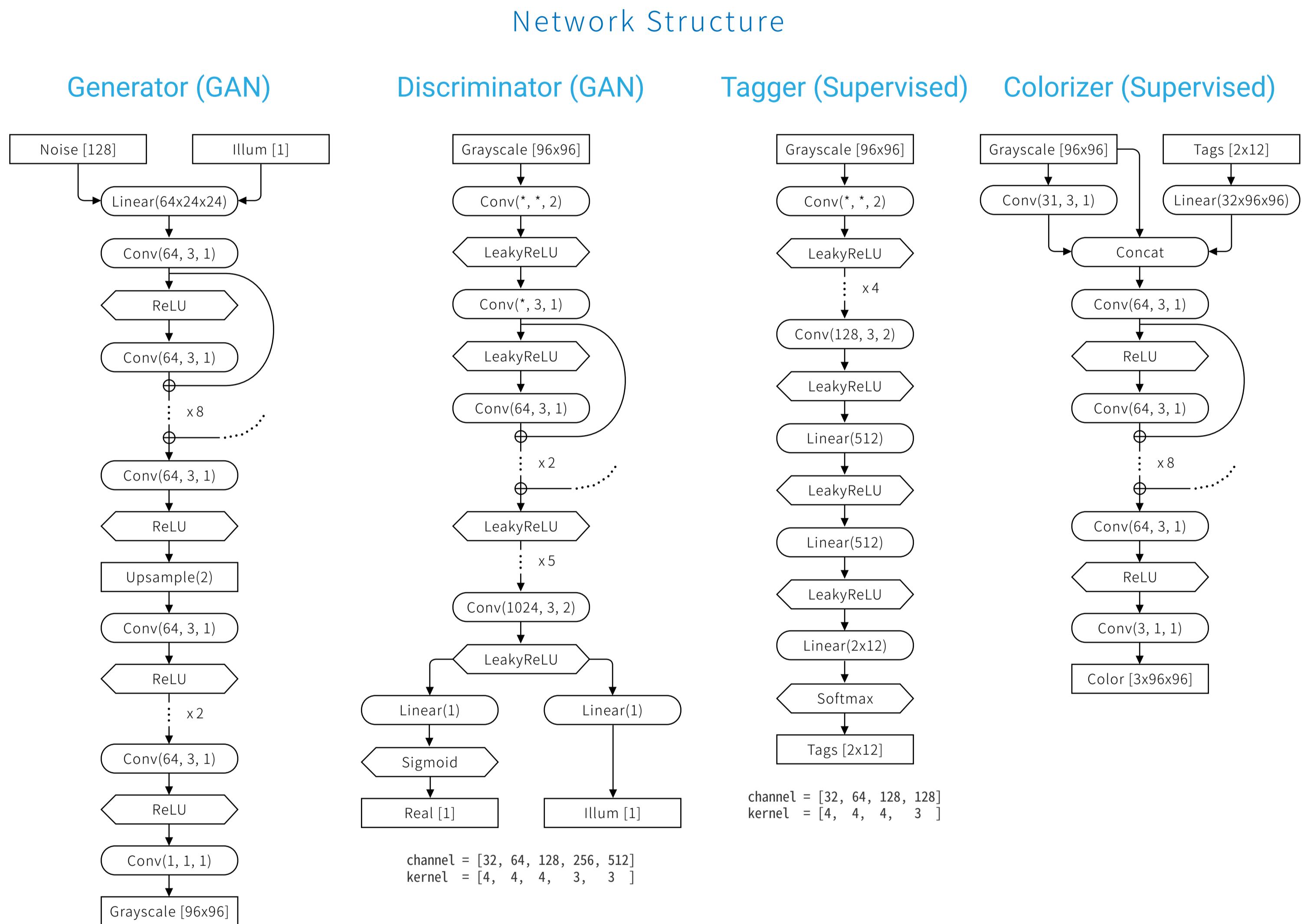


ADL Homework #4 Report

b04902013 鄧逸軒

1. Model Description

GAN 的架構是以 make.girls.moe ver. Camellia (arxiv 1708.05509) 的架構為基礎做改進的，基本上就是用 ResNet 組成的 ACGAN. GAN 的部份輸入條件換為頭髮顏色的深度，產生出灰階的圖像，之後交給 Colorizer 進行上色. Tagger 是為了用來重新標記資料，使用髮色與瞳色都只有一種的資料進行訓練，約有 11000 筆.



