CPSC 8430-Deep Learning

Harshini Gaddam

Homework-4

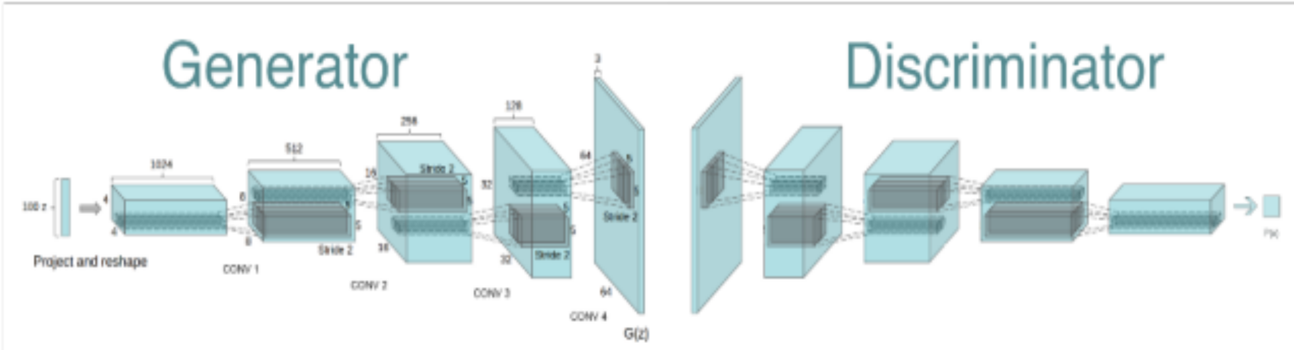
Introduction:

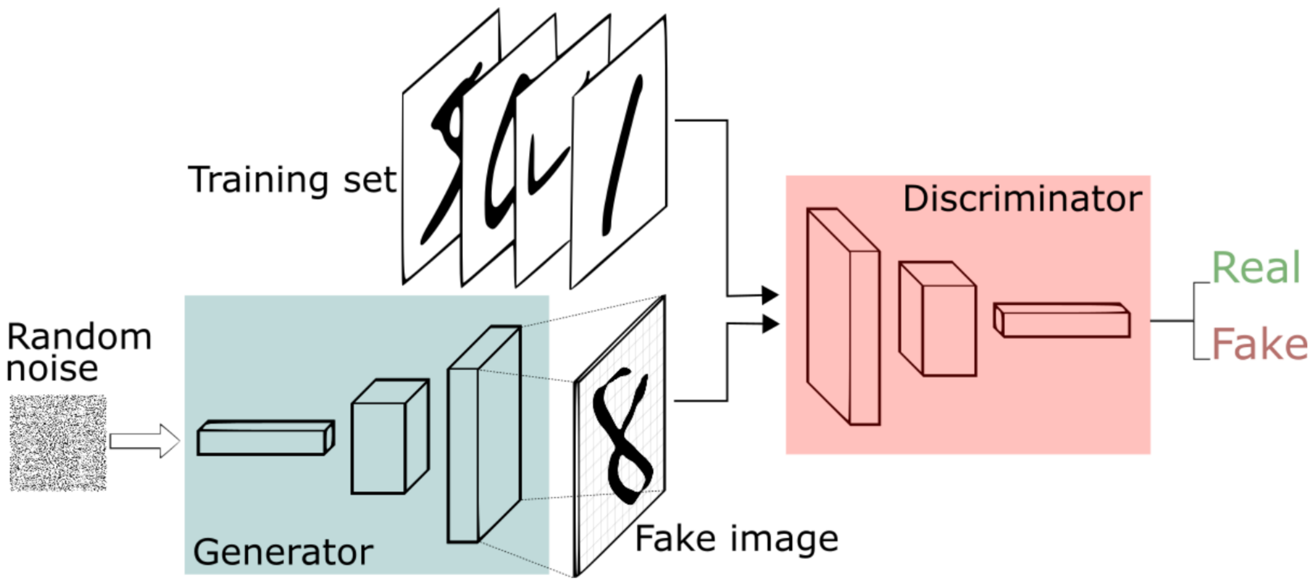
In this assignment, discriminator, and generator pairs of GAN (Generative Adversarial Network) are trained on CIFAR10 dataset utilizing techniques from DCGAN, Wasserstein GAN and ACGAN.
A GAN consists of two neural networks: a discriminator and a generator. While the discriminator attempts to distinguish between the real and created samples, the generator generates new examples that are comparable to the training data. The discriminator tries to recognize the generated samples while the generator tries to make samples that can trick the discriminator during training. Until the generator generates samples that are identical to the real data, or until the target level of quality is reached, the training procedure is repeated.

