Exploring Deep Learning Techniques for Artist and Style Recognition in the Paintings-100 Dataset

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 a challenging interdisciplinary research problem in computer vision. With more fine-art paintings being available in the form of high-resolution digital scans, the development of effective classification algorithms has become vital. Such algorithms would have numerous applications, including but not limited to museum curation, several different industries, painting theft and forgery investigation, and art education. While some progress has been done in this field, accurately identifying the painter or the artistic style from the painting remains a complex task. Towards that end, we present an enhanced image dataset comprising high-resolution painting images from 100 diverse artists across 14 distinct styles. This dataset builds upon the Painting-91 dataset originally created by Khan et al. Our main contributions in this work are three-fold. First, we improve the older dataset by correcting errors, enhancing image resolution, and expanding it with more images, artists, and styles. Second, we perform an extensive evaluation of this newly constructed Paintings-100 dataset using several different convolutional neural network (CNN)-based classification techniques for both artist and style recognition tasks. Finally, we explore the different stylistic characteristics that the networks focus on to recognize the specific artists and styles of paintings, and 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; style classification; artist classification; CNN ensemble.

I. Introduction

The current work expands upon our previous work [1], where we present a new high-resolution dataset of paintings, and explore image classification on it.

In the last decade, a significant quantity of artwork has been digitized. That fact, combined with the substantial progress in the area of computer vision, has opened up the interesting research area of automated painting classification [2]. Automated painting classification can be broken down into two subtasks: artist identification and style categorization. The former task involves identifying a painting as the work of a specific artist, while the second task involves labeling paintings by art movement or style. Identifying artists and styles in fine-art paintings has numerous applications in several industries, such as tourism and movie-making, art education, and investigation of art forgery (though the system proposed here does not claim to be suitable for this last application). For instance, a user may take the photo of a painting or reproduction somewhere and want to know more about the painting, such as the artist's name and style. A system that can autonomously classify art is, therefore, of great interest. However, both tasks pose

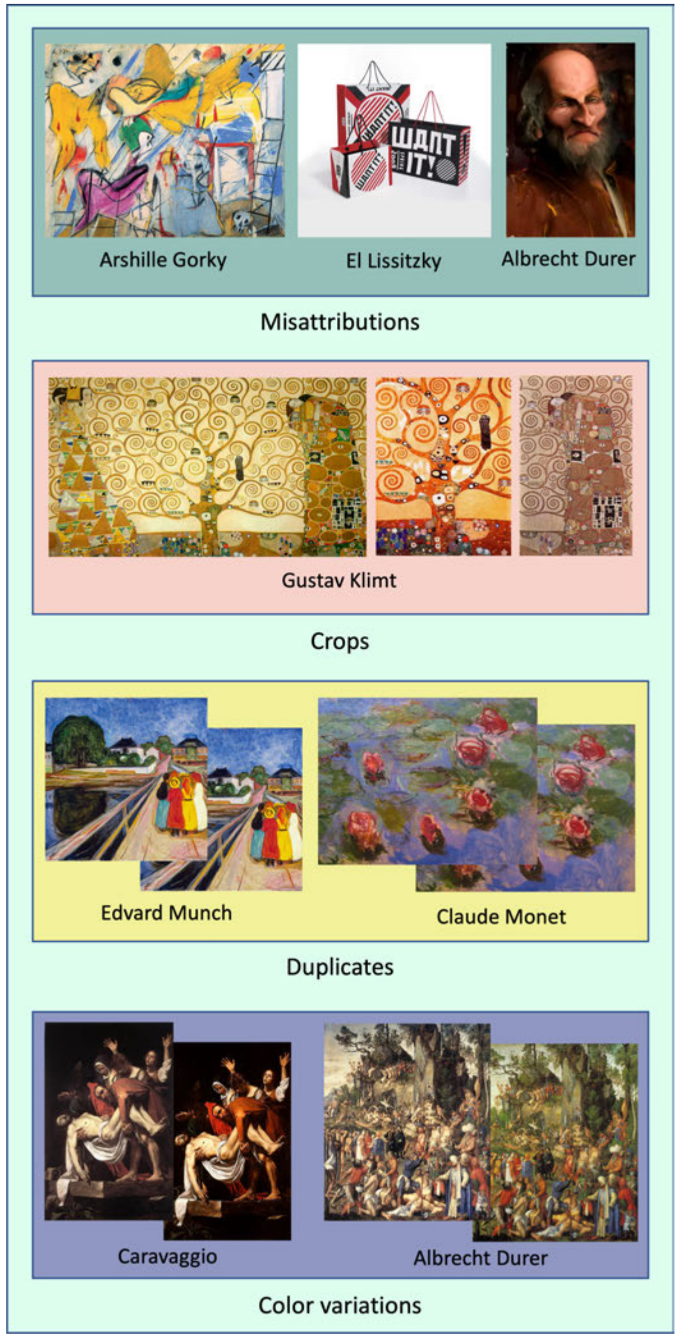

Figure 1. Some errors that exist in the old Painting-91 dataset.

