DOI: 10.1111/cgf.15146
Eurographics Symposium on Rendering 2024
E. Garces and E. Haines
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 43 (2024), Number 4

Realistic Facial Age Transformation with 3D Uplifting

X. Li[1,2] G. C. Guarnera[2,3] A. Lin[1,2] A. Ghosh[1,2]

[1]Imperial College London, UK
[2]Lumirithmic Ltd.
[3]University of York, UK

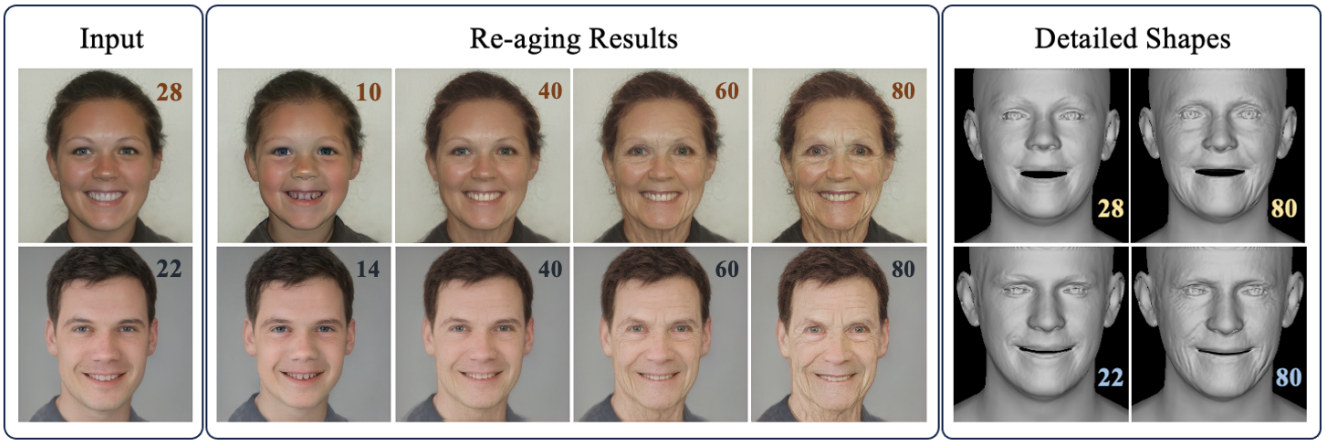


Figure 1: *Given a single input image (column 1), our method can produce realistic re-aging results (10 years old to 80 years old) in both facial appearances with skin color changes (column 2 – 5) and detailed shapes with high-frequency geometric changes, such as skin wrinkles (column 6 – 7).*

Abstract

While current facial re-ageing methods can produce realistic results, they purely focus on the 2D age transformation. In this work, we present an approach to transform the age of a person in both facial appearance and shape across different ages while preserving their identity. We employ an α-(de)blending diffusion network with an age-to-α transformation to generate coarse structure changes, such as wrinkles. Additionally, we edit biophysical skin properties, including melanin and hemoglobin, to simulate skin color changes, producing realistic re-ageing results from ages 10 to 80 years. We also propose a geometric neural network that alters the coarse scale facial geometry according to age, followed by a lightweight and efficient network that adds appropriate skin displacement on top of the coarse geometry. Both qualitative and quantitative comparisons show that our method outperforms current state-of-the-art approaches.

CCS Concepts

• ***Computing methodologies*** → ***Texturing;*** *Reflectance modeling; Shape modeling;*

1. Introduction

Facial age editing has a long history in computer graphics and vision due to its various applications. Recent films, such as *Indiana Jones and the Dial of Destiny* (2023), apply digital age transformation to the actors to represent the past, present, and future; for example, they de-aged the actor Harrison Ford from 80 to 40 years old. This has stimulated facial age transformation to become a popular and significant research direction.

Several challenges exist in this field. It is impossible to find a large number of paired facial photos spanning a long period. Consequently, some studies employ image-to-image translation techniques, such as age-conditioned Generative Adversarial Networks (GANs), to convert faces from source ages to target ages [OESF*20, HKSC19, LLS19]. While these methods can

produce satisfactory re-aging results, they often struggle to preserve identity, especially in transitions from childhood to adulthood. Hence, many of these approaches focus only on age transformation within adulthood [YPN*20, ZCS*22].

Another significant challenge is generating plausible 3D shape reconstructions with detailed face re-aging effects, such as wrinkles. Although some works successfully reconstruct high-quality 3D coarse shapes using model-free regression approaches [BV99], statistics facial models [CWZ*14, LBB*17, PKA*09], or geometric neural networks [RBSB18, ZBT22], they often overlook age as a controllable parameter for head size.

Moreover, simulating changes in optical skin properties due to ageing remains challenging. Most related works in computer vision and graphics focus on structural changes, such as producing wrinkles [OESF*20, HKSC19, LLS19, YPN*20, ZCS*22], while neglecting changes in skin properties. For instance, concentrations of principal chromophores (melanin and hemoglobin) typically decrease with age, resulting in lighter and paler skin. In contrast, [IGAJG15] proposes a comprehensive biophysical skin model to simulate the optical effects of ageing, which informs our work. However, their method does not include structural changes in skin.

In this work, we aim to transform the facial age of a person in both facial appearance and shape across different ages while preserving their identity. To edit facial appearances, we first employ a skin model to simulate skin color changes by adjusting the concentrations of the two primary chromophores (melanin and hemoglobin). We further employ an α-(de)blending diffusion network [HBC23] to edit faces within a paired dataset from ages 10 to 80 years, controlling the parameter α. However, the relationship between age and α is non-linear; hence, we introduce an age-α transformation method to achieve more plausible re-aging results.

For the age-dependent transformation of facial geometry, we build upon part of the trained DECA [FFBB21] to capture accurate latent space information, such as camera, lighting, and facial geometry, from a single input image. The geometry information can be fed into FLAME [LBB*17] to produce high-quality 3D facial reconstructions. However, current methods related to 3D reconstruction generally ignore the effects of age. Therefore, we employ a geometric neural network to incorporate the age factor into the FLAME model, thus altering the coarse scale facial geometry according to age. This network is complemented by a refine network that adds age-appropriate skin displacement on top of the coarse geometry. Compared to some methods with heavy neural networks, our network structure is simple and efficient, enabling the production of age-appropriate facial geometry, including fine-scale wrinkles. Qualitative and quantitative evaluations show that our method outperforms current state-of-the-art methods.

The main contributions of this paper are: (i) a diffusion-based facial age editing network that effectively preserves identity; (ii) a biophysical skin editing approach that simulates age-related skin color changes; (iii) an age-conditioned geometric network that can adjust the head size for children; (iv) a lightweight and efficient refine network that produces realistic high-frequency geometric changes, such as skin wrinkles.

2. Related Work

Here, we give a brief overview of two related topics: face age transformation and reconstruction.

Face age transformation Face re-aging is one of the most challenging face manipulation tasks due to missing datasets of real people spanning their whole lives. Additionally, the apparent age of a face can differ from the actual chronological age due to various factors such as genetics, lifestyle, environmental influences, and the presence of makeup. Early works in this field focus on building a physics-based model [LTC02, SCS*12] or a prototype-based method [KSSS14, TGMM12, TBP01] to simulate aging effects and facial details, such as wrinkles.

More recent approaches are mostly data-driven methods using deep learning techniques. Some employ age-conditioned Generative Adversarial Networks(GAN)-based [GPAM*20] methods to generate re-aged images by editing latent representation [OESF*20, HKSC19, LLS19]. While these approaches can produce realistic results, it is difficult for them to preserve identity information. To address this problem, some work design a network containing an Encoder and a generative Decoder [HLY*21, GTLMC22] or directly apply an Encoder-Decoder-based network [YPN*20, APCO21]. These methods can produce high-quality re-ageing results. However, they overlook the optical changes in skin due to ageing.

In recent years, several biophysical skin appearance models have been proposed [DJ06, JSB*10, GGD*20, AXX*23, LGLG24]. These methods mainly focus on simulating accurate skin appearance using practical models, but do not account for ageing changes. In contrast, [IGAJG15] proposes a complex biophysical skin model to simulate appearance changes due to ageing, yet it does not incorporate any structural changes.

Aside from appearance changes, human faces undergo changes in shape as they age. Therefore, some research, such as [HLY*21], also focuses on extracting shape features. However, these studies primarily concentrate on 2D image editing. In contrast, our work performs face re-ageing by considering both the 2D image space and 3D head shape. Furthermore, unlike other approaches that categorize ages into specific groups, our method allows for the editing of faces on a continuous scale to achieve any target age within the extensive range of $10-80$ years old.

