Mini-Project 1. Feature Extraction and Transfer Learning

Zhenghui Chen, Lei Kang
Department of Chemical and Biological Engineering
HKUST
{zchenef, lkangaa} @connect.ust.hk

Abstract

In this project, we extract the feature of image on MNIST dataset and Raphael's paintings dataset by scattering net and pretrained deep neural networks VGG19 respectively. Then we use different unsupervised learning methods such as PCA/MDS, Manifold learning (Diffusion Map, ISOMAP, LLE) and t-SNE to visualize the feature vectors extracted from Raphael's painting dataset. The scatter plots of the feature vector can give a hint on the classification of the dataset. Finally, supervised learning methods such as SVM are used to achieve the image classification in Raphael's paintings and MNIST dataset.

1 Image dataset

1.1 MNIST dataset

The MNIST database is a dataset of handwritten digits. It has 60,000 training samples, and 10,000 test samples. Each image is represented by 28x28 pixels, each containing a value 0-255 with its grayscale value. The dataset is downloaded from this website.

https://github.com/datapythonista/mnist

The digits have been size-normalized and centered in a fixed-size image. And the corresponding label are included in the dataset. Figure 1 shows that one handwritten digit image from dataset.

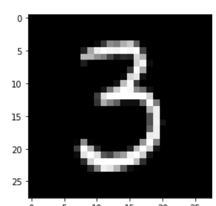

Figure 1: Handwritten digit image

1.2 Raphael's paintings

The Raphael's paintings dataset contains 28 images with the format of TIFF, TIF, JPG. The dataset is downloaded from this website.

https://drive.google.com/folderview?id=0B-yDtwSjhaSCZ2FqN3AxQ3NJNTA&usp=sharing

The images of TIFF and TIF format contains RGBA channels, we convert all the pictures to JPG format with RGB channels. We label the known Raphael image and Not Raphael image as training set, and the disputed image as test set. The training set contains 21 images and the test set contains 7 images.

2 Feature extraction

1.1 Invariant Scattering Convolution Network

We use the invariant scattering convolution network to extract the feature vector of the MNIST data. Figure 2 shows the result of different feature maps after achieving morlet wavelet convolution on Figure 1.

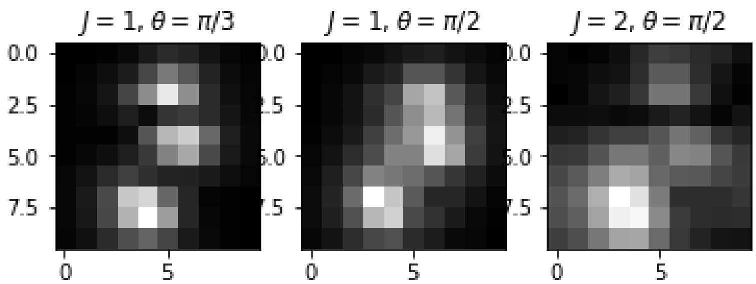

Figure 2: different morlet wavelet convolution feature maps

We choose 1000 image samples from the MNIST dataset to do the feature extraction, and the size of each image is 28x28. After the fast scattering transform, the feature vector is represented as 14416 variables of each input image.

1.2 VGG19

VGG was proposed by the Visual Geometry Group of Oxford. It is proved that increasing network depth can improve the accuracy of the network to some extent. VGG19 is a deeper convolutional neural network and use 3 × 3 filters to capture the notion of left/right, up/down, center. Figure 3 shows that the structure diagram of VGG19.

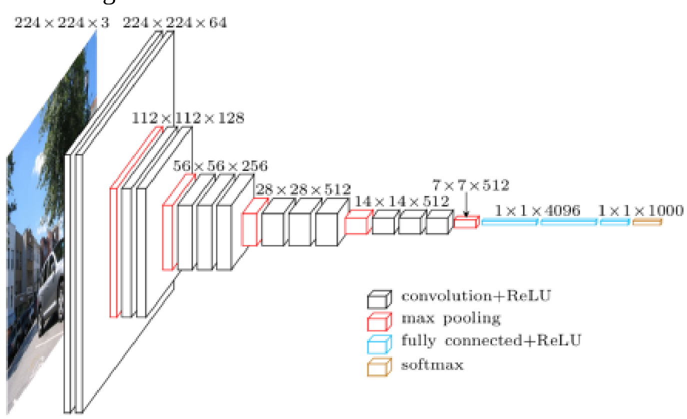

Figure 3: The structure diagram of VGG19

We use the pretrained deep neural networks VGG19 to extract the feature vector of the Raphael's paintings. As we can see from the figure 3, the input image size of the VGG19 net is fixed as [224, 224, 3] (height, width, channels). Therefore, we need to resize the image size into [224, 224], and convert all the images into 4-dimensional tensor [28, 224, 224, 3]. Figure 4(a) shows one resized painting from the dataset, Figure 4(b) shows that the output feature maps of layer conv1_1 and Figure 4(c) shows that the output feature maps of layer conv3_1. The output of layer fc7 is the feature vector with 4096 variables. We use this feature vector to do image classification.

