

- Model description(2%) Must include model structure, objective function for G and D
  - Preprocessing
 

```
img_shape = (96,96)
```

one-hot feature vector

```
hair_color = [
    'orange hair', 'white hair', 'aqua hair', 'gray hair',
    'green hair', 'red hair', 'purple hair', 'pink hair',
    'blue hair', 'black hair', 'brown hair', 'blonde hair'
]
```

```
eyes_color = [
    'gray eyes', 'black eyes', 'orange eyes',
    'pink eyes', 'yellow eyes', 'aqua eyes', 'purple eyes',
    'green eyes', 'brown eyes', 'red eyes', 'blue eyes'
]
```

```
hair_style = [long hair', short hair']
```
  - Generator:
 

```
seq_len = 100
noise_len = 100
Concat(seq, noise)
Deconv2D(512, 6, 1, bias=False, act=relu)
Deconv2D (256, 4, 2, bias=False, act=relu)
Deconv2D (128, 4, 2, bias=False, act=relu)
Deconv2D (64, 4, 2, bias=False, act=relu)
Deconv2D (3, 4, 2, bias=False, act=tanh) // (96,96,3)
```
  - Discriminator
 

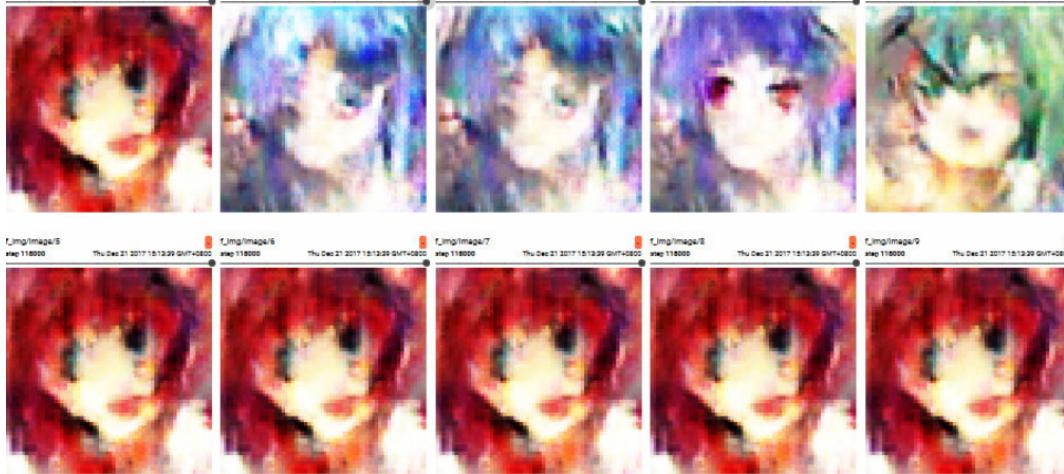
```
Image(shape=(96,96,3))
Conv2D(64, 4, 2, bias=False, act=lrelu)
Conv2D(128, 4, 2, bias=False, act=lrelu)
Conv2D(256, 4, 2, bias=False, act=lrelu)
Conv2D(512, 3, 2, bias=False, act=lrelu)
ConcatCondition(seq_len=100)
Conv2D(512, 1, 1, bias=False, act=lrelu)
Conv2D(1, bias=False, act=sigmoid)
```
  - Objective Function
 

```
g_loss = log(1-D(G(z)))
d_loss = log(1-D(real_img, real_seq))
    + log(D(G(z), real_seq))/3
    + log(D(wrong_img, real_seq))/3
    + log(D(real_img, wrong_seq))/3
```
  - Optimizer
 

```
g_optimizer = Adam(lr=0.0002, beta1=0.5)
d_optimizer = Adam(lr=0.0001, beta1=0.5)
```

- How do you improve your performance (2%)

一開始使用 64\*64 的圖片和 noise dimension=20 來 train，雖然能夠 train 起來但是圖片相當的模糊，且多樣性很低。如下



(圖，上面五張為不同 condition，下面五張為相同 condition)

嘗試使用 96\*96 的圖片來訓練，輸出的時候再 resize 到 64\*64，並且把 generator 第一層 hidden 由 256 改為 512 後，訓練出來的圖片變得相當清楚，但是仍然有多樣性不足的問題，推測可能為 noise 影響太小。如下

