Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 45 (2024), No. 2, pp. 469–480
DOI : 10.47974/JIOS-1564

Conditional GAN for face aging

Anil Kumar Bhardwaj *
School of Electronics and Communication Engineering
Shri Mata Vaishno Devi University
Katra
Jammu and Kashmir
India

Charu Arora †
Department of Applied Science
Bharati Vidyapeeth's College of Engineering
Guru Govind Singh Inderprashtha University
New Delhi
India

Nisha Malhotra §
Payal Malik ‡
Arvind Rehalia @
Department of Information Technology
Bharati Vidyapeeth's College of Engineering
Guru Govind Singh Inderprashtha University
New Delhi
India

Abstract

Face Ageing is a method for converting a person's image into a certain age group. Recent research has shown that generative adversarial networks are capable of producing artificial pictures. In the paper, a conditional GAN- based model of facial ageing is proposed. Finally, a cutting-edge "Identity Preserving" optimization strategy is initiated. The suggested technique has tremendous potential, as shown by an objective analysis of the photos of aged

* *E-mail:* **anil.bhardwaj@smvdu.ac.in (Corresponding Author)**
† *E-mail:* **charuarora099@gmail.com**
§ *E-mail:* **nisha.jan89@gmail.com**
‡ *E-mail:* **malik.88.p@gmail.com**
@ *E-mail:* **rehaliaarvind@gmail.com**

and rejuvenated faces produced utilising cutting-edge facial recognition and age estimation technologies.

Subject Classification: *94-XX.*

Keywords: *Face aging, GAN, Face synthesis.*

1. Introduction

Translating a person's face into a certain age groupis known as facial ageing, also known as human face ageing synthesis [1].

The issue of facial ageing has gained a lot of attention from academics due to its applicability in several fields [2], including face recognition across age groups [3,4] and entertainment. Aneffective face ageing mesh consists of following:

1. Produce pictures of a high caliber.
2. Maintain facial identity.
3. Produce translations across several domains.

Facial ageing just affects aperson's appearance as they age or get younger. It does not affect who they are as a person. Incorporating facial ageing is now feasible thanks to a technique called Generative Adversarial Networks. A family of neural networks for unsupervised learning is called Generative Adversarial Networks (GANs) [5]. In 2014, Goodfellow raised the idea of GANs. GANs can analyze, capture, and duplicate changes within a dataset since they are basically a system of two competing neural network models that competewith one another. It has been shown that adding even a tiny amount of noise to the original data may readily fool the majority of conventional neural networks into misclassifying objects. Surprisingly, when noise isincluded, the model's confidence in a false positiveis larger than it is in a true positive. Due to the fact that most machine learning models only learn from a little quantity of data, which is extremely disadvantageous and prone to overfitting, this enemy exists. Input and output are also almost linearly mapped. Even a slight change in a point in the feature space can cause the data tobe misclassified because, although the separation borders between various classes may appear to be linear, they actually consist of linearities.

Three categories can be used to categorize generative adversarial networks (GANs):

- Generative: Acquaint yourself with a generative model that explains how data is produced using a probabilistic framework.
- Adversarial: The model is being instructed in a hostile setting.
- Networks: For training, use deep neural network (AI) techniques.

We have a discriminator and a generator neural network in a GAN. The generator tries to trick the discriminator by creating phony samples of data. The discriminator aims at distinguishing between real and false ones. The generator and the discriminator compete with each other throughout the training stage. The process is carried out in repetition and each time, the generator and discriminator get better at doing their respective jobs.

At the limit, the generator consistently creates flawless copies of the input domain, and the discriminator consistently misclassifies and forecasts "uncertain" outcomes (such as 50% for true and false). We don't need to get to this point in order to get at a functional generator model; thisis only an illustration of an idealized situation.GANs frequently use convolutional neural networks or CNNs as generator and discriminator models when working with picture data.

The fact that the first description of the technique used CNNs and image data from the computer vision field provides an explanation for this.

