ITM Web of Conferences **78**, 02016 (2025) https://doi.org/10.1051/itmconf/20257802016
CSEIT 2025

Research on Face Age Synthesis Method

Wanru Zhu

School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi, China

Abstract. Face age synthesis technology simulates the natural change process of the human face as it grows or decreases with age by generating a model, which is of great value in practical applications such as cross-age identification and missing person search. This paper presents a systematic review of research advances in face age synthesis. In the current research context, methods for face age synthesis can be classified into classical methods and deep learning methods. Early classical methods can be divided into physical model-based methods, prototype-based methods and reconstruction-based methods; with the development of deep learning, the current stage of methods is centered on Generative Adversarial Networks (GAN) and their variants. Conditional GAN (cGAN) improves generative controllability through target age labels, disentangled GAN separates shape, texture and identity features, and Attention-based GAN and StyleGAN family of models can further optimize detail fidelity. However, identity drift in long-span synthesis, insufficient generalization across datasets and high computational cost remain key challenges. In the future, lightweight architectures, physiological a priori fusion and ethical compliance mechanisms need to be explored to promote applications in security, medical, and other scenarios.

1 Introduction

Face age synthesis is one of the core topics in computer vision and pattern recognition, aiming to simulate the natural process of face change with age by generating models. As a key technology in the fields of cross-age face recognition, missing person search and digital entertainment, its core challenge is to balance the authenticity of age features with the consistency of identity information.

For example, in the search for missing persons, the relevant personnel can generate the possible adult appearance of a missing person based on his/her childhood photo, thus assisting the police in cross-age identification and improving the efficiency and accuracy of the search. In film and television special effects, the dynamic aging effect of characters needs to maintain the natural transition and realistic effect in the vision, which also relies on the ability of high-fidelity image generation. However, face aging involves non-linear physiological changes such as skin texture degradation and skeletal remodeling, and is affected by individual differences such as genetics and environment, so achieving accurate and controllable age synthesis still faces serious challenges.

zhuwanru@stu.xidian.edu.cn

Early studies were mainly based on statistical models with linear transformation methods. For example, statistical facial models based on principal component analysis (PCA) decompose faces into shape and texture components, and predict aging trajectories through linear functions [1]. Although such methods introduce an age normalization strategy to improve recognition robustness, the generated images have low realism and rely on strong parametric assumptions, making it difficult to model complex nonlinear features. With the breakthrough of deep learning technology, generative adversarial network (GAN) has become a mainstream framework. For example, conditional GAN (cGAN) controls the generation process through target age labeling, which significantly improves image quality and controllability. CycleGAN achieves unsupervised image-to-image transformation by introducing cyclic consistency loss, which enables it to learn the mapping relationship between source and target domains without pairwise training data. StyleGAN, by innovatively introducing a style feature vector, achieves fine control over different scale features (e.g., pose, hairstyle, and facial expression, etc.) in the generated images and enhances the realism of the generated face images.

This paper provides a systematic review of the technology evolution in the field of face age synthesis, covering both classical statistical models and deep learning frameworks, with a focus on comparing the performance differences and applicability scenarios of different methods. By quantitatively analyzing indicators such as age accuracy and mean absolute error (MAE), the paper reveals the advantages of disentangled design and attention mechanism in identity preservation, and explores the challenges of segmentation such as child face synthesis and cross-ethnicity generation, which will provide theoretical and technological references for future research.

2 Classification of Face Age Synthesis Methods

In the current research context, methods for face age synthesis can be classified into classical methods and deep learning methods. Early face age synthesis was mainly based on classical methods, including physical model-based methods, prototype-based methods and reconstruction-based methods [2].

In recent years, with the emergence and development of deep learning technology, the methods of face age synthesis have gradually shifted to the direction of deep learning, among which, GAN-based face age synthesis has set off a widespread concern and research boom in this field. Therefore, this paper will focus on the GAN-based face age synthesis technology and develop an in-depth analysis.

2.1 Classical Methods

Early research on face age synthesis was mainly based on traditional statistical models with linear transformation methods. According to the difference between specific models, the classical methods can be divided into three categories: physical model-based methods, prototype-based methods and reconstruction-based methods.

