**附件三、**

**2018年（第一批）Google支持教育部产学合作协同育人项目**

**大学生创新训练项目申报书**

申报项目名称：基于深度强化学习与通用机械臂的智能棋类机器人

主申报人UR KEY：\_\_\_\_\_\_\_\_\_\_\_\_\_\_S18A07\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

（请务必准确填写您的UR Key。UR Key是合作高校教师便捷、安全参与Google中国教育合作项目的唯一标识，如果您还没有UR Key，请访问Google中国教育合作项目登记表网址 <http://services.google.cn/fb/forms/ur_user_register/> 进行登记）

主申报人姓名：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_黄舒帆\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

主申报人所属学校：\_\_\_\_\_\_\_\_\_\_\_\_\_\_华东师范大学\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Google中国教育合作部**

**2018年4月**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 项目名称 | 基于深度强化学习与通用机械臂的智能棋类机器人 | | | | | |
| 主负责人 | 姓名 | 学院 | 专业 | 毕业时间 | 学号 | 项目分工 |
| 黄舒帆 | 数学科学学院 | 数学与应用数学 | 2020年6月 | 10151900315 | 图像识别算法 |
| 团队成员  （如有） | 姓名 | 学院 | 专业 | 毕业时间 | 学号 | 项目分工 |
| 陈伟文 | 计算机科学与软件工程学院 | 软件工程 | 2019年6月 | 10152510217 | 机械臂精度控制 |
| 华 赟 | 计算机科学与软件工程学院 | 软件工程 | 2019年6月 | 10152510157 | 算法模型训练 |
| 指导教师 | 姓名 | 金博 | | 联系电话 | 18616956260 | |
| 职称 | 讲师 | | Email | bjin@sei.ecnu.edu.cn | |
| 主要研究领域 | 机器学习、深度学习、强化学习、在线学习理论与算法，及其在计算机视觉、医疗大数据领域的应用 | | | | |
| 项目简介  (200字内) | 传统棋类机器人的算法往往以暴力搜索或监督式学习为主，前者需要消耗大量计算资源，后者依赖海量人工数据。而机器人往往局限于软件，缺乏实体机器人。  本设计方案包含基于非监督的深度强化学习训练的棋类对弈模型，同时利用计算机视觉技术完成棋盘状态识别，另外采用通用机械臂配合一套控制算法控制棋子，达到模拟真实对弈场景。最后利用3D打印技术，完成了一系列部件的整合。预期实现一个如同与人类对弈的棋类机器人。    相机及  配套sdk  视觉算法  强化学习  Dobot机械臂  +控制系统 | | | | | |
| 项目主要内容 | **深度强化学习系统**  通过同CNN实现策略价值网络，用强化学习的方法来拟合得到最优的action，以此作为模型的输出。 策略价值网络 这里的策略价值网络指的是在给定当前局面s的情况下，返回当前局面下每一个可行action的概率以及当前局面评分的模型，整体结构类似于一个蒙特卡洛搜索树。策略价值网络的输出结果是作为蒙特卡洛树搜索的依据，这个网络是作为蒙特卡洛树搜索算法的一部分。而在收集数据对策略价值网络进行训练的过程中，也是包括了 选择，拓展，模拟和反向传播这四个步骤。  其中选择步骤是选择当前状态可以转移到的状态，而拓展则是将当前状态转换到之前得到的可以转移的状态，模拟则是不断执行之前的操作，直到达到结束状态（5子相连）为止，反向传播的过程则是先对结束状态评分（胜为1，负为-1，多次对局取平均值），再将评分一次传递给前向的状态。  在实际的网络描述中，使用了一个4\*9\*9\*1的张量来描述**状态s**，其中4指的是包含当前步骤的前溯4步局势，9\*9则是描述局势的三值矩阵，其中0表示未落子，1表示落了黑子，-1表示落了白子，最后一个1则是表达了当前的player（即当前状态为黑方或者白方）。而**action**则是从当前状态转移到下一个状态要对当前矩阵做出的修改，简而言之就是在某个未落子点落子，表现为将该点的值从0变为1或者-1（取决于当前player）。 