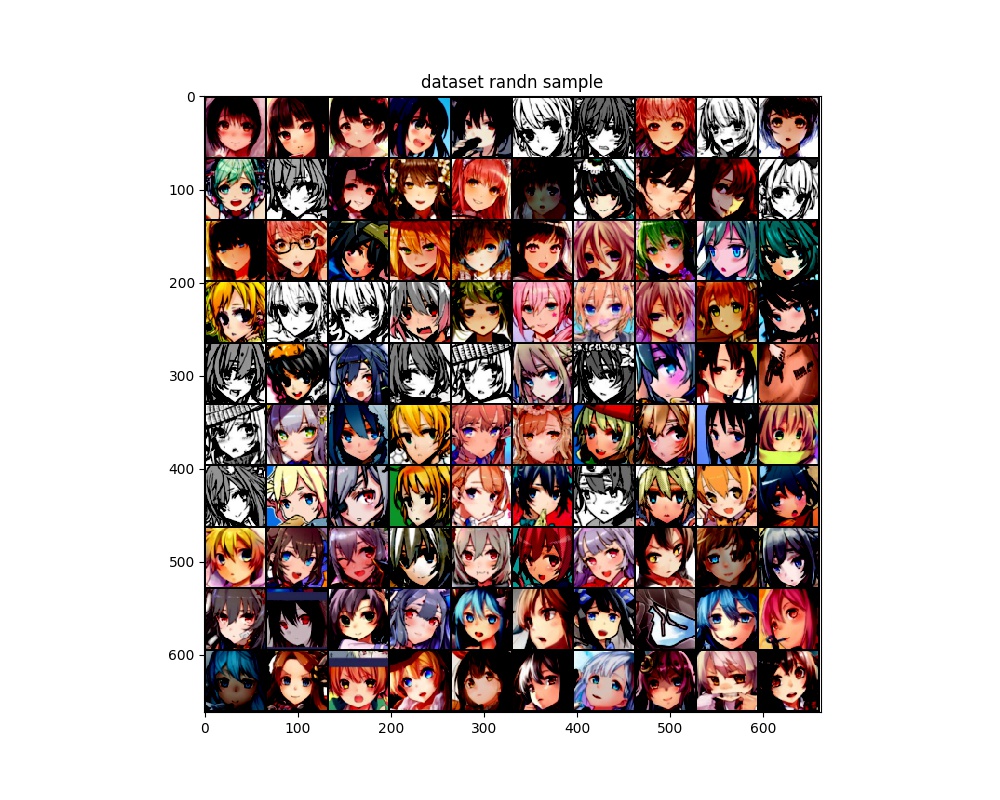
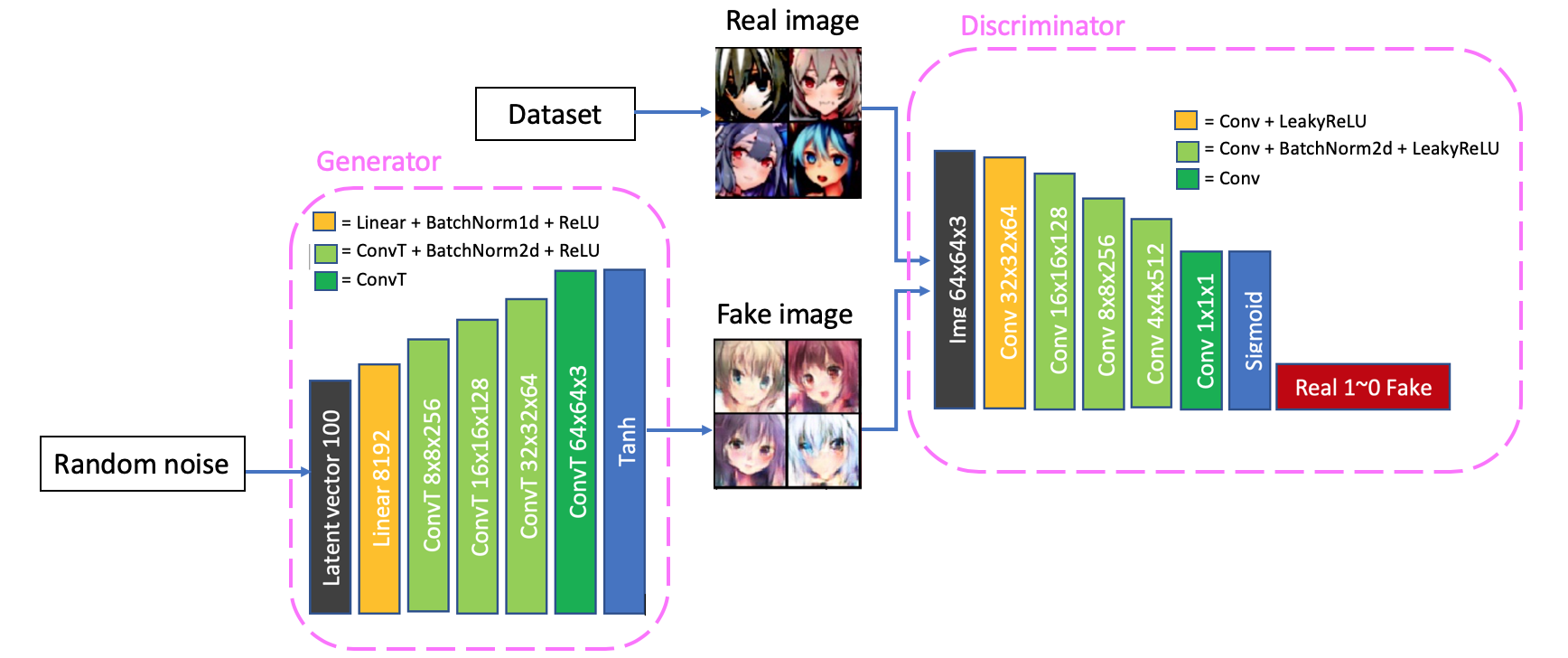
學號：r07921001 系級：電機所二 姓名：李尚倫

