IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

Artistic Style Generation using Cycle GAN

Venkatesh Bellale[1], Suryakant Kumar Kashyap[2], Vinay Rawat[3], Nikhil Shinde[4], Prof. Rashmi Kale5
Students, Department of Computer Engineering[1,2,3,4]
Faculty, Department of Computer Engineering[5]
Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India
bellale15dm@gmail.com[1], suryakantk94@gmail.com[2], vinay.rawat058@gmail.com[3], nikhilshinde662@gmail.com[4]

Abstract: *Every once in a while we have always dreamt about being great artists and create famous artworks like M.F Hussain, Leonardo Da Vinci, Vincent Van Gogh etc. Unfortunately not everyone can produce such great artworks. However nowadays, such art styles can be generated creatively due to enhancement of various multiple computer vision techniques, especially like GAN (Generative Adversarial Network). In this project, our main motive is to focus on generating Artistic images majorly Indian art style. Also, our project aims to improve the accuracy of the model through use of various Deep Learning techniques and models a much more cleaner and pre-processed data set.*

Keywords: Machine Learning, Generative Adversarial Net- works, CGAN, Deep Neural Networks, Web Application

I. Introduction

For generating artificial output images with certain artistic style, we have come across a artistic style generating algorithm: CycleGAN. CycleGAN has the capability of predicting the relationship amongst multiple given datasets which consists of unpaired images, further creating a way for unpaired image-to-image translation.It is majorly used in facial recognition tasks in terms of transfer learning as well as generaluse of image classification tasks.When given any content withan image as an input, our CGAN algorithm outputs an image with a general Artistic-style filter.

DNN (Deep neural networks) have recently played a major trans-formative role in advancing AI (Artificial Intelligence) across various computing domains. In general, different gen- erative deep networks have been proposed by many researchers that have the ability to generate images to match or surpassa given training distribution. Generative Adversarial Networks (GAN) have been quite helpful in achieving this goal .How- ever, we argue that such networks are do not have a lot of potential to generate more creative products in their original design.

We demonstrate that this approach is effective at synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images, among other tasks. Eventually, our goal is to find out more about Cycle GAN method in terms of the architecture, the quality of the generated artworks and the efficiency of the whole generating process to figure out how Cycle GAN can help us produce paintings like those of famous artists.

II. Data And Pre-Processing

Pre-processing generally means the transformations that are applied to our dataset before feeding it to the Machine Learning/Deep Learning algorithm. Data Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis.

We collected Indian Style traditional paintings from variouswebsites over the internet.In total we collected more than 2000 images out of which 80 percent were used to train the Cycle GAN model and 20 percent were used to test the model.

We have checked the quality of datasets, and removed duplicated images or ruptured ones, which is good for data pre-processing. For our CGAN model, to resolve the problem of having varied resolution across images, we converted our input images into size of 128 x 128 pixels, in JPG format. For pre-processing purposes we have used different Deep learningand Machine Learning libraries like Keras and Numpy.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

III. METHODS

3.1 Architecture

A CycleGAN model architecture consists of two generator and two discriminator models. Both the models are trained simultaneously. It is a unsupervised machine learning technique using unpaired images dataset. The basic structureis same for our CycleGAN model to generate the paintings similar to that of Artists. CycleGAN generator in our model comprises of convolutional layers and resnet blocks with activation function as ReLu which takes input as a data set of images and generates the output image while the discriminatormodel is made up of convolutional layers only with activation function as LeakyReLu which takes the generated image as input and checks that image is a real or fake image.

The first generator model takes input as paintings from the dataset and generates them similar to photos dataset while the second one, takes input as normal pictures taken through camera as dataset and generates them similar to paintings. The final output shows real image converted into a generated image similar to painting and the painting converted into a normal picture (reconstructed). Each generator receives feedback from their respective discriminator which trains them to become better by every iteration. One generator receives additional data from the other generator throughout the training phase. This type of feedback ensures that a generator's image is cycle consistent, meaning that running two generators on the same image should yield the same result.