(a) (b) (c)

Figure 4: The structure diagram of VGG19

3 Visualize the features

For each painting in Raphael's paintings dataset, we can extract one feature vector with the size of 4096x1 by VGG19 and we got 28 feature vectors in total for the whole dataset.
Unsupervised learning methods such as PCA/MDS, Manifold learning (Diffusion Map, ISOMAP, LLE) and t-SNE are used to visualize these 28 painting features and these methods can help to visualize the distance among different kinds of paintings.

3.1 Unsupervised learning methods

3.1.1 PCA

PCA (Principal Component Analysis), is one of the most widely used data dimensionality reduction algorithms. It uses orthogonal transformation to transform a set of possible variable correlation data into a set of linear uncorrelated variables, and the converted variables are called principal components.

3.1.2 MDS

Multidimensional scaling (MDS) is a means of visualizing the level of similarity of individual cases of a dataset.

3.1.3 Diffusion map

Diffusion maps is a dimension reduction technique that can be used to discover low dimensional structure in high dimensional data. It assumes that the data points, which are given as points in a high dimensional metric space, actually live on a lower dimensional structure.

3.1.4 ISOMAP

ISOMAP is an extension of MDS, where pairwise euclidean distances between data points are replaced by geodesic distances, computed by graph shortest path distances.

3.1.5 Locally linear Embedding (LLE)

The algorithm assumes that any date point in a high dimensional ambient space can be a linear combination of data points in its neighborhood.

3.1.6 T-distributed Stochastic Neighbor Embedding (t-SNE)

It is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions.

3.2 Result in Raphael's paintings

After projecting the painting feature vector of each painting into the first and second principle eigenvectors, we can get the following scatters plots. The red dots represent the painting feature vectors extracted from non-Raphael's paintings while the green dots represent the painting feature vectors extracted from Raphael's paintings. The yellow dots are the representative of painting feature vectors extracted from paintings which are disputed. In the Category, we assign different value for each class, 0 means non-Raphael, 0.5 means disputed and 1 means Raphael.

