2023 International Conference on Information Technologies (InfoTech-2023), Proceedings of the IEEE Conference, Rec # 58664
20-21 September 2023, Bulgaria

Performance Analysis of GANs for Denoising Images

Presented By:-
VARSHA SINGH
Research Scholar (IT)
Indian Institute of Information Technology Allahabad, Prayagraj, U.P. India.

Authors

VARSHA SINGH	rsi2018002@iiita.ac.in
THANMAI DEEPTHI	iit2020184@iiita.ac.in
JAYESH GINNARE	iit2020191@iiita.ac.in
KONTHAM PAVANI	iit2020193@iiita.ac.in
MOHIT RATHOUR	iit2020194@iiita.ac.in
INDER SONU	iit2020195@iiita.ac.in
PROF. UMA SHANKAR TIWARY	ust@iiita.ac.in

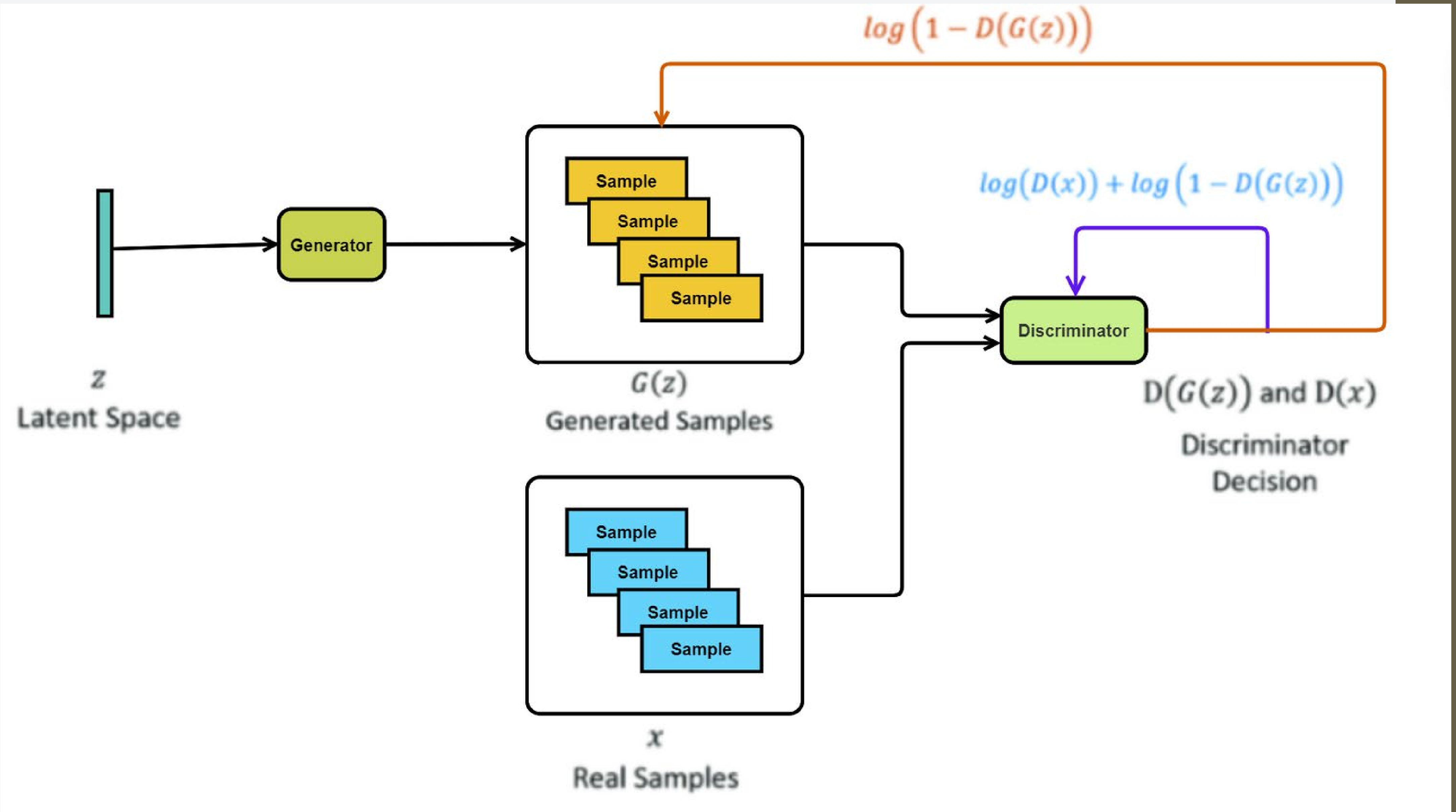
INTRODUCTION

- Denoising is a crucial task in image processing, where the aim is to remove unwanted noise from an image while maintaining its important features.
- Generative Adversarial Networks, or GANs, have emerged as a promising solution for this task, by learning to generate clean images from noisy ones.
- This presentation will explore the performance and efficiency of different GAN models, including SRGAN, DCGAN, Auto encoding GAN, and LSGAN.
- We will evaluate their denoising performance on a common dataset using metrics such as Peak Signal-to-Noise Ratio (PSNR) and analyze their ability to handle different types of noise.

