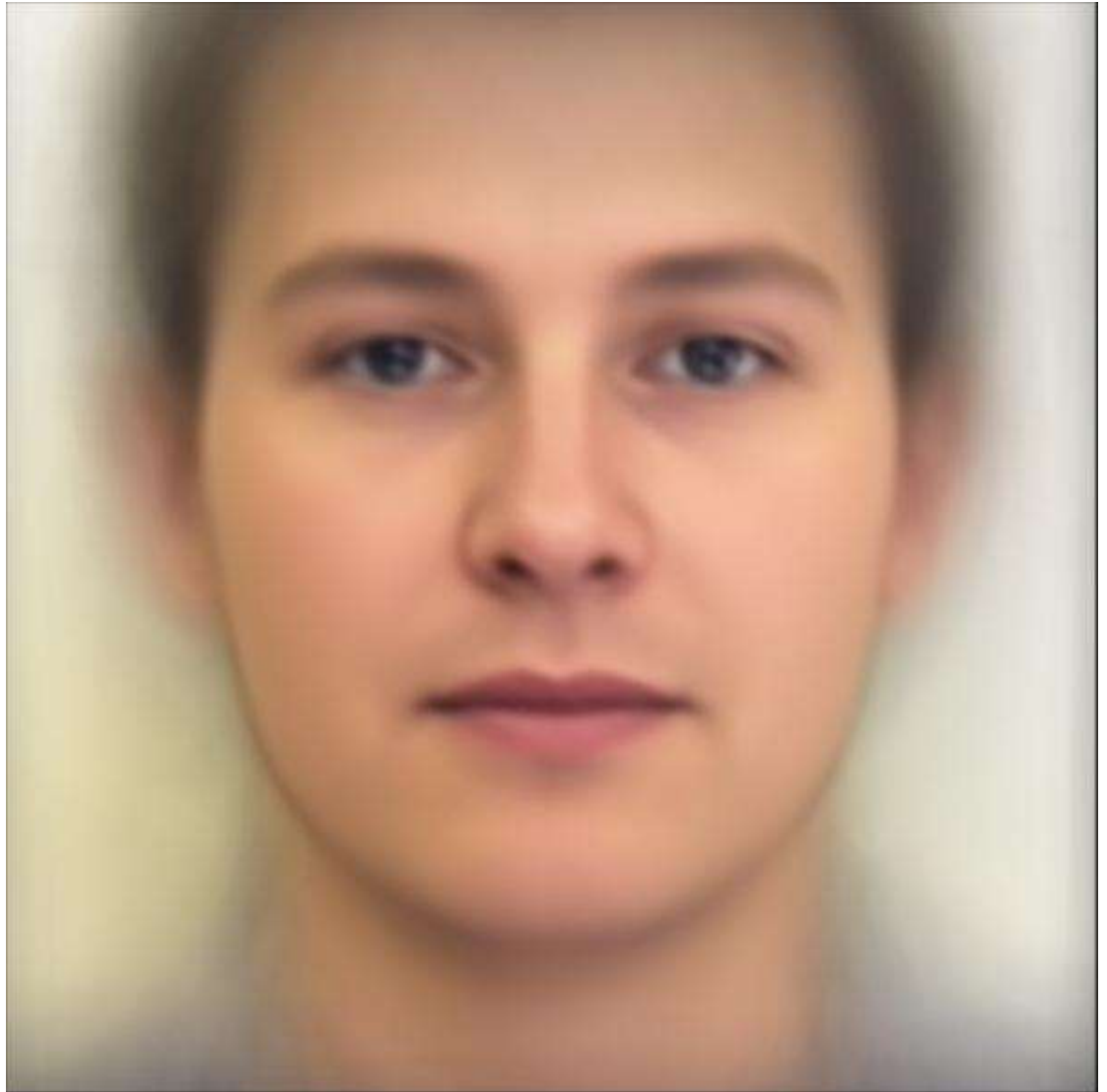


#### HW4

學號： b05901041 系級： 電機二 姓名： 蘇家軒

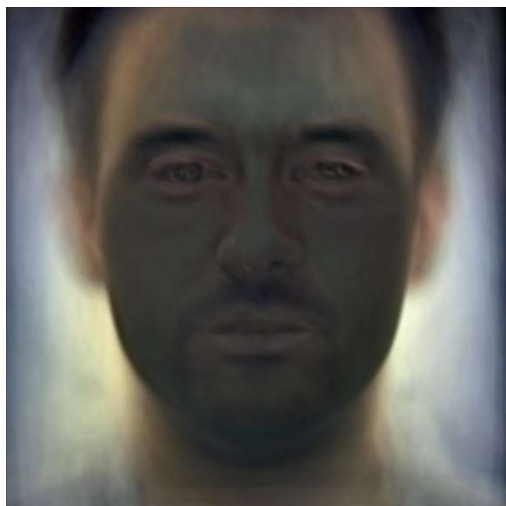
### A. PCA of colored faces

1. (.5%) 請畫出所有臉的平均。



