

基于改进 YOLOv5 的旋转目标检测算法及其应用研究

沈中华¹, 陈万委¹, 甘增康^{2*}

(1. 桂林理工大学 机械与控制工程学院, 广西 桂林 541006;

2. 深圳职业技术学院 智能制造研究院, 广东 深圳 518055)

摘要: **目的** 提高工业分拣上常见的纹理多样、随机堆放盒体的检测与识别能力。**方法** 提出一种基于 YOLOv5 的旋转目标检测算法, 该算法包括目标分类、位姿角度识别和边界框位置定位 3 个模块功能。目标分类模块中, 通过自建数据集和设计 8 种目标分类标签以供模型分类学习; 位姿角度识别模块对 YOLOv5 头网络中添加角度预测分支, 引入环形平滑标签角度分类方法, 实现分拣盒体的旋转角度精准检测; 在边界框位置定位模块上, 使用 L_{CIoU} 回归框损失函数来计算旋转框回归损失, 得到紧密包裹目标位置的边界框。**结果** 改进的 YOLOv5 算法在自建数据集上检测精度达到 95.03%, 在机器人多物体分拣实验中的准确率可达 100%。**结论** 本文算法在盒体处于散乱堆放、密集堆放、堆叠堆放场景下均具有较高的定位与识别精度。

关键词: 杂乱盒体; YOLOv5; 环形平滑标签; 旋转角度检测

中图分类号: TP391 文献标识码: A 文章编号: 1001-3563(2023)19-0229-09

DOI: 10.19554/j.cnki.1001-3563.2023.19.030

Research on Rotating Target Detection Algorithm and Application Based on Improved YOLOv5

SHEN Zhong-hua¹, CHEN Wan-wei¹, GAN Zeng-kang^{2*}

(1. College of Mechanical and Control Engineering, Guilin University of Technology, Guangxi Guilin 541006, China;

2. Institute of Intelligent Manufacturing Technology, Shenzhen Polytechnic University, Guangdong Shenzhen 518055, China)

ABSTRACT: The work aims to improve the detection and recognition ability of boxes with various textures and scattered stacking, which are common in industrial sorting. A rotating target detection algorithm based on YOLOv5 was proposed. This algorithm included three modules: target classification, pose angle recognition and boundary box location. In the target classification module, self-built data sets and eight target classification labels were designed for model classification learning; In the pose angle recognition module, an angle prediction branch was added to the YOLOv5-head network. The angle classification method of circular smooth label was introduced to realize accurate detection of rotation angle of sorting boxes; In the boundary box location module, the L_{CIoU} regression box loss function was used to calculate the regression loss of the rotating box, and the boundary box that tightly wrapped the target position was obtained. The detection accuracy of the improved YOLOv5 algorithm in the self-built data set reached 95.03%. In the robot multi object sorting experiment, the accuracy rate reached 100%. The algorithm in this paper has high positioning and identification accuracy when boxes are in scattered, dense and stacked conditions.

KEY WORDS: messy Box; YOLOv5; ring smooth label; rotation angle detection

在当今电子商务盛行的背景下,仓储物流需要频繁地对商品进行保管、存储、配送及装卸搬运。如何使用机器人智能化和柔性化地完成商品盒的分拣、堆垛和拆垛操作,以此减轻仓储物流领域的人工劳动强度,成为各大电商和物流公司急需解决的问题^[1]。

利用视觉感知技术赋予分拣机器人随机分拣物品的能力是当今智能化物流的趋势。目前,大多数视觉机器人抓取系统仅能抓取相互分离的商品盒,而针对紧密堆放或遮挡的商品盒抓取还需进一步的研究。在随机堆放场景下,商品盒具有尺度变化大、排列紧密、角度方向多样、背景环境复杂等特点,使得检测系统难以获取目标平面坐标、旋转角度等位姿信息,导致分拣的准确性差,容易出现漏拣、误拣。

位姿估计是机器人实现随机抓取的关键问题,现有对位姿估计的研究方法大致可以分为三大类:第1类是基于模板匹配的方法^[2-4], Hinterstoisser 等^[2]提出的 linemod 算法是此类方法的代表,但该算法只考虑颜色梯度和深度信息构建特征,而对具有丰富纹理特征和遮挡的商品盒抗干扰性差;第2类是基于点云的方法,如郭交通^[5]使用 RANSAC 算法、法线聚类 and 欧式聚类来获取商品盒平面上的点云,以此来计算出目标物的位姿信息,但是使用点云重建的方法因存在光线、遮挡和相机性能差异等因素,重建出的结果不可避免存在或多或少的孔洞,影响最终位姿估计的精度;第3类是基于二维图像在分类、检测与分割任务中具有良好的鲁棒性,如将 Faster R-CNN、R-FCN、Mask-RCNN、YOLO 等神经网络迁移到抓取目标的位姿估计中^[6-10],但该类神经网络都是基于水平框的方式实现目标位置定位,无法获取随机堆放盒体旋转角度的信息。针对此问题诸多学者提出解决目标旋转角度预测问题的算法,如朱煜等^[11]提出一种粗调与细调相结合的旋转框检测网络 R²-FRCNN; Ding 等^[12]提出了 ROI-Transforme 检测网络; Yang 等^[13-14]提出旋转检测器 SCRDet、R3Det。但上述的算法都无法避免在旋转框回归过程中存在角度周期性和边缘交换性的问题,因此 Yang 等^[15]提出将目标旋转角度预测转化为角度分类问题,并设计环形平滑标签作为模型的预测结果,但该方法多用于遥感场景下的图像处理,对分拣目标的角度预测还需进一步研究。

