

DOI:10.19651/j.cnki.emt.2518845

基于改进 YOLOv8n 的快递包裹缺陷检测方法研究^{*}

杨慧敏 高小雯 李瑞涛 王汉霞

(东北林业大学土木与交通学院 哈尔滨 150040)

摘要:为解决快递包裹缺陷检测中对复杂包裹类型和细节特征的识别能力有限,以及现有模型在精度和实时性方面的不足,提出一种基于改进 YOLOv8n 的快递包裹缺陷检测算法。首先,将网络中的 C2f 模块融合频率自适应空洞卷积设计了 C2f-FADC 模块,在处理多尺度、多频率缺陷检测任务时灵活调整,优化特征提取过程和提高表征能力;其次,引入 SimSPPF 模块替代原有 SPPF 模块,简化结构的同时增强多尺度特征融合能力,改善对小尺寸目标的感知效果;最后,将边界框回归损失函数替换为 Shape-IoU,以更精准地建模预测框与 GT 框之间的形状与尺度差异,优化检测定位性能。在自制的包裹缺陷数据集上,改进后的算法检测精度为 96.3%,与原算法相比 mAP50 提高了 4.4%,检测速度达到 98 帧,综合考量较其他算法具有明显优势,验证了该方法的有效性和优越性。

关键词: 缺陷检测;快递包裹;YOLOv8n;频率自适应空洞卷积(FADC);SimSPPF;Shape-IoU

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:** 520.6020

Research on an improved YOLOv8n-based method for defect detection in express packages

Yang Huimin Gao Xiaowen Li Ruitao Wang Hanxia

(College of Civil Engineering and Transportation, Northeast Forestry University, Harbin 150040, China)

Abstract: To address the limited recognition capability of complex package types and fine-grained features in package defect detection, as well as the shortcomings in precision and real-time performance of existing models, this paper proposes an improved YOLOv8n-based algorithm for defect detection in express packages. First, the C2f module in the network is integrated with frequency-adaptive dilated convolution (FADC) to design the C2f-FADC module, which dynamically adjusts when handling multi-scale and multi-frequency defect detection tasks, optimizing the feature extraction process and improving the representational ability. Secondly, the SimSPPF module is introduced to replace the original SPPF module, simplifying the structure while enhancing multi-scale feature fusion capability and improving the perception of small-sized targets. Finally, the bounding box regression loss function is replaced with Shape-IoU to more accurately model the shape and scale differences between the predicted and ground-truth boxes, optimizing the detection localization performance. On a self-constructed package defect dataset, the improved algorithm achieved a detection accuracy of 96.3%, with a 4.4% in-increase in mAP50 compared to the original algorithm, and a detection speed of 98 FPS. Considering both precision and speed, the proposed method shows significant advantages over other algorithms, validating its effectiveness and superiority.

Keywords: defect detection; express package; YOLOv8n; frequency-adaptive dilated convolution (FADC); SimSPPF; Shape-IoU

0 引言

在现代物流行业中,包裹缺陷检测已成为确保运输过程中货物完好无损、提升服务质量和客户满意度的关键环节。随着电商行业的迅猛发展,包裹缺陷检测的准确性和

实时性在保障货物安全、提升供应链效率以及降低运营成本方面变得尤为重要。因此,开发高效且准确的检测方法,不仅能够提升物流效率,还能增强企业在行业中的竞争力。包裹缺陷检测分为传统方法和基于深度学习的方法。传统方法依赖人工,易受视觉疲劳影响,检测结果不全面^[1]。而

收稿日期:2025-05-18

^{*} 基金项目:黑龙江省自然科学基金(LH2021C016)项目资助

深度学习技术及其强大的抗干扰能力使基于图像的物体检测算法^[2]在包裹缺陷检测中得到了广泛应用。通过优化深度学习模型的准确性和实时性,可以有效提升检测效率,减少错误率,对于保障物流安全和提升服务质量具有重要意义^[3]。

基于深度学习的缺陷检测算法分为两类:两阶段算法(如 Fast R-CNN^[4]、Faster R-CNN^[5]、Mask R-CNN^[6])和单阶段算法(如 SSD^[7-8]、YOLO^[9-15])。两阶段算法精度高但速度慢,不适合边缘设备;单阶段算法处理速度较快,但精度稍低。深度学习模型通常依赖大量计算资源,而边缘设备计算资源有限,复杂的参数设计增加了部署难度,且实时性也是应用中的挑战。

针对现阶段存在的问题和挑战,研究者们设计了多种缺陷检测模型。在提升精度方面,杨明旭等^[16]基于YOLOv8设计主干网络UNetV2,采用轻量化ADown卷积层和引入ContextAggregation注意力机制,将原颈部网络与检测头替换为Slim-Neck和SAttention,检测精度提升至95.7%,但导致模型计算量增至11.9 G。Liu等^[17]在YOLOv5中集成CBAM模块增强特征关注,通过金字塔聚合网络实现跨尺度融合,并优化骨干网络添加ASFF结构混合多尺度特征,实现平均精度0.96、召回率0.91及F1分数0.94,但计算负载增加,推理速度仅25 FPS,难以满足实时检测需求。Mao等^[18]基于CenterNet提出用ResNeXt替换ResNet,集成SKNet注意力机制与YOLOv8的C2f模块形成多分支结构,并引入椭圆高斯核优化尺寸回归损失,提升模型检测精度达95.2%。然而参数量仍高达32.498 M,计算量达90.86 G。王静蕾等^[19]基于YOLOv10在C2f模块后引入CBAM注意力机制,采用全维动态卷积强化特征提取能力,优化头部层结构及损失函数,实现检测准确率>98%,但FPS仅52.35,无法匹配高速产线实时需求。这部分学者虽然有效提升模型性能,但在降低模型复杂度和提高推理速度方面仍有改进空间。在多类缺陷检测方面,Li等^[20]基于YOLOX提出CSPCrossLayer模块替换骨干网络中的CSPLayer,在骨干输出端增加SA模块,并用PSblock模块来替代特征融合网络中的CSPLayer,提升跨尺度特征融合效率。但不同缺陷类别检测精度差异显著,Burn缺陷检测精度达93.5%,而pitting缺陷检测精度仅72.1%。周耀威等^[21]基于YOLOv7使用BoT3替换主干网络,引入CBAM注意力机制与MP-2模块融合以增强特征表达,但包裹缺陷数据集类别不平衡,且完好无损检测精度达94.7%,而包装褶皱检测精度仅83.1%。在轻量化方面,Liu等^[22]在SSD基础上引入MobileNetv3作为骨干网络,采用下采样堆叠分别对最大池化和常规3×3卷积处理的特征图进行融合,将扩张卷积整合到Inceptionv3中,形成多分支并行扩张卷积模块,使模型参数数量减少了14.16 M,检测速度达到48 fps,但在模型的轻量化,以及推理速度方面仍然具有较大的改

