

基于改进 YOLOv11 的初加工小龙虾外观品质无损检测方法

王许佳 王淑青 胡浩宇

(湖北工业大学电气与电子工程学院, 湖北 武汉 430070)

摘要: [目的] 实现初加工小龙虾外观品质的快速、无损在线检测, 提升复杂场景下的检测精度与实时性。[方法] 构建面向初加工产线复杂场景下的小龙虾外观品质数据集, 提出一种基于 YOLOv11n 的小龙虾外观品质检测算法 YOLO-CQ。引入 DualConv 改进 C3k2 模块, 在优化计算资源分配的同时增强模型的特征提取能力。在 C2PSA 模块中引入可变形注意力机制 DAT, 使模型能自适应聚焦关键区域。引入基于 IoU 动态加权的 SlideLoss 优化分类损失函数, 以平衡不同难度样本在模型训练过程中的梯度贡献。[结果] 改进模型 YOLO-CQ 的精确率、召回率和 mAP@0.5 分别为 86.6%, 84.3%, 91.3%, 较原始 YOLOv11n 分别提升 2.5%, 2.3%, 3.4%, 推理速度达到 211.5 帧/s。[结论] 所提 YOLO-CQ 模型在保持模型轻量化与实时性的同时可显著提升初加工小龙虾外观品质检测精度, 满足初加工产线在线分级需求。
关键词: 小龙虾; 初加工产线; 深度学习; 注意力机制; 外观品质检测; 目标检测; YOLOv11

A non-destructive appearance quality detection method for primary-processed crayfish based on improved YOLOv11

Wang Xujia Wang Shuqing Hu Haoyu

(School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, Hubei 430070, China)

Abstract: [Objective] To achieve rapid and non-destructive online detection of the appearance quality of crayfish during primary processing and to improve detection accuracy and real-time performance in complex scenarios. [Methods] A dataset for crayfish appearance quality is constructed for complex scenarios on primary-processing production lines, and a crayfish appearance quality detection algorithm YOLO-CQ based on YOLOv11n is proposed. DualConv is introduced to improve the C3k2 module, enhancing the model's feature extraction capability while optimizing the allocation of computational resources. Deformable attention (DAT) is integrated into the C2PSA module, enabling the model to adaptively focus on key regions. In addition, an IoU-based dynamically weighted SlideLoss is adopted to optimize the classification loss, thereby balancing the gradient contributions of samples with different difficulty levels during training. [Results] YOLO-CQ achieves the precision of 86.6%, the recall of 84.3%, and the mAP@0.5 of 91.3%, which represent improvements of 2.5%, 2.3%, and 3.4% over those of YOLOv11n, respectively. YOLO-CQ reaches an inference speed of 211.5 FPS. [Conclusion] YOLO-CQ significantly improves the detection accuracy of crayfish appearance quality while maintaining model lightness and real-time performance, meeting the requirements of online grading on primary-processing production lines.

Keywords: crayfish; primary processing line; deep learning; attention mechanism; appearance quality detection; object detection; YOLOv11

近年来,小龙虾产业蓬勃发展,已成为中国渔业产业需求提升与产品形态多样化,外观品质逐渐成为影响产业链最完整、综合产值最高的产业之一^[1]。随着小龙虾消费产品定级与市场价值的关键因素。目前,小龙虾分拣加工

基金项目:国家自然科学基金(编号:62306107)

通信作者:王淑青(1969—),女,湖北工业大学教授,博士。E-mail:755163916@qq.com

收稿日期:2025-11-16 改回日期:2026-04-28

引用格式:王许佳,王淑青,胡浩宇. 基于改进 YOLOv11 的初加工小龙虾外观品质无损检测方法[J]. 食品与机械, 2026, 42(9): 83-93.

Citation: Wang Xujia, Wang Shuqing, Hu Haoyu. A non-destructive appearance quality detection method for primary-processed crayfish based on improved YOLOv11[J]. Food & Machinery, 2026, 42(9): 83-93.

环节仍以人工目测进行品质判定为主。然而,在小龙虾加工规模扩大和生产标准化推进的背景下,该方法的局限日益凸显:检测效率低、质量标准依赖工人经验,且长时间重复作业易引发视觉疲劳与注意力衰减,导致漏检与误判增加^[2-3]。此外,小龙虾产业存在明显的旺季与淡季^[4],峰值需求导致临时用工增加与成本攀升,人员流动带来学习曲线反复与质量波动,使人工分拣逐渐成为制约生产效率的瓶颈环节。

在食品加工产业中,对产品外观品质的自动化、智能化检测是保障产品质量、提升生产效率的关键环节^[5]。Ponce 等^[6-7]采用基于几何特征和手工设计特征的机器视觉方法,通过颜色、纹理、形状与尺寸等参数构建判别指标,并结合阈值分割或基于机器学习的传统分类器实现外观分级与缺陷检测。该方法对光照变化、背景干扰以及目标姿态和堆叠遮挡较为敏感,特征参数需要随产线条件重新标定,在复杂工业场景下难以同时兼顾检测精度与泛化能力^[8-9]。近年来,基于深度学习的计算机视觉技术因其高精度、高效率以及泛化性强的特点,已成为实现该目标的主流方法,并被广泛应用于肉类^[10]、水果^[11]、蔬菜^[12]等多个领域。

针对小龙虾养殖过程中夜间人工巡塘效率低、成本高的问题,李路等^[13]提出了一种基于 RGB 与红外图像多模态融合的轻量化目标检测网络 FAS-YOLOv8n,有效提升了夜间爬岸上草小龙虾的检测精度与实时性,缓解了弱光条件下单一模态图像对检测性能的限制。Ye 等^[14]在 YOLOv5 基础上引入 Transformer 结构与坐标注意力机制,针对传送带分级场景中小龙虾样本在图像中占比较小且类间差异细微、网络难以聚焦有效区域的问题,实现了小龙虾大小与成熟度的自动分拣,在保持模型轻量化的同时显著提升了检测性能。耿春新等^[15]针对小龙虾各部位在检测过程中因重叠、遮挡及姿态变化而导致的边界框回归精度下降问题,引入 Inner-MPDIoU 损失函数改进 YOLOv8n 算法,实现了虾体、螯足和尾部的精准定位。王子豪等^[16]利用 DeepLabV3+ 语义分割模型,通过分析虾头、虾钳等部位的像素占比与整虾质量的关系估算小龙虾的品质等级,但单视角平面图像在等级分界阈值易受到成像误差与分割误差的影响,实际应用存在一定的局限性。Shi 等^[17]面对小龙虾虾头、虾尾、螯足与步足等小尺度部位在密集堆叠、姿态变化与相互遮挡条件下的精细分割需求,提出一种面向边缘设备的轻量化实例分割模型 YOLOv11n-DHBC,在显著压缩模型体积的同时,提升了模型的多尺度特征建模能力,实现了对小龙虾各部位的高精度分割。鲁濠等^[18]提出一种基于 YOLOv5 的改进算法用于检测初加工小龙虾品质是否合格,表明目标检测模型可以用于半成品品质筛选,然而其数据集规模有限,分类标准也较为单一。王淑青等^[19]针对初加工生产线上小龙虾密集堆叠、遮挡且存在设备噪声干扰的

