

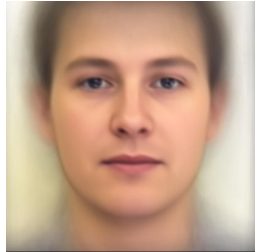
Machine Learning HW7 Report

學號：b06705058 系級：資管二

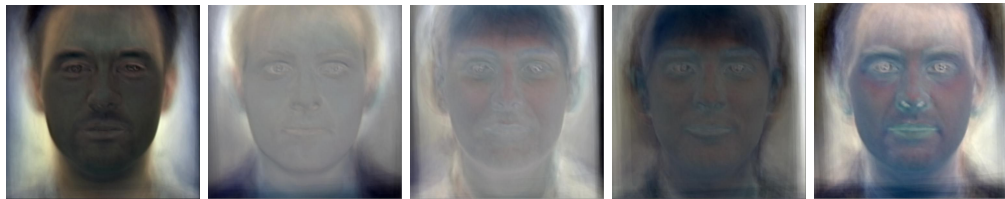
姓名：劉品桷

1. PCA of color faces:

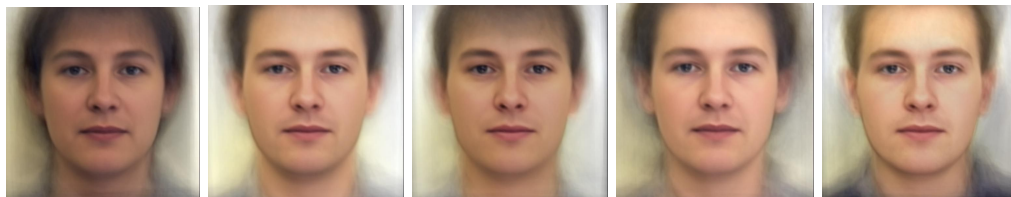
- a. 請畫出所有臉的平均。



- 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。



- 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。

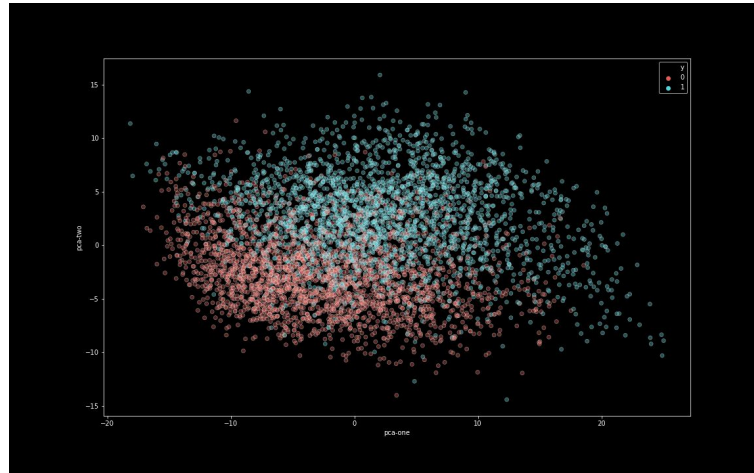
4.1、2.9、2.4、2.2、2.1

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)

測試過兩種降維方法, 一種是auto-encoder, 一種是pca, 單用 auto-encoder的表現不好, reconstruction loss使用MSE的話為0.023, Accuracy = 0.66, 如果使用pca的accuracy可以到0.961, 兩個一起使用時與單純使用pca效果差不多。

- 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。



預測的很不太準, 只有0.679的準確率。

- 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。

因為使用MSE作為loss function, adadelta作為optimizer, auto-encoder使用四層convolution layer, filter = 128(3, 3)、128(3, 3)、64(3, 3)、64(3, 3), 四次maxpooling(2, 2), decoder reconstruct之後圖片像是有打馬賽克。

