

DOI:10.19651/j.cnki.emt.2518768

改进 YOLOv8n 的托盘目标检测算法^{*}

刘晓非¹ 薛瑞雷¹ 钟华刚² 刘彦君¹

(1. 新疆大学智能制造现代产业学院(机械工程学院) 乌鲁木齐 830017; 2. 中国机械总院集团青岛分院有限公司 青岛 266300)

摘要: 针对现实工厂环境下,光线不足、障碍物较多等因素的干扰,时常会对托盘造成漏检、误检等问题,提出一种基于改进 YOLOv8n 的托盘目标检测方法。首先,将结合 Transformer 的 BRA 稀疏注意力模块加入到 YOLOv8n 模型的主干网络特征提取环节,以减少障碍物遮挡对托盘检测的干扰;其次,引入 Shape-IoU 损失函数,进一步增强了模型在光线不足以及背景干扰严重情况下对托盘的识别能力;最后,利用基于 GSConv 的 Slim-neck 结构重构 YOLOv8n 的特征融合网络,实现轻量化颈部网络。实验结果表明,改进后的算法在测试集上的平均精度均值达到 89.6%,相较于原模型提升 2.8%,漏检率和误检率分别下降 2% 和 2.2%,有效改善了光线不足和障碍物遮挡情况下托盘识别的漏检和误检问题,同时检测帧率达到 330.52 fps,可以快速精准地进行托盘检测识别,适合部署在智能叉车上,以提高运营效率并提升仓库智能化水平。

关键词: 托盘检测; YOLOv8n; BRA 稀疏注意力模块; Shape-IoU; 基于 GSConv 的 Slim-neck 结构

中图分类号: TP391; TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.20

Improved YOLOv8n algorithm for pallet target detection

Liu Xiaofei¹ Xue Ruilei¹ Zhong Huagang² Liu Yanjun¹

(1. School of Mechanical Engineering Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi 830017, China;

2. China Academy of Machinery Science & Technology Qingdao Branch Co., Ltd., Qingdao 266300, China)

Abstract: Addressing the issues of missed and false detections of pallets in real factory environments, often caused by factors such as insufficient lighting and numerous obstacles, a tray detection method based on an improved YOLOv8n is proposed. Firstly, the Bi-Level Routing Attention (BRA) sparse attention module combined with Transformer is incorporated into the backbone network feature extraction phase of the YOLOv8n model, to reduce the interference from obstacle occlusion on pallet detection. Secondly, the Shape-IoU loss function is introduced, further enhancing the model's ability to recognize pallets in conditions of insufficient lighting and severe background interference. Finally, the feature fusion network of YOLOv8n is reconstructed using the GSConv-based Slim-neck structure, achieving a lightweight neck network. Experimental results indicate that the improved algorithm achieves a mean Average Precision (mAP) of 89.6% on the test set, representing a 2.8% improvement compared to the original model. The missed detection rate and false detection rate decrease by 2% and 2.2%, respectively. This effectively mitigates the problems of missed and false detections of pallets in situations of insufficient lighting and obstacle occlusion. Additionally, with a detection frame rate of 312.5 fps, the method enables rapid and accurate pallet detection and recognition, making it suitable for deployment on smart forklifts to enhance operational efficiency and elevate warehouse intelligence levels.

Keywords: pallet detection; YOLOv8n; BRA sparse attention module; Shape-IoU; GSConv-based Slim-neck structure

0 引言

在智能制造体系中,无人工业车辆是仓库运作的核心组件,负责托盘的自动化运输任务。但在实际操作中,这些

车辆主要因障碍物密布、光线不足等干扰,而遭遇漏检、误检等问题^[1-2]。托盘识别技术作为无人工业车辆的关键所在,对于提升搬运效率具有重大意义。当前,如何提高托盘识别的精度与速度,已成为亟待攻克的技术难题^[3]。

收稿日期:2025-05-08

*基金项目:新疆维吾尔自治区自然科学基金(2022D01C391)、天山英才培养计划(2022TSYCLJ0044)项目资助

传统托盘图像检测依靠手工构造被检测对象的特征^[4-5],例如,周涛等^[6]依据托盘的尺寸大小,并结合托盘的颜色信息以及底部边角点的特征,运用一种从宽泛到精确的检测方法来识别托盘。但这种方法需要提取大量的特征样本才能实现对托盘的有效识别,耗费大量的时间成本与人力成本,同时算法鲁棒性差,难以实际应用。

近年来,计算机技术以惊人的速度飞速发展,与此同时,深度学习理论也在不断成熟和完善^[7]。深度学习算法能够高效排除无关信息,显著减少检测识别所需处理的数据量,从而加快识别速度。在这样的背景下,越来越多的研究者开始利用深度学习方法研究和实现对托盘的检测识别。目标检测算法分两类:两阶段方法如 R-CNN、Fast-RCNN,精度高但计算量大、速度慢,难以满足实时需求。Mohamed 等^[8]利用区域 Faster RCNN 模型与卷积神经网络级联来实现托盘识别与检测。该方法通过深度挖掘图像特征,增强了自动引导叉车在光线不足情况下对托盘的识别效果。单阶段方法如 YOLO、SSD,无需候选框,单次前向传递直接预测,计算量小、速度快,适合移动终端部署。单阶段方法兼具高识别精度和速度的优势,因此它在实际研究中得到了广泛应用。王平凯等^[9]提出了一个基于改进 YOLOv3 网络的托盘识别与检测方案,采用 K-Means+十聚类对 Anchor Box 进行重新聚类,以更好地适应托盘检测的需求。该方案在一定程度上减轻了阴暗条件下托盘信息模糊对识别精度造成的影响;Rocha 等^[10]提出了一种深度学习模型优化与多目标评估方法来对托盘进行检测识别。针对 YOLOv5+UNet 在自定义托盘数据集上自动优化超参数,混合模型在有障碍物遮挡的情况下表现良好;肖施睿^[11]利用改进 YOLOv5s 网络对托盘进行检测。采用剪枝技术简化模型结构并针对微小物体增设专用识别模块,优化了检测系统的性能指标;刘继泽^[12]在原模型基础上引入了自适应目标区域提取网络(adaptive target area extraction network,At)和强化的多尺度特征融合网络结构(enhanced multi-scale feature fusion network structure,Em),改进后的算法 YOLO-AtEm 增强了托盘特征表示。胡海霞等^[13]提出一种改进 YOLOv8n-CGDG 的托盘识别检测方法,用 C2f-SCConv 替换原 C2f 模块,同时引入 DAT 特征选择模块并采用 GIoU 评估预测框与真实框的匹配程度,有效提升了改进模型对被遮挡目标的敏感度。Vu 等^[14]在 YOLOv8n 网络中添加了几何感知的视觉特征重加权 GAV-FR 模块来增强托盘目标区域特征响应并抑制背景干扰,新模型 YOLOv8n-GAVFR 显著提高了精度。李泽胤等^[15]通过将 SimAM 注意力机制与 Inner-CIoU 损失函数引入 YOLOv8n 模型,解决了复杂背景下目标检测的漏检和误检问题。