As a consequence, considerable progress has been seen in the more widespread usage of CNNs to generate cutting-edge results in a range of computer vision applications, including object identification and face recognition, in recent years. The latent space—the generator's input—provides a compressed representation of the assortment of pictures or images that were utilized to train the model while modeling picture data. This indicates that the generator creates original images of pictures and provides output that is straightforward for model users or developers to inspect and evaluate. It's possible that the focus on computer vision applications with CNNs and the significant advancements in GAN capabilities over other generative models, whether or not they use deep learning, are both outcomes of the capacity to visually assess the caliber of the generated output. A notable advancement in the technology is the use of GANs to produce conditional output. Agenerative model may be taught to generate new instances from the input domain when the input, a random vector from the latent space, is conditioned (depending) on another input. A class value, such as male or female when creating pictures of individuals, or a digit when creating images of handwritten numbers, may be used as anextra input. If both the generator and the discriminator are dependent on some extra information y, generative adversarial networks may be expanded to a conditional model. y is any additional data, such as class labels or information from other modalities. By adding y as another input layer to the discriminator and the generator, we may make a change.

The discriminator is also conditional, which meansthat it is given a second input in addition to an image that may either be true or false. The discriminator would therefore anticipate the input to come from that class in the case of conditional input of a classification label type and instead trainthe generator to produce samples of that class to trick the discriminator.

In this manner, samples from the domain of a given type may be produced using a conditional GAN. Going one step further, GAN models may be trained using examples from a particulardomain, like images. This makes it possible to use GAN for applications like text-to-image or image-to-image translation. This allows some of the mostamazing uses of GANs, such as transferring styles, colouring pictures, changing summery pictures into wintery ones, turning daytime pictures into nighttime ones, and so on.

Examples of genuine and produced night pictures, as well as (conditionally) real day photos, are provided as input in the case of conditional GANs for image-to-image translation, such as day-to- night transformation. A random vector drawn from the latent space and (perhaps) actual everyday images are supplied to the generator.

One of the many important advancements in the use of deep learning algorithms in areas like computer vision is data augmentation. Improved model proficiency and a regularization impact from data augmentation, which reduces generalization error and leads to better model performances. It works by generating fresh, believable instances from the input domain that serve as the foundation for training the model. In the case of picture data, the procedures are basic and consist of cropping, zooming, and straightforward changes of the existing images in the training dataset.

A successful generative modelling approach to data augmentation offers an alternate and maybe more domain-specific method. Although it is not often acknowledged, data augmentation is really a more straightforward kind of generative modeling. In complex domains or regions where there is a dearth of data, generative modeling provides a technique to more effectively train models. GANs have seen a lot of success in this use case in areas like deep learning. Most other models (including GANs) lose 50% of the subject's identity or appearance since they ignore facial emotions and insignificant accoutrements like spectacles and beards.

Age-cGANs take into account each of these factors, making them superior than simple GANs in this area. Four networks make up Age- cGAN: the encoder, FaceNet, generator, anddiscriminator networks.

The encoder teaches how to use the z0 latent vector to inversely map the input face pictures and age condition. FaceNet is a face recognition network that learns to distinguish between an image that is being input (image x) and one that has been rebuilt. Generator network is responsible for creating a picture from a hidden (latent) representation made up of a facial image and a state vector. Differentiation between the authentic images and false images is done by the discriminator network.

2. Related Work

Three categories of research on face ageing exist: prototype-based approaches, physical model-basedmethods, and deep learning methods. A non- parametric model is used in the prototype-based approach [6,7]. However, the fundamental problem with these methods is that since the outcomes are based on typical features, little nuances are missed. Physical models provide intricate simulations of face appearance to simulate ageing processes and pinpoint a set of markers or statistical factors. The biggest disadvantage is that they need several faces with the same ID under various age groups.

Generative Adversarial Networks (GANs): Inimage-generation tasks, image-to-image conversion, etc., GANs have shown tremendous performance. The basic GAN developed by Ratford et al. is expanded, and it is shown that fully convolutional architecture is effective for GAN training. To obtain high-resolution results in face photographs, Karras et al. suggest a progressively expanding generator and discriminator. A self-awareness module is expanded by SA-GAN to encourage attention to detail. Cycle GAN's pix2pix image translationmodel is one of the first to provide promising image transformation outcomes in both supervised and unsupervised picture situations between two domains.

