MATH 6380P Mini-Project 1: Feature extraction and classification for MNIST

Ganghua FAN {gfanab}@connect.ust.hk
Department of Mathematics, HKUST

1. Abstract

In this project, I firstly apply the scattering net and pre-trained VGG19 to extract features of the MNIST dataset. By using the Principal Component Analysis (PCA), I visualize the data with reduced dimensionality. Support Vector Machine (SVM) and Logistic Regression is then trained to do the classification based on extracted features.
The result shows that ScatNet+SVM combination achieves highest accuracy up to 99.07%. The explanation about the phenomena is given at the end of this report.

2. Introduction

MNIST database is a large handwritten-digit dataset of 10 digits, which contains 60,000 training images and 10,000 testing images [1].
The images in the dataset are 28x28 grayscale images. It is widely used for training and testing in the field of machine learning.

Neural Collapse of features
(NC1) Variability collapse: As training progresses, the within-class variation of the activations becomes negligible.
(NC2) Convergence to Simplex ETF: The vectors of the class-means (after centering by their global-mean) converge to having equal length, forming equal-sized angles between any given pair, and being the maximally pairwise-distanced configuration constrained to the previous two properties.

In this project, I compute the global mean, class-means and covariance matrix etc. of features extracted by ScatNet using matlab. It turns out the relative difference of the contraction of within class variation and equal-norms of class-means is 0.92, the relative difference of equal-angularity and maximal-angle equiangularity is 0.99.

3. Feature Extraction

ScatNet: Wavelet scattering network (ScatNet) generates a representation that's invariant to data rotation/translation while being stable to deformation. It requires no training and is suitable to extract features before classification.
In this project, I use 2-layer ScatNet with 8 angles for the wavelet transform. The shape of the features from the wavelet scattering transform is (70000, 81, 7, 7) which is flattened to (70000, 3969). The ScatNet code was adapted from the Kymatio. [2]
VGG19: VGG19 is a convolutional neural network with 19 layers.
The pretrained network on ImageNet has default input size of 224x224 pixels and classify images into 1000 object categories. Here, I resize the MNIST database images to 224x224, then use weights pre-trained on ImageNet. By deleting the fully-connected layer, I get a feature extractor. The shape of the features from the VGG19 is flattened to (70000, 25088).

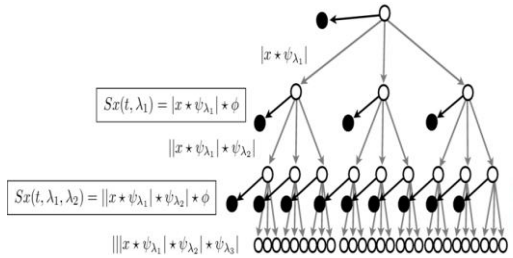

Fig 1. scattering network structure

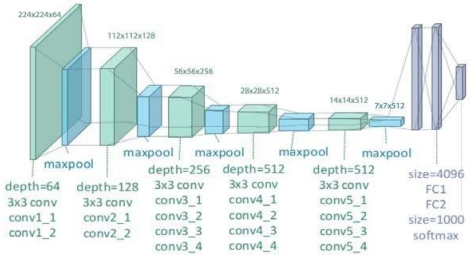

Fig 2. VGG19 structure

4. Visualization

In this project, I combine PCA (pick first 20 components) and t-SNE (reduce to 2d-space) for dimensionality reduction. The visualization of the first 1000 test data is shown in Fig 3 and Fig 4. The visualization results tell us ScatNet and VGG19 can extract useful features for classification.

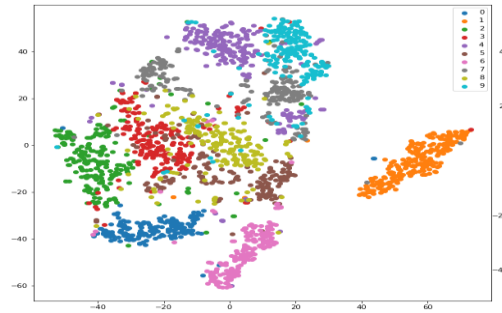
Fig 3. visualization of ScatNet

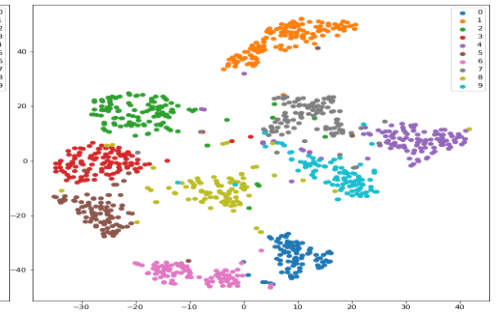
Fig 4. visualization of VGG19

5. Classification

I use SVM with Radial basis function kernel and Logistic Regression with 30 epochs to classify the images based on the features extracted above. The results of this project is listed in Table 1. The ScatNet+SVM achieves the best performance, with accuracy up to 99.07%. The VGG19+SVM also performs good while ScatNet+Logistic Regression only has 90.5% accuracy.

Table 1. Accuracy comparison

Network	Classifier	Accuracy
ScatNet+SVM	SVM	99.07%
ScatNet+Logistic	Logistic Regression	90.5%
VGG19+SVM	SVM	97.8%

6. Conclusion

The visualization of the feature extracted from pre-trained VGG19 net shows distinct clusters, which indicates the transfer learning of pre-trained VGG19 is useful on MNIST dataset. So, we can expect the fine-tune network of pre-trained VGG19 will be a good classifier of MNIST database. One thing need to mention is that VGG19 contains so many parameters, which leads to short of memory when implement on personal PC. Actually, the VGG19+SVM is only trained on part of training data in my code, so the real accuracy might be higher.

Based on the same feature from ScatNet, the accuracy difference between SVM and Logistic Regression might illustrate linear classifier is not very suitable for handwritten digits classification.

7. References

[1] Gangaputra, Sachin. "Handwritten digit database". Retrieved 17 August 2013.
[2] Andreux M., Angles T., Exarchakis G., Leonarduzzi R., Rochette G., Thiry L., Zarka J., Mallat S., Andén J., Belilovsky E., Bruna J., Lostanlen V., Hirn M. J., Oyallon E., Zhang S., Cella C., Eickenberg M. (2019). Kymatio: Scattering Transforms in Python. arXiv preprint arXiv:1812.11214.