Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 3, No. 1, 2023

Cross-age face synthesis based on conditional adversarial autoencoder

Yifan Yan [*], Weiwei Yu

College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China

[*] **Corresponding author**: Yifan Yan (Email: 316218099@qq.com)

Abstract: Face aging aims to render face images with desired age attribute. It has tremendous impact to a wide-range of applications, e.g., criminal investigation, entertainment. The rapid development of generative adversarial networks (GANs) has shown impressive results in face aging. Among them, the Conditional Adversarial Autoencoder (CAAE) proposed in 2017 has achieved good results in face aging. However, the generated faces still have the problems that the aging features are not obvious and the identity information are not well maintained. In addition, research have shown that the human aging process is affected by genes. Different races have different external characteristics of aging. However, the current research does not take the racial factor into account, ignores the racial differences in the aging process. It affects the accuracy of transformation. To solve the above problems, this paper proposes a cross-age face synthesis based on conditional adversarial autoencoder: First, a conditional adversarial autoencoder is used as the infrastructure to build a cross-age face synthesis model based on race constraints. Secondly, the discriminator is composed of a discriminant network and a classification network, and a category loss function is designed to generate a real face that matches the target age; Finally, the model uses a identity feature extractor and a discriminator of the multi-scale architecture. Through multi-level discrimination from pixel values to high-level semantic information, the loss of personal identity features is minimized. UTKFace and MegaAge-Asian datasets are used in the experiment. Three comparative experiments are designed for the above improvements. The results show that the racial constraints make the generated images effectively maintain the racial characteristics, such as skin color and texture; The classification function of the discriminator improves the aging effect; The design of the multi-scale discriminator enables the generated face to have more stable local structure and identity information. Through qualitative and quantitative analysis, it is shown that this method has higher aging accuracy and identity retention than the CAAE method.

Keywords: Face aging; Generative adversarial networks; Conditional adversarial autoencoder; Deep learning; Multi-scale discriminator.

1. Introduction

Face aging is a task of synthesizing faces of a certain person under a given age. Age synthesis includes age progression (prediction of future looks)) and regression (estimation of previous looks). Cross age face synthesis technology has a wide range of applications. For example, it could be applied to help find lost children or to predict what someone will look like in the criminal investigation. In terms of film and television entertainment, it could help actors achieve cross-age performance and develop interesting commercial applications, etc. It is attracting more and more researchers' attention because of its various applications. However, the rigid requirement to the training and testing datasets makes face aging still an extremely challenging task. Most existing works require the training samples for a given person over a long range of years. In addition, they also require the query image to be labeled with the true age. The scarcity of datasets makes it difficult to achieve the optimal training.

Traditional face aging methods can be categorized into physical model-based approaches and prototype-based approaches. Physical model-based approaches model the shape and texture changes with age in terms of skin, muscle, and wrinkle, etc. Wu characterized the ageing degree by controlling the facial deformation and skin folds. Sadick explored the commonalities of aging by studying changes in eye skin and bones. Suo achieved face aging by building a facial muscle model. Although the above method can achieve face aging, it has two disadvantages: They require lots of training sample and are computationally expensive. Prototype-based approaches compute an average face. They divide training data by age into different groups. Each group is represented by the average face. The aging effect is realized by synthesizing the prototype face into the target face. Burt studied different age groups in 1995. Kemelmacher introduced optical flow to aging. By processing the prototype face, the final synthesized face texture is clearer. Wang designed models to capture intermediate states between adjacent age groups. The method still requires paired data for training. It is difficult to preserve person-specific information, which makes the synthetic faces look unrealistic.

Recently, Generative Adversarial Networks (GANs) based approaches have been achieved success in generating high quality images. Wang proposed an Identity-Preserved Conditional Generative Adversarial Networks (IPCGANs). IPCGANs introduced a perceptual loss to keep identity information. Zhang proposed a conditional adversarial autoencoder (CAAE) network. They assume that the face images lie on a high-dimensional manifold. Given a query face, we could find the face on the manifold. Stepping along the direction of age changing, we will obtain the face images of different ages and identity information will be retained. This method realizes age progression and regression at the same time. Before that, very few works focus on face regression. Simonyan K proposed a pyramid architecture of GANs (PA-GANs), emphasizing the importance of forehead and hair in aging effect. They design an age extractor and a discriminator of pyramid architecture to extract aging features.

