

Comic Generation

- Model description (下一頁有圖示)

- pre-trained hair and eyes classifier

- 先用監督式學習訓練兩個分類器 (hair_classifier and eyes_classifier)，透過四層conv2d最後輸出的是12種頭髮顏色的機率 (或是11種眼睛顏色的機率)。objective function是cross entropy。這兩個分類器訓練完之後會接在G的後面幫助G的學習，但這兩個分類器的參數在G的學習中不會再被更新。

- G

- Model Structure

- 將頭髮和眼睛顏色個變成one-hot vector在和Gaussian noise 結合起來當作input，經過四層conv2d_transpose輸出 (64,64,3)的圖片，在通過D,hair_classifier,eyes_classifier。

- objective function

- 包含三個部分 (加總)：
 - 通過D的值 (同improved WGAN 越大越好)
 - 通過hair_classification的cross entropy的值 (越小越好)
 - 通過eyes_classification的cross entropy的值 (越小越好)

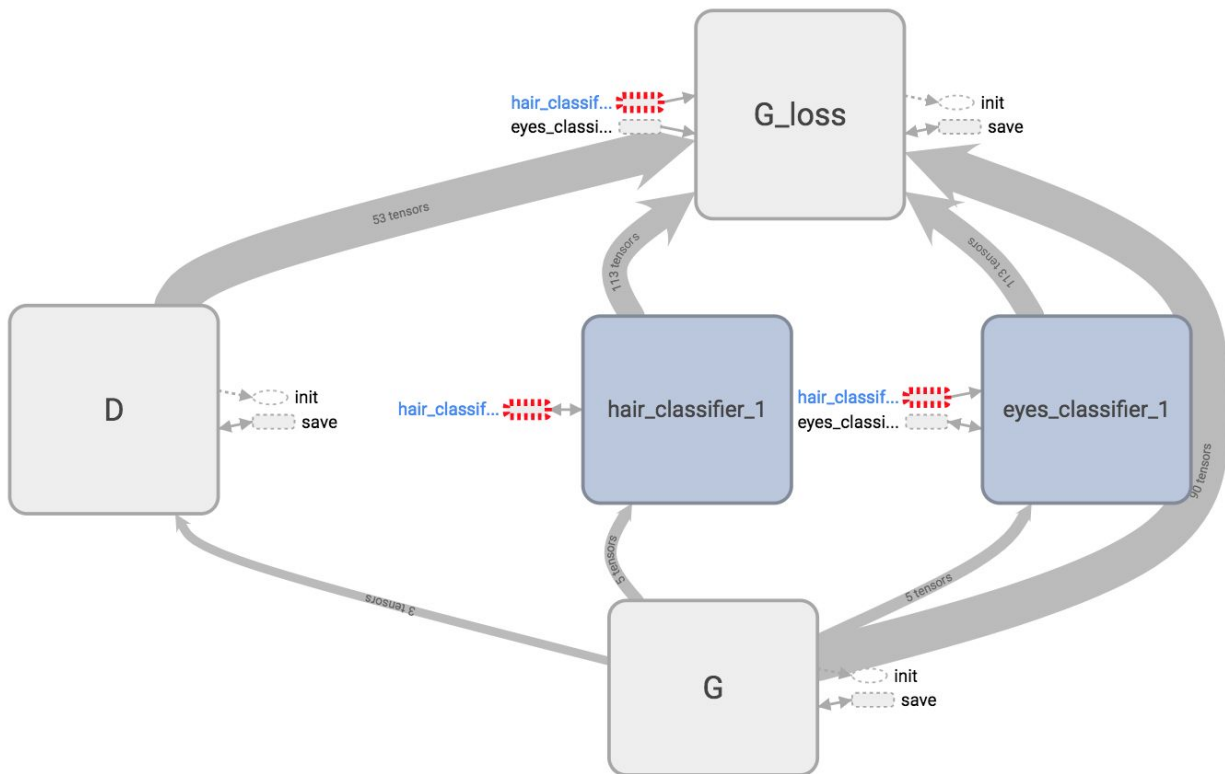
- D

- Model Structure

- 將G產生的image和real image with wrong text和wrong image with right text 視為negative的資料，和real image with right text的 positive 的資料以及同時仿照Improvcd WGAN，將negative和 positive的資料取中間值，一起送進四層conv2d。

- Objective function

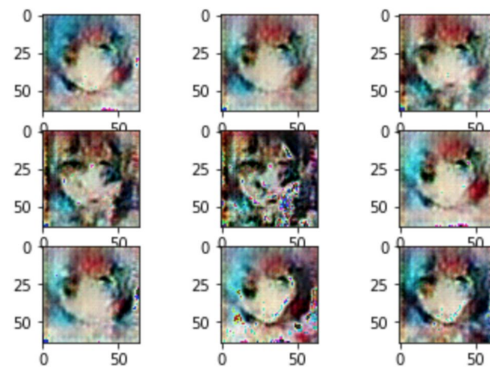
- 同 Improved WGAN



● How do you improve your performance (2%)

