

# 基于改进 YOLOv5s 的自动导引运输车 托盘孔位视觉定位方法

崔 升<sup>1</sup> 唐芳丽<sup>1</sup> 郑亮宇<sup>1</sup> 曾伟理<sup>1</sup> 曲伟伟<sup>2</sup>

(1. 广西中烟工业有限责任公司柳州卷烟厂, 广西 柳州 545006;

2. 扬州市天宝自动化工程有限公司, 江苏 扬州 225800)

**摘要:** [目的] 自动导引运输车在搬运过程中, 需定位的托盘孔位中存在的细小、形变、低对比度孔位的视觉定位不准的问题。因此, 提出一种基于改进 YOLOv5s 的自动导引运输车托盘孔位视觉定位方法。[方法] 结合 ShuffleNetV2 的通道混洗操作改进和 CBAM 注意力机制改进, 对基本 YOLOv5s 框架进行改进, 使其聚焦于形变关键区域中亚像素级边界模糊的孔位区域; 基于 SIoU 损失函数关注微小孔位, 并计算托盘孔位在相机坐标系下的空间三维坐标, 得到相机坐标系到孔位区域坐标系的变换关系, 采用改进的 YOLOv5s 框架输出 AGV 机械臂坐标系下的托盘孔位三维坐标。[结果] 试验方法可有效捕捉亚像素级定位精度边界, 绝对误差  $<0.03$  cm, 相对误差  $<0.83\%$ ;  $F_1$  分数、mAP 指标分别为 95.2%、94.8%; 浮点运算次数、参数量和模型体积分别为 4.8 G、2.6 M、4.28 MB。[结论] 试验方法有效解决了需定位托盘孔位中存在的细小、形变、低对比度孔位的视觉定位难题, 提升了自动导引运输车托盘搬运效率。

**关键词:** YOLOv5s; 自动导引运输车; 托盘孔位定位; 深度可分离卷积; CBAM 注意力

## Visual positioning method for AGV tray holes based on improved YOLOv5s

CUI Sheng<sup>1</sup> TANG Fangli<sup>1</sup> ZHENG Liangyu<sup>1</sup> ZENG Weili<sup>1</sup> QU Weiwei<sup>2</sup>

(1. Guangxi Zhongyan Industrial Co., Ltd., Liuzhou Carette Factory Guangxi Zhuang Autonomous Region, Liuzhou, Guangxi 545006, China; 2. Yangzhou Tianbao Automation Co., Ltd., Yangzhou, Jiangsu 225800, China)

**Abstract:** [Objective] In the handling process of the automatic guided vehicle (AGV), accurate visual positioning of tray holes, particularly those that are small, distorted, and of low contrast, poses a significant challenge. Therefore, a visual positioning method for tray holes in AGVs based on improved YOLOv5s is proposed. [Methods] Combining the channel shuffling operation improvement of ShuffleNetV2 and the CBAM attention mechanism improvement, the basic YOLOv5s framework is improved to focus on the sub-pixel level boundary blurred hole position areas in the deformation key region. Based on the SIoU loss function, attention is paid to the micro hole positions, and the spatial three-dimensional coordinates of the tray hole positions in the camera coordinate system are calculated to obtain the transformation relationship from the camera coordinate system to the hole position area coordinate system. The improved YOLOv5s framework is utilized to output the three-dimensional coordinates of the tray hole positions in the AGV robotic arm coordinate system. [Results] The experimental results show that the experimental method can effectively capture sub-pixel level positioning accuracy boundaries, with an absolute error of less than 0.03 cm and a relative error of less than 0.83%. The  $F_1$  score and mAP index are 95.2% and 94.8%, respectively. The number of floating-point operations, parameter count, and model volume are 4.8 G, 2.6 M and 4.28 MB, respectively. [Conclusion] The experimental method effectively solves the visual positioning problem of small, deformed, and low contrast

**基金项目:** 广西壮族自治区工信厅技术创新项目(编号:2009016677)

**通信作者:** 唐芳丽(1997—), 女, 广西中烟工业有限责任公司柳州卷烟厂工程师。E-mail:13321727467@163.com

**收稿日期:** 2025-06-11 **改回日期:** 2025-11-23

**引用格式:** 崔升, 唐芳丽, 郑亮宇, 等. 基于改进 YOLOv5s 的自动导引运输车托盘孔位视觉定位方法[J]. 食品与机械, 2026, 42(1): 79-85.

**Citation:** CUI Sheng, TANG Fangli, ZHENG Liangyu, et al. Visual positioning method for AGV tray holes based on improved YOLOv5s [J]. Food & Machinery, 2026, 42(1): 79-85.

holes in the tray holes that need to be positioned, and improves the efficiency of AGV tray handling.

