

ADLxMLDS 2017 Fall

HW4 - Anime Generation

B05901189 吳祥勸

January 7, 2018

1 Model Description

使用經典的 conditional DCGAN network，結構大致如下圖所示。

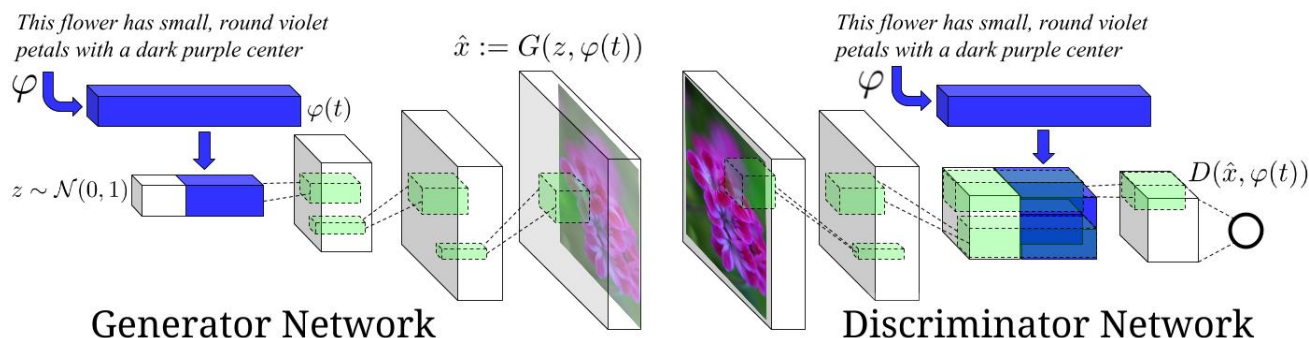


Figure 1: Conditional DCGAN

1. Generator:

使用 100 維的 Gaussian Noise，配上 one hot 的 condition 當作輸入。先通過一個 Dense Layer 變成 $[4*4*128]$

Reshape 成 $[1, 4, 4, 128]$

通過四個 transpose convolution，使用 kernel size = 5，stride = 2。

filter 數目分別為 128, 64, 32。

其中前三層都使用 batch normalization，以 lrelu 當作 activation function。

最後一層使用 tanh 作為 activation function。

2. Discriminator:

因為使用 tanh 當作 Generator 最後一層的 activation function，真實圖片也必須經過 normalize，把值都線性轉換到 $[-1, +1]$ 的區間內。convolution 層和 Generator 基本上完全相反。不過輸出前會將 convolution 完的圖和 condition 相接，在通過兩層 Dense Layer 才得到輸出。

3. Objective Function:

$$\min_G \max_D E_{x \sim P_{real}} [\log(D(x))] - E_h [\log(1 - D(G(h)))]$$

2 Performance Improvements

1. Discriminator 上增加 Dense 層:

原先我在 Discriminator 上只有一層 Dense 層，這樣產生出來的圖片會沒辦法對應到給定的 Condition，加上了一層 512 個 neuron 的 Dense 層後很明顯的產生出來的圖幾乎都有符合 Condition 了。這個是 Generator 和 Discriminator 唯一不對稱的地方。

2. Discriminator 輸入加入 noise:
算是 Augmentation 的一種方式，也可以弱化 Discriminator 一些，實驗上發現可以讓訓練比較穩定，比較快開始符合 Condition 也比較慢 overfit。
3. 用 Leaky Relu 取代 Relu:
GAN 的訓練常常會有 Gradient Vanishing 的問題，為了解決這個問題可以用 lrelu 取代 relu，而且要用大一點的 α (大約 $1e-3 \sim 1e-2$)，這樣才不會像 relu 有些地方沒有梯度。

3 Experiment Settings and Observation

1. Hyperparameters:
learning rate = $2e-4$
Optimizer: Adam
beta1 = 0.5
noise dimension: 100 sampled from standard normal distribution.
discriminator input noise: noise sampled from normal distribution with dev=0.1
2. 處理 Label 的方法:
這次的 label 很雜亂，有時候會有髮色或瞳色會同時有很多 label，或是完全沒有 label。我使用的是 one-hot label，處理怪異的 label 是假如有多個 label 就從中隨機選一個，如果完全沒有 label 就從全部可能的顏色裡面隨機選一個。
3. 比較有無使用無 label 的圖:
嘗試拿掉所有沒有髮色也沒有瞳色 label 讀圖，比較兩種方法在約 200 個 epoch 後的表現，如下所示。總共圖片約 31k 張，去除無 label 的圖剩下 17k 張，換算出來的 epoch 數也不同。可以發現沒有使用的話圖片品質較差，顏色也較不統一。可見即使沒有 label 也有助於產出更真實的圖片。



Figure 2: 沒有使用

Figure 3: 有使用

4. 比較 discriminator input noise 大小的影響: discriminator 的輸入加上噪音可以達到 augmentation 的效果，但是噪音的大小也是要經過嘗試我比較了兩種大小，

分別是 standard deviation = 0.5 及 0.1 的 random normal noise. 因為圖片都是 normalize 到 $[-1, +1]$ 區間的, 可以知道 std=0.5 已經算是很大的 noise 了。實驗上也可以發現這麼大的 noise 反而會讓圖片變得怪怪的。



Figure 4: std=0.5

Figure 5: std=0.1