

Enhancing Classification Performance through Example Generation

Anna Baumeister (w270)
anna.baumeister@tum.de

Nicholas Gao (w348)
nicholas.gao@tum.de

Matthias Mayer (w213)
matthias.mayer@tum.de

Frédéric Rackwitz (w515)
rackwitz@in.tum.de

Yannick Spreen (w231)
y.spreen@tum.de

Project Proposal

Evidence shows, that GANs can provide a viable method to enhance classification performance and accuracy by giving the NN more examples to train on. We find this approach very promising and want to explore further into this direction. The field for this enhancement is broad, and we want to try some new approaches on the CIFAR-10 and 100 datasets.

1. Introduction

While there were numerous generative networks developing in the past, recent results promise new ways of generating examples which improve training even more. Our goal is to fundamentally understand generative models. With this knowledge we will be able to assess the performance of semi-supervised learning for CIFAR-10/100. These datasets have not been explored as deeply yet in recent research, so we hope to gain additional insight into the topic.

1.1. Related Works

- Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (DCGAN) [4]
- An introduction into GANs: Generative adversarial networks [1]
- Improved techniques for training GANs by OpenAI [5]
- Together with the blog-post-writeup [3]

2. Dataset

- CIFAR-10 [2], a labeled dataset with natural images. It was so far used less frequently in training GANs but could prove quite potent due to small image sizes (short training time) and available labels which can help the discriminator learn.

- CIFAR-100 [2] is an advanced test for our newly developed skills. It provides a lot of classes which in turn require a lot of different models for generation. It also contains a lot fewer examples per class than other sets, so generation of new datapoints is an even bigger incentive for CIFAR-100 than for CIFAR-10 for example.

3. Methodology

Since we do not have much experience in the field of generative neural networks, if any at all, we need to familiarize ourself with the topic first. We will look into DCGANs to generate new data for CIFAR-10 training. Afterwards we will attempt to use our new knowledge to enhance classification performance on the bigger CIFAR-100 dataset.

1. Apply DCGAN to CIFAR-10 to generate new images from the 10 classes
2. Apply extensions from [5] to our GAN, such as feature matching and minibatch discrimination
3. Recover classifier from discriminator part of the different CIFAR-10 GANs
4. Investigate amount of labeled data needed and utility of proposed extensions to train a classifier for CIFAR-10
5. Scale to CIFAR-100

4. Outcome

Our final goal would be to have trained a network which has good performance in classifying CIFAR-100 images, possibly without giving it the full labeled dataset. In theory it should perform better than usual training without GAN generation of additional examples on an equal amount of labeled data.

In practice, we cannot assume that we will surely reach this goal in time, since it's been shown to be difficult. For this reason, we will look into CIFAR-10 first.

References

- [1] I. Goodfellow. Nips 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*, 2016. 1
- [2] A. Krizhevsky, V. Nair, and G. Hinton. The cifar-10 and -100 datasets. online: <http://www.cs.toronto.edu/kriz/cifar.html>, 2014. 1
- [3] OpenAI. Generative Models. online: <https://blog.openai.com/generative-models/>, 2016. [accessed 19-Dec-2017]. 1
- [4] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *CoRR*, abs/1511.06434, 2015. 1
- [5] T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. Improved techniques for training gans. *CoRR*, abs/1606.03498, 2016. 1