

大连理工大学本科毕业设计开题报告

基于深度学习的机器人抓取方法研究与实现

Research and Implementation of Robot Grasping Method Based on Deep Learning

学 院(系): 电子信息与电气工程学部

专 业: 计算机科学与技术

学 生 姓 名: 范瀚文

学 号: 201981303

指 导 教 师: 杨鑫 教授

开题报告日期: _____

大连理工大学

Dalian University of Technology

说 明

一、开题报告应包括下列主要内容：

1. 课题来源及研究的目的和意义；
2. 主要设计要求；
3. 国内外在该方向的研究现状及分析；
4. 主要研究内容；
5. 研究方案及预期达到的目标；
6. 为完成课题已具备和所需的条件和经费；
7. 主要参考文献。

二、开题报告由指导教师填写意见、签字后，统一交所在院系保存，以备检查。

指导教师评语：本次开题报告，内容丰富，结构合理，论述清晰，数据准确，引用规范，符合学科要求和研究方向。该学生对于相关文献进行了广泛的阅读和分析，对研究问题和方法有了明确的思路 and 计划。该学生具有较强的自主学习能力和创新意识，展现了良好的学术素养和专业水平。本次开题报告达到了预期的目标和效果，值得肯定和表扬。建议在后续的研究过程中继续保持认真负责的态度，注意跟进最新的文献动态和技术进展，完善研究设计和实验方案，及时和指导教师沟通过交流，解决可能遇到的问题和困难。相信该学生能够顺利的完成毕业设计，并取得优秀的成果。

指导教师签字：_____

检查日期：_____

1 课题来源及研究目的和意义

近几十年来，全球大量研究者投入到机器人的研究中，工业领域的机器人技术获得巨大的进步，大量的工业机器人被应用在工业生产过程中。目前，应用于日常生活中的智能服务机器人领域方兴未艾^[1]。但是，目前投入市场的服务机器人智能化程度较低^[2]。能够实现快速、精准抓取的服务机器人仍然处于实验室研究阶段，还未能大量应用于日常生活。随着中国开始进入老龄化阶段，社会对服务机器人的需求不断增加。为了护理照料残疾人和老年人，服务机器人需要能对日常环境下的物体，比如水杯，牙膏桶等生活用品进行抓取^{[3][4]}。因此，研究服务机器人抓取技术具有现实意义。同时由于在实际场景中物体千变万化，建立模板搜索库通常费时费力。鉴于服务机器人的应用场景，研究一种基于视觉的、能够在非结构环境下实现的机器人自主抓取技术至关重要，同时对于提升服务机器人的场景交互能力、推广服务机器人的应用也大有裨益。

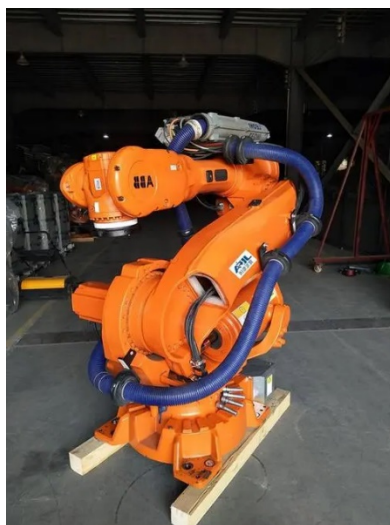


图 1.1 ABB 工业机器人

随着机器学习技术的普及，目前的自主抓取方法逐渐将抓取规划算法转为抓取检测算法进行研究，因此自主抓取的研究重点是建立抓取分类器并基于分类器检测最优抓取。虽然目前的研究中已经建立了一些抓取分类器，但是这些分类器通常需要人为设计特征，特征的设计是经验式、启发式的，并且针对不同抓取条件可能需要设计不同的特征。如今，随着大数据时代的到来，以及深度学习技术的飞速发展，人类向智能时代更进一步，深度学习中的卷积神经网络(Convolutional Neural Network,CNN)，是一种有监督的特征学习方法，在大数据的驱动下能够学习到事物或行为中的深层特征，尤其是在二维特征

的学习与表达方面，如图像分类、目标检测、行为识别等领域，卷积神经网络已经展现出了十分优异的效果，目前的一些研究甚至已经超过了人类的识别能力。将深度学习引入机器人自主抓取领域，如基于卷积神经网络建立抓取分类器，无需人为设计特征，其本身可以学习到目标物体区域的可抓取特征，如果再将不同模态的传感数据分别输入给卷积神经网络，使其学习不同传感数据的抓取特征，将会产生更高的抓取分类精度，对于机器人在实际场景中的稳定抓取将带来质的提高。因此，采用卷积神经网络建立的机器人抓取分类器能够学习多种模态的抓取特征，对于机器人自主抓取领域的研究具有重要的启发意义。

2 主要设计要求(指标或参数)

