“FACE AGING WITH CONDITIONAL GENERATIVE ADVERSARIAL NETWORKS”

(ORAL PRESENTATION @ ICIP 2017)

Authors: G. Antipov, M. Baccouche, J.-L. Dugelay
Date / place: September 20th 2017, Beijing, China

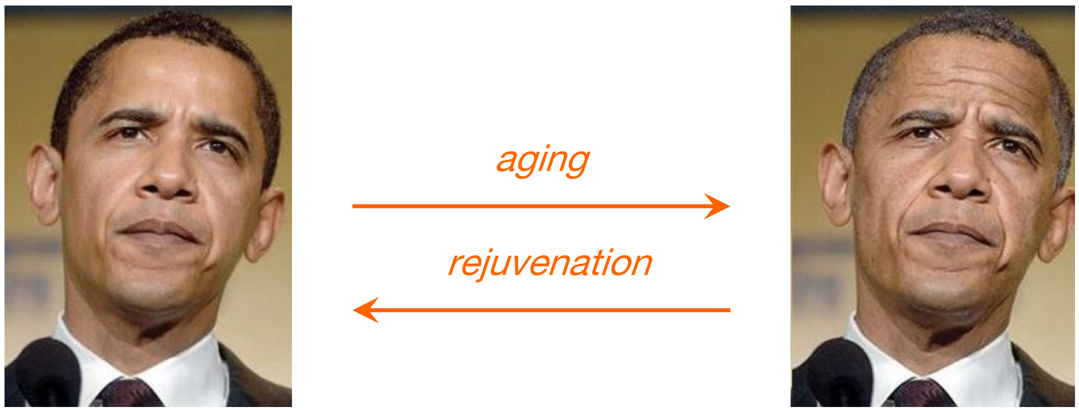
orange™

EURECOM
Sophia Antipolis

Face Aging/Rejuvenation: Introduction and Motivation

Objective:

Aesthetically rendering a face image with natural aging and rejuvenating effects on the given face of an individual.



(43 y.o.) (55 y.o.)

Motivation to Study:

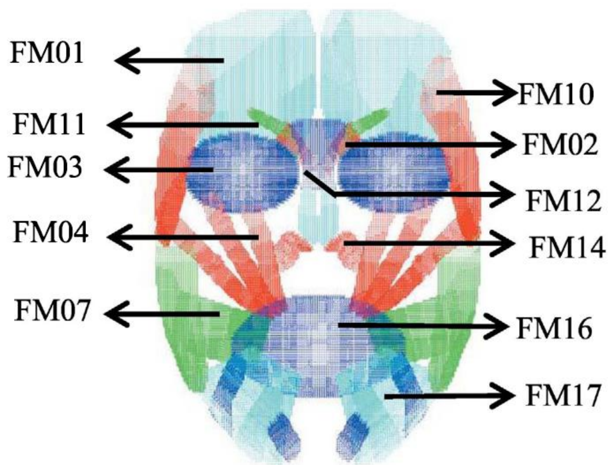
1. <u>Biometrics:</u> age normalization prior to face verification.
2. <u>Police & law enforcement:</u> research of fugitives and lost people.
3. <u>Film making:</u> special effects.
4. <u>Very challenging ill-posed problem.</u>

Previous Work on Face Aging/Rejuvenation

Two families of face aging/rejuvenation studies

Modelling-based

Employ parametric models to simulate the physical aging mechanism of muscles, skin and skull of an individual.

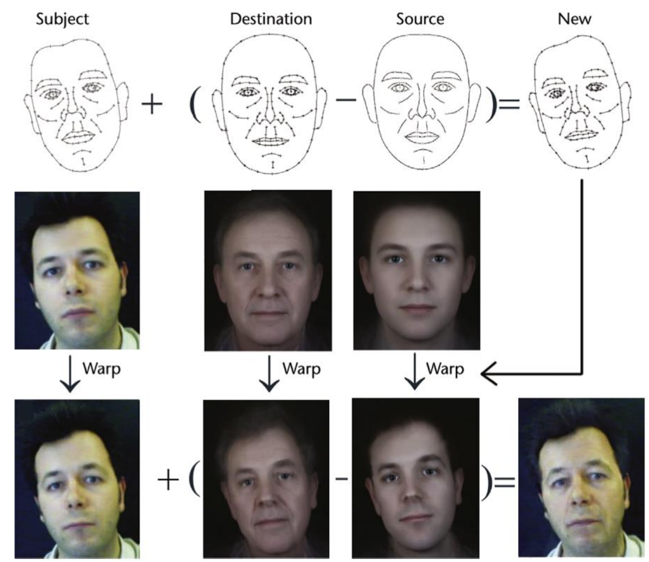


Shihfeng et al. "Aging simulation using facial muscle model"

- \+ Better preserves identity
- − Require training faces of the same person at different ages

Prototype-based

Define average faces calculated on training images of certain age categories as their prototypes.



Burt and Perrett. "Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information"

- \+ Fast; do not require same-person training data
- − Prone to lose identity information

− Designed uniquely for aging/rejuvenation (cannot be adapted to model other face variations)

Discriminative and Generative Models

observations (x)

Discriminative Model

"cat"

Discriminative Model

"dog"

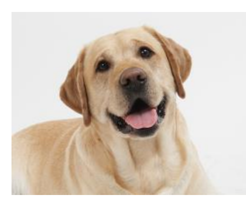
labels (y)

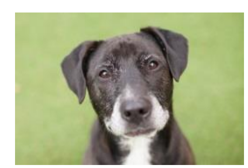
Discriminative: estimation of $p(y|x)$

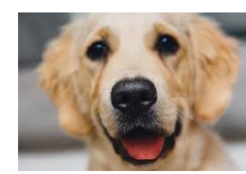
Generative: estimation of $p(x, y)$

Generative Model (of cats)

Generative Model (of dogs)

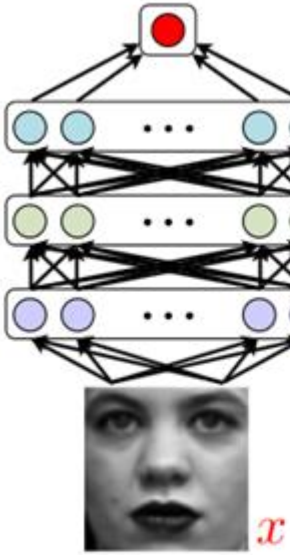






Generative Adversarial Networks (GANs) [1]

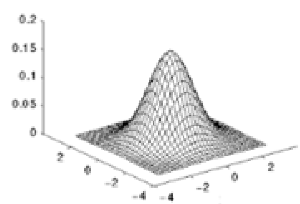
Latent space: N^z
Space of natural faces: N^x



