Deep Learning (Homework 3)

資管所碩一 309706029 謝敦凱

1.

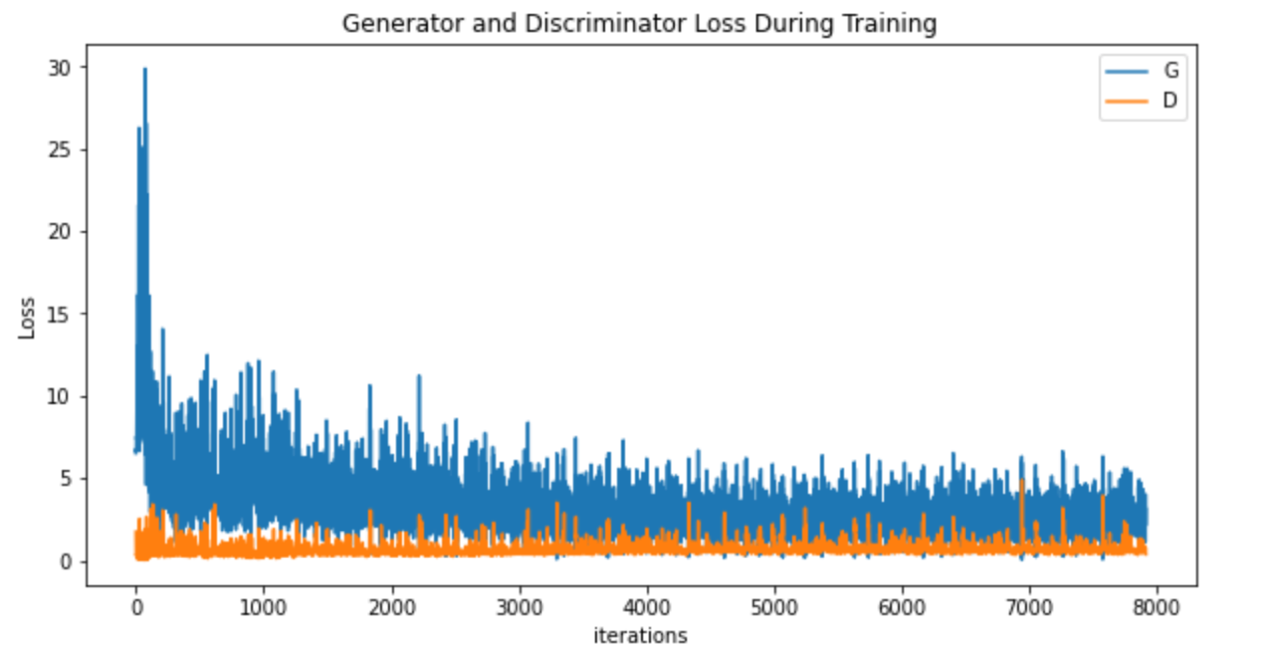
1. Describe how you preprocess the dataset (such as resize, crop, rotate and flip) and explain why. (5%)

Resize圖片大小為64。

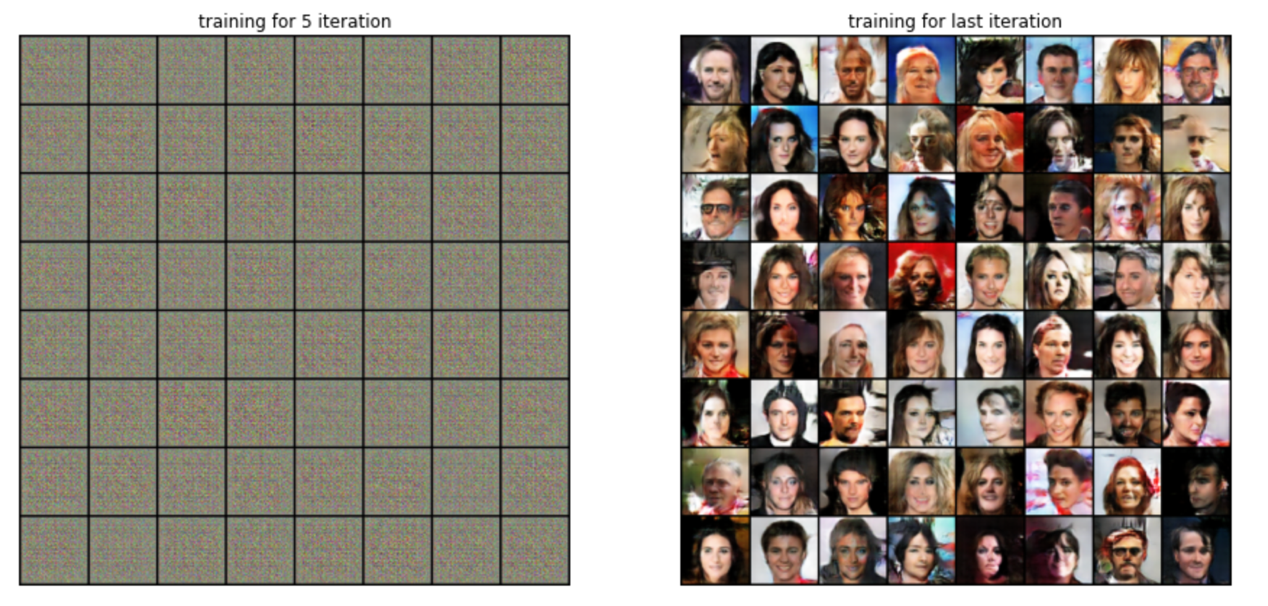
Crop:用centercrop，以中心點去cropping圖片，size = 64。

都是為了讓圖片大小一致。

1. Plot the learning curves for both generator and discriminator. (15%)



(c) Draw some samples generated from your generator at different training stages. For example, you may show the results when running at 5th and final learning iteration. (5%)



2.

