Temporal face feature progression with cycle GAN

Roy Ashish¹, B G Prasad²

¹ Department of Computer Science & Engineering, B.M.S. College of Engineering, Bengaluru, India
² Department of Computer Science & Engineering, B.M.S. College of Engineering, Bengaluru, India
Email: ashishroytech@gmail.com
bgprasad.cse@bmsce.ac.in

Abstract. The aging process creates significant changes in the appearances of people's faces. When compared to other causes of variation in face imaging, aging-related variation has specific distinct properties. Facial Aging variations, for example, is unique for each person; it occurs gradually and is significantly influenced by other characteristics including health, gender, and life-style. As a result, the proposed effort will use Generative Adversarial Networks to address these critical concerns (GANs). Generative Adversarial Networks (GAN's) is made up of a generator and a discriminator network. The generator model generates images that a discriminator model analyses to determine if they are real or fake. This paper provides a Temporal Face Feature Progressive framework with Cycle GAN, which maintains the initial appearance and identity in the elderly aspect of their facial structure. To address aging concerns, our goal is to transform an initial age category image into a targeted age with age progression. We show that our temporal face features progressive cycle GAN learns and transfers facial traits from the source group to the targeted group by training various images. The IMDB-WIKI Face dataset has been used to obtain the results for the same.

Keywords—Generative Adversarial Networks, GAN’s, Face aging, Cycle GAN.

1. Introduction

Deep learning has advanced significantly in a plethora of disciplines in recent years, including traffic, medicine, and the arts. In some manner, they're all associated with digital image processing or artificial intelligence. In this paper, we look into facial aging, which is a unique and essential topic. The human face is the most complicated structure and the characteristic that most clearly defines a person. Face aging is the process of altering the age of a person’s facial structure to match that of another. The characteristics of the final generated image should remain the same, and the age of the produced image should match that of the reference image for a correct stylization.

Additionally, the final facial image should be similar in age to the pictures which were gathered initially. Age transition methods are utilized in a number of ways, including image processing, editing, and content development. Due to the lack of expressive feature representations, traditional stylization methods are mainly based on correlating color statistics and are often restricted to one-to-one transformation.

In the last year, various face aging models have been proposed, all of which are witnesses to multiple advancements in the human face aging model. Nevertheless, the most complicated problem in practice is still facial aging, attracting increasing academic attention due to various applications in recognition, cross-age verification, and amusement. It could be utilized to locate a missing individual or anticipate someone's future appearance. Over the last year, researchers have used GAN's to successfully produce high-resolution images [1] [2].
Consistent GAN (Generative Adversarial Networks) is GAN's classes that receive preliminary information from a source image and generate an output with a specific desired characteristic. Following Cycle GAN [3], it imposes a limitation on the actual image to ensure that our target image is not out of shape.

We developed the Temporal Face Feature Progression for facial aging based on the Cycle GAN [11]. Our work includes image transformation, in which an image is assisted in transforming from a sample age group to an intended age group while maintaining the person's original identity.

The following is a summary of our work in this paper:

1. We use Cycle GAN to simulate Temporal Face Feature Progression, a GAN that produces a first-rate image of the chosen age group as they get older.
2. We developed a progressive approach that kept the input image's original facial identification throughout the procedure.

The paper consists of a total of 6 sections. Section 2 covers the work which is relevant in the field of facial aging. In Section 3, we present our temporal face feature progression with the cycle GAN methodology. In Section 4, we outline our data gathering, taking into account the critical element of gender. Section 5 provides the outcomes of our efforts, and we examine them from several perspectives. Section 6 summarizes our findings and addresses the potential of face aging, which can be done in the future.

2. Related Work

Creating a photorealistic facial aging image has long been a challenge [5]. Recently, a start-up company released an app that can slow down the aging process; however, it is mainly used for entertainment. An essential thing that facial aging can accomplish is to predict how young children will seem when they grow older and provide current photographs to identify missing people.

