ARTISTIC AUTHENTICATION OF RAPHAEL PAINTINGS

YUE LI , JINGLIN CHEN, SHIJIE CUI, YUXUAN ZHOU   SCHOOL OF MATHEMATICS SCIENCES, PEKING UNIVERSITY

INTRODUCTION

Technique advances in the past decade make it possible to perform artistic identification using mathematical analysis. Our task is to discriminate genuine Raphael paintings from forgeries.Thanks to Prof. Wang from HKUST, we obtain 28 high-quaily scans of Raphael style paintings. After using geometric tight frame and Gabor wavelet to extract features from the paintings, we develop several models to classify the fakes and genuine. Leave-one-out cross validation shows that wavelet method with neutral network and support vector machine perform best among all methods, reaching an accuracy of 95.2% and 90.5% respectively. From our analysis, we also know that the genuine and fake pictures vary most in certain directions, reflecting the unique painting style of the painter. Finally, we make predictions about the seven disputed pictures in the dataset.

DATA DESCRIPTION

The data set provided by Prof.Yang Wang from HKUST consists of high resolution scans of 28 paintings, scaled to a uniform density of 200 dots per painted inch. The picture size vary from 1192*748 to 6326*4457 pixels. Of the 28 paintings, 12 have been attributed to Raphael, 9 have been known to be non-Raphael, and others are currently questioned by experts. For easier studying and feature capturing, we change them into grey-scale scans.

REFERENCES

[1]Trevor Hastie, Robert Tibshirani, Jerome Friedman.*The Elements of Statistical Learning : Data Mining, Inference, and Prediction.* Springer. 2009

[2]Haixia Liua, Raymond H. Chana, Yuan Yao. *Geometric tight frame based stylometry for art authentication of vanGogh paintings.* Elsevier. 2015

[3]C.R. Johnson Jr., E. Hendriks, I.J. Berezhnoy, E. Brevdo, S.M. Hughes, I. Daubechies, J. Li, E. Postma, J.Z. Wang. *Image processing for artist identification*, IEEE Signal Process. Mag. 25 (2008)37-48.

FEATURE EXTRACTION

Tight Frame Method With Moment Statistics

First, we use one-level geometric tight frame [2] to capture the first- and second-differences in the horizontal, vertical and diagonal directions in every small neighborhood of the paintings, as is proposed in [2]. The geometric tight frame we use has 18 filters, so there are 18 corresponding matrices for each painting after the decomposition, see[2]. To capture the characteristics of the decomposition, we use three statistics as proposed in[2];they are, specifically,

$$\mu^{(i,j)} = \frac{1}{m_i n_i} \sum_{l=1}^{m_i} \sum_{k=1}^{n_i} a_{l,k}^{(i,j)}$$

$$\sigma^{(i,j)} = \left(\frac{1}{m_i n_i - 1} \sum_{l=1}^{m_i} \sum_{k=1}^{n_i} (a_{l,k}^{(i,j)} - \mu^{(i,j)})^2\right)^{\frac{1}{2}}$$

$$p^{(i,j)} = \#\hat{A}^{(i,j)}/(m_i n_i)$$

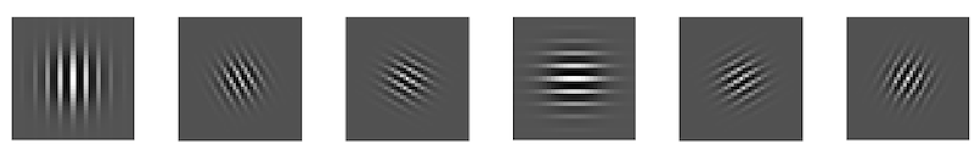
Gabor Wavelet Transformation With Energy Values

Next, we apply Gabor wavelet transformation[3] in feature extraction. Gabor wavelet is similar to human visual system in terms of response mechanism, and can capture local features in multiple orientations and scales.

The definition of Gabor wavelet filters come in pairs: $G_{\text{even}}(x, y, \sigma, \alpha, \omega)$ and $G_{\text{odd}}(x, y, \sigma, \alpha, \omega)$, which are the real and imaginary parts of the function below

$$e^{2\pi i\omega(x \sin\alpha + y\cos\alpha)} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

We use six orientations ($\alpha = k\pi/6$, for each $k \in \{0, 1, 2, 3, 4, 5\}$) and four scales for analysis, as shown below. For each of the 24 filter, we use the Gabor wavelet energy value, i.e., the sum of the squared values obtained by convolving both components with an image patch, and the moment statistics, to extract the features, see[3].



