MNIST classification

Luyu Cen, Jingyang Li, Zhongyuan Lyu and Shifan Zhao

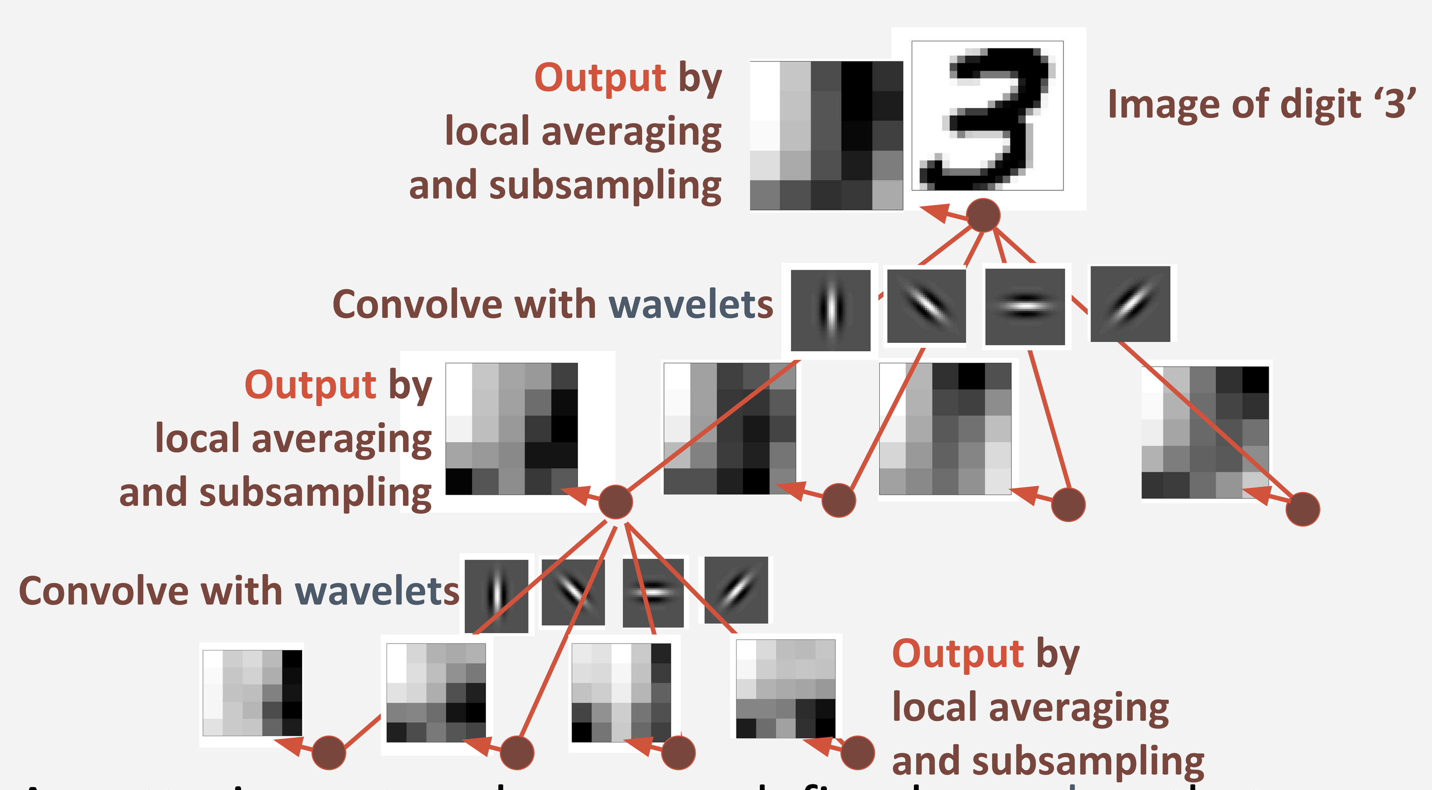
Introduction

With 60000 training examples, 10000 test examples, MNIST dataset of handwritten digits witnessed the development of classification methods since 1998.