上述研究均基于神经网络实现托盘识别,且都进行了算法的优化,使模型精度在一定程度上得到了提升。但这些方法在光线不足或障碍物遮挡下的托盘识别效果较为一

般或计算量相对较大,导致识别过程较缓慢,因此达不到工厂环境中对识别速度和准确度的要求。

为优化上述问题,本研究基于 YOLOv8n 网络(you only look once version 8 nano)^[16],首先,在 YOLOv8n 模型的主干网络特征提取环节,引入结合 Transformer 的 BRA (bi-level routing attention)稀疏注意力模块,以降低障碍物遮挡对托盘检测造成的不利影响。其次,采用 Shape-IoU 损失函数,进一步增强了模型对外来目标的识别能力,提高特征提取的深度与广度。最后,利用基于 GSConv 的 Slim-neck 架构对 YOLOv8n 的特征融合网络进行重新设计,以达到构建轻量级颈部网络的目的。根据以上贡献,将改进后的算法命名为 YOLOv8n-BSG,B 表示 BRA 稀疏注意力模块,S 表示 Shape-IoU 损失函数,G 表示基于 GSConv 的 Slim-neck 架构。新算法在自制数据集中平均精度均值提升了 2.8%,漏检率和误检率分别下降 2% 和 2.2%,检测帧率达到 330.52 fps,有效改善了现实工厂下光线不足和障碍物遮挡对托盘造成的漏检和误检问题,满足现场使用要求。

1 YOLOv8n 算法简介

YOLOv8n 在 YOLOv5^[17]的坚实基础上进行了全面的扩展与优化,构建了一个集对象检测、实例分割及图像分类功能于一体的统一训练框架。该框架在继承 YOLOv5 目标检测方面的高效能特性的基础上,还进一步引入了实例分割与图像分类的新模块。其核心架构如图 1 所示,YOLOv8n 在主干网络部分进行了多项优化:通过整合 CBS 单元、C2f 组件以及 SPPF 模块构建新型特征提取体系,其中特别借鉴了 YOLOv7 的 ELAN 拓扑结构理念。相较于前代版本,YOLOv8n 用梯度更丰富的轻量级 C2f 模块替代原有 C3 组件,并依据模型规模差异实施通道参数动态调节机制以适配各类目标检测任务的需求。

特征融合阶段,在保留 YOLOv5 的 PAN-FPN 基础上,YOLOv8n 去除了特征图升维阶段的卷积操作,并将网络中传统的 C3 模块全面升级为具备梯度流优化的 C2f 结构,以进一步提升效率。在 Head 部分,YOLOv8n 对解耦头结构进行了革新,实现了分类与检测头的分离,这一改变增强了模型的灵活性和可扩展性。此外,YOLOv8n 还将传统的 Anchor-Based 检测机制升级为 Anchor-Free 架构,从而摒弃预设锚框依赖,简化了模型设计。

为了进一步优化检测精度和鲁棒性,YOLOv8n 采用任务对齐分配器(task aligned assigner)优化训练过程中的样本匹配机制,并引入了分布聚焦损失(distribution focal loss)作为监督信号,通过概率分布建模实现样本权重的自适应平衡,以实现更精准的样本学习与优化。综上所述,本文决定在 YOLOv8n 的基础上进行进一步的改进与优化,以充分发挥其强大的功能与性能优势。

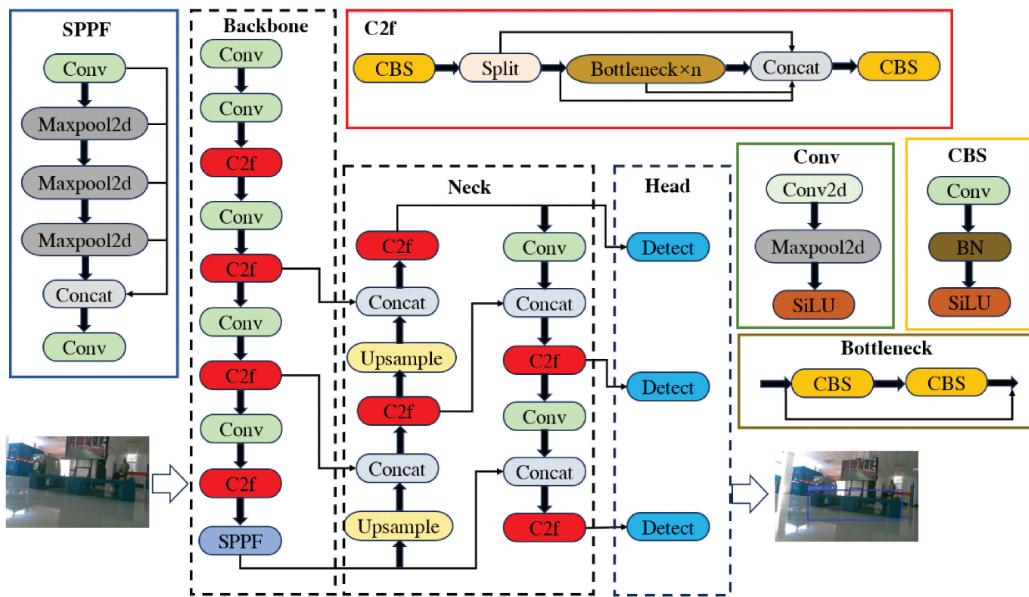


图 1 YOLOv8n 网络结构

Fig. 1 YOLOv8n network structure

2 改进 YOLOv8n 模型

尽管 YOLOv8n 算法在检测精度与速度方面表现出色^[18],然而在面对复杂场景中的托盘目标时,因光线不足,且极易遭受遮挡等问题的干扰,致使该算法常常出现漏检、误检等情况。鉴于此,本文拟从 3 个方面对模型加以改进。

1) 在 YOLOv8n 骨干网络的第 9 层,即 C2f 和 SPPF 之间增加一个 BRA_nchw 模块,可优先筛选 C2f 输出的关键中级语义特征,避免冗余信息干扰金字塔融合,以增强网络对通道信息的长期依赖,从而提升模型在障碍物遮挡场景下对托盘检测的精确度。

2) 鉴于数据集中存在光线不足以及托盘背景干扰严重等样本质量较低的情况,为优化模型性能,特引入 Shape-IoU 损失函数。该损失函数能够在训练过程中平衡不同质量样本的贡献权重,从而加速模型的收敛并提高算法的泛化性能,使得模型在应对各类托盘检测任务时展现出更强的稳定性与适应性。

3) 利用 GSConv(group shuffle convolution)轻量级卷积模块和 VoV-GSCSP(group shuffle cross stage Partia)轻量级集成模块重构原 YOLOv8n 模型的特征融合网络,在几乎没有精度损失的情况下降低了模型的复杂度。

2.1 添加 BRA_nchw 注意力模块

在工厂环境里,托盘经常受到各种障碍物的遮挡,成为了目标检测任务面临的难题。为了解决这一问题,本文提出了一种改进的 transformer 自注意力机制 BRA_nchw^[19],改善了由于障碍物遮挡导致的漏检和误检问题。在 YOLOv8n 网络对输入托盘图像进行特征提取的过程