Face reconstruction There are many excellent 3D face reconstruction methods. Several works apply model-free approaches to regress 3D shapes, such as 3D Morphable Model (3DMM) [BV99]. There are also some well-known statistical models, such as FaceWarehouse [CWZ*14], Basel Face Model (BFM) [PKA*09], and FLAME [LBB*17]. Recently, several works have relied on deep learning methods to reconstruct 3D mesh, such as COMA using convolutional mesh autoencoder [RBSB18] and MICA [ZBT22]. Unlike our work, these works focus on reconstructing a good coarse shape, but not on re-ageing faces. Therefore, they do not incorporate an age factor to control the head shape.

After obtaining the coarse shape, some works further add refinement methods to obtain more detailed shape features. Some employ or capture a 3D high-quality face dataset, and then compute facial details map from images [CCZ*19, GZC*17, LMG*20, THM*17]. On the other hand, some methods design neural networks to predict facial detail maps, such as displacement maps [FFBB21]. More

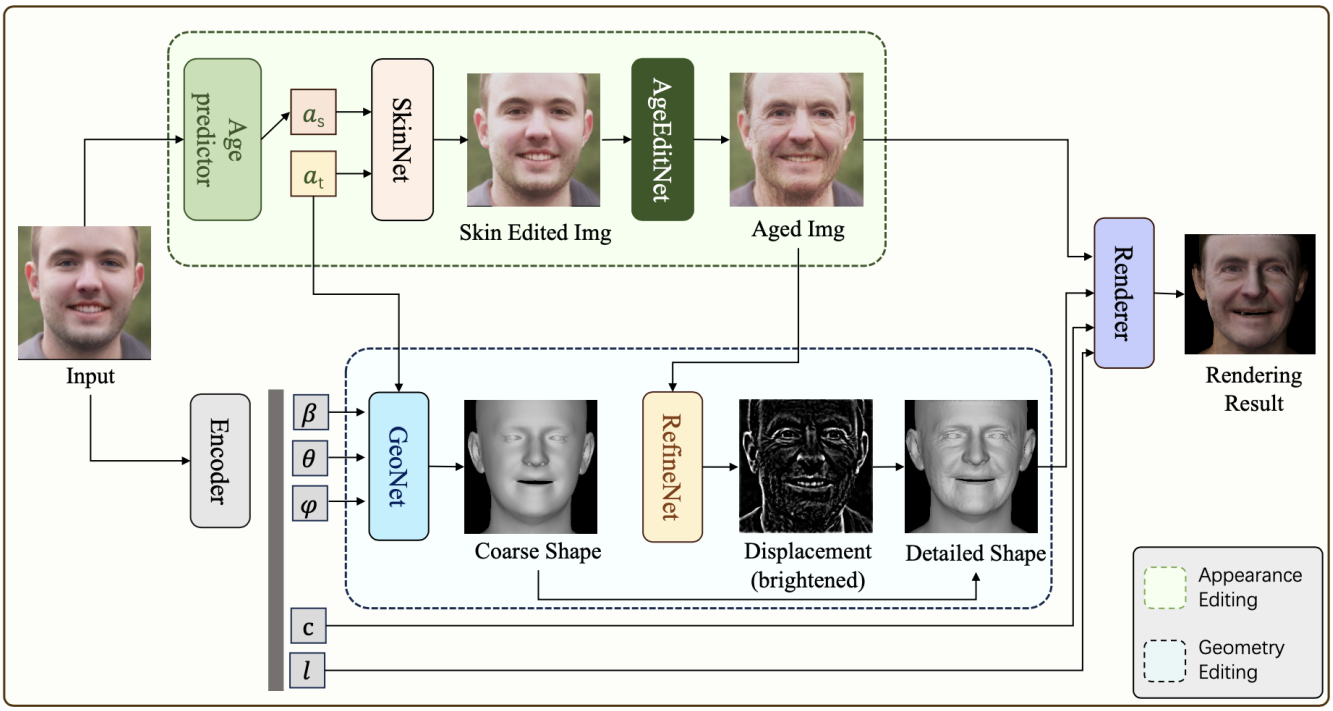


Figure 2: *Overview. Starting with a single image I as input, our method first employs an age predictor [RTG15], to estimate the source age a_s. With a given target age a_t, SkinNet edits the biophysical skin appearance due to ageing based on a skin model while AgeEditNet changes the structural aging effects iteratively from the source age a_s to the target ages a_t. Subsequently, we use the trained Encoder of DECA [FFBB21] to predict lighting, camera, and shape information. The next step involves GeoNet, which processes shape (β), expression (ψ), pose (θ), and age (a) parameters to reconstruct a coarse shape. This is followed by RefineNet, a lightweight network designed to predict a displacement map. Finally, we render the resulting image using the re-aged texture and detailed shape, incorporating the camera (c) and lighting (l) parameters. Here, we brighten the displacement map to show the details more clearly.*

related to our work, DECA [FFBB21] presents a state-of-the-art method to reconstruct an animatable detailed 3D Face Model from a single in-the-wild image. Furthermore, Danecek et al. extended DECA for emotion reconstruction with a novel deep perceptual emotion consistency loss [DBB22]. We follow a similar strategy to obtain detailed shape reconstruction, but our refinement network is lightweight, consisting of only two convolutional layers.

3. Method: 3DFAT

We propose 3DFAT, a novel method to generate a 3D age-editable model from a single image. We leverage the trained *Encoder* of *DECA* [FFBB21] to extract lighting, camera, and shape information from a single image. For texture editing, we first apply a simple biophysical skin model to simulate changes in skin albedo due to ageing, followed by a novel diffusion network, *AgeEditNet*, which alters the age of the input image (Section 3.1). In the 3D shape reconstruction phase, our novel geometric network, *GeoNet*, uses shape, expression, pose, and age factors to reconstruct a coarse shape. This is followed by a lightweight network, *RefineNet*, which predicts a displacement map to achieve a detailed shape (Section 3.2). Finally, using the same camera and lighting settings, we render the face with the re-aged texture and the detailed shape.

3.1. Texture Editing

3.1.1. Biophysical Skin Editing

Ageing causes coarse structure changes like wrinkles, as well as changes to optical properties due to chromophores such as melanin and hemoglobin, which primarily affect skin color. In recent years, several biophysically-based spectral skin reflectance models have been proposed [DJ06, JSB*10, GGD*20, AXX*23, LGLG24].

In this work, we employ the recently proposed skin model

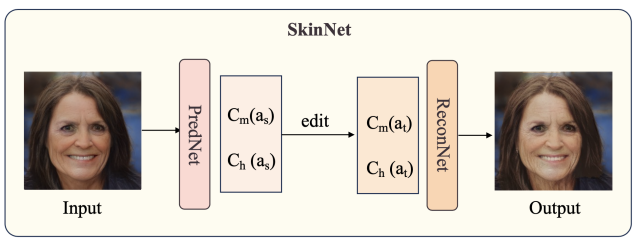


Figure 3: ***Biophysical skin editing.*** *The SkinNet consists of PredNet and ReconNet. Given an input image, PredNet predicts chromophore concentration, which is further edited and fed into ReconNet to reconstruct re-aged skin appearance.*

of [LGLG24], which is based on diffusion theory and includes five choromophore parameters. This model has been demonstrated to reproduce skin appearance with high accuracy. We simulate skin changes due to ageing by adjusting melanin and hemoglobin concentrations (C_m and C_h). Figure 3 illustrates the network structure of *SkinNet*, which consists of *PredNet* and *ReconNet*. *PredNet* predicts the concentrations of the two chromophores, and *ReconNet* reconstructs skin appearance from chromophores concentrations. Both networks are Multilayer Perceptrons (MLPs).

Between these two networks, we compute the new chromophore parameters by taking into account the source age a_s and target age a_t as show in the following equation:

$$\begin{aligned} C_m(a_t) &= (1 + \frac{k_m(a_s - a_t)}{10}) \times C_m(a_s), \\ C_h(a_t) &= (1 + \frac{k_h(a_s - a_t)}{10}) \times C_h(a_s) \end{aligned} \tag{1}$$

where $k_m = -0.08$ and $k_h = -0.06$ represent the decrease in concentration of each chromophore per decade as a function of age [IGAJG15]. An ablation study for this step is reported in Section 4.5.

Training Details: The training dataset for these two networks is the augmented Lookup Table generated by the skin appearance model [LGLG24]. This Lookup table covers a wide number of skin colors, allowing for the simulation of appearances across different ethnicities and genders. The training process is fast and efficient since we do not train on a large image dataset. The testing dataset is the same image dataset used by other networks in our pipeline.

The loss function for both networks is the Mean Squared Error(MSE). The *PredNet* and *ReconNet* are separately optimized using Adam solver with a weight decay of 1e-6 and learning rates of 3e-4.

