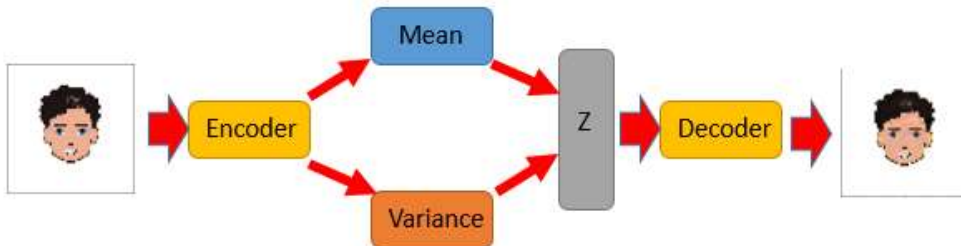


Deep Learning (Homework 3)

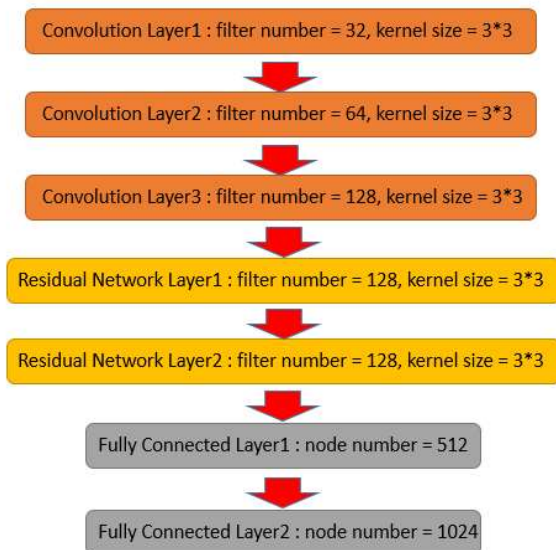
Prob1

這次使用 VAE 來做圖像重建，首先將輸入的圖像 data 將過 encoder 得到 mean 與 variance，變數 z 藉由這兩個參數做取樣得到，接著將變數 z 做 decoder 可以還原回原本的圖像。

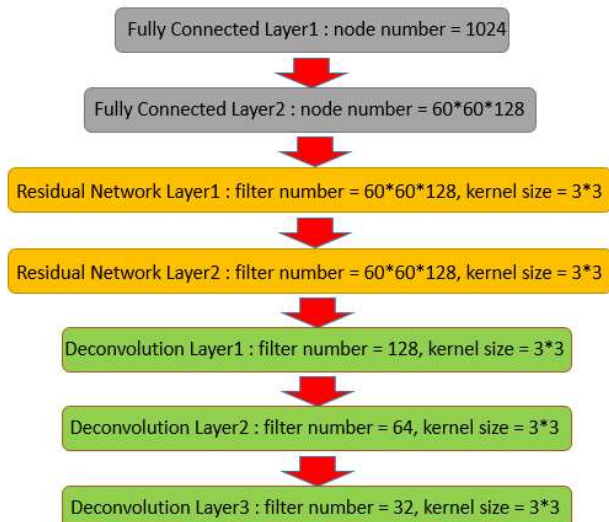
其 VAE 架構如下：

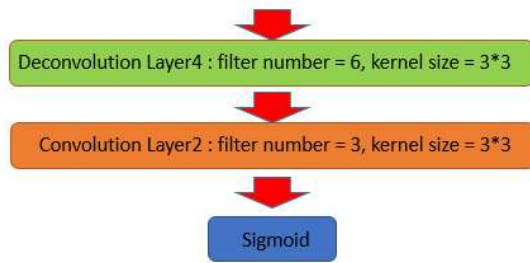


其中 Encoder 架構如下：



其中 Decoder 架構如下：





VAE 之損失函數推導如下：

來源：变分自动编码器(VAE)理解和实现(Tensorflow)

$$\begin{aligned}
 KL(q(z|x)||p(z|x)) &= \int q(z|x) \log \frac{q(z|x)}{p(z|x)} dz \\
 &= \int q(z|x) \left(\log q(z|x) - \log \frac{p(z,x)}{p(x)} \right) dz \\
 &= \int q(z|x) (\log q(z|x) - \log p(z,x) + \log p(x)) dz \\
 &= \int q(z|x) (\log q(z|x) - \log p(z,x)) dz + \log p(x) \\
 &= E_{z \sim q(z|x)} \log \frac{q(z|x)}{p(z,x)} + \log p(x)
 \end{aligned}$$