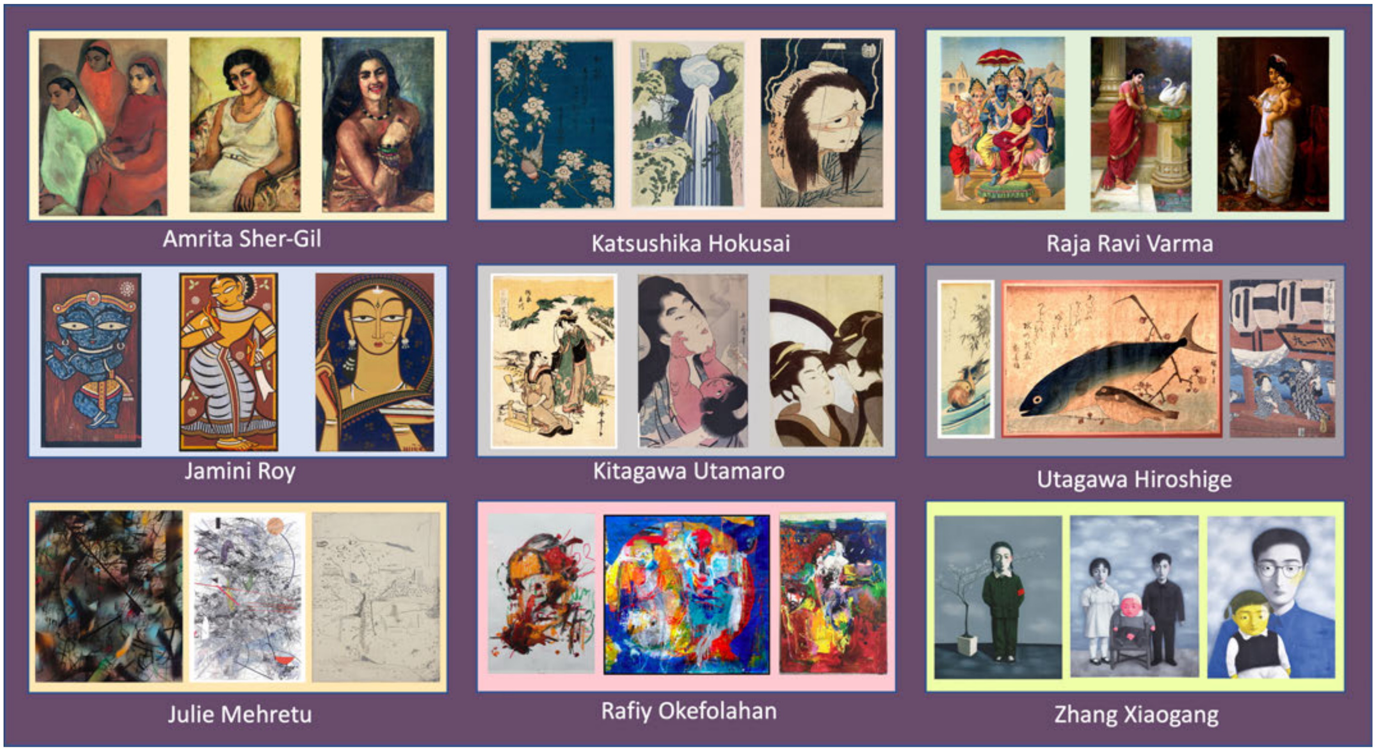

Figure 2. Some paintings of the newly added nine artists that are included in the Paintings-100 dataset.

significant challenges 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 due to the presence of stylistic variations that can occur even within a single artist's work [3] [4] [5].

In the current work, our contribution is threefold: first, we do a detailed discussion of the new Paintings-100 image dataset curated by us [1], highlighting our improvements to the Painting-91 Dataset [3]. Second, we have done an extensive evaluation of this rich and varied dataset using several convolutional neural network (CNN)-based methods on whole images as well as random image patches for both the artist and style classification tasks, finally showing that an ensemble of multiple models performs best. Last, but not the least, we attempt to identify the salient features of some of the style classes by examining the CNN response maps, and also identify the points of confusion between classes.

The rest of this paper is organized as follows. Section II lists some of the other work in this field. Section III and its subsections outline in detail the construction of the new dataset used in this work. Section IV describes our proposed methodology in detail, and Section V reports the classification performance on this dataset, the experiments performed, the results obtained, and discusses the outcomes. Finally, we list our conclusions and directions for future research in Section VI.

II. Related Work

Painting styles encompass the unique techniques, methods, and characteristics artists use to express themselves. Over the last two decades, computer vision, machine learning, and artificial intelligence have been successfully applied to analyzing and interpreting fine-art paintings and drawings [6] [7] [8] [9], offering innovative tools for art experts and scholars. Unlike artist categorization, which focuses on individual artists, style classification recognizes that multiple painters can share a common style, making it a distinct and challenging problem [3].

The traditional computer vision techniques for image classification use either color [10], shape [11] or texture [12] features. But feature extraction also results in the loss of semantic information from the painting image, thus increasing the challenge of identifying its style [13]. Techniques such as detecting and recognizing the artist's signature are not universally applicable due to the signature often getting cropped out of digital reproductions. That is why in recent years, computer vision researchers have explored the painting style recognition problem using CNNs, which are better at preserving semantic information. One of the major challenges of being able to effectively use CNNs for painting classification is the need for large hand-labeled datasets [5]. The limited availability of training data has led to a reliance on pre-trained models. Thus, instead of training a neural network from scratch, existing approaches either fine-tune pre-trained models, utilize them for feature extraction, or opt for non-neural network-based methods altogether.

In [5], the authors explored the applicability of CNNs for art-related image classification tasks by performing extensive CNN fine-tuning experiments and consolidating the results for five different art-related classification tasks. They also showed that fine-tuning networks pre-trained for scene recognition and sentiment prediction produced better results than those pre-trained for object recognition, demonstrating the effectiveness of leveraging scene and sentiment knowledge for style recognition. Rodriguez et al. [7] used a weighted sum of the individual-patch classification outcomes to provide the final stylistic label of the analyzed painting. Lately, [14] employed a framework to compare the performance of six pre-trained CNN architectures (Xception, ResNet50, InceptionV3, InceptionResNetV2, DenseNet121, and EfficientNet B3) for style classification using transfer learning, and studied the effect of different optimizers with learning rates on each model.

In our previous work [15], we explored the use of pre-trained 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 [3], the WikiArt dataset [16], and the Painting dataset consisting of ten classes of fine-art paintings from the PASCAL VOC [17]. But even though these datasets exist, the number of hand-labeled paintings available for effectively using CNNs is very limited [5]. To that end, we have worked on expanding the existing Painting-91 dataset [3] to construct a bigger dataset called the Paintings-100 dataset [1]. While constructing this dataset, we worked on improving the existing Painting-91 dataset [3] 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. We also enhanced image resolutions from the previous dataset, and augmented certain artist categories, which had fewer images in the previous dataset, with more images. Finally, we did extensive experiments with several CNN models to address both the artist classification and style classification tasks.

