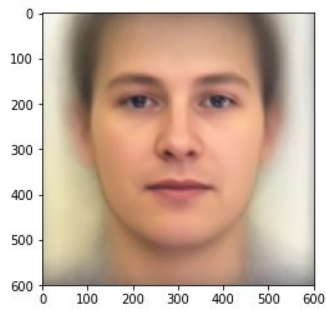


# HW4

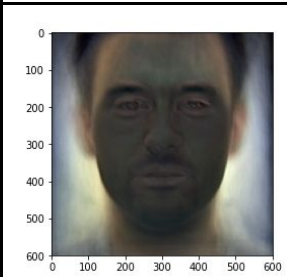
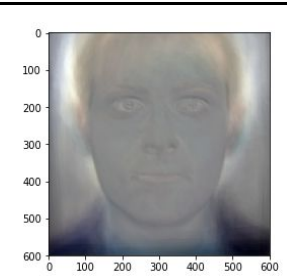
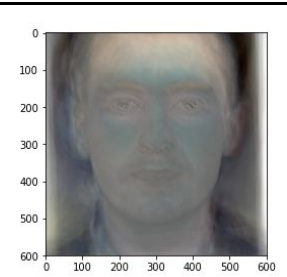
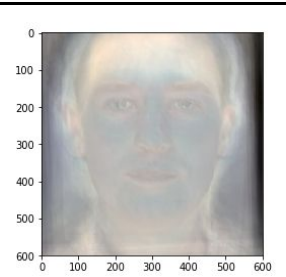
學號：b04705026 系級：資管三 姓名：林彥廷

## A. PCA of colored faces


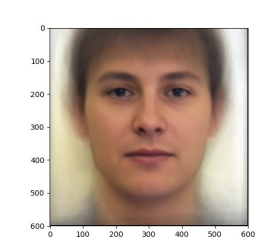

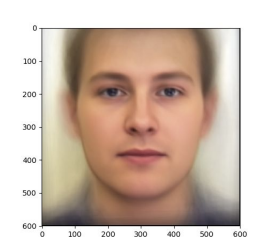
### A.1. mean face



### A.2. top-4 eigen-faces

第一大	第二大	第三大	第四大
			

### A.3. recontructed faces with 4 comp.

ID #70	ID #332	ID #357	ID #58
			

### A.4. Table for explained variance

第一大	第二大	第三大	第四大
4.1%	2.9%	2.4%	2.2%

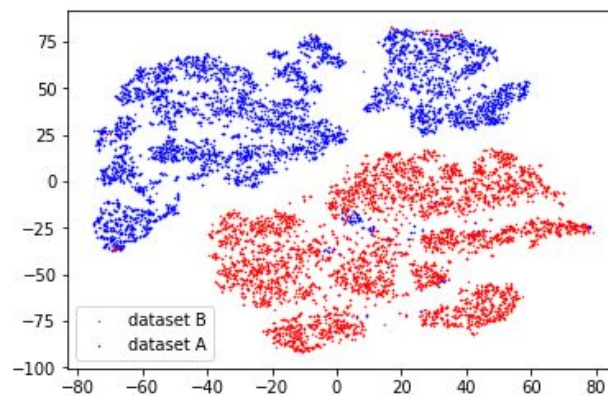
## B. Image clustering

### B.1.

- Feature Extraction方法
  - PCA(250 compn.) vs Autoencoder(64 compn.)  
(256-128-64-128-256)
  - Clustering方法皆是Kmeans (2 compn.)
- 討論
  - Autoencoder可以使用非線性的把特徵抽出，相較於PCA可以考慮一些複雜的關係，使用高壓縮比，找出好的特徵。
  - PCA單純找出單範正交的特徵組合，對於非線性的關係無法掌握。對圖片有空間性質的關係表現並不好。
- 模型表現

Dim. Reduction	Autoencoder	PCA
public score	.99074	.51690
private score	.99082	.51767

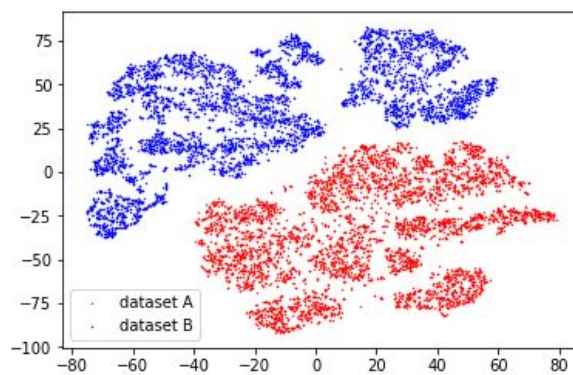
### B.2.



從我的模型預測label再由TSNE投影到2D空間。

結果有些資料並沒有分的很開（中下）

### B.3.

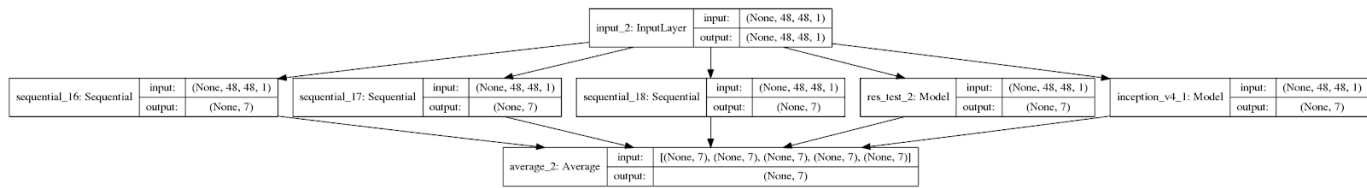


我使用兩層的autoencoder進行降維到64維，再透過TSNE轉換到2D空間來畫圖。以這個投影的圖來說，圖中下的地方，我的模型並沒有把兩個資料集分的很開，可能是因為高維空間的feature並沒有抓好，也有可能需要更複雜的關係或轉換。

## C. Ensemble learning

### C.1. 作業三 Ensemble Learning

- Ensemble模型架構與方法



- 5個小model皆為不同概念的CNN(VGG, ResNet, Inception)為了讓每個模型都有強項，並用”VOTING”的方式決定最終預測結果。
- “Ensemble Model”最後使用 ADD層，把5個模型softmax轉換出的機率相加，挑選最大機率的當成最終結果。
- 子模型分開訓練，因為要用不同的optimizer和其他參數。

- 子模型架構

- 前3個為相同架構：類似Vgg，但深度不同。
- 第四個，概念來自residual net，差別在於減少模型深度，用了12個residual module。
- 第五個，概念來自inception v4，差別在於在減少模型深度，網路最後加上Dense的分類層。
- 每個模型每層後都加上Batchnormalization，已經在模型尾端的分類層加上Dropout(0.5)。
- 第四，五個模型activation func=relu
- 每個模型接使用Data augmentation。

- 模型表現

	ensemble model	single model
public score	0.72137	0.68514
private score	0.70855	0.68180