

基于聚类与 SDC 算法的智能仓储系统 分类识别技术研究

侯晓辉, 王金山, 吴艳明, 程 嵩, 齐佳丽

(大唐(赤峰)新能源有限公司, 内蒙古 赤峰 024000)

摘要: 该文以 YOLOv8 为基本框架, 在主干框架中添加 SDC 增加图像特性提取能力, 并在检测头中添加改进损失函数强化位置识别能力; 使用聚类算法和神经网络耦合来输出货物信息。该文提出的改进模型在 5 种不同类别货物检测中, 相比于无锚框算法 YOLOv8, PA 值分别提升了 8.04%, 8.26%, 8.18%, 7.49% 和 7.41%, mPA 值提升了 7.35%。在图像特征采集后, 基于概率的聚类分析展现出良好地分类效果, 对于货架层、地面层和地面货架混合层的货物进行测试。货架层和地面层的识别准确率分别为 97.1% 和 96.8%。而混合环境下, 准确率为 94.2%。

关键词: 智能仓储; 聚类分析; SDC; 机器视觉; 分类识别

中图分类号: TP39

文章编号: 1000-0682(2026)02-0084-06

文献标识码: A

DOI: 10.19950/j.cnki.CN61-1121/TH.2026.02.016

Research on classification and recognition technology of intelligent warehouse system based on clustering and SDC algorithm

HOU Xiaohui, WANG Jinshan, WU Yanming, CHENG Song, QI Jiali

(Datang(Chifeng) Renewable Co., Ltd., Inner Mongolia Autonomous Region Chifeng 024000, China)

Abstract: This paper takes YOLOv8 as the basic framework, adds SDC in the main framework to enhance the ability to extract image features, and adds an improved loss function in the detection head to strengthen the position recognition ability. The clustering algorithm and the neural network coupling algorithm are used to output the goods information. The improved model proposed in this paper has increased the PA values by 8.04%, 8.26%, 8.18%, 7.49% and 7.41% respectively compared with the anchor-free algorithm YOLOv8 in the detection of five different types of goods, and the mPA value has increased by 7.35%. After image feature collection, probability-based cluster analysis demonstrated excellent classification results, and tests were conducted on goods in the shelf layer, ground layer, and mixed ground and shelf layers. The accuracy rates of goods identification on the shelf layer and the ground layer were 97.1% and 96.8% respectively. In a mixed environment, the accuracy rate was 94.2%.

Keywords: intelligent warehouse; cluster analysis; SDC; machine vision; classification and recognition

0 引言

能源转型是实现双碳战略目标的重要手段^[1]。新能源发电得到大规模发展, 而新能源发电厂具有

站点多面广、空间跨度大、备件种类多的特点, 这给仓储管理增加了难度。传统仓储管理主要依靠人工进行操作, 而随着仓储物品种类和数量的增加, 逐渐暴露出人工操作效率低和准确度不高的问题。随着自动化控制和信息处理技术的发展, 传统仓储向智能仓储转型成为未来发展的趋势^[2-3]。

智能仓储将自动化控制设备与人工智能算法相结合, 可减少人工成本, 实现物品自动化识别分类等功能^[4]。而智能仓储的核心便是机器视觉在自动

收稿日期: 2025-11-25

基金项目: 大唐(赤峰)新能源有限公司科技项目“新能源智能仓储管理体系建设”(DTNM-2024-20340)

第一作者: 侯晓辉(1974—), 男(蒙古族), 内蒙古赤峰人, 本科, 高级经济师, 研究方向为智能仓储。

化中的应用,通过高清摄像头实时捕捉货物图像信息,并结合深度学习对不同货物进行识别和分类^[5-7]。郑建彬等^[8]学者通过高清摄像头收集盒装乳制品图像,并结合改进的图像特征提取算法进行分析,发现可以有效降低识别错误率。但是该方法适用性较窄,只能对盒装物品进行识别。Yilmazer R等^[9]学者对模型进行改进,提出半检测的自动化检测模型,实例验证发现可以有效识别混合状态下的货物信息,准确判断出货架状态,但是面对货物种类较多或仓库布局复杂的情况,识别能力明显下降。可再生发电厂仓库存储物品种类繁多,仓库布局复杂,如何针对形状、大小、种类各异的物品进行准确识别分类成为研究的难点。