进空间。

考虑到现有研究的局限性,本文提出一种基于改进YOLOv8n的快递包裹缺陷检测方法,旨在实现包裹缺陷检测的高定位精度和快速识别。首先,引入频率自适应空洞卷积(frequency adaptive dilated convolution, FADC)设计C2f-FADC,使得网络能够在处理多尺度、多频率缺陷检测任务时灵活调整,优化特征提取过程和提高表征能力,提高检测精度和推理速度。其次,采用简化版空间金字塔池化模块(simplified spatial pyramid pooling-fast, SimSPPF)替代原有的SPPF模块,增强了对细小缺陷的捕捉能力,并优化了多尺度特征融合,提升了检测精度。最后,采用Shape-IoU损失函数替代传统的回归损失函数,以更精确地建模预测框与GT框之间的形状和尺度差异,从而提高了包裹缺陷的定位精度。

1 YOLOv8n模型

YOLOv8是由Ultralytics开发的目标检测模型,提升了速度、精度和易用性。YOLOv8使用了CSPDarknet作为其骨干网络,将CSP模块更改为C2f模块,C2f模块通过交叉阶段部分瓶颈与两个卷积操作将上下文信息与高级特征相结合,以提高检测性能。在损失函数方面,增加了DFL loss以补充边界框损失,能够处理连续的标签,并且可以优化分类和边界框回归,进一步提高对小目标的检测性能^[23-24]。YOLOv8提供5种不同尺寸的模型,本文选用YOLOv8n作为基准模型,其结构包括输入层(Input)、主干网络(Backbone)、特征融合模块(Neck)和输出层(Head),分别负责图像处理、特征提取与融合、以及类别预测和边界框回归。

2 改进YOLOv8n模型

为提升包裹缺陷检测的准确性和效率,本文提出了一种改进YOLOv8n的网络模型,首先,将FADC与C2f模块相融合,并用融合后的C2f-FADC模块替换Backbone和Neck部分的C2f模块,提升模型在处理复杂频率信息和多尺度特征时的性能,增强对不同尺度、形状和姿态的缺陷的感知能力。其次,用SimSPPF模块替换SPPF模块,通过简化设计,减少冗余计算,优化多尺度特征的融合,在提升检测精度的同时加快检测速度。最后用Shape-IoU损失函数替换原有的CIoU损失函数,将几何关系、边框的形状和尺度差异同时考虑,改善形状不规则和尺度差异较大的目标的检测性能,进一步提高定位精度。改进YOLOv8n网络结构如图1所示。

2.1 C2f-FADC模块

在中转站的传送带上进行包裹缺陷检测时,系统需在毫秒级别内完成图像采集、处理和缺陷识别,同时确保高准确率。YOLOv8n的C2f模块通过跨阶段特征融合优化了上下文信息,但其标准卷积操作仍存在明显局限性,包括感

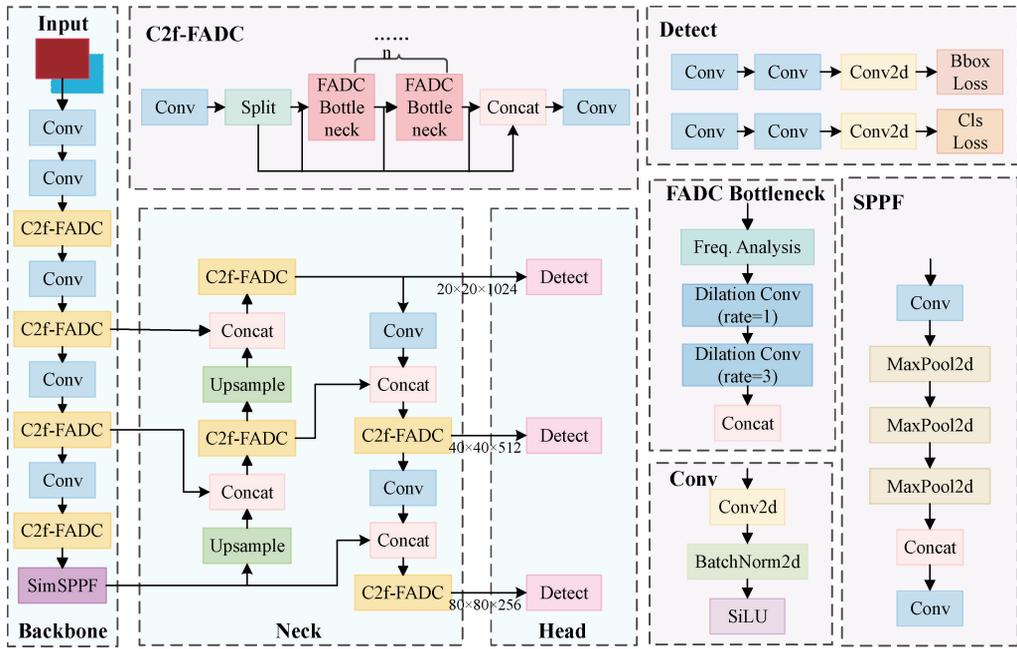


图 1 改进 YOLOv8n 网络结构

Fig. 1 Improved YOLOv8n network structure

受野单一,固定尺度的卷积核难以兼顾污损和褶皱的多尺度特征;频率响应不足,对高频细节(如 $<1\text{ mm}$ 的封口裂痕)的捕捉能力较弱,易导致漏检。

为了使模型能够更精准的识别和定位目标缺陷,本文引入一种通过自适应捕捉不同频率特征的 FADC 与 C2f 结合。FADC 包括 3 个关键策略:自适应扩张率(adaptive

data rate, AdaDR)、自适应内核(adaptive kernel, AdaKern)和频率选择(frequency selection, FreqSelect),它们分别在空间膨胀率、卷积核权重以及频率平衡方面优化,增强了卷积操作的有效性,改进了特征提取的精度^[25]。FADC 的结构如图 2 所示。

