技创新培育专项资金 (pdjh2024a469)

一定的困扰和挑战。在确保检测精度的前提下, 如何降低深度学习模型的复杂度和计算量成为当前亟须解决的关键问题。当前主流的目标检测算法众多, 本文特别介绍了一阶段的 YOLO (You Only Look Once) 算法。与两阶段的检测算法不同, 该算法基于回归, 将定位与分类任务合并, 获得更快的检测速度, 在实时检测方面具有显著优势。在电容表面缺陷检测领域, 基于深度学习的目标检测模型受到终端检测设备计算能力的限制, 模型部署依然面临巨大挑战。如何对计算量庞大的目标检测模型进行轻量化处理, 使其能够部署在计算资源有限的设备上, 成为当前电容表面缺陷检测领域研究的热点问题。在实际应用中, 该方法取得了较好的效果, 提高了检测的准确性和效率。杜宝侠等人提出了一种基于改进 YOLOv8 的苹果检测方法<sup>[2]</sup>, 该方法主要应用于苹果的表面缺陷检测。通过改进 YOLOv8 模型, 提高了对苹果表面缺陷的识别率, 为后续的分类和筛选提供了有力支持。陈道贺等人提出了一种基于机器视觉的玻璃瓶口缺陷检测方法<sup>[3]</sup>。该方法利用图像处理技术对玻璃瓶口进行预处理、特征提取和分类, 实现了瓶口的缺陷检测。实验结果表明, 该方法具有较高的准确性和稳定性, 能够满足实际生产的需要。Li 等人基于改进 YOLOv5 的目标检测算法开发设计了一种能够快速、准确地检测航空发动机部件表面缺陷的方法<sup>[4]</sup>。此方法在传统的 YOLOv5 基础上, 进行了一系列的改进, 包括增加特征提取模块、使用注意力机制、优化损失函数等, 以提高对微小缺陷的检测能力。PHAN Q B 等人将 YOLOv8 目标检测算法与粒子群优化算法相结合设计了一种快速高效检测光伏电池故障的系统<sup>[5]</sup>。结果显示结合了粒子群优化算法的 YOLOv8 在故障检测方面表现优异, 大大提高了检测的准确性和效率。李俊杰等人在 2023 年提出了一种基于改进 YOLOv3 (You Only Look Once version 3) 的电容表面缺陷检测方法<sup>[6]</sup>。该方法利用 YOLOv3 算法的高效性和准确性, 通过对电容表面图像进行训练和学习, 实现对电容表面缺陷的自动检测。与传统的检测方法相比, 基于深度学习的方法具有更强的鲁棒性和泛化能力, 能够应对更复杂多变的缺陷类型<sup>[7]</sup>。虽然 YOLOv3 在目标检测领域已经取得了很好的效果, 但随着深度学习技术的不断发展, 更新的算法也在不断涌现。YOLOv8 作为 YOLO 系列的最新版本, 在检测精度和速度上都有了进一步的提升。由上述文献可看出, 缺陷检测需对作业对象开展特定位置识别外, 例如在苹果、玻璃瓶口缺陷检测等物质检测领域也有相应的应用和需求, 但针对电容的识别和定位研究较少, 且现有成果

在识别率及定位精度上有较大提升空间, 因此本文在 YOLOv8 的基础上通过自建数据集训练铝电解电容缺陷检测模型并探讨 YOLOv8 模型的工作原理与数据集构建方法。

## 2 基于 YOLOv8 的铝电解电容外观缺陷检测

### 2.1 YOLOv8 模型

YOLOv8 是 Ultralytics 公司于 2023 年推出的一款新型开源的计算机视觉任务模型, 该模型具备强大的图像分类、物体检测和实例分割的能力。相较于其前身 YOLOv5, YOLOv8 引入了许多重要的改进, 在目标检测领域取得了显著的性能提升。作为一种 SOTA 模型, YOLOv8 在基于之前版本成功经验的基础上, 融入了新的功能和优化, 使得其结构更为高效和准确。该模型主要由三个关键部分组成: 主干网络 (Backbone)、颈部网络 (Neck) 和检测 (Head)。这种结构使得信息能够在不同层之间流畅传递, 从而提高特征提取和物体检测的准确性<sup>[8]</sup>。