该文针对复杂环境下多尺度图像特征如何提取的难题,基于机器视觉仪器捕捉图像的方法,提出以YOLOv8为基本框架,在主干框架中添加空间可形变卷积(SDC)来对复杂图像进行特征提取来提高识

别准确度,同时在检测头中采用混合损失函数来增强对物品位置的确定。最后采用聚类算法处理图像信息,对不同物品进行分类并结合神经网络输出物品信息,最终实现复杂环境下多尺度物品的识别和分类。

1 多尺度图像特征提取

智能仓储的系统框架如图1所示,整体通过在仓库区域布置硬件对货物信息进行实时的捕获,然后通过计算机网络将数据信息进行分类管理。供应链协同平台根据传递的信息进行操作反馈。而在框架中最重要的便是准确识别货物信息,并根据货物特征进行分类^[10]。在传统仓储中通常是由人来进行识别并录入分类信息,这一过程需要大量时间和物力,并且难以保证准确度^[11-12]。机器视觉设备可以代替人眼,有效监测仓库状态,通过对捕捉到的图像进行分析得到图像特征,并自动实现识别和分类^[13]。

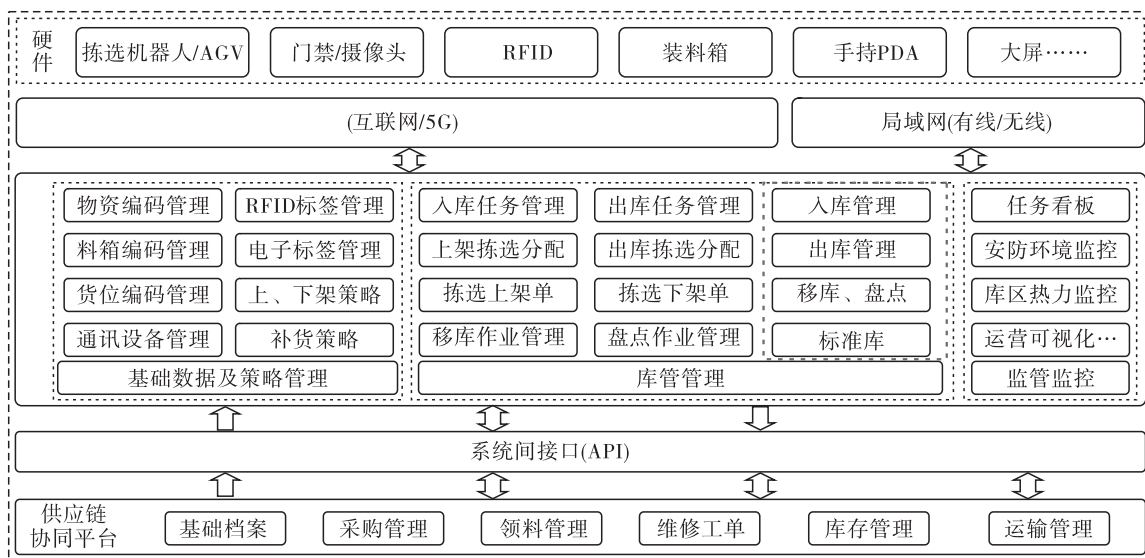


图1 智能仓储的系统框架图

图像特征提取过程如图2所示。在YOLOv8主干上通过3×3标准卷积(CNN)对图像特性进行采样,经过初步简化后再通过3×3SDC进行进一步地特征提取。

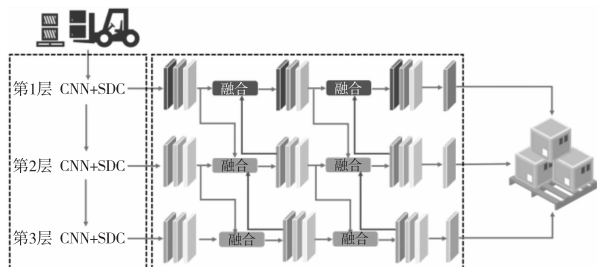


图2 图像特征的提取过程

CNN是常用的图片特征提取算法,在目前图像特征提取中展现出了良好的效果。CNN是一个典型的人工神经网络模式,在隐藏层中通过人为设置来训练调控样本数据,实现非线性关联。但是新能源发电厂仓库布局必须根据现场实际进行划分,具有复杂性和多样性,同时仓库需要存储从大到小、形状各异的货物,这对图像特征提取提出了更高的要求。CNN针对多尺度和复杂环境图像特征提取效果并不是很好。因此该文提出SDC来提高分辨能力。SDC是CNN改进算法,其提取图像特征主要包括了位置确认和可形变卷积提取两个过程。为了确认仓库货物位置,首先将三维空间信息进行定义设

置,如图 3 所示。

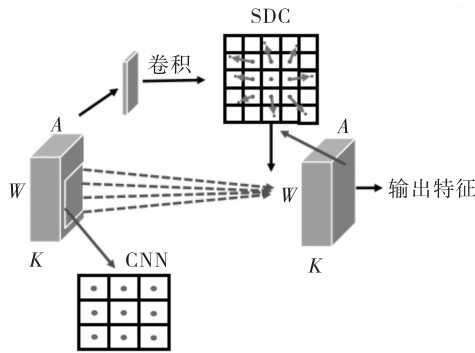


图 3 CNN - SDC 特征提取过程

A, K 和 W 分别表示高清摄像头捕捉到图像的通道数、高度和长度。因此图像位置信息可以由式(1)表示:

$$X' = \text{Connet}(X, C_i, C_j) \quad (1)$$

式中:Connet 代表数据信息连接处理; X 为特征图; C_i 和 C_j 为坐标通道; i 为 K 和 W 矩阵,第 1 行设置为 0,随着行数增加不断增加。 j 同样为 K 和 W 矩阵,矩阵列为常数。然后对矩阵进行线性化处理,保证区间在 $-1 \sim 1$ 。

明确图像位置信息后,需要对图像信息进行卷积提取,可形变卷积对复杂图像信息的提取具有良好效果,特征提取过程如式(2)所示:

$$Y(p) = \sum_{p_n \in r} w(p_n) \cdot X'(p + p_n + \Delta p_n) \quad (2)$$

式中: $Y(\cdot)$ 为输出特征; p 为图像具体的位置; $w(p_n)$ 为位置 p_n 的占比权重; p_n 通过枚举法,将 r 中位置一一进行表示。 Δp_n 代表可形变卷积过程中的偏差值。通过数据库学习来不断调整偏差值,可形变卷积算法可以对不同形状的货物进行信息提取。

2 边界损失函数的改进

YOLO 中的损失函数通常采用 CIoU 损失函数^[10]。CIoU 损失函数重的边界损失通常采用交并比来进行表示。交并比损失主要反映了物体的长宽比,并通过长宽比不同来进行位置确定。但是缺乏对具体位置与预测区间的实际差距描述。为了解决这一问题,一种新的方法归一化瓦瑟斯坦距离被提出。该方法利用高斯分布和瓦瑟斯坦距离来定量表征预测区间和实际区间的相似性,并且有效捕捉小目标的位置信息。为了提高多尺度复杂情况下货物边界的定位准确度,将传统的交并比损失和归一化瓦瑟斯坦距离进行结合。交并比损失可以由式(3)

表示:

$$L_{\text{CIoU}} = 1 - R_{\text{IoU}} + \frac{\rho(b, b^{Bt})}{c^2} + au \quad (3)$$

式中: ρ 为预测区间中心点和实际区间中心点的距离; c 为货物对角线距离; a 和 u 为不同货物长宽比的差异; R_{IoU} 为交并比。

归一化瓦瑟斯坦距离损失如式(4)所示:

$$L_{\text{NWD}} = \exp\left(-\frac{\sqrt{A_2^2(M_B, M_{Bt})}}{\tau}\right) \quad (4)$$