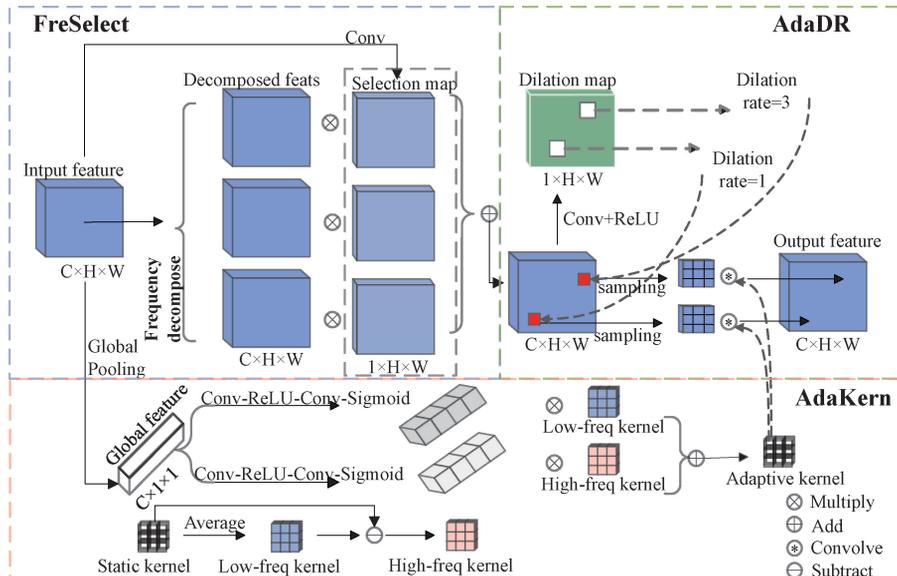


图 2 FADC 结构图

Fig. 2 Architecture of FADC

AdaDR 策略动态调整每个像素的扩张率,平衡大接受野和有效带宽。传统膨胀卷积在高膨胀率下易丢失高频

信息。AdaDR 通过在低频区域增加扩张率、在高频区域抑制扩张率,减少高频信息丢失,提升感受野捕捉能力。具

体计算过程为式(1):

$$\hat{X} = \sum_{i,j} W_{i,j} \times X_{i,j} \quad \text{where} \quad W_{i,j} = f(AdaDR) \quad (1)$$

其中, $W_{i,j}$ 表示卷积核权重; $X_{i,j}$ 为输入特征图的像素值; $f(AdaDR)$ 是根据输入特征图的空间信息动态计算的膨胀率函数。

AdaKern 通过将卷积核分解为低频和高频部分,并动态加权,提高卷积操作的有效带宽。低频部分通过均值滤波器平滑,捕捉全局信息,高频部分通过残差部分提取局部细节。通过轻量级全局池化和卷积操作预测加权系数,优化特征提取。具体计算公式为式(2):

$$W = W_{avg} + \lambda_l \times W_l + \lambda_h \times W_h \quad (2)$$

其中, W 为最终的卷积核参数,包含低频部分 W_l 和高频部分 W_h ,并通过动态权重系数 λ_l 和 λ_h 加权。 W_{avg} 为卷积核的均值部分,表示低频部分的平滑处理。 λ_l 和 λ_h 是通过网络学习得到的动态权重系数,分别表示低频和高频部分的加权值。

FreqSelect 策略通过傅里叶变换对输入特征图转换到

频率域,并为每个频带生成掩码 M_b ,选择不同的频率成分。每个频带的频率成分通过掩码提取并加权,最终得到平衡后的频率特征图。具体为式(3)~(4):

$$X_b = M_b \times F(X) \quad (3)$$

$$X^{(i,j)} = \sum_{b=1}^B A_b^{(i,j)} \times X_b^{(i,j)} \quad (4)$$

其中, X_b 表示第 b 个频带的频率特征,通过傅里叶变换 $F(X)$ 得到; M_b 是二进制掩码,负责选择特定频带的频率成分; $A_b^{(i,j)}$ 是频带 b 上的加权系数,决定了该频带在空间域中的重要性; B 是频带的总数。

将 FADC 与 C2f 结合,其中 Bottleneck 模块中的卷积层替换为 FADC 模块,从而形成了新的 FADC Bottleneck 模块,最终构建出 C2f-FADC 模块。C2f-FADC 模块将 Backbone 和 Neck 部分的 C2f 模块替换掉。与基准模型相比,C2f-FADC 模块在增强特征提取的同时,改善了对细节缺陷的捕捉精度,且在保证不增加计算量的同时,提高了模型对复杂缺陷的识别准确率和鲁棒性,使得该模型在快速地进行缺陷检测时具有更好的性能。C2f-FADC 模块如图 3 所示。

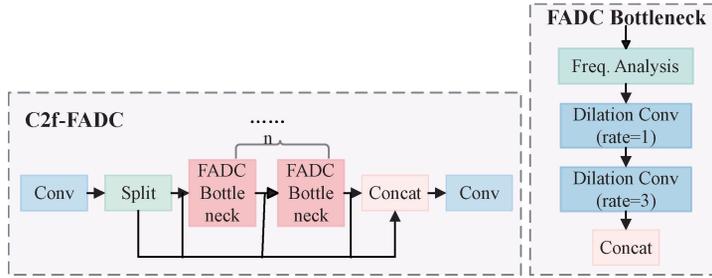


图 3 C2f-FADC 结构

Fig. 3 Architecture of C2f-FADC

2.2 SimSPPF

SPPF 模块通过多尺度池化增强模型对不同尺寸特征的感知能力,旨在提高对目标的检测精度。然而,在快速包裹缺陷检测中,包裹上的缺陷通常较小且具有高频特征,如裂痕和褶皱,SPPF 在捕捉这些微小缺陷时存在局限。同时,SPPF 的感受野和卷积核固定,难以灵活应对多种形态和姿态的缺陷。此外,SPPF 设计较为复杂,计算量较大,导致处理速度较慢。

为解决上述问题,本文引入一种简化的空间金字塔池化模块 SimSPPF,该模块通过多个卷积操作和池化技术优化特征提取,增强了对细小缺陷和高频信息的捕捉能力,并有效融合不同尺度的特征,提高了对多尺度缺陷的识别精度^[26]。同时,SimSPPF 简化设计,减少冗余计算,使得模型在保持高精度的同时,能够更快速地进行缺陷检测。SimSPPF 结构如图 4 所示。