本文针对现有算法对分拣盒体在随机堆放场景下的状态多变、随机性强,且密度较高时检测与识别能力差,提出一种基于 YOLOv5 的旋转框定位目标检测算法,用于分拣盒体的旋转角度预测。该算法将目标物的角度回归问题转化为分类问题,并引入环形平滑标签角度分类方法来实现分拣盒体角度的精准预测。使用 L_{Clou} 回归框损失函数衡量旋转框回归损

失,解决旋转目标检测模型收敛速度慢和稳定性差的问题。

1 算法描述

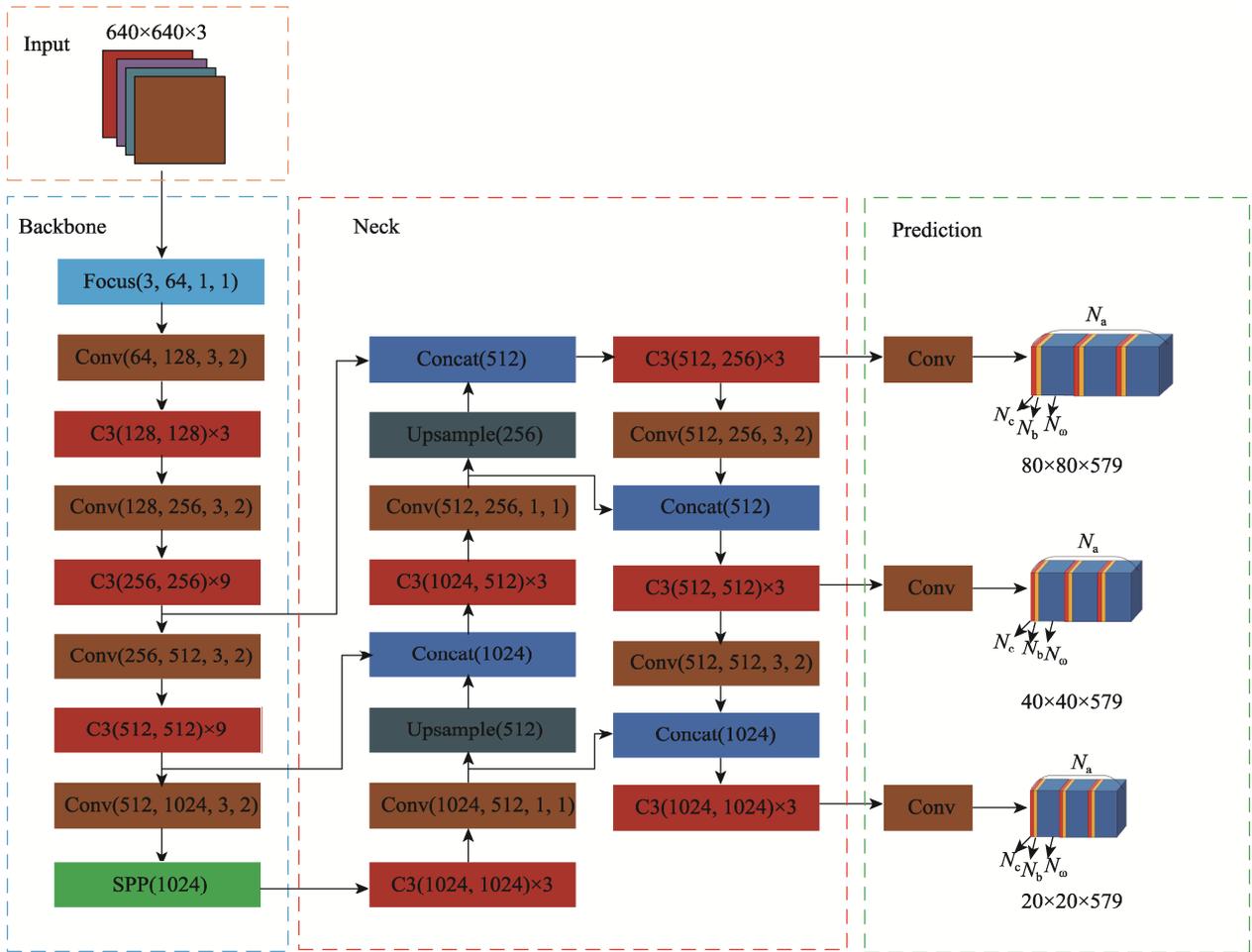
1.1 YOLOv5 目标检测算法

YOLOv5 主要由输入端、主干网络 Backbone、Neck 和预测端 Prediction 等 4 个部分组成^[16],其中输入端通过自适应图片缩放算法将图片统一缩放到一个标准尺寸。主干网络 Backbone 由 Focus、CBL (Convolution-Batch Normalization-Leak ReLU)、BottleneckCSP 以及空间金字塔池化模型 (Spatial Pyramid Pooling, SPP)^[17]组成,用以提取输入图像的特征。Focus 模块将图像 W、H 的信息集中到通道上,再使用卷积模板对其进行特征提取,减少图像特征信息的丢失。BottleneckCSP 模块基于残差网络的思想,设计了 2 条分支,一条分支用于卷积操作,另一条分支用于不同层次的特征融合,在保证网络预测准确率的同时降低了计算量。SPP 模块由不同卷积核大小的池化层组成,作用是实现多重感受野融合,提升网络在复杂环境下的检测能力。Neck 部分由 PAN 网络^[18]构成,通过从上向下和从下向上的路径实现了不同尺寸特征图的融合,丰富了特征信息。预测端 Prediction 输出 3 个不同大小的特征向量用于对目标的预测。

