

本文引用格式: 曹海文,陈德平,杨丹妮,等.基于改进 YOLOv8 的瓶装产品褶皱检测模型[J].自动化与信息工程,2025,46(4):11-21.

CAO Haiwen, CHEN Deping, YANG Danni, et al. Wrinkle detection model for bottled products based on improved YOLOv8[J]. Automation & Information Engineering, 2025,46(4):11-21.

基于改进 YOLOv8 的瓶装产品褶皱检测模型*

曹海文^{1,3} 陈德平^{3,4} 杨丹妮³ 王楠³ 钟震宇³ 段二强²

(1.广东工业大学, 广东 广州 510006

2.佛山市云米电器科技有限公司, 广东 佛山 528308

3.广东省科学院智能制造研究所, 广东 广州 510070

4.汕头大学, 广东 汕头 515063)

摘要:针对瓶装产品褶皱的旋转角度随机、多尺度,以及瓶体背景对比度低等特性,提出基于改进 YOLOv8 的瓶装产品褶皱检测模型。该模型利用 Sobel 算子提取边缘特征,通过融合边缘特征与空间特征,增强低对比度区域的边缘特征提取能力;通过融合卷积与加性注意力机制,使 YOLOv8 模型能够同时提取图像的全局特征与局部特征,改进模型对不同尺度褶皱的特征融合能力,从而增强对旋转目标的鲁棒性;利用基于注意力的尺度内特征交互模块替换快速空间金字塔池化模块,增强骨干网络在同一尺度内的褶皱特征提取能力,解耦特征提取与多尺度特征融合任务;利用双向特征金字塔网络改进路径聚合网络,通过加权特征融合的方式动态调整特征权重,提升模型的多尺度特征融合能力;基于该模型构建了一套瓶装产品褶皱检测系统。实验结果表明,改进的 YOLOv8 模型的 AP50 为 82.3%,相较于 YOLOv8-OB 模型提升了 6.6%,能更好地完成瓶装产品褶皱的检测任务。经实际生产线验证,该瓶装产品褶皱检测系统的检测精度、推理速度和稳定性均符合工业应用要求。

关键词:瓶装产品褶皱检测;改进 YOLOv8;Sobel 算子;加性注意力机制;基于注意力的尺度内特征交互模块;双向特征金字塔网络

中图分类号: TP391

文献标志码: A

文章编号: 1674-2605(2025)04-0002-11

DOI: 10.12475/aie.20250402

开放获取

Wrinkle Detection Model for Bottled Products Based on Improved YOLOv8

CAO Haiwen^{1,3} CHEN Deping^{3,4} YANG Danni³

WANG Nan³ ZHONG Zhenyu³ DUAN Erqiang²

(1.Guangdong University of Technology, Guangzhou 510006, China

2.Foshan VIOMI Electrical Technology Co., Ltd., Foshan 528308, China

3.Institute of Intelligent Manufacturing, Guangdong Academy of Sciences, Guangzhou 510070, China

4.Shantou University, Shantou 515063, China)

Abstract: Aiming at the random rotation angles, multi-scale characteristics, and low contrast with the bottle background of wrinkles in bottled products, this paper proposes an improved YOLOv8-based wrinkle detection model for bottled products. The model utilizes the Sobel operator to extract edge features and enhances edge extraction capability in low-contrast regions by fusing edge features with spatial features. By integrating convolution and additive attention mechanisms, the YOLOv8 model can simultaneously extract global and local features of the image, improving its ability to feature fusion wrinkles of different scales and enhancing robustness against rotated targets. The fast spatial pyramid pooling module is replaced with an intra-scale feature interaction module based on attention, strengthening the backbone network's ability to extract wrinkle features within the same scale and decoupling

* 基金项目: 佛山市顺德区科技创新项目(2130218002519);
广东省科学院青年人才专项(2024GDASQNR-0327)。

feature extraction from multi-scale feature fusion tasks. The path aggregation network is improved using a bidirectional feature pyramid network, which dynamically adjusts feature weights through weighted feature fusion to enhance the model's feature fusion capability. Based on this model, a wrinkle detection system for bottled products is constructed. Experimental results show that the improved YOLOv8 model achieves an AP50 of 82.3%, a 6.6% improvement over the YOLOv8-OBb model, demonstrating superior performance in wrinkle detection for bottled products. Verified on actual production lines, the system meets industrial requirements in terms of detection accuracy, inference speed, and stability.

Keywords: bottled products wrinkle detection; improved YOLOv8; Sobel operator; additive attention mechanism; attention-based intra-scale feature interaction module; bidirectional feature pyramid network

0 引言

瓶装产品褶皱检测是指通过视觉检测技术对瓶体外观进行检测并识别表面褶皱,以确保包装质量完整,避免因褶皱导致内容物泄漏或污染。近年来,基于深度学习的视觉检测技术高速发展,已广泛应用于医疗影像分析、人脸识别和缺陷检测等领域^[1],成为褶皱检测的重要技术手段。

旋转目标检测技术主要用于倾斜文本^[2]和遥感目标^[3]的检测。瓶装产品褶皱的视觉特征与遥感目标相似,具有明显的旋转角度和多尺度特性,因此采用旋转目标检测技术可以有效提取褶皱的几何特征,从而提高褶皱的检测精度。旋转目标检测可分为单阶段和两阶段方法。R3Det^[4]、S²Anet^[5]等单阶段检测方法提取小目标特征能力有限,导致细小褶皱难以被有效检测。RRPN^[6]和 RoI Transformer^[7]等两阶段检测方法虽然检测精度较高,但依赖预设的角度锚框和复杂的候选区域生成网络;由于瓶装产品褶皱的旋转角度随机,角度锚框无法充分覆盖其所有几何变化,导致模型泛化性能受限;同时,复杂的候选区域生成网络结构也降低了模型的检测速度。相比之下,YOLOv8 模型不依赖先验锚框,通过特征提取网络和多尺度特征融合机制即可实现瓶装产品褶皱检测。但 YOLOv8 模型仅通过普通卷积进行特征提取,难以有效处理褶皱区域与瓶体包装高度融合引起的目标边界模糊问题;且当褶皱的旋转角度超出卷积感受野时,会丢失褶皱的局部特征,影响模型的检测精度。

