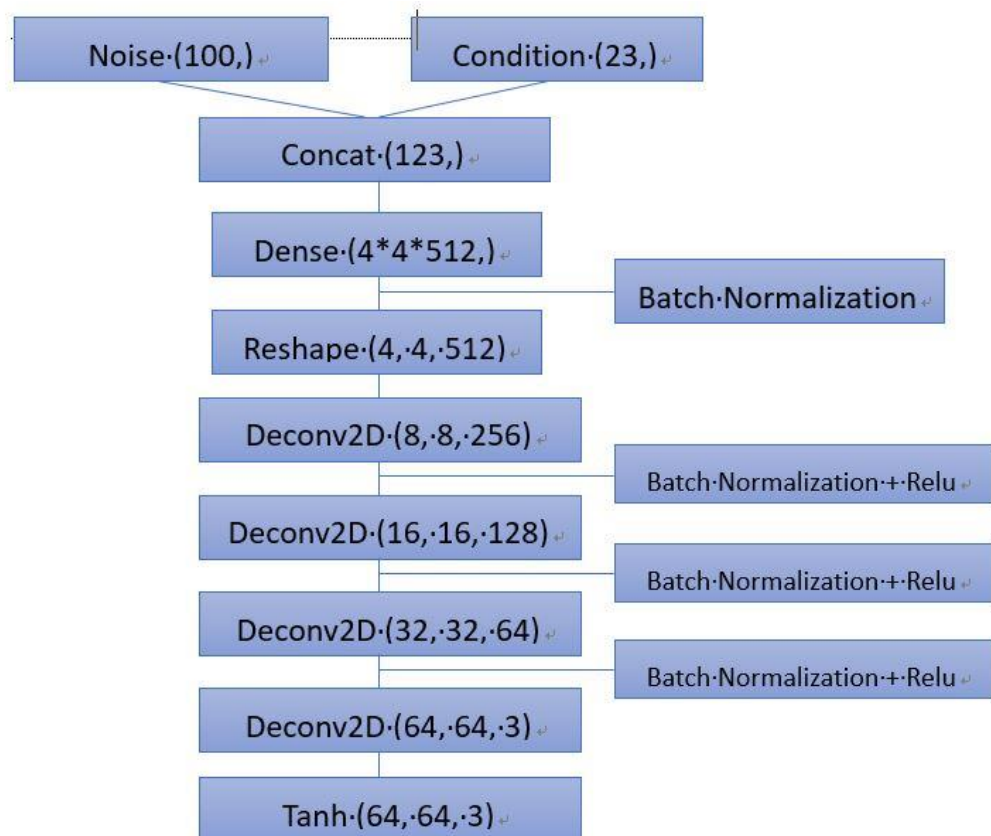


Report

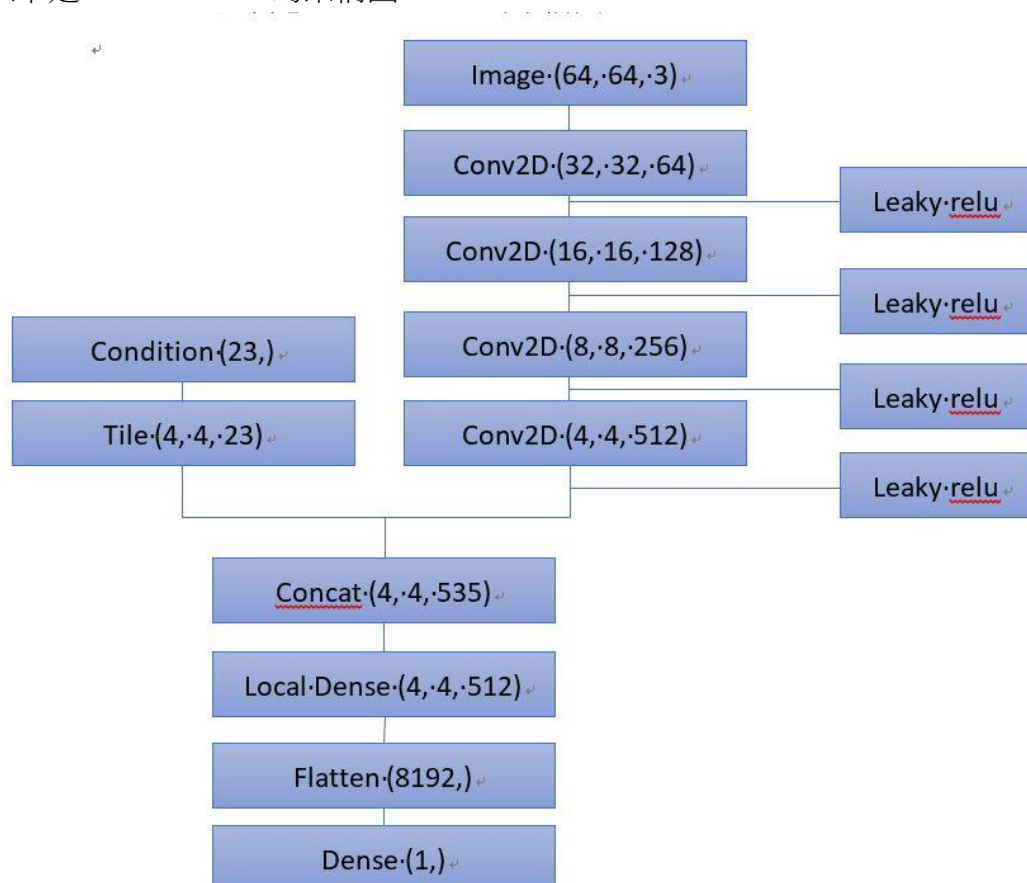
R06922074 吳柏威

- Model description(2%)

- 在 model 上我們採用的是 DCGAN 的架構在給予改良，參考的 github 是 <https://github.com/paarthneekhara/text-to-image>
- 經過測試發現 skip thought 在本 dataset 上效果不彰，所以我們將 tag 裡面有提及頭髮和眼睛顏色各做一個 array 的 onehot，接著將兩個 array 接起來作為 condition input。舉例來說，頭髮的 array size 是 12，若 tag 中有 blue hair，則我們將 array 中代表頭髮的 element 設為 1 其餘設為 0，眼睛也是一樣，若該 data 並未提及頭髮或是眼睛，則兩個 array 全設為 0
- 以下是 Generator 的架構圖



■ 以下是 Discriminator 的架構圖



■ Objective function

我們計算 loss 時使用的方式是最原始 GAN 採用的方式

訓練 D 時將 D 輸出的 value 做 sigmoid 之後與對應的 label 做 binary crossentropy 之後加起來。

Fake image 就是使用 generator 產出的假圖片

Wrong image 就是拿下一個 data 的 image

Wrong tag 就是拿下一個 data 的 tag

Image Input	Tag Input	Label
Real Image	tag	1
Fake Image	tag	0
Wrong Image	tag	0
Real Image	Wrong tag	0

訓練 G 時，將 D 的 weight 固定住，loss 計算的方式是，將 Fake image 以及 tag 輸入 D，輸出時做 sigmoid 與 1 做 binary crossentropy

- How do you improve your performance (2%)
 - 比起使用 skip thought 讀取語意換轉換成 vector 的方式，利用兩個 one hot 的 array concat 起來的方式更可以讓 model 吃到 condition 並且改變輸出動漫臉頭髮以及眼睛的顏色