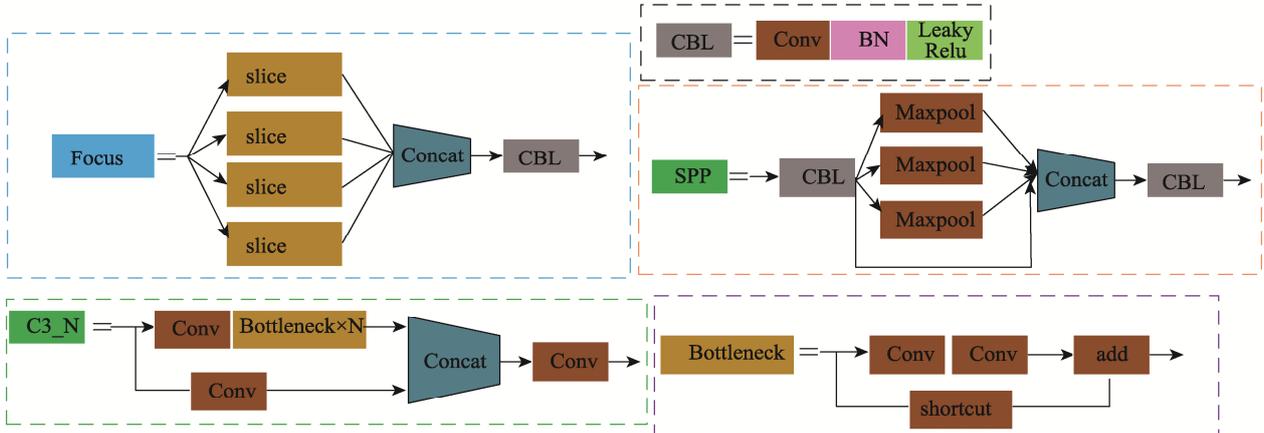
YOLOv5 虽然广泛应用于各种目标检测任务中,但将其应用于随机堆放盒体的分拣场景中仍存在一些问题:分拣盒体目标的长宽比例一般较大并且具有方向性,若采用水平框会忽略分拣盒体的方向信息,无法对分拣物进行精确定位;在流水线作业过程中,待分拣盒体之间距离较近,使用水平框回归在后处理环节中会造成各盒体间的像素相互抑制,容易造成漏检;基于深度学习的随机分拣场景中,神经网络在学习过程中采样点容易落在目标物外,不利于模型的收敛。针对上述问题,本文以 YOLOv5 为基础引入了角度预测分支和旋转框回归损失计算方法,改进后的 YOLOv5 模型如图 1 所示。

1.2 增加角度预测分支

针对机器人在分拣过程需要获取分拣物的位姿信息,在 YOLOv5 的基础上添加角度预测分支。本实验算法采用五参数 (x, y, w, h, θ) 来表示分拣目标旋转矩形框,旋转矩形框采用长边定义法,其中 (x, y) 表示预测目标中心点坐标,矩形框的长边定义为 w ,短边定义为 h 。以图片的左上角为原点,水平向右的方向为 x 轴,垂直向下的方向为 y 轴,建立坐标系。旋转角 θ 定义为 x 轴绕顺时针方向旋转至与商品盒长边平行所转过的角度,旋转角 θ 的定义图如图 2 所示。



a 本文算法总体框架



b 算法中不同模块计算流程

图 1 改进后 YOLOv5 网络结构
Fig.1 Improved YOLOv5 network structure

根据圆形平滑标签(Circular Smooth Label, CSL)方法的启示^[15],可以将分拣箱体旋转角度预测视为分类问题。由于深度学习分类任务中多使用独热编码(one-hot)的方法来标注数据,但在本实验中直接使用独热编码会造成数据稀疏问题,因此通过生成服从高斯函数的标签来代替独热编码后的标签。以 91°为例,角度编码示意图如图 3 所示,角度分布图像呈环形分布,0°和 180°的结果相同,则 CSL 的角度范围

为 $[0^\circ, 179^\circ]$, CSL 的具体公式如下:

$$CSL(x) = \begin{cases} \exp\left(-\frac{(x-\mu)^2}{2\sigma_\lambda^2}\right) & \mu-r < x \leq \mu+r \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

