





# How to Generate Synthetic Paintings to Improve Art Style Classification

Sarah Pires Pérez<sup>(✉)</sup>  and Fabio Gagliardi Cozman 

Escola Politécnica, Universidade de São Paulo, São Paulo, Brazil

**Abstract.** Indexing artwork is not only a tedious job; it is an impossible task to complete manually given the amount of online art. In any case, the automatic classification of art styles is also a challenge due to the relative lack of labeled data and the complexity of the subject matter. This complexity means that common data augmentation techniques may not generate useful data; in fact, they may degrade performance in practice. In this paper, we use Generative Adversarial Networks for data augmentation so as to improve the accuracy of an art style classifier, showing that we can improve performance of EfficientNet B0, a state of art classifier. To achieve this result, we introduce Class-by-Class Performance Analysis; we also present a modified version of the SAGAN training configuration that allows better control against mode collapse and vanishing gradient in the context of artwork.

**Keywords:** Computer vision · GAN · Art style classification

## 1 Introduction

No society, however low it may have been its level of material accomplishment, has ceased to produce art [15]. Art has been, through the ages, a mirror reflecting society; the purpose of art has varied but its importance in capturing society has been constant. One way to discuss the content of artwork is through style categories. Such categories indicate connections between paintings and help us to better understand their meanings. Even if many art theorists look at style categories with reservations,<sup>1</sup> categories provide an important guideline for beginner art aficionados and the general crowd within museums and galleries. Besides, style categories can be used by recommendation tools to improve a museum visitor's experience, either in a traditional physical museum location or in one of the increasingly common digital galleries.

Alas, the current quantity of online art has now surpassed our manual indexing abilities. The field of computer vision has developed artwork style classification tools using a diversity of techniques from feature-based machine learning [2, 32] to deep learning [3, 10, 23]. Beyond the challenge of class imbalance,

---

<sup>1</sup> This feeling can be found in articles on the study of art: <https://www.artsy.net/article/alina-cohen-art-movements-matter>.

style classification is further complicated by aesthetic diversity, a problem that is not usually met in most image classification models. For instance, the Baroque is directly linked to the historic moment of the Catholic Church struggling against the Protestant Reformation; in an attempt to combat the Protestant faith, the Church announced during the Council of Trent (1545–1563) the dissemination of religious ideas through images to awaken religious fervor—therefore, the Baroque works are very concentrated on the religious theme. Minimalism, on the other hand, does not have a single theme, but it employs a few elements in a distinctive manner [15].

The use of image augmentation is well established in computer vision so as to enhance the content of a training dataset. But usual techniques do not produce good results when applied to artwork; for instance, many of the common image augmentation techniques modify the resulting color of the image, which can be problematic in the case of artwork as colors often have or convey particular meaning. In this paper we propose an artwork style classifier that resorts to Generative Adversarial Networks (GANs) as an image augmentation tool for oversampling.

Artwork generation with GANs is not a direct matter; even high performance GAN architectures did not do well in preliminary experiments, so we had to propose techniques to enhance image quality and diversity. We describe in this paper a methodology for data augmentation that is tuned to the requirements of artwork classification; we also describe a specialized network architecture and an adjusted loss function that both contribute to the overall performance of our proposal. In short, our contributions are both in adopting GANs for data augmentation in the particular context of artwork classification, and in proposing novel techniques that may be of broad interest whenever similar classification problems are met.

The paper is organized as follows. Section 2 analyzes related work on art style classification and Generative Adversarial Networks. Section 3 describes our approach; results and analysis are shown in Sect. 4. Finally, conclusions are presented in Sect. 5.

## 2 Related Work

In this section, we explore relevant efforts on artwork classification. We also present the architectural development of Generative Adversarial Networks (GANs), its use in the art domain and its relevance to image augmentation.

### 2.1 Artwork Classification

Initial research on artwork classification was focused on feature-based machine learning [2, 32]. Since the initial success of a convolutional neural network (CNN) in object classification for ImageNet [25], several studies in artwork classification have explored these networks [3, 10, 16, 23]. To improve the performance of the

classification of art styles with CNNs, researchers have investigated the benefits of pre-trained weights from the ImageNet Challenge instead of random startups [26, 36].