D tries to output 1

Discriminator D

$x \in N^x$
$x \sim p_{data}$



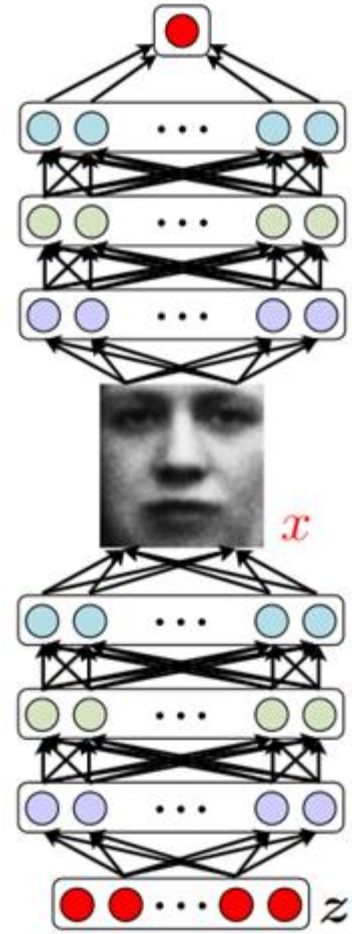
D tries to output 0

Discriminator D

Synthetic faces

Generator G

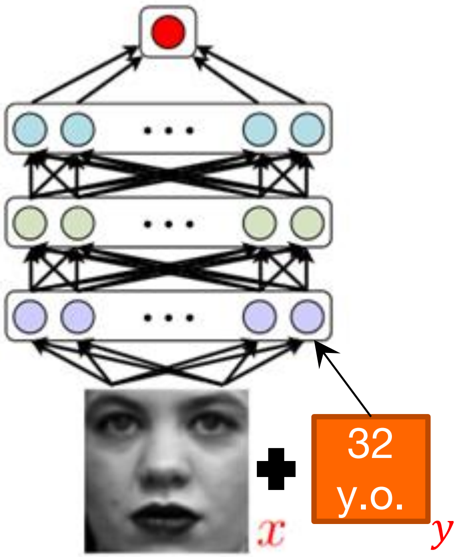
$z \in N^z$
$z \sim N(0, I)$



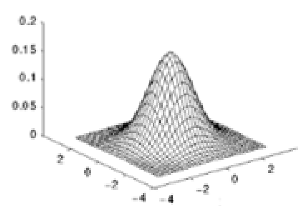
$$\min_G \max_D V(D,G) = \mathbb{E}_{x \sim p_{data}}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))]$$

Conditional Generative Adversarial Networks (cGANs) [2]

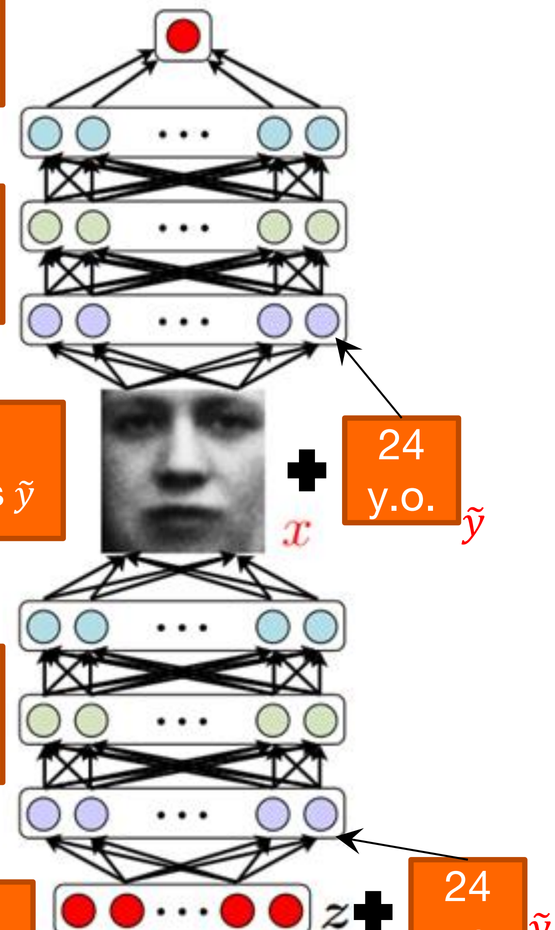
Latent space: N^z
Space of natural faces: N^x
Space of conditions: N^y



$x \in N^x, y \in N^y$
$x, y \sim p_{data}$
→ Discriminator D
→ D tries to output 1



$z \in N^z, \tilde{y} \in N^y$
$z \sim N(0, I), \tilde{y} \sim p_y$
→ Generator G
→ Synthetic faces + conditions $\tilde{y}$
→ Discriminator D
→ D tries to output 0

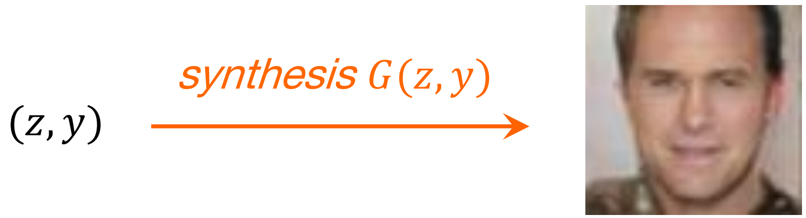


$$\min_G \max_D V(D, G) = \mathbb{E}_{x,y \sim p_{data}}[\log D(x, y)] + \mathbb{E}_{z \sim p_z(z), \tilde{y} \sim p_y}[1 - \log D(G(z, \tilde{y}), \tilde{y})]$$