**Keywords:** improve YOLOv5s; automated guided vehicle; positioning of tray holes; depthwise separable convolution; CBAM attention

在食品工业的仓储及车间物流环节,自动导引运输车(automated guided vehicle, AGV)凭借其高效的自主导航与柔性搬运能力,成为提高整体运营效率的关键设备<sup>[1-3]</sup>。然而,在实际运行过程中,高精度、低延迟的托盘定位技术已成为无人叉车大规模落地的关键瓶颈<sup>[4-6]</sup>,直接影响 AGV 叉车拣选操作的稳定性和整体物流效率。如果定位不准确,很容易导致叉偏差、货物掉落或重复调整等问题,严重影响自动化过程的连续性。

在实际应用中,托盘可能因地面上不平或叉取操作而产生微小位移。此外,AGV 自身的定位误差、运动控制误差以及在执行动作过程中产生的机械偏差,存在细小、形变、低对比度孔位的视觉定位不准的问题,会在整个流程中不断累积。刘蕾<sup>[7]</sup>使用安装在 AGV 上的二维摄像头捕捉目标区域的形变关键区域,通过模板匹配方法定位目标的中心点,通过手眼校准确定目标在机器人坐标系中的三维坐标。但在托盘倾斜时,模板匹配失败概率较大,存在托盘孔位形变关键区域模糊、形变现象。邹斌等<sup>[8]</sup>构建了一个多尺度目标定位模型,以提高密集人群定位范围的覆盖率。然而,当多尺度目标定位模型部署在嵌入式设备上时,会受到两级定位架构串行计算特性的限制,推理速度慢。李哲等<sup>[9]</sup>通过生成抓取点坐标,经坐标转换后映射至机械臂关节空间,结合 AGV 路径规划与逆运动学求解,完成了抓取动作。然而,该方法对小尺寸工件的抓取效果较差,存在孔位中心坐标漂移现象。伍子嘉等<sup>[10]</sup>通过视觉定位技术,设计了机械臂工件抓取技术,但仍存在微小目标定位效果不理想、定位精度受限的问题。

与传统的“先选择候选区域,然后分类”的两阶段算法不同,YOLO 是一种先进的单级目标检测算法,其核心创新是将目标定位视为一个回归问题。而 YOLOv5s 为 YOLO 的第 5 代,通过逻辑划分图片区域进行预测,用处众多<sup>[11-13]</sup>,例如 PCB 板缺陷定位、半导体芯片定位、物流包裹分拣等任务,均与托盘孔位定位同属“高精度、小目标”的定位问题。YOLOv5s 算法被用来精准定位这些细小缺陷,可以有效替代人眼和显微镜,同样,在解决食品工业的仓储及车间物流环节中高精度、高效率视觉定位任务过程中潜力巨大。

在物料搬运的实际作业场景中,AGV 托盘孔位尺寸较小,且通常处于倾斜、旋转等非刚性工况中,导致托盘孔位视觉定位常面临非刚性工况的两大核心挑战:动态运动导致的模糊效应(如 AGV 急停/转向时的托盘瞬时位移)和机械载荷引发的形变现象(如托盘承载重物后的结构性弯曲),超出了 YOLOv5s 原始架构对微小目标的亚像素级定位精度边界,进一步加剧了孔位定位难度。因

此,试验拟提出一种基于改进 YOLOv5s 的 AGV 托盘孔位视觉定位方法。针对待定位孔位中心坐标漂移、托盘孔位视觉定位效果受限问题,分析其因固定感受也难以捕捉模糊边界问题,通过 ShuffleNetV2 的通道混洗操作改进和 CBAM 注意力机制改进两种手段改进基本 YOLOv5s 框架,旨在满足细小、形变、低对比度孔位的视觉定位需求,进而提升 AGV 的智能化作业水平。

## 1 AGV 托盘孔位视觉定位方法设计

### 1.1 考虑微小目标模糊、形变现象的 YOLOv5s 改进

YOLOv5s 作为经典的单阶段目标定位网络<sup>[14]</sup>,凭借其 CSPDarknet 主干架构、高效的路径聚合网络(PANet)以及自适应锚框计算机制,在实时目标定位任务中展现出卓越的性能优势。基本 YOLOv5s 框架如图 1 所示。

1.1.1 ShuffleNetV2 的通道混洗操作改进 AGV 的加减速( $\geq 1.5 \text{ m/s}^2$ )会导致托盘在相机视野中产生非均匀模糊,而普通卷积的固定感受野会放大这种非均匀模糊程度,导致待定位孔位中心坐标漂移,故改进 ShuffleNetV2 的通道混洗操作,通过混洗建立当前帧与历史帧孔位坐标的通道级关联,抑制单帧抖动,控制孔位视觉定位过程中的中心坐标漂移现象,解决动态运动导致的模糊效应问题。

