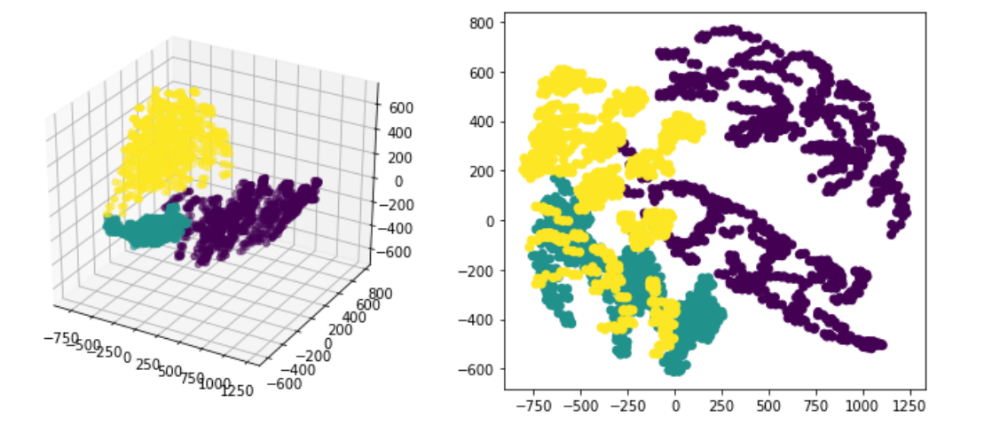
|  |
| --- |
| 訓練資料 |

* 這次使用的是總數3000張的人臉資料，分為Class1 Class2 Class3，我把其中2000切割為訓練集，剩下1000張作為測試集。



* 之後進行資料前處理，使用的演算法是PCA降維



如圖，左圖是取3個主成分，右圖只取2個主成分，可以發現經過PCA後只有3維可以完美切割，2維具有重疊區域，推測訓練的結果會有天花板。

|  |
| --- |
| 模型結構 |

* Two Layer
  + Hidden Size 10
  + Iteration 5000
  + Learning Rate 0.01
* Three Layer With Sigmoid
  + Hidden Size 10
  + Iteration 5000
  + Learning Rate 0.01
* Three Layer With Relu
  + Hidden Size 10
  + Iteration 5000
  + Learning Rate 0.01

訓練的條件盡量設定相同，以利比較不同模型的差距。

|  |
| --- |
| 訓練結果 |

* Accuracy

|  |  |  |  |
| --- | --- | --- | --- |
|  | Two Layer | Three Layer(Sigmoid) | Three Layer(Relu) |
| ACC | 0.818 | 0.853 | 0.881 |

這個表示從testset 測試出的結果，可以知道層數越多，Acc應該會越好，而activation 則是以Relu 為佳。

* Learning Curve & Decision Boundary

由學習曲線可以看出，兩層結構需要較久時間才能收斂，而三層可以迅速的收斂，

其中又以Relu 學習速度最快。另一方面，從decision boundary 來看，三者其實差不多。