本次设计的目标是，通过 Kinect 获得未知物体的单一视角图像，建立抓取表达与抓取参数之间的映射准则，基于卷积神经网络建立抓取分类器，最终将最优抓取表达映射至抓取参数实现对目标物体的成功抓取。为此需要在以下几点展开研究：

(1)基于点云的目标分割

目标分割的目的在于准确区分物体区域与桌面区域，从而使抓取采样算法可以集中在物体区域附近，避免大量无用检测，且提升抓取检测算法的实时性。

(2)基于卷积神经网络的抓取分类器

基于卷积神经网络建立抓取分类模型，将构建像关于原始颜色特征与原始深度特征的两个卷积神经网络分类器，最后再将两个分类器作为特征提取器，分别提取颜色特征与深度特征，最终构造出基于多模态卷积神经网络的抓取分类模型。

(3)自主抓取系统实验研究

为了验证抓取分类模型的有效性，本文将在数据集上进行检测以评估分类模型的分类性能，同时基于 Baxter 双臂机器人进行 Gazebo7.0 模拟环境下的抓取实验，包括 Kinect 位姿标定，物体抓取等内容。

3 国内外研究的历史和现状

从工业机器人向服务型机器人转变，国外和我国都进行了不少尝试，且都有不错的产品诞生。国外研究人员较早开展服务机器人领域的研究，目前取得一些成果。图 3.1 左侧是一款名为 PR2 的服务机器人，配备有一对机械臂，末端分别有一对二指夹具。PR2 搭载机器人专用操作系统；配有高性能硬件，包括中央处理器，各类视觉和导航定位传感器。目前，PR2 主要是用于科学研究中，尚未运用到日常生活中。研究人员已经实现 PR2 的自主抓取、自主开门等功能。

如图 3.1 右侧所示是一款名为 Home Exploring Robotic Bulter(HERB)的服务机器人, HERB 采用轮式自主移动, 配有的一对机械臂各有七个自由度, 工作范围较大, 它配有 RGB-D 深度相机传感器, 可实现物体识别与定位^[5]。目前, HERB 主要运用于科学研究。美国卡内基梅隆大学等研究者们实现了 HERB 的物体抓取搬运等功能^[5]。

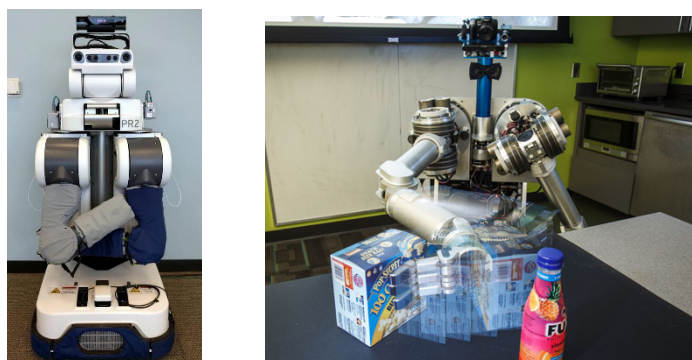


图 3.1 国外服务型机器人

国内对于服务型机器人的研究也不落下风, 北京理工大学黄强教授带领的团队研发了第五代“汇童”系列移动机器人 BHR-5, 如图 3.2 所示。BHR-5 仿造了人体结构, 全身关节较多, 自由度较高, 头部可移动角度较大, 安装了视觉传感器, 视野广;四肢配有各类力学传感器, 可模仿人的各种动作, 该机器人已经实现了乒乓球对打功能。

上海交通大学的曹其新教授团队研发了一款全向移动服务机器人, 该机器人头部可以转动, 并配有视觉传感器, 视野广, 能够实现目标物体的识别和定位功能; 配有一对机械臂, 各有 6 个自由度, 可以实现对放在水平操作平台上的竖直放置的物体进行定位与抓取^[6]。

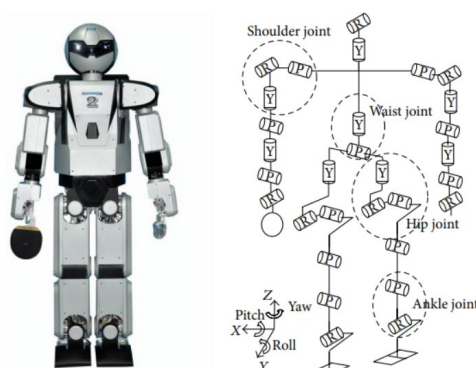


图 3.2 BRH-5 机器人

3.1 深度学习研究现状

深度学习是由机器学习的一个方向发展而来的。在机器学习中，深度学习是一种基于数据驱动算法。深度学习的概念是相对于如支持向量机和逻辑回归等浅层学习而言的。深度学习的本质是具有许多层的神经网络，它使用前一层的输入作为下一层的输出来学习高度抽象的数据特性。与传统的神经网络相比，深度学习的表征能力更强。经过对深层神经网络的训练和对图像分类等任务的适当调整，可以提高特征提取能力，逐渐取代了手工设计特征，大大节省了开发时间。并且深度学习也具备可塑性。通常，对于传统算法，如果需要调整模型，可能需要对代码进行大量更改。对于深度学习的网络框架，如果需要调整模型，只需调整参数，因此深度学习具备非常高的灵活性。深度学习的应用也具有广泛性，它可以根据问题建模，而不局限于一个固定的问题。

