Image Synthesis with Conditional Generative Adversarial Networks

Machine Learning Course Project Presentation

Shreya Roy[1] Soumyajit Gangopadhyay[2]
Project Mentor: Shubham Gupta[3]

[2]Department of Electrical Engineering
Indian Institute of Science (IISc), Bengaluru

[1]Department of Computational and Data Sciences
Indian Institute of Science (IISc), Bengaluru

[3]Department of Computer Science and Automation
Indian Institute of Science (IISc), Bengaluru

E0 270: Machine Learning

Overview

- Introduction
- Architecture with Objective function for unconditional GAN
- Experimental Observations on Synthetic Data
- Experimental Results on MNIST Data using unconditional GAN
- Architecture with Objective function for conditional GAN
- Experimental Results on MNIST Data using conditional GAN
- Architecture for Deep Convolutional GAN
- Experimental Results on MNIST Data using Deep Convolutional GAN
- Comparison between conditional GAN and Deep Convolutional GAN
- Implementation of CGAN in realistic Face Image synthesis
- Our proposed architecture
- Future Work
- Conclusion

Introduction

- Generative Adversarial Networks (GANs) have been used in the recent past for image generation, training data augmentation and many other applications.
- Lot of variants like cGAN (Conditional GAN), DCGAN (Deep Convolutional GAN), ACGAN (Auxiliary Classifier GAN) etc.
- Adversarial training between two neural networks (Generator and Discriminator) forms the basic underlying architecture)
- Hardly converges during training in practice since there is no stable equilibrium point but there is a saddle point as in a minimax game between 2 players)

Architecture with Objective function for unconditional GAN

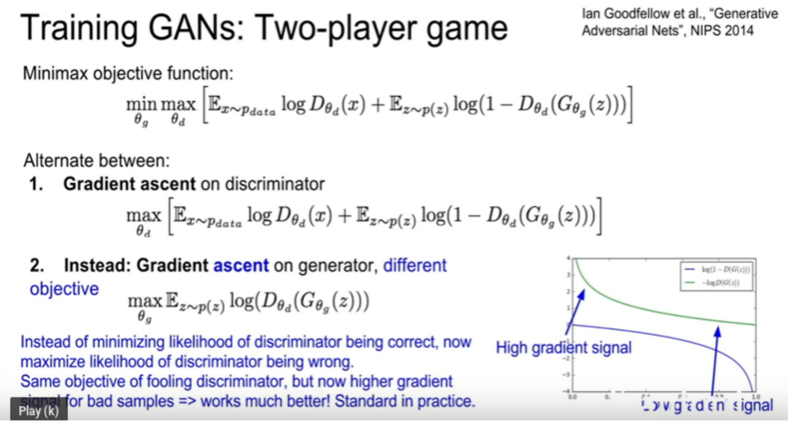

Figure: Architecture of Basic GAN (No conditioning used)

Architecture with Objective function for unconditional GAN (contd.)

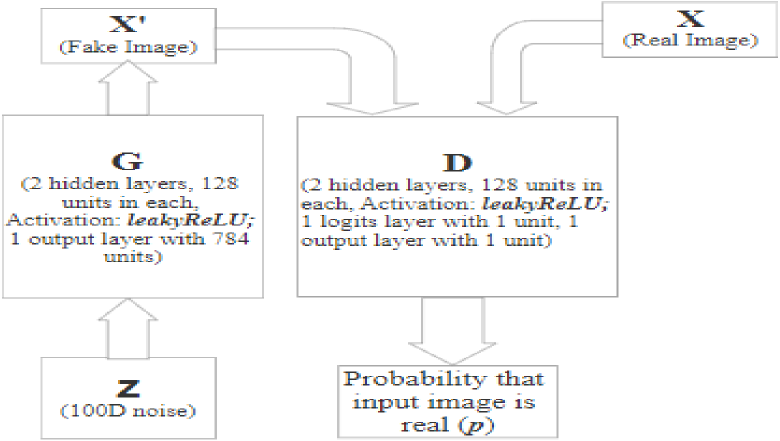

Figure: Architecture of GAN used for training on MNIST Dataset (No conditioning used)

Experimental Observations on Synthetic Data

- 10000 examples of Synthetic Data (x,y) where

$$X \sim \mathcal{N}(0, 1)$$

  and

$$y = 10 + x^2.$$

- The generator initially gave erroneous data while the discriminator was at its very best generating a very low probability less than 0.1.
- After a few epochs, the generator was close to the actual data while the discriminator was found to give probability as high as 0.9.
- However, immediately after this epoch, the generator failed miserably in generating real data. **Also it started diverging from the real distribution.** The discriminator was also found to output low probability under that situation as expected but it was not at its best during the training since it was also found that discriminator was giving probability as high as 0.7 under that undesirable situation.

