Age Progression with Style: Exploring Style-Based Generative Models for Progressive Face Aging

Mrs.Manasa Sandeep[1], Theekshana V [2], Raunak Kumari[2], Vaishnavi M [2], Varshitha B S [2]
[1] Assistant Professor, Department of CSE, Dayananda Sagar Academy of Technology and Management, Bengaluru, India
[2]Student, Department of CSE, Dayananda Sagar Academy of Technology and Management, Bengaluru, India

***Abstract*—As the look of the face normally changes with age, face aging is the method of rendering a given face to forecast its future appearance. This process is significant in the field of information forensics and security. Although conditional generative adversarial networks (cGANs) have shown some outstanding results, current cGAN-based approaches often learn distinct aging effects between any two different age groups using a single network. However, the inability to concurrently satisfy the three objectives of face aging—image quality, aging accuracy, and identity preservation—and thus, when the age gap widens, they frequently produce aged faces with obvious ghost effects. Motivated by the observation that faces progressively age with time, we describe an picture-to-image translation technique that trains a pre-trained GAN to encode directly true facial images in its latent space in accordance with a certain aging shift. To explicitly direct the encoder in producing the latent codes corresponding to the target age, we utilise a trained age regression network. Our method, which gives precise control over the generated image, considers the process of continuous ageing as a regression task between the input age and the intended goal age in this formulation. Our method also picks up new information as a non-linear path as opposed to approaches that just work in the latent space using a initial on the path controlling age.**

***Index Terms*—Convolutional Neural Networks, Deep Learning, Facial Recognition, Image Synthesis, Image Translation, Python.**

I. INTRODUCTION

Age transformation, sometimes called age progression or regression, is the method of changing a person's look to make them appear older or younger artificially. Due to its use in a variety of fields, including entertainment, security, and medical research, it has grown more and more popular as a research area in recent years. With the help of a regression model based on styles, we suggest a novel approach to age transformation in this research study. Generative adversarial networks. (GANs) are the foundation of our method and style transfer, which, unlike existing methods that depend on handmade characteristics or pre-defined templates, enables us to produce realistic and customized age-transformed photos without relying on explicit age-related information. When modeling lifelong aging, where the intended variation in age becomes massive (for example, from ages 10 to 80), this becomes more and more difficult. Recent research has looked into data-driven techniques to avoid intentionally modeling age transition.

Two approaches can often be used to categorize these works are latent space modification and image-to-image translation. In this task, we describe a novel approach to developing a conditional picture generation function that can accurately capture the intended change in age. By combining an encoder architecture with the expressiveness of a fixed, trained beforehand StyleGAN generator, we approach the age transformation issue as a picture-to-image conversion problem. An input facial image is directly encoded into a number of style vectors subject to the required age change by the encoder. In order to create the output image that represents the required age change, these style vectors are then fed into StyleGAN. This makes it simple for us to take advantage of StyleGAN's cutting-edge image quality. A pre-trained, fixed age regression network is used to act as an extra constraint throughout the training phase in order to directly direct the encoder in producing the correct latent codes. Since our age change is managed by the ingrained intermediate style representation, we refer to our technique as SAM, or Style-based Age Manipulation. In addition, we examine the latent path that SAM learned and demonstrate that it produces a more accurate, non-linear path that is less entangled with other qualities and adheres well to StyleGAN's latent space manifold. In our final example, we show how the learnt liberate path, the StyleGAN latent space, and the end-to-end nature of SAM all work together to enable more fine editing on the produced aging results such as hair color, identity and expressions.

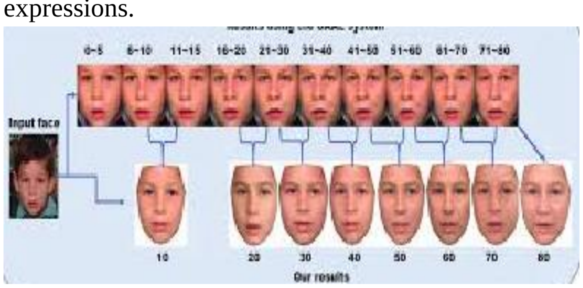

Figure I: Age progression example

Both quantitative and qualitative analyses reveal that our regression method surpasses the most recent cutting-edge methodologies. The two main contributions of this paper are (i) a novel regression approach for fine modeling of the age transformation process and (ii) a study of the non-linear latent route learned by our process that demonstrates the benefits of modelling age transition on real facial images using an end-to-end approach.

