HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 1 (2023)

A Study of Assistive Robotics in Simulation Environment

Ava Shi

Palo Alto High School, Palo Alto, CA, USA

ABSTRACT

As the demand for caregiving increases, robots hold significant promise to provide improvements in the quality of life for millions worldwide. While previous research has provided baseline learning environments, this experiment investigates the adaptability of control policies in new environments that are more realistic to real world applications. Using Assistive Gym, an open-source physics simulation framework for the study, the 6 commercial robots were trained with control policies, then in a customized bed/bathing environment with varying bed positions and multi robot reaching environment with cooperative reinforcement learning.

Introduction

In the United States, 26% of the population, roughly 64 million adults, have some type of disability [1]. About 13.7% have serious difficulties walking or climbing stairs and 3.7% have difficulty with dressing or bathing. Robotics that offers versatile physical assistance present the opportunity to positively impact the lives of people who require support for their everyday tasks. Compared to real-world robotics systems, physics simulations provide a safe environment where robots are able to act, make decisions, and learn from their mistakes without putting real individuals at risk. The simulation accounts for realistic population diversity through a wide spectrum of human body shape, weight, and physical capability enabled by a variety of powerful hardware, including CPU, GPU and DPU; thousands of human-robot trials can be performed in a few hours. Wildly available physics simulators, including Bullet [8], Webot [13], Gazebo [12], and MuoJoCo[14] lower the barrier to build and research robot tasks in a safe simulation environment. Many robotic assistive task focused simulation frameworks and applications have emerged, noticeable ones include Assistive Gym [2] and RcareWorld [3] .]. This study uses Assistive Gym thanks to its open source, established baseline of robot assistive tasks and quick start features.

Related Works

A. Simulation Environment

OpenAI Gym is a reinforcement learning framework for learning control policy for simulated agents. In this study, both robots and humans are the agents. Assistive Gym [2] is open-source physics-based simulation framework built on OpenAIA Gym for physical human-robot interactions and robotic assistance. Some commonly used physics engines used for simulating robotic environments in OpenAI Gym are PyBullet [10], and MuJoCo [14]. Pybullet, which is used to build Assistive Gym, is a python module for the open-source Bullet physics engine that has been used for training and validating real robots using physics simulation.

B. Assistive Tasks (Environments)

Assistive Gym released a suite of simulation environments for six tasks [Fig .1] associated with activities of daily living, including: Itch Scratching, Bed Bathing, Drinking Water, Feeding, Dressing and Arm Manipulation.