式中: x 和 μ 分别为在 $[0^\circ, 179^\circ]$ 间的连续变量和箱体实际旋转角度; r 为窗口函数的半径(本次实验选用高斯函数)。

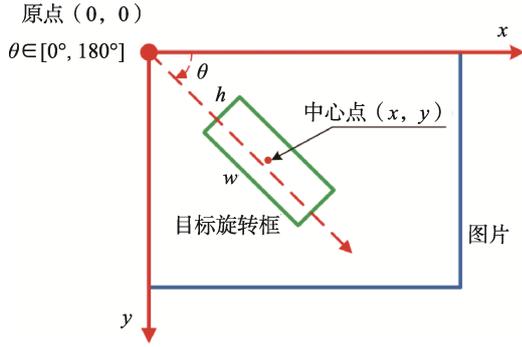


图2 旋转角 θ 定义
Fig.2 Diagram for definition of rotation angle θ

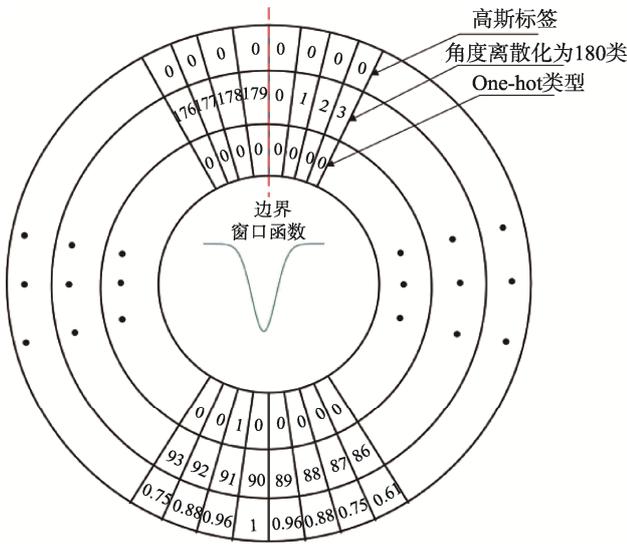


图3 角度编码示意图
Fig.3 Angle coding diagram

对于类别损失计算问题,本文采用二分类交叉熵损失函数对180个角度类别进行分类,定义如下:

$$\begin{cases} y_{n,i} = \text{csf}(x) \\ l_{n,i} = -[y_{n,i} \ln(\alpha_{n,i}) + (1 - y_{n,i}) \ln(1 - \alpha_{n,i})] \\ l_n = \sum_{i=0}^{179} l_{n,i} \\ \text{loss}(z, y) = \frac{1}{n} \sum_{i=0}^{n-1} l_n \end{cases} \quad (2)$$

式中: n 为模型在一个batch_size条件下的训练样本数量; i 为角度类别,由于 0° 和 180° 的结相同, $i \in [0^\circ, 179^\circ]$; $\sigma_{n,i}$ 为模型将第 n 个样本预测为第 i 个角度的概率; $y_{n,i}$ 为第 n 个样本下第 i 个角度的标签。

YOLOv5的网络结构主要分为主干网络、颈部网络和头部网络,添加角度预测分支后,模型在头部网络中输出通道变为 M 个, M 的计算见式(3)。

$$M = (N_c + N_b + N_\omega) N_a \quad (3)$$

式中: N_a 为锚框数量; N_c 为盒体类别数; N_b 为旋转矩形框参数(x, y, w, h, c),即旋转矩形框的中心点(x, y)、宽高(w, h)和置信度(c); N_ω 为角度类别,共180个类别。文中 N_a 为3, N_c 为8, N_b 为5, N_ω 为180,因此 $M=579$ 。

1.3 旋转框回归损失计算方法

在YOLOv5的头部网络中添加角度预测分支后,还需要明确目标旋转框损失计算方法。它是用来度量YOLOv5预测值与真实标签值的差距,预测值越接近真实标签值,损失函数值越小。损失函数的选择关系到模型训练的稳定性 and 收敛速度。本文选用 L_{Clou} 损失函数作为目标旋转框回归损失函数,并通过下文的对照组实验来验证其性能。损失函数 L_{Clou} 既考虑到标注矩形框 A 与预测矩形框 B 的交并比(Intersection over Union, I_{IoU}),还将中心点距离、宽高比加入损失计算中,如图4所示。 L_{Clou} 损失函数的计算见式(4)和式(5)。

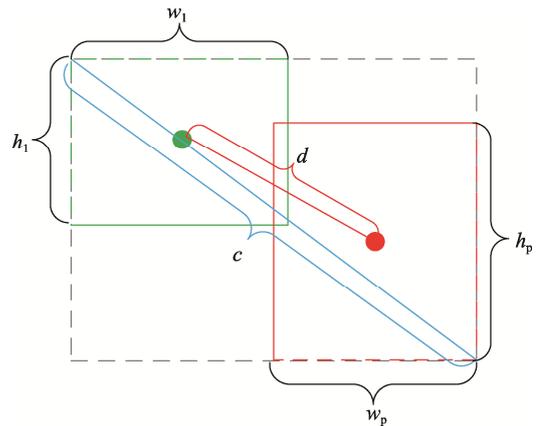


图4 L_{Clou} 损失函数
Fig.4 L_{Clou} loss function

$$I_{\text{IoU}} = \frac{A \cap B}{A \cup B} \quad (4)$$

式中: I_{IoU} 为预测框和真实框的IoU值; A 为真实框; B 为预测框。

$$\begin{cases} L_{\text{Clou}} = 1 - I_{\text{IoU}} + \frac{\rho^2(a, b)}{c^2} + \sigma v \\ \sigma = \frac{v}{1 - I_{\text{IoU}} + v} \\ v = \frac{4}{\pi^2} (\arctan \frac{w_1}{h_1} - \arctan \frac{w_p}{h_p}) \end{cases} \quad (5)$$

式中: a 为真实框 A 的中心点坐标; b 为预测框 B 的中心点坐标; ρ 为计算 a 点与 b 点的欧式距离; c 为最小对角线距离; w_1 和 h_1 分别为真实框的宽和高; w_p 和 h_p 分别为预测框的宽和高。

