Identification of Artist of Paintings

Gyuhak Kim[*1], YingYing Shi[†1] and Yue Xiang[‡2]

[1]Department of Mathematics, New York University
[2]Department of Computer Science, New York University

Abstract

This project adopted machine learning algorithms to classify the authors of paintings. We extracted features of paintings by convolutional neural network. The result shows that the algorithms are strong at identifying authors between different art movements such as Impressionism and Surrealism. The algorithms were still able to distinguish the most similar painters with fairly high precision.

1   Introduction

Recognizing the author of a painting is not easy for many people. Our project was inspired by the difficulty, and we hoped that machine learning could assist us on understanding arts. This report describes an approach to predicting the artist of a painting. This identification problem consists of two stages: feature extraction and classification. During our project, we perform an evaluation of possible feature extraction methods, as well performance of different classification algorithms. We believe that this technique could be extended to the field of aesthetics analysis and authentication.

2   Dataset

Our dadtaset consists of 100 paintings for each painters. We have 11 painters, thus the total number of paintings is 1,1000. We separated the dataset into 65% training, 15% validation and 20% testing. All paintings are downloaded from open source website WikiArt [1]. We removed frames or distorted area in a painting. We further labeled our dataset into three subgroups by styles: Renaissance, impressionism and surrealism. The main reason for labeling the styles is to see the performance of the algorithms on differentiating similar paintings within the styles or between different styles. Our 11 painters are the following:

Impressionist: Claude Monet, Vincent van Gogh, Camille Pissarro, Pierre-Auguste Renoir, Alfred Sisley
Renaissance: Pietro Vannucci, Tintoretto, Titian, Sandro Botticelli
Surrealism: Rene Magritte, Salvador Dali

In Figure 1, we presented an example of paintings to emphasize the similarity and difference within and between the styles.

---

[*]gk1326@nyu.edu

[†]gloriashiyy@nyu.edu

[‡]yue.xiang@nyu.edu

Figure 1: From left to right, *Argenteuil*, *Sand on the Quayside (Port Marly)*, and *Madonna and Child with Six Saints* by Claude Monet, Alfred Sisley, Sandro Botticelli, respectively. Monet and Sisley are impressionists. They share similar themes and painting styles. On the other hands, Botticelli is from Renaissance. He usually illustrated stories from the Bible.

3 Feature Extraction

We chose Caffe [2], an open-source framework, and tried several convolutional neural network (CNN) architectures trained by different organizations. The architectures we tried include BVLC GoogleNet [2], Fine-tuning on Flickr Style [4], BVLC AlexNet [5] since they aimed to classify by content, texture, or style. We chose the the architecture and layers with best performance, and it was evaluated by multiclass SVM.

We have chose BVLC GoogleNet as our final model. The BVLC GoogleNet of Caffe was trained on 1.2 million images into 1,000 leaf nodes. University of California, Berkeley, trained it to classify pictures by content. We investigated features from the second last level of the network, referred to as pool5/7x7s1. The reason to use the second last level is because it is of high dimension and well-trained after dozens of layers. Each picture is transferred to a 1024-dimensional vector and is computed from images center-cropped and re-sized to 256 by 256 pixels.

In order to illustrate features derived by deep learning, we performed principle component analysis (PCA) on our training set. The PCA result with three leading factors is presented in Figure 2. The first graph reveals that painters from different styles are highly separable. For example, we could expect that the artworks by Monet and Botticelli would be easily distinguishable as we saw in Figure 1. On the other hands, the second graph shows that painters from the same styles are hard to separate. This is because painters from the same art movements would share similar themes or styles. On the last graph, we can see the five impressionists painters are assorted, from which we could expect non-linear kernels or training models would be required for a better performance.

There were other feature extraction methods such as histogram of oriented gradients (HOG) or SIFT. Although deep learning is regarded as the state of the art method for image classification, we may find a combination of several feature extraction methods to improve the performance.

