Exploring the Powerful Models of Face Aging and Rejuvenation

Zhizhong Huang
School of Computer Science

November 7, 2019

Content

- Starting up
  - Introduction to Face Aging
  - Prior works of Face Aging
- Powerful Models
  - Conditional Adversarial Autoencoder(CAAE)[1]
  - Learning Face Age Progression: A Pyramid Architecture of GANs(PyramidGAN)[2]
  - Identity-Preserved Conditional GAN(IPCGAN)[3]
  - Attribute-aware Wavelet-based GAN(AWGAN)[4]
  - GAN with Spatial Attention Modules[5]

Introduction

- Synthesizing faces of a certain person under a given age.
- Face aging is of great importance for cross-age recognition, entertainment, forensics, *e.g.* , *help find lost children or predict someone looks like in the future.*
- Tree basic requirements of Face Aging Methods:
  - Photorealistic synthesized faces with natural aging effect, without ghost artifacts or blurry.
  - Preserve the identity of the faces.
  - Esimated age of synthesized faces should lie in the target age.

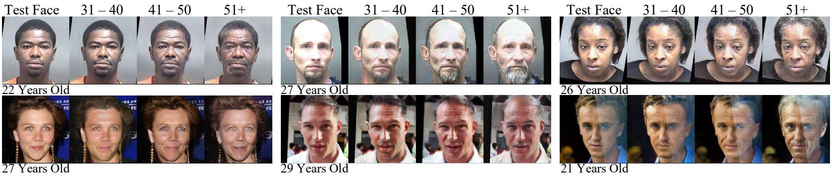

Figure: Illustration of generation results by [5]

Introduction

Most used datasets in Face Aging:

Database	Number of images	Number of subjects	Number of images per subject	Time lapse per subject (years)	Age span (years old)	Average a (years ol
MORPH [6]	52,099	12,938	1 - 53 (avg. 4.03)	0 - 33 (avg. 1.62)	16 - 77	33.07
CACD [7]	163,446	2,000	22 - 139 (avg. 81.72)	7 - 9 (avg. 8.99)	14 - 62	38.03
FG-NET	1,002	82	6 - 18 (avg. 12.22)	11 - 54 (avg. 27.80)	0 - 69	15.84

Table: Statistics of face aging databases used for evaluation [2].

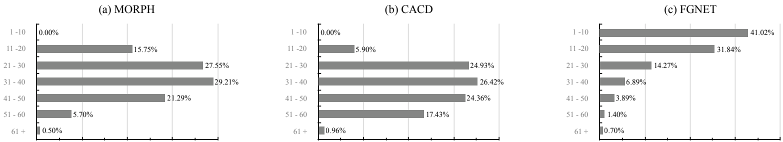

Figure: Age distributions of (a) MORPH, (b) CACD, and (c) FGNET [2].

Prior works

- Traditional face aging methods can roughly be categorized into (1) prototype-based [8] and (2) physical model-based approaches [9].
- Prototype-based approaches compute an average face of each age group as aging pattern which is used for synthesizing an aged face [8], while personalized features of the face tend to get lost.
- Physical model-based approaches create parametric models for aging various parts of the face thus suffer from high complexity and computational cost [9].
- There is a significant breakthrough since *cGANs-based* methods has been introduced into Face Aging [1, 3, 2, 4, 5], which will be discussed here.

Prior works

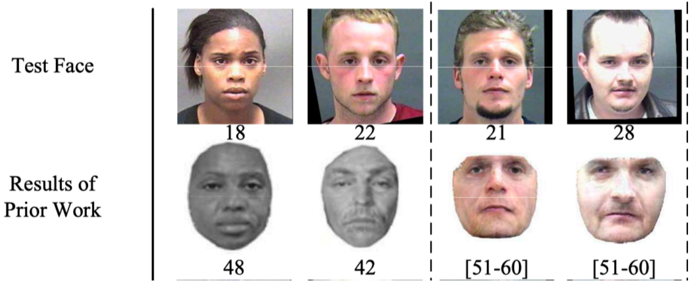

Figure: Performance comparison with prior work on Morph [4, 6, 10, 11]

- CAAE: Age Progression/Regression by Conditional Adversarial Autoencoder [1].
- Before CAAE, Most existing face aging works require paired samples, which is extremely difficult to collect.
- CAAE first introduces Conditional GANs into Face Aging via *traversing in the latent space on the face manifold $\mathcal{M}$*.
- Colors indicate correspondence of personality. Arrows and circle points denote the traversing direction and steps, respectively. Solid arrows direct traversing along the age axis while preserving the personality. The dotted arrow performs a traversing across both the age and personality axes.