There has been significant activity in artwork classification in recent years. Cetinic *et al.* expanded artwork classification beyond style, genre, artist and time period classification, examining also nationality and testing scene recognition and sentiment analysis techniques [6]. Chu *et al.* studied the style classification task focusing on describing image texture with deep learning. They investigated the intra-layer and inter-layer correlations in order to create deep features for style classification. [9]. The work of Elgammal *et al.* analyzed the learned representations of a fine-tuned ResNet-152, noting that some of the style patterns designed by Heinrich Wölfflin (1846–1945) correlate with the PCA decomposition of these learned representations [14]. In Rodriguez *et al.*, five image patches of painting were used for training and weights for each patch were optimized in order to improve accuracy of the final model [29]. Sandoval *et al.* also worked with image patches, but in a two-stage deep learning approach, in which these five patches are trained independently at a first step. At the second stage, the outcome of these patches are fused to a second shallow neural network for the final decision [31]. Zhong *et al.* presented a two-channel dual path network and two inputs are used: the RGB image and four-directional gray-level co-occurrence matrix for detecting the brush stroke information [44]. The work of Zhu *et al.* not only trained the Inception V3 network for classifying nine artistic movements, but also used Grad-CAM heat map for visualizing the areas of the images the model was focusing for class prediction [46]. Chen *et al.* presented an adaptive cross-layer correlation for artwork classification, in which it adaptively weights features in different spatial locations based on similarity [7]. Bianco *et al.* studied the advantages of training a model with the full image of the paintings and also its crop at different resolutions [4].

## 2.2 Generative Adversarial Networks

Generative Adversarial Networks (GANs) are built by training two networks with a minimax game framework, where we have a model capable of generating synthetic data (model G) and another model that evaluates whether the data is real or synthetic (model D) [19]. In the work of Mirza and Osindero, an extra parameter was introduced in the GAN architecture so as to allow the generator model to create images according to class labels [27]. Chen *et al.* developed a GAN architecture to learn disentangled representations in an unsupervised manner; they introduced a representation learning algorithm called Information Maximizing Generative Adversarial Networks (InfoGAN), in which an information-regularized minimax game is used in order to train a multi class generative model without the label information [8].

Some work on GANs has focused on the impact of the loss function. Arjovsky *et al.* used the concept of the Earth Mover (EM) distance, also known as Wasserstein-1 distance, as a loss function [1]. Gulrajani *et al.* improved the Wasserstein function loss adding a gradient penalty [20].

An important concept in the evolution of the GAN architecture is the self-attention mechanism. The convolutional architecture processes information in local neighborhoods and it has no mechanism to deal with long distance dependencies. The self-attention enables both the generator and the discriminator to deal with widely separated spatial regions. The work of Zhang *et al.* presented this concept as Self-Attention GAN (SAGAN) [43] and other proposals have used this architecture as a reference [5, 11].

The use of synthetic images to improve classifiers has been adopted particularly in the medical field, for instance in simulating lung nodules [21], ECG [42], liver lesions [17], chromosomes [40], skins lesions [28] and Covid-19 results [12, 38]. We also mention the work of Suh *et al.*, where classifier loss is included in the GAN training process in order to reduce ambiguity between classes: their classification enhancement generative adversarial networks (CEGAN) consist of three independent networks – a discriminator, a generator and a classifier – using WGAN-GP for classification under imbalanced data conditions [34].

Generative adversarial networks have already been used in the art domain. For instance, one can find a generator for image style transfer (the CycleGAN [45]), an Image-to-Image translator from art to real images [18, 37], and a model specialized in creating Chinese landscapes [41]. The first, and to the best of our knowledge the only, work aimed at creating artwork with the WikiArt dataset appeared in Ref. [13]. There, Elgammal *et al.* developed the Creative Adversarial Network to creatively generate artwork by maximizing deviation from established styles and by minimizing deviation from art distribution [13]. The authors found that the way to encourage the generator “to be creative” was to penalize it any time in which it was too easy for the Discriminator to identify the synthetic image as being art from a certain style.

### 3 Our Proposal

Simply put, our main goal is to enhance the accuracy of artwork style classification. In pursuing this, we were led to study the potential benefits of synthetic images generated by Generative Adversarial Networks. As noted before, synthetic art generation has been employed only by Elgammal *et al.* (2017), but not with the purpose of enhancing classification; to the best of our knowledge, the latter task has not been investigated yet. In short, our specific goal here is not to make art, but to improve art style classification.

We introduce a strategy that we refer to as Class-by-Class Performance Analysis. In order to start up learning, a baseline model is trained without image augmentation techniques. The next step is to explore the benefit of geometric transformations. We then concentrate on classes with the lowest performance; we want to maximize the information about these classes. Such information is enhanced by a version of GANs with a self-attention mechanism for image diversity that is trained with the Wasserstein with Gradient Penalty loss function for avoiding vanishing gradient.

In addition, classes with low performance may or may not contain a small number of images. Hence, we suggest two strategies when we sample either:

- low quantity classes: add a multiple of the number of original images;
- high quantity classes: add a fraction of the number of original images.

There are many decisions that must be set in implementing this strategy; some of them depend on the particular artwork collection is dealing with. In the following subsections, all relevant aspects of our strategy are discussed: our study of image augmentation, our choices regarding the GAN architecture and the loss function and, finally, the model architecture for our classification task.