Main Idea: Face Aging/Rejuvenation with cGANs

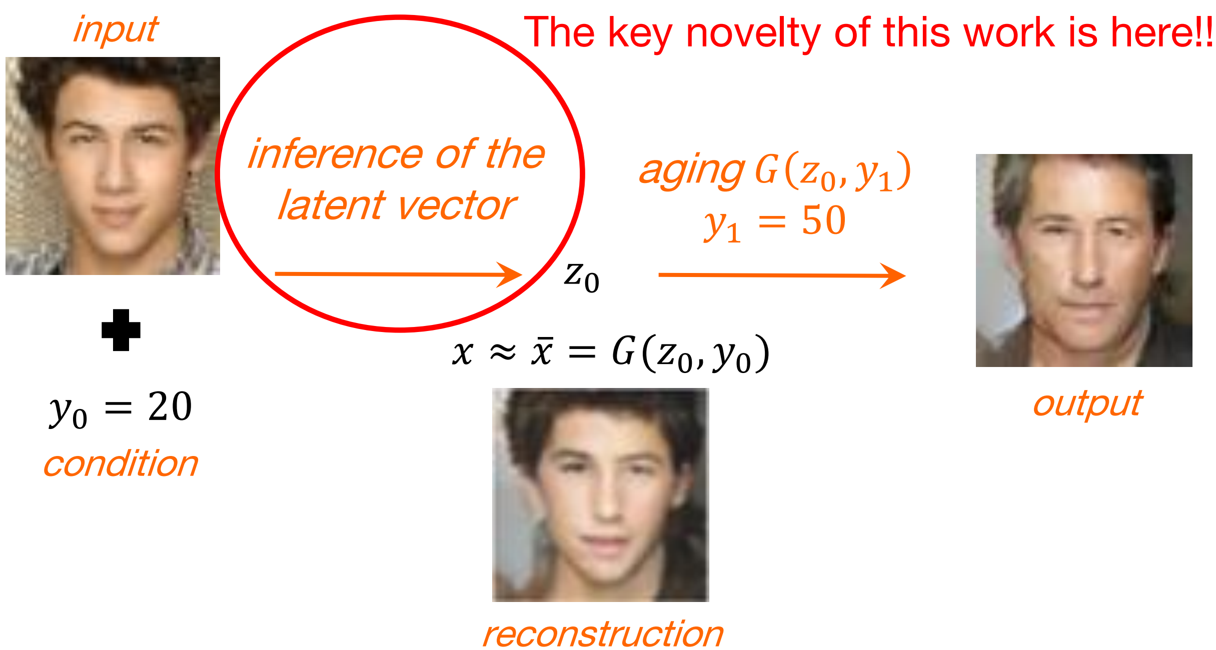
Step 1 (offline).

Training of Age-cGAN so that its generator G can be used for synthesis of random faces of the given age.



Step 2 (online).

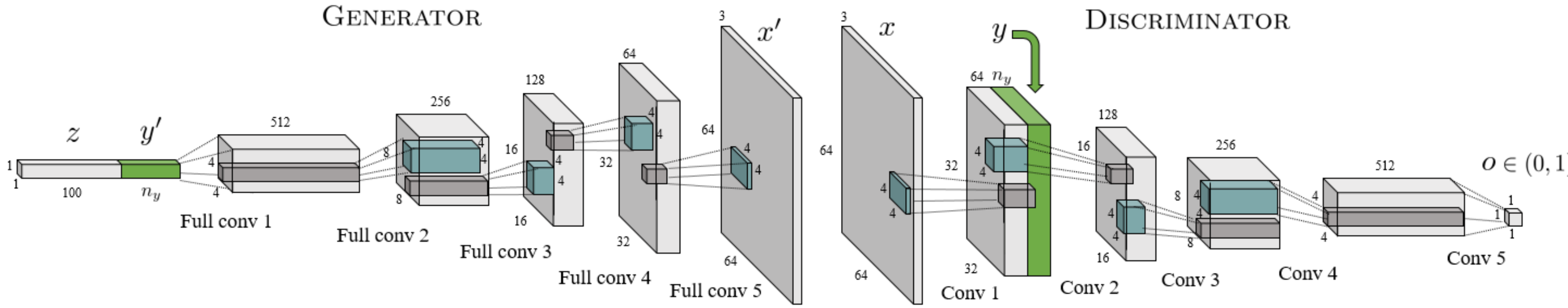
Given an input face x of age y_0, inferring a latent vector z_0 which allows to approximate the input with the designed generator G: $x \approx \bar{x} = G(z_0, y_0)$. Then switching the input age y_0 by the target one y_1 to produce the resulting aged/rejuvenated synthetic output face $G(z_0, y_1)$.



orange™

EURECOM
Sophia Antipolis

Design of Age-cGAN



Radford et al. "Unsupervised representation learning with deep convolutional generative adversarial networks"

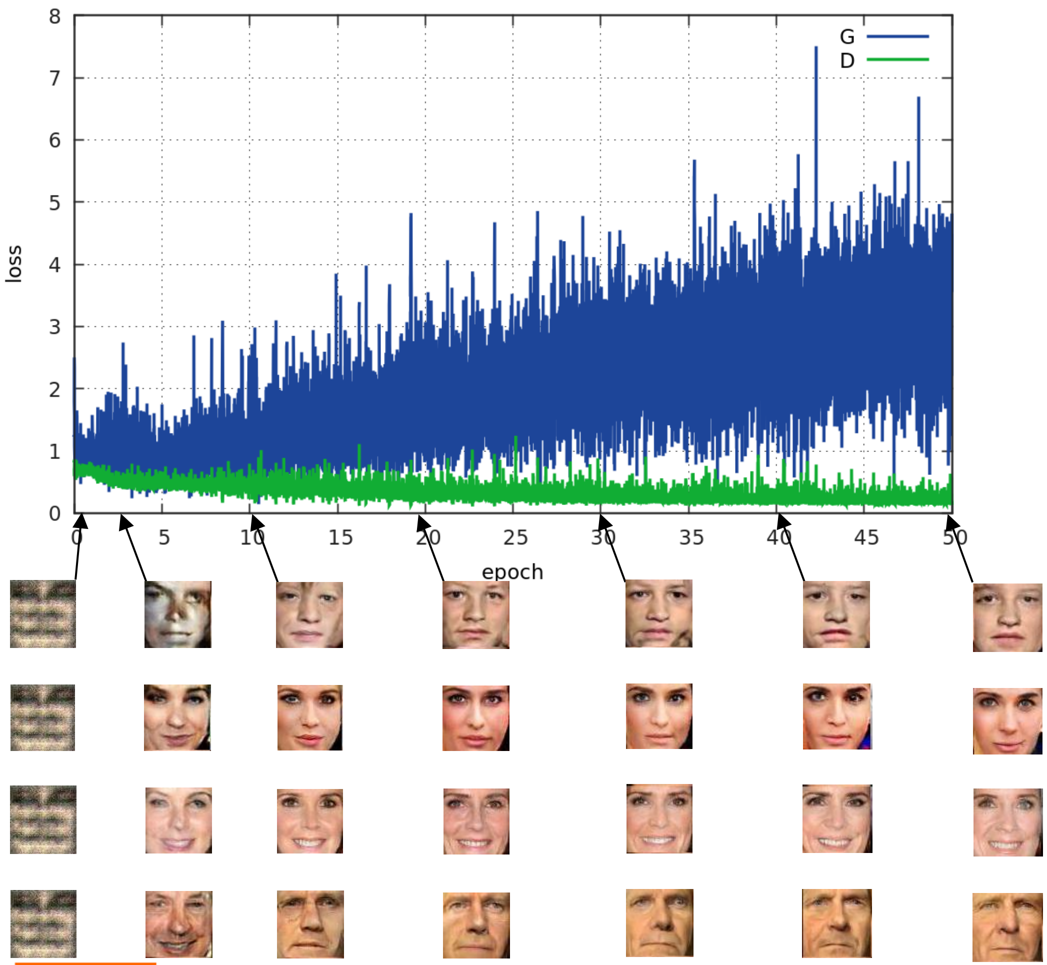
Design of Age-cGAN:

1. Architectures of G and D are taken from DCGAN [3] to ensure stable convergence.
2. The conditions y are fed to G and D according to the conclusions of [4].
3. Conditions y are one-hot binary vectors of the dimension 6, because there are 6 age categories: "0-18", "19-29", "30-39", "40-49", "50-59" and "60+".
4. 120K from the IMDB-Wiki dataset [5] were used for training.
5. Two types of face crops have been tested, namely, "face-only" and "face+40%":

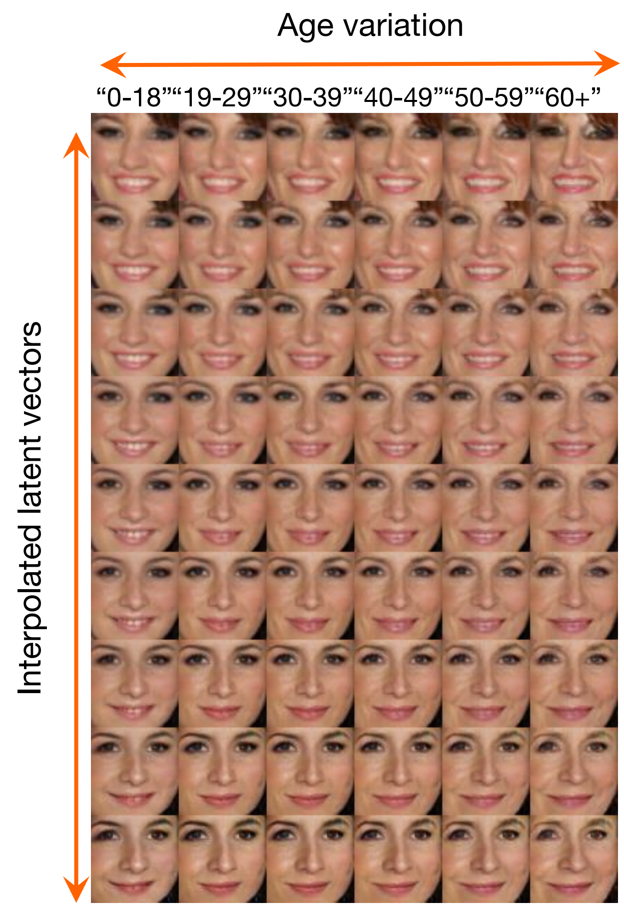
 

Learning of the Synthetic Face Manifold

Training progress of Age-cGAN



Part of the learned synthetic face manifold

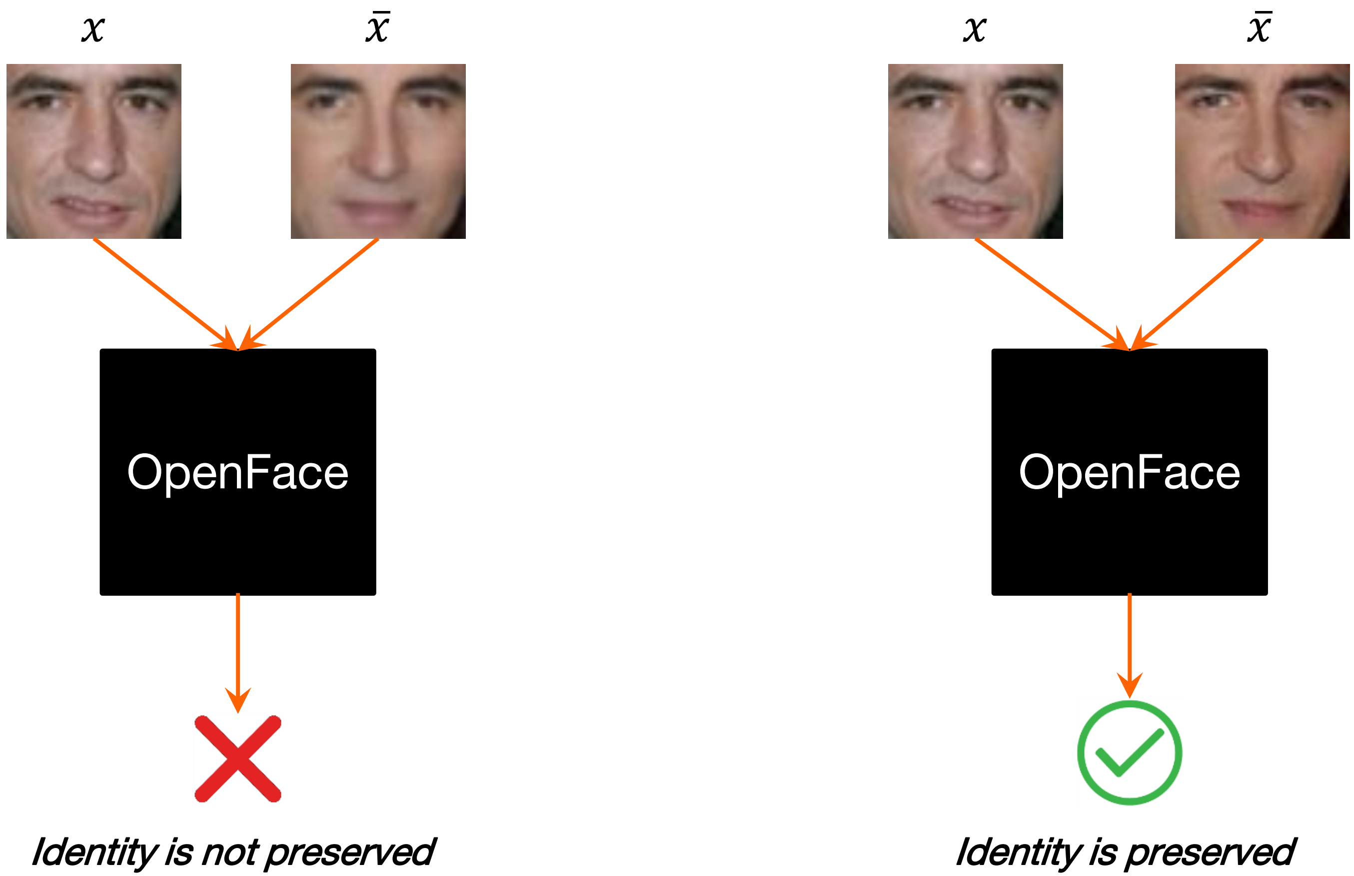


orange™

EURECOM
Sophia Antipolis

Objective Evaluation of the Quality of Face Reconstruction

OpenFace face recognition software [6] is used as a **black box**: given a pair of an input face x and its reconstruction $\bar{x}$, the software estimates whether the original image identity is preserved in the reconstruction or not.

