

A Novel Technique to Regenerate Sculpture Using Generative Adversarial Network

S A K Jainulabudeen^{a,1}, H Shalma^b, S Gowri Shankar^c, D Anuradha^d,
K Soniya^e

^{a,b,d}Assistant Professor, Panimalar Engineering College, Chennai

^cAssociate Professor, ^eIEEE Member

Abstract. Dancing, music or any format of art has been a prominent thing from the past centuries. The many dynasties ruled the nation for centuries but every king encouraged the art one way or the other. The present day is just a minute part of the finest part of that era of art; the art of any form had been lost in the shadows to redeem the lost art we are going to use the latest technology like machine learning and artificial intelligence. The art lovers of the present age can seek the knowledge of lost art through this modern day technology. The retrieval of this art can only be done if there is a possibility to learn their language which helps in reading the old sculptures or the paintings on the walls of the ancient architecture. Now using the present day technology we are going to recoup that lost art through reading the walls of those structures where the art has been hidden for centuries. So at present we do not allow the art to continue to fall into shadow and extinguish later on, thus in this paper we present a DC-GAN model which has been created to inherit all the artistic skills of our ancestors by training from the key images of art designed as sculptures by our forefathers.

Keywords. Sculpture, Machine Learning, Artificial Intelligence, DC-GAN model

1. Introduction

The Ancient sculpture has its own prominence in the present day world as each artifact of the past age is the key or solution for an immense knowledge. The ancient era used the sculpture or paintings as the mode of communication with the future generations about the immense knowledge they have gained in the field of art the retrieval and study of each sculpture can provide colossal collection to treasure of the art which has been grounded since a stretch of years. Therefore the utilization of technology not solely preserves these sculptures script given to us by our ancestors however additionally create humans obtain from the sculpture within the same means the forefathers tend to convey the data of art for us.

A man-made neural network is an Associate Interconnected node cluster that cherishes the large neuron network within an exceedingly high brain. Each circular node as shown in Figure 1 represents an artificial neuron and an arrow represents a link between one artificial neuron's output and another's input. The prototype computer is called model DC-GAN [1,3]. The key and primary goal of this DC-GAN model is to inherit all the artistic skills that our ancestors had through training with the

images of the sculptures that they had made. We need the additional technical update to which we use the principle of machine learning to modulate ANN to CNN [1,2] (Convolutional neural network) in order to construct this model fully, also stated as ConvNet.

The CNN is a class of deep as shown in Figure 2, feed-forward artificial neural networks used for the analysis of visual images. The CNNs use a variety of multilayer perceptron's that were only designed to reduce pre-processing. These are also called artificial neural networks as shift variant or space variant, where the name suggests the role of such networks. Such networks are therefore based on the characteristics of their shared-weight design and invariance of translation. Compared to the other image classification algorithms, CNNs also use relatively far less pre-processing. This says the network is learning the filters that have been hand-made in traditional algorithms.

2. Related Works

In this paper the system proposed is to inherit the artistic skills of our ancestors who have preserved that in the form of sculptors in the ancient structures. In paper [1] Jiale Zhi a tool namely pixel brush has been inherited onto this proposed system due to its efficiency in paper [2] Kavitha et al. proposed generating artwork from text descriptions using conventional deep learning and also the effective photo-realistic image generations. The main note of paper [3] Ahmed Elagammal et al. is CAN (Creative adversarial networks) is the major system which generates art by looking at art sculptured on the ancestral structures and also learns about styles. The problem in the paper [4] M. Arjovsky et al proposes is about the unsupervised learning, nothing but it states about seeking probability distribution and also solution stating for the problem as problem density which results to decline the KL divergence. In the paper [5] I.J. Goodfellow et al the author proposes a new framework for estimating the generative models through an adversarial process, in which we train two models simultaneously. The stacked principle of generative adversarial network was discussed by author in paper [6] X.Huang et al. The author has suggested a novel generative model called the SGAN model, which has a main role and is also equipped to invert bottom-up disseminative network hierarchical representations. The paper Han Zhang et al [7] clearly explains about the introduction of the SGANs which are Stack GAN. The main reason behind the introduction of these SGANs is to overcome the challenges faced in generating in high quality images in GANs. In paper K. Cho et al [8] the learning phrase representation using RNN Encoder-Decoder for statistical machine translation has been proposed making it a new additional type of other neural network model. The concept proposed in the paper P. Isola et al [9] is about the image to image translation with conditional adversarial networks. This paper comes up with a solution for image to image translation problems as conditional adversarial networks. The up gradation to the available system has been proposed by the paper A. Nguyen et al [10] plug and play generative networks.

3. Proposed Method

The method stated in this paper is to bring back the artistic skills which our ancestors

had done exceptionally well in olden days by replicating the pattern and knowledge of the art which had been completely saved till date in the form of inscriptions on the walls great structures, monuments, etc. The basic working of proposed method is primarily to read the sculpture and work it with the dataset available, if any of the art in the dataset matches then it revises and joins to fulfill the complete artistic form, if not it puts the newly found art into the dataset for future purposes.