YOLOv8 模型中主干网络主要由 Conv 模块、C2f 模块、SPPF 模块组成, 该模块中的 Conv 负责输入图像做卷积运算提出图像的特征信息。C2f 通过引入残差连接和更多的跳层连接, 使得网络更好的学习全局信息和局部语义信息<sup>[9-10]</sup>, 从而增强模型的特征信息提取能力。SPPF 模块是一种用于特征提取的结构, 通过引用空间金字塔池化机制, 增强模型对不同尺度目标的感知能力, 进而提高模型的感受野。颈部网络 (Neck) 通常负责将骨干网络 (Backbone) 提取的特征与上采样后的特征进行融合, 以生成最终用于检测的输出。移除某些连接层可能会改变这一过程, 但具体的影响取决于其他网络部分的调整。通过这些优化措施, YOLOv8 进一步提高了网络的性能和准确性。在 YOLOv8 检测 (Head) 模块中, 与 YOLOv5 相比, 该模型在检测头部分进行了重大改进<sup>[11]</sup>。它采用了目前主流的解耦头结构, 将分类任务和检测任务在结构上进行了分离。这种巧妙设计使得模型能够更专注地处理各自的任務, 从而更精确地执行分类和检测任务。通过这种优化, YOLOv8 显著提升了性能和精确度, 在目标检测任务中展现出出色的表现。除了检测头的改进, YOLOv8 还从原先的 Anchor-Based 方法迁移到了 Anchor-Free 方法, 这种改进允许模型可以直接学习预测物体的边界框, 无须预先定义锚点框。从而减小了锚点框数量和尺寸选择对模型的干扰, 使模型能够更加灵活地适应不同场景中物体形状和大小的变化<sup>[12]</sup>。

YOLOv8 还提供了五种不同大小的模型来满足在不同的场景需求, 分别是 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x。这些不同大小的模型可以根据实际应用的需求进行选择, 从而在性能和计算效率之间取得平衡。通过灵活调整模型大小, YOLOv8 能够更好地适应不同场景的目标检测任务<sup>[13]</sup>。YOLOv8 在目标检测领域中取得了重大突破。通过引入新的功能和改进, 该模型在准确性、鲁棒性和计算效率方面表现出色。从解耦头结构到 Anchor-Free 方法的迁移, 再到骨干网络和颈部网络的优化, YOLOv8 展现了强大的性能和潜力, 网络结构如图 1 所示。