复杂背景,构建轻量化目标检测模型 YOLO-HDR,在 YOLOv8 框架中引入 PP-HGNetV2 骨干并结合动态卷积 (DynamicConv) 与 RepNCSPeLAN4 模块以增强多尺度特征表达和融合能力,提升了复杂背景下缺陷品质检测的精度与实时性。然而其数据集主要来源于单一工厂,样本多样性仍有进一步提升的空间,在不同产线与环境下的泛化能力仍需进一步验证。

目前,基于深度学习的小龙虾检测与分级方法研究主要集中在视角固定、光照条件单一或个体相对分散的环境中,对实际初加工生产线多视角、多光照以及密集堆叠、遮挡等复杂工况下的识别方法研究不足。此外,初加工生产线上的小龙虾往往存在目标小、姿态多变等问题,进一步增加了品质分级难度。试验拟面向真实初加工生产线场景,构建不同视角、不同光照条件下的初加工小龙虾外观品质数据集,并在 YOLOv11n 基础上提出一种适合工业流水线部署的高精度检测模型 YOLO-CQ,为小龙虾初加工智能化分级设备提供依据。

1 小龙虾数据集构建

1.1 小龙虾分级标准分析

小龙虾初加工产品主要有虾仁、虾尾、整只的原味虾以及调味小龙虾^[20]。加工企业在对小龙虾进行清洗后,通常采用水煮或蒸制将小龙虾加工至半熟状态,并在此基础上开展分级作业。煮熟后小龙虾的蜷曲程度可反映其新鲜度:鲜活个体在受热后通常呈蜷缩状,尾部弯曲至胸前;而死亡个体则伸展僵直。通过连接头部、头尾交界处及尾端 3 点形成的夹角可量化其蜷曲度,蜷曲度 $<90^{\circ}$ 判定为鲜活合格,反之为不新鲜^[19]。此外,甲壳残缺或头尾连接松弛的个体易出现肉质松散问题。生产线上脱落的断肢、虾头等易干扰自动分拣系统,应作为独立类别进行检测与排除。基于新鲜度与外观缺陷的组合,产线中小龙虾可划分为新鲜有缺陷、新鲜无缺陷、不新鲜有缺陷和不新鲜无缺陷 4 类。此外,还有一类干扰样本:脱落的残肢与虾体。

1.2 数据采集与处理

试验构建的小龙虾外观品质数据集由两部分图像构成。第一部分采集自湖北黄冈、潜江等地若干小龙虾加工企业的初加工生产线。采集系统由机械臂与工业相机协同组成,采用 Intel RealSense D435 相机进行俯视动态拍摄,图像分辨率为 1 280 像素 $\times$ 924 像素。视频帧率设为 48 帧/s,每隔 6 帧提取 1 张静态图像,共获得 2 370 张图片。经人工筛选,去除模糊及高相似度样本后,保留 1 486 张有效图像。

第二部分为补充试验采集数据,用于增强数据的多样性与代表性。于 2025 年 5 月 1—7 日,在湖北潜江中国虾谷多次购入不同批次的小龙虾样品,经水煮处理后进行图像采集。采集设备为红米 K70 Ultra 手机的后置摄像

头,图像分辨率为4 096像素×3 042像素。采集时选取自然光、LED照明及漫反射3种光照条件,并在不同俯视角($0^\circ \pm 15^\circ$)下进行拍摄,经筛选后择优保留2 238张图像。为了消除不同设备带来的分辨率差异与尺度不一致问题,将图像分辨率统一缩小至1 280像素×960像素。最终共获得3 724张有效图像。

1.3 数据集标注与增强

使用LabelImg软件对数据集进行人工标注。标准类别共5类:新鲜无缺陷、不新鲜无缺陷、新鲜有缺陷、不新鲜有缺陷和干扰样本(脱落的残肢和虾体)。对整只个体和脱落部件分别进行标注,确保目标区域完整覆盖虾体轮廓。图像边缘处无法明确辨别其类别的目标样本,将其标注为干扰样本。标注完成后按8:1:1比例划分数据集,得到训练集2 980张、验证集372张和测试集372张。数据集的基本信息见表1,除干扰外的4类样本实例基本平衡。此外,为了增强模型的泛化能力与数据的真实性和有效性,使用旋转变换及添加噪声等策略对训练集进行数据增强,将训练集中样本数量扩充至5 908张。

2 小龙虾图像识别方法

2.1 YOLOv11网络

YOLOv11^[21]是一种较新的一阶段目标检测算法,在检测精度与推理速度之间实现了优异平衡。该模型由输入层、骨干网络(Backbone)、颈部网络(Neck)和检测头(Head)组成。作为YOLOv8的直接继任者,YOLOv11针对性地引入多项架构改进:引入C3k2模块以提升特征提取效率;结合位置敏感空间注意力(PSA)构建C2PSA模块,强化模型对遮挡目标的特征感知能力;采用快速空间金字塔池化模块(SPPF)以进一步优化多尺度特征融合,显著提升小目标检测精度。YOLOv11n作为该系列的轻量化版本,在保持实时检测能力的同时,相比以往YOLO模型具备更高的精确率和计算效率,适用于资源受限的边缘部署场景。因此,选择以YOLOv11n作为基准模型进行改进。

2.2 YOLOv11网络改进

针对初加工小龙虾外观品质检测任务中目标尺寸普遍较小、堆叠遮挡以及不同品质等级间外观差异细微等问题,试验在YOLOv11n模型基础上,从特征提取、注意力建模及损失函数优化3个方面进行改进,构建YOLO-CQ模型。首先,在C3k2模块中引入DualConv优化其C3k结构,以增强模型的局部特征提取与跨通道信息交互能力。其次,在C2PSA模块中引入可变形注意力(DAT)机制替换原有注意力单元,使模型能够通过目标形变与尺度变化自适应聚焦关键区域。最后,在损失函数部分引入SlideLoss优化分类损失,通过基于IoU的动态加权策略平衡样本难度分布,强化模型对复杂与边界样本的学习能力。改进后的YOLO-CQ网络结构如图1所示。

表1 数据集基本信息

Table 1 Basic information of the dataset

类别	图像数量/张			图像标注总数
	训练集	验证集	测试集	
新鲜无缺陷	1 234	156	154	1 778
不新鲜无缺陷	1 008	126	126	1 482
新鲜有缺陷	1 076	136	134	1 576
不新鲜有缺陷	860	108	108	1 220
干扰样本	1 310	164	164	2 996

2.2.1 改进C3k2模块 C3k2模块基于跨阶段部分连接(CSP)思想设计,是一种兼顾特征复用与梯度流动的特征提取结构。它将输入特征图划分为两支,一支直接传递以保留原始特征,另一支通过若干C3k单元或瓶颈单元处理,使用残差连接以进一步增强特征表达能力。随后两支特征在通道维度进行拼接,并经 1×1 卷积调整通道数后输出,形成浅层与深层特征的互补表示,从而在特征提取能力与计算复杂度方面取得平衡。

