Time Traveler: a real-time face aging system

Lejian Ren
SKLOIS, IIE, CAS
renlejian@outlook.com

Si Liu
SKLOIS, IIE, CAS
liusi@iie.ac.cn

Yao Sun
SKLOIS, IIE, CAS
sunyao@iie.ac.cn

Jian Dong
Qihoo 360 AI Institute, Beijing, China
Dongjian-iri@360.cn

Luoqi Liu
Qihoo 360 AI Institute, Beijing, China
llq667@gmail.com

Shuicheng Yan
Qihoo 360 AI Institute, Beijing, China
eleyans@nus.edu.sg

ABSTRACT

In this demo we introduce a novel sysetm that can generate face aging results in real-time. It contains sevel important techniques, including face detection, alignmetn, parsing, and a novel Contextual Generative Adversarial Nets (C-GANs) based face aging modules. The C-GANs consists of a conditional transformation network, an age discriminative network and a transition pattern discriminative network.

CCS CONCEPTS

• **Computing methodologies → Pattern Analysis and Machine Intelligence**;

KEYWORDS

Face Aging, Generative Adversarial Nets, Contextual Modeling

Figure 1: The *Time Traveler* sysetem. The users stand in front of a monintor and a camera. His/her aging results are instantly shown in the monitor.

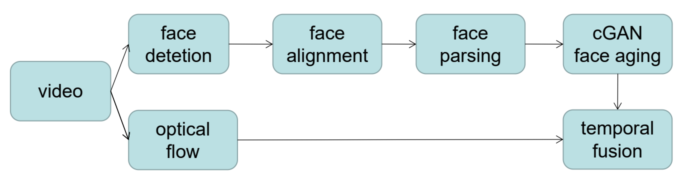

Figure 2: The framework of the *Time Traveler* system.

1 INTRODUCTION

Face aging, also known as age progression, is attracting more and more research interests. It has plenty of applications in various domains including cross-age face recognition, finding lost children, and entertainments. In recent years, face aging has witnessed various breakthroughs and a number of face aging models have been proposed [3]. Face aging, however, is still a very challenging task in practice for various reasons. First, faces may have many different expressions and lighting conditions, which pose great challenges to modeling the aging patterns. Besides, the training data are usually very limited and the face images for the same person only cover a narrow range of ages.

In this paper, we mainly consider the cross-age transition pattern. Specifically, transition pattern contains two aspects. One is the identity consistency, and the other is the appearance changes. Identity preserving is critical in face aging based applications, e.g., cross-age face verification. Appearance changes include texture and shape alterations. Transition pattern is age-aware. For example, when one grows from baby to teenagers, the main appearance difference is the face becomes larger. When one grows from the age of 50 to 60, the main facial changes lie on the texture alteration, such as the gradually developed eye bag, senile plaques and wrinkle. Different from traditional GANs which only model the real data distribution of each individual age, we focus on the higher-order cross-age correlations, which will make the face aging results more appealing. To model the above-mentioned transition patterns, we propose a Contextual Generative Adversarial Nets (C-GANs). Figure 1 illustrates C-GANs briefly. For an input face, C-GANs can generate faces for any target age group. To ensure the generated images real, C-GANs uses two discriminative networks to model the distribution of each individual age group as well as the transition patterns of two adjacent groups respectively.

The contributions of this paper are summarized as follows. First, we design an effective and efficient contextual GANs [5] based face aging system whose aging results are significantly better than existing face aging methods. The source code of our method will be released to the academic area for further research. Second, we introduce a novel transition pattern discriminative network in the C-GANs to regularize the synthesized faces satisfying the cross-age face aging rules. Third, the conditional face transformation network in C-GANs is different with existing GANs generators in that it is much deeper, with several specially designed skip layers to preserve both the high-level semantics and low-level information. It makes the generated images more natural and real.

2 THE SYSTEM

The system architecture is shown in Figure 2. It contains six modules, including face detection, face alignment, face parsing, cGAN face aging, optical flow estimation and temporal fusion modules.

We train our *face detection* model on the AFLW and Wider Face datasets with a framework similar to Faster R-CNN [6]. To speed up, we use face tracking MCPF [8]. However, images from both data sets are from the Internet, mostly with good lighting conditions. The training set is expanded by adding extra face images collected by ourselves based on application requirements. The input image x is aligned via the *face alignment* techniques [1] which locates 68 points on the faces. The landmark is used to align the faces. Then we use a face parsing model [7] to conduct *face parsing* to generate facial and non-facial regions. The non-facial region containing the background, the hair and clothes, are masked with gray color to facilitate the GANs training.

Then the image is paired with an arbitrary age label to feed into the *C-GANs face aging* module, which consists of three networks. The conditional transformation network transforms the input face to the desired age; the age discriminative network assists generating images indistinguishable with the real ones; the transition pattern discriminative network regularize the generated images satisfying the cross-age aging rules. The proposed C-GANs can be trained end-to-end and very easy to reproduce.