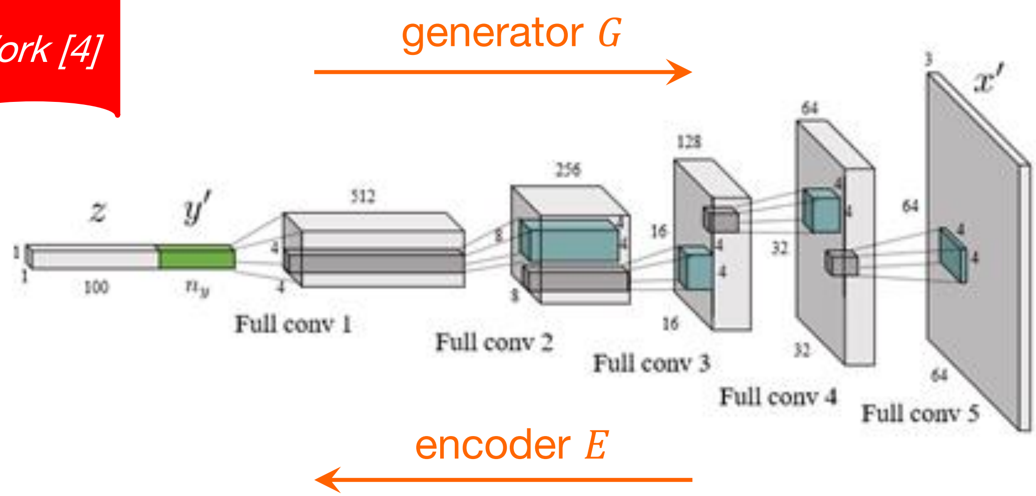


orange™

EURECOM
Sophia Antipolis

Face Reconstruction $\bar{x}_0$ via Training of an Encoder CNN (Theory)

Previous Work [4]



The simplest way to reconstruct a given face x of age y_0 with a cGAN (i.e. to infer the latent vector z_0 so that $x \approx \bar{x} = G(z_0, y_0)$) is **to train an encoder CNN $\boldsymbol{E}$ which approximates inverse the mapping learned by $\boldsymbol{G}$: $\boldsymbol{E} \approx \boldsymbol{G}^{-1}$.**

1. E is trained on 100K of synthetically generated triples $\{z_i, y_i, x_i\}$ where $x_i = G(z_i, y_i)$.

2. E is optimized to minimize $||x_i - G(E(x_i), y_i)||_{L_2}$.

3. Once E is trained, $z_0 = E(x)$, $\bar{x}_0 = G(z_0, y_0)$.

Face Reconstruction $\bar{x}_0$ via Training of an Encoder CNN (Evaluation)

Original x | Reconstruction $\bar{x}_0$

Original x | Reconstruction $\bar{x}_0$

Reconstruction Approach	Face Verification Score
Only encoder E: $\bar{x}_0$	89,0%

Reconstruction Approach	Face Verification Score
Only encoder E: $\bar{x}_0$	53,2%

orange™

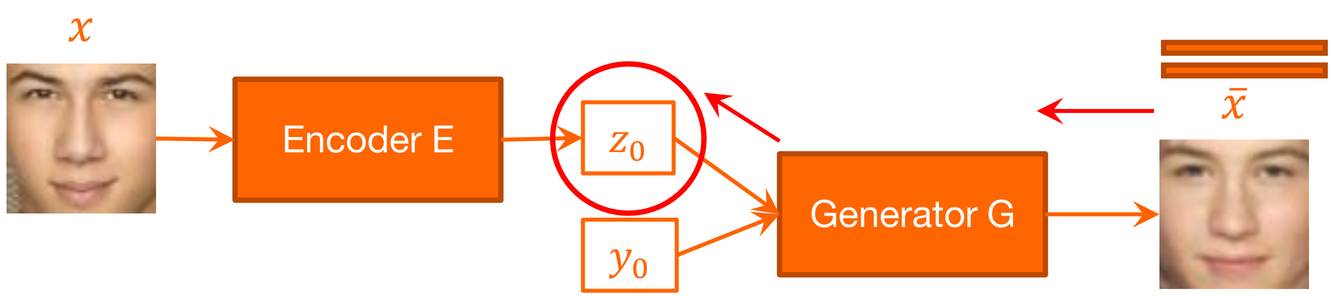
EURECOM Sophia Antipolis

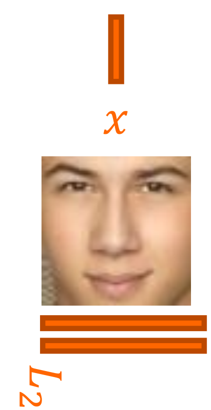
Improved Face Reconstruction $\bar{x}_{pixel}$ via Latent Vector Optimization (Theory)

Previous Work [7]

!

<u>Problem:</u> the reconstructed image is too different from the original one.



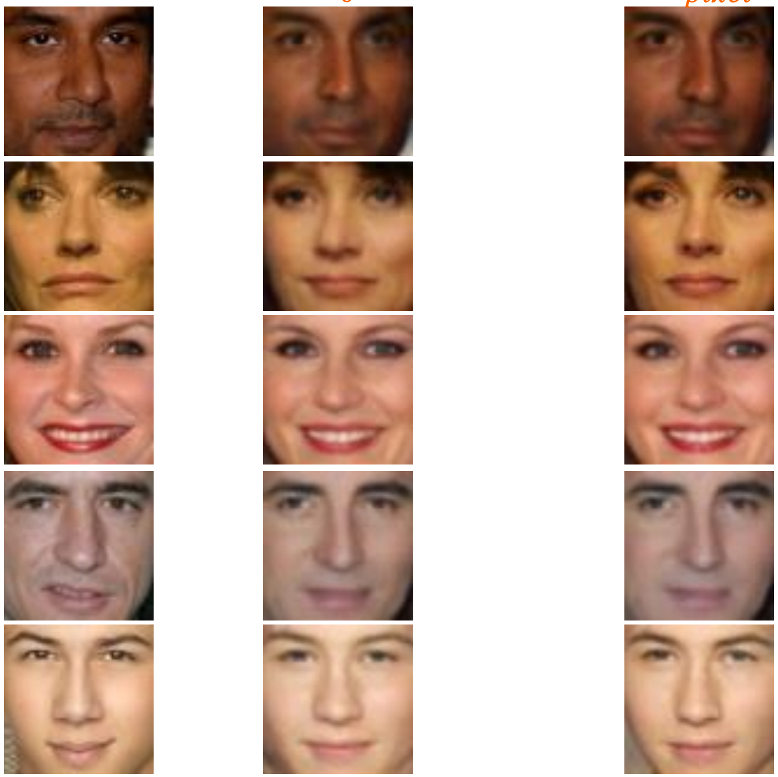


✓ *<u>Solution:</u> due to the fact that G is differentiable, the latent vector z can be optimized to minimize the pixel-level difference between the original and the reconstructed faces using L-BFGS-B [8].*