深度学习在图像领域中的运用，主要是采用卷积神经网络 (Convolutional Neural Networks, CNN) 的方法。卷积神经网络核心结构是卷积层。卷积层通过卷积核在图像上滑动提取图像中的特征。因此，卷积神经网络不断地取代了手工设计特征，被广泛地运用到图像的特征提取中。

目前，被广泛使用的卷积神经网络的结构主要是由 LeNet 卷积神经网络确定的。YanLeCun 于 1994 年提出了 LeNet 卷积神经网络。LeNet 网络主要被用于识别手写字符，网络结构比较简单，只包含了 2 个卷积层用于提取手写字符的特征。紧接着 AlexNet 的诞生，卷积神经网络迎来了新的研究高潮。AlexNet 在 LeNet 网络结构的基础上增加了几个卷积层，获得更好的特征提取能力。它表明运用复杂的神经网络模型可以有效地提高图像处理的能力。至此，确定了卷积神经网络在图像处理领域的绝对优势。由此，研究人员不断增加卷积神经网络的层数，以期获得更好的特征提取能力。VGGNet 将网络层数增加到 10 至 19 层，典型的层数是 16 和 19。VGGNet 提出卷积块的概念，由几个卷积层构成一个卷积块，可以有效地减少参数量，并且使感受范围更大^[7]。VGGNet 开始采用 ReLU 函数作为激活函数，非线性表示能力增强，可以训练更深的网络层。

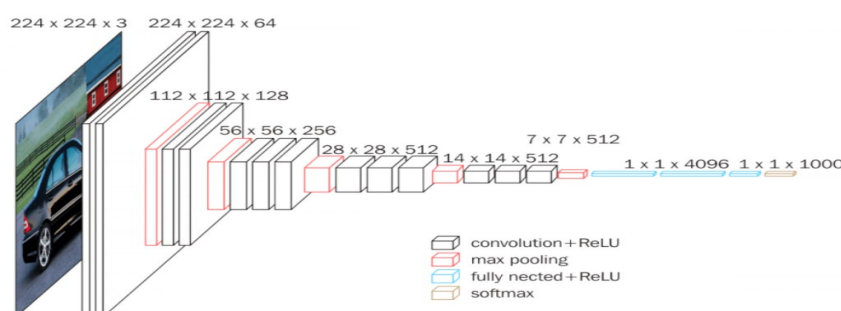


图 3.3 VGG16 层数示意图

从 LeNet 到 AlexNet 和 VGGNet 不断增加卷积神经网络的层数,提高了提取特征的能力,但是也导致了梯度消失,无法训练深层网络的问题。ResNet 则采用残差学习单元,通过拟合残差的方式有效地减少了梯度消失的可能,获得了既增加了网络深度又不影响网络训练的效果^[8]。

以上的卷积神经网络架构在图像分类任务中获得了很高的分类准确率。除此之外,将深度学习运用到目标检测领域也取得了快速的发展和显著的成果。目标检测任务主要是对图像中的目标进行分类和定位,一般用一个矩形边界框标注目标在图像中的位置。目前,基于深度学习的目标检测方法可以分为两大类:一类是基于区域提议的目标检测方法,另一类是基于回归的目标检测方法^[9]。基于区域提议的目标检测方法首先生成区域提议,然后在每个提议区域上进行分类和预测边界框。基于区域提议的两阶段目标检测方法主要包括 R-CNN、Fast R-CNN、Faster R-CNN 等。基于回归的目标检测方法采用回归的思想,直接在整个图像上回归出目标类别和预测边界框。因为这个方法是端到端直接输出目标类别和预测边界框,所以这类方法也被统称为一阶段目标检测方法。基于回归的一阶段目标检测方法主要包括 YOLO (You Only Look Once)系列和 SSD (SingleShot Multibox Detector)系列。而最新图像定位算法则开始尝试两阶段深度学习的方法,这类算法和之前的算法相比往往具有很好的速度优势和精度优势,其中比较著名的是轻量级的残差挤压和激励网络 GGCNN^[10], GGCNN2 这两个和机器人相结合的目标定位检测与抓取算法。

3.2 抓取技术研究现状

抓取检测模块、抓取规划模块和控制模块组成了机器人抓取系统^[11]。其中,抓取检测模块是整个系统的重点部分,抓取检测是否准确,直接影响到后续的抓取是否成功。抓取检测模块可以分为三个部分:目标定位和识别、位置和姿态估计和确定抓取点。