1.1 Training Objective

GAN的 Loss 大致與 ACGAN + DRAGAN 一樣. 唯一不一樣的地方是我拿掉了訓練 Auxiliary Classifier 時，判斷 Generator 產生圖片種類的 loss. 因為我發現分類器會很容易抓錯特徵，Generator 準確率很快就到 95% 可是分類其實是錯的.

Tagger 的部份是 Categorical XENT，而 Colorizer 是 MSE.

2. Data Collection

Extra Dataset 來自 Pixiv, 利用他的推薦系統對畫師進行排序，篩出約前 1000 名畫風較好且一致的畫師後，載下推薦的圖片，用 lbpcascade_animeface 切出頭像的範圍，最後約有70000張頭像. Download: <https://goo.gl/ZJ6yb5>

3. How to improve your performance

3.1 Grayscale

產生灰階圖像再另外上色使得效果變好許多，我猜想是因為連續的條件(頭髮深度)更容易共享訓練出來的權重，而且同種類的訓練資料也變多，最後的成果細節豐富許多。如果直接分顏色進行生成，常常會因為資料分佈不均，造成某些顏色的組合發生 collapse 或不好看。之所以要有髮色深度的條件，是因為 Colorizer 會根據灰階的深度來決定顏色的深度(Sec. 5)，黑灰白髮要靠輸入深度來控制，訓練集的髮色深度是抓 [8:24, 16:-16] 範圍的中位數。