$$z^*_{pixel} = \underset{z}{argmin}\|x - \bar{x}\|_{L_2}$$

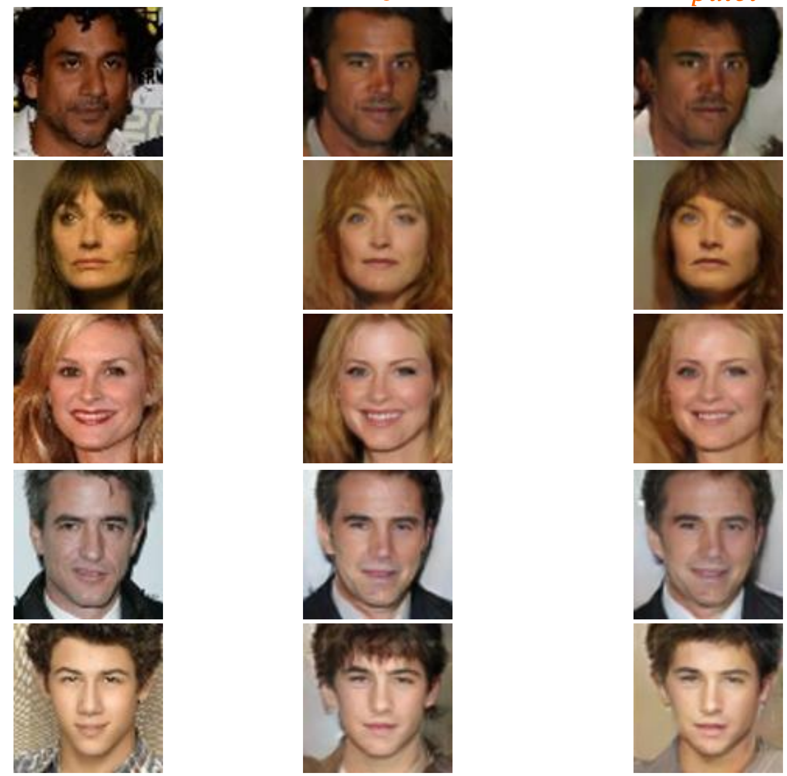
Improved Face Reconstruction $\bar{x}_{pixel}$ via Latent Vector Optimization (Evaluation)

Original x | Reconstructions: $\bar{x}_0$, $\bar{x}_{pixel}$



Reconstruction Approach	Face Verification Score
Only encoder E: $\bar{x}_0$	89,0%
Pixel-level optimization: $\bar{x}_{pixel}$	94,5%

Original x | Reconstructions: $\bar{x}_0$, $\bar{x}_{pixel}$



Reconstruction Approach	Face Verification Score
Only encoder E: $\bar{x}_0$	53,2%
Pixel-level optimization: $\bar{x}_{pixel}$	59,8%

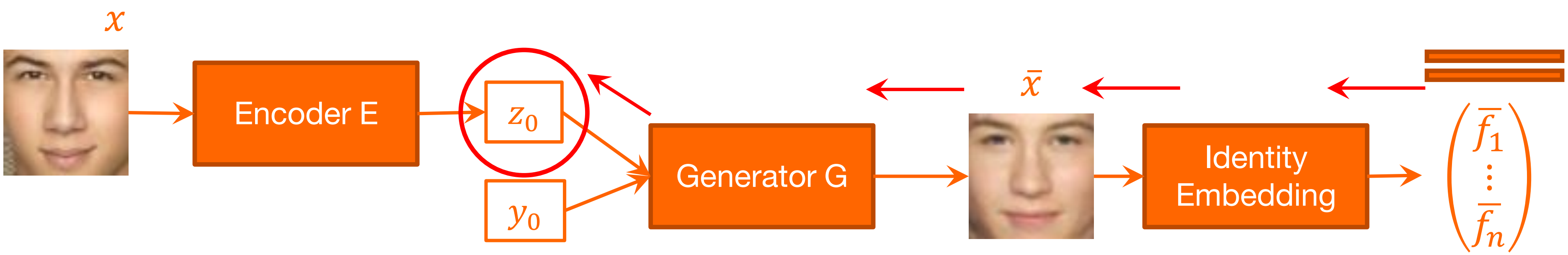
orange™

EURECOM
Sophia Antipolis

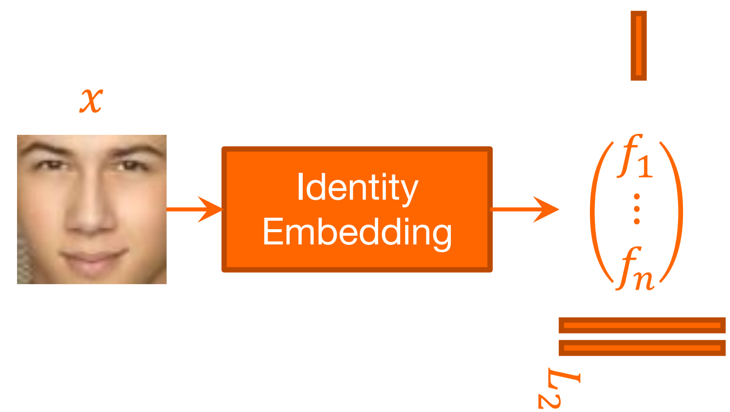
Proposed "Identity-Preserving" Face Reconstruction $\bar{x}_{IP}$ (Theory)

Proposed Approach

<u>Problem:</u> the pixelwise optimization increases the blurriness of faces, but does not really improve the identity preservation.



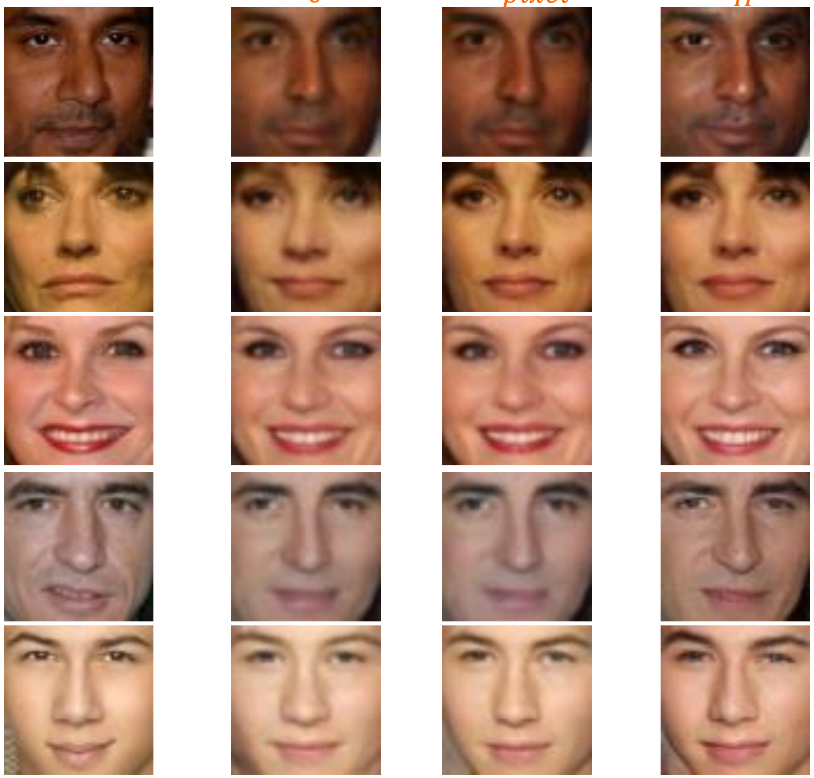
<u>Solution:</u> instead of pixelwise optimization perform an optimization of the face recognition embeddings FR.



