



SRDP—创新引导专项 结题报告书

项目名称：一套基于机器视觉和强化学习的废电池分拣方法的设计

起止时间：2019.10-2021.3

负责人：杨澍

院系专业：电子信息工程

联系电话：17860767305

电子信箱：ys9018@stu.ouc.edu.cn

指导教师：俞智斌

年 月 日

填表须知

一、项目结题报告书参照《中国海洋大学本科生研究发展计划实施办法》的相关要求，逐项认真填写，填写内容必须实事求是，思路清晰、论证充分、表达明确严谨。空缺项要填“无”。

二、格式要求：表格中的字体应为小四号宋体，18磅行距；需签字部分由相关人员以黑色钢笔或水笔签名。均用A4纸双面打印，于左侧装订成册。

三、项目结题报告由项目组成员集体填写。

四、项目结题除提交本报告书外，还应提供以下附录：

1. 附录一：SRDP—创新引导专项经费使用情况表
2. 附录二：项目成果支撑材料

一、项目组成员（成果排名以所列序号为准，姓名栏本人签字）

序号	学号	姓名	学院	项目中分工
1	18130041042	杨澍	信息科学与工程学院	项目负责人, 虚拟环境负责人
2	18130041043	张昊	信息科学与工程学院	强化学习及运动学逆解算法负责人
3	18020024016	寇一笑	信息科学与工程学院	目标检测负责人
4	18130061017	李博林	信息科学与工程学院	深度估计算法负责人
5	18090021020	宋奕聪	工程学院	柔性机械手模具建模和制作

二、本项目经费使用情况（明细见附录一）

1. 项目计划经费: 1 万元 2. 项目共使用经费: 0 万元

三、项目成果形式及数量（需另行提供附件）

1. ☒文献资料综述（无）份； 2. ☐调查报告（无）份； 3. ☐研究论文（无）份；
4. ☐发表论文（无）份； 5. ☐实验报告（无）份； 6. ☒软件设计 1 份；
7. ☒硬件研制 1 份； 8. ☐获得专利（无）份； 9. ☐其他（无）件，名称：
附件清单：

四、项目基本情况及取得的成果

1. 项目背景

如今的控制系统从算法、硬件设计到落地实施都极其复杂，需要具体工作具体分析，如快递分拣机器人，自动驾驶汽车等等。然而作为人类，我们在执行大多数此类工作时，尤其是在限定场景下，主要依赖的感官只有视觉，我们靠眼睛就可以在大部分场景下完成驾驶、搬运等工作，而且我们可以学习使用诸如叉车吊车一类的工具，这意味着我们对于工具的使用更为灵活。因此，我们应该也可以找到一种算法，使得机器人可以仅用光学传感器完成这些工作，并且可以轻松迁移到各种型号不同的机器人上。这样不仅可以极大地简化了系统，降低了设计难度，还减小了成本。事实上，得益于 GPU 的飞速发展，越来越多的研究人员正致力于利用深度学习实现这个目标。如汽车自动驾驶领域，Waymo 和 Uber 公司都使用了激光雷达模块来创建汽车周围环境的高精度三维地图，以此弥补纯视觉感知的不足，提供更多的安全冗余。而特斯拉创始人兼 CEO 埃隆·马斯克多次声称特斯拉汽车的自动驾驶系统中，识别交通场景将只使用摄像头而完全摆脱激光雷达，依靠基于摄

像机的纯视觉算法来识别交通场景，通过深度神经网络从安装在车辆周围的八个摄像头的视频源中检测道路、汽车、物体和人。这样无疑会让自动驾驶系统的成本直接下降一个数量级，Waymo 汽车的配备有 3 个激光雷达，单个激光雷达价值 4000 美元，而特斯拉的完全自动驾驶系统只需 1 万美元。自动驾驶产业动辄数百亿美元，物流分拣又有一套成熟的体系，相比之下，垃圾分拣是一个有需求，预算低，却对自动化有较高需求的领域。

