Image Enhancement using GANs II

Group 7

Submitted By:

Soumya Ranjan Naik
Ansuman Samal
Azazul Haque
Hrushikesh Sahu
Md.Azharuddin

Guided By:

Dr. Sanjay Kumar Patra
Dr. Sangita Pal

Recap...

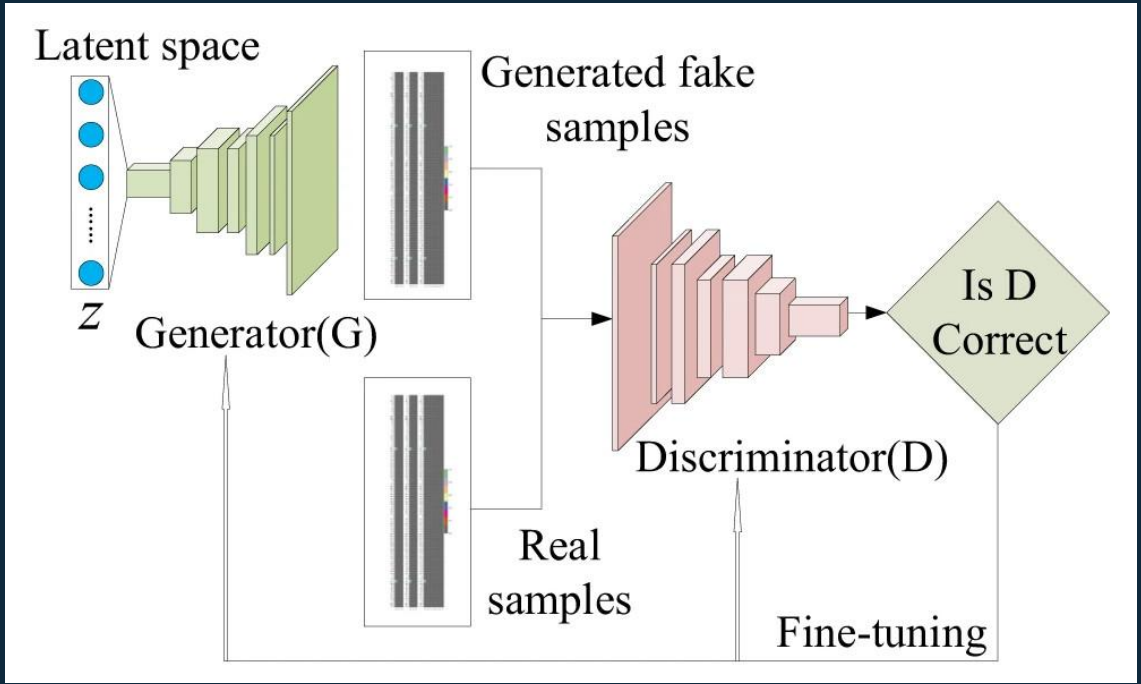


Image Source: https://www.nature.com/articles/s41524-020-00352-0

Latent Space

- N-dimensional input vector to Generator model.
- Each variable is drawn from a Gaussian distribution with a mean of zero and a standard deviation of one.
- Generator learns to map points onto the latent space with specific output images. This mapping varies each time the model is trained.
- Points in the latent space can be kept and used in simple vector arithmetic to create new points in the latent space.