在初加工小龙虾外观品质检测任务中,目标在图像中的占比整体偏小,品质判定依赖于整体弯曲形态与局部缺陷特征,且会受到传送带上密集堆叠、遮挡等因素的影响。C3k2模块中连续堆叠的卷积结构单元容易导致特征冗余与信息混叠。一方面,浅层纹理与边缘信息在多次卷积与下采样过程中容易被过度平滑或压缩,使得模型难以区分细微外观差异;另一方面,普通卷积在特征提取时感受野受限,模型难以充分融合全局与局部特征。因此,试验在对C3k2模块的C3k子模块中引入双通路卷积DualConv^[22],构造C3k2_DCB模块,其结构如图2所示。

DualConv是结合了组卷积与异构卷积思想的轻量化卷积算子,其结构如图3所示。在DualConv设计中,卷积滤波器被划分为 G 个组以处理输入特征图。对于每一组而言,输入通道被分为两部分:占总数的 $1/G$ 的输入通道通过并行的 3×3 组卷积与 1×1 逐点卷积处理,随后在对应位置逐元素相加;其余通道仅由 1×1 卷积处理。基于这种思想,DualConv中并行的 3×3 卷积核与 1×1 卷积核可以提取到更丰富的全局特征与局部特征并将其融合,从而有效地捕捉不同尺度图像特征。连续施加在全通道上的 1×1 卷积核保留了输入特征图的原始特征,完成通道维度上的特征重组与语义整合,从而增强模型的跨通道信息交互能力。同时,DualConv继承了组卷积策略的高效特性,通过对卷积滤波器的稀疏化分配,显著降低了模型参数量与计算量。

假设输入特征图尺寸为 $H_i \times W_i \times C_i$,输出特征图尺寸为 $H_o \times W_o \times C_o$,DualConv中组卷积核大小为 $K \times K$,DualConv分组数为 G ,则DualConv的总计算量 F_{DC} 为混合卷积部分 F_{mix} 和纯逐点卷积部分 F_{point} 的计算量之和:

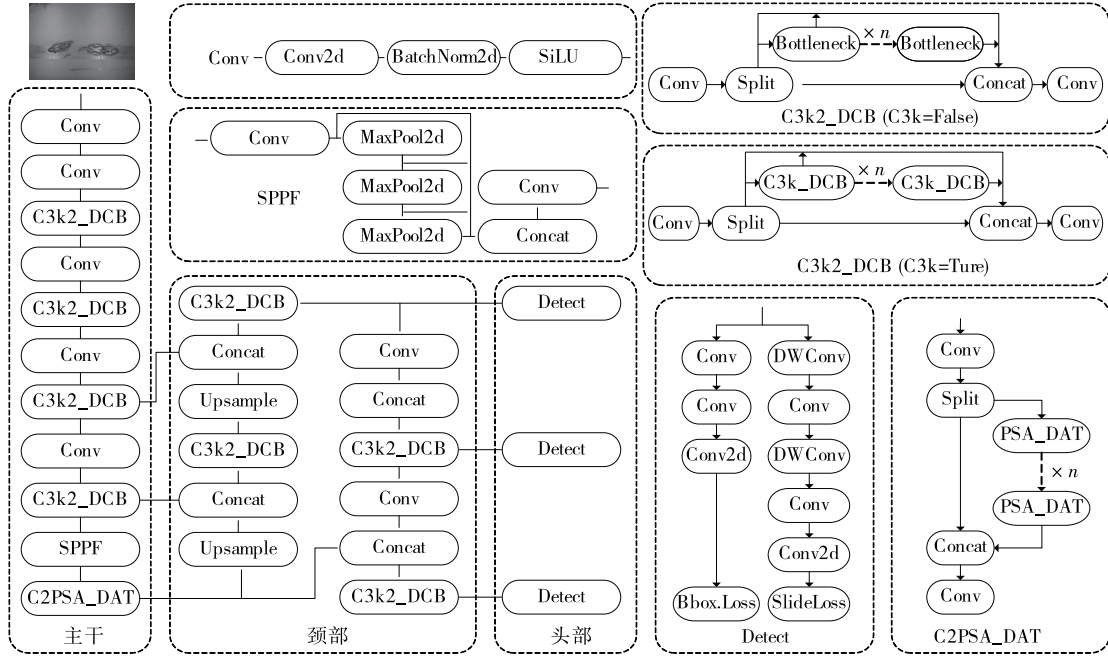


图 1 YOLO-CQ 整体结构图

Figure 1 Overall architecture of YOLO-CQ

$$F_{DC} = F_{\text{mix}} + F_{\text{point}} = H_o W_o C_i C_o \left(\frac{K^2 + 1}{G} + 1 - \frac{1}{G} \right) = H_o W_o C_i C_o \left(1 + \frac{K^2}{G} \right), \quad (1)$$

式中:

H_o 、 W_o 、 C_o ——输出特征图的高度、宽度、通道数;

C_i ——输入特征图的通道数;

K ——卷积核大小;

G ——DualConv 的分组数。

在标准卷积层中,每个输出通道均需与全部输入通道进行卷积操作,其浮点运算量 F_{sc} 可表示为:

$$F_{sc} = H_o W_o K^2 C_i C_o. \quad (2)$$

则 DualConv 与标准卷积层的浮点运算量的比值为:

$$\frac{F_{DC}}{F_{sc}} = \frac{1}{K^2} + \frac{1}{G}. \quad (3)$$

由式(3)可知,设 DualConv 中的组卷积核大小为 3×3 , G 较大时, DualConv 的计算量为标准卷积层的约 1/9。如图 2 所示,试验构建了基于 DualConv 的 C3k2_DCB 模