中,随着网络的深入,深层特征图变得越来越抽象,这可能会导致浅层网络的位置信息和深层网络的语义信息之间的依赖性减弱。特别是对于工厂内部受障碍物遮挡的托盘,随着网络的深入,很容易弱化整体特征,导致托盘识别出现漏检或误检。

自注意力机制是 transformer 的核心模块,它可以帮助网络捕捉远距离的上下层依赖关系。这种结构存在内存消耗高和计算成本高的缺点。因此,本文提出一种新的稀疏动态注意力机制,即 BRA_nchw 注意力机制,该算法首先在粗粒度区域过滤掉最不相关的键值对,然后计算剩余区域中的 token-to-token 注意力。该操作通过引导每个查询去处理与其语义信息最为相关的键值对,可以实现更灵活的计算力分配。BRA_nchw 自注意力模块示意图如图 2 所示(图 2 中 mm 表示矩阵相乘)。

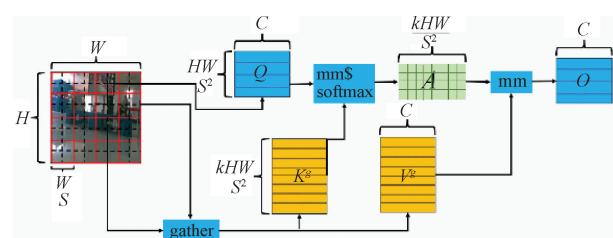


图 2 BRA_nchw 自注意力模块示意图

Fig. 2 The schematic chart of BRA_nchw self-attention module

1) 假设输入的托盘图像 $X \in R^{H \times W \times C}$,首先将其划分为 $S \times S$ 不同的区域,因此,每个区域包含 $\frac{HW}{S^2}$ 向量。也就是说,将 X 转换成 $X^r \in R^{\frac{S^2 \times HW}{S^2} \times C}$,那么 $Q, K, V \in$

$R^{S^2 \times \frac{HW}{S^2} \times C}$ 能够通过线性映射获得,如式(1)所示。

$$\begin{cases} Q = X^r W^q \\ K = X^r W^k \\ V = X^r W^v \end{cases} \quad (1)$$

式中: $W^q, W^k, W^v \in R^{C \times C}$ 分别为 query、key、value 的线性映射权值。 W^q, W^k, W^v 可以在模型训练过程中不断更新。

2) 构建有向图结构定位区域间的强关联节点。首先对各区域的 Q 和 K 特征进行均值处理,分别提取 Q^r 和 K^r 参数 ($Q^r, K^r \in R^{S^2 \times C}$),然后基于这两个特征向量构建表征区域关联强度的邻接矩阵,计算公式如式(2)所示。

$$A^r = Q^r (K^r)^T \quad (2)$$

保留每个区域的前 k 个连接来修剪关联图,并将前 k 个连接的索引保存到路由索引矩阵 $I^r \in N^{S^2 \times k}$ 。其公式如式(3)所示 (I^r 的第 i 行包含第 i 个区域的前 k 个最相关区域的索引)。

$$I^r = topkIndex(A^r) \quad (3)$$

3) 跨区域路由索引矩阵 I^r 用于量化分析细粒度单元间(token-to-token)注意力关联。具体来说,区域 i 中的每个 query 元素对 k 个路由区域中的所有键值对执行矩阵运算,结果分别记录为 $I_{(i,1)}^r, I_{(i,2)}^r, \dots, I_{(i,k)}^r$ 。首先需要采集键值张量,如式(4)所示。

$$\begin{cases} K^s = gather(K, I^r) \\ V^s = gather(V, I^r) \end{cases} \quad (4)$$

其中, K^s 和 V^s 分别是采集后的键和值的张量。然后对收集到的键值对进行 Attention 操作,如式(5)所示。

$$O = Attention(Q, K^s, V^s) + LCE(V) \quad (5)$$

$LCE(V)$ 是一个局部上下层增强器,它能够用深度可分离卷积来参数化^[20]。

2.2 损失函数优化

目标检测算法的性能深受损失函数的影响^[21]。传统的损失函数如 CIoU、GIoU、EIoU 和 DIoU,通常仅考虑 GT(真实边界)框与预测边界框之间的几何关系。损失是根据边界框之间的相对位置和相对形状来计算的,而在进行边界框回归的过程中,边界框本身的固有属性(如形状和尺度)往往被忽略。在托盘的检测识别中,由于光线不足、障碍物遮挡和物体形状的可变性等因素,原有的 YOLOv8n 模型检测性能较差,为了解决这一问题,本文将原始模型的 CIoU^[22]损失函数替换为 Shape-IoU^[23]损失函数,如图 3 所示。

Shape-IoU 的计算考虑了数据集中目标对象的尺度因子,并引入了一个与真实边界框(ground truth, GT)形状相关的权重系数。这些权重系数用于在计算损失函数时调整水平和竖直方向上的偏差的影响。该损失函数侧重于根据边界框自身独有的形状特性以及尺度大小来计算损失,因此,在目标检测任务中,该方法能够显著提高目标边界框定位的精确性,从而提高边界框回归性能。其计算

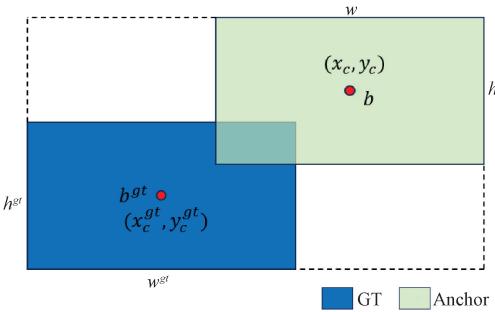


图 3 Shape-IoU 损失函数

Fig. 3 Shape-IoU loss function

公式如下:

$$W_w = \frac{2 \times (w^g)^s}{(w^g)^s + (h^g)^s} \quad (6)$$

$$h_h = \frac{2 \times (h^g)^s}{(w^g)^s + (h^g)^s} \quad (7)$$

$$d^s = h_h \times \frac{(x_c - x_c^g)^2}{c^2} + W_w \times \frac{(y_c - y_c^g)^2}{c^2} \quad (8)$$

$$\Omega^s = \sum_{t=w, h} (1 - e^{-w_t})^\theta \quad (9)$$

$$\begin{cases} \omega_w = h_h \times \frac{|w - w^g|}{\max(w, w^g)} \\ \omega_h = W_w \times \frac{|h - h^g|}{\max(h, h^g)} \end{cases} \quad (10)$$