1. (2.5%) 訓練一個 model。
2. (1%) 請描述你使用的 model（可以是 baseline model）。包含 generator 和 discriminator 的 model architecture、loss function、使用的dataset、optimizer 參數、以及訓練 step 數（或是 epoch 數）。

Dataset為71314張96x96x3的二次元妹子圖片, batch size = 64, sampling前100張如下圖：



Discriminator ＆ Generator architecture為：(baseline model, dcgan)



Disc Loss =

Gen Loss =

Optimizer為Adam, lr = 1e-4, betas=(0.5, 0.999)

1. (1.5%) 請畫出至少 16 張 model 生成的圖片。

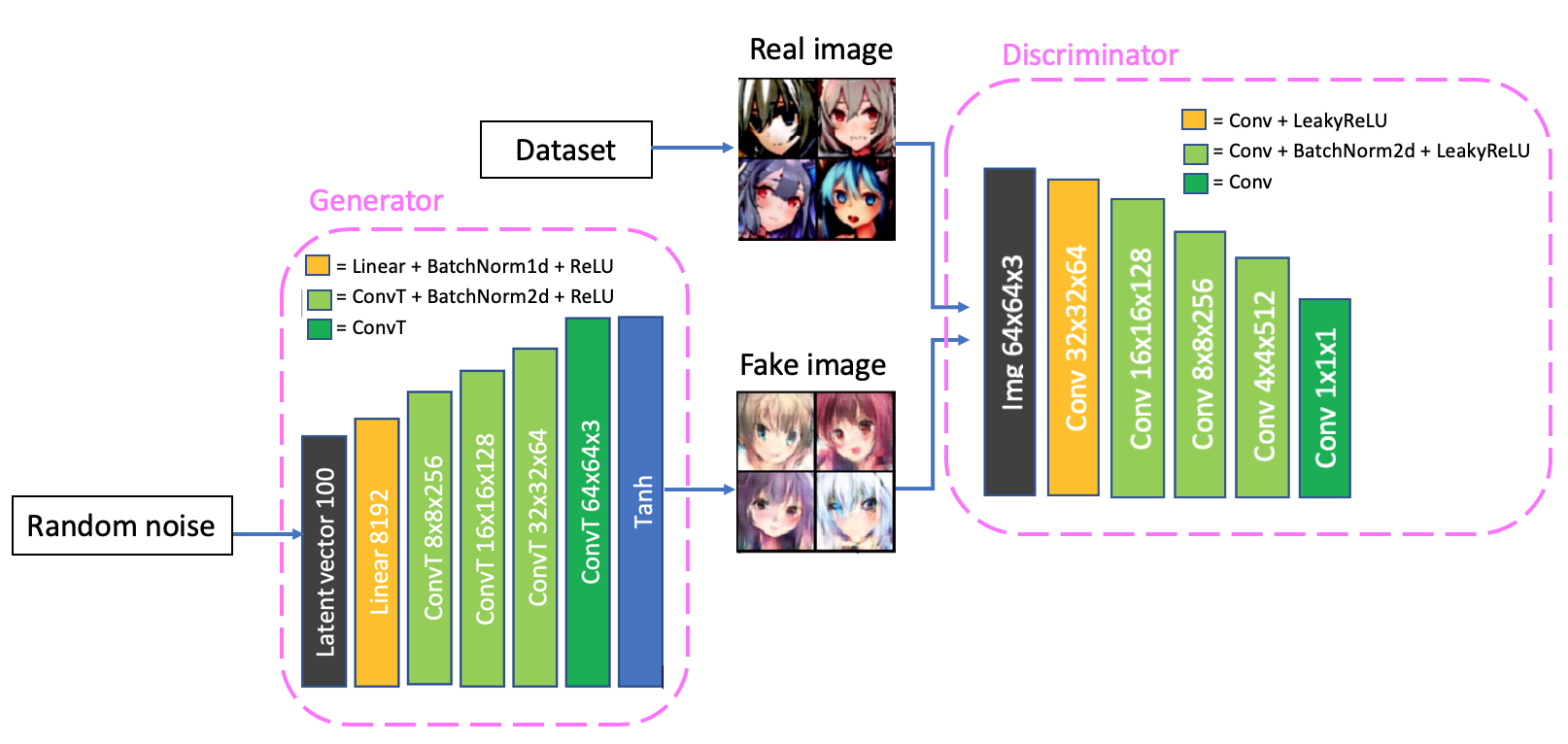
下方為隨機生成的16張圖片，是在第9個epoch取的，randn seed為34。