The High-quality image generation model (HQGM) further improves the image quality with the Boundary Equilibrium Generative Adversarial Network (BEGAN). Cheng introduced the attention mechanism and feature loss network into the cross-age synthesis model. Zhang designed a progressive age transformation algorithm. Generative Adversarial network-based approaches have been promoted the work of face aging. However, aging characteristics still not obvious and identity information are not preserved well. In addition, the study [18-20] found that, affected by genes, diet, daily life and other factors, the degree of skin aging varies among different races. The current study does not take race into account. Ignoring the differences of skin aging between different races will inevitably lead to inaccurate age synthesis.

In this paper, we propose a cross-age face synthesis based on conditional adversarial autoencoder. The face is first mapped to a latent vector through a convolutional encoder. The latent vector preserves the identity information of the input face. Then the deconvolutional generator takes a vector, a age code and a ethnic code as its input and generates a face with the target age. Two discriminators are imposed on the encoder and generator forcing to generate more photo-realistic faces. To keep identity information, we introduce the pretrained VGG-FACE network in front of the discriminator to extract the identity-related features, and design a multi-scale discriminator to distinguish the input and output at multiple scales. Finally, the classification function of the discriminator is added to make the age conversion more accurate.

The benefit of the proposed model can be summarized from four aspects:

(1) We emphasize that different races have different aging characteristics. The proposed approach introduces racial constraints on the conditional adversarial autoencoder to achieve age changes while retaining racial characteristics. Finally, adding the classification function of the discriminator further improves the accuracy of age conversion.

(2) Identity-specific features are estimated by a discriminator at multiple scales, which retain the face identity information better and increase the face details.

(3) Adding the classification function of the discriminator further improves the accuracy of age conversion.

2. Related Work

2.1. Traversing on the Manifold

CAAE assumes the face images lie on a high-dimensional Manifold. Traversing along a certain direction on this manifold can realize age conversion while keeping the identity information unchanged. High-dimensional manifold is difficult to model. Therefore, CAAE maps the high-dimensional manifold to the low-dimensional space to manipulate. Faces are mapped to the latent space (low-dimensional space) which extracts the identity information. Concatenating with the age labels ,they are mapped to the manifold– generating a series of face images. Conditional adversarial autoencoder is used to learn this mapping. The whole framework is illustrated in Figure 1. CAAE is the combination of Adversarial Autoencoder and Conditional Generative Adversarial Nets (CGAN). Face x is mapped to latent space by encoder E to extract identity feature Z. Z is connected with the age label l, and mapped to the high-dimensional manifold by the generator G, generating a series of face images y. Discriminator d1 and d2 are imposed on the encoder and generator forcing to generate realistic faces.

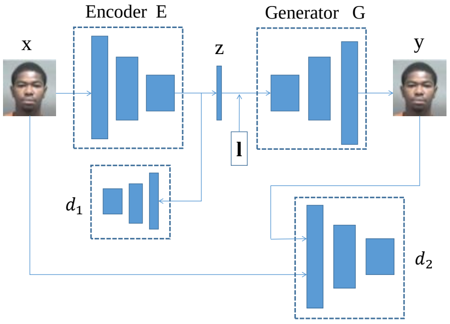

Fig. 1 Conditional adversarial autoencoder

2.2. Racial differences

Aging is an inevitable and complex process. Aging is affected by many factors, such as dietary habit and environment. Studies have found that skin effects are also influenced by genes. Aging characteristics vary across ethnic groups due to underlying structural and functional differences. Scholars measured and analyzed the skin color and color heterogeneity in four different ethnic groups living in the same environment (African-American, Caucasian, Chinese and Mexicans). The results showed that Caucasians had lower hue, which means more red skin. The Chinese volunteers showed signs of yellowing of their skin. African-American showed the least variation in skin color. Venkatesh S found differences in skin aging among different ethnic groups. The study analyzed changes in skin function and structure during aging in different ethnic groups. Studies have found that Caucasian skin is thinner, and the skin loses elasticity and wrinkles earlier in the individual aging process. On the contrary, African have more collagen in their skin, which can delay aging and keep the skin smooth and elastic. Asians are less prone to wrinkles in the early stages of aging. Vashi N A found that skin color has many characteristics that make the aging process unique. Asians, Hispanics, and African-American have different facial structures. Differences in epidermal melanin concentration make people with darker skin more prone to depigmentation. And thicker, denser dermis makes facial wrinkles less noticeable. There are large differences in skin between African, Asian, and Caucasian races. There are different aging characteristics between different races.

