Deep Learning Homework 3
0860908 李少琪

1. Generative Adversarial Network (GAN)

   ☐ Data Augmentation

   Resize : 32 pixel

   > 減少資料量，加速 GAN 運算，但可能缺失資料

   Center Crop: 32 pixel

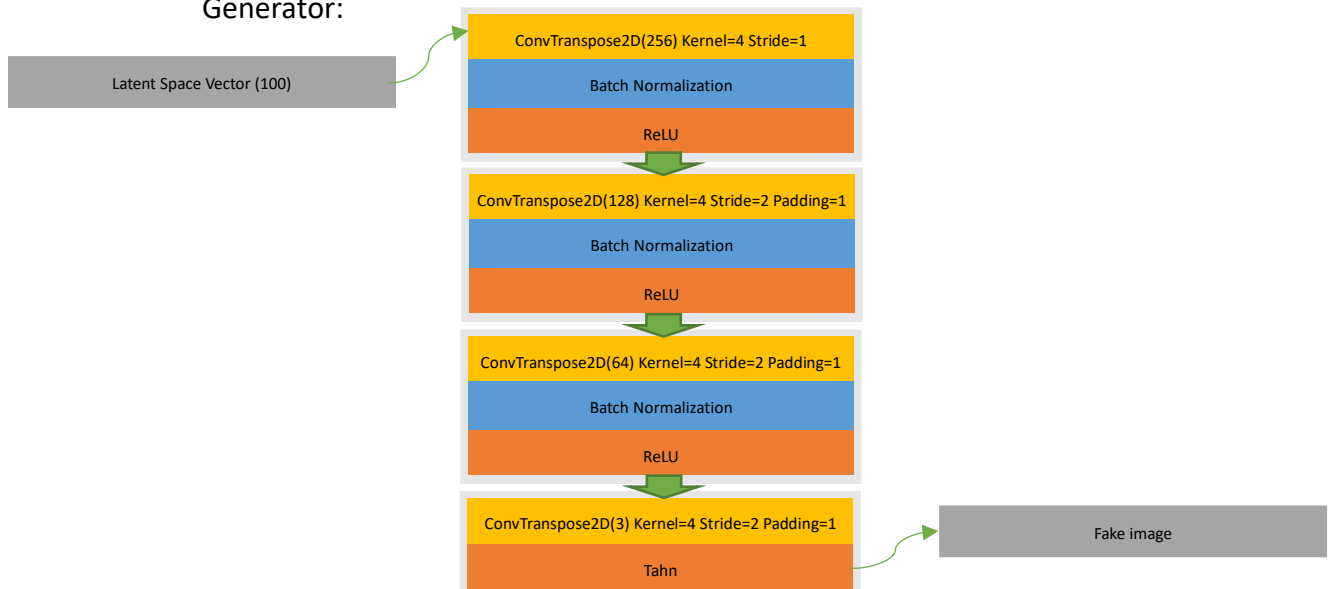   Random Horizontal Flip 隨機水平翻轉 : probability = 0.5

   > 因為人臉為對稱的，左右相反依然很容易就看出是人臉

   Normalize 歸一化 -> [-1, 1]