2019 年 4 月，Bulk Handling Systems 在南卡罗来纳州伯克利县的垃圾回收中心配置了 9 台智能分拣机器人。不过受限于高昂的价格(20 万美元/台)，HBS 的产品覆盖范围有限，尚未进入国内市场。同样位于美国的 AMP Robotics(简称“AMP”)，该公司 2019 年对外宣布完成了 1600 万美元的 A 轮融资。不过 AMP 拒绝透露智能分拣机器人的价格，至今也尚未进入中国市场。北京爱分类垃圾分拣中心利用弓叶科技人工智能分拣机器人，进行了可回收垃圾的精准分类，这是人工智能垃圾分拣机器人的首次落地，虽然不清楚价格，但是从机械臂的关节灵活度以及体积来看价格很高。

显然垃圾分拣庞大的应用场景使得成本成为了无法回避的问题。因此，我们借鉴了约翰霍普金斯大学的 CRAVES 方法，只使用单目摄像头抓取物体。基于图片估计机械臂姿态，识别和定位目标物体，生成抓取策略，价格低廉，操作简单，也更容易部署到新的环境。

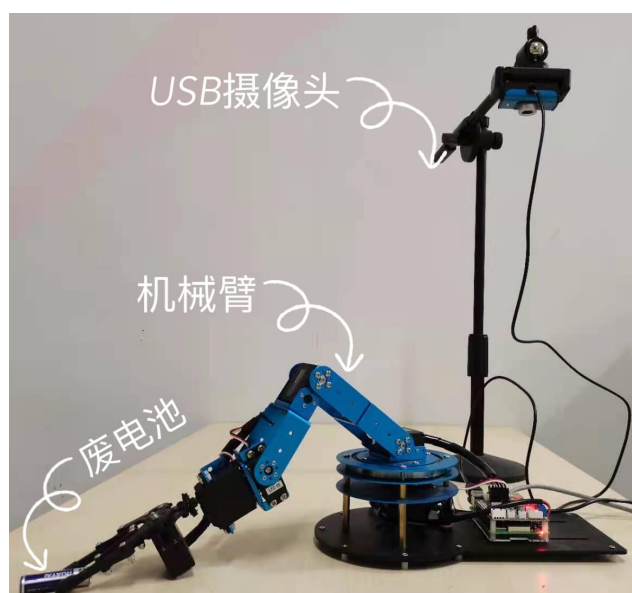


图 1 基于视觉的抓取系统示意图

2. 项目目标

① 在 2019 年 CVPR 论文 *CRAVES: Controlling Robotic Arm with a Vision-based, Economic System* 和 2019 年 ICRA 论文 *A Vacuum-driven Origami “Magic-ball” Soft Gripper* 的基础上，将其视觉控制机械臂的思想，以及 MIT 的新型柔性机械手应用于一个实际场景。

② 研究与设计一套基于机器视觉和强化学习的废电池分拣方法的原型机，实现简单常见的有害垃圾辅助分类，即在人的监管下实现初步的自动分拣。

3. 项目简介

虚拟环境：

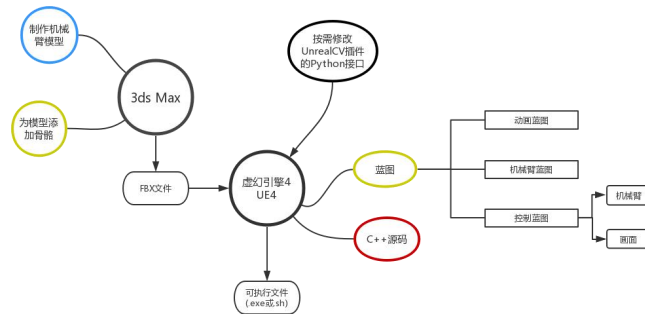


图 2 虚拟环境的结构图

虚拟环境的搭建使用了虚幻 4 引擎，虚幻 4 引擎是一款开源的游戏开发引擎，相比其他引擎，虚幻引擎不仅高效、全能，还能直接预览开发效果。用虚幻 4 引擎搭建用以训练机械臂的网络大致需要 3 个步骤。

