Deep Learning Lab 2018

Srdjan Milojevic and Maciek Janowski

Exercise 4

1 Overview of the task

The main goal of this exercise is to implement a reinforcement learning algorithm and create a DQN agent and evaluate its performance on the CartPole and CarRacing control tasks from the OpenAI Gym benchmark suite. Additionally, students will find out more about experience replay, target networks and ε-greedy exploration.

2 CartPole Environment

The students first extend the starter code to work on CartPole environment. Using ε-annealing as a more sophisticated method than ε-greedy exploration, the created DQN agent "solves" the CartPole game (fulfill the condition for solving as stated in documentation) in only 60 training epochs as seen on Figure 1. On Figure 2 students present training episode rewards.

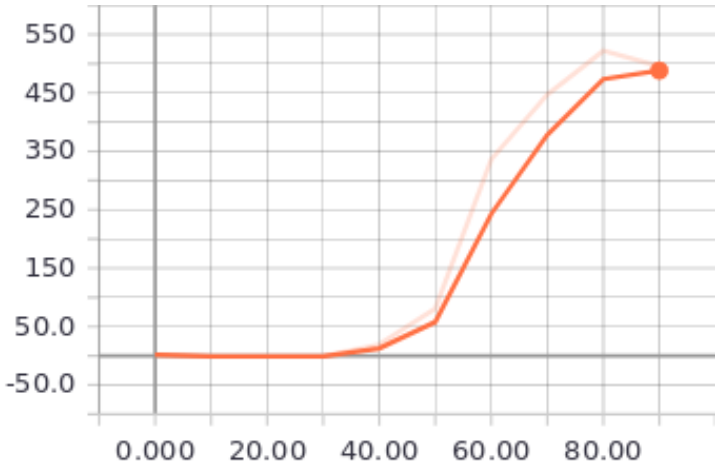

Figure 1: Averaged validation episode rewards.

3 CarRacing Environment

After students successfully solved CartPole, next task was to adapt the code for CarRacing environment. This means changing the neural network of two fully connected layers to a convolutional one. Other than that, input data needed to be reprocessed and the exploration phase had to be set to have certain biases towards accelerate and straight actions. With a setup like this, the students create a DQN agent which scores an average of 725 reward point on 10 runs. The results are displayed on Figure 3 and Figure 4.

The final code can be found here.

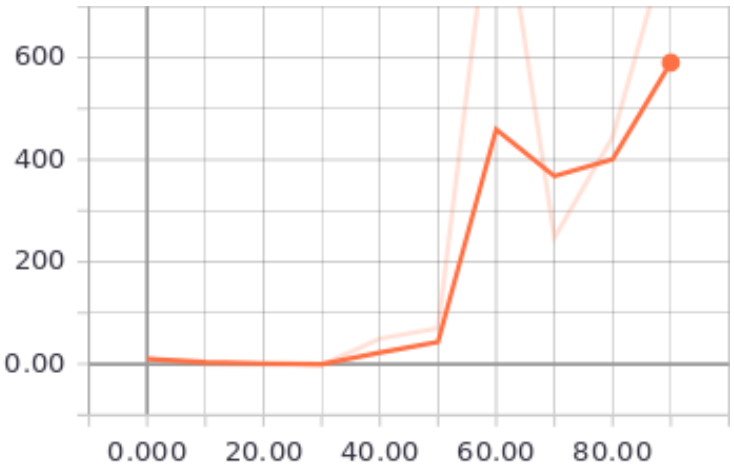

Figure 2: Training episode rewards.

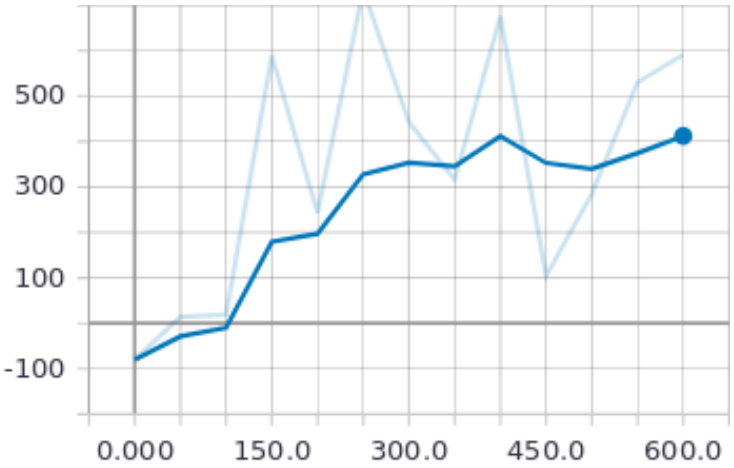

Figure 3: Averaged validation episode rewards.

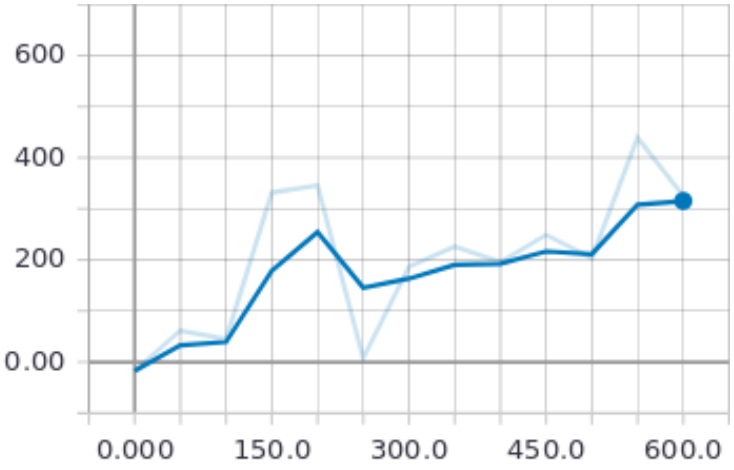

Figure 4: Training episode rewards.