

本文引用格式：熊治国,周恒旭,冯煜升.基于 CNN 的机器人环境点云分类研究[J].自动化与信息工程,2023,44(2):16-21.

XIONG Zhiguo, ZHOU Hengxu, FENG Yusheng. Research on point cloud classification of robot environment based on CNN[J]. Automation & Information Engineering, 2023,44(2):16-21.

基于 CNN 的机器人环境点云分类研究*

熊治国 周恒旭 冯煜升

(北京理工大学珠海学院航空学院, 广东 珠海 519000)

摘要: 为解决目前三维点云识别算法存在训练模型计算量较大、算法网络结构复杂的问题,进行基于卷积神经网络(CNN)的机器人环境点云分类研究。首先,在机器人仿真环境中搭建家庭相关物品模拟服务场景,并使用模拟三维激光雷达生成环境点云;然后,将环境点云聚类分割出单个物体的点云簇;接着,利用二维投影法将三维点云数据转化为二维图像信息;最后,利用数据增强技术扩展数据集,并结合改进的 LeNet-5 卷积神经网络训练识别模型,得到相较于经典的 LeNet-5 模型更高的精度。实验结果表明,将环境点云先分割成单个物体,再进行分类识别是可行的,并具有一定的应用价值。

关键词: 机器人环境理解;激光点云;数据增强;卷积神经网络;物体分类

中图分类号: TP242

文献标志码: A

文章编号: 1674-2605(2023)02-0004-06

DOI: 10.3969/j.issn.1674-2605.2023.02.004

Research on Point Cloud Classification of Robot Environment Based on CNN

XIONG Zhiguo ZHOU Hengxu FENG Yusheng

(School of Aviation, Beijing Institute of Technology, Zhuhai 519000, China)

Abstract: To solve the problems of large training model computation and complex algorithm network structure in current 3D point cloud recognition algorithms, a research on point cloud classification in robot environments based on convolutional neural networks (CNN) is conducted. Firstly, build a simulation service scenario for household related items in a robot simulation environment, and use a simulated 3D LiDAR to generate environmental point clouds; Then, the environmental point cloud is clustered and segmented into point cloud clusters of individual objects; Next, the 2D projection method is used to convert the 3D point cloud data into 2D image information; Finally, by utilizing data augmentation technology to expand the dataset and combining it with the improved LeNet-5 convolutional neural network to train the recognition model, higher accuracy was obtained compared to the classic LeNet-5 model. The experimental results indicate that it is feasible to segment the environmental point cloud into a single object before recognition, and it has certain application value.

Keywords: robot environment understanding; laser point cloud; data enhancement; convolutional neural network; object classification

0 引言

复合机器人对周围环境的理解是实现智能化的基础,也是实现其他高级功能的前提条件。智能复合机器人需要传感器加持来实现具体功能,如 PR2 机器人利用二维激光雷达(light detection and ranging, LiDAR)实现即时定位与地图构建(simultaneous localization and mapping, SLAM);利用深度摄像头

实现物体检测^[1]等。基于 LiDAR 的机器人成像系统广泛应用于视觉测量、检测、分类识别、引导和自动化装配等领域^[2]。在利用 LiDAR 进行机械臂避障的过程中^[3],三维环境信息经地面分割、体素网格滤波、点云语义分割和包围盒提取后,可得到完整的避障所需信息,同时分割后的点云簇具有分类识别物体的特征信息。

16 * 基金项目:北京理工大学珠海学院校级科研发展基金(XK-2018-10)

近些年,利用深度学习进行点云处理和物体识别成为研究热点^[4-6],主要分为基于点云图的算法和基于鸟瞰图的算法2大类。其中,基于点云图的算法结合卷积神经网络(convolutional neural network, CNN)提出一种在三维空间中与点云顺序无关的算子^[7]。文献[8]将CNN用于三维体素化后的点云匹配,证明深度学习具有较强的点云特征提取能力。文献[9]利用深度学习进行3D点云的目标检测,但对远距离稀疏点云的检测效果不佳。文献[10]利用图卷积网络对点云数据进行分类,在标准数据集上取得良好的效果,但模型较为复杂,需要大量调试网络参数。文献[11]提出一种融合CNN和多重参考系模型(multiple reference frame, MRF)的点云分割方法,操作复杂。基于鸟瞰图的算法先将点云数据转换为鸟瞰图,再利用RCNN、Fast RCNN、Faster RCNN、YOLO和SSD等算法^[12-18]进行目标检测。