III. Dataset Construction

When we worked on [15], we realized that the images in the original Painting-91 dataset [3] 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 described in the subsections below, along with the improvements made by us.

A. Low Resolution

This was the main motivation for constructing the new dataset. The mean size of an image in the Painting-91 dataset is 268×263 pixels. These dimensions are smaller than the input sizes of many modern 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 [18]. We were successful in this task for about 97% of the images, but we also ran into other errors as detailed next.

Figure 3. Some paintings of the 14 different style categories that are included in the Paintings-100 dataset.

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

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 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 a painter's identity or style 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.

G. 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.

We also added the style movement Ukiyo-e, into this new collection for style classification task. Examples of paintings from each of these 14 style classes is shown in Figure 3. Table II displays a list of the various painting styles used by the different artists that are included in the Paintings-100 dataset.

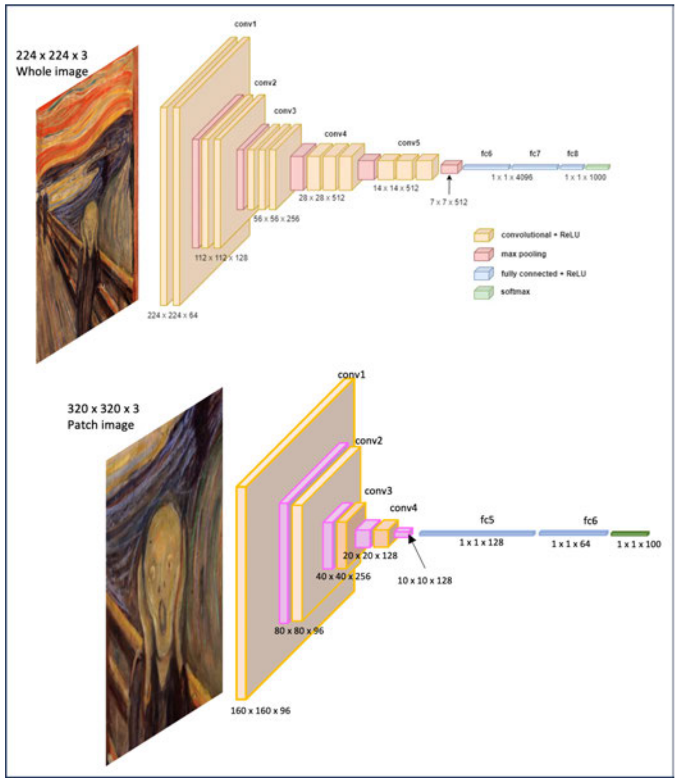

Figure 4. For artist classification task, we used both whole images as well as random patches from the images to feed into different CNN models.

IV. Methodology

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 [3]. In the current work, we have analyzed the dataset with respect to both these problems.

A. Artist Classification

While CNN-based models have largely outperformed other techniques for various classification tasks, the artist classification problem is somewhat challenging for these models. This is mainly due to two reasons. 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 limits 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 on its own, we decided to use a bit of both.

TABLE II
DIFFERENT PAINTING STYLES INCLUDED IN THE PAINTINGS-100 DATASET.

Style	Artists	# Images
Abstract Expressionism	Jackson Pollock, Mark Rothko, Willem De Kooning	167
Baroque	Caravaggio, Diego Velazquez, Jan Vermeer, Nicolas Poussin, Peter Paul Rubens, Rembrandt Van Rijn	304
Constructivism	El Lissitzky, Kazimir Malevich, Wassily Kandinsky	153
Cubism	Fernand Leger, Georges Braque, Piet Mondrian, Picasso	157
Impressionism	Claude Monet, Edgar Degas, Edouard Manet, Pierre-Auguste Renoir	205
Neo-classical	Jacques-Louis David, Jean-Auguste-Dominique Ingres	106
Pop Art	Andy Warhol, David Hockney, Roy Lichtenstein	153
Post Impressionism	Amedeo Modigliani, Georges Seurat, Paul Cezanne, Paul Gauguin, Vincent Van Gogh	255
Realism	Camille Corot, Gustave Courbet, James McNeill Whistler, Jean Francois Millet, Raja Ravi Varma	256
Renaissance	Raphael, Sandro Botticelli, Titian	172
Romanticism	Caspar David Friedrich, Eugene Delacroix, Francisco De Goya, John Constable, Joseph Mallord William Turner, William Blake	310
Surrealism	Georgia Okeefe, Joan Miro, Max Ernst, Rene Magritte, Salvador Dali, Zhang Xiaogang	314
Symbolism	Gustave Moreau, Gustav Klimt	105
Ukiyo-e	Katsushika Hokusai, Kitagawa Utamaro, Utagawa Hiroshige	150

1) Preprocessing: To address the problem of too few images and too much detail, we used an ensemble of multiple CNN models that use both downsampled whole images and full-size patches cropped out of the high-resolution images. These patches were randomly selected square patches of size 224×224 pixels or larger. In both cases (downsampled whole image and cropped patches), we used 24 images per class with augmentation (variations created by slight rotation, translation, shear, scaling, and horizontal mirroring) for training, 6 per class for validation, and the rest for testing. The whole and cropped images were histogram normalized and preprocessed for their corresponding CNN models.

However, using only whole images for training poses another problem. Even though each style class has three or more artists, the total number of training images is still quite low for training CNNs properly. Because of this, we use image augmentation techniques to increase the size of our training set. We use translation, rotation, shear, zoom and horizontal flip operations on our images for augmentation. The images are also resized to 224×224 pixels. Finally, each image is passed through a preprocessing function specific to each pre-trained network before passing through the network itself.

2) Model Selection: Classifying whole images and classifying patches are two different problems. 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 [19] trained on the ImageNet image dataset [20] since we had far fewer images. These two models are shown in Figure 4. The models were chosen empirically. We used decision fusing based on the labels predicted by the two models.