GENERATIVE ADVERSARIAL NETWORKS

- It is a type of deep learning model that consists of two neural networks: a generator network and a discriminator network.
- The generator network generates fake data, while the discriminator network tries to distinguish between real and fake data.
- The two networks are trained together in an adversarial manner, where the generator tries to generate data that is increasingly difficult for the discriminator to distinguish from real data.

GENERATIVE ADVERSARIAL NETWORKS



GENERATIVE ADVERSARIAL NETWORKS

- The log loss function is commonly used in GANs as the adversarial loss function.
- It has two parts: the generator loss and the discriminator loss.
- The generator loss is the negative log probability of the discriminator correctly classifying fake data.
- The generator tries to minimize this loss by generating data that is more difficult for the discriminator to classify as fake.

$$V(G, D) = E_{x \sim Pdata} logD(x) + E_{z \sim Pdata} log(1 - D(G(z))$$

PROBLEM STATEMENT

- Generative Adversarial Networks (GANs) have emerged as a promising approach for denoising tasks.
- Given a set of noisy images, the task is to evaluate and compare the performance of different GAN-based models for image denoising.
- Our aim is to identify the strengths and weaknesses of different GAN architectures and training strategies in terms of their ability to remove noise while preserving image details and structures.
- However, the effectiveness of different GAN models for denoising, such as SRGAN, DCGAN, Auto Encoding GAN, and LSGAN, remains uncertain, as each model possesses unique strengths and weaknesses.
- To address this issue, a comprehensive comparative study is necessary to determine the optimal GAN model for denoising under various conditions.

SRGAN

SRGAN (Super-Resolution GAN) is a type of GAN that is designed to enhance the resolution of images. The architecture of SRGAN consists of a generator network and a discriminator network.

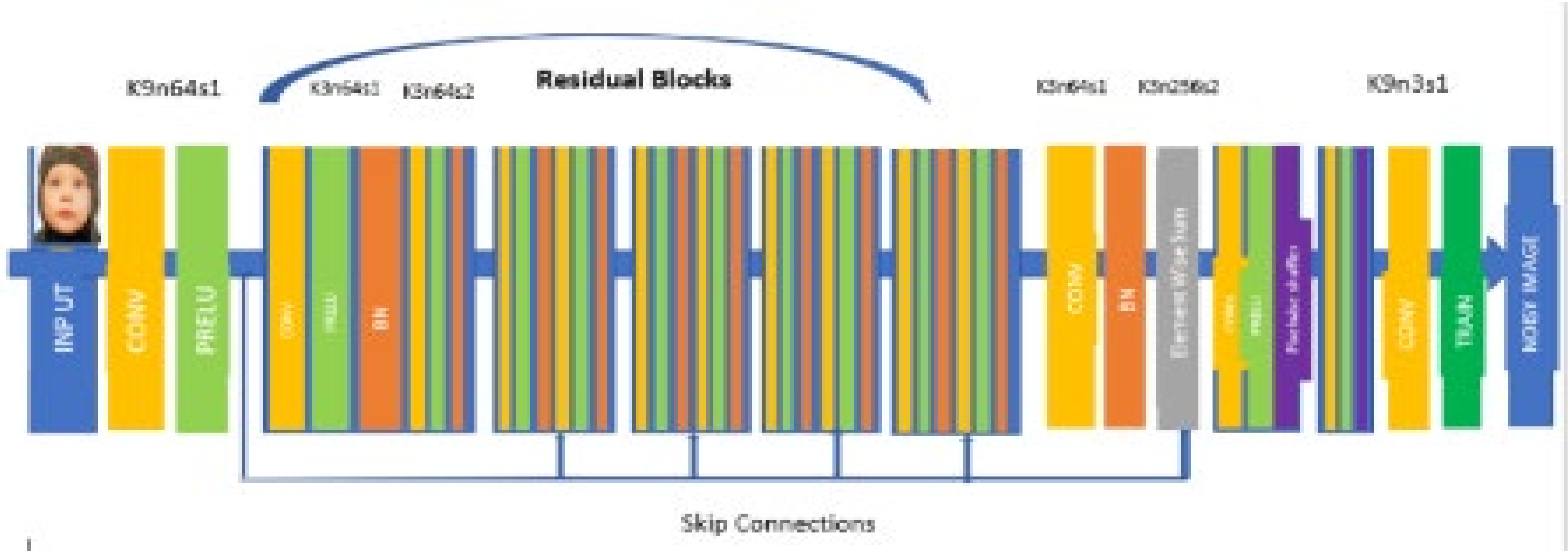
Generator:
- Deep convolutional neural network which gives high resolution image as output.
- Has several residual blocks to preserve important features of the input image
- Uses a skip connection to preserve fine details in the image.

