

MLDS HW3 Report

Members

D04944007 高瑞宗
R05922139 林子芃
R04921106 陳彥谷
B02902105 廖瑋中

Environment

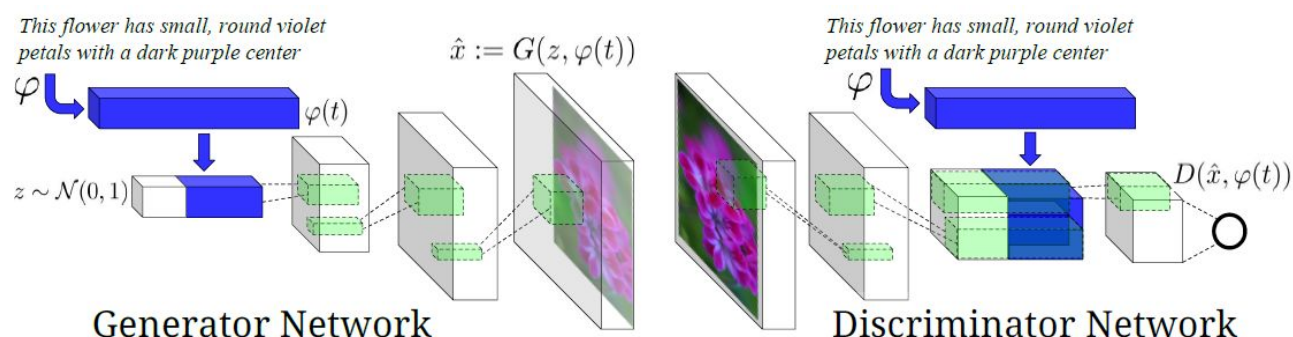
我們用自己實驗室的Azure雲端計算資源：

Name	Cores	Memory	OS	GPU
mslabgpu2	6	56G	Ubuntu 16.04	Tesla K80, 12G Ram
mslabgpu34	12	128G	Ubuntu 16.04	Tesla K80, 12G Ram * 2

Library We Use

1. skipthoughts
2. h5py
3. requests
4. Theano (for skipthoughts)
5. scikit-learn (for skipthoughts)
6. scikit-image (for skipthoughts)
7. NLTK (for skipthoughts)

Model Description



我們的 model structure 如上圖。DCGAN 就是把 GAN 中的 generator 和 discriminator 換成兩個CNN。generator 中去掉 pooling，使用 transposed convolutional layer，ReLU 激活，然後最後使用 tanh。discriminator 中用 stride 代替 pooling，然後使用 leakyReLU 激活。我們的 input 都使用了 batch normalization 避免 collapse。

GAN 的第一層採用 normal noise distribution Z 為 input，接著重組成一個 4 dimensional tensor，最後通過 sigmoid function 輸出。

我們的 generator 的 objective function 為

$$\tilde{V} = \frac{1}{m} \sum_{i=1}^m \log \left(1 - D \left(G(z^i) \right) \right)$$

而 discriminator 的 objective function 為

$$\tilde{V} = \frac{1}{m} \sum_{i=1}^m \log D(x^i) + \frac{1}{m} \sum_{i=1}^m \log \left(1 - D(\tilde{x}^i) \right)$$

How to improve performance

1. LSGAN

我們將 objective function 改成 least square，對於 discriminator 來說，他的 loss function 為

$$\min_G \left\{ \frac{1}{2} E_{x \sim p_{data}(x)} [(D(x) - b)^2] + \frac{1}{2} E_{z \sim p_z(z)} [(D(G(z)) - a)^2] \right\}$$

而對於 generator 來說，他的 loss function 為

$$\min_G \frac{1}{2} E_{z \sim p_z(z)} [(D(G(z)) - c)^2]$$

對於 a、b、c 的選擇，經過我們在網路上的查詢，發現助教的配置 (a,b,c) = (0, 1, -1) 是最多人使用，且據說是最好的。因此，我們也使用這組參數。

但是因為時間不足的關係，我們沒有 train 到足夠的 epoch 數量，因此效果看起來並不是很好。

Experiment Settings and Results

1. DCGAN

Tags for images	“eye”+“hair” + top 3 tags	“eye”+“hair”
Epoch	445	375
Training time	~3days	~3days
Optimizer	Adam	Adam
Learning rate	0.0002	0.0002
Noise dimension	100	100
momentum	0.5	0.5

我們最後選擇使用 “eye” + “hair” + top 3 tags。

“eye” “hair” 是所有包含這兩個字的tag，然後加上排名前三的tag。

主要原因是因為 eye+hair+top3 training 時間比較長，所以效果也比較好。

此外，我們在一開始時也嘗試在96*96的image上面做training，但後來發現需要的resource太多，training時間又長，train出來的效果沒有想像中的好因此放棄，決定使用64*64的image做training。

2. Results

black hair black eyes



blue hair blue eyes



red hair red eyes



green hair green eyes



long hair black eyes



short hair black eyes



long hair red eyes



short hair red eyes



3. Bonus results

None

Observations

在實作的過程中，我們發現 batch normalization 真的很重要。在其他框架大致相同的情況下，缺少了 batch normalization，便會讓整個 model 壞掉。而再加了 batch normalization 之後，我們的 model 就可以正常運作了。

在generator生成圖片的時候，noise的大小對產生出來的圖片差異度影響很大。雖然noise很大的時候圖片差異度很大，但也很容易崩壞。

Team Division

Member	Work
D04944007 高瑞宗	Data preprocessing + basic LSGAN
R05922139 林子芃	Data preprocessing + DCGAN + report
R04921106 陳彥谷	討論
B02902105 廖瑋中	Data preprocessing + LSGAN + report