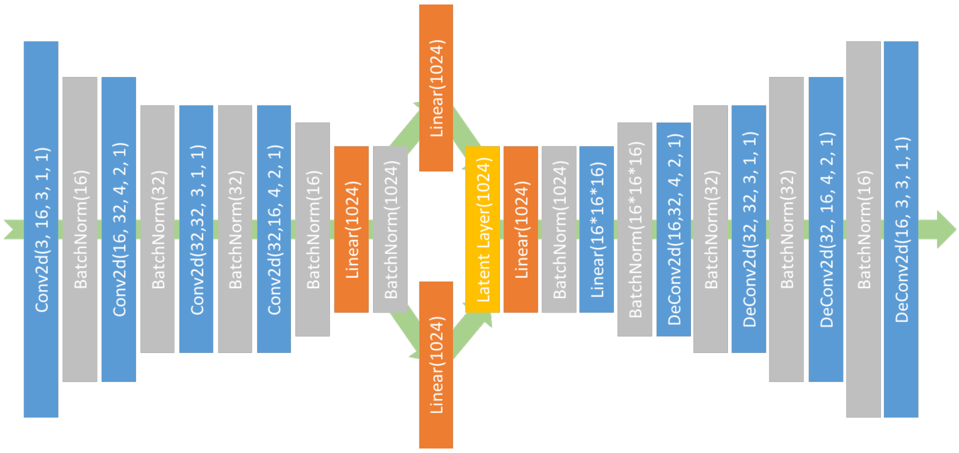
系所級：電機碩一　　　　　學號：r06921076　　　　　姓名：林益璟

Problem 1. VAE (6%)

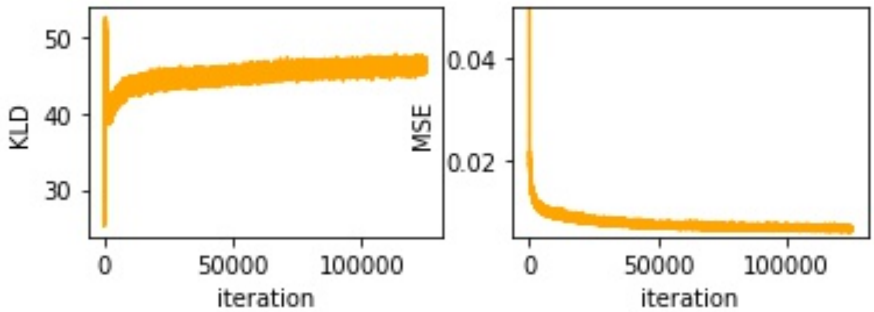
1 Describe the architecture & implementation details of your model (1%)

