

# Survey

Yuhang Li

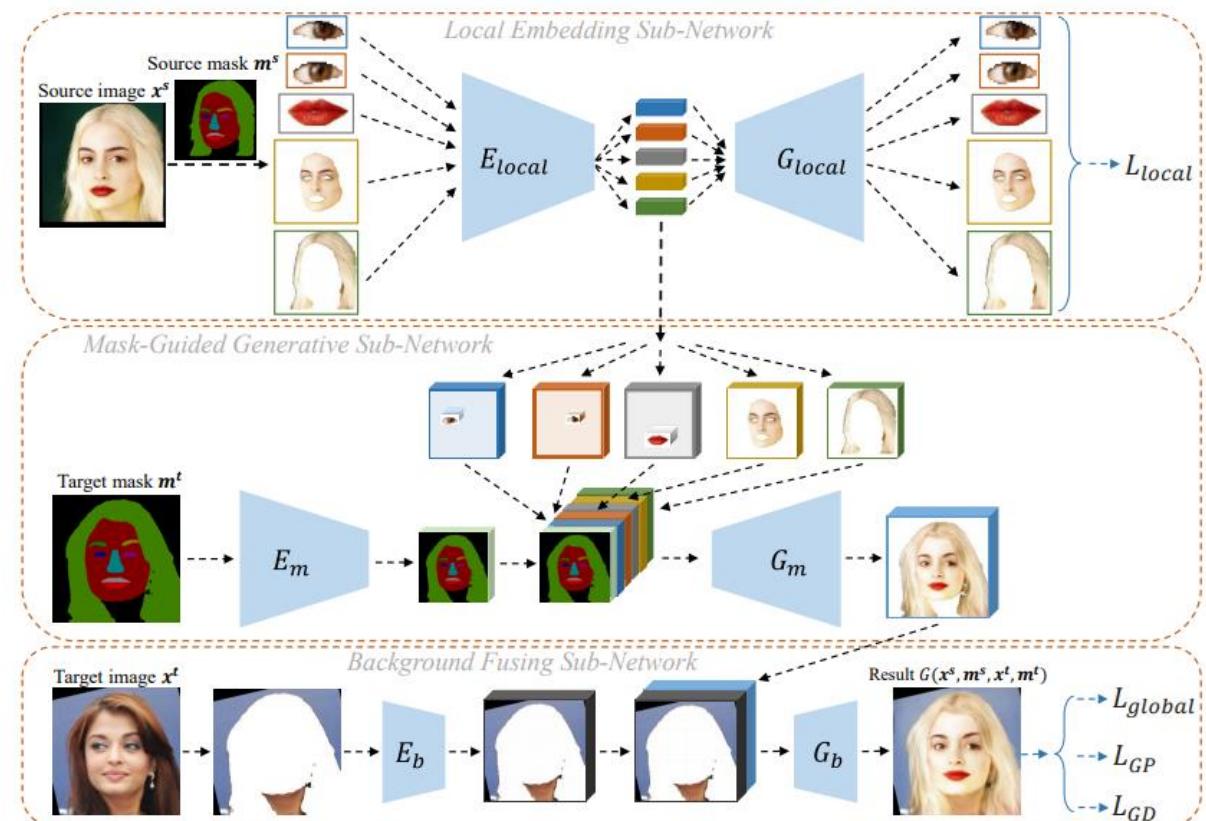
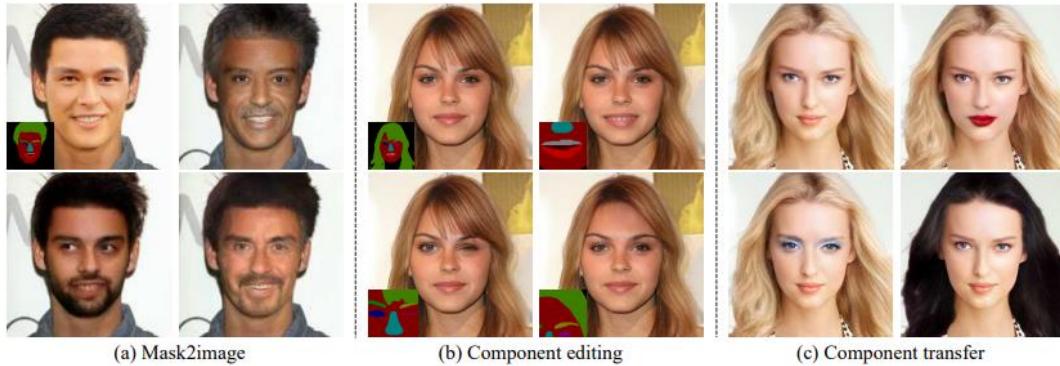
2019-9-9

# Summary

- Task: single/multiple attribute transfer/exchanging/editing/restoring/rotating/aging
- Architecture: En-De, Unet, multiple branches, cycle
- Map: one-to-one, one-to-many, many-to-many
- Branch: solo, dual, cycle
- Attribute carrier: input code, learned code, learned disentangled code, exemplar
- Generated form: rgb image, residual, mask(attention map) + color, 3D face model
- Identity: preserving or not
- Guide: mask, landmarks, edge, exemplar
- 2D/3D
- Other subjects: face detection, face reconstruction, face morphable model

# Mask-Guided Portrait Editing with Conditional GANs

Shuyang Gu<sup>1</sup> Jianmin Bao<sup>1</sup> Hao Yang<sup>2</sup> Dong Chen<sup>2</sup> Fang Wen<sup>2</sup> Lu Yuan<sup>2</sup>  
<sup>1</sup>University of Science and Technology of China      <sup>2</sup>Microsoft Research  
[{gsy777,jmbao}@mail.ustc.edu.cn](mailto:{gsy777,jmbao}@mail.ustc.edu.cn)      [{haya,doch,fangwen,luyuan}@microsoft.com](mailto:{haya,doch,fangwen,luyuan}@microsoft.com)



# GazeCorrection: Self-Guided Eye Manipulation in the wild using Self-Supervised Generative Adversarial Networks

Jichao Zhang \*  
Shandong University

Meng Sun \*  
Shandong University

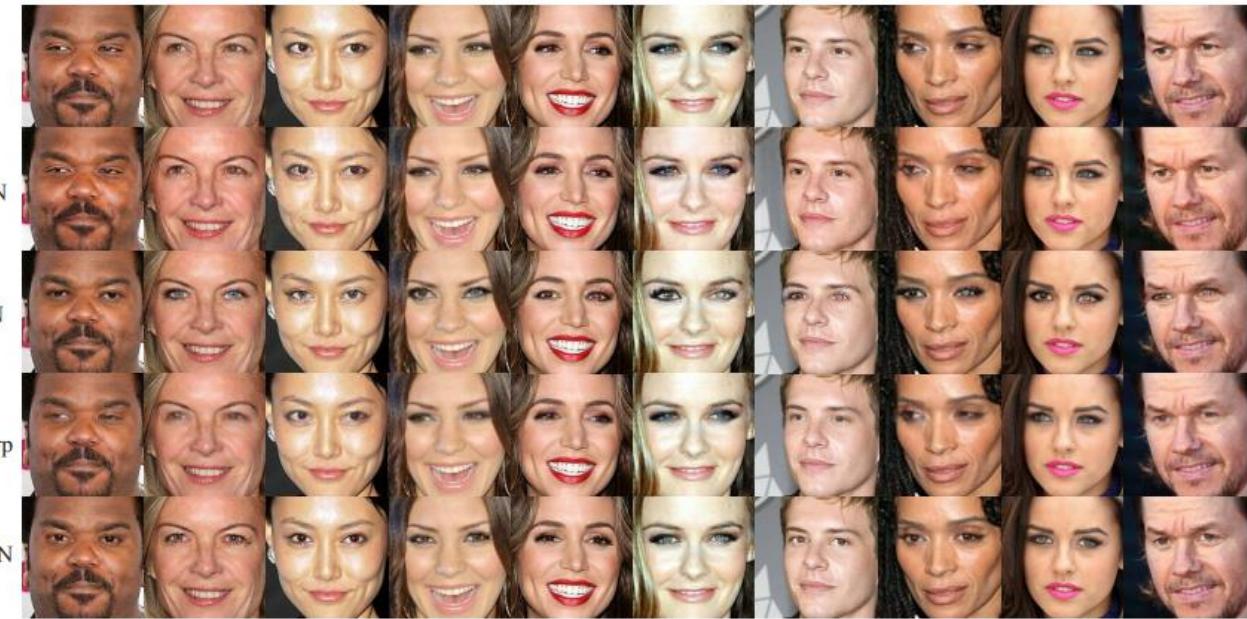
Jingjing Chen \*  
Shandong University

Hao Tang  
University of Trento

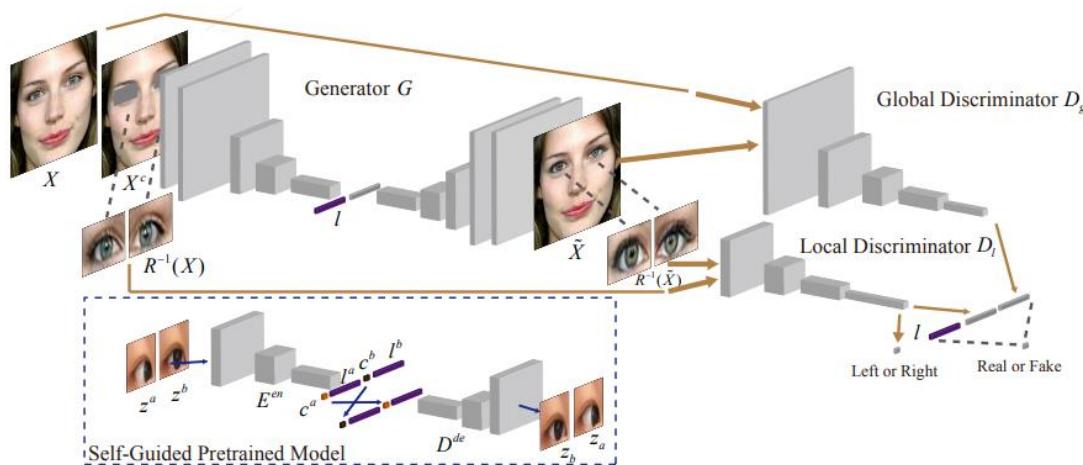
Yan Yan  
Texas State University

Xueying Qin  
Shandong University

Nicu Sebe  
University of Trento



## Eye inpainting



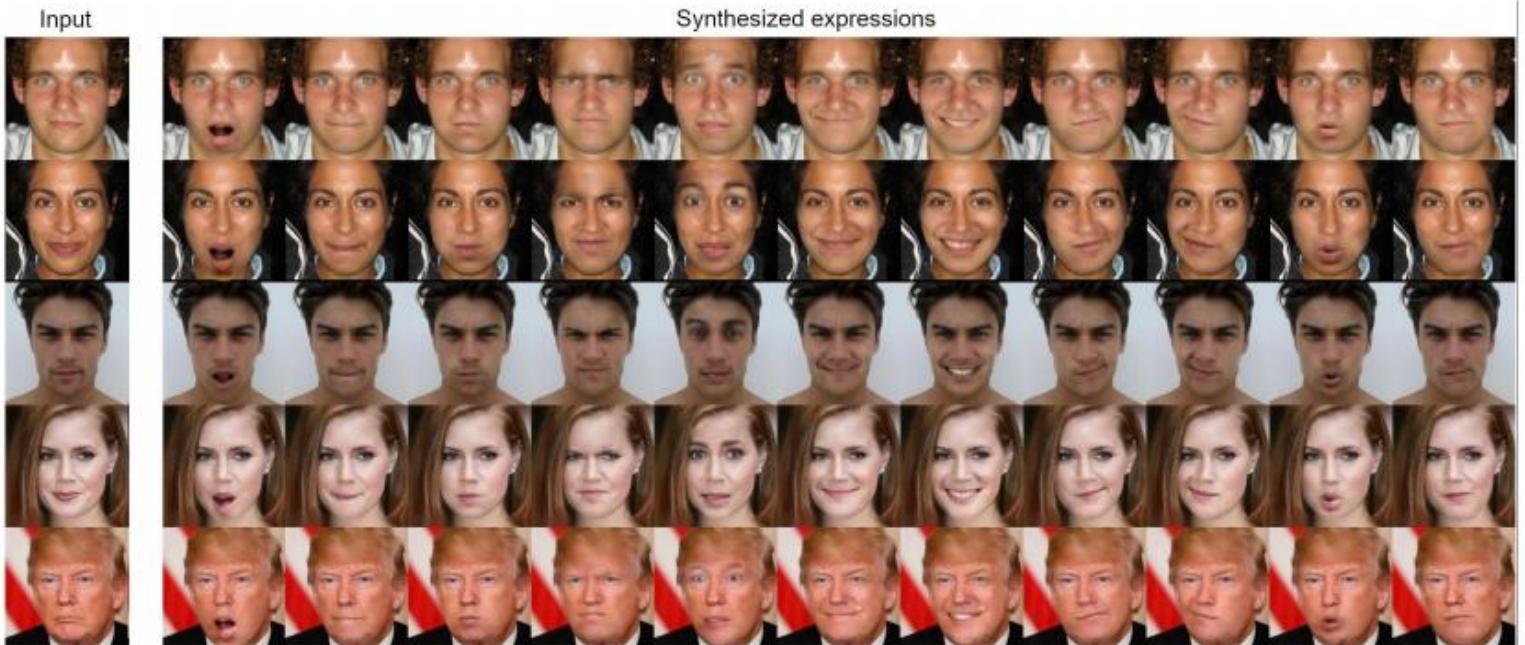
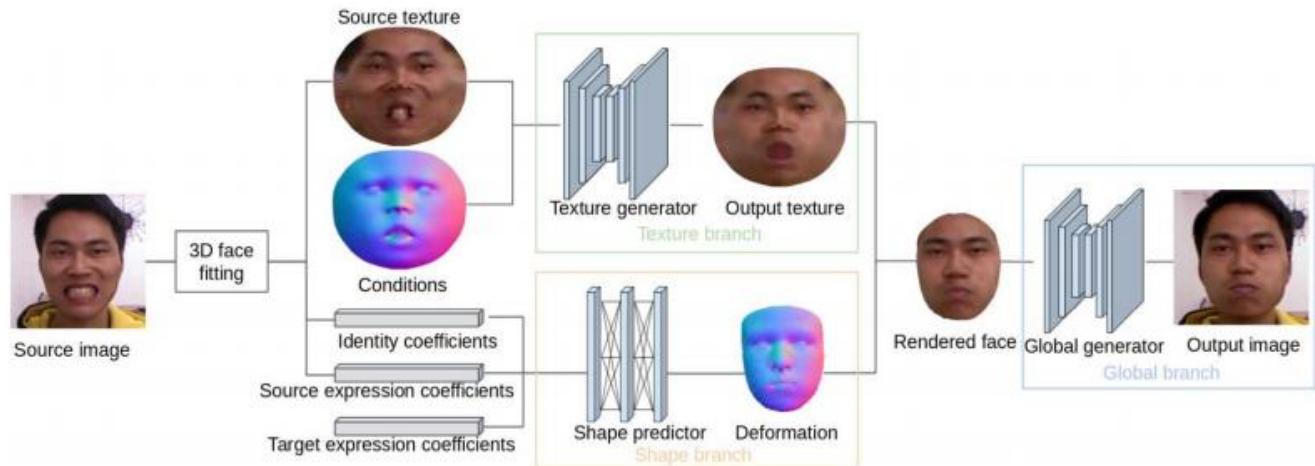
# 3D Guided Fine-Grained Face Manipulation