3.1.2. Ageing Structural Editing

After editing skin properties, we employ *AgeEditNet*, a denoising-diffusion network as outlined in [HBC23], to iteratively edit the age in images. The α-blending diffusion method, known for its simplicity and efficiency, serves as a robust denoising diffusion framework, with several powerful applications. In fact, it enables the mapping between arbitrary image densities. Using this network architecture (Figure 4), we achieve the mapping of facial features from source ages to target ages by adjusting the parameter α.

Age-alpha transformation: However, the relationship between the parameter α and age a may not be linear. To address this issue, we use a pre-trained age predictor [RTG15], to estimate the age of subjects in images in the training dataset at each iteration. We then model this relationship using polynomial regression, represented as $\alpha = \Theta(a)$. This allows for the accurate prediction of actual input values for both the source and target ages. For further details on this approach, please refer to the ablation study in Section 4.5.

Age editing method: Using the age-transformation parameter α, our network, denoted as D, is capable of iteratively editing faces to achieve a target age, α_t, starting from the source face age, α_s. The sampling algorithm is detailed below.

Generating facial dataset $'Lifespan\ Dataset'$ with paired ages: We selected the FFHQ-Aging dataset [OESF*20], an extension of the Nvidia FFHQ dataset [KLA21], which includes

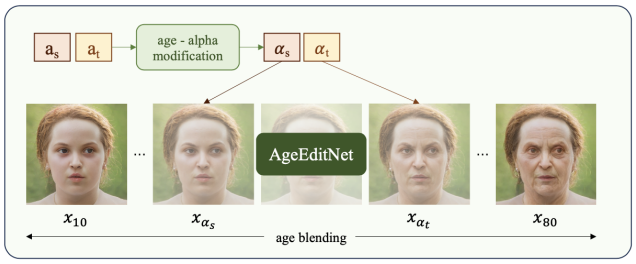


Figure 4: *Ageing structural editing. AgeEditNet is an α-blending diffusion network that re-ages the image from the source age a_s to the target age a_t with age-α modification.*

Age editing algorithm

Require: I_s, $\alpha_s := \frac{s}{T}$, $\alpha_t := \frac{t}{T}$
  if $\alpha_t > \alpha_s$ **do**
    for $i = s, ..., t$ **do**
      $I_{\alpha_{i+1}} = I_{\alpha_i} + D(x_{\alpha_i}, \alpha_i)$
  else do
    for $i = t, ..., s$ **do**
      $I_{\alpha_{i+1}} = I_{\alpha_i} - D(x_{\alpha_i}, \alpha_i)$

70,000 images with age and gender information. To create a paired datasets, essential for our work, we follow a similar approach to FRAN [ZCS*22], by applying SAM [APCO21] to the FFHQ-Aging dataset. This process generates a high-quality *Lifespan Dataset* with paired facial images of individuals aged from 10 to 80 years old.

Training Details: Our network's training requires paired synthetic data of individuals aged 10 and 80 years old, along with the original images. The input data consist of images representing a 10 years old age state, while the reference data are images corresponding to an 80 years old state. Additionally, we use predictions of images with the real age to compare with the original images, ensuring the network's quality. The loss function used in this process is defined as follows:

$$\mathcal{L}_{age} = ||O_{80} - I_{80}||_2 + ||O_{\alpha_r} - I_{\alpha_r}||_2 \tag{2}$$

where O_{80} and I_{80} represent the predicted and synthetic images at 80 years old, respectively. Similarly, O_{α_r} and I_{α_r} denote the predicted and reference images corresponding to the actual age, respectively.

To optimize this loss function, we use the Adam optimizer with a learning rate of $4e-3$; the training of *AgeEditNet* is conducted over 200 epochs.

3.2. Shape Reconstruction

3.2.1. Coarse Shape Reconstruction

FLAME is a statistical 3D face model capable of reconstructing high-quality 3D coarse shapes [LBB*17]. However, it is limited in its ability to modify the head shape in accordance with age changes. To address this issue, we incorporate the age factor along with shape, expression, and pose information into the geometry network

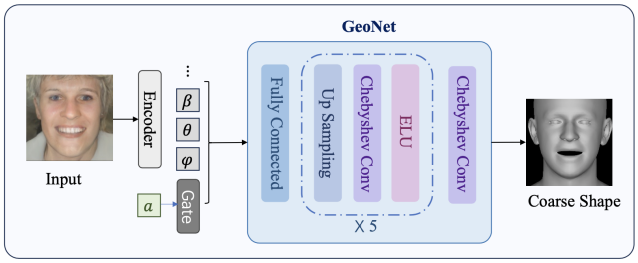


Figure 5: ***Coarse shape reconstruction.*** *Given the shape (β), expression (ψ), pose (θ) parameters predicted by the pre-trained Encoder, along with an age parameter a, GeoNet reconstructs a coarse shape of the input image.*

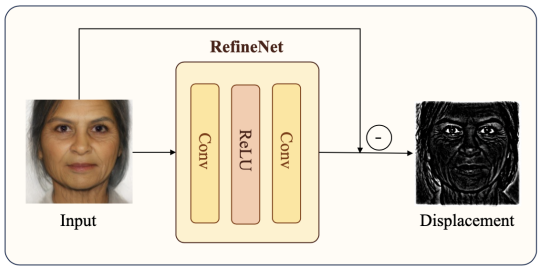


Figure 6: ***Detailed shape reconstruction.*** *RefineNet extracts the high-frequency information from the input image as a displacement map.*

GeoNet. This allows for the reconstruction of a 3D coarse shape that can be altered with age. For more details on the impact of the age factor, please refer to the ablation study in Section 4.5.

GeoNet is a convolutional mesh decoder, similiar to the one described in [RBSB18]. It consists of a fully connected layer that maps input facial data, followed by five blocks and a Chebyshev Convolution layer [DBV16]. Each block includes an Up-Sampling layer, a Chebyshev Convolution layer with $K = 3$ Chebyshev polynomials, and an Exponential Linear Unit(ELU) [CUH16] activation function. Regarding the age parameter, we use a gating mechanism for adults to ensure the head size remains constant during adulthood. The architecture of this part is shown in Figure 5.

Datasets: Reconstructing a high-quality 3D shape from a single input image presents significant challenges. To address this, we use the Facescape dataset [YZW*20], which multi-view images of 359 subjects with 20 expressions and three different views (front, left, and right). This dataset helps us in reconstructing accurate 3D coarse shapes. Additionally, we employ *Lifespan Dataset* for training. However, instead of using three different views, we use synthetic images representing three distinct ages. Training on this mixed dataset forces the network to maintain age consistency while keeping pose and expression factors constant.

Training details: Leveraging the trained *Encoder* of DECA, we accurately predict shape (β), expression (ψ), pose (θ), camera (c), and lighting (l) parameters. The first three parameters form a suitable latent space for the FLAME model [LBB*17], while the last two are used for rendering. The total loss function is

$$\mathcal{L}_{shape\_c} = \lambda_v \mathcal{L}_v + \lambda_{pho} \mathcal{L}_{pho} + \lambda_{age} \mathcal{L}_{age} + \lambda_{reg} \mathcal{L}_{reg} \quad (3)$$

where $\mathcal{L}_v$ is the vertices loss, $\mathcal{L}_{pho}$ is the photometric loss, and $\mathcal{L}_{age}$ is the age prediction loss; each loss component is weighted by a factor λ_x

• **Vertices loss:** it computes the difference between the predicted vertices $\widetilde{V}$ from *GeoNet* and those V from FLAME, as $\mathcal{L}_v = ||\widetilde{V} - V||_1$.

• **Photometric loss:** defined as $\mathcal{L}_{pho} = ||M \odot (\widetilde{R} - R)||_2$, where M is the face mask. R are the reference images, which vary based on the dataset source, ensuring shape consistency and improved re-aged results. $\widetilde{R}$ are the rendering results given by $\mathcal{R}(G,S,c)$ where G is the coarse geometry, S is the shaded face computed as

$$S(T,l,N)_{i,j} = T_{i,j} \odot \sum_{k=1}^{9} l_k H_k(N_{i,j}) \quad (4)$$

where T is diffuse texture, l and H are the SH coefficients and basis [RH01], and N is the normal map of pixel (i,j) in the UV coordinate space.

• **Age prediction loss:** $\mathcal{L}_{age} = ||\widetilde{a} - a||_2$, measures the accuracy of predicted ages a against reference ages $\widetilde{a}$. The same age predictor as in the age editing part is used.

• **Regularization:** to regularize the predicted mesh we use $\mathcal{L}_{reg} = \mathcal{L}_{normal} + \mathcal{L}_{laplacian}$, where $\mathcal{L}_{normal}$ is the mesh normal consistency, computed for each pair of neighboring faces and $\mathcal{L}_{laplacian}$ for mesh smoothing [NISA06].

We train this network using the Adam optimizer with a learning rate of $1e-4$ and a learning rate decay of 0.99 every epoch.

