

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

本科生校企合作毕业设计 开题报告（2024 版）



项目名称： 3D 视觉识别及其在机器人抓取中的应用

学生姓名： 代贤晨

杨帆

崔城铭

指导教师： 李广晔

学 院： 机械与动力工程学院

企业导师： 张云飞

企业名称：深圳慧智星晨科技有限公司

填表说明

开题报告应包含以下内容（报告全文使用宋体小四号字体填写，1.5倍行间距）：

1. 《上海交通大学毕业设计开题报告》封面：
应包含项目的名称、赞助企业名称、学生姓名、指导老师等。
2. 《保密协议》：
指导老师和每位学生都需要在保密协议上签字。
3. 摘要：
对项目内容、设计规范、问题分析、项目计划以及项目成果和意义的总结，限定在一页纸内，应包含项目的关键技术。
4. 问题描述和简介（包含文献综述）：
描述项目背景、项目内容以及预期成果。描述现有相关产品、技术的发展及应用情况，完成相关文献综述，并描述项目的创新性。当引用外部资源时，应对所引用的文献、技术基准标注来源。
5. 项目需求与设计规范书：
描述项目需求以及将需求转化为设计规范的详细步骤，说明项目目标的确定过程。
6. 概念设计方案拟订：
概念设计方案初步拟订，可以描述多个设计方案。
7. 项目进度计划及预算：
项目重要时间节点（用甘特图等方式表示项目计划），小组成员的工作量以及预算。
8. 结论：
总结报告中的重要内容。
9. 参考文献：
列出完整的外部信息来源，并在文中标注参考文献

上海交通大学

校企合作毕业设计保密协议

项目名称：3D视觉识别及其在机器人抓取中的应用。

资助企业：深圳慧智星晨科技有限公司。

鉴于上述企业资助上海交通大学机械与动力工程学院全日制本科生开展毕业设计工作，企业将以口头、书面、数码、或其它形式向导师及学生提供或批露某些工业机密、贸易机密、业务信息、技术信息、以及其它信息（以下统称“**保密信息**”）。

因此，项目指导老师和学生需遵守如下规定：

1. 保密信息为企业所有，项目指导老师和学生应对保密信息严格保密，并且不向任何第三方透露该信息。
2. 除非企业明确书面授权，项目指导老师和学生不得将保密信息披露给任何第三方。
3. 项目指导老师和学生不得利用保密信息(i)与企业直接或间接地竞争，或(ii)服务于当事人之间潜在的业务合作、交易和商业机会之外的任何目的。
4. 保密义务不适用于项目指导老师和学生通过下列任一方式所获得的信息：
 - (a) 属于公开领域的，或
 - (b) 由项目指导老师和学生合法拥有，而非直接或间接地从与企业相关的某方，或承担保密义务的第三方所获得的，或
 - (c) 由第三方基于诚信或权利而提供的，或
 - (d) 由于非项目指导老师和学生的行为而为公众所知的。
5. 本协议项下义务应自本协议签署之日起立即生效，且持续有效三年。
6. 企业披露保密信息并不构成就保密信息之任一部分向项目指导老师和学生授予许可或任何其它权利。
7. 根据本协议而建立的关系应为保密性质，且应根据本协议条款视为保密信息的一部分。除非企业事前书面许可，项目指导老师和学生不得通过任何媒介、渠道和方式披露其参与本协议、或本协议内容、或任何项目指导老师和学生因本协议而制作和开发的信息、材料、报告或其它工作内容。
8. 如企业书面要求，项目指导老师和学生应根据本协议收到的全部保密信息，包括但不限于保密信息的概要、摘录和复制品送还给企业。

9. 由毕业设计项目中产生的知识产权为双方共有财产，拟公开发表的论文（期刊和会议）或专利需事先征得双方同意。
10. 本协议的所有条款应被视为可分割的，如任一条款被认定为无效的、非法的或不可执行的，本协议其它条款并不受任何影响且应维持有效。
11. 本协议根据中华人民共和国法律解释，受其管辖。
12. 任何因执行本协议所发生的，或与本协议有关的争议，应提交学校所在地人民法院管辖。

签名

指导教师： 

学生： 

日期： 2023.12.29

毕业设计开题报告

项目名称	3D 视觉识别及其在机器人抓取中的应用
赞助企业	深圳慧智星晨科技有限公司
<p>摘要： 对项目内容、设计规范、问题分析、项目计划以及项目成果的总结，限定在一页纸内。</p> <p>在当今医疗领域，智能化的机器人技术正逐渐成为提升工作效率和精确度的关键。特别是在处理如血液、基因和代谢产物等珍贵生物样本的存取任务中，这一技术的重要性更是显著。典型的生物样本库通常将这些样本保存在深低温条件下以保持其完整性和活性。在这种环境下，自动化设备成为了提高样本处理效率和减少人为错误的关键。然而，深低温环境为机器人抓取操作带来了独特的挑战。由于反复的移动和定位，冻存盒表面逐渐结霜，导致盒体轻微滑动和孔位识别困难，增加了夹取失败的风险。为解决此问题，本项目提出了一套创新的低温视觉辅助挑管系统，旨在提高挑管系统的视觉定位精度，从而确保更加精准和安全的操作。</p> <p>项目的核心在于应用深度学习技术，利用卷积神经网络进行物体识别、图像分割以及最佳抓取位姿的预测。深度学习算法通过多层非线性变换提取图像中的复杂特征，从而大幅提升目标检测的准确性。在普通情况下，机械臂能通过 3 轴移动稳定地运送样本管。但在异常情况下，系统需要通过机械臂视觉系统来检测样本管的具体位姿，并据此调整抓取和放置动作。这要求机械臂具备 6 自由度的动作能力，以及基于环境理解的智能抓取功能。</p> <p>此外，为了进一步提升抓取操作的成功率和安全性，项目计划在末端执行器上加装传感器，并结合基于力的强化学习算法。通过实时的力反馈数据，系统能够不断学习并优化其抓取策略，从而在动态和复杂的环境中实现更高效和安全的操作。</p> <p>综上所述，本项目旨在通过结合先进的深度学习技术和力反馈强化学习，为低温环境下的生物样本库自动化存取提供一种创新解决方案。该方案不仅提升了操作的精度和效率，还增强了系统的鲁棒性和安全性，对未来医疗领域的自动化设备发展具有重要的指导意义。</p>	

问题描述和简介（包含文献综述）：

描述项目背景、项目内容以及预期成果。描述现有相关产品、技术的发展及应用情况，完成相关文献综述，描述项目的创新性。当引用外部资源时，应对所引用的文献、技术基准标注来源。

1. 项目背景

生物样本库通常用来保存血液、基因、代谢产物等珍贵的生物样本。生物样本保存在冻存盒内的冻存管中，并整盒存放于 -80° 甚至 -196° 的深低温环境中。为提高存取效率，当用户需要提取样本时，由自动化设备将存有目标样本的冻存盒取出，并使用机械臂夹取目标样本到空冻存盒中完成样本出库，并在 -30° 环境下使用挑管夹爪夹取目标样本到空冻存盒中完成样本出库。深低温保存条件下，冻存盒体表面会逐渐结霜。在零度以下的工作环境中反复移动、定位和出入库的过程，会导致结霜表面逐渐增厚，盒体发生轻微滑动，盒底孔位结霜干扰识别。长期使用中会导致挑管时盒体定位精度下降，夹爪取放位置误差增大，造成挑管成功率下降，严重时可能损坏样本，造成重大的损失。另外，当出现样本管倾倒等异常情况时，难以通过孔位识别进行定位抓取，也会导致挑管失败。

因此，项目需要一套视觉辅助的机械臂挑管系统，能够实现精准的抓取与稳定的路径规划，同时在检测到样本管异常情况时控制机械臂调整样本管位姿并放置到目标位置。在正常工作环境下，机械臂 3 轴移动即可实现样本管的平稳运送。但在样本管异常放置情况下，倾斜样本管的位姿调整与移送要求机械臂实现 6 自由度动作，并能够基于环境理解进行智能抓取。

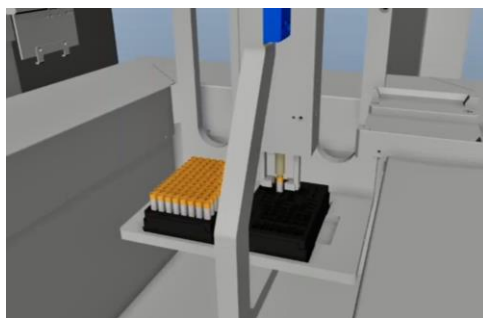


图 1 样本管出库拣选场景

2. 文献综述

针对将医药试剂取出并移放至试剂盒的这一应用场景，周硕研发了一种实用新型专利，基于机器视觉获取载料区域的二维图像，提取孔位灰度值等状态信息获得显著抓取目标，据此控制机械臂组件带动试管出入孔位^[1]。该装置能够自动

识别孔位中是否存在试管，并能够实现对试管的抓取和存放，但是，由于使用无分类的二维目标定位方法仅对孔位进行定位，没有试管的三维图像输入，当孔位中的试管出现斜倚、横躺等异常位姿时，机械臂根据孔位信息定位抓取，往往无法实现有效抓取。

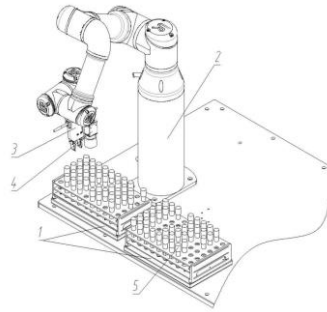


图 2 基于机器视觉的试管进出料装置^[1]

2.1 机器视觉模块

传统二维显著目标检测方法通过低级特征表示或颜色对比、背景先验等辅助方法，对输入图像中视觉最显著的目标区域进行定位和分割^[2]。然而，深低温环境下孔位结霜，孔位变形、边界模糊，传统的检测方法边界算法固定，泛化能力较差。自 2015 年以来，基于深度学习的目标检测方法有了更为优异的表现，通过将图像输入卷积神经网络实现物体识别、图像分割与最佳抓取位姿预测^[3-5]。深度学习通过一系列非线性变换，从原始数据中提取多层特征并进行表示。深度学习网络对接收的图像进行多层卷积、池化操作，能够得到类别、坐标等信息，极大程度提高了目标检测的准确性。

2.1.1 图像分割深度学习算法

图像分割在计算机视觉中起着至关重要的作用，是计算机视觉最关键的领域之一，它也是物体识别、医学成像、自动驾驶和图像编辑等广泛应用的基础。图像分割的主要目标是提取连贯的和有意义的部分或区域，为实现这一目标，多年来发展了大量的技术，包括基于传统的图像学方法，如阈值分割、边缘检测和区域增长，以及最近在深度学习中利用包括卷积神经网络（CNN）或深度学习模型（GAN、Transformer）。

(1) 卷积神经网络 CNN

卷积神经网络（CNN）是具有三维神经元排列的神经网络，包括宽度、高度、深度 3 个维度，其结构可分为以下几部分：卷积层、激活函数、池化层和全连接

层。其中卷积层产生了主要的计算量，也被称作 CNN 的核心，它相当于图像处理中的滤波器，能够捕捉到局部的特征点，例如图像的边缘或斑点等。卷积层的输出要经过一个激活函数，让输出转换为非线性，同时允许模型学习更复杂的特征。池化层中的下采样用于降低输出的维度，防止过拟合，也能够减少参数量。经典的 CNN 网络中经过两次卷积、池化后的输出进入全连接层，在全连接层中映射到样本标记空间，其中每一层相当于序列神经元的平铺。

