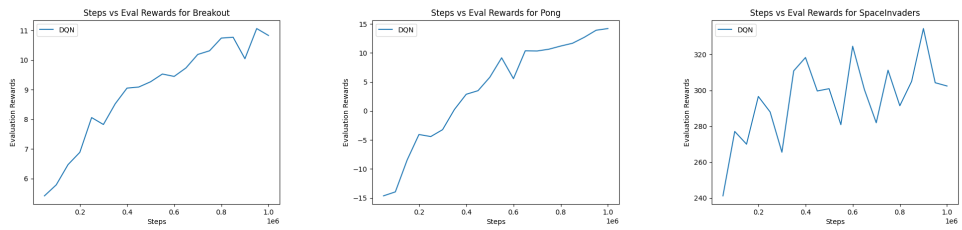
[CSCI-GA 3033-090]
Special Topics: Deep Reinforcement Learning
Homework - 2 Deep Q Learning++

Xu Cao
xc2057@nyu.edu

13 oktober 2021

Question 1

Finished, please check the code. See Fig.1.



(a) (b) (c)

Figur 1: DQN for (a) Breakout; (b) Pong; (c) Space Invaders

Question 2

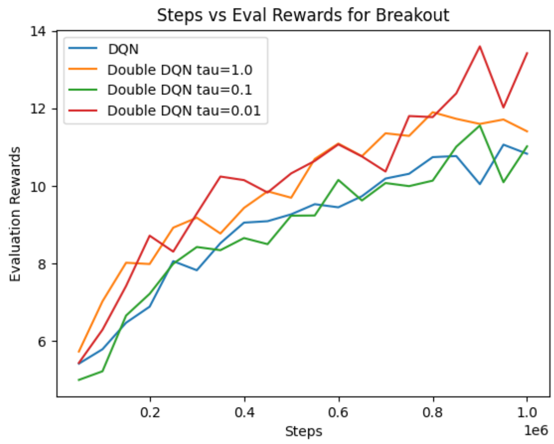
Finished, please check the code. First, I use Breakout to select the hyper-parameter tau for Double DQN[4]. I test tau = 1, tau = 0.1, tau = 0.01, and find that tau = 0.01 shows the best performance. See Fig.2.

Then, I use tau = 0.01 for all experiment that using Double DQN. The result for question 2 is showed in See Fig.3.

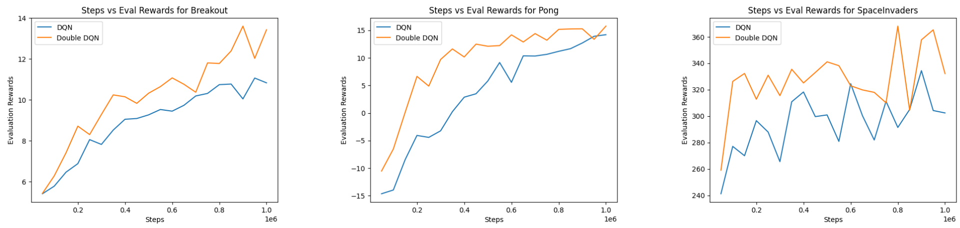
Question 3

Finished, please check the code. See Fig.4.

My finding is: for Breakout, the performance of Double DQN + prioritized experience replay is worse than Double DQN; for Pong and Space Invades, Double DQN + prioritized



Figur 2: Tau selection for Double DQN



(a) (b) (c)

Figur 3: DQN, Double DQN for (a) Breakout; (b) Pong; (c) Space Invaders. The orange line denotes Double DQN.

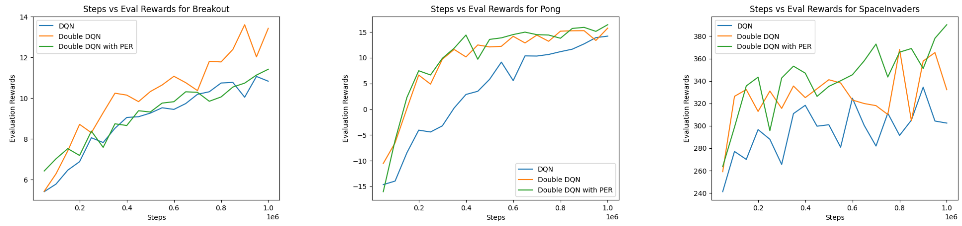
experience replay is better than Double DQN. And for all 3 games, Double DQN + prioritized experience replay converge faster than both Double DQN and DQN. I also find similar result from [3]'s Table 6. Double DQN + prioritized experience replay outperform DQN with uniform replay on most of games, but it still has 8 games that Double DQN + prioritized experience replay perform worse than DQN. I think this phenomenon is related to the propriety of the game. If the game is hard to train, it will need more exploration. Thus, prioritized experience replay would be not helpful for those games.

Question 4

Finished, please check the code. See Fig.5.

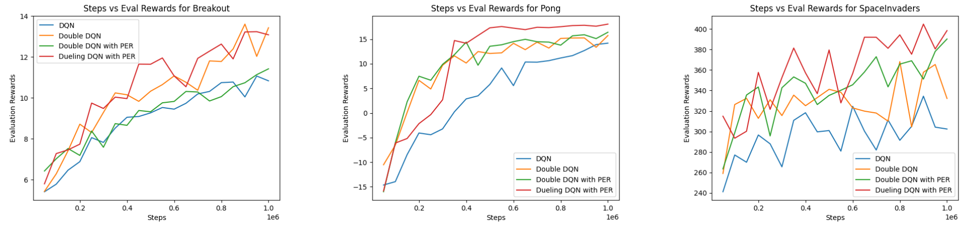
For Dueling DQN, I use the mean function as Rainbow suggest[1].