设定 ShuffleNetV2 网络输入视觉特征图的高度、宽度、通道数分别为  $H$ 、 $W$ 、 $C_{\text{in}}$ ,利用  $K \times K$  标准卷积核进行处理,可通过式(1)计算出参数量  $P$ 。

$$P = C_{\text{in}} C_{\text{out}} K K \circ \quad (1)$$

运算量  $Q$  的计算式为:

$$Q = C_{\text{in}} C_{\text{out}} H W K K \circ \quad (2)$$

式中:

$C_{\text{out}}$ ——卷积处理后的通道数。

深度可分离卷积  $P_{\text{DW}}$  的计算式为:

$$P_{\text{DW}} = C_{\text{in}} K K + C_{\text{in}} C_{\text{out}} \circ \quad (3)$$

深度可分离卷积的参数量为:

$$Q_{\text{DW}} = C_{\text{in}} H W K K + C_{\text{in}} C_{\text{out}} H W \circ \quad (4)$$

联立式(1)~式(4),可得:

$$\frac{P_{\text{DW}}}{P} = \frac{Q_{\text{DW}}}{Q} = C_{\text{out}}^{-1} + K^{-2} \circ \quad (5)$$

根据式(5)可知,将深度可分离卷积引入 ShuffleNetV2 网络中,进行两阶段改进。

(1) 可变形分组卷积:在 SFU 第 2 分支中,3×3 卷积核动态调整感受野形状。例如,当检测到托盘四角下垂时,卷积核自动拉伸以匹配孔位的椭圆边界。

(2) 多尺度特征融合:通过步长为 2 的 SFU 模块捕获

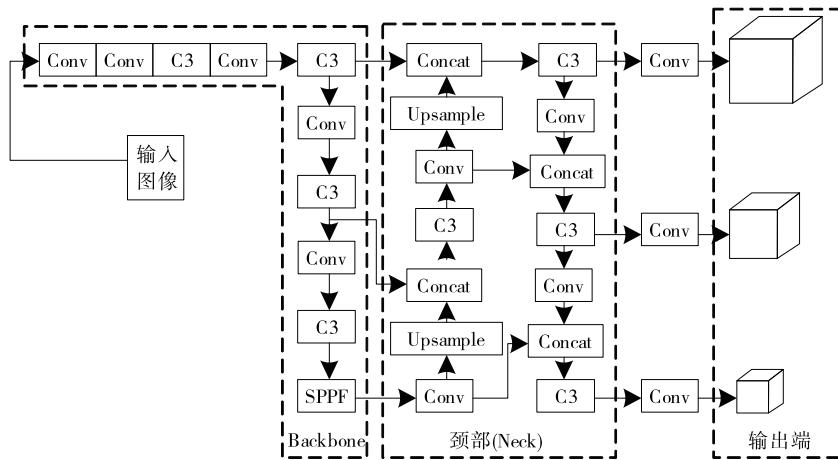


图1 基本YOLOv5s框架

Figure 1 Basic YOLOv5s framework

托盘整体弯曲度(如中间下陷5 cm),结合局部孔位形变数据,反向推算最佳叉齿插入角度。

ShuffleNetV2的通道混洗操作改进结构如图2所示。

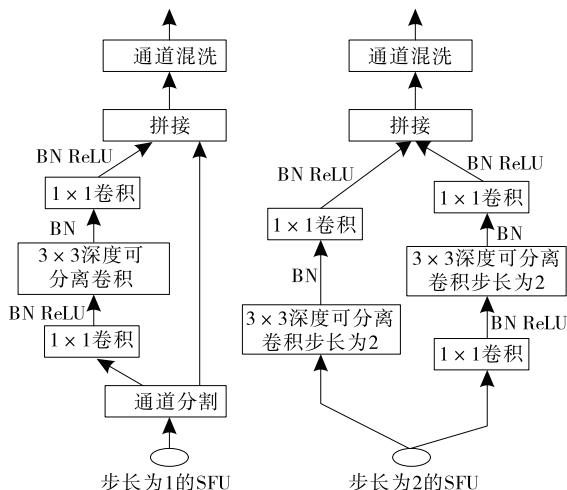


图2 ShuffleNetV2的通道混洗操作改进结构

Figure 2 Improved structure of channel mixing operation in ShuffleNetV2

1.1.2 CBAM注意力机制改进 托盘中部承载重物时,形变量呈梯度分布(如边缘形变0.5 mm→中心形变3 mm),此时,金属托盘孔位可能从圆形变为椭圆(长短轴比达1.2:1),塑料托盘则出现局部塌陷,而对所有区域平等处理,无法聚焦形变关键区域,故改进CBAM注意力机制,可实现从“看到形变”到“理解形变”再到“补偿形变”的闭环提升,进一步扩大感受野,解决孔位视觉定位过程中的模糊边界问题。CBAM注意力机制结构如图3所示。

