

基于多模态的无序堆叠快递包裹机械臂视觉抓取系统

吴蓬勃¹, 张金燕², 李莉¹, 孙会泽¹, 王帆³

(1.石家庄邮电职业技术学院 智能工程系, 石家庄 050021; 2.河北工程技术学院 网络与通信学院, 石家庄 050091; 3.石家庄邮电职业技术学院 速递物流系, 石家庄 050021)

摘要: **目的** 为实现快递交叉带分拣机的自动供包, 设计一款基于多模态的无序堆叠快递包裹机械臂视觉抓取系统。**方法** 将堆叠快递包裹 RGB 图像和深度图多模态信息送入全卷积残差网络模型进行识别, 经过模型后处理、抓取策略设计、模型训练与部署, 实现快递包裹抓取区域精确识别; 结合包裹姿态估计, 控制机械臂完成快递包裹的抓取和搬运。**结果** 实验证明, 文中提出的视觉识别方法实现了快递包裹抓取区域的精确识别, 测试准确率达到 95.73%。**结论** 基于该方法构建的快递包裹机械臂视觉抓取系统, 可以完成快递交叉带分拣机的自动供包操作。

关键词: 多模态; 堆叠; 快递包裹; 机械臂视觉抓取

中图分类号: TP391.0 **文献标识码:** A **文章编号:** 1001-3563(2022)15-0068-09

DOI: 10.19554/j.cnki.1001-3563.2022.15.008

Robot Arm Visual Grasping System of Disorderly Stacking Express Package Based on Multimode

WU Peng-bo¹, ZHANG Jin-yan², LI Li¹, SUN Hui-ze¹, WANG Fan³

(1. Department of Intelligent Engineering, Shijiazhuang Post & Communication Technology College, Shijiazhuang 050021, China; 2. School of Network and Communication, Hebei Polytechnic Institute, Shijiazhuang 050091, China; 3. Department of Express and Logistics, Shijiazhuang Post & Communication Technology College, Shijiazhuang 050021, China)

ABSTRACT: The work aims to design a robot arm visual grasping system of disorderly stacked express package based on multimode to realize the automatic package feeding of express cross belt sorter. The multi-modal information of RGB image and depth map of stacked express packages was sent to the residual network model of the full convolution for recognition. After model post-processing, grasping strategy design, model training and deployment, the accurate identification of the grasping area of express packages was realized. Combined with the package attitude estimation, the grasping and handling of express packages were completed through the robot arm. Experiments showed that the visual recognition method proposed in this paper realized the accurate recognition of the grab area of express package, and the test accuracy was 95.73%. The robot arm visual grasping system of express parcel based on this method can complete the automatic package feeding operation of express cross belt sorter.

KEY WORDS: multimodal; stacking; express package; robot arm visual grasping

收稿日期: 2022-04-02

基金项目: 河北省大中学生科技创新能力培育专项 (22E50226D); 河北省高等学校科学技术研究项目 (ZC2021252)

作者简介: 吴蓬勃 (1980—), 男, 硕士, 石家庄邮电职业技术学院教授, 主要研究方向为计算机视觉技术、机器人技术。

通信作者: 张金燕 (1979—), 女, 本科, 高级工程师, 主要研究方向为物联网技术和人工智能技术。

近年来,随着电商和互联网技术的飞速发展,快递包裹的业务量与日俱增,这对快递分拣效率提出了更高的要求^[1-2]。交叉带分拣机作为快递分拣中心的核心设备,其分拣效率直接影响快递的时限。目前,交叉带分拣机的供包操作主要依靠人工完成,即人工将快递包裹从传输带上逐个取出,扫描条码后,再逐个放置到交叉带分拣机的入口,这极大地阻碍了分拣效率的提高。

随着深度学习技术的快速发展^[3-9]和机械臂智能化水平的不断提高^[10-11],基于视觉的机械臂抓取技术逐渐成为研究热点之一。王栋^[12]基于深度学习目标检测技术,使用双目相机识别物体并获取其坐标,通过双指夹具抓取传输带上物体。秦志民等^[13]采用 YOLOv3 进行目标分类检测,采用 DarkNet-53 五参数法进行目标位姿的检测,取得了较好的效果。随着传感器技术的发展,基于多种传感器技术融合的多模态识别方法优势凸显^[14-16]。因为 RGBD 相机可同时采集物体的 RGB 图像和深度图像^[17],可获取物体的三维位置信息^[18-19],所以 RGBD 相机逐渐被应用于机械臂视觉抓取领域。针对物体 6D 姿态识别问题,Zeng 等^[20]提出了基于多视角自监督深度学习的 6D 姿态估计模型。该模型需要采集场景中每个物体 15~18 个视角的图像,以获取物体 6D 姿态,使得基于该模型的机械臂视觉抓取效率较低。Song 等^[21]提出了一种利用关键点、边缘向量和对称对应关系的 6D 姿态估计方法。Qi 等^[22]采用 2D RGB 图像与 3D 点云相结合的方式,通过投票方式缩小搜索空间,提高了 3D 目标检测的精度。Zeng 等^[23]针对杂乱场景下未知物体的抓取问题,提出了“先抓取再识别”^[23]的