式中: s 表示尺度因子,它与数据集中检测到的目标的尺度有关。 w^g 和 h^g 分别对应 GT 边界框的长度和宽度, (x_c^g, y_c^g) 对应 GT 边界框的中心坐标。 w 和 h 分别为锚框的长度和宽度, (x_c, y_c) 为其中心坐标。 W_w 和 h_h 分别为水平方向和竖直方向上的权系数,其值与 GT 边界框的形状有关。 d^s 和 Ω^s 分别表示距离和形状损失分量。 θ 反映了对形状损失的关注,防止以预测框的移动为代价过度强调形状损失。Shape-IoU 边界框回归损失如下:

$$L_{Shape-IoU} = 1 - L_{IoU} + d^s + \frac{\Omega^s}{2} \quad (11)$$

2.3 基于 GSConv 的轻量级颈部网络 Slim-neck

自动引导叉车叉取托盘的效率取决于算法模型的运行效率。除了检测精度之外,检测速度也是影响算法模型运行效率的关键因素。一般来说,拥有较少参数量(Parameters)和浮点运算次数(GFLOPs)的模型将更有利干算法的快速运行。为了得到检测精度高、检测速度快的轻量化模型,最终实现在计算能力有限的边缘设备上部署,本研究使用 GSConv 模块^[24]对 YOLOv8n 模型的颈部网络进行重构。GSConv 的结构图如图 4 所示。

首先对输入特征图进行普通卷积降采样,然后进行深度卷积(DWConv),并将两次卷积的结果进行拼接。最后,执行特征重排序 Shuffle 操作,将前两次卷积操作中紧密相关的语义信息的通道重新组合在一起。GSConv 的目的是使深度可分离卷积(DSC)的输出尽可能接近普通卷积。

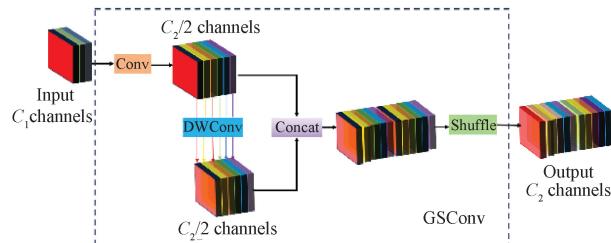


图 4 GSConv 卷积模块结构图

Fig. 4 The structure chart of GSConv convolution module

它采用特征重排序 Shuffle 操作将普通卷积生成的信息渗透到 DSC 卷积生成的信息的每一部分中,在保持较高精度的同时降低了模型复杂度。这种方法可以使普通的卷积信息完全混合到 DSC 卷积的输出中,更好地平衡模型的精度和速度。本研究使用轻量级卷积 GSConv 代替普通卷积,其计算成本约为普通卷积的 60%~70%,但对模型学习能力的贡献与普通卷积相似。接下来,本文基于 GSConv 引入新的 GSbottleneck 模块,如图 5(a)所示。利用 GSbottleneck 模块重构 YOLOv8n 模型的 C2f 模块,所提出的模块命名为 VoVGSCSP,如图 5(b)所示。

GSConv 模块与 VoVGSCSP 轻量化结构的合理结合,

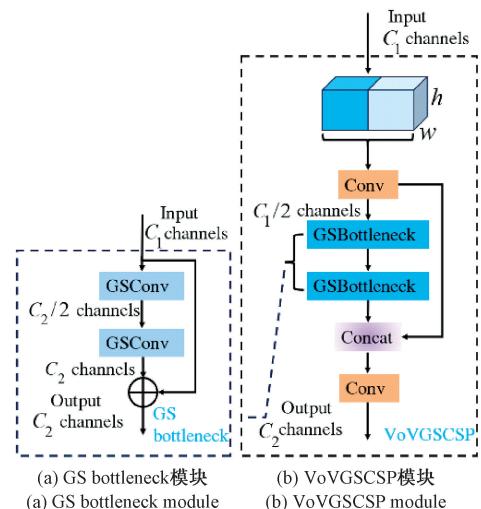


图 5 Slim-neck 网络结构图

Fig. 5 The structure chart of Slim-neck network

可以在保证精度损失最小或不损失的情况下实现轻量化(这种组合结构称为 Slim-neck 网络)。本研究采用 Slim-neck 重建 YOLOv8n 的颈部网络,一方面降低了 YOLOv8n 模型的复杂度;另一方面,保持了托盘识别的高检测精度。改进后的模型结构如图 6 所示。

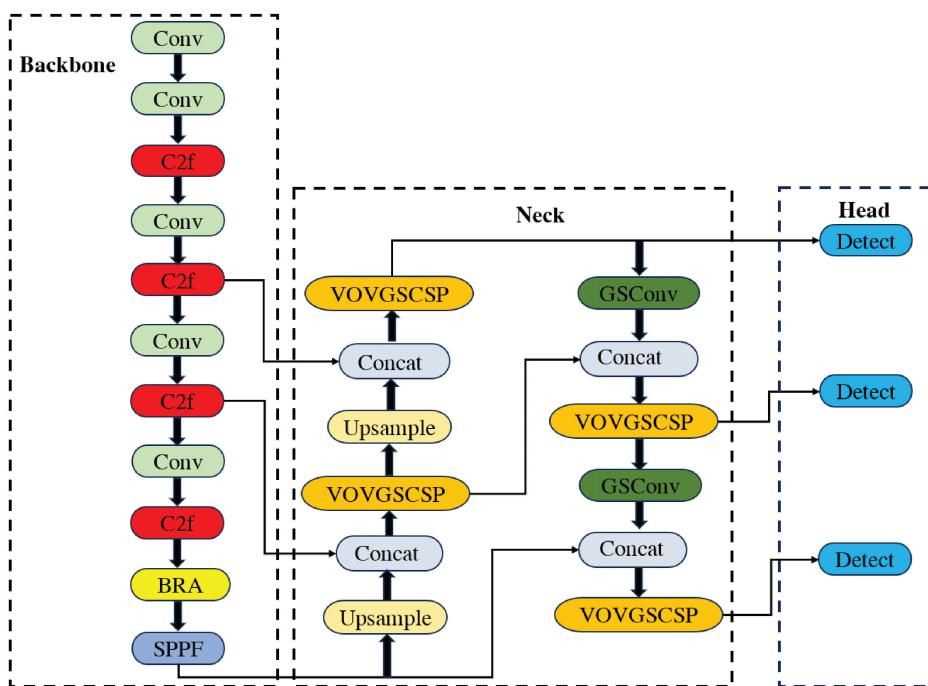


图 6 YOLOv8n-BSG 模型架构

Fig. 6 Network structure of YOLOv8n-BSG

3 实验结果和分析

3.1 实验环境

表 1 为本文实验环境配置,模型训练需历经 300 轮迭代

,使用的优化算法是 SGD。具体超参数设置包括:批量大小设置为 16,起始学习率设为 0.01,动量参数为 0.937,以及权重衰减参数为 0.0005。动量机制能够有效提升随机梯度下降(SGD)算法在特定优化方向上的收敛效率,同

时显著降低训练过程中的参数波动。具体而言, 较大的动量参数有助于加速模型在训练过程中向最优解方向收敛; 适度的权重衰减系数(L2 正则化)能够有效控制模型复杂度, 从而防止过拟合现象的发生。学习率的设置需要权衡收敛速度与稳定性: 较大的学习率虽能有效缩短训练周期, 却可能引发梯度下降路径的振荡现象, 进而影响收敛轨迹稳定性, 在极端情况下甚至导致算法无法准确锁定全局最优解区域; 而过小的学习率则会显著降低收敛效率。此外, 迭代次数的设置也需谨慎: 过多的迭代次数会增大模型过拟合的风险, 而过少的迭代次数则会导致模型无法充分学习数据特征。

