University of Strathclyde Glasgow

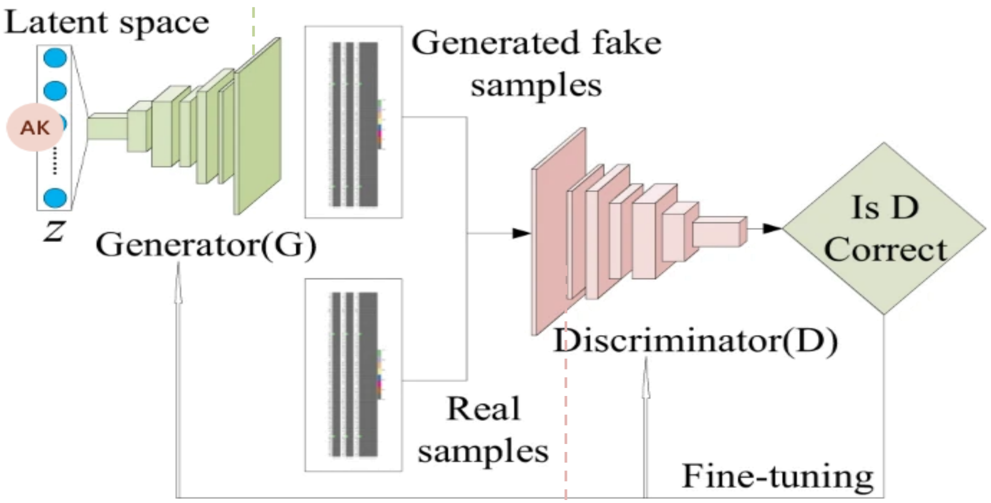
An Evaluation Of Diffusion Models Vs Generative Adversarial Networks On Image Synthesis

By: Abhishek Gowda K S

Supervisor: Dr. Feng Dong

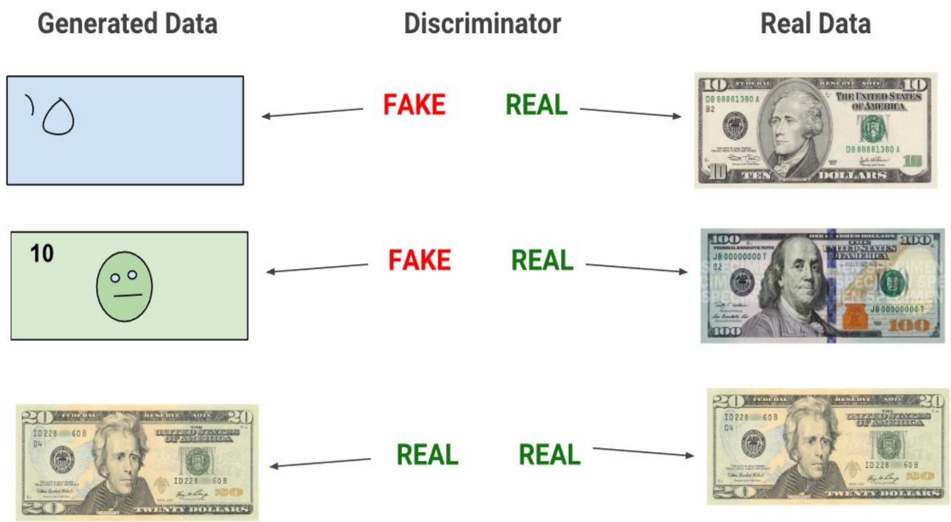
Generative adversarial networks

The generator learns to generate fake data by seeking feedback from the discriminator and convinces the discriminator that its output is real. It creates a sample image using a fixed-length random vector drawn from a Gaussian distribution. After training, points in this multidimensional vector space will correspond to points in the real image, forming a compressed representation of the data distribution.



The discriminator model predicts a binary class label of real or fake using an example from a real and generated picture as input. The real example comes from the training dataset. The generated examples are output by the generator model.

When the discriminator classifies the image as fake the generator loss **penalizes** the **generator** for producing such a sample and the weights are then updated through **backpropagation** by the generator loss to **maximize error**.

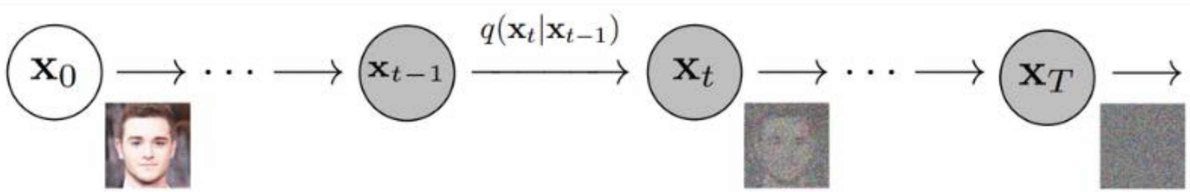


The discriminator is penalised for misclassification of data, and the weights are then updated through **backpropagation** by the discriminator loss to **minimize error**.