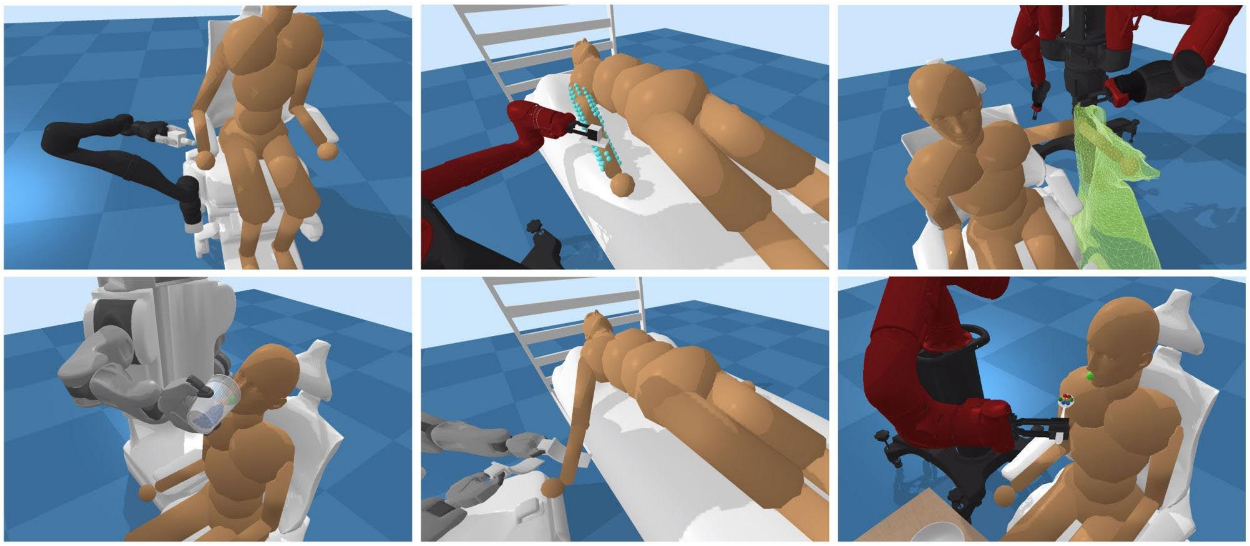

Fig. 1 six tasks from Assistive Gym [17]

The human models in the Assistive Gym contain body size, weight, and joint limits. Self-collision between various limbs and body parts of humans is enabled, human mobility limitations, such as head arm tremor and joint weakness, are also supported. In total, the human model in Assistive Gym has 40 controllable joints, including an actuated head, torso, arms, waist and legs. Using Pybullet, the human model can be programmatically generated.

Assistive Gym provides support for six collaborative robots that are commonly used for physical human-robot interaction. These robots include the PR2, Jaco, Baxter, Sawyer, Stretch and Panda. The first part of this paper examines the capabilities of these robot platforms to physically assist people and which robot platform performs the best.

Methodology

A. Control Policy

To train a robot to learn control policy for each assistance task, I used the same proximal policy optimizations (PPO) from the Assistive Gym baseline. PPO is a policy gradient algorithm used across a number of several contexts, from Atari games to real world quadruped robot locomotion [7]. When training a robot, control policy, all the parameters and settings for PPO and environment are held constant. Due to the limitations in hardware resources, the robot control policy for each task is only trained with 1M time steps. All the control policies are trained using a local Linux machine with an Intel Xeon® 40 cores CPU and a NVIDIA GeForce RTX 3080 GPU. Training time of each task varied from 0.5 hours to 5.5 hours.

B. Bedding Position Environment

In the Assistive Gym bed bathing environment, a person lies on the bed in a resting position, however, in a real medical application setting, patients are put into standard procedure positions, including prone, fowler, semi-fowler and lateral [16]. To evaluate a robot's adaptability to variation of environments, I modified the baseline of the bed bathing environment to introduce four standard bedding positions [Figure 3a.1]. In the new environment,, the angles of human joints angles, shoulders, knee and hip are adjusted accordingly to match the standard clinic bedding position. A new hospital bed model is selected and the bed angle radian is also adjusted to match the corresponding person position.