方法,将抓取后的物体进行识别,大大降低了识别的难度;将 RGB 图像和深度图多模态信息送入全卷积残差网络,输出抓取区域标签图,同时将深度图转换为 3D 点云,计算抓取面法向量,输出抓取角度信息。

针对交叉带分拣机的自动供包问题,文中综合运用 RGB 图像和深度图多模态信息,通过全卷积残差网络模型^[23]提取图像特征,结合模型后处理、抓取策略设计实现快递包裹抓取区域精确识别;结合包裹姿态估计,控制机械臂完成快递包裹的抓取和搬运。

1 系统设计

系统以快递交叉带分拣机供件台为适配对象,将位于送料传输带和等待上车传输带之间的人工供件装置,替换为机械臂视觉抓取装置,系统组成见图 1^[24]。

绿色拣货区左端的光电探测器可以检测快递包裹的进入。机械臂视觉抓取系统将快递包裹逐个抓取、扫码,并搬运到等待上车传输带上,为快递包裹逐个进入交叉带分拣机小车做好准备。

机械臂视觉抓取系统以 NVIDIA JETSON AGX XAVIER (以下简称 AGX) 平台为控制与图像处理单元;通过顶部的 RGBD 三维相机识别拣货区中的快递包裹,获取包裹表面中心位置的三维坐标和包裹姿态;通过变换矩阵将三维坐标转换为机械臂坐标,控制机械臂运转到相应位置抓取包裹,将包裹搬运到条码扫描区,完成条码扫描;调整机械臂姿态,将包裹面单向上摆放到等待上车传输带上。

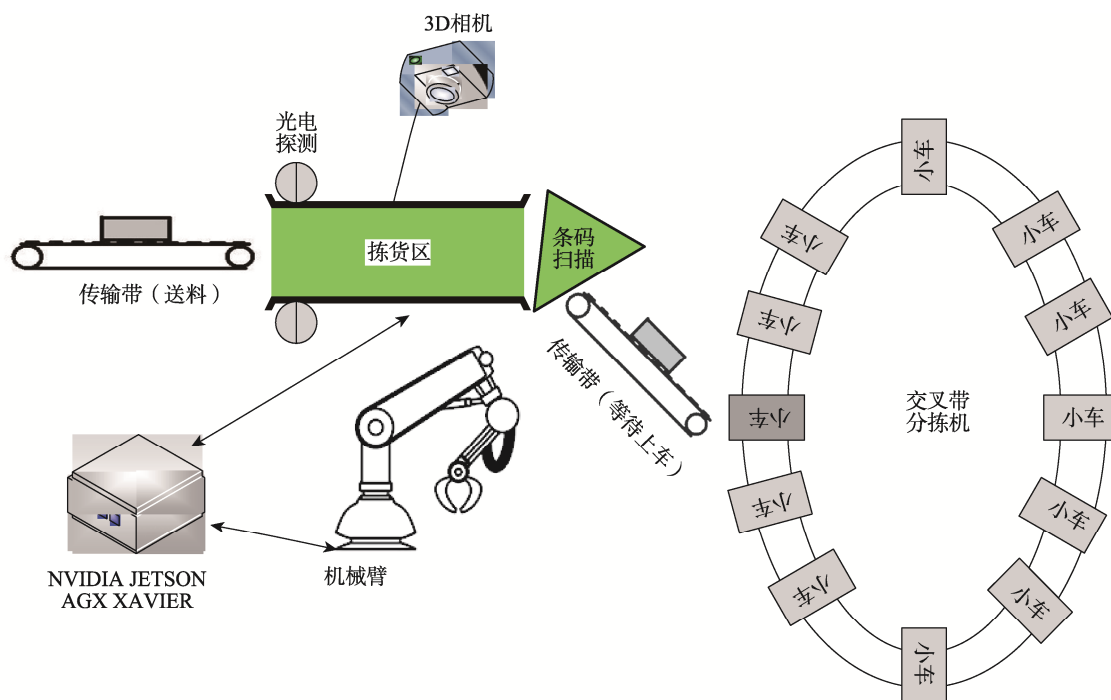


图 1 系统组成简图

Fig.1 Diagram of system components

2 快递包裹抓取策略

快递包裹的抓取策略见图2。首先进行图像数据采集,然后通过图3的模型进行抓取区域识别,生成热点区域(可抓取区域);通过轮廓提取,判断当前区域是否存在热点区域;如果没有热点区域,则控制送料传输带送入包裹;如果有热点区域,则将热点区