2. (.5%) 請畫出前四個 Eigenfaces，也就是對應到前四大 Eigenvalues 的 Eigenvectors。

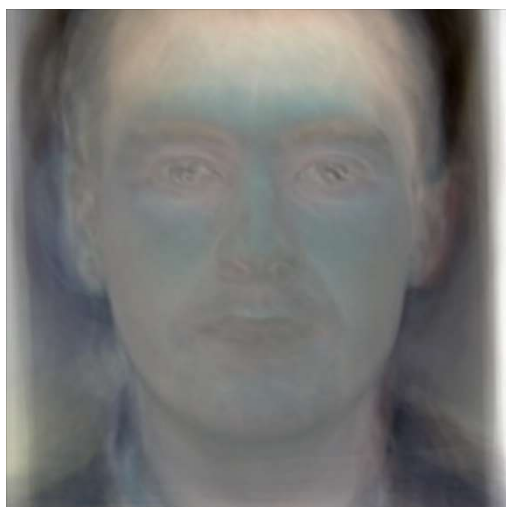
(1)



(2)



(3)

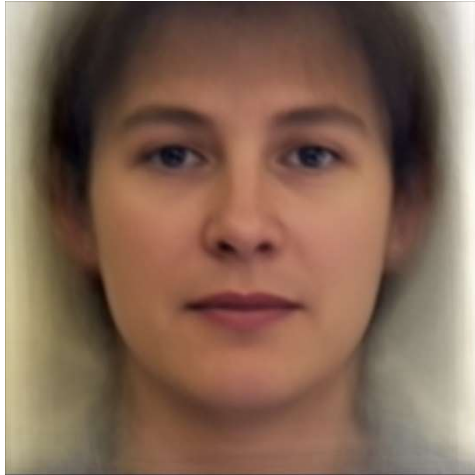


(4)