2 实验结果与分析

2.1 分拣系统实验平台

机器人分拣实验平台主要包括 IntelRealSense D435 深度感知相机、敬科六轴机械臂、吸盘、服务端主机和机械臂控制系统的客服端主机等。将敬科六轴机械臂固定在实验平台一侧, 并将深度感知相机固定在机械臂末端的气动执行器上。相机通过 USB 线与服务端主机相连, 实现图像数据的传输。服务端主机用于部署基于 pytorch 深度学习框架构建的旋转目标检测算法, 其硬件配置为 i7-11700 CPU, NVIDIA T1000, 4 GB GPU。选用 YOLOv5x 作为模型的基础网络, 并使用 COCO 数据集上的预训练权重对模型参数进行初始化, 提高计算效率。对模型添加角度预测分支后, 模型的深度为 407 层, 模型参数量为 8.97×10^7 , 计算量为 2.24×10^{11} 的浮点运算数 (Floating Point of Operations, FLOPs)。服务端的电脑与客服端以 TCP 的方式建立通信, 机械臂根据盒子中心点的位置坐标, 并按照分拣顺序执行抓取操作。

2.2 实验数据集

当分拣箱体处于随机堆放的状态时, 具有状态多变、随机性强、密度高等特点。针对此情况, 以牙膏盒、饼干盒、剃须刀盒、相机盒、茶盒、药盒、灰度盒及小型箱体为研究对象, 模拟实际分拣物的多种随机堆放状态 (散乱、密集、遮挡、堆叠等), 利用 RealSense D435 相机进行拍摄, 拍摄角度尽可能模拟 eye-in-hand 手眼系统分拣场景, 采集的单张图片中包含多个目标, 拍摄过程中随意改变目标摆放角度, 使得数据集的目标旋转角度能覆盖 $[0^\circ, 179^\circ]$ 。不同长

宽比对旋转角度变化的敏感程度有所差异, 如图 5 所示, 实线框代表目标物的真实框, 虚线框代表模型的预测框, 斜线代表两框的交集。由图 5 可知, 长宽比为 4 的物体旋转角度由 30° 变为 60° 时, 交并比降低了 49.5%, 而长宽比为 1.42 的物体旋转角度由 30° 变为 60° 时, 交并比仅降低了 10.6%。由此可知, 算法对长宽比较大的商品盒的敏感程度更高。牙膏盒、饼干盒、药盒、灰度盒及小型盒的长宽比较大, 算法对其角度预测较为敏感, 而饼干盒、剃须刀盒、相机盒的长宽比较小, 算法对其角度预测敏感性欠佳。多种类箱体混淆, 更能体现模型的目标分类和旋转框回归性能。数据集共包含 1 667 张分辨率为 $1\ 280 \times 720$ 的 RGB 图像, 采用 6:2:2 的比例划分训练集、验证集、测试集。

YOLOv5 属于监督型神经网络, 在网络训练前使用 rolabelimg 标注软件对数据集图片进行旋转框标注, 并使用标签名 (toothpaste、cookie、razor...) 对箱体进行分类识别与区分, 标注后的标签类别和旋转框参数保存到与原图相对应的 xml 格式文件中。后期再将 xml 格式文件转换为 YOLOv5 标签所需的 txt 格式文件, 并对旋转框参数 (x 、 y 、 w 、 h) 基于图片的宽高进行归一化处理。为了降低模型过拟合的风险和增强图像样本中的背景复杂度, 本文采用 Mosaic 数据增强方法, 对传入的图像进行随机翻转、缩放、裁剪等处理, 将得到的图像再进行依次拼接成新图输入模型训练。

训练参数设置: 图片分类个数为 8; 图片输入模型前大小更改为 640×640 ; 采用 SGD 优化算法进行参数的更新; 批量大小为 1; 最大迭代次数为 200; 使用余弦退火策略动态调整每次迭代的学习率, 初始学习率设为 0.01。

