# Continuous Face Aging via Self-estimated Residual Age Embedding

Zeqi Li*
ModiFace
lizeqi@cs.toronto.edu

Ruowei Jiang
ModiFace
irene@modiface.com

Parham Aarabi
ModiFace
parham@modiface.com

## 1. Supplementary

### 1.1. Network Architecture and Optimization Settings

During training, we use Adam optimizer with the learning rate of 0.0002 and batch size of 20 and 5 for 256 and 512 model respectively. The model is trained for 200 epochs and learning rate is linearly decayed over last 100 epochs.

| Layer | Stride | Act. | Norm | Output Shape |
|---|---|---|---|---|
| Input | - | - | - | 256x256x3 |
| Conv. 7 x 7 | 1 | ReLU | Spectral | 256x256x64 |
| Conv. 3 x 3 | 2 | ReLU | Spectral | 128x128x128 |
| Conv. 3 x 3 | 2 | ReLU | Spectral | 64x64x256 |
| Res. Block | 1 | ReLU | Spectral | 64x64x256 |
| Res. Block | 1 | ReLU | Spectral | 64x64x256 |
| Res. Block | 1 | ReLU | Spectral | 64x64x256 |
| Res. Block | 1 | ReLU | Spectral | 64x64x256 |
| Res. Block | 1 | ReLU | Spectral | 64x64x256 |
| Res. Block | 1 | ReLU | Spectral | 64x64x256 |

Table 1. Identity Encoder **E** specification. Spectral means spectral normalization [4] is applied after each convolutional layer.

| Layer | Stride | Act. | Norm | Output Shape |
|---|---|---|---|---|
| Encoding | - | - | - | 64x64x256 |
| Res. Block | 1 | ReLU | Instance | 64x64x256 |
| Res. Block | 1 | ReLU | Instance | 64x64x256 |
| Res. Block | 1 | ReLU | Instance | 64x64x256 |
| Deconv. 3 x 3 | 2 | ReLU | Instance | 128x128x128 |
| Deconv. 3 x 3 | 2 | ReLU | Instance | 256x256x64 |
| Conv. 7 x 7 | 1 | Tanh | - | 256x256x3 |

Table 2. Generator **G** specification. Instance means instance normalization [7] is applied after each convolutional layer.

Detailed network architectures for **E**, **G** and **C** are presented in **Table** 1, 2 and 3 respectively.

*This work is done during Zeqi Li's full-time employment at ModiFace.

| Layer | Norm | Output Shape |
|---|---|---|
| Encoding | - | 64x64x256 |
| GAP | - | 1x1x256 |
| Flatten | - | 256 |
| Linear | Weight | 100 |

Table 3. Age estimator **C** specification. GAP means global average pooling. Weight means weight normalization [6] is applied to linear layer (bias term are set to zero).

### 1.2. Pair-wise Identity Preservation Results

Here, we provide the complete pair-wise identity preservation comparison using Face++ in **Table** 4 and 5 for CACD2000 [1] and FFHQ [3], respectively. As can be seen, our model achieves the highest verification rate in every aspects compared to prior works.

### 1.3. More Aging Results

**Continuous Aging.** We generate the complete continuous aging results of a person from age 20 to age 69 and the results are displayed in **Fig.** 3. As shown, aging proceeds in a natural and gradual manner.

**Enlarged Comparison of group 50+** In **Fig.** 1, we show the enlarged generated images of age group 50+. Our model is able to generate fine aging details aligned with the target age group.

### 1.4. Limitations

While our work can generate natural face aging, we also observe some failure cases when generating outputs for input image with hats and glasses or faces with heavy make-ups (in **Fig.** 2). The model also does not work well for extreme target age like 95-year-old, where the corresponding exemplar-face aging basis is hardly trained due to lack of data for those minority classes.

