GAN: How does a mode collapse look like?

11. Februar 2020

1 Our Data

We used a set of 40 number of images taken from a white color digit of 5 printed on a black background. The images are augmented by a set of scaling and a small angles of rotation in ± 5 degrees.

2 Our Model

- **Discriminator**: It takes as input a 28×28 grayscale image and produces a binary output, namely 0 if the input image is fake and 1, if it is real. The discriminator is implemented in form of a convolutional neural network with the LeakyReLU activation function and batch normalization. The used optimizer is the adam version of stochastic gradient descent.
- **Generator**: It takes as input a *flexible-size* vector from a latent space and produces a gray 28×28 image. A fully connected layer converts the taken vector from the latent space to another activated vector that will be reshaped into a set of low-resolution gray images. Later, we upsample such low-resolution images while adopting LeakyReLU activation. The final output of the generator will be of size 28×28, as we already mentioned.

3 Mode Collapse

As stated in [1], a GAN is said to be in a mode collapse if its generator is not able to produce the entire spectrum of possible outputs. In other words only a small subset of possible outcomes are produced by the generator.

4 Identifying a Mode Collapse

To produce a mode collapse scenario, we trained our GAN two times with the same data. The only difference was the length of the taken vector from

the latent space, that was considered to have a convenient size of 128 for a normal training and later reduced to 1 so that a mode collapse can be observed.

In case, the latent-vector has a length of 128, the GAN training is performed well as shown in Fig. 1. Here, the loss and the accuracy of the fake and real images obtained from the discriminator are shown in orange and blue plots. As one observes, the discriminator losses with respect to the fake and real images stay almost constantly around the value of 0.5, in epoch period of 10 to 100 (See Fig. 1-Up). It clearly means, the produced fake images by the generator are in a good quality so that the discriminator fails to label them as fake. By looking at the Fig. 1-Bottom, we observe that the accuracy of the discriminator remains once again around 0.5, conveying that the fake images are as good as real images so that in almost half of the time the real images are labeled by discriminator as fake ones. The horizontal axis in Fig. 1 represents the number of epoches.

To observe a mode collapse, as already mentioned, the latent-vector size is reduced to 1. This results to the generator loss shown as green curve in Fig. 2-Up to keep on oscilating during the entire time of the training.

Literatur

[1] Goodfellow I.J., NIPS 2016 Tutorial: Generative Adversarial Networks, 2016; arXiv:1701.00160.

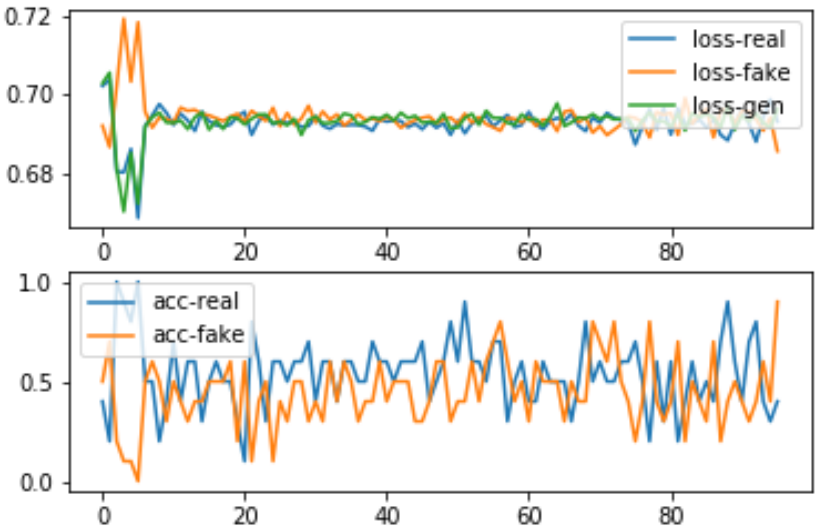

Abbildung 1: The loss and the accuracy of the discriminator with respect to the fake images shown as orange. The same loss and accuracy obtained from the discriminator and this time with respect to the real images are also shown in blue. **(Up)** The loss values of the discriminator with respect to the fake and real images. The loss values stay around 0.5 meaning the discriminator can not properly distinguish between the reak and the fake images. In addition, the generator loss is shown in green and stays consistent from the epoch 10 till 100. **(Bottom)** The plot shows the accuracies of the discriminator that stay almost around 0.5 with respect to the real and the fake images. It means half of the time, the real images are labeld as fake by the discriminator, and vise versa.

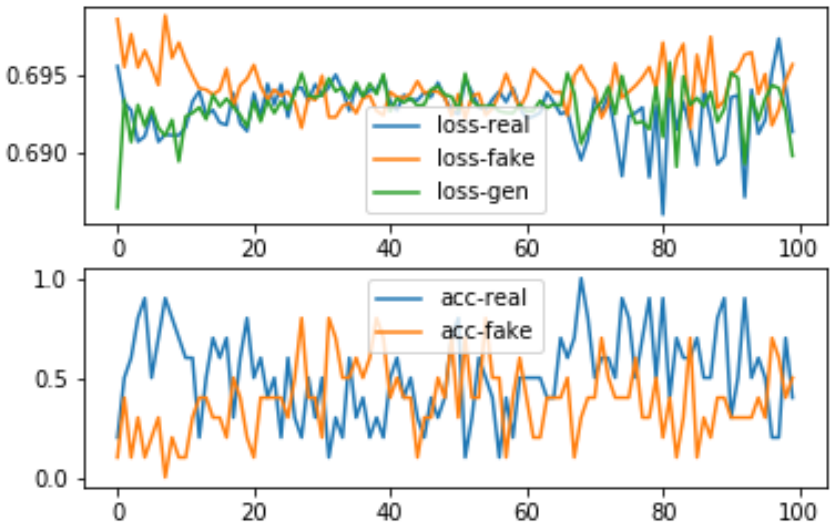

Abbildung 2: **(Up)** The generator loss shown in green keeps on oscilating during the entire period of training till the last epoch of 100. In addition, the discriminator loss with respect to the fake and real images shown as orange and blue also oscilate, specifically after the epoch number 70. Such oscilation in loss values are considered as the sign of a mode collapse. **(Bottom)** The loss values of the discriminator shown in orange and blue correspond to the fake and real images. Though they both oscilate around the value of 0.5, but it may not result to a variety of created images by the generator. In fact looking at the created image by the generator in this case, we some repeated patterns are created. Such repeatation is the result of reducing the latent-vector size to 1. In case of taking latent vectors of sitze of 128, one sees a variety of generated fake images.