$$z_{IP}^{*} = \underset{z}{argmin} \| FR(x) - FR(\bar{x}) \|_{L_2}$$

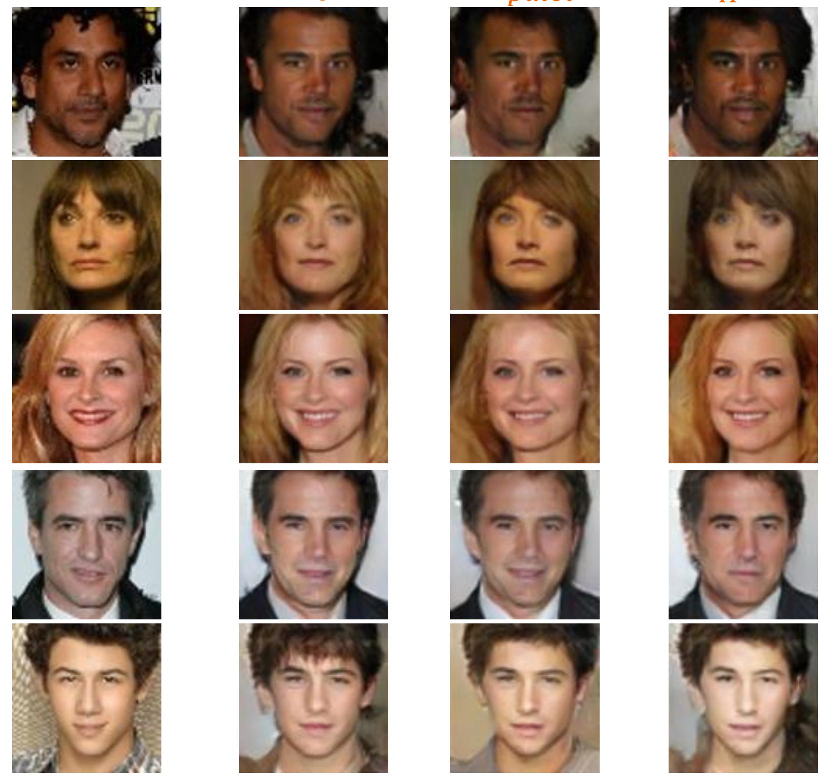
Proposed "Identity-Preserving" Face Reconstruction $\bar{x}_{IP}$ (Evaluation)

Original x | Reconstructions: $\bar{x}_0$, $\bar{x}_{pixel}$, $\bar{x}_{IP}$



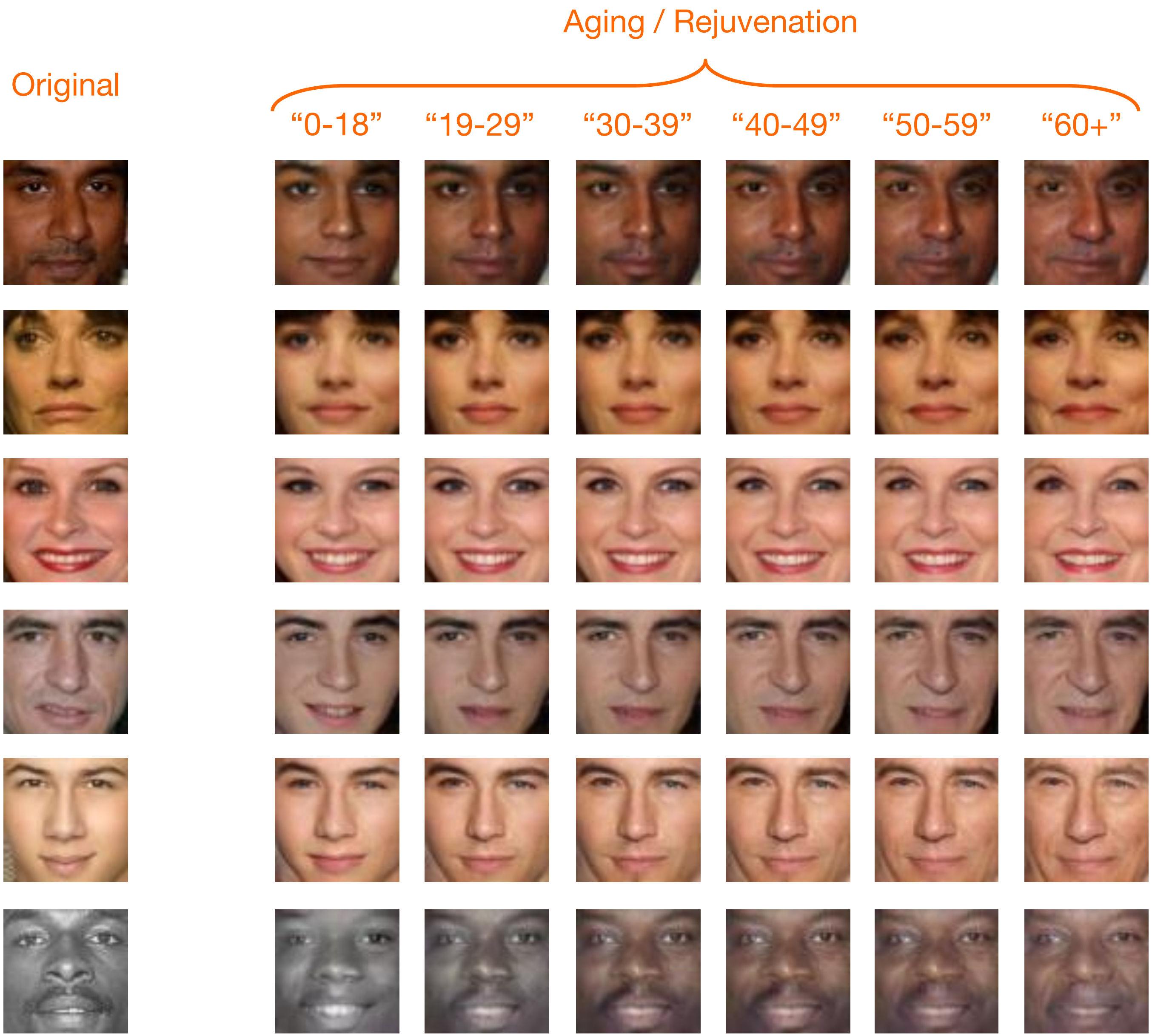
Reconstruction Approach	Face Verification Score
Only encoder E: $\bar{x}_0$	89,0%
Pixel-level optimization: $\bar{x}_{pixel}$	94,5%
"Identity-Preserving" optimization: $\bar{x}_{IP}$	**97,6%**