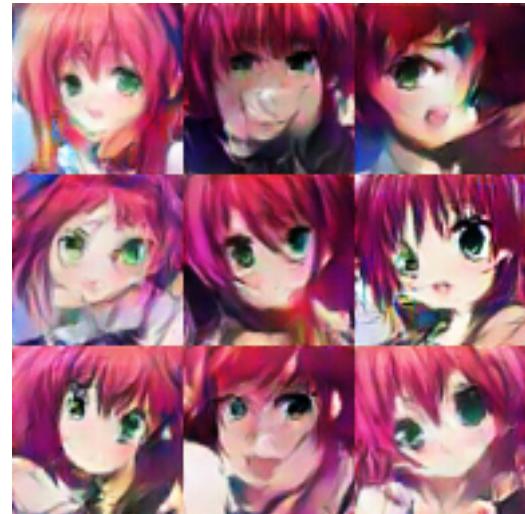


(圖，九張相同 condition)

後來嘗試把 noise dimension 加大到 100 維後 collapse mode 的問題就解決了，基本上五張圖片能產生出足夠的差異。



(圖，condition 紅髮綠眼)



(圖，condition 紅髮綠眼)

- Experiment settings and observation (2%)

實驗過換成 WGAN 的 Objective function 來做訓練，但是一直 train 不太起來，圖片的臉一直只有輪廓，但是五官很模糊，不太確定是什麼原因，可能寫壞了...

另外還有嘗試過使用不同的 optimizer 來訓練，做了兩種不同的實驗，替換為 Adam(beta=0.9) 跟 RMSProp(lr=0.0001) 但是肉眼看最終產生出來的圖片感覺是差不多的，收斂速度也沒有明顯的差異。

實驗過 generator 最後一層使用 tanh 和 sigmoid，發現使用 tanh 的話會稍微收斂的比較快一點，應該是因為 tanh 會讓輸出在 [-1,1] 之間，且平均為 0。

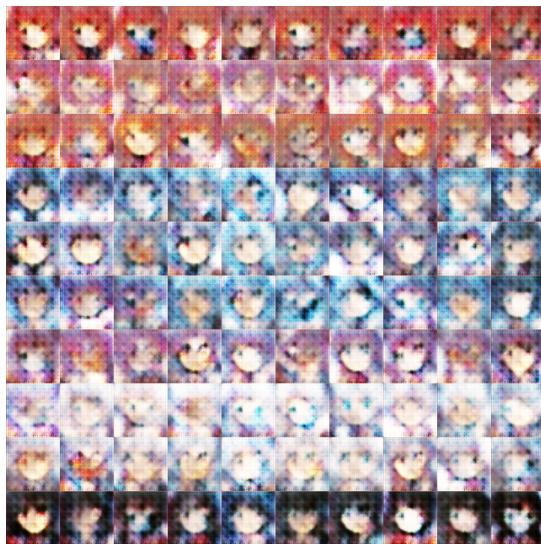
另外也實驗了加入 long hair 跟 short hair 兩個 condition 去 train，但發現似乎沒有辦法很有效的學到這兩個 condition，有時候輸入 long hair 產生出來的頭髮仍然很像短髮，感覺由於 training data 只有臉的部分，頭髮被截掉很多，所以很難區分出來長短髮。

最後調整以後得到產生效果最好的模型參數如第一題，batch size 為 64，訓練時間約 4 小時。

以下是圖片是 training 過程產生的圖片。

Row 為不同顏色的頭髮 condition，Col 為不同顏色的眼睛 condition。

可以看到大概在 1000 次 update 時就能產生出模糊的人臉圖案。10000 次 update 後產生出來的圖片已經相當不錯，但如果訓練太久到了大約 30000 update 後就會出現嚴重的 collapse mode，雖然產生的圖片依然滿足 condition 但是幾乎所有人物都長得一模一樣。



(圖 , 1000 update)



(圖 , 3000 update)



(圖 , 5000 update)



(圖 , 10000 update)



(圖 , 20000 update)



(圖 , 30000 update)