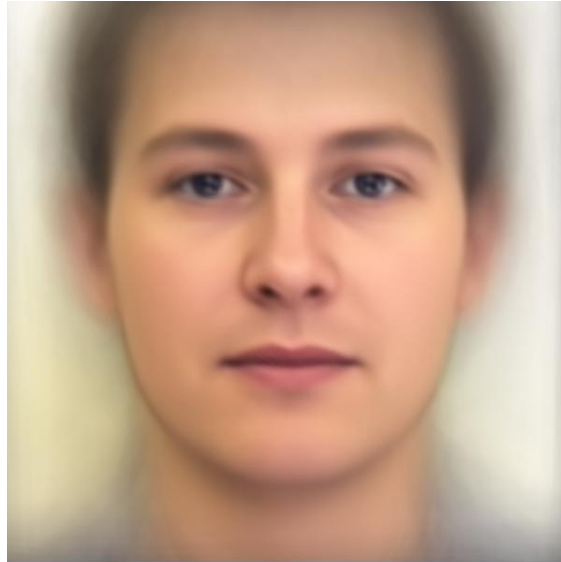


## Machine Learning HW7 Report

學號：B06902067 系級：資工二 姓名：許育銘

### 1. PCA of color faces:



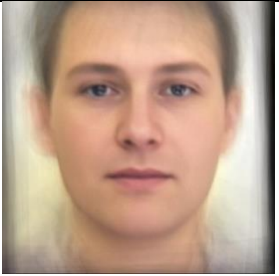


- a. 請畫出所有臉的平均。



- b. 請畫出前五個 Eigenfaces，也就是對應到前五大 Eigenvalues 的 Eigenvectors。



- c. 請從數據集中挑出任意五張圖片，並用前五大 **Eigenfaces** 進行 **reconstruction**，並畫出結果。

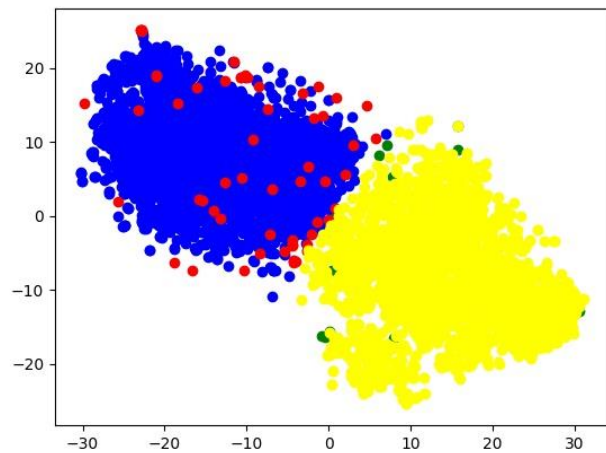
76_reconstruction	80_reconstruction	92_reconstruction
		
100_reconstruction	85_reconstruction	
		

- d. 請寫出前五大 **Eigenfaces** 各自所佔的比重，請用百分比表示並四捨五入到小數點後一位。

1	2	3	4	5
4.1%	2.9%	2.4%	2.2%	2.1%

## 2. Image clustering:

- a. 請實作兩種不同的方法，並比較其結果(reconstruction loss, accuracy)。(不同的降維方法或不同的 **cluster** 方法都可以算是不同的方法)  
我使用同一個 **model**，分別用 **PCA** 和 **t-SNE** 降到 2 維，再用 **kmeans** 分類。  
PCA accuracy: public score: 0.60838, private score: 0.60842  
t-SNE accuracy: public score: 0.96212, private score: 0.96240
- b. 預測 **visualization.npy** 中的 **label**，在二維平面上視覺化 **label** 的分佈。(用 **PCA**, **t-SNE** 等工具把你抽出來的 **feature** 投影到二維，或簡單的取前兩維 2 的 **feature**)  
其中 **visualization.npy** 中前 2500 個 **images** 來自 **dataset A**，後 2500 個 **images** 來自 **dataset B**，比較和自己預測的 **label** 之間有何不同。  
藍色代表預測 **A** 正確，綠色代表將 **A** 誤認成 **B**，黃色代表預測 **B** 正確，紅色代表將 **B** 誤認成 **A**。  
少數的資料在 **t-SNE** 後會跑到另一邊，不過大部分都分得很明顯。



- c. 請介紹你的 model 架構(encoder, decoder, loss function...)，並選出任意 32 張圖片，比較原圖片以及用 decoder reconstruct 的結果。