综上所述,一方面,使用大型深度学习网络直接处理点云或点云鸟瞰图时,对数据集和算力的要求较高,部署阶段需压缩模型。考虑到二维图像的深度学习算法更加成熟稳定,较小的网络结构也能实现良好的识别效果,且家庭服务场景中的物体种类有限,故先将单个物体的三维点云簇通过不同角度投影得到多角度图像,再进行深度学习构建物体分类识别模型是一种可行的策略。相较于输入完整的点云,点云投影图像可灵活地进行分类处理,通用性强。另一方面,因为二维图像不含景深信息,若仅使用二维图像需要大量训练才能实现语义理解。若利用三维雷达获取点云数据,经过处理形成单个物体的点云簇^[3],可避免从单一视角理解复杂景深环境信息的特点,解决不同景深下的物体相互遮挡造成的语义理解难题。

本文针对复合机器人对家庭服务场景存在的理解难题,首先,基于前期工作^[3]将仿真环境中家庭服务场景的点云处理成单个物体的点云簇,并通过投影法将物体的三维图像转化为二维图像;然后,利用数据增强技术扩充图像数据形成数据集;最后,利用数据集对卷积神经网络模型进行训练,并通过不断调整

网络参数来提高准确率,从而达到准确分类识别物体的效果。

1 场景搭建和物体点云获取

本文通过CoppeliaSim机器人仿真软件进行智能复合机器人的家庭服务场景搭建。家庭服务场景为常见的室内非结构场景,主要包括衣柜、沙发、椅子、盆栽、桌子、人等6类,如图1所示。

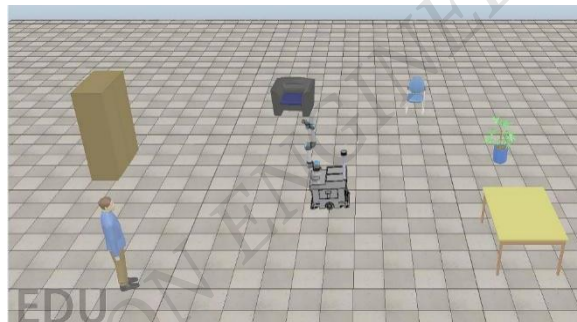


图1 仿真环境中的家庭服务场景

为获取家庭服务场景的激光点云数据,在仿真环境中采用模拟的VLP-16三维激光雷达传感器插件进行仿真,利用CoppeliaSim的远程API与Python通信,获取复合机器人三维激光雷达的点云单帧数据,并保存为.ply格式文件,经MATLAB可视化后如图2所示。

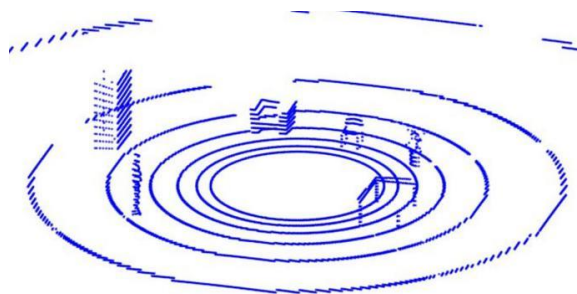


图2 家庭服务场景的点云图

2 数据集制作

利用文献[3]的方法,在包含各类物体的环境点云基础上,首先,利用基于横截面积的算法对环境点云数据进行地面分割,将大面积的地面点云分离,避免聚类时出现干扰;然后,采用体素网格滤波算法对环境点云进行压缩,以减少点云数据量,便于后续算法

