

Deep Learning (Homework 3)

0753420 郭家瑄

I. Cartoon Character Generation

i. Data preprocessing

這邊對 cartoon 資料集進行的前處理有兩項：

1. Resize

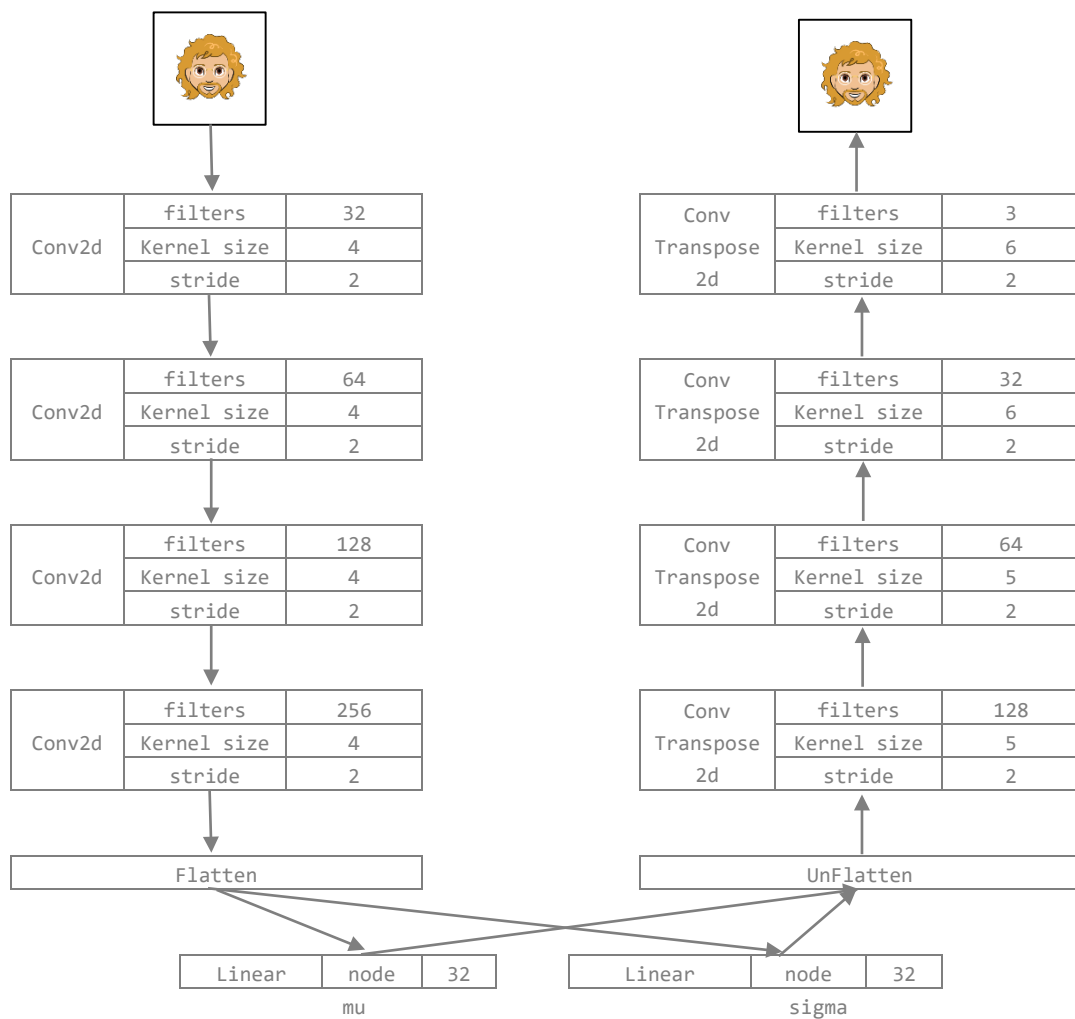
因為原圖解析度比較高，所以將所有圖片都 resize 成大小為 64*64。

2. Normalize

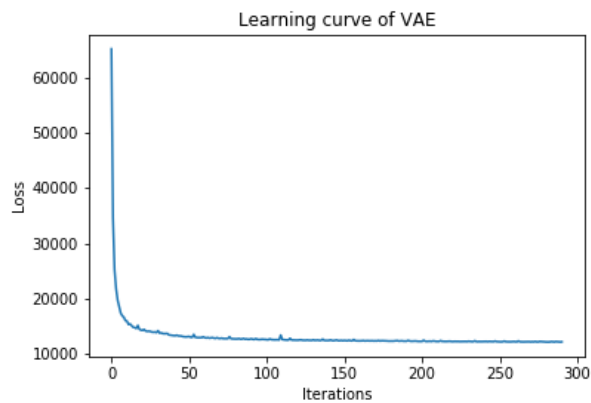
資料集的圖片有 RGB 三個通道，每個通道的值域都是界於 0~255 間，這邊將所有通道的值都壓縮到 0~1 之間。

ii. Model detail and learning figure

1. Network structure



2. Learning curve



右圖是 VAE 在 training 過程中 Loss 的 Curve。

Loss 的計算有兩項，一項是重建的圖片與原始圖片間的 MSE，另一項是 Encoder 結果的 μ 跟 σ 與 Normal distribution 之間的 KL divergence。