为此,本文提出基于改进 YOLOv8 的瓶装产品褶皱检测模型。该模型利用图像边缘增强模块 (edge

image enhancement module, EIEM) 替代骨干网络 C2f 模块中的 Bottleneck 模块,以增强低对比度区域的轮廓感知能力,缓解因褶皱区域边界模糊导致的特征提取困难问题;引入加性注意力 (additive attention, AddA) 机制^[8],增强颈部网络 C2f 模块的多尺度检测能力和褶皱旋转变化的鲁棒性;利用基于注意力的尺度内特征交互 (attention-based intra-scale feature interaction, AIFI) 模块替换快速空间金字塔池化 (spatial pyramid pooling-fast, SPPF) 模块,以增强骨干网络在同一尺度内的褶皱特征提取能力,解耦特征提取与多尺度特征融合任务;利用双向特征金字塔网络 (bidirectional feature pyramid network, BiFPN)^[9]改进路径聚合网络 (path aggregation network, PANet)^[10],以动态调整特征权重,从而提升模型的多尺度特征融合能力。最后,基于该模型构建一套瓶装产品褶皱检测系统,可满足工业环境的实时检测需求。

1 YOLOv8 模型的改进

1.1 YOLOv8 模型

YOLOv8 模型主要由骨干网络 (Backbone)、颈部网络 (Neck) 和检测头 (Head) 组成,其结构如图 1 所示。

YOLOv8 模型通过骨干网络提取输入图像的特征;利用颈部网络进行多尺度特征融合;采用检测头实现目标分类和边界框回归任务。其中,C2f 模块^[12]主要用于多层次图像特征提取与融合;SPPF 模块主要通过不同大小的池化层得到不同尺寸的特征,以增强模型的多尺度特征融合能力。

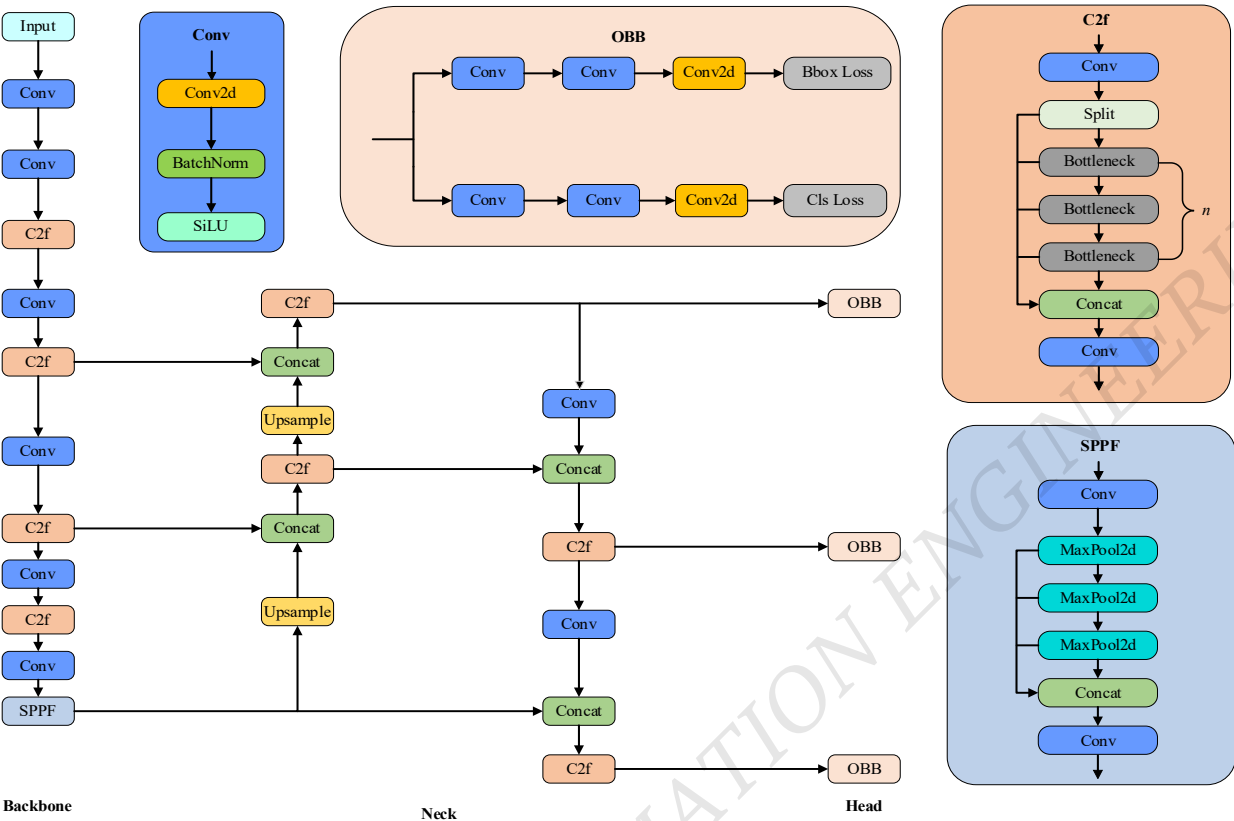


图 1 YOLOv8 模型结构图

YOLOv8 模型的整体设计高效合理，但仍存在以下问题：

- 1) C2f 模块通过普通卷积进行图像特征提取，而单一维度的图像特征难以有效分离褶皱与背景区域；同时，有限的感受野不利于捕获全局特征，导致 YOLOv8 模型在处理多尺度褶皱时性能受限；
- 2) SPPF 模块通过不同的池化层进行多尺度特征融合，其部分功能与颈部网络重叠，不仅增加了计算开销，还可能导致特征冗余传递，影响颈部网络的特征融合；
- 3) 颈部网络的多尺度特征融合能力不足，且融合方式固定，难以适应褶皱尺度多变、形态复杂的特性。

