A COMPARATIVE STUDY OF CYCLE GAN AND PROGRESSIVE GROWING GAN FOR SYNTHETIC DATA GENERATION

Yash Burad
Pune Institute of Computer Technology
Information Technology

Kushal Burad
P.V.G College of Engineering, Pune
Computer Science

*Abstract-***Modern deep learning applications require a lot of data. Most of the times the data is insufficient and can be very costly to collect. This problem can be solved by synthetic data generation using. Generative Adversarial Networks.**
We have performed a detailed comparative study of two most popular GANs for synthetic data generation and when each of them is used. Using the following approaches alone would solve most modern deep learning data insufficiency problems.

I. INTRODUCTION

Training deep learning models often require huge datasets and a lot of time. Often large datasets are not available or can be expensive and time-consuming to generate. This problem of lack of data has bothered data scientists for a long time.

As the field of data science evolved some modern approaches were developed in order to generate data from existing datasets like Generative Adversarial Networks also known as GANS.

Synthetic data generative methods such as GANs are learning the true data distribution of the training dataset and are capable of generating new data points from this distribution with some variations and are not merely reproducing the old (training) data the model has been trained on. We have confirmed this finding using a range of non-parametric statistical significance tests.

GANs were introduced in the year 2014 in a paper written by Ian Goodfellow and other researchers at the University of Montreal, including Yoshua Bengio. Referring to GANs, Facebook's AI research director Yann LeCun called adversarial training "the most interesting idea in the last 10 years in machine learning as it has numerous applications.

GANs' potential for both good and evil is huge as they can learn to mimic any distribution of data at a very high accuracy which means GANs can be taught to create worlds eerily similar to our own in any domain: images, music, speech, prose. They are robot artists in a sense, and their output is impressive as it is difficult to differentiate between original and output generated by GANs. But they can also be used to generate fake media content but as its upside its plays exceedingly important role in generating synthetic data.

II. MOTIVATION

Deep learning can generate interestingly accurate results given sufficient data. Networks like inception, vgg16, resnet are trained on millions of images in imagenet dataset. This shows that models trained on larger datasets perform better.

Huge training data is not available for all tasks. For some tasks, it is even a challenge to gather thousands of example images. This is usually the case for medical images such as mammography for breast cancer detection and localization, chest x-rays for lung cancer detection, or MRI scans to locate brain tumors.

Having large datasets also reduces the chances of overfitting where model performs well on input data but is not able to generalize the training on test set and hence performs poorly on test data. Often this problem occurs when training data is not sufficient.

Certain techniques like image augmentation can be used to generate training data. This technique is not very effective for larger models. GAN based image data generation has proven to be very effective in such cases.

III. DETAILED COMPARISION

A. CYCLE GAN-

1. Introduction

Cycle GAN is a method of performing image-to-image translation. Image to image transfer is a process of transforming an image into another. Cycle GANs are slightly different from other image to image translation techniques as they do not require input in the form of paired samples.
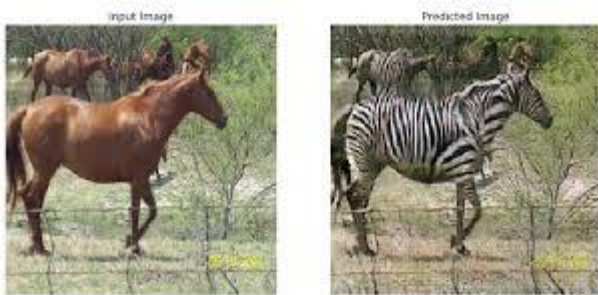
Fig 1: Image to image translation using Cycle GAN[1].

In the above image the horse was transformed into zebra using image to image translation.