2.1.1 Physical Model-Based Methods

Physical model-based methods to face age synthesis model physiological changes in the face such as facial structure, skin wrinkles, bone growth, etc. through parametric models, but the generation effect is limited by computational complexity and linear assumptions, which makes it difficult to capture non-linear aging features. Suo et al. proposed CONGRE, which constructs local aging models through spatial anatomical decomposition, and connects short-term modes in conjunction with a Markov process to realize the long-term evolution, and introduce regional consistency constraints to guarantee overall coordination. This method synthesizes aging sequence age estimation errors close to the real data on some dense databases, however, it still suffers from insufficient anatomical modeling refinement, limited retention of individual features, and strong data dependence [3].

2.1.2 Prototype-Based Methods

Face age synthesis using prototype-based method utilizes a non-parametric model by constructing an average face of a specific age group as a prototype face, calculating its difference in facial shape and texture with the prototype face of the target age group as an aging pattern, and then applying this pattern to the target face to synthesize the aged face. Since prototype-based methods have the property of relying on static templates, identity information is easily lost, e.g., the generated image is transformed from the input image in terms of ethnicity and gender. Kemelmacher-Shlizerman et al. proposed Illumination-Aware Age Progression (IAP), which enables continuous aging modeling of a single child's photo to an adult by constructing a relightable age subspace and using cross-age optical flow alignment to maintain illumination consistency without 3D reconstruction. Experiments show that the method is close to the real data in terms of user selection rate in long-span age simulation, but it is insufficient for modeling wrinkles and hair color changes in old age and relies on large-scale age-aligned data. This study enables fully automated age progression in an uncontrolled environment for the first time, providing new ideas for forensic applications [4].

2.1.3 Reconstruction-Based Methods

The reconstruction-based approach decomposes personalized and age-related features through coupled dictionary learning (CDL), which improves the personalized modeling of aging trajectories but still relies on multi-age pairwise data. Lanitis et al. propose a PCA-based statistical face model that decomposes faces into shape and texture components and predicts age change through a linear aging function. The method introduces an age normalization strategy for the first time, aligning the perceived age of the samples during the training and testing phases, and improves the accuracy of the face recognition system from 51% to 66%. However, its reliance on linear assumptions makes it difficult to capture complex nonlinear aging features and produces images with low fidelity [1].

2.2 Deep Learning Methods

With the development of deep learning techniques, GAN and its variants have become the mainstream framework for face age synthesis. According to the technical characteristics, the existing methods can be classified into cGAN, CycleGAN, StyleGAN, Disentangled GAN, Attention-based GAN, and GAN for children.

2.2.1 cGAN

The cGAN takes the target age as the conditional input, controls the generation process through the target age label, and optimizes the image quality by combining with the adversarial loss to generate the age-transformed image.

Chen et al. proposed attention-aware conditional GAN (ACGAN), which integrates the channel attention and spatial attention mechanisms, and optimizes the detail generation by combining with the local discriminator to improve details such as wrinkles and features by focusing on the central region of the face to enhance the authenticity. The model verifies the effectiveness of the attention mechanism in age synthesis, resulting in a significant improvement in the quality of the generated images. However, the model relies on high-quality age labels, is weakly adapted to unlabeled data, and lacks generalization ability [5].

2.2.2 CycleGAN

CycleGAN achieves unpaired data training through cyclic consistency loss.

Sun et al. proposed noise-semantic guided GAN (NSG-GAN). The model adopts a biphasic age-transition framework, introduces ProjectionNet to generate “soft” latent maps, ConstraintNet to disentangle semantic features, and incorporates an attention mechanism to optimize the detail fidelity. NSG-GAN supports unpaired data training and solves the problem of identity consistency and detail distortion, but the model structure is complex and the training convergence speed is slow [6].

2.2.3 StyleGAN

StyleGAN utilizes its semantic latent space to achieve continuous age editing.

Hsieh et al. propose an improved model of StyleGAN2 with an attention channeling mechanism. The model combines channel attention (CBAM) with the StyleGAN2 generator and introduces an identity discriminator to constrain identity consistency. It supports fine-grained generation over wide age spans and reduces identity retention errors, but it has high computational resource requirements and poor real-time performance [7]. Falkenberg et al. proposed a combination of StyleGAN3 and InterFaceGAN, which uses StyleGAN3 to generate an adult image, which is inverted to a child via InterFaceGAN. Based on the potential spatial linear operation, the large-scale synthetic dataset HDA-SynChildFaces is constructed for advancing child age synthesis research, however, the model suffers from identity drift when generating long-span images of children [8].