CBAM注意力机制模块中包含两个基本单元,其中通道注意力子模块(CAM)通过动态评估视觉特征图各通

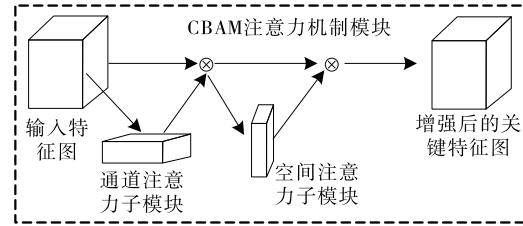


图3 CBAM注意力机制结构图

Figure 3 CBAM attention mechanism structure

道的贡献度,对关键通道赋予更高权重,以强化孔位特征表达。通过在通道维度上并行执行最大/平均池化操作,分别捕获特征向量  $F_{\text{avg}}^c, F_{\text{max}}^c$ ,应用多层感知机(MLP)学习各通道权重,从而抑制不重要的通道信息。CAM在对输入视觉特征图进行处理后,生成通道注意力图  $M_c(F)$ :

$$M_c(F) = \sigma(MLP(\text{AvgPool}(F)) + MLP(\text{MaxPool}(F))) = \sigma(W_1(W_0 F_{\text{avg}}^c) + W_1 W_0 F_{\text{max}}^c), \quad (6)$$

式中:

$\sigma$ —sigmoid函数;

$F$ —输入到CBAM注意力机制的原始视觉特征图;

$W_0, W_1$ —MLP的权重矩阵。

空间注意力子模块(SAM)能够使模型聚焦于视觉特征图中的孔位区域。SAM处理后的空间注意力图  $M_s(F)$ 为:

$$M_s(F) = \sigma(f^{(7 \times 7)})([\text{AvgPool}(F); \text{MaxPool}(F)]) = \sigma(f^{(7 \times 7)})(F_{\text{avg}}^c; F_{\text{max}}^c), \quad (7)$$

式中:

$f^{(7 \times 7)}$ —7×7卷积操作<sup>[15-17]</sup>;

[; ]—通道拼接运算。

## 1.2 应用损失函数的微小孔位改进 YOLOv5s 定位

AGV 托盘孔位尺寸微小, 像素占比低<sup>[18]</sup>, 为了在充分保留 AGV 托盘孔位中托盘孔位区域边缘、纹理等细节信息的同时, 更加关注微小孔位, 应用较低层级的 P2 层用于捕捉亚像素级定位精度边界, 较高层级 P2 能够提取托盘孔位与周围托盘结构的相对位置关系。将两个不同 P2 层提取的视觉特征图与 Backbone 网络对应的微小尺度视觉特征图进行拼接, 结合 SIoU 损失函数(角度损失、距离损失、形状损失、IoU 损失), 可实现孔位视觉定位。

(1) 角度损失: 关注预测框 A 与托盘孔位目标框 B 中心点连线和水平轴的角度偏差, 计算式为:

$$\Lambda = \cos \left( 2 \sin \left( \arcsin \left( \frac{c_h}{r} \right) - \frac{\pi}{4} \right) \right), \quad (8)$$

式中:

$c_h, r$ —B 中心点与 A 中心点之间的高度差和欧式距离。

通过三角函数运算, 角度损失能根据角度偏差大小动态调整惩罚权重, 从而有效约束托盘孔位目标框中心点的偏差方向。

(2) 距离损失: 用于衡量 B 中心点与 A 中心点之间的绝对距离差异, 计算式为:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) = 2 - e^{-\gamma \rho_x} - e^{-\gamma \rho_y}, \quad (9)$$

$$\begin{cases} \rho_x = \left( \frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2 \\ \rho_y = \left( \frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2, \\ \gamma = 2 - \Delta \end{cases} \quad (10)$$

式中:

$(b_{c_x}^{gt}, b_{c_y}^{gt})$ —B 中心点坐标;

$c_w, c_h$ —B 的最小外接矩形的宽、高;

$(b_{c_x}, b_{c_y})$ —A 中心点坐标。

(3) 形状损失: 专注于 B 与 A 的宽、高比例差异, 计算式为:

$$\begin{cases} \Omega = (1 - e^{-w_w})^\theta + (1 - e^{-w_h})^\theta \\ w_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \\ w_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases}, \quad (11)$$

式中:

$w^{gt}, h^{gt}$ —B 的宽、高;

$w, h$ —A 的宽、高;

$\theta$ —形状损失关注程度系数。

(4) IoU 损失: 用于计算预测框 A 与 B 的交并比, 衡量

两者在面积上的重叠程度, 计算式为:

$$L_{IoU} = \frac{S(A \cap B)}{S(A \cup B)}, \quad (12)$$

式中:

S—面积参数。

综上, SIoU 损失函数为:

$$L_{SIoU} = 1 - L_{IoU} + 2^{-1}(\Lambda + \Delta + \Omega). \quad (13)$$

结合 SIoU 损失函数后, 改进 YOLOv5s 框架如图 4 所示。

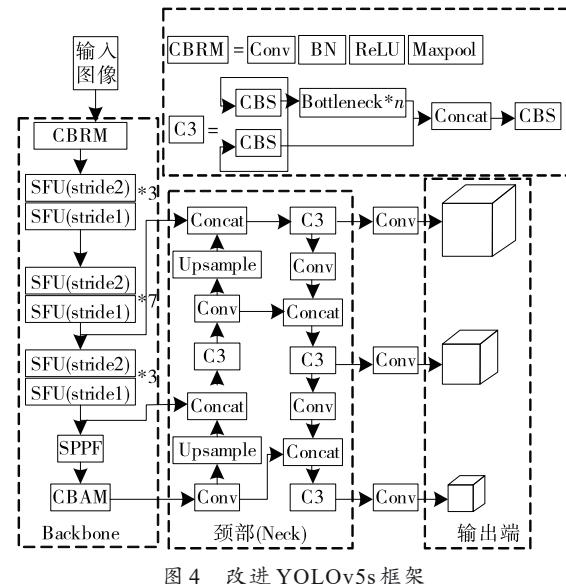


Figure 4 Improved YOLOv5s framework

根据托盘孔位在关键区域中的二维坐标信息( $u, v$ ), 即托盘孔位在左视图关键区域中的位置。使用已标定的双目摄像头和 SGBM 算法进行立体匹配, 得到( $u, v$ )在右视图中的对应像素( $u', v'$ ), 并根据 SIoU 损失函数, 计算视差  $\phi = u - u'$ 。

对微小尺度视觉特征图进行拼接, 则托盘孔位在相机坐标系下的空间三维坐标( $x, y, z$ )计算式为:

$$\begin{cases} x = u \cdot B / \phi \\ y = v \cdot B / \phi, \\ z = B \cdot \delta / \phi \end{cases} \quad (14)$$

式中:

$B, \delta$ —双目摄像头光心距离以及焦距。

通过手眼标定, 确定转换矩阵  $T_{arm-cam}$ , 用于表示相机坐标系到孔位区域坐标系的变换关系:

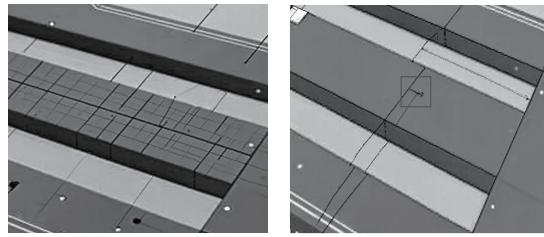
$$G_{arm} = T_{arm-cam} \cdot G_{cam}, \quad (15)$$

式中:

$G_{cam} = (x, y, z, 1)^T$ —托盘孔位在相机坐标系下的齐次坐标;

$G_{\text{arm}}$ ——托盘孔位在AGV机械臂坐标系下的三维坐标。

根据双目定位原理定位微小孔位,拼接后的微小孔位定位示意图如图5所示。



(a) 微小尺度视觉特征图拼接 (b) 微小孔位定位结果  
图5 拼接后的微小孔位定位示意图

Figure 5 Microhole positioning after splicing

将托盘孔位坐标 $(x, y, z)$ 代入式(15),可得到AGV机械臂坐标系下的托盘孔位三维坐标:

$$G_{\text{arm}} = T_{\text{arm-cam}} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (16)$$

由此即可实现AGV托盘孔位视觉定位。

## 2 验证实验

以某托盘叉车AGV为试验对象,开展托盘孔位视觉定位研究,AGV视觉定位平台如图6所示。

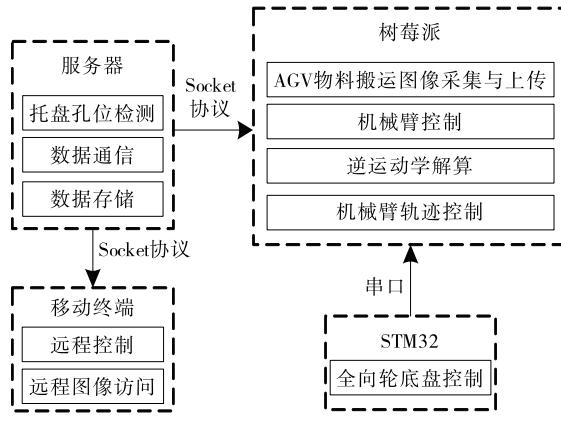


图6 AGV视觉定位平台  
Figure 6 AGV visual positioning platform

树莓派通过串口通信与STM32进行数据交互,形成完整的控制闭环;移动终端设备实施远程控制功能,并实时监控AGV物料搬运过程的视频流数据;服务器作为系统的核心处理单元,完成数据中转与存储管理,调用树莓派搭载的双目摄像头主动采集AGV作业区域内的托盘孔位图像,接收服务器返回的托盘孔位坐标数据后,结合双目视觉定位原理精确计算托盘孔位在AGV机械臂坐