1. (3.5%) 請選擇下列其中一種 model： WGAN, WGAN-GP, LSGAN, SNGAN（不要和 1. 使用的model 一樣，至少 architecture 或是 loss function 要不同）
2. (1%) 同 1.a ，請描述你選擇的 model，包含 generator 和 discriminator 的 model architecture、loss function、使用的dataset、optimizer 參數、及訓練 step 數（或是 epoch 數）。

Dataset為71314張96x96x3的二次元妹子圖片, batch size = 64, 同第一題。

Discriminator ＆ Generator architecture為：



Disc Loss =

Gen Loss =

Gradient Penalty =

= 250, n\_critic = 3

Optimizer為Adam, lr = 20\*1e-4, betas=(0.5, 0.999)

1. (1.5%) 和 1.b 一樣，就你選擇的 model，畫出至少 16 張 model 生成的圖片。

下方為隨機生成的16張圖片，是在第39個epoch取的，randn seed為4。



1. (1%) 請簡單探討你在 1. 使用的 model 和 2. 使用的 model，他們分別有何性質，描述你觀察到的異同。

首先wgangp生成的影像清晰度稍微比dcgan來的偏模糊一點點點，因為前者是屬於wgan系列的loss的算法完全不一樣，以結果來看dcgan的minimize binary cross entropy loss (=maximize JS divergence) 看起來會使圖片在清晰度上有比較好的表現，但wgangp的訓練穩定度就比dcgan來的好一點，使用的是wasserstein distance (with 1 – 𝐿𝑖𝑝𝑠𝑐h𝑖𝑡𝑧 weight clipping)，多勸幾個epoch也不會導致model collapse，詳細結果可見第三題。而圖片的品質部分，兩個model生成的妹子在隨機sample下都有眼睛，都沒有到整個爛掉，可愛的也很可愛，我個人是覺得沒有差太多。

1. (4%) 請訓練一個會導致 mode collapse 的 model。
2. (1%) 同 1.a ，請描述你選擇的 model，包含 generator 和 discriminator 的 model architecture、loss function、使用的dataset、optimizer 參數、及訓練 step 數（或是 epoch 數）。

使用與第一題的model (dcgan) 與parameter，詳見第一題，取第18 epoch：

17ep (如下圖，尚未collapse)， 18ep(如下圖，已完全collapse)

1. (1.5%) 請畫出至少16張 model 生成且具有mode collapse現象的圖片。

下方為隨機生成的16張圖片，是在第18個epoch取的，randn seed為4。



1. (1.5%) 在不改變 optimizer 和訓練 step 數的情況下，請嘗試使用一些方法來減緩 mode collapse。說明你嘗試了哪些方法，請至少舉出一種成功改善的方法，若有其它失敗的方法也可以記錄下來。

法一：使用wgan系列的model可以避免collapse的問題，如用第二題的wgangp(左)取一樣的optimizer parameter在第18ep下的比較原本dcgan(右)

但由於第二題的wgan\_gp要能勸起來的參數不太一樣，因此雖然沒有collapse但在同樣的參數下很模糊，因此調整了一下wgangp的參數成和第二題一樣的參數，ep18的結果如下圖(左)，ep39的結果如下圖(右)，基本上並不會像原本有勸越多個epoch越容易collapse的問題。

