Computer Vision HW2 Report

B04507009 電機四 何吉瑞

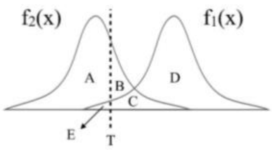
Problem 1

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.

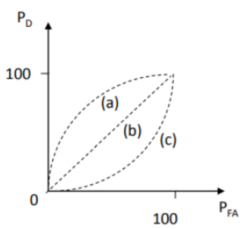


False Positive: B and C. Instances in these two regions are classified as positive. However, they actually follows the distribution f2, which means false positive occurs.
False Negative: E. Instances in this region are classified as positive. However, they actually follows the distribution f1, which means false negative occurs.

Problem 1(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)?



(a) and (b). Reasonably, the probability of true positive is higher or equal than that of false positive. When the Distribution of positive and negative are identical, ROC (b) occurs. However, the ROC like (c) will never occur in reasonable cases.

Problem2

(a)

Mean face:

First five eigenfaces:

From left to right, the first to the fifth.

Reconstruction:

n=5

MSE: 695.5818128077133

n=50

MSE: 119.11436096683154

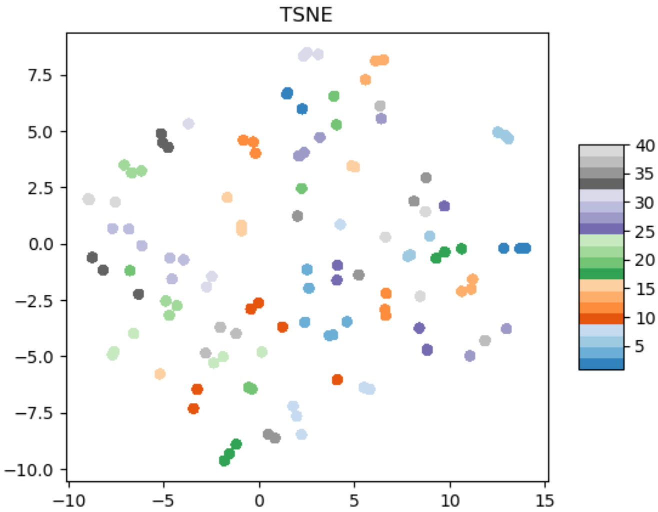
n=150

MSE: 40.40392373321021

n=all(279)

MSE: 0.002851286463538221

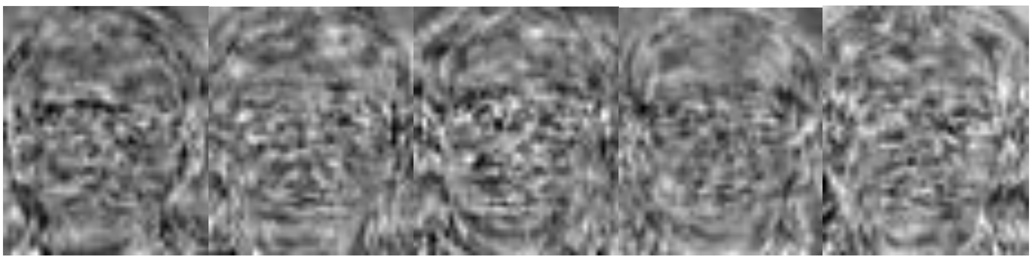
TSNE results on testing images



As we can see, the samples with different labels are separated and with same labels are clustered.