For robust and more natural face aging result, We adopt the Flownet [2] to estimate the cross-frame correpondence. The optcal flow as well as the sequential face aging results are fed into the *temporal fusion* module to produce the final face aging result.

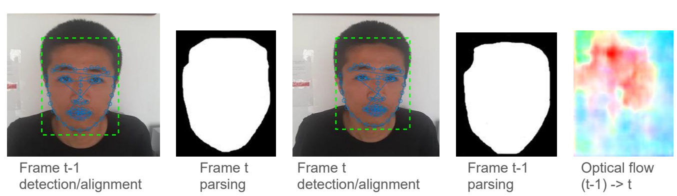

Figure 3: step by step result

3 DEMOSTRATION

The demonstration will consist of a laptop, a video camera and and a large screen. The laptop processes the photos caputured by the video camera. The face aging results are shown on the screen. A poster detailing the our research will also be shown.

We show the face aging results in Figure 2. We can see that C-GANs can generate quite appealing results. We can draw the following observations. First, the faces are quite real and natural. Second, the generated faces can change gradually when getting older. For example, in the output faces of the last 5 age groups (in red boxes) of the first and second rows, the beards appear and become white. Third, C-GANs can synthesize images with large age gaps. For example, the input face of the third row is quite young, but the synthesized faces in the 60+ group (in yellow box) is still quite real. For another example, the face in the fourth row is a senior lady. We can produce very child-looking young face (in yellow box) for the 0 – 10 age group. Fourth, C-GANs can produce very detailed texture changes. For example, in the fifth, sixth and seventh rows, the synthesized faces in the red boxes contain frighteningly real convincing enough crow's feet, canthus wrinkles and eye bags. Fifth, the shapes of face and facial features also change during the aging/progression. For example, in the last two rows, when the seniors are transformed to babies, their face become smaller, and their eyes and ocular distance become larger.

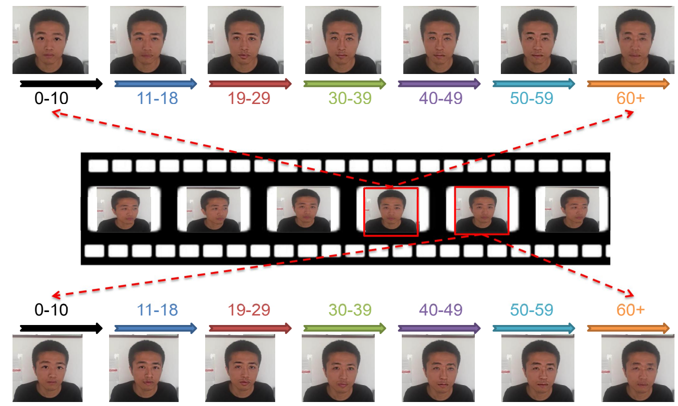

Figure 4: The input face and the generated faces for 7 age groups.

4 CONCLUSIONS

In this paper, we propose a contextual generative adversarial nets to tackle the face aging problem. Different from existing generative adversarial nets based methods, we explicitly model the transition patterns between adjacent age groups during the training procedure. From baby to teenagers period, the transition patterns is shown in the way that the face becomes bigger, while from the ages of 30 to the age of 50, the transition patterns include the gradually developed wrinkle. To this end, the C-GANs consists of two discriminative networks, i.e., an age discriminative network and a transition pattern discriminative network. They are collaboratively contribute to the appealing results. In future, we plan to use the technique for other facial applications, such as automatic makeup [4].

5 ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (No.61572493, Grant U1536203).

REFERENCES

[1] Dlib C++ Library. http://dlib.net/. (????).
[2] Alexey Dosovitskiy, Philipp Fischer, Eddy Ilg, Philip Hausser, Caner Hazirbas, Vladimir Golkov, Patrick van der Smagt, Daniel Cremers, and Thomas Brox. 2015. Flownet: Learning optical flow with convolutional networks. In *ICCV*.
[3] Y. Fu, G. Guo, and T. S. Huang. 2010. Age synthesis and estimation via faces: a survey. *TPAMI* (2010).
[4] Si Liu, Xinyu Ou, Ruihe Qian, Wei Wang, and Xiaochun Cao. 2016. Makeup like a superstar: Deep localized makeup transfer network. *IJCAI* (2016).
[5] Si Liu, Yao Sun, Wei Wang, Defa Zhu, Xiangbo Zhu, and Shuicheng Yan. 2017. Face Aging with Contextual Generative Adversarial Nets. In *ACM MM*.
[6] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In *NIPS*.
[7] Zhen Wei, Yao Sun, Jinqiao Wang, Hanjiang Lai, and Si Liu. 2017. Learning Adaptive Receptive Fields for Deep Image Parsing Network. In *CVPR*.
[8] Tianzhu Zhang, Changsheng Xu, and Ming-Hsuan Yang. 2017. Multi-task Correlation Particle Filter for Robust Visual Tracking. In *CVPR*.