A. Face Aging

There are two types of modern facial aging concepts: prototype and physical methods. The prototype method uses the average number of faces in a category within a specific demographic group to establish an aging pattern, which is then used to build a picture for the chosen age group. Physical methods for face transformation in the country rely on physical structures, facial muscles, and other facial shapes.

The main issue with these types of age progression approaches is that they ignore individualized information in face photos, resulting in unrealistic face images. To avoid this, a few representation-based techniques that take person-specific facial information into account to simulate an aged face have been developed [6] [7].
These strategies can help to preserve a person's identity to some extent, but the image will suffer from
ghost artifacts as a result of the reconstruction techniques used. Significant efforts to address this issue
have been made in recent years, with correctness in aging and the preservation of one's identity being
widely recognized as the two fundamental foundations of its effectiveness.

[8] presented a recurrent facial aging algorithm on the basis of RNN (Recurrent Neural Networks),
which results in smoother faces among neighboring aged groups, although this study lacks
identification information. [9] discusses a conditional GAN-based auto-encoder that encrypts the input
image before generating an aged image as output.

Fig. 1. Shows the representation of DX and DY discriminators, which are used to distinguish between
authentic and fake images. Moreover, it shows the Translation of domain X as the source to the intended
domain Y using the binding function G and the reverse mapping function, which performs mapping on
domain Y to map domain X, where G represents the mapping of X to Y (X→Y) and F represent the
mapping of Y to X (Y→X).

B. Face Feature Progression

Face progression is a type of image regeneration issue wherein the source facial picture is transformed
into the desired image. At the same time, the original characteristics of the person are preserved [10].
Age-conditional GAN [11] was designed to retain the individual's features in age synthesis.

First and foremost, GAN's were used to track the age progression of a particular age group. GAN's core
idea is built on the process called the min-max game, which is used in order to distinguish between fake
and actual photos. In addition, two losses, Cycle consistency loss and adversarial loss, play an integral
role since adversarial loss correlates with the generated image sample distribution to the desired field. In
contrast, Cyclic consistency loss prevents the already learned G and F mappings from contradicting one
another.

\[ L_{GAN}(G, D_V, U, V) = E_{v \sim Q_{data(v)}} [\log D_V(v)] + E_{u \sim Q_{data(u)}} [\log (1 - D_V(G(u))] \]

\[ L_{cyc}(G, F) = E_{u \sim Q_{data(u)}} [F(G(u) - u] + E_{v \sim Q_{data(v)}} [G(F(v) - v)] \]
Cycle loss is defined as consistent cyclic loss, which attempts to recover the original image from newly produced samples. The cyclic consistency loss in the onward direction is represented by each and every picture $u$, which is from class $U$, where $u$ is mapped to $G(u)$ and $F$ is mapped to $(G(u)) \sim u$. Two generators $G$ & $F$, where $F(v) \& G(u)$ produces a fictitious image. $v \sim Q_{data}(v)$ and $u \sim Q_{data}(u)$ are domain $V$ and $U$ data distributions, with $u$ denoting distribution in the $U$ domain and $v$ denoting samples in the $V$ domain. (Equation 2)

Cyclic consistent losses force $G$ (Generator) to avoid needless variations and subsequently generates an image with a structural system identical to the source. Compared to all other models, GAN creates crisp and clean photos; yet, it lacks adequate image resolution. However, progressive GAN eventually overcame this issue and generated a very high-resolution result [12].

C. Generative Adversarial Networks (GAN’s)

Recently, it was discovered that GAN’s (Generative Adversarial Networks) produce significantly high-resolution pictures. There are two models in this system: a generator and a discriminator. To produce an image $x$, provide $z$ as the noise vector to the generative Model $G$ selected from the homogeneous distribution of $P_z(z)$ maps to $z$.

The discriminator's goal is to differentiate among authentic and false images using either $x$ or $\sim x$ (from the probability of data $P(x)$, images $x$ is selected). The discriminator has been trained to provide the proper label more often while continuously training the parameter $G$ to optimize the log function given by log $[1- D(G(z))]$.

