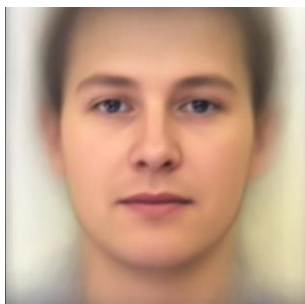


HW4

學號：B04505026 系級：工海三 姓名：蔡仲閔

A. PCA of colored faces

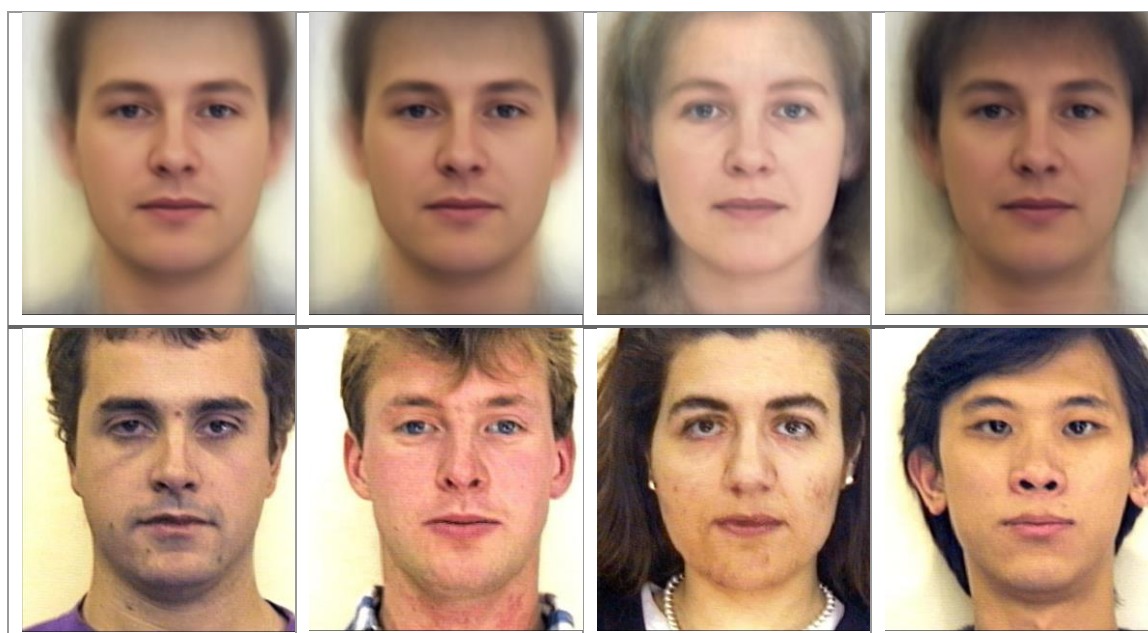
A.1. (.5%) 請畫出所有臉的平均。



A.2. (.5%) 請畫出前四個 Eigenfaces



A.3. (.5%) 用前四大 Eigenfaces 進行 reconstruction，並畫出結果。



A.4. (.5%) 請寫出前四大 Eigenfaces 各自所佔的比重

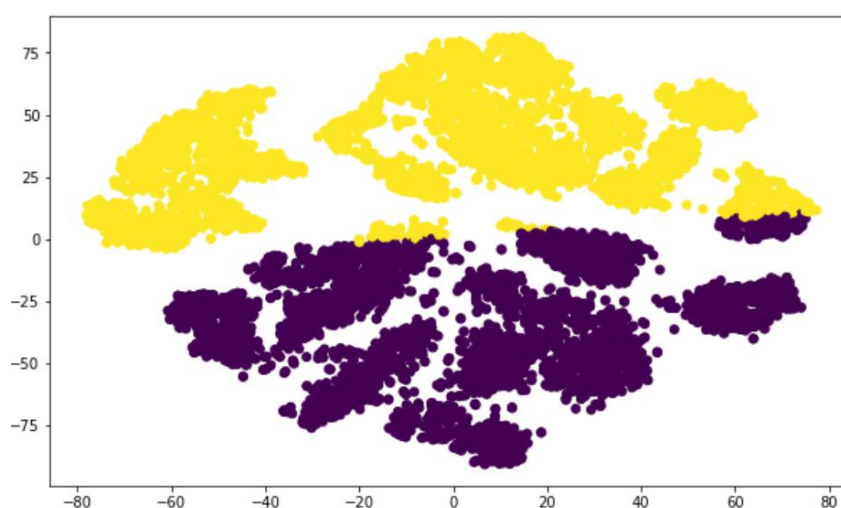
4.1%	2.9%	2.4%	2.2%
------	------	------	------

B. Image clustering

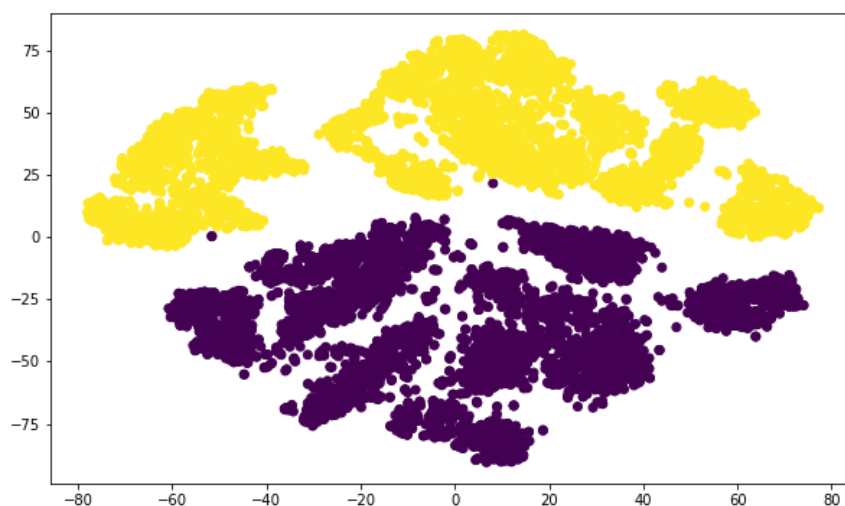
- A.5. (.5%) 請比較至少兩種不同的 **feature extraction** 及其結果。在這邊我們分別使用 **Autoencoder** 降至 32 維以及 **PCA** 來進行，並使用 **K-means** 進行分類，可以看到兩種方法的在最後的表現上都相當優異，**PCA** 達到 100% 的正確率，兩者花費時間則約略相同。

	Time Cost	Private Score	Public Score
Autoencoder	121.99 sec	0.98	0.98
PCA	130.94 sec	1.00	1.00

- A.6. (.5%) 預測 **visualization.npy** 中的 **label**，在平面上視覺化 **label** 的分佈。



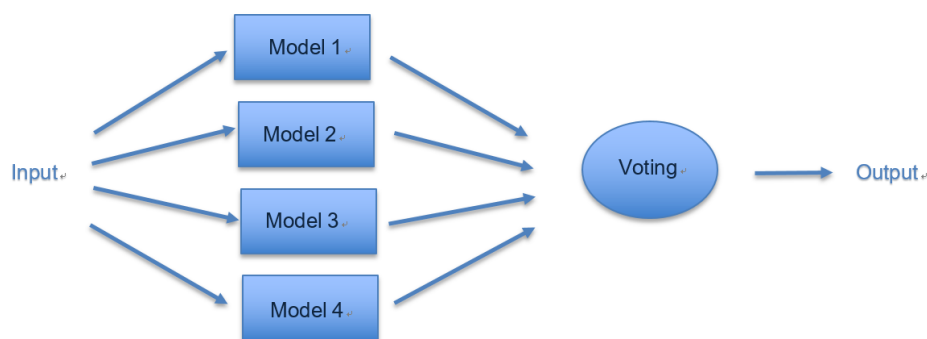
- A.7. (.5%) 請根據這個資訊，在二維平面上視覺化 **label** 的分佈，接著比較和自己預測的 **label** 之間有何不同。



在預測 **visualization.npy** 的過程中，我選擇用 **t-sne** 來進行降維，並直接降到二維平面，在圖中看到牠將兩組資料有效的分開，比較 **true label** 作圖，算是相當精確，僅有少部分點出現明顯誤差。

B. Ensemble learning

B.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）



Model	Loss	Private Score	Public Score
Model 4	0.34	0.64	0.67
Model 3	0.39	0.58	0.58
Model 2	0.36	0.62	0.63
Model 1	0.37	0.63	0.64
Ensemble Model	0.31	0.68	0.68

在這次的實驗中，我選擇 hw3 的 CNN model 來做 ensemble，我參考 (<https://towardsdatascience.com/ensembling-convnets-using-keras-237d429157eb>) 的做法，並加入不同的 model，也實作了 Image generator，從上表可以看到，上面四個 model 的表現都並不起眼，在 Private Score 上最高僅有 0.64，然後透過 ensemble 結合四個 model 後，這邊使用 keras 的 maximum layer 選出最高分的 class，最後得到相當好的成果，甚至超過 hw3 時使用的架構極大的 model，證實 ensemble 相當有效。