   > Mean = (0.5, 0.5, 0.5)
   >
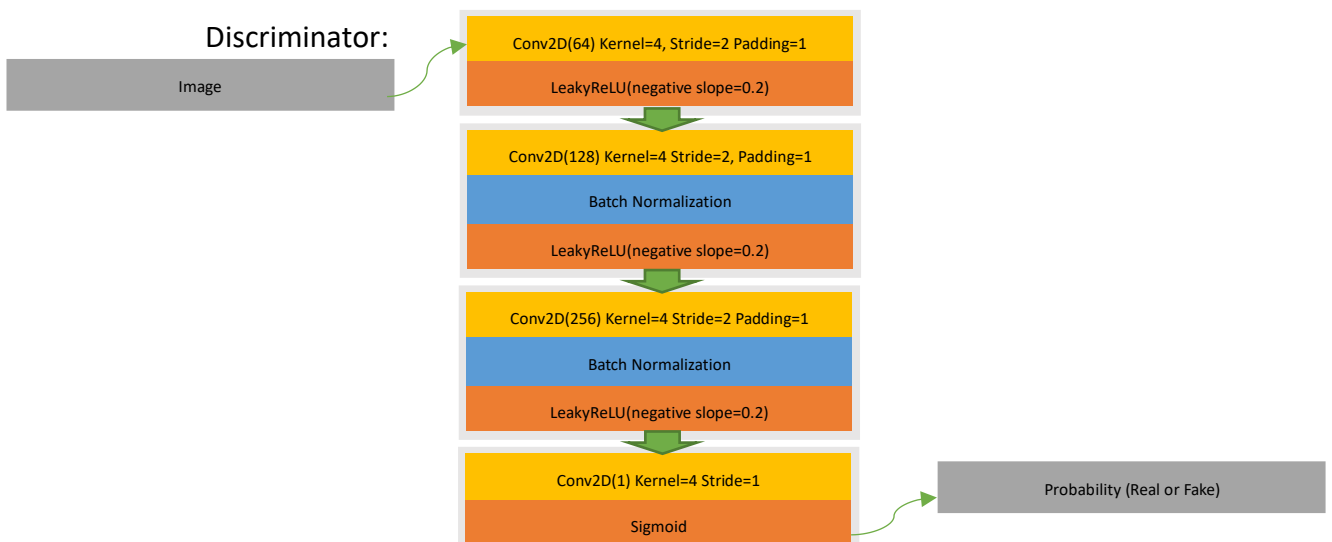   > Std = (0.5, 0.5, 0.5)

   ☐ Model Architecture

   Generator:

將原始網路架構減少一層，因為 image size 減少(feature 減少), 所以減少深度，也加速運算。

   Discriminator:
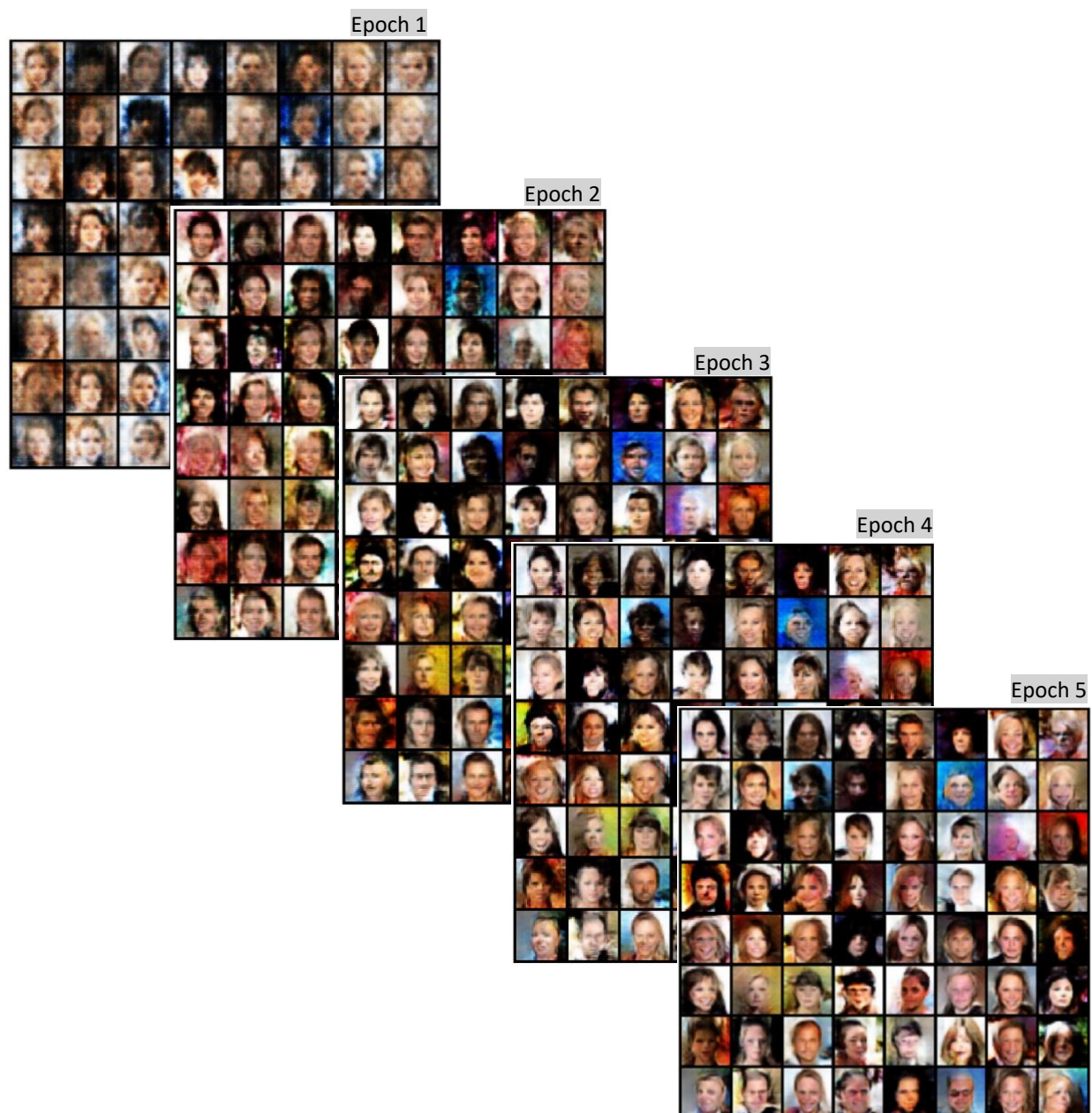
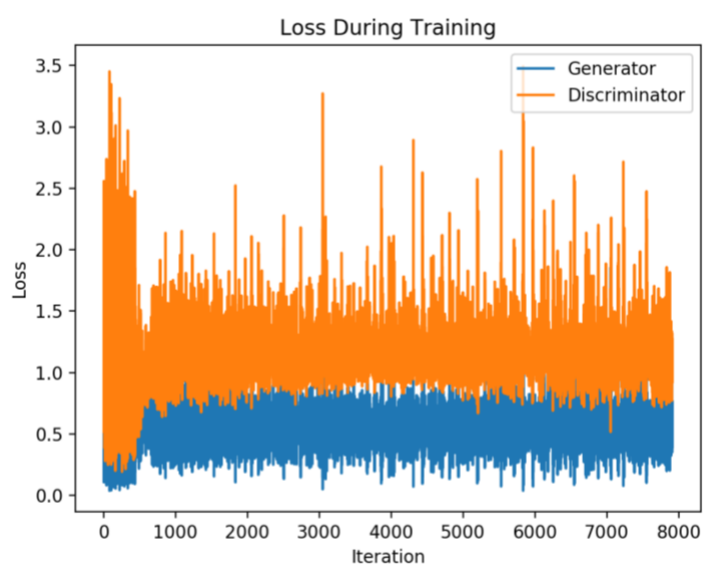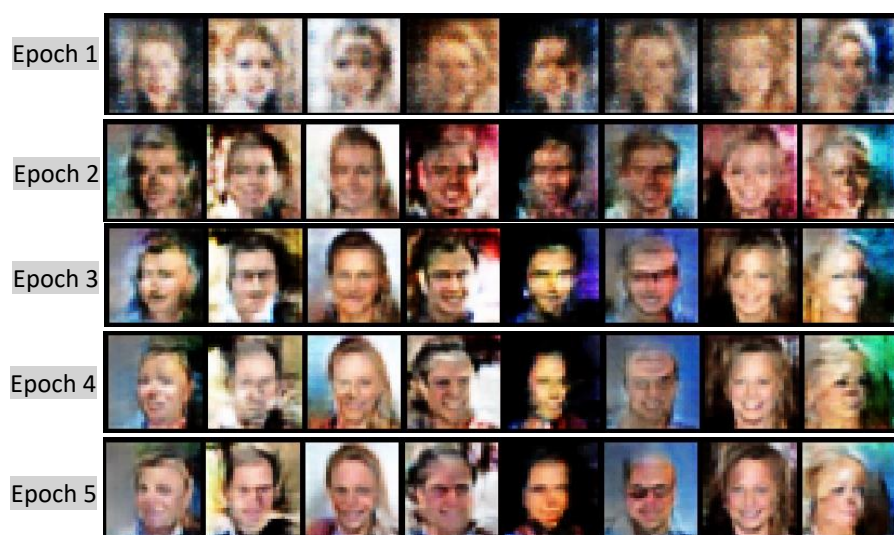也是將原始網路架構減少一層，減少深度，也加速運算。

