Towards Automated Classification of Fine-art Painting Style: a Comparative Study

Ravneet Singh Arora and Ahmed Elgammal
Department of Computer Science, Rutgers University, NJ, USA
{ravarora ,elgammal}@cs.rutgers.edu

Abstract

This paper presents a comparative study of different classification methodologies for the task of fine-art genre classification. 2-level comparative study is performed for this classification problem. 1st level reviews the performance of discriminative vs. generative models while 2nd level touches the features aspect of the paintings and compares Semantic-level features vs low-level and intermediate level features present in the painting.

1. Introduction

In the last decade there have been impressive advances in developing computer vision algorithms for different object recognition-related problems including: instance recognition, categorization, scene recognition, pose estimation, etc. When we look into an image we not only recognize object categories, and scene category, we can also infer various cultural and historical aspects. For example, when we look at a fine-art paining, an expert or even an average person can infer information about the genre of that paining (e.g. Baroque vs. impressionism) or even can guess the artist who painted it. This is an impressive ability of human perception for learning complex visual concept. In this paper we look into the problem of automatic classification of the genre of fine-art paintings.

Besides the scientific merit of the problem from the perception point of view, there are various application motivations. With the increasing sizes of digitized art databases on the Internet comes the daunting task of organization and retrieval of paintings. There are millions of paintings present on internet. To manage properly the databases of these paintings, it becomes very essential to classify paintings into different categories and sub-categories. This classification structure can be utilized as an index and thus can really improve the speed of retrieval process. Also it will be of great significance if we can infer new information about an unknown painting using already existing database of paintings. Also, with the explosive use of smart phones, there is a need for developing applications that automatically recognizes the genre, era, artist, and identity of paintings for tourism and museum industry.

In this paper we report the outcome of a comparative study of different classification methodologies for the task of automated classification of the genre of fine-art paintings. We defined a classification task between seven fine-art genres: *Renaissance, Baroque, Impressionism, Cubism, Abstract, Expressionism, and Popart.* Figure 1 shows an example of each of the seven categories. This is quite a challenging task since it takes an expert to define what characterizes such genres and the boundary between them. It can be combination of many features: low-level features such as color, texture, shading, stroke pattern; mid-level features such as line styles, geometry, perspective; or high-level features such as objects presence, or scene decomposition. How hard is this recognition task for computers? what is the performance of the state-of-the-art algorithms in the field of computer vision and pattern recognition on this task.

We approach the problem from a supervised learning perspective. The goal is to evaluate generative vs discriminative models as well as low to intermediate-level vs. semantic-level features. We evaluated different classification methodologies, namely: 1) A Discriminative model using a Bag-of-Words (BoW) approach; 2) A Generative model using BoW; 3) Discriminative model using Semantic-level features. We also evaluated various feature descriptors within these models. It is worth noting that, instead of using low-level features like color, light, shades and texture our study is focused on Intermediate level features (BoW features) and Semantic-level features. **Figure 2** illustrates the organization of our comparative study.

Preprint submitted to 21st International Conference on Pattern Recognition.
Received April 7, 2012.

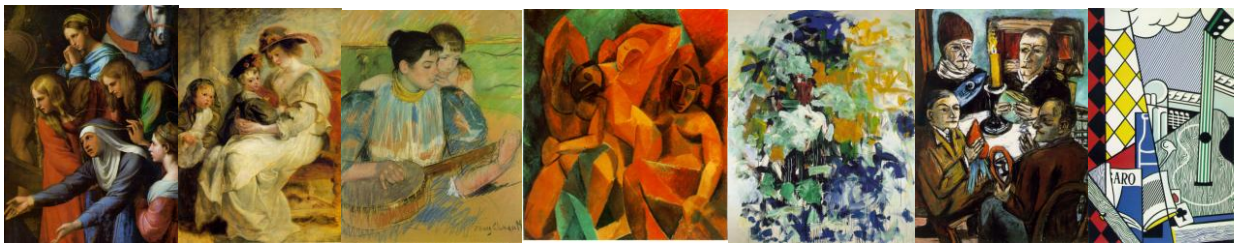

Figure 1 Example painting used in our evaluations. Left to right: Renaissance, Baroque, Impressionism, Cubism, Abstract, Expressionism, and Popart.