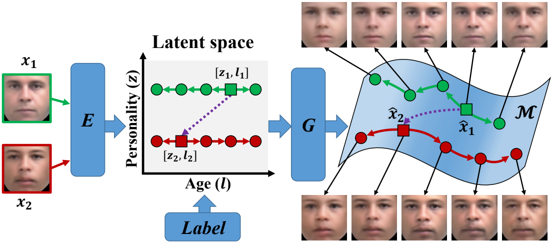

Figure: Illustration of traversing on the face manifold $\mathcal{M}$ [1].

CAAE consists of:

- encoder E maps the input face to a vector z.
- generator G outputs synthesized faces with z concatenated one-hot label l as input.
- discriminator D_z imposes the uniform distribution on z.
- discriminator D_{img} forces the output face to be photo-realistic and plausible for a given age label.

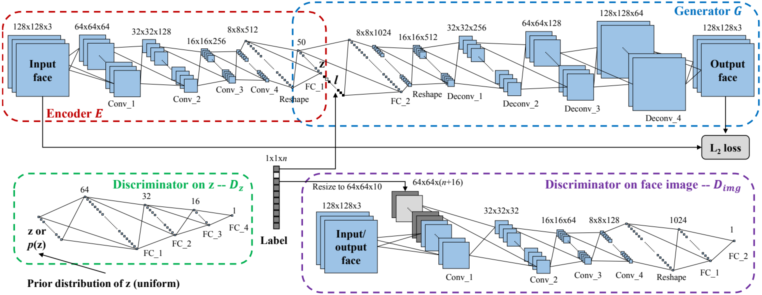

Figure: Structure of the proposed CAAE network [1].

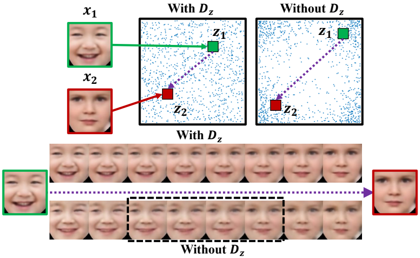

Figure: Effect of D_z [1].

- illustrated z in a 2-D space, which indicates traversing in manifold $\mathcal{M}$.
- D_z forces z to a uniform distribution. Without D_z, z may present sparse distribution along the path of traversing, *i.e.* "holes", causing the generated face to look unreal.

CAAE

Drawbacks: synthesized faces tend to loss their identity information and become blurry due to

- downsampling into a feature vector;
- total variation loss: $\mathcal{L}_{\mathcal{TV}} = \sum_{i,j} |y_{i+1,j} - y_{i,j}| + |y_{i,j+1} - y_{i,j}|$ is used to smooth output and remove the ghosting artifacts.

Figure: Comparison to prior works of face aging [1].

- using Pixel-wise L2 loss to preserve identity, which causes strong ghosting artifacts.
- cannot capture other changes of faces, *e.g.* hair color.

PyramidGAN

- generator G learns the age transformation.
- Discriminator D encourages generated faces to be indistinguishable which is a pyramid hierarchical architecture based on ϕ_{age}: pre-trained VGG-16 structure on CACD. Features from multi-level are used to train Least Squares GAN [12]. ϕ_{age} captures the properties gradually from exact pixel values to high-level age-specific semantic information, handling aging effect generation in a fine-grained way.
- Pixel Level Loss, *i.e.* L2 Loss is applied to preserve background and identity information while Pre-trained deep face descriptor [13] keeps the person-dependent properties.