标系下的三维空间坐标;STM32控制器通过PID算法精确调控全向轮底盘的运动状态。

试验样本数为6780个,按照9:1划分为训练集和测试集,在Pytorch框架下构建托盘孔位定位模型,设定学习率初值为 $1 \times 10^{-3}$ ,每批次处理样本数为32,权重衰减因子为 $1 \times 10^{-3}$ ,最大训练次数为500。分别应用试验方法、文献[8]方法(基于Faster R-CNN的定位算法)、文献[9]方法(基于新型RRT算法的定位方法)对托盘孔位进行视觉定位。不同方法下的待定位孔位中心坐标漂移量 $F_1$ 分数曲线走势规律如图7所示。

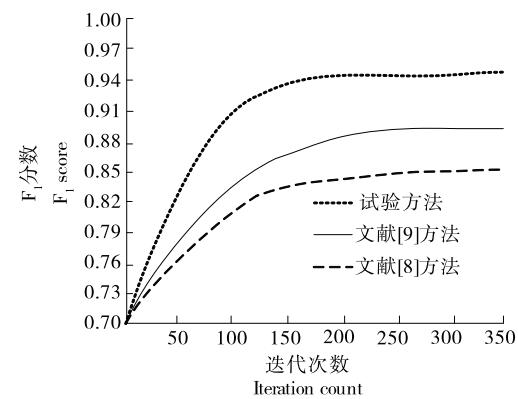


图7 待定位孔位中心坐标漂移量 $F_1$ 分数曲线对比结果

Figure 7 Comparison results of  $F_1$  score curves for center coordinate drift of the hole position to be located

由图7可知,随着迭代次数的增加, $F_1$ 分数曲线总体呈持续上升趋势。文献[8]方法采用Faster R-CNN模型进行托盘孔位视觉定位,因对小目标定位效果不佳,且视觉特征图空间分辨率有限,难以精准捕捉孔位细节信息,致使托盘孔位视觉定位的 $F_1$ 分数最高值仅为0.858。文献[9]方法采用原始YOLOv5s模型,利用CSPDarknet骨干网络和PANet特征金字塔结构,能够更精准地捕捉托盘孔位的细节信息,使得 $F_1$ 分数明显增大。

综上,CBAM改进的YOLOv5s通过双注意力机制协同作用,优先保留对低对比度敏感的卷积核(如Gabor滤波器等效响应),得到了渲染带形变(仿射变换)、模糊(高斯核 $\sigma=1.5$ )的虚拟托盘图像,解决了细小、形变、低对比度孔位的定位难等问题。

AGV需确保叉齿水平对准孔位中心,平面坐标 $(x, y)$ 直接决定横向定位精度,而深度距离反映叉齿与托盘的垂直间隙(通常需控制5~10 mm防碰撞)。故假设输出托盘孔位三维坐标为 $(x, y, z)$ ,将孔位中心点至相机基线中心的距离 $d$ 作为实际值,深度距离 $z$ 与实际距离之间的绝对误差、相对误差见表1。

由表1可知,试验方法的绝对误差均 $\leq 0.03$  cm,其中大约90%的样本误差 $\leq 0.02$  cm,表明该方法在深度方向

表 1 托盘孔位视觉定位结果

Table 1 Analysis of visual positioning results for tray hole positions

试验号	<i>z</i> /cm	<i>d</i> /cm	绝对误差/cm	相对误差/%
1	2.08	2.09	0.01	0.48
2	2.51	2.50	0.01	0.40
3	3.15	3.14	0.01	0.32
4	3.58	3.61	0.03	0.83
5	4.24	4.26	0.02	0.46
6	5.17	5.16	0.01	0.19
7	5.56	5.55	0.01	0.18
8	6.41	6.43	0.02	0.31
9	7.24	7.26	0.02	0.27
10	7.81	7.79	0.02	0.25

(*Z*轴)的定位精度达到亚毫米级,满足对高精度定位的需求。试验方法的相对误差为0.18%~0.83%,验证了方法在复杂工况下的稳定性。试验方法的误差与距离呈弱相关性(如3.58 cm时误差最大为0.03 cm,而7.81 cm时误差为0.02 cm),说明试验方法通过改进的SloU损失函数有效抑制了深度增加导致的误差累积效应。

综上,在AGV托盘孔位视觉定位中,深度距离*z*与实际距离*d*的误差控制是可靠性工作的核心。通过捕捉亚像素级定位精度边界,可将绝对误差控制在有效范围内,降低相对误差,AGV叉齿不会撞击到托盘底部,满足工业场景的严苛要求。