The suggested approach was expanded upon by Choi et al. depicted in Figure 1 to include multi-domain picture manipulation utilizing only two networks. Similar to cyclegan, pix2pix IPC models use cyclic reconstruction losses to achieve unsupervised and supervised image translation with promising results in face ageing on low resolution. Fang et al. presented a novel triple translation loss for creating correlation between various age groups. Amore recent model, HRFAE, classifies the FFHQ dataset in order to get over the problem of the lack of an open-source high resolution face ageing dataset. It has shown to be quite effective in creating images.

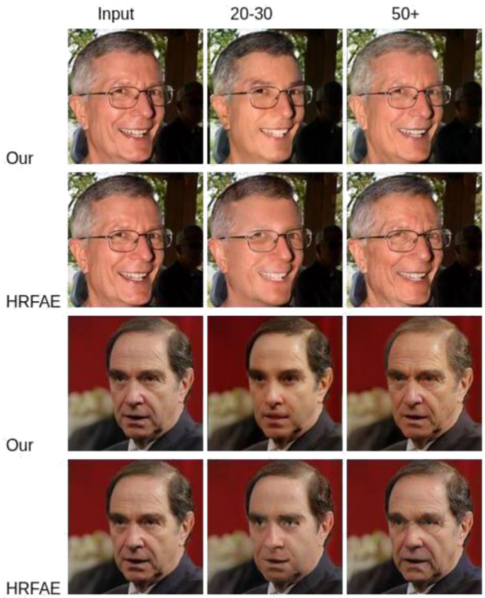

Figure 1

Comparison of 2-Domain Age translation

3. Proposed methodology

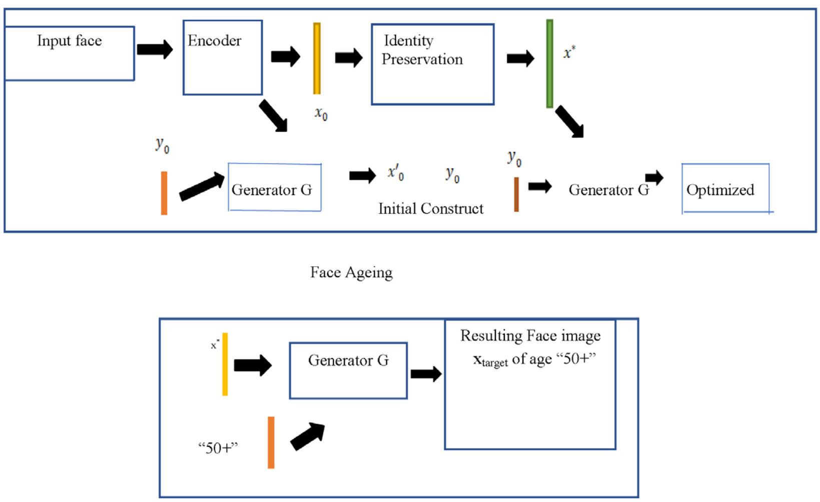

Figure 2

Our methodology

Our approach to face ageing is based on Age- cGAN, a generative model for synthesizing humanfaces in the relevant age ranges depicted in Figure 2.

1. To create a reconstructed face using the formula*G (z*, y0),* select the best latent vector z* for an input face picture x with an age of 0.
2. To produce the final face picture = *G(z*, ytarget)* for a specific target age, just alter the agesat the generator input.

The crucial stage is the first one, which involves reconstructing the input face. We includeconditional information into G's input and D's first convolutional layer in the Age-cGAN modeldescribed in this study. Age: The ADAM method is used to optimise the cGAN. We created six age categories—0–18, 19–29, 30-39, 40–49, 50–59, and 60+—to encode a person's age.

Initial approximation of the latent vector Unlike auto-encoders, cGAN lacks an explicit method for inversely translating an input picture (x) with an attribute (y) to the latent vector (z),which is x = G(z,y). By refining the encoder E, a neural network that roughly approximates the inverse mapping, we areable to get around this issue [8]. Even while GAN is undoubtedly the most powerful generative model available today, it is unable to faithfully replicate all genuine face photos, which include many options for facial features, accessories, backdrops, etc. In most cases, Age - cGAN may estimate the input natural face photos rather than reconstructing them precisely.