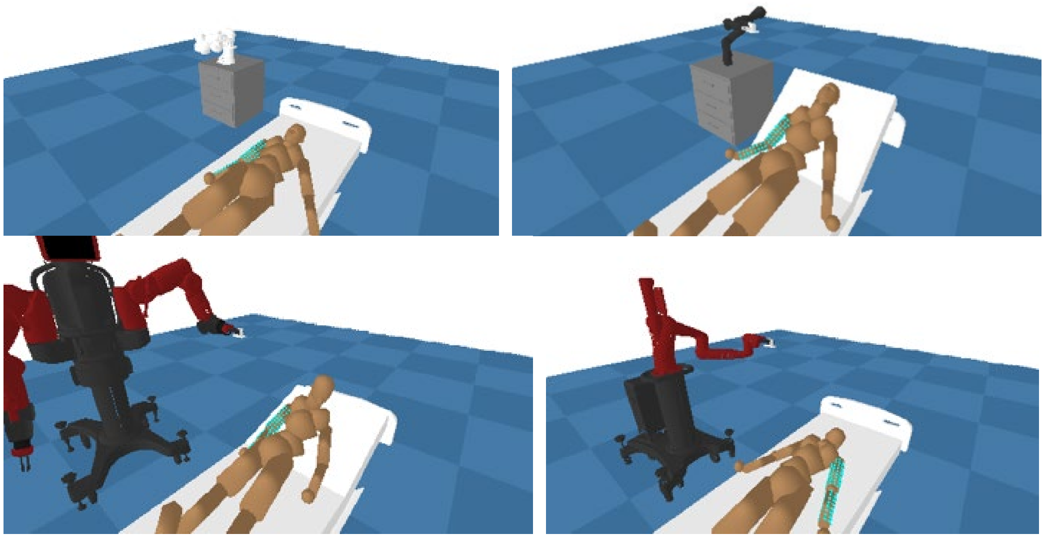

Fig. 3a.1. standard bedding positions: Lateral, Fowler, Semi-Fowler and Spine

C. Multi Robot Reaching Environment

Autonomous cooperative robotics have many applications in health care and medicine, including robotic surgery. Reaching a common target is the first step in many robots' cooperative tasks. The environment released in Assistive gym has only a single robot. Inspired by the examples of Create a New Reaching Assistive Environment [18] and multi-robot control [19], I built a new environment that is multi robot reaching to explore the control policy for multi robot corporations.

Fig 4a.1 Multi Robots Reaching Environment

The goal of two robots in this environment is to reach a person's mouth cooperatively. Fig 4b.2 illustrates the high-level architecture. In this environment, a super-agent controls two robots, the human is set to static and the same PPO algorithms from Assistive Gym feeding tasks are reused.

To simplify, the reward function for the agent is defined as : config('distance_weight')*reward_distance_mouth_target + config('action_weight')*reward_action + preferences_score

Each robot observes the status of its own and human. These statuses include end_effector_pos, end_effector_orient, joint angles, human head pos, human head orient. The action that the robot can take are the motors on its all-controllable joints. After collecting the actions and observations from individual robots, The Agent then concatenates them and pass into PPO algorithm.

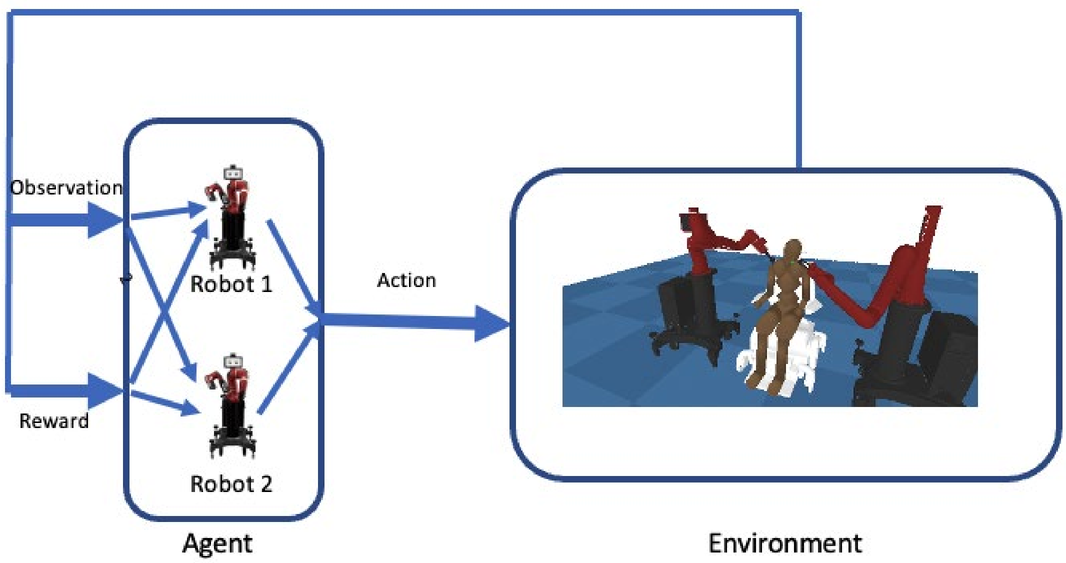

Fig 4b.2 high level architecture

Results and Findings

The charts in Fig 1a-e1-5 show the learning progress of each robot when was trained to execute an assistive task, the. Comparisons of the rewards that each robot received and its success rate during the training are listed in [Table-1], [Table-2] compare the success rate of six robots when associated with five assistance tasks.