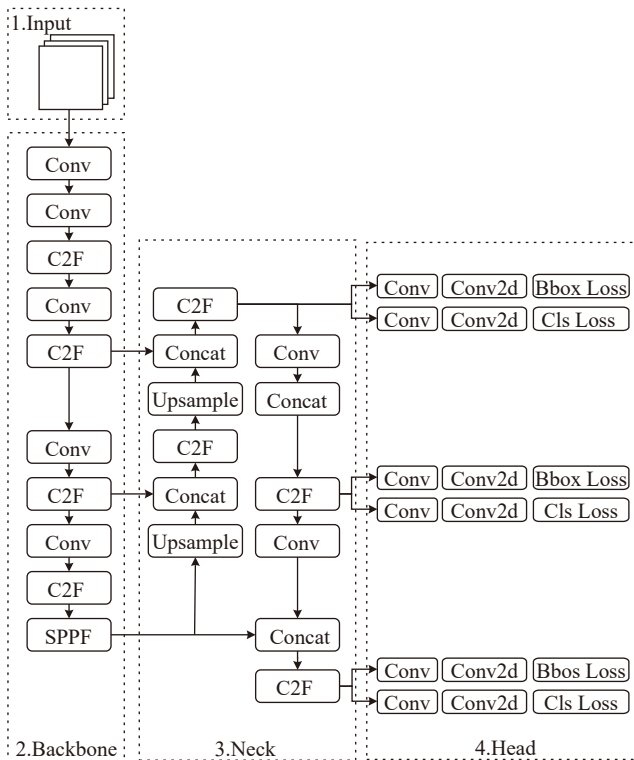


图 1 YOLOv8 模型检测网络

2.2 总体系统框架设计

总体系统框架以 YOLOv8 模型为基础, 系统总体框架如图 2 所示, 为了实现铝电解电容外观缺陷检测需通过 OpenCV 调用工业相机并将视频流逐一分解成单张照片, 为后续模型推理做准备。基于 YOLOv8 模型的目标检测能力为系统提供了高效而准确的缺陷检测工具, 该模型已经在训练阶段学习了铝电解电容外观的特征, 使其能够在图像中识别和定位潜在的缺陷区域。在推理运算阶段, YOLOv8 模型将应用于单张照片, 进行预测并标定图像中检测到的缺陷, 通过 YOLOv8 内置 plot() 函数绘制检测结果并返回注释图像的 NumPy 数组, 使得检测到的缺陷信息可以被

轻松地集成到进一步的分析或处理流程中, 为融合下游任务做准备。

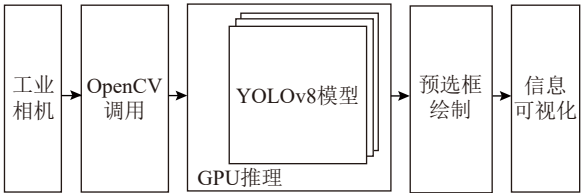


图 2 系统总体框架

3 实验数据

3.1 数据集获取

本文的数据通过收集不同缺陷样本的铝电解电容, 并在严格控制条件对电容进行筛选, 选定划痕 (Scratched)、破损 (Dented)、凹陷 (Damaged) 作为铝电解电容的缺陷样本检测类型, 构建一个多样化的电容缺陷数据集。在数据集获取和制作的过程中, 利用工业相机对每个电容进行多方位、多角度的拍摄, 特别要求对光线的要求, 确保拍摄出的图像能够真实反应电容器的实际状况, 保证每个电容的缺陷特征和程度能够准确记录。

3.2 数据集预处理

在数据集预处理过程中, 对图像进行了大小调整。通过根据工业相机拍摄电容器的实际尺寸和距离, 对图像进行缩放处理, 确保每一张实验数据集图像具有相同的尺寸和格式, 提高数据的一致性。同时, 去除了空白、模糊等无效样本图片。预处理后的图像在 Labeling 软件中接受标注处理, 涉及对凹陷、划痕、破损缺陷类型的标注, 以确保对后续模型训练。

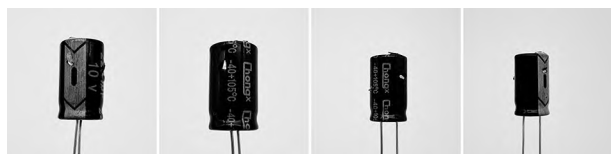
3.3 数据集制作

在数据集制作过程中, 首先将经过处理的电容图像与相应的缺陷标签进行精确匹配组合, 形成一个完整的数据集。该过程需要借助高效的数据库管理, 确保每个图像与对应的标签能够精准无误。为进一步扩充数据集并提高模型的泛化能力和鲁棒性, 采用了交叉验证的方法, 分别构建了训练集 (1 476 张图像)、验证集 (368 张图像), 以全面评估模型的性能。最后, 为了增加样本数量, 使用数据增强技术, 对电容图像进行了旋转、翻转和缩放等操作。进一步优化数据集, 为模型训练提供多样和丰富的样本, 如图 3 所示。



(a) 电容划痕样本





(b) 电容破损样本



(c) 电容凹陷样本

图 3 电容缺陷数据集

## 4 实验结论分析

### 4.1 实验环境

在本次实验中, 本文选择了 NVIDIA GeForce RTX 3060 作为操作计算平台, 操作系统选择 Ubuntu 20.04。使用 Python 作为编程语言, 版本为 3.10, 并采用深度学习框架 PyTorch, 版本为 Torch 1.11.0。此外, CUDA 版本为 11.3, 用于加速深度学习计算。实验环境的详细配置如表 1 所示。

