機器學習 HW#8 Report　資管碩一 R10725018 郭宇雋

1. *Make a brief introduction about variational autoencoder (VAE). List one advantage comparing with vanilla autoencoder and one problem of VAE*

VAE 在 Encoder 的輸出中加入了 Noise。具體方法是讓 Encoder 輸出兩組向量 m 與 σ，再從高斯分配抽樣一組向量 e，將 exp(σ) 乘上 e，最後加上 m，以得到最終的 Code。最後再將這個加過 Noise 的 Code 餵進 Decoder 產生最終輸出。

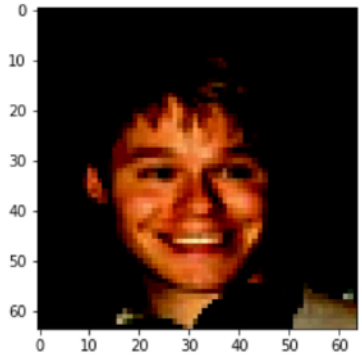
因為 VAE 在 Code 上面加入了 Noise，因此對於某個 Input，他對應的 Code 範圍會變得更廣，不單單只是 Code Space 上的一個點，因此 Reconstruct 的結果會比較好；此外，也能讓 Code Space 上兩輸入中間的位置能同時被兩個輸入的 Code 涵蓋到，藉此產生真正介於兩輸入之間的 Reconstruct 結果。

由上述可知，因為加入了 Noise，因此從 VAE 的 Code Space 上面抽樣 Code，會產生較好且較真實的 Reconstruct (Decode) 結果；而若是使用 Vanilla Auto-encoder，會比較容易產生不真實的結果。

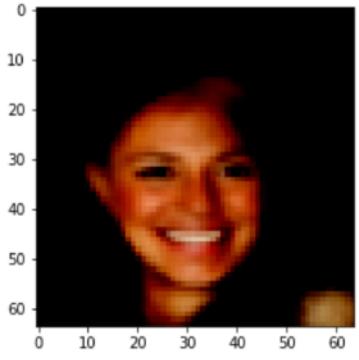
而 VAE 一個很大的問題就是，他的目標只是想生成一個跟訓練資料中某個資料很像的結果，而非生成一個真正真實且全新的結果，也就是 VAE 其實只學會了「模仿」這件事。因此生成的圖片中任何一個 Pixel 有錯，對 VAE 來說的傷害都是一樣的，但真實世界的應用並非如此。

2. *Train a fully connected autoencoder and adjust at least two different element of the latent representation. Show your model architecture, plot out the original image, the reconstructed images for each adjustment and describe the differences*

原始圖片：

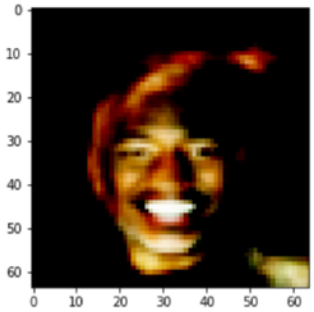


原始結果：



第一種調整：把 Latent Repr. (Dim=64) 第零維的數值加上 20

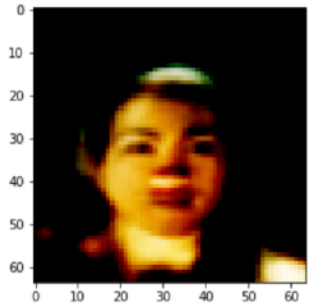
結果：



差異：從小男孩變成戴頭巾的黑人

第二種調整：把 Latent Repr. (Dim=64) 最後一維的數值加上 20

結果：



差異：從小男孩變成看起來像是戴一頂小帽子的女人

模型架構：

```
self.encoder = nn.Sequential(
    nn.Linear(64 * 64 * 3, 1024),
    nn.LeakyReLU(0.1),
    nn.Linear(1024, 512),
    nn.LeakyReLU(0.1),
    nn.Linear(512, 256),
    nn.LeakyReLU(0.1),
    nn.Linear(256, 128),
    nn.LeakyReLU(0.1),
    nn.Linear(128, 64),
)

self.decoder = nn.Sequential(
    nn.Linear(64, 128),
    nn.LeakyReLU(0.1),
    nn.Linear(128, 256),
    nn.LeakyReLU(0.1),
    nn.Linear(256, 512),
    nn.LeakyReLU(0.1),
    nn.Linear(512, 1024),
    nn.LeakyReLU(0.1),
    nn.Linear(1024, 64 * 64 * 3),
    nn.Tanh()
)
```