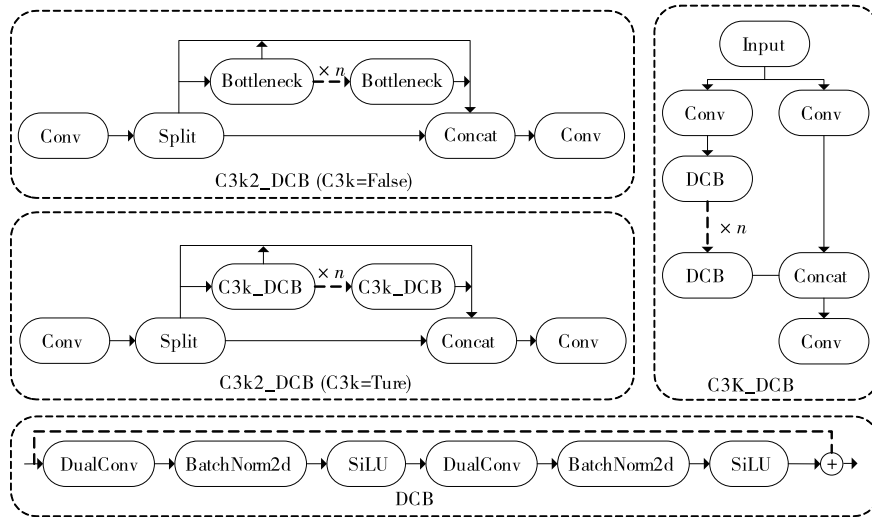


图 2 C3k2_DCB 模块结构图

Figure 2 Architecture of the C3k2_DCB module

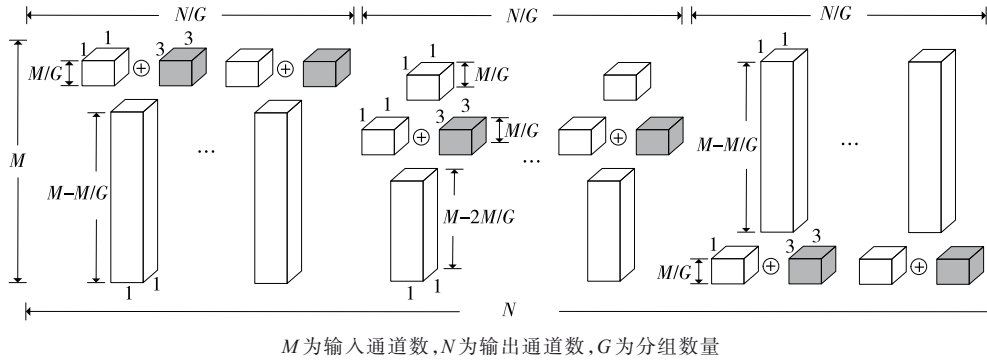


图3 DualConv 结构图

Figure 3 Architecture of DualConv

块。在该模块的C3k_DCB子单元中,采用DualConv替换了Bottleneck层内的部分普通卷积结构。这一改进的核心目的在于优化计算资源分配并提升特征传播效率。DualConv独特的并行结构在Bottleneck层内建立了多尺度特征流通道,有助于抑制标准卷积在深层特征映射中可能出现的特征模糊现象,从结构层面缓解小目标与细粒度缺陷检测中信息压缩过度的问题。

2.2.2 改进C2PSA模块 为了使模型在空间维度上更充分地关注图像中的关键区域,YOLOv11在原有CSP结构基础上嵌入PSA空间注意力机制,构造了C2PSA模块。C2PSA的核心设计是在一条分支中引入由多头自注意力(MHSA)与前馈网络(FFN)构成的PSA子模块,在另一条分支通过残差连接保留输入特征,最终在通道维度融合两路信息,以在保持梯度流动和特征复用的同时增强全局建模能力。在C2PSA中,MHSA的注意力权重计算基于固定采样窗口,在全局范围内采样密集且非自适应,难以针对输入目标的形变与位置变化进行动态调整。在小龙虾外观质量检测任务中,由于目标尺度较小且外观差异主要集中于局部结构与纹理细节,YOLOv11在复杂光照、姿态变化及遮挡条件下难以充分聚焦关键区域,对小目标和细粒度特征的识别效果容易受到局部纹理分布与背景噪声的影响,导致特征表达存在不稳定性。此外,MHSA在高分辨率特征图上计算量相对较大,增加了网络整体的推理负担并限制了其在实时检测场景下的应用效率。因此,将C2PSA模块中的MHSA注意力模块替换为可变形注意力(DAT)模块,构成C2PSA_DAT模块,其结构如图4所示。

DAT是一种将自注意力的全局依赖建模与可变形卷积^[23]的空间自适应性相融合的注意力机制,其核心思想是通过可学习偏移与连续插值实现特征采样位置的动态调整,使模型能够根据目标的形变与位置变化自适应地聚焦关键区域,从而增强对局部结构与细粒度特征的代表能力^[24]。DAT注意力机制的整体结构如图5所示。

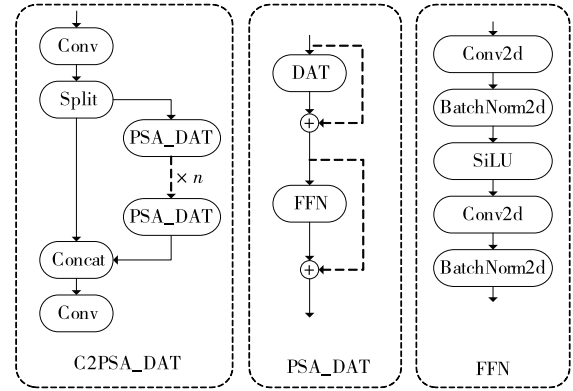


图4 C2PSA_DAT模块结构图

Figure 4 Architecture of the C2PSA_DAT module

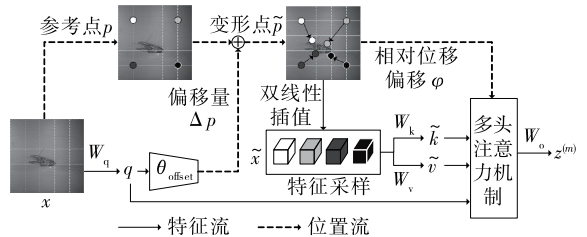


图5 可变形注意力机制结构图

Figure 5 Architecture of the deformable attention mechanism

首先,DAT在输入特征图 x 上生成均匀分布的参考点 p ,用于提供空间采样的初始位置,以确保模型在不同尺度的特征层中共享统一的空间参考体系。随后通过线性映射 W_q 得到输入特征图 x 的查询向量 q ,并将 q 作为偏移预测网络 $\theta_{\text{offset}}(\cdot)$ 的输入,用于生成每个参考点位置 p 对应的可学习偏移量 Δp 。为了防止偏移预测过大导致训练过程出现剧烈波动,DAT使用 $\tanh$ 函数与缩放比例系数 s 限制偏移量大小,以保证采样过程的平滑性与数值稳定性。

$$\Delta p \leftarrow s \cdot \tanh(\Delta p)$$

$$\Delta p = \theta_{\text{offset}}(q)$$
(4)

然后,DAT在原始特征图 x 上以偏移后的参考点 $p +$

Δp 作为变形点 $\tilde{p}$, 并将其作为参考位置执行双线性插值采样, 得到融合空间动态信息的偏移特征图 $\tilde{x}$ 。每个相邻格点根据其采样位置的距离获得不同权重, 距离越近, 权重越大, 所得到的特征值在空间上连续且可微, 使得模型能够平滑地学习注意力的空间分布, 避免了采样突变或梯度不连续问题。采样过程为

$$\tilde{x} = \phi(x; p + \Delta p), \quad (5)$$

式中:

ϕ ——双线性插值函数。

变形特征 $\tilde{x}$ 再分别经过线性映射生成键 $\tilde{k}$ 和值 $\tilde{v}$, 从而将插值采样后的特征映射到与原始特征图 p 统一的语义子空间, 使偏移采样特征键 $\tilde{k}$ 和值 $\tilde{v}$ 在通道维度上与查询特征 q 保持一致。线性映射关系为

$$q = xW_q, \tilde{k} = \tilde{x}W_k, \tilde{v} = \tilde{x}W_v, \quad (6)$$

式中:

W_q, W_k, W_v ——可学习的投影参数矩阵。

最后, 通过多头注意力机制将查询特征 q 、偏移后的特征键 $\tilde{k}$ 和特征值 $\tilde{v}$ 拼接整合, 并通过线性投影矩阵 W_o 得到融合局部细节与全局上下文的多头注意力 $z^{(m)}$ 。输出形式为

$$z^{(m)} = \sigma(\delta + \varphi(\tilde{p}; B))\tilde{v}^{(m)} \\ \delta = \frac{q^{(m)}\tilde{k}^{(m)}}{\sqrt{d}}, \quad (7)$$