Zhenglin Geng<sup>1,2</sup>, Chen Cao<sup>2</sup>, and Sergey Tulyakov<sup>2</sup>

<sup>1</sup>Stanford University, <sup>2</sup>Snap Inc.

zhenglin@stanford.edu, {chen.cao, stulyakov}@snapchat.com

Unlimited amount of expressions



Methods	ACD	User Preference Ours / Others
Texture mapping [5]	0.6194	<b>69.8</b> / 30.2
StarGAN [6]	0.5981	<b>86.8</b> / 13.2
GANimation [21]	0.5595	<b>86.2</b> / 13.8
<b>Ours</b>	<b>0.5107</b>	N/A
Ground truth	0.4608	<b>53.4</b> / 46.6

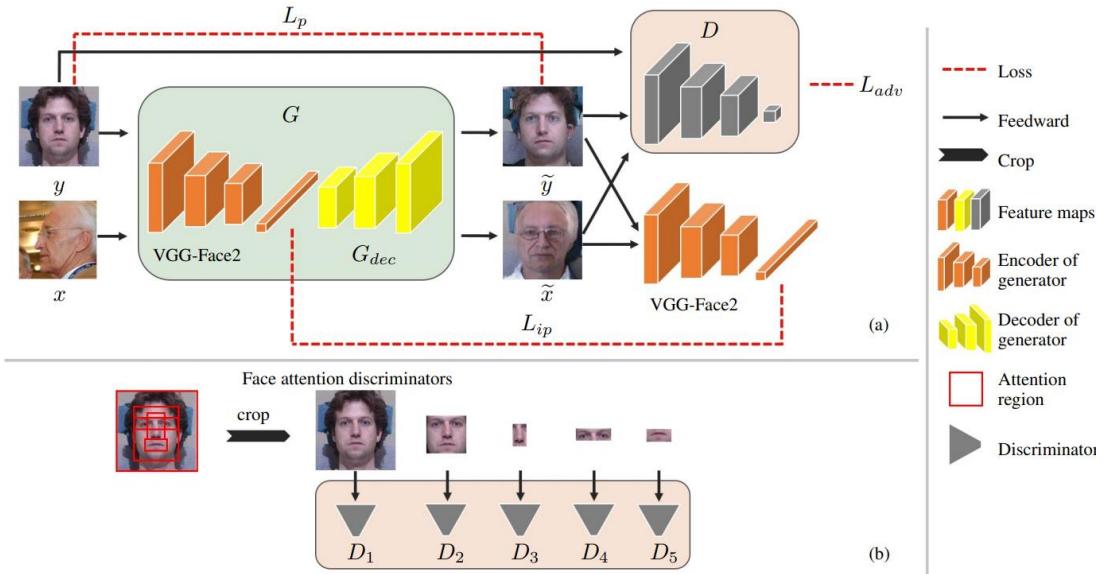
# Unsupervised Face Normalization with Extreme Pose and Expression in the Wild

Yichen Qian<sup>12</sup>, Weihong Deng<sup>1\*</sup>, Jiani Hu<sup>1</sup>

<sup>1</sup>Beijing University of Posts and Telecommunications

<sup>2</sup>AI Labs, Didi Chuxing, Beijing 100193, China

{mx54039q, whdeng, jnhu}@bupt.edu.cn



- Face normalization instead of frontalization



# Semantic Component Decomposition for Face Attribute Manipulation

Ying-Cong Chen<sup>1</sup> Xiaohui Shen<sup>4</sup> Zhe Lin<sup>3</sup> Xin Lu<sup>3</sup> I-Ming Pao<sup>3</sup> Jiaya Jia<sup>1,2</sup>

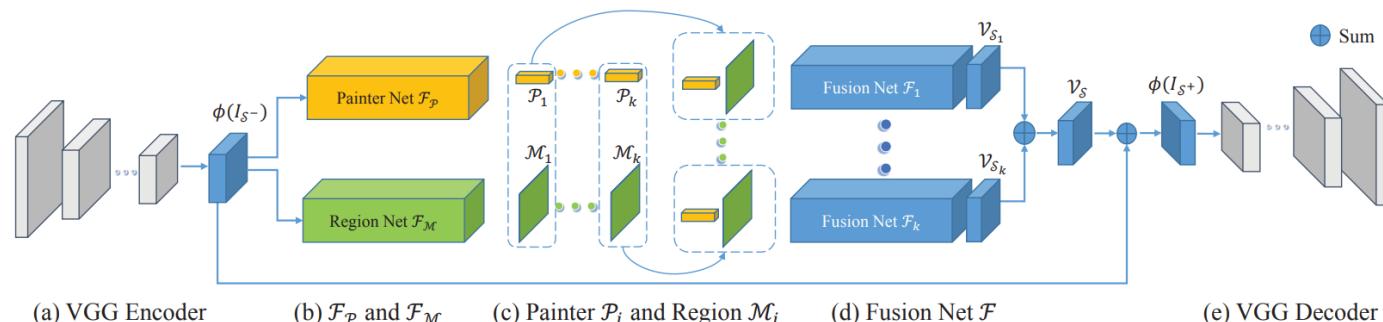
<sup>1</sup>The Chinese University of Hong Kong <sup>2</sup>Tencent YouTu Lab <sup>3</sup>Adobe Research <sup>4</sup>ByteDance AI Lab

{ycchen, leo{jia}@cse.cuhk.edu.hk

shenxiaohui@gmail.com

{zlin, xinl, pao}@adobe.com

Decompose facial attributes  
Edit strength and remove unwanted edit effect



(a) Input



(a) Input



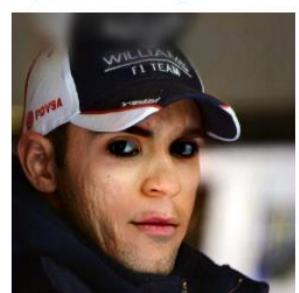
(b) All Components



(c) Unwanted Removed



(d) Input



(e) All Components



(f) Unwanted Removed



(a) Input

(b) Initial Result

(c) Edited Result

# Automatic Face Aging in Videos via Deep Reinforcement Learning

Chi Nhan Duong <sup>1</sup>, Khoa Luu <sup>2</sup>, Kha Gia Quach <sup>1</sup>, Nghia Nguyen <sup>2</sup>,  
Eric Patterson <sup>3</sup>, Tien D. Bui <sup>1</sup>, Ngan Le <sup>4</sup>

<sup>1</sup> Computer Science and Software Engineering, Concordia University, Canada

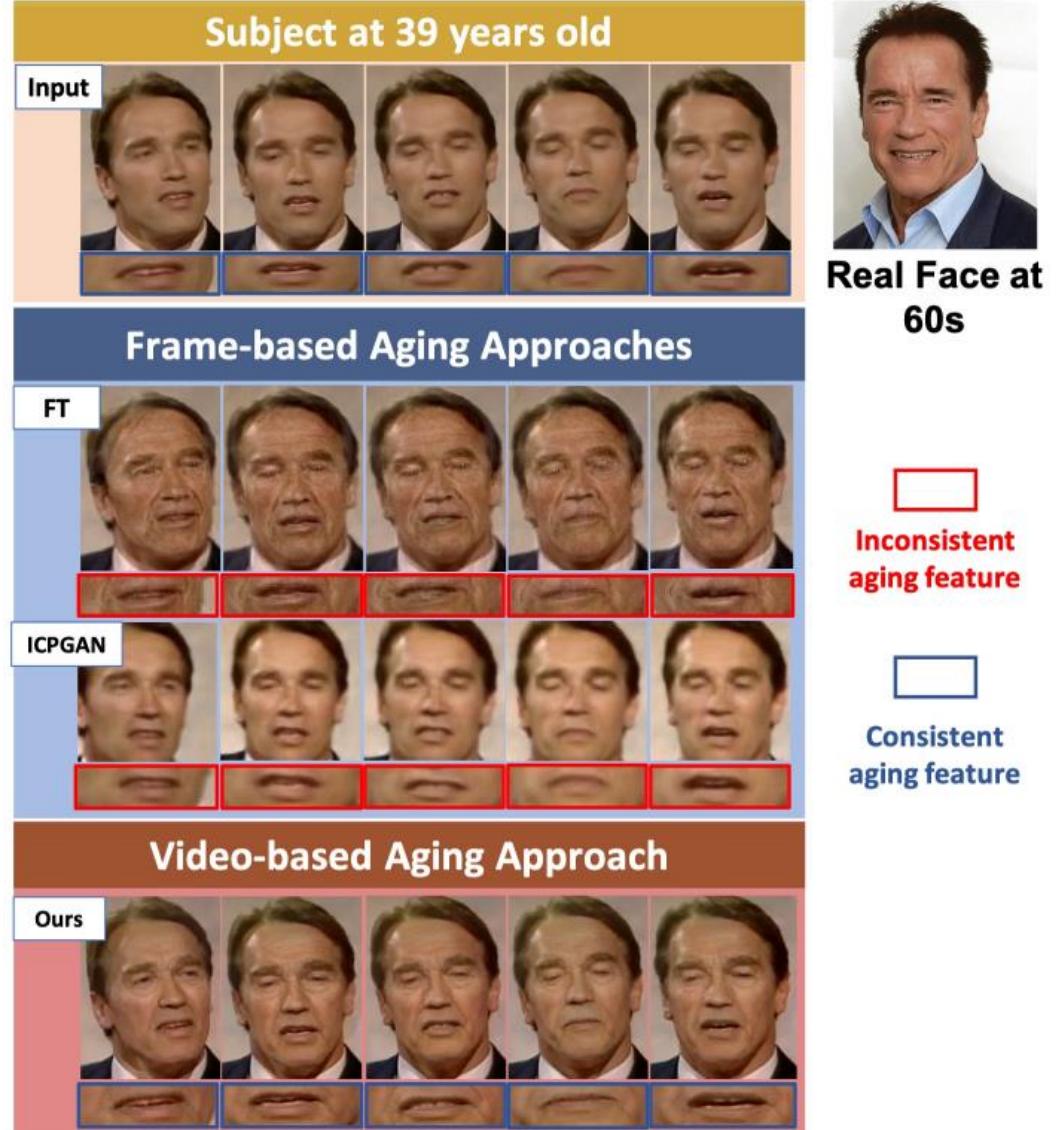
<sup>2</sup> Computer Science and Computer Engineering, University of Arkansas, USA

<sup>3</sup> School of Computing, Clemson University, USA

<sup>4</sup> Electrical and Computer Engineering, Carnegie Mellon University, USA

<sup>1</sup>{dcnhan, kquach}@ieee.org, bui@encs.concordia.ca, <sup>2</sup>{khoaluu, nhnguyen}@uark.edu,