- 最後使用的 VAE 架構如下：

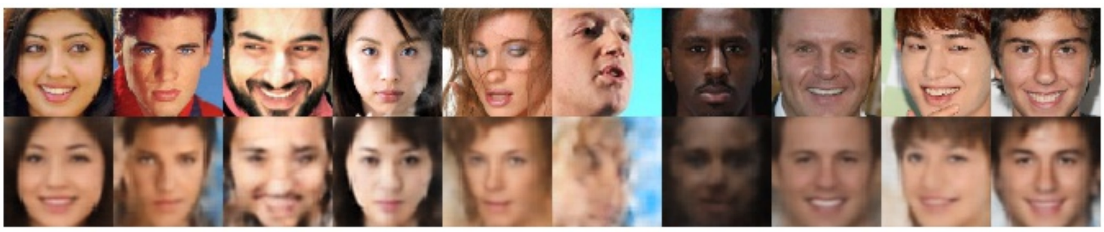


encoder & decoder 分別以 4 組 Conv/BN/ReLu 組成，並在 bottleneck 以 Linear(1024)相接。

2 Plot the learning curve (reconstruction loss & KL divergence) of your model [fig1_2.jpg](1%)

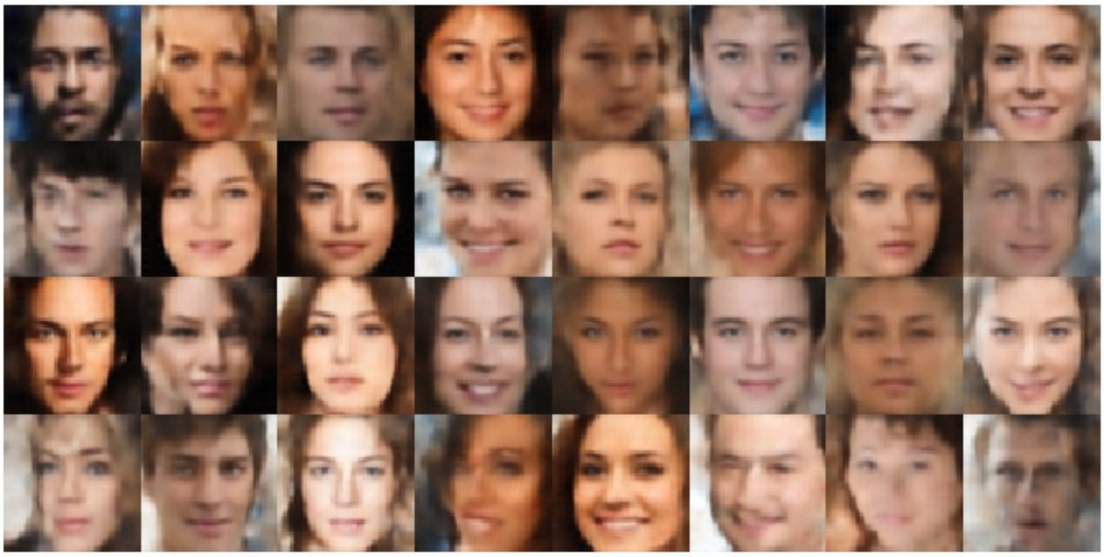


3 Plot 10 testing images and their reconstructed result of your model [fig1_3.jpg], and report your testing MSE of the entire test set (1%)



- testing MSE of the entire test set is: 0.010567353

4 Plot 32 random generated images of your model [fig1_4.jpg] (1%)



5 Visualize the latent space by mapping test images to 2D space (with tSNE), and color them with respect to an attribute of your choice [fig1_5.jpg] (1%)

Pass

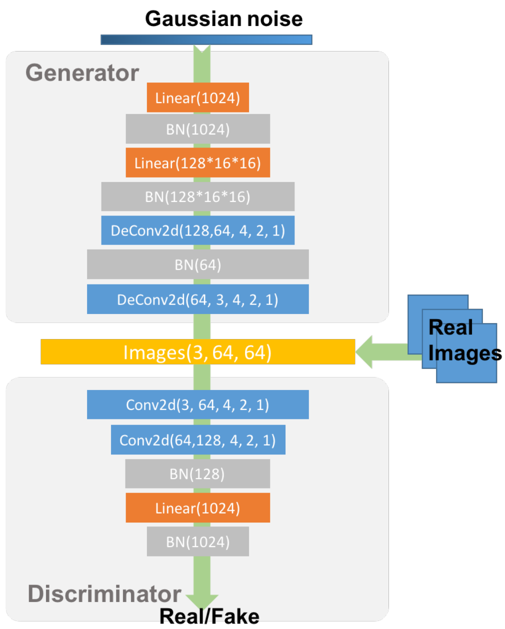
6 Discuss what you've observe and learned from implementing VAE (1%)

- 比起 AutoEncoder，我們可以掌握其中 latent vector 的分佈情形，使得我們能自行產生特定分布的噪音(fig1_4 使用 standard normal distribution)，經過 decoder 來產生圖片。
- 實作過程中發現 BatchNorm 的有無對 training 有些許的影響，bottleneck dimension 設定小於 128 也會使得 decoder 的資訊量不足以回復原圖的特徵。

Problem 2. GAN (5%)

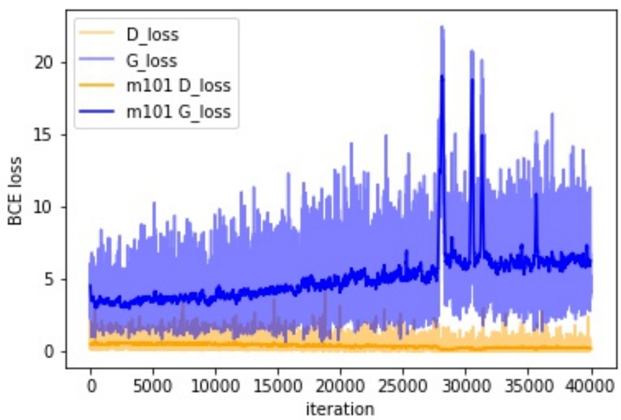
1 Describe the architecture & implementation details of your model (1%)

- GAN 的 noise_dim=62，以 Linear/BN/ReLu 為一組，疊兩次，之後再以 2 次 Conv Tranpose 將圖片轉回(3,64,64)；decoder 則倒過來處理，先經過兩次 Conv，再接 2 層 linear 輸出一個值來決定 real/fake。



- 訓練階段使用 3 個 tips：
  1. Gaussian noise
  2. BN for Discriminator input
  3. Training times ratio (D/G)=2

2 Plot the learning curve (in the way you prefer) of your model and briefly explain what you think it represents [fig2_2.jpg] (1%)



因為 training 期間, Generator 的浮動極大，且 iteration scale 遠大於 BCE loss，使得 learning curve 看起來很雜亂且無法看出趨勢。解決方法是將每 101 個 iteration 的值。

