

Pace University CS 627: AI Computer Science
Generating images of faces using Generative Adversarial Networks(GANS)

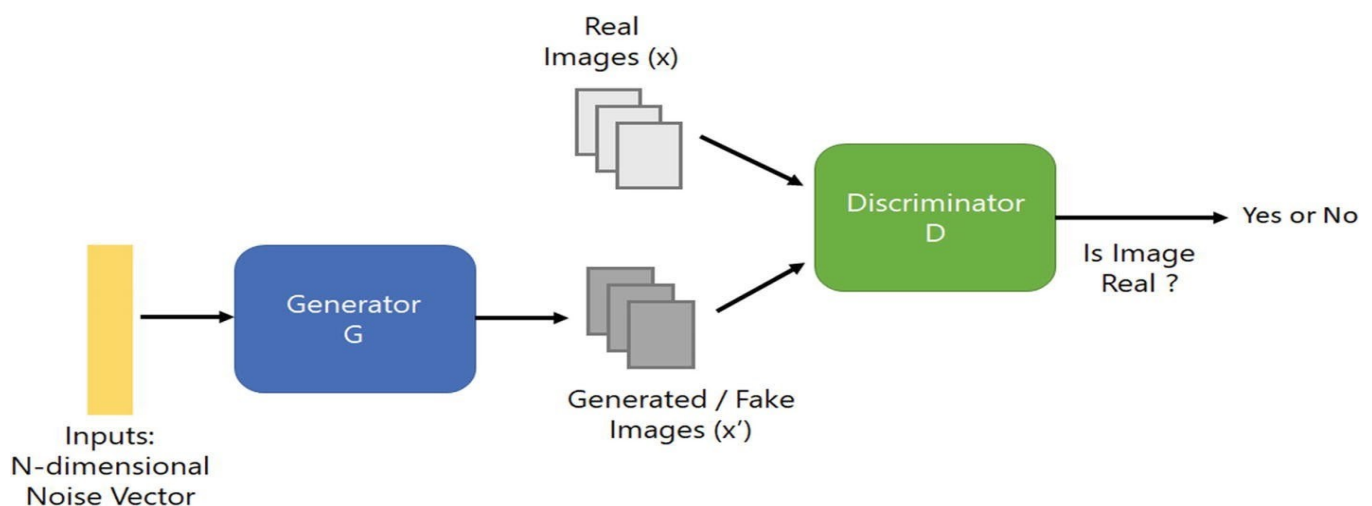
UID:-U01758611

Name:- Yashwanth Varre

Introduction:-

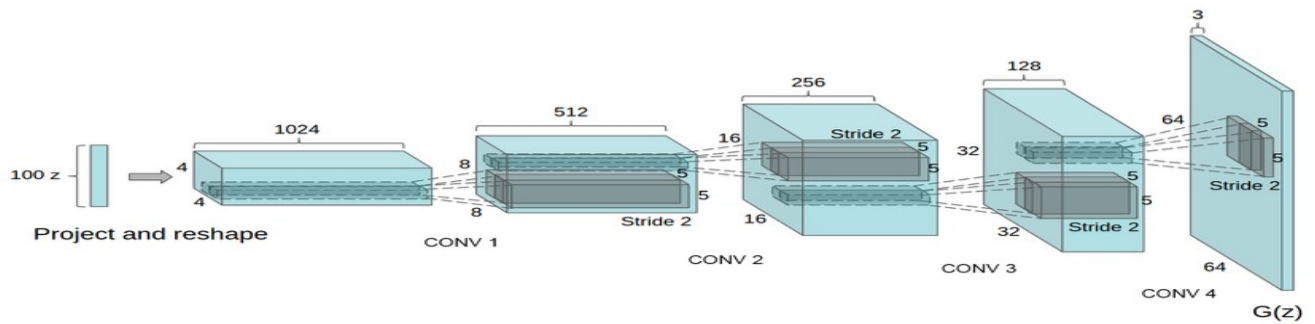
A GAN is a generative model in which two neural networks are competing in a typical game theory scenario. The first neural network is the generator, responsible of generating new synthetic data instances that resemble your training data, while its adversary, the discriminator tries to distinguish between real (training) and fake (artificially generated) samples generated by the generator. The mission of the generator is to try fooling the discriminator, and the discriminator tries to resist from being fooled. That's why the system as a whole is described as adversarial. Gans was designed by Ian Goodfellow in 2014. This is a combination of supervised and unsupervised algorithm.