There has been a great deal of research in this domain but many approaches involved supervised training which required paired data i.e. the image with corresponding output for training the network. It was in 2017 the concept of cycle GANs was introduced through a research paper [1].
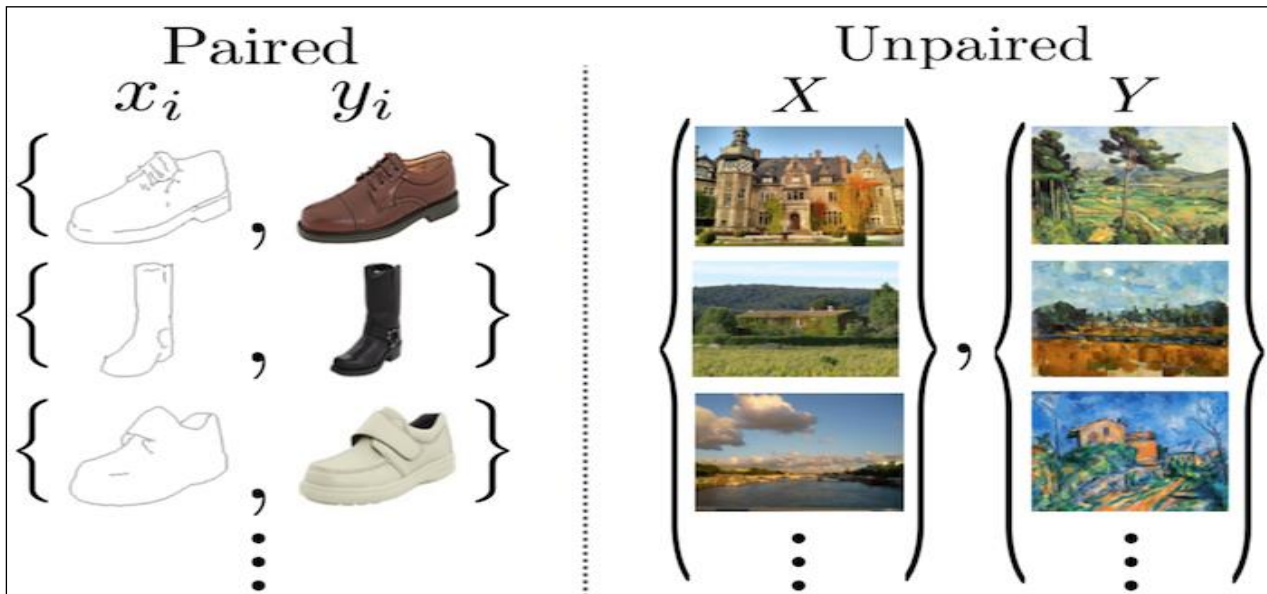

Fig 2: Difference between paired and unpaired data[1]

Cycle GANs were different from other techniques in a way that they did not require paired training data. This means that the input and output images may not be entirely similar as shown the above picture taken from the same research paper.

2. Theory

Let's say that we want to transform image X into Y. We use a network consisting of 2 GAN networks where,

- GAN1-Performs transformation from X->Y .

- GAN2-Performs transformation from Y->X.

We know that each GAN has a generator and a discriminator.

- Generator1(G)- generates images in Y set.

- Discriminator1(Dy)- tries to predict original image from the images in Y set and the one generated by generator1.

- Generator2(F)- generates images in X set.

- Discriminator2(Dx)- tries to predict original image from the images in X set and the one generated by generator2.

It is based on the principle that an image translated from one domain to another and back again should look like the original image.

Adversarial Loss-

Adversarial loss can in theory help to generate images with the same distribution as that in the target domain. So adversarial losses help in learning mapping G and F. a network can map the same set of input images to any random permutation of images in the target domain, where any of the learned mappings can induce an output distribution that matches the target distribution. Thus, an adversarial loss alone cannot guarantee that the learned function can map an individual input x_i to a desired output y_i. This is why cycle consistency loss is introduced.