随着光学传感器的发展,机器人已经能够通过视觉感知系统进行自动抓取。传统上,基于视觉的机器人抓取系统由一系列组件构成,包括目标定位、目标姿态估计、抓取检测和抓取规划。目标定位是在输入图像中确定目标物体的位置。目标姿态估计是目标物体相对于参考点的旋转和平移。抓取检测是关于目标对象的抓取配置,如抓取点、抓取角度和抓取器的开度大小。抓取规划是指为了安全地抓取物体将物体从其静止表面托起所需的路径规划过程。

为了简化待抓取目标物体的定位问题,最早的一些方法假设待抓取目标被放置在单一的背景里。因此,当背景环境稍微有一些复杂时,这些方法就会无效。一些目标检测方法利用机器学习的方法来训练基于二维手工标定的分类器。然而,由于手工标定的限

制，这些分类器只有有限的性能。近年来，深度学习已经开始主导与图像相关的任务，如目标检测和目标分割。除此之外，训练数据从 RGB 图像扩充到 RGB-D 深度图像，并提出了输入二维图像或三维图像的深度学习网络，极大地提高了目标定位的性能，促进了机器人抓取技术的快速发展。

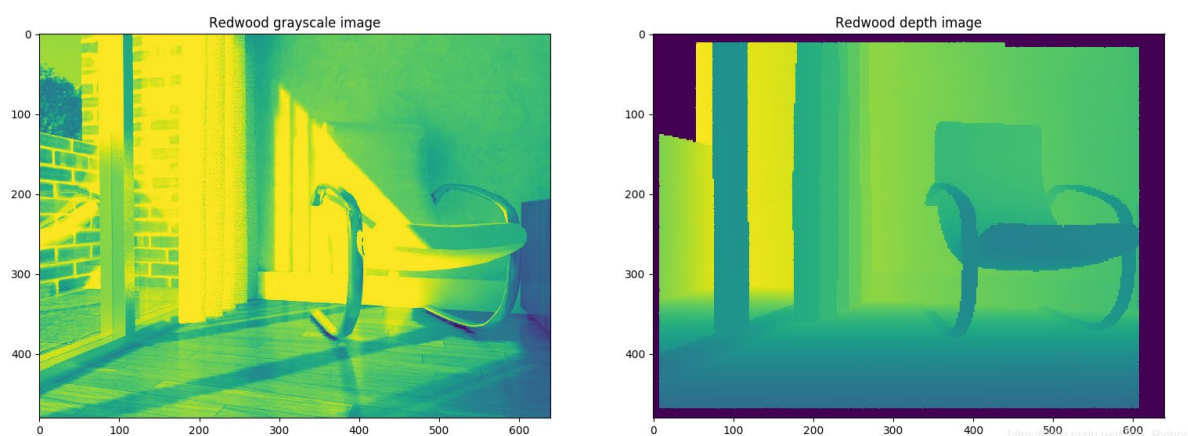


图 3.4 RGB-D 深度图像示例

6D 姿态表示方法是一种表示三维位置信息的方法，它可以表示物体三维坐标 (x,y,z) 和在三维坐标上的旋转信息。6D 姿态表示方法大致可分为四种，分别是基于对应、基于模板、基于投票和基于回归。基于对应关系的目标位姿估计方法主要是针对具有丰富纹理的物体进行二次估计，以实现二维空间里面的特征点的匹配。首先从不同角度投影已有的三维模型，得到多幅图像。通过寻找图像上的 2D 特征点与渲染图像之间的匹配，建立 2D 像素点到 3D 像素点的对应关系。基于对应关系的方法很难计算不含纹理的物体的姿态，而基于模板的姿态估计可以实现这一点。在基于模板的姿态估计方法中，通常使用梯度信息，将现有三维模型从不同角度投影生成的多幅图像作为模板。基于投票的目标姿态估计方法主要用于计算遮挡物体的姿态。对于这些对象，每个图像块通常用于对输出进行投票。Wang 等人提出了一种通用的框架，称为 DenseFusion，可以实现直接从彩色图像和深度图像中评估出已知抓取目标物体的 6D 姿态。DenseFusion 是一种独立处理两个数据源（RGB 图像和深度图像）的异构体系结构，它使用一种新的网络来提取像素的密集特征，并从中对其进行估计。基于回归的方法同时完成目标检测和 6D 姿态估计。与采用多阶段策略从输入图像中估计目标姿态的其他方法不同，这种方法学习从输入图像到姿态参数表示的即时映射方法，从而将 6D 目标姿态估计与目标检测结合起来。基于回归的方法可以分为两种：一种是直接进行回归 6D 物体姿态，另一种是回归关键点位置，提供二维和三维对应关系。