SimSPPF 通过一系列卷积操作和池化技术来优化特征提取过程。首先,输入特征图通过一个初始的 SimConv 卷积层进行处理,通道数被降低为原来的一半。SimConv

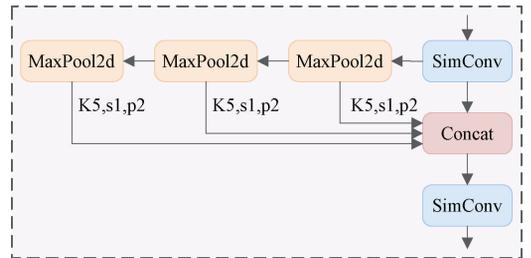


图 4 SimSPPF 结构

Fig. 4 Architecture of SimSPPF

是一个自定义卷积模块,包含卷积操作 (nn.Conv2d)、批归一化 (nn.BatchNorm2d) 和 ReLU 激活函数。卷积操作用于提取输入特征图的关键信息,批归一化加速训练过程并提升模型的稳定性,而 ReLU 激活函数引入非线性,增强了模型的表达能力。然后,SimSPPF 通过多次最大池化和拼接操作,融合不同尺度的特征。最后,经过一个 SimConv 卷积层,将融合后的特征图转换为指定的输出通道数,从而进一步优化特征表示并增强检测精度。

2.3 Shape-IoU

YOLOv8 模型的损失函数由分类损失 BCE Loss, 边框回归损失 DFL Loss 和 CIoU Loss 组成。CIoU Loss 主要考虑了预测框与 GT 框之间的重叠区域、中心点距离和长宽比, 但却忽视了边界框回归样本自身形状和尺度等固有属性在边界框回归过程中的作用。在目标检测中, 边框回归损失是定位分支的核心, 它直接影响检测精度。原损失函数的不足可能导致在处理形状或尺度差异较大的目标时, 定位精度降低。

为了解决这一问题, 本文引入了 Shape-IoU 损失函数, 其不仅考虑到预测框和 GT 框之间的重叠程度, 还考虑边框自身形状和尺度的差异, 能够更精准地计算 IoU 值, 从而提高边框回归的精度。相比传统的损失函数, Shape-IoU 不仅关注边框之间的几何关系, 还能有效处理形状不规则和尺度差异较大的目标, 能更准确地体现边界框回归效果^[27]。

Shape-IoU 损失函数由 IoU 损失 (IoU cost)、距离损失 (distance cost) 和形状损失 (shape cost) 组成, 其计算参数如图 5 所示。

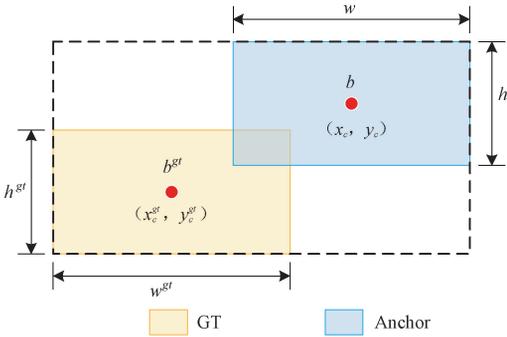


图 5 计算参数示意图

Fig. 5 Schematic diagram of calculation parameters

其计算公式如式(5)~(10)所示。

$$IOU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (5)$$

$$\omega\omega = \frac{2 \times (\omega^{gt})^{scale}}{(\omega^{gt})^{scale} + (h^{gt})^{scale}} \quad (6)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(\omega^{gt})^{scale} + (h^{gt})^{scale}} \quad (7)$$

$$distance^{shape} = hh \times (x_c - x_c^{gt})^2 / c^2 + \omega\omega \times (y_c - y_c^{gt})^2 / c^2 \quad (8)$$

$$\Omega^{shape} = \sum_{t=\omega, h} (1 - e^{-\omega_t})^\theta, \theta = 4 \quad (9)$$

$$\begin{cases} \omega_\omega = hh \times \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})} \\ \omega_h = \omega\omega \times \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases} \quad (10)$$

其中, B 为预测框; B^{gt} 为 GT 框; ω 和 h 分别是预测框的宽度和高度, h^{gt} 、 ω^{gt} 分别表示 GT 框的长和宽, x_c 、 y_c

表示预测框的中心点坐标, x_{gtc} 、 y_{gtc} 分别表示 GT 框的中心点坐标; $scale$ 是比例因子, 与数据集中目标的比例相关; $\omega\omega$ 和 hh 分别是水平和垂直方向的权重系数; c 为包含 GT 框和预测框的最小框的对角线距离。 Ω^{shape} 表示形状因子, 用于衡量预测框和 GT 框形状的不一致性, 确保预测框的形状尽可能接近真实的 GT 框形状。参数 θ 用于调节模型对形状损失的重视程度, 以防止过分关注形状损失而忽视预测框的定位精度, 通过遗传算法得知该参数的最优值为 4; ω_t 为宽/高损失。形状损失是通过计算预测框与 GT 框之间宽度和高度的差异, 再根据最大值进行归一化获得。对应的边界框回归损失如式(11)所示。

$$L_{shape-IoU} = 1 - IOU + distance^{shape} + 0.5 \times \Omega^{shape} \quad (11)$$

3 实验结果与分析

3.1 数据集介绍

鉴于当前开源的包裹缺陷数据集稀缺且类别单一, 难以满足本研究需求, 实验数据为自主采集的包裹缺陷数据集, 研究对象选定为瓦楞纸包装盒。图像采集使用 Kinect V1 相机实地拍摄包裹, 原始图像分辨率为 $1\,280\text{ pixel} \times 960\text{ pixel}$ 。为模拟实际应用场景并确保数据的多样性与现实性, 采集过程涵盖了不同拍摄距离、多角度取景以及不同程度的遮挡情况等多种变量条件。

本文在实验过程中采集了 4 500 张包裹图像, 涵盖缺陷包裹 3 750 张和完整包裹 750 张, 缺陷包裹类别包括封口损坏、侧边损坏、包装褶皱及污损, 图 6 展示了 5 种不同包裹类型的图像示例。同时, 使用 LabelImg 工具对所有包裹图片中的目标进行了标注, 标注类别分别是完整 (completion)、封口损坏 (seal)、侧边损坏 (broadside)、包装褶皱 (fold)、包装污损 (stained)。