表 1 实验环境参数

Table 1 Experimental environment parameters

名称	配置
操作系统	Windows11
CPU	Intel(R)Core(TM)i9-13700H
GPU	GeForce RTX4060
学习框架	pyTorch2.0.0
加速环境	CUDA11.8
编程语言	Python3.8

3.2 数据集及评价指标

由于缺乏官方托盘数据集, 本文在不同光照条件(正常、光线不足)、不同场景(实验工厂和车间工厂)下, 使用不同材质的托盘(塑料、木质)共拍摄了 3 000 张托盘图片, 其中还包含了一些受遮挡的托盘图片集, 如图 7 所示。数据集包含正常托盘图片 1 500 张, 以及在光线不足和障碍物遮挡条件下拍摄的托盘图片 1 500 张。并按照 8:1:1 的比例划分为训练集、验证集和测试集, 并使用“labelimg”对图片进行标注, 仅设定一个类别, 命名为“pallet”。

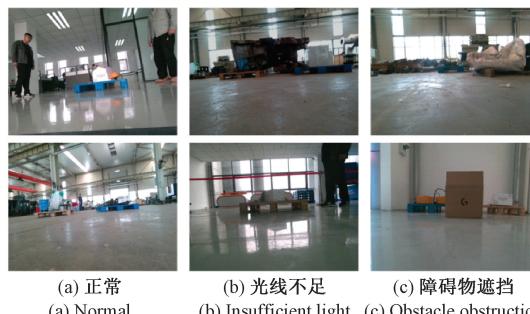


图 7 部分样本数据

Fig. 7 Partial sample data

本文采用漏检率(miss rate, MR)、误检率(false alarm rate, FR)、平均精度均值(mean average precision, mAP), 本试验的评价标准为 mAP@0.5, 即将 IoU 阈值设置为 0.5 时所求得的平均精度均值, 参数量(Parameters)以及浮点运算次数(GFLOPs)作为实验的评价指标。部分计算

公式如下:

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$AP = \int_0^1 P_{(r)} dr \quad (14)$$

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (15)$$

$$MR = \frac{M_{(IL+OO)}}{A_{(IL+OO)}} \quad (16)$$

$$FR = \frac{F_{(IL+OO)}}{A_{(IL+OO)}} \quad (17)$$

式中: TP 表示在所有的检测样本当中, 正确的样本被算法检测为正确样本的个数; FP 表示真实结果为错误但是算法预测结果为正确, 即错误的样本但是被算法检测为正确的样本; FN 是真实结果为错误并且算法预测结果也为错误, 即错误的样本被算法检测为错误的样本。 P 表示在所有的检测样本当中, 正确的样本被算法检测为正确样本的个数与所有检测物体的比值。 R 代表正确的样本被算法检测为正确样本的个数与测试集的比值。 AP 为平均精度, 它是基于 P/R 值曲线与坐标轴所围成的面积使用积分计算后得到的。 C 表示类别数目, mAP 是对 AP 进行加权平均得到的, 由于托盘识别是单目标分类问题, 托盘的平均精度均值即为平均精度。 IL 是光线不足下的托盘图片, OO 是障碍物遮挡下的托盘图片, $M_{(IL+OO)}$ 表示在所有的检测样本当中已确认的漏检图片里, 光线不足和障碍物遮挡的托盘图片的个数。 $F_{(IL+OO)}$ 表示在所有的检测样本当中已确认的误检图片里, 光线不足和障碍物遮挡的托盘图片的个数。 $A_{(IL+OO)}$ 表示在所有的检测样本当中, 光线不足和障碍物遮挡的托盘图片的个数。

3.3 不同损失函数的对比

为了评估 Shape-IoU 损失函数的性能, 于相同的实验条件下, 选择现今流行的损失函数 GIoU、DIoU、EIoU 和 WIoU, 并将其分别集成到 YOLOv8n 模型中, 替换原模型本身的 CIoU 损失函数。实验对比结果如表 2 所示, 数据显示, 当模型边框损失函数为 Shape-IoU 时, 模型的精度、漏检率和误检率均得到了更显著的改善。这表明 Shape-IoU 损失函数通过其智能的权重调整机制, 可更为有效地处理图像质量差异带来的梯度波动, 减少特征存在缺陷的托盘样本所引发的负面影响, 增强模型对托盘检测任务的适应性。

3.4 注意力机制替换对比

为充分展现所选 BRA_nchw 注意力模块的优越性, 本文对比实验了几种当下流行的注意力机制, 包括 CBAM、SE、SimAM、GAM。

实验结果如表 3 所示, 数据表明, 增加所选用的几种注意力机制普遍能够改善原有模型的精度、漏检率和误检

表2 损失函数对比实验

Table 2 Loss function comparison experiment %

损失函数	MR	FR	mAP
CIoU(v8n)	8.2	3.8	86.8
DIoU	8.0	3.5	87.0
GIoU	8.0	3.1	87.2
EIoU	7.7	3.2	87.5
WIoU	7.5	3.0	87.9
Shape-IoU	7.3	2.6	88.1

率,其中BRA_nchw注意力模块对托盘检测性能的正向影响最为显著。具体来说,BRA_nchw注意力模块的引入使得模型精度提高了1.7%,漏检率和误检率分别下降了1.2%和1.1%,同时其计算量为8.6 G,参数量为3.35 M。CBAM和SE注意力机制的计算量和参数量相似,相较于BRA_nchw注意力模块,精度分别落后了0.8%和1.2%,漏检率分别增加了0.6%和1.1%,误检率分别增加了0.3%和0.6%。虽然SimAM注意力机制在计算量和参数量上更少,但精度仅提升0.1%,漏检率和误检率分别下降了0.3%和0.5%。与BRA_nchw注意力模块相比,GAM注意力机制的精度降低了0.3%,漏检率和误检率均增高了0.2%,同时其计算量和参数量分别增多2.7 G和4.01 M。从整体上权衡检测精度、漏检率和误检率以及计算资源的投入,BRA_nchw注意力模块在保持较高精度、较低漏检率和误检率的同时,还能有效控制计算量和参数量,从而确保检测识别速度的稳定性,因此更具优势。

表3 注意力机制对比实验

Table 3 Comparative experiment of attention mechanism

模型	MR/ %	FR/ %	mAP/ %	GFLOPs/ G	Params/ M
YOLOv8n	8.2	3.8	86.8	8.1	3.01
CBAM	7.6	3.0	87.7	8.3	3.06
SE	8.1	3.3	87.3	8.2	3.05
SimAM	7.9	3.3	86.9	8.2	3.03
GAM	7.2	2.9	88.2	11.3	7.36
BRA_nchw	7.0	2.7	88.5	8.6	3.35

