International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 282-285 www.ijaem.net ISSN: 2395-5252

Anybody Can Draw: Generating Face Photographs From Free-hand Sketches.

Shikha K [1], Bhavana S [2]
[1], [2] Students
Department of Information Technology
B.K. Birla College of Arts, Science and Commerce (Autonomous) Kalyan, Thane, India.

Submitted: 30-08-2021 Revised: 03-09-2021 Accepted: 05-09-2021

ABSTRACT: If you've ever tried to draw a face, you'll know it can be a complicated process. Unlike animals or objects, people are unique and have very defined features that make them recognizable. Movies have improved the way people think about faces but still can't compare to what humans see from another person's perspective. The world of art is seeing an evolution in technology called deep learning that is finally starting to produce the most realistic human images yet. In this paper, we focus on creating realistic face photographs from free hand sketches using RNN and DiscoGAN. RNN will be completing the half drawn face sketch and DiscoGAN will be translating the face sketch to the realistic face photograph. The RNN consists of encoder and decoder in which the encoder takes the half drawn sketch and passes it to the decoder and the decoder being autoregressive completes the sketch by taking the previous inputs. After completing the sketch the Sketch is fed into the DiscoGAN, whose circuit consists of a generator and a discriminator. The DiscoGAN will translate the sketch into real face photograph. It also follows transformation function in which if x is translated into y then after reversing the model y should be able to form x.

KEYWORS: RNN, DiscoGAN, encoder, decoder, generator, discriminator, Realistic face photographs, hand drawn sketches.

I. INTRODUCTION:

Face sketch synthesis has been a buzz in computer vision and pattern recognition area, because of its significance in entertainment industry and Law enforcement agencies. For recognition of the face sketch many deep learning algorithms are used widely. Here we will be using Generative Adversarial Networks, GANs, for generating realistic face images from its corresponding sketch. Generative Adversarial Networks are generative modeling approach which uses deep learning methods like convolution neural networks. The generative model is categorized into two sub-models: The Generator model and the discriminator model. The generator model would be the model which we will train to generate new examples, whereas, the Discriminator model will try to categorize examples into real or fake i.e. from the domain or generated. The different types of GANs which are widely used are Pix2Pix GANs, deep con-evolutional GAN, etc. there we use DiscoGAN (Discover Cross-domain Relations). Recently, DiscoGAN became very popular because it could learn cross-domain relations even by giving unsupervised data. For us i.e. humans, learning cross-domain relations are very natural. We can easily figure out how two different things are related to each other's by their images. But for a Machine Learning model to figure such unrelated images or cross- domain relations is a different task. Still, in recent times Disco GAN had gave trustworthy results in learning cross-domain relations. DiscoGAN's and cycleGAN both learn two individual transformation functions. One architecture learns a transformation from domain X to domain Y and the other architecture learns a reverse mapping. But they both use reconstruction loss as a measure of how perfect the original image is reconstructed after two times transforming across domains. Both, DiscoGAN and CycleGAN follow the principle which states that if we transform an image from domain one to domain two and then back to domain one, then it should match the original image.

II. PROPOSED METHOD:

We will first complete an uncompleted or half drawn sketch to a relevant sketch matching the facial attributes such as eyes, mouth, and nose using neural networks and then will use the network architecture having generator, discriminator by DiscoGAN.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 282-285 www.ijaem.net ISSN: 2395-5252

a) Half drawn sketch to complete sketch: For completing the sketch we will use RNN i.e. Recurrent Neural networks. RNN is the first algorithm which remembers its input, because of an internal memory.

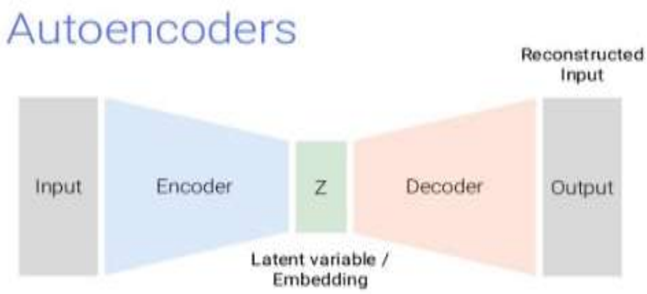

Fig. 1: Sketch-RNN

Auto encoder of the network is bidirectional RNN.

