Age Progression/Regression by Conditional Adversarial Autoencoder

Zhifei Zhang, Yang Song, and Hairong Qi

THE UNIVERSITY OF TENNESSEE KNOXVILLE

IEEE 2017 Conference on Computer Vision and Pattern Recognition

CVPR July 21-26 HONOLULU 2017

Motivation:

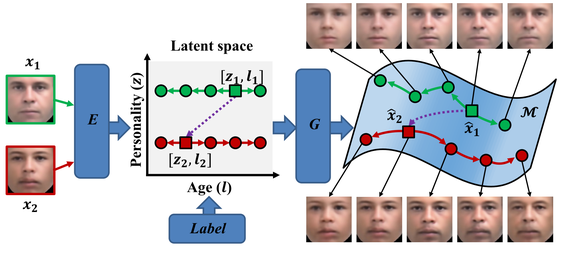
Existing works:

- Group-wised learning
- Query with label
- Step-to-step transition
- Progression by adding texture

Ours:

- Joint learning
- Query without label
- One-step transition
- Bidirectional generation

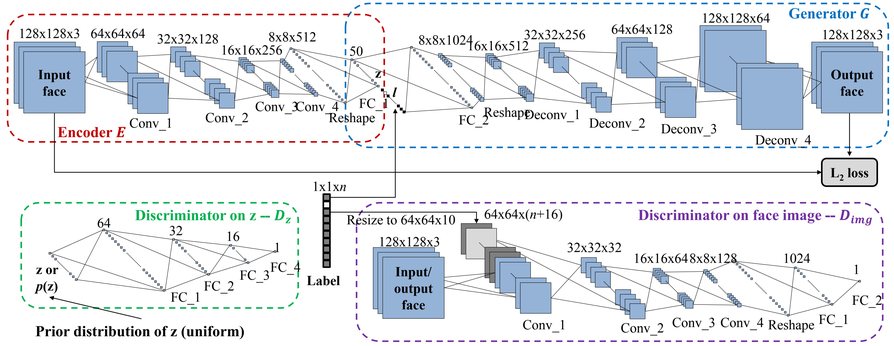
Main Idea --- Manifold Traversing:



Assumptions:

- The faces lie on a manifold (M)
- Ages and personality change smoothly on the manifold
- Traversing on the manifold corresponds to age/personality transformation.

Approach:



Training --- mainly three losses:

- Reconstruction loss $L_{const}(E, G)$ $\quad L_{const}(E,G) = \|x - G(E(x), l)\|_2$
- Adversarial loss on latent variable $L_{adv\_z}(E, D_z)$

$$L_{adv\_z}(E, D_z) = \mathbb{E}_{x \sim p_{data}}\left[\log\left(1 - D_z(E(x))\right)\right] - \mathbb{E}_{z \sim noise}[\log D_z(z)]$$

- Adversarial loss on image $L_{adv\_img}(G, D_{img})$

$$L_{adv\_img}(G, D_{img}) = \mathbb{E}_{x \sim p_{data}}\left[\log\left(1 - D_{img}\left(G(E(x), l)\right)\right) - \log D_{img}(x, l)\right]$$

Testing --- face with arbitrary labels (target ages):

- Generate older or younger faces conditioned on the personality of input faces and given labels

$$\hat{x} = G(E(x), l), \qquad l = 0, 1, \cdots, 9$$

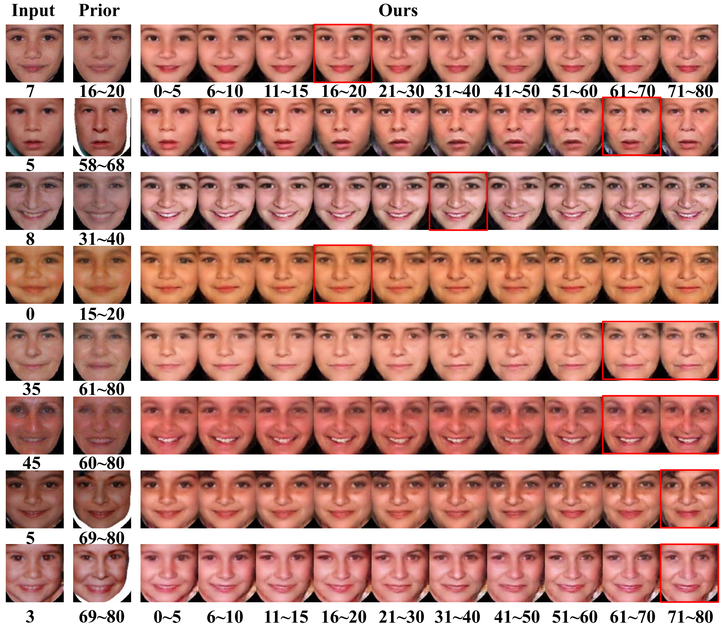
Results:

Data Collection:

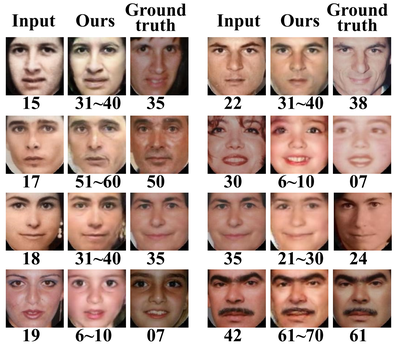
- We create a dataset named **UTKFace**, consisting of over 20,000 faces with long age span (0~80)
- MORPH dataset. 55,000 faces of 13,000 subjects from 16 to 77 years old
- CACD dataset. 163,446 faces of 2,000 subjects from 16 to 62 years old
- We randomly select 10,670 faces with a uniform distribution on gender and age.
- We use the face detection algorithm with 68 landmarks to crop out and align the faces

Qualitative Comparison:

- Compare to prior --- the BEST result achieved by existing works



- Compare to ground truth



Quantitative Comparison:

- Totally, 124 subjects, and 4716 valid votes
- Each survey has 45 questions randomly selected from over 1000 questions

	Better than prior?	Same to ground truth?
Yes	52.77%	48.38%
No	28.99%	29.58%
Not sure	18.24%	22.04%