为提高模型泛化能力, 解决原始的图像数据存在类别不平衡问题, 本文使用 Albumentations 库中的一系列随机的数据增强策略, 包括随机平移、缩放、旋转、亮度对比度调整、高斯噪声、灰度化等方式将数据集扩充到 11 250 张, 数据增强前后每类缺陷的数量如图 7 所示。将处理后的数据集按 8 : 1 : 1 的比例随机划分训练集 (9 000 张)、验证集 (1 125 张) 与测试集 (1 125 张)。

3.2 实验环境及参数设置

本文实验操作系统为 Windows 11, CPU 处理器为 Intel(R) Core(TM) i7-13650HX, 内存 16 GB; GPU 为 NVIDIA GeForce RTX 4060。在 Visual Studio Code 中进行实验, 深度学习框架为 Pytorch (版本 2.0.0), Python 版本为 3.8。训练时, 输入图片的分辨率为 640×640 , 训练轮次为 200 轮, Batch_size 为 8, 采用 SGD 优化器训练网络, 网络初始学习率为 0.001。

3.3 评价指标

本实验使用了 6 个主要评价指标: mAP、Parameters、



图6 5种缺陷类型的图像示例

Fig. 6 The image examples of five types of defects

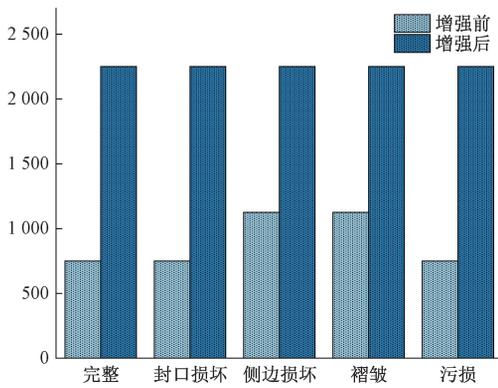


图7 各类缺陷数量图

Fig. 7 Quantity diagram of various defects

$$Recall = \frac{TP}{TP + FN} \quad (16)$$

其中, TP 、 FP 和 FN 分别表示正确检测的真阳性、假阳性和假阴性。

3.4 不同卷积模块性能对比实验

为验证并选择最适宜的卷积模块, 本文基于YOLOv8n目标检测架构, 采用近期先进的9种不同卷积模块进行性能评估, 包括全维度动态卷积(ODConv)、动态蛇形卷积(DSConv)、动态卷积(DynamicConv)、可切换空洞卷积(SAConv)、空间通道重组卷积(SCCConv)、空间深度转换卷积(SPDCConv)、可变核卷积(AKConv)、感受野自注意力卷积(RFACConv)、FADC, 不同卷积模块对比结果如表1所示。

表1 不同卷积网络实验对比

Table 1 Comparison of experiments with different convolutional networks

| 卷积网络 | mAP50/ % | mAP50-95/ % | Params/ M | GFLOPs | FPS |
|-------------|-------------|----------------|--------------|--------|-----|
| YOLOv8n | 91.9 | 86.9 | 3.0 | 8.1 | 84 |
| ODConv | 81.5 | 72.0 | 2.5 | 5.6 | 81 |
| DSConv | 86.6 | 79.8 | 3.2 | 8.4 | 45 |
| DynamicConv | 88.3 | 82.8 | 3.3 | 6.5 | 56 |
| SAConv | 88.9 | 85.4 | 3.6 | 7.0 | 53 |
| SCCConv | 90.2 | 85.9 | 3.1 | 8.2 | 85 |
| SPDCConv | 87.0 | 81.9 | 2.7 | 7.4 | 87 |
| AKConv | 85.3 | 80.3 | 2.7 | 7.4 | 46 |
| RFACConv | 88.2 | 83.2 | 3.0 | 8.3 | 45 |
| FADC | 95.1 | 90.2 | 3.0 | 8.1 | 92 |

通过表1实验结果可知, 选择FADC为卷积模块时, 模型检测性能最优, 在mAP50和mAP50-95分别达到了95.1%和90.2%; 相较于其他卷积网络, 在mAP50上展现出了13.6%、8.5%、6.8%、6.2%、4.9%、8.1%、9.8%和

FLOPs、FPS、Precision和Recall, 分别用于衡量模型的检测性能、速度、准确度及全面性。各个指标的定义和计算方法如下:

1)mAP: 综合反映模型的综合性能与定位精度。计算为各类别AP的平均值, 公式为:

$$AP = \int_0^1 P(R) dR \quad (12)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (13)$$

其中, N 为类别数量, AP 表示 P-R 曲线与横轴围成的面积。

2)FPS: 衡量模型每秒处理的图像帧数。

$$FPS = \frac{Framenum}{ElapsedTime} \quad (14)$$

其中, Framenum 为处理的图片总数; ElapsedTime 为完成所有检测任务的时间。

3)Precision和Recall: 分别衡量正类预测比例和正确检测比例, 计算公式分别为:

$$Precision = \frac{TP}{TP + FP} \quad (15)$$

6.9%的相对优势,在 mAP50-95 上展现出了 18.2%、10.4%、7.4%、4.8%、4.3%、8.3%、9.9%和 7%的相对优势。同时,FADC 在提升检测精度的同时,并未增加参数量和计算量,实现了性能与计算效率之间的美好平衡。FADC 的优势在于能够灵活地根据输入数据的特性调整卷积操作,增强对不同尺度和频率信息的敏感度,从而优化特征提取过程并提升检测精度。因此,本研究选择 FADC 作为首选卷积网络。

3.5 不同损失函数性能实验

不同损失函数对模型预测值与实际值之间差距的度量和优化方向不同,为探究不同损失函数与模型性能之间的关系,实验对比了 CIoU、MP-DIoU、Power-IoU、SD Loss、Focaler-CIoU、Focaler-DIoU、Focaler-GIoU、Slide Loss、Shape-IoU 九种损失函数在自建数据集上的性能情况,不同损失函数对比结果如表 2 所示。

表 2 不同损失函数对比结果

Table 2 Comparison results of different loss functions

| 损失函数 | mAP50/% | mAP50-95/% | Params/M | FPS |
|--------------|-------------|-------------|----------|-----|
| CIoU | 91.9 | 86.9 | 3 | 84 |
| SD Loss | 91.6 | 86.9 | 3 | 58 |
| MP-DIoU | 90.2 | 86.4 | 3 | 59 |
| Powerful-IoU | 91.3 | 86.6 | 3 | 86 |
| Focaler-GIoU | 90.3 | 85.4 | 3 | 69 |
| Focaler-CIoU | 91.9 | 86.8 | 3 | 63 |
| Focaler-DIoU | 91.4 | 85.0 | 3 | 86 |
| Slide Loss | 91.3 | 86.6 | 3 | 85 |
| Shape-IoU | 92.3 | 87.4 | 3 | 85 |

