The results of resnet18 and vgg16 with different learning rate and do experiment on Cifar10.

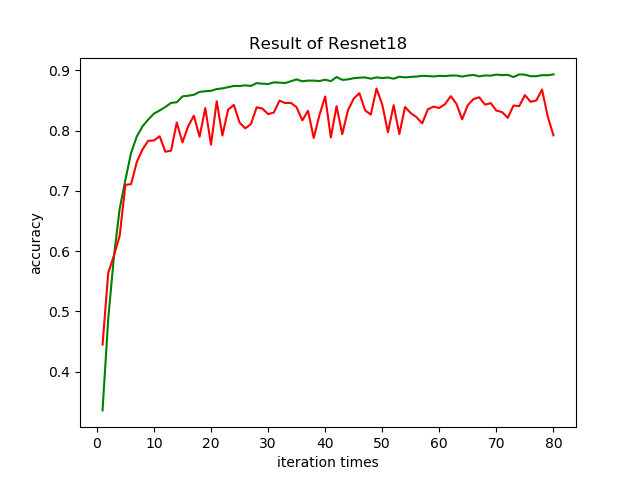


Fig 1. Learning rate = 0.1

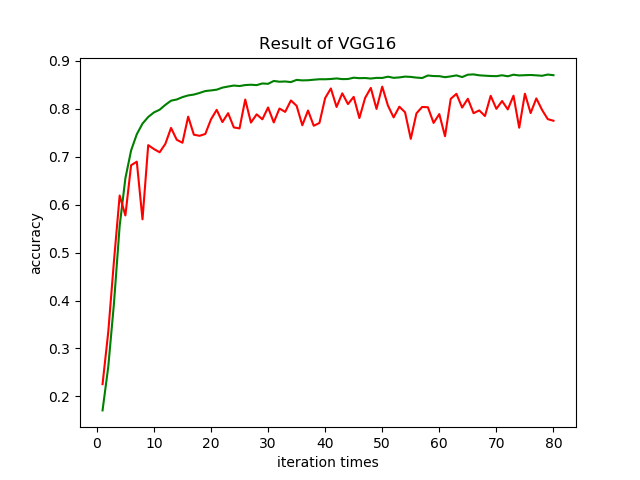


Fig 2. Learning rate=0.1

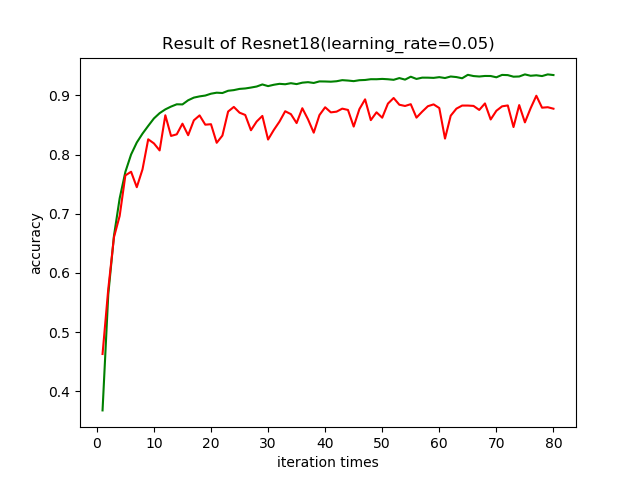


Fig 3. Learning rate =0.05

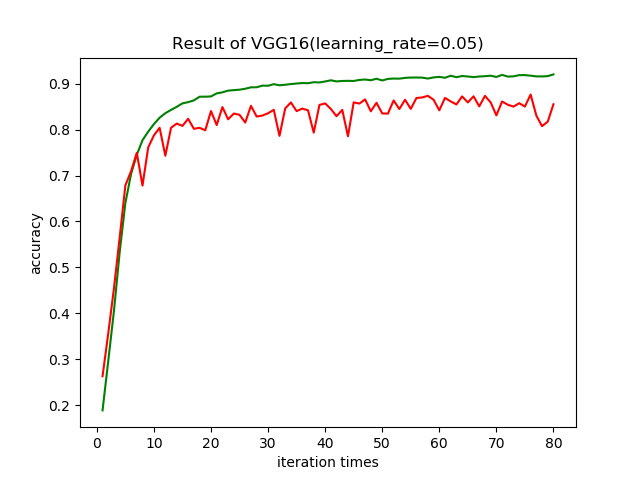


Fig 4. Learning rate=0.05

Here, we do experiment on Cifar10. The number of iterations is 80.

We use two different learning rate. We show the result of 80-th iteration.

TABLE 1. The result of 80-th iteration.

|  |  |  |
| --- | --- | --- |
|  | VGG16 | Resnet18 |
| 0.1(train) | 87.03% | 89.31% |
| 0.1(test) | 77.52% | 79.20% |
| 0.05(train) | 92.02% | 93.42% |
| 0.05(test) | 85.52% | 87.74% |

We show the result of the average accuracy from 70th to 80th iteration.

