A

Summer Internship Project Report on

Learning Face Age Progression: A Pyramid Architecture of GANs

Submitted by:

Shivani Kushwaha

Shaheed Bhagat Singh State Technical
Campus, Ferozepur-152004, Punjab, (INDIA)

Under The Guidance of

Dr. Satish Kumar Singh

Associate Professor

Department of Information Technology

Indian Institute of Information Technology Allahabad

Devghat, Jhalwa, Allahabad-211015, (U.P. ,INDIA)

UNDERTAKING

I declare that my 2 months Summer Internship report titled "**Learning Face Age Progression: A Pyramid structure of GANs**", submitted to Dr. Satish Kumar Singh, Associate Professor, Department of Information Technology , IIIT- Allahabad, as a part of my internship is original. I have not plagiarised or submitted the report anywhere else.

Date: 16 August 2019

Place: Allahabad

Shivani Kushwaha

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and deepest thanks to my supervisor **Dr. Satish Kumar Singh**, **Associate Professor** at IT-Department, Indian Institute of Information Technology, Allahabad for his continuous support, supervision, motivation and guidance throughout the tenure of my internship in spite of his hectic schedule, he truly remained driving spirit in my project and his experience gave me the light in handling this project and helped me in clarifying the abstruse concepts.

Lastly, I would like to thank Nand sir, God almighty and my parents for always supporting me.

Shivani Kushwaha

Dr. Satish Kumar Singh
Associate Professor,
IT-Department, IIIT- Allahabad

ABSTRACT

Since the face age progression deals with addressing the two main aspects, i.e. aging accuracy and identity permanence, a Generative Adversarial Network is employed to model restraint for the inherent subject specific characteristics as well as the age- specific changes in faces with time.

For generating more realistic facial details pyramidal adversarial discriminator is utilised which models the effects of aging in a more advanced manner.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CACD	Cross-Age Celebrity Dataset
CNN	Convolution Neural Network
NN	Neural Network
MLPs	Multi Layer Perceptron
GANs	Generative Adversarial Networks
VAEs	Variational Autoencoders

INTRODUCTION

Age progression is the method of altering the facial images to demonstrate the consequences of aging on the face of an individual.

RELATED WORK

While solving the problem of face age progression innate complexity of physical aging and interferences occurring due to the other factors makes it harder. There are many attempts that have been employed for tackling the above problem:

- A compositional and dynamic model was employed where compositional model represent the faces in distinct age groups by hierarchy of And-Or-Graph(And nodes fragments the face into finer details i.e. hair, wrinkles, etc; Or nodes depicts the variety of face by substitutes). It is modelled as Markov process or the parse graph on the basis of two important aspects of face age progression i.e. accuracy and identity preservation [4].
- Analyzing the skin's anatomical structure, a 3 layered dynamic skin model to counterfeit wrinkles[5],
- Conditional Adversarial Autoencoder (CAAE) is also employed but 48.38% age progressed faces can be verified in human based evaluation on 856 synthetic pairs ; subjective preference votes (52.77%) are higher than that of prior work (28.99%) on 235 aged face pairs.
- Other methods are also used but could not meet the identity preservation and aging accuracy were not properly fulfilled.

There are many other methods that were employed for face age progression, but in our work we will be implementing face age progression with the help of GANs.

A typical GAN comprises of a generator G and a discriminator D which is trained with the help of adversarial nets in iterations. The GANs work on implicit density distribution. The generator method G tries to creates a probability distribution from the noise samples fed into it and confuses discriminator method D, where the method D aims to distinguish the ground truth (real face images) with the fake aged images generated by the generator G, using implicit probability density. In other words, GANs works on the principle of Mini- Max two player game algorithm of AI. Both the neural networks, G and D are based upon the MLP (Multi- Layer Perceptron).

The objective function of Mini-Max player game :

$$min_G max_D \mathrm{E}_{x\ P_{data}} \log[D(x)] + \mathrm{E}_{x\ P_z} \log[1 - D(G(x))] \quad (1)$$

where z is the noise sample from prior probability distribution P_z and x denotes ground truth (real face images) based upon a specific probability distribution P_{data}. When both the generator G and discriminator D converges the probability distribution of P_G (synthesized images) becomes equivalent to the $P_{data.}$

The main reason behind opting for GAN in age progression is as follows:

- It facilitates age estimation and face verification while keeping the issues of aging effect generation as well as identitypreservation,
- In the previous study forehead and hair components of face is not kept under considerations.
- Other neural network models face the problem of intractability while training; the intractable density function with latent z of Variational Autoencoders is as follows:

$$p\Theta(x) = \int p\Theta(z) p\Theta(x|z) dz$$

so ,they were incapable of computing the p(x|z) for every z .