### 3.1 Image Augmentation

Data Augmentation is a natural solution to the problem of limited data [33]. With respect to images, augmentation techniques consist of geometric transformations, such as rotation, image cropping, flipping and color conversions [34]. We studied these techniques in the context of art style classification, a setting with many classes and serious class imbalance.

Our first experiment was to ascertain the benefits of rotation, image cropping and flipping. Results can be seen in Table 6. Improvement was obtained by rotating the image between -10 and 10 degrees, horizontal flips and random crops. Alas, traditional image augmentation techniques often generate simple and redundant copies of the original data in many cases [34]. In some cases, data augmentation may not be a profitable idea. For many domains, color conversions offer image diversity; for instance, the object class “bicycle” can be represented with red or blue bicycles. In artwork, however, colors are meaningful and altering them without care can result in an image that lies in a category different from the original one. For example, Pablo Picasso’s Blue Period paintings (1901–1904) should always be represented by gloomy shades of grayish blue. The true atmosphere color for Édouard Manet should always be violet and many of his Impressionist colleagues used violet as their main color [24].

### 3.2 Generative Adversarial Network

One of the most challenging tasks for Generative Adversarial Networks is producing a diverse set of synthetic images. Wang *et al.* compare the most influential GAN architectures with respect to image quality, performance against vanishing gradient and ability for mode diversity. In order to create synthetic artwork, where shape must greatly vary and colors are very relevant, we chose to focus on a mechanism proven to excel in diversity; we adopted Self-Attention Generative Adversarial Network (SAGAN) with some modifications which will be explained later. We chose this architecture as the basis of our GAN as the self-attention mechanism helps the GAN to learn global and long-range dependencies across multi-class images [39].

Figure 1 shows a self-attention module. Transformers are used to create the key  $f(x)$ , query  $g(x)$  and value  $h(x)$ :

$$f(x) = W_f x, \quad g(x) = W_g x, \quad h(x) = W_h x.$$

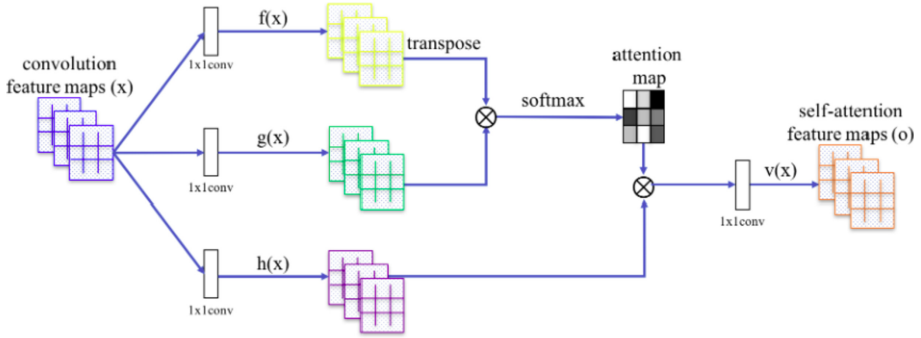


Fig. 1. The self-attention mechanism [43]

The attention map is created after applying a softmax to the dot product of the key and the query (Eq. 1). Another dot product is taken between the attention map and the value; the attention map is applied to the value in order to create a self-attention map ( $o_j$ ) (Eq. 2).

$$\alpha_{j,i} = \text{softmax}(f(x_i)^\top g(x_j)), \tag{1}$$

$$o_j = W_v \left( \sum_{i=1}^N \alpha_{j,i} h(x_i) \right). \tag{2}$$

The parameters  $W_f$ ,  $W_g$ ,  $W_h$  and  $W_v$  are the learned weight matrices.

### 3.3 Adversarial Loss Function

The original version of the SAGAN training configuration used hinge loss. For our dataset, we experienced frequent mode collapse and vanishing gradient with this loss, likely due to our dataset’s diversity. Moreover, in order to guarantee this to be  $k$ -Lipschitz, as needed to prevent mode collapse, a Gradient Penalty was added to the loss function. For this reason, we used the Wasserstein with Gradient Penalty (Wasserstein-GP) loss function so as to have better control of the values for feedback. That is, the loss is:

$$L = \mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [D(\tilde{x})] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)] + \lambda \mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2], \tag{3}$$

where  $\mathbb{P}_r$  is the real images distribution,  $\mathbb{P}_g$  is the generated images distribution and  $\mathbb{P}_{\hat{x}}$  is the sampling uniformly along straight lines between pairs of points sampled from the data distribution  $\mathbb{P}_r$  and the generator distribution  $\mathbb{P}_g$ . In our experiments, the interpolation between a batch of real images and fake images was enforced and the gradient norm of its output was limited at 1. The value of the penalty coefficient ( $\lambda$ ) was 10, following the original paper [20].

### 3.4 EfficientNet

The EfficientNet B0 architecture was chosen for these experiments due to its high performance and relatively small size. EfficientNet B0 belongs to a family of models known as EfficientNet, presented by Tan and Le [35]. The EfficientNet B0 is the EfficientNet family’s baseline, from which all the models from B1 to B7 are scaled up uniformly in dimensions of depth, width and resolution. Table 1 summarizes the stages in the EfficientNet B0. Each stage consists of one or more layers built with blocks of MBCConv [30], which are combined with a Squeeze-and-Excitation optimization [22].

**Table 1.** Summary of the composition of the EfficientNet B0

Stages	Operators	Resolution	# of channels	# of layers
1	Conv3 × 3	224 × 224	32	1
2	MBCConv1, k3 × 3	112 × 112	16	1
3	MBCConv6, k3 × 3	112 × 112	24	2
4	MBCConv6, k5 × 5	56 × 56	40	2
5	MBCConv6, k3 × 3	28 × 28	80	3
6	MBCConv6, k5 × 5	14 × 14	112	3
7	MBCConv6, k5 × 5	14 × 14	192	4
8	MBCConv6, k3 × 3	7 × 7	320	1
9	Conv1 × 1 & Pooling & FC	7 × 7	1280	1

## 4 Experimental Results

The methods described previously are only useful if they do lead to improvement in realistic circumstances. We developed and tested them by dealing with real artwork and existing datasets. In this section, we present the dataset we used and its characteristics, the training configurations for both GAN and classifier and the results for each step of the Class-by-Class Performance Analysis.

### 4.1 The Wikiart Dataset

We used the Wikiart dataset in our experiments. In fact, the version we used was the one discussed by Elgammal *et al.* [14], from which we adopted the following conventions:

- New Realism and Contemporary Realism were added to Realism;
- Action Painting was added to Abstract-Expressionism;
- Synthetic Cubism and Analytical Cubism were added to Cubism.

**Table 2.** Dataset used in experiments.

Art movement	Total of images	Training images
Abstract expressionism	2,783	2,283
Art nouveau	4,292	3,442
Baroque	4,241	3,448
Color field painting	1,615	1,308
Cubism	2,417	1,942
Early renaissance	1,391	1,134
Expressionism	6,720	5,457
Impressionism	13,060	10,566
Minimalism	1,258	1,009
Naïve art	2,340	1,917
Northern renaissance	2,552	2,084
Pop art	1,460	1,205
Realism	11,400	9,188
Romanticism	6,963	5,640
Ukiyo-e	1,167	940

A total of 63,659 images are available there; 10% of them were used for testing and 10% of the remaining dataset was used for validation. The training volumetry of the image distribution is presented in the last column of Table 2. The training dataset is used both in classifier training and in GAN training. The test dataset is the same for all experiments to ensure comparability of results. The class imbalance is apparent with the volumetry varying between 10,566 and 940 images.

## 4.2 GAN Training Configuration

In order to accommodate the training process in 2 GPUs GeForce GTX 1080 Ti (12 GB) to generate images of  $128 \times 128$ , we lowered the batch size to 32 images (the SAGAN was trained with batch size of 256 images). The process ran for 200.000 epochs (approximately 34 h). Each class was trained independently, using the equivalent training dataset of the classifier. The optimizer setup follows the original SAGAN article: Adam optimizer with  $\beta_1 = 0$  and  $\beta_2 = 0.9$ . The learning rate is constant but specific for each model: for the discriminator is 0.0004 and for the generator is 0.0001 [43].

## 4.3 EfficientNet B0 Training Configuration

Using the ImageNet pretrained model, the fine-tuning for our classification purpose was done by unfreezing each block until there was no further improvement



on the validation dataset. The Stochastic Gradient Descent optimizer was used with decay 0.9 and momentum 0.9; initial learning rate of 0.01 with decay after the fifth epoch ( $lr = lr * e^{-0.1}$ ). Images were resized to  $224 \times 224$  and batch size was 32 images.

#### 4.4 Baseline Results

The result of the EfficientNet B0 baseline trained model and the EfficientNet B0 trained model with geometric augmentation are shown in Table 6 (first and second lines). The performance for each class of the latter is shown in Fig. 2. It is important to highlight that low performance is not correlated to image quantity: the Ukiyo-e movement – the Japanese style – had the least amount of images and the best f1-score. Still, the worst performance was obtained by the Pop Art class with only 1205 images. On the other hand, the second and third worst performances belong to two classes with the largest number of images - Expressionism and Romanticism.