- 在D的訓練過程中，原本若只在image input上面加上gradient penalty，訓練下來的結果會使得D loss很低，G loss很高而產生的圖不近理想（如下圖）

red hair blue eyes



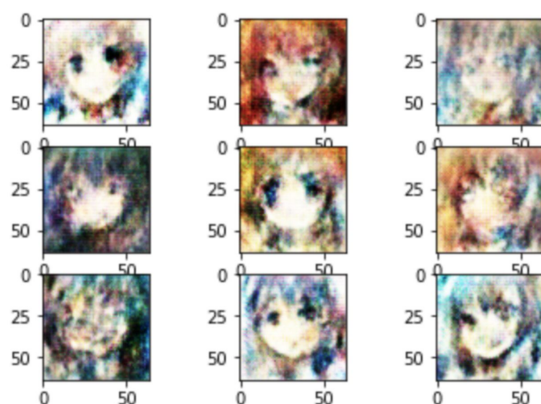
因此在頭髮和眼睛的input也加上gradient penalty。

EX :

real : (1,0,0, ... , 0) fake : (0,1,0, ... , 0)

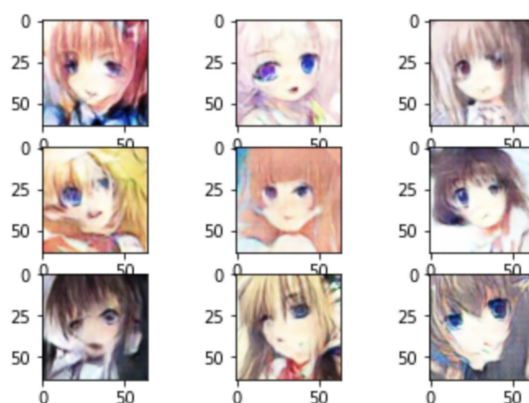
→(ε,1-ε,0, ... , 0) 去算gradient penalty

- 在real image 和 fake image的比例上，若fake image 比例太大，會有不好的影響（如下圖），因此最後是用1:1。



- 由於單用Improved WGAN，要學習到頭髮和眼睛的顏色變化並不是很容易的事情，因此特別加上兩個分類器來幫助學習。
before :

pink hair brown eyes



after :

pink hair orange eyes



○

- Experiment settings and observation (2%)

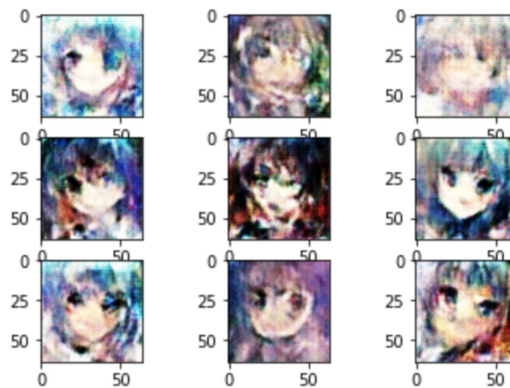
- setting

- batch_size = 16
 - kernel of conv2d = (5,5)
 - strides of conv2d = (2,2)
 - filters of 4 conv2d of classifier are 32,32,32,32
 - filters of 4 conv2d_transpose of G are 512,256,128,3
 - filters of 4 conv2d of D are 128,256,512,1024
 - activation是用leaky_relu (alpha=0.2)
 - optimizer is Adam (lr=0.0001 beta1=0.0,beta2=0.9)

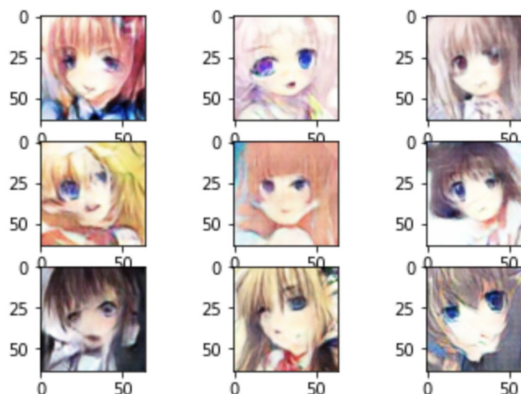
- observations

- batch若太大，則圖案較類似，且圖案也較模糊。選擇較小的batch_size，結果相對較好，下圖為batch_size的比較情況。

batch_size = 128



batch_size = 16



- 有嘗試過非常簡單的conv2d（如filter都設為32），但結果圖案滿模糊的，所以可能不太能夠用太過於簡化的模型來做動漫圖片的生成。
- 有時train到某個程度後，圖片的清晰度和頭髮眼睛顏色的準確度好像就不會變得更好，甚至有可能會更差，
- 雖然兩個分類器的accuracy只能達到50%，但仍有不錯的效果，可能原因有以下幾點：
 - 很多顏色非常接近，而在光影變化下，會導致顏色相對變化，如亮光中的紅色和暗光中的粉色非常相近。
 - 有些標示可能錯誤。
 - 因此儘管正確度不高，但生成的圖仍有不錯的表現。
- rmsprop和adam都可以跑得動，但有點難比較誰優誰劣。