Weight 初始化:

    Randomly initialization from normal distribution with mean=0, stdev=0.02

☐ Loss Function

Binary cross entropy: 讓 Discriminator 學習判斷真的照片(1)和假的照片(0)，讓 Generator 往真的照片(1)學習

☐ Result

Epoch 1

Epoch 2

Epoch 3

Epoch 4

Epoch 5

Loss During Training

□ Discussion

Real Image

Generator → Fake Image → Discriminator → Label
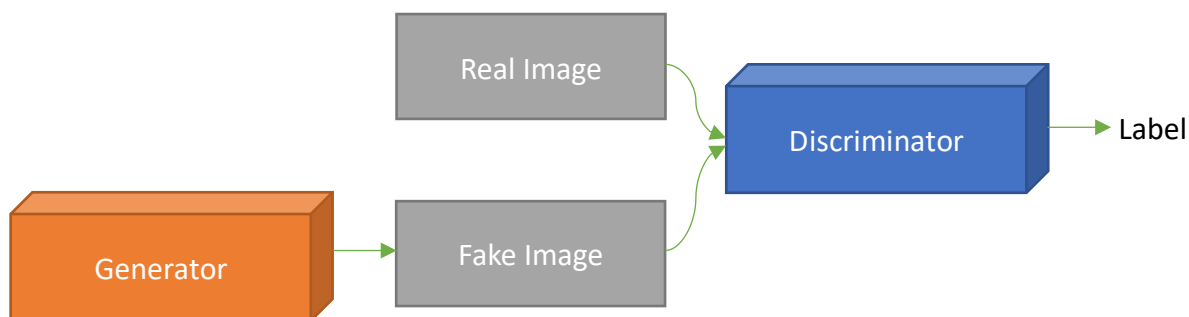
在做 Back propagation 時，對於哪些部分需要作反向傳播，而哪些不需要，搞混了好一陣子，導致 Generator 完全沒被學習到。訓練時，我發現 Loss 的走向不太理想，Generator 的 loss 幾乎沒變，Discriminator 也是，但其中 Discriminator 對於 Real images 的 Loss 下降快速，但是對於