Although the style classification task takes the same input as the artist recognition task, it has a different output. Specifically, here the challenge lies in comprehending the artistic style of the artwork, which is often more complex and subtle than merely recognizing the painting's content or the artist. For this task, we use deep learning as well. The styles present in our dataset are Abstract expressionism, Baroque, Constructivism, Cubism, Impressionism, Neo-classical, Pop art, Post-impressionism, Realism, Renaissance, Romanticism, Surrealism, Symbolism and Ukiyo-e. Each of these styles offers a unique perspective on how artists interpret the world and their experiences. For example, abstract expressionism that flourished in the mid-twentieth century, emphasizes spontaneous, automatic, or subconscious creation, with Jackson Pollock and Mark Rothko as key figures. On the other hand, Ukiyo-e is a genre of Japanese art that flourished from the 17th to the 19th centuries, admired for its beauty, craftsmanship, and cultural significance. All the style categories being used

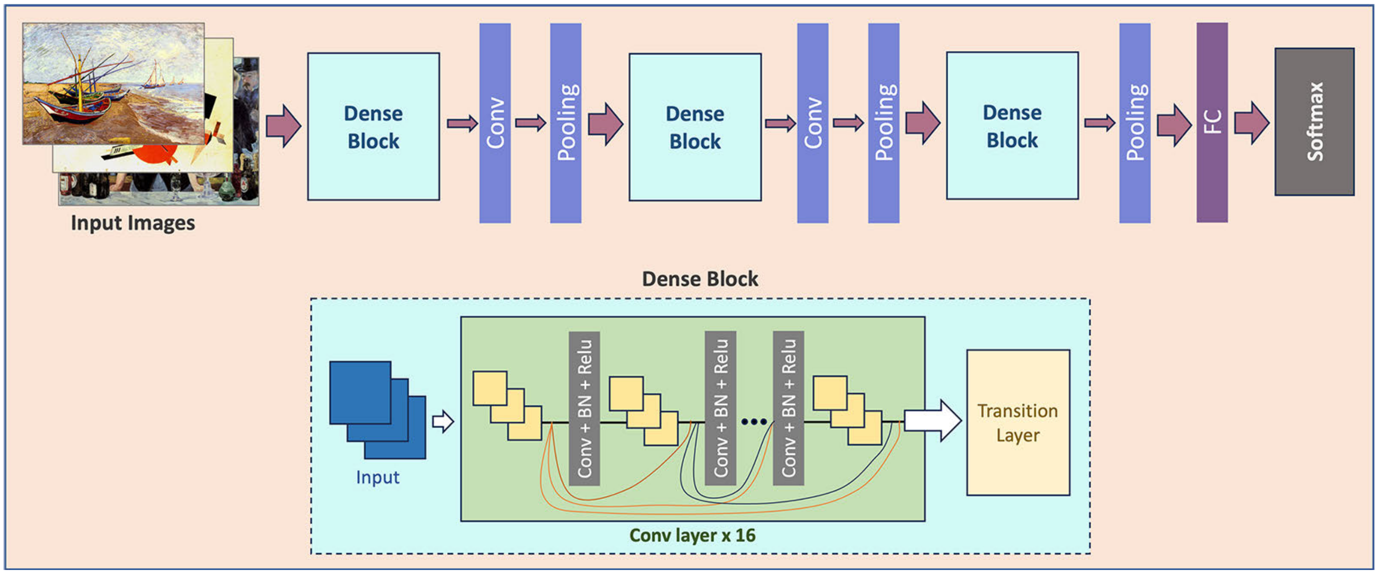

Figure 5. The internal architecture of DenseNet201 [21].

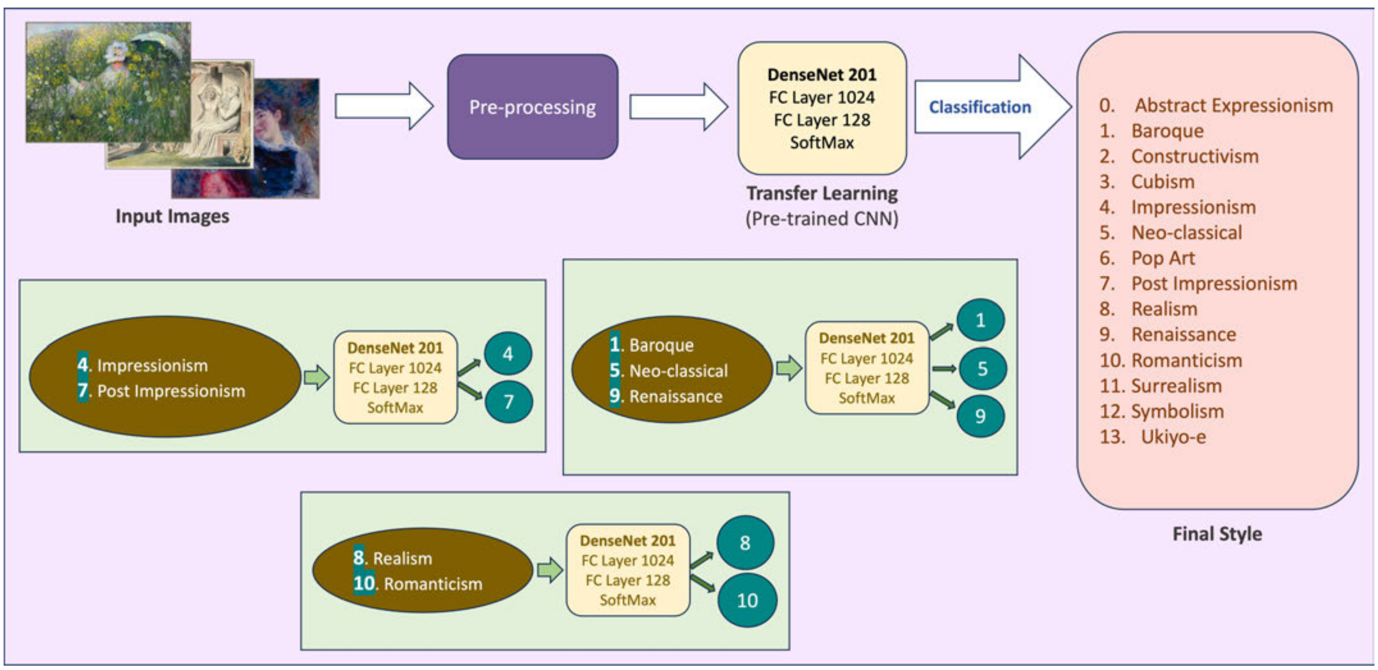

Figure 6. Our proposed method for style recognition. We employ an ensemble of four DenseNet201 models to do style classification in two stages.

for this work are listed in Table II. It should be noted that only 55 out of the 100 artists are included in this task since some of the other artists painted multiple styles, or were sole representatives of their style.