(a) Please describe the meaning of the following four pictures during training of GAN, where blue dashed line indicates the discriminator, green solid line indicates the generator. The answer should include the following: (Note: Each step should be discussed.)

− what is the meaning of black dashed line, x and z

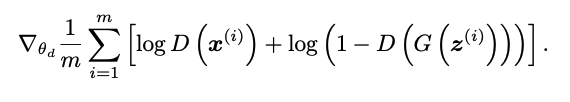
Black dashed line : real image的資料分佈

X : 透過G(z)產出的fake image分佈

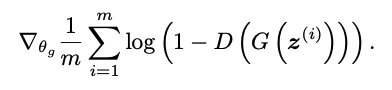
z : noise的分佈，有些z會被抽出來去generator內生成fake image

− which step is to train the generator or discriminator and show the corresponding objective function

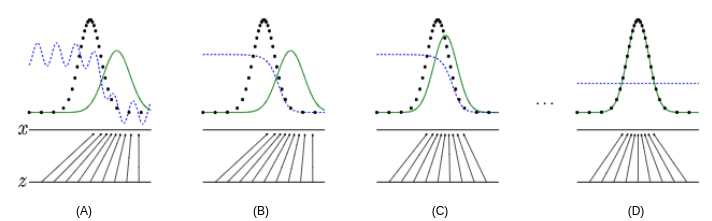
(a), (b),(c) 是train generator and discriminator的階段

Loss of discriminator : 

Loss of generator:

  
− why D(x) equals to 1/ 2 in ideal case when the training is finished

因為到了(d)階段，可以看到Pg == Pdata，這時generator已經的產出圖片的分佈已經很接近real image的分佈，所以discriminator只能瞎猜，D(X)的機率才會是1/2，因為分辨X來自training data的機率為1/2其實就在亂猜，理想應該是D(X) = 0.9，或是越高越好，這樣discriminator會更準。



(b)The Helvetica Scenario often happens during training procedure of GAN. Please explain why this problem occurs and how to avoid it. (5%)

Generator如果把不同的noise z都投射到同一張fake image，然後騙過discriminator，那generator就會一直把不同的noise z投射到那張fake image，反正騙得過，所以就無法訓練generator生成更多圖片了。

避免的方法：

改進成Wasserstein GAN

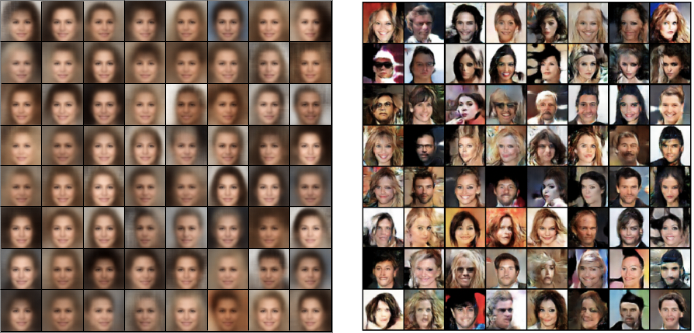
1.把discriminator最後一層拿掉sigmoid。

2.Discriminator and Generator的Loss不取log。

3.Discriminator的參數更新後，取絕對值且截斷到不超過一個固定常數

4.不要用動量的優化算法，改用SGD

(c) Both VAE and GAN are generative models. The following figures are random generated results by using VAE (left) and GAN (right). Please compare two results and describe the pros and cons of two models. (10%) (Hint: You can compare the loss function and training method using these two models.)



VAE：生成的圖比較模糊。

GAN：生成的圖比較創新，且不模糊。

Loss fumction:

VAE最大化ELBO，目的是為了做maximun likelihood estimation，等價於最小化KL，但這個KL不是數據跟噪音的KL，而是model給出的p(x)和數據所展示的p(x)間的KL。

GAN是最小化Jensen Shannon Divergence，這個JS是從model來的p(x)數據所展示的p(x)之間的。

Training model:

VAE里面z到x的depencence是随機的，而GAN是確定性的。這使得VAE的inference相對来说是well-defined的，而GAN的inferernce相對來說是病態的。

這個建模上的差别也是來自於兩者建模目的的不同。GAN就是為了生成新圖片，而VAE除此之外，更想做的事情是建模數據的分布以及學到數據的隱含表示。

對於建模數據的分布，通常來說也就是density estimation，雖然VAE和GAN都有在x的空間上定義概率密度，但GAN定義的密度更難計算。