The mechanism used for development of this device is GAN which is known as General Adversarial Networks (GAN). GAN architecture proves to be unique and efficient enough for the retrieval process. Thus we prefer GAN over all the other methods to be implemented. The GAN has its own identical as well as unique architecture to build. Generative adversarial networks (GANs) [5,6] are deep neural network architectures comprising of two nets, denting one against the other. One neural network, called as generator [5,6], generates new data instances, while the other, called as discriminator [5,6], evaluates them for authenticity or correctness or also accuracy; that is the discriminator decides whether or not each instance of the data assessed is from the actual training dataset [7] or belongs to it. We're trying to do something more banal here than imitating the Mona Lisa, an art designed by great painter Leonardo Da vinci. Now we will generate handwritten numerals like those found in the MNIST dataset from the real world. The generator, meanwhile, does its job of creating new images [9] which it passes on to the discriminator. In this method, it does say that it was said to be evaluating to prove that the retrieval is authentic, although it is fake. The generator's goal is to generate passable or traceable handwritten digits, to lie without being caught. The discriminator's goal is to identify images that come from the generator as fake. In a field such as Computer Vision [8], which has been explored and studied for a long time, a recent addition was GAN which immediately became a new standard for training machines.

Here are the steps a GAN takes which have been discussed above in detail as shown in Figure 1.

1. The generator takes in numbers at random and returns an image.
2. The discriminator receives this generated image alongside a stream of images taken from the actual dataset.
3. The discriminator takes both real and fake images and returns probabilities; probability is the number between 0 and 1, with 1 representing a prediction of authenticity or accuracy and 0 representing fake or unrealistic image.

So for more precise information regarding the image retrieval we have designed a double feedback loop:

4. The discriminator is with the ground truth of the images in a feedback loop, which we know is already available.
5. The generator is in an interlocked feedback loop with the discriminator.

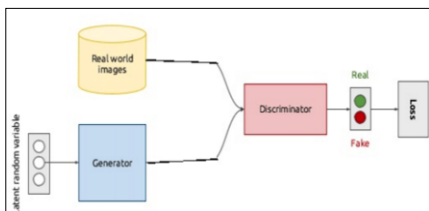


Figure 1. GAN Architecture

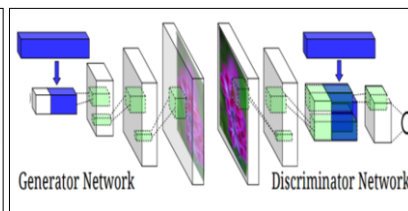


Figure 2. GAN Structure

As shown in Figure 2, the generator is an inverse convolutional network, in a sense declaring itself as the identity of the images to be authenticated. While a regular convolution classifier takes samples of an image and down as a dataset to generate a probability, the generator takes a random noise vector and samples it to a picture. The first throws down data through sampling techniques such as max pooling, and the second generates new, unique data from above.

4. System Design – GAN Structure

A Generative Adversarial Network (GAN) has two parts as shown in the Figure 3. The Generator which learns to generate conceivable data, generated data becomes deleterious training examples for the discriminator; it learns to distinguish the generators forged data from tangible data.

4.1 Discriminator

Through training, the discriminator network learns a function to tell the difference between the real and generated data [10,11]. The first features the discriminator learns to look for may be relatively obvious aspects of the data which easily separate the real from the fake. Once the discriminator has learned something about what to use to tell the two apart, the generator can take advantage of what the discriminator has learned in order to learn for itself.

Layer 1: Convolutional 2D layer having 64 filters with strides 2 and kernel size as 5 and having input shape 100x100x1. The activation feature used here is Leaky ReLU and 0.4 dropout function. It is termed as input layer.

Layer 2: Convolutional 2D layer having 128 filters with strides 2 and having input shape 50x50x128. Activation feature used here is Leaky ReLU and a 0.4 dropout function.

Layer 3: Convolutional 2D layer having 256 filters with strides 2 and having input shape 25x25x1. Leaky ReLU and a dropout function of 0.4 is the activation function used here.

Layer 4: Convolutional 2D layer having 512 filters with strides 2 and having input shape 13x13x512. The activation feature used here is the Leaky ReLU and a 0.4 dropout function.

Layer 5: Fully Connected layer having 86528 neurons which is flattened from the previous layer.

Layer 6: Dense layer having 1 neuron with an activation function as Sigmoid [4].

4.2 Generator

The generator will attempt to generate fake images which will fool the discriminator into thinking they are real. And when an image is fed, the discriminator will try to distinguish between a real image and a generated one as best as it can.

Layer 1: A dense layer contains 160000 neurons with input dimension 100 that is Noise vector and Batch Normalization is carried out and activated using ReLU. The output of this layer is reshaped into 25x25x256. Dropout of 0.4 is applied. Here too this layer is termed as input layer.