3. Approach

3.1. Overview

The whole framework of our networks is illustrated in Figure 2. It contains a encoder E, a generator G, two discriminator Dimg and Dz, a VGG-FACE network. Face is mapped to latent space by encoder E to extract identity feature. Identity feature is connected with the age label and race label.Then they mapped to the high-dimensional manifold by the generator G.Due to the lack of training data of the same individual with different ages, the age label is consistent with the age of the input face,and make the generated face as close as possible to the input face. Pretrained VGG-FACE network extract the identity-related features of the input face and generated face. These different levels of feature maps are put into the multi-scale discriminator Dimg to be distinguished and age estimation. The discriminator Dz guarantees a smooth transition of the generated faces.

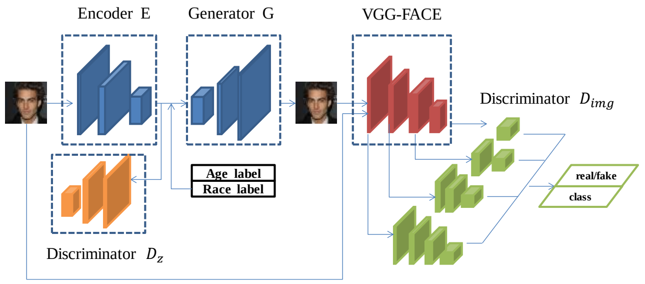

Fig. 2 Network framework

3.2. The encoder and generator networks

According to the characteristics of the skin, this paper divides the research objects into four types:(1) African-American (2) Asian, Latino, Hispanic; (3) Europe, North Africa and Southwest Asia; (4) Indian. Age is divided into ten groups. Both race information and age information are represented by one-hot encoding.

The age synthesis process is as follows: The input face images $\mathrm{X_{ijk}}$ are 224 × 224 RGB images. The output of encoder $\mathrm{E(X_{ijk}) = Z_i}$ preserves the high-level personal feature of the input face. The output face conditioned on certain age and race can be expressed by $\mathrm{G(Z_i,\ l_j, l_k) = \widehat{X}_{ijk}}$, where $\mathrm{l_j}$ denotes the one-hot age label, $\mathrm{l_k}$ denotes the one-hot race label. Let Convk represents a 5 × 5 convolution-Batchnorm-Relu layer with stride 2 and k output channels. The architecture of encoder is Conv64−Conv128− Conv256 − Conv512 − Conv512. The architecture of encoder is Conv64 −Conv128 − Conv256 − Conv512 − Conv512. Let deconvk represents a 4 × 4 deconvolution-Batchnorm-Relu layer with stride 2 and k output channels. The architecture of generator is deconv512 −deconv256 – deconv128 – deconv64 – deconv64 – deconv3. We inject the conditions before deconv512.

3.3. Multi-scale discriminator

The discriminator $\mathrm{D_{img}}$ receives the real image $\mathrm{X_{ijk}}$ and the synthetic image $\mathrm{\widehat{X}_{ijk}}$, and is responsible for distinguishing the two pictures. Natural images exhibit multi-scale characteristics. Inputting multi-scale feature maps into the discriminator can improve the ability to capture details. Therefore, the VGG-FACE network is introduced to extract different levels of identity features. VGG-FACE network is pre-trained on large datasets and has excellent performance in feature extraction. Along the hierarchical architecture, VGG-FACE captures different levels of identity information from pixel values to high-level semantics. Taking the outputs of the 2nd, 4th, 7th, and 10th convolutional layers, the final facial features are jointly estimated by the discriminator at the four scales. Compared with the original discrimination of a single image, the generated face is easier to distinguish from the real face. The local structure of the face is better maintained. $\mathrm{f_1, f_2, f_3, f_4}$ denote the outputs of the 2nd, 4th, 7th, and 10th convolutional layers.

Corresponding to the above feature maps, the discriminator is designed as a multi-scale architecture.The scores of four pathways are finally concatenated and jointly estimated by the discriminator Dimg . The detailed structure of Dimg is shown in Fig. 3. The size of the convolution kernel is 4×4. The stride is 2, followed by the BatchNorm layer, and the activation function is LeakyReLU.

