MLDS 4-2

組員：陳泓均、陳欽安、詹書愷、丁昱升

Model

DNN...

      Dense(84*84*4,128)

      ReLU()

      Dense(128, num_action)

Fail QAQ

      Reward ≅ 2

Model

Conv2d(4, 32, kernel_size=8, stride=4)

Conv2d(32, 64, kernel_size=4, stride=2)

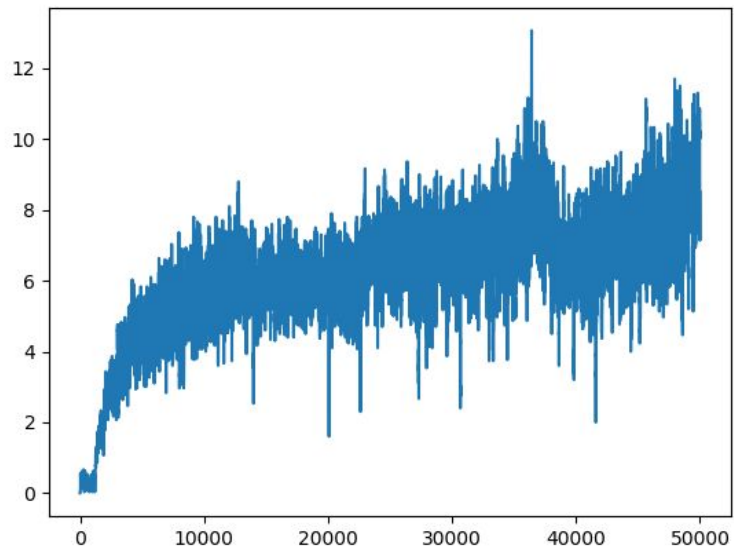Conv2d(64, 64, kernel_size=3, stride=1)

Flatten()

Dense(7*7*64, 512)

ReLU()

Dense(512, num_action)

(GAMMA = 0.999, Adam optimizer)
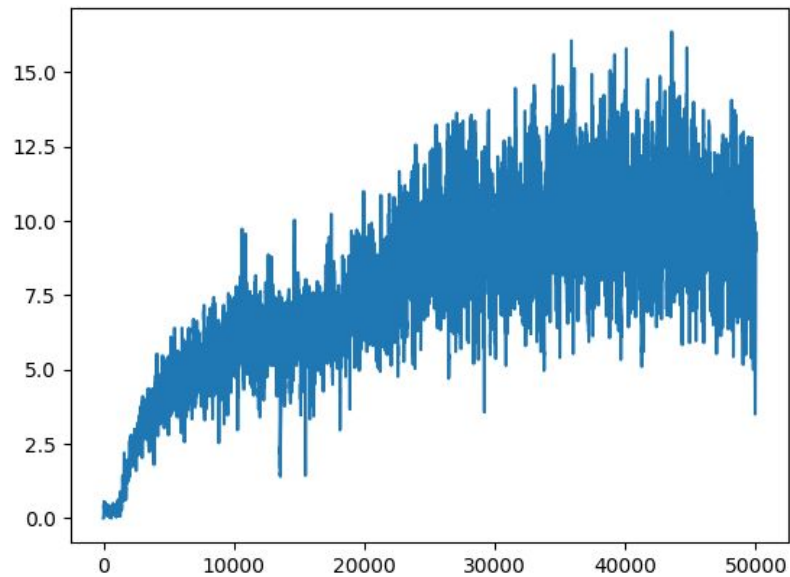
Reward

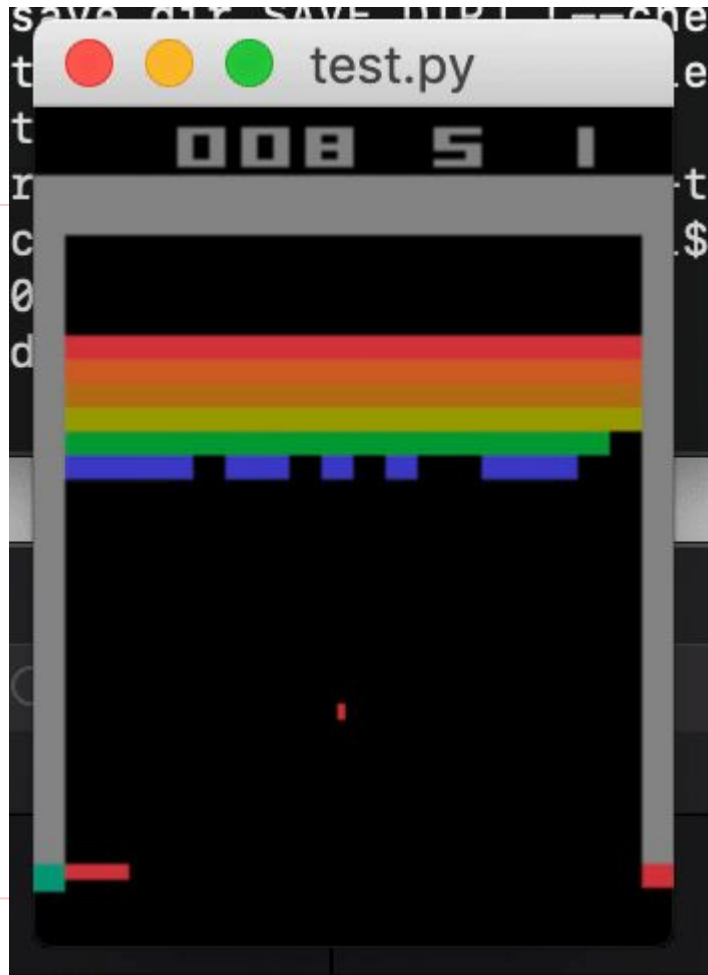# Test (50000 episode)