式中:

$\sigma(\cdot)$ ——Softmax 函数;

$\tilde{p} = p + \Delta p$ ——经偏移后的采样位置;

$\varphi(\tilde{p}; B)$ ——相对位置偏移函数(用于刻画采样位置之间的集合关系);

δ ——查询 q 与偏移键 $\tilde{k}$ 的点积相似度(用于权衡查询与偏移键的语义相关性);

m ——遍历当前位置的若干采样点。

综上, C2PSA_DAT 模块使用可变形注意力机制 DAT 替换了原始 C2PSA 模块中基于固定采样方式的 MHSA 注意力。通过可学习偏移与连续插值实现对采样位置的动态调整, DAT 实现了对空间特征的动态偏移建模与自适应特征聚焦, 从而使模型能够在不同尺度与形变条件下保持对关键区域的关注。

2.2.3 优化损失函数 在初加工小龙虾外观品质检测任务中, 简单样本数量通常远多于困难样本, 这会导致数据集存在样本难易不平衡的问题。位于难易边界的模糊样本往往包含更多噪声、遮挡、光照干扰或姿态变化等复杂因素, 具有更高的识别难度与更丰富的判别信息, 对模型性能提升至关重要。然而, YOLOv11 采用的分类损失函数为带有交并比 (IoU) 标签的二元交叉熵损失函数 (BCE-WithLogitsLoss)。该函数在计算时对所有样本赋予相同权重, 未能根据预测框质量或样本难度自适应调整梯度

贡献。这导致模型在训练中更容易学习到简单样本的特征分布, 而对复杂样本的特征表征与识别能力相对不足。为解决此问题, 引入 SlideLoss^[25] 损失函数:

$$L_{\text{slide}} = f(I_i) \cdot L_{\text{base}}(x) \\ f(I_i) = \begin{cases} 1 & I_i \leq \mu - 0.1 \\ e^{1-\mu} & \mu - 0.1 < I_i < \mu \\ e^{1-x} & I_i \geq \mu \end{cases}, \quad (8)$$

式中:

$f(I_i)$ ——动态加权函数;

L_{base} ——基础分类损失函数 (BCEWithLogitsLoss);

I_i ——第 i 个样本预测框与真实框之间的交并比 IoU;

μ ——所有样本 IoU 的平均值。

SlideLoss 提出了一种基于样本 IoU 的动态加权函数, 通过对不同难度的样本赋予差异化权重, 从而自适应调整训练中不同难度样本的梯度贡献。如图 6 所示, SlideLoss 以所有样本 IoU 的平均值作为动态阈值 μ , 将高 IoU 样本划分为简单样本, 低 IoU 样本划分为困难样本。在此基础上, SlideLoss 进一步以 $\mu - 0.1$ 为界, 将困难样本细分为 IoU 接近于阈值 μ 的模糊样本和 IoU 更低的较困难样本。对于简单样本, SlideLoss 赋予其随 IoU 增大而指数衰减的权重, 防止大量简单样本主导梯度方向, 避免模型陷入局部最优。位于阈值附近的模糊样本通常涉及到噪声、遮挡、光照变化等复杂因素, 具有更高的判别难度, 是提升模型性能上限的关键。因此, SlideLoss 赋予此类样本更高的优先级, 以引导模型学习复杂特征。而对于 IoU 更低的困难样本, 考虑到其中可能混杂大量非典型特征与噪声, SlideLoss 仅对其保持基础权重, 以避免异常样本引发梯度振荡。综上, SlideLoss 在不同难度样本间建立了平滑的动态权重过渡机制, 有效平衡了模型的训练焦点, 增强了算法在复杂场景下的稳定性与泛化能力。

3 试验与分析

3.1 试验环境

试验平台使用 Windows 11 操作系统, 编程语言为 Python 3.10.16, 深度学习框架为 PyTorch2.2.1, 深度学习加速库为 CUDA 12.1。硬件环境包括 CPU 为 Intel(R)

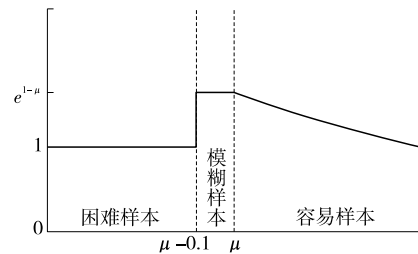


图 6 SlideLoss 动态加权函数

Figure 6 Dynamic weighting function of SlideLoss

Core i5-14600KF, GPU 为 NVIDIA GeForce RTX 4070Ti Super 16 GB, 内存为 2×16 GB。训练过程中, 批处理尺寸为 32, 训练轮次为 200, 初始学习率为 0.001 111, 采用 AdamW 优化器, 动量为 0.9, 权重衰减为 0.000 5。为了保证结果的公正性, 试验均在上述软硬件配置下完成, 训练参数保持一致。此外, 采用精确率(P)、召回率(R)、平均精度均值(mAP@0.5)、参数量(Params)、浮点运算量

(FLOPs)和帧率(FPS)作为模型评价指标。

3.2 主流模型对比试验

为了验证试验模型在小龙虾外观品质检测任务中的性能表现, 将其与当前主流目标检测算法进行对比试验。对比模型包括经典的 YOLO 系列、基于 Transformer 的 RT-DETR 系列、轻量化模型 NanoDet 系列以及 YOLO 轻量化变体 YOLO-MS^[26]。试验结果见表 2。

表 2 不同模型检测性能对比

Table 2 Comparison of detection performance of different models

模型	P/%	R/%	mAP@0.5/%	Params/M	GFLOPs	FPS
YOLOv5n	82.9	83.7	87.5	2.50	7.1	233.9
YOLOv8n	81.2	81.6	86.5	3.00	8.1	239.2
YOLOv10n	79.4	81.1	86.1	2.27	6.5	250.7
YOLOv11n	84.1	82.0	87.9	2.58	6.3	223.1
YOLOv12n	78.9	80.2	85.1	2.56	6.3	205.9
NanoDet-m	67.2	66.2	69.7	0.91	1.3	31.5
NanoDet-plus-m	61.1	67.8	73.4	1.15	1.7	32.1
YOLO-MS	78.4	81.3	85.7	2.36	7.2	217.4
RT-DETR-r18	82.1	80.7	78.4	19.88	57.0	96.6
RT-DETR-r34	83.1	82.6	82.2	31.11	88.8	79.8
RT-DETR-l	83.3	80.4	78.4	28.45	100.6	110.0
YOLO-CQ	86.6	84.3	91.3	2.83	7.6	211.5