Cycle Consistency Loss-

Since adversarial loss alone cannot guarantee that the generated image will look similar to the target image

we try to reduce the total possible mappings by arguing that an image after going through the translating cycle should look like the input image. This loss is the difference between input image and the generated image after a cycle of transformations[F(G(xi))].
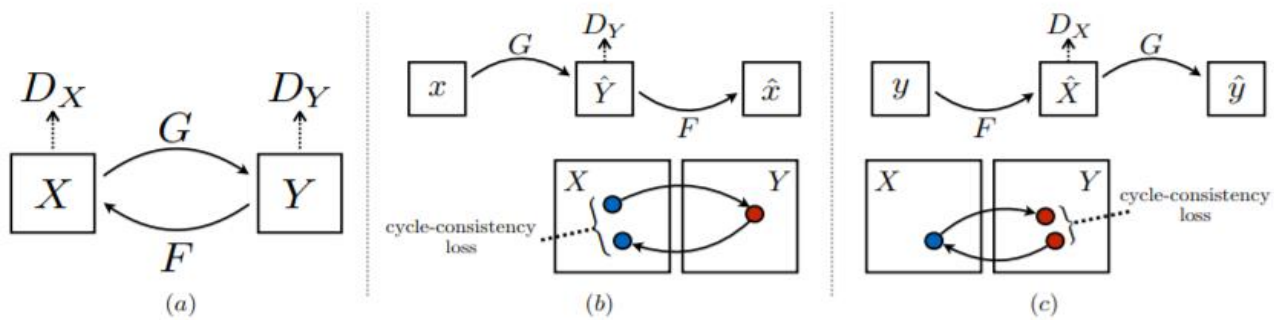

Fig 3: a) Our model contains two mapping functions G: X → Y and F: Y → X, and associated adversarial discriminators DY and DX. DY encourages G to translate X into outputs indistinguishable from domain Y, and vice versa for DX and F. To further regularize the mappings, we introduce two cycle consistency losses that capture the intuition that if we translate from one domain to the other and back again we should arrive at where we started: (b) forward cycle-consistency loss: x → G(x) → F(G(x)) ≈ x, and (c) backward cycle-consistency loss: y → F(y) → G(F(y)) ≈ y[image source-original research paper[1]]

3. Use Cases-

Cycle GAN should generally be used to generate data sets for scene classification as they can reproduce the style of the image very effectively and take a lot of time to capture the details. For example converting a winter scene to a summer scene can be done easily but using it for medical data generation which involves a lot of details is often difficult for Cycle GANS.

B. PG GANS-

1. Introduction-

GANS are used to generate synthetic data for various industrial applications from creating synthetic faces to generating medical data to train models. However, GANS are not much useful as they are unable to generate high resolution images as it affects the stability of the model.

Progressive growing GANS are very useful in such scenarios as they can generate real looking, accurate and high resolution images. It involves starting with a very small input and incrementally adding that increases the output size of the generator model and the input size of the discriminator model until the desired size is achieved

All GANs are effective at generating sharp images but on the downside they are limited to small image sizes to maintain model stability.

For generating high resolution image and maintaining the original size of image Progressive growing GAN come into light. It a stable approach to training GAN models to generate large high-quality images that involves incrementally increasing the size of the model during training.

Progressive growing GAN also known as PG GANs are capable of generating photorealistic synthetic faces and objects at high resolution that are remarkably realistic that's why we decided to use Progressive growing GAN which will produce remarkably high quality images. Faces given below do not exist in real life they are only use to show exceptional capabilities of Progressive growing GAN .

2. Working-

PG GANS involve incrementally adding layers to the network progressively during training. This approach was first used by Tero Karras et.al(2018) in his paper titled "Progressive Growing of GANs for Improved Quality, Stability, and Variation" .

The addition of layers incrementally also helps reduce the total training time of the entire network as the most of the iterations are done at lower resolutions and thus are faster.

Since the layers are added progressively, it allows the model to learn coarse details and as the training progresses it learns finer details.