聚类；最后，采用改进的基于密度的噪声应用空间聚类（density-based spatial clustering of applications with noise, DBSCAN）算法对分割压缩后的环境点云进行单个物体点云的聚类分割，将家庭服务场景的单帧点云数据聚类分割成6个点云簇，即6个不同类别物体，其中人的点云簇如图3所示。



图3 人的点云簇

通过不同角度将6类物体的三维点云簇投影，每类物体获得80幅二维投影图像。由每类物体80幅图像形成的数据集数据数量较少，训练的识别模型鲁棒性不强。本文通过缩小放大、旋转(45°、90°、180°、270°)、像素平移、翻转(水平、垂直)、明亮度改变和添加噪声(椒盐、高斯)等方法对图像进行数据增强，使1幅图像衍生出132幅图像，6类物体的数据集共有63360幅图像。数据增强后的人点云投影图集如图4所示。

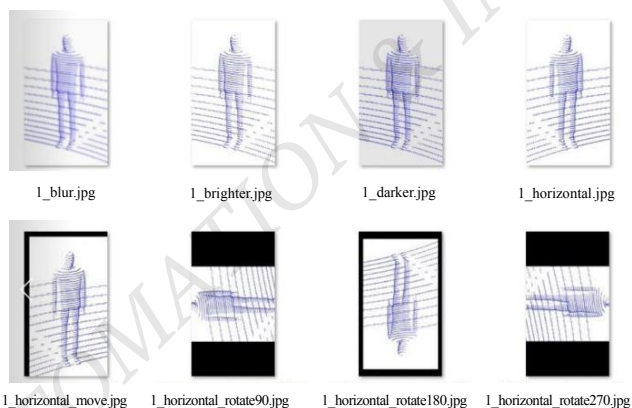


图4 数据增强后的人点云投影图集

3 卷积神经网络

物体识别问题可视为分类问题。通过机器学习方法将环境点云分割成单个物体点云，故在物体点云投影图分类后无需再锚框。通过先验知识和前期的实验

积累，在经典的LeNet-5卷积神经网络模型的基础上设计图像分类模型，并根据数据集和图像大小、图像中物体的分辨率和外观属性，调整LeNet-5卷积神经网络的池化方式、卷积核大小，增大全连接层的神经元个数和层数，如图5所示。

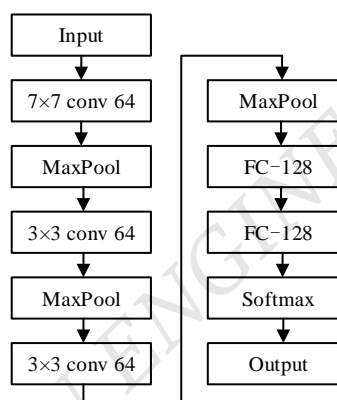


图5 改进的LeNet-5卷积神经网络

改进的LeNet-5卷积神经网络包含3个卷积层、3个池化层、2个全连接层和1个Softmax层。其中，Softmax层用于分类预测。

4 实验

4.1 算法实施过程

物体分类识别算法流程如图6所示。

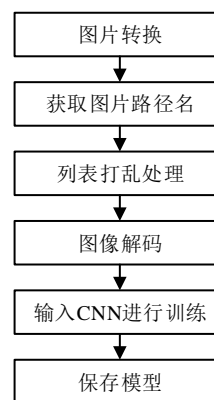


图6 物体分类识别算法流程

首先，将数据集中的图像转换成64×64像素的小图像，并获取所有图像的路径名，存放到对应的图像路径列表中，同时将有图像的标签存放到标签列表中；然后，对图像路径列表和标签列表打乱处理，并

将乱序的列表分为 2 部分，一部分用来训练，另一部分用来测试；接着，将图像进行解码操作，生成图像 batch（批），对训练变量声明；最后，将图像 batch 输入到改进的 LeNet-5 卷积神经网络进行训练，保存训练后的模型。

