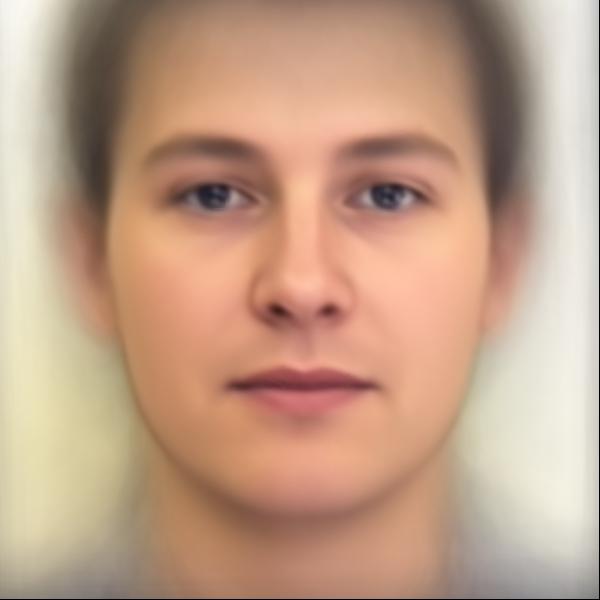
HW4

學號：B04505026 系級：工海三 姓名：蔡仲閔

1. PCA of colored faces
   1. (.5%) 請畫出所有臉的平均。



* 1. (.5%) 請畫出前四個 Eigenfaces

|  |  |  |  |
| --- | --- | --- | --- |
| Top 1C:\Users\qa276\Downloads\eigen_face_0.jpg | Top 2C:\Users\qa276\Downloads\eigen_face_1.jpg | Top 3C:\Users\qa276\Downloads\eigen_face_2.jpg | Top 4C:\Users\qa276\Downloads\eigen_face_3.jpg |

* 1. (.5%)用前四大 Eigenfaces 進行 reconstruction，並畫出結果。

|  |  |  |  |
| --- | --- | --- | --- |
| C:\Users\qa276\Downloads\reconstruction_11.jpg | C:\Users\qa276\Downloads\reconstruction_22.jpg | C:\Users\qa276\Downloads\reconstruction_29.jpg | C:\Users\qa276\Downloads\reconstruction_146.jpg |
| C:\Users\qa276\AppData\Local\Microsoft\Windows\INetCache\Content.Word\11.jpg | C:\Users\qa276\Downloads\Aberdeen\22.jpg | C:\Users\qa276\Downloads\Aberdeen\29.jpg | C:\Users\qa276\Downloads\Aberdeen\146.jpg |

* 1. (.5%) 請寫出前四大 Eigenfaces各自所佔的比重

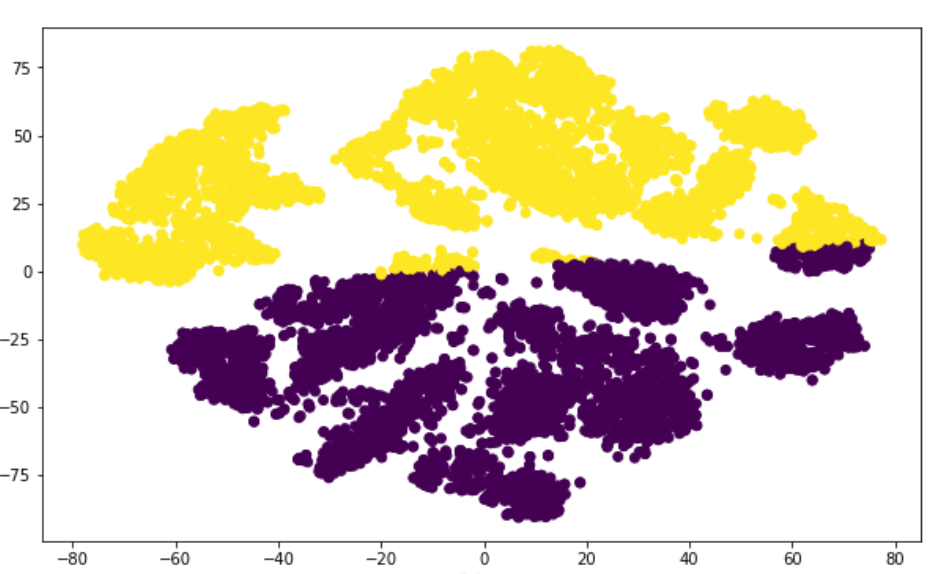
|  |  |  |  |
| --- | --- | --- | --- |
| 21.4% | 2.6% | 2.3% | 1.8% |