3.2.2. Detailed Shape Reconstruction

A primary challenge in reconstructing a 3D ageing model is the generation of facial details, such as wrinkles. The coarse shapes produced by *GeoNet* contain low- and middle-frequency information about faces. Early work in this area often use some low-pass filters like Difference of Gaussians (DoG) to blur the images, then extracting high-frequency details by subtracting the blurred image [CBZB15]. While these methods produce high-quality results, they require the selection of user-defined parameters. To overcome this limitation, we propose *RefineNet*, a lightweight network using two convolutional layers. We initialize the network's weights with a Gaussian filter ($\mu = 0$, $\sigma = 3$), thus eliminating the need for manual parameters tuning and reducing both training time and complexity compared to deeper networks [FFBB21]. The architecture of *RefineNet* is shown in Figure 6, where the displacement map is brightened for clarity.

Training details: Given that our datasets consists only of 2D images, the loss function for this network compares rendering results with reference images. Inspired by [THM*17], we incorporate the 2D gradient of predictions and references into the loss function, to retain high-frequency details while suppressing noise. The Loss function is

$$\mathcal{L}_{shape\_d} = \lambda_{pho} \mathcal{L}_{pho} + \lambda_{dx} \mathcal{L}_{dx} + \lambda_{dy} \mathcal{L}_{dy} \quad (5)$$

with photometric loss $\mathcal{L}_{pho}$ and derivative losses $\mathcal{L}_{dx}$ and $\mathcal{L}_{dy}$ on the x and y axis, respectively.

• Photometric loss: After extracting the displacement map D, we convert the coarse shape G_c to normal map N_c. We then obtain

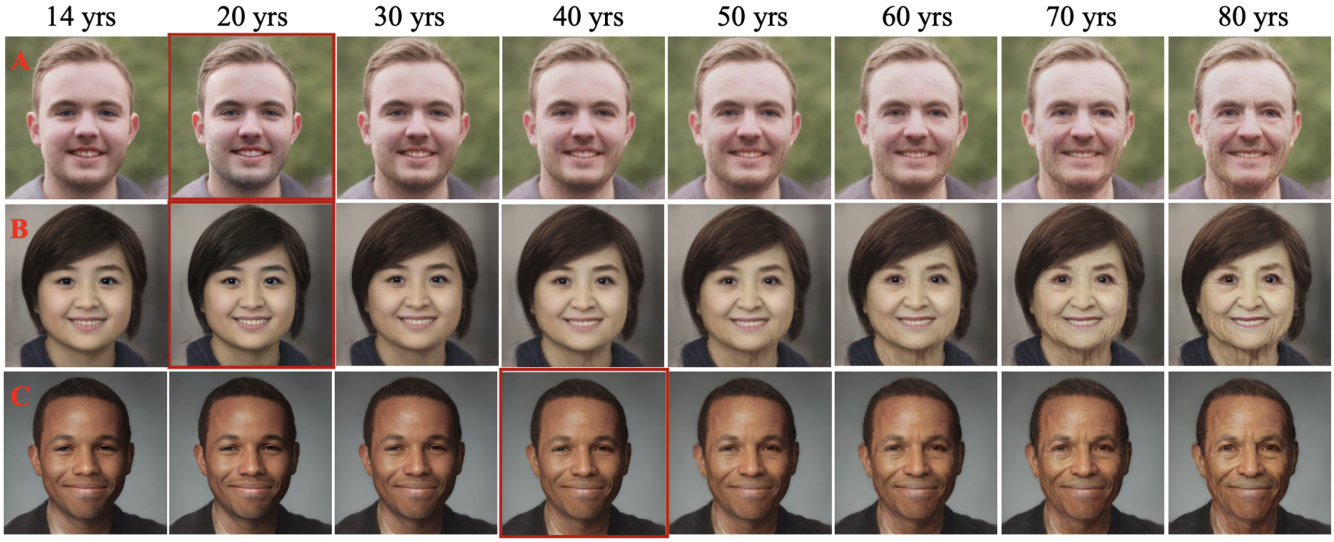


Figure 7: *Age transformation results from 14 to 80 years old. The input images are outlined with a thin red border. Our methods generates realistic facial structural changes, such as wrinkles, with good identity preservation, as well as skin color changes due to ageing.*

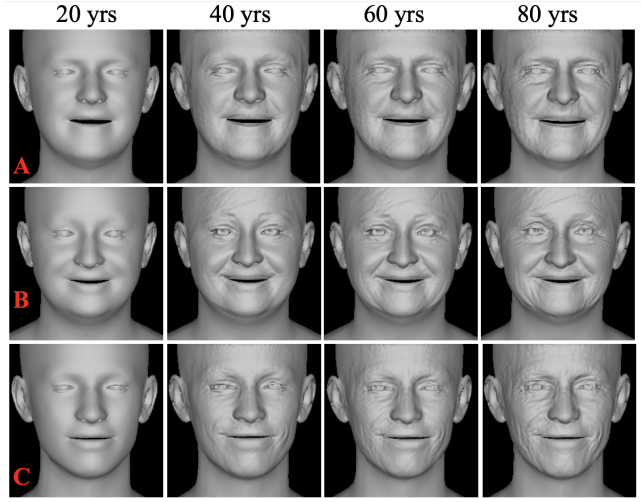


Figure 8: *Coarse shape reconstructions with 20 years old and detailed shape reconstructions of 40, 60, and 80 years old.*

the detailed shape G_d as

$$G_d = G_c + D \odot N_c \tag{6}$$

We convert G_d to the detailed normal map N_d, compute detailed rendering $\mathcal{R}(G_d, S(T,l,N_d),c)$, and then compute the photometric loss with the same equation as in Section 3.2.1.

• **Derivative loss:** Following the approach in [THM*17], the derivative loss is calculated as

$$\mathcal{L}_{dx} = ||\frac{\partial R}{\partial x} - \frac{\partial \widetilde{R}}{\partial x}||_1, \quad \mathcal{L}_{dy} = ||\frac{\partial R}{\partial y} - \frac{\partial \widetilde{R}}{\partial y}||_1 \tag{7}$$

We train this network using an Adam optimizer with a learning rate of $1e-5$ and a 0.99 decay rate per epoch.

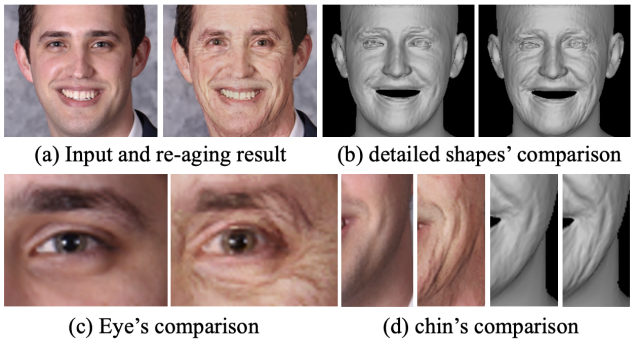


Figure 9: *Comparison of the input (20 years old) and re-ageing result (60 years old) and detailed comparison of the right eye and chin in both facial appearance and detailed shape.*

4. Results and Evaluation

In this section, we present the results of our work (Section 4.1), qualitative and quantitative evaluation (Section 4.2 and Section 4.3), a user study (Section 4.4), ablation study (Section 4.5), and limitations (Section 4.6).

4.1. Results

Here, we sequentially display the results of 2D re-ageing results and 3D coarse and detailed shape reconstruction of our work.

We present the lifespan changes of the whole face of three subjects (A, B, C) from 14 to 80 years old in Figure 7. As can be seen, our method produces realistic structural changes, such as wrinkles, and performs well in identity preservation. Furthermore, the results also illustrate the skin color changes due to ageing.

