International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

Semantic Image to Image Translation using Machine Learning Algorithms

G.Pranavi, M.Pranava Brunda, P.Jashwanth, V.Naga Chaitanya, D.Rajeswara Rao

Abstract: In semantic image-to-image translation, the goal will be to learn mapping between an input image and the output image. A model of semantic image to image translation problem using Cycle GAN algorithm is proposed. Given a set of paired or unpaired images a transformation is learned to translate the input image into the specified domain. The dataset considered is cityscape dataset. In the cityscape dataset, the semantic images are converted into photographic images. Here a Generative Adversarial Network algorithm called Cycle GAN algorithm with cycle consistency loss is used. The cycle GAN algorithm can be used to transform the semantic image into a photographic or real image. The cycle consistency loss compares the real image and the output image of the second generator and gives the loss functions. In this paper, the model shows that by considering more training time we get the accurate results and the image quality will be improved. The model can be used when images from one domain needs to be converted into another domain inorder to obtain high quality of images.

Keywords: Cycle GAN, Cycle consistency loss, Semantic, Translation

I. INTRODUCTION

The process of translation of one image to another image has made progress in recent times. It involves the process of translating the image from one domain to the other domain. In this paper Cycle GAN algorithm is used to translate the images. The semantic images to be translated may be either paired or unpaired. Paired data is the process of translating the sketch into its corresponding real time image. Unpaired data involves the translation of image of one kind to another kind.

In this paper, the proposed model uses cityscape dataset taken from Kaggle to illustrate the process of semantic image-to-image translation. Since the semantic image may not be clear at times, the translation of semantic image to real image is necessary. The objective of this model is to convert the semantic image which is in the form of painting or sketch into real or photographic image with more accuracy.

Revised Manuscript Received on April 30, 2020.

* Correspondence Author

G.Pranavi*, Department of CSE, VR Siddhartha Engineering College, Vijayawada, India, Email:pgogineni7799@gmail.com

M.Pranava Brunda, Department of CSE, VR Siddhartha Engineering College, Vijayawada, India, Email:mpranava4u@gmail.com

P.Jashwanth, Department of CSE, VR Siddhartha Engineering College, Vijayawada, India, Email:pothinajashwanth@gmail.com

V.Naga Chaitanya, Department of CSE, VR Siddhartha Engineering College, Vijayawada, India, Email:nagachaitanya.valeti@gmail.com

D.Rajeswara Rao, Department of CSE, VR Siddhartha Engineering College, Vijayawada, India, Email:rajeshpitam@gmail.com

For this purpose, the Cycle GAN algorithm along with cycle consistency loss is used. The cycle consistency loss is used for calculating the difference between the original image given as input and the output image obtained from the second generator. High accuracy is obtained by obtaining minimum generator and discriminator loss.

This model can be tested on various datasets like cityscape, maps, human faces etc. Compared to other approaches our paper has better accuracy and the translated images are clear. This model outperforms other models in a way that the translated images have minimum loss and the images are clear. The architecture of the model proposed in this paper is simple to understand and implement.

II. RELATED WORK

Ziqiang Zheng et.al.,[1] stated that a GAN algorithm can achieve the image translation between the images of various domains. He introduced a denoising image-to-image translation model for enhancing the process of image creation and maximize the model's robustness. The cascade loss and perturbed loss will enhance the content organising behaviour of the generator. There exists a gap between the photographic images and the semantic images. He included one hot semantic label map that forces the generator to give most importance for the generation of objects that overlap.

Pramuditha Perera et.al.,[2] stated that the model of image-to-image translation have made some important changes with the invention of GAN. Considering an input image from a specific domain, the objective of this model is to translate the image into another given or considered domain. Given several images, across n various domains, the objective of In2I is to yield the corresponding image in a particular domain. Recent works in research domain have learned the transformation across various tasks such as night to day time images, greyscale to colour images and so on.

Ting Chung Wang et.al.,[3] stated a new method for translating the real images from semantic using Conditional GAN. Here two main issues addressed are the complexity in generating images of high resolution with GAN's and the lack of details and realistic texture in obtained the high resolution images. He proposed a method to produce multiple results by giving the similar input, and accessing the users to modify the object appearance.

Feng Xiong et.al.,[4] stated that the image to image translations have gained awareness in many computer articles because of it's performance. It maps an image from specified kind of domain to the required domain. The goal of GAN is to generate images that trick the discriminator from distinguishing the real or fake images. The GANs are mostly used in image to image translations. To describe the effect of redundancy and noise in the images generated, the discriminator distinguishes if the image is real or not in the latent space. It is a challenge to obtain pixel level translation.

III. METHODOLOGY

In this paper, the semantic image-to-image translation is done on the cityscape dataset. The dataset is taken from Kaggle. The Cycle GAN algorithm is used for translating the images from one kind of domain to another kind of domain. The programming language used is Python. The platform used is Google Colab. As Google Colab can work with large workloads it is preferred for working with machine learning models.

A. Preprocessing

Compression technique is used to compress the images into similar size. Image resizing is done, so that all the images in the dataset are resized into same size of 256x256. Histogram equalization is used to increase the contrast of the images. The images are cropped and resized so that all the images in the dataset are of same size. Transformation of images is done inorder to convert the grayscale images to RGB images. This transformation is performed because images with better quality can be obtained. The resized images are then normalised. Normalisation is a process in which data is projected into a predefined range of either [-1,1] or [0,1].

B. Architecture of the model

The model is developed using pytorch. Here an image buffer is used which stores the previously generated images. The image buffer is used to update the discriminator using generated images. In this paper, convolution and deconvolution blocks and resnet blocks are used along with activation functions. Activation functions such as tanh, ReLu and sigmoid are used. A convolution network is able to capture the dependencies namely spatial and temporal dependencies in images. The convolution network reduces the images such that they are easier to process and do not loose features that are important for good prediction. The convolution layer has kernel which is matrix of weights. The kernel performs element wise multiplication and then sums up the result into a single output pixel. Convolution layer also uses two other techniques- padding and strides. Padding is a process of adding zeroes to the input matrix. Stride denotes the number of steps we are moving in each step. By default the stride value is one.
The type of activation functions utilized include Sigmoid, tanh and ReLu. The sigmoid activation function is in the form of a S shaped curve and it's range is between 0 and 1.Tanh function is called as Hyperbolic Tangent function. It's range is between -1 to 1. ReLu stands for Rectified Linear Units. ReLu is most commonly used function and it's range is from 0 to infinity. ReLu function learns faster than sigmoid and tanh functions.

C. Cycle Consistency Loss

The Cycle GAN algorithm involves training of two generators and two discriminators. The weights and biases are calculated using convolution and batch normalisation in the models of generators and discriminators. An Adam optimiser having a learning rate of 0.0002 is used for both the generator and discriminator. The activation function used in the discriminator model is Leaky ReLu.

In the Cycle GAN algorithm, one generator converts the images from first domain A to second domain B. The other generator will convert the images from second domain B to first domain A. Each generator has a corresponding discriminator which distinguishes the real images from synthesized images. The Cycle GAN has cycle consistency loss. In this paper, the cycle consistency loss for both the generators and discriminators model is determined. The generator loss is represented as G_AB and G_BA. The discriminator loss is represented as D_A and D_B.

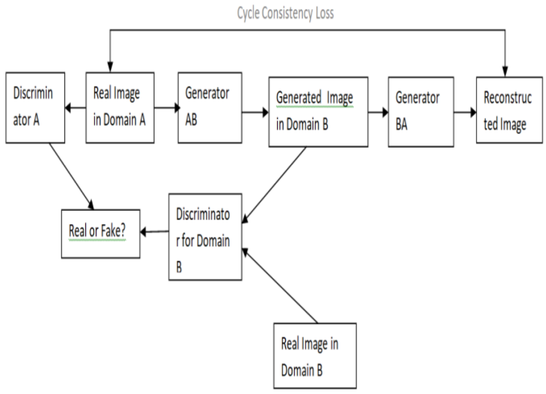
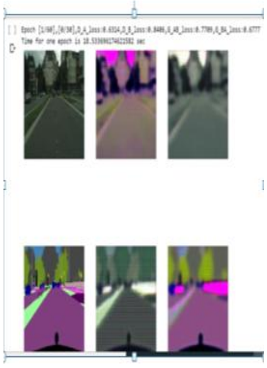
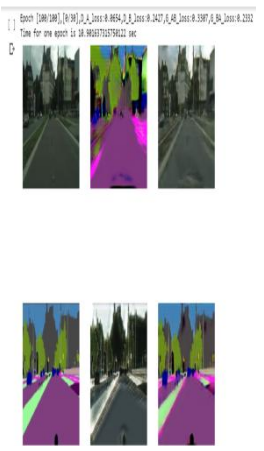
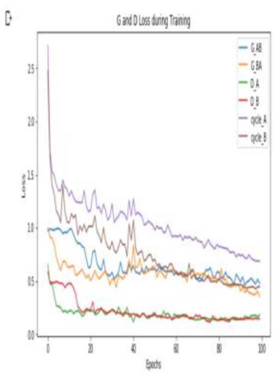