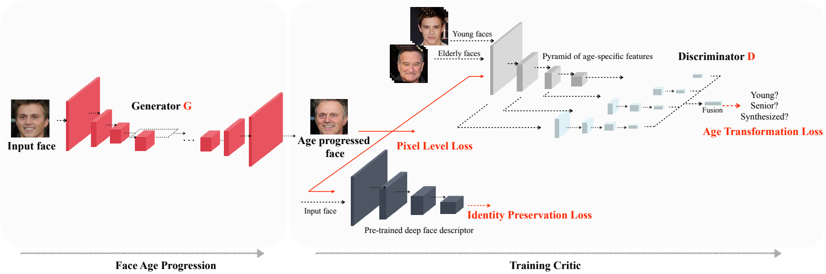

Figure: Framework of PyramidGAN [2].

IPCGAN

- Different from the two methods above-mentioned, IPCGAN [3] adpots a pre-trained AlexNet $h(\cdot)$, chooses a proper layer *i.e. conv*5 of AlexNet, and then applys L2 loss between features of ground truth and synthesized faces:

$$L_{identity} = \sum_{x \in p_x(x)} \left\| h(x) - h\left(G(x|C_t) \right\|^2$$

- Age classifier is same AlexNet as $h(\cdot)$, which is trained on CACD and forces synthesized faces into target age group. Additionally, Discriminator D is simliar with [14] while injecting the conditional feature maps into first convlution layer.

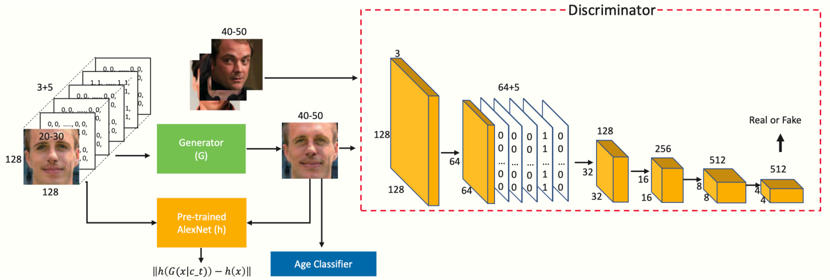

Figure: Pipeline of our proposed IPCGANs for face aging [3].

- AWGAN: Attribute-aware Face Aging with Wavelet-based Generative Adversarial Networks [4].
- Motivation: existing approaches mainly focus on face aging itself while neglecting other critical conditional information of the input (*e.g.* facial attributes), thus would mislead the model into learning translations other than aging, causing serious ghosting artifacts and even incorrect facial attributes in generation results.

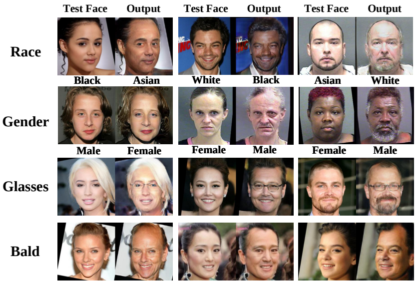

Figure: mismatched facial attributes generated by face aging model [4].

- Facial attribute are embedded into both generator *G* and discriminator *D*, and an resiual connection between input and synthesized faces is established since face aging could be considered as rendering aging effects conditioned on the input young face image.
- Wavelet packet transform(WPT) is adpoted to capture age-related textural features in multi-level like [2].

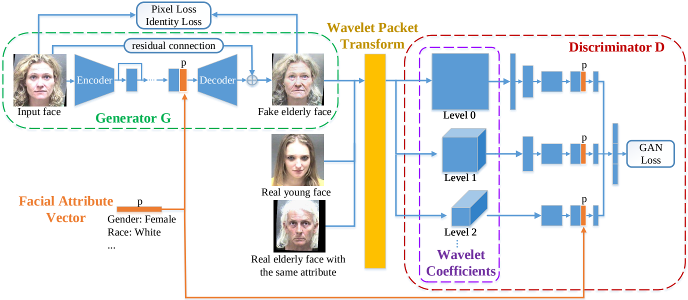

Figure: Overview of the proposed Attribute-aware Wavelet-based GAN [4].