Figure 8 displays the corresponding coarse and detailed shape

Figure 10: *An example of a round-trip transformation, aging a face from 22 years old (a) to 62 years old (b), then de-aging back to 22 years old (c).*

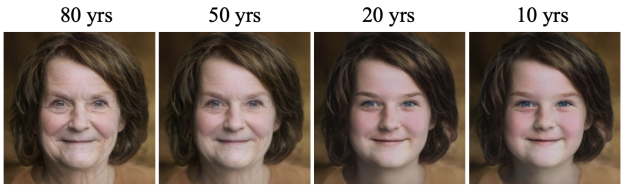


Figure 11: *An extreme example of a subject from 80 to 10 years old.*

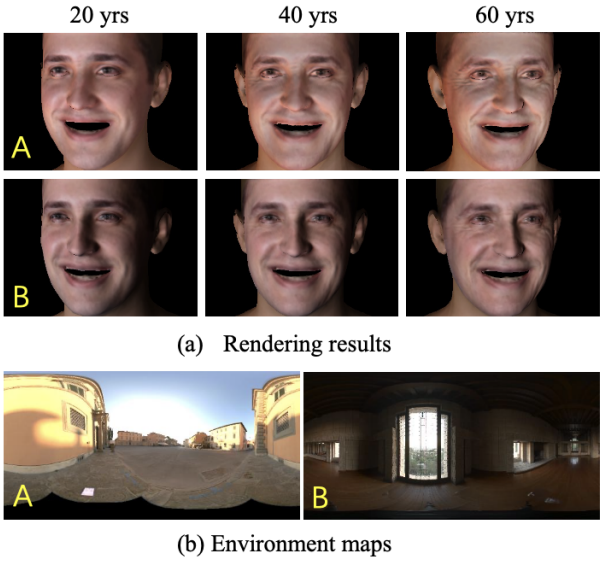


Figure 12: *Multi-view rendering results of 20, 40, 60 years old with two environment maps.*

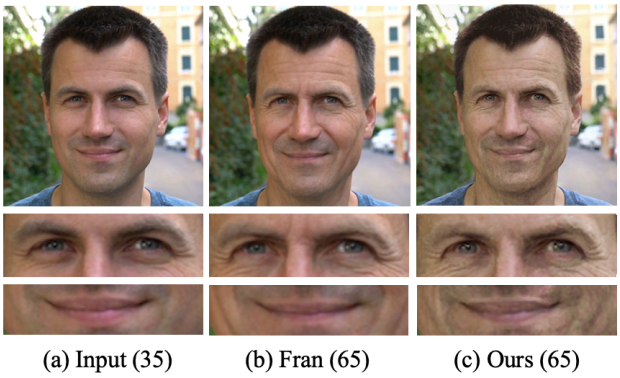


Figure 13: *Comparison of re-aged results using FRAN [ZCS*22] and our method. Input image from FRAN [ZCS*22].*

reconstructions for ages 40, 60, and 80 years. As shown, *GeoNet* generates high-quality reconstructions from a single input image, followed by *RefineNet* which produces detailed shapes, including a wrinkle map. To highlight the details, we show a comparison of a subject between ages 20 and 60, focusing on both facial appearance and detailed shape, including cropped areas the right eye and chin, as shown in Figure 9.

Our method generates realistic ageing effects, altering facial details and subtly changing eye color in addition to skin color. Figure 10 shows the process of ageing the face from 22 years old (a) to 62 years old (b), then de-ageing it back to 22 years old (c). As can be seen, our method is robust and consistent for both the ageing and de-ageing processes. We also show an extreme example of age reversal from 80 to 10 years old in Figure 11.

These examples demonstrate that our method effectively handles subjects of various ages, ethnicities, and genders, producing realistic re-ageing results with good identity preservation. Please see more comparisons and the ablation study of the age factor in Section 4.5.

Figure 12 shows rendering results of a face after age transformation and 3D upliting to obtain detailed shape. Here, we employ a simple normal-based rendering technique using the Phong model to relight faces under different environment maps. We render the faces of a subject of 20, 40, 60 years old with diffuse albedo obtained from the transformed 2D texture and shading according to the transformed geometry; for specular roughness and specular reflectance we choose empirical values that are constant over the face.

4.2. Qualitative Evaluation

To the best of our knowledge, there are no public 3D re-ageing methods in computer graphics and vision, therefore we compare 2D rendering results with recent state-of-the-art re-ageing methods, namely SAM [APCO21], LATS [OESF*20], HRFAE [YPN*20], FRAN [ZCS*22], as well as 3D detailed shape reconstruction methods close to us, namely DECA [FFBB21] and EMOCA [DBB22]. Please note that for 3D comparison, those methods can only reconstruct 3D shapes without age editing.

Age transformation: FRAN is a recent re-ageing approach that demonstrates good identity preservation [ZCS*22]. As their code is not publicly available, we borrow an image from their paper which is representative of their best result to compare with us. As can be seen in Figure 13, our method performs slightly better in identity preservation, such as mouth shape (row 3), and produces more plausible facial details, such as a grayer eye color for the old subject (row 2), as well more realistic lightening of skin color with age. Additionally, FRAN manipulates facial ages from 18 to 80 years old, which is narrower range than our method.

Figure 14 shows the results of several recent methods for two

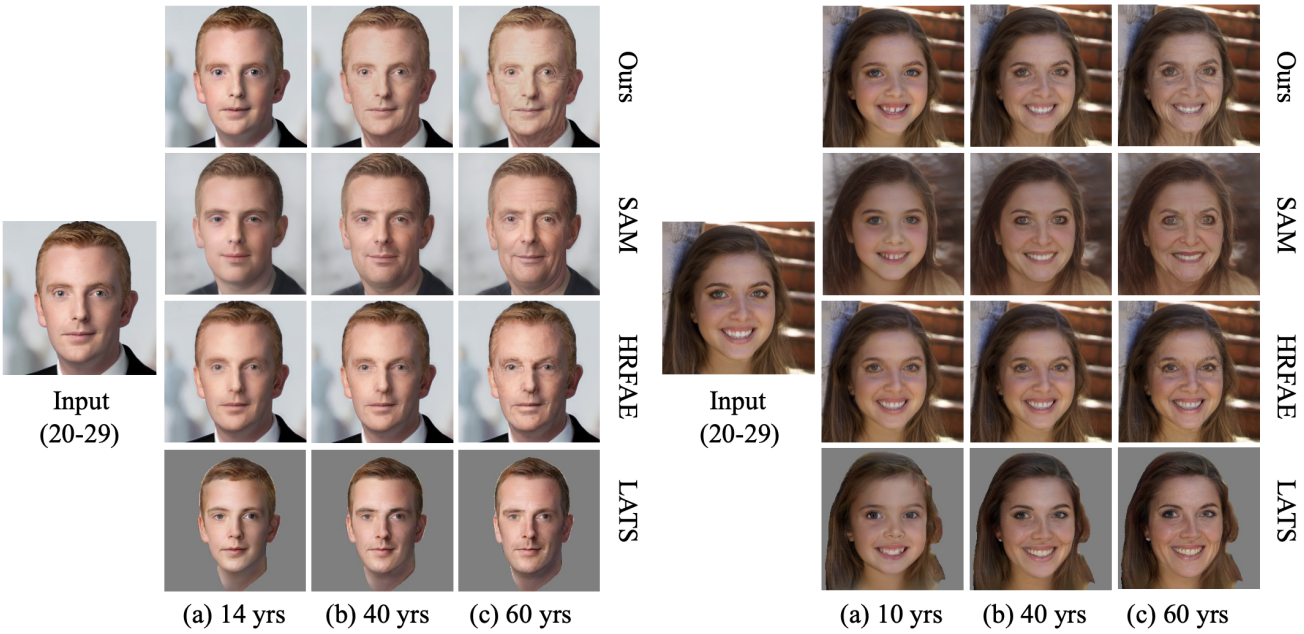


Figure 14: *Qualitative comparison among our method and three SOTA ageing methods, namely SAM [APCO21], HRFAE [YPN*20], and LATS [OESF*20]. Notably, SAM shows a large identity loss compared to the other three methods, particularly for the male subject. In contrast, HRFAE demonstrates improved identity preservation, but this method is constrained by the limited age range (20 – 69 years old). Additionally, challenges in producing significant changes for the oldest age group are observed in both HRFAE and LATS although LATS is the best method among these four to de-age images to the children. Aside from the re-ageing quality, SAM and LATS can NOT preserve the background in the input image, especially for the right subject with a complex background.*

subjects, displayed side by side. SAM has significant identity loss compared to the other methods, particularly failing to reconstruct a similar chin pose for the male subject. Compared with SAM, HRFAE demonstrates improved identity preservation; however, this method is constrained by its limited age range (20-69 years old). Additionally, we observe that HRFAE does not produce significant changes for the oldest group, a problem shared by LATS, although LATS is the most effective method among these four for de-ageing images to childhood. Aside from the re-ageing quality, SAM and LATS cannot preserve the background in the input image, especially for the right subject with a complex background. Please note that despite SAM poor performance in identity preservation, likely due to some pre-processing steps, it performs well in rendering results with different ages. Therefore, similar to [ZCS*22], we choose SAM to generate our training dataset. Overall, our method outperforms the other four in this task.