Block Diagram:-



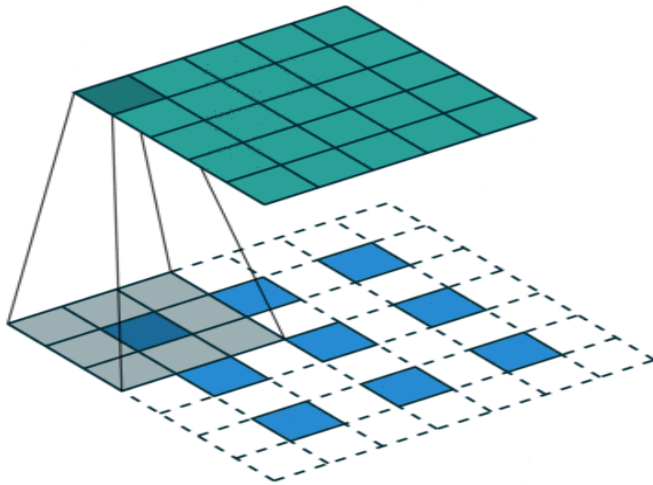
figure_1: block diagram of gans

Generator:-



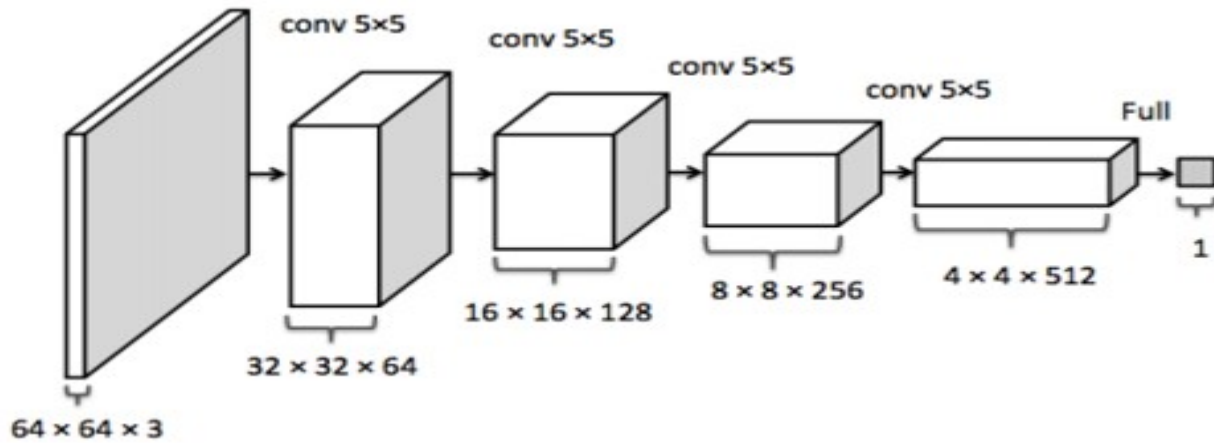
figure_2: block diagram of generator network

- Random noise is given as input of shape[batch_size, 1,1,100].
- The input is passed through each layer where transpose convolution followed by batch normalization , leaky relu is performed.
- In the final layer the input is given to tanh activation function we get 64*64 image.



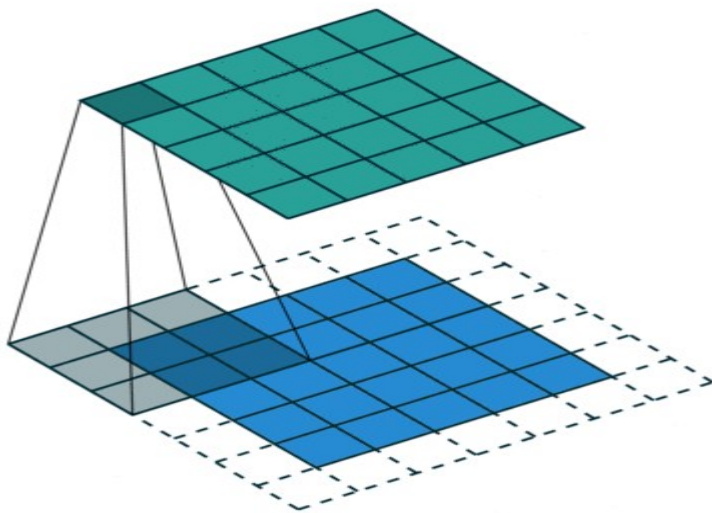
- A gif of transpose convolution.

