MATH 6380O FInal Project

Generating Images via Generative Models

Weizheng ZHU, Hongming ZHANG, Min FAN, Jianhui ZHANG

Introduction & Objectives

This is the final project of MATH 6380O. The final objective of this project is generating images via generation models.

In this project, we trained one generative models to generate images. More specifically, we reproduced the current state-of-art generative model introduced by (Rie Johnson and Tong Zhang, 2018) and achieved comparable performance.

Generative adversarial network (GAN) is a powerful class of generative models that case generative modeling as a game between two networks. It has become popular for generating data that mimin observations by learning a suitable variable transformation from a random variable. To tackle the problem of instability, Rie Johnson and Tong Zhang(2018) proposes a different theoretical foundation for generative adversarial methods with does not rely on the minimax formulation.

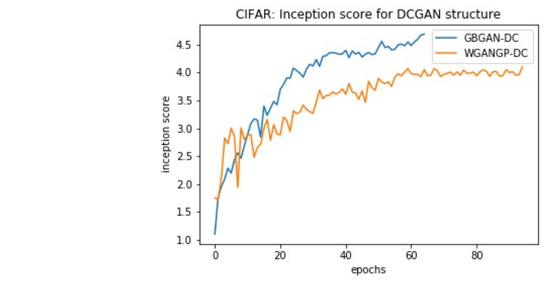
We reproduce this model as the final project of MATH 6380O. And the Empirical results demonstrate the effectiveness of this new method.

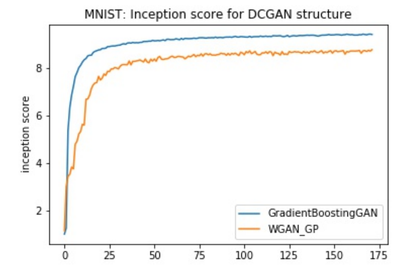
Hyper-parameters

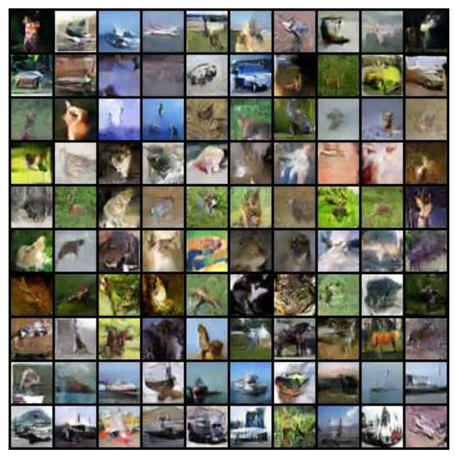
The hyper-parameter is set as used in the code. But decay of learning rate in approximation procedure is used when inception score stop increases.

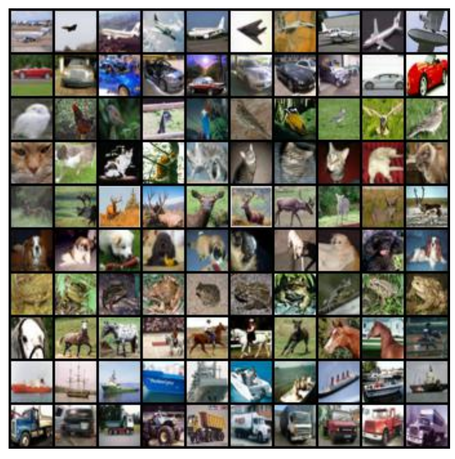
We also try decay of learning rate of discriminator and update rate \eta. It should be really careful when balance these two parameters.

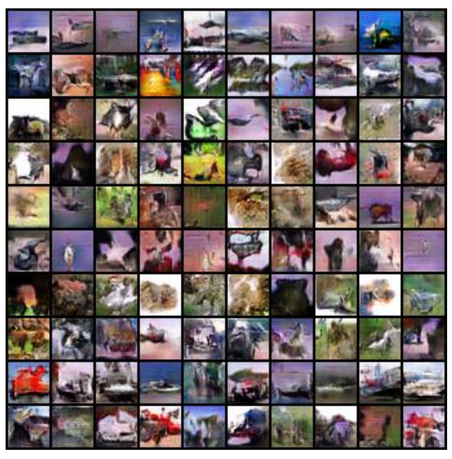
Empirical Results

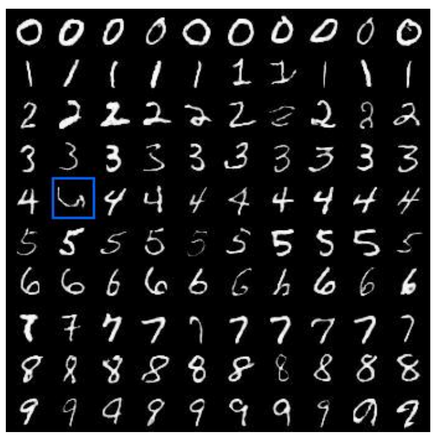












Gradient Boosting GAN

Real Image

WGAN-GP GAN

Dataset

In our experiments, we use two datasets as the training data as follows:

(1) MNIST: The MNIST database of handwritten digits, available from this page, has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger set available from NIST. The digits have been size-normalized and centered in a fixed-size image. Models are typically evaluated based on how well they can distinguish the integer from the images. We downloaded the MNIST dataset from the following webpage: http://yann.lecun.com/exdb/mnist/

(2) Cifar-10: Cifar-10 was firstly introduced by Krizhevsky et al. (2009). It consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images. Models are typically evaluated based on how well they distinguish the images into the correct categories. We downloaded the Cifar-10 dataset from the following webpage: https://www.cs.toronto.edu/~kriz/cifar.html

Training Tips

DCGAN: using DeConvolution GAN structure (shallow than resnet), GradientBoosting GAN will outperform WGAN-GP and achieve higher inception score and (non-collapsed) images in both MNIST and CIFAR10.

ResNet: But when it turns to ResNet structure, I haven't make it as successful as DCGAN. It's unstable in training process comparing with same structure network for WGAN-GP, which has stable and progressive growing in Inception Score. Especially in some cases, the inception score will suddenly dropped and the whole network will soon break down.

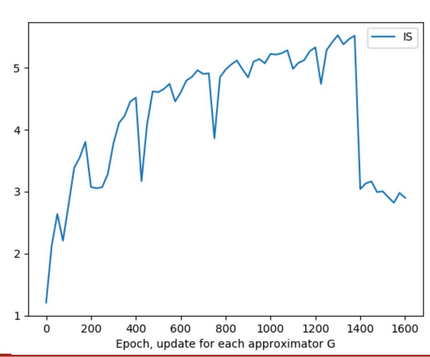
Training Tips for My Experience:

- No batch normalization in discriminator (as opposed to structure listed in appendix). A reason is the gradient $\frac{\partial D}{\partial G}|G = G(z)$ shouldn't be affected by other images in same batch.
- Balance discriminator learning rate and η is like to balance discriminator and Generator.
- Learning rate decay is useful: Decay learning rate of disc and η. Decay of learning rate for approximating in compressing.
- Kaiming Initialization seems work better.
- Using built-in function for each gradient computation.

Results and Disscussion

Results shown above is based on DC-GAN framework. For MNIST dataset, weak GANs will tend to collapse to 'bold' front and for some particular number, they will collapse to one (or few) mode(s) . However, both WGAN-GP and GB-GAN performs well in our experiment. For CIFAR-10 dataset, DCGAN is not enough to produce clear images due to its limitation in complexity. With same structure GBGAN could generate visually better images, especially in horse class(8th line), ship class(9th line) and truck class (10th line). On the contrary, images generated from WGAN-GP could not be identified easily and collapse in terms of color in our experiment.

When it turns to ResNet structure, where WGAN-GP could produce benchmark results. GB-GAN suffers from unstability in the training process. In particular, failure mode below, that is the network breaks down suddenly, frequently happened and we could not train a satisfying network so far. If we monitor the loss of discriminator and generator, we find that when this failure happens, the discriminator dominates suddenly but we haven't figure out the reason.




```
3 [1350], Inception Score:5.4632, D loss:1.3594, G loss:0.7399, G approx:0.0014 ,Ave Update:0.0018, G norm:0.6644, D norm:2.4471
9 [1375], Inception Score:5.5165, D loss:1.3683, G loss:0.7343, G approx:0.0013 ,Ave Update:0.0019, G norm:0.6523, D norm:2.4519
9 [1400], Inception Score:3.0413, D loss:0.0172, G loss:8.4761, G approx:0.0929 ,Ave Update:0.0096, G norm:1.6531, D norm:2.4769
1 [1425], Inception Score:3.1341, D loss:0.0142, G loss:10.7213, G approx:0.1029 ,Ave Update:0.0120, G norm:1.8654, D norm:2.4990
2 [1450], Inception Score:3.1648, D loss:0.0271, G loss:12.2959, G approx:0.1018 ,Ave Update:0.0114, G norm:2.0922, D norm:2.5054
```

References

1. Johnson, R. and Zhang, T., 2018. Composite Functional Gradient Learning of Generative Adversarial Models. *arXiv preprint arXiv:1801.06309*.
2. Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pp.1189-1232.
3. Arjovsky, M., Chintala, S. and Bottou, L., 2017. Wasserstein gan. *arXiv preprint arXiv:1701.07875*.
4. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. and Courville, A.C., 2017. Improved training of wasserstein gans. In *Advances in Neural Information Processing Systems* (pp. 5769-5779).
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).