Latent Vector Analysis

The face ageing task's end objective presupposes that a person's age needs to be changed while theiridentity is preserved. Although the initial latent approximation produced by E results in aesthetically appealing face reconstructions, around 50% of the time the actual identity is lost.. As a result, we must enhance the initial latent approximation. Here, we introduce a fresh "Identity Preserving" approach. The basic premise is that the difference between identities in the original and reconstructed images can be expressed as the Euclidean distance between the corresponding embeddings of the original and reconstructed image, given a face recognition neural network (FR) that can identify aperson's identity in an input face image (x). Reducing the aforementioned distance may thus improve identification preservation in the resultingimage.

4. Experiments

Age - cGAN was trained using the approximately 110,000 picture IMDB-Wiki_cleaned dataset. Age-cGAN completely decodes and makes independent the picture data represented by latent vector z and constraint y. More particularly, we see that the latent vector z uniquely encodes age and the latent vector y uniquely encodes a person'sidentity, facial posture, hairstyle, etc. To formally assess how well Age - cGAN is effective at producing faces that fall into precise age groups; we utilized the most recent age estimate from CNN as in [9]. We assess how well CNN's projected age performs when compared to actual photographs. Despite the synthetic materials' estimated age, the average age prediction accuracy of the resultant synthetic photos is only 16% poorer than the natural images—images that the CNN never saw during training. This demonstratesthat our technology may be used to produce generations of age-appropriate, realistic face pictures mentioned in Table 1 and Table 2.

Ageing and Face Reconstruction with Identity Preservation. It can be noticed that the optimised reconstructions are more accurate than the first ones and closest to the real (original) ones. The decision-making process gets more challenging when comparing two latent vector optimization strategies, however. The "per pixel" more accurately depicts surface-level facial characteristics like hair colour. On the other hand, the "preserve identity" optimisation better captures identity attributes. We use the OpenFace [10] programme to compare the two methods for identity-preserving face reconstruction objectively. One of the top open source facial recognition programmes right now is OpenFace. OpenFace detects whether or not two face photographs are ofthe same person if there are two of them. Table 1 displays the proportion of successful Open Face outputs (where the programme concludes that the face picture and its reconstruction belong to the same individual). Only 50% of the test scenarios result in OpenFace being able to identify the original subject in the initial reconstruction. The "per pixel" optimisation marginally raises this proportion. In comparison, the identity-preserving optimisation strategy significantly improves face recognition, with a performance of 83.01%. Once we have a face reconstruction of the original picture that preserves identification, we replace the original age category with the desired age to getthe output.

Accuracy = number of true predictions / total predictions

Table 1

OpenFace FR Score

Reconstruction Type	FR Score
Initial Reconstruction	52.9%
"Pixelwise" Optimization	58.7%
"Identity-Preserving" Optimization	83.01%

Table 2

Epoch-wise accuracy

Training	Accuracy					
Epoch	40%	50%	60%	70%	80%	90%
Epoch=10	94.7562	95.8987	95.9624	95.9884	96.0365	96.0846
Epoch=20	96.0789	96.1204	96.1327	96.1686	96.2168	96.2650
Epoch=30	96.2589	96.3005	96.3132	96.3488	96.3971	96.4454
Epoch=40	96.4389	96.4806	96.4937	96.5290	96.5774	96.6258
Epoch=50	96.6605	96.6742	96.7090	96.7575	96.8060	96.9030

5. Conclusion

In this study, we presented a novel, efficient technique for artificially ageing the human face using the age-conditional GAN (Age - cGAN). There are two phases in the method:

1. Enter face reconstruction to get the best latent approximation possible.
2. The act of face ageing is carried out bychanging condition y at the generator's input.

The innovative "Identity- Preserving" optimizationstrategy forms the basis of our methodology. This enables the identity of the original individual to be maintained throughout the rebuilding. This strategy is used everywhere. This implies that it may be utilized for a variety of face changes, including beard growth and the usage of sunglasses, in addition to facial ageing.