域按照 RGBD 相机到热点区域中心的深度值(h)和面积进行排序,优先抓取 h 更小(包裹与 RGBD 相机距离最近,在堆积包裹的上侧)、面积更大(更加适宜抓取的包裹)的快递包裹。最后,将热点区域坐标(x_0, y_0)和深度值 h 转化为机械臂坐标(x, y, z);对待抓取区域包裹姿态进行估计,得出姿态角为(α, β);将三维坐标和姿态数据(x, y, z, α, β)送入机械臂控制程序,完成包裹抓取。

3 无序堆叠包裹抓取区域识别模型设计

3.1 模型设计

无序堆叠快递包裹抓取区域识别模型见图3^[23]。在进行分拣业务前,通过 RGBD 相机采集拣货区背景的 RGB 图和深度图用于模型后端的背景滤除。在图3的无序堆叠快递包裹多模态识别模型中,分别将拣货区包裹的 RGBD 图像和深度图像送入模型。模型首先使用微调的残差网络模型 ResNet 分别提取2个图像特征;然后,经过3次空间卷积、上采样等操作后输出结果图。与普通的检测检测有所不同,该模型的输出结果为一幅灰度图,白色区域为可抓取区域,黑色为背景区域。

这幅灰度图中存在背景干扰,首先需要通过后面的 BackgroundFilter 结合背景 RGB 图和深度图进行背景滤除。然后,根据像素值生成相应的热图。在热图中,通过不同的像素值区间来标识可抓取区域、不可抓取区域和背景区域。通过这种方式,实现了标签类别的泛化,即不去判断目标物体为哪种类别的包裹,而是去判断哪些区域为可抓取区域。最后,通过可抓取区域的轮廓提取方法 FindContours、深度数据相似度判定方法 DepthSimilar 二次滤除背景,得到最终的待抓取区域图。

图2 快递包裹抓取策略
Fig.2 Express package capture strategy

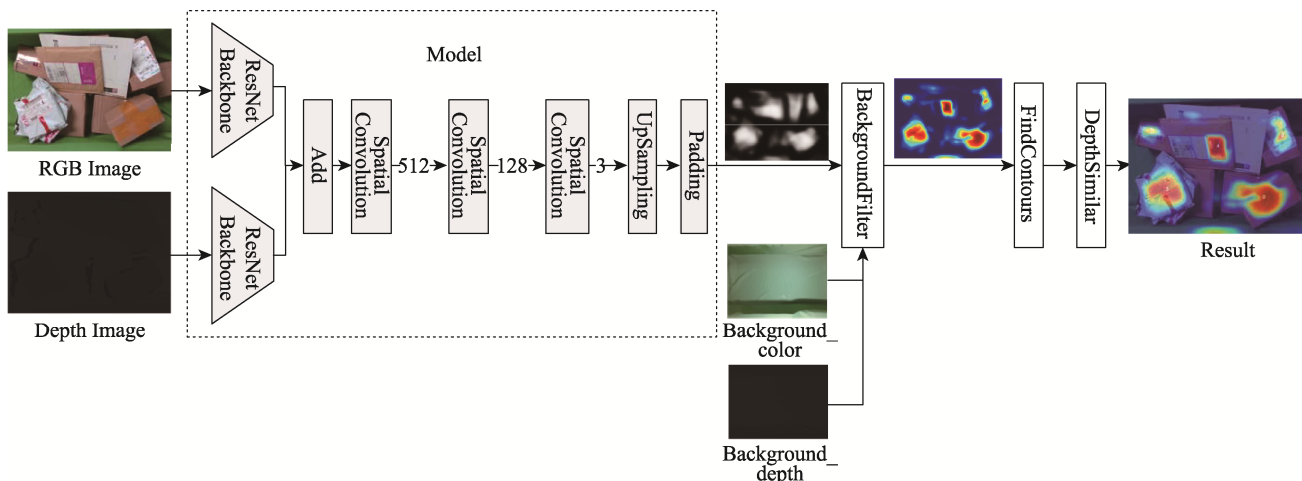


图3 无序堆叠快递包裹多模态识别模型
Fig.3 Multimodal identification model of disorderly stacked express packages