3.5 消融实验结果

为了充分证实本文所提出的一系列改进模块方法的有效性,采用了YOLOv8n作为基础模型,在统一实验条件下逐步添加BRA_nchw注意力模块、Shape-IoU损失函数和GSConv+VoV-GSCSP模块,开展消融实验,相关的实验数据如表4所示(A表示BRA_nchw注意力模块,B表示Shape-IoU损失函数,C表示GSConv+VoV-GSCSP模块)。表4中“√”标记意味着相应的模块已被纳入至模型中。模型2、3和4分别对原始模型增加BRA_nchw注意力模块,使用Shape-IoU损失函数替换原有的损失函数,引入GSConv+VoV-GSCSP模块,在此过程中均为单一模块的改动。模型5同时融入BRA_nchw和Shape-IoU损失函数,模型6同时融入Shape-IoU损失函数和GSConv+VoV-GSCSP模块。模型7同时融入BRA_nchw模块和GSConv+VoV-GSCSP模块。模型8同时融入三个模块。由实验数据可知,在相同的试验条件下,添加BRA_nchw模块,模型漏检率和误检率分别下降1.2%和1.1%,平均精度均值上升1.7%。更换损失函数为Shape-IoU后,模型漏检率和误检率分别下降0.9%和1.2%,平均精度上升1.3%。轻量级网络GSConv+VoV-GSCSP模块单独使用会造成一定的精度损失,但计算量和参数量分别降低了0.8 G和0.25 M。在模型5中,当BRA_nchw注意力模块和Shape-IoU损失函数共同引入YOLOv8n原模型时,其精度、漏检率和误检率相比于模型2和3均有所改进,且计算量和参数量与模型2相同,此结果表明,Shape-IoU损失函数的融入未增加模型的计算负担。引入BRA_nchw注意力模块与Shape-IoU损失函数,分别结合GSConv+VoV-GSCSP模块至YOLOv8n原模型中,均提升了模型的精度,降低了漏检率或误检率,参数量与计算量却明显减少,表明轻量级网络模块的加入减轻了模型改进过程中的复杂性且没有对模型的检测性能产生负面影响,也提高了检测速度。相比YOLOv8n模型,改进后的模型漏检率和误检率分别下降2%和2.2%,平均精度上升2.8%。使用原YOLOv8n算法和本文改进的YOLOv8n-BSG算法对相同的图像进行检测,检测效果如图8所示。从检测效果上看,改进后的模型漏检和误检减少,且成功识别遮挡严重的托盘。

表4 YOLOv8n消融实验

Table 4 YOLOv8n ablation experiment

模型	A	B	C	MR/%	FR/%	mAP/%	GFLOPs/G	Params/M
模型-1(v8n)	×	×	×	8.2	3.8	86.8	8.1	3.01
模型-2	√	×	×	7.0	2.7	88.5	8.6	3.35
模型-3	×	√	×	7.3	2.6	88.1	8.1	3.01
模型-4	×	×	√	8.3	3.7	86.5	7.3	2.76
模型-5	√	√	×	6.0	1.9	89.7	8.6	3.35
模型-6	×	√	√	7.2	2.8	88.2	7.3	2.76
模型-7	√	×	√	7.1	2.5	88.5	8.2	3.04
模型-8	√	√	√	6.2	1.6	89.6	8.2	3.04

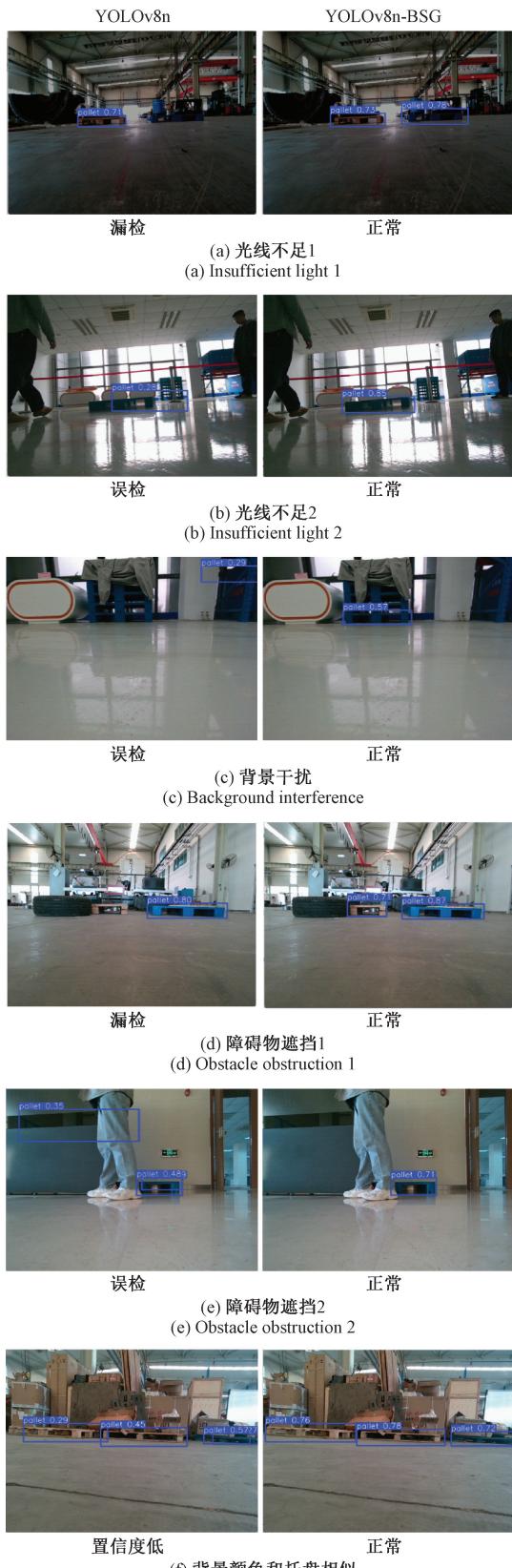


图 8 两种模型检测效果对比

Fig. 8 Comparison of detection effects between two models

3.6 对比实验

本文采用对比实验对基准 YOLOv8n 模型、设计的 YOLOv8n-BSG 增强型算法、主流深度检测网络和引言中提到的 YOLO-AtEm、YOLOv8n-CGDG、YOLOv8n-GAVFR 进行横向性能评估。实验结果如表 5 以及图 9 和图 10 所示。

表 5 主流算法对比实验结果

Table 5 Mainstream algorithm comparison experimental results

模型	MR/ %	FR/ %	mAP/ %	GFLOPs/ G	Params/ M
Faster-RCNN	19.5	19.8	71.2	150.2	54.27
SSD	17.7	19.3	72.6	111.3	28.32
YOLOv5s	7.5	3.1	88.2	16.0	7.05
YOLOv7-tiny	5.8	1.4	89.9	13.5	6.12
YOLOv8n	8.2	3.8	86.8	8.1	3.01
YOLOv8s	6.9	2.1	88.7	11.07	27.60
YOLO-AtEm ^[12]	7.6	2.9	88.3	5.6	16.45
v8n-CGDG ^[13]	6.4	2.0	89.3	9.1	4.00
v8n-GAVFR ^[14]	7.8	3.5	87.9	11.7	3.85
YOLOv10n	7.0	3.4	88.1	7.6	2.70
YOLOv11n	6.8	3.3	88.3	6.5	2.60
Ours	6.2	1.6	89.6	8.2	3.04