Fake images 的 Loss 變化幾乎不變，才發現在 Fake images 這部分並沒有被 Update，導致只有對 Discriminator 訓練 real images。
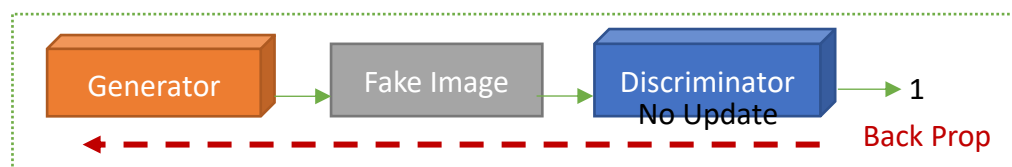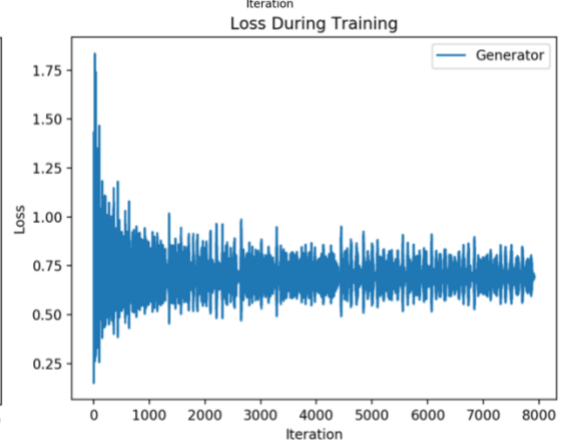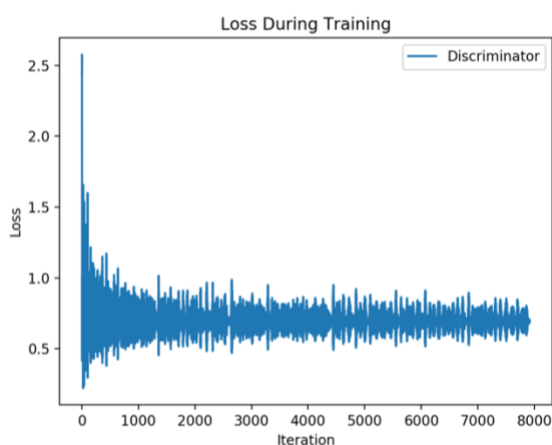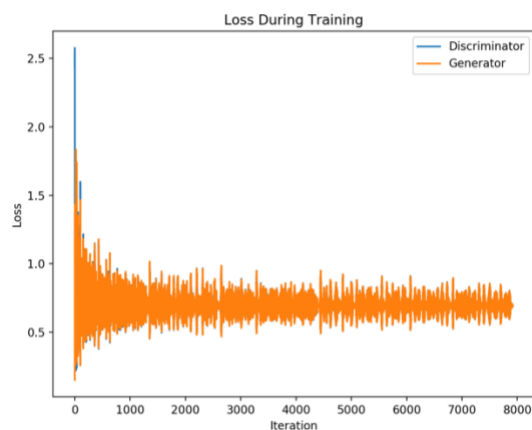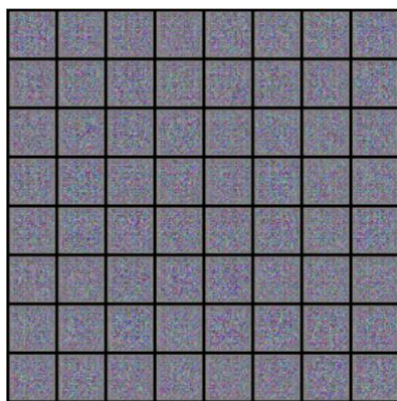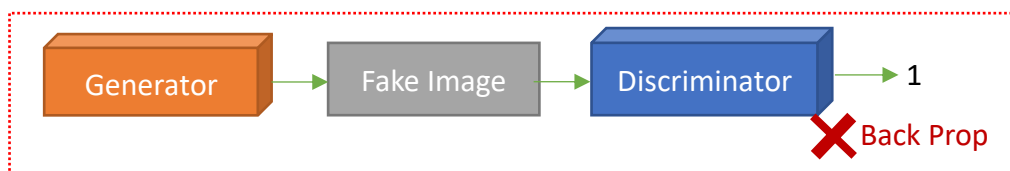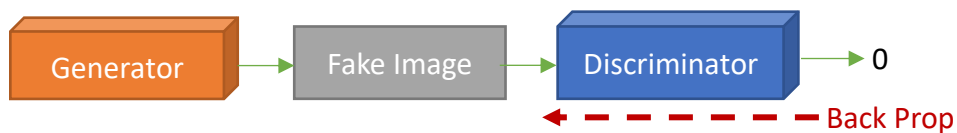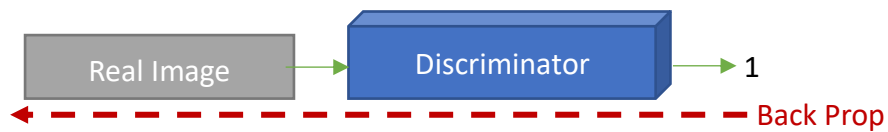
2. Deep Q Network (DQN)