3.2 模型后处理

原始图像经过图 3 中的模型处理、背景过滤后, 仍然会存在一些背景干扰, 例如: 光照变化、阴影、背景污渍等都会导致背景被判定为可抓取区域, 因此, 文中引入了深度数据相似度判定, 以滤除这些干扰。

深度数据相似度判定在轮廓提取后进行, 主要判断当前轮廓区域深度数据与背景深度数据是否相似, 如果相似, 则判断为背景, 否则认定为可抓取区域。具体流程见图 4。

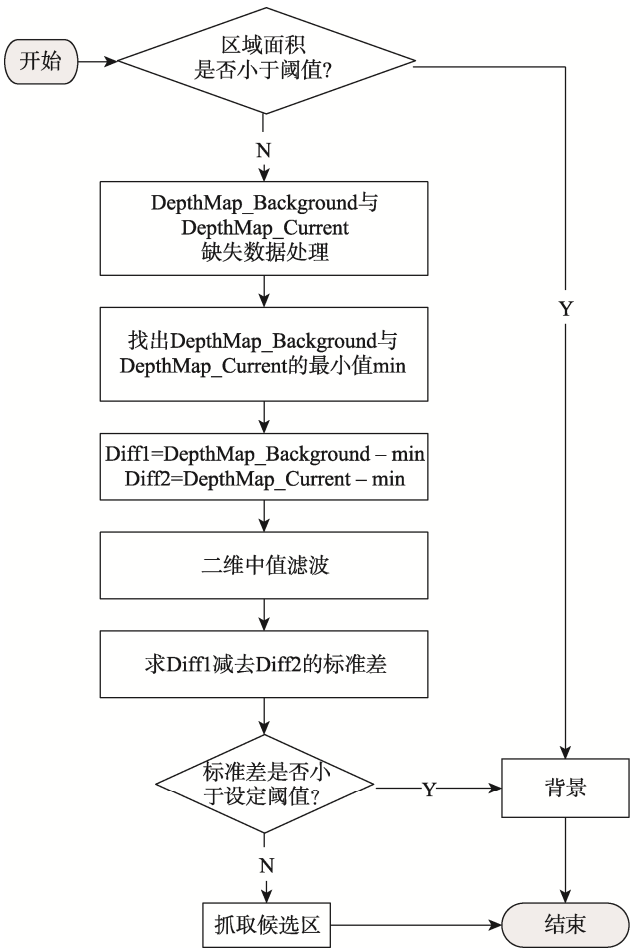


图 4 深度数据相似性判断流程
Fig.4 Judgment process of depth data similarity

首先, 如果轮廓区域面积过小, 则不适宜抓取, 直接忽略掉。由于 RGBD 相机采集的深度数据存在一定的噪声, 所以需要背景深度图 (DepthMap_Background) 和当前深度图 (DepthMap_Current) 中的缺失数据进行处理, 对零值数据不做比较。然后, 求出 DepthMap_Background 和 DepthMap_Current 中的最小非零值 (min), 并减去 min, 实现数据的归一化; 对两者的差值 Diff1 和 Diff2 分别进行中值滤波, 消除孤立噪声点; 通过求 Diff1 减去 Diff2 的标准差, 获取 Diff1 和 Diff2 的偏离程度; 如果标准差小于设

定阈值, 则判断该区域为不可抓取的背景区域, 否则, 判定为抓取候选区域。表 1 列举了在强光、阴影等干扰下, 模型识别的结果和加入深度数据相似度处理后的结果。可以看出, 在强光或者阴影干扰下, 在模型识别的热图中, 背景区域出现了若干可抓取的深色区域, 通过深度数据相似度处理后, 这些背景的抓取区域被很好地过滤掉了, 最终在 RGB 图上通过方框显示出可抓取的区域。

表 1 加入深度数据相似度判定后的识别效果
Tab.1 Recognition effect after adding depth data similarity judgment

干扰源	RGB 图	RGB+模型识别结果热图	深度数据相似性处理后的图像
强光 阴影			
强光			
强光			
强光			

3.3 数据采集与标注

Zeng 等^[23]的数据集中包含 1 837 张经过标注的图像数据, 主要以料框中的各形态商品为主。文中系统对 Zeng 等^[23]数据集进行了增扩, 重新采集了 17 个种类 (7 种纸箱、3 种气泡扁平件、3 种纸质扁平件、4 种软包装袋) 的快递包裹图像, 见图 5。通过快递包裹不同数量、种类、姿态的组合, 共计采集 800 张图像 (见图 6), 并设计了图像标注软件进行图像数据标注 (见图 7)。