Fig 1. Cycle Consistency Loss

The cycle consistency loss is described as follows. Here, the output of the first generator is given as an input to the second generator, and the output image of the second generator must match the original image. This loss calculates the difference between the image given as input to the first generator and the second generator's output image. This is called as forward cycle of cycle consistency loss. The reverse can also be done to calculate the loss and this is called as backward cycle. The discriminator is fed with inputs by generator. The discriminator compares the image from the generator and the original input image, and distinguishes between the real and fake image. The cycle consistency loss is depicted in figure 1. In this paper, the model built uses the backward cycle to calculate the cycle consistency loss.
In this paper, the proposed model is built by considering the cityscape dataset. The dataset contains many images and we use 50 images for training.
A total of 100 epochs were run to train the model appropriately. As the number of epochs increases, the quality of image being translated also increases thereby reducing the loss. If we obtain minimum consistency loss, the accuracy of the model increases. In this model, at the end of the epochs the generator and discriminator loss obtained are D_A=0.0654,D_B=0.2427,G_AB=0.3307, G_BA=0.2332.
The epochs and the image translation are depicted in Table I.
After all the given epochs are completed, the loss functions and the translated images are depicted in the Table-I.
After all the epochs are completed the semantic image translated is most near to the real image. The generator and discriminator loss can also be depicted in the form of graph with loss on Y-axis and epochs on X-axis. The graph obtained is depicted in the third row of Table-I .

Table –I: Results obtained after execution of the model

Image	Description	Loss
	Epoch 1	D_A=0.6314 , D_B=0.8406 , G_AB=0.7709 , G_BA=0.6777
	Epoch 100	D_A=0.0654 , D_B=0.2427, G_AB=0.3307, G_BA=0.2332
	Generator and discriminator loss	

D. Experiments

The proposed model is also trained with the face dataset taken from CUHK website. Here, a total of 60 epochs were taken, to obtain the results. The generator and discriminator loss obtained were G_AB=0.3363,G_BA=0.5318,D_A=0.1054, D_B=0.3863. The translation is depicted in the figure 2.

Fig 2. Face dataset after 60 epochs

IV. CONCLUSION

In this paper, our proposed model is implemented using Cycle GAN algorithm. The cycle GAN algorithm uses the cycle consistency loss to calculate the loss function. This algorithm is best suited for translating images from one domain to another domain. If the loss obtained is minimum ,the quality of images translated will be maximum. In this paper, the proposed model has minimum amount of loss thereby increasing the quality of the image translated. There are many applications to the proposed system. Translating the Google map images to satellite images is one of the applications. The proposed model can be used where high quality of images are to be produced.

REFERENCES

1. Ziqiang Zheng, Chao Wang, Zhibin Yu, Haiyong Zheng,Bing Zheng, “INSTANCE MAP BASED IMAGE SYNTHESIS WITH A DENOISING GENERATIVE ADVERSIAL NETWORK”,IEEE Access,Volume 6, PP:33654-33665,IEEE 2018
2. [2] Pramuditha Perera, Mahdi Abavisani, and Vishal M. Patel, “IN2I : UNSUPERVISED MULTI-IMAGE-TO-IMAGE TRANSLATION USING GENERATIVE ADVERSARIAL NETWORKS”, 24th International Conference on Pattern recognition,PP:140-146, IEEE 2018
3. [3] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz,Bryan Catanzaro , “HIGH-RESOLUTION IMAGE SYNTHESIS AND SEMANTIC MANIPULATION WITH CONDITIONAL GANs”, Conference on Computer Vision and Pattern Recognition,PP:8798-8807,IEEE 2018
4. [4]Feng Xiong, Qianqian Wang, Quanxue Gao, “CONSISTENT EMBEDDED GAN FOR IMAGE-TO-IMAGE TRANSLATION”,IEEE Access, Volume 4, PP: 1-11, IEEE 2018

AUTHORS PROFILE

G.Pranavi, 4/4 B.Tech, Depatment of CSE, VR Siddhartha Engineering College, Areas of interest: Python, Machine Learning, IOT

M.Pranava Brunda, 4/4 B.Tech, Department of CSE, VR Siddhartha Engineering College, Areas of interest: Machine Learning, Data Science, IOT

P.Jashwanth, 4/4 B.Tech, Department of CSE, VR Siddhartha Engineering College, Areas of interest: Machine Learning, IOT

V.Naga Chaitanya, 4/4 B.Tech, Department of CSE, VR Siddhartha Engineering College, Areas of interest: Machine Learning, IOT

D.Rajeswara Rao, Faculty in Department of CSE, VR Siddhartha Engineering College, Areas of research include Data mining and data warehousing, Machine Learning, Soft Computing. A member of CSI, ISTE, ACM Amaravathi Chapter Foundation Secretary Year 2016. An author of so many national and international publications.