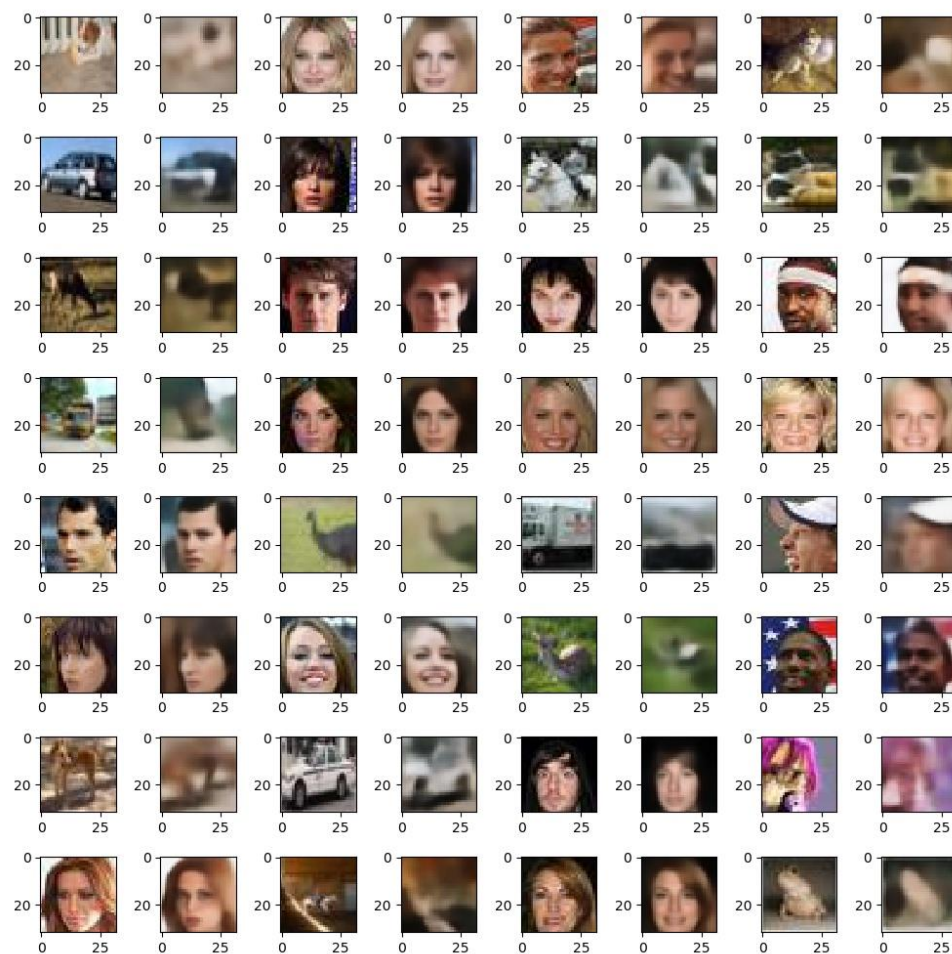
encoder:

input_1 (InputLayer)	(None, 32, 32, 3)	0
conv2d_2 (Conv2D)	(None, 32, 32, 64)	1792
max_pooling2d_1 (MaxPooling2)	(None, 16, 16, 64)	0
conv2d_3 (Conv2D)	(None, 16, 16, 128)	73856
conv2d_4 (Conv2D)	(None, 16, 16, 128)	147584
max_pooling2d_2 (MaxPooling2)	(None, 8, 8, 128)	0
reshape_1 (Reshape)	(None, 8192)	0
dense_1 (Dense)	(None, 128)	1048704

decoder

dense_2 (Dense)	(None, 8192)	1056768
reshape_2 (Reshape)	(None, 8, 8, 128)	0
up_sampling2d_1 (UpSampling2)	(None, 16, 16, 128)	0
conv2d_transpose_1 (Conv2DTr	(None, 16, 16, 128)	147584
conv2d_transpose_2 (Conv2DTr	(None, 16, 16, 128)	147584
up_sampling2d_2 (UpSampling2)	(None, 32, 32, 128)	0
conv2d_transpose_3 (Conv2DTr	(None, 32, 32, 64)	73792
conv2d_transpose_4 (Conv2DTr	(None, 32, 32, 64)	36928
conv2d_transpose_5 (Conv2DTr	(None, 32, 32, 3)	1731

loss function: mse  
comparison



左側為原圖，右側為 **reconstruct** 後的圖。