

HW4 Report

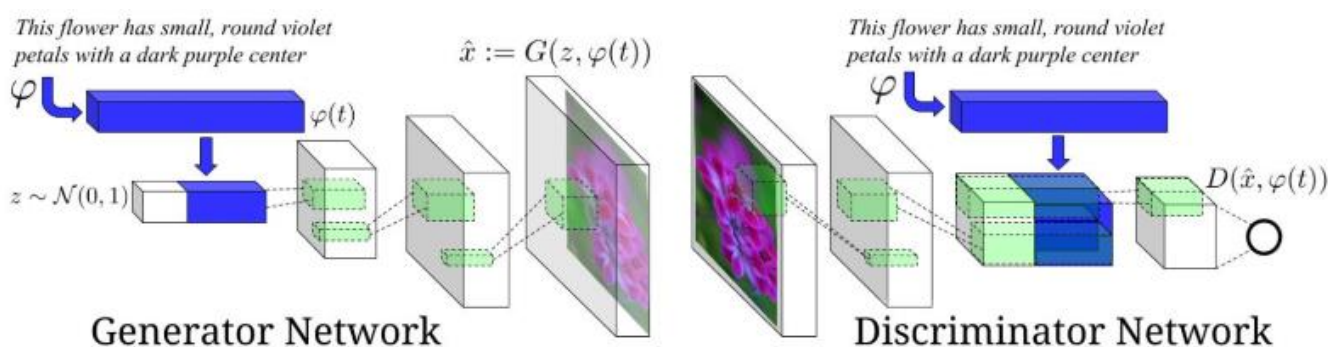
2017 ADLxMLDS

B03901163 電機三 鄭 煦

Problem description

利用 conditional GAN 來訓練 machine 根據眼睛和頭髮顏色來產生不同 comic girl 的圖片

Model description



這次的作業我參照 DCGAN 的 paper 來實做 DCGAN，架構跟 paper 的基本上相同。首先會先用 skip_thoughts 把 input 的 text encode 成 2400 維的向量，再把這個向量經過一層

dense 降維成 256，然後跟 100 維的 noise 接在一起，作為 generator 的 input。

Generator 由四層的 deconvolution 組成，而 discriminator 由四層的 convolution 組成。

Discriminator 的 input 是張 64x64 的影像。下圖是整個 model 的 objective function。

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x}|\mathbf{y})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z}|\mathbf{y})))].$$

以下是我最終所使用的參數:

Noise dim: 400, reduced text_embedding: 256, image_size: 64x64, gf_dim: 64(第一層 conv 的 output), df_dim: 64(第一層 deconv 的 input), gfc_dim: 1024(generator fully connected 的 output), dfc_dim: 1024(discriminator fully connected 的 output), epochs: 500

How to improve performance

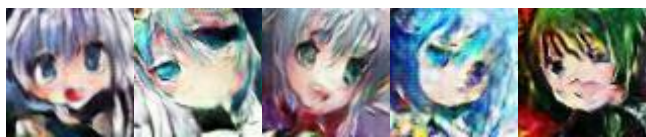
這次的作業裡面，我有試著去 implement wgan 來改善 model collapse 的問題，雖然是有改善產生重複圖的問題，可是產生的圖品質都蠻差的，下面是 wgan 在 500 個 epoch 產生的圖片品質。所以最後我還是採用 DCGAN 來當最後上傳的 model。



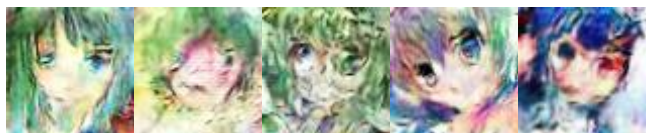
Experiment settings and observation

這次的作業我有試過調整 optimizer, noise dimension,和 text embedding 的向量長度。

這是 adam optimizer 的結果

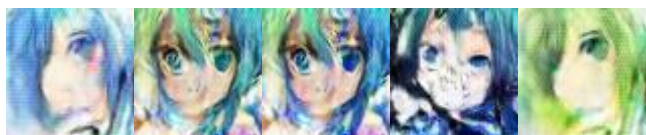


這是 rms prop 的結果

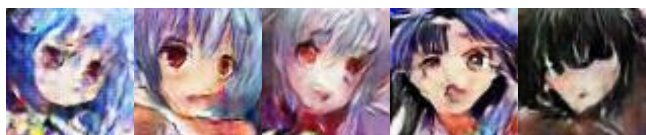


可以明顯發現 adam 的圖片品質比較好一些。

這是 text embedding 130 維的結果



這是 text embedding 256 維的結果

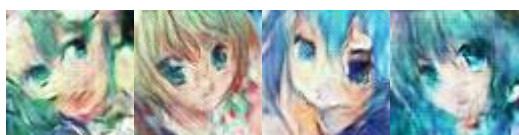


這個部份我覺得差異沒有很大，不過 256 的看起來好一點點。

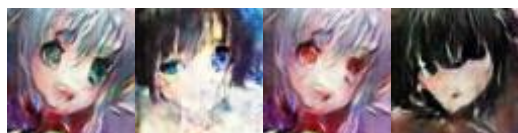
這是 noise dim=100 的結果



這是 noise dim=500 的結果



這是 noise dim=400 的結果



一開始我本來只是想試看看提高 noise 維度是不是可以解決 model collapse 的問題，然而並沒有效，不過很湊巧地發現提高了 noise dimension 後，可以產生品質在好一些的圖片。