Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 369–376

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Painter Classification Over the Novel Art Painting Data Set via The Latest Deep Neural Networks

Mehmet Oguz Kelek, Nurullah Calik*[1], Tulay Yildirim

Abstract

Painters can be affected during their life simultaneously from different movements, the same movements or a few different movements. This situation makes the problem of identifying, and classifying painters more difficult. In this paper, we have tested the latest deep neural networks on this problem. There are 17 painters who lived in different term and influenced by different art movements, an average of 46 paintings per painters in our data set to test this problem. GoogleNet, DenseNet, ResNet50, ResNet101 and Inceptionv3 networks are applied to this data set. Although DenseNet gives the highest result, considering the cost parameters such as training time and file size, the Inceptionv3 and ResNet50 which provide near to DenseNet results is the optimum networks.

Keywords: *Convolutional Neural Network, Art Painter Classification, Deep Neural Networks, Painting*

1. Introduction

In recent years, the transfer of the art work to digital environment has accelerated, so large libraries have begun to be created on the internet. These libraries brought together paintings from museums, churches or collectors in different place of the world to a wide range of people, such as educators, curious painting viewers and art students. However, museums, churches and collectors have hundreds of thousands of paintings, and digitization is not an easy task so a painter classifier will become more purposive and practical, for museum curators, curious painting viewers and art students. We hope that creating an painter classifier allows curators, educators, curious painting viewers and art students to

[1] Corresponding author. Tel.: +90 212 383 59 10
E-mail address: ncalik@yildiz.edu.tr

10.1016/j.procs.2019.06.053

automatically tag objects and allow visitors to browse paintings more freely.

Convolutional Neural Networks(CNN) is a mixture of biology and computer science. CNN is a very effective mechanism used for image recognition and classification. In order to classify or define the images given, CNN uses features that distinguish between living things or things in the picture. For instance, when we look at the picture of a dog, we can define the dog by examining its characteristics such as head structure, nose structure, ear structure. CNN also performs the same sort of classification or recognition by detecting lower level features such as curves and edges. If the number of objects in a image or video is small, CNN may not be used to perform feature extraction, recognition and classification. However, it is inevitable to use CNN if the number of objects in the image or video is hundreds or more.

1.1. Related work

In study1, a variety of models ranging from a simple CNN designed from scratch to ResNet-18 network with transfer learning had trained. The data set includes of 300 paintings per artist from 57 well-known artists. The network achieves pretty higher accuracy than previous studies. The results show that CNNs are the most accurate tool for painters identification with the study.

In the another work2, evaluated deep CNN is up to 19 weight layers for large- scale image classification. Results were showed that the representation depth is useful for the accuracy of classification, and that state-of-the-art performance on challenge data set can be achieved using a CNN architecture with quite increased depth. The main contribution of this study to the literature was to evaluate and improve the networks of increased depth using 3×3 convolution filters.

In addition to other works, a data set which contain total of 4266 paintings and 91 painters was investigated. This study aimed to demonstrate three applications of the data set: artist categorization, style classification, and clarity detection. The results showed that latest art computerized vision methods can classify 50% of invisible paintings to the painter in data set and can relate their artistic style to more than 60%of cases correctly in the study3.

It is important to accurately identify the artistic style of the paintings and to folder large artistic databases. In the study4, the existing approaches to data set for 25 different styles (AlexNet, ResNet) were used to solve the problem of detecting the artistic style of a picture. To obtain the results, the network is first pre-trained on ImageNet, and deeply retrained for artistic style. The results of styles are between 89% and 99%. With the increase of the demands of digital contemporary painting collections, automatic theme stylization has grown in academic and commercial area. The latest and significant in deep neural networks has provided efficient visual features that achieve state-of-the-art results in various visual classification tasks.

In study5, the perceptiveness of these features in identifying artistic styles in paintings and suggest a compact binary representation of the paintings were examined. Integrated with the PiCodes descriptors, these features demonstrate high classification outcome on a large scale collection of paintings.

In addition to previous works, each art painting is distorted by various operations to simulate various case that it would show up on photographs or videos in the study6. The distortions are implemented by computer to create large data sets for CNNs to train. The method was compared with the Scale-Invariant Feature Transform (SIFT) and the result demonstrated the proposed method outperform SIFT by 13.6% as test error rate.

1.2.Paper organization

This paper includes four sections: In Section II, detailed information is given about the our novel data set and utilize latest CNNs which are GoogleNet, DenseNet, Resnet50, Resnet101, Inceptionv3. The performance results of the CNNs were explained, compared to these outputs in Section III. All results are interpreted and the results were summarized in Section IV.

2. Material and Methods

2.1 Data set

Art painters are influenced by various art movements throughout their lives and they paint according to these influences, which makes it difficult to classify painters. Our primary contribution in this study is to create a new and challenging set of data by selecting works of painters in different styles. The paintings are collected from the various internet websites. There are 17 painters who lived in different term and influenced by different art movements, an average of 46 paintings per painters and a total of 797 paintings in our data set. This situation makes the problem of painters classification challenging.

Fig. 1. A total of six painters were selected randomly from our data set consisting of 17 different painters. two work of these painters were selected (upper and bottom). Class distributions are (a) Edvard Munch, (b) Caspar David Friedrich, (c) Frida Kahlo, (d) Pablo Picasso, (e) Vincent Van Gogh, and (f) Caravaggio

Our data set includes 17 artists from 34 to 63 samples from each painter. Our data set includes paintings which each painter has of different objects, human bodies, human faces and creatures of nature. Some sample paintings from the data set are shown in Fig.1. The artist with 34 paintings is Johannes Vermeer. The artist, who has 63 paintings, is Gustav Klimt. In addition to these two artists, there are artists like Caravaggio (45) , Caspar David Friedrich (43), Diego Velázquez (44), Edvard Munch (46), Frida Kahlo (43), Paul Cézanne (43), Pablo Picasso (46), Rembrandt Harmenszoon Van Rijn (44), Vincent Van Gogh (50), Albrecht Dürer (49), Claude Monet (50), Édouard Manet (48), Michelangelo Buonarroti (39), Paul Gauguin (55), Raffaello Sanzio (55). Parentheses next to painters show how many paintings are found in painter classes in the novel data set.

There are art movement or movements represented by artists in our data set. For instance, the main representatives of the movement of classicism; Leonardo da Vinci, Michelangelo Buonarroti and Raffaello, the main representatives of the Baroque movement; Caravaggio, Rubens, Rembrandt and Valezquez, the main representatives of the impressionism movement; Claude Monet, Auguste Renoir, Vincent van Gogh, Cezzanne, Toulouse Leatrec, Sisley, Camille Pissarro, the main representatives of the Cubism movement; Pablo Picasso and Georges Braque are the main representatives of the surrealism movement; Vincent van Gogh, the principal representative of the symbolism movement; Paul Gauguin.These movements and painters show us how difficult our novel data set is.

2.2 Applied deep learning networks

Pre-trained models are often used in classification problems in the studies1,3,7,8,9. This is because pre-trained networks are trained in large scale data sets, so the filter coefficients obtained are highly descriptive. For this reason, pre-trained network is taken for different problems, and fully connected (FC), softmax, classifier layers are discarded. Instead of these layers, the FC layer, which has neurons as new problem class number N_c, and softmaxlog loss layer (1) are added up to network. p_c, and N_{batch} refer probabilistic result of network corresponding with class c, and mini-batch size for training, respectively. Then, the new designed network is trained for the new data set. In this study, GoogleNet, Inceptionv3, ResNet50, ResNet101, DenseNet networks which have received a degree in ImageNet competition have been used.

$$p_c = \frac{e^{x_c}}{\sum_{i=1}^{N_c} e^{x_i}} \qquad\qquad L_{batch} = -\sum_{i=1}^{N_{batch}} log(p_i) \qquad (1)$$

2.2.1. Googlenet

In 2014, the winner of the ILSVRC 2014 competition was GoogleNet. It reached a top-5 error rate of 6.67%. This accuracy was very close to human level performance and After training, Andrej Karpathy was able to achieve a top-5 error rate of 5.1% in single model and 3.6% in ensemble. GoogleNet used a CNN inspired by LeNet, however applied a novel element which is entitled an inception module. GoogleNet's architecture occured a 22 layer deep CNN but reduced the number of parameters from 60 million to 4 million in the work10.

2.2.2. ResNet50 and ResNet101

ResNet is abbreviation for Residual Network. Deep CNNs have made a breakthrough for image classification and visual identification tasks have also benefited from very deep models. There is a trend to go more deeper, to solve more complex tasks and to improve the classification accuracy. However, as we go deeper; the training of the neural network becomes more difficult and accuracy start to feed and then deteriorates. Residual learning tries to solve these problems by using shortcut mechanism. ResNet50, ResNet101 have 50 and 101 layer, respectively in the study11.

2.2.2. Densenet

CNNs with K layers have K connections, while between each layer and its subsequent layer - our network has $K \times (K+1)/2$ direct connections. For each layer, the property maps of all previous layers are used as input and their property maps are used as input to all subsequent layers. In the study12, DenseNet has several stronger advantages: it attenuates vanishing-gradient problem, strengthens feature propagation, substantially reduce the number of parameters and countenance feature reuse. Unfortunately, training time takes a long time.

2.2.2. Inceptionv3

Inceptionv3 is a network model that has been made by improvement Inceptionv1 and Inceptionv2. The difference of Inceptionv3 from Inceptionv2 is that it integrates the FC layer and batch-normalized as auxiliary classifier as well as convolutional layers in those studies13,14.

3. Experimental Results

In this section, the data set is divided into different training and testing ratios and the performance metrics of the networks are evaluated. Since there are different paintings of different painters, the data set is imbalanced. Thus, we have calculated the macro F-measure value in addition to the accuracy (2). In this way, a robust evaluation protocol is prepared for networks.

$$Acc_i = \frac{\sum_i \mathbf{M}_{ii}}{\sum_{i,j} \mathbf{M}_{ij}} \tag{2}$$

$M \in \Re^{17\times17}$ is the confusion matrix. Class number is 17. To procure the Macro F-Measure, first of all precision (pre_i) and recall (rcl_i) values are calculated for ith class in (3) for multi-class sets. After calculating the F-Measure of all classes, F_i values are averaged (4).

$$pre_i = \frac{\mathbf{M}_{ii}}{\sum_i \mathbf{M}_{ij}} \qquad rcl_i = \frac{\mathbf{M}_{ii}}{\sum_j \mathbf{M}_{ij}} \tag{3}$$

$$F_i = \frac{2 \times pre_i \times rcl_i}{pre_i + rcl_i} \qquad F\text{-}measure = \frac{1}{N_c} \sum_{i=1}^{N_c} F_i \tag{4}$$

Networks are trained in the same hyper parameter environment. These values are given in Table 1. Stochastic gradient descent (SGD) was chosen as an optimizer. The initial learning rate: 10^{-3}, the maximum epoch size: 20, the mini batch size: 20 and the learning rate drop factor: 0.8 are defined. The graphics card used to implement these tests is NVIDIA GeForce GTX 1080 Ti, CPU speed is 2.6 GHz. MATLAB program was used for simulation.

Table 1. Hyperparameter for Training Process of Networks

Optimizer	SGD
Momentum	0.9
Init. Learning Rate	0.001
Batch Size	20
Epoch	20
Init. Weigths	pre-trained

The data set was firstly allocated as 90% training and 10% tests randomly. Achievement of all networks in this test environment was calculated. After this process is done 5 times, average performance of the networks and standard deviation are evaluated.

Table 2. Performance Metrics and Specificity of Networks: 90%Training 10% Test Ratios

Deep Nets	Size	Parameter	Depth	Input Size	Train Time [min]	Test Time[sec]	Accuracy (%)	F-Measure (%)
GoogleNet	27 MB	7.0 M	22	224×224	3.72	0.34	69.90 ± 3.13	69.46 ± 3.31
Inceptionv3	89 MB	23.9 M	48	299×299	18.36	0.94	79.26 ± 2.65	79.81 ± 3.22
Resnet50	96 MB	25.6 M	50	224×224	9.57	0.58	76.48 ± 5.46	74.48 ± 5.34
Resnet101	167 MB	44.6 M	101	224×224	21.92	0.72	78.43 ± 3.84	77.02 ± 4.15
DenseNet	77 MB	20 M	201	224×224	59.48	2.07	79.36 ± 7.18	78.49 ± 7.05

The results in Table 2 show the specifications and performances of the networks that worked under 90%training and 10%test rate. According to these results; the lowest accuracy is provided by GoogleNet network and the highest accuracy is DenseNet.

Despite it's high accuracy, training and test time is much longer than other networks. It has also the highest standard deviation 7.18. Accuracy results close to DenseNet network have also given Inceptionv3, Resnet50 and ResNet101 networks. Considering parameters such as training time, test time, and size, DenseNet network is a costly network. When we examine the results of Inceptionv3 and ResNet50 based on the requirements of DenseNet and ResNet101, the optimal results in terms of size and training time parameters, and the second best result in terms of accuracy and test time parameters. Under these the circumstances, the optimal networks for our data set is Inceptionv3 and ResNet50 network. The comparison of the training speeds of the networks is shown in Figure 1 We tested them in the selected iterations. It is important that after the 200th iteration, the training of networks becomes stationary.

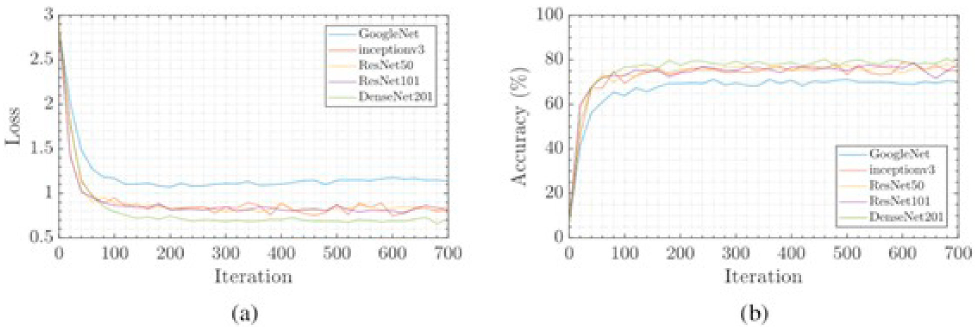


Fig. 2. Training speeds of networks,(a): Loss Function,(b): Accuracy: The networks are tested in selected iterations. Their learning speeds are significant. At the end of training process, the highest accuracy and the lowest loss are achieved with DenseNet. The results are obtained under 90% training and 10% test ratio.

Table 3.Performance Criteria of Networks at Different Test Ratios

	%10	%10	%50	%50	%90	%90
Deep Nets	ACC	F1	ACC	F1	ACC	F1
GoogleNet	69.90 ± 3.13	69.46 ± 3.31	63.71 ± 4.12	63.21 ± 4.00	45.21 ± 1.48	44.66 ± 1.47
Inceptionv3	79.26 ± 2.65	79.81 ± 3.22	70.51 ± 3.12	70.46 ± 3.01	47.55 ± 2.59	47.26 ± 2.30
Resnet50	76.48 ± 5.46	74.48 ± 5.34	76.30 ± 2.79	75.99 ± 2.68	56.34 ± 2.10	56.56 ± 2.11
Resnet101	78.43 ± 3.84	77.02 ± 4.15	73.61 ± 3.01	73.17 ± 2.70	56.52 ± 2.11	56.68 ± 2.49
DenseNet	79.36 ± 7.18	78.49 ± 7.05	77.69 ± 3.49	77.39 ± 3.44	59.16 ± 1.56	59.12 ± 1.53

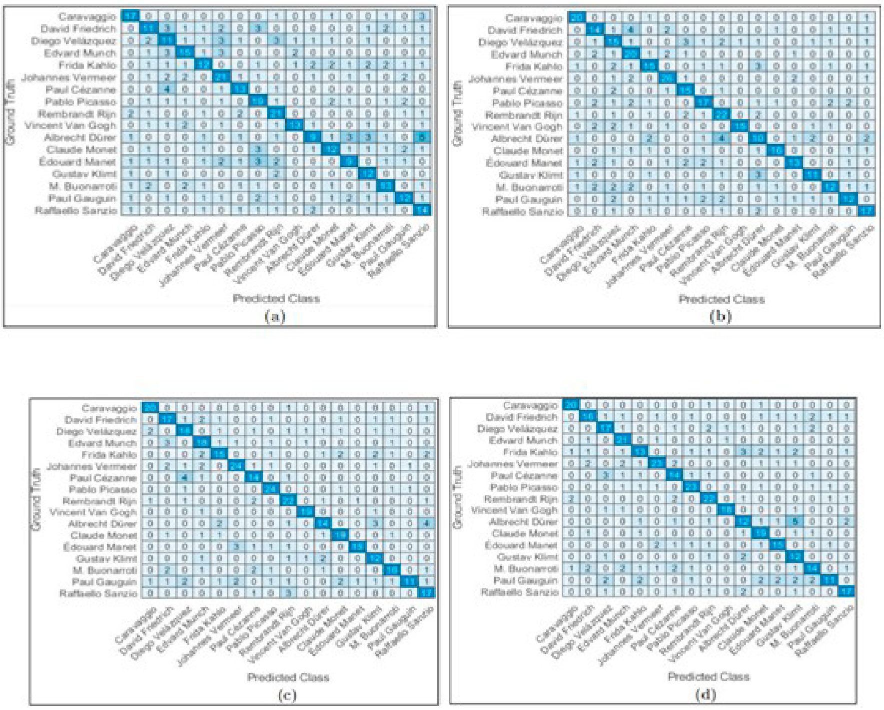
We did not want to examine the CNN networks that used under 90% training and 10% test rate. How the CNN networks used under different training and testing rates would give results is wondered and our networks are examined in three different scenarios 10%,50%, and 90% test ratios. For the scenarios results are shown in Table 3. The decreasing of networks' performances is stand out as the testing rates increasing. When 50%, and 90% test rate are dealt, there is a serious decrease in Inceptionv3 and GoogleNet in comparison to ResNet50, ResNet101 and DenseNet.