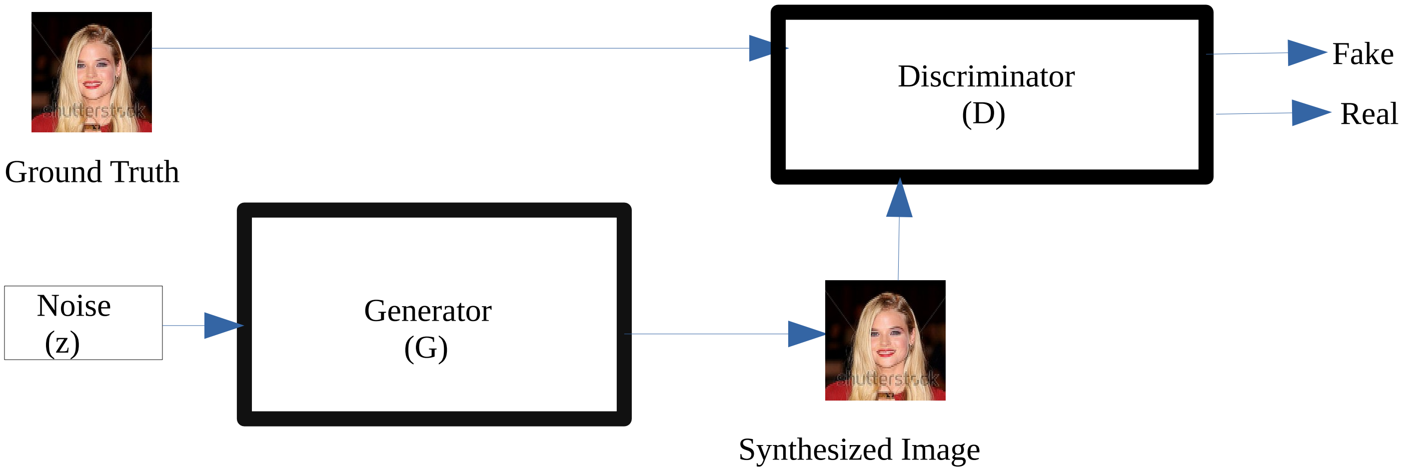


Fig 1.1. A classic GAN structure

GOALS/OBJECTIVES

Research in the field of face age progression can be employed in finding out the subjects that went missing, in our case we will be analyzing on images ranging from 0-50 years of age, which can further be used by the cops to search for the missing subjects even after years of he/she went missing.

PLATFORMS USED

Hardware Requirements	Software Requirements
RAM - 8 GB	Operating system dependency- Ubuntu 16.04.6 LTS (Xenial Xerus)
Processor used- Intel core i7 6^{th} gen	Programming Language- Python
	Version of Language used- Python3.7
	Anaconda (Integrated Development Environment) IDE
	Libraries of Python- Keras, Pytorch, Tensorflow, sklearn, scikit, numpy, pandas, matplotlib etc.

METHODOLOGY

Dataset

In our work the images are taken from CACD[6] and UTKFace dataset[1]. The CACD dataset contains 1,63446 images of 2,000 celebrities, and is available in many versions out of which Original face images (detected and croped by openCV face detector) is taken, whereas, in UTKFace dataset is large-scale face dataset having wide age span (from 0 to 116 years old). It consists of over 20,000 face images with annotations of age, gender, and ethnicity. It is available in 3 versions out of which Aligned and Cropped Faces are taken into consideration.

Generator

We applied a CNN based generator that takes younger face image along with the target age label as inputs and generates an elder face image conditioned on them. The generative network comprises of both encoder and decoder.

When the younger face image is given as an input to the generator it first exploits various convolution layers followed by a non-linear activation function ReLU along with Instance Normalization to encode it to a latent space.

