Facial Image Generation

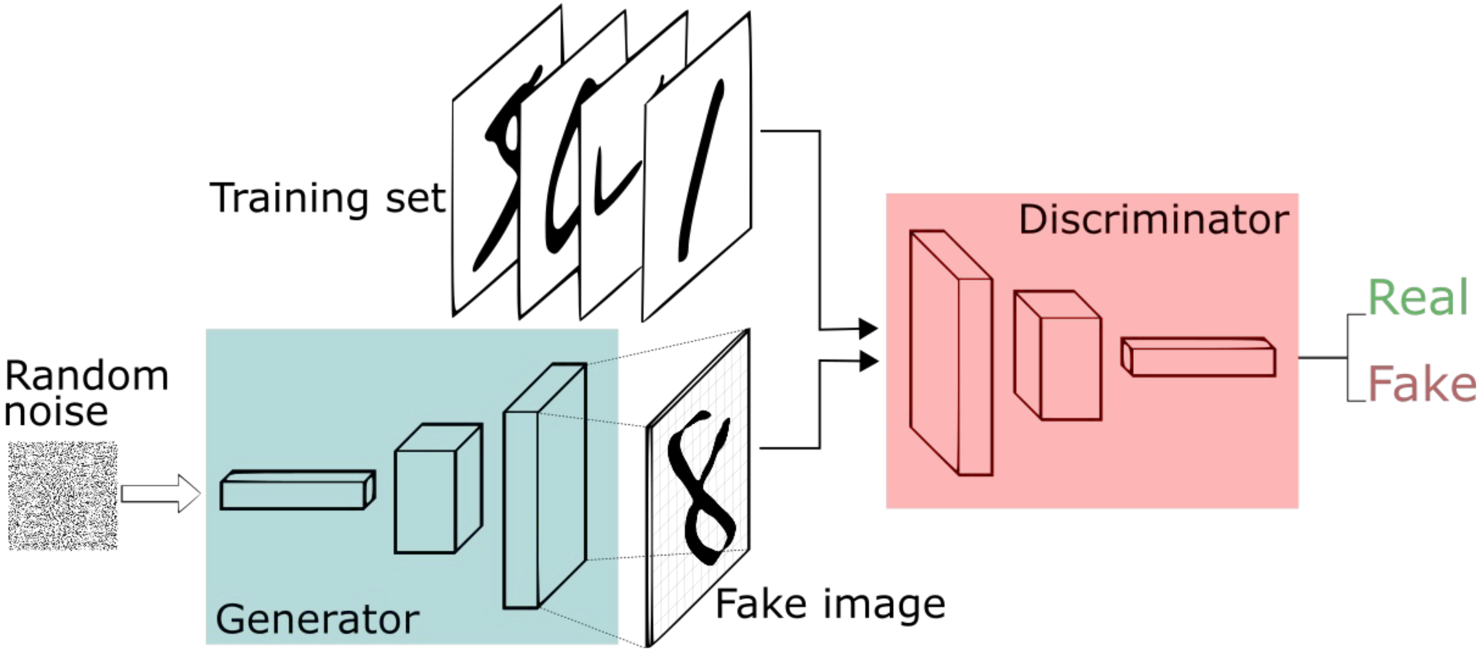
Suhas Buravalla,
Mohammad Maaz,
Carter Rogers

Dataset

- The CelebA dataset is a widely used dataset for generative deep learning
- Consists of over 200,000 high-resolution images of celebrity faces