① 在 3ds Max 中完成机械臂的建模和骨骼设置，建好后以 FBX 格式导入 UE4。

② 在 UE4 环境下编写控制和动画系统。UE4 支持两种编程方式，称为蓝图的可视化编程和 C++，二者的功能有一定重叠，但不完全相同，各自都有其不可替代之处，使用者可根据自身能力和需求决定功能重叠部分两种编程方式的使用比例。在我们的环境中，我们定义了一个动画蓝图（等价于 C++ 中一个动画类），一个控制蓝图，和一个自定义蓝图（以下简称 Arm 蓝图）。动画蓝图用来定义机械臂的运动范围和方式，控制蓝图用来定义各种控制函数，Arm 蓝图则是机械臂本身的抽象。

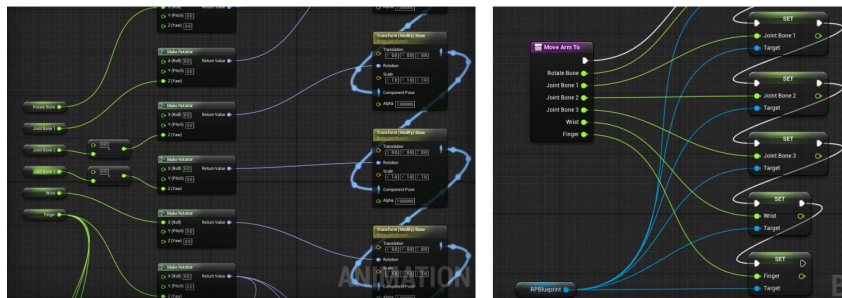


图 3 左图为机械臂运动的动画蓝图定义，右图则为控制机械臂运动的蓝图函数定义

③ 按需修改 UnrealCV 插件。UnrealCV 是一个帮助机器视觉研究者用 UE4 构建虚拟世界的开源项目，它可以提供与 Python 的通信，也可以利用它编写用于和虚拟环境交互的指令，以便在 Python 中调用。我们在原插件的基础上，根据 CRAVES 作者的公开数据独立复现了 10 个指令，并根据需要增加了诸多监控功能。

```
vset /arm/pose [float] [float] [float] [float] [float] [float] 设置机械臂姿态
vget /arm/pose 获取机械臂当前姿态
vset /arm/moveto [float] [float] [float] [float] [float] [float] 平滑移动到某姿态
vget /arm/keypoints 获取关键点的 3D 坐标
vget /arm/query_collision 查询是否碰撞
```

```

vget /arm/reset_hit 重置碰撞标志
vset /env/random_lighting 设置随机灯光
vset /env/floor/texture [str] 设置地板纹理
vset /env/sky/texture [str] 设置天空纹理
vset /data_capture/capture_frame [str] 生成一个用于姿态估计的数据
vset /data_capture/capture_masks [str] 生成一张语义分割图
vget /actor/velocity [str] 获取某物体当前速度

```

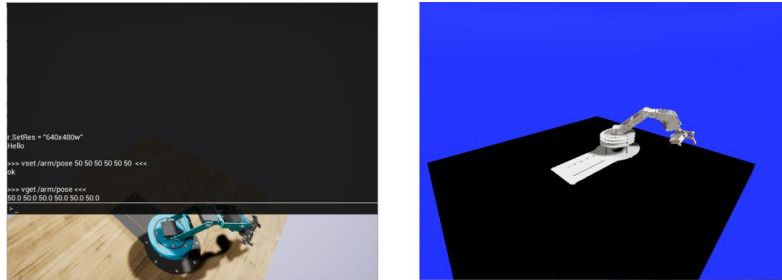


图 4 部分命令举例，左图为设置和获取机械臂关节姿态，右图为语义分割图

使用 UE4 搭配 UnrealCV 制作的虚拟环境可扩展性强，适用于绝大部分计算机视觉问题，从目标检测、语义分割、图像转译到强化学习、自主导航均可应用。同时该环境具有极强的兼容性，主流版本的 Windows 系统、Linux 系统和 MacOS 系统均可使用，同时支持 C++ 和 Python 通信。