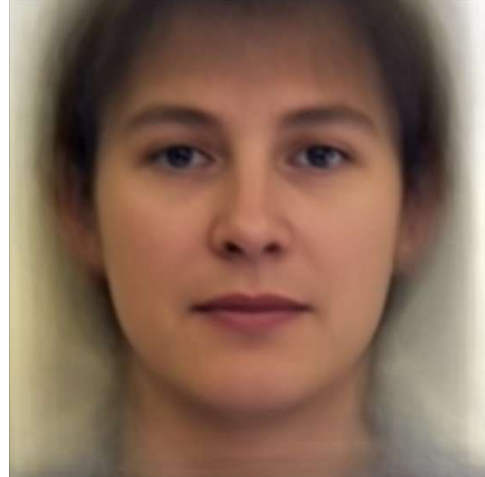


3. (.5%) 請從數據集中挑出任意四個圖片，並用前四大 Eigenfaces 進行 reconstruction，並畫出結果。

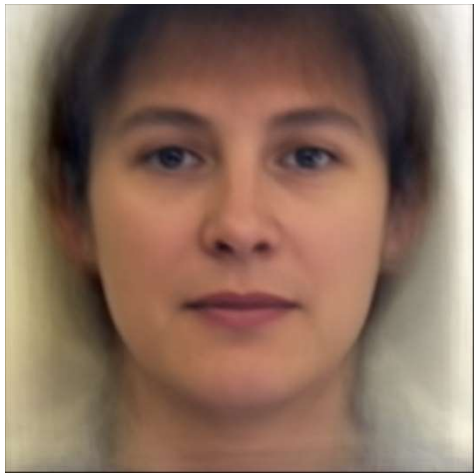
0.jpg



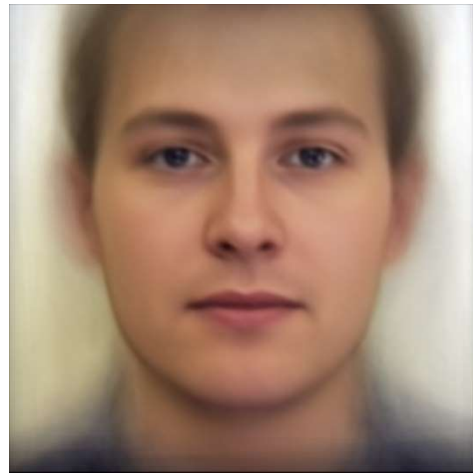
1.jpg



2.jpg



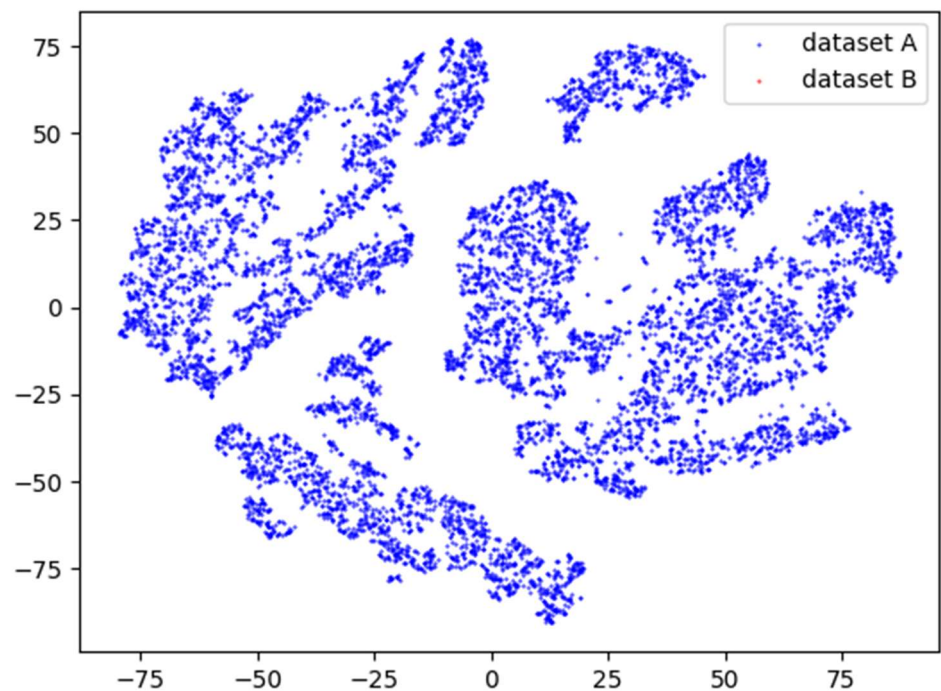
3.jpg



4. (.5%) 請寫出前四大 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。  
4.2%, 3.0%, 2.4%, 2.2%