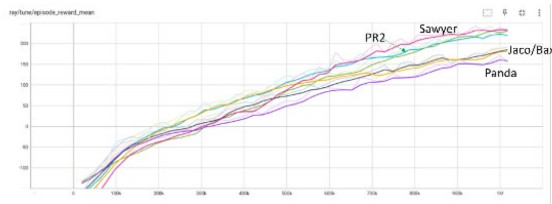

Fig. 2a.1. Bed Bathing task

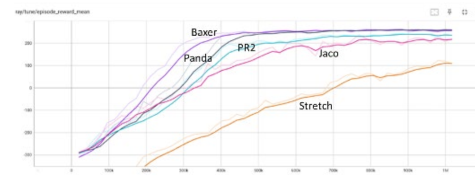

Fig. 2b.2 Feeding task

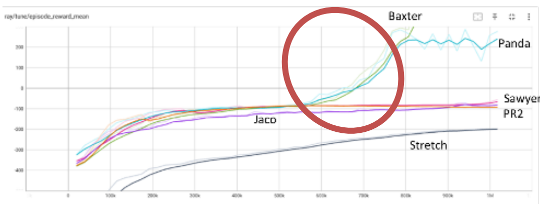

Fig. 2c.3. Drinking task

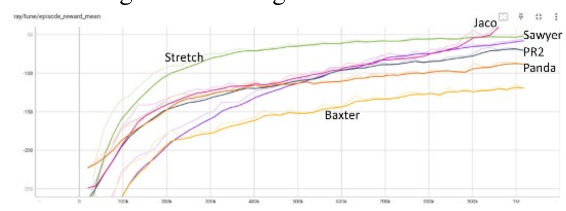

Fig. 2d.4. Dressing task

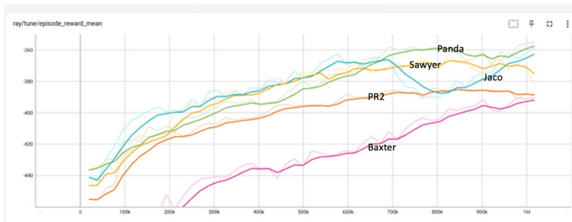

Fig .2e.5. Arm Manipulation task

Table 1 - Average reward for 100 trials with an Active Human model

Task	**PR2**	**Jaco**	**Baxter**	**Stretch**	**Panda**	**Sawyer**
Feeding	118.3	109.3	131.7	58.4	130.8	129
Drinking	-18.6	389	407.3	-99	202.9	-45
Bed Bathing	119.21	92.9	116.02	96.9	77.3	123
Dressing	-31	-9.04	-58	-26.7	-45	-15.5
Manipulation	-190	-170	-194	-201	-180	-187

Table 2 - Success rate for 100 trials with an Active Human model

Task	**PR2**	**Jaco**	**Baxter**	**Stretch**	**Panda**	**Sawyer**
Feeding	99%	91%	100%	70%	100%	0%
Drinking	0%	72%	65%	0%	39%	0%
Bed Bathing	30%	4%	13%	32%	1%	24%
Dressing	0%	32%	0%	0%	1%	0%
Manipulation	0%	0%	0%	0%	2%	1%