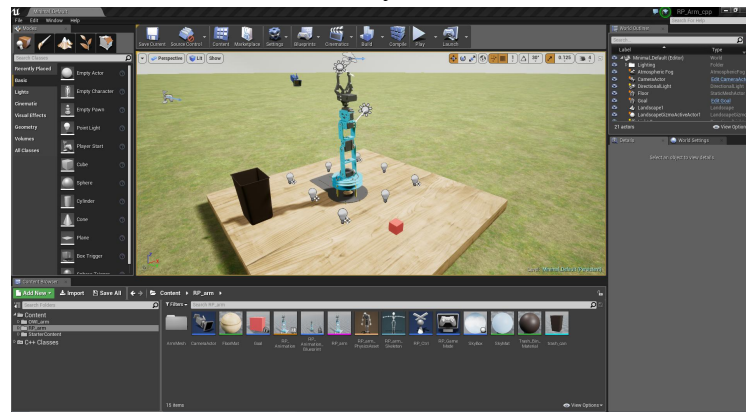
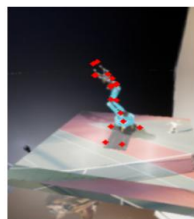
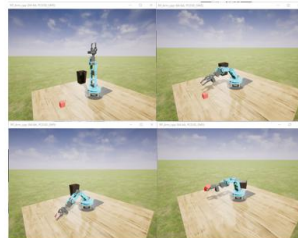


图 5 虚拟环境的编辑器

我们在虚拟环境中进行了姿态估计的训练。我们在机械臂上标注了 18 个关键点（如图 6）用以定位机械臂的姿态，利用虚拟环境标注了 5000 张图片作为数据集。制作数据集时，我们随机生成一个机械臂姿态，随机设置灯光，并从 COCO2017 的验证集里随机挑选两张图片，一张作为地板的纹理，另一张作为天空的纹理，以达到随机背景的目的。我们还在虚拟环境里测试了强化学习家族的 DDPG 算法。结果如图 6(b)所示



(a)



(b)

图 6 图(a)为姿态估计算法在虚拟环境中的效果，
图(b)为虚拟环境中强化学习算法练习抓取

我们的终极目标是能使用端到端的强化学习解决任务，即从图片推理出运动策略（当前方案为从图片->关节旋转角度->运动策略）这样不仅将两个网络变成了一个网络，缩短了推理时间，而且还减少了第一个环节产生的误差对决策造成的影响，因此在这个方向上我们使用 V-REP 环境测试了 DDPG 算法在这一方向上的可行性。

首先通过机器人通过深度相机获得图像的深度信息，然后用目标检测神经网络实时给出目标在图片中的位置，两相结合就能确定物体相对于机器人的位置，它对机器人进行后续强化学习抓取等方法做了必要铺垫。由于时间原因我们只做了小球模拟垃圾，让机器人进行自动检测和抓取。该方案仅对于实验进行仿真确认可能性，并为后续的强化学习研究做铺垫。

控制系统：

CRAVES 中，作者先用姿态估计网络得到机械臂的状态，然后将目标位置直接传给强化学习网络，由强化学习网络生成策略。经试验，该方法是可行的，问题在于无法准确估计待抓物体的位置。因此我们决定将该问题留给下一个团队解决，初步考虑可以用双目或多个摄像头解决该问题。

我们采用了垂直于机械臂的拍摄方法，可以省去目标物体一个维度的预测。当前三个网络均使用 Python 编程语言在 PC 上运行，因此为了提高速度，我们使用运动学逆解代替了强化学习生成抓取策略。