由表 2 可知, YOLO-CQ 在精确率、召回率与 mAP@0.5 上领先于其他模型, 与基准模型 YOLOv11n 相比分别提升了 2.5%, 2.3%, 3.4%。尽管 YOLO-CQ 的参数量与浮点运算量与基准模型相比分别增加了 0.25 M、1.3 GFLOPs, 但整体计算开销变化较小, 仍保持较高的轻量化特性。基于 Transformer 的 RT-DETR 系列随着模型规格的增大, 其计算与存储开销显著增加。RT-DETR-l 模型的参数量与浮点运算量分别为 YOLO-CQ 的 10, 13 倍, 难以在资源受限的实时场景中高效部署。轻量化模型 NanoDet 系列虽然在参数量与计算量上具有明显优势, 但其在试验任务中检测精度较低, mAP@0.5 仅为 69.7% 和 73.4%, 且推理速度远低于 YOLO 系列, 难以满足高精度实时检测的需求。此外, 与致力于轻量化设计的 YOLO-

MS 相比, YOLO-CQ 在参数量略增的情况下, 实现了 mAP@0.5 提升 5.6% 的显著优势。在实时性方面, YOLO-CQ 的帧率为 211.5 帧/s, 虽然较基准模型 YOLOv11n 下降了 11.6 帧/s, 但仍远高于 RT-DETR 系列和 Nanodet 系列, 基本满足工业检测需求。综上, YOLO-CQ 在保持较高推理速度与较低计算成本的同时, 取得了更优的检测精度, 展现出良好的综合性能。

3.3 C3k2 改进对比试验

为了验证试验提出的 C3k2_DCB 模块的有效性, 在保持 YOLOv11n 整体结构与训练参数一致的条件下, 分别采用多种典型卷积形式进行对比试验, 结果见表 3。其中, 模型 C3k2_XXX 表示在 C3k2_DCB 模块中以对应的卷积结构替换 DualConv 所构成的变体模型。

表 3 C3k2 对比试验

Table 3 Comparative experiments of C3k2

模型	P/%	R/%	mAP@0.5/%	Params/M	GFLOPs	FPS
C3k2_DCB	84.7	82.9	88.9	2.83	7.5	208.3
C3k2_DSConv ^[27]	84.7	80.0	88.3	3.20	7.9	192.3
C3k2_RFACConv ^[28]	86.6	78.4	88.6	2.95	7.8	185.4
C3k2_DynamicConv ^[29]	76.3	85.2	87.2	3.55	7.2	222.2
C3k2_SACConv ^[30]	77.0	85.7	85.7	3.17	7.2	204.0

由表 3 可知,C3k2_DCB 模型在各项指标上表现更为均衡,mAP@0.5 达到 88.9%、参数量为 2.83 M、浮点运算量为 7.5 GFLOPs、FPS 为 208.3 帧/s。与之相比,C3k2_RFACnv 在精确率上略高,为 86.6%,但召回率较低,且 FPS 下降至 185.4 帧/s。C3k2_DynamicCnv 在召回率上略有优势,达到了 85.2%,且 FPS 为 222.2 帧/s,但精确率下降明显,仅为 76.3%。C3k2_DSConv 在精确率、召回率与 mAP@0.5 上均与 C3K2_DCB 的较为接近,但其参数量与浮点运算量更高,在模型实时性与轻量化方面仍存在一定差距。C3k2_SACnv 的召回率最高,达到了 85.7%,但其精确率与 mAP@0.5 下降明显。综合来看,试验提出的 C3k2_DCB 模块在精度、速度与复杂度之间实现了更优的平衡,验证了试验所提改进在特征提取方面的有效性。

3.4 C2PSA 改进对比试验

为了验证在 C2PSA 模块中 PSA 子模块内部使用 DAT 替换原注意力单元的合理性与有效性,在相同位置分别嵌入 EMA^[31]、SEAM^[25]、MSDA^[32] 和 LSKA^[33] 4 种代表性注意力机制作为对照方案,结果见表 4。

由表 4 可知,DAT 在 mAP@0.5 与精确率方面表现最佳,分别达到了 89.4%与 81.8%。虽然其召回率与其他注意力机制的相近,且参数量略有增加,但整体仍在合理范围内。SEAM 在召回率上表现突出,达到了 88.1%,mAP@0.5 与 DAT 也相近,但精确率偏低,仅为 77.8%。EMA 与 MSDA 在试验任务场景下并未表现出显著优势,mAP@0.5 与精确率相对较低,分别为 86.2%、75.6% 和 86.2%、77.1%。LSKA 在精确率与召回率上与 DAT 的相近,分别达到了 81.4%和 85.1%,且具有较好的实时性,帧

率达到了 238.1 帧/s,但其平均精度均值仍与 DAT 的存在差距。总体来看,DAT 注意力机制的改进效果更优,能够帮助算法更好地提高整体检测性能。

3.5 消融试验

为了验证各改进模块对整体检测性能的实际贡献,以 YOLOv11n 为基准算法,在自建小龙虾外观品质数据集上开展消融试验。所有试验的训练方法与训练环境完全一致,仅在网络结构中依次引入 C3k2_DCB、C2PSA_DAT 与 SlideLoss 3 个改进模块,以分析各模块对模型性能的独立与协同影响。消融试验结果见表 5,其中 Baseline 表示未引入改进模块的 YOLOv11n。

由表 5 可知,单独引入 C3k2_DCB 后,模型的精确率、召回率、mAP@0.5 分别提高了 0.6%、0.9%、1.0%,而参数量与浮点运算量分别上升了 0.25 M 和 1.2 GFLOPs,FPS 下降了 6.6%。说明引入 DualConv 可有效增强模型的特征跨层交互与局部聚合能力,尽管该结构略微增加了计算开销,但模型总体上仍保持轻量化特性。单独引入 C2PSA_DAT 后,模型的召回率与 mAP@0.5 分别提升了 3.6%和 1.5%。表明在加入可变形注意力机制后,模型能更灵活地聚焦目标的局部细节与形变区域,从而有效增强模型的特征感知能力。虽然精确率降低了 2.3%,但模型的实时性明显提升,FPS 上升了 9.3%。单独引入 SlideLoss 后,模型的 mAP@0.5 与召回率分别提升了 1.4%和 1.2%,但精确率下降了 3.2%。说明 SlideLoss 使模型在训练中能更充分学习复杂与边界样本,从而提升模型的召回性能,但同时降低了对简单样本的判别约束,导致精度略有波动。在同时引入 C3k2_DCB 与 C2PSA_DAT 后,模型的平均精度均值提升至 90.4%,召回率提升至

表 4 C2PSA 对比试验
Table 4 Comparative experiments of C2PSA

模型	P/%	R/%	mAP@0.5/%	Params/M	GFLOPs	FPS
C2PSA_DAT	81.8	85.6	89.4	2.61	6.3	243.9
C2PSA_EMA	75.6	85.1	86.2	2.53	6.3	192.3
C2PSA_SEAM	77.8	88.1	88.0	2.55	6.3	196.0
C2PSA_MSDA	77.1	86.5	86.2	2.60	6.3	166.7
C2PSA_LSKA	81.4	85.1	87.9	2.55	6.3	238.1

表 5 消融试验
Table 5 Ablation experiment

Baseline	C3k2_DCB	C2PSA_DAT	SlideLoss	P/%	R/%	mAP@0.5/%	Params/M	GFLOPs	FPS
✓				84.1	82.0	87.9	2.58	6.3	223.1
✓	✓			84.7	82.9	88.9	2.83	7.5	208.3
✓		✓		81.8	85.6	89.4	2.61	6.3	243.9
✓			✓	80.9	83.2	89.3	2.58	6.3	225.1
✓	✓	✓		82.2	84.2	90.4	2.83	7.6	203.5
✓	✓	✓	✓	86.6	84.3	91.3	2.83	7.6	211.5

