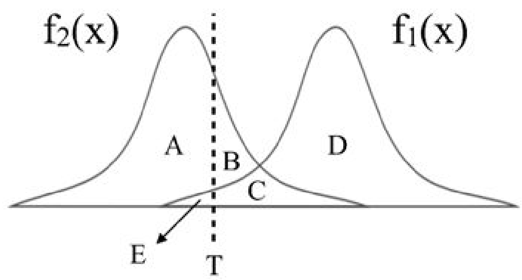
Computer Vision: from Recognition to Geometry
HW2

Name: 張邵瑀　Department:電機碩二　Student ID: r06921081

Problem 1

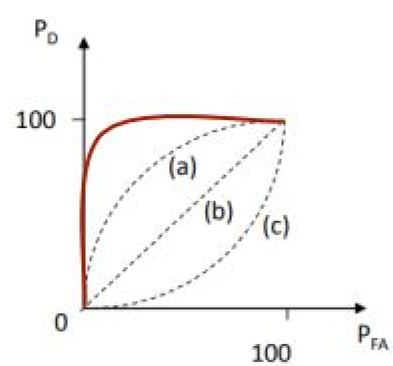
(a) Assume X is a continuous random variable that denotes the estimated probability of a binary classifier. The instance is classified as positive if $X > T$ and negative otherwise. When the instance is positive, X follows a PDF $f_1(x)$. When the instance is negative, X follows a PDF $f_2(x)$. Please specify which regions (A ~ E) represent the cases of *False Positive* and *False Negative*, respectively. Clearly explain why.



B: False Positive for $f_2(x)$ 因為X大於T，但其實正確分佈應該為negative，但判成為Positive
C: False Positive for $f_2(x)$ 因為X大於T，但其實正確分佈應該為negative，但判成為Positive
E: False Negative for $f_1(x)$ 因為X小於T，但其實正確分佈應該為Positive，但判成為negative

(b) There are three ROC curves in the plot below. Please specify which ROC curves are considered to have reasonable discriminating ability, and which are not. Also, please answer that under what circumstances will the ROC curve fall on curve (b)?

Ans: (a)是最合理的ROC curves，因為若是兩個PDF可以分的越開則會愈接近紅線的部份，而完全重合時就是(b)那條線，(c)則不可能發生，因為(b)已經是最糟情況了。

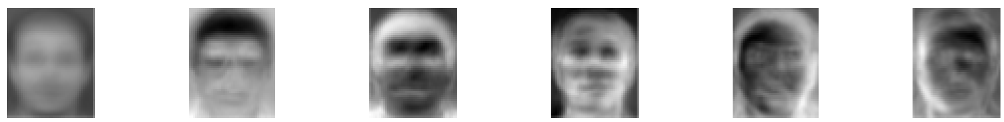


Problem 2

(a) PCA

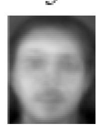
In this task, you need to implement PCA from scratch, which means you cannot call PCA function directly from existing packages.