<sup>3</sup>ekp@clemson.edu, <sup>4</sup>thihoanl@andrew.cmu.edu



# APDrawingGAN: Generating Artistic Portrait Drawings from Face Photos with Hierarchical GANs

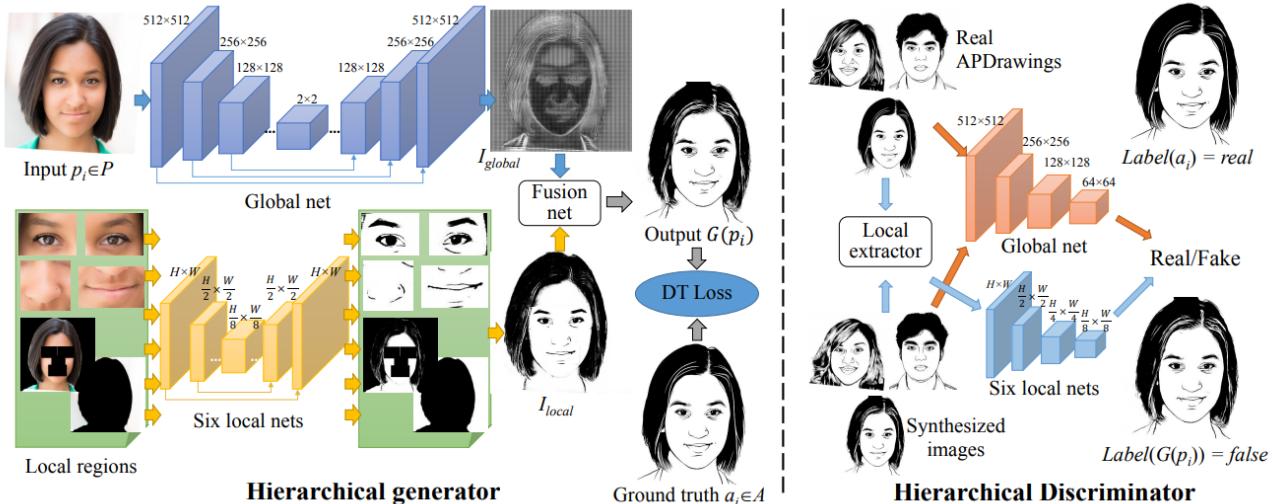
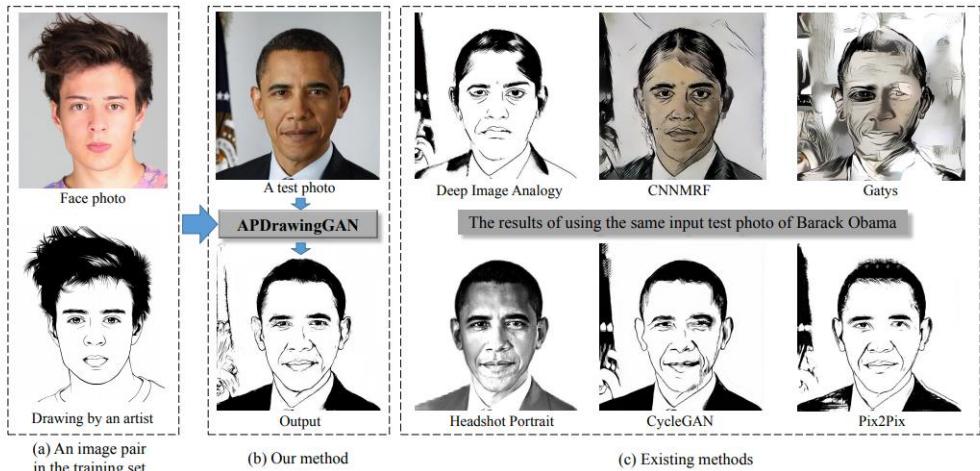
Ran Yi, Yong-Jin Liu\*

CS Dept, BNRist

Tsinghua University, China

{yr16, liuyongjin}@tsinghua.edu.cn

Yu-Kun Lai, Paul L. Rosin  
 School of Computer Science and Informatics  
 Cardiff University, UK  
 {LaiY4, RosinPL}@cardiff.ac.uk



# Attribute-aware Face Aging with Wavelet-based Generative Adversarial Networks

Yunfan Liu<sup>1\*</sup> Qi Li<sup>1,2,3\*</sup> Zhenan Sun<sup>1,2,4</sup>

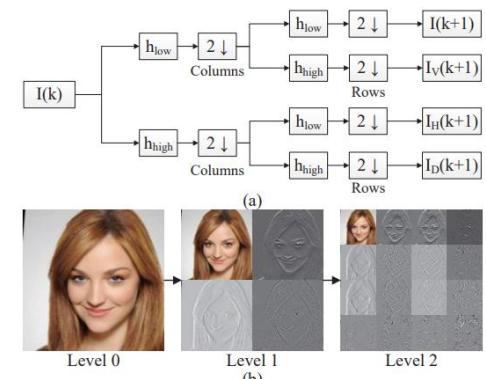
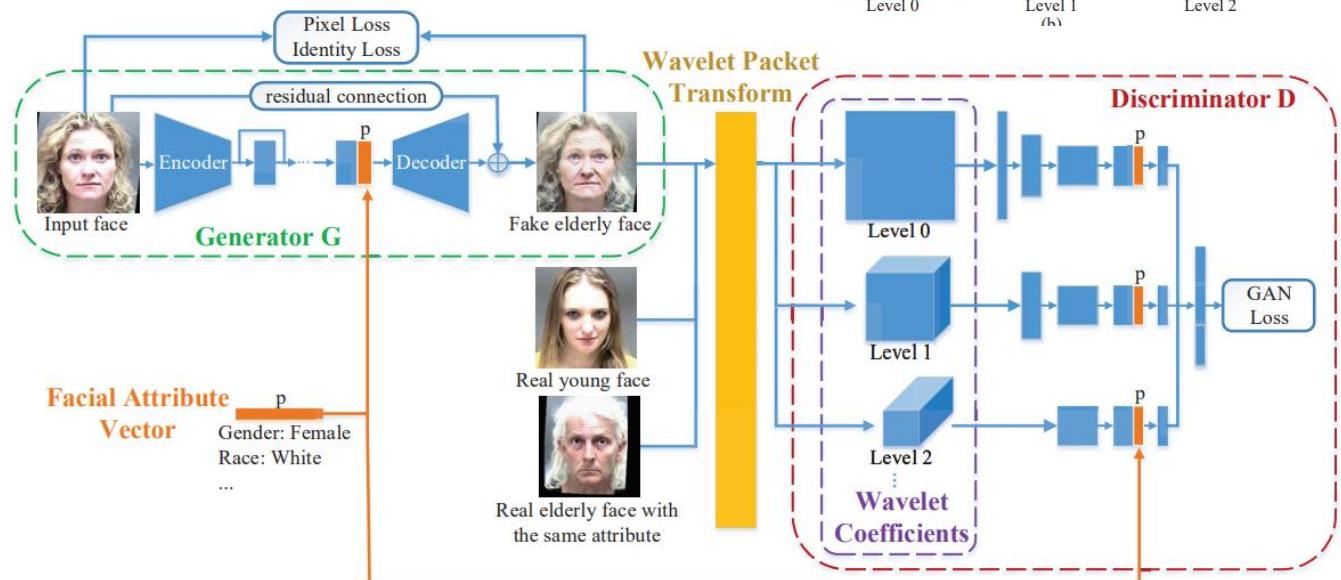
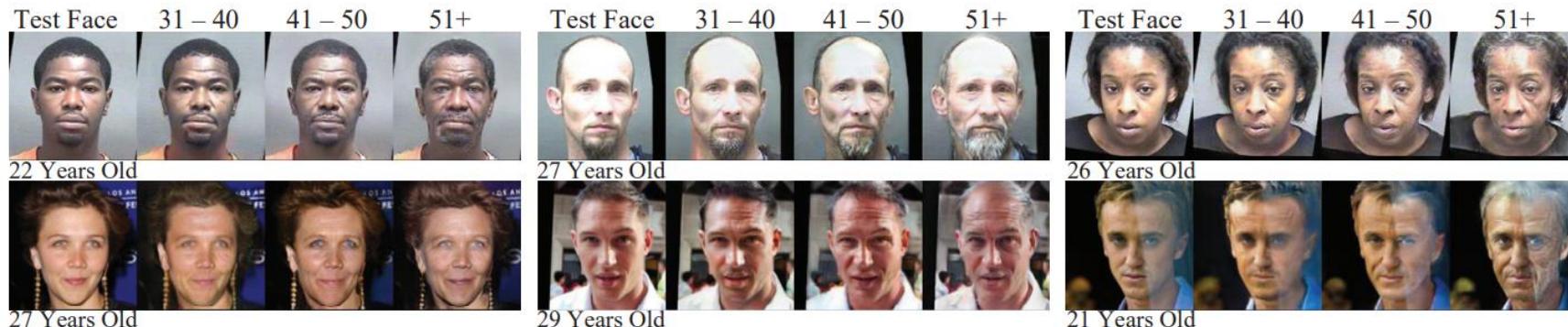
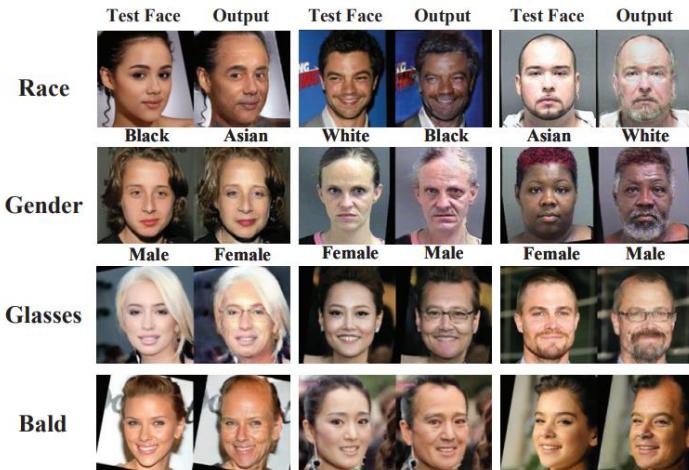
<sup>1</sup> Center for Research on Intelligent Perception and Computing, CASIA

<sup>2</sup> National Laboratory of Pattern Recognition, CASIA

<sup>3</sup> Artificial Intelligence Research, CAS, Jiaozhou, Qingdao, China

<sup>4</sup> Center for Excellence in Brain Science and Intelligence Technology, CAS

yunfan.liu@cripac.ia.ac.cn, {qli, znsun}@nlpr.ia.ac.cn



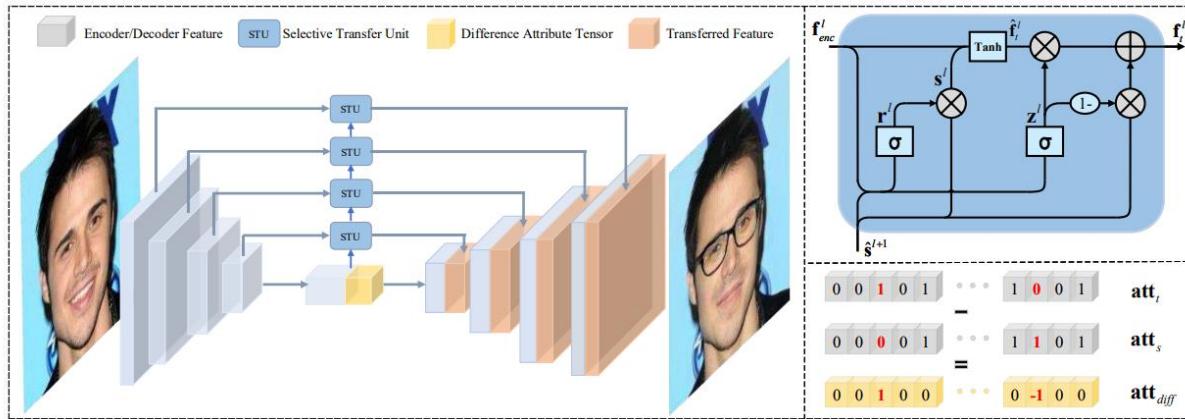
Level 0      Level 1 ( $I_h$ )      Level 2

# STGAN: A Unified Selective Transfer Network for Arbitrary Image Attribute Editing

Ming Liu<sup>1\*</sup>, Yukang Ding<sup>2</sup>, Min Xia<sup>1</sup>, Xiao Liu<sup>2</sup>, Errui Ding<sup>2</sup>, Wangmeng Zuo<sup>(✉)1,3</sup>, Shilei Wen<sup>2</sup>

<sup>1</sup>Harbin Institute of Technology, <sup>2</sup>Department of Computer Vision Technology (VIS), Baidu Inc., <sup>3</sup>Peng Cheng Laboratory, Shenzhen

csmliu@outlook.com, csmxia@gmail.com, wmu@hit.edu.cn,  
{dingyukang, liuxiao12, dingerrui, wenshilei}@baidu.com



- bottleneck layer: blurry and low quality editing result
- skip connections: weaken attribute manipulation ability
- selectively takes the difference between target and source attribute vectors as input