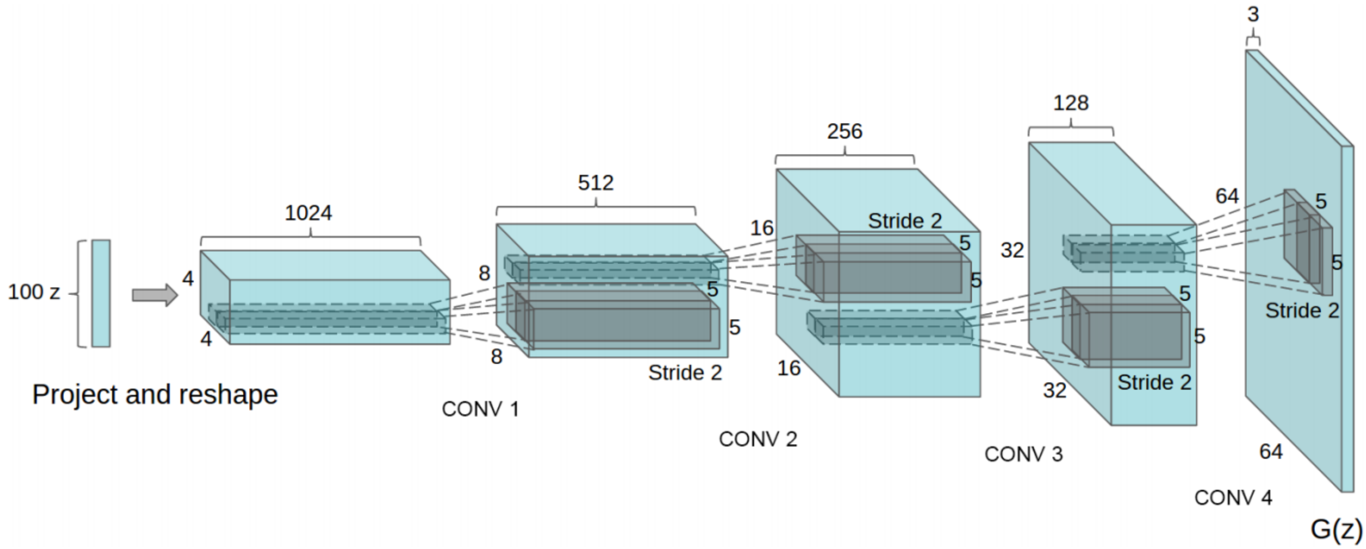
Github link: https://github.com/Harshini-Gaddam/CPSC-8430-DL_HW4.git

Implementation:

1. **DCGAN:**
   It stands for Deep Convolutional Adversarial Network. The generator tries to produce images during training that will make the discriminator believe they are real images from the training set. The discriminator, meantime, strives to accurately determine if an image is authentic or fraudulent. The discriminator gets stronger at spotting phony images as the two networks continue to learn and develop. The generator gets better at producing realistic images. To produce realistic images, it employs convolutional neural networks (CNNs) in both the generator and the discriminator.


Generator
Discriminator
100 z
Project and reshape
1024
512
256
128
Stride 2
CONV 1
CONV 2
CONV 3
CONV 4
G(z)