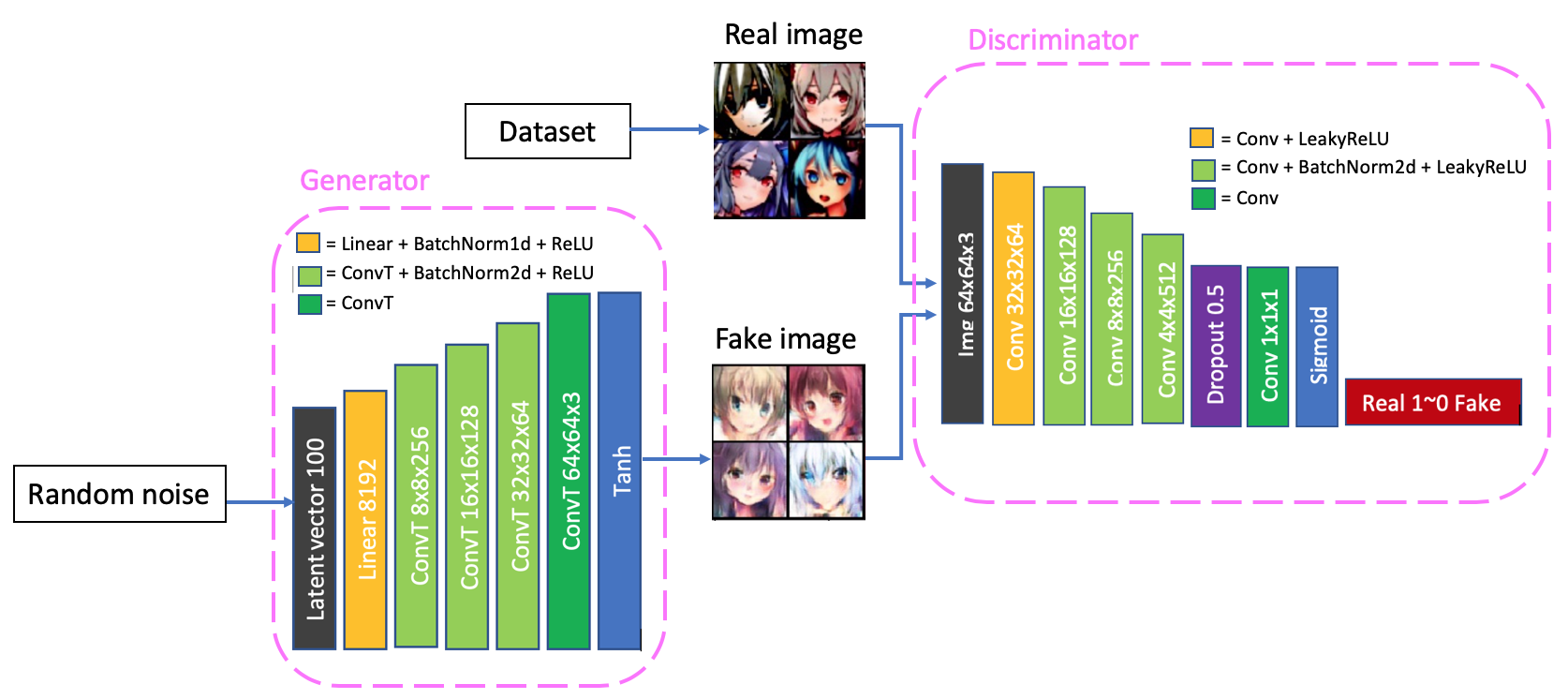
而這是因為wgan改進了dcgan等基本gan的假設，即假設Pdata和Pgen這兩個高維空間中的低維manifold有重疊，因此可以用maximize LS divergence (minimize binary cross entropy) 的方式求loss，而wgan認為Pdata和Pgen在高維空間中可能不會重疊，因此要用wasserstein distance來求loss，效果就是能夠比較穩定的收斂而不會導致collapse。

法二：用與原本第一題相同model，相同的參數，但Discriminator的最後一個Conv層之前加一個dropout=0.5層，來減少overfitting的導致collapse的機會，一樣勸18個epoch結果如下(左)，勸30個epoch的結果如下(右)：

而其他嘗試如：把dropout也加到generator的最後一個conv之前，但這麼做會導致model勸不太起來，而只加入到discriminator會有幫助可以理解為，因為discriminator太overfitting於某種臉型，才會導致generator也致力於生成該種臉型導致collapse，因此dropout僅需要加在discriminator即可解決問題。

Network architecture:



Raw figures:

Img 64x64x3

Conv 32x32x64

Conv 16x16x128

Conv 8x8x256

Conv 4x4x512

Dropout 0.5

Conv 1x1x1

Sigmoid

Real 1~0 Fake

Img 64x64x3

Conv 32x32x64

Conv 16x16x128

Conv 8x8x256

Conv 4x4x512

Conv 1x1x1

Sigmoid

Real 1~0 Fake

Img 64x64x3

Conv 32x32x64

Conv 16x16x128

Conv 8x8x256

Conv 4x4x512

Conv 1x1x1

Latent vector 100

Linear 8192

ConvT 8x8x256

ConvT 16x16x128

ConvT 32x32x64

ConvT 64x64x3

Tanh