在两阶段抓取检测方法中,滑动窗口策略通常用于检测二维抓取^[12]。Lenz 等人提出了一个由两个深度卷积神经网络组成的两阶级的级联系统,其中第一个深度卷积神经网络输出的检测结果由第二个深度卷积神经网络重新评估^[13]。第一个深度卷积神经网络具有较少的特征,运行速度更快,并且可以有效地减少错误的候选抓取。第二个深度卷积神经网络,功能更多,速度较慢,只能对几个候选抓取进行检测。尽管它们达到了很高的精度,但迭代扫描过程非常缓慢。夏晶等人提出一种基于目标检测的级联卷积神经网络的机器人平面抓取位姿快速检测方法,检测速度明显快于滑动窗口策略的两阶段抓取检测方法^[14]。

上海理工大学的余玉琴等人提出一种改进的 YOLOv2 抓取检测算法。此算法只利用 RGB 图像,对图像中的待抓取目标物体进行识别和定位,然后再采用 K-means 聚类方法对深度图像中的物体进行聚类分析,以确定待抓取目标物体到摄像机的大致距离^[15]。东南大学赵敏等人进行了基于机器人与六维力传感器的抓取系统的设计与实现。采用导纳控制算法使原本只有位置控制能力的机器人同时具有位置和力两种控制能力。

4 主要研究内容

4.1 机器人操作环境的模拟搭建

现实中的机器人抓取往往在复杂条件下进行,其抓取精度和抓取方式往往会受到实际环境中光线环境和实际物体摆放位置等因素的严重干扰。所以为了研究机器人抓取中的精度、速度等问题,首先要将机器人所处的环境进行简化建模来达到进行测试的效果。因此,将机器人进行抓取的环境要求如下:

- (1) 机器人所处的环境视为光线均匀且亮度充足,不会发生突变;
- (2) 抓取物体较小,且位置摆放合理,在机器人一次抓取的能力范围内;
- (3) 机器人只需要进行机械臂的移动就可以完成对物体的抓取;
- (4) 整个过程中认为机器人夹爪的抓力足够强,被抓取物品表面摩擦力足够。

4.2 深度学习算法构建

机器人抓取的流程可划分为:抓取检测、抓取规划和抓取控制三个子模块。其中抓取控制模块往往由机器人自身规划完成,不需要额外自主进行编写。抓取控制模块则根据不同的机器人种类各有差别,不同机器人使用基于机器人操作系统(ROS)的自主控制方式和控制参数,这些参数可以通过不同机器人的开发文档中查阅得到。

抓取检测根据不同的划分标准可划分为不同的类别。如按照抓取表示划分,可以分为 2D 平面抓取和 6DoF 空间抓取,前者适合工业场景,后者适合更灵活的场景。按照

抓取数据划分,可以分为基于 RGB 图像、基于深度图像或基于 RGB-D 图像的方法,不同的数据源提供不同的信息量和噪声水平。按照抓取任务划分,可以分为具有已知定位和姿态、具有已知定位和无姿态、无定位和无姿态的方法,不同的任务难度和要求不同。按照抓取策略划分,可以分为分析方法或数据驱动方法,前者根据物理学原理选择抓取点,后者根据大量数据学习选择抓取点。而按照深度学习标准对于机器人抓取方式进行划分,可大致划分为如下种类:

(1) 基于回归的方法:基于回归的方法可以分为两类。一种是直接回归方法,这类方法直接从输入图像中输出一个或多个抓取位姿,包括位置、角度和宽度。例如,GGCNN2^[15]是一种端到端的全卷积网络,它可以从深度图中预测出每个像素点处的抓取成功率、抓取宽度和抓取角度。另一种是间接回归方法:这类方法先从输入图像中提取出一些特征或候选区域,然后再对每个特征或候选区域进行抓取位姿的回归。例如,Deep6DPose 是一种先采样后评估的方法,它先在深度图上采样出多个抓取位姿以及与其对应的图像块,然后再将图像块和采样出来的抓取深度作为网络输入,输出给定抓取的置信度。基于回归的方法有很多优点:它们不需要预先定义或训练一个抓取模型或模板,而是通过端到端地学习来适应不同形状和大小的物体。它们可以同时完成目标检测和 6D 姿态估计等任务,减少了计算复杂度和误差累积。常见的基于回归的方法有 BB8、SSD6D^[17]、PoseCNN^[18]、Deep6DPose 等。