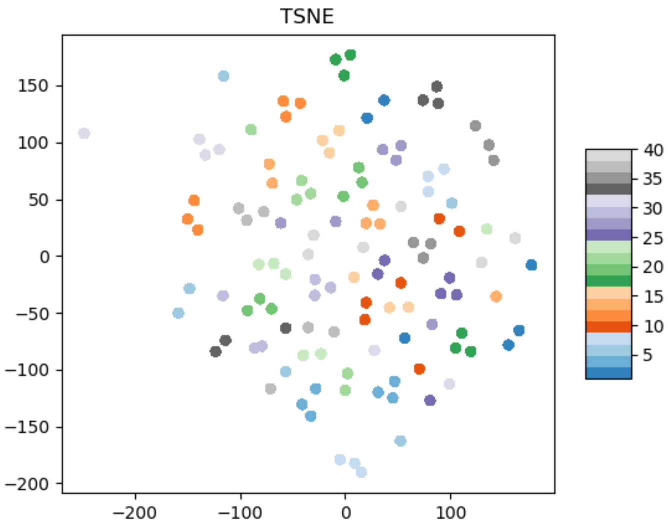
(b)

First five fisherfaces:



It seems that the results are noisy, but the contour of faces can be roughly visualized.

TSNE results on testing set:



As we can see, most samples with different labels are separated and with same labels are clustered.

(c)

Cross validation: I split the data according to their index. Specifically, fold 1, 2, 3 are with label [0, 1], [2, 3], [4, 5, 6], respectively.

In each test, I fit the PCA on **selected training set only instead of the whole training set.**

KNN for PCA (%)

k/n	3	10	39
1	71.25/77.5/55.83	93.75/91.25/81.67	93.75/93.75/88.33
3	71.25/65/50.83	81.25/83.75/72.5	91.25/87.5/81.67
5	57.5/66.25/44.16	77.5/76.25/65.83	83.75/82.5/73.33

Selected parameter: k=1, n=39

Accuracy on testing set: 95.83%

Test k=1, n=10 and k=3, n=39 since the performance is close to k=1, n=39

k=1, n=10: 94.16%

k=3, n=39: 93.33%

The selected parameters still yield best performance.

KNN for LDA (%)

k/n	3	10	39
1	45/45/38.33	82.5/82.5/75.83	96.25/91.25/89.17
3	45/43.75/40	83.75/82.5/75.83	96.25/93.75/88.33
5	45/43.75/40	83.75/82.5/75	97.5/93.75/90

Selected parameters: k=5, n=39

Compared parameters: k=3, n=39 and k=1, n=39

Results on testing set:

k/n	39
1	91.66
3	92.5
5	92.5

Discussion:

The performance of LDA doesn't outperform that of PCA. A possible reason is that the existing samples for each label is insufficient to train a LDA well.

Problem3

(a)Model architecture

Feature extractor				
Name	kernel	strides	chn. I/O	size I/O
conv0	5×5	1×1	1/8	$28 \times 28/24 \times 24$
Maxpooling0	2×2	2×2	8/8	$24 \times 24/12 \times 12$
conv1	5×5	1×1	8/16	$12 \times 12/8 \times 8$
Maxpooling1	2×2	2×2	16/16	$8 \times 8/4 \times 4$
Fully-connected network				
Name	shape I/O		Activation function	
fc0	256/128		ReLU	
fc1	128/64		ReLU	
fc2	64/10		SoftMax	

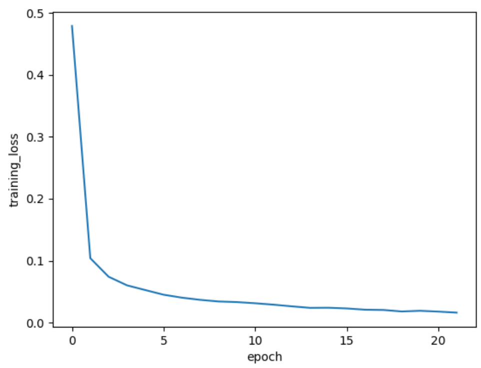
I use two convolutional layers with activation function ReLU to extract features and a fully-connected network for classification.