3D detailed reconstruction: To the best of our knowledge, there is no publicly available 3D facial re-ageing method in computer graphics and vision. Therefore, we only compare with DECA [FFBB21] and EMOCA [DBB22], which are closer to our method. These two methods can NOT edit ages, so we first employ our method to alter the ages in the input images for them and then produce detailed shapes for the aged images using their methods. In contrast, we can obtain the re-aged image and shape directly using our method. As shown in Figure 15, our method outperforms

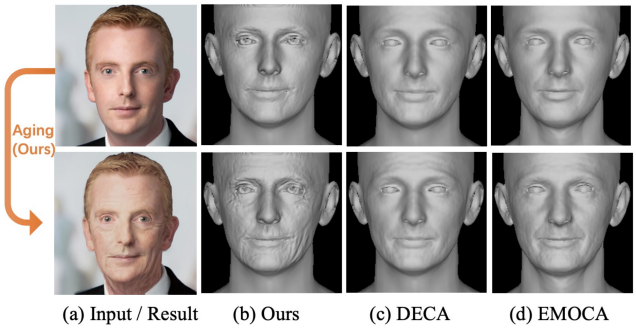


Figure 15: *Comparison of 3D detailed reconstruction among DECA [FFBB21],EMOCA [DBB22], and ours. DECA and EMOCA can NOT alter ages, therefore we feed them the re-aged images using our methods. The input image is taken from the CelebA dataset.*

the other two methods in generating ageing effects on the facial geometry, such as wrinkles around the eyes and chins.

4.3. Quantitative Evaluation

Quantitative evaluation of re-ageing tasks is still a challenge due to the missing photos of people. Similar to recent works [YPN*20, APCO21, LLS19, HKSC19, YHWJ17], we employ an online face recognition API InsightFace [DGXZ19] to evaluate the results

of SAM [APCO21], LATS [OESF*20], HRFAE [YPN*20], and our method. Here, we choose 100 images of the CelebA-HQ dataset [KALL18]. Since the age labels of most test data are "young", we assume the average ages are 25 years old and re-age images to 60 years old to compare. Among these four methods, LATS [OESF*20] is the only method to split ages into many classes, but it can interpolate images between two classes. For the age class 50-69, we choose the middle interpolation of this group to represent 60 years old.

We predict the age and identity accuracy of the re-aged images. The results of average age prediction in Table 1 align with our qualitative evaluation that LATS and HRFAE have under-age issues to different extents. For identity accuracy, we predict the embedding of the original faces using the work of [DGXZ19] and compute the cosine similarity of two embeddings. In Table 1, we note that LATS [OESF*20] and HRFAE [YPN*20] receive slightly better identity preservation than ours while all these three surpass SAM [APCO21]. However, taking their under-age problems into account, 3DFAT (ours) outperforms the other three methods.

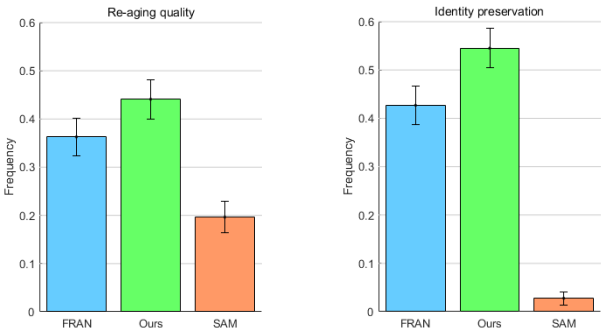
Table 1: *Quantitative evaluation*

	avg. age pred.	identity preservation
SAM [APCO21]	56.93	0.6173
LATS [OESF*20]	43.5681	0.9230
HRFAE [YPN*20]	49.0181	0.9607
Ours	59.17	0.8518

4.4. User Study

To further evaluate the performance of our method, we conducted two user studies with 10 participants each, comparing our method with SAM [APCO21] and FRAN [ZCS*22]. In the first study, participants were instructed to choose the method that best represented how the person in the input photograph might look at a specified target age. We selected 12 input images, including people of different ages (26-61 years old) and genders, with target ages ranging from 20 to 80 years old. In the second study, participants were asked to select the method whose output best preserved the identity of the person at the target age. The experiment proceeded by presenting the reference image on the left and three re-aged images on the right, arranged in random order. Participants engaged in a ranking task, making sequential selections: first choosing the most accurate representation, then the second best from the remaining images. For each input image, the same outputs were presented multiple times, with the order of presentation changed each time to exclude any potential bias. The selections were aggregated across multiple subjects. All experiments were conducted in a darkroom using a calibrated screen to ensure consistent viewing conditions. We implemented our experiments in Matlab utilizing the Psychophysics Toolbox extensions (PsychToolbox v3.0.19) [Bra97, Pel97, KBP*07].

Overall, we collected 576 rankings per user study, including a small pilot run. Following the ranking task, we aggregated the selections across participants and performed statistical analyses using Fisher's exact test to determine the significance of the differences in method preferences. In the first user study, our method significantly outperformed both SAM and FRAN in being selected as the most accurate representation of the person at the target age. Our method was selected as the first choice 44.1% of the time, FRAN 36.3%, and SAM 19.6%. The p-values for the first choices were $p = 0.0082$ (FRAN vs. ours), $p < 1 \times 10^{-9}$ (FRAN vs. SAM), and $p < 1 \times 10^{-18}$ (ours vs. SAM), indicating strong preferences. In the second user study, which focused on identity preservation, the results favored our method, with our method being selected 54.5% of the time, FRAN 42.7%, and SAM 2.8%. The p-values for first choices were $p < 1 \times 10^{-4}$ (FRAN vs. ours), $p < 1 \times 10^{-66}$ (FRAN vs. SAM), and $p < 1 \times 10^{-96}$ (ours vs. SAM). These results demonstrate that our method outperforms FRAN and SAM in both accurately re-aging and preserving the identity of the person. Figure 16 shows the frequencies with which each method was chosen as the first choice in each user study.



(a) Re-aging quality (b) Identity preservation

Figure 16: *Bar plot showing the frequency with which each method was chosen as the first choice, along with 95% confidence intervals.*

4.5. Ablation Study

This section presents three ablation studies: the incorporation of the age factor in *GeoNet*, the age-alpha transformation in *AgeEditNet*, and the placement of biophysical skin editing within our pipeline using *SkinNet*.

Age factor in *GeoNet*: Figure 17 compares 3D coarse shape reconstruction with different ages as input for the *GeoNet*, and the 3D reconstruction result from FLAME [LBB*17]. Our result of facial reconstruction with 20 years old is similar to that of FLAME. However, in contrast to FLAME, our method can manipulate the head shape to a smaller size when given a younger age, and preserve the identity shape, expressions, and poses at the same time.

Age-alpha transformation in *AgeEditNet*: It is well-established that facial changes vary across different age intervals. For example, facial changes are generally more pronounced between the ages of 10 and 20 years than between 20 and 30 years. Therefore, it is not feasible to re-age the input by linearly controlling the parameter α linearly from the source age to the target age as described in Section 3.1.2. Here, we introduce an age-alpha transformation in *AgeEditNet* to address this issue. Figure 18 compares the results with and without this operation. The input image sources from the CelebA dataset with the "young" label. We assume the subject is around 20 years old, and re-age their face to 10 years old and 30 years old, respectively. As shown in this figure, the facial

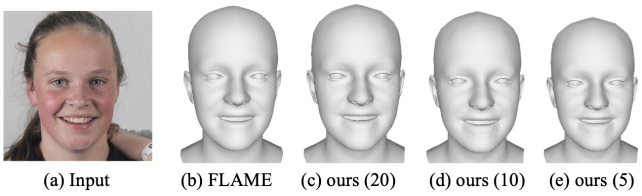


Figure 17: *Comparison between FLAME [LBB*17] and 3D coarse shapes with different ages from GeoNet.*

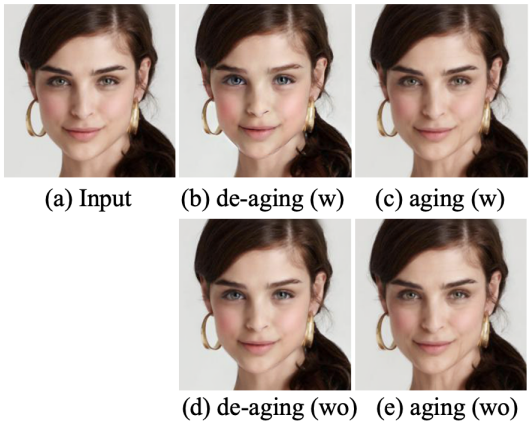


Figure 18: *Ablation study of age-alpha transformation in AgeEditNet. The image is taken from the CelebA dataset.*

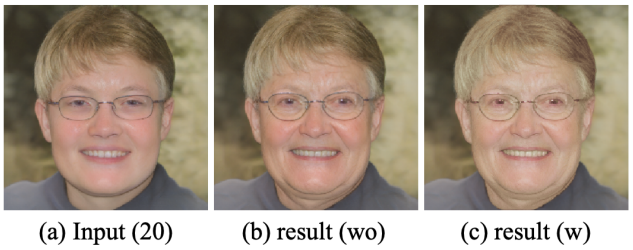


Figure 19: *Comparison between re-ageing results with and without biophysical skin editing.*

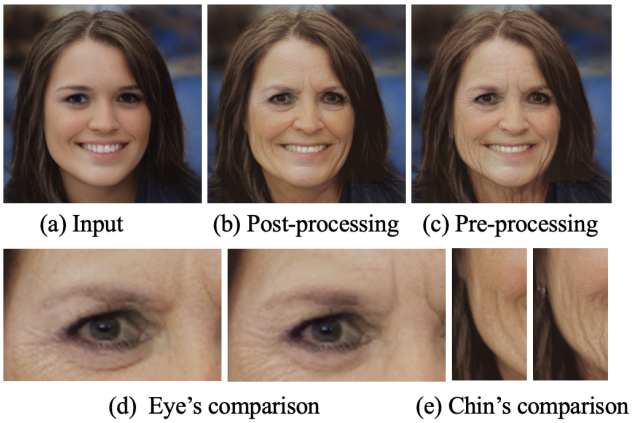


Figure 20: *Ablation study of the biophysical skin editing before and after the structural changes.*

changes from $20-10$ years old are larger than those from $20-30$ years old, consistent with what we would expect for a face.

