

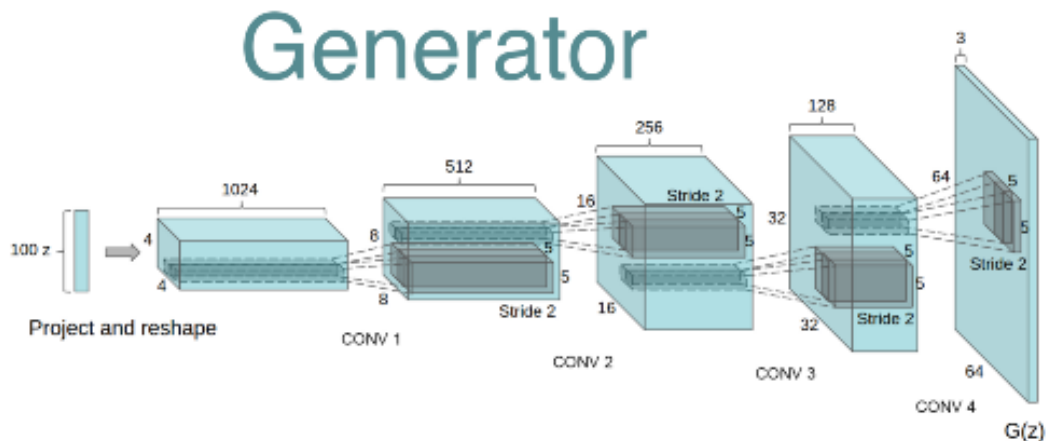
ADLxMLDS HW4 Report

R06922022 資工所碩一 曹燁文

(一) Model Description (2%)

Generator 示意圖

http://gluon.mxnet.io/chapter14_generative-adversarial-networks/dcgan.html



input 是一般 DCGAN 的做法，將 100 維 random vector z 和文字接在一起輸入。我的 model 和圖片一樣是五層，但大小不太一樣，首先通過 $4 \times 4 \times 256$ 的 dense，channel (filter 數量) 那維是 128, 64, 32, 3，都是圖中數字的 1/4，kernel size 跟 stride 與圖片相同。

Discriminator 示意圖

http://gluon.mxnet.io/chapter14_generative-adversarial-networks/dcgan.html



維度與 Generator 相同，只是在倒數第二層接上文字的資訊，再通過一層 Conv 輸出。

Loss 使用助教給的四種：

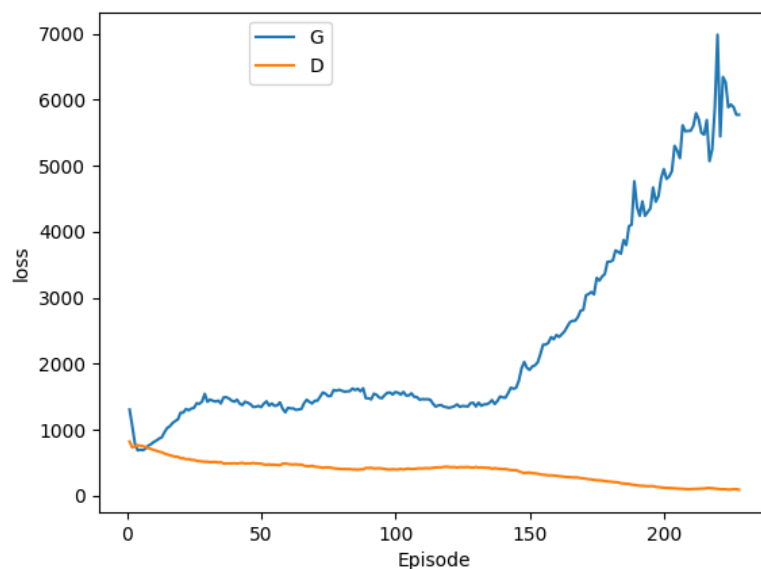
d0: (real img, right text), d1: (fake img, right text), d2: (real img, wrong text), d3: (wrong img, right text)，使用 generator 的 loss 是用 d1 與 1 計算 cross entropy，因為它希望能騙過 discriminator，所以希望 fake img 能跟 right text 配在一起。discriminator 的 loss 是 d0 與 1 以及 d1,d2,d3 與 0 計算 cross entropy，且後三者相加之後除以三，讓 right 跟 wrong/fake 的比例一樣。文字處理的部份我使用 one-hot encoding，只使用助教會 test 的 12 種 hair color 及 11 種 eyes color，共 23 維。

(二) Improve Performance (2%)

我實驗了兩種 improve 方法，一種是對照片進行更進一步的預處理，包含只留下同時擁有髮色、眼睛顏色的圖片、對圖片進行左右翻轉、順時針和逆時針旋轉特定的微小角度；另一種是更新 Generator 的次數是 Discriminator 的兩倍。第一種方法會讓圖片看起來清楚一點，但是使用之後會有圖片 test 的時候有幾張幾乎長得一樣的問題(mode collapse)。第二種方法效果很差，所以沒有使用。最後使用第一種方法但沒有挑除不完整 label 的圖片，圖見第三部分。

(三) Experiment and Observation (2%)

參數使用助教提供的版本，adam with $lr=2e-4$ ，momentum=0.5，gaussian noise dim 100，batch size 64。至於 epoch 我發現他到 130~150 epoch 就差不多收斂，再 train 下去會有類似 overfitting 的現象，也就是 G 和 D 的 loss 會越差越大，最後用眼睛判斷選擇第 68 個 epoch 的結果。



最一般的設定(未進行預處理、未挑除 label 不全的圖片、G:D=1:1)



進行預處理、挑除 label 不全的圖片、G:D=2:1



進行預處理、挑除 label 不全的圖片、G:D=1:1



最終結果：進行預處理、G:D=1:1