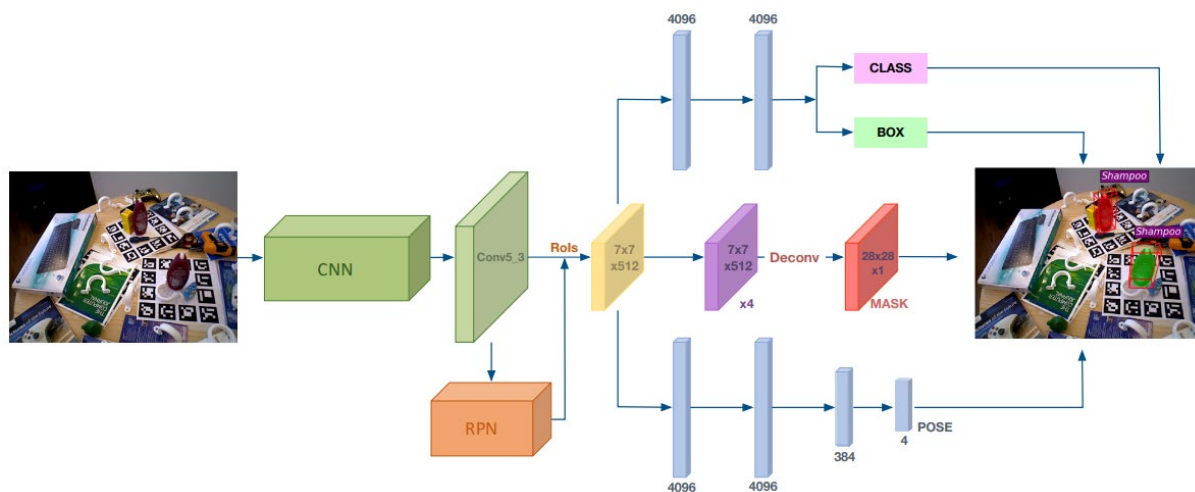


图 4.1 Deep6DPose 算法流程示意图

(2) 基于分类的方法:基于分类的机器人抓取算法是一种利用深度学习网络从输入图像中预测出抓取位姿的可行性或置信度的方法。基于分类的方法可以分为两类,一是基于采样和评估的方法:这类方法先从输入图像中采样出一些抓取位姿候选,然后再

对每个候选进行分类，输出其抓取成功率或置信度。例如，Dex-Net 2.03 是一种先采样后评估的方法，它先在深度图上采样出多个抓取位姿以及与其对应的图像块，然后再将图像块和采样出来的抓取深度作为网络输入，输出给定抓取的置信度^[19]。另一种是基于直接预测的方法：这类方法直接从输入图像中输出一个或多个抓取位姿及其可行性或置信度。例如，MultiGrasp 是一种基于 YOLO 框架的方法，它可以从 RGB-D 图中预测出每个物体上多个定向矩形框，并输出每个框代表的平行夹持器抓取位姿和成功率。基于分类的方法有很多优点，它们可以利用大量数据集进行训练，并提高泛化能力和鲁棒性。它们可以根据不同任务或场景选择最合适或最可靠的抓取位姿基于分类的方法。比较常见的算法有 Dex-Net、GG-CNN、GraspNet^[20]等。

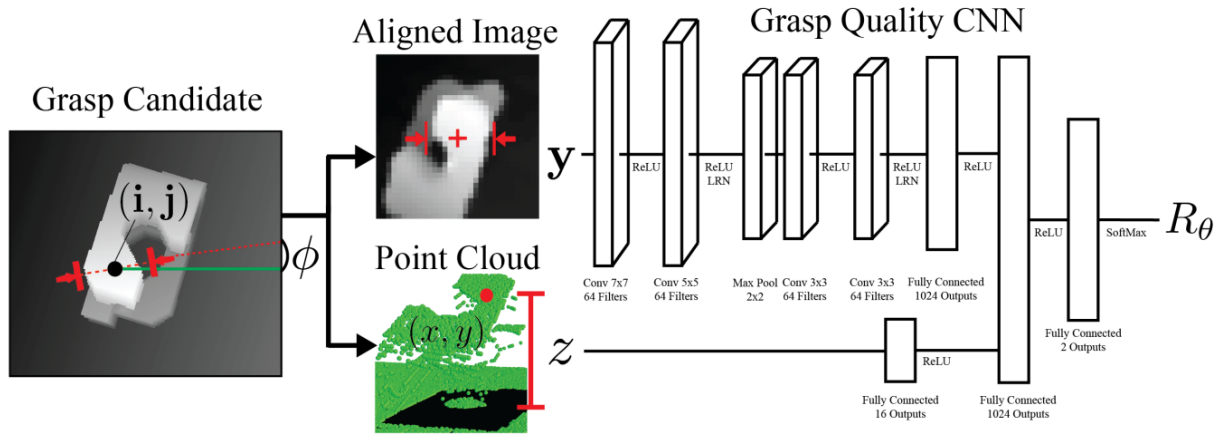


图 4.2 Dex-Net 网络示意图

(3) 基于强化学习的方法：基于强化学习的机器人抓取算法是一种利用智能体与环境交互，通过不断尝试和反馈来学习最优的抓取策略的方法。基于强化学习的方法可以分为以下两类：一是基于值函数的方法：这类方法先定义一个状态-动作对应的价值函数，然后通过迭代或近似来更新价值函数，从而找到最大化累积回报的动作。例如，Deep Q-learning (DQL) 是一种基于值函数的方法，它使用深度神经网络来拟合 Q 值，并利用经验回放和目标网络来提高稳定性和收敛性^[21]。另一种是基于策略梯度的方法：这类方法直接优化一个参数化的策略函数，使得期望回报最大化。例如，DDPG 是一种基于策略梯度的方法，它结合了 Actor-Critic 结构和 DQN 中的技巧，可以处理连续动作空间和高维状态空间^[22]。基于强化学习的方法有以下优点：它们可以从零开始自主地探索和学习抓取策略，而不需要人工标注或专家示范。它们可以同时考虑视觉信息和触觉信息，并进行多模态感知和决策。常见的算法有 Q-Learning、Policy Gradient、Actor-Critic 等。