New blocks are added to the output of the generator and the input of the discriminator using skip connections. The weight of the added block is controlled by a

parameter alpha which starts with zero and linearly increases over time to 1.
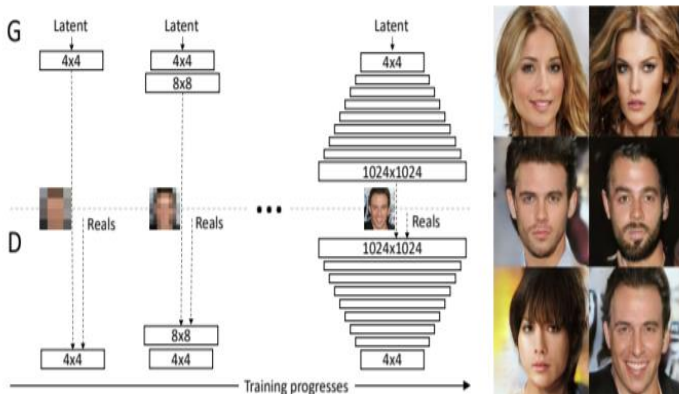

Fig 4. The training starts with low spatial resolution generators(G) and discriminators(D) and as the training progresses the layers are incrementally added to G and D for higher resolution output.[2]

3. Use Case-

PG GAN is preferred to used when required synthetic data should be of high resolution and it should preserve important traits of original data like in case of medical where minute details are very important and images are required to preserve all the details.

IV. LIMITATIONS OF GANs

There are few problems associated with GANs. These problems are generally concern with the training process of GANs and include mode collapse, internal covariate shifts, and vanishing gradients problem .

A. Mode collapse-

Mode collapse is a problem that is very common during training of GANs that refers to a situation in which the generator network starts generating samples that have little variety or when a model starts generating the same images. This happens because sometimes a probability distribution is multimodal and very complex in nature. This means there can be multiple peaks for different sub-graphs of samples. Sometimes, GANs fail to model a multimodal probability distribution which results in mode collapse. A situation in which all the generated samples are one and the same or for solving this problem data set with diverse characteristics should be used.

B. Vanishing gradients-
This takes place when gradient flows backward, from the final layer to the first layer. As it flows backward it tents to gets increasingly smaller. In few cases the gradient is so small that the initial layers learn very

slowly or stop learning completely. In this case, the gradient doesn't change the weight values of the initial layers at all which results in the training of the initial layers in the network is effectively stopped. This is known as the vanishing gradients problem. For solving this problem many techniques have come into light in recent years like gradient clipping.

V. CONCLUSION

What distinguishes GANs from other synthetic data generation models is that GANs aim to learn the true data distribution of the training dataset with a view to generate new data points from this distribution with some variations and not just reproducing the old data the model has been trained on. Since it is difficult to always learn the exact true distribution of the training data, GANs try to use the power of neural networks to learn a function to approximate the mapping to model a distribution as close as possible to the true data distribution.

VI. REFERENCES

[1] Jun-YanZhu and Taesung Park (2017) , Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, (pp. 2223-2232)

[2] Tero Karras and Timo Aila (2018) ,'Progressive Growing of GAN for Improved quality , stability and variation'. arXiv:1710.10196

[3] Ian J. Goodfellow, Jean Pouget-Abadie (2014), 'Generative Adversarial Networks' . arXiv:1406.2661

[4] Goodfellow I. (2016) ,NIPS 2016 tutorial, arXiv:1701.00160v4

[5] Aytar.Y, Castrejon.L, Vondrick.C, Pirsiavash.H, and Torralba.A. Cross-modal scene networks. PAMI, 2016. 3

[6] Bousmalis.K, Silberman.M, Dohan.D, Erhan.D, and Krishnan.D. Unsupervised pixel-level domain adaptation with generative adversarial networks. In CVPR, 2017. 3

[7] Brislin.R.W. Back-translation for cross-cultural research. Journal of cross-cultural psychology, 1(3):185–216, 1970. 2, 3

[8] Cordts.M, Omran.M, Ramos/S, Rehfeld.T, Enzweiler.M, Benenson.R, Franke.U, Roth.S, and Schiele.B. The cityscapes dataset for semantic urban scene understanding. In CVPR, 2016. 2, 5, 6, 18

[9] Deng.J,, Dong.W, Socher.R, L. Li.J, Li.KI, and Fei-Fei.L. Imagenet: A large-scale hierarchical image database. In CVPR, 2009. 8, 13, 18