基于以上分析，本文根据褶皱检测任务的需求对 YOLOv8 模型进行改进。

1.2 基于信息融合的 C2f 模块改进

为了增强骨干网络 C2f 模块的特征提取能力，利

用 EIEM 替代 C2f 模块中的 Bottleneck 模块，采用 Sobel 算子提取边缘特征，以增强褶皱边界的可见性。EIEM 结构如图 2 所示。

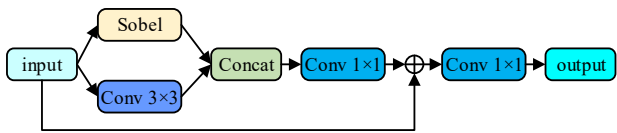


图 2 EIEM 结构图

首先，通过 Sobel 算子和 3×3 卷积分支分别提取输入图像的边缘特征和空间特征；然后，利用 concat 模块融合边缘特征与空间特征；接着，通过 1×1 卷积降低通道数；最后，经残差连接和 1×1 卷积处理后，输出兼具边缘细节和空间结构的综合特征表示。

Sobel 算子一般用于图像处理中的边缘检测任务，其通过计算图像像素值的梯度来检测图像边缘，即利用 2 个 3×3 卷积核 G_x 和 G_y 分别计算图像在水平和垂直方向上的梯度：

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad G_y = \begin{bmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (1)$$

尽管瓶装产品褶皱与瓶体背景的对比度较低，但其边缘轮廓通常存在一定的梯度变化。本文通过融合边缘特征与空间特征，增强了低对比度区域的可见性，缓解了因褶皱边界模糊导致的特征提取困难问题。

1.3 基于 AddA 机制的 C2f 模块改进

针对瓶装产品褶皱尺度不一、旋转角度随机的特性，本文通过引入 AddA 机制来增强颈部网络 C2f 模块的多尺度检测能力和褶皱旋转变化的鲁棒性。AddA 机制在视觉变换器（vision transformer, ViT）^[11] 的基础上简化了特征交互方式，并在保留全局特征获取能力的同时提升了计算效率，其流程如图 3 所示。其中， Q 、 K 、 V 分别表示 Query（查询）、Key（键）和 Value（值），均由输入经过线性变换而来。

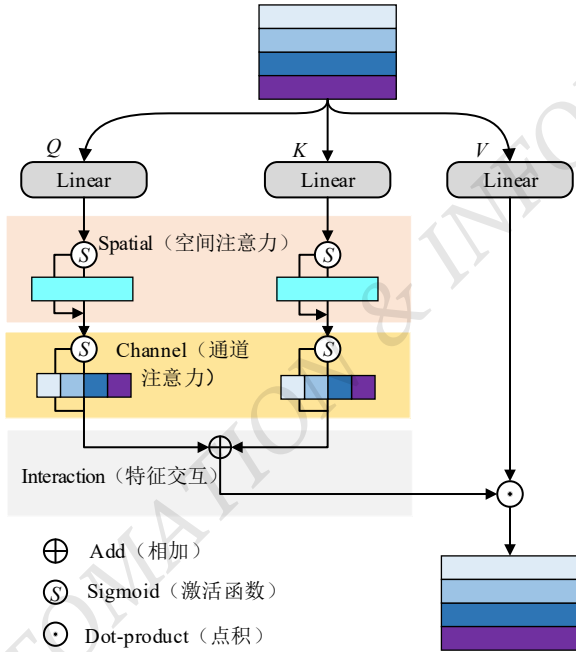


图 3 AddA 机制流程图

首先， Q 和 K 分别在空间维度和通道维度上进行注意力操作，提取特征图的通道特征和空间特征；然后，通过相加的方式实现 Q 和 K 之间的特征交互（见公式(2)），生成注意力权重矩阵；最后，利用注意力

权重矩阵对 V 进行加权，从而完成注意力操作。

$$Sim(Q, K) = \Phi(Q) + \Phi(K) \quad \text{s.t.} \quad \Phi(Q) = C(S(Q)) \quad (2)$$

式中： $Sim(Q, K)$ 为 Q 和 K 之间的相似度， $\Phi(\cdot)$ 为映射操作，包括空间注意力 $S(\cdot)$ 和基于 Sigmoid 的通道注意力 $C(\cdot)$ 。

空间注意力 $S(\cdot)$ 的处理流程与深度可分离卷积^[12] 基本相同。首先，利用 depthwise 卷积独立提取每个通道的局部空间特征（见图 4(a)）；然后，通过 pointwise 卷积对通道维度进行压缩，生成空间注意力加权图并进行加权处理（见图 4(b)）。

