

电池字符缺陷检测 DDP-YOLOv8 模型方法^{*}

李绪涛 邓耀华

(广东工业大学机电工程学院 广州 510000)

摘要:针对消费电池产品表面字符缺陷检测中存在的缺陷位置动态分布、多尺度适应性差和细小缺陷识别困难等关键技术难题,提出了一种创新性的可变形大核卷积注意力、动态采样和 P2-动态检测头的 YOLOv8 (DDP-YOLOv8) 检测模型框架。首先,针对 YOLOv8 在特征提取过程中无法有效调整特征图权重的问题,设计 DCNv3-LKA 注意力模块,通过融合动态卷积网络与大核注意力机制,在特征提取阶段实现空间权重自适应调整。其次,针对 YOLOv8 颈部网络在字符缺陷检测中的采样位置固定和多尺度适应性差的问题,对 YOLOv8 的颈部网络结构进行重构,采用跨尺度特征融合模块 (CCFM) 架构并提出一种引入了动态偏移量与可学习采样权重双驱动机制的动态采样器 DS (DS-CCFM 模块),突破传统特征金字塔的固定几何约束。最后,针对消费电池产品表面字符小尺度及 YOLOv8 检测头使用普通卷积层导致的特征表达不足与信息丢失问题,增加 P2 小目标检测层并在检测头融入 DynamicHead 多个自注意力机制 (P2-DynamicHead 模块),提升对微小缺陷的捕获能力。实验结果表明,DCNv3-LKA、DS-CCFM 和 P2-DynamicHead 模块分别使模型在字符缺陷数据集上的平均精度均值 (mAP) mAP@0.5 达到 91.8%、91.2% 和 92.4%,相较于 YOLOv8n 分别提高了 1.7%、1.1% 和 2.3%。DDP-YOLOv8 最终实现了 94.0% 的 mAP@0.5,相较于基准模型 YOLOv8n 提升了 3.9%,模型检测速度为 85.1 fps,满足电池大规模定制生产中字符缺陷检测对高精度与实时性的需求。

关键词: 字符缺陷; YOLOv8; 注意力; 动态采样; 动态检测头

中图分类号: TP391.4; TN98

文献标识码: A

国家标准学科分类代码: 520.6040

DDP-YOLOv8 model for battery character defect detection

Li Xutao Deng Yaohua

(School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510000, China)

Abstract: To address the critical challenges in surface character defect detection of consumer batteries, including dynamic defect localization, multi-scale adaptability, and fine-scale defect recognition, this paper proposes an innovative DDP-YOLOv8 framework. Firstly, to resolve the limitation of YOLOv8 in effectively adjusting feature map weights during feature extraction, we design a DCNv3-LKA attention module to achieve adaptive spatial weight adjustment through dynamic convolution and large-kernel attention fusion. Secondly, aiming to overcome the fixed sampling positions and poor multi-scale adaptability of YOLOv8's neck network in character defect detection, we restructure the neck architecture by adopting a CCFM framework and propose a dynamic sampler (DS-CCFM module) incorporating dual-driven dynamic sampling mechanism. Finally, to mitigate the insufficient feature representation and information loss caused by standard convolution layers in YOLOv8's detection head when handling small-scale battery characters, we introduce a P2 small-target detection layer and integrate multiple self-attention mechanisms from DynamicHead into the detection head (P2-DynamicHead module) to improves small defect recognition. Experimental results demonstrate that the DCNv3-LKA, DS-CCFM, and P2-DynamicHead modules achieve mean average precision (mAP) mAP@0.5 of 91.8%, 91.2%, and 92.4% respectively on the character defect dataset, representing improvements of 1.7%, 1.1%, and 2.3% over baseline YOLOv8n. DDP-YOLOv8 achieves a final mAP@0.5 of 94.0%, representing a 3.9% improvement over the baseline model YOLOv8n. With an FPS of 85.1, the model

收稿日期: 2024-10-21 Received Date: 2024-10-21

^{*} 基金项目: 国家自然科学基金 (52175457)、广东省基础与应用基础研究基金 (2022B151520053)、东莞市重点领域研发项目 (20221200300042) 资助

meets the requirements of high accuracy and real-time performance for character defect detection in large-scale customized battery production.

Keywords: character defect; YOLOv8; attention; dynamic sampling; dynamichead

0 引言

消费电池表面的字符文本不仅传递产品信息,还承载了关键的制造和质量控制信息。然而,字符印刷过程易受生产工艺、材料质量和环境等因素的影响,导致字符缺印、模糊和污染等常见缺陷的出现。因此,准确检测产品表面字符缺陷对于保障产品质量与生产效率至关重要。

传统的字符缺陷检测方法主要分为人工抽检和机器视觉检测两类。人工抽检方式由于检测效率低、标准不统一、人工成本高等缺点,正逐步被现代化企业淘汰。基于机器视觉的字符缺陷检测方法则通过字符定位和分割,再与标准图像进行对比,从而判断字符的质量。熊继淙等^[1]采用基于形状的模板匹配算法快速定位目标区域,结合灰度值差影匹配算法,实现了锂电池标签字符缺陷的检测。裴炜冬^[2]则提出通过多模板分区域的方法检测锂电池表面字符的缺陷,相较于传统的图像像素差值法,其检测的鲁棒性得到了提升。然而,基于机器视觉的字符缺陷检测方法在图像质量和字符样式变化时,通常需要更新标准对比图像和调整算法参数。这一过程导致较高的维护成本和时间开销,影响大规模定制生产中快速响应和灵活调整的需求。