The generator G architecture is described below in Table.1.1.[2]

Layer	Kernel	Stride	Padding	Activation Size
Conv	9 x 9	1	4	32 x w x h
Conv↓	3 x 3	2	1	64 x w/2 x h/2
Conv↓	3 x 3	2	1	128 x w/4 x h/4
Resnet	3 x 3	1	2	128 x w/4 x h/4
Resnet	3 x 3	1	2	128 x w/4 x h/4
Resnet	3 x 3	1	2	128 x w/4 x h/4
Resnet	3 x 3	1	2	128 x w/4 x h/4
De-Conv↑	3 x 3	2	1	64 x w/2 x h/2
De-Conv↑	3 x 3	2	1	32 x w/2 x h/2
De-Conv	9 x 9	1	4	3 x w x h

Table.1.1. : Generator Architecture

Discriminator

Discriminator also consists of strided convolution layer along with padding, to distinguish the real face images and fake face images probability distribution. The discriminator consists of various pathways for training the overall GAN, which is based on the VGG16 architecture for all the four paths. In discriminator the pyramidal architecture is employed, actual faces and the output generated by the Generator is fed to the discriminator for obtaining more intense age-specific facial details, throughout training. Meanwhile, Batch Normalization and LeakyReLU is utilized at each convolutional layer, except the last one in each pathway.

The architecture of discriminator D is depicted below in Table.1.2. [2]

Pathway	1	2	3	4
Input	512	256	128	64
Layers+	Conv-512 Conv-512 Conv-1	Conv-512 Conv-512 Conv-512 Conv-1	Conv-256 Conv-512 Conv-512 Conv-512 Conv-1	Conv-128 Conv-256 Conv-512 Conv-512 Conv-512 Conv-1

+ Layers are denoted as : Conv- <output>; kernel =4, stride=2, padding =1

Table 1.2: Discriminator Architecture

Training GANs

After preprocessing of images from datasets .i.e. classifying the images in accordance with their ages and normalizing them the images are fed to the generator as a noise using a .pickle file and the discriminator D is given the ground truth images along with the images generated by the generator G, The discriminator D finds out the variation in probability density of both the ground truth with fake images generated by the generator.

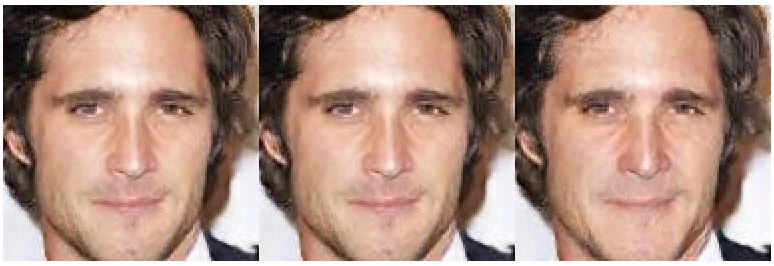
The training process decreases the loss of Generator to Discriminator along with loss of Generator to Generator, to generate more photorealistic image in a fine-grained manner.

For identity preservation we will be using the work done in **Towards Open-Set Identity Preserving Face Synthesis [3]**, as it incorporates asymmetric loss function for training the input face image to generate an identity vector along with any other input face image to withdraw an attribute vector containing pose, emotions and illumination,etc.

RESULTS

Age 15

Age 30



Age 25

Age 30

Age 40

CONCLUSION

From the results it can be concluded that the model works better for ages above 18 - 50 years. In this study, presents an effective solution for identity preservation and aging accuracy

FUTURE WORK

In the future we will be focusing on generating age progressed images of infants in account.

REFERENCES

[1] *UTKFace dataset.*[*http://aicip.eecs.utk.edu/wiki/UTKFace*](http://aicip.eecs.utk.edu/wiki/UTKFace).

[2] Hongyu Yang, Di Huang, Yunhong Wang, Anil K. Jain. 2019 Learning Continuous Face Age Progression: A Pyramid of GANs:32-38

[3] Jianmin Bao, Dong Chen, Fang Wen, Houqiang Li, Gang Hua. 2018 Towards Open- Set Identity Preserving FaceSynthesis.

[4] J. Suo, S. C. Zhu, S. Shan, and X. Chen. A compositional and dynamic model for face aging. IEEE PAMI, 32(3):385–401, Mar. 2010. 1,5

[5] Y. Wu, N.M. Thalmann, and D. Thalmann. A plastic-visco-elastic model for wrinkles in facial animation and skin aging. In PG, pages 201,214, 1994. 1,2

[6] *CACD dataset.*https://bcsiriuschen.github.io/CARC/ .