与传统目标检测的标注方法有所不同, 文中模型的标注结果不是坐标位置, 而是一幅灰度图, 通过 3 种颜色 (黑、灰、白) 来标识 3 种标签 (不可抓取区域、可抓取区域、背景区域)。其中, 不可抓取区域一般为包裹边沿或者不适合抓取的包裹, 例如大部分埋在包裹堆下面的包裹。

3.4 模型训练

将 2 637 张经过标注的数据和对应的标签按照 8 : 2 的比例划分成训练集和测试集, 送入模型进行

训练,在配有4块英伟达 RTX TITAN 24 G 的 GPU 服务器上训练。经过3 001个批次训练后,模型训练的损失率降到了13%(见图8),将测试集的527

张图片在训练好的模型上进行测试,平均准确率达到95.73%(见图9,深色标识为判断正确,白色标识为判断错误)。



图5 不同种类的快递包裹
Fig.5 Different kinds of express packages

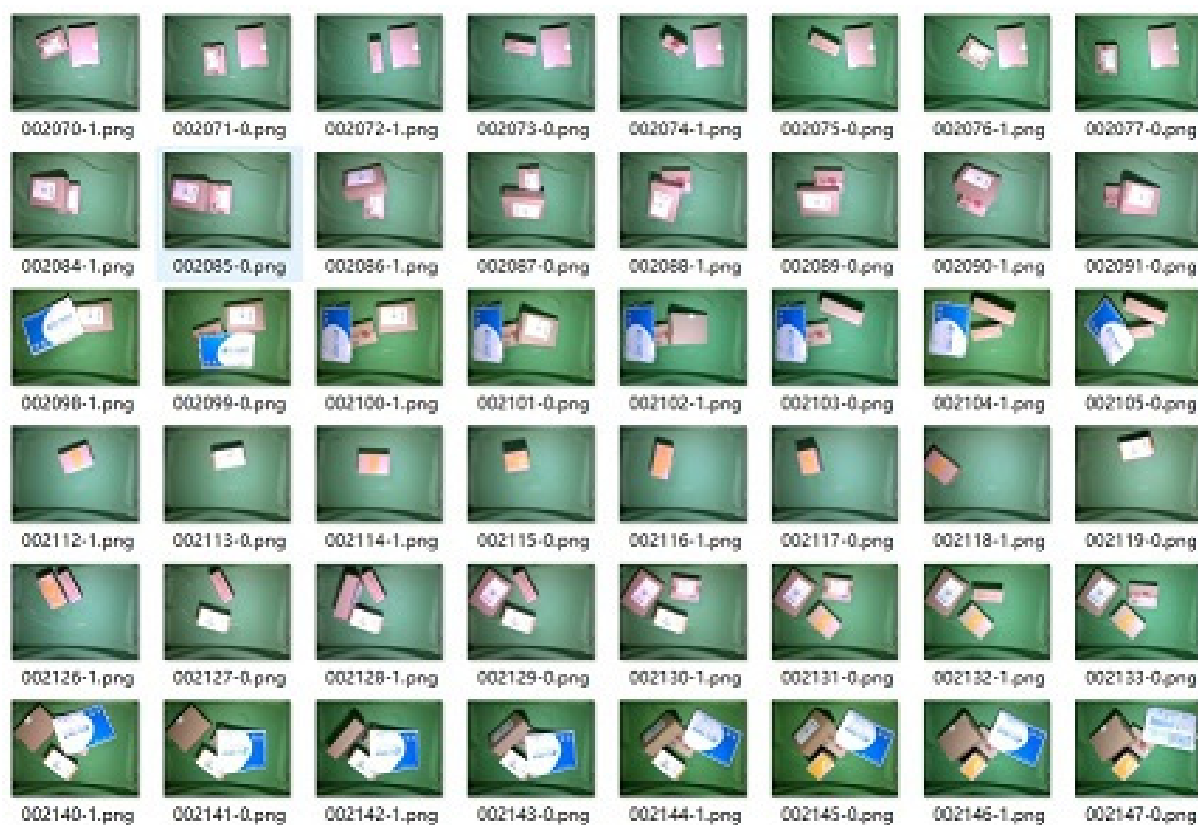


图6 采集的图像
Fig.6 Collected images



图 7 标注后的图像
Fig.7 Annotated images

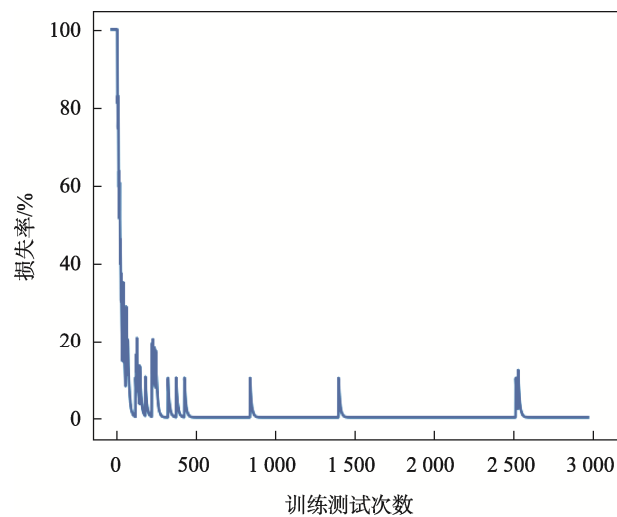


图 8 模型损失率曲线
Fig.8 Model loss rate curve

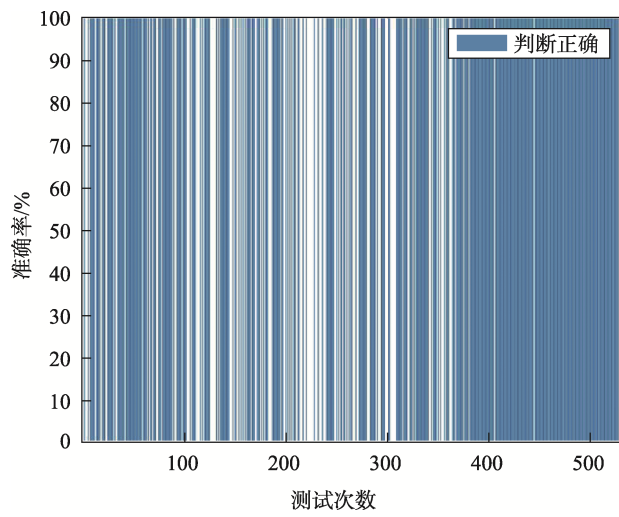