Layer 2: 2D up sampling is done and Inverse 2D Convolution is applied with 128 filters and kernel size of 5 and out shape is 50x50x128. Batch Normalization and ReLU activation function is applied.

Layer 3: 2D up sampling is done and Inverse 2D Convolution is applied with 64 filters and kernel size of 5 and out shape is 100x100x64. Batch Normalization and ReLU activation function is applied.

Layer 4: Inverse 2D Convolution is applied with 32 filters and kernel size of 5 and out shape is 100x100x32. Batch Normalization and ReLU activation function is applied.

Layer 5: Inverse 2D Convolution is applied with 1 filters and kernel size of 5 and out shape is 100x100x1. Sigmoid activation function is applied. This layer produces a generated grayscale image of size 100x100 pixels. Here this layer makes itself as output layer.

5. Experimental Result

Typically, the concept of machine learning works with two sets that are said to be mostly Training and Testing [11,14]. The result generated will verify that the network correctly recognizes images or at least recognizes a percentage of them if you don't get accurate predictions, go back to the training set, look at the hyper parameters you used to tuning the network [12,13]. As well as your data quality and look at the pre-processing techniques as result shown in Figure 3. The data set contains 800 grayscale images of 100x100 pixels in size. To make learning simple and easy, the images are all uniform between 0 and 1.

5.1 Activation Function – Sigmoid, Rectified Linear Unit (ReLU), Leaky ReLU

Sigmoid is used because this function exists between (0 and 1) [11,12], it is therefore particularly used for models to predict the likelihood as associated output. Since there is only a chance of anything between 0 and 1, sigmoid is the right choice.

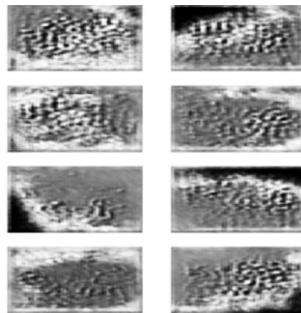


Figure 3. Generated Image

6. Conclusion and Future Work

The stated paper will ensure inheriting the artistic skills had been preserved by our ancestors in the form of sculptures designed by sculptors on the great monumental walls. The model creates its own sculpture by the learning from the previous sculpting

works by the ancestors. On given noise the sculpture is generated, though the generated image is not that clear containing noise in it, on optimizing hyper-parameters a stable model can be created thus a clear picture of sculptures can be generated. A same model of making can be made using Conditional GAN in which an image is given as an input and can expect a model to give a sculpted version of the given image. This project can be enhanced by convolutional 3D and a 3D model of the sculpture can be generated and can be 3D printed.

References

- [1] Jiale Zhi “PixelBrush: Art Generation from text with GANs”, Stanford University, pg: 1-9, jz2@stanford.edu, 2017.
- [2] Kavitha Subramani, Murugavalli S “Recognizing Ancient Characters from Tamil Palm Leaf Manuscripts using Convolution Based Deep Learning”, IJRTE, Vol.10, pp. 6873-6880, Sep 2019.
- [3] Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, Marian Mazzone, “CAN: Creative Adversarial Networks Generating “Art” by Learning About Styles and Deviating from Style Norms”, 2017 ICC conference.
- [4] M. Arjovsky, S. Chintala, and L. Bottou. “Wasserstein GAN”, Jan 2017.
- [5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A.C. Courville, and Y. Bengio. “Generative adversarial nets”, 2014 Advances in neural information processing systems, Jun 2014.
- [6] X. Huang, Y. Li, O. Poursaeed, J. Hopcroft, and S. Belongie. “Stacked generative adversarial networks”, Dec 2016.
- [7] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, XiaoLei Huang, Dimitris N. Metaxas “StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks”, 2018 IEEE Transactions on Pattern Analysis and Machine Intelligence, July 2018.
- [8] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. “Learning Phrase Representations using RNN EncoderDecoder for Statistical Machine Translation”, Jun 2014.
- [9] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. “Image-to-image translation with conditional adversarial networks”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [10] A. Nguyen, J. Yosinski, Y. Bengio, A. Dosovitskiy, and J. Clune. “Plug & play generative networks: Conditional iterative generation of images in latent space”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [11] A. Odena, C. Olah, and J. Shlens. “Conditional image synthesis with auxiliary classifier GAN”, Volume 1, Sept 2017.
- [12] C. Dong, C. C. Loy, K. He, and X. Tang. “Learning a deep convolutional network for image super-resolution”, 2014 European Conference on Computer Vision (ECCV), pp 184-199, volume 8692.
- [13] A. Dosovitskiy and T. Brox. “Generating images with perceptual similarity metrics based on deep networks”, Feb 2016.
- [14] S. Gu, W. Zuo, Q. Xie, D. Meng, X. Feng, and L. Zhang. “Convolutional sparse coding for image super-resolution”, 2015 IEEE Conference on Computer Vision (ICCV), 2015, pp: 1823-1831.