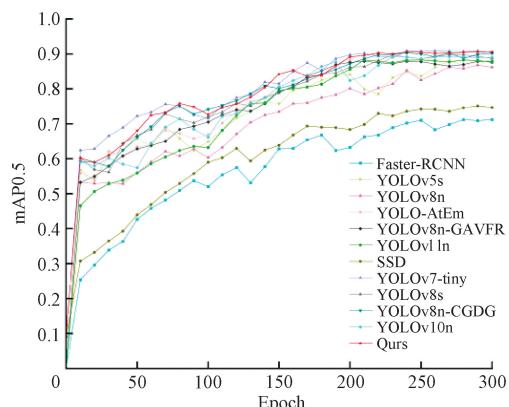


图 9 不同模型 mAP@0.5 曲线对比

Fig. 9 Comparison of mAP@0.5 curves among different models

实验对比分析表明,本文提出的 YOLOv8n-BSG 算法相较于 SSD^[25]、Faster-RCNN^[26] 两种经典检测算法,在计算量、参数量和精度方面均表现出超越的效果。YOLOv8n 和 YOLOv8s 是 YOLOv8 的不同版本,相较于轻量化的 YOLOv8n, YOLOv8s 通过扩展网络深度和宽度,显著增强了模型的多尺度特征表征能力。将 YOLOv8n-BSG 算法与 YOLO 系列轻量化模型(YOLOv5s、YOLOv7-tiny(SiLU)^[27] 和 YOLOv8s^[28] 算法)

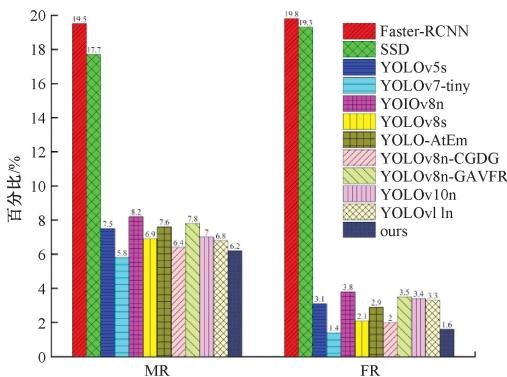


图 10 不同模型漏检率和误检率对比

Fig. 10 Comparison of MR and FR among different models

进行比较, YOLOv8n-BSG 算法在计算量和参数量方面均显著优于对比算法, 其计算量和参数量甚至不到其余算法的一半, 同时在检测精度方面, YOLOv8n-BSG 算法在轻量级网络中保持了较高的水平。YOLOv8n-CGDG 算法虽然在漏检率、误检率和精度方面逼近本文改进的 YOLOv8n-BSG 算法, 但其计算量和参数量相比 YOLOv8n-BSG 算法分别增加了约 1 G 和 1 M。YOLOv11n^[29]是 YOLOv11 的一个超小型版本, 它在对比算法中拥有较少的计算量和参数量, 但其漏检率和误检率较 YOLOv8n-BSG 算法分别增加 0.6% 和 1.7%, 精度落后 1.3%。YOLOv7-tiny(SiLU) 算法在这些算法中展现了最高的精度、最低的漏检率和误检率, 然而, 本文改进的算法在仅略微牺牲漏检率、误检率和精度的条件下, 显著降低了模型计算量和参数量。以上结果表明, YOLOv8n-BSG 算法在模型精度和检测速度方面均取得了令人满意的效果, 在同类算法中具有较强的竞争力。

3.7 泛化性实验

为了研究 YOLOv8n-BSG 在实际部署场景中的泛化性能, 本研究在公司附近其他工厂里(与前述拍摄数据集的环境不同)实地拍摄了 100 张在光线不足和存在障碍物遮挡条件下的图片, 并对这些图片进行了标注。随后, 使用 YOLOv8n-BSG 和 YOLOv8n 模型对这些标注好的图片进行了测试验证, 测试结果如表 6 所示。

表 6 泛化性实验结果

Table 6 Generalization experimental results

模型	MR/ %	FR/ %	mAP/ %	GFLOPs/ G	Params/ M
v8n	8.3	3.6	86.9	8.2	3.01
ours	6.3	1.1	89.8	8.2	3.03

在处理实际拍摄所得数据时, YOLOv8n-BSG 模型的平均精度均值(mAP)达到了 89.8%, 相较于 YOLOv8n 模型, 其性能提升了 2.9%。在实际应用场景中, YOLOv8n-

BSG 对托盘的预测准确性相较于改进前有了显著提高。此外, 与 YOLOv8n 相比, YOLOv8n-BSG 的漏检率和误检率分别降低了 2% 和 2.5%, 在实际工厂环境中, 特别是在光线不足或有障碍物遮挡的情况下, 有效解决了托盘的漏检和误检问题。

4 结 论

针对托盘检测系统在光线不足和障碍物遮挡的情况下识别检测精度和速度不足问题, 本研究通过加入 BRA-nchw 注意力模块, 将损失函数替换为 Shape-IoU, 利用基于 GSConv 的 Slim-neck 架构对 YOLOv8n 的特征融合网络进行重新设计等改进方式有效提升了模型的性能。改进后的轻量化检测模型在保持计算量与参数量基本不变的前提下, 将平均精度均值(mAP)提升至 89.6%, 相较于原 YOLOv8n 架构实现 2.8% 的性能突破, 适用于无人工业车辆等需要低功耗边缘计算的场景。在实地采集的数据集上进行的测试中, 单幅图像的检测帧率高达 330.52 fps, 充分满足了托盘搬运任务对于检测速度的严格要求。然而其内存占用量在某些微型嵌入式设备上显现出部署瓶颈。针对此问题, 后续将通过整合参数量化与分层蒸馏机制, 系统性压缩模型存储空间, 在确保关键特征表达能力的前提下突破更轻量化的难题。

参考文献

- [1] 赵天裔. 基于深度相机的托盘识别和定位系统设计[D]. 成都: 电子科技大学, 2023.
- [2] ZHAO T Y. Design of pallet recognition and positioning system based on depth camera [D]. Chengdu: University of Electronic Science and Technology of China, 2023.
- [3] KNITT M, SCHYGA J, ADAMANOV A, et al. Estimating the pose of a euro pallet with an rgb camera based on synthetic training data[J]. ArXiv preprint arXiv:2210.06001, 2022.
- [4] 向雷, 蒋文波. 改进 YOLOv8 的城市行车道路障碍物检测算法研究[J]. 电子测量技术, 2025, 48(1): 29-38.
- [5] XIANG L, JIANG W B. Research on improved YOLOv8 urban driving road obstacle detection algorithm[J]. Electronic Measurement Technology, 2025, 48(1): 29-38.
- [6] 詹燕, 陈志慧, 朱宝昌, 等. 基于自适应颜色快速点特征直方图的托盘识别方法[J]. 上海交通大学学报, 2023, 57(3): 297-308.
- [7] ZHAN Y, CHEN ZH H, ZHU B CH, et al. A pallet recognition method based on adaptive color fast point feature histogram[J]. Journal of Shanghai Jiao Tong University, 2023, 57(3): 297-308.
- [8] WANG Q, CHEN W X, PENG H Y. An