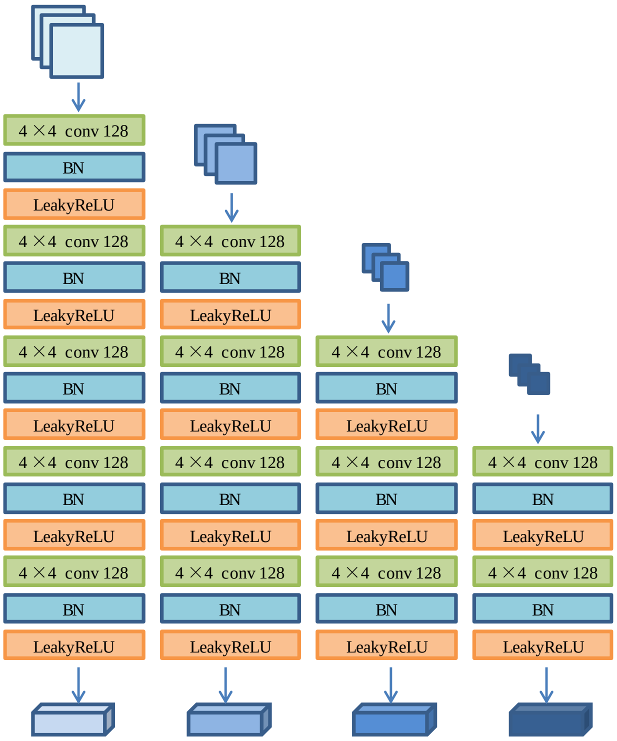

Fig. 3 Part of the $\mathrm{D_{img}}$ network framework

In addition to the four-layer feature maps, the input to the discriminator also includes age and ethnicity information. The feature maps and the conditional information are concatenated together and fed into the fully connected layer.

In addition to making the synthetic image as realistic as possible, we also want it to have the characteristics of the target age. Therefore, the function of the discriminator to classify the age of the picture is increased, and the category loss function is increased to force the synthetic face to have the texture features of the desired age. The discriminative part and the classification part share the network structure and parameters, and are only separated in the last layer, which are respectively mapped into two outputs: probabilities of the real data and age class. The Sigmoid function is used for discrimination and the Softmax function is used for classification. The number of categories is 10 (10 age groups).

3.4. Objective function

The input and output faces are expected to be similar. Train the encoder and generator with the L2 norm:

$$L_2 = minL(\mathbf{X_{ijk}}, G(E(\mathbf{X_{ijk}}), \mathbf{l_j}, \mathbf{l_k})) \tag{1}$$

In order to ensure smooth transition of faces, z is forced to obey a uniform distribution in the latent space through D_z . $p_{data}(x)$ denotes the distribution of the training data. $q(z|x)$ denotes the distribution of z. Assuming $p(z)$ is a prior distribution, and $z^* \sim p(z)$ denotes the random sampling process from $p(z)$. A objective function can be used to train E and D_z :

$$\min_{E} \max_{D_z} L_{E\,D_z} = E_{z^* \sim p(z)}[log D_z(z^*)] + E_{x \sim p_{data}(x)}[log(1 - D_z(E(x)))] \tag{2}$$

The discriminator helps synthesize realistic faces. D_{img} and G with condition can be trained by

$$E_{X_{ijk} \sim p_{data}(x)}[log(1 - D_{img}(G(E(\mathbf{X_{ijk}}), \mathbf{l_j}, \mathbf{l_k})))] \tag{3}$$

In addition to generating realistic face, it should also ensure that the age of the synthesized image is consistent with the age label. The discriminator D_{img} forces the generator to generate faces of the desired age group by age-classifying:

$$L_{age} = E_{X_{ijk}\sim p_{data}(x)}[\log P(C=j|\mathbf{X_{ijk}})] + E_{X_{ijk}\sim p_{data}(x)}[\log P(C=j|G(E(\mathbf{X_{ijk}}),\mathbf{l_j},\mathbf{l_k}))] \quad (4)$$

C denotes the age category.

L_{TV} denotes the total variation which is effective in removing the ghosting artifacts.

$$L_{TV} = TV(G(E(\mathbf{X_{ijk}}),\mathbf{l_j},\mathbf{l_k})) \quad (5)$$

Finally the objective function becomes:

$$L_{total} = \alpha L_2 + L_{GD_z} + L_{GD_{img}} + L_{age} + \beta L_{TV} \quad (6)$$

The coefficients α and β balance the smoothness and high resolution.