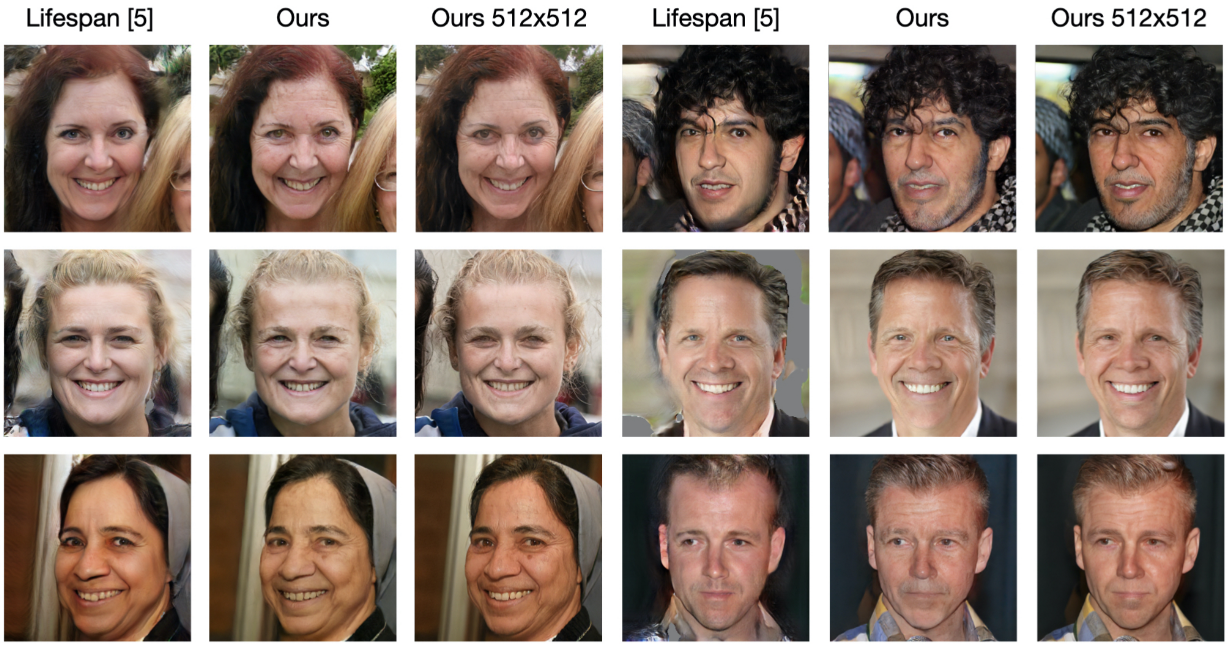


Figure 1. Enlarged examples of generated age 50+. Our model demonstrate better details in aging effects such as wrinkles and beard change.