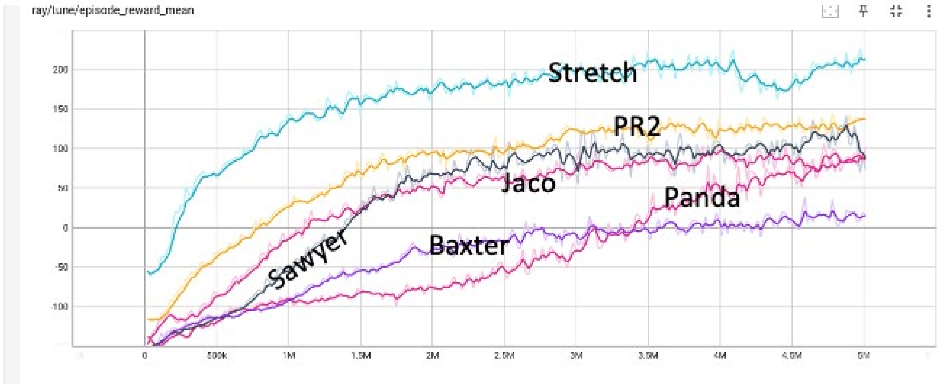

Fig.3b.2. The learning rate of bed bathing in fowler position

Table 3 - Average reward from for 100 trials in evaluation

Bedding Position	**PR2**	**Jaco**	**Baxter**	**Stretch**	**Panda**	**Sawyer**
Assistive Gym Baseline	119.21	92.9	116.02	96.9	77.3	123
New Position-Fowler	59.25	38.52	9.09	112.12	45.22	44.25

Table 4 - Success rate from for 100 trials in evaluation

Bedding Position	**PR2**	**Jaco**	**Baxter**	**Stretch**	**Panda**	**Sawyer**
Assistive Gym Baseline	30%	4%	13%	32%	1%	24%
New Position-Fowler	2%	4%	0%	45%	0%	8%

Of all three tasks, feeding, drinking and bed bathing, most robots were able to learn to reasonable control policy, with varying levels of performance between the robots robot. Dressing is the most challenging task for all robots, with no robot reaching more than 2% success rate. PPO training for dressing assistance took an average of 6 hours, 3 times more than the rest of the tasks. Long, long training time in dressing tasks is due to simulating dynamic cloth.

Under the new bed bathing environment, each robot was still able to learn reasonable control policy [Fig 3a.1] [Table 3] [Table 4]. For multi robots reaching tasks, I used PPO to train three robots, Saewyer, PR2 and Baxter. The rendering [Fig 4a.1] from trained control policy shows that robots can achieve the goal of reaching the human mouth cooperatively.

Limitations

Compared with the published results [2], my reproduced performance evaluation results [Fig 2a.1-e5] [Table 1] [Table 2] are lower. The discrepancy may be due to the lack of sufficient training steps: in my study, 1M time steps were trained, while the paper [2] reports using 10M time steps. Some robot limitations were also discovered. Baxter's short arm makes some tasks more difficult, such as arm manipulation, which requires reaching around the person's arm. In the drinking task learning, starting from 600K steps, both Baxter and Panda received a significant boosting reward as highlighted in Fig.2c.3. The reasons for sudden reward boosting had yet to be examined.

Consequently, the data presented in this work might only reflect part of the picture. The multi robot reaching environment presented is a basic use case, cooperative robots are the same type in each task, humans are modeled as static, reward function is also simplified. A more complex and realistic environment could be explored in the future.

Next Steps

Training multi agent/robots for cooperative tasks remains challenging and exciting. In this study, the multi robot reaching environment was built in this study could be enriched and enhanced. The improvement areas include adding an active human model, to enhance reward function, to use a dedicated agent for each robot and to experiment different reinforcement learning algorithms.

Acknowledgement

I would like to thank my advisor Matthew Giamou for his guidance in this project. I would also like to thank the team from Georgia Institute of Technology for publishing the open-source assistive gym platform and providing friendly documents.