☐ Learning

a. Probability of random agent: NOOP(0.33), UP(0.33), DOWN(0.33)

```
Episode:        1, interaction_steps:   2048, reward:  0, epsilon: 0.999493
[Info] Save model at './drive/My Drive/Colab Notebooks/DL_HW3/model' !
Evaluation: True, episode:        1, interaction_steps:    2048, evaluate reward:  0
Episode:        2, interaction_steps:   4096, reward:  0, epsilon: 0.998986
Episode:        3, interaction_steps:   6144, reward:  0, epsilon: 0.998479
Episode:        4, interaction_steps:   8192, reward:  0, epsilon: 0.997972
Episode:        5, interaction_steps:  10240, reward:  0, epsilon: 0.997466
Episode:        6, interaction_steps:  12288, reward:  0, epsilon: 0.996959
Episode:        7, interaction_steps:  14336, reward:  0, epsilon: 0.996452
Episode:        8, interaction_steps:  16384, reward:  0, epsilon: 0.995945
Episode:        9, interaction_steps:  18432, reward:  0, epsilon: 0.995438
Episode:       10, interaction_steps:  20480, reward:  0, epsilon: 0.994931
Episode:       11, interaction_steps:  22528, reward:  0, epsilon: 0.994424
Evaluation: True, episode:       11, interaction_steps:  22528, evaluate reward:  0
```
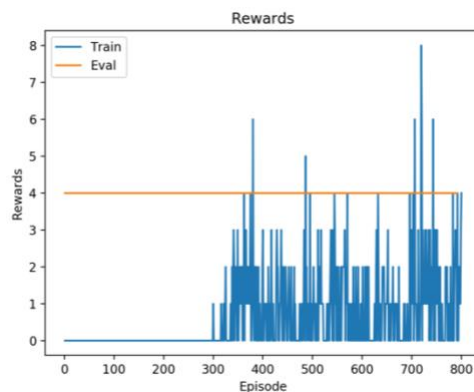

三種動作機率一樣時，因隨機移動難在遊戲中得分，所以學習較緩慢。

b. Probability of random agent: NOOP(0.3) UP(0.6) DOWN(0.1)

epsilon_decay = 1000000

```
Use device: cuda
Episode:        1, interaction_steps:   2048, reward: 12, epsilon: 0.997972
[Info] Save model at './drive/My Drive/Colab Notebooks/DL_HW3/model' !
Evaluation: True, episode:        1, interaction_steps:    2048, evaluate reward:  0
Episode:        2, interaction_steps:   4096, reward: 12, epsilon: 0.995945
Episode:        3, interaction_steps:   6144, reward: 12, epsilon: 0.993917
Episode:        4, interaction_steps:   8192, reward: 11, epsilon: 0.991890
Episode:        5, interaction_steps:  10240, reward: 11, epsilon: 0.989862
Episode:        6, interaction_steps:  12288, reward: 10, epsilon: 0.987835
Episode:        7, interaction_steps:  14336, reward: 11, epsilon: 0.985807
Episode:        8, interaction_steps:  16384, reward: 12, epsilon: 0.983780
Episode:        9, interaction_steps:  18432, reward: 11, epsilon: 0.981752
Episode:       10, interaction_steps:  20480, reward:  9, epsilon: 0.979725
Episode:       11, interaction_steps:  22528, reward:  9, epsilon: 0.977697
Evaluation: True, episode:       11, interaction_steps:  22528, evaluate reward:  0
```