表 1 实验环境配置

实验环境	配置
操作系统	Ubuntu-20.04
GPU	NVIDIA GeForce RTX 3060
Python	3.10
PyTorch	1.11.0
CUDA	11.3

### 4.2 实验评判标准

在目标检测训练任务中常以准确率 (Precision)、召回率 (Recall)、mAP@50 和 mAP@50-90 作为评估模型性能的标准。其中 TP (True Positives) 表示模型正确预测为正例的数量, FP (False Positives) 为模型错误预测为正例的数量, FN (False Negatives) 为模型错误预测为负例的数量, 准确率 (Precision) 和召回率 (Recall) 计算公式如下:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

mAP 表示评估目标检测模型整体性能指标, 本文选取 mAP@50 (在 IoU 阈值为 0.50 时的平均精度均值) 和 mAP@50-95 (在 IoU 阈值从 0.50 到 0.95 范围内的平均精度均值) 阈值来计算 mAP。在计算 mAP@50, 首先对于每个目标类别计算在 IoU 阈值为 0.50 时的精度 - 召回率曲线下面积, 即 AP@50。然后,

将所有类别的 AP@50 取平均, 得到 mAP@50, 其中 C 表示类别的数量, AP 表示第 i 个类别的 AP@50, 如下所示:

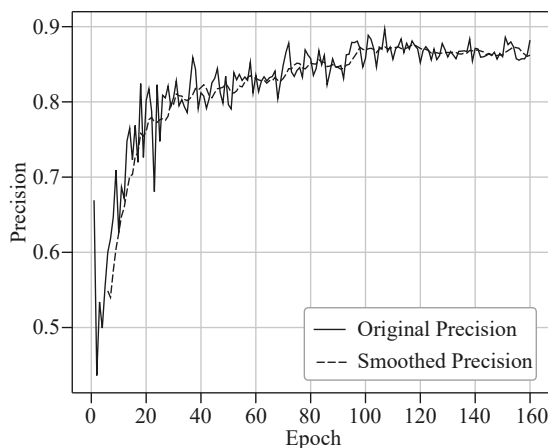
$$\text{mAP}_{50} = \frac{1}{C} \sum_{i=1}^C \text{AP}_i \quad (3)$$

在计算 mAP@50-95 时, 考虑了 IoU 阈值从 0.50 到 0.95 的范围, 即对于每个目标类别, 计算在 IoU 阈值从 0.50 到 0.95 的精度 - 召回率曲线下面积的平均值, 将所有类别的 AP@50-95 取平均得到 mAP@50-95。在式 (4) 中, C 表示类别的数量, AP 表示第 i 个类别在 IoU 阈值从 0.50 到 0.95 的 AP 的平均值, 计算公式如下所示:

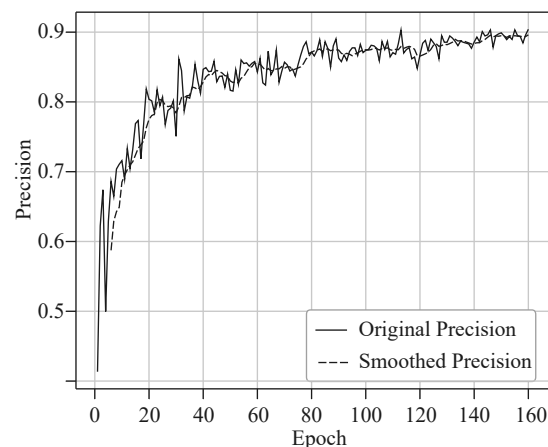
$$\text{mAP}_{50 \sim 95} = \frac{1}{C} \sum_{i=1}^C \text{AP}_{50 \sim 95}^i \quad (4)$$

