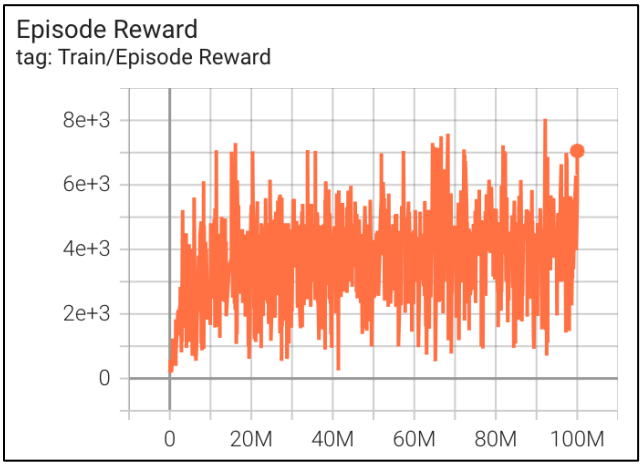
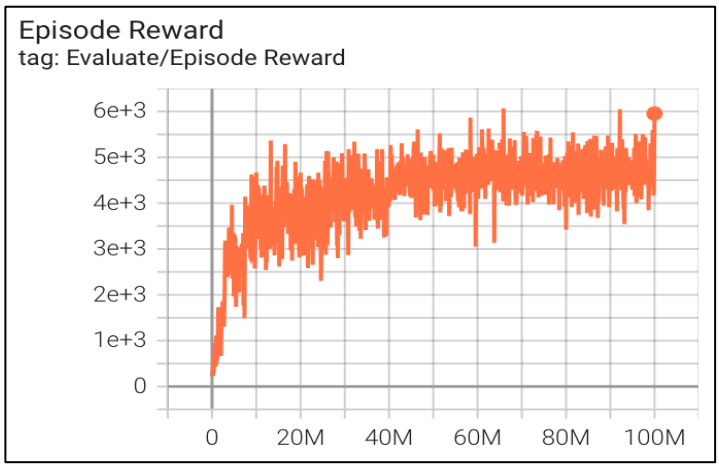
Deep Q-Network
Lab Report # 2

By
312581020
許瀚丰

Selected Topics in Reinforcement Learning
Fall 2023
Date Submitted: October 15, 2023

- Screenshot of Tensorboard training curve and testing results on DQN. (30%)
  - Training curve