Primary GAN's min-max game is displayed using the following function:

$$\min_{G} \max_{D} \mathbb{E}_{x \sim p_{data}} \log[D(x)] + \mathbb{E}_{z \sim p_{z}} \log[1 - D(G(z))]$$

(Equation 3)

GAN’s does not use a sophisticated sample distribution model, yet the data generation is so unguided that it occasionally produces undesirable results. This challenge is tackled using Conditional GAN (cGAN), which directs the data created using conditional input on both the discriminator and the generator [13]. It performs well in various translation methods, including text to picture [13] and picture to picture [14].

Increasing the loss functions while developing the mapping from the source images to the targeted image, the pixel to pixel (pix2pix) model [15] gives a direct solution to this issue. As a result, the model provides the key to avoid issues in machine learning, which previously required separate tools and made charts from colorized versions of black-and-white pictures and aerial photographs.

Furthermore, GANs first generate substrate plans (e.g., intensity charts) before mapping organic locations to pictures. Cycle GAN [16] expands on these attempts by including a powerful loss function, the cyclic loss, that can recover the original image after forwarding and inverse transformation.

D. Style Transfer

In face aging, the notion of generating a targeted age group picture from input base age group photographs is known as style transfer [17]. The purpose is to send the source picture (the picture to be converted into another facial feature) and the targeted picture (the themed picture) to the system to recreate the creative form of the targeted view on the composition of the source image.

Content and style loss occurs when images are transferred from one domain to another. This approach retrieves information from a Neural network that has already been trained, resulting in content and aesthetic loss, also known as perceptual loss.
The style transfer approach is primarily focused on learning the mapping functions that move images from one domain to another. As a result, it has various applications, including style transfer on pictures captured by satellite to Google Maps visuals of traffic, RGB to Grayscale conversions and vice-versa, pictures to paintings, etc.

![IMDB-Wiki Dataset](image)

**Figure 2.** IMDB-Wiki Dataset.

### 3. Methodology

#### A. Overview

Faces were divided across two gender-based groups, i.e., males and females for the age group of 20-25. Our goal is to create a learning algorithm to provide samples for the age groups X (20-25) and obtain output Y (25-30). The received image Y is again fed into the model to go through the same process for five iterations. Our approach uses two mapping parameters to map G and F, from X to Y (X → Y) and from Y to Z, i.e., Y → Z respectively, and so on goes in a cycle. Additionally, the proposed method uses two discriminant parameters D_X and D_Y, where discriminator D_X discriminates among pictures X and generates images F(y); similarly, discriminator D_Y distinguishes among image y and generates pictures G(z), and this process continues for five iterations in order to get age-progressive facial images. Figure 3 depicts the transition of an image from one domain to another. The dataset is separated into age groups with at least 50,000 training photos in each. We trained it over 400 iterations with a learning constant of 0.002.

#### B. Temporal Face Feature Progression with Cycle GAN

The proposed topology is based on [3] Johnson et al.'s work; it has shown promising results in enhanced resolution and style transfers. Figure 3 depicts the entire architecture of our Temporal Face Feature Progression with Cycle GAN (Generative Adversarial Networks), consisting of two variations:

1) The 1st model, which produces target face images of varying ages while considering the parameter x to make them look genuine.
2) The discriminator, which distinguishes between actual and fake photos.

Face aging follows several typical aging patterns from a physiological standpoint. The age progression of people's heads, for example, caused a shift in human appearance. However, as we can see in our everyday lives, even though they were naturalized in the same group, everyone's aging effect is distinct. Because of the unique facial features and aspects of the human faces, age progression is influenced by
various elements such as race, gender, heredity, health status, and chronological age. Face aging progresses in a variety of ways, each with its own set of challenges.

We discovered that the genders of the source face image influence the functioning of our model in our tests. Models developed on female character photos, for particular, perform much better on age progression than models trained on men's facial image, which bodes well for the future.