 - 比起使用 WGAN, WGAN-GP 作為計算 loss 的方式，使用原本 GAN 計算 loss 的方式，更能夠吃到 condition 並且 condition 帶來的改變更為明顯。雖然原本 GAN loss 收斂較慢，但是在 WGAN 以及 WGAN-GP 下，每一回合 D 要跑 5 次 G 要跑一次共 6 次，而使用原本的 GAN 計算 loss，D 要跑一次 G 要跑 2 次共 3 次，並且在 gradient policy 以及 clip weight 上計算需要消耗時間，所以在計算上可以快上 WGAN 非常多，所以在訓練上是比 WGAN 還要快的
 - 取消 D 的 batch normalization layer，在使用原本 GAN 的 loss 時在 D 都會加上 batch normalization layer，但是我發現 D 把 bn 層拿掉之後才 train 得起來，performance 是從原本一片雜訊提升到有臉出來
 - 固定 noise，選取輸出較完整，較符合 condition 的 noise 記錄下來，降低輸出崩壞圖的可能性
- Experiment settings and observation (2%)
 - Learning 太大($2e-4$):
 - ◆ 臉部形成很快，但是不好看，臉歪嘴斜
 - 每一回合 D 跑一次 G 跑兩次
 - ◆ 除此之外 train 不起來
 - ◆ D 跑太多次 D 的 loss 會收斂太快 G 學不到東西
 - ◆ G 跑太多次 D 和 G 的 loss 會一直膨脹
 - 計算 cross entropy 的 axis 要設定為最後一維不可以全部一起求平均
 - ◆ 除此之外 train 不起來
 - Leaky relu 的 alpha 設定為 0.2
 - ◆ 除此之外 train 不起來

- Style Transfer



- 進行的 style transfer 目標是將黑白動漫圖片彩色化
- 使用的架構是 Cycle GAN
- 使用的 Dataset 是 anime face 以及經過灰階化的 anime face
- Code 存放在 Cycle_gan.py
- 結果如上圖所示

1~3 排意思為:

原彩色圖	經過 generator x2y 以及 generator y2x 後 cycle image	經過 generator x2y 輸出結果(黑白照)
------	---	----------------------------

4~6 排意思為:

原黑白圖	經過 generator y2x 以及 generator x2y 後 cycle image	經過 generator y2x 輸出結果(彩色照)
------	---	----------------------------

7~9 排以及 10~12 排意思對應 1~3 排以及 4~6 排