Discriminator:
- Deep convolutional neural network, distinguishes between real and fake images.
- Uses a patch-based approach

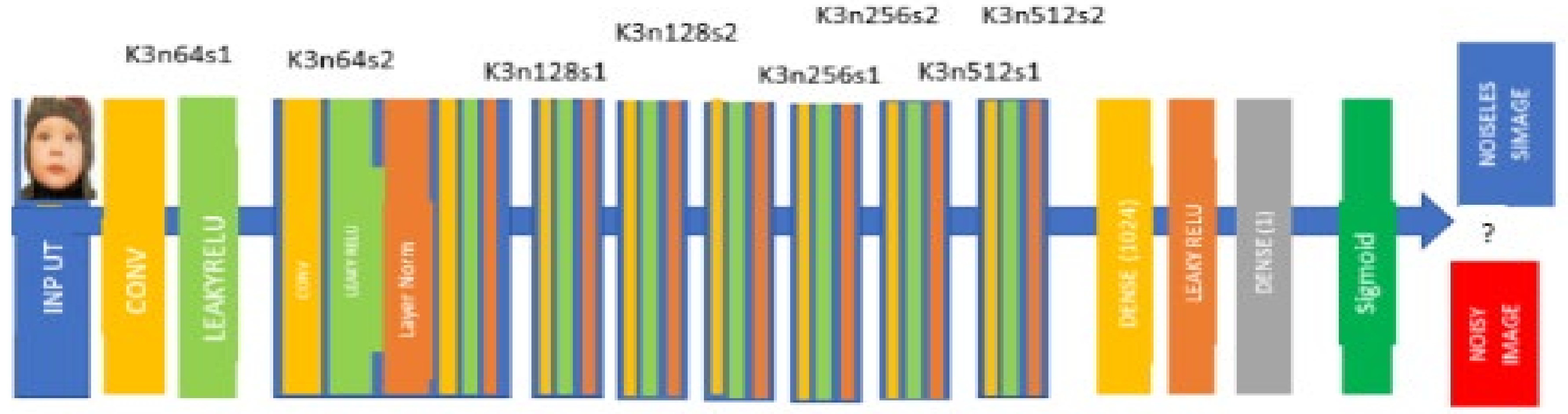
Loss function:
- Content loss
- Adversarial loss
- Perceptual loss

LSGAN

LSGAN (Least Squares GAN) is a type of GAN that uses a least squares loss function instead of the traditional binary cross-entropy loss function.



Generator



Discriminator

LSGAN

Generator:

- Consists of several layers of transposed convolutions or upsampling followed by convolutional layers
- Layers gradually increase the resolution and complexity of the image until it reaches the desired size and complexity.

Discriminator:

- It typically consists of several convolutional and pooling layers followed by a fully connected layer that outputs a scalar value representing the probability that the input data is real.

Loss function:

- uses a least squares loss function
- The discriminator is trained to minimize this loss function, while the generator is trained to maximize it.

VANILLA GAN

Generator:

- Consists of several layers offully connected, convolutional, and/or transposed convolutional layers.
- Takes a random noise vector as input and generates fake data, such as images or text.

