

1. (2.5%) 訓練一個 model。

a. (1%) 請描述你使用的 model (可以是 baseline model)。包含 generator 和 discriminator 的 model architecture、loss function、使用的 dataset、optimizer 參數、以及訓練 step 數 (或是 epoch 數)。

訓練的模型為 DCGAN, generator 架構為一層 linear layer + 四層 ConvTranspose2d layer, 最後經過 tanh function 得到 output image. Discriminator 為四層 Conv2d layer, 最後經過 Sigmoid 得到預測結果. 使用作業提供之 face dataset. 訓練參數部分, 所採用得 loss 為 BCELoss, optimizer 為 Adam, learning rate 為  $1e-4$ , 共訓練 10 epoch.

b. (1.5%) 請畫出至少 16 張 model 生成的圖片。



2. (3.5%) 請選擇下列其中一種 model： WGAN, WGAN-GP, LSGAN, SNGAN (不要和 1. 使用的 model 一樣, 至少 architecture 或是 loss function 要不同)

a. (1%) 同 1.a, 請描述你選擇的 model, 包含 generator 和 discriminator 的 model architecture、loss function、使用的 dataset、optimizer 參數、及訓練 step 數 (或是 epoch 數)。

訓練的模型為 WGAN, 這裡主要想討論 loss function 對訓練結果的影響模型架構, 因此模型架構沒有改變. generator 架構為一層 linear layer + 四層 ConvTranspose2d layer, 最後經過 tanh function 得到 output image. Discriminator 為四層 Conv2d layer, 最後經過 Sigmoid 得到預測結果. 使用作業提供之 face dataset.

訓練參數部分, 所採用得 loss 對 generator 來說為  $-\text{mean}(D(G(Z_i)))$ , 對 discriminator 來說為  $-\text{mean}(D(X_i) + \text{mean}(D(G(Z_i))))$ , 其中  $X_i$  代表真實的圖片,  $G(Z_i)$  代表產生的圖片, 我們希望最後 generator 產生得 image 能夠有最小得 Wasserstein Distance, 另外為了要滿足 1-Lipschitz 的性質, 需

要做 weight clipping, 這裡 clip 參數設定為 0.01, optimizer 為 Adam, learning rate 為  $1e-4$ , 共訓練 20 epoch.

b. (1.5%) 和 1.b 一樣, 就你選擇的 model, 畫出至少 16 張 model 生成的圖片。



c. (1%) 請簡單探討你在 1. 使用的 model 和 2. 使用的 model, 他們分別有何性質, 描述你觀察到的異同。

使用這兩者產生出來的圖片看起來都算不錯, 差別為使用 dcgan 產生的圖片模糊的現象比較不嚴重, 另外在訓練過程中發現 wgan 要得到跟 dcgan 差不多水準的圖片需要訓練比較多個 epoch, 在 dcgan 我訓練了 10 個 epoch, 而在 wgan 我訓練了 20 個 epoch. 另一個現象是如果把 dcgan 也訓練 20 個 epoch, 會發現產生出來的圖片沒有比較進步, 反而是有 mode collapse 的情況發生.

3. (4%) 請訓練一個會導致 mode collapse 的 model。

a. (1%) 同 1.a, 請描述你選擇的 model, 包含 generator 和 discriminator 的 model architecture、loss function、使用的 dataset、optimizer 參數、及訓練 step 數 (或是 epoch 數)

使用與 1.a 同樣的模型架構, 差別在 epoch 訓練 30. 訓練的模型為 DCGAN, generator 架構為一層 linear layer + 四層 ConvTranspose2d layer, 最後經過 tanh function 得到 output image. Discriminator 為四層 Conv2d layer, 最後經過 Sigmoid 得到預測結果. 使用作業提供之 face dataset. 訓練參數部分, 所採用得 loss 為 BCELoss, optimizer 為 Adam, learning rate 為  $1e-4$ , 共訓練 30 epoch.

b. (1.5%) 請畫出至少 16 張 model 生成且具有 mode collapse 現象的圖片。



c. (1.5%) 在不改變 optimizer 和訓練 step 數的情況下，請嘗試使用一些方法來減緩 mode collapse。說明你嘗試了哪些方法，請至少舉出一種成功改善的方法，若有其它失敗的方法也可以記錄下來

使用了兩種方法，第一種為使用如 wgan 的 loss function，這種方法在訓練同樣 epoch 數量的情況只有些微的 mode collapse，而另一種方法為在 wgan 的基礎上，再使用了 wgan-gp 的技巧，不做 weight clipping，而是做 gradient penalty。這種方法也只會有些微的 mode collapse，而收斂速度較 wgan 快一些。