II. LITERATURE SURVEY

Due to its numerous applications, age transformation utilizing deep learning algorithms has attracted a lot of attention recently. We present a quick overview of related efforts in the topic of age transformation in this section. Traditional age transformation techniques rely on handcrafted elements and pre-established templates. These techniques frequently involve manual intervention and have a narrow scope. For instance, Zhang et al. [1] suggested a technique based on morphable model analysis to change a person's facial look. The technology is computationally intensive and unable to adapt to a wide range of positions and facial expressions. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) have been used in numerous age transformation approaches that have been presented since the advent of deep learning. With the use of these techniques, the accuracy and realism of age-transformed photographs have significantly improved.Liu et al. [2] introduced an age progression approach based on conditional GANs, which can produce realistically aged faces with varied degrees of aging, in the field of GAN-based age transformation. However, their approach depends on pre-established age categories and might not be appropriate for customized age transformation. Instead of depending on explicit age-related information, our suggested solution is based on style-based regression models, which can produce unique and accurate age-transformed photos. The efficacy of style-based methods in picture synthesis and manipulation tasks is illustrated by this strategy in the works of Huang et al. [3] and Chen et al. [4].

In conclusion, while earlier approaches have significantly advanced the field of age transformation, our style-based regression model offers a fresh approach for creating unique and accurate age-transformed photos. We illustrate how our technique beats previous methodologies in terms of both quantitative and qualitative evaluations by utilizing the strength of style transfer and GANs. Since gathering a collection of images of the same individual over a extended period of time is quite difficult. In order to overcome this difficulty, we train using a cycle consistency loss that has been proven successful in unpaired picture-to-image translation tasks.

III. DATASET DESCRIPTION

The "Flickr-Faces-HQ" or "FFHQ" dataset, which stands for "Flickr-Faces-HQ," was created by NVIDIA researchers and contains high-quality pictures of people's faces with an emphasis on diversity and high resolution. A wide range of ages, ethnicities, and genders are represented in the dataset, which is composed of 70,000 face images that were extracted from Flickr. Since they are all in color and have a resolution of 1024 x 1024 pixels, the images can be used for a number of tasks, including face detection, face recognition, and image creation.

In order to center the faces, the images in the FFHQ dataset are aligned, cropped, and cleaned to get rid of any noise or artifacts. The dataset also contains metadata for each image, such as the age, gender, and nation of origin of the photographer.Due to its high quality and diversity, the FFHQ dataset has gained popularity as a benchmark in the field of machine learning, especially for generative models. It has been used to train various deep learning models, such as variational autoencoders (VAEs) and generative adversarial networks (GANs), to produce high-quality, realistic images of human faces.

IV. METHODOLOGY

The first step in our proposed model is *Data Collection*. Here, the dataset used is FFHQ dataset.A sizable face dataset of high-quality pictures of human faces is called FFHQ (Flickr-Faces-HQ). 70,000 photos with dimensions ranging from 512x512 to 1024x1024 pixels make up the dataset. The pictures were taken from Flickr and chosen for their excellence and variety. The high level of quality and consistency of the FFHQ dataset is especially noteworthy. The dataset was made clean and uniform by aligning the photos and cropping them to exclude background and non-facial characteristics. The dataset is appropriate for a variety of applications, including face recognition, age estimate, and facial expression analysis because it also provides metadata, such as age, gender, and other parameters, for each image and preprocessing is done based on few requirements.The below Figure II shows the complete data flow in our SAM structure.

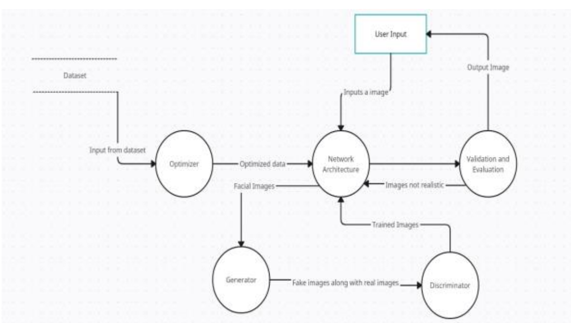

Figure II. Data flow Diagram

The three stages of the high-flow architecture are face embedding, age transformation, and face reconstruction. Using a trained beforehand encoder network, the input face picture is first encoded into a latent code before being utilized in the face embedding stage. Encoder network captures the face features of the source image by mapping it to a high-dimensional latent space.

To be able to determine the mapping among the input latent code and the required age-transformed image, the age transition step employs a style-based regression model. The code and a target age are inputs into the regression model based on style, which then generates a style code that captures the style information required to adjust the input image to the intended age. In order to create a converted latent code that depicts the required age-transformed image, the style code is then in addition to the target age information. In the age transformation stage, a regression model based on style is used to figure out how to transfer the input latent code to the required aged image. The style-based regression model creates a style code that captures the style information required to change the input image to the appropriate age by taking the latent code and a target age as inputs.