Discriminator:

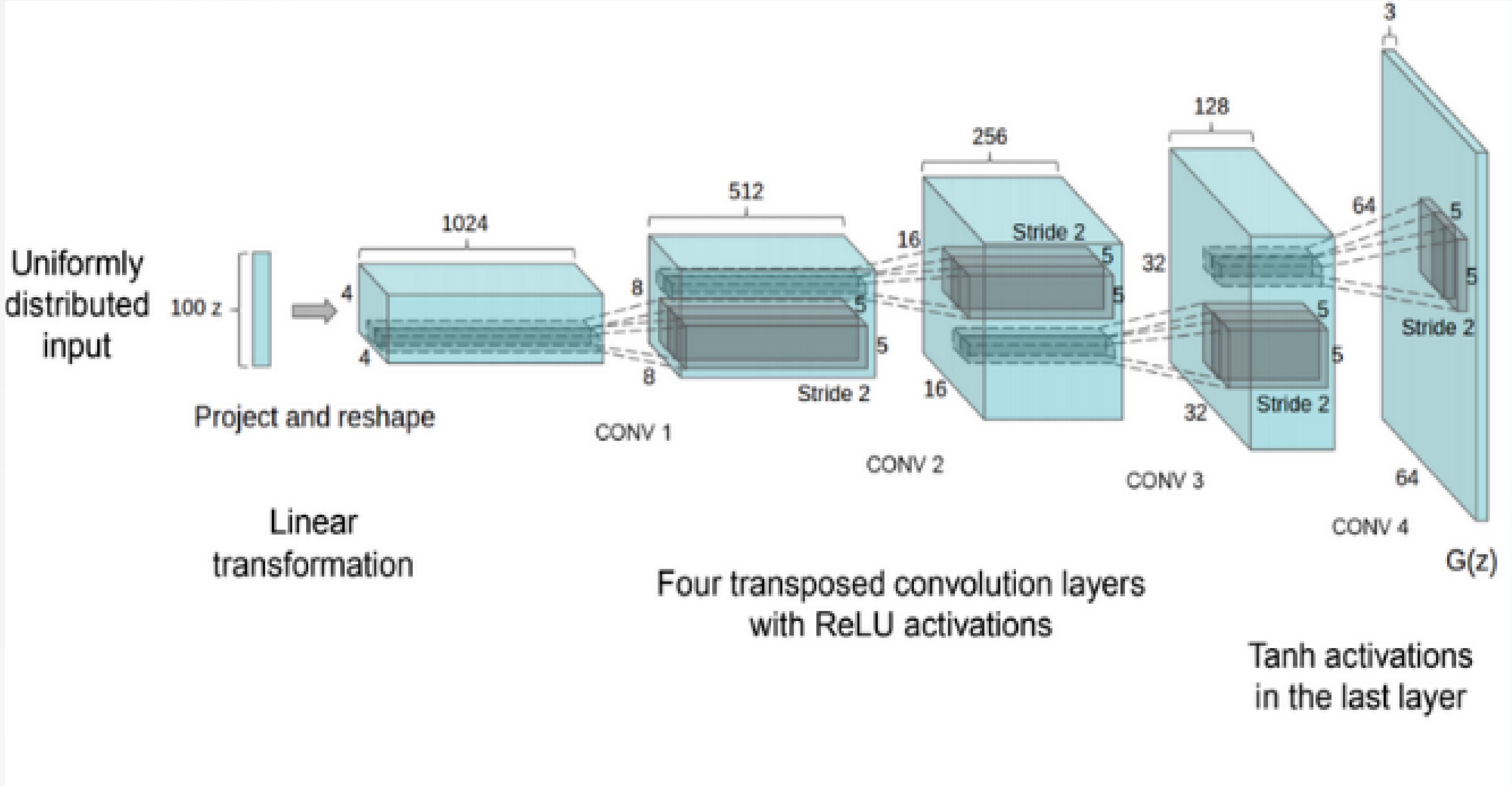
- It typically consists of several convolutional and pooling layers followed by a fully connected layer that outputs a scalar value representing the probability that the input data is real.

Loss function:

- The loss function used in Vanilla GAN is the binary cross-entropy loss function, which measures the difference between the discriminator's output for real data 1, and the discriminator's output for fake data 0.

DCGAN

DCGAN stands for Deep Convolutional Generative Adversarial Network. It is a type of GAN that uses deep convolutional neural networks.



DCGAN

Generator:

- Consists of several layers of fully connected, convolutional, and/or transposed convolutional layers.
- Takes a random noise vector as input and generates fake data, such as images or text.

Discriminator:

- Consists of a series of convolutional layers, followed by batch normalization and a Leaky ReLU activation function.
- The output is a scalar value indicating whether the input image is real or fake.

Loss function:

- The loss function used in Vanilla GAN is the binary cross-entropy loss function, which measures the difference between the discriminator's output for real data 1, and the discriminator's output for fake data 0.

AUTO ENCODING GAN

Autoencoding GAN (Adversarial Autoencoder) is a type of GAN that combines the architecture of an autoencoder and a GAN. It consists of an encoder network, a generator network, and a discriminator network.