2.2.4 Disentangled GAN

Disentangled GAN decouples face features into independent components such as shape, texture and identity to model the aging process separately.

The disentangled lifespan face synthesis (DLFS) model proposed by He et al. separates the shape and texture features, and controls the shape transformations through conditional convolution, and the channel attention module controls the texture transformations, thus achieving biologically rational aging. This model is the first to achieve biologically rational shape and texture decoupling and supports nonlinear aging trajectory generation, however, it relies on high-quality paired data and has limited ability to generalize across datasets [9]. The DLAT^{+} proposed by Xie et al. adopts a diversity generation mechanism combined with IDAG a metric to quantify the rationality of identity changes, and the model can generate multimodal rational aging patterns, which solves the problem of unitary synthetic results [10].

2.2.5 Attention-based GAN

Attention-based GAN focuses on key regions, such as wrinkles and features, through the attention mechanism to enhance the authenticity of the generated details.

Huang et al. proposed MTLFace using a multitasking framework, which is based on an attention-based feature decomposition (AFD) to separate the identity and age features, and supports age synthesis and recognition task co-optimization. The model's synthetic data contributes to enhancing the performance of the recognition model, but its synthesis effectiveness under extreme age spans has not been fully validated [11].

2.2.6 GAN for children

GAN for children optimizes identity preservation with age-sensitive balancing by targeting the rapidly changing nature of children's faces.

ChildGAN proposed by Chandaliya et al. combines SAGAN with a multi-component loss function to design conditional encoders to generate children's images, which pushes forward standardization of children's age synthesis, however, its identity consistency decreases in long span synthesis, and identity drift significantly [12]. Falkenberg et al. proposed a combination of StyleGAN3 and InterFaceGAN to generate cross-racial child images through latent space manipulation, but the authenticity of the synthetic data in the lower age group needs to be improved [8].

3 Comparison of methods

This section presents a comparative analysis of GAN based face age synthesis models. Since the control of experimental variables used varies across the literature, representative data from them can be selected for analysis.

Table 1 gives a comparison of the performance of each representative model. The evaluation metrics include mean values of age accuracy (%), MAE, and identity preservation (%). Among them, mean values of age accuracy (%) indicate the proportion of the age of the generated face image falling into the target age group, MAE indicates the mean absolute error between the predicted age and the actual age, and identity preservation (%) indicates the identity preservation rate between the input face and the generated face.

Table 1 Performance comparison of representative GAN-based face age synthesis methods

Method	Year	Dataset	Mean Values of Age Accuracy (%)	MAE	Identity Preservation (%)
DLFS [9]	2021	FFHQ	65.6	3.53±2.81	96.47
NSG-GAN [6]	2022	MORPH	-	1.20±6.81	99.99
MTLFace [11]	2023	MORPH	67.07	2.98	-
DLAT+ [10]	2024	FFHQ-Aging	-	2.6	83.68

For age accuracy metrics, MTLFace performs better on the MORPH dataset (67.07%), thanks to the explicit modeling of age features in the multitasking framework; DLFS has a slightly lower age accuracy on FFHQ (65.6%) due to the decoupled texture and shape design.

For identity retention metrics, DLFS has higher identity similarity (96.47%) due to its decoupling strategy to reduce feature interference; NSG-GAN achieves 99.99% identity verification rate through semantic bootstrapping, but relies on high threshold filtering.

However, existing methods face challenges in long-span age synthesis, cross-dataset adaptation, and computational efficiency.

In terms of long-span age synthesis, most of the current models significantly degrade identity consistency over the span of 30 years (e.g., ChildGAN [12] has a 12% reduction in Rank-1 recognition rate).

In terms of cross-dataset adaptation, DLFS [9] performs well on FFHQ, but identity similarity decreases when migrating to MORPH.

In terms of computational efficiency, StyleGAN2+Attention [8] requires 24GB of GPU memory and is difficult to deploy to mobile.