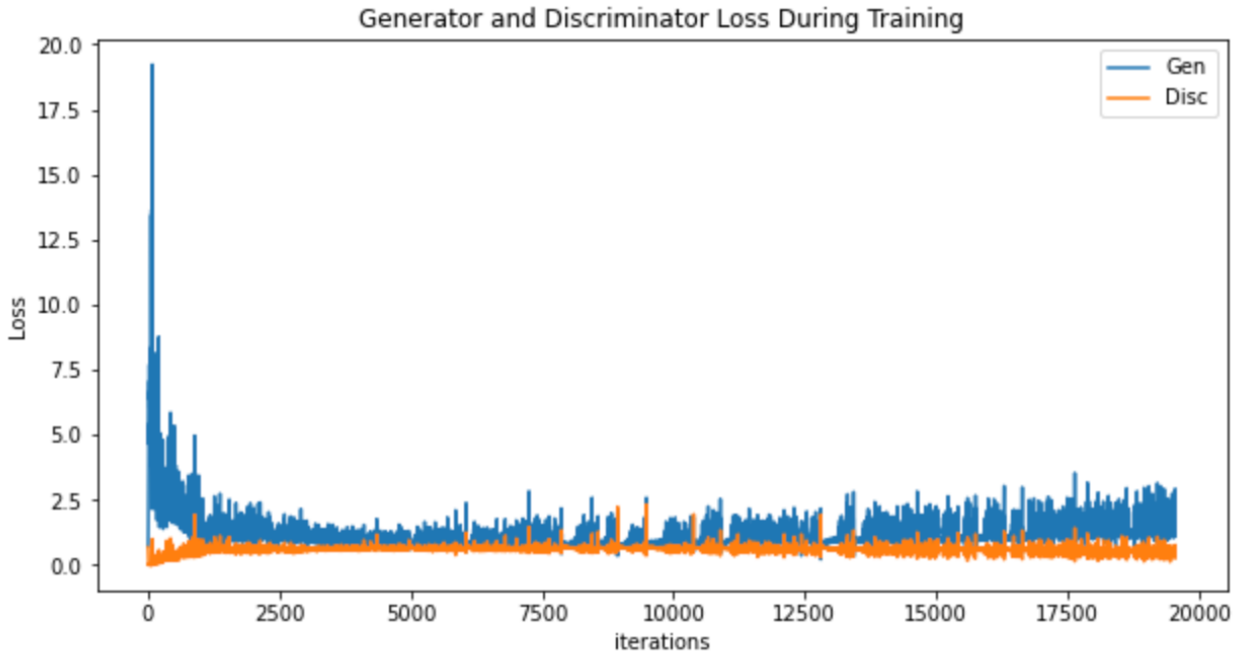

Training set
Random noise
Generator
Fake image
Discriminator
Real
Fake



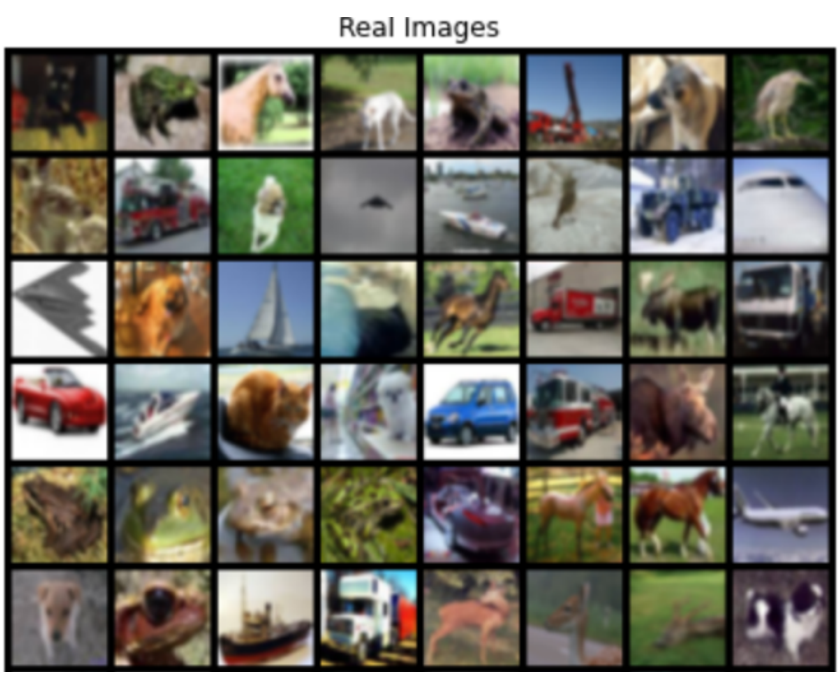
Parameters:

Learning_rate= 2e-4
Batch_size= 128
Image_size= 64
Channels_image= 3
Noise_Dimensions= 100
Number_of_epochs= 50
Features_of_discriminator= 64
Features_of_generator= 64
Beta= 0.5
Optimizer=Adam
Momentum=0.9

When CIFAR10 dataset is trained using 50 epochs, the generator and discriminator loss observed during training is plotted for 20000 iterations.