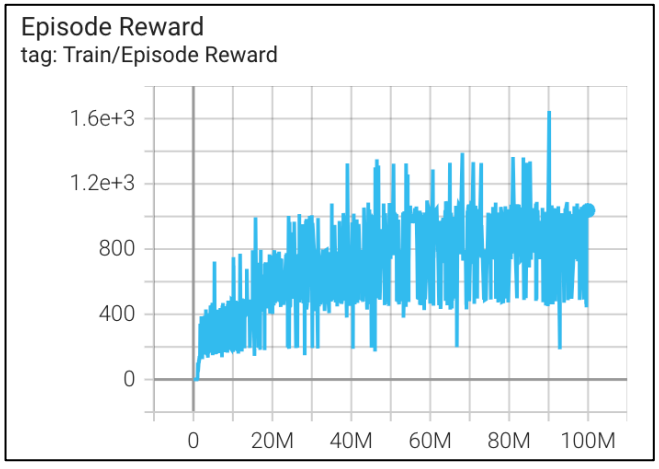


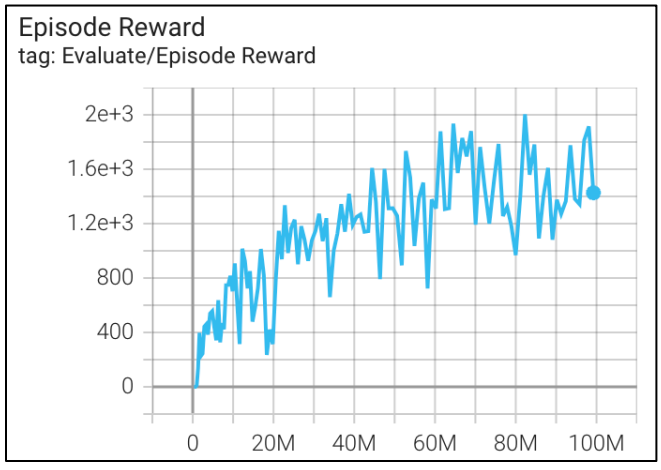


  - Testing results (5 games)

```
episode 1 reward: 5240.0
episode 2 reward: 4920.0
episode 3 reward: 5670.0
episode 4 reward: 4910.0
episode 5 reward: 7360.0
average score: 5620.0
```

- Screenshot of Tensorboard training curve and testing results on Enduro-v5 (10%)
  - Training curve

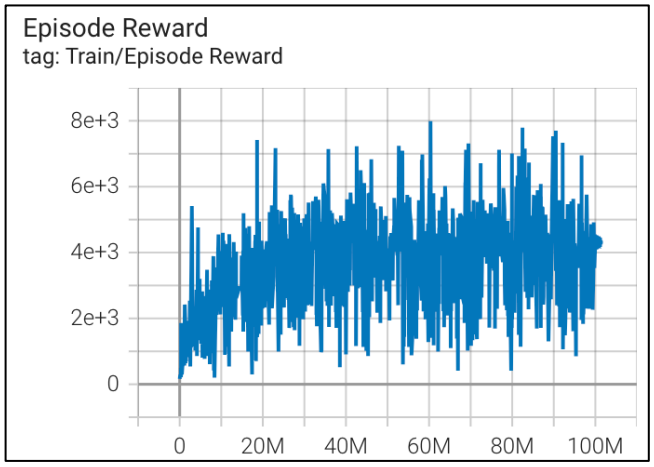


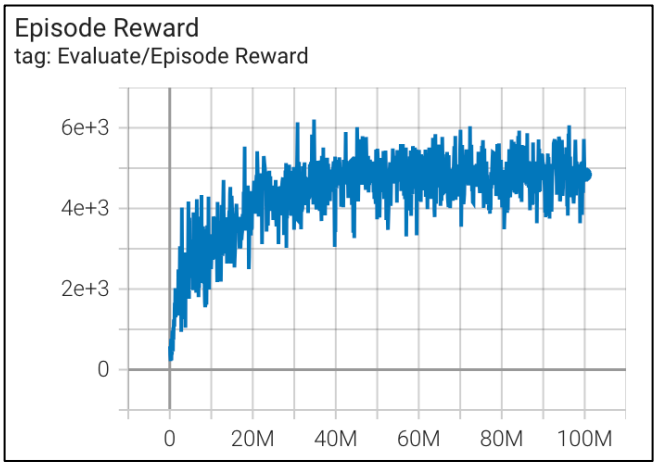


- Testing results (5 games)

```
episode 1 reward: 1977.0
episode 2 reward: 1687.0
episode 3 reward: 1375.0
episode 4 reward: 1995.0
episode 5 reward: 2340.0
average score: 1874.8
```

- Screenshot of Tensorboard training curve and testing results on DDQN, and discuss the difference between DQN and DDQN
  - Training curve