式中: M_B 为预测区间的高斯分布情况; M_{Bt} 为与预测区间对应的实际区间的高斯分布情况。 A_2^2 为预测与实际区间的瓦瑟斯坦距离,它的具体含义为预测与实际区间的高斯分布距离; τ 为图像数据集的系数值。

改进后的边界函数如式(5)所示:

$$L = \gamma L_{\text{CIoU}} + (1 - \gamma) L_{\text{NWD}} \quad (5)$$

式中: L 为混合损失函数; γ 为权重系数,经过超参数调试后,选取为 0.4。

边界损失被用来确定目标图像位置。改进后的损失函数可以有效确定尺寸形状复杂的图像,提高了模型对多尺度复杂环境下货物的辨别能力。

3 基于概率的聚类分析

通过高清摄像头捕捉仓库图像信息,然后通过 CNN + SDC 对图像特征进行提取后,该文采用聚类分析法对特征信息进行处理,实现对货物的自动识别和分类。图像的特征信息首先经过高斯混合近似处理,如式(6)所示:

$$p(x) = \sum_{i=1}^i \pi_i N(x | u_i, \Sigma_i) \quad (6)$$

式中: u_i 为第 i 个分量的均值; Σ_i 为第 i 个分量的协方差矩阵; π_i 为第 i 个分量生成的概率。

货物信息的输入 x 可由第 i 个高斯混合近似分量表示,则目标值 y 便由第 i 个线性模型所表示, y 的概率分布如式(7)所示:

$$p(y|x, \theta) = N(y | w_i^T x, \beta_i^{-1}) \quad (7)$$

式中: w_i^T 为第 i 个变量的线性回归权重; β_i^{-1} 为零方差的高斯噪声; θ 为模型参数。

基于概率的聚类分析模型可由图 4 所示。变量的经验分布如式(8)所示:

$$p(z_{k,i} = 1) = \pi_i \quad (8)$$

式中: $z_{k,i}$ 为一组变量; π_i 在 $0 \sim 1$ 区间内。

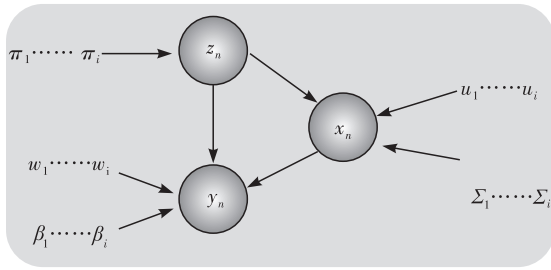


图 4 基于概率的聚类分析模型

x_n 和 y_n 为训练调控阶段的输入变量, z_n 为潜在变量, π, w, β, u 和 Σ 为聚类模型参数。

分析图 4, 将所有变量整合得到完整模型, 如公式(9):

$$p(Z, X, y | \theta) = p(Z | \theta) p(X | Z, \theta) p(y | Z, X, \theta) \quad (9)$$

聚类分析后的信息通过神经网络进行输出。神经网络是深度学习的主要模型, 通过模拟人脑神经网络结构, 经过一层层信息传递和处理后输出有效信息。一般为三层结构, 输入层中数据输入, 隐藏层中数据处理和输出层中信息输出, 最主要的是在隐藏层进行数据处理过程, 可由公式(10)所示:

$$y(t) = f\left(\sum_{i=1}^n x_i(t) v_i(t) + a_i(t)\right) \quad (10)$$

式中: $x_i(t)$ 为输入数据; $v_i(t)$ 为隐藏层系数; $a_i(t)$ 为调控误差。

神经网络在训练过程中通过实际值与预测值的差值来对隐藏层系数进行修正, 逐渐缩小预测与实际的误差。更新系数的过程如式(11)~式(13)所示:

$$E = \|y'(t) - y(t)\|^2 \quad (11)$$

$$v_i(t+1) = v_i(t) - \frac{\partial E}{\partial v_i(t)} \quad (12)$$

$$a_i(t+1) = a_i(t) - \frac{\partial E}{\partial a_i(t)} \quad (13)$$

式中: E 是反馈的误差值; $v_i(t+1)$ 为迭代后调整的系数; $a_i(t+1)$ 为迭代后调整的误差。

4 实际案例分析

研究数据由某可再生能源发电厂提供。该电厂在仓库布置高清摄像头, 共收集货物图像 3000 张, 构成货物检测的数据样本集。对图像进行分类, 可以分为小型箱体、中型箱体、大型箱体、中型桶体和大型桶体五种。数据样本按照 8:1:1 的比例进行划分成训练集、测试集、验证集。输入图像的初始分辨率为 640×640 , 迭代次数为 800, 初始学习率为

0.01, 优化器为随机梯度下降算法, NMS 阈值为 0.75。实验环境基于 Ubuntu 22.04 操作系统。实验数据的训练、测试和验证是在配备 Intel i5-1035G1 CPU 和 Nvidia Tesla T4 的 GPU 上完成, 并使用 CUDA 10.1 和 CUDNN 7.6.5 进行加速。

为了对提出的模型性能进行定量评估, 采用精准率 P , 召回率 R , 平均精度 AP 和 mAP 来对提出模型的性能进行量化, P 表示检测过程中将正类预测为正的样本数量占所有预测为正的样本数量的比例, R 表示检测过程中将正类预测为正的样本数量占所有真正的正类样本数量的比例, 如式(14)和式(15)所示:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (14)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (15)$$

式中: N_{TP} 为正类预测并实际为正的样本数量; N_{FP} 为假类预测但实际为正的样本数量; N_{FN} 为正类预测但实际为负的样本数量。

平均精度 AP 通过 A-R 曲线得到, 其计算如式(16)所示:

$$PA = \int_0^1 P dR \quad (16)$$

但是 AP 主要针对特定类别样本, 当检测过程中存在多种类别时, 通常采用 mAP 来进行定量表征, mAP 的计算如式(17)所示:

$$mPA = \frac{1}{N} \sum_{i=1}^N PA_i \quad (17)$$

为了直观展示模型识别货物的准确度, 绘制了混淆矩阵, 如图 5 所示。

小型箱体	134	6	1	1	0
中型箱体	0	33	0	1	0
大型箱体	0	2	57	0	1
中型桶体	0	2	0	40	0
大型桶体	0	0	1	0	21

小型箱体 中型箱体 中型桶体 中型桶体 大型桶体

图 5 混淆矩阵

从图 5 可以看出, 对于小型箱体, 正确识别出了 134 个, 仅有 8 个误检测; 对于中型箱体, 正确识别 33 个, 1 个误检测; 对于大型箱体, 正确识别 57 个, 误检测 2 个; 对于中型桶体, 正确识别 40 个, 误检测 2 个; 对于大型桶体, 正确识别 21 个, 误检测 1 个。

分析发现将小型箱体检测成中型箱体较多,这是因为两者形状和尺寸接近,容易发生误检测现象。而大型箱体和小型箱体有明显的尺寸差异,箱体和桶体具有明显的形状差异。通过标准-可形变卷积相结合可以对图像特征进行有效提取,保证了输出结果的准确度。

为了验证该文提出模型性能,进行了消融实验。分别验证了 SDC 对图像特征提取和改进损失函数在货物识别中的作用。为了进行直观展示,损失函数简写成 L。同时本研究采用无锚框算法 YOLOv8 作为基础框架。消融实验的结果如表 1 所示。

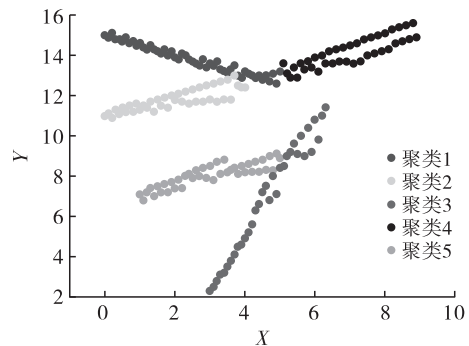
表 1 消融实验的结果

SDC	L	PA/%					mPA/%
		小型箱体	中型箱体	大型箱体	中型桶体	大型桶体	
		82.14	82.25	83.09	84.04	83.45	83.59
•		85.44	86.29	86.54	86.25	85.95	85.27
	•	83.65	83.88	84.34	84.21	84.02	84.21
•	•	90.18	90.51	91.27	91.53	90.86	90.94