不同消融方案下的平均精度均值、模型体积、浮点运算次数和参数量见表2。

表 2 消融试验对比分析

Table 2 Comparative analysis of ablation experiments

消融方案	浮点运算次数/G	参数量/M	模型体积/MB	平均精度均值/%
基本YOLOv5s框架	15.3	7.21	13.61	85.3
引入ShuffleNetV2的通道混洗操作改进环节	4.5	2.43	4.13	86.9
引入CBAM注意力机制改进环节	4.7	2.52	4.22	89.5
引入SloU损失函数环节	4.8	2.58	4.25	92.4
整体托盘孔位视觉定位框架	4.8	2.60	4.28	94.8

由表2可知,在AGV托盘孔位视觉定位场景下,各消融方案所展现出的性能差异非常明显。浮点运算次数和模型体积分别为15.3 G、13.61 MB,将其运行在资源受限的AGV嵌入式设备中,难以满足托盘孔位实时性定位要求,在面对复杂工况时,极易出现漏检、误检,平均精度均

值仅为85.3%;引入ShuffleNetV2的通道混洗操作改进后,有效提升了轻量化水平,但在托盘孔位定位精度上未有明显改善;引入CBAM注意力机制改进后,增强了对托盘孔位视觉特征图的捕捉能力,虽小幅度增加了计算负担,但在平均精度均值上,相比基本YOLOv5s框架增加了4.2%;引入SloU损失函数后,进一步优化了对微小孔位的视觉定位效果;整体托盘孔位视觉定位框架后,进一步增强了托盘孔位定位效果,平均精度均值上升到最大值(94.8%)。

综上,基本YOLOv5s框架的计算资源消耗较大,而单纯轻量化可能牺牲关键特征提取能力,同时,对微小目标(如直径<5 cm的孔位)的定位误差敏感,SloU可减少机械臂抓取时的位姿偏差,这种综合性的改进,可以为工业AGV视觉系统提供一种“高精度—低延迟”的可行方案,可直接应用于智能仓储、柔性生产线等场景。

### 3 结论

研究提出的自动导引运输车托盘孔位视觉定位方法,创新性应用了一种改进YOLOv5s框架的双重优化策略,即通过轻量化结构改进(ShuffleNetV2),在Neck部分嵌入CBAM模块,通过通道—空间双注意力聚焦形变区域的亚像素级孔位,并结合融合角度损失(防止旋转偏移)、距离损失(中心点精度)、形状损失(适应孔位形变)、IoU损失(边界框贴合度)的复合损失函数,实现了对模糊、亚像素级的孔位视觉定位。结果表明,相比基本YOLOv5s框架,改进后的浮点运算次数从15.3 G降至4.8 G,参数量从7.21 M减少至2.60 M,模型体积压缩至4.28 MB。试验方法提升了对于微小孔位的定位能力,平均精度均值提升至92.4%,精准聚焦孔位区域,平均精度均值进一步提升至94.8%。深度方向(*Z*轴)定位绝对误差≤0.03 cm,相对误差≤0.83%。

综上,试验方法显著提升了自动导引运输车叉齿与托盘的对接成功率,可扩展至多托盘堆叠、动态避障等复杂场景,进一步释放自动导引运输车系统的自动化潜力。但该方法在多目标交互场景中的应用局限性较大,例如强反光环境、极端形变等场景的应用效果尚不可知。后续可探索事件相机+SNN架构,进一步分析极端工况适应性,有效改善因运动规划与视觉定位的时序不同步导致的响应延迟问题,最终建立一套适用于极恶劣工况的实时视觉定位系统。

### 参考文献

- [1] 闫冠宇,蔡泽平,梁家海,等.一种基于全局视觉的AGV目标检测与定位算法[J].北部湾大学学报,2024,39(4): 68-75, 86.  
YAN G Y, CAI Z P, LIANG J H, et al. An AGV object detection and localization algorithm based on global vision[J]. Journal of Beibu Gulf University, 2024, 39(4): 68-75, 86.