Under the 50% train and test rate, confusion matrices of each network are shown in Figure 2. The confusion matrices show which painter's paintings are confused other paintings of the painters. Mostly confused painters are Albrecht Dürer and Paul Cézanne in confusion matrix of ResNet50, ResNet101, DenseNet ve GoogleNet. The most confused painters in the confusion matrix of Inceptionv3 are Albrecht Dürer, Diego Velázquez, Gustav Klimt and Frida Kahlo. Albrecht Dürer is confused with artists Gustav Klimt, Raffaello Sanzio and Rembrandt Harmenszoon Van Rijn in different networks. Another quite confused painter, Paul Cézanne, is confused with Diego Velázquez.

4. Conclusions

In this paper, we tested CNNs performances in a challenging data set which painters from different art movements were involved, painters affected by the same art movements, and painters affected by different art movements.When we look at the results obtained, the network giving the highest accuracy was DenseNet. However, when we look at the duration of the training period, we see that the cost is high. When we look at the training time of Inceptionv3 network which is very close with DenseNet accuracy as a 79.26%, we see that it is plenty short compared to DenseNet training time. When the cost constraints are taken into account, the network provides that the most optimum result is Inceptionv3 in tested CNNs.

As a future work, we plan to design a novel network for this topic.We will design a network that shows how the artists are influenced by movements and how much are affected by movements. We will also design a network that calculates how much the painters of the same period or different periods are affected. This application will be useful for art lovers.



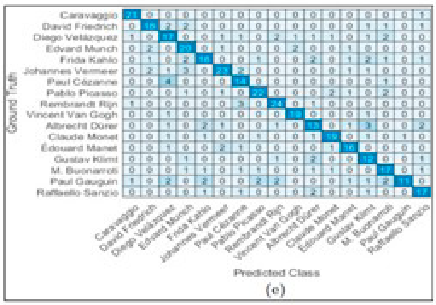


Fig. 3.The confusion matrix of the networks: (a) GoogleNet,(b) Inceptionv3,(c) ResNet50, (d) ResNet101 and (e) DenseNet. Albrecht Dürer's paintings are the most confused painter in all networks except DenseNet. Paul Cézanne's paintings are riddling for DenseNet.

5. References

1. N. Viswanathan, Artist identification with convolutional neural networks, Standford193CS231N Report.
2. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.
3. S. . W. J. . F. M. Khan, Fahad Beigpour, Painting-91: A large scale database for computational painting categorization, Machine Vision and Applications (2014) 1385–1981397.
4. A. Lecoutre, B. N´egrevergne, F. Yger, Recognizing art style automatically in painting with deep learning, in: ACML, 2017.
5. Y. Bar, N. Levy, L. Wolf, Classification of artistic styles using binarized features derived from a deep neural network, in: ECCV Workshops, 2014.
6. Y. Hong, J. Kim, Art painting identification using convolutional neural network, International Journal of Applied Engineering Research 12 (4) (2017) 532–539.
7. B. Saleh, A. M. Elgammal, Large-scale classification of fine-art paintings: Learning the right metric on the right feature, CoRR abs/1505.00855. arXiv:1505.00855.207URLhttp://arxiv.org/abs/1505.00855
8. T. E. Lombardi, The classification of style in fine-art painting, Ph.D. thesis, New York, NY, USA (2005).
9. S. Banerji, A. Sinha, Painting classification using a pre-trained convolutional neural network, in: International Conference on Computer Vision, Graphics, and Image processing, Springer, 2016, pp. 168–179.
10. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Computer Vision and Pattern Recognition (CVPR), 2015. URL : http://arxiv.org/abs/1409.4842
11. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016) 770–219778.
12. G. Huang, Z. Liu, L. v. d. Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2261–2269. doi:10.1109/CVPR.2017.
13. X. Xia, C. Xu, B. Nan, Inception-v3 for flower classification, in: 2017 2nd International Conference on Image, Vision and Computing (ICIVC), 2017, pp. 783–787. doi:10.1109/ICIVC.2017.7984661.
14. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, CoRR abs/1512.00567. arXiv:1512.00567. URL : http://arxiv.org/abs/1512.00567