### 4.3 实验结果分析

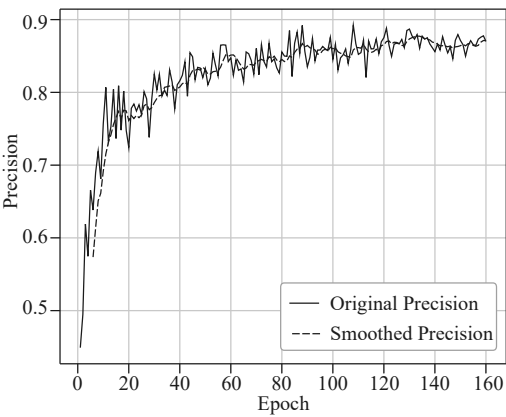
将实验结果可视化处理绘制精确度、平均精度均值、召回率, 如图 4 所示, 实验结果表明 YOLOv8 模型在 Epoch 大于 100 之后模型开始收敛, 表现出良好的性能。



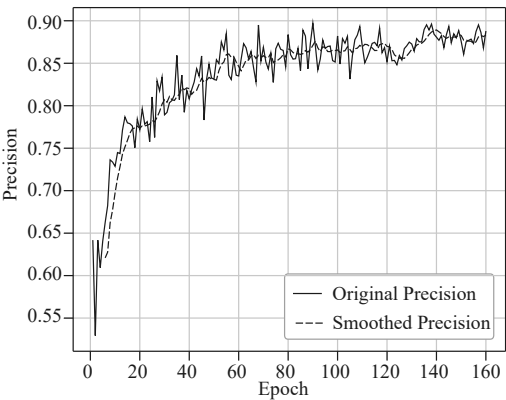
(a) YOLOv8l 模型精度折线图



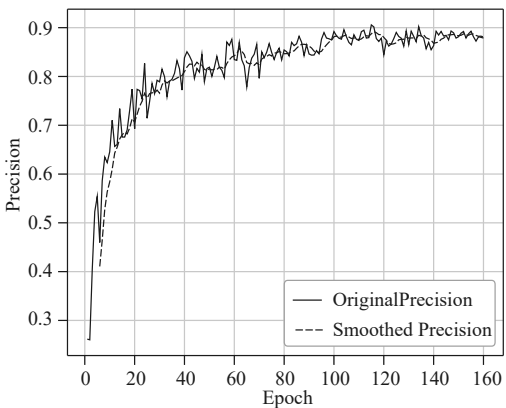
(b) YOLOv8m 模型精度折线图



(c) YOLOv8n 模型精度折线图



(d) YOLOv8s 模型精度折线图



(f) YOLOv8x 模型精度折线图

图 4 YOLOv8 模型精度训练结果

本文通过对 YOLOv8 模型的训练实验, 针对不同版本的模型 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x 进行性能评估。研究关注模型的精确度、召回率、mAP@50 和 mAP@50-95 等指标。从表 2 中可以分析可知, 在电容缺陷样本数据集中, YOLOv8m 模型表现最优, 显示该模型在目标预测方面更为准确。相反, YOLOv8s 在召回率上表现出色, 即更多真实目标被成功检测。在平均精度 (mAP@50 和 mAP@50-95) 指标上分析, YOLOv8s 和 YOLOv8x 在不同的 IoU 阈值下均呈现出较高的性

能。模型网络层数的不同也反映了各版本在模型结构上的差异, YOLOv8l 和 YOLOv8x 共享较大的网络深度, 这表示该模型更注重对抽象特征的学习。综合对比实验结果表明, 不同版本的模型为了适应使用场景的需求, 对模型网络结构做出相应的结构调整以支持不同终端的算力需求。

表 2 YOLOv8 模型训练结果 单位 /%

Model	Precision	Recall	mAP@50	mAP@50-95
YOLOv8n	88.6	78.7	88.1	56.1
YOLOv8s	89.1	81.5	88.3	57.8
YOLOv8m	89.3	79.2	87.5	57.1
YOLOv8l	86.6	80.6	87.4	56.7
YOLOv8x	88.1	82.2	88.2	57.4

4.4 检测结果可视化

在模型测试过程中, 随机选取划痕、凹陷、破损样本进行实验测试, 如图 5 所示, 在测试过程中 YOLOv8 模型均能捕捉到电容缺陷样本并绘制出相应的预选框。为了进一步验证模型, 使用交叉缺陷样本进行缺陷检测, 图 5 所示的测试结果表明模型能够正确判断样本缺陷类型, 说明该方法具有合理性和实际应用价值。