4. Experimental Evaluation

4.1. Dataset

The experiment selects the UTKFace dataset and the MegaAge-Asian dataset. UTKFace contains 23,600 face images, annotated with age and ethnicity information, ranging from 1 to 78 years old. Since there are relatively few Asian faces in the UTKFace dataset, this paper selects 20,000 face images in MegaAge-Asian as a supplement. Classify face images by age. Since the face changes greatly before the age of 20, it is a group every 5 years before the age of 20, and a group every 10 years after the age of 20. Divide the dataset into ten groups: 0-5, 6-10, 11-15, 16-20, 21-30, 31-40, 41-50, 51-60, 61-70, 70+.

4.2. Experimental environment

The graphics card model is NVIDIA's Tesla P100-PCIE, and the memory is 16GB. Programming language is Python. Epoch is 50, and batch size is 100. In this paper, the ADAM learning strategy is used to dynamically adjust the learning rate. Learning rate is 0.0002, beta1 is 0.5. α=100, β=10.

4.3. Qualitative comparison

In order to show that the race constraints, age loss, and multi-scale discrimination proposed in this paper are effective, three sets of comparative experiments are designed. Select people of different ages, genders, and races to conduct experiments. For each input image, generate the corresponding ten ages.

4.3.1. Racial Constraint

Adding race labels to training allows people of different races to maintain their own aging characteristics. Figure 4 is a comparative experiment of race restraint. The first line is the input image. The left column of images has race labels added during training, the right column is not. Observing the four sets of comparative experiments, it can be seen that the effect of adding ethnic labels is better: The skin color of the face in columns 1, 3, and 7 is closer to the input image; In addition, the faces in columns 1, 3, and 7 developed wrinkles earlier than those in columns 2, 4, and 8, which is consistent with research showing that Caucasians have earlier wrinkles; The facial skin in column 5 was smoother than in column 6, in line with the findings that African Americans are less prone to wrinkling. In general, the experimental results of adding ethnic labels during training better restore ethnic skin color, and the generation effect of skin texture is more reasonable.

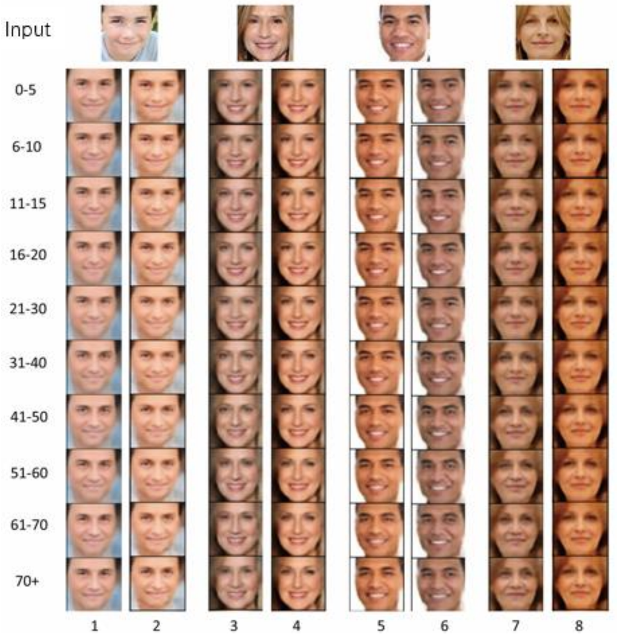

Fig. 4 Racial constraint comparison experiment

4.3.2. Age classification

The age classification function of the discriminator is added during training to make the generation effect of each age group more accurate. Figure 5 is age classification comparison experiment. The first row is the input image. The left column images are trained with a class loss added and the right column uses the original discriminator. Observing the four sets of comparative experiments, it can be seen that the effect of adding category loss is better: At the age of 0-5, the face in the left column image is more youthful than the right column image; After the age of 50, the left column has obvious eye bags, prominent cheekbones, obvious canthus wrinkles and nasolabial folds, the skin texture is clearer, and the aging effect is better. Compared with the right column, the aging characteristics are not obvious. it can be concluded that the addition of category loss can more accurately maintain the skin texture characteristics of each age group.