2012 年基于 CNN 改进的 AlexNet 成功面世并引起了巨大的反响^[6]。在 AlexNet 网络中包含 5 个卷积层、3 个池化层和 2 个全连接层，它的网络更加复杂，但也意味着它能更准确地处理图像。由此，CNN 在图像分割中具有出色特征提取和表达能力被人们发现。CNN 在语义分割模型中的应用有着巨大的多样性，为了克服 CNN 具有固有归纳偏差、分割利用率低等限制，发展出了如 R-CNN、FCN、UNet、DeepLab 等经典模型，这些模型在相当长的一段时间里占据了语义分割模型的重要地位。AliF. Khalifa 等人在阐述图像分割在医学图像领域中的重要性的同时重点介绍了 CNN 的架构改进，针对于 CNN 的大量图像分割技术仍然在不断涌现^[7]。

Huaijia Lin 等人提出了一种新的范式——Propose-Reduce 和一个 Seq Mask R-CNN 框架，通过一步生成输入视频的完整序列，他们进一步在现有的图像级实例分割网络上建立了一个序列传播头来进行长期传播^[8]。Seq Mask R-CNN 通过在 Mask R-CNN 上增加一个额外的序列传播头，建立帧间的时间关系。Hongliang He 等人提出了一种新的核实例分割的向心方向网络（CDNet）^[9]。具体来说，他们将向心方向特征定义为一类指向核中心的相邻方向，来表示核内像素之间的空间关系。为了使网络更加轻量化，研究人员引入了超体素的概念，即将具有相似特征的体素云聚类来减少三维空间的计算量，Shi-Sheng Huang 等人探索了一种基于超体素的深度学习解决方案，形成一个有效的基于超体素的卷积神经网络

（Supervoxel-CNN）^[10]。Supervoxel-CNN 可以有效地融合在线三维重建过程中投影在超体素上的多视图 2D 特征和 3D 特征，期望在效率和分割精度之间取得平衡。Guoping Xu 等人提出了一种将 LeViT Transformer 模块集成到 UNet 架构中的 LeViT-Unet，用于进行快速而准确的医学图像分割，可以有效地重用特征映射的空间信息并平衡了分割的准确性和效率^[11]。Jinming Cao 等人介绍了一个形状感知

卷积层 (ShapeConv) 处理深度特征, 深度特征首先分解成形状组件和基本组件, 下两个可学习的权重与他们独立合作, 最后一个卷积应用于这两个组件的重加权组合^[12]。CNN 网络对图像特征的有效提取在以上工作中起了关键支撑作用, 非线性的处理能力让 CNN 在不同的维度中提取图像特征时显得非常优越。

(2) 生成对抗网络 GAN

生成对抗网络 (GAN) 属于生成式模型, 该类模型在无监督深度学习方面占据着重要位置^[13]。GAN 的提出是受到博弈论中的“零和博弈”的启发, 它通过框架中的生成模型网络和鉴别模型网络不断互相博弈学习来获得更好的分布收敛, 二者一方进行伪造一方进行侦辨, 使得模型在这种博弈过程中互相训练, 两个模型不断竞争对抗, 直到达到纳什均衡。虽然 GAN 能够在自我博弈中获取较高的性能, 但像素间的关系易被忽略而造成特征的缺失, 分割结果也可能因此而不够连续或产生较大形变。为了实现提升 GAN 进行图像分割的性能、补全全局关系、加强卷积分割网络等目标, GAN 的变体 SegAN、SCAN、PAN 等网络模型应运而生。

X. He 等人提出了一种基于 GAN 的分割方法 GANSeg, 它可以生成基于潜在掩膜条件的图像, 减轻了现有的自动编码网络中对图像对和预定义图像变换的依赖, 从而减轻了以往方法所需要的完整或弱注释^[14]。Dahye Kim 等人提出了一种无监督图像分解算法, 以获得一种基于乘法图像模型的内在表示, 该表示对不良偏差场具有良好的鲁棒性^[15]。分割模型进一步设计为在 GAN 框架中加入几何约束, 使分割函数分布与先验形状分布之间的差异最小化。

(3) Transformer

深度架构的改进已经成为一个研究热点, Transformer 也逐渐成为基于深度学习的图像分割方法中的主流架构。Transformer 最早由 Vaswanid 等人提出, 由于其编解码器的独特设计, 能够轻松处理不确定输入、捕获长距依赖关系, Transformer 也因此逐渐成为自然语言处理等人工智能领域的主要深度学习模型^[16]。

原始的 Transformer 由编码器和解码器组成, 在编解码器中都使用了自注意力机制加上前馈神经网络的结构, 在 Transformer 中还使用了多层的自注意力机制构成的多头注意力机制, 提升了模型对不同信息之间相关度的捕获, 这种特性很有利于分析上下文信息。Robin Strudel 等人提出的 Segmenter 是一种用于语义分割的

Transformer 模型，与基于卷积的方法相比，它允许在第一层和整个网络中对全局上下文进行建模^[17]。而 SegFormer 反思了 Segformer 其中出现的一些问题，比如需要位置编码却失去了图像大小的宽容性，在大尺度训练中这一点会成为计算复杂度上较大的障碍，SegFormer 则做到了架构轻量化和特征图层次化^[18]。

M. Pu 等人提出了一种基于 Transformer 的边缘检测器，边缘检测变压器 (EDTER)，同时利用全图像上下文信息和详细的局部线索，提取清晰的物体边界^[19]。H. Liu 等人提出了一种创新的基于 Transformer 的框架，即基于高级特征和低级线索来识别虚线曲线^[20]。这两个工作都聚焦于对象边界检测和分割问题，在这类问题中，被分割目标的边界通常由背景的混合组成，而在其他情况下，分割目标是整个图像中的小部分。此外，边界可以表现出可变性和小尺度特征，这些特征导致目标边界检测和分割的精度较低。Transformer 上下文依赖的特点能够很好的将边界特征利用起来，取得很好的分割效果。Z. Li 等将 Transformer 与高分辨率 CNN 特征图相结合，以实现精确的图像补丁定位^[21]。X. Huang 等也借鉴了 CNN 的思路，基于 U 型结构重新设计了 Transformer 块中的前馈网络，命名为 ReMix-FFN，通过重新整合局部上下文和全局依赖关系来探索全局依赖关系和局部上下文，从而更好地识别特征^[22]。

2.1.2 机械臂抓取姿态估计

在机器人抓取任务中，准确的抓取姿态是稳定抓取的重要因素。抓取姿态估计的目的是对待抓取物体预测出一个有效的抓取姿态，以提供给机器人控制单元对机器人夹爪进行合理控制。抓取估计方法主要可分为由模型机理驱动的解析法和由数据驱动的经验法。传统由模型机理驱动的解析法主要通过建立抓取操作的物理模型，分析夹爪和物体之间相互作用的空间以生成最优的抓取姿态，具体的可以通过力封闭或形封闭来评价不同抓取姿态的稳定程度^[23]。虽然解析法可以通过严格的物理模型表述抓取姿态的本质，但是由于这种方法过度依赖假设条件和简化模型，并且解析法需要提前知道物体的三维模型，导致在抓取姿态估计中只能对指定的目标进行抓取姿态估计，不适应于对未知物体的抓取姿态估计。由数据驱动的经验法，目前主要利用深度学习技术构建判别模型或拟合映射函数并在大量已经标注好的物体抓取姿态数据集上进行训练，使得模型可以逼近人类的抓取姿态行为。

近年来,随着低成本的深度相机和深度学习技术的不断成熟,推动了深度学习在抓取姿态估计领域的研究。Lenz 等人首次将深度学习引入抓取姿态估计模型中,利用滑动窗口穷举大量矩形框,并使用支持向量机对矩形框进行分类,置信度高的矩形框为抓取区域^[24]。Chu 等人借鉴两阶段目标检测算法,提出了抓取区域建议网络,利用 ResNet50 作为特征提取网络完成抓取姿态的预测^[25]。Redmon 等人提出单阶段抓取姿态估计模型,使用 AlexNet 为骨干网络提取整张图片的特征并进行全局抓取姿态估计,避免了采用滑动窗口方法出现不同尺寸图像块的重复计算问题^[26]。Kumra 等人在 Redmon 的研究基础上引入深度图像信息作为输入,提出一种多模态姿态估计模型,使用更加复杂的残差神经网络 ResNet50 代替 AlexNet 作为特征提取网络,并在原有的彩色三通道 (RGB) 特征提取网络上并行 ResNet50 网络用于对深度图像的特征提取,最后将两个模型提取的特征进行融合^[27]。Jefferey 等人提出抓取质量评估卷积神经网络 GQ-CNN,结合点云信息对抓取位置图像进行分割最终得到对应抓取点的概率,选择抓取概率最高的位置作为最终的抓取姿态^[28]。Morrison 等人提出一种仅使用深度图像进行抓取姿态估计的模型,并首次提出以每个像素为单位进行抓取姿态预测,预测速度得到明显提升^[29]。

在上述经典的基于深度学习的抓取姿态估计算法基础上,后人做出了大量的改进工作。H. Zhang 等为了缩小待抓取目标姿态检测的范围,提出一种基于感兴趣区域的抓取姿态估计模型,取代了之前在整张图像上进行抓取姿态估计的方法,完成了在单目标场景和多目标场景下的抓取姿态估计^[30]。张立洲利用 Faster R-CNN 对待抓取目标进行检测确定目标的位置,由获取的位置信息对输入的 RGB 图片和深度图片进行裁剪,随后分别送入利用 AlexNet 构建的抓取姿态检测网络,从而预测出场景中所有目标的抓取姿态^[31]。李秀智等在 Chu 等人^[25]提出的抓取姿态估计基础上,使用 MobileNetV2 作为主干网络设计了轻量级多目标抓取姿态估计模型,为了提高预测准确率,利用改进 Yolov3 回归的目标框与抓取姿态框做交并比得到最终的抓取姿态^[32]。李正明等在 Faster R-CNN 的基础上加入 InceptionV2 网络模型,实现了对不同尺度、不同种类以及多个目标的快速姿态估计和定位^[33]。孙先涛等提出一种基于 VGG16 网络的最小面积外接矩形算法,解决了姿态估计角度存在的离散性问题^[34]。张云洲等利用深度数据替换 RGB 图像的 B 通道,生

成 RGD 图像,并通过浅层特征和深层特征融合提高网络模型对尺度多变物体的适应能力^[35]。俞松提出一种基于 Yolo 神经网络的抓取姿态估计模型,引入旋转锚框机制,以适应任意角度摆放物体的抓取姿态预测,并使用 Ghost 模块对 EfficientNet 网络进行重构以减少计算量^[36]。郝宸等在 GG-CNN 研究的基础上,设计了一种新型注意力机制,并将其应用于 GG-CNN 网络中,使得预测准确率得到明显提升^[37]。H. Cao 等人提出了一种轻量级的抓取姿态检测模型,利用高斯核的抓取表示对训练样本进行编码以体现中心点抓取置信度最大的原则,同时通过在抓取检测的瓶颈处加入感知域模块提高特征的识别力^[38]。

2.2 机械臂运动模块

正常工作时,机械臂基于 2D 孔位图像信息,进行 3 轴移动运送样本管;在样本管异常放置情况下,需要通过机械臂上配备的 3D 视觉系统检测样本管位姿并由机械臂进行 6 自由度动作完成抓取与归置,基于环境理解进行智能抓取。同时,为减少夹取失败样本损坏的情况,需要考虑额外的安全措施。