(a) 电容缺陷样本检测效果图



(b) 电容缺陷样本检测效果图

图 5 电容缺陷样本检测

5 结 论

本文以铝电解电容缺陷检测为背景, 针对在工业生产环节对铝电解电容良品率的需求, 本文提出了一种基于 YOLOv8 的电容缺陷检测系统。实验结果表明, YOLOv8n 的精度 mAP@50 为 88.1%, YOLOv8s 为 88.3%, YOLOv8m 为 87.5%, YOLOv8l 为 87.4%, YOLOv8x 为 88.2%, 平均值为 87.9%。实验数据表明该模型能够准确识别出缺陷的电容, 相较于传统人工目检, 该系统具有更高的准确率, 显著提升了电容工业生产缺陷检测环节的准确率, 降低了电容的残次率。基于 YOLOv8 的电容缺陷检测系统在实际生产环节中, 具备广泛的应用价值, 该研究为工业自动化检测领域的发展提供理论基础。

## 参考文献:

[1] 伍济钢, 梁谋, 曹鸿, 等. 基于改进 YOLOv5 的 PCB 小目标缺陷检测研究 [J]. 光电子·激光, 2024, 35 (2): 155-163.

[2] 杜宝侠, 唐友, 辛鹏, 等. 基于改进 YOLOv8 的苹果检测方法 [J]. 无线互联科技, 2023, 20 (13): 119-122.

[3] 陈道贺, 姜文斌. 基于机器视觉的玻璃瓶口缺陷检测方法 [J]. 自动化应用, 2021 (4): 141-143.

[4] LI X, WANG C, JU H J, et al. Surface Defect Detection Model for Aero-engine Components Based on Improved YOLOv5 [J/OL]. Applied Sciences, 2022, 12 (14): 7235 (2022-07-18). <https://www.mdpi.com/2076-3417/12/14/7235>.

[5] PHAN Q B, NGUYEN T T. A Novel Approach for PV Cell Fault Detection Using YOLOv8 and Particle Swarm Optimization [EB/OL]. (2023-04-26). <https://www.techrxiv.org/doi/full/10.36227/techrxiv.22680484.v1>.

[6] 李俊杰, 周骅, 唐纲浩. 基于改进 YOLOv3 的电容表面缺陷检测方法 [J]. 智能计算机与应用, 2023, 13 (3): 235-241.

[7] 罗俊. 基于深度学习的混凝土结构表面病害分割算法 [D]. 石家庄: 石家庄铁道大学, 2022.

[8] 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.

[9] 陈晨, 徐慧英, 朱信忠, 等. 基于 YOLOv8 改进的室内行人跌倒检测算法: FDW-YOLO [J/OL]. 计算机工程

与科学, 2023: 1-14 (2023-12-28). <http://kns.cnki.net/kcms/detail/43.1258.TP.20231227.1421.002.html>.

[10] 程焕新, 矫立浩, 骆晓玲, 等. 改进 YOLOv8 的遥感图像检测算法 [J]. 无线电工程, 2024, 54 (5): 1155-1161.

[11] 王金鹏, 何萌, 甄乾广, 等. 基于 COF-YOLOv8n 的油茶果静、动态检测计数 [J/OL]. 农业机械学报, 2024: 1-15 (2024-01-17). <http://kns.cnki.net/kcms/detail/11.1964.S.20240117.0910.002.html>.

[12] 吴磊, 储钰昆, 杨洪刚, 等. 面向铝合金焊缝 DR 图像缺陷的 Sim-YOLOv8 目标检测算法 [J/OL]. 中国激光, 1-15 (2024-02-22). <http://kns.cnki.net/kcms/detail/31.1339.tn.20240220.1338.050.html>.

[13] 戴康佳, 徐慧英, 朱信忠, 等. 基于轻量化目标检测网络的 RGB-D 视觉 SLAM 系统 [J/OL]. 计算机工程与科学, 2024: 1-12 (2024-01-30). <http://kns.cnki.net/kcms/detail/43.1258.TP.20240129.1701.002.html>.