Fig.5 also show the comparison and improvement over vanilla DQN in different games.



(a) (b) (c)

Figur 4: DQN, Double DQN for (a) Breakout; (b) Pong; (c) Space Invaders. The green line denotes Double DQN + prioritized experience replay.



(a) (b) (c)

Figur 5: DQN, Double DQN for (a) Breakout; (b) Pong; (c) Space Invaders. The red line denotes Dueling DQN + prioritized experience replay.

Question 5 Bonus points

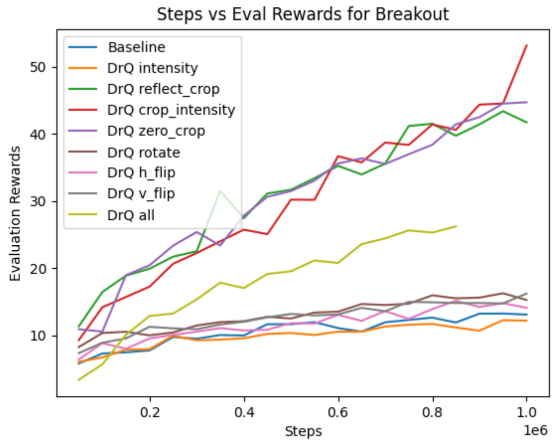
To use DrQ on top of DQN, we should obey three separate regularization mechanisms.

- transformations of the input image
- averaging the Q target over K image transformations
- averaging the Q function itself over M image transformations

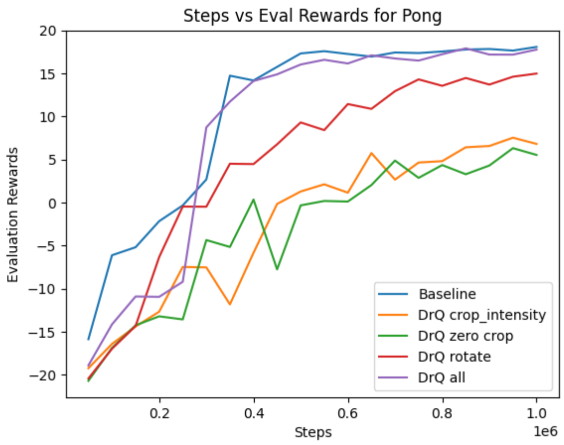
For our code, we should test the first part: transformations of the input image. The author test different transformation methods including (1) Random Shift; (2) Cutout; (3) Horizontal/Vertical Flip; (4) Rotate; and (5) Intensity. Finally, they concluded that random shifts strike a good balance between simplicity and performance[2]. Therefore they limit the choice of augmentation to this transformation.

For my experiment, I also choose Random Shift (name crop in the code) as my first choice for DrQ, and compare it with the Dueling DQN and several augmentation method mentioned in the paper. It shows that after using random shift as the data augmentation method, the final evaluation results improves a lot for Breakout (See Fig.6). Both reflect crop, crop intensity and zero crop perform well and finally achieve total reward >= 40.

But for Pong and Space Invaders, adding random shift only improve a little (see Fig.7, Fig.8). For Pong, it may become worse than the baseline model (Dueling DQN). I think it is related to the propriety of the game. If a game has many edge states like the agent stay at the edge of the game window, DrQ with random shift may cause some trouble.



Figur 6: Steps vs Evaluation Rewards for DrQ-Breakout

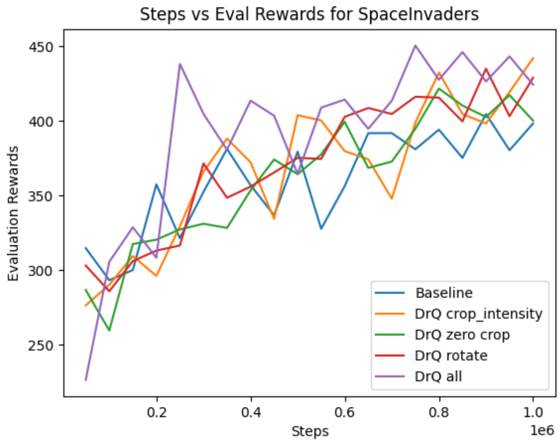


Figur 7: Steps vs Evaluation Rewards for DrQ-Pong

Referenser

[1] Matteo Hessel, Joseph Modayil, Hado Van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Azar, and David Silver. Rainbow: Combining improvements in deep reinforcement learning. In *Thirty-second AAAI conference on artificial intelligence*, 2018.

[2] Ilya Kostrikov, Denis Yarats, and Rob Fergus. Image augmentation is all you need: Regularizing deep reinforcement learning from pixels. *arXiv preprint arXiv:2004.13649*, 2020.



Figur 8: Steps vs Evaluation Rewards for DrQ-Space Invaders

[3] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay. *arXiv preprint arXiv:1511.05952*, 2015.

[4] Hado Van Hasselt, Arthur Guez, and David Silver. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30, 2016.