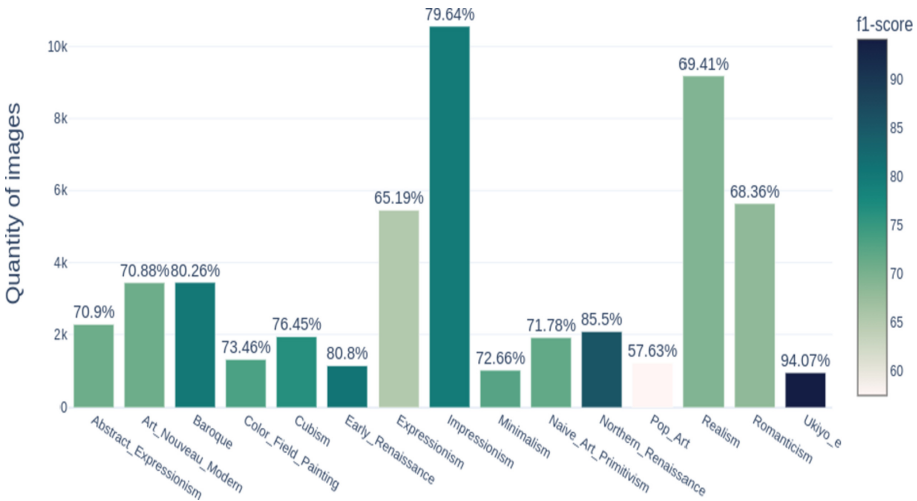


Fig. 2. Analysis of the trained EfficientNet B0 with geometric augmentation

#### 4.5 Sampling Low Quantity Classes

Pop Art is the fourth lowest class in terms of quantity and images, so the low quantity class sampling strategy was applied. The results are shown in Table 3. The classifier that obtained the best accuracy score is not the same as that obtained the best f1-score for the Pop Art class. This shows that other classes have benefited from the information obtained from the images generated (Tables 4 and 5).

**Table 3.** Summary of experiments with added synthetic Pop Art images

Synthetic images	class f1-score	Classifier accuracy
1205 (1×)	60.00%	74.00%
2410 (2×)	62.07%	73.83%
3615 (3×)	<b>65.52%</b>	74.05%
4820 (4×)	62.81%	<b>74.07%</b>

#### 4.6 Sampling High Quantity Classes

Both Expressionism and Romanticism classes have a similar behavior to that of our reference model: they have many more images than most of the classes – more than four times Pop Art’s image quantity – but low performance. It is interesting to observe that for both classes, the quantity of images that generated the best results was 1/4 of its original training data quantity. The synthetic Romantic class had a better f1-score when 3/4 of the quantity of the original data was added, but still didn’t produce a better classifier.

**Table 4.** Summary of experiments with added synthetic Expressionist images

Synthetic images	class f1-score	Classifier accuracy
682 (1/8)	64.61%	73.81%
1364 (1/4)	<b>66.77%</b>	<b>74.40%</b>
2728 (1/2)	65.62%	73.96%
4092 (3/4)	65.67%	74.11%

**Table 5.** Summary of experiments with added synthetic Romantic images

Synthetic images	class f1-score	Classifier accuracy
705 (1/8)	69.67%	73.69%
1410 (1/4)	70.38%	<b>74.27%</b>
2820 (1/2)	69.93%	73.83%
4230 (3/4)	<b>70.69%</b>	74.18%

#### 4.7 Summary of Results

Figure 3 shows how each class performed in the best experiment for each class (Pop Art, Expressionism and Romanticism). Although the classification problem has many classes to be able to point out correlations, it is possible to observe that the Art Nouveau class had its best performance with synthetic images from

the Pop Art artistic movement. It is also observed that the Pop Art class itself performs even better when we have the information of the synthetic images of the artistic movements Romanticism and Expressionism.

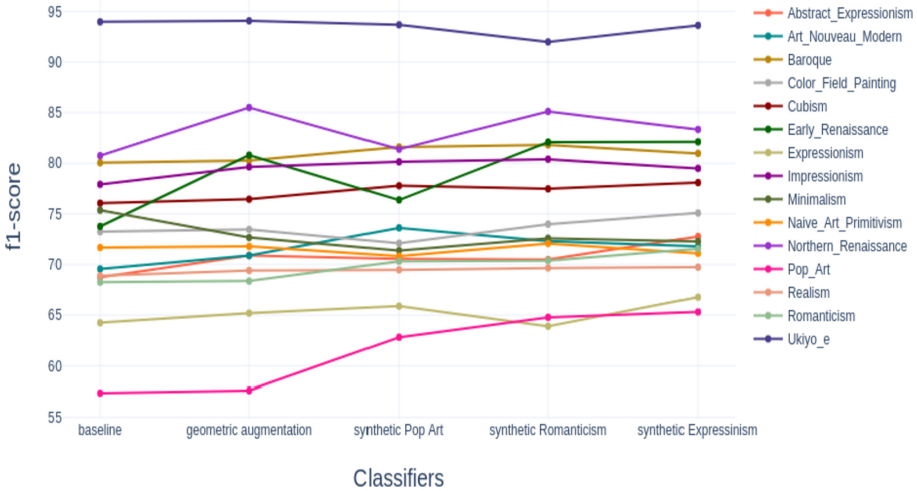


Fig. 3. The evolution of performance for each class during experiments.

