Evaluation of Spectral Normalization for GANs Using Inception Score

Eysteinn GUNNLAUGSSON, Egill VIGNISSON, Charles HAMESSE

Group 16

Abstract. The abstract should summarize the contents of the paper. LNCS guidelines indicate it should be at least 70 and at most 150 words. It should be set in 9-point font size and should be inset 1.0 cm from the right and left margins. . . .

Keywords: Generative adversarial networks, generative models, image generation

1 Introduction

Motivate the problem you are trying to solve, attempt to make an intuitive description of the problem and also formally define the problem. (1-2 pages including title, authors and abstract)

The purpose of this project is to investigate the performance of the different types of Generative Adversarial Networks (GANs) [1] for image generation as well as possible improvement options. The project was originally defined based on varying levels of priority where everything assigned priority 1 was promised to be completed.

- Implement Deep Convolutional Generative Adversarial Network with original loss [2] (priority 1)
- Implement the inception score metric citesalimans2016improved (priority 1)
- Implement Spectral Normalization (priority 1)
- Evaluate all our GANs on our reptile dataset (priority 1)
- Implement other losses (LSGAN, WGAN) [3] (priority 2)
- Evaluate GANs on CIFAR-100 (priority 2)
- Implement mini-batch discrimination and or other improvements [4]. (priority 3)

In order to evaluate the performance of these GANs, the evaluation metric known as the inception score, as described in [4], will be implemented. Furthermore a data set consisting of roughly 30K animal pictures (mostly reptiles), was fetched from the Flickr API while other well known options such as a subset of CIFAR-100 or ImageNet were thought of as possible replacements in the case of unsatisfactory results.

2   Background

Summarize a few notable approaches/papers tackling the same problem. The selection should cover different possible techniques that can be (have been) used for the same task with success. Also, it is good to mention other recognition/synthesis tasks that use the same deep learning technique as yours. (1-2 pages)

2.1   Generative Adversarial Networks

2.2   Inception Score

2.3   Spectral Normalization

3   Approach

Describe the final approach you are take for this problem. For instance, here you would describe the details of the network's architecture. What training parameters and techniques you have used. The computational complexity of your model. And similar questions. To help explain your approach please make figures to accompany your text description. (1-3 pages)

3.1   Data Collection

Images were collected using Flickr's API. Due to varying amounts of quality pictures of objects that would be interesting to base the image generation on, a collection that included mostly reptiles with a fair amount of arachnid's thrown into the mix was ultimately settled upon. In total the data set is made up of approximately 30k color images all re-sized to dimensions of 108x108x3 before training was conducted.

3.2   Frame of Reference

In order to measure the performance of each implementation a frame of reference had to be established. As the golden standard, that every implementation would be compared to, the inception score of the actual data set was calculated resulting in a score of **1.5295098 ± 0.06244182**.

4   Experiments

In this section, you should present the results you achieved with various experiments. The results can be presented in tables, plots, etc.

The purpose of the project is to analyze the nature and effectiveness of different GAN architectures as well as different improvement options. The different models but in order to do so a frame of reference was needed

4.1   Inception Score Considerations

We present a challenge related to the computation and evaluation of the inception score. Most authors evaluate the inception score on 50K GAN-generated images, as recommend the authors of the original paper [4]. By running a few preliminary experiments, we quickly realize that on top of the actual training of the network, sampling and computing the inception score are also resource-intensive tasks, and sampling 50K images is simply not possible with the time or resources available for this project.

Now, the number of images considered for evaluating the inception score has an impact on this score, as shows Table 1. This is due to the fact that the inception score not only evaluates the content of a given image but also the distribution of categories amongst the whole set of images resulting from the split. In other words, the score is sensitive to the number of images divided by the number of splits.

Images	Splits	Inception score	Images	Splits	Inception score
256	5	8.13 +- 0.41	256	10	6.72 +- 0.55
512	5	8.04 +- 0.54	512	10	7.92 +- 0.56
1024	5	9.79 +- 0.36	1024	10	

Table 1: Inception score for various number of samples of the cifar10 dataset.

We choose to stick with 1024 generated images and 5 splits for all of our experiments. With this configuration, we have a target inception score of 9.79. As expected, this is below the claimed inception score of the whole cifar10 dataset, 11.24 [4]. Thus we won't reach state of the art results in terms of inception score, but this isn't an issue since our purpose is to compare various improvements of GAN networks, which isn't affected by this choice.

Other considerations on the inception score are explained in [5].

4.2   DCGAN

In order to establish a frame of reference a plain vanilla version of DCGAN using (Fig. 3 shows an example).