0	-0.5	0.456	1	-0.3

1-D Latent vector example

Arithmetic

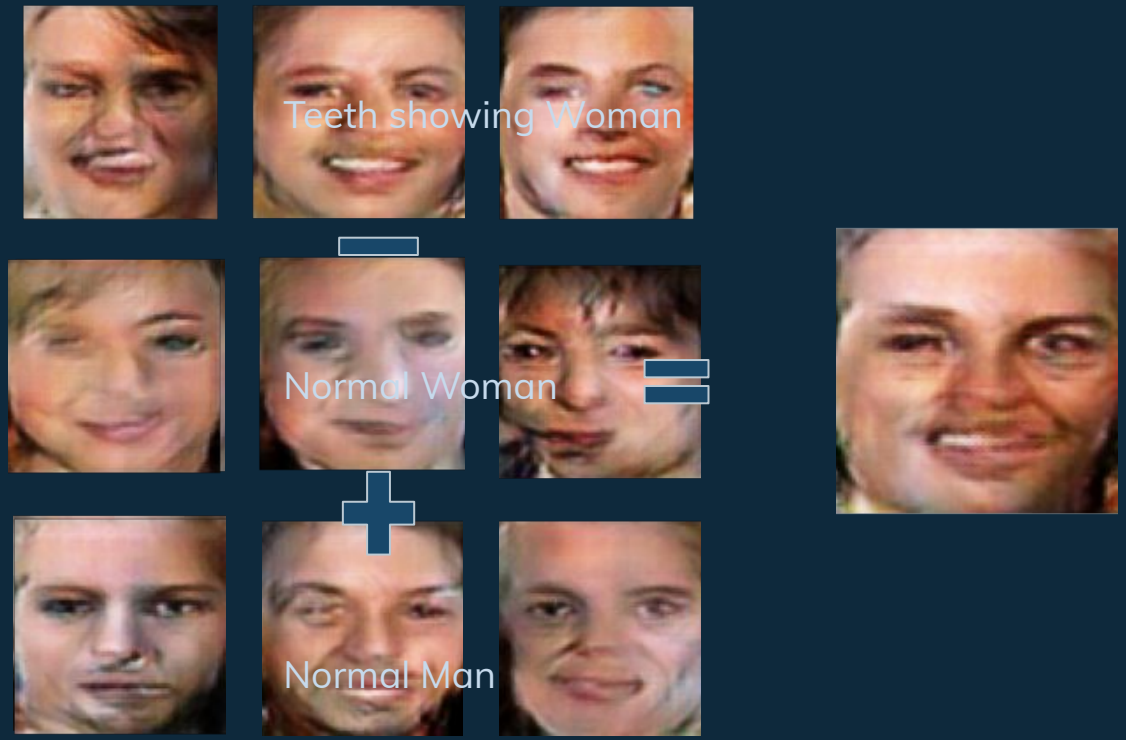


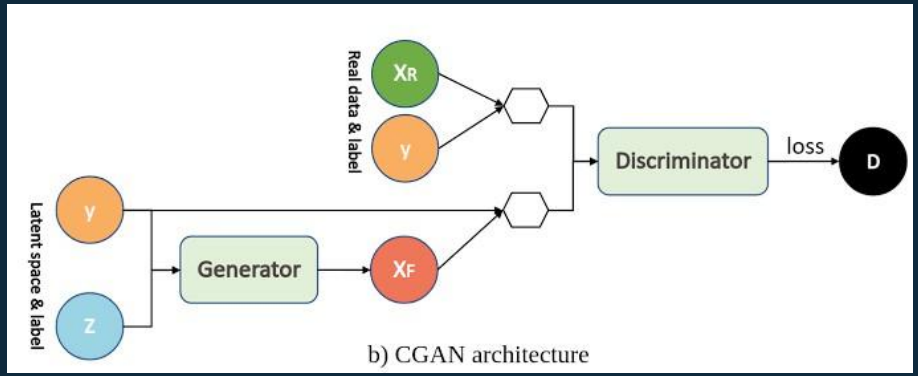
Image Source: Our Laptop

Transition

Image Source: Our Laptop

Conditional GAN (cGAN)

- **Problem it overcomes:** No control over output
- Additional information is provided to Generator and Discriminator that is correlated with the input images, such as class labels.
- By conditioning the model on additional information it is possible to direct the data generation process.