- The encoder network is a deep neural network that takes input data and produces a latent code
- The generator network is another deep neural network that takes the latent code as input and generates fake data, such as an image or a text
- The discriminator network is a deep neural network that takes both real data and fake data as input, and tries to distinguish between them.
- The loss function used in Autoencoding GAN includes two parts: the reconstruction loss and the adversarial loss.

DATASETS

CelebA dataset:

The dataset consists of over 200,000 celebrity images, captured from the internet. The images cover a wide range of individuals, including actors, actresses, musicians, and other public figures.

MNIST Fashion dataset :

Similar to the original MNIST dataset, the MNIST Fashion dataset contains 60,000 training images and 10,000 testing images. Each image is grayscale and has a resolution of 28x28 pixels, resulting in a total of 784 pixels per image.

RESULTS

CELEBA DATASET:

GAN TYPE	Generator Loss	Discriminator Loss	Evaluation PSNR	Evaluation MSE
Auto encoding GAN	0.010754	0.000380	28.145482	99.722172
SRGAN	0.337	0.579	32.875262	82.398437
DCGAN	0.692	0.748	15.3267	183.2709

RESULTS

MNIST FASHION DATASET :

GAN TYPE	Generator Loss	Discriminator Loss	Evaluation PSNR	Evaluation MSE
Vanilla gan	1.2906	0.8752	51.5603	0.5117
LSGAN	0.4605	0.3149	50.8515	0.6175
DCGAN	0.000044	0.007917	32.9509	33.1110

OBSERVATIONS

- Tuning the hyperparameters yields better results.
- Less batch size and more number of epochs increases the efficiency of the model in most of the cases.
- CelebA Dataset is large enough which can train model efficiently but to train the model in systems having less specification, one must break down the dataset into smaller sample.
- ReLU prevents the overfitting of the model.
- Batch Normalization in the architecture stabilizes the training model and accelerates convergence.

CONCLUSION

- Both visual and quantitative results clearly show that Auto encoding GAN performs better than DCGAN.
- After running different datasets on the models, we conclude that Auto encoding GAN worked better than the other models on CelebA dataset, while on MNIST Fashion Dataset, Vanilla Gan proved to be better.
- With the increasing epochs, the reconstruction images by Auto encoding GAN become even better.
- The resulting images are less blurry and usable in practice.
- The overall impression of the Auto encoding GAN model is that it yields entirely satisfactory results.
- Considering the minimal improvement of the results in tests with increased layers, potential networks that are built even deeper is likely to yield the same results.
- However, the generated results by DCGAN are of poor quality and does not meet the hypothesis.

REFERENCES

- Vint, David & Anderson, Matthew & Yang, Yuhao & Ilioudis, Christos & Di Caterina, Gaetano & Clemente, Carmine. Automatic Target Recognition for Low Resolution Foliage Penetrating SAR Images Using CNNs and GANs. Remote Sensing. 13. 596. 10.3390/rs13040596, February 2021
- Yu, Timothy & Ma, Da & Cole, Jayden & Ju, Myeongjin & Beg, Mirza Faisal & Sarunic, Marinko. Spectral Bandwidth Recovery of Optical Coherence Tomography Images using Deep Learning. 67-71. 10.1109/ISPA52656.2021.9552122, September 2021.
- Radford, A., Metz, L., & Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, November 2015.
- P. Gong, J. Liu and S. Lv, "Image denoising with gan based model," Journal of Information Hiding and Privacy Protection, vol. 2, no.4, pp. 155–163, September 2020.
- Tran, L.D. Nguyen, S.M. and Arai, Masayuki. “GAN-based noise model for denoising real images”. In Proceedings of the Asian Conference on Computer Vision, November 2020.
- Yorozu, T., Hirano, M., Oka, K. and Tagawa, Y. Electron spectroscopy studies on magneto-optical media and plastic substrate interface. IEEE translation journal on magnetics in Japan, 2(8), pp.740-741, August 1987.
- Wei, B., Zhang, L., Wang, K. et al. Dynamic scene deblurring and image de-raining based on generative adversarial networks and transfer learning for Internet of vehicle. EURASIP J. Adv. Signal Process.121 (2021). https://doi.org/10.1186/s13634-021-00829-0, December 2021.
- Singh, Varsha, and Uma Shanker Tiwary. "Visual content generation from textual description using improved adversarial network." Multimedia Tools and Applications 82.7 : 10943-10960, September 2022.

THANK YOU!