2.2.1 机械臂抓取方法

在结构化的工业操作环境中,使用机器人抓取工件的传统方式是通过示教编程,人为设定机器人的运动轨迹与模式^[39]。由于机器人运动固定,当工件位姿或机器人起始位置发生一定变化时抓取就会失败。近年来,由于深度学习在图像识别与检测领域展露出的优异表现,机器人抓取任务中也逐渐采用了深度学习方法。Lenz 等提出在包含对象场景的 RGB-D 视图中应用深度学习算法检测机器人抓取从而避免人工进行特征设计,同时能够生成机器人抓取点位姿^[24]。

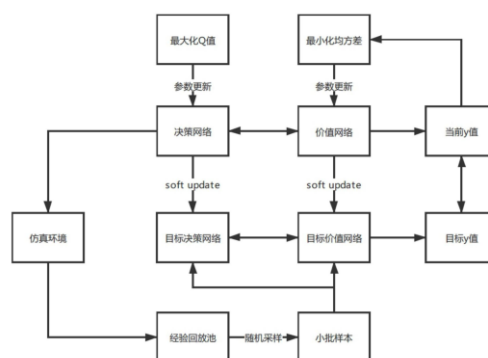


图 3 机械臂深度强化学习框架^[40]

为减少创建抓取数据集的人力投入,许多学者开始研究基于深度强化学习算法的机器人控制。L. Pinto 和 A. Gupta 通过控制两个机器人手臂训练卷积神经网络

络, 收集了超过 50K 的数据点, 从而预测出输入图像区域的最优抓握方向^[41]。上海交通大学的王贺升等也根据环境的三维点云状态特征, 对两台机械臂进行对抗式深度强化学习训练, 得到高鲁棒性的机器人操作系统^[40]。其强化学习框架如图 3 所示。

强化学习是一种基于环境交互的机器学习范式, 侧重于在多样环境下寻求解决问题的多种策略。智能体不断与环境进行交互并收集样本, 通过样本学习最优策略以实现累积回报的最大化。智能体与环境的交互通常被建模为马尔可夫过程 (markov decision process, MDP) ^[42]。一个 MDP 可由五元组 $\langle S, A, P, R, \gamma \rangle$ 表示, 其中 S 为状态空间, A 为智能体动作空间, P 为状态转移函数, R 为奖励函数空间, γ 为折扣因子。智能体为环境选择动作, 环境接收后状态变化, 并将产生的奖励与惩罚等强化信息反馈给智能体, 智能体基于强化信息和环境状态按照一定策略选择下一个动作, 重复以上过程, 利用奖励不断调整动作直至结束。通过将深度强化学习应用于机器人控制, 机器人能够获得更多的有效学习样本, 提高学习速度, 缩短了训练时间。

针对物体密集的抓取环境, 单一抓取动作难以很好地完成抓取任务。许多研究者提出让机器人执行推动等动作来分离密集物体, 采用推动-抓取协同系统以提高抓取成功率^[43-45]。I. Sarantopoulos 提出了 Split DQN 方法, 分别用两个 Q 网络学习推动目标物体和推动障碍物体的任务, 并在 1000 个随机场景测试集中运行测试推送任务。Split DQN 算法的成功率达到 88.6%, 能够有效地在密集场景中分离物体^[46]。

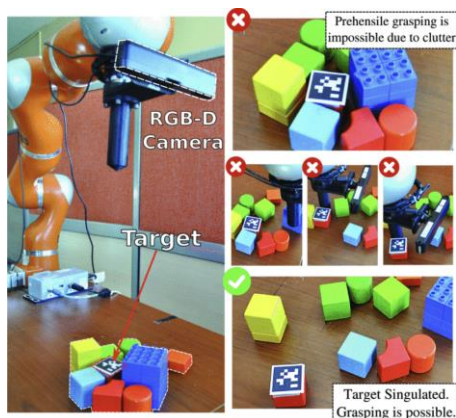


图 4 基于 Split DQN 算法分离密集物体^[46]

A. Zeng 等提出了一个无模型深度 Q 网络算法学习框架, 在该框架中训练两

个全卷积网络，分别推断推动和抓取动作对末端执行器方向和位置的效用。将相机处理后得到的 RGB-D 高度图输入全卷积网络中，对网络输出的像素级 Q 值图进行 Q 值预测，选择 Q 值最高的位置执行推抓动作，在仿真环境和真实实验场景中，对比仅使用抓取动作的全卷积网络，抓取成功率均有提升^[47]。

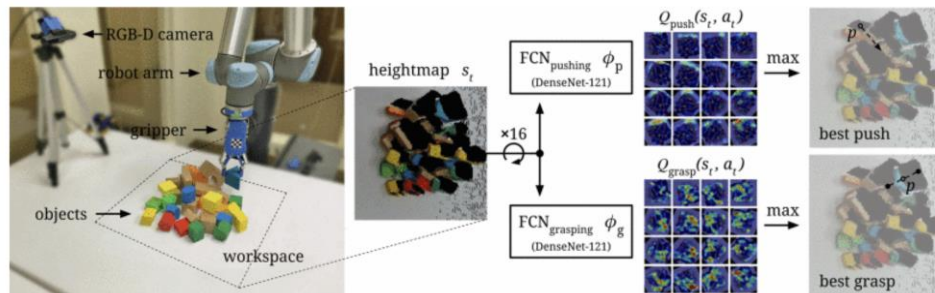


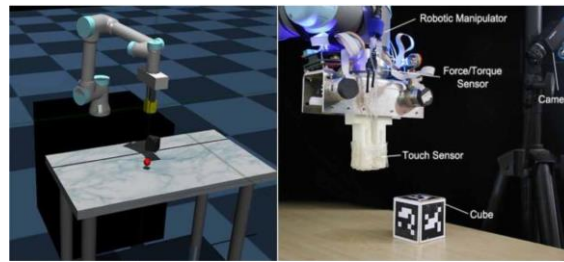
图 5 基于无模型深度 Q 网络算法的机械臂推抓协同系统^[47]

2.2.2 机械臂抓取策略优化

虽然将强化学习应用于机器人控制能够提供更通用而高效的策略，但根据环境反馈自我修正的过程也降低了机器人工作的稳定性与可靠性。针对在机器人上应用强化学习的安全问题，J. Garcia 和 F. Fernández 提出引入 PI-SRL 算法，定义风险函数并设置状态行动空间的基线行为^[48]。S. P. Coraluppi 和 S. I. Marcus 提出在马尔可夫决策过程中建立风险敏感标准从而在奖励设置中规避风险^[49]。然而，在执行抓取的动态环境中，这些策略可能并不适用。

在协作机器人领域，常通过碰撞检测与相应的反应策略来保证人机交互的安全性^[50]。但类似的力控制策略在基于强化学习的机械臂控制中较少被使用。Lin N 等提出将力和触觉感知引入强化学习中。强化学习的算法实现是将深度确定性策略梯度算法（Deep Deterministic Policy Gradient, DDPG）和 Hindsight Experience Replay（HER）相结合。DDPG 的基本框架是演员-评论家（Actor-Critic）算法，使用双重神经网络架构，加快了收敛速度^[51]。HER 算法对于机械臂抓取任务奖励稀疏的情况，设置 replay buffer 经验池，学习失败经验，重塑稀疏奖励表达式，从而提高训练速度和效率^[52]。Lin N 等在机械臂上加装力/扭矩和触觉传感器来感知外部环境，通过离散化的力数据判断机械臂是否与待抓取物体接触，以及是否与障碍物发生碰撞，并相应给予奖励或纠正性反馈。在与没有力反馈控制的机械臂进行推送物体的对比实验中，具有力反馈控制的实验组在模拟和真实实验环境中的任务成功率都远高于另一组，表明基于力的强化学习算法能够优化机械臂控

制策略^[53]。



(a)模拟环境

(b)真实环境

图 6 基于力的强化学习算法控制机械臂进行实验^[53]

2.3 仿真实验平台模块

机器人仿真在机器人设计和研究过程中起重要作用，通过仿真实验来验证机器人的各种性能和特点，已经是机器人理论研究的必备方法^[54]。目前，国内外较通用的机器人仿真软件有 MATLAB、ADAMS、V-REP 以及 ROS 系统等。MATLAB 具有强大的矩阵计算能力，能够进行机器人的运动学、动力学、轨迹规划等仿真，但可视化性能较差。ADAMS 在机器人虚拟样机领域具有快速简便建模、交互式仿真和动画等功能^[55]，但不适合本科生使用。ROS 是一个开源的机器人操作系统，提供了很多机器人的硬件抽象和常用功能，使得机器人的开发更加便捷^[56]。

在医疗样本库管理中，确保机械臂系统在复杂冷冻环境下高效、稳定地抓取和归置样本管是至关重要的。由于样本库操作的复杂性和特殊性，仿真实验平台成为研究人员优化机械臂控制和视觉系统的利器。通过在仿真环境中进行大量实验，研究人员能够更好地理解机械臂在不同情景下的表现，从而为医疗样本库管理带来新的解决方案。为贴合本项目的背景，我们查阅了相关文献，并根据一个使用 ROS2 和多个仿真器进行机器人抓取物品的模拟实验找到了以下几种较为常见的仿真平台^[57]。

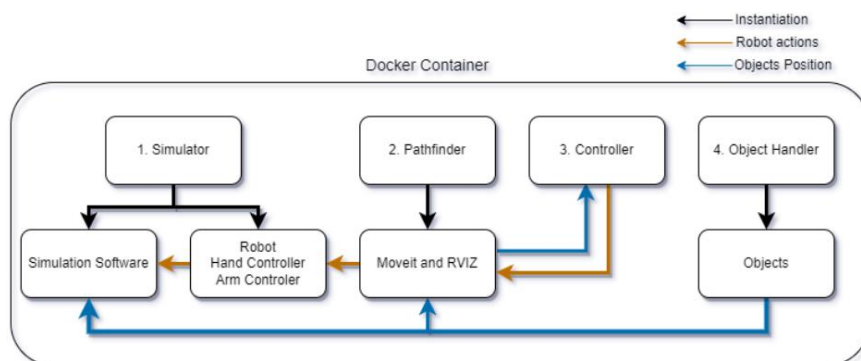


图 7 多个仿真器实验流程^[57]

首先，NVIDIA Isaac Sim 以其基于 Unreal Engine 的仿真环境而脱颖而出，其能够提供逼真的场景模拟，模拟真实感观察和复杂任务场景。多机器人协同工作的支持以及与 ROS 的紧密集成，使得该平台下仿真实验更加贴近实际场景。故而 Isaac Sim 在需要高度真实感观察和多机器人协同工作的医疗机械臂操作仿真中具备显著优势^[58]。

其次，V-REP (Virtual Robot Experimentation Platform) 作为一款灵活可扩展的仿真平台，通过提供多样的机器人和传感器模型，使研究人员能够在仿真环境中验证不同的机械臂抓取策略。其丰富的 API 支持允许研究人员进行算法开发，并通过实验评估各种视觉定位和机械臂运动策略的性能^[59]。