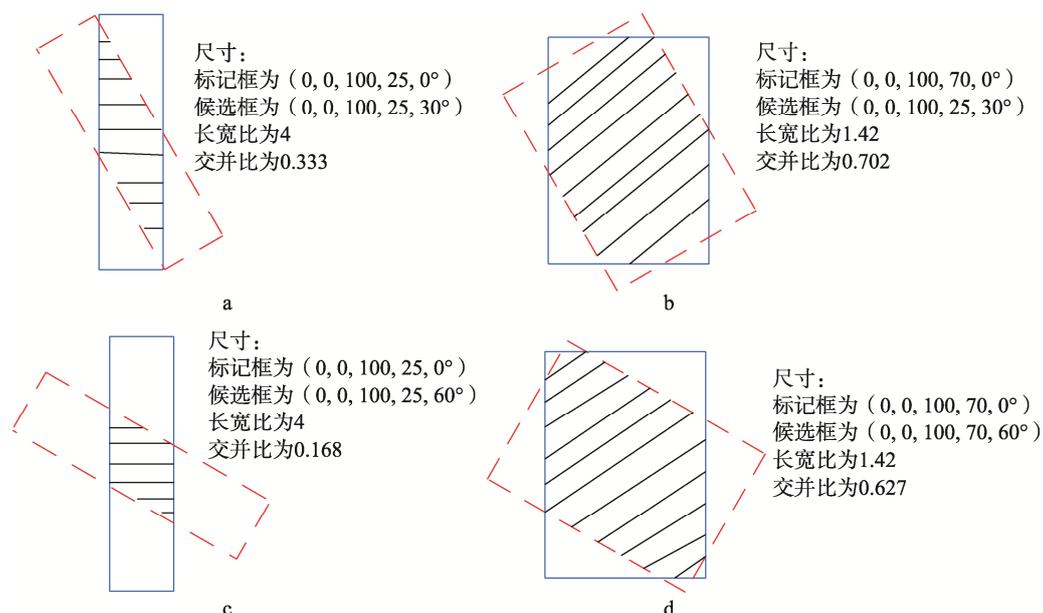


图 5 不同长宽比物体的 I_{IoU}
Fig.5 IoU of objects with different aspect ratios

2.3 实验结果

2.3.1 模型训练结果

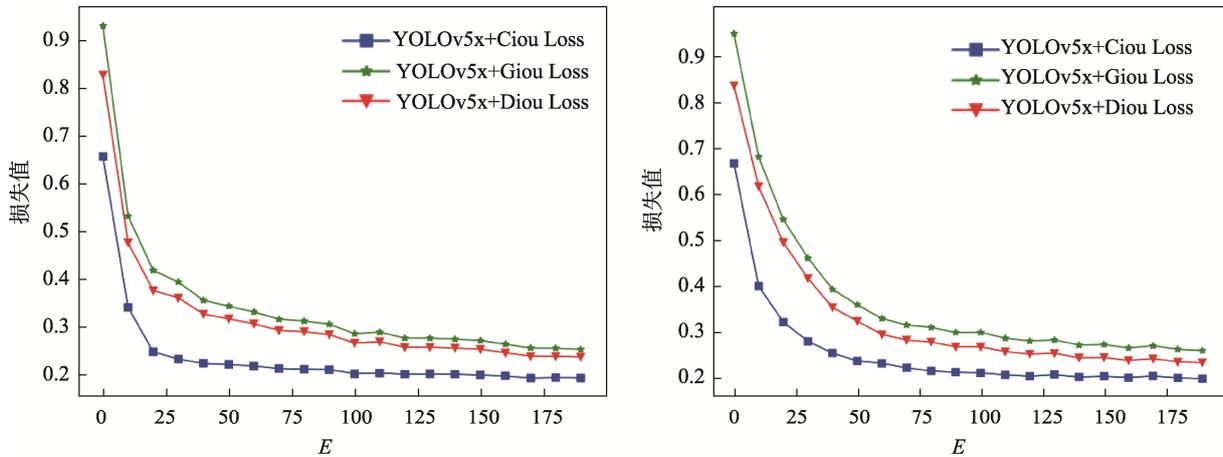
模型性能对机器人分拣的识别与定位精度有直接的影响。为了提高模型的收敛速度与检测精度，本文分析角度预测分支与回归框损失函数对 YOLOv5 模型的影响。

为了分析损失函数对模型训练的稳定性及收敛速度的影响，在不同损失函数条件下，开展对照组实验。本文选择 L_{GIoU} 损失函数、 L_{DIoU} 损失函数和 L_{CIoU} 损失函数建立对照组实验。如图 6 所示为不含 CSL 角度预测分支与含 CSL 角度预测分支的 YOLOv5 模型在不同回归框损失条件下的训练结果曲线。图 6 中横轴 E 为模型训练的迭代次数，纵轴 loss 为模型训练损失值。从训练曲线可知，在模型中添加角度预测分支后损失计算过程会加入角度损失函数，模型迭代 100 次左右趋于平稳，相较于原模型增加了 50 次迭代。使用 L_{CIoU} 函数作为模型回归框损失函数，会使得模型的稳定性和收敛性提高，收敛迭代次数相较于 L_{GIoU} 损失函数和 L_{DIoU} 损失函数减少了 25 次迭代。从

表 1 模型检测精度可知，添加角度预测分支后，模型对 8 类盒体的检测精度均有效提高，评价指标平均精度均值 (mean Average Precision, mAP) 提升了 4.92%。其中，对剃须刀盒、茶盒和药盒的检测性能提升最明显，剃须刀盒的平均精度 (Average Precision, AP) 提升了 6.9%，茶盒和药盒的 AP 值分别提升了 7.63%、7.58%。

2.3.2 模型检测效果

使用自建数据集的模型检测结果如图 7~9 所示。本文算法 (图中紧贴目标轮廓的旋转框) 和 YOLOv5 算法 (图中水平框) 的检测结果进行比较表明，使用本文算法更能清楚地区分随机堆放场景下每个盒体的位置和方向。无论目标物是散乱堆放，还是密集堆放或堆叠堆放时，基于 YOLOv5 的旋转目标检测算法的回归框都能紧密包裹目标物，有效避免邻近目标间的像素干扰，从而提升此场景下的箱体特征提取能力，提高模型的检测精度。无添加角度预测分支的原 YOLOv5 算法在散乱堆放时回归框还能各自分离，但当目标物处于密集堆放或堆叠堆放时，相邻回归框出现严重的交叠和像素干扰。



a 无添加CSL分支模型训练损失曲线 b 添加CSL分支的模型训练损失曲线

图 6 各模型训练曲线
Fig.6 Training curve of each model

表 1 不同模型检测精度比较
Tab.1 Comparison of detection precision of different models.

