Deep RL Arm Manipulation

Kun Zhao

Abstract—Deep RL Arm Manipulation project! For this project, My goal is to create a DQN agent and define reward functions to teach a robotic arm to carry out two primary objectives:
Have any part of the robot arm touch the object of interest, with at least a 90Have only the gripper base of the robot arm touch the object, with at least a 80

Index Terms—Robot Software, Udacity, Gazebo,Jetson TX2.

1 Introduction

THE A key difference between RL and Deep RL is the use of a deep neural network. Think of the collection of value-action pairs that define what actions an agent should take in any situation as a function of the observations that the agent receives from its environment.

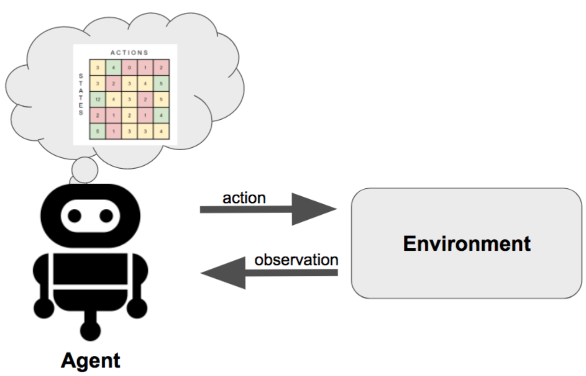

Fig. 1. RL.

2 Background

A deep Q-network (DQN), which is able to combine reinforcement learning with a class of artificial neural network16 known as deep neural networks.in Fig.3 the gazebo environment loads up, there are the robotic arm, a camera sensor, and an object in the environment.

3 Reward Functions

1. Robot arm touch the object.

Because only arm tourching the object and base joint is fixed, So just reward function, reward = -abs(gripperbox.min.z - probbox.max.z).

2. Arm's gripper base touch the object. base on the distance between gripper and object,compute the reward "distGoal = BoxDistance(propBBox,gripBBox)",and then compute the smoothed moving average of the delta of the distance to the goal,"avgGoalDelta = (distDelta * alpha) + (distGoal * (1.0f - alpha))" ; final,reward function is "10*avgGoalDelta - distGoal"

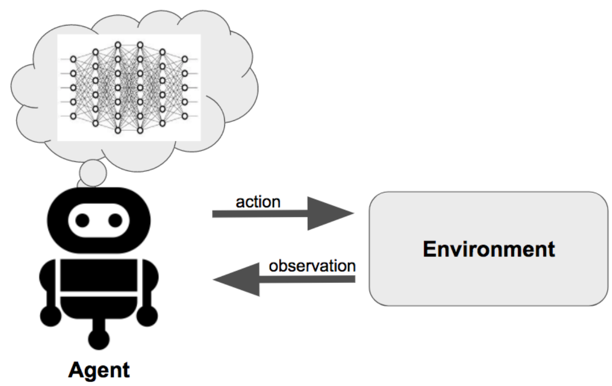

Fig. 2. DRL.

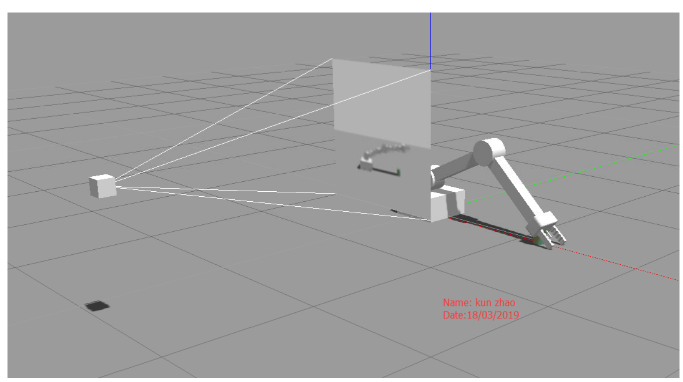

Fig. 3. Arm touching the object1

4 Hyperparameters

I choose same hyperparameters for both objectives, but using different reward function to reach different objectives.camera picture input resolution is 64*64,optimizer is "Adam", learning rate is 0.0001 , due to limit computation resource in cloud , I set replay buffer size to 2000(default 10000),batch size is 4. use LSTM is true, LSTM's size is 32. Because arm have 3 joint, every joint have two actions , so the action number is 6. if arm reach the goal(arm touching or gripper touching the objective),give reward to 2.0,or -2.0

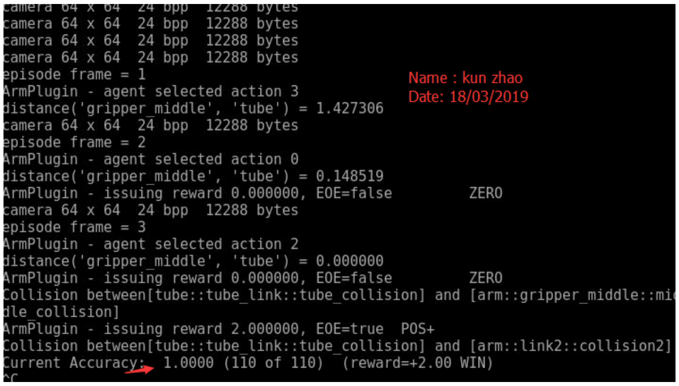

Fig. 4. Arm touching the object2

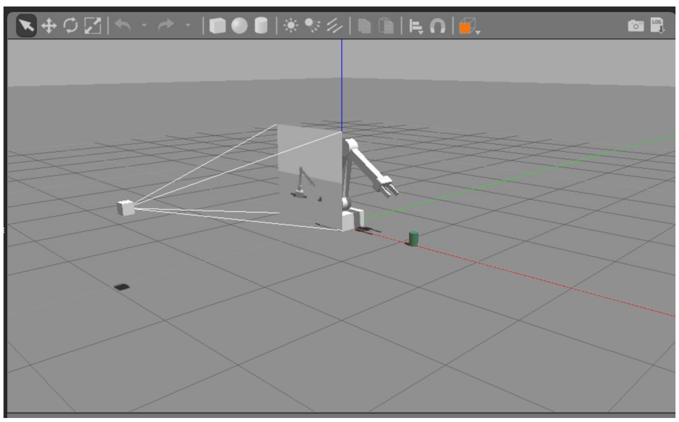

Fig. 5. gripper base touch the object1

5 Results

The robot arm touching the object with 100% accuracy for more than 100 runs.Please see Fig.3 and Fig.4 The Arm's gripper base touch the object with 100% accuracy for more than 100 runs.Please see Fig.5 and Fig.6

6 Future Work

Above result is base on the LOCK base joint setting. if unlock the base joint,it will harder to train than LOCK-BASE,because this is increase the free degrees of arm. more action space , more state space.

Fig. 6. gripper base touch the object2