4 Learning Algorithms

We implement several learning algorithms to distinguish between the whole classes: 11 artists, between different subclasses: 5 impressionist painters, 4 Renaissance painters and finally, between style: Impressionism, Renaissance and Surrealism. Based on our best performance of different method, we implement hard ensemble method.

For linear classifier, we tried support vector machine (SVM) and linear discriminant analysis (LDA). Inspired by the paper [3], we tried SVM with stochastic gradient descent to improve training speed. We chose hinge loss function and l_2 regularization. Since the features rarely had
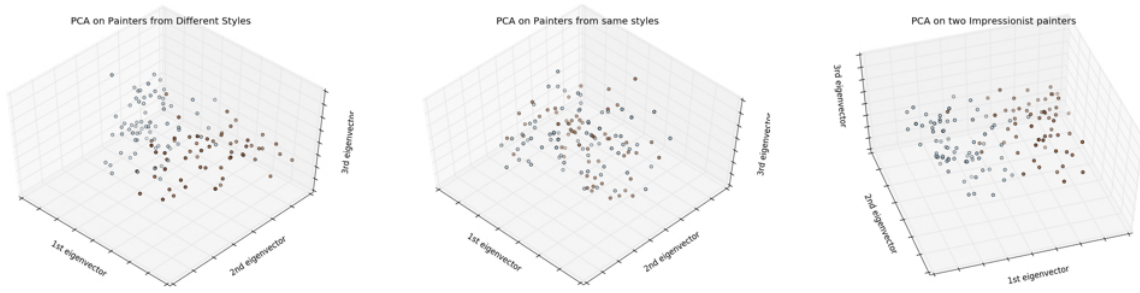
Figure 2: The first graph shows that it is separable between different styles. The second and third graph show that it is hardly separable within styles.

	SGD	SVM	kNN	LDA	QUA	RF	Ensenble
Whole dataset	0.39	0.39	0.54	0.56	0.91	0.44	0.38
Impressionist	0.24	0.32	0.42	0.37	0.76	0.27	0.24
Renaissance	0.21	0.18	0.27	0.3	0.76	0.33	0.18
Style	0.2	0.16	0.24	0.2	0.72	0.2	0.2

Table 1: Error rate of each method. Impressionist (Renaissance) represents classification within the style. Style represents classification between styles.

sparsity problem, l_2 regularization is used over l_1 regularization. We chose polynomial kernel of degree 2. The choice of degree can be explained by the PCA graph that shows quadratic boundary between different classes. Polynomials of degree higher than 2 were sensitive to overfitting.

For non-linear classifier, we used random forest (RF), k-nearest neighborhood (kNN), and quadratic discriminant analysis (QUA). For random forest, we restricted the choice of splitting variable to be the square root of 1024. Also, we chose the maximum depth of the tree to be 40. The depth of tree less than 40 were sensitive to overfitting. Also, 1000 number of trees were used in the forest.

Finally, hard voting ensemble method is used to improve prediction. Hard voting method decides the final prediction by the majority. For example, if two models vote for label A and one vote for label B, the final decision will be label A. We have chosen the three best classification models as the voting participants. They are SGD, SVM, and random forest. The other methods were discarded because their error rate was higher than 50 percent.

The ensemble method gives a slightly better result than the average of the participating models. We could obtain 2 percent lower error rate after ensemble method. Hard voting did not always give a better result than the single classification model. If a majority votes the wrong label, the prediction will be wrong. This explains why the ensemble method gave higher error rate than SVM when it classifies styles.