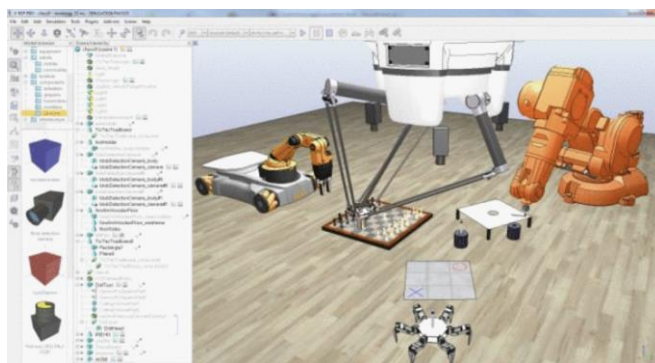


图 8 V-REP 仿真场景示例^[59]

此外，Gazebo 作为一款强大的开源 3D 机器人仿真平台，通过提供友好的用户界面和广泛的 API，使得研究人员能够轻松构建仿真场景。其与 ROS 的无缝集成为实验提供了更加灵活的控制和数据采集手段，使得研究人员能够更深入地了解机械臂在真实操作中的表现^[60]。

Webots 作为通用的机器人仿真软件，以其跨平台性和对多种编程语言的支持而脱颖而出。Webots 提供了友好的用户界面和广泛的 API，使得研究人员能够轻松构建仿真场景。其通用性使其成为研究人员探索不同机械臂抓取策略的理想选择，同时为整个仿真实验提供了强大的支持^[61]。

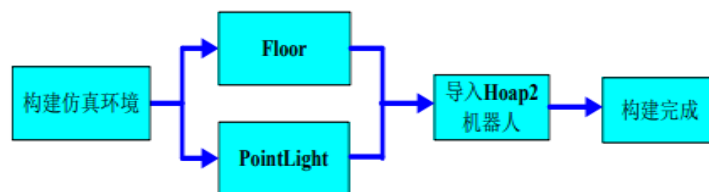


图 9 Webots 构建仿真环境的一般流程图^[61]

最后，MORSE (Modular OpenRobots Simulation Engine) 作为一个模块化的开源仿真引擎，专注于支持多机器人系统和传感器仿真。在仿真实验中，MORSE 的模块化设计允许用户根据需要定制仿真场景，以满足不同研究需求。研究人员通过 MORSE 能够在仿真环境中进行全面的机械臂抓取实验，评估不同策略在医疗样本库操作中的可行性和效果^[62]。

基于上述几种方法综合考虑后，Isaac Sim 在逼真性、ROS 集成、深度学习支持、传感器模拟以及强化学习方面的多重优势，我们选定 Isaac Sim 为我们项目虚拟实验的仿真软件。

在选择了 Isaac Sim 作为虚拟实验的仿真软件后，我们针对如何将在仿真环境中学到的策略和模型成功地迁移到真实环境开展了进一步的文献查阅，并选择了 Sim2Real 算法。

Sim2Real 是一种在虚拟环境中学到的模型或策略成功应用于真实环境的方法。由于提供了在真实环境中进行实验之前，在仿真环境中测试和优化算法的机会，该方法在机器学习和机器人领域变得越来越重要^[63]。一种常见的 Sim2Real 方法是通过在仿真中生成大量的合成数据，然后使用这些数据进行训练。这些合成数据可以涉及到虚拟相机图像、传感器数据以及机器人的运动轨迹。深度学习算法在这些合成数据上进行训练，学到的模型被认为可以在真实环境中泛化^[64]。

然而，虚拟环境和真实环境之间存在许多差异，这导致 Sim2Real 的迁移并非总是顺利的。这些差异可能涉及到感知数据的噪音、物理特性的不同、光照条件的变化等。因此，研究者们提出了一系列方法来提高 Sim2Real 的性能。

一种常见的方法是领域自适应 (Domain Adaptation)，通过在虚拟环境和真实环境之间建立映射，使得在虚拟环境中学到的知识可以更好地适应真实环境^[58]。领域自适应方法可以通过对抗训练、数据增强等技术来实现。

另一种方法是通过使用增强学习算法，使得机器人可以在真实环境中不断地与环境交互、学习和适应^[65]。这种方法通常需要更多的实际实验数据，但可以更好地应对真实环境的复杂性。

综合考虑 Sim2Real 的研究进展，选择 Isaac Sim 作为仿真软件有助于克服虚拟环境和真实环境之间的差异。其逼真性、ROS 集成、深度学习支持以及在传感器模拟和强化学习方面的优势提供了更好的 Sim2Real 迁移性能，并有助于更有效

地将虚拟实验的结果应用于真实世界的机器人系统。

3. 预期成果

基于上述背景，针对深低温环境中传统定位方法难以适应高精度的机械臂工作需求与样本管异常情况处理的问题^[2, 66]，本项目拟实现一套低温视觉辅助挑管系统，以 3D 视觉系统辅助机械臂实现抓取。控制系统包括基于深度学习的图像分割与姿态估计算法、基于视觉的机械臂轨迹规划与持续学习，以及机械臂仿真实验平台。聚焦于从冻存盒中存取试剂的应用场景，本项目创新性地引入深度学习算法增加视觉系统识别与定位的精度，并通过力反馈的强化学习增加抓取成功率与安全性。

项目需求与设计规范书：

描述项目需求，以及将需求转化为设计规范的详细步骤，说明项目目标的确定过程。

为了描述客户需求以及转化成设计规范，经过小组讨论与头脑风暴，搭建出本产品的 QFDI 质量屋如下图，搭建流程：

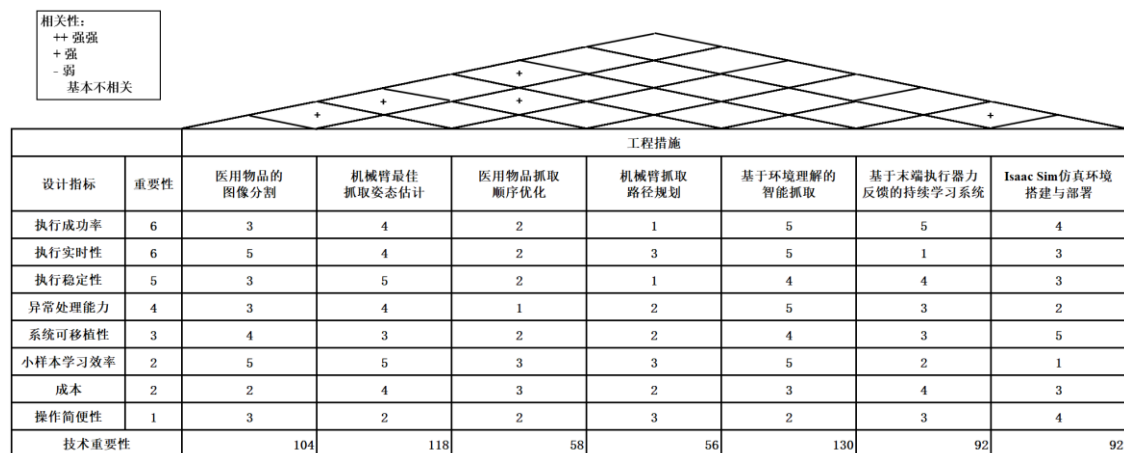


图 10 QFDI 质量屋

1. 调查客户要求(VOC)，转化成相应的质量要求，画出质量展开表，并量化其重要度。---建左墙

经过调研和集体讨论，我们列出 9 种设计指标，分别是执行成功率、执行实时性、执行稳定性、异常处理能力、系统可移植性、小样本学习效率、成本、操作简便性，并根据重要性依次评分，重要性评分标准如下：

- 1 分：不影响功能实现的需求；
- 2 分：不影响主要功能的需求；
- 3 分：表示比较重要地影响功能实现的需求；
- 4 分：表示重要地影响功能实现的需求；
- 5 分：表示基本的,涉及安全的,特别重要的需求。

2. 客户要求(VOC)和需求的质量，转换成相应的工程措施,用符号或数据表明相互之间的关系程度。--同步建天花板房间

为了避免限制产品设计方案，影响创造力的发挥，我们从产品整体出发，得出了相互独立的七项工程措施，分别是医用物品的图像分割、机械臂最佳抓取姿态估计、医用物品抓取顺序优化、机械臂抓取路径规划、基于环境理解的智能抓取、基于末端执行器力反馈的持续学习系统、Issac Sim 仿真环境搭建与部署。并且工程措施的组合是全面系统的，可据此生成完整的设计方案。

质量屋的房间代表各客户需求和设计要求间的相互关系，通过矩阵图将相互关系用符号或数字表示出来。

标准如下：

- (1) 若某项客户质量需求与所有设计要求间关系是空白，则应重新评估，或可增加满足些客户需求的设计要求。若某项设计要求与所有客户质量需求的关系也是空白，应检查此设计要求是由哪一项推导出来，能否取消。
- (2) 若一项顾客需求与大多数设计要求都有很强的关系，应重新分析，可能是该需求是高一级的需求，此时应将其分成几个子项目，同样每一项设计要求与大多数客户需求有较强关系，也要分析，分成几个设计要求。
- (3) 客户需求与设计要求间的关系用数字表示（5分制，表示关系的密切程度）。

3. 确定各工程措施间的交互作用和关系，并用符号表示。---建屋顶

根据 QFD 图，我们可以得出我们的设计目标：

- (1) 尽可能平衡图像分割算法和姿态估计算法的实时性和准确性；
- (2) 环境理解和力反馈对机械臂抓取的稳定性和鲁棒性至关重要；
- (3) 仿真环境搭建与部署为工业应用提供数据支撑和辅助验证。

概念设计方案拟订：

概念设计方案初步拟订，可以描述多个设计方案。

根据前述的项目目标，本项目“3D 视觉识别及其在机器人抓取中的应用”的设计方案可以归纳为以下三个方面：

- (1) 通过综合应用 3D 视觉技术和深度学习算法，解决不同姿态的试剂管在复杂环境中的抓取问题；
- (2) 机械臂轨迹规划组件根据抓取姿态信息生成最优运动路径。机械臂控制系统基于强化学习，通过 3D 视觉输入和末端执行器上传感器测量的力反馈数据改善抓取策略；
- (3) 构建一个基于 Isaac Sim 的机器臂拣选任务仿真系统，实现仿真到现实算法的研发。

1. 3D 图像分割、姿态估计和环境理解在医用物品抓取中的应用

1.1 抓取物品的图像分割

本项目的应用场景为在深低温环境中操作生物样本，要求图像分割算法具有高度的准确性、稳定性和对环境变化（如结霜）的鲁棒性。以下是几种适合该应用场景的深度学习图像分割算法。

Mask R-CNN 是一种用于目标检测和实例分割的深度学习模型，于 2017 年由 He 等提出^[67]。Mask R-CNN 在 Faster R-CNN 的基础上进行扩展，引入了额外的分支网络，用于预测每个目标实例的掩码，在语义分割的基础上实现了实例分割。同时，Mask R-CNN 仍能够实现准确的目标检测。



图 11 Mask R-CNN 在 COCO 数据集的测试效果^[67]

Mask R-CNN 的整体结构由两个主要部分组成：共享的卷积特征提取网络和 3 个并行的分支网络。共享的特征提取网络通过卷积层和池化层提取图像特征；分支网络包括目标分类分支、边界框回归分支和掩码分支，分别用于预测目标的类别、边界框和精确掩码。Mask R-CNN 在多个视觉挑战数据集上取得了优异的

