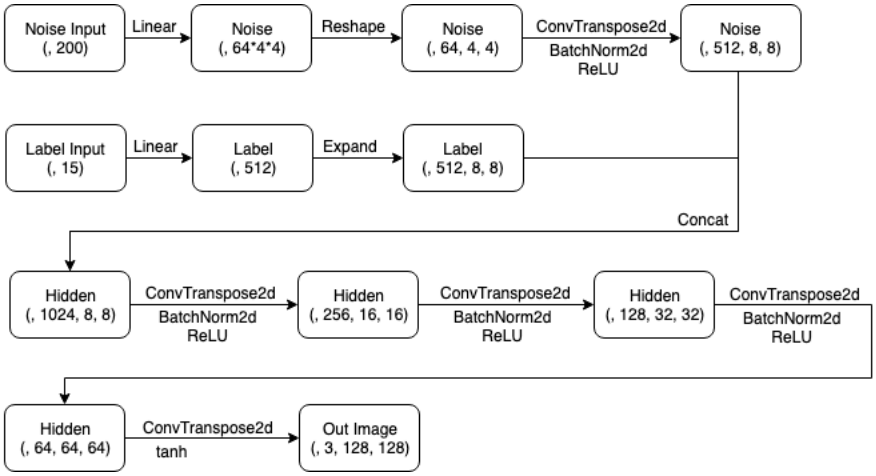
ADL Homework 4 Report

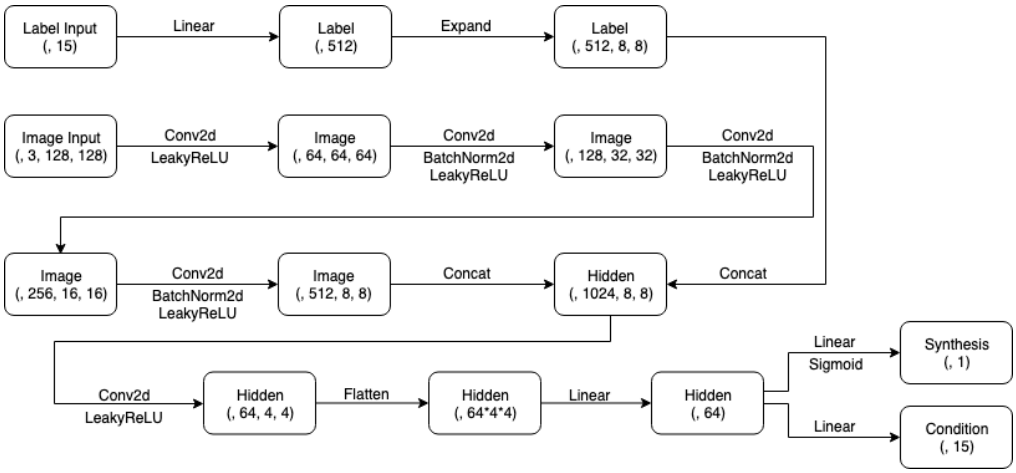
b05902002 資工三 李栢淵

1. Explain the structure of your networks and loss terms in detail

Structure of Generator



Structure of Discriminator



Other Properties

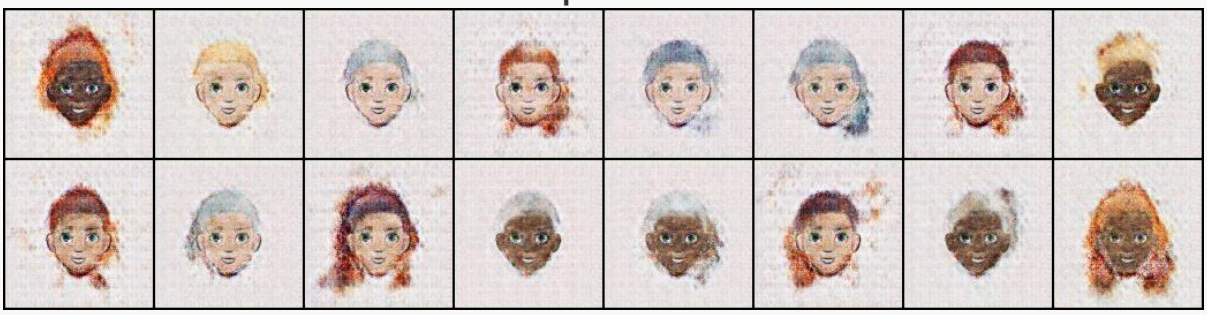
- Loss
  - Binary cross entropy on synthesis
  - Binary cross entropy on condition
- Optimizer: Adam
  - learning rate: 2e-4