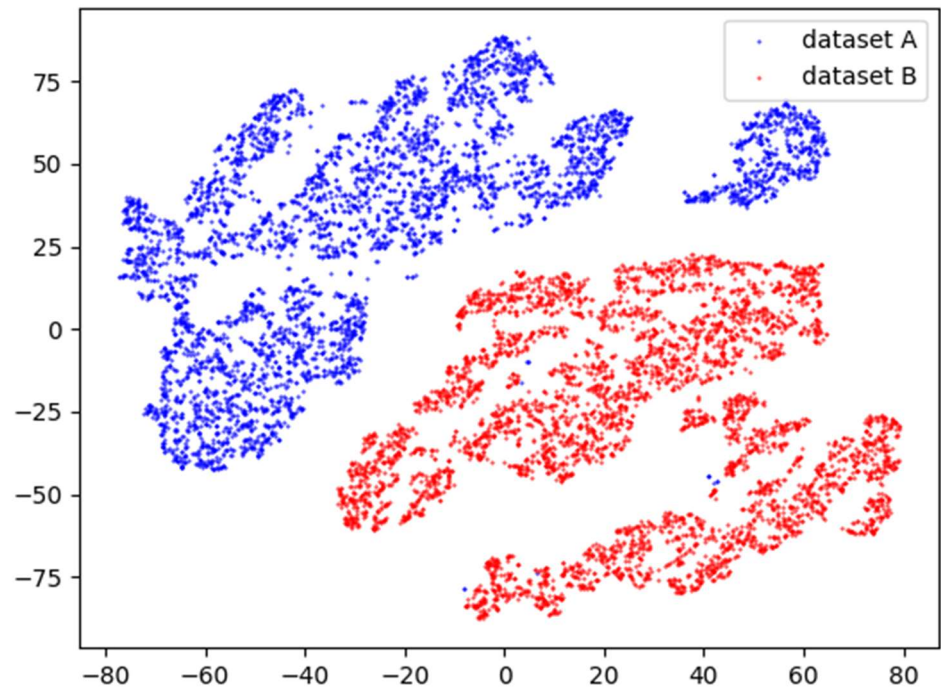
## B. Image clustering

1. (.5%) 請比較至少兩種不同的 feature extraction 及其結果。(不同的降維方法或不同的 cluster 方法都可以算是不同的方法)
  - (1) auto encoder:  
使用助教提供之 sample code，正確率約 99.3%
  - (2) pca + kmeans:  
使用 pca 降維並用 kmeans 做分類，正確率為 99.7%
2. (.5%) 預測 visualization.npy 中的 label，在二維平面上視覺化 label 的分佈。



3. (.5%) visualization.npy 中前 5000 個 images 跟後 5000 個 images 來自不同 dataset。請根據這個資訊，在二維平

面上視覺化 label 的分佈，接著比較和自己預測的 label 之間有何不同。



實際上我在做 label 時只有 309 個被分到 dataset b，造成上一題幾乎沒有見到紅色的分布。

## C. Ensemble learning

1. (1.5%) 請在 hw1/hw2/hw3 的 task 上擇一實作

ensemble learning，請比較其與未使用 ensemble method 的模型在 public/private score 的表現並詳細說明你實作的方法。（所有跟 ensemble learning 有關的方法都可以，不需要像 hw3 的要求硬塞到同一個 model 中）

我以 neuron network 和 generative model 實作 hw2 的 ensemble model, 以 neuron network 之結果為主, 但若 generative model 之值 $>0.2$  或 $<0.2$  時則以 generative 之結果為主。原先 generative model 正確率大概在 84%, neuron network 則在 85.5%, ensemble model 之正確率卻來到 76%, 檢查後發現絕大多數之 prediction 皆為 0, 因此修改策略, 若 neuron network 判定為 1 之資料即直接輸出 1, 若 generative model 之值超過 0.8 也判定為 1, 正確率來到了 85.7%, 略有進步。