4.2 评价指标

本文选择的模型评价指标为误差损失和精度。误差损失越低，精度越高，模型性能越佳。

误差损失采用交叉熵计算，如公式(1)所示。

$$E(p, q) = -\sum_x p(x) \log q(x) \quad (1)$$

式中： x 为输入， p 为真实分布， q 为模型的预测分布。

精度（ACC）是真实标签值与模型预测值的差的平方和，如公式(2)所示。

$$Acc = \sum_0^n (y - \bar{y})^2 / n \quad (2)$$

式中： y 为真实标签值， \bar{y} 为模型预测值。

4.3 模型训练及对比分析

模型训练的硬件环境：英特尔 I5-8300H 处理器、英伟达 GTX1050Ti 显卡和 500 GB 固态硬盘；软件环境：Python3.6.5、tensorflow1.8.0、cuda9.0 和 cudnn7.6.5。训练参数设置：BATCH_SIZE 为 32，MAX_STEP 为 6 000，learning_rate 为 0.000 001，激活函数为 ReLU 函数。训练过程误差损失的变化如图 7 所示。

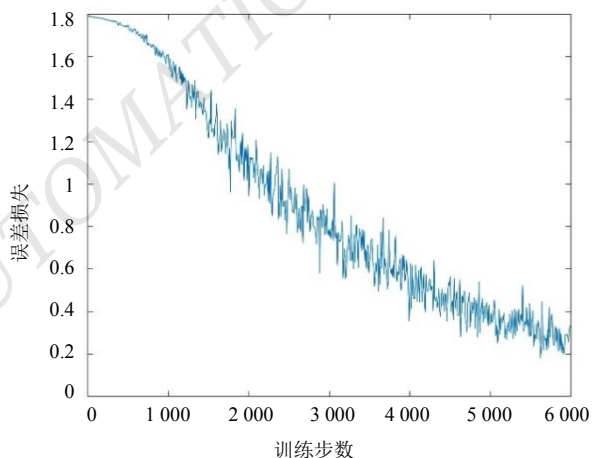


图 7 训练过程误差损失的变化

由图 7 可知，在不断训练过程中，误差损失由 1.8 降低至 0.2，说明模型训练输出不断接近真实标签值。训练过程精度的变化如图 8 所示。

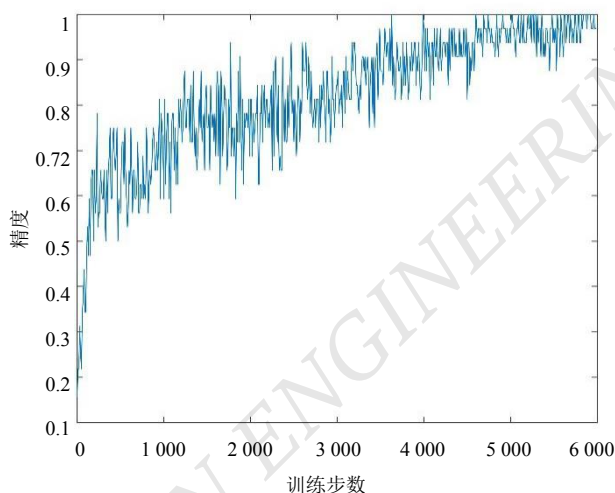


图 8 训练过程精度的变化

由图 8 可知，在不断训练过程中，精度不断提高，在训练步数为 6 000 时，精度接近 0.98。

为进一步测试改进的 LeNet-5 卷积神经网络性能，将该网络模型与经典的 LeNet-5 网络模型进行对比测试。对比测试采用相同的数据集，设置相同的训练步数，重复进行 3 次训练，误差损失结果如表 1 所示，精度结果如表 2 所示。

表 1 误差损失对比表

模型	第 1 次	第 2 次	第 3 次	平均
LeNet-5	0.283	0.226	0.249	0.252
本文模型	0.213	0.190	0.221	0.208

表 2 精度对比表

模型	第 1 次	第 2 次	第 3 次	平均
LeNet-5	0.901	0.896	0.924	0.907
本文模型	0.972	0.977	0.964	0.971