近年来,深度学习技术为字符缺陷检测提供了有效工具。卢丹等^[3]改进了 LeNet-5 网络以应用于汽车手柄字符缺陷检测,通过改变输入图像像素增加缺陷字符的缺陷特征信息,并引入 Adam 优化器和批量归一化操作,提升了检测的准确性。赵卫东等^[4]结合 Faster RCNN 与 RetinaNet,提出了端到端的电表字符检测算法,先通过 Faster RCNN 提取字符区域,再利用 RetinaNet 精确检测缺陷。Peng 等^[5]基于 EfficientNet 的核心模块构建了名为 BBE 的网络,主要由特征提取网络 BUNet、特征融合网络 BWNet、分类和回归网络组成,用于复杂背景下喷码字符的缺陷检测工作。然而,这些改进方法和模型在复杂场景、大规模图像处理或需要高精度时,存在扩展性不足、计算开销较大或对参数敏感的问题,影响了模型的通用性和实际应用效果。

随着目标检测技术的快速发展,YOLO^[6]系列算法凭借其精度与速度优势,在实时对象检测中表现突出。特别是 YOLOv8 算法,因其在高效定位和分类各种缺陷方面的优势,已被广泛应用于缺陷检测领域。李冰等^[7]针对复杂环境下风机桨叶缺陷检测精度较低的问题,提出

了一种改进 YOLOv8 的风机桨叶缺陷检测算法。通过引入 LSKA 注意力机制并采用加权双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)结构,增强了多尺度特征融合能力,提升了小目标检测的精度。杜昌都等^[8]通过在 YOLOv8 中加入卷积块注意力模块(convolutional block attention module, CBAM)机制,提升了贴片芯片引脚缺陷检测的精度。王安静等^[9]则引入 GFPN 特征融合模块,加强了相邻层级间的跨尺度连接和同尺度下的跨层连接,改善了鼓形滚子表面缺陷特征在卷积网络中的传递效果。

消费电池产品表面的字符缺陷往往较小且复杂,位置多变,类型繁多,且常伴随细微变化。当前基于 YOLOv8 改进的缺陷检测方法在字符缺陷检测场景中存在局限,即常规注意力机制难以应对背景干扰,固定采样策略导致细微缺陷特征丢失,检测头对小目标检测能力弱。故在 YOLOv8 的基础框架上,提出了一种基于可变形大核卷积注意力、动态采样和 P2-动态检测头的 YOLOv8 (deformable large kernel convolutional attention, dynamic sample, and P2-dynamichead YOLOv8, DDP-YOLOv8)模型。1)主干网络部分,创新性地设计结合大核注意力机制的可变形卷积网络(deformable convolutional networks v3 with large kernel attention, DCNv3-LKA)。2)颈部网络部分,对 YOLOv8 的颈部网络结构进行重构,采用跨尺度特征融合模块(cross-scale feature fusion module, CCFM)架构并创造性地设计一种引入了动态偏移量和可学习采样权重的动态采样器(dynamic sampler, DS)。3)检测头部分,增加 P2 小目标检测层并在检测头融入 DynamicHead 多个注意力机制。

1 字符缺陷检测 DDP-YOLOv8 模型构建

1.1 DDP-YOLOv8 模型总体架构

YOLOv8 模型主要包括输入模块、主干网络、颈部网络和检测头 4 个部分。其中,主干网络部分采用 C2f^[10]模块和 SPPF^[11]模块,颈部网络结合特征金字塔网络(feature pyramid network, FPN)和路径聚合网络(path aggregation network, PAN)两种主要的网络结构形成了 PAN-FPN 架构^[12],检测头使用解耦头,并采用无锚框的思想。

针对消费电池产品表面字符缺陷的特点,现有的 YOLOv8 模型存在一定的局限性。由于字符缺陷在形

状、排列、大小及空间分布上具有较强的多样性,设计了一种注意力机制模块——DCNv3-LKA,该模块结合大核卷积和改进的可变形卷积,从而使得模型能够更加聚焦于图像中的关键区域,灵活适应不同字符缺陷类型的变换。在此基础上,对模型的颈部结构进行重要优化,提出 DS-CCFM 模块,该模块采用 CCFM 网络架构,有效融合了来自不同通道的特征,并设计了一种高效的动态采样

器 DS,使模型具备了同时调整采样点位置和权重的能力,从而提升了字符检测任务的表现。考虑到字符缺陷通常为小尺寸目标,设计了 P2-DynamicHead 模块,增加 P2 小目标检测层以捕捉更多的微小特征,同时采用了 DynamicHead 检测头提供的多个注意力机制和池化层,使模型能够充分表达复杂的目标特征,进一步提高检测准确率。DDP-YOLOv8 网络模型具体结构如图 1 所示。