$$L(x) = E_{z \sim q(z|x)} \log \frac{p(z,x)}{q(z|x)} = \log p(x) - KL(q(z|x)||p(z|x))$$

$$\begin{aligned}
 L(x) &= E_{z \sim q(z|x)} \log \frac{p(z,x)}{q(z|x)} \\
 &= E_{z \sim q(z|x)} \log \frac{p(x|z)p(z)}{q(z|x)} \\
 &= \int q(z|x) (\log p(z) - \log q(z|x) + \log p(x|z)) dz \\
 &= - \int q(z|x) \left(\log \frac{q(z|x)}{p(z)} \right) dz + \int q(z|x) \log p(x|z) dz \\
 &= -KL(q(z|x)||p(z)) + E_{z \sim q(z|x)} (\log p(x|z))
 \end{aligned}$$

$$E_{z \sim q(z|x)} (\log p(x|z)) \approx \frac{1}{L} \sum_{l=1}^L \log p(x|z_l), z_l \sim q(z|x)$$

$$-KL(q(z|x)||p(z)) = \frac{1}{2} \sum_{j=1}^J (1 + \log((\sigma_j)^2) - (\mu_j)^2 - (\sigma_j)^2)$$

$$\begin{aligned}
 L(x) &= -KL(q(z|x)||p(z)) + E_{z \sim q(z|x)} (\log p(x|z)) \\
 &= \frac{1}{2} \sum_{j=1}^J (1 + \log((\sigma_j)^2) - (\mu_j)^2 - (\sigma_j)^2) + \frac{1}{L} \sum_{l=1}^L \log p(x|z_l)
 \end{aligned}$$

以下為此次我 Training 的參數：

Learning rate= 1e-04 → 3.0336164e-06

Epoch times= 50

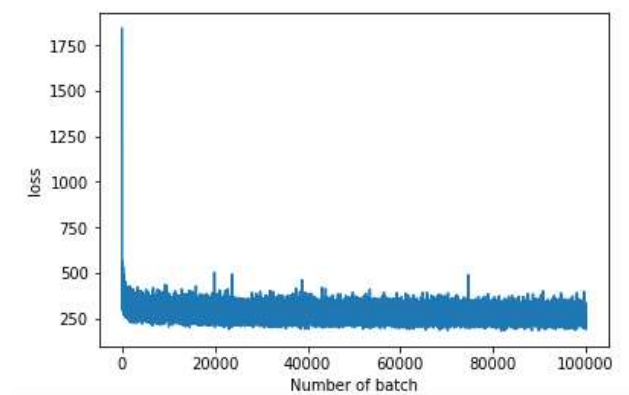
Batch size=5

Data number=10000 (cartoon 10000 images)


Result:

Loss= 267.9562744140625

Learning curve per batch 如下:



以下我會將 Train 好的 model 丟入 15 筆圖片，每張圖片會先 resize 成 60*60 pixel 的大小。

原圖片	經過 encoder 再經 decoder
	
	
	







接著換做下部分的測試

以下為此次我 Training 的參數：

Learning rate= $1e-04 \rightarrow 7.1793643e-06$

Epoch times= 20

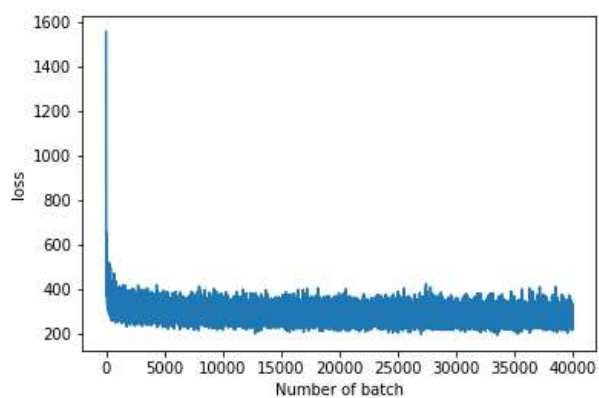
Batch size=5

Data number=10000 (cartoon 10000 images)





Result:

Loss= 264.4379638671875


Learning curve per batch 如下：



我隨機取了 2 張 image 當測試







原圖片	經過 encoder 再經 decoder
	
	