由表 1 可知，2 种网络模型的误差损失相差不大，平均数值较小，说明 2 个网络模型都进行了有效的反馈学习和网络参数的自动调整。

由表 2 可知，改进的 LeNet-5 卷积神经网络模型相较于经典的 LeNet-5 模型，精度明显提升。

编写 Python 和 GUI 界面，载入训练后的分类模

型,使用未经过训练的图像作为推理测试。以盆栽植物点云为例,测试效果如图 9 所示。

图像分类器
zhiwu: 0.905155



图 9 改进的 LeNet-5 卷积神经网络模型测试效果图

改进的 LeNet-5 卷积神经网络模型正确识别了盆栽植物(标签为 zhiwu),置信度为 0.905。作为对比,同样的图像利用经典的 LeNet-5 网络模型进行训练预测,测试效果如图 10 所示。

图像分类器
zhiwu: 0.449495



图 10 经典的 LeNet-5 模型测试效果图

经典的 LeNet-5 网络模型对盆栽植物的置信度为 0.449,说明对于点状成型的特征图片,经典的 LeNet-5 模型难以有效分类识别。

5 结论

本文利用改进的 LeNet-5 卷积神经网络对单物体点云二维图像进行分类识别,相较于经典的 LeNet-5 网络模型具有更高的精度。同时验证了卷积神经网络可用于点云的二维图像识别,相较于完整的物体直观

图像信息,物体点云图像包含物体特征信息较少,识别难度较大,需要改进网络结构以提高识别精度。本文在研究过程中,使用了数据增强技术对数据进行扩充,提高了训练模型的鲁棒性,为后续复合机器人进行更高级的功能开发提供基础条件。

参考文献

- [1] BOHREN J, RUSU R B, JONES E G, et al. Towards autonomous robotic butlers: lessons learned with the PR2[C]//2011 IEEE International Conference on Robotics and Automation. IEEE, 2011:5568-5575.
- [2] 卢荣胜,史艳琼,胡海兵.机器人视觉三维成像技术综述[J].激光与光电子学进展,2020,57(4):9-27.
- [3] 周恒旭,何志敏,周燕.面向机械臂避障的非结构场景感知研究[J].机械与电子,2020,38(9):74-80.
- [4] 俞斌,董晨,刘延华,等.基于深度学习的点云分割方法综述[J].计算机工程与应用,2020,56(1):38-45.
- [5] 张佳颖,赵晓丽,陈正.基于深度学习的点云语义分割综述[J].激光与光电子学进展,2020,57(4):28-46.
- [6] 许安宁.基于深度学习的三维点云语义分割方法综述[J].长江信息通信,2021,34(1):59-62.
- [7] QI C R, SU H, MO K, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:652-660.
- [8] 梁振斌,熊风光,韩燮,等.基于深度学习的点云匹配[J].计算机工程与设计,2020,41(6):1689-1693.
- [9] 敖建锋,苏泽锴,刘传立,等.基于点云深度学习的 3D 目标检测[J].激光与红外,2020,50(10):1276-1282.
- [10] 王旭娇,马杰,王楠楠,等.基于图卷积网络的深度学习点云分类模型[J].激光与光电子学进展,2019,56(21):56-60.
- [11] 蒋腾平,王永君,张林淇,等.融合 CNN 和 MRF 的激光点云层次化语义分割方法[J].测绘学报,2021,50(2):215-225.
- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[J].IEEE Computer Society, 2013.
- [13] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015:1440-1448.
- [14] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Advances in Neural Information Processing Systems, 2015:91-99.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the

IEEE Conference on Computer Vision and Pattern Recognition, 2016:779-788.

[16] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:7263-7271.

[17] REDMON J, FARHADI A. YOLOv3: an incremental

improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[18] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Springer, Cham, 2016:21-37.

[19] 魏溪含,涂铭,张修鹏著.深度学习与图像识别:原理与实践[M].北京:机械工业出版社, 2019.