由表 2 可知,在参数量不变的情况下,不同的损失函数对模型的检测性能造成轻微影响。改进模块 mAP50、mAP50-95 和 FPS 分别提升了 0.4%、0.5% 和 1,表明

Shape-IoU 在平衡检测精度和计算效率方面具有优势,能有效降低包裹缺陷的漏检和误检率。

不同损失函数收敛曲线对比图如图 8 所示,可以看出模型采用 Shape-IoU 后损失下降平滑且损失值达到最低,表明模型在学习包裹缺陷特征时更加稳定。这主要是因为 Shape-IoU 损失函数不仅考虑了 GT 框和预测框之间的距离,还涉及了它们的形状和方向相似性,使得模型在训练过程中能够得到更加全面的损失度量,有助于模型更快地调整到正确的预测方向,进而加速模型收敛。

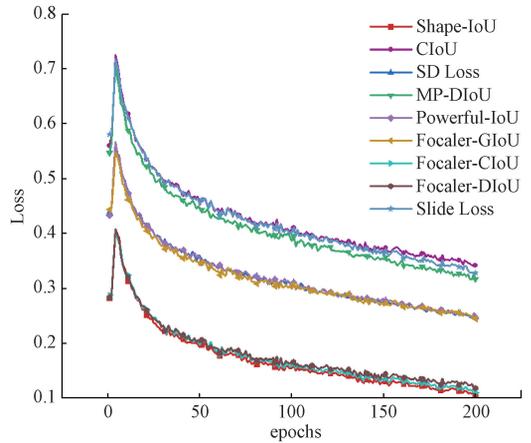


图 8 不同收敛曲线对比图

Fig. 8 Comparison of different convergence curves

3.6 消融对比实验

本研究基于 YOLOv8n 模型进行了 3 项改进,设计了改进 YOLOv8n 模型。为评估每项改进的效果,本文在自制包裹数据集上进行 4 组消融实验。其中,M1 表示将 C2f 模块替换为 C2f-FADC 模块,M2 表示将 SPPF 替换为 SimSPPF,M3 表示将基础模型的 CIoU 替换为 Shape-IoU,相关实验结果如表 3 所示。

表 3 消融实验对比结果

Table 3 Comparison results of ablation experiments

| 模型 | C2f-FADC | SimSPPF | Shape-IoU | mAP50/% | Params | GFLOPs | FPS |
|------------|----------|---------|-----------|---------|-----------|--------|-----|
| YOLOv8n | | | | 91.9 | 3 006 623 | 8.1 | 84 |
| +M1 | ✓ | | | 95.1 | 3 006 623 | 8.1 | 92 |
| +M2 | | ✓ | | 92.8 | 3 007 007 | 8.1 | 87 |
| +M3 | | | ✓ | 92.3 | 3 006 623 | 8.1 | 85 |
| +M1+M2 | ✓ | ✓ | | 96.1 | 3 007 007 | 8.1 | 96 |
| +M1+M3 | ✓ | | ✓ | 95.3 | 3 006 623 | 8.1 | 93 |
| +M2+M3 | | ✓ | ✓ | 93.2 | 3 007 007 | 8.1 | 89 |
| +M1+M2+ M3 | ✓ | ✓ | ✓ | 96.3 | 3 007 007 | 8.1 | 98 |

从表 3 结果可知,将 Backbone 和 Neck 部分的 C2f 模块替换为 C2f-FADC 模块后,mAP 值提升了 3.2%,检测

速度达到了 92 帧,且参数量和计算复杂度保持不变,证明了 FADC 模块通过自适应调整扩张率和卷积核,增强了网

络在多尺度、多频率缺陷检测中的特征提取能力,从而提升了检测精度和速度;引入 SimSPPF 模块替换 SPPF 后, mAP 值提升了 0.9%,检测速度达到了 87 帧,且参数数量和计算复杂度基本保持不变,说明 SimSPPF 简化了空间金字塔池化操作,增强了多尺度特征融合,提升了缺陷检测的精度和速度;将 CIoU 替换为 Shape-IoU 后, mAP 值提升了 0.4%,检测速度达到了 85 帧,且参数数量和计算复杂度保持不变,证明 Shape-IoU 更精准地建模预测框与 GT 框之间的形状和尺度差异,提升了模型的定位精度;同时引入 C2f-FADC 与 SimSPPF, mAP 值提升了 4.2%,进一步增强了多尺度特征融合和细节缺陷的感知能力,提升了模型的检测精度;同时引入 C2f-FADC 与 Shape-IoU, mAP 值提升了 3.4%,这两者的结合提升了模型的定位精度,同

时保留了较高的 FPS,改善了复杂缺陷的检测性能。最终,将上述三部分改进整合进 YOLOv8n 模型后,检测性能表现最佳, mAP 值提升 4.4%、检测速度达到了 98 帧、参数数量和计算量基本不变,检测结果验证了本文算法在检测精度和计算效率之间达到了良好的平衡,验证了这三项改进的综合优势。

3.7 可视化效果对比实验

为了更直观对比本文算法与 YOLOv8n 在快递包裹缺陷检测中的效果差异,本文针对自制数据集的五类缺陷进行了对比实验,并展示了相应的可视化检测效果。如图 9 所示,图 9(a) 为 YOLOv8n 算法的直观检测效果,图 9(b) 为本文算法的直观检测效果。



(a) YOLOv8n算法的直观检测效果
(a) Intuitive detection results of the YOLOv8n algorithm



(b) 本文算法的直观检测效果
(b) Intuitive detection results of the proposed algorithm

图 9 可视化快递包裹缺陷对比图

Fig. 9 Visualization of express package defect comparison

实验结果表明, YOLOv8n 存在定位不精准和误检现象,难以满足中转站缺陷包裹的检测需求。相比之下,本文算法具有更强的浅层特征提取能力,在识别缺陷包裹时能够精准识别缺陷类型,尤其在特征提取、定位精度和误检抑制方面有效提升,展现出明显的性能优势。可视化检测效果对比图还显示,本文算法在提升完整、封口损坏、侧边损坏、包装褶皱、包装污损 5 类缺陷检测方面均表现出色。此外,如图 10 所示,本文算法与 YOLOv8n 在训练过程中分别对比了 mAP50 和 mAP50-95 指标,从图中的数据也可以明显看出,本文算法在这两个指标上均超过了基准模型。