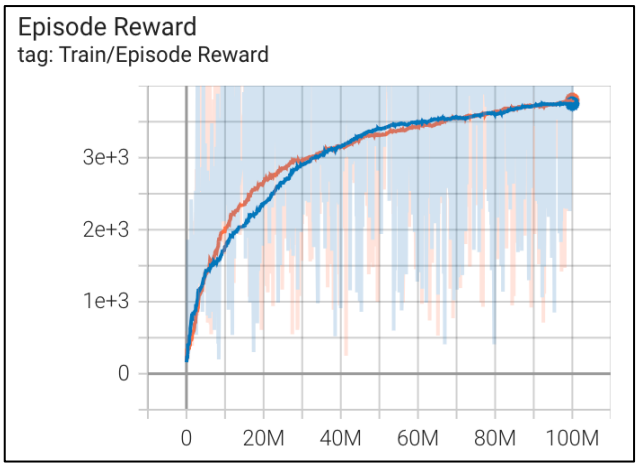


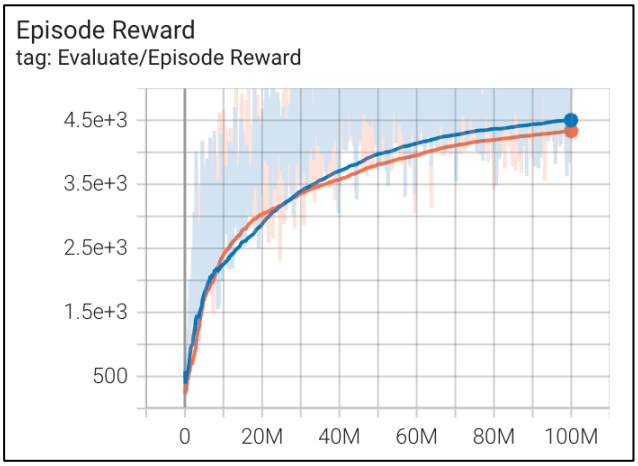


  - Testing results (5 games)

```
episode 1 reward: 5340.0
episode 2 reward: 4800.0
episode 3 reward: 4300.0
episode 4 reward: 4840.0
episode 5 reward: 7990.0
average score: 5454.0
```

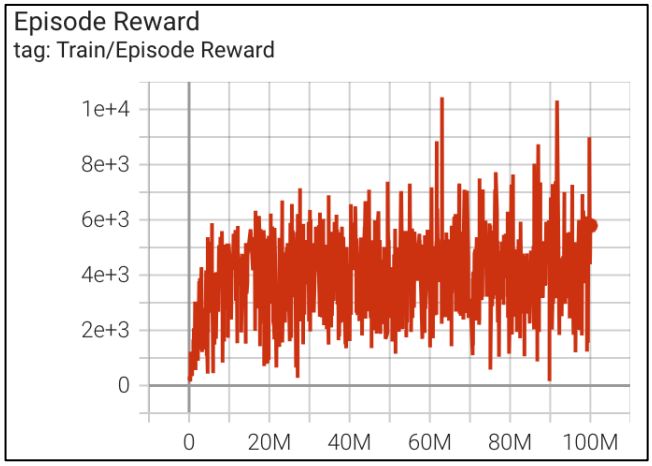
- Training curve (comparison)
  - DQN(Orange) and DDQN(Blue), with 0.999 smoothing