B. Style Classification

In this section, we discuss our experiments regarding the style classification task and our results in detail.

1) Preprocessing: While addressing the artist recognition problem previously, we found that dividing the image into small patches and using them for training the classifier worked well [1]. However, that technique did not work well with the style recognition task. Our intuitive understanding of this is, the style class of an image is much more dependent on the whole image rather than finer details. That is why, our models cannot reliably learn the style patterns with small patches of the images.

However, using only whole images for training poses another problem. Even though each style class has three or more artists, the total number of training images is still quite low

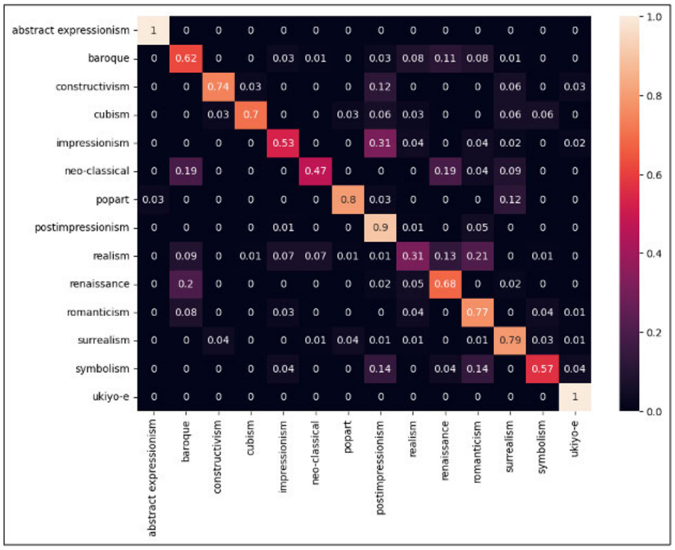

Figure 7. Confusion Matrix of Style Classification using DenseNet201. The rows indicate actual class labels while the columns indicate predicted class labels.

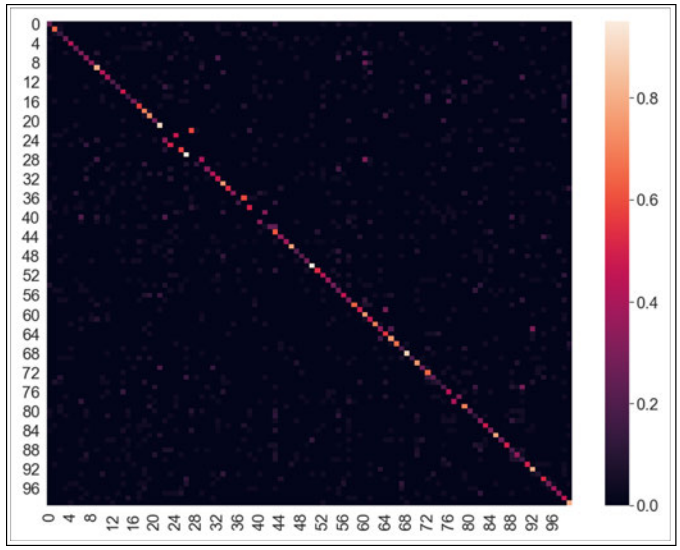

Figure 8. The confusion matrix for the artist classification experiment using the combined decision of two CNN models. The rows indicate actual class index values while the columns indicate predicted index values.

for training CNNs properly. Because of this, we use image augmentation techniques to increase the size of our training set for fine-tuning pre-trained networks (as detailed in the next section). We use translation, rotation, shear, zoom and horizontal flip operations on our images for augmentation. The images are also resized to 224×224 pixels. Finally, each image is passed through a preprocessing function specific to each pre-trained network before passing through the network itself.

2) Model Selection: The lack of labeled training images that makes painting classification so challenging for CNNs is somewhat less acute for style classification, but it is still very much present. The limited number of training images makes this problem particularly suited for transfer learning. For this work, we test five well-known CNN architectures on our data. Out of the five, three performed well on the style classification problem. All CNN models were pre-trained on ImageNet data [20].

Since the number of images, even after augmentation, is not sufficient to train a deep neural network from scratch, we did not create our own model for this task. We tried several pre-trained CNN models and compared their performance. The models that performed reasonably well were the VGG16 network [19], the DenseNet121 network [21], and the DenseNet201 network [21]. The tested models that did not perform well were the InceptionV3 [22] and the EfficientNetB3 [23] models. Their performances are detailed in Section V. Since the DenseNet201 was our best-performing model, we selected this model for all further classification experiments.

DenseNet201 [21] is a deep convolutional neural network with 201 layers, where each layer is connected to every other layer in a feed-forward manner. It connects all its 201 layers directly, without skipping any connections. This allows each layer to learn not just from the previous layer, but also from all the layers that came before, mitigating the vanishing gradient problem and enhancing feature propagation. Additionally, it promotes feature reuse while achieving a compact architecture with a reduced parameter count. The internal architecture of this network is shown in Figure 5.