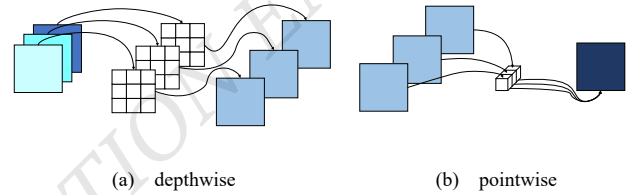


图 4 深度可分离卷积

基于 Sigmoid 的通道注意力 $C(\cdot)$ 的结构参考了 SE-Net^[13]，其处理流程如图 5 所示。

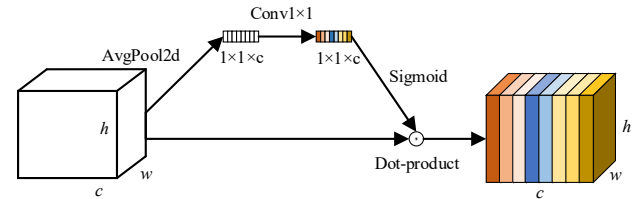


图 5 基于 Sigmoid 的通道注意力 $C(\cdot)$ 处理流程图

首先，通过全局平均池化（AvgPool）操作在通道维度上提取输入特征；然后，利用 Sigmoid 函数将卷积输出值限制在 0~1 之间，从而得到各通道权重；最后，在通道维度上将通道权重与输入进行加权处理，完成基于 Sigmoid 的通道注意力 $C(\cdot)$ 操作。

综上所述，通过卷积与 AddA 机制的融合，C2f 模块能够同时提取图像的全局特征与局部特征，增强对不同尺度褶皱特征的融合能力；且 AddA 机制通过对长程依赖关系的捕捉，有效补偿了因褶皱旋转导致的局部特征丢失问题，从而提升了 YOLOv8 模型对瓶装产品褶皱旋转变化的鲁棒性。

1.4 AIFI 模块

本文利用实时检测 Transformer (real-time detection Transformer, RT-DETR) 模型^[14]中的 AIFI 模块替换 SPPF 模块。AIFI 模块结构如图 6 所示。其中, S3、S4 和 S5 分别为特征提取网络中的最后 3 幅特征图; F5 为完成注意力操作并重构后的特征图。

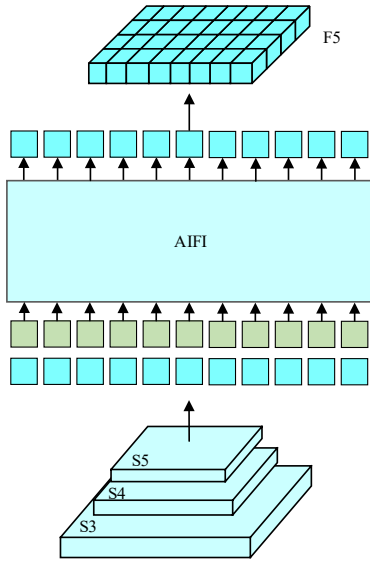


图 6 AIFI 模块结构图

首先, 骨干网络最后一层的输出被展平 (flatten) 为一维向量; 然后, 利用 AIFI 模块中的注意力操作 (attention) 对该一维向量进行特征交互; 最后, 将被展平的一维向量还原 (reshape) 为二维向量。AIFI 模块的整体流程可归纳为

$$Q = K = V = F_{\text{flatten}}(x), \quad (3)$$

$$O_{\text{AIFI}} = R_{\text{reshape}}(A_{\text{AIFI}}(Q, K, V))$$

AIFI 模块的功能与 SPPF 模块相反, 其主要通过注意力机制在同一尺度内进行特征交互。深层特征图具有更高级的语义信息和更大的感受野, 在深层特征图中进行自注意力操作, 能够在减少计算量的同时, 更好地提取全局特征。

利用 AIFI 模块替代 SPPF 模块, 使特征提取网络更专注于单一尺度的特征提取; 同时, AIFI 模块可对全局特征进行提取, 以弥补卷积感受野受限的问题,

从而增强 YOLOv8 模型对褶皱旋转形变的适应能力。

1.5 BiFPN

颈部网络中的 PANet 是在特征金字塔网络^[15]的基础上增加了一个方向融合, 实现了深层特征与浅层特征之间的互补。然而, 其本质上仍是一个单向的特征金字塔网络, 特征融合能力有限。同时, 仅通过 concat 操作进行特征融合, 缺乏对不同层级特征贡献的动态调整。为此, 本文利用 BiFPN 对 PANet 进行改进。PANet 与 BiFPN 的结构分别如图 7、8 所示。

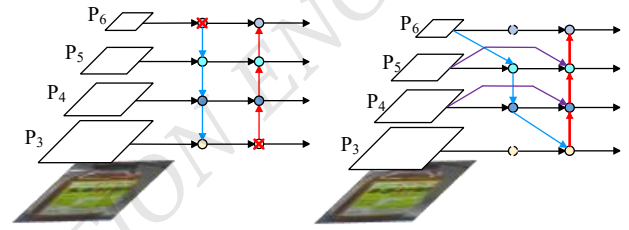


图 7 PANet 结构图

图 8 BiFPN 结构图

图 8 中, BiFPN 先删除了只有单向输入而没有进行特征融合的冗余节点; 再在中间层的原始输入和输出节点之间添加一条跨层连接路径, 使原始输入特征能够在自上向下和自下向上的两条路径中同时传播, 从而实现真正双向特征连接。

BiFPN 还在特征融合方式上进行了优化, 其采用加权特征融合的方式对特征图的贡献进行动态调整:

$$O_{\text{BiFPN}} = \sum_i \frac{w_i}{\varepsilon + \sum_j w_j} \cdot I_i \quad (4)$$

式中: O_{BiFPN} 为融合后 BiFPN 模块的输出特征, I_i 为第 i 个输入特征, w_i 为第 i 个输入特征对应的可学习权重, ε 用于避免分母为零, $\sum_j w_j$ 为所有权重之和。