In this project, we applied the scattering net to extract features of MNIST images. Based on these features, we used traditional supervised learning methods LDA, SVM, Random Forest to classify the images. We also tested neural networks VGG19 and Resnet on this dataset.

Feature Extraction

Scattering Net



A scattering network uses predefined wavelets that can convolve with the original image to get invariant features. By rotating the wavelet function 8 times with angle $k\pi/8$, $k=0,1,2,\ldots,7$ and scaling it by factors 2^0, 2^1, 2^2, we obtain 24 wavelets that capture different features. For illustration, we only show 4 wavelets above. In the second layer, the images generated from the first layer's convolution convolves with the wavelets again.

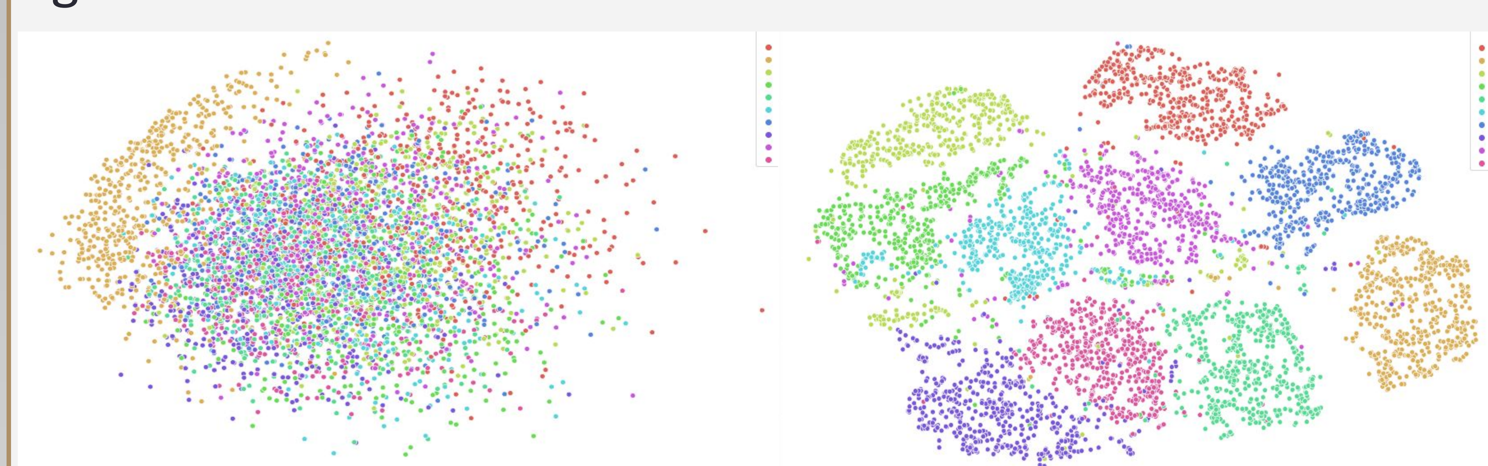
For each image in the MNIST dataset, we collected all the outputs from three layers of nodes and reshape them to be a 1×5425 row vector for later visualization and classification use.

Visualization

We implement PCA and t-SNE on the features extracted by Scattering Net for visualization in two dimensions.

In **PCA**, using only the first two PCs can hardly distinguish the different labels from each other, owing to the low percentage of variance explained by first two PCs (around 0.3) and even the first three PCs represented in 3D does not improve much.