Biophysical skin editing: We report three ablation studies on biophysical skin editing. The first study aims to show the importance of this step. The second study further discusses the order of this step in the whole pipeline. The third one compares the editing effects based on the concentrations of two chromophores.

Figure 19 (b, c) reports an ablation study of the ageing results for an input image aged 20 years (a) with and without biophysical skin editing, showing the subject at 60 years old. As can be seen, the result with biophysical skin editing appears more realistic than the one without editing.

Figure 20 compares the effects of biophysical skin editing (*SkinNet*) applied before (c) and after (b) the structural changes (*AgeEditNet*). As can be observed, the pre-processing result more effectively produces ageing effects, such as wrinkles, especially in the areas of the eyes and chin. The primary reason is that skin editing reduces the skin color intensity as it ages, resulting in lighter skin tones, including areas around wrinkles. Therefore, we perform skin editing before making changes to facial structures.

The last ablation study for biophysical skin editing examines the different values for the decrease in the concentrations of melanin and hemoglobin, demoted as k_m and k_h, respectively. Biophysical skin editing affects different skin types based on the estimation of these two chromophores. Here is an example showing the input image (60 years old) and de-aged results (30 years old) with different k_m and k_h (see Figure 21). The range for k_m is 0, -0.04, -0.08, -0.16, while the range for k_h is 0, -0.03, -0.06, -0.12.

As shown in this Figure 21, we de-aged the input image (60 yrs) (a) to the result (30 yrs) (e) with $k_m = -0.08, k_h = -0.06$ [IGAJG15]. Images in the top row show that faces get darker as k_m decreases, due to increased melanin concentration. Similarly, images in the second row show that faces get more reddish as k_h decreases, due to increased hemoglobin concentration. These changes are in line with what we would expect from de-aging the subject.

4.6. Limitations

Although our method can produce realistic 3D re-ageing results by editing in both texture and geometry, it still has some limitations. First, we employ a diffusion-based method to edit ages between $10-80$ years old, which excludes the infant and child stages. Second, while our method outperforms many others in identity preservation, it cannot produce certain generative ageing effects, such as gray hair and skin sagging. Third, a halo effect is present in some final results. One possible solution is to employ alpha matting instead of a binary facial mask for better performance around facial boundaries. Additionally, using an eye mask can help mantain accurate eye colors. Moreover, we currently train our method using a 256×256 resolution dataset; however, it can be trained with a higher-resolution dataset. Finally, we do not separate environment

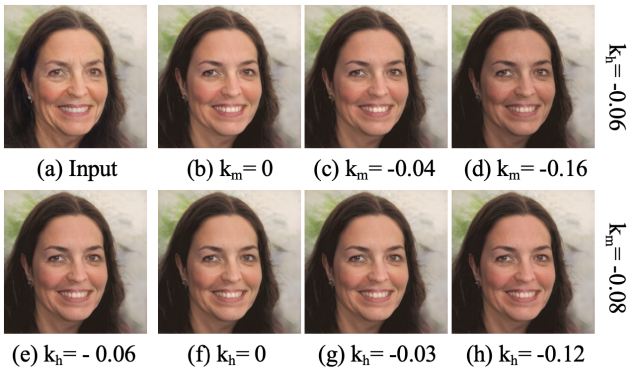


Figure 21: *Ablation study of the biophysical skin editing with different decreases of chromophore concentrations k_m and k_h. The subject in (a) is 60 years old, and the target age is 30 years old, so we expect an increase in chromophores concentration.*

maps from the input images, which impacts the relightability of the generated age-transformed 3D face.

5. Conclusion

This paper presents a neural pipeline for realistic editing of facial age in terms of both facial appearance and 3D shape with age-appropriate details. Unlike previous works that either focused on purely 2D age transformation of facial images, or purely focused on 3D geometry estimation for a given 2D facial photograph, our method is the first to propose a joint pipeline that uplifts facial age transformation from 2D to 3D. Furthermore, we demonstrate state-of-the-art results for both the aspects of age-appropriate facial appearance transformation with identity preservation, as well as facial geometry transformation including reduction in head size for young age and fine scale geometry for skin wrinkles for older age. In future work, it would be interesting to extend our method to handle changes in hair appearance as well as larger scale changes in facial shape due to ageing using generative techniques. Additional useful direction of extension would be predicting de-lit albedo textures with age-transformations for realistic relighting applications.

Acknowledgements

We thank the volunteer subjects for participating in the user study. This work was partly supported by the EPSRC grant EP/X011364/1: GNOMON, and by the BBSRC grant BB/X01312X/1.

References

[APCO21] ALALUF Y., PATASHNIK O., COHEN-OR D.: Only a matter of style: Age transformation using a style-based regression model. *ACM Trans. Graph. 40*, 4 (jul 2021). doi:10.1145/3450626.3459805. 2, 4, 7, 8, 9

[AXX*23] ALIAGA C., XIA M., XIE H., JARABO A., BRAUN G., HERY C.: A Hyperspectral Space of Skin Tones for Inverse Rendering of Biophysical Skin Properties. *Computer Graphics Forum* (2023). doi:10.1111/cgf.14887. 2, 3

[Bra97] BRAINARD D. H.: The psychophysics toolbox. *Spatial Vision 10*, 4 (1997), 433 – 436. doi:10.1163/156856897X00357. 9

[BV99] BLANZ V., VETTER T.: A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques* (USA, 1999), SIGGRAPH '99, ACM Press/Addison-Wesley Publishing Co., p. 187–194. doi:10.1145/311535.311556. 2

[CBZB15] CAO C., BRADLEY D., ZHOU K., BEELER T.: Real-time high-fidelity facial performance capture. *ACM Trans. Graph. 34*, 4 (jul 2015). doi:10.1145/2766943. 5

[CCZ*19] CHEN A., CHEN Z., ZHANG G., MITCHELL K., YU J.: Photo-realistic facial details synthesis from single image. In *Proceedings of the IEEE International Conference on Computer Vision* (2019), pp. 9429–9439. 2

[CUH16] CLEVERT D., UNTERTHINER T., HOCHREITER S.: Fast and accurate deep network learning by exponential linear units (elus). In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings* (2016), Bengio Y., LeCun Y., (Eds.). 5

[CWZ*14] CAO C., WENG Y., ZHOU S., TONG Y., ZHOU K.: Facewarehouse: A 3d facial expression database for visual computing. *IEEE Transactions on Visualization and Computer Graphics 20*, 3 (2014), 413–425. doi:10.1109/TVCG.2013.249. 2

[DBB22] DANECEK R., BLACK M. J., BOLKART T.: EMOCA: Emotion driven monocular face capture and animation. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2022), pp. 20311–20322. 3, 7, 8

[DBV16] DEFFERRARD M., BRESSON X., VANDERGHEYNST P.: Convolutional neural networks on graphs with fast localized spectral filtering. In *Proceedings of the 30th International Conference on Neural Information Processing Systems* (Red Hook, NY, USA, 2016), NIPS'16, Curran Associates Inc., p. 3844–3852. 5

[DGXZ19] DENG J., GUO J., XUE N., ZAFEIRIOU S.: Arcface: Additive angular margin loss for deep face recognition. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 4685–4694. doi:10.1109/CVPR.2019.00482. 8, 9

[DJ06] DONNER C., JENSEN H. W.: A spectral bssrdf for shading human skin. In *Proceedings of the 17th Eurographics Conference on Rendering Techniques* (Goslar, DEU, 2006), EGSR '06, Eurographics Association, p. 409–417. 2, 3

[FFBB21] FENG Y., FENG H., BLACK M. J., BOLKART T.: Learning an animatable detailed 3d face model from in-the-wild images. *ACM Trans. Graph. 40*, 4 (jul 2021). doi:10.1145/3450626.3459936. 2, 3, 5, 7, 8