Generative Adversarial Networks

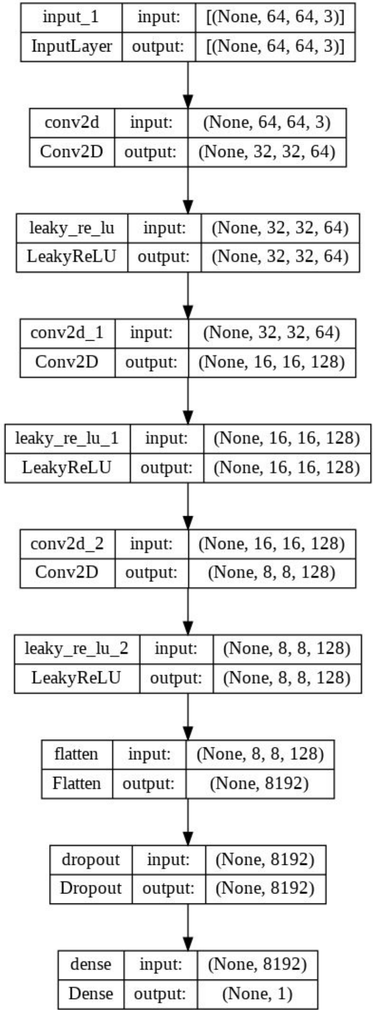
Network Architecture



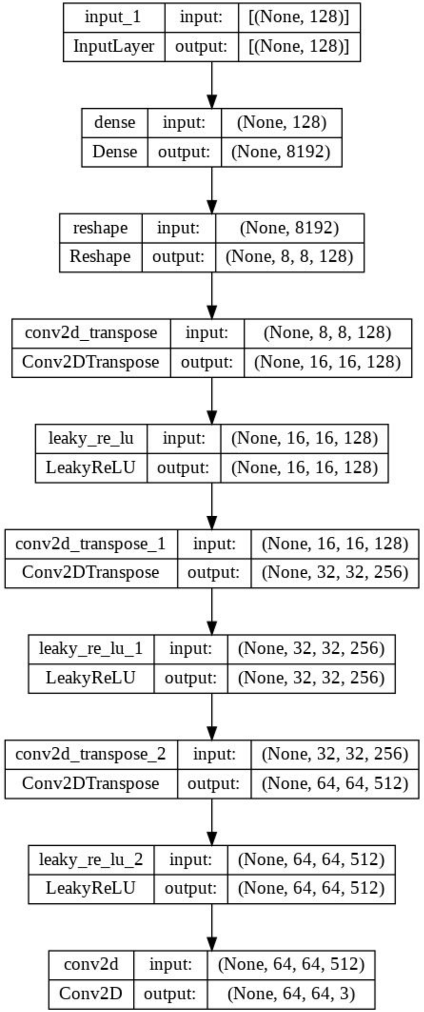
Algorithm

1. Initialize the generator and discriminator networks
2. For each batch of real images:
   a. Generate a batch of fake images using the generator
   b. Train the discriminator on the fake and real data
   c. Train the generator on a batch of fake images, using the output of the discriminator as feedback
3. Repeat this process until the generator produces data that is indistinguishable from the real data