- [2] 王辉,袁斌,王伟博,等.基于视觉和RFID复合导航的AGV设计与研究[J].机床与液压,2022,50(7):87-91.
- WANG H, YUAN B, WANG W B, et al. Design and research of AGV based on vision and RFID integrated navigation[J]. Machine Tool & Hydraulics, 2022, 50(7): 87-91.
- [3] 吕开旺,王明睿,刘振国,等.视觉引导机器人协作AGV上下料系统研究[J].制造业自动化,2022,44(1):123-126.
- LYU K W, WANG M R, LIU Z G, et al. Research on vision guided robot cooperating with AGV loading and unloading system[J]. Manufacturing Automation, 2022, 44(1): 123-126.
- [4] 雷杏子,王树才,龚东军,等.基于YOLOv5s的筐装禽蛋上料机器人视觉定位方法[J].华中农业大学学报,2024,43(3):302-310.
- LEI X Z, WANG S C, GONG D J, et al. A method for visually positioning loading robot of basket-packed poultry eggs based on YOLOv5s[J]. Journal of Huazhong Agricultural University, 2024, 43(3): 302-310.
- [5] 余杰先,张中华,陈凌玲,等.移动机器人在自动化生产线上下料中的定位方法[J].机电工程技术,2023,52(10):196-200.
- YU J X, ZHANG Z H, CHEN L L, et al. Positioning method of mobile robots in the process of automatic production line loading and unloading[J]. Mechanical & Electrical Engineering Technology, 2023, 52(10): 196-200.
- [6] 刘泽平,刘明兴,麻方达,等.基于移动目标检测和目标追踪的全局视觉AGV的定位算法[J].青岛大学学报(自然科学版),2022,35(1):85-92.
- LIU Z P, LIU M X, MA F D, et al. A positioning algorithm for global vision guided AGV based on moving target and target tracking[J]. Journal of Qingdao University (Natural Science Edition), 2022, 35(1): 85-92.
- [7] 刘蕾.基于2D视觉的复合AGV机器人上下料的定位方法[J].信息技术与信息化,2022(7): 58-61.
- LIU L. A positioning method for loading and unloading of composite AGV robots based on 2D vision[J]. Information Technology and Informatization, 2022(7): 58-61.
- [8] 邹斌,张聪.基于Faster R-CNN的密集人群检测算法[J].计算机应用,2023,43(1): 61-66.
- ZOU B, ZHANG C. Dense crowd detection algorithm based on Faster R-CNN[J]. Journal of Computer Applications, 2023, 43 (1): 61-66.
- [9] 李哲,于海生,吴贺荣,等.基于新型RRT算法与视觉定位的机械臂工件抓取[J].组合机床与自动化加工技术,2022(12): 148-153.
- LI Z, YU H S, WU H R, et al. Workpiece grasping of manipulator based on improved RRT algorithm and visual positioning[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2022(12): 148-153.
- [10] 伍子嘉,陈航,彭勇,等.动态环境下融合轻量级YOLOv5s的视觉SLAM[J].计算机工程,2022,48(8): 187-195, 205.
- WU Z J, CHEN H, PENG Y, et al. Visual SLAM with lightweight YOLOv5s in dynamic environment[J]. Computer Engineering, 2022, 48(8): 187-195, 205.
- [11] 王泷,纪元霞,吴红刚,等.基于改进YOLOv5s的白酒摘酒酒度检测方法[J].食品与机械,2025,41(7): 65-71.
- WANG L, JI Y X, WU H G, et al. Detection of alcohol content in liquor gathering based on improved YOLOv5s[J]. Food & Machinery, 2025, 41(7): 65-71.
- [12] 张海峰,芦新春,冯博,等.基于改进YOLOv5s的海水鱼种类识别[J].食品与机械,2024,40(8): 84-92.
- ZHANG H F, LU X C, FENG Y R. Marine fish species recognition based on improved YOLOv5s[J]. Food & Machinery, 2024, 40(8): 84-92.
- [13] 魏晶鑫,陈中举,许浩然.融合Fasternet与YOLOv5模型的鸡蛋外观检测[J].食品与机械,2024,40(8): 105-112, 165.
- WEI J X, CHEN Z J, XU H R. Detection of egg appearance based on Fasternet and YOLOv5 model[J]. Food & Machinery, 2024, 40(8): 105-112, 165.
- [14] 黄友锐,王照峰,韩涛,等.结合轻量化YOLOv5s的动态视觉SLAM算法[J].电子测量技术,2024,47(11): 59-68.
- HUANG Y R, WANG Z F E, HAN T, et al. Dynamic visual SLAM algorithm combined with lightweight YOLOv5s[J]. Electronic Measurement Technology, 2024, 47(11): 59-68.
- [15] LIU Q, TANG X Y, HUO J. Attitude measurement of ultraclose-range spacecraft based on improved YOLOv5s and adaptive Hough circle extraction[J]. Applied Optics, 2024, 63 (5): 1 364-1 376.
- [16] 李爽,曹阳,沈琴琴,等.改进YOLOv5s的小目标车辆检测算法[J].计算机仿真,2025,42(2): 155-161.
- LI S, CAO Y, SHEN Q Q, et al. Small target vehicle detection algorithm based on improved YOLOv5s[J]. Computer Simulation, 2025, 42(2): 155-161.
- [17] 刘锡琳,潘文松,张爱军.基于改进YOLOv5s的轮毂气门孔检测算法[J].电子设计工程,2024,32(19): 140-144, 149.
- LIU X L, PAN W S, ZHANG A J. Hub valve hole detection algorithm based on improved YOLOv5s[J]. Electronic Design Engineering, 2024, 32(19): 140-144, 149.
- [18] 蒋畅江,刘朋,舒鹏.基于改进YOLOv5s的动态视觉SLAM算法[J].北京航空航天大学学报,2025,51(3): 763-771.
- JIANG C J, LIU P, SHU P. Dynamic visual SLAM algorithm based on improved YOLOv5s[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(3): 763-771.