    - beta1: 0.5
- Noise size: 200
- Batch size: 16
- MAGIC random seed: 1126

註：上述的model模型及參數為最早通過 strong baseline之版本

2. Plot your training progress (10 pics)

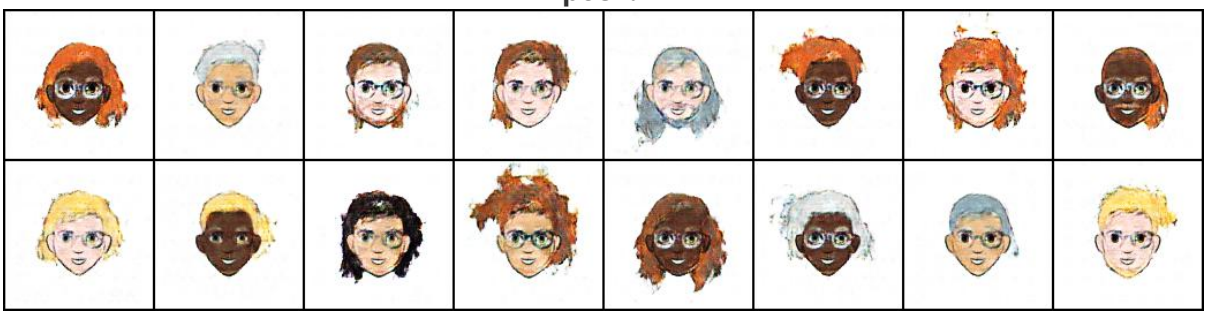
Epoch: 1



Epoch: 2

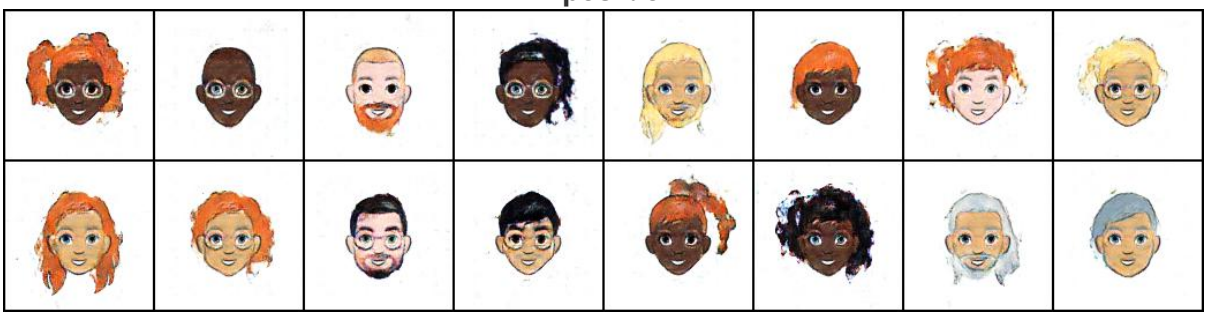
Epoch: 3

Epoch: 4

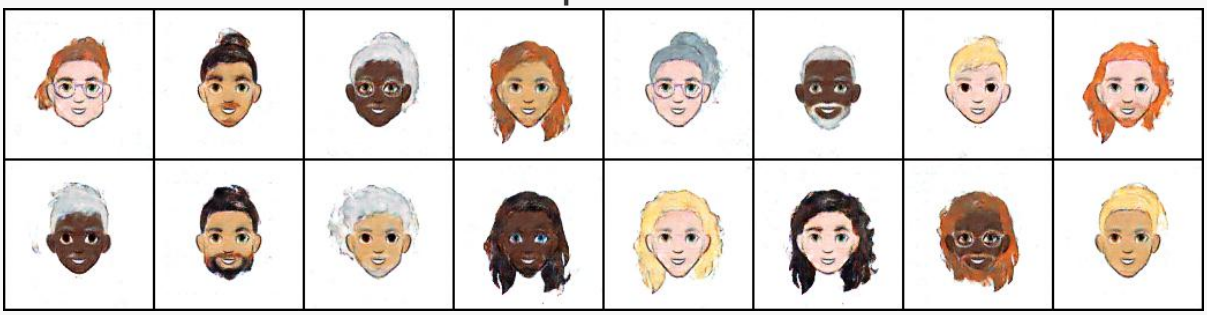


Epoch: 5

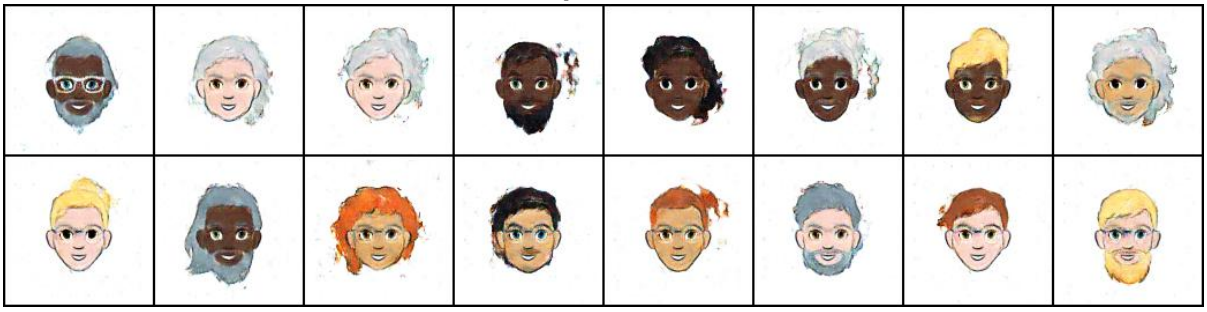
Epoch: 6



Epoch: 7



Epoch: 8



Epoch: 9

Epoch: 10

3. Design at least 3 different experiments. Describe your settings, making comparisons and report your observations.

- You may be deducted some points if TAs consider your experiments “太過敷衍”