It is seen that after 2400 iterations, there is no increase or decrease seen in the plot. It denotes that the generator and discriminator are no longer significantly improving, indicating that the model has reached an equilibrium state. It means that the models have developed to the point where the generator can produce samples that are convincing enough to deceive the discriminator, and the discriminator can reliably spot the fake samples.



These are the real images from the data.

Fake images generated after the last epoch (epoch=50).

2. WGAN:

The Wasserstein distance, also known as the Earth Mover's distance, is a distinct loss function that is used by the "Wasserstein Generative Adversarial Network," a subtype of GAN, to calculate the separation between real and produced samples. Some of the training stability problems that classic GANs frequently experience is addressed by WGAN. Along with other uses like style transfer and data augmentation, WGAN has been used to generate images. WGAN uses a different loss function to address some of the training stability problems in conventional GANs.
The discriminator loss function's negative sign is crucial since it motivates the critic to give actual samples greater ratings than artificially produced ones. Without the warning, the critic would be enticed to give real samples lower ratings, which would not reduce the Wasserstein distance.

Parameters:
Z_dimension= 128
Weight_clip= 0.01
Critical_iteration= 5

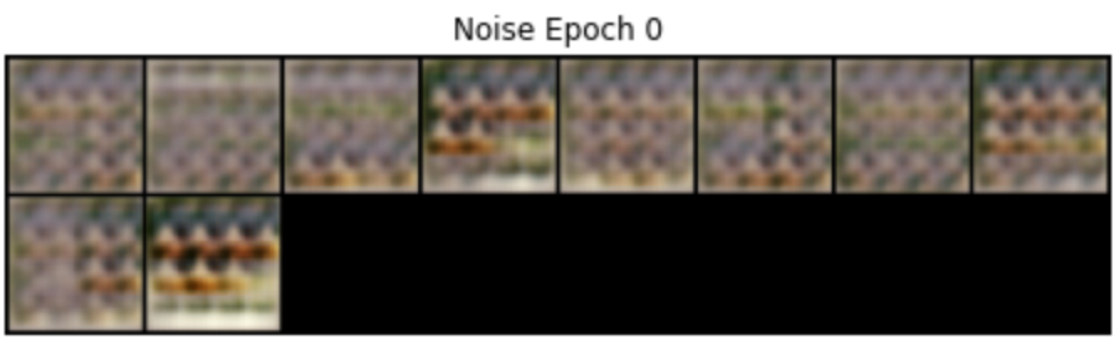


Image generated at epoch=0

Image generated at epoch=20

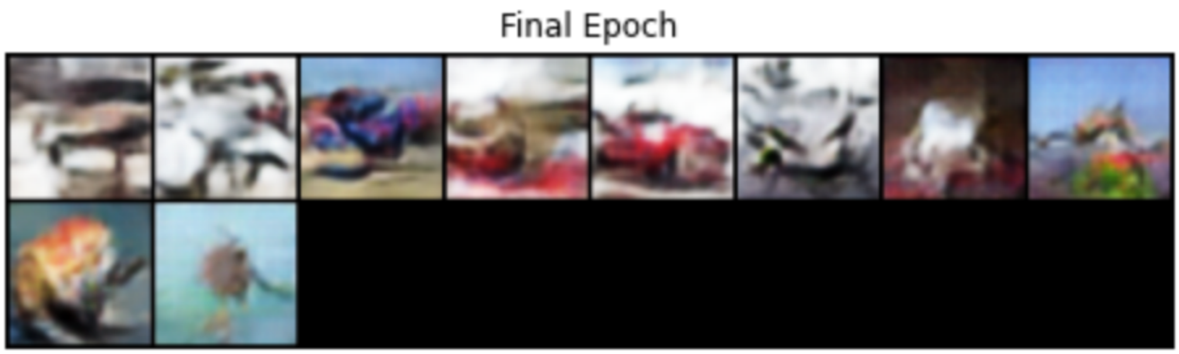
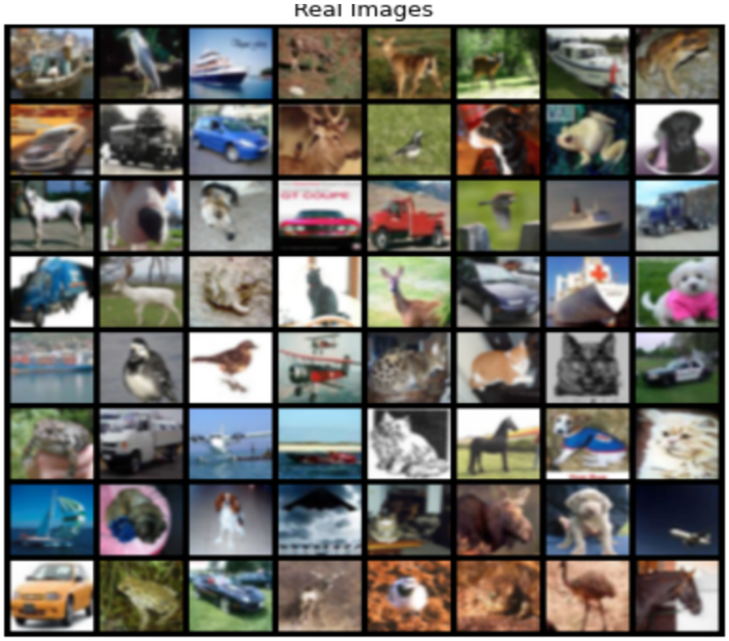
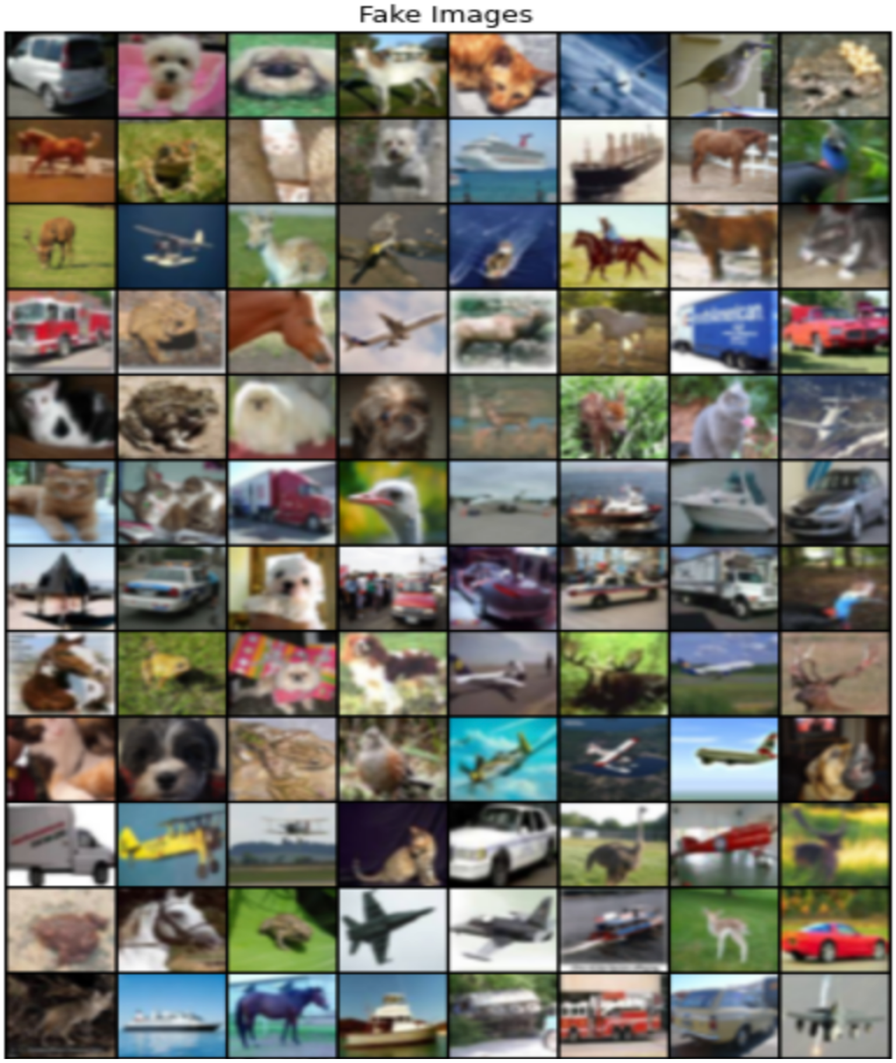