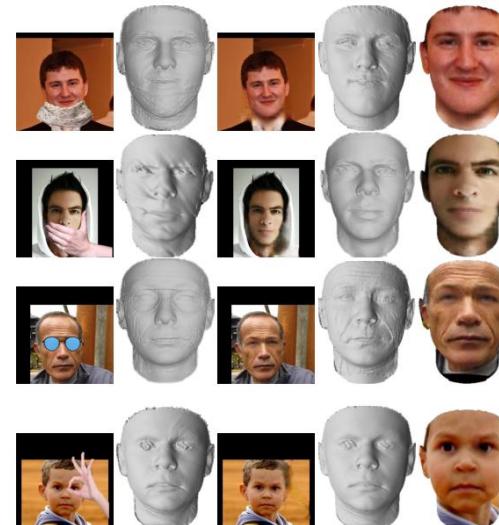
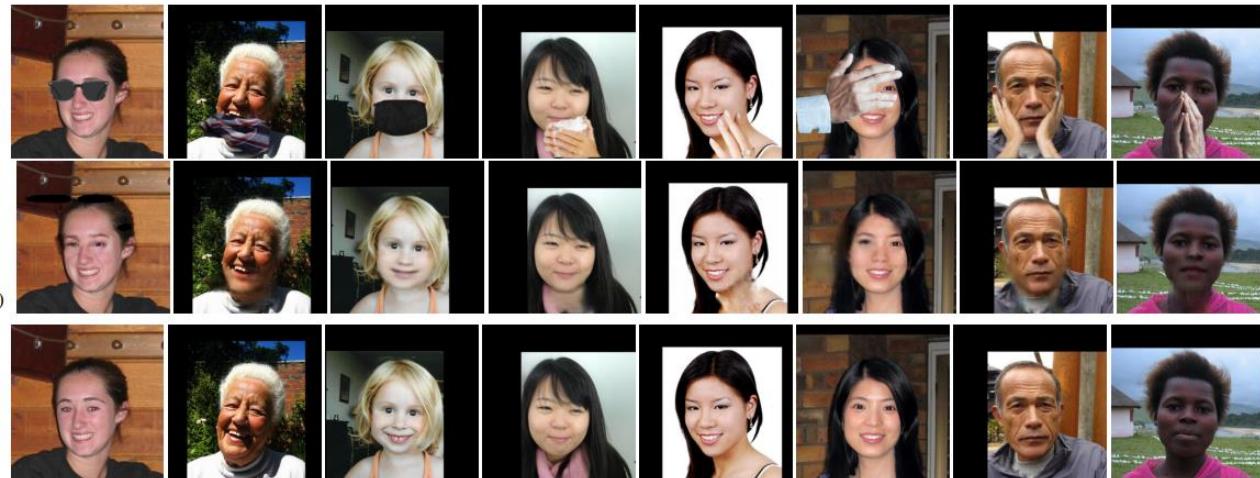
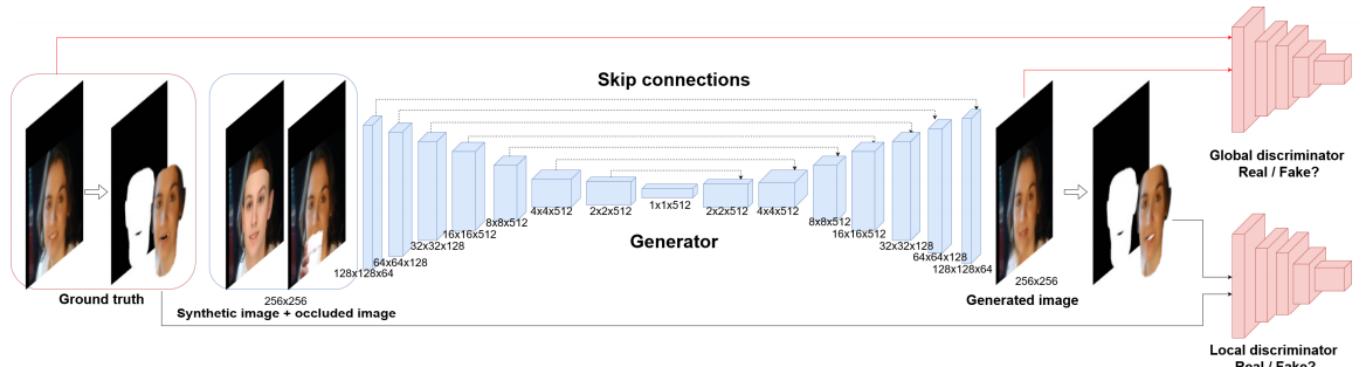
# Face De-occlusion using 3D Morphable Model and Generative Adversarial Network

3DMM + occluded image

Xiaowei Yuan and In Kyu Park

{xiaoweichn@qq.com pik@inha.ac.kr}

Dept. of Information and Communication Engineering, Inha University, Incheon 22212, Korea



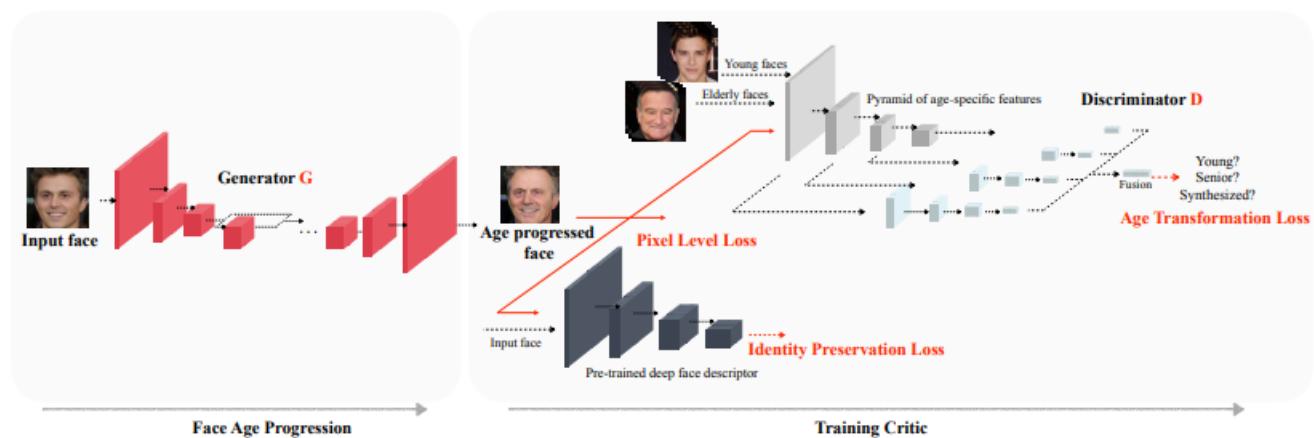
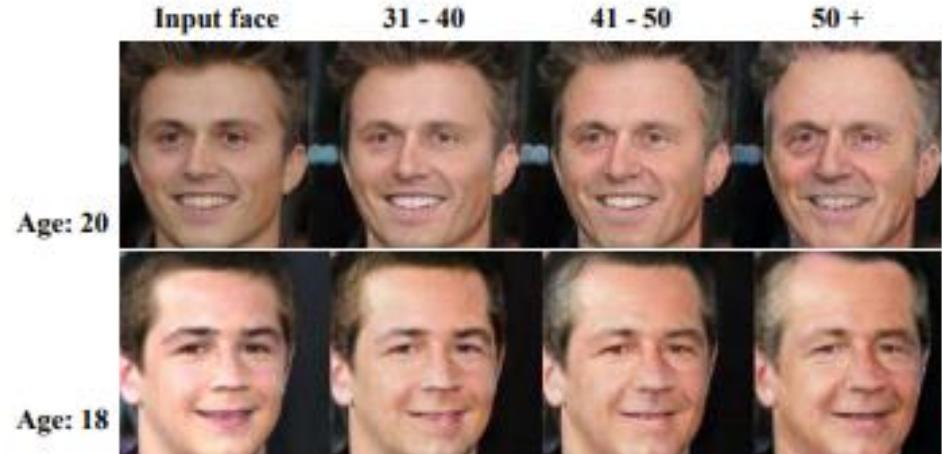
# Learning Face Age Progression: A Pyramid Architecture of GANs

Hongyu Yang<sup>1</sup> Di Huang<sup>1\*</sup> Yunhong Wang<sup>1</sup> Anil K. Jain<sup>2</sup>

<sup>1</sup>Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, China

<sup>2</sup>Michigan State University, USA

{honguyuyang, dhuang, yhwang}@buaa.edu.cn jain@cse.msu.edu



# FaceID-GAN: Learning a Symmetry Three-Player GAN for Identity-Preserving Face Synthesis

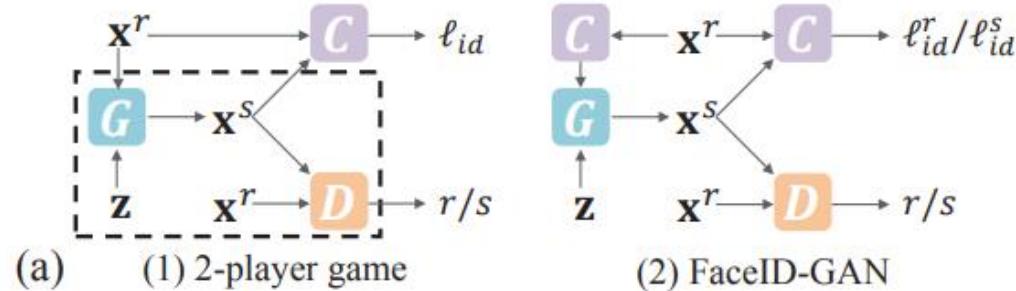
Yujun Shen<sup>1</sup>, Ping Luo<sup>1,3</sup>, Junjie Yan<sup>2</sup>, Xiaogang Wang<sup>1</sup>, Xiaoou Tang<sup>1</sup>

<sup>1</sup>CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong

<sup>2</sup>SenseTime Research

<sup>3</sup>Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

{sy116, pluo, xtang}@ie.cuhk.edu.hk, yanjunjie@sensetime.com, xgwang@ee.cuhk.edu.hk



# Exploring Disentangled Feature Representation Beyond Face Identification

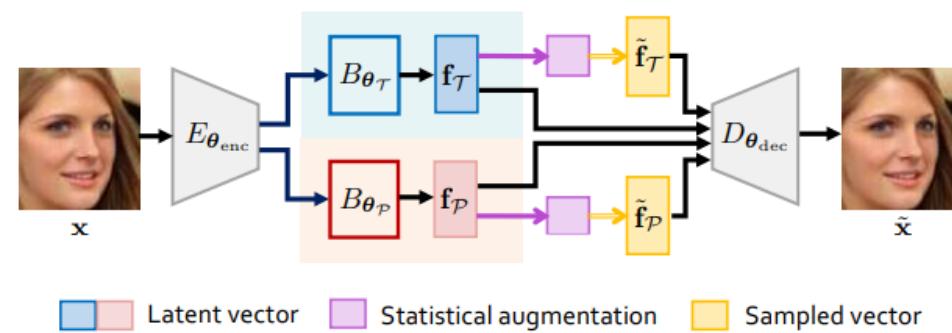
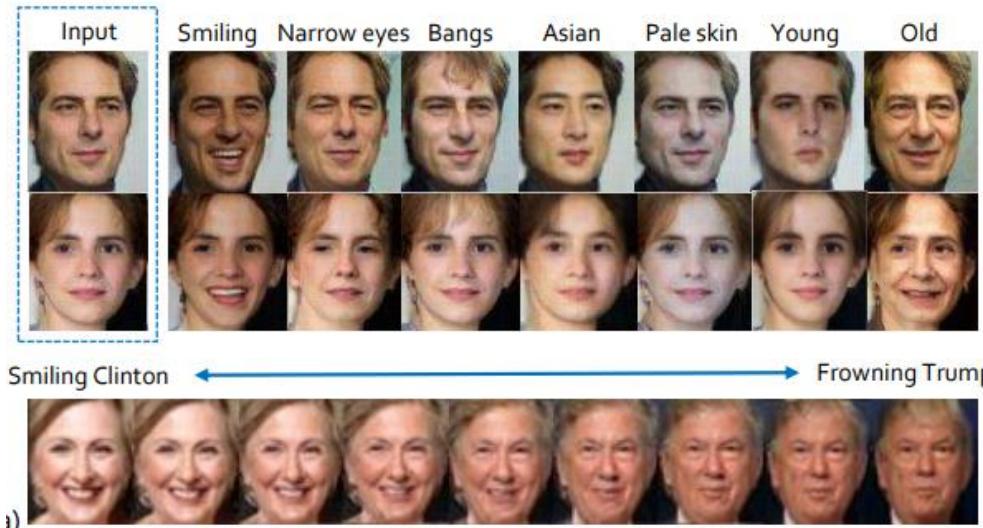
Yu Liu<sup>1\*</sup>, Fangyin Wei<sup>2\*</sup>, Jing Shao<sup>2\*</sup>, Lu Sheng<sup>1</sup>, Junjie Yan<sup>2</sup>, Xiaogang Wang<sup>1</sup>

<sup>1</sup>CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong

<sup>2</sup>SenseTime Group Limited

{yuliu, lsheng, xgwang}@ee.cuhk.edu.hk, weifangyin@pku.edu.cn,  
{shaojing, yanjunjie}@sensetime.com