The reason why ensemble method does not give much higher prediction is that we only used three classification models in the voting process. In particular, if some of the models are similar, they are likely to miss the same pattern in the data, and the ensemble method will not be so different from the single classification model. We used SGD, SVM, and random forest in the voting. Since SGD and SVM are similar in techniques, they might overlook the same problem. The detailed error rate is presented in Table 1.
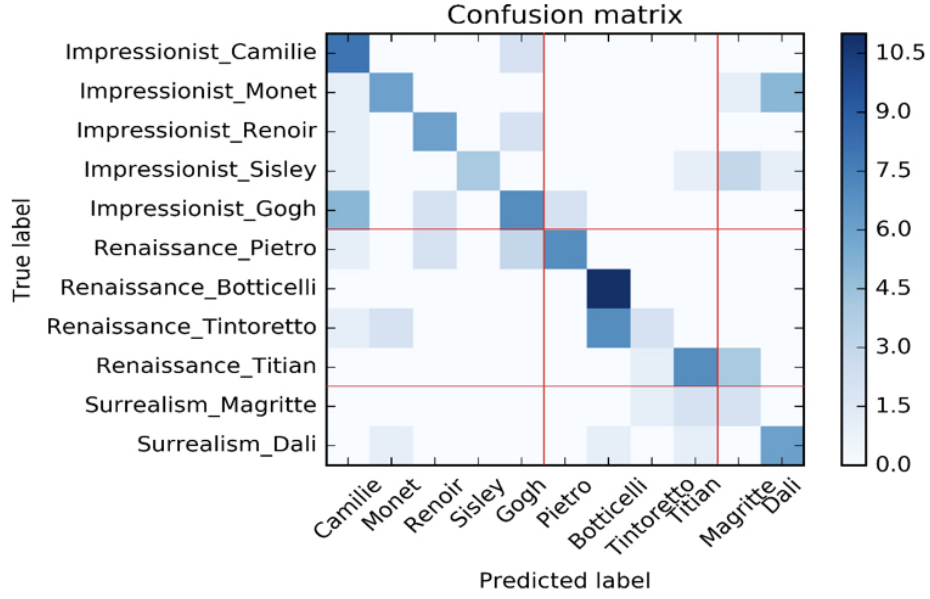
3

Figure 3: The diagonal entries are distinctly darker than the off-diagonals. This indicates that the prediction was fairly accurate.

5 Analysis

We analyze our result by drawing a confusion matrix based on the prediction rate given by the ensemble method. A confusion matrix M summarizes the performance of the final model and it is presented on right. For example, the entry M_{21} detailed the frequency that a Monet painting is classified as Camille. Darker entry indicates a higher frequency.

The diagonal entries are relatively darker than the off-diagonal entries. This implies that each artist is mostly identified correctly by the final model. The confusion matrix is separated into 6 blocks by red lines. The diagonal block matrices show the model prediction within the same style while the off-diagonal block matrices show the prediction between different styles. Entries on the diagonal blocks are relatively darker than those on the off-diagonal blocks. This implies that the wrong predictions are usually occurred within the same style.

The entry M_{78} of the confusion matrix shows that the paintings by Tintoretto were misclassified frequently as those by Botticelli. The misclassification is explained by their similar painting styles and the themes they portray. Both artists drew religious stories or figures such as the Madonna or Evangelists from the Bible. The contents of the other Renaissance painters are slightly different. Titian usually drew portraits and Pietro painted the Madonna and child.

6 Future Work

There are three directions to improve the classification accuracy. The first is to increase the size of training set by collecting more paintings for each artist. The second is to try more classification models, then implement the ensemble method again. If we had more models in the ensemble method, the prediction accuracy would improve. The third is to try classification using part of the picture or picture with shade or distortions. We can apply this technique to practical recognition of paintings.

References

[1] Wikiart. https://www.wikiart.org/. Accessed: 2016-03-23.

[2] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. *CoRR*, abs/1310.1531, 2013.

[3] S. Karayev, A. Hertzmann, H. Winnemoeller, A. Agarwala, and T. Darrell. Recognizing image style. *CoRR*, abs/1311.3715, 2013.

[4] S. Karayev, M. Trentacoste, H. Han, A. Agarwala, T. Darrell, A. Hertzmann, and H. Winnemoeller. Recognizing image style. *arXiv preprint arXiv:1311.3715*, 2013.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.