性能，被广泛应用于目标检测、实例分割等计算机视觉任务中。

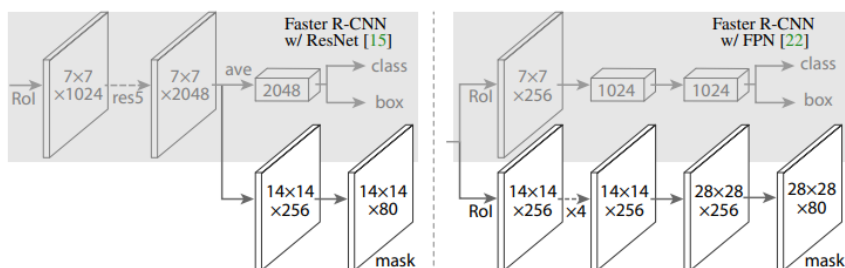


图 12 Mask R-CNN 的两种头部架构^[67]

DeepLab 系列模型是 Chen 等提出的深度卷积神经网络模型，其核心是空洞卷积的使用^[68]。空洞卷积在实时分割领域很受欢迎，包括 DeepLab 家族、密集上采样卷积和混合空洞卷积(DUC-HDC)、密集连接的 Atrous 空间金字塔池(DenseASPP)等。

DeepLabv1 使用了空洞卷积来增大感受野，以更好地捕捉图像中的上下文信息。DeepLabv2 引入了空洞空间金字塔池化(ASPP)模块，用于在多个尺度上提取特征^[69]。DeepLabv3 将扩张卷积的级联模块和并行模块相结合，改进了 ASPP，提出了更通用的框架，适用于更多网络^[70]。

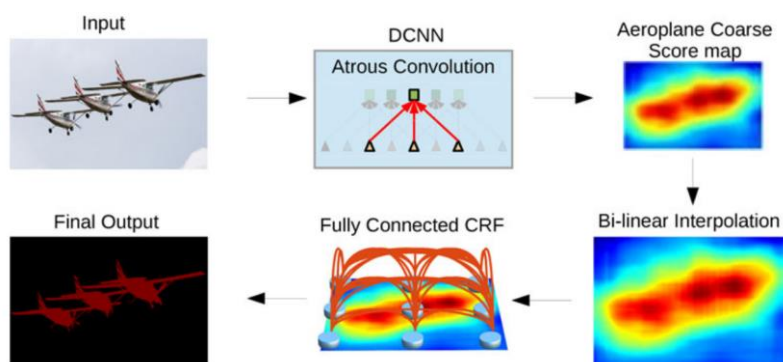


图 13 DeepLab 系列算法原理图^[68]

DeepLabv3+在 DeepLabv3 的基础上进行改进，将 DeepLabv3 作为编码器输入图像，同时添加一个解码器模块，这种编解码结构可以通过空洞卷积来平衡精度和运行时间。它使用改进的 Xception 作为主干，使用了深度可分离卷积，速度更快。

1.2 机械臂最佳抓取姿态估计

根据图像分割结果，系统需要进一步估计机械臂的最佳抓取姿态。在基于深度学习的物体抓取姿态估计领域，利用图像的高维特征与抓取位姿之间的关系可

回归出已知物体的 6D 姿态信息，并具备高精度和鲁棒性。

与大多数直接回归出旋转和平移矩阵的姿态估计算法不同，基于实例语义分割的 PVN3D 算法框架通过深度神经网络检测物体的三维关键点，然后以最小二乘拟合的方式估计 6D 姿态参数^[71]。

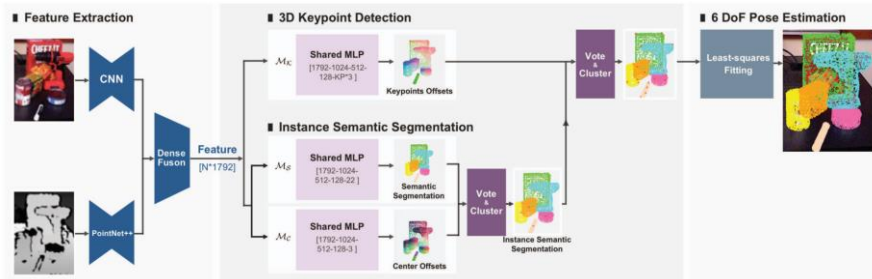


图 14 PVN3D 网络架构^[71]

RGB-D 图像作为输入，经由特征提取模块，学习到的特征通过 3D 关键点检测模块 M_k 获得每个点的偏移关键点。此外还有一个用于多对象处理的实例分割模块，其中语义分割模块 M_s 预测每点语义标签，中心投票模块 M_c 预测到对象中心的每点偏移。利用学习到的每点偏移量，聚类算法被应用于区分具有相同语义标签的不同实例，并且同一实例上的点为其目标关键点投票。最后，对预测的关键点应用最小二乘拟合算法计算得到 6D 姿态参数。因此，PVN3D 是一个两阶段方法，其核心为两大部分：一个是三维关键点检测模块，一个是 6D 姿态参数拟合模块。

另一种具有代表性的姿态估计算法是基于内部解耦的卷积神经网络 PoseCNN^[72]。由于旋转矩阵 R 和平移矩阵 T 这两个参数具有不同的视觉特性，因此在内部解耦 R 和 T 的估计具有可行性。

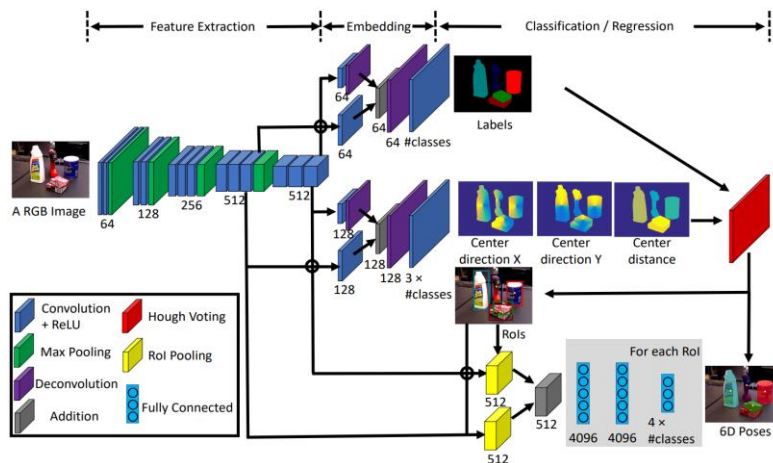


图 15 PoseCNN 网络架构^[72]

对于输入的一张 RGB-D 图片, PoseCNN 网络首先通过语义分割, 预测输入图像中每个像素的对象标签。其次, 它通过定位 2D 图像的目标中心并估计目标到相机的距离来估计三维平移量。为了找到目标的二维中心 c , 设计了一个 Hough 投票层并将其集成到网络中。投票层以像素级语义标注结果和中心回归结果作为输入。对于每个对象类, 它首先计算图像中每个位置的投票得分。投票得分表示对应的图像位置在类中的目标中心的可能性。最后, 利用 Hough 投票层预测的目标边界框, 利用两个 RoI 池化层对网络第一阶段生成的视觉特征进行剪裁和池化, 进行 3D 旋转回归。因此, PoseCNN 包含两个阶段: 特征提取模块、执行语义分割与回归任务。

1.3 医用物品抓取顺序的计算

确定合理的抓取顺序对于提高整体抓取效率和成功率具有至关重要的意义。考虑到各种医疗物品之间的空间布局和相互关联, 系统需要计算出一个最优的抓取顺序。这一过程包含复杂的逻辑和规划算法。

传统的方法包括启发式算法和基于规则的系统, 例如, 使用贪婪算法快速生成抓取顺序, 优先抓取最容易对其他物品造成遮挡的物品, 辅以根据任务紧急性来调整抓取顺序; 也可以设计一个基于优先级策略的机械臂抓取算法, 通过一个基于优先级的位置损失函数, 抑制多物体杂乱堆叠场景下底层物品的特征表达, 使网络的注意力集中在顶层物体, 在提升抓取性能的同时获得更高的实时性能。

若以上列举的方法仍无法适应复杂的挑管环境, 也可利用机器学习和优化算法, 例如遗传算法^[73]和蚁群算法^[74]。这些方法通过模拟自然现象, 能够解决复杂的顺序问题和优化问题。

2. 基于视觉的机械臂轨迹规划与持续学习系统

2.1 自适应机械臂平台与路径规划

本项目中拟通过一套基于 NVIDIA Jetson 开发平台的非夕拂晓 7 轴自适应机器人实现样本管的抓取与归置。该机器人具有 7 个自由度, 重复定位精度在 $\pm 0.05\text{mm}$, 臂展 941mm, 力控能力强, 能够实现对样本管的精准抓取。

Grav 力控型夹爪是非夕机器人的配套产品, 最小力控为 0.5N, 可实现顺从抓取和接触式抓取, 尺寸是 $202\text{mm} \times 159\text{mm} \times 84\text{mm}$ 。由于冻存盒孔位密集, Grav 夹爪的夹持运动可能受到空间限制, 因此需要考虑装配定制的试管夹爪,

通过使用细长夹爪来缓解试管间空间局限的问题。后续项目开展过程将对两种夹爪实际工作效果进行确定。



图 16 非夕 7 轴自适应机器人



图 17 Grav 力控型夹爪

为提高机械臂在定位、抓取与归置样本管过程中的操作效率，机械臂控制系统需要考虑多种因素规划最优路径。目前在机械臂中较为广泛使用的路径规划算法是 RRT-connect 算法^[75]。原始的 RRT 算法以起点作为根节点构建一棵搜索树，通过随机采样增加叶子节点，以获得从起点位置到目标位置的路径。RRT-connect 算法则在此基础上，分别从起点和目标点同时扩展随机树，两棵树产生连接时路径规划成功，这种双向扩展加快了搜索速度，实现了在 2-3 秒内规划一次运动。规划出无碰撞几何路径后，机械臂导航通常会在本地对路径进行平滑优化。

但是，研究表明，轨迹优化除了平滑路径，已经实现了更广泛的应用^[76]。CUDA 加速库 cuRobo 就将机器人导航问题视为全局运动优化问题，利用大规模并行 GPU 加速运动规划和优化，在 30ms 内生成全局动作。CuRobo 库基于 CUDA 内核，采用 PyTorch 实现，将目标姿态和初始关节配置作为输入，通过运行并行逆运动学优化实例计算出目标关节配置，基于初始关节构型到目标关节构型的线性插值生成用于轨迹优化的种子。启动并行优化实例，从优化解决方案中选择最优轨迹，批处理优化求解器执行数次基于粒子的优化迭代后，将种子移入良好区

间，用基于 GPU 批处理的 L-BFGS 优化器快速收敛到最小值，从而比 SOTA 轨迹优化方法更快地生成更短的路径^[77]。CuRobo 在 NVIDIA Jetson 平台上运行效果优秀，因此本项目拟采用 cuRobo 库的算法对机械臂从定位、抓取样品管到放置进冻存盒的过程进行路径规划。

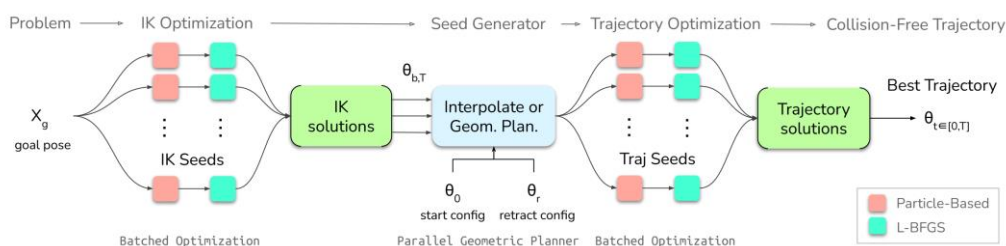


图 18 cuRobo 轨迹优化实现^[77]