3.2 Extra Dataset

換了新的 dataset 也獲得了一些進步。訓練出來的圖片畫風較新，而且因為解析度滿 96x96，細節比較清楚(e.g. 髮絲)。

3.3 Noise Intensity

生成的時候調低雜訊的強度，可以讓產生分佈比較集中，降低圖片壞掉機率，同時因為圖片比較接近，放在一起看的時候比較和諧。

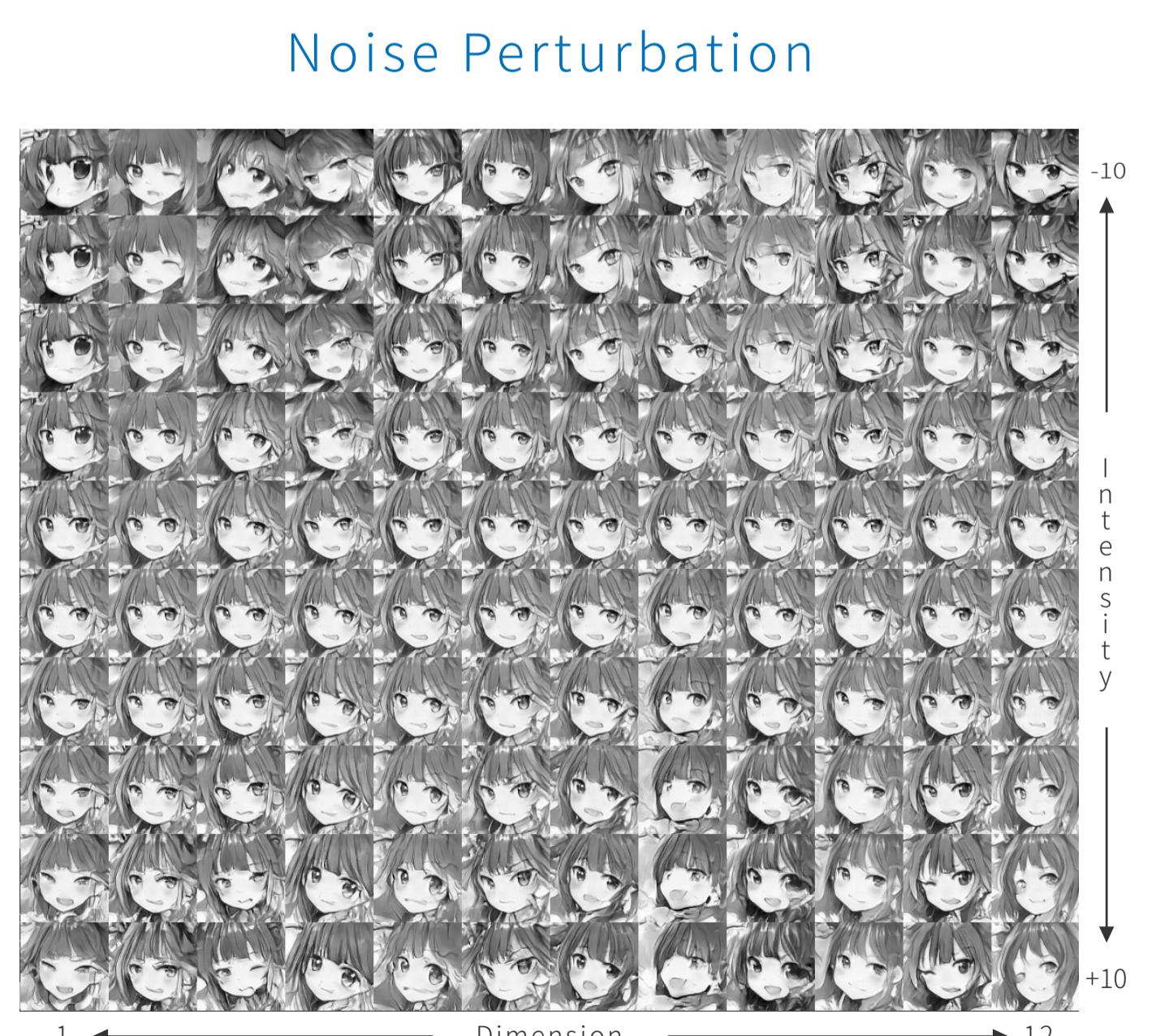
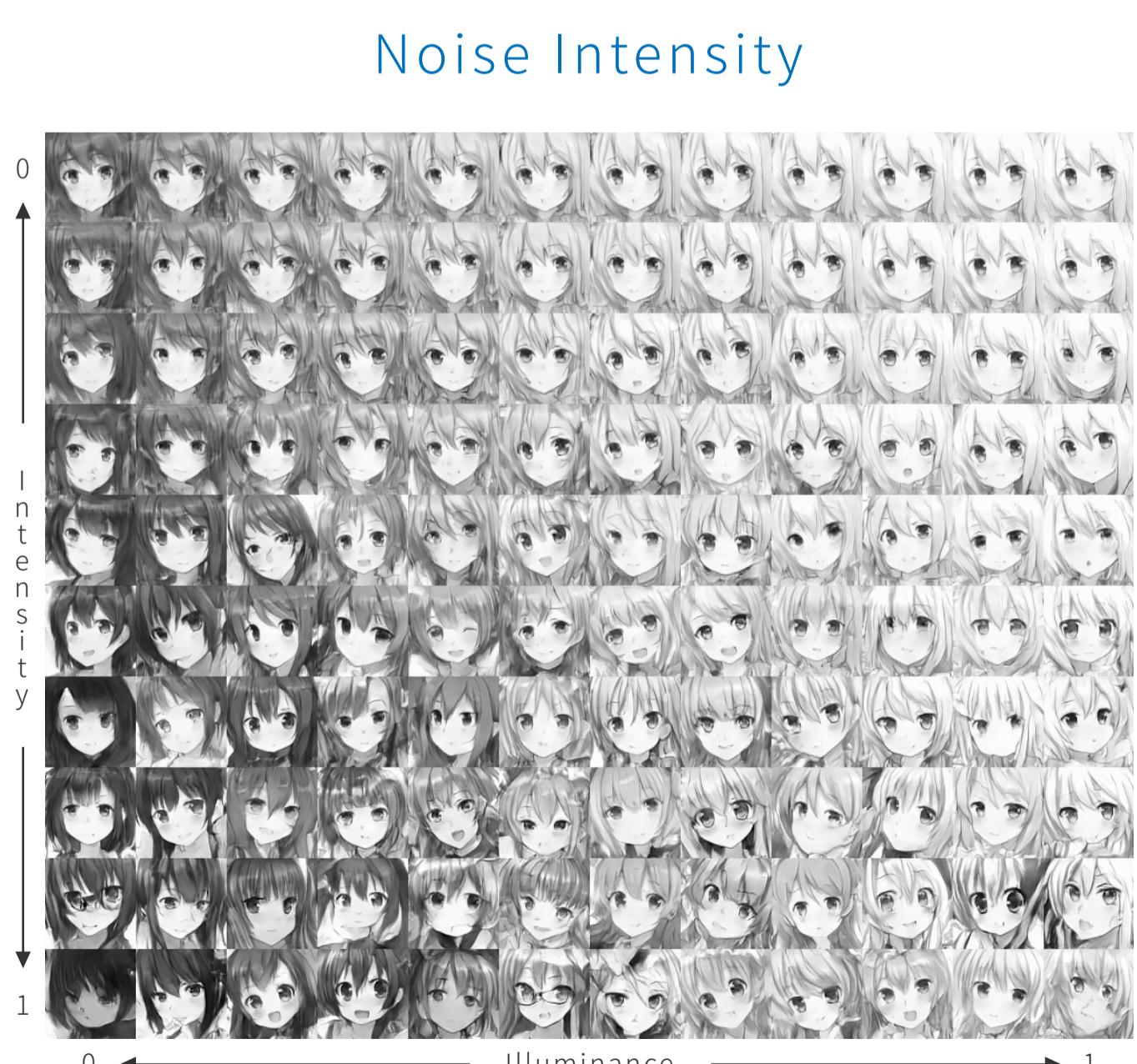
3.4 Experience Replay