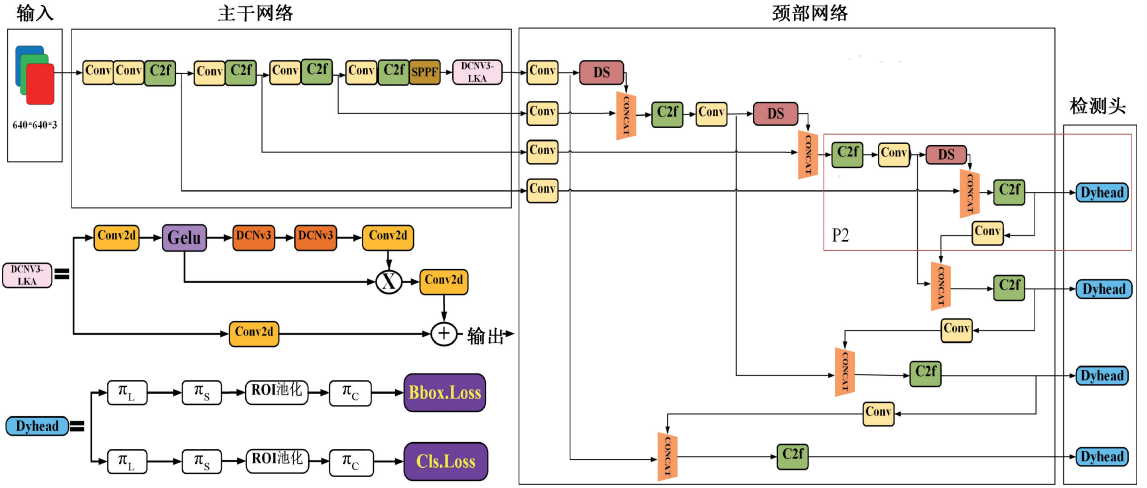


图 1 DDP-YOLOv8 网络架构

Fig. 1 DDP-YOLOv8 network architecture diagram

1.2 DDP-YOLOv8 模型各模块设计

1) DCNv3-LKA 模块设计

DCNv3-LKA 模块的核心是带有偏移量的大核变形卷积^[13]。大卷积核注意力 (large kernel attention, LKA)^[14]是一种用于捕捉图像中广泛上下文信息的机制,模仿了自注意力机制^[15]的感受野 (receptive field, RF),但是使用更少的计算量。假设输入特征图为 X , 其尺寸为 $H \times W \times C$, 卷积核的尺寸为 $K \times K$, 自注意力机制的计算过程如下:

首先,生成查询 (query, Q), 键 (key, K) 和值 (value, V) 矩阵:

$$Q = XW_Q \quad (1)$$

$$K = XW_K \quad (2)$$

$$V = XW_V \quad (3)$$

式中: W_Q, W_K, W_V 是可训练的权重矩阵, 尺寸均为 $C \times C$ 。

然后计算注意力分数:

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (4)$$

式中: softmax 函数是一种将一个包含任意实数的向量转换为一个概率分布的激活函数; d_K 为键向量的维度, 通常为 C 。注意力权重的计算结果 $\text{softmax}\left(\frac{QK^T}{\sqrt{d_K}}\right)V$ 将每

个位置的输出特征都考虑了输入特征图上所有位置的信息,从而实现了全局感受野。然而,自注意力机制的计算复杂度为 $O(H^2W^2C)$, 在高分辨率特征图上计算量非常大。LKA 通过大核卷积将计算复杂度降为 $O(HWK^2C)$, 由于 K^2 通常远小于 $H \times W$, 故 LKA 的计算复杂度相较于自注意力机制大幅减少。

在 LKA 的基础上,本模块通过优化特征提取网络,提高了模型对字符缺陷的自适应。如图 2 所示,首先引入大小和方向都可以任意变化的偏移量来动态地调整卷积核的采样位置,对输入特征图进行可变形卷积^[16]。设输入特征图为 X , 卷积核为 W , 卷积核的采样位置为 p , 偏移量为 Δp , 则可变形卷积的输出为:

$$Y(p_0) = \sum_{p_n \in R} W(p_n) X(p_0 + p_n + \Delta p_n) \quad (5)$$

式中: p_0 是输出特征图的位置; p_n 是卷积核的相对位置; Δp_n 是动态偏移量。

输出特征图通过区域建议网络 (region proposal network, RPN) 生成一组候选区域 (region of interest, RoI), 接着使用 RoI 池化层^[17] 和 RoI 对齐层对特征进行提取。定义分别为:

$$Y_{\text{RoIPool}} = \text{RoIPool}(X, R, \Delta R, H_{\text{out}}, W_{\text{out}}) \quad (6)$$

$$Y_{\text{RoIAlign}} = \text{RoIAlign}(X, R, \Delta R) \quad (7)$$

式中: R 是区域候选框; ΔR 是偏移量; $H_{\text{out}}, W_{\text{out}}$ 是输出特征图的高度和宽度。可变形 RoI 池化层通过引入动态偏

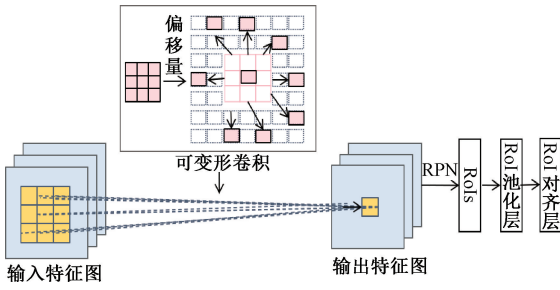


图 2 DCNv3 结构

Fig. 2 Structure diagram of DCNv3

移量, 实现在对象形状不规则时的高效特征提取, 从而提高了模型的鲁棒性。RoI 对齐层则通过双线性插值的方法解决了量化误差问题, 从而提高对目标的检测精度。

2) DS-CCFM 模块设计

CCFM^[18] 网络架构的核心思想是利用跨尺度的特征融合来增强对尺度层次信息的感知能力。CCFM 通过将来自不同层次的特征图进行融合, 传递网络中不同尺度的信息, 从而利用多分辨率的特征图在细节信息和低分辨率特征间搭建全局上下文。这样的跨尺度特征融合机制能够提高网络对于尺度变化的适应性。