- EX: Normalizing the input image to [-1,1] is considered too simple and basic
- 幾個參考方向: GAN的不同objective、model condition on label 的方式、各種tricks .

Experiment1: Different batch size

- Batch size: 16
  - Gradient Explode: @69 epoch
  - FID score: 63.257 (65 epoch)
- Batch size: 8
  - Gradient Explode: @62 epoch
  - FID score: 58.160 (60 epoch)
- Batch size: 4
  - Gradient Explode: 未發生
  - FID score: 27.616 (100 epoch)

發現：

1. **Batch Size: 16 v.s. Batch Size: 8**

   如下表格所示：

Batch Size: 16	Batch Size: 8
	

可以發現，兩者最大的差異在於：眼鏡的明顯程度。

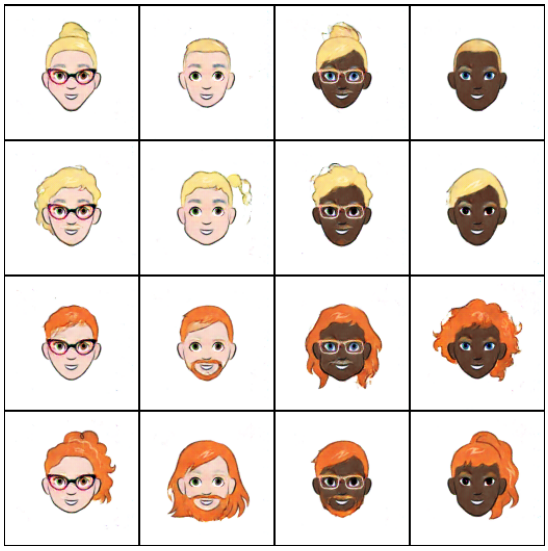
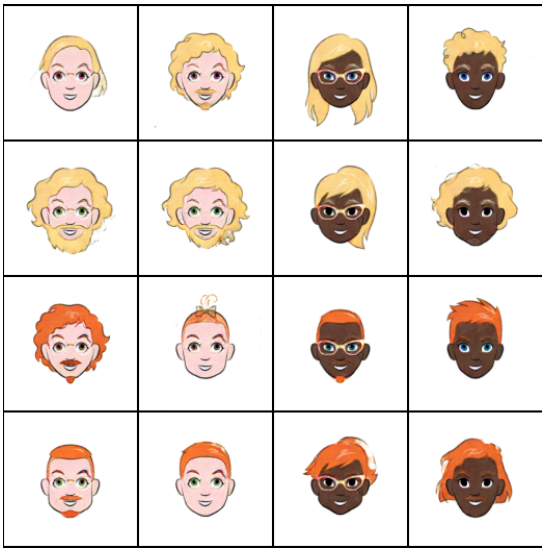
對於這種比較小的feature，比較小的batch size在train的過程迭代較慢，但卻有比較精細的結果。