EXPERIMENTS

We use leave-one-out cross-validation(CV) procedure to measure the accuracy of our models. It should be noted that, in the model training process of foward stage-wise feature selection and neutral network, we use the original gray-scale scans; in the training of support vector machine and decision tree, we split each painting into 16 patches to increase the sample size.

Feature Extraction	Model	TPR (TP)	TNR (TN)	Classification Accuracy
Tight Frame	Forward Stagewise	75%	88.9%	81.0%
	SVM	83.3%	77.8%	81.0%
	Decision Tree	67.7%	77.8%	71.4%
	Neural Network	83.3%	88.9%	85.7%
Wavelet	Forward Stagewise+SVM	83.3%	100%	90.5%
	Neural Network	91.7%	100%	95.2%

Table 1: Experiment Results

MODEL FITTING

In the feature selection and classification process, the following models are used.

- **Forward Stage-wise Feature Selection**. We do forward stage-wise feature selection according to the area under ROC curve. The result for tight frame situation is shown in Figure 1.
- **Support Vector Machine**. Linear and quadratic kernel are used in tight frame situation and wavelet situation respectively.
- **Decision Tree**. The result of feature selection based on tight frame is shown in Figure 2.
- **Neutral Network**. We use 1-hiden-layer neutral network with 6 hidden units. The connection between layers are linear, sigmoid and linear respectively.

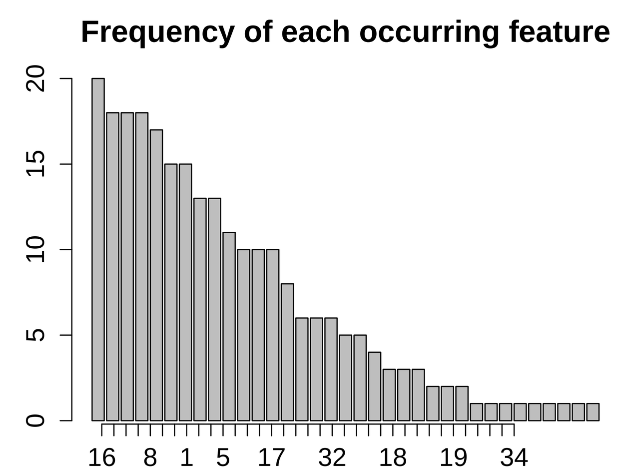


Figure 1: Forward stage

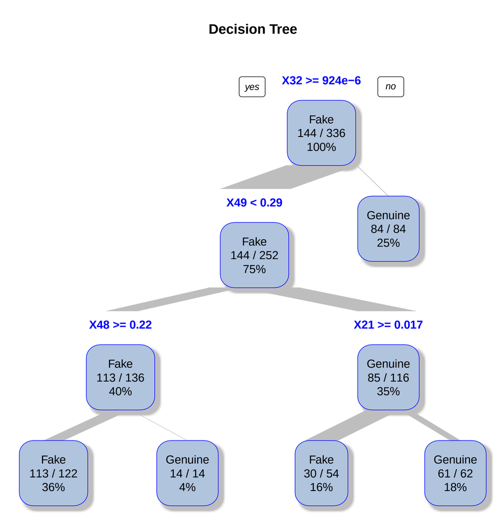


Figure 2: Decision tree

CONCLUSION

We use both tight frame method and wavelet transformation to extract brushwork features. After that, several kinds of models are used. From Table 1 we can see that the results based on wavelet is better. Thus wavelet method may be preferred for this problem.

In features selecting, Forward stage-wise and Decision tree pick up 9 and 4 features respectively. τ_{16} and τ_{17} rank high in both methods, indicating the genuine and fake pictures can be separated in vertical direction. Through models built by wavelet and forward stage-wise method, we can tell that at vertical, $\pi/3$ and $2\pi/3$ direction, the difference between genuine and fake is notable.

Upon all the models, we give our prediction to the 7 pictures remain disputed (Pic 1/7/10/20/23/25/26). Most experiments show that Picture 1/10/20/26 are genuine, and Picture 7/23 are counterfeit. For picture 25, our results varies, so we have reservations about it.

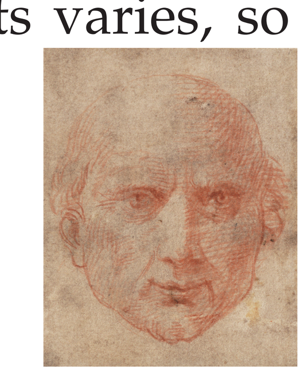


Figure 3: Classified as counterfeit by all (No.23)

Figure 4: Remain controversial (No.25)