2.2 基于环境理解的智能抓取系统

在样本管出现斜倚遮挡、横躺重叠等密集异常情况时，机械臂的单一抓取动作难以实现对样本管的成功抓取，机械臂推动-抓取协同系统的建立成为了该场景下的一种可行解决方案^[44]。

为了获得最优的推动和抓取位置，需要利用强化学习对机械臂进行智能控制，根据环境反馈的奖励值训练控制器，得到最优策略。强化学习算法中常使用状态-动作二元函数 $Q(s,a)$ ，用于评估在当前状态下采取动作所获得的累计期望奖励。将状态和动作空间组成一张 Q 表，每个状态和动作对应一个 Q 值，根据 Q 值在每个状态下选择最佳动作，从而最大化长期奖励，这就是 Q 学习算法的主要思想。深度 Q 网络算法 (deep Q -network, DQN) 则是将 Q 学习与神经网络相结合，使用一个深度神经网络 Q 网络近似表示 Q 函数^[78]。

在文献综述部分我们提到，A. Zeng 等提出了一个无模型深度 Q 网络算法学习框架来构建推抓协同系统，并在测试中有着较为优秀的抓取表现^[47]。本项目拟基于该深度 Q 学习框架，由控制系统持续接收 3D 视觉输入实时检测试管的异常放置状态，将检测图像输入全卷积网络中进行特征提取，将特征向量分别输入抓取网络和推动网络中进行像素级 Q 值预测，据此选择机械臂动作位置。本项目拟使用的全卷积神经网络是 G. Huang 等提出的 DenseNet 模型^[79]。对推动和抓取动作设置奖励函数，在仿真环境中训练机械臂使得机械臂能够在样本管放置异常的情况下成功抓取，并能够在实际操作反馈中不断迭代改进。

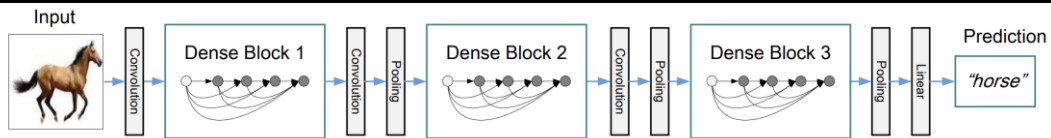


图 19 DenseNet 网络结构^[79]

2.3 基于末端执行器力反馈的持续学习系统

在仿真环境中通过强化学习训练机械臂时，由于机械臂学习的策略是高度非线性的，并且现实环境存在不可避免的扰动，需要通过末端执行器的力反馈优化持续学习系统，避免执行策略失控，提高抓取成功率。力控制策略常使用各种触觉传感器、测力传感器等测量所抓取物体的压力和力矩变化，监测机械臂是否成功抓取以及是否存在冲击、碰撞等异常情况。触觉传感器主要有电阻式、电容式、光电式、压电式等类型，其各自的优缺点如表 1 所示^[80]。由于触觉传感器装配在末端执行器上，体积应较小，故项目中可选用压电式或电阻式传感器检测末端执行器处所受压力，安装测力传感器在关节处来测量夹持物体的力矩。

表 1 触觉传感器各类型优缺点对比

触觉传感器类型	优点	缺点
电阻式	结构简单，动态范围宽	迟滞、单调响应非线性
电容式	灵敏度高，受温度影响小	体积较大，检测范围受限
光电式	不受电磁干扰	体积较大，存在滞后
压电式	灵敏度高，动态范围宽	不易检测静止的力

本项目中拟采用 Lin N 等提出的基于力控制的强化学习算法从而更好地检测抓取情况^[53]。将获得的力反馈数据离散化进行学习，在仿真环境和真实实验中进行训练与测试，设置奖励函数对动作进行纠正性反馈。机械臂控制算法能够通过样品管与末端执行器的接触情况识别抓取事件成功与否，当触觉传感器值为零时，机械臂应重新进行抓取动作；通过给关节扭矩设定合理阈值避免与冻存盒或其他样品管碰撞发生危险，如若产生碰撞则计为一次故障事件，重新调整抓取姿态。力反馈数据使系统能够动态地调整抓取力度和姿态，从而在抓取散乱样品管与放置进冻存盒的过程中均能实现稳定且可靠的抓取。

3. 机器臂拣选任务仿真系统搭建

在本项目中，我们致力于构建一个基于 NVIDIA Isaac Sim 的机器臂拣选任

务仿真系统，并实现仿真到现实（sim2real）算法的研发。该仿真系统的目标是通过在虚拟环境中模拟生物样本挑管任务，提供一个有效的研究平台，以验证和优化我们提出的视觉辅助挑管系统。

3.1 Isaac Sim 仿真环境搭建及云端部署

NVIDIA Omniverse™ Isaac Sim 是适用于 NVIDIA Omniverse™平台的机器人仿真工具包，具有构建虚拟机器人世界和实验的基本功能。它为研究人员和从业人员提供了创建稳健、物理精确的模拟和合成数据集所需的工具和工作流程。Isaac Sim 通过 ROS/ROS2 支持导航和操作应用程序，能模拟来自 RGB-D、激光雷达和 IMU 等传感器的传感器数据，用于各种计算机视觉技术，例如域随机化、地面实况标记、分割和边界框。

在本项目中，我们拟在 NVIDIA Isaac Sim 环境中搭建了一个复杂的多机器人仿真场景。该场景包括了高度逼真的工作环境、机器人操作臂、工作台和拣选目标物体。通过仔细构建和设定仿真场景，我们可以模拟真实世界中机器臂的操作流程，包括挑管、定位和异常处理等各个环节。为了确保仿真结果的准确性和稳定性，我们将系统部署在云端，以保证高强度的仿真计算在云端服务器集群上执行，以确保数据采集的效率和规模。

Isaac Sim 环境深度集成了 NVIDIA 的 Isaac SDK，为我们提供了先进的计算能力。这使得我们能够利用 GPU 进行仿真中的深度学习、路径规划和视觉感知任务，为视觉辅助挑管系统的算法开发提供强大的支持^[81]。

Isaac Sim 还包括最新的 PhysX，可用于进行快速，稳定的 GPU 物理模拟。提供了可扩展的功能和用户自定义的 UI，用于场景管理的 python 脚本，提供了用于构建虚拟机器人世界和实验的基本功能。

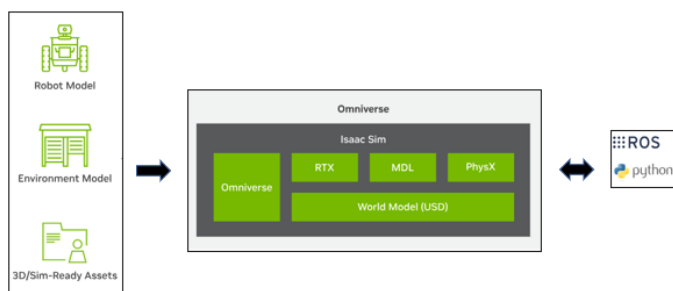


图 20 Isaac Sim 解决流程

由于 Isaac Sim 使用了实时光线追踪技术，如下图所示，仿真场景中的光照

及物体反射都十分真实，真实的环境可以为摄像头积累大量的有价值数据。由于是在仿真环境中，每个物体的位置和姿态都存储在计算机中，可以方便获取物体的数据，这样就可以得量带标签的数据，方便进行物体的检测、识别、分割、深度信息获取^[82]。此外，Isaac Sim 支持多种传感器的模拟，包括摄像头、激光雷达等。在我们的项目中，我们将使用这些模拟传感器来准确模拟在低温环境中挑选生物样本的视觉感知，为算法的训练和测试提供真实的感知输入。



图 21 实时光线追踪技术^[82]

Isaac Sim 不仅为仿真提供了强大的功能，还允许我们将将在仿真中验证的算法和模型直接部署到实际机器人平台中。这种无缝的集成能力使得我们能够在仿真中进行全面的测试和验证，然后将其应用于真实环境，最大程度地减少试错的时间和成本。

3.2 sim2real 算法开发与现实对齐

在视觉辅助挑管系统的 Sim2Real 算法中，我们着眼于有效地将仿真环境学到的参数和策略迁移到真实世界的机器臂操作。这一算法的核心目标是在仿真和真实环境之间建立一个稳健的转移桥梁，以确保在仿真中获得的结果能够在实际应用中具有较好的性能。对于操纵任务，该策略开始使用具有估计模型的粗糙虚拟环境进行模拟训练。然后，采用模拟到真实的迁移来比较虚拟机器人和真实机器人的性能。最后，根据模拟世界和现实世界之间的差异更新模拟中的模型^[64]。

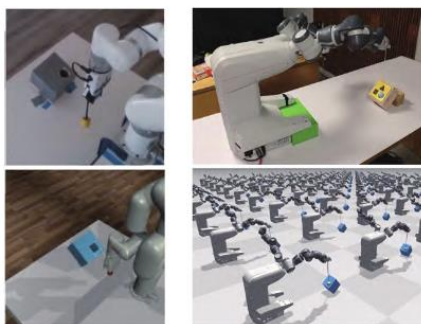


图 22 真实机械臂（上）与虚拟环境（下）^[64]

Sim2Real 算法首先涉及到对仿真中机器臂拣选任务的参数学习。我们使用先进的深度学习技术，如强化学习和深度神经网络，以在仿真环境中学习到机器臂在挑管过程中的控制策略。这包括了对挑管路径、抓取力度、以及异常情况处理等方面的学习。通过大量仿真实验，我们逐渐优化模型的参数，确保在仿真环境中的仿真机器臂表现出色。学习到的参数需要进行有效的迁移，以适应真实世界的机器臂。这一过程可能涉及到对模型的微调和校正，以适应潜在的仿真与真实环境之间的差异。这可能包括传感器误差、物体质地的差异、以及环境动力学等因素的考虑。

该算法中关键的一步是对视觉感知系统的适应性设计。在仿真中训练的视觉模型可能在真实环境中受到光照、遮挡等因素的干扰。因此，我们引入了对抗训练和领域自适应等技术，以提高模型对真实感知场景的鲁棒性^[83]。

Sim2Real 算法也是一个不断迭代的过程。在机器臂进行挑管任务时，系统会实时收集反馈信息，包括挑管的准确性、异常处理的效果等。这些反馈信息被用于在线调整模型参数，实现在真实场景中的自适应。通过这种方式，算法具备较强的动态调整和适应性，能够有效解决挑管系统在真实环境中面临的挑战。