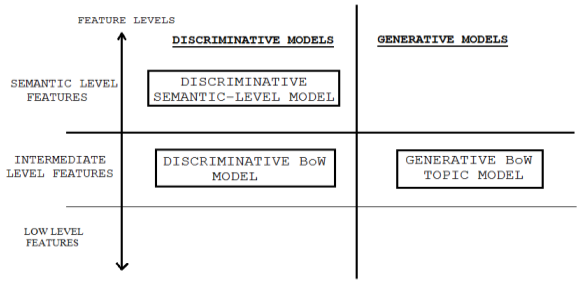

Figure 2 Classification Methodologies

2. Previous Work Overview

There is very little work done in the area of automated fine-art classification. Most of the work done in the problem of paintings classification utilizes low-level features such as color, shades, texture and edges. Lombardi [5] presented a comprehensive study of the performance of such features for paintings classification. In that work the genre of the painting was identified as a result of recognizing the painter. Sablatnig et. al. [11] uses brush-strokes patterns to define structural signature to identify the artist style. Khan et. al. [12] uses a BoW approach with low–level features of color and shades to identify the painter among eight different artists. In [13] and [24] also similar experiments with low-level features were conducted. Unlike most of the previous work that focused on inferring the artist from the painting, our goal is to directly recognize the genre of the painting, which is a more challenging task.

3. Models Summary

In this section we present details of the 3 models that we used in this study. As shown in **Figure 2** these 3 models differ in terms of the classification methodology as well as the type of features used to represent the painting.

3.1. Discriminative Semantic-level model

In this approach a discriminative model (SVM [4] in our paper) is employed on top of semantic-level features. We used the Classeme descriptors [3], As described in [3], Classeme consists of the output of various pre-defined basic classifiers trained on a large number of object categories (2659 categories). These classifiers use a large number of low-level features through multi-kernel learning. The Classifier functions of these pre-trained classifiers are then utilized to produce a feature vector for a given image (dimension 2659). We used such feature vectors to train an SVM classifier for each painting genre. We hypothesize that Classeme features are suitable for representing and summarizing the overall contents of a painting since it captures semantic-level information about object presence in a painting encoded implicitly in the output of the pre-trained classifiers.

3.2. Discriminative Bag-of-Words model

BoW[9] is a very popular model in text categorization used to represent documents where order of the words does not matter. BoW was successfully adapted for object categorization [15]. Typical application of BoW on an image involves several steps which includes - 1) locating interest points in an image, 2) representation of such points/regions using descriptors, 3) Codebook formation using K-Means clustering, 4) Quantization of image using Codebook clusters and 5) representation of an image using histogram of these words. Thus the end result of Bag of Words model is a histogram of words which is used as an intermediate level feature to represent a painting. In our study we applied an SVM classifier on a code-book trained on images from our dataset. We used both Color SIFT (CSIFT) [1] and opponent SIFT (OSIFT) [2] as local features.

3.3. Generative Bag-of-Words Topic model

Generative topic model uses Latent Dirichlet Allocation (LDA [8]). In studies [6] and [7], LDA and Probabilistic Latent Semantic Analysis (pLSA) topic models have been applied for object categorization, localization and scene categorization. This paper is the first evaluation of such models in the domain of fine-art categorization.

Preprint submitted to 21st International Conference on Pattern Recognition.
Received April 7, 2012.

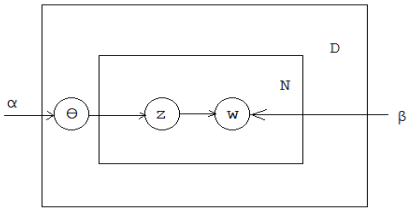

Figure 3 Graphical Model for Image Generation

For the purpose of our study, we used Latent Dirichlet Allocation (LDA [8]) topic model and applies it on BoW representation of paintings using both CSIFT and OSIFT features. In LDA, each item is represented by finite mixture over a set of topics and each topic is characterized by a distribution over words. **Figure 3** shows a graphical model the image generation process. As shown in the model, parameter Θ defines the topic distribution for each document (Total number of documents being D). Θ is determined by Dirichlet parameter α. β represents the word distribution for each topic. The total number of words is N.

To use LDA for classification task, we build model for each of the seven painting categories. First step is to represent each training image by a quantized vector using Bag of Words model described earlier. This vector quantized representation of each image is used for parameter estimation using Variational Inference. Thus, we will get LDA parameters Θ_c and β_c for each category. Once we have a new test image d we can infer the parameter Θ_{cd} for each category and $p(d|\ \Theta_{cd}, \beta_c)$ is used as the likelihood of the image to belong to a particular class c.

4. Data Set and Experiment Setup

Our dataset contains 7 categories of paintings namely Abstract, Baroque, Renaissance, Popart, Expressionism, Impressionism and Cubism. Each class consists of 70 paintings. The paintings are part of the Artchive fine-art dataset [16]. For each experiment 5-fold cross-validation is performed with 20% of the images (98) chosen for testing purpose in each fold.