6. Limitations

While the proposed facial aging method, leveraging a conditional Generative Adversarial Network (GAN) and an "Identity Preserving" optimization strategy, shows promise, there are inherent limitations to consider. One potential challenge is the accurate representation of subtle aging nuances, such as fine lines and wrinkles. The model's performance might exhibit variability across diverse demographic groups, posing challenges in achieving consistent and precise age transformations for individuals from different backgrounds. Addressing these limitations is crucial for refining the method and ensuring its reliability across a broader spectrum of facial characteristics and demographics.

7. Future work

To advance the proposed facial aging method, future work can focus on refining the model's sensitivity to nuanced aging features. Exploring additional contextual information or incorporating advanced features may enhance the model's capability to generate realistic age transformations consistently across diverse individuals. Generalization improvements, especially across various demographic characteristics, could be a key area for research. Additionally, investigating alternative optimization techniques to bolster the model's robustness and accuracy would contribute to its reliability in generating accurate and nuanced facial age transformations.

References

[1] Fu, Y., Guo, G., & Huang, T. S., Age synthesis and estimation via faces: A survey. IEEE transactions on pattern analysis and machine intelligence, 32(11), 1955-1976.Diederik P Kingma and Max Welling, "Auto-encoding variational bayes," in Proceedings of International Conference on Learning Representations, Banff, Canada, 2014 (2010).

[2] Ramanathan, N., & Chellappa, R., Face verification across age progression. *IEEE transactions on image processing,* 15(11), 3349-3361 (2006).

[3] Amos, B., Ludwiczuk, B., & Satyanarayanan, M., Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science*, *6*(2), 20 (2016).

[4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y., Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27 (2014).

[5] Wang, Z., Tang, X., Luo, W., & Gao, S., Face aging with identity-preserved conditional generative adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7939-7947 (2018).

[6] Fang, H., Deng, W., Zhong, Y., & Hu, J., Triple-GAN: Progressive face aging with triple translation loss. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 804-805 (2020).

[7] Perarnau, G., Van De Weijer, J., Raducanu, B., & Álvarez, J. M., Invertible conditional gans for image editing (2016). arXiv preprint arXiv:1611.06355.

[8] Antipov, G., Baccouche, M., Berrani, S. A., & Dugelay, J. L., Apparent age estimation from face images combining general and children-specialized deep learning models. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 96-104 (2016).

[9] Amos, B., Ludwiczuk, B., & Satyanarayanan, M., Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science*, 6(2), 20 (2016).

[10] Kingma, D. P., & Welling, M., Auto-encoding variational bayes (2013). arXiv preprint arXiv:1312.6114.

[11] Schroff, F., Kalenichenko, D., & Philbin, J., Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 815-823 (2015).

[12] Ling, H., Soatto, S., Ramanathan, N., & Jacobs, D. W., Face verification across age progression using discriminative methods. *IEEE Transactions on Information Forensics and Security*, 5(1), 82-91 (2009).

[13] Xu, C., Tang, Y., Toyoura, M., Xu, J., & Mao, X., Generating users' desired face image using the conditional generative adversarial network and relevance feedback. IEEE Access, 7, 181458-181468 (2019).

[14] Xu, C., Tang, Y., Toyoura, M., Xu, J., & Mao, X., Generating users' desired face image using the conditional generative adversarial network and relevance feedback. IEEE Access, 7, 181458-181468 (2019).

[15] Grimmer, M., Ramachandra, R., & Busch, C., Deep face age progression: A survey. *IEEE Access*, *9*, 83376-83393 (2021).

[16] Ramanathan, N., & Chellappa, R., Modeling age progression in young faces. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), Vol. 1, pp. 387-394 (2006, June). IEEE.

[17] Zhang, Z., Song, Y., & Qi, H., Age progression/regression by conditional adversarial autoencoder. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5810-5818 (2017).

[18] Fu, Y., Guo, G., & Huang, T. S., Age synthesis and estimation via faces: A survey. IEEE transactions on pattern analysis and machine intelligence, *32*(11), 1955-1976 (2010).