YOLOv8 模型通过加权特征融合的方式动态调整特征权重, 能够突出重要特征表达, 抑制次要特征表达, 从而提升特征融合能力。

1.6 改进的 YOLOv8 模型

改进的 YOLOv8 模型结构如图 9 所示。

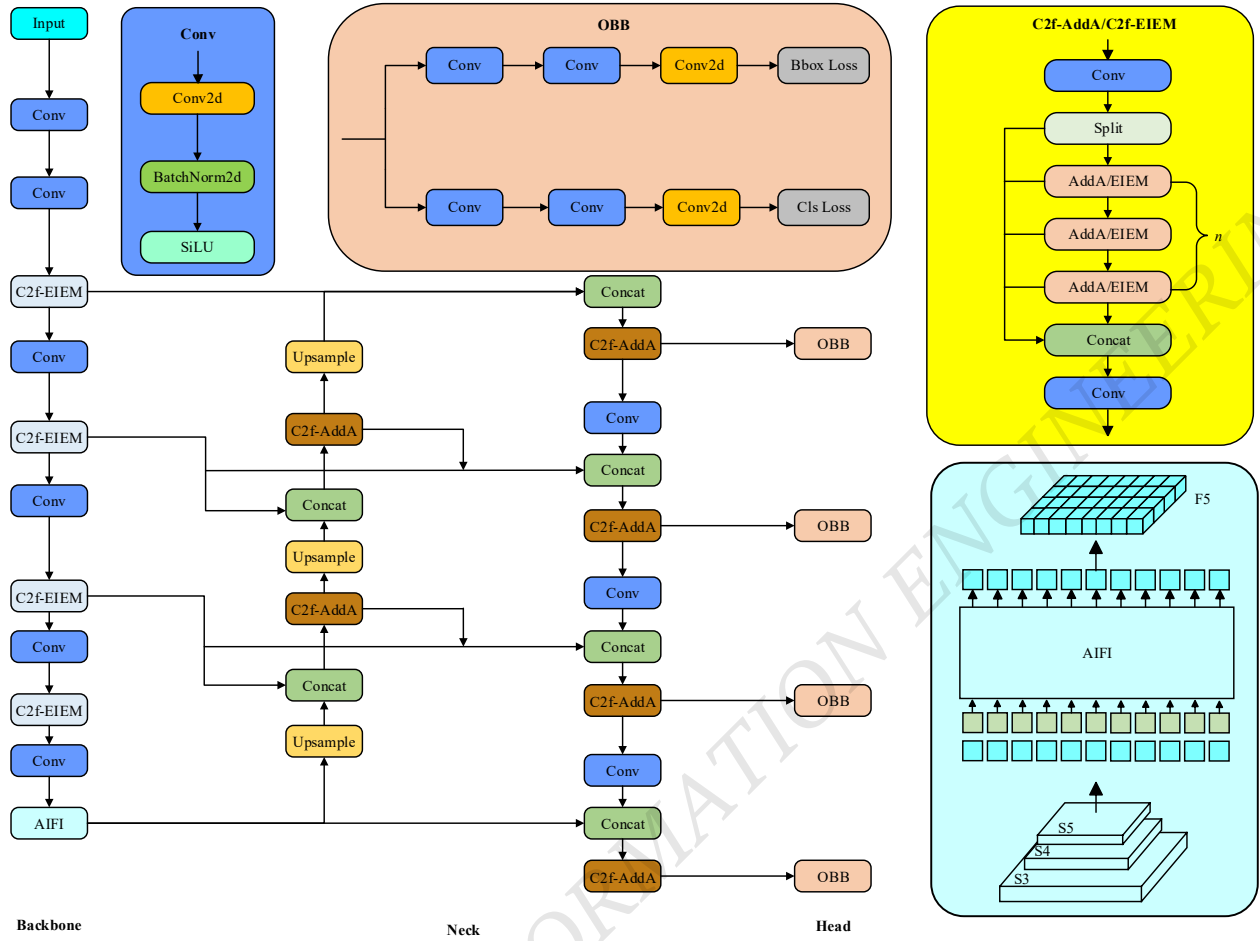


图9 改进的YOLOv8模型结构图

2 实验

为了评估基于改进 YOLOv8 的瓶装产品褶皱检测模型的性能,本实验在自建的瓶装产品褶皱数据集上进行训练和测试。该数据集包括 5 个品牌、8 种不同外观的 2942 幅瓶子图像和 12727 个标注实例,并按 9:1 的比例划分为训练集和测试集。

2.1 实验环境与关键参数设置

软件环境:操作系统 Ubuntu 18.04, Python 3.8, PyTorch 2.1.0, CUDA 11.8, MMDetection 2.25.1, MMRotate 0.3.3。

硬件环境:NVIDIA DGX-1 平台,Intel(R) Xeon(R) E5-2698 v4 @ 2.20 GHz CPU, 8 × Tesla P100-SXM2 16 GB GPU。

关键参数设置:为了保证对比模型的原始性能,

训练时遵循模型的默认参数设置;对比模型的学习率为 0.0025,动量为 0.9,权重衰减为 0.0001,优化器为 SGD,批尺寸为 2,在有预训练权重下训练轮次为 100;YOLOv8n-OBB 模型的学习率为 0.01,动量为 0.937,权重衰减为 0.0005,优化器为 SGD,批尺寸为 32,在无预训练权重下训练轮次为 300。

2.2 评估指标

本实验评估指标包括参数量 (params)、浮点计算量 (floating-point operations, FLOPs)、AP50、帧每秒 (frame per second, FPS)。其中, params、FLOPs 分别用于衡量模型的大小和计算复杂度;AP50 是指交并比为 0.5 时某一类别的平均精度 (averager precision, AP),用于衡量模型的检测精度;FPS 用于衡量模型的检测速度。