References

[1] Admin. (2022, March 21). *Bullet real-time physics simulation*. Bullet RealTime Physics Simulation. Retrieved January 2, 2023, from https://pybullet.org/wordpress/

[2] *Anatomy, patient positioning - StatPearls - NCBI Bookshelf*. (n.d.). Retrieved January 3, 2023, from https://www.ncbi.nlm.nih.gov/books/NBK513320/

[3] *Bullet real-time physics simulation*. Bullet RealTime Physics Simulation. (n.d.). Retrieved January 2, 2023, from https://pybullet.org/wordpress/index.php/forum-2/

[4] Centers for Disease Control and Prevention. (2022, October 28). *Disability impacts all of us infographic*. Centers for Disease Control and Prevention. Retrieved January 2, 2023, from https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html

[5] Erickson, Z., Gangaram, V., Kapusta, A., Liu, C. K., & Kemp, C. C. (2019, October 10). *Assistive gym: A physics simulation framework for assistive robotics*. arXiv.org. Retrieved January 2, 2023, from https://arxiv.org/abs/1910.04700

[6] Gazebo. (n.d.). Retrieved January 2, 2023, from https://gazebosim.org/

[7] Google. (n.d.). *Google colaboratory*. Google Colab. Retrieved January 2, 2023, from https://colab.research.google.com/drive/1NPWZNFpB9NCgTQpbwM78jVHJAC7q_0oR?usp=sharing

[8] Healthcare-Robotics. (n.d.). *6. creating a new assistive environment · Healthcare-Robotics/Assistive-Gym Wiki*. GitHub. Retrieved January 2, 2023, from https://github.com/Healthcare-Robotics/assistive-gym/wiki/6.-Creating-a-New-Assistive-Environment

[9] Healthcare-Robotics. (n.d.). *Healthcare-Robotics/Assistive-Gym: Assistive Gym, a physics-based simulation framework for physical human-robot interaction and robotic assistance.* GitHub. Retrieved January 2, 2023, from https://github.com/Healthcare-Robotics/assistive-gym

[10] *Isaac Sim*. NVIDIA Developer. (2022, December 30). Retrieved January 2, 2023, from https://developer.nvidia.com/isaac-sim

[11] Kwiatkowski, A., Alvarado, E., Kalogeiton, V., Liu, C. K., Pettré, J., van de Panne, M., & Cani, M.-P. (2022, March 7). *A survey on reinforcement learning methods in character animation*. arXiv.org. Retrieved January 2, 2023, from https://arxiv.org/abs/2203.04735

[12] Körber, M., Lange, J., Rediske, S., Steinmann, S., & Glück, R. (2021, March 8). *Comparing popular simulation environments in the scope of Robotics and Reinforcement Learning*. arXiv.org. Retrieved January 2, 2023, from https://arxiv.org/abs/2103.04616

[13] *Mujoco*. MuJoCo. (n.d.). Retrieved January 2, 2023, from https://mujoco.org/

[14] *Ray 2.2.0*. RLlib: Industry-Grade Reinforcement Learning - Ray 2.2.0. (n.d.). Retrieved January 2, 2023, from https://docs.ray.io/en/latest/rllib/index.html

[15] *Reinforcement learning for robot research: A comprehensive review and ...* (n.d.). Retrieved January 3, 2023, from https://journals.sagepub.com/doi/abs/10.1177/17298814211007305

[16] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017, August 28). *Proximal policy optimization algorithms*. arXiv.org. Retrieved January 2, 2023, from https://arxiv.org/abs/1707.06347

[17] *Webots*. Webots: robot simulator. (n.d.). Retrieved January 2, 2023, from https://cyberbotics.com/

[18] *Welcome to spinning up in deep rl!*¶. Welcome to Spinning Up in Deep RL! - Spinning Up documentation. (n.d.). Retrieved January 2, 2023, from https://spinningup.openai.com/en/latest/

[19] Wikimedia Foundation. (2022, December 12). *Reinforcement learning*. Wikipedia. Retrieved January 2, 2023, from https://en.wikipedia.org/wiki/Reinforcement_learning

[20] Ye, R., Xu, W., Fu, H., Jenamani, R. K., Nguyen, V., Lu, C., Dimitropoulou, K., & Bhattacharjee, T. (2022, October 19). *RCareWorld: A human-centric simulation world for Caregiving Robots*. arXiv.org. Retrieved January 2, 2023, from https://arxiv.org/abs/2210.10821