1. Perform PCA on the training data. Plot the mean face and the first five eigenfaces and show them in the report.

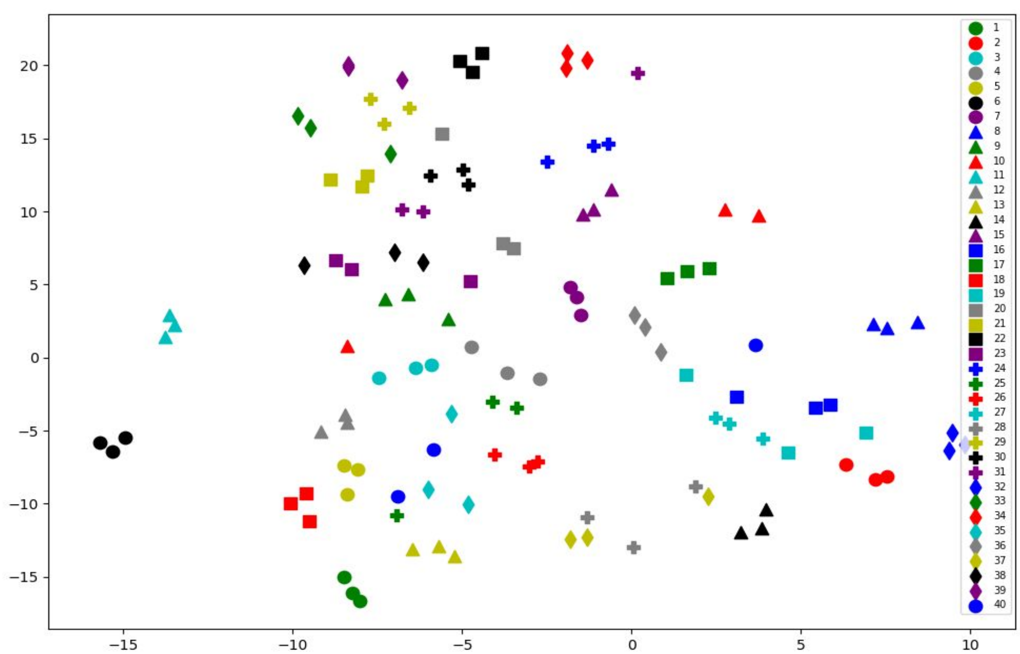


mean face | eigenfaces 1 | eigenfaces 2 | eigenfaces 3 | eigenfaces 4 | eigenfaces 5

2. Take $person_8\_image_6$, and project it onto the above PCA eigenspace. Reconstruct this image using the first n = { 5, 50, 150, all } eigenfaces. For each n, compute the mean square error (MSE) between the reconstructed face image and the original $person_8\_image_6$. Plot these reconstructed images with the corresponding MSE values in the report.

n	5	50	150	all
Reconstructed Image				
MSE	692.70	119.20	40.39	1.24e-26

3. Reduce the dimension of the image in testing set to dim = 100. Use t-SNE to visualize the distribution of test images.

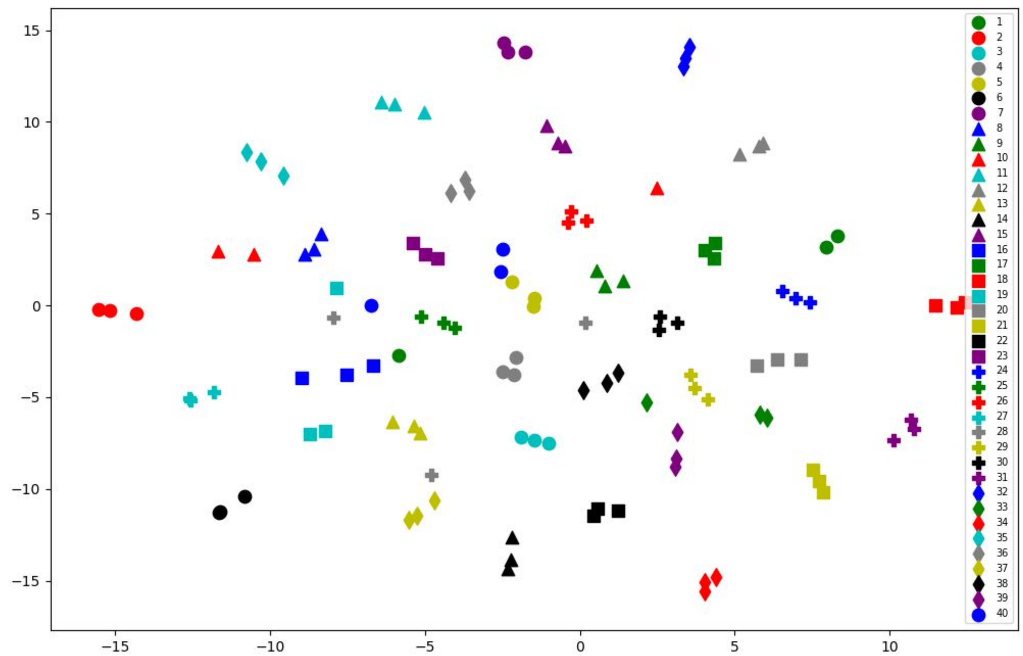


(b) LDA

In this task, you need to implement LDA from scratch, which means you cannot call LDA function directly from existing packages.

1. Implement LDA and plot first 5 Fisherfaces.

2. Use t-SNE to visualize the distribution of the projected testing data, which has the dimension of 30.



(c) To apply the k-nearest neighbors (k-NN) classifier to recognize the testing set images, please determine the best k and n values by 3-fold cross-validation.

For simplicity, the choices for such hyper-parameters are:

k = {1, 3, 5} and n = {3, 10, 39}.

Please show the cross-validation results and explain your choice for (k, n). Also, show the recognition rate on the testing set using your hyper-parameter choice. Please apply the above comparing method on both PCA and LDA.