Score : 32.79   :(

Tips - Double DQN

$$Y_t^{\text{DoubleDQN}} \equiv R_{t+1} + \gamma Q(S_{t+1}, \operatorname*{argmax}_a Q(S_{t+1}, a; \boldsymbol{\theta}_t), \boldsymbol{\theta}_t^-).$$
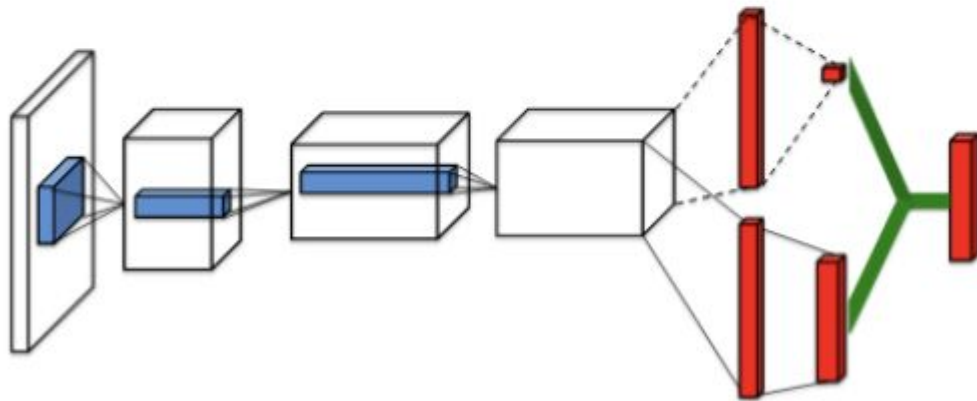
Reward

Test (50000 episode)

Score : 49.2

Tips - Dueling DQN

Reward

Test

Score : 73.47 !!!

- 卡住?!