## Identity distilling and dispelling

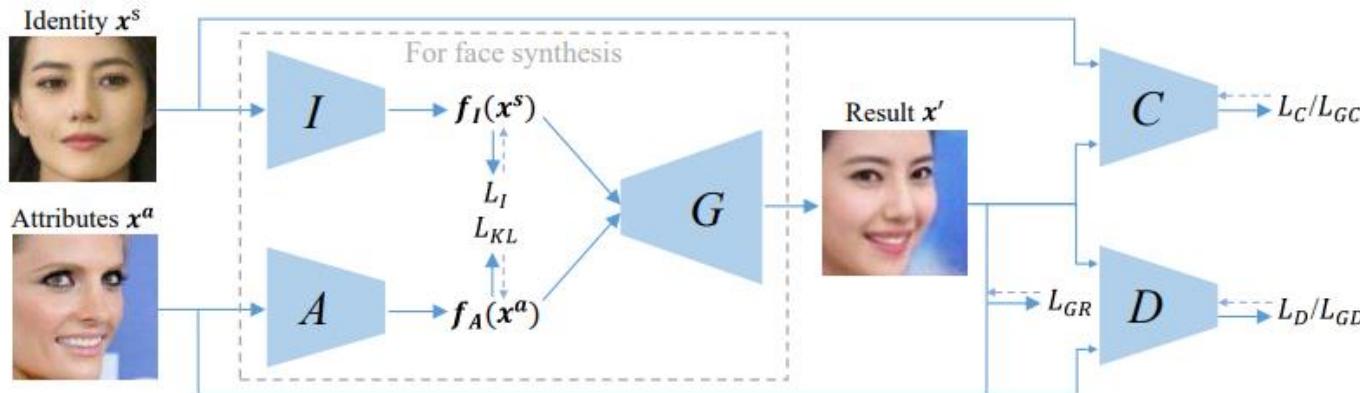


# Towards Open-Set Identity Preserving Face Synthesis

Jianmin Bao<sup>1</sup>, Dong Chen<sup>2</sup>, Fang Wen<sup>2</sup>, Houqiang Li<sup>1</sup>, Gang Hua<sup>2</sup>

<sup>1</sup>University of Science and Technology of China      <sup>2</sup>Microsoft Research

jmbao@mail.ustc.edu.cn {doch, fangwen, ganghua}@microsoft.com lihq@ustc.edu.cn



# Label Denoising Adversarial Network (LDAN) for Inverse Lighting of Faces

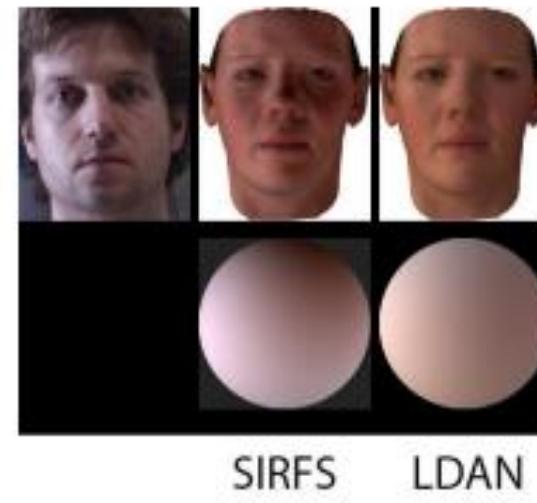
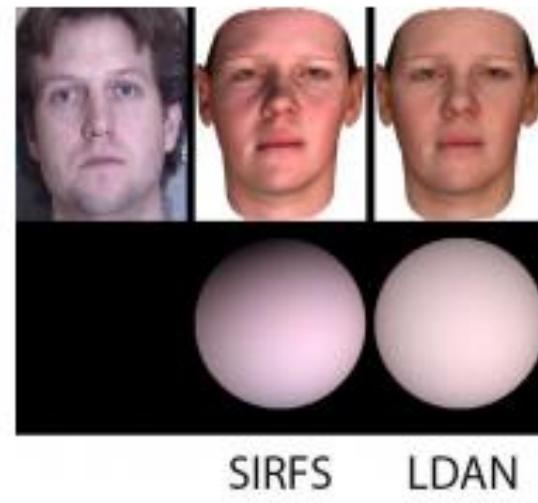
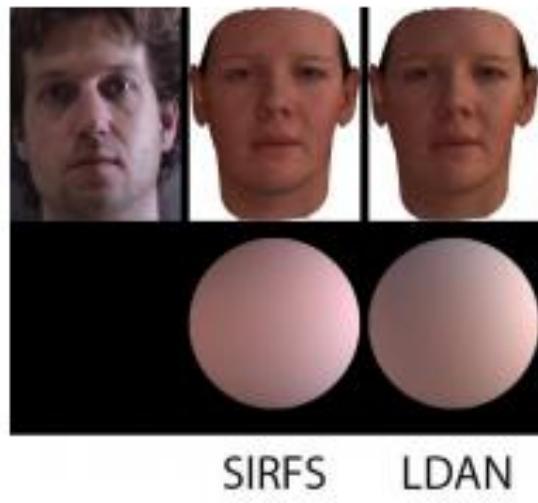
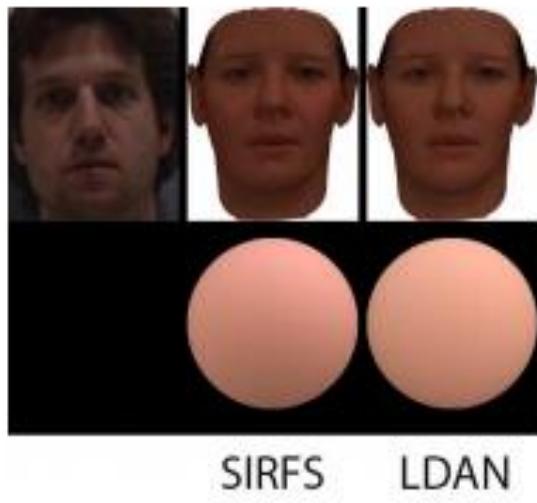
Hao Zhou \* Jin Sun\* Yaser Yacoob David W. Jacobs

University of Maryland, College Park, MD, USA

{hzhou, jinsun, yaser, djacobs}@cs.umd.edu

Lighting estimation

One attribute distilling and dispelling



# A Face-to-Face Neural Conversation Model

Hang Chu<sup>1,2</sup> Daiqing Li<sup>1</sup> Sanja Fidler<sup>1,2</sup>

<sup>1</sup>University of Toronto <sup>2</sup>Vector Institute

{chuhang1122, daiqing, fidler}@cs.toronto.edu



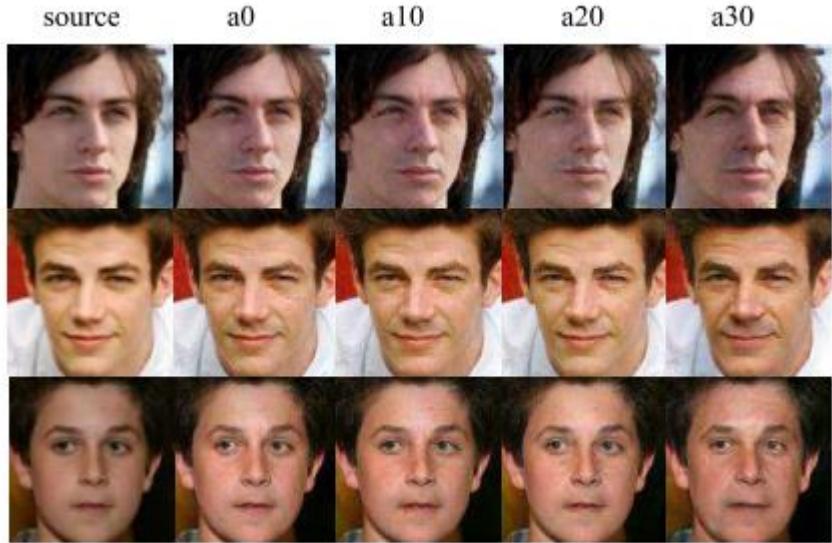
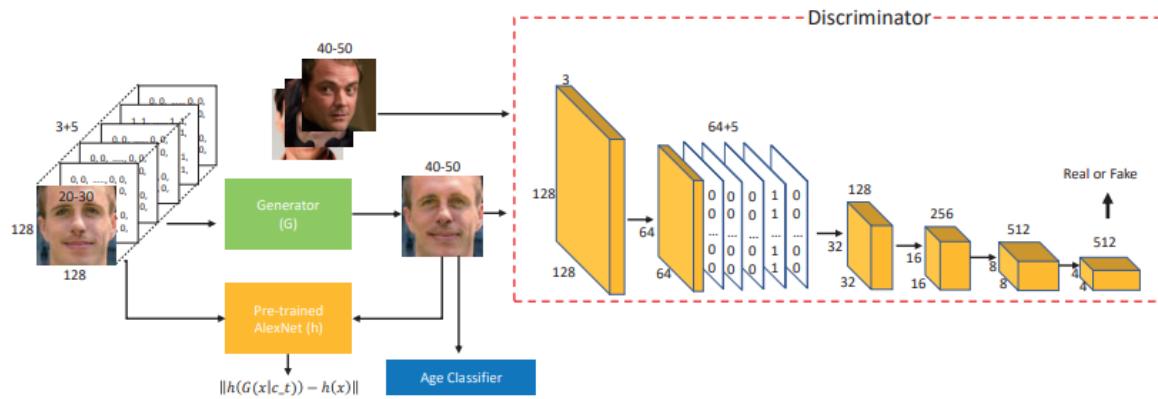
source text	source face sequence	true target text	text only [12, 27]	text+face
we went to the hickory stick, we had a drink, two drinks. she doesn't know where he is.		and then? and then i went home alone. i don't know where he is.	we drank a bottle of champagne. i'm sorry.	and then i went to bed. i don't know where she is.
and he sleeps only one hour a night. a night that marked the opening of a new chapter. i hope you're not a hothead like sonny.		he's a great man. in world history. he's a good kid.	he sleeps in the same bed. for the future.	he's a good man. in the history of the world. he's a good kid.
i guess they was worried they wouldn't find a vein in my arm. oh, he's so cute.		what's that number? oh, my god.	what's that?	i don't think so. oh, my god he's so cute.
can you hear me? i'm still here. scott. stop. so i don't really remember, yeah. i can't feel my legs.		i'm here. scott. stop. yeah, right. stupid. i can't feel my legs.	i'm sorry. yeah, yeah, yeah. and i can't	what the f*** are you doing here? well, you know what? i'm sorry. it's too much.

# Face Aging with Identity-Preserved Conditional Generative Adversarial Networks

Zongwei Wang  
Shanghaitech University  
[wangzw@shanghaitech.edu.cn](mailto:wangzw@shanghaitech.edu.cn)

Xu Tang  
Baidu  
[tangxu02@baidu.com](mailto:tangxu02@baidu.com)

Weixin Luo, Shenghua Gao\*  
Shanghaitech University  
[{luowx, gaoshh}@shanghaitech.edu.cn](mailto:{luowx, gaoshh}@shanghaitech.edu.cn)



# Deep Semantic Face Deblurring

Ziyi Shen<sup>1</sup>

Wei-Sheng Lai<sup>2</sup>

Tingfa Xu<sup>1\*</sup>

Jan Kautz<sup>3</sup>

Ming-Hsuan Yang<sup>2,4</sup>

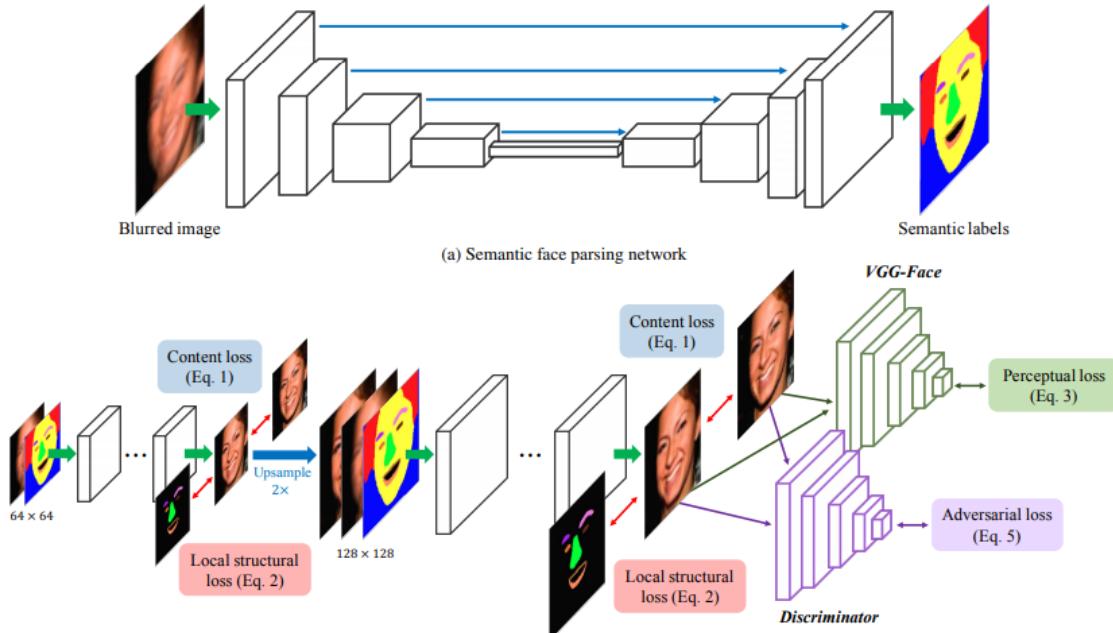
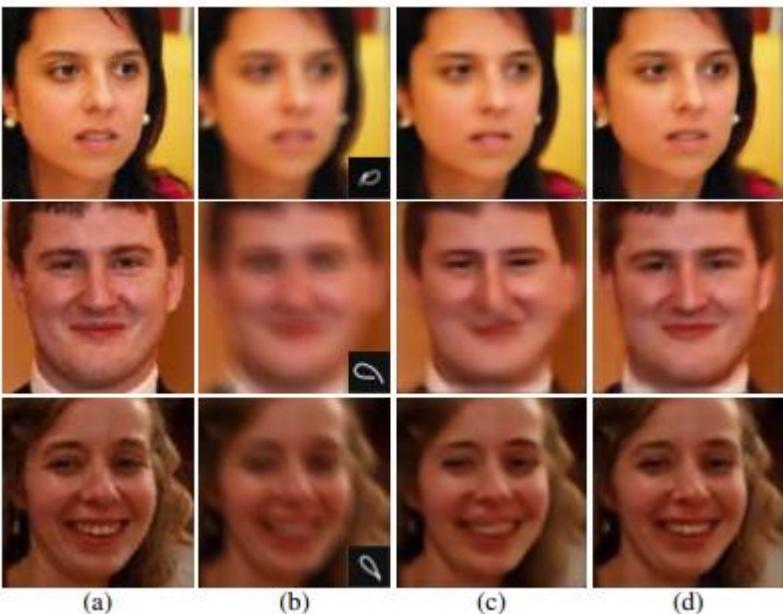
<sup>1</sup>Beijing Institute of Technology