YOLOv8 的采样器虽然包含了内嵌的特征选择, 它对目标各位置的采样仍有很大的局限性, 并未能充分利用跨尺度信息进行有效调整。为解决该问题, 提出了 DS, 以应对采样位置固定、局部特征不敏感等问题。DS 结构如图 3 所示。设输入特征图为 X , 其尺寸为 $H \times W \times C$, 首先定义一个标准的初始采样网络 S_{init} , 将 X 通过线性层生成每个位置的偏移量 ΔS :

$$h = \frac{-scale + 1}{2}, \dots, \frac{scale - 1}{2} \quad (8)$$

$$S_{init} = \frac{1}{scale} \cdot meshgrid(h, h) \quad (9)$$

$$\Delta S = 0.5\sigma \cdot linear(X) \cdot linear(X) \quad (10)$$

式中: $scale$ 决定了采样的密度; $meshgrid$ 函数用于生成一个二维网格坐标系; σ 是用于缩放偏移量的比例因子, 以确保偏移量不会过大。

X 通过卷积层和 Sigmoid 函数生成可学习的动态权重 W , 将 ΔS 进行权重调整后与 S_{init} 生成采样集 S_{weight} 。最后, 将 S_{weight} 进行归一化以匹配 $grid_sample$ 函数, 然后通过网格采样函数生成新的特征图 X' :

$$W = Sigmoid(Conv(X)) \quad (11)$$

$$S_{weight} = S_{init} + \Delta S \times W \quad (12)$$

$$S_{norm} = 2 \cdot \frac{S_{weight}}{[H, W]} - 1 \quad (13)$$

$$X' = grid_sample(X, S_{norm}) \quad (14)$$

其中, $grid_sample$ 是 Pytorch 的函数, 根据每个采样

点的位置, 通过双线性插值的方法获取输入特征图在该位置的值, 并重新组合生成新的特征图。

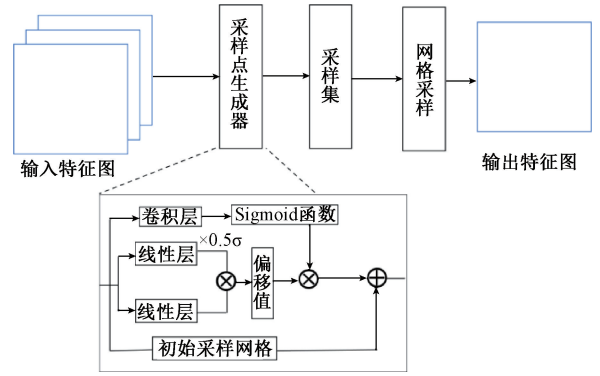


图 3 DS 结构原理

Fig. 3 Schematic diagram of DS

将 DS 嵌入到 CCFM 网络架构中, 使模块具备了可调谐采样点位置和范围的能力。动态偏移量能够灵活调整采样点的位置^[19], 适应输入特征的几何变化, 达成较优解和性能的提升。通过引入区域权重的调整机制, 该模块不仅能捕捉全局信息, 同时还加强了对局部特征的关注。DS 与 CCFM 的跨尺度特征融合机制协同作用, 进一步提升了模型在复杂场景中的表现。

3) P2-DynamicHead 模块设计

模块在原有检测层的基础上, 通过与网络的较浅层进行特征融合, 增加一个具有更高分辨率的检测层 P2。较高分辨率的特征图能够提供更多的空间信息, 能够更好地捕捉到小尺寸目标的细节, 有助于检测小尺寸的字符缺陷。同时, 在 4 个检测层中应用 DynamicHead 检测头的多个注意力机制和 RoI 池化层进行目标分类和位置回归。

DynamicHead 检测头^[20] 结构如图 4 所示, 将输入特征图视为一个三维的张量, $L \times S \times C$, 其中, L 是特征层级、 S 是特征图的宽高乘积 (HW)、 C 是通道数, 在每一个维度上分别采用尺度感知注意力机制 π_L 、空间感知注意力机制 π_S 和任务感知注意力机制 π_C , 并在 π_C 之前采用 RoI 池化层进行压缩, 该检测头可表示为:

$$W(\Gamma) = \pi_C(RoIPool(\pi_S(\pi_L(\Gamma) \cdot \Gamma) \cdot \Gamma)) \cdot \Gamma \quad (15)$$

其中, $\Gamma \in \mathbf{R}^{L \times S \times C}$, π_L 在特征层级维度上执行, 不同层级的特征图对应了不同的目标尺度, 在层级增加注意力, 可以增强目标检测的尺度感知能力。 π_S 聚焦于不同空间位置的判别能力, 其参考了可变形卷积的原理, 采用位置偏移以聚焦于有判别力的区域, 能够考虑到目标在不同视角下呈现出不同的形状、旋转和位置。 π_C 专注于通道 C , 可以动态开关特征通道以辅助不同任务。