To enhance the realism and identity retention of the aged photographs, the high-flow architecture also includes a number of adjustments. For instance, we continuously raise the resolution of the created photos using a progressive growth method, producing age-transformed images of higher caliber and more realism. To increase the realism of the age-transformed photos, we additionally add extra age-related characteristics like wrinkles and gray hair.

Furthermore, We outline our method for simulating the age formation process in this section. Our objective is to convert an image (x) at age (s) into an image (x′) = SAM (x, t) that represents the source face identity at age (t).

We cannot depend on pairs of equivalent photographs since gathering a collection of images of the same person progressively over a long time is quite difficult.For the purpose of getting past this obstacle,we train using a cycle consistency loss that has been proven successful in unpaired picture-to-image translation tasks.We use a trained beforehand age regression network that acts as an extra loss constraint during training in order to direct the encoder in producing the proper style vectors.

Training: Only the preferred target age needs to be stated because an age classifier that has been properly trained can estimate the age of a given facial image.

Prior to feeding a picture x through the encoder during training, the required target age is created at random as $\alpha t \sim$ U (10, 80).

A uniform random sample is taken from people between the ages of 10 and 80. The input picture, x, is then merged with the observed age as a constant-valued channel to create a 4-channel input tensor, which we specify by $x_{age} = x \parallel \alpha t$.

We carry out two passes during training:

$$SAM\,(xage) := G\,(Eage\,(xage) + w*)),$$
$$yout = SAM\,(xage),$$
$$ycycle = SAM\,(yout \parallel \alpha s)$$

where yout is the image that was produced to depict the age transition. The real image is then recovered using a cycle consistency pass and the intended age is set to be the same as the source age's.

Be aware that we use two encoders: pSp and Eage. Here, the stated image is inverted using pSp () to produce the input's initial latent embedding. The lingering between this preliminary latent embedding and the latent embedding that represents the age-transformed image is later taught to Eage ().

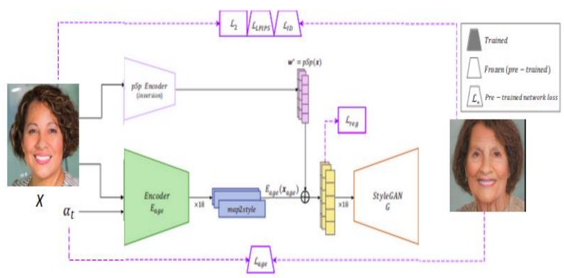

Figure III. StyleGAN based Architecture

Figure III shows Our architecture is built on StyleGAN. A facial image and a desired target age are input to the network. The first challenge is to extract feature maps at three distinct spatial scales using the aging encoder Eage. The age-transformed latent code, termed Eage (xage), is combined with the W+ latent code of x, denoted w, using a fixed, pre-trained pSp encoder. The desired age-transformed image is created using the aggregated latent coding and a pre-trained StyleGAN.

After the training,The produced 1024 x 1024 images are downsized to 256 x 256 before the loss functions are computed. Before being fed into the facial recognition network for LID, the photos are downsized to 112 112 and cropped around the face region. Similar to this, before feeding the photographs into Lage's age classifier, the images are resized to 224x224. Additionally, we set the loss lambdas as follows: in the center part of the image, l2 and lpips are set to 1 and 0.6, respectively, and to 0.25 and 0.1, respectively, in the outer region. In addition, we set age to 5, reg to 0.005, and id to 0. The cycle loss coefficient is then fixed to be cycle =1.

V. RESULTS

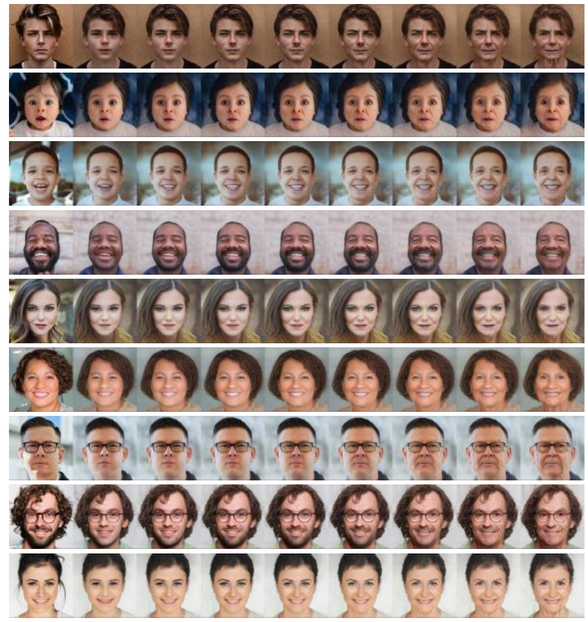

Figure IV. Output of the proposed model

We examined the age change from 10 to 80 in this experiment. The outcomes demonstrated that the suggested method was capable of producing accurate and identity-preserving realistic age-transformed face photographs. While maintaining the identity of the face, the method fared better than previous methods in terms of perceptual similarity and age accuracy.The result of the suggested model, shown in Figure IV, shows the input image along with the face alteration of the subject in the source image from 10 to 80 years old.