Experimental Results on MNIST Data using unconditional GAN

MNIST data: 60000 training images, 10000 test images

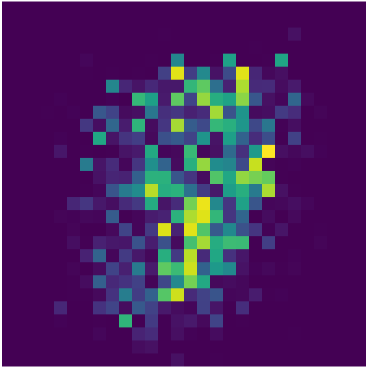

Figure: Number of Epochs = 400

Experimental Results on MNIST Data using unconditional GAN (contd.)

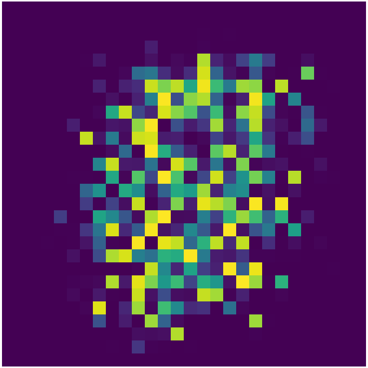

Figure: Number of Epochs = 500

Experimental Results on MNIST Data using unconditional GAN (contd.)

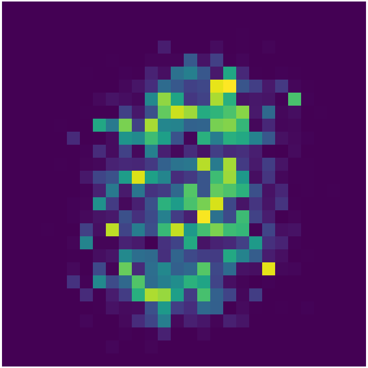

Figure: Number of Epochs = 1000

Architecture with Objective function for conditional GAN

- Minimax Objective function[2] :

$$V(D,G) = \mathop{\mathbb{E}}_{x \sim p_{data}(x)}[\log D(x|y)] + \mathop{\mathbb{E}}_{z \sim p_z(z)}[\log(1 - D(G(z|y))]$$

- Aim :

$$\min_G \max_D V(D,G)$$

- Here G and D represent the parameters of the generator and the discriminator networks respectively.

[2] [3]

Architecture with Objective function for conditional GAN (contd.)

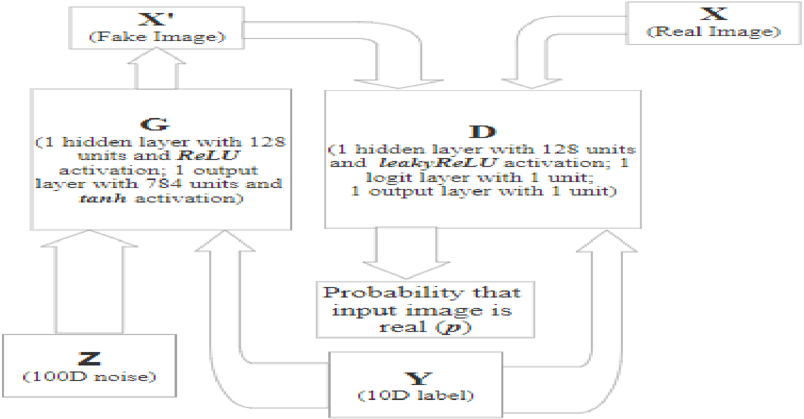

Figure: Architecture of GAN used for training on MNIST Dataset (Conditioning used)

