

# Machine Learning HW7 Report

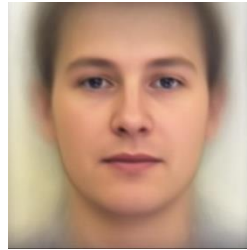
學號：r07942092

系級：電信碩一

姓名：白佳灝

## 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。(選擇第 1,30,50,70,90 張)

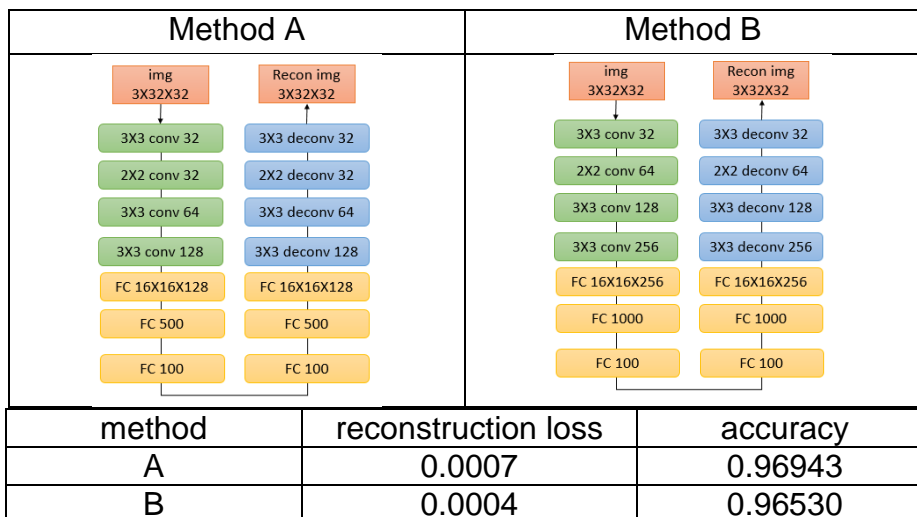


- d. 請寫出前五大 Eigenfaces 各自所佔的比重，請用百分比表示並四捨五入到小數點後一位。

1st	2nd	3rd	4th	5th
4.1%	2.9%	2.4%	2.2%	2.1%

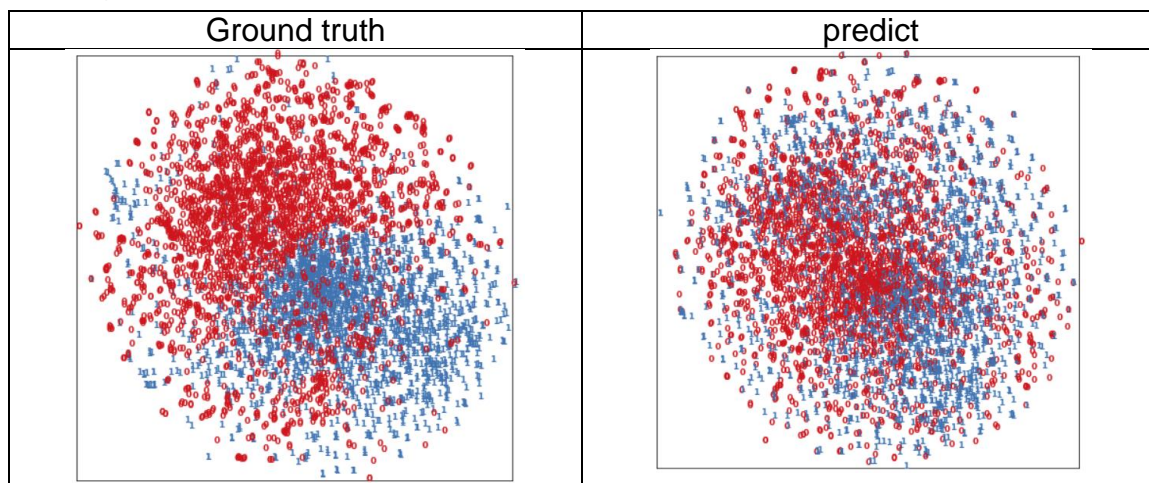
## 2. Image clustering:

- 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)



- 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)

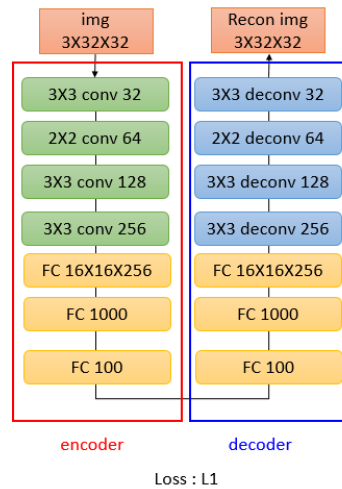
其中 visualization.npy 中前 2500 個 images 來自 dataset A，後 2500 個 images 來自 dataset B，比較和自己預測的 label 之間有何不同。





從圖中可以發現，我預測的圖並沒有將 dataset A 和 dataset B 的特徵分得很開，相對之下，可以發現助教給的答案所畫出的圖，不同 dataset 的特徵是有明顯分開的。

- 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。

model 架構



原圖	decoder reconstruct 的結果
	

Reconstruct 的圖與原圖挺像的，表示 auto encoder 學的應該不錯，但還是有點誤差，像是顏色深淺有點不依，且重建的圖有點模糊，但每張圖基本的特徵有抓到。