

隊名：深的不要不要的

- **Environment (1%)**

- OS 2016 x86_64 GNU/Linux
- CPU Intel(R) Core(TM) i7-4790K CPU @ 4.00GHz
- Memory
- GPU NVIDIA Corporation GM200 [GeForce GTX 980 Ti]
- Libraries:

Tensorflow 1.0

Nltk 3.2.2

Genism 1.0.1

Pandas 0.19.2

- **Model description (3%)**

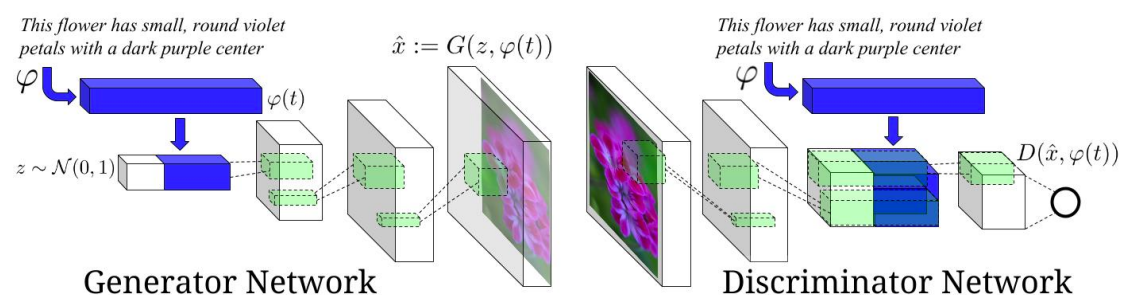
- *Word-embedding*

先 filter 過與五官相關的 tag，'eye', 'ear', 'nose', 'mouth', 'hair', 'tear', 'hat', 'male', 'female'

再用 skipthoughts 轉成 4800 維，取前面的 2400 維

- *Model*

跟 paper Generative Adversarial Text-to-Image Synthesis 一樣，由 text feature 訓練一個 DC GAN。將 Generator model 產生的圖片丟進 Discriminator Network，讓其分辨是否符合 tag



- **How do you improve your performance (5%)**

1. 圖片正規化：除以 255，會從原本 train 不起來的狀況，進步到 train 出人像
2. tag 降維：Skipthought 從原本的取 4800 維，改成取 2400 維，原本的 4800 會將雜訊蓋掉，因此產生出來的動漫人物的臉會都一樣，但是若取 2400 為的話，會有臉的變化。

- **Experiment settings and observation (5%)**

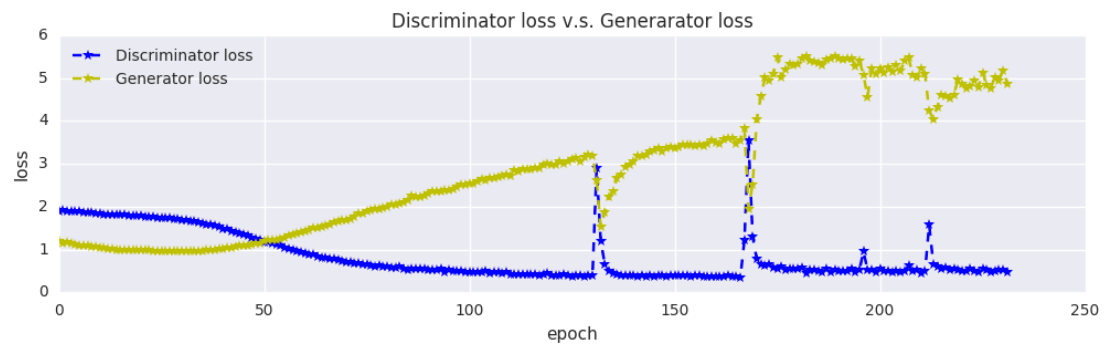
- *Parameters:*

$z_dim = 100$ / $t_dim = 256$ / $image_size = 64$ / $gf_dim = 64$ / $df_dim = 64$ / $gfc_dim = 1024$ / $caption_vector_length = 2400$ / $batch_size = 64$ /
至 $epoch = 300$ / $learning\ rate = 0.0002$

- 時間（一次 *epoch*）：

→ 約 28.8 分鐘

- *Loss*



- *Observation*

Model train epoch 大約到 160 的時候會壞掉，但因為 train 一次要花 5 天的時間，所以來不及解決這個問題

- **Team division (1%)**

r05521608 土木碩一 任精瑋: model design

r04521603 土木碩二 魏仰廷: model analysis

r05631018 生機碩一 曾秋旺: model trial

- **Reference:** [Generative Adversarial Text-to-Image Synthesis](#) Scott Reed, Zeynep

Akata, Xincheng Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee