Paintings-100: A Diverse Painting Dataset for Large Scale Classification

Erica M. Knizhnik*†, Brian Rivera*‡, Atreyee Sinha§, and Sugata Banerji*¶
*Department of Mathematics and Computer Science, Lake Forest College, Lake Forest, IL, USA.
†Email: knizhnikem@lakeforest.edu
‡Email: riverab@lakeforest.edu
§Computing and Information Sciences, Edgewood College, Madison, WI, USA. Email: asinha@edgewood.edu
¶Corresponding author. Email: banerji@lakeforest.edu

***Abstract*—Painting classification is an interesting cross-disciplinary research problem in computer vision. With the increased accessibility of digitized collections of fine-art paintings, development of effective painting classification algorithms has become vital as they have many potential applications in museums, various industries, painting theft investigation, forgery detection, art education, etc. However, the availability of large scale annotated benchmark datasets with high-resolution authentic painting images still remains a challenge. Towards that end, in this work, we develop an image dataset consisting of high-resolution painting images from 100 different artists spanning 14 different styles. This dataset is an extension of the Painting-91 dataset constructed by Khan et al. Our contribution towards extending this dataset are threefold. First, we address the limitations of the dataset by removing errors and enhancing image resolutions. Second, we add more images to augment some of the artist categories with fewer images. Third, we include the works of nine more painters from diverse backgrounds and styles for creating a more representative and inclusive database of fine-art paintings. We also perform a preliminary evaluation of this newly constructed Paintings-100 dataset using several different Convolutional Neural Network (CNN)-based classification techniques for artist recognition. Furthermore, we demonstrate that our proposed and improved dataset is more suitable for patch-based models than the earlier published Painting-91 dataset due to larger image resolutions.**

Keywords*—*painting classification; image dataset; CNN

I. Introduction and Background

Due to the substantial digitization of artworks in recent years, along with the impressive development in the area of computer vision, automated painting classification is an interesting and crucial research problem. Fine-art painting classification, which includes identifying the artist and the artistic style from a painting, has many applications in museums, various industries, painting theft investigation, forgery detection, art education, etc. However, this is not a trivial task due to the complexities of artistic styles, subjectivity in the interpretation of paintings, varied image quality, lack of fine details, and context of the visual images [1] [2].

In the last decade, several studies have focused on employing computer vision techniques to analyze paintings and other forms of visual art [3]–[6]. In our previous work [7], we explored the use of pre-trained Convolutional Neural Network (CNN) models as a feature extraction tool for painting classification. Some of the popular painting datasets that are available publicly for artist and style classification include the Painting-91 dataset by [8], the WikiArt dataset [9], and the Painting dataset consisting of ten classes of fine-art paintings from the PASCAL VOC [10].

Figure 1. Some errors that exist in the old Painting-91 dataset.

While CNNs have proven to be versatile and effective in various image-related tasks, one of the major challenges of being able to effectively use CNNs for painting classification is the need for large hand-labeled datasets [2]. To that end, in this paper, we have worked on improving the existing Painting-91 dataset [8] to not only include newer artists and painting styles, but also carefully remove different mis-attribution and other human errors that existed in that dataset. Some of these errors are shown in Figure 1. Our newly constructed dataset, called the Paintings-100 dataset, also has enhanced image resolutions than the previous dataset. We have augmented certain artist categories, which had fewer images in the previous dataset,

Figure 2. Some paintings of the newly added nine artists that are included in the Paintings-100 dataset.

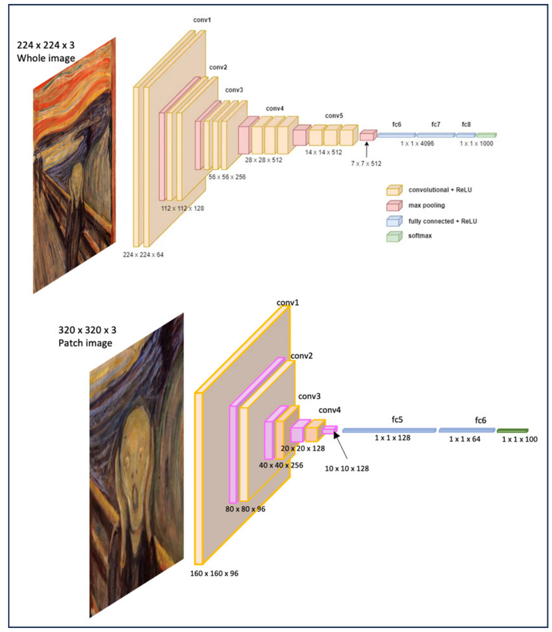

Figure 3. For artist classification task, we used both whole images as well as random patches from the images to feed into different CNN models.

with newer images. Last, but not the least, we have also done a preliminary evaluation of this dataset using several CNN based methods on both whole images as well as random image patches for the artist classification task.

The rest of this paper is organized as follows. Section II and its subsections outline in detail the dataset constructed in this work. The techniques used to evaluate the classification performance on this dataset, the experiments performed, and results obtained are described in detail in Section III. Finally, we list our conclusions and directions for future research in Section IV.

II. Dataset Construction

When we worked on [7], we realized that the images in the original Painting-91 dataset [8] are too small for learning meaningful features using deep learning. While trying to replace the images with their high-resolution versions, we found several kinds of human errors and other limitations in the original dataset which needed to be fixed. These issues, some of which are shown in Figure 1, are as follows:

A. Low resolution

This was the main motivation for this work. The mean size of an image in this dataset is 268×263 pixels. These dimensions are smaller than the input sizes of many CNN models. So, to improve the quality of the data, we started replacing the images with high-resolution versions downloaded from the Internet via Google Reverse Image Search [11]. We were successful in this task for about 97% of the images, but we also ran into other errors as detailed next.

B. Mis-attributions

These are images labeled with a painter's name that are not painted by that painter. Some of these mis-attributed images are deliberate attempts to copy the attributed painter's style, some are created using image editing software by making collages of existing paintings, and some others have simply been downloaded from a source on the Internet which also had the wrong label.

C. Duplicates

Several of the images in most artist classes are duplicates of other images also in the class. The number of images per class varies from 30 to 51, which is already very small for training deep learning models, and the presence of duplicate and mislabeled images further reduces this number. For instance, the painter class Hieronymus Bosch has 50 paintings, out of which 25 are duplicates (exact or slightly variant copies), and a further 5 are wrongly attributed, thus bringing the actual number of usable images down to 20.

D. Cropped images

These are images which show only part of a painting, the whole of which may or may not be present in the dataset. Since the overall composition bears as much information about

TABLE I
New artists whose paintings were added to the dataset, along with their nationality and style.

Artist	Nationality	Style
Amrita Sher-Gil	Hungarian-Indian	Several
Jamini Roy	Indian	Indian folk art
Julie Mehretu	Ethiopian American	Several
Katsushika Hokusai	Japanese	Ukiyo-e
Kitagawa Utamaro	Japanese	Ukiyo-e
Rafiy Okefolahan	Cape Verdean	Contemporary multimedia
Raja Ravi Varma	Indian	Indian realism
Utagawa Hiroshige	Japanese	Ukiyo-e
Zhang Xiaogang	Chinese	Surrealism

a painter's identity as details do, just having a small cropped portion of a painting in the dataset is not ideal.

E. Color variations

These are also copies of other images in the dataset. However, instead of being exact duplicates, these images have a different color palette. There is no way of knowing which of the copies has a more accurate color palette, and so, color cues lose their significance in classification. To further complicate matters, some artists (such as Andy Warhol) themselves produced multiple copies of the same painting with slight differences in details and color, which count as different images in the dataset.

F. Lack of diversity

While the original dataset contains an impressive collection of works from 91 painters and 13 style categories, this collection focuses exclusively on Europe and the Americas. There are no painters representing the rich artistic heritage of Asia and Africa. This is not exactly an error, but an omission in the dataset that needed to be addressed for overall improvement.

Improvements

We took several steps to address the above issues. First, we replaced most images with their high-resolution versions wherever such a version was available in the public domain. The mean image size in the new dataset is $1,523 \times 1,493$ pixels. This amounts to a 32-fold increase in the number of pixels per image, on average. Second, we replaced wrongly labeled images with their correct counterparts, or new images by the same artist. Third, wherever possible, we also added new paintings to all artist categories that had less than 50 paintings. Fourth, we reduced the number of duplicates by replacing them with new paintings wherever possible. Last, but not the least, we added 50 paintings each by 9 more painters spanning a diverse array of styles representing Asian and African art (shown in Figure 2 and Table I). This makes our new Paintings-100 dataset a more diverse, inclusive and representative database of fine-art paintings. The presented Paintings-100 dataset has $5,357$ images which is an impressive 25% increase from the $4,266$ image Painting-91 dataset.

III. Experiments and Discussion

The original Painting-91 dataset, and by extension, the proposed Paintings-100 dataset, are both designed for two classification tasks. These tasks are artist classification and style classification. The first task is straightforward as every image has an artist class label, and the artist classes are roughly equal in size. For the second task, the dataset contains 14 style class labels in addition to the 100 artist class labels. This is a slight increase from 13 style classes in the Painting-91 dataset (Ukiyo-e is the new style class introduced). Each style class contains works from more than one artist, but not all artists have a style class label [8]. In the current work, we have only analyzed the dataset with respect to the first problem.

The artist classification problem is somewhat challenging for CNN-based models. This is mainly due to two reasons.

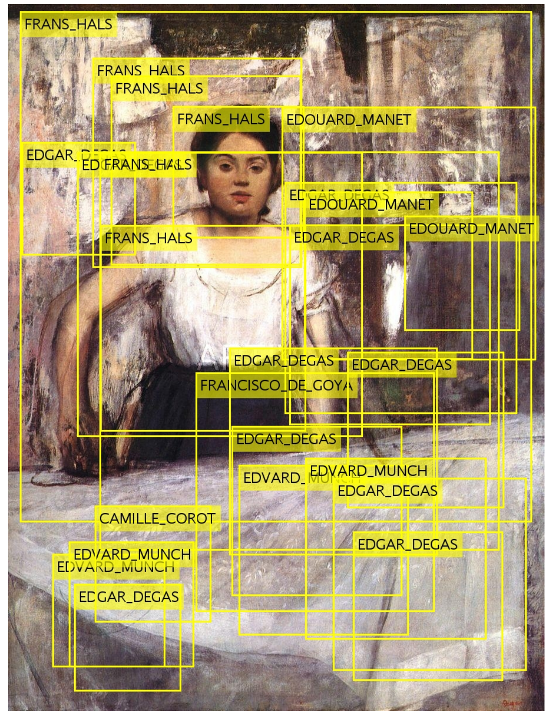

Figure 4. Painting by Edgar Degas. When the whole image is used for artist recognition, the CNN identified it as a Frans Hals painting, whereas by using random patches, it is correctly classified as an Edgar Degas artwork.

Firstly, deep learning is data-hungry, and very few artists manage to paint more than a few dozen completed paintings in their lifetime. This severely reduces the images available for training. Secondly, CNN models take fairly low-resolution images as their input. This means, we either need to downsample the images and lose all detail, or crop the images and lose all sense of composition and context. Since neither solution was fully acceptable to us, we decided to use an ensemble of multiple CNN models that use both downsampled whole images and full-size patches cropped out of the high-resolution images. For classifying the patches, we designed our own CNN from scratch and trained it using 25 random square patches from each training image. For the whole image classification, we fine-tuned the VGG-16 network [12] trained on the ImageNet image dataset [13]. These two models are shown in Figure 3. In both cases, we used 24 images per class with augmentation for training, 6 per class for validation, and the rest for testing. We used decision fusing based on the labels predicted by the two models. Our initial results were promising, with the patch-based model yielding a 32% accuracy on the test set, the whole image model yielding 33%, and the fused accuracy at 38%. The confusion matrix for this result is shown in Figure 5. Figure 4 illustrates the effectiveness of such a fusion. In this example, although the whole image classifier predicts the label to be Frans Hals, different patches vote for different labels and the true artist, Edgar Degas, gets the most votes.

We also did a visualization of the responses from the

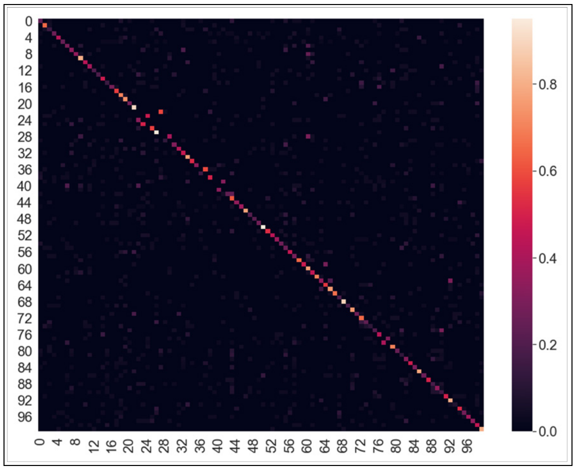

Figure 5. The confusion matrix for the artist classification experiment using the combined decision of two CNN models. The rows indicate actual class index values while the columns indicate predicted index values.

last convolutional layer of our patch-image classifying CNN using the Grad-CAM technique [14]. This ”heatmap” analysis highlights the regions of an image that are key identifiers for artist recognition. While this is a work in progress, the results demonstrated in Figure 6 show some of the characteristics of artists that the network can identify correctly. For instance, bold outlines are a signature characteristic of Jamini Roy and these outlines are highlighted in the topmost example in Figure 6. Similarly, dotted patterns and certain kinds of brush strokes are recognized as characteristic features of Roy Lichtenstein and Vincent Van Gogh, respectively.