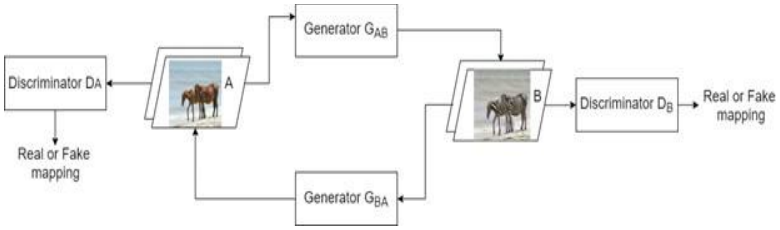

Figure 1: Our CycleGAN architecture model

3.2 Training

We utilised a learning rate of 0.0002, a mini-batch size of 30, and around five epochs in the training. The L1 loss is used to improve the generator by measuring the mean absoluteerror between input and output components. The resultant Generator Loss is a weighted loss that includes the validation, reconstruction, and identification losses. These numbers reflect how successfully our generator produces stylised pictures. Weused MSE Loss to calculate the mean squared error (squared L2 Norm) between the input components and the goal forour discriminators. We weighted the validation loss and the generator loss to get our discriminator 2 loss. This shows how successfully our discriminators distinguish between actual andmanufactured photos.

For our generative networks, we use the architecture pro- posed by Johnson et al. [12], who have demonstrated outstanding results in neural style transfer and super resolution. Three convolutions, many residual blocks [13], two fractionally- strided convolutions with stride 1 to 2, and one convolution that maps features to RGB are included in this network. For 128 * 128 pictures, we utilise 6 blocks, and for 256 * 256 and higher-resolution training photos, we use 9 blocks. We employinstance normalisation [14] in the same way as Johnson etal. [12] did. We employ 70 PatchGANs [15, 16, 17] for the discriminator networks, which are designed to determine whether 70 overlapping picture patches are real or fraudulent. A patch-level discriminator architecture has fewer parameters than a full-image discriminator and can fully convolutionally work on pictures of any size.

$$L(G, F, DX, DY) = LGAN(G, DY, X, Y) + LGAN(F, DX, Y, X) + Lcyc(G, F),$$

where controls the relative importance of the two objectives. We aim to solve:

$$G, F = argminG, FmaxDx, DY\, L(G, F, DX, DY).$$

To stabilise our model training approach, we use two strategies from recent research. To begin, we substitute the negative log likelihood goal in LGAN (Equation 1) with a least-squares loss. This loss is more consistent during train-ing and produces higher-quality results. We train the G to minimise $Expdata(x)[(D(G(x))1)2]$ and the D to minimise $Eypdata(y)[(D(y)1)2] + Expdata(x)[D(G(x))2]$ for a GANloss LGAN(G, D, X, Y). Second, we update the model according to Shrivastava et. al. technique's [11] to decreasemodel oscillation.

Instead of using pictures created by the most recent gen- erators, discriminators use images from the past. The 50

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

previously produced photos are stored in an image buffer. Witha batch size of 1, we employ the Adam solver. With a learning rate of 0.0002, all networks were trained from start.

IV. Results and Analysis

We have trained the CGAN(CycleGAN) architecture for more than 100 epochs, which takes about 10 hours to com- plete. In general, we can say that the images generated are good combinations of arts and reality. As training progresses, this model keeps optimizing the color and the texture of the generated image. Surprisingly, the generated images from our model still possess a bit of realism.

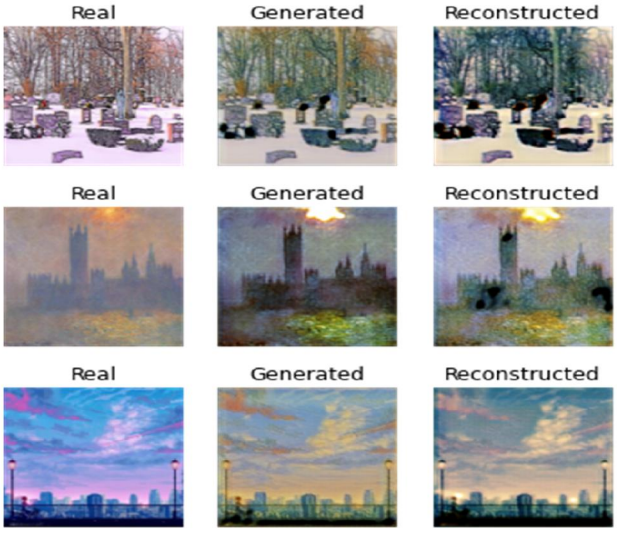

Figure 2: Photo to Monet and Monet to Photo Conversion using our CycleGAN model

We can see the results of our model in the images above(Fig.2 and 3).The generated image has an artistic style flavor init.