Tips - Double + Dueling DQN

$$Y_t^{\text{DoubleDQN}} \equiv R_{t+1} + \gamma Q(S_{t+1}, \underset{a}{\arg\max} Q(S_{t+1}, a; \boldsymbol{\theta}_t), \boldsymbol{\theta}_t^-).$$
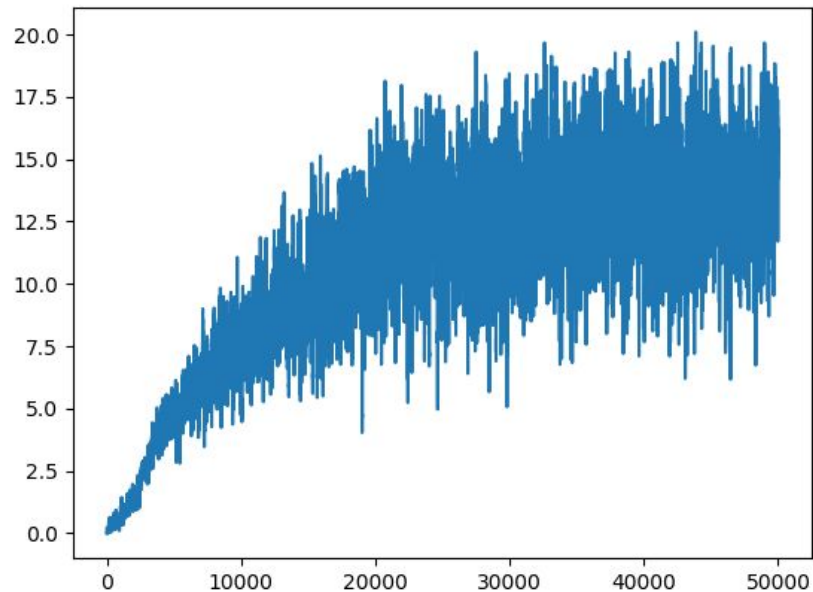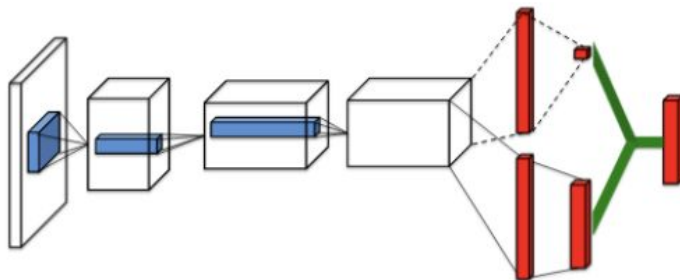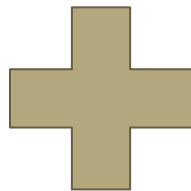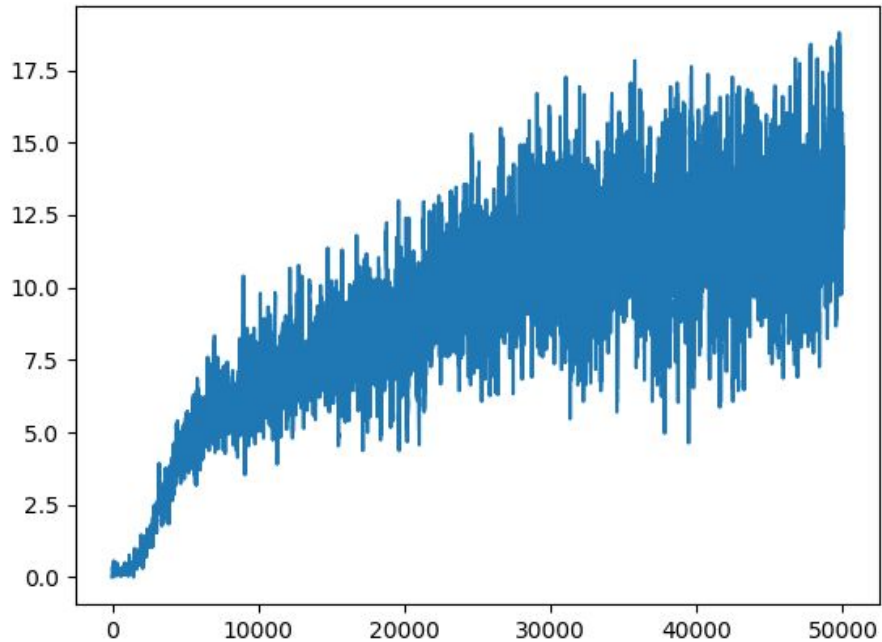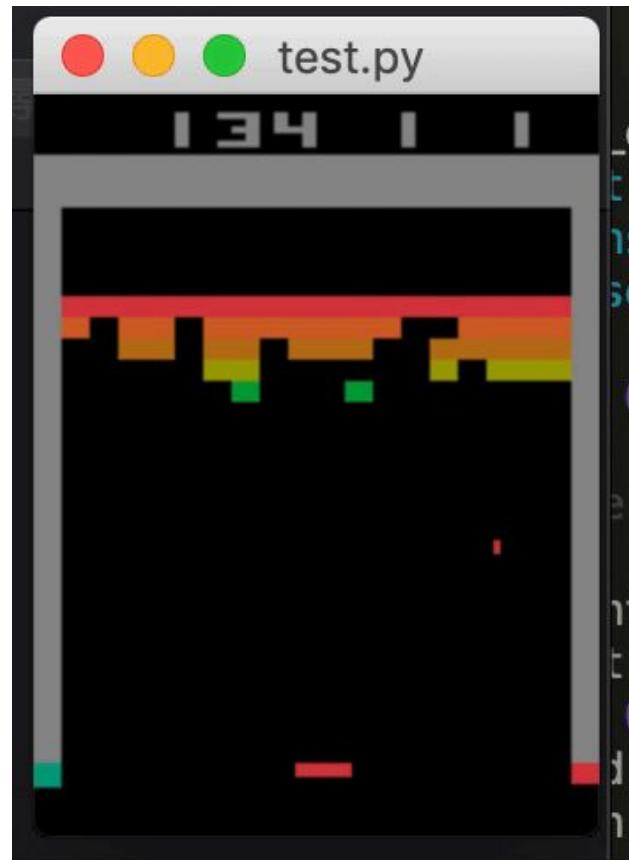
Reward

Test (50000 episode)

Score : 69.25 !

Conclusion

- DNN sucks!
  - Without subtraction
  - Need to find features in images
- Tips really help!
  - But reward oscillate more?
- Dueling is better than Dueling + Double ?
  - Not significant, maybe just error

Thanks for Listening