Table 6 also shows the best result of each EfficientNet B0 model trained with synthetic artwork class. The strategy of Class-by-Class Performance Analysis allowed us to improve the classifier accuracy by almost 2%.

Table 6. EfficientNet B0 trained models.

Experiment	Precision	Recall	Accuracy
Baseline	73.58%	73.47%	72.43%
Geometric Augmentation	74.62%	74.46%	73.59%
Geometric Augmentation + synthetic Pop Art	74.89%	74.50%	74.07%
Geometric Augmentation + synthetic Expressionism	<b>76.12%</b>	<b>75.41%</b>	<b>74.40%</b>
Geometric Augmentation + synthetic Romanticism	75.30%	75.55%	74.27%

### 4.8 Generated Images

Figures 3, 4 and 5 show a sample of the images generated in our experiments. To help in understanding the quality of these generated images, real images of corresponding artistic styles were added above. By visual inspection, it is noticeable that the generated images retain general properties of the styles. For the three depicted artistic styles, it is clear that the models are flawed in terms

of shape definition, but they are competent in the choice of colors, even in the variety of colors that a style contains and correctly separating grayscale images from colored images (examples in Figs. 4 and 5) (Fig. 6).



Fig. 4. First row of images contains real Pop Art painting and second row contains generated Pop Art painting.

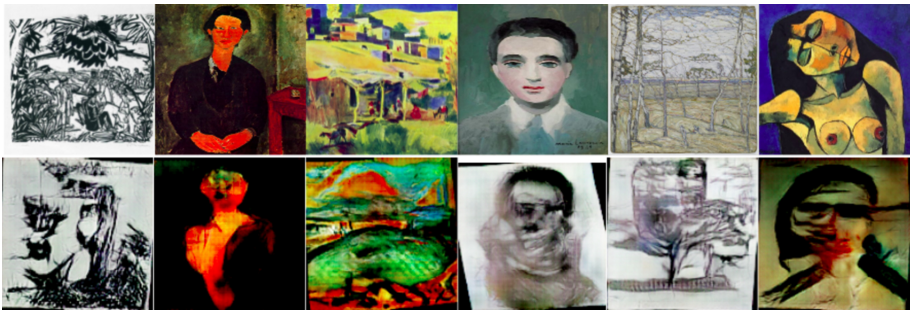


Fig. 5. First row of images contains real Expressionist painting and second row contains generated Expressionist painting.

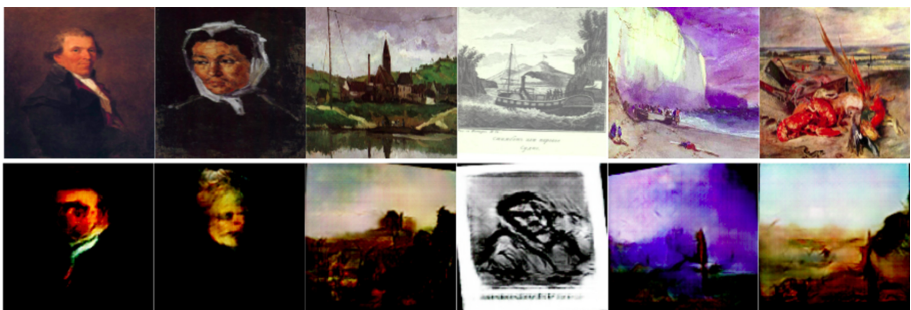


Fig. 6. First row of images contains real Romantic painting and second row contains generated Romantic painting.

## 5 Conclusion

Artwork style classification is quite challenging: class imbalance and high diversity within classes and similarities between classes are inevitable given existing art styles. Besides, classification performance does not correlate with dataset size. It is thus natural to look for data augmentation strategies; however, straightforward techniques do not generate images that can help distinguish between class styles.

In this paper we explored GANs for data augmentation in this setting; this is a little explored avenue and one that cannot be taken without some care. To be able to actually use GANs to generate artwork-like images with the desirable properties and diversity, we had to develop a working methodology, a specialized network architecture, and an adapted loss function. These technical contributions should be valuable in other settings where artwork is to be classified; moreover, they should be useful in data augmentation scenarios with many classes that are hard to differentiate and that are sensitive to popular existing techniques such as coloring and rotation.