PCA

k \ n	3	10	39
1	0.5416	0.7500	0.8416
test	0.5833	0.9416	0.9583
3	0.4749	0.6083	0.7000
test	0.5333	0.8500	0.9333
5	0.3750	0.5250	0.5833
test	0.4833	0.7750	0.9083

LDA

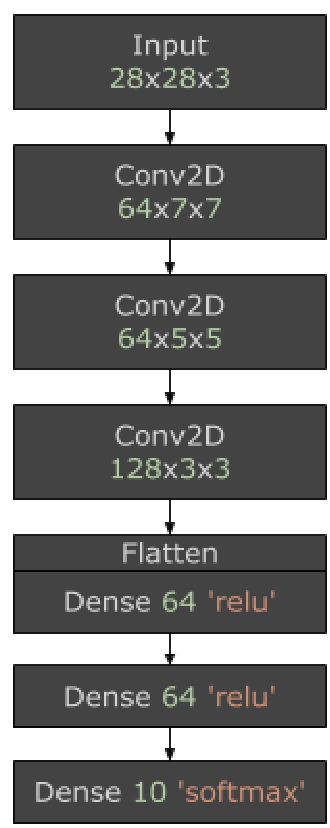
k \ n	3	10	39
1	0.3499	0.8416	0.9083
test	0.3833	0.8083	0.9166
3	0.2500	0.7250	0.8500
test	0.3916	0.8166	0.9166
5	0.2500	0.6583	0.7583
test	0.3916	0.8166	0.9166

Do you observe an improved recognition rate using fisherfaces (compared to eigenfaces obtained by PCA)? If so (or if not), what might be the possible explanation?

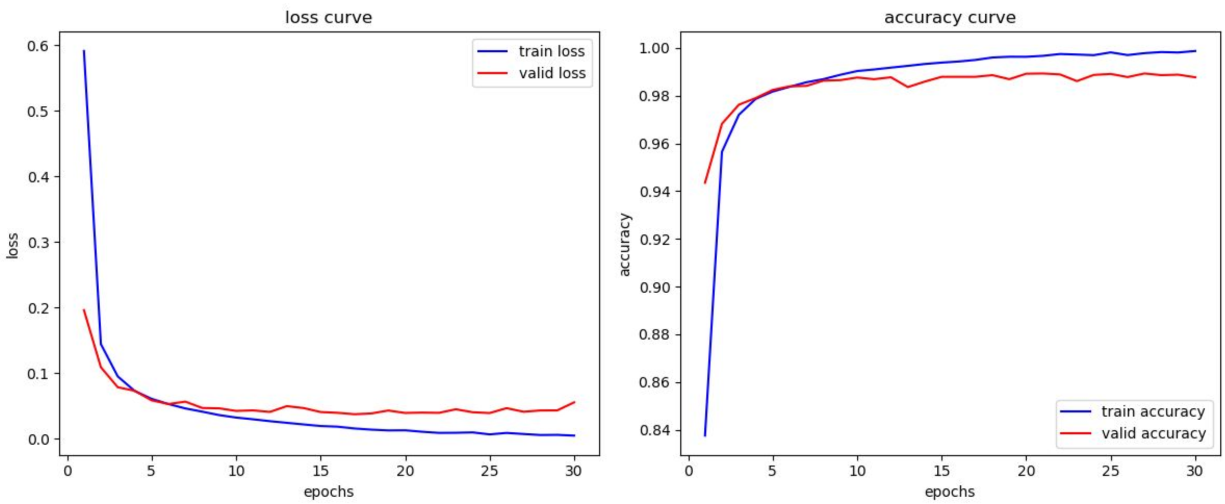
使用fisherfaces沒有比較準確，是因為LDA是對PCA取出的eigenvector做降維，但在training data中重要的特徵有可能並不是真正重要的特徵，其中重要的資訊可能分佈在testing data當中，所以LDA可能會產生對於training data的特徵依賴進而產生overfitting，由上表也可以很明顯的看出結果LDA，中在n較大的時候training的結果都比PCA好，但testing結果卻恰恰相反，而在低維度的時候甚至因為資訊不足而有非常嚴重的underfitting，間接證實了這個論點。