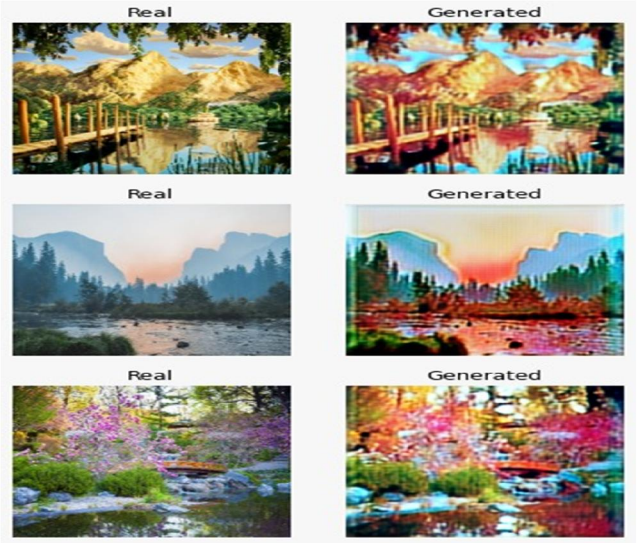

Figure 3: Photo to Rajasthani Style Conversion using our CycleGAN model

What we can analyse from the results is that the model produces an good accuracy for the generated image. The reconstructed images also have a glimpse of the original imagewhich have been reconstructed from the generated ones. In figure 3 we can see the images that have been generated onour Rajasthani images dataset. The images generated have an great accuracy and on this basis we can say that our modelwas trained successfully.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

V. Conclusion And Future Works

We implemented the Cycle GAN model which targeted on producing Indian style artworks which we pre-trained for almost 10 hours on a dataset of more than 2000 images. We noticed that, pictures generated by our CGAN model look more realistic and are regarded as artistic photos by most people during the survey which we conducted among our Batch mates. In a nutshell we can conclude that Cycle GAN model can serve as a good artistic style generator for photos based on various styles and filters.

For future work, we look forward to explore more different CV (Computer Vision) architectures, most probably some other GAN(Generative Adversarial Network) architecture. Along with it, we would like to change the dataset to some different art forms of different artists like M.F. Hussain, Da Vinci or even some domestic painters and various traditional art forms. If the dataset is relatively small in size we may try using transfer learning to convert the image. For better and simpler use of the model we will also be creating a web application with a easier User Interface. Finally, we would be focusing on improving our model and accuracy.

References

[1]. Best Artworks of All Time. https://www.kaggle.com/ikarus777/best- artworks-ofall-time. Accessed: 2020-09-29.
[2]. ID score for PyTorch https://github.com/mseitzer/pytorch-fid. Accessed: 2020-11-10.
[3]. Leon Gatys, Alexander S. Ecker, and Matthias Bethge. A Neural Algo- rithm of Artistic Style. 2015. arXiv: 1508.06576 [cs.CV].
[4]. Generate Art using Fast Style Transfer in a second. https://www.kaggle.com /tarunbisht11/generate-art-using-fast-style-transfer-in-a- second.
[5]. I'm Something of a Painter Myself. https : / / www . kaggle . com / c / gan - getting - started/data. Accessed: 2020-10-30.
[6]. Phillip Isola et al. Image-to-Image Translation with Conditional Adver- sarial Networks. 2018. arXiv: 1611.07004 [cs.CV].
[7]. Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2015. arXiv: 1409.1556 [cs.CV].
[8]. Yaxing Wang et al. Transferring GANs: Generating Images from Limited Data. Vol. 6. ECCV, 2018, pp. 220–236. arXiv: 1805.01677 [cs.CV].
[9]. Jun-Yan Zhu et al. Unpaired Image-to-Image Translation using Cycle- Consistent Adversarial Networks. 2020. arXiv: 1703.10593 [cs.CV].
[10]. Martin Heusel et al. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. 2018. arXiv: 1706.08500v6 [cs.LG].
[11]. A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb. Learning from simulated and unsupervised images through adversarial training. In CVPR, 2017.
[12]. J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In ECCV, 2016.
[13]. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016.
[14]. D. Ulyanov, A. Vedaldi, and V. Lempitsky. Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022, 2016.
[15]. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Imageto-image translation with conditional adversarial networks. In CVPR, 2017
[16]. C. Li and M. Wand. Precomputed real-time texture synthesis with markovian generative adversarial networks. ECCV, 2016.
[17]. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cun- ´ ningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image superresolution using a generative adversarial network. In CVPR, 2017. 5