P2-DynamicHead 模块通过增加小目标检测层 P2, 并

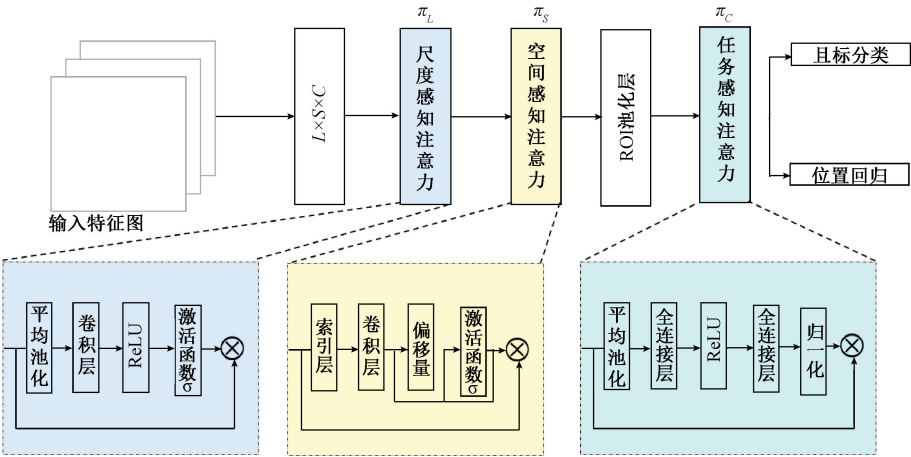


图 4 DynamicHead 结构

Fig. 4 Structure diagram of DynamicHead

在特征层级之间、空间位置之间以及输出通道内部结合多个自注意力机制,实现了尺度感知、空间感知和任务感知的统一,提高了目标检测头的性能。

2 实验部分

2.1 实验环境和模型超参数

为了验证所构建模型的有效性,以 Windows11 为操作系统,以 PyTorch 为深度学习框架,搭建了一个实验平台,所有的实验均在该平台上进行。实验平台的配置如表 1 所示。

表 1 实验工作站具体配置

环境参数	数值
操作系统	Windows11
GPU	NVIDIA GeForce RTX 4060(8 GB)
CUDA	12.1
CPU	Intel Core i7-12650H
RAM	32 GB
深度学习框架	Pytorch 2.1.0
编程语言	Python 3.8

同时,在所有实验的整个训练过程中应用了一致的训练参数和超参数。具体的训练参数和超参数如表 2 所示。

2.2 实验数据集

数据集来自某电池制造加工企业丝印字符检测机的缺陷图像,如图 5 所示,共计 1 618 张图像,采用 labeling 工具对图像进行标注。缺陷类型有 3 种,分别是污染(pollution)、缺印(omission)和模糊(unevenness),数量分别为 805、912、507,采用训练集:验证集:测试集=8:1:1 对数据集进行划分。

表 2 训练参数和超参数

Table 2 Training parameters and hyperparameters

	参数	数值
训练参数	图像大小	640×640
	批次大小	4
	轮数	300
	工作线程数	8
超参数	学习率	0.01
	动量	0.937
	权重衰减	0.000 5



图 5 常见的产品表面字符缺陷

Fig. 5 Common surface character defects of products

为了防止数据集规模小导致的过拟合问题,在所有实验中采用数据增强技术,通过色彩增强、图像翻转、图像缩放、图像拼接等多种数据增强方法,增强了模型的鲁棒性和泛化性,使其能够在避免过拟合的同时处理不同

场景下的检测任务,从而提高检测的准确性和可靠性,其效果如图 6 所示。

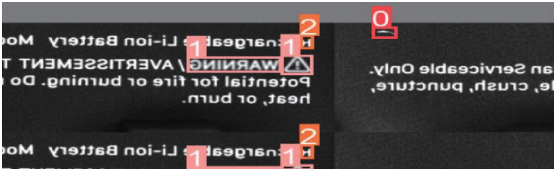


图 6 数据增强效果

Fig. 6 Data augmentation effects

2.3 实验评价指标

实验中,采用准确率 (precision, P)、召回率 (recall, R)、平均精度均值 (mean average precision, mAP)、模型参数量 (parameters, Params) 和帧率客观评价标准来评估训练后的字符缺陷检测模型的性能。 P 、 R 、 mAP 计算公式如式 (16)~(18) 所示。

$$P = \frac{TP}{TP + FP} \tag{16}$$

$$R = \frac{TP}{TP + FN} \tag{17}$$

$$mAP = \frac{1}{n} \sum_{k=1}^n \int_0^1 P_k dR_k \tag{18}$$

式中: TP 表示实际为正、预测为正的样本数量; FP 表示实际为负、预测为正的样本数量; FN 表示实际为正、预测为负的样本的数量; n 是数据集类别的数量。

2.4 实验及结果

1) 消融实验

为验证 3 种模块的有效性,进行多组消融实验。实验主要通过 在 YOLOv8n 中逐步加入改进模块,并测试其在字符缺陷数据集上的性能。图 7 所示为实验结果,表 3 为实验结果数据,其中√表示添加了该模块。