Problem 3

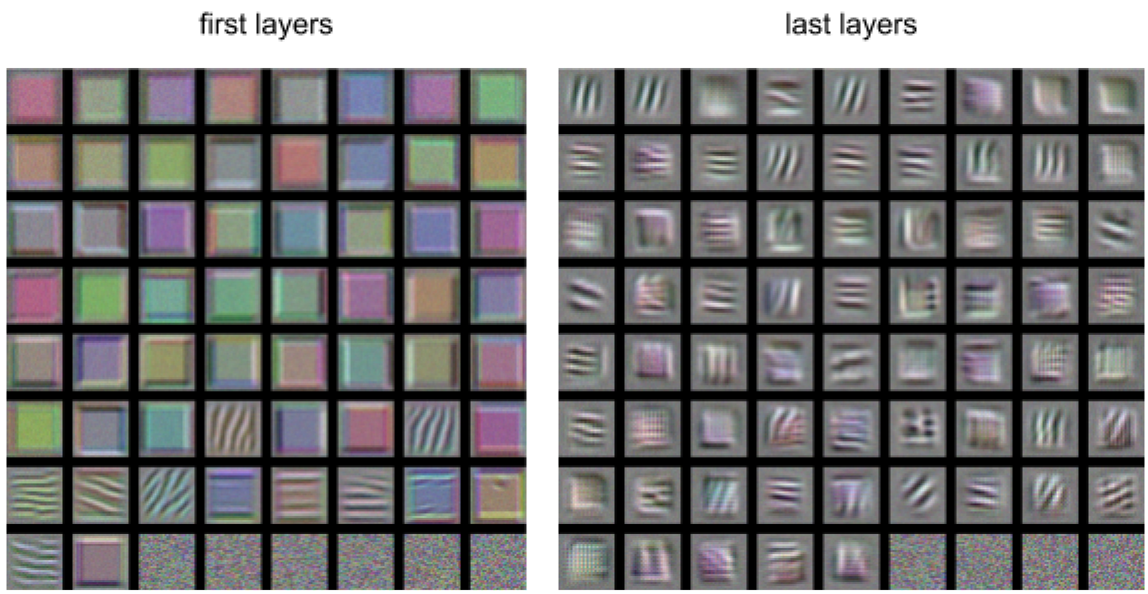
(a) Build a CNN model and train it on the given dataset. Show the architecture of your model in the report.
(b) Report your training / validation accuracy, and plot the learning curve (loss, accuracy) of the training process.



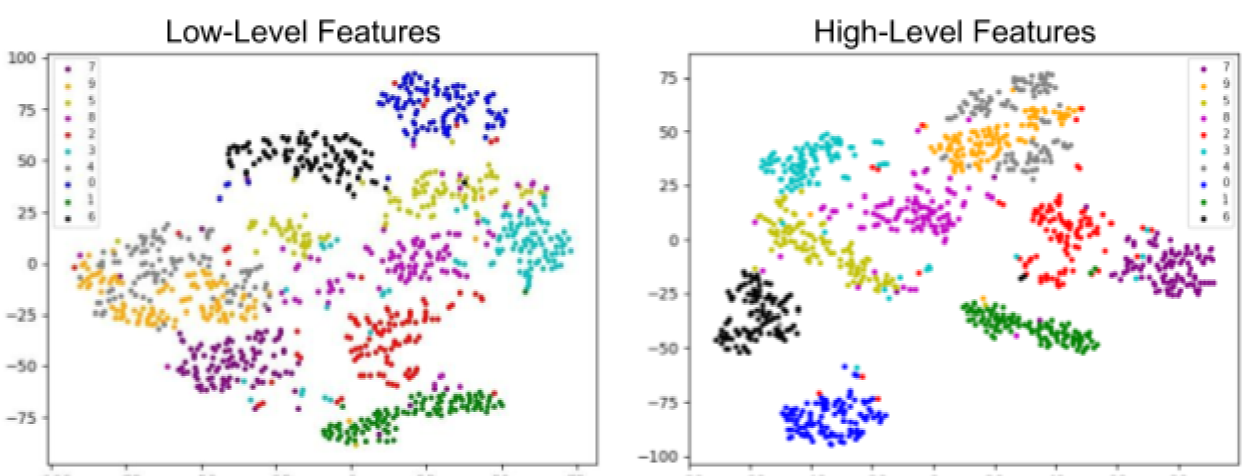
	training	validation
accuracy	0.9987	0.9877
loss	0.0051	0.0558



(c) Visualize at least 6 filters on both the first and the last convolutional layers.



(d) Visualize high-level and low-level features of 1000 validation data (100 for each class) extracted from different layers, and explain what you have observed from the two t-SNE plots.



可以看出來在洋紅色，天青色，草綠色分別為8，3，5在手寫上此3個數字的確是形狀最相近的，所以他們不管在high或是low都靠的較近，甚至在Low的時候三者有互相混雜的情況，而5跟6在兩者都很接近這也很合理其兩者也差在開口，而最難分辨的應該就是黃色(9)灰色(4)兩者在位置上市重疊在同一個區塊的，但在High feature的時候區塊還是分界明顯，但在Low feature的時候兩者很混雜，但在這個問題上還是做的不錯。