Project Proposal - ECE 285

Haonan Peng
Department of ECE
A14765890

LinXiao Zhang
Department of ECE
A59013665

Abstract

DCGAN is one of the popular and successful network design for GAN. It mainly composes of convolution layers without max pooling or fully connected layers. It uses convolutional stride and transposed convolution for the downsampling and the upsampling. We want to reimplement the DCGAN model and train it on a new dataset(large-scale CelebFaces Attributes Dataset). The main objective of the model is to get a Generator Network to generate new images of fake human faces that look as realistic as possible.

1 Problem Definition

DCGAN is the standard convolutional baseline that many GAN architectures were based upon, it can generate high quality images by using strided convolutional layers in the discriminator to downsample the images and fractionally-strided convolutional layers to unsample the images. The importance of DCGAN is that it contributes significantly to balancing GAN training with its convolutional architecture, GAN and naturally DCGAN have an unsupervised network structure. In order to better understand how DCGAN works we plan to reimplement the DCGAN model and train it on a new dataset to generate new images of fake human faces.

2 Tentative Method

Although GANs bridge the gap between supervised learning and unsupervised learning, GAN are also known to be unstable in the process of training which usually results in producing absurd outputs from the generator. To mitigate that, DCGANs propose a more robust network architecture to train with relative-faster convergence speed in most settings.

DCGANs contain two neural networks: discriminator and generator. The discriminator is trained to classify images generated by the generator from the natural images, while the goal of the generator is to produce novel images to increase the error rate of classification done by the discriminator. In our method, the generator is a five layers architecture containing convolutional-transpose layers, batch norm layers, ReLU activations, and Tanh activation. The input is a latent vector, z, that is drawn from a standard normal distribution and the output is a 3x64x64 RGB image. The discriminator also contains five layers that are made up of strided convolution layers, batch norm layers, and LeakyReLU activations. At last, rather than using momentum to accelerate training, our method chose the Adam optimizer for both networks.

3 Experiments

CelebFaces Attributes Dataset (CelebA) is a large-scale face attributes dataset with more than 200K celebrity images, each with 40 attribute annotations. The images in this dataset cover large pose variations and background clutter. CelebA has large diversities, large quantities, and rich annotations, including 10,177 number of identities, 202,599 number of face images and 5 landmark locations,

40 binary attributes annotations per image. This dataset can be employed as the training and test sets for the face generation task. We plan to reimplement the DCGAN model and train it to generate new images of fake human faces. Besides, we want to try different loss functions and network architectures to study their effect on the performance of DCGAN model.

References

Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.