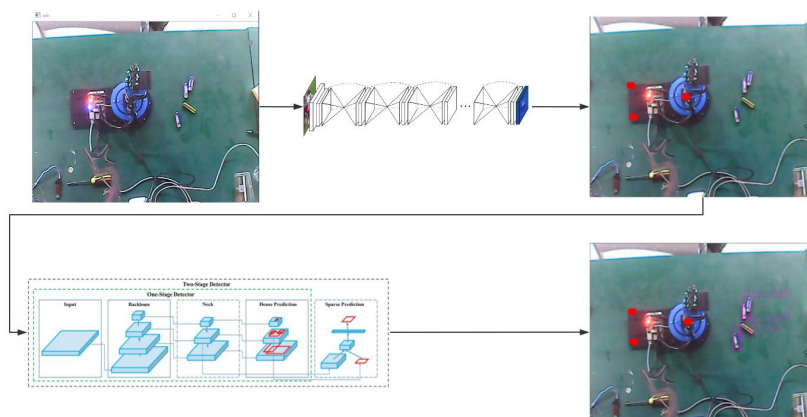


图 7 图片处理过程示意图



图 8 抓取系统框架图

姿态估计部分我们使用了经典的堆叠沙漏网络，堆叠沙漏网络是一款经典姿态估计网络模型，它开创性地使用了多尺度特征图，极大地提高了准确率。

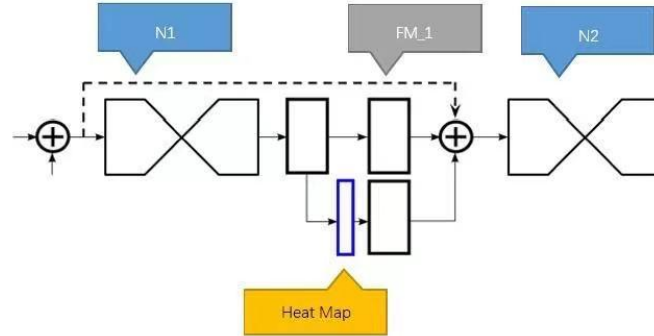


图 9 堆叠沙漏网络结构图

如上图，N1 代表第一个沙漏网络，提取出的混合特征经过 1 个 1×1 全卷积网络后，分成上下两个分支，上部分支继续经过 1×1 卷积后，进入下一个沙漏网络。下部分支先经过 1×1 卷积后，生成 heat map，就是图中蓝色部分。

上图中蓝色方块比其他三个方块要窄一些，这是因为 heat map 矩阵的 depth 与训练数据里的节点数一致，比如 $[1 \times 64 \times 64 \times 16]$ ，其他几个则具有较高的 depth，如 $[1 \times 64 \times 64 \times 256]$ 。

heat_map 继续经过 1×1 卷积，将 depth 调整到与上部分支一致，如 256，最后与上部分支合并，一起作为下一个沙漏网络的输入。

我们在虚拟环境中预训练了沙漏网络（结果如图 6(a) 所示），然后人工标注了 300 张真实图片，用真实环境对网络进行微调，结果如图 10 所示。图中黄色点为沙漏网络对机械臂关键点的估计。右边两个黄点之间的距离 d_0 已知，用于计算比例尺。左侧黄点为坐标系原点，坐标系原点到右侧两黄点重点的连线即为坐标系 x 轴，电池到坐标系远点的真实距离 d 即为

$$d = p \bullet \frac{d_0}{p_0}$$

其中 p_0 为右侧两参考点之间的像素距离， p 为电池检测框中心点到坐标系原点的像素距离。由此可以计算出每个电池相对于 x 轴的角度 θ_i ，机械臂抓取每个电池的策略就可以由运动学逆解公式求出。

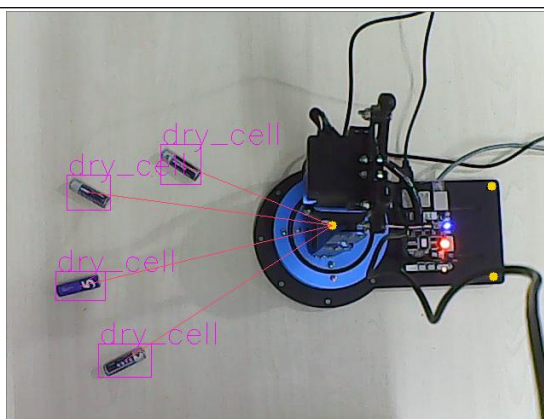


图 10 途中电池框 YOLOv4 的识别结果，黄点为沙漏网络对机械臂关键点的估计。

检测目标物体我们使用了 YOLOv4，从表 1 中的数据基本上超过了大部分的目标检测器，同时具有超快的速度，因此选择 YOLOv4 作为我们首选的目标检测算法。检测器通常由两部分组成：backbone 和 head。前者在 ImageNet 上进行预训练，后者用来预测类别信息和目标物体的边界框。在最近几年，目标检测器在 backbone 和 head 之间会插入一些网络层，这些网络层通常用来收集不同的特征图。我们将其称之为目标检测器的 neck。通常，一个 neck 由多个 bottom-up 路径和 top-down 路径组成。使用这种机制的网络包括 Feature Pyramid Network (FPN), Path Aggregation Network (PAN), BiFPN 和 NAS-FPN。