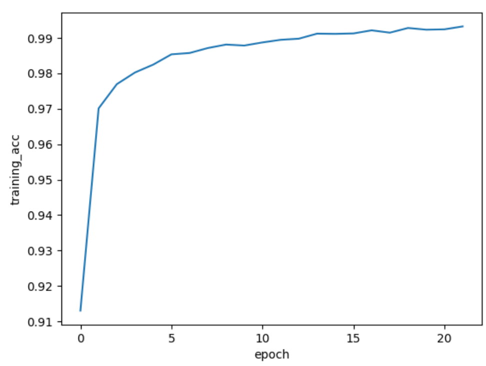
(b) Training/Validation accuracy

Highest accuracy on training/validation set: 99.25%/99.08%

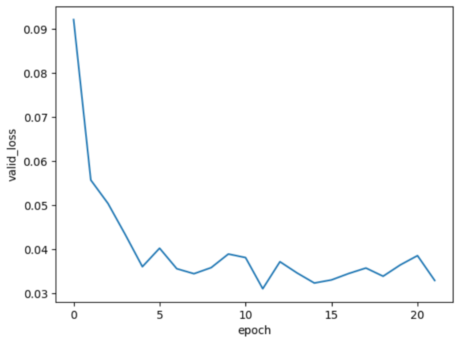
Learning curve for accuracy and loss for training and validation set:

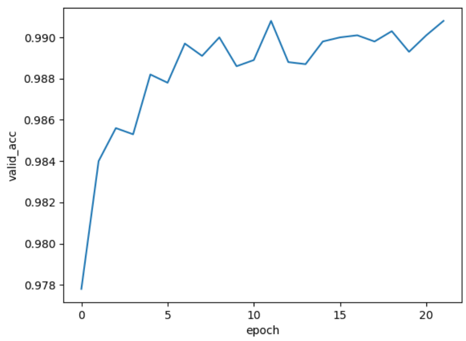
Training loss and accuracy:





Validation loss and accuracy:





(c)Visualize filters

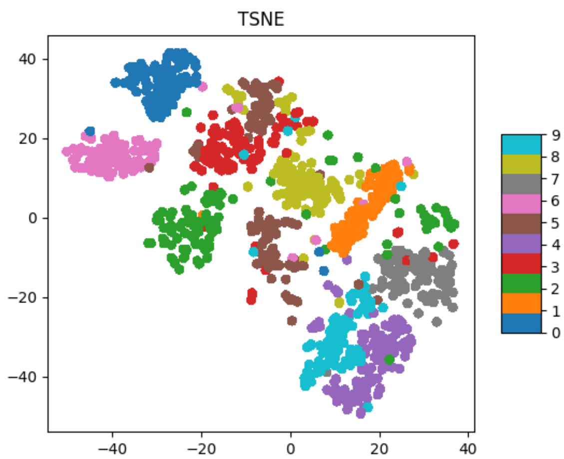
Six filters at the first convolutional layer are shown as follows:

Six filters at the second convolutional layer are shown as follows:

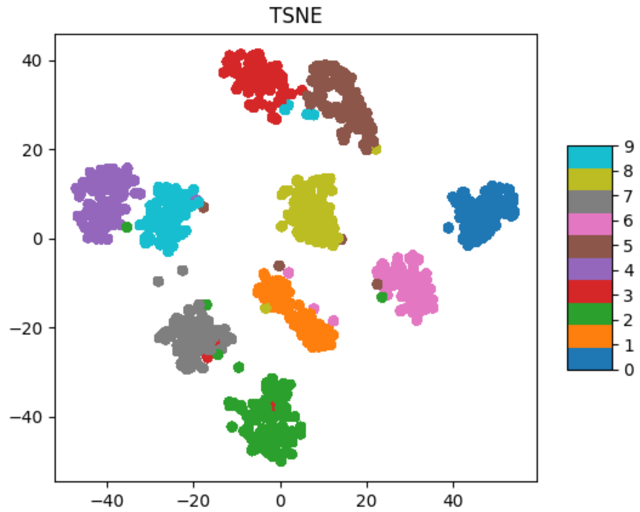
t seems that this network uses filters in the first convolutional layer to crop some region and use filters in the second convolutional layer to extract some features like textures in different shape.

(d)TSNE on CNN features

Low level:



High level:



As we can see, the clustering results is clearer in high level features. In the results of low level features, the difference between each classes is not clear. However, in the results of high level features, 10 clusters can be clearly observed.