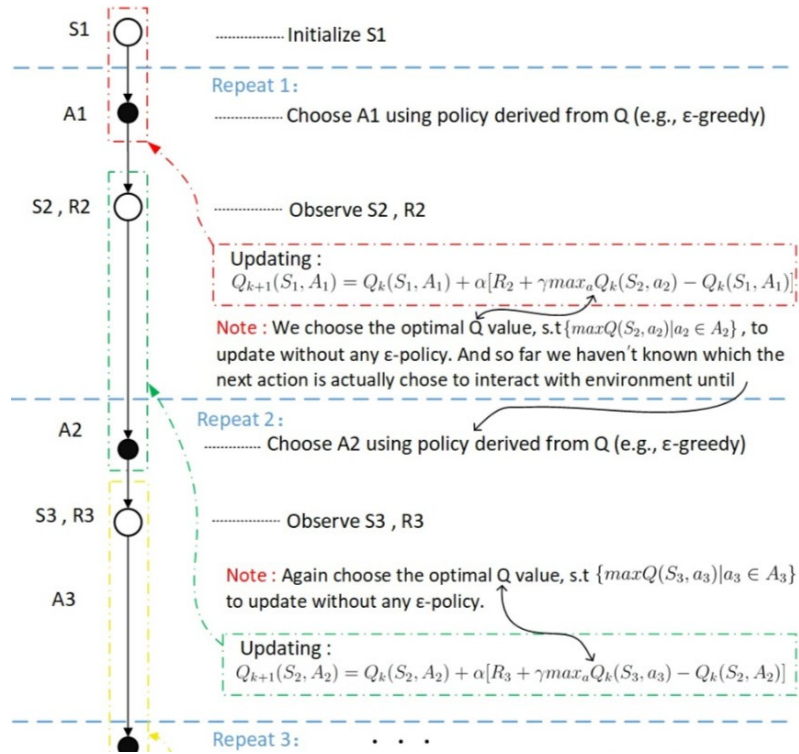


图 4.3 Q-Learning 算法流程示意图

不同的算法在针对不同条件的抓取过程中有各自的优势，本次实验计划对常见的深度学习算法进行实验和比较，最终构建出一个满足速度和精度要求的，具有可操作性的轻量化抓取检测算法。由于本次实验的抓取过程在模拟环境下抓取距离较近的小型物体，在本设计中采用 GG-CNN 作为主要使用算法。

5 研究方案及预期达到的目标

通过前面的分析，本课题研究方案可确定如下：首先在虚拟环境中对机器人抓取环境进行建模模拟；学习深度学习的目标检测和机器人操作系统的控制方法，完成机器人的虚拟环境操作；针对目标抓取的需要，构建符合需求的目标检测和机器人抓取算法，以期在抓取精度和抓取速度上取得优势；以 Gazebo7.0 作为模拟平台，进行系统仿真，并验证最终效果。

预期目标：

- (1) 构建满足基本环境的机器人抓取环境建模。
- (2) 掌握机械臂操作流程唤醒算法。

(3) 测试现有的深度学习机器人抓取算法，构建满足速度和精度的深度学习机器人抓取算法。

(4) 实现从目标识别到机器人抓取整个流程的仿真。

6 为完成课题已具备和所需的条件和经费

本课题是基于 Ubuntu16.04 系统下、基于机器人编程语言 ROS 的 Gazebo 7.0 模拟环境下进行的实验与算法研究工作。其中整个 ROS 生态由通信(Plumbing)、工具(Tools)、功能(Capabilities)与社区(Community)四大部分组成，如下图 6.1。和现实环境相比，Gazebo7.0 仿真环境具有仿真环境自由，可选择机器人不受经济条件限制，试错成本低和机器人操作不受现实硬件条件限制等优点，Gazebo 模拟器显示如下图 6.2 所示，所使用机器人实物如图 6.3 所示。