YOLOv4 使用了 CSPResNeXt+PANet+SPP+SAM，在精度和速度上都优于项目进行时最快的和最准确的检测器。

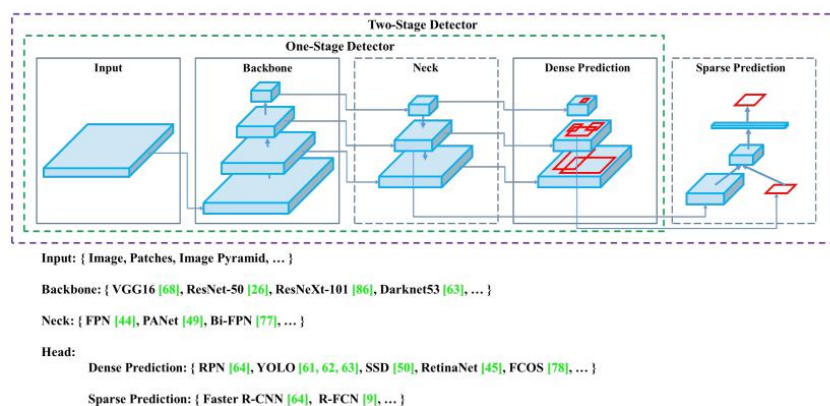


图 11 YOLOv4 原理图

机械臂：

我们的机械臂主控基于树莓派 4B，摄像头分辨率为 640*480，所有的算法均由 PC 完成，运算结果通过 SSH 传到树莓派控制机械臂。其结果如图 12 所示。

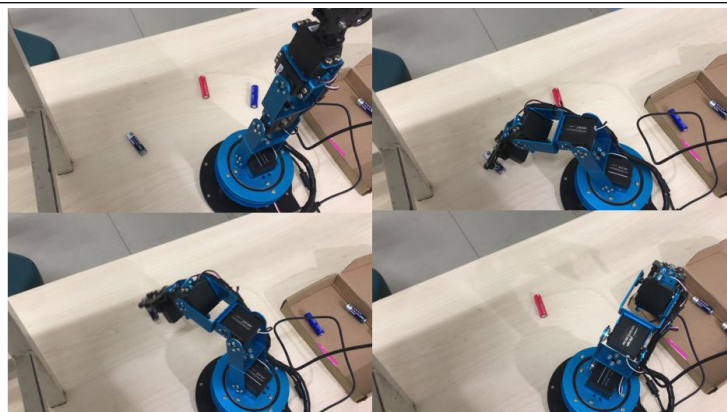


图 12 机械臂真机操作抓取

机械手模型：

一个鲁棒性良好的机械手也是提高抓取成功率的关键，因此我们为该系统选择了 2019 年 ICRA 会议中一篇名为 *A Vacuum-driven Origami “Magic-ball” Soft Gripper* 文章中的柔性机械手。该机械手使用浇筑的方法制作，我们完成了浇筑所用模具的 3D 模型。

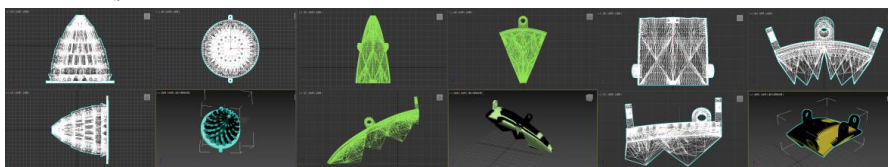


图 13 机械手模具的模型

4. 项目成果

我们项目的成果共有三部分：

1. 使用虚幻 4 引擎制作了一个便于测试算法和标注数据的虚拟环境，制作了一个姿态估计数据集（图 6(a)），并在虚拟环境中测试了强化学习算法（图 6(b)）。
2. 设计了一套基于纯视觉的基于视觉的控制系统，不依赖其他传感器，并使用真机进行了测试（图 12）。
3. 制作了柔性机械手模具的 3D 模型（图 13）。
4. 使用 V-Rep 测试了端到端的强化学习算法。