Therefore, future research could explore lightweight architectures, cross-domain adaptive strategies, and generative models with joint physiological prior.

4 Conclusion

As an important research direction in the field of computer vision and pattern recognition, face age synthesis technology has made significant progress in recent years driven by generative models. This paper systematically reviews the technical evolution of classical statistical models and deep learning frameworks, focusing on various types of GAN-based methods and their performance differences in age accuracy, identity preservation and image quality. Although the classical methods laid the foundation for early research, their shortcomings such as insufficient linear modeling capability and low generative realism limit the practical applications. In contrast, GAN-centered deep learning methods significantly improve the controllability and fidelity of synthetic images by introducing strategies such as decoupling design and attention mechanism. It demonstrates significant advantages in terms of identity retention rate and image quality.

However, identity drift in long-span age synthesis, insufficient model generalization across datasets, and high computational costs remain core challenges. In addition, child facial synthesis still suffers from low generation authenticity and significant identity drift due to drastic physiological changes and data scarcity, while the adaptability of cross-race synthesis needs to be further verified.

Future research needs to seek a balance between technological breakthroughs and ethical compliance. At the technical level, lightweight architectures can be explored to reduce computational costs, physiological a priori knowledge can be introduced to enhance synthetic rationality, and cross-domain adaptive strategies can be developed to enhance model generalization. Meanwhile, a more comprehensive evaluation system needs to be constructed to unite indicators such as age accuracy, identity retention, and diversity to prevent the existing methods from over-relying on a single dataset or evaluation criteria. At the ethical level, a data privacy protection mechanism and a traceability framework for synthesized images should be established to prevent technical abuse.

In conclusion, the further development of face age synthesis needs to take into account the needs of technological innovation and application scenarios, and promote its reliable landing in the fields of security, medical care, and entertainment.

References

1. A. Lanitis, C. J. Taylor and T. F. Cootes, "Toward automatic simulation of aging effects on face images," IEEE Trans. Pattern Anal. Mach. Intell., pp. 442–455, 2002.
2. A. Kale and O. Altun, "Face age synthesis: A review on datasets, methods, and open research areas," Pattern Recognit., Art. no. 109791, 2023.
3. J. Suo, X. Chen, S. Shan, W. Gao and Q. Dai, "A concatenational graph evolution aging model," IEEE Trans. Pattern Anal. Mach. Intell., pp. 2083–2096, 2012.

4. I. Kemelmacher-Shlizerman, S. Suwajanakorn and S. M. Seitz, "Illumination-aware age progression," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, USA, 2014.
5. X. Chen, Y. Sun, X. Shu and Q. Li, "Attention-aware conditional generative adversarial networks for facial age synthesis," Neurocomputing, vol. 451, pp. 167–180, 2021.
6. M. Sun et al., "A unified framework for biphasic facial age translation with noisy-semantic guided generative adversarial networks," IEEE Trans. Inf. Forensics Secur., vol. 17, pp. 1513–1527, 2022.
7. T. K. Hsieh, T. J. Liu and K. H. Liu, "Facial age synthesis by StyleGAN2 with attention," in Proc. Int. Conf. Consumer Electron. Taiwan (ICCE-Taiwan), pp. 93–94, July 2024.
8. M. Falkenberg, A. B. Ottsen, M. Ibsen and C. Rathgeb, "Child face recognition at scale: synthetic data generation and performance benchmark," Front. Signal Process., vol. 4, Art. no. 1308505, 2024.
9. S. He et al., "Disentangled lifespan face synthesis," in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Montreal, QC, Canada, 2021.
10. J. C. Xie et al., "Diverse and lifespan facial age transformation synthesis with identity variation rationality metric," arXiv preprint, arXiv:2401.14036, 2024.
11. Z. Huang, J. Zhang and H. Shan, "When age-invariant face recognition meets face age synthesis: A multi-task learning framework and a new benchmark," IEEE Trans. Pattern Anal. Mach. Intell., vol. 45, no. 6, pp. 7917–7932, 2023.
12. P. Chandaliya and N. Nain, "ChildGAN: Face aging and rejuvenation to find missing children," Pattern Recognit., Art. no. 108761, 2022.