Unfortunately we can hardly interpret the losses of GANs, since the generator and discriminator are in a situation of competition where an improvement on the one leads to a deterioration on the other. Consequently, there exists a number of approaches to try and make these networks reach an equilibrium quicker. **Todo: insert citations or examples if not already done in background theory.** However, this aspect of the art of training GANs goes below the scope of our work.
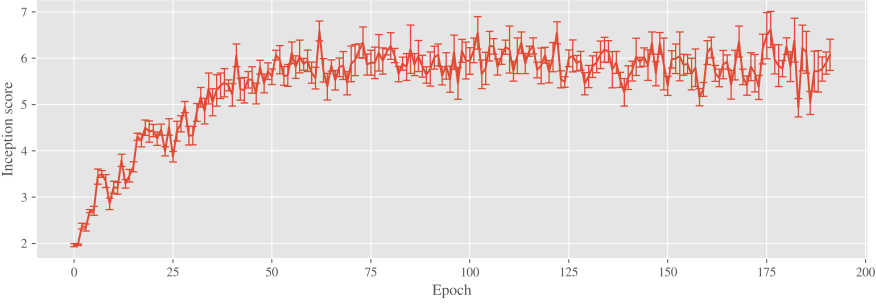
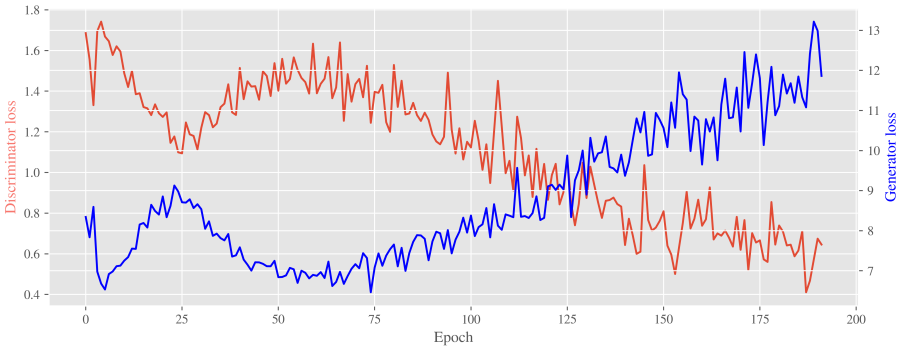Fig. 1: Inception score over 190 epochs with Vanilla DCGAN.

Fig. 2: Losses over 190 epochs with Vanilla DCGAN.

4.3 DCGAN with Spectral Normalization

4.4 LSGAN

4.5 WGAN

5 Conclusions

Explain what conclusions you can draw from these set of experiments? The set of experiments and results reported here should justify some of the design choices described in the previous sections. (3-6 pages)

References

1. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in neural information processing systems. (2014) 2672–2680

2. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR **abs/1511.06434** (2015)

3. Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., Smolley, S.P.: Least squares generative adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE (2017) 2813–2821

4. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training gans. In: Advances in Neural Information Processing Systems. (2016) 2234–2242

5. Barratt, S., Sharma, R.: A note on the inception score. arXiv preprint arXiv:1801.01973 (2018)

A   Template examples

Table 2: Font sizes of headings. Table captions should always be positioned *above* the tables. The final sentence of a table caption should end without a full stop

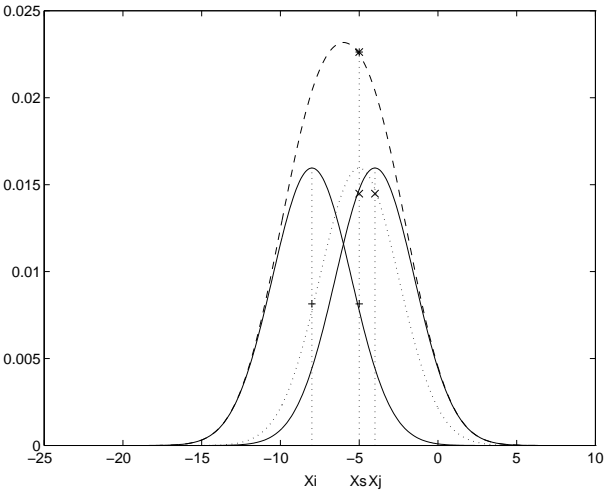| Heading level	Example	Font size and style
Title (centered)	**Lecture Notes ...**	14 point, bold
1st-level heading	**1 Introduction**	12 point, bold
2nd-level heading	**2.1 Printing Area**	10 point, bold
3rd-level heading	**Headings.** Text follows ...	10 point, bold
4th-level heading	*Remark.* Text follows ...	10 point, italic

(Fig. 3 shows an example).

Fig. 3: wuzup

$$\psi(u) = \int_0^T \left[\frac{1}{2} \left(\Lambda_0^{-1} u, u \right) + N^*(-u) \right] dt \tag{1}$$
$$= 0? \tag{2}$$

Please punctuate a displayed equation in the same way as ordinary text but with a small space before the end punctuation.