Image generated at epoch=50

Real Images

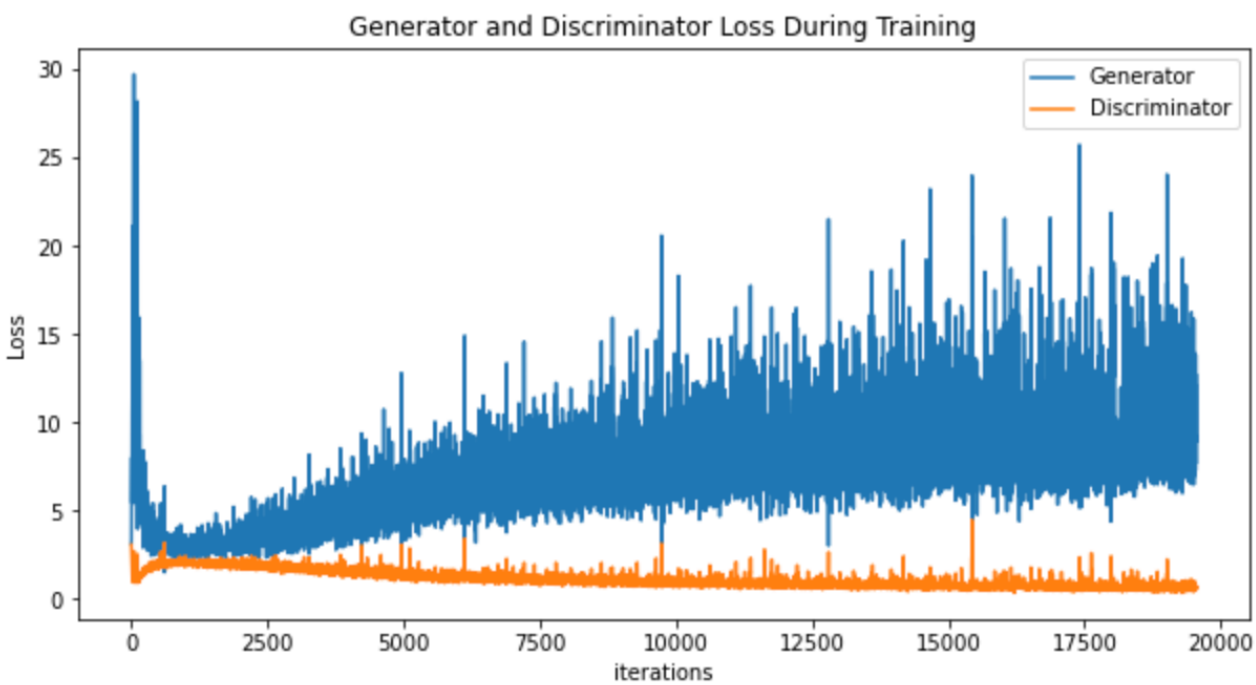


Fake Images



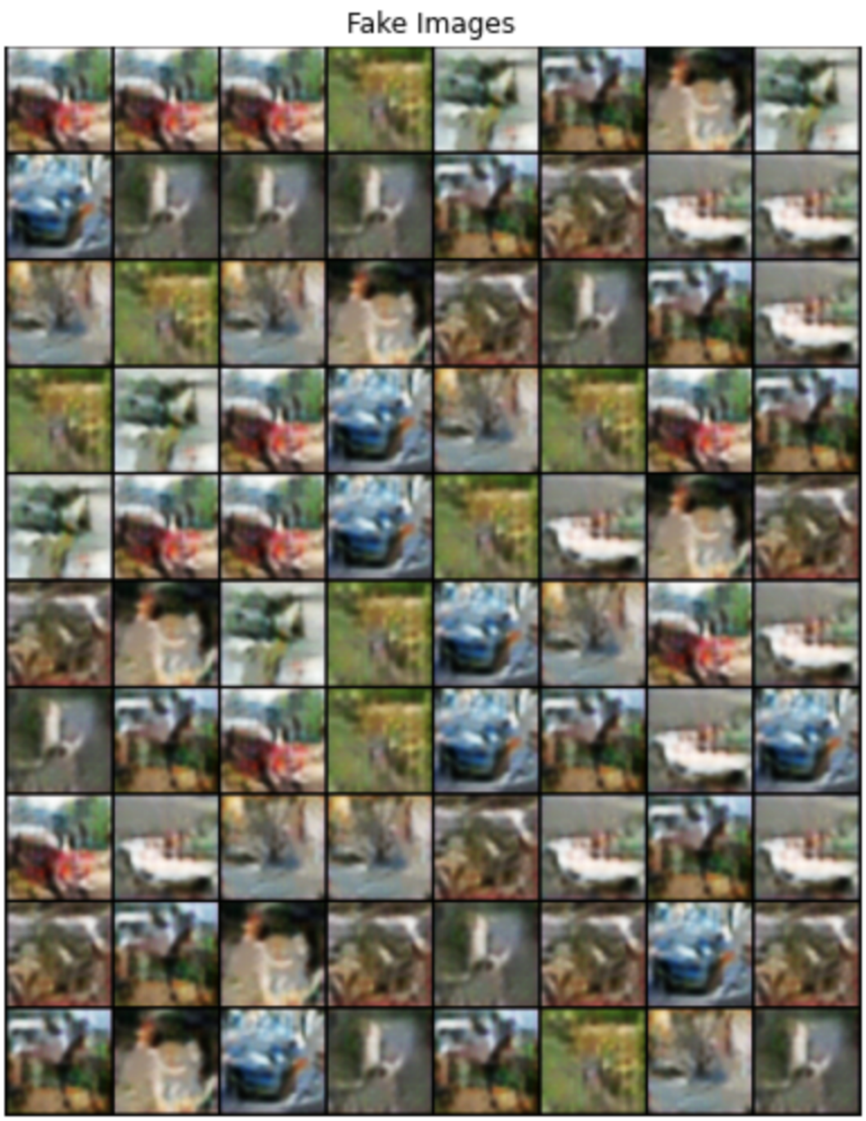
3. ACGAN:

ACGAN, or "Auxiliary Classifier Generative Adversarial Network," is a subtype of GAN that enhances the discriminator with an auxiliary classifier. The auxiliary classifier enables the discriminator to categorize the samples into other groups in addition to determining which samples are real and which are fake. ACGAN has been used to produce images that are conditioned on class labels, such as images of various animals, objects, or humans with particular characteristics. It uses additional factor called class labels.



It is seen that generator loss is increasing which means it is producing the samples that are dissimilar to the real samples in the given class. The discriminator loss is decreasing which indicates that it is becoming more effective at identifying fake samples generated by generator.


Real Images


Fake Images

Comparison between DCGAN, WGAN and ACGAN:

All GANs are trained using the same hyperparameters and trained for 50 epochs under the same Palmetto environment.

Architectural Difference:

DCGAN is a foundational architecture for GANs that uses CNNs to produce realistic images. ACGAN adds an auxiliary classifier to the discriminator to classify the samples into different categories, while WGAN uses a different loss function to address some of the training stability problems in conventional GANs.

Computational time:

The computational time taken by DCGAN, WGAN, ACGAN to run 50 epochs under the same Palmetto environment is 392.96 sec, 2485.05 sec, 477.55 sec. DCGAN takes less time compared to WGAN and ACGAN takes less time compared to WGAN.

Images Quality:

From the above observations, it is seen that DCGAN generated photorealistic and highly detailed images. Whereas WGAN generated diverse and variable images.

Training Stability:

Since WGAN uses a different loss function compared to DCGAN, so it has more stable training than DCGAN.

Convergence speed:

From the above Generator and discriminator plots, it has been seen that DCGAN converges faster than WGAN and WGAN converges faster than ACGAN.

FID score:

Frechet Inception Distance(FID). To determine the activations for the real images and the fake ones, a pre-trained Inception-v3 model was used. Pass each image through the Inception-v3 model in particular, then extract the activations at a chosen layer. Lower FID score indicates higher performance.