Cool GANs and Fake Celebrities

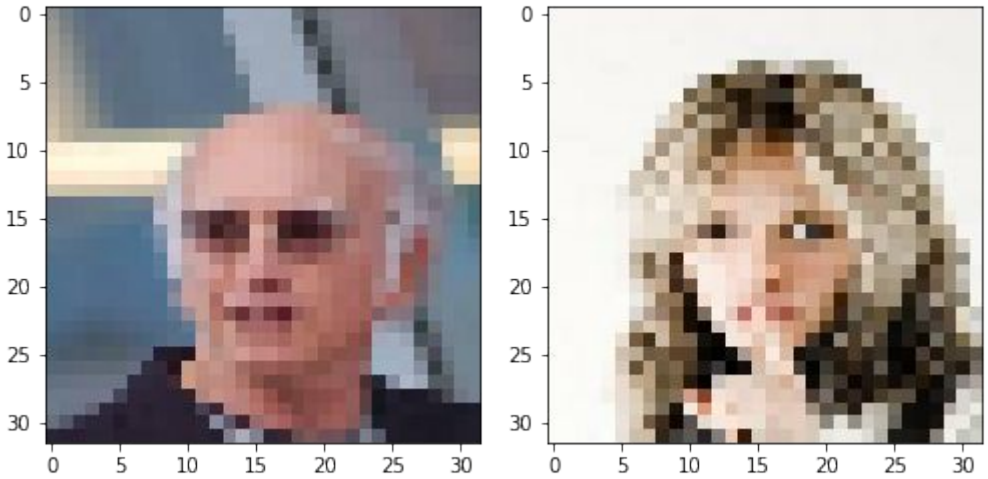
Project in DD2424 Deep Learning in Data Science

Diogo Pinheiro Jakob Lindén Márk Csizmadia Patrick Jonsson

KTH Royal Institute of Technology

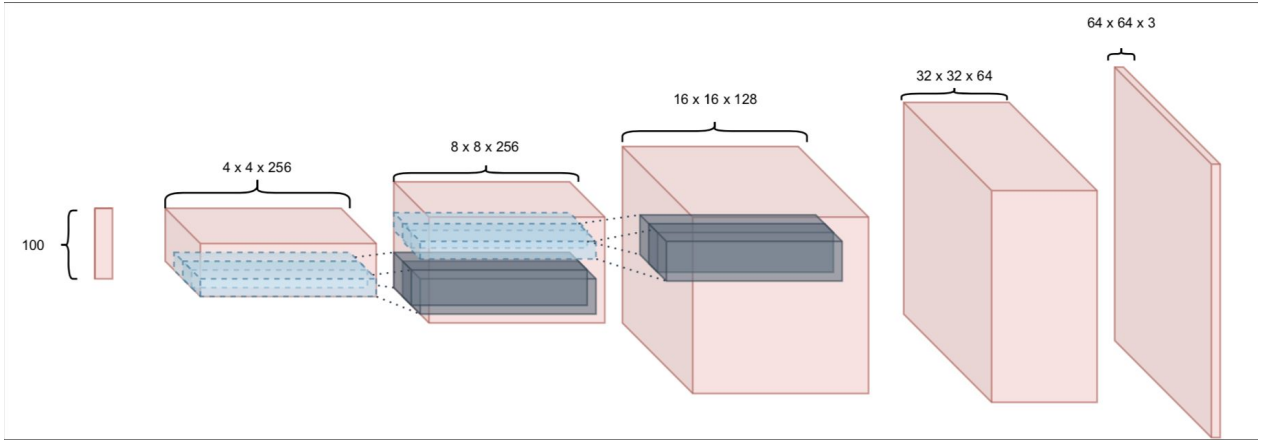
May, 2021

Data Overview



Method - DCGAN

- Replaces all deterministic spatial pooling functions (e.g. max pooling) with strided convolutions;
- Eliminates fully connected hidden layers.



