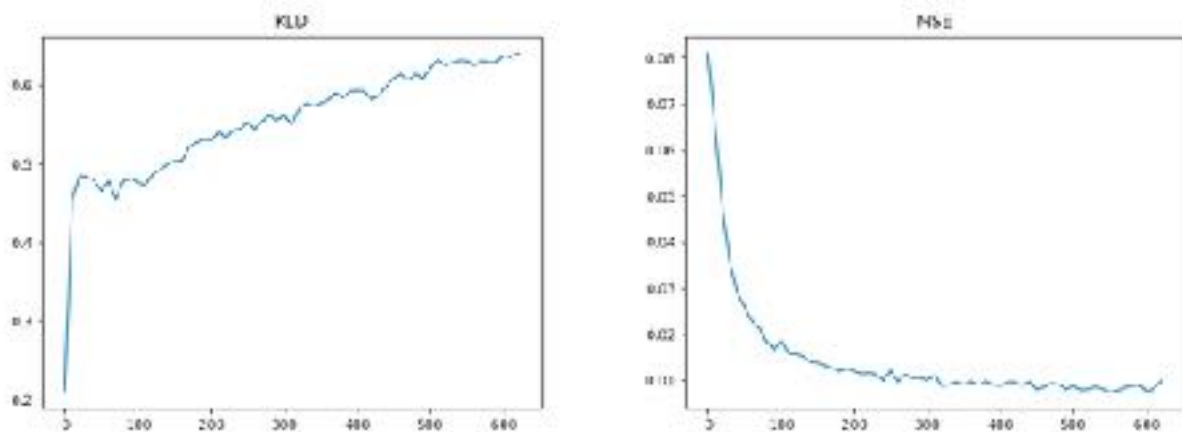


### Problem 1.VAE

1. Describe model : 我的encoder net用 4層 2d convNet , dim = 64 每層倍增 128, 256, 512, 每層都會做batch normalization 再加上 leaky relu, 使得網絡的效果最好, 最後reshape成  $z\_dim = 512$  , 在接上linear function 別分算  $z\_mean$  ,  $z\_log\_sigma\_sq$  generator model, 經過四層的deconv net 在最後一層output時用sigmoid 其他都用 relu 作為 activation function.

### 2. fig1\_2



### 3. fig1\_3



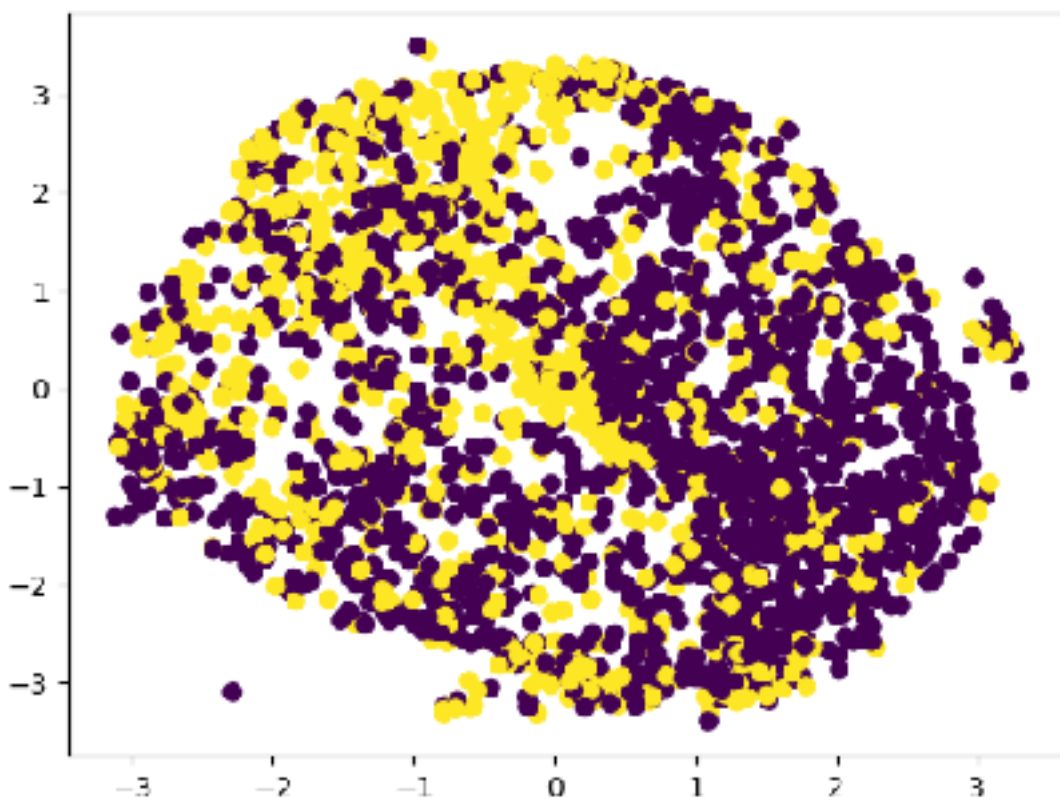
4. fig1\_4.