五、项目执行情况

1. 人员分工

杨澍：使用虚幻 4 引擎搭建了机械臂虚拟环境，制作了虚拟环境数据集，测试了强化学习算法（未使用），代码实现了姿态估计算法，算法整合

张昊：参与拍摄与标注了目标检测以及姿态估计的数据集，代码实现了运动学逆解算法，负责了强化学习算法文献的查找

李博林：参与拍摄与标注了目标检测以及姿态估计的数据集，YOLOv5 的使用，负责了目标检测算法文献的查找

寇一笑：参与拍摄与标注了目标检测以及姿态估计的数据集，测试了 Kinect 相机，并负责深度估计算法的实现（未使用），负责了专利文献、深度估计算法文献的查找

宋奕聪：为机械手模具建模

2. 工作进度

十月：立项，填写材料

十一月：拍摄、搜寻目标检测数据集并标注

十二月：训练目标检测算法，查找强化学习文献

一月：期末考试以及过年

二月——三月：搭建机械臂虚拟环境

四月：继续训练目标检测算法，提高精度

五月：在虚拟环境中测试强化学习算法

六月：期末考试

七月：在机械臂真机上测试抓取以及目标检测，效果不完美

八月：项目目标改为分拣三类有害垃圾，提出不含强化学习的技术路线

九月：期末考试，开学事宜

十月：重新拍摄并标注目标检测以及姿态估计的数据集

十一月：整合算法，测试算法

十二月——一月：期末考试

二月：过年

3. 工作表现

杨澍：代码能力优秀，擅长钻研学习新技术

张昊：代码能力优秀，对算法有很好的理解能力

李博林：代码能力优秀，各项工作都能良好完成

寇一笑：代码能力优秀，工作态度认真，可以良好完成各项工作

宋奕聪：专业能力十分优秀，工作态度非常认真，是团队唯一一个有能力为机械手模具建模的人，团队不可或缺的人才。

六、技术积累传承情况

1. 技术积累传承主要形式推荐给感兴趣的学弟学妹。
2. 后续团队人员组成：无。
3. 后续团队研究方向：
 - ①使用 2 个或以上相机作为传感器
 - ②使用端对端的强化学习算法

七、项目实施心得及建议

1. 项目总结（对取得的成果进行总结）我们项目贡献主要有三部分：
 - ① 使用虚幻 4 引擎制作了一个便于控制和数据标注的虚拟环境，并制作了一个姿态估计数据集。
 - ② 设计了一套基于纯视觉的基于视觉的控制系统，不依赖其他传感器。
 - ③ 制作了柔性机械手模具的 3D 模型
2. 存在的主要问题
 - ① 未能利用端对端的强化学习算法。
 - ② 未能制作出柔性机械手。
 - ③ 未能使相机位置可以任意摆放。
3. 未来工作的设想（未来工作设想包括但不限于：竞赛、毕业设计、创业项目等）
 - ① 将柔性机械手用于 URPC。
 - ② 将项目用于毕业设计。
4. 收获和体会

项目的立项目标技术上是完全可以实现的，但落实到行动上却很难，我们的项目对人力资源的需求远大于对资金的需求，而学生能投入在项目上的时间相对有限，极大限制了我们的技术能力发挥。尽管如此，团队的所有成员还是在这个过程中积累了大量的经验，涉及编程能力，软件设计，文献检索与阅读，各种硬件的使用等多个领域，受益匪浅。
5. 对创新引导专项和创新教育实践中心的建议等

立项之初我们对于项目可能遇到的困难认识不足，因此希望今后可以向新项目组介绍往年项目的失误和失败，以帮助他们少走弯路，节省宝贵的时间，我们也非常愿意整理项目过程中的失误，传授经验。

八、项目指导教师意见

签字：年 月 日	
九、创新教育实践中心意见	
负责人签字： 年 月 日	盖章
十、大学生创新创业训练计划管理委员会意见	
负责人签字： 年 月 日	盖 章：