b) CGAN architecture

Image Source: google.com

Fashion MNIST

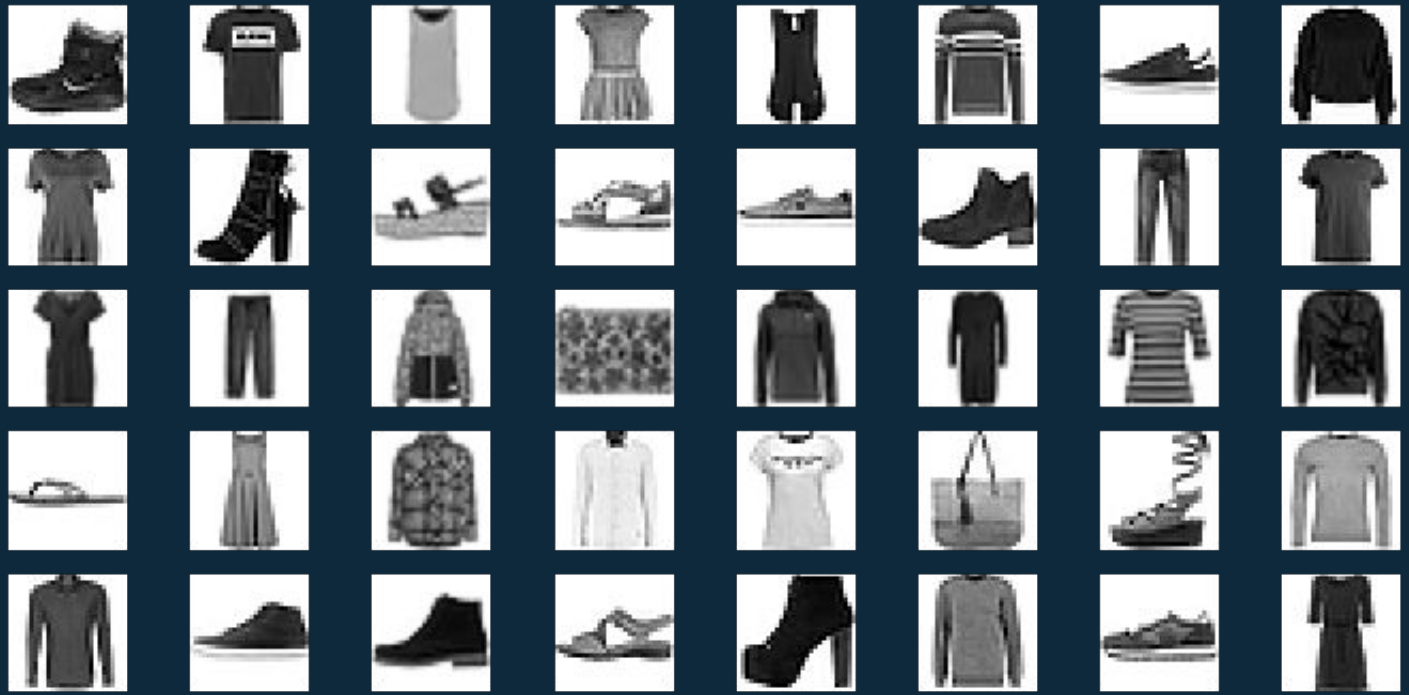
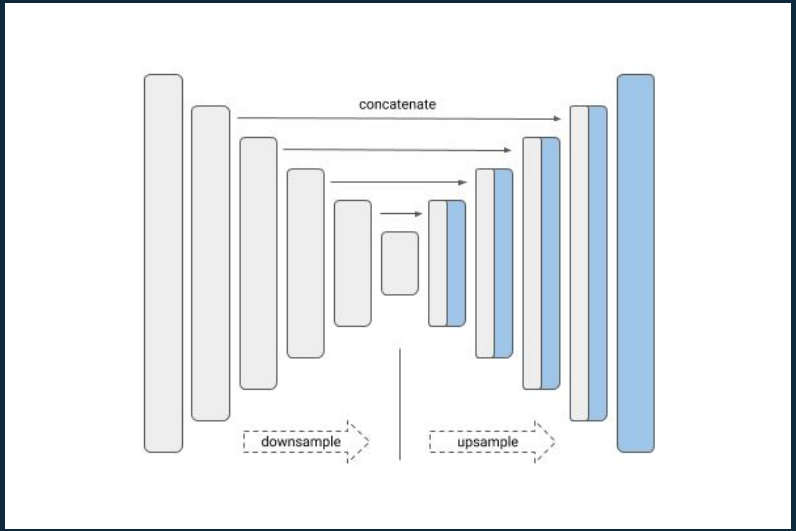


Image Source: Our Laptop

Our cGAN Output

Image Source: Our Laptop

U-Net Architecture



- A replacement of common Encoder-Decoder model
- It contains links or skip-connections between layers of the same size in the encoder and the decoder.
- This gives the generator a means to circumvent the bottleneck for information.

Image Source: kaggle.com

PatchGAN Architecture

- Not actually GAN, rather a Discriminator model.
- Introduced in paper “*Image-to-Image Translation with Conditional Adversarial Networks, 2016*” .
- It is designed to classify patches of an input image as real or fake, rather than the entire image.
- The output is a single feature map of real/fake predictions that can be averaged to give single score.
- A patch size of 70 × 70 was found to be effective across a range of image-to-image translation tasks.

Pix2Pix GAN

- Pix2Pix stands for Pixel to Pixel
- It is an approach to train deep CNN for image-to-image translation.
- It is a type of cGAN where the condition is the input image itself.
- The U-Net architecture is prefered for generator.
- The PatchGAN architecture is preferred for discriminator.

References

- Alec Radford, Luke Metz & Soumith Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks”
- Shuo Yang, Ping Luo, Chen Change Loy & Xiaoou Tang, “From Facial Parts Responses to Face Detection: A Deep Learning Approach”
- Mehdi Mirza & Simon Osindero, “Conditional Generative Adversarial Nets”
- Phillip Isola, Jun-Yan Zhu, Tinghui Zhou and Alexei A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks”
- Generative Adversarial Networks with Python by Jason Brownlee

Thanks!

Any questions?

http://bit.ly/igit36_group7