强化学习方法 始终由最优模型来生成数据，在生成的数据的基础上，其中生成的数据是自我对弈过程中产生的每一步的局面以及经由反向传播得到的每一个局面的评分，在单独一盘中，胜为1，负为-1，然后为了在训练的过程中拟合到尽可能多的局面，根据评分的高低决定选择该action的概率，而在初始状态下，每个action的被选择的概率是等可能的，但是随着训练的进行，那些容易取得胜利的状态action被选择的概率会显著增加。 深度神经网络 最开始是公共的3层全卷积网络，分别使用32、64和128个 3\*3的filter，使用ReLu激活函数。然后再分成policy和value两个输出，在policy这一端*，先使用4个 1*1的filter进行降维，再接一个全连接层，使用soft max非线性函数直接输出棋盘上每个位置的落子概率；在value这一端， 先使用2个1\*1的filter进行降维，再接一个64个神经元的全连接层，最后再接一个全连接层，使用tanh非线性函数直接输出[-1,1]之间的局面评分。整个策略价值网络的深度只有5~6层  **图像识别系统**  图像获取使用了微视图像的工业用摄像头，用在官方给出的sdk基础上用C#二次开发了一个可以以一定时间间隔拍摄图像的后台程序，将得到的照片保存 在了本项目目录下的camera文件夹中  07FCCD29EC0C80B938FB381E257477F6.png  图像处理的设计流程是：每隔一定时间通过摄像头读取棋盘图像，再对读取到的棋盘图像做裁剪操作，截取出棋盘区域，接着将截取出的棋盘区域与上一个时间片得到的棋盘区域last进行比对，得到落子的颜色以及位置，并更新last，在截取功能的实现中，先对原始图片做了中值模糊进行降噪，然后通过Canny操作获取图像边缘，最后只保留最大边缘内部的区域，该区域即为棋盘区域，在比对功能的实现中，先对两张图片都采用了高斯核的自适应二值化操作，得到它们的线框图，然后对线框图进行异或操作并降噪，得到落子的位置，最后在原始图像上，根据得到的落子位置，对像素点的值做均值操作，若大于设定阈值，则为白子，反之，则为黑子，并求得落子区域的中心点，再将中心点的坐标与经过截取操作后的棋盘区域进行映射得到逻辑位置即落子位置的[x,y]表示最后返回落子位置的[x, y]表示以及落子方。  A25898EBDB75E050D4476D1D9BD52396.png  本项目组已完成上述的算法实现并进行了初步调试，该算法可以有效地避免由于光照不均匀带来的识别错误，但是该算法依然没有办法解决摄像头或者棋盘发生较大抖动或者移位时的正确识别，因此，本项目组认为可以在之后使用YOLO算法对棋盘区域中的格点做目标检测，然后分析每个格点是否落子以及判断落子的颜色，这样可以解决移位带来的识别困难，进一步增加系统的识别可靠性。  **机械臂控制系统**  采用DOBOT魔术师机械臂，用机械臂末端的吸盘吸棋子和落子。俯视图如下：    用官方提供的一套易用的API的python接口，实现一个机械臂控制系统，选用机械臂的吸盘模块，用于移动棋子。标定棋盘的四个角，随后测量出棋盘格边长，棋盘逻辑坐标与真实物理坐标使用左下角+xy轴偏移量的形式来实现精确定位。由于机械臂长度有限，这里使用9\*9的棋盘。即使如此，依然需要将棋盘垫高才能使机械臂到达所有可落子区域。为此，侧视图如下：    棋盘为9\*9共81个棋盘格，棋盘区凹槽准备6\*7个棋子格，令 tot\_idx = 42，cur\_idx = 0（初始化），每一个idx(index)对应一个xyz坐标，表示该棋盘格所在空间相对机械臂位置，棋盘坐标[X, Y]也对应一个xyz（同理）。  **3D打印**  为控制各个模块之间位置的精确定位，3D打印出一套定位的架子或者底板。  3D打印中其实原理简单，过程中也遇到不少问题，比如打印材料托盘过重，无法自动旋转加料，这里我们采用前期手动旋转加料，后期涂油减小摩擦来使其自动化。  3D打印定位十字架，该模块用于控制机械臂、棋盘、棋子凹槽区的相对位置，本来是一个四个方向都很长的架子，由于打印机尺寸限制，不得不做些妥协，砍去延长部分，只打印控制相对位置的十字。    3D打印棋子凹槽，该凹槽为一个原型，用于放置棋子，因为直接从棋子盒中取棋子有一定困难，这个凹槽采用一个矩形采用布尔减法减去一个球的一部分获得，具体计算过程如下：    110^2 + (r-42)^2 = r^2  解得 r = 165  凹槽设计过程  众所周知，3D打印过程异常缓慢，同时为了节省材料采用了一定空心化，即在打印物品内部不完全实心，用蜂巢一样的六边形。  打印机不算小，但是我们需要打印的物件挺庞大的，于是分块打印，将大物件拆成各个小物件 | | | | | |
| 项目目标 | 用强化学习算法，实现棋力强大的模型，完成一个通用化棋类对弈模型。  在使用任何摄像设备的情况下，高效，高可靠地准确识别出当前局势，达到到百分百的判断精准。  精准控制机械臂对棋子的操作，使误差达到当前工业机械臂可以达到的最小误差。  最终实现全自动的棋类博弈机器人，体验与同人对弈无异。 | | | | | |