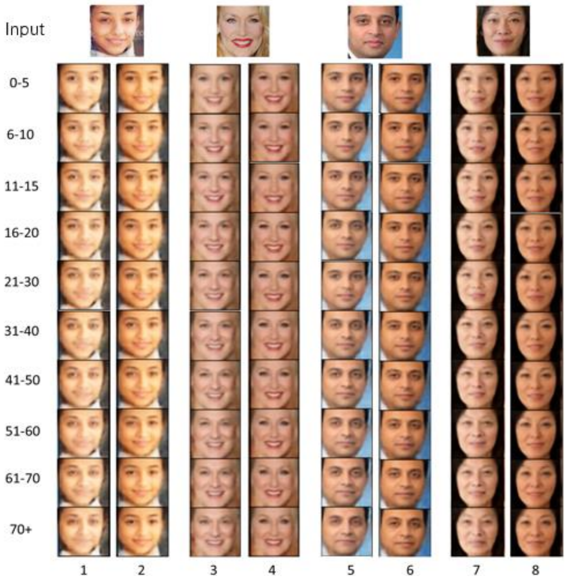

Fig. 5 Age classification comparison experiment

4.3.3. Multi-scale discriminator

The discriminator is designed as a multi-scale structure during training, and the facial features can be better captured by judging the feature maps of different levels. Figure 6 is a comparison experiment between the multi-scale structure discriminator and the ordinary discriminator. The first row is the input image. The multi-scale discriminator is used for training on the images in the left column, and the normal discriminator is used in the right column. Observing the four sets of comparative experiments, it can be seen that the effect of the multi-scale discriminator is better: The faces in columns 2 and 4 are rough and distorted, especially the red frame, while the faces in columns 1 and 3 are relatively smooth. The features of the 6th and 8th columns are blurred, especially the nose and facial texture, while the 5th and 7th columns have clear facial features and maintain more details. It can be concluded that the identification of multi-layer feature maps can better preserve identity information.

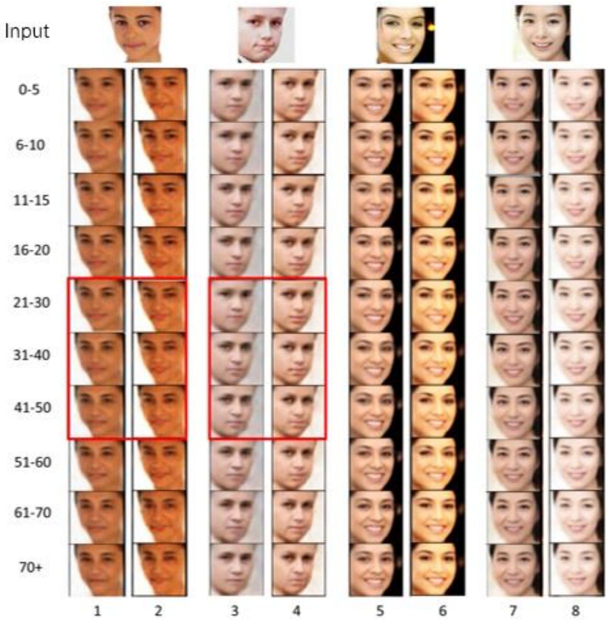

Fig. 6 Multi-scale discriminator comparison experiment

4.3.4. Compared with CAAE

The images in the left column are our method, and the right column is the CAAE method. Observing the four groups of comparative experiments, it can be seen that the method in this paper is more effective: In terms of ethnic characteristics, the aging characteristics in column 3 are not as obvious as those in columns 5 and 7, which is consistent with the fact that African-American skin contains more fibroblasts and wrinkles appear later, and Caucasian skin is thin and prone to wrinkles. , and the left column better restores the skin tone; In terms of aging effect, the left column has sagging face, prominent cheekbones, clear wrinkles, and more obvious aging characteristics, while the aging effect of the right column is not good, especially the fourth column has no obvious change; In terms of facial features, the first column has more details than the second column, and the facial features are more clear. Therefore, compared with the traditional method, the method in this paper retains the racial characteristics such as skin color and skin folds, the aging effect is more prominent, and the face details are more abundant.