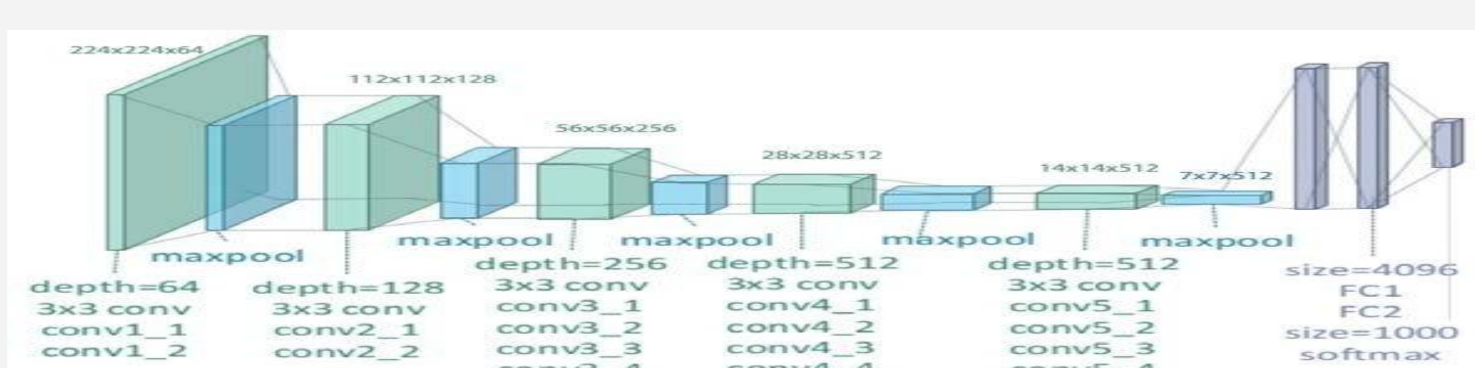
t-SNE performs much better in this case since it can better capture the non-linear structure than PCA which is essentially a linear projection. Briefly speaking, t-SNE translates Euclidean distance into similarity (conditional probability) among all points, and minimizes the total KL divergence. However, the computing cost of t-SNE is greater than PCA.



PCA t-SNE

VGG19

VGG 19 is a pre-trained model of convolution networks. This network can classify objects into 1000 classes. However, we need to do some modifications to classify MNIST datasets. So we removed 8 layers of CNN (conv3_1 to conv4_4) and adjust the network parameters. For example, the pre-trained network outputs a 1000 dimension vector, so we changed the parameter to 10 in the last layer of FC and also changed some other parameters so that the MNIST data fit the network and there are fewer parameters to train. The figure shows an original VGG 19 network. The test accuracy on the MNIST dataset is 98.9%.



Resnet

It is believed that deeper networks perform better. But aftering inserting some identity layers into CNN, people find it usually becomes less efficient. That's because of the well-known gradient explosion and vanishing. This problem was beautifully solved by Resnet proposed by Kaiming He and his collaborators. In their paper, they introduced residual learning blocks. By using feedforward neural networks and "shortcut connection", their network actually is learning residual mapping rather than the original mapping. In our context, we used the Resnet-18 for MNIST dataset. Since the original channel for Resnet is 3, we have to modify the channel of MNIST to be 10. The accuracy 98.81% after five epochs.

Results

Method\Set	Test
LDA	**0.9935**
SVM	0.9482
RandomForest	0.9833
VGG19	0.9899
Restnet	0.9881

In LDA, SVM and RandomForest we directly utilize the features from Scattering Net. LDA achieves amazing accuracy, partly due to the coincidence of the data structure with the LDA model assumption. In comparison to traditional methods, deep learning network VGG19 and Resnet achieve similar accuracy.

Conclusion

The combination of Scattering Net and traditional machine learning methods can achieve performance as good as deep learning methods. From the visualization by t-SNE, 10 classes are clearly distinguished. This might be due to the simplicity of MNIST dataset.

Contribution

Luyu CEN: Feature extraction by the scattering net.
Jingyang Li: Test VGG19 on MNIST
Zhongyuan Lyu: Visualization and classification by supervised learning methods
Shifan Zhao: Test Resnet on MNIST