젠더보존에 기반한 얼굴 합성 모델 탐구

이소려, 이효종*
전북대학교컴퓨터공학부
lisuli19880104@gmail.com, hlee@jbnu.ac.kr

Exploring the Aged Face Synthesize Model Based on Gender Preservation

Suli Li, Hyo Jong Lee*
Dept. of Computer Science and Engineering, Jeonbuk National University
*Corresponding author

Abstract

Face aging aims to synthesize future face images by reflecting the age factor on given faces. In recent years, deep learning-based approaches have made outstanding progress in simulating the aging process of the human face. However, generating accurate and high-quality aging faces is still intrinsically difficult. We propose a new method that incorporates gender information into the model, which achieves comparable and stable performance. Experimental results demonstrate that our method can preserve the identity well and generate diverse aged faces.
Keywords: convolutional neural networks; generative adversarial networks; Face aging

1. Introduction

Face aging is an image synthesis task that renders face images to different ages with natural aging effects while maintaining the subject's identity and distinctive facial features. It has broad applications, such as finding long-lost children, identifying fugitives, cross-age face verification, etc., to help create a safe and secure society. In recent years, face aging has made remarkable progress, and many methods have been proposed. Compared to the earlier traditional solutions, these methods produce faces with more pleasing aging effects and fewer ghosting problems.

However, the problem has not been essentially solved. Generating accurate and high-quality aging faces is still intrinsically difficult. The generated faces from these methods cannot simultaneously meet three important requirements for face aging: aging accuracy, gender preservation, and identity preservation. They cannot guarantee a smooth aging result because of the inherent complexity of face aging. When the original face images are of young children, this is especially true, where age progression beyond a few years is considered impractical. Sometimes, the aged faces appear younger than those of older age groups. There also has a problem is the erroneous appearance of a male beard on the face of a generated aging female. When the original face images are of young children, this is especially true. For instance, it's common for infants and toddlers to lack obvious gender characteristics. The learning model might inaccurately predict a child's sex and provide an inaccurate simulation.

For this reason, our goal is to achieve high natural and gender preservation face aging. In this paper, we proposed a new deep learning model to solve the mentioned problems.

The rest of the paper is organized as follows. We present a brief literature survey of related work in Section II. We then illustrate our proposed model in Section III. We present our experimental methods and results in Section IV. Lastly, we conclude our research in Section V.

2. Related work

The published methods of face aging can be divided into three categories: physical model-based methods, prototype-based methods, and deep learning methods. Physical model-based methods simulate the changes in facial appearance over time by designing a complex parametric model. However, the main problem of physical model-based methods is that they are computationally expensive and do not generalize well due to the mechanical aging rules. The prototype-based methods first divide the faces into groups according to different ages. Then compute the average faces of people in the same age group as the prototypes. The testing face can be aged by adding the differences between the prototypes of any two age groups. However, prototype-based approaches may miss the personalized information because the age pattern is determined by the average face.

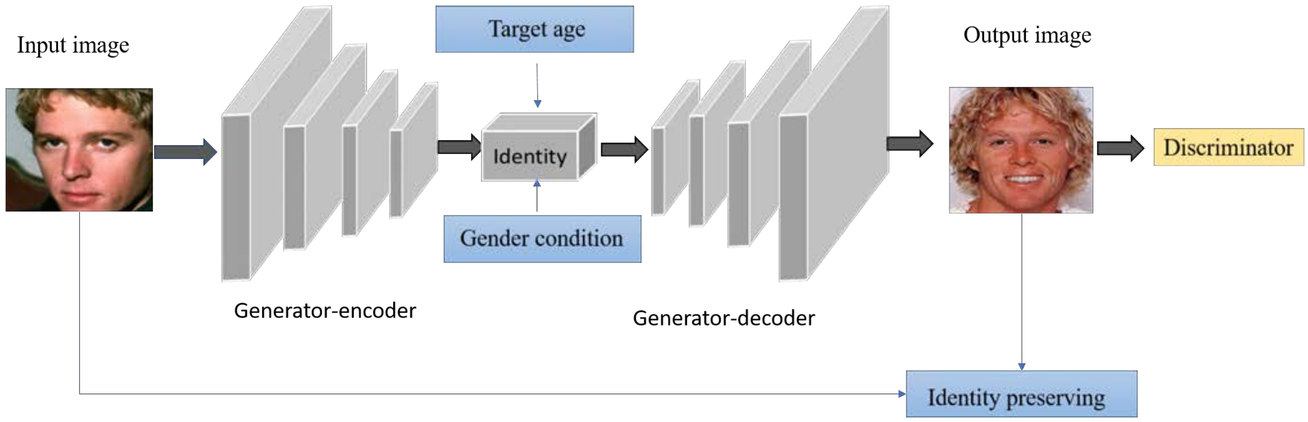

Figure 1. The architecture of our method.

Deep learning methods have recently demonstrated their potential for face aging. Wang et al. [1] utilized a recurrent neural network to represent transformation patterns across different ages smoothly. These models are effective, but they require massive paired faces of the same subject over a long period for training, which is difficult to collect. The cGANs-based methods [2]-[8] combined the concepts of C-GAN [9] and auto-encoder to learn different age mappings by one or two models with a target age group as the condition.

In this work, we are more interested in deep learning based methods. We add gender-preserving components to the model to improve the quality of the generated faces.