3.8 各网络模型性能对比

为了验证本文算法对于包裹缺陷检测的优异性能,本文在相同的实验条件下,将改进算法与多种先进算法

进行比较,包括 Faster-RCNN、SSD、YOLOv5s、YOLOX-s、YOLOv7-tiny、YOLOv8n、YOLOv9t、YOLOv10n、YOLOv11n 和 YOLOv12n 算法。实验结果如表 4 所示。

从表 4 实验结果表明,本文算法相较于两阶段算法 Faster-RCNN,在 mAP50 上提升了 8.8%,FPS 提升了 89;与 SSD 相比,在 mAP50 上提高了 8.9%,FPS 提升了 73。与近几年 YOLO 先进算法相比,本文算法在 mAP50 分别提高了 9.3%、7.3%、20.8%、5.9%、6%、4.5% 和 5.9%。同时在参数数量和计算量与 YOLOv8n 几乎相同的情况下,本文算法的检测精度提升了 4.4%,FPS 达到了 98 帧,表明该算法在精度与效率间取得了良好平衡。本文算法在检测精度和速度方面优于其他对比算法,证明了其对包裹缺陷检测的价值。

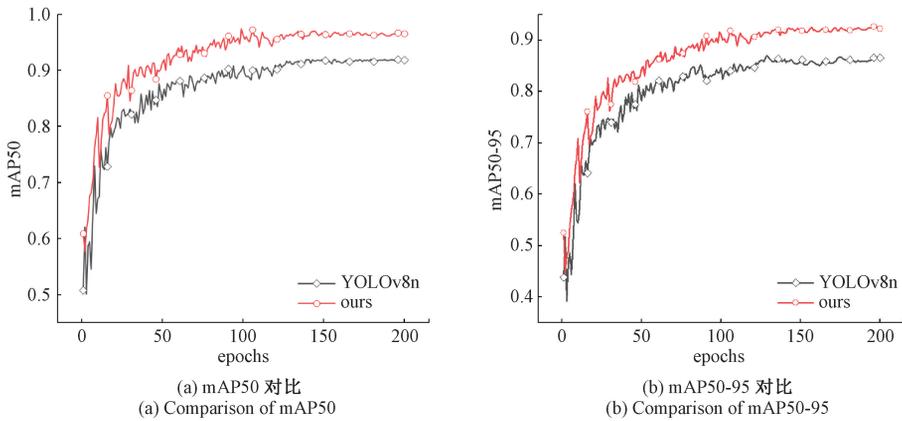


图 10 模型训练精度对比图

Fig. 10 Comparison chart of model training accuracy

表 4 各网络模型性能对比结果

Table 4 Performance comparison results of different network models

| 算法 | completion | seal | broadside | Fold | stained | mAP50/% | Params/M | GFLOPs | FPS |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|----------|--------|-----------|
| Faster-RCNN | 94.5 | 84.3 | 83.0 | 86.5 | 89.2 | 87.5 | 41.3 | 254.1 | 9 |
| SSD(vgg16) | 97.6 | 83.2 | 78.2 | 85.3 | 93.1 | 87.4 | 25.8 | 137.2 | 25 |
| YOLOv5s | 98.1 | 81.4 | 78.3 | 85.9 | 91.7 | 87.0 | 7.02 | 15.8 | 37 |
| YOLOX-s | 93.7 | 90.1 | 85.3 | 87.5 | 88.7 | 89.0 | 9 | 26.77 | 35 |
| YOLOv7-tiny | 95.7 | 64.3 | 66.2 | 74.1 | 77.1 | 75.5 | 6 | 13.1 | 41 |
| YOLOv8n | 98.9 | 89.7 | 86.1 | 89.6 | 95.1 | 91.9 | 3 | 8.1 | 84 |
| YOLOv9t | 98.9 | 85.1 | 85.8 | 89.9 | 92.4 | 90.4 | 2 | 7.6 | 52 |
| YOLOv10n | 98.6 | 88.1 | 82.5 | 89.6 | 92.6 | 90.3 | 2.7 | 6.7 | 53 |
| YOLOv11n | 99.2 | 87.8 | 86.7 | 90.1 | 95.1 | 91.8 | 2.6 | 6.3 | 74 |
| YOLOv12n | 99.1 | 89.3 | 83.0 | 86.6 | 94.1 | 90.4 | 2.6 | 6.5 | 63 |
| 本文算法 | 99.9 | 95.0 | 93.7 | 95.1 | 97.9 | 96.3 | 3 | 8.1 | 98 |

4 结 论

为解决现有包裹缺陷检测模型在检测精度、速度和轻量化之间难以平衡的问题,本文提出一种基于改进YOLOv8n的快递包裹缺陷检测方法。引入频率自适应空洞卷积设计了C2f-FADC,使得网络能够在处理多尺度、多频率缺陷检测任务时灵活调整,优化特征提取过程和提高表征能力,增强了模型的泛化性能和鲁棒性,从而提高了检测精度和推理速度;采用SimSPPF模块替代SPPF模块,增强了对细小缺陷的捕捉能力,并在提高检测效率的同时,保留了高精度的特征提取能力,从而提高了对多尺度缺陷的检测能力。引入Shape-IoU损失函数,替代了传统的回归损失函数,以更精确地建模预测框与GT框之间的形状和尺度差异,显著提升了包裹缺陷的定位精度。为评估本文算法的性能,将其与多种先进算法在相同数据集上进行了对比实验,结果证实本文算法在检测精度和速度上有明显提升,与YOLOv8n的参数数量和计算复杂度基本相同的情况下,实现了检测精度提高了4.4%,检测速度达

到了98帧。因此,本文算法能够实现高精度的快速检测,对提升包裹缺陷检测具有重要影响。该模型目前尚处于实验阶段,未来研究将在保持高精度的情况下,进一步降低模型计算量和参数量,使其能够应用到更多的快递包裹缺陷检测场景中。