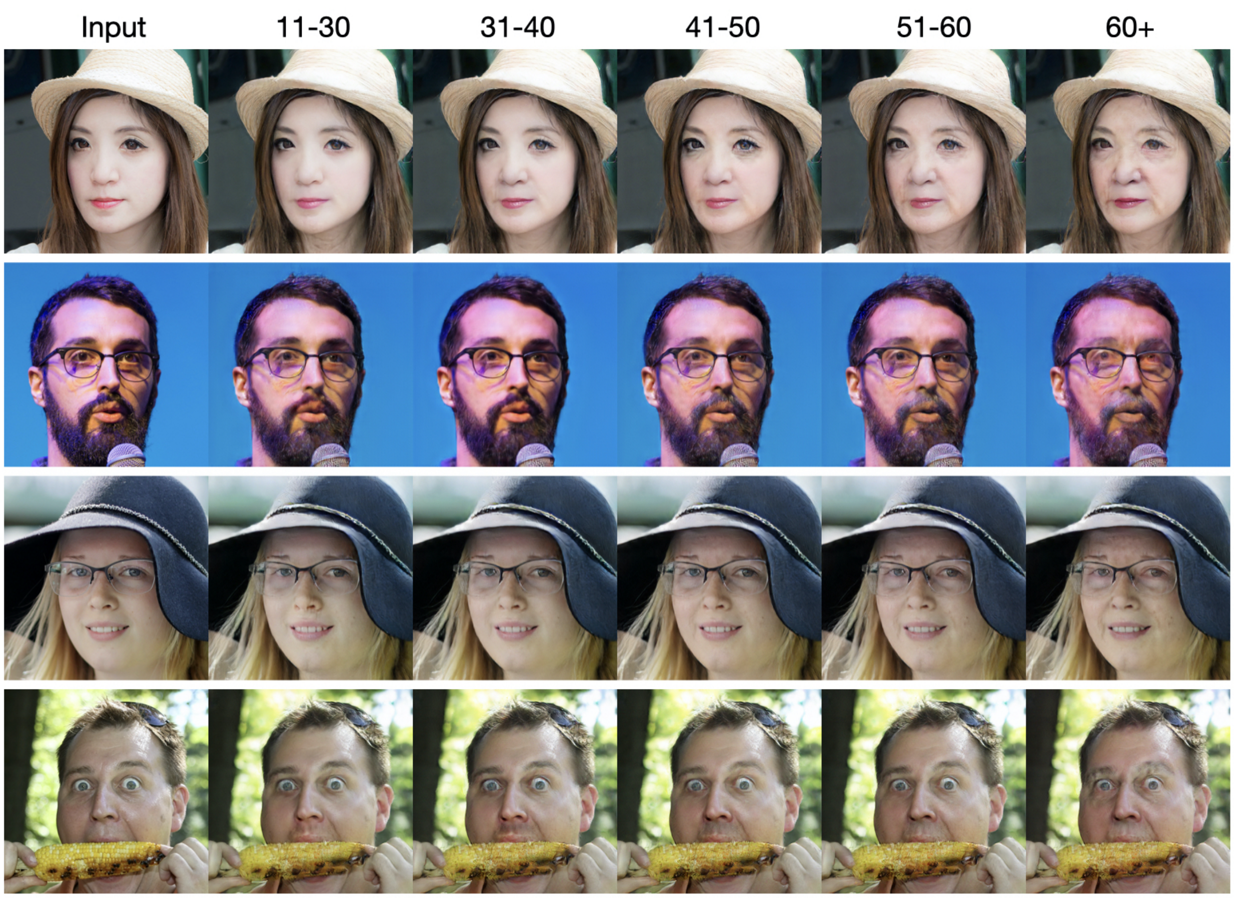


Figure 2. Failure cases: heavy makeup with hat, glasses with bad lighting. Our model could not best capture the personalized information such as skin texture in these cases.