Experimental Results on MNIST Data using conditional GAN

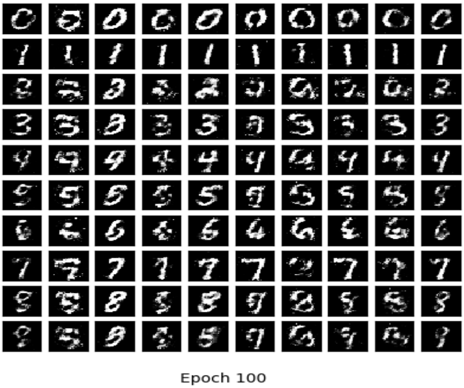

Figure: Output obtained at 100^{th} epoch using conditional GAN

Architecture for Deep Convolutional GAN

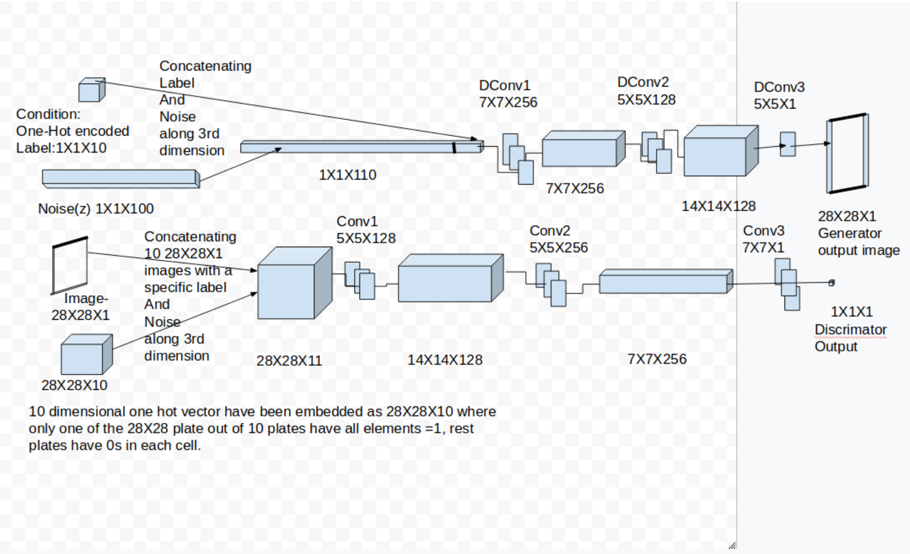

Figure: Architecture of conditional DCGAN used for training on MNIST Dataset

Experimental Results on MNIST Data using Deep Convolutional GAN

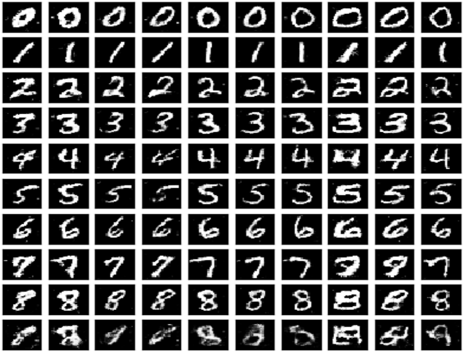

Figure: Generated MNIST Data at 30^{th} epoch with conditional DC-GAN

Comparison between conditional GAN and Deep Convolutional GAN

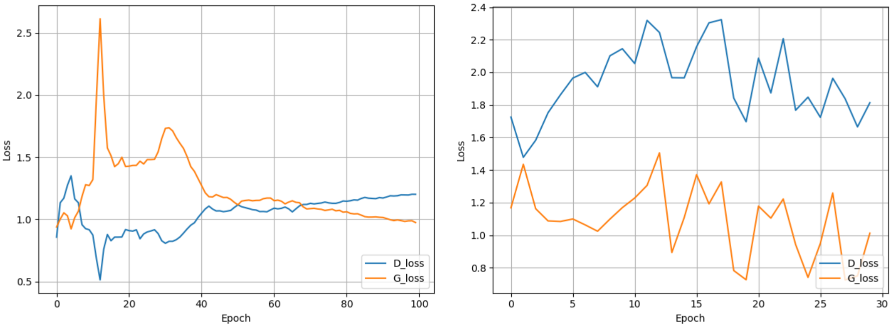

Figure: cGAN vs DCGAN training history

Application of CGAN in realistic face image synthesis

- In the problem of face recognition with limited training data, CGAN can be used to augment the existing training set[3]. The generated faces of each identity should be composed of various combinations of attribute vectors such as pose, facial expression, lighting condition, etc.[4] to make the augmented training set diverse and balanced.
- To keep identity constant and allow other attributes to vary, we can provide attribute vector to both the Generator and Discriminator.
- To train the Discriminator, we will use another dataset (CelebA was used in the paper) which is already enriched with various combinations of face attributes such as pose, expression, lighting, age, etc.
- Although the identities present in the target face recognition dataset may not be present in the dataset which is diverse and balanced and will teach the discriminator about the various face attributes only.