3) Ensemble-based Classification: Now, we will discuss the ensemble-based classification method shown in Figure 6. The confusion matrix of the style classification task as done by the DenseNet201 model is shown in Figure 7. As can be seen in the matrix, there are several areas of inter-class confusion. The biggest of these is between Impressionism and Post-impressionism. Other large confusion rates are between Realism and Romanticism, and between Baroque, Neo-classical and Renaissance. To handle these particularly difficult classification problems, we train three more DenseNet201 models. The first of these is trained only on Impressionism and Post-impressionism images, the second only on Baroque, Neo-classical and Renaissance images, and the third only on Realism and Romanticism images. While testing, we first pass each test image through the network trained to classify between all 14 classes. If the output label is one of the classes with high confusion, that image is passed through the CNN model specialized for that class and the output label from this second model is assigned to it. So, if the first model outputs one of the labels shown in the green circles in Figure 6, the image in question is passed through a second CNN model which assigns the final label to it. This leads to considerable improvement in the classification accuracy as detailed in Section V-B.

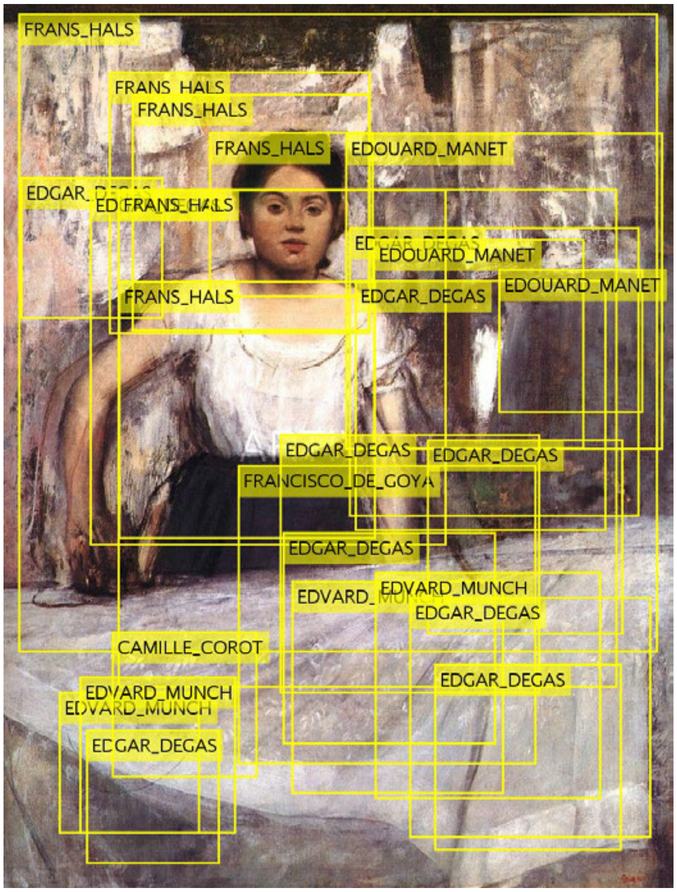

Figure 9. Painting by Edgar Degas. When the whole image is used for artist recognition, the CNN identified it as a Frans Hals painting, whereas by using random patches, it is correctly classified as an Edgar Degas artwork.

V. Experiments and Discussion

In the following subsections we describe the two sets of experiments that we performed on the Paintings-100 dataset. These two sets of experiments were done for the artist recognition and style recognition tasks, respectively.

A. Artist Classification Experiments

For the artist classification task, 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 8. Figure 9 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 last convolutional layer of our patch-image classifying CNN using the Grad-CAM technique [24]. 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 10 show some of the characteristics of artists that the network can identify correctly. For example, bold outlines are a signature characteristic of Indian painter Jamini Roy and these outlines are highlighted in the topmost example in Figure 10. Similarly, dotted patterns and certain kinds of brush strokes are recognized as characteristic features of Roy Lichtenstein and Vincent Van Gogh, respectively.

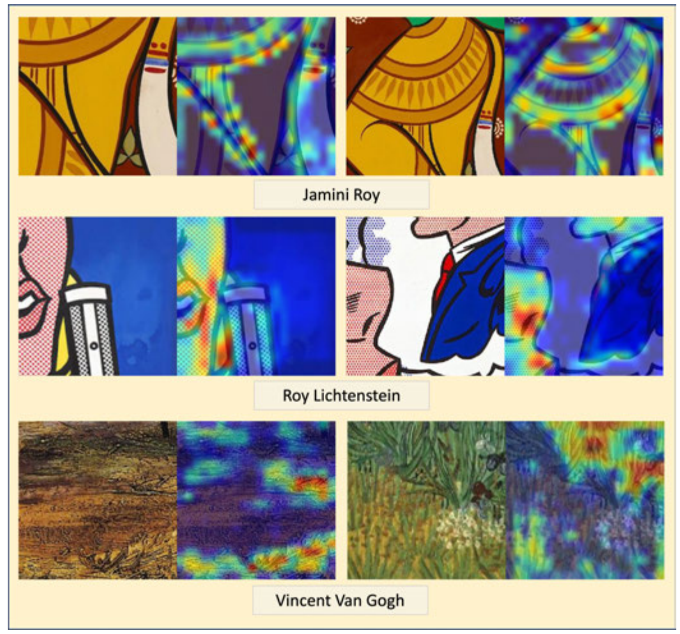

Figure 10. A few paintings and their Grad-CAM response maps showing regions of interest for artist recognition as detected by the CNNs.

B. Style Classification Experiments

For the style classification task, we first ran the same experiment once for each CNN architecture that we tested. This was a single 14-class classification of all style images. Out of the total number of images shown in Table II, we used 80% from each class for training and the rest for testing. 80% of the 80% used for training are used for true training and the other 20% are used for validation. The models that performed reasonably well on this experiment were the VGG16 network [19], the DenseNet121 network [21], and the

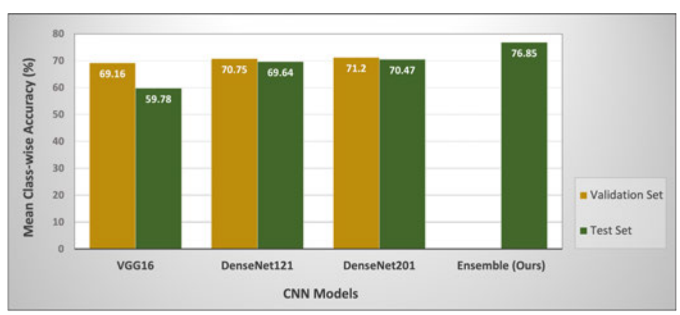

Figure 11. A comparison of the validation and test set accuracies of the three different CNN architectures that we tested, along with the ensemble accuracy on the test set.