Discriminator



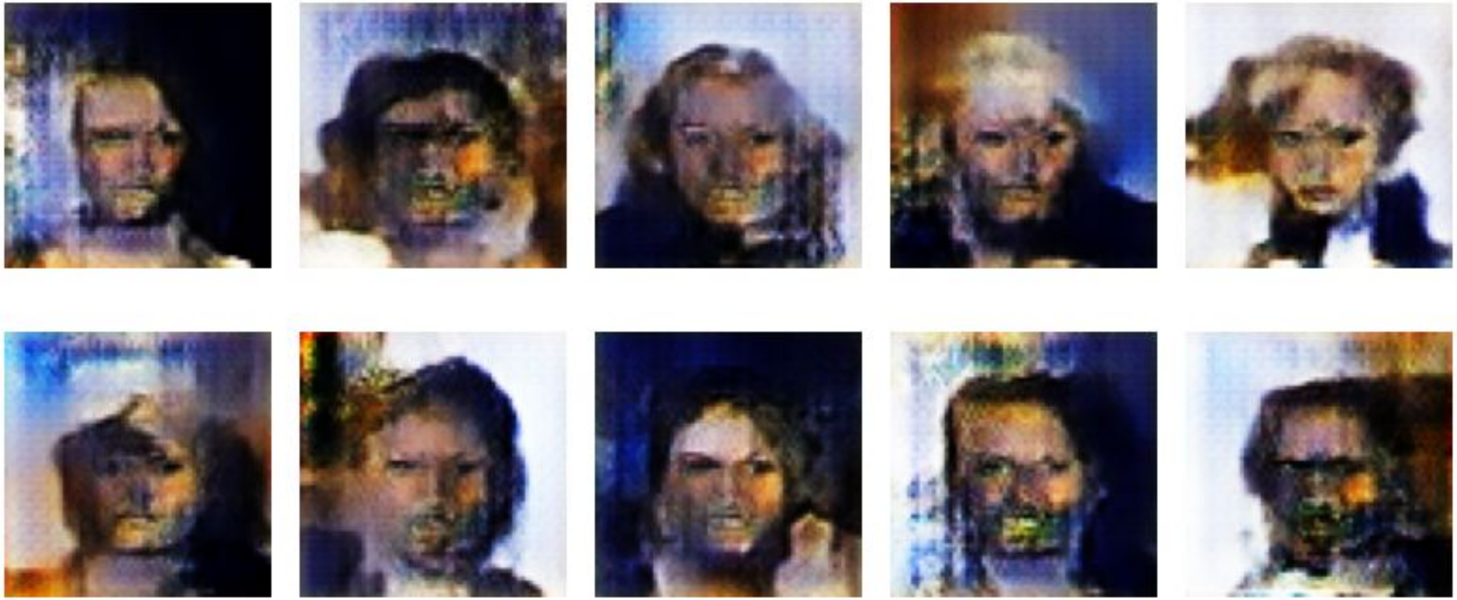
Generator



Training


Epoch: 0 Disc loss: 0.33 Gen loss: 2.55
Epoch: 10 Disc loss: 0.65 Gen loss: 1.06
Epoch: 20 Disc loss: 0.60 Gen loss: 0.99
Epoch: 30 Disc loss: 0.55 Gen loss: 0.85

Results



Variational Autoencoders

VAE Algorithm

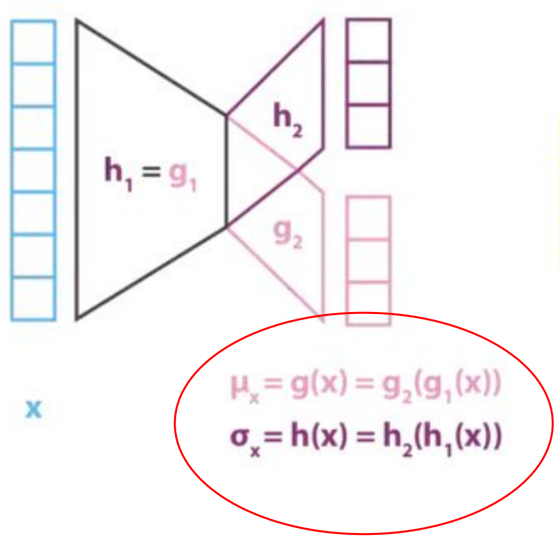
Autoencoder (similar to GANs) - encoder and decoder

Dimensionality reduction using PCA - preserve as much information from input as possible

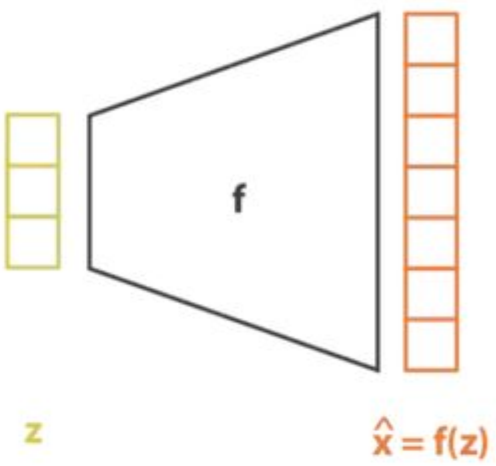
What makes VAEs different from GANs?

- Latent (encoded) space must be regularized
- Without this, model WILL overfit (only trained to minimize loss)
- Minimize reconstruction error
- Encoder returns distribution over latent space (instead of single point)