We decided to consider making a few changes to our model; to the best of our understanding, to realize face aging with Cycle GAN [19], we need to segregate the facial image dataset into multiple age groups. We split each age sample into two components, male and female, resulting in a dual number of age categories than usual age categories.

4. Experiment

This section explains how we prepare our data and how we put our proposed facial aging approach into practice. Collecting facial photos of one individual with a broad spectrum of biological age groups is difficult and expensive, as we indicated previously. We're attempting to collect facial photographs with gender and age information to use in our current study; also, the demographic breakdown of the facial images should be uniform across all ages. The empirical concepts and settings are based on our goal in this project, whose objective is to look into how gender affects face ageing models.

![Figure 3. Temporal Face Feature Progression Proposed Model.](image)

A. Dataset Overview

To determine the output of our model, we have used the clean IMDB-WIKI data samples. Due to the absence of age and gender details in the source facial image, we chose to use a single dataset that contains age and gender labels for every facial appearance, which is why the IMDB-WIKI dataset is used. Moreover, the sourced age labels can't be used directly because many facial images consist of two or more people in one embodiment, so two methods are proposed to address this issue [18]. Finally, [18] provides us with clean IMDB-Wiki data samples, a subclass of the larger 250 K facial images, IMDB-WIKI data samples. The age distribution in IMDB-WIKI is illustrated in Figure 2 and Figure 4 depicts the age distribution in cleaned IMDB-Wiki.
We split the dataset into two different categories of gender for the age group 20-25, with each class including at least 50000 photographs of the appropriate ages for training and 30% for testing.

B. Implementation of Proposed Methodology

Our approach primarily focuses on the current Generative Adversarial Network; therefore, we trained our model using the most up-to-date state-of-the-art technology. Our training parameter is identical to the Cycle GANs since our framework is similar to the Cycle GANs.

On the NVIDIA GTX Titan V GPU, we train our data sets over 500 iterations at a training rate of 0.002 in roughly 26 hours. It requires less than 0.30 seconds to test by establishing a log of test datasets in the form of a .html file and comparing the input and output images.

The Cycle GAN [3] methodology trains our Temporal face feature progression-based GAN model on every other pair of predetermined age category. The generator design of our model is based on [3], which is established to be an excellent method for high-performance NST (Neural Style Transfer) and super-resolution. Two fully connected layers of two strides and numerous residual blocks [20] are accompanied by two perceptibly stride convolutions with ½ stride in this configuration. We use nine blocks for 256 x256 face photos.

In addition, as proposed by Johnson et al., instance normalization is used [3]. To tackle issues of training that are not in stable stages and due to which the network collapses, we have used the WGAN-GP loss function instead of the initial Cycle GAN's most minor square loss function. In WGAN-GP, the Gradient Penalty model is implemented, which solves the problems of exploding and vanishing gradients. WGAN-GP can dramatically improve learning rate, convergence, & data quality. A Cycle GAN model is trained five times over each set of facial pictures for the training aspect to get the facial images in a progression or a multiple of 5.

5. Evaluation And Results

Our temporal face feature progression is a cyclic model which converts pictures from one age category to another and repeats N number of times for the same image, where N equals 5. Figure 5 and 6, depicts the translation from source domain A (Female Category) to intended domain B (Female Aged Progressed Category) and the translation of Source domain X (Male Category) to the intended domain.

![Cleaned IMDB-Wiki Dataset](image_url)

**Figure 4.** Cleaned IMDB-Wiki Dataset.
Y (Male Progressed Aged Category) respectively. The training of one image is repeated N number of times in order to get progressive aged images.