2.3 实验结果与分析

2.3.1 对比实验

将本文提出的改进 YOLOv8 模型与单阶段旋转目标检测模型（RetinaNet-O^[16]、R3Det、S²ANet、YOLOv8n-ORB）、两阶段旋转目标检测模型（Faster R-CNN-O、RoI Transformer、Gliding Vertex^[17]、Oriented R-CNN^[18]）在自建的瓶装产品褶皱数据集上进行对比实验，结果如表 1 所示。

表 1 9 种旋转目标检测模型对褶皱的检测结果

模型	FLOPs/G	params/M	AP50/%	FPS/ (f/s)
RetinaNet-O	65.63	36.15	60.7	30.2
R3Det	102.88	41.60	73.0	22.0
S ² Anet	61.35	38.54	74.5	25.6
RoI Transformer	89.49	55.03	75.9	16.3

续表

模型	FLOPs/G	params/M	AP50/%	FPS/ (f/s)
Gliding Vertex	75.59	41.13	75.2	21.9
Oriented R-CNN	75.63	41.13	75.6	21.8
YOLOv8n-ORB	8.30	3.10	75.7	135.6
本文模型	12.30	2.75	82.3	38.2

由表 1 可知：本文模型的参数量最少，检测精度最高，计算复杂度、检测速度仅次于 YOLOv8-ORB 模型，表明本文模型的褶皱检测能力得到明显改善，性能优于其他对比模型。

9 种旋转目标检测模型对 3 种瓶装产品（奶茶、柠檬茶、东鹏特饮）的褶皱检测效果如图 10 所示，其中第一列为标注。



图 10 各模型褶皱检测效果对比图

由图 10 可知：RetinaNet-O 模型的褶皱检测效果最差，奶茶瓶、柠檬茶瓶上存在多个漏标；R3Det 模型虽在奶茶瓶上存在漏标，但整体检测效果较好；RoI Transformer 等其他对比模型在 3 种瓶装产品上均存在不同程度的误标，以及预测框与标注框重合度较低的问题；本文模型在奶茶瓶和柠檬茶瓶上的预测框与标注框的重合度最高，表明本文模型对瓶装产品褶皱的检测性能有明显改善。

2.3.2 消融实验

为分析各模块之间的配合情况，以及对 YOLOv8 模型的改进贡献，通过在基线模型中逐步添加改进模块的方式进行消融实验，结果如表 2 所示。

表 2 消融实验结果

模块				FLOPs/ GB	params/ MB	AP50/ %	FPS/ (f/s)
BiFPN	AIFI	C2f- AddA	C2f- EIEM				
—	—	—	—	8.30	3.10	75.7	135.6
✓	—	—	—	13.0	3.04	79.1	98.0
✓	✓	—	—	12.9	2.96	80.7	66.5
✓	✓	✓	—	12.4	2.85	81.0	30.8
✓	✓	✓	✓	12.3	2.75	82.3	38.2

由表 2 可知：与基线模型相比，采用 BiFPN 作为多尺度特征融合颈部网络后，AP50 提高了 3.4%，但由于 BiFPN 额外增加了 1 个检测头，导致 YOLOv8 模型的检测速度下降；AIFI 模块仅处理分辨率较低的特征图，不仅没有显著增加模型的参数量和浮点计算量，还将 AP50 提高了 1.6%；虽然 AddA 机制相较于传统注意力机制降低了浮点计算量，但在所有 C2f 模块中使用，明显拖慢了模型的检测速度；将骨干网络中的 C2f 替换为 C2f-EIEM，在提升模型检测速度的同时，AP50 提高了 1.3%，表明 Sobel 算子能够有效提升模型对褶皱边缘特征的提取能力。

热力图通过颜色变化可以直观地展示各改进模块加入后，YOLOv8 模型在瓶装产品褶皱检测任务上的性能变化。消融实验热力图如图 11 所示。

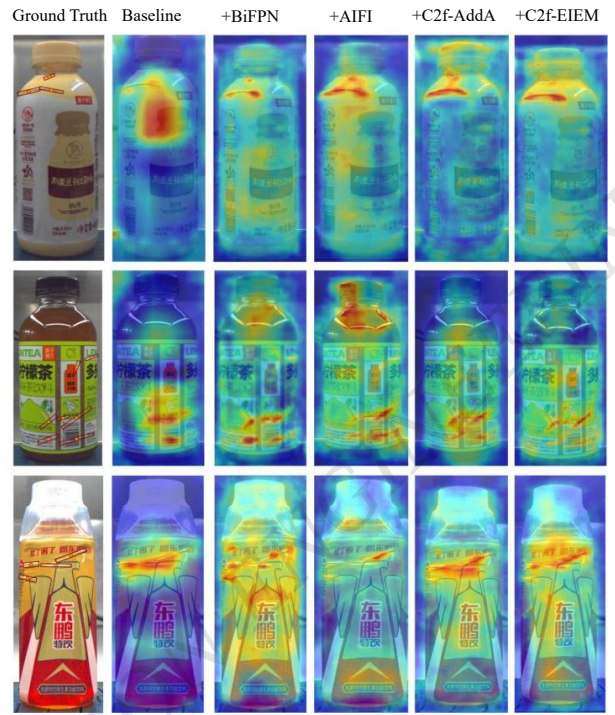


图 11 消融实验热力图

由图 11 可知：基线模型的关注区域较为分散，只能大致定位褶皱位置，表明在背景复杂或目标尺度较小的情况下，其特征提取能力有限，难以准确定位目标；加入 BiFPN 后，YOLOv8 模型的关注区域集中在褶皱上，表明多尺度特征融合改进能够增强模型对褶皱尺度变化的适应性；引入 AddA 机制后，褶皱区域的颜色明显变深，表明 YOLOv8 模型在抑制背景噪声的同时聚焦关键区域；Sobel 算子的加入增强了 YOLOv8 模型的边缘特征提取能力，使褶皱轮廓更加清晰。综上所述，随着各改进模块的依次加入，YOLOv8 模型的褶皱检测性能得到了逐步提升。