We found that our approach and methodology can generate better classifiers. In particular, the combination of self-attention for mode diversity and the Wasserstein-GP loss function against vanishing gradient enabled us, in our experiments, to improve the augmentation of the images and consequently increase the accuracy of the model by almost 2 %.

**Acknowledgements.** This work is part of the Center for Data Science with funding by Itaú Unibanco. The first author thanks Itaú Unibanco for its generosity in authorizing research activities that led to this work. The second author was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grant 312180/2018-7, and by São Paulo Research Foundation (FAPESP), grant 2019/07665-4.

## References

1. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein gan. arXiv (2017)
2. Arora, R.S., Elgammal, A.: Towards automated classification of fine-art painting style: a comparative study. In: Proceedings - International Conference on Pattern Recognition, pp. 3541–3544 (2012)
3. Bar, Y., Levy, N., Wolf, L.: Classification of artistic styles using binarized features derived from a deep neural network. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8925, pp. 71–84. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-16178-5\\_5](https://doi.org/10.1007/978-3-319-16178-5_5)
4. Bianco, S., Mazzini, D., Napoletano, P., Schettini, R.: Multitask painting categorization by deep multibranch neural network. *Expert Syst. Appl.* **135**, 90–101 (2019). <https://doi.org/10.1016/j.eswa.2019.05.036>
5. Brock, A., Donahuey, J., Simonyan, K.: Large scale gan training for high fidelity natural image synthesis, pp. 1–35 (2018). arXiv
6. Cetinic, E., Lipic, T., Grgic, S.: Fine-tuning convolutional neural networks for fine art classification. *Expert Syst. Appl.* **114**, 107–118 (2018). <https://doi.org/10.1016/j.eswa.2018.07.026>

7. Chen, L., Yang, J.: Recognizing the style of visual arts via adaptive cross-layer correlation. In: MM 2019 - Proceedings of the 27th ACM International Conference on Multimedia, pp. 2459–2467 (2019). <https://doi.org/10.1145/3343031.3350977>
8. Chen, X., Duan, Y., Houthoof, R., Schulman, J., Sutskever, I., Abbeel, P.: Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In: Neural Information Processing Systems, pp. 2180–2188 (2016)
9. Chu, W.T., Wu, Y.L.: Image style classification based on learnt deep correlation features. *IEEE Trans. Multimedia* **20**, 2491–2502 (2018). <https://doi.org/10.1109/TMM.2018.2801718>
10. Condorovici, R.G., Florea, C., Vertan, C.: Automatically classifying paintings with perceptual inspired descriptors. *J. Visual Commun. Image Represent.* **26**, 222–230 (2015). <https://doi.org/10.1016/j.jvcir.2014.11.016>
11. Daras, G., Odena, A., Zhang, H., Dimakis, A.G.: Your local gan: designing two dimensional local attention mechanisms for generative models. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 14519–14527 (2020). <https://doi.org/10.1109/CVPR42600.2020.01454>
12. Eldeen, N., Khalifa, M.: Detection of coronavirus (covid-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest x-ray dataset (2020). <http://www.egyptscience.net>
13. Elgammal, A., Liu, B., Elhoseiny, M., Mazzone, M.: Can: creative adversarial networks, generating “art” by learning about styles and deviating from style norms, pp. 1–22 (2017). arXiv
14. Elgammal, A., Liu, B., Kim, D., Elhoseiny, M., Mazzone, M.: The shape of art history in the eyes of the machine. In: 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, pp. 2183–2191 (2018)
15. Farthing, S.: *Tudo sobre Arte*. 2 edn. (2018)
16. Florea, C., Toca, C., Gieseke, F.: Artistic movement recognition by boosted fusion of color structure and topographic description. In: Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017, pp. 569–577 (2017). <https://doi.org/10.1109/WACV.2017.69>
17. Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., Greenspan, H.: Gan-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* **321**, 321–331 (2018). <https://doi.org/10.1016/j.neucom.2018.09.013>
18. Gao, X., Tian, Y., Qi, Z.: RPD-GAN: learning to draw realistic paintings with generative adversarial network. *IEEE Trans. Image Process.* **29**, 8706–8720 (2020). <https://doi.org/10.1109/TIP.2020.3018856>
19. Goodfellow, I., et al.: Generative adversarial networks. *Commun. ACM* **63**, 139–144 (2014). <https://doi.org/10.1145/3422622>
20. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.: Improved training of wasserstein gans, vol. 2017-Decem, pp. 5768–5778 (2017)
21. Han, C., et al.: Synthesizing diverse lung nodules wherever massively: 3d multi-conditional gan-based ct image augmentation for object detection. In: Proceedings - 2019 International Conference on 3D Vision, 3DV 2019, pp. 729–737 (2019). <https://doi.org/10.1109/3DV.2019.00085>
22. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 7132–7141 (2018). <https://doi.org/10.1109/CVPR.2018.00745>
23. Karayev, S., et al.: Recognizing image style. In: BMVC 2014 - Proceedings of the British Machine Vision Conference 2014, pp. 1–20 (2014). <https://doi.org/10.5244/c.28.122>