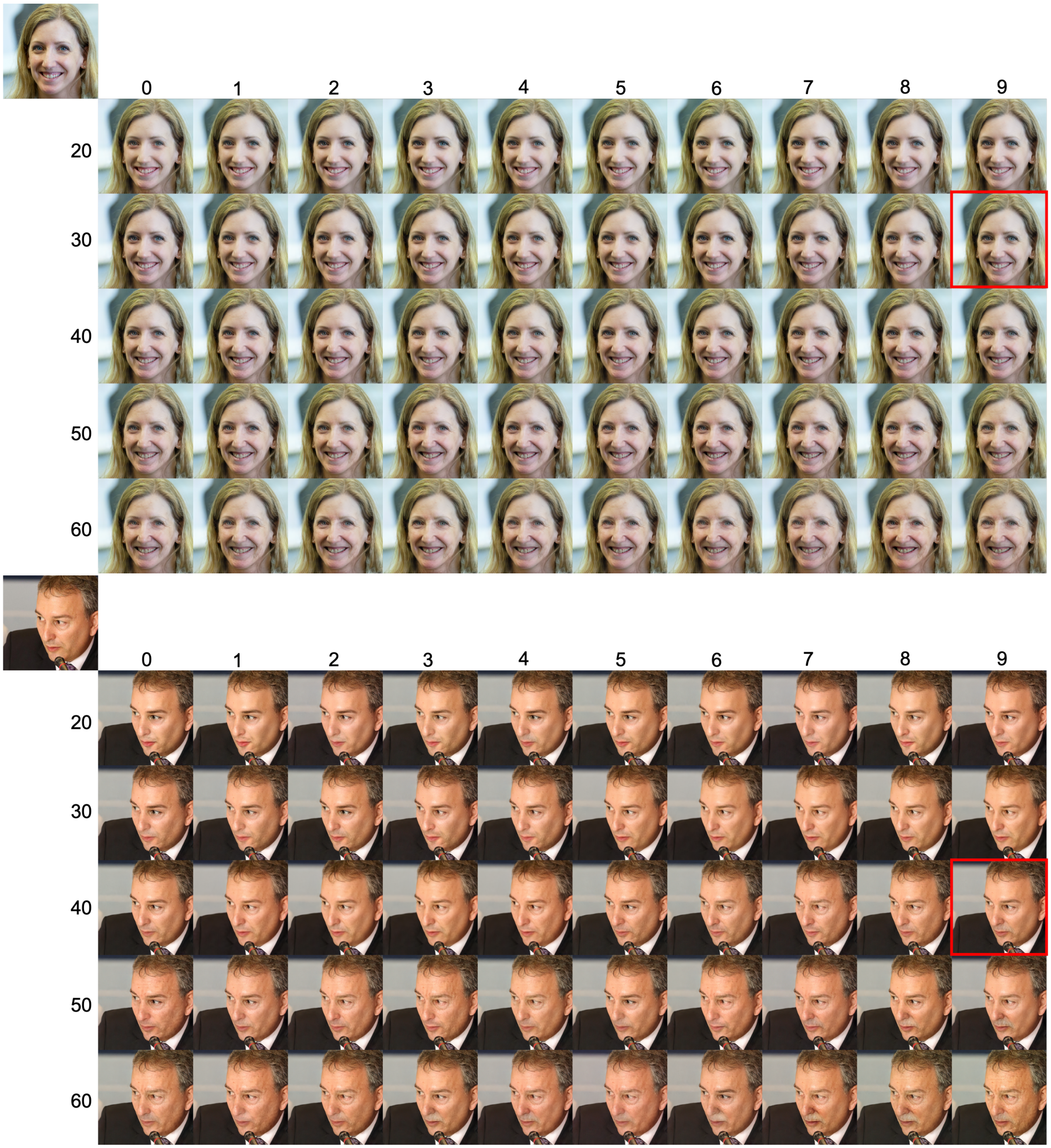


Figure 3. Complete continuous aging results from age 20 to 69. Input is at the top left corner of each image grid. Generated image of real age is in the red box.

| | Average of All Pairs | Hardest Pair | Easiest Pair |
|---|---|---|---|
| CAAE [9] | 60.88% | (test, 50+): 2.0% | (40-49, 50+): 99.97% |
| IPCGAN [8] | 91.40% | (10-29, 50+): 62.98% | (40-49, 50+): 99.98% |
| $S^2$GAN [2] | 98.91% | (10-29, 40-49): 94.08% | (40-49, 50+): 99.96% |
| Lifespan [5] | 93.25% | (test, 50-69): 80.94% | (30-39, 50-69): 99.75% |
| Ours | **99.97%** | (test, 40-49): 99.96% | (test, 30-39): 100.00% |

Table 4. Complete evaluation of identity preservation in terms of face verification rates on CACD2000 [1].

| | Average of All Pairs | Hardest Pair | Easiest Pair |
|---|---|---|---|
| Lifespan [5] | 87.11% | (test, 50-69): 72.32% | (30-39, 50-69): 98.85% |
| Ours | **99.98%** | (test, 60+): 99.96% | (test, 30-39): 100.00% |

Table 5. Complete evaluation of identity preservation in terms of face verification rates on FFHQ [3].

## References


[1] Bor-Chun Chen, Chu-Song Chen, and Winston H Hsu. Cross-age reference coding for age-invariant face recognition and retrieval. In *European conference on computer vision*, pages 768–783. Springer, 2014. 1, 4

[2] Zhenliang He, Meina Kan, Shiguang Shan, and Xilin Chen. S2gan: Share aging factors across ages and share aging trends among individuals. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 9440–9449, 2019. 4

[3] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4401–4410, 2019. 1, 4

[4] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. In *International Conference on Learning Representations*, 2018. 1

[5] Roy Or-El, Soumyadip Sengupta, Ohad Fried, Eli Shechtman, and Ira Kemelmacher-Shlizerman. Lifespan age transformation synthesis. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. 4

[6] Tim Salimans and Durk P Kingma. Weight normalization: A simple reparameterization to accelerate training of deep neural networks. In *Advances in neural information processing systems*, pages 901–909, 2016. 1

[7] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016. 1

[8] Zongwei Wang, Xu Tang, Weixin Luo, and Shenghua Gao. Face aging with identity-preserved conditional generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7939–7947, 2018. 4

[9] Zhifei Zhang, Yang Song, and Hairong Qi. Age progression/regression by conditional adversarial autoencoder. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5810–5818, 2017. 4