Loss Function : $\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$.

Diffusion Model

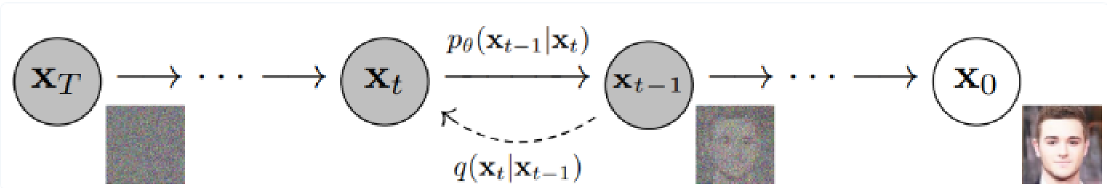
Diffusion Models are generative models that utilize the Markov chain to generate data comparable to the data on which they are trained by adding gaussian noise in succession to the training data and then recovering the data by reversing the noising process. comparable to the data on which they are trained by adding gaussian noise in succession to the training data and then recovering the data by reversing the noising process.



The forward process:

The model gradually introduces a certain amount of noise into the data in each step producing a sequence of noisy samples.

The data gradually loses its features with each step until it resembles a pure gaussian noise image



Reverse process samples from a pure gaussian noise distribution by reversing the forward process.

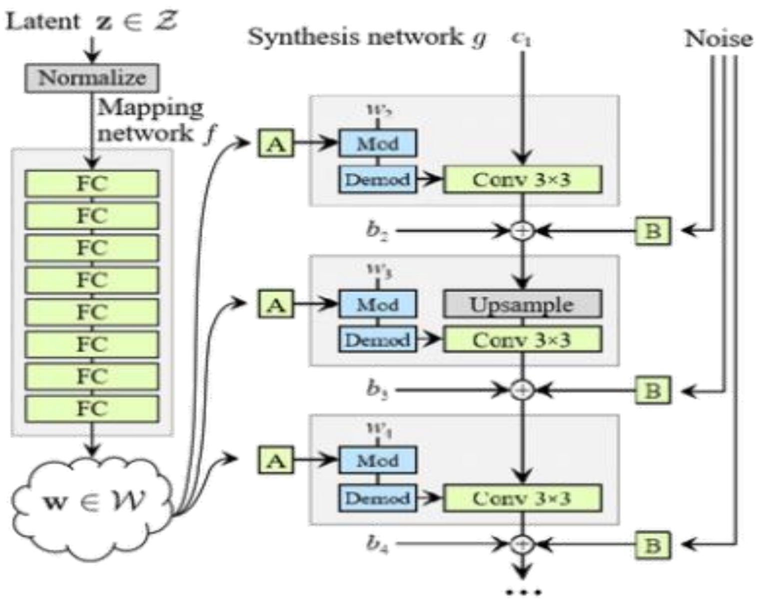
The sampling begins with noise and progresses to generate less noisy samples where parameters of gaussian noise transitions are learned until a final sample matching the original data is obtained.

Each step t corresponds to a certain noise level and x_T can be thought of as a mixture of a signal x_0 and some noise ε.

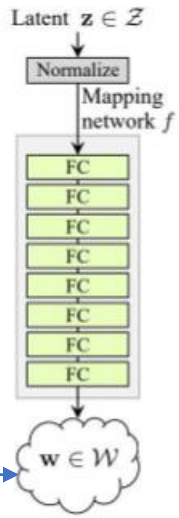
The first image is generated by converting the real image to a pure noisy image using the forward method. The second image is generated using the opposite procedure, in which the model begins eliminating the gaussian noise and begins rebuilding the image in the manner of the original image.

STYLE GAN

The Style Generative Adversarial Network can independently vary the input of each level without influencing other features, allowing it to generate photorealistic, high-resolution images by giving control over the resulting image's properties by adjusting the style vectors and noise.



The Mapping Network transforms the latent vector z into the intermediate latent space w using 8- fully-connected layers.



"Style" relates to important data aspects such as stance and identity. In order to produce stochastic variation, StyleGAN inserts noise into the spatial data; the additional noise is monitored to create variations in the feature. Noise introduces local alterations at the pixel level and seeks stochastic variation in order to generate local versions of features.

Weight demodulation performs the scale and shifts parameters out of a sequential computation path, instead of scaling into the parameters of convolutional layers.