Training 的不夠，來不及把 lambda KL 參數條大，所以 variance 不夠高，太依靠 input 圖片。



5. tSNE1\_5

我選的attribute是male or not，左上角黃色是male，右下角紫色是female

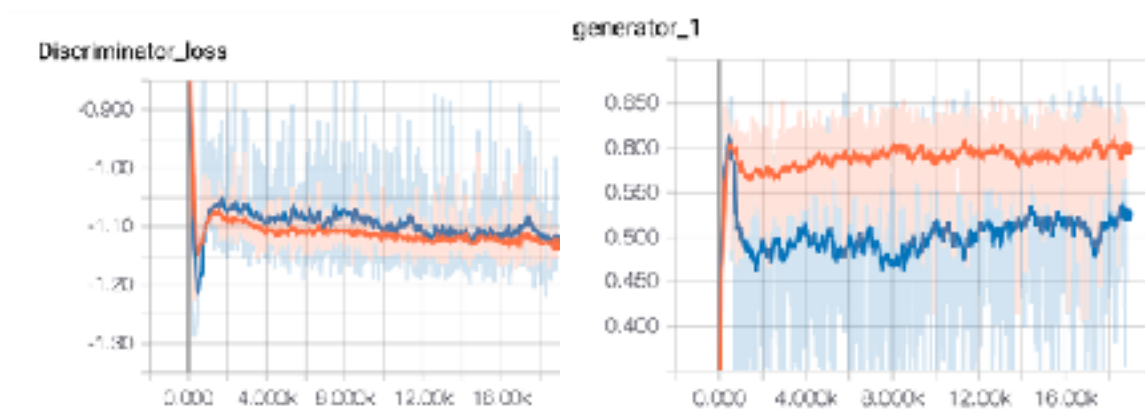


6.

VAE 原先只是 auto encoder 跟 decoder 但是加入隨機變量，使得產生圖片的能力變強，從 data 學到的是在 latent space 的 region，而不是在單個點，encode 學到一個概率分佈，引入 KL divergence 使得 learning variances 增加，比只用 reconstruction loss 好

## Problem 2 GAN

1. 我使用的model是dcgan + wgan，因為gan存在著困難，generator和discriminator很難透過loss來觀察training的好壞，必須邊train 邊看著圖片是不是變好，而wgan就解決了這個問題，首先discriminator最後一層去掉sigmoid, generator loss不取log, 每次更新參數之後把絕對值階段到不超過(-0.01, 0.01)區間，每層都加入batch normalize, activation function使用leaky relu，Optimizer用RMSProp，generator等價變換成真實與生成間的 J S 散度當 discriminator越好，generator也最好，loss 近似於最小化和之間的 J S 散度
- 2.



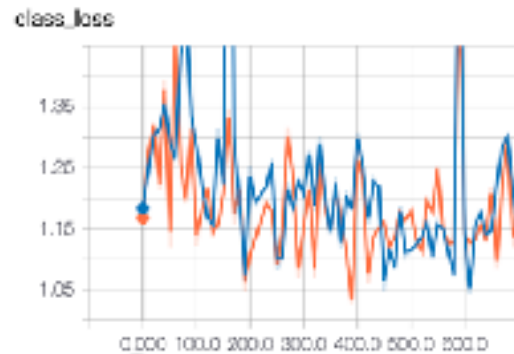
藍色是generator，橘色是discriminator，training初期loss變動大，大概進入穩定之後，可以觀察discriminator loss變小，代表d\_net變強，而g\_net loss會上升，是因為d\_net 變強的關係，但可以觀察在16.00K處，d loss上突一個角，相對g loss是下降的

## 3. fig2\_3



4. gan很難train，儘管已經加上wgan讓臉變的很清楚，但還是有許多壞掉的情況出現，可能用BEgan等其他gan能做的越來越好，使用leaky relu 跟batch normalization 都讓model training 更穩，還有用RMSProp學習效果都比adam SGD明顯要好
5. 可能因為有對抗網路，所以人臉生成比V A E 還要好，V A E 的臉非常模糊甚是看不出是人臉，因為他只有透過部分隨機變量，kl與mse 作為loss，變化性還是不夠高，而G A N 是完全透過random生成每一次的圖片，並且經過對抗，使圖片變得更好。

## Problem 3



1. 在原先的 G A N 上，跟 discriminator 共用變數，接出一段新 dense net classifier，我選用 smile or not smile，將有笑的圖片標記成 1 其餘為 0，sigmoid cross entropy 算 loss，並且將 label 接在 random z 的最後面，generator 出來的圖片就會分成兩個 class。
2. 橘色是 real，藍色是 fake，從圖中觀察因為我共用了 discriminator 的變數，所以在 discriminator 浮動的時候，classification loss 也會跟著大幅浮動。