作者简介:

熊治国, 男, 1977 年生, 博士, 副教授, 主要研究方向: 无人机飞行控制。E-mail: xiongzhi guo416@163.com

周恒旭(通信作者), 男, 1992 年生, 硕士, 助教, 主要研究方向: 机器人智能控制等。E-mail: 1281859757@qq.com

~~~~~

(上接第 10 页)

根据表 6 的分析结果并结合表 7 优化起落架后的重量, 本文选用起落架优化斜角和上端厚度的方式, 提升多旋翼无人机的固有频率及其续航时间。

## 5 结论

针对多旋翼无人机装载 5G 模块后, 导致其飞行不稳定的问题, 本文利用 SolidWorks 和 ANSYS Workbench 软件对多旋翼无人机进行模态分析。通过具体分析多旋翼无人机的固定板厚度、5G 模块装载位置和起落架升高后的固有频率, 得知起落架结构对多旋翼无人机的固有频率有显著影响。

为了避免多旋翼无人机的固有频率和激励频率相近导致共振现象的发生, 对多旋翼无人机的起落架进行结构优化设计, 提高多旋翼无人机的固有频率, 进而保证其飞行稳定性。

#### 参考文献

[1] FENG Licheng, KATUPITIYA Jay. UAV-based persistent full area coverage with dynamic priorities[J]. Robotics and Autonomous Systems, 2022, 157: 104244.

[2] 陈梅香,张瑞瑞,陈立平,等.无人机农林业应用全球研究态势

分析[J].智慧农业(中英文),2021,3(3):22-37.

[3] 李永红,洪小刚,邢涛.无人机在桥梁检测中的实用研究[J].测绘与空间地理信息,2022,45(7):81-84.

[4] 梁宇锋,潘满清,霍翔,等.基于无人机多光谱图像的荔枝叶片氮磷钾含量检测[J].机电工程技术,2022,51(5):138-142.

[5] 张红,宋凯廷.5G 与超高清视频融合应用关键技术探讨[J].数字通信世界,2022(9):100-102.

[6] 郑杰,陈高亮,时海龙,等.5G 环境下无人机高清视频数据实时传输组网方式研究[J].通信技术,2020,53(4):923-926.

[7] 王筱淇,刘震磊,李颖,等.四轴无人机总体结构模块化设计[J].中国科技信息,2020(2):21-23.

[8] 田卫军,李郁,何扣芳,等.四轴旋翼飞行器结构与模态分析[J].制造业自动化,2014,36(4):37-39.

[9] 刘峰,高鸿渐,喻辉,等.基于有限元的四旋翼无人机碳纤维结构优化设计与固有模态分析[J].玻璃钢/复合材料,2017(4):17-23.

[10] 练伟,王华山,余嘉.基于有限元分析的 3D 打印脊柱侧弯矫形器局部优化设计[J].医用生物力学,2021,36(6):855-861.

[11] 段国晨,赵景丽,赵伟超.先进复合材料在无人机结构的应用[J].纤维复合材料,2022,39(2):105-114.

[12] 张坤金,郑忠才,高岩,等.不同网格划分对机体有限元模态分析结果的影响[J].小型内燃机与摩托车,2009,38(5):69-71.

[13] 魏文菲,张春元,李超,等.某四旋翼飞行器机架的模态分析[J].兵器装备工程学报,2017,38(2):40-42.

#### 作者简介:

刘诚, 男, 1998 年生, 硕士研究生, 主要研究方向: 复杂系统的集成与设计。E-mail: 805327372@qq.com

刘忆森(通信作者), 女, 1985 年生, 博士研究生, 副研究员, 主要研究方向: 智能传感及深度学习。E-mail: ys.liu@giim.ac.cn

周松斌, 男, 1978 年生, 博士研究生, 研究员, 主要研究方向: 智能传感与检测。E-mail: Sb.zhou@giim.ac.cn

赵路路, 男, 1995 年生, 硕士研究生, 助理工程师, 主要研究方向: 复杂系统的集成与设计。E-mail: 282731537@qq.com

袁飞, 男, 1984 年生, 博士, 讲师, 主要研究方向: 智能数据处理。E-mail: eric\_f\_y@foxmail.com