我取出這兩張圖經過 encoder 並且取樣後的向量 $z1$ 與 $z2$
我想將這兩個向量做線性組合，觀察輸出的圖片

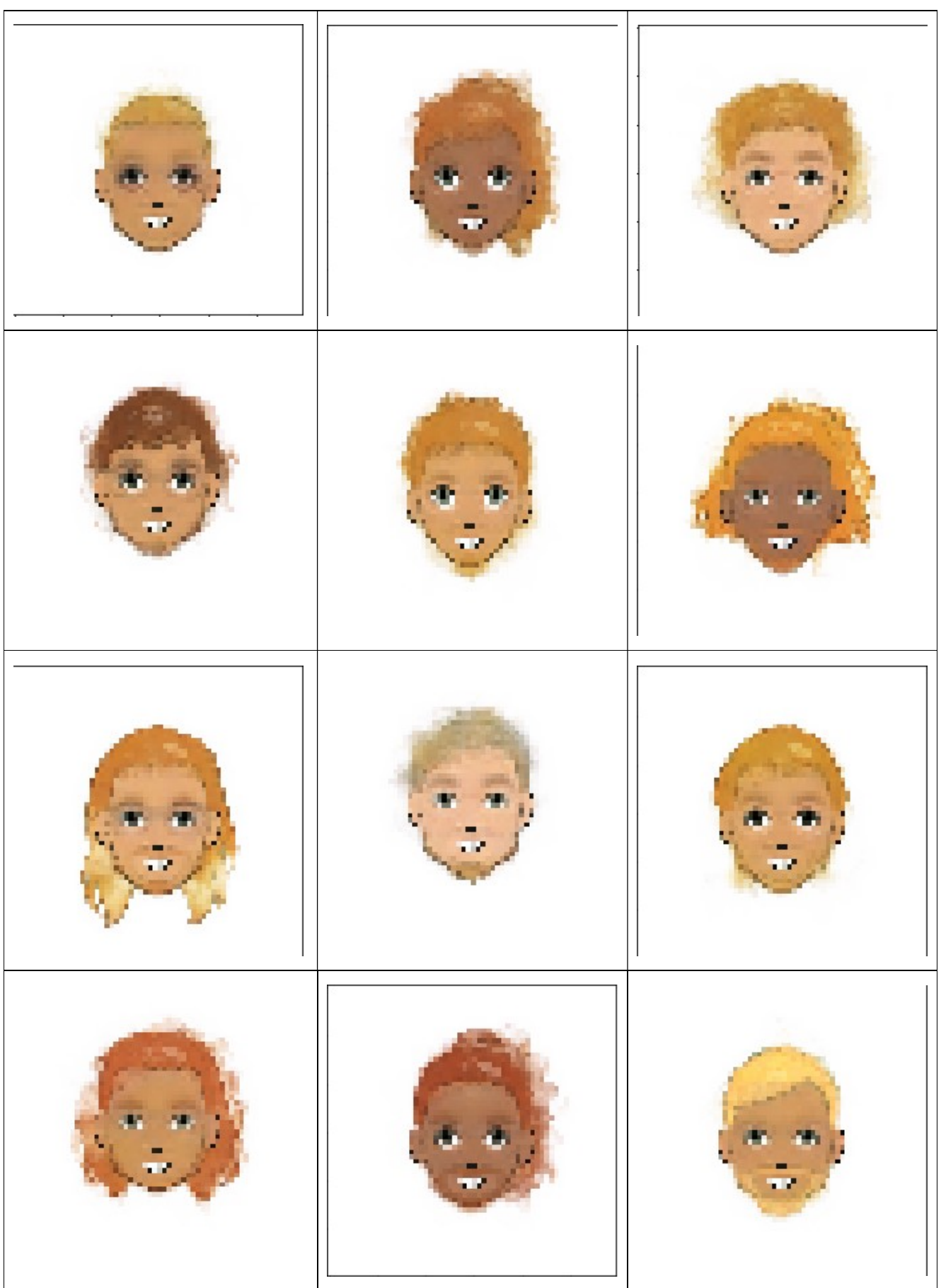
$Z=0.75*z1+0.25*z2$	$Z=0.5*z1+0.5*z2$	$Z=0.25*z1+0.75*z2$
		

可以看出來並不能使用向量線性組合得到在兩張片之間的圖片，而是產生了新的圖片。

以下為我輸入高斯分布(平均值=0, 標準差=0.8)的隨機向量 z 產生的結果：





膚色以皮膚色居多、頭髮以棕色居多、有戴墨鏡的人很少且生成的不是很明顯、圖片有點模糊像用蠟筆畫的