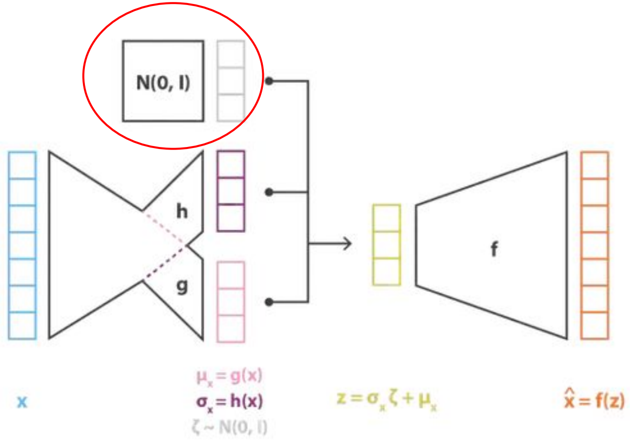
VAE Network Architecture



Encoder



Decoder

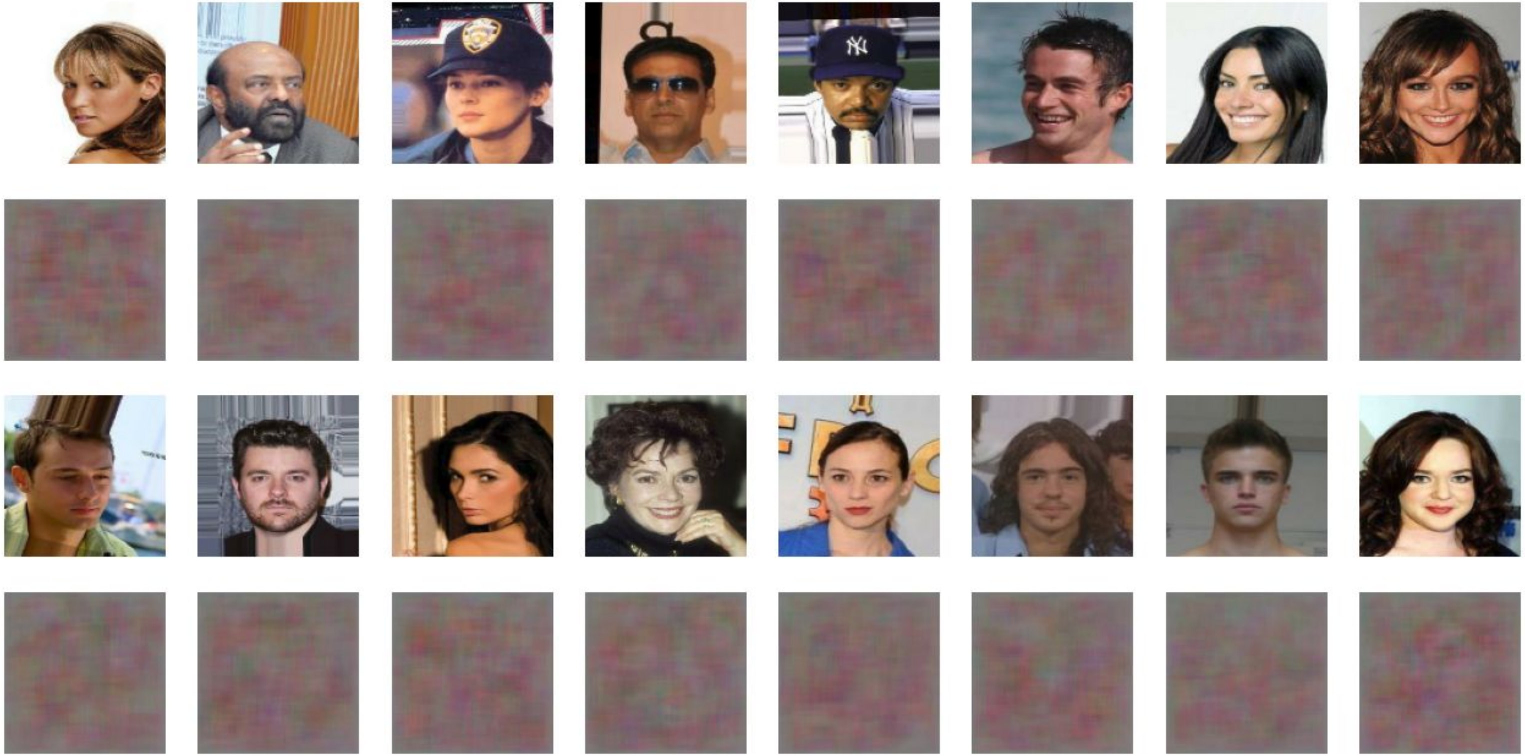


Full network

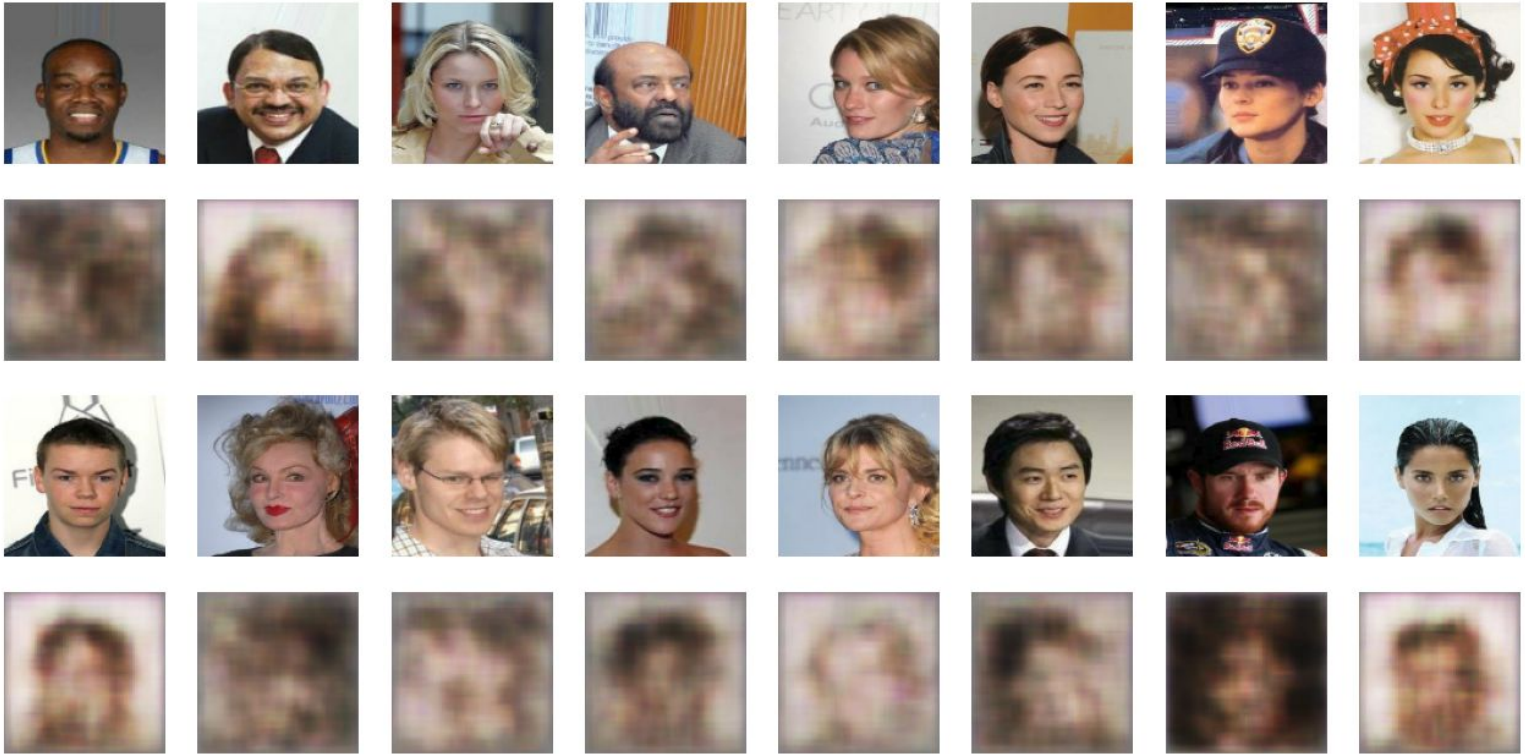
VAE Training

- 112x112 images, 50 epochs
- Less computationally expensive than GANs
- Batch size of 128, learning rate of .001
- Train/test split (unlike GANs and WGANs)
- Kullback-Leibler Divergence (kl loss) and BCE loss
    - *kl loss used to calculate difference between actual and observed probability distribution*