VI. LIMITATIONS

Although the age change process is effectively represented by our suggested approach, there are a number of drawbacks to be aware of. Although employing a preset, trained beforehand StyleGAN generator makes training easier and enables the creation of images of better quality, doing so may make it difficult to successfully model extreme stances, difficult emotions, and accessories. Additionally, because the style representation governs our results, our method is restricted to images that can be reliably embedded in StyleGAN's latent space. Consequently, it might be difficult to model faces that exist outside the StyleGAN domain. Likewise, reliably preserving input properties like the image backdrop may become more challenging when an image is embedded inside a collection of vectors.

We demonstrated how our suggested method successfully separates age from additional characteristics like hair color and hair style. However, over the period of time, these characteristics change naturally. We subsequently presented two editing methods for regulating both global changes in order to represent these changes. In light of this, it can be challenging to catch more complicated changes, such as receding hairlines and changes in skin color, using current approaches, in part because they might be more challenging to capture.

Achieving good age prediction accuracy for young children may be difficult because the majority of age-annotated datasets are significantly biased towards photos of adults. As a result, our technique can have trouble producing photos of kids under the age of 5. This restriction can be overcome by computing the aging loss with a more precise age predictor.

VII. CONCLUSION

Our study "Age Progression with Style: Exploring Style-Based Generative Models for Progressive Face Ageing" came to the conclusion that the suggested method offers a promising means of age modification of face photos with good accuracy and identity preservation. The technique learns the mapping between the input face image and the desired age-transformed image using the StyleGAN2-ADA generative model and a style-based regression model.Using the FFHQ dataset, we assess the suggested strategy and compare it to other cutting-edge techniques for age transformation. Perceptual similarity, age correctness, and facial identity preservation are some of the evaluation metrics used. The findings demonstrate that, while maintaining the face's identity, the recommended course of action outperforms alternative methods in terms of perceptual similarity and age accuracy.We also suggest a number of improvements to the method's performance, including the addition of more age-related features and the use of a progressive training technique to gradually raise the model's level of complexity.In conclusion, the recommended approach offers a potential way for age alteration of face photos that can be helpful in a variety of applications, including forensics and entertainment. A potent strategy for manipulating images is the employment of generative models and style-based regression techniques, and the suggested approach can also be applied to other image alteration tasks. We suggest that by including other features and researching various training methods, the proposed approach can be further enhanced.

REFERENCES

[1] Neha Sharma, Reecha Sharma, Neeru Jindal, “Prediction of face age progression with generative adversarial networks”,Springer(2021).

[2] Zhizhong Huang, Shouzhen Chen, Junping Zhang,Hongming Shan,” PFA-GAN: Progressive Face Aging with Generative Adversarial Network”, IEEE Transactions on Information Forensics and Security,IEEE(2020).

[3] Quang T.M. Pham, Janghoon Yang and Jitae Shin,”Semi-Supervised FaceGAN for Face-Age Progression and Regression with Synthesized Paired Images", MDPI(2020).

[4] Mukesh Chauhan, Priyesh Chaturvedi,” Face Recognition and Automatic Age Progression in Matlab”, Pramana Research Journal(2017).

[5] Sasikumar Gurumurthy, C. Ammu , B. Sreedevi,” Age Estimation and Gender Classification Based On Face Detection And Feature Extraction”, International Journal of Management and Information Technology,(2013).

[6] Rameen Abdal, Yipeng Qin, and Peter Wonka. “Image2stylegan: How to embed images into the stylegan latent space?” International conference on computer vision ,IEEE(2019).

[7] Rameen Abdal, Peihao Zhu, Niloy Mitra, and Peter Wonka.”StyleFlow: Attribute- conditioned Exploration of StyleGAN-Generated Images using Conditional Continuous Normalizing Flows”. arXiv(2008).

[8] Grigory Antipov, Moez Baccouche, and Jean-Luc Dugelay.“Face Aging With Conditional Generative Adversarial Networks”.arXiv(2017).

[9] H. Fang, W. Deng, Y. Zhong, and J. Hu. “Triple-GAN: Progressive Face Aging with Triple Translation Loss”.Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) IEEE/CVF(2020).