TABLE 2. The result of the average accuracy from 70th to 80th iteration.

|  |  |  |
| --- | --- | --- |
|  | VGG16 | Resnet18 |
| 0.1(train) | 87.01% | 89.15% |
| 0.1(test) | 80.00% | 83.73% |
| 0.05(train) | 91.74% | 93.36% |
| 0.05(test) | 84.60% | 87.62% |

From the above results, resnet18 with learning rate equal to 0.05 has the best performance.

Table3. The memory of two kinds of models.

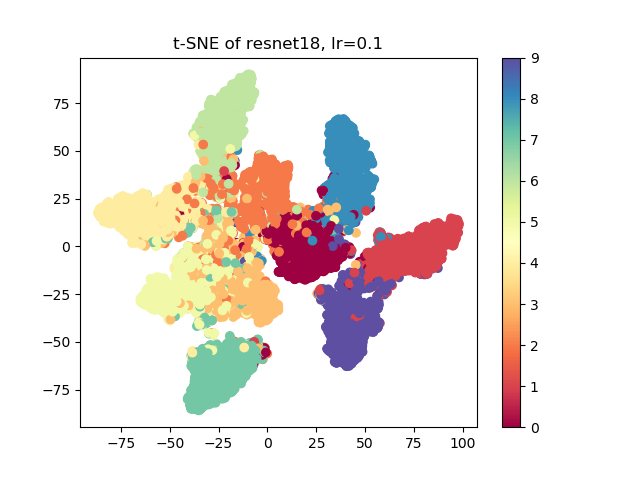
|  |  |  |
| --- | --- | --- |
|  | VGG16 | Resnet18 |
| Memory | 57.58M | 43.71M |

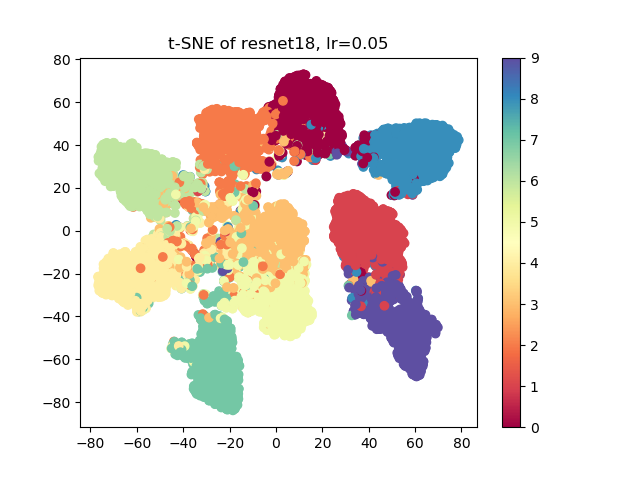
The parameters of Resnet18 are less than VGG16 (the main difference is from the convolution kernels). From the results, the results and memory of Resnet18 are better than vgg16.

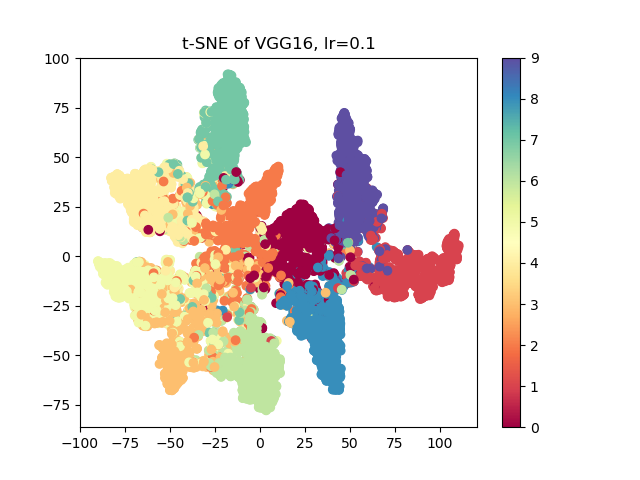
Visualization

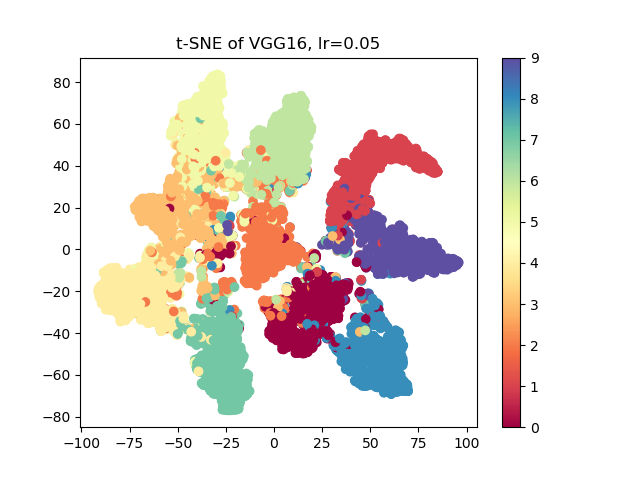
Here, the test data set is passed through the network, and the output 10-dimensional vectors are visualized with t-SNE and PCA.