訓練 Discriminator 的時候混合以前生成的圖片，可以使得目標較為穩定。

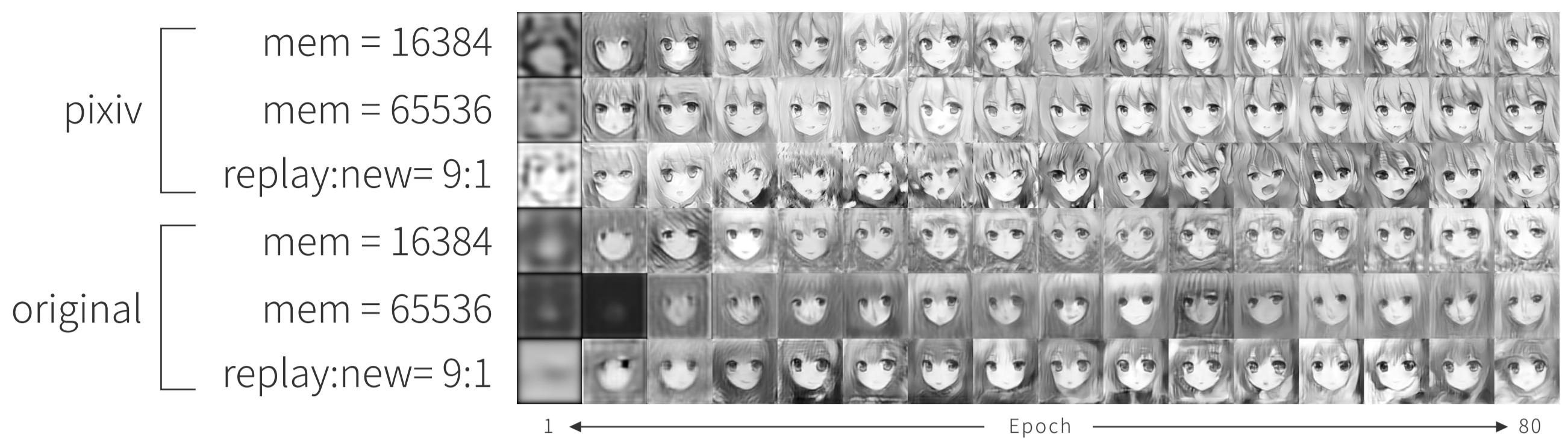
4. Experimental Settings and Observations

4.1 Generator

Noise 包含了除了髮色的所有資訊，可以看出對 noise 做擾動可以改變眼睛的樣式，臉型，角度等等。



Generator Training Progress



4.2 Colorizer

很多 Image-to-Image translation 都是使用 L1, 不過我訓練的時候發現 L1 會在訓練一陣子後突然壞掉, 可能是因為底部不連續的部份. 改用 MSE 之後就能穩定的訓練.

Colorizer Training Progress



5. Style Transfer

原本的模型中, 最後上色的部份就是一種灰階轉彩色的 style transfer 任務, 我另外實做了一個 GAN 的版本, 架構上輸入多加上灰階的圖片, 並且 Generator 對於真實 tag 加上 MSE loss. 從下面的圖片可以看出 GAN 可以處理髮色深度不對的問題, 但是部份類別出現了不管輸入圖片的問題.

GAN Colorizer



Supervised Colorizer



6. Final Result

GAN Result (Original Dataset)
GrayScaled Output + Colorizer



GAN Result (Pixiv Dataset)
GrayScaled Output + Colorizer



GAN Result (Pixiv Dataset)
RGB Output



Samples from Pixiv Dataset

