School of Computer Science and Applied Mathematics
UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

COMS4047A DQN Project

Angus Mackenzie - 1106817

October 1, 2019

1 Introduction

This lab explores the implementation of a Deep Q-Learning (DQN) agent to play pong through the use of OpenAI [1]. Initially we were provided source code, which acted as the base, and then a solution was built on top to the specification of [2] and [3].

2 Method

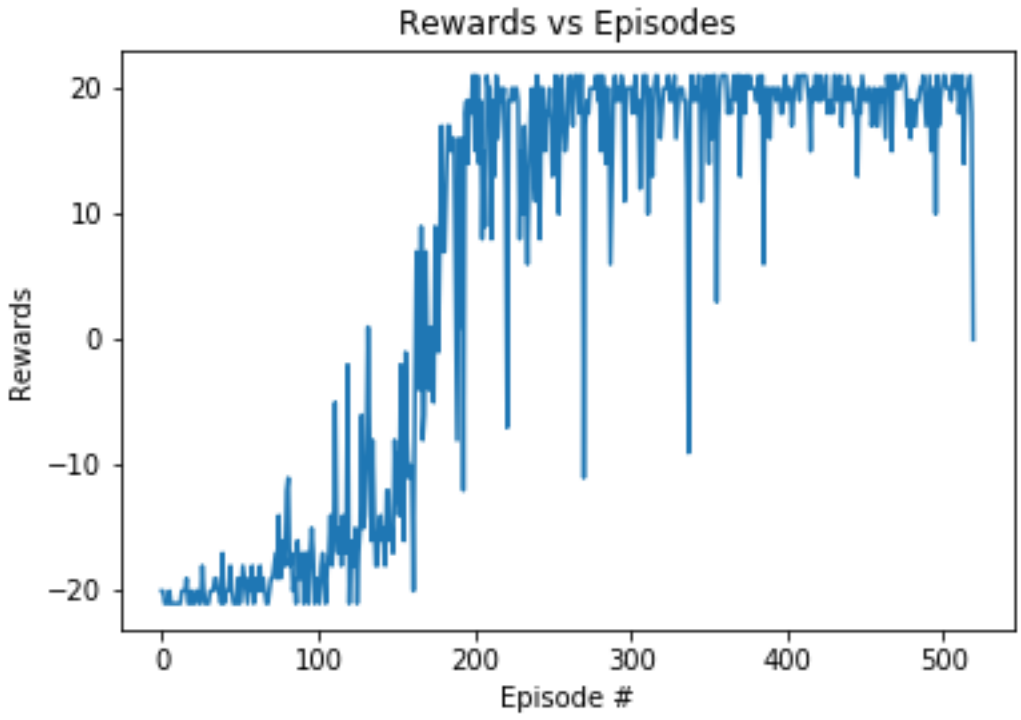
Before creating the models, various aspects of gym allow the input from the atari gameplay to be manipulated into a 84x84 frame that can be fed into our model [1].

A model was created in order to train the DQN, this model used the same architecture as that of [3]. This network takes in an input of 4 frames of 84x84 pixels each as the state representation, and then has an output for each possible action. The Q-network used 3 convolutional layers made up of *Conv2d*, that were all passed through a ReLu before continuing to the next layer.

- The first layer of the network has a stride of 4, and an input of 16 8x8 filters
- The second layer has an input of 32 with a stride of 2 and 4x4 filters.
- The Third layer has an input of 64, and a filter of 3x3 with a stride of 1.
- We flatten the output, and then linearise it to ensure it is of the correct size.
- The output is simply the linear layer that has an output for each action.

The model was then trained for 1 million steps, and the rewards for every 100 steps tracked. The model was also saved every 100 steps.

3 Results



H

As shown above, the model started learning slowly, with initial rewards per episode beginning around -20. After training for roughly 2 hours the model increases the reward per episode to roughly 20.

References

[1] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym, 2016.

[2] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin A. Riedmiller. Playing atari with deep reinforcement learning. *CoRR*, abs/1312.5602, 2013.

[3] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.