3 部署与系统优化

3.1 部署环境与系统架构

为实现基于改进 YOLOv8 的瓶装产品褶皱检测模型的高效部署，以 Jetson Orin Nano(4 GB)嵌入式组件为深度学习核心计算平台，开发了一套低功耗、高性能的瓶装产品褶皱检测系统，以满足工业环境的实

时检测需求。瓶装产品褶皱检测系统架构示意图如图 12 所示。

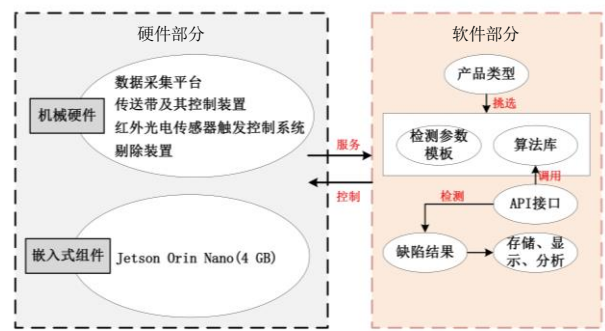


图 12 瓶装产品褶皱检测系统架构示意图

3.1.1 硬件部分

瓶装产品褶皱检测系统硬件部分主要由数据采集平台、传送带及其控制装置、红外光电传感器触发控制系统、剔除装置和嵌入式组件 Jetson Orin Nano（4 GB）等组成。

数据采集平台主要包括 4 台 Huaray MV-13CG-E 工业相机，负责对瓶装产品进行拍摄，其中 2 台相机正对瓶装产品，另外 2 台相机以 30° 倾角安装（见图 13），以确保拍摄角度覆盖瓶装产品的各个表面。每台相机均配备 1 个 LED 条形光源，以保证在不同角度和光照条件下均能清晰采集瓶装产品的表面信息。

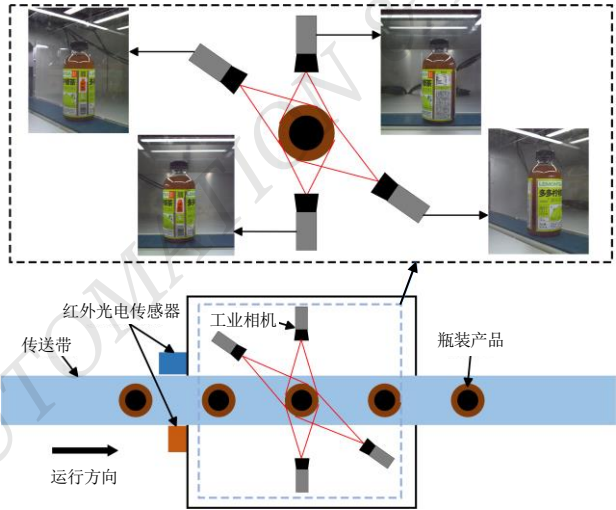


图 13 数据采集平台示意图

瓶装产品褶皱检测系统工作流程如下：1）当瓶

装产品通过数据采集平台时，会遮挡红外光电传感器发出的光信号，传感器向工业相机输出触发信号；2）工业相机接收到触发信号并经过一段延时后，对瓶装产品进行全表面图像采集；3）采集的瓶装产品图像数据输入至嵌入式组件，通过基于改进 YOLOv8 的瓶装产品褶皱检测模型进行瓶装产品褶皱检测，并将检测结果发送至 PLC；4）若检测到瓶装产品存在褶皱缺陷，PLC 发出剔除信号，剔除装置将缺陷瓶装产品从生产线中剔除到瑕疵品回收箱内，避免其继续流入后续工序，至此完成瓶装产品褶皱的完整检测。

3.1.2 软件部分

软件前端采用基于 Flask 开发的 Web 可视化界面，通过消息队列遥测传输（message queuing telemetry transport, MQTT）协议、WiFi 6 IoT 芯片实现检测数据的高效传输及实时显示，利用 MySQL 数据库存储检测结果，便于后期数据分析与回溯。

软件后端以本文提出的基于改进 YOLOv8 的瓶装产品褶皱检测模型为核心，依托嵌入式组件 Jetson Orin Nano（4 GB）在本地进行模型推理，以降低对云端的依赖，从而提高系统的响应速度。

3.2 系统优化

为了进一步提高基于改进 YOLOv8 的瓶装产品褶皱检测模型在边缘设备上的推理速度，采用以下优化策略：

- 1）模型量化，将基于改进 YOLOv8 的瓶装产品褶皱检测模型的权重数据格式由 FP32 转换为 INT8，以提高检测速度并降低内存占用；
- 2）TensorRT 加速，利用 NVIDIA TensorRT 优化基于改进 YOLOv8 的瓶装产品褶皱检测模型，将 PyTorch 模型(.pt)转换为 TensorRT 引擎(.engine)，以提高推理吞吐量并降低延迟，从而提升检测速度；
- 3）多线程并行处理，利用嵌入式组件 Jetson Orin Nano（4 GB）实现多线程并行计算，以提高计算资源利用率；
- 4）缓存优化与数据预处理，采用图像缓存机制