3. Methods

We design a short-term aging generator to high natural and gender preservation face. We illustrate our proposed architecture in Figure 1. To learn an age-invariant identity representation for AIFR and synthesize a face image with the same identity, we assume a real face image x with two labels $\{y_{identity}, y_{age}\}$, where $y_{identity}$ represents the label for identity and y_{age} for the target age. Our method is to train a generator G (including encoder and decoder) that can generate a face image x' of a particular age yage. The target age aims to control the transformation process. Considering the huge range of ages in our dataset, we divide different ages into 4 groups: 21-30, 31-40, 41-50, and 51+. Specifically, we introduce a gender modulator within G to reshape the identity features map by considering the target age and utilizing it as self-guiding information. Our gender condition is represented as a two-dim one-hot gender vector. Each dim represents a specific gender. The decoder network takes an identity feature map alongside the gender vector and produces an aged face image. Identity preservation also should go along with a change in age. In addition, we adopt the discriminator presented in [10] as our discriminator D.

Loss Function: It comprises two parts: adversarial loss for matching the distribution of outputs to the data distribution and identity preservation loss.

$$L = L_{adv}(G, D) + L_{ip} \tag{7}$$

4. Experiments

We initialized the model with He initialization [11].

Figure 2. Examples of the data used. The top row shows the original images, and the bottom row shows the aligned and normalized images.

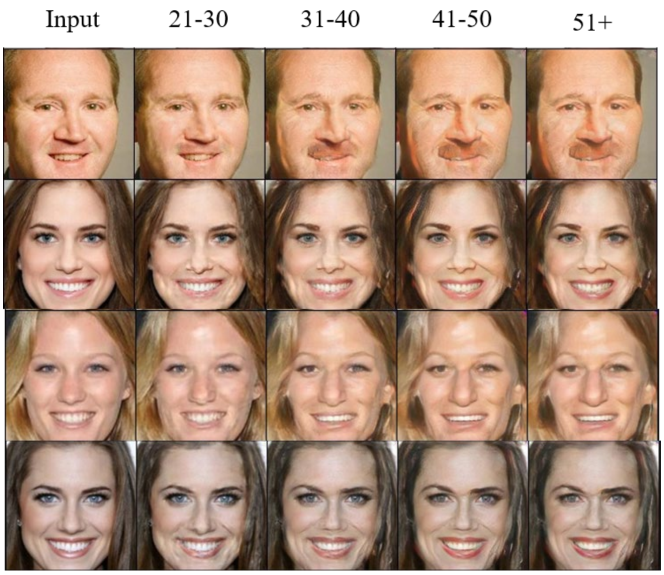

Figure 3. Sample results of our method on CACD dataset for face aging.

The model was implemented based on PyTorch v1.4.0. During training, we trained all models with a maximum of 120,000 iterations on four Nvidia Titan X Pascal GPUs with a mini-batch size of 12. The model was optimized by adam optimization method [12], and the learning rate is 1.0×10^{-4}.

We use the cross-age celebrity dataset (CACD) [13] as the training data. Input images were preprocessed and normalized. We used the multi-task cascaded convolutional network to detect face areas and facial landmarks in the training images. After detecting the eye position, we applied an affine transformation to the data to align the face images based on the detected eye coordinates. All faces were globally cropped to 112 × 112 based on five facial landmarks (two mouth corners, nose center, and two eyes) and a similarity transformation. Figure 2 shows some original and preprocessed face images from the CACD.

Figure 3 shows the sample results for face age progression. Even though some of the synthesized faces look unrealistic, the results show that our method can preserve the identity well and generate diverse aged faces.

5. Discussion

Since this paper mainly presents the face aging framework, the invertible deep learning network for face progression and rejuvenation deserves studying in the future.

Acknowledgment

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant 2019R1D1A3A03103736 and in part by project for Joint Demand Technology R&D of Regional SMEs funded by Korea Ministry of SMEs and Startups in 2021 (No. S3035805).

References

[1] W. Wang, Y. Yan, Z. Cui, J. Feng, S. Yan, and N. Sebe, "Recurrent face aging with hierarchical autoregressive memory," IEEE transactions on pattern analysis and machine intelligence, vol. 41, no. 3, pp. 654-668, 2018.

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

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

[4] H. Yang, D. Huang, Y. Wang, and A. K. Jain, "Learning face age progression: A pyramid architecture of gans," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 31-39.

[5] J. Despois, F. Flament, and M. Perrot, "High-resolution controllable face aging with spatially-aware conditional gans," ed: Google Patents, 2021.

[6] Z. Huang, S. Chen, J. Zhang, and H. Shan, "PFA-GAN: Progressive face aging with generative adversarial network," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 2031-2045, 2020.

[7] J. Song, J. Zhang, L. Gao, X. Liu, and H. T. Shen, "Dual Conditional GANs for Face Aging and Rejuvenation," in IJCAI, 2018, pp. 899-905.

[8] Q. Li, Y. Liu, and Z. Sun, "Age progression and regression with spatial attention modules," in Proceedings of the AAAI conference on artificial intelligence, 2020, vol. 34, no. 07, pp. 11378-11385.

[9] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.

[10] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125-1134.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1026-1034.

[12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[13] B.-C. Chen, C.-S. Chen, and W.H. Hsu, "Cross-age reference coding for age-invariant face recognition and retrieval," European Conference on Computer Vision, pp. 768–783: Springer, 2014.