# Robotic Arm Manipulation Using Deep Q-Learning

Vijayasri Iyer

**Abstract**—This project report is a brief overview of a serial manipulator which manipulates itself to touch, grasp and pick up objects in a simulated ROS (Robot Operating System) environment. The arm uses Deep Reinforcement Learning; particularly a technique called Deep Q-learning (DQN) to achieve this feat.

**Index Terms**—DQN, IEEEtran, Udacity RoboND, Serial Manipulator

## 1 Introduction

THE past few decades have seen robots being used as assistive labour to do heavy industrial assembly line jobs which are dangerous and repititive. In these cases, robots were mostly programmed by hand with a detailed description of the task. But now since robots are moving into industries such as healthcare and autonomous transport where it isn't possible to hardcode all of the features required for the task. This led to the use of machine learning. Machine learning is a set of algorithms that allows the computer to learn the features involved in a task. A particular paradigm of machine learning known as reinforcement learning where an agent will learn from the environment by interacting with it and receiving rewards for performing actions. Reinforcement algorithms that incorporate deep learning, have shown significant promise in this avenue. One such algorithm called Deep Q-learning was introduced in 2015 by Google's DeepMind which beat the world's best players at complex games like Atari and Go. In this paper, the DQN algorithm is used to teach a robotic arm to manipulate its surrounding objects using just 2D images of the arm and the object. Performing this task, will help fuel further development of robots to perform pick and place tasks using DeepRL. An illustration of the DQN algorithm is shown below in fig 1.

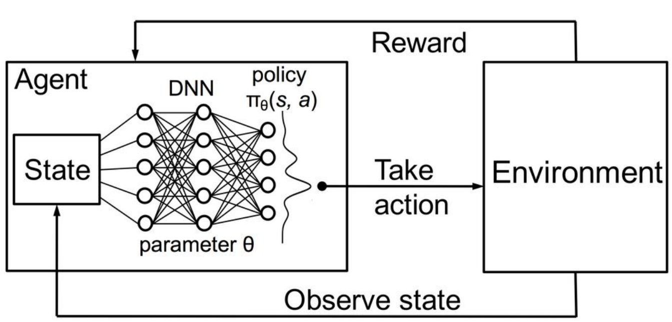


Fig. 1. Deep Q-Learning

### 1.1 Deep Q-Networks

The DQN algorithm is a spin-off of the original Q-Learning algorithm which uses deep learning instead of the Q-table (lookup table). DNNs (Deep Neural Networks) have been known to have excellent capability for extracting feature representations through backpropagation and the use of multiple layers for representation learning. After a sufficient amount of training iterations, DNN knows which information like color or shape is important to do the task. The DNN performs the task of mapping the 2D images to appropriate actions that the agent should take in order to reach the goal state. This is especially useful in an continuous state space, since the DNN can output several possible actions in one training pass. The DQN takes in the state of the agent, inputs it to the DNN, and as per a certain policy the DNN gives possible actions that the agent can perform to manipulate its environment.

## 2 Reward Policy

At this stage, you should begin diving into the technical details of your approach by explaining to the reader how parameters were defined, what type of network was chosen, and the reasons these items were performed. This should be factual and authoritative, meaning you should not use language such as I think this will work or Maybe a network with this architecture is better... Instead, focus on items similar to, A 3-layer network architecture was chosen with X, Y, and Z parameters Explain why you chose the network you did for the supplied data set and then why you chose the network used for your robotic inference project. [?]

- example

1) example

## 3 Hyperparameters

This section should discuss the data set. Items to include are the number of images, size of the images, the types of images (RGB, Grayscale, etc.), how these images were collected (including the method). Providing this information is critical if anyone would like to replicate your results. After all, the intent of reports such as these are to convey information and build upon ideas so you want to ensure others can validate your process. Justifying why you gathered data in this way is a helpful point, but sometimes this may be omitted here if the problem has been stated clearly in the introduction. It is a great idea here to have at least one or two images showing what your data looks like for the reader to visualize.

## 4 RESULTS

This is typically the hardest part of the report for many. You want to convey your results in an unbiased fashion. If you results are good, you can objectively note this. Similarly, you may do this if they are bad as well. You do not want to justify your results here with discussion; this is a topic for the next session. Present the results of your robotics project model and the model you used for the supplied data with the appropriate accuracy and inference time For demonstrating your results, it is incredibly useful to have some charts, tables, and/or graphs for the reader to review. This makes ingesting the information quicker and easier.

## 5 DISCUSSION

This is the only section of the report where you may include your opinion. However, make sure your opinion is based on facts. If your results are poor, make mention of what may be the underlying issues. If the results are good, why do you think this is the case? Again, avoid writing in the first person (i.e. Do not use words like I or me). If you really find yourself struggling to avoid the word I or me; sometimes, this can be avoid with the use of the word one. As an example: instead of : I think the accuracy on my dataset is low because the images are too small to show the necessary detail try: one may believe the accuracy on the dataset is low because the images are too small to show the necessary detail. They say the same thing, but the second avoids the first person. Reflect on which is more important, inference time or accuracy, in regards to your robotic inference project.

## 6 CONCLUSION / FUTURE WORK

This section is intended to summarize your report. Your summary should include a recap of the results, did this project achieve what you attempted, and is this a commercially viable product? For Future work,address areas of work that you may not have addressed in your report as possible next steps. For future work, this could be due to time constraints, lack of currently developed methods / technology, and areas of application outside of your current implementation. Again, avoid the use of the first-person.

#### *6.0.1 Subsubsection Heading Here*

TABLE 1
Table

| | |
|---|---|
| One | Two |
| Three | Four |