

# MLDS HW4

## REPORT

R04943151 梁可擎

### MODEL DESCRIPTION

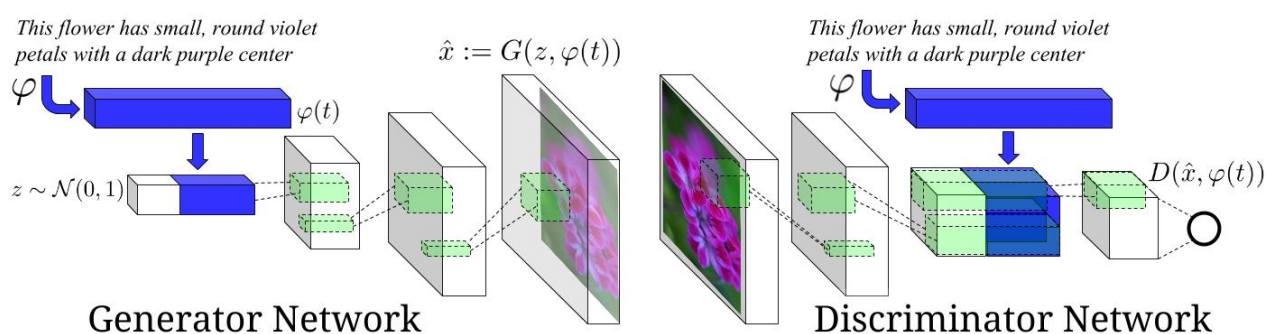


Fig. 1 Text to image model structure.

根據圖一這篇 paper，主要 model 結構為 DCGAN，condition texts 經過 skipthoughts encoded 之後 concat 在 noise 後面 ( $z$ )。這次的 dataset 其實描述很單純，可以不需要用到 skipthoughts，但由於 overfitting 會很嚴重，給予 condition 更多 data loss 對訓練其實也有幫助。在 objective function 計算部分，Discriminator 會給出三種分數，分別為  $s_r$  (real image, right text)、 $s_w$  (real image, wrong text)、 $s_f$  (fake image, right text)，Discriminator 的 loss 為  $\mathcal{L}_D \leftarrow \log(s_r) + (\log(1-s_w) + \log(1-s_f))/2$ ，generator 的 loss 為  $\mathcal{L}_G \leftarrow \log(s_f)$ 。

# HOW TO IMPROVE PERFORMANCE

## ITERATION

在訓練的過程中，為了避免 discriminator 收斂太快，會將 G 更新兩個 iteration 之後再更新一次 D。這樣的做法發現還是在 10 個 epoch 之後 D 就收斂了，大約在 50 epochs 之後 generator 產生出的影像全部變成同一個人物但顏色不同，即雜訊完全被忽略，判斷是 condition 權重太強造成 overfitting。

## L2 REGULARIZATION

為了降低 condition weights 造成的影響，generator loss 計算時加上 L2 regularized 過的 condition weights，這樣可以讓 discriminator loss 在約 40 epochs (10000 batches) 之後才收斂，如圖二。在 100 epochs 時仍能在同個 caption 下產生不同的人物，而 L2 beta 值的調整將在實驗結果敘述。

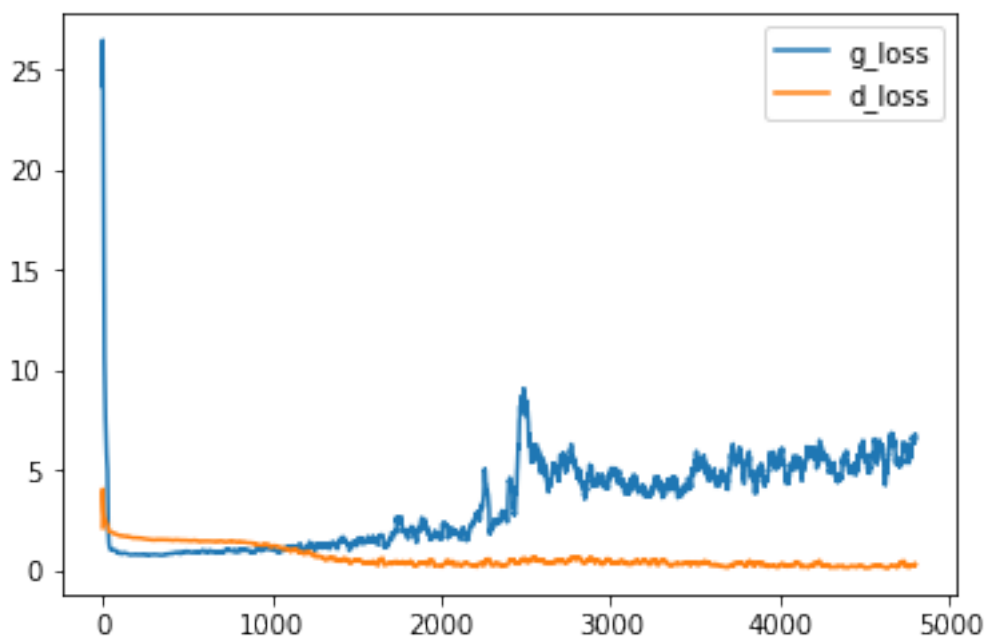


Fig. 2 Training curve with L2 regularization.

## HELP GENERATOR CHEAT

因為 discriminator 收斂之後，generator 的品質也會大幅下滑，可以觀察到 generator 一開始能逐漸描繪出人像，卻為了符合 condition 以至於最後產生色塊而沒有人物的結果。所以嘗試了幫助 generator 作弊的方法，就是有 10% 的機率將 fake image 的分數變成 real image 餵給 D loss。這個方法實驗結果雖然 90 epochs 之後就崩壞了，但搭配前面兩個方法可以在 80 epochs 時產生比之前更細緻的圖片。

## EXPERIMENTS



Fig 3 Pink hair pink eyes: beta=0.1

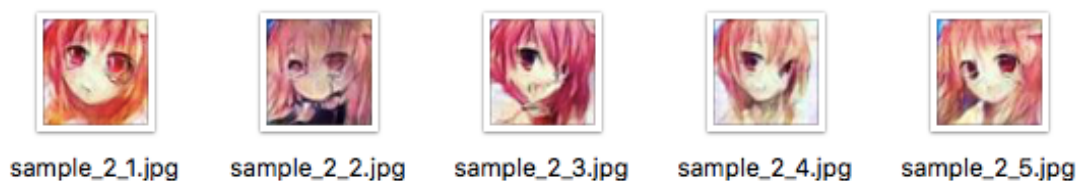


Fig 4 Pink hair pink eyes: cheating with beta=0.01



Fig 5 Pink hair pink eyes: cheating with beta=0.1