For codebook formation Harris Laplace detector is used to find interest point. For efficient computation the number of interest points for each painting is restricted to 3000. Standard K-means Clustering algorithm is used to build a Codebook of size 600 words. SVM classifier is trained on both local-level and semantic-level descriptors. For SVM, Radial Basis function (RBF) is being used as Kernel. To determine parameters C and γ, grid search algorithm implemented by [4] is employed. Grid search algorithm uses cross-validation to pick up the optimum parameter values. Also this process is preceded by scaling of data set descriptors.

For experiments with LDA, David Beli's C-code [10] is used for the task of parameter estimation and inference. This C-code uses variational inference technique, which tries to estimate parameters β and Θ using similar simple model. For parameter estimation alpha is set to be 0.1 and LDA code is set to estimate the value of α during estimation process.

5. Experimental Results

We evaluated and tested the three models on our dateset and calculated and compared the classifying accuracy for each of them. **Table 1** shows the confusion matrix of the Discriminative Semantic Model over the five fold cross validation. The overall accuracy percentage achieved is 65.4% (64/98).

Table 1. Discriminative Semantic Model

ACCURACY(%)	Ba	Ab	Re	Po	Ex	Imp	Cub
Baroque	87.5	0	14.3	0	5.3	17.8	1.78
Abstract	0	64	0	7.1	7.1	1.8	1.9
Renaissance	5.4	0	64.3	5.35	14.3	3.5	0
Popart	0	1.78	1.8	73.1	0	3.5	1.8
Expressionism	1.8	20.2	7.1	3.6	48.2	17.8	12.9
Impressionism	5.36	8.	9	5.3	17.8	48.2	9.2
Cubism	0	6	3.5	5.3	7.1	7.1	72.4

Table 2 and **3** show the confusion matrices for the discriminative BoW model with CSIFT and OSIFT features respectively. Overall accuracy achieved is 48.47% (47/98) and 56.7% (55/98) respectively.

Table 4 and **5** shows the confusion matrices for the generative topic model using CSIFT and OSIFT features, with average accuracy of 49% (48/98) and 50.3% (49/98) respectively.

Table 6 summarizes the overall results for all the experiments. **Figure 4** shows the accuracies for classifying each genre using all the models evaluated.

Table 2. Discriminative BoW using CSIFT

ACCURACY(%)	Ba	Ab	Re	Po	Ex	Imp	Cub
Baroque	71.4	0	12.9	0	8.5	17.1	0
Abstract	0	48	5.8	10	8.5	5.7	7.1
Renaissance	18.6	6.7	41.4	0	5.8	9.3	18.5
Popart	0	15	0	70	11.5	9.3	15.7
Expressionism	0	15	18.6	2.8	28.5	12.9	13
Impressionism	8.5	8.6	3.7	8.6	17.2	45.7	11.4
Cubism	1.5	6.7	17.6	8.6	20	0	34.3

Table 3. Discriminative BoW using OSIFT

ACCURACY(%)	Ba	Ab	Re	Po	Ex	Imp	Cub
Baroque	82.1	0	10.7	0	14.3	17.9	3.6
Abstract	0	54.2	3.6	7.1	7.1	3.6	7.1
Renaissance	3.6	0	64.3	3.6	21	0	7.1
Popart	0	12.5	3.6	75	0	0	17.9
Expressionism	0	16.7	0	3.6	36	10.7	28.6
Impressionism	14.3	8.33	7.2	3.6	10.7	57.1	7.1
Cubism	0	4.2	10.8	7.1	14.3	10.7	28.6

Preprint submitted to 21st International Conference on Pattern Recognition.
Received April 7, 2012.

Table 4. Generative BoW topic model using CSIFT

ACCURACY(%)	Ba	Ab	Re	Po	Ex	Imp	Cub
Baroque	86.6	0	14.3	0	14.3	7.1	7.1
Abstract	0	58.3	7.1	26.6	0	7.1	14.3
Renaissance	6.6	8.3	42.8	20	14.3	0	7.1
Popart	0	0	7.1	13.3	0	0	14.3
Expressionism	0	8.3	7.1	6.6	36	14.3	7.1
Impressionism	6.6	25	14.3	13.3	21.4	71.4	14.3
Cubism	0	0	7.1	20	14.3	0	35.7

.