项目进度计划及预算：

项目重要时间节点（用甘特图等方式表示项目计划），小组成员的工作量以及预算。

1. 项目时间节点

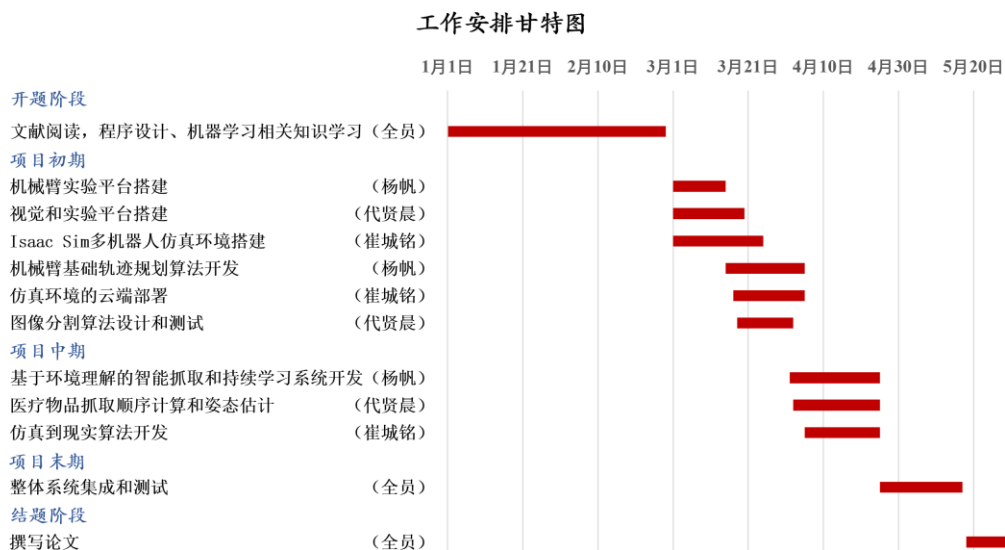


图 23 项目甘特图

2. 小组成员工作量

代贤晨部分主要负责医用物品抓取中 3D 视觉的图像分割及姿态估计和环境理解部分，具体包括以下内容：

(1) 抓取物品的图像分割

在复杂的医疗物品抓取任务中，对于各种形态和尺寸的医用物品（如疫苗管、试管等），利用 3D 视觉技术将这些物品精确地从背景中分割出来。这一过程需要考虑物体间遮挡、光照变化以及物体本身的反光材质等问题。通过对 3D 点云数据的深度学习处理，实现对各种医用物品和它们的空间位置的精确识别，从而为后续的抓取姿态估计和序列规划提供可靠的前处理数据。

(2) 机械臂最佳抓取姿态估计

根据图像分割结果，使用 3D 视觉技术构建目标物体的高精度立体模型，结合机器学习算法对抓取点的稳定性和可行性进行评估。抓取姿态涉及到物体的方向、机械臂的接近角度和路径规划等多个要素，采用优化算法来确保最终推导出的抓取姿态达到抓取成功率、运动时间和能耗的平衡。姿态估计模块必须实时工作，以适应动态变化的抓取环境和任务需求。

(3) 医疗物品抓取顺序的计算

考虑到各种医疗物品之间的空间布局和相互关联,系统需要计算出一个最优的抓取顺序以提高整体抓取效率和成功率。这一过程包含复杂的逻辑和规划算法,例如需要优先抓取会对后续操作造成遮挡的物品,或者根据任务紧急性来调整抓取顺序。适配边缘计算的算法需要考虑到计算资源和响应时间的限制,确保在有限的硬件资源下也能够快速给出最优解决方案。

杨帆部分主要针对机械臂的轨迹规划和持续学习,具体包含以下内容:

(1) 基于 3D 视觉信息的抓取/归置轨迹规划

针对由视觉系统所提供的抓取姿态信息和空间信息,机械臂轨迹规划组件负责生成最优的抓取和放置路径。轨迹规划算法需考虑多种因素,包括机械臂安全规范、运动学限制以及预期路径中的潜在障碍物。由于操作空间的高度约束性,算法必须能够适应各种抓取角度和位置,同时保持高效和精确。此外,由于物品的放置位置通常具有特定的空间要求,轨迹规划也需确保归置过程的精准性。

(2) 基于环境理解的智能抓取系统

在物体受到遮挡或重叠的情况下,3D 视觉合并算法来提升系统对于环境的感知能力。通过持续接收的视觉输入,机械臂控制系统学习如何重新排列物体,从而获得更好的视角以执行抓取任务。学习过程通过模拟和实际操作的反馈进行迭代改进,从而确保在不断变化的工作环境中仍能保持卓越的抓取表现。

(3) 开发基于末端执行器力反馈的持续学习系统

通过使用各种机械臂末端执行器上安装的传感器测量所抓取物体的力矩变化,机械臂控制算法能够根据这些数据识别并区分成功与失败的抓取事件,并且能动态地调整抓取力度和姿态,适应不同物体的重量和材质,从而在不同情形下均能实现稳定且可靠的抓取。

崔城铭部分主要负责机器臂拣选任务仿真系统搭建,具体内容为下:

(1) IsaacSim 多机器人仿真环境搭建以及云端部署

为有效支持机器臂拣选任务的仿真研究,本设计在 NVIDIA Isaac Sim 环境中搭建了一个复杂的多机器人仿真场景,包含了高度逼真的工作环境、机器人操作臂及其相应的工作台和拣选目标物体。为了保证仿真结果的准确性和稳定性,系统还将可在云端进行部署,这使得高强度的仿真计算可以在云端服务器集群上

执行，确保数据采集的效率与规模。

(2) 仿真到现实的算法开发和现实对齐

仿真到现实（sim2real）算法旨在将仿真环境中学习到的参数和策略有效迁移到真实世界的机器臂操作中。通过结果对比和模型调整，仿真中习得的视觉识别与控制策略将会不断优化，以补偿仿真与现实之间的潜在差异。此外，开发的仿真到现实传递算法将侧重于减少真实机器臂验证时所需要的数据量，加快算法的迭代速度，使得从仿真环境到真实操作环境的过渡更为平滑，加速机器臂应用的开发和部署。

3. 项目预算

表 2 项目预算

序号	项目模块	金额（元）
1	摄像设备	3000
2	软件许可	2500
3	硬件组件	2000
4	实验耗材	2500
总金额	10000	

结论:

总结报告中的重要内容

在本毕设项目中，我们致力于解决医用物品抓取和机械臂轨迹规划两个关键问题，通过综合应用 3D 视觉技术和深度学习算法，以及构建仿真环境，实现在复杂环境中的高效抓取和持续学习。本项目涉及的三个子课题分别为 3D 视觉图像分割、机械臂轨迹规划以及仿真环境搭建，其中每个子课题都涵盖了多个具体任务和评价标准。

首先，在医用物品抓取方面，我们需要通过 3D 视觉技术实现了医用物品的精确图像分割，同时考虑复杂的医疗环境中可能出现的挑战，如物体遮挡、光照变化等。通过深度学习处理 3D 点云数据，我们要实现对医用物品及其空间位置的准确识别，为后续的抓取姿态估计和序列规划提供可靠的前处理数据。抓取物品的图像分割不仅关注效率和执行速度，还注重成功率、稳定性以及可推广性。

其次，在机械臂轨迹规划方面，我们需要基于 3D 视觉信息实现抓取和归置轨迹的规划。考虑到机械臂操作空间的高度约束性，需设计轨迹规划算法以适应各种抓取角度和位置，并确保路径的高效和精确。此外，为了应对物体遮挡问题，需开发智能抓取系统，通过持续学习重新排列物体，不断优化抓取策略。末端执行器的力反馈数据被引入，以实现抓取成功率的最大化。

最后，在仿真环境搭建方面，我们计划构建基于 Isaac Sim 的机器臂拣选任务仿真系统，并致力于 sim2real 算法的研发。通过在云端进行仿真环境的构建，我们希望能够模拟复杂多机器人操作环境，同时通过优化仿真参数与真实操作的对齐算法，提高操作策略的迁移效率。仿真到现实的算法开发包括对仿真环境中学习到的参数和策略的迁移，以及减少真实验证所需数据量，加速算法的迭代速度。

综合而言，本项目通过解决医用物品抓取和机械臂轨迹规划中的多个关键问题，以及搭建仿真环境并开发 sim2real 算法，来为自动化领域中的机器人操作任务提供新的解决方案。通过整合 3D 视觉技术、深度学习算法和仿真技术，我们期望在自动化医疗场景中展示 3D 视觉技术的实用价值，并为机器人操作的智能化发展奠定坚实的基础。

参考文献:

列出完整的外部信息来源，并在文中标注参考文献。

- [1] 周硕. 基于机器视觉的试管进出料装置: 中国, CN219807434U [P]. 2023-10-10.
- [2] DU G, WANG K, LIAN S, et al. Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review [J]. *Artificial Intelligence Review*, 2021, 54(3): 1677-1734.
- [3] CALDERA S, RASSAU A, CHAI D. Review of Deep Learning Methods in Robotic Grasp Detection [J]. *Multimodal Technologies and Interaction*, 2018, 2(3): 57.
- [4] 贾松敏, 鞠增跃, 张国梁, 等. 一种基于深度学习的目标识别与抓取定位方法: 中国, CN108648233A [P]. 2018-10-12.
- [5] 周滢慇. 基于机器视觉的生活垃圾智能分拣系统的设计与实现 [D]. 哈尔滨: 哈尔滨工业大学, 2019.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Advances in neural information processing systems*, 2012, 25.
- [7] KHALIFA A F, BADR E. Deep Learning for Image Segmentation: A Focus on Medical Imaging [J]. *Comput Mater Contin*, 2023, 75(1): 1995-2024.
- [8] LIN H, WU R, LIU S, et al. Video instance segmentation with a propose-reduce paradigm [C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 1739-1748.
- [9] HE H, HUANG Z, DING Y, et al. Cdnet: Centripetal direction network for nuclear instance segmentation [C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 4026-4035.
- [10] HUANG S-S, MA Z-Y, MU T-J, et al. Supervoxel convolution for online 3d semantic segmentation [J]. *ACM Transactions on Graphics (TOG)*, 2021, 40(3): 1-15.
- [11] XU G, WU X, ZHANG X, et al. Levit-unet: Make faster encoders with transformer for medical image segmentation [J]. *arXiv preprint arXiv:210708623*, 2021.
- [12] CAO J, LENG H, LISCHINSKI D, et al. Shapeconv: Shape-aware convolutional layer for indoor rgb-d semantic segmentation [C]// *Proceedings of the IEEE/CVF international conference on computer vision*, 2021: 7088-7097.
- [13] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [14] HE X, WANDT B, RHODIN H. Ganseg: Learning to segment by unsupervised hierarchical image generation [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 1225-1235.
- [15] KIM D, HONG B-W. Unsupervised segmentation incorporating shape prior via generative adversarial networks [C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 7324-7334.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. *Advances in neural information processing systems*, 2017, 30.
- [17] STRUDEL R, GARCIA R, LAPTEV I, et al. Segformer: Transformer for semantic segmentation [C]// *Proceedings of the IEEE/CVF international conference on computer vision*, 2021: 7262-7272.
- [18] XIE E, WANG W, YU Z, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [J]. *Advances in Neural Information Processing Systems*, 2021,