![Female Face Feature Progression Result](image1)

**Figure 5.** Female Face Feature Progression Result.

![Male Face Feature Progression Result](image2)

**Figure 6.** Male Face Feature Progression Result.

6. Conclusion And Future Works

The adversarial network's remarkable picture-generating result prompted us to utilize GANs to solve our face aging problems. In this paper, we employ Cycle GAN, which converts images through one category to another, holding for the other way around. Older approaches to facial aging did not place a strong emphasis on identity retention. Based on GAN's (Generative Adversarial Networks), this study suggests an innovative and impactful facial aging technique. Age categories are set based on gender data, as needed in the study results, and then the networks are trained independently based on gender details and specified age categories.

The framework we utilize is used in most experimental contexts, although the loss function has been changed with WGAN-GP. This method has a broad application in that it can be used to preserve one's natural identity while also generating a picture appropriate for a specific age group. We decided to compile a single facial feature dataset with a well-balanced age group and a precise age label for future work. The age classifications in the existing facial pictures collection are estimated by special approaches that may not be notably accurate. We can further work on improving the quality of the source images, due to which the accuracy of the model is impacted.

References

[1] Mehdi Mirza and Simon Osindero. 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 (2014).

[2] Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015).

[3] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” arXiv preprint arXiv:1703.10593, 2017]
[4] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv preprint(2017).

[5] C. Vondrick, H. Pirsiavash, and A. Torralba, “Generating videos with scene dynamics,” in Advances In Neural Information Processing Systems, 2016, pp. 613–621.

[6] Wei Wang, Yan Yan, Stefan Winkler, and Nicu Sebe. 2016. Category specific dictionary learning for attribute specific feature selection. IEEE Transactions on Image Processing 25, 3 (2016), 1465–1478.

[7] Hongyu Yang, Di Huang, Yunhong Wang, Heng Wang, and Yuanyan Tang. 2016. Face aging effect simulation using hidden factor analysis joint sparse representa-tion. IEEE Transactions on Image Processing 25, 6 (2016), 2493–2507.

[8] Wei Wang, Zhen Cui, Yan Yan, Jiashi Feng, Shuicheng Yan, Xiangbo Shu, and Nicu Sebe. 2016. Recurrent face aging. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2378–2386.

[9] Zhifei Zhang, Yang Song, and Hairong Qi. 2017. Age progression/regression by conditional adversarial autoencoder. In The IEEE Conference on Computer Vision And Pattern Recognition (CVPR), Vol. 2.

[10] Liu, M.-Y., Breuel, T., & Kautz, J. (2017). Unsupervised image-to-image translation networks. In Proceeding advancement neural information process system (pp. 700–708).

[11] Antipov, G., Baccouche, M., & Dugelay, J. L. (2017). Face aging with conditional generative adversarial networks. In Computer vision and pattern recognition. arXiv:1702.01983v1.

[12] Karras, T., Samuli Laine, T. A., & Lehtinen, J., (2018). Progressive growing of GAN for improved quality, Stability and variation. In Conference ICLR, arXiv:1710.10196v3.

[13] Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative adversarial text to image synthesis. In Proceeding international conference on machine learning (pp. 1060–1069).

[14] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2018). Unpaired image to image translation using cycle-consistent adversarial networks. arXiv:1703.10593v6.

[15] Isola, P., Zhu, J., & Efros, A. (2017). Image to image translation with conditional adversarial network. In Proceeding IEEE international conference computer vision and pattern recognition (pp. 1125–1134).

[16] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2018). Unpaired image to image translation using cycle-consistent adversarial networks. arXiv:1703.10593v6.

[17] Xu, Z., Wilber, M., Fang, C., Hertzmann, A., & Jin, H. (2018). Learning from multi-domain artistic images for arbitrary style transfer. arXiv preprint arXiv:1805.09987.

[18] Grigory Antipov, Moez Baccouche, Sid-Ahmed Berrani, Jean-Luc Dugelay, Apparent age estimation from face images combining general and children-specialized deep learning models, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops 2016, pp. 96–104.

[19] Sveinn Palsson, Eirikur Agustsson, Radu Timofte, Luc Van Gool, Generative adversarial style transfer networks for face aging, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops 2018, pp. 2084–2092.

[20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 770–778.] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 770–778.