3 Plot 32 random generated images of your model [fig2_3.jpg] (1%)

4 Discuss what you've observed and learned from implementing GAN (1%)

如同 Q2-2 所說，training 時 loss 的浮動極大，且多次實驗下發現，若 D_loss 低於 0.01 時，可能會使得 Generator 無法再進步，甚至壞掉(只會輸出雜訊)。在多次嘗試下發現，Training times ratio (D/G)設定為 2(每一次 iteration，D 先 train 2 次，G 才 train 1 次)，比較不會發生 Generator 壞掉的情形，但在 iteration 超過 150,000 次以後仍有高機率壞掉，故後面的實驗(包含 ACGAN)都只在 40,000~80,000 個 iteration 內完成。

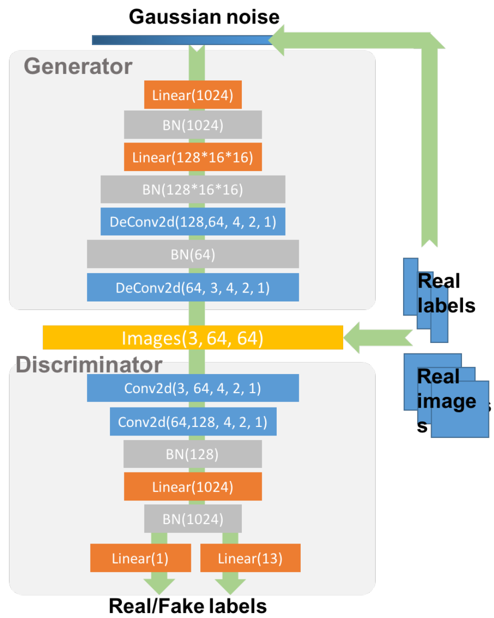
5 Compare the difference between image generated by VAE and GAN, discuss what you've observed (1%)

首先是 VAE，我們發現大部分產生的圖片都有 mean face 的效果，個人猜測是因為經過 bottleneck 後的有限資訊要滿足 MSE 作 optimization，且 model 最後是有正確答案的詳細資訊供優化，故最後會得到 mean 的效果 ；以 GAN 而言，Generator G 只有 discriminator 的批改分數來作為優化依據，他缺少以 pixelwise 取得的 MSE 來對每個 generated pixel 做優化，故 train 出來的圖會比起 VAE 更銳利，且常常出現無法辨識的圖片 。

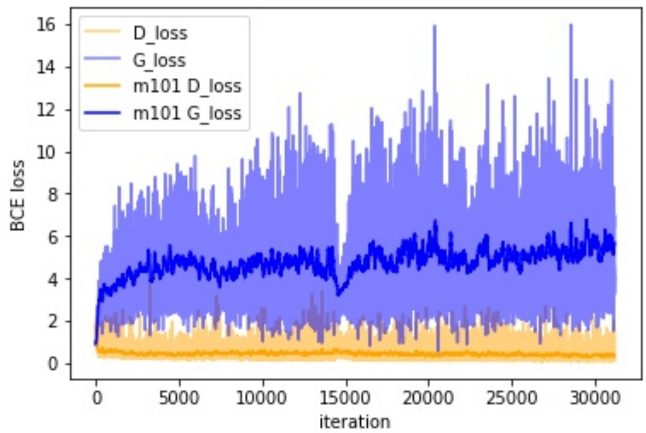
Problem 3. ACGAN (4%)

1 Describe the architecture & implementation details of your model (1%)

- 與 GAN 的大同小異，只是多了 labels input/ output。Label output 採用連續行變數處理，故不同於一般使用 CroossEntropy，我這裡使用 MSE 算 loss。
- 訓練階段也使用 3 個 tips：
  1. Gaussian noise
  2. BN for Discriminator input
  3. Training times ratio (D/G)=2



2 Plot the learning curve (in the way you prefer) of your model and briefly explain what you think it represents [fig3_2.jpg] (1%)



同 GAN 的理由ＱＱ，但與 GAN 相比，ACGAN 的 G_loss 平均約在 5 左右，比 GAN 低了一點(且 GAN learning curve 中，G_loss 還有上升的趨勢)。

3 Plot 10 pair of random generated images of your model, each pair generated from the same random vector input but with different attribute. This is to demonstrate your model’s ability to disentangle feature of interest. [fig3_3.jpg] (2%)

上圖是預期印出有微笑跟沒有微笑的圖片，但看起來 model 並沒有學到居互 attribute 呈現不同的特徵，我想可能是因為我用 MSE 作為各 attributes 的 loss function，加上 normalization 後的灰階值得到的 MSE 與 CrossEntropy 差異太大，所以重要性被縮小所致。