機器學習 hw8 report

R11922131 資工碩一 周柏諺

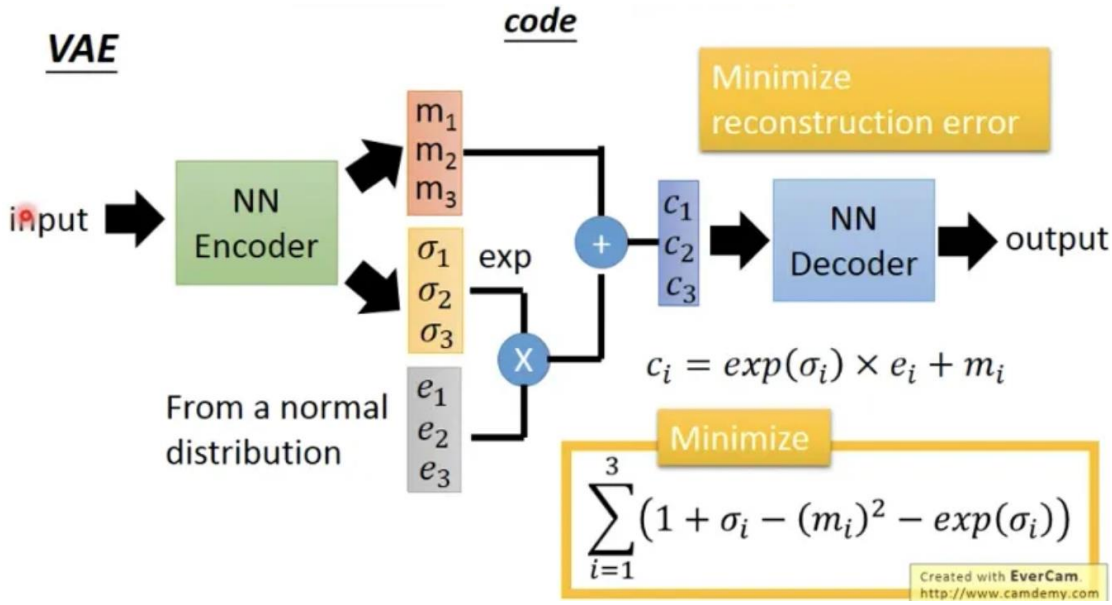
==========

Problem 1:

Choose a variation of autoencoder. Show an image of the model architecture. Then, list an advantage and a disadvantage comparing with vanilla autoencoder. Also, put on the paper link as reference. Eg, denoising autoencoder, variational autoencoder, etc.

Ans:

variational autoencoder:



優:

在 code 上加入 noise，使得在 reconstruction 時在 code space 重疊處能夠同時 minimize mean square error，連續的 code space 讓最後產生的圖介於其中，得到較好的圖片。若是使用 vanilla autoencoder，可能產生的圖片都不像任一真實得圖片。

缺:

同時也因為會加入 noise，映射後是一個機率分布，而非特定值，所以當面對較複雜的資料分布時，重建的結果可能會較差。

Ref: https://arxiv.org/abs/1312.6114

Ref :

https://www.youtube.com/watch?v=YNUek8ioAJk&list=PLJV_el3uVTsPy9oCRY30oBPNLCo89yu49&index=27

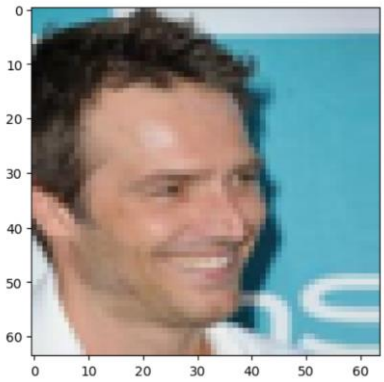
Problem 2:

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.