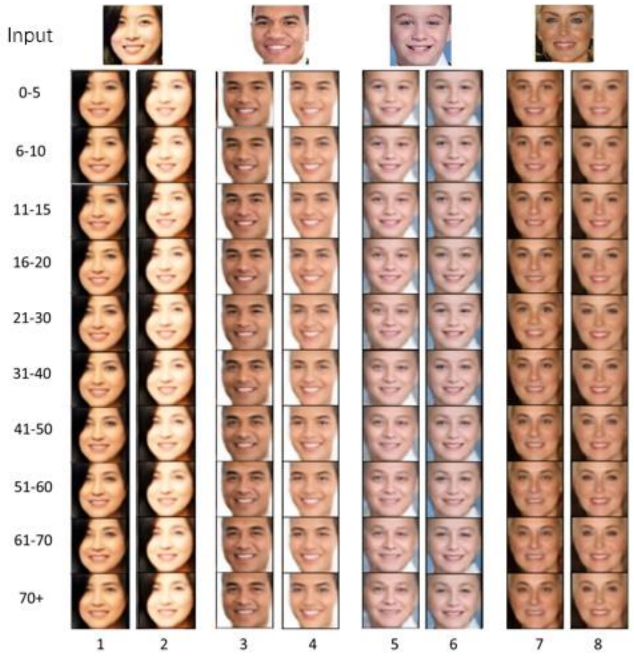

Fig. 7 Compared with CAAE

4.4. Quantitative comparison

In order to objectively evaluate the performance of our method, quantitative experiments are carried out with this method and the CAAE method. In cross-age face synthesis work, two important criteria are the accuracy of age synthesis and the retention of personal identity features. In terms of age accuracy, scholar Dong Liuyin chose the public API measurement of Face++, and this paper also chooses Face++ for evaluation. In terms of identity accuracy, the Cosine Similarity index and the Structural Similarity index are selected for evaluation.

4.4.1. Age estimation

Select 30 face images and generate their faces corresponding to ten age groups respectively. Estimate the mean absolute error (MAE) between actual effect and expected age. The lower the value of MAE, the more accurate the age synthesis. The face pictures obtained by the two methods are input into Face++ respectively, and the average value of the MAE index of ten age groups is obtained: The MAE value of the method in this paper is smaller than that of the CAAE method, indicating that the age synthesis effect of this method is more accurate.

4.4.2. Cosine Similarity

Cosine Similarity evaluates similarity by measuring the cosine of the angle between vectors A and B. Calculated as follows:

$$cos(\theta) = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}} \tag{7}$$

A_i and B_i represent the components of vector A and vector B. The closer the cosine similarity value is to 1, the higher the similarity between the two. Select 30 face images to be tested, calculate the cosine similarity between the input image and the face images generated by the two models, and average the obtained cosine similarity. The results are shown in Table 2. It can be seen that the cosine similarity index of the model proposed in this paper is higher than that of the CAAE model, indicating that this method has better identity information retention effect in cross-age face synthesis work.

4.4.3. Structural Similarity

Structural Similarity (SSIM) is an index to measure the similarity of image structure. SSIM measures the similarity between images from three aspects: structure, brightness, and contrast. The structural similarity between two images x and y is calculated as follows:

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \tag{8}$$

μ_x, μ_y represent the mean value of x and y, σ_x^2, σ_y^2 represent the variance of x and y, and σ_{xy} is the covariance of x and y. $C_1=(k_1L)^2, C_2=(k_2L)^2$ is a constant used to maintain stability. L represents the dynamic range of pixel values, $k_1=0.01, k_2=0.03$. The closer the SSIM value is to 1, the higher the structural similarity between the two images. Select 30 face images to be tested, calculate the structural similarity between the input image and the face images synthesized by different models, and finally calculate the average of the obtained structural similarities. The results are shown in Table 3. It can be seen that the structural similarity index of the model proposed in this paper is higher than that of the CAAE model, indicating that this method has advantages in maintaining the facial structure. Compared with the CAAE method, it can effectively avoid the distortion of the face and has a better identity. Information remains effective.

Table 1. Age estimation

age	0-5	5-10	10-15	15-20	20-30	30-40	40-50	50-60	60-70	70+	average
CAAE	9.77	8.43	5.57	3.73	3.63	2.77	2.07	3.27	7.97	11.2	5.841
ours	10	6.2	5.27	3.6	2.9	2.17	3.23	2.23	6.07	12.13	5.38

Table 2. Cosine Similarity

age	0-5	5-10	10-15	15-20	20-30	30-40	40-50	50-60	60-70	70+	average
CAAE	0.984	0.986	0.988	0.986	0.985	0.990	0.989	0.989	0.985	0.985	0.986
ours	0.989	0.990	0.989	0.987	0.990	0.990	0.993	0.991	0.990	0.987	0.990

Table 3. Structural Similarity

age	0-5	5-10	10-15	15-20	20-30	30-40	40-50	50-60	60-70	70+	average
CAAE	0.562	0.512	0.583	0.551	0.574	0.541	0.568	0.536	0.568	0.550	0.559
ours	0.654	0.664	0.684	0.678	0.593	0.660	0.604	0.614	0.604	0.625	0.638

