MATH 6038o Mini-Project 1: Spotting The Real Raphael's Sketches

香港科大 HKUST

XU SIAO, TAN JUNHONG, HE YUNJIE, LU JIHUA; Department of Mathematics

Introduction

As we all know, art fraud is very common, and identification is very difficult. With the development of artificial intelligence, we may no longer need such trouble. A digital approach via deep learning has been shown to work efficiently. In our project, we use the model to identify whether the 6 paintings are from Raphael.

Data

The training and test sets are made up of 28 labeled digital paintings of Raphael or forgeries either in jpeg or tiff format,12 of which are from Raphael, 10 are not and 6 remaining disputed.

Methodology

Models

- **Geometric Tight Frame**: This method extracts features by applying 18 pre-selected filters. For classification, we use 4 traditional machine learning methods, including logistic, Naïve Bayes, support vector machine and random forest.
- **Res-Net**: We use the output before fully connected layer for feature extraction, and change the structure of fully connected layer for fine-tuning.

Evaluation

- **Cross--one-out –validation** : We remove the disputed pictures. Each time we select a known picture as test data and use the remaining pictures as training set. We repeat this process in order.
- **Prediction**: We take the whole 22 labeled ones as our training set and train the model, after which we predict the 6 unlabeled ones whether they are forgeries or not.

Feature Extraction

Geometric Tight Frame

Based on geometric structure of data, we extract 18 filters to represent transformation towards 18 directions. They are then used to calculate the mean, variance and tail in the next step, where the three indices are extracted as features.

Res-Net

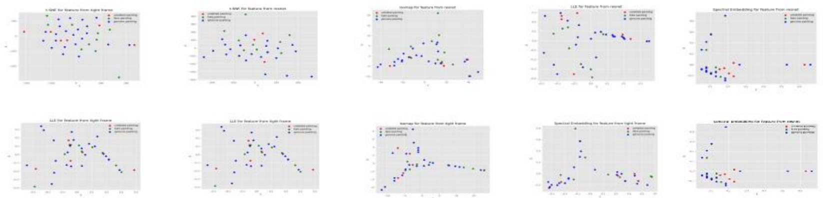
Res-Net means residual network, and is used in both feature extraction and classification. To implement this, we first divide every image into 100 small images to get more samples, then we resize images to 96*96 as input. For conveniently dealing with data, we build our own Res-Net instead of using package. The output of convolution layer and pooling layer can be seen as the feature of image.

Scattering Network

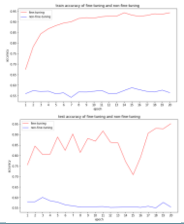
A wavelet scattering network computes a translation invariant image representation, which is stable to deformations and preserves high frequency information for classification. In order to run the ScatNet package in MATLAB successfully, we change all the images to 8 bit-depth and resize them to 512 pixels * 512 pixels. With ScatNet package, we are able to look at the filters and perform scattering display.

Visualization

Manifold learning recovers low-dimensional manifold structures from high-dimensional sampling data, that is, to find low-dimensional manifolds in high-dimensional space, and to find the corresponding embedded mapping in order to achieve dimension reduction or data visualization. We implemented five Manifold learning methods (Isomap, LLE, Spectral Embedding, t-SNE, PCA) to reduce from 35 * 54 to 35 * 2 (GTF) and 35 * 25088 to 35 * 2 (RN).



Fine Tuning



In transfer learning, we use fine-tuning to re-train our Res-Net. First, change output dim of fully connected layer from 10(CIFAR10) to 2(Belongs to Raphael or not), then we can re-train Res-Net. Compared to non-fine-tuning(directly input images to train Res-Net), fine-tuning gets higher accuracy and uses less time. We believe the small samples (1000) are not enough to train Res-Net directly and will result in getting bad accuracy.

Image Classification

Image 1

Image 7

Image 10

Image 23

Image 25

Image 26

Feature Extraction	Classification	CV Accuracy
GTF	Logistic	67.52%
	Naïve Bayes	67.52%
	SVM	67.52%
	Random Forest	53.72%
RN	RN	95.19%

Table 1 : Model Accuracy

ID	Logistic	Naïve Bayes	SVM	RF	RN
1	1	1	1	1	1
7	0	0	0	1	1
10	1	0	0	1	1
23	1	1	1	1	1
25	1	1	1	1	1
26	1	1	1	1	0

Table2 : Predictions of unlabeled images. "1"represents the image as Raphael's painting and "0" the opposite.

Conclusion & Analysis: From the table above, we find that features extracted from Res-Net are more accurate and provide more accuracy when we fine-tune the deep neural network for image classification compared to the ones from GTF and classification using traditional machine learning methods. In image recognition, the traditional method is to extract handcrafted features and then feed them to traditional machine learning methods. We believe that the reason why deep neural networks have much better performances is that they learn by creating a more abstract representation of data as the networks grow deeper, as a result, the model automatically extracts features and yields higher accuracy results.

Contribution

Every team member evenly contributes to all the tasks, including math modeling, data preprocessing, code writing and poster producing.

Reference

- St´ephane Mallat E´cole Normale Sup´erieure, CNRS, PSL .Understanding Deep Convolutional Networks.2016.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren. Deep Residual Learning for Image Recognition.2015