Original x | Reconstructions: $\bar{x}_0$, $\bar{x}_{pixel}$, $\bar{x}_{IP}$



Reconstruction Approach	Face Verification Score
Only encoder E: $\bar{x}_0$	53,2%
Pixel-level optimization: $\bar{x}_{pixel}$	59,8%
"Identity-Preserving" optimization: $\bar{x}_{IP}$	**82,9%**

Aging/Rejuvenation by the Proposed Approach (1)



orange™

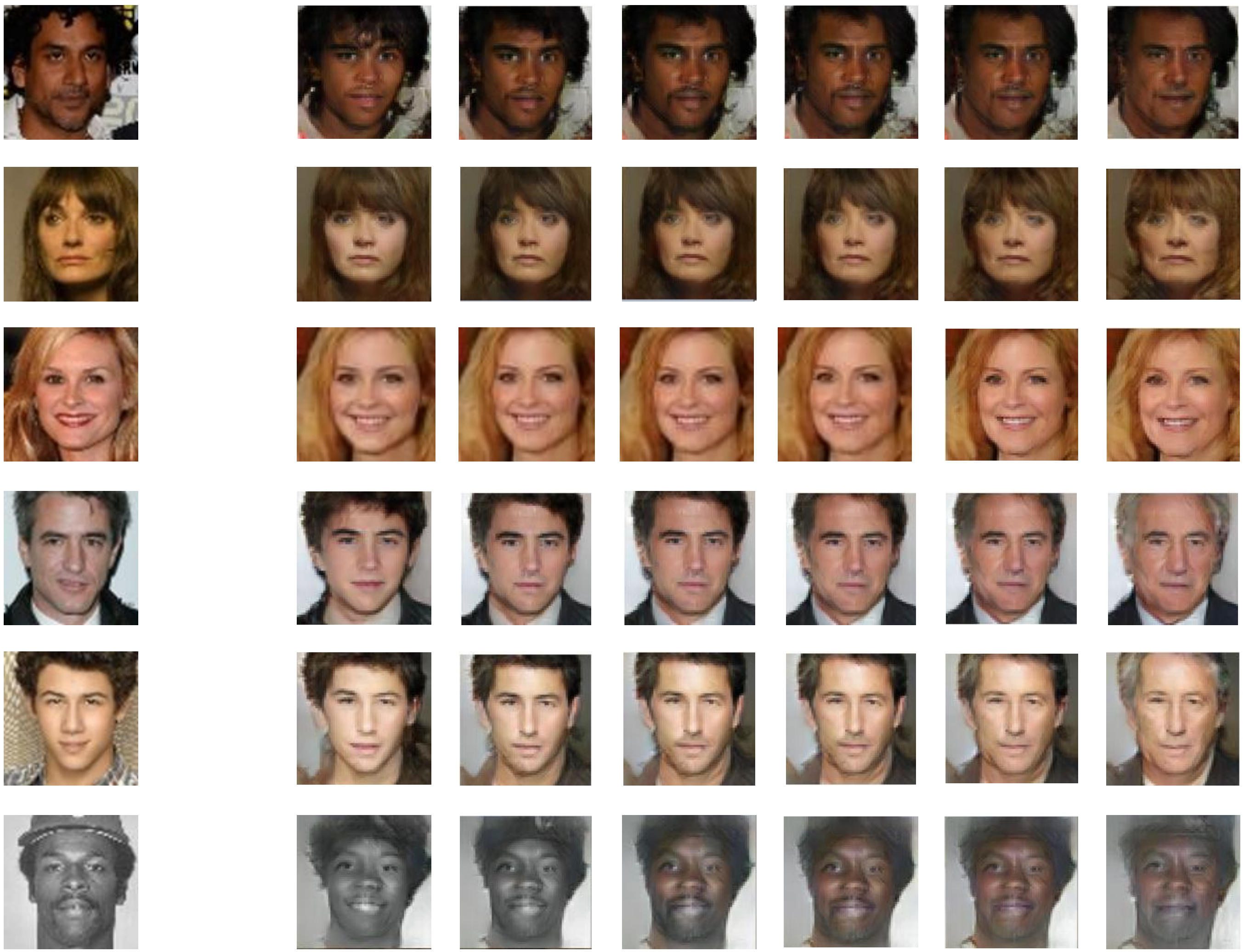
EURECOM
Sophia Antipolis

Aging/Rejuvenation by the Proposed Approach (2)

Aging / Rejuvenation

Original

“0-18” “19-29” “30-39” “40-49” “50-59” “60+”



orange™

EURECOM
Sophia Antipolis

Conclusions

1. We have designed Age-cGAN, the first GAN for age-conditioned face synthesis.

2. We have proposed "Identity-Preserving" approach for face reconstruction with GANs without losing the original person's identity. Our approach is universal in the sense that it can help in all face editing applications, and not just for aging/rejuvenation.

3. The resulting aging/rejuvenation method demonstrates visually convincing results. It has been successfully adapted for improving the cross-age face verification.

4. In our recent paper Boosting Cross-Age Face Verification via Generative Age Normalization" which will be presented at International Joint Conference on Biometrics 2017, we extend Age-cGAN and show that it be applied to improve cross-age face verification.

Thank you! Questions?

Contact: grigory.antipov@orange.com

References

[1] Goodfellow et al. “Generative Adversarial Nets”, NIPS, 2015.

[2] Jon Gauthier. “Conditional generative adversarial nets for convolutional face generation”. Class Project for Stanford CS231N, 2014.

[3] Radford et al. “Unsupervised representation learning with deep convolutional generative adversarial networks”. ICLR, 2016.

[4] Perarnau et al. “Invertible Conditional GANs for image editing”. NIPS Workshop, 2016.

[5] Rothe et al. “Deep expectation of real and apparent age from a single image without facial landmarks”. IJCV, 2016.

[6] Amos et al. “OpenFace: A general purpose face recognition library with mobile applications”. Technical report CMU-CS-16-118, 2016.

[7] Zhu et al. “Generative visual manipulation on the natural image manifold”. ECCV, 2016.

[8] Byrd et al. “A limited memory algorithm for bound constrained optimization”. JSC, 1995.