[3] [2]
[4] [1]

Our Auxilliary GAN architecture to generate realistic face images conditioned on various attributes

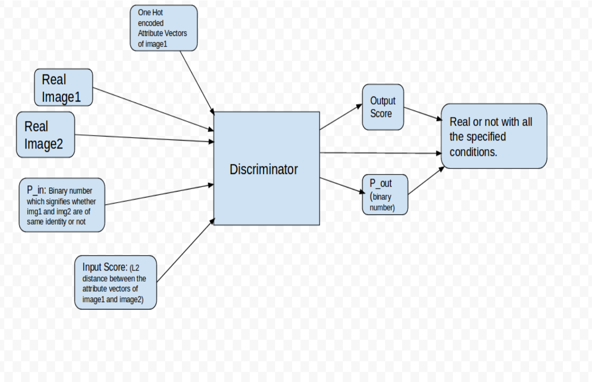

Figure: Flow Diagram of training the Discriminator (with the aim to minimize the difference between input score and output score; and minimize the cross entropy between P_{in} and P_{out})

Our Auxiliary GAN architecture to generate realistic face images conditioned on various attributes

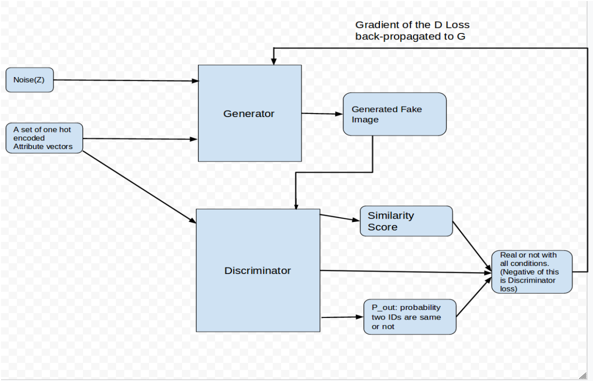

Figure: Flow Diagram of training the Generator: Use the gradient of discriminator loss to update its parameters with the aim of minimizing the D loss on the generated fake data

Future Work

- Although by varying the conditional information provided to this extended GAN, the resulting generative model can generate faces with specific attributes from random noise, it requires strong supervision of the label information provided with the annotated training data. Hence, to overcome this problem the IVI (Intra-class variation isolation) GAN has recently been proposed by some researchers from France which provides the ability to learn realistic models directly from data in an unsupervised fashion . This formulation is able to learn realistic models with continuous, semantically meaningful input parameters and needs only the weak supervision of binary attribute labels. We can implement the IVI-GAN in future.

Conclusion

- A Generative Adversarial Network (GAN) takes the idea of using a generator model to generate fake examples and the discriminator model tries to decide if the image it receives is a fake (i.e. from the generator) or a real sample.
- The conditional GAN which is an extension of generative adversarial networks (GANs) to a conditional setting, attempts to guide the data generation process in a better way by providing certain contextual information about the data.

Conclusion

- Convolutional nets, in general, find areas of correlation within an image, that is, they look for spatial correlations. This means a DCGAN would likely be more fitting for image/video data, whereas the general idea of a GAN can be applied to wider domains.
- As discussed above, in the ordinary conditional gan (C-GAN) we feed the network with one conditional information. to provide the network with more side-information apart from the identity label such as age, lighting condition, facial expression in order to generate more photorealistic new face of the given identity with sufficient intra-class variations an additional task-specific auxiliary classifier to the discriminator is used to optimize the model on the original tasks as well as the additional task

Bibliography

[1] J. Gauthier. Conditional generative adversarial nets for convolutional face generation. 2015.

[2] H. Huang, P. S. Yu, and C. Wang. An introduction to image synthesis with generative adversarial nets. *CoRR*, abs/1803.04469, 2018.

[3] M. Mirza and S. Osindero. Conditional generative adversarial nets. *CoRR*, abs/1411.1784, 2014.