Method - Fréchet Inception Distance (FID)

- Goal: Compare images generated by the model and the real images in the dataset;
- Inception v3 model computes computer vision statistics adequate for comparison;
- Compares two Gaussians, one representing the real images and one representing the generated images;
- Distance is measured between the mean and covariance associated to the two Gaussians.

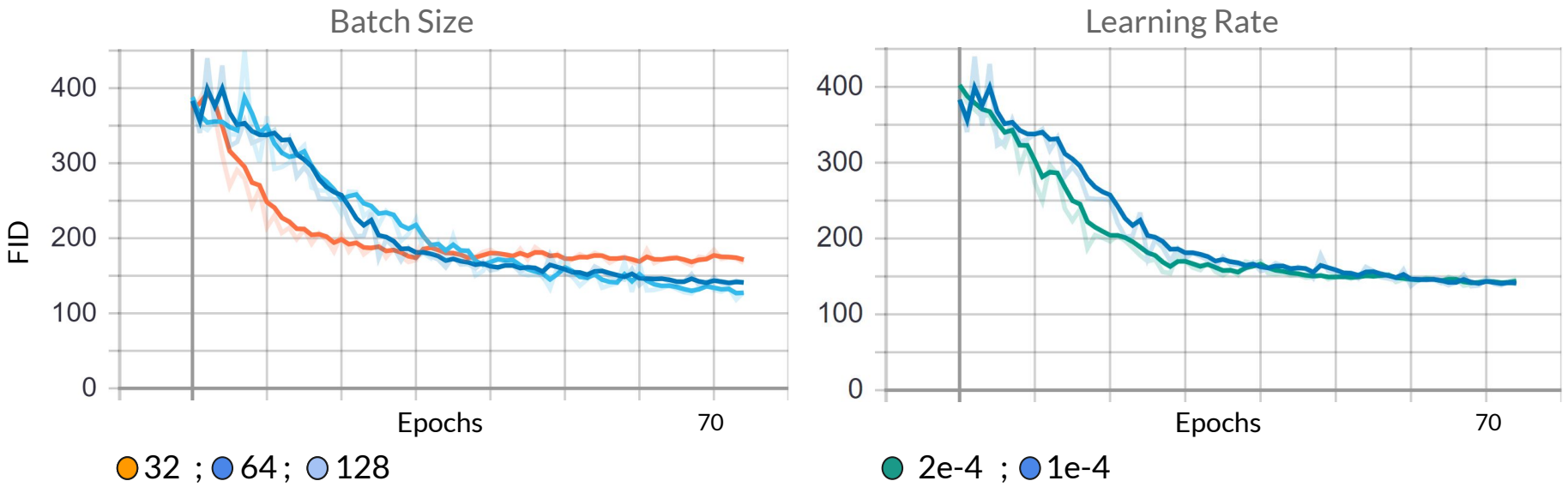
$$d^2\left((m, C), (m_w, C_w)\right) = \|m - m_w\|_2^2 + Tr\left(C + C_w - 2\left(CC_w\right)^{1/2}\right)$$

Guidelines [1]

- Replace any pooling layers with strided convolutions (discriminator) and fractional-strided convolutions (generator);
- Use batchnorm in both the generator and the discriminator;
- Remove fully connected hidden layers for deeper architectures;
- Use ReLU activation in generator for all layers except for the output, which uses Tanh;
- Use LeakyReLU activation in the discriminator for all layers;
- Use learning rate of 2e-4 instead of 1e-3;
- Use ß1=0.5 instead of 0.9