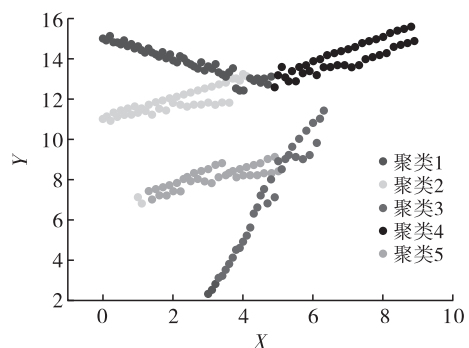
从表 1 可知,作为对比的无锚框算法 YOLOv8 的 mPA 值最低,为 83.59%,说明传统无锚框算法面对复杂图像的识别准确度不高。采用 SDC 对图像特征进行提取后,发现对各类特征的识别都得到提升,最终 mPA 值提升到了 85.27%。这主要是因为 SDC 具有对复杂图像特征的提取能力,这极大提升了对多尺寸复杂图像的识别。采用改进损失函数训练后来对不同类别图像进行识别时,相比较于无锚框算法,各类特征识别都得到提升,mPA 值提升至 84.21%,但精准度仅提升 0.62%,这表明损失函数对识别精度略有增加。将 SDC 加入主干、改进混合函数加入检查探头后,在货物 5 种不同类别检测中,PA 值分别提升了 8.04%、8.26%、8.18%、7.49% 和 7.41%,mPA 值提升了 7.35%。对不同尺寸、形状货物的识别准确度得到了大幅度提升,这表明本研究提出的改进方法可以有效提高对多尺度复杂条件图像的识别能力。

采用基于概率的聚类分析对处理后的图像信息进行分类,为了进行对比,将无锚框算法 YOLOv8 处理的图像信息作为数据库 1,该文提出改进算法处理的图像信息作为数据样本库 2。图 6 为对不同数据样本库聚类分析后的合成数据集。在分析过程中权重、偏差值和观测噪音被设置为随机值,以确保实验具有代表性。从图 6 可以看出,对于不同的数据样本库,基于概率的聚类算法都可以实现种类的准

确分类,将数据对应分为 5 种。对比图 6(a)和图 6(b),在不同聚类的交界处,很容易发生误判断。图 6(b)使用数据库 1 作为输入数据样本,在聚类 1、聚类 2 和聚类 4 具有相似性的货物识别过程中,容易发生误判情况。在聚类 3 个和聚类 5 中,也发生了误判。同时可以发现对底层结构的描述不够充分,会出现误判。而使用数据库 2 作为输入数据样本时,明显误判现象减少。3×3 的 CNN-SDC 耦合算法可以有效提取图像特征信息,改进损失函数可以有效增强算法对货物位置的确定,最终实现对多尺度复杂环境中货物的识别和分类。



(a) 数据库 1



(b) 数据库 2

图 6 不同数据样本库聚类分析后的合成数据集

为了验证该文提出模型的适用性,分别对货架层、地面层和地面货架混合层的货物进行测试。图 7 为识别各种环境下货物准确率随迭代次数的变化情况。分析图 7 可知,3 种环境下基本在 35 次迭代都能实现收敛,说明模型计算过程不会出现过拟合现象。对比 3 种环境,货架层和地面层的货物识别准确率分别为 97.1% 和 96.8%。两者的准确率值比较接近,表明模型对这 2 种环境可以实现稳定地识别和分类。而混合环境下,准确率有所下降,为 94.2%,当环境变得复杂时,会降低模型的分辨能力,但是该模型依然保持在一个较高的准确率。总的来说,对于单一环境下的货物,该文提出模型可实现 96% 以上的准确识别,而在处理复杂环境下货物