84.2%。表明两个模块在特征提取与注意力聚焦方面形成了互补关系。在同时引入C3k2_DCB、C2PSA_DAT与SlideLoss后,模型的精确率、召回率、mAP@0.5分别达到86.6%, 84.3%, 91.3%, 较基准模型分别提升了2.5%, 2.3%, 3.4%。表明结构改进与损失函数优化在特征提取与模型学习层面实现了协同增益,从而进一步提高了整体检测性能。

综上,试验所提方法在参数量与计算量仅小幅增加的前提下显著提升了检测精度,且保持了较高水平的推理速度,具备出色的实时检测能力。

3.6 复杂场景检测效果可视化分析

为了更加直观地展示YOLO-CQ在不同场景下的性能,选取代表性样本图像进行可视化对比试验,结果如图7所示。图7(a)为正常检测情况下不同模型对小龙虾的识别效果:RT-DETR-l模型受到边缘信息不足的影响,将边缘的新鲜小龙虾误检为死亡小龙虾;YOLOv11n漏检

了图像中央的小龙虾样本;YOLOv8n与YOLO-CQ均识别出了所有目标,但YOLO-CQ的置信度更高。图7(b)为局部遮挡条件下不同模型对小龙虾的识别效果:虽然所有模型均检测到了左边的未遮挡小龙虾,但对于局部遮挡下的小龙虾样本,RT-DETR-l、YOLOv8n、YOLOv11n均出现漏检情况。图7(c)为堆叠检测场景下不同模型对小龙虾的识别效果:RT-DETR-l、YOLOv8n、YOLOv11n均漏检了小龙虾上方的残肢。受到干扰样本的影响,所有模型均检出了图像中的两只小龙虾样本,但YOLO-CQ的置信度表现更高。总体而言,YOLO-CQ在各种场景下的目标定位能力更强且置信度更高,具备出色的检测能力。

3.7 热力图可视化分析

为直观展示DAT替换前后模型对目标区域的特征响应变化,采用Grad-CAM++热力图进行可视化分析,结果如图8所示。红色区域代表模型主要关注的区域,红色越深表明模型对该区域的关注程度越高。基准模型的关

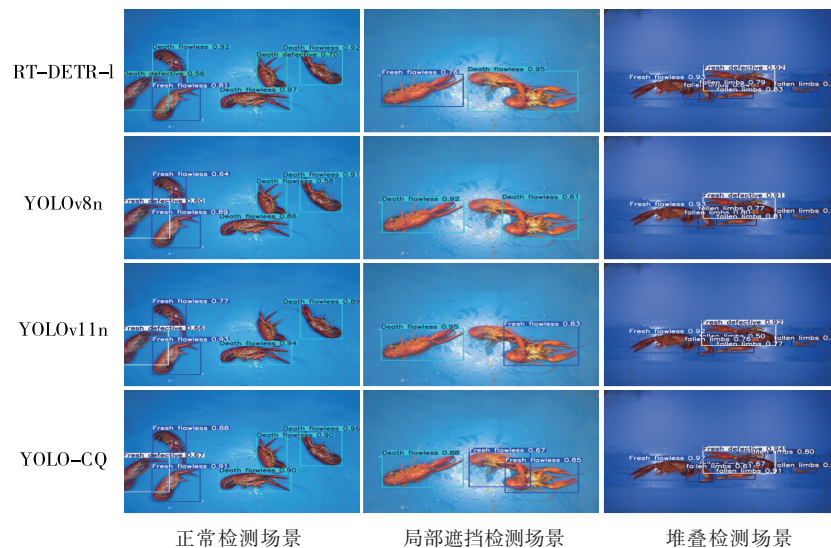


图7 复杂场景下检测效果对比

Figure 7 Comparison of detection performance in complex scenarios

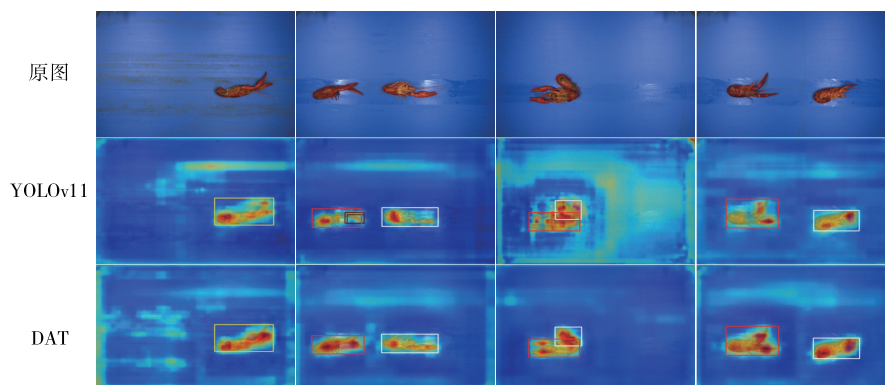


图8 热力图可视化结果

Figure 8 Visualization results of heatmaps

键区域响应分布相对分散,对小龙虾关键部位与局部缺陷的聚焦程度不足。引入 DAT 后,模型在目标区域的响应更集中,背景和非目标区域的特征响应明显减弱,表明模型对关键区域的关注程度更高,有利于提升检测性能。

4 结论

试验构建了面向复杂工业场景的小龙虾外观品质数据集,在 YOLOv11n 的基础上提出了改进模型 YOLO-CQ,用以实现初加工小龙虾外观品质的快速、无损自动检测。结果表明,YOLO-CQ 在自建数据集上精确率、召回率和平均精度均值分别达到 86.6%,84.3%,91.3%,相较基准模型 YOLOv11n 分别提升了 2.5%,2.3%,3.4%,同时仍保持较高的推理速度。在边缘目标、局部遮挡及堆叠干扰等复杂场景下,YOLO-CQ 的目标定位更稳定、置信度更高,能够满足初加工小龙虾生产线在线分级的应用需求。后续将进一步扩展数据多样性,探索结合多元传感信息的多模态检测方法,以进一步提升模型在复杂工业环境中的泛用性。