24. Kastan, D.S., Farthing, S.: *On Color*. Yale University Press, New Haven (2018)
25. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2012). <https://doi.org/10.1145/3065386>
26. Lecoutre, A., Negrevergne, B., Yger, F.: Recognizing art style automatically in painting with deep learning. *J. Mach. Learn. Res.* **77**, 327–342 (2017)
27. Mirza, M., Osindero, S.: Conditional generative adversarial nets, pp. 1–7 (2014)
28. Qin, Z., Liu, Z., Zhu, P., Xue, Y.: A gan-based image synthesis method for skin lesion classification. *Comput. Methods Prog. Biomed.* **195** (2020). <https://doi.org/10.1016/j.cmpb.2020.105568>
29. Rodriguez, C.S., Lech, M., Pirogova, E.: 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 - Proceedings, pp. 1–7 (2019). <https://doi.org/10.1109/ICSPCS.2018.8631731>
30. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv 2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 4510–4520 (2018)
31. Sandoval, C., Pirogova, E., Lech, M.: Two-stage deep learning approach to the classification of fine-art paintings. *IEEE Access* **7**, 41770–41781 (2019)
32. Shamir, L., Macura, T., Orlov, N., Eckley, D.M., Goldberg, I.G.: Impressionism, expressionism, surrealism: automated recognition of painters and schools of art. *ACM Trans. Appl. Percept.* **7** (2010). <https://doi.org/10.1145/1670671.1670672>
33. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J. Big Data* **6**, 1–48 (2019)
34. Suh, S., Lee, H., Lukowicz, P., Lee, Y.O.: Cegan: classification enhancement generative adversarial networks for unraveling data imbalance problems. *Neural Netw.* **133**, 69–86 (2021)
35. Tan, M., Le, Q.V.: Efficientnet: rethinking model scaling for convolutional neural networks. In: 36th International Conference on Machine Learning, ICML 2019 2019-June, pp. 10691–10700 (2019)
36. Tan, W.R., Chan, C.S., Aguirre, H.E., Tanaka, K.: Ceci n’est pas une pipe: a deep convolutional network for fine-art paintings classification. In: Proceedings - International Conference on Image Processing, ICIP 2016-August, pp. 3703–3707 (2016). <https://doi.org/10.1109/ICIP.2016.7533051>
37. Tomei, M., Cornia, M., Baraldi, L., Cucchiara, R.: Art2real: unfolding the reality of artworks via semantically-aware image-to-image translation, vol. 2019-June, pp. 5842–5852 (2019). <https://doi.org/10.1109/CVPR.2019.00600>
38. Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., Pinheiro, P.R.: Covidgan: data augmentation using auxiliary classifier GAN for improved covid-19 detection. *IEEE Access* **8**, 91916–91923 (2020). <https://doi.org/10.1109/ACCESS.2020.2994762>
39. Wang, Z., She, Q., Ward, T.E.: Generative adversarial networks: a survey and taxonomy, pp. 1–41 (2019). arXiv
40. Wu, J., Huang, Z., Thoma, J., Acharya, D., Van Gool, L.: Wasserstein divergence for GANs. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11209, pp. 673–688. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01228-1\\_40](https://doi.org/10.1007/978-3-030-01228-1_40)
41. Xue, A.: End-to-end chinese landscape painting creation using generative adversarial networks, pp. 3863–3871 (2020). arXiv

42. Özal Yıldırım, Pławiak, P., Tan, R.S., Acharya, U.R.: Arrhythmia detection using deep convolutional neural network with long duration ecg signals (2018)
43. Zhang, H., Goodfellow, I., Metaxas, D., Odena, A.: Self-attention generative adversarial networks. In: Chaudhuri, K., Salakhutdinov, R. (eds.) Proceedings of the 36th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 97, pp. 7354–7363. PMLR (2019)
44. Zhong, S., Huang, X., Xiao, Z.: Fine-art painting classification via two-channel dual path networks. *Int. J. Mach. Learn. Cybern.* **11**, 137–152 (2020)
45. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks, vol. 2017-October, pp. 2242–2251 (2017)
46. Zhu, Y., Ji, Y., Zhang, Y., Xu, L., Zhou, A.L., Chan, E.: Machine: the new art connoisseur (2019). <http://arxiv.org/abs/1911.10091>