t-SNE：

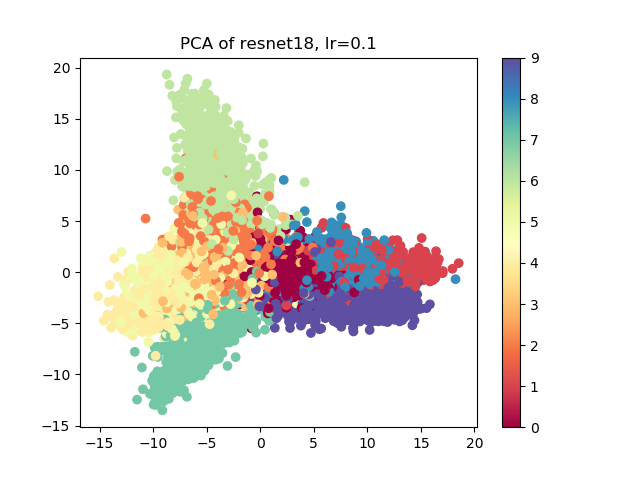


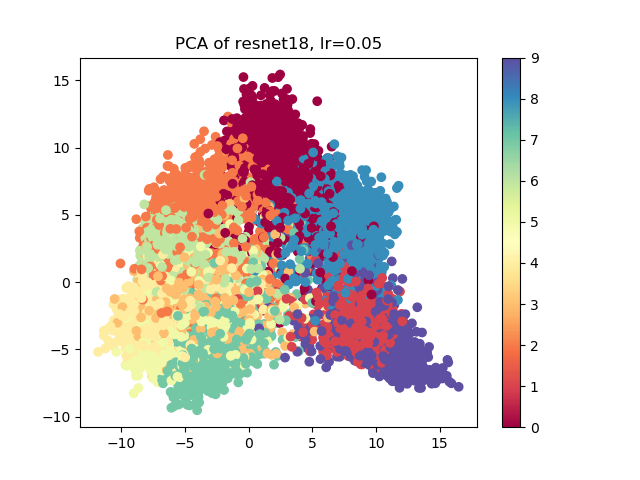


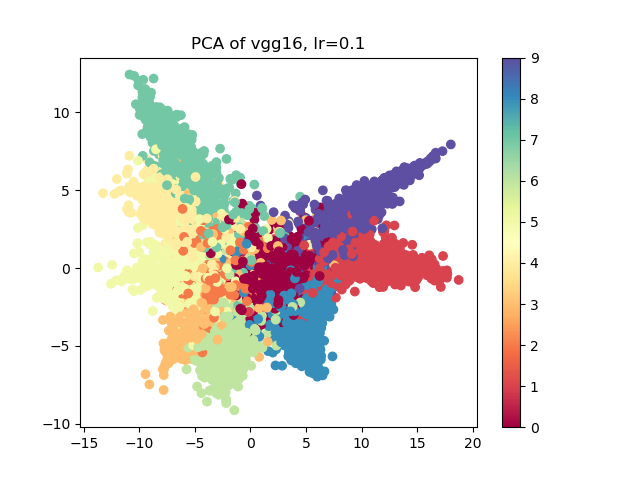


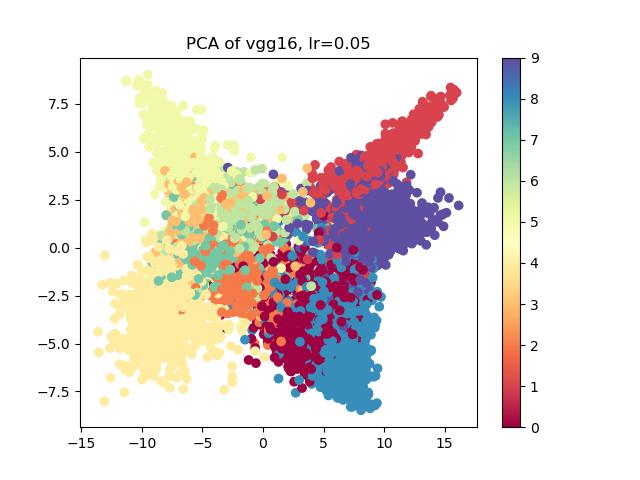


PCA：









From the above results, the visualization effect of t-SNE is much better than that of PCA, and the aliasing phenomenon is a little less.