Discriminator:-



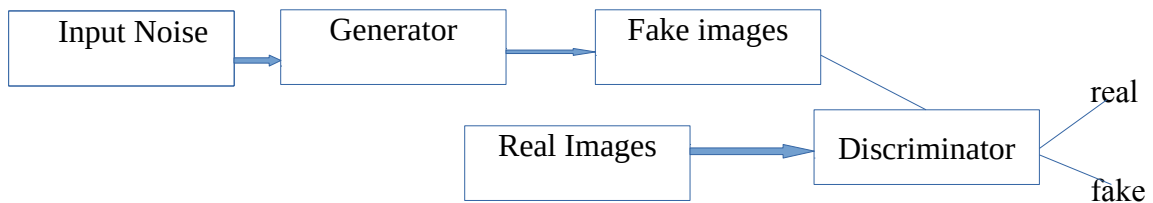
figure_3: block diagram of discriminator network

- Every layer of the discriminator contains a Convolutional layer, followed by batch normalization and relu activation function.
- Using discriminator network the real and fake images are trained.
- During the training the images are passed through sigmoid function to check the probability of image.
- If the probability is greater than 0.5 discriminator considers it as a real image.



- A gif of convolution.

Loss Function:-



Generator $\rightarrow P(\text{fake}) \uparrow$

Discriminator $\rightarrow P(\text{real}) \uparrow, P(\text{fake}) \downarrow$

P represents probability

$$\text{Loss function}(L) = \log(D(\text{real})) \uparrow + \log(D(\text{fake})) \uparrow \downarrow$$
$$= \log(D(\text{real})) \uparrow + \log(1 - D(\text{fake})) \uparrow \downarrow$$
$$\max_D \min_G L = \sum_N \log(D(\text{real})) + \log(1 - D(\text{fake}))$$

figure_4: Loss function

Steps:-

- Load the dataset and resize the data to 64*64(<https://www.kaggle.com/jessicali9530>)
- Take random noise as input for the generator for generating of samples.
- In generator each layer the input is passed through:
 - transpose convolution layer
 - Batch normalization layer
 - Activation function
- Finally the input is passed through tanh activation function.
- Similarly the real images are also trained in the similar fashion.
- The samples generated from generator and the trained real images are passed through discriminator.
- A sigmoid function is applied to the data and if the probability is greater than 0.5 they are considered to be real.

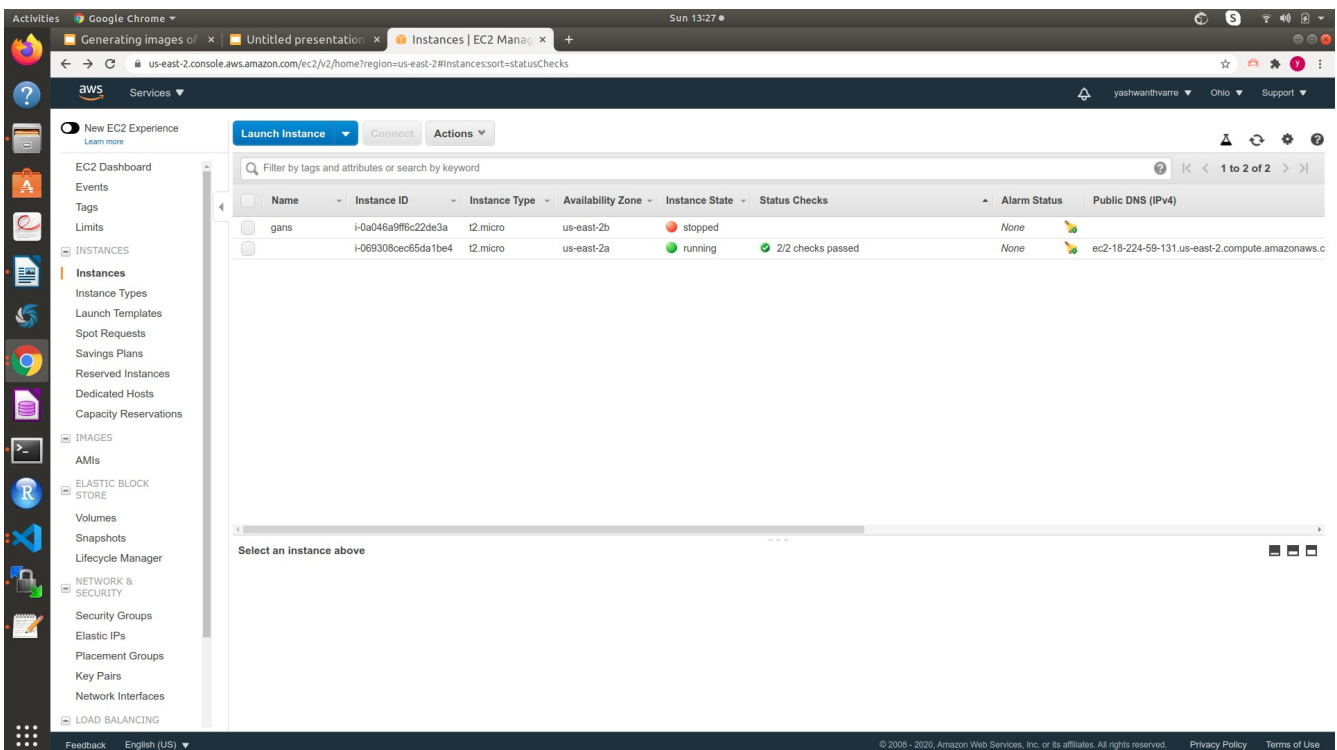
- A loss function is used to measure the loss between discriminator and generator. Generator tries to minimize the loss and discriminator tries to maximize the loss.

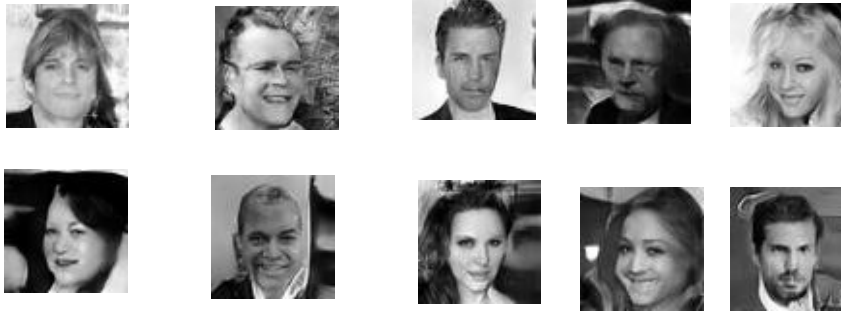
Deploying the code in cloud platform:-

- Amazon Web Services (AWS) is a secure cloud services platform, offering compute power, database storage, content delivery and other functionality to help businesses scale and grow. Running web and application servers in the cloud to host dynamic websites.
- AWS provides many services for deploying machine learning models, deep learning models and much more.
- Login to AWS using your credentials.
- Search EC2 in the search .
- One of the service provided by AWS is EC2 instance of Ubuntu 18.04.
- EC2 instance is a virtual server in Amazon's Elastic Compute Cloud for running applications on the cloud.
- I have created a EC2 t2micro instance which is free tier applicable and a pem file is given for connecting to the server.
- We use putty generator for creating a private key using the generated pem file.
- “ssh -i .‘pemfile’ public Dns” By using the following command in terminal we can connect the server and the host system.
- We transfer the code using Windows Secure copy(winscp) software.
- Install the necessary software for the code, train the model and save it.
- Deploy the code via terminal to get the output of images.

Screenshots of Output:-

Screenshot_1 running of EC2 instance in aws:





- Fake images which are evaluated by discriminator

Advantages:-

- GANs generate data that looks similar to original data.
- They can generate different versions of text, video and audio,

Disadvantages:-

- You need to provide different types of data continuously to check if it works accurately or not.
- Generating results from text or speech is very complex

Applications:-

- Can help predict next frame in the video.
- Enhancing the resolution of images.

References:-

- A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
- P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. arXiv preprint arXiv:1611.07004, 2016.
- M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein gan. arXiv preprint arXiv:1701.07875,2017.
- I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Advances in neural information processing systems, pages 2672–2680, 2014.
- [https://www.kaggle.com/jessicali9530/celeba-dataset\(dataset\)](https://www.kaggle.com/jessicali9530/celeba-dataset(dataset))

Questions

1) What is tensorflow?

A) Tensorflow is a open source software library for machine learning and deep learning models.

2) Why we use Relu?

A) This mainly ensures linearity among the data.

3) Why convolution is used?

A) It is used for blurring, sharpening, embossing, edge detection, and more. This is accomplished by doing a convolution between a kernel and an image.

4) Why batch normalization is used?

A) This is used to normalize the data so that data redundancy is reduced.

5) What are Sigmoid functions used?

A) when we need to calculate the probability value as its range is between 0 to 1 we use sigmoid function.

6) What is padding in CNN?

A) padding means we just add zeros while performing convolution so that the dimensions of the image remains same even after convolution.

7) What is a stride in CNN?

A) The steps by which a kernel has to move over an image.

8) What is tf.placeholder in tensorflow?

A) This lets you assign the input later.

9) What is the use of generator?

A) Generator provides sample data from noise which are similar to original data.

10) When do we use tanh activation function?

A) when we don't want to lose the output of neuron when they give -ve values we use tanh function as it range is (-1,1).