**作者简介:** 李泽沁 (2001—), 男, 汉族, 广东湛江人, 本科在读, 研究方向: 嵌入式系统; 赵子荣 (2002—), 男, 汉族, 广东云浮人, 本科在读, 研究方向: 嵌入式系统; 盛磊 (2003—), 男, 汉族, 广东广州人, 本科在读, 研究方向: 嵌入式系统; 曾良涛 (2003—), 男, 汉族, 广东韶关人, 本科在读, 研究方向: 嵌入式系统; 姜丽 (1985—), 女, 汉族, 湖北武汉人, 教师, 硕士, 研究方向: 嵌入式开发, 数字信号处理; 谭德立 (1962—), 男, 汉族, 湖南长沙人, 副研究员, 硕士, 研究方向: 嵌入式系统。

(上接 54 页)

[12] 龚永罡, 裴晨晨, 廉小亲, 等. 基于 Transformer 模型的中文文本自动校对研究 [J]. 电子技术应用, 2020, 46 (1): 30-33+38.

[13] 吕云飞. 电容式微机械超声换能器建模与仿真技术研究 [D]. 太原: 中北大学, 2018.

[14] 李清格, 杨小冈, 卢瑞涛, 等. 计算机视觉中的 Transformer 发展综述 [J]. 小型微型计算机系统, 2023, 44 (4): 850-861.

[15] 唐潘, 汪学明. 融合时间感知与兴趣偏好的推荐模型研究 [J]. 计算机工程与应用, 2023, 59 (24): 268-276.

[16] 刘敬瑾. 基于图卷积网络的跨被试脑电情绪识别研究 [D]. 汕头: 汕头大学, 2022.

[17] 张静. 融合用户和产品信息的情感分类研究 [D]. 北京: 北京工业大学, 2020.

[18] CARION N, MASSA F, SYNNAEVE G, et al. End-to-End Object Detection with Transformers [C]//Computer Vision-ECCV 2020.Glasgow: Springer, 2020: 213-229.

[19] DAI Z H, LIU H X, LE Q V, et al. CoAtNet: Marrying Convolution and Attention for All Data Sizes [J/OL]. arXiv:2106.04803 [cs.CV].[2024-03-01].<https://doi.org/10.48550/arXiv.2106.04803>.

[20] BEAL J, KIM E, TZENG E, et al. Toward Transformer-Based Object Detection [J/OL]. arXiv:2012.09958 [cs.CV].[2024-02-26].<https://arxiv.org/abs/2012.09958>.

[21] REN S Q, HE K M, 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.

[22] YAN H T, LI Z, LI W J, et al. ConTNet: Why not Use Convolution and Transformer at the Same Time? [J/OL]. arXiv:2104.13497 [cs.CV].[2024-02-25].<https://arxiv.org/abs/2104.13497>.

[23] YOO J, KIM T, LEE S, et al. Enriched CNN-Transformer Feature Aggregation Networks for Super-Resolution [J/OL]. arXiv:2203.07682 [cs.CV].[2024-02-26].<https://arxiv.org/abs/2203.07682>.

[24] XIAO T, SINGH M, MINTUN E, et al. Early Convolutions Help Transformers See Better [J/OL]. arXiv:2106.14881 [cs.CV].[2024-02-20].<https://arxiv.org/abs/2106.14881>.

[25] LI Y W, ZHANG K, CAO J Z, et al. LocalViT: Bringing Locality to Vision Transformers [J/OL]. arXiv:2104.05707 [cs.CV].[2024-02-15].<https://arxiv.org/abs/2104.05707>.

**作者简介:** 于祉祺 (1996—), 女, 汉族, 黑龙江富锦人, 学术型硕士, 研究方向: 无损检测; 刘皓源 (1997—), 男, 汉族, 辽宁鞍山人, 沈阳工业大学, 学术型硕士, 研究方向: 无损检测; 何璐瑶 (1994—), 女, 汉族, 辽宁沈阳人, 副教授, 博士, 研究方向: 在线检测; 杨理践 (1957—), 男, 汉族, 辽宁沈阳人, 教授, 博士, 研究方向: 在线检测; 刘斌 (1981—), 男, 汉族, 辽宁沈阳人, 教授, 博士, 研究方向: 在线检测。