It will first take a half drawn face sketch as input and will pass it to the encoder and the auto encoder will scan the half drawn sketch. The sketch sequence (S) and its reverse sequence (Sr) are given to the two encoding RNNs. Simultaneously this makes the bidirectional network. We will get two hidden states as the result as h and hr. These hidden states will be projected to two Vectors μ and $\sigma^{\wedge}$. Considering z as an arbitrary parameter will be formulated with μ and $\sigma^{\wedge}$.

$$\boldsymbol{\sigma = \exp(\sigma^{\wedge}/2)}$$

Where, $\boldsymbol{\mu = W\mu\, h + b\mu}$

& $\boldsymbol{\sigma = W\sigma h + b\sigma}$

Here, z is conditioned on the input sketch. The decoder is an autoregressive RNN. The first hidden state of RNN is :

$$\boldsymbol{h_0 = \tanh(W_z Z + b_z)}$$

Now, we will feed so, h_0 & Z together to the decoder circuit as the first input. Now, the decoder being the regressive circuit will beefed with the output of the first one i.e. S, as the input for the second block. This will continue till the completion of the sketch.

b) Sketch to realistic photograph: After completion of sketch we need to translate that sketch into realistic photograph. For this, we will use DiscoGAN. A DiscoGAN is a generative adversarial network. In GANs, the generator and discriminator receive some extra conditional input data. In GANs, the output image that is generated with the generator network is random, which means, it might generate images of any face that was there in dataset. But with DiscoGAN, we can generate images what we want. If we want to generate a face it'll generate an image of their face. This is done by conditioning the network. DiscoGAN learns the mapping from the detected image x and the noise vector z, to the output y: G :{ x, y} -> y.

c) Architecture of the network: The input i.e. the incomplete sketch is fed into the sketch RNN which consists of an encoders and decoders; it will complete the half drawn sketch and pass it to the desired output i.e. realistic face photograph.

d) RNN (Recurrent Neural Networks): For completing the half drawn face sketch, the RNN decoder is used as the main stand-alone model. After giving the input once, we will remove the encoder and the decoder becomes the auto regressive model. The decoder first converts the sketch into a hidden state h. later, the other points of the sketches are found by keeping h as the initial hidden state.

e) DiscoGAN: A DiscoGAN is a generative adversarial network that generates face images in domain B given a sketch in domain A. Following diagram is an architectural diagram of a DiscoGAN which illustrates the above mentioned.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 9 Sep 2021, pp: 282-285 www.ijaem.net ISSN: 2395-5252