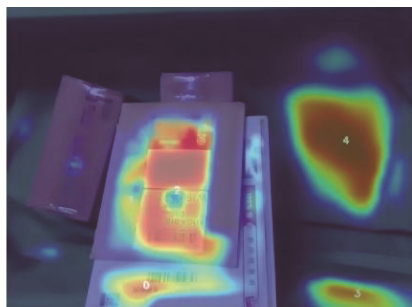


图 9 模型测试柱状图
Fig.9 Model test histogram

4 包裹姿态估计

由于快递包裹是无序堆叠到一起的,为实现机械臂高效准确的抓取,系统采用 3D 点云方式获取包裹抓取区域的 3D 姿态信息。首先通过 RGB 和深度图像数据生成 3D 点云;然后根据包裹抓取区域热点图

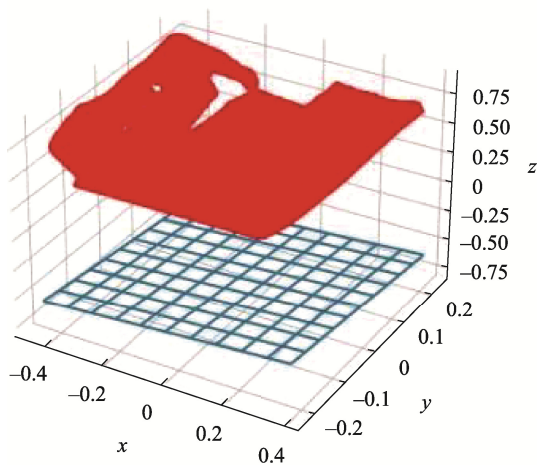
(图 10a)对点云进行裁剪,仅保留待抓取区域的 3D 点云(图 10b)。由于快递包裹表面一般为平面或者有较少的凹凸面,同时考虑到机械臂抓取吸盘带有缓冲装置,所以,将快递包裹表面按平面进行处理。将待抓取区域的 3D 点云进行主成分分析,获取最小二乘拟合平面(图 10c),根据平面方程计算出待抓取区域的姿态角。



a 包裹抓取区域热点图



b 点云裁剪图



c 点云估计平面图

图 10 包裹姿态估计

Fig.10 Package pose estimation

5 系统测试

将训练好的快递包裹抓取区域识别模型和姿态估计算法部署到 AGX 硬件平台上, 通过 RealsenseD435i 摄像头以 640×480 分辨率采集 RGB 图像和深度图像, 并送入模型进行识别, 模型的识别速率达到了每秒 1.53 帧。由于图像识别与机械臂搬运可以并行进行, 而机械臂的搬运时间远大于 1 s, 所以模型的识别速率可以满足系统实时性的要求。