减少重复的 I/O 访问, 并进行图像归一化与增强, 提高褶皱检测的稳定性。

4 结论

本文结合瓶装产品褶皱的视觉特征, 对 YOLOv8 模型的结构进行改进, 以提高其褶皱检测性能。首先, 在骨干网络的 C2f 模块中引入 Sobel 算子, 通过融合边缘特征和空间特征, 增强模型对褶皱边缘的检测能力; 然后, 在颈部网络的 C2f 模块中引入了 AddA 机制, 提升了 C2f 模块的多尺度特征融合能力和对旋转目标的鲁棒性; 接着, 利用 AIFI 模块替代 SPPF 模块, 解耦特征提取与多尺度特征融合任务, 使特征提取网络专注于单一尺寸的特征提取; 最后, 引入 BiFPN 改进颈部网络结构, 使特征融合方式更为合理。实验结果表明, 在模型参数量和浮点计算量几乎不变的情况下, 改进的 YOLOv8 模型的检测效果明显优于 YOLOv8-OBb 模型, 能够有效完成瓶装产品褶皱检测任务。但 YOLOv8 模型^[19-20]引入不同改进模块后, 其浮点计算量上升, 检测速度下降, 不利于模型的实际部署。后续工作将对模型进行轻量化改进, 以提升模型的检测速度。

基于改进 YOLOv8 的瓶装产品褶皱检测模型构建了一套瓶装产品褶皱检测系统。经实际生产线验证, 该系统在检测精度与实时性方面满足工业质检要求, 具备规模化部署条件, 为包装缺陷自动化检测提供了一种新的可复用解决方案。

©The author(s) 2024. This is an open access article under the CC BY-NC-ND 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

参考文献

- [1] ZOU Z, CHEN K, SHI Z, et al. Object detection in 20 years: A survey[J]. Proceedings of the IEEE, 2023,111(3):257-276.
- [2] LIAO M, SHI B, BAI X. Textboxes++: A single-shot oriented scene text detector[J]. IEEE Transactions on Image Processing, 2018, 27(8): 3676-3690.
- [3] AZIMI S M, VIG E, BAHMANYAR R, et al. Towards multi-class object detection in unconstrained remote sensing imagery [C]//Asian Conference on Computer Vision. Cham: Springer International Publishing, 2018:150-165.
- [4] YANG X, YAN J, FENG Z, et al. R3Det: Refined single-stage detector with feature refinement for rotating object[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2021,35(4):3163-3171.
- [5] HAN J, DING J, LI J, et al. Align deep features for oriented object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60:1-11.
- [6] MA J, SHAO W, YE H, et al. Arbitrary-oriented scene text detection via rotation proposals[J]. IEEE Transactions on Multimedia, 2018, 20(11):3111-3122.
- [7] DING J, XUE N, LONG Y, et al. Learning RoI transformer for detecting oriented objects in aerial images[J]. arXiv preprint arXiv:1812.00155, 2018.
- [8] ZHANG T, LI L, ZHOU Y, et al. Cas-vit: Convolutional additive self-attention vision transformers for efficient mobile applications[J]. arXiv preprint arXiv:2408.03703, 2024.
- [9] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [10] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [11] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [12] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [14] ZHAO Y, LV W, XU S, et al. Dets beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

- [15] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [17] XU Y, FU M, WANG Q, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(4): 1452-1459.
- [18] XIE X, CHENG G, WANG J, et al. Oriented R-CNN for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 3520-3529.
- [19] 赵朝,岳龙旺.基于 YOLOv8 改进的机器人多物体抓取检测算法[J].机电工程技术,2025,54(2):132-137.
- [20] 王众玄,邹光明,顾浩文,等.基于改进 YOLOv8 的堆叠零件实例分割研究[J].机床与液压,2024,52(19):9-16.

作者简介:

曹海文,男,1999年生,硕士研究生,主要研究方向:机器视觉。E-mail: a13710660939@163.com
陈德平,男,1999年生,硕士研究生,主要研究方向:机器视觉。E-mail: cdp152418@126.com
杨丹妮,女,1998年生,硕士研究生,主要研究方向:机器视觉应用、项目管理。E-mail: dn.yang@giim.ac.cn
王楠,女,1989年生,博士研究生,高级工程师,主要研究方向:机器视觉。E-mail: n.wang@giim.ac.cn
钟震宇,男,1971年生,博士研究生,研究员,主要研究方向:机器视觉。E-mail: zy.zhong@giim.ac.cn
段二强(通信作者),男,1980年生,本科,主要研究方向:物联网技术。E-mail: duaneq@viomi.com

《自动化与信息工程》第十届编辑委员会

顾问: 朱学峰 程韬波

主编: 陈小奇 周雪峰

副主编: 刘桂雄(常务副主编) 谢巍 张正涛 邓耀华 邹涛

委员(以姓氏笔画为序):

丁其川 王迪 王立辉 王佐勋 王清辉 方灶军 尹晓燕 邓耀华 卢旭 叶廷东 冯颖
毕贵军 朱海飞 朱萍玉 伍锦荣 刘丹 刘伟 刘乙奇 刘国良 刘振忠 刘桂雄 江励
苏晓杰 杜胜 杜玉晓 李震 李杰浩 吴伟林 吴俊君 吴黎明 何力 邹涛 邹湘军
冷杰武 张正涛 张华伟 陈小奇 陈启愉 陈浩耀 陈海永 林若波 周松斌 周雪峰 赵志甲
胡耀华 钟勇 洪晓斌 姚屏 姚锡凡 徐大波 徐向荣 徐智浩 黄凯 黄从智 黄创绵
黄国健 黄宝山 曹东 曹志兴 符长虹 梁桥康 韩鹏 景奉水 温淑焕 谢巍 蔡延光
管贻生 魏世民