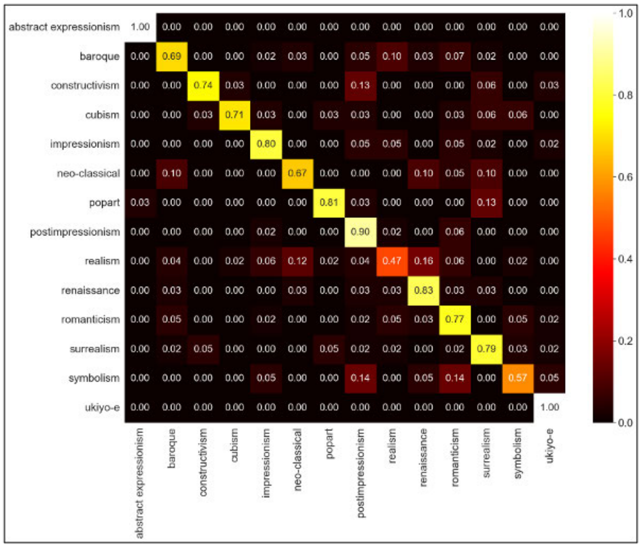

Figure 12. Confusion Matrix of Style Classification using our ensemble-based method. The rows indicate actual class labels while the columns indicate predicted class labels.

DenseNet201 network [21]. Since the DenseNet201 with a validation accuracy of 71.20% was the best performer, we proceeded to the next stage with this as our primary architecture. The validation and test set classification accuracies of all the models we tested are shown in Figure 11. It should be noted that the validation accuracy is not shown for our proposed ensemble model since we are combining the decisions of different trained models to get this result, and the concept of validation is not meaningful here.

For the next stage, we trained three more DenseNet201 networks. The first was trained on Impressionism and Post-impressionism images and gave us an accuracy of 92.16%. The second model was trained on three classes, namely, Baroque, Neo-classical, and Renaissance, and gave us a validation accuracy of 77.42%. The last model was trained on images from two style classes - Realism and Romanticism, and yielded a validation accuracy of 82.22%.

Subsequently, we created a two-level ensemble of CNN models as described in Section IV and passed all test images through this ensemble. This method gave us a test accuracy of 76.85%, which was an improvement of about 6% over a single DenseNet201 handling all 14 style categories on the test set. It should also be noted that this result is even higher than the results shown by [14] on the Painting-91 dataset [3] which contains one style class and five artists less (for this specific problem) than our Paintings-100 dataset. A comparison of the class-wise classification accuracies using the proposed ensemble method can be found in Figure 12.

Next, we used the Grad-CAM method [24] to view the response maps of a classification network. For this experiment, we used the VGG16 network instead of the DenseNet201 since VGG16 is a linear model and easier to combine with Grad-CAM. A small sample set from the results is shown in

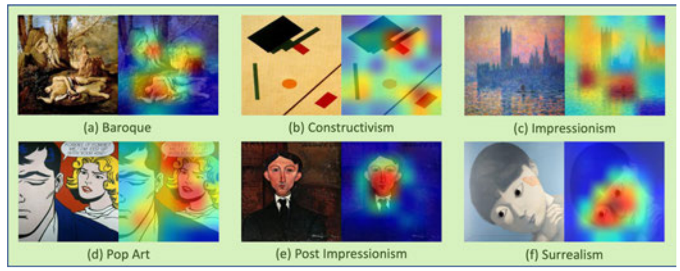

Figure 13. A few paintings and their Grad-CAM response maps showing regions of interest for style recognition as detected by the CNNs.

Figure 13. This gives us some insight into what the networks are looking for to correctly classify the images. For instance, in Figure 13(a), the network clearly recognizes the Baroque style by focusing on the three human figures in the painting. In (d) and (e), the network focuses on the faces of the human figures to recognize their respective styles, while in (f), the focus is primarily on the eyes of the subject. These black eyes are a defining characteristic of surrealist painter Zhang Xiaogang, and the network learns to recognize them during training.

Figure 14 shows some of the confusing images that were misidentified by the system. But it is easy to see that these images are actually confusing to label. While (a), (c), and (f) are labeled as Surrealism, Cubism, and Constructivism, respectively, all of them look somewhat similar to the post-impressionist pastoral landscapes by Van Gogh or Paul Gauguin. Similarly, (b), (d) and (e) have features of the style they are labeled as by the network, along with the style they are originally annotated with.

Finally, we wanted to see how good our model was in

Figure 14. Some confusing images that were misclassified for style recognition as detected by the CNNs.

Figure 15. A few samples of images from outside this dataset, whose styles were accurately predicted by our system.

recognizing the styles of images from outside the dataset. To test this, we created a second small dataset of 140 images (10 from each style class). All of these images were by artists who were not among our style recognition training data. In fact, most of these artists are not even in the Paintings-100 dataset. Our system performed quite well with this completely unseen data as well, predicting 44.28% of the styles correctly. Some of the correctly predicted images from these external images are shown in Figure 15, and some of the wrongly classified images are shown in Figure 16. It should be noted that classifying painting style often involves a subjective decision and the same painting may sometimes reflect the properties of two or more different styles, which makes classifying the work of unseen artists very challenging.