- Compared to extracting multi-scale features as in [2], the advantage of using WPT is that the computational cost is significantly reduced since calculating wavelet coefficients could be regarded as forwarding through a single convolutional layer.

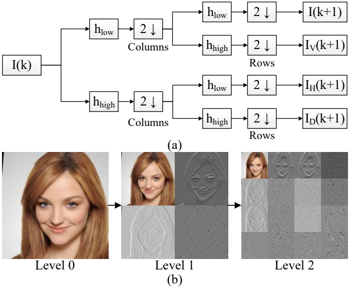

Figure: Demonstration of wavelet packet transform [4].

- facial attribute *e.g.* moustache and ghosting artifacts.
- incorporating FAE suppresses the undesired facial attribute drift by reducing the matching ambiguity.
- WPT reduces the ghosting artifacts and discrepancies with target age.

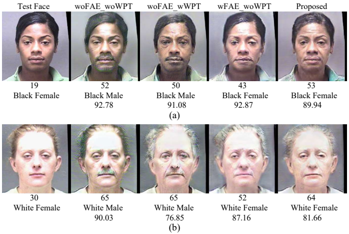

Figure: Ablation study of facial attribute embedding (FAE) and wavelet packet transform (WPT) [4].

Spatial Attention GAN

- Spatial Attention GAN: Age Progression and Regression with Spatial Attention Modules [5].
- L2 Loss usually causes synthesized faces to be blurry and strongly ghosted [14], thus in [15] proposed an attention-based method to improve facial expression synthesis.
- The final output is combined by an attention mask $\mathcal{A}$ and a color mask $\mathcal{C}$ like:

$$\mathcal{I}_{y_f} = (1 - \mathcal{A}) \cdot \mathcal{C} + \mathcal{A} \cdot \mathcal{I}_{y_o}$$

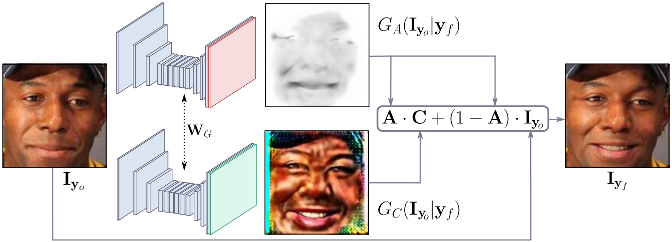

Figure: Illustration of Attention-based generator [15].

Spatial Attention GAN

[5] designs an GAN with Spatial Attention Modules for Face Aging to do both age progression and regression.

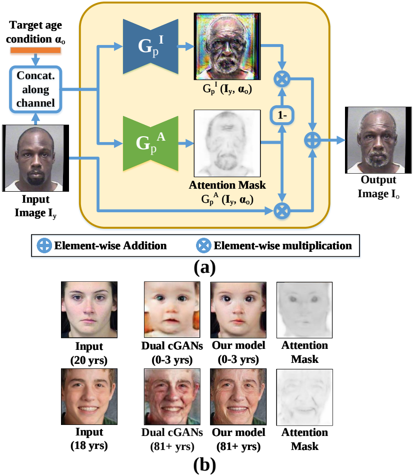

Figure: The age progressor G_p. (a) The detailed structure of G_p. (b) Sample results generated by Spatial Attention GAN [15] and Dual cGANs[16].

To preserve identity information in face generation, some prior works introduce CycleGAN [17], that's to say:

$$G\left(G\left(x_0, k\right), -k\right) \approx x_0$$

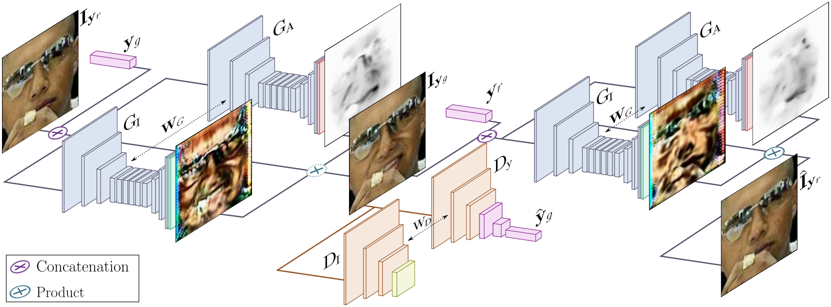

Figure: Example of CycleGAN [17] in face generation [15].