图 6.1 ROS 系统组成

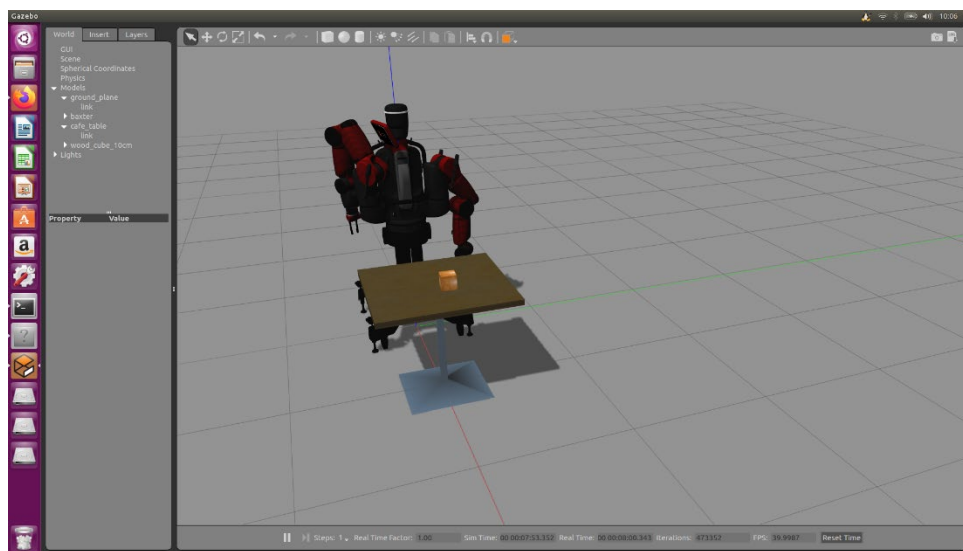


图 6.2 Gazebo7.0 模拟抓取环境



图 6.3 Baxter 双臂协同操作机器人

参 考 文 献

- [1] Barakeh Z A , Alkork S , Karar A S , et al. Pepper Humanoid Robot as a Service Robot: a Customer Approach[C]// 2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART). 2019.
- [2] 颜云辉, 徐靖, 陆志国,等. 仿人服务机器人发展与研究现状[J]. 机器人, 2017(4):14.
- [3] Kimmo J V, Salin S E.A Need for Service Robots Among Health Care Professionals in Hospitals and Housing Services[C]. International Conference on Social Robotics,2017,10652:178-187.
- [4] Chivarov N, Chikurtev D, Markov E, et al. Cost Oriented Tele-Controlled Service Robot for Increasing the Quality of Life of Elderly and Disabled-ROBCO[C]. 18th International Federation of Automatic Control Conference on Technology Culture and International Stability, 2018:192-197.
- [5] 机器人管家:HERB 机器人[J]. 机械工程师, 2013(6):1.
- [6] 余蕾斌.全方位移动仿人型机器人运动规划研究[M].上海交通大学,2010.
- [7] Simonyan K , Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. arXiv:1409.1556, 2014.
- [8] He K , Zhang x , Ren S , et al. Deep Residual Learning for Image Recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas,N2016:770-778.
- [9] 李章维,胡安顺,王晓飞.基于视觉的目标检测方法综述[J].计算机工程与应用,2020,56(08):1-9.
- [10] H. Cao, G. Chen, Z. Li, J. Lin and A. Knoll, et al. Residual Squeeze-and-Excitation Network with Multi-scale Spatial Pyramid Module for Fast Robotic Grasping Detection[C]. 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 13445-13451
- [11] Sulabh K, Christopher K. Robotic grasp detection using deep convolutional neural networks[C]. International Conference on Intelligent Robots and Systems(IROS),2017:769-776.
- [12] 李华健. 基于深度学习目标检测的服务机器人抓取技术研究[D].哈尔滨工程大学, 2020. DOI:10.27060/d.cnki.ghbcu.2020.000650.
- [13] Lenz I, Lee H, Saxena A. Deep Learning for Detecting Robotic Grasps[J]. International Journal of Robotics Research, 2014,34(4-5):705-72.
- [14] 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(06):28-36.
- [15] 余玉琴,魏国亮,王永雄.基于改进 YOLOv2 的无标定 3D 机械臂自主抓取方法[J].计算机应用研究,2019,37(5):1-8.
- [16] Hu W , Wang C , Liu F , et al. A grasps-generation-and-selection convolutional neural network for a digital twin of intelligent robotic grasping[J]. Robotics and Computer Integrated Manufacturing: An International Journal of Manufacturing and Product and Process Development, 2022(77-):77.
- [17] Kehl W , Manhardt F , Tombari F , et al. SSD-6D: Making RGB-based 3D detection and 6D pose estimation great again[J]. IEEE Computer Society, 2017.
- [18] Xiang Y , Schmidt T , Narayanan V , et al. PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes[J]. 2017.

- [19] Mahler J , Liang J , Niyaz S , et al. Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics[J]. 2017.
- [20] Fang H S , Wang C , Gou M , et al. GraspNet: A Large-Scale Clustered and Densely Annotated Dataset for Object Grasping[J]. 2019.
- [21] Gu S , Lillicrap T , Sutskever I , et al. Continuous Deep Q-Learning with Model-based Acceleration[J]. JMLR.org, 2016.
- [22] Hou Y , Liu L , Wei Q , et al. A novel DDPG method with prioritized experience replay[C]// IEEE International Conference on Systems. IEEE, 2017.