角度预测	AP 值/%								mAP 值/%
	T_box	C_box	R_box	Ca_box	Te_box	P_box	B_box	S_box	
-	91.67	88.17	89.26	89.82	88.12	89.27	93.09	91.46	90.11
√	94.22	91.52	95.42	92.84	94.85	96.04	98.29	97.05	95.03

注: T_box、C_box、R_box、Ca_box、Te_box、P_box、B_box、S_box 分别表示牙膏盒、饼干盒、剃须刀盒、相机盒、茶盒、药盒、灰度盒及小型盒体; mAP 值表示 I_{IoU} 在 0.5 时, 模型能识别的所有类 AP 值的均值; “-”表示不使用 CLS 角度预测分支, “√”表示使用 CLS 角度预测分支。

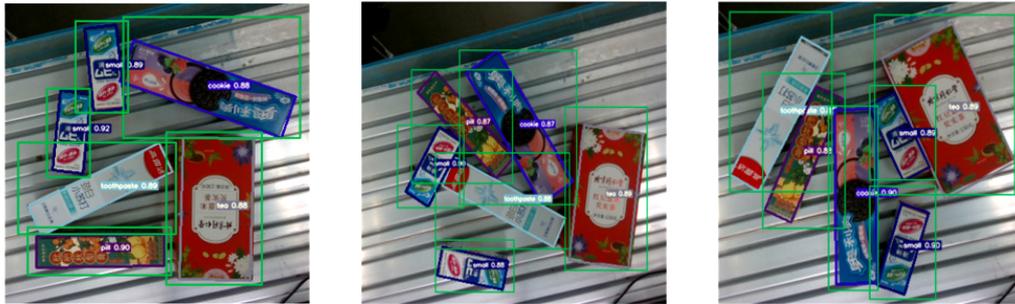


图 7 散乱堆放盒体检测效果
Fig.7 Detection effect of scattered boxes

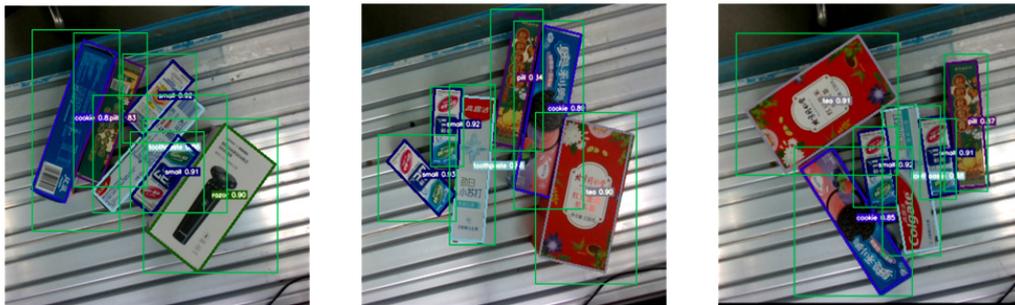


图 8 密集堆放盒体检测效果
Fig.8 Detection effect of densely stacked boxes



图 9 堆叠堆放盒体检测效果
Fig.9 Detection effect of blocked or stacked boxes

2.3.3 机器人分拣实验

为了验证旋转目标检测网络在机器人分拣场合下的真实性能,采用如图 10 所示的机器人实验平台进行算法模型的部署。本文以散乱堆放、密集堆放、堆叠堆放 3 种场景开展无序分拣实验。每类盒子各进行 10 次实验,规定在全部盒体按照预测得分高低的顺序抓取,且不造成其他盒体跌落时认为本次分拣实验成功。结果统计如表 2 所示,可以看出,本文所构建的旋转目标检测网络在散乱堆放、密集堆放场景下 8 类盒体的分拣成功率均为 100%。在堆叠堆放时,茶盒和灰度盒准确率有所下降,原因为茶盒体积过大抓取过程中有时候会造成其他盒体跌落,而灰度盒由于表面纹理特征单一,重度堆叠时有时候出现漏拣。使用水平框回归的原 YOLOv5 模型在散乱堆放场景下的成功率为 100%,但在密集堆放和堆叠堆放场景分拣的准确率不高,其原因为水平框回归无法紧密包裹分拣物,而机器人抓取点为水平框的中心点,当中心点落在背景或其他盒体上

时,机器人无法有效抓取到目标。在添加了角度预测分支后,回归框紧密贴合分拣物,很好地把分拣物特征与背景分割出来。因此,相比于原 YOLOv5 模型本,文算法分拣成功率更高,识别与定位精度更准。

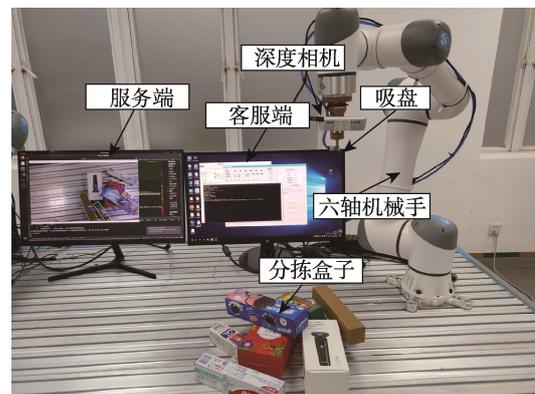


图 10 机器人分拣实验平台
Fig.10 Experimental platform of robot sorting

表2 箱体成功抓取率统计
Tab.2 Successful grab rate of boxes

%

堆放情形	算法	T_box	C_box	R_box	Ca_box	Te_box	P_box	B_box	S_box
散乱堆放	YOLOv5	100	100	100	100	100	100	100	100
	本文算法	100	100	100	100	100	100	100	100
密集堆放	YOLOv5	100	90	100	100	100	100	100	90
	本文算法	100	100	100	100	100	100	100	100
堆叠堆放	YOLOv5	80	90	100	90	100	100	80	70
	本文算法	100	100	100	100	90	100	90	100