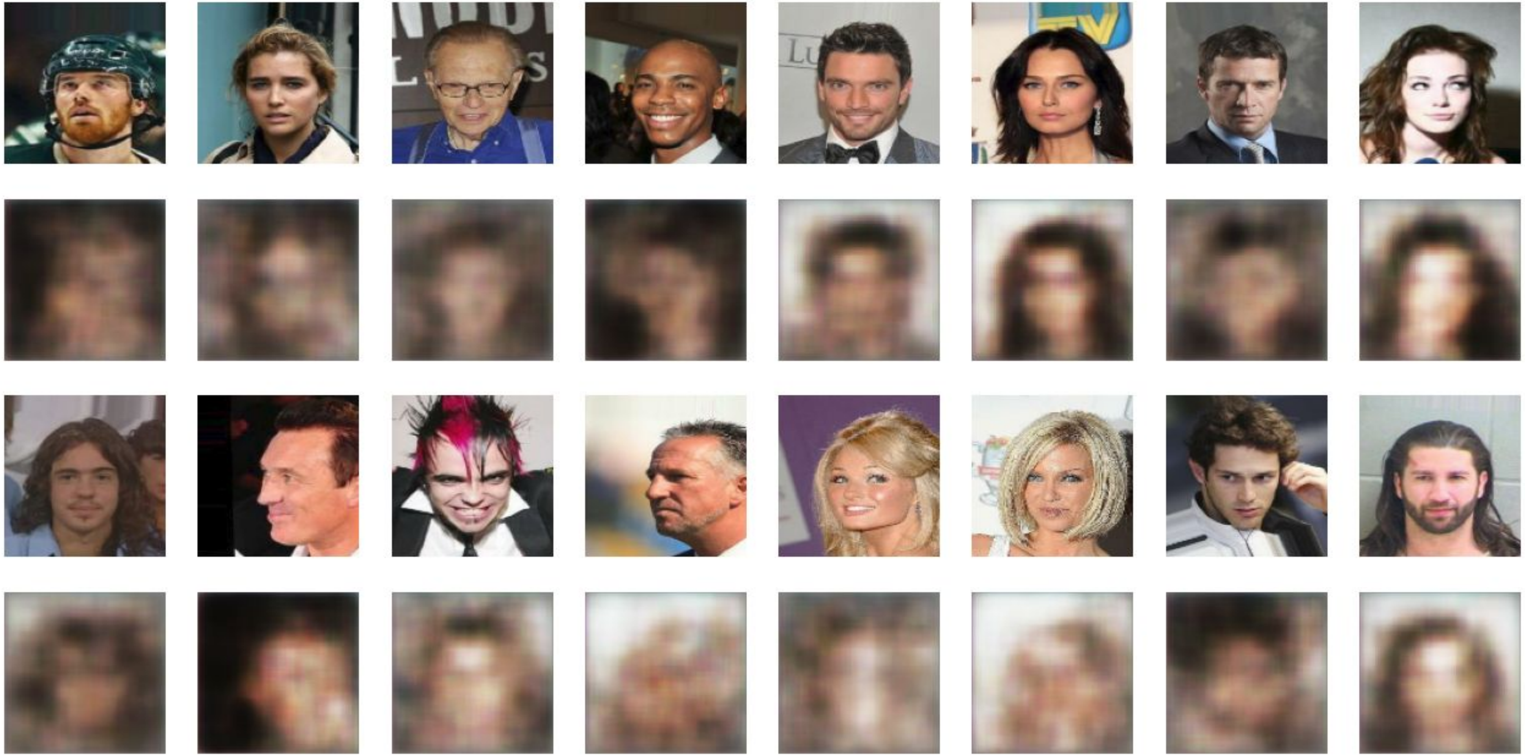
VAE Results - 1 epoch



VAE Results - 20 epochs



VAE Results - 50 epochs



VAE Results

- Strengths: able to identify outlines and shapes of faces, less computationally expensive than GANs
- Weaknesses: lack of definition in facial features (better with 50 epochs)
- Why are VAEs less popular than GANs?
    - “Generative Adversarial Networks is the most interesting idea in the last 10 years in Machine Learning.” — Yann LeCun, Director of AI Research at Facebook AI
    - However, VAEs are a viable alternative if you don’t need photorealistic images