VI. Conclusion and Future Work

In this work, we expanded our experimentation upon the large scale diverse high-resolution image dataset that we recently presented for artist and style classification. While this was based on the existing Painting-91 dataset, the improvements were significant enough for the Paintings-100 dataset to be considered a new dataset. We have explored both the artist recognition and the painting style classification problems by conducting extensive experiments using several CNN architectures, and found that ensembles of CNN models showed more promising results for both tasks. As the experiments show, our proposed ensemble methods perform better than any one single CNN model tested by us. Some of these methods cannot be applied on the original Painting-91 dataset because of low-resolution images. The focus of our work was exploring the suitability of the newly introduced Paintings-100 dataset for the artist and style classification problems, and we can safely say that it is indeed suitable for these tasks.

There are many different ideas that we would like to try out on this dataset in the near future. Currently, we are selecting the patches for the artist recognition task randomly. In future, we want to try selecting patches with face detectors and object detectors to see how that affects our results. Photographic conditions such as ambient lighting and camera model create big differences in the color maps of the digitized paintings. We plan to use some color normalization techniques to reduce the effect of photographic conditions on the paintings. In a later work, we would like to extend this work by including other painting datasets and other CNN models, since the generalization performance of our method still has room for improvement. We would also like to expand upon the response maps portion of this work to better understand and explain the functioning of our models.

References

[1] E. Knizhnik, B. Rivera, A. Sinha, and S. Banerji, “Paintings-100: A diverse painting dataset for large scale classification,” in The Ninth International Conference on Advances in Signal, Image and Video Processing (SIGNAL 2024), 2024, pp. 12–15.

[2] D. Zhu, “The application of computer image processing technology in the fine arts creation,” in 2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA), 2014, pp. 790–792.

[3] 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.

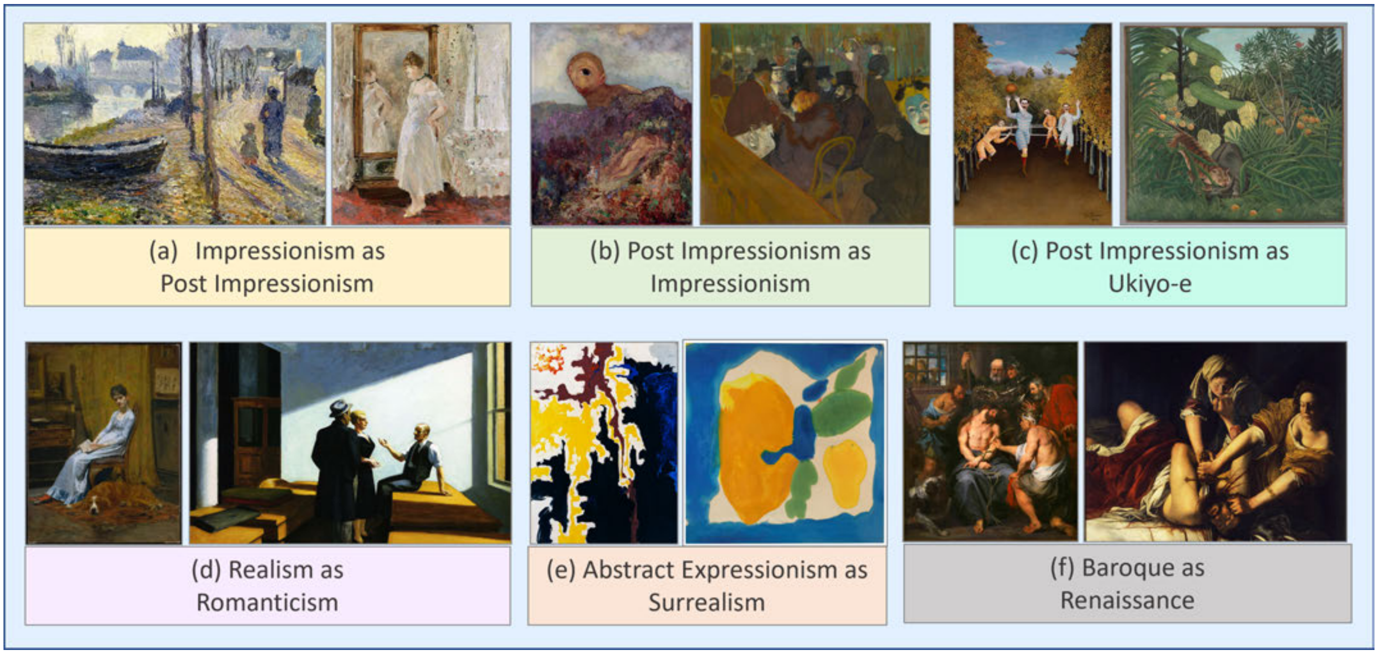

Figure 16. A few samples of images from outside this dataset, whose styles confused our system, leading to wrong predictions.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] I. Nunez-Garcia, R. Lizarraga-Morales, U. Hernandez-Belmonte, V. jimenez Arredondo, and A. Lopez-Alanis, "Classification of paintings by artistic style using color and texture features," Computación y Sistemas, vol. 26, 12 2022.

[11] R. G. Condorovici, C. Florea, and C. Vertan, "Automatically classifying paintings with perceptual inspired descriptors," Journal of Visual Communication and Image Representation, vol. 26, 2015, pp. 222–230. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1047320314001977

[12] Q. You, J. Luo, H. Jin, and J. Yang, "Robust image sentiment analysis using progressively trained and domain transferred deep networks," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 29, 09 2015.

[13] Q. Zhao and R. Zhang, "Classification of painting styles based on the difference component," Expert Systems with Applications, vol. 259, 2025, p. 125287. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417424021547

[14] B. Menai and M. Babahenini, "The effect of optimizers on CNN architectures for art style classification," International Journal of Computing and Digital Systems, vol. 31, 2023, pp. 353–360.

[15] 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.

[16] 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.

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

[18] "Google reverse image search," https://www.google.com/imghp, accessed: May 29, 2025.

[19] 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.

[20] 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.

[21] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818–2826.

[23] A. Batool and Y.-C. Byun, "Lightweight efficientnetb3 model based on depthwise separable convolutions for enhancing classification of leukemia white blood cell images," IEEE Access, vol. 11, 2023, pp. 37 203–37 215.

[24] 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.