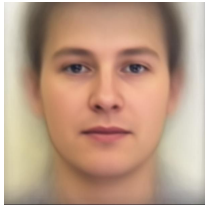


HW4

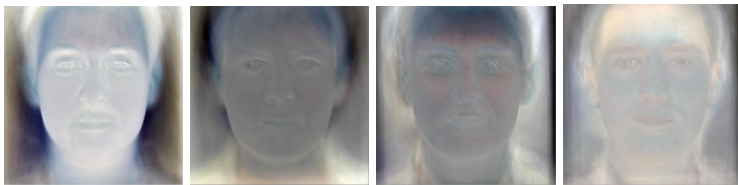
學號：r06725028 系級：資管碩一 姓名：黃于真

A. PCA of colored faces

A.1. (.5%) 請畫出所有臉的平均。

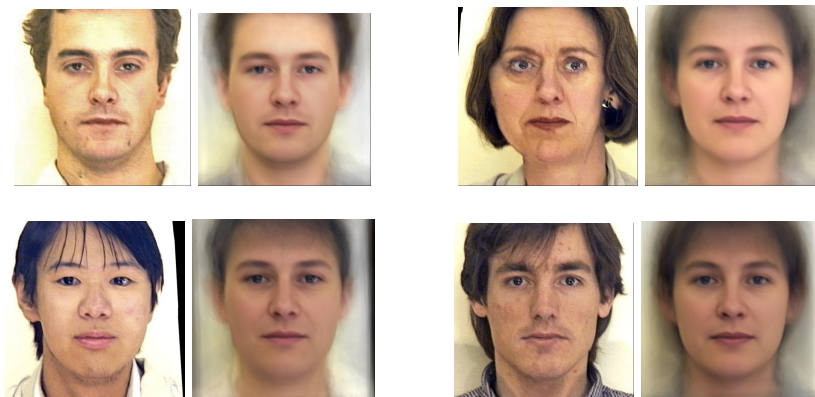


A.2. (.5%) 請畫出前四個 Eigenfaces，也就是對應到前四大 Eigenvalues 的 Eigenvectors。



(由左至右排序)

A.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。



(左原圖，右重建)

(.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。

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

B. Image clustering

(Collaborators: r06725053 韓宏光)

B.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)

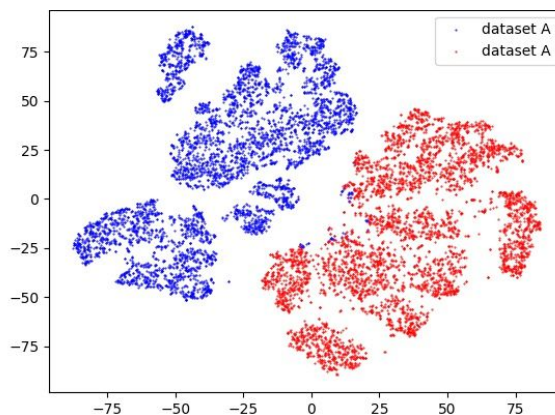
(reference:https://github.com/thtang/ML2017FALL/blob/master/hw6/image_cluster.py)

第一種方法我使用sample code的降維方法，經過測試epochs為120時有最好結果，第二種方法是sklearn中的pca套件，附圖是pca的參數設定，至於n_com部分，經過測試，400時會有最好的結果，如下表格，可以看到使用pca降維的效果較好，可以大幅提升分數到接近1，至於sample code的降維方法，如果epochs設太大，雖然val_loss可以降到很低，但可能學到的encoder太複雜，反而降維效果變差，也使得分數下降到0.97多而已。

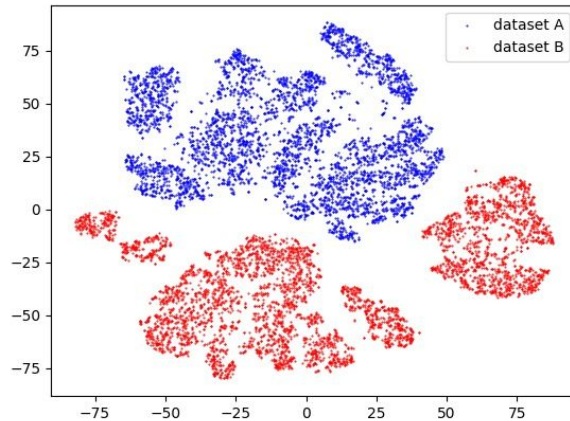
```
PCA(n_components=n_com, copy=True, whiten=True, svd_solver='full', tol=0.0, iterated_power='auto', random_state=0)
```

	pca360	pca400	pca450	sample150	sample80	sample120
public	0.99996	1.00000	0.99998	0.97266	0.98393	0.99226
private	0.99997	1.00000	0.99998	0.97250	0.98395	0.99221

B.2. (.5%) 預測 visualization.npy 中的 label，在二維平面上視覺化 label 的分佈。



B.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。



比較兩圖會發現似乎有點角度轉向的差別，但整體形狀是差不多的，而第二小題的圖是用sample code的降維方法，在kaggle上的private分數是0.99221，所以可以看到有些混淆的地方，有些藍色點混在紅色群的邊邊位置，但第三題的圖則基本上紅色和藍色點都是各自分開的。

C. Ensemble learning

C.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）

(reference:<https://github.com/WindQAQ/ML2017/blob/master/hw3/test.py>)

我選擇的是作業三來做ensemble learning，我選擇當時訓練出來的三個val_acc在0.7以上的模型，把這三個模型分別預測出的機率直接相加起來之後再取argmax，結果如下表格，可以看到做ensemble learning之後的分數比原本三個model各自結果都還高，可見使用三個model來預測可以起到互補的作用。

model	model_1	model_2	model_3	ensemble model
public	0.69322	0.69490	0.69713	0.70716
private	0.68626	0.69824	0.69239	0.69991