<sup>2</sup>University of California, Merced

<sup>3</sup>Nvidia

<sup>4</sup>Google Cloud

[https://sites.google.com/site/ziyishenmi/cvpr18\\_face\\_deblur](https://sites.google.com/site/ziyishenmi/cvpr18_face_deblur)



# Pose-Guided Photorealistic Face Rotation

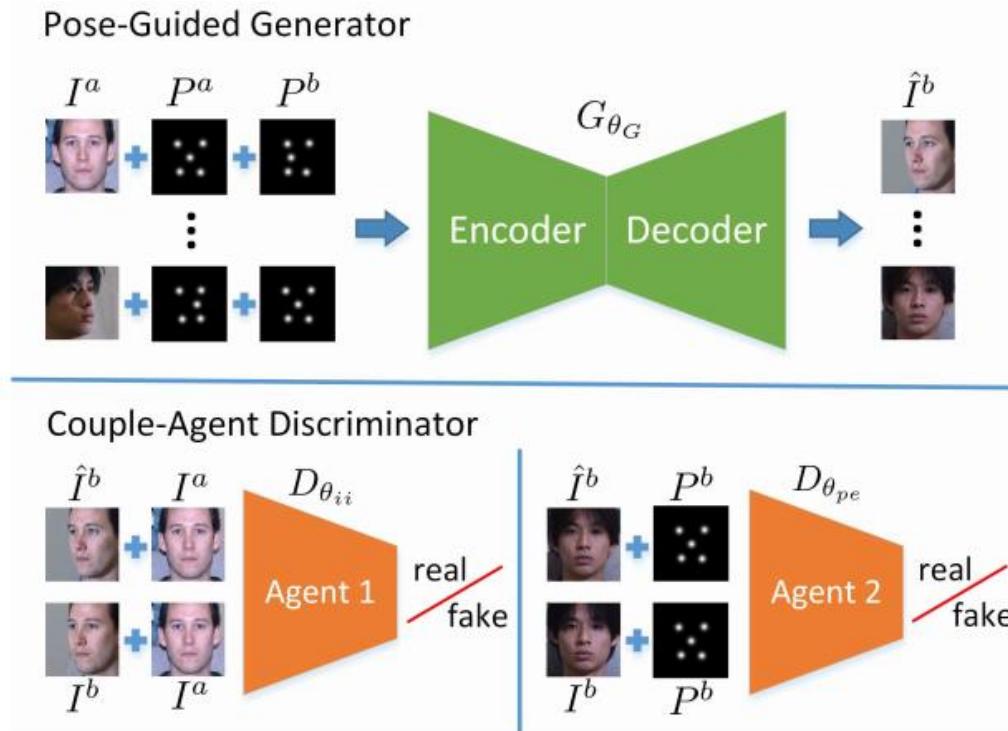
Yibo Hu<sup>1,2</sup>, Xiang Wu<sup>1</sup>, Bing Yu<sup>3</sup>, Ran He<sup>1,2\*</sup>, Zhenan Sun<sup>1,2</sup>

<sup>1</sup>CRIPAC & NLPR & CEBSIT, CASIA

<sup>2</sup>University of Chinese Academy of Sciences

<sup>3</sup>Noah's Ark Laboratory, Huawei Technologies Co., Ltd.

{yibo.hu, xiang.wu}@cripac.ia.ac.cn, yubing5@huawei.com, {rhe, znsun}@nlpr.ia.ac.cn



## Autoencoding beyond pixels using a learned similarity metric

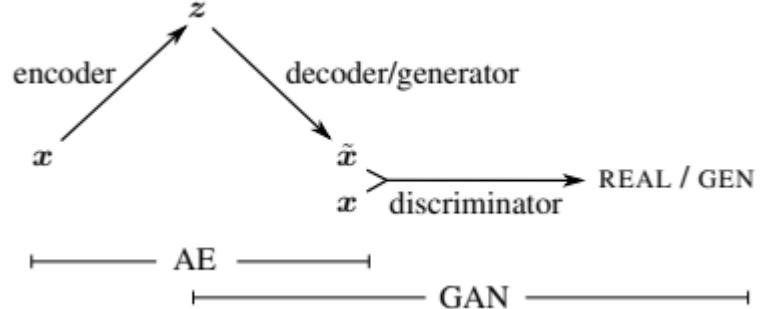
Anders Boesen Lindbo Larsen<sup>1</sup>  
 Søren Kaae Sønderby<sup>2</sup>  
 Hugo Larochelle<sup>3</sup>  
 Ole Winther<sup>1,2</sup>

ABLL@DTU.DK  
 SKAAE SONDERBY@GMAIL.COM  
 HLAROCHELLE@TWITTER.COM  
 OLWI@DTU.DK

<sup>1</sup> Department for Applied Mathematics and Computer Science, Technical University of Denmark

<sup>2</sup> Bioinformatics Centre, Department of Biology, University of Copenhagen, Denmark

<sup>3</sup> Twitter, Cambridge, MA, USA



- Combination of VAE and GAN
- Learned similarity metric: switch (3) into (7)

$$\mathcal{L}_{\text{VAE}} = -\mathbb{E}_{q(z|x)} \left[ \log \frac{p(x|z)p(z)}{q(z|x)} \right] = \mathcal{L}_{\text{llike}}^{\text{pixel}} + \mathcal{L}_{\text{prior}}$$
(2)

with

$$\mathcal{L}_{\text{llike}}^{\text{pixel}} = -\mathbb{E}_{q(z|x)} [\log p(x|z)]$$
(3)

$$\mathcal{L}_{\text{prior}} = D_{\text{KL}}(q(z|x)\|p(z)) ,$$
(4)

$$\mathcal{L}_{\text{llike}}^{\text{Dis}_l} = -\mathbb{E}_{q(z|x)} [\log p(\text{Dis}_l(x)|z)]$$
(7)



Figure 3. Samples from different generative models.



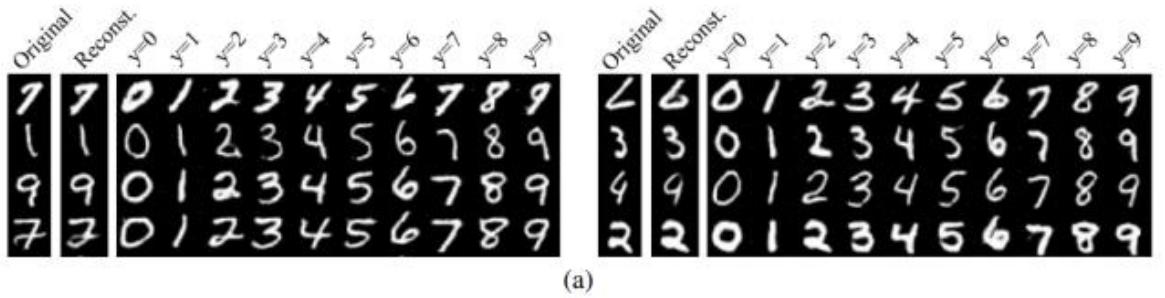
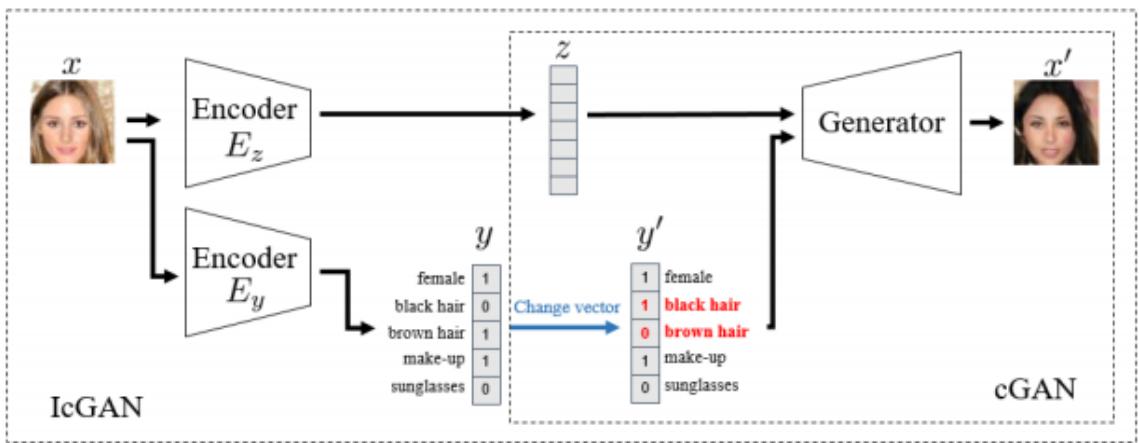
Figure 4. Reconstructions from different autoencoders.

# Invertible Conditional GANs for image editing

Guim Perarnau, Joost van de Weijer, Bogdan Raducanu  
 Computer Vision Center  
 Barcelona, Spain  
 guimperarnau@gmail.com  
 {joost,bogdan}@cvc.uab.es

Jose M. Álvarez  
 Data61 @ CSIRO  
 Canberra, Australia  
 jose.alvarez@nicta.com.au