5. Conclusion

Aiming at the problems that the current ageing work does not take ethnic factors into account, and the aging characteristics are not obvious and the identity characteristics are not well maintained, this paper proposes Cross-age face synthesis based on conditional adversarial autoencoder. The method uses conditional adversarial autoencoders as the basic architecture, firstly adding race as conditional information into training, taking into account the differences in the aging process of different ethnic groups. This improvement enables the transformed face to retain features such as race-specific skin tone and skin texture, and the aging effect is more accurate and reasonable. Secondly, the discriminator is composed of a discriminant network and a classification network, and an age-related category loss function is designed to further improve the aging effect. Finally, this paper introduces the VGG-FACE network to extract identity features at different levels, and designs the discriminator as a multi-scale architecture. Through multi-scale discrimination, the generated face shows superiority in maintaining local structure and has a clearer texture. The face synthesized by this method not only retains the ethnic characteristics, but also improves the aging effect and the preservation of identity information. In future work, one needs to consider gender differences in the aging process, and the other is to use a more effective loss function to further enrich face details.

References

[1] Wu Y, Kalra P, Moccozet L, et al. Simulating wrinkles and skin aging[J]. VisualComputer, 1999, 15(4):183-198.

[2] Sadick N S, Karcher C, Palmisano L. Cosmetic dermatology of the aging face[J]. Clinics in Dermatology, 2009, 27(3-supp-S).

[3] Suo J, Zhu S C, Shan S, et al. A compositional and dynamic model for face aging[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 32(3): 385-401.

[4] Burt D M, Perrett D I. Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information[J]. Proc Biol Sci, 1995, 259(1355):137-143.

[5] Kemelmacher-Shlizerman I , Suwajanakorn S , Seitz S M . Illumination-Aware Age Progression[J]. Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014.

[6] WANG W, CUI Z, Yan Y, et al. Recurrent Face Aging[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE Press, 2016: 2378-2386.

[7] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in neural information processing systems. 2014: 2672-2680.

[8] Wang Z, Tang X, Luo W, et al. Face Aging with Identity-Preserved Conditional Generative Adversarial Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7939-7947.

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

[10] Simonyan K, Zisserman A. Very deep convolutional net-works for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[11] SONG H Z,WU X J. High-quality image generation model for face aging / processing[J]. Journal of Image and Graphics. 2019,24(04):592-602. 10.19678/j.issn.1000-3428.0062018

[12] Berthelot D , Schumm T , Metz L . BEGAN: Boundary Equilibrium Generative Adversarial Networks[J]. arXiv, 2017.

[13] CHENG Z K,SUN R,SUN Q J,ZHANG X D. Cross-age face synthesis based on conditional adversarial autoencoder[J/OL]. Computer Engineering:1-13[2022-0428].

https://doi.org/10.19678/j.issn.1000-3428.0062018.

[14] ZHANG J Q. Face Aging and Rejuvenation Based on Generative Adversarial Network[D]. University of Electronic Science and Technology of China,2021.

[15] Qawaqneh Z, Mallouh A A, Barkana B D . Deep convolutional neural network for age estimation based on VGG-face model［EB/OL］.［20180423］.

[16] MAKHZANI A, SHLENS J, JAITLY N, et al. Adversarial autoencoders [EB/OL]. (2015-11-18)[2021-06-06]

[17] Mirza M, Osindero S. Conditional generative adversarial nets[J]. ar Xivpreprint ar Xiv:1411.1784, 2014.

[18] The effect of age on skin color and color heterogeneity in four ethnic groups[J]. Skin Research & Technology, 2010, 16(2):168-178.

[19] Venkatesh S , Maymone M , Vashi N A . Aging in Skin of Color[J]. Clinics in Dermatology, 2019, 37(4):351-357.

[20] Vashi NA, de Castro Maymone MB, Kundu RV. Aging Differences in Ethnic Skin. J Clin Aesthet Dermatol. 2016 Jan;9(1):31-8.

[21] Zhang Z, Song Y, Qi H. Age progression/regression by conditional adversarial autoencoder[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017:4352-4360

[22] Yunxuan Zhang, Li Liu, Cheng Li, Chen Change Loy. Quantifying Facial Age by Posterior of Age Comparisons, In British Machine Vision Conference (BMVC), 2017.

[23] KINGMA D P, BA J. ADAM: A method for stochastic optimization. [J]. Computer Science, 2014.

[24] DONG L Y. Research on the synthesis of face aging combined with age estimation[D]. Shanghai Normal University.2021

[25] Face++ research toolkit. https://www.faceplusplus.com/.