Face Aging using DCGANs

Moiz Sajid MOIZSAJID619@GMAIL.COM
National University of Computer and Emerging Sciences

Abstract

In the past, performing face aging on the user's input image has been really difficult. Previous methods have failed to preserve the spatial structure of the image. In this report, I propose a new method for performing face aging using Deep Convolution Generative Adversarial Networks (DCGANs). The method has not yet been fully development but initial results show very promising results. Further research and experiments needs to made inorder to make this method more mature.

1. Introduction

Face aging methods basically consists of two methods: parametric and non-parametric. The images produced by parametric methods look very unrealistic while the parametric methods are generally not very fast. There are many uses cases of the face aging. Face aging itself is very useful particularly for the entertainment industry. Face aging can also be used to find people who might have got lost at young age.

GANs were first introduced by Goodfellow et al. Since then they have surprised the world because of their ability to produce visually pleasing images. Unlike other generative models like Boltzmann Machine and Variational Autoencoders (VAEs), GANs can produce high quality images without blurriness. The generator and the discriminator network works in an mini-max way where the generator tries to produce new images similar to the original data distribution while the discriminator tries to distinguish whether the image produced by the generator is fake or real.

2. Methodology

The main approach that I used in this report was proposed by Radford et al in their paper about DCGANs. Their paper was one of first attempts in using Convolutional Neural Networks inside GANs. As a result, the images produced by this GAN were of very high visual quality. For generator network, fractionally strided convolution was used. For the discriminator network, strided convolution was used which was mainly due to the fact the authors wanted to avoid using max pooling for reducing the paramters. The generator network only used one fully connected layer at the start where it takes input from the noise distribution Z. The discriminator network used fully connected only at the last layer where its input is flattened.

Batch Normalization was used in both the generator and the discriminator since the authors of the DCGANs paper reported better generated images. Their paper was the first one to use Batch Normalization, since Batch Normalization did not existed when Goodfellow et al first proposed the GANs. Batch Normalization helps to promote gradient flow in a network and also somewhat solves the problem of bad weight initialization. The generator network uses RELU activation function except for the final output layer which uses Tanh activation function. The discriminator network used Leaky RELU activation function for all its layers. For regularizing both the generator and discriminator network, Dropout was also used.

The main approach that I used for performing face aging is through Vector Arithmetic of the latent representation Z. Mikolov et al first demonstrated the learned representation for their Word2Vec model using simple arithmetic operations. I applied this vector arithmetic on our problem as well. However, the results produced from it were not exactly what I expected it to output. I averaged the three representation images for better stability before performing vector arithmetic.

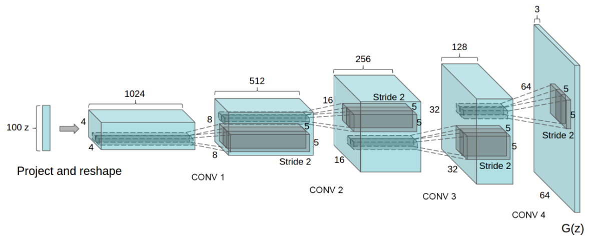

Figure 1. Encoder Architecture Diagram

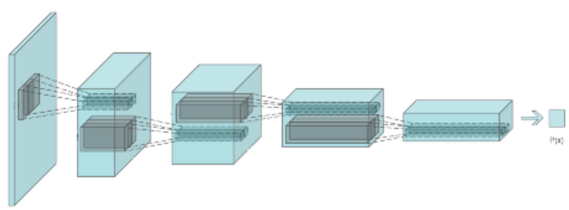

Figure 2. Decoder Architecture Diagram

3. Experiments

For training my generative models, I used the IMDB-Wiki and CelebA datasets. Unlike the IMDB-Wiki dataset, CelebA dataset provides much more attributes of the images hence is more useful for generating images using conditioning. Because of its richness, CelebA is now the defacto dataset for reporting the performance of different GANs.

For the IMDB-Wiki dataset, I only used the Wiki images which contains only 62,328 images while the CelebA dataset contains more than 200,000 images. No preprocessing on the images was performed like mean normalization. For the IMDB-Wiki, the images were of unequal length. Therefore, I cropped out the face from images and rescaled the images to 64 by 64. Furthermore, no data augmentation whatsoever was applied to the images.

4. Conclusion and Future Work

In this report, we looked at a method that leverages DCGANs for face aging. However, my method fails short of actually doing face aginging; however, further improvements can be made which can improve its results. These approaches have been briefly summarized below.

Firstly, we can add an encoder into our architecture which will basically encode the image to latent representation Z. One of the main advantages of this will

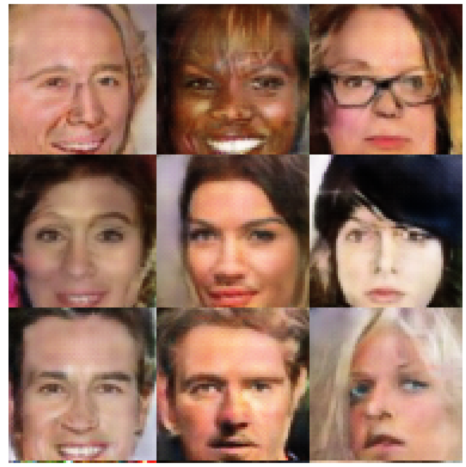

Figure 3. Image Generation during 25th epoch of training

Figure 4. Vector Arithmetic using Latent Representation Z

be that the user will be able to provide his/her own image for face aging. Secondly, Wasserstein GANs could also have been used because they allow for better training of GANs. Recent paper about Wasserstein GANs make the training of GANs much more stable and smooth. Finally, Conditional GANs could have been used because they would have allowed us to take the full advantage of our diverse dataset with many attributes.

5. References

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in Proceedings of advances in Neural Information Processing Systems, Montreal, Canada, 2015

Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in Proceedings of International Conference on Learning Representations, San Juan, Puerto Rico, 2016.

Guim Perarnau, Joost van de Weijer, Bogdan Raducanu,and Jose M lvarez, "Invertible conditional gans for image editing," in Proceedings of advances in Neural Information Processing Systems Workshops, Barcelona, Spain, 2016.