我也發現右邊的藍色眼睛Train的比較明顯。

2. **Batch Size: 8 v.s. Batch Size: 4**

結果出來的時候讓我感到非常驚訝，FID score竟然降到快過Bonus baseline。

隨後把圖印出來之後，我發現越小的batch size對於細節train得越好，在Batch Size: 4時，原本train不太好的不同眼睛顏色（尤其綠色、棕色），也變得相當明顯，如下表格：

Batch Size: 8	Batch Size: 4
	

註：後來的實驗都是建立在Batch size: 4的條件下

Experiment2: Different Distribution of Noise

On the model described in Problem 1

Settings

- Batch size: 4
- Learning Rate: 2e-4
- Run 100 epochs

Compared

- Normal Noise (Original)
  - FID Score: 27.616
- Bernoulli Noise
  - FID Score: 164.304

The output of image looks really not okay ...

Experiment3: Different Normalization

3-1. Spectral Normalization

Result: Good 感覺沒有特別的imprvement，但是能train起來，圖片還蠻漂亮

FID Score: 44.702

3-2. Layer Normalization

Result: Fail

圖片完全生出一樣的，Train不起來，附上偏爛的圖，爛到不想算FID。

Spectral Normalization	Layer Normalization
	

Experiment4: Different Loss

Settings

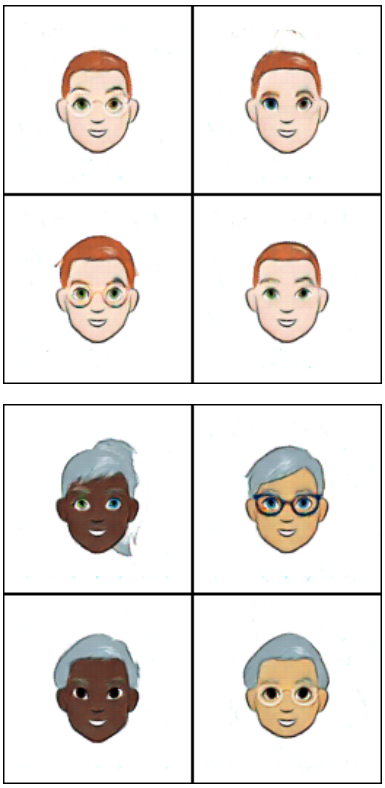
- Batch size: 4
- Learning Rate: 2e-4
- Both used Gradient Penalty

Compared

- DCGAN (Original)
    - FID Score: 27.616 (100 epochs)
- WGAN
    - FID Score: 57.134 (170 epochs)

The output looks okay

在Run 100 epochs版本的WGAN Loss時，發現其FID score只有82.667。其中蠻特別的是有生出異色瞳的照片，覺得還蠻有趣的。



Wgan 的Loss其實不太好train，很努力的讓他好一點了QQ。

Experiment5: Improve Generator

這部分指簡略敘述，因為也Train不起來。我嘗試把Generator的架構換成Residual Network，想看看有沒有比較好的結果。

雖然最後圖片是還可以，只是condition的feature整個出不來，出現很多多重髮色，整個performance差很多。

但是model的結構沒有多試幾種，可能調一下有機會變好。

3. Summary

總結來說，其他的improvement都沒有得到比最初model還要好的結果。

原本的model神秘地把DCGAN的Loss跟Gradient Penalty結合在一起...這樣得到了極好的結果，而且把GP拔掉還train不太起來...

而且我還leak了label的imformation給Discriminator，等於讓Discriminator變得超強。

結果就還蠻鬧的，大概總結了一下最好的model的方法。

- ACGAN
- Gradient Penalty
- Batch Size: 4
- Noise with normal distribution