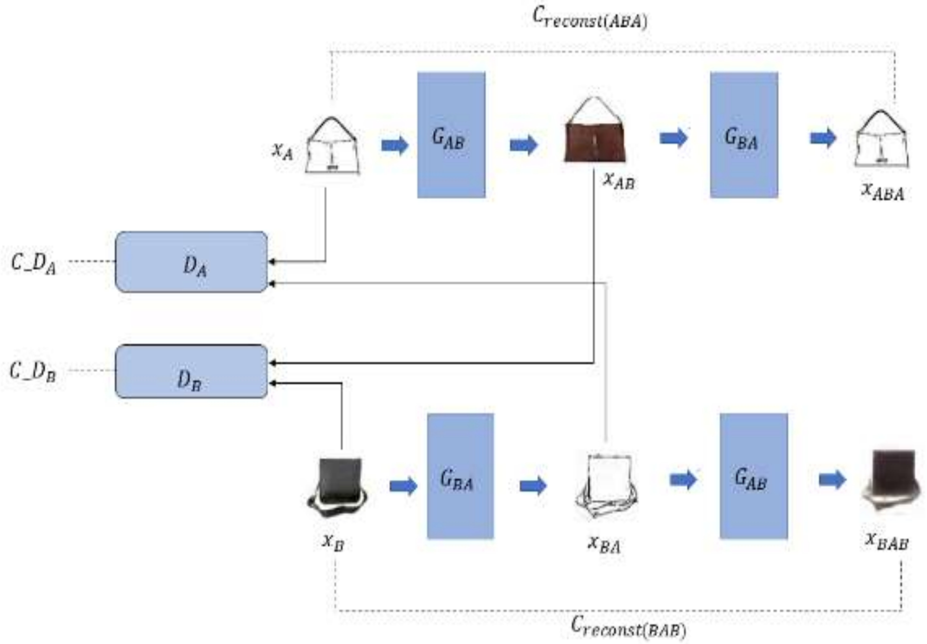

Fig.2: DiscoGAN Architecture

The face image generated in domain B resembles the face Sketches in domain A in both style and pattern. This relation can be learned without explicitly pairing images from the two domains during training. Now, as this is quite a powerful capability, pairing of items is a time consuming task. DiscoGAN tries to learn two generator functions of the form of neural networks gab and GBA. So that the face sketch X_A where fed through the generator G_A produces an images X_{AB} , that looks real in domain B. Also, where this image X_{AB} is fed through the other generator network $G_{BA,}$ it should exactly look like X_A.

III. LOSS FUNCTION:

DiscoGAN learns two individual transformation functions. One architecture learns a transformation from domain X to domain Y and the other architecture learns a reverse mapping. But it uses reconstruction loss function as a measure of how perfect the original image is reconstructed after two times transforming across domains. So to evaluate the loss functions we have, the total generator loss of DiscoGAN is given by: The total generator loss of DiscoGAN is given by,

generator loss = gan_loss + λ L1_loss

Were, L1_loss is that the mean absolute error between generated image and target image. λ is associate degree capricious worth that is taken as one hundred. The total differentiator loss is,

discriminator Loss = real_loss + generated_loss

Where, real_loss could be a sigmoid cross entropy loss of the real images associate degreed an array of ones. generated_loss could be a sigmoid cross entropy loss of the generated pictures associate degreed an array of zeros.

IV. THE GOOD SIDE:

Generating realistic images from sketches has been used for many purposes, like detecting drawings with bigoted content or recognizing people from photograph. Another important use is to recognize paintings from famous artists to determine their authenticity. There are many cases where people have used technology to take an artist's work; this system will eliminate human error in judging authenticity.

V. RESULT:

Our research has the following parts as mentioned earlier.

a) Sketch Completion: For completing the sketch we use recurrent neural networks which create the networks with loops in there, which allows it to persist the information. As we Require the content from previous input, we make use of RNN. RNN feeds the data from previous input into all stages; as a result it helps in the sketch completion process. At the output of the first stage, that is the RNN sketch completion, the input face sketch is completed. The completed sketch is received as the output.

b) Face sketch to realistic photograph : For translating the face sketch into realistic Photograph, we made use of DiscoGAN, which is generative adversarial network. We trained many face sketches. Later, a random incomplete hand drawn face sketches was given to the network which was first completed and then translated into its corresponding realistic face photograph.

VI. LIMITATION:

Although, we had images processing as our core concept of generating photographs, we could not process corresponding color coding for the photos. We only could shape the structure is black & white.

VIII. CONCLUSION:

Deep learning is showing a very promising future for computer vision or computer aided sketching and may be applied to many other problems in the future. Our system illustrates the image processing if any incomplete hand drawn face sketch is given. Further, our work can be improved by more accurate sketch completion and image processing units.

FUTURE WORK:

With further training, we can solely include the color coding in the photograph also. For this we will require a graphics card implemented in the system.

REFRENCES:

[1]. Yuhang Li, Xuejin Chen, Binxin Yang, Zihan Chen, and Zheng-Jun Zha. “DeepFacePencil: Creating Face Images from Freehand Sketches,” October 12, 2020, 9. https://doi.org/10.1145.
[2]. Jian Zhao, Xie Xie, Lin Wang, Meng Cao, and Miao Zhang. “Generating Photographic Faces from the Sketch Guided by Attribute Using GAN.” Research Gate XX (n.d.).
[3]. Xiaogang Wang and Xiaoou Tang. “Face Photo-Sketch Synthesis and Recognition” 31 (n.d.).
[4]. Nice Maria, Jayalakshmi V, J Ananthakrishnan, Tania Sam, and Deepa P L. “Sketch to Image Translation” 07, no. 07 (n.d.).
[5]. Mingming Hu and Jingtao Guo. “Facial Attribute-Controlled Sketch-to-Image Translation with Generative Adversarial Networks.” EURASIP Journal on Image and Video Processing, 2020. https:// doi.org/10.1186/ s13640- 020-0489-5.