IV. Conclusion and Future Work

In this work, we presented a large scale diverse high-resolution image dataset for artist and style classification. While this was based on the Painting-91 dataset, the improvements were significant enough for the Paintings-100 dataset to be considered a new dataset. Although we need to run many more experiments, initial results based on an ensemble of CNN models showed promising results for the artist classification task.

There are many different ideas that we would like to try out on this dataset in the near future. Currently, we select the patches randomly. In future, we want to try selecting patches with face detectors and object detectors to see how that affects our results. We plan to use some color normalization techniques to reduce the effect of photographic conditions on the paintings. Last, but not the least, we still need to run style classification experiments on this dataset and see the results.

References

[1] J. M. Silva, D. Pratas, R. Antunes, S. Matos, and A. J. Pinho, “Automatic analysis of artistic paintings using information-based measures,” Pattern Recognition, vol. 114, 2021, p. 107864.

[2] E. Cetinic, T. Lipic, and S. Grgic, “Fine-tuning convolutional neural networks for fine art classification,” Expert Systems with Applications, vol. 114, 2018, pp. 107–118.

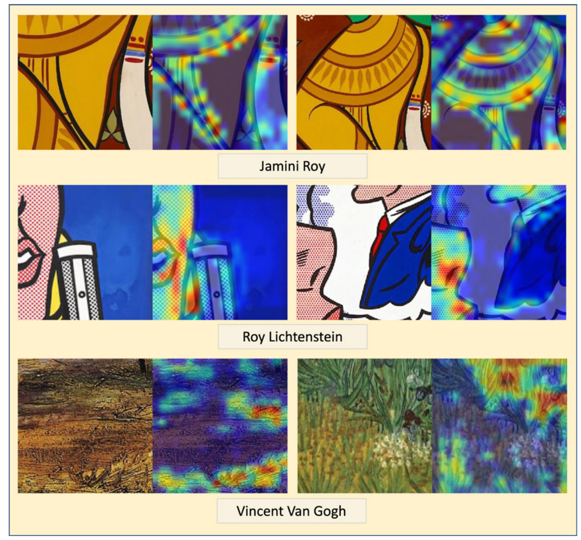

Figure 6. Some paintings and their heatmaps showing regions of interest for artist recognition as detected by the CNNs.

[3] R. S. Arora and A. Elgammal, “Towards automated classification of fine-art painting style: A comparative study,” in Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), 2012, pp. 3541–3544.

[4] C. S. Rodriguez, M. Lech, and E. Pirogova, “Classification of style in fine-art paintings using transfer learning and weighted image patches,” in 2018 12th International Conference on Signal Processing and Communication Systems (ICSPCS), 2018, pp. 1–7.

[5] S. Smirnov and A. Eguizabal, “Deep learning for object detection in fine-art paintings,” in 2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo), 2018, pp. 45–49.

[6] C. Sandoval, E. Pirogova, and M. Lech, “Two-stage deep learning approach to the classification of fine-art paintings,” IEEE Access, vol. 7, 2019, pp. 41 770–41 781.

[7] S. Banerji and A. Sinha, “Painting classification using a pre-trained convolutional neural network,” in Proceedings of the 2nd Workshop on Computer Vision Applications (WCVA) at the Tenth Indian Conference on Vision, Graphics and Image Processing (ICVGIP), 2016, pp. 168–179.

[8] F. S. Khan, S. Beigpour, J. V. de Weijer, and M. Felsberg, “Painting-91: A large scale database for computational painting categorization,” Machine and Vision Application (MVAP), vol. 25, no. 6, 2014, pp. 1385–1397.

[9] B. Saleh and A. Elgammal, “Large-scale classification of fine-art paintings: Learning the right metric on the right feature,” International Journal for Digital Art History, no. 2, Oct. 2016, pp. 70–93.

[10] E. J. Crowley and A. Zisserman, “In search of art,” in Workshop on Computer Vision for Art Analysis, ECCV, 2015, pp. 54–70.

[11] “Google reverse image search,” https://www.google.com/imghp, accessed: December 29, 2023.

[12] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, 2015.

[13] J. Deng et al., “Imagenet: A large-scale hierarchical image database,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.

[14] R. R. Selvaraju et al., “Grad-cam: Visual explanations from deep networks via gradient-based localization,” in 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626.