时,可实现 94% 以上的准确识别。

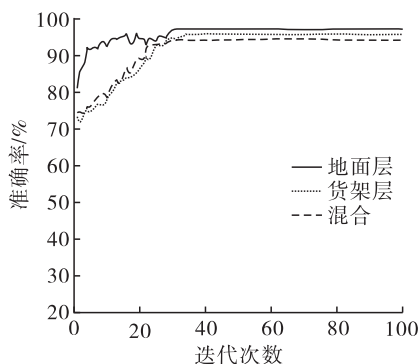


图7 识别各种环境下货物准确率随迭代次数的变化情况

5 结论

新能源发电装机容量逐渐增涨,其相应的仓储系统也日趋复杂化。传统的仓储系统依靠人工进行货物识别和分类耗费大量时间和人力,同时又存在准确度低的缺点。随着自动控制和深度学习的发展,基于机器视觉设备捕捉仓库货物图像信息,并结合深度学习算法进行特征提取的自动识别分类方法成为智能仓储的重点。而目前面对新能源发电厂仓储系统布局复杂,存储货物大小、种类差异较大的特点,如何实现准确识别和分类成为研究的难点。该文提出以 YOLOv8 为基本框架,在主干框架中添加 SDC 来增加对多尺度图像特性提取能力,并在检测头中添加改进损失函数来提高对位置的识别能力。最后采用聚类算法和神经网络相结合输出货物信息。进行消融实验来验证 SDC 和混合函数的作用。该文提出的改进模型在 5 种不同类别货物检测中, PA 值分别提升了 8.04% ,8.26% ,8.18% ,7.49% 和 7.41% , mPA 值提升了 7.35%。在图像特征采集后,基于概率的聚类分析展现出良好地分类效果,对于货架层、地面层和地面货架混合层的货物进行测试。

可知货架层和地面层的货物识别准确率分别为 97.1% 和 96.8%。而混合环境下,准确率下降至 94.2% ,环境越复杂,模型识别能力越低,但该文提

出的模型依然实现 94% 以上的准确率。

参考文献:

- [1] 李晖,刘栋,姚丹阳.面向碳达峰碳中和目标的我国电力系统发展研判[J].中国电机工程学报,2021,41(18):6245-6259.
- [2] 徐慧剑.基于物联网 RFID 技术的智能仓储系统的设计与实现[J].制造业自动化,2012,34(07):139-141.
- [3] 李秉津.基于物联网技术的电力应急物资智能仓储管理模式设计[J].网络安全和信息化,2024(09):71-73.
- [4] 林健光.基于 ZigBee 的智能仓储物流系统设计研究[J].装备维修技术,2023(1):38-43.
- [5] 廖仕军.基于机器视觉的工业机器人定位系统研究与应用[J].无线互联科技,2021,18(7):63-64.
- [6] 王成军,韦志文,严晨.基于机器视觉技术的分拣机器人研究综述[J].科学技术与工程,2022,22(3):893-902.
- [7] 徐青青.基于机器视觉的工业机器人智能分拣系统设计[J].仪表技术与传感器,2019(8):92-95.
- [8] 郑建彬,白雅贤,詹恩奇,等.基于改进 SIFT 匹配方法的货架乳制品识别[J].计算机科学,2017,44(9):315-319.
- [9] Yilmazer R, Birant D. Shelf auditing based on image classification using semi-supervised deep learning to increase on-shelf availability in grocery stores [J]. Sensors, 2021, 21(2): 327.
- [10] 李云娟,樊雪双.基于数学建模的 PSO-DE 算法在机器人智能拣货过程中的应用研究[J].自动化与仪器仪表,2024(03):197-200.
- [11] 吴海彬,林宜.基于改进 Voronoi 图的移动机器人在线路径规划[J].中国工程机械学报,2007(01):117-121.
- [12] 唐媛红,刘月云.基于 PLC 的水果分拣机器人动作控制优化[J].农机化研究,2021,43(12):233-237.
- [13] 黄雨晨,张利平,唐秋华,等.带 AGV 开放车间调度的析取图模型与求解[J].机械设计与制造,2023(6):10-15.