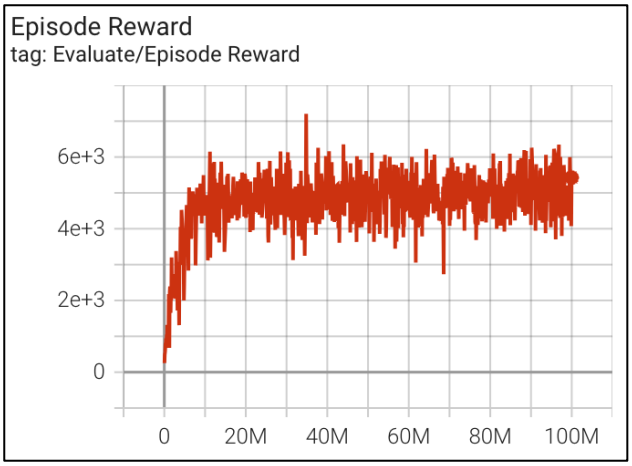




- Discuss between DQN(Orange) and DDQN(Blue), with 0.999 smoothing
  - DDQN 與 DQN 最主要的差別是 DDQN 是使用 behavior net 預估出來最高的 action，在將其用來取得 target net 的 Q(s, a)來更新參數，而不是單純使用 target net 預估出來最高的 Q(s, a)來更新，以此方式就可以防止在更新時由於預估誤差導致一直取到被高估的值，讓最後的整個模型訓練後整體皆會高估其結果。
  - 而在此任務上，兩者的 training reward 相差不多，然而在後期 Evaluate 時 DDQN 的表現更好，我認為是因為 DDQN 相比於 DQN，其於沒有經過訓練的畫面上預估的 Q(s, a) 更加準確，而 DQN 對於沒有被 sample 過的資料在估計上更加容易失準(易高估)，因此在 Evaluate 上 DDQN 的效果才會更好。

- Screenshot of Tensorboard training curve and testing results on Dueling DQN, and discuss the difference between DQN and Dueling DQN.
  - Training curve

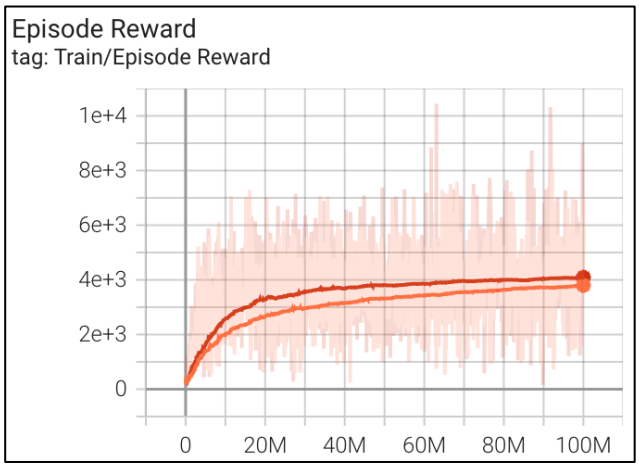


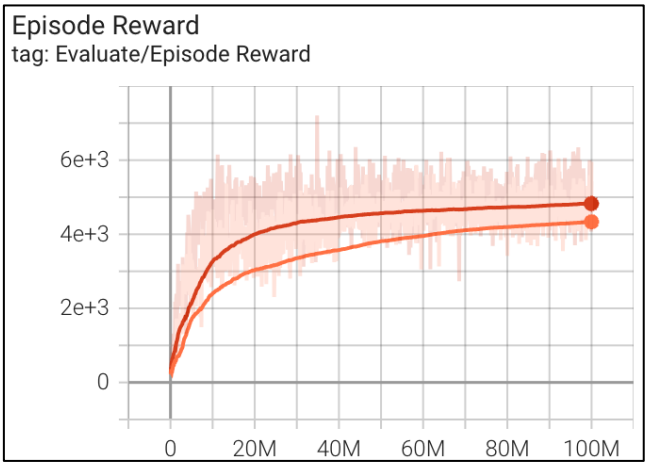


  - Testing results (5 games)

```
episode 1 reward: 5140.0
episode 2 reward: 4440.0
episode 3 reward: 5260.0
episode 4 reward: 8010.0
episode 5 reward: 6880.0
average score: 5946.0
```

  - Training curve (comparison)
    - DQN(Orange) and Dueling DQN(Red), with 0.999 smoothing





- Discuss between DQN and Dueling DQN
  - Dueling DQN 與 DQN 的差別僅在神經網路的設計，Dueling 將 Q value 分為了 V value 與 A value， 並用類似 regularization 的方式盡量讓所有 action 的 A 總和為 0，以此來防止模型可能只改動 A 的值而使 V 為 0，讓模型可以更有效率的學到哪個 action 在此 state 下更好。
  - 而在此實驗中，也如預期的，Dueling DQN 不管是在 training 與 evaluate 的結果相比於 DQN 皆好上不少，也代表使用 Dueling DQN 嘗試將學習的目標分為 V 與 A 真的加強了模型的學習，因此讓 reward 能夠更高。