因為連線斷掉為能記錄 log，但從 output 中可以發現 reward 並沒有上升反而下降。

```
Evaluation: True, episode:      691, interaction_steps: 1415168, evaluate reward:  0
Episode:      692, interaction_steps: 1417216, reward:  0, epsilon: 0.010000
Episode:      693, interaction_steps: 1419264, reward:  0, epsilon: 0.010000
Episode:      694, interaction_steps: 1421312, reward:  0, epsilon: 0.010000
Episode:      695, interaction_steps: 1423360, reward:  6, epsilon: 0.010000
Episode:      696, interaction_steps: 1425408, reward:  1, epsilon: 0.010000
Episode:      697, interaction_steps: 1427456, reward:  0, epsilon: 0.010000
Episode:      698, interaction_steps: 1429504, reward:  0, epsilon: 0.010000
Episode:      699, interaction_steps: 1431552, reward:  0, epsilon: 0.010000
Episode:      700, interaction_steps: 1433600, reward:  3, epsilon: 0.010000
Episode:      701, interaction_steps: 1435648, reward:  4, epsilon: 0.010000
```

猜測是 epsilon 下降太快，導致 Policy Network 還未能在隨機得分中學習到得分方式，就一直利用 Target Network 中選擇下一個 action，因此我將 epsilon decay 調大為 4000000。

```
Use device: cuda
Episode:      1, interaction_steps:   2048, reward: 14, epsilon: 0.997972
[Info] Save model at './drive/My Drive/Colab Notebooks/DL_HW3/model' !
Evaluation: True, episode:      1, interaction_steps:   2048, evaluate reward:  0
Episode:      2, interaction_steps:   4096, reward: 10, epsilon: 0.995945
Episode:      3, interaction_steps:   6144, reward: 13, epsilon: 0.993917
Episode:      4, interaction_steps:   8192, reward: 13, epsilon: 0.991890
Episode:      5, interaction_steps:  10240, reward: 11, epsilon: 0.989862
Episode:      6, interaction_steps:  12288, reward: 10, epsilon: 0.987835
Episode:      7, interaction_steps:  14336, reward:  9, epsilon: 0.985807
Episode:      8, interaction_steps:  16384, reward: 11, epsilon: 0.983780
Episode:      9, interaction_steps:  18432, reward:  9, epsilon: 0.981752
Episode:     10, interaction_steps:  20480, reward: 10, epsilon: 0.979725
Episode:     11, interaction_steps:  22528, reward: 11, epsilon: 0.977697
```

但跑了好幾個小時，Colab 依然又斷線了，以下為跑到一半的 output。到了六白多 episode，epsilon 仍然有 0.68。

```
Evaluation: True, episode:    601, interaction_steps: 1230848, evaluate reward:  0
Episode:    602, interaction_steps: 1232896, reward:  5, epsilon: 0.694858
Episode:    603, interaction_steps: 1234944, reward:  5, epsilon: 0.694351
Episode:    604, interaction_steps: 1236992, reward:  5, epsilon: 0.693844
Episode:    605, interaction_steps: 1239040, reward:  4, epsilon: 0.693338
Episode:    606, interaction_steps: 1241088, reward:  6, epsilon: 0.692831
Episode:    607, interaction_steps: 1243136, reward:  7, epsilon: 0.692324
Episode:    608, interaction_steps: 1245184, reward:  4, epsilon: 0.691817
Episode:    609, interaction_steps: 1247232, reward:  5, epsilon: 0.691310
Episode:    610, interaction_steps: 1249280, reward:  8, epsilon: 0.690803
Episode:    611, interaction_steps: 1251328, reward:  5, epsilon: 0.690296
Evaluation: True, episode:    611, interaction_steps: 1251328, evaluate reward:  3
Episode:    612, interaction_steps: 1253376, reward:  5, epsilon: 0.689789
Episode:    613, interaction_steps: 1255424, reward:  6, epsilon: 0.689283
Episode:    614, interaction_steps: 1257472, reward:  9, epsilon: 0.688776
Episode:    615, interaction_steps: 1259520, reward:  6, epsilon: 0.688269
Episode:    616, interaction_steps: 1261568, reward:  5, epsilon: 0.687762
Episode:    617, interaction_steps: 1263616, reward:  6, epsilon: 0.687255
Episode:    618, interaction_steps: 1265664, reward:  7, epsilon: 0.686748
```

已經用了好幾個不同的帳號跑，但都因為連續跑了太多次的 GPU，導致無法分配到 GPU。