| 已有基础 | 三维打印的设计过程是：先通过计算机建模软件建模，再将建成的三维模型“分区”成逐层的截面，即切片，从而指导打印机逐层打印。设计软件和打印机之间协作的标准文件格式是STL文件格式。一个STL文件使用三角面来近似模拟物体的表面。  本项目组已经购置了Makerbot Replicator 2 桌面3D打印机。具备100微米的打印精度，410立方英寸的打印体积。本项目采用3D打印机打印固定机械臂、棋盘与相机的底座，以及放置旗子的凹槽区，目前以及完成了CAD设计，并实现了初版模具的打印。  DOBOT是由深圳越疆科技有限公司研发的机械臂，号称是全球首款高精度消费机桌面机械臂。DOBOT机械臂可以灵活翻转挪移，能进行画、写、移动、抓握东西，比如在电脑上打字、写字、画画、使用照相机等。2018年春晚该型号机械臂表演了书法春联。  本项目目前有一台dobot机械臂，用于棋子的移动。目前已完成了棋盘区域的定位与整体逻辑代码的编写，棋子格区域由于刚刚打印完成，部分数据坐标还未标定，不过逻辑代码已经完成，还需标定数据再稍加调试理论上便可跑通。 | | | | | |
| 预期成果 | 对于一种棋类，可以通过输入其游戏规则和实际道具（棋盘、棋子等），完成一个可以和人类对弈的机器人，机器人由眼睛（照相机），手（机械臂），大脑（深度强化学习算法训练出的模型）组成，整个机器人集成在NVIDIA tx2上，模拟与真实人类下棋体验。根据算法的训练数据多寡以及算法复杂度，将训练完成的棋类AI分为入门与高级两个等级，分别满足娱乐以及实际训练的需求。        注：请合理估计项目预期成果。若获得资助，请按照项目预期成果完成项目，如果最终无法按照承诺完成项目预期成果，我们会在Google大学合作部网站公布未结项名单并在汇报给教育部高教司的年度汇报总结中列出未结项名单；同时对于同一指导老师的其他项目申报，我们将谨慎评审，谢谢理解。 | | | | | |
| 项目特色与创新 | 集成了深度强化学习、图像识别、机械臂控制等诸多模块，物理集成上使用3d打印打印出连接、定位的支架。  深度强化学习：使用基于无监督的强化学习的Alpha Zero算法，这样可以不需要收集大量的真实对弈数据，由此减少了获取数据的成本，并且Alpha Zero算法可以在不断地自我对弈中逐渐提高棋类AI的智能性。因此，可以通过不断地让模型进行自我训练使得AI的智能性逐步提高。相较于传统的启发式搜索算法需要进行复杂的启发函数设计等，深度强化学习在数据有效性、模型简洁性以及结果可靠等方面均表现出了较大优势。  图像识别：使用了轻量级的传统方法对图像进行处理，避免了大量训练产生的时间成本以及算力成本，同时通过将三值化问题转换为二值化问题的过程，有效避免了光照不均衡问题。  机械臂控制：使用了通用机械臂实现了功能，极大地降低了开发成本，并且由于可以使用通用的软件SDK，极大地提升了开发效率，并且易于维护和复用。在开发商提供的SDK基础上，手动拓展出了跨平台环境。  3D打印：使用了轻量级建模软件123D，利用分治和布尔运算构建出了模型。在建模的过程中，通过合理地空心化处理，节省了材料与时间 | | | | | |
| 实施进度 | 2018年1月  查阅相关文献，基本了解强化学习原理（同时了解蒙特卡洛方法以及深度神经网络）。  了解主流机械臂，综合考量做出选择  2018年3月  分析对比已有的开源模型（包含围棋、五子棋、黑白棋等），在此基础上构建一套适用于当前场景的模型，并进一步了解了构建强化学习模型的有效方法。  开始了解图像处理基本原理，了解open cv常用接口，提出一种分割棋盘区域加权平均的图像识别思路。  2018年4月  修改之前图像处理思路，变为基于自适应二值化与图像异或的图像算法  构建机械臂控制系统规划，开始模块编写  为了整个系统的相对稳定，计划3D打印出地板或是定位器，开始3d建模  2018年5月  强化学习模型完成初版，在人机对弈中取得较好成绩  机械臂控制系统基本完成，开始控制算法调试  3D打印完成，系统准备联调 | | | | | |
| 经费使用规划（请给出具体明细） | 工业相机+镜头+支架，大约4000  3d打印耗材，大约1000  纸质材料打印复印费用，300  外出交流差旅交通费，大约2000  专利申请，大约1000  五子棋实体棋盘棋子，200  办公用品（图书、尺笔、鼠标、键盘、固态硬盘等）：1000  总计约9500元 | | | | | |
| 指导教师意见 | 指导教师签字：                                              年 月    日 | | | | | |
| 院系审核意见 | 负责单位（公章）：  年 月 日 | | | | | |