Image Generation from the Facial Composite using Variational Auto-Encoders, Generative Adversarial Networks & Bayesian Optimization

Ayush Thada (16BCE1333) Sanket Prashant Deshpande (16BCE1084)
School of Computer Science and Engineering, VIT University, Chennai Campus

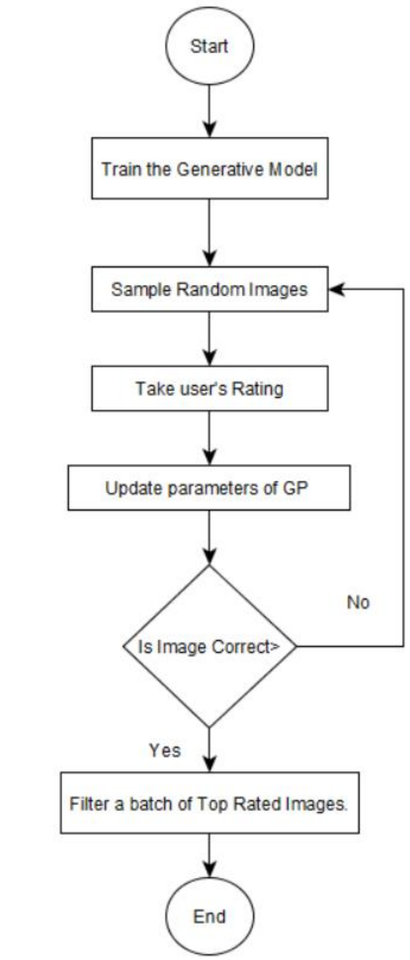
Abstract

In many criminal cases to identify the image of the personal law agencies used facial composites. A facial composite is a graphical representation of one or more eyewitnesses' memory of a face, as recorded by a composite artist. But the image which is drawn totally depends on how effectively the suspect was observed by a person and usually ordinary people don't observe things with such great details that they can asset an artist to draw a photorealistic image. So in this process, we are proposing a system in which user can directly provide rating to the random images shown to them and on the basis of that subsequent images are generated. This leverages the effectiveness of the photographic memory as well as the ability to differentiate between the images of people and animals that have been evolved over a period of thousands of years.

Introduction

The core problem that we have to solve to realize the system is image generation as well as the ability to find the generalize the pattern over a few images. TO tackle the image generation issue we are using the two most popular methods which are most widely used in the medical field, statistics, and other related domains. The first one is Variational Auto-Encoder which maps the dataset of the image to a normal distribution. The second methods are Generative Adversarial Networks. It is an image generation state of the art method in which two networks image generators and discriminator play a min-max game which objective to each other over the course of training. After a training generator is capable of generating sharp photorealistic images from random noises. Our model is nothing but a sequential improvement in image generations over time with the action from the user. It can be viewed as a stochastic process and in fact, we are treating this as a Gaussian Process. The Bayesian Optimization allows us to move forward in this time series on the action of the user. At time t, computer samples some images randomly from distribution and user rating on images act as the data and previous state is prior knowledge which helps users to yield the posterior distribution which is a state at time t+1.

Methodology



Datasets:

- For Variational Auto-Encoders, Lost Faces in the Wild **(LFW)** dataset is used.
- For Generative Adversarial Network, Flickr-Faces-HQ Dataset **(FFHQ)** dataset is used.

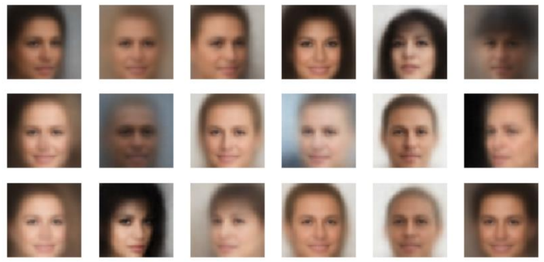
Bayesian Optimization:

- GPyOpt library is used.
- Gaussian Processes is used.

Results

For Variational Auto-Encoder Model:

1. Generated Images:



2. Query:

- Widest Smile

- Darkest Hair:

For Style-GAN model:

1. Generated Image:

2. Query:

- Smiling Teenager Girl:

Discussion & Conclusion

The model was able to extract the image according to our query only from the user rating. During the testing of the model, the state of the Gaussian Process at <t=0> is decided by 10 images only. And at each subsequent time interval, only 1 image was generated and asked for users' ratings. The reason we decide to go with only 1 image to make sure even I the cases od very such large uncertainty how to do model performs and as per results we observed results are decent. The query was kept simple because we were working under very high uncertainty. In real cases, we can increase the no of images required to get the posterior at time <t> using the distribution at <t-1. as prior.

In the case of Variational Autoencoders, the mean and standard deviation encoding size was 8. While in the case of the Style-GAN model we've used an encoding size of length 512 . This is also the reason the sharpness in images generated in both cases is different from each other.

As future work, we'd like to work on the issue to address the problem of local minima in great detail. For now, in case of local optima appears we assign a very low rating to the image and it throws us out from the local optima and to automate this model further we'll be using Reinforcement learning.

Acknowledgements

[1] Auto-Encoding Variational Bayes, Diederik P Kingma, Max Welling. Availabel [Online] https://arxiv.org/pdf/1312.6114.pdf

[2] A Style-Based Generator Architecture for Generative Adversarial Networks, Tero Karras, Samuli Laine, Timo Aila. Availbael [Online] https://arxiv.org/pdf/1812.04948.pdf

[3] Bayesian Optimization using the Gaussain Processes. Availabel [Online] https://github.com/SheffieldML/GPyOpt

[4] Pretrained Mdoel. Availabel [Online] https://drive.google.com/drive/folders/1MASQyN5m0voPcx7-9K0r5gObhvvPups7