参考文献

- [1] 全国水产技术推广总站,中国水产学会.中国小龙虾产业发展报告(2024)[J].中国水产,2024(7):14-20.
National Fisheries Technology Extension Center, China Society of Fisheries. Development report on China's crayfish industry (2024) [J]. China Fisheries, 2024(7): 14-20.
- [2] 高竞博,李晔,杜闯.基于深度学习的小龙虾分级算法[J].现代计算机,2020(26):40-46.
Gao J B, Li Y, Du C. Crawfish classification algorithm based on deep learning[J]. Modern Computer, 2020(26): 40-46.
- [3] 焦鸣.基于视觉引导的小龙虾分级系统设计与试验[D].武汉:武汉轻工大学,2024:1-2.
Jiao M. Design and test of crayfish grading system based on visual guidance[D]. Wuhan: Wuhan Polytechnic University, 2024: 1-2.
- [4] Li Y F, Ku X C, Wang S D, et al. Evaluation of annual production and economic benefits of tropical rice-crayfish coculture system in China[J]. Agricultural Systems, 2024, 217 (C): 103936.
- [5] Xiao Z F, Wang J L, Han L, et al. Application of machine vision system in food detection[J]. Frontiers in Nutrition, 2022, 9: 888245.
- [6] Ponce J M, Aquino A, Millan B, et al. Automatic counting and individual size and mass estimation of olive-fruits through computer vision techniques[J]. IEEE Access, 2019, 7: 59 451-59 465.
- [7] Jia Z X, Li M, Shi C, et al. Determination of salmon freshness by computer vision based on eye color[J]. Food Packaging and Shelf Life, 2022, 34: 100984.
- [8] 唐彦嵩,徐锐豪,王凤加.机器视觉在食品无损检测中的应用研究进展[J].中国食品学报,2024,24(12):13-27.
Tang Y S, Xu R H, Wang S J. Research progress in the application of machine vision in food nondestructive detection [J]. Journal of Chinese Institute of Food Science and Technology, 2024, 24(12): 13-27.
- [9] 贾志鑫,杨霖,史策,等.农产品品质在线感知技术应用研究进展[J].农业机械学报,2025,56(6):17-32.
Jia Z X, Yang L, Shi C, et al. Research progress on application of online perception technology for agricultural product quality [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 17-32.
- [10] 王炼,柳军,皮杰,等.基于YOLOv8n的猪肉新鲜度图像识别算法[J].食品与机械,2025,41(5):98-104.
Wang L, Liu J, Pi J, et al. Image recognition algorithm for pork freshness based on YOLOv8n[J]. Food & Machinery, 2025, 41 (5): 98-104.
- [11] 李俊,曹博涛,彭新东.基于改进YOLOv7-tiny的苹果缺陷识别方法[J].食品与机械,2025,41(8):100-108.
Li J, Cao B T, Peng X D. An apple defect identification method based on improved YOLOv7-tiny[J]. Food & Machinery, 2025, 41(8): 100-108.
- [12] 施利春,边可可,王松伟,等.基于改进U-Net和IWOA-LSSVM的番茄综合品质检测方法研究[J].食品与机械,2025,41(8):109-117.
Shi L C, Bian K K, Wang S W, et al. Research on tomato comprehensive quality detection method based on improved U-Net and IWOA-LSSVM[J]. Food & Machinery, 2025, 41(8): 109-117.
- [13] 李路,孙超奇,周玉凡,等.基于FAS-YOLOv8n的爬岸上草小龙虾多源图像融合识别方法[J].农业机械学报,2025,56(8):526-534.
Li L, Sun C Q, Zhou Y F, et al. Multi-source image fusion recognition of crawling bank grass crayfish based on FAS-YOLOv8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(8): 526-534.
- [14] Ye X H, Liu Y X, Zhang D D, et al. Rapid and accurate crayfish sorting by size and maturity based on improved YOLOv5[J]. Applied Sciences, 2023, 13(15): 8 619.
- [15] 耿春新,王爱民,阎天宇,等.基于改进YOLOv8n损失函数的克氏原螯虾体特征识别算法[J].渔业现代化,2024,51(5):72-80.
Geng C X, Wang A M, Yan T Y, et al. Research on the identification algorithm of crayfish body features based on the improved YOLOv8n loss function[J]. Fishery Modernization, 2024, 51(5): 72-80.
- [16] 王子豪,胡志刚,付丹丹,等.基于DeepLab V³⁺估测小龙虾虾头虾钳占比及分级[J].食品与机械,2024,40(5):81-87,218.
Wang Z H, Hu Z G, Fu D D, et al. Grading crayfish by estimating the proportion of crayfish head and pincers based

- on DeepLab V³⁺ [J]. Food & Machinery, 2024, 40(5): 81-87, 218.
- [17] Shi W, Zhang J, Fu Y F, et al. Lightweight detection and segmentation of crayfish parts using an improved YOLOv11n segmentation model[J]. Scientific Reports, 2025, 15: 25634.
- [18] 鲁濠, 王淑青, 鲁东林, 等. 基于改进YOLOv5的小龙虾品质检测方法[J]. 湖北工业大学学报, 2023, 38(4): 76-81.
- Lu H, Wang S Q, Lu D L, et al. Quality detection method of crayfish based on improved YOLOv5[J]. Journal of Hubei University of Technology, 2023, 38(4): 76-81.
- [19] 王淑青, 陈开元, 周森, 等. 基于YOLO-HDR的小龙虾缺陷品质检测方法[J]. 食品与机械, 2025, 41(3): 100-107.
- Wang S Q, Chen K Y, Zhou M, et al. A method for detecting quality defects of crayfish based on YOLO-HDR[J]. Food & Machinery, 2025, 41(3): 100-107.
- [20] 高晓光, 吕蒙, 臧芳波, 等. 小龙虾加工与保鲜技术研究进展[J]. 保鲜与加工, 2021, 21(12): 126-131, 139.
- Gao X G, Lyu M, Zang F B, et al. Research progress of crayfish processing and preservation technology[J]. Storage and Process, 2021, 21(12): 126-131, 139.
- [21] Khanam R, Hussain M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024-10-23) [2025-08-19]. <https://arxiv.org/abs/2410.17725>.
- [22] Zhong J C, Chen J Y, Mian A. DualConv: dual convolutional kernels for lightweight deep neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(11): 9 528-9 535.
- [23] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 764-773.
- [24] Xia Z F, Pan X R, Song S J, et al. Vision transformer with deformable attention[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA: IEEE, 2022: 4 784-4 793.
- [25] Yu Z P, Huang H B, Chen W J, et al. YOLO-FaceV2: a scale and occlusion aware face detector[J]. Pattern Recognition, 2024, 155: 110714.
- [26] Chen Y M, Yuan X B, Wang J B, et al. YOLO-MS: rethinking multi-scale representation learning for real-time object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(6): 4 240-4 252.
- [27] Qi Y L, He Y T, Qi X M, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation[C]// 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris: IEEE, 2024: 6 047-6 056.
- [28] Zhang X, Liu C, Yang D G, et al. RFConv: receptive-field attention convolution for improving convolutional neural networks[EB/OL]. (2023-04-06) [2025-09-11]. <https://arxiv.org/abs/2304.03198>.
- [29] Han K, Wang Y H, Guo J Y, et al. ParameterNet: parameters are all you need[EB/OL]. (2023-06-26) [2025-10-03]. <https://arxiv.org/abs/2306.14525>.
- [30] Qiao S Y, Chen L C, Yuille A. DetectoRS: detecting objects with recursive feature pyramid and switchable atrous convolution[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN: IEEE, 2021: 10 208-10 219.
- [31] Ouyang D L, He S, Zhang G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]// 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island: IEEE, 2023: 1-5.
- [32] Jiao J Y, Tang Y M, Lin K Y, et al. DilateFormer: multi-scale dilated transformer for visual recognition[J]. IEEE Transactions on Multimedia, 2023, 25: 8 906-8 919.
- [33] Lau K W, Po L M, Rehman Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [J]. Expert Systems with Applications, 2024, 236: 121352.