[1] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016

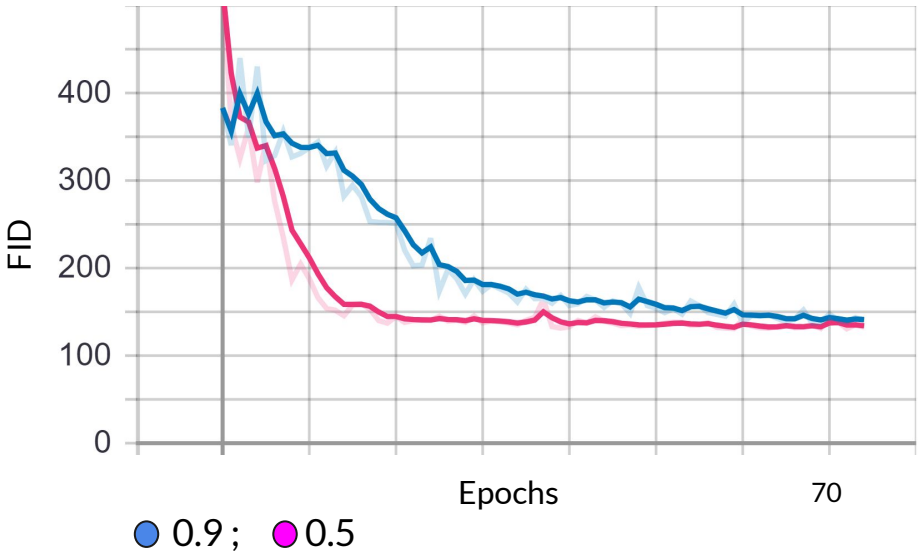
Experiments



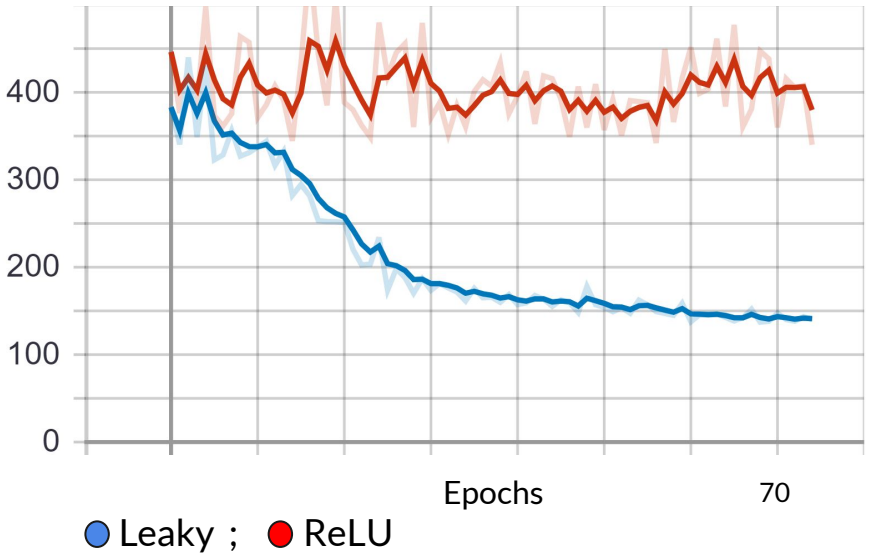
Experiments

Sparse Gradients

Beta1

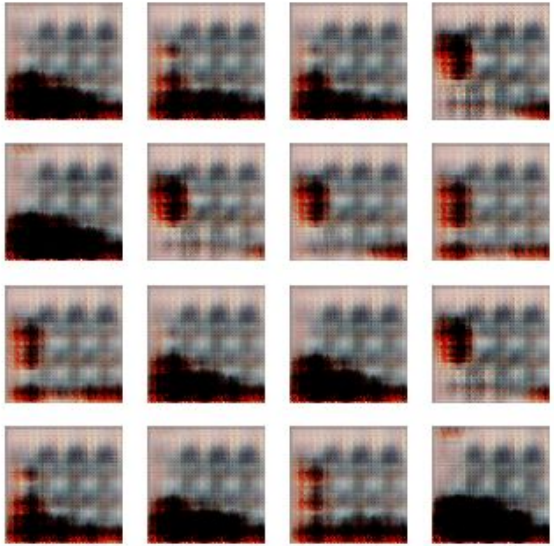


Leaky vs ReLU in generator



Final Result

32x32



64x64

Final Result

32x32

64x64

Future Work

- Optimize hyperparameters;
- Increase image size;
- Vector arithmetic for visual concepts [1]

[1] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016