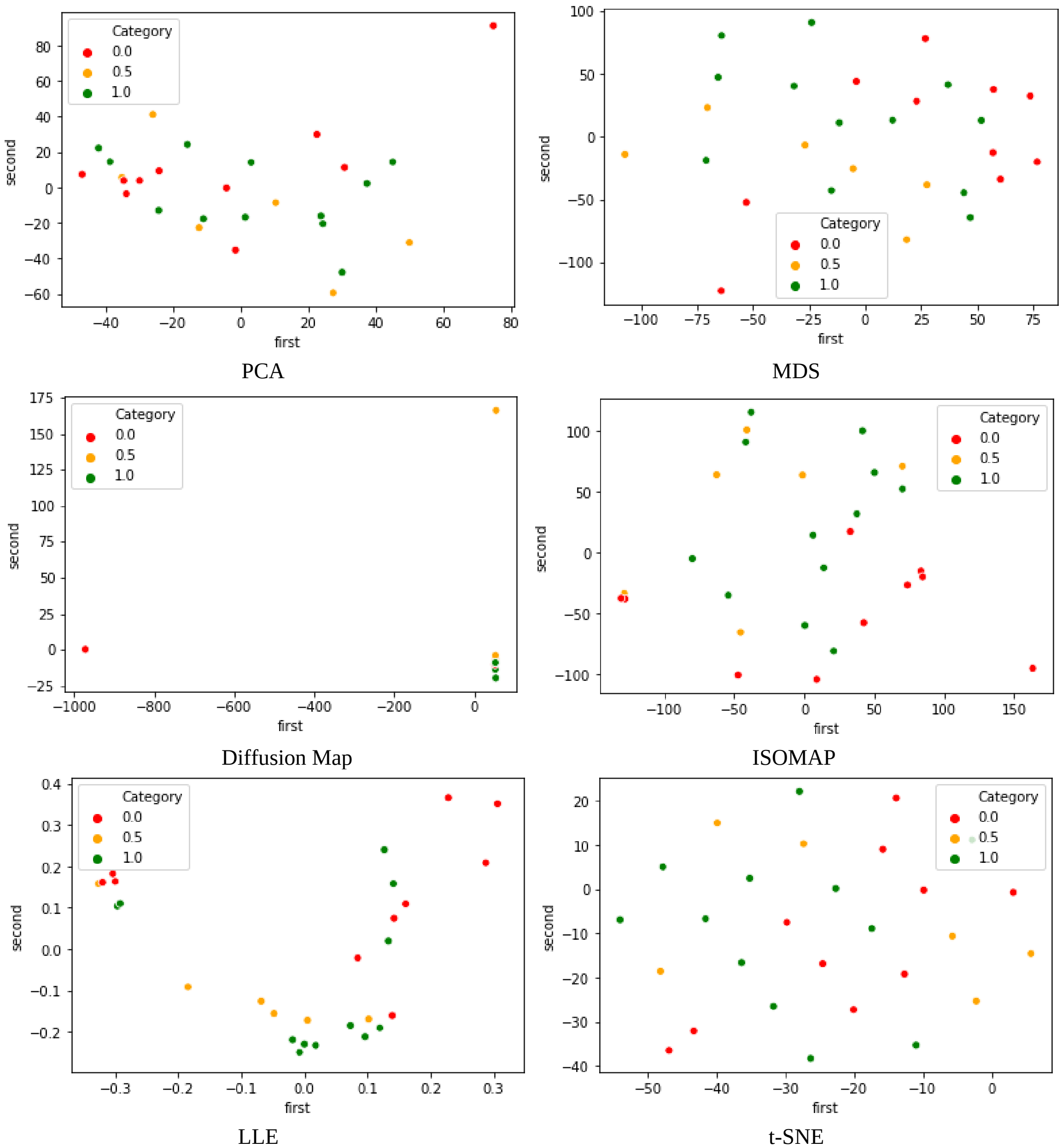

Figure 5: The scatter plots of different methods

From the six scatter plots above, it can be observed that MDS, Diffusion Map, ISOMAP, and LLE can differentiate the three classes while PCA and t-SNE fail to tell the difference. In the MDS figure, most of the non-Raphael painting feature vectors are on the upper right corner while the Raphael paintings are in the middle. In the Diffusion Map figure, the non-Raphael paintings cluster on the lower left corner while true Raphael painting are on the lower right corner. In the ISOMAP figure, the non-Raphael paintings are on the lower right corner while the true painting are in the middle. In the LLE figure, the non-Raphael painting lies on the two sides while the true and disputed paintings are in the middle.

4 Image classification

4.1 SVM

Support Vector Machine（SVM）is a supervised learning method that widely used in classification problem. In this project, we use the RBF kernel function and grid search method to find the best parameters C and gamma. C is the penalty factor, gamma determines the distribution of data after mapping to a new feature space. Finally, the best model is established by parameter optimization.

4.1.1 Result in MNIST

In MNIST, we use SVM method to classify both the original dataset without feature extraction [1000, 784] and feature vector dataset [1000, 14416]. We split the dataset into training set and test set randomly. The training set contains 700 samples and the test set contains 300 samples. The training label and test label are from 0 to 9 corresponding to the handwritten digits from 0 to 9.

Table 1: SVM model based on original dataset and feature vector

	Accuracy		Parameters	
feature vector	Train	1	gamma	10^{-9}
	Test	0.963	C	100
original dataset	Train	1	gamma	10^{-7}
	Test	0.913	C	100

Table 1 shows the results of the two models. The model based on feature vector achieves the train accuracy of 100% and test accuracy of 96.3%. The model based on original dataset achieves the train accuracy of 100% and test accuracy of 91.3%. The model based on feature vector performs better than the model based on original dataset. It shows that scattering convolution network can extract the valid feature and improve the classification result based on SVM model.

4.1.2 Result in Raphael's paintings

In Raphael's paintings, we use training set (contain 21 known paintings) to establish the SVM model and predict 7 unknown images by the model. We both use the original dataset without feature extraction [21, 224×224×3] and feature vector dataset extracted by VGG19 [21, 4096] to train the model. Then we use the model to predict the 7 unknown paintings including No.1, No.7, No.10, No.20, No.23, No.25, No.26 paintings.

Table 2: SVM model based on original dataset and feature vector

Number	Model based on	
	feature vector	original dataset
1	Raphael	Raphael
7	Raphael	Raphael
10	Raphael	Raphael
20	Raphael	Raphael
23	Raphael	Raphael
25	Raphael	Raphael

26	Raphael	Not Raphael

Table 2 shows the prediction result of the unknown paintings. The model based on feature vector predicts the No.1, No.7, No.10, No.20, No.23, No.25, No.26 disputed paintings are all Raphael paintings, while the model based on original dataset predicts No.26 painting is not Raphael painting and the other paintings are all Raphael paintings. From these two models, the No.1, No.7, No.10, No.20, No.23, No.25 disputed paintings are confirmed as Raphael paintings.

5 Conclusion

In this project, the scattering net was used to extract the feature of MNIST dataset and VGG19 was used to extract the feature of Raphael's painting dataset. Then we visualize the features of Raphael's painting dataset which can provide a hint on the classification of the paintings. Many unsupervised learning methods such as PCA/MDS, manifold learning (Diffusion map, ISOMAP, LLE), and t-SNE are used to visualize the features. We find MDS, Diffusion map and ISPMAP can successfully separate different kinds of paintings. Finally, we perform SVM on both MNIST dataset and Raphael's painting dataset to make the classification. The SVM model based on MNIST dataset achieves the 96.3% accuracy of image classification. The SVM model based on the feature vector of Raphael's painting dataset predicts the 7 unknown paintings are all Raphael paintings.

6 Contribution

Zhenghui Chen: Code for the feature extraction, image classification
Zhenghui Chen: Report writing for the feature extraction, image classification part
Lei Kang: Code for visualize the features
Lei Kang: Report writing for the abstract, visualize the features and conclusion part

References

[1] J. Bruna and S. Mallat, "Invariant scattering convolution networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 1872–1886, 2013.
[2] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.