3 结语

在仓储物流行业,需要频繁地对商品进行分销、拆箱存储、定量配送。为了降低人工劳动强度,提高工作效率,基于视觉感知的机器人技术在仓储物流领域应用越来越广泛。本文针对非结构化场景中堆叠或遮挡的多类商品盒算法定位与识别精度差的问题,提出了一种基于YOLOv5的旋转目标检测算法,在原有YOLOv5模型的基础上,增加旋转角度预测分支和寻找最优边界框回归损失函数的方式,达到了模型的最优训练效果。本文基于训练好的模型搭建机器人多物体分拣系统,该系统利用旋转目标检测网络获得商品盒类别、位置和旋转角度,将模型预测得分作为机器人分拣顺序,实现了机器人在散乱堆放、密集堆放以及堆叠堆放场景下自主、安全和稳定的分拣。证明了本文算法在复杂工况下具有较高的定位与识别精度。

参考文献:

- [1] 黄冰鹏,林义忠,杨中华,等.码垛机器人的研究与应用现状[J].包装工程,2017,38(5):82-87.
HUANG Bing-peng, LIN Yi-zhong, YANG Zhong-hua, et al. Research and Application Situation of Palletizing Robots[J]. Packaging Engineering, 2017, 38(5): 82-87.
- [2] HINTERSTOISSER S, LEPETIT V, ILIC S, et al. Model Based Training, Detection and Pose Estimation of Texture-Less 3d Objects In Heavily Cluttered Scenes[C]// Asian Conference on Computer Vision, Springer, Berlin, Heidelberg, 2012: 548-562.
- [3] RIOS-CABRERA R, TUYTELAARS T. Discriminatively Trained Templates for 3d Object Detection: A Real Time Scalable Approach[C]// Proceedings of the IEEE International Conference on Computer Vision, 2013: 2048-2055.
- [4] HODANĚ T, ZABULIS X, LOURAKIS M, et al. Detection and Fine 3D Pose Estimation of Texture-Less Objects in RGB-D Images[C]// Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, IEEE, 2015: 4421-4428.
- [5] 郭交通.基于点云的目标识别和位姿估计方法研究[D].南京:南京航空航天大学,2019.
GUO Jiao-tong. Research on Object Recognition and Pose Estimation Based on Point Cloud[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019.
- [6] 伍锡如,黄国明,孙立宁.基于深度学习的工业分拣机器人快速视觉识别与定位算法[J].机器人,2016,38(6):711-719.
WU Xi-ru, HUANG Guo-ming, SUN Li-ning. Fast Visual Identification and Location Algorithm for Industrial Sorting Robots Based on Deep Learning[J]. Robot, 2016, 38(6): 711-719.
- [7] 杜学丹,蔡莹皓,鲁涛,等.一种基于深度学习的机械臂抓取方法[J].机器人,2017,39(6):820-828.
DU Xue-dan, CAI Ying-hao, LU Tao, et al. A Robotic Grasping Method Based on Deep Learning[J]. Robot, 2017, 39(6): 820-828.
- [8] 夏晶,钱堃,马旭东,等.基于级联卷积神经网络的机器人平面抓取位姿快速检测[J].机器人,2018,40(6):794-802.
XIA Jing, QIAN Kun, MA Xu-dong, et al. Fast Planar Grasp Pose Detection for Robot Based on Cascaded Deep Convolutional Neural Networks[J]. Robot, 2018, 40(6): 794-802.
- [9] 杨劲松.基于深度学习的多物体分拣系统研究[D].长春:吉林大学,2021.
YANG Jin-song. Research on Multi-Object Sorting System Based on Deep Learning[D]. Changchun: Jilin University, 2021.

- [10] 董红召, 方浩杰, 张楠. 旋转框定位的多尺度再生物品目标检测算法[J]. 浙江大学学报(工学版), 2022, 56(1): 16-25.
DONG Hong-zhao, FANG Hao-jie, ZHANG Nan. Multi-Scale Object Detection Algorithm for Recycled Objects Based on Rotating Block Positioning[J]. Journal of Zhejiang University (Engineering Science), 2022, 56(1): 16-25.
- [11] 朱煜, 方观寿, 郑兵兵, 等. 基于旋转框精细定位的遥感目标检测方法研究[J]. 自动化学报, 2023, 49(2): 415-424.
ZHU Yu, FANG Guan-shou, ZHENG Bing-bing, et al. Research on Detection Method of Refined Rotated Boxes in Remote Sensing[J]. Acta Automatica Sinica, 2023, 49(2): 415-424.
- [12] DING J, XUE N, LONG Y, et al. Learning Roi Transformer for Oriented Object Detection in aerial Images[C]// 2019 IEEE Conference on Computer Vision and Pattern Recognition, Long Bench: IEEE, 2019: 2849-2858.
- [13] YANG X, YANG J, YAN J, et al. Scrdet: Towards More Robust Detection for Small, Cluttered and Rotated Objects[C]// 2019 IEEE/CVF International Conference on Computer Vision. South Korea: IEEE, 2019: 8232-8241.
- [14] YANG X, YAN J, FENG Z, et al. R3det: Refined Single-Stage Detector with Feature Refinement for Rotating Object[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(4): 3163-3171.
- [15] YANG X, YAN J. On the Arbitrary-Oriented Object Detection: Classification Based Approaches Revisited[J]. International Journal of Computer Vision, 2022, 130(5): 1340-1365.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [17] HE Kai-ming, ZHANG Xiang-yu, REN Shao-qing, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [18] LI H, XIONG P, AN J, et al. Pyramid Attention Network for Semantic Segmentation[J]. arXiv Preprint, 2018: 11-12.

责任编辑: 曾钰婵