Spatial Attention GAN

- (a) G_p and G_r perform age progression and regression given the conditional age vector α_o and α_y , respectively.
- Reconstruction loss is used to ensure that personalized features in the input image is preserved in the output.
- (b) D_p and D_r are discriminators designed to distinguish real images from synthetic ones and estimate the age of the input face image, and they are involved in the age progression cycle and regression cycle, respectively.

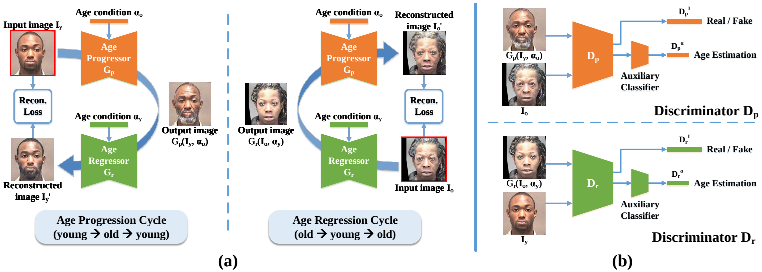

Figure: The framework of Spatial Attention GAN [15]

Thank You

- 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*, 2017, pp. 5810–5818.

- Hongyu Yang, Di Huang, Yunhong Wang, and Anil K Jain,
  "Learning face age progression: A pyramid architecture of gans,"
  in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 31–39.

- 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*, 2018, pp. 7939–7947.

- Yunfan Liu, Qi Li, and Zhenan Sun,

"Attribute-aware face aging with wavelet-based generative adversarial networks,"
in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11877–11886.

Qi Li, Yunfan Liu, and Zhenan Sun,
"Age progression and regression with spatial attention modules,"
arXiv preprint arXiv:1903.02133, 2019.

Karl Ricanek and Tamirat Tesafaye,
"Morph: A longitudinal image database of normal adult age-progression,"
in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*. IEEE, 2006, pp. 341–345.

Bor-Chun Chen, Chu-Song Chen, and Winston H Hsu,
"Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset,"

IEEE Transactions on Multimedia, vol. 17, no. 6, pp. 804–815, 2015.

Ira Kemelmacher-Shlizerman, Supasorn Suwajanakorn, and Steven M Seitz,
"Illumination-aware age progression,"
in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3334–3341.

Jinli Suo, Song-Chun Zhu, Shiguang Shan, and Xilin Chen,
"A compositional and dynamic model for face aging,"
IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 3, pp. 385–401, 2009.

Jinli Suo, Xilin Chen, Shiguang Shan, Wen Gao, and Qionghai Dai,
"A concatenational graph evolution aging model,"
IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 11, pp. 2083–2096, 2012.

- Hongyu Yang, Di Huang, Yunhong Wang, Heng Wang, and Yuanyan Tang,
  "Face aging effect simulation using hidden factor analysis joint sparse representation,"
  IEEE Transactions on Image Processing, vol. 25, no. 6, pp. 2493–2507, 2016.
- Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley,
  "Least squares generative adversarial networks,"
  in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2794–2802.
- Omkar M Parkhi, Andrea Vedaldi, Andrew Zisserman, et al.,
  "Deep face recognition.,"
  in *bmvc*, 2015, vol. 1, p. 6.
- Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros,
  "Image-to-image translation with conditional adversarial networks,"

in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

Albert Pumarola, Antonio Agudo, Aleix M Martinez, Alberto Sanfeliu, and Francesc Moreno-Noguer,
"Ganimation: Anatomically-aware facial animation from a single image,"
in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 818–833.

Jingkuan Song, Jingqiu Zhang, Lianli Gao, Xianglong Liu, and Heng Tao Shen,
"Dual conditional gans for face aging and rejuvenation.,"
in *IJCAI*, 2018, pp. 899–905.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros,
"Unpaired image-to-image translation using cycle-consistent adversarial networks,"
in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.