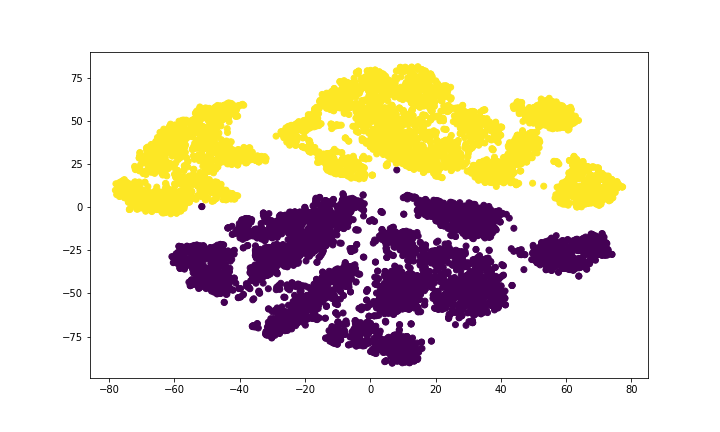
B. Image clustering

* 1. (.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 |

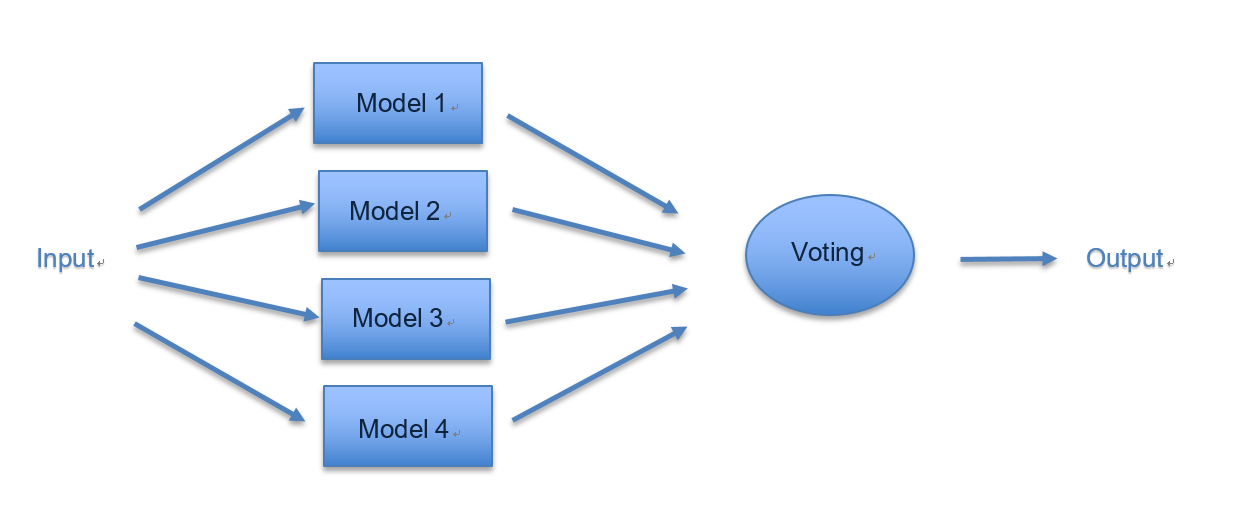
* 1. (.5%) 預測 visualization.npy 中的 label，在平面上視覺化 label 的分佈。



* 1. (.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，在圖中看到他將兩組資料有效的分開，比較ture label作圖，算是相當精確，僅有少部分點出現明顯誤差。

1. 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）



|  |  |  |  |
| --- | --- | --- | --- |
| 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。