参考文献

- [1] 王震洲,杨榕,宿景芳.改进YOLOv8的烟盒缺陷检测算法[J].电子测量技术,2024,47(13):110-119.
WANG ZH ZH, YANG R, SU J F. Improved YOLOv8-based cigarette box defect detection algorithm[J]. Electronic Measurement Technology, 2024, 47(13): 110-119.
- [2] 窦智,高浩然,刘国奇,等.轻量化YOLOv8的小样本钢板缺陷检测算法[J].计算机工程与应用,2024,60(9):90-100.
DOU ZH, GAO H R, LIU G Q, et al. Lightweight YOLOv8 small sample steel plate defect detection algorithm [J]. Computer Engineering and

- Applications, 2024, 60(9): 90-100.
- [3] 周耀威, 孔令军, 李慧刚, 等. 基于通道注意力机制与金字塔池化的包裹破损检测算法[J]. 无线电工程, 2023, 53(11): 2626-2634.
- ZHOU Y W, KONG L J, LI H G, et al. Package damage detection algorithm based on channel attention mechanism and pyramid pooling [J]. Radio Engineering, 2023, 53(11): 2626-2634.
- [4] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [5] GHOSH R. Product identification in retail stores by combining faster R-CNN and recurrent neural network[J]. Multimedia Tools and Applications, 2024, 83(3): 7135-7158.
- [6] SAPKOTA R, AHMED D, KARKEE M. Comparing YOLOv8 and mask R-CNN for instance segmentation in complex orchard environments [J]. Artificial Intelligence in Agriculture, 2024, 13: 84-99.
- [7] 李波, 李志康, 周钰彬. 结合特征融合和注意力机制的 SAR 舰船检测[J]. 电子测量技术, 2024, 47(10): 134-140.
- LI B, LI ZH K, ZHOU Y B. SAR ship detection algorithm combining feature fusion and attention mechanism[J]. Electronic Measurement Technology, 2024, 47(10): 134-140.
- [8] LI X SH, WANG CH L, ZENG ZH. WS-SSD: Achieving faster 3D object detection for autonomous driving via weighted point cloud sampling[J]. Expert Systems with Applications, 2024, 249(1): 123805.
- [9] WANG J X, LIU M, DU Y R, et al. PG-YOLO: An efficient detection algorithm for pomegranate before fruit thinning [J]. Engineering Applications of Artificial Intelligence, 2024, 134(1): 108700.
- [10] HE X J, SONG X N. Improved YOLOv4-tiny lightweight target detection algorithm[J]. Journal of Frontiers of Computer Science & Technology, 2024, 18(1): 138.
- [11] CHENG ZH, GAO L P, WANG Y, et al. EC-YOLO: Effectual detection model for steel strip surface defects based on YOLO-v5[J]. IEEE Access, 2024, 12: 62765-62778.
- [12] WANG CH Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [13] LIANG J B. A review of the development of YOLO object detection algorithm [J]. Applied and Computational Engineering, 2024, 71(1): 39-46.
- [14] HUSSAIN M. YOLOv1 to v8: Unveiling each variant-a comprehensive review of YOLO[J]. IEEE Access, 2024, 12: 42816-42833.
- [15] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [16] 杨明旭, 张俊宁, 张志强, 等. 基于改进YOLOv8的药品泡罩包装缺陷检测算法[J]. 包装工程, 2025, 46(1): 145-154.
- YANG M X, ZHANG J N, ZHANG ZH Q, et al. Defect detection algorithm for blister packaging of tablets based on improved YOLOv8 [J]. Packaging Engineering, 2025, 46(1): 145-154.
- [17] LIU G Y, ZHANG SH, WANG L X, et al. Research on mechanical automatic food packaging defect detection model based on improved YOLOv5 algorithm[J]. PLoS One, 2025, 20(4): e0321971.
- [18] MAO Y ZH, ZHENG SH B, LI L M, et al. Research on rail surface defect detection based on improved CenterNet[J]. Electronics, 2024, 13(17): 3580.
- [19] 王静蕾, 赵英杰, 李峰, 等. 基于改进YOLOv10的盒装食品生产线包装缺陷检测方法[J]. 食品与机械, 2025, 41(4): 236-241.
- WANG J L, ZHAO Y J, LI F, et al. Packaging defect detection method for boxed food production line based on improved YOLOv10 [J]. Food and Machinery, 2025, 41(4): 236-241.
- [20] LI CH F, XU AO, ZHANG Q B, et al. Steel surface defect detection method based on improved YOLOX[J]. IEEE Access, 2024, 12: 37643-37652.
- [21] 周耀威, 孔令军, 戴琪, 等. 基于多头自注意力机制的轻量级包裹破损检测算法[J]. 无线电通信技术, 2025, 51(1): 161-170.
- ZHOU Y W, KONG L J, DAI Q, et al. Lightweight package damage detection algorithm based on multi-head self-attention mechanism [J]. Radio Communication Technology, 2025, 51(1): 161-170.
- [22] LIU SH H, HUANG L M, ZHAO Y B, et al. Lightweight single shot multi-box detector: A fabric defect detection algorithm incorporating parallel dilated convolution and dual channel attention [J]. Textile Research Journal, 2024, 94(1-2): 209-224.
- [23] 米增, 连哲. 面向通用目标检测的YOLO方法研究综述[J]. 计算机工程与应用, 2024, 60(21): 38-54.
- MI Z, LIAN ZH. A review of YOLO methods for general object detection [J]. Computer Engineering

- and Applications, 2024, 60(21): 38-54.
- [24] TALAAT F M, ZAINELDIN H. An improved fire detection approach based on YOLO-v8 for smart cities[J]. Neural Computing and Applications, 2023, 35(28): 20939-20954.
- [25] CHEN L W, GU L, ZHENG D ZH, et al. Frequency-adaptive dilated convolution for semantic segmentation [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 3414-3425.
- [26] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications [J]. ArXiv preprint arXiv: 2209.02976, 2022.
- [27] ZHANG H, ZHANG SH J. Shape-IoU: More accurate

metric considering bounding box shape and scale[J].

ArXiv preprint arXiv:2312.17663, 2023.

作者简介

杨慧敏, 博士, 高级工程师, 主要研究方向为目标检测、智慧物流。

E-mail: 18845636729@163.com

高小雯, 硕士研究生, 主要研究方向为目标检测、智慧物流。

E-mail: 19930531545@163.com

李瑞涛, 硕士研究生, 主要研究方向为目标检测、交通运输。

E-mail: 15229677679@163.com

王汉霞, 硕士研究生, 主要研究方向为目标检测、交通运输。

E-mail: 550704713@qq.com