[GGD*20] GITLINA Y., GUARNERA G. C., DHILLON D. S., HANSEN J., LATTAS A., PAI D., GHOSH A.: Practical measurement and reconstruction of spectral skin reflectance. *Computer Graphics Forum 39* (07 2020), 75–89. doi:10.1111/cgf.14055. 2, 3

[GPAM*20] GOODFELLOW I., POUGET-ABADIE J., MIRZA M., XU B., WARDE-FARLEY D., OZAIR S., COURVILLE A., BENGIO Y.: Generative adversarial networks. *Commun. ACM 63*, 11 (oct 2020), 139–144. doi:10.1145/3422622. 2

[GTLMC22] GOMEZ-TRENADO G., LATHUILIÈRE S., MESEJO P., CORDÓN Ó.: Custom structure preservation in face aging. In *European Conference on Computer Vision* (2022), Springer, pp. 565–580. 2

[GZC*17] GUO Y., ZHANG J., CAI J., JIANG B., ZHENG J.: Cnn-based real-time dense face reconstruction with inverse-rendered photo-realistic face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 41* (2017), 1294–1307. 2

[HBC23] HEITZ E., BELCOUR L., CHAMBON T.: Iterative -(de)blending: a minimalist deterministic diffusion model. In *ACM SIGGRAPH 2023 Conference Proceedings* (New York, NY, USA, 2023), SIGGRAPH '23, Association for Computing Machinery. doi:10.1145/3588432.3591540. 2, 4

[HKSC19] He Z., Kan M., Shan S., Chen X.: S2gan: Share aging factors across ages and share aging trends among individuals. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (2019), pp. 9439–9448. doi:10.1109/ICCV.2019.00953. 1, 2, 8

[HLY*21] He S., Liao W., Yang M., Song Y., Rosenhahn B., Xiang T.: Disentangled lifespan face synthesis. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* (Los Alamitos, CA, USA, oct 2021), IEEE Computer Society, pp. 3857–3866. doi:10.1109/ICCV48922.2021.00385. 2

[IGAJG15] Iglesias-Guitian J. A., Aliaga C., Jarabo A., Gutierrez D.: A biophysically-based model of the optical properties of skin aging. *Comput. Graph. Forum 34*, 2 (may 2015), 45–55. doi:10.1111/cgf.12540. 2, 4, 10

[JSB*10] Jimenez J., Scully T., Barbosa N., Donner C., Alvarez X., Vieira T., Matts P., Orvalho V., Gutierrez D., Weyrich T.: A practical appearance model for dynamic facial color. *ACM Transactions on Graphics 29*, 6 (2010), 1–10. doi:10.1145/1866158.1866167. 2, 3

[KALL18] Karras T., Aila T., Laine S., Lehtinen J.: Progressive growing of GANs for improved quality, stability, and variation. In *International Conference on Learning Representations* (2018). 9

[KBP*07] Kleiner M., Brainard D., Pelli D., Ingling A., Murray R., Broussard C.: What's new in psychtoolbox-3. *Perception 36*, 14 (2007), 1–16. 9

[KLA21] Karras T., Laine S., Aila T.: A style-based generator architecture for generative adversarial networks. *IEEE Trans. Pattern Anal. Mach. Intell. 43*, 12 (dec 2021), 4217–4228. doi:10.1109/TPAMI.2020.2970919. 4

[KSSS14] Kemelmacher-Shlizerman I., Suwajanakorn S., Seitz S. M.: Illumination-aware age progression. In *2014 IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 3334–3341. doi:10.1109/CVPR.2014.426. 2

[LBB*17] Li T., Bolkart T., Black M. J., Li H., Romero J.: Learning a model of facial shape and expression from 4d scans. *ACM Trans. Graph. 36*, 6 (nov 2017). doi:10.1145/3130800.3130813. 2, 4, 5, 9, 10

[LGLG24] Li X., Guarnera G., Lin A., Ghosh A.: Practical measurement and neural encoding of hyperspectral skin reflectance. In *2024 International Conference on 3D Vision (3DV)* (Los Alamitos, CA, USA, mar 2024), IEEE Computer Society, pp. 1301–1309. doi:10.1109/3DV62453.2024.00116. 2, 3, 4

[LLS19] Liu Y., Li Q., Sun Z.: Attribute-aware face aging with wavelet-based generative adversarial networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Los Alamitos, CA, USA, jun 2019), IEEE Computer Society, pp. 11869–11878. doi:10.1109/CVPR.2019.01215. 1, 2, 8

[LMG*20] Lattas A., Moschoglou S., Gecer B., Ploumpis S., Triantafyllou V., Ghosh A., Zafeiriou S.: Avatarme: Realistically renderable 3d facial reconstruction "in-the-wild". In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Los Alamitos, CA, USA, jun 2020), IEEE Computer Society, pp. 757–766. doi:10.1109/CVPR42600.2020.00084. 2

[LTC02] Lanitis A., Taylor C., Cootes T.: Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 4 (2002), 442–455. doi:10.1109/34.993553. 2

[NISA06] Nealen A., Igarashi T., Sorkine O., Alexa M.: Laplacian mesh optimization. In *Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia* (New York, NY, USA, 2006), GRAPHITE '06, Association for Computing Machinery, p. 381–389. doi:10.1145/1174429.1174494. 5

[OESF*20] Or-El R., Sengupta S., Fried O., Shechtman E., Kemelmacher-Shlizerman I.: Lifespan age transformation synthesis. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI* (Berlin, Heidelberg, 2020), Springer-Verlag, p. 739–755. doi:10.1007/978-3-030-58539-6_44. 1, 2, 4, 7, 8, 9

[Pel97] Pelli D. G.: The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision 10*, 4 (1997), 437 – 442. doi:10.1163/156856897X00366. 9

[PKA*09] Paysan P., Knothe R., Amberg B., Romdhani S., Vetter T.: A 3d face model for pose and illumination invariant face recognition. In *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance* (2009), pp. 296–301. doi:10.1109/AVSS.2009.58. 2

[RBSB18] Ranjan A., Bolkart T., Sanyal S., Black M. J.: Generating 3d faces using convolutional mesh autoencoders. In *Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part III* (2018), Springer-Verlag, p. 725–741. doi:10.1007/978-3-030-01219-9_43. 2, 5

[RH01] Ramamoorthi R., Hanrahan P.: An efficient representation for irradiance environment maps. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2001), SIGGRAPH '01, Association for Computing Machinery, p. 497–500. doi:10.1145/383259.383317. 5

[RTG15] Rothe R., Timofte R., Gool L. V.: Dex: Deep expectation of apparent age from a single image. In *IEEE International Conference on Computer Vision Workshops (ICCVW)* (December 2015). 3, 4

[SCS*12] Suo J.-L., Chen X., Shan S., Gao W., Dai Q.: A concatenational graph evolution aging model. *IEEE Transactions on Pattern Analysis and Machine Intelligence 34* (2012), 2083–2096. 2

[TBP01] Tiddeman B., Burt M., Perrett D.: Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications 21*, 5 (2001), 42–50. doi:10.1109/38.946630. 2

[TGMM12] Tazoe Y., Gohara H., Maejima A., Morishima S.: Facial aging simulator considering geometry and patch-tiled texture. In *ACM SIGGRAPH 2012 Posters* (New York, NY, USA, 2012), SIGGRAPH '12, Association for Computing Machinery. doi:10.1145/2342896.2343002. 2

[THM*17] Tran A., Hassner T., Masi I., Paz E., Nirkin Y., Medioni G. G.: Extreme 3d face reconstruction: Seeing through occlusions. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2017), 3935–3944. 2, 5, 6

[YHWJ17] Yang H., Huang D., Wang Y., Jain A. K.: Learning face age progression: A pyramid architecture of gans. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2017), 31–39. 8

[YPN*20] Yao X., Puy G., Newson A., Gousseau Y., Hellier P.: High resolution face age editing. *2020 25th International Conference on Pattern Recognition (ICPR)* (2020), 8624–8631. 2, 7, 8, 9

[YZW*20] Yang H., Zhu H., Wang Y., Huang M., Shen Q., Yang R., Cao X.: Facescape: A large-scale high quality 3d face dataset and detailed riggable 3d face prediction. 598–607. doi:10.1109/CVPR42600.2020.00068. 5

[ZBT22] Zielonka W., Bolkart T., Thies J.: Towards metrical reconstruction ofnbsp;human faces. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIII* (Berlin, Heidelberg, 2022), Springer-Verlag, p. 250–269. doi:10.1007/978-3-031-19778-9_15. 2

[ZCS*22] Zoss G., Chandran P., Sifakis E., Gross M., Gotardo P., Bradley D.: Production-ready face re-aging for visual effects. *ACM Trans. Graph. 41*, 6 (nov 2022). doi:10.1145/3550454.3555520. 2, 4, 7, 8, 9