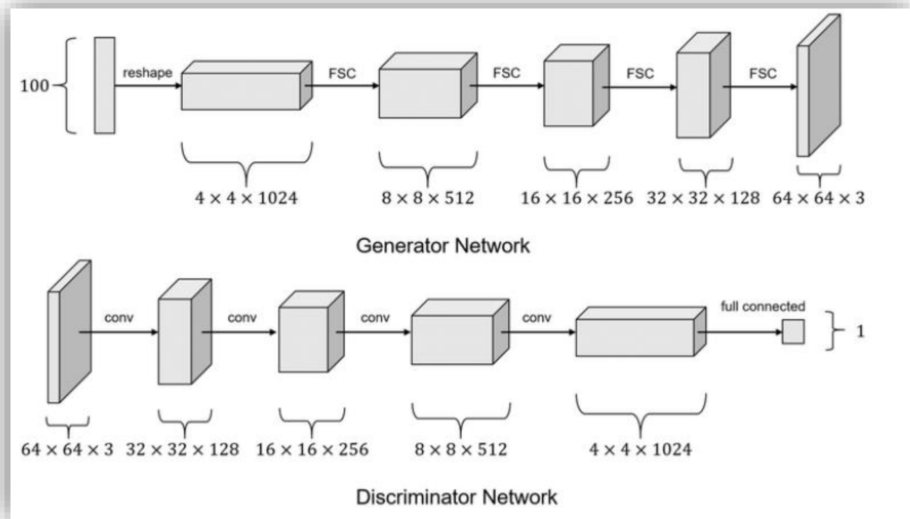
IMAGES GENERATED BY THE MODEL

The first image is obtained after training the model for 1000 steps, the resulting image appeared to be blurry due to the presence of noise.

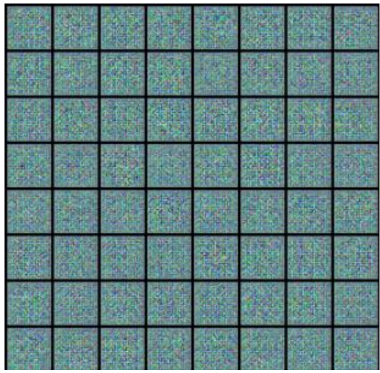
The second image is generated after training the model for 50,000 steps; the resulting image was clear and resemble that of images in the real dataset

Deep Convolution GAN

It is a class of Convolutional Neural networks that consists of convolutional and transposed convolutional layers that are deliberately embedded into its architecture.



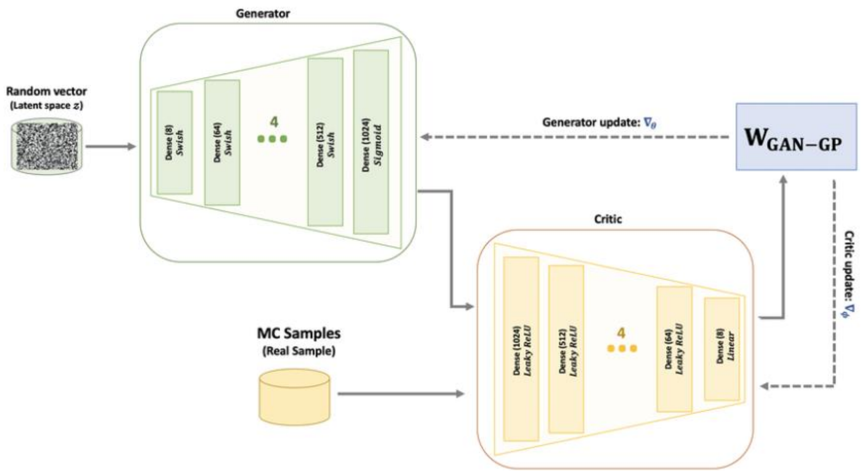
- Uses Batch normalisation
- To achieve deeper architectures fully connected hidden layers are removed.
- In the generator, hidden layers use the ReLU activation function whereas the output layer uses tanh.
- The discriminator uses LeakyReLU in all the layers.



Wasserstein GAN

The primary contribution of the WGAN model is the use of a new loss function called Wasserstein distance that encourages the discriminator to act as a critic that predicts a score of how real or fake a given input looks.

The idea of Gradient Penalty is to enforce a constraint such that the gradients of the critic's output w.r.t the inputs to have unit norm



FID

We can obtain a fair idea of how good the generator is at producing fake images by studying Fidelity and Diversity. The most often used metric to evaluate performance is Fréchet Inception Distance (FID), which estimates the distance between feature vectors produced for actual and generated images.

The score shows how comparable the two groups are in terms of statistics on raw picture characteristics derived using the inception v3 image classification model. Lower scores imply that the two sets of photographs are more comparable and that higher-quality images correspond well.

Models	FID	Steps
DC-GAN	74.46	200k
W-GAN GP	82.45	200k
StyleGAN2	75.89	10k
Diffusion	61.23	250k

GANs in healthcare

Drug discovery is another area in healthcare where generative adversarial networks might help. The networks may be used to generate molecular structures for drugs employed in disease targeting and treatment.

Researchers may utilise the existing database to train the generator to uncover novel chemicals that might possibly be used to treat new ailments.

The researchers do not have to manually scan the full database for chemicals that might help fight new diseases because the system automatically finds such compounds, reducing the time necessary for medication research and development.