Conclusion: VAEs easier to train, but underperform GANs as a generator

Wasserstein Generative Adversarial Networks With Gradient Penalty

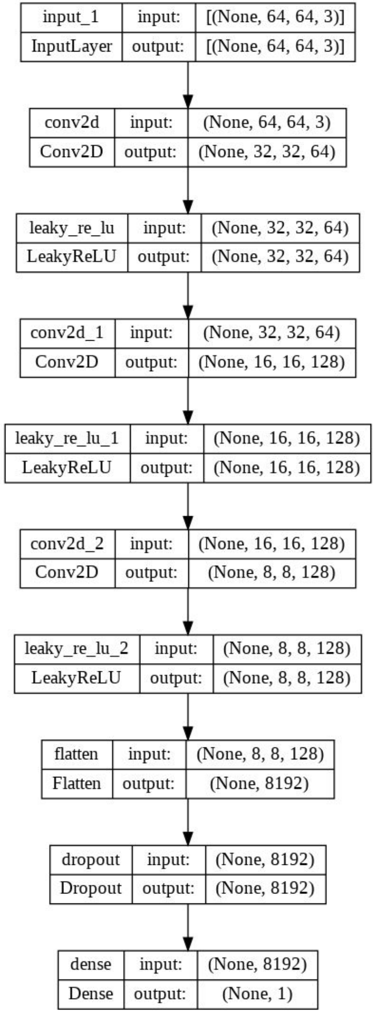
Why WGAN?

- Improve the stability of training and prevent errors like mode collapse
- GANs were viewed to have high potential but they did not perform as optimally as expected when it came to really large computations.
- WGAN discriminator provides a better learning signal to the Generator when compared to the GAN discriminator.

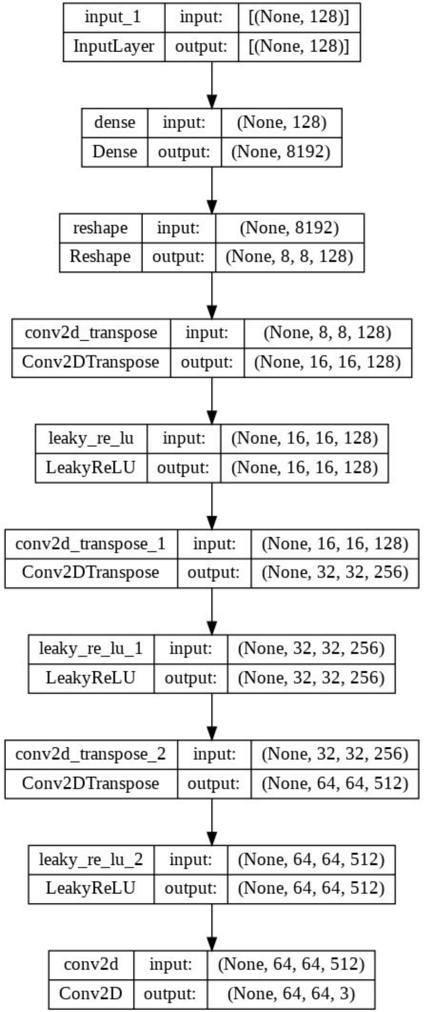
WGAN with GP

- WGAN was first implemented with weight clipping method
- Improved version of WGAN was introduced with Gradient Penalty method
- Instead of clipping weights we introduce a loss term called “Gradient penalty” which is added to the norm of discriminator gradient

Discriminator



Generator



Training

- Number of epochs = 30
- Training time was roughly about 70 minutes for each epoch.
- Learning rate = 0.002
- Wasserstein distance used to calculate the distance between real and generated probability distributions.

Results

Epoch: 10 Disc loss: 0.10 Gen loss: -0.43

Epoch: 20 Disc loss: 0.30 Gen loss: -0.36

Epoch: 30 Disc loss: -0.08 Gen loss: -0.49

Results

- Features such as nose, eyes, hair, etc are prominent in the results produced by WGAN compared to the other networks.
- The clarity of images is better, leading to a better quality generated image.
- Distortion is observed in the generated images
- Relatively more accurate compared to the images generated by GAN model and VAE model

References

- Cheong, Soon Yau. Hands-On Image Generation with TensorFlow: A Practical Guide to Generating Images and Videos Using Deep Learning. Packt Publishing, 2020.
- Chollet, Francois. “DCGAN to generate face images.” Keras, 29 April 2019, https://keras.io/examples/generative/dcgan_overriding_train_step/. Accessed 12 December 2022.
- Chollet, Francois. “WGAN-GP overriding `Model.train_step`.” Keras, 9 May 2020, https://keras.io/examples/generative/wgan_gp/. Accessed 12 December 2022.
- Joseph, Rocca. “Understanding Variational Autoencoders (VAEs) | by Joseph Rocca.” Towards Data Science, 23 September 2019, https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73. Accessed 12 December 2022.
- Liu, Ziwei. “CelebA Dataset.” MMLab@CUHK, https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html. Accessed 12 December 2022.