

23-24 学年春季强化学习第 2 次实验作业

(机器人方向)

教师：赵冬斌 朱圆恒

教师助教：陈亚冉

2024 年 4 月

作业任务-强化学习机器人视觉导航

利用课堂讲述的强化学习算法实现机器人导航任务，环境提供机器人的第一视角图像信息，设计强化学习策略控制机器人移动速度到达指定位置，抓取目标点。

我们提供了 StableBaseline3 框架下的 PPO 简单实现，有一定学习效果但不能稳定学习到有效的策略，可能由于以下原因：状态维度高，仅依靠图像输入通常需要较长的学习时间和良好的参数设置；环境奖励稀疏，仅在到达目标位置时有奖励等等。

针对以上问题，自己定义合适的辅助密集奖励，或使用更适合连续动作空间的强化学习算法（DDPG，SAC 等）来实现有效的导航策略学习（300 个 episode 内），相比于给定的 PPO 有提升，完成机器人导航任务。

将整个过程整理成完整的报告，包括奖励函数设计、网络结构设计、训练使用算法、算法参数选择以及结果分析，整个小组研究过程中的分工与贡献等，在规定的时间内提交报告和源代码，完成实验作业。

注意：严禁抄袭代码和报告，若发现，本部分为零分！

平台搭建

具体环境下载、搭建和基线算法，可参考：<https://github.com/DRL-CASIA/EpMineEnv>