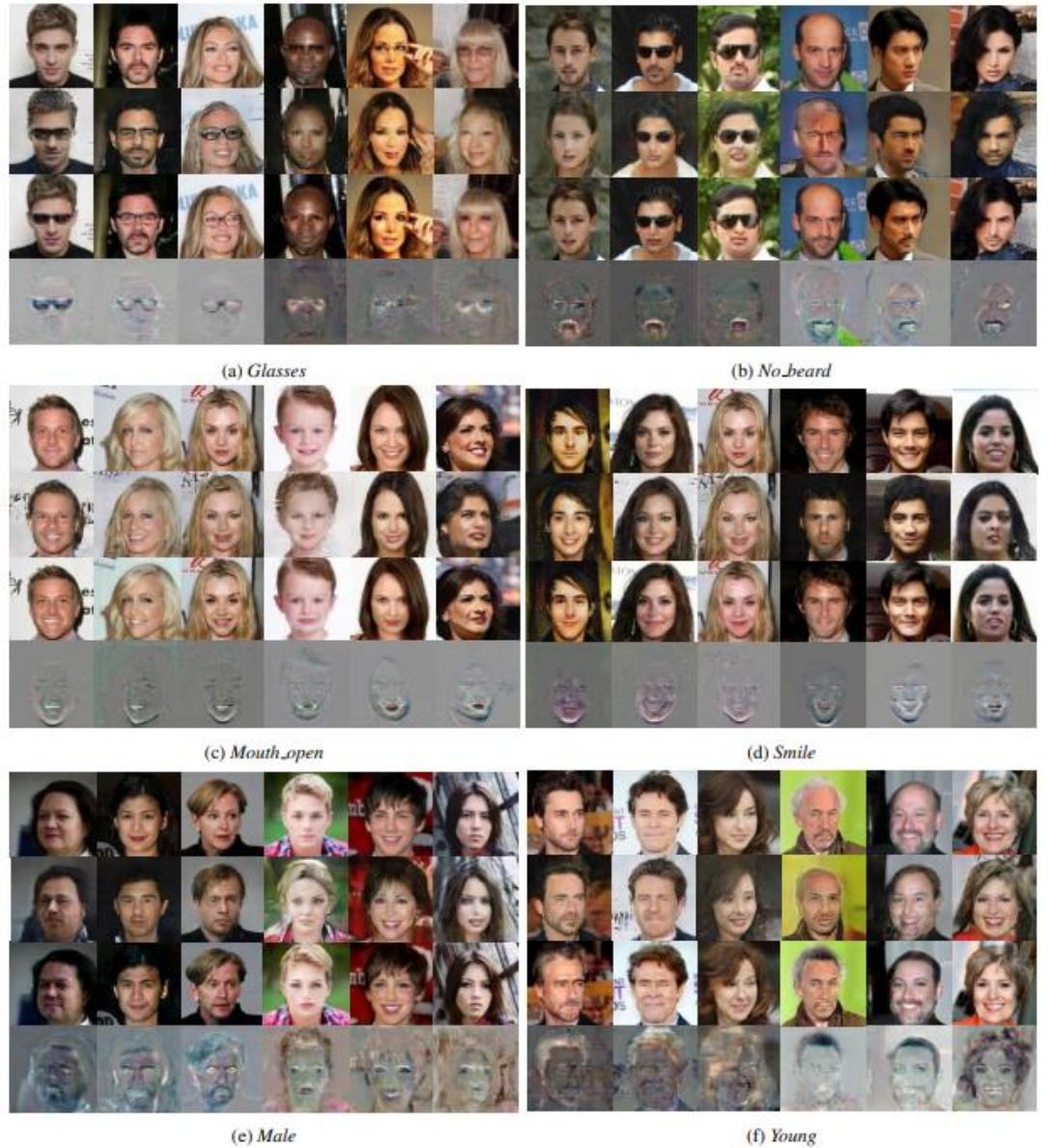
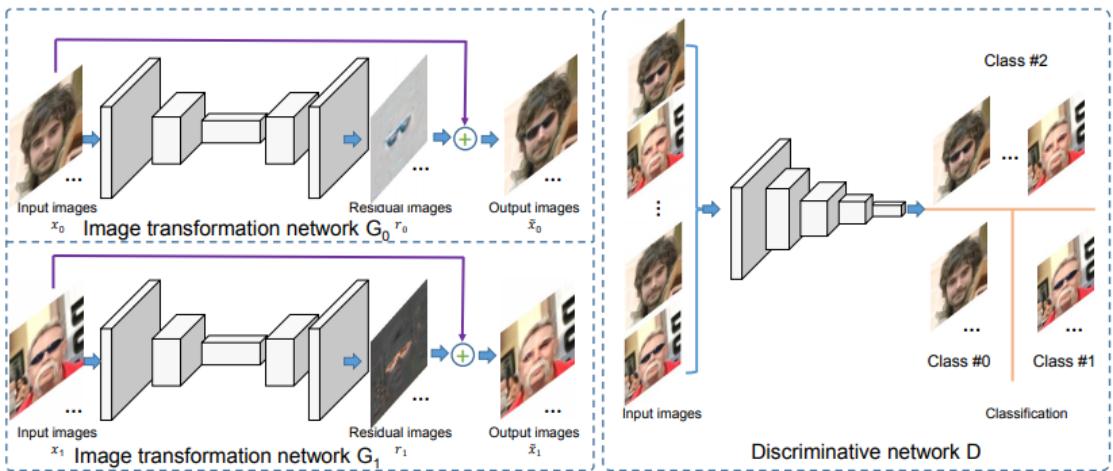
- Inference: image to z
- Encoder: image to z



# Learning Residual Images for Face Attribute Manipulation

Wei Shen Ruijie Liu  
Fujitsu Research & Development Center, Beijing, China.  
[{shenwei, rjliu}@cn.fujitsu.com](mailto:{shenwei, rjliu}@cn.fujitsu.com)

- Learn **residual** and then sum up
- Discriminator: real, G1, G2
- One attribute to another attribute
- **Dual** training, but not cycle



# Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis

Rui Huang<sup>1,2\*†</sup> Shu Zhang<sup>1,2,3\*</sup> Tianyu Li<sup>1,2</sup> Ran He<sup>1,2,3</sup>

<sup>1</sup>National Laboratory of Pattern Recognition, CASIA

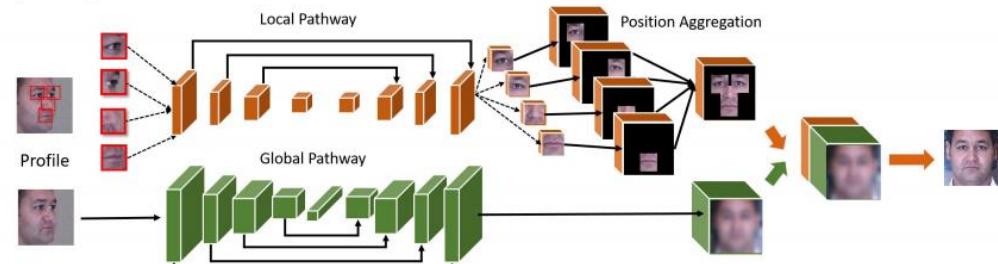
<sup>2</sup>Center for Research on Intelligent Perception and Computing, CASIA

<sup>3</sup>University of Chinese Academy of Sciences, Beijing, China

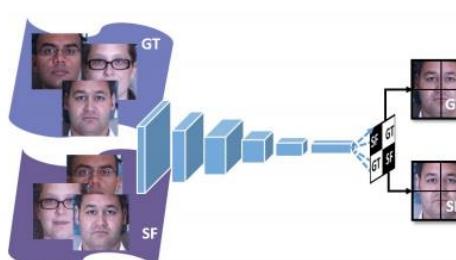
huangrui@cmu.edu, tianyu.lizard@gmail.com, {shu.zhang, rhe}@nlpr.ia.ac.cn

Facial parts  
Identity

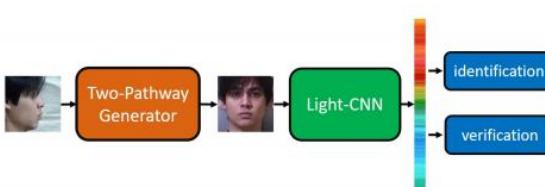
**Two-pathway Generator Network**



**Discriminator Network**



**Recognition via Generation**



(a) Profile (b) Ours (c) [33] (d) [41] (e) [8] (f) [44] (g) [12] (h) Frontal



# StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation

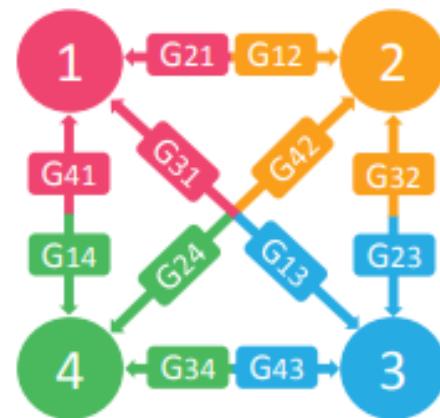
Yunjey Choi<sup>1,2</sup> Minje Choi<sup>1,2</sup> Munyoung Kim<sup>2,3</sup> Jung-Woo Ha<sup>2</sup> Sunghun Kim<sup>2,4</sup> Jaegul Choo<sup>1,2</sup>

<sup>1</sup> Korea University <sup>2</sup> Clova AI Research, NAVER Corp.

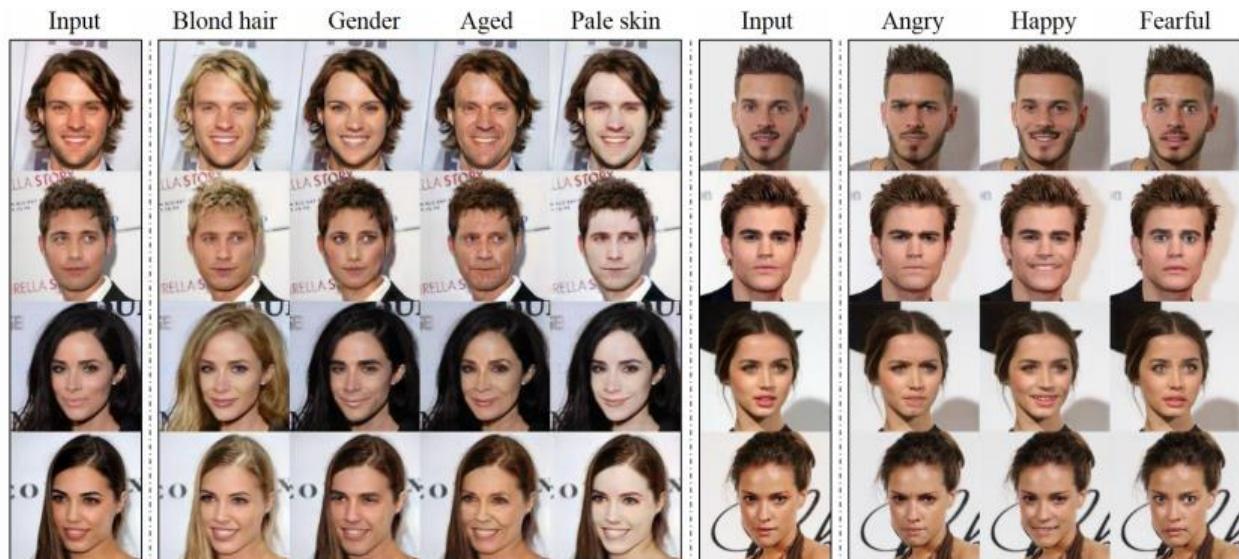
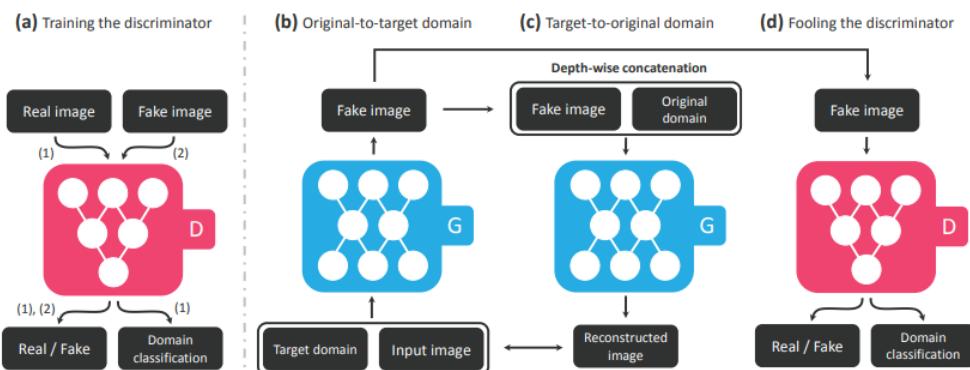
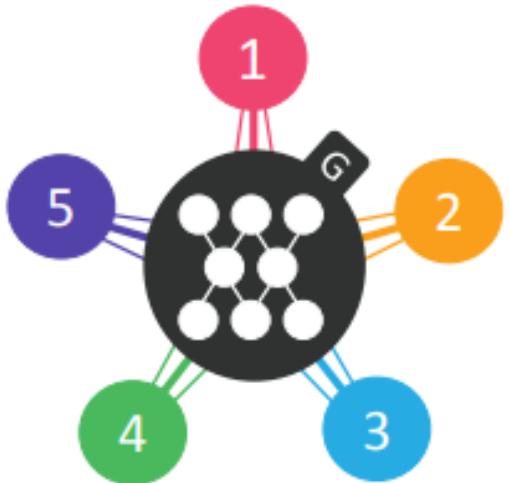
<sup>3</sup> The College of New Jersey <sup>4</sup> Hong Kong University of Science & Technology