Table 5. Generative BoW topic model using OSIFT

ACCURACY(%)	Ba	Ab	Re	Po	Ex	Imp	Cub
Baroque	75.5	0	14.3	0	3.6	10.7	7.1
Abstract	0	62.5	3.5	27.3	3.6	3.6	0
Renaissance	7.1	4.2	39.2	3.3	7.1	3.6	10.7
Popart	0	8.3	0	28	3.6	0	7.1
Expressionism	7.1	0	17.8	14	36	3.6	10.7
Impressionism	10.2	25	10.7	10.2	32	68	21.4
Cubism	0	0	14.3	16.9	14.3	10.7	42.9

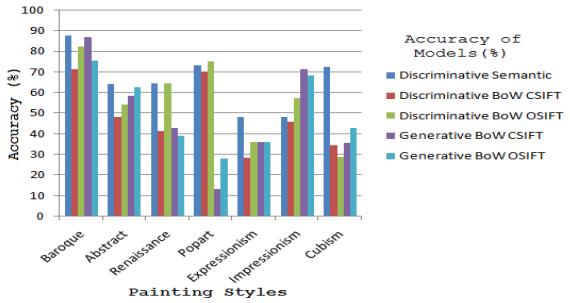

Figure 4 Classification Accuracy for each approach on each genre.

6. Discussion

As can be seen from the results the Discriminative model with Semantic-level features achieved the highest accuracy followed by Discriminative BoW with OSIFT, Generative BoW with OSIFT, Generative BoW with CSIFT and Discriminative BoW CSIFT.

These results are inline with our hypothesis that Semantic-level information would be more suitable for the task of fine-art genre classification. Both the discriminative and generative models gave comparable results. The OSIFT features outperformed the CSIFT features in the discriminative case, however the difference is not significant in the generative case.

By examining the results we can notice that the Baroque style is always classified with the highest accuracy in all techniques. It is also interesting to notice that the Popart genre is classified with accuracy over 70% in all the discriminative approaches while the generative approach performed poorly in that genre.

Table 6. Summary of Classification Results

Overall Summarized Results				
Discriminative Semantic Model	Discriminative BoW CSIFT	Discriminative BoW OSIFT	Generative BoW CSIFT	Generative BoW OSIFT
65.4%	48.47%	56.7%	49%	50.3%
Std 4.8%	Std 2.45%	Std 3.26%	Std 2.43%	Std 2.46

References

[1] A. E. Abdel-Hakim and A. A. Farag, “CSIFT: A SIFT descriptor with color invariant characteristics,” in IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, 2006, pp. 1978–1983.

[2] Koen E. A. van de Sande, Theo Gevers and Cees G. M. Snoek, Evaluating Color Descriptors for Object and Scene Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence

[3] Lorenzo Torresani, Martin Szummer, Andrew Fitzgibbon. Efficient Object Category Recognition Using Classemes. European Conference on Computer Vision, 2010

[4] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[5] Thomas Edward Lombardi, The classification of style in fine-art painting (January 1, 2005). ETD Collection for Pace University. Paper AAI3189084.

[6] L. Fei-Fei and P. Perona. A Bayesian Hierarchical Model for Learning Natural Scene Categories.IEEE Comp. Vis. Patt. Recog. 2005

[7] Josef Sivic, Bryan C. Russell, Alexei A. Efros, Andrew Zisserman, William T. Freeman “Discovering Objects and their location in images” In ICCV 2005.

[8] D. Blei, A. Ng, and M. Jordan. Latent dirichlet allocation. Journalof Machine Learning Research, 3:993–1022, 2003.

[9] Two bag-of-words classifiers ICCV 2005 short courses on Recognizing and Learning Object Categories .

[10] David Blei LDA C-code available at http://www.cs.princeton.edu/~blei/lda-c/

[11] R. Sablatnig, P. Kammerer, and E. Zolda. Hierarchical Classification of Paintings Using Face- and Brush Stroke Models. ICPR 1998

[12] Fahad Shahbaz Khan, Joost van de Weijer, & Maria Vanrell. (2010). Who Painted this Painting?

[13] T. Melzer, P. Kammerer, and E. Zolda. Stroke Detection of Brush Strokes in Portrait Miniatures using a Semi-Parametric and a Model Based Approach

[14] I. Widjaja, W. Leow, and F. Wu. Identifying Painters from Color Profiles of Skin Patches in Painting Images.

[15] G. Csurka, C. Dance, L.X. Fan, J. Willamowski, and C. Bray (2004). "Visual categorization with bags of keypoints". Proc. of ECCV International Workshop on Statistical Learning in Computer Vision.

[16] http://www.artchive.com/