- 34: 12077-12090.
- [19] PU M, HUANG Y, LIU Y, et al. Edter: Edge detection with transformer [C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022: 1402-1412.
- [20] LIU H, LI C, LIU X, et al. Neural recognition of dashed curves with gestalt law of continuity [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 1373-1382.
- [21] LI Z, WANG W, XIE E, et al. Panoptic segformer: Delving deeper into panoptic segmentation with transformers [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 1280-1289.
- [22] HUANG X, DENG Z, LI D, et al. Missformer: An effective transformer for 2d medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2022.
- [23] 肖传杰. 基于深度学习的机械臂抓取位姿检测方法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2020.
- [24] LENZ I, LEE H, SAXENA A. Deep learning for detecting robotic grasps [J]. The International Journal of Robotics Research, 2015, 34(4-5): 705-724.
- [25] CHU F-J, XU R, VELA P A. Real-world multiobject, multigrasp detection [J]. IEEE Robotics and Automation Letters, 2018, 3(4): 3355-3362.
- [26] REDMON J, ANGELOVA A. Real-time grasp detection using convolutional neural networks [C]// 2015 IEEE international conference on robotics and automation (ICRA): IEEE, 2015: 1316-1322.
- [27] KUMRA S, KANAN C. Robotic grasp detection using deep convolutional neural networks [C]// 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS): IEEE, 2017: 769-776.
- [28] 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]. arXiv preprint arXiv:170309312, 2017.
- [29] MORRISON D, CORKE P, LEITNER J. Learning robust, real-time, reactive robotic grasping [J]. The International journal of robotics research, 2020, 39(2-3): 183-201.
- [30] ZHANG H, LAN X, BAI S, et al. Roi-based robotic grasp detection for object overlapping scenes [C]// 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS): IEEE, 2019: 4768-4775.
- [31] 张立洲. 基于深度学习的目标识别及抓取定位研究 [D]. 天津: 天津科技大学, 2019.
- [32] 李秀智, 李家豪, 张祥银, 等. 基于深度学习的机器人最优抓取姿态检测方法 [J]. 仪器仪表学报, 2020, 41(05): 108-117.
- [33] 李正明, 章金龙. 基于深度学习的抓取目标姿态检测与定位 [J]. 信息与控制, 2020, 49(02): 147-153.
- [34] 孙先涛, 程伟, 陈文杰, 等. 基于深度学习的视觉检测及抓取方法 [J]. 北京航空航天大学学报, 2023, 49(10): 2635-2644.
- [35] 张云洲, 李奇, 曹赫, 等. 基于多层次特征的机械臂单阶段抓取位姿检测 [J]. 控制与决策, 2021, 36(08): 1815-1824.
- [36] 俞松. 基于轻量化深度网络的目标检测与物体抓取位置研究 [D]. 扬州: 扬州大学, 2022.
- [37] 郝宸, 田瑾, 韩华, 等. 基于注意力机制的实时性抓取检测算法 [J]. 传感器与微系统, 2022, 41(01): 131-134.
- [38] CAO H, CHEN G, LI Z, et al. Lightweight convolutional neural network with Gaussian-based grasping representation for robotic grasping detection [J]. arXiv preprint arXiv:210110226,

- 2021.
- [39] EKVALL S, KRAGIC D. Interactive grasp learning based on human demonstration [C]// IEEE International Conference on Robotics and Automation, 2004 Proceedings ICRA'04 2004: IEEE, 2004: 3519-3524.
- [40] 王贺升, 韩兴耀, 吴文华, 等. 基于对抗式深度强化学习的高鲁棒性机械臂操作方法及系统, CN113076615B [P]. 2022-07-15.
- [41] PINTO L, GUPTA A. Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours [C]// 2016 IEEE international conference on robotics and automation (ICRA): IEEE, 2016: 3406-3413.
- [42] BERTSEKAS D. Dynamic programming and optimal control: Volume I [M]. Nashua: Athena scientific, 2012: 56-62.
- [43] COSGUN A, HERMANS T, EMELI V, et al. Push planning for object placement on cluttered table surfaces [C]// 2011 IEEE/RSJ international conference on intelligent robots and systems: IEEE, 2011: 4627-4632.
- [44] DOGAR M R, SRINIVASA S S. A planning framework for non-prehensile manipulation under clutter and uncertainty [J]. Autonomous Robots, 2012, 33: 217-236.
- [45] CHANG L, SMITH J R, FOX D. Interactive singulation of objects from a pile [C]// 2012 IEEE International Conference on Robotics and Automation: IEEE, 2012: 3875-3882.
- [46] SARANTOPOULOS I, KIATOS M, DOULGERI Z, et al. Split deep q-learning for robust object singulation [C]// 2020 IEEE International Conference on Robotics and Automation (ICRA): IEEE, 2020: 6225-6231.
- [47] ZENG A, SONG S, WELKER S, et al. Learning synergies between pushing and grasping with self-supervised deep reinforcement learning [C]// 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS): IEEE, 2018: 4238-4245.
- [48] GARCIA J, FERNÁNDEZ F. Safe exploration of state and action spaces in reinforcement learning [J]. Journal of Artificial Intelligence Research, 2012, 45: 515-564.
- [49] CORALUPPI S P, MARCUS S I. Risk-sensitive and minimax control of discrete-time, finite-state Markov decision processes [J]. Automatica, 1999, 35(2): 301-309.
- [50] HADDADIN S, ALBU-SCHAFFER A, DE LUCA A, et al. Collision detection and reaction: A contribution to safe physical human-robot interaction [C]// 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems: IEEE, 2008: 3356-3363.
- [51] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [J]. arXiv preprint arXiv:150902971, 2015.
- [52] ANDRYCHOWICZ M, WOLSKI F, RAY A, et al. Hindsight experience replay [J]. Advances in neural information processing systems, 2017, 30.
- [53] LIN N, ZHANG L, CHEN Y, et al. Reinforcement learning for robotic safe control with force sensing [C]// 2019 WRC Symposium on Advanced Robotics and Automation (WRC SARA): IEEE, 2019: 148-153.
- [54] 李硕, 强华. 工业机器人实验教学平台 [J]. 实验技术与管理, 2018, 35(04): 166-170.
- [55] 赵川. 工业机器人轨迹规划及运动仿真系统研究 [D]. 汉中: 陕西理工大学, 2017.
- [56] 刘佩森, 靳杏子, 郑翔鹏, 等. 基于 ADAMS 的工业机器人建模与动力学仿真 [J]. 成都工业学院学报, 2018, 21(04): 10-13+59.
- [57] AUDONNET F P, HAMILTON A, ARAGON-CAMARASA G. A Systematic Comparison of Simulation Software for Robotic Arm Manipulation using ROS2 [C]// 2022 22nd International

- Conference on Control, Automation and Systems (ICCAS): IEEE, 2022: 755-762.
- [58] KIH MUHAMAD F, KIM J-S, PARK J-H. Learning Robust Perception Based Controller for Quadruiped Robot [J]. IEEE Access, 2023.
- [59] ROHMER E, SINGH S P, FREESE M. V-REP: A versatile and scalable robot simulation framework [C]// 2013 IEEE/RSJ international conference on intelligent robots and systems: IEEE, 2013: 1321-1326.
- [60] 秦吉军. 基于 GAZEBO 的半虚拟化机械臂及应用场景的设计 [D]. 青岛: 青岛科技大学, 2020.
- [61] 兰淼淼, 王英侠, 陈新新, 等. Webots 平台的仿人机器人仿真的研究 [J]. 福建电脑, 2020, 36(05): 5-8.
- [62] ECHEVERRIA G, LASSABE N, DEGROOTE A, et al. Modular open robots simulation engine: Morse [C]// 2011 IEEE International Conference on Robotics and Automation: IEEE, 2011: 46-51.
- [63] HÖFER S, BEKRIS K, HANDA A, et al. Sim2Real in robotics and automation: Applications and challenges [J]. IEEE transactions on automation science and engineering, 2021, 18(2): 398-400.
- [64] MIAO Q, LV Y, HUANG M, et al. Parallel learning: Overview and perspective for computational learning across Syn2Real and Sim2Real [J]. IEEE/CAA Journal of Automatica Sinica, 2023, 10(3): 603-631.
- [65] ARNDT K, HAZARA M, GHADIRZADEH A, et al. Meta reinforcement learning for sim-to-real domain adaptation [C]// 2020 IEEE International Conference on Robotics and Automation (ICRA): IEEE, 2020: 2725-2731.
- [66] 万松峰, 吴靖明. 基于机器视觉的智能抓取系统的设计 [J]. 机械管理开发, 2022, 37(10): 228-229.
- [67] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]// 16th IEEE International Conference on Computer Vision (ICCV), Venice, ITALY, 2017: 2980-2988.
- [68] CHEN L-C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs [J]. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [69] YANG M, YU K, ZHANG C, et al. DenseASPP for Semantic Segmentation in Street Scenes [C]// 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, 2018: 3684-3692.
- [70] CHEN L-C, ZHU Y, PAPANDREOU G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation [C]// 15th European Conference on Computer Vision (ECCV), Munich, 2018: 833-851.
- [71] HE Y, SUN W, HUANG H, et al. Pvn3d: A deep point-wise 3d keypoints voting network for 6dof pose estimation [C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020: 11632-11641.
- [72] XIANG Y, SCHMIDT T, NARAYANAN V, et al. PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes [J]. arXiv preprint arXiv:171100199, 2017.
- [73] DEB K, PRATAP A, AGARWAL S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II [J]. Ieee Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
- [74] DORIGO M, DI CARO G, GAMBARDELLA L M. Ant algorithms for discrete optimization [J]. Artificial Life, 1999, 5(2): 137-172.

- [75] KUFFNER J J, LAVALLE S M. RRT-connect: An efficient approach to single-query path planning [C]// Proceedings 2000 ICRA Millennium Conference IEEE International Conference on Robotics and Automation Symposia Proceedings (Cat No 00CH37065): IEEE, 2000: 995-1001.
- [76] TOUSSAINT M. Logic-Geometric Programming: An Optimization-Based Approach to Combined Task and Motion Planning [C]// IJCAI, 2015: 1930-1936.
- [77] SUNDARALINGAM B, HARI S K S, FISHMAN A, et al. Curobo: Parallelized collision-free robot motion generation [C]// 2023 IEEE International Conference on Robotics and Automation (ICRA): IEEE, 2023: 8112-8119.
- [78] 贺笑. 基于深度强化学习的机器人随机抓取研究 [D]. 南京: 东南大学, 2022.
- [79] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks [C]// Proceedings of the IEEE conference on computer vision and pattern recognition, 2017: 4700-4708.
- [80] 包玉龙, 徐斌, 舒昊鑫, 等. 触觉传感器研究现状与展望 [J]. 装备制造技术, 2019, (11): 17-21.
- [81] 张珂. 全向轮分拣系统的路径规划与控制仿真研究 [D]. 哈尔滨: 哈尔滨工业大学, 2022.
- [82] 温强. 基于深度神经网络的机器人运动控制技术研究 [D]. 秦皇岛: 燕山大学, 2022.
- [83] 戈振鹏. 基于主动视觉和强化学习的机械臂装配研究 [D]. 成都: 电子科技大学, 2023.

指导教师意见（课题难度是否适中、工作量是否饱满、进度安排是否合理、工作条件是否具备、是否同意开题等）：

本毕业设计课题聚焦于三维视觉引导的机器人抓取技术，开展 3D 视觉算法开发，并整合机器人操作系统，实现精准的物体定位、姿态识别以及机器人灵巧抓取。这一研究有望在工业自动化、物流搬运以及复杂环境操作等领域的机器人应用中产生深远影响，具有明确的方法论研究与实际应用价值。

课题难度适中，三位组员的工作量分配均衡且充实，项目团队制定了合理的进度规划。目前，用于研究的先进 3D 摄像头、机器人手臂平台以及高性能计算设备等硬件资源均已到位，这些硬件设施将为我们的设计与研究提供坚实的基础。鉴于以上准备工作充分，我们提议启动此关于基于 3D 视觉的机器人抓取技术的研究课题。

指导教师签名： 李海

2023 年 12 月 29 日

答辩小组意见：

审查结果： 同意 不同意

答辩组长签名： 王东

2024 年 1 月 12 日