- intelligentimage detection method using improved canny edge detection operator[J]. International Journal of Information Technology and Management, 2022, 21(4):369-381.
- [6] 周涛,高承帅,蒋久红,等.基于单目视觉的叉车托盘识别与测距方法[J].机电工程技术,2024,53(6):245-249.
ZHOU T, GAO CH SH, JIANG J H, et al. A forklift pallet recognition and distance measurement method based on monocular vision[J]. Mechanical & Electrical Engineering Technology, 2024, 53 (6): 245-249.
- [7] 苏怡萱,李智,盘书宝.基于改进 YOLOv8 的输电线路绝缘子缺陷检测方法[J].电子测量与仪器学报,2025,39(1):14-23.
SU Y X, LI ZH, PAN SH B. Improved YOLOv8-basedinsulator defect detection method for transmission lines [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39 (1): 14-23.
- [8] MOHAMED I S, CAPITANELLI A, MASTROGIOVANII F, et al. Detection, localisation and tracking of pallets using machine learning techniques and 2D range data[J]. Neural Computing and Applications, 2020, 32: 8811-8828.
- [9] 王平凯,孙光泽,朱芮萱.基于改进的 YOLOv3 的托盘检测方法研究[J].机电工程技术,2021,50(7): 29-32.
WANG P K, SUN G Z, ZHU R X. Research on pallet detection method based on improved YOLOv3 [J]. Mechanical & Electrical Engineering Technology, 2021,50(7): 29-32.
- [10] ROCHA E D S, CHEVTCHENKO S F, CAMBUIM L F S, et al. Optimized pallet localization using rgb-d camera and deep learning models[C]. 2023 IEEE 19th International Conference on Intelligent Computer Communication and Processing (ICCP), IEEE, 2023: 155-162.
- [11] 肖施睿.复杂环境下物流托盘视觉识别技术研究[D].成都:西南交通大学,2022.
XIAO SH R. Research on visual recognition technology of logistics pallet in complex environment [D]. Chengdu: Southwest Jiaotong University,2022.
- [12] 刘继泽.面向工业移动机器人的轻量级托盘检测算法设计与实现[D].成都:电子科技大学,2024.
LIU J Z. Design and implementation of a lightweight pallet detection algorithm for industrial mobile robots [D]. Chengdu: University of Electronic Science and Technology of China, 2024.
- [13] 胡海霞,向岑阳,范维芹.基于改进 YOLOv8n 模型的托盘识别方法研究[J/OL].重庆工商大学学报(自然科学版), 1-11 [2024-12-21]. <https://kns.cnki.net.webvpn.xju.edu.cn:8040/kcms/.detail/50.1155.N.20240914.1645.007.html>.
HU H X, XIANG C Y, FAN W Q. Study on pallet recognition method based on improved YOLOv8n[J/OL]. Journal of Chongqing Technology and Business University(Natural Science Edition), 1-11 [2024-12-21]. <https://kns.cnki.net.webvpn.xju.edu.cn:8040/kcms/detail/50.1155.N.20240914.1645.007.html>.
- [14] VU V D, HOANG D D, TAN P X, et al. Occlusion-robust pallet pose estimation for warehouse automation[J]. IEEE Access, 2024, 12: 1927-1942.
- [15] 李泽胤,李栋,房建东,等.改进的 YOLOv8n 遥感图像轻量化检测模型[J].电子测量技术,2025,48(6): 130-142.
LI Z Y, LI D, FANG J D, et al. Improved YOLOv8n lightweight detection model for remote sensing images[J]. Electronic Measurement Technology, 2025, 48 (6): 130-142.
- [16] TERVEN J, CÓRDOVA-ESPARZA D M, ROMERO-GONZÁLEZ J A. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS[J]. Machine Learning and Knowledge Extraction, 2023, 5 (4): 1680-1716.
- [17] ZHU X K, LYU SH CH, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]. IEEE/CVF International Conference on Computer Vision,2021: 2778-2788.
- [18] 王德伟,刘小芳.基于改进 YOLOv8n 的带钢表面缺陷检测方法[J].国外电子测量技术,2024,43 (7): 158-169.
WANG D W, LIU X F. Surface defect detection method of strip steel based on improved YOLov8n [J]. Foreign Electronic Measurement Technology, 2024, 43 (7): 158-169.
- [19] ZHU L, WANG X J, KE ZH H, et al. Bifomer: Vision transformer with bi-level routing attention[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023: 10323-10333.
- [20] REN S CH, ZHOU D Q, HE SH F, et al. Shunted self-attention via multi-scale token aggregation [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition,2022: 10853-10862.
- [21] 李慧琴,宋赵铭,刘存祥,等.基于 YOLOv8n 的番茄果

- 实检测模型改进[J/OL]. 河南农业大学学报, 1-14 [2024-12-21]. <https://doi.org/10.16445/j.cnki.1000-2340.20240511.002>.
- [21] LI H Q, SONG ZH M, LIU C X, et al. Improvement of tomato fruit detection model based on YOLOv8n[J/OL]. Journal of Henan Agricultural University, 1-14 [2024-12-21]. <https://doi.org/10.16445/j.cnki.1000-2340.20240511.002>.
- [22] YUAN X P, FU ZH B, ZHANG B W, et al. Research on lightweight algorithm for gangue detection based on improved Yolov5[J]. Scientific Reports, 2024, 14(1):6707.
- [23] ZHANG H, ZHANG SH J. Shape-IoU: More accurate metric considering bounding box shape and scale[J]. ArXiv preprint arXiv:2312.17663, 2023.
- [24] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles[J]. ArXiv preprint arXiv:2206.02424, 2022.
- [25] 胡焱, 原子昊, 涂晓光, 等. 基于对比学习的改进 SSD 目标检测算法[J]. 红外技术, 2024, 46(5):548-555.
- HU Y, YUAN Z H, TU X G, et al. Improved SSD object detection algorithm based on contrastive learning [J]. Infrared Technology, 2024, 46(5):548-555.
- [26] ZHAO Q H, LIU Y Q. Design of apple recognition model based on improved deep learning object detection framework Faster-RCNN[J]. Advances in Continuous and Discrete Models, 2024, 2024(1):49.
- [27] WANG CH Y, BOCHKOVSKIY A, LIAO H Y. 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.
- [28] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems(ADICS), 2024.
- [29] JEGHAM N, KOH C Y, ABDELATTI M, et al. Evaluating the evolution of YOLO(You Only Look Once) models: A comprehensive benchmark study of YOLO11 and its predecessors[J]. ArXiv preprint arXiv:2411.00201, 2024.

作者简介

刘晓非,硕士研究生,主要研究方向为移动机器人、托盘的识别与定位。

E-mail: lxf682@163.com

薛瑞雷(通信作者),本科,高级工程师,硕士生导师,主要研究方向为智能机器人研发和机器视觉图像处理。

E-mail: 1945073312@qq.com