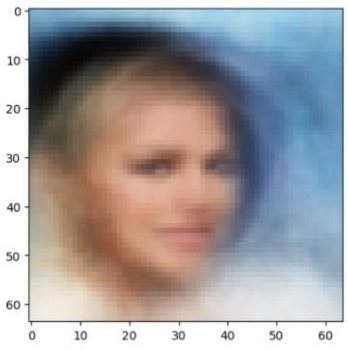
```
model
```

```
[18]: fcn_autoencoder(
        (encoder): Sequential(
          (0): Linear(in_features=12288, out_features=512, bias=True)
          (1): ReLU()
          (2): Linear(in_features=512, out_features=24, bias=True)
        )
        (decoder): Sequential(
          (0): Linear(in_features=24, out_features=512, bias=True)
          (1): ReLU()
          (2): Linear(in_features=512, out_features=12288, bias=True)
          (3): Tanh()
        )
      )
```

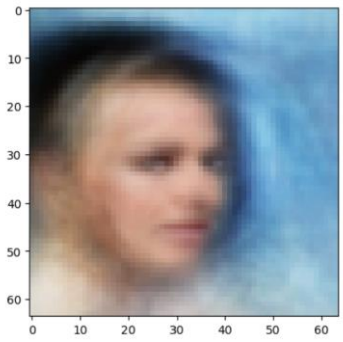
```
if model_type in ['fcn']:
    img = img.reshape(img.shape[0], -1)
    x = model.encoder(img)
    x[0][0] = x[0][0]/5
    output = model.decoder(x)
    output = output.reshape(3,64,64)
```



x[0][0] = x[0][0]*5

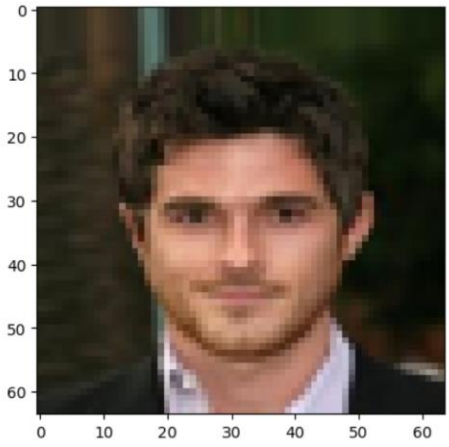


x[0][0] = x[0][0]/5

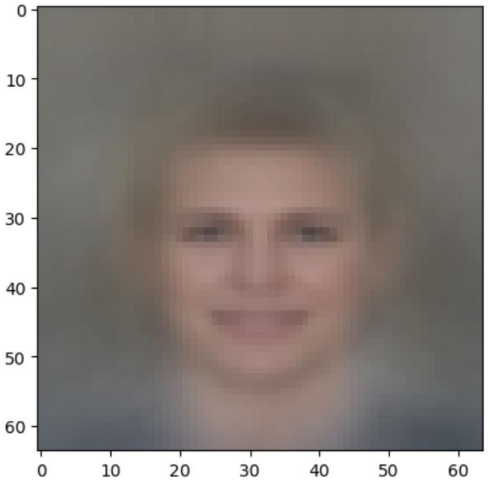


可以看出 x[0][0]*5 的嘴角有微笑的感覺，且鼻子等五官皆比較清楚，另一種則較為嚴肅，鼻子的地方比較模糊。

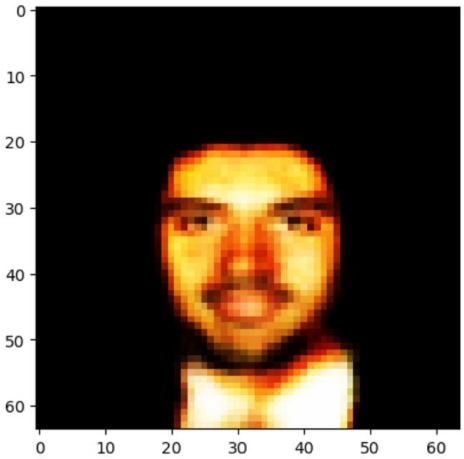
```
if model_type in ['fcn']:
    img = img.reshape(img.shape[0], -1)
    x = model.encoder(img)
    x[0] = x[0]*5
    output = model.decoder(x)
    output = output.reshape(3,64,64)
```



x[0] = x[0]/5



x[0] = x[0]*5



可以看出 x[0] = x[0]/5 的整體偏白，且背景幾乎不見了，但是五官比較清楚，膚色也較接近原圖，另一種則直接沒有背景，且整體偏暗，偏紅。