One model, multiple domain  
Attribute code

(a) Cross-domain models



(b) StarGAN

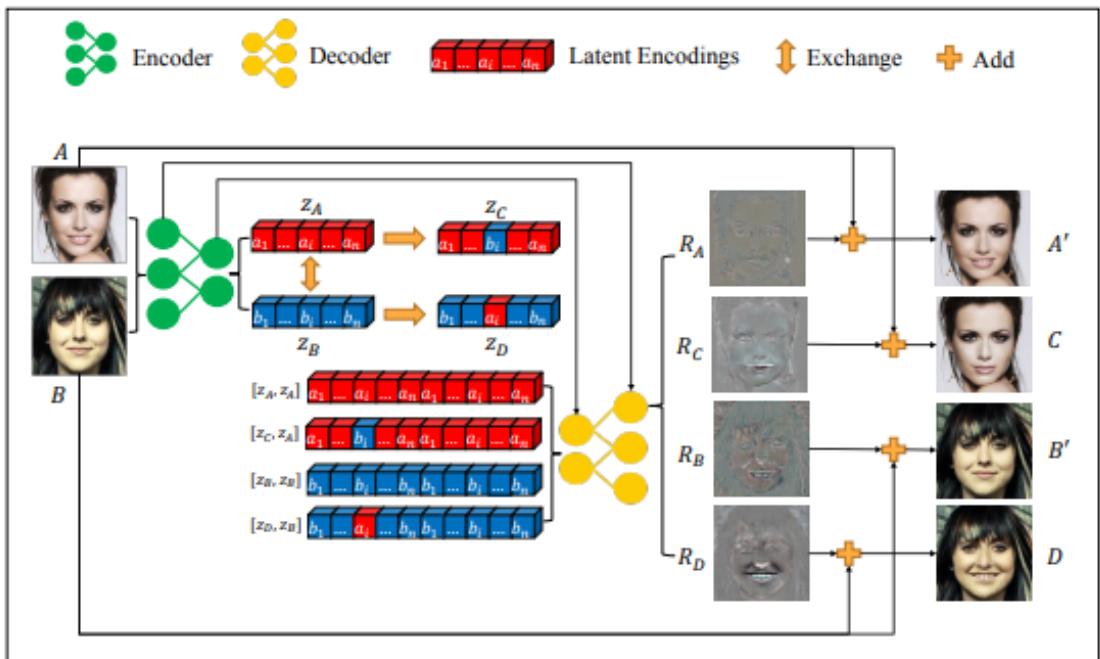


# ELEGANT: Exchanging Latent Encodings with GAN for Transferring Multiple Face Attributes

Taihong Xiao, Jiapeng Hong, and Jinwen Ma

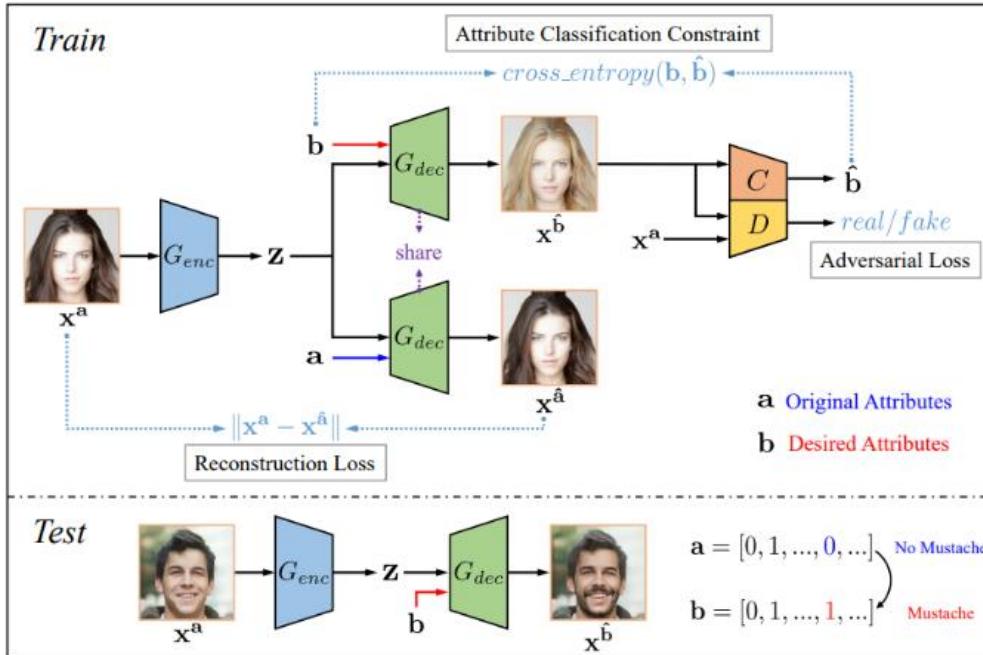
Department of Information Science, School of Mathematical Sciences  
and LMAM, Peking University, Beijing, 100871, China  
[{xiaotaihong, jphong}@pku.edu.cn](mailto:{xiaotaihong, jphong}@pku.edu.cn), [jwma@math.pku.edu.cn](mailto:jwma@math.pku.edu.cn)

- By exemplars
- Multiple attributes
- High quality



# AttGAN: Facial Attribute Editing by Only Changing What You Want

Zhenliang He, Wangmeng Zuo, *Senior Member, IEEE*, Meina Kan, *Member, IEEE*,  
Shiguang Shan, *Senior Member, IEEE*, and Xilin Chen, *Fellow, IEEE*



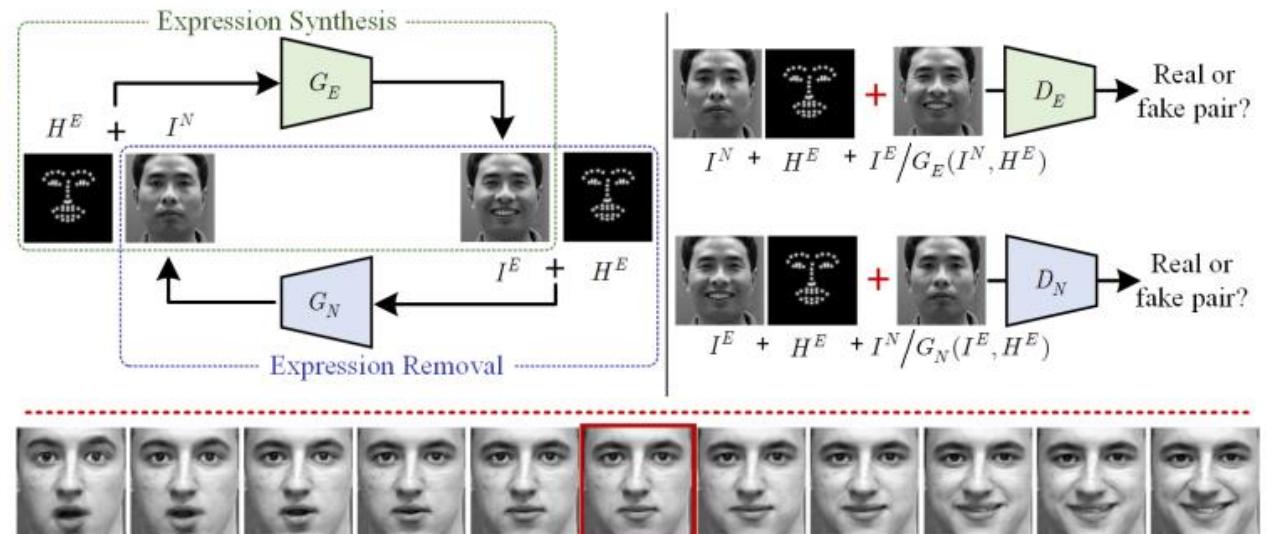
## Geometry Guided Adversarial Facial Expression Synthesis

Lingxiao Song<sup>1,2</sup> Zhihe Lu<sup>1,3</sup> Ran He<sup>1,2,3</sup> Zhenan Sun<sup>1,2</sup> Tieniu Tan<sup>1,2,3</sup>

<sup>1</sup>National Laboratory of Pattern Recognition, CASIA

<sup>2</sup>Center for Research on Intelligent Perception and Computing, CASIA

<sup>3</sup>Center for Excellence in Brain Science and Intelligence Technology, CAS



# GANimation: Anatomically-aware Facial Animation from a Single Image

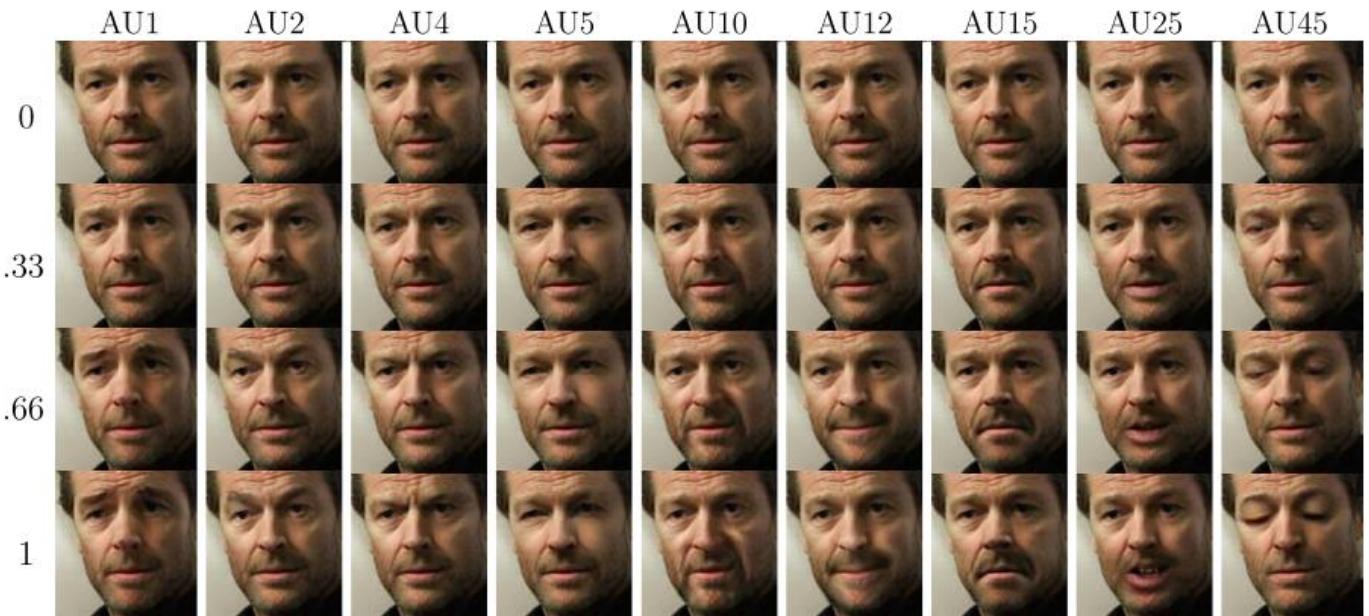
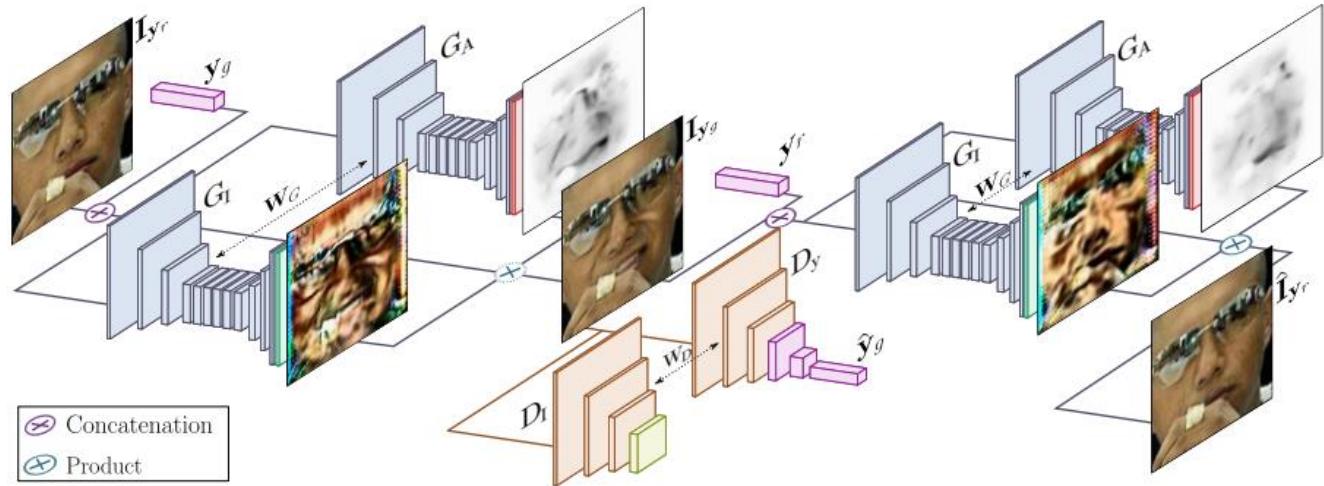
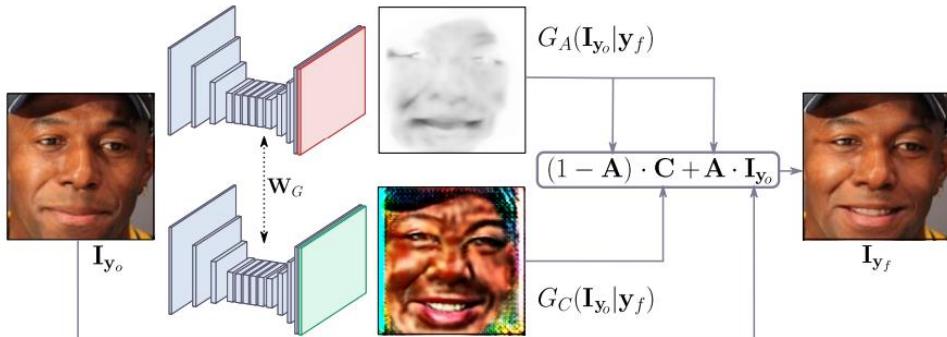
Albert Pumarola<sup>1</sup>, Antonio Agudo<sup>1</sup>, Aleix M. Martinez<sup>2</sup>,  
Alberto Sanfeliu<sup>1</sup>, Francesc Moreno-Noguer<sup>1</sup>

<sup>1</sup>Institut de Robòtica i Informàtica Industrial, CSIC-UPC, 08028, Barcelona, Spain

<sup>2</sup>The Ohio State University, Columbus, OH 43210, USA

- Action Units (Aus)
- Continues facial movements.
- Instead of residual, learn two mask

$$\mathbf{I}_{\mathbf{y}_f} = (1 - \mathbf{A}) \cdot \mathbf{C} + \mathbf{A} \cdot \mathbf{I}_{\mathbf{y}_o}$$



## PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup

Huiwen Chang  
Princeton University

Jingwan Lu  
Adobe Research

Fisher Yu  
UC Berkeley

Adam Finkelstein  
Princeton University