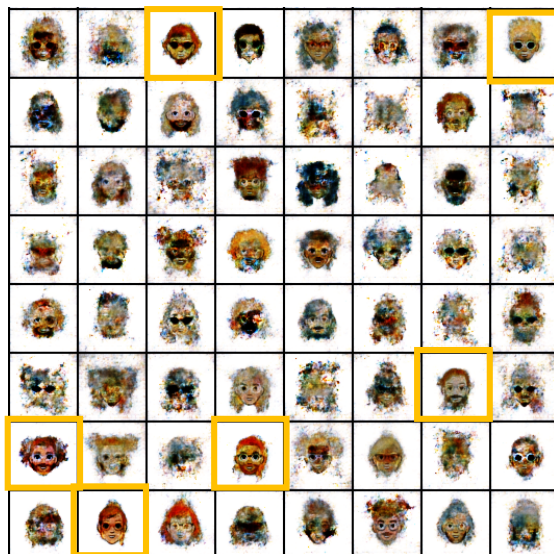
iii. Result

1. Reconstructed examples



右圖是隨機抽取圖片，進 model 重建完的結果，每個小框框的右邊是原始圖片，左邊是重建後結果。可以很明顯的看的出來，model 重建的效果非常好，幾乎分不太出來是原圖還是重建後圖片。

2. Generated examples

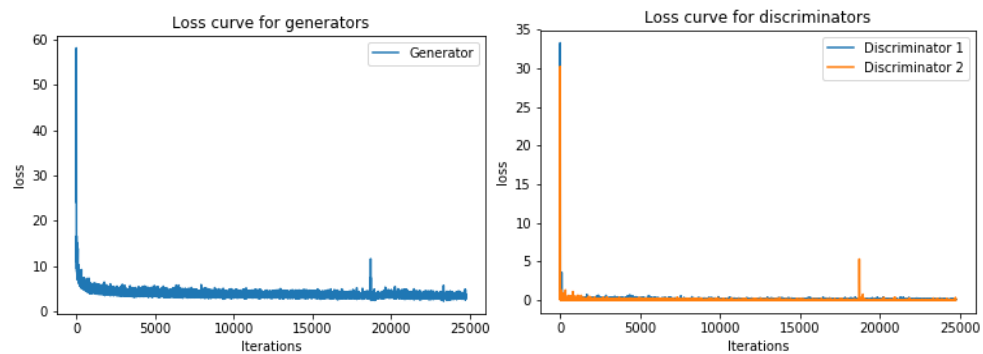


右圖是隨機產生亂數進入 model 生成的圖片，因為 VAE 的 training 相當花時間，因此我認為我還沒將 model 訓練到收斂，許多生成的結果不是很好，但仍有幾張是看的出來臉的輪廓，如圖上標註的記號。

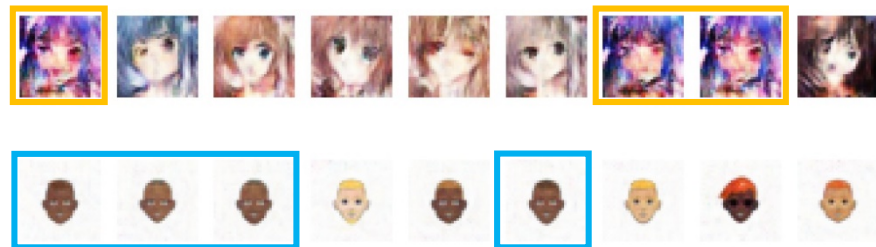
II. Style Transfer

i. Learning figure

建完 cycleGAN 的 loss function 後，得到的 generators 和 discriminators 的 Loss curve 如下：



ii. Result



上面是利用 train 好的 model 隨機採樣出的狀況，首先轉換後的圖片其實有點崩壞，人臉有種扭曲的感覺，像是拼湊上去，另一個觀察到的現象是蠻多 sample 好像是長的一樣的。

iii. Discussion

1. What's mode collapse?

Mode collapse 是指不管給 generator 怎樣的 input，它生成出的樣本多樣性相當侷限，甚至是指產生同一種樣本的現象。

2. Is mode collapse issue serious in this task? Why?

我認為 mode collapse 在這個 model 蠻嚴重的，從 Result 上標出的框框可以看的出來，轉換後相似的圖片過多，