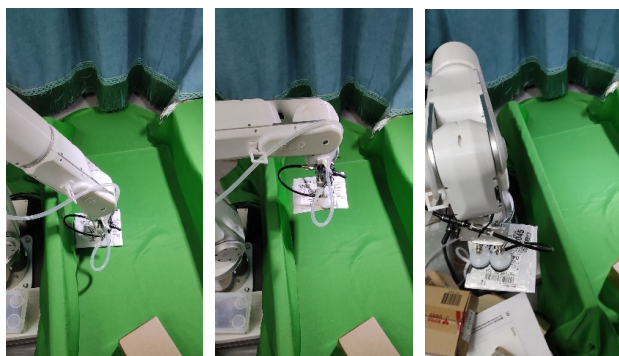
通过机械臂和 RGBD 相机手眼标定, 将抓取区域三维坐标转换为机械臂运动坐标, 结合包裹姿态角, 控制机械臂完成包裹抓取动作。快递包裹抓取与搬运的实际测试效果见图 11。



a 包裹箱



b 纸质扁平件



c 气泡扁平件

图 11 快递包裹的抓取与搬运

Fig.11 Grabbing and handling of express package

6 结语

面向交叉带分拣机供件台, 设计了一款基于多模态的无序堆叠快递包裹识别和抓取系统。基于 RGBD 相机采集的 RGB 图像和深度图像, 在 ResNet 模型的基础上, 进行网络微调、多级卷积和采样获取抓取区域图; 通过后端的背景过滤、深度数据相似性判别等处理, 实现了堆叠快递包裹抓取区域的识别; 通过 3D 点云获取包裹姿态; 通过数据集增扩、标注和模型训练部署, 在 AGX 嵌入式平台上实现了每秒 1.53 帧的识别速度; 通过坐标转换, 控制机械臂实现了快递包裹的高效抓取。

参考文献:

- [1] 白文杰. 基于快递面单深度学习的物流分拣规划系统研究[D]. 淮南: 安徽理工大学, 2021: 1-6.
BAI Wen-jie. Research on Logistics Sorting Planning System Based on Express Single Target Detection[D]. Huainan: Anhui University of Science & Technology, 2021: 1-6.
- [2] 纪浩林. 交叉带式物流快递自动分拣系统设计与实现[D]. 阜新: 辽宁工程技术大学, 2016: 1-2.
JI Hao-lin. Implementation and Design of Logistics and Express Cross-Belt Automatic Sorting System[D]. Fuxin: Liaoning Technical University, 2016: 1-2.
- [3] ZHANG Li-ping, LI Wei-jun, SUN Lin-jun, et al. GmFace: An Explicit Function for Face Image Representation[J]. Displays, 2021, 68(1): 102022.
- [4] NING X, GONG K, LI W, et al. Feature Refinement and Filter Network for Person Re-Identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(9): 3391-3402.
- [5] 吴蓬勃, 张金燕, 王帆, 等. 快递暴力分拣行为视觉识别系统[J]. 包装工程, 2021, 42(15): 245-252.
WU Peng-bo, ZHANG Jin-yan, WANG Fan, et al. Visual Recognition System of Violent Sorting Behavior in Express Delivery[J]. Packaging Engineering, 2021, 42(15): 245-252.
- [6] ZHANG Jian-ming, XIE Zhi-peng, SUN Juan, et al. A Cascaded R-CNN with Multiscale Attention and Imbalanced Samples for Traffic Sign Detection[J]. IEEE Access, 2020, 8: 29742-29754.
- [7] WANG C, BAI X, WANG X, et al. Self-Supervised Multiscale Adversarial Regression Network for Stereo Disparity Estimation[J]. IEEE Transactions on Cybernetics, 2021, 51(10): 4770-4783.
- [8] WANG Chen, WANG Xiang, BAI Xiao, et al. Self-Supervised Deep Homography Estimation with Invertibility Constraints[J]. Pattern Recognition Letters, 2019, 128(1): 355-360.
- [9] ZHOU Lei, BAI Xiao, LIU Xiang-long, et al. Learning Binary Code for Fast Nearest Subspace Search[J]. Pattern Recognition, 2020, 98(1): 107040.
- [10] 房国栋, 高军伟, 朱晨曦, 等. 基于机器视觉的机械臂智能分拣系统[J]. 仪表技术与传感器, 2020(12): 72-76.
FANG Guo-dong, GAO Jun-wei, ZHU Chen-xi, et al. Intelligent Sorting System for Manipulator Based on Machine Vision[J]. Instrument Technique and Sensor, 2020(12): 72-76.
- [11] 李恩宇. 基于单目视觉的六自由度机械臂分拣系统设计[D]. 抚顺: 辽宁石油化工大学, 2019: 1-4.
LI En-yu. Design of a Six-Degree-of-Freedom Mechanical Arm Sorting System Based on Monocular Vision[D]. Fushun: Liaoning Petrochemical University, 2019: 1-4.
- [12] 王栋. 基于视觉的机械臂目标跟踪与抓取方法研究[D]. 青岛: 青岛科技大学, 2021: 1-7.
WANG Dong. Research on Object Tracking and Grasping System of Robot Arm Based on Vision[D]. Qingdao: Qingdao University of Science & Technology, 2021: 1-7.
- [13] 秦志民. 视觉机械臂目标识别与抓取技术研究及系统开发[D]. 北京: 北京印刷学院, 2021: 1-7.
QIN Zhi-min. Research on Target Recognition and Grasping Technology and System Development of Vision Machine Arm[D]. Beijing: Beijing Institute of Graphic Communication, 2021: 1-7.
- [14] 甄军平, 张平. 旅客多模态识别在机场安检中的应用[J]. 现代计算机, 2021(13): 60-65.
ZHEN Jun-ping, ZHANG Ping. Application of Passenger Multimodal Identification in Airport Security Inspection[J]. Modern Computer, 2021(13): 60-65.
- [15] 吕葛梁. 手指多模态识别技术的研究[D]. 杭州: 杭州电子科技大学, 2020: 1-6.
LYU Ge-liang. Research on Finger Multimodal Identification Technology[D]. Hangzhou: Hangzhou Dianzi University, 2020: 1-6.
- [16] 刘小雨. 基于多模态深度学习的驾驶行为识别方法研究[D]. 南京: 南京邮电大学, 2021: 1-6.
LIU Xiao-yu. Research on Driving Behavior Recognition Method Based on Multi-Modal Deep Learning[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2021: 1-6.
- [17] WANG Yang-fan, WANG Chen, LONG Peng, et al. Recent Advances in 3D Object Detection Based on RGB-D: A Survey[J]. Displays, 2021, 70: 102077.
- [18] RAHMAN M, TAN Y, XUE J, et al. Notice of Violation of IEEE Publication Principles: Recent Advances in 3D Object Detection in the Era of Deep Neural Networks: A Survey[J]. IEEE Transactions on Image Processing, 2020, 29(1): 2947-2962.
- [19] ARNOLD E, AL-JARRAH O, DIANATI M. A Survey on 3d Object Detection Methods for Autonomous Driving Applications[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(10): 3782-3795.
- [20] ZENG A, YU K T, SONG S, et al. Multi-View Self-Supervised Deep Learning for 6D Pose Estimation

- in the Amazon Picking Challenge[C]// 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore: 2017: 1386-1383.
- [21] SONG C, SONG J, HUANG Q. Hybridpose: 6d Object Pose Estimation under Hybrid Representations[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Online, 2020: 428-437.
- [22] QI C R, CHEN XINLEI, LITANY O, et al. ImVoteNet: Boosting 3D Object Detection in Point Clouds with Image Votes[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Online, 2020: 4403-4412.
- [23] ZENG A, SONG Shu-ran, YU Kuan-ting, et al. Robotic Pick-and-Place of Novel Objects in Clutter with Multi-Affordance Grasping and Cross-Domain Image Matching[C]// 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, 2017: 3750-3757.
- [24] 高国亮, 巩雪. 基于智能交叉带分拣系统的机械臂设计[J]. 哈尔滨商业大学学报(自然科学版), 2021, 37(2): 195-199.
- GAO Guo-liang, GONG Xue. Design of Manipulator Based on Intelligent Cross Belt Sorting System[J]. Journal of Harbin University of Commerce (Natural Sciences Edition), 2021, 37(2): 195-199.

责任编辑: 曾钰婵