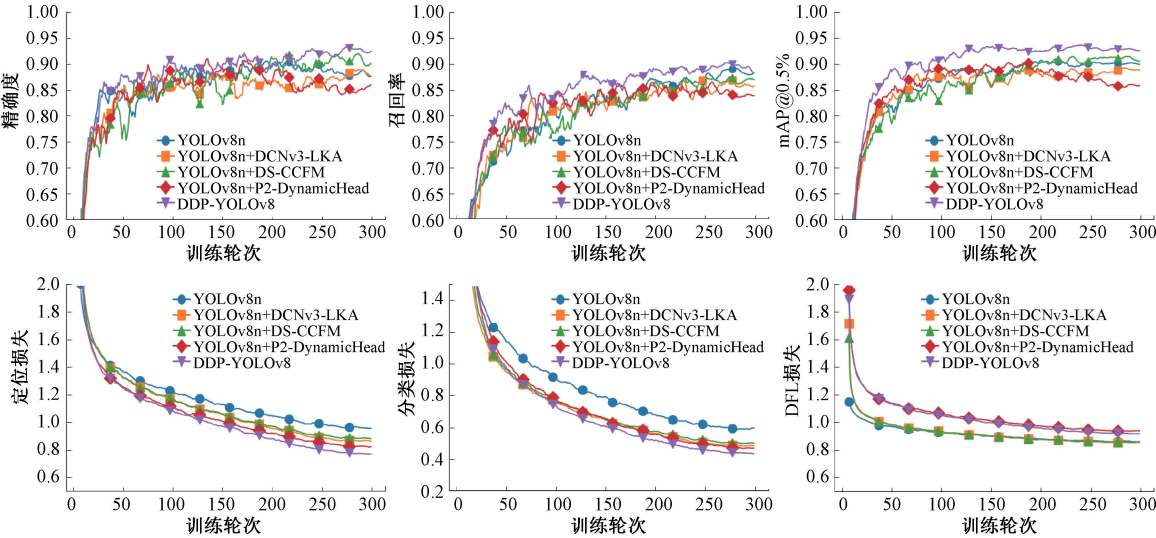


图 7 消融实验结果

Fig. 7 Ablation study results

表 3 消融实验结果数据
Table 3 Ablation study results

YOLOv8n	DCN-LKA	DCNv3-LKA	CCFM	DS-CCFM	DynamicHead	P2-DynamicHead	mAP@ 0. 5/%	Params/(×10 ⁶)	帧率/fps
√							90. 1	3. 01	147. 4
√	√						91. 2	2. 63	114. 3
√		√					91. 8	2. 59	114. 6
√			√				90. 8	1. 97	132. 5
√				√			91. 2	1. 96	156. 6
√					√		91. 3	2. 14	93. 7
√						√	92. 4	2. 87	84. 0
√		√		√			92. 0	2. 02	132. 7
√		√				√	93. 3	3. 39	89. 0
√			√	√		√	93. 0	2. 76	89. 1
√		√		√		√	94. 0	3. 28	85. 1

消融实验结果表明,加入 DCNv3-LKA 模块后,模型的 mAP@0.5 达到 91.8%,较加入 DCN-LKA 模块后的模型提高了 0.6%,这一结果表明,该模块在优化特征提取网络方面取得了良好的效果。在 DS-CCFM 模块的加入下,模型的 mAP@0.5 达到 91.2%,相较于仅采用 CCFM 架构的模型提高了 0.4%,这些结果验证了动态采样模块在提升多尺度特征融合方面的有效性。同时,模型的参数量较 YOLOv8n 下降约 35%,帧率提高了 9.2 fps,这表明 DS-CCFM 模块在提高检测精度的同时,还提高了推理速度。此外,加入 P2-DynamicHead 模块后,模型的 mAP@0.5 相比于加入 DynamicHead 模块提升了 0.9%,充分证明了在 4 个检测层中采用注意力机制和池化层策

略的有效性。在加入提出的 3 种模块后,模型的 mAP@0.5 达到了最高值 94.0%,比 YOLOv8n 提高了 3.9%。

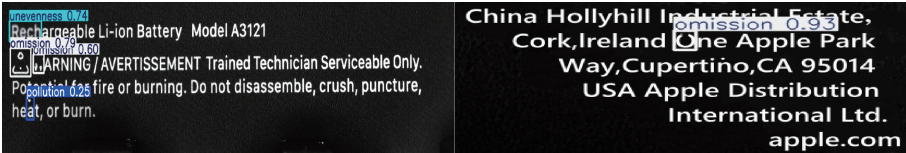
2) 对比实验

为了进一步验证所提出的改进模型在字符缺陷检测中的性能优势,选取了几种主流目标检测算法作为对照模型,分别为 RT-DETR^[18]、YOLOv3^[21]、YOLOv5n 和 YOLOv8n。此外,为了进一步评估改进模块的有效性,实验还添加了最新的基于 YOLOv8 改进的模块(CBAM^[8]、LSKA^[7]、GPFN^[9]、BIFPN^[7])进行对比实验,以检验 DDP-YOLOv8 模型的表现。实验结果数据如表 4 所示,图 8 所示为检测结果对比。

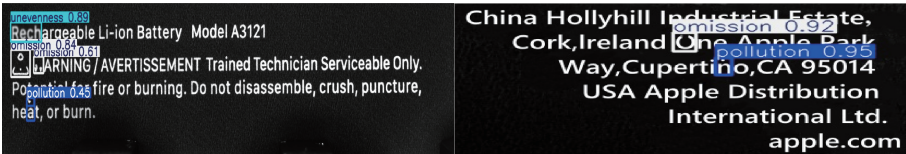
表 4 对比实验结果数据

Table 4 Comparative experimental results

模型	准确率/%	召回率/%	污染 Ap/%	模糊 Ap/%	缺印 Ap/%	mAP@0.5/%	Params/(×10 ⁶)	帧率/fps
YOLOv5n	77.6	92.0	85.9	86.3	85.4	85.9	2.51	181.6
YOLOv3	83.9	92.0	87.7	87.1	91.0	87.6	103.69	31.5
RT-DETR	85.8	96.0	85.7	84.3	97.2	89.1	32.81	38.9
YOLOv8n	85.5	94.0	89.8	94.8	85.8	90.1	3.01	147.4
YOLOv8n+CBAM	79.8	97.0	88.3	94.2	84.4	89.0	2.75	99.7
YOLOv8n+LSKA	88.8	95.0	88.4	94.8	87.5	90.2	3.08	106.3
YOLOv8n+GPFN	84.9	96.0	87.4	96.5	87.0	90.3	2.75	98.7
YOLOv8n+BIFPN	86.9	95.0	89.6	94.2	86.9	90.2	1.67	121.7
DDP-YOLOv8	86.3	97.0	94.4	95.1	92.5	94.0	3.28	85.1



(a) YOLOv8n检测结果置信度评分
(a) Confidence score diagram of YOLOv8n detection results



(b) DDP-YOLOv8检测结果置信度评分
(b) Confidence score diagram of DDP-YOLOv8n detection results

图 8 检测结果对比

Fig. 8 Comparison of detection results

由表 4 可知, YOLOv5n、YOLOv3、RT-DETR 和 YOLOv8n 模型的 mAP@0.5 分别为 85.9%、87.6%、89.1%和 90.1%,说明这些主流的目标检测算法在字符缺陷数据集上的检测效果并不理想。YOLOv8n+CBAM、YOLOv8n+LSKA、YOLOv8n+GPFN 和 YOLOv8n+BIFPN 模型的 mAP@0.5 分别为 89.0%、90.2%、90.3%和 90.2%,相比原始 YOLOv8n 的 90.1%并没有显著提升,说明这些模块并未显著提升模型在字符缺陷检测中的性

能。最终, DDP-YOLOv8 模型在 mAP@0.5 上达到了 94.0%,显著优于其他对比模型,说明了提出的改进模块的有效性,尤其在提升字符缺陷检测的精度方面表现突出。由图 8 可以看出,在不同段落字符的检测任务中, DDP-YOLOv8 均能有效识别字符缺陷,且在置信度评分上均优于 YOLOv8n,特别是在细小缺陷的检测任务中, DDP-YOLOv8 表现出更优的性能,进一步验证了模型的有效性和泛化能力。

此外,还将 DDP-YOLOv8 模型与最新发布的 YOLOv10^[22] 基础模型进行了对比,利用已有数据集进行测试。如表 5 所示,DDP-YOLOv8 模型的实验测试指标与 YOLOv10 基础模型的实验测试指标相当,进一步证明 DDP-YOLOv8 模型具有一定的先进性和实用性。

表 5 DDP-YOLOv8 与 YOLOv10n 性能对比

Table 5 Performance comparison between DDP-YOLOv8 and YOLOv10n

模型	准确率/ %	召回率/ %	mAP@ 0.5/%	Params/ ($\times 10^6$)	帧率/fps
DDP-YOLOv8	86.3	97.0	94.0	3.28	85.1
YOLOv10n	86.1	96.0	93.8	2.71	106.1

3 结 论

针对电池表面字符缺陷检测中存在的目标尺度微小、缺陷分布随机性强以及特征表达能力不足等关键问题,提出了一种全新的 DDP-YOLOv8 模型,从特征提取、特征融合和检测头 3 个核心环节进行了系统性优化,提升了模型在消费电池字符缺陷检测场景中的性能与实用性。首先,针对 YOLOv8 在特征提取过程中全局信息捕捉能力弱、权重分配不均的问题,创新性地设计了 DCNv3-LKA 模块,通过融合可变形卷积与大核注意力机制,实现了特征图空间权重的自适应分配,使模型能够更精准地聚焦于关键特征区域,mAP@0.5 提升了 1.7%。针对 YOLOv8 在字符缺陷检测中采样位置固定、局部特征敏感性差以及多尺度特征融合效率低的问题,提出了 DS-CCFM 模块,通过动态采样机制与可学习权重相结合,显著增强了模型对多尺度特征的适应能力,mAP@0.5 提高了 1.1%,且帧率提高了 9.2 fps,在保证精度的前提下进一步优化了计算效率。最后,针对 YOLOv8 检测头对小尺寸字符特征表达能力不足的问题,创新性地引入了 P2-DynamicHead 模块,通过增加高分辨率检测层并集成自适应注意力机制,有效提升了模型对小目标的检测能力,mAP@0.5 提升了 2.3%。实验结果表明,DDP-YOLOv8 模型在检测性能与计算效率之间实现了平衡,其 mAP@0.5 达到 94.0%,参数量为 3.28×10^6 ,帧率为 85.1 fps,说明该模型在满足工业实时检测需求的同时,显著提升了消费电池表面字符缺陷的检测精度。未来研究将致力于 DDP-YOLOv8 模型的持续优化,重点探索两个方向,通过将所提出的改进模块迁移至 YOLOv10 等先进架构,进一步提升模型的特征提取能力和检测精度;采用模型剪枝等轻量化技术,优化模型的计算效率与推理速度,为智能制造场景下的字符缺陷检测提供更高效、更精确的解决方案。

参考文献

[1] 熊继淙,邢旭朋,马军. 基于 Halcon 的锂电池标签表面缺陷检测[J]. 科技视界,2020(3):36-37.
XIONG J C, XING X M, MA J. Surface defect detection of lithium battery labels based on Halcon[J]. Science and Technology Vision, 2020(3): 36-37.

[2] 裴炜冬. 基于机器视觉的锂电池丝印 LOGO 缺陷检测系统研究与实现[D]. 广州:广东工业大学,2018.
PEI W D. Research and implementation of lithium battery screen printing LOGO defect detection system based on machine vision[D]. Guangzhou: Guangdong University of Technology, 2018.

[3] 卢丹,刘红,刘轩,等. 基于深度学习的汽车手柄字符缺陷检测[J]. 长春理工大学学报(自然科学版), 2022,45(6):52-58.
LU D, LIU H, LIU X, et al. Defect detection of automotive handle characters based on deep learning[J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2022, 45(6): 52-58.

[4] 赵卫东,陈凤,黄汉成,等. 基于深度学习的电表字符缺陷检测方法[J]. 安徽工业大学学报(自然科学版), 2021,38(1):76-81.
ZHAO W D, CHEN F, HUANG H CH, et al. A method for defect detection of electric meter characters based on deep learning [J]. Journal of Anhui University of Technology (Natural Science Edition), 2021, 38(1): 76-81.

[5] PENG J, ZHU W, LIANG Q, et al. Defect detection in code characters with complex backgrounds based on BBE[J]. Mathematical Biosciences and Engineering, 2021, 18(4): 3755-3780.

[6] JIANG P, ERGU D, LIU F, et al. A review of YOLO algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066-1073.

[7] 李冰,张易牧,魏乐涛,等. 基于改进 YOLOv8 的风机桨叶缺陷检测算法[J]. 电子测量技术, 2024, 47(13):89-99.
LI B, ZHANG Y M, WEI L T, et al. A defect detection algorithm for wind turbine blades based on improved YOLOv8 [J]. Electronic Measurement Technology, 2024,47 (13):89-99.

[8] 杜昌都,徐雷,陈俊,等. 基于轻量化 YOLOv8 网络的贴片芯片引脚缺陷检测[J]. 组合机床与自动化加工技术,2024(9):113-117.
DU CH D, XU L, CHEN J, et al. Defect detection of chip pin based on lightweight YOLOv8 network [J]. Modular Machine Tool & Automatic Manufacturing

- Technique, 2024 (9): 113-117.
- [9] 王安静,袁巨龙,朱勇建,等. 基于改进 YOLOv8s 的鼓形滚子表面缺陷检测算法[J]. 浙江大学学报(工学版), 2024, 58(2): 370-380, 387.
WANG AN J, YUAN J L, ZHU Y J, et al. A surface defect detection algorithm for drum rollers based on improved YOLOv8s[J]. Journal of Zhejiang University (Engineering Science Edition), 2024, 58(2): 370-380, 387.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv preprint arXiv:2004.10934, 2020.
- [11] 刘熹,陈晨,双丰. 基于改进 YOLOv7-tiny 的多种类绝缘子检测算法[J]. 仪器仪表学报, 2024, 45(9): 101-110.
LIU X, CHEN CH, SHUANG F. A multi-category insulator detection algorithm based on improved YOLOv7-tiny[J]. Chinese Journal of Scientific Instrument, 2024, 45(9): 101-110.
- [12] WANG J, BI L, MA X, et al. An efficient YOLOX-based method for photovoltaic cell defect detection [J]. Instrumentation, 2024, 11(2): 83-95.
- [13] AZAD R, NIGGEMEIER L, HÜTTEMANN M, et al. Beyond self-attention: Deformable large kernel attention for medical image segmentation[C]. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024: 1287-1297.
- [14] DING X, ZHANG X, HAN J, et al. Scaling up your kernels to 31x31: Revisiting large kernel design in CNNs [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11963-11975.
- [15] PAN X, GE C, LU R, et al. On the integration of self-attention and convolution[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 815-825.
- [16] CHEN F, WU F, XU J, et al. Adaptive deformable convolutional network[J]. Neurocomputing, 2021, 453: 853-864.
- [17] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9308-9316.
- [18] LYU W, XU S, ZHAO Y, et al. Detrs beat YOLOs on real-time object detection [J]. ArXiv preprint arXiv: 2304.08069, 2023.
- [19] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 6027-6037.
- [20] 曾勇杰,范必双,杨涯文,等. 改进 YOLOv8 算法在风机叶片缺陷检测上的应用[J]. 电子测量与仪器学报, 2024, 38(8): 26-35.
ZENG Y J, FAN B SH, YANG Y W, et al. Application of improved YOLOv8 algorithm in wind turbine blade defect detection[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(8): 26-35.
- [21] FARHADI A, REDMON J. YOLOv3: An incremental improvement [C]. Computer Vision and Pattern Recognition. Berlin/Heidelberg: Springer, 2018, 1804: 1-6.
- [22] WANG A, CHEN H, LIU L, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.

作者简介



李绪涛, 2022 年于华南农业大学获得学士学位, 现为广东工业大学硕士研究生, 主要研究方向为机器视觉。

E-mail: 1962452168@qq.com

Li Xutao received his B. Sc. degree from South China Agricultural University in 2022.

Now he is a M. Sc. candidate at Guangdong University of Technology. His main research interest includes machine vision.



邓耀华(通信作者), 2012 年于华南理工大学获得博士学位, 现为广东工业大学教授, 博士生导师, 中国仪器仪表学会精密机械分会委员, 主要研究方向为机器视觉和人工智能。

E-mail: dengyaohua@gdut.edu.cn

Deng Yaohua(Corresponding author) received his Ph. D. degree from South China University of Technology in 2